



1<sup>st</sup> International  
**Applied Statistics**  
Conference **UYİK**

01 - 04 October 2020 TOKAT - TÜRKİYE



*The conference is held on the honor of  
Prof. Dr. Yüksel BEK*



**1<sup>st</sup> INTERNATIONAL  
APPLIED STATISTICS CONFERENCE**

**( UYİK 2020 )**

**ONLINE**



**PROCEEDINGS BOOK**

ISBN: 978-975-7328-78-0

**October 01 - 04, 2020  
TOKAT - TÜRKİYE**



1<sup>st</sup> International  
**Applied Statistics**  
Conference **UYIK**  
01 - 04 October 2020 TOKAT - TÜRKİYE



# 1<sup>st</sup> INTERNATIONAL APPLIED STATISTICS CONFERENCE

## Proceedings of the UYIK-2020

ISBN: 978-975-7328-78-0

ONLINE - TOKAT / TURKEY

**Note:** All administrative, academic and legal responsibilities of the departments belong to their authors

# Proceedings of the 1<sup>st</sup> International Applied Statistics Conference (UYIK-2020)

**Publisher:**

Tokat Gaziosmanpasa University

**Editors:**

Yalcin TAHTALI, Samet Hasan ABACI, Lutfi BAYYURT

**E-Book Layout, Preparation and Composition:**

Yalcin TAHTALI, Samet Hasan ABACI, Lutfi BAYYURT

All published articles were peer-reviewed by Scientific Committee

The organizers do not have any legal liability for to contents of the presentation texts

**Organized by**

Tokat Gaziosmanpasa University, Tokat, Turkey

## COMMITTEES

### HONOURARY COMMITTEE

Dr. Ozan Balcı, Governor of Tokat

Prof. Dr. Bünyamin Şahin, Rector of Tokat Gaziosmanpaşa University

### CONFERENCE CO-CHAIRMAN

Prof. Dr. İsa Gökçe, Tokat Gaziosmanpaşa University

Prof. Dr. Orhan Kavuncu, Kastamonu University

Prof. Dr. Ahmet Şekeroğlu, Niğde Ömer Halisdemir University

Dr. Yalçın Tahtalı, Tokat Gaziosmanpaşa University

### CONFERENCE SECRETARIAT

Dr. Saniye Demir, Tokat Gaziosmanpaşa University

Dr. Samet Hasan Abacı, Ondokuz Mayıs University

PhD. Abdulkadir Keskin, İstanbul Medeniyet University

### ORGANIZING COMMITTEE

Prof. Dr. Ali Reza Vaezi, University of Zanjan

Prof. Dr. Aşir Genç, Necmeddin Erbakan University

Prof. Dr. Daniel Liberacki, University of Life Sciences in Poznan

Prof. Dr. Ercan Efe, Kahramanaraş Sütçü İmam University

Prof. Dr. Halil Aydoğdu, Ankara University

Prof. Dr. Hasan Önder, Ondokuz Mayıs University

Prof. Dr. Husna Leila Yusran, Trisakti University, Indonesia

Prof. Dr. İrfan Oğuz, Tokat Gaziosmanpaşa University

Prof. Dr. İsmail Akyol, Ankara University

Prof. Dr. Jacek Dlugosz, UTP University, Poland

Prof. Dr. Kadri Yürekli, Tokat Gaziosmanpaşa University

Prof. Dr. Kenan Kılıç, Niğde Ömer Halisdemir University

Prof. Dr. Kubilay Özyer, Tokat Gaziosmanpaşa University

Prof. Dr. Muhip Özkan, Ankara University

Prof. Dr. Murat Türkeş, Boğaziçi University

Prof. Dr. Mustafa Şahin, Kahramanaraş Sütçü İmam University

Prof. Dr. Piotr Stachowski, University of Life Sciences in Poznan

Prof. Dr. Rosman Bin Md Yusoff, Tun Hussien Onn University, Malaysia

Prof. Dr. Soner Çankaya, Ondokuz Mayıs University

Prof. Dr. Şenay Sarıca, Tokat Gaziosmanpaşa University

Prof. Dr. Tahir Rashid, University of Salford, England  
Prof. Dr. Taşkın Öztaş, Atatürk University  
Prof. Dr. Xun-Chang (John) Zhang, ARS Grazinglands Research Laboratory  
Prof. Dr. Tamer Kayaalp, Çukurova University  
Prof. Dr. X.C, Zhang, Agricultural Research Service. United States  
Prof. Dr. Zahide Kocabaş, Ankara University  
Prof. Dr. Zeynel Cebeci, Çukurova University  
Assoc. Prof. Dr. Adnan Ünal, Niğde Ömer Halisdemir University  
Assoc. Prof. Dr. Amar Iqbal Anwar, Cape Breton University, Canada  
Assoc. Prof. Dr. Asawer Alwasiti, University of Technology, Baghdad, Iraq  
Assoc. Prof. Dr. Çağrı Özgür Özkan, Kahramanmaraş Sütçü İmam University  
Assoc. Prof. Dr. Esen Oruç, Tokat Gaziosmanpaşa University  
Assoc. Prof. Dr. Fatih Yılmaz, Tokat Gaziosmanpaşa University  
Assoc. Prof. Dr. Mustafa Al-Mukhtar, University of Technology Baghdad, Iraq  
Assoc. Prof. Dr. Roman Rolbiecki, UTP University, Poland  
Assoc. Prof. Dr. Songül Gürsoy, Dicle University  
Assoc. Prof. Dr. Taner Tunç, Ondokuz Mayıs University  
Assoc. Prof. Dr. Velibor Špalevic, University of Montenegro  
Dr. Ahmet Pekgör, Necmettin Erbakan University  
Dr. Emine Berberoğlu, Tokat Gaziosmanpaşa University  
Dr. Derviş Topuz, Niğde Ömer Halisdemir University  
Dr. Kadir Karakaya, Selçuk University  
Dr. Saniye Demir, Tokat Gaziosmanpaşa University  
Dr. Kadir Karakaya, Selçuk Üniversitesi  
Dr. Yunus Akdoğan, Selçuk University  
Dr. Zafer Yıldız, Cumhuriyet University  
Dr. Abdulkadir Bektaş, Ministry of Energy and Natural Resources  
Dr. Bilal Habeşi Özkaynar, Turkish Statistical Institute  
Dr. Mesut Demircan, General Directorate of Meteorology  
Dr. Samet Hasan Abacı, Ondokuz Mayıs University  
PhD. Coşkun Parım, Yıldız Teknik University  
PhD. Mehmet Şamil Güneş, Yıldız Teknik University  
PhD. Lütfi Bayyurt, Tokat Gaziosmanpaşa University  
Salih Babagiray, General Directorate of Water Management  
Kadir Aytaç Özaydın, General Directorate of Agricultural Research and Policies

**SCIENTIFIC COMMITTEE**

- Prof. Dr. Abdullah Yeşilova, Van Yüzüncü Yıl University  
Prof. Dr. Asep Hermawan, Trisakti University, Indonesia  
Prof. Dr. Cemal Atakan, Ankara University  
Prof. Dr. Çiğdem Takma, Ege University  
Prof. Dr. Coşkun Kuş, Selçuk University  
Prof. Dr. Ecevit Eydurhan, Iğdır Üniversitesi  
Prof. Dr. Erkan Işığışok, Bursa Teknik University  
Prof. Dr. Emin Özköse, Kahramanmaraş Sütçü İmam University  
Prof. Dr. Ertuğrul Kardeş, Eskişehir Osmangazi University  
Prof. Dr. Güzin Yüksel, Çukurova University  
Prof. Dr. Halil Aydoğdu, Ankara University  
Prof. Dr. Handan Ankaralı, İstanbul Medeniyet University  
Prof. Dr. Hasan Bal, Gazi University  
Prof. Dr. Hasan Örkü, Gazi University  
Prof. Dr. Hikmet Orhan, Suleyman Demirel University  
Prof. Dr. Hüdaverdi Bircan, Cumhuriyet University  
Prof. Dr. Hülya Atıl, Ege University  
Prof. Dr. İsmail Keskin, Selçuk University  
Prof. Dr. İsmail Kınacı, Selçuk University  
Prof. Dr. Kabir Haruna Danja, Federal College of Education Zaira, Nigeria  
Prof. Dr. Kamile Kula, Kırşehir Ahi Evran University  
Prof. Dr. Kenan Yıldız, Tokat Gaziosmanpaşa University  
Prof. Dr. Mehmet Ali Cengiz, Ondokuz Mayıs University  
Prof. Dr. Mehmet Gürcan, Fırat University  
Prof. Dr. Mehmet Mendes, Çanakkale Onsekiz Mart University  
Prof. Dr. Mehmet Ziya Fırat, Akdeniz University  
Prof. Dr. Mohammad Masood Tariq, Balochistan University, Pakistan  
Prof. Dr. Muhammed Bekci, Cumhuriyet University  
Prof. Dr. Nazif Çalış, İskenderun University  
Prof. Dr. Ömer Cevdet Bilgin, Atatürk University  
Prof. Dr. Özgür Yeniay, Hacettepe University  
Prof. Dr. Özkan Görgülü, Kırşehir Ahi Evran University  
Prof. Dr. Özlem İlk Dağ, Orta Doğu Teknik University  
Prof. Dr. Rasim Koçyiğit, Tokat Gaziosmanpaşa University

Prof. Dr. Rüştü Yayar, Tokat Gaziosmanpaşa University  
Prof. Dr. Sait Ekinci, Kahramanaraş Sütçü İmam University  
Prof. Dr. Seyit Ali Kayış, Karabük University  
Prof. Dr. Sıddık Keskin, Yüzüncü Yıl University  
Prof. Dr. Turgut Aygün, Yüzüncü Yıl University  
Prof. Dr. Türkan Erbay Dalkılıç, Karadeniz Teknik University  
Prof. Dr. Willy Arafah, Trisakti University, Indonesia  
Prof. Dr. Yalçın Karagöz, Cumhuriyet University  
Prof. Dr. Yavuz Sanişoğlu, Yıldırım Beyazıt University  
Prof. Dr. Yılmaz Akdi, Ankara University  
Prof. Dr. Yüksel Terzi, Ondokuz Mayıs University  
Prof. Dr. Zafer Küçük, Karadeniz Teknik University  
Assoc. Prof. Dr. Alper Sinan, Akdeniz University  
Assoc. Prof. Dr. Özgür Koşkan, Isparta University of Applied Sciences  
Assoc. Prof. Dr. Abdulkerim Karaarslan, Atatürk University  
Assoc. Prof. Dr. Atıf Evren, Yıldız Teknik University  
Assoc. Prof. Dr. Bilge Gözener, Tokat Gaziosmanpaşa University  
Assoc. Prof. Dr. Emel Ceyhun Sabır, Çukurova University  
Assoc. Prof. Dr. Fatih Üçkardeş, Adıyaman University  
Assoc. Prof. Dr. Kübra Yazıcı, Tokat Gaziosmanpaşa University  
Assoc. Prof. Dr. Leman Tomak, Ondokuz Mayıs University  
Assoc. Prof. Dr. M. Çağatay Korkmaz, Artvin Çoruh University  
Assoc. Prof. Dr. Neslihan İyit, Selçuk University  
Assoc. Prof. Dr. Nursel Koyuncu, Hacettepe University  
Assoc. Prof. Dr. Ömer Alkan, Atatürk University  
Assoc. Prof. Dr. Pakize Taylan, Dicle University  
Assoc. Prof. Dr. Sinan Şahin, Namık Kemal University  
Assoc. Prof. Dr. Songül Akın, Dicle University  
Assoc. Prof. Dr. Şengül Cangür, Düzce University  
Assoc. Prof. Dr. Ülkü Erişoğlu, Necmettin Erbakan University  
Assoc. Prof. Dr. Ünal Korkmaz, Sakarya University  
Assoc. Prof. Dr. Yakut Gevrekçi, Ege University  
Assoc. Prof. Dr. Yüksel Öner, Ondokuz Mayıs University  
Dr. Ersin Uysal, Dicle University  
Dr. Abdurrahman Kara, Dicle University

- Dr. Adem Tüzemen, Tokat Gaziosmanpaşa University
- Dr. Ahmet Pekkör, Necmettin Erbakan University
- Dr. Ali Murat Tatar, Dicle University
- Dr. Aydın Karakoca, Necmettin Erbakan University
- Dr. Birol Topcu, Namık Kemal University
- Dr. Doğan Yıldız, Yıldız Teknik University
- Dr. Hande Küçükönder, Bartın University
- Dr. Harun Kınacı, Erciyes University
- Dr. Mustafa Zeybek, Artvin Çoruh University
- Dr. Onur Toka, Hacettepe University
- Dr. Selim Gündüz, Adana Bilim University
- Dr. Serdar Genç, Kırşehir Ahi Evran University
- Dr. Şirin Çetin, Tokat Gaziosmanpaşa University
- Dr. Tark Yılmaz, Aksaray University
- Dr. Emel Özgümüş Demir, Ministry of Family, Labor and Social Services
- Dr. Yener Ünal, Cumhuriyet University



## PROF. DR. YÜKSEL BEK'İN ÖZGEÇMİŞİ



Prof. Dr. Yüksel Bek, 1949 yılında Alaca (Çorum)'da doğdu. Atatürk Üniversitesi Ziraat Fakültesi'nden 1969 yılında mezun oldu. İki yıl asistanlık yaptıktan sonra, 1971 yılında devlet bursu ile lisansüstü öğrenimi için İngiltere'ye gitti. University of Reading, Applied Statistics (Department of Mathematics and Statistics) bölümünde lisans tamamlama eğitimini başarı ile bitirerek, University of Reading, Applied Statistics Department' da yüksek lisans öğrenimini tamamladı. Aynı yıl Atatürk Üniversitesi Ziraat Fakültesi Zootekni Bölümü'ne dönerek o zamanki ismi ile İstatistik ve Planlama bilim dalında göreve başladı.

Doktora öğrenimini 1976 yılında tamamladı. 1976-1981 yılları arasında Atatürk Üniversitesi Ziraat Fakültesi'nde Dr. Asistan olarak görev yaptı. 1981 Yılında Fen Bilimleri, Uygulamalı İstatistik (Biyostatistik) bilim dalında doçent oldu. YÖK yasası ile birlikte yeniden yapılanma sonrasında 1982 yılında Çukurova Üniversitesi Ziraat Fakültesi (ÇÜZF), Zootekni Bölümü, Biyometri ve Genetik Anabilim Dalı'nda doçent olarak göreve başladı. Aynı anabilim dalında 1989 yılında Biyometri profesörü oldu. 2000 yılında Ondokuz Mayıs Üniversitesi (OMÜ) Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı'nın kuruluşu ile birlikte Kurucu Anabilim Dalı Başkanı ve Biyoistatistik profesörü olarak göreve başladı. Bu görevini 2016 yılına kadar sürdürerek aynı yıl emekliye ayrıldı.

1982-2000 yılları arasında ÇÜZF Biyometri ve Genetik Anabilim Dalı başkanlığı, 1991- 1994 yıllarında ÇÜZF dekan yardımcılığı, 1994-1996 yıllarında Mustafa Kemal Üniversitesi İskenderun Mühendislik Fakültesi Dekanlığı, 1996-1999 yıllarında Çukurova Üniversitesi Bilgisayar Bilimleri Uygulama ve Araştırma Merkezi Müdürlüğü, 2000 yılında OMÜ Sağlık Hizmetleri MYO Müdürlüğü, 2000-2006 yıllarında OMÜ Bilgisayar Uygulama ve Araştırma Merkezi Müdürlüğü, 2000-2016 yıllarında OMÜ Tıp Fakültesi Biyoistatistik Anabilim Dalı başkanlığı görevlerini yürüttü. İstatistik alanında 100 civarında orijinal araştırma makalesi, 2 kitap ve konferans bildirileri şeklinde çok sayıda bilimsel eser sahibidir.

## PROF. DR. YÜKSEL BEK'İN MESAJI

Değerli konuklar, saygıdeğer hocalarım ve meslektaşlarım,

Ekonomik kriz ve pandemi koşulları altında fevkalade güç şartlara rağmen büyük bir çaba göstererek bu kongrenin gerçekleşmesini sağlayan Tokat Gaziosmanpaşa Üniversitesi yöneticilerine, kongre düzenleme ve bilim kurulu üyelerine, kongrenin gerçekleşmesine katkıda bulunan, desteklerini esirgemeyen tüm kişi, kurum ve kuruluşa teşekkürlerimi sunuyorum.

Ayrıyeten çok kapsamlı olarak düzenlenmesinden onur duyduğum bu kongrenin (I. Uluslararası Uygulamalı İstatistik Kongresi) adıma atfedilmesi beni ziyadesiyle memnun etmiştir. Bu kadirşinaslık hususunda anlayış gösteren, başta Tokat Gaziosmanpaşa Üniversitesi Rektörüne ve emeği geçen tüm düzenleme ve yürütme kurulu üyelerine şükranlarımı arz ediyorum.

Ben kendimi şanslı öğretim üyelerinden birisi olarak kabul ediyorum, çünkü dünden bugüne 45 yıllık akademik hayatımda birlikte çalıştığım, bilimsel birliktelik kurduğum, önemli katkılarımla olduğumu düşündüğüm genç, dinamik ve birbirinden değerli arkadaşlarımla hemen hemen tümü benimle amaçta, beklentilerde, görüş ve yaklaşımlarında bilimselliği yürekten paylaştılar. Çalıştığım tüm arkadaşlarımla, akademik yaşantılarında Victor Hugo'nun **“Kendi ışığına güvenen başkasının parlamasından rahatsızlık duymaz”** prensibi ile hareket etmişlerdir. Tüm arkadaşlarımla, her türlü eleştiriye açık, bilimsel açıdan olgunluğa erişse de oldum demeden, danışmayı ve itidal ile tartışmayı ilke edinmeleri, her türlü ortamda temsil yeteneğine ve deneyime sahip olmaları, mesleğin geleceği için beni ümitlendiriyor.

İstatistik bilimi, yöntem geliştirmeye yönelik çalışmalar dışında, okuyucu veya bizzat yürütücü olarak, bilimsel araştırmaya gönül vermiş tüm insanların bir noktasından dokunduğu bir fil misali, çok geniş bilimsel analiz yöntemler topluluğunu barındıran, istisnasız tüm bilim alanlarına çok önemli katkılar sunan disiplinler arası bir bilim dalıdır. Biyoloji, ziraat, tıp, mühendislik, ekonomi, eğitim, hukuk, diğer tüm fen ve sosyal bilimlerinde araştırma yapan, yazılan bilimsel makaleleri okuyup anlamaya çalışan, kanıta dayalı uygulamalar yapmak isteyen herkesin kullanmak zorunda olduğu, etki alanı oldukça geniş ancak her alana özgü farklı yönleri ile tanınan bir bilim dalıdır.

İstatistik yöntemlerin kullanılması o kadar genişlemiş ve gelişmiştir ki, uygulamalı istatistiğin (applied statistics) çeşitli alanlar için özelleştirilmiş alt disiplinleri ortaya çıkmıştır. Örneğin biyoloji, tıp, ziraat gibi biyolojik bilimler için biyometri (biometry) veya biyoistatistik (biostatistics), çevrebilimleri için çevresel istatistik (environmetry), jeoloji ve maden mühendisliği için jeometri (geometry), ekonomi bilimleri için ekonometri (econometry), psikoloji için psikometri (pyscometry), sosyometri, teknometri, arkeometri, yani diğer bir ifade ile başında bir alan ismi sonunda “metri” eki bulunan tüm kelimeler uygulamalı istatistiğin kapsamı içindedir. Temel istatistik prensipler aynı olsa dahi, bu alanların her birinde sıklıkla kullanılan, bazı diğer alanlarda pek sık kullanılmayan farklı istatistik yöntemler vardır. İstatistiğin alana özgü anılması, o alanda sık kullanılan istatistik yöntemlerin diğer alanda kullanılanlardan farklı olmasıdır. Ama hepsinin üst çatısı uygulamalı istatistiktir. Temeli bilgisi güçlü olan bir istatistikçi için bu alanlardan birinden diğerine geçiş belki, kısa bir süre o alandaki uygulamalarda deneyim kazanmayı gerektirebilir.

Diğer taraftan bilgisayar ve bilişim teknolojilerindeki gelişmeler ve veri yapılarındaki değişimler, veri işleme ve sonuçların yorumlamasında da bazı önemli değişikliklere neden olmaktadır. Son dönemlerde büyük verilerin (big data) bir platformdan başka bir platforma taşınması, yapay zekâ (AI) ve veri analiz teknolojilerinin değişmesi, tahminleme ve yapay zekâda kullanılacak alt yapının birbirinden ayrı olması veri işlemeyi zor ve farklı bir boyuta taşımıştır. Bu sorunların üstesinden gelebilmek, gigabayt, terabayt hatta peta baytlar büyüklüğündeki verilerin yarattığı problemleri çözmek için günümüzün en popüler araştırma ve uygulama alanlarının başında yer alan “Veri Bilimi” alanı doğmuştur. Doğası gereği tıpkı istatistik gibi veri bilimi alanı da disiplinler arası bir niteliğe sahiptir. Uygulamalı istatistik alanında çalışanların yeni doğan bu veri bilimi (data sciences) alanına ilgi göstermeleri gelecek gelişmeler açısından oldukça önemlidir.

Veri bilimi ile istatistik ne kadar iç içedir, bu tartışma konusudur. Bu konuya girmeden veri biliminin neleri hedeflediğini, hedefine ulaşmak için hangi bilim dallarından yararlandığını iyi anlamak gerekmektedir.

İstatistik araştırmacılar genelde kilobayt, megabayt, çok nadirde olsa gigabayt büyüklüğünde küçük verilerle muhatap olurken veri bilimciler genelde büyük veri (big data) ile muhataptırlar.

Küçük ve büyük verinin toplanması, düzenlenmesi, ayıklanması, işlenmesi arasında büyük farklar vardır. Veri hacmi belirli limitlerin üzerinde olduğunda, geleneksel sistem ve yöntemler bu tip verileri işlemede yetersiz kalmaktadır. Büyük ve karmaşık yapıdaki verinin öncelikle uygun işlenebilir formata dönüştürülmesi gerekmektedir. Küçük veri daha kontrollü ortamlarda ve kontrollü formatta toplanır, bu nedenle erişim ve işleme daha kolaydır. Aslında küçük veri hacim bakımından ulaşılabilir, işlenebilir ve bilgi edinilebilir boyuttaki veridir. Geleneksel veri işleme büyük ve karmaşık veri ile uğraşmaz.

Google gibi, büyük kurum ve kuruluşlara ait büyük verileri gibi, çoğu zaman gerçek zamanlı akıp gelen verilerdir. Modern teknoloji gittikçe daha fazla miktarda bilginin oluşturulabilmesini ve depolanabilmesini sağladığından mevcut veri hacminde büyük artışlar görülmektedir. Dünyadaki verilerin yüzde 90'ının son iki yılda oluşturulduğu tahmin ediliyor. Örneğin, Facebook kullanıcıları her saat 10 milyon fotoğraf yüklediği biliniyor. Dünya genelinde bağlantı halindeki cihazların sayısının, Internet of Things (IoT), 2025'e kadar 75 milyardan daha fazla büyümesi öngörülmüyor. Veri bilimi, verilerden değer elde etmek üzere bilimsel yöntemleri, süreçleri, algoritmaları ve sistemleri kullanan disiplinler arası bir alandır. Veri bilimi, her türlü ortamdan web, akıllı telefon, müşteriler, sensorlar ve diğer kaynaklardan toplanan farklı yapılarıdaki büyük verileri analiz etmek için bilgisayar bilimi, istatistik, iş bilimi ve bazı özel becerileri bir arada kullanmayı gerektiren yeni bir alandır. Veri bilimi uzmanları aynı ortamda kod yazma, çalıştırma, verileri görselleştirme ve görmek için birçok aracı aynı anda kullanırlar. Bunları yapabilmek için birçok programlama dillerinden yararlanılmaktadır. Python, R, Julia, Java, Scala ve Go gibi. Bunlardan her biri farklı yerlerde araç olarak kullanılabilir. Yani **veri bilimci** bir programcıdan çok istatistik bilen, bir istatistikçi'den çok programlama bilen bir kişidir.

Yeni teknolojiler sayesinde toplanan ve depolanan büyük veriler çok fazla çeşitliliğe ve zenginliğe sahip olmaktadır. Veri bilimciler genellikle kara kutu (black box) modelleri, sınıflama ve regresyon teknikleri, yapay zeka, makine öğrenmesi, derin öğrenme, genetik algoritmalar, bulanık mantık vb yöntemleri kullanmaktadır, olasılık modellerini (probabilistic models) kullanan yorumlayıcı istatistik yöntemler pek kullanılmamaktadır. Araştırmacı, geliştirmeci veri bilimciler ile veri bilim işletmeni olarak çalışanlar arasında da eğitim açısından bazı önemli farklılıklar vardır.

Teknoloji çağında Dünya genelinde organizasyonların ve toplumların dönüşümü ancak bu büyük verilerin yorumlanması sayesinde mümkündür. Bunların teferruatına burada girmek olanaksız, ama Uygulamalı istatistikçiler olarak hemen yanımızda gelişen bir "Veri Bilimi" gerçeğinin olduğunu artık unutmamak gerekir. Veri Bilimi, ilk kez 2002 yılında "International Council for Science" tarafından "Data Science Journal" duyuruldu: Bilim ve Teknoloji Verileri Komitesi. 2008 yılına gelindiğinde "veri uzmanı" unvanı ortaya çıktı ve kısa süre içinde alanda büyük ilerlemeler kaydedildi. Gittikçe daha fazla sayıda kolej ve üniversitede veri bilimi dereceleri sunulmaya başlansa da o zamandan bu yana yeterince veri uzmanı bulunamıyor, Türkiye'de bu sayı çok daha azdır.

İstihdam şeklinde de geçen zaman içerisinde büyük değişimler olmakta, büyük verilere sahip kurumlar ve şirketler verilerini değerlendirmek amacıyla istatistikçi yerine çoğu zaman veri bilimci aramaktadırlar. İstatistikçiler, özellikle uygulamalı istatistik dalında çalışanlar, büyük verilerle çalışabilme yeteneğini kazanabilmeleri için kendilerini veri bilimi teknikleri yönünden yetiştirmeleri mesleğimize önemli katkı getirecek ve istihdam içinde kolaylık sağlayacaktır.

Kongreler yoğun güncel bilgi aktarılması yanında, genelde genç bilim insanlarının açılım ve tanıtımı için fırsat sunması açısından da büyük önem ve değer taşımaktadır. Bu kongrenin uygulamalı istatistik alanında çalışan tüm akademisyenlere, sonuçlarını kullanacak olan diğer tüm araştırmacılara hayırlı olmasını ve önemli bilimsel katkılar sağlamasını diliyorum, herkesi saygı ile selamlıyorum.

Prof. Dr. Yüksel BEK

## PREFACE

Dear Scientists, Distinguished Researchers and Dear Friends,

With the auspices and support of Tokat Gaziosmanpaşa University, the “International Applied Statistics Conference”, which we organized for the first time between 1-4 October 2020, was held online with the contribution and participation of a large group of academicians. While organizing our congresses; we have aimed loyalty to our professors who are devoted to the science of investment and statistics for young people. For this purpose, we held our first congress on behalf of "Prof. Dr. Yüksek BEK". During the pandemic process that affected the whole world, our congress, which was opened face-to-face and then completed online, was attended from nine different countries, and presentations were made in Turkish and English and took place in a truly international congress climate. We would like to thank everyone who contributed to our congress.

We are planning to hold the second of our congress “2<sup>nd</sup> International Applied Statistics Congress”, again with your active contribution and participation, in the memory of our late teacher "Prof. Dr. Orhan DÜZGÜNEŞ" between 24-27 June 2021 in Tokat, in the atmosphere of an academic feast. We will be honored and happy to host you in Tokat which is one of Turkey's history-steeped cities.

In our information feast, which we plan to hold with a multidisciplinary perspective in our second congress, as in this congress, it is aimed to discuss and record topics from every branch of applied statistics in a scientific environment. We await your contribution and participation and wish you convenience in your work.

### Conference Co-Chairmans

## ÖNSÖZ

Çok Değerli Bilim İnsanları, Kıymetli Araştırmacılar ve Değerli Dostlarımız,

Tokat Gaziosmanpaşa Üniversitesi himayesi ve destekleriyle 1-4 Ekim 2020 tarihleri arasında birincisini düzenlediğimiz “Uluslararası Uygulamalı İstatistik Kongresi” geniş bir akademisyen grubunun katkı ve katılımları ile çevrimiçi olarak gerçekleştirilmiştir. Kongrelerimizi düzenlerken; gençlere yatırım ve istatistik bilimine gönül vermiş hocalarımıza vefayı amaç edindik. Bu amaç doğrultusunda ilk kongremizi Prof. Dr. Yüksek Bek adına gerçekleştirdik. Tüm dünyayı etkisi altına alan pandemi sürecinde yüz yüze açılış ve devamında çevrimiçi olarak tamamlanan kongremize dokuz farklı ülkeden katılım sağlanmış Türkçe ve İngilizce sunumlar yapılmış ve gerçek anlamda uluslararası kongre ikliminde gerçekleşmiştir. Kongremize katkı veren herkese gönülden teşekkür ederiz.

Kongremizin ikincisini “2. Uluslararası Uygulamalı İstatistik Kongresi”, yine sizlerin etkin katkı ve katılımlarıyla, merhum hocamız Prof. Dr. Orhan DÜZGÜNEŞ anısına 24-27 Haziran 2021 tarihleri arasında, Tokat’ta akademik bir şölen havasında gerçekleştirmeyi planlıyoruz. Sizleri, yılın en güzel dönemlerinin birinde, Türkiye’nin tarih kokan şehirlerinden biri olan Tokat’ta ağırlamaktan onur ve mutluluk duyacağız.

Bu kongrede olduğu gibi ikinci kongremizde de multidisipliner bir bakış açısı ile gerçekleştirmeyi planladığımız bilgi şölenimizde, uygulamalı istatistiğin her dalından konuların bilimsel bir ortamda tartışılıp, kayıt altına alınması hedeflenmektedir.

Ulaşılabilecek bilgi birikiminin ISBN numaralı bildiri kitabında ve akademik dergilerde yayınlanması planlanmaktadır. Katkı ve katılımlarınızı bekler, çalışmalarınızda kolaylıklar dileriz.

### Kongre Eş Başkanları

## TABLE OF CONTENT

COMMITTEES.....	iii
PROF. DR. YÜKSEL BEK'İN ÖZGEÇMİŞİ .....	viii
PROF. DR. YÜKSEL BEK'İN MESAJI.....	ix
PREFACE .....	xi
TABLE OF CONTENT .....	xii
ONLINE PRESENTATIONS.....	1
ONLINE PRESENTATIONS CONTENT .....	2
POSTER PRESENTATIONS.....	298
POSTER PRESENTATIONS CONTENT .....	299
NATIONALITY OF PARTICIPATIONS .....	304

## **ONLINE PRESENTATIONS**

## ONLINE PRESENTATIONS CONTENT

Estimation of Lognormal Renewal Function under Progressively Censored Data .....	5
Fuzzy Estimator for the Median and an Application .....	6
Statistical Modeling of Strong Earthquake Occurrences by Binary Dependent Hidden Markov Model .....	7
Cereal Data Analysis to Compare the Mean Square Error Criterion of the Predictors in Linear Mixed Models.....	8
Koronavirüs Anksiyete Ölçeği Sivas İli Uygulaması.....	9
Determination of Best Variance-Covariance Structure in Mixed Model (Sas Proc Mixed) with Various Parameter Estimation Methods.....	16
Determination of Difference between Heavy Metal Values in Vegetables Growing in Soils Composed of Erciyes Strato Volcano by Fuzzy and Classical Variance Analysis .....	17
Using Path Analysis to Predict Egg Weight from Egg Physical Traits of Potchefstroom Koekoek Chicken Layer Breed.....	18
Tarım İşletmelerinde Üretim Planında Değişikliğe Etki Eden Faktörlerin Belirlenmesi.....	19
Bread Consumption Habits of Selcuklu District of Konya Province and Willingness to Pay .....	26
Sulama Organizasyonlarının Yönetim Şekillerine Göre Etkinliklerinin Belirlenmesi .....	35
Solar Radyasyondaki Değişimin Farklı İstatistik Yaklaşımlarla Analizi.....	44
Reliability Evaluation of Circular $m$ -Consecutive- $k$ -out-of- $n$ : F System with Non-Overlapping Runs Subjected to Shocks .....	45
PM10 and SO <sub>2</sub> predictions for the most polluted station in Ankara .....	46
Prioritization of Ergonomics Projects with Integration of CRITIC and CODAS Methods .....	47
Examining the Relationship between Organizational Commitment and Business Motivation - A Practice in the Public Sector.....	49
Normal Dağılım Varsayımı Sağlanmadığı Durumda Süreç Ortalamasını Takip etmek için Çarpıklık ve Basıklık Düzeltmesi Yöntemi.....	51
Diagnostic Accuracy of PCR Test and Symptoms of the Disease for Covid-19 by Multivariate Roc Analysis.....	62
Karar Ağaçları Yöntemi ile Operatör Değişikliği Nedenlerinin Araştırılması.....	63
The Use of New Robust Approaches for Nonlinear Quantile Regression When There are Outlying Observations.....	73
Consumption and Preferences of Cooperative Branded Products in Tokat Provincial Center .....	74
Zara Honey That a Local Product Consumption of Individuals and Reasons for Preference (Sivas Provincial Center District) .....	76
Optimization of Power Conservation and Pressure Drop of Nanofluid Two-Phase Flow in the Pipeline.....	78
Using Entropy Measures in Weighted Least Squares with Non-Linear Datasets .....	92
Baz İstasyonlarında Risk Değerlendirmesi .....	93
Bound Approximations for Reliability of Circular Consecutive- $k$ -out-of- $n$ Systems.....	94
Investigation of Relations between Economic Growth, Energy Use and Human Capital with Seemingly Unrelated Regression Model.....	95
Simulation of Reorder Point with Variable Demand .....	96
Automation of Experimental Design with MS Excel VBA .....	100
Bilgisayar ve Telekom Sektörleri Ürün Gruplarının Üzerine Bir Uygulama.....	102
A New Unit Probability Distribution Approach with the Lindley Distribution Example.....	115
Yaşam Standartları Faktörünün Gelişmemişlik Faktörü Üzerinden Doğuştan Yaşam Beklentisine Etkisi: Yapısal Bir Model.....	116
Türkiye’de Ekonomik Büyüme, İşsizlik ve Enflasyonun Kayıtdışı İstihdam Üzerindeki Etkisi: ARDL Sınır Testi Yaklaşımı.....	122
Uzaktan Eğitim-Öğretim (Online) Sisteminde Etik Sorunlar ve Çözüm Önerileri Üzerine Bir Araştırma ve Etik Yemini .....	123
Determination of G20 Indices Affecting BIST with Hybrid Regression Models .....	129
Investigation of the Relation of Gold Prices and Crypto Money in the Pandemia Process with Machine Learning Methods .....	130

Statistical Inference for the Quasi-Exponential Distribution Based on Progressively Type-II Censored Samples .....	131
Türkiye’den New Jersey’e Göç ve Tersine Göç Üzerine Araştırma .....	132
Şekerli Su Kullanımının Ross Erkek Broilerlere Etkilerinin Yaşam Analiziyle Belirlenmesi .....	133
Regression Analysis of Morphological Traits on Body Weight in South African Non- Descriptive Indigenous Goats.....	140
Determination of Best Fitted Regression Model for Estimation of Body Weight in Nguni Cattle..	141
Forecasting for Tea Production and Yield in Turkey .....	142
Bağlantı Dengesizliği Katsayısının Örnekleme Dağılımı Üzerine Bir Simulasyon Çalışması.....	143
Estimating Claims Reserve Splitting IBNER and IBNYR in the Chain Ladder Method .....	151
Determination of Direct and Indirect Effects of Weaning Weight in Anatolian Buffaloes .....	156
Changes in Statistics in the Big Data Age.....	157
Factors Influencing the Educational Achievements of the Students in the University of Peshawar, Pakistan .....	158
The Cohort Component Method Applied in the Albanian Population Projection .....	163
Clustering Using Genetic Algorithm in Data Mining .....	173
Clustering Analysis Using Ensemble Methods in Machine Learning.....	174
Using Stochastic Differential Equations Modelling for Agricultural Products on Real Data.....	175
Increasing the Power of Goodness of Fit Tests for Cauchy Distribution Versus Symmetrical Distributions for Small Samples by Means of Characterization .....	176
Detecting the Speaker Using Artificial Neural Networks on the Audio Signal .....	177
Flexicurity: A Comparison to OECD Countries .....	178
Tünel Açma Makineleri Üreten Bir Firmanın Kesici Disk Sert Dolgu Sürecinde İstatistiksel Süreç Kontrolü Çalışmaları .....	179
Estimation of the Ridge Parameter Using the Jackknife Approach .....	180
Statistical Modelling of Lamb Culling During the Rearing Period with the Use of Classification Trees.....	181
Normality Tests of Morphological Measurements of Biological Data of Sheep .....	182
Nonlinear Dimensional Reduction to with the Use of Artificial Neural Networks in Patients with Breast Tumor.....	183
Application of Gompertz Model on Broilers Growth .....	184
Sample Size in Clinical Researches: Power of the Test and Effect Size.....	185
Modification of the Schnute Growth Model .....	187
A Comparative Study of Excel Spreadsheet and JMP Statistical Software’s for Regression Analysis in Agricultural Studies: A Case Study of Chickpea Yield Management by Land Rolling .....	188
Finite Element Models Applied to Predict Soil Compaction in Agricultural Areas .....	194
Öğretmenlerin Örgütsel Sinizm Algısının İncelenmesi .....	198
Problems Encountered in Organic Almond Growing; in Case of Dicle and Eğil .....	204
Almond Producers' Perspectives towards Agricultural Organizations: In Case of Dicle and Eğil ..	205
Rüzgar Hızının İstatistiksel Analizi .....	206
Türkiye’nin Farklı Bölgelerindeki Aylık ve Yıllık Yağışların Dağılım Fonksiyonlarının Belirlenmesi .....	207
Türkiye’nin Farklı Bölgelerindeki Yağışların Mevsimsel ve Yıllık Yağış Trendleri .....	209
Hipotez Kontrolünde, Çeşitli Dağılımlardan Çekilen Örneklerin Genişliği Kaç Olmalıdır?.....	210
Tavuklarda Yumurta Ağırlığı ile İlişkili Genlerin Farklı Çoklu Hipotez Test Yöntemleri ile Araştırılması.....	211
Sinop İli’nde Hava Sıcaklıklarının İstatistiksel Olarak Değerlendirilmesi .....	212
Producing Short-Term Forecasts of Turkish Gross Domestic Product: A Bottom-Up Approach from the Demand Side .....	213
PISA 2018 Testi Türkiye Örnekleme Göre Öğrencilerin Eğitim ve Sosyal Faaliyetlerinde Dijital Cihaz ve İnternet Kullanımlarının Değerlendirilmesi .....	214
Experimental Design and Use of SPSS in Agricultural Research.....	224
Drug Eluting Stents versus Bare Metal Stents in Patients with ST-Elevated Myocardial Infarction	225
Uncertainty of the Claims Development Result in the Chain Ladder Method for Compulsory Motor Insurance .....	230



Etlik Piliçlerde Bazı Ağırlık ve Kan Parametreleri Arasındaki İlişkinin Kanonik Korelasyon Analizi ile İncelenmesi .....	231
Türkiye'de Süt Sığırlarından Üretilen Süt Miktarının Bulanık K-Ortalamaları Analiziyle Değerlendirilmesi .....	232
Hisse Senedi Fiyat Hareketlerini Prophet Zaman Serisi Modeli İle Öngörüleme.....	233
Independent and Dependent Proportion Comparisons .....	235
Canonical Correlation and a Livestock Application .....	236
Altın Fiyatlarındaki Değişimin Farklı Regresyon Yöntemleri ile Karşılaştırılması: Covid-19 Pandemisi Etkisi.....	237
Laboratuvar Hata Miktarlarının İstatistiksel Kalite Kontrol ve Delphi Yöntemi ile Azaltılması ...	238
Covid 19 Pandemisi Sürecinde BİST'te İşlem Gören Perakende Ticaret Şirketlerinin Finansal Performanslar Bağlamında Critic ve Copras Yöntemleriyle Karşılaştırılması .....	240
COVID-19 Hastalığı için Ülkeler Bazında Virüs Bulaşma Katsayısı ( $R_0$ ) ile Nüfus Yoğunluğu ve Ortalama Hanehalkı Büyüklüğü Arasındaki İlişki .....	242
Bireylerin İnternet Kullanım Alışkanlıklarının Genelleştirilmiş Doğrusal Modeller ve Sınıflandırma Analizi ile İncelenmesi.....	243
Modeling the Frequency of Auto Insurance Claims in the Albanian Market .....	245
Rüzgâr Hızı için SETAR Modeli ile Öngörü Elde Edilmesi.....	246
Türkiye'nin İklimsel Değişikliği ve Mevsimsel Ekstrem Yağışların İstatistiksel Olarak Değerlendirilmesi .....	247
Increasing the Power of Goodness of Fit Tests for Cauchy Distribution Versus Symmetrical Distributions for Small Samples by Means of Characterization .....	248
Kantil Regresyon Analizi ile Saanen Keçilerinde Bazı Vücut Ölçümleri Kullanılarak Canlı Ağırlık Tahmini .....	257
Bazı Büyüme Modellerinin Lineer Eğik Asimptotlu Olduklarındaki Uyum Performansları .....	258
Evaluation of Twitter Messages Related to Covid-19 Pandemic Using .....	259
Text Mining Technique.....	259
Meralarda Homojen Alanlarına Göre İndikatör Türlerin Belirlenmesinin İstatistiksel Açıdan Değerlendirilmesi .....	260
Comparison of Some Different Clustering Methods in Double Dendrogram Heat Maps .....	270
Solvability of Stochastic Fractional Differential Equation with Q-Winer Processes.....	271
Solvability System of Stochastic Fractional Differential Equations with Multiplicative Random Noise in Probability Space .....	272
Use of Biostatistics to Assess Performance of Balochistan Camel Breeds.....	273
Çekirdek Yoğunluk Tahmininde Kullanılan Bant Genişliği Seçim Yöntemleri.....	274
Nonparametric Estimation of Renewal Function under Progressively Censored Data.....	275
Data Visualization and Data Analysis Examples from Textile Engineering .....	276
Kadının İş Hayatındaki Rolünün Çok Değişkenli İstatistiksel Yöntemler Kullanılarak Değerlendirilmesi .....	277
Statistical Inference for Doubly Geometric Process with Lognormal Interarrival Times.....	297

## Estimation of Lognormal Renewal Function under Progressively Censored Data

Ömer ALTINDAĞ<sup>1</sup>, Halil AYDOĞDU<sup>2</sup>

<sup>1</sup>Department of Statistics and Computer Sciences, Bilecik Şeyh Edebali University, Bilecik, Turkey

<sup>2</sup>Department of Statistics, Ankara University, Ankara, Turkey

### Abstract

Renewal process, a counting process based on a sequence of independent and identically distributed random variables, is an important stochastic model being used in fields of applied probability such as reliability theory, inventory theory, queuing theory, warranty and risk analysis. In many applications involving renewal process knowledge of its mean value function, which is also called as renewal function, is required. This function gives expected number of events occurred in a specified period of time. Let  $X_1, X_2, \dots, X_n$  be a realization of a renewal process. Then renewal function can be estimated based on this sample. Estimation of renewal function is well studied in the literature if the sample is complete. Frees 1986a and 1986b suggested several parametric and nonparametric estimators for renewal function and investigated their statistical properties. Lin 1988 studied estimation of renewal function when the sample is not complete and includes right censored observations. Later, Baxter and Li 1995 dealt with nonparametric interval estimation of renewal function under right censored data. Estimation problem of renewal function is still being studied in the literature for different cases. For instance, see Alvarez 2006, Bebbington et al 2007 and Aminzadeh 2011. Recently, Altındağ 2017 have considered estimation problem of renewal function when the sample is progressively censored. In this study, we consider parametric estimation of renewal function when the distribution of inter-renewal times is lognormal and the sample is progressively censored. A plug-in estimator for the lognormal renewal function is proposed and its statistical properties such as consistency, asymptotic unbiasedness and asymptotic normality are established. A simulation study is also performed to evaluate small sample behavior of the estimator.

**Key words:** Estimation, progressive censoring, renewal function, lognormal distribution

## Fuzzy Estimator for the Median and an Application

Derviş TOPUZ<sup>1</sup>

<sup>1</sup>Department of Health Services Science, Niğde Zübeyde Hanım Vocational School of Health Services, Niğde Ömer Halisdemir University, Niğde, Turkey

### Abstract

A single value in a statistic that can represent a data set is called an average. In the calculations made with different methods, sometimes all values or classes belonging to the data set are not taken into account. In addition, in order to make valid and reliable decisions about any event (birth, death, disease), the probability distribution of the variables is examined. If the data distribution is distorted due to unsymmetrical outliers, for a continuous probability distribution, a difference between the median number value and the average number value may occur equal to a standard deviation. The median method developed with the fuzzy logic approach for these cases can be a more valid and reliable average measure compared to the arithmetic mean. In this study, the steps of estimating and interpreting the median value with a fuzzy approach were performed on a clinical dataset. The median of the sample dataset,  $\tilde{m} = 24.2$  value was calculated between 0.0% and 99.6% confidence. It has been tried to be shown with this study that the calculated median value represents the values of the dataset with contradictory observations more validly and reliably.

**Key words:** Median, confidence, Fuzzy estimator, non-parametric

## Statistical Modeling of Strong Earthquake Occurrences by Binary Dependent Hidden Markov Model

Ozgur DANISMAN, Umay UZUNOGLU KOCER

### Abstract

Hidden Markov models are stochastic processes that have mathematical applications in many areas such as engineering, speech recognition, signal processing or seismology. It can be considered as an extension of classical Markov models and can be defined as a two-dimensional stochastic process. The first stochastic process is related to the hidden (latent) states and it is modeled by finite-state Markov chains. The second stochastic process, on the other hand, is related to the observations emitted from the hidden states of the Markov chain. The observations appear explicitly to observer after each hidden state transition. Hence the name ‘hidden Markov model’ is used. The main assumption of hidden Markov models is that; observation symbols are independent and identically distributed random variables. They depend only on the hidden state in which they are emitted at the same time. There exist some cases where this assumption may not be valid in practice. That is, an observation that occurs in the current state may depend on the previous observation that occurred in the previous state. In this study, a new type of hidden Markov model is introduced in which the current hidden state and emitted observation, and the previous hidden state and emitted observation have a first-order Markov dependency. The model is capable of capturing a possible first-order Markov dependency between the last and the previous steps of the system. It is an alternative model to the classical hidden Markov model for revealing the Markov dependency between the current and the previous binary information of the system. An experimental study is conducted to show the performance of the proposed model compared to the classical hidden Markov model. Besides, occurrences of strong earthquakes are modeled with both the classical hidden Markov model and the proposed model, and the results are compared.

**Key words:** Hidden Markov model, Markov process, binary dependency, earthquake geology, strong earthquake, parameter estimation

**Cereal Data Analysis to Compare the Mean Square Error Criterion of the Predictors in  
Linear Mixed Models**

Özge KURAN<sup>1</sup>

<sup>1</sup>Department of Statistics, Dicle University, Diyarbakır, 21280, Turkey

**Abstract**

Linear mixed models are commonly preferred in the examination of grouped data include longitudinal data, repeated measures, blocked designs and multilevel data. These data species appear in many ranges as various as agriculture, biology, economics, manufacturing and geophysics. Although the most widely employed prediction procedures are the best linear unbiased estimator and the best linear unbiased predictor in linear mixed models, the procedures exist unstable and ascribe misleading outcomes under multicollinearity condition. To avoid these undesirable outcomes, alternative prediction procedures are introduced for linear mixed models. Some of these are the ridge predictors, the Liu predictors and the modified ridge predictors in linear mixed models. The goal of this study is, under multicollinearity, to confront the mean square error criterion performances of these predictors with cereal data analysis.

**Key words:** Linear mixed model, Liu predictor, modified ridge predictor, multicollinearity, Ridge predictor

## Koronavirüs Anksiyete Ölçeği Sivas İli Uygulaması

Şebnem ZORLUTUNA<sup>1</sup>

<sup>1</sup>Sivas Cumhuriyet Üniversitesi, İ.İ.B.F. Ekonometri Bölümü, Sivas, Türkiye

### Özet

Dünya Sağlık Örgütü tarafından 30 Ocak 2020’de, Covid-19 hastalığının uluslararası kamu sağlığı acil durumu olduğu açıklanmış ve 11 Mart 2020 tarihinde ise virüs pandemi, yani küresel salgın hastalık olarak ilan edilmiştir. Ardından birçok önlem ve uygulamalar hızlı bir şekilde hayatımıza girmiş, hükümetler tarafından uygulanan sokağa çıkma kısıtlamaları, sıkı sosyal mesafe kuralları, okul ve işyerlerinin kapatılması ve sürekli artan enfeksiyon sayıları ile bütün dünyada yaşam alt üst olmuştur. Covid-19’a bağlı ölüm sayısı ve sıkıntılar artmaya devam ettikçe yüksek düzeyde uzun süreli korku ve anksiyete yaşayan insanların sayısında da bir artış gözlemleneceği daha önce yaşanan salgınlardan bilinmektedir. Yaşanılan süreçte de koronavirüs enfeksiyonu olan kişileri tespit etmek için alınacak önlemlere büyük önem verilirken pandemiden etkilenen kişilerin ruh sağlığı bakım ihtiyaçlarının belirlenmesi nispeten ihmal edilmektedir. Bu çalışma, COVID-19 kriziyle ilişkili olası disfonksiyonel anksiyete vakalarını tanımlamak için geliştirilen Koronavirüs Anksiyete Ölçeğini (KAÖ) Sivas halkına uygulamayı amaçlamaktadır. Böylece koronavirüsün Sivas halkı üzerinde yarattığı anksiyetin ölçülmesi ve gerekli kişilerce toplum ruh sağlığının geliştirilmesinin sağlanabileceği düşünülmektedir.

**Anahtar Kelimeler:** Anksiyete, Covid-19, Koronavirüs Anksiyete Ölçeği, pandem-i, Sivas.

### Abstract

It was announced by the World Health Organization on January 30, 2020 that Covid-19 disease is an international public health emergency, and on March 11, 2020, it was declared as a virus pandemic, that is, a global epidemic. Later, many measures and practices quickly entered our lives, and life all over the world was turned upside down with the curfew restrictions imposed by governments, strict social distance rules, closure of schools and workplaces, and the ever-increasing number of infections. It is known from previous epidemics that as the number of deaths and troubles due to Covid-19 continues to increase, the number of people experiencing high levels of long-term fear and anxiety will also increase. While the measures to be taken to detect people with coronavirus infection are given great importance in the process, determining the mental health care needs of people affected by the pandemic is relatively neglected. This study aims to apply the Coronavirus Anxiety Scale developed to identify possible cases of dysfunctional anxiety associated with the COVID-19 crisis to the people of Sivas. Thus, it is thought that the anxiety created by the coronavirus on the people of Sivas can be measured and the necessary people can improve the mental health of the community.

**Keywords:** Anxiety, Covid-19, Coronavirus Anxiety Scale, pandemic, Sivas

### Giriş

Koronavirüs hayvanlarda bolca rastlanan büyük bir virüs ailesi olarak tanımlanmaktadır. Önce hayvandan insana bulaşan virüsün insandan insana da yayılabildiği zamanla anlaşılmıştır. COVID-19, yeni koronavirüs hastalığının ismi olup, ilk kez 2019 yılının son aylarında Çin’de tespit edilmiştir. COVID-19 salgını yaygın bulaşma potansiyeli, yüksek ölüm oranı, öngörülemezliği ve alışılmadık korunma yolları (sosyal mesafelenme, izolasyon) nedeniyle pek çok boyutta krize neden olan küresel

bir afettir. Travmatik etkiler yaratma potansiyeline sahip her afet gibi bu salgın da psikososyal etkileri açısından ele alınmalıdır (Türk Psikologlar Derneği Travma, Afet ve Kriz Birimi).

Dünya Sağlık Örgütü verilerine göre Eylül ayı sonu itibari ile dünyadaki vaka sayısı 33.249.563'e, ölüm sayısı ise 1.000.040'a ulaşırken, Türkiye'deki vaka sayısı 315.845'e, ölüm sayısı ise 8.062'ye ulaşmıştır. Covid-19'a bağlı ölüm sayısı ve sıkıntılar artmaya devam ettikçe yüksek düzeyde uzun süreli korku ve anksiyete yaşayan insanların sayısında da bir artış gözlemlenebilmektedir (Guan ve ark., 2020).

Anksiyete, kişinin bir problemle başa çıkamama konusundaki aşırı uyarılmasını ifade eder. Belirli bir odak noktası olan ve tehdidin ortadan kalkmasıyla hızlıca çözülen korkunun aksine, anksiyete belirsizdir ve korkuya oranla çok daha yavaş kaybolur (Kalat vd., 2007). Kovid-19 salgını nedeniyle süreç boyunca ve süreç sonrasında insanlar ruh sağlığı konusunda sıkıntı yaşayabileceklerdir. Fakat koronavirüs enfeksiyonu olan kişileri tespit etmek için alınacak önlemlere büyük önem verilirken pandemiden etkilenen kişilerin ruh sağlığı bakım ihtiyaçlarının belirlenmesi nispeten ihmal edilmektedir (Xiang ve diğerleri, 2020).

Bu çalışmada, Koronavirüs Anksiyete Ölçeği Sivas halkına uygulanmış ve uygulama sonucunda örneklemin %11,7'sinde anksiyete varlığı görülmüştür. Ayrıca, yapılan analizler anksiyete belirtisi ile cinsiyet, yaş, gelir düzeyi, meslek, oturlan ev türü, kişide kronik bir hastalığın var olması, evde kronik hastalığı olan kişi varlığı ve son 2 hafta içinde koronavirüs hakkında düşünme veya medya izleme süresi arasında anlamlı bir ilişki olduğunu göstermiştir.

### **Materyal ve Metot**

Araştırma örneklemini, basit rastgele örneklem yöntemi ile seçilen Sivas'ta yaşayan 214 kişiden oluşmaktadır. Katılımcılara koronafobiyi ölçmek için kullanılan Koronavirüs Anksiyete Ölçeği ve demografik özellikleri içeren sorulardan oluşan bir anket uygulanmıştır.

Koronavirüs Anksiyete Ölçeği, kişilerin koronavirüsle ilgili düşünce ya da bilgilere maruz kaldıklarında fizyolojik olarak buna dayalı korku ve anksiyete belirtilerini ne sıklıkla yaşadıklarına ilişkin 5 soru ve tek boyuttan oluşan 5'li likert türünde bir ölçektir. Puanlanması "0" "hiçbir zaman", "1" "Nadir, bir veya iki günden az", "2" "Birkaç gün", "3" "7 günden fazla" ve "4" "son iki haftada neredeyse her gün" şeklindedir (Lee vd., 2020-a)

KAÖ, Lee (2020) tarafından COVID-19 kriziyle ilişkili olası disfonksiyonel anksiyete vakalarını tanımlamak için geliştirilmiştir. Alan dizinde yer alan üç farklı (Biçer ve ark., 2020; Evren ve ark., 2020; Akkuzu ve ark., 2020) uyarılma çalışmalarının sonucunda KAÖ'nin psikometrik özelliklerinin orijinal çalışma ile paralellik gösterdiği, Türkçe'ye uyarlamada gerekli koşulların sağlandığı, güvenilir ve geçerli bir ölçüm aracı olarak Türk popülasyonunda kullanılabilir olduğu görülmüştür.

### **Bulgular ve Tartışma**

Araştırma örneklemini, basit rastgele örneklem yöntemi ile seçilen Sivas'ta yaşayan 214 kişiden oluşmaktadır. Tablo 1'de katılımcıların demografik özellikleri ve cevaplarına ilişkin dağılımlar görülmektedir.

Tablo 1. Katılımcıların Demografik Özellikleri ve Cevaplarına İlişkin Dağılım

YAŞ	f	%	CİNSİYET	f	%
18-40	144	67,3	Kadın	94	47,5
41-64	65	30,4	Erkek	104	52,5
65 ve üzeri	5	2,3			
<b>Toplam</b>	214	100,0	<b>Toplam</b>	198	100,0
EĞİTİM	f	%	MESLEK	f	%
İlköğretim	5	2,4	Ev Hanımı	10	4,7
Ortaöğretim	19	9,0	Memur	100	47,2
Mesleki Eğitim	10	4,7	Esnaf	2	0,9
Önlisans	15	7,1	Serbest Meslek	14	6,6
Lisans	98	46,2	İşçi	27	12,7
Lisansüstü	65	30,7	Öğrenci	12	5,7
<b>Toplam</b>	212	100,0	Diğer	47	22,2
			<b>Toplam</b>	212	100,0
MEDENİ DURUM	f	%	SAĞLIK PERSONELİ	f	%
Evli	157	73,4	Evet	43	20,1
Bekâr	57	26,6	Hayır	171	79,9
<b>Toplam</b>	214	100,0	<b>Toplam</b>	214	100,0
GELİR	f	%	ÇOCUK SAYISI	f	%
0-1500 TL	13	6,8	Yok	77	36,0
1501-3000 TL	22	11,5	1 tane	51	23,8
3001-4500 TL	38	19,9	2 tane	71	33,2
4500 TL'den fazla	118	61,8	3 veya daha fazla	15	7,0
<b>Toplam</b>	191	100,0	<b>Toplam</b>	214	100,0
YAŞANILAN EV TÜRÜ	f	%	EVDE YAŞAYAN KİŞİ SAYISI	f	%
Apartman	127	63,2	1 kişi	19	8,9
Site	47	23,4	2 kişi	38	17,8
Müstakil	27	13,4	3 kişi	51	23,8
<b>Toplam</b>	214	100,0	4 kişi	88	41,1
			5 veya daha fazla	18	8,4
			<b>Toplam</b>	214	100,0
KRONİK HASTALIK VARLIĞI	f	%	EVDE KRONİK HASTALIĞI OLAN	f	%
Var	23	10,7	Var	43	20,1
Yok	191	89,3	Yok	171	79,9
<b>Toplam</b>	214	100,0	<b>Toplam</b>	214	100,0
EVDE 65 YAŞ ÜSTÜ YAŞAYAN	f	%	COVID-19 TANISI KONDU MU?	f	%
Var	19	8,9	Evet	6	2,8
Yok	195	91,1	Hayır	208	97,2
<b>Toplam</b>	214	100,0	<b>Toplam</b>	214	100,0
COVID-19 OLAN TANIDIK/AKRABA	f	%	DAHA ÖNCE ANKSİYETE TEDAVİSİ GÖRDÜNÜZ MÜ?	f	%
Var	139	65,0	Evet	11	5,2
Yok	75	35,0	Hayır	202	94,8
<b>Toplam</b>	214	100,0	<b>Toplam</b>	213	100,0
SON 2 HAFTADA CORONAVİRÜS HAKKINDA KAÇ SAAT DÜŞÜNDÜNÜZ/MEDYA İZLEDİNİZ				f	%
1-3 saat				64	29,9
3-5 saat				51	23,8
5-7 saat				36	16,8
7 saatten daha fazla				63	29,4
<b>Toplam</b>				214	100,0



Tablo 2’de 5 sorudan oluşan koronavirüs anksiyete ölçeğine verilen cevapların dağılımı görülmektedir. Dağılımlara bakıldığında en çok 2.soruda belirtilen koronavirüs hakkında düşünüldüğünde uykuya dalma ve uykuda kalma konusunda sorun yaşandığı görülmektedir. İkinci sırada, Koronavirüs ile ilgili haberler okuduğunda veya dinlediğinde baş dönmesi, sersemleşme veya baygın hissetme sorunu yer almaktadır. Üçüncü sırada, Koronavirüs hakkında düşünüldüğünde veya bilgiye maruz kalındığında mide bulantısı hissetme ve mide problemleri yaşama; dördüncü sırada, Koronavirüs hakkında düşünüldüğünde veya bilgiye maruz kalındığında yemek yemeye ilgiyi kaybetme ve beşinci sırada, Koronavirüs hakkında düşünüldüğünde veya bilgiye maruz kalındığında felç olmuş veya donmuş gibi hissetme sıkıntısının yaşandığı görülmektedir.

Tablo 2. Ölçek Sorularına Verilen Cevapların Dağılımı

Son 2 Hafta Boyunca Aşağıdaki Aktiviteleri Ne Sıklıkta Yaşadınız?	Hiç	Nadiren Bir Veya İki Günden Az	Birkaç Gün	7 Günden Fazla	Son 2 Haftada Neredeyse Her Gün
1.Koronavirüs ile ilgili haberleri okuduğumda veya dinlediğimde başım döndü, sersemleştim veya baygın hissettim.	155 (%72,4)	35 (%16,4)	19 (%8,9)	3 (%1,4)	2 (%0,9)
2.Koronavirüs hakkında düşündüğüm için uykuya dalma veya uykuda kalma konusunda sorun yaşadım.	133 (%62,1)	41 (%19,2)	32 (%15)	7 (%3,3)	1 (%0,5)
3. Koronavirüs hakkında düşündüğümde veya bilgiye maruz kaldığımda felç olmuş veya donmuş gibi hissettim.	190 (%88,8)	19 (%8,9)	2 (%0,9)	3 (%1,4)	0
4. Koronavirüs hakkında düşündüğümde veya bilgiye maruz kaldığımda yemek yemeye ilgimi kaybettim.	168 (%78,5)	27 (%12,6)	16 (%7,5)	2 (%0,9)	1 (%0,5)
5. Koronavirüs hakkında düşündüğümde veya bilgiye maruz kaldığımda mide bulantısı hissettim ve mide problemleri yaşadım.	166 (%77,6)	23 (%10,7)	21 (%9,8)	3 (%1,4)	1 (%0,5)

Orijinal koronavirüs anksiyete ölçeğindeki  $\geq 9$ ’luk kesme puanına göre, örneklerin %2,3’ü disfonksiyonel endişeli olarak sınıflandırılmıştır. Ancak KAÖ’ni geliştiren Lee(2020-b),  $\geq 9$ ’un orijinal kesme puanının çok sıkı olmasından dolayı, KAÖ’nin makul bir duyarlılık oranına sahip olabilmesi için kesme puanının  $\geq 5$ ’e düşürülmesi gerektiğini belirtmiştir. Çünkü orijinal KAÖ araştırması sadece koronavirüs hakkında anksiyete hisseden kişilerden oluşurken, diğer çalışmaların analizinde kullanılan örneklerde herhangi bir anksiyete önkoşulu aranmamaktadır.

Kesme puanı  $\geq 5$  alındığında ise, örneklerin %11,7’si kaygı bozukluğu olanlar olarak sınıflandırılmıştır. %88,3’ü ise kaygı bozukluğu olmayan kişiler olarak sınıflandırılmıştır. Cronbach  $\alpha$  değeri 0,835’ tir.

Yapılan istatistiksel analizler sonucu koronavirüse bağlı anksiyete belirtisinin cinsiyete, yaşa, gelir düzeyine, mesleğe, oturulan ev türüne, kişide kronik bir hastalığın var olmasına, evde kronik hastalığı olan kişi varlığına ve son 2 hafta içinde koronavirüs hakkında düşünme veya medya izleme süresine göre farklılaştığı ( $p < 0,05$ ) görülmüştür. Anksiyete varlığının anlamlı bulunan değişkenlere göre dağılımlarına bakılmıştır. Tablo 3’te görüldüğü gibi kadınlarda anksiyete varlığı daha fazladır.

Tablo 3. Anksiyetenin Cinsiyete Göre Dağılımı

Anksiyete	Cinsiyet		Toplam
	Kadın	Erkek	
Var	14	9	23
Yok	80	95	175
<b>Toplam</b>	94	104	198

Yaşa göre bakıldığında, anksiyete görülenlerin %48'i 41-64 yaş aralığında, %40'ı 18-40 yaş aralığında, %12'si ise 65 yaş ve üzerinde olanlar olduğu Tablo 4'te görülmektedir.

Tablo 4. Anksiyetenin Yaşa Göre Dağılımı

Anksiyete	Yaş			Toplam
	18-40	41-64	65 ve üzeri	
Var	10	12	3	25
Yok	134	53	2	189
<b>Toplam</b>	144	65	5	214

Tablo 5, anksiyetenin gelir düzeyine göre dağılımını göstermektedir. Buna göre, gelir düzeyi 4500 TL'den fazla olanlarda en fazla anksiyete görülmektedir.

Tablo 5. Anksiyetenin Gelir Düzeyine Göre Dağılımı

Anksiyete	Gelir Düzeyi				Toplam
	0-1500	1501-3000	3001-4500	4500den fazla	
Var	5	0	8	10	23
Yok	8	22	30	108	168
<b>Toplam</b>	13	22	38	118	191

Tablo 6'ya bakılarak, mesleklere göre anksiyete varlığının en fazla memurlarda görüldüğü söylenebilir.

Tablo 6. Anksiyetenin Mesleğe Göre Dağılımı

Anksiyete	Meslek							Toplam
	Ev Hanımı	Memur	Esnaf	Serbest Meslek	İşçi	Öğrenci	Diğer	
Var	4	12	1	2	3	1	2	25
Yok	6	88	1	12	24	11	45	187
<b>Toplam</b>	10	100	2	14	27	12	47	212

Anksiyete varlığı oturlan ev türüne göre Tablo 7'ye bakarak sıralandığında müstakil ev, apartman ve site şeklindedir.

Tablo 7. Anksiyetenin Oturlan Ev Türüne Göre Dağılımı

Anksiyete	Oturlan Ev Türü			Toplam
	Apartman	Site	Müstakil	
Var	8	5	10	23
Yok	119	42	17	178
<b>Toplam</b>	127	47	27	201

Tablo 8'de görüldüğü gibi, kronik hastalığı olan kişilerde anksiyete varlığı yaklaşık %35 iken, kronik hastalığı olmayanlarda ise anksiyete varlığı yaklaşık %9'dur.

Tablo 8. Anksiyetenin Kronik Hastalık Varlığına Göre Dağılımı

Anksiyete	Kronik Hastalık		Toplam
	Var	Yok	
Var	8	17	25
Yok	15	174	189
Toplam	23	191	214

Evde kronik hastalığı olan kişilerde anksiyete varlığı %23, evde kronik hastası olmayanlarda ise anksiyete varlığı yaklaşık %9 olduğunu Tablo 9'a bakarak söyleyebiliriz.

Tablo 9. Anksiyetenin Evde Kronik Hastalık Varlığına Göre Dağılımı

Anksiyete	Evde Kronik Hasta		Toplam
	Var	Yok	
Var	10	15	25
Yok	33	156	189
Toplam	43	171	214

Tablo 10'dan anksiyetenin en fazla son iki haftada 3-5 saat arası covid-19 hakkında düşününen veya medya izleyenlerde görüldüğü söylenebilir.

Tablo 10. Anksiyetenin Son İki Hafta Kaç Saat Covid-19 Hakkında Düşünüldüğü veya Medya İzlendiğine Göre Dağılımı

Anksiyete	Son iki hafta kaç saat covid-19 hakkında düşündünüz veya medya izlediniz				Toplam
	1-3 saat	3-5 saat	5-7 saat	7 saatten daha fazla	
Var	1	11	6	7	25
Yok	63	40	30	56	189
Toplam	64	51	36	63	214

## Sonuç

COVID-19'un psikolojik etkileriyle ilgili temel verileri bildiren ve sağlık uzmanları ile araştırmacılara bu küresel sağlık kriziyle mücadele çabalarına yardımcı olacak kısa bir akıl sağlığı taraması sağlayan KAÖ, bu çalışmada Sivas halkına uygulanmıştır. Uygulama sonucunda coronavirus ile ilgili olarak anksiyete yaşayanların örneklemin % 11,7 si olduğu belirlenmiştir.

Günümüzde uygulanan etkili ve yaygın olarak kullanılan sosyal mesafe önlemleri (Toplantıların iptal edilmesi - Yüz yüze eğitime ara verilmesi - İşyerlerinin kapatılması - Seyahatin kısıtlanması), karantinalarda toplumun genelinin yararı için toplumu oluşturan bireylerin kimi hak ve özgürlüklerinin geçici olarak kısıtlanması durumları ruh sağlığını etkiler ve mümkün olan en kısa zamanda ve azami özenle bireysel ve toplumsal müdahaleleri içeren ruhsal destek programlarının geliştirilmesini gerektirir (Türkiye Psikiyatri Derneği Ruhsal Travma ve Afet Çalışma Birimi).

COVID-19 salgınının, tüm dünya için mevcut akıl sağlığı hizmeti sunumunda ekstra bir sosyo-ekonomik yük haline geleceği, ayrıca küresel mental hastalık yükünü de artıracığı ifade edilebilir. Bu nedenle yapılacak daha geniş kapsamlı çalışmalar ile toplumların salgının neden olduğu ruh sağlığı durumları belirlenmeli ve ilgili kişilerce gereken politikalar üretilmelidir.

Salgın sürecinde alınan önlemlerin insanlar üzerinde yaratmış olduğu psikolojik baskıyı hafifletmek için farklı çözüm yollarına gidilebilir. Örneğin, Çin COVID-19'a yakalanan kişilerin

psikolojik sorunlarıyla başa çıkmak amacıyla, internet teknolojisi kullanımı yoluyla yeni bir psikolojik kriz müdahale modeli geliştirmiştir. Batı Çin Hastanesi'nin bu yeni modeli, hastalara, ailelerine ve sağlık personeline psikolojik müdahale yapmak için doktorları, psikiyatristleri, psikologları ve sosyal hizmet uzmanlarını internet platformlarına entegre ederek destek sağlamaktadır (Zhang vd., 2020).

### Teşekkür

Sherman A. Lee, “Koronavirüs Anksiyete Ölçeği'nin, klinik değerlendirme ve araştırmalarda kullanımını teşvik etmek amacıyla çalışmaya uygun şekilde atıfta bulunulmasının ötesinde, başkaları tarafından çoğaltılması ve kullanılması için resmi bir iznin gerekli olmadığını” belirtmektedir. Yazar bu açıklaması ile araştırmacılara ölçeğin kullanım iznini vermiştir. Bu sebeple kendisine teşekkür ederim.

### Kaynaklar

- Akkuzu, H., Yumuşak, F. N., Karaman, G., Ladikli, N., Türkkan, Z., & Bahadır, E., 2020. Koronavirüs Kaygı Ölçeği'nin Türkçe Güvenirlik Ve Geçerlik Çalışması. *Kıbrıs Türk Psikiyatri ve Psikoloji Dergisi*, 2(2), 63-67.
- Bıçer, İ., Çakmak, C., Demir, H., & Kurt, M. E. Koronavirüs Anksiyete Ölçeği Kısa Formu: Türkçe Geçerlik ve Güvenirlik Çalışması. *Anadolu Kliniği Tıp Bilimleri Dergisi*, 25(Supplement 1), 216-225.
- Evren, C., Evren, B., Dalbudak, E., Topcu, M., & Kutlu, N. (2020). Measuring anxiety related to COVID-19: A Turkish validation study of the Coronavirus Anxiety Scale. *Death Studies*, 1-7.
- Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., Liu, L., Shan, H., Lei, C., Hui, D. S. C., Du, B., Li, L., Zeng, G., Yuen, K.-Y., Chen, R., Tang, C., Wang, T., Chen, P., Xiang, J., ... Zhong, N. (2020). Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*, 382(18), 1708-1720.
- Kalat JW, Shiota MN. Emotion. Belmont, CA: Thompson Wadsworth, 2007.
- Lee, S.A., (2020-a). Coronavirus Anxiety Scale: A brief mental health screener for COVID-19 related anxiety. *Death Stud.* 44 (7). <https://doi.org/10.1080/07481187.2020.1748481>.
- Lee SA. Koronavirüs Anksiyete Ölçeği replikasyon analizi. *Düşünen Adam The Journal of Psychiatry and Neurological Sciences* 2020;33:00-00.
- Lee, S. A., Mathis, A. A., Jobe, M. C., & Pappalardo, E. A. (2020-b). Clinically significant fear and anxiety of COVID-19: A psychometric examination of the Coronavirus Anxiety Scale. *Psychiatry Research*, 113112.
- Türk Psikologlar Derneği Travma, Afet ve Kriz Birimi, <https://www.psikolog.org.tr/tr/blog/www/covid-19-salgini-sirasinda-ruh-sagligi-x656> [erişim tarihi: 01.09.2020]
- Xiang, Y., Yang, Y., Li, W., Zhang, L., Zhang, Q., Cheung, T., Ng, C.H., 2020. Timely mental health care for the 2019 novel coronavirus outbreak is urgently needed. *Lancet* 7, 228–229. [https://doi.org/10.1016/S2215-0366\(20\)30046-8](https://doi.org/10.1016/S2215-0366(20)30046-8).
- Zhang, J., Wu, W., Zhao, X., & Zhang, W. (2020). Recommended psychological crisis intervention response to the 2019 novel coronavirus pneumonia outbreak in China: a model of West China Hospital. *Precision Clinical Medicine*, 3(1), 3-8.

**Determination of Best Variance-Covariance Structure in Mixed Model (Sas Proc Mixed)  
with Various Parameter Estimation Methods**

Adile TATLIYER TUNAZ<sup>1</sup>

<sup>1</sup>Kahramanmaras Sutcu Imam University, Agriculture Faculty, Department of Animal Science  
Kahramanmaras, Turkey

**Abstract**

The aim of this study was to compare the covariance structures by using maximum likelihood (ML), restricted maximum likelihood (REML) and Minimum Variance Quadratic Unbiased Estimator (MIVQUE) in the estimation methods in repeated measures design with mixed model approach. In the study, live weight (birth, 30th, 60th, 90th, 120th day) values of 60 head Kilis goats from birth to 120 days old were used as research data. For the purpose of evaluate of the relationship among the data, Compound symmetry (CS), Variance components (VC), (First-order autoregressive (AR(1)), Unstructured (UN), Toeplitz (TOEP), Heterogenous compound symetry (CSH), Heterogenous first-order autoregressive (ARH(1)), Heterogenous toeplitz (TOEPH), First-Order Autoregressive Moving-Avarege (ARMA(1,1)), Toeplitz With Two Bands (TOEP(2)), First-order factor analytic (FA(1)), Equal Diagonal Factor Analytic (FA1(1)), Unstructured correlations (UNR), Banded Unstructured (UN(1)), Ante-Depence (ANTE(1)) covariance structures were used. The most appropriate covariance structure was selected according to  $2\text{Ln}(L)$ , AIC, AICC and BIC information criteria in modeling the relationship between data in all three estimation methods (ML, REML and MIVQUE0), UN and UNR covariance structures were determined as the most appropriate covariance structures, although they gave the same results.

**Key words:** General Lineer Mixed model, parameter estimations, variance-covariance structure

## Determination of Difference between Heavy Metal Values in Vegetables Growing in Soils Composed of Erciyes Strato Volcano by Fuzzy and Classical Variance Analysis

Kenan KILIÇ<sup>1</sup>, Derviş TOPUZ<sup>2</sup>

<sup>1</sup>Department of Environmental Engineering, Engineering Faculty, Niğde Ömer Halisdemir University, 51240, Niğde, Turkey

<sup>2</sup>Department of Health Services Science, Niğde Zübeyde Hanım Vocational School of Health Services, Niğde Ömer Halisdemir University, Niğde, Turkey

### Abstract

In the study, we aimed to determine whether there is a difference between heavy metals found in vegetables taken from different points within an area of approximately 2400 km<sup>2</sup> covering the east of Erciyes strato volcano. Due to the lack of homogeneous trial material with sufficient number of repetitions, we applied the fuzzy and classical random blocks trial plan method. Thus, we calculated based on the criterion of reducing errors based on uncertainties arising from differences between treatments and material inhomogeneity. The calculations made with the fuzzy method were calculated at the tolerance level of  $h = 0.5$  and under the uncertainties  $Z(x) = 218$ . According to classical and fuzzy variance analysis, it was found that the effect between products (pepper, tomato, bean, pumpkin, eggplant) was statistically significant ( $P < 0.05$ ) and between heavy metals (Cd, Co, Cr, Cu, Mn, Ni, Pb, Zn) It has been calculated that the efficacy for / products is significant at 5% error level compared to both methods. As a result of the calculations made by both methods, a positive correlation was found between potentially toxic zinc, chromium, copper and manganese elements in soils and vegetables, while a negative statistical relationship was found between cadmium at  $r_{Zn\text{vegetable-Soil}} = -0.53$ ,  $r_{Cr\text{vegetable-soil}} = 0.29$ ,  $r_{Cu\text{vegetable-soil}} = 0.40$  and  $r_{Mn\text{vegetable-soil}} = 0.71$  levels. There is a relationship between potential toxic (Cd, Co, Cr, Cu, Mn, Ni, Pb, Zn) elements related to the study area and potential toxic (Cd, Co, Cr, Cu, Mn, Ni, Pb, Zn) elements in vegetables. Fuzzy applications related to determination of absence are discussed for the first time.

**Key words:** Heavy metal, block, experiment plan, Fuzzy two-way analysis of Variance

**Using Path Analysis to Predict Egg Weight from Egg Physical Traits of Potchefstroom  
Koekoek Chicken Layer Breed**

T.L TYASI\*, V.R HLOKOE, K MOKOENA, K.M MOLABE, M.C MATHAPO, L.J SELALA,  
M MASWANA, L.T RASHIJANE

School of Agricultural and Environmental Sciences, Department of Agricultural Economics and  
Animal Production, University of Limpopo, Private Bag X1106, Sovenga 0727, Limpopo, South  
Africa

**Abstract**

Path analysis is a mathematical tool which is utilized to examine the cause effect relationship between independent and dependent variables. The objective of the study was to examine the direct and indirect effects of egg physical traits viz; egg length (EL), egg width (EWD), yolk weight (YW), albumen weight (AW) and shell weight (SW) on egg weight (EW). About 200 eggs of Potchefstroom Koekoek layer breed were used. Pearson correlation and path analysis were used. Correlation results indicated that EW had a positive highly significant correlation ( $P < 0.05$ ) with EWD, AW, SW and EL. Path analysis showed that shell weight (2.14) had a highest direct effect while egg length had a higher indirect effect on egg weight. In conclusion, path analysis results suggest that shell weight and egg length might be used as a selection criterion during breeding to increase egg weight of Potchefstroom Koekoek layer breed. Further studies are needed to be done using path analysis in large number of Potchefstroom Koekoek layer breed eggs or other chicken layer breed eggs.

**Key words:** Correlation, direct effect, egg length, indirect effect, shell weight

## Tarım İşletmelerinde Üretim Planında Değişikliğe Etki Eden Faktörlerin Belirlenmesi

Fatma ÇİFTÇİ<sup>1\*</sup>, İsmail ÇİFTÇİ<sup>2</sup>

<sup>1</sup>Konya PTT Baş Müdürlüğü, Konya, Türkiye

<sup>2</sup>Konya İl Tarım ve Orman Müdürlüğü, Konya, Türkiye

### Özet

Çalışmanın temel amacı tarım işletmelerinde üretim planında değişikliğe etki eden faktörlerin belirlenmesidir. Bu amaca yönelik olarak Konya ili Karapınar ilçesinde faaliyet gösteren tarım işletmeleri toplam popülasyonu oluşturmakta olup popülasyondan örnek çekmede tabakalı tesadüfi örnekleme yöntemi kullanılmış ve 72 anket yapılmıştır. Yapılan anketler sonucunda tarım işletmecilerinin üretim planında değişiklik yapmasına etki eden faktörlerin belirlenmesinde geçerlilik ve güvenilirlik testleri yapılmış olup 4 faktör belirlenmiştir. Bu faktörler arasındaki ikili değişkenler arasındaki birlikte değişim derecesini belirlemek için korelasyon analizi yapılmış olup faktörler arasında farklı önem seviyelerinde pozitif yönlü anlamlı ve yüksek bir ilişki olduğu belirlenmiştir. Son olarak belirlenen faktör yükleri ile üretim kararları arasındaki ilişki lojistik regresyon ile incelenmiş ve regresyon analizi sonucunda üretim kararları üzerinde iklimsel, ekonomik, sosyal ve politik faktörlerin etkili olduğu belirlenmiş ve öneriler hazırlanmıştır.

**Anahtar Kelimeler:** Regresyon Analizi, Üretim Planı, Selçuklu, Konya

### Abstract

The study is to determine the factors that affect the change in the production plan in the main agricultural enterprises. For this purpose, agricultural enterprises operating in Konya province Karapınar district constitute the total population, and the stratified random sampling method was asked and a survey was conducted to draw samples from the population. The questionnaires to be conducted are 4 questions that have been used to help determine the factors that affect the operators to make changes in the production plan and to be tested for reliability. There is a high positive and significant relationship between these factors at different levels of significance, and correlation analysis was performed to determine the difference between mutual variables among other factors at different significance levels. Finally, the relationship between their loads and production decisions was examined with logistic reg, and regression analysis revealed Programs and suggestions have been prepared in which climatic, economic, and political influences and policies are effective on decision making.

**Key words:** Regression analysis, production plan, Selçuklu, Konya

### Giriş

Bir tarım işletmesinin müteşebbisi için en önemli karar, bir üretim dönemi içerisinde işletmesinde hangi üretim faaliyetlerine yer vermesidir (Başar, 2004; Gökçe, 2004; Arıkan, 2016). Bu karar işletmenin karşılaştığı ekonomik şartlar, pazar koşulları, üretimde kullanılan girdilerin temini açısından önem arz etmektedir. Kâr amacı güden bir işletmenin elinde bulundurduğu sınırlı üretim faktörleri ile maksimum kar yapmayı amaçlaması hem sosyolojik hem de ekonomik bir düşünce yapısıdır. Bir üretim dönemi içerisinde bir işletmenin mevcut şartlarda en yüksek gelire ulaşabilmesi için hangi üretim faaliyetine ne kadar yer verileceği önemlidir. Bunun için işletme planlamasına ihtiyaç vardır (Günden ve Miran, 2008)



Tarım işletmelerinin planlanması, belirli bir periyotta, işletmede mevcut arazi, işgücü, sermaye gibi üretim faktörlerinin optimal şekilde kullanılması için faaliyetlerle ilgili kararları kapsar ve o işletmenin ne şekilde üretimde bulunması gerektiğinin yollarını ortaya koymaktadır (Rehber ve Tipi, 1993; Erkuş ve Demirci, 1996). İşletmelerin planlanması ve faaliyetlerin analizinde amaç, mevcut şartlarda mümkün olan en yüksek geliri elde edebilmek için teşebbüslerin kombinasyonunu ve işletme yöntemini tespitidir. İşletmecilikte ilk yapılacak iş olan bu belirlemeyi toplu bir şekilde gösteren programa plan denilmektedir (Şahin ve Miran, 2008a). İşletme planlamasında ya yeni kurulan bir işletmenin tam organizasyonu veya mevcut bir işletmenin reorganizasyonu söz konusudur. Ancak uygulamada işletme planı denilince, mevcut bir işletmenin reorganizasyonu akla gelmektedir. Çünkü yeni bir işletmenin kurulması durumuna sık rastlanılmamaktadır.

İyi bir işletme planı; çiftçi ve ailesinin işgücünün, işletmenin sermayesinin, üreticinin işletmecilik kabiliyetinin etkin biçimde kullanılmasına yardım eder. Planın, mevcut şartlarda mümkün olan en yüksek geliri elde edebilmek için, işletmede hangi ürünlerin ne ölçüde ve hangi yöntemlerle yetiştirileceğini, bu yetiştirme sürecinde üretim faktörlerinin optimal kombinasyonunu, yani kaynakların en ekonomik olarak nasıl değerlendirileceğini gösteren önceden hazırlanmış bir program olması zorunludur (Çiftçi ve ark., 2019). Kısaca planlama, üretim faktörlerinin organizasyonunu kapsamaktadır (Bölek ve ark., 2007; Şahin ve Miran, 2008b). Bu nedenle çalışmada üreticilerin üretim planlarına etki eden faktörlerin belirlenmesi amaçlanmıştır.

#### Metot

Çalışmanın temel amacı tarım işletmelerinde üretim planında değişikliğe etki eden faktörlerin belirlenmesidir. Bu amaca yönelik olarak Konya ili Karapınar ilçesinde faaliyet gösteren tarım işletmeleriyle anket yapılmıştır. Anket yapılırken tabakalı tesadüfi örnekleme yönteminden yararlanılmış olup işletmeler 3 tabakaya göre ayrılarak analizler yapılmıştır. Buna göre 0-50 dekar işletmelerde 6, 51-150 dekar işletmelerde 24 ve 151-+ dekar işletmelerde 42 anket olmak üzere toplamda 72 anket yapılmıştır.

Tarım işletmecilerin üretim planında değişikliğine etki eden faktörlerin belirlenmesi için öncelikle ölçeklerin geçerlilik ve güvenilirlik analizleri yapılmıştır. Güvenirlik analizi kapsamında Cronbach's Alpha değeri kullanılmış olup değer 1'e yakın olması istenmektedir. Geçerlilik analizi kapsamındaysa faktör analizi kullanılmıştır. Faktör analizinin sonucunda elde edilen veriler teoriye uygun olarak maddelenmiş ve dört faktör altında bir araya getirilmiştir (Bayramoğlu ve ark., 2019). Faktör analizinden sonra üretim planında değişikliğine etki eden faktörler ile üretim planı arasında korelasyon analizi yapılmış ve ikili değişkenler arasındaki ilişki incelenmiştir. Korelasyon katsayısının (r) hesaplanmasında aşağıdaki formülden yararlanılmaktadır. Korelasyon analizi sonucunda elde edilen katsayı 0.20'den küçük ise çok zayıf korelasyon, 0.20 ile 0.40 arasında ise zayıf korelasyon, 0.40-0.60 arası orta şiddetli korelasyon, 0.60-0.80 arası yüksek korelasyon ve 0.80'den büyük ise çok yüksek korelasyon olduğu yorumu yapılmaktadır (Kalaycı, 2010).

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{\sum X^2 - \frac{(\sum X)^2}{n}} \sqrt{\sum Y^2 - \frac{(\sum Y)^2}{n}}}$$

Lojistik regresyon analizi, bağımlı değişkenin kategorik bir yapıya sahip olduğu bağımsız değişkenlerin ise farklılık gösterdiği durumlarda kullanılmaktadır (Aktaş ve Erkuş, 2009). Bu çalışmada bağımlı değişken olarak ele alınan üretim planında değişiklik iki şıklı olmasından dolayı ikili (binary) lojistik regresyon analizi yapılmıştır. Üretim planında değişiklik 1: evet, 0: hayır olmak üzere bağımlı değişken sınıflandırılmış olup, kullanılan değişkenlerin kesikli olmasından dolayı normallik varsayımın

sağlanması zorunlu olmadığından kullanılması rahat olmaktadır. Lojistik regresyon analizi olasılık oranını temel olarak alırken, olasılık oranı bir olayın gerçekleşme veya gerçekleşmeme ihtimali üzerinde durmaktadır. Olasılık oranının doğal logaritması alınarak elde edilen lojistik regresyon modelinin parametrelerini tahmin ederken en yüksek olabilirlik (maximum likelihood) yöntemi yaygın olarak kullanılır (Berenson ve Levine, 1996). Böylece iki değişkenli lojistik regresyon modeli,

$$P(Y) = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)} = \frac{1}{1 + \exp(-\beta_0 - \beta_1 X)}$$

Şeklinde yazılır. Lojistik regresyon modeli yazıldıktan sonra modeldeki katsayılar,

$$\ln\left(\frac{P(Y)}{Q(Y)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

$$\frac{P(Y)}{Q(Y)} = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n} = e^{\beta_0} e^{\beta_1 X_1} e^{\beta_2 X_2} e^{\beta_n X_n}$$

şeklinde hesaplanır (Çokluk, 2010; Şenel ve Alatlı, 2014). Burada Q(Y), Q(Y) = 1- P(Y) şeklinde hesaplanır. Olasılık oranının OR = P(Y) / Q(Y) hatırlanacak olursa her bir parametrenin odds ratio değeri olasılık oranı olarak dikkate alınır. Bu değer bağımlı değişkenin, bağımsız değişkenin etkisiyle kaç kat daha fazla ve ya yüzde kaç ihtimalle görülmesi olasılığını açıklamaktadır (Özçomak ve ark., 2005).

## Tartışma

Çalışma kapsamında tarım işletmecilerinin üretim planında değişiklik yapmasına etki eden faktörlerin belirlenmesinde öncelikle kullanılan soruların geçerlilik ve güvenilirlik testlerinin yapılması gerekmektedir. Literatürde genel anlamda çoktan seçmeli veya likert soru tiplerinde güvenilirlik, Cronbach Alpha testiyle ölçülmektedir (Santos, 1999; Ercan ve ark., 2007; Kartal ve Dirlik, 2016). Bu nedenle tarım işletmecilerinin vermiş oldukları cevaplara ilişkin güvenilirlik testi yapılmıştır. Cronbach's Alpha testi kullanılan ölçekte yer alan soruların tarım işletmecileri tarafından doğru anlaşılıp anlaşılmadığını kontrol etmede yardımcı olmaktadır. Yapılan Cronbach Alpha testine ait sonuçlar tabloda verilmektedir. Kullanılan ölçeğin güvenilirlik analizinin yapılmasından sonra geçerlilik analizinin yapılması gerekmektedir. Literatürde yapısal geçerliliğin kullanılmasında faktör analizi kullanılmaktadır (Büyüköztürk, 2002; Aytaç ve Öngen, 2012). Fakat faktör analizi yapılmadan önce kullanılan ölçekteki verilerin uygun olup olmadığını belirlemek için Kaiser-Maier Olkin (KMO) testi yapılmalıdır. KMO testine ait sonuçlar tabloda verilmektedir.

Tablo 1. KMO ve Bartlett's Testi

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0,895
Bartlett's Test of Sphericity	Approx. Chi-Square	1024,031
	df	120
	Sig.	0,000

KMO testinin sonuçlarına göre verilerin faktör analizi için uygun olduğu anlaşılmıştır. Ayrıca KMO testi sonucunun %1 önem seviyesinde anlamlı olması kullanılan verilerin normal dağıldığını göstermektedir. KMO testi sonucunda kullanılan ölçeğin faktör analizine uygun olduğu belirlenmiş olup, faktör analizi ile toplam faktörleri faktör boyutlarına ayırarak daha az veri ile çalışılmasını sağlaması açısından önemlidir.

Faktör analizi sonucunda tarım işletmecilerinin üretim planında değişiklik yapmasına etki eden faktörlere ait faktör yükleri, varyans oranları ve güvenilirlik katsayıları tabloda verilmektedir. Yapılan

faktör analizi sonucunda tarım işletmecilerinin üretim planında değişiklik yapmasına etki eden faktörler iklimsel faktörler, ekonomik faktörler, sosyal faktörler ve politik faktörler olarak dört gruba ayrılarak incelenmiştir. Tabloya göre iklimsel faktör boyutuna ait faktör yükleri 0,773 ile 0,520 arasında değişmekte ve bağımlı değişkeni açıklama gücü ise %18,86 olarak hesaplanmıştır. Boyutun güvenilirlik katsayısı ise %0,857 olarak hesaplanmıştır. Ekonomik faktörler boyutunun faktör yükleri 0,772 ile 0,542 arasında değişmekte ve bağımlı değişkeni tek başına %17,464'ünü açıklamaktadır. Bu boyutun Cronbach's Alpha değeri ise 0,831 olarak belirlenmiştir. Sosyal faktörler boyutunda faktör yükleri 0,855 ile 0,528 arasında değişim göstermektedir. Bağımlı değişkeni tek başına açıklamadaki oranı %15,72'dir. Bu boyutun güvenilirlik katsayısı ise 0,818 olarak belirlenmiştir. Politik faktörler ise toplamda 2 sorudan oluşmakta olup, faktör yükleri 0,890 ile 0,593 arasında farklılık göstermektedir. Bu boyutun bağımlı değişkeni tek başına açıklama gücü %13,514 olurken, Cronbach's Alpha değeri 0,737 olarak bulunmuştur.

Tablo 2. Faktör Analizi

		Faktör Yükleri	Varyans Oranı	Güvenirlilik Katsayıları
<b>İklimsel Faktörler</b>	Düzensiz Yağış	0,789	19,269	0,927
	Dolu	0,776		
	Kuraklık	0,736		
	Dolu	0,721		
<b>Ekonomik Faktörler</b>	Ürün Fiyatında Dalgalanma	0,832	18,496	0,849
	Girdi Fiyatında Dalgalanma	0,812		
	Faiz Oranında Değişiklikler	0,789		
	İşletme Borçları	0,781		
	Finansman Bulabilme	0,774		
<b>Sosyal Faktörler</b>	Aile İçi İlişkiler	0,841	16,421	0,832
	İş Bölümü	0,798		
	Çiftçi Aile Nüfusundaki Değişim	0,771		
	Yabancı İşgücü Bulma Güçlüğü	0,742		
<b>Politik Faktörler</b>	Hükümet Politikalarında Değişiklik	0,890	14,423	0,764
	Destek ve Teşviklerin Yetersizliği	0,821		

Faktör analizi sonucunda tarım işletmecilerin üretim planını etkileyen faktörler 4 faktör boyutu altında oluşturulmuştur. Bu boyutlar arasındaki ikili değişkenlerin değişim derecesini belirlemek için korelasyon analizi yapılmıştır (Ağızan ve ark., 2019). Korelasyon analizinde sonuçlar +1 ile -1 arasında değişmekte ve işaretin pozitif olması iki değişken arasında pozitif ve aynı yönlü bir ilişki olduğunu, tersi durumunda ise negatif veya ters yönlü bir ilişki olduğu söylenmektedir.

Tablo 3. Korelasyon Analizi ile Üretim Planını Etkileyen Faktörlerin Belirlenmesi

		Üretim Planı	İklimsel Faktörler	Ekonomik Faktörler	Sosyal Faktörler	Politik Faktörler
<b>Üretim Planı</b>	Pearson Korelasyon	1				
	Önem Düzeyi (2-tailed)					
<b>İklimsel Faktörler</b>	Pearson Korelasyon	0,121*	1			
	Önem Düzeyi (2-tailed)	0,023				
<b>Ekonomik Faktörler</b>	Pearson Korelasyon	0,270**	0,705**	1		
	Önem Düzeyi (2-tailed)	0,004	0,000			
<b>Sosyal Faktörler</b>	Pearson Korelasyon	0,301	0,532**	0,587**	1	
	Önem Düzeyi (2-tailed)	0,099	0,000	0,000		
<b>Politik Faktörler</b>	Pearson Korelasyon	0,482*	0,569**	0,632**	0,531**	1
	Önem Düzeyi (2-tailed)	0,003	0,020	0,003	0,001	
**. Correlation is significant at the 0.01 level (2-tailed).						
*. Correlation is significant at the 0.05 level (2-tailed).						
***. Correlation is significant at the 0.10 level (2-tailed).						

Yapılan korelasyon analizi sonucunda üretim planını etkileyen faktörler arasında farklı önem seviyelerinde pozitif yönlü anlamlı ve yüksek bir ilişki olduğu belirlenmiştir. Tarım işletmecilerinin üretim planı kararı ile bu kararın değişmesinde etkili olan faktörler arasındaki ilişki incelendiğinde ise boyutlar arasında %1 ile %10 arasında anlamlı ve pozitif bir yönlü ilişki belirlenmiştir.

Tarım işletmelerinde üretim planlarına etki eden faktörler belirlendikten sonra ilişkilerin tespit edilmesi için lojistik regresyon analizi yapılmıştır. Lojistik regresyon analizi, bağımlı değişkenin kategorik bir yapıda olduğu zamanda kullanılmakta olup, çalışma kapsamında bağımlı değişkenin 0 ve 1 olmak üzere 2 şıklı olmasından dolayı ikili(binary) lojistik regresyon analizi kullanılmıştır (Kendirli ve Çankaya, 2016). Buna göre modelde yer alacak bağımlı değişken tarım işletmelerinin üretim planında değişiklik olarak belirlenmiştir. Bağımsız değişkenler ise iklimsel, ekonomik, sosyal ve politik faktörler olarak belirlenmiştir. Bağımlı değişken olarak alınan üretim planında değişikliklerin yapılması “1” ve üretim planında değişiklik yapmayanlar “0” olarak kodlanmıştır. Yapılan lojistik regresyon analizi sonucunda bağımsız değişkenlerin bağımlı değişkeni açıklama gücü olan R2 %68 olarak belirlenmiştir. Elde edilen bu değer ile bağımsız değişkenlerin bağımlı değişkeni açıklama gücü %68 olarak bulunmuştur. Bulunan bu değer anlamlı olup olmadığını belirlemek için F testi yapılmaktadır. F testi modelin geçerli olup olmadığını göstermektedir. Yapılan F testi sonucunda modelin %1 anlam düzeyinde istatistiki açıdan önemli olduğu belirlenmiştir.

Tablo 4. Modeldeki Değişkenlerin Katsayıları

	Katsayılar	Standart Hata	Wald	df	Önem Düzeyi	Veri Olasılık Değeri	
Step 1 <sup>a</sup>	İklimsel Faktörler	2,258	1,546	1,232	1	0,098	1,232
	Ekonomik Faktörler	1,221	0,752	2,554	1	0,093	0,875
	Sosyal Faktörler	2,845	0,842	2,245	1	0,045	2,523
	Politik Faktörler	1,143	0,745	1,235	1	0,008	6,423
	Constant	-23,215	6,802	25,112	1	0,000	0,000

Modelde yer alan değişkenlere ait katsayıların ve bunların önem düzeyleri tabloda verilmiştir. Buna göre bağımsız değişkenler iklimsel, ekonomik, sosyal ve politik faktörler göstergelerinin hepsi istatistiki olarak anlamlı belirlenmiştir. Ayrıca modelde yer alan bağımsız değişkenlerde meydana gelecek olan değişmelerin bağımlı değişkeni nasıl etkileyeceği veren olasılık değeri Exp(B) ile gösterilmektedir. Örneğin; iklimsel faktörler değişkeninin olasılık değeri 1,232 olarak belirlenmiştir. Yani iklim değişikliğinde meydana gelecek %100'lük bir artış ile üretim planının değişme olasılığı %23 oranında artmaktadır. Ya da iklim değişikliği bir kat artarsa, üretim planında değişiklik olasılığını da 2,880 kat artacaktır. Nitekim tarım sektörü iklime yüksek bağlılık gösteren bir sektör olup iklimsel faktörlere bağlı olarak rekolte kayıpları yaşanmakta ve dolayısıyla da fiyat dalgalanmaları sıklıkla görülmektedir. Bir diğer değişken olan ekonomik faktörde ise durum benzerlik göstermektedir. Buna göre ekonomik faktörlerde meydana gelecek %100'lük bir artış ile üretim planının değişme olasılığı %14 artması beklenmektedir. Özellikle girdi maliyetlerinde yükselme ve işletmelerdeki sermaye devir hızlarının düşük olması tarımsal üretimdeki planlamaları sınırlandırmaktadır. Bu kapsamda potansiyel sermayeye uygun ve girdi maliyetlerini minimize edecek üretim faaliyetlerine yer verilmesi gerekmektedir. Sosyal ve politika faktörlerin de benzer özellikler görülmekte olup sosyal faktörlerde meydana gelecek %100'lük bir artış üretim planındaki değişikliği %152 artırma olasılığı bulunmaktadır. Aynı şekilde politik faktörlerde de meydana gelecek %100'lük bir artış üretim kararlarını %542 oranında değiştireceği sonucuna ulaşılmaktadır.

Sonuç olarak üretim kararları üzerinde politik faktörlerin daha etkili olduğu söylenebilmektedir. Özellikle devlet desteklemelerinin, teşviklerin ve yatırım kararlarının bölge ve işletme planlamaları için

önemli olduğu belirlenmiş olup, işletme ölçeklerine, sermayesine, iklimine ve toprak şartlarına uygun üretim planlamalarının yapıldığı ifade edilmektedir.

## Sonuç

Planlama, Yönetimin bir fonksiyonu olarak işletmelere müdahale etmenin bir şeklidir. İşletmelerin amaçlarına ulaşması için birden fazla seçenek olabilir. İşletmenin amacına ulaşmasında en kısa zamanda ve en etkin kaynak kullanımı ile hangi seçeneğin uygulanacağını belirlemek planlama olarak ifade edilmektedir. Tarım işletmeleri de ekonomik bir birim olarak tanımlanmakta olup, kar amacına yönelik faaliyet göstermektedirler. Nitekim karı maksimize etmek için tarım işletmeleri sahip oldukları kaynakları üretim faaliyetlerine göre dağılımını yapması gerekmektedir. Hangi üretim faaliyetine ne kadar yer verileceği planlama fonksiyonu içerisinde yer almaktadır.

Üretim planlaması yapılabilmesi için öncelikle üretim maliyetlerini hesaplanması gerekmektedir. Bunun yanında kredi kullanımı ve geri ödemesinin planlanması, yatırımların planlanmasında üreticiler üretim maliyetlerine ihtiyaç duymaktadırlar. Bu kapsamda illere ve bölgelere göre üretim maliyetlerinin yanı sıra işletme bazlı üretim maliyetlerinin hesaplanması gerekliliği ortaya çıkmaktadır.

Üretim maliyetlerinin yanı sıra; iklimsel, ekonomik, sosyal ve politik faktörleri bir bütün olarak dikkate alarak üretim bölgelerinin kararlaştırılması gerekmektedir. Bu bölgelerde ekolojik istiklaliklerine ve ekonomik getirisi yüksek olan ürünlerin yetiştiriciliği planlanması gerekmektedir. Ayrıca bu bölgelerde faaliyet gösteren işletmelere yöneltilebilir nitelikte desteklemelerin ve yatırım desteklerinin verilmesi gerekmektedir.

Son olarak üretim planlamalarında yardımcı olacak şekilde üreticilerin yönetim, finansman, girdi kullanımı, teknik bilgi vb. konularda eğitim ve yayım faaliyetlerine katılımlarının sağlanması gerekmektedir. Bu kapsamda bu bölgelerde ihtiyaçların karşılanması doğrultusunda üreticilere eğitim çalışmalarını düzenlenmelidir.

## Kaynaklar

- Ağızhan, K., Bayramoğlu, Z., Bozdemir, M. , Ağızhan, S. (2019). Beş Faktör Kişilik Özelliklerinin Kadın Girişimciliğine Etkisi: Konya İli Örneği. XI. IBANESS Congress Series – Tekirdag/Turkey. Edirne.
- Aktaş, C., Erkuş, O. (2009). Lojistik Regresyon Analizi ile Eskişehir'in Sis Kestiriminin İncelenmesi, İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 8 (16), 47-61.
- Arıkan, C. (2016). Kadın Girişimcilikte Başarı ve Başarıyı Etkileyen Faktörler: Bursa Örneği, Yönetim ve Ekonomi Araştırmaları Dergisi, 14 (3), 138-156.
- Aytaç, M. , Öngen, B., (2012). Doğrulayıcı faktör analizi ile yeni çevresel paradigma ölçeğinin yapı geçerliliğinin incelenmesi, İstatistikçiler Dergisi: İstatistik ve Aktüerya, 5 (1), 14 22.
- Başar, M., (2004). Girişimcilik ve Girişimcinin Özellikleri, Eskişehir, Açık Öğretim Fakültesi Yayını No:824.
- Bayramoğlu, Z., Eroğlu, O. , Ağızhan, K., (2019). Tarım İşletmesi Yöneticilerinin Geleceğe Yönelik Beklentilerinin Analizi. I.International Science and Academic Congress. Konya: 249-259.
- Berenson, M. , Levine, D., (1996). Estadística básica en administración, concepto y aplicaciones, MÃ©xico Editorial Prentice-Hall Hispanoamérica.
- Bölek, Y., Oğlakçı, M. , Kılılı, F., (2007). Pamukta (Gossypium spp.) erkenciliği belirleyen faktörler ve üretim planlaması, KSÜ Fen ve Mühendislik Dergisi, 10 (1), 116-125.
- Büyüköztürk, Ş., (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı, Kuram ve Uygulamada Eğitim Yönetimi Dergisi, 8 (4), 470-483.
- Çiftçi, K., Miran, B. , Kanberoğlu, Z., (2019). Çiftçilerin Planlı Üretim Yapma Eğilimlerini Artıran Yaklaşımların Belirlenmesi İzmir İli Örneği, Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi, 29 (1), 82-87.

- Çokluk, Ö., (2010). Lojistik regresyon analizi: Kavram ve uygulama, Kuram ve Uygulamada Eğitim Bilimleri, 10 (3), 1357-1407.
- Ercan, I., Yazici, B., Sigirli, D., Ediz, B., Kan, I., (2007). Examining Cronbach alpha, theta, omega reliability coefficients according to sample size, Journal of modern applied statistical methods, 6 (1), 27.
- Erkuş, A. , Demirci, R., (1996). Tarımsal İşletmecilik ve Planlama (Geniştirilmiş ve Gözden Geçirilmiş II. Baskı), Ankara Üniv. Ziraat Fak. Ders Kitabı, Yayın, 1435.
- Gökçe, O., (2004). Tarımsal Girişimcilik, Çiftçi ve Köy Dünyası Dergisi, Sayı.
- Günden, C., Miran, B., (2008). Bölge Bazında Uygulamaya Elverişli Esnek Üretim Planlarının Bulanık Doğrusal Programlama Yöntemiyle Elde Edilmesi Üzerine Bir Araştırma, Ege Üniversitesi Ziraat Fakültesi Dergisi, 45 (2), 113-124.
- Kalaycı, Ş., (2010). SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri, Ankara, Asil Yayın Dağıtım, p. 426.
- Kartal, S.K., Dirlık, E.M., (2016). Geçerlik kavramının tarihsel gelişimi ve güvenilirlikte en çok tercih edilen yöntem: Cronbach Alfa Katsayısı, Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 16 (4).
- Kendirli, S., Çankaya, M., (2016). Bist Turizm Endeksindeki Şirketlerde İşletme Sermayesi Yönetiminin Karlılık Üzerindeki Etkisini Ölçmeye Yönelik Bir Araştırma, International Review of Economics and Management, 4 (2), 46-68.
- Özçomak, M.S., Oktay, E., Özer, H., (2005). Erzurum ilinde potansiyel doğalgaz talebini etkileyen faktörlerin tespiti, VII. Ulusal Ekonometri ve İstatistik Sempozyumu, 26-27.
- Rehber, E., Tipi, T., (1993). Tarımsal işletmecilik ve planlama, Uludağ Üniversitesi Güçlendirme Vakfı Yayın (84), 31.
- Santos, J.R.A., (1999). Cronbach's alpha: A tool for assessing the reliability of scales, Journal of extension, 37 (2), 1-5.
- Şahin, A., Miran, B., (2008a). Tarım İşletmelerinde Üretim Dallarına Göre İşgücü Kullanımının Çok Boyutlu Analizi: Bayındır İlçesi Örneği, Ege Üniversitesi Ziraat Fakültesi Dergisi, 45 (1), 37-47.
- Şahin, A., Miran, B., (2008b). Risk koşullarında tarım işletmelerinin planlanması: oyun teorisi yaklaşımı, Hayvansal Üretim, 51 (1).
- Şenel, S., Alatlı, B., (2014). Lojistik regresyon analizinin kullanıldığı makaleler üzerine bir inceleme, Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 5 (1), 35-52.

## **Bread Consumption Habits of Selcuklu District of Konya Province and Willingness to Pay**

Fatma ÇİFTÇİ<sup>1</sup>

<sup>1</sup>Konya PTT Baş Müdürlüğü, Konya, Türkiye

### **Abstract**

The main purpose of this study is to determine the bread consumption and habits of the households living in Konya. For this purpose, a survey was conducted with 101 households determined in the Selcuklu district of Konya province according to the proportional sampling method. The classification made by taking into consideration the in- come criteria that was interpreted by analyzing the survey results. Accordingly, the distribution of households by income groups is similar. At the same time, the average price of bread was 1,10 TL with a rate of 68%, and the preferred bread weight was 200 g with a ratio of 43%. Although 58% of consumers change according to income groups, they stated that bread is of good quality. It has been determined in the scope of the study that daily bread consumption amount is 1-3, with an average rate of 78%, 40% ovens are preferred in the selection of bread purchase areas, and special attention is paid to quality when buying bread.

**Key words:** Bread, bread consumption, consumption habits, Selcuklu, Konya

### **Introduction**

Bread was first produced in Egypt in the 4000s BC and has been one of the most consumed foods with an increasing consumption until today (Parimala and Sudha, 2015). Bread, which has an important place in our tables, is a basic food item due to its features such as simple production technology, being cheaper and easier to provide than other foodstuffs, and also being nutritious and satisfying. According to the Turkish Food Codex Regulation "Bread and Bread Types Communiqué, the definition of bread; It is a product made to bread wheat flour with drinkable water, salt, yeast and additives permitted in the "Turkish Food Codex Regulation, if necessary (Anonymous, 2012).

Bread consumption varies depending on some factors such as gender, age, profession, income levels and habits of individuals (Dağlıoğlu, 1998). Bread consumption varies according to different regions, age and income groups. Turkey in regional consumed in bread quantity of the Black Sea region of 356 g / person / day, Thrace, Marmara, Aegean Coast of 391,2 g / person / day, in Central Anatolia 407,2 g / person / day, South Anatolia and Eastern Anatolia It has been determined as 442,4 g / person / day in the region (Doğan, 2003). Soil Products Office report results according to his research center in 12 provinces in Turkey in 2008 and the average per capita bread consumption in Turkey was determined to be approximately 333 g (TMO, 2008).

The main purpose of this study is to determine the bread consumption and habits of households in Konya Selçuklu District, and to determine the willingness of consumers to pay for bread . A survey has been conducted for this main purpose and the results obtained have been evaluated by being subjected to statistical analysis. In the study, the consumption amount of bread, which has a great place and importance in our daily life, and the criteria taken into consideration in bread preference were examined.

## Material and Method

This study was carried out in 2019 to determine the bread consumption habits of households living in Konya Province Selcuklu District. The data used in the study belong to 101 households. The following proportional sampling method was used to determine the sample volume. According to these criteria, the sample volume was found to be 92 with 95% confidence interval and 10% margin of error. However, the number of surveys was completed to 101 in order to increase the accuracy of the data. Within the scope of the study, the socio-economic characteristics of the consumers, their income levels, bread consumption behaviors were determined according to the percentages method, and the factors that consumers pay attention to while buying bread and the factors that negatively affect individuals in the sale of bread were examined with Likert-type questions. In addition, the Conditional Assessment method was used to determine the willingness of consumers to pay extra against competitive products. The conditional assessment method used in the research is basically a survey method. In the application of the method, by creating a hypothetical market for any environmental goods or services that cannot be bought or sold in the market, the benefits to be obtained by the people from the said goods or services are presented to the people selected for the survey with a scenario and in return for the benefit obtained from the use or consumption of this good or service. It is learned how much they want to pay (Carson, 2000; Yildirim, 2014). In this context, Willingness To Pay (WTP) method was used to determine the willingness of consumers to pay. In general, the function of the Willingness To Pay (WTP) curve developed in the EBM studies is as shown below (Maalouf et al., 2004).

$$WTP = \beta'x + \varepsilon$$

$$\begin{aligned}
 WTP &= \\
 &= 1 \text{ ise } \mu_0 < WTP \leq \mu_1, & (1) \\
 &= 2 \text{ ise } \mu_1 < WTP \leq \mu_2, \dots\dots\dots \\
 &= 10 \text{ ise } \mu_9 \leq WTP
 \end{aligned}$$

Here  $\mu$  is a parameter representing a threshold that separates categories in the observed variable. Probability can be written after the equations.

$$\text{Prob}(WTP = 0) = \Phi(-\beta'x) \dots\dots\dots$$

$$\text{Prob}(WTP = 10) = 1 - \Phi(\mu_{10} - \beta'x) \quad (2)$$

Willingness to pay is listed in eleven (11) categories in the study. Accordingly, those who do not want to pay at all (0), those who want to pay 10% (1), those who want to pay 20% more (2), those who want to pay 30% more (3), those who want to pay 40% more (4), those who want to pay 50% more (5), those who want to pay 60% more (6), those who want to pay 70% more (7), those who want to pay 80% more (8), those who want to pay 90% more (9), and those who want to pay 100% more (10).

## Research Findings and Discussion



### **Socio-Economic Characteristics of the Families Examined**

Bread is among the most consumed food products by households. Bread has an important place in the households' budget in both macro and micro aspects. Households public budget important a place must be the inflation bread causes you to have the highest weight in the basket. As a matter of fact, in many studies, the share of bread in household budget was determined as 11.93% (DİE, 2004). In this context, the bread consumption behavior of households by income in work were examined.

First, social characteristics of households were examined in the study. As a matter of fact, social characteristics are among the most important factors determining consumption behavior. In this context, the average household size of the households in Selcuklu district of Konya province was determined as 3.41 persons. 48.55% of the population is men and 51.45% is women. In addition, when the household population structures by income groups are examined, it was determined that there are differences between the groups. As a matter of fact, the population has increased in households with a monthly income of up to 6,500 TL, while the population has decreased in those with a monthly income of more than 6,500 TL.

One of the most important parameters of the modern world is education. The importance of education is increasing day by day. However, educational status positively affects consumption habits. As the educational status of the consumers increases, their consumption habits become more conscious and their desire for healthy eating increases. In addition, the most important factor determining income is education. As can be seen in Table 2, consumers with a high level of education have a high income level. Within the scope of the study, it was determined that the higher the education level, the higher the income level and the parallel increase in the income level decreased the bread consumption. While the average daily bread consumption in the household is 2 if the population is mainly university and high school graduates, the average daily bread consumption is more than 6 if the population is mainly primary and secondary school graduates.

Depending on the income level of consumers, the price paid for bread differs. It has been determined that the higher the income level, the higher the price paid for bread. Households (70%) with a monthly income of up to 6.500 TL pay 1,1 TL for bread, while households with a monthly income higher than 6.500 TL (50%) pay 1,25 TL (Figure 1). Households with high education and income prefer to pay 13,63% more for bread than households with low income. It is known that food safety (quality, packaging, hygiene, health, etc.) is the main reason why households with a monthly income higher than 6.500 TL pay 1,25 TL or more for bread. In addition, when the weight of the preferred breads was examined, it was determined that the consumers preferred maximum 200 grams of bread.

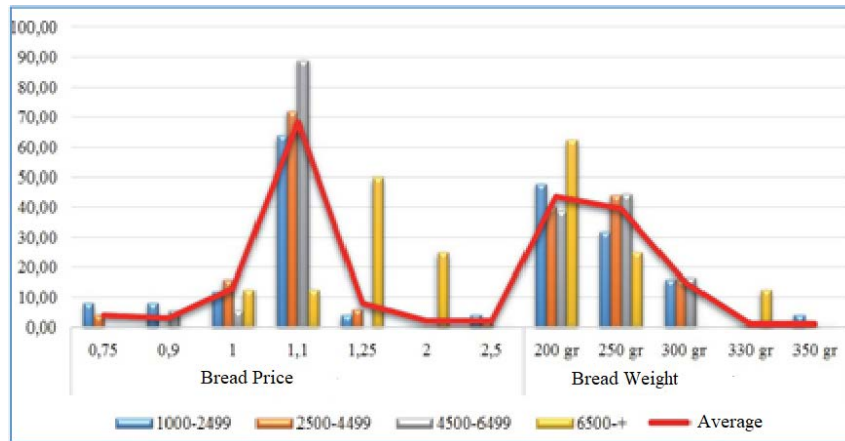


Figure 1. Price and Weight of Bread Purchased

As in every product, quality is very important in bread consumption. Consumers' desire to consume healthy products increases and quality bread production is encouraged accordingly. However, the opinions of the households regarding the quality of the breads produced under the current conditions differ. Within the scope of the study, 58,42% of households thought that the quality of bread was sufficient, while 41,58% of them thought that the quality of the bread was not sufficient (Table 1). As a matter of fact, according to the results of a similar study, 47,12% of consumers think that the quality of the bread is sufficient, while 52,88% think that it is poor quality (Yurdatapan, 2014).

Table 1. Quality Assessment of Bread (Number, %)

Income Groups (TL)	Yes		No		Total	
	Number	Rate	Number	Rate	Number	Rate
1000-2499	12,00	48,00	13,00	52,00	25,00	100
2500-4499	35,00	70,00	15,00	30,00	50,00	100
4500-6499	6,00	33,33	12,00	66,67	18,00	100
6500+	6,00	75,00	2,00	25,00	8,00	100
Average	21,84	58,42	12,94	41,58	34,78	100

Bread consumption has a very important place in human nutrition. When 200 grams of bread is consumed a day, on average, 35 percent of the energy required for a person, 25 percent of protein, 66 percent of vitamin B1 and 55 percent of the fiber need are met. At the same time, the most vital nutritional sources of a person such as magnesium, potassium, sodium, vitamin A, vitamin E, vitamin B1, vitamin B2, vitamin B6, niacin, copper, zinc, iron and phosphate are provided by bread consumption (Anonymous, 2020). Within the scope of the study, the amount of bread consumption of consumers with high income levels is decreasing. While the amount of bread consumed per person per day in low and middle income households is on average 220 g, the daily bread consumption of consumers in high-income households is 180 g. It is known that a healthy individual should consume an average of 6-7 slices of bread and each slice is 28 grams. In this context, it is said that a healthy individual should consume an average of 182 gr between 168 and 196 gr (Anonymous, 2019). In addition, it is determined that households with high income levels meet the nutrients contained in bread from other foods, and the rate of encountering health problems caused by excessive bread consumption is also low.

Table 2. Daily Amount of Bread Consumed in Households (Number, %)

Factors		Income Groups (TL)				
		1000-2499	2500-4499	4500-6499	6500-+	Average
1-3 (Piace)	Number	17,00	38,00	16,00	8,00	26,50
	Rate	68,00	76,00	88,89	100,00	78,22
4-6 (Piace)	Number	6,00	11,00	2,00	-	7,29
	Rate	24,00	22,00	11,11	-	18,81
7-10 (Piace)	Number	1,00	1,00	-	-	0,74
	Rate	4,00	2,00	-	-	1,98
10-+ (Piace)	Number	1,00	-	-	-	0,25
	Rate	4,00	-	-	-	0,99
Total	Number	25,00	50,00	18,00	8,00	34,78
	Rate	100,00	100,00	100,00	100,00	100,00

One of the most important criteria affecting the consumption habits of consumers is the places where they buy the products. Reasons such as hygiene conditions and healthy production affect the purchasing decisions of individuals. Within the scope of the study, bakeries are the leading places where households prefer to buy bread with a rate of 40,50%. The most important reason to prefer ovens is quality. In other studies, it was determined that 64,1% of consumers bought bread from bakeries, 15,4% from grocery stores, 14,6% from markets and 1% from supermarkets (İnan, 2013).

Table 3. Where to Buy Bread (Number, %)

Where to Buy Bread		Income Groups (TL)				
		1000-2499	2500-4499	4500-6499	6500-+	Average
Bakery	Number	7,00	17,00	11,00	6,00	12,58
	Rate	28,00	34,00	61,11	75,00	40,59
Grocer	Number	12,00	17,00	3,00	1,00	12,00
	Rate	48,00	34,00	16,67	12,50	32,67
Market	Number	5,00	11,00	2,00	1,00	7,12
	Rate	20,00	22,00	11,11	12,50	18,81
Supermarket	Number	-	3,00	2,00	-	1,84
	Rate	-	6,00	11,11	-	4,95
I'm Doing it Myself	Number	-	1,00	-	-	0,50
	Rate	-	2,00	-	-	0,99
Other	Number	1,00	1,00	-	-	0,74
	Rate	4,00	2,00	-	-	1,98
Total	Number	25,00	50,00	18,00	8,00	34,78
	Rate	100,00	100,00	100,00	100,00	100,00

The reasons why consumers prefer bread purchasing places are given in Table 6. In terms of accessibility, it was determined that being close to home was the most preferred reason at 42,57%, while price and habits constituted the reasons to prefer it at 7,92%. In a study conducted in the province of Istanbul, it is seen that it constitutes the reason for habituation at the rate of 19,60% and the reason for choosing price at the rate of 3,50% (Ertürk, et al., 2010).

Table 4. The Reason For Choosing Where To Buy Bread (Number, %)

Factors		Income Groups (TL)				
		1000-2499	2500-4499	4500-6499	6500-+	Average
Price	Number	4,00	2,00	1,00	1,00	2,24
	Rate	16,00	4,00	5,56	12,50	7,92
Quality	Number	6,00	19,00	11,00	4,00	13,17
	Rate	24,00	38,00	61,11	50,00	39,60
Habit	Number	1,00	5,00	2,00	-	3,08
	Rate	4,00	10,00	11,11	-	7,92
Close to Home	Number	14,00	22,00	4,00	3,00	15,31
	Rate	56,00	44,00	22,22	37,50	42,57
Other	Number	-	2,00	-	-	0,99
	Rate	-	4,00	-	-	1,98
Total	Number	25,00	50,00	18,00	8,00	34,78
	Rate	100,00	100,00	100,00	100,00	100,00

Among the most important criteria affecting the bread preference of households is the relationship between price and health. White bread is preferred more than other bread types by households with low and middle income because it contains easily accessible nutrients in terms of raw materials and is cheaper than other bread types in terms of price. Whole wheat bread is healthier than other bread types with its natural nutrients. It was determined within the scope of the study that it is preferred more by high-income households because it is more expensive than other bread types in terms of price.

Table 5. Preferred Bread Types (Number, %)

Factors		Income Groups (TL)				
		1000-2499	2500-4499	4500-6499	6500-+	Average
White bread	Number	21,00	40,00	10,00	4,00	27,10
	Rate	84,00	80,00	55,56	50,00	74,26
Wholemeal bread	Number	-	4,00	1,00	-	2,16
	Rate	-	8,00	5,56	-	4,95
Whole White Bread	Number	2,00	4,00	7,00	4,00	4,04
	Rate	8,00	8,00	38,89	50,00	16,83
Tandoori Bread	Number	-	2,00	-	-	0,99
	Rate	-	4,00	-	-	1,98
Other	Number	2,00	-	-	-	0,50
	Rate	8,00	-	-	-	1,98
Total	Number	25,00	50,00	18,00	8,00	34,78
	Rate	100,00	100,00	100,00	100,00	100,00

Bread consumption status varies according to the consumption habits of individuals. When the non-consumption of bread is examined, it is seen that mold with the rate of 54,46% is the cause of the most mold and the least distortion of its appearance with the rate of 2,97%. In a research conducted in the province of Isparta, molding, change in taste, hardening or crumbling, in turn, were determined as indicators that the bread was too stale to be eaten (Ertürk, A et al., 2010).

Table 6. The Condition of Not Consuming Bread (Number, %)

Factors		Income Groups (TL)				
		1000-2499	2500-4499	4500-6499	6500+	Average
If Moldy	Number	13,00	28,00	8,00	6,00	18,98
	Rate	52,00	56,00	44,44	75,00	54,46
If the taste has changed	Number	4,00	4,00	1,00	-	3,15
	Rate	16,00	8,00	5,56	-	8,91
If Hardened and Crumbled Easily	Number	6,00	13,00	7,00	1,00	9,25
	Rate	24,00	26,00	38,89	12,50	26,73
Image Is Distorted	Number	1,00	1,00	-	1,00	0,82
	Rate	4,00	2,00	-	12,50	2,97
If the Day It Is Not Wasted	Number	1,00	4,00	2,00	-	2,58
	Rate	4,00	8,00	11,11	-	6,93
Total	Number	25,00	50,00	18,00	8,00	34,78
	Rate	100,00	100,00	100,00	100,00	100,00

Households peoples while bread attention that matters Table 9 5's indentation scaling with consumers when evaluated (4,79/5) for the best cooked not be, (4,60/5) of them hygienic attention to be sold under the conditions determined they were.

Table 7. Considerations When Buying Bread (Number, %)

Factors	5	4	3	2	1	Average	Rate (%)
When hygienic conditions are produced	3,76	0,63	0,09	0,08	0,02	4,58	91,68
When sold in hygienic conditions	3,71	0,71	0,12	0,04	0,02	4,60	92,08
Well cooked	4,06	0,71	-	0,02	-	4,79	95,84
To be warm and fresh	2,33	1,35	0,30	0,14	0,03	4,14	82,77
Appearance	3,07	1,23	0,12	0,06	0,01	4,49	89,70
Without additives	3,27	0,75	0,15	0,14	0,04	4,35	86,93
To be nutritious	3,12	1,15	0,15	0,06	0,01	4,49	89,70
To be cheap	1,29	0,71	0,30	0,36	0,29	2,94	58,81

(5: It Bothers so much, 4:Bothers, 3:No Idea, 2:Insignificant, 1: It doesn't matter at all)

The reasons disturbing the consumers in the sale of bread When evaluated with 5-point Likert scaling in Table 10, (4,56/5) of the consumers are uncomfortable with the dust because the bread is open to the top, (4,54/5) is seen.

Table 8. Reasons that Disturb Consumers in the Sale of Bread (Number, %)

Factors	5	4	3	2	1	Average	Rate (%)
Sale of Breads	3,66	0,67	0,12	0,06	0,03	4,54	90,89
Contact with dust due to its open top	3,76	0,67	0,06	0,04	0,03	4,56	91,29
Bread crates open top shipping way	3,56	0,71	0,06	0,10	0,04	4,48	89,51
Bakers not paying attention to hygiene	3,66	0,63	0,18	0,00	0,05	4,52	90,50

(5: It Bothers so much,4:Bothers,3:No Idea,2:Insignificant,1: It doesn't matter at all)

Table 9. Consumer Willingness to Pay for Bread Probit Analysis (Number, %)

	Estimate	Std. Error	Wald	Degree of Freedom	Level of Importance	95% Confidence Interval	
						Lowest Limit	Highest Limit
[WTP = 0,00]	-0,161	0,473	0,116	1	0,733	-1,088	0,766
[WTP = 1,00]	2,052	0,369	30,886	1	0,000	1,328	2,775
[WTP = 2,00]	2,718	0,374	52,789	1	0,000	1,985	3,451
[WTP = 3,00]	3,432	0,382	80,589	1	0,000	2,683	4,182
[WTP = 4,00]	3,775	0,387	95,099	1	0,000	3,016	4,533
[WTP = 5,00]	4,034	0,391	106,357	1	0,000	3,267	4,801
[WTP = 6,00]	4,547	0,401	128,335	1	0,000	3,76	5,334
[WTP = 7,00]	4,638	0,404	132,128	1	0,000	3,847	5,429
[WTP = 8,00]	4,717	0,405	135,347	1	0,000	3,922	5,512
[WTP = 9,00]	4,758	0,407	136,987	1	0,000	3,961	5,555
Household Size	-0,063	0,031	3,974	1	0,046	-0,124	-0,001
Age	0,175	0,040	18,916	1	0,000	0,096	0,253
Marital status	0,328	0,066	24,557	1	0,000	0,198	0,458
Income status	0,115	0,057	4,020	1	0,045	0,003	0,227
Profession	0,027	0,019	2,062	1	0,151**	-0,010	0,064
Education	0,285	0,038	56,648	1	0,000	0,211	0,360
Gender	0,177	0,114	2,419	1	0,120*	-0,046	0,400

According to the results of the probit analysis, all parameters in the model except occupation and gender were considered important because the significance levels of the in t statistics values were less than 10 % . According to the results of the probit analysis, an inverse relationship was determined between the size of the household and the willingness to pay for bread. As a matter of fact, as the number of households increases, bread consumption will increase, thus household expenses will increase. A positive correlation was found between age, marital status, income and education parameters among other variables and willingness to pay for bread.

### Conclusions and Recommendations

Bread consumption habits of households in Selcuklu District of Konya Province, which is the research area, were examined. According to the findings of the study, the average price of bread was found to be 1,15 TL, and it was determined that 68,32% of the consumers bought bread for 1,10 TL. It has been determined that consumers prefer maximum 200 grams of bread from existing breads sold in the market, and it is seen that they consume the most white bread compared to other bread groups. In the research, it was determined that the quality of the bread was sufficient at the rate of 58,42%, and it was determined that the place where consumers preferred to buy bread the most due to this criterion was the bakeries. The main thing that households pay attention to when buying bread is that it is cooked well and that it is sold under hygienic conditions. As a result of the information obtained within the scope of the study, it is revealed that bread is a staple food, regular inspection of bread production and sales places, encouragement of quality bread production, and awareness of healthy bread consumption. It is thought that a great contribution to the economy of the country will be made if the public is made aware of the preservation and utilization methods of bread and the awareness of preventing bread waste is brought to society and individuals.

### References

- Anonymous, (2012). Turkish Food Codex Bread and Bread Types Notification. 17/02/1999 Official Gazette No. 23614
- Anonymous (2020). 25 percent of protein needs will be met from bread, <https://www.hurriyet.com.tr/> ( Access Date 17.02.2020).

- Anonymous, (2019). Never Cut It Completely, [https:// www.hurriyet.com.tr/](https://www.hurriyet.com.tr/) (Access Date 18.02.2020).
- Carson, R.T. 2000. Contingent Valuation: A User's Guide, *Environmental Science and Technology*, 34(8) 1413-1418.
- Dağlıoğlu, O. (1998). Importance of Bread and Its Place in Our Nutrition. *Bakery World*, 27 (2):38-40.
- Doğan, İ.S. (2003). "Grain Processing Technology. Yüzüncü Yıl University Faculty of Agriculture, Department of Food Engineering Lecture Notes, Van. " Ataturk Univ. Faculty of Agriculture. *Derg.*, 42 (2): 165-180, 2011
- Elgün, A. And Ertugay, Z. (1995). Grain Processing Technology. Ataturk University. Faculty of Agriculture, Publication No: 297 (2nd Edition), Erzurum.
- Ertürk, A., Arslan, N., Sarıca, D., Demircan, V. (2010). "Bread Consumption and Waste of Families in the Urban Area of Isparta Province " Süleyman Demirel University, Faculty of Agriculture, Department of Agricultural Economics, Research Paper, Isparta. " *Academic Food* 13 (4) (2015) 291-298
- İnan, M. (2013). 'Bread Consumption Habits and Determination of Producer-Consumer Dynamics in the Center of Karaman Province'. Karamanoğlu Mehmetbey University, Institute of Science, Biology Department, Master's Thesis, Karaman.
- Parimala, K.R. and Sudha, M.L. (2015). Wheat-Based Traditional Flat Breads of India. *Critical Reviews in Food Science and Nutrition*, 55 (1): 67-81.
- Maalouf, N. M., Sakhaee, K., Parks, J. H., Coe, F. L., Adams-Huet, B., & Pak, C. Y. (2004). Association of urinary pH with body weight in nephrolithiasis. *Kidney international*, 65(4), 1422-1425.
- Ministry of Agriculture and Forestry, [https://www .tarimorman.gov.tr/](https://www.tarimorman.gov.tr/) ( Access Date 01.02.2020).
- SIS, (2004). Household Consumption Expenditure Results, 2003. State Institute of Statistics News Bulletin, <http://www.die.gov.tr> .
- TMO, (2008). Turkish Grain Board, <http://www.tmo.gov.tr> (Access Date 12.12.2019).
- Yıldırım, R. (2014). Aydın ilinde karasu sorunu ve zeytinyağı işletmelerinin çözüme yönelik tercihlerinin değerlendirilmesi (Master's thesis, Adnan Menderes Üniversitesi).
- Yurdatan, S. (2014). 'Bread Bread Industry and Consumer Trends in Turkey: What Edir Central County Example'. Namık Kemal University Institute of Science, Department Agricultural Economics, Master's Thesis, Tekirdağ

## Sulama Organizasyonlarının Yönetim Şekillerine Göre Etkinliklerinin Belirlenmesi

Süheyla AĞIZAN<sup>1</sup>, Zeki BAYRAMOĞLU<sup>1</sup>, Kemalettin AĞIZAN<sup>1</sup>, Merve BOZDEMİR<sup>1</sup>

<sup>1</sup>Selçuk Üniversitesi Ziraat Fakültesi Tarım Ekonomisi Bölümü, Konya, Türkiye

### Özet

Su; kıt ve stratejik öneme sahip doğal kaynaklardan biri olup canlıların yaşamı ve tarım sektörü için vazgeçilmezdir. Kullanım alanının (tarım, sanayi, evsel kullanım, enerji üretimi, balıkçılık, vb.) geniş olması ve nüfusun her yıl artması su kaynaklarına olan talebi artırmaktadır. Dünyada ve Türkiye’de su tüketiminin yaklaşık %70’i tarım sektörüne ait olup sektörde sürdürülebilir su kullanımının sağlanması su yönetimiyle mümkündür. Su yönetimi ise yerel düzeyde tarım sektörüne su sağlayan sulama kooperatifleri, sulama birlikleri, belediyeler ve köy tüzel kişiliklerinin etkinliklerinin belirlenmesi ve artırılmasıyla gerçekleştirilebilir. Bu nedenle çalışmada sulama organizasyonlarının yönetim şekillerine göre etkinliklerinin karşılaştırmalı olarak belirlenmesi amaçlanmıştır. Bu amacın gerçekleştirilmesinde “Malmquist Toplam Faktör Verimliliği” analizinden yararlanılmış olup sulama organizasyonlarının tahsil ettikleri su ücretleri ve diğer gelirlerinin toplamından oluşan “toplam gelir” çıktı olarak, organizasyonlarda çalışan daimî-geçici personel, sulama masrafları, personel giderleri, diğer giderler, sulanan arazi, sulama hattı, şebeke tipleri ile sulanan arazi miktarı ise girdi olarak modele dâhil edilmiştir. Çalışma birincil ve ikincil verilerden oluşmakla birlikte birincil veriler Konya Kapalı Havzası’nda faaliyet gösteren ve örneğe seçilen 67 adet sulama organizasyonundan anket uygulamasıyla elde edilmiştir. Elde edilen veriler sonucunda 2016-2018 yılları arasında sulama kooperatiflerinin teknik etkinliklerinin 1, diğer organizasyonların ise 1’in altında olduğu belirlenmiştir. Nitekim sulama kooperatiflerinin iki zaman periyodu arasındaki etkinlik değişimlerinde toplam faktör verimliliği değişimi 2017 yılında bir önceki yıla göre 0,52 ve 2018 yılında ise bir önceki yıla göre %29’luk bir artış ile 0,67 olarak belirlenmiştir. Teknik etkinlik açısından 2016-2018 yılları arasında belediyelerin 0,63-0,64 arasında, köy tüzel kişiliklerinin 0,82-0,88 arasında değer aldıkları belirlenmiştir. Bir diğer organizasyonda (sulama birliklerinde) ise teknik etkinlik 0,25 olarak tespit edilmiştir. Analiz sonucunda sulama birliklerinin, organizasyonlar arasında en düşük etkinliğe sahip olduğu belirlenmiştir. Sonuç olarak sulama organizasyonlarının kapasitelerini istenilen seviyede kullanamadıkları ve bu nedenle ivedilikle yapısal önlemlerin alınması gerektiği tespit edilmiştir.

**Anahtar kelimeler:** Etkinlik, Malmquist toplam faktör verimliliği, su yönetimi, sulama organizasyonları, sürdürülebilir su kullanımı

## Determining the Efficiency of According to the Management Types of Irrigation Organizations

### Abstract

Water; it is one of the scarce and strategically important natural resources and is indispensable for the life of living things and the agricultural sector. The wide usage area (agriculture, industry, domestic use, energy production, fishing, etc.) and the increase of the population every year increase the demand for water resources. Approximately 70% of water consumption in the world and Turkey belong to the agricultural sector, ensuring sustainable water use in agriculture, water management is possible. Water



management can be realized by determining and increasing the efficiency of irrigation cooperatives, irrigation unions, municipalities and village legal entities that provide water to the agricultural sector at the local level. For this reason, it is aimed to comparatively determine the effectiveness of irrigation organizations according to their management styles within the scope of this study. In the realization of this goal, "Malmquist Total Factor Efficiency" analysis was used and the "total income" consisting of the sum of water fees and other revenues collected by irrigation organizations was included in the model as output, while permanent-temporary personnel working in organizations, irrigation expenses, personnel expenses, other expenses, irrigated land, irrigation line, network types and the amount of irrigated land were included in the model as inputs. Although the study consists of primary and secondary data, the primary data were obtained through a survey application from 67 irrigation organizations operating in Konya Closed Basin. As a result of the data obtained, it was determined that the technical efficiency of irrigation cooperatives was 1, and other organizations were below 1 between 2016-2018. As a matter of fact, the total factor productivity change in the efficiency changes of irrigation cooperatives between the two time periods was determined as 0.52 in 2017 compared to the previous year and 0.67 in 2018 with an increase of 29% compared to the previous year. In terms of technical efficiency, it was determined that between the years 2016-2018, municipalities received values between 0.63-0.64 and village legal entities between 0.82-0.88. In another organization (irrigation unions), the technical efficiency was determined as 0.25. As a result of the analysis, it was determined that irrigation unions have the lowest efficiency among organizations. As a result, it was determined that irrigation organizations could not use their capacity at the desired level and therefore structural measures should be taken urgently.

**Key words:** Efficiency, Malmquist total factor efficiency, water management, irrigation organizations, sustainable water use

## Giriş

Canlıların yaşaması için hayati öneme sahip olan su; insanlar, bitkiler ve hayvanlar için temel besin maddelerindedir. İnsan vücudunda kan dolaşımının olabilmesi ve zararlı maddelerin dışarı atılabilmesi için suya ihtiyaç duyulurken, bitkilerde ise tohumun çimlenmesi, gübre-ilaç gibi girdilerin etkinliğinin artırılması, birim alanda verimliliğin sağlanması için suya ihtiyaç duyulmaktadır. Dolayısıyla su, diğer doğal kaynaklar gibi ekonomik değere sahip olmakla birlikte fizyolojik açıdan da önemlidir. Su kaynaklarının sınırlı olması ve kullanım alanının geniş olması nedeniyle üzerindeki baskı nüfus artışının da etkisiyle giderek artmakta, temiz-kullanılabilir suya erişim zorlaşmaktadır. Bu nedenle su sıkıntısının yaşanması tarım sektörü başta olmak üzere tüm kullanım alanlarını olumsuz yönde etkilemektedir. Nitekim tarım sektörü insanların besin ihtiyacını karşılaması açısından stratejik öneme sahiptir. Dolayısıyla suyun en fazla (%70) kullanım alanı olan tarım sektöründe, su olmadan nüfusun ihtiyaçlarını karşılamak imkansızdır (FAO AQUASTAT, 2019). Yaşamın devamlılığı için ihtiyaçları karşılamının yolu ise sürdürülebilir su kullanımı ile mümkündür.

Sürdürülebilir su kullanımı; tarım sektöründe suyun etkin kullanılması, tasarrufun dikkate alınması, optimal fiyatlandırma yaklaşımlarının benimsenmesi ve su yönetiminin sağlanması ile gerçekleştirilebilir. Su yönetimi; mevcut su kaynaklarının planlamasını, geliştirilmesini, dağıtımını ve optimum kullanımını ifade etmektedir (Evsahibioglu ve ark., 2010). Tarım sektöründe ise su yönetimi ulusal düzeyde Devlet Su İşleri (DSİ), Çölleşme ve Erozyonla Mücadele, Çevre Yönetimi, Doğa Koruma ve Milli Parklar, Su Yönetimi ve Meteoroloji Genel Müdürlükleri ile Türkiye Su Enstitüsü, Enerji Piyasası Düzenleme Kurumu tarafından gerçekleştirilmektedir. Yerel düzeyde ise özellikle suyun arazilere dağıtılmasında ve optimum kullanılmasında sulama birlikleri, sulama kooperatifleri, belediyeler ve köy tüzel kişilikleri (KTK) görev almaktadırlar. Bu nedenle su kullanımında büyük paya sahip olan tarım sektöründe faaliyet gösteren sulama organizasyonlarının teknik, sosyal, ekonomik, vb.

açından değerlendirilmeleri; su yönetiminin etkinliği ve sürdürülebilir su kullanımı açısından önemlidir. Bu doğrultuda çalışma kapsamında Konya Kapalı Havzası'nda faaliyet gösteren sulama organizasyonlarının yönetim şekillerine göre etkinliklerinin karşılaştırılması amaçlanmıştır. Amacın gerçekleştirilmesinde ise toplam faktör verimliliği (TFV) analizinden yararlanılmıştır.

### **Materyal ve Metot**

Çalışma birincil ve ikincil verilerden oluşmaktadır. Birincil veriler araştırma alanı olarak seçilen Konya Kapalı Havzası'nda DSİ tarafından inşa edilen ve sulama organizasyonlarına devredilen tesislerden anket uygulamasıyla elde edilmiştir. Konu ile ilgili literatür çalışmaları, kurum/kuruluş rapor ve istatistikler ise ikincil verileri oluşturmaktadır.

Anket uygulaması havza genelinin oluşturan Konya, Aksaray, Karaman ve Niğde illerinde gerçekleştirilmiş olup havzada DSİ tarafından inşa edilmiş ve farklı organizasyonlara devredilmiş tesisler ana çerçeveyi oluşturmuştur. Bu doğrultuda 18 adet sulama birliğine, 14 adet belediyeye, 11 adet KTK'ya ve 299 adet sulama kooperatifine tesis devri gerçekleştirilmiştir. Sulama birlikleri, belediye ve KTK'lar için tam sayım yöntemi kullanılmıştır. Sulama kooperatifleri için toplam sulama kooperatifi varlığının %10'u örneğe seçilmiştir. Ancak çeşitli nedenlerden dolayı 4 adet KTK ve 2 adet belediye ile görüşme sağlanamamış olup toplam 67 adet organizasyon ile görüşülmüştür.

Çalışma kapsamında karar verme birimleri olarak sulama organizasyonları diğer bir ifadeyle sulama kooperatifleri, belediyeler, KTK'lar ve sulama birlikleri ele alınmıştır. Karar verme birimleri olan organizasyonların TFV'lerinin hesaplanabilmesi için girdi ve çıktı değişkenlerinin belirlenmesi gerekmektedir. TFV hesaplarında genellikle toplam gelir, brüt üretim değeri vb. gelirler çıktı olarak ele alınmaktadır. Bu nedenle çalışma kapsamında sulama organizasyonlarının tahsil ettikleri su ücretleri ve diğer gelirlerinin toplamından oluşan "toplam gelir" çıktı olarak modele dâhil edilmiştir. Girdi içerisinde ise organizasyonlarda çalışan daimî ve geçici personel, sulama masrafları, personel giderleri, diğer giderler, sulanan arazi, sulama hattı ve şebeke tipleri ile sulanan arazi miktarı ele alınmıştır.

Çalışma kapsamında sulama organizasyonlarının etkinlikleri "Malmquist-TFV" ile belirlenmiştir. TFV analizi; üretimde kullanılan girdilerin ne kadar etkin ve yoğun bir şekilde kullanıldığının seviyesi ile ilgilidir (Comin, 2006). Üretimde kullanılan kaynaklardan birinin önemli olmasının yanı sıra diğerlerinin de aynı derecede önem taşımalarını ve bu kaynakların birbirlerini etkileme derecelerini ölçmektedir. Kaynaklardan herhangi birinde meydana gelecek değişim, kaynağın verimliliğinde de değişime neden olacaktır. Böylece tüm faktörlere göre kaynağın verimliliğini ölçerek genel verimlilik düzeyi hakkında bir fikir elde edilebilecektir. TFV; toplam çıktının, çıktının üretilmesinde kullanılan girdilere oranlanması ile belirlenmektedir. Dolayısıyla üretim sürecinde kullanılan tüm girdilerin etkinliği önem derecelerine göre sıralanabilmektedir. Artışların hesaplanması verimliliğin ne yönde ve hangi hızda değiştiği konusunda açık bir şekilde fikir vermektedir.

İsveçli istatistikçi Malmquist tarafından geliştirilen Malmquist-TFV endeksi (González-Rodríguez ve ark., 2015), TFV'deki değişimleri ölçmek ve bu değişimlerin niteliğini analiz etmek için sıklıkla kullanılan yöntemlerden birisidir (Özcan, 2011). Veri zarflama tabanlı Malmquist verimlilik endeksi zaman içerisindeki verimlilik değişimini ölçmektedir (Chen ve Ali, 2004). Bir başka ifade ile Malmquist Endeksi; belirli bir yıl için bir önceki yıla göre bir etkinlik ölçüsü oluşturmaktadır. Bu etkinlik ölçüsüne göre verimliliğin zaman içerisinde ne kadar değişim gösterdiği (arttığı veya azaldığı) ortaya çıkmaktadır.

Genel itibarıyla uzaklık fonksiyonuna göre belirlenen Malmquist-TFV aşağıdaki gibi ifade edilmektedir:

$$m_0(y^t, x^t, y^{t+1}, x^{t+1}) = \left( \frac{d_0^t(y^t, x^{t+1})}{d_0^t(y^t, x^t)} * \frac{d_0^{t+1}(y^{t+1}, x^{t+1})}{d_0^t(y^t, x^t)} \right)^{1/2}$$

Yukarıda endekslerin geometrik ortalaması verilmiş olup birincisi t dönemi teknolojisini ifade ederken, ikincisi t+1 dönemi teknolojisini göstermektedir.

Malmquist-TFV endeksine göre verimlilikteki değişiklikler teknik etkinlik değişimi (TED) ve teknolojik değişim (TD) olarak iki kısımda ele alınmaktadır (Chen ve Ali, 2004; Barros, 2005).

$$\text{Etkinlikteki Değişim (TED)} = \left( \frac{d_0^{t+1}(y^{t+1}, x^{t+1})}{d_0^t(y^t, x^t)} \right)$$

$$\text{Teknolojik Değişim (TD)} = \left( \frac{d_0^{t+1}(y^{t+1}, x^{t+1})}{d_0^t(y^t, x^t)} * \frac{d_0^t(y^t, x^t)}{d_0^{t+1}(y^t, x^t)} \right)^{1/2}$$

TD şimdiki (t) ve sonraki (t+1) periyotlar arasındaki teknolojideki değişimi hesaplar; TED, etkinlik değişimini ölçmektedir. Ayrıca TFV'de bir iyileşme hem teknik etkinlik hem de teknolojik değişimde bir iyileşme anlamı taşımamaktadır (González-Rodriguez ve ark., 2015). TED, “saf teknik etkinlik (STE)” ve “ölçek etkinliği (ÖE)” olmak üzere iki kısımda değerlendirilmektedir. STE, ölçeğe göre değişken getiri varsayımı altında girdilerin etkin kullanımı anlamına gelirken, ÖE; optimal ölçek büyüklüğünde veya ölçeğe göre sabit getiride faaliyet gösterilmesini ifade etmektedir (Anderson ve ark., 2002). Bu doğrultuda TFV = TED x TD veya TFV = (STED x ÖED) x TD şeklinde de hesaplanabilmektedir (Yen ve Othman, 2011).

Bu çalışmada da sulama organizasyonlarının etkinliklerinde meydana gelen değişimleri ölçmek amacıyla Malmquist-TFV Endeksi uygulanmış ve analize ilişkin hesaplamalar DEAP 2.1 programıyla gerçekleştirilmiştir.

## Bulgular ve Tartışma

Bugüne kadar birçok araştırmada TFV konusunda çalışmalar yapılmıştır. Özellikle bankacılık, sigortacılık (Akhisar ve Tezergil, 2014), tekstil (Gelmez ve ark., 2018), turizm (Benli, 2012), çimento (Kula ve ark., 2009), otomotiv (Lorcu, 2010) ve imalat (Özcan, 2011) gibi sektörler üzerinde Malmquist-TFV analizi kullanılarak sektörlerin veya işletmelerin verimlilikleri araştırılmıştır. Tarım sektöründe ise Malmquist-TFV analizi kullanılarak hem sektörün hem de işletmelerin verimliliklerini inceleyen birçok çalışma gerçekleştirilmiştir. Özellikle TİGEM (Candemir ve Deliktaş, 2006), süt üretim kooperatifleri (Bektaş, 2016), fındık fabrikaları (Doğan ve Bulut, 2014), bitkisel üretim faaliyetleri (Özden, 2005) ve su ürünleri (Seki ve Akbulut, 2015) konusunda Malmquist-TFV ölçülmüştür. Mikro düzeyde yapılan bu çalışmaları yanı sıra makro düzeyde Türkiye tarımını Malmquist-TFV endeksine göre inceleyen çalışmalar da bulunmaktadır (Cankurt ve ark., 2010; Mollavelioğlu ve Ceylan, 2010; Kaya ve Aktan, 2011; Bozoğlu ve Eroğlu, 2012; Karaman ve Özalp, 2017; Mollavelioğlu ve ark., 2017). Sulama işletmeciliği üzerine ise sadece Aydın ve ark. (2017) tarafından Kırklareli, Edirne, Tekirdağ ve Çanakkale illerinde yapılan çalışma literatürde yer almaktadır.

Yapılan Malmquist-TFV analizinde incelenen sulama organizasyonlarına ilişkin 2016-2018 yıllarına ait veriler kullanılmış, her yıla ait teknik etkinlik değerleri hesaplanmıştır (Çizelge 1). Teknik etkinlik hesaplanırken 1-3 yıl içerisinde değerler hesaplanmıştır. Çalışmada yer alan üç çizelgede de “t” dönemleri incelenecek olup, sulama kooperatiflerinin teknik etkinlerinin üç yıl içinde de “1” olduğu belirlenmiştir. Buna göre diğer organizasyonlar sulama kooperatiflerine göre değerlendirilmiştir. Yıllara göre belirlenen teknik etkinlikler sulama organizasyonlarına göre değişmekle birlikte genel ortalama 2016 ile 2017 yıllarında 0,71 ve 2018 yılında ise 0,72 olarak belirlenmiştir (Çizelge 1). Buna göre sulama

organizasyonlarının ortalama 0,71-0,72 etkinlik düzeyinde çalıştıkları söylenebilir. Bu durum sulama organizasyonlarının teknolojiyi etkin kullanımında gerilediklerini göstermektedir. Ayrıca yıllar itibarıyla değişmemekle birlikte en etkin çalışan organizasyonlar sırasıyla; sulama kooperatifleri, KTK'lar, belediyeler ve sulama birlikleridir. Türkiye'nin Batı Marmara bölgesinde yapılan bir çalışmada da benzer sonuçlar bulunmuş ve sulama kooperatiflerinin sulama birliklerine göre daha etkin çalıştıkları belirlenmiştir (Aydın ve ark., 2017). Coşkun (2006) tarafından 27 tane sulama birliği üzerine yapılan bir etkinlik çalışmasında sulama birliklerinden sadece 6 tanesinin tam etkin çalıştığı belirlenmiştir. Frija ve ark. (2009) ise yapmış oldukları çalışmada sulama organizasyonlarının önemli bölümünün ölçek ekonomisine uygun çalışmadığını tespit etmiştir. Özdemir ve Armağan (2009) tarafından yapılan çalışmada Aydın yöresinde faaliyet gösteren 8 adet sulama birliğinden sadece 2 tanesinin tam etkin çalıştığı, Sayın (2011)'in Antalya ilinde yaptığı bir çalışmada sulama birliklerinin teknik etkinliklerinin ortalama 0,77 olduğunu belirtmişlerdir. Bektaş (2016) ise çalışmasında Çanakkale'de faaliyet gösteren kooperatiflerin teknolojik donanımlarının yetersiz olmasından dolayı etkin çalışmadıklarını tespit etmiştir.

Çizelge 1. Sulama Organizasyonlarının Yıllara Göre Etkinlikleri

Yıllar	Organizasyon Türü	Dönemler			vrs
		t-1	t	t+1	te
2016	Sulama Birliği	0,00	0,25	0,50	1,00
	Belediye	0,00	0,63	1,25	1,00
	KTK	0,00	0,82	1,57	1,00
	Sulama Kooperatifi	0,00	1,00	2,00	1,00
	<b>Genel Ortalama</b>	<b>0,00</b>	<b>0,71</b>	<b>1,42</b>	<b>1,00</b>
2017	Sulama Birliği	0,13	0,25	0,38	1,00
	Belediye	0,41	0,63	0,98	1,00
	KTK	0,59	0,82	1,30	1,00
	Sulama Kooperatifi	0,54	1,00	1,51	1,00
	<b>Genel Ortalama</b>	<b>0,41</b>	<b>0,71</b>	<b>1,09</b>	<b>1,00</b>
2018	Sulama Birliği	0,17	0,25	0,00	1,00
	Belediye	0,46	0,64	0,00	1,00
	KTK	0,63	0,88	0,00	1,00
	Sulama Kooperatifi	0,68	1,00	0,00	1,00
	<b>Genel Ortalama</b>	<b>0,50</b>	<b>0,72</b>	<b>0,00</b>	<b>1,00</b>

Çizelge 2'de TED, TD, saf teknik etkinlik değişimi (STED), ölçek etkinliği değişimi (ÖED) ve toplam faktör verimliliği değişimi (TFVD) değerleri hesaplanarak organizasyon türüne göre ortalama Malmquist endeksleri ve yıllara göre değişimler verilmiştir. Değişimler ölçege sabit getiri varsayımı altında hesaplanmıştır.

Çizelge 2. Sulama Organizasyonlarının Yıllara Göre Toplam Faktör Verimliliği Değişimi

Yıllar	Organizasyon Türü	TED	TD	STED	ÖED	TFVD
2017	Sulama Birliği	1,00	0,51	1,00	1,00	0,51
	Belediye	1,00	0,55	1,00	1,00	0,55
	KTK	1,00	0,60	1,00	1,00	0,60
	Sulama Kooperatifi	1,00	0,52	1,00	1,00	0,52
	<b>Genel Ortalama</b>	<b>1,00</b>	<b>0,53</b>	<b>1,00</b>	<b>1,00</b>	<b>0,53</b>
2018	Sulama Birliği	1,01	0,67	1,00	1,01	0,67
	Belediye	1,02	0,68	1,00	1,02	0,69
	KTK	1,08	0,67	1,00	1,08	0,72
	Sulama Kooperatifi	1,00	0,67	1,00	1,00	0,67
	<b>Genel Ortalama</b>	<b>1,01</b>	<b>0,67</b>	<b>1,00</b>	<b>1,01</b>	<b>0,68</b>

Sulama organizasyonlarının iki zaman periyodu arasındaki etkinlik değişimi sonucunda Malmquist endeks değerlerinin yorumlanmasında; elde edilen m değeri >1'den ise bir önceki döneme

göre Malmquist-TFV’de bir artış olduğu, eğer m değeri =1 ise bir önceki döneme göre Malmquist-TFV’de herhangi bir değişim olmadığı ve m değeri <1 ise bir önceki döneme göre Malmquist-TFV’de bir azalış olduğu söylenebilir (Akhisar ve Tezergil, 2014). Ayrıca TED, etkinlikteki değişimi göstermekte olup, bu değer 1’den büyük olması organizasyonun kapasite sınırını yakaladığını göstermektedir. TD, kullanılan teknolojideki değişimi ifade etmekte olup STED ölçeğe değişen getiri koşulunda teknik etkinlikteki değişimi göstermektedir. ÖED ise maksimum kapasite kullanım derecesini ifade etmektedir. STED ve ÖED değerlerinin 1’den büyük olması organizasyonların kurumsal olarak etkin yönetildiklerini ve uygun ölçekte hizmet sağladıklarını göstermektedir. Son olarak TFVD ise teknik ve teknolojik etkinlikteki değişimlerinin toplamıdır (Lorcu, 2010; Akhisar ve Tezergil, 2014).

Çıktıların değerlendirilmesi yapılırken ikinci ve üçüncü yılın değişimlerinin hesaplandığına dikkat edilmelidir. Nitekim değişimin hesaplandığı için ilk yıl modele dâhil edilmemektedir. Bu nedenle 2017-2018 yıllarına ait değişimler gösterilmektedir. Çizelge 2’de sulama organizasyonlarının 2017 ve 2018 yılı TFVD gösterilmiştir. TFVD 1’den büyük ise TFV’nde bir artış olduğu, 1’den küçük ise TFV’nde bir azalma olduğu söylenebilir. Bu durum sulama organizasyonlarına göre değişmemekle birlikte 2017 yılında ortalama TFVD bir önceki yıla göre 0,53 olarak hesaplanmıştır. Bu değere göre organizasyonlar 2017 yılında 2016 yılına göre yılda %47’lik bir verimlilik düşüşü yaşamışken, 2018 yılında ise kayıp bir önceki yıla (2017) göre %32 olarak belirlenmiştir. Sulama organizasyonlarına göre TFVD incelendiğinde 2017 yılında sulama birliklerinin 0,51, belediyelerin 0,55, KTK’ların 0,60 ve sulama kooperatiflerinin 0,52 değerlerine sahip oldukları belirlenmiştir. 2018 yılında ise TFVD sulama birliklerinde %31 artış yaşanmasına rağmen bir önceki yıla göre TFV %33 azalmıştır. Bu durum belediyelerde %31, KTK’larda %28 ve sulama kooperatiflerinde ise %33 olarak hesaplanmıştır. Görüldüğü üzere TFVD’nin en yüksek olduğu organizasyon KTK (0,72) olurken, artış hızı bakımında sulama birlikleri (%31) ve sulama kooperatifleri (%28) ön plandadır. Nitekim 2018 yılında sulama birliklerinin DSİ’ye devredilmesi ile yönetim etkinliğinin sağlanması sulama birliklerinde etkinliği artıran önemli bir unsur olmuştur. Ayrıca pompajlı ve kapalı sistemlere geçişlerin en yüksek olduğu organizasyonlar olan sulama kooperatiflerinde ise %29’luk verim artışı önemli görülmektedir. Genel itibariyle her iki dönemde de STED ve ÖED değerlerinde herhangi bir değişiklik olmaması nedeniyle verimlilik artışının temelinde teknolojinin olduğu görülmektedir.

Günümüze kadar yapılan çalışmalarda da TFVD’nin en önemli sebebinin TD olduğunu gösteren çalışmalar literatürde yer almaktadır (Aydın ve ark., 2017). Su ürünleri endüstri firmaları üzerinde yapılan bir çalışmada da benzer sonuçlar görülmüş olup, TFVD %20,5 oranında bir artış göstermiş ve bunun en önemli sebebinin teknolojik ilerlemenin olduğu ifade edilmiştir (Seki ve Akbulut, 2015). Fındık satış kooperatiflerine üzerine yapılan bir çalışmada ise TFVD’nin en önemli sebebinin teknolojik değişim olduğu belirlenmiştir (Candemir ve ark., 2011). Bununla birlikte makro düzeyde tarımsal TFV’deki artışın büyük ölçüde teknik gelişmelerden kaynaklandığını gösteren çalışmalar olduğu gibi (Mollavelioğlu ve ark., 2017), TFVD’nin nedeni olarak teknolojik değişimi neden olarak gösteren çalışmalarda yer almaktadır (Kaya ve Aktan 2011).

Çizelge 2’de aynı zamanda teknolojik değişim (TD) gösterilmiş olup 2017 yılında TD incelendiğinde en düşük değere sulama birliklerinin sahip olduğu görülmektedir. Birliklerin bu dönemde yönetim etkinsizliği nedeniyle altyapının yeterince kullanılamaması, su iletim tipi olarak cazibeli sistemlerin kullanılıyor olması ve iletim hattı tipi olarak da açık kanallarının yaygın olması bu sonucu beraberinde getirmiştir. Fakat 2018 yılında kapalı sistemlere geçişin hızlanması, tahsilat oranının yükselmesi ve bakım ihtiyaçlarının karşılanabilmesi TD’yi %31 oranında artırarak diğer organizasyonlarla aynı seviyeye gelmesini sağlamıştır. TD açısından 2017 yılında en düşük değere sahip bir diğer organizasyon ise sulama kooperatifleridir. Kooperatiflerin 2017 yılında TD açısından düşük değere sahip olmasının temel nedeni ön ödemeli (kartlı) sistemin yaygınlaşmamış olmasıdır. Bu sistemin 2018 yılında organizasyonların çoğunda faaliyete geçirilmesiyle TD değeri 0,67’ye yükselmiştir.

Dolayısıyla ön ödemeli sistemde kooperatiflerin tahsilat oranları artmıştır. Fakat her iki organizasyonda da teknolojik değişimde artış yaşanmasına rağmen her ikisinde de teknolojik değişimin verimlilik kaybına etkisi %33 olarak hesaplanmıştır.

Çizelge 3. Yıllık Ortalamalar Göre TFV ve Bileşenlerindeki Değişim

Yıllar	TED	TD	STED	ÖED	TFVD
2	1,00	0,53	1,00	1,00	0,53
3	1,01	0,67	1,00	1,01	0,68
<b>Ortalama</b>	<b>1,01</b>	<b>0,60</b>	<b>1,00</b>	<b>1,01</b>	<b>0,60</b>

Çizelge 3'te yıllara göre ortalama TFV ve bileşenlerindeki değişimler verilmiştir. Çizelgede sulama organizasyonlarının son 3 yılda ortalama %40 verimlilik kaybı yaşadığı ve bu verim kaybının en önemli sebebinin teknolojik değişim olduğu görülmektedir. Bu durum organizasyonlar itibariyle benzerlik göstermektedir. Çizelge 4'te ise yıllara göre ortalama TFV ve bileşenlerindeki değişimler organizasyonlar itibariyle gösterilmiştir. Elde edilen 3 yıllık veriler doğrultusunda sulama birliklerinde %41, belediyelerde %38, KTK'larda %34 ve sulama kooperatiflerinde ise %40 oranında bir verimlilik azalışı söz konusudur. Daha önceki verilerde sulama kooperatiflerinin dönem içerisinde tam etkin çıkmasına karşı, yıllar içerisinde değişim yüzdesi belediye ve KTK'ların gerisinde kalmıştır. Nitekim 2018 yılında yerel seçimlerden dolayı belediye ve KTK'ların hizmet götürme amaçları ve siyasi faktörler nedeniyle dağıtılan su miktarını artırdıkları ve ücretsiz su dağıtımını yaptıkları tespit edilmiştir. Bu nedenle belediye ve KTK'ların etkinlik değişim yüzdelerindeki artış oranı diğerlerine göre daha yüksektir. Çizelgelere göre göreceli olarak son 2 yılda etkinliğini en çok artıran sulama organizasyonunun KTK'lar olduğu, bu organizasyonu sırasıyla belediye, sulama kooperatifleri ve sulama birliklerinin takip ettiği görülmektedir. 2018 yılında bir önceki yıla göre tüm organizasyonlarda TFV'de bir artış sağlanmıştır. Verimlilik artışı sulama işletmelerinin rekabet güçlerini ve sürdürülebilirliklerini artırmaktadır.

Çizelge 4. Sulama Organizasyonlarının Yıllık Ortalamalara Göre TFV ve Bileşenlerindeki Değişim

Organizasyon Türü	TED	TD	STED	ÖED	TFVD
Sulama Birliği	1,01	0,59	1,00	1,01	0,59
Belediye	1,01	0,62	1,00	1,01	0,62
KTK	1,04	0,64	1,00	1,04	0,66
Sulama Kooperatifi	1,00	0,60	1,00	1,00	0,60
<b>Genel Ortalama</b>	<b>1,01</b>	<b>0,60</b>	<b>1,00</b>	<b>1,01</b>	<b>0,60</b>

Bugüne kadar konu ile ilgili birçok çalışmada sulama kooperatiflerinin, birliklerinin, belediyelerin etkinlikleri ayrı ayrı hesaplanmış olup organizasyonları bütüncül yaklaşımla inceleyen çalışmanın literatürde eksik olduğu görülmüştür. Bu doğrultuda yapılan çalışma ile organizasyonların TFV'leri hesaplanmış ve verimlilik azalışlarının teknoloji kullanımından kaynaklandığı tespit edilmiştir. Yıllar itibariyle sulama organizasyonlarının teknoloji değişimine bağlı olarak TFVD'lerinin artmış olmasına rağmen bu değer hâlen 1'in altında olması organizasyonların kapasitelerini istenilen seviyede kullanamadıklarının göstergesidir. Ayrıca organizasyonların teknik etkinlik ortalamalarının 1'in altında olması da yapısal önlemlerin alınması gerektiğinde bir ispatıdır. Dolayısıyla organizasyonların etkinliklerini ve kapasitelerini artırmak için alınması gereken önlemlerin başında maliyetlerin azaltılması yer almaktadır. Ayrıca verimli teknolojinin kullanılarak (damla, toprak altı sulama sistemler, sayaç sistemleri, kartlı sistemler vb.) toplam etkinliğin artırılması söz konusu olabilir. Sulama organizasyonlarının rekabetlerini artırıcı ve ihtiyaç duydukları teknolojik gelişmelere yönelik kamu desteklemelerinin sağlanması da TFV'yi artırıcı bir unsurdur. Bununla birlikte başta sulama birlikleri olmak üzere tüm organizasyonlarda personel verimliliğinin artırılmasıyla da etkinlik artışı

sağlanabilir. Nitekim organizasyon yöneticilerinin izleme, değerlendirme, bakım ve onarım faaliyetlerinden sorumlu olması ve bu yöneticilerin kişisel, finansal ve teknik anlamda beceri sahibi olması organizasyonların sürdürülebilirliği açısından önemlidir. Bu nedenle organizasyon yöneticilerinin belirli bir eğitim seviyesine sahip olması veya konu ile ilgili eğitimlere katılmaları gerekmektedir. Eğitimlerde yöneticilerin kurumsallaşmaya, yönetim becerilerine ve teknik bilgilerine yönelik eksiklerinin giderilmesi önceliklendirilmelidir. Ayrıca yapısal sorunlara yönelik (arazi eğimi, sulama sistemleri, şebeke tipleri, tahsilat vb.) kamu denetimlerinin ve desteklemelerinin artırılarak organizasyonların desteklenmesi de önem arz eden bir diğer etkidir. Sulama organizasyonlarında su kapasitelerine uygun üretim deseninin belirlenmesi, üretim desenine göre sulama zamanlarının belirlenmesi ve bir plan dâhilinde su dağıtım programlarının oluşturularak dağıtılması teknik açıdan organizasyonların etkinliklerinin artırılmasını sağlayacaktır.

### Teşekkür

Çalışma Selçuk Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü tarafından desteklenen 19401016 no'lu projeden elde edilmiş olup desteklerinden dolayı teşekkür ederiz.

### Kaynaklar

- Akhisar İ, Tezergil S. 2014. Malmquist Toplam Faktör Verimlilik Endeksi: Türk Sigorta Sektörü Uygulaması. Finansal Araştırmalar ve Çalışmalar Dergisi, 5 (10), 1-14.
- Anderson RI, Fok R, Springer T, Webb J. 2002. Technical Efficiency and Economies of Scale: A Non-Parametric Analysis of REIT Operating Efficiency. European Journal of Operational Research, 139 (3), 598-612.
- Aydın B, Özkan E, Hurma H, Aktaş E, Azabağaoğlu Ö, Özdemir G. 2017. Sulama İşletmeciliğinde Etkinlik Analizi (Kırklareli, Edirne, Tekirdağ ve Çanakkale İlleri Örneği). Türk Tarım ve Doğa Bilimleri Dergisi, 4 (1), 70-78.
- Barros CP. 2005. Evaluating The Efficiency of A Small Hotel Chain with A Malmquist Productivity Index. International Journal of Tourism Research, 7 (3), 173-184.
- Bektaş S. 2016. Biga İlçesinde Süt Üretim Kooperatiflerinin Etkinlik ve Verimlilik Analizi. Çanakkale Onsekiz Mart Üniversitesi Sosyal Bilimler Enstitüsü İktisat Anabilim Dalı, 139.
- Benli YK. 2012. Veri Zarflama Analizi (VZA) ve Malmquist Toplam Faktör Verimliliği (TFV): Konaklama İşletmelerinde Bir Uygulama. Ege Academic Review, 12 (3), 369-382.
- Bozoğlu M, Eroğlu NA. 2012. Farklı Politika Dönemlerinde Türkiye'nin Tarımsal Verimlilik, Etkinlik ve Teknolojisinin Analizi. Türkiye X. Tarım Ekonomisi Kongresi. Konya: 113-120.
- Candemir M, Deliktaş E. 2006. TİGEM İşletmelerinde Teknik Etkinlik, Ölçek Etkinliği, Teknik İlerleme, Etkinlikteki Değişme ve Verimlilik Analizi: 1999-2003. Tarım ve Köy İşleri Bakanlığı, Tarımsal Ekonomi Araştırma Enstitüsü, Yayın no:141.
- Candemir M, Özcan M, Güneş M, Deliktaş E. 2011. Technical Efficiency and Total Factor Productivity Growth in The Hazelnut Agricultural Sales Cooperatives Unions in Turkey. Mathematical and Computational Applications, 16 (1), 66-76.
- Cankurt M, Miran B, Günden C, Şahin A. 2010. AB ve Türkiye'nin Tarımsal Üretim Etkinlik ve Verimliliği Üzerine Küresel Krizlerin Etkileri. Türkiye IX. Tarım Ekonomisi Kongresi, Şanlıurfa: 221-228.
- Chen Y, Ali AI. 2004. DEA Malmquist Productivity Measure: New Insights with An Application To Computer Industry. European Journal of Operational Research, 159 (1), 239-249.
- Comin D. 2006. Total Factor Productivity. New York University and NBER, Pridobljeno iz. <http://www.people.hbs.edu/dcomin/def.pdf>.
- Coşkun Z. 2006. Aşağı Ceyhan ve Aşağı Seyhan Ovaları Sulama Projeleri Alanındaki Sulama Birliklerinin Etkinlik Analizi. Türkiye VII. Tarım Ekonomisi Kongresi. Antalya: 442-451.
- Doğan H, Bulut, Y. 2014. Giresun'daki Fındık Fabrikalarının Etkinlik Analizi: Malmquist-TFV Endeksi Uygulaması. Adnan Menderes Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 1 (3), 76-88.
- FAO AQUASTAT. 2019. Food and Agriculture Organization. <http://www.fao.org/land-water/en/>, Access Date: 27.06.2020

- Frija A, Speelman S, Chebil A, Buysse J, Van Huylenbroeck G. 2009. Assessing The Efficiency of Irrigation Water Users' Associations and Its Determinants: Evidence From Tunisia. *Irrigation and Drainage: The Journal of The International Commission on Irrigation and Drainage*, 58 (5), 538-550.
- Gelmez E, Göral R, Çağlıyan V. 2018. İşletmelerin Malmquist-TFV Endeksi ile Verimliliklerinin Analizi: Tekstil Sanayi Örneği. *Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi*, 21 (1), 18-29.
- González-Rodríguez MdR, Martín-Samper R, Giuliani AC. 2015. Evaluating The Efficiency Progress with Technology in A Spanish Hotel Chain. *Revista de Administração de Empresas*, 55 (5), 551-562.
- Karaman S, Özalp A. 2017. Türkiye Tarım Sektörü Bölgesel Toplam Faktör Verimliliğinin Malmquist Endeks ile Belirlenmesi. *Gaziosmanpaşa Üniversitesi Ziraat Fakültesi Dergisi*, 34 (1), 209-217.
- Kaya P, Aktan HE. 2011. Türk Tarım Sektörü Verimliliğinin Parametrik Olmayan Bir Yöntemle Analizi. *Journal of Alanya Faculty of Business/Alanya İşletme Fakültesi Dergisi*, 3 (1), 261-282.
- Kula V, Kandemir T, Özdemir L. 2009. VZA Malmquist Toplam Faktör Verimlilik Ölçüsü: İMKB'ye Koteli Çimento Şirketleri Üzerine Bir Araştırma. *Selçuk Üniversitesi İktisadi ve İdari Bilimler Fakültesi Sosyal ve Ekonomik Araştırmalar Dergisi*, 9 (17), 187-202.
- Lorcu F, 2010. Malmquist Toplam Faktör Verimlilik Endeksi: Türk Otomotiv Sanayi Uygulaması. *Istanbul University Journal of the School of Business Administration*, 39 (2).
- Mollavelioğlu MŞ, Ceylan R. 2010. Türkiye ve AB Ülkelerinde Tarımsal Toplam Faktör Verimliliği ve Yakınsama Analizi. *Akdeniz Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 10 (20), 86-103.
- Mollavelioğlu MŞ, Kanberoğlu Z, Oğuz A. 2017. AFRO-Avrasya Ülkelerinde Tarımsal Toplam Faktör Verimliliği Büyümesi: Malmquist Endeks Analizi. *Yüzüncü Yıl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 1 (1).
- Özcan A. 2011. Türkiye İmalat Sanayinde Toplam Faktör Verimliliği ve Beşeri Sermaye İlişkisi. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 20 (2), 223-242.
- Özdemir K, Armağan G. 2009. Aydın İlindeki Sulama Birliklerinin Faaliyetlerinin Değerlendirilmesi ve Etkinliklerinin Belirlenmesi. *Adnan Menderes Üniversitesi Ziraat Fakültesi Dergisi*, 7 (2), 75-83.
- Özden A. 2005. Aydın İli Tarım İşletmelerinde Bitkisel Üretim Faaliyetlerinin Verimliliklerinin Belirlenmesi. *Adnan Menderes Üniversitesi Tarım Ekonomisi ABD*, 139.
- Sayın B. 2011. Antalya'da Sulama İşletmeciliği Faaliyetleri, Üreticilerin Sulama Suyu Talebi ve Sulama İşletmeciliği Faaliyetlerine Katılım Düzeyinin Değerlendirilmesi. *Akdeniz Üniversitesi Fen Bilimleri Enstitüsü, Doktora Tezi. Antalya*, 166s.
- Seki İ, Akbulut M. 2015. Su Ürünleri Endüstrisindeki Firmaların Etkinlik ve Verimliliklerini Belirleyen Faktörlerin TR22 Bölgesi İçin Analizi. *Tarım Ekonomisi Dergisi*, 21 (1 ve 2), 1-6.
- Yen FL, Othman M. 2011. Data Envelopment Analysis to Measure Efficiency of Hotels in Malaysia. *SEGi Review*, 4 (1), 25-36.



## Solar Radyasyondaki Değişimin Farklı İstatistik Yaklaşımlarla Analizi

Müberra ERDOĞAN<sup>1</sup>, Mehmet Murat CÖMERT<sup>1</sup>, Kadri YÜREKLI<sup>1</sup>

<sup>1</sup>Biyosistem Mühendisliği Bölümü, Gaziosmanpaşa Üniversitesi Ziraat Fakültesi, Taşlıçiftlik, 60100, Tokat, Türkiye

### Özet

Trend analizi hidro-meteorolojik verilerdeki değişimin saptanması, gelecekte yapılması gerekli kararları alma açısından oldukça önemlidir. Bu çalışmada Tokat, Amasya, Çorum ve Samsun illerine ait solar radyasyon verilerindeki değişim trend analizi ile araştırılmıştır. Çalışmada aylık, mevsimlik ve yıllık solar radyasyon değişimleri incelenmiştir. Çalışmada trend analizi için Mann-Kendal (MK), Spearman Rho (SR) ve yenilikçi trend analiz (ITA) yöntemleri kullanılmıştır. Tokat ili için MK testine göre Haziran ayında, SR testine göre Mart ayında azalan trend saptanmıştır. Amasya ili için tüm yaklaşımlar Kasım ayı ve Mevsim-1 (M-1) sezonunda artan trend ortaya çıkarmıştır. Çorum ili için tüm trend testlerine göre ocak ayında azalan trend saptanmıştır. Samsun ili için MK testine göre Nisan ayı, Mevsim-3 (M-3) ve Mevsim-4 (M-4) sezonları, SR testine göre Nisan ayı ve Mevsim-4 (M-4) sezonunda artan trend saptanmıştır. Tüm istasyonlar için yıllık solar radyasyon verilerinde trend görülmemiştir. Yenilikçi trend analizi ile diğer trend analizi sonuçlarını destekleyen grafikler elde edilmiştir.

**Anahtar kelimeler:** Solar radyasyon, Mann-Kendal, Spearman's Rho, yenilikçi trend analizi

## Analysis of the Change in Solar Radiation Using Different Statistical Approaches

### Abstract

Trend analysis is very important in determining the changes in hydro-meteorological data in order to make the necessary decisions in the future. In this study, the variability in solar radiation data of Tokat, Amasya, Çorum and Samsun provinces was investigated by trend analysis. Mann-Kendall, Spearman Rho and innovative trend analysis methods were used for trend analysis in the study. For the province of Tokat, a decreasing trend was detected in June according to the MK test and in March according to the SR test. All approaches for Amasya province have revealed an increasing trend in November and Season-1 (S-1) season. According to all trend tests for Çorum province, a decreasing trend was determined in January. For Samsun province, an increasing trend was detected in April, Season-3 (S-3) and Season-4 (S-4) seasons according to MK test, April and Season-4 (S-4) season according to SR test. There were no trends in the annual solar radiation data for all stations. Graphs supporting other trend analysis results were obtained with innovative trend analysis.

**Key words:** Solar radiation, Mann-Kendal, Spearman's Rho, innovative trend analysis

**Reliability Evaluation of Circular m-Consecutive-k-out-of-n: F System with Non-Overlapping Runs Subjected to Shocks**

Fahrettin Özbey<sup>1</sup>

<sup>1</sup>Department of Statistics, Bitlis Eren University, 13000, Bitlis, Turkey

**Abstract**

In the working environment, most engineering system or unit is subject to external shocks. Reliability evaluation under shock models has attracted great deal of attention in the literature. In general, shock models can be classified into five groups: the cumulative shock model, extreme shock model, run shock model,  $\delta$ -shock model, and mixed shock model. These shock models focused on the magnitudes of shocks which were defined as random variable. Some recent works on shock models is that one of the main random variables is the number of components affected by successive shocks. In this study focused on the number of components destroyed by the shock, and phase-type distributions are also used for modeling the intervals between successive shocks, because phase-type distributions are useful and suitable for modeling the intervals between events. An F system is a system that fails when the number of failure components reaches a certain number and/or layout. Similarly, the circular m-consecutive-k-out-of-n: F system with non-overlapping runs consists of n components, which can have a circular order, and fails if and only if at least m non-overlapping runs of k consecutive components fail ( $m \leq k \leq n$ ). The main goal of this paper is to show how phase-type distributions can be used to determine the reliability of a circular m-consecutive-k-out-of-n: F system with non-overlapping runs that is subjected to shocks that destroy a random number of components.

**Key words:** Circular system, F system, Phase-type distribution, reliability

## PM10 and SO<sub>2</sub> predictions for the most polluted station in Ankara

Hatice ONCEL CEKIM<sup>1</sup>

<sup>1</sup>Department of Statistics, Hacettepe University, Ankara, Turkey

### Abstract

PM10 and SO<sub>2</sub> are the leading gases that cause many diseases and deaths caused by air pollution. Controlling these gases is one way to ensure acceptable living conditions in cities. Therefore, it is essential to identify the current situation to ensure future air pollution control. In this study, the singular spectrum analysis (SSA), the time series method, is used to both analyze the present air pollution of Ankara and make predictions, covering January 2010 / August 2020 periods. With SSA methods, the forecasts of PM10 and SO<sub>2</sub> measurements for the most polluted station in Ankara have obtained during 5, 10, 15, 20 and 25 horizons. According to the results, the monthly means of PM10 concentrations will be 55.28, 59.43, 64.62 and 69.47 in September, October, November and December 2020, respectively. On the other hand, these values are expected to be 7.48, 7.27, 6.93 and 6.53 for SO<sub>2</sub> in the four months of 2020. In the two-year average forecasts, according to the Air Quality standards determined by the European Commission, while the PM10 value will exceed the limit, it is estimated that the SO<sub>2</sub> value will remain below the limit.

**Key words:** Air pollution, forecast, PM10, singular spectrum analysis, SO<sub>2</sub>

## **Prioritization of Ergonomics Projects with Integration of CRITIC and CODAS Methods**

Gülin Feryal CAN<sup>1</sup>, Elif KILIÇ DELİCE<sup>2</sup>

<sup>1</sup>Başkent Üniversitesi, Mühendislik Fakültesi Endüstri Mühendisliği Bölümü, Ankara, Türkiye

<sup>2</sup>Atatürk Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Erzurum, Türkiye

### **Abstract**

Manual material handling tasks performed in workplaces are one of the most important causes of occupational musculoskeletal disorders. These disorders bring along many problems such as the inability of employees to continue working, having chronic health problems, decreasing their economic power, emergence of treatment and insurance costs for employers, and the emergence of new employee employment. For these reasons, it is important to improve ergonomic conditions manual material handling tasks in workplaces. In this direction, many ergonomics projects for different tasks are carried out in companies. However, each project brings with labor, time and capital requirements. Accordingly, it is important that companies should determine which ergonomics project has priority for implementation by making a business plan. In the study, Multi Criteria Decision Making (MCDM) structure was used in order to prioritize ergonomics projects. In this context, six different ergonomics projects were prioritized by considering the criteria as, complaints reduction rates, current risk level of the task, estimated annual earnings, estimated return period, number of people performing the task, level of back strain, level of shoulders and arms strain, the number of days not attending work to improve manual material handling task conditions. In the study, the CRITIC method was used to determine the weight of the risk criteria, and the CODAS method was used in ordering the ergonomics projects. In the CRITIC method, the level of knowledge about the criteria is determined by considering the level of relationship between these criteria and the amount of differentiation between the criteria values. In the CODAS method, by using Euclidean and Taxicab distances; the ergonomics project to be performed firstly can be defined. Both methods are objective methods that can be applied without decision maker's effect. For this reason, the rankings obtained appear as more sensitive results, far from subjectivity. The proposed integrated approach has been applied to six different manual material handling tasks performed in a machining company to improve ergonomically these tasks' conditions. According to the obtained results, estimated return period were emerged as the most important criterion for prioritizing ergonomics projects. Additionally, the projects which should be performed at first was determined as the fourth ergonomic project.

**Key words:** CRITIC, CODAS, ergonomics, multi criteria decision making, manual material handling

## **CRITIC ve CODAS Yöntemlerinin Entegrasyonu ile Ergonomi Projelerinin Önceliklendirilmesi**

### **Özet**

Çalışma alanlarında gerçekleştirilen elle malzeme taşıma görevleri, mesleki kas-iskelet sistemi rahatsızlıklarının en önemli nedenlerinden birisidir. Söz konusu rahatsızlıklar, çalışanların işe devam edememeleri, kronik sağlık problemlerine sahip olmaları, ekonomik güçlerinde düşüşlerin yaşanması,

işverenler açısından tedavi ve sigorta maliyetlerinin ortaya çıkması, yeni işçi istihdamının gündeme gelmesi gibi birçok problemi de beraberinde getirmektedir. Bu nedenlerle, çalışma alanlarında elle malzeme taşıma görev koşullarının ergonomik açıdan iyileştirmesi önem taşımaktadır. Firmalarda bu doğrultuda, farklı görevlere yönelik birçok ergonomi projesi gerçekleştirilmektedir. Ancak her bir proje, beraberinde iş gücü, zaman ve sermaye ihtiyacını da getirmektedir. Buna göre, firmaların bir iş planı yaparak; hangi ergonomi projesine öncelik vereceklerini belirlemeleri önem taşımaktadır. Çalışmada, ergonomi projelerinin önceliklendirilmesi amacıyla Çok Kriterli Karar Verme (ÇKKV) yapısından faydalanılmıştır. Bu kapsamda, elle malzeme taşıma görev koşullarının iyileştirilmesine yönelik altı farklı ergonomi projesi, şikâyetlerdeki azalma oranı, göreve ait mevcut risk düzeyi, tahmin edilen yıllık kazanç miktarı, tahmin edilen geri dönüş periyodu, görevi gerçekleştiren kişi sayısı, sırta oluşan zorlanma düzeyi, omuzlar ve kollarda oluşan zorlanma düzeyi, işe devam edilemeyen gün sayısı kriterleri dikkate alınarak önceliklendirilmiştir. Çalışmada, risk kriterlerinin ağırlıklarını belirlemek amacıyla CRITIC yöntemi uygulanmış, ergonomi projelerinin sıralanmasında ise, CODAS yöntemi kullanılmıştır. CRITIC yönteminde kriterlere ilişkin bilgi düzeyi, söz konusu kriterler arasındaki ilişki düzeyi ve kriter değerlerindeki değişim miktarı dikkate alınarak belirlenmektedir. CODAS yönteminde ise, Öklid ve Taksikab uzaklıkları kullanılarak; öncelikle gerçekleştirilmesi gereken ergonomi projesi tanımlanabilmektedir. Her iki yöntem de, karar verici etkisi olmaksızın uygulanabilen, objektif yöntemlerdir. Bu nedenle, elde edilen sıralamalar da öznellikten uzak, daha hassas sonuçlar olarak ortaya çıkmaktadır. Önerilen entegre yaklaşım, talaşlı imalat yapılan bir firmada gerçekleştirilen altı farklı elle malzeme taşıma görevinin ergonomik açıdan iyileştirilmesi amacıyla uygulanmıştır. Elde edilen sonuçlara göre, ergonomi projelerinin önceliklendirilmesinde en önemli kriter; tahmin edilen geri dönüş periyodu olarak ortaya çıkmıştır. Ayrıca, öncelikle gerçekleştirilmesi gereken ergonomi projesinin ise; dördüncü proje olduğu belirlenmiştir.

**Anahtar kelimeler:** CRITIC, CODAS, ergonomi, çok kriterli karar verme, elle malzeme taşıma

## Examining the Relationship between Organizational Commitment and Business

### Motivation - A Practice in the Public Sector

Zübeyir ÖZÇELİK<sup>1</sup>, Gaye MAT ÇELİK<sup>2</sup>

<sup>1</sup>Gazi Üniversitesi, Ankara, Türkiye

<sup>2</sup>Dicle Üniversitesi, Diyarbakır, Türkiye

#### Abstract

In this study, it is aimed to examine the organizational commitment and job motivation of the personnel working in a public institution. For this purpose, 768 permanent employees participated in the research. The study, which was examined by considering the organizational commitment model of Allen and Meyer, was examined in the context of emotional commitment, continuance commitment and normative commitment dimensions. In terms of work motivation, the levels in Herzberg's double factor theory were discussed in the context of extrinsic and intrinsic work motivation, and the relationship between organizational commitment levels was revealed by the correlation method. The personnel participating in the research show a commitment to the institution at the level of 87.5% and it has been concluded that this commitment is predominantly in the direction of emotional attachment. On the other hand, the lowest tendency was observed in the continuation commitment, which is a cost-based type of commitment. Moreover, intrinsic motivation factors have a significant relationship on emotional commitment, continuance commitment and normative commitment, which are sub-dimensions of organizational commitment; It is possible to say that extrinsic motivation factors have a significant relationship only on emotional commitment. Based on these results, some suggestions about increasing organizational commitment and work motivation in the public sector are included in our study.

**Key words:** Continuity commitment, emotional commitment, job motivation, normative commitment, organizational commitment

## Örgütsel Bağlılık ve İş Motivasyonu Arasındaki İlişkilerin İncelenmesi - Kamu Sektöründe Bir Uygulama

#### Özet

Bu çalışmada, bir kamu kuruluşunda görev yapmakta olan personelin örgütsel bağlılık ve iş motivasyonunun incelenmesi amaçlanmaktadır. Bu amaçla, 768 kadrolu çalışan araştırmaya katılmıştır. Allen ve Meyer'in örgütsel bağlılık modeli ele alınarak incelenen çalışma duygusal bağlılık, devam bağlılığı, normatif bağlılık boyutları bağlamında ele alınmıştır. İş motivasyonu konusunda da Herzberg'in çift faktör kuramındaki düzeyler dışsal ve içsel iş motivasyonu bağlamında ele alınarak örgütsel bağlılık düzeyleri ile arasındaki ilişki korelasyon yöntemiyle ortaya koyulmuştur. Araştırmaya katılan personel kurumuna %87,5 düzeyinde bağlılık göstermekte olup söz konusu bu bağlılığın ağırlıklı olarak duygusal bağlılık yönünde olduğu sonucuna ulaşılmıştır. En düşük eğilimin ise maliyete dayalı bir bağlılık türü olan devam bağlılığı boyutunda olduğu görülmüştür. Ayrıca, içsel motivasyon faktörlerinin örgütsel bağlılığın alt boyutları olan duygusal bağlılık, devam bağlılığı ve normatif bağlılık üzerinde anlamlı bir ilişkisi olduğunu; dışsal motivasyon faktörlerinin ise yalnızca duygusal bağlılık

üzerinde anlamlı bir ilişkisi olduğunu söylemek mümkündür. Bu sonuçlardan hareketle çalışmamızda kamu sektöründe örgütsel bağlılığın ve iş motivasyonunun artırılması konularına yönelik birtakım önerilere yer verilmiştir.

**Anahtar kelimeler:** Devam bağlılığı, duygusal bağlılık, iş motivasyonu, normatif bağlılık, örgütsel bağlılık

## Normal Dağılım Varsayımı Sağlanamadığı Durumda Süreç Ortalamasını Takip etmek için Çarpıklık ve Basıklık Düzeltmesi Yöntemi

Sevgi YURT ÖNCEL<sup>1</sup>, Muhammed USLU<sup>1</sup>

<sup>1</sup>Kırıkkale Üniversitesi Fen-Edebiyat Fakültesi, İstatistik Bölümü, Yahşihan, Kırıkkale, Türkiye

### Özet

Bu çalışmada, Normal dağılım varsayımının sağlanamadığı durumlar için, süreç ortalamasını takip etmek için kullanılan klasik Shewhart grafiklerinin yerine alternatif grafikler önerilmiştir. Örneklem dağılımının önemli derecede çarpık olduğu durumda çarpıklık düzeltmesine dayalı kontrol grafiklerinin çizimi, örneklem dağılımı simetrik ama önemli derecede basık olduğu durumda ise basıklık düzeltmesi yöntemine dayalı kontrol grafiklerinin çizimi üzerinde durulmuştur. Gerçek bir sanayi ürününe ait veri seti üzerinde uygulama yapılmıştır.

### Anahtar kelimeler:

### Abstract

In this study, alternative charts are proposed to replace the classical Shewhart charts used to monitoring the process mean for cases where the normal distribution assumption cannot be achieved. The drawing of control charts based on skewness correction when the sample distribution is significantly skewed, and drawing of control charts based on kurtosis correction method when the sample distribution is symmetrical but significantly kurtosis is emphasized. Application has been made on a data set of a real industrial product.

### Key words:

### Giriş

İstatistiksel süreç kontrolünde grafik yöntemi oldukça tercih edilen bir yöntemdir. Kalite kontrol grafiklerinin kullanılması fikri 1926 yılında Shewhart tarafından ortaya atılmıştır. Eğer bu grafiklerin kontrol limitleri iyi belirlenirse bir sürecin takibinde çok etkili araçlardır. Klasik Shewhart kontrol grafikleri, kalite değişkeninin normal dağılıma sahip olduğu varsayımına dayanır. Ancak kalite değişkeni her zaman normal dağılıma sahip olmayabilir. Özellikle küçük örneklem hacmi ile çalışıldığından Normal dağılım varsayımında ısrarcı olmak yanlış alarm oranının artmasına yol açar ve süreç hakkında hatalı karar vermeye neden olur.

Bu problemin ortadan kaldırılması için literatürde önerilen bazı grafik yöntemleri dayanıklı kontrol grafiği, medyan kontrol grafiği, ön istatistiklere dayalı kontrol grafiği, ağırlıklı varyans grafiğidir. Bu çalışmada kalite değişkeni çarpık ya da basık bir dağılıma sahip olduğunda Cornish-Fisher açılımına dayalı olan çarpıklık düzeltmesi ve basıklık düzeltmesi yöntemi tartışılacaktır. Örneklem alınıldığı kitlenin dağılımının ve parametrelerinin bilinmediği durumda çarpıklık ve basıklık katsayısı tahmin edilerek Cornish-Fisher açılımına dayalı kontrol grafiği limitleri elde edilecektir.



$X_1, X_2, \dots, X_n \sim N(\mu_0, \sigma^2)$  iken  $H_0 : X \sim N(\mu_0, \sigma^2)$  hipotezine karşılık  $H_1 : X \sim N(\mu_1 = \mu_0 + \delta\sigma, \sigma^2)$  hipotezi ile Normal dağılıma sahip olan bir kitlenin ortalamasındaki değişim (kaymalar) araştırılabilir. Süreç kontrol altında iken örneklem ortalaması  $\bar{x}$  'nin kontrol limitlerinin dışına çıkması olasılığı  $\alpha = 1 - P(AKL \leq \bar{X} \leq ÜKL | \mu = \mu_0)$  dir. Burada yer alan  $\mu_0$  ve  $\mu_1$  parametreleri sırasıyla süreç kontrol altında ve süreç kontrol dışında iken  $X$  rasgele değişkeninin kitle ortalamasını göstermektedir.  $X$  rasgele değişkeninin dağılımı normal dağılım olduğunda  $\alpha = 2\Phi(-\delta)$  'dir. Burada  $\Phi$ , standart normal dağılımın dağılım fonksiyonunu göstermektedir. Süreç kontrol altındayken kitle ortalamasının  $\mu_0$  'dan  $\mu_1$  'e kayması halinde bu kaymayı birbirini izleyen noktalarda tespit edememe olasılığı  $\beta = P(AKL \leq \bar{X} \leq ÜKL | \mu_1 = \mu_0 + \delta\sigma)$  dir. Kontrol grafiğinde yer alan örneklem ortalamalarının kontrol limitleri dışına çıkmadan önce grafikte işaretlenmesi beklenen nokta sayısı, ortalama çalışma uzunluğu ( $OÇU$ ) olarak adlandırılır ve kontrol grafiğinin performansını değerlendirmek için kullanılır. Süreç kontrol altındayken ortalama çalışma uzunluğu  $OÇU = \frac{1}{\alpha}$  olarak

hesaplanır ve süreç kontrol dışında iken  $OÇU = \frac{1}{1-\beta}$  dir. Örneklemin dağılımı Normal dağılım iken ortalama için Shewhart Kontrol grafiğinin ortalama çalışma uzunluğu  $\alpha = 0.0027$  için  $1/0.0027 = 370$  'dir. Yani süreç kontrol altındayken ortalama 370 noktada bir kez kontrol grafiği, doğal nedenlerden dolayı süreç kontrol dışında sinyali verecektir. Süreç kontrol altındayken  $OÇU$ 'nun büyük olması istenir. Süreç kontrol dışına çıkmışsa erken uyarı almak için  $OÇU$ 'nun küçük olması istenir.

### Klasik Shewart Kontrol Grafikleri

Bir üretim sürecinin, istatistiksel kalite kontrolünün yapılabilmesi için uygun bir kalite özelliği ve bu özelliğe uyumlu kontrol grafiği seçilmelidir. Ölçülebilen kalite değişkenlerin dağılım ortalamasında olabilecek kaymaları takip edebilmek amacıyla için  $\bar{x}$  kontrol grafiği çizilir. Bir üretim sürecinin istatistiksel olarak kontrol altında olup olmadığını incelemek için Faz I'de ilgilenilen kalite değişkenine ait gözlemler elde edilir ve değişkenin dağılımının bilinmeyen parametreleri varsa tahmin edilir ve kontrol limitleri elde edilir. Faz II aşamasında ise süreçte bir değişimin olup olmadığı takip edilir.

Ortalaması  $\mu$  standart sapması  $\sigma$  olan normal dağılımın kitle ortalamasındaki değişimi takip etmek için çizilecek Shewhart kontrol grafiği limitleri,

$$ÜKL_{\bar{x}} = \mu_{\bar{x}} + Z_{\alpha/2}\sigma_{\bar{x}}$$

$$OÇ_{\bar{x}} = \mu_{\bar{x}}$$

$$AKL_{\bar{x}} = \mu_{\bar{x}} - Z_{\alpha/2}\sigma_{\bar{x}}$$

biçimindedir.  $P(Z > Z_{\alpha/2}) = \frac{\alpha}{2}$  eşitliğinde yer alan  $\alpha = 0.0027$ , iken  $Z_{\alpha/2} = Z_{0.00135} = 3$  değeri kontrol limiti katsayısı olarak kullanılmaktadır. Bu nedenle klasik Shewhart kontrol grafiklerinde  $\mu_{\bar{x}} \mp 3\sigma_{\bar{x}}$  sınırlarında  $\alpha = 0.0027$  'dir [15].

Süreç ortalamasındaki kaymaları tespit edebilmek için çizilen kontrol grafikleri kitle parametreleri biliniyorken ve bilinmiyorken durumları için farklı hesaplamalarla çizilir.

$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$  iken  $\bar{X} \sim N\left(\mu_{\bar{x}}, \frac{\sigma_{\bar{x}}^2}{\sqrt{n}}\right)$  olduğundan dolayı kitle parametreleri  $\mu$  ve  $\sigma$  biliniyorken grafiğin limitleri  $\bar{x}$  grafiğinin limitleri

$$\dot{U}KL_{\bar{x}} = \mu_x + 3 \frac{\sigma_x}{\sqrt{n}}$$

$$OC_{\bar{x}} = \mu_x$$

$$AKL_{\bar{x}} = \mu_x - 3 \frac{\sigma_x}{\sqrt{n}}$$

biçimindedir. Eğer kitle parametreleri  $\mu$  ve  $\sigma$  bilinmiyorsa  $\hat{\mu} = \bar{\bar{X}} = \frac{1}{m} \sum_{i=1}^m \bar{X}_i$  ve  $\hat{\sigma}_{\bar{x}} = \frac{\bar{R}}{d_2 \sqrt{n}}$  tahmin edicileri kullanılarak

$$\dot{U}KL_{\bar{x}} = \bar{\bar{X}} + 3 \frac{\bar{R}}{d_2 \sqrt{n}} = \bar{\bar{X}} + 3A_2$$

$$OC_{\bar{x}} = \bar{\bar{X}}$$

$$AKL_{\bar{x}} = \bar{\bar{X}} - 3 \frac{\bar{R}}{d_2 \sqrt{n}} = \bar{\bar{X}} - 3A_2$$

formülleri ile kontrol grafiğinin limitleri bulunur. Burada  $A_2 = \frac{3}{d_2 \sqrt{n}}$  dir.  $\bar{X}_1, \bar{X}_2, \bar{X}_3, \dots, \bar{X}_m$  alt

örneklemere ait ortalamalar olmak üzere,  $i$ ' inci örneklemin ortalaması,  $\bar{X}_i = \frac{1}{n} \sum_{j=1}^n X_{ij}$  ile hesaplanan

genel ortalama  $\bar{\bar{X}} = \frac{1}{m} \sum_{i=1}^m \bar{X}_i$ ,  $\mu$ ' nün yansız bir tahmin edicisidir ve  $\bar{R} = \frac{1}{m} \sum_{i=1}^m R_i$ , her alt örnek grubundan hesaplanan genişlik değerlerinin aritmetik ortalamasıdır.

### Çarpıklık Düzeltmesi Yöntemi ile Shewhart Kalite Kontrol Grafiklerinin Oluşturulması

İlgilenilen kalite değişkeninin dağılımı için normal dağılım varsayımı sağlanamazsa I. tip hata olasılığını büyütmeyen uygun diğer yöntemlerle grafik çizmek gerekir. Çarpık bir dağılıma sahip örneklem için klasik Shewhart kontrol grafikleri yerine Cornish-Fisher açılımına dayanan çarpıklık düzeltmesi yöntemi kullanılabilir. Çarpıklık ölçüleri simetrik dağılıma sahip olmayan bir veri setinin normal dağılımdan hangi düzeyde ve ne yönde uzaklaştığını saptamaya yarayan ölçülerdir. Büyüklüğü çarpıklığın kuvvetini, işareti ise yönünü göstermektedir. Örneklemin alındığı kitlenin dağılımı ve parametreleri bilinmiyorken dağılımın çarpıklık katsayısı

$$\hat{k}_3 = \frac{1}{nr} \sum_{i=1}^r \sum_{j=1}^n \left( \frac{x_{ij} - \bar{\bar{x}}}{\hat{\sigma}} \right)^3$$

formülü ile tahmin edilir. Burada  $\bar{\bar{X}} = \frac{1}{nr} \sum_{i=1}^r \sum_{j=1}^n X_{ij}$  olmak üzere  $\sigma$ ' nin tahmin edicisi olarak

$\hat{\sigma}_x = \sqrt{\frac{1}{nr-1} \sum_{i=1}^r \sum_{j=1}^n (X_{ij} - \bar{\bar{X}})^2}$  kullanılır. Bir dağılımın çarpıklığı,  $k_3$ ' ten yararlanarak  $k_3=0$  ise

dağılım simetriktir,  $k_3 > 0$  ise dağılım sağa çarpıktır,  $k_3 < 0$  ise dağılımın sola çarpıktır biçiminde değerlendirilir.

Çarpıklık düzeltmesi yöntemi Cornish Fisher açılımına dayanır.  $X$  ortalaması 0, standart sapması 1 olan standartlaştırılmış rasgele değişken olsun.  $X_\alpha$ ,  $X$  rasgele değişkeninin dağılımının  $\alpha$  yüzdeliği olsun.  $Z_\alpha$  standart normal dağılıma sahip rasgele değişken olsun.  $k_r$ ,  $X$ 'in  $r$ . kümülanı ( $r \geq 3$ ) olsun.

O zaman  $X_\alpha$ ,

$$X_\alpha = Z_\alpha + \frac{1}{6}(Z_\alpha^2 - 1)k_3 + \frac{1}{24}(Z_\alpha^3 - 3Z_\alpha)k_4 - \frac{1}{36}(2Z_\alpha^3 - 5Z_\alpha)k_5^2 + \dots$$

biçiminde Cornish-Fisher açılımına sahiptir (Cornish ve Fisher, 1938).  $3\sigma$  sınırları için,  $\alpha_1 = 1 - 0.00135 = 0.99865$  ve  $\alpha_2 = 1 - \alpha_1 = 1 - 0.99865 = 0.00135$  olarak alınırsa  $U_{\alpha_1} = 3$ ,  $U_{\alpha_2} = -3$  bulunarak (3.2) eşitliğinde yerine koyulduğunda  $X_{\alpha_1}$  ve  $X_{\alpha_2}$ 'nin Cornish-Fisher açılımı,

$$X_{\alpha_1} = 3 + \frac{4}{3}k_3 + \frac{3}{4}k_4 + \dots$$

$$X_{\alpha_2} = -3 + \frac{4}{3}k_3 - \frac{3}{4}k_4 + \dots$$

biçiminde elde edilir ancak  $X$ 'in bir çok kümülanını içeren karmaşık forma sahiptir. Çarpıklık katsayısı  $k_3$ 'ün kalite kontrol grafikleri üzerindeki etkisi ile ilgilendiğimizden dolayı

$$X_{\alpha_1} = 3 + \frac{4}{3}k_3 - \frac{13}{12}k_3^2 + Y_1$$

$$X_{\alpha_2} = -3 + \frac{4}{3}k_3 + \frac{13}{12}k_3^2 + Y_2$$

formülleri kullanılır. Burada yer alan  $Y_1$  ve  $Y_2$  terimlerinin pratikte hesaplanması ve kullanılması zor olduğundan kalite kontrol grafiklerinin limitleri

$$\dot{ÜKL} = X_{\alpha_1} = 3 + \frac{4}{3}k_3 - \frac{13}{12}k_3^2 + Y_1 \approx 3 + \frac{4}{3}k_3h(k_3)$$

$$AKL = X_{\alpha_2} = -3 + \frac{4}{3}k_3 + \frac{13}{12}k_3^2 + Y_2 \approx -3 + \frac{4}{3}k_3h(k_3)$$

formülleri ile yaklaşık olarak hesaplanır. Burada yer alan  $h(\cdot)$  fonksiyonunun özellikleri,  $h(k_3) = -h(k_3)$ ,  $\lim_{|k_3| \rightarrow \infty} k_3h(k_3) = 0$  ve  $h(0) = 1$  dir. Chan ve Cui (2003), çalışmalarında basitlik sağlaması için  $h$  fonksiyonunu  $h(k_3) = \frac{1}{1 + \delta k_3^2}$  biçiminde seçmiştir ve pek çok çarpık dağılım (Weibull, Log-normal, Burr ... gibi) için yapılan simülasyon çalışmalarında  $\delta$  parametresini 0,2'ye çok yakın bulduklarından dolayı  $h(k_3) = \frac{1}{1 + 0.2k_3^2}$  olarak almışlardır. Dolayısıyla bireysel veriler (individual observation) kalite kontrol grafiklerinin üst ve alt kontrol limiti

$$\dot{ÜKL} = 3 + \frac{\frac{4}{3}k_3}{1 + 0.2k_3^2}$$

$$AKL = -3 + \frac{\frac{4}{3}k_3}{1 + 0.2k_3^2}$$

biçiminde oluşturulur.

**Süreç dağılımı ve parametreleri biliniyorken çarpıklık düzeltmeli  $\bar{X}$  kontrol kartı limitleri**

$X_1, X_2, X_3, \dots, X_n$ , ortalaması  $\mu_X$ , standart sapması  $\sigma_X$  ve çarpıklığı  $k_3$  olarak bilinen dağılımın alt örneklem grubu olsun. Bu durumda Chan ve Cui (2003)'nin önerdiği çarpıklık düzeltmeli  $\bar{X}$  ve  $R$  kontrol grafiği limitleri

$$\text{ÜKL}_{\bar{X}} = \mu_{\bar{X}} + \left( 3 + \frac{\frac{4}{3}k_3(\bar{x})}{1 + 0.2k_3^2(\bar{x})} \right) \sigma_{\bar{X}} = \mu_X + (3 + c_4^*) \frac{\sigma_X}{\sqrt{n}}$$

$$\text{OC}_{\bar{X}} = \mu_{\bar{X}}$$

$$\text{AKL}_{\bar{X}} = \mu_{\bar{X}} + \left( -3 + \frac{\frac{4}{3}k_3(\bar{x})}{1 + 0.2k_3^2(\bar{x})} \right) \sigma_{\bar{X}} = \mu_X + (-3 + c_4^*) \frac{\sigma_X}{\sqrt{n}}$$

biçimindedir. Burada  $c_4^* = \frac{\frac{4}{3}k_3(\bar{x})}{1 + 0.2k_3^2(\bar{x})}$  dir ve eğer dağılım simetrik ise  $c_4^* = 0$  olur ve  $\bar{X}$  grafiği klasik Shewhart grafiğinin limitlerine sahip olur.

**Süreç dağılımı ve parametreleri bilinmiyorken çarpıklık düzeltmeli  $\bar{X}$  kontrol kartı limitleri**

Eğer dağılımın ortalaması ve varyansı bilinmiyorsa kontrol grafiklerinin oluşturulabilmesi için bu parametrelerin tahmin edicileri kullanılır. Eğer örneklemin alındığı kitlenin çarpıklık katsayısı  $k_3$  ise  $\bar{X}$  'nin dağılımının çarpıklık katsayısı  $\frac{k_3}{\sqrt{n}}$  dir. Bu durumda çarpıklık düzeltmesiyle elde edilen  $\bar{X}$  kontrol grafiği limitleri,

$$\text{ÜKL}_{\bar{X}} = \bar{\bar{X}} + \left( 3 + \frac{4\hat{k}_3 / (3\sqrt{n})}{1 + 0.2\hat{k}_3^2 / n} \right) \frac{\bar{R}}{d_2^* \sqrt{n}} = \bar{\bar{X}} + A_U^* \bar{R}$$

$$\text{OC}_{\bar{X}} = \bar{\bar{X}}$$

$$\text{AKL}_{\bar{X}} = \bar{\bar{X}} + \left( -3 + \frac{4\hat{k}_3 / (3\sqrt{n})}{1 + 0.2\hat{k}_3^2 / n} \right) \frac{\bar{R}}{d_2^* \sqrt{n}} = \bar{\bar{X}} - A_L^* \bar{R}$$

olarak hesaplanır. Burada,  $A_U^* = \left( 3 + \frac{4\hat{k}_3 / (3\sqrt{n})}{1 + 0.2\hat{k}_3^2 / n} \right) \frac{1}{d_2^* \sqrt{n}}$  ve  $A_L^* = \left( -3 + \frac{4\hat{k}_3 / (3\sqrt{n})}{1 + 0.2\hat{k}_3^2 / n} \right) \frac{1}{d_2^* \sqrt{n}}$  tablo değerleri  $k_3$  değerine ve örneklem büyüklüğüne göre Chan ve Cui (2003) 'nin çalışmasında verilen tabloda olarak yer almaktadır.

Çarpıklık düzeltmesi yöntemiyle elde edilmeye çalışılan kontrol grafikleri için gerekli katsayılar olan  $d_2^*$  ve  $d_3^*$  klasik Shewhart kontrol grafikleri için gerekli katsayılar olan  $d_2$  ve  $d_3$ , ün yerine geçer.

$d_2^*$  sabitinin değerleri Chan ve Cui (2003,s.564)'de yer almaktadır. Eğer hesaplanan  $k_3$  değeri tabloda yer almıyorsa interpolasyon formülü ile hesaplanır.

### Basıklık Düzeltmesi Yöntemi ile Shewhart Kalite Kontrol Grafiklerinin Oluşturulması

Bir veri setinin dağılımının basıklığı dördüncü moment yardımıyla hesaplanabilir. Dağılımın dördüncü dereceden momenti, standart sapmasının dördüncü kuvvetine bölünürse standart bir basıklık ölçüsü elde edilmiş olur. Yani, momentler yardımıyla basıklık katsayısı  $k_4 = \frac{\mu_4}{\sigma^4}$  şekilde formüle edilir. Bir dağılımın basıklığı  $k_4$ 'ten yararlanarak  $k_4=3$  ise dağılımın basıklığı normal dağılıma uygundur,  $k_4 < 3$  ise dağılımın basıklığı normal dağılıma göre daha basıktır,  $k_4 > 3$  ise dağılımın basıklığı normal dağılıma göre daha sivridir biçiminde değerlendirilir.

$X_{i1}, X_{i2}, \dots, X_{in}$ ,  $i = 1, 2, \dots, r$  ortalaması  $\mu$ , standart sapması  $\sigma$ , çarpıklığı 0, basıklığı  $k_4$  olan kitle dağılımının  $n$  hacimli  $r$  tane örneklem grubu olsun. Eğer dağılımın parametreleri bilinmiyorsa kontrol kartlarının oluşturulabilmesi için basıklık katsayısının tahmin edilmesi gerekmektedir.  $k_4^*$ , basıklık katsayısı  $k_4$  için bir tahmin edicidir ve

$$\hat{k}_4 = \frac{1}{nr} \sum_{i=1}^r \sum_{j=1}^n \left( \frac{x_{ij} - \bar{x}}{\hat{\sigma}_x} \right)^4$$

biçiminde hesaplanır. Süreç dağılımı simetrik ancak sıfırdan büyük bir basıklığa sahip olduğunda kalite kontrol grafikleri için basıklık düzeltmesi yöntemi kullanılabilir. Bu yöntem Tadikamalla ve Popescu (2007) tarafından geliştirilmiş olup Chan ve Cui (2003)'e benzer şekilde Cornish-Fisher açılımına dayalı basıklık düzeltmesi yapılmıştır.  $X$  rasgele değişkeninin dağılımının simetrik bir dağılım olduğunu varsayalım. Cornish-Fisher açılımına göre,  $X_{\alpha_1}$  ve  $X_{\alpha_2}$ ,  $k_4$  basıklık katsayısının ve bazı hata terimlerinin bir fonksiyonu olarak

$$X_{\alpha_1} = 3 + \frac{3}{4}k_4 + E_1 \approx 3 + g(k_4)$$

$$X_{\alpha_2} = -3 - \frac{3}{4}k_4 + E_2 \approx -3 - g(k_4)$$

biçiminde yazılabilir.

Tadikamalla ve Popescu (2007), çalışmalarında polinom, doğrusal, üstel, sigmoidal, büyüme, güç yasası ve verim yoğunluğu modelleri dahil olmak üzere farklı modeller denediklerini ve verileri

açıklamak için en iyi modelin  $g(k_4) = \frac{bk_4}{1 + ck_4}$  formunda olduğunu gördüklerini belirtmişler ve lineer

olmayan regresyon yöntemiyle katsayıları  $b \approx 1$  ve  $c \approx 0.33$  olarak hesaplamışlardır. Bu regresyon denklemi için  $R^2 = 0.99997$  ve standart hata 0,00254'tür. Buna göre kitle dağılımının  $\alpha$ -kuantilleri kullanılarak bireyler (individual observation) için kontrol limitleri

$$\hat{ÜKL} = X_{\alpha_1} \approx 3 + \frac{k_4}{1 + 0.33k_4}$$

$$OÇ = 0$$

$$AKL = X_{a_2} \approx -3 - \frac{k_4}{1 + 0.33k_4}$$

biçiminde olur. Simetrik leptokurtik dağılımlar olan Burr dağılımı, t dağılımı ve Exponential Power dağılımının basıklık değerleri için bu denklemlerin sağlam sonuçlar verdiği görülmüştür. Ancak rasgele değişkenlerin aldığı değerler sınırlı veya kesilmiş (tuncated) dağılımlar için bu yöntemin iyi sonuç vermediği görülmüştür. (Tadikamalla ve Popescu, 2007). Normal dağılıma sahip veriler için bu formüller klasik Shewhart kontrol grafiklerine indirgenir.

### ***Süreç dağılımı ve parametreleri biliniyorken basıklık düzeltmeli $\bar{X}$ kontrol kartı limitleri***

Kalite karakteristiği  $X$ 'in ortalaması  $\mu_X$ , standart sapması  $\sigma_X$  ve basıklığı  $k_4$  olan simetrik bir dağılıma sahip olduğunu varsayalım. Bu durumda Tadikamalla ve Popescu (2007), Chan ve Cui'ye benzer şekilde  $\bar{X}$  kontrol kartı limitleri,

$$\bar{U}KL_{\bar{X}} = \mu_X + \left( 3 + \frac{k_4(\bar{X})}{1 + 0.33k_4(\bar{X})} \right) \frac{\sigma_X}{\sqrt{n}}$$

$$OC_{\bar{X}} = \mu_X$$

$$AKL_{\bar{X}} = \mu_X - \left( 3 + \frac{k_4(\bar{X})}{1 + 0.33k_4(\bar{X})} \right) \frac{\sigma_X}{\sqrt{n}}$$

(5.3)'de gösterilen kontrol kartı limitlerini önermişlerdir. Burada  $k_4(\bar{X})$ , örneklem ortalaması  $\bar{X}$ 'nin basıklık katsayısıdır ve  $k_4(\bar{X}) = \frac{k_4(X)}{n}$  dir.

### ***Süreç dağılımı ve parametreleri bilinmiyorken basıklık düzeltmeli $\bar{X}$ kontrol kartı limitleri***

Eğer kitle dağılımı ve parametreleri bilinmiyorsa kontrol kartı limitleri Shewhart kontrol kartı limitlerine benzer şekilde aşağıdaki gibi olur.

$\bar{X}$  kontrol kartı limitleri,

$$\bar{U}KL_{\bar{X}} = \bar{\bar{X}} + \left( 3 + \frac{k_4(\bar{X})}{1 + 0.33k_4(\bar{X})} \right) \frac{\bar{R}}{d_2^* \sqrt{n}} \equiv \bar{\bar{X}} + A_2^* \bar{R}$$

$$OC_{\bar{X}} = \bar{\bar{X}}$$

$$AKL_{\bar{X}} = \bar{\bar{X}} - \left( 3 + \frac{k_4(\bar{X})}{1 + 0.33k_4(\bar{X})} \right) \frac{\bar{R}}{d_2^* \sqrt{n}} \equiv \bar{\bar{X}} - A_2^* \bar{R}$$

$d_2^*$ ,  $A_2^*$  sabit değerlerinin nasıl hesaplandığı ve farklı  $n$  değerlerine göre tablo hali Tadikamalla ve Popescu (2007)'nin makalelerinde gösterilmiştir.

Tablo 1. Basıklık Düzeltmesi için  $A_2^*$  Sabit Değerleri

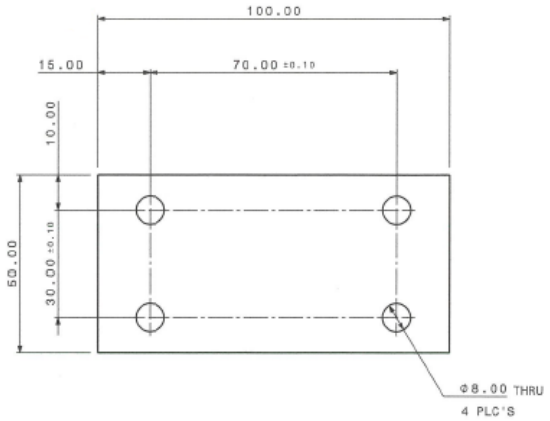
n	Basıklık						
	0	1	2	3	4	5	6
2	1.881	2.187	2.425	2.613	2.766	2.893	3.001
3	1.023	1.146	1.248	1.333	1.406	1.469	1.524
4	0.729	0.796	0.854	0.905	0.949	0.988	1.023
5	0.577	0.619	0.657	0.69	0.721	0.748	0.773
6	0.483	0.511	0.538	0.562	0.584	0.604	0.622
7	0.419	0.439	0.458	0.476	0.493	0.508	0.522
8	0.373	0.386	0.401	0.415	0.428	0.44	0.451
9	0.337	0.347	0.358	0.368	0.379	0.389	0.398
10	0.308	0.315	0.324	0.332	0.341	0.349	0.356
12	0.266	0.269	0.274	0.279	0.285	0.291	0.296
15	0.223	0.222	0.225	0.228	0.231	0.235	0.238
20	0.18	0.176	0.176	0.177	0.178	0.18	0.182
25	0.153	0.148	0.147	0.147	0.147	0.148	0.149

### Uygulama

Bu kesimde yukarıda verilen formüllerin ve tabloların nasıl kullanılacağını açıklamak için bir veri seti üzerinde uygulama yapılmıştır. Bunun için bir fabrikadan alınan ayar plakası verileri kullanılmıştır.



Şekil 1. Ayar Plakası



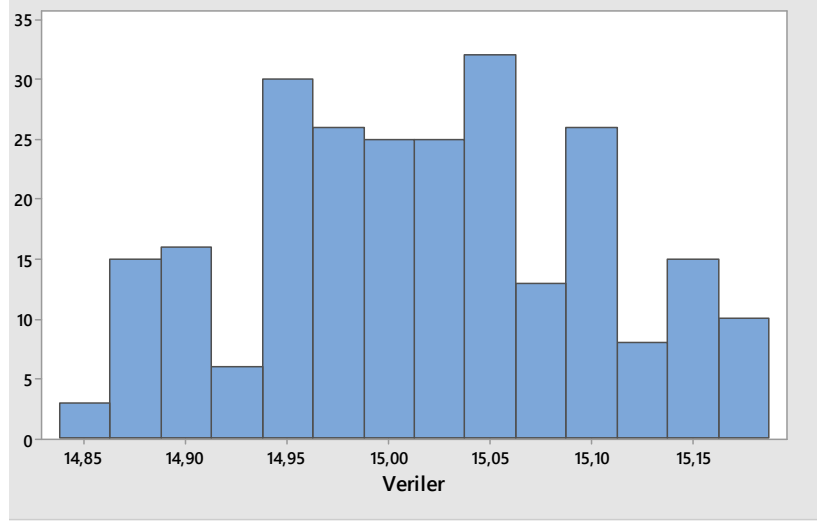
Şekil 2. Ayar Plakası Teknik Resmi

Tablo 2. Veri Seti - Ayar Plakası

Grup	1	2	3	4	5	$\bar{X}$	$R$
1	15.1	15.05	15.05	15.05	15.06	15.062	0.05
2	15.06	15.05	15	15.03	15.03	15.034	0.06
3	15.05	15.1	15.11	15.09	15.08	15.086	0.06
4	15.06	14.95	14.99	14.9	15.02	14.984	0.16
5	15	15.03	15.03	15.06	15.05	15.034	0.06
6	15.16	15.11	15.14	15.14	14.93	15.096	0.23
7	15.12	14.95	15	14.88	15.04	14.998	0.24
8	14.97	14.91	15.1	15.03	15.16	15.034	0.25
9	15.02	14.96	15.14	15.16	14.97	15.05	0.2
10	14.96	14.86	15.08	15.13	14.98	15.002	0.27
11	14.89	14.94	15.18	15.1	15.14	15.05	0.29
12	15.17	14.85	14.96	14.95	15.17	15.02	0.32
13	14.89	15.04	14.89	15.1	14.85	14.954	0.25
14	15.13	15.11	15.15	15.1	14.95	15.088	0.2
15	14.98	14.88	15.11	14.97	15.09	15.006	0.23
16	14.91	14.97	15.11	14.9	15.18	15.014	0.28
17	15.1	15.14	15.14	15.1	14.96	15.088	0.18
18	15.07	14.98	14.94	14.91	14.88	14.956	0.19
19	14.97	14.98	15.18	14.99	15.14	15.052	0.21
20	14.87	14.92	15.08	14.97	15.15	14.998	0.28
21	14.87	14.9	14.87	14.91	15.01	14.912	0.14
22	15.11	14.98	14.87	14.99	15.1	15.01	0.24
23	14.99	14.97	14.91	14.87	14.88	14.924	0.12
24	15.13	15.18	14.94	15.01	15.18	15.088	0.24
25	15.07	14.88	15.18	15.1	15.1	15.066	0.3
26	15.18	14.9	14.97	14.96	14.98	14.998	0.28
27	14.87	15.03	14.9	15.13	14.96	14.978	0.26
28	15.07	15.06	14.95	14.98	15.13	15.038	0.18
29	15.14	15.08	15.02	15.04	14.99	15.054	0.15
30	15.07	15.14	15.18	14.87	15.07	15.066	0.31
31	15.11	15	15.01	15.01	15.04	15.034	0.11
32	14.99	14.99	14.97	14.98	15.05	14.996	0.08
33	15.04	15.03	15.08	15.05	15.03	15.046	0.05
34	15.03	14.97	14.95	14.95	14.98	14.976	0.08
35	14.94	14.92	14.93	14.88	15.04	14.942	0.16
36	14.97	14.91	15.1	15.03	15.06	15.014	0.19
37	15.1	15.07	15.01	15.05	14.99	15.044	0.11
38	14.9	14.91	14.87	14.88	14.92	14.896	0.05
39	15.02	15.1	15.12	15.08	15.08	15.08	0.1
40	15.05	15.09	15.09	15.03	15.04	15.06	0.06
41	15.1	15.04	15.04	14.99	14.95	15.024	0.15
42	14.92	14.96	14.96	14.99	15.02	14.97	0.1
43	15.05	15.05	15.12	15.14	15.05	15.082	0.09
44	15.02	14.94	15.02	15.03	14.99	15	0.09
45	15.02	15.05	15.06	15.02	14.97	15.024	0.09
46	14.98	15	14.96	15.05	14.98	14.994	0.09
47	14.94	14.94	15.02	14.94	14.97	14.962	0.08
48	14.96	14.94	15.03	15.02	15.05	15	0.11
49	15.02	15.01	14.96	14.95	14.94	14.976	0.08
50	14.97	14.97	15.01	14.99	15	14.988	0.04
						$\bar{X} = 15.0169$	$\bar{R} = 0.1628$

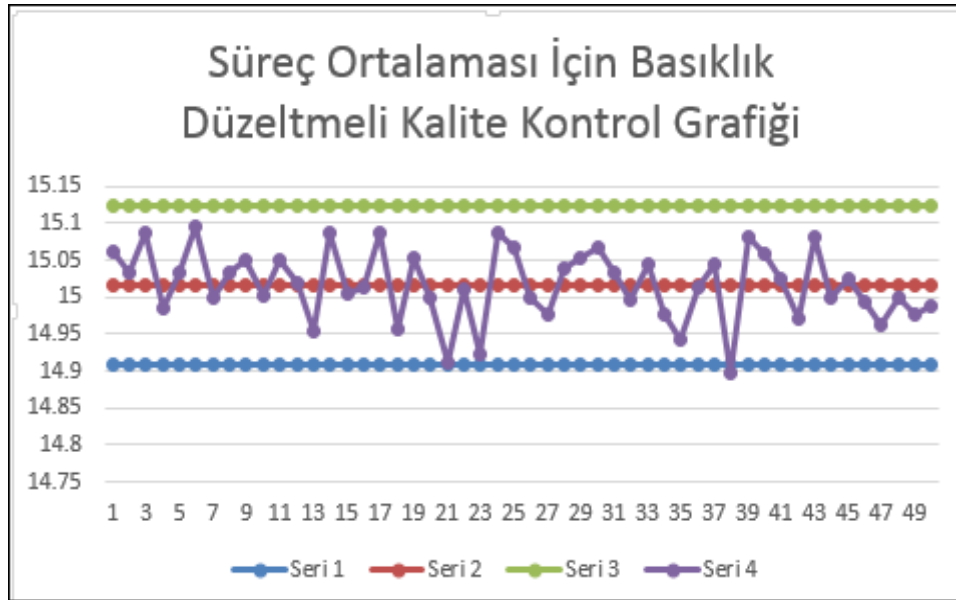


Bu veri seti için yapılan analizler sonucunda  $\bar{X} = 15.0169$ ,  $\bar{R} = 0.1628$ ,  $s_x = 0.08294$ ,  $\hat{k}_3 = 0.046864$ ,  $\hat{k}_4 = 2,199797$  olarak hesaplanmıştır. Minitab 17’de yapılan Anderson Darling uyum iyiliği sonuçlarına göre veri setinin Normal dağılıma uymadığı görülmüştür (AD=1.379, p-değeri <0,005). Dolayısıyla bu verilere klasik Shewhart kontrol kartlarını çizmek uygun değildir. Verilerin histogramı Şekil 3 ile verilmiştir.



Şekil 3. Ayar Plakası Verilerinin Histogramı

Elde edilen istatistiklere ve çizdirilen histogram grafiğine göre dağılımın simetrik ancak Normal dağılıma göre daha basık bir dağılım olduğu görülmüş ve basıklık düzeltmeli kontrol grafiği çizilmiştir.



Şekil 4. Basıklık Düzeltmeli Kontrol Grafiği

Süreç dağılımının ve parametrelerin bilinmediği durum için çizilen basıklık düzeltmeli kontrol grafiği için verilen formüller ve  $\hat{k}_4 = 2,199797$  için Tablo 1 den okunan  $A_2 = 0.657$  değeri kullanılarak

$$\dot{U}KL = 15.12392$$

$$O\dot{C} = 15.01696$$

$$AKL = 14.91$$

biçiminde hesaplanmıştır.

### Sonuçlar ve Öneriler

Bir üretim sürecinde normal dağılım varsayımının şüpheli olduğu durumda, çarpık ya da ağır kuyruklu dağılım söz konusu olduğunda klasik Shewhart kontrol kartı yerine çarpıklık ya da basıklık düzeltmeli kontrol kartlarının kullanması önerilir. Bir imalat hattında sürecin istatistiksel kontrolü için yapılan denetlemeler ile verimliliği artırmak, yüksek kar elde etmek ve zaman tasarrufu elde etmek açısından doğru kontrol kartının seçilmesi çok önemlidir.

Normal dağılımını varsayımını sağlamayan, simetrik ama normal dağılıma göre daha sivri bir dağılıma sahip olan bu verilerimiz için basıklık düzeltmeli kontrol grafikleri çizildiğinde süreç kontrolü daha başarılı sonuçlar vermiştir.

### Kaynaklar

- Bai, D.S., Choi, I.S., 1995. “ $\bar{X}$  and R control charts for skewed population”, Journal of Quality Tecnology, 27(2), 305-701.
- Chan, L.K., Cui, H.J., 2003. “Skewness correction  $\bar{X}$  and R charts for skewed distributions”, Naval Research Logistics, 50, 1-19.
- Chang, Y.S., Bai, D.S., 2001. “Control charts for positively- skewed populations with weighted wstandard deviations”, Quality and Reliabilite Engineering İnternational, 17(5), 397-406, 2001.
- Cornish, E.A., Fisher, R.A., 1938. “ Moments and cumulants in the specification of distributions”, Revue de l'Institut international de Statistique, 307-320, 1938.
- Montgomery, 1995. Introduction to Statistical Quality Control, New York, John Wiley & Sons.
- Öncel, S.Y., Özarslan H., 2020. “Çarpık Dağılımlar için Çarpıklık Düzeltmesi Yöntemine Dayalı  $\bar{X}$  ve R Kontrol Grafikleri”, Akademik Platform Mühendislik ve Fen Bilimleri Dergisi, basım aşamasında.
- Özarslan, H., 2018. “Çarpık Dağılımlar için Shewhart Kontrol Grafikleri”, Yüksek Lisans Tezi, Kırıkkale Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Kırıkkale, 2018.
- Shewhart. W.A., 1926 "Quality control charts." The Bell System Technical Journal vol.5, no 4, pp. 593-603.
- Tadikamalla, P.R., Popescu, D.G., 2007. “Kurtosis correction metod for  $\bar{X}$  and R control charts for long-tailed symmetrical distributions”, Naval Research Logistics, 54(4), 371-383, 2007.
- Turhan, N., Öncel, S.Y., 2019. “Süreç ortalamasının izlenmesi için sıra istatistiklerine dayalı kalite kontrol kartları”, İstatistikçiler Dergisi: İstatistik ve Aktüerya, 12(2), 72-89.

## Diagnostic Accuracy of PCR Test and Symptoms of the Disease for Covid-19 by Multivariate Roc Analysis

Aycan Mutlu YAĞANOĞLU<sup>1</sup>, Mehmet TOPAL<sup>2</sup>

<sup>1</sup>Ataturk University, Faculty of Agriculture, Department of Animal Science, Erzurum, Turkey

<sup>2</sup>Kastamonu University, Faculty of Medicine, Department of Biostatistics Kastamonu, Turkey

### Abstract

The new coronavirus disease (Covid-19) is a new viral respiratory tract disease that was first identified with high fever and shortness of breath in Wuhan, China, on January 13, 2020. It is known that the disease is transmitted by droplets and contact. It is defined as a pandemic due to the global epidemic. The new coronavirus is a type of coronavirus family that can affect both animals and humans. In retrospect, some different viruses from the coronavirus family seem to cause severe respiratory diseases such as MERS (Middle East Respiratory Syndrome) and SARS (Severe Acute Respiratory Syndrome). A total of 160 individuals, 89 men and 71 women, who applied to the hospital with suspicion of Covid-19, were asked about the possible complaints of the patients in line with the results of tomography and PCR, and the answers received were recorded. As a result of the tomography test, which is the gold standard test, a total of 14 cases were positive and 146 cases were negative. Considering the areas under the curve as a result of ROC analysis, PCR = 0.871, Loss of Taste or Smell = 0.813, Shortness of Breath or Difficulty Breathing = 0.762, Fever = 0.712, Cough = 0.613 and Sore Throat = 0.598 values were obtained respectively.

**Key words:** Covid-19, diagnostic, accuracy, ROC analysis

## Çok değişkenli ROC Analizi ile Covid-19 için PCR Testi ve Diğer Semptomların Tanısal Doğruluğu

### Özet

Yeni koronavirüs hastalığı (Covid-19), ilk olarak 13 Ocak 2020'de Çin'in Wuhan kentinde yüksek ateş ve nefes darlığı ile tanımlanan yeni bir viral solunum yolu hastalığıdır. Hastalığın damlacıklar ve temas yoluyla bulaştığı bilinmektedir. Küresel salgın nedeniyle pandemi olarak tanımlanmaktadır. Yeni koronavirüs, hem hayvanları hem de insanları etkileyebilen bir tür koronavirüs ailesidir. Geçmişte bakıldığında, koronavirüs ailesinden bazı farklı virüslerin, MERS (Orta Doğu Solunum Sendromu) ve SARS (Şiddetli Akut Solunum Sendromu) gibi ciddi solunum yolu hastalıklarına neden olduğu görülmektedir. Çalışmamızda, hastaneye Covid-19 şüphesiyle başvuran 89 erkek ve 71 kadın olmak üzere toplam 160 kişiye tomografi ve PCR sonuçları doğrultusunda hastaların olası şikayetleri sorulmuş ve alınan cevaplar kayıt altına alınmıştır. Altın standart test olan tomografi testi sonucunda toplam 14 vaka pozitif, 146 vaka negatif çıktı. ROC analizi sonucunda eğri altında kalan alanlar göz önüne alındığında PCR = 0.871, Tat veya Koku Kaybı = 0.813, Nefes Darlığı veya Nefes Almada Zorluk = 0.762, Ateş = 0.712, Öksürük = 0.613 ve Boğaz Ağrısı = 0.598 değerleri sırasıyla elde edilmiştir.

**Anahtar kelimeler:** Covid-19, teşhis, doğruluk, ROC analizi

## Karar Ağaçları Yöntemi ile Operatör Değişikliği Nedenlerinin Araştırılması

Merve ÇATAL<sup>1</sup>, Sevgi YURT ÖNCEL<sup>1</sup>

<sup>1</sup>İstatistik Bölümü, Fen-Edebiyat Fakültesi, Kırıkkale Üniversitesi, 71450 Yahşihan, Kırıkkale, Türkiye

### Özet

Bilişim teknolojisi çok büyük miktarlardaki veriyi saklamayı mümkün kılmaktadır. Veri madenciliği, büyük veri yapılarından değerli bilgilerin çıkarılma sürecidir. Teknolojinin gelişmesiyle beraber bankacılık, alışveriş, sosyal iletişim ve haberleşme gibi günlük pek çok işler dijital platformlarda bilgisayar, akıllı cihaz gibi araçlar yardımıyla gerçekleştirilmektedir. Bu işlemler dijital ortamlarda gerçekleştirildikçe artan veri yığınlarından önemli bilgilerin elde edilmesi veri madenciliğinin gelişmesine neden olmuştur. Karar ağaçları ise, veriler arasında gizli kalmış ilişkileri, örüntüleri ortaya çıkarabilmek amacıyla kullanılan bir yöntemdir. Anlaşılır ve kolay yorumlanabilir olması, veri tabanı sistemleri ile kolayca entegre edilebilmeleri ve güvenilirliklerinin iyi olması nedenleri ile veri madenciliği yöntemleri içerisinde en yaygın kullanılan yöntemlerden biridir. Bu çalışmada, karar ağaçları yöntemleri ile kullanıcıların operatör değişikliğine neden olan etkenler ele alınmış, kullanıcıların hangi durumlarda operatör değişikliğine gittikleri üzerine anket yapılmıştır. C4.5, CART ve CHAID en popüler algoritmalar. C5.0, CART ve CHAID algoritmaları incelenmiş ve C4.5 algoritması ile veriler analiz edilip karar ağacı oluşturulmuştur.

**Anahtar kelimeler:** C4.5 algoritması, GSM operatör değişikliği, karar ağaçları, veri analizi, veri madenciliği

## Investigation of Reasons for Operator Change with Decision Trees Method

### Abstract

Information technology is productively scooping massive amounts of data. Data mining is the process of making talk about big data sites. With the development of technology, many daily businesses such as banking, shopping, social communication and communication will be realized on digital platforms, such as computers and smart devices. As these operations are performed in digital environments, learning important from the increasing data mass has led to the development of data mining. Decision Aging is a method used to reveal hidden relationships and patterns among data. It is one of the most widely used methods of being understandable and easy to interpret, being easily integrated with database systems and testing their reliability well. On this page, the factors that cause response operator change with decision transformation are questioned as they change the operator in style. C4.5, CART and CHAID are the most popular algorithms. C5.0, CART and CHAID algorithms were examined and the data was analyzed with C4.5 algorithm and decision tree was formed.

**Key words:** C4.5 algorithm, decision trees, data analysis, data mining, GSM operator change

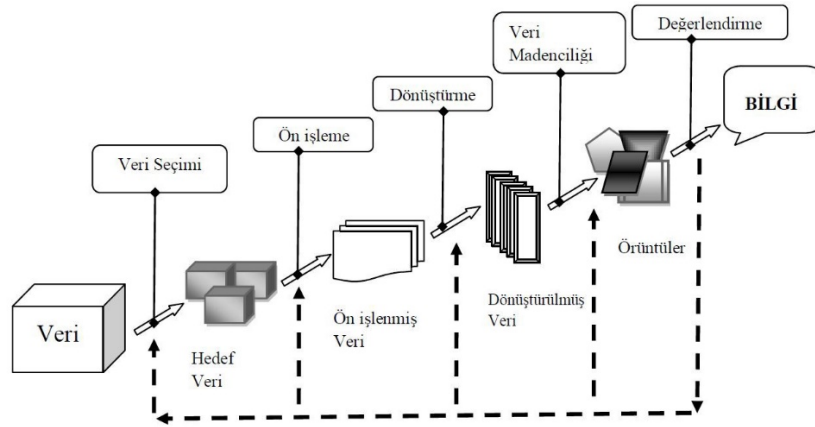
### Giriş

Gelişen ve gelişmeye devam eden teknoloji vasıtasıyla dijital platformlar her geçen gün daha fazla etkili olmaya başlamıştır. İnsanlar dijital platformlar sayesinde akıllı telefonları ile birçok günlük işlerini,

Borsa, sigortacılık, bankacılık, telekomünikasyon, pazarlama, perakendecilik, mühendislik, biyoloji, istihbarat, tıp, genetik, kriminoloji, endüstri vb. pek çok dalda, zaman ve mekân kavramı olmadan halledilmektedirler. Hastanelerde, belediyelerde, marketlerde hemen hemen artık her yerde her işlem veri tabanında yer alır. Gelişen teknoloji ile akıllı telefonların, bilgisayarların, tabletlerin işlemcilerinin ve kapasitelerinin artması bu araçlarda daha fazla ve daha hızlı veri saklanabilmesine olarak sağlamaktadır. Tüm bu saklanan veriler depolandıkları yerde bilgi olabilmek için beklemektedirler.

Veri madenciliğinin temeli, ilk sayısal bilgisayar olan “ENIAC’a (Electrical Numerical Integrator And Calculator)” dayanmaktadır. 1946 yılında geliştirilmiş olan ve günümüzde kullandığımız bilgisayarların atası sayılabilen ENIAC, ABD’li bilim insanları J. Presper Eckert ile John Mauchly tarafından, II. Dünya Savaşı’nda ABD ordusunun kullanması için geliştirilmiştir. Bu “ilk” bilgisayar, 30 ton ağırlığında, 170 m<sup>2</sup>’lik alanı kaplayan bir cihazdır (Öğüt, 2014).

Veri madenciliği, veri içerisindeki yeni, gizli kalmış veya beklenmeyen örüntüleri bulmak için analizler yapar. Büyük bir süreç olarak adlandırılan bilgi keşfi sürecinin bir bölümüdür. Şekil 1’de veri madenciliği sürecine yer verilmiştir.



Şekil 1. Veri Madenciliği Süreci

“Uydu Verilerinden Karar Ağaçları Kullanarak Orman Yangını Tahmini” adlı çalışmada uzmanlar uydulardan alınan veriler ile orman yangınlarının tahmini yapmışlardır. Uydudan alınan verilerden hesaplanan Normalize Edilmiş Fark Bitki Örtüsü İndeksi (NVDI), Arazi Yüzeyi Sıcaklığı (LST) ve Termal anomali (TA) verileri kullanılarak orman yangınları tahmin edilmiştir. Bahsedilen verilerden tahmin yapmak için karar ağaçları kullanılmışlardır. Gerçekleştirilen denemelerde ortalama %98,62 duyarlılık oranı ile gerçekleşen yangınlar doğru tahmin edilmiştir. Tüm denemelerde yapılan tahminler için ortalama %93,11 doğruluk ile gerçek durum belirlemişlerdir. (Nurettin ve Mehmet Emin, 2020)

“İnsani Gelişmişlik Endeksinin Karar Ağacı Algoritmaları ile Modellenmesi: BM’de Bir Uygulama 2010-2017 Dönemi” adlı çalışmada UNDP’nin 2010-2017 yıllarını kapsayan 79 ülkenin verileri kullanılarak veri madenciliğinin karar ağacı tekniklerinden C5.0 ve Gini algoritmaları ile karar ağaçları oluşturulmuştur. Karar ağaçları ile birlikte İGE’ye etki eden faktörler belirlenmiş ve ülkeler çok yüksek, yüksek, orta ve düşük düzey gelişmiş ülkeler olarak sınıflandırılarak kurallar elde edilmiştir. Yapılan analizler sonucunda C5.0 algoritması ile %97,94 ve Gini algoritması ile %91,93’lük doğru sınıflandırma başarısı elde edilmiştir. Bunun dışında duyarlılık ve belirleyicilik istatistikleri de hesaplanmıştır. İGE’ye en fazla etki eden değişkenlerin eğitim, istihdam ve sağlık göstergelerindeki değişkenler olduğu tespit edilmiştir. (Emre ve Ali, 2020)

“Veri Seti-Sınıflandırma İlişkisinde Performansa Etki Eden Faktörlerin Değerlendirilmesi” adlı çalışmada 32 farklı veri seti kullanılarak deneyler gerçekleştirilmiştir. 9 farklı sınıflandırma algoritması da kullanılarak bu algoritmaların veri setleri üzerindeki başarımları da kıyaslanmıştır. Elde edilen sonuçlar, algoritmaların etkinliğinin yanı sıra test yöntemlerinin başarıma etkilerini de ortaya çıkarmıştır. Veri setine ait parametrelerin sınıflandırma başarımına etkileri makalede detaylı olarak ortaya konulmuştur. (Abdullah ve Murat, 2020)

“Müşteri Kayıp Analizi İçin Sınıflandırma Algoritmalarının Karşılaştırılması” adlı çalışmada uzmanlar kaggle.com üzerinden yayımlanan toplam iki veri seti ile çalışmışlardır. İlk veri seti bir banka müşterilerine, ikinci veri seti ise telefon operatörü müşterilerine aittir. Kullanılan iki veri seti için de çeşitli nitelikler incelenerek müşteri kayıp analizi yapmışlardır. Çalışmada Logistic Regression, Naive Bayes, Decision Tree, K-NN, SVM ve LDA sınıflandırma modelleri kullanılmıştır. İncelemede toplam 13.000 müşteri bilgisi üzerinden belirlenen yöntemler ile çalışma yapılmıştır. Algoritmanın uygulanması kolay ve literatür çalışması fazla olduğu Python Programlama Dili’ni seçmişlerdir. Kullanılan sınıflandırma yöntemleri her iki veri seti için de uygulanmış ve doğruluk oranları karşılaştırılmışlardır. En kararlı sonucu karar ağaçları algoritması olduğu gözlemlenmiştir. (Alper ve Yelda, 2020)

“Kredi Kartı Kullanan Müşterilerin sosyo-ekonomik Özelliklerinin Kümeleme Analiziyle İncelenmesi” adlı çalışmada, uzmanlar kredi kartı kullanan müşterilerin sosyo-ekonomik özelliklerinin gruplanması amaçlamışlardır. Çalışmada, öncelikle bireysel bankacılık ve onun bir işlevi olan kredi kartlarının tanımlanmasına, bu kavramların ülkemizdeki yeri ve öneminin belirlenmesine yer verilerek, kredi kartı kullanan banka müşterileri kümeleme analiziyle gruplandırılmıştır. Uygulamada, verilere en uygun teknik olduğu için kümeleme analizinin hiyerarşik olan yöntemlerinden ortalamalar bağlantı tekniği tercih edilmiştir. Bu yöntemle ilgili banka müşterileri sosyo-ekonomik özelliklerine göre üç kümede gruplanmışlardır. İlk kümede en yoğun müşteri topluluğu bulunurken, İkinci kümede daha az müşteri topluluğu yer almış, üçüncü kümede ise azınlıkta olan müşteri grubu yer almıştır. Bu üç kümeye göre, müşterilerin on adet sosyo-ekonomik değişkene göre farklılık gösterdiğini gözlemlenmiştir. (Aşan, Zerrin. 2015)

Bu çalışmanın önemi; günümüzde kendisi oldukça önemli ve de ihtiyaç duyulan bir alan haline gelen veri madenciliğinin sınıflandırma yöntemlerinden olan karar ağaçlarının ele alınmasıdır. Uygulamada, IBM SPSS Modeler 18.0 programı kullanılmıştır. Uygulanan ankete ait veriler Excel’de düzenlenmiştir. Teknolojinin oldukça yaygınlaştığı, telefon kullanımının oldukça fazla olduğu bu dönemde operatör kullanıcılarının beklentilerinin neler olduğunu ve kullanıcıların hangi durumlarda operatör değişikliği yaptıklarının belirlenmesi amacı ile Google Form üzerinden 32 soru içeren anket hazırlanmıştır. Araştırma 2020 yılının Ocak-Ağustos aylarında yapılmıştır ve 765 operatör kullanıcısı ankete katılmıştır. Anket verileri mail yoluyla ve kişilerle birebir görüşülerek elde edilmiştir. Kişilere anket soruları yöneltilmeden önce etik kurul onayı alınmıştır.

## **Materyal ve Yöntem**

Veri madenciliği modelleri üç temel başlıkta incelenebilir. Sınıflandırma ve regresyon, kümeleme, birliktelik kuralları. (Akpınar, 2000). Karar ağaçları bir sınıflandırma modelidir.

Karar Ağaçları,

- Düşük maliyet,
- Anlaşılmasının, yorumlanmasının ve veri tabanları ile entegrasyonun kolaylığı,
- Güvenilirliklerinin iyi olması,

- Hem kategorik hem de sayısal verilerle kullanılabilmesi nedenleri ile yaygın şekilde kullanılan veri madenciliği yöntemidir.

ID3 (Iterative Dichotomiser 3), C4.5 & C5.0 (ID3 karar ağacı algoritmasının ileri versiyonları), CART (Classification And Regression Trees), CHAID (Chi-Squared Automatic Interaction Detector), QUEST, SLIQ (Supervised Learning in Quest), SPRINT (Scalable Parallelizable Induction of Decision Tree) algoritmaları en yaygın kullanılan karar ağacı yöntemleridir.

ID3 Algoritması, 1986 yılında J.R. Quinlan tarafından kategorik verilerin sınıflandırılması amacıyla geliştirilmiş, en temel sınıflama yöntemlerinden biridir. ID3 Algoritması kendisinden sonra bulunan sınıflandırma algoritmalarının da temelini oluşturduğu için oldukça önemlidir. ID3 algoritmasında genel entropi değeri hesaplanır.

$$H(y) = H(p_1, p_2, \dots, p_k) = \sum_{j=1}^k (p_j \log_b(1/p_j))$$

Entropi, belirsizliği ve beklenmeyen durumun ortaya çıkma olasılığını göstermektedir. 0 ile 1 arasında değerler almaktadır. Eşit olasılıklı durumlar belirsizliğin yüksek olduğunu göstermektedir. Eğer ihtimal eşitse entropi 1 değerini alır.

Entropi hesaplandıktan sonra her bir değişken için kazanım değeri hesaplanır.

$$Gain(x_i) = H(Y) - \sum_{j=1}^k P(X_{ij})H(Y/X_{ij}), \quad i = 1, 2, \dots, m$$

Kazanım değeri en büyük olan değişken en belirleyici özellik olarak seçilir. Belirlenen ilk değişken ise kök düğüm olarak alınır. Her bir düğüm saf (yaprak) olana kadar 1. aşamaya geri dönülerek ağacın dalları tamamlanır.

ID3 algoritmasında kategori sayısı arttıkça hata oranı artmaktadır. Özellikle büyük veri setlerini sınıflandırabilmek amacıyla J.R. Quinlan 1993 yılında C4.5 algoritmasını geliştirmiştir. C5.0 algoritması ise C4.5 ile aynı sonuçları verse de görünüş olarak daha iyi sonuç vermektedir ve daha hızlıdır. C4.5 algoritması dallandırma yaparken kazanım değil kazanım oranı (gain ratio) kullanır.

$$GainRatio(X_i) = \frac{Gain(X_i)}{H(X_i)}, \quad i = 1, 2, \dots, m$$

Cart Algoritması, 1984 yılında Breiman, Friedman, Olshen ve Stone tarafından geliştirilmiştir. Hem sayısal hem de kategorik veri türlerini sınıflandırmada kullanılmaktadır. Sınıflar kategorik verilerden oluşuyor ise gini indeksi, sayısal verilerden oluşuyor ise regression trees kullanılarak ağaç oluşturulur. Gini indeks değeri en küçük olan özellik en fazla belirleyiciliğe sahip özellik olarak alınır ve dallandırma en küçük gini indeksi ile yapılır.

$$Gini(X_{ij}) = 1 - \sum_{m=1}^k [P(Y_m/X_{ij})]^2$$

Chaid Algoritması, 1980 yılında Kass tarafından geliştirilmiştir. Kategorik değişkenlerde kullanılır. Sınıflandırma yaparken Ki-Kare istatistiğinden faydalanmaktadır. Quest Algoritması, 1997 yılında Loh and Shih tarafından geliştirilmiştir. Quest algoritması, ağacın oluşturulması sırasında, Chaid ve Cart'ın aksine tüm değişkenler ile ayrı ayrı ilgilenir. Bu algoritma hesaplama maliyetinin düşürülmesi amacıyla geliştirilmiştir.

## Uygulama

Bu makalede, 765 operatör kullanıcıya operatör değişikliği ile ilgili sorular yöneltilmiş olup, elde edilen verilerin karar ağacı yöntemlerinden C4.5 algoritması ile değerlendirilmesi sağlanmıştır. Kullanıcıları kullandıkları operatör ve memnuniyetleri hakkında sorular sorulmuş ve bu sorular kategorize edilerek sistemdeki karşılıkları Tablo 1’de verilmiştir.

Tablo 1. Operatör Kullanıcılarına Sorulan Sorular ve Demografik Verileri

Karakteristiğın				%
S. No	İsmi	Açıklaması	Aldığı Değerler	
1	Soru 1	Kişilerin cinsiyet bilgisi	1 Kadın	400 (%52,3)
			2 Erkek	365 (%47,7)
2	Soru 2	Kişilerin yaş bilgisi	1 2-14	0
			2 15-19	112 (%14,6)
			3 20-24	219 (%28,6)
			4 25-39	336 (%43,9)
			5 40-59	91 (%11,9)
			6 60 yaş üzeri	7 (%0,9)
3	Soru 3	Kişilerin medeni durum bilgisi	1 Bekar	438 (%57,3)
			2 Nişanlı	38 (%5)
			3 Evli	286 (%37,4)
			4 Boşanmış	3 (%0,4)
4	Soru 4	Kişilerin eğitim durumu bilgisi	1 İlkokul	10 (%1,3)
			2 Ortaokul	23 (%3)
			3 Lise	194 (%25,4)
			4 Önlisans	78 (%10,2)
			5 Lisans	391 (%51,1)
			6 Yüksek Lisans	63 (%8,2)
			7 Doktora	6 (%8)
5	Soru 5	Kişilerin aylık gelir bilgisi	1 Herhangi bir gelirim yok	172 (%22,5)
			2 Öğrenciyim. Öğrenim kredisi/bursu alıyorum.	139 (%18,2)
			3 Asgari ücret	106 (%13,9)
			4 2.500-4.000	118 (%15,4)
			5 4.000-7.000	194 (%25,4)
			6 7.000 ve üzeri	36 (%4,7)
6	Soru 6	Kişilerin çalıştıkları sektör bilgisi	1 Çalışmıyorum	85 (%11,1)
			2 Öğrenci	234 (%30,6)
			3 Kamu kurumu	261 (%34,1)
			4 Özel kurum	140 (%18,3)
			5 Kendi işim	15 (%2)
			6 Serbest meslek	22 (%2,9)
			7 Emekli	8 (%1)
			8 Malulen emekli	591 (%77,3)
7	Soru 7	Kişilerin yaşadıkları idari bölge bilgisi	1 İl	156 (%20,4)
			2 İlçe	18 (%2,4)
			3 Köy	1 (%1)
8	Soru 8	Kullanmakta olduğunuz telefon akıllı ise işletim sistemi nedir?	1 Akıllı Telefon Kullanmıyorum	576 (%75,3)
			2 Android	182 (23,8)
			3 İOS	6 (%8)
			4 Windows	563 (%73,6)
9	Soru 9	Operatör Değişikli Yaptınız mı?	1 Evet	202 (%26,4)
			2 Hayır	337 (%44,1)
10	Soru 10	Kullandığınız ilk operatör hangisi?	1 Turkcell	179 (%23,4)
			2 Türk Telekom	249 (%32,5)
			3 Vodafone	166 (%21,7)



11	Soru 11	İlk kullandığınız operatörden sonra hangisine geçiş yaptınız?	1 Turkcell	166 (%21,7)
			2 Türk Telekom	340 (%44,4)
			3 Vodafone	259 (%33,9)
12	Soru 12	Sosyal medya uygulamalarını kullanırken operatörünüzün hizmet kalitesi ihtiyacınızı karşılıyor mu?	1 Evet	414 (%54,1)
			2 Hayır	84 (%11)
13	Soru 13	Herhangi bir sıkıntı yaşadığınızda müşteri hizmetleri ile işinize yarayan süre zarfında iletişime geçebiliyor musunuz?	1 Evet	267 (%34,9)
			2 Hayır	379 (%49,5)
14	Soru 14	Müşteri hizmetleri sorununuzu çözmeye yeterli olabiliyor mu?	1 Evet	122 (%15,9)
			2 Hayır	264 (%34,5)
15	Soru 15	Yaşadığınız idari bölgede kullandığınız operatöre ait operatör işlem merkezi mevcut mu?	1 Evet	351 (%45,9)
			2 Hayır	159 (%20,8)
16	Soru 16	Olabilecek bir doğal afet anında kullanmakta olduğunuz operatörün çekim gücünün yeterli gelebileceğini düşünüyor musunuz? Daha önce böyle bir durumla karşılaştıysanız yeterli oldu mu?	1 Daha önce böyle bir durumla karşılaşmadım. Karşılaşırsam yeterli olacağını düşünüyorum.	256 (%33,5)
			2 Daha önce böyle bir durumla karşılaşmadım. Karşılaşırsam yeterli olacağını düşünmüyorum.	299 (%39,1)
			3 Daha önce böyle bir durumla karşılaştım. Yeterli oldu.	31 (%4,1)
			4 Daha önce böyle bir durumla karşılaştım. Yeterli olmadı.	179 (%23,4)
17	Soru 17	Gelişen altyapı teknolojisi operatör seçiminizde ve operatör değiştirmenizde etkili oldu mu?	1 Evet	474 (%62)
			2 Hayır	291 (%38)
			3 Yeterli bilgiye sahip değilim.	455 (%59,5)
18	Soru 18	Kullandığınız operatörün çekim gücünden memnun musunuz?	1 Evet	310 (%40,5)
			2 Hayır	262 (%34,2)
19	Soru 19	Faturalı hat kullanıyorsanız, ödediğiniz miktar belirlediğiniz tutarı aşıyor mu?	1 Evet	319 (%41,7)
			2 Hayır	184 (%24,1)
			3 Faturalı hat kullanmıyorum	3 (%4)
20	Soru 20	Hat değişikliği yapma sebebiniz nedir?	1 Konuşma paketi yetersiz	1 (%1)
			2 Mesajlaşma paketi yetersiz	297 (%38,8)
			3 İnternet paketi yetersiz	194 (%25,4)
			4 Çekim gücü yetersiz	216 (%38,2)
			5 Paket fiyatları oldukça yüksek	50 (%6,5)
			6 Faturam anlaştığım tarife fiyatını aşıyor	4 (%5)
21	Soru 21	Yaşadığınız idari bölgede kullandığınız operatöre ait operatör işlem merkezi mevcut mu?	1 Evet	606 (%79,2)
			2 Hayır	159 (%20,8)
22	Soru 22	Önceki operatörünüzü kullanma süreniz	1 0-1 Yıl	203 (%26,5)
			2 1-3 Yıl	254 (%33,2)
			3 3-5 Yıl	136 (%17,8)
			4 5 yıldan daha fazla	172 (%22,5)

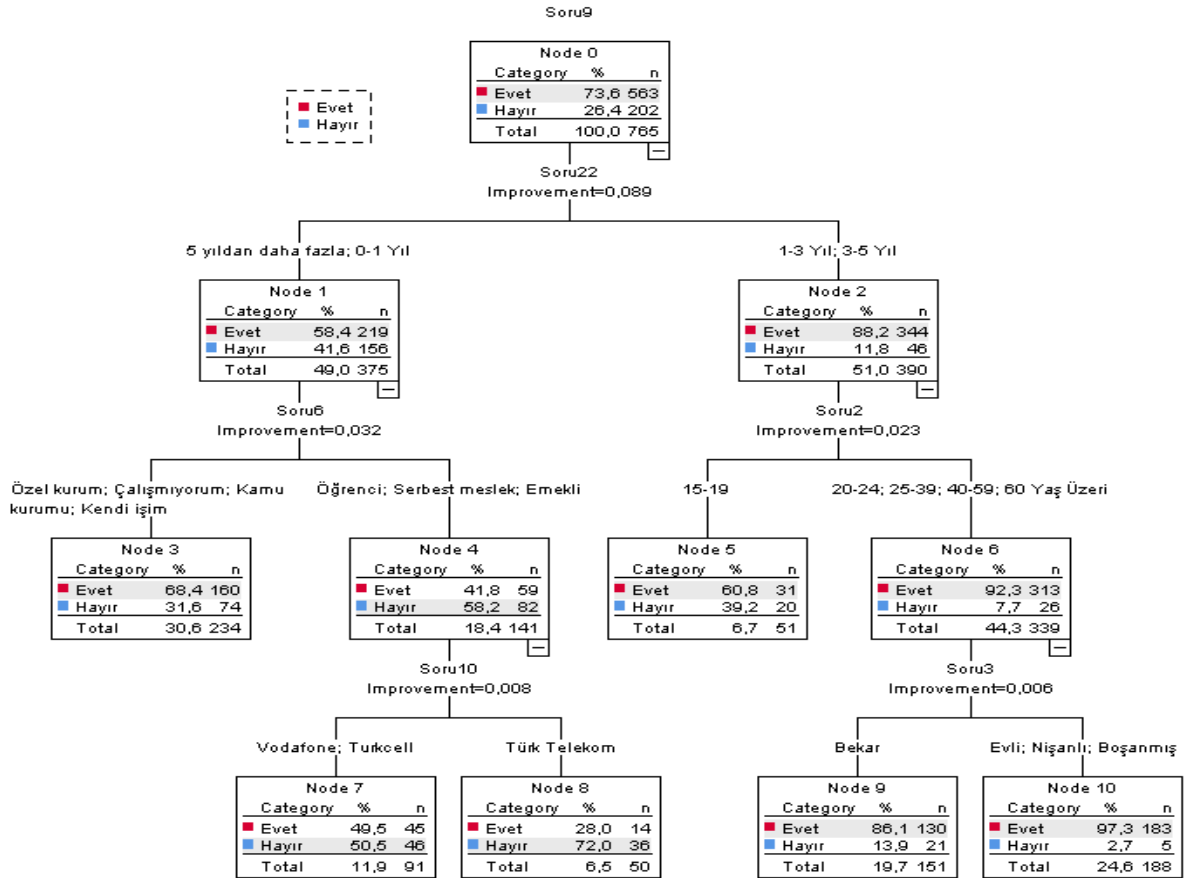
Karar ağacı uygulanmadan önce kişilerin operatör değişikliği yapıp yapmadığı sorusuna verdiği evet ve hayır yanıtları ile diğer değişkenler arasında anlamlı bir farklılık bulunup bulunmadığına dair Ki-Kare testleri yapılmıştır ve anlamlı ilişki bulunan değişkenler Tablo 4’te gösterilmiştir.

Tablo 2. Ki-Kare Testi Sonuçları

Karşılaştırılan Testler	Exact sig. (p value)
Soru9*Soru1	0.037
Soru9*Soru2	0.00
Soru9*Soru3	0.00
Soru9*Soru4	0.02
Soru9*Soru5	0.018
Soru9*Soru6	0.00
Soru9*Soru7	0.01
Soru9*Soru8	0.00
Soru9*Soru10	0.00

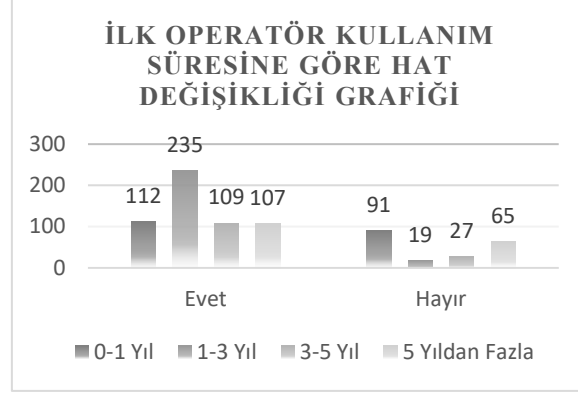
### Soru9\* Operatör Değişikliği yaptınız mı?

Ki-Kare testi sonuçlarına bakıldığında cinsiyet, yaş bilgisi, medeni durum, eğitim durumu, aylık gelir bilgisi, çalışılan sektör, yaşanan idari bölge, kullanılan işletim sistemi ve tercih ettikleri ilk operatör, operatör değişikliği yapmalarında etkili olan değişkenlerdir.



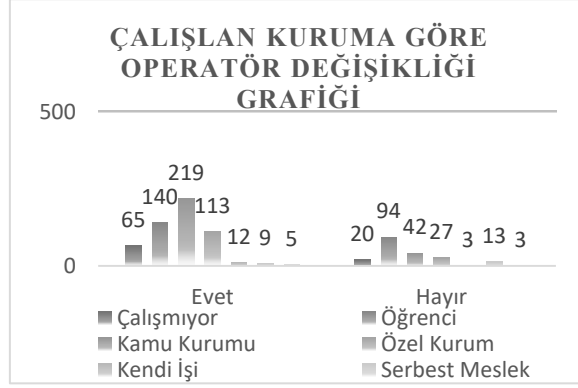
Şekil 2. Karar Ağacı

Karar ağacında bulunan değişkenleri grafiklerle de gösterelim. Öncelikle kök düğüm olan önceki operatörü kullanma süresine bakıyoruz.



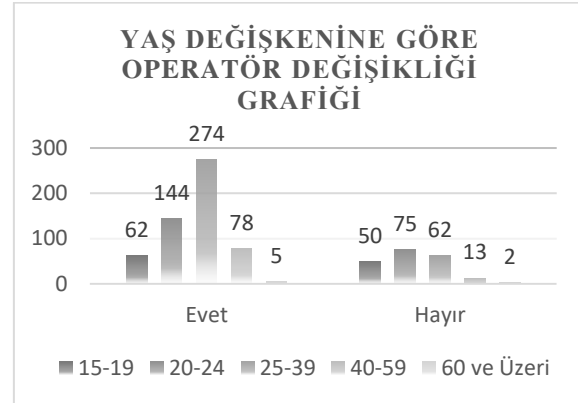
Şekil 3. İlk Operatörün Kullanım Süresi Grafiği

Grafiğe bakıldığında operatörünü 1-3 yıl arasında kullanan kullanıcılar daha fazla hat değişikliği yapmıştır.



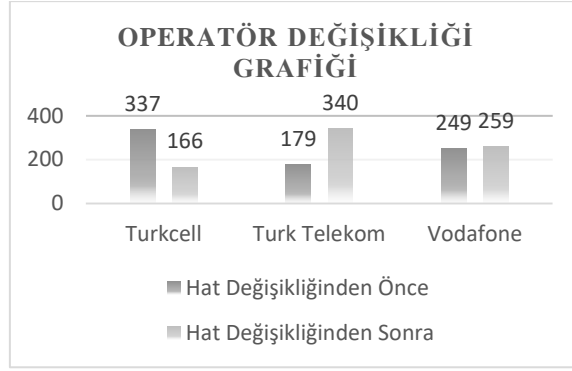
Şekil 4. Çalışılan Kurum Bilgileri için Operatör Değişikliği

Grafiğe bakıldığında kamu kuruma çalışanları diğer sektör çalışanlarına göre daha fazla hat değişikliği yapmıştır.



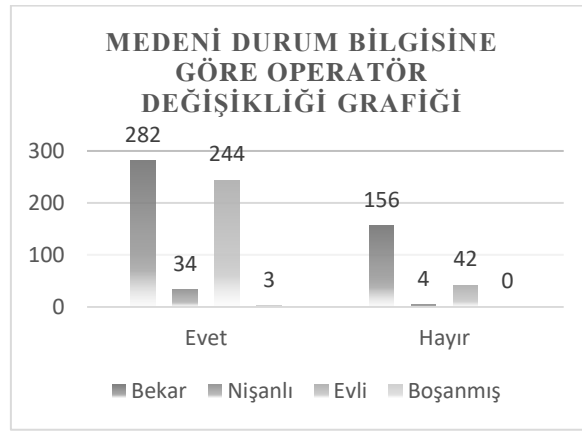
Şekil 5. Yaş için Operatör Değişikliği

Grafiğe bakıldığında 25-39 yaş aralığındaki operatör kullanıcıları diğer kullanıcılara göre daha fazla hat değişikliği yapmıştır.



Şekil 6. Operatör Değişikliği Grafiği

Grafiğe bakıldığında Turkcell kullanıcıları diğer operatör kullanıcılarına göre daha fazla oranda hat değişikliği yapmıştır.



Şekil 7. Medeni durum için operatör değişikliği

## Sonuç ve Tartışmalar

Karar ağacı CART (Classification And Regression Trees) algoritması kullanılarak oluşturulmuştur. Kategorik verilerden oluşan değişkenler gini indeksi kullanılarak hesaplanmış olup, gini indeksi en küçük değişken olan 'operatörün kullanım süresi' en belirleyici özellik seçilmiş ve kök düğüm olarak kabul edilmiştir. Şekil 2'ye bakıldığında, operatörünü 5 yıldan fazla ve 0-1 yıl aralığında kullanan kullanıcıların 1 ile 5 yıl arasında kullananlara göre daha düşük oranda operatör değişikliği yaptığı görülmüştür. Operatörler hat operatör değişikliğini yeni yapan ve uzun süreli kullanan kullanıcılarına özel tarifeler uygulamaktadır. Bu duruma meslek durumu da neden olmaktadır. Özel kurumda çalışan, kamu kurumunda çalışan ve çalışmayan bireylerin hat değişikliğine daha fazla eğilimli olduğu görülmektedir. Serbest meslekle uğraşan emekli ve öğrenci olanlarda ise hat geçişi kullanılan operatöre göre değişmektedir. Vodafone ve Turkcell kullanıcıları %49,5 oranında operatör değişikliği yaparken Türk Telekom kullanıcıları daha düşük bir oranla %28 olarak operatör değişikliğine gitmektedirler. Operatörünü 1-5 yıl arasında kullanan kullanıcılara bakıldığında ise %88 oranında operatör değişikliği yapıldığı görülmektedir. Bu duruma neden olan bir başka değişken incelendiğinde yaş faktörünün etkili olduğu görülmüştür. 15-19 yaş arası kullanıcılar %60,8 oranında iken bu oran 20 yaş ve üzeri kişilerde %92,8 olmaktadır. 20 yaş üzeri kişilerde etkili olan bir diğer faktör de kişilerin medeni durum olarak saptanmıştır. Bekâr bireyler %86,1 oranında operatör değişikliği yaparken, evli nişanlı ve boşanmış kişiler %97,3 oranıyla bekâr kişilere göre operatör değişikliğine daha eğilimlidir.

Grafığe bakıldığında bekâr ve evli operatör kullanıcıları diğer kullanıcılara göre daha fazla hat değışikliğı yapmıştır.

İnternetin bir hayli fazla kullanıldığı bu dönemde kullanıcılara yöneltilen sorulara verilen cevaplardan yola çıkarak, kullanıcıların internet paketleri ile operatörleri kullanan kişilerin büyük bir kısmının çekim gücünden şikâyetçi olduğu gözlenmiştir. Kullanıcıların %39,1'i olabilecek bir doğal afet durumunda kullandıkları operatörün yeterli gelmeyeceğini düşünmektedir. Kişilere hat değışikliğı yapma nedenleri sorulduğunda ise kişilerin %38,2'si paket fiyatlarının çok pahalı olduğunu belirtmişlerdir.

### Kaynaklar

- Akpınar, H. (2000). "Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliğı," İstanbul Üniversitesi İşletme Fakültesi Dergisi, cilt 29, s. 1-22
- Albayrak, AS. (2009). "Türkiye'de Yerli ve Yabancı Ticaret Bankalarının Finansal Etkinliğe Göre Sınıflandırılması: Karar Ağacı, Lojistik Regresyon ve Diskriminant Analizi Modellerinin Bir Karşılaştırılması," Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, sayı 14, s. 113-139.
- Albayrak AS. (2009). Koltan Yılmaz Ş. Veri madenciliğı: Karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. 14(1): 31-52.
- Aslı, Sema Kayapınar, Tahsin Çetinyokuş. (2014) "Veri Madenciliğinde Karar Ağacı Algoritmaları İle Bilgisayar Ve İnternet Güvenliğı Üzerine Bir Uygulama." Journal Of Industrial Engineering (Turkish Chamber Of Mechanical Engineers) 25
- Altunkaya, Bülent. (2019) " Veri Madenciliğı Yöntemleri ve R Uygulamaları", Ankara
- Balaban ME. (2015). Kartal E. Veri madenciliğı ve makine öğrenmesi temel algoritmalar ve R dili ile uygulamaları. İstanbul.
- Çakır, Ö. (2008). "Veri Madenciliğinde Sınıflandırma Yöntemlerinin Karşılaştırılması: Bankacılık Müşteri Veri Tabanı Üzerinde Bir Uygulama," Doktora Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Çalış, A., Kayapınar, S., Çetinyokuş, T., (2014). Veri madenciliğinde karar ağacı algoritmaları ile bilgisayar ve internet güvenliğı üzerine bir uygulama. Endüstri Mühendisliğı Dergisi, 25 (3-4): 2-19.
- Han, J., ve Kamber, M. (2001). Data mining concepts and techniques. Academic Press.
- Orhan, U. (2012). Makine Öğrenmesi Dersi Entropi, Karar Ağaçları (ID3 ve C4.5 algoritmaları), Sınıflandırma ve Regresyon Ağaçları. Ders Notları. Müh.Mim. Fakültesi Bilgisayar Mühendisliğı. Çukurova Üniversitesi. Adana. Retrieved from <http://bmb.cu.edu.tr/uorhan/DersNotu/Ders03.pdf>
- Savaş, S., Topaloğlu N., Yılmaz, M. (2012). "Veri Madenciliğı ve Türkiye'deki Uygulama Örnekleri," İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, sayı 21, s. 1-23.
- Tosun, T. (2006). "Veri Madenciliğı Teknikleriyle Kredi Kartlarında Müşteri Kaybetme Analizi," Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Quinlan, J. R. (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, USA.

## The Use of New Robust Approaches for Nonlinear Quantile Regression When There are Outlying Observations

Burak DİLBER<sup>1</sup>, A. Fırat ÖZDEMİR<sup>1</sup>

<sup>1</sup>Department of Statistics, Dokuz Eylül University, İzmir, Turkey

### Abstract

Iterative algorithms are used for parameter estimation in nonlinear regression models. The iterative method used in nonlinear quantile regression obtains the best parameter estimation by minimizing the sum of the absolute deviations of the residuals. In this study, two different approaches for nonlinear quantile regression are proposed. The first approach minimizes the median of the absolute deviations of the residuals and the second approach minimizes the median of the squares of the residuals. Newly proposed NO quantile estimator is used for the estimation of the median of residuals. Nonlinear quantile regression and new robust approaches are compared with a simulation study by using theoretical data sets consisting of outliers. Furthermore, the methods are examined graphically to understand how they can model different types of relationship between variables. New robust approaches outperformed nonlinear quantile regression against outliers.

**Key words:** Nonlinear quantile regression, NO quantile estimator, nonlinear regression models, outlying observations

## Consumption and Preferences of Cooperative Branded Products in Tokat Provincial Center

Rüveyda YÜZBAŞIOĞLU<sup>1</sup>

<sup>1</sup>Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Tarım Ekonomisi Bölümü, Tokat, Türkiye

### Abstract

Marketing activities carried out through cooperatives protect not only the interests of the producers but also the consumers. A marketing approach centered on producer and consumer interests plays a key role in combating rural poverty. If rural poverty reduction in countries where the majority live in rural areas in developing countries such as Turkey, would be the economy. The place of local products is important for the recognition of the region or for economic development. Tokat province is among the advanced provinces in terms of agricultural products. In this context, the purpose of this study is to examine the consumption and preference of cooperative branded products by individuals who are substituting at the slap center. For this purpose, face-to-face interviews were made with 383 individuals residing in the center. The consumers interviewed within the scope of the research are young and working population. The monthly income of the consumers is 6 120.10 TL / month. More than half of the consumers are male and married. Most of the education level is high school and undergraduate graduates. In Tokat province, the replacement period is long and the rural replacement rate is low. More than half of the consumers do not know the definition of cooperative branded product, but it has been observed that the cooperative branded products are consumed excessively (such as Torku, Marmarabirlik, etc.).

One of the most effective factors in consumers' preference of cooperative brands over other brands is that they think that cooperative branded products are of higher quality. This shows that the society's opinion towards the cooperatives is positive. It has been seen that the society has a positive attitude and behavior towards cooperatives. The society has a desire to support cooperatives. Consumers are aware that cooperative branded products are of higher quality and more affordable, and they expect to see more product options, advertisements and promotions in cooperative brands.

The factors that are econometrically effective in the consumption of cooperative branded products are the consumer's gender, age, income, being of rural origin, knowing the product, price and quality.

Cooperatives can implement policies that encourage and support their branded products in order to protect both the producer and the consumer.

**Key words:** Binary logit, cooperative product, preference, consumption

## Tokat İl Merkezinde Kooperatif Markalı Ürünlerin Tüketimi ve Tercihleri

### Özet

Kooperatifler kanalıyla gerçekleşen pazarlama faaliyeti sadece üreticinin değil tüketicinin de menfaatinin korur. Üretici ve tüketici menfaati merkezli bir pazarlama anlayışı kırsal yoksullukla mücadelede kilit rol oynar. Kırsal yoksulluk azalır da Türkiye gibi çoğunluğun kırsalda yaşadığı ülkelerde ülke ekonomisi kalkınmış olur. Yörenin tanınması ya da ekonomik kalkınmayı sağlaması için yöresel

ürünlerin yeri önemlidir. Tokat ili tarımsal ürünler anlamında ileri iller arasındadır. Bu bağlamda tokat merkezde ikame eden bireylerin kooperatif markalı ürünleri tüketimi ve tercihi incelenmesi bu araştırmanın amacını oluşturmaktadır. Bu amaç için merkezde ikame eden 383 birey ile yüz yüze görüşülmüştür. Araştırma kapsamında görüşülen tüketicilerin genç ve çalışan nüfustur. Tüketicilerin aylık gelir 6 120.10 TL/ay'dır. Tüketicilerin yarısından fazlası erkek ve evlidir. Eğitim seviyesi çoğunluğunun lise ve lisans mezunu oluşturmaktadır. Tokat ilinde ikame etme süreleri uzun ve kırsal ikame etme oranı düşüktür. Tüketicilerin yarısından fazlası kooperatif markalı ürün tanımını bilmeyip, ancak kooperatif markalı ürünleri fazla oranda tüketildiği (Torku, Marmarabirlik, gibi) gözlenmiştir.

Tüketicilerin kooperatif markalarını diğer markalara göre tercih etmelerinde en etkili faktörden biri de kooperatif markalı ürünlerin daha kaliteli olduğunu düşünmeleridir. Bu da toplumun kooperatiflere karşı görüşünün olumlu olduğunu göstermektedir. Görülmüştür ki toplum kooperatiflere ilişkin olumlu tutum ve davranış içerisindedir. Toplumun kooperatifleri destekleme isteği bulunmaktadır. Tüketiciler kooperatif markalı ürünlerin daha kaliteli ve uygun fiyatlı olduğu bilincinde olup kooperatif markalarında daha fazla ürün seçeneği, reklam ve promosyon çalışması görmeyi beklemektedirler.

Kooperatif markalı ürünleri tüketiminde ekonometrik olarak etkili olan faktörler tüketicinin cinsiyeti, yaş, geliri, kırsal kökenli olması, ürünü bilme durumu, fiyat ve kalitesidir.

Hem üreticiyi korumak hem de tüketiciyi korumak adına kooperatifler markalı ürünlerini teşvik edici ve destekleyici politikalar uygulanabilir.

**Anahtar kelimeler:** Binary logit, kooperatif ürün, tercih, tüketim



**Zara Honey That a Local Product Consumption of Individuals and Reasons for Preference (Sivas Provincial Center District)**

Rüveyda YÜZBAŞIOĞLU<sup>1</sup>

<sup>1</sup>Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Tarım Ekonomisi Bölümü, Tokat, Türkiye

**Abstract**

The status of honey consumption, which is very important for human health and also for animal products, and some socio-demographic characteristics that affect consumption have been tried to be revealed with this research. Situated in the front row at the local market in Turkey production is quite high, patented count for much more than the absence of such a honey work well for consumer research has brought to the fore.

The questionnaire, which is the main material of the study, was collected by interviewing 272 individuals living in the central district of Sivas province. In the first part of the study, questions were asked to reveal the socio-economic structures of the individuals. 50.37% of the individuals interviewed within the scope of the research are women, the average age is 33.67 and the average monthly income is 2 034.03 TL. In the second part of the study, honey consumption of individuals was examined. 98.53% of individuals consume honey and more than half (57.09%) prefer to consume honey comb. In the third part, the reasons of preference for the local product, dice honey, are examined. Individuals' reasons for preference for dice were measured with a 5-point Likert scale under 17 factors. Factor analysis was performed to reduce the factor loadings and the cronbach's alpha value was found 0.897 and KMO 0.792, thus 17 variables were summarized under 4 factors. In the last part of the study, some features that affect the reasons of individuals' preference for pesticides have been tried to be determined and binary logit analysis has been applied for this. As a result of the analysis, the monthly income of significant individuals at the 10% significance level and the last factor in the factor analysis negatively affect the pest honey consumption. As a result of the analysis, it was revealed that individuals who were found to be significant at the 5% significance level had a positive effect, while individuals at the 1% significance level had a negative effect.

**Key words:** Honey, Binary logit, Factor analysis, preference, consumption

**Bireylerin Yöresel Ürün Olan Zara Balı Tüketimi ve Tercih Nedenleri (Sivas İli Örneği)**

**Özet**

İnsan sağlığı için çok önemli olan, hayvansal ürünler için de yer alan bal tüketim durumu ve tüketimi etkileyen bazı sosyo-demografik özellikler bu araştırma ile ortaya konulmaya çalışılmıştır. Üretimde ön sıralarda yer alan Türkiye’de ve yerel pazarı oldukça yüksek, patentli sayısı oldukça fazla olan bal için tüketicilere yönelik böyle bir çalışma olmaması araştırmayı iyice ön plana çıkarmıştır. Araştırmanın ana materyali olan anket Sivas ili merkez ilçede yaşayan 272 bireyle görüşme yolu ile toplanmıştır. Araştırmanın birinci bölümünde bireylerin sosyo ekonomik yapılarını ortaya koyacak sorular sorulmuştur. Araştırma kapsamında görüşülen bireylerin %50.37’si kadın, ortalama 33.67 yaş ve aylık ortalama geliri 2 034,03 TL’dir. Araştırmanın ikinci bölümünde bireylerin bal tüketimi incelenmiştir.

Bireylerin %98.53'ü bal tüketmekte ve yarısından fazlası (%57.09) petek bal tüketmeyi tercih etmektedir. Üçüncü bölümde yörenin ürünü olan zara balı tercih nedenleri incelenmiştir. Bireylerin zara balı tercih nedenleri 17 faktör altında 5'li likert ölçeği ile ölçülmüştür. Faktör yüklerini azaltmak için faktör analizi yapılmış ve cronbach's alpha değeri 0.897 ve KMO 0.792 bulunmuş böylelikle 17 değişken 4 faktör altında özetlenmiştir. Araştırmanın son bölümünde bireylerin zara balı tercih nedenlerini etkileyen bazı özellikler belirlenmeye çalışılmış ve bunun için binary logit analizi uygulanmıştır. Analiz sonucunda %10 önem seviyesinde anlamlı bireylerin aylık geliri ve faktör analizindeki son faktör zara balı tüketimini negatif yönlü etkilemektedir. %5 önem seviyesinde anlamlı bulunan bireylerin pozitif yönlü etkilerken %1önem seviyesinde bireylerin negatif yönlü etkilediği analiz sonucunda ortaya konulmuştur.

**Anahtar kelimeler:** Bal, Binary logit, Faktör analizi, tercih, tüketim

## Optimization of Power Conservation and Pressure Drop of Nanofluid Two-Phase Flow in the Pipeline

Asawer A. ALWASITI<sup>1\*</sup>, Zainb Y.SHNEEN<sup>2</sup>, Abbas K.AI SHALAL<sup>3</sup>

<sup>1</sup>Chemical Engineering Department, University of Technology, Baghdad, Iraq,

<sup>2</sup>Chemical Engineering Department, University of Technology, Baghdad, Iraq

<sup>3</sup>Training Center, Baghdad, Iraq

### Abstract

Power consumption in pipeline transportation is one of the most challenging issues in recent days for its economic importance, especially in two-phase flow. Recently, nano additives are used for enhancing flowability and reduce the power conservation in pipelines. Although, other factors like; velocity, physical properties, temperature...etc can affect the flowability, the optimization of these factors was never performed. Hence, in this work, a multifactorial design method, Design of Experiments (DOE) was used to evaluate the interaction effect of mixture velocity, temperature, and nanoparticle concentration on the power conservation and pressure drop of two-phase flow in a horizontal pipe. The optimization of factors was also performed using both experimental and analytical analyses.

The effects of mixture velocity of (oil-water) flow (0.8-2.4) m/s, temperature (20-60) °C, and nano-silica concentration (0-300) ppm and their interactions were investigated experimentally using a horizontal pipe and numerically using a central composite design. The significance of the results was evaluated by statistical analysis of variance (ANOVA)

The results showed that the interaction effects between the mentioned factors were found to be highly significant in influencing power conservation and the predicted values were found to be in good agreement with the experimental values. This study shows that central composite design could efficiently be used for the modeling of power conservation, and it is an economical way of obtaining the maximum amount of information with the fewest number of experiments.

**Key words:** Power conservation, pressure drop, two-phase flow, statistics, central composite design

### Introduction

Petroleum is considered the main income source of a lot of countries due to its importance in exporting issues and industry. It is transferred from one country to another or another place by many methods. The transportation through pipelines system is considered as one of the operatives and economical way to transport petroleum in exporting issue and transferring to storages or refineries. The presence of water with crude oil is so common in pipelines that make the flow is a two-phase flow (liquid-liquid flow). This type of flow is accompanied by an increase in pressure drop and fiction. (Asawer, et al. 2020). Several pumps are used to overcome this issue leading to high power consumption as well as its reversible effect on pipelines transportation by producing pressure more than the maximum allowable operating pressure. Hence, several methods have been used to lower the pressure drop and power consumption as well as lowering friction factors. One of these methods is adding some effective additives. These additives are used to lower the pressure drop and enhance the flowability of fluid like; polymers, surfactants, and nanoparticles.

Recently, nanotechnology is used in reducing the energy required to transport fluid in the pipeline, especially in the petroleum industry. Hascakir et al. (2008) used iron powder in micron-sized to reduce the viscosity of heavy oil. Hamedi (Hamedi and Babadagli, 2010) used a microwave oven to the mixture of heavy oil and different nano-sized metals. Their goal was to attain an effectual aquathermolysis without steam existence, using the photochemical reactions of microwaves to get reactive hydrogen radicals, which would work as hydrogen donors for the reaction of aquathermolysis.

Esteban (2013) studied the effect of SiO<sub>2</sub> on the viscosity of crude oil at steady-state conditions and different concentrations and temperatures. He found that the reduction of viscosity increases as nanoparticle size decreases. Also, they stated that there is a large interaction between nanoparticles interacting and asphaltene aggregates, which causes an increase in the contact area causing a further heavy hydrocarbons fragmentation.

Nada (Nada, 2017) used nanoparticles to water-oil mixture flow in the pipeline. She studied the influence of nanoparticle and nanoparticles-surfactant on the phase inversion point (oil-in-water) into (water-in-oil) flows in a horizontal pipe. She found that at low mixture velocity, there is no effect of additive and velocity on the phase inversion point, while by increasing mixture velocities and using Nano additives, the phase inversion point was delayed to an upper value of the dispersed phase volume fraction (oil volume fraction).

The effect of nanoparticles on the pressure drop in the piping system was studied by the work of (Asawer and Raheeq, 2019). They added locally prepared nanoparticles of SiO<sub>2</sub> at a different concentration to Iraqi oil. They measured the pressure drop experimentally and calculate it theoretically. They declared that 100ppm nanoparticle concentration was the optimum concentration that gives a maximum reduction in pressure drop.

It can be observed from many researches that temperature, velocity a well as additive concentration are all significant parameters that affected pressure drop and power consumption. Although numerous investigations were performed by varying the mentioned factors, the optimization of pressure drop and power consumption between these variables and their interaction was never performed, according to the author's knowledge. Therefore, the present work used a multifactorial design consisting of factors of temperature, nano additives concentration, and velocity to study and evaluate their respective interaction effects on pressure drop as well as power consumption.

One of the methods used to investigate the performance of the multi-leveled fractional design and optimize the desired factors (to obtain either a minimized or maximized response) is the Design of Experiment DOE. It is adapted by many researchers in recent years. Gunaraj (Gunaraj and Murugan 1999) performed Central Composite Design CCD and Response Surface Methodology RSM methods to optimize the effect of interaction between the open-circuit voltages, wire feed rate, welding speed, and nozzle-to-plate distance in the submerged arc welding of pipes. The optimization of the operating variables versus pH value and coagulant dosage was investigated by Ghafari (Ghafari et al. 2009) using CCD and RSM techniques. Furthermore, Zabeti (Zabeti et al. 2009) also made optimization of CaO catalyst activity for biodiesel production using CCD and RSM, they investigated the maximizing of yield production with the interaction effects of precursor dosage and calcination. Ghadimi (Ghadimi 2009) also applied CCD, RSM, and Box-Behnken Design (BBD) in optimizing the effect and interaction between nanoparticle volume concentration and power on the stability and thermal conductivity of nanofluids.

In the present study, the optimization of three variables temperature, nano additive concentration, and velocity of two-phase flow (oil-water) was performed to determine a minimum pressure drop and maximum power conservation in pipeline transportation. CCD and RSM were both chosen as methods

of optimization in this investigation and were both performed to analyze and evaluate the obtained results.

### Optimization Technique

The first step in the optimization technique is identifying the independent factor and desired product. In this work, temperature, nanoparticle concentration, and velocity of the two-phase fluid are the factors, while the pressure drop and power conservation are the desired product. After, it is necessary to determine the effect of these factors, so many experiments should be done to cover the effect of these factors and their interactions on the desired product. The (Experimental Design) is considered as the best method used for this purpose and to minimize the number of experiments required to test the system. In this method, the information from pre-existing data is extracted by using a statistical method to interpret the results in a regular form with the minimum number of observations. This method consists of two parts, the first part is making a design of the experiments according to a certain plan. The Central Composite Design CCD is used for this purpose as it is the proper technique. In this method, the number of experiments is calculated according to the following equation Mahsa et al. (2012), and Marcin et al. (2015).

$$n = 2k + 2k + nc \quad (1)$$

where  $k$  is the number of experimental studied factors. Each experimental factor is tested at five levels: maximum (+1), middle (0), minimum (-1), and two in-between ( $\pm 2$ ).

The second part is completing the regression analysis for the specified set of runs. The analysis of variance (ANOVA) was used to evaluate the adequacy of the models. ANOVA determines which factors have a significant effect on the responses by using Fisher's statistical test (F-test). Effects calculated using ANOVA that falls below a confidence level of 95% were discarded to improve the regression model.

### Material and Method

The schematic diagram of the experimental setup composes of 6 m length with an inner diameter 12.7 mm of acrylic horizontal pipe, pressure gauges, and pumps, as shown in figure 1. The experimental procedure is cited in the work of (Asawer and Nada 2017).

Tap water and crude oil with a fraction (15% water and 85% crude oil) were used as a two-phase mixture. The experiments were carried out at different temperatures ranged from 20°C to 60°C and under different mixture velocity ranged from 0.8 to 2.6 m/s. The effect of nano-silica (NS) addition on the flow behavior was studied at concentrations ranged from 100 to 300 mg/L. These samples were prepared using a magnetic stirrer and an ultrasonic device. The conservation of power was calculated using equation 2.

$$P_S = \frac{P - P_D}{P} * 100\% \quad (2)$$

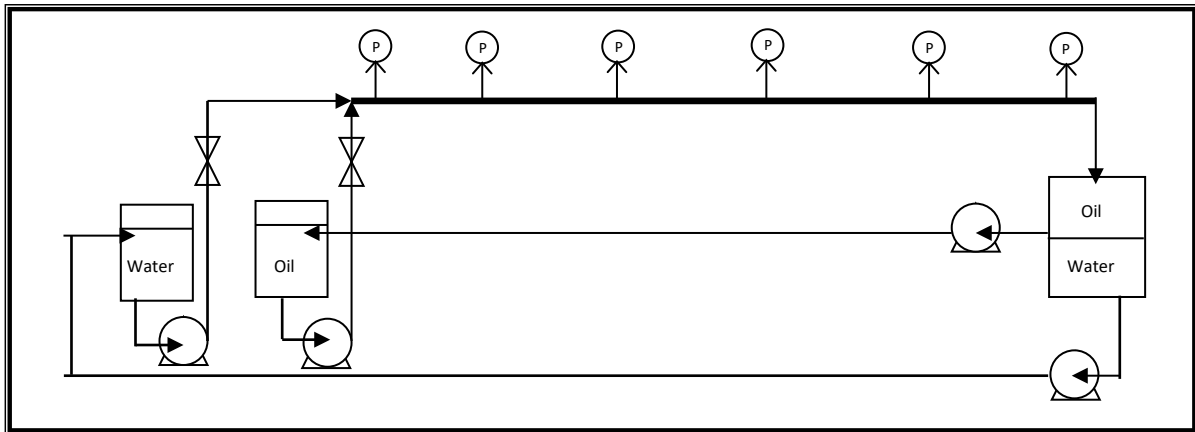


Figure 1. The Schematic Diagram

## Results and Discussion

According to the central composite rotatable design (CCRD) method, twenty experiments were designed to exam the effect of the three operating variables (temperature, nanoparticle concentration, and velocity of the mixture) as well as their interaction effects. Table 1 shows the variable and the value associated with each level

Table 1. Variables and Values Associated with Levels

Variable	Unit	Level				
		-2	-1	0	+1	+2
Temperature	°C	35	20	40	60	55
Nano concentration	ppm	60	0	150	300	230
Speed of mixture	m/s	1.3	0.8	1.7	2.6	2

The experimental results of the twenty runs performed with the corresponding of pressure drop and power conservation are shown in Table 2.

Table 2. Experimental Results According to CCRD

No of experiment	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	Pressure Drop (Pa)	Power conservation (%)
1	-1	-1	-1	4.2	18
2	1	-1	-1	2.9	24.5
3	-1	1	-1	2.2	47.6
4	-1	-1	1	6.3	26.3
5	1	-1	1	4.3	29
6	-1	1	1	5.3	36.1
7	1	1	-1	1.2	45.4
8	1	1	1	2.7	46
9	-2	0	0	3.5	16.6
10	2	0	0	2.5	18.2
11	0	-2	0	3.4	5.55
12	0	2	0	2.6	27.7
13	0	0	-2	5.2	21.8
14	0	0	2	2.7	30
15	0	0	0	3	16.6
16	0	0	0	3	16.6
17	0	0	0	3	16.6
18	0	0	0	3	16.6
19	0	0	0	3	16.6

**Pressure drop**

The Statistical analysis of the experimental results was performed using STATISTICS Software V12 and Expert Design Software. To derive a mathematical correlation of pressure drop, a least-squares method was by fitting a response surface to the measured values. Various models have been used as shown in Table 3.

Table 3: Summary of the model used to fit the average pressure drop

Model	P-value	R <sup>2</sup>	DF	Remark
Linear	0.000	0.844	4	Low R <sup>2</sup>
2FI	0.000	0.986		suggested
Quadratic	0.000	0.937		The p-value of the estimated coefficient are > 0.05
Polynomial	0	0.893		The p-value of the estimated coefficient are > 0.05

The above table reveals that the linear function gives the lower value of R<sup>2</sup> while both quadratic and cubic functions are good enough to fit the data, but all and most of the p-value of the estimated coefficient are > 0.05 respectively. The sequential sum of squares for the two-factor interaction was the best to fit the data, so it was adopted.

Based on the experimental results of pressure drop expressed in and regression analysis of the mathematical model of the sequential sum of squares for the two-factor interaction in CCD with the best value of R<sup>2</sup> is

$$\Delta p = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_1 X_2 + b_5 X_2 X_3 + b_6 X_1 X_3 + b_7 X_1^2 + b_8 X_2^2 + b_9 X_3^2 + b_{10} X_1 X_2 X_3 \quad (3)$$

This model considers linear effects, quadratic effects, two-way interactions as well as the square of two factors between the studied factors.

The analysis of variance (ANOVA) was used to indicate the accuracy of the generated model and to determine the statistically significant factors. Tables 4 and 5 show the performance of ANOVA test. The table shows that some terms (concentration and the quadratic form of velocity) have a higher p-value (>0.05, terms that were not statistically significant).

Table 4. The Regression Analysis Results of Equation 2

	Coefficient estimate	Standard Error	DF	t-value df=9	p-value
Intercept	4.962	1.027	1	4.827	0.0009
Temperature (X <sub>1</sub> )	-0.228	0.062	1	-3.645	0.005
Concentration (X <sub>2</sub> )	0.015	0.009	1	1.695	0.124
Velocity (X <sub>3</sub> )	4.217	1.850	1	2.279	0.0486
X <sub>1</sub> X <sub>2</sub>	0.00001	0.000021	1	3.412	0.0077
X <sub>2</sub> X <sub>3</sub>	-0.017	0.0049	1	-3.475	0.00698
X <sub>1</sub> X <sub>3</sub>	-0.020	0.0034	1	-5.821	0.00025
X <sub>1</sub> <sup>2</sup>	0.0024	0.00076	1	3.175	0.0112
X <sub>2</sub> <sup>2</sup>	-0.000071	0.00003	1	-2.376	0.0414
X <sub>3</sub> <sup>2</sup>	-0.439	0.541	1	-0.811	0.437
X <sub>3</sub> X <sub>2</sub> <sup>2</sup>	0.00005	0.000016	1	3.058	0.013
The proportion of variance (R <sup>2</sup> )	0.995				
R	0.996				

Table 5. Results for ANOVA for the Model

	Sum of Square	DF	Mean Square	F-value	p-value
Regression	262.66	11	23.87	769.68	0.000000
Residual	0.27	9	0.031		
Total	262.94	20			
Corrected Total	42.49	19			
Regression vs Corrected	262.6	11.	23.87	10.67	0.000006

The above model was improved by removing terms with p-values greater than 0.05, considered as statistically insignificant at a 95% confidence level. These removed terms were not taken into account in the regression analysis. The enhanced model is defined by Equation (4)

$$\Delta p = b_0 + b_1 X_1 + b_3 X_3 + b_4 X_1 X_2 + b_5 X_2 X_3 + b_6 X_1 X_3 + b_7 X_1^2 + b_9 X_2^2 + b_{10} X_3 X_2^2 \quad (4)$$

This model predicts pressure drop based on two linear effects, two quadratic effects, three two-way interaction, and one sum of the square of two factors interaction. The statistical summary of the improve the model is shown in Tables 6.

Table 6. The Regression Analysis Results of Equation 3

	Coefficient estimate	Slandered Error	DF	t-value df=11	p-value
Intercept	5.004	0.895	1	0.00	0.00
Temperature ( $x_1$ )	-0.165	0.051	1	0.00	0.00
Concentration( $X_2$ )	2.687	0.184	1	0.00	0.00
$X_1 X_2$	0.00007	0.000023	1	0.00	0.00
$X_2 X_3$	-0.008	0.0020	1	0.00	0.00
$X_1 X_3$	-0.020139	0.003781	1	0.00	0.00
$X_1^2$	0.001654	0.000633	1	0.00	0.00
$X_2^2$	-0.000021	0.000000	1	0.00	0.00
$X_3 X_2^2$	0.000020	0.000000	1	0.00	0.00
Proportion of variance ( $R^2$ )	0.99				
R	0.995				

The summary of ANOVA, after having removed terms with a higher P-value (terms that were not statistically significant) is shown in Table 7. The result shows that the suggested model is best to fit the simulation data. The model's F- and P-values are 787.19 and < 0.0001, respectively.

The model F-value compares the variance of the model to residual's variance, while the P-value measures the probability of obtaining the null hypothesis (none of the input variables affects the pressure drop). From the statistical point of view, these values validate that the model is significant.

Although the improved model didn't change the regression coefficient  $R^2$  (it remains at the same range of 0.98) but all the terms in this model are significant (p-value < 0.05).

Table 7. Results for ANOVA for the Model

	Sum of Square	DF	Mean Square	F-value	p-value
Regression	262.532	9	29.17	787.19	0.000000
Residual	0.407	11	0.037		
Total	262.94	20			
Corrected Total	42.49	19			
Regression vs Corrected	262.53	9	29.17	13.04	0.000002



**Power conservation**

Table 8 shows the suitability of several models to detect power conservation. The best model is the sequential sum of squares for the two-factor interaction

Table 8. Summary of the Model Used to Fit Power Conservation

Model	P-value	R	R <sup>2</sup>	DF	Remark
Linear	0.000	0.956	0.932	4	The p-value of the estimated coefficient are > 0.05
2FI	0.000	0.99	0.986		suggested
Quadratic	0.000	0.971	0.943		The p-value of the estimated coefficient are > 0.05
Polynomial	0	0.939	0.969	6	

Applying equation 3 which gives the best fitting for power conservation, the coefficient estimated as well as its ANOVA test are shown in tables 9 and 10

Table 9. The Regression Analysis Results of Equation 3

	Coefficient estimate	Slandered Error	DF	t-value df=9	p-value
Intercept	386.676	21.26622	1	18.1826	0.000002
Temperature (X <sub>1</sub> )	-0.530	0.41230	1	-1.2863	0.245728
Concentration(X <sub>2</sub> )	-1.034	0.06885	1	-15.0260	0.000005
Velocity (X <sub>3</sub> )	-488.338	16.96930	1	-28.7777	0.000000
X <sub>1</sub> X <sub>2</sub>	0.000	0.00036	1	0.8708	0.417326
X <sub>2</sub> X <sub>3</sub>	1.507	0.07051	1	21.3798	0.000001
X <sub>1</sub> X <sub>3</sub>	0.168	0.00938	1	17.9102	0.000002
X <sub>1</sub> <sup>2</sup>	0.003	0.00390	1	0.7906	0.459263
X <sub>3</sub> <sup>2</sup>	154.196	5.13973	1	30.0008	0.000000
X <sub>2</sub> <sup>2</sup>	-0.488	0.02116	1	-23.0722	0.000000
X <sub>2</sub> X <sub>3</sub> <sup>2</sup>	386.676	21.26622	1	18.1826	0.000002
The proportion of variance (R <sup>2</sup> )	0.999				
R	0.999				

Table 10: Results for ANOVA for the Model

	Sum of Square	DF	Mean Square	F-value	p-value
Regression	12178.94	10.00000	1217.894	10673.26	0.000000
Residual	0.68	6.00000	0.114		
Total	12179.62	16.00000			
Corrected Total	2450.27	15.00000			
Regression vs Corrected	12178.94	10.00000	1217.894	7.46	0.000325

Again the model of equation 2 is modified by removing all factors that give higher values of p (p> 0.05). The modified equation is expressed as follows

$$\Delta p = b_0 + b_1 X_1 + b_3 X_3 + b_4 X_1 X_2 + b_5 X_2 X_3 + b_6 X_1 X_3 + b_7 X_1^2 + b_9 X_2^2 + b_{10} X_3 X_2^2 \quad (5)$$

This model predicts power conservation based on two linear effects, two quadratic effects, two two-way interaction, and one sum of the square of two factors interaction. The statistic and Annova summary of the improve the model are shown in Tables 10 and 11.

Table 11. The Regression Analysis Results

	Coefficient estimate	Slandered Error	DF	t-value df=11	p-value
Intercept	366.247	43.20874	1	8.47622	0.000014
Concentration( $X_2$ )	-0.993	0.14525	1	-6.83894	0.000076
Speed ( $X_3$ )	-479.868	52.70341	1	-9.10506	0.000008
$X_2X_3$	1.474	0.17810	1	8.27514	0.000017
$X_1X_3$	0.078	0.01526	1	5.10606	0.000640
$X_3^2$	152.768	16.08778	1	9.49592	0.000005
$X_2X_3^2$	-0.478	0.05430	1	-8.80842	0.000010
The proportion of variance ( $R^2$ )	0.994				
R	0.997				

Table 12: Results for ANOVA for the modified model

	Sum of Square	DF	Mean Square	F-value	p-value
Regression	12165.76	7.00000	1737.965	1128.203	0.000000
Residual	13.86	9.00000	1.540		
Total	12179.62	16.00000			
Corrected Total	2450.27	15.00000			
Regression vs Corrected	12165.76	7.00000	1737.965	10.639	0.000080

The F-value of the models is large enough to validate the suitability of the model. The  $R^2$  value is 0.994.

### ***Model validity***

In order to ascertain the accuracy of the model and its validity, the normal probability plot and the residual test were examined which was produced by the STATICS software. Figures 2 and 3 show the normality plot of pressure drop and power conservation. The figures also reveal the residuals were close to the straight line and were scattered randomly around it with no particular pattern, showing no abnormal behavior against noise.

The plot of the predicted values of pressure drop and power conservation against the actual.

Values are shown in figures 4 and 5. The data points are evenly scattered on and around the line, showing that the model gave a good prediction of the pressure drop and pow conservation.

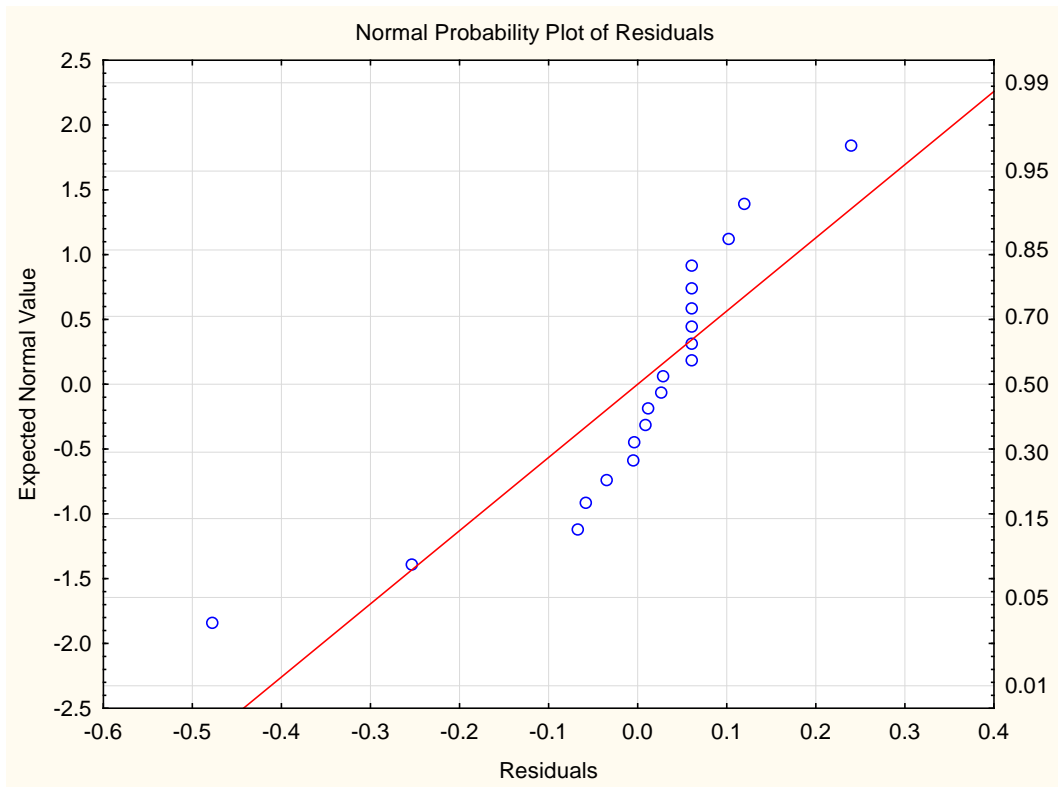


Figure 2. Normal Plots of Residuals for Pressure Drop

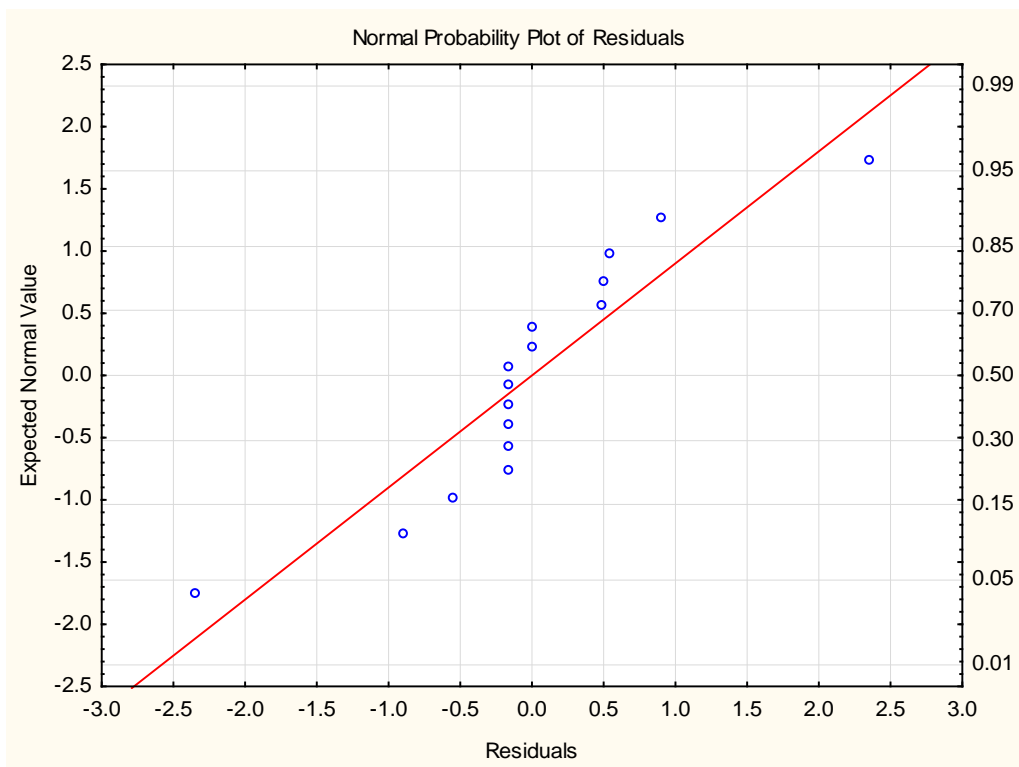


Figure 3. Normal Plots of Residuals for Power Conservation

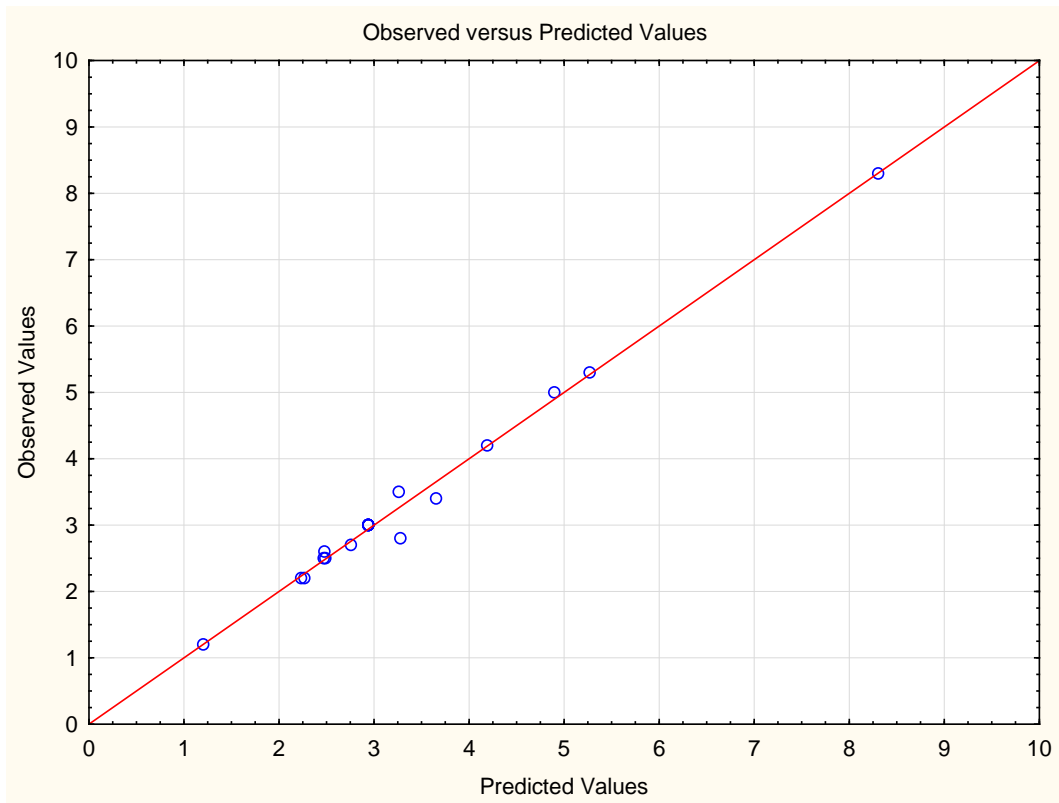


Figure 4. Predicted Values of Pressure Drop

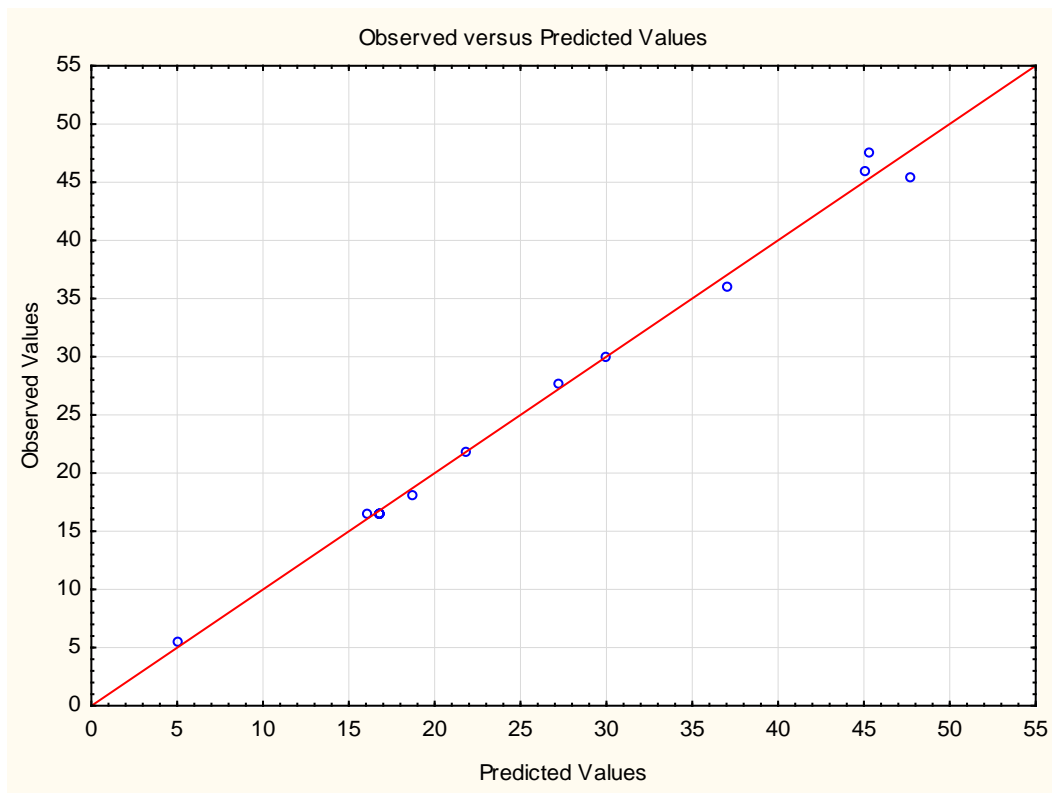


Figure 5: Predicted values of power conservation

**Variables sensitivity**

The sensitivity analysis is a technique adopted by scientists and engineers to detect how the different values of an independent variable can affect the observable output variables under a given set of conditions (Griewank et al. 2008). After obtaining the regression model for the observable output, the sensitivity analysis was carried out to determine the effect of each input parameter on the output variable.

The sensitivity analysis was done by taking the partial derivative of the observable output variable concerning individual input variables. The partial derivative of the regression models for the pressure drop and power conservation concerning the four input parameters are given as:

$$\Delta p = b_0 + b_1X_1 + b_3X_3 + b_4X_1X_2 + b_5X_2X_3 + b_6X_1X_3 + b_7X_1^2 + b_9X_2^2 + b_{10}X_3X_2^2 \quad (4)$$

$$\frac{dp}{dX_1} = -0.165 + 0.0007X_2 - 0.0201X_3 + 0.003X_1$$

$$\frac{dp}{dX_2} = 2.687X_2 + 0.0007X_1 - 0.000801 X_3 + 0.00004X_2 + 0.00004X_2X_3$$

$$\frac{dp}{dX_3} = -0.000801 X_2 - 0.02X_1 + 0.00002X_2^2$$

The partial derivatives of power conservation model are shown below:

$$\frac{dp}{dX_1} = 0.078 X_3$$

$$\frac{dp}{dX_2} = -0.993 + 1.474 X_3 - 0.478X_3^2$$

$$\frac{dp}{dX_3} = -479X_3 + 1.474X_2 + 0.078X_1 + 305.4X_3 - 0.956X_2X_3$$

Tables 13 and 14 show the sensitivity of the input variables (temperature, concentration, and speed) against the pressure drop and power conservation for different configurations. The negative value of the sensitivity denotes the decreasing of observable output variable concerning increment in the input variable, while a positive value implies the opposite.

Table 13 shows that for all configurations, the temperature has a negative sensitivity indicating that increasing temperature causes a decrease in pressure drop and this is expected since that increasing temperature results in decreasing pressure. The positive value of speed in all configuration except one indicates that increasing speed causes an increase in pressure drop values. Nanoparticle concentration shows negative and positive values. The same behavior, of negative and positive values, also appears in table 14 for all variables. This is expected due to the combined effect of variables.

Table 13. The Sensitivity Results of Pressure Drop

Temperature	Nanoparticle concentration	speed	Temperature	Nanoparticle concentration	speed
-1	-1	-1	-0.1486	-2.67965	0.02802
1	-1	-1	-0.1426	-2.67825	-0.01198
-1	1	-1	-0.1472	2.69427	0.01202
-1	-1	1	-0.1888	-2.69575	0.02802
1	-1	1	-0.1828	-2.69435	-0.01198
-1	1	1	-0.1874	2.67833	0.01202
1	1	-1	-0.1412	2.69567	-0.02798
1	1	1	-0.1814	2.67973	-0.02798
-2	0	0	-0.171	-0.0014	0.04
2	0	0	-0.159	0.0014	-0.04
0	-2	0	-0.1664	-5.374	0.01608
0	2	0	-0.1636	5.374	-0.01592
0	0	-2	-0.1248	0.01602	0
0	0	2	-0.2052	-0.01602	0

Table 14. The Sensitivity Results of Power Conservation

Temperature	Nanoparticle concentration	speed	Temperature	Nanoparticle concentration	speed
-1	-1	-1	-0.078	-2.945	171.092
1	-1	-1	-0.078	-2.945	171.248
-1	1	-1	-0.078	-2.945	175.952
-1	-1	1	0.078	0.003	-174.196
1	-1	1	0.078	0.003	-174.04
-1	1	1	0.078	0.003	-173.16
1	1	-1	-0.078	-2.945	176.108
1	1	1	0.078	0.003	-173.004
-2	0	0	0	-0.993	-0.156
2	0	0	0	-0.993	0.156
0	-2	0	0	-0.993	-2.948
0	2	0	0	-0.993	2.948
0	0	-2	-0.156	-5.853	347.2
0	0	2	0.156	0.043	-347.2

### *Optimization results*

The optimization was performed to investigate the optimum operating conditions. In this process the optimum desirability was determined depending on the boundary conditions, or “goals” set for each of the factors and responses. In this work, the goal for temperature, nano concentration, and velocity were set to “in range”. Indeed, the pressure drop and power conservation were set at goals of " minimize" and “maximize”, respectively. RSM in Expert Design Software was used to get the optimum conditions. The results are illustrated in table 15.

Table 15. The Optimization Results

Factor	Range	Lower limit	Upper limit	Predicted	Experimental
Temperate	In range	20	60	56	60
Nanoparticle concentration	In range	60	300	230	300
Mixture velocity	In range	0.8	2.6	0.8	0.8
Pressure drop	Minimize	1.2	8.3	1.1	1.2
Power conservation	Maximize	15	55.6	55	55.6

## Conclusions

The lowest pressure drop and highest power conservation were reported at values of 1.2 Pa and 56.7%, respectively.

From ANOVA, the interaction effects between temperature concentration and velocity was found to be highly significant in influencing the pressure drop as well as power conservation. Therefore, the interaction effects between parameters should not be neglected.

Through optimization, the optimum desired operating condition in this range of study in accordance to the desired goals set was found at a temperature of 56 °C and a velocity of 0.8 m/s, and a concentration 230 ppm.

## References

- Asawer A. Alwasiti, Raheek I. Ibrahim (2019) Improving the Flowability of Heavy Crude Oil in Pipelines Using Prepared Nanosilica: Experimental Investigation and CFD Simulation, *Petroleum and Coal*, 61(6) 1455-1467
- Asawer A. Alwasiti, Zainb Y. Shneen, Raheeq I. Ibrahim, Abbas K. Al (2020) Energy Analysis and Phase Inversion Modeling of Two-Phase Flow with Different Additives, *Ain Shams Engineering Journal*, available On line.
- Asawer, A. Alwasiti, N. M. Farhan (2017) Effect of Surfactant Additives on Phase inversion in Oil-Water Pipe flow System, *Engineering and Technology Journal*, Vol. 35, Part A. No. 2
- Esteban Alberto Taborda Acevedo. (2013) Viscosity reduction of heavy crude oil through the addition of nanofluids on the non-thermal process, Ph.D thesis, National University of Colombia.
- Ghadimi, A. (2013). Stability and thermal conductivity of low concentration titania nanofluids. Doctoral dissertation, University Malaya, Malaysia.
- Ghafari, S.; Aziz, H.; Isa, M.; and Zinatizadeh, A. (2009). Application of response surface methodology (RSM) to optimize coagulation–flocculation treatment of leachate using poly-aluminum chloride (PAC) and alum. *Journal of Hazardous Materials*, 163(2-3), 650-656.
- Griewank A., Walther A. (2008) Evaluating derivatives: principles and techniques of algorithmic differentiation. SIAM, Philadelphia.
- Gunaraj, V.; and Murugan, N. (1999). Application of response surface methodology for predicting weld bead quality in submerged arc welding of pipes. *Journal of Materials Processing Technology*, 88(1-3), 266-275.
- Hamedi Shokrlu, Y., and Babadagli, T., (2010) Effects of Nano-Sized Metals on Viscosity Reduction of Heavy Oil/Bitumen during Thermal Applications”, presented at the Canadian Unconventional Resources and International Petroleum Conference, 19-21 October 2010, Calgary, Alberta, Canada
- Hascakir, B., Babadagli, T., and Akin, S. (2008) Experimental and Numerical Modeling of Heavy-Oil Recovery by Electrical Heating, presented at the 2008 SPE International Thermal Operations and Heavy Oil Symposium, 20-23 October 2008, Calgary, Alberta, Canada
- Mahsa, Azami, Bahram, Morteza, Nouri, Sirous, Naseri, Abdolhosein, (2012). A central composite design for the optimization of the removal of the azo dye, methyl orange from waste water using the Fenton reaction. *J. Serb. Chem. Soc.* 77 (2), 235–246.

- Marcin, Dutka, Ditaranto, Mario, Løvås, Terese, 2015. Application of a central composite design for the study of NO<sub>x</sub> emission performance of a low NO<sub>x</sub> burner. *Energies* 8, 3606–3627. <http://dx.doi.org/10.3390/en8053606>.
- Nada M. Farhan, (2017) Effect of Nanoparticles and Surfactant on Phase Inversion of Two Phases, *Iraqi Journal of Chemical and Petroleum Engineering*, Vol.18 No.1 121 - 128.
- Nada M. Farhan, (2017) Effect of Nanoparticles and Surfactant on Phase Inversion of Two Phases, *Iraqi Journal of Chemical and Petroleum Engineering*, Vol.18 No.1 121 - 128.
- Zabeti, M.; Daud, W.; and Aroua, M. (2009). Optimization of the activity of CaO/Al<sub>2</sub>O<sub>3</sub> catalyst for biodiesel production using response surface methodology. *Applied Catalysis A: General*, 366(1), 154-159.



## Using Entropy Measures in Weighted Least Squares with Non-Linear Datasets

Hıncal TOPÇUOĞLU<sup>1</sup>, Atif EVREN<sup>2</sup>

<sup>1</sup>Yildiz Technical University, Institute of Science, Department of Statistics, İstanbul, Turkey

<sup>2</sup>Yildiz Technical University, Faculty of Science, Department of Statistics, İstanbul, Turkey

### Abstract

For nonlinear data sets, the problem of heteroscedasticity becomes more complicated due to model misspecification. However, in literature, the same remedies like transformation on variables, or the weighted least squares are applied whether the data is linear or nonlinear. In classic weighted least squares methodology, the data points showing higher variation are given smaller weights. On the other hand determining the weights is not straightforward. In this study, we use different entropy measures in determining the weights, since entropy is a measure of uncertainty like standard deviation. Besides, as the generalizations of Shannon entropy, Tsallis, and Rényi entropies provide researchers more flexibility owing to their parametric nature. In the end, we compare the goodness of fit statistics of each weighting alternative assuming different entropy measures. We also emphasize that at least for some of nonlinear data sets, giving weights by some entropy measures provided the best goodness of fit.

**Key words:** Heteroscedasticity, entropy, weighted least squares, Shannon entropy, Tsallis entropy, Rényi entropy

## Baz İstasyonlarında Risk Değerlendirmesi

Muhammed Resul KOZAN<sup>1</sup>

<sup>1</sup>Feridun Çelik Mahallesi 1652. Cadde No: 21/15 Altındağ, Ankara, Türkiye

### Özet

Hayatımızın bir parçası haline gelen bilgisayarlar ve mobil cihazlar, kablosuz iletişim teknolojilerinin hızla gelişmesi sonucu, ticari ve sosyal alanlarda insan hayatını büyük ölçüde kolaylaştırmıştır. Baz istasyonları, Küresel Mobil İletişim Sistemi haberleşmesinin vazgeçilmez elemanlarından biridir. Telekom operatörleri, yüksek iletişim kapasitesine ulaşmak, geniş kapsama alanına sahip olmak ve gelişen yeni teknolojilere uyum sağlamak için baz istasyonlarının sayısını arttırmak zorundadır. Baz istasyonlarının yaydıkları elektromanyetik ışıımından dolayı bu sistemlerin civarında yaşayan insanlar tedirgin olmuş ve istasyon sayılarının arttırılması ile birlikte basında yer alan olumsuz haberlerden ötürü bu tedirginlik daha da artmıştır. Bu makalede tedirginliğe sebep olan baz istasyonlarının insan sağlığı üzerindeki riskleri değerlendirilmiştir. Konu hakkında yazılmış yerli ve yabancı tezlerinde incelenmesi sonucu “istatistiksel anlam taşıyacak bir sağlık riskinin olmadığı, ancak bu konuda çalışmaların sürmesi gerektiği...” ifade edilmektedir. İnsanlarda büyük bir tedirginliğe sebep olmasına rağmen uluslararası tezlerde yapılan risk değerlendirmeleri dikkate alınarak baz istasyonlarının insan sağlığı üzerinde olumsuz bir etkisi bulunmadığı sonucuna varılmaktadır.

**Anahtar kelimeler:** Baz istasyonu, insan sağlığı, mobil cihaz, risk değerlendirme, telekom

## Bound Approximations for Reliability of Circular Consecutive- $k$ -out-of- $n$ Systems

Ahmet DEMİRALP<sup>1</sup>, Mehmet GÜNGÖR<sup>1</sup>

<sup>1</sup>Inonu University, Department of Econometrics, Malatya, Turkey

### Abstract

This paper provides reliability bounds of circular consecutive- $k$ -out-of- $n$  systems with equal component probabilities. This system consists of  $n$  circularly ordered components  $\{1, 2, \dots, n\}$ . The system is failed if at least  $k$  consecutive components fail and sometimes it is hard to determine exact reliability for such coherent system. So, several methods have been developed to determine the useful bounds of reliability for last years. Thus, we investigate the reliability bounds of the circular consecutive- $k$ -out-of- $n$  systems and also, unlike these examinations, some of the boundary approaches applied for the linear failure systems are adapted to circular failure systems. In order to compare the applied approaches, the results obtained for some  $n$  and  $k$  values are given in the tables.

**Key words:** Circular Consecutive- $k$ -out-of- $n$  systems, Coherent system, reliability bounds

**Investigation of Relations between Economic Growth, Energy Use and Human Capital  
with Seemingly Unrelated Regression Model**

Esra CANPOLAT GÖKÇE<sup>1</sup>, Ahmet DEMİRALP<sup>1</sup>

<sup>1</sup>Inonu University, Department of Econometrics, Malatya, Turkey

**Abstract**

The aim of this study is to examine the relationship between the economic growth, energy use and human capital variables of seven developed member countries with the help of seemingly unrelated regression model. If there is a relationship between the errors of the equations in the equation system due to the seemingly unrelated regression models, estimates that take into account this relationship can be made. In this study, in the seemingly unrelated regression model, gross domestic product per capita as an indicator of economic growth, amount of energy used per capita for energy use variable, and education and health expenditure data per capita for countries were used as indicators of human capital. The variables consist of annual observations including the years 2000-2015 for seven countries. This relationship is created from the data to be used for testing of our model are explored with seemingly unrelated regression. According to the results obtained, the coefficients obtained for the independent variables in the model where the gross domestic product variable is the dependent variable appear to be statistically significant. Health expenditure and population variables have positive effects on the gross domestic product variable and the effect of education expenditures is negative. In the model where the energy consumption variable is a dependent variable, the coefficients of all variables are statistically significant. While health spending and population growth increase energy consumption, the increase in education expenditures reduces energy consumption. The results obtained from the econometric analyzes show that the effect of education expenditure is negative when we show that the effect of health expenditure on human capital investment and population variable is positive. This means that policies on education expenditures need to be made more efficient.

**Key words:** Economic growth, energy use, seemingly unrelated regression

## Simulation of Reorder Point with Variable Demand

Emre ASLAN<sup>1</sup>

<sup>1</sup>Tokat Gaziosmanpaşa University, Dept. of Business Adm., Tokat, Turkey

### Abstract

In this study, it is aimed to test whether the reorder point value calculated with determined service level equals calculated service level. For this purpose, a simulation model is built in MS Excel, reorder point is calculated and run for the average demand, standard deviation, lead time and service level, and as a result, the number of orders and backorders are counted during this process and these are proportioned to each other. If this value, which shows the backorder percentage in order cycles, is subtracted from 1, it will give the service level, which is the percentage of not running out of stock in all order cycles. The service level calculated as a result of the experiment was compared with the service level determined at the beginning. Since it is a stochastic experiment, different results occur each time the experiment is repeated. For this reason, it is necessary to repeat the experiment until the average results come to a steady state. The experiment is replicated for 10 times. The average value of calculated service levels of 10 replications is found almost 81% which is fairly close to desired service level 80%.

**Key words:** Reorder point, service level, simulation, variable demand

### Introduction

The problem of determining the reorder point (ROP) is fairly easy when the demand ( $d$ ) and lead time ( $L$ ) for a single product are constant. When the lead time is multiplied by the demand, the amount of product needed in the period after the order is placed until it is supplied is calculated as the reorder point. In this way, when the order is delivered, the inventory level decreases to zero and increases to order amount ( $Q$ ) at the same time. However, when the demand is variable, if it is distributed normally, while it is a reasonable approach to multiply the average demand by the lead time to determine the reorder point, out of stock may arise if the demand is above average during the lead time. For this reason, when the demand varies, it is necessary to calculate the safety stock while calculating the reorder point. Considering the number of standard deviations ( $z$ ) for a specified service level and the standard deviation of the demand during the lead time ( $\sigma_d\sqrt{L}$ ), it is necessary to calculate a safety stock and add it to the demand during the lead time. In other words, in a situation where the demand is normally distributed, the reorder point is calculated by  $ROP = \bar{d}L + z\sigma_d\sqrt{L}$  this way (Krajewski et al., 2016: 352-355; Heizer et al., 2017: 508-511).

Service level is a probability value for not running out of stock in an order cycle. For instance if the desired probability for not running out of stock is 50%, in other words service level is 50%,  $z$  value corresponding to 50% of the area under normal distribution curve is zero. If the service level is 90%,  $z$  value corresponding to 90% of the area under normal distribution curve is 1,2815. The desired and predetermined service level affects the safety stock, as service level increases, safety stock increases, so reorder point increases. The question is whether the reorder point provides the service level. In this study the aim is .....

## Material and Method

A hypothetical example was constructed for comparing desired and calculated service levels. Demand ( $d$ ) for a product is normally distributed with a mean ( $\bar{d}$ ) 100 and standard deviation ( $\sigma_d$ ) 15. Lead time ( $L$ ) is 4 days and desired service level is 80%.  $z$  value for 80% service level is 0.84162. According to this example reorder point ( $ROP$ ) is calculated as in the formula;

$$ROP = 100 \times 4 + 0.84162 \times 15 \times \sqrt{4} = 425.2486 \approx 425$$

So when the inventory level decreases to 425 a new order must be placed. Quantity of order ( $Q$ ) is determined equal to  $ROP$  which is 425 and initial stock level is determined equal to  $\bar{d} \times L$  which is 400 units. All the values in the example are given in Table 1.

Table 1. Parameter Values of Example

Parameter	Value	
Average demand per period ( $\bar{d}$ )	100	
Standard deviation of demand ( $\sigma_d$ )	15	
Lead time ( $L$ )	4	
Service level and $z$ value	80%	0.84162
Reorder point ( $ROP$ )	425	
Initial stock level	400	
Quantity of order ( $Q$ )	425	
Number of replications	10	

After calculating  $ROP$  according to the given values, the question is “Does the calculated reorder point really provide desired service level?” To answer this question a simulation model is built in MS Excel, which works according to the numbers in Table 1. The variables used and algorithm of the simulation are as follows:

$i$	: period	$BO_i$	: backorders on period $i$
$d_i$	: demand on period $i$	$RO_i$	: received order on period $i$
$\sigma_{di}$	: standart deviation on period $i$	$TLD_i$	: time left to delivery on period $i$
$L$	: lead time	$SIML$	: simulation length
$Q$	: quantity of order	$NOR$	: number of replications
$ROP$	: reorder point		
$IPB_i$	: inventory position before an order on period $i$		
$IPA_i$	: inventory position after an order on period $i$		
$OH_i$	: on-hand inventory on period $i$		
$SRB_i$	: scheduled receipts before an order on period $i$		
$SRA_i$	: scheduled receipts after an order on period $I$		

Set initial parameters $d, \sigma_d, L, Q, ROP$ , service level, initial stock level, $SIML, NOR$
Starting from 1 up to $NOR$
$OH_0$ =initial stock
$IPA_0 = OH_0$
Starting from $i=1$ up to $i= SIML$ ;
Determine demand $d_i$ according to norm. dist. with mean $d$ and std. dev. $\sigma_d$
Control Received Order ( $RO_i$ )
Calculate on hand inventory ( $OH_i = OH_{i-1} + RO_i - d_i - BO_{i-1}$ )
If $OH_i \geq 0$ then $OH_i = OH_i$ else $OH_i = 0$
Calculate backorder ( $BO_i = OH_{i-1} + RO_i - d_i - BO_{i-1}$ )
If $BO_i < 0$ then $BO_i =  BO_i $ else $BO_i = 0$
Calculate inventory position ( $IPB_i = IPA_{i-1} - d_i$ )
Calculate scheduled receipt before an order
If $TLD_{i-1} < 1$
then $SRB_i = 0$
else $SRB_i = (\text{Count delivery dates If}(\text{from period 1 to } i-1) > i) \times Q$

```

Determine order;
    If  $IPB_i \leq ROP$ 
    then
        set an order  $Q$ 
        calculate scheduled receipt after an order
            If  $Q > 0$  then  $Q$  else 0
        calculate inventory position ( $IPA_i = OH_i + SRB_i + SRA_i - BO_i$ )
        determine delivery date
            If random number < service level
            then  $(i + L)$ 
            else  $(i + L + 1)$ 
        calculate day left ( $TLD_i = L$ )
    else
         $Q = 0$ 
        calculate inventory position ( $IPA_i = OH_i + SRB_i - BO_i$ )
        calculate day left ( $TLD_i = (TLD_{i-1} - 1)$ )
Calculate service level  $(1 - \Sigma BO / \Sigma RO)$ 
    
```

Simulation length is determined as 1500 days (periods) which is long enough for adequate number of orders and calculated values to become steady. The experiment is replicated for 10 times. An image of model may be seen in Figure 1.

Average d	Standart d	Average d	Standart d	Lead Time	Initial stoc	Reorder P	Order Qua	Service lev	Service lev	# of Receiv	# of Backorders				
100	15	99,85533	15,38609	4	400	425	425	0,8	0,809659	352	67				
		Average d		Standart d		Number of Received Orders		Backorders		Number of Service level of experiment					
		99,85533		15,38609		352		67		0,809659					
Day	D	OH	RCV	SR before	SR after	ar	BO	IP before	aQ	IP after	orc	Due Date	Day Left	Order Nun	ABO
0			400					0	400	0	400			0	0
1	106	294	0	0	425	0	0	294	425	719	5	4	1	0	
2	96	198	0	425	0	0	0	623	0	623	-	3	1	0	
3	89	109	0	425	0	0	0	534	0	534	-	2	1	0	
4	96	13	0	425	0	0	0	438	0	438	-	1	1	0	
5	98	340	425	0	425	0	0	340	425	765	9	4	2	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1497	74	80	0	425	0	0	0	505	0	505	-	1	353	0	
1498	86	419	425	0	425	0	0	419	425	844	1502	4	354	0	
1499	104	315	0	425	0	0	0	740	0	740	-	3	354	0	
1500	79	236	0	425	0	0	0	661	0	661	-	2	354	0	

Figure 1. Simulation Model in MS Excel

### Findings

After 10 replications of 1500 days length simulation model calculated service level values are found as in Table 2. The average value of calculated service level values of 10 replications is determined as 0.810307. Almost 81% service level is calculated according to 10 replications and it is almost equal to desired service level value 80%. The explanation of the difference between desired and calculated service levels may arise from number of replications to be small. If number of replications increase the difference may decrease. Another reason may be the structure of inventory system and simulation model. The inventory system is a continuous review system however simulation model built in Excel is discrete. Determination of delivery date is discrete in the simulation model, but continuous inventory system places an order as inventory level decreases to  $ROP$ . A more sensitive mechanism that imitates continuity for delivery date determination may overcome this problem.

Table 2. Calculated Service Level Values of 10 Replications

<b>Replication Number</b>	<b>Calculated Service Level</b>
Replication 1	0.818182
Replication 2	0.798867
Replication 3	0.830028
Replication 4	0.810198
Replication 5	0.798867
Replication 6	0.822034
Replication 7	0.806818
Replication 8	0.810734
Replication 9	0.807910
Replication 10	0.799435
<b>Average value</b>	<b>0.810307</b>

### **Conclusion**

In this study, it is aimed to test whether the reorder point value calculated with desired service level equals to calculated service level. For this purpose, a simulation model is built in MS Excel, reorder point is calculated and run for the average demand, standard deviation, lead time and service level, and as a result, the number of backorders and orders are counted during this process and these are proportioned to each other to calculate service level of the simulation. This process is replicated for 10 times and the average of 10 replications is found to be almost 81%, while desired service level is 80%. The small difference may arise from small number of replications or different structures of inventory system and simulation model. In further research the it would be helpful to make experiments with a bigger number of replications, with different level of parameters, for instance 90% service level. And also variable lead time with variable demand may be studied.

### **References**

- Heizer, J., Render, B., & Munson, C. (2017). Principles of Operations Management: Sustainability and Supply Chain Management (Global editon). Pearson.
- Krajewski, L. J. R., Malhotra, M.K., & Ritzman, L.P. (2016). Operations management: processes and Supply chains. 11th Ed. Pearson prentice–Hall. New Jersey.



## Automation of Experimental Design with MS Excel VBA

Emre ASLAN<sup>1</sup>

<sup>1</sup>Tokat Gaziosmanpaşa University, Dept. of Business Adm., Tokat, Turkey

### Abstract

A designed experiment is a test or set of tests in which intentional changes are made in the input variables and changes in the responses of the corresponding output variables are observed and defined. An experiment can be designed to determine the variables that have the greatest impact on the output, to adjust these variables so that the variability of the output is small or the effect of uncontrollable variables is minimal. As the number of parameters, level of parameters increase, the number of experiments to be done also increase. In this study, it is aimed to automate the repetition of experiments for many parameters and levels with MS Excel Visual Basic Applications (VBA).

**Key words:** Excel, experimental design, VBA

### Introduction

An experiment is defined as test or series of tests in which intentional changes are made to the input variables of a process or system so that changes may be observed in the output response (Montgomery, 2001: 1). An experiment may be run for one or more of the following reasons: to determine the principal causes of variation in a measured response, to find the conditions that give rise to a maximum or minimum response, to compare the responses achieved at different settings of controllable variables, to obtain a mathematical model in order to predict future responses (Dean et al., 2017: 1).

There are three common strategy of experimentation. First one is best-guess approach in which the experimenter makes a guess about the factor depending on his/her technical knowledge. Another strategy is one-factor-at-a-time. It consists of varying each factor from the lower to the upper level while the other factors being constant at their lower level. This strategy is missing the interaction between factors. Making a factorial experiment while dealing with several factors is right strategy (Montgomery, 2001: 2-4).

There are two aspects to any experimental problem: the design of the experiment and the statistical analysis of the data (Montgomery, 2001: 11). The stages of an experimental design are; definition of the problem, selection of parameters and levels, choice of response variable, selection of experimental design, experimentation, data analysis and conclusion (Montgomery, 1997: 506-507).

If a stochastic rather than deterministic experiment is performed, each experiment will have to be repeated as a different result may be obtained in each experiment. These processes will also be quite time consuming as the parameter, level and number of replications increase. While  $p_i$  indicating parameter  $i$ ,  $l_i$  indicating level of parameter  $i$  the total number of experiments to be performed is calculated as  $(\prod l_i) \times \text{rep}$ . From here, as the number of parameters, the levels in each parameter and the number of replications increases, the number of experiments to be performed will increase.

### Material and Method

In the experimental design made in MS Excel, parameter combinations for all levels from the lower limit to the upper limit of each parameter are created and printed in the target cells, the parameter values in these cells are considered as input and spreadsheet automatically calculates the new experiment values

as much as the number of replications and prints the output values to the required cells and in this way designed number of experiments and results are obtained. The variables used and algorithm that automates experiments are as follows.

<p><math>p_i</math> : parametre i  <math>P</math> : Total number of parameters  <math>l_i</math> : level of parameter i  <math>low_i</math> : lower level of parameter i  <math>high_i</math> : upper level of parameter i  <math>rise_i</math> : increase in parameter i at each levels  <math>rep</math> : number of replications</p>
<p>Starting from 1 up to P (total number of parameters) for i              Starting from <math>low_i</math> up to <math>high_i</math> for <math>p_i</math> with an increase <math>rise_i</math>                  Print input parameters to the target cells              Starting from 1 up to <math>rep</math> for r                  Print whole input parameters and response variable to the target cells</p>

In the experimental design made in MS Excel, parameters, lower and upper levels and increment may be seen in Table 1. Input cell in the fifth column of the table is filled with the value to make the experiment after running VBA. In the example there are 5 parameters, which have different number of levels. By multiplying all number of levels total number of experiments can be found as  $10 \times 9 \times 5 \times 19 \times 5 = 42,750$ . If 10 replications are desired for all experiments than total number of experiments can now be found as 427,500. Duration of the whole experiments depends on time of each experiment. In the example time for 10 replications takes about 4 seconds. Time for whole experiments to be done takes about 2 days (47,4 hours). The first experiment starts with following values;  $P_1 = 3$ ,  $P_2 = 2$ ,  $P_3 = 0$ ,  $P_4 = 0.5$ ,  $P_5 = 200$  and replicated for 10 times; in the second experiment all parameter values remain constant except  $P_5$  rises to 250 and replicated for 10 times, this goes on and final experiment is done with values  $P_1 = 30$ ,  $P_2 = 10$ ,  $P_3 = 400$ ,  $P_4 = 0.95$ ,  $P_5 = 400$  and replicated for 10 times.

Table 1. Parameters, Lower and Upper Levels

Parameters	Lower	Upper	Increase by	Input Cell	Number of levels
P1	3	30	3		10
P2	2	10	1		9
P3	0	400	100		5
P4	0.5	0.95	0.5		19
P5	200	400	50		5

This study aims to automate the experimentation stage of experimental design. Data analysis which is the following stage of experimental design is out of this study's scope.

### Conclusion

In this study, it is aimed to automate the experimentation stage of experimental design by the help of MS Excel VBA. The calculations of the experiment is made in a spreadsheet. As the number of parameters, level of parameters and number of replications increase in an experiment, number of experiments to be done also increases. It becomes time consuming to make all those experiments manually. By the help of VBA, input parameter values are determined within lower and upper levels of parameters, and experiments are made and output values are recorded to the target cells automatically.

### References

- Dean, A., Voss, D., & Draguljić, D. (2017). Design and analysis of experiments. 2nd Ed. New York: Springer.
- Montgomery, D.C. (2001). Design and Analysis of Experiments. 5th Ed. USA: John Wiley & Sons.
- Montgomery, D. C. (1997). Introduction to statistical quality control. USA: John Wiley & Sons.

## Bilgisayar ve Telekom Sektörleri Ürün Gruplarının Üzerine Bir Uygulama

Aylin UÇAN<sup>1</sup>, Doğan YILDIZ<sup>1</sup>

<sup>1</sup>Yıldız Teknik Üniversitesi, Fen Edebiyat Fakültesi, İstatistik, İstanbul, Türkiye

### Özet

Çok boyutlu ölçekleme (ÇBÖ) analizinin kullanımını konu alan bu çalışma, ÇBÖ analizinin teorik zeminde incelenmesi, bilgisayar ve telekom sektörlerinin temel ürün gruplarının, örneklem olarak seçilen teknoloji marketlerinin 2019 satış adetlerinin ürün, bölge ve markasal dağılımları hakkında örneklendirme olmak üzere üç temel amaca hizmet etmeye çalışmaktadır. Bu çalışma sonucunda sektör bazında en çok satılan ürün gruplarını belirleyip, bölgesel ÇBÖ analizi çıktıları, işletmelere gerek rekabet yarışında ayrıcalıklı bir konum kazanmada gerekse pazardaki boşlukları tespit ederek, bu boşluklarda da konumlanabilme hakkında bakış açıları kazandırabilecektir. Marka ÇBÖ analizi çıktıları ile hangi marka ve ürün grubunun diğerlerine göre daha fazla satış adeti olduğu belirlenmeye çalışılacaktır. Benzer ve farklı markaları konumlandırmaya yardımcı olabilecektir. (Marka gizliliğinden dolayı analiz boyunca marka isimleri gizli tutulacaktır).

**Anahtar kelimeler:** Konumlandırma, çok boyutlu ölçekleme, marka, bölge, ürün

### Abstract

This study, which focuses on the use of multidimensional scaling (MDS) analysis, serves three main purposes, namely, analyzing the MDS analysis on a theoretical ground, exemplifying the product, regional and brand distributions of the main product groups of the computer and telecom sectors, and the 2019 sales numbers of the technology markets selected as a sample. works. As a result of this study, by determining the best selling product groups on the basis of the sector, the outputs of the regional MDS analysis will be able to provide the enterprises with a view to gain a privileged position in the competition race and to identify the gaps in the market and to position them in these gaps. With the outputs of the Brand MDS analysis, it will be tried to determine which brand and product group has more sales than others. It will be able to help position similar and different brands. (Brand names will be kept confidential throughout the analysis due to brand privacy).

**Key words:** Positioning, multidimensional scaling, brand, region, product

### GİRİŞ

Gelişen teknoloji çağı ile elektrik-elektronik ürün gruplarında pek çok değişiklik meydana gelmiştir ve hala gelmeye devam etmektedir. Tüketicinin daha çok günlük hayatta sosyal medya, alışveriş, iletişim, eğlence gibi alanlarda kullandığı bilgisayar ve telekom ürün grupları vazgeçilemez araç gereç haline gelmişlerdir. Firmalar, tüketicilerin artan ilgisinin ne yönde, nasıl olduğunu anlayabilmek için pek çok araştırma yapmaktadırlar.

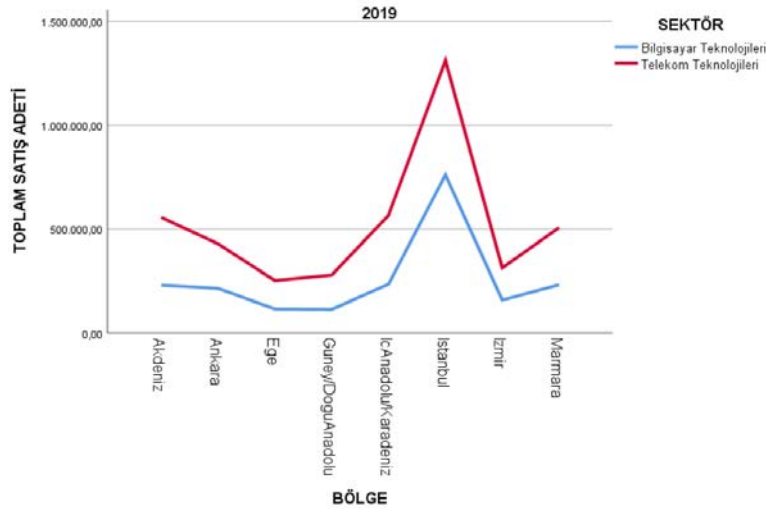
Pazar; en geleneksel tanımıyla, alıcı ve satıcının bulunduğu ortamdır. Bu ortam, satıcının ürün veya hizmetlerinin alıcıya ulaştırılma faaliyetlerinin tümünü kapsar. Pazar payı ise işletmenin pazar içindeki etkinlik alanını yani payını anlatan bir tabirdir. Bir malın toplam satış miktarı ya da gelirleri içinde bir firmaya ait olan pay anlamına gelen pazar payı, bir şirketin pazardaki gücünün ve karlılığının da en önemli göstergesidir. Pazarlama, pazar araştırması ve reklamcılık dahil olmak üzere ürün veya

hizmetlerin tanıtımını ve satışını yapma eylemidir. Konumlandırma hedef pazardaki tüketicilerin zihninde uygun bir yere sahip olmak ve rakiplerden farklılaşmak için; ürün hakkındaki iletişim bilgisi olarak tanımlanır. Konumlandırma hedef pazarlama işleminin temel sürecidir. İşletmelerin pazarda başarılı olabilmek için ürün, fiyat, dağıtım kanalları ve konumlandırma konusunda hem kendini hem de rakipleri tanıması gerekir.

Bu çalışmanın amacı, ürünlerin satış hacimlerini konumlandırmaya, tüketicinin daha çok hangi ürün grubuna yönlendiğini anlamaya uygun bir metod olarak hizmet edebilen çok boyutlu ölçekleme analizini varsayımları, parametreleri ve sunduğu görsel çıktılar açısından incelemektir. Bilgisayar ve telefon teknolojilerinde kullanılan ürünler benzer özelliklere sahip oldukları kadar pazarda farklı satış hacimlerine sahip olabilmektedirler. Araştırmaya, Türkiye’de satış yapan teknoloji marketlerinin bilgisayar ve telekom teknolojileri temel ürün grupları dahil edilmiştir. Sektörlere göre ürün grupları sırası ile aşağıdaki gibidir:

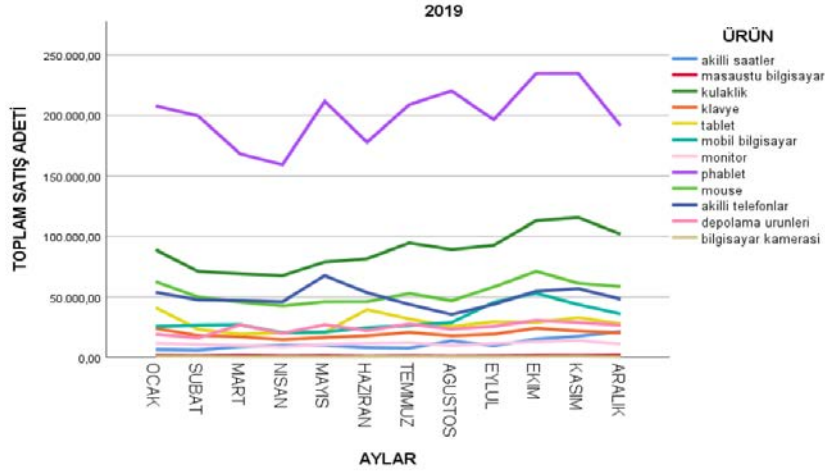
- Bilgisayar teknolojileri temel ürün grupları: depolama ürün grupları, klavye, bilgisayarlar, mouse, mediatablet, mobil bilgisayar, monitörler, kameralar
- Telefon teknolojileri temel ürün grupları: akıllı saatler, akıllı telefonlar, kulaklıklar, phablet ürün gruplarıdır.

Çalışmaya 2019 yılı aylarının ve markaların satış adetleri dâhil edilmiştir. Satış adetlerinin benzerliğine göre iller 8 bölgede gruplanmıştır. Elde edeceğimiz çıktılar ürün gruplarının bölgesel durumlarını anlayıp, tüketici tarafından en çok tercih edilen ürün gruplarını belirleyebilmemize ve ürünler arasında ilişki kurabilmemize yardımcı olacaktır. Marka isimleri veri gizliliğinden dolayı gizli tutulacaktır.



Şekil 1. Sektörlerin 2019 Yılı Toplam Satış Adetleri Bölge Çizgi Grafiği

Her iki sektörün 2019 yılı toplam satış adetleri bölge bazlı olarak Şekil 1’deki gibidir. 2019 yılında telekom sektöründe belirtilen ürün gruplarının satış adetleri daha fazladır. Sektörlerin bölgelere göre göstermiş oldukları satış trendleri benzerlik göstermektedir. En çok satış adetleri İstanbul bölgesinde bulunmaktadır.



Şekil 2. Ürün Gruplarının 2019 Yılı Aylara Göre Toplam Satış Adetleri Çizgi Grafiği

Şekil 2’de aylara göre benzer satış trendine sahip ürün grupları mevcuttur. Phablet, kulaklık, akıllı telefon ürün grupları diğer ürün gruplarına göre daha çok satış adetine ve farklı satış trendine sahiptir.

Tablo 1. Bilgisayar ve Telekom Teknolojileri Sektörlerinin Ürün Gruplarının 2019 Yılı Toplam Satış Adetleri %’lik Oranları

Bilgisayar Teknolojileri 2019	
Ürün Adı	% Satış Adeti Oranı
bilgisayar kamerası	0%
masaüstü bilgisayar	1%
monitor	7%
klavye	11%
depolama ürünleri	14%
tablet	17%
mobil bilgisayar	18%
mouse	31%

Telekom Teknolojileri 2019	
Ürün Adı	% Satış Adeti Oranı
akıllı saatler	3%
akıllı telefonlar	14%
kulaklık	25%
phablet	57%

2019 yılında satışlardaki %’lik oranlar Tablo 1’deki gibidir. 2019 yılında bilgisayar teknolojileri sektörünün en çok satış hacmi olan ürünü mouselardır. Masaüstü bilgisayarlar ve kameralar ise diğer ürün gruplarına göre en az satış adeti olan ürünlerdir. Benzer şekilde telekom ve teknolojileri ürün gruplarında örnekleme en çok satış payına sahip olan ürün grubu %57 ile phablet ürün grubudur. Telefon ile en çok kullanılan kulaklık ürün grubu da bu sektörde %25 ile phabletten sonra en çok satış adeti olan ürün grubudur.

## Literatür

Literatür taraması sonucunda, pazarlama, marka, bölge ve ürün konusunda ülkemizde şu ampirik çalışmalara rastlanılmıştır:

- Ustaahmetoğlu ve Say (2005) araştırmalarında Türkiye’de faaliyet gösteren 9 gazlı içecek firmasının markalarının, tüketicilerle algılanılan konumunu araştırabilmek için tüketicilerle bir

anket gerçekleştirip elde edilen verileri çok boyutlu ölçekleme analizi ile analiz etmişlerdir. Kola tüketiminin kişinin gelirinden çok alışkanlıkları ile ilgili olduğunu, eğlenme alışkanlığı ile kola tüketimi arasında da ilişki olduğunu ve aynı zamanda kolanın susuzluk giderici bir içecek olarak düşünüldüğünü tespit etmişlerdir. Markalara susuzluk giderici bir içecek ile piyasaya ürün çıkarabilmeleri hakkında öneride bulunulmuştur. Ek olarak en çok tercih edilen markanın Coca-Cola ve en az popülerliği olan markanın Zemzem markası olduğunu tespit etmişlerdir.

- Yenidoğan (2008) araştırmasında üniversite öğrencilerinin spor ayakkabı markaları üzerinde bir marka algısı çalışmasını gerçekleştirmiştir. Çalışmada markalar gizli tutulmuştur. Çalışma sonunda markaların avantajlı fiyat sunumu ile benzer algılandıkları tespit edilmiştir. Ayrıca markalardaki farklılığın spor ayakkabının tarzına bağlı olduğuna dair sonuçlarda ortaya çıkmıştır. Çalışma sonunda düşük ve yüksek fiyat satışı olan markalar belirli bir şekilde gruplanmıştır. Çalışma, piyasaya yeni girecek olan bir spor ayakkabı markası için pazarda kendisine doğru konumu bulabilmesi açısından bir bakış açısı olmuştur.

### Çok Boyutlu Ölçekleme Analizi

Çok boyutlu ölçekleme analizi çok değişkenli ve keşfedici bir veri analiz tekniğidir. Çok boyutlu ölçeklemenin öncelikli amacı orijinal verilerin mümkün olduğunca az boyutlu bir uzayda gösterim ile yorumlama kolaylığı sağlamak ve bir koordinat sistemine yerleştirilmesidir (Johnson ve Wichern, 1999:761). Böylece farklılıklar bir grafik ile görselleştirilmekte ve açıkça görülebilmektedir. Bu haritada nesnelere noktalar olarak gösterilmektedir. Analizde, nesnelere benzerlikleriyle nesnelere arasındaki uzaklıklar karşılaştırılmaktadır. Noktalarla temsil edilen benzer nesnelere birbirine daha yakın, farklı nesnelere ise birbirine daha uzaktır (Wickelmaier, 2003). Çok boyutlu ölçekleme analizi, veriler ile ilgili dağılım varsayımı gerektirmeyen bir yöntemdir (Özdamar, 2004: 502).

Borg ve Groenen (1997) 'e göre ÇBÖ analizinin genel amaçları aşağıdaki gibidir:

1. Veriyi görsel olarak keşfedilebilir yapmak amacıyla benzerlik/benzemezlik verisini daha az boyutlu bir uzaydaki uzaklıklar olarak yansıtan bir methodur.

2. Farklı nesnelere arasındaki farkı ortaya koymaya yarayan bir teknik olarak bu nesnelere ampirik farklarını yansıtmakta kullanılır.

3. Benzerlik/benzemezlik yargılarının altında yatan boyutları keşfetmek için kullanılan veri analitik bir yöntemdir.

Analizde temel ayrım, metrik ve metrik olmayan modeller arasındaki farktır. Metrik ve metrik olmayan modeller, çok boyutlu ölçekleme modeli tarafından tahmin edilebilen koordinatlarla hesaplanan, veri ve uzaklıklar arasındaki ilişkiler hakkında farklı varsayımlara sahiptir. Metrik modeller ilişkinin en az aralık ölçeği ile ölçülmüş niteliklere sahip olduğunu varsayarken, metrik olmayan modeller sadece ordinal ölçekle ölçülmüş nitelikleri gerektirmektedir (Mackay ve Zinnes, 1986:325-344). Çok boyutlu ölçekleme tipi benzerlik verisinin niteliksel (metrik olmayan ÇBÖ) veya niceliksel (metrik ÇBÖ) olup olmadığına göre özel bir takım türlere de ayrılmaktadır. Bu özel modeller, Klasik ÇBÖ (Classical MDS), Tekrarlı ÇBÖ (Replicated MDS) ve Ağırlıklı ÇBÖ (Weighted MDS) olarak sınıflandırılabilir (Young, 2007).

Klasik ÇBÖ, matris bağlantılı ortalaması alınan veriyi veya ham matrisi kullanan bir modeldir. Bu modeli kullanırken algoritma, mümkün olduğunca orijinal veriye uyan hipotetik bir öklidyen nesne uzayı üretmektedir (Giguere, 2006:27-38). Tekrarlı ÇBÖ, benzerlik verisinin birden fazla matrisinin eş zamanlı analizine olanak tanımaktadır (Young, 2007). Matrisler farklı bireylerden veya çoklu zamanlarda tek bir bireyden elde edilebilir. Tek bir ölçüm sonucuna ulaşılır. Ağırlıklı ÇBÖ ise bireysel farklar ölçeği olarak bilinir. Analizin bu tipinde SPSS algoritmaları hem nesne uzayı hem de modeller,

her bir katılımcının ortak nesne uzayındaki boyutlar için ağırlıklı farkları gösteren bir katılımcı uzayı sunmaktadır (Giguere, 2006:27-38). Tüm bu bilgiler ışığında, ÇBÖ modellerini, sahip oldukları karşılaştırmalı özellikleri bakımından aşağıdaki şekilde gruplandırmak mümkündür.

Tablo 2. Çok Boyutlu Ölçekleme Analizinde Kullanılan Modeller (Kaynak: <http://www.acarindex.com/dosyalar/makale/acarindex-1423868778.pdf>, 2008)

	Metrik ÇBÖ	Metrik Olmayan ÇBÖ	Özellik
Klasik ÇBÖ	Aralık, Oran	Nominal, Ordinal	Veri Düzeyi
	Tek	Tek	Matris Sayısı
	Öklidyen $\rightarrow l\{S\} = D + E$	Minkowski $\rightarrow m\{S\} = D + E$	Uzaklık Modeli*
	Ağırlıksız	Ağırlıksız	Model Tipi
Tekrarlı ÇBÖ	Aralık, Oran	Nominal, Ordinal	Veri Düzeyi
	Birkaç	Birkaç	Matris Sayısı
	Öklidyen $\rightarrow l_k\{S_k\} = D + E_k$	Minkowski $\rightarrow m_k\{S_k\} = D + E_k$	Uzaklık Modeli*
	Ağırlıksız	Ağırlıksız	Model Tipi
Ağırlıklı ÇBÖ	Aralık, Oran	Nominal, Ordinal	Veri Düzeyi
	Birkaç	Birkaç	Matris Sayısı
	Öklidyen $\rightarrow l_k\{S_k\} = D_k + E_k$	Minkowski $\rightarrow m_k\{S_k\} = D_k + E_k$	Uzaklık Modeli*
	Ağırlıklı	Ağırlıklı	Model Tipi

\* $l\{S\}$  = benzerliklerin doğrusal transformasyonu, D=uzaklıklar, E=harita matrisi

\* $m\{S\}$  = benzerliklerin monotik transformasyonu

\* $l_k\{S_k\}$  = 3 uzaklıklarına en iyi uyan k. Benzerlik matrisi  $S_k$ 'nın doğrusal transformasyonu

\* $m_k\{S_k\}$  = Matris D'deki uzaklıklara uyan en küçük karelerin , benzerlik matrisi  $S_k$ 'nın , monotik transformasyonu

\* $D_k$  = Tekrarlı ÇBÖ tek bir uzaklık matrisi (D) oluştururken , Ağırlıklı ÇBÖ her bir veri seti  $S_k$  için bir tane olmak üzere m tane uzaklık matrisi  $D_k$  oluşturur.

Sıklıkla kullanılan 5 uzaklık ölçüsü şunlardır: Öklidyen (Euclidean) uzaklığı, Mahalanobis uzaklığı, Minkowski uzaklığı, Block uzaklığı (City Block veya Manhattan olarak bilinir), Chebyshev uzaklığı (www.itl.nist.gov, 2007).

- 1) Öklid uzaklığı, iki nokta arasındaki doğrusal uzaklıktır. Her bir değişkenin farkının karesinin ve bu kareli farkların toplamının karekökünün alınmasıyla hesaplanan bir geometrik uzaklık ölçüsüdür. (Giguere, 2006: 27-38).
- 2) Mahalanobis uzaklığı, değer vektörü ile ortalamaların farkının, kovaryans matrisinin tersi ve yine değer vektörü ile ortalamaların farkının tersyüzünün (transpose) çarpımlarının karekökü olarak hesaplanır.
- 3) Chebyshev uzaklığı, iki nesne arasındaki mutlak maksimum uzaklığa eşittir.
- 4) City Block (Manhattan) uzaklığı boyutlar arasındaki ortalama farka eşittir. Bu ölçüt kullanıldığında farkın karesi alınmadığı için sıradışlıkların etkisi azalır.

- 5) Minkowski mesafesi Öklid uzayı'nda bir metrik'tir. P=2 olduğu zaman Öklidyen mesafesi ve p=1 olduğu zaman Manhattan mesafesi'nin genelleştirilmesi ile oluşturulur.

Tablo 3. Sıklıkla Kullanılan 5 Uzaklık Ölçüsü Formülleri (Kaynak: <http://www.acarindex.com/dosyalar/makale/acarindex-1423868778.pdf>, 2008)

Öklidyen (Euklidean) Uzaklığı	$d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$
Mahalanobis Uzaklığı	$d = \sqrt{(\bar{x} - \bar{y})^T P^{-1} (\bar{x} - \bar{y})}$
Minkowski Uzaklığı	$d = \left[ \sum_{i=1}^n  x_i - y_i ^p \right]^{\frac{1}{p}}$
Block Uzaklığı	$d = \sum_{i=1}^n  x_i - y_i $
Chebyshev Uzaklığı	$d = \text{Max}  x_i - y_i $

### Uygulama

Uyumun veya uyum iyiliğinin bir ölçüsü olan stress (gerginlik) ölçüsü, ÇBÖ analizinde geniş bir kullanıma sahiptir ve analiz sonucu elde edilen grafiksel düzenlemede kullanılan boyut sayısının uygun olup olmadığının belirlenmesinde önemli bir ölçüt olarak karşımıza çıkmaktadır (Filiz ve Çemrek, 2005). Stress ölçüsü;

$$\text{STRESS} = \left[ \frac{\sum (d_{ij} - \hat{d}_{ij})^2}{\sum \hat{d}_{ij}} \right]^{1/2}$$

$\hat{d}_{ij}$  = i. ve j. bireyler arasındaki veri uzaklığı,

$D_{ij}$  = i. ve j. bireyler arasındaki konfigrasyon uzaklığı

olarak yansıtılmaktadır (Doğan, 2003:33-37). Stress oranı, ÇBÖ çözümünün uygunluğuna karar vermede bir kriter olarak kullanılmaktadır. Kruskal 1964 yılında gerginlik değerinin yorumlanması için tablodaki gibi çözümün uygunluğunu yansıtan bir rehber sunmuştur (Wickelmaier, 2003):

Tablo 4: Stress değer aralıkları ve uygunluk durumu

STRESS	UYGUNLUK
$\geq 0,20$	Uyumsuz gösterim
0,1	Düşük uyum
0,05	İyi uyum
0,025	Mükemmel uyum
0	Tam uyum



Çalışmanın amacı ÇBÖ Analizi ile ürünlerin satış adetlerinin iki boyutlu bir uzayda gösteriminin elde edilmesidir. Böylece, birimler ve değişkenler arasındaki ilişkiler belirlenmeye çalışılarak, Türkiye’de satış adetlerinin hangi bölge, marka ve ürün gruplarında benzerlik ve farklılık gösterdiği anlaşılmasına çalışılacaktır. ÇBÖ Analizleri SPSS 25 programı ile yapılmıştır ve klasik çok boyutlu ölçekleme yöntemi kullanılmıştır.

Sektörler için bölge bazlı hesaplanan boyutlara ilişkin değerler Tablo 5’teki gibi elde edilmiştir.

Tablo 5. Bölgeler için Boyut Değerleri

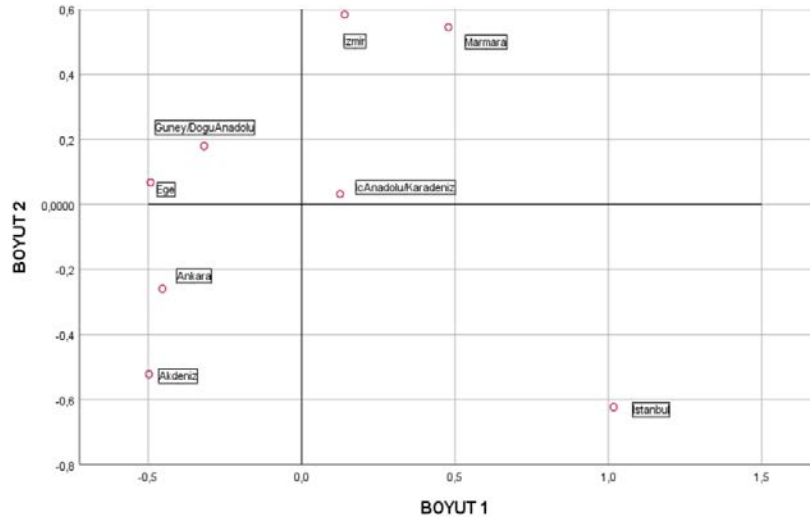
BİLGİSAYAR VE TEKNOLOJİLERİ				TELEKOM VE TEKNOLOJİLERİ			
Sıra No	Bölge	Boyutlar		Sıra No	Bölge	Boyutlar	
		1	2			1	2
1	Akdeniz-Adana+4İL	-,666	-,366	1	Akdeniz-Adana+4İL	-,627	-,403
2	Akdeniz-Antalya/Isparta/Burdur	-,609	-,170	2	Akdeniz-Antalya/Isparta/Burdur	-,659	-,121
3	Ankara	-,277	-,229	3	Ankara	-,291	-,217
4	Ege	-,320	,064	4	Ege	-,351	,092
5	Güney/Dogu Anadolu	-,192	,192	5	Güney/Dogu Anadolu	-,194	,194
6	İç Anadolu/Karadeniz	,143	,132	6	İç Anadolu/Karadeniz	,240	,046
7	İstanbul	1,151	-,681	7	İstanbul	1,119	-,656
8	İzmir	,264	,514	8	İzmir	,218	,559
9	Marmara	,507	,543	9	Marmara	,546	,505

Bu sonuçların güvenilirliği için analiz sonucu stress değerleri Tablo 6’daki gibidir. Her iki değer de 0,025’ten küçük ve 0’a yakın oldukları için tam uyum söz konusudur diyebiliriz.

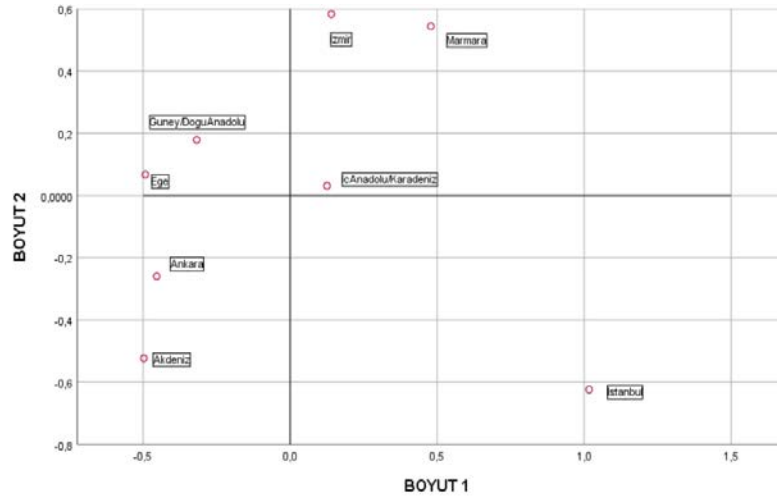
Tablo 6. Bölge Bazlı Stress Değerleri

Bilgisayar ve teknolojileri için bölge stress değeri	,00014
Telekom ve teknolojileri için bölge stress değeri	,00015

Boyut değerleri sonucu her iki sektör içinde elde edilen koordinat düzlemi Şekil 3 ve 4’teki gibidir.



Şekil 3. Bilgisayar Ürün Grupları Bölgelerinin İki Boyutlu Uzayda Gösterimi



Şekil 4. Telekom Ürün Grupları Bölgelerinin İki Boyutlu Uzayda Gösterimi

Her iki sektör için benzer koordinat düzlemi çıktıkları elde edilmiştir. Merkezden uzaklaştıkça farklılıklar artarken, merkeze yakınlaştıkça benzerlikler artmaktadır. Elde edilen koordinat düzleminde İstanbul, Marmara ve Akdeniz bölgeleri merkezden oldukça uzakta konumlanmışlardır. Bu konumlanmanın sebebi diğer bölgelerden farklı olarak daha fazla satış hacmine sahip olmalarıdır şeklinde yorum yapabiliriz. Benzer şekilde Güney/Doğu Anadolu, İç Anadolu/Karadeniz ve Ege bölgeleri merkeze oldukça yakın konumlanmışlardır. İdeal konum hakkında yorum yapacak olursak; ideal konumları İç Anadolu/Karadeniz, Ege ve Güney/Doğu Anadolu olarak düşünebiliriz. İstanbul, Akdeniz ve Marmara bu bölgelere göre daha konumlanma açısından daha riskli bölgelerdir diyebiliriz

Sektörler için ürün bazlı belirlenen boyutlara ilişkin değerler Tablo 7’teki gibi elde edilmiştir.

Tablo 7. Ürünler için Boyut Değerleri

BİLGİSAYAR VE TEKNOLOJİLERİ			
Sıra No	ÜRÜN	Boyutlar	
		1	2
1	Masaüstü Bilgisayar	-,920	-,088
2	Klavye	-,363	-,220
3	Mediatablet	-,085	-,324
4	Mobil Bilgisayar	,084	-,300
5	Monitör	-,197	,237
6	Mouse	,805	-,535
7	Depolama ürünleri	,502	,309
8	Kamera	,173	,921

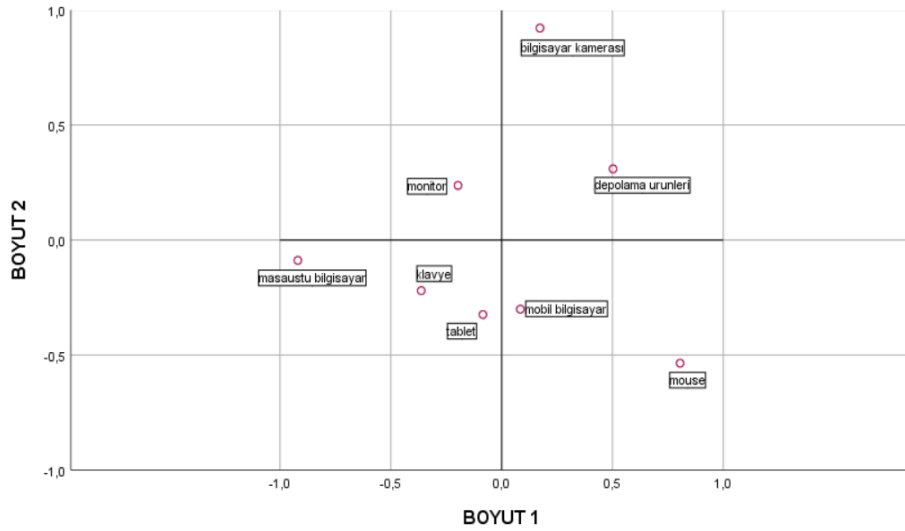
TELEKOM VE TEKNOLOJİLERİ			
Sıra No	ÜRÜN	Boyutlar	
		1	2
1	Giyilebilir Teknoloji	,711	,066
2	Kulaklık	,207	,176
3	Phablet	-,689	,303
4	Akıllı Telefon	-,229	-,545

Bu sonuçların güvenilirliği için analiz sonucu stress değerleri Tablo 8’deki gibidir. Her iki değer de 0,025’ten küçük ve 0’a yakın oldukları için tam uyum söz konusudur şeklinde yorum yapabiliriz.

Tablo 8. Ürün Bazlı Stress Değerleri

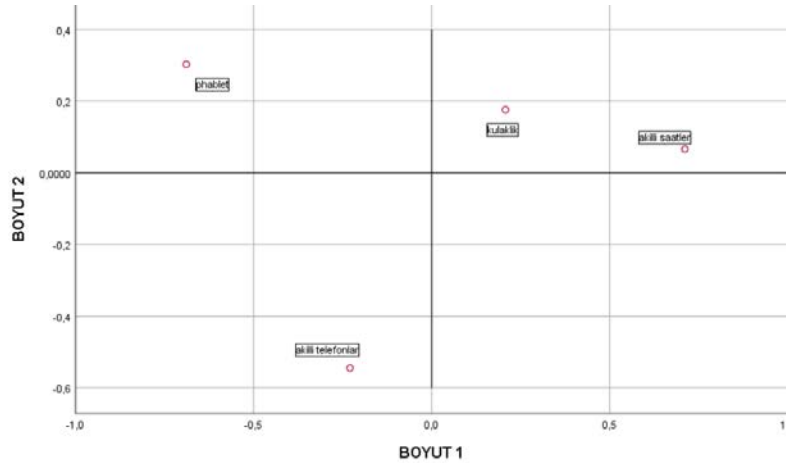
Bilgisayar ve teknolojileri için ürün grupları stress değeri	0,0002
Telekom ve teknolojileri için ürün grupları stress değeri	0,00026

Tablolardaki iki boyutlu ürünler için elde edilen koordinatlar, koordinat sistemine yerleştirildiğinde, bilgisayar ürün grupları için ürünler arasındaki ilişkileri iki boyutlu uzayda gösteren Şekil 5 aşağıdaki gibi elde edilmiştir.



Şekil 5. Bilgisayar Ürün Gruplarının İki Boyutlu Uzayda Gösterimi

Şekil 5'te görüldüğü üzere mouse ürün grubu en yüksek, bilgisayar kamerası ürün grubu en düşük satış hacmine sahip olduğu için diğer ürün gruplarından farklılaşarak merkezden oldukça uzakta konumlanmışlardır şeklinde yorum yapabiliriz. Masaüstü bilgisayar, monitör ve klavye benzer satış adetlerine sahip oldukları için birbirlerine yakın konumlanmış olabilirler. Ayrıca bu ürünlerin bu kadar yakın konumlanmasının ve benzer satış adetlerine sahip olmalarının nedeni birbirlerini tamamlayıcı ürün olmaları olabilir. Diğer ürün gurplarına göre aralarındaki ilişkinin daha yüksek olduğu söylenebilir.



Şekil 6. Telekom Ürün Gruplarının İki Boyutlu Uzayda Gösterimi

Şekil 6 da görüldüğü gibi phablet ve akıllı telefon ürün gruplarının kendilerine has bir satış adetleri oldukları için ayrı bölgelerde tek başlarına konumlanmış olabilirler. Phabletin bu şekilde marjinal bir konum edinebilmesinin bir nedeni akıllı telefonların yaptığı tüm işlevi yapabilmesi olabilir.

Sektörler için marka bazlı belirlenen boyutlara ilişkin değerler Tablo 9'daki gibi elde edilmiştir.

Tablo 9. Markalar için Boyut Değerleri

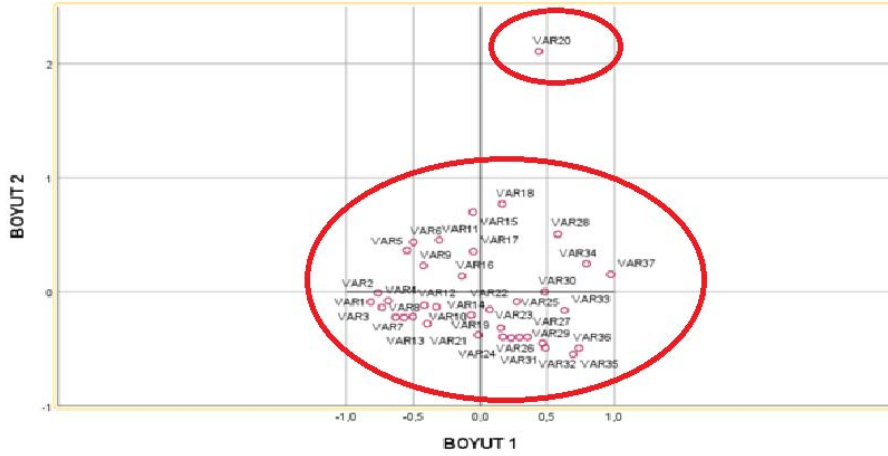
BİLGİSAYAR VE TEKNOLOJİLERİ				BİLGİSAYAR VE TEKNOLOJİLERİ				TELEKOM VE TEKNOLOJİLERİ				TELEKOM VE TEKNOLOJİLERİ			
Sıra No	MARKA	Boyutlar 1		Sıra No	MARKA	Boyutlar 2		Sıra No	MARKA	Boyutlar 1		Sıra No	MARKA	Boyutlar 2	
		1	2			1	2			1	2			1	2
1	MARKA 1	-,818	-,091	23	MARKA 34	,154	-,318	1	MARKA 1	-,911	,196	23	MARKA 28	-,079	-,151
2	MARKA 2	-,764	-,010	24	MARKA 35	,167	-,397	2	MARKA 3	-,844	,189	24	MARKA 29	,034	-,027
3	MARKA 4	-,736	-,139	25	MARKA 36	,272	-,088	3	MARKA 4	-,787	,236	25	MARKA 30	-,022	-,187
4	MARKA 5	-,688	-,080	26	MARKA 37	,227	-,405	4	MARKA 6	-,068	1,351	26	MARKA 31	,007	-,205
5	MARKA 6	-,548	,362	27	MARKA 39	,290	-,400	5	MARKA 7	-,708	,170	27	MARKA 32	,048	-,201
6	MARKA 7	-,500	,433	28	MARKA 40	,578	,504	6	MARKA 9	-,641	,155	28	MARKA 33	,247	,060
7	MARKA 8	-,632	-,225	29	MARKA 41	,353	-,398	7	MARKA 10	-,619	,124	29	MARKA 34	,127	-,202
8	MARKA 10	-,570	-,228	30	MARKA 42	,480	-,002	8	MARKA 11	-,517	,249	30	MARKA 35	,121	-,276
9	MARKA 11	-,425	,227	31	MARKA 45	,464	-,452	9	MARKA 13	-,527	,084	31	MARKA 36	,177	-,251
10	MARKA 12	-,505	-,220	32	MARKA 46	,486	-,493	10	MARKA 14	-,496	,069	32	MARKA 37	,178	-,313
11	MARKA 13	-,307	,452	33	MARKA 48	,628	-,165	11	MARKA 15	-,367	,240	33	MARKA 38	,312	-,165
12	MARKA 14	-,417	-,119	34	MARKA 49	,791	,244	12	MARKA 16	-,436	,040	34	MARKA 39	,234	-,350
13	MARKA 16	-,396	-,281	35	MARKA 53	,693	-,549	13	MARKA 18	-,341	,072	35	MARKA 40	1,857	1,525
14	MARKA 17	-,328	-,135	36	MARKA 54	,736	-,495	14	MARKA 19	-,344	-,003	36	MARKA 43	,352	-,415
15	MARKA 19	-,055	,696	37	MARKA 55	,974	,151	15	MARKA 20	-,315	-,020	37	MARKA 44	,452	-,326
16	MARKA 21	-,138	,136					16	MARKA 21	,503	1,478	38	MARKA 45	,524	-,278
17	MARKA 22	-,051	,352					17	MARKA 22	-,255	-,051	39	MARKA 47	,443	-,521
18	MARKA 25	,163	,768					18	MARKA 23	-,183	,008	40	MARKA 49	,549	-,485
19	MARKA 26	-,067	-,206					19	MARKA 24	,226	,643	41	MARKA 50	,632	-,425
20	MARKA 27	,436	2,106					20	MARKA 25	-,163	-,093	42	MARKA 51	,560	-,586
21	MARKA 29	-,016	-,380					21	MARKA 26	-,136	-,115	43	MARKA 52	,575	-,623
22	MARKA 30	,069	-,158					22	MARKA 27	-,083	-,091	44	MARKA 53	,684	-,529

Bu sonuçların güvenilirliği için SPSS analizi sonucu stress değerleri Tablo 10'daki gibidir. Her iki değer de 0,025'ten küçük ve 0'a yakın oldukları için tam uyum söz konusudur diyebiliriz.

Tablo 10. Marka Bazlı Stress Değerleri

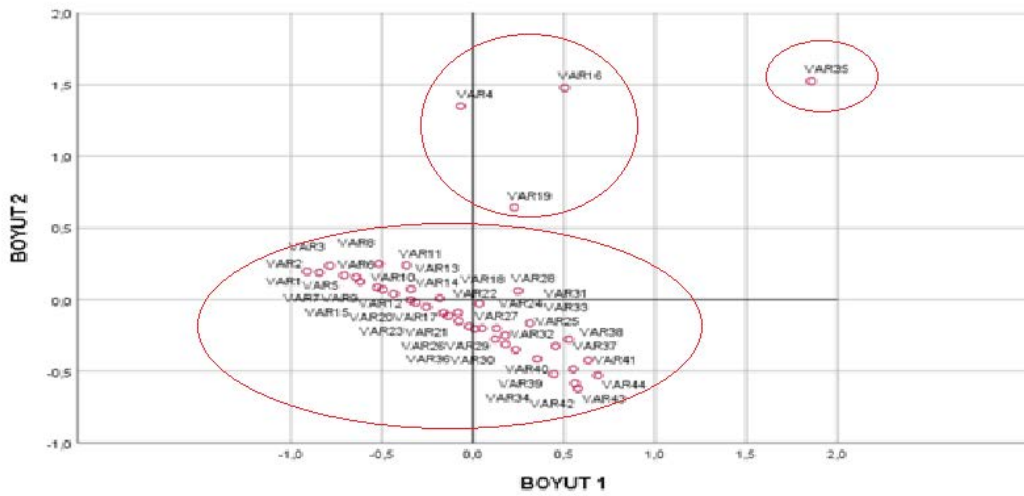
Bilgisayar ve teknolojileri için marka stress değeri	0,00016
Telekom ve teknolojileri için marka stress değeri	0,0006

Markalar için elde edilen koordinat düzlemleri Şekil 7 ve 8'deki gibidir.



Şekil 7. Bilgisayar Ürün Grupları Satan Markaların İki Boyutlu Uzayda Gösterimi

Şekil 7'de görüldüğü üzere markaları iki grup olarak düşünecek olursak, VAR 20 olarak belirtilen marka diğer tüm markalardan ve merkezden uzakta konumlanmıştır. Veri setindeki satış adetlerinden yola çıkarak bu marka en yüksek satış hacmine sahip olduğunu ve pazarda ciddi bir payı olduğunu söyleyebiliriz. Merkezin etrafında konumlanmış olan markaların oluşturduğu kümedeki markalar ise daha benzer satış adetlerine sahiptirler. Ve aynı zamanda birbirine rakip markalardır şeklinde yorum yapabiliriz.



Şekil 8. Telekom Ürün Grupları Satan Markaların İki Boyutlu Uzayda Gösterimi

Şekil 8'e göre koordinat düzleminde elde edilen çıktıyı üç gruba ayırabiliriz. 35 numaralı gözlemi temsil eden marka telekom teknolojileri sektöründe diğer tüm markalardan farklı olarak pazarda ciddi bir payı olduğu söylenebilir. 4, 16 ve 19. gözlemlerin temsil ettikleri markalar diğer gruptaki markalardan farklı olarak gruplanarak 35 numaralı markaya rakip olma eğiliminde olduğu düşünülebilir. Diğer tüm markalar benzer satış hacimlerine sahip oldukları için merkezin etrafında, birbirlerine yakın gruplanmışlardır.

### Sonuç ve Öneriler

Yapılan analiz sonucunda ürün gruplarının satış adetlerinde farklılık olduğu ve bu farklılığın, bölge, ürün ve markaya göre değiştiği gözlemlenmiştir. Koordinat düzlemi olarak elde ettiğimiz çıktılarda,

- ✓ Birbirini tamamlayan ürün gruplarının aynı boyutta, aynı işleve sahip rakip olarak düşünülebilecek ürün gruplarının farklı boyutlarda konumlanabildiğini,
- ✓ Benzer satış hacimlerine sahip nesnelerin daha yakın, farklı satış hacimlerine sahip noktaların daha uzak konumlandığını gözlemleyebildik.

2019 yılı satış verilerine telekom sektörünün pazar payı, bilgisayar sektöründen oldukça fazladır. Günlük hayatta bilgisayarlar ile yapılan işlemlerin telefonlardaki teknolojik gelişme ile daha kolay ve hızlı yapılabilmesi bu sektörün önem kazanmasının bir sebebi olabilir. Bu noktada pazarda var olan bir markanın farklılık yaratmak için her iki sektörde de faaliyet gösteriyor ise telekom sektörüne karşı phablet ve akıllı telefon satışlarında stratejiler geliştirmesi önerilebilir. Sadece bilgisayar sektöründe var olan markalar ise mobil ürünler üzerine stratejiler gerçekleştirip, bu mobil bilgisayar üzerine yenilikler düşünebilirler. Aynı şekilde bu sektörde en çok satılan mouse ürün grupları da sektörün önde gelen bir ürünüdür. Mouse ürün grubunda yapılabilecek ve tüketici tarafından beğenilecek bir yenilik satışlarda fark yaratmaya yardımcı olabilir. Bölgesel olarak, İstanbul diğer illere göre farklı bir konuma sahip olduğu için bu bölgede konumlanın diğer bölgelere göre daha riskli olduğu söylenebilir. İç anadolu/Karadeniz, Güney/Doğu Analou ve Ege bölgeleri en ideal konumlardır şeklinde yorum yapabiliriz.

Ayrıca Çok Boyutlu Ölçekleme analizi verideki nesnelere konumlanma için görsel anlamda önemli bir analizdir. Farklı istatistiksel analizler ile ürünler arasındaki ilişki daha da güçlendirilebilir.

### Kaynaklar

- Alpar R., 2011. Çok Değişkenli İstatistiksel Yöntemler.  
Borg I., Groenen P., 1997. Modern Multidimensional Scaling.  
Doğan, İ., 2003. Kuzularda büyümenin çok boyutlu ölçekleme yöntemi ile değerlendirilmesi, Uludağ Üniversitesi Dergisi.  
Dunder E., 2011. Çok Boyutlu Ölçekleme Analizi ve SPSS, Rproject Uygulamaları, <http://emredunder.blogspot.com/2011/07/cok-boyutlu-olcekleme-analizi-ve-spss-r.html>  
Filiz, Z., Çemrek, F., 2005. Avrupa Birliği'ne üye ülkeler ile Türkiye'nin karşılaştırılması, 7. Ulusal Ekonometri ve İstatistik Sempozyumu, İstanbul.  
<http://www.acarindex.com/dosyalar/makale/acarindex-1423868778.pdf>  
Johnson RA., Wichern DW., 1999. Applied Multivariate Statistical Analysis.  
Özdamar K., 2004. SPSS Uygulamalı Temel İstatistik.  
Perret JK., 2018. SPSS – Multidimensional Scaling, [https://www.youtube.com/watch?v=\\_CiDwyYYDSU](https://www.youtube.com/watch?v=_CiDwyYYDSU)  
Sharma R., 2016. How to perform Multidimensional scaling in SPSS?- Therahul RTV, <https://www.youtube.com/watch?v=7ZCFZOm14c0>  
Tatlıdil H., Uygulamalı Çok Değişkenli İstatistiksel Analiz, Ankara, 1992.  
Tüzüntürk S., (2009). Çok Boyutlu Ölçekleme Analizi: Suç İstatistikleri Üzerine Bir Uygulama, Uludağ Üniversitesi.

- Usrataahmetođlu E., 2005. Marka Konumlandırma da Çok boyutlu Ölçekleme Tekniđinin Kullanılması.
- Wilkermaier, F., 2003. An introduction to MDS, Sound Quality Research Unit, 2003.  
www.itl.nist.gov, 2007
- Yenidođan TG., 2008. Pazarlama Arařtırmalarında Çok Boyutlu Ölçekleme Analizi: Üniversite Öğrencilerinin Marka Algısı Üzerine Bir Arařtırma, Akdeniz Üniversitesi.
- Young W.Y. 2007. Multidimensional scaling, <http://forrest.psych.unc.edu/teaching/p208a/mds.html>,  
Eriřim Tarihi: 30 Mayıs 2007.

## A New Unit Probability Distribution Approach with the Lindley Distribution Example

Kadir KARAKAYA<sup>1</sup>, Selim GÜNDÜZ<sup>2</sup>, Christophe CHESNEAU<sup>3</sup>, Mustafa Ç. KORKMAZ<sup>4</sup>

<sup>1</sup>Department of Statistics, Selçuk University, 42250, Konya, Turkey

<sup>2</sup>Department of Business, Adana Alparslan Türkeş Science and Technology University,  
01250, Adana, Turkey

<sup>3</sup>LMNO, University of Caen-Normandie, Caen, 14032, France

<sup>4</sup>Department of Measurement and Evaluation, Artvin Çoruh University, City Campus,  
Artvin, 08000, Turkey

### Abstract

In this study, a new one-parameter distribution called unit-Lindley distribution is introduced. Some statistical properties of the introduced distribution are examined. The maximum likelihood estimation method is used to estimate the model parameter. An extensive Monte Carlo simulation study is carried out to interpret the mean squared error and bias of the maximum likelihood estimate. Also, real data application is considered to demonstrate the usefulness of new distribution.

**Key words:** Data analysis, maximum likelihood estimate, Monte Carlo simulation, unit distributions



## Yaşam Standartları Faktörünün Gelişmemişlik Faktörü Üzerinden Doğuştan Yaşam Beklentisine Etkisi: Yapısal Bir Model

Firdevs Berrak CABİ<sup>1</sup>, Taner TUNÇ<sup>2</sup>

<sup>1</sup>Ondokuz Mayıs Üniversitesi Lisansüstü Eğitim Enstitüsü, Samsun, Türkiye

<sup>2</sup>Ondokuz Mayıs Üniversitesi İstatistik Bölümü, Samsun, Türkiye

### Özet

Doğuştan yaşam beklentisi, belirli bir yaşa ulaşan kişilerin henüz yaşamadıkları ortalama yıl sayısıdır. Beklenen yaşam süresine etki eden birçok faktör olmakla birlikte bu etkilerin incelenmesinde kullanılan istatistiksel yöntemler bulunmaktadır. Bu çalışmada, regresyonel ilişkilerin var olduğu değişkenler ve değişkenlerin oluşturduğu faktörler arasında değişkenlerin hatalarını da modele katan yapısal eşitlik modellemesi kullanılmak suretiyle doğuştan yaşam beklentisi üzerine etki eden yapılar ortaya çıkarılmıştır. Yapısal eşitlik modellemesi, karmaşık veri yapılarındaki değişkenlerin birbirlerine etkisini inceleyerek bir model oluşturulmasında ve oluşturulan bu modelin uygun olup olmadığını test edilmesinde kullanılan çok değişkenli istatistiksel bir yöntemdir. Modelin uyumluluğuna dair sonuçlarının yorumlanmasına olanak sağlayan uyum iyiliği testlerini içinde barındırarak oluşabilecek hataları ortadan kaldırır ve en iyi modelin seçilmesinde önemli bir araçtır. Yapısal eşitlik modellemesinde modele ait değişkenlerdeki dolaylı ve doğrudan ilişkiler ortaya konulabilmektedir. Dolaylı ilişkilerin ortaya çıkarılmasında, aracı değişkenin istatistiksel olarak anlamlılık ölçülmesi için sobel testi kullanılmaktadır. Bu çalışmada gelişmemişlik ve yaşam standartları faktörlerinin doğuştan beklenen yaşam süresi üzerine etkisini inceleyen, uyum iyiliği ölçütlerine göre ilişkileri en iyi açıklayan model oluşturulmuştur. Modelde, gelişmemişlik faktörünün aracı değişken görevi üstlendiği gözlemlenmiştir. Buna göre, yaşam standartları faktörünün beklenen yaşam süresi üzerinde direkt etkisi bulunmamaktadır; fakat gelişmemişlik faktörü üzerinden dolaylı bir etkisi mevcuttur. Oluşturulan yapısal denklem için ise, tahmin edilen tüm parametrelere ait t değerleri istatistiksel olarak anlamlı ve açıklama katsayısı %71 ile yeterlidir.

**Anahtar kelimeler:** Aracı değişken, doğuştan yaşam beklentisi, Sobel testi, uyum iyiliği ölçütleri, yapısal eşitlik modellemesi

## The Effect of The Living Standards Factor on the Life Expectancy at Birth Through the Undevelopment Factor: A Structural Model

### Abstract

Life expectancy at birth is average number of years that people who have reached a certain age have not yet lived. There are many factors that affect life expectancy, statistical methods can be used to examine these effects. In this study, by using structural equation modeling, the structures that affect the life expectancy at birth have been revealed. Structural equation modeling is a multivariate statistical method used to create a model by examining the effects of variables in complex data structures on each other

and to test whether this model is compatible. It includes goodness of fit tests that allow the interpretation of the results of the model's compatibility, eliminating the errors that may occur, and is an important tool in selecting the best model. Indirect and direct relationships in the variables of the model can be revealed in structural equation modeling. The sobel test is used for statistical significance measurement of the mediator variable in revealing indirect relationships. In this study, a model that examines the effects of undevelopment and living standards factors on life expectancy at birth and explains the relationships best according to the criteria of goodness of fit was created. The effect of the living standards factor on the undevelopment factor was examined and a statistically significant structural model was established. It has been observed that the undevelopment factor acts as the mediator variable. Accordingly, the factor of living standards does not have a direct effect on life expectancy; but it has an indirect effect on the undevelopment factor. For the structural equation created, t values of all estimated parameters are statistically significant and explanation coefficients of 71% are sufficient.

**Key words:** Life expectancy at birth, mediator variable, structural equality modeling, the goodness of fit Criteria, Sobel test

## Giriş

Doğuştan yaşam beklentisi, yeni doğan bireylerin mevcut ölüm oranları ile yaşaması beklenen ortalama yıl sayısıdır. Hayat tablolarından yeterli ölüm verilerinin olduğu durumlarda hesaplanabilir. Nüfus ve ölüm verilerine ulaşılabilirlik yok ise, istatistiksel yöntemlere başvurularak da hesaplama yapılabilmektedir (Mahdian vd, 2016). Doğuştan beklenen yaşam süresi tahminleri, düşük ve yüksek ölüm oranlarından oldukça etkilenmektedir (Preston vd, 2001). Bebek ölüm oranları cinsiyete göre farklılık göstermekle beraber, fazla olması o ülkenin sağlık ve kaynak eşitsizliği durumu hakkında bilgi vermektedir. Kız bebekler, erkek bebeklere göre doğumda daha dayanıklıdır. Eğer bir ülkedeki kız bebek ölüm oranı fazlaysa bu durum o ülkedeki kaynak eşitsizliğini ifade etmektedir (World Bank, 2012). Bununla beraber, sağlığı etkileyen gelir ve gıda gibi öğelerin tüketim oranı üzerindeki etkisi ölüm oranı üzerinde de dolaylı bir etkisi bulunmaktadır (Hertz vd, 1994). Beklenen yaşam süresi, bir ülkenin sosyoekonomik kalkınma derecesi ile de yakından ilişkilidir (Delavari vd, 2016).

Bu araştırmada, insan yaşamını etkilediği düşünülen değişkenler bir araya getirilerek ilişki düzeyleri incelenmiş ve anlamlı bir yapısal model oluşturulmuştur. Yapılan literatür taramasında ise beklenen yaşam süresine etki eden değişkenlerin incelendiği, yapısal eşitlik modellemesi yöntemi kullanılarak yapılmış benzer herhangi bir çalışmaya rastlanılmamıştır.

## Materyal ve Metot

Çalışmada kullanılan 2015 yılına ait 21 değişkenli veri seti, Dünya Bankası veri tabanında yer alan “World Development Indicators” sekmesinde verinin ulaşılabilirliğine göre doğuştan beklenen yaşam süresinin genellikle yüksek olduğu ülke rakamlarından oluşmaktadır. Veri setindeki bazı değişkenlerde kayıp veri fazlalığı olduğu görülmüş; tahmin edilecek parametre değerlerinde sapmalara yol açacağından hareketle ve modelin oluşturulmasında göz ardı edilmiş ve kullanılmamıştır (Kline, 2015). Bu şartlar altında, eksik veri bulunmayan 13 değişken ve 78 ülkeden oluşan veri setiyle istatistiksel olarak anlamlı bir modelleme yapılmıştır.

Doğuştan beklenen yaşam süresi hesaplanan yılın öncesindeki yıllara ait değerlerden oldukça etkilenmektedir. Yapılan çalışmada kullanılacak verinin ardıl olarak oluşturulması ve analizde kullanılması daha sağlıklı sonuçlar elde etmeye imkan sağlamaktadır (Shaw vd., 2005). Bu duruma göre; modelde yer alacak doğuştan beklenen yaşam süresi değişkeni hesaplandığı yıldan önceki son beş yılı kapsayacak şekilde (2010, 2011, 2012, 2013, 2014 yılları dahil) kullanılmıştır.

Değişkenlerin doğuştan yaşam beklentisi üzerindeki etkilerini incelemek amacıyla kullanılan yapısal eşitlik modellemesi, genel olarak tek bir olguyu ölçmek için geliştirilmiş ölçekler yardımıyla kesikli değişkenlerden oluşan veri setlerinde kullanılmaktadır. Bu çalışmada ise, veri seti sürekli değişkenlerden oluştuğundan, yapılan analizlerdeki hatayı ve oluşan çarpıklığı azaltmak için özgün düzenlemelerde bulunulmuştur. Öncelikle, verinin ortalama ve varyansı değişmeyecek şekilde, kayıp verinin yeri bulunduğu değişkenin kendi ortalaması ile tamamlanmıştır. İkinci adımda çarpıklığı önlemek amacıyla veri standart normal dağılıma uygun hale getirilmiş ve standartlaştırılmıştır. Son olarak, elde edilen yeni değişkenlerin ortalamalarından birim standart sapma uzaklıklarına göre değişken değerleri buldukları bölgelere göre beşli likert ölçek haline dönüştürülmüştür. Dönüştürülmüş değişkenler yoluyla da yapısal eşitlik modeli parametreleri tahmin edilmiştir. Modelde bulunan dolaylı etkileri ortaya çıkarmak için aracı değişken analizine başvurulmuştur. Yapısal eşitlik modellemesinde aracı değişkenler, bağımsız değişkenler ele alındığında bağımlı değişken, bağımlı değişkenler ele alındığında ise bağımsız değişken olarak tanımlanmaktadır (Sümer, 2000). Kullanılan aracı değişkenin istatistiksel olarak anlamlılığı ise, Sobel testi ile ölçülmektedir. Aracılık etkisi incelenirken kullanılan üç farklı regresyon denklemi eşitlik (1.1), (1.2) ve (1.3) ile verilmiştir.

$$\text{Model 1: } Y_{dep} = \alpha_1 + \tau X_{Ind} + \varepsilon_1 \quad (1.1)$$

$$\text{Model 2: } X_{med} = \alpha_2 + a X_{Ind} + \varepsilon_2 \quad (1.2)$$

$$\text{Model 3: } Y_{dep} = \alpha_3 + \tau' X_{Ind} + b X_{med} + \varepsilon_3 \quad (1.3)$$

Aracı değişken etkisi ise,  $ab = (\tau - \tau')$  eşitliği tespit edilmektedir (Sobel, 1982; Sobel, 1986). Bu dolaylı etkilerin istatistiksel olarak anlamlılığını test etmek için Sobel testi, Z test istatistiğini kullanarak ölçümün standart hatalarını tahmin eder. Bu sayede, modelde bulunan dolaylı etkilerin büyüklüğü eşitlik (1.4) ile hesaplanır.

$$Z = \frac{(\tau - \tau')}{\text{Standart Hata}} = \frac{(\tau - \tau')}{\sqrt{a^2 \sigma_b^2 + b^2 \sigma_a^2}} \quad (1.4)$$

Eşitlik (1.4)'e göre, a aracı değişken ile bağımsız değişken arasındaki ilişkiyi,  $\sigma_a^2$  ise katsayıların standart hatasını ifade etmektedir. Bağımlı değişken ile aracı değişken arasındaki ilişki ise b ile gösterilmektedir. Bu katsayıların standart hatası ise,  $\sigma_b^2$  ile verilir (Sobel, 1982).

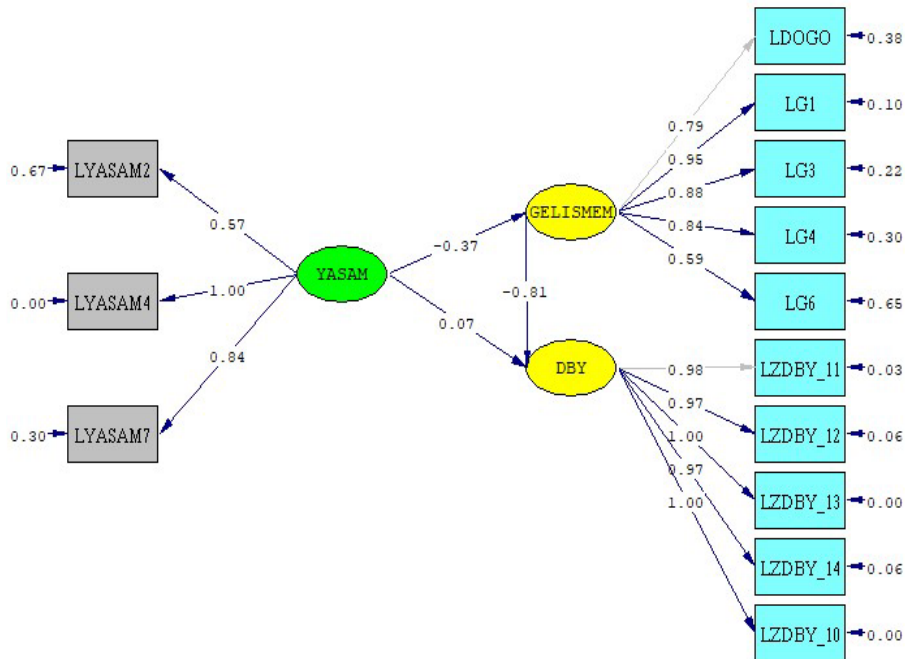
## Bulgular ve Tartışma

Bu çalışmada, doğuştan yaşam beklentisini etkileyen değişkenler ile 2015 yılına ait yapısal bir model oluşturulmuştur. Önerilecek olan yapısal modelin oluşturulması için, açıklayıcı faktör analizi ile faktör yapısı oluşturulmuştur. Bu faktörleşme, 3 boyutlu yapı, doğrulayıcı faktör analizi uygulanarak doğrulanmıştır. Oluşan yapı içinde doğuştan beklenen yaşam süresi bağımlı değişken kabul edilerek bu değişkene etki eden 2 faktörün varlığı ortaya konmuştur. Bu faktörler; gelişmemişlik ve yaşam standartları olarak adlandırılmıştır. Gizil değişkenler altında tanımlanan gözlenen değişkenlerin ilgili oldukları yapıları kapsayıp kapsamadığını değerlendirmek için gizil faktörlere ait güvenilirlik ve açıklanan varyans ölçütleri hesaplanmıştır. Doğuştan beklenen yaşam süresi faktörünün yapı güvenilirliği 0,99; Gelişmemişlik faktörünün yapı güvenilirliği 0,90 ve yaşam standartları faktörünün yapı güvenilirliği 0,85 olarak elde edilmiştir. Ayrıca faktörlerin tek boyut altında açıklanan varyansları ise sırasıyla 0,96; 0,67; 0,67 olarak bulunmuştur. Faktörlerin yapı geçerliliklerinin 0,70'den; açıklanan varyansların ise 0,50'den yüksek olması önerilen modelin anlamlı ve önemli olduğunun göstergesidir (Nunnally, 1978; Hair vd., 1998). Önerilen modelin istatistiksel olarak anlamlılığına ait tüm ölçütler Tablo 1 ile verilmiştir.

Tablo 1. 2015 Yılına Ait Doğrulayıcı Faktör Analizi Sonuçları

Faktörler	Standart Yükler	t-değeri	Yapı Güvenirliği	R <sup>2</sup>	Açıklanan Varyans
<b>DBY</b>					
LZDBY_10	1,00	48,79	0,99	1,00	0,96
LZDBY_11	0,98	6,11		0,97	
LZDBY_12	0,97	47,26		0,94	
LZDBY_13	1,00	48,79		1,00	
LZDBY_14	0,97	28,60		0,94	
<b>GELISMEMIS</b>					
LG1	0,95	9,29	0,90	0,90	0,67
LG3	0,88	8,35		0,78	
LG4	0,84	14,35		0,70	
LG6	0,59	5,38		0,35	
LDOGO	0,79	5,28		0,62	
<b>YASAM STANDARDI</b>					
LYASAM2	0,57	5,50	0,85	0,33	0,67
LYASAM4	1,00	12,41		1,00	
LYASAM7	0,84	9,11		0,70	

Modelde 2015 yılı için doğuştan beklenen yaşam süresini en çok açıklayan değişkenlerin 2010 ve 2013 yıllarına ait oldukları (LZDBY\_10, LZDBY\_13) görülmüştür. Gelişmemişlik faktörünün en çok etkileyen değişken kız bebek ölüm oranıdır (LG1). Yaşam standartları faktörünü ise en çok kişi başı gayri safi milli gelir (LYASAM4) etkilemektedir. Önerilen yapısal modele ait path diagramı Şekil 1’de verilmiştir.



Chi-Square=75.13, df=59, P-value=0.07670, RMSEA=0.060

Şekil 1. Önerilen Modele Ait Path Diyagramı

Önerilen modele ait uyum ölçütü değerleri ise Tablo 2 ile verilmiştir.

Tablo 2. 2015 Yılına Ait Yapısal Eşitlik Uyum İndeksleri

Uyum Ölçütleri	İyi Uyum	Kabul Edilebilir Uyum	Önerilen Model	Uyum
$\chi^2/sd$	$0 \leq X^2/sd \leq 2$	$2 \leq X^2/sd \leq 3$	1,27	İyi Uyum
$p$ değeri	$0,05 \leq p \leq 1,00$	$0,01 \leq p \leq 0,05$	0,0767	İyi Uyum
RMSEA	$0 < RMSEA \leq 0,05$	$0,05 < RMSEA \leq 0,08$	0,06	Kabul Edilebilir
NFI	$0,95 \leq NFI \leq 1,00$	$0,90 \leq NFI \leq 0,95$	0,96	İyi Uyum
NNFI	$0,95 \leq NNFI \leq 1,00$	$0,90 \leq NNFI \leq 0,95$	0,99	İyi Uyum
CFI	$0,97 \leq CFI \leq 1,00$	$0,95 \leq CFI \leq 0,97$	0,99	İyi Uyum
GFI	$0,90 \leq GFI \leq 1,00$	$0,89 \leq GFI \leq 0,85$	0,87	Kabul Edilebilir

Doğuştan beklenen yaşam süresine ait yapısal modelin açıklanan varyansı ( $R^2$ ), %71 olarak ölçülmüş ve anlamlı olduğu görülmüştür. Ayrıca gelişmemişlik faktörünün aracı değişken görevi üstlendiği tespit edilmiştir. Yaşam standartlarını oluşturan değişkenlerin, doğuştan yaşam beklentisi üzerinde direkt bir etkisi yoktur. Fakat gelişmemişlik faktörü üzerinden dolaylı bir etkisi mevcuttur.

Yapısal denklemlere göre, gelişmemişlik faktöründeki 1 birimlik artış doğuştan beklenen yaşam süresinde 0,80 birimlik bir azalışa neden olmaktadır. Yaşam standartları değişkeni ise gelişmemişlik değişkenindeki değişimin %13'ünü karşılamaktadır. Yaşam standartları değişkenindeki 1 birimlik artış, gelişmemişlik değişkeninde 0,37'lik bir azalışa sebep olmaktadır.

$$DBY = -0,80 * GELISME + 0,072 * YASAM \quad R^2 = 0,71$$

$$S_b \quad (0,10) \quad (0,074)$$

$$t \text{ değerleri } -7,74 \quad 0,98$$

$$GELISMEM = -0,37 * YASAM \quad R^2 = 0,13$$

$$S_b \quad (0,12)$$

$$t \text{ değerleri } -3,07$$

Gelişmemişlik faktörünün doğuştan yaşam beklentisi üzerinde, ters yönlü anlamlı ve direkt etkisi mevcuttur. Barındırdığı değişkenlerin, genel olarak gelişmemiş ülkelerde değerlerinin fazla olduğu görülmektedir. Yaşam standartları faktörü doğuştan beklenen yaşam süresi üzerinde direkt etkisi bulunmasa da, gelişmemişlik faktörü üzerinden dolaylı bir etkisi bulunmaktadır. Bu durumda gelişmemişlik faktöründeki azalma, yaşam standartları ve doğuştan yaşam beklentisini yükseltmektedir. Başka bir ifade ile yaşam standartları faktörü ile doğuştan yaşam beklentisi arasında dolaylı pozitif bir etki bulunmaktadır.

## Kaynaklar

- Delavari S, Zandian H, Rezaei S, Moradinazar M, Delavari S, Saber A, Fallah R. 2016. Life Expectancy And Its Socioeconomic Determinants In Iran. Electronic Physician, 8:10, 3062-3068. DOI: 10.19082/3062
- Hair, J. , Anderson, R.E. And Tapham, R.L. 1998. Multivariate Data Analysis. Prentice Hall, New York.
- Hertz E, Hebert JR, Landon J. 1994. Social And Environmental Factors And Life Expectancy, Infant Mortality And Maternal Mortality Rates: Results Of A Cross-National Comparison. Social Science & Medicine, 39:1, 105-14.

- Kline RB. 2015. Principles And Practice Of Structural Equation Modeling (Fourth Edition). Guilford Press, New York.
- Mahdian M, Sehat M, Fazel RM. 2016. Life Expectancy At Birth İn Aran-Bidgol Region, Iran, 2012: A Study Based On Corrected Health Houses Data. *International Journal Of Epidemiologic Research*, 3:3, 259-267
- Nunnally JC. 1978. Psychometric Theory. Second Edition. Mcgraw-Hill Collage.
- Preston SH, Heuveline P, Guillot M. 2001. Demography, Measuring And Modelling Population Processes. Blackwell Publishers, Oxford.
- Shaw JW, Horrace WC, Vogel RJ. 2005. The Determinants Of Life Expectancy: An Analysis Of The OECD Health Data. *Southern Economic Journal*, 768-783.
- Sobel ME. 1982. Asymptotic Confidence Intervals For İndirect Effects İn Structural Equation Models. *Sociological Methodology*, 13, 290-312. DOI:10.2307/270723.
- Sobel ME. 1986. Some New Results On İndirect Effects And Their Standard Errors İn Covariance Structure. *Sociological Methodology*, 16, 159-186. DOI:10.2307/270922.
- Sümer N. 2000. Yapısal Eşitlik Modelleri: Temel Kavramlar Ve Örnek Uygulamalar. *Türk Psikoloji Yazıları*, 3:6.49-73.
- The World Bank. 2012. World Development İndicators 2012, United States Of America.

**Türkiye’de Ekonomik Büyüme, İşsizlik ve Enflasyonun Kayıtdışı İstihdam Üzerindeki Etkisi: ARDL Sınır Testi Yaklaşımı**

Rümeysa ÇELİK<sup>1</sup>, Abdurrahman KESKİN<sup>1</sup>, Abdulkadir KESKİN<sup>2</sup>

<sup>1</sup>İstanbul Üniversitesi, İktisat Fakültesi, Çalışma Ekonomisi Bölümü, İstanbul, Türkiye

<sup>1</sup>İstanbul Medeniyet Üniversitesi, Siyasal Bilgiler Fakültesi, İşletme Bölümü, Sayısal Yöntemler Anabilim Dalı, İstanbul, Türkiye

**Özet**

Ekonominin küreselleşmesi, birçok sektörde ve ülkede işgücünün kayıt dışı olmasına sebep olmuştur. Bu problemin çözülmesine yönelik politikalar ile önlemler alınmaya çalışılsa da dünya genelinde olduğu gibi Türkiye’de de kayıt dışı istihdam, çözülmesi zor bir problem haline gelmiştir. Kayıt dışı istihdama karşı uygulanan politikaların yanı sıra makroekonomik faktörlerdeki değişimler, kayıt dışı istihdamı etkilemektedir. Bu çalışmanın amacı da, makroekonomik faktörler içerisinde yer alan ekonomik büyüme, işsizlik ve enflasyonun kayıt dışı istihdam ile ilişkisini ortaya koymaktır. Bu amaç doğrultusunda ekonomik büyüme, işsizlik ve enflasyon göstergelerinin, kayıt dışı istihdam üzerindeki uzun dönemli ve kısa dönemli ilişkisi incelenmiştir. Yapılan çalışmada, TCMB ve TÜİK’ten elde edilen 2004-2020 yıllarına ait çeyreklik veriler kullanılmıştır. Kayıt dışı istihdam üzerinde belirlenen faktörlerin uzun ve kısa dönem etkileri ARDL sınır testi kullanılarak incelenmiştir. Ampirik sonuçlar, kayıt dışı istihdam ile işsizlik, ekonomik büyüme ve enflasyon arasında ilişki olduğunu göstermektedir. Uzun dönem katsayılarının tamamı istatistiksel olarak %1, %5 ve %10 düzeylerinde anlamlı olup, ekonomik büyüme, işsizlik ve enflasyon ile kayıtdışı istihdam arasında uzun dönemli bir ilişki görülmektedir. Kısa dönemde ise meydana gelen dengesizliklerin CointEq(-1) %45’i ortadan kalkmaktadır.

**Anahtar kelimeler:** Kayıt dışı istihdam, ekonomik büyüme, işsizlik, enflasyon, ARDL sınır testi

## Uzaktan Eğitim-Öğretim (Online) Sisteminde Etik Sorunlar ve Çözüm Önerileri Üzerine Bir Araştırma ve Etik Yemini

Adnan MAZMANOĞLU<sup>1</sup>, Ramazan MACIT<sup>2</sup>, Nur Selin ÖZEN<sup>3</sup>, Selin SARAÇ<sup>4</sup>

<sup>1</sup>Toros Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Mersin, Türkiye

<sup>2</sup>Toros Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Mersin, Türkiye

<sup>3</sup>Toros Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Mersin, Türkiye

<sup>4</sup>Toros Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Mersin, Türkiye

### Özet

Bu çalışma orta büyüklükteki bir vakıf üniversitesinde, uzaktan eğitimde kendiliğinden beliren akademik etiğe uymayan sınav yapılış şekli ve sınav değerlendirmede karşılaşılan durumlar hakkında aydınlatıcı sonuçlar ortaya çıkarmak için tasarlanmıştır. Buna dayanarak gerek mühendislik fakültesi öğrencileri gerekse herhangi bir üniversitenin bir fakültesinde eğitim- öğretim yapan öğretim elemanlarına rastgele yöntemi ile anket uyguladık. Dijital ortamda ve İstatistiksel Analiz Yöntemleriyle (IBM SPSS Statistics 26) elde edilen verilerin değerlendirilmesinden sonra elde edilen sonuçları tüm online eğitim-öğretimde görev alan herkesle paylaşmayı düşünüyoruz. Bu sonuçlardan da öğrencilerimizin online şekilde verilen derslerin verilmiş, uygulanış ve sınav yöntemlerine karşı duydukları memnuniyetinin ölçülmesinde %50'den yüksek bir oran elde edilmiştir. Denetimsiz sınavlarda karşılaşılan ahlaki olmayan, akademik dürüstlüğe sığmayan (aşırma, intihal, başkasının çalışmasını kopya etmek, ders notu, kitap ve bunun gibi dokümanları kullanmak, konu uzmanlarından yardım almak vb.) davranışlardan kaçınmanın öğrenciler üzerinde ne gibi etkilerin olduğu ve yukarıda sözü edilen etik dışı davranışların öğrenciler tarafından yapılmayacağına dair herhangi bir taahhütte (söz vermek) bulunmayı ankete katılan 103 kişilik bir örnekleme dayanarak kabul edenlerin oranının %79.61, taahhütte bulunmayanların oranı %14.56, kısmen buna onay verenlerin ise %5.83 olduğu saptanmıştır.

**Anahtar kelimeler:** Akademik etik, etik yemini, online ders yöntemi, online sınav yöntemi, online değerlendirme sistemi

### Abstract

This study was designed to reveal enlightening results about the way of taking exams that do not comply with academic ethics and the situations encountered in exam evaluation in a medium-scaled foundation university. Based on this, we randomly applied a questionnaire to both engineering faculty students and lecturers working in a faculty of an university. We intend to share the results obtained after the evaluation of the data obtained in the digital environment and by Statistical Analysis Methods (IBM SPSS Statistics 26) with everyone involved in online education. From these results, a rate higher than 50% was obtained in measuring the satisfaction of our students with the teaching, implementation and exam methods of online courses. Based on the 103 students (sample) who participated in the questionnaire, it was determined that the rate of students who are agree to make any commitment for behaviors encountered in uncontrolled exams, which are immoral and do not fit into academic honesty (plagiarism, copying someone else's work, using lecture notes, books and similar documents, getting



help from subject experts, etc.) that will not be done by the student was 79.61%. The rate of those who were not make a commitment was 14.56%, and those who partially gave consent to this were found to be 5.83%.

**Key words:** Academic ethics, ethical oath, online course method, online evaluation system, online exam method

## Giriş

Bu çalışma, birtakım üniversitelerde kendiliğinden beliren akademik etik dışı davranışlar (akademik sahtekârlık) hakkında bir araştırma yaparak bu konuda dünyanın değişik üniversitelerinde bu konu hakkında yapılan incelemelerle kıyaslama yapmak için düşünülmüştür. Yaptığımız inceleme ve araştırma sonunda bazı sonuçları paylaşmak istiyoruz. Bir Çalışmada; kullanılan yöntemi açıklamadan önce, *Sekiz Etik Kuralları* olarak (Ortalama Öğrenci Sahtekârlık Bildirimi) aşağıdaki gibi oluşturulmuştur; Öğrenciler,

1. Üniversitedeki dürüstlük ve sorumluluk kurallarını desteklemeyi kabul ediyorum. (%34,97)
2. Bir üniversite öğrencisi olarak, bu görev üzerinde yetkisiz bir yardımın ne bana verilmediğini ne de almadığımı kabul ediyorum. Bu ifadeyi kabul ederek, etik yeminini bozmanın sonuçlarının farkında olduğumu ifade ediyorum. (%32,30)
3. Hiçbir üniversite öğrencisi kendi akademik performansını adil olmayan bir şekilde ilerletmeyecek ve üniversite topluluğunun diğer öğrencilerinin akademik takiplerini hiçbir şekilde sınırlamayacak veya engellemeyecektir. (%28,60)
4. Üniversite öğrenci biriminin bir üyesi olarak, akademik takiplerimle ilgili hiçbir emek içinde yalan söylemeyeceğim, kopya çekmeyeceğim veya çalmayacağım ve diğerlerinde bu davranışa müsamaha göstermeyeceğim. Eğer bu ifadeye uymazsam hareketlerimi izleyecek olan sonuçları anlıyorum. Buna, sınav notum olarak 0 almak ve ihraç olasılığıyla akademik sonuçlar dâhildir (%27,75).
5. Her öğrenci, bir üniversite öğrencisi için yakışıksız olan ve öğrenciyi ve/veya Üniversiteyi itibarsızlaştıran davranışlardan etik olarak kaçınmalıdır. Aşağıdakilerin hepsi olacağımı kabul ediyorum:
  - Etik olarak kopya çekmekten kaçınan;
  - Etik olarak hileden kaçınan;
  - Etik olarak aşırı maktan kaçınan;
  - Etik olarak Üniversite Etik Yeminini bozmakla karşılaşmaktan kaçınan.
  - Etik olarak Üniversite Etik Yemininin bozulmasını bildirmekten kaçınan (%30,40).
6. Bir üniversite öğrencisi olarak aşağıdakileri yapmayacağımı kabul ediyorum:
  - Bir başkasının çalışmasını kendi çalışmam gibi sunmak.
  - Herhangi bir akademik çalışmayla/sınavla ilgili yetkisiz yardım vermek, almak veya aksi halde yararlanmak.
  - Bir başka kişiye, bir başka öğrencinin akademik çalışmasının zarar görmesi niyetiyle herhangi bir konuda yanlış bilgi vermek.

- Bir başka öğrencinin akademik çalışmasının zarar görmesi niyetiyle bir başka öğrencinin akademik aktivitelerle meşgul olma denemelerine engel olmak.
- Öğrenci, hasta bakımıyla direkt olarak ilişkili aktivitelerle veya akademik aktivitelerle meşgul olurken onu artan bir yaralanma veya hastalık riskine sokmak.
- Öğrencinin sınıf sırasını, not ortalamasını veya diğer herhangi bir akademik başarısını veya çabasını maddi olarak yanlış sunmak.
- Bir Etik Yemininin bozulmasını yanlış biçimde bildirmek.
- Bir Etik Yeminin bozulması iddiasının araştırmasını veya soruşturmasını, Etik Konseyine yanlış bilgi veya ifade vererek, bir vakayla ilgili bir tanığı veya suçlayan kişiyi veya bir Etik Kuralının bozulmasının soruşturmasıyla ilgili herhangi bir kişiyi taciz ederek veya bir Etik Konseyi tutanağındaki kanıtları yok ederek, saklayarak veya kanıt üreterek veya başka bir şekilde engellemek.
- Yukarıdaki ifadelere uymamamın bir akademik duruşmayla sonuçlanacağını anlıyorum (%30,54).

7. Üniversite, bilginin takibine adanmış akademik bir topluluktur. Bu takip için temel olan akademik bütünlüktür. Üniversite topluluğuna katılarak, Üniversitenin Akademik Bütünlük Açıklamasını kabul ediyorum ve bunun içindeki dürüstlük ve nezaket kurallarını destekleyeceğime yemin ediyorum. Kendim, doğruluk ve dürüstlük idealleriyle uyumlu olarak davranacağım ve bu idealleri bozacak hareketlere samimi biçimde karşı çıkacağım (%30,94).

8. Bu sınav sırasında kimseye yardım etmediğime ve kimseden yardım almadığıma şerefim üzerine yemin ederim. Öğrenciliğim boyunca, her kopya şüphesini öğrencilerden oluşan Etik Komitesine bildireceğime söz veriyorum. Etik Yeminin bozulması, cezalandırıcı hareketlerin en güçlüleri olan okuldan uzaklaştırma veya atılmayla karşı karşıya bırakır. Ders dışı alıştırmalar Etik Komitenin yetkisinin dışındadır. Bu durumlarda, öğrencilerin çalışmalarını aşımadıklarını beyan etmek için bu yemini makaleleri üzerinde imzalamaları beklenir. Bu makale, benim kendi çalışmamı Üniversite yönetmelikleriyle uyumlu biçimde sunmaktadır. (%28,80). Not: 'Üniversite' kelimesinin kullanıldığı yerlere öğrencilerin okulunun adı girilmiştir.

Yönteme gelince, ifadeler ABD'de farklı üniversitelerde kullanılan etik yeminlerinden çıkarılmıştır ve uzunluk, resmîlik ve ifadelerin akademik sahtekârlık için sonuçlar içermeye derecesine göre değişkenlik göstermektedir. Daha uzun, resmi olan ve sonuçları olan etik yeminleri daha düşük bir kopya çekme olasılığıyla ilişkilendirilmiştir. Sonuçlar önemli bir üçlü etkileşimi göstermekte ve etik yeminlerinin en iyi biçimde nasıl tasarlanacağını önermektedir. Tüm dünyanın ortak görüşü olarak, Akademik sahtekârlığın epidemi seviyesine yaklaştığı düşüncesinde birleşmişlerdir. Bu fikri ilk ortaya atan (Davis, Drinan & Gallant, 2009) dir. Birçok farklı sahtekârlık biçimi vardır ve çalışmalar kopya çekme oranlarının %68' den (sınav sırasında bir kopya sayfası kullanmak) %98' e (bir başka öğrenciden kopya çekmek; Davis ve diğerleri, 2009) kadar değiştiğini göstermektedir. İlk çoklu kampus çalışmalarının birinde, McCabe, Trevino ve Butterfield (1996), lisans öğrencilerinin %52 sinin önceki yıl en az bir kez bir sınavda kopya çektiklerini iddia ettiklerini gösterdiler. Online eğitimdeki artışa paralel olarak daha fazla fakülte online sınavlar yapmaktadırlar (Hay & Bull, 2002; Kamenetz, 2010; Meloncon, 2007; Pulliam, 2005). Öğrenciler online sınavları daha tatmin edici olarak değerlendirmekte ve bunları yaparken daha az sınav stresi bildirmektedirler (Dorrian & Wache, 2009; Stowell & Bennett, 2010). Ne yazık ki, bir sınav online olarak girmek öğrencilerin akademik olarak sahtekâr olma (yani kopya çekme) konusundaki teşvikini arttırabilir. Üniversitelerin birçok fakültesinde online sınavlar sırasında akademik sahtekârlığı azaltmak için birçok yöntem denenmiştir. Bazıları web kamerasıyla sınıfta disiplini sağlamayı keşfetmiş (Mirza & Staples, 2010), aşırma riskini azaltmak için sınav merkezleri ve etik kuralları kullanmışlardır (Khare & Lam, 2008). McCabe ve arkadaşları, etik kuralların

ve dış etkilerin akademik bütünlüğü nasıl etkilediğini görmek için bir çalışma yürütmüşlerdir. Akademik sahtekârlığın, akranın akademik sahtekârlığının anlayışlarıyla pozitif olarak ilişkili ve akademik bütünleştirme politikalarının *hem öğrenciler hem de fakülte tarafından görülen anlayışıyla ve kabulüyle ters ilişkili* olduğunu bulmuşlardır (McCabe ve diğerleri, 2002). Ayrıca, akademik sahtekârlığın, bir akran (yaşdaş) tarafından kopya çekilmenin bildirilmesinin *anlaşılan kesinliğiyle ve cezaların anlaşılan ciddiliğiyle ters ilişkili olduğunu* da bulmuşlardır. ***En önemli bulgu, etik yemini olan okullardaki akademik sahtekârlık seviyesinin asla bir etik yemini kullanmayan okullardakinden önemli ölçüde daha düşük olduğudur (Mc Cabe ve diğerleri, 2002).***

Bu çalışmada, akademik sahtekârlıktan vazgeçirmesi en olası olan bir etik yemininin elemanları bulunmaya çalışılmıştır. İlk olarak, 30'dan fazla okuldan etik yeminleri değerlendirilmiş ve yeminlerin uzunluğa, resmiliğe (örn., daha büyük kelimelerin ve resmi dilin kullanımı) ve yeminin bozulmasının sonuçlarından bahsedilip edilmediğine göre değişkenlik gösterdiği fark edilmiştir. Bu üç etkenin her birinin bilinçli biçimde değiştirilmesine karar verilmiş ve sonuçları olan daha uzun, resmi etik yeminlerin daha düşük bir kopya çekme olasılığıyla ilişkili olacağı hipotezlenmiştir.

### **Materyal Metot**

Tüm kampüse (6033 öğrenci) bu anonim çalışmaya katılmaları için bir e-posta daveti gönderilmiş olup, Qualtrics kullanılarak yaratılan soruşturmaya bir link verilmiştir. Katılımcılar bir izin formunu ve daha sonra etik yemini ifadelerinin her birini okumuşlardır. Katılımcılar, her etik kuralı/yeminine dayanarak kopya çekmelerinin (soruşturmada notları, ders kitabını veya bir başka kişinin yardımını kullanmak olarak tanımlanmıştır) mümkün olacağı, %0'le %100 arasındaki yüzdeyi belirtmek için kaydırıcı bir ölçek kullanmışlardır. Öğrencilere soruşturmayı tamamlamak için 2 haftalık bir süre verilmiştir.

Aynı yöntemle, IBM SPSS Statistics Version 26, yazılım programıyla gerçekleştirilen anket değerlendirme sonuçları biz de orta büyüklükte bir vakıf üniversitesi olarak Türkiye genelinde rastgele örnekleme yöntemiyle seçtiğimiz 12 (orta ve büyük ölçekte) üniversiteye uyguladığımız bir ankette 25 soruya verilen yanıtların sonuçlarını etik kurallar açısından bilhassa akademik sahtekarlığa öğretim elemanlarının nasıl baktığına ilişkin görüşlerinin diğer çalışmalarla karşılaştırıldığında bölgesel olarak farklılıkların ne olduğunu görmek açısından değerlendirmeyi düşündük. Çalışmamızda ankete ait verilere Kaiser-Meyer- Olkin Measure (KMO) testi uygulandı ve KMO testi %60 dan büyük bulunmuştur. Bunun anlamı istatistiksel tutum ölçeği maddelerine faktör analizi yapılabileceğini kontrol etmektir, bunun için KMO ve Bartlett testinden yararlanılmıştır. Sonuç, verinin faktör analizi için uygun olduğunu göstermiştir.  $p=0000$  dir.

### **Bulgular**

Ankette 8. soru “*Örgün eğitimin sınav koşullarının dışında öğrencilerden ahlaki olmayan, akademik dürüstlüğe sığmayan (aşırma, intihal, başkasının çalışmasını kopya etmek, izinsiz kitap kopyalayıp kullanmak, denetimsiz bir ortamda öğrenciye verilen ödevleri başkalarına yaptırmak vb.) davranışlardan kaçınmaları için onlara yemin gibi bir taahhüt önerir misiniz?*” şeklinde hazırlanmıştır. 71 öğretim elemanının verdiği yanıtta %62 sinin akademik dürüstlüğe sığmayan “*akademik sahtekarlık*” davranışlardan kaçınmaları için öğrencilere bir “*etik yemini*” taahhüt etmeyeceklerini yanıtlamışlardır. Halbuki kendi öğrencilerimiz üzerinde ne gibi etkilerin oluştuğu ve yukarıda sözü edilen etik dışı davranışların öğrenciler tarafından yapılmayacağına dair herhangi bir taahhütte (söz vermek) bulunmayı ankete katılan 103 kişilik bir örnekleme dayanarak kabul edenlerin oranının %79,61, taahhütte bulunmayanların oranı %14,56, kısmen buna onay verenlerin ise %5,83 olduğu saptanmıştır. Yine yukarıda “Wisconsin Üniversitesinde” gerçekleştirilen ankette öğrencilerin “Üniversitedeki dürüstlük ve sorumluluk kurallarını desteklemeyi kabul ediyorum” diyenlerin oranının (%34,97) olduğu gözlemlenmiştir. Öğrenci ve öğretim elemanı arasındaki farklılıkların hangi neden/nedenlerden

kaynaklandığını araştırmak çok önemlilik arz etmektedir ve araştırılması ilginç sonuçlar doğuracaktır. Ayrıca “*Uzaktan Eğitim araçlarıyla yapılan sınav veya ödevlerin yeterli ve ölçme/değerlendirmenin adil olduğuna inanmaktayım*” diyen öğretim elemanları %47 olasılıkla inandıkları ortaya çıkmıştır. Bu da ikinci ilginç bir sonuç olup öğrencilerimiz ise %66 olasılıkla *ölçme/değerlendirmenin adil olduğuna inanmakta oldukları söylemişlerdir*. Buna rağmen “*uzaktan eğitim sınavlarında öğrencilerin etik dışı (kopya çekme, ahlaki olmayan davranış, akademik dürüstlüğe aykırı) davranışlardan uzak durduğunu düşünüyorum*” sorusuna %37 olasılıkla etik dışı davranışlardan uzak durmayacaklarını ifade etmişlerdir. Bu soruya, Wisconsin Üniversitesindeki ankette öğrencilerin %30,40 olasılıkla katıldıklarını beyan etmişlerdir. Kendi üniversitemizde değerlendirdiğimiz ankette “*Uzaktan eğitim sınavlarında kopya çekmekten kaçınıyorum*” sorusuna %83,3 olasılıkla, “*Uzaktan eğitim sınavlarında hileden, suiistimalden kaçınıyorum*” sorusuna yine %83.1 ve yine ilgili diğer sorumuz olan “*Uzaktan eğitim ödevlerinde aşırıktan (intihalden), başkasının çalışmasını kendi çalışmam gibi sunmaktan kaçınıyorum*” sorusuna ise %86,6 gibi yüksek bir olasılıkla yanıt vermişlerdir. Bu çalışmamızda da **salgın** dolayısıyla uygulanan “*Salgın (Pandemi) dönemi için uygulanan uzaktan eğitim ile ilgili genel memnuniyetinizi derecelendiriniz*” sorusuna öğretim elemanları %55 olasılıkla tam karar veremedikleri yani “*kararsız olanların*” oranı oluşmuştur. Aynı soruya öğrencilerimiz ise %70 olasılıkla “*memnunum*” seçeneğini işaretlemişlerdir.

### **Tartışma**

Bu çalışmanın sonuçları, bildirilen akademik sahtekârlığın daha düşük seviyeleriyle ilişkili olduğunu göstermiştir. Uzaktan eğitim süresinde derslerinde eğitimi verebilmek için gerekli olan tüm doküman ve malzemeleri paylaşmak (%88), Uzaktan Eğitim sürecinde öğrencilere danışmanlık yapmak ve her soruyu yanıtlamak (%78), Uzaktan Eğitim derslerinin süresi yeterlidir (%67), Uzaktan Eğitimle ilgili sunulan ek kaynakların yeterliliği de önemli unsurlardır. Bunlar öğretim üyeleriyle yapılan ankette çıkan sonuçlardır. Bu sonuçlar öğrencilerin anketlere verdikleri yanıtlarla da örtüşmektedir. Bu durumda sonuçlar etik yeminlerinin daha düşük öğrenci sahtekârlığı seviyeleriyle ilişkili olduğunu gösteren önceki çalışmaları bir adımda ileriye götürerek hangi etkenlerin değiştirilmesi gerektiğini önermektedir. Uzaktan eğitimlerin (Online) testlerinin birçok kazancı da vardır ve bunlardan akademik sahtekârlık korkusu yüzünden kaçınmak talihsizlik olacağını düşünürüz.

Örneğin, yaptığımız incelemelerde birçok fakülte, öğrencilerin kopya çekecekleri korkusu yüzünden, daha ucuz ve yönetilmesi daha kolay olmasına rağmen online sınavlar yapmaz. Buzzetto-More (2008), öğrencilerin %72,8 inin online sınavları uygun gördüğünü ve %89.7 sinin notlar üzerinde hemen geri bildirim almayı sevdiğini bulmuştur (Buzzetto-More, 2008). Online sınavlar ders zamanından, kâğıttan ve baskı masraflarından tasarruf sağlar. Ayrıca, online sınavlar aktif öğrenim alıştırmaları ve ilginç konuları daha fazla deşmek için daha fazla ders süresi verilmesine olanak sağlar.

Bu çalışmanın sonuçları, akademik sahtekârlığı, resmi, daha uzun ve sonuçları olan bir etik yemini kullanarak en aza indirmek için yollar önermektedir. Bu çalışma gerçek bir kopya çekme davranışını incelemiş olmasa da akademik sahtekârlığı durdurmak için yollar tasarlamaya doğru iyi bir başlangıç olacağını düşünmektedir. Sonuçta, üniversiteniz birine bağlı olmasa bile dersler veya sınıflar için bir etik yemini tesis etmek akademik sahtekârlığı önemli ölçüde azaltabileceğini düşünmenizi önermekteyiz.

### **Kaynaklar**

Bilgen Kıral, Seda Saracaloğlu, 11-14 Mayıs 2017 tarihli IX. Uluslararası Eğitim Araştırmaları Kongresinde Ordu, YYÜ Eğitim Fakültesi Dergisi (YYU Journal of Education Faculty), 2018; 15(1):323-359.

- Buzzetto-More, N. (2008). Student perceptions of various e-learning Components. *Interdisciplinary Journal of Knowledge & Learning Objects*, 4, 113-135.
- Davis, S., F., Drinan, P. F., & Gallant, T. B. (2009). *Cheating in school: What we know and what we can do*. Malden, MA: Wiley-Blackwell.
- Dorrian, J., & Wache, D. (2009). Introduction of an online approach to flexible learning for on-campus and distance education students: Lessons learned and ways forward. *Nurse Education Today*, 29(2), 157-167.
- Ethics & Behavior*, 11, 219-232. doi:10.1207/S15327019EB1103\_2
- Hay, I., & Bull, J. (2002). Passing online exams. *Journal of Geography in Higher- Education*, 26, 239-244. doi:10.1080/03098260220144757.
- Kamenetz, A. (2010). *DIY U: Edupunks, edupreneurs, and the coming transformation of higher education*. White River Junction, VT: Chelsea Green.
- Khare, A., & Lam, H. (2008). Assessing student achievement and progress with online examinations: Some pedagogical and technical issues. *International Journal on E-Learning*, 7, 383-402.
- Mazmanoğlu, A. "Bilimsel Araştırma Yöntemleri ve Etik", *Toros Üni., Ders Notları*, 2019, MERSİN
- McCabe, D. L., Treviño, L., & Butterfield, K. D. (2001). Cheating in academic institutions: A decade of research.
- McCabe, D., Treviño, L., & Butterfield, K. (2002). Honor codes and other contextual influences on academic integrity: A replication and extension to modified honor code settings. *Research in Higher Education*, 43, 357-378.
- McCabe, D. L., Treviño, L., & Butterfield, K. D. (1996). The influence of collegiate and corporate codes of conduct on ethics-related behavior in the workplace. *Business Ethics Quarterly*, 6(4), 461-476.
- Meloncon, L. (2007). Exploring electronic landscapes: Technical communication, online learning, and instructor preparedness. *Technical Communication Quarterly*, 16, 31-55.
- Mirza, N., & Staples, E. (2010). Webcam as a new invigilation method: Students' comfort and potential for cheating. *Journal of Nursing Education*, 49, 116-119.
- Pulliam, L. (2005). Online testing: Promise or practice? *Media & Methods*, 41(4), 29.
- Regan A. R. Gurung, Tiffany M. Wilhelm, and Tonya Filz, "Optimizing Honor Codes for Online Exam Administration", Department of Psychology, University of Wisconsin-Green Bay, *ETHICS&BEHAVIOR*, 22(2), 158-162, Taylor & Francis Group, ISSN: 1050-8422, 2012, USA
- Roig, M., & Marks, A. (2006). Attitudes toward cheating before and after the implementation of a modified honor code: A case study. *Ethics & Behavior*, 16, 163-171. doi:10.1207/s15327019eb1602\_6
- Stowell, J., & Bennett, D. (2010). Effects of online testing on student exam performance and test anxiety. *Journal of Educational Computing Research*, #2, 161-171.

## Determination of G20 Indices Affecting BIST with Hybrid Regression Models

Ahmet SEL<sup>1</sup>

<sup>1</sup>Milli Eğitim Bakanlığı

### Abstract

Index forecasts are very difficult but attractive because stock market indices are influenced by several macroeconomic variables, such as political developments, general economic outlook, investor perceptions and investment preferences and movements of other indices. Using hybrid machine learning models, this study aims to predict the relationship between the BIST 100 index and the G20 index. The closing values of 2432 days were taken as the data set by considering the common days when the index days between 2010/01/01 and 2019/12/31 are open. The indices expected to have an impact on the BIST index among the G20 countries are Bovespa, Italy40, KOSPI, Nikkei 225, BMVIIPC, Shanghai and Tadawul, with correlation values (below 0.80) being considered. In the analysis, the lowest MAE = 0.0650, MSE = 0.0082 and RMSE = 0.0903 values for the 30 segmentation for  $R^2 = 0.9918$  were found using data mining software, with the aid of the Knime program, taking into account 70 percent of the data and 30 percent of the test. Based on the analysis results, it was found that the BIST-100 index was positively influenced by the Bovespa, Italy40, KOSPI, Nikkei 225 and BMVIIPC indices, while the Shanghai index had a negative impact. The Tadawul index was found to have no effect on BIST. The results obtained are thought to assist the decision-makers in estimating the course of the BIST index and in guiding the investments.

**Key words:** Data mining, big data, hybrid machine learning, regression

**Investigation of the Relation of Gold Prices and Crypto Money in the Pandemia Process  
with Machine Learning Methods**

Ahmet SEL<sup>1</sup>

<sup>1</sup>Milli Eğitim Bakanlığı

**Abstract**

In the decision-making process for investments, forecasting techniques and models are highly relevant for individuals and organizations. The benefit that investors get is maximized by accurate forecasts. Fluctuations were observed in all market instruments and investment instruments during the pandemic process. The influence of selected cryptocurrency forms on gold prices during the pandemic process is investigated in this research. The Python programming language is used to render applications. Via training and test sets, the relation between the independent variables Bitcoin, EOS, Tether, TRON and Ripple on gold prices was investigated by machine learning. The multivariate decision tree, help vector machines and random forest regression models were used together for machine learning models As a result; with the help of support vector machines with the highest explanation rate ( $R^2 = 0.91$ ), it has been determined that there are Bitcoin +0.755, EOS -0.596, Tether -0.122 and Tron + 0.220 effects on gold prices. The results obtained will assist decision-makers in directing their investments in times of high market volatility.

**Key words:** Gold, crypto money, pandemic, investment, machine learning

**Statistical Inference for the Quasi-Exponential Distribution Based on Progressively  
Type-II Censored Samples**

Sümeyra SERT<sup>1</sup>, İhab A.S. ABUSAIF<sup>1</sup>, Ertan AKGENÇ<sup>1</sup>, Kadir KARAKAYA<sup>1</sup>, Coşkun KUŞ<sup>1</sup>

<sup>1</sup>Department of Statistics, Selçuk University, 42250, Konya, Turkey

**Abstract**

In this study, the point and interval estimation on the parameter of quasi exponential distribution are discussed under a progressively Type-II censoring scheme. The maximum likelihood, modified likelihood, pivotal type and modified pivotal type estimates are obtained. In addition, interval estimation is also constructed based on asymptotic normality of maximum likelihood estimate and a pivotal quantity. A Monte Carlo simulation study is performed to discuss the mean squared error and bias of estimates. The coverage probabilities and mean length of the confidence intervals are also simulated. A numerical example is also provided.

**Key words:** Maximum likelihood estimator, modified maximum likelihood estimator, Monte Carlo simulation, pivotal inference, progressively Type II censoring scheme, Quasi exponential distribution



## Türkiye’den New Jersey’e Göç ve Tersine Göç Üzerine Araştırma

Abdurrahman KESKIN<sup>1</sup>, Recep Uğurcan ŞAHİN, Recep UZUNALIOĞLU

<sup>1</sup>İstanbul Üniversitesi, İktisat Fakültesi, Çalışma Ekonomisi Bölümü, İstanbul, Türkiye

### Özet

İnsanlık tarihi ile başladığı kabul edilen göç olgusunun, küreselleşmenin etkisiyle uluslararası hareketlilik boyutu ivme kazanmış ve görece basitleşmiştir. Uluslararası göç hareketliliğinde dönemsel olarak birçok faktör etkili olmaktadır. Günümüzde ülkelerin yaşadıkları iç karışıklıklar, siyasal problemler, doğa koşulları ve ekonomik zorluklar gibi itici faktörlerin yanı sıra iletişim ve ulaşım teknolojilerinin gelişmesi ile birlikte uluslararası göç artış göstermiştir. Türkiye’de ise son yıllardaki istatistiklere göre yurtdışına göç eden kişi sayısının arttığı görülmektedir. Bu araştırmanın amacı Türkiye’den Amerika Birleşik Devletleri’ndeki New Jersey Eyaleti’nin Paterson Bölgesi’ne göç eden bireylerin, göç etmesindeki temel belirleyici faktörleri ve tersine göç etme eğilimlerinin incelenmesidir. Araştırmamız uluslararası göç niteliği taşıyan bu örneklem üzerinden sınırlandırılmıştır. Bu amacı gerçekleştirmek için Türkiye’den göç eden 142 gönüllü katılımcıya internet üzerinden ve birebir olmak üzere bir anket çalışması uygulanmıştır. Elde edilen veriler correspondence ile incelenmiş, sonrasında ilişkili değişken ve düzeyleri yorumlanmıştır. Çalışmadan elde edilen bulgulara göre; eğitim amacı ile göç eden ve eğitim düzeyi yüksek olan bireylerin büyük bölümü Türkiye’ye geri dönmeyi düşünmemektedirler. Ayrıca nitelikli işlerde çalışan ve gelir seviyesi yüksek olan bireylerin büyük bir bölümünün Türkiye’ye geri dönmeyi düşünmediği veya kararsız oldukları sonucuna ulaşılmıştır. Ekonomik amaçla göç eden, eğitim düzeyi lise seviyesinde olan, gelir seviyesi düşük olan ve nitelsiz işlerde çalışan bireylerin büyük bir kısmında ise Türkiye’ye dönme düşüncesi bulunmaktadır.

**Anahtar kelimeler:** Göç, göç nedenleri, uluslararası göç, tersine göç, correspondence analizi

## Şekerli Su Kullanımının Ross Erkek Broilerlere Etkilerinin Yaşam Analiziyle Belirlenmesi

Emre ASLAN<sup>1</sup>, Özgür KOŞKAN<sup>1</sup>, Yasin ALTAY<sup>2</sup>, Sedat AKTAN<sup>1</sup>

<sup>1</sup>Isparta Uygulamalı Bilimler Üniversitesi, Zootekni Bölümü, Isparta, Türkiye

<sup>2</sup>Eskişehir Osmangazi Üniversitesi, Zootekni Bölümü, Eskişehir, Türkiye

### Özet

Yaşam analizi yöntemi yapılacak çalışmalarda başarısızlık olarak adlandırılan ve çoğunlukla ölüm gibi karşımıza çıkan olayların meydana gelmesine kadar geçen sürenin analizidir. Bu çalışma da farklı iki muamele uygulanan hayvanların Yaşam Tablosu ve Kaplan-Meier Yöntemi yöntemi ile ortalama yaşam süreleri hesaplanmıştır. Materyal olarak 450 adet ross erkek broiler hayvan kullanılmıştır. Deneme tek faktörlü ve faktörün iki seviyesi (kontrol ve muamele) olacak şekilde rastgele örneklenerek hayvanlar her bir grupta 225 adet olacak şekilde ortalama 42 gün boyunca yürütülmüştür. Deneme sonunda kontrol ve muamele gruplarının ortalama yaşama süreleri hesaplanmıştır. Şekerli su verilen hayvanların normal su verilen hayvanlara göre ölüm oranı daha az olduğu tespit edilmiştir. Sonuç olarak şekerli su kullanımının civciv döneminde ölüm oranı üzerine bir etkisi olmadığı sonucu ortaya çıkmıştır ( $P>0.05$ ). Yaşam analizi metodu ortalama süre tahmini ortaya koyduğu için bir çok hayvancılık araştırmasında ekstra bir bilgi vereceğinden dolayı tercih edilmesi ile literatüde daha fazla katkı sağlanacağı düşünülmektedir. Ayrıca yaşam analizi yöntemi hayvancılıkta alanında kullanımına sahip olup bir çok uygulamada ve materyalde denenerek kullanılabilir.

**Anahtar kelimeler:** Broiler yaşama süresi, Kaplan-Meier, şekerli su, yaşam tablosu

### Determining the Effects of Sugar Water Use on Ross Male Broilers by Survival Analysis

#### Abstract

Survival analysis is the analysis of the time until the occurrence of events such as death, which is called a failure in studies to be carried out. In this study, the average life expectancy of animals treated with two different treatments was calculated using the Life Table and Kaplan-Meier method. 450 Ross male broiler animals were used as material. The experiment was carried out by randomly sampling with one factor and two levels of the factor (control and treatment), and using the animals for an average of 42 days, with 225 animals in each group. At the end of the experiment, the average survival times of the control and treatment groups were calculated. The animals that were given the sugar water had less mortality than the normal watered animals. Although the usage of sugared water has no effect on mortality during the 42 day period, it is statistically insignificant in practice ( $P> 0.05$ ). Since the method of survival analysis gives an average time estimate, it will provide extra information in many animal husbandry studies, and it is thought that it will make more contribution to the literature with its preference. Besides, the life analysis method has its use in the field of animal husbandry and can be used in many applications and materials.

**Key words:** Broiler life time, Kaplan-Meier, sugary water, life tables

## Giriş

Yaşam tablosu yöntemi en eski yöntemler arasında kullanılan bir istatistik metottur (Lee ve Wang, 2003). Yaşam tablosu metodu, zaman aralıklarına göre yaşam süresi verilerini tablolaştırarak ele alıp ve bu aralıklarda yaşam işlevlerini hesaplayan bir yöntemdir (Altman 1991; Parmar & Machin 1995).

Özellikle de tıp, ziraat, mühendislik, veterinerlik, ormancılık, nüfus bilimcileri ve sigortacılar tarafından yaygın kullanılmaktadır (Cox, 1972). Çoğunlukla sigortacılar farklı yaş gruplarındaki nüfusun ne zamana kadar hayatta kalma sürelerini tahminin de kullanılmaktadır. Tıp alanında ise yaptıkları çalışmaların genelinde Cohort yaşam tablosu denilen aynı grubun belirli bir zaman izlendiği tablolar kullanılmaktadır (Özdemir, 1994). Klinikte ve özellikle de kanserden ölüm oranlarının hesaplanmasında ve yaşam süresinin tahminlerinde kullanılmıştır. Yaşam tablosu analizleri 1940'lı yılların sonlarında işaretlenmiş hayvan gruplarının yıllık belirli yaşlardaki ölüm oranlarının tahmini içinde kullanılmıştır (Morris, vd., 1988).

Yaşam analizleri bir olayın zamana bağlı olarak oluşması ve sonucunda bu olaya etki eden faktörlerin belirlenmesi ve eldeki değişkenlere bağlı olarak ortalama yaşam sürelerini tahmin etmekte kullanılan bir analiz metotudur. Bu metotun kullanılmaya başlaması kaba ölüm hızları kullanılarak yapılan analizlerden daha güvenilir sonuçlar vermesi bakımından önemli bir gelişmedir. Bu metot bireylerin aynı bir şekilde çalışmaya alınmasını gerektirmeme özelliğine de sahiptir. Bunun yanı sıra da çalışma dışında kalanların verilerini de kullanılabilir (Sümbüloğlu ve Akdağ, 2009).

Diğer bir yöntem ise Kaplan-Meier yöntemi olup, Kaplan ve Meier tarafından 1958 yılında geliştirilen bu yöntem başlangıçtan itibaren zaman aralıklarını kullanmadığı için az sayıda veriler içinde kullanışlıdır. Kaplan-Meier metodu ile yaşam eğrisi, medyan, yaşam süresi gibi değerlerde bulunabilir (Clark, vd., 2003).

Kaplan-Meier yönteminde zaman içinde her olay yeni bir ihtimali tahmin edilmektedir. Bunun gereğinde elde edilen her bir şekil aşama gibi görünmektedir. Bu durumda da belli bir sıra olay süreleri arası izlenmemektedir. Çoğu analizde Kaplan-Meier yönteminin son aşama olabileceği veya bu adımın, gelişmiş analizlerde ilk veya başlangıç olarak istatistik model seçiminde yardımcı olabileceği vurgulanmıştır (Harrel, 2001). Araştırmacı tarafından belirlenen örnekte incelediği olayın yaşanma yoğunluğu, oranının veya ihtimalinin zamana ait değişimini analiz eden bir yöntemdir. Bu metot yaşam tablolarına göre az sayıda veri ile analiz yapılabilir (Özdamar, 2003).

Bu çalışmanın amacı yaşam analizi metodunu tanımlayarak, ross broiler erkek kanatlı hayvanların normal su ve şekerli su kullanımının Yaşam analiz metodunu kullanarak ortalama yaşam sürelerinin belirlenmesidir.

## Materyal ve Metot

Çalışmada 450 adet erkek ross broiler kullanılmıştır. Bu hayvanların grup faktörünün 2 seviyesi olup grup 1 şekerli su, grup 2 ise normal su olarak dikkate alınarak yaşam analizleri uygulanmıştır.

Yaşam analizi (Sağkalım) için uygulanan Yaşam Tablosu ve Kaplan-Meier Yöntemi olmak üzere 2 yöntem bulunmaktadır (Özdamar, 2003). Bu analizler İstatistik paket programından SPSS 24 yardımı ile yapılmıştır.

### *Yaşam analizinde yaşam tablosu metodu*

Yaşam Tablosu (Life table, actuarial) metodu, izlem (yaşam) süresini eşit zaman aralıklı frekans tablosuna çevirerek her bir sınıf değeri için yaşam fonksiyonlarını tahmin etmeyi amaçlar. Bu metot, yaşam (izlem) süresinin 6' dan fazla ( $k > 6$ ) eşit aralıklı sınıf oluşturabilecek ve birim sayısının 100 den

fazla ( $n > 100$ ) olduğu verilerde kullanılabilir. Yöntem ilk başta % 100 olan yaşama ihtimalinin belirli aralıklarla ile vardığı yeniyaşama ihtimal düzeylerini belirler (Şenocak, 1992). Ayrıca bu metotta sınıf orta değerlerine göre tahmin yapılır (Özdamar, 2003).

İzleme periyodu, incelenen verinin özelliğine göre, 3'er aylık, 6'şar aylık veya 1'er yıllık aralıklara bölünür ve elde edilen her aralık için sağkalım hesaplamaları yapılır. Çözümlemede kullanılan simgeler aşağıdaki gibidir (Dişçi, 2011).

$L_i$  : i. aralığın başında izlenmekte olan hasta sayısı,

$D_i$  : i. aralıkta, sonlanan ('olay' tanımına göre, ölen ve/veya nüks eden) hasta sayısı,

$U_i$  : i. aralıkta çeşitli nedenlerle izlenemeyen (kayıp) hasta sayısı,

$E_i$  : i. aralıkta risk altındaki hasta sayısı

$Q_i$  : i. aralıkta sonlanma hızı

$$Q_i = L_i - \frac{D_i}{E_i}$$

$P_i$  : i. aralıkta sağkalım hızı

$$P_i = 1 - Q_i$$

$R_i$  : başlangıçtan 1. aralığın sonuna kadar olan dönemde birikimli sağkalım hızı

Tablo 1. Survival Analizinde Yaşam Tablosu

(1) (Interval) j	(2) No. Alive beginning of interval $O_i$	(3) No. lost to follow up or withdrawn alive during interval $W_i$	(4) No. Died during interval $\lambda_i$	(5) No. Exposed to risk of dying during interval) $O_i$	(6) Proportion dying $q_i$	(7) Proportion surviving $P_i$	(8) Cumulative proportion surviving $P_i$
1	$O_1$	$W_1$	$\lambda_1$	$O_1$	$q_1$	$P_1$	$P_1$
2	$O_2$	$W_2$	$\lambda_2$	$O_2$	$q_2$	$P_2$	$P_2$
3	$O_3$	$W_3$	$\lambda_3$	$O_3$	$q_3$	$P_3$	$P_3$
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
n	$O_n$	$W_n$	$\lambda_n$	$O_n$	$q_n$	$P_n$	$P_n$

Yaşam analizinin hesaplanması için Tablo 1'de hazırlanan yaşam tablosu oluşturulmuştur. Burada ilk sütun aralık sayısını, 2. sütun her aralığın başındaki hasta sayısını bulunmaktadır. 3. sütun aralıkta gözlemden çıkan veya kayıp olan hasta sayısı yer almaktadır. 4. sütun aralıktaki ölüm sayısı bulunmaktadır. 5. sütun aralıktaki risk altındaki hasta sayısını hesaplamaktadır. 6. sütunda aralıktaki ölme sayısını, 7. sütun aralıktaki sağkalım ihtimalini ve son olarak 8. sütun ise yaşam dağılımının yaşam tablosu ihtimalini göstermektedir.

### **Kaplan-Meier yöntemi**

Kaplan-Meier yöntemi sağkalım dağılımlarının şekil ile gösterilmesinde yararlanır ve çok kullanışlıdır. Yaşam analizinde en yaygın olarak kullanılan yöntem olmakla birlikte bazı önemli özellikleri de bulunmaktadır. Bu özellikler; sağkalım (yaşam) analizindeki özel istatistiklerde en çok kullanılan medyan süresi olmaktadır. Diğer tahminleyicilere benzer şekilde, Kaplan-Meier tahminleyicisinin

standart hatası, bu tahminleyicinin potansiyel hatasını vermektedir. Güven aralığı nokta tahmininden daha güvenilirdir (İnceoğlu, 2013).

Bu yöntemde N birim için (yaşam süreleri)  $t_1 < t_2 < t_3 < \dots < t_n$  olmak üzere yaşam olasılık fonksiyonu aşağıda verilmiştir.

Yaşam fonksiyonu (Y(t)), standart hatası (SE(Y(t))) ve eklemeli ölüm fonksiyonu ( $\Lambda(t)$ ) sırasıyla aşağıda verilmiştir (Özdamar, 2003).

$$Y(t) = \prod [(N-i+1 - \delta_i) / (N-i+1)]$$

$$SE(Y(t)) = Y(t) [\sum (\delta_i / ((N-i)(N-i+1)))]^{0.5}$$

$$\Lambda(t) = -\ln Y(t)$$

Ayrıca bu metot ile farklı tedavi yöntemlerinin yapıldığında veya farklı grupların karşılaştırmasında farklı yaşam fonksiyonları ve buna bağlı olarak farklı yaşam süreleri elde edilir.

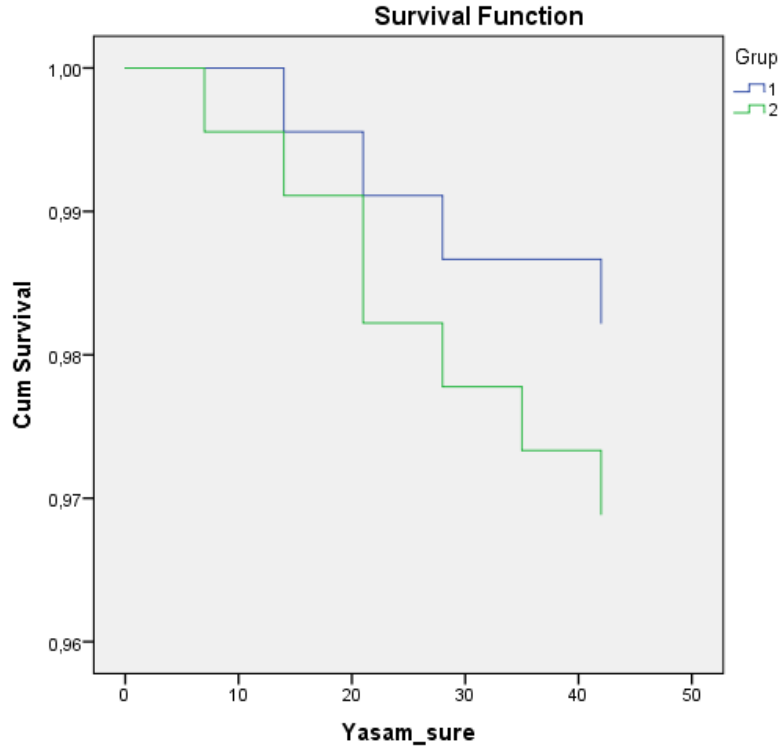
### Bulgular ve Tartışma

Çalışmaya alınan 450 adet ross broiler erkek kanaltı hayvan üzerinde uygulanan yaşam analizinde, yaygın olarak kullanılan analiz metotlarından olan Yaşam Tablosu ve Kaplan-Meier metodu kullanılmıştır.

Tablo 2. Yaşam Tablosu Metodu ile Yapılan Yaşam Analizi Sonuçları

First-order Controls	Interval Start Time	Number Entering interval	Number Withdrawing during interval	Number Exposed to Risk	Number of Terminal Events	Proportion Terminating	Proportion Surviving	Cumulative Proportion Surviving at End of Interval	Std. Error of Cumulative Proportion Surviving at End of Interval	Probability Density	Std. Error of Probability Density	Hazard Rate	Std. Error of Hazard Rate
Şekerli Su	0	225	0	225	0	0.00	1.00	1.00	0.00	0.000	0.000	0.00	0.00
	7	225	0	225	1	0.00	1.00	1.00	0.00	0.001	0.001	0.00	0.00
	14	224	0	224	1	0.00	1.00	0.99	0.01	0.001	0.001	0.00	0.00
	21	223	0	223	1	0.00	1.00	0.99	0.01	0.001	0.001	0.00	0.00
	28	222	0	222	0	0.00	1.00	0.99	0.01	0.000	0.000	0.00	0.00
	35	222	0	222	1	0.00	1.00	0.98	0.01	0.001	0.001	0.00	0.00
42	221	221	110	110	0	0.00	1.00	0.98	0.01	0.000	0.000	0.00	0.00
Normal Su	0	225	0	225	1	0.00	1.00	1.00	0.00	0.001	0.001	0.00	0.00
	7	224	0	224	1	0.00	1.00	0.99	0.01	0.001	0.001	0.00	0.00
	14	223	0	223	2	0.01	0.99	0.98	0.01	0.001	0.001	0.00	0.00
	21	221	0	221	1	0.00	1.00	0.98	0.01	0.001	0.001	0.00	0.00
	28	220	0	220	1	0.00	1.00	0.97	0.01	0.001	0.001	0.00	0.00
	35	219	0	219	1	0.00	1.00	0.97	0.01	0.001	0.001	0.00	0.00
42	218	218	109	109	0	0.00	1.00	0.97	0.01	0.000	0.000	0.00	0.00

Yaşam analizi sonuçları Tablo 2’de yaşam tablosu metodu kullanılarak ilk satırda gözlem süresi sıfırdan başlatılarak 7 şer gün ara ile 42. güne kadar şekerli su ve normal su verilmiştir. Şekerli su verilen kanatlı hayvanlarda ölüm 7, 14, 21 ve 35 günler de görülür iken, normal su verilen kanatlı hayvanlar da ise ilk günden itibaren 42. gün haricinde diğer günlerde ölüm gözlenmektedir. Şekerli su verilen hayvanlarda ölüm oranı, normal su verilen hayvanlara göre daha az gözlenmektedir.



Şekil 1. Normal ve Şekerli Su Verilen Hayvanların Ortalama Yaşam Süreleri

Şekil 1 incelendiğinde, ortalama yaşam fonksiyonu normal su verilen grupta ölümler daha erken başlamış olup, ilerleyen zamanla kümülatif ölüm sayısı oranı artmaktadır. Bu durum şekerli su verilen grupta daha geç başlamış olup ölüm sayıları ve oranları daha az olduğu görülmektedir.

Tablo 3. Normal ve Şekerli Su Verilen Hayvanların Durum İşleme Özeti

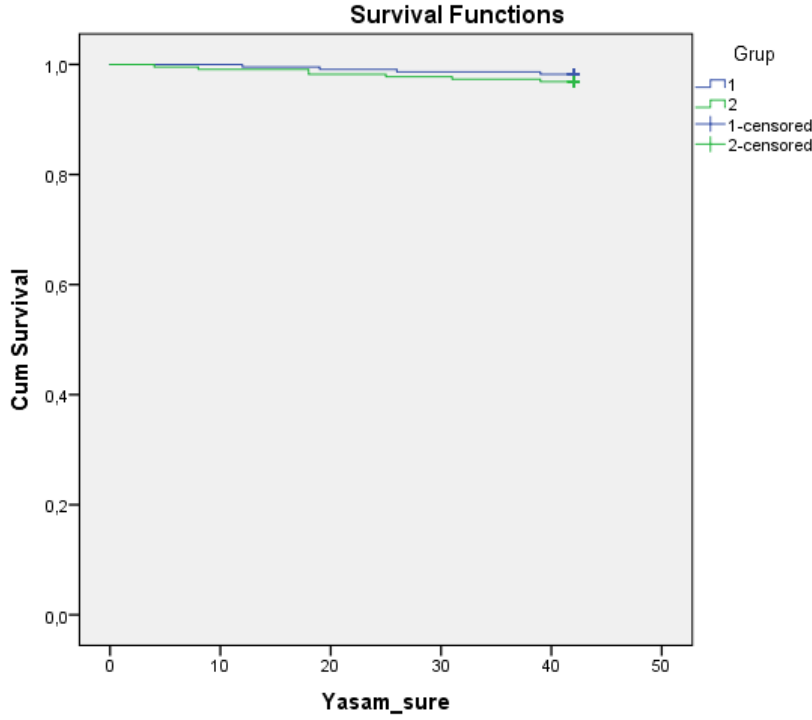
Grup	Total N	N of Events	Censored	
			N	Percent
Şekerli Su	225	4	221	98.2%
Normal Su	225	7	218	96.6%
Overall	450	11	439	97.6%

Kaplan-Meier metodu kullanılarak yapılan yaşam analizinde Tablo 3 incelendiğinde, şekerli su verilen grupta sadece 4 adet ölüm olurken, normal su verilende ise 11 adet ölüm olmuştur. Son durumda şekerli su verilen hayvan sayısı 221 iken, normal su verilen hayvan sayısı ise 218 toplamda 439 adet hayvan kalmıştır.

Tablo 4. Normal ve Şekerli Su Verilen Hayvanların Ortalama Yaşam Süreleri Tahmini

Grup	Mean			
	Estimate	Std.Error	%95 Confidence Interval	
			Lower Bound	Upper Bound
Şekerli Su	41.680	0.182	41.324	42.036
Normal Su	41.329	0.284	40.773	41.885
Overall	41.504	0.169	41.174	41.835

Tablo 4'te incelendiğinde, her iki grup içinde ortalama yaşam tahminleri şekerli su verildiğinde 41.68 ve normal su verildiğinde 41.33 olarak belirlenmiştir.



Şekil 2. Normal ve Şekerli Su Verilen Hayvanların Ortalama Yaşam Süreleri

Yaşam fonksiyonun grafiksel olarak sunulması bilgi edinmeyi kolaylaştırır. Bu durum da Şekil 2 'de Kaplan-Meier eğrisi yaşam eğrisinin bir tahminidir. Şekilde dik kısım yaşam olasılıklarını, yatay kısım ise süre yi göstermektedir. Grup 1 şekerli su kullanan hayvanların yaşam olasılıkları, grup 2 normal su kullanan hayvanlara göre daha yüksek olduğu görülmektedir. Şekerli su kullanan hayvanların yaşam olasılıklarının diğer gruba göre daha yüksek olduğu söylenebilir.

### Sonuç

Çalışma genel olarak değerlendirildiğinde, yaşam tablosu ve Kaplan-Meier metotlarının kullanımı sağlık alanın yanı sıra hayvancılığın gerek yetiştirme gerekse de besleme alanında da uygulanılabileceği görülmektedir. Broiler yetiştiriciliğinde şekerli su kullanımının ölüm oranlarını azaltmada bir etkisinin olmadığı ve istatistik olarak önemli olmamasına rağmen ticari yetiştiricilikte hayvan sayısının (örnek içerisindeki deney ünitesi sayısı) fazlalığı göz önünde bulundurulduğunda ekonomik olarak önemli olduğu vurgulanabilir. Yaşam analizleri ortalama süre tahminini ortaya koyduğu için bir çok hayvancılık araştırmasında ekstra bir bilgi vereceğinden dolayı tercih edilmesi ile çalışmaların literatüre daha fazla katkı sağlanacağı düşünülmektedir.

### Kaynaklar

- Altman, D. G., 1991, Practical Statistics For Medical Research, Chapman and Hall, Cornwall.  
Brian, S. E., and Sophia, R. H., 2007, A Handbook of Statistical Analyses Using Stata, 4th Edition.  
Clark, TG., Bradburn MJ., Love, SB., Altman, DG., 2003. Survival Analysis Part I: Basic Concepts and First Analyses. British Journal of Cancer, 89: 232-238.  
Cox, D. R., 1972, Regression Models and Life Tables, Journal of Royal Statistical Society, Series 8, 34, 187-220.  
Cox, D. R., and Oakes, D., 1984, Analysis of Survival Data: Chapman and Hall, London, 201 p.  
Dişçi, R., 2011. Temel ve Klinik Biyoistatistik. İstanbul Medikal Yayıncılık, 221-234s, İstanbul.  
Harrel, F., 2001, Regression Modelling Strategies, New York, Springer, 568p.

- İnceođlu, F., 2013. Sađkalım Analiz Yöntemleri ve Karaciđer Nakli Verileri İle Bir Uygulama. İnönü Üniversitesi, Sađlık Bilimleri Enstitüsü, Yüksek Lisans Tezi, 1-58s, Malatya.
- Lee, ET., Wang, J., 2003. Statistical Methods for Survival Data Analysis, Wiley.
- Morris, A. B., Day, A. M., Macmillan, K. L., 1988. Studies of Evolution Rate Embryo Survival and Calving Rate in a Twin Breeding Experiment. 11 th International Congress on Animal Reproduction an Artificial İnsemination, Vol.2, 109, University College Dublin, Ireland.
- Özdamar, K., 2003. SPSS ile Biyoistatistik. Kaan Kitabevi, 506s, Eskişehir.
- Özdemir, A., 1994, Hayatta Kalabilme Analizi Yöntemleri, Yüksek Lisans Tezi, Ankara Üniversitesi, Ankara.
- Parmar, M., Machin, D., 1995, Survival Analysis, John Wiley and Sons Ltd.
- Sümbülođlu, K., Akdađ, B., 2009. İleri Biyoistatistiksel Yöntemler. Hatibođlu Yayınları, 224- 245s, Ankara.
- Şenocak M., 1992, Özel Biyoistatistik Epidemiyolojide Sayısal Çözümleme, İstanbul, Çađlayan Kitabevi, 221s.



## Regression Analysis of Morphological Traits on Body Weight in South African Non-Descriptive Indigenous Goats

K. MOKOENA<sup>1</sup>, R. VHLOKOE<sup>1</sup>, L.T RASHIJANE<sup>1</sup>, K.M MOLABE<sup>1</sup>, M.C MATHAPO<sup>1</sup>,  
L. J SELALA<sup>1</sup>, M. MASWANA<sup>1</sup>, T.L TYASI<sup>1\*</sup>

<sup>1</sup>School of Agricultural and Environmental Sciences, Department of Agricultural Economics and Animal Production, University of Limpopo, Private Bag X1106, Sovenga 0727, Limpopo, South Africa

### Abstract

Regression is the statistical technique used for estimation of the relationships between one dependent variable and one or more independent variables. The objectives of the study are to: 1) investigate the relationship between body weight (BW) and morphological traits such as; withers height (WH), body length (BL), rump length (RL), rump height (RH) and heart girth (HG) and 2) to develop a model for the prediction of BW. One hundred South African non-descript indigenous goats (does = 80 and bucks = 20) between age one to four years were used as experimental animals. The best prediction equation for estimation of body weight was selected based on the coefficient of determination ( $R^2$ ) and mean square error (MSE). Pearson correlation and simple linear regression analysis were used for data analysis to achieve the study objectives. Pearson correlation findings revealed that BW had a positive highly statistical significant ( $P < 0.01$ ) correlation with WH ( $r = 0.70$ ), RH ( $r = 0.40$ ) and BL ( $r = 0.54$ ) in does while BW of bucks had a negative statistical significant ( $P < 0.05$ ) correlation with RL ( $r = -0.26$ ), positive correlation with WH ( $r = 0.28$ ), positive highly statistical significant correlation with BL ( $r = 0.37$ ) and low correlation with HG ( $r = 0.06$ ). Regression analysis developed models for each morphological trait and the findings indicated that withers height had a highest coefficient of determination ( $R^2 = 0.49$ ) and lowest mean square error (MSE = 46.17) in does, while rump height had a highest  $R^2$  (0.21) with lowest MSE (88.44) in bucks. In conclusion, association findings suggest that by improving WH, RH and BL might result the improvement of BW in South African non-descript indigenous goats. Simple linear regression results suggest that WH and RH could be used as a selection criterion throughout breeding to improve BW of indigenous goats. Findings of the current study might be used by indigenous goat farmers as selection criteria and for creation of an appropriate breeding program for improvement of BW based on morphological traits.

**Key words:** Withers height, body length, rump height, rump length, heart girth

## Determination of Best Fitted Regression Model for Estimation of Body Weight in Nguni Cattle

V.R HLOKOE<sup>1</sup>, K. MOKOENA<sup>1</sup>, L.T RASHIJANE<sup>1</sup>, K.M MOLABE<sup>1</sup>, M.C MATHAPO<sup>1</sup>,  
L.J SELALA<sup>1</sup>, M. MASWANA<sup>1</sup>, T.L TYASI<sup>1</sup>

<sup>1</sup>School of Agricultural & Environmental Sciences, Department of Agricultural Economics and Animal Production, University of Limpopo, Private Bag X1106, Sovenga 0727, Limpopo, South Africa

### Abstract

Stepwise regression is the statistical technique involving step-by-step construction of regression model in which an automatic selection of independent variables occurs. The study was conducted to determine the relationship between body weight (BW) using biometric traits such as head width (HW), head length (HL), ear length (EL), body length (BL), rump height (RH), withers height (WH), sternum height (SH), rump width (RW), and heart girth (HG) from Nguni and to determine the best fitted regression model to be used for prediction of live weight of Nguni cattle. A total of sixty female and twenty male Nguni cattle between the ages of one to four years were used. Pearson correlation and stepwise regression were used for data analysis. The correlation results showed that BW had a positive highly statistical significant correlation ( $P < 0.01$ ) with WH ( $r = 0.93$ ), HG ( $r = 0.90$ ), RH ( $r = 0.90$ ), SH ( $r = 0.89$ ), RW ( $r = 0.70$ ), and a positive statistical significant correlation ( $P < 0.05$ ) with EL ( $r = 0.45$ ), and BL ( $r = 0.43$ ) in males, whereas BW had a positive highly statistical significant correlation ( $P < 0.01$ ) with HG ( $r = 0.76$ ), RH ( $r = 0.70$ ) and WH ( $r = 0.55$ ), and a positive statistical significant correlation ( $P < 0.05$ ) with BL ( $r = 0.29$ ), and HW ( $r = 0.26$ ) in female Nguni cattle. Stepwise regression models were formed with BW as dependent variable and HW, HL, EL, BL, RH, WH, SH, RW and HG as independent variables. The best fitted regression model for estimation of body weight was determined by using mean square error (MSE) and coefficient of determination ( $R^2$ ). The model including WH, HG, RH, SH, RW and EL was the best fitted model ( $R^2 = 0.91$ ,  $MSE = 1454.16$ ) for estimation of body weight in males, and the model including HG, RH, WH, SH and BL was the best fitted model ( $R^2 = 0.66$ ,  $MSE = 4540.86$ ) for estimation of body weight in female Nguni cattle in this study. Correlation findings suggests that improvement of WH, HG, RH, SH, RW, EL, BL and HW might result in the increase in BW of Nguni cattle. Regression results suggests that body weight can be estimated accurately by combination of two or more biometric traits.

**Key words:** Body length, rump height, withers height, sternum height, rump width, heart girth

## Forecasting for Tea Production and Yield in Turkey

Nur İlkay ABACI<sup>1</sup>

<sup>1</sup>Ondokuz Mayıs University, Faculty of Agriculture, Department of Agricultural Economics, 55139, Samsun, Turkey

### Abstract

When tea is evaluated in terms of consumption, it has many positive effects on the human body with its antioxidant properties and is the most consumed beverage in the world after water. In terms of the amount of tea consumed per capita in the world, Turkey ranks first. When evaluated countries in the world in terms of production, according to the FAO 2019 data, China 2.8 million tons, while in first place with black tea production, Turkey 1407448 tons of fresh tea leaves and 251327 tons, ranks fifth with black tea production. However, it ranks 8th after Zimbabwe in terms of black tea production (319.88 kg/da) obtained from the decare. This study is aimed to make forecasting about the future of tea to be produced in Turkey in 2023. This study is aimed to forecasting the tea to be produced in Turkey in 2023. For this purpose, tea production area (da), fresh and black tea production amount (ton), fresh tea and black tea production per decare (kg) and black tea processing the loss values (%) per decare between 1988-2019 years were used, which were obtained from TURKSTAT and FAO. Forecastings were made in SPSS using Brown's Double Exponential Smoothing method. According to the findings, in the next 4 years until 2023 compared to 2019, an average of 1.42% in the tea production area, 6.22% in fresh tea production, 6.96% in black tea production, 5.36% in fresh tea production, and 6.39% in black tea production per decare increase was estimated. However, when the loss per 1 decare, in black tea processing is examined, an increase of 0.02% is expected. It is thought that it will be beneficial to minimize this loss especially by using planned fertilization and using technological tea processing devices during the production phase.

**Keywords:** Black tea, fresh tea, production, time series, yield

## Bağlantı Dengesizliği Katsayısının Örneklem Dağılımı Üzerine Bir Simülasyon Çalışması

Selma KÖKÇÜ<sup>1</sup>, Orhan KAVUNCU<sup>1</sup>, Yasemin GEDİK<sup>2</sup>

<sup>1</sup>Kastamonu Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Genetik ve Biyomühendislik Bölümü, Kastamonu, Türkiye

<sup>2</sup>Eskişehir Osmangazi Üniversitesi, Ziraat Fakültesi, Zootečni Bölümü, Eskişehir, Türkiye

### Özet

Bu çalışmanın amacı bağlantı dengesizliği katsayısı sıfır olmayan popülasyonlardan çekilecek örneklerden hesaplanan bağlantı dengesizliği katsayılarının kuadratik formlarının nasıl bir dağılım göstereceğinin belirlenmesidir. Bağlantı dengesinde olan popülasyonlardan çekilen örneklerden hesaplanan bağlantı dengesizliği katsayısının (D) ortalamasının sıfır olduğu ( $\mu_D=0$ ) ve kuadratik bir form olarak

$$\chi_0^2 = \frac{n D^2}{p_1 \cdot p_2 \cdot q_1 \cdot q_2}$$

istatistiğinin (metinde 2 numaralı istatistik) bir Khi-Kare dağılımı gösterdiği bilinmektedir. Bu çalışma teorik bir çalışma olup canlı materyal ile deneme kurulmamıştır.  $\mu_D=0$  ve  $\mu_D=0.15$  bağlantı dengesizliği katsayılarına, iki lokus iki allel sayılarına sahip popülasyonlardan simülasyon yöntemiyle çekilen  $n=10, 20, 50$  genişlikteki örneklerden hesaplanan birisi yukarıdaki kuadratik bir form olan istatistik olmak üzere üç ayrı kuadratik istatistiğin dağılımı çalışılmıştır.

Sonuçta bağlantı dengesizliği katsayısının  $0.15$  olduğu popülasyondan çekilen örneklerden hesaplanan bu kuadratik formun 3 s.d.'li noncentral Khi-Kare dağılımı gösterdiği bulunmuştur. Paydaki D'nin, ortalamadan sapmasının karesi olarak hesaplanan kuadratik form (metinde 3 numaralı istatistik) ise beklendiği gibi bağlantı dengesizliği katsayısı  $\mu_D=0.15$  olduğu zaman da 1 s.d.'li Khi-Kare dağılımı göstermektedir. Klasik bir Khi-Kare formülü olan ve metinde 4 numara ile verilen istatistik de bağlantı dengesizliği katsayısı  $\mu_D = 0.15$  olduğu zaman da yine beklendiği gibi 3 s.d.'li Khi-Kare dağılımı göstermektedir.

**Anahtar kelimeler:** Bağlantı dengesizliği katsayısı, Khi-kare dağılımı, noncentral Khi-kare dağılımı

### Abstract

The aim of this study is to determine the distribution of the quadratic forms of the linkage disequilibrium coefficients calculated from the samples drawn from populations with non-zero linkage disequilibrium coefficient. It is known that the average of coefficient of linkage disequilibrium (D) calculated from samples drawn from populations in linkage equilibrium is zero ( $\mu_D = 0$ ) and the statistic

$$\chi_0^2 = \frac{n D^2}{p_1 \cdot p_2 \cdot q_1 \cdot q_2}$$

(statistic number 2 in the text), which is a quadratic form, shows a chi square distribution. This study is a theoretical study and no trial has been established with live material. It is studied the distribution of

three different quadratic statistics (one is given above, for the other two see the methods) calculated from samples of size  $n = 10, 20, 50$  generated through simulation from populations with linkage disequilibrium coefficients of  $\mu_D = 0$  and  $\mu_D = 0.15$  for a model of two alleles, two loci.

The results show that the first quadratic form (statistic number 2) which is much known statistics in literature, has found to be fitted noncentral chi-square distribution with 3 df in the samples drawn from the populations of  $\mu_D = 0.15$ . The second quadratic form (statistic number 3 in the methods) was stable in being fitted to the central chi-square distributions of 1 df in the existence of linkage disequilibrium. The third quadratic form (statistic number 4) has a distribution fitted to chi square with 1 df.

**Key words:** Linkage disequilibrium, Chi- square distribution, noncentral Chi-square distribution

## Giriş

Bağlantı dengesizliği üzerine çalışmalar 1950’li yıllarda başlamıştır (Kavuncu 1987). İki allelli iki lokus için bağlantı dengesizliği katsayısı rastgele çiftleşen bir popülasyonda

$$\mu_D = P(AB).P(ab)-P(Ab).P(aB) \quad (1)$$

şeklinde hesaplanır. Burada bir lokustaki alleller A ve a, diğer lokustaki alleller ise B ve b harfleriyle gösterilmiştir.  $P(AB)$ , oluşacak gametlerden rastgele birinin AB olma ihtimalidir. Popülasyon bağlantı dengesindeyse, bu gamet popülasyonundan çekilen n gametlik bir örnekte bağlantı dengesizliği katsayısı istatistiği D’nin sıfır olması beklenir ve böyle bir örnekte hesaplanan kuadratik formdaki

$$\chi_0^2 = \frac{n D^2}{p_1 \cdot p_2 \cdot q_1 \cdot q_2} \quad (2)$$

(2) numaralı istatistik 1 serbestlik dereceli Khi-Kare dağılımı gösterir (Hartl ve Clark 2007, Gillespie 2004, Nielsen ve Slatkin 2013). Paydadaki  $p_1$  A geninin,  $q_1$  a geninin,  $p_2$  B geninin ve  $q_2$  ise b geninin frekanslarıdır. Lokuslarda allel sayısı ikiden fazla olduğunda da (1) numaralı istatistiğin çok allelli duruma uyarlanmasıyla, yine bağlantı dengesi varsayımı altında,  $(k_1-1)(k_2-1)$  serbestlik dereceli bir Khi-Kare dağılımı gösteren bir istatistik elde edilir. Burada  $k_1$  ve  $k_2$  birinci ve ikinci lokustaki allellerin sayısıdır. (2) numaralı istatistiğin Khi-Kare dağılımı göstermesi, popülasyonun bağlantı dengesinde olduğu varsayımına, yani örneklerden hesaplanan D’lerin ortalamasının sıfır olduğu ( $\mu_D=0$ ) varsayımına bağlıdır.

Ancak popülasyon bağlantı dengesinde değilse örneklerden hesaplanan bağlantı dengesizliği katsayısının nasıl bir dağılım göstereceğini bilmek gerekir. Ne var ki bağlantı dengesinde olmayan popülasyonlardan çekilen örneklerden hesaplanacak bağlantı dengesizliği katsayısı ve (2) numaralı kuadratik formdaki gibi istatistiklerin örnekleme dağılımı üzerine bir çalışmaya rastlanamamıştır. DNA verilerinin istatistik analizinden bağlantı dengesizliğini ölçmeye yönelik teorik çalışmalara kadar taranan literatürde bu durum geçerlidir (Weir 1996).

Bu çalışmada bu eksikliği gidermeye yönelik bir çalışmanın bir bölümü sunuma hazırlanmıştır. Böyle bir çalışmanın Teorik popülasyon genetiğinin ilgili alanlarında yararlı olması beklenmektedir. Ayrıca bu çalışmadan elde edilecek sonuçların, genetik varyans unsurlarının tahminindeki isabeti artıracığı, bu yüzden de kantitatif genetik ve bu tahminlere çokça ihtiyaç duyulan bitki ve hayvan ıslahı çalışmalarında kullanılma olanağı bulacağı düşünülmektedir. Çünkü eğer popülasyon bağlantı dengesinde değilse lokusların etkileri arasındaki kovaryans sıfır olmayacak, dolayısıyla bu kovaryans, genetik varyans unsurlarının tahmin edilmesinde sapmalara yol açacaktır (Kavuncu 1987). Bunun yanında nükleotid sitelerinin birlikte dağılımını dikkate alan moleküler ölçekli (Crow ve Kimura 1970)

çalışmalarda ve gerek polimorfizm ve gerekse protein varyasyonları gibi biyoinformatik alanlarda da katkı sağlaması ümit edilmektedir.

Bu tebliğde, iki lokusta iki allel olan ve çeşitli bağlantı dengesizliği katsayılarına sahip olan sanal popülasyonlardan simülasyon yöntemiyle çekilen farklı genişlikteki örneklerden hesaplanan bağlantı dengesizliği katsayısı ile ilgili istatistiklerin dağılımı belirlenmeye çalışılacaktır.

Çalışmada bağlantı dengesizliği katsayıları ve örnek genişlikleri olmak üzere iki faktörün seviyelerine göre sanal popülasyonlar belirlenmiş olacaktır. Bütün bu örneklerin simülasyonu, hesaplamalar ve örnekleme dağılımları R yazılımı kullanılarak geliştirilecek bilgisayar programları ile bulunacaktır.

## Yöntem

### *Bağlantı dengesizliği katsayısından hesaplanan kuadratik istatistikler*

Bağlantı dengesizliği katsayıları gamet frekanslarına bağlıdır. Buna göre, birinci lokustaki A allelinin gittiği lokusa ikinci lokusta B veya b allelinin gitmesi tamamen şansa bağlı ise popülasyon bu iki lokusta bağlantı dengesindedir. İki lokus bağlı, yani aynı kromozom üzerinde olsa da iki lokustaki allellerin gametlere dağılımı birbirinden bağımsız olabilir. Bu durumda bağlantı dengesizliği katsayısı sıfır olacaktır.

Örnekten hesaplanan bağlantı dengesizliği katsayısının sıfır olduğu hipotezini kontrol etmek için (2) numaralı istatistikle hesaplanan değer kullanılır. Ancak örneğin çekildiği popülasyonda bağlantı dengesizliği katsayısı sıfır değilse o zaman (2) numaralı istatistik iyi bir istatistik olmayabilir. Bunun yerine

$$\chi_1^2 = \frac{n (D - \mu_D)^2}{p_1 \cdot p_2 \cdot q_1 \cdot q_2} \quad (3)$$

istatistiği kullanılabilir. Burada  $p_1$  A geninin frekansı,  $p_2$  ise B geninin frekansıdır.  $q_1 = 1 - p_1$ ,  $q_2 = 1 - p_2$

Bunun yanında popülasyon bağlantı dengesizliği katsayısından bağımsız olarak

$$\chi_2^2 = \sum_{i=1}^4 \frac{(x_i - np_i)^2}{np_i} \quad (4)$$

(4) numaralı istatistik de bir çeşit kontrol grubu üretmek üzere denenebilir. Burada  $x_i$ 'ler,  $i=1,2,3$ , ve 4 için sırasıyla AB, Ab, aB ve ab gametlerinin örnekteki sayısı,  $p_i$ 'ler de aynı sırada gametlerin popülasyondaki frekanslarıdır. Bu durumda bu istatistiğin serbestlik derecesi 3 olacaktır.

Bu üç istatistiğin dağılımı Khi-Kare veya merkezi olmayan (noncentral) Khi-Kare olabilir. Merkezi olmayan Khi-Kare dağılımı için merkezi olmama parametresi, (3) numaralı istatistik için,

$$\lambda = n \cdot \mu_D^2 \quad (5)$$

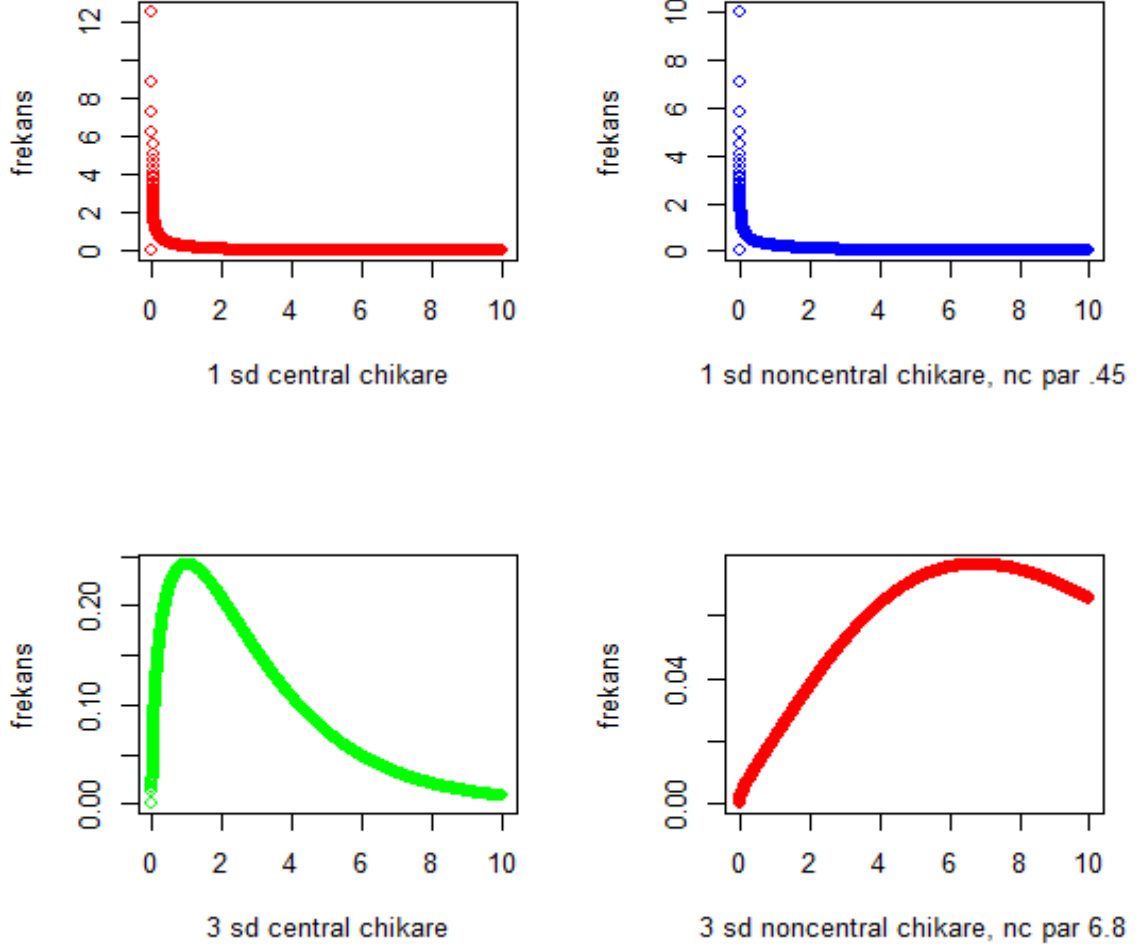
(4) numaralı istatistik için ise

$$\lambda = n \sum_{i=1}^4 p_i^2 \quad (6)$$

alınmıştır. Merkezi olmayan Khi-Kare için Seber (1963)'e bakınız.

Bu (2), (3), (4) numaralı üç istatistik simülasyon yöntemi ile üretilen n genişliğindeki k örnek için hesaplanmış ve ilk iki istatistiğin (2 ve 3 numaralı istatistikler) merkezi (central) ve merkezi olmayan 1 s.d.'li Khi-Kare dağılımına uygunluğu, üçüncü istatistiğin (4 numaralı istatistik) de merkezi ve merkezi olmayan 3 s.d.'li Khi-Kare dağılımına uygunluğu, Kolmogorov-Smirnov testi ile kontrol edilmiştir. İşlem 20 kere tekrarlanmıştır. 1 ve 3 s.d.'li merkezi ve merkezi olmayan Khi-Kare dağılımlarının grafiği ve merkezi olmama parametreleri 1 s.d.'li Khi-Kare için 5 numaralı istatistiğe göre 0.45, 3 s.d.'li Khi-Kare için 6 numaralı istatistiğe göre 6.8 olarak hesaplanmıştır. Üretilen örneklerden hesaplanan

istatistiklerin dağılımına örnek olarak örnek genişliği  $n=20$  ve bağlantı dengesizliği katsayısı  $\mu_D=0.15$  kombinasyonu için bir tekerrür Şekil 1’de gösterilmiştir.



Şekil 1. Üretilen Örneklem Dağılımlarının Uygunluğu Araştırılan Khi-Kare Dağılımları

### *Simülasyonla örnek üretilen popülasyonlar*

Aslında bir yüksek lisans tezi olan çalışmanın bu tebliğde sunulan kısmında iki allelli iki lokus için biri bağlantı dengesinde, diğeri bağlantı dengesizliğinde ( $\mu_D=0.15$ ) olan iki popülasyondan  $n=10, 20$  ve  $50$  genişliğinde  $10,000$ 'er örnek çekilmiştir. Her simülasyon işlemi  $20$  defa tekrarlanmıştır. Buna göre araştırmada ele alınan faktörler iki bağlantı dengesizliği katsayısı ( $0$  ve  $0.15$ ) ve üç örnek genişliği ( $10, 20$  ve  $50$ ) için olmak üzere  $2*3=6$  popülasyonun her biri simülasyonla  $10000$  defa üretilmiş ve bu altı simülasyonun her biri  $20$  defa tekrarlanmış; deneme tekniği tabiriyle simülasyon  $20$  tekerrürlü yapılmıştır. Çalışma için yazılan algoritma  $\mu_D=0.15$  ve  $n=20$  için Tablo 2.1’de verilmiştir.

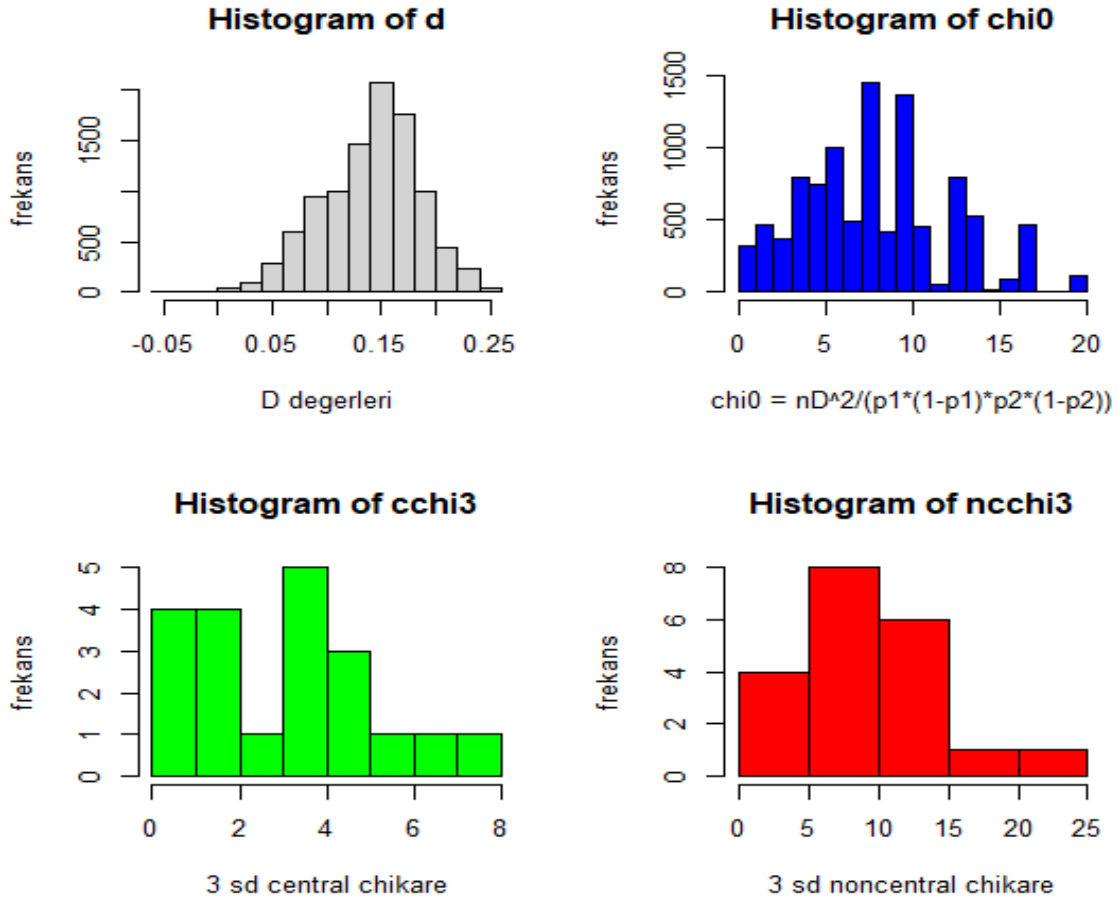
Tablo 1. Çalışmada Kullanılan R Programındaki Algoritma

<pre>#parametreler: q1=P(AB), q2=P(Ab), # q3=P(aB), q4=P(ab) nrep=deneme sayisi, # ngen=örnek genişliği q1=.4 q2=.1 q3=.1 q4=.4 nrep=10000 ngen=20 dort=q1*q4-q2*q3 ncpar=dort^2*ngen ncpar3=ngen*(q1^2+q2^2+q3^2+q4^2) dort ncpar ncpar3 d=c(rep(seq(0,0),nrep)) chi1=c(rep(seq(0,0),nrep)) chi2=c(rep(seq(0,0),nrep)) chi0=c(rep(seq(0,0),nrep)) for(j in 1:nrep){ x1=0 x2=0 x3=0 x4=0 y=runif(ngen) ncchi3=rchisq(ngen,3,ncpar3) ncchi1=rchisq(ngen,1,ncpar) cchi1=rchisq(ngen,1) cchi3=rchisq(ngen,3) par(mfrow=c(4,2)) for (i in 1:ngen){ ifelse(y[i]&lt;=q1,(x1=x1+1),ifelse (y[i]&lt;=(q1+q2),(x2=x2+1),ifelse (y[i]&lt;=(q1+q2+q3),(x3=x3+1),(x4=x4+1)))) } d[j]=(x1*x4-x2*x3)/ngen^2 p1=(x1+x2)/ngen p2=(x1+x3)/ngen chi0[j]=ngen*d[j]^2/(p1*(1-p1)*p2*(1-p2))</pre>	<pre>chi1[j]=ngen*(d[j]-dort)^2/(p1*(1-p1)*p2*(1- p2))chi2[j]=(x1-ngen*q1)^2/(ngen*q1)+(x2- ngen*q2)^2/(ngen*q2)+(x3- ngen*q3)^2/(ngen*q3)+(x4-ngen*q4)^2/(ngen*q4) } d=round(d,digits=3) d.table=table(d) d.table # d degerlerinin frekansini yazdir mean(d) # chi0 = nD^2/(p1*(1-p1)*p2*(1-p2)) # chi1 = n*(d-dort)^2/(p1*(1-p1)*p2*(1-p2)) # chi2 = tam chi-kare #histogramlar hist(d,ylab="frekans",xlab="D degerleri") hist(chi0,ylab="frekans",xlab="noncentral chikare") hist(chi1,ylab="frekans",xlab="central chikare") hist(chi2,ylab="frekans",xlab=" ??? chikare") hist(ncchi1,ylab="frekans",xlab="1 s.d noncentral chi-kare") hist(ncchi3,ylab="frekans",xlab="3 s.d noncentral chi-kare") hist(cchi1,ylab="frekans",xlab="1 s.d central chi- kare") hist(cchi3,ylab="frekans",xlab="3 s.d central chi- kare") ks.test(chi0,cchi1) ks.test(chi0,ncchi1) ks.test(chi0,cchi3) ks.test(chi0,ncchi3) ks.test(chi1,cchi1) ks.test(chi1,ncchi1) ks.test(chi1,cchi3) ks.test(chi1,ncchi3) ks.test(chi2,cchi1) ks.test(chi2,ncchi1) ks.test(chi2,cchi3) ks.test(chi2,ncchi3)</pre>
---	--

### Araştırma Sonuçları

Simülasyondan önce, bağlantı dengesizliği katsayısı sıfır olan popülasyondan üretilen örneklerden hesaplanan (2) ve (3) numaralı kuadratik istatistiklerin örnekleme dağılımının 1 serbestlik dereceli Khi-Kare dağılımına uygunluğu öngörülürken (4) numaralı kuadratik istatistik için 3 s.d.'li Khi-Kare dağılımına uygunluk öngörülüyordu. Tablo 3. 1, 3. 2 ve 3. 3'te bu üç kuadratik formun örnekleme dağılımının, denenen altı popülasyonda ele alınan Khi-Kare dağılımlarına uygunluk için yapılan 20'şer Kolmogorov-Smirnov testinden kaç tanesinde hipotezin kabul edildiği ( $p > .05$ ) verilmiştir. Üretilen örneklerin histogramları da yapılmış ve Kolmogorov-Smirnov testi yanında gözle de gözlem yapmak niyetiyle 20 tekerrürün her birinde Şekil 1'de verilen teorik dağılımlarla karşılaştırılmıştır. Bu histogramlara örnek olarak  $n=20$  ve  $\mu_D=0.15$  için tekerrürlerin birisinde üretilen 10000 örnekten elde edilen D değerleri ve (1), (2) ve (3) numaralı istatistiklere ait histogramlar Şekil 2'de gösterilmiştir.





Şekil 2.  $\mu_D=0.15$  ve  $n=20$  için Üretilen 10000 Örnekten Hesaplanan Değerlerin Grafikleri

D değerlerinin dağılımı (gri), (2) numaralı istatistikle elde edilen değerlerin dağılımı (mavi). 3 s.d.'li Khi-Kare dağılımı (yeşil). 3 s.d.'li noncentral Khi-Kare dağılımı (kırmızı).

Tablo 2’de (2) numaralı kuadratik istatistik için  $n=10, 20, 50$  örnek genişliğine sahip bağlantı dengesizliği katsayısı  $\mu_D=0$  ve  $\mu_D=0.15$  olan sanal popülasyonlardan hesaplanan dağılımlardan  $\mu_D=0$  ve  $n=10$  örnek genişliğine sahip bir popülasyonda 20 kez tekrarlanan simülasyon sonucunda 1 serbestlik dereceli central Khi-Kare dağılımına uygunluk gösteren tekrarların sayısı 19, 1 serbestlik dereceli noncentral Khi-Kare 18, 3 serbestlik dereceli central Khi-Kareye 1 ve 3 serbestlik dereceli noncentral Khi-Kareye 0 tekrar bulunmuştur.

Buna karşılık  $\mu_D=0.15$  ve  $n=10$  için 3 s.d.’li central ve noncentral Khi-Kare 20 tekrarın 15’inde uygun bulunmuştur.  $n=20$  ve 50 için bulunan sonuçlar aynı bulguyu teyit etmektedir.

Tablo 2. Altı Popülasyondan Elde Edilen Sonuçlar (2 Numaralı İstatistik İçin)

Bağlantı Dengesizliği	Örnek Genişliği(n)	Khi Kare uygunluk			
		1 s.d central	1 s.d noncentral	3 s.d central	3 s.d noncentral
$\mu_D=0$	10	19	18	1	0
	20	19	19	0	0
	50	19	20	0	0
$\mu_D=0.15$	10	0	2	15	15
	20	0	0	0	17
	50	0	0	0	16

Tablo 3’de ikinci kuadratik form olan (3) numaralı istatistiğin sonuçları görülmektedir. Bu istatistiğin örnekleme dağılımı  $\mu_D=0$  için 1 s.d.’li central ve noncentral Khi-Kareye uyum göstermektedir. Noncentralite parametresi  $\mu_D=0$  için sıfır, dolayısıyla noncentral Khi-Kare ile central Khi-Kare’nin aslında aynı dağılım olduğu dikkate alınır bu sonuç olması beklenendir.

Tablo 3. Altı Popülasyondan elde edilen sonuçlar (3 Numaralı İstatistik İçin)

Bağlantı Dengesizliği	Örnek Genişliği(n)	Khi Kare uygunluk			
		1 s.d central	1 s.d noncentral	3 s.d central	3 s.d noncentral
$\mu_D=0$	10	19	18	1	0
	20	20	19	0	0
	50	19	19	0	0
$\mu_D=0.15$	10	17	17	0	0
	20	16	9	0	0
	50	14	0	0	0

Tablo 4’te üçüncü kuadratik form (4) numaralı istatistik için bulunan sonuçlar verilmiştir. Bu kuadratik form üretilen gamet frekansları ile aynı varsayımlara göre beklenen gamet frekansları arasındaki farkın karesine dayandığından bulunan sonuçlar normaldir.

Tablo 4. Altı Popülasyondan elde edilen sonuçlar (4 numaralı istatistik için)

Bağlantı Dengesizliği	Örnek Genişliği(n)	Khi Kare uygunluk			
		1 s.d central	1 s.d noncentral	3 s.d central	3 s.d noncentral
$\mu_D=0$	10	0	0	19	5
	20	0	0	19	0
	50	0	0	20	0
$\mu_D=0.15$	10	1	3	19	3
	20	0	2	20	0
	50	0	0	16	0

## Tartışma ve Öneriler

Çalışmada bulunan sonuçlara göre (Tablo 2) bağlantı dengesizliği katsayısı sıfır olmayan popülasyonlarda birinci kuadratik formun (2) numaralı istatistiğin noncentral Khi-Kare dağılımı göstermektedir. Ancak 1 değil de 3 s.d.’li Khi-Kare sonucu beklenmeyen bir sonuçtur. Bunun üzerinde biraz daha çalışmak gerektiği sonucuna varılmıştır.

Sonuçları Tablo 3’de verilen ikinci kuadratik form (3) numaralı istatistiğin beklendiği gibi, bağlantı dengesizliği parametresinden farkın karesi olarak hesaplandığından 1 s.d.’li Khi-Kare dağılımı göstermektedir.

Tablo 4’te verilen sonuçlar bir çeşit kontrol grubu oluşturmak üzere elde edilmiş sonuçlardır. Beklendiği gibi burada sonuçlar her durumda 3 s.d.’li central Khi-Kare dağılımına uygunluk göstermektedir.

$n=10$  için bulunan sonuçlar, bu örnek genişliğinin Khi-Kare dağılımının ön şartlarını sağlamaya yetmemesinden kaynaklanıyor olabilir. Çünkü bu kadar küçük bir örnek genişliğinde, neredeyse bütün sınıflarda beklenen frekanslar 5’ten küçük çıkacaktır. Bu durumda popülasyon genetiği ile ilgili parametre tahmini veya hipotez kontrolü çalışmalarında örnek genişliğinin 10’dan büyük olması gerektiği anlaşılmaktadır.

Sonuç olarak çalışmanın, ilerleyen dönemde, ikiden fazla allelli modelleri de kapsamına alması düşünülmektedir. Uygunluk testi için burada kullanılan Kolmogorov-Smirnov testi yanında başka testlerin de denenmesi yararlı olacaktır.

**Kaynaklar**

- Crow J and Kimura M, 1970, An Introduction to Population Genetics, Harper and Row Publisher Inc.,  
Gillespie, J. H., 2004, Population Genetics. Baltimore and London. The Johns Hopkins University  
Press., ISBN 0-8018-8008-4.
- Hartl D. L. and Clark A.G., 2007 4th ed., Principles of Population Genetics, Sinauer Assoc., USA.
- Kavuncu O, 1987, Effect of Linkage Disequilibrium on the Additive and Dominance Variances in a  
Random Mating Small Population, Communication (Journal Of Science Published by Ankara  
Univ.), Series A1, 36: 131–142.
- Nielsen R. and M. Slatkin, 2013. An Introduction to Population Genetics, Sinauer Associates, Inc.  
Publishers, Sunderland, Massachussets U. S.A.
- Seber, G. A. 1963, The Non-Central Chi-Squared and Beta Distributions. *Biometrika*,50(3/4), 542-544.
- Weir, B.S, 1996, Genetic Data Analysis II. Sinauer Assoc., Inc., Sunderland, MA.

## Estimating Claims Reserve Splitting IBNER and IBNYR in the Chain Ladder Method

Kleida HAXHI<sup>1</sup>, Endri RAÇO<sup>1</sup>, Etleva LLAGAMI<sup>1</sup>, Oriana ZAÇAJ<sup>1</sup>

<sup>1</sup>Polytechnic University of Tirana, Faculty of Mathematical Engineering and Physical Engineering,  
Albania

### Abstract

Sometimes the incurred claim cost contains more information than paid claims. Instead of using a chain ladder estimate we can split IBNER and IBNYR along the lines of Schnieper. This split, into IBNER and IBNYR reserves gives valuable insight into the development process that is significant for the final estimates. The most important feature of the method is that it allows to use a prior volume measure, such as premiums, as exposure for unknown claims. In contrast to the Chain Ladder Method that takes incurred claims as exposure for the combined IBNYR and IBNER reserve.

**Key words:** Claims reserving, RBNS, IBNR, chain ladder method

### Introduction

We consider non-life claims reserving based on incurred claims costs. IBNR (incurred but not reported) are the unknown claims and RBNS (reported but not settled) are the claims that are known but not paid. IBNR is divided into IBNER (incurred but not enough reported) and IBNYR (incurred but not yet reported).

The most important feature of the method is that it allows to use a prior volume measure, such as premiums, as exposure for unknown claims. In contrast to the Chain Ladder Method that takes incurred claims as exposure for the combined IBNYR and IBNER reserve.

### Chain Ladder Method

The chain ladder method is one of the most popular used method for estimating outstanding claims reserves. The main reason is its simplicity and the fact that it is distribution free. This does not mean that under this method no statistical assumptions.

$C_{ik}$  denote the accumulated total claims amount of accident year  $i$ ,  $1 \leq i \leq n$ , paid or incurred up to development year  $k$ ,  $1 \leq k \leq n$ . The values of  $C_{ik}$  for  $i + k \leq n + 1$  are known and we want to estimate the values of  $C_{ik}$  for  $i + k > n + 1$ , in particular the ultimate claims amount  $C_{in}$  for each accident year  $i=2, \dots, n$

The calendar years are on the diagonal. The “CL idea”, in the wording of is that all accident years behave similarly and that we have approximately

$$C_{ik} = C_{i,k-1} \cdot f_k$$

**Table 1. Incurred Claims Triangle**

<i>Accident</i>	<i>Development year</i>				
<i>year</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>...</i>	<i>n-1</i>
<i>1</i>	$C_{10}$	$C_{11}$	$C_{12}$	$\dots$	$C_{1,k-1}$
<i>2</i>	$C_{20}$	$C_{21}$	$C_{22}$	$\dots$	
<i>3</i>	$C_{30}$	$C_{31}$	$C_{32}$	$\dots$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
<i>n-1</i>	$C_{k-1,0}$	$C_{k-1,1}$			
<i>n</i>	$C_{k,0}$				

*Mack's assumptions*

- For some parameters  $f_k > 0$ , and all  $i = 1, \dots, n$  and  $k = 1, \dots, n-1$ , we have:

$$E(C_{i,k+1} | C_{i1}, \dots, C_{ik}) = C_{ik} f_k$$

- The random variables for the accident years are independent, i.e. the random vectors  $\{C_{i1}, \dots, C_{in}\}$  are independent of each other.

The Chain Ladder estimator is:

$$\hat{f}_k = \frac{\sum_{i=1}^{n-k} C_{i,k}}{\sum_{i=1}^{n-k} C_{i,k-1}}$$

This estimator is unbiased and uncorrelated for different  $k$ .

By multiplying by the successive development factors, we get an estimate of the ultimate claim cost  $C_{i,n-1}$ .

$$\hat{C}_{i,n-1} = C_{i,n-i} \cdot \hat{f}_{n+1-i} \cdot \dots \cdot \hat{f}_{n-1}$$

**Schnieper's Method**

For Schnieper method we need to split the cumulative incurred claims triangle in two: the new claims triangle and the run-off triangle. Let  $N_{ik}$  be the incurred claim cost for claims incurred in accident year  $i$  and reported during development year  $k$ , as recorded by the end of that year. This is the status by the end of the year when the claims were reported. We can compute the new claims triangle if our claims data base contains information about the reporting date for each claim. Schnieper use an exposure measure for IBNYR,  $E_i$ . It would be earned premium, sum insured or duration for the entire portfolio year  $i$ .

We denote by  $D_{ik}$  the change in the incurred cost for existing claims.

$$D_{ik} = C_{ik} - C_{i,k-1} - N_{ik}$$

It can be written as the collection of any two of our three triangles:

$$D_k = \{N_{ik}, D_{ik}; i + k \leq n\}$$

Table 2. New Claims Triangle

<i>Accident</i>	<i>Development year</i>				
<i>year</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>...</i>	<i>n-1</i>
<i>1</i>	$N_{10}$	$N_{11}$	$N_{12}$	$\dots$	$N_{1,k-1}$
<i>2</i>	$N_{20}$	$N_{21}$	$N_{22}$	$\dots$	
<i>3</i>	$N_{30}$	$N_{31}$	$N_{32}$	$\dots$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
<i>n-1</i>	$N_{k-1,0}$	$N_{k-1,1}$			
<i>n</i>	$N_{k,0}$				

Table 3. Run-off Triangle

<i>Accident</i>	<i>Development year</i>				
<i>year</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>...</i>	<i>n-1</i>
<i>1</i>	$D_{10}$	$D_{11}$	$D_{12}$	$\dots$	$D_{1,k-1}$
<i>2</i>	$D_{20}$	$D_{21}$	$D_{22}$	$\dots$	
<i>3</i>	$D_{30}$	$D_{31}$	$D_{32}$	$\dots$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
<i>n-1</i>	$D_{k-1,0}$	$D_{k-1,1}$			
<i>n</i>	$D_{k,0}$				

*Schnieper's assumptions*

a) For parameters  $\lambda_k \geq 0$ ,  
 $E[N_{ik} | D_{i+k-1}] = E_i \lambda_k ; \quad i = 1, 2, \dots, n ; \quad k = 1, 2, \dots, n - 1$

b) For some parameters  $\delta_k$  ,  
 $E[C_{i,k-1} + D_{ik} | D_{i+k-1}] = C_{i,k-1} \delta_k ; ; \quad i = 1, 2, \dots, n ; \quad k = 1, 2, \dots, n - 1$

c) The random variables for the accident years are independent, the random vectors  $\{N_{i0}, D_{i0}, \dots, N_{i,k-1}, D_{i,k-1}\}$  are independent of each other.

It follows from (a) and (b) that

$$E[C_{i,k} | D_{i+k-1}] = E[C_{i,k-1} + D_{ik} + N_{ik} | D_{i+k-1}] = C_{i,k-1} \delta_k + E_i \lambda_k \quad *$$

This is in contrast to the Chain ladder in the notation,

$$E[C_{i,k} | D_{i+k-1}] = E[C_{i,k-1} + N_{i,k} + D_{i,k} | D_{i+k-1}] = C_{i,k-1} f_k$$

It means that the expected cost for new claims is not influenced by the observed claims, but instead proportional to the earned premium, while in the Chain ladder the expected value of both N and D are dependent on  $C_{i,k-1}$ .

From assumptions (a) and (b) it is the following estimators are unbiased, given  $D_{i+k-1}$ .

$$\widehat{\lambda}_k = \frac{\sum_{i=1}^{n-k} N_{i,k}}{\sum_{i=1}^{n-k} E_i} \quad k = 1, \dots, n - 1$$

$$\widehat{\delta}_k = \frac{\sum_{i=1}^{n-k} (N_{i,k-1} + D_{i,k})}{\sum_{i=1}^{n-k} C_{i,k-1}} = 1 + \frac{\sum_{i=1}^{n-k} D_{i,k}}{\sum_{i=1}^{n-k} C_{i,j-1}} \quad k = 1, \dots, n - 1$$

While accident year  $i = 1$  is fully developed, for year  $i = 2$  we can find that

$$E[C_{2,k-1}|D_n] = C_{2,k-2} \delta_{k-1} + E_2 \lambda_{k-1}$$

An estimator  $\hat{C}_{2,k-1}$  of the ultimate claim cost is found by linking up in the estimates  $\hat{\lambda}_k$  and  $\hat{\delta}_{k-1}$ .

By using \* and iterated expectation

$$\begin{aligned} E[C_{i,k+1}] &= E[E[C_{i,k+1} | D_{i+k} | D_{i+k-1}]] \\ &= E[C_{i,k} \delta_{k+1} + E_i \lambda_{k+1} | D_{i+k-2}] \\ &= C_{i,k-1} \delta_k \delta_{k+1} + E_i [\lambda_k \delta_{k+1} + \end{aligned}$$

This procedure can be used repeatedly to find an expression for the conditional expectation of the ultimate claim cost  $C_{i,k-1}$  for the accident years  $i > 2$

$$\begin{aligned} (**) \quad E[C_{i,k-1}|D_k] &= C_{i,k-i} \delta_{k-i+1} \cdots \delta_{k-1} \\ &+ E_i(\lambda_{k-i+1} \delta_{k-i+2} \cdots \delta_{k-1} \\ &+ \lambda_{k-i+2} \delta_{k-i+3} \cdots \delta_{k-1} \\ &\quad + \lambda_{k-2} \delta_k \\ &\quad + \lambda_{k-1}) \end{aligned}$$

According to Schnieper the expression  $E_i$  in (\*\*) gives the IBNYR reserve and the first term represents the known claims. The IBNER adjustment to the incurred claims reserve is obtained by subtracting  $C_{i,k-1}$  with  $C_{i,k-1}(\delta_{k-i+1} \cdots \delta_{k-1} - 1)$ . Schnieper's claim is true under the additional assumption: the incurred claims of accident year  $i$  are reported in development year  $k$ ,  $N_{ik}$  have the same expected development from  $k+1$  and onward as the incurred claims reported previously  $C_{i,k-1}$  have.

### Application to Motor Third Party Liability

We applied Chain Ladder Method and Schnieper Method using earned premiums as exposure for IBNYR. The data are based on the claims of Domestic Third Party Liability (DMTPL) in Albania. The figures are in Albanian currency.

Table 3. Incurred Cumulative Claims Triangle C

	0	1	2	3	4	5
2014	30,273,132	67,385,047	79,032,966	87,088,089	87,463,089	87,595,089
2015	55,611,007	121,915,467	163,464,758	168,784,419	179,238,100	
2016	62,814,708	139,145,260	184,552,320	192,870,185		
2017	57,729,322	120,551,662	145,054,785			
2018	67,727,288	128,383,880				
2019	84,944,477					

**Exposures**

Year	Earned Premium
2014	805,230,671
2015	857,010,533
2016	921,918,702
2017	892,394,125
2018	821,252,582
2019	831,252,582

**Results**

Year	Chain Ladder			Schnieper Method		
	Latest incurred	IBNR	Ultimate	IBNER	IBNYR	Ultimate
2014	50,370,320	-	50,370,320	-	-	50,370,320
2015	45,787,742	260,000	46,047,742	(442,442)	702,702	46,048,002
2016	48,814,766	(694,694)	48,120,072	(6,264,258)	5,569,564	48,120,072
2017	47,143,096	(108,108)	47,034,988	(3,817,814)	3,709,706	47,034,988
2018	43,349,306	1,107,106	44,456,412	(1,825,824)	2,932,930	44,456,412
2019	26,362,336	14,413,399	40,775,735	2,406,404	12,006,995	40,775,735
<i>Total</i>	<i>261,827,566</i>	<i>14,977,703</i>	<i>276,805,269</i>	<i>(9,943,934)</i>	<i>24,921,897</i>	<i>276,805,529</i>

**Conclusions**

The ultimate claim cost is the same for both methods. IBNER is mostly negative for the effect of claims handler costs. Since we have the same results the split of IBNER and IBNYR is not necessary. The need of having a tail is significant for the final best estimate. The result of the split is a reason to look into reserves and try to make them unbiased.

**References**

- Wüthrich, M.V., Merz, M. 2008. Stochastic Claims Reserving Methods in Insurance, Wiley
- Mack, T. 1997. Measuring the variability of chain–ladder reserve estimates. In: Claims Reserving Manual, vol. 2. London: Institute of Actuaries.
- Shniepper, R. 1991. Separating true IBNR and IBNER Claims. ASTIN Bulletin, Vol 21, No. 1.
- Liu, H., Verrall, P. 2009. Predictive distributions for reserves which separate true IBNR and IBNER claims. ASTIN Bulletin, 39(1), No. 35-60.
- Verrall, R., Nielsen J.P. and Jessen A. 2010. Prediction of RBNS and IBNR claims using claim amounts and claim counts, ASTIN Bulletin, 40(2), 871-887
- Ohlsson, E., 2015. Using separate exposure for IBNYR and IBNER in the Chain Ladder method



## **Determination of Direct and Indirect Effects of Weaning Weight in Anatolian Buffaloes**

Aziz ŞAHİN<sup>1</sup>, Zafer ULUTAŞ<sup>2</sup>, Arda YILDIRIM<sup>3</sup>, Yüksel AKSOY<sup>4</sup>

<sup>1</sup>Department of Animal Science, Faculty of Agriculture, Kırşehir Ahi Evran University, Kırşehir, Turkey

<sup>2</sup>Department of Animal Production and Technologies, Ayhan Sahenk Faculty of Agricultural Sciences and Technologies, Omer Halis Demir University, Niğde, Turkey

<sup>3</sup>Department of Animal Science, Faculty of Agriculture, Tokat Gaziosmanpasa University, Tokat, Turkey

<sup>4</sup>Department of Animal Science, Faculty of Agriculture, Osmangazi University, Eskişehir, Turkey

### **Abstract**

This research was conducted to determine the indirect and direct effects of birth weight, dam age and daily milk yield on weaning weight of Anatolian buffaloes. For this purpose, 2291 records of Anatolian buffaloes. The data used in the study were collected from Anatolian buffaloes that calved in Tokat province and its districts between 2012-2017. In the study, weaning weight as dependent variable, birth weight, dam age and daily milk yield as independent variables were considered. The correlation coefficients between weaning weight with birth weight, daily milk yield and dam age were determined as 0.248, 0,087 and 0.117, respectively. In the study, direct effects of birth weight, dam age and daily milk yield on weaning weight were determined as 0.229, 0.038 and 0.113 respectively. Indirect effects of independent variables on weaning weight were determined as 0.0524, 0.0014, and 0.0128 respectively. In present research, it was determined that weaning weight was significantly effected from birth weight, daily milk yield and dam age.

**Key words:** Direct effect, indirect effect, dependent variable, independent variable

## Changes in Statistics in the Big Data Age

Sadullah ÇELİK<sup>1</sup>

<sup>1</sup>Aydın Adnan Menderes University, Nazilli Faculty of Economics and Administrative Sciences,  
Department of Econometrics, Aydın, Turkey

### Abstract

After the 1990's, the internet, mobile phones and laptops started to enter our lives. With the introduction of these three technologies into our lives, there has been an exponential increase in the amount of information produced. This increase in the amount of information has created a new trend called Big Data. Big Data refers to large volume, velocity and diverse format data that is difficult or impossible to process and analyze using traditional computational methods. Big Data has brought new opportunities for organizations and governments and new challenges for statisticians. On the one hand, Big Data makes big promises to discover patterns, patterns and trends that cannot be detected with small datasets. On the other hand, the volume, velocity, dimension and variation of data bring along some computational and statistical difficulties such as scaling, storage, spurious correlation, random errors, measurement errors and random internality. The purpose of this study is to provide an overview of how the properties of Big Data (volume, velocity, variety, value and veracity) lead to a paradigm change in statistics and computational methods. In addition, new perspectives emerging in the field of big data analysis and computing are presented. Also, for Big Data, the exogenous assumption used in most statistical methods is not correct due to the random internality characteristic. This can lead to false statistical inferences and scientific results.

**Key words:** Population, big data, computational science, statistics, sampling

## Büyük Veri Çağında İstatistikte Yaşanan Değişimler

### Özet

1990'lı yıllarda sonra internet, cep telefonu ve dizüstü bilgisayarlar hayatımıza girmeye başladı. Bu üç teknolojinin hayatımıza girmesiyle beraber üretilen bilgi miktarında üstel bir artış yaşandı. Bilgi miktarındaki bu artış Büyük Veri denen yeni bir trendi karşımıza çıkardı. Büyük Veri, geleneksel hesaplama yöntemleri kullanarak işlenmesi ve analiz edilmesi zor veya imkânsız olan büyük, hızlı ve çeşitli formattaki verileri ifade eder. Büyük Veri, kuruluşlar ve hükümetler için yeni fırsatlar ve istatistikçiler için yeni zorluklar getirdi. Bir taraftan Büyük Veri, küçük veri setleri ile saptanamayan desenleri, kalıpları ve eğilimleri keşfetmek için büyük vaatler veriyor. Diğer taraftan, verinin büyüklüğü, hızı, boyutu ve çeşitliliği, ölçekleme, depolama, sahte korelasyon, rassal hatalar, ölçüm hataları ve tesadüfi içsellik gibi birtakım hesaplama ve istatistik zorlukları beraberinde getirmektedir. Bu çalışmanın amacı, Büyük Verinin özelliklerinin (veri büyüklüğü, hız, çeşitlilik, değer ve doğruluk) istatistik ve hesaplama yöntemlerinde nasıl bir paradigma değişikliğine yol açtığına dair genel bir bakış açısı sunmaktır. Bununla birlikte büyük veri analizi ve hesaplama alanında ortaya çıkan yeni perspektifler sunulmuştur. Ayrıca, Büyük Veri için, çoğu istatistiksel yöntemde kullanılan dışsal varsayımı tesadüfi içsellik özelliğinden dolayı doğru değildir. Bu yanlış istatistiksel çıkarımlara ve bilimsel sonuçlara yol açabilmektedir.

**Anahtar Kelimeler:** Anakütle, büyük veri, hesaplamalı bilim, istatistik, örneklem

**Factors Influencing the Educational Achievements of the Students in the University of Peshawar, Pakistan**

Muhammad IJAZ<sup>1</sup>, Rida BATOOL<sup>1</sup>, Syed Muhammad ASIM<sup>1</sup>

<sup>1</sup>Department of Statistics, University of Peshawar, Pakistan

**Abstract**

The main goal of the paper is to determine the factors affecting the student's academic achievements at the University of Peshawar. This study would help the students and teachers to identify the factors that improve the academic achievements of the students. A descriptive type questionnaire was designed to gather information from 264 students studying in various departments at the University of Peshawar, Pakistan. The study focuses on different factors including socio-economic factors, teacher's effectiveness and methodology of their teaching, social media, and the use of mobile phones, and parental factors. Chi-square statistics were used to study the impact of factors affecting the student academic achievements. The analysis concluded that except for the gender under socio-economic factors, teacher experience under teacher effectiveness factors, parent relationship with the student and their family income under parental factors, and the use of social media and mobile phones all the factors have an impact on the student academic achievements.

**Key words:** Students, socio-economic factors, parental factors, teacher effectiveness, social media, and mobile phones.

**Introduction**

No one should argue that a country's well-trained graduates have a key role to play in its economic stability and social growth. The universities are responsible for delivering the highest educational standard and producing well-trained graduates. The universities should undertake research studies to highlight the factors influencing the academic achievements of the students, thereby placing the students on the right track and the students themselves to earn good grades in order to catch incentives and meet business demands.

Many research studies were conducted to highlight the factors influencing student academic achievements, for example, Abubakar & Oluwatoyin (2017) worked on the effects of teacher motivational techniques on students and academic achievements. The strategies for motivation included scholarship, awards, lesson assessment, encouragement, good guidance and counseling, and teaching skills. The study concluded that a significant relationship exists between teaching strategies and academic success between students. Akahome, Rilwani, & Gbakeji (2014) outlined the factors affecting the regional attrition of secondary school students. The study of Akiri (2013) determined that the teacher effectiveness is not only the factor to influence the student grades. Alos & David (2015) concluded that the teacher-based variables are strongly related to student success relative to research behaviors and student parental influences. Ayuba & Mohammed (2014) considered course curriculum, department role, and support facilities to study its effect on academic performance among students. The study revealed that departmental position is highly linked to academic achievements of students. For other studies the references are there given at the end of the paper.

The current study focuses on the factors effecting the student academic achievements of the University of Peshawar. There are four types of factors are included in the study such as socio-economic factors, teacher related aspects, social media and mobile use, and parental factors.

### **Methodology**

The data were collected through a well defined descriptive type questionnaire from total of 264 respondents consisting of male and female students. The data were gathered on various factors, in socio-economics, the factors are include gender of the respondent, study program, marital status, interest field, relax in co-education, and class regularity. Teacher related aspects are the teaching methodology, teaching methodology effect, teacher experience, and class environment. The parental and family factors are include parent education, parent involve in their child education, student relation with their parents, and family income. The last factors are the social media and the use of mobile phones. The data were analyzed through SPSS software.

### **Research Objectives**

The following are the fundamental goals of the research study

1. To study the impact of socio-economic factors on student academic achievements
2. To explore the impact of Teacher related aspects on student academic achievements
3. To identify the impact of parental factors on student academic achievements
4. To study the relationship of the social media and mobile phones on student academic achievements

### **Research Hypothesis**

The current research study will focuses on the following research hypothesis

1. There is an insignificant relationship between the socio-economic factors and student academic achievements
2. There is an insignificant relationship between the Teacher related factors and student academic achievements
3. There is an insignificant relationship between the Parental factors and student academic achievements
4. There is an insignificant relationship between the use of social media and student academic achievements

### **Results and Discussion**

#### ***Socio-economic factors affecting student academic achievements***

Table 1 demonstrates the Chi-square analysis of the socio-economic factors and its impact on student academic achievements. The result showed that there is no significant relationship between the gender ( $\chi^2=0.881$ ,  $p=0.830$ ), co-education ( $\chi^2=2.638$ ,  $p=0.451$ ) and student academic performance at 5% level of significance. It means that there is no role of the gender and co-education in improving the student academic achievements. However, the program of study that is Bs or Ms ( $\chi^2=8.602$ ,  $p=0.035$ ), marital status ( $\chi^2=9.655$ ,  $p=0.022$ ), field of interest ( $\chi^2=8.848$ ,  $p=0.031$ ), and class regularity ( $\chi^2$

=11.808, 0.008) shows a significant relationship with the student academic achievements at 5% level of significance which implies that these factors play a key role in improving the grades of the students.

Table1. Socio-economic Factors Affecting Student Academic Achievements

Variable	$\chi^2$	P-value	Sig/Insig
Gender	0.881	0.830	insignificant
Program of study	8.602	0.035	significant
Marital status	9.655	0.022	significant
Field of interest	8.848	0.031	significant
Co-education	2.638	0.451	insignificant
Class regularity	11.808	0.008	significant

### *Teacher related factors affecting student academic achievements*

Table 2 defines the impact of teacher related factors on the student grades. Among others the teacher experience ( $\chi^2=0.709$ ,  $p=0.871$ ) shows an insignificant relationship with the performance of the students at 5% level of significance. Whether the teacher is less experienced or more but if their teaching methodology is easy to understand then this will improve the grades of the students. While the teaching methodology ( $\chi^2=15.557$ ,  $p=0.001$ ), effect of teaching methodology ( $\chi^2=9.058$ ,  $p=0.029$ ), and class environment ( $\chi^2=9.134$ ,  $p=0.028$ ) play a significant role in improving the student grades. In the University of Peshawar, most of the teachers even in numerical sciences use multimedia for teaching due to which the students face many problems in their theoretical work. This is the main reason that the students of this discipline gain low grades and those students who are engaged with lectures presented on a white-board gain high grades.

Table2. Teacher Related Factors Affecting Student Academic Achievements

Variable	$\chi^2$	P-value	Sig/Insig
Teaching methodology	15.557	0.001	significant
Effect of teaching methodology	9.058	0.029	significant
Teacher experience	0.709	0.871	insignificant
Class environment	9.134	0.028	significant

### *Parental factors affecting student academic achievements*

The results given in table 3 shows that there is no effect of the student parents relation ( $\chi^2=3.883$ ,  $p=0.275$ ) and family income ( $\chi^2=5.256$ ,  $p=0.511$ ) on the academic achievements of the students. Meanwhile, the parent education ( $\chi^2=40.17$ ,  $p=0.000$ ) plays a significant role in improving their children education performance. Similarly, those students got more grades whose parents are involved in their children education ( $\chi^2=12.97$ ,  $p=0.005$ ) activities. The performance of the students whose families are well-educated and their parents are involved in their children academic activities are better than of those who families are illiterate.

Table 3. Parental Factors Affecting Student Academic Achievements

Variable	$\chi^2$	P-value	Sig/Insig
Parent education	40.17	0.000	significant
Parent involve in their children education	12.97	0.005	significant
Student parents relation	3.883	0.275	insignificant
Family income	5.256	0.511	insignificant

### *Social media and mobile phones affecting student academic achievements*

The results given in table 4 implies that there is no relationship between social media ( $\chi^2=2.103$ ,  $p=0.551$ ) and the use of mobile phones ( $\chi^2=2.870$ ,  $p=0.420$ ) with the student academic performance at the 5% level of significance. The use of social media and mobile phones would not guarantee the students low grades but the time spent on it which is not included in this study. Those students are going to bear low grades that have spent more time on social media and Smartphone.

Table 4. Social Media and Mobile phones Affecting student academic achievements

Variable	$\chi^2$	P-value	Sig/Insig
Social Media	2.103	0.551	insignificant
Mobile Phones	2.870	0.420	insignificant

### **Conclusion**

The result of the analysis can be summarized as follows.

The study program, marital status, area of interest, and class regularity are the significant determinants of the academic achievements of the student in socio-economic factors.

In teacher related factors, teaching methodology, effect of teaching methodology, and the class environment are the reasons which upgrades the student performance.

Those students got high grades whose parents are well-educated and who are also active in educational activities for their children.

Based on current data, social media and mobile phone use have no impact on academic performance of the students.

### **References**

- Abubakar, Y. A., & Oluwatoyin, F. C. (2017). Effect of Teachers' motivational Strategies on Students' academic Achievements: Experience From Nigeria. *MOJEM: Malaysian Online Journal of Educational Management*, 3(2), 70-81.
- Akahomen, D. O., Rilwani, M. L., & Gbakeji, J. O. (2014). Factors Influencing Secondary School Students' Attrition In Geography In Esan West Local Government Area, Edo State, Nigeria. *Journal Of Emerging Trends In Educational Research And Policy Studies*, 5(2), 154-161.
- Akiri, A. A. (2013). Effects Of Teachers' Effectiveness On Students' Academic Performance In Public Secondary Schools; Delta State-Nigeria. *Journal Of Educational And Social Research*, 3(3), 105.
- Alos, S. B., Caranto, L. C., & David, J. J. T. (2015). Factors Affecting The Academic Performance Of The Student Nurses Of BSU. *International Journal Of Nursing Science*, 5(2), 60-65.
- Ayuba, B., & Mohammed, N. (2014). Investigating The Factors Affecting Performance Of Business Management Students In Nigerian Universities. *IOSR Journal Of Business And Management (IOSR-JBM)*, 16(8), 11-20.
- Das, S., Das, M., & Das, T. (2014). Impact Of Socio-Economic And Academic Background On Students' Achievement In Higher Education: A Study Of Assam University. *Journal Of Global And Scientific Issues*, 2(4), 5-12.

- Diaz, A. L. (2003). Personal, Family, And Academic Factors Affecting Low Achievement In Secondary School. *Electronic Journal Of Research In Educational Psychology And Psychopedagogy*, 1(1), 43-66.
- Ekperi, P. (2018). Impact Of Teacher Characteristics On Students' Academic Performance In Public Secondary Schools. *International Journal Of Research And Innovation In Social Science (IJRISS)*, 2(12), 514-519.
- Harb, N., & El-Shaarawi, A. (2006). Factors Affecting Students' Performance (Munich Personal Repec Archive Paper No. 13621). Accessed On November, 2, 2011.
- Martha, K. (2009). Factors Affecting Academic Performance Of Undergraduate Students At Uganda Christian University. Unpub. Master Of Arts In Educational Management Dissertation, Makerere University, Uganda.
- Olufemioladebinu, T., Adediran, A. A., & Oyediran, W. O. (2018). Factors Influencing The Academic Achievement Of Students' In Colleges Of Education In Southwest, Nigeria. *Journal Of Education And Human Development*, 7(3), 109-115.
- Omotayo, B. K. (2014). Teachers' Characteristics And Students' Performance Level In Senior Secondary School Financial Accounting. *Journal Of Empirical Studies*, 1(2), 48-53.
- Rasul, S., & Bukhsh, Q. (2011). A Study Of Factors Affecting Students' Performance In Examination At University Level. *Procedia-Social And Behavioral Sciences*, 15, 2042-2047.
- Umoh, P. (2020). Teachers Characteristics And Academic Performance Of Students With Hearing Impairment In Special Education Centre Ibom Layout, Cross River State. *Cross River State* (April 1, 2020).

## The Cohort Component Method Applied in the Albanian Population Projection

Brikena XHAJA<sup>1</sup>, Kleida HAXHI<sup>2</sup>, Eglantina KALLUÇI<sup>3</sup>

<sup>1</sup>Polytechnic University of Tirana, Faculty of Mathematical Engineering, Albania

<sup>2</sup>Polytechnic University of Tirana, Physical Engineering, Albania

<sup>3</sup>University of Tirana, Faculty of Natural Sciences, Albania

### Abstract

Most developed countries in Europe have experienced a series of huge demographic changes since the second quarter of the 20th century. The demographic situation in our country and in our region lately has been studied with particular interest. In general, the purpose of demographic projections is to review three major demographic trends: births, mortality, fertility, and migration, which will shape the size, age structure, and distribution of tomorrow's population. It then examines key policy issues that will be impacted by these trends: global poverty, public finance and climate change. The Cohort component method is one of the most common methods for population projection. It is widely used and it provides flexibility in population projections. A Cohort is defined as a group of people who experience the same demographic event over a period of time. This paper aims to design the population projection in Albania using the Cohort component method.

**Key words:** Cohort component method, population projection, mortality, mortality rates, fertility rate, migration, survival rate

### Introduction

Demographic features are of key importance to development, but this link sometimes is ignored. Debates about population policy are stirring, with columnists and academics arguing about what lies ahead if global population challenges aren't actively integrated into policy and planning processes. The study of people using the tool of demography is now appearing across development discourse, with policy implications that reach far beyond family planning and reproductive health.

Population IS undeniably important, but how, for whom, and with what consequences is a complex story. One thing is certain that population issues in the 21st century are different from those in the last century.

This work presents an overview of the impact that demographic changes will have on development over the first half of the 21st century by taking a close look at three demographic trends: fertility, mortality, and immigration; and examining how these will touch policy issues including poverty, public finance and infrastructure, and climate change.

In this work also examines key policy issues that will be impacted by these trends: global poverty, public finance and climate change.

From 1950 to 1990, the Albanian population experienced decades of rapid population growth, although the pace of this growth has declined in a progressive way. The fertility rate in spite of the fall was much higher than anywhere else in Europe. Mortality rate was relatively favorable, in line with a Mediterranean- type pattern, exchanges with the foreign countries through migration were almost null and movements within the country were severely limited. It was easy to forecast the future of the population by an extrapolation of ongoing trends at any stage of this smooth process.



The 1990s dramatically modified this picture. Fertility decline accelerated to levels much closer to the European standards, external migration has been massive and has strongly affected the age pyramid whereas internal migration has deeply modified the balance between the country's regions.

Besides these difficulties, population projections remain highly reliable due to the inertia that characterizes most demographic trends. In 2021, twenty years after the most recent Albanian census, one may estimate that some 70% of the population in the country at that time must have been already present and counted. There would be few uncertainties about this large nucleus of the future population and in particular, we will know for sure the age of these persons.

Population projections offer a basic tool for social and economic planning. The future population of Albania is not described in a unique forecast that would give the false and comfortable feeling of certainty. A diversity of future trends is proposed, among which there will be given some priority to the trends of demographic growth and ageing, but their multiplicity tell us that the future is still insecure.

### ***Demographic features***

The number of inhabitants is modified each year by population movements. Persons who die and those who emigrate all contribute to the reduction of the population whereas newborns and immigrants lead towards its growth. The number of deaths in a year depends simultaneously on population size and mortality risk. Exposed to similar health conditions, a larger population experiences more deaths than a smaller one. But sanitary conditions also play a part: deaths are more numerous when these conditions are bad rather than good.

The very role of demographic analysis is to identify the impact of the two factors through the calculation of mortality risks: the ratio of the number of deaths to the number of inhabitants.

And in the same way is calculated also the risk posed by emigration: the ratio of the number of emigrants to the number of inhabitants. By nature, these ratios are “independent from population size”. They can be compared from one country to another; they measure then differences in behaviors regarding emigration or the differences in health conditions.

Population projections put things upside down in order to calculate the future number of births, deaths, and emigrants on the basis of information on the population size on the one hand and a prognosis of the rates or risks, on the other.

The question of migration needs a specific treatment: Albania in the last years has faced the problem of internal migration, i.e. persons leaving one place for another one inside the country. Calculations are conducted on an annual basis for the whole country starting from the last census date, where the present population for the whole country is known disaggregated by districts, sex, age and, if needed, by a lot of other characteristics. The number of inhabitants has been calculated as of the 1st of January 2001 (by sex, age and district) to be the actual departure point. Afterwards it is moved forward in two steps:

1. The present population of each age (and sex) gets one year older and drops the persons that died and emigrated, according to the mortality and emigration rates;
2. Newborn boys and girls in the current year are added, except for those who died or emigrated in their first year before the 1st of January 2002.

They are calculated on the basis of fertility rates and the number of women in fertile ages on one hand and infant mortality and emigration rates, on the other.

Table1. Population Change 2001-2017

<b>Town</b>	<b>2001</b>	<b>2017</b>	<b>Absolut Change</b>	<b>Relative change</b>
Berat	192,667	131,942	-60,725	-32%
Dibër	189,275	125,579	-63,696	-34%
Durrës	244,753	284,823	40,070	16%
Elbasan	361,980	283,822	-78,158	-22%
Fier	381,841	302,507	-79,334	-21%
Gjirokastrë	112,726	65,939	-46,787	-42%
Korçë	264,902	214,321	-50,581	-19%
Kukës	110,858	79,559	-31,299	-28%
Lezhë	158,920	129,019	-29,901	-19%
Shkodër	255,898	207,924	-47,974	-19%
Tiranë	596,704	862,361	265,657	45%
Vlorë	192,796	188,795	-4,001	-2%
<b>Total</b>	<b>3,063,320</b>	<b>2,876,591</b>	<b>-186,729</b>	<b>-6%</b>

### ***Fertility***

The term fertility refers to the occurrence of a live birth (or births) to an individual, a group, or an entire population. It is determined by a combination of biological, social, psychological, economic, and cultural factors. Biological factors affect fecundity, whereas social, psychological, economic, and cultural factors affect choices regarding whether to have children, how many to have, and when to have them. Although biological factors set an upper limit on a woman’s lifetime fertility, most women bear children at levels far below that limit. The broad array of factors that affect personal choice is thus paramount in the study of fertility.

Fertility rates vary considerably among individuals and populations and change over time.

The size of the global population is most affected by fertility rates in countries with moderate to high population growth. At the global level, fertility rates have been falling since the last decades of the 20th century as a result of the (largely) voluntary choices of couples to limit the number of children born. The average number of children per woman fell from 5 in 1950 to 2.7 in 2005 (Bogart’s, John and Rodolfo Balata, editors. 2000).

The total fertility rate (TFR) indicates the number of children a woman if she passed through the childbearing years at the current age - specific fertility rates.

**TFR= number of births per month / number of marriages per month**

The age-specific fertility rates  $f_{(x,t)}$  can be calculated in this way, putting in the denominator only the number of women: at age  $x$  in year  $t$ , the births  $B_{(x,t)}$  to mothers aged  $x$  in year  $t$  are related to the average of the numbers  $F_{(x,t)}$  and  $F_{(x,t+1)}$  of women aged  $x$  on the first of the two Januarys surrounding  $t$ .

$$f_{(x,t)} = \frac{B_{(x,t)}}{(F_{(x,t)} + F_{(x,t+1)}) / 2}$$

The TFR (or TPF—total period fertility rate) is a better index of fertility than the Crude birth rate (annual number of births per thousand population) because it is independent of the age structure of

the population, but it is a poorer estimate of actual completed family size than the total cohort fertility rate, which is obtained by summing the age-specific fertility rates that actually applied to each cohort as they aged through time. In particular, the total fertility rate (TFR) does not necessarily predict how many children young women now will eventually have, as their fertility rates in years to come may change from those of older women now. However, the TFR is a reasonable summary of current fertility levels.

The total fertility rate for Europe has been estimated to be 1.5 children per woman (Wikipedia).

Actually it takes a total fertility rate of 2.1 or 2.2 to replace each generation, this number is called the replacement rate, because some children will die before they grow up to have their own two children. In countries with low life expectancies, the replacement rate is even higher 2.2–3.

Fertility in Albania underwent major changes during the communist era. Total fertility rose during the 1950s, reaching a peak of almost seven children per woman by 1960, before beginning a steady decline to little over five in 1970, to less than four in 1980, and to just over three in 1990. While the reduction of fertility was not as rapid as in some other countries, it was nevertheless substantial, especially when compared to countries with a similar level of economic development.

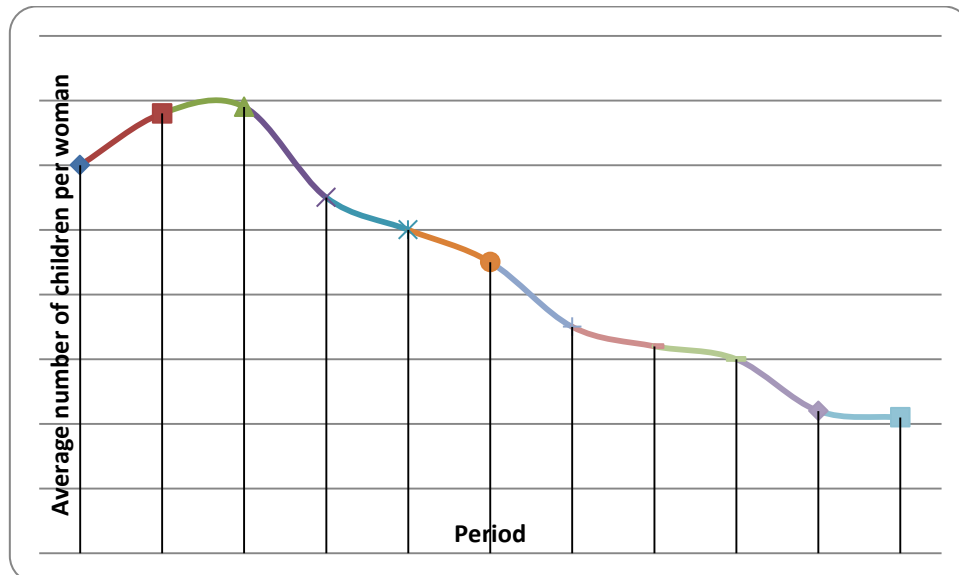


Figure 1. Total Fertility Rate, 1950-2000 INSTAT

The early years after the Second World War were marked by increased fertility. This might be compared with other postwar baby – booms, but it might also be a pre-decline rise in fertility, which was a widespread aspect of fertility transition.

Fertility fell mostly at older ages and for teenagers. At older ages has happened because at that time the couples had starting using family planning to prevent childbearing and for teenagers the reason is because the age of getting married has increased. Taken together, the patterns in age – specific and total fertility rates suggest a significant increase in the use of deliberate birth control, in spite of the absence of any organized family planning programs. This was of particular importance for fertility reduction in Albania during the communist period.

Despite this reduction Albania still has the highest fertility rate in Europe. This might not continue for long. If fertility continues to decrease in the coming years at the same rate, Albania will experience the same reduction of fertility below replacement level found in other south – eastern

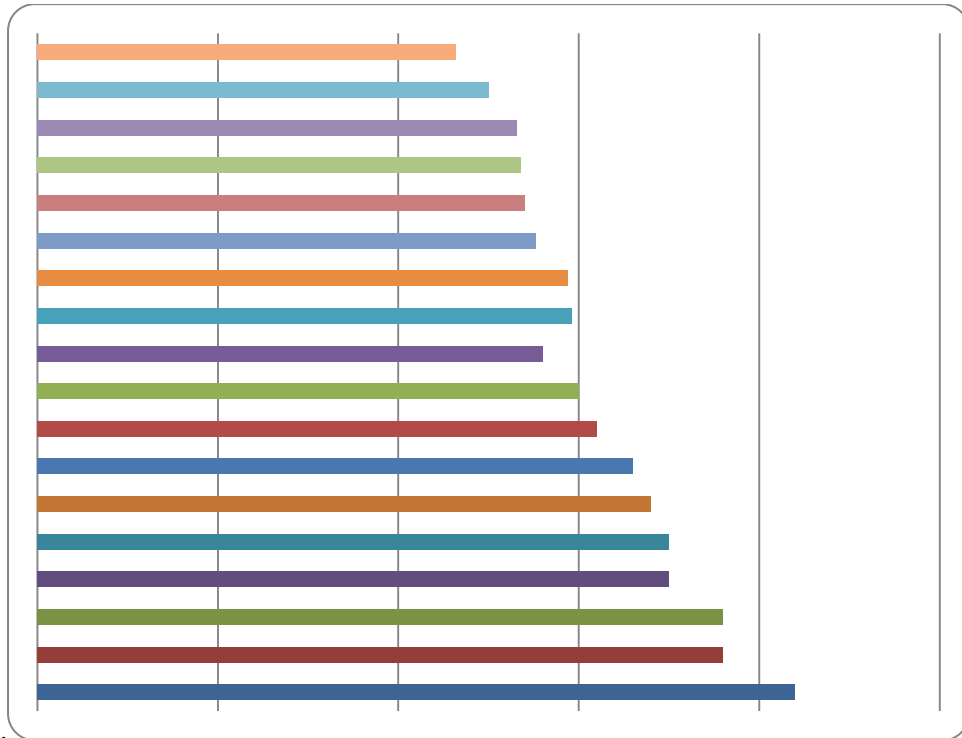


Figure 2. Total Fertility Rate in European Countries, 2000 (INSTAT)

In summary, it may be said that compared with almost all other European countries, the decline in fertility was late to arrive. On the other hand, compared with many non European countries at similar levels of economic development the rate of change was reasonably rapid.

We came to the conclusion that the number of marriages is not a direct variable that affects the number of births in Albanian population but we can say that based on this parameter, the total fertility rate in Albania has decreased and for the coming years it will be stationary with seasonal patterns.

In the figure below it is shown the gap that currently exists, but is projected to narrow, between total fertility in Europe and total fertility in every other major area. A gap in fertility levels also exists between the other major areas of the world, but by 2050 levels are expected to converge in a narrow band between 1.84 and 1.92 children per woman. Europe will take the longest to enter this band, and will do so through rising fertility, in contrast to falling fertility in other major areas.

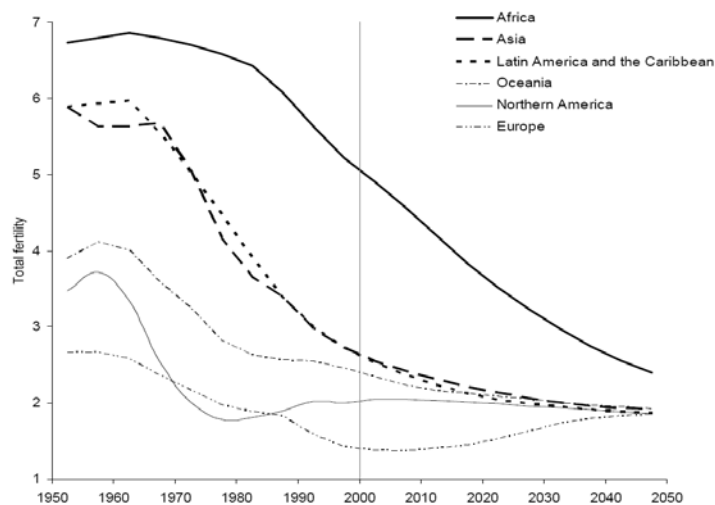


Figure 3. Total Fertility Rates, 1950 – 2050

Mortality exerts some additional influence on the growth rate. Across major areas, its effect on growth largely it works against that of fertility, since where fertility is higher, mortality also tends to be higher. Over time, life expectancy is expected to rise fairly smoothly.

Furthermore, fertility depends on three elements: women's reproductive capacity, men's reproductive capacity, and the meeting of these two aptitudes through couple formation, with all three elements depending on age.

### ***Mortality***

Mortality rate is a measure of the number of deaths (in general, or due to a specific cause) in some population, scaled to the size of that population, per unit time. Mortality rate is typically expressed in units of deaths per 1000 individuals per year; thus, a mortality rate of 9.5 in a population of 100,000 would mean 950 deaths per year in that entire population, or 0.95% out of the total. It is distinct from morbidity rate, which refers to the number of individuals in poor health during a given time period (the prevalence rate) or the number of newly appearing cases of the disease per unit of time (incidence rate).

Mortality modeling has a very long history. Numerous models have been proposed since Gompertz published his law of mortality in 1825. Mortality forecasting is a more recent effort.

Transformation of the data is used to assist modelling. The use of logarithms of rates is common in mortality modelling, one advantage in forecasting being that rates are necessarily positive. Mortality reduction (or improvement) factors<sup>1</sup>, traditionally used by actuaries, offer the same advantage. Brass (1971) used the logit transformation

$$\log it(1 - s_x) = \frac{1}{2} \ln\left(\frac{1 - s_x}{s_x}\right)$$

where  $s_x$  is the probability of surviving from birth to exact age  $x$ , to achieve approximate linearity in relating observed to standard age patterns. Heligman & Pollard (1980) experimented with various measures. Wilmoth (1990) used a transform based on age-specific probabilities of death,  $q_x$ , that approximates log hazard rates:

$$f_x = \log\left(\frac{q_x}{1 - \frac{1}{2}q_x}\right) \approx \log \mu_{x+\frac{1}{2}}$$

Other useful transformations, or link functions, are considered by Renshaw & Haberman (2000). In an innovative approach, de Jong & Marshall (2007) adopt the Wang transform

$$z_x = \Phi^{-1}(s_x)$$

where  $\Phi$  denotes the cumulative normal distribution (Wang, 2000).

### ***Albanian Mortality***

The analysis about Albanian mortality is based on the life tables calculated for the period 1950 to 2000. The table shown below will give the mortality indicators. Life expectancy at birth is taken as an indicator for the overall mortality. The probability of dying at age 0 – 1 is taken as an indicator for the infant and child mortality and life expectancy at the age 15 is taken as an indicator for the adult mortality. Looking at the table below it is clear that the period from 1950 to 2000 was significant in terms of reducing mortality in Albania. So it can be said that in a short period of 50 years life expectancy at birth is increased by 23 years. Most of the improvements in life expectancy at birth occurred in the first decade,

when life expectancy increased by 10.4 years, which means that the increase is more than one year for each year of the decade. The reasons for this major improvement within a decade is that during this period there was a large reduction of mortality rates at all ages from infectious and parasitic diseases, which were the main causes of death before the war in Albania. Mortality continued to improve during the following four decades, although not at the same rate as it is seen for the first decade: from 1969 to 1989 life expectancy at birth improved by 3.5 years.

Although it is not a major improvement compared to the first two decades, it is still a significant and moreover a continuing trend. This trend becomes even more significant when compared to the experiences of other eastern European countries during the eighties, which saw an increase in mortality, especially among adults: for instance in Russia life expectancy fell below 70 years in 1990.

Table 2. Mortality Indicators, 1950 -2000 (INSTAT)

Years	Life Expectancy at Birth	Life expectancy at 15 years	Infant Mortality rate	Child mortality rate
1950	51.6	53.9	143.1	105.3
1955	55	56.3	103.9	124.3
1960	62	58.6	96.6	58.6
1965	64.1	58.6	92.9	34.2
1970	66.5	60.1	89.1	22.2
1975	67	60	84.8	18.9
1980	68	60.1	74.1	17.4
1985	70.2	60	46.4	14.4
1990	70.7	60.6	45.4	15.5
1995	73.1	60.6	26.4	12
2000	74.6	62.4	22.8	9.1

International comparisons show that in 2000 Albania has very high life expectancy with 71.5 years for males and 78.1 for females despite its level of economic development. For example in Hungary the respective values are 67.7 years for males and 75.7 years for females. And even in the countries richer and wealthier than Albania show a life expectancy not much higher than Albania, such as Ireland with 73.0 years for males and 75.7 for females.

The two components of mortality that have experienced these major improvements are infant and child mortality. Infant and child mortality in post- Second World War Albania were very high: infant mortality was 143.1 deaths per thousand and child mortality 105.3 per thousand. In 1950 infant as well as child mortality were very high as compared to neighboring countries: Bulgaria and former Yugoslavia had respectively an infant mortality of 82 (in 1953) and 43 (in 1950) per thousand.

### ***Migration***

Migration has an effect on health and survival. It is quite often motivated by the desire to improve living conditions and, in the case of refugees and asylum seekers, even to escape from life threatening situations. The act of migrating itself often involves an important part of the risk, especially in the case of forced (refugees) or clandestine migration. The risks undertaken are sometimes enormous, whether they concern long journeys plagued with difficulties.

The problem is that the data required calculating these rates are rarely available. In most cases, the data are from retrospective questions asked during a census. Even if they are of good quality, that only gives, for international migration, numbers of arrivals but not of departures.

### Implementing the Cohort Component Method

We have now discussed mortality, fertility, and migration—the three components of population change. We provide a combination of these components in a complete projection model. Cohort component method is based in demographic features > births, deaths and migration. The procedure for making population projections with the cohort component was developed by Whelpton in the 1930s. It can be thought of as an elaboration of the ideas contained in the balancing demographic equation:

$$P(t+n) = P(t) + B(t) - D(t) + I(t) - E(t)$$

Where:

- P(t) population in time t
- B(t) and D(t) birth and deaths occurred in time t and t+n.
- I(t) and E(t) immigrants and emigrants during the period t and t+n.

The Kohort component method is used when population projections by age and gender are required for 5 years, 10 years or more time periods. This projection tool allows planners to consider the future needs of different segments of the population including the needs of children, women in their reproductive years, people in the workforce, and the elderly. It also allows planners to project the overall size of the population. The results can be used in all aspects of local and regional development plans. What my thesis topic consists of is the implementation of population projection with the Cohort component method in practice. The necessary data for the population projection for 2015 are:

- Population by age and gender for the years 1995, 2005 and 2015.

Based on these data, the birth rate, mortality rate and survival rate will be calculated

Age	Population		Survival Rate		Birth Rate	Births	2000		Births 2000-2004	2005	
	Male	Female	Male	Female			Male	Female		Male	Female
0-4	141,686	132,263	0.993095937	0.99299562			94,263	90,567		102,637	98,612
5-9	157,555	149,497	0.997852299	0.997988776			140,708	131,337		93,612	89,932
10-14	165,151	158,979	0.997981753	0.997895202			157,217	149,196		140,406	131,072
15-19	143,573	149,532	0.995158293	0.995310355	0.0197122	14,738	164,818	158,644	15,636	156,899	148,882
20-24	110,284	120,252	0.992767818	0.992493471	0.078144965	46,985	142,878	148,831	58,152	164,020	157,900
25-29	102,024	110,902	0.994113831	0.994434773	0.112830945	62,566	109,486	119,349	67,331	141,845	147,714
30-34	104,197	111,234	0.994113831	0.994434773	0.077353408	43,022	101,423	110,285	42,655	108,842	118,685
35-39	110,207	111,978	0.992291378	0.993651135	0.027549981	15,425	103,584	110,615	15,237	100,826	109,671
40-44	107,921	105,509	0.992291378	0.993651135	0.003714	1,959	109,357	111,267	2,066	102,785	109,913
45-49	86,712	82,235	0.986154926	0.987067315	0.000326707	134	107,089	104,839	171	108,514	110,561
50-54	71,810	67,977	0.986154926	0.987067315			85,511	81,171		105,606	103,483
55-59	62,506	58,945	0.97782362	0.964411228			70,816	67,098		84,328	80,122
60-64	57,960	54,997	0.97782362	0.964411228			61,120	56,847		69,245	64,710
65-69	43,936	41,800	0.893360979	0.874100154			56,675	53,040		59,764	54,824
70-74	31,495	33,155	0.893360979	0.874100154			39,251	36,537		50,631	46,362
75-79	16,486	22,055	0.280400648	0.471374532			28,136	28,981		35,065	31,937
80-84	9,370	14,271	0.280400648	0.471374532			4,623	10,396		7,889	13,661
85+	4,625	10,241	0.011824316	0.329523957			7,252	10,102		1,382	8,229
					<b>Total</b>	<b>184,830</b>		<b>Total</b>	<b>201,249</b>		

**1<sup>st</sup> International Applied Statistics Conference (UYIK-2020)**  
**Online – Tokat / Turkey, 1-4 October 2020**

Population Projection 2015		
Age	Male	Female
0-4	121,588	110,343
5-9	135,646	127,115
10-14	136,007	128,170
15-19	141,925	137,721
20-24	110,185	116,064
25-29	104,082	101,306
30-34	99,732	103,116
35-39	82,412	87,378
40-44	92,936	95,647
45-49	84,354	87,446
50-54	78,590	80,386
55-59	76,766	76,296
60-64	66,084	68,275
65-69	51,421	52,829
70-74	40,089	47,017
75-79	31,558	38,006
80-84	15,339	18,174
85+	6,257	8,603
<b>Total</b>	<b>1,474,971</b>	<b>1,483,892</b>

Albanian population (INSTAT) 2015		
Age	Male	Female
0-4	89,354	81,733
5-9	90,100	81,581
10-14	107,351	98,587
15-19	124,592	119,875
20-24	136,017	125,216
25-29	118,551	99,299
30-34	94,216	86,874
35-39	79,858	82,432
40-44	83,375	89,921
45-49	92,343	97,947
50-54	102,602	104,160
55-59	95,342	95,922
60-64	74,293	74,993
65-69	58,547	59,621
70-74	49,491	50,658
75-79	37,061	39,077
80-84	19,144	22,506
85+	8,962	14,195
<b>Total</b>	<b>1,461,199</b>	<b>1,424,597</b>

Age	Popullsia		Normat e mbijetesës		Birth rate	Births	2010		Births 2010-2014	Projection 2015	
	Male	Female	Male	Female			Male	Female		Male	Female
0-4	123,302	111,990	0.993798768	0.992614839			88,159	84,701		100,923	96,965
5-9	141,647	132,971	0.997593479	0.998061725			122,537	111,163		87,612	84,076
10-14	154,170	148,295	0.99813419	0.997289296			141,306	132,713		122,242	110,947
15-19	157,782	154,250	0.996131302	0.996062006	0.0197122	15,203	153,882	147,893	14,576	141,042	132,354
20-24	120,918	126,654	0.99206762	0.993720764	0.078144965	49,487	157,172	153,643	60,032	153,287	147,311
25-29	90,002	96,342	0.991548859	0.991775766	0.112830945	54,352	119,959	125,859	71,004	155,925	152,678
30-34	89,629	97,028	0.991548859	0.991775766	0.077353408	37,527	89,241	95,550	36,955	118,945	124,824
35-39	94,751	102,285	0.99177657	0.993184657	0.027549981	14,090	88,872	96,230	13,256	88,487	94,764
40-44	107,580	109,985	0.99177657	0.993184657	0.003714372	2,043	93,972	101,588	1,887	88,141	95,574
45-49	99,669	97,111	0.987455129	0.988037099	0.000326707	159	106,695	109,235	178	93,199	100,896
50-54	78,840	75,941	0.987455129	0.988037099			98,419	95,949		105,357	107,929
55-59	63,910	61,616	0.966632991	0.973445936			77,851	75,033		97,184	94,801
60-64	60,076	59,945	0.966632991	0.973445936			61,778	59,980		75,253	73,040
65-69	51,469	49,266	0.936814418	0.947291773			58,071	58,353		59,716	58,387
70-74	36,318	38,102	0.936814418	0.947291773			48,217	46,669		54,402	55,278
75-79	21,453	25,734	0.367144914	0.434189397			34,023	36,094		45,170	44,209
80-84	10,737	16,163	0.367144914	0.434189397			7,876	11,173		12,491	15,672
85+	4,636	9,067	0.027325602	0.299418232			4,069	9,733		3,003	7,766
<b>Total</b>						172,860	<b>Total</b>	197,888			



	<b>2015 Population</b>
<b>Total population projection provided by Cohort component method</b>	<b>2,958,863</b>
<b>Actual population in Albania</b>	<b>2,885,796</b>
<b>Change (under/ over estimated)</b>	<b>73,067</b>

Once the population projection with the Cohort component model is completed compare the projected population and that of 2015 obtained from INSTAT in two ways. The first by making the difference between the projected and the real population. The second way is by building the population and by making a comparison of the projected population and the real one.

### **Conclusions**

The cohort-component method is a mainstay in the demographer’s toolbox. It provides a theoretically complete model that accounts for the individual components of growth and for the impact of changes in demographic composition over time.

It can incorporate many different application techniques, types of data, and assumptions regarding future trends. It provides projections of the components of growth and changes in demographic composition, as well as projections of total population.

A final limitation of the cohort-component method is that—although it provides the mathematical framework for making projections for cohorts and components of growth—it provides no guidance regarding the choice of assumptions that will lead to reasonable forecasts. How much will mortality rates decline over the next 20 years? Will fertility rates go up or down? Will migration follow the patterns observed over the last 10 years or revert to the patterns observed during the previous 10 years? What economic, social, psychological, political, or biological factors might cause recent demographic trends to change course?

Nothing in the cohort-component method itself provides answers to these questions.

### **References**

- Stanley S. Jeff T. David S. (2002) State and Local Population Projection  
 Barsotti O. Bonaguidi A. (1999), Previsioni demografiche: ipotesi sulle migrazioni  
 Bongaarts J., Burch T. e Wachter K. , Family Demography: Methods and Their Application  
 Lutz, W. (1996), The Future Population of the World. What Can We Assume Today?  
 Preston S., Heuveline P. e Guillot M. (2001), Demography. Measuring and modelling population processes.  
<https://www.measureevaluation.org/resources/training/online-courses-and-resources/non-certificate-courses-and-mini-tutorials/population-analysis-for-planners/lesson-8>  
<http://www.instat.gov.al/al/home.aspx>  
<http://www.worldbank.org/en/country/albania>

## Clustering Using Genetic Algorithm in Data Mining

Erkut TEKELİ<sup>1</sup>, Özlem AKAY<sup>2</sup>, Güzin YÜKSEL<sup>3</sup>

<sup>1</sup>Vocational School of Kozan, Çukurova University, 01530 Kozan, Adana, Turkey

<sup>2</sup>Department of Statistics, Çukurova University, 01330 Sarıçam, Adana, Turkey

### Abstract

Commonly used in data mining, cluster analysis is the process of separating  $n$  units with  $m$  features into  $k$  subset in accordance with various algorithms and available information. These various algorithms used in cluster analysis can give different cluster results for the same data set. In this case, it is necessary to find the most appropriate algorithm where the units are clustered correctly. In this study, in order to ensure that the units are in the right clusters, cluster analysis was performed with genetic algorithm, which is an intuitive algorithm that can find exact or approximate results in the optimization and search problem. The suitability function has been defined for use in genetic algorithm and applied on real data sets. Moreover, the results were supported by executing a simulation study. Clustering performance was calculated with different cluster validity indexes for both simulation study and real data, and the results were compared with the most used clustering algorithms. The obtained results showed that clustering with genetic algorithm gives more accurate results.

**Key words:** Clustering, genetic algorithm, validity indexes

## Veri Madenciliğinde Genetik Algoritma Kullanarak Kümeleme

### Özet

Veri madenciliğinde yaygın şekilde kullanılan, kümeleme analizi,  $m$  adet özelliğe sahip  $n$  adet birimin çeşitli algoritmalar ve elde bulunan bilgiler doğrultusunda  $k$  adet alt kümeye ayırma işlemidir. Kümeleme analizinde kullanılan bu çeşitli algoritmalar aynı veri seti için farklı kümeleme sonuçları verebilmektedir. Bu durumda birimlerin doğru kümelendiği en uygun algoritmayı bulmak gerekmektedir. Bu çalışmada birimlerin doğru kümelerde bulunmasını sağlamak amacıyla optimizasyon ve arama probleminde tam ya da yaklaşık sonuçlar bulabilen ve sezgisel bir algoritma olan genetik algoritma ile kümeleme analizi yapılmıştır. Genetik algoritmada kullanılmak üzere uygunluk fonksiyonu tanımlanmış ve gerçek veri setleri üzerinde uygulanmıştır. Dahası bir simülasyon çalışması yapılarak sonuçlar desteklenmiştir. Hem simülasyon çalışması hem de gerçek veriler için farklı küme geçerlilik indeksleri ile kümeleme performansı hesaplanmış ve bulunan sonuçlar en çok kullanılan kümeleme algoritmaları ile karşılaştırılmıştır. Elde edilen sonuçlar genetik algoritma ile kümelemenin daha doğru sonuçlar verdiğini göstermiştir.

**Anahtar kelimeler:** Kümeleme, genetik algoritma, geçerlilik indeksleri

## Clustering Analysis Using Ensemble Methods in Machine Learning

Erkut TEKELİ<sup>1</sup>

<sup>1</sup>Vocational School of Kozan, Çukurova University, 01530 Kozan, Adana, Turkey

### Abstract

In recent years, multiple classification systems also called Ensemble systems, have gained increasing interest in the computational intelligence and machine learning community. This interest has been deserved as the Ensemble systems have proved to be very effective and extremely versatile in a wide variety of problem areas and real-world applications. Initially, Ensemble Systems developed to reduce variance and thus improve the accuracy of an automated decision-making system have been successfully used to address a variety of machine learning problems such as feature selection, confidence prediction, missing feature, incremental learning, error correction, and imbalanced class data. This study provides an overview of ensembled systems, their features, and how they can be applied to such a wide range of applications. In this study, the computer application of bagged clustering methods, which is an ensembled method, has been done by using the e1071 package in the R programming language. For this purpose, models were created and analyzed using the Bagged clustering method that uses k-means as a community method for a real data set. In order to compare with the bagged clustering method, the data set was clustered with 3 different hierarchical methods in addition to the k-means method. Clusters were estimated with each method and cluster performance was measured with 3 different criteria. Although the bagged clustering method is the best clustering method according to the Rand index and classification rate criteria, the average linked hierarchical method has shown the best performance according to the silhouette criterion. Moreover, the study was supported by a Monte Carlo simulation.

**Key words:** Ensemble, clustering, machine learning

## Using Stochastic Differential Equations Modelling for Agricultural Products on Real Data

Nihal İNCE<sup>1</sup>, Sevil ŞENTÜRK<sup>1</sup>

<sup>1</sup> Eskişehir Technical University, Eskişehir, Turkey

### Abstract

In recent years, stochastic differential equations (SDD) have been included in the literature and become an area of interest with the introduction of randomness to the structure of the problem. The desire of obtaining more realistic models emerges the need for modeling with the stochastic differential equations instead of ordinary differential equations in the fields of financial mathematics, biology, physics, biology, chemistry, economics, and so on. Unlike the ordinary differential equations, most of the stochastic differential equations don't have an exact solution process; therefore, there are different numeric solution techniques. In this study, the Euler-Maruyama (EM) method is used to numerically solve the SDEs. Also, approximate probability density functions of random variables of solutions of SDE model are found by using Generalized Entropy Optimization Methods (GEOM). In application, the mentioned method is fulfilled using the pumpkin quantity between 1990 and 2019 in Turkey. After obtained results, the fact that farmers have reached the estimated number of crops to be produced in the coming years will help them make timely decisions and significantly increase the intensity of production.

**Key words:** Stochastic differential equations, probability density function, generalized entropy optimization methods, random variable, stochastic process

**Increasing the Power of Goodness of Fit Tests for Cauchy Distribution Versus Symmetrical Distributions for Small Samples by Means of Characterization**

Ahmet PEKGÖR<sup>1</sup>

<sup>1</sup>Necmettin Erbakan University, Faculty of Science, Department of Statistics, Konya, Turkey

**Abstract**

In this study, a new algorithm is proposed to develop well-known goodness-of-fit tests for Cauchy distribution against symmetrical distributions in small samples. The characterization theorem on the Cauchy distribution is used in this new algorithm. Critical values were obtained by simulation and power analysis was performed to compare the performance of the newly introduced algorithm. After this power analysis, it was shown that the results obtained with the new algorithm have a better power against symmetrical distributions with small sample sizes than classical goodness of fit tests in the literature.

**Key words:** Cauchy distribution, characterization, goodness of fit, Monte Carlo simulation, robustness

## Detecting the Speaker Using Artificial Neural Networks on the Audio Signal

Hasan Alp ŞAHİN<sup>1</sup>, Hasan ÖNDER<sup>1</sup>

<sup>1</sup>Ondokuz Mayıs University, Agricultural Faculty, Department of Animal Science, 55139, Atakum, Samsun, Turkey

### Abstract

This study aimed to determine the effectiveness of the identification of the speaker and his/her gender in artificial neural networks (ANN) through voice signals, which are rapidly increasing in security and cyber security. As the experimental material, the word "book", which was recited by 5 women, 5 men with 10 repetitions between the ages of 19-30, was used as a sound recording. Voice recordings were recorded with a Sony Icd-Px470 tape recorder from a fixed distance and were recorded as 128kb / sec. MATLAB software was used to make analysis. According to the results, the coefficient of determination for the identity of the speaker according to the ANN classification was 0.90, and the coefficient of determination for the identity the speaker's gender was 1.0. Results showed that Artificial Neural Networks can be easily and reliably used to determine the gender and identity of individual person by using audio signals.

**Key words:** Artificial Neural Networks, gender, identity, speaker, voice recognition

## **Flexicurity: A Comparison to OECD Countries**

Berna YAZAR ASLAN<sup>1</sup>

<sup>1</sup>ÇASGEM

### **Abstract**

Within the framework of the indicators used by the European Commission to monitor the success of flexicurity policies in labor markets, the changes on employment and unemployment rate due to change of flexicurity indicators of 35 OECD countries in the 2008-2017 period were estimated with panel data analysis models. According to the estimation results, it can be said that the unemployment rates are high in countries where the protective legislation for employment is strict. The relatively flexible implementation of labor force protection practises, increased spending on lifelong learning and active labor market policies stand out as factors that support the increase in employment rates. Countries with a high level of welfare protect individuals under the roof of a strong social security system.

**Key words:** Guaranteed flexibility, employment, labor market, unemployment, panel data analysis

## **Güvenceli Esneklik: OECD Ülkelerine Yönelik Bir Karşılaştırma**

### **Özet**

Avrupa Komisyonu tarafından işgücü piyasalarında güvenceli esneklik politikalarının başarılarını izlemede kullanılan göstergeler çerçevesinde 2008-2017 döneminde 35 OECD ülkesinin güvenceli esneklik göstergelerinde yaşanan değişimlerin işgücü piyasası göstergelerinden işsizlik ve istihdam oranları üzerindeki etkileri panel veri analizi modelleri ile tahmin edilmiştir. Tahmin sonuçlarına göre, istihdamın korunmasına yönelik koruyucu mevzuatın katıldığı ülkelerde işsizlik oranlarının yüksek seyrettiği söylenebilir. İşgücünü koruyucu mevzuat uygulamalarının nispeten daha esnek olması, hayat boyu öğrenmeye ve aktif işgücü piyasası politikalarına yapılan harcamamaların artması istihdam oranlarında artışı destekleyen unsurlar olarak öne çıkmıştır. Yüksek refah seviyesine ulaşan ülkeler güçlü sosyal güvenlik sistemi çatısı altında bireyleri koruma altına almaktadır.

**Anahtar kelimeler:** Güvenceli esneklik, istihdam, işgücü piyasası, işsizlik, panel veri analizi

## Tünel Açma Makineleri Üreten Bir Firmanın Kesici Disk Sert Dolgu Sürecinde İstatistiksel Süreç Kontrolü Çalışmaları

Rıfkiye Gamze GENÇ<sup>1</sup>, Pelin TOKTAŞ<sup>2</sup>

<sup>1</sup>E-Berk Makina Metalurji İnş. Taah. San. Tic. A.Ş., Malıköy Anadolu Osb Mah., Ankara, Türkiye

<sup>2</sup>Başkent Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Ankara, Türkiye

### Özet

Tünel açma makineleri, çeşitli toprak ve kaya tabakalarını dairesel kesit çizerek kazmaya yarayan makinelerdir. Ön kısmında yer alan döner bir kesici kafa ile kazı yaparken, arka bölümünde yer alan itki silindirleri ile de kazı esnasında ilerleme işlemini gerçekleştirmektedir. Bu makineler tünel yüzeyine taş döşeme işlemi de yaptığından, açılan tüneller, yüzeyi döşenmiş halde tamamlanmaktadır. Bu çalışmada, tünel açma makineleri üreten bir firmada, kesici kafa üzerinde bulunan kesici disklerinin ve kesicilerinin üretim sürecinde aşınmazlığını sağlayan dolguların sertliğinin istenilen ölçüde olmadığı tespit edilmiştir. Öncelikle, kesici kafa üzerinde bulunan kesici disk ve kesicilere uygulamış olduğu sert dolgu sürecinden elde edilen sertlik derecelerine ait alınan ölçümlerin yeterliliği kontrol edilmiştir. Ardından süreç için belirlenmiş spesifikasyon limitleri göz önünde bulundurularak süreç yeterlilik analizi yapılmış ve uygun kontrol grafikleri oluşturulmuştur. Yapılan tüm analiz ve çalışmalardan sonra sürecin kontrol altında olmasını engelleyen özel nedenler tespit edilmiş, süreç kontrolünde kullanılmak üzere Faz II kontrol limitleri oluşturulmuştur. Daha önce istatistiksel süreç kontrolü yapılmayan bu firma için yürütülen bu çalışma, bir örnek niteliğindedir.

**Anahtar kelimeler:** İstatistiksel süreç kontrolü, ölçüm yeterlilik analizi, süreç yeterlilik analizi, kontrol grafikleri, sert dolgu süreci



## **Estimation of the Ridge Parameter Using the Jackknife Approach**

Ismail SHAH<sup>1</sup>, Faiza SAJID<sup>1</sup>

<sup>1</sup>Department of Statistics, Quaid-i-Azam University, Islamabad, 45320, Pakistan

### **Abstract**

Regression techniques are generally used to predict a response variable using one or more predictor variables. In many fields of study, the regressors can be highly intercorrelated. Due to this correlation, the problem of multicollinearity occurs and consequently, the ordinary least squares estimates become inconsistent and lead to wrong inferences. To handle the problem, the ridge regression approach is commonly suggested. In this thesis, we revisit the problem of estimating the ridge parameter “**k**” by proposing some new estimators using the Jackknife method and compare them with some existing estimators. The performance of the proposed estimators to the existing ones is evaluated using extensive Monte Carlo simulations as well as two real-life examples. The results suggested that the proposed estimators outperform the existing estimators.

**Key words:**

## Statistical Modelling of Lamb Culling During the Rearing Period with the Use of Classification Trees

Dariusz PIWCZYŃSKI<sup>1</sup>, Julia PIWCZYŃSKA<sup>2</sup>, Magdalena KOLENDA<sup>1</sup>, Hasan ONDER<sup>3</sup>

<sup>1</sup>Faculty of Animal Breeding and Biology, UTP University of Science and Technology, Bydgoszcz, Poland

<sup>2</sup>Faculty of Veterinary Medicine, Wrocław University of Environmental and Life Sciences, Wrocław, Poland

<sup>3</sup>Faculty of Agriculture, Ondokuz Mayıs University, Samsun, Turkey

### Abstract

The aim of the presented research was to statistically analyse the survival of 44 133 Polish Merino and Polish Merino in Old-Type lambs between birth and the 100<sup>th</sup> day of their life, using classification trees and logistic regression. The study included lambs born between 2008 and 2017 and used in 43 flocks in Poland. The results showed that 9.27% of all controlled lambs did not survive till the 100<sup>th</sup> day of life. The statistical analysis of the case of lambs' death during their first 100 days of life was carried out using multiple logistic regression as well as classification trees, using two algorithms CART and CHAID (Enterprise Miner 15.1, SAS package). The calculated quality criteria (percentage of misclassifications, average squared error and the area under the ROC curve) for tree models that were created during the research suggested that the classification trees formed based on CART algorithm were the most accurate in defining the variability of studied characteristics, i.e. survival of lambs up to the 100<sup>th</sup> day of age. For the best available classification model, the ranking of variable importance, developed based on the "Importance" measure, allowed to conclude that the type of lamb's birth, season, following by year of birth, subsequent lambing, lamb's sex and its breed were the most significantly differentiating factors. It was noted that the tree build with the use of CART algorithm was composed of 30 leaves. It was also shown that the highest frequency of lamb's death during the rearing period was to be expected among triplets born in winter or summer (37.14% of all deaths), while the highest chance (98.42%) of surviving till the 100<sup>th</sup> day of life showed singletons, born from their mother's 3<sup>rd</sup> to 6<sup>th</sup> litter, in the spring-winter season in the last year of the present research.

**Key words:** CART, CHAID, lamb culling

## Normality Tests of Morphological Measurements of Biological Data of Sheep

Ansar ABBAS<sup>1</sup>, Aman ULLAH<sup>2</sup>, Abdul WAHEED<sup>3</sup>

<sup>1</sup>Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan

<sup>2</sup>Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan

<sup>3</sup>Faculty of Veterinary Science, Bahauddin Zakariya University, Multan, Pakistan

### Abstract

Generally, it is assumed that the biological data follow normal distribution otherwise the inference breaks down. Prior to statistical analysis, we therefore test the normality of the biological data. The main focus of this study was to apply different statistical techniques of normality on the biological data of sheep in order to see which biological body measurement followed normal distribution. A total of 164 sheep (85 female and 79 male) aged 28-365 days were included randomly in this study. Two renowned statistical tests of normality i.e. Kolmogorov-Smirnov test and Shapiro-Wilk test were applied on different body measurements of both male and female sheep, viz. withers height(WH), body length (BL), head length (HL), head width (HW), ear length (EL), ear width (EW), neck length (NL), neck width (NW), heart girth (HG), rump length (RL), rump width (RW), tail length (TL), barrel depth (BD), sacral pelvic width (SPW), birth weight (BW) and present weight (PW). A graphical technique quantile-quantile (Q-Q) plot was also used to see the data pattern of said variables. The Mean ( $\pm$  SD) of WH, BL, HL, HW, EL, EW, NL, NW, HG, RL, RW, TL, BD, SPW, BW and PW in all sheep were: 59.25 ( $\pm$  8.37), 57.21 ( $\pm$  9.61), 20.81 ( $\pm$  3.29), 8.74 ( $\pm$  2.09), 25.61 ( $\pm$  2.79), 11.25 ( $\pm$  1.41), 22.17 ( $\pm$  3.86), 14.37 ( $\pm$  2.39), 62.14 ( $\pm$  9.82), 11.49 ( $\pm$  2.22), 14.65 ( $\pm$  3.52), 10.37 ( $\pm$  2.54), 37.56 ( $\pm$  6.91), 66.68 ( $\pm$  12.16), 3.05 ( $\pm$  0.687) and 17.35 ( $\pm$  4.95) respectively. The mean WH, BL, HL, HW, EL, EW, NL, NW, HG, RL, RW, BD, SPW, BW and PW of male sheep were larger than female sheep. Among various body measurements of both male and female sheep, we observed that WH, BL, HG, BD and SPW body measurements were normally distributed by applying Kolmogorov-Smirnov test. While, body measurements WH, BL, NL, HG and BD were normally distributed by applying Shapiro-Wilk test. However, WH and BL of both male and female sheep were normally distributed in both of the tests. The results of this study suggest that the normality assumption should be checked for drawing robust inference of the data set.

**Key words:** Thalli sheep, normality testing, Kolmogorov-Smirnov, Shapiro-Wilk tests, Q-Q plot.

**Nonlinear Dimensional Reduction to with the Use of Artificial Neural Networks in  
Patients with Breast Tumor**

Canan DEMİR<sup>1</sup>

<sup>1</sup>Van Yuzuncu Yil University, Van, Turkey

**Abstract**

Breast cancer, which is the most common among women in the world and constitutes approximately 30% of all cancers, takes places near the top among the diseases that threaten women's health. Nonlinear Principal component analysis is a technique used to linearly map multidimensional data to lower dimensions with minimal information loss. The aim of this study is to investigate whether NLPCA is effective in reducing the number of variables in patients with breast tumor. 30-variable data sets of 569 patients with breast tumor were used in the implementation process. Nonlinear principal component analysis and artificial neural networks (ANN) were combined in the study. In the first stage, NLPCA was used in artificial neural networks for dimension reduction. In the second step, prediction models were created in artificial neural networks with both dimension reduced and non-dimensionally reduced data set. The results were interpreted by comparing the performance of these two prediction models with each other. The two basic components obtained with NLPCA explained 96.37% of the total variance.  $R^2$  values of the prediction models created for the dimension reduced data set and non-dimensionally reduced data set are 87% and 91%, respectively. This research reveals the benefits of NLPCA and ANN combination in patients with breast tumor. It is thought that using principal components as inputs can reduce unnecessary information and complexity.

**Key words:** Artificial neural networks, breast tumor, dimensional reduction, nonlinear principal component analysis, variance

## Application of Gompertz Model on Broilers Growth

Hafiz Muhammad ISHAQ<sup>1</sup>, Abdul WAHEED<sup>1</sup>, Asim FARAZ<sup>1</sup>, Abdul GHAYAS<sup>2</sup>

<sup>1</sup>Department of Livestock and Poultry Production, Bahauddin Zakariya University Multan, Pakistan

<sup>2</sup>Department of Poultry Production, University of Veterinary and Animal Sciences Lahore, Pakistan

### Abstract

Biostatistical tools are essential in analyzing biological data and are commonly employed in studying growth in different animal and bird species. Current study encompassed application of Gompertz model to explain growth of broilers maintained under standard conditions to underline the exact cut point or decision about harvesting them for meat. Gompertz model parameters estimates were obtained for broiler growth. A total of five hundred day old chicks of Cobb-500 were procured from local commercial hatchery and were reared at Avian Ward Faculty of Veterinary Sciences, Bahauddin Zakariya University, Multan chicks were reared for 35 days period under standard management practices by providing them a temperature of 32.5°C for first three days then dropping the temperature until it reaches 25°C and maintained it for rest of the life. The feed and water was offered on *ad-libitum* basis, broiler feed contained 22% crude protein during first 10 days and 20% crude protein from day 11 to day 30. The chicks were divided into five groups (each containing 100 chicks) based on their day old weight. The data about various body measurements and live weight were recorded on weekly basis where 10 birds randomly chosen from each group were measured. Colour markings were used for group identification. In order to obtain body weight for growth curve adjustment, the birds of each replicate were weighted weekly upto four weeks using weighing balance with accuracy of  $\pm 0.01$  gram. The body confirmation traits (in cm) were also measured on weekly basis including body length, keel length, body girth, wing length and shank length. Body length was measured from tip of the beak to the tip of toe while holding the calm bird on the floor. For measuring the keel length, keel bone was made prominent and measurement was noted from cranial pointed end of bone to caudal pointed end. For body girth, measuring tape was rapped around the body of bird at chest area. Wing length was recorded from tip of one wing to the tip of other side wing while spreading the both wings. Shank length as measured as distance from hock joint to foot pad. Three parametric Gompertz model was fitted on mean weights and body dimensions. Parameters (a, b and c) were estimated by using Statistix Version 8.0 software. Gompertz model showed the best fit for all groups and it was concluded that further expanded study would be planned to check the reproducibility of the experimental results and application the Gompertz model.

**Key words:** Gompertz model, broiler, biological growth

## Sample Size in Clinical Researches: Power of the Test and Effect Size

Adnan ÜNALAN<sup>1</sup>

<sup>1</sup>Niğde Ömer Halisdemir University, Faculty of Medicine, Department of Basic Medical Sciences,  
Division of Biostatistics and Medical Informatics, Niğde, Turkey

### Abstract

The approval of local ethics committees is required for clinical researches. In order to obtain approval, how the sample size is determined, whether power analyze is done or not and under what assumptions these analyzes are made, are important questions/problems. In hypothesis tests, it is possible two types of errors (type 1 error denoted by  $\alpha$  and type 2 error denoted by  $\beta$ ), of which  $\alpha$  is the probability of rejecting the null hypothesis that is actually true, and is the probability of accepting the actually false null hypothesis. These errors also determine the reliability of the test ( $1-\alpha$ ) and the power of test ( $1-\beta$ ). While  $\alpha$  is directly determined by the researchers and generally as taken 5% (in some cases 1%),  $\beta$  cannot be determined directly. Because  $\beta$ , hence the power of test ( $1-\beta$ ) depends on the  $\alpha$  (negatively correlated with  $\beta$ ) the variation in the population (positively correlated with  $\beta$ ) and sample size ( $n$ ; negatively correlated with  $\beta$ ). In clinical researches, it is required that  $\beta$  does not exceed 10% (in some cases 5%) so the power of test should be at least 90% and above. In this study, the sample sizes required for some statistical tests (independent sample t-test, one-way ANOVA and Chi-square) which are widely used in clinical research, were calculated with the G\*Power program and some evaluations were made. As a result, as expected in the statistical tests, it was observed that decreasing both  $\alpha$  and effect size and increasing the power of the test significantly increased the sample size. However, it was also observed that increasing effect on the sample size of increasing the power of test decreased (5-11%) in the smaller values of  $\alpha$  in the independent sample t-test, decreased (nearly 5%) when increasing the number of compared groups in one-way ANOVA and decreased (10-15%) when increasing degree of freedom of Chi-square test.

**Key words:** Clinical researches, test of hypothesis, sample size, power analysis, effect size

## Klinik Araştırmalarda Örneklem Büyüklüğü: Testin Gücü ve Etki Büyüklüğü

### Özet

Klinik araştırmaların yapılabilmesi için yerel etik kurulların onayı gerekmektedir. Onay alınabilmesi için ise araştırmada öngörülen örneklem büyüklüğünün nasıl belirlendiği, kullanılacak istatistik teste ilişkin güç analizinin yapılıp-yapılmadığı ve bu analizlerin hangi varsayımlar altında yapıldığı konuları önemli birer soru/sorun olarak karşımıza çıkmaktadır. Hipotez testlerinde iki tip hata (1. tip hata  $\alpha$  ile 2. tip hata  $\beta$  ile gösterilir) ile karşılaşmak mümkündür ve bunlardan  $\alpha$  gerçekte doğru olan sıfır/yokluk hipotezinin yapılan istatistik test sonucunda reddedilmesi olasılığı,  $\beta$  ise gerçekte yanlış olan sıfır hipotezinin kabul edilmesi olasılığıdır. Bu hatalar aynı zamanda analizlerde kullanılacak istatistik testin güvenilirlik düzeyini ( $1-\alpha$ ) ve testin gücünü ( $1-\beta$ ) de belirlemektedir.  $\alpha$ 'nın değeri araştırmacılar tarafından genellikle %5 (bazı durumlarda %1 olarak da alınır) olarak doğrudan belirlenirken,  $\beta$ 'nin değeri araştırmacı tarafından doğrudan belirlenemez. Çünkü  $\beta$ 'yi dolayısıyla da testin gücünü ( $1-\beta$ )  $\alpha$ 'nın değeri ( $\alpha$  ile  $\beta$  arasındaki ters yönde ilişki nedeniyle  $\alpha$  arttıkça  $\beta$  azalacak dolayısıyla testin gücü artacaktır) anakütledeki değişkenlik (değişkenlik ile  $\beta$  arasındaki doğru ilişki nedeniyle değişkenlik

azaldıkça  $\beta$  da azalacak dolayısıyla testin gücü artacaktır) ve örneklem büyüklüğü ( $n$ ;  $n$  ile  $\beta$  arasındaki ters yönde ilişki nedeniyle  $n$  artıkça  $\beta$  azalacak dolayısıyla testin gücü artacaktır) gibi faktörler belirlemektedir. Günümüzde yürütülen klinik araştırmalarda  $\beta$ 'nın da %10'u (bazı durumlarda %5'i) geçmemesi dolayısıyla da testin gücünün de en az %90 ve üzerinde olması istenmektedir. Bu çalışmada, G\*Power programıyla klinik araştırmalarda da yaygın olarak kullanılan bazı istatistik testler (bağımsız örneklem t-testi, tek yönlü ANOVA ve Ki-kare testi) için gerekli olan örneklem büyüklükleri hesaplanmış ve bazı değerlendirmeler yapılmıştır. Sonuç olarak beklenen şekilde incelenen testlerde  $\alpha$  ve etki büyüklüğünün düşürülmesi ile testin gücünün artırılmasının örneklem büyüklüğünü önemli oranda artırdığı görülmüştür. Bununla birlikte testin gücünün artırılmasının örneklem büyüklüğü üzerindeki artırıcı etkisinin; bağımsız örneklem t-testinde  $\alpha$ 'nın daha küçük değerlerinde düştüğü (%5-11), tek yönlü ANOVA'da karşılaştırılacak grup sayısının artması durumunda düştüğü (yaklaşık %5), Ki-kare testinde ise serbestlik derecesinin artması durumunda yine düştüğü (%10-15) görülmüştür.

**Anahtar kelimeler:** Klinik araştırma, hipotez testi, örneklem büyüklüğü, güç analizi, etki büyüklüğü

## Modification of the Schnute Growth Model

Olgun DURAN<sup>1</sup>, Deniz ÜNAL<sup>2</sup>

<sup>1</sup>Çukurova Üniversitesi, Tıp Fakültesi Balcalı Hastanesi Radyoloji A.B.D, Adana, Türkiye

<sup>2</sup>Çukurova Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, Adana, Türkiye

### Abstract

Statistical modeling with growth data is the most efficient way to get more objective results and understand biological development. Schnute growth model is a large scale model that includes many nonlinear growth models. It offers more optimum parameter estimates compared to other growth models, especially regardless of whether the growth curve has an asymptotic feature. In recent years, in the literature studies, there are approaches in the theoretical and application fields in which the equations are modified to make sense of growth equations biologically. The purpose of modifying the equations is to convert growth parameters  $a, g, h$ , into biologically meaningful parameters such as maximum value ( $A$ ), spike maximum specific growth rate ( $\mu_m$ ) and lag time ( $\lambda$ ). In this study, it is demonstrated by which mathematical operations the parameters of the Schnute growth model are converted to biologically significant parameters and a new modified Schnute growth model is presented to the literature.

**Key words:** Growth models, modified equation, Schnute

## Schnute Büyüme Modelinin Modifikasyonu

### Özet

Büyüme verileriyle istatistiksel modelleme yapmak, daha nesnel sonuçlar elde etmede ve biyolojik gelişmeyi anlamada en verimli yolu oluşturur. Schnute büyüme modeli, doğrusal olmayan pek çok büyüme modelini de kapsayan geniş çaplı bir modeldir. Özellikle büyüme eğrisinin asimptotik özelliğinin olup olmamasına bakmadan diğer büyüme modellerine göre parametre tahminlerini daha optimal sunar. Son yıllarda literatür çalışmalarında, büyüme denklemlerine ait parametreleri, biyolojik olarak anlamlandırmak için denklemlerin modifiye edildiği teorik ve uygulama alanlarında yaklaşımlar bulunmaktadır. Denklemleri modifiye etmedeki amaç, büyüme parametreleri olarak temsil edilen  $a, g, h$  gibi parametrelerin biyolojik olarak anlamı bulunan maksimum değer ( $A$ ), ani spesifik büyüme hızı ( $\mu_m$ ), gecikme süresi ( $\lambda$ ) gibi parametrelere dönüştürülmesidir. Bu çalışmada, Schnute büyüme modelinde yer alan parametrelerin biyolojik anlamlı parametrelere dönüşümlerinin hangi matematiksel işlemlerle yapıldığı gösterilmiş ve yeni bir modifiye edilmiş Schnute büyüme modeli literatüre sunulmuştur.

**Anahtar kelimeler:** Büyüme modelleri, modifiye denklemler, Schnute



**A Comparative Study of Excel Spreadsheet and JMP Statistical Software's for  
Regression Analysis in Agricultural Studies: A Case Study of Chickpea Yield  
Management by Land Rolling**

Songül GÜRSOY<sup>1</sup>

<sup>1</sup>Department of Agricultural Machinery and Technology Engineering, Faculty of Agriculture, Dicle University, 21280, Sur, Diyarbakır, Turkey

**Abstract**

Excel spreadsheet and JMP statistical software can be used for regression analysis in agricultural studies. However, there can be some difference between the regression analysis results of excel and JMP since Excel treats numeric data differently than either JMP or SAS and it also has its own regression algorithms. Therefore, this study compares the regression analysis results of Excel and JMP statistical software's showing the relationship between grain yield of chickpea and the penetration resistance at 5 cm soil depths in a study which the different ground pressures of land roller (0, 20, 25, 30, 35, 40 kPa) were tested at different times (pre-emergence and post-emergence) under field conditions. The results of the study showed that in Excel, the most popular spreadsheet package, the linear regression analysis could be performed, and a regression equation and the coefficient of determination could be also created by using scatterplot with trendline. Therefore, it can be used as powerful yet flexible tool that can be applied to a wide range of common problems in agricultural experiment.

**Key words:** Regression analysis, Excel, JMP, agriculture.

**Introduction**

Regression analysis is a set of statistical methods used for the estimation of relationships between a dependent variable and one or more independent variables. It can be utilized to assess the strength of the relationship between variables and for modeling the future relationship between them. Regression analysis includes several variations, such as linear, multiple linear, and nonlinear. The most common models are simple linear and multiple linear. Nonlinear regression analysis is commonly used for more complicated data sets in which the dependent and independent variables show a nonlinear relationship (Soto, 2013). Regression analysis uses a chosen estimation method, a dependent variable, and one or more explanatory variables to create an equation that estimates values for the dependent variable. The regression model includes outputs, such as R<sup>2</sup> and p-values, to provide information on how well the model estimates the dependent variable (Agrawal, 2018). Charts, such as scatter plot matrices, histograms, and point charts, can also be used in regression analysis to analyze relationships and test assumptions (Chatterjee and Sarkar, 2009).

All major statistical software packages perform least squares regression analysis and inference. Simple linear regression and multiple regression using least squares can be done in some spreadsheet applications and on some calculators. While many statistical software packages can perform various types of nonparametric and robust regression, these methods are less standardized; different software packages implement different methods, and a method with a given name may be implemented differently in different packages. Excel spreadsheet is mostly used to perform the simple linear regression analysis. Also, we can create a regression equation using scatterplot with trendline in Excel (Brown, 2001; Mason, 2012; Carlberg, 2016).

The aim of this study is to compare the regression analysis results of Excel and JMP statistical software's showing the relationship between grain yield of chickpea and the penetration resistance at 5 cm soil depths in a study which the different ground pressures of land roller (0, 20, 25, 30, 35, 40 kPa) were tested at different times (pre-emergence and post-emergence) under field conditions.

### Material and Methods

A field experiment was conducted in February-June 2018 at Dicle University Field Crop Production area in Diyarbakır, Turkey.

The empty weight of the land roller used in the experiment was 550 kg with roller diameter of 50 cm and roller length of 250 cm. The empty weight of the land roller included the frame design, the transport wheels, the hydraulics, etc. The weight of the land roller was increased by filling the 25, 50, 75, and 100% of the roller volume with water to 673, 795, 918 and 1040 kg.

The ground pressures of the roller was calculated by dividing the roller weights to the soil contact area of the roller as 0, 20, 25, 30, 35, 40 kPa. The roller was pulled by the MF 7240 tractor at 6.4 km h<sup>-1</sup>. The speed of 6.4 km h<sup>-1</sup> was chosen to be the most suitable speed in matching the tractor to the land roller.

Soil penetration resistance were measured using a digital cone penetrometer FieldScout SC 900 (Spectrum Technologies, Aurora, IL) recording the pressure applied in Pascals every 2.5 cm, to a depth of 45 cm. Penetration resistance was also recorded at three locations of each plot 45 days after planting.

Grain yield was calculated as kg ha<sup>-1</sup> from grain weight, which was cut by hand and machine threshed in a 10.0 m<sup>2</sup> harvest area for each plot.

To determine the relationship between chickpea grain yield and penetration resistance at 5 cm soil depth, the regression analysis was performed by using excel and JMP statistical software.

### Results and Discussion

Figure 1 shows the summary output table of linear regression after performing the regression analysis in Excel. In Figure 1, Multiple R. is the Correlation Coefficient that measures the strength of a linear relationship between two variables. The larger the absolute value, the stronger the relationship. R Square is the Coefficient of Determination, which is used as an indicator of the goodness of fit. It shows how many points fall on the regression line. In our example, R<sup>2</sup> is 0.099 (rounded to 2 digits), which is very weak for linear relationship. In this case, we must perform the non-linear regression analysis. Adjusted R Square is the R square adjusted for the number of independent variable in the model. You will want to use this value instead of R square for multiple regression analysis. Standard Error is another goodness-of-fit measure that shows the precision of your regression analysis - the smaller the number, the more certain you can be about your regression equation. While R<sup>2</sup> represents the percentage of the dependent variables variance that is explained by the model, Standard Error is an absolute measure that shows the average distance that the data points fall from the regression line. Also, the ANOVA part is seen in Figure 1 for a simple linear regression analysis in Excel. The df is the number of the degrees of freedom associated with the sources of variance.

SS is the sum of squares. The smaller the Residual SS compared with the Total SS, the better your model fits the data. MS is the mean square. F is the F statistic, or F-test for the null hypothesis. It is used to test the overall significance of the model. The significance F is the P-value of F. The Significance F value gives an idea of how reliable (statistically significant) your results are. If Significance F is less than 0.05 (5%), your model is OK. If it is greater than 0.05, you'd probably better choose another independent variable or non-linear analysis.

Basically, it splits the sum of squares into individual components that give information about the levels of variability within your regression model: The standard formula for linear regression with one independent variable is:

$$Y(X) = a + bX$$

where:

Y(X) = Predicted value of the dependent variable

a = Y-intercept

b = Coefficient of the independent variable

X = Any value of the independent variable. Thus, using Excel terminology, the regression formula in this study could be written as: Yield = 29.94-0.00494 x penetration resistance. In this study, we could not perform non-linear regression analysis in Excel. However, Kucukelci and Evren, (2017) stated that by Microsoft Excel solver package, one can solve some optimization problems including nonlinear regression problems precisely. Also, Mason (2012) reported that Excel performs the regression analysis and writes the regression results to a summary output table. Included in the summary output table are R<sup>2</sup>, the standard error of the Y estimate, the Y-intercept, the standard error and t-statistic of the Y intercept, the coefficient of each independent variable, and the standard error and t-statistic of each coefficient.

ÖZET ÇIKIŞI									
<i>Regresyon İstatistikleri</i>									
Çoklu R	0,314818								
R Kare	0,099111								
Ayarlı R Kare	0,072614								
Standart Hata	4,999953								
Gözlem	36								
<i>ANOVA</i>									
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Anlamlılık F</i>				
Regresyon	1	93,51027	93,51027	3,740481	0,061468				
Fark	34	849,984	24,99953						
Toplam	35	943,4943							
<i>Katsayılar standart Hata</i>									
	<i>t Stat</i>	<i>P-değeri</i>	<i>Düşük %95</i>	<i>Üyüksek %95</i>	<i>Düşük 95,0%</i>	<i>Üyüksek 95,0%</i>			
Kesişim	29,94673	1,412118	21,20696	3,51E-21	27,07696	32,8165	27,07696	32,8165	
penetratic	-0,00494	0,002552	-1,93403	0,061468	-0,01012	0,000251	-0,01012	0,000251	
<i>FARK ÇIKIŞI</i>									
	<i>Gözlem</i>	<i>Ölülen yield</i>	<i>Farklar</i>						
	1	29,60612	-6,44586						
	2	29,60612	-3,57086						
	3	29,43334	-7,02629						

Figure 1. Spreadsheet for Linear Regression in Excel.

The scatter plots for this study are shown in Figure 2. The regression line demonstrates the relationship between the independent variable (penetration resistance) and dependent variable (chickpea yield). In the scatter plot, Excel create regression equation and R2 value as quadratic regression (using polynomial, order 2), cubic regression (using polynomial, order 3), and other polynomial regression functions (orders 4, 5, and 6), logarithmic regression, exponential regression, and power regression. The highest value of R2 was selected as the quadratic regression equation of best fit. This results shows that the relationship between penetration resistance and chickpea yield was non-linear because it fit the best as polynomial.

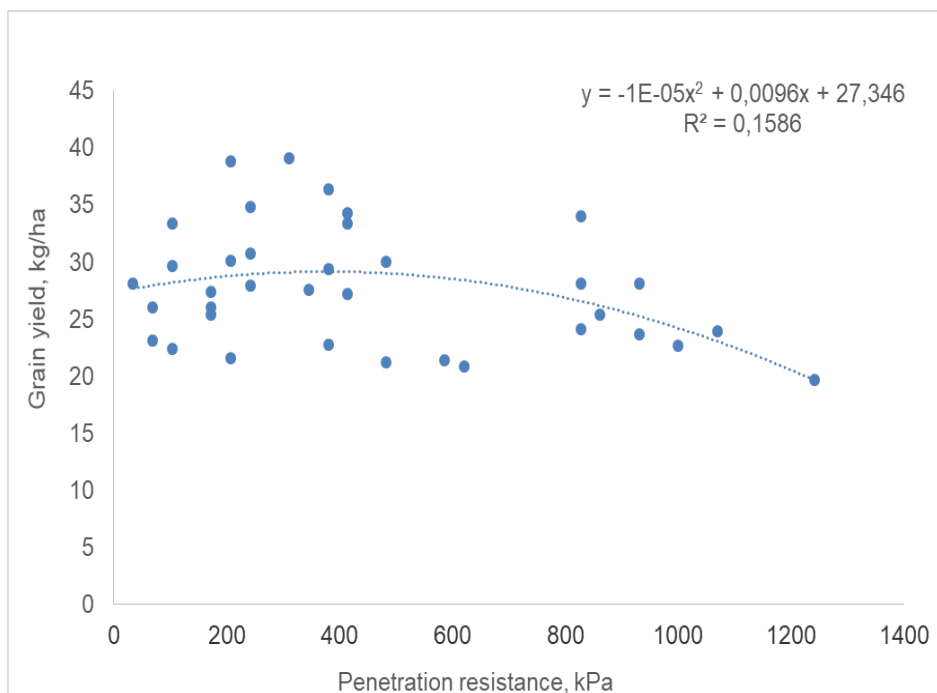


Figure 2. Non-Linear Regression. An X-Y Scatter Plot Illustrating the Difference Between The Data Points and the Polynomial Fit.

The results of linear and non-linear regression analysis performed in the JMP package software to determine the relationship between the penetration resistance and the chickpea yield are shown in Figures 3 and 4. The  $R^2$  and P values were similar in Excel and JMP software for linear analysis. Also, the model equation was  $Yield = 29.94 - 0.00494 \times \text{penetration resistance}$  in JMP software as well as in Excel. Since a simple linear regression model inadequately characterized the relationship between penetration resistance and yield response variables, a Polynomial Regression analysis was performed. The  $R^2$  and P value in the Polynomial Regression model was higher and lower than that in the simple linear regression model, respectively. Therefore, we can select the Polynomial Regression model to predict the relationship between penetration resistance and yield. However, the low relationship between the penetration resistance and yield could result from the variation of the field data as shown by the high standard deviations and error sum of square.

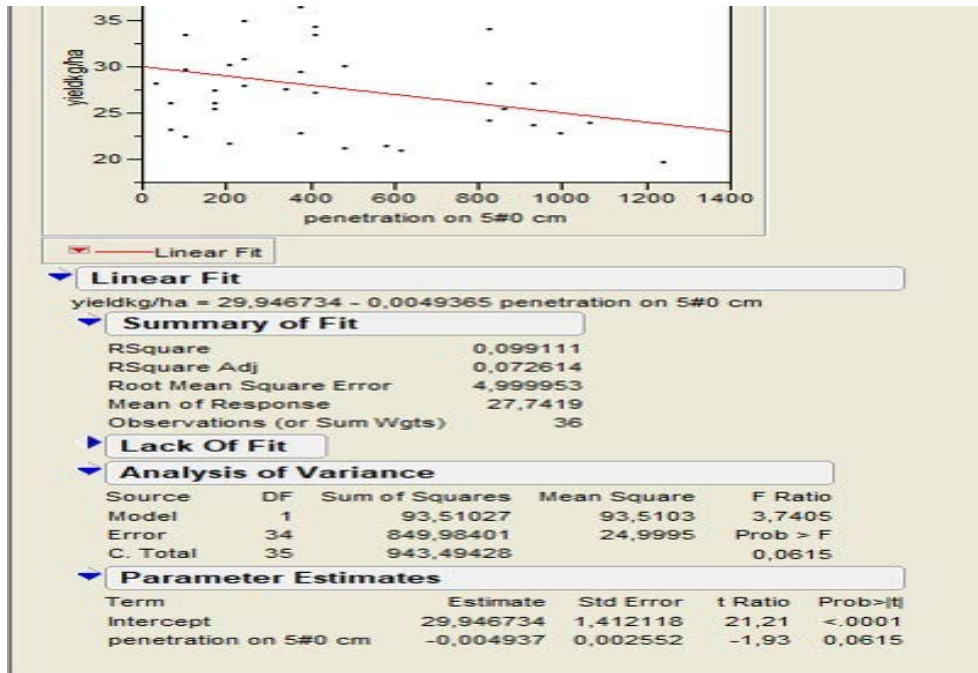


Figure 3. Spreadsheet for Linear Regression in JMP

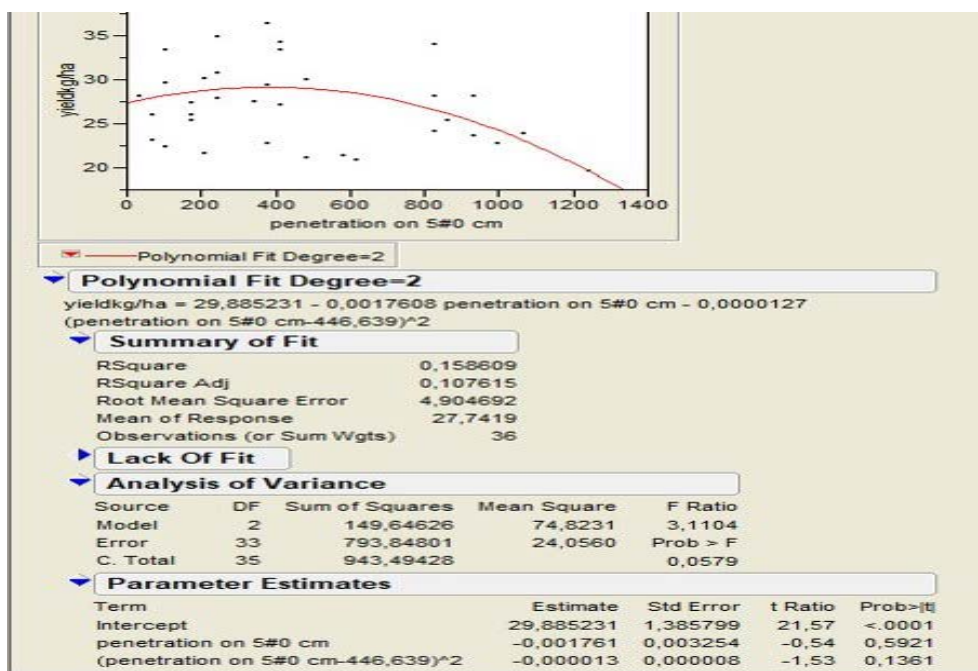


Figure 4. Spreadsheet for Non-Linear Regression in JMP

## Conclusion

The results of this study showed that Excel, the most popular spreadsheet package, could be used as powerful yet flexible tool that can be applied to a wide range of common problems in agricultural experiment. However, in Excel, only linear regression analysis could be performed although a non-linear regression equation and  $R^2$  was created using scatterplot with trendline in Excel.

## References

- Agrawal S. 2018. P value and R squared. <https://www.rpubs.com/shubh2565/pvalue-rsquared>.
- Brown AM. 2001. A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. *Computer Methods and Programs in Biomedicine*. 65 (3): 191-200.
- Carlberg C. 2016. *Regression Analysis Microsoft® Excel*. Pearson Education, Inc., USA
- Chatterjee C, Sarkar RR. 2009. Multi-step polynomial regression method to model and forecast malaria incidence. *PLoS ONE* 4(3): e4726.
- Kucukelci, T. D., Evren, A. 2017. A Comparison Between Microsoft Excel Solver and NCSS, SPSS Routines for Nonlinear Regression Models. *Journal of Engineering Technology and Applied Sciences* 2 (2) : 91-100.
- Mason JO. 2012. Performing regression analysis using microsoft excel. *International Journal of Arts and Commerce*. 1 (5): 192-208.
- Soto T. (2013) Regression Analysis. In: Volkmar F.R. (eds) *Encyclopedia of Autism Spectrum Disorders*. Springer, New York, NY.

## Finite Element Models Applied to Predict Soil Compaction in Agricultural Areas

Songül GÜRSOY<sup>1</sup>

<sup>1</sup>Department of Agricultural Machinery and Technology Engineering, Faculty of Agriculture, Dicle University, 21280, Sur, Diyarbakır, Turkey

### Abstract

Soil compaction can be commonly defined as the increase of bulk density or the decrease in porosity of soil due to externally or internally applied loads. Soil compaction in agriculture areas is a complex problem in which soil, crops, weather and machinery interact. External pressure due to the use of heavy machinery and inappropriate soil management can lead to the compaction. Soil compaction has an important effect on water, infiltration into soil, crop emergence, root penetration, crop nutrient and water uptake, and over all crop yield.

During the past few decades, soil compaction has received increasing attention in the field of agricultural production due to the increasing size of farms, farm equipment and the time needed to complete farm operations at seeding and harvest. The majority of research efforts are focused on soil-machine/tire and field testing. Some analytical, numerical and empirical approaches were investigated to predict soil compaction in agriculture areas. Specially, the use of numerical simulations has been widespread in many engineering fields and related areas. One of the main numerical methods used in modeling of soil compaction is the finite element method (FEM).

This paper aims to present a review of the FEM applications in predicting soil compaction in agricultural areas.

**Key words:** Soil compaction, soil-tire interaction, modeling, FEM

### Introduction

Soil compaction can be commonly defined as the increase of bulk density or the decrease in porosity of soil due to externally or internally applied loads. In other words, soil compaction decreases the number and size of large pores in soil, and also increases the mechanical resistance of the soil because of pressing soil particles more closely together (Nawaz et al., 2013). Soil compaction has an important effect on water, infiltration into soil, crop emergence, root penetration, crop nutrient and water uptake, and over all crop yield (Unger and, Kaspar, 1994; Parlak and Özasan Parlak, 2011). Soil compaction in agriculture areas is a complex problem in which soil, crops, weather and machinery interact. External pressure due to the use of heavy machinery and inappropriate soil management can lead to the compaction. Nawaz et al. (2013) stated that a complete understanding of effects of soil compaction on soil properties and plant growth is necessary to meet the future global challenge of food security. The authors, who reviewed the advances in understanding, quantification, and prediction of the effects of soil compaction, found that soil compaction has direct effects on soil physical properties such as bulk density, strength, and porosity; therefore, these parameters can be used to quantify the soil compactness.

Many field and laboratory studies have been conducted to better understand the soil compaction process (Botta et al., 2009; Etana et al., 2013; Soracco et al., 2015). However, the field and laboratory studies on soil compaction are generally labor intensive, time consuming, and expensive. Recently, computing developments have made it possible to predict soil compaction. Rather, it can be used effectively to better understand the mechanics of soil compaction and the factors affecting it (Defossez and Richard, 2002). Several models, until nowadays, are developed to assess the soil compaction due to

traffic load. Among these models, the Finite Element Method (FEM) is the most advanced mathematical tool, which can be mostly used to study tyre–soil interaction and soil compaction (Cueto et al., 2016). Therefore, the objective of this study is to review the Finite Element Models developed to predict the soil compaction and stress-strain behaviors in agricultural areas.

### Finite Element Analysis of Soil-Tire Interaction

The finite element method (FEM) is a numerical procedure for solving physical problems governed by differential equations. The FEM results in determining approximate values of desired parameters at specific points such as force, velocity, displacement and stress, called nodes, from linear or non-linear equation systems. A continuous approach function is assumed to represent the solution at each node (Velloso et al., 2018). Finite element models apply continuum mechanics and may have the potential to describe the mechanical behavior of soil.

The finite element tool is widely used in studying and prediction of soil compaction and tire-soil interaction for the tire running on agriculture area. Raper et al. (1994) who incorporated a soil compaction model that uses both normal and shearing stresses into a finite element model and compared the predicted stress levels against laboratory test, found that all predicted values at final deformed depths less than 03 m were within the 95% confidence intervals of the measured values. Also, Cueto et al. (2013) developed a model for soil compaction simulation caused by agricultural tire traffic by a linear elastic constitutive model. Predictions of deflection and tire contact area in rigid surface were compared with experimental results and the tire model was validated. In their study, a three dimensional model of the soil was created and the soil properties were represented with an Extended Drucker Prager material model by using ABAQUS/STANDARD 6.8-1 code. They found that simulated and observed stresses after wheel traffic under different inflation pressures and tire loads agree well. Liu and Wong (1996) reported that the finite element method (FEM) had a better potential to obtain results that are more precise because a smaller number of suppositions and simplifications are introduced. Xia K. (2011) developed a Drucker–Prager/Cap model implemented in ABAQUS for predicting the soil compaction as showed in Figure 1. The researcher stated that developed numerical model can serve as a robust tool on predicting soil compaction and tire mobility.

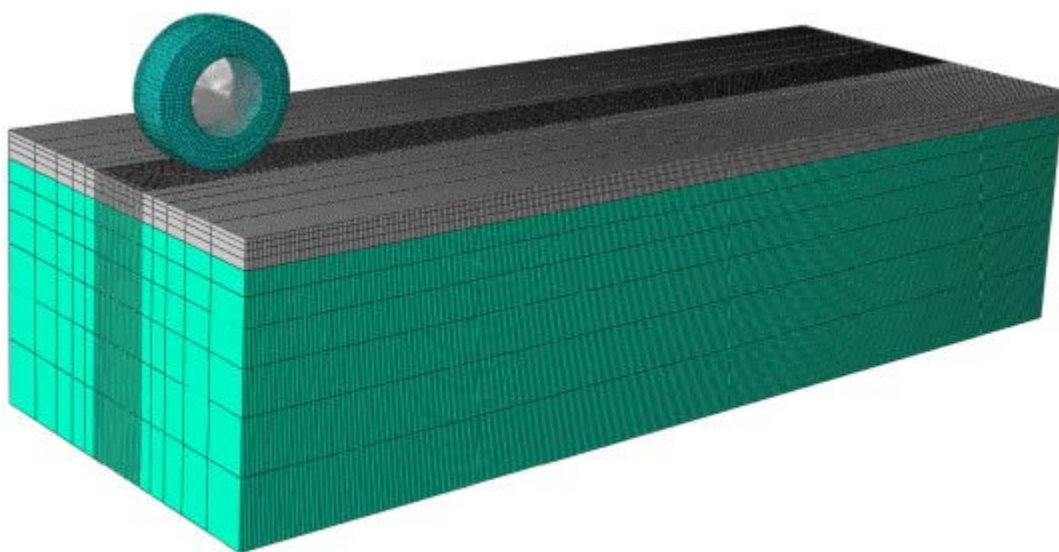


Figure 1. Finite Element Model Of Tire/Terrain Interaction (Xia, 2011).



Biriş et al. (2009) stated that for mathematical modelling the soil should be considered as a homogeneous and isotropic material, and the Drucker-Prager plasticity model can be used to simulate the behaviour of agricultural soil. They developed a model for prediction of the stress state in agricultural soil below agricultural tyres in the driving direction and perpendicular to the driving direction, which are different from one another, using the Finite Element Method. Nasiri et al. (2013) used the numerical and empirical methods for determining of soil compaction after tractor passing.. In empirical method, a Massy Ferguson tractor with a 3 moldboard mounted plough was used. In the study, the resistance of soil was determined using penetrometer and determination of cone index (CI). In numerical method, the finite element was utilized for simulation of soil and penetrometer interaction by ABAQUS software. They supposed the soil as Elastic-perfectly plastic material and linear elasticity and applied the Drucker-Prager plasticity model for indicating the elastic and plastic behavior of soil, respectively. Ungureanu et al., (2018) assessed the soil compaction depth by modelling the distribution of stresses and displacements in sandy soil under wheel loads applied by an agricultural trailer. For this purpose, a Finite Element Model was performed in the Quickfield Student program using the 2D geometric model of plane load.

Silva (2018) evaluated the mechanical behavior of the soil after a sugarcane crop preparation subjected to vehicle traffic by using the Finite Element Method (FEM) of the compaction in the soil layer 0–100 cm subdivided into 0–20 and 20–100 cm. The Modified Cam Clay elastoplastic constitutive model was used in the numerical analysis. The properties of the soil samples collected in each layer assessed through oedometer tests were used, considering the moistures of 10%, 15% and 20%. They stated that through the pre-consolidation stress, elastoplastic variable of the Modified Cam Clay model, it was possible to identify the region where soil compaction occurs. Poodt et al. (2003) used FEM to evaluate a soil profile under traffic of beet harvesters considering four types of tires, using the Cam Clay model. Berli et al. (2004) evaluated soil compaction by field traffic using FEM, also using a Modified Cam Clay model. Horn (1988) stated that the phenomenon of the soil compaction can be explained in the classical elasto-plastic conception of stress–strain phenomena by considering the soil as a material that reacts elastically up to a certain limit of stress; beyond that limit, any incremental stress results in the plastic deformation and this stress threshold for a given soil, under given climatic conditions, depended on soil texture, degree of aggregation, and matric potential.

### **Conclusion**

Considerable work is being done for different applications of numerical modeling in the field of soil compaction. These studies showed that The Finite Element Method has been currently the most advanced mathematical tool that can be used to modeling the process of artificial compaction of soil under the action of rolling bodies of tractors and agricultural machinery.

### **References**

- Berli M, Kirby JM, Springman SM, Schulin R. 2004. Modelling compaction of agricultural subsoils by tracked heavy construction machinery under various moisture conditions in Switzerland. *Soil Tillage Res.*, 73: 57-66
- Biriş SŞ, Vlăduţ V, Ungureanu N, Paraschiv G, Voicu G. 2009. Development and Experimental Testing of a FEM Model for the Stress Distribution Analysis in Agricultural Soil due to Artificial Compaction. *Agriculturae Conspectus Scientificus*, 74 (1): 21-29.
- Botta GF, Tolon Becerra A, Bellora Tourn F. 2009. Effect of the number of tractor passes on soil rut depth and compaction in two tillage regimes. *Soil Till Res.*103:381-386.
- Cueto OG, Coronel CEI, Bravo EL, Morfa CAR, Suárez MH. 2016. Modelling in FEM the soil pressures distribution caused by a tyre on a Rhodic Ferralsol soil. *Journal of Terramechanics*, 63: 61–67.

- Cuetoa OG, Corone CEI, Morfac CAR, Sosa GU, Gómez LHH, Calderón GU, Suárez MH. 2013. Three dimensional finite element model of soil compaction caused by agricultural tire traffic. *Computers and Electronics in Agriculture* 99: 146–152
- Defossez P, Richard G. 2002. Models of soil compaction due to traffic and their evaluation. *Soil and Tillage Research*, 67 (1): 41-64.
- Etana A, Larsbo M, Keller T, Arvidsson J, Schjønning P, Forkman J, Jarvis N. 2013. Persistent subsoil compaction and its effects on preferential flow patterns in a loamy till soil. *Geoderma.*;192:430-436.
- Horn R. 1988. Compressibility of arable land. Drescher J, Horn R., de Boodt M. (Eds.), *Impact of Water and External Forces on Soil Structure*, Catena Suppl., 11: 53-71
- Liu, C.H., Wong, J.Y., 1996. Numerical simulations of tire–soil interaction based on critical state soil mechanics. *J. Terramech.* 33, 209–221.
- Nasiri, M., M. Soltani, and A. M. Motlagh. 2013. Determination of agricultural soil compaction affected by tractor passing using 3D finite element. *Agric Eng Int: CIGR Journal*, 15(3): 11 – 16.
- Nawaz MF, Bourrié G, Trolard F. 2013. Soil compaction impact and modelling. A review. *Agron. Sustain. Dev.*, 33: 291–309
- Parlak M, Özaslan Parlak A. 2011. Effect of soil compaction on root growth and nutrient uptake of forage crops. *Journal of Food, Agriculture & Environment*, 9 (3-4): 275 - 278.
- Poodt MP, Koolen AJ, Van der Linden JP. 2003. FEM analysis of subsoil reaction on heavy wheel loads with emphasis on soil preconsolidation tension and cohesion. *Soil Tillage Res.*, 73: 67-76
- Raper RL, Johnson CE, Bailey AC. 1994. Coupling normal and shearing stresses to use in finite element analysis of soil compaction. *Transactions of the ASAE*, 37(5): 1417-1422
- Silva RP, Rolim MM, Gomes IF, Pedrosa EMR., Tavares UE., Santos AN. 2018. Numerical modeling of soil compaction in a sugarcane crop using the finite element method. *Soil and Tillage Research*, 181: 1-10.
- Soracco CG, Lozano LA, Villarreal R, Palancar TC, Collazo DJ, Sarli GO, Filgueira RR. 2015. Effects of compaction due to machinery traffic on soil pore configuration. *R. Bras. Ci. Solo*, 39: 408-415.
- Unger PW, Kaspar TC. 1994. Soil Compaction and Root Growth: A Review. *Agron. J.*, 86: 759-766.
- Ungureanu N, Vlăduț V, Biriș SȘ, Paraschiv G, Dincă M, Zăbavă BȘ, Ștefan V, Gheorghiu NE. 2018. FEM modelling of machinery induced compaction for the sustainable use of agricultural sandy soils. 46. Symposium "Actual Tasks on Agricultural Engineering", Opatija, Croatia., 201-211.
- Velloso NS, Costa ALG, Magalhães RR, Santos FL, De Andrade ET. 2018. The finite element method applied to agricultural engineering: A Review. 6 (3): 286-299.
- Xia K. 2011. Finite element modeling of tire/terrain interaction: Application to predicting soil compaction and tire mobility. *Journal of Terramechanics*, 48 (2): 113-123

## Öğretmenlerin Örgütsel Sinizm Algısının İncelenmesi

Murat AKYÜZ<sup>1</sup>, Taner TUNÇ<sup>1</sup>

<sup>1</sup> Ondokuzmayıs Üniversitesi, Fen Edebiyat Fakültesi, Samsun, Türkiye

### Özet

Örgütte sinizm kavramı iki farklı şekilde kullanılmaktadır. İlki, bireyin kendi kişiliğinden kaynaklanan ve hayata olan bakış açısını yansıtan genel sinizmdir. İkincisi, bireyde sinik davranış oluşmasına neden olan örgütsel sinizmdir. Örgütsel sinizm, kişinin örgütüne karşı olan negatif tutumudur. Örgütsel sinizm sadece olumsuz insanların örgüte kattığı duygular değildir, aynı zamanda bu tutumların çalışma ortamlarında deneyimler yoluyla paylaşılmasıdır. İnsanlar sinik olmaya karar vermezler. Örgütlerde sinizm, ilgisizlikten, yok sayılmaktan kaynaklı doğan bir tepkidir. Eğitim gibi girdisi çıktısı çok olan ve odak noktasında insan olan eğitim örgütlerinde örgütsel sinizm kavramı çok önem taşımaktadır. Bu çalışmanın amacı, eğitim çalışanlarının çalıştığı organizasyonlara karşı besledikleri negatif tutumlardaki algının ölçülmesidir. Bu amaçla Tokat ilinde ve ilçelerindeki liselerde görevli öğretmenlerden oluşan bir denek grubu ile çalışılmıştır. Deneklere 13 maddelik Örgütsel Sinizm Ölçeği ile demografik değişkenlerden oluşan soru formu verilmek suretiyle istatistiksel analiz verisi toplanmıştır. İstatistiksel analizler sonucunda öğretmenlerin örgütsel sinizm düzeylerinde cinsiyet, branş, meslekte aldığı ödül, mesleki kıdem yılı, görev yaptığı okuldaki çalışma yılı, görev yaptığı okul türü ve yerleşim birimi değişkenlerine göre anlamlı farklılıklar bulunmaktadır. Bununla beraber yapısal eşitlik modellemesi bilişsel sinizmin ve duyuşsal sinizmin davranışsal sinizim üzerinde etkili olduğunu göstermiştir.

**Anahtar kelimeler:** Eğitim, lise, örgütsel sinizm, öğretmen, yapısal eşitlik modellemesi

## Examination of Teachers' Perception of Organizational Cynicism

### Abstract

The concept of cynicism is used in two different ways in the organization. The first is general cynicism that arises from the individual's own personality and reflects his outlook on life. The second is organizational cynicism, which causes cynical behavior in the individual. Organizational cynicism is a negative attitude towards one's organization. Organizational cynicism is not only the emotions that negative people bring to the organization, but also the sharing of these attitudes through experiences in the work environment. People don't decide to be cynical. cynicism in organizations is a reaction arising from indifference and ignoring. The concept of organizational cynicism is very important in educational organizations such as education, which have a lot of input and output and focus on people. The aim of this study is to measure the perception of education staff in negative attitudes towards the organizations they work for. For this purpose, we worked with a sample group consisting of teachers working in high schools in Tokat and its districts. Statistical analysis data were collected by giving the subjects a 13-item Organizational Cynicism Scale and a questionnaire consisting of demographic variables. As a result of the statistical analysis, there are significant differences in the organizational cynicism levels of teachers according to the variables of gender, branch, awards in the profession, professional seniority, the year they work, the type of school they work and the place of residence. In addition, structural equation modeling has shown that cognitive cynicism and affective cynicism affect behavioral cynicism.

**Key words:** Education, high school, organizational cynicism, teacher, structural equation modeling

## Giriş ve Kuramsal Çerçeve

Örgütsel sinizmi ilk keşfeden araştırmacılardan biri olan Niederhoffer örgütsel sinizmi, kuralsızlığın bir türü olarak kendi kolluk gücü ve kamuya yönelik polislerin olumsuz tutumları olarak tanımlamakta ve polis teşkilatı üzerinde yaptığı çalışmalardan yola çıkarak sinizmi, polis dünyasının değerler sistemi üzerine kurulmuş ideolojik bir kalkan olarak ifade etmektedir (Niederhoffer, 1967: 9). Polislerin sinik davranışlarıyla ilgili olan bu durumun Durkheim'in anomi kavramından yola çıkarak açıklandığı bahsedilmektedir (Brandes, 1997: 13). Mino (2002) çalışmasında, sinik davranış gösteren insanların işe yönelik motivasyonları ve bağlılıklarında düşme, şikayetlerinde ve devamsızlıkta artma görüldüğünü ifade etmiştir.

Örgütsel sinizm özellikle 1990'lı yılların başından günümüze işletme yönetimi, örgütsel davranış, insan kaynakları yönetimi, iş etiği, halkla ilişkiler gibi farklı alanlardan araştırmacıların ilgisini çeken bir konu olmuştur (Tokgöz ve Yılmaz, 2008: 291).

Goldner vd. (1977), örgütsel sinizm çalışmalarında hedef olarak örgütü alırken, bunun başlıca psikolojik yöneliminin organizasyona ilişkin bireysel inançlara odaklanmak olduğunu söylemiştir. Buna karşılık sinizmi anlamak için sosyolojik bir yaklaşım benimsemişlerdir. Bu yaklaşım Merton'un adaptasyon paradigmasıyla uyumludur. Burada örgütsel sinizm yapısı doğrudan incelenemezken, tanımlama yapılarak sinik bilgi kavramından söz edilmektedir (Goldner vd., 1977: 541).

Sinizim, pek çok kez farklı açılardan incelenmiş ve farklı farklı tanımları ortaya çıkmıştır, bu tanımlardan bazıları; bir inanç olduğu, bir duygu olduğu, karakteristik bir özellik veya küresel bir özellik olduğu yönündedir. Sinizm kavramına mesleki yönden bakıldığında, polislik ve işçilik gibi yüksek stresi olan işler ilk olarak göz önüne alınmaktadır. Ancak en yaygın görüş, örgütsel sinizmin; örgüt hakkındaki negatif inanışlar, (örneğin, örgütün bütünlükten yoksunluğu), negatif duygular ve olaylara bağlı ortaya çıkan davranışlar olduğudur (örneğin, dürüstlük yoksunluğuna karşı somut davranışlar). Bu noktada, negatif davranışların örgüte karşı olduğunu belirtmek gerekir (Scott ve Zwig, 2008: 96).

### *Örgütsel sinizmin boyutları*

Örgütsel sinizmin bir tutum olarak kavramsallaştırılması, tutumlara uygulanan kuralların örgütsel sinizme de uygulanması anlamına gelmektedir. Bu kurallardan birisi de, tutumun; düşünmekle başlayan ve davranışla sona eren düşünce-davranış sürecinin bir parçası olmasıdır (Delken, 2004: 11).

Tutum, ortamsal faktörlerle etkileşim halinde davranışı oluşturmaktadır. Tutum, davranışın öncü uyarıcısı, bireyi davranışa hazırlayan bilgi birikimi olarak tanımlanabilir (Erdoğan, 1996: 33). Bireylerin inançları, tutumları ve değerleri, işe yönelik tavırlarını ve örgütsel ilişkilerini etkilemektedir (Sabuncuoğlu ve Tüz, 2001: 340).

Sinizim alt boyutlarına ait tutumlar bilişsel, duyuşsal ve davranışsal olmak üzere üç temel boyutta incelenmektedir. Tutumun oluşması için bu üç boyut arasında örgütsel ve uyumlu bir ilişki ve koordinasyon olması gerekmektedir (İnceoğlu, 2004: 25).

### *Bilişsel boyut*

Örgütsel sinizmin bilişsel boyutu örgütteki uygulamaların adalet, dürüstlük ve içtenlikten yoksunluğuna olan inancı içerir. Sinikler örgütlerinde bu ilkelerin sıklıkla ilkesizliğe kurban edildiğine ve ahlaksızca davranışların norm haline geldiğine inanırlar. Yine, sinikler insanların davranışlarında tutarsız ve güvenilmez olduğuna da inanabilirler. Sinikler sıklıkla eylemlerin arkasında gizli motiflerin olduğuna inanırlar ve açık olandan ziyade hileyi görmeye çalışırlar (Dean vd., 1998: 345-346).

### *Duyuşsal Boyut*

Örgütsel sinizm, sinik inançlarla bağlantılı düşüncelerden kaynaklanan yoğun duyguları kapsamaktadır (Brandes ve Das, 2006: 237).

Örgütsel sinizmin duyuşsal boyutu sinizmin düşünölen bir şey olmasının yanı sıra hissedilen bir şey olduğunu da gösterir. Sinizm örgüte ilişkin güçlü duygusal tepkileri de içerebilir. Bu olumsuz güçlü duygulara deęişik örnekler verilebilir. Sinikler örgütlerine karşı saygısızlık ve öfke hissedebilirler. Yine bu insanlar örgütlerini düşündüklerinde rahatsızlık, nefret ve hatta utanç duyabilirler. Dolayısıyla sinizm farklı olumsuz duygularla ilgilidir. Bununla birlikte sinikler örgütlerini kendi standartları ile tartarak kendilerinin örgütten üstün olduklarını düşünmekten gizli bir biçimde zevk alabilirler (Dean vd., 1998: 346).

### *Davranışsal Boyut*

Örgütsel sinizmin üçüncü ve son boyutu, örgüte yönelik aşağılayıcı ve eleştirel davranışlar gösterme eğilimi anlamına gelen davranış boyutudur. Örgütsel sinizmin son boyutu olan davranış boyutu, olumsuz ve çoğunlukla kötölemek yönündeki davranışlara karşı bir eğilimi ifade eder. Sinik kişilerde görölen davranışlar açısından en belirgin davranışsal eğilim örgüte yönelik sert eleştirilerdir. En belirgin şekilde ifade edilen türleri, örgütün dürüstlük, samimiyet vb. özelliklerden yoksun olduğu şeklindedir. Bunların yanında çalışanlar, mizahı özellikle de alaycı mizahı sinik tutumlarını ifade etmek için çok kullanırlar. Örgütsel sinikler, örgütte geleceęe dair eylemler hakkında kötümser tahminlerde bulunmaya eğilimli olabilirler. Örneğın, bir kalite girişiminin örgüt için pahalı olmaya başladığı anda hemen bırakılacağını tahmin edebilirler (Dean vd., 1998: 346).

### **Amaç ve Kapsam**

Örgütsel sinizm bireylerin ve örgütlerin verimliliğini düşürdüğü ve sinizmin örgütlerde bireyler arasında çok hızlı yayıldığı sonucuna varıldığı düşünölmektedir. Bu çalışmada özellikle Bilişsel Sinizmin ve Duyuşsal Sinizmin Davranışsal Sinizime ne derece etkisi olduğunu belirlemek ve öğretmenlerin demografik bilgilerine göre sinizm algılarının farklılıklarını ve dereceleri ölçölmeye çalışmıştır. Eğitimde örgütsel sinizim düzeyinin düşürölmesi için çözüm yolları aramak ve önerilerde bulunmak ikincil amaçtır.

Bu çalışmanın kapsamı dahilinde eğitimde örgütsel sinizm alt boyutlarının birbirlerini ne derecede etkilediğı tespit edilmiştir. Örgütsel sinizm ve alt boyutları teorik bazda irdelenmiş ve konuya ilişkin lise öğretmenlerine yönelik bir araştırma yapılmıştır. Araştırmada Tokat ili merkez ve ilçelerinde bulunan Milli Eğitim Bakanlığına bağılı liselerde çeşitli branşlarda görev yapan öğretmenler yer almaktadır. Öğretmenlerden gelen görüşler doğrultusunda çözüm önerileri de verilmiştir.

### **Yöntem**

Araştırmanın evrenini, Tokat İl ve İlçelerinde bulunan Milli Eğitim Bakanlığına bağılı 93 lisede görev yapan 2301 öğretmenden oluşmaktadır. Araştırmanın örneklemini 430 katılımcıdan oluşmaktadır. Veriler, kullanıcı adı ile kısıtlanmış, her kullanıcının bir kez doldurabildiğı on-line anket yöntemiyle toplanmıştır. Verilerin toplanması için 10 sorudan oluşın kişisel bilgi formu, Brandes vd. (1999) tarafından geliştirilmiş olan 13 ifadeden oluşın örgütsel sinizm ölçeğı kullanılmıştır. Araştırma kapsamında anket yöntemiyle elde edilen veriler SPSS ve LISREL paket programları ile analiz edilmiştir. Demografik özelliklerin analizi için temel düzeyde parametrik ve parametrik olmayan istatistiksel yöntemler ile çok deęişkenli bir istatistiksel yöntem olan yapısal eşitlik modellemesi ile sinizmin alt boyutları arasında istatistiksel olarak anlamlı yapısal bir model önerilmiştir.

## Bulgular

Çalışmada yer alan değişkenlerin oluşturduğu faktörlere ait güvenilirlik göstergesi olan Cronbach Alfa iç tutarlılık katsayıları Tablo 1 ile verilmiştir.

Tablo 1. Alt Boyutların Güvenilirlik Katsayıları

Örgütsel Sinizm Ölçeği (Alpha =0,933 ) Alt Boyutları	Cronbach's Alpha Değerleri
Bilişsel Sinizm Alt Boyutu	<b>0,927</b>
Duyuşsal Sinizm Alt Boyutu	<b>0,964</b>
Davranışsal Sinizm Alt Boyutu	<b>0,815</b>

Tablo 1'e göre tüm alt boyutların iç tutarlılık katsayılarının yüksek düzeyde güvenilir olduğu, maddelerin bir araya gelerek oluşturdukları alt boyutlarda aynı olguyu ölçtüğü görülmektedir.

Tüm alt boyutlarda ( Örgütsel sinizm, bilişsel, duyuşsal, davranışsal ) öğrenim durumuna göre anlamlı bir fark bulunmamıştır ( $p>0,05$ ). Ancak duyuşsal sinizm alt boyutunda kadınlarla erkekler arasında anlamlı bir farkın olduğu ( $p<0,05$ ) ve kadınların duyuşsal sinizm algısının erkeklere göre daha yüksek çıktığı görülmüştür. Bununla birlikte alt boyutlar mesleki kıdem açısından değerlendirildiğinde 1-5 yıl arası çalışan öğretmenlerin 6-10 yıl, 11-15 yıl, 16-20 yıl, 21 yıl ve üstü çalışan öğretmenlerle aralarında anlamlı bir fark olduğu ( $p<0,05$ ); mesleki kıdemi 1-5 yıl arası olan öğretmenlerin tüm boyutlarda ( Örgütsel sinizm, bilişsel, duyuşsal, davranışsal) sinizm düzeyinin yüksek olduğu görülmüştür. Ayrıca güzel sanatlar öğretmenlerinin diğer branşlardaki öğretmenlere göre sinizm puanlarının daha yüksek olduğu tespit edilmiştir ( $p<0,05$ ). Ancak; Okul türlerine göre bakıldığında Güzel sanatlar liselerinde çalışan görsel sanatlar öğretmenlerinin diğer liselerde çalışan görsel sanatlar öğretmenlerine göre sinizm düzeylerinin daha düşük olduğu görülmüştür( $p<0,05$ ). Bu ise Güzel Sanatlar Liselerinde görev yapan Görsel Sanatlar öğretmenlerinin bu okullarda daha mutlu ve iyi bir ortamda çalıştıklarının göstergesidir. Teşekkür ve takdir belgesi alan öğretmenlerin hiç ödül almayanlara göre sinizm düzeyleri daha düşüktür ( $p<0,05$ ). Öğretmenlerin görev yaptıkları okulların yerleşim yerlerine göre duyuşsal sinizm puanları açısından merkez ilçe, gelişmiş ilçe ve kırsal ilçeler arasında istatistiksel olarak anlamlı bir farklılık mevcut olup duyuşsal sinizm düzeyi en yüksek olan öğretmenler gelişmiş ilçelerde çalışanlardır ( $p<0,05$ ).

Örgütsel sinizmin alt boyutları olan bilişsel ve duyuşsal sinizmin davranışsal sinizm üzerindeki etkisini belirlemek için yapılan yapısal eşitlik modellemesinde önerilen yapısal modele ait uyum indeksleri Tablo 2'de verilmiştir.

Tablo 2. Yapısal Modele Ait Uyum İndeks Değerleri

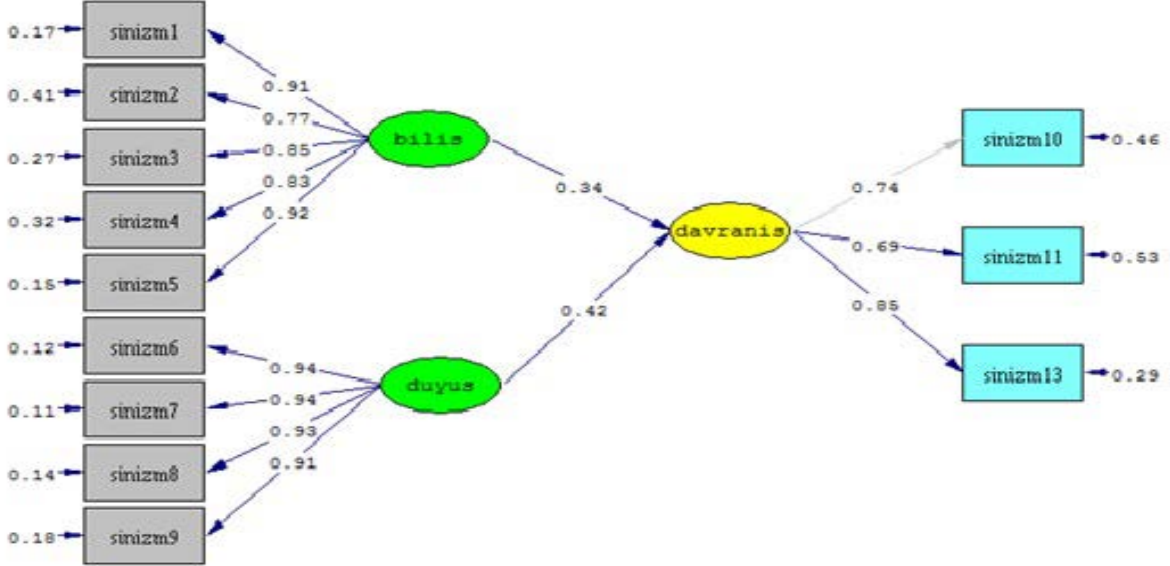
Uyum Ölçütleri	İyi Uyum	Kabul Edilebilir Uyum	Modele Ait Değerler
$X^2/sd$	$0 \leq X^2/sd \leq 2$	$2 \leq X^2/sd \leq 3$	2,908
RMSEA	$0 < RMSEA \leq 0,05$	$0 < RMSEA \leq 0,08$	0,078
NFI	$0,95 < NFI \leq 1,00$	$0,90 < NFI < 0,95$	0,960
NNFI	$0,95 < NFI \leq 1,00$	$0,90 < NFI < 0,95$	0,950
CFI	$0,95 < CFI \leq 1,00$	$0,90 < CFI < 0,95$	0,970
IFI	$0,95 < IFI \leq 1,00$	$0,90 < IFI < 0,95$	0,910
SRMR	$SRMR < 0,5$	$SRMR < 1$	0,047

Tablo 2'te önerilen yapısal modelin anlamlılığına ait uyum indekslerinin tümü model için iyi uyum değerleri verdiği ve kabul edilebilir düzeyde olduğu görülmüştür. Önerilen yapısal modele ait standart katsayılar t değerleri ve açıklama oranları Tablo 3 ile verilmiştir.

Tablo 3. Yapısal Modele Ait Standart Yükler ve Model Geçerlilik Ölçütleri

Maddeler	Standart Yükler	t-değeri	Yapı Güvenilirliği ve (Açıklanan Varyans)	R <sup>2</sup>
SinizmBilişsel1	0,91	7,31	0,9327 (0,7359)	0,83
SinizmBilişsel2	0,77	13,57		0,59
SinizmBilişsel3	0,85	12,01		0,73
SinizmBilişsel4	0,83	10,30		0,68
SinizmBilişsel5	0,92	6,72		0,85
SinizmDuyuşsal6	0,94	9,49	0,9618 (0,8628)	0,88
SinizmDuyuşsal7	0,94	8,64		0,89
SinizmDuyuşsal8	0,93	9,62		0,86
SinizmDuyuşsal9	0,91	11,72		0,82
SinizmDavranışsal10	0,74	10,63	0,828 (0,6179)	0,64
SinizmDavranışsal11	0,69	9,89		0,57
SinizmDavranışsal13	0,85	5,90		0,71
Bilişsel → Davranışsal	0,34	4,79		0,12
Duyuşsal → Davranışsal	0,42	5,76		0,18

Tablo 3'e göre tüm alt boyutlara göre önerilen modele ait yapı güvenilirliği 0,70 den yüksektir. Ayrıca her boyutta açıklanan varyans değerleri de 0,50 den yüksektir. Bu ise önerilen modelin başarılı bir olduğunu göstermektedir (Nunnally, 1978; Hair ve ark. 1998). Önerilen modele ait bilişsel ve duyuşsal sinizmin direk olarak davranışsal sinizme etkisinin görüldüğü path diyagramı Şekil 1'ile verilmiştir.



Şekil 1. Önerilen Yapısal Eşitlik Modeli

Path diyagramından elde edilen eşitlik;

$$\text{Davranışsal Sinizm} = 0,42 * (\text{Duyuşsal Sinizm}) + 0,34 * (\text{Bilişsel Sinizm}) \quad R^2 = 0,51$$

olarak verilebilir. Yapısal denkleme göre; Duyuşsal Sinizm ve Bilişsel Sinizm alt boyutları birlikte Davranışsal Sinizmdeki değişimin %51 'ini açıklamaktadır. Yani bu alt boyutlar Davranışsal sinizm üzerinde istatistiksel olarak önemli bir etkiye sahiptir.

## Sonuçlar

Toplular için önemli bir kaynak olan eğitimde öğretmenlerin önemi yadsınamayacak kadar fazladır. Öğretmenlik mesleği hem eğitim sisteminde hem de öğrencilerle ve velilerle olan iletişimiyle adeta 3 boyutlu bir örgütte var olmaktadır. Bu nedenle öğretmenlerin moral motivasyon gibi faktörlerle bilgilerini öğrencilere aktarması çok önemlidir. Ayrıca öğretmenlerin eğitim öğretim sürecinde sinizme muhatap olmaları kendi örgütlerine olan inançlarının kaybolmasına ve verimliliklerinin düşerek öğrencilere verdikleri eğitimin ve öğretimin kalitesini düşürmektedir. Özellikle sinizmin son halkası olan davranışsal sinizmde öğretmenin artık sinizm algısının davranışlarına yansması, bunun fiili olarak bulunduğu çevre tarafından gözlemlenmesi son derece olumsuz bir durumdur. Çünkü öğretmen bilişsel ve duyuşsal olarak sinizme kapılmış olabilir ama bu sinik durumun davranışlarına yansması endişe vericidir. Bu çalışmada bilişsel ve duyuşsal sinizmin davranışsal sinizm üzerinde önemli derece etki yaptığı ve hatta tetiklediği tespit edilmiştir. Davranışsal sinizm düzeyini düşürmek için özellikle duyuşsal sinizmin azaltılması gerekmektedir. Bunun için kurum yöneticilerine, olumsuz durumların önlenmesi noktasında; öğretmenlerin fikirlerine değer vermesi, yetki ve sorumluluk paylaşımı ile yönetime katılımlarının sağlanması, öğretmenlerin fikirlerini rahatça ifade etmesi ve yeni fikirler sunmasının teşvik edilmesi, öğretmenlerle samimi ve güvenilir bir iletişimin kurulması yoluyla bilgi paylaşımının sağlanması, göreve yeni başlayan öğretmenlere oryantasyon imkanı sağlanması gibi bir çok tedbir alınmalıdır. Öğretmenlerin motive edilerek aitik duygusunun kazandırılması, öğretmenlerin kendilerini geliştirebilmelerine imkan sağlanması ve bu bağlamda eğitimler verilmesi, öğretmenler arasında ayrımcılık yapılmaması ve etik değerlere uygun davranılması eğitim kalitesinin yükselmesi için gereklidir.

## Kaynaklar

- Brandes P.; Organizational Cynicism: Its Nature, Antecedents, and Consequences, Doktoral Thesis (Unpublished); University of Cincinnati, 1997.
- Brandes, P.; D. Das; “Locating Behaviour Cynicism at Work: Construct Issues and Performance Implications”, Employee Health, Coping and Methodologies (Edt. Pamela L. Perrewe, Daniel C. Ganster), JAI Press, New York, 2006, pp.233-266.
- Delken, M.; Organizational Cynicism: A Study Among Call Centers, A Thesis Submitted in Partial Fulfillment of The Requirements for The Degree of Master of Economics University of Maastricht, 2004.
- Dean, J. W.; P. Brandes; R. Dharwadkar; “Organizational Cynicism”, Academy of Management Review, Vol: 23 (2), 1998, pp.341-352.
- Erdoğan, İ.; İşletme Yönetiminde Örgütsel Davranış, İ.Ü. İşletme Fakültesi İşletme İktisadi Enstitüsü, Yayın No: 266/158, İstanbul, 1996.
- Goldner, F. H.; R. R. Ritti; T. P. Ference; “The Production of Cynical Knowledge in Organizations”, American Sociological Review, Vol: 42(4), 1977, pp.539-551.
- Hair, J. F., Anderson, R. E., Tatham, R. L. ve Black, W. C. 1998, Multivariate Data Analysis, Prentice Hall, New Jersey.
- İnceoğlu, M.; Tutum Algı İletişim, Elips Yayınevi, Ankara, 2004.
- Mino, C. E.; Organizational Trust, Organizational Cynicism and Organizational Commitment During A Change Initiative, Unpublished Doctoral Dissertation, Alliant International University, Los Angeles California, 2002.
- Niederhoffer, A.; Behind The Shield: The Police in Urban Society, Garden City, NJ: Anchor, 1967.
- Nunnally, J. C.; Psychometric theory (2. Baskı). New York: McGraw-Hill.1978.
- Sabuncuoğlu, Z.; M. Tüz; Örgütsel Psikoloji, Ezgi Kitabevi, Bursa, 2001.
- Scott, K. A.; D. Zweig; Dispositional Predictors of Organizational Cynicism, Department of Management University of Toronto Scarborough, 2008.
- Tokgöz, N.; H. Yılmaz; “Örgütsel Sinizim: Eskişehir ve Alanya’daki Otel İşletmelerinde Bir Uygulama”, Anadolu Üniversitesi Sosyal Bilimler Dergisi, Cilt: 8 (2), 2008, s.283-305.



## **Problems Encountered in Organic Almond Growing; in Case of Dicle and Eğıl**

Songül AKIN<sup>1</sup>, Abdurrahman KARA<sup>1</sup>

<sup>1</sup>Department of Agricultural Economics, Faculty of Agriculture, Dicle University, 21280, Sur,  
Diyarbakır, Turkey

### **Abstract**

Almond, which is one of the hard-shelled fruits, is very rich in its nutrient content. Therefore, it is very important in healthy nutrition. The number of almond trees Diyarbakır, Turkey constitutes 3.9% of the assets of the almond tree. Dicle and Eğıl districts selected as study areas constitute 13.8% of the Diyarbakır almond tree presence. In order to spread organic almond cultivation, various institutions have been supported in these districts since 2011. In the study, the problems encountered in organic almond cultivation in these districts were investigated by conducting a questionnaire. It has been determined, the almond stone wasp (*Eurytoma amygdali*) has been the most important problem causing economic loss. In addition, it was determined that the most limiting factor for organic almond cultivation is to find workers.

**Key words:** Organic agriculture, almond producers, problems

**Almond Producers' Perspectives towards Agricultural Organizations: In Case of Dicle  
and Eğil**

Songül AKIN<sup>1</sup>

<sup>1</sup>Department of Agricultural Economics, Faculty of Agriculture, Dicle University, 21280, Sur,  
Diyarbakır, Turkey

**Abstract**

Agricultural organizations are the organizations operating according to the principle of farmers' self-help. Organizational success depends mainly on their members because they literally rely on democratic governance and members' voluntary participation. Thus, it is a bottom-up movement, and so share holders' or members' organizational perspectives and expectations have generally been positive and high. In this study was examined the organizational perspectives of the almond farmers who were members of the organic producers' association and operating in Dicle and Eğil districts of Diyarbakır province, Turkey. To this end, in total of 52 almond farmers in both districts were included in the study using complete enumeration sampling method. Study data were collected with questionnaires through face-to-face farmer interviews and qualified conversations. In analysis of the data were used descriptive statistics and Chi-Square analysis. In ordinal data obtained with Likert scales, an internal consistency analysis was also performed calculating Cronbach's alpha statistics. It was revealed that of all the members of Organic Almond Producers' Association, 78.9% became member just to access the government supports as 42.3% never visited the Association in one year of period and generally not agreed to the proposal saying "farmer associations have important roles in marketing." According to the results we concluded that organic producers' associations in the study area hitherto could never manage to accomplish their marketing functions and hence there was not any expectation formed in farmers' mind that associations might have an active role in marketing

**Key words:** Agricultural organization, organizational function, expectation

## Rüzgar Hızının İstatistiksel Analizi

Abdulkadir KESKİN<sup>1</sup>, Saniye DEMİR<sup>2</sup>, Bilal Habeşi ÖZKAYNAR<sup>3</sup>, Mehmet Şamil GÜNEŞ<sup>4</sup>,  
Kadir Aytaç ÖZAYDIN<sup>5</sup>, Coşkun PARİN<sup>4</sup>, Abdülkadir BEKTAŞ<sup>7</sup>

<sup>1</sup>İstanbul Medeniyet Üniversitesi, Siyasal Bilgiler Fakültesi, İşletme Bölümü, Sayısal Yöntemler  
Anabilim Dalı, İstanbul, Türkiye

<sup>2</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Toprak Bilimi ve Bitki Besleme Bölümü, Tokat,  
Türkiye

<sup>3</sup>Türkiye İstatistik Kurumu, Çankaya, Ankara

<sup>4</sup>Yıldız Teknik Üniversitesi, Tıp Fakültesi, İstatistik Bölümü, İstatistik Teorisi Anabilim Dalı, İstanbul,  
Türkiye

<sup>5</sup>Tarımsal Araştırmalar ve Politikalar Genel Müdürlüğü

<sup>6</sup>Enerji ve Tabii Kaynaklar Bakanlığı

### Özet

Yaşadığımız yerküre üzerinde, enerji kaynaklı yakıtlar hızla azalmaktadır. Kömür, fosil yakıtlar gibi yenilenemeyen kaynaklardan enerji elde etme sürecinde salınan sera gazları ve radyoaktif yüzeyler, yerkürenin sıcaklığının artmasına neden olmakta ve iklim değişikliğine yol açmaktadır. Tüm bu faktörler, enerji oluşturmak için yenilenebilir kaynakları açıklayan araştırmacıları zorlamaktadır. Rüzgar, yenilenebilir bir kaynaktır ki; doğal olarak bulunur ve nerdeyse tükenmez bir kaynaktır. Bu çalışmada; Trabzon- Merkez meteoroloji istasyonuna ait 1967-2018 yılları arasındaki aylık, mevsimsel ve yıllık rüzgar verilerinin analizi yapılmıştır. Trend analizi için linear regresyon. Mann-Kendall ve Spearman Rho testi kullanılırken; değişim noktasının belirlenmesi için Pettitt's test, von Neumann ratio test, Buishand's range test ve standard normal homogeneity (SNH) test kullanılmıştır. Trend analiz sonuçları ise; mart ayında yükseliş ve kasım ayında ise düşüş eğilimindedir. Sadece Pettitt's testi için, 2013 yılında değişim noktası belirlenmiş, diğer yıllarda değişim gözlenmemiştir.

**Anahtar kelimeler:** İklim değişikliği, iklim değişkenleri, Trend analizi, değişim noktası, önemlilik düzeyi

## Statistical Analysis of Wind Speed

### Abstract

The fuel energy resources around the globe are declining at a faster rate. The greenhouse gases emitted in the process of harnessing energy from the non-renewable resources such as coal, fossil-fuel and radioactive substances cause the global warming and there is a danger of climate change. All these factors compel the researchers to explore renewable sources for power generation. Wind is a renewable resource, which is available naturally and is practically inexhaustible. In this study, the monthly, seasons and annual wind data collected between the years 2013 to 2015 is evaluated. The change detection analysis has been conceded using Pettitt's test, von Neumann ratio test, Buishand's range test and standard normal homogeneity (SNH) test, while non-parametric tests including linear regression, Mann-Kendall and Spearman rho tests have been applied for trend analysis. Trend analysis results are; It tends to increase in March and decrease in November. Only for Pettitt's test, the change point was determined in 2013, no change was observed in other years.

**Key words:** Climate change, climatic variables, Trend analysis, change point, significance level

## **Türkiye'nin Farklı Bölgelerindeki Aylık ve Yıllık Yağışların Dağılım Fonksiyonlarının Belirlenmesi**

Saniye DEMİR<sup>1</sup>, Yunus AKDOĞAN<sup>2</sup>, Orhan KAVUNCU<sup>3</sup>, Aşır GENÇ<sup>4</sup>, Bilal Habeşi ÖZKAYNAR<sup>5</sup>,  
Kenan KILIÇ<sup>6</sup>

<sup>1</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Toprak Bilimi , Taşlıçiftlik 60240, Tokat, Türkiye

<sup>2</sup>Selçuk Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 42250, Konya, Türkiye

<sup>3</sup>Kastamonu Üniversitesi, Mühendislik Fakültesi, Genetik ve Biyomühendislik Bölümü, Kastamonu, Türkiye

<sup>4</sup>Necmettin Erbakan Univ Fen Fakültesi, İstatistik Bölümü 42090, Konya, Türkiye

<sup>5</sup>Türkiye Enstitüsü Kurumu, Çankaya, 06570, Ankara, Türkiye

<sup>6</sup>Niğde Ömer Halisdemir Üniversitesi, Çevre Mühendisliği, Niğde, Türkiye

### **Özet**

Olasılık dağılım fonksiyonları ve yağış frekans analizleri, istatistiksel yağış modellerinde ve iklim değişikliğinin değerlendirilmesinde önemli bir rol oynamaktadır. Özellikle ağır yağışlar, nehir kenarlarında bulunan yerleşim alanlarındaki zararları engellemek ya da minimize etmek için hidrolik yapıların dizaynında kullanılmaktadır. Türkiye, çok çeşitli iklimlere sahiptir. Herbir bölgede, özel iklim karakteristikleri bulunmaktadır. Bundan dolayı, her bölgede meydana gelen yağışların değişimi ayrı ayrı değerlendirilmelidir. Bu çalışmada Karadeniz, Geçit Bölgesi ve İç Anadolu Bölgesi'ne ait 1967-2018 yılları arasındaki 52 yıllık zaman aralığında yer alan yağış verileri kullanılmıştır. Yağış verilerinin değişim ve trendi normal (N), log-normal (LN), three-parameter log-normal (LN III), Gamma (GAM), Pearson type III (P III), Gumbel, three-parameter Weibull (WE III) and logarithmic logistic (LLO) dağılımlar kullanılarak belirlenmiştir. Bu dağılımların uygunluğu Kolmogorov-Smirnov (K-S), Anderson-Darling and Cramér -von Mises testleri ile değerlendirilmiştir. Bu uyum iyiliği testlerinin kullanılmasıyla, her bir bölge için aylık ve yıllık yağışlar için en uygun dağılımlar belirlenmiştir.

**Anahtar kelimeler:** Frekans analizi, olasılık dağılımları, tahmin, uyum iyiliği testleri, Türkiye

## **Determination of the Distribution Function of Months and Annual Rainfall in Different Regions of Turkey**

### **Abstract**

Probability density functions and precipitation frequency analyses play an important role in statistical precipitation models and climate change assessments. Especially heavy precipitations are used in the design of hydraulic structures to minimize or prevent the damages in settlements along the riversides. Turkey has a diverse range of climates. Each region has specific climate characteristics. Therefore, change in precipitations should be assessed separately for each region. In this study, 52-year precipitation data covering the years 1967-2018 of precipitation stations in the Black Sea, Transitional Area and Central Anatolia Climate Zone of Turkey were used. Change and trends in precipitation data were assessed through the use of normal (N), log-normal (LN), three-parameter log-normal (LN III), Gamma (GAM), Pearson type III (P III), Gumbel, three-parameter Weibull (WE III) and logarithmic logistic (LLO) distributions. The fit of these distributions to available data was checked with the use of

various goodness of fit tests and Kolmogorov-Smirnov (K-S), Anderson-Darling and Cramér -von Mises test. With the use of this goodness of fit tests, best fitting distributions were determined for monthly and annual precipitations of each region.

**Key words:** frequency analyses, probability distribution, prediction, the goodness of fit tests, Turkey.

## Türkiye'nin Farklı Bölgelerindeki Yağışların Mevsimsel ve Yıllık Yağış Trendleri

Bilal Habeşi ÖZKAYNAR<sup>1</sup>, Yunus AKDOĞAN<sup>2</sup>, Saniye DEMİR<sup>3</sup>

<sup>1</sup>Türkiye Enstitüsü Kurumu, Çankaya, 06570, Ankara, Türkiye

<sup>2</sup>Selçuk Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 42250, Konya, Türkiye

<sup>3</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Toprak Bilimi, Taşlıçiftlik 60240, Tokat, Türkiye

### Özet

Türkiye iklim değişikliğinden en çok etkilenen Akdeniz havzasında bulunmaktadır. Özellikle, kurak ve yarı-kurak alanlar iklim değişikliği ve su eksikliğini içeren ciddi çevresel problemlere maruz kalmaktadır. Tüm ülkelerde yağışlarda görülen değişimler farklı olmaktadır. Bazı yerlerde çok, bazı yerlerde ise azdır. En fazla değişimler mevsimsel yağış dağılımlarını etkileyen ekstrem yağışlarda görülmektedir. Bölgelere bağlı olarak bu değişiklikler kuraklık ya da sel şeklinde görülmektedir. Gelecekteki iklim şartlarında bunlar daha sık ve daha yıkıcı bir etkiye sahip olacaktır. Bu çalışmada, Türkiye'nin farklı iklim kuşağındaki 6 istasyona ait 1980-2018 uzun yıllar yağış verileri kullanılmıştır. Mevsimsel ve yıllık yağışları analizinde Mann-Kendall ve Sen's slope testi uygulanmıştır. Bu analizler sonucunda, ilkbahar mevsiminde artış, sonbahar ve kış mevsiminde azalmalar gözlenmiştir.

**Anahtar kelimeler:** Mann-Kendall testi, Sen's slope testi, yağış trendi, iklimsel değişiklik,

## Seasonal and Annual Precipitation Trends in Rainfall in Different Regions of Turkey

### Abstract

Turkey is located in the Mediterranean basin most affected by climate change. In particular, arid and semi-arid areas are subject to serious environmental problems including climate change and water shortages. Changes in precipitation are different in all countries. It is less in some places while it is high in some places. The greatest changes are observed in extreme precipitation affecting seasonal precipitation distributions. Depending on the regions, these changes are seen as drought or flood. In future climatic conditions, these will have a more frequent and more destructive impact. In this study, the precipitation data of 6 meteorological stations between 1980-2018 were analyzed in different climate zones in Turkey. Mann-Kendall and Sen's slope test was applied in the analysis of seasonal and annual precipitation. As a result of these analyzes, it was observed that there was an increase in the spring season and a decrease in the autumn and winter season.

**Key words:** Mann-Kendall test, Sen's slope test, precipitation trend, climate change

## Hipotez Kontrolünde, Çeşitli Dağılımlardan Çekilen Örneklerin Genişliği Kaç Olmalıdır?

Tahsin KESİCİ<sup>1</sup>, Orhan KAVUNCU<sup>2</sup>, Zahide KOCABAŞ<sup>3</sup>, Saniye DEMİR<sup>4</sup>

<sup>1</sup>TOBB Üniversitesi, Ankara, Türkiye

<sup>2</sup>Kastamonu Üniversitesi, Mühendislik Fakültesi, Genetik ve Biyomühendislik Bölümü, Kastamonu, Türkiye

<sup>3</sup>Ankara Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Kastamonu, Türkiye

<sup>4</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Toprak Bilimi, Tokat, Türkiye

### Özet

Hipotez kontrolü yaparken kullanılan örneğin mümkün olduğunca geniş olması istenir. Ancak masraf ve zaman gibi faktörler araştırmacıyı örnek genişliğini sınırlı tutmaya zorlar. Ayrıca bazı durumlarda örnek genişliğini gereğinden büyük tutmak deney hatasının da büyümesine yol açabilir. Bu yüzden bilimsel araştırmalarda örneklerin optimal bir genişlikte olması arzu edilir. Örnek genişliğini belirlerken örneğin çekildiği popülasyonun ihtimal yoğunluk dağılımı, birinci tip hata, ikinci tip hata, testin gücü ve etki büyüklüğü (örnekleme hatası) gibi faktörler dikkate alınır. Bu çalışmada örnek genişliğini belirlemede dikkate alınan bu faktörler tartışılmış, bilhassa hipotezle tavsif edilen örnekleme dağılımının normal dağılım göstermesi ön şartı üzerinde durulmuştur. Merkezi limit teoremine göre bu ön şart örnek genişliği sonsuza giderken gerçekleşir. Ancak normal dağılım gösteren popülasyonlardan çekilen örneklerde daha küçük genişliklerle bu ön şartın sağlandığı bilinmektedir. Bu çalışmada çeşitli dağılımlardan simülasyon yöntemiyle üretilen örneklerin ortalamalarının dağılımının normal dağılım gösterip göstermediği, değişen örnek genişlikleri için araştırılmıştır.

**Anahtar kelimeler:** Hipotez kontrolü, örnekleme hatası, merkezi limit teoremi, simülasyon

### What Should be the Width of Samples Taken from Various Distributions in Hypothesis Control?

#### Abstract

In testing hypotheses, the sample size is desired to be large as much as possible. However, some factors like money and time force the researcher to work with the smaller samples. Besides sometimes too large samples may cause the experimental error also too big. Therefore, in scientific researches the samples are tried to have an optimal size. To assign the sample size there are some factors such as the probability density function of the hypothetic population, the probabilities of the first and second type errors, the power of the test and the sampling error (the size of the effect) to be taken into account.

In this study, these determining factors have been discussed, especially it is focused on the assumption that hypothetic sampling distribution should be normal. According to the central limit theorem, this assumption is realized when the sample size is infinitely large. However, even while the random samples drawn from normal populations are smaller, the assumption is consistent. In this study, the sampling distributions of the means of the random samples drawn from various distributions are simulated for different sizes and detected if they are distributed normal.

**Key words:** Hypothesis checking, sampling error, central limit theorem, simulation

## **Tavuklarda Yumurta Ağırlığı ile İlişkili Genlerin Farklı Çoklu Hipotez Test Yöntemleri ile Araştırılması**

Ümit BİLGİNER<sup>1</sup>, Burak KARACAÖREN<sup>1</sup>

<sup>1</sup>Akdeniz Üniversitesi, Ziraat Fakültesi, Antalya, Türkiye

### **Özet**

Tavukçuluk sektörü günümüzde önemli bir endüstri olmuş ve halen ilerlemeye devam etmektedir. Türkiye’de tavukçuluk sektörü sürekli büyüyen, ihracatını hızlı bir şekilde artıran, yoğun iş istihdamı sağlayan ve tarımı destekleyen yapısıyla Türkiye ekonomisine büyük ölçüde katkı sağlayan sektörlerden biridir. Bu çalışmada tavuklarda yumurta ağırlığı ile ilişkili genler farklı çoklu hipotez düzeltme yöntemleri ile incelenmiştir. Literatürde tavuklarda yumurta ağırlığından sorumlu genetik ve genomik faktörlerle ilgili sınırlı sayıda çalışma mevcuttur. Bu çalışmanın materyalini Rhode Island ırkı tavuklardan ( $n=1078$ ) toplanan zamana dayalı (başlangıç, 28, 36, 56, 66, ve 72 haftalık ) yumurta ağırlıkları ve 294.705 SNP oluşturmaktadır. Genomik çalışmalarda çok fazla sayıda hipotez testinin yapılması TİP 1 ve TİP2 hatalarının miktarının artmasına neden olmaktadır. Bu çalışmada çoklu hipotez test yöntemlerinden: yanlış keşif oranı, hochberg, bonferroni gibi yöntemlere ait sonuçlar genomik analiz için tartışılmıştır.

**Anahtar kelimeler:** Genomik seleksiyon, genomik ilişki çalışması, yumurta ağırlığı, tavukçuluk sektörü

## **Investigation of Genes Associated with Egg Weight in Chickens Using Different Multiple Hypothesis Test Methods**

### **Abstract**

The poultry sector has been an important industry today and still continues to progress. The chicken industry is constantly growing, increase exports quickly, providing employment intensive business and the structure supporting the agricultural sector which is one of Turkey's economy on a large scale contributed. In this study, genes associated with egg weight in chickens were examined using different multiple hypothesis correction methods. There are limited studies in the literature on genetic and genomic factors responsible for egg weight in chickens. The material of this study consists of time-based (baseline, 28, 36, 56, 66, and 72 weeks old) egg weights and 294.705 SNPs collected from Rhode Island breed hens ( $n = 1078$ ). Excessive hypothesis testing in genomic studies leads to an increase in the amount of TYPE 1 and TYPE 2 errors. In this study, the results of multiple hypothesis testing methods: false discovery rate, hochberg, bonferroni, were discussed for genomic analysis.

**Key words:** Egg weight, genomic association study, genomic selection, poultry sector



## **Sinop İli'nde Hava Sıcaklıklarının İstatistiksel Olarak Değerlendirilmesi**

Bilal Habeşi ÖZKAYNAR<sup>1</sup>, Saniye DEMİR<sup>2</sup>, Burhan GÜRLEYEN<sup>2</sup>, Hayati GÖNÜLTAŞ<sup>2</sup>

<sup>1</sup>Türkiye İstatistik Kurumu, Çankaya, 06570, Ankara, Türkiye

<sup>2</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Toprak Bilimi, Taşlıçiftlik 60240, Tokat, Türkiye

### **Özet**

Sinop iline ait 1971-2018 yılları arasındaki mevsimsel, aylık ve yıllık sıcaklık verileri analiz edilmiştir. Mann- Kendall analizi ile değişimin yönü, Sen's slope ile olasılık dağılımları belirlenmiştir. 2040, 2065 ve 2090 yılları için projeksiyon analizi yapılmıştır. Analiz sonuçlarına göre sıcaklık giderek artmaktadır.

**Anahtar kelimeler:** Mann\_Kendall, Sen's slope, sıcaklık, iklim değişikliği, projeksiyon

### **Statistical Evaluation of Air Temperature in Sinop Province**

#### **Abstract**

Seasonal, monthly and annual temperature data of Sinop province between 1971-2018 were analyzed. Probability distributions are determined with Sen's slope; while determining the direction of change with Mann-Kendall analysis. Projection analysis has been made for the years 2040, 2065 and 2090. According to the analysis results, the temperature is gradually increasing.

**Key words:** Mann\_Kendall, Sen's slope, temperature, climate change, projection

**Producing Short-Term Forecasts of Turkish Gross Domestic Product: A Bottom-Up  
Approach from the Demand Side**

Barış SOYBILGEN<sup>1</sup>

<sup>1</sup>Istanbul Bilgi University, Santral İstanbul Kazım Karabekir Cad. No:2 Eyüpsultan 34060 İstanbul,  
Türkiye

**Abstract**

Gross domestic product is a widely tracked indicator by all economic actors to assess the current situation of the economy. However, national account statistics are released with a 10 to 13 weeks delay in Turkey. Therefore, it is crucial to produce short-term predictions of national accounts using indicators that are released with a shorter delay. Unlike the previous literature that only focuses on predicting headline gross domestic product, we aim to predict gross domestic product through its demand-side subcomponents indirectly in this study. To predict gross domestic product's subcomponents, we use traditional linear methodologies and machine learning models such as lasso regressions, ridge regressions, elastic net, bagged decision trees, random forests, boosted decision trees, dynamic factor models, and bridge equations. After producing short-term predictions of each subcomponent, we combine them according to their share in gross domestic product. We also predict headline gross domestic product directly for comparison. Our results show that forecasting gross domestic product directly yields slightly better results than forecasting indirectly. However, indirect predictions provide more information about the current and future state of the economy. Our results also show that machine learning models have better forecasting performance compared to traditional linear methods.

**Key words:** Dynamic factor model, forecasting, gross domestic product, machine learning, Turkey

## **PISA 2018 Testi Türkiye Örneğine Göre Öğrencilerin Eğitim ve Sosyal Faaliyetlerinde Dijital Cihaz ve İnternet Kullanımlarının Değerlendirilmesi**

Gülin DELİKANLI<sup>1</sup>, Doğan YILDIZ<sup>1</sup>

<sup>1</sup>Yıldız Teknik Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, İstanbul, Türkiye

### **Özet**

Araştırma kapsamında 2018 PISA Testi Türkiye örneğinin bilişim ve iletişim teknolojilerine yatkınlık anket cevapları baz alınarak öğrencilerin dijital cihaz ve internet kullanımları ilk olarak cinsiyete, okul türlerine göre ve daha sonrasında ise genel olarak ele alınmıştır. Analizler aşamasında ki-kare bağımsızlık testi, güvenilirlik analizi, açıklayıcı faktör analizi ve son olarak çoklu faktör regresyon analizi uygulanmıştır. Analizler sonucunda, erkek öğrencilerin kız öğrencilere kıyasla dijital cihazlara ve internete daha kolay eriştikleri, dijital cihazları ve interneti kullanmayı ise kız öğrencilerin daha çok sevdikleri ortaya çıkmıştır. Kız öğrenciler daha çok kullanmayı isterken erişim oranlarının düşük olması ise toplumsal cinsiyet eşitsizliğinin göstergesidir. Fen Lisesi ve Sosyal Bilimler Lisesi öğrencilerinin dijital cihaz ve internete erişim oranlarının yüksek olmasının, eğitimdeki başarılarının da yüksek olmasını etkileyen etmenlerden biri olduğu sonucuna ulaşılmıştır. Öğrencilerin dijital cihazlara olan sevgilerini etkileyen faktörlerin de dijital cihazlara olan ilgi, dijital cihazlardan ve internetten edinilen tecrübe, özgüven ve sosyal çevreleri ile dijital cihazlar ve internet hakkındaki paylaşımları olduğu sonucuna varılmıştır.

**Anahtar kelimeler:** Açıklayıcı faktör analizi, çoklu faktör regresyon analizi, dijital cihaz, internet, PISA

### **Assessment Of Digital Device And Internet Use In Students' Educational And Social Activities According To PISA 2018 Test Turkey Sample**

#### **Abstract**

In the scope of the research, 2018 PISA Test Turkey's survey results about information and communication technologies are taken as base and students' digital device and internet usage inclinations examined according to sexes and school type, and later generally. In the process, chi-square independence test, reliability analysis, exploratory factor analysis and finally multi-factor regression analysis were carried out. As a result of analyses, it is revealed that male students have more easy access to the digital devices and internet, and female students love using digital devices and internet more. Also, female students want to use digital devices more but the low access rates indicate gender inequality. It is determined that students with high rate access to internet and digital devices at Science High School and Social Sciences High School is one of the factors affecting the high success of education. It is concluded that the factors affect students' love for the digital devices are the interest, the experience and self-confidence gained from digital devices and the internet, and the communication they have with their social circle through them.

**Key words:** Digital devices, exploratory factor analysis, internet, multi-factor regression analysis, PISA

## Giriş

Ekonomi ve toplumun ortak ürünü olan bilgi birikiminin teknoloji sayesinde artış göstermesi Dijital Çağ'ı yani Bilgi Çağı'nı başlatmıştır. 1900'lü yıllarla başlayan teknik keşifler önce elektriğin ve elektrikli aletlerin keşfi daha sonrasında ise internetin keşfedilmesiyle kendini geliştirmiş ve günümüzde de geliştirmeye devam etmektedir.

Tarım, ekonomi, bankacılık, tıp, astronomi ve daha birçok sektörde etkisini gösteren dijital dönüşüm, ülkelerin kalkınmasında ve gelişmesinde büyük rol oynamaktadır. Dijital çağa öncülük eden ve yenilikleri yakından takip eden ülkeler fark yaratmakta ve bu farkı artırmak için dijital yatırımlarını güçlendirmektedir. Bu noktada özellikle dijital çağda büyüyen çocukları anlamak hem toplum bilimi hem de ülkelerin kalkınma planlarını yönlendirmeleri açısından önemlidir. Bir öğrenci dijital dünyaya kendini ne kadar yakın hissederse, yaşadığı toplumda ve dolayısıyla ülkede buldukları çağa ve geleceğe yönelik o derece gelişim göstermektedir.

Dijital cihaz ve internet kullanımının kitlelerce yaygınlaşmasıyla hemen hemen yediden yetmiş her kesimden insanın hayatı dijital dünyadan etkilenmektedir. Bu bağlamda en çok etkilenen kesimlerin başında, dijital dünyadaki yenilikler aracılığıyla yaşadığı çağı ve dünyayı algılamaya çalışan, kendini ifade etmek isteyen genç bireyler gelmektedir. Gençlerin kendini tanıma ve keşfetmede dijital ürün ve interneti bir araç olarak görmeleri ile bilgiye ulaşma aşamasında da sanal kaynakları temel referans olarak algılamaları, onları dijital çağın en önemli özneleri konumuna getirir. Dijital ürüne ve internette ulaşılabilirlik durumlarına göre ev yaşantıları, okul yaşantıları ve sosyal yaşantıları gibi hayatlarının her alanında dijital dünyadan faydalanmaktadırlar. Dijital ürün ve internete ulaşılabilirlik konusunda ise her genç bireyin eşit şartlara sahip olmadığını görüyoruz. Bu eşit olmayan şartların nedeniyse; sosyoekonomik, coğrafi konum, ebeveynlerin yaşadıkları çağa karşı tutumları ve benzeri faktörlerdir. Genç bireyler arasındaki bu eşitsizlik eğitimlerine yansımakta ve toplumsal olarak kalkınmayı sağlayan gençlerin yeni dünyaya adapte olamamasıyla ülkelerin gelişmişlik seviyelerinin belirlenmesinde rol oynamaktadır.

Ülkelerin geleceği olarak görülen gençler ve gençlerin içinde bulunduğu dijital dönüşüm ülkelerin gelişimine yön veren etmenlerdendir. Özellikle son dönemlerde uluslararası geniş ölçekli test uygulamaları, genç bireyler üzerinde yapılan araştırmalarla ülkelerin eğitim kalitesini ilgilendiren etmenleri belirleyerek eğitim politikası ve yatırımlarına yön vermelerini sağlar. Değişen ve gelişen stratejileri ile ülkelerin dijitalleşen dünyada nerede olduklarını belirlemeleri ve güçlü ülkeler konumunda yer alabilmeleri için bu testleri düzenli olarak yapmaları ve analiz etmeleri gerekmektedir. Ayrıca testleri yorumlamak ve bu doğrultuda kararlar almak analiz etmek kadar önemlidir. Ülkelerin bu noktada sıkça başvurduğu testlerden birisi olan PISA testi bu araştırmada ele alınacaktır (Dönmez, 2019).

Bu çalışmada, en son yayınlanan PISA uygulaması olan PISA 2018 testindeki Türkiye örnekleminin bilişim ve iletişim teknolojilerine yatkınlık anket sonuçlarından elde edilen veriler kullanılarak, Türkiye'deki genç bireylerin dijital cihazlara ve internete olan yaklaşımlarını değerlendirmek ve dijital çağa ne kadar adapte olabildiklerini belirlemek amaçlanmıştır.

## Literatür Taraması

**Dijital Teknolojiye Yönelik Tutum Ölçeği (Emine CABI, 2015):** Makalede, Dijital Teknolojiye Yönelik Tutum (DTYT) ölçeğinin geçerlilik ve güvenilirlik çalışması konu olarak alınmıştır. Örneklemi 2013-2014 eğitim-öğretim yılında Ankara ili Yenimahalle ilçesine bağlı beş farklı ortaöğretim kurumundaki 689 öğrenci oluşturmaktadır. Analiz aşamasında DTYT ölçeğine faktör analizi uygulanmış ve 39 maddeyi içeren sekiz faktörlü bir yapıdan oluştuğu görülmektedir. Makalenin sonucunda sekiz faktörlü yapıyı oluşturan faktörler “yetkinlik”, “sosyal ağlar”, “derste teknoloji

kullanımı”, “teknolojiye yönelik ilgi”, “benim için teknoloji”, “olumsuz yönler”, “eğlence amaçlı kullanım” ve “bilinçli kullanım” olarak adlandırılmıştır. Son yıllarda, öğrencilerin boş zamanlarını çoğunlukla bilgisayar aktiviteleri ile geçirdikleri görülmüştür (Cabı, 2015a).

**Z Kuşağının Dijital Teknolojiye Yönelik Tutumları (Pınar ERTEN, 2019):** Makalede, 2000 yılından sonra dünyaya gelen ve teknoloji ile iç içe olan gençlerin oluşturduğu Z kuşağının dijital teknolojiye yönelik tutumları ele alınmıştır. Örneklemi 2017-2018 öğretim yılında, Bingöl ili Merkez ilçesindeki farklı devlet liselerinde öğrenim gören 318 öğrenci oluşturmaktadır. Araştırmada öğrencilerin dijital teknolojilere olan tutumlarını etkileyen faktörlerden cinsiyet, ailenin gelir düzeyi ve ebeveyn eğitim durumları ele alınmıştır. Aynı zamanda öğrencilerin kullandıkları dijital araçları hangi amaç doğrultusunda kullandıkları belirlenmiştir. Araştırma sonucunda ele alınan faktörlerden, ailenin gelir düzeyi faktörüne ve ebeveyn eğitim durumu faktörüne bağlı olarak öğrencilerin dijital teknolojiye yönelik tutumlarının değiştiği, cinsiyet faktörüne bağlı olarak herhangi bir değişimin olmadığı belirlenmiştir. Öğrencilerin müzik dinlemek ve indirmek, TV izlemek, internette video izlemek ve indirmek, sosyal ağlarda gezinmek gibi aktiviteleri çoğunlukla akıllı cep telefonu, tablet ve dizüstü bilgisayar ile yaptıkları görülmüştür. Aynı zamanda öğrenciler dijital teknolojilerin kullanım amacına göre hem hayatı kolaylaştırabileceğini hem de hayatlarını olumsuz yönde etkileyebileceğini görüş olarak belirtmişlerdir (Erten, 2019).

**Dijital Çağın Öğreneni (Ebru ARDIÇ, Adnan ALTUN, 2017):** Makalede araştırmacılar, teknoloji ile iç içe, dijital çağın öğreneni olan Z kuşağı öğreneninin özellikleri ile ilgili genellemeler ortaya çıkarmayı amaçlamışlardır. Araştırma sonucunda elde edilen bulgular ışığında, Z kuşağı öğreneninin teknolojiyi bireysel yaşam alanlarında kullanmalarının yanında, araştırma yapmak, bilgi edinmek, yaratıcılık, yenilik, kariyer ve teknoloji becerilerini geliştirmek amacıyla da kullandıkları ortaya çıkmıştır (Ardıç ve Altun 2017).

**Dijital Yerli Çocukların Dijital Teknolojiye Yönelik Görüşlerinin İncelenmesi (Emine CABI, 2015):** Makalede dijital yerli olarak tanımlanan öğrencilerin teknolojiye bakış açıları, sosyal ve eğitim hayatlarında yer alan teknoloji ile ilgili görüşlerinin ortaya çıkarılması amaçlanmıştır. Örneklemi 2013-2014 eğitim- öğretim yılında Ankara ilindeki dört farklı ortaöğretim kurumundan 103 öğrenci oluşturmaktadır. Makalede içerik analiz yöntemi kullanılmıştır. Analizler sonucunda Ticaret Meslek Lisesi öğrencilerinin %95’inin okullarında kullanılmak üzere tablet ve akıllı tahta talep ettikleri görülmüştür. Buna ek olarak dijital yerli çocuklar eğitimde teknolojinin daha çok kullanılması gerektiğini, Youtube ve Twitter erişiminin engellenmesini büyük bir sorun olarak gördüklerini belirtmişlerdir (Cabı, 2015b).

**Geleceğin İşgücü Olarak Z Kuşağının Dijital Teknolojiye Yönelik Tutumları (Benan YÜCEBALKAN, Barış AKSU):** Makalede araştırmacılar, Z kuşağının dijital teknolojiye yönelik tutumlarını, yeni işgücünden beklenen yetkinlikler çerçevesinde anket yöntemi ile çeşitli kategorilere göre belirlemeyi amaçlamışlardır. Araştırma sonucunda, Z kuşağının dijital teknolojiye karşı yetkinliklerinin, sosyal ağları aktif kullanma oranının, derslerde teknoloji kullanımına sıcak bakma oranının, teknolojiye yönelik ilgilerinin, teknolojinin bireysel gelişimlerine katkısının, teknolojiyi eğlence amaçlı ve bilinçli kullanma oranının yüksek olduğu sonucuna varılmıştır (Yücebalkan ve Aksu, 2018).

## **Materyaller ve Yöntemler**

### ***Materyaller***

Araştırmada, OECD’nin (Ekonomik İşbirliği ve Kalkınma Örgütü) 2018 yılında düzenlemiş olduğu PISA (Uluslararası Öğrenci Değerlendirme Programı) kapsamında gerçekleştirilen testteki anketlerden

biri olan bilişim ve iletişim teknolojilerine yatkınlık anketinden elde edilen veriler ile çalışılmıştır. Çalışılan veriler OECD PISA internet sitesinden (<https://www.oecd.org/pisa/>) elde edilmiştir.

### **Kullanılan Analizler**

Çalışmada Ki-Kare Testi, Güvenilirlik, Açıklayıcı Faktör ve Çoklu Doğrusal Regresyon Analizleri SPSS programı kullanılarak yapılmıştır.

**Ki-Kare Bağımsızlık Testi;** Parametrik olmayan istatistiksel yöntemlerden ki-kare bağımsızlık testinde ele alınan değişkenlerin, farklı ölçütlere ya da belirli bir amaca göre iki ya da daha çok yönlü çapraz tablo biçiminde sınıflandırılması halinde, değişkenlerin belirlenen özellikleri arasında ilişkinin olup olmadığı test edilir (Uzgören ve Uzgören, 2007).

**Güvenilirlik Analizi;** Bir ölçekte bulunması gereken en önemli özelliklerden birisi olan güvenilirlik, aynı ölçme aracına sahip ölçümlerde elde edilen ölçüm değerlerinin tutarlılığının göstergesidir (Kılıç, 2016).

**Açıklayıcı Faktör Analizi;** Açıklayıcı faktör analizinde, çok sayıda olan değişkenler arasındaki ilişkileri baz alarak az sayıda faktör bulmak ve bu faktörler ile teoriler üretmek amaçlanır (Büyüköztürk, 2002).

**Çoklu Doğrusal Regresyon Analizi;** Çoklu doğrusal regresyon analizinde, bir bağımlı değişken ile birden çok bağımsız değişkenin arasındaki işlevsel ilişki incelenmektedir. Analizin amacı, bağımlı değişkeni etkilediği düşünülen bağımsız değişkenlerin her birinin bağımlı değişkeni ne kadar etkilediklerini bulmak ve aralarındaki ilişkiyi incelemektir (Bekiroğlu ve ark., 2013).

### **Bulgular**

Çalışmadan elde edilen bulguların bir kısmı bu bölümde gösterilmiştir.

### **Ki-Kare Bağımsızlık Testleri**

Çalışmada ‘Cinsiyet’ ile ‘Okul Türleri’ değişkenlerine ki-kare bağımsızlık testleri uygulanmıştır.

Tablo 1. ‘Cinsiyet’ ile ‘Evde Taşınabilir Dizüstü Bilgisayara Ulaşılabilirlik’ İçin Ki-Kare Testi

		Evinizde kullanmanız için taşınabilir dizüstü bilgisayar var mı?						
		Var ve kullanıyorum	Var fakat kullanmıyorum	Yok	Total	Value	df	Asymptotic Significance (2-sided)
Cinsiyetiniz nedir?	Kadın	Count	1618	312	1394	3324		
		% within Cinsiyetiniz nedir?	48,7%	9,4%	41,9%	100,0%		
		% of Total	24,2%	4,7%	20,8%	49,7%		
	Erkek	Count	1598	427	1337	3362		
		% within Cinsiyetiniz nedir?	47,5%	12,7%	39,8%	100,0%		
		% of Total	23,9%	6,4%	20,0%	50,3%		
Total	Count	3216	739	2731	6686			
	% within Cinsiyetiniz nedir?	48,1%	11,1%	40,8%	100,0%			
	% of Total	48,1%	11,1%	40,8%	100,0%			

Pearson Chi-Square	18,994	2	,000
Likelihood Ratio	19,067	2	,000
Linear-by-Linear Association	,198	1	,656
N of Valid Cases	6686		

H<sub>0</sub>: Cinsiyet değişkeni ile evde taşınabilir dizüstü bilgisayara ulaşılabilirliği ele alan değişken arasında istatistiksel açıdan anlamlı bir ilişki yoktur.

Tablolara bakıldığında, kız öğrencilerin %58,1’inin, erkek öğrencilerin ise %60,2’sinin evlerinde taşınabilir dizüstü bilgisayara ulaşabildiği ve istatistiksel açıdan yaklaşıldığında da iki değişken arasında anlamlı bir ilişki olduğu görülmektedir. Sonuç olarak, erkek öğrencilerin evlerinde taşınabilir dizüstü bilgisayara ulaşma imkânı kız öğrencilere göre daha fazladır.

Tablo 2. ‘Cinsiyet’ ile ‘Evde Masaüstü Bilgisayara Ulaşılabilirlik’ İçin Ki-Kare Testi

		Evinizde kullanmanız için masaüstü bilgisayar var mı?							
		Var ve kullanıyorum	Var fakat kullanmıyorum	Yok	Total				
Cinsiyetiniz nedir?	Kadın	Count	1095	396	1830	3321			
		% within Cinsiyetiniz nedir?	33,0%	11,9%	55,1%	100,0%			
		% of Total	16,3%	5,9%	27,3%	49,5%			
Erkek	Erkek	Count	1585	380	1418	3383			
		% within Cinsiyetiniz nedir?	46,9%	11,2%	41,9%	100,0%			
		% of Total	23,6%	5,7%	21,2%	50,5%			
Total	Total	Count	2680	776	3248	6704			
		% within Cinsiyetiniz nedir?	40,0%	11,6%	48,4%	100,0%			
		% of Total	40,0%	11,6%	48,4%	100,0%			

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	141,619	2	,000
Likelihood Ratio	142,254	2	,000
Linear-by-Linear Association	139,978	1	,000
N of Valid Cases	6704		

H<sub>0</sub>: Cinsiyet değişkeni ile evde masaüstü bilgisayara ulaşılabilirliği ele alan değişken arasında istatistiksel açıdan anlamlı bir ilişki yoktur.

Tablolara bakıldığında, kız öğrencilerin %44,9’unun, erkek öğrencilerin %58,1’inin evlerinde masaüstü bilgisayara ulaşabildiği ve istatistiksel açıdan yaklaşıldığında da iki değişken arasında anlamlı bir ilişki olduğu görülmektedir. Sonuç olarak, erkek öğrencilerin evlerinde masaüstü bilgisayara ulaşma imkânı kız öğrencilere göre daha fazladır.

Tablo 3. ‘Cinsiyet’ ile ‘İnternet Büyük Bir Kaynaktır’ İçin Ki-Kare Testi

		İlgilendiğim bilgileri bulmak için internet büyük bir kaynaktır.								
		Kesinlikle katılmıyorum		Kesinlikle katılıyorum		Total				
Cinsiyetiniz nedir?	Kadın	Count	222	365	1709	871	3167			
		% within Cinsiyetiniz nedir?	7,0%	11,5%	54,0%	27,5%	100,0%			
		% of Total	3,5%	5,7%	26,8%	13,7%	49,7%			
Erkek	Erkek	Count	301	542	1408	893	3204			
		% within Cinsiyetiniz nedir?	9,4%	16,9%	45,8%	27,9%	100,0%			
		% of Total	4,7%	8,5%	23,0%	14,0%	50,3%			
Total	Total	Count	523	907	3177	1764	6371			
		% within Cinsiyetiniz nedir?	8,2%	14,2%	49,9%	27,7%	100,0%			
		% of Total	8,2%	14,2%	49,9%	27,7%	100,0%			

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	64,818	3	,000
Likelihood Ratio	65,102	3	,000
Linear-by-Linear Association	20,446	1	,000
N of Valid Cases	6371		

H<sub>0</sub>: Cinsiyet değişkeni ile ‘İnternet büyük bir kaynaktır’ değişkeni arasında istatistiksel açıdan anlamlı bir ilişki yoktur.

Tablolara bakıldığında, kız öğrencilerin %81,5’i interneti büyük bir kaynak olarak görürken, erkek öğrencilerin ise %73,7’sinin interneti büyük bir kaynak olarak gördüğü ve istatistiksel açıdan yaklaşıldığında da değişkenler arasında anlamlı bir ilişki olduğu görülmektedir. Sonuç olarak kız öğrenciler erkek öğrencilere göre daha fazla internetin büyük bir kaynak olduğunu düşünmektedirler.

Tablo 4. ‘Cinsiyet’ ile ‘Dijital Cihazları Kullanmayı Seviyorum’ İçin Ki-Kare Testi

		Dijital cihazları kullanmayı seviyorum.								
		Kesinlikle katılmıyorum	Katılmıyorum	Katılıyorum	Kesinlikle katılıyorum	Total				
Cinsiyetiniz nedir?	Kadın	Count	226	345	1765	834	3170			
		% within Cinsiyetiniz nedir?	7,1%	10,9%	55,7%	26,3%	100,0%			
		% of Total	3,6%	5,4%	27,8%	13,1%	49,9%			
Erkek	Erkek	Count	270	430	1485	988	3179			
		% within Cinsiyetiniz nedir?	8,7%	13,5%	46,7%	31,1%	100,0%			
		% of Total	4,3%	6,8%	23,4%	15,6%	50,1%			
Total	Total	Count	502	775	3250	1822	6349			
		% within Cinsiyetiniz nedir?	7,9%	12,2%	51,2%	28,7%	100,0%			
		% of Total	7,9%	12,2%	51,2%	28,7%	100,0%			

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	51,430	3	,000
Likelihood Ratio	51,502	3	,000
Linear-by-Linear Association	,208	1	,647
N of Valid Cases	6349		

H<sub>0</sub>: Cinsiyet değişkeni ile ‘Dijital cihazları kullanmayı seviyorum’ değişkeni arasında istatistiksel açıdan anlamlı bir ilişki yoktur.

Tablolara bakıldığında, kız öğrencilerin %82’si dijital cihazları kullanmayı seviyorken, erkek öğrencilerin ise %77,8’inin sevdiği ve istatistiksel açıdan yaklaşıldığında da değişkenler arasında anlamlı bir ilişki olduğu görülmektedir. Sonuç olarak, kız öğrenciler erkek öğrencilere göre dijital cihazları kullanmayı daha çok seviyorlardır.

Tablo 5. ‘Okul Türleri’ ile ‘Okulda Projeksiyon Cihazına Ulaşılabilirlik’ İçin Ki-Kare Testi

		Okulunuzda kullanmanız için projeksiyon var mı?				
		Var ve kullanıyorum	Var fakat kullanmıyorum	Yok	Total	
Okul türleri 2	Anadolu Lisesi	Count	1227	659	1021	2907
		% within Okul türleri 2	42,2%	22,7%	35,1%	100,0%
		% of Total	18,5%	10,0%	15,4%	43,9%
Mesleki ve Teknik Anadolu Lisesi		Count	961	425	673	2059
		% within Okul türleri 2	46,7%	20,6%	32,7%	100,0%
		% of Total	14,5%	6,4%	10,2%	31,1%
Çok Programlı Anadolu Lisesi		Count	94	54	110	258
		% within Okul türleri 2	36,4%	20,9%	42,6%	100,0%
		% of Total	1,4%	0,8%	1,7%	3,9%
Ortaokul		Count	4	6	8	18
		% within Okul türleri 2	22,2%	33,3%	44,4%	100,0%
		% of Total	0,1%	0,1%	0,1%	0,3%
Fen Lisesi		Count	115	63	44	222
		% within Okul türleri 2	51,8%	28,4%	19,8%	100,0%
		% of Total	1,7%	1,0%	0,7%	3,4%
Sosyal Bilimler Lisesi		Count	129	70	24	223
		% within Okul türleri 2	57,8%	31,4%	10,8%	100,0%
		% of Total	1,9%	1,1%	0,4%	3,4%
Anadolu İmam Hatip Lisesi		Count	373	203	324	900
		% within Okul türleri 2	41,4%	22,6%	36,0%	100,0%
		% of Total	5,6%	3,1%	4,9%	13,6%
Anadolu Spor Lisesi/Anadolu Güzeli Sanatlar Lisesi		Count	15	7	12	34
		% within Okul türleri 2	44,1%	20,6%	35,3%	100,0%
		% of Total	0,2%	0,1%	0,2%	0,5%
Total		Count	2918	1487	2216	6621
		% within Okul türleri 2	44,1%	22,5%	33,5%	100,0%
		% of Total	44,1%	22,5%	33,5%	100,0%

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	98,925	14	,000
Likelihood Ratio	111,307	14	,000
Linear-by-Linear Association	3,340	1	,068
N of Valid Cases	6621		

H<sub>0</sub>: Okul türleri değişkeni ile okulda projeksiyon cihazına ulaşılabilirliği ele alan değişken arasında istatistiksel açıdan anlamlı bir ilişki yoktur.

Tablolara bakıldığında, okullarında projeksiyon cihazına ulaşma imkânı en yüksek iki okul türü öğrencileri sırasıyla %89,2 oranıyla Sosyal Bilimler Lisesi öğrencileri ve %80,2 oranıyla Fen Lisesi öğrencileri olduğu, istatistiksel açıdan yaklaşıldığında da değişkenler arasında anlamlı bir ilişki olduğu görülmektedir.

Tablo 6. ‘Okul Türleri’ ile ‘Okulda Etkileşimli Tahtaya Ulaşılabilirlik’ İçin Ki-Kare Testi

		Okulunuzda kullanmanız için etkileşimli tahta var mı?				
		Var ve kullanıyorum	Var fakat kullanmıyorum	Yok	Total	
Okul türleri 2	Anadolu Lisesi	Count	2027	281	605	2913
		% within Okul türleri 2	69,6%	9,6%	20,8%	100,0%
		% of Total	30,5%	4,2%	9,1%	43,9%
Mesleki ve Teknik Anadolu Lisesi		Count	1234	264	565	2063
		% within Okul türleri 2	59,8%	12,8%	27,4%	100,0%
		% of Total	18,6%	4,0%	8,5%	31,1%
Çok Programlı Anadolu Lisesi		Count	152	28	82	262
		% within Okul türleri 2	58,0%	10,7%	31,3%	100,0%
		% of Total	2,3%	0,4%	1,2%	3,9%
Ortaokul		Count	3	3	12	18
		% within Okul türleri 2	16,7%	16,7%	66,7%	100,0%
		% of Total	0,0%	0,0%	0,2%	0,3%
Fen Lisesi		Count	191	18	13	222
		% within Okul türleri 2	86,0%	8,1%	5,9%	100,0%
		% of Total	2,9%	0,3%	0,2%	3,3%
Sosyal Bilimler Lisesi		Count	202	16	8	226
		% within Okul türleri 2	89,4%	7,1%	3,5%	100,0%
		% of Total	3,0%	0,2%	0,1%	3,4%
Anadolu İmam Hatip Lisesi		Count	511	93	299	903
		% within Okul türleri 2	56,6%	10,3%	33,1%	100,0%
		% of Total	7,7%	1,4%	4,5%	13,6%
Anadolu Spor Lisesi/Anadolu Güzeli Sanatlar Lisesi		Count	24	4	7	35
		% within Okul türleri 2	68,6%	11,4%	20,0%	100,0%
		% of Total	0,4%	0,1%	0,1%	0,5%
Total		Count	4344	707	1591	6642
		% within Okul türleri 2	65,4%	10,6%	24,0%	100,0%
		% of Total	65,4%	10,6%	24,0%	100,0%

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	231,552	14	,000
Likelihood Ratio	259,082	14	,000
Linear-by-Linear Association	8,097	1	,004
N of Valid Cases	6642		

H<sub>0</sub>: Okul türleri değişkeni ile okulda etkileşimli tahtaya ulaşılabilirliği ele alan değişken arasında istatistiksel açıdan anlamlı bir ilişki yoktur.



Tablolara bakıldığında, okullarında etkileşimli tahtaya ulaşma imkânı en yüksek iki okul türü öğrencileri sırasıyla %96,5 oranıyla Sosyal Bilimler Lisesi öğrencileri ve %94,1 oranıyla Fen Lisesi öğrencileri olduğu, istatistiksel açıdan yaklaşıldığında da değişkenler arasında anlamlı bir ilişki olduğu görülmektedir.

### Güvenilirlik Analizi

Çalışmada 20 değişkene güvenilirlik analizi uygulanmıştır.

Tablo 7. Güvenilirlik ve Değişken-Toplam İstatistikleri

	Değişken Silindiğinde Ölçek Ortalaması	Değişken Silindiğinde Ölçek Varyansı	Düzeltilmiş Korelasyon	Değişken Silindiğinde Cronbach's Alpha
G11	52,293	109,808	,510	,927
G22	52,075	109,908	,555	,926
G33	52,304	110,182	,552	,926
G44	52,313	109,584	,546	,926
G55	52,371	110,314	,490	,927
G77	52,531	109,516	,555	,926
G88	52,297	108,018	,668	,923
G99	52,072	108,984	,647	,924
G1010	52,179	108,372	,681	,923
G1111	52,180	108,119	,676	,923
G1212	52,481	108,063	,589	,925
G1313	52,355	108,613	,628	,924
G1414	52,172	108,755	,652	,924
G1515	52,238	108,280	,666	,924
G1616	52,376	108,408	,606	,925
G1717	52,282	108,415	,645	,924
G1818	52,347	108,820	,638	,924
G1919	52,370	108,780	,577	,925
G2020	52,264	108,905	,640	,924
G2121	52,356	109,251	,587	,925
Cronbach's Alpha	Değişken Sayısı			
,928	20			

Tablolara bakıldığında, örneklem büyüklüğünün ve soru sayısının yeterli büyüklükte olmasından kaynaklı Cronbach's Alpha değerinin bire yakın bir değer alması dolayısıyla ölçümlerin tutarlılığının mükemmel olduğu ve Cronbach's Alpha değerinin herhangi bir değişken silindiğinde önceki değerle aralarında analizi etkileyecek boyutta bir fark olmadığı görülmektedir.

### Açımlayıcı Faktör Analizi

Güvenilirlik analizinde ölçümlerin tutarlılığının mükemmel olması sonrası aynı 20 değişkene açımlayıcı faktör analizi uygulanmıştır.

Tablo 8. Bartlett Küresellik Testi ve Ortak Faktör Varyansları

		Ortak Faktör Varyansları		
		Başlangıç	Çıkarılmış	
		G11	1,000	,609
		G22	1,000	,612
		G33	1,000	,617
		G44	1,000	,601
		G55	1,000	,594
		G77	1,000	,520
		G88	1,000	,695
		G99	1,000	,693
		G1010	1,000	,752
		G1111	1,000	,739
		G1212	1,000	,699
		G1313	1,000	,692
		G1414	1,000	,623
		G1515	1,000	,708
		G1616	1,000	,698
		G1717	1,000	,629
		G1818	1,000	,698
		G1919	1,000	,603
		G2020	1,000	,742
		G2121	1,000	,693
Kaiser-Meyer-Olkin Örnekleme Yeterliliğinin Ölçüsü	,940			
Bartlett Küresellik Testi	Yaklaşık Ki-Kare Değeri	76218,900		
	Serbestlik Derecesi	190		
	Anlamlılık	,000		

H<sub>0</sub>: Tüm korelasyon katsayıları sıfırdır.

Tablolara bakıldığında, KMO yeterlilik ölçüsünün bire çok yakın bir değer alması faktör analizinin mükemmel biçimde kullanılabileceğini ayrıca Bartlett küresellik testinin istatistiksel açıdan anlamlı çıkması değişkenlerin aralarında ilişki olduğunu ve değişkenlerin faktörlenebilmesinde hiçbir sorun olmadığını, her bir değişkenin çıkarılmış ortak yüklerinin 0,50'den yüksek olması ise değişkenlerin her birinin genel varyansı açıklayabilirliğinin yeterli düzeyde olduğunu göstermektedir.

Tablo 9. Açıklanan Toplam Varyans

Bileşen	Açıklanan Toplam Varyans								
	Başlangıç Özdeğerleri			Faktörleştirme Sonrası Değerler			Rotasyon Sonrası Değerler		
	Toplam	Varyans Yüzdəsi	Kümülatif Yüzde	Toplam	Varyans Yüzdəsi	Kümülatif Yüzde	Toplam	Varyans Yüzdəsi	Kümülatif Yüzde
1	8,585	42,924	42,924	8,585	42,924	42,924	3,369	16,843	16,843
2	2,004	10,022	52,946	2,004	10,022	52,946	3,351	16,757	33,600
3	1,418	7,092	60,038	1,418	7,092	60,038	3,284	16,418	50,018
4	1,210	6,052	66,091	1,210	6,052	66,091	3,214	16,072	66,091
5	,692	3,462	69,552						
6	,641	3,203	72,756						
7	,548	2,741	75,497						
8	,534	2,669	78,167						
9	,510	2,550	80,716						
10	,463	2,313	83,029						
11	,453	2,265	85,294						
12	,421	2,103	87,397						
13	,365	1,827	89,224						
14	,362	1,810	91,034						
15	,354	1,768	92,802						
16	,324	1,620	94,422						
17	,306	1,530	95,952						
18	,298	1,489	97,442						
19	,277	1,386	98,827						
20	,235	1,173	100,000						

Tabloya bakıldığında, 4 faktörün etkili olduğu ve bu faktörlerden, birinci faktörün %42,92'lik, ikinci faktörün %10,02'lik, üçüncü faktörün %7,09'luk ve son olarak dördüncü faktörün %6,05'lik açıklanabilirliğe sahip olduğu ayrıca faktörlerin toplam varyansın %66,09'unu açıkladığı görülmektedir.

Tablo 10. Rotasyonlu Bileşen Matrisi

	Rotasyonlu Bileşen Matrisi			
	Bileşen			
	1	2	3	4
G2020	,795	,214	,194	,162
G2121	,784	,186	,163	,133
G1818	,751	,280	,191	,140
G1919	,713	,192	,175	,168
G1717	,680	,276	,226	,198
G1212	,216	,783	,182	,084
G1616	,269	,767	,177	,083
G1313	,263	,754	,192	,133
G1515	,223	,742	,282	,168
G1414	,185	,662	,247	,300
G1010	,200	,252	,776	,213
G1111	,232	,242	,768	,189
G88	,239	,207	,735	,234
G99	,127	,193	,723	,341
G77	,195	,184	,648	,169
G11	,103	,112	,184	,743
G55	,159	,063	,144	,737
G33	,157	,136	,197	,732
G44	,211	,121	,162	,718
G22	,077	,190	,252	,711

Döndürülmüş faktör yapısında faktörlerin arasında bir ilişki beklenmediği için ortogonal rotasyon çeşitlerinden varimax metodu kullanılmış olup, rotasyonlu bileşen matris tablosuna bakıldığında değişkenlerin birbirinden bağımsız 4 faktöre dağıldığı görülmektedir.

- 1.Faktör: Paylaşım Faktörü: Sosyal çevre ile dijital cihaz ve internet paylaşımları
2. Faktör: Özgüven Faktörü: Dijital cihazlar ve yazılım programları ile ilgili özgüven
3. Faktör: Tecrübe Faktörü: Dijital cihazlar ile ilgili edinilen tecrübe
4. Faktör: İlgi Faktörü: Dijital cihazlara ve internete duyulan ilgi

### Çoklu Doğrusal Regresyon Analizi

Faktör analizinde oluşan 4 faktöre regresyon analizi uygulanmıştır.

Tablo 11. Anova ve Katsayılar

ANOVA						Katsayılar								
Model	Kareler Toplamı	df	Kareler Ortalaması	F	Anlamlılık	Model	Standardize Olmayan Katsayılar		Standardize Katsayılar		Collinearity Statistics			
							B	Std. Hata	Beta	t	Anlamlılık	Tolerans	VIF	
1	Regresyon	2439,087	4	609,772	1938,602	.000	(Sabit)	,285	,035		8,177	,000		
	Artık	2165,622	6885	,315			özgüven	,036	,013	,031	2,734	,006	,539	1,856
							tecrübe	,166	,014	,136	11,755	,000	,509	1,966
							ilgi	,748	,012	,613	59,980	,000	,655	1,528
	Toplam	4604,709	6889				paylaşım	,031	,013	,026	2,372	,018	,563	1,777

$$H_0: \beta_0=\beta_1=\beta_2=\beta_3=\beta_4=0$$

Anova tablosuna bakıldığında istatistiksel açıdan modelin anlamlı olduğu görülmektedir.

Standartlaştırılmamış katsayılara bakıldığında bağımsız değişkenlerin her birinin bağımlı değişkeni pozitif yönde etkilediği ve her birinin istatistiksel açıdan anlamlı olduğu buna ek olarak dört değişkenin de VIF değeri 5'ten küçük olduğu için değişkenlerde çoklu doğrusal bağlantı sorunu olmadığı görülmektedir.

$$\text{Model: } \text{Sevgi} = 0,285 + 0,036 X \text{Özgüven} + 0,166 X \text{Tecrübe} + 0,748 X \text{İlgi} + 0,031 X \text{Paylaşım}$$

Tablo 12. Model Özeti

Model	R	R Kare	Düzeltilmiş R Kare	Tahmin Standart Hatası
1	,728	,530	,529	,5608

### Sonuç

Erkek öğrencilerin kız öğrencilere kıyasla dijital cihazlara ve internete daha kolay eriştikleri, dijital cihazları ve interneti kullanmayı ise kız öğrencilerin daha çok sevdikleri ortaya çıkmıştır. Kız öğrenciler daha çok kullanmayı isterken erişim oranlarının düşük olması ise toplumsal cinsiyet eşitsizliğinin göstergesidir.

Fen Lisesi ve Sosyal Bilimler Lisesi öğrencilerinin dijital cihaz ve internete erişim oranlarının yüksek olmasının, eğitimdeki başarılarının da yüksek olmasını etkileyen etmenlerden biri olduğu sonucuna ulaşılmıştır.

Öğrencilerin dijital cihazlara olan sevgilerini etkileyen faktörlerin de dijital cihazlara olan ilgi, dijital cihazlardan ve internetten edinilen tecrübe, özgüven ve sosyal çevreleri ile dijital cihazlar ve internet hakkındaki paylaşımları olduğu sonucuna varılmıştır.

## Kaynaklar

- Ardıç E., Altun A., 2017. “Dijital Çağın Öğreneni”, Uluslararası Sosyal Bilgilerde Yeni Yaklaşımlar Dergisi, 1, sayı 1, s.n. 12–30
- Bekiroğlu N., Konyalıoğlu R., v.d., 2013. “Çoklu Doğrusal Regresyon Sonuçlarının Jackknife Tekniği İle Tekrarlanabilirliğinin Değerlendirilmesi”, Marmara Medical Journal Dergisi, sayı 26, s.n. 63–67
- Büyüköztürk Ş., 2002. “Faktör Analizi: Temel Kavramlar ve Ölçek Geliştirmede Kullanımı”, Kuram ve Uygulamada Eğitim Yöntemi Dergisi, sayı 32, s.n. 470–483
- Cabı E., 2015a. “Dijital Teknolojiye Yönelik Tutum Ölçeği”, Kastamonu Eğitim Dergisi, 24, sayı 3, s.n. 1229–1244
- Cabı E., 2015b. “Dijital Yerli Çocukların Dijital Teknolojiye Yönelik Görüşlerinin İncelenmesi”, International Journal of Innovative Research in Education, 2, sayı 1, s.n. 10–15
- Dönmez G., 2019. “Lise Öğrencilerinin Bilgi Güvenliği Farkındalığı İle Dijital Okuryazarlığın Arasındaki İlişkinin İncelenmesi”
- Erten P., 2019. “Z Kuşağının Dijital Teknolojiye Yönelik Tutumları”, Gümüşhane Üniversitesi Sosyal Bilimler Enstitüsü Elektronik Dergisi, 10, sayı 1, s.n. 190–202
- Kılıç S., 2016. “Cronbach’ın Alfa Güvenirlik Katsayısı”, Journal of Mood Disorders, 6, sayı 1, s.n. 47–48
- Uzgören N., Uzgören E., 2007. “Dumlupınar Üniversitesi Lisans Öğrencilerinin Memnuniyetini Etkileyen Bireysel Özelliklerin İstatistiksel Analizi”, Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, sayı 17, s.n. 173–192
- Yücebalkan B., Aksu B., 2018. “Geleceğin İşgücü Olarak Z Kuşağının Dijital Teknolojiye Yönelik Tutumları”, Sosyal Bilimlerinde Güncel Akademik Çalışmalar, 1, sayı 1, s.n. 1

## Experimental Design and Use of SPSS in Agricultural Research

Meral DOĞAN<sup>1</sup>

<sup>1</sup>Harran University, Graduate School of Natural and Applied Sciences, Department of Horticulture  
Sanliurfa, Turkey

### Abstract

Experiment can be defined as a process or process applied to test the effect of various conditions under control on unknown characteristics of experimental units. Experimental design is the regulation of these controlled conditions to which experimental units will be exposed. Experimental design is used in many different areas. In scientific research in the agricultural field, it is very important to conclude and sustain the study in line with its objectives. Therefore, in the studies to be carried out, planning and design of the experiment, conducting the experiment and obtaining the data, data analysis and interpretation of the results are required. Otherwise, the scientific aspects of the research remain weak. Various statistical programs are used for data analysis. SPSS is one of the statistics programs widely used in agricultural research. Working on data is a complex and time-consuming process, but SPSS can easily process and use information with the help of some techniques. These techniques are used to analyze, transform, and generate the characteristic pattern between different data variables. In addition, the output can be obtained by graphical representation so that the user can easily understand the result. In this study, it is aimed to be a resource for future studies by using various agricultural researches, examining experimental designs in agricultural research and determining the possibilities of using SPSS.

**Key words:** Agricultural research, statistics, experiment design, SPSS

## **Drug Eluting Stents versus Bare Metal Stents in Patients with ST-Elevated Myocardial Infarction**

Dr.Etleva LLAGAMI<sup>1</sup>, Dr.Endri RAÇO<sup>1</sup>, Dr.Kleida HAXHI<sup>1</sup>, Dr.Oriana ZAÇAJ<sup>1</sup>

Polytechnic University of Tirana, Faculty of Mathematical Engineering and Physical Engineering,  
Albania

### **Abstract**

Drug eluting stents (DES) were developed to reduce the growth of scar tissue. In order to compare the benefits and harms of DES versus bare metal stents, BMS, a study was performed on 636 patients of the University Medical Center of Tirana “Mother Teresa”, who were diagnosed with ST-elevated myocardial infarction (STEMI). Statistical analysis showed no evidence of a difference between stent types regarding all-cause mortality and major cardiovascular in-hospital events. Follow-up procedures took place 3 months after discharge, using a standardized protocol that included outpatient visits, telephone contacts, and recordings of cardiac events. The primary end point was all-cause mortality. Cardiovascular deaths, re-infarctions, and repeated target vessel revascularizations (TVRs) were defined as major adverse cardiac events (MACEs). During long-term follow-up, any events of stroke, heart failure-related re-hospitalizations, and MACEs were noted. The risk of re-infarction was slightly lower in those treated with a DES compared to a BMS (6.3% compared to 5.1%). A similar trend was observed in cases of target vessel revascularization (5.5% compared to 4.3%). So, this current study suggests that using DES rather than BMS may reduce the need for future coronary procedures. However, there is no evidence to support claims, that it saves lives or reduces major adverse cardiovascular events.

**Key words:** DES, BMS, STEMI, Kaplan-Meier, log-rank test

### **Introduction**

Two types of stents are used when dealing with coronary problems: Bare Metal Stents (BMS) and Drug Eluting Stents (DES). Nowadays DES designs have rendered the rate of major cardiac events so low that physicians generally lean towards using DES to treat patients previously considered at higher risk. A definitive randomized study was done in Albania, where not only DES, but also BMS were implanted, so as to give more clear-cut answers on this topic.

### **Methods**

636 consecutive patients with STEMI\* (531 men and 105 women with a mean age of 63.82, range: 30-85 y.), who were admitted to the Cardiology Center of the “Mother Teresa” University Medical Center of Tirana between April 2019 and December 2019, took part in this prospective study. The study inclusion criteria were:

ST-segment elevation  $\geq 1$  mm in at least 2 contiguous electrocardiographic (ECG) leads or new left bundle branch block with increasing cardiac enzymes twice the upper limit of normal within 12 hours of symptom onset,

Treatment with primary stent implantation (PSI).

\*ST-elevated myocardial infarction (STEMI).

**Coronary angiography, primary angioplasty and stenting**

Initial medical therapy during STEMI consists of oxygen administration, antiplatelet therapy (aspirin, thienopyridines and glycoprotein IIb/IIIa inhibitors), anticoagulation (heparin or bivalirudin), anginal pain relief with nitrates and morphine, and beta-blockade. Medical therapy upon hospital discharge may include ACE inhibitors, angiotensin receptor blockers, aldosterone antagonists and HMG CoA reductase inhibitors.

All patients received chewable 300 mg aspirin and 600 mg clopidogrel before coronary angiography. The patients' angiographic data were obtained from the cardiac catheterization laboratory records. Emergency coronary angiography and angioplasty were performed by the percutaneous femoral approach. The contralateral artery was first injected, and after visualizing the coronary arteries, 2.5 mg of nitrate was injected into the infarct-related artery (IRA) to rule out possible coronary spasm.

After angioplasty, all patients were admitted to the intensive care unit (ICU), where 100 mg of aspirin and 75 mg of clopidogrel were continued in all patients. The use of glycoprotein IIb/IIIa inhibitors was left to the discretion of the operator. Concomitant medical treatment with statins, angiotensin-converting enzyme inhibitors, and  $\beta$ -blockers was given.

**Follow - up**

Follow-up procedures took place 3 months after discharge, using a standardized protocol that included outpatient visits, telephone contacts, and recordings of cardiac events. The primary end point was all-cause mortality. Cardiovascular deaths, re-infarctions, and repeated target vessel revascularizations (TVRs) were defined as major adverse cardiac events (MACEs). During long-term follow-up, any events of stroke, heart failure-related re-hospitalizations, and MACEs were noted.

**Statistical Analysis and Results**

The studied population was divided into 2 groups according to the type of stent implanted. Patients with a recent DES stent placement were defined as DESgroup (n =351), those with a recent BMS stent placement as BMSgroup (n =285).

To examine which stent is used more often in older patients with STEMI (with higher bleeding risks, warranting a minimized duration of dual antiplatelet therapy after percutaneous coronary intervention), the age variable was divided in two categories: patients younger than 65; and those older.

Parametric comparisons of numbers of patients with a recent BMS stent placement or DES in two age-groups were performed with a 2-tailed Student *t* test.

Table 1. Types of stent and age-group

		Types of stents and age-groups			
		Age < 65		Age > 65	
		(A)		(B)	
		Count	Column N %	Count	Column N %
Stent	DESgroup	189	55.8%	162	54.5%
	BMSgroup	150	44.2%	135	45.5%
<p><i>Results are based on two-sided tests. For each significant pair, the key of the category with the smaller column proportion appears in the category with the larger column proportion.</i></p> <p><i>Significance level for upper case letters (A, B, C): .01</i></p> <p><i>Significance level for lower case letters (a, b, c): .05<sup>a</sup></i></p> <p><i>a. Tests are adjusted for all pairwise comparisons within a row of each innermost subtable using the Benjamini-Hochberg correction.</i></p>					

Two types of stents had statistically the same rate of use in each age-groups (p-value=0.687).

To study the effect of type of stents in mortality and in major adverse cardiac events during the hospitalization, Pearson Chi-Square Tests were used (Table2, Cardiovascular death and MACE in hospitalization). In Table2, cardiac death and MACE variables were expressed as percentages of cases and 95% confidence interval of percentages.

$\chi^2= 1.316$ , p-value=0.251 in test of independence of cardiac death by stent type and  $\chi^2= 0.802$ , p-value=0.371 in test of independence of major adverse cardiac events by stent type proved there is no significant role of this factor in mortality or in MACE hospitalizations.

Table 2. Cardiovascular Death and MACE in Hospitalization

		Stent							
		BMS				DES			
Count		Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %	Count	Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %	
Cardiac death	12	3.4%	1.9%	5.7%	15	5.3%	3.1%	8.3%	
Stroke and re-hospitalization	30	9.14%	6.9%	13.1%	19	8.6%	7.3%	10.7%	

During follow-up procedures, any events of stroke, heart failure-related re-hospitalizations, and MACEs were noted, thus creating two new variables: cardiovascular death (follow-up) and MACE (follow-up). Once again, using Pearson Chi-Square Tests on these variables, Table 3(Cardiovascular death and MACE follow-up) showed us that the type of stent in major adverse cardiac events during long-term follow-up ( $\chi^2= 0.012$ , p-value=0.914) does not play a significant role.

Table 3. Cardiovascular Death and MACE Follow-up)

		Type of stents							
		BMS				DES			
		Count	Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %	Count	Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %
Cardiovascular death f.u	1.00	9	2.7%	1.3%	4.8%	3	1.1%	0.3%	3.0%
MACE f.u	1.00	27	7.8%	5.3%	10.9%	24	8.5%	5.7%	12.2%

The risk of re-infarction was slightly lower in those treated with a DES compared to a BMS ( $\chi^2=3.7925$ , p -value =0.054). The percentage of re-infarction was 5.1% of all cases of DES implantation, and it was 6.3 % in the cases of BMS implantation. The percentage of target vessel revascularization (TVR) was statistically lower in DESgroup (4.3 %), compared to 5.5% of TVR in BMSgroup ( $\chi^2=3.76$ , p -value =0.059).

How do these two types of stents affect the probability of cardiovascular problems?

To answer this question, it is necessary to define "lifetime" (a variable that defines the period of time, during which a STEMI patient has not had any cardiovascular problems after stent implantation). Using Kaplan Meier curves of Log Survival Function (Figure 1) and log-rank tests (Table 4), we noticed that there was not any significant difference between the two graphs. So, we concluded that the type of stent does not play a notable role in the health conditions of cardiac patients.



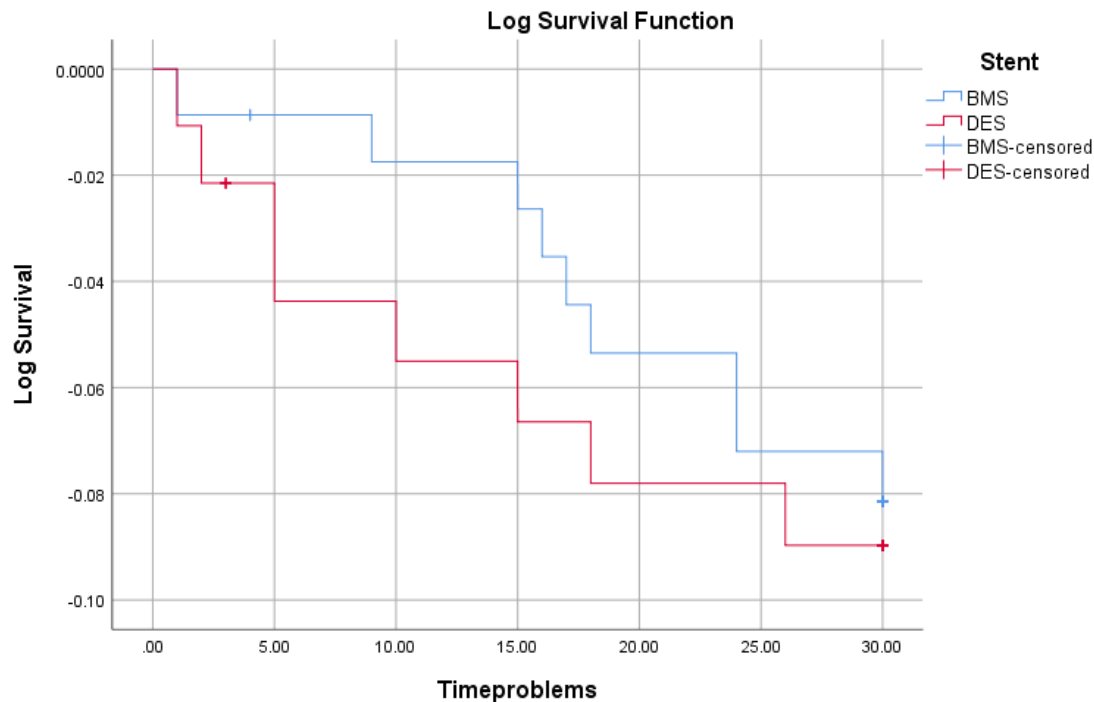


Figure 1. Kaplan-Meier Graph of Log Survival Function by Type of Stent

Table 4. Three Tests for Difference in Survival Data

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	.169	1	.681
Breslow (Generalized Wilcoxon)	.228	1	.633
Tarone-Ware	.198	1	.657
<i>Test of equality of survival distributions for the different levels of Stent DES.</i>			

## Conclusion

These findings indicate that there is no evidence to support the claim, that Drug Eluting Stents (DES) save lives or reduce major adverse cardiovascular events in the hospital or in follow-up outcomes, when treating patients with STEMI who were undergoing primary PCI.

The risk of re-infarction was slightly lower in those treated with a DES compared to a BMS. A similar trend was observed in cases of target vessel revascularization. Thus this study suggests that using DES rather than BMS may slightly reduce the need for future coronary procedures.

## References

- Which Stent for Your Older Patient with Coronary Artery Disease?  
 Pascal Meier, MD reviewing Varenne O et al. Lancet 2017 Nov 1 Gerber RT and Gershlick AH. Lancet 2017, NEJM Journal Watch Cardiology Nov 1 December 12, 2017.
- Ayhan E, Uyarel H, Ergelen M, et al. Primary angioplasty in women with ST-elevation myocardial infarction: in-hospital and long-term clinical results. Turk Kardiyol Dern Ars. 2011; 39(2): 114-121.
- Kastrati A, Dibra A, Spaulding C, Laarman GJ, Menichelli M, Valgimigli M, Di Lorenzo E, Kaiser C, Tierala I, Mehilli J, Seyfarth M, Varenne O, Dirksen MT, Percoco G, Varricchio A, Pittl U, Syvanne M, Suttorp MJ, Violini R, Schomig A. Meta-analysis of randomized trials on drug-eluting stents vs. bare metal stents in patients with acute myocardial infarction. Eur Heart J 2007; 28 (22):2706–2713.

- Ibanez B, James S, Agewall S, Antunes MJ, Bucciarelli-Ducci Ch, Bueno H, Caforio ALP, Crea F, Goudevenos JA, Halvorsen S, Hindricks G, Kastrati A, Lenzen MJ, Prescott E, Roffi M, Valgimigli M, Varenhorst C, Vranckx P, Widimsky P. The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation
- A Practical Guide To Understanding Kaplan-Meier Curves Jason T. Rich, Md, J. Gail Neely, Md, Facs, Randal C. Paniello, Md, Facs, Courtney C. J. Voelker, Md, D. Phil (Oxon), Brian Nussenbaum, Md, Facs, And Eric W. Wang, Md. Published in final edited form as: Otolaryngol Head Neck Surg. 2010 Sep; 143(3): 331–336.

## Uncertainty of the Claims Development Result in the Chain Ladder Method for Compulsory Motor Insurance

Kleida HAXHI<sup>1</sup>, Endri RAÇO<sup>1</sup>, Etleva LLAGAMI<sup>1</sup>, Oriana ZAÇAJ<sup>1</sup>

<sup>1</sup>Polytechnic University of Tirana, Faculty of Mathematical Engineering and Physical Engineering,  
Albania

### Abstract

The analysis of the uncertainty of the observable claim development result is important for the claims calculation reserve. The distribution free chain ladder method estimates the total ultimate claims amounts at time  $n$  after updating the information at time  $n+1$ . The observable claims development at time  $n$  for the accounting year  $(n, n+1)$  is the difference between the two successive best estimate prediction for the ultimate claims. In this paper we'll analyze the uncertainty of the observable claims development for the compulsory motor insurance.

**Key words:** Claim development, Chain Ladder Method, analysis of the uncertainty

## Etlik Piliçlerde Bazı Ağırlık ve Kan Parametreleri Arasındaki İlişkinin Kanonik Korelasyon Analizi ile İncelenmesi

Emine BERBEROĞLU<sup>1</sup>, Eylem YILDIZ KURNAZ<sup>2</sup>

<sup>1</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Taşçıiftlik, Tokat, Türkiye

<sup>2</sup>Bolu İl Tarım ve Orman Müdürlüğü, Mudurnu İlçe Müdürlüğü, Mudurnu, Bolu, Türkiye

### Özet

Bu çalışmada etlik piliçlerde bazı ağırlık ve kan parametreleri arasındaki ilişkiler kanonik korelasyon analizi ile incelenmiştir. Bu amaçla Gaziosmanpaşa Üniversitesi Tarımsal Uygulama ve Araştırma Merkezinde sıcaklık stresi üzerine yapılan bir çalışmada kullanılan Ross 308 ırkına ait etlik piliçin bazı ağırlık özellikleri ile X değişken kümesi canlı ağırlık, kalp ağırlığı, karaciğer ağırlığı ve abdominal yağ ağırlığı; kan parametreleri ile Y değişken kümesi glikoz düzeyi, kolesterol düzeyi ve hematokrit değerleri olmak üzere iki kümeye ayrılmıştır. Yapılan çalışmada, ele alınan özellikler arasında hesaplanan en büyük Pearson korelasyon katsayısı 0.469 olarak karaciğer ve kalp arasında hesaplanmıştır. İlk kanonik değişken çiftleri arasında hesaplanan en büyük kanonik korelasyon katsayısı ise birinci kanonik değişken çifti arasında 0.600 olarak hesaplanmıştır ( $P < 0.01$ ). Araştırma bulgularına göre karaciğer ağırlığı kanonik ağırlıklara (0.821) ve kanonik yüklere (0.804) en büyük katkıyı karaciğer ağırlığı değeri yaptığı için kan parametrelerini tahmin etmede bir kriter olabileceği sonucuna ulaşılmıştır.

**Anahtar kelimeler:** Değişken kümesi, etlik piliç, kanonik korelasyon, kanonik değişken, kan parametreleri

## Investigation of the Relationship Between Some Weight and Blood Parameters in Broiler Chicken by Canonical Correlation Analysis

### Abstract

In this study, the relationships between some weight and blood parameters in broiler chickens were examined by canonical correlation analysis. For this purpose, Ross 308 data obtained from a study on temperature stress in Gaziosmanpaşa University Agricultural Research Center were used. Data were divided into two clusters: some weight characteristics of broiler chickens and X variable set (live weight, heart weight, liver weight and abdominal fat weight) blood parameters and Y variable set (glucose level, cholesterol level and hematocrit) values. In the study, the largest Pearson correlation coefficient calculated between the properties was calculated as 0.469 between liver and heart. The largest canonical correlation coefficient calculated between the first pair of canonical variables was found to be 0.600 between the first pair of canonical variables ( $P < 0.01$ ). According to the findings of the study, canonical weights (0.821) and canonical loads (0.804) made the highest contribution to liver weight. Therefore, it could be a criterion for estimating blood parameters.

**Key words:** Blood parameters, broiler, canonical correlation, canonical variable, variable set

## Türkiye'de Süt Sığırlarından Üretilen Süt Miktarının Bulanık K-Ortalamaları Analiziyle Değerlendirilmesi

Yalçın TAHTALI<sup>1</sup>, Soner ÇANKAYA<sup>2</sup>, Lütfi BAYYURT<sup>1</sup>, Samet Hasan ABACI<sup>3</sup>

<sup>1</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Tokat, Türkiye

<sup>2</sup>Ondokuz Mayıs Üniversitesi, Yaşar Doğu Spor Bilimleri Fakültesi, Samsun, Türkiye

<sup>3</sup>Ondokuz Mayıs Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Samsun, Türkiye

### Özet

Bu çalışmanın amacı Türkiye'nin 12 farklı TR bölgesinde süt sığırcılığında üretilen süt miktarının Bulanık K-ortalama kümeleme algoritmalarını kullanarak kümeleme yapmaktır. Çalışmada kullanılan veriler 2004-2019 yılları arasında Türkiye İstatistik kurumundan alınmıştır. Analiz sonucunda oluşan 4 küme sırasıyla; Güneydoğu Anadolu-TRC, Kuzeydoğu Anadolu-TRA, Orta Anadolu-TR7 ve Batı Anadolu-TR5 1. kümede; Ege-TR3, Doğu Karadeniz-TR9, Akdeniz-TR6, Ortadoğu Anadolu-TRB ve Batı Karadeniz-TR8 2. kümede; İstanbul-TR1 3. kümede; Batı Marmara-TR2 ve Doğu Marmara-TR4 ise 4. kümede oluşmaktadır.

**Anahtar kelimeler:** Süt sığırını, süt miktarı, kümeleme analizi, Bulanık K- ortalama

## Hisse Senedi Fiyat Hareketlerini Prophet Zaman Serisi Modeli İle Öngörüleme

Zeynep ÖZTÜRK<sup>1</sup>

<sup>1</sup>Artvin Çoruh Üniversitesi, Hopa İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Artvin, Türkiye

### Özet

Bir zaman serisi modeli kurulumunda tatiller, insandan kaynaklanan kriz, savaş gibi şok etki tarihleri, eksik veriler, trenddeki değişiklikler ve aykırı değerler gibi etkenler en uygun modeli bulma ve tahmin yapma sürecini zorlaştırmaktadır. Türkiye’de de her geçen gün artan volatilité ve spekülâtif olaylar neticesinde hisse senedi fiyat hareketlerini tahmin etme ve öngörüleme yapmak oldukça güçtür. Bu nedenle, bu çalışmada, tatilleri ve istenilen tarihleri önsel bilgi olarak ele alan ve genelleştirilmiş eklemeli bir model olan Prophet zaman serisi modeli metodoloji ve sonuçları ile sunulacaktır. R analitik platformunu kullanarak 21.09.2018-21.09.2020 yılları arasında günlük alınan verilerle hisse senedi fiyat hareketlerinin (XU100) gelecek tahmini yapılacaktır. Farklı tahmin modelleri ile karşılaştırmak için istatistiksel performans ölçüsü kullanılacaktır.

**Anahtar kelimeler:** Hisse senedi fiyatları, öngörüleme, Prophet, zaman serisi modeli, R programı, zaman serisi modelleri

## **The Impact of the Insurance Market to the Financial Stability in Albania**

Oriana ZAÇAJ<sup>1</sup>, Kleida HAXHI<sup>1</sup>, Endri RAÇO<sup>1</sup>, Etleva LLAGAMI<sup>1</sup>

<sup>1</sup>Polytechnic University of Tirana, Faculty of Mathematical Engineering and Physical Engineering,  
Albania

### **Abstract**

Since its liberalization in 1999, the insurance market has grown rapidly. Initially, the growth related to the Motor TPL market segment. Later on, although not at the same pace, the other market segments grew considerably. In 2019 the insurance market growth (GWP) was approximately 4.1% compared to the year 2018. This continuous growth is reflected in both life and nonlife insurance business. The consumer has started to feel attracted to other insurance like property or life insurance, which is understandable in the view of the economic growth of Albania. But how does the insurance market growth affect the Albanian financial market? In this presentation we'll try to present the data reflecting this impact.

**Key words:** Insurance market, GDP, financial market, Albania

## Independent and Dependent Proportion Comparisons

Özgür KOŞKAN<sup>1</sup>, Asena KORKMAZ<sup>1</sup>, Duygu KAŞIKÇI<sup>1</sup>

<sup>1</sup>Department of Animal Science, Faculty of Agriculture, Isparta University of Applied Sciences,  
Isparta, Turkey

### Abstract

Z test is generally used to compare two independent proportions. Two dependent proportions are the comparison of the proportions that are reobtained from the same experimental unit at different times. Proportions should be obtained from the same experimental units. In addition to the z test used for comparing two dependent proportions, McNemar z test is also used for comparing two dependent proportions. In the method of this study, z test was used to compare two independent proportions. Twin births by districts were used as the independent variable. The twinning rate between 2013 and 2017 was used as a dependent variable. For the comparison of two dependent proportions, McNemar z test was used. The twinning rate of the same sheep in different years can be given as an example of two dependent proportions. The result of this study showed that McNemar Z test statistics give very small p values (significance), even though there is a little difference between proportions. McNemar test statistic values were greater than normal because McNemar test evaluates the goodness of fit of the year. In McNemar z test while proportions in different years are calculated, lots of data are lost. Therefore instead of existing methods, new methods should be developed. These new methods should prevent very small dependent proportions to be statistically significant. In the results, we found McNemar z value of 164.204, which is unrealistic and this shows that the development of new methods is necessary.

**Key words:** Compare two dependent proportions, compare two independent proportions, McNemar Z test, Pirit sheep, Z test.



## Canonical Correlation and a Livestock Application

Özgür KOŞKAN<sup>1</sup>, M. Burcu KOLCUOĞLU<sup>1</sup>

<sup>1</sup>Department of Animal Science, Faculty of Agriculture, Isparta University of Applied Sciences,  
Isparta, Turkey

### Abstract

The purpose of this study was to apply and interpret canonical correlation analysis, one of the multivariate statistical methods, in the field of animal husbandry. The variables used in the study were slaughter weight, hot carcass weight, cold carcass weight, bone weight, total meat yield, minced meat + batch yield, valuable parts yield and processing loss, respectively. For each variable, considering the values in the literature, a multivariate normal distribution data set resembling the original was created by simulation method. In the study, the left variable set consisted of slaughter weight, hot carcass weight, cold carcass weight, bone weight, whereas right variable set consisted of total meat yield, minced meat + batch yield, valuable parts yield and processing loss. Canonical correlations were found as 0.905, 0.330, 0.272, 0.016. Of the canonical correlations, 0.905 which is the highest of the canonical correlation values was significant ( $p < 0.01$ ).

**Key words:** Multivariate, canonical correlation, analysis

**Altın Fiyatlarındaki Değişimin Farklı Regresyon Yöntemleri ile Karşılaştırılması:  
Covid-19 Pandemisi Etkisi**

Adem TÜZEMEN<sup>1</sup>, Selim ÇAM<sup>2</sup>

<sup>1</sup>Tokat Gaziosmanpaşa Üniversitesi, İİBF İşletme Bölümü, Tokat, Türkiye

<sup>2</sup>Sivas Cumhuriyet Üniversitesi Hastanesi İstatistik Birimi, Sivas, Türkiye

**Özet**

Yakın dönemde başlayıp hala devam etmekte olan Covid-19 pandemisi küresel çapta tüm finansal olayları etkilemiştir. Ülkelerin sahip oldukları para birimleri (dünya çapında dolaşım hakimiyeti olan), enerji kaynakları ve değerli madenlerde de beklenmeyen değişimler yaşanmıştır. Covid-19 pandemisinin Türkiye'deki altın fiyatları üzerinde yaptığı değişimi modellemek adına yapılan bu çalışmada 01.2014-09.2020 dönemindeki aylık veriler kullanılmıştır. Analiz aşamasında parçalı regresyon ve lineer olmayan regresyon yöntemleri kullanılmıştır. Böylelikle her iki yöntemin altın fiyatı tahminlemedeki sonuçları ve model oluşturmadaki başarısı incelenmiştir. Araştırmanın sonucunda, parçalı regresyon modeli ile lineer olmayan regresyon modelleri arasında farkın olduğu görülmüştür. Karşılaştırmada temel alınan  $R^2$ , MSE, F istatistikleri kullanılmıştır. Parçalı regresyon modeli sonuçları incelendiğinde  $R^2$  ve F istatistiklerinin lineer olmayan regresyon modeline göre daha yüksek, buna karşılık MSE değerinin ise daha düşük hesaplanmıştır.

**Anahtar kelimeler:** Parçalı regresyon, lineer olmayan regresyon, altın, düğüm

**Comparison of the Change in Gold Prices with Different Regression Methods: Covid-19  
Pandemic Effect**

**Abstract**

The Covid-19 pandemic, which started recently and is still ongoing, has affected all financial events globally. Unexpected changes were also experienced in the currencies of the countries (which are dominated by worldwide circulation), energy resources and precious metals. Covid-19 of the pandemic, to model the changes made on the price of gold in Turkey as requested in this study, monthly data was used in the period 01.2014-09.2020. Piecewise regression and nonlinear regression methods were used in the analysis phase. Thus, the results of both methods in forecasting gold prices and their success in modeling were analyzed. As a result of the study, it was seen that there is a difference between piecewise regression models and nonlinear regression models. Statistics based on  $R^2$ , MSE, F were used for comparison. When the piecewise regression model results were examined, the  $R^2$  and F statistics were calculated higher than the nonlinear regression model, whereas the MSE value was found to be lower.

**Key words:** Piecewise regression, non-linear regression, gold, knot

## Laboratuvar Hata Miktarlarının İstatistiksel Kalite Kontrol ve Delphi Yöntemi ile Azaltılması

Adem TÜZEMEN<sup>1</sup>, Çağdaş YILDIZ<sup>1</sup>

<sup>1</sup>Tokat Gaziosmanpaşa Üniversitesi, İİBF İşletme Bölümü, Tokat, Türkiye

### Özet

Sağlık hizmetlerindeki hata miktarı yaklaşımının, genel kalite yaklaşımındakinden daha hassas olduğu açıktır. Laboratuvar hataları sınıflandırma sistemine 2020 yılında yeni bir hata maddesi eklenmiştir. bu maddenin, hastane bilgi sistemine eklenmesinden hemen sonraki hata miktarları ile iyileştirme sonrasındaki hata miktarındaki değişimi incelenmiştir. Her ne kadar laboratuvar numunelerini hastane bilgi sistemleri ile takip etmek mümkünse de yeni bir değişkenin mevcut sisteme entegre edilmesi ayrı bir dikkat gerektirmektedir. Çalışmada kullanılan veriler 01.01.2020-20.09.2020 tarihleri arasındaki dönemi kapsamaktadır. Yeni hata maddesinin haftalık miktarları ile yapılan çalışmada, hatanın azaltılmasına yönelik direnç noktaları bulunmuştur. Bireysel yargıların ve davranışların etkilediği düşünülen hata varlığına yönelik Delphi yöntemi kullanılarak odak grup görüşmeleri yapılmıştır. Üç aşamalı Delphi çalışmasının, ilk aşamasında olası sebepler odak gruba sorulmuştur. İkinci aşamada yüz yüze yapılan görüşmelerde olası direnç sebepleri daraltılmıştır. Üçüncü aşamada ise önceki aşamalarda yer alan bilgiler doğrultusunda, hata miktarındaki azalmaya direnç oluşturan hususlar netleştirilmiştir. Delphi analizi sonucunda ortaya çıkan sebeplere yönelik iyileştirme çalışması yapılmıştır. Çalışma süreci boyunca veriler alınarak istatistiksel kalite diyagramları ile süreç incelenmiştir. İyileştirme sonrası hata miktarlarındaki düşüşün istatistiksel olarak anlamlı olduğu hesaplanmıştır. Sonuç olarak, istatistiksel kalite kontrol yaklaşımı ve Delphi yönteminin bir arada kullanılması yardımıyla, laboratuvar hataları sınıflandırma sistemine yeni dahil edilen bir maddenin hata miktarının düşürülmesi sağlanmıştır.

**Anahtar kelimeler:** İstatistiksel kalite kontrol, Delphi yöntemi, laboratuvar hataları sınıflandırma sistemi

### Reduction of Laboratory Error Amounts with Statistical Quality Control and Delphi

#### Method

#### Abstract

It is clear that the amount of error approach in health care is more sensitive than the general quality approach. A new error item was added to the laboratory errors classification system in 2020. The change in the amount of error immediately after this item was added to the hospital information system and the amount of error after improvement was examined. Although it is possible to follow laboratory samples with hospital information systems, integrating a new variable into the existing system requires special attention. The data used in the study covers the period between 01.01.2020-20.09.2020. In the study with the weekly amounts of the new error item, resistance points were found to reduce the error. Focus group interviews were conducted using the Delphi method for the presence of errors that were thought to be affected by individual judgments and behaviors. In the first stage of the three-stage Delphi study, possible reasons were asked to the focus group. In the second phase, possible reasons for resistance were narrowed in face-to-face interviews. In the third stage, in line with the information in the previous stages,

the issues that resisted the reduction in the amount of error were clarified. Improvement work has been carried out for the reasons revealed as a result of the Delphi analysis. Throughout the study process, the data were collected and the process was examined with statistical quality diagrams. The decrease in error amount after improvement was calculated to be statistically significant. As a result, using the statistical quality control approach and Delphi method together, it was ensured that the amount of error of a new item included in the laboratory error classification system was reduced.

**Key words:** Statistical quality control, Delphi method, laboratory defect classification system

## **Covid 19 Pandemisi Sürecinde BİST’te İşlem Gören Perakende Ticaret Şirketlerinin Finansal Performanslar Bağlamında Critic ve Copras Yöntemleriyle Karşılaştırılması**

Ayşe Elif YAZGAN<sup>1</sup>, Özlem KARADAĞ AK<sup>1</sup>

<sup>1</sup>Necmettin Erbakan Üniversitesi, Uygulamalı Bilimler Fakültesi, Bankacılık Bölümü, Konya, Türkiye

### **Özet**

Çalışmanın amacı Borsa İstanbul (BİST) Ticaret Endeksi (XTCRT)’nde yer alan en tanınmış altı perakende firmasının 2018-2020 dönemleri arasındaki finansal performanslarının pandemi sürecinde ne derece etkilendiklerini incelemektir. Pandemi süreci 2019 yılı sonunda başladığından analizde firmaların 2. dönem verileri esas alınmıştır. Veriler, internet ortamında firmaların web sitelerinden ve Kamuyu Aydınlatma Platformu (KAP) web sayfasından elde edilmiştir. Araştırmada kullanılan dokuz adet finansal oran (Cari Oran, Likidite Oranı, Aktif Devir Hızı Oranı, Stok Devir Hızı Oranı, Aktif Kârlılık Oranı, Öz sermaye Kârlılık Oranı, Net Kâr Marjı, Kaldıraç Oranı, Borçlanma Oranı) mevcut alan yazından faydalanılarak seçilmiştir. Çalışmanın amacı doğrultusunda bu firmaların finansal performansları Critic ve Copras yöntemleri ile değerlendirilmiştir. Çalışmada öncelikle 1995 yılında Diakoulaki ve arkadaşları tarafından geliştirilen, çok kriterli karar verme problemlerinde kriterlerin önem ağırlıklarının objektif olarak belirlenmesi için Critic yöntemi kullanılmıştır. Daha sonra Zavadskas ve Kaklauskas tarafından 1996 yılında geliştirilmiş, kriter değerlerinin belirlenmesinde fayda kriterinin üst düzeye çıkartılması ve maliyet kriterinin en düşük seviyeye indirilerek değerlendirilmesi için kullanılan Copras yönteminden faydalanılmıştır. Yapılan değerlendirme sonucunda Critic yöntemi ile elde edilen kriter ağırlıklarına göre en önemli 3 kriter 2018 yılında sırasıyla Kaldıraç Oranı, Aktif Devir Hızı Oranı, Stok Devir Hızı Oranı; 2019 yılında Aktif Kârlılık Oranı, Borçlanma Oranı, Aktif Devir Hızı Oranı; 2020 yılında ise Borçlanma Oranı, Likidite Oranı ve Kaldıraç Oranı olarak belirlenmiştir. Copras yöntemine göre belirlenen finansal performansı en yüksek olan 3 firma 2018 yılında sırasıyla Carrefoursa, Bim ve Şok; 2019 yılında Migros, Şok ve Adese; 2020 yılında ise Carrefoursa, Migros ve Adese olarak tespit edilmiştir.

Sonuç olarak özellikle birçok ülkede belirli aralıklarla getirilen sokağa çıkma yasakları ile karantina süreci gibi alınan birtakım önlemler, tüketici davranışlarında değişikliğe neden olmuştur. Bu durum ise birçok sektörün yeniden şekillenmesine neden olmuştur. Bu sektörlerin başında gelen perakende ticaret sektöründe yer alan birçok firma ise pandemi sürecinde müşterilerine sunmuş oldukları online alışveriş imkânıyla durumu lehlerine çevirmişlerdir.

**Anahtar kelimeler:** BIST, Critic, Copras, Covid 19 pandemisi, perakende sektörü

### **Comparison of Retail Trade Companies with Critic and Copras Methods in Terms of Financial Performances in the Period of Covid-19 Pandemic**

#### **Abstract**

The purpose of the study is to examine the extent to which the financial performances of the six most well-known retail companies in the Borsa Istanbul (BIST) Trade Index (XTCRT) between 2018-2020 were affected during the pandemic process. Since the pandemic process started at the end of 2019, the

2nd period data of the companies were taken as basis in the analysis. The data were obtained from the websites of the companies and the Public Disclosure Platform (KAP) website in the internet environment. The nine financial ratios used in the study (Current Ratio, Liquidity Ratio, Asset Turnover Ratio, Stock Turnover Ratio, Asset Profitability Ratio, Equity Profitability Ratio, Net Profit Margin, Leverage Ratio, Borrowing Ratio) were selected using the current literature. In line with the purpose of the study, the financial performances of these companies were evaluated using Critic and Copras methods. In the study, the Critic method, which was developed by Diakoulaki et al. In 1995, was used to objectively determine the importance weights of the criteria in multi-criteria decision making problems. Later, the Copras method, which was developed by Zavadskas and Kaklauskas in 1996, was used to increase the utility criterion and to evaluate the cost criterion by minimizing it. According to the criteria weights obtained by the Critic method as a result of the evaluation, the most important 3 criteria are Leverage Ratio, Asset Turnover Rate, Stock Turnover Rate in 2018, respectively; Asset Profitability Rate, Borrowing Rate, Asset Turnover Rate in 2019; In 2020, the Borrowing Ratio, Liquidity Ratio and Leverage Ratio were determined. The 3 companies with the highest financial performance determined according to the Copras method were respectively Carrefoursa, Bim and Şok in 2018; Migros, Şok and Adese in 2019; In 2020, it was identified as Carrefoursa, Migros and Adese.

As a result, certain measures, such as curfews imposed at certain intervals in many countries and the quarantine process, have caused changes in consumer behavior. This situation has reshaped many sectors. Many companies in the retail trade sector, which is one of these sectors, turned the situation into their favor with the online shopping opportunity they offered to their customers during the pandemic process.

**Key words:** BIST, Critic, Copras, Covid 19 pandemic, retail industry

## COVID-19 Hastalığı için Ülkeler Bazında Virüs Bulaşma Katsayısı ( $R_0$ ) ile Nüfus Yoğunluğu ve Ortalama Hanehalkı Büyüklüğü Arasındaki İlişki

E. Burcu MAMAK EKİNCİ<sup>1</sup>

<sup>1</sup>Bilgisayar Mühendisliği Bölümü, Başkent Üniversitesi, Ankara, Türkiye

### Özet

2020 yılının ilk günlerinde Çin’de görülen ve mart ayından itibaren tüm Dünya’yı etkisine alan salgına neden olan yeni tip koronavirüs, COVID-19 hastalığı olarak adlandırılmıştır. Son verilere göre COVID-19 salgını dünyada 32.3 milyon insana bulaşmış ve yaklaşık 990 bin insanın ölümüne sebep olmuştur. Yapılan çalışmalarda COVID-19 virüsünün damlacık yoluyla ve doğrudan temas ile geçebildiği bilinmektedir. Hastalığın bulaşma hızının önemli bir göstergesi ise virüs bulaştırma katsayısı  $R_0$ ’dır. Bu çalışmada, virüs yayılımına etki eden önemli faktörlerden ikisi olan nüfus yoğunluğu ve ortalama hane halkı büyüklüğü göz önüne alınmıştır. Ülkeler bazında virüs bulaşma katsayısı ( $R_0$ ) ile nüfus yoğunluğu ve ortalama hanehalkı büyüklüğü arasındaki ilişki korelasyon ve regresyon analizi ile değerlendirilmiştir. Değişkenler arasındaki var olan ilişkinin yapısı ve miktarı açıklanmaya çalışılmıştır.

**Anahtar kelimeler:** Covid-19, nüfus yoğunluğu,  $R_0$  değeri, korelasyon, regresyon analizi

## Bireylerin İnternet Kullanım Alışkanlıklarının Genelleştirilmiş Doğrusal Modeller ve Sınıflandırma Analizi ile İncelenmesi

Burcu DURMUŞ<sup>1</sup>, Öznur İŞÇİ GÜNERİ<sup>1</sup>

<sup>1</sup>Muğla Sıtkı Koçman Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Kötekli Kampüsü, Muğla, Türkiye

### Özet

Bu çalışmanın temel amacı internetin hangi kesim tarafından hangi amaçlar için kullanıldığının incelenmesidir. Bu amaçla 2019 yılında farklı illerde 304 katılımcı ile 18-60 yaş aralığında bir anket çalışması yapılmıştır. Araştırma kapsamında internet kullanımına etki ettiği düşünülen 17 faktör belirlenmiştir. Çalışmadan elde edilen cevaplar binary ve sıralı lojistik regresyon analizi, korelasyon analizi, ortalama, frekans tablosu gibi istatistiksel metotlar yardımıyla incelenmiştir. Sıralı lojistik regresyon analizi ile elde edilen bulgulara göre; yaş, eğitim düzeyi, gelir durumu, yaşanan yer, çocuk sayısı gibi demografik değişkenlerin yanı sıra ev sahibi olma durumu, internetten dizi ya da film izleme alışkanlığı, interneti okul ya da iş için kullanma durumu bireyin interneti ne zamandan beri kullandığına etki etmektedir. Sınıflandırma analizi sonuçlarına göre J48, LMT, Simple Logistic, Random Forest ve Bayes Net algoritmaları için sırasıyla %99.01, %98.36, %98.36, %98.03 ve %97.37 doğru sınıflandırma oranları elde edilmiştir. Algoritma başarısını etkileyen Kappa istatistik değeri, ortalama mutlak hata, kök görelî kare hatası gibi değerler de bu oranlar ile paralellik göstermektedir. En yüksek başarıyı sağlayan J48 algoritmasına göre; bireyin yaşı günlük internet kullanım süresini belirleyen tek faktördür. Diğer algoritmalara göre ise; temel özellik yaş olmakla birlikte bunun yanında eğitim seviyesi, gelir düzeyi, çocuk sayısı, iş ya da okulundan memnuniyet durumu, ne zamandır internet kullandığı günlük internet kullanım süresini belirleyen diğer faktörler olarak bulunmuştur. Bu sonuçlar, bireylerin günlük yaşamlarına dair ipuçları vermekte ve internetin bireyler üzerindeki işlevsel etkileşimini ortaya koymaktadır. Ayrıca karar ağacı algoritmalarının bireylerin günlük internet kullanım süresini yüksek bir başarıyla belirlediğini ortaya koymakta ve internet kullanım süresine etki eden faktörleri belirlemede yardımcı olmaktadır.

**Anahtar kelimeler:** Binary lojistik regresyon, internet kullanımı, karar ağaçları, sınıflandırma analizi, sıralı lojistik regresyon.

## Investigation of Internet Usage Habits of Individuals with Generalized Linear Models and Classification Analysis

### Abstract

The main purpose of this study is to examine which segments use the internet for what purposes. For this purpose, a survey was conducted with 304 participants between the ages of 18-60 in different provinces in 2019. Within the scope of the research, 17 factors that were thought to have an effect on internet usage were determined. The answers obtained from the study were analyzed with the help of statistical methods such as binary and ordered logistic regression analysis, correlation analysis, mean and frequency table. According to the findings obtained by ordered logistic regression analysis; In addition to demographic variables such as age, education level, income status, place of residence,



number of children, the status of being a home owner, the habit of watching TV series or movies on the internet, using the internet for school or work affect how long an individual uses the internet. According to the results of the classification analysis, 99.01%, 98.36%, 98.36%, 98.03% and 97.37% correct classification rates were obtained for J48, LMT, Simple Logistic, Random Forest and Bayes Net algorithms, respectively. Values such as Kappa statistical value, mean absolute error, root relative square error, which affect the success of the algorithm, are also parallel to these ratios. According to the highest performance J48 algorithm; The age of an individual is the only factor that determines the duration of daily internet usage. According to other algorithms; Although the main feature is age, education level, income level, number of children, satisfaction with work or school, and how long the internet has been used are other factors that determine the duration of daily internet usage. These results give clues about the daily lives of individuals and reveal the functional interaction of the internet on individuals. In addition, it reveals that decision tree algorithms determine the daily internet usage time of individuals with high success and help to determine the factors affecting internet usage time.

**Key words:** Binary logistic regression, internet usage, decision trees, classification analysis, ordered logistic regression

## **Modeling the Frequency of Auto Insurance Claims in the Albanian Market**

Endri RAÇO<sup>1</sup>, Kleida HAXHI<sup>1</sup>, Etleva LLAGAMI, Oriana ZAÇAJ<sup>1</sup>

<sup>1</sup>Polytechnic University of Tirana, Faculty of Mathematical Engineering and Physical Engineering,  
Albania

### **Abstract**

The appropriate insurance pricing system allows non – life insurance companies to cover expected losses, expenses and technical reserves. The valuation of claims frequency or the count data is an important issue on the determination of the insurance premium. The analysis of count regression identifies the risk factors and the prediction of the expected frequency of claims. For this purpose the most significant steps are the modeling of claims frequency, the methodology of count data models and the risk factors to explain the frequency of claims. The best model is chosen by means of the log-likelihood ratio and the information criteria.

**Key words:** Count data models, risk factors, log-likelihood ratio

## Rüzgâr Hızı için SETAR Modeli ile Öngörü Elde Edilmesi

Ümran M. KAHRAMAN<sup>1</sup>, İsmail ARSEL<sup>2</sup>, Aşır GENÇ<sup>1</sup>, Galip OTURANÇ<sup>2</sup>

<sup>1</sup>Necmettin Erbakan Üniversitesi, Konya, Türkiye,

<sup>2</sup>Karamanoğlu Mehmetbey Üniversitesi, Karaman, Türkiye

### Özet

Bu çalışmada, Weibull dağılımına benzer bir dağılım gösteren rüzgâr hızı zaman serisinin kendinden uyarımlı eşiksel otoregresif (SETAR) modeli ile tahmini elde edilmeye çalışılmıştır. Eşiksel model, gözlemlerin eşik değer(ler)ine göre sıralanarak farklı ARMA modelleri ile ifade edilebilmesine imkan tanımaktadır. Weibull dağılımından üretilerek Bootstrap örnekleme yapılan n=1000 gözlemlili veri seti için bir SETAR modeli elde edilmiştir. Karaman iline ait 2000-2018 yılları arasındaki günlük rüzgâr hızı verisi için SETAR modeli belirlenmeye çalışılmıştır. Rüzgâr hızındaki mevsimsellik etkisinden dolayı haftalık verilerle elde edilen model bilgi kriterlerine göre daha iyi sonuç vermiştir. Haftalık veriler için iki rejimli SETAR modeli artıklarının GARCH etkisine sahip olduğu belirlenmiş ve modelin öngörü aralığı tahmin edilmiştir.

**Anahtar kelimeler:** Eşiksel otoregresif model, GARCH, rüzgar hızı, SETAR öngörü, Weibull dağılımı

### Abstract

In this study, the estimation of the wind speed time series, which shows a similar distribution to the Weibull distribution, was obtained by using the self-excited threshold autoregressive (SETAR) model. The threshold model allows observations to be expressed in different ARMA models by sorting them according to their threshold value (s). A SETAR model was obtained for a data set with n = 1000 observations, which was sampled from Bootstrap by generating from Weibull distribution. A SETAR model was tried to be determined for daily wind speed data between 2000 and 2018 in the province of Karaman. Due to the seasonality effect on wind speed, the model obtained with weekly data gave better results than the information criteria. For weekly data, it was determined that the residuals of the two regime SETAR model have GARCH effect and the prediction interval of the model was estimated.

**Key words:** Threshold autoregressive model, GARCH, wind speed, SETAR prediction, Weibull distribution

## Türkiye'nin İklimsel Değişikliği ve Mevsimsel Ekstrem Yağışların İstatistiksel Olarak Değerlendirilmesi

Yunus AKDOĞAN<sup>1</sup>, Saniye DEMİR<sup>2</sup>, Orhan KAVUNCU<sup>3</sup>, Bilal Habeşi ÖZKAYNAR<sup>4</sup>

<sup>1</sup>Selçuk Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 42250, Konya, Türkiye

<sup>2</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Toprak Bilimi, Taşlıçiftlik 60240, Tokat, Türkiye

<sup>3</sup>Kastamonu Üniversitesi, Genetik ve Biyomühendislik Fakültesi, Kastamonu, Türkiye

<sup>4</sup>Türkiye İstatistik Kurumu, Çankaya, 06570, Ankara, Türkiye

### Özet

Ortalama ve standart sapma değerindeki çok küçük değişiklikler, ekstrem olayların olasılıklarında çok büyük değişimlere neden olmaktadır. Bu çalışmada; Türkiye'nin farklı bölgelerindeki 5 meteorolojik istasyona ait yağış verileri kullanılmıştır. 1980-2018 yılları arasındaki yağışlar mevsimsel olarak değerlendirilmiştir. Ortalama ve standart sapmaya bağlı olarak ıslak ve kuru gün sıklığının nasıl değiştiği incelenmiştir. 1980-1999 yılları referans olarak alınmıştır.  $P_{10}$ ,  $P_{25}$ ,  $P_{75}$  ve  $P_{90}$  yüzdeliklerine göre, ortalama ve standart sapmaya bağlı olarak ekstrem mevsimlerin olasılıkları analiz edilmiştir. Ortalama, standart sapma ve ekstrem mevsimlerin olasılıkları arasında non-lineer bir ilişki bulunmuştur. Eşik değer, ortalama değer değişmesiyle yakından etkilenmektedir.

**Anahtar kelimeler:** Islak-kuru gün, ortalama, standart sapma, Türkiye

### Statistical Evaluation of Climate Change and Seasonal Rainfall Extreme of Turkey

#### Abstract

Small changes in the mean and standart deviation values can produce relatively large changes in the probability of extreme events. In this study; precipitation data of five meteorological stations in different regions of Turkey are used. The precipitation between 1980-2018 has been evaluated as seasonal.  $P_{10}$ ,  $P_{25}$ ,  $P_{75}$  and  $P_{90}$  percentiles based on the probability of extreme seasons as function of mean and standart deviation is analysed. The main conclusion is a non-linear relationship between changes in mean and standart deviation values and extreme seasons probability. With these threshold values, the main influence corresponds to changes in mean value.

**Keywords:** Wet/dry day, mean, standart deviation, Turkey

## Increasing the Power of Goodness of Fit Tests for Cauchy Distribution Versus Symmetrical Distributions for Small Samples by Means of Characterization

Ahmet PEKGÖR<sup>1</sup>

<sup>1</sup>Necmettin Erbakan University, Faculty of Science, Department of Statistics, Konya, Turkey.

### Abstract

In this study, a new algorithm is proposed to develop well-known goodness-of-fit tests for Cauchy distribution against symmetrical distributions in small samples. The characterization theorem on the Cauchy distribution is used in this new algorithm. Critical values were obtained by simulation and power analysis was performed to compare the performance of the newly introduced algorithm. After this power analysis, it was shown that the results obtained with the new algorithm have a better power against symmetrical distributions with small sample sizes than classical goodness of fit tests in the literature.

**Key words:** Cauchy distribution, characterization, goodness of fit, Monte Carlo simulation, robustness

### Introduction

Goodness of fit testing is one of the most important statistical tool in statistical inference. Most statistical analysis assume that the data come from certain distribution. In this case, goodness of fit(GOF) tests are used to determine whether the data come from certain distribution or not. Some of well-known GOF tests are Kolmogorov-Smirnov(KS), Anderson-Darling(AD) and Cramer-von Mises(CVM). Note that, these tests can be used to GOF testing for the all continuous distribution. In this study, these tests are used to test whether the data come from Cauchy distribution or not and their power are increased with an algorithm.

The Cauchy distribution is one of the most popular model in statistics. A random variable  $X$  is said to have the Cauchy distribution (denoted by  $C(\mu, \sigma)$ ) with location parameter  $\mu \in \mathbb{R}$  and scale parameter  $\sigma > 0$  if its probability density function is given by

$$f(x; \mu, \sigma) = \frac{1}{\pi\sigma \left(1 + \left(\frac{x - \mu}{\sigma}\right)^2\right)}, \quad x \in \mathbb{R} \quad (1)$$

the corresponding distribution function is also takes the form

$$F(x; \mu, \sigma) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left( \frac{x - \mu}{\sigma} \right), \quad x \in \mathbb{R}.$$

It is interesting that all raw moments are not defined. In addition, it is symmetric around median with bell-shaped density function and it is member of location-scale family. The Cauchy distribution is the first distribution that comes to mind in the long-tailed sense for a statistician. It is used as an error term distribution in some statistical models. See Johnsan et al. (1995) for an extensive bibliography. The Cauchy distribution is also used for modelling real data. For example, see also Kagan (1992); Min et al. (1996); Nolan (2014); Stapf et al. (1996) and Winterton et al., 1992).

In this paper, a characterization theorem and well-known GOF tests are combined to test whether the data come from Cauchy or not. In Section 2, a new algorithm is described. In Section 3, critical values and power of GOF test are predicted by simulation. Some concluding remarks are provided in Section 4.

## A New Algorithm

In this section, an algorithm is proposed to increase the power of KS, AD and CVM tests for the goodness of fit to Cauchy distribution. Given a random sample  $X_1, X_2, \dots, X_n$  from a population having a distribution function  $F(x)$ , consider the problem of testing  $H_0: F(x) = F_0(x; \mu, \sigma)$  for some  $\mu \in \mathbb{R}$  and  $\sigma > 0$ . The alternative hypothesis is  $H_1: F(x) \neq F_0(x; \mu, \sigma)$  for any  $\mu \in \mathbb{R}$  and  $\sigma > 0$ . The test statistics of these tests are given, respectively, by

$$KS = \max_{i=1,2,\dots,n} \left[ \max \left\{ \frac{i}{n} - m, m - \frac{i-1}{n} \right\} \right],$$

$$AD = -n - \frac{2}{n} \sum_{i=1}^n [(i-0.5)\log\{m\} + (n-i+0.5)\log\{1-m\}]$$

and

$$CVM = \frac{1}{12n} + \sum_{i=1}^n \left( m - \frac{i-0.5}{n} \right)^2,$$

where  $m$  is  $F_0(X_{(i)}; \hat{\mu}, \hat{\sigma})$ ,  $\hat{\mu}$  and  $\hat{\sigma}$  are reasonable estimates of parameters and  $X_{(i)}$  denotes  $i$ th order statistic.

In practice, null hypothesis is constructed as  $H_0$ : Data come from  $F_0 \equiv C(\hat{\mu}, \hat{\sigma})$ , where  $\hat{\mu}$  and  $\hat{\sigma}$  are the estimates of parameters  $\mu$  and  $\sigma$ . In general, maximum likelihood estimates are preferred in the analysis. Since the Cauchy distribution is long-tailed, robust estimates are used in the analysis. This null hypothesis can be viewed as the data come from any Cauchy distribution. However, asymptotic null distribution of the these test statistics are derived for specified distribution such as  $F_0 \equiv C(\mu_0, \sigma_0)$ , where  $\mu_0$  and  $\sigma_0$  are given and they are not estimates. In this regards, some problems are arisen related to critical values and  $p$  values when using KS, AD and CVM tests for goodness of fit. In this case, null distribution should be obtained by simulation for different sample sizes for testing  $H_0$ : Data come from  $C(\hat{\mu}, \hat{\sigma})$  or equivalently  $H_0$ : Data come from any Cauchy distribution. Critical values of these tests are predicted in the next section.

In this section, a new algorithm is introduced to improve the power of KS, AD and CVM tests by using a characterization lemma for Cauchy distribution. Let  $X$  follows from  $C(\mu, \sigma)$  distribution with pdf (1), then

$$W = \frac{X - \mu}{\sigma} \quad (2)$$

has standart Cauchy distribution with parameters  $\mu = 0$  and  $\sigma = 1$ . Let us consider following lemma given in Bansal et al. (1999) which helps to construct the algorithm.

**Lemma 1.** (*Cauchy Characterization*). The random variable  $W$  is  $C(0,1)$  if and only if  $(W-1)/(W+1)$  is  $C(0,1)$ .

The improving algorithm for GOF test is given below:

### Algorithm 1

**A1.** Calculate the estimates of parameters  $\mu$  and  $\sigma$ , say  $\hat{\mu}$  and  $\hat{\sigma}$ , based on the sample  $X_1, X_2, \dots, X_n$ , where

$$\hat{\mu} = \begin{cases} \frac{1}{2} \left( X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)} \right), & n \text{ is even} \\ X_{(\frac{n+1}{2})}, & \text{otherwise} \end{cases}$$

$$\hat{\sigma} = \frac{1}{2} (\xi_{0.75} - \xi_{0.25})$$

and  $X_{(i)}$  is  $i$ th order statistic ( $i = 1, 2, \dots, n$ ), and  $\xi_p$  ( $0 < p < 1$ ) is the sample  $p$ th quantile.

**A2.** Calculate the  $W_i = (X_i - \hat{\mu})/\hat{\sigma}$ ,  $i = 1, 2, \dots, n$  (Recall transformation 2: If  $X_i$  follows from  $C(\mu, \sigma)$ , then  $W_i$  can be treated as approximately distributed as  $C(0, 1)$  under  $H_0$ .)

**A3.** Calculate the  $Z_i = (W_i - 1)/(W_i + 1)$ ,  $i = 1, 2, \dots, n$  (Recall Lemma 1: If  $W_i$  follows from  $C(0, 1)$ , then  $Z_i$  follows also from  $C(0, 1)$ ).

**A4.** Calculate the estimates of parameters  $\mu$  and  $\sigma$ , say  $\tilde{\mu}$  and  $\tilde{\sigma}$ , based on the sample  $Z_1, Z_2, \dots, Z_n$ , where

$$\tilde{\mu} = \begin{cases} \frac{1}{2} \left( Z_{(\frac{n}{2})} + Z_{(\frac{n}{2}+1)} \right), & n \text{ is even} \\ Z_{(\frac{n+1}{2})}, & \text{otherwise} \end{cases}$$

$$\tilde{\sigma} = \frac{1}{2} (\xi_{0.75}^* - \xi_{0.25}^*)$$

and  $Z_{(i)}$  is  $i$ th order statistic ( $i = 1, 2, \dots, n$ ) and  $\xi_p^*$  ( $0 < p < 1$ ) is the sample  $p$ th quantile.

**A5.** Calculate the  $W_i^* = (Z_i - \tilde{\mu})/\tilde{\sigma}$ ,  $i = 1, 2, \dots, n$

**A6.** Null hypothesis is  $H_0: W_i^*$  come from  $C(\tilde{\mu}, \tilde{\sigma}) \equiv X_1, X_2, \dots, X_n$  come from any Cauchy distribution.

**A7.** Calculate the KS(or AD or CVM) test statistic based on  $W_i^*$ ,  $i = 1, 2, \dots, n$  and compare with critical values.

The test procedures used Algorithm 1 are called modified KS(M-KS), modified AD(M-AD) and modified CVM(M-CVM). In this paper, it is investigated that the step A3 may improve the power of the existence tests such as KS, AD or CVM tests. The critical values and the power of the introduced methodology are given in next section.

### Simulation Study

In the simulation study, 200 000 trials are conducted to obtain critical values for GOF tests and subsequently to obtain power predictions. R programming language is used to obtain critical values and estimate power. The quantiles of GOF tests and modified GOF tests under null distributions are given in table 1 for  $\alpha = 0.01$  and  $\alpha = 0.05$ . The sample sizes are considered as  $n = 10: 5: 100$ . From figure 1, first three panel show that the test procedure with Algorithm 1 has desired significance level of  $\alpha$ . The power of the tests are also given tables 2-3 and figures 1-2. In the power analysis, symmetric alternatives are chosen as normal(mean, sd), logistic(location, scale) and t(df) distributions which have pdf similar to pdf of Cauchy(location, scale). According to the these tables and figures, the powers of introduced methodology(M-KS and M-CVM) are higher than the power of existence tests for symmetric alternative distributions when sample size less than 100.

It should be noted that same patterns are observed for  $n \geq 100$ . The power of M-AD is higher than the power of AD test for symmetric alternative distributions when the sample size less than 50 in general. In this regards, the practitioner should get the histogram of the data. If a symmetrical pattern is observed, the suggested method is suitable to test GOF for small or moderate sample sizes.

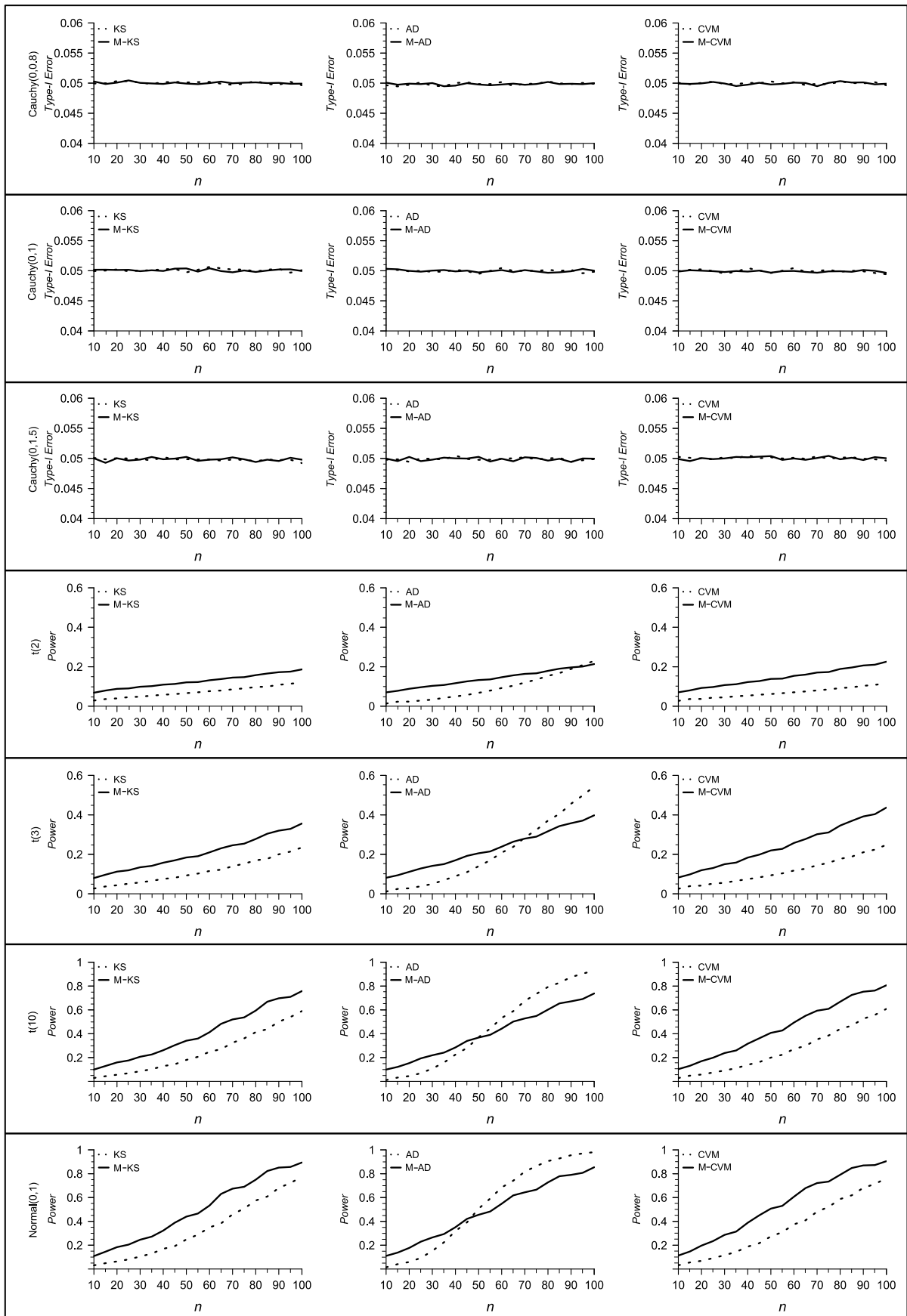


Figure 1. Type-I Errors and the Power of Tests for Symmetric Alternatives



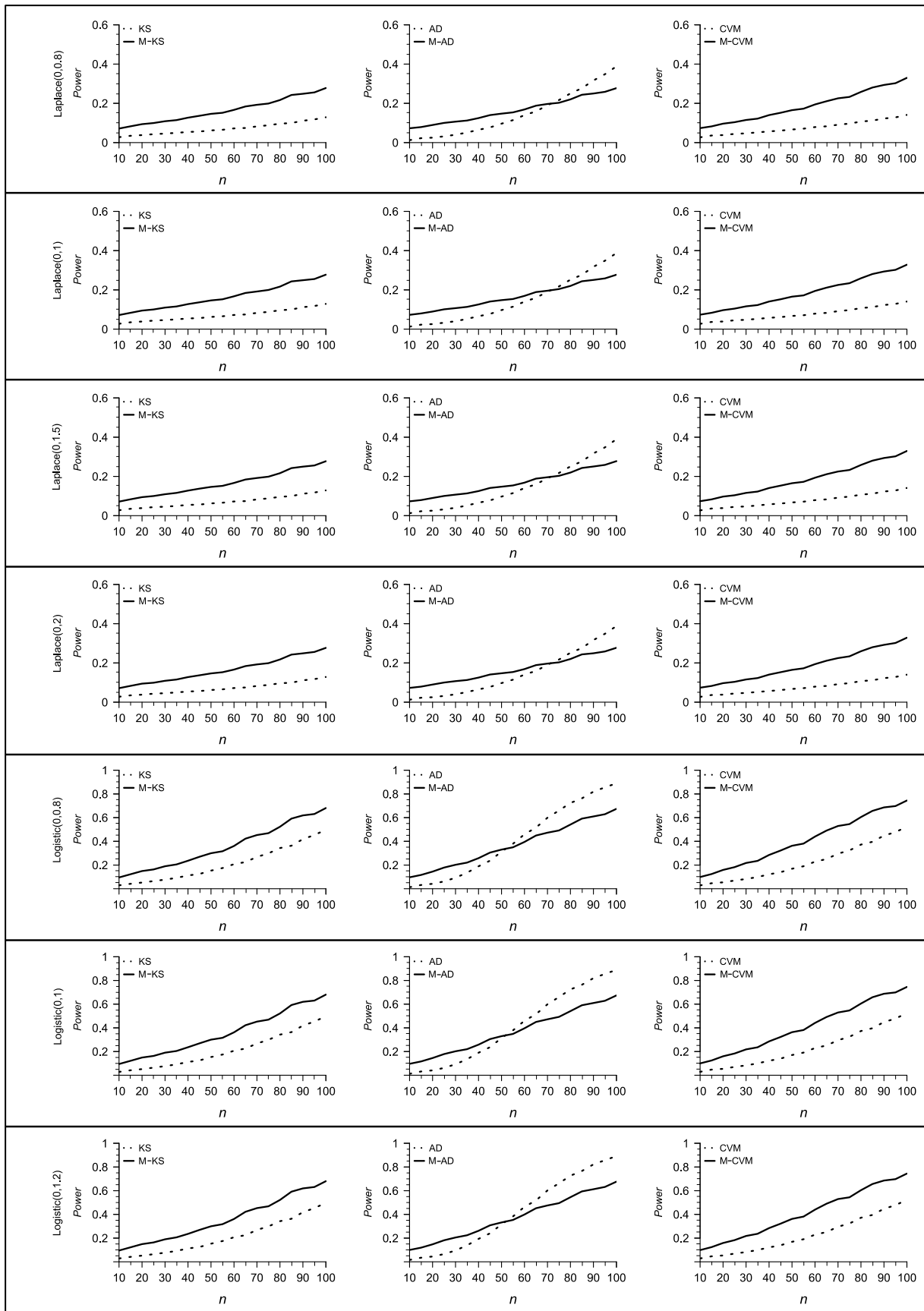


Figure 2. The power of Tests for Symmetric Alternatives

**1<sup>st</sup> International Applied Statistics Conference (UYIK-2020)**  
**Online – Tokat / Turkey, 1-4 October 2020**

Table 1. Critical values for  $\alpha = 0.01$  and  $\alpha = 0.05$

$\alpha$	Statistics	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$	$n = 55$	$n = 60$	$n = 65$	$n = 70$	$n = 75$	$n = 80$	$n = 85$	$n = 90$	$n = 95$	$n = 100$
0.01	<i>KS</i>	0.3391	0.2848	0.2423	0.2207	0.1980	0.1843	0.1708	0.1630	0.1527	0.1461	0.1390	0.1349	0.1285	0.1246	0.1202	0.1175	0.1132	0.1105	0.1073
	<i>M-KS</i>	0.3434	0.2953	0.2523	0.2283	0.2043	0.1896	0.1753	0.1666	0.1558	0.1487	0.1414	0.1369	0.1303	0.1263	0.1216	0.1189	0.1145	0.1116	0.1084
	<i>AD</i>	1.3221	1.4515	1.4279	1.5263	1.4807	1.5140	1.4921	1.5360	1.5072	1.5252	1.5085	1.5378	1.5124	1.5275	1.5163	1.5396	1.5173	1.5309	1.5204
	<i>M-AD</i>	1.3378	1.5018	1.4964	1.5827	1.5417	1.5699	1.5491	1.5888	1.5582	1.5709	1.5557	1.5836	1.5571	1.5712	1.5573	1.5799	1.5600	1.5698	1.5591
	<i>CVM</i>	0.2115	0.2328	0.2177	0.2308	0.2187	0.2264	0.2180	0.2259	0.2181	0.2228	0.2173	0.2228	0.2165	0.2209	0.2165	0.2213	0.2161	0.2195	0.2160
	<i>M-CVM</i>	0.2231	0.2461	0.2332	0.2428	0.2294	0.2349	0.2264	0.2322	0.2244	0.2279	0.2221	0.2273	0.2203	0.2242	0.2200	0.2242	0.2193	0.2224	0.2186
0.05	<i>KS</i>	0.2700	0.2314	0.1962	0.1820	0.1630	0.1530	0.1417	0.1363	0.1276	0.1226	0.1167	0.1136	0.1084	0.1052	0.1015	0.0994	0.0958	0.0936	0.0910
	<i>M-KS</i>	0.2820	0.2407	0.2038	0.1867	0.1667	0.1561	0.1443	0.1384	0.1293	0.1242	0.1181	0.1148	0.1093	0.1062	0.1023	0.1002	0.0965	0.0943	0.0915
	<i>AD</i>	0.9239	1.0098	0.9844	1.0631	1.0218	1.0473	1.0292	1.0671	1.0408	1.0568	1.0435	1.0682	1.0490	1.0591	1.0500	1.0678	1.0530	1.0613	1.0538
	<i>M-AD</i>	0.9655	1.0510	1.0342	1.1027	1.0631	1.0832	1.0686	1.1015	1.0756	1.0885	1.0769	1.0989	1.0786	1.0877	1.0785	1.0963	1.0808	1.0865	1.0800
	<i>CVM</i>	0.1295	0.1480	0.1364	0.1486	0.1396	0.1463	0.1403	0.1473	0.1414	0.1457	0.1418	0.1466	0.1422	0.1451	0.1422	0.1459	0.1426	0.1449	0.1428
	<i>M-CVM</i>	0.1417	0.1570	0.1456	0.1545	0.1448	0.1502	0.1444	0.1502	0.1442	0.1479	0.1441	0.1484	0.1439	0.1465	0.1437	0.1472	0.1438	0.1458	0.1438

1<sup>st</sup> International Applied Statistics Conference (UYIK-2020)  
Online – Tokat / Turkey, 1-4 October 2020

Table 2. The Power of Tests for Symmetric Alternatives for  $\alpha = 0.05$

Distributions	Statistics	n = 10	n = 15	n = 20	n = 25	n = 30	n = 35	n = 40	n = 45	n = 50	n = 55	n = 60	n = 65	n = 70	n = 75	n = 80	n = 85	n = 90	n = 95	n = 100	
Cauchy(0,0.8)	KS	0.0499	0.0499	0.0503	0.0503	0.0500	0.0498	0.0501	0.0502	0.0501	0.0501	0.0502	0.0499	0.0497	0.0499	0.0503	0.0498	0.0497	0.0502	0.0496	
	M-KS	0.0502	0.0498	0.0500	0.0504	0.0500	0.0499	0.0498	0.0501	0.0499	0.0498	0.0500	0.0502	0.0502	0.0499	0.0500	0.0501	0.0500	0.0500	0.0498	0.0499
	AD	0.0498	0.0496	0.0499	0.0502	0.0497	0.0496	0.0502	0.0503	0.0500	0.0500	0.0503	0.0497	0.0500	0.0500	0.0504	0.0501	0.0500	0.0500	0.0502	0.0500
	M-AD	0.0502	0.0499	0.0500	0.0500	0.0501	0.0496	0.0497	0.0501	0.0499	0.0498	0.0499	0.0500	0.0500	0.0499	0.0500	0.0503	0.0500	0.0500	0.0500	0.0501
	CVM	0.0500	0.0498	0.0500	0.0503	0.0499	0.0499	0.0501	0.0500	0.0503	0.0500	0.0503	0.0497	0.0497	0.0499	0.0498	0.0503	0.0500	0.0500	0.0502	0.0496
Cauchy(0,1)	M-CVM	0.0499	0.0499	0.0500	0.0502	0.0499	0.0495	0.0497	0.0501	0.0498	0.0499	0.0501	0.0500	0.0495	0.0501	0.0503	0.0501	0.0501	0.0501	0.0498	0.0499
	KS	0.0499	0.0500	0.0501	0.0498	0.0500	0.0500	0.0502	0.0502	0.0497	0.0501	0.0506	0.0503	0.0501	0.0501	0.0498	0.0501	0.0502	0.0502	0.0496	0.0500
	M-KS	0.0501	0.0501	0.0501	0.0501	0.0499	0.0500	0.0499	0.0503	0.0503	0.0498	0.0503	0.0499	0.0503	0.0499	0.0500	0.0497	0.0500	0.0501	0.0502	0.0499
	AD	0.0504	0.0502	0.0500	0.0500	0.0499	0.0499	0.0502	0.0500	0.0495	0.0501	0.0505	0.0500	0.0501	0.0500	0.0502	0.0500	0.0502	0.0500	0.0502	0.0496
	M-AD	0.0504	0.0503	0.0500	0.0499	0.0501	0.0502	0.0500	0.0501	0.0498	0.0500	0.0502	0.0498	0.0502	0.0499	0.0497	0.0498	0.0498	0.0500	0.0504	0.0501
Cauchy(0,1.5)	CVM	0.0500	0.0502	0.0503	0.0500	0.0496	0.0499	0.0505	0.0501	0.0497	0.0501	0.0505	0.0499	0.0500	0.0502	0.0499	0.0501	0.0501	0.0500	0.0497	0.0495
	M-CVM	0.0499	0.0502	0.0501	0.0500	0.0499	0.0500	0.0499	0.0501	0.0498	0.0500	0.0500	0.0499	0.0499	0.0497	0.0500	0.0500	0.0499	0.0502	0.0501	0.0497
	KS	0.0502	0.0500	0.0501	0.0501	0.0500	0.0498	0.0503	0.0499	0.0500	0.0499	0.0499	0.0497	0.0499	0.0499	0.0497	0.0499	0.0499	0.0498	0.0499	0.0493
	M-KS	0.0501	0.0494	0.0501	0.0498	0.0499	0.0503	0.0500	0.0501	0.0504	0.0497	0.0499	0.0500	0.0503	0.0500	0.0495	0.0499	0.0499	0.0497	0.0502	0.0499
	AD	0.0499	0.0499	0.0495	0.0501	0.0500	0.0499	0.0506	0.0500	0.0499	0.0499	0.0500	0.0499	0.0501	0.0503	0.0499	0.0501	0.0502	0.0495	0.0498	0.0500
t(2)	M-AD	0.0501	0.0497	0.0503	0.0496	0.0499	0.0502	0.0501	0.0501	0.0503	0.0496	0.0500	0.0496	0.0503	0.0502	0.0497	0.0500	0.0495	0.0501	0.0500	
	CVM	0.0502	0.0500	0.0499	0.0498	0.0501	0.0499	0.0504	0.0499	0.0501	0.0499	0.0500	0.0498	0.0503	0.0499	0.0500	0.0501	0.0501	0.0498	0.0498	0.0495
	M-CVM	0.0498	0.0495	0.0500	0.0498	0.0499	0.0502	0.0501	0.0502	0.0503	0.0497	0.0499	0.0497	0.0500	0.0503	0.0498	0.0500	0.0500	0.0496	0.0501	0.0499
	KS	0.0294	0.0370	0.0406	0.0468	0.0491	0.0545	0.0589	0.0623	0.0671	0.0714	0.0770	0.0801	0.0864	0.0914	0.0982	0.1003	0.1093	0.1147	0.1214	
	M-KS	0.0693	0.0803	0.0889	0.0910	0.0993	0.1027	0.1100	0.1136	0.1217	0.1232	0.1326	0.1386	0.1460	0.1481	0.1582	0.1665	0.1731	0.1762	0.1875	
t(3)	AD	0.0142	0.0238	0.0241	0.0295	0.0341	0.0432	0.0503	0.0584	0.0687	0.0801	0.0930	0.1047	0.1209	0.1358	0.1545	0.1687	0.1897	0.2090	0.2308	
	M-AD	0.0711	0.0787	0.0890	0.0971	0.1040	0.1084	0.1175	0.1268	0.1337	0.1368	0.1479	0.1574	0.1646	0.1675	0.1801	0.1909	0.1974	0.2020	0.2146	
	CVM	0.0282	0.0375	0.0376	0.0444	0.0456	0.0514	0.0544	0.0589	0.0625	0.0669	0.0713	0.0758	0.0812	0.0860	0.0927	0.0967	0.1041	0.1090	0.1171	
	M-CVM	0.0715	0.0812	0.0933	0.0979	0.1081	0.1123	0.1233	0.1282	0.1392	0.1408	0.1547	0.1609	0.1715	0.1739	0.1897	0.1971	0.2073	0.2111	0.2265	
	KS	0.0272	0.0378	0.0438	0.0523	0.0582	0.0669	0.0761	0.0823	0.0937	0.1033	0.1163	0.1231	0.1406	0.1533	0.1702	0.1783	0.2009	0.2147	0.2357	
t(10)	M-KS	0.0805	0.0976	0.1133	0.1201	0.1353	0.1423	0.1583	0.1702	0.1848	0.1913	0.2109	0.2312	0.2465	0.2548	0.2788	0.3053	0.3209	0.3297	0.3572	
	AD	0.0121	0.0257	0.0289	0.0390	0.0507	0.0699	0.0900	0.1106	0.1407	0.1707	0.2094	0.2385	0.2860	0.3248	0.3721	0.4058	0.4573	0.4989	0.5445	
	M-AD	0.0820	0.0946	0.1119	0.1293	0.1422	0.1513	0.1707	0.1931	0.2061	0.2153	0.2397	0.2648	0.2804	0.2903	0.3175	0.3439	0.3585	0.3717	0.3985	
	CVM	0.0262	0.0398	0.0423	0.0526	0.0566	0.0666	0.0750	0.0840	0.0942	0.1042	0.1186	0.1276	0.1455	0.1575	0.1780	0.1888	0.2112	0.2247	0.2491	
	M-CVM	0.0835	0.0990	0.1202	0.1312	0.1509	0.1589	0.1843	0.1993	0.2203	0.2284	0.2585	0.2782	0.3028	0.3117	0.3468	0.3701	0.3931	0.4042	0.4381	
Normal(0,1)	KS	0.0291	0.0441	0.0560	0.0691	0.0857	0.1029	0.1294	0.1441	0.1809	0.2067	0.2490	0.2707	0.3248	0.3611	0.4139	0.4392	0.5014	0.5408	0.5907	
	M-KS	0.0996	0.1296	0.1598	0.1765	0.2075	0.2260	0.2623	0.3042	0.3418	0.3602	0.4131	0.4839	0.5209	0.5377	0.5946	0.6693	0.6979	0.7099	0.7588	
	AD	0.0133	0.0344	0.0467	0.0708	0.1095	0.1602	0.2272	0.2829	0.3731	0.4456	0.5349	0.5920	0.6782	0.7368	0.7950	0.8304	0.8776	0.9064	0.9324	
	M-AD	0.1015	0.1237	0.1557	0.1954	0.2216	0.2437	0.2874	0.3422	0.3704	0.3931	0.4451	0.5040	0.5304	0.5513	0.6045	0.6572	0.6738	0.6934	0.7402	
	CVM	0.0291	0.0490	0.0582	0.0757	0.0923	0.1118	0.1403	0.1622	0.2004	0.2249	0.2730	0.2990	0.3537	0.3859	0.4412	0.4683	0.5270	0.5586	0.6098	
Normal(0,1)	M-CVM	0.1038	0.1317	0.1708	0.1993	0.2388	0.2607	0.3168	0.3621	0.4080	0.4279	0.4945	0.5517	0.5943	0.6096	0.6694	0.7259	0.7539	0.7628	0.8077	
	KS	0.0313	0.0492	0.0643	0.0813	0.1055	0.1299	0.1698	0.1924	0.2484	0.2888	0.3495	0.3835	0.4596	0.5087	0.5755	0.6089	0.6805	0.7230	0.7740	
	M-KS	0.1088	0.1456	0.1840	0.2052	0.2463	0.2724	0.3230	0.3890	0.4402	0.4660	0.5333	0.6307	0.6749	0.6906	0.7499	0.8231	0.8518	0.8567	0.8943	
	AD	0.0149	0.0410	0.0591	0.0933	0.1520	0.2247	0.3194	0.3925	0.5066	0.5921	0.6851	0.7418	0.8183	0.8643	0.9063	0.9285	0.9555	0.9701	0.9813	
	M-AD	0.1106	0.1379	0.1774	0.2299	0.2649	0.2936	0.3512	0.4217	0.4555	0.4833	0.5484	0.6185	0.6440	0.6667	0.7273	0.7795	0.7907	0.8078	0.8544	
Normal(0,1)	CVM	0.0316	0.0559	0.0686	0.0913	0.1162	0.1439	0.1880	0.2168	0.2740	0.3096	0.3747	0.4092	0.4818	0.5211	0.5879	0.6188	0.6817	0.7132	0.7632	
	M-CVM	0.1133	0.1477	0.1971	0.2348	0.2866	0.3148	0.3887	0.4507	0.5072	0.5309	0.6072	0.6793	0.7219	0.7340	0.7917	0.8482	0.8698	0.8728	0.9059	

**Table 3.** The power of tests for symmetric alternatives for  $\alpha = 0.05$

Distributions	Statistics	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$	$n = 55$	$n = 60$	$n = 65$	$n = 70$	$n = 75$	$n = 80$	$n = 85$	$n = 90$	$n = 95$	$n = 100$
		Laplace(0,0.8)	<i>KS</i>	0.0278	0.0348	0.0384	0.0430	0.0460	0.0495	0.0542	0.0563	0.0620	0.0652	0.0722	0.0743	0.0826	0.0882	0.0964	0.0995	0.1114
<i>M-KS</i>	0.0716		0.0829	0.0942	0.0997	0.1090	0.1147	0.1269	0.1367	0.1467	0.1520	0.1666	0.1842	0.1928	0.1991	0.2168	0.2427	0.2488	0.2559	0.2783
<i>AD</i>	0.0125		0.0227	0.0253	0.0320	0.0400	0.0516	0.0651	0.0774	0.0963	0.1146	0.1403	0.1591	0.1911	0.2169	0.2509	0.2770	0.3167	0.3484	0.3886
<i>M-AD</i>	0.0730		0.0792	0.0896	0.1004	0.1068	0.1126	0.1250	0.1403	0.1475	0.1542	0.1688	0.1882	0.1970	0.2030	0.2205	0.2437	0.2503	0.2587	0.2775
<i>CVM</i>	0.0271		0.0362	0.0383	0.0444	0.0471	0.0520	0.0571	0.0612	0.0666	0.0709	0.0792	0.0831	0.0912	0.0968	0.1072	0.1122	0.1226	0.1283	0.1412
Laplace(0,1)	<i>M-CVM</i>	0.0737	0.0826	0.0967	0.1040	0.1154	0.1222	0.1402	0.1520	0.1655	0.1727	0.1943	0.2108	0.2256	0.2330	0.2585	0.2805	0.2939	0.3026	0.3300
	<i>KS</i>	0.0276	0.0348	0.0388	0.0434	0.0462	0.0496	0.0546	0.0563	0.0621	0.0653	0.0723	0.0746	0.0828	0.0879	0.0962	0.0998	0.1118	0.1177	0.1285
	<i>M-KS</i>	0.0717	0.0833	0.0943	0.1000	0.1097	0.1154	0.1276	0.1369	0.1464	0.1519	0.1672	0.1843	0.1918	0.1996	0.2166	0.2432	0.2494	0.2554	0.2778
	<i>AD</i>	0.0125	0.0228	0.0255	0.0326	0.0403	0.0521	0.0656	0.0775	0.0970	0.1145	0.1403	0.1593	0.1910	0.2178	0.2523	0.2765	0.3177	0.3485	0.3876
	<i>M-AD</i>	0.0728	0.0799	0.0897	0.1008	0.1070	0.1131	0.1257	0.1406	0.1477	0.1535	0.1692	0.1887	0.1955	0.2040	0.2201	0.2440	0.2506	0.2584	0.2771
Laplace(0,1.5)	<i>CVM</i>	0.0269	0.0365	0.0387	0.0450	0.0476	0.0523	0.0576	0.0613	0.0671	0.0707	0.0790	0.0829	0.0916	0.0969	0.1073	0.1124	0.1228	0.1285	0.1407
	<i>M-CVM</i>	0.0735	0.0832	0.0968	0.1043	0.1161	0.1226	0.1406	0.1520	0.1654	0.1718	0.1945	0.2111	0.2250	0.2341	0.2591	0.2807	0.2935	0.3022	0.3286
	<i>KS</i>	0.0277	0.0349	0.0387	0.0435	0.0458	0.0497	0.0550	0.0563	0.0624	0.0658	0.0722	0.0743	0.0830	0.0881	0.0962	0.0994	0.1117	0.1178	0.1288
	<i>M-KS</i>	0.0718	0.0835	0.0945	0.0999	0.1092	0.1159	0.1276	0.1377	0.1465	0.1520	0.1669	0.1843	0.1920	0.1988	0.2163	0.2422	0.2498	0.2559	0.2776
	<i>AD</i>	0.0125	0.0228	0.0253	0.0324	0.0400	0.0520	0.0658	0.0777	0.0972	0.1148	0.1407	0.1595	0.1909	0.2175	0.2513	0.2769	0.3172	0.3483	0.3895
Laplace(0,2)	<i>M-AD</i>	0.0729	0.0796	0.0901	0.1004	0.1071	0.1135	0.1253	0.1412	0.1478	0.1540	0.1680	0.1887	0.1958	0.2030	0.2196	0.2432	0.2507	0.2587	0.2778
	<i>CVM</i>	0.0270	0.0365	0.0385	0.0448	0.0469	0.0521	0.0577	0.0617	0.0674	0.0711	0.0790	0.0830	0.0917	0.0971	0.1070	0.1119	0.1230	0.1284	0.1411
	<i>M-CVM</i>	0.0738	0.0830	0.0971	0.1038	0.1159	0.1229	0.1409	0.1528	0.1654	0.1723	0.1933	0.2113	0.2249	0.2328	0.2583	0.2803	0.2938	0.3024	0.3298
	<i>KS</i>	0.0279	0.0349	0.0386	0.0434	0.0458	0.0498	0.0545	0.0565	0.0622	0.0654	0.0720	0.0749	0.0827	0.0877	0.0967	0.0991	0.1115	0.1176	0.1284
	<i>M-KS</i>	0.0716	0.0830	0.0946	0.0992	0.1092	0.1153	0.1277	0.1371	0.1465	0.1526	0.1661	0.1841	0.1926	0.1990	0.2175	0.2426	0.2488	0.2560	0.2770
Logistic(0,0.8)	<i>AD</i>	0.0125	0.0228	0.0254	0.0321	0.0400	0.0519	0.0650	0.0778	0.0973	0.1148	0.1410	0.1591	0.1911	0.2174	0.2520	0.2760	0.3165	0.3487	0.3879
	<i>M-AD</i>	0.0728	0.0792	0.0899	0.1000	0.1071	0.1129	0.1254	0.1411	0.1475	0.1539	0.1687	0.1891	0.1967	0.2031	0.2202	0.2436	0.2492	0.2586	0.2772
	<i>CVM</i>	0.0270	0.0366	0.0385	0.0446	0.0473	0.0521	0.0569	0.0617	0.0674	0.0713	0.0793	0.0829	0.0919	0.0969	0.1077	0.1115	0.1226	0.1286	0.1404
	<i>M-CVM</i>	0.0738	0.0828	0.0971	0.1036	0.1159	0.1230	0.1406	0.1526	0.1649	0.1727	0.1939	0.2111	0.2249	0.2336	0.2595	0.2802	0.2923	0.3019	0.3287
	<i>KS</i>	0.0286	0.0419	0.0525	0.0644	0.0773	0.0916	0.1120	0.1244	0.1529	0.1745	0.2073	0.2251	0.2686	0.2982	0.3431	0.3631	0.4185	0.4529	0.4982
Logistic(0,1)	<i>M-KS</i>	0.0951	0.1223	0.1493	0.1628	0.1896	0.2051	0.2364	0.2695	0.2997	0.3163	0.3610	0.4226	0.4533	0.4689	0.5227	0.5918	0.6186	0.6307	0.6811
	<i>AD</i>	0.0130	0.0318	0.0419	0.0618	0.0930	0.1350	0.1894	0.2366	0.3132	0.3792	0.4594	0.5160	0.5996	0.6598	0.7240	0.7627	0.8192	0.8555	0.8898
	<i>M-AD</i>	0.0965	0.1163	0.1448	0.1785	0.2020	0.2203	0.2579	0.3050	0.3301	0.3496	0.3950	0.4491	0.4732	0.4921	0.5421	0.5918	0.6100	0.6286	0.6739
	<i>CVM</i>	0.0284	0.0465	0.0538	0.0693	0.0820	0.0985	0.1216	0.1391	0.1687	0.1900	0.2284	0.2505	0.2962	0.3233	0.3725	0.3959	0.4488	0.4771	0.5248
	<i>M-CVM</i>	0.0987	0.1241	0.1589	0.1823	0.2171	0.2350	0.2848	0.3219	0.3624	0.3806	0.4401	0.4909	0.5294	0.5451	0.6049	0.6570	0.6862	0.6978	0.7445
Logistic(0,1.2)	<i>KS</i>	0.0283	0.0422	0.0518	0.0643	0.0772	0.0917	0.1125	0.1247	0.1537	0.1746	0.2071	0.2253	0.2681	0.2990	0.3426	0.3641	0.4192	0.4533	0.4977
	<i>M-KS</i>	0.0950	0.1221	0.1497	0.1629	0.1904	0.2046	0.2363	0.2695	0.3003	0.3160	0.3622	0.4219	0.4526	0.4693	0.5216	0.5926	0.6197	0.6304	0.6807
	<i>AD</i>	0.0128	0.0318	0.0412	0.0619	0.0933	0.1351	0.1902	0.2370	0.3146	0.3788	0.4602	0.5152	0.5988	0.6599	0.7240	0.7634	0.8193	0.8552	0.8893
	<i>M-AD</i>	0.0965	0.1165	0.1452	0.1786	0.2023	0.2202	0.2583	0.3053	0.3300	0.3485	0.3964	0.4496	0.4721	0.4925	0.5412	0.5911	0.6094	0.6283	0.6736
	<i>CVM</i>	0.0281	0.0466	0.0532	0.0695	0.0822	0.0992	0.1216	0.1392	0.1695	0.1900	0.2288	0.2503	0.2950	0.3237	0.3724	0.3957	0.4484	0.4768	0.5247
Logistic(0,1.2)	<i>M-CVM</i>	0.0989	0.1238	0.1589	0.1825	0.2173	0.2350	0.2850	0.3218	0.3623	0.3803	0.4405	0.4905	0.5287	0.5455	0.6041	0.6571	0.6868	0.6977	0.7442
	<i>KS</i>	0.0283	0.0420	0.0522	0.0641	0.0768	0.0919	0.1126	0.1243	0.1530	0.1751	0.2076	0.2248	0.2690	0.2984	0.3425	0.3634	0.4182	0.4533	0.4978
	<i>M-KS</i>	0.0949	0.1224	0.1491	0.1633	0.1900	0.2060	0.2359	0.2693	0.2996	0.3176	0.3617	0.4218	0.4531	0.4686	0.5218	0.5918	0.6193	0.6312	0.6804
	<i>AD</i>	0.0128	0.0320	0.0417	0.0620	0.0928	0.1356	0.1901	0.2372	0.3130	0.3795	0.4612	0.5151	0.5990	0.6593	0.7238	0.7634	0.8192	0.8545	0.8892
	<i>M-AD</i>	0.0970	0.1164	0.1444	0.1785	0.2019	0.2210	0.2583	0.3047	0.3291	0.3503	0.3957	0.4487	0.4724	0.4918	0.5421	0.5909	0.6096	0.6285	0.6735
Logistic(0,1.2)	<i>CVM</i>	0.0281	0.0464	0.0536	0.0692	0.0821	0.0994	0.1219	0.1390	0.1689	0.1906	0.2293	0.2508	0.2962	0.3232	0.3721	0.3956	0.4479	0.4771	0.5248
	<i>M-CVM</i>	0.0991	0.1241	0.1585	0.1826	0.2180	0.2363	0.2846	0.3214	0.3622	0.3813	0.4403	0.4900	0.5293	0.5445	0.6043	0.6564	0.6860	0.6975	0.7444

## Results and Discussion

A new algorithm is introduced for goodness of fit testing on Cauchy distribution. The critical values are simulated for  $\alpha = 0.01$  and  $\alpha = 0.05$ . According to the results of simulation section, if the data exhibit symmetrical pattern and sample size is small or moderate, new methodology should be used to goodness of fit to Cauchy distribution. When the histogram based on given data has skewed form, the new algorithm is not recommended.

## References

- Bansal, N., Hamedani, G. G., Key, E., Volkmer, H., Zhang, H., & Behboodian, J. (1999). Some characterizations of the normal distribution. *Statistics & probability letters*, 42(4), 393-400.
- Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions*. Vol.2 (2nd. ed), New York: John Wiley & Sons, Ltd. 0-471-58494-0.
- Kagan, Y. Y. (1992). Correlations of earthquake focal mechanisms. *Geophysical Journal International*, 110(2), 305-320.
- Min, I. A., Mezić, I., & Leonard, A. (1996). Levy stable distributions for velocity and velocity difference in systems of vortex elements. *Physics of fluids*, 8(5), 1169-1180.
- Nolan, J. P. (2014). Financial modeling with heavy-tailed stable distributions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(1), 45-55.
- Stapf, S., Kimmich, R., Seitter, R. O., Maklakov, A. I., & Skirda, V. D. (1996). Proton and deuteron field-cycling NMR relaxometry of liquids confined in porous glasses. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 115, 107-114.
- Winterton, S. S., Smy, T. J., & Tarr, N. G. (1992). On the source of scatter in contact resistance data. *Journal of electronic materials*, 21(9), 917-921.

## Kantil Regresyon Analizi ile Saanen Keçilerinde Bazı Vücut Ölçümleri Kullanılarak Canlı Ağırlık Tahmini

Lütfi BAYYURT<sup>1</sup>, Samet Hasan ABACI<sup>2</sup>, Hasan ÖNDER<sup>2</sup>, Yalçın TAHTALI<sup>1</sup>

<sup>1</sup>Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Zootečni Bölümü, 60240 Tokat, Türkiye

<sup>2</sup>Ondokuz Mayıs Üniversitesi, Ziraat Fakültesi, Zootečni Bölümü, 55139 Samsun, Türkiye

### Özet

Bu çalışmanın amacı regresyon analizinde hata terimlerinin normal dağılıma sahip olmaması ve modelin aykırı değerler içermesi durumunda işlevini yitiren en küçük kareler (EKK) metodu ile alternatif regresyon yöntemlerinden olan Kantil regresyonu karşılaştırmaktır. Bu amaçla Samsunun Bafra ilçesinde özel bir işletmede yetiştirilen 92 baş Saanen oğlaklarından alınmış doğum ağırlıkları (CA) ve bazı vücut özelliklerinden vücut uzunluğu (VU), cidago yüksekliği (CY), sağrı yüksekliği (SY) ve göğüs çevresi (GÇ) ölçümleri kullanılmıştır. EKK ile Kantil regresyon modelinin tahmin performansını karşılaştırmak için düzeltilmiş belirtme katsayısı ( $R^2_{Adj}$ ) ve hata kareler ortalamasının karekökü (HKOK) (Root Mean Square Error-RMSE), ortalama mutlak sapma (OMS) (Mean Absolute Deviation-MAD) ve ortalama mutlak yüzde hata (OMYH) (Mean Absolute Percentage Error-MAPE) performans kriterleri kullanılmıştır. İncelenen yöntemlerin  $R^2$ , RMSE, MAD ve MAPE değerleri karşılaştırıldığında; Saanen oğlaklarının canlı ağırlık tahmininde Kantil regresyon yaklaşımının EKK metodundan daha iyi sonuç verdiği belirlenmiştir.

**Anahtar kelimeler:** Canlı ağırlık tahmini, en küçük kareler, Kantil regresyon analizi, Saanen, vücut ölçümleri

## Bazı Büyüme Modellerinin Lineer Eğik Asimptotlu Olduklarındaki Uyum Performansları

Mehmet KORKMAZ<sup>1</sup>, Yeliz KAŞKO ARICI<sup>2</sup>

<sup>1</sup>Ordu University, Faculty of Science and Letters, Department of Mathematics, Ordu, Türkiye

<sup>2</sup>Ordu University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Ordu, Türkiye

### Özet

Bu çalışmada, çoğu genellikle bir üs yatay asimptota sahip olan bazı iyi bilinen büyüme modelleri üs doğrusal eğik asimptotlu büyüme modellerine dönüştürülmüştür. Lineer eğik asimptotlu bir büyüme fonksiyonu elde etmek için  $[F(t)=m(t)f(t)]$ , büyüme fonksiyonu  $[f(t)]$  doğrusal artan fonksiyonla çarpıldı  $[m(t)=pt+q ; p>0$  ve  $q\geq 0$  olmak üzere]. Sayısal örnekler üzerinden model parametreleri hesaplanmış ve modellerin performansları yaygın olarak kullanılan uyum kriterlerine ( $R^2$ ,  $R^2_{adj}$ , HKO, AIC) göre karşılaştırılmıştır. Sonuç olarak, kullanılan uyum kriterleri bakımından lineer eğik asimptotlu modellerin klasik yatay asimptotlu modellere göre oldukça iyi olduğu görülmüştür.

**Anahtar kelimeler:** Büyüme modeli; sigmoidal model, lineer eğik asimptot, yatay asimptot, uyum kriteri

### Fit Performances of Some Growth Models When They Have Linear Oblique Asymptote

#### Abstract

In this study, some well known growth models, most of which generally have an upper horizontal asymptote, were turned into the growth models with an upper linear oblique asymptote. To obtain a growth function with a linear oblique asymptote ( $F(t)=m(t)f(t)$ ), the growth function ( $f(t)$ ) was multiplied by a linear increasing function ( $m(t)=pt+q; p>0$  and  $q\geq 0$ ). The model parameters were calculated and the models' performances were compared based on commonly used goodness-of-fit criteria ( $R^2$ ,  $R^2_{adj}$ , MSE, AIC). As a results, it was observed that the models with linear oblique asymptote were quite better than the models with classical horizontal asymptote in terms of the goodness-of-fit criterias used.

**Key words:** Growth model, sigmoidal model, linear oblique asymptote, horizontal asymptote, goodness of-fit criteria

## **Evaluation of Twitter Messages Related to Covid-19 Pandemic Using Text Mining Technique**

Ebru KESKİN<sup>1</sup>, Mehmet MENDEŞ<sup>1</sup>

<sup>1</sup>Çanakkale Onsekiz Mart University, Faculty of Agriculture, Department of Animal Science,  
Biometry and Genetics Unit, Çanakkale, Turkey

### **Abstract**

This study has been carried out to generate summarized information about the public's perspective and thoughts on Covid-19 pandemic by evaluating the posts and messages sent on social media about Covid-19 pandemic with the Text Mining Technique. For this purpose, 40.000 twitter messages have sent regarding the Covid-19 outbreak were evaluated with the Text Mining Technique. As a result of the Text Mining analysis, it was determined that the tweets of the twitter users included in the study focused mainly on four words, Mask, Covid-19, Ankara Wuhan and the Scientific Committee, and their posts were shaped around these four words. At the same time, Sentiment Analysis was also performed after Text Mining analysis for four words that occupy the agendas of twitter users. Thus, it was possible to produce information about how the words Mask, Covid-19, Ankara Wuhan and Scientific Committee, which constitute an important place in the agenda of Twitter users, affect their emotions. As a result of the Sentiment Analysis, it was concluded that individuals especially hearing the word Covid-19 or mentioning this word led to have a negative emotion in themselves.

**Key words:** Covid-19, text mining, sentiment analysis, mask



## Meralarda Homojen Alanlarına Göre İndikatör Türlerin Belirlenmesinin İstatistiksel Açıdan Değerlendirilmesi

Kadir Aytaç ÖZAYDIN<sup>1</sup>, Arife AVAĞ<sup>2</sup>, Ufuk KARADAVUT<sup>3</sup>, Emre Süreyya DUMLU<sup>4</sup>, Sabahaddin  
ÜNAL<sup>5</sup>, Celalettin AYGÜN<sup>6</sup>

<sup>1</sup>Tarla Bitkileri Merkez Araştırma Enstitüsü Müdürlüğü, Ankara, Türkiye

<sup>2</sup>Bahçelievler İlçe Tarım Müdürlüğü, İstanbul, Türkiye

<sup>3</sup>Ahi Evran Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Kırşehir, Türkiye

<sup>4</sup>Doğu Anadolu Tarımsal Araştırma Enstitüsü Müdürlüğü, Erzurum, Türkiye

<sup>5</sup>Abant İzzet Baysal Üniv., Ziraat Fak., Tarla Bitkileri.Bölümü, Bolu, Türkiye

<sup>6</sup>Geçit Kuşağı Tarımsal Araştırma Enstitüsü Müdürlüğü, Eskişehir, Türkiye

### Özet

Meralar çok yönlü kullanım özelliğine sahip alanlardır. Ancak temel olarak hayvancılık sektörü tarafından otlatılarak değerlendirilmeleri söz konusudur. Meralar farklı bitki türlerini ve diğer canlı ırklarını bir arada bulundurarak biyoçeşitliliğin zenginleşmesine katkıda bulunur. Doğal veya yapay meralar doğru yönetildiği takdirde sürdürülebilir olarak daha fazla fayda sağlamaktadır. Hayvansal üretim faaliyetlerinde hayvan otlatma amaçlı olarak planlamalar yapılmaktadır. Mera kullanım ve yönetimi açısından ekolojiye uygun bitki türleri doğru tespit edilmelidir. 2007-2017 yılları arasında Tarım ve Orman Bakanlığı'na bağlı Tarımsal Araştırmalar ve Politikalar Genel Müdürlüğü (TAGEM) tarafından iki proje yapılmıştır. Projeler Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) ve Bitkisel Üretim Genel Müdürlüğü (BÜGEM) tarafından desteklenmiştir. Bu projeler ile Türkiye'de mera alanlarının özellikleri, homojen alanları ve mevcut farklı bitki türlerinin yayılma alanları belirlenmiştir. Meralarda ekolojiye uygun olan indikatör bitki türlerinin belirlenmesi amacıyla farklı modelleme çalışmaları yapılmıştır. İndikatör bitki türlerinin belirlenmesi bilgisayarda Excel ortamında sql veri tabanı ile bağlantılı olarak hazırlanan bir uygulama programıyla daha pratik ve daha kolay bir şekilde yapılabilmektedir. Bu analiz işlemi, seçilen bir bölge, iklim veya topoğrafyaya göre sınıflandırılmış bir homojen alanda yapılabilmektedir. Seçilen bölgenin ekolojisine en uygun ortalama 4 – 6 arasında farklı indikatör türler seçilerek planlanmakta, mera durum ve sağlık sınıfı iyi yapıda sürdürülebilir bir mera yönetimi uygulanabilmektedir.

**Anahtar kelimeler:** Homojen alan, indikatör tür, ISPD, mera, model, REDIS

## Statistical Evaluation of Determination of Indicator Species According to Their Homogeneous Areas in Rangelands

### Abstract

Rangelands are areas with versatile use. However, they are basically grazed and utilized by the livestock sector. Rangelands contribute to the enrichment of biodiversity by containing different plant species and

other living races together. Natural or artificial rangelands provide more benefits if managed correctly. In animal production activities, plans are made for grazing. In terms of rangeland use and management, plant species suitable for ecology should be determined correctly. Two projects were carried out by the General Directorate of Agricultural Research and Policies (GDAR) affiliated to the Ministry of Agriculture and Forestry between 2007-2017 years. This Projects has been supported by Scientific and Technological Research Council of Turkey (TUBITAK) and the General Directorate of Plant Production (TUGEM). This project was determined by the properties of rangeland fields in Turkey and spread areas of the different plant species. Different modelling studies were carried out in order to determine the indicator plant species suitable for ecology in rangelands. With an application program prepared in connection with the SQL database in Excel environment on the computer, it is more practical and easier to determine the indicator plant species. This analysis process can be performed in a homogeneous area classified according to a selected region, climate or topography. A sustainable rangeland management with good rangeland condition and health class in the selected region can be applied by selecting different indicator species between 4 and 6 on average for a region that best suits its ecology.

**Key words:** Homogeneous area, indicator species, ISPD, model, pasture, REDIS

## Giriş

Meralar küresel anlamda, otsu ve çalimsı bitkilerin oluşturduğu doğal veya yarı doğal vejetasyona sahip, otlatılarak değerlendirilen tabii kaynaklar olarak tarif edilebilir. Ülkemiz toplam alanın %16,8'ini oluşturan meraların (TÜİK, 2018), hemen her bölgede şiddetli bir bozulma sürecinden geçtiği bilinmektedir. Meralar sadece hayvanlara yem sağlayan alanlar olarak değil, biyolojik çeşitliliğin kaynağı, yaban hayatının barınağı, havza geliştirme, erozyonu önleme, biyolojik mücadele, organik tarım, rekreasyon ve dinlenme alanları ve kırsal mirasın korunması açısından çok yönlü olarak değerlendirilmektedir.

Doğal kaynaklar içerisinde insanoğluna değişik alanlarda pek çok hizmetler sağlayan meraların bitki örtüsündeki değişimlerin izlenmesi, yorumlanması ve sürdürülebilir kullanımı için indikatör (anahtar) bitki türlerinin belirlenmesi gerekmektedir. Anahtar türler meraların durumu ve trendi hakkında bilgi verirken diğer yandan uygulanacak mera ıslah ve amenajman metotlarının türler üzerindeki etkisini inceleme fırsatı vermektedir. Mera bitki örtülerindeki değişimin izlenmesi, bitki örtüsünün sürdürülebilir kullanımına yönelik kararların isabetli verilmesine katkı sağlamaktadır. Vejetasyon etütleri meranın durumunu, olumlu veya olumsuz yöndeki seyrini, kullanım durumunu, yem kalitesini, verimini, toprak ve su koruma özelliklerini belirlemede önemli olduğu bilinmektedir (Gökbulak 2003). Meralarda bozulmaya ait eşik değeri vejetasyonun artık geri döndürülemez olduğu noktadır. Bu eşik değeri çitle çevirme ve hayvanların meradan uzaklaştırılması, meranın otlatılmaması gibi normal idari tedbirlerle iyileşmenin artık mümkün olmayacağı bir noktadır. Ayrıca ağır otlatmanın yanında erozyon, toprak bozulması, iklim gibi bir çok etken bozulma sürecini etkilemektedir.

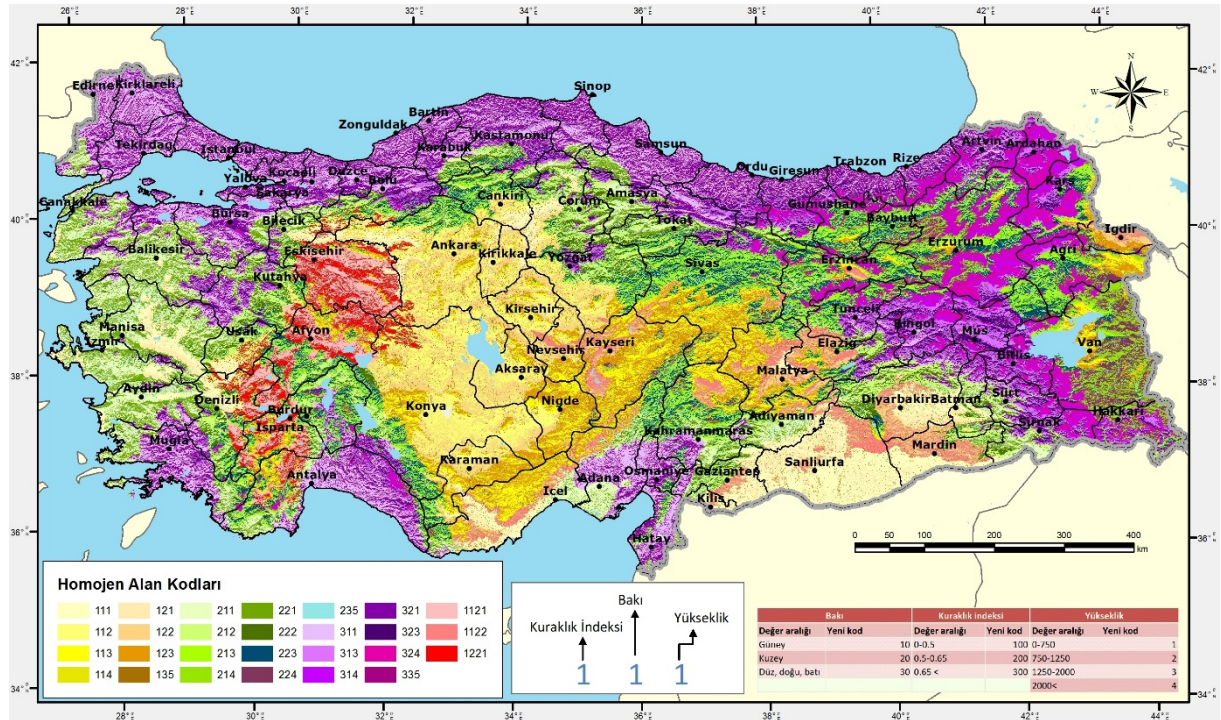
Tarımsal Araştırmalar ve Politikalar Genel Müdürlüğünde (TAGEM) ülkesel bazda mera etüt çalışmaları yapılmıştır. Bu sayede mera alanlarımızda bulunan bitki türleri tanımlanmış, sürdürülebilir mera yönetimi alt yapı çalışmaları hazırlanmıştır.

Mera durum sınıfının belirlenmesinde vejetasyondaki bitkilerin otlanmaya karşı tepkileri dikkate alınarak Azalıcı, Çoğalıcı ve İstilacı olmak üzere sınıflandırılmıştır (Koç vd., 2003). Bu sınıflar toplanarak bu türlerin kompozisyondaki oranlarına göre mera durum sınıfı belirlenebilir. Örneğin familya gurubu olarak; baklagil, buğdaygil ve diğer familya gurubu bitki türlerinden; *Trifolium pratense* Azalıcı Baklagil, *Dactylis glomerata* Azalıcı Buğdaygil, *Poa bulbosa* Çoğalıcı Buğdaygil, *Cynodon dactylon* Çoğalıcı Buğdaygil, *Centaurea iberica* İstilacı Diğer familya olarak sınıflanmıştır.

Azalıcı Bitkiler otlanmaya karşı en hassas olan klimaks türler olup otlatma baskısı arttıkça bu türlerin kompozisyondaki oranları devamlı olarak azalır. Bunlar aynı zamanda bitki örtülerinin en lezzetli türleri olup otlatma esnasında özellikle ve çoğunlukla tercih edilirler. Özellikle seçici otlatmanın varlığında, ideal mera kapasitelerinde dahi, otlatma baskısı azalıcı türler üzerinde yoğunlaşır. Otlatma şiddetinin artması ile birlikte vejetasyonda bu türler yavaş yavaş azalarak merada bozulma, yani klimaks durumdan uzaklaşma başlar.

Vejetasyondaki çoğalıcı bitkiler I. ve II. tip çoğalıcı bitkiler olarak ikiye ayrılırlar. I. Tip çoğalıcılar orta derecede lezzetli türlerdir, ikinci derecede tercih edilirler. Orta seviyeli otlatma şartlarında ya bir miktar artarlar yada mevcut durumlarını koruyabilirler. Otlatma baskısı arttıkça veya meralar orta duruma doğru ilerledikçe bu türler de azalmaya başlar. Klimaks vejetasyonda lezzetsiz olan II. Tip çoğalıcı bitkiler ise otlatma baskısı arttıkça veya arazinin bozulması durumunda artabilirler. İstilacı Türler; Bu türler klimaks topluluklarda bulunmazlar. Bitki örtüleri bozuldukça azalıcı ve çoğalıcı türlerden boşalan yerlere yerleşirler. Lezzetsiz ve yabancı ot özelliğindeki bu bitkiler orta veya zayıf mera durum sınıfına sahip meralarda bulunurlar.

TAGEM tarafından desteklenen ve bizim de aktif olarak araştırmacı olarak katıldığımız çalışmada mera ıslah yönetim çalışmasında Homojen alanları belirlenmesinde iklim ve topoğrafya verileri kullanılmış, kuraklık baki ve rakım değerlerine göre sınıflama yapılmıştır. Meteoroloji Genel Müdürlüğünden tüm Türkiye'ye için koordinatları bilinen iklim istasyonlarına ait uzun yıllar günlük iklim verileri alınmıştır. İstasyon verilerinden faydalanarak *Food and Agriculture Organization (FAO) FAO 56 Penman-Monteith* yöntemine göre (Allen vd., 1998) yıllık Potansiyel Evapotranspirasyon (referans bitki su tüketimi) değerleri hesaplanmıştır. Noktasal iklim değerleri topografik etkiyi dikkate alan enterpolasyon modelleri (Hutchinson, 1995) ile tüm alana yayılarak Coğrafi Bilgi Sistemleri (CBS) ortamında raster tabanlı iklim yüzey katmanları elde edilmiştir. Modelde 90 m çözünürlüklü Sayısal Yükseklik Modeli (SYM) verileri kullanılmıştır. Sayısal yükseklik verisinden yükseklik, baki, eğim haritaları oluşturulmuştur. Elde edilen bu veriler iklim parametrelerinden üretilen katmanlarla CBS ortamında modellenerek örnekleme yapılacak homojen alanlar (Şekil 1) belirlenmiştir (Avağ vd., 2012).



Şekil 1. Kuraklık, Bakım ve Rakım Faktörlerine Göre Türkiye'nin 27 Sınıflı Homojen Alanları

Ulusal Mera verileri etüd çalışmaları kapsamında botanik kompozisyonu belirlemek için (Koç ve Çakal, 2004) tarafından açıklanan *Modifiye Edilmiş Tekerlekli Nokta Metodu* kullanılmış ve vejetasyon etüdü meradaki hâkim bitkilerin çiçeklendiği dönemde doğu, batı, kuzey ve güney yöneyleri esas alınarak 4 hatta toplam 400 noktada bitki okunmak suretiyle yapılmıştır. Okuma yapılırken lup içerisinde okunan her bir bitkitürüne ait değerler toplam bitki sayısına oranlanarak türlerin botanik kompozisyondaki oranları tespit edilmiş olup içerisinde bitki düşmüyorsa ya da taş ise çıplak alan olarak değerlendirilmiştir. Mera kesimlerinde, bitki örtüsünün toprağı kaplama oranı (*TKO*) vejetasyon etüdü sırasında bitkiye rastlanan nokta sayısının ölçülen toplam nokta sayısına oranlanması ile belirlenmiştir (Gökkuş vd., 2000). Ölçümlerde bitki örtüsünün dip kaplama alanı esas alınmıştır. *Mera Kalite Derecesi (MKD)* hesabında bitki örtüsünün durum sınıfını ortaya koymak için azalıcı türlerin kompozisyondaki oranına ilaveten çoğaltıcıların hesaba katılma payları Koç ve ark. (2003) tarafından önerilmiştir. Hesaba Katılan Türlerin Oranı (%) sınıflamasında; Çok İyi (76–100), İyi (51–75), Orta (26–50) ve Zayıf (0–25) olarak sınıflanmıştır.

*The Resource and Environmental Data Interpretation System (REDIS)* programı kullanılarak mera etüdüleri modellenabilir. Modelin kuvvetli yönleri, ekolojiye uyarlanabilmesi, mera durum sınıfını belirleme ve ileriye dönük izleme değerlendirme çalışmalarındaki güvenilirliği ve pratikliğidir (Bosch ve Gauch, 1991). Modelin özü anahtar türlere dayanır ve her homojen alan için bir model ve her model için anahtar türler belirlenir. Anahtar türler bölgesel değişiklik göstermez.

Entegre Bitki Dinamikleri Sistemi *integrated system plant dynamics (ISPD)* sisteminin temelinde *Detrended Cosrrespondance Analysis (DCA)*, *Principle Component Analysis (PCA)*, *Korelasyon Analizi* ve *D-İstatistik* yöntemleri bulunmaktadır. Bozulma modeli inşası *ISPD* paket programı içerisinde *Degradation Model Construction (DMOC)* ile yapılmaktadır. Bu işlemle veriler x eksenini boyunca çevresel faktörlere göre kontrol edilmektedir. *DMOC*'ta verilerin test edilmesinin amacı, x eksenindeki bozulma eğilimini tanımlamak ve eğimin gerçekten bir bozulma eğilimi olduğunu test etmek ve doğrulamak bakımından gereklidir. Bunun için, çevresel veriler yine *DMOC* analizinin sıralama sonuçları üzerine eklenerek ve bozulma eğilimi doğrulandığında model kaydedilmektedir. Bozulma eğilimi grafiğinde y eksenini bozulma eğiliminden olan sapmaları ve toplam hata payını göstermektedir. Herhangi bir nokta gruptan çok fazla sapma göstermişse, o nokta değerlendirme dışı bırakılmaktadır.

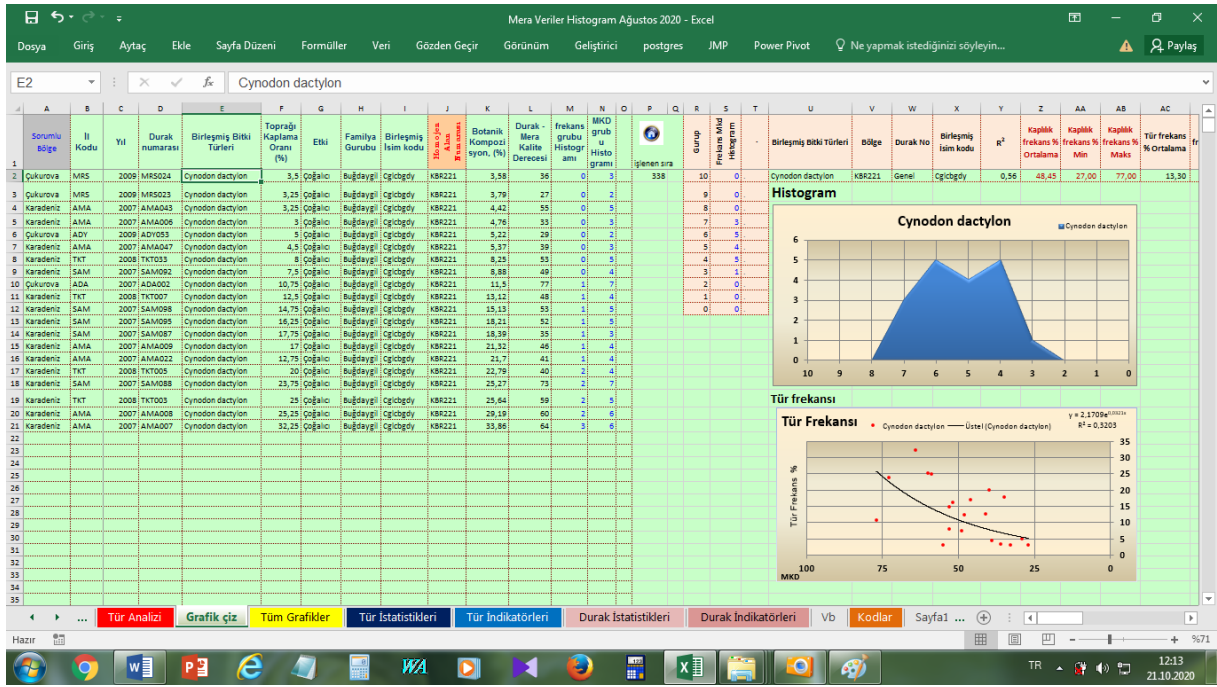
## Materyal ve Metot

Mera etüd verileri TAGEM mera projeleri veri tabanına aittir. Veriler *Excel* ve açık kaynak kodlu *Postgres* veri tabanında bulunmaktadır. İndikatör türleri belirlemek için özel olarak *Excel Visual Basic* Makro kodları hazırlanmıştır.

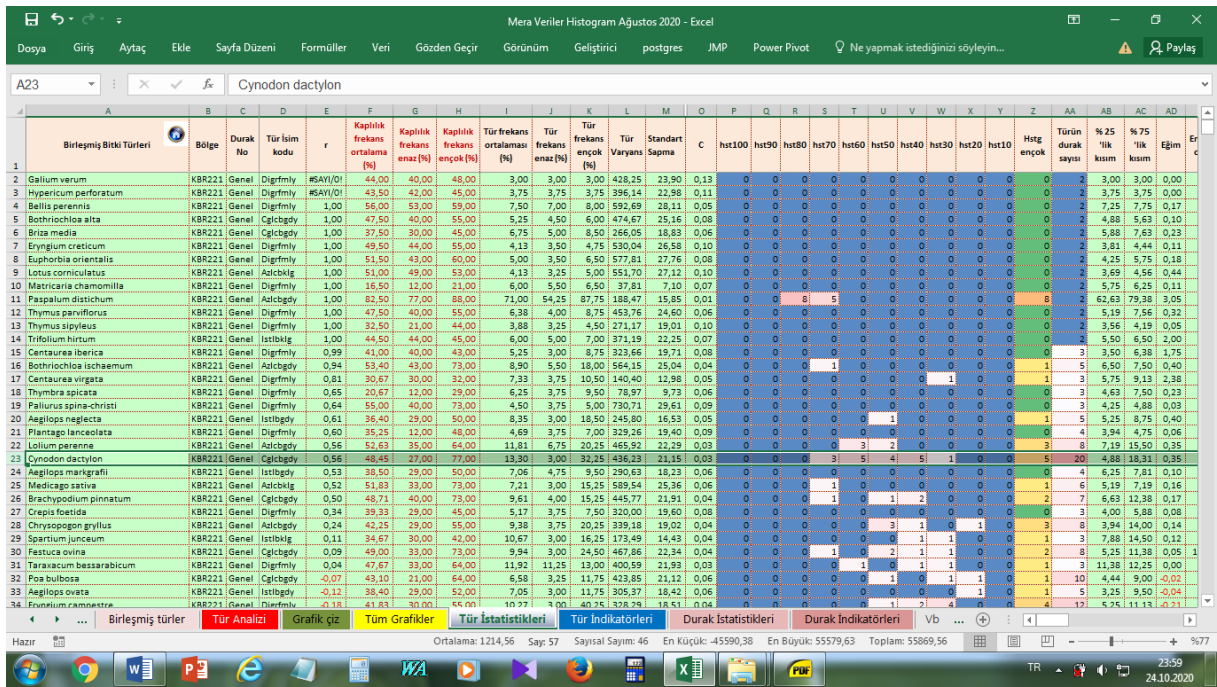
Mera verilerinde duraklara ait konum bilgileri, arazi özellikleri, rastlanılan bitki adları ve bunların frekanslarına ait bilgiler bulunmaktadır. Veri tabanındaki veriler kullanılarak indikatör türlerin analiz edildiği özel *excel* makro programı hazırlanmıştır. *Visual Basic* makroları ile *sql* veri tabanı veya *excel* içerisine kopyalanan durak -tür verileri ile çalışılabilmektedir (Şekil 2). Bitki türlerini tür bazında çalışmak mümkün olduğu gibi benzer özellik gösteren bitki türlerini gruplayarak ta çalışmalar yapılabilir. *Ulusal Mera Veri Tabanı* çalışmalarında bitki türlerinin farklı gruplama seçenekleri tasarlanmış ve bu seçenekler türler ile ilişkili olarak veri tabanına dahil edilmiştir. Örneğin isteğe bağlı olarak Azalıcı baklagiller, İstilaıcı baklagiller şeklinde basit gruplamalar olabileceği gibi, Buğdaygiller Kaliteli, Buğdaygiller Kalitesiz, Anthemis +Alyssum+Senecio+Ajuga +Stachys, Astragalus Dikenli, Dikenli Diğer familyalar, Zehirli Bitkiler, Çalımsılar, MedicagoTekyillik, Potentilla +Taraxacum +Ranunculus, Malva +Salvia +Nepeta, Poterium +Cichorium, Thymus +Origanum, TragopogonSpecies gibi birleşmiş bitki grupları çalışmalarda kullanılabilir. Veri tabanında etüdü yapılmış bir bölgeye ait (il, ilçe, köy, mevki, bir durak veya homojen alan vb. olabilir) durakların verileri kullanılarak indikatör tür



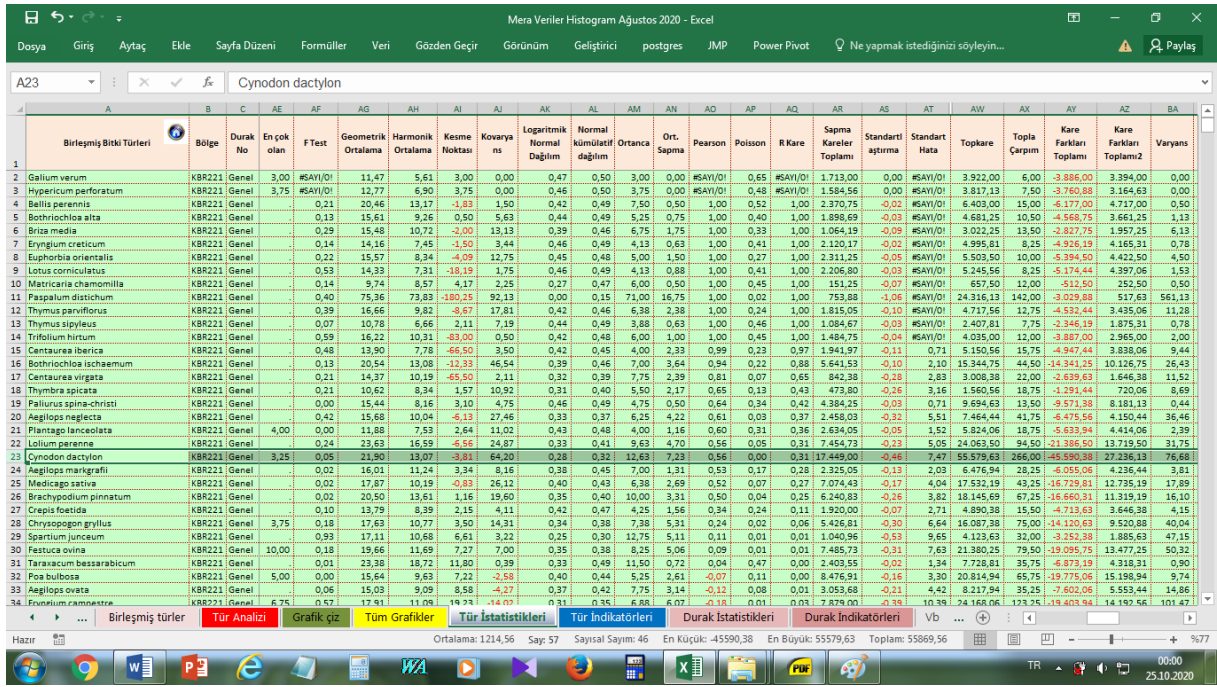
1<sup>st</sup> International Applied Statistics Conference (UYİK-2020)  
Online – Tokat / Turkey, 1-4 October 2020



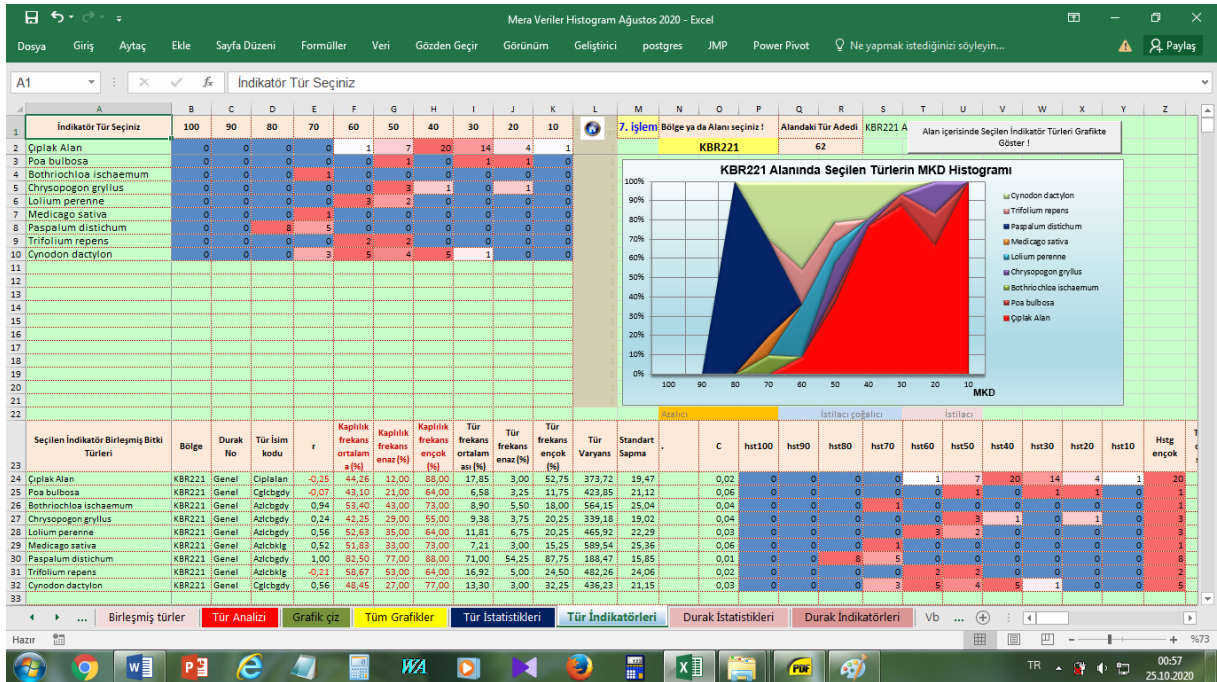
Şekil 3. KBR221 Alanındaki Her Bir Tür İçin Hesaplanan Farklı İstatistik Hesaplamaları ile Grafik Çizimleri Örneği



Şekil 4. KBR221 Alanındaki Türlerin Tanımlayıcı İstatistik Değerleri Örneği



Şekil 5. KBR221 Alanındaki Türlerin Diğer Tanımlayıcı İstatistik Değerleri Örneği



Şekil 6. KBR221 Alanında İndikatör Seçilen Türlerin Bazı Tanımlayıcı İstatistik Analizleri ve MKD Histogramı

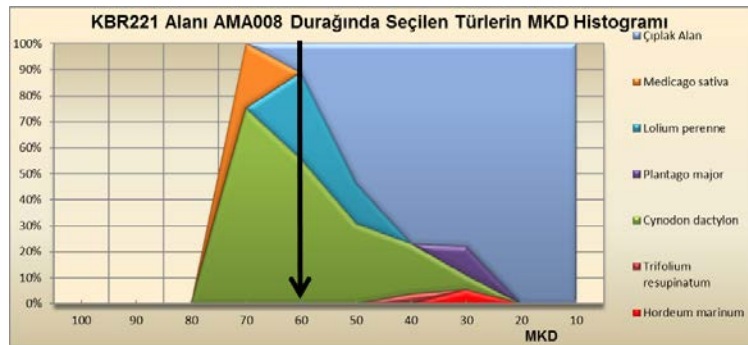
İki değişkenli istatistik analizleri 45 adet olup istenilen istatistik analizleri öne çıkarılarak analiz işlemlerini yapmak mümkündür. Bu analizler arasından önemli görülen birkaç tanesine göre verilerde sıralama yapılmaktadır. Öne çıkan istatistik analizi korelasyon (r) değeridir. Korelasyon değerinin  $\pm 1$  e en yakın değerler olması istenir. Standart sapma değerinin düşük olması, farklı durak sayısının sayısının (örnek, n) fazla olması istenir. Histogram verilerine göre MKD'nin her %10 luk sınıftaki tür frekansları incelenir. PCA analizi excelin eklentisi olup lisanslı yazılım modülleri olarak ulaşılabildiğinden dolayı düzenli olarak PCA analizleri yapılamamıştır. PCA analizi imkanlar dahilinde yapılabilmektedir. Özellikle iyi meralarda yüksek frekans değeri gösteren yem değerleri yüksek Azalıcı veya Çoğaltıcı türlerin olması istenir. Çalışılan bölge veya durak bazında incelenen türlerin farklı mera kalite derecelerinde dağılımları incelenerek aralarından en uygun türlerin seçimleri yapılır.

## Bulgular ve Tartışma

İlk aşamada alan bazında yapılan çalışmada Kuraklık Bakı Rakım' a göre 27 homojen alan arasından (Şekil 1) örnek olarak 221 nolu alan seçilmiştir. KBR221 nolu homojen alanda Adana, Adıyaman, Amasya, Çorum, Hatay, Kahramanmaraş, Samsun ve Tokat (8 il) illeri seçilmiş olup bu illerde 39 farklı durakta 251 adet farklı bitki türü bulunmaktadır. Bu farklı türlerden tür frekansı % 3'ün üzerinde olan 126 farklı tür seçilmiştir. Alanda MKD değeri 12 ile 88 arasındadır. Tüm türlerin duraklardaki frekans değerleri ile MKD değerleri arasındaki farklı istatistik verileri hesaplanmıştır. Örneğin 10 farklı türün korelasyon değerleri; *Çıplak Alan*: -0,25, *Poa bulbosa*: -0,07, *Bothriochloa ischaemum*: 0,94, *Chrysopogon gryllus*: 0,24, *Lolium perenne*: 0,56, *Medicago sativa*: 0,52, *Medicago minima* - 0,49, *Paspalum distichum*: 1, *Trifolium repens*: -0,21 ve *Cynodon dactylon*: 0,56 olarak bulunmuştur.

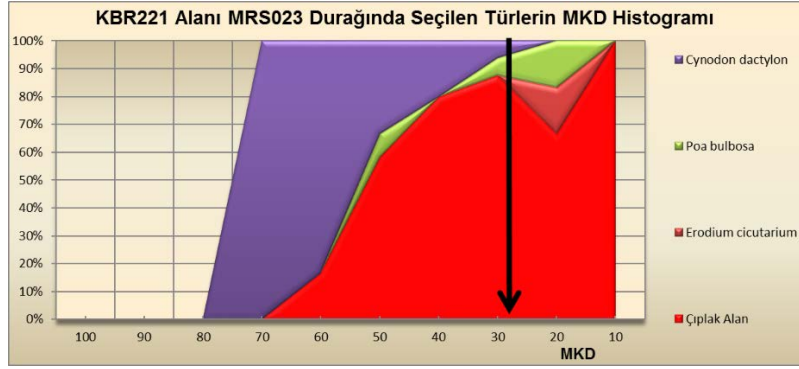
KBR221 nolu homojen alanında 39 durakta *Çıplak Alan* 'ın (Örnek sayısı, n= 39) korelasyon değerinin -0,25 olması *Çıplak Alan* ile MKD değerleri arasında eksi (-) yönde ilişki olduğu, çıplak alan arttıkça MKD değerinin azaldığı anlaşılmaktadır. Aralarındaki ilişkinin değeri 0,25 olarak bulunmuştur. Aynı homojen alanda *Cynodon dactylon* türünün 22 durakta (Örnek sayısı, n= 22) korelasyon değerinin 0,56 olması *Cynodon dactylon* türü ile MKD değerleri arasında eksi (+) yönde ilişki olduğu, tür frekansı arttıkça MKD değerinin de arttığı anlaşılmaktadır. Aralarındaki ilişkinin değeri 0,56 olarak bulunmuştur.

KBR221 nolu homojen alanda farklı türler arasından örnek olarak seçilen *Cynodon dactylon* 22 farklı durakta görülmüştür. Duraklardaki türün frekans değerleri ile durakların MKD değerleri arasındaki bazı istatistik değerleri: Korelasyon (r): 0,56, Kaplılık frekans ortalama (%): 48,45, Kaplılık frekans enaz (%): 27,00, Kaplılık frekans ençok (%): 77,00, Tür frekans ortalaması (%): 13,30, Tür frekans enaz (%): 3,00, Tür frekans ençok (%): 32,25, Tür Varyansı: 436,23, Standart Sapma: 21,15 olarak bulunmuştur. Türlerin alansal dağılımları *excel 3B harita uygulaması* modülü yardımı ile kontrol edilebilmektedir. Şekil 9 da koordinat bilgilerine göre türlerin alansal dağılımlarının incelenmesi görülmektedir.

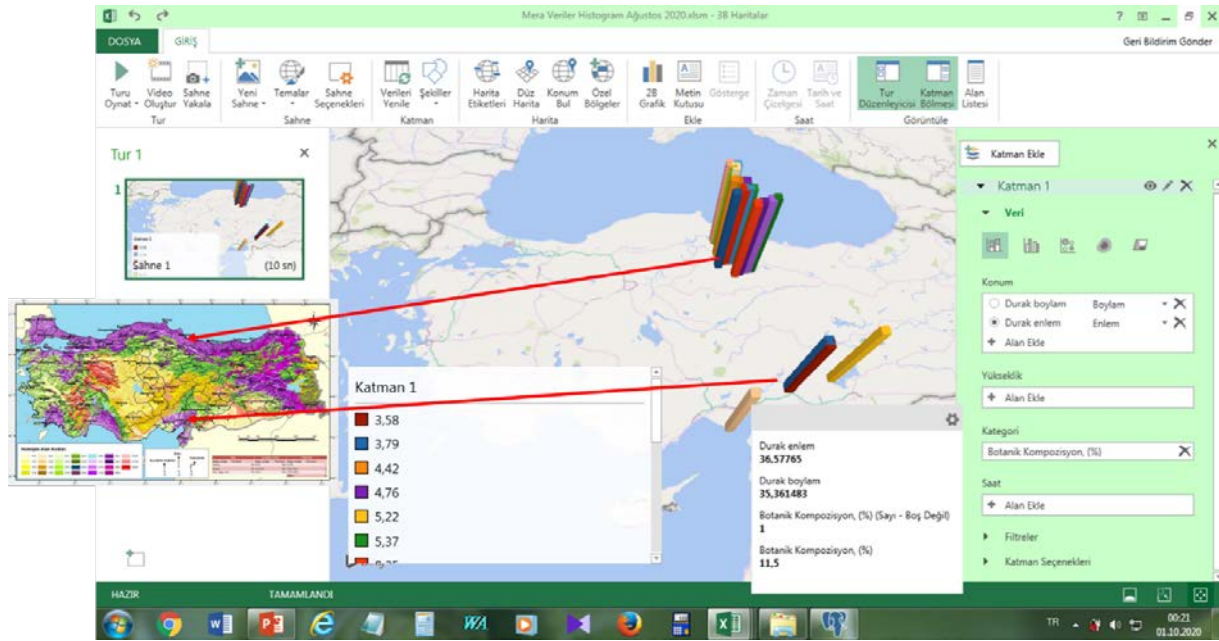


Şekil 7. KBR221 Alanı AMA008 Durağında Seçilen Türlerin MKD Histogramı ve Durağın MKD (60) değeri





Şekil 8. KBR221 Alanı MRS023 Durağında Seçilen Türlerin MKD Histogramı ve Durağın Mkd (27) Değeri



Şekil 9. KBR221 Nolu Homojen Alanda Seçilen *Cynodon Dactylon* Türünün Alansal Dağılımının İncelenmesi

İncelenen mera alanında vejetasyondaki değişim trendinin değerlendirilmesi ve izlenmesi için KBR221 nolu homojen alandaki 8 ilde seçilen *Lolium perenne*, *Cynodon dactylon*, *Chrysopogon gryllus*, *Festuca ovina* türleri indikatör olarak kullanılabilir (Şekil 6). Durak bazında; AMA008 durağında seçilen *Lolium perenne*, *Cynodon dactylon* ve *Medicago sativa* türleri indikatör olarak kullanılabilir (Şekil 7). MRS023 durağında seçilen *Cynodon dactylon* türü indikatör olarak kullanılabilir (Şekil 8).

ISPD, REDIS gibi programlar ile mera etüdlerinin değerlendirilmesi, indikatör türlerin belirlenmesi işlemleri eskiden kullanılan işletim sistemleri (DOS, Windows xp) aracılığı ile yapılmaktadır. Etüd verilerinin not defteri aracılığı ile hazırlanıp DAT dosyası ile veri girişlerinin yapılması önemli derecede zaman, ekip, maliyet ve deneyim isteyen işlemlerdir. İndikatör tür belirleme işlemleri için özel hazırlanan Postgresql veri tabanı altyapısı ve Excel makro programı sayesinde karşılaşılan sorunlar önemli ölçüde azalmıştır. Excel programında Visual Basic makrosu ile çalışan program yaygın etkinin artırılması amacıyla web üzerinden çalışacak şekilde altyapısı hazırlanmaktadır. Konu uzmanı teknik elemanların internet tarayıcısı ile HTML sayfaları üzerinden SQL veri sorgulamaları yapılarak indikatör tür çalışmaları yapılacak şekilde altyapı çalışmaları halen devam

etmektedir. Bu sayede standart mera veri tabanı alt yapısı üzerinden Türkiye'nin sürdürülebilir, sağlıklı mera kullanım ve yönetimi çalışmalarına pratik ve hızlı olarak katkıda bulunulacaktır.

### **Teşekkür**

Meteoroloji Genel Müdürlüğü (MGM), Tarımsal Araştırmalar ve Politikalar Genel Müdürlüğü'nce (TAGEM) TÜBİTAK tarafından Kamu Araştırmaları Destek Grubu (KAMAG) 1007 kapsamında desteklenen 106G017 nolu "Ulusal Mera Kullanım ve Yönetim Projesi" ve BÜGEM tarafından desteklenen "Mera Varlığının ve Mera Durum Sınıflarının Belirlenmesi Projesi" 'nde emeği geçen tüm personel ve danışmanlara teşekkür ederiz.

### **Kaynaklar**

- Allen R.G., Pereira, L.S., Raes, D., Smith, M., 1998, FAO, Potential Evapotranspiration, Crop evapotranspiration, Guidelines for computing crop water requirement, FAO irrigation and drainage paper 56, ROMA, (ISSN 0 24-5284), (ISBN 92-5-104219-5)
- Avağ A., Mermer A., Yıldız H., Ünal E., Urla Ö., Aydoğdu M., Dedoğlu F., Aydoğmuş O., Torunlar H., Tuğaç M. G., Ünal S., Mutlu Z., Özaydın K. A., Özgöz M. M., Aksakal E., Dumlu, E.S., Kara A., Uzun M., Çakal Ş., Yıldırım T., Aksoyak Ş., Tezel M., Aygün C., Kara İ., Erdoğan İ., Sever A. L., Atalay A., Yavuz T., Avcı M., Çınar S., İnal İ., Yücel C., Cebel H., Keçeci M., Başkan O., Depel G., Palta Ç., Çarkacı A., Karadavut U., Şimşek U., Sürmen M., Odabaşı G., Gül D., Koç A., Erkovan H. İ., Güllap K., Kendir H., Şahin N., 2012, Ulusal Mera Kullanım Ve Yönetim Projesi Sonuç Raporu, Gıda Tarım ve Hayvancılık Bakanlığı, Tarımsal Araştırmalar ve Politikalar Genel Müdürlüğü, Tübitak Proje No : 106G017, 483 sayfa
- Bosch, OJH, Gauch, HG, 1991, 'The use of degradation for the assessment and ecological interpretation of range condition', The grassland society of South Africa, vol. 8, no. 4, pp. 138-46.
- Gökbulak, F., 2003. Selected physical properties of heavily trampled soils livestock trails. İstanbul Üniversitesi Orman Fakültesi Dergisi, Seri: A, 53(1):39-40
- Gökkuş, A., Koç A., Çomaklı B., 2000. Çayır Mera Uygulama Kılavuzu. Atatürk Üni. Ziraat Fak. Ofset Tesisi, Basım Sayısı: 3, 157s.
- Hutchinson M.F. 1995. Interpolating mean rainfall using thin plate smoothing splines. Int. J. Geogr. Info. Systems 9: 385-403.
- Koç, A., M. Güven, B. Çomaklı, Ö. Menteşe ve A. Bilgili, 2003. Ardahan meralarında gübreleme ve otlatma sistemi uygulamalarının sığırlarda canlı ağırlık artışına etkisi. Türkiye 5. Tarla Bitkileri Kongresi, 13-17 Ekim 2003, Diyarbakır.
- Koç A., Çakal Ş., 2004. Comparison of some rangeland canopy coverage methods. Int. Soil Cong. On Natural Resource Manage. For Sust. Develp., June 7-10, 2004, Erzurum, Turkey, D7, 41-45.
- TÜİK, 2018. Tarım İstatistikleri. Türkiye İstatistik Kurumu, Ankara.

## Comparison of Some Different Clustering Methods in Double Dendrogram Heat Maps

Samet Hasan ABACI<sup>1</sup>, Yalçın TAHTALI<sup>2</sup>, Ahmet ŞEKEROĞLU<sup>3</sup>

<sup>1</sup>Ondokuz Mayıs University, Faculty of Agriculture, Department of Animal Science, 55139, Samsun, Turkey.

<sup>2</sup>Tokat Gaziosmanpaşa University, Faculty of Agriculture, Department of Animal Science, 60240 Tokat, Turkey

<sup>3</sup>Niğde Ömer Halisdemir University, Faculty of Agricultural Sciences and Technologies, Department of Animal Production and Technologies, 51240 Niğde, Turkey

### Abstract

Heat maps can be used to visualize large amounts of multidimensional data and define sets of rows of similar values. Combining the heat maps with hierarchical clustering based on the distance or similarity between them, the result of the clustering calculation is displayed in a heat map with a dendrogram, the tree structure of the hierarchy. Both rows and columns can be clustered in the heat map. In this study, it was aimed to compare some different clustering methods in double dendrogram heat maps for classification for some quality properties of eggs produced in different weeks. For this purpose, egg weight, shape index, albumen index, shell thickness, yolk index, Haugh unit, shell weight and surface area values of 150 eggs obtained from eggs produced between 24-42 weeks (biweekly) at Gaziosmanpaşa University research and application farm were used. Single Linkage, Complete Linkage, Simple Average, Group Average and Ward's minimum variance methods were used to comparing the correct classification performance in double dendrogram heat maps. NCSS package program was used for data analysis. According to the Cophenetic Correlation Coefficient (CP) values results that show the goodness of fit clustering methods, Ward's minimum variance method showed the worst fit (CP for rows:0.625 – CP for variables: 0.753), while the Group average method showed the best fit (CP for rows:0.861 – CP for variables: 0.759). According to the findings, it was determined that the Group Average method fits better for double dendrogram heat maps. For the rows, weeks 24 and 26 formed a cluster, while weeks 28 and 42 formed a cluster. For the variables, egg weight, shape index, yolk index, Haugh unit and surface area formed a cluster, while the albumen index, shell thickness and shell weight formed another cluster. As a result, it was also understood that a double dendrogram could be used for such data.

**Key words:** Cluster, dendogram, egg, egg quality, heat map

**Solvability of Stochastic Fractional Differential Equation with Q-Winer Processes**

Habib DJOURDEM<sup>1</sup>

<sup>1</sup>Laboratory of Fundamental and Applied Mathematics of Oran (LMFAO), University of Oran1,  
Ahmed Benbella, Algeria

**Abstract**

In this paper, we will apply stochastic analysis techniques with the help of Krasnoselskii's fixed point theorem to solving stochastic fractional differential equation driven by Q-Winer processes with linear bounded covariance operator Q.

**Key words:** Stochastic equation, mild solution, fractional derivative, Krasnoselskii's fixed point

**Solvability System of Stochastic Fractional Differential Equations with Multiplicative  
Random Noise in Probability Space**

Noureddine BOUTERAA<sup>1</sup>

<sup>1</sup>Laboratory of Fundamental and Applied Mathematics of Oran (LMFAO), University of Oran1,  
Ahmed Benbella, Algeria

**Abstract**

This paper is devoted to discuss the existence of solutions to system of stochastic fractional differential equations with multiplicative random noise in probability space under suitable assumptions based on the Leray-Schauder nonlinear alternative.

**Key words:** Stochastic equation, mild solution, fractional derivative, fixed point theorem

## Use of Biostatistics to Assess Performance of Balochistan Camel Breeds

Abdul FATIH<sup>1</sup>, Muhammad Masood TARIQ<sup>1</sup>, Masroor Ahmad BAJWA<sup>1</sup>,  
Majed RAFEEQ<sup>1</sup>, Asim FARAZ<sup>2</sup>, Abdul WAHEED<sup>2</sup>, Ecevit EYDURAN<sup>3</sup>, Illahi Bakhsh  
MARGHAZANI<sup>4</sup>, Jameel AHMED<sup>5</sup>, Mudassar JEHAN<sup>5</sup>

<sup>1</sup>Center for Advanced studies in Vaccinology and Biotechnology, University of Balochistan,  
Quetta, Pakistan

<sup>2</sup>Department of Livestock and Poultry Production, Bahauddin Zakariya University, Multan

<sup>3</sup>Faculty of Agriculture. Department of Animal Science, Biometry & Genetics Unit, Iğdır  
University, Iğdır, Turkey

<sup>4</sup>Lasbela University of Agriculture, Water and Marine Sciences, Uthal, Balochistan, Pakistan

<sup>5</sup>Livestock and Dairy Development Department Balochistan, Quetta, Pakistan

### Abstract

This study is aimed to evaluate the quantitative performance of camels found in Balochistan province in Pakistan. Camel is major livestock species of the province with 41% of total camel population of Pakistan. There are eight recognized breeds namely Brahvi, Kachhi, Kharani, Lassi, Makrani, Rodbari, Pishin and Kohi breeds. It was revealed that Kachhi breed was a riverine camel breed while others were mountainous camels. Rodbari breed had highest birth weight (48.58kg) among other breeds while the weaning weight (118.83kg) is high in Kharani breed and Adult weight (712kg) is high in Pishin breed. The milk yield per lactation (2049liter) is high in Kachhi breed and lactation length (579.18 days) is high in Brahvi breed. Age of puberty female camel (1282.6 days) and age of first breeding (1575.7 days) are delay in Brahvi breed. Gestation period (405.53 days) is high in Makrani breed and dry period (369.12 days) is long in Brahvi breed. Calving interval (787.65 days) is long in Rodbari breed. Age at puberty for male camel is long in Pishin breed and peak rutting vigor (128.54 days) is more in Kohi breed. Duration for copulation (25.38 mins) is high in Brahvi breed. The study is concluded that camel is most important species arid, semi-arid and desert areas like in Balochistan. The camel of this province can be improved with application of specific latest molecular techniques regarding selection breeding programs which can cover the delay maturity, low conception rate and long calving interval period.

**Key words:** Biostatistical data, camel, performance, pastoral, breeds, Balochistan

## Çekirdek Yoğunluk Tahmininde Kullanılan Bant Genişliği Seçim Yöntemleri

Şule KARAKOÇ<sup>1</sup>, Necla GÜNDÜZ<sup>1</sup>

<sup>1</sup>Gazi Üniversitesi Fen Fakültesi İstatistik Bölümü, Ankara, TURKEY

### Özet

Çekirdek yoğunluk tahmin yöntemi, parametrik olmayan yoğunluk tahmini konusunda yaygın olarak kullanılan bir yöntemdir. Bilinmeyen olasılık yoğunluk fonksiyonu için çekirdek tahmin edicisi ile ilgili rasgele değişkene ilişkin olasılık yoğunluk yada dağılım fonksiyonu tahmininde bulunulur. Veriye dayalı olarak uygulanan bu tahmin yönteminde, çekirdek tahmin edicisinin performansı, düzgünleştirme parametresi olarak da bilinen bant genişliği parametresine bağlıdır. Bir başka ifade ile, bant genişliği, olasılık yoğunluk fonksiyonu yada dağılım fonksiyonu tahmininde önemli bir rol oynamaktadır. Örneğin, küçük bant genişlikleri dağılımın kuyruklarında pürüzler oluştururken, büyük bant genişlikleri dağılımın kuyruklarında aşırı düzleşme meydana getirerek dağılımın özelliklerini görmemizi sağlamaz. Çekirdek düzgünleştirmesinde kullanılan çeşitli bant genişliği seçim yöntemleri vardır. Bu çalışmanın amacı, basit bant genişliği seçimi (normal ölçek), yerine koyma ve en küçük kareler çapraz doğrulama gibi en yaygın olarak kullanılan bant genişliği seçme yöntemlerini incelemektir.

**Anahtar kelimeler:** Çekirdek yoğunluk tahmini, bant genişliği seçimi, en küçük kareler çapraz doğrulama, yerine koyarak bant genişliği seçimi, normal ölçek bant genişliği seçimi

### Bandwidth Selection Methods in Kernel Density Estimation

#### Abstract

Kernel density estimation method is a widely used method for nonparametric density estimation. For the unknown probability density function, the probability density or distribution function of the related random variable is estimated by the kernel estimator. In this data-based estimation method, the performance of the kernel estimator depends on the bandwidth parameter, also known as the smoothing parameter. In other words, bandwidth plays an important role in the probability density function or distribution function estimation. For example, while small bandwidths create roughness in the tails of the distribution, large bandwidths do not allow us to see the properties of the distribution by creating excessive flattening in the tails of the distribution. There is various bandwidth selection methods used in kernel smoothing. The aim of this study is to examine most popular bandwidth selection methods such as the simple bandwidth selection (normal scale), plug-in and least square cross-validation.

**Key words:** Kernel density estimation, bandwidth selection, least squares cross-validation, plug-in bandwidth selection, normal scale bandwidth selection

## Nonparametric Estimation of Renewal Function under Progressively Censored Data

Ömer ALTINDAĞ<sup>1</sup>, Halil AYDOĞDU<sup>2</sup>

<sup>1</sup>Department of Statistics and Computer Sciences, Bilecik Şeyh Edebali University, Bilecik, Turkey

<sup>2</sup>Department of Statistics, Ankara University, Ankara, Turkey

### Abstract

Let  $(X_n)_{n=1,2,\dots}$  be a sequence of independent and identically distributed positive random variables representing successive failure times with common distribution function  $F$ . Define  $N(t) = \max\{n: S_n \leq t\}$ , for  $t \geq 0$  where  $S_0 = 0$ ,  $S_n = X_1 + X_2 + \dots + X_n$ ,  $n = 1, 2, \dots$ . Then the stochastic process  $\{N(t), t \geq 0\}$  is said to be a renewal process. Renewal process is a generalization of Poisson process and it gives number of successive events occurred in interval  $(0, t]$ . The renewal process is an important model utilized in stochastic modeling especially for reliability theory and warranty analysis. For application fields of renewal process see Finkelstein 2008, Blischke and Murthy 2011. Mean value function of the renewal process is crucial in many of applications. It is defined as  $M(t) = E(N(t))$ ,  $t \geq 0$ . Mean value function of the renewal process is also called as renewal function. It gives expected number of events occurred up to time  $t$ . Let  $X_1, X_2, \dots, X_n$  be a sample from a renewal process. Then the renewal function  $M(t)$  can be estimated for a fixed  $t \geq 0$  based on this sample. Frees 1986a and 1986b suggested some parametric and nonparametric estimators for the renewal function and investigated their statistical properties under complete sample. Lin 1988 studied estimation of renewal function when the data includes right censored observations. Later, Baxter and Li 1995 proposed nonparametric confidence interval for renewal function under right censored data. Estimation problem of renewal function is still being studied in the literature for different cases. For instance, see Alvarez 2006, Bebbington et al 2007 and Aminzadeh 2011. Recently, Altındağ 2017 have studied estimation problem of renewal function when the data is obtained under progressive censoring which is widely used in failure analysis to accelerate testing process. In this study, we consider nonparametric estimation of renewal function when the data is progressively censored. A plug-in estimator for renewal function is proposed and its statistical properties such as consistency, asymptotic unbiasedness and asymptotic normality are established. A simulation study is also performed to evaluate small sample behavior of the estimator.

**Key words:** Nonparametric estimation, progressive censoring, renewal function



## Data Visualization and Data Analysis Examples from Textile Engineering

Mehmet Erdem İNCE<sup>1</sup>

<sup>1</sup>Gaziantep University, Textile Engineering Department, 27310, Gaziantep, Turkey

### Abstract

In this proceeding, I would like to share the data visualization and data analysis examples that I use in my academic studies. In my experimental studies, since my input variable is generally a categorical variable and the output is a numerical one, the type of graph that I use frequently for data visualization is “box and whiskers plot”. The drawing of green diamonds showing the 95 percent confidence interval on the box and whiskers plot and the addition of comparison circles to the right side of the plot make the plot telling the whole story and standing alone. The story is completed by placing the table containing summary statistics and comparison results (p-values) after the box and whiskers plot. Another type of graph and analysis that I enjoy using are the histogram and “Goodness of Fit Test - Shapiro-Wilk W Test” tests that numerically analyze whether the given data show normal distribution or not. I also plan to share my experimental data on normal distribution analysis in this proceeding. To summarize, statistics is of great importance in the sensitive analysis and effective presentation of my data.

**Key words:** Data visualization, data analysis, box and whiskers plot, normal distribution, normality check.

### Özet

Bu bildiride akademik çalışmalarında kullandığım veri görselleştirme ve veri analizi örneklerini sizlerle paylaşmak istiyorum. Deneysel çalışmalarında genellikle girdi değişkenim kategorik, çıktı değişkenim ise sayısal bir değişken olmasından ötürü; veri görselleştirme işleminde sıkça kullandığım grafik tipi “kutu ve bıyıklar” grafiği olmaktadır. Kutu ve bıyıklar grafiği üzerine % 95 güven aralığını gösteren yeşil elmasların çizimi ve grafiğin sağ tarafına karşılaştırma dairelerinin eklenmesi grafiği tek başına ayakta durup tüm hikâyeyi anlatan hale getirmektedir. Kutu ve bıyıklar grafiğinden sonra özet istatistiğini ve karşılaştırma değerlerini (p-değerlerini) içeren tablonun eklenmesi ile de hikâye tamamlanmaktadır. Kullanmaktan haz ettiğim bir başka grafik ve analiz tipi ise; eldeki verilerin görsel ve numerik olarak normal dağılım gösterip göstermediğinin tespit edildiği, histogram çizimi ve “Uygunluk Testi – Shapiro-Wilk W Testi” dir. Bu bildiride normal dağılım analizi uyguladığım deneysel verilerimi de paylaşmayı planlıyorum. Özet olarak istatistik bilimi verilerimin hassas analizi ve etkin sunumunda önemli rol oynamaktadır.

**Anahtar kelimeler:** Veri görselleştirme, veri analizi, kutu ve bıyıklar grafiği, normal dağılım, normallik kontrolü

## **Kadının İş Hayatındaki Rolünün Çok Değişkenli İstatistiksel Yöntemler Kullanılarak Değerlendirilmesi**

Sinem ESER<sup>1</sup>, Doğan YILDIZ<sup>1</sup>

<sup>1</sup>Yıldız Teknik Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, İstanbul, Türkiye

### **Özet**

İnsan doğduğu andan itibaren toplumun ona yüklediği rollerle yaşamaktadır. Bu yüklenen roller özellikle bazı toplumlarda kadınlar için çok daha fazladır. Bu çalışmada, 'Kadının iş hayatındaki rolü' araştırılmıştır. Veri anket yöntemi kullanılarak elde edilmiştir. Anket kişilere internetten gönderilmiştir. İnternet kullanamayan kişilerle yüz-yüze anket yapılmıştır. Ankette 600 gözlem değerlendirmeye alınmıştır. Anket değerlendirilmeden önce giriş ve literatür bölümleriyle konu ele alınmıştır ve yapılan analizlerle ilgili bilgiler verilmiştir. Ankette yer alan 'kadınlar neden iş hayatında olmalı/olmamalı' açık uçlu sorularına alınan cevaplara yer verilmiştir. Ankete verilen cevapların grafiksel olarak gösterilmesinde Python yazılımı kullanılmıştır. 'Ankete katılan katılımcıların çoğuna göre kadınların iş hayatında yeri vardır ve bunun en önemli nedeni maddi özgürlüktür.' sonucuna ulaşılmıştır. Ayrıca ki-kare, faktör, regresyon analizleri yapılmıştır. Yapılan bu analizlerde SPSS paket programı kullanılmıştır. Sonuç olarak; Regresyon denkleminde göre, kadının iş hayatında olmasıyla ilgili olan deneyim ve performans değişkeninin değeri arttıkça ve kadınların çalışmasının nedenleri olan motivasyon değişkeninin değeri arttıkça aile hayatı düzeni artar.

**Anahtar kelimeler:** Faktör analizi, iş hayatı, kadın, maddi özgürlük, regresyon analizi

## **Evaluation of the Role of Women in Business Life Using Multivariate Statistical Methods**

### **Abstract**

From the moment a person is born, they have been living with the roles that society imposes on them. These attributed roles are much higher for women, especially in some societies. In this study, "The role of women in business life" was investigated. Data were obtained using the survey method. The survey was sent to the people on the internet. A face-to-face survey was conducted with people who could not use the internet. In the survey 600 observations were taken into consideration. Before the survey was evaluated, the subject was discussed with the introduction and literature sections and information about the analysis was given. The answers to the open-ended questions of "why women should / should not be in business" in the survey are included. The Python software was used in graphical data analysis. It is concluded that according to most of the participants, women have a role in business life and the most important reason is financial freedom. Also, chi-square, factor, regression analyses were done. In these analyses, the SPSS package program was used. As a result; According to the regression equation, family life order increases as the value of the experience and performance variable increases and the value of the motivation variable, which is the reason why women work, increases.

**Key words:** Factor analysis, business life, woman, financial freedom, regression analysis

## Giriş

Tarih boyunca insanları anlamlandırabilmek için bireyler; yaş, cinsiyet, statü, ekonomik güç ve ırk gibi çeşitli kategorilere yerleştirilerek varoluş süreçlerini tanımlamaya çalışmışlardır. Bu sınıflandırmalardan birisi de insan hayatı için önemli bir işleve sahip olan, onun davranış biçimlerini, yaşam standartlarını belirleyen cinsiyet olgusu içindeki “kadın” ve “erkek” kavramlarıdır (Akkaş, 2019).

Kadın veya erkek, günlük dildeki yaygın kullanımıyla hem bireyin biyolojik anlamda dişi veya er oluşunu, hem de toplumun bireye sunduğu roller sistemi dâhilinde anlam kazanan kadın veya erkek oluşu ifade eden iki terimdir. Bu terimlerde anlam kazanan biyolojik boyut ile biyolojik yapıda temellenen toplumsal boyut birbirlerinden çok farklı şeylerdir (Vatandaş, 2011).

Biyolojik bir kavram olan cinsiyet bireylerin doğumuyla başlayan süreci ifade eder. Yani bireylerin biyolojik anlamda kadın ya da erkek olarak doğmasıdır (Akkaş, 2019).

Toplumsal cinsiyet sosyo-kültürelendir, insan icadıdır, her toplumun kendi kültürleriyle bağlantılı olarak kadınlar ve erkeklere yüklediği rollerdir. Doğduğumuz andan itibaren toplumun belirlediği rollere göre yaşamaya başlıyoruz. Toplumsal cinsiyet kavramı, renklerden giysilere, jest ve mimiklerden davranışlara kadar yaşamın her alanında varlığını hissettiriyor. Toplumsal cinsiyet özellikle kadınlara roller yükler. Kadına yüklenen birinci rol, annelik ve eş rolüdür. Erkek güç, akıl, otorite, bağımsızlık, iktidar ile özdeşleştirilirken, kadın anlayışlı, uyumlu, sabırlı olarak özdeşleştirilir. Kadının öncelikleri evi ve kocası olarak görülür (Şentürk, 2019).

Kadınların toplumsal bağlamda baskı altında olmaları ve ayrımcılığa maruz kalmaları yazılı tarihin başından beri olmasına rağmen bunu değiştirmek için harekete geçmeleri daha ileri tarihlere dayanıyor (Kaçar, 2020).

On dokuzuncu yüzyıldan itibaren hızla gelişen endüstrileşme, geleneksel tarım toplumlarında ailenin ücretsiz işçisi olarak ağır iş yükü taşıyan kadına, eğitim görme ve ev dışında ücretli çalışma olanakları sağlamıştır. Eğitim düzeyinin yükselmesi ve toplumdaki işlevinin çeşitlenmesi ile kadın sosyal haklarını genişletmek için mücadeleye başlamıştır (Kuzgun ve Sevim, 2004).

Demokratik Türkiye Cumhuriyeti’nin gerekliliği ve Mustafa Kemal Atatürk’ün temennisi; kadınların yalnızca anne ve evlerinin kadını olmaması aksine tüm sosyal haklar ile donatılmış, erkekler ile eşit ve Türkiye’nin geleceği için emek harcayan çağdaş bireyler olmasıdır. Mustafa Kemal Atatürk, “Şuna inanmak lazımdır ki, dünya üzerinde gördüğünüz her şey kadının eseridir” sözleriyle kadının ne kadar değerli bir varlık olduğunu duyurmuştur (Kaçar, 2020).

Cumhuriyetin ilanından sonra gerçekleştirilen devrimler çerçevesinde Türk kadınına önemli toplumsal haklar verilmiştir. Seçme seçilme hakkı, yasa önünde erkekle eşit konumda olmak gibi haklar yanında 1936 İş Kanunu ile çalışma hayatında bazı düzenlemeler yapılmıştır. Türkiye’de 1950 yılından itibaren sanayileşme süreci geliştikçe çalışan kadınların oranı da artmıştır (Kuzgun ve Sevim, 2004).

Türkiye’de ailelerin yaşam biçimi çoğunlukla erkeklerin eve gelir getirdiği, tek gelirli aile modeline dayanmaktadır. Ama erkeklerin bazı nedenlerden (ekonomik kriz, iş kazaları, hastalık, ölüm vb.) dolayı işsiz kalacağı ve bu durumda aile gelirinin kimin sağlayacağı düşünülmemektedir. Kadınların da çalışmasıyla çift gelirli aile modeline geçilmesi, gelir düzeyini yükseltecektir, yoksulluk riskini azaltacaktır. Ayrıca kadınların yaratıcılıklarını, becerilerini ve yeteneklerini geliştirecektir. Birçok araştırmaya göre çalışan kadınların çoğunun elde ettikleri gelirleri kendileri için değil, daha çok aileleri için harcadıkları görülmektedir. Erkekler kazançlarını daha çok eğlence, alkol, sigara vb., kadınlar ise çocukların isteklerine, eğitimlerine ve sağlıklarına harcamaktadır. Böylece çalışan kadın sayısının artması genç kuşakların daha iyi eğitilmesine ve daha sağlıklı yetişmelerine olumlu etki yapmaktadır. Daha eğitilmiş genç kadın kuşakları yetiştikçe, çalışan kadın oranı da yükselecektir.

Çalışma hayatı kadınlara ekonomik özgürlük sağlamakla beraber özgüvenlerinin ve toplum saygınlıklarının artmasını da sağlar (Küçük, 2015). Ancak Türk toplumunda kadınların çalışma hayatı açısından istenilen düzeye ulaşamamış olmasındaki önemli etkenlerden birisi, yapılan tüm reformlara karşın toplumun baskısından kurtulamamış olmalarıdır (Kuzgun ve Sevim, 2004). Kadınların uzun çalışma saatlerine maruz bırakılması, düşük ücretle erkeklerin alternatifi olarak görülmesi, sigortasız çalıştırılarak kadın emeğinin sömürülmesi, çalışma yaşamı içerisinde işe alınmada eşitlik ilkesinin uygulanmaması, annelik ve aile yaşamının kadının çalışmasını kısıtlaması veya sınırlandırması, eşit işe eşit ücret ödenmemesi, eğitimsizliği bahane edilerek terfi ettirilmemesi, taciz, sosyal çevre faktörleri vb. olarak bu örnekleri artırmak mümkündür. Bu durumlar 21. Yüzyılda hala kadınların zorlu çalışma koşullarının yanı sıra sırf kadın olmaktan kaynaklanan birçok sorunla karşılaştığını göstermektedir (Küçük, 2015).

Bu çalışmada veri anket yöntemiyle elde edilmiştir. Ankette yer alan kadının toplumsal cinsiyet rolüyle alakalı ifadelerle ve kadın neden iş hayatında olmalı/olmamalı ifadeleriyle, kadının iş hayatındaki rolünü nasıl gördüklerini öğrenmek amaçlanmıştır.

### Literatür Taraması

Dünya’da eşitliklerle ilgili araştırmalara bakıldığında;

- 2018’de Dünya Ekonomik forumu tarafından 149 ülke arasında, kadınların ekonomiye katılımı, fırsat eşitliği, eğitim imkanları, sağlık ve siyasi gücü gibi kriterlerle cinsiyet eşitliği sıralaması yapılmış. Kadınlar için cinsiyet ayrımcılığının en az olduğu ilk 3 ülke: İzlanda, Norveç, İsveç. Türkiye bu sıralamada 130’uncu olarak sonlarda yer almış (Euronews, 2018).

- 2018 yılında yapılan araştırmada, LinkedIn ile yapılan iş birliğine dayanarak, dünya genelinde Yapay Zekâ profesyonellerinin sadece %22’sinin kadın olduğunu, %78’inin erkek olduğu görülmüş. Bu, son yıllarda sabit kalan ve şu anda olumlu bir gelecek eğilimi göstermeyen %72’lik bir cinsiyet farkını oluşturduğuna ulaşılmış (Dünya Ekonomi Forumu, 2018).

- 2019 yılı AB üye devletlerinin cinsiyet eşitliği sıralamasına göre AB’nin ortalama puanı 100 üzerinden 67,4 olarak hesaplanmış. AB ortalamasına göre erkeklerle aynı işi yapan kadınlar, erkeklere göre yılda %16 daha az para kazanıyor. 4 Kasım’dan sonrası yılın %16’sına denk geldiği için bugüne ‘**Eşit Ödeme Günü**’ (Equal Pay Day) adı verilmiş (Euronews, 2019).

Türkiye’de 15-65 yaş arası cinsiyetlere göre işgücü ve kazanç istatistiklerine bakıldığında;

- 2019 yılında işgücüne katılma oranları kadınlarda %38,7, erkeklerde %78,2’dir, istihdam oranı kadınlarda 32,2, erkeklerde 68,3’tür. 2019 yılında bile kadınların işgücüne katılma oranları ve istihdam oranları erkeklere göre çok düşüktür. İşsizlik oranları kadınlarda 16,8, erkeklerde ise 12,7’dir. Kadınların işgücüne katılma oranları erkeklerden daha az olmasına rağmen işsizlik oranları daha fazladır (TUİK, 2019a).

- Eğitim düzeyi arttıkça yıllık ortalama brüt kazanç artmıştır. Tüm eğitim seviyelerinde kadın için ortalama brüt kazanç erkeklerden düşüktür. Cinsiyetler arası ücret farkı en fazla meslek lisesi mezunlarındadır (TUİK, 2019b).

Kahramanmaraş, Ankara ve Diyarbakır şehirlerinde anket yöntemiyle yapılan araştırmalar incelendiğinde;

- Bu 3 araştırmaya ve bu araştırmaya göre; kadınların iş hayatında olmasının en önemli nedeni olarak ‘maddi özgürlük’ sonucuna varılmıştır.

- Kahramanmaraş ilinde yapılan anketten, girişim kararı almakta en önemli etkenin ekonomik özgürlük, işletmeyi kurarken karşılaştıkları en önemli sorun toplumun kadınlar hakkındaki değeri, işletmeyi kurduktan sonra karşılaşılan en büyük sorunun rol çatışması olduğu sonucuna ulaşmışlar (Soysal, 2010).
- Ankara ilinde yapılan anketten, iş hayatında da çalışma hayatında da karşılaştıkları sorunların sözlü tacize maruz kalmaları, cinsiyet ayrımı gibi durumlar olduğu sonucuna ulaşmışlar (Küçük, 2015).
- Diyarbakır ilinde yapılan anket, kadınların %53,8'i kadının iş hayatında yeri olmadığını söylemişler ve bunun nedeni olarak kadının asıl görevinin çocuk ve ev işleri olması gerektiğini söylemişler. İş hayatında olması gerektiğini düşünenler bunun nedeninin ekonomik özgürlük ve aile bütçesine katkıda bulunmak olduğunu söylemişler. İş hayatındaki kadın çalışan sayısının, erkek çalışan sayısına göre az olmasının nedenleri sorulduğunda aile büyüklerinin karşı çıkması en yüksek neden olmuş. İşverenlere yapılan ankette, işverenlerin %94,4'ü kadının iş hayatında yeri olduğunu düşünmüşler. Buradan işverenlerin kadınlara iş hayatında yer vermesine rağmen kadınların kendilerini o yerde göremedikleri düşünülebilir. Kadınların iş hayatında yeri olduğunu düşünen işverenler bunun nedeninin ekonomik özgürlük olduğu düşüncesini en fazla belirtmiş. Aynı konumdaki kadın maaşı ile erkek personel maaşı arasında fark olup olmadığını belirlemek amacıyla sorulan sorulara verilen cevaplara göre işletmelerin %77,3'ü eşit bir dağılım uygularken %22,7'si eşit ücret vermediği sonucuna ulaşılmış. Buradan kadınlara karşı adaletsizlik yapıldığına ulaşılmış (Anonim, 2013).

## **Materyaller ve Yöntemler**

### ***Materyal***

- Veri anket yöntemi kullanılarak elde edildi. Anket soruları 4 alt bölümden oluşturuldu.
- İlk bölümde kişisel bilgiler soruldu. İlk bölümün son sorusu olan 'Kadınlar iş hayatında olmalı mı?' sorusunun cevaplarına göre anket 2'ye ayrıldı.
- İkinci bölümde evet cevabı verenler için 3 soru soruldu. Bu soruların biri açık uçlu soru, diğer ikisi 5'er ifadeden oluşan sorudur. Bu ifadeler için düşünce seviyesinin belirlenmesi istendi. Cevap şıkları 5'li likert ölçeği olarak oluşturuldu.
- Üçüncü bölümde hayır cevabı verenler için 3 soru soruldu. Bu soruların biri açık uçlu soru, diğer ikisi 5'er ifadeden oluşan sorudur. Bu ifadeler için düşünce seviyesinin belirlenmesi istendi. Cevap şıkları 5'li likert ölçeği olarak oluşturuldu.
- Son olarak dördüncü bölümde verilen 12 ifade için düşüncelerini belirtmeleri istendi. Cevap şıkları 5'li likert ölçeği şeklinde oluşturuldu.
- Anket link şeklinde internet ortamından yollandı ve ayrıca yüz yüze anket yöntemi uygulandı. Yaklaşık eşit sayıda veri elde edildi.
- 600 kişiyle yapılan ankette kararsal örnekleme kullanıldı. Bu yüzden sonuçlarda sistematik hata olabileceği öngörüldü.

### ***Kullanılan analizler***

- Anket sorularının cevap dağılımları tablo ve grafiklerle gösterildi. Bu dağılımlar Python yazılımıyla elde edildi.
- Birçok ki-kare testi yapıldı.

Ki-Kare Testi: Çok kategorili 2 değişkenin oluşturduğu hücrelerdeki, gözlenen frekans ile beklenen frekansı karşılaştırarak, değişkenler arası ilişkiyi sorgulayan parametrik olmayan bir testtir (Yıldırım, 2016).

- Uygulamada güvenilirlik analizi, faktör analizi ve regresyon analizi SPSS programı ile yapıldı. Bu analizlerde likert ölçekli soruların uygun olanları kullanıldı.

**Güvenilirlik Analizi:** Güvenilirlik analizi, daha önceden belirlenmiş bir ölçek türüne göre hazırlanmış ankete verilen yanıtların tutarlılığını ölçer. Güvenilirlik analizi için kullanılan temel analiz Cronbach Alpha ( $\alpha$ ) değerinin bulunmasıdır. 0,7'den düşük  $\alpha$  değerleri anketin zayıf güvenilirliği olduğunu gösterir,  $\alpha > 0,8$  olması ise anketin yüksek güvenilirliğe sahip olduğunu göstermektedir (Ekonomi Analiz, 2013).

**Faktör Analizi:** Faktör analizi, aynı yapıyı ölçen çok sayıda değişkenden, daha az sayıda ve tanımlanabilir şekilde anlamlı değişkenler elde etmeye yarayan çok değişkenli bir istatistik yöntemidir (Büyüköztürk, 2002).

**Regresyon Analizi:** Aralarında sebep-sonuç ilişkisi bulunan 2 veya daha fazla değişken arasındaki ilişkiyi belirlemede ve bu ilişkiyi kullanarak konu ile ilgili tahminler ya da kestirimler yapmak amacıyla kullanılır (Gültekin, 2020).

## Bulgular

### Verinin incelenmesi ve Ki-Kare testleri

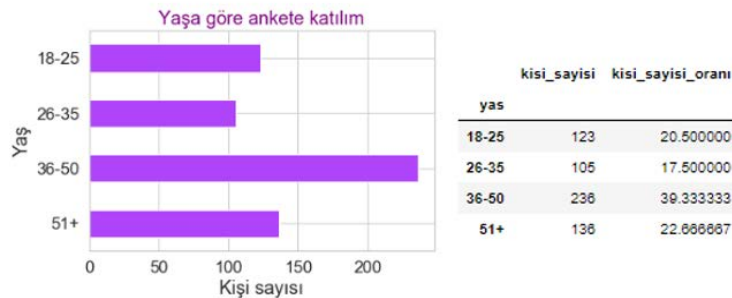
Ankette doğduğunuz yer, ikametiniz, annenizin ve babanızın memleketi sorularının cevapları şehirden, 7 bölgemiz ve yurtdışı olarak değiştirilmiştir ve yaş değişkeni kategorik hale getirilmiştir.

Ankete katılan kişilerin sorulara hangi cevapları hangi sayıda ve oranda verdiğine bakılmıştır ve verilen cevapların sayısı grafiklerle de gösterilmiştir. Bu incelemeler Python yazılımında yapılmıştır. Ankete katılan 600 kişinin 221(%36,83)'i erkek, 379(%63,16)'u kadındır.



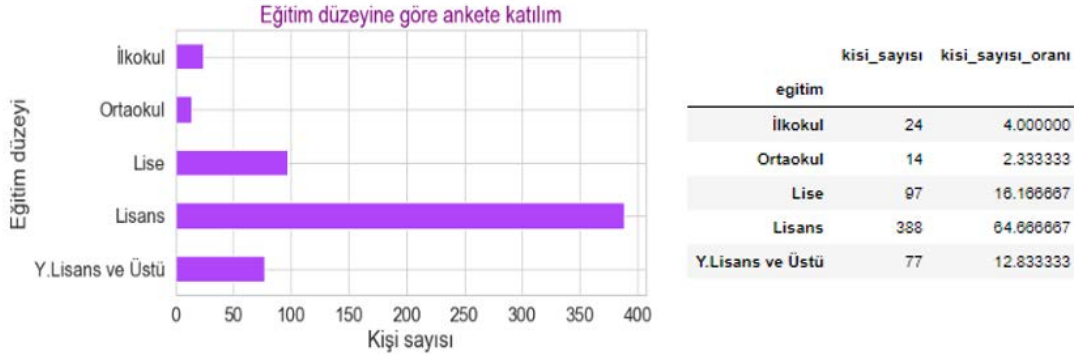
Şekil 1 Cinsiyete Göre Ankete Katılım

Ankete 18-85 yaş arası kişiler katılmıştır. Yaşa göre dağılımda en fazla katılım 36-50 yaş arası olmuştur.



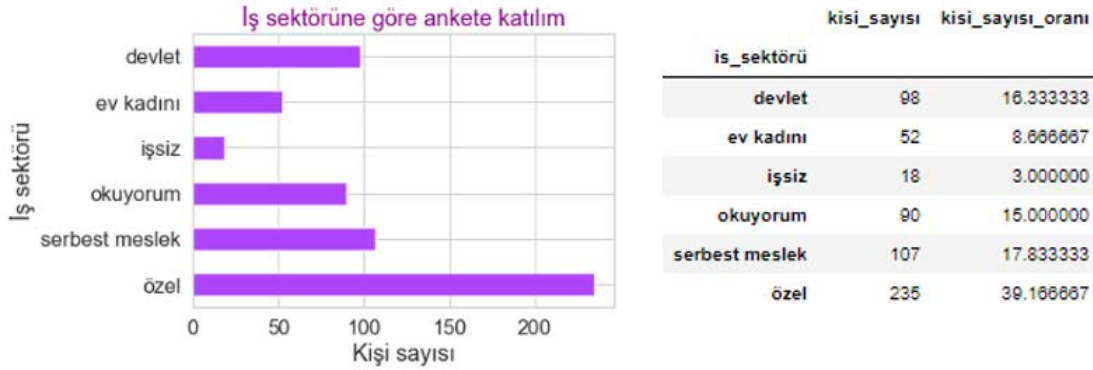
Şekil 2. Yaşa Göre Ankete Katılım

Eğitim düzeyine göre en fazla katılım lisans olmuştur.



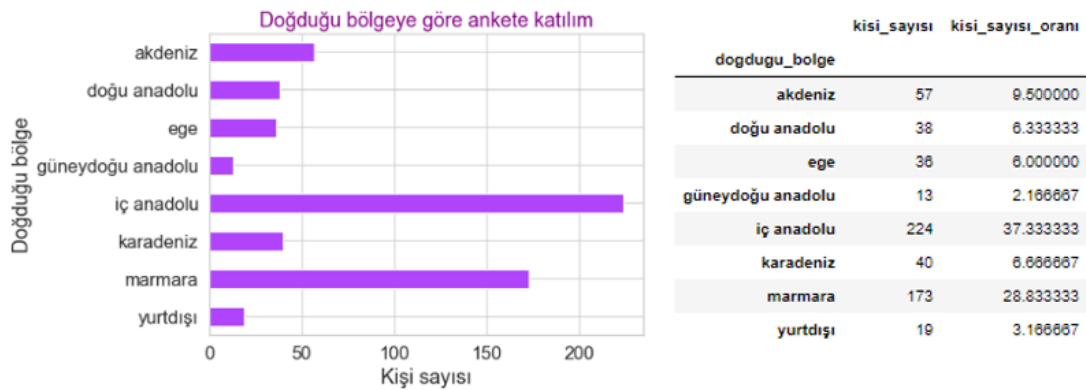
Şekil 3. Eğitim Düzeyine Göre Ankete Katılım

Ankete katılanların iş sektörü en fazla özel sektördür.

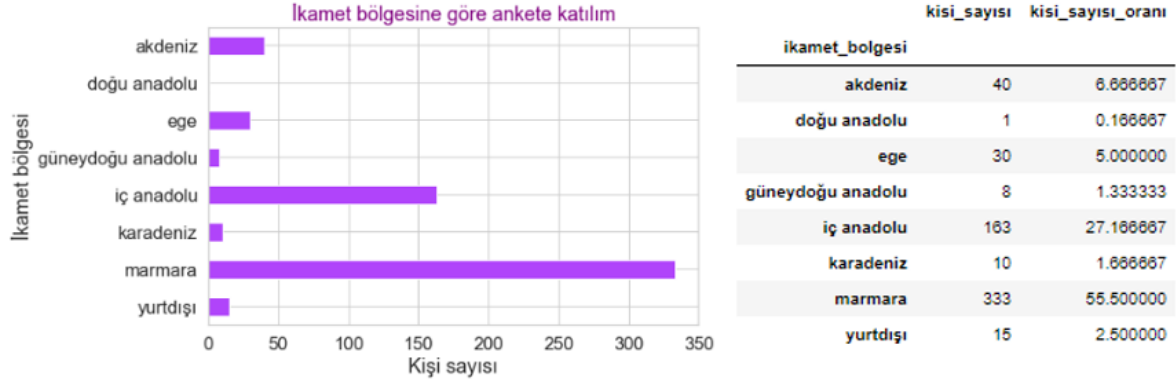


Şekil 4. İş Sektörüne Göre Ankete Katılım

Ankete katılanlardan İç Anadolu bölgesinde doğan sayısı fazladır ve büyük bir çoğunluğu Marmara bölgesinde ikamet etmektedir. Doğdukları bölgelere göre Marmara bölgesi dışındaki tüm bölgelerde azalma olmuştur. Marmara bölgesinde olan bu artışın sebebi İstanbul'da bulunan üniversitelerin daha çok tercih edilmesi ve İstanbul'da olan çeşitli iş imkânları olabilir.



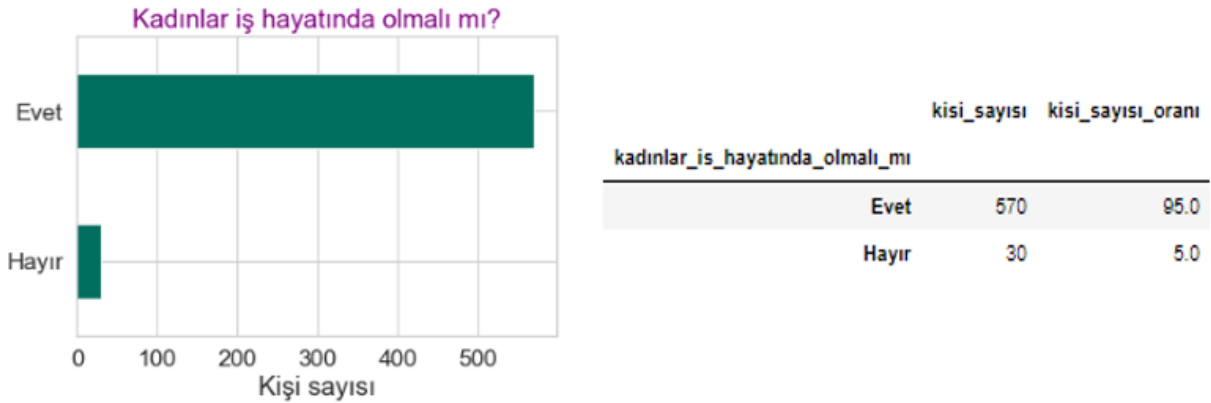
Şekil 5. Doğduğu Bölgeye Göre Ankete Katılım



Şekil 6. İkamet Bölgesine Göre Ankete Katılım

Anne ve babalarının memleketleri büyük çoğunluk İç Anadolu Bölgesidir. Annelerinin eğitim düzeyinin çoğunluğu ilkökuldür. 600 kişiden 348 kişinin yani %58'inin ilkökul mezunu olması iyi bir oran değildir. Ankete katılanların annelerin eğitim düzeyinde olduğu kadar olmasa bile babaların eğitim düzeyinde de en fazla ilkökul mezunu vardır. Annelerinin %75,9'u ev hanımıdır. Çalışan sayısı çok azdır. Babalarının çoğunluğu serbest meslekte çalışmaktadır.

'Sizce Kadınlar İş Hayatında Olmalı Mıdır?' sorusuna 570 (%95) kişi 'evet', 30 (%5) kişi 'hayır' cevabını vermiştir.



Şekil 7. Kadınlar İş Hayatında Olmalı Mı?

'Sizce Kadınlar İş Hayatında Olmalı Mıdır?' sorusuyla, cinsiyet ve eğitim düzeyi değişkenlerine ayrı ayrı ki-kare analizi yapılmıştır.

H<sub>0</sub>: 'Kadınlar iş hayatında olmalı mı?' sorusu değişkeniyle cinsiyet arasında anlamlı bir ilişki yoktur.

H<sub>1</sub>: 'Kadınlar iş hayatında olmalı mı?' sorusu değişkeniyle cinsiyet arasında anlamlı bir ilişki vardır.

2x2 lik çapraz tabloda yapılacak ki-kare testine karar verirken ki-kare testinin altında çıkan en küçük beklenen değere bakılır. Bu tabloda (5 < en küçük beklenen değer (11,05) < 25) olduğu için Yates' ki-kare teste bakılmıştır.

Yates' ki-kare tablosundaki anlamlılık değeri 0,004 < 0,05 olduğu için H<sub>0</sub> reddedilir. Yani değişkenler arasında anlamlı ilişki vardır. Araştırmaya katılanların 'Kadınlar iş hayatında olmalı mı?' sorusuna %95'i evet, %5'i hayır cevabını vermiştir. Bu soruya kadınların %2,9'u erkeklerin %8,6'sı hayır cevabı vermiştir. Hayır, cevabı çok olmasa da erkeklerin, kadınların iş hayatında olmasını istememe oranı kadınların istememe oranından fazladır.



Tablo 1. ‘Kadınlar İş Hayatında Olmalı Mı?’ ve Cinsiyet için Ki-Kare Testi

			kadınlarishayati		Toplam
			Evet	Hayır	
cinsiyet	Kadın	Miktar	368	11	379
		cinsiyete göre %	97.1%	2.9%	100.0%
	Erkek	Miktar	202	19	221
		cinsiyete göre %	91.4%	8.6%	100.0%
Toplam		Miktar	570	30	600
		cinsiyete göre %	95.0%	5.0%	100.0%

Ki-Kare Testi

	Değer	serbestlik derecesi(df)	Asimptotik Anlamlılık (Sig.) (2-taralı)	Kesin Anlamlılık (2-taralı)
Pearson Ki-Kare	9.531 <sup>a</sup>	1	.002	
Yates' Ki-Kare <sup>b</sup>	8.370	1	.004	
Olabilirlik Oranı	9.110	1	.003	
Fisher Kesin Olasılık Testi				.003
Geçerli N sayısı	600			

a. 0 hücre(.0%) 5'ten az beklenen değere sahiptir. En küçük beklenen değer 11.05'tir.  
b. Yalnızca 2x2 tablo için hesaplanır.

H<sub>0</sub>: ‘Kadınlar iş hayatında olmalı mı?’ sorusu değişkeniyle eğitim düzeyi arasında anlamlı bir ilişki yoktur.

H<sub>1</sub>: ‘Kadınlar iş hayatında olmalı mı?’ sorusu değişkeniyle eğitim düzeyi arasında anlamlı bir ilişki vardır.

Tablo 2. ‘Kadınlar İş Hayatında Olmalı Mı?’ ve Eğitim Düzeyi için Ki-Kare Testi

			kadınlarishayati		Toplam
			Evet	Hayır	
eğitim_yeni ve altı	ortaokul	Miktar	31	7	38
		eğitim düzeyine göre %	81.6%	18.4%	100.0%
	lise	Miktar	92	5	97
		eğitim düzeyine göre %	94.8%	5.2%	100.0%
	lisans ve lisans üstü	Miktar	447	18	465
		eğitim düzeyine göre %	96.1%	3.9%	100.0%
Toplam		Miktar	570	30	600
		eğitim düzeyine göre %	95.0%	5.0%	100.0%

Ki-Kare Testi

	Değer	serbestlik derecesi(df)	Asimptotik Anlamlılık (Sig.) (2-taralı)	Kesin Anlamlılık (2-taralı)
Pearson Ki-Kare	15.663 <sup>a</sup>	2	.000	.001
Olabilirlik Oranı	10.167	2	.006	.008
Fisher Kesin Olasılık Testi	11.166			.003
Geçerli N sayısı	600			

a. 2 hücre(33.3%) 5'ten az beklenen değere sahiptir. En küçük beklenen değer 1.9'dur.

Ki-Kare testi tablosunun altında %33,3 oranında 5'ten az beklenen değere sahip hücre olduğu sonucuna ulaşılmıştır. Bu oran %20'den fazla olduğu için exact (kesin) testi yapılmıştır. Ki-kare tablosundaki kesin anlamlılık değeri 0,001 < 0,05 olduğu için H<sub>0</sub> reddedilir. Yani değişkenler arasında anlamlı ilişki vardır.

Ankete katılanlardan eğitim düzeyi ortaokul ve altı olanların %18,4'ü, lise olanların %5,2'si, lisans ve lisansüstü olanların %3,9'u ‘Kadınlar iş hayatında olmalı mı?’ sorusuna hayır cevabı vermişlerdir. Yüzelere göre kadınların iş hayatında olmasını istemediğini belirten en fazla ortaokul ve

altı eğitim düzeyine sahip olanlardır. Yüzde olarak en az lisans ve lisansüstü çıkmış olsa da bunun sebebi ankete eğitim düzeyi lisans ve lisansüstü olan katılımın fazla olmasıdır.

Kişi sayısı olarak en fazla 18 kişiyle lisans ve lisansüstü eğitime düzeyine sahip olanlar, kadınların iş hayatında olmasını istememiştir.

‘Sizce Kadınlar İş Hayatında Olmalı Mıdır?’ sorusuna Evet cevabı verenler için;

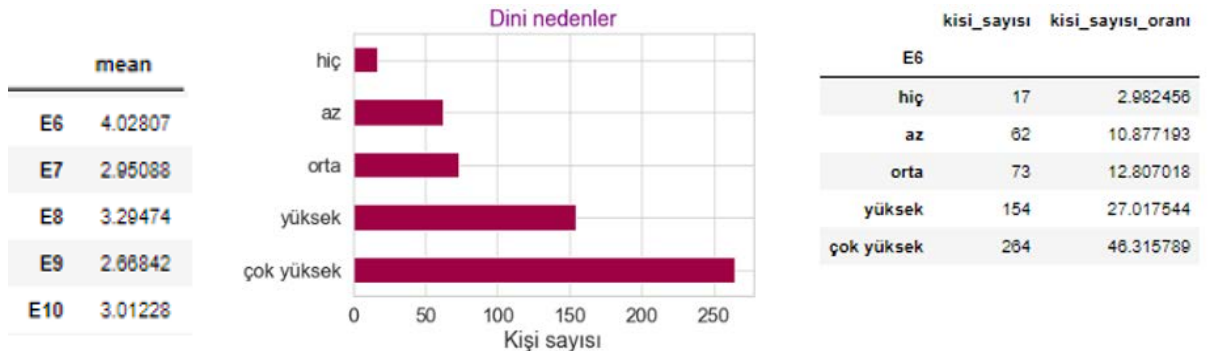
‘Aşağıda kadınların iş hayatında olmasının istenmesi ile ilgili bazı sebepler verilmiştir. Lütfen her bir sebep için kendi görüşünüze uygun ifadeyi seçiniz’, ‘Kadınların iş hayatında olmasını istemeniz ile ilgili yukarıdaki ifadeler dışında başka düşünceniz varsa yazabilir misiniz?’ ve ‘Sizce bazı insanlar neden kadınların iş hayatında olmasını istemiyor olabilir? Lütfen her bir neden için uygun ifadeyi seçiniz’ soruları sorulmuştur. İlk sorunun ifadeleri; ‘E1=Detayları anlamada iyi olmaları’, ‘E2=Çalışanlara moral verirler ve motive ederler’, ‘E3=Maddi özgürlüğe sahip olmalılar’, ‘E4=Çoğu konuda daha yeteneklidirler’ ve ‘E5=Geçim zorlukları’ şeklindedir. Genel olarak ifadeler kadınların iş hayatında olması için yüksek derecede sebep olarak görülmüştür. Özellikle ‘maddi özgürlüğe sahip olmalılar’ ifadesinin en önemli sebep olarak görüldüğü söylenebilir.



Şekil 8. Maddi Özgürlüğe Sahip Olmalılar

İkinci soruda 570 kişiden 135 cevap alınmıştır. Birbirine benzeyen ifadeler çıkarılmıştır ve 31 ifade açık uçlu sorular maddesine yazılmıştır. Örnek olarak; ‘2020 de bile hala kadınların çalışmasını isteyip istememeyi tartışıyor olmaktan utanıyorum. Kadının yeri ev değildir. Kadın kamusal alanda var olmalıdır.’. Çok doğru bir düşünce.

Üçüncü sorunun ifadeleri; ‘E6=Dini nedenler’, ‘E7=Karar almada duygusal davrandıklarını düşünmeleri’, ‘E8=Özel hayattaki sorumluluklarından dolayı işlerine odaklanamayacaklarını düşünmeleri’, ‘E9=Detaylarda kaybolacaklarını düşünmeleri’, ‘E10=Gerektiğinde sert olmakta zorlandıklarını düşünürler’ şeklindedir. ‘Dini nedenler’ ifadesi kadınların iş hayatında olmasının neden istenmiyor olabileceğinde en fazla etkiyi göstermiştir.



Şekil 9. Dini Nedenler

‘Sizce Kadınlar İş Hayatında Olmalı Mıdır?’ sorusuna Hayır cevabı verenler için;

‘Aşağıda kadınların iş hayatında olmasının istenmemesi ile ilgili bazı sebepler verilmiştir. Lütfen her bir sebep için kendi görüşünüze uygun ifadeyi seçiniz.’, ‘Kadınların iş hayatında olmasını istememeniz ile ilgili yukardaki ifadeler dışında başka düşünceniz varsa yazabilir misiniz?’ ve ‘Sizce bazı insanlar neden kadınların iş hayatında olmasını istiyor olabilir? Lütfen her bir neden için uygun ifadeyi seçiniz.’ soruları sorulmuştur. İlk sorunun ifadeleri; ‘H1=Dini nedenler’, ‘H2=Karar almada duygusal davranmaları’, ‘H3=Özel hayattaki sorumluluklarından dolayı işlerine odaklanamazlar’, ‘H4=Detaylarda kaybolurlar’, ‘H5= Gerektiğinde sert olmakta zorlanırlar’ şeklindedir. Ortalamalara göre en az sebep olarak ‘Dini nedenler’ ifadesine ulaşılmıştır.



Şekil 10. Dini Nedenler

İkinci soruda 30 kişiden 7 cevap alınmıştır. Örnek olarak; ‘Mantıksız kararlar almaları.’. Bazı kişilerin kadınlar hakkında böyle düşünmesi üzücü.

Üçüncü sorunun ifadeleri; ‘H6=Detayları anlamada iyi olduklarını düşünürler’, ‘H7=Çalışanlara moral verdiklerini ve onları motive ettiklerini düşünürler’, ‘H8=Maddi özgürlüğe sahip olmaları gerektiğini düşünürler’, ‘H9=Çoğu konuda daha yetenekli olduklarını düşünürler’, ‘H10=Geçim zorlukları yaşadıkları için’ şeklindedir. En yüksek ortalama ‘Maddi özgürlüğe sahip olmaları gerektiğini düşünürler’ ifadesidir.



Şekil 11. Maddi Özgürlük

Evet-hayır sorularının içeriklerinin aynı olduğunu yukarıdaki soru ifadelerinden görüyoruz. Bundan sonraki analizlere evet-hayır cevapları birleştirilerek devam edilmiştir. Yeni değişkenler B1(E1-H6), B2(E2-H7), B3(E3-H8), B4(E4-H9), B5(E5-H10), B6(E6-H1), B7(E7-H2), B8(E8-H3), B9(E9-H4), B10(E10-H5) şeklindedir.

Yeni oluşan tüm değişkenlerin cinsiyet ve eğitim düzeyi değişkenleriyle ki-kare testleri yapılmıştır. Analizlerden önce yeni oluşan değişkenlerin kategori sayıları ve eğitim düzeyi kategori sayısı 5'ten 3'e düşürülmüştür. Analizler sonunda değişkenler arası fark olanların çıktıklarına yer verilmiştir ve yorumlanmıştır.

Bu kısımda yapılan ki-kare testlerine örnek olarak;

‘Kadınların neden iş hayatında olması isteniyor?’ sorusu ‘Detayları anlamada iyiler(B1)’ ifadesinin, cinsiyet değişkeniyle Ki-Kare Testi;

Tablo 3. Cinsiyet ve Maddi Özgürlük (B3) İfadesi için Ki-Kare Testi

			‘Kadınların neden iş hayatında olması isteniyor?’ Maddi özgürlük			Toplam
			hiç ve az	orta	yüksek ve çok yüksek	
cinsiyet	Kadın	Miktar	5	29	345	379
		cinsiyete göre %	1.3%	7.7%	91.0%	100.0%
	Erkek	Miktar	15	22	184	221
		cinsiyete göre %	6.8%	10.0%	83.3%	100.0%
Toplam		Miktar	20	51	529	600
		cinsiyete göre %	3.3%	8.5%	88.2%	100.0%

**Ki-Kare Testi**

	Değer	serbestlik derecesi(df)	Asimptotik Anlamlılık(Sig.) (2-tarafli)
Pearson Ki-Kare	14.349 <sup>a</sup>	2	.001
Olabilirlik Oranı	13.881	2	.001
Geçerli N sayısı	600		

a. 0 hücre (.0%) 5'ten az beklenen değere sahiptir. En küçük beklenen değer 7.37'dir.

H<sub>0</sub>: Cinsiyet ile B3 değişkeni arasında anlamlı ilişki yoktur.

H<sub>1</sub>: Cinsiyet ile B3 değişkeni arasında anlamlı ilişki vardır.

Ki kare tablosundaki (Tablo 3) anlamlılık değeri 0,001 <0,05 olduğu için H<sub>0</sub> reddedilir. Yani değişkenler arasında anlamlı fark vardır.

Kadınların %91'i erkeklerin %83,3'ü ‘Maddi özgürlük(B3)’ ifadesine yüksek seviyede katılmıştır. Kadınların bu düşünceye erkeklerden daha fazla katıldığı sonucuna ulaşılır.

‘Kadınların neden iş hayatında olması isteniyor?’ sorusu ‘Detayları anlamada iyiler(B1)’ ifadesinin, eğitim düzeyi değişkeniyle Ki- Kare Testi;

H<sub>0</sub>: Eğitim düzeyi ile B3 değişkeni arasında ilişki yoktur.

H<sub>1</sub>: Eğitim düzeyi ile B3 değişkeni arasında anlamlı ilişki vardır.

Tablo 4. Eğitim Düzeyi ve Maddi Özgürlük (B3) İfadesi için Ki-Kare Testi

			‘Kadınların neden iş hayatında olması isteniyor?’ Maddi özgürlük			Toplam
			hiç ve az	orta	yüksek ve çok yüksek	
eğitim_yeni	ortaokul ve altı	Miktar	7	7	24	38
		eğitime göre %	18.4%	18.4%	63.2%	100.0%
	lise	Miktar	2	10	85	97
		eğitime göre %	2.1%	10.3%	87.6%	100.0%
	lisans ve lisans üstü	Miktar	11	34	420	465
		eğitime göre %	2.4%	7.3%	90.3%	100.0%
Toplam		Miktar	20	51	529	600
		eğitime göre %	3.3%	8.5%	88.2%	100.0%

**Ki-Kare Testi**

	Değer	serbestlik derecesi(df)	Asimptotik Anlamlılık (Sig.) (2-tarafli)	Kesin Anlamlılık(2-tarafli)
Pearson Ki-Kare	36.218 <sup>a</sup>	4	.000	.000
Olabilirlik Oranı	22.156	4	.000	.000
Fisher Kesin Olasılık Testi	23.701			.000
Geçerli N sayısı	600			

a. 3 hücre (33.3%) 5'ten az beklenen değere sahiptir. En küçük beklenen değer 1.27'dir.

Ki-Kare testi tablosunun (Tablo 4) altında %33,3 oranında 5’den az beklenen değere sahip hücre olduğu sonucuna ulaşılmıştır. Bu oran %20’den fazla olduğu için exact (kesin) testi yapılmıştır. Ki-kare tablosundaki kesin anlamlılık değeri  $0,0 < 0,05$  olduğu için  $H_0$  reddedilir. Yani değişkenler arasında anlamlı ilişki vardır.

Ortaokul ve altı eğitime sahip olanların %63,2’si, lise eğitimine sahip olanların %87,6’sı, lisans ve lisansüstü eğitime sahip olanların %90,3’ü ‘Maddi özgürlük(B3)’ ifadesine yüksek seviyede katılmıştır. Lisans ve lisansüstü eğitime sahip olanların bu düşünceye daha fazla katıldığı sonucuna ulaşılar.

‘Kadınların neden iş hayatında olması istenmiyor?’ sorusu ‘Dini nedenler(B6)’ ifadesinin, eğitim düzeyi değişkeniyle Ki- Kare Testi;

$H_0$ : Eğitim düzeyi ile B6 değişkeni arasında anlamlı ilişki yoktur.

$H_1$ : Eğitim düzeyi ile B6 değişkeni arasında anlamlı ilişki vardır.

Tablo 5. Eğitim Düzeyi ve Dini Nedenler(B6) İfadesi için Ki-Kare Testi

			‘Kadınların neden iş hayatında olması istenmiyor?’ Dini nedenler			Toplam
			hiç ve az	orta	yüksek ve çok yüksek	
eğitim_yeni	ortaokul ve altı	Miktar	7	13	18	38
		eğitime göre %	18.4%	34.2%	47.4%	100.0%
lise		Miktar	24	15	58	97
		eğitime göre %	24.7%	15.5%	59.8%	100.0%
lisans ve lisans üstü		Miktar	58	49	358	465
		eğitime göre %	12.5%	10.5%	77.0%	100.0%
Toplam		Miktar	89	77	434	600
		eğitime göre %	14.8%	12.8%	72.3%	100.0%

#### Ki-Kare Testi

	Değer	serbestlik derecesi(df)	Asimptotik Anlamlılık(Sig.) (2-tarafli)
Pearson Ki-Kare	31.237 <sup>a</sup>	4	.000
Olabilirlik Oranı	26.600	4	.000
Geçerli N sayısı	600		

a. 1 hücre (11.1%) 5'ten az beklenen değere sahiptir. En küçük beklenen değer 4.88'dir.

Ki kare tablosundaki (Tablo 5) anlamlılık değeri  $0,00 < 0,05$  olduğu için  $H_0$  reddedilir. Yani değişkenler arasında anlamlı ilişki vardır.

Ortaokul ve altı eğitime sahip olanların %47,4’ü, lise eğitimine sahip olanların %59,8’i, lisans ve lisansüstü eğitime sahip olanların %77’si ‘Dini nedenler(B6)’ ifadesine yüksek seviyede katılmıştır. Lisans ve lisansüstü eğitime sahip olanların bu düşünceye daha fazla katıldığı sonucuna ulaşılar.

Anketin son bölümünde olan ‘Lütfen aşağıdaki her bir ifade için sizin görüşünüze en uygun olan seçeneği işaretleyiniz.’ Sorusu için 12 ifade vardır. Bu ifadeler, ‘K1 = Çalışan bir anne çocukları ile çalışmayan bir annede olduğu kadar sıcak ve güvenli bir ilişki kurabilir.’, ‘K2 = Bir kadının, kocasının kariyerine destek olması, kendisinin kariyer sahibi olmasından daha önemli değildir.’, ‘K3 = İş hayatında erkekler, kadınlardan daha iyi bir yönetici değildir.’, ‘K4 = Evlendikten ve çocuk sahibi olduktan sonra kadının davranışlarını değiştirmesi ve daha geleneksel roller üstlenmesi zorunlu değildir.’, ‘K5 = Ev işleri ve çocuk bakımı bir kadının en temel sorumluluğu değildir, istenirse başkasına bırakılabilir.’, ‘K6 = Eğer annesi çalışıyorsa küçük yaşta bir çocuğun sorun yaşama olasılığı yüksek değildir.’, ‘K7 = Erkeğin işi para kazanmak, kadının işi ev işi ve aile ile ilgilenmek değildir.’, ‘K8 = Kadınların çalışma hayatına girmesi, erkekler arasında işsizliğe yol açmaz.’, ‘K9 = Kadınlar yüksek öğrenim görmeliler ve isterlerse çalışmalılar.’, ‘K10 = Çalışan kadının aile hayatı düzensiz değildir.’, ‘K11 = Kadınların matematiksel ve mekanik konularda erkeklerden daha az yetenekli oldukları söylenemez.’, ‘K12 = Genelde çalışan kişiler patronlarının kadın olmasından rahatsız olmazlar.’

şeklinde (Berkman, 2005; Kuzgun ve Sevim, 2004). İfadelere verilen cevapların ortalamalarına göre ifadeler katılmışlardır. Sadece ‘Genelde çalışan kişiler patronlarının kadın olmasından rahatsız olmazlar’ ifadesi için kararsız kalınmıştır. Bu ifadelerin grup sayısı 3’e düşürülerek cinsiyet, yaş ve eğitim düzeyi değişkenleriyle ki-kare analizi yapılmıştır.

Bu kısımda yapılan ki-kare testlerine örnek olarak;

‘Çalışan bir anne çocukları ile çalışmayan bir annede olduğu kadar sıcak ve güvenli bir ilişki kurabilir. (K1)’ ifadesinin, cinsiyet değişkeniyle Ki-Kare Testi;

H<sub>0</sub>: Cinsiyet ile K1 değişkeni arasında anlamlı ilişki yoktur.

H<sub>1</sub>: Cinsiyet ile K1 değişkeni arasında anlamlı ilişki vardır.

Tablo 6. Cinsiyet ve K1 için Ki-Kare Testi

			Çalışan bir anne çocukları ile çalışmayan bir annede olduğu kadar sıcak ve güvenli bir ilişki kurabilir			Toplam
			kesinlikle katılmıyorum ve katılmıyorum	kararsızım	kesinlikle katılıyorum ve katılıyorum	
cinsiyet	Kadın	Miktar	39	29	311	379
		cinsiyete göre %	10.3%	7.7%	82.1%	100.0%
Erkek	Miktar	61	29	131	221	
		cinsiyete göre %	27.6%	13.1%	59.3%	100.0%
Toplam	Miktar	100	58	442	600	
	cinsiyete göre %	16.7%	9.7%	73.7%	100.0%	

#### Ki-Kare Testi

	Değer	serbestlik derecesi(df)	Asimptotik Anlamlılık(Sig.) (2-tarafli)
Pearson Ki-Kare	39.259 <sup>a</sup>	2	.000
Olabilirlik Oranı	38.255	2	.000
Geçerli N sayısı	600		

a. 0 hücre (.0%) 5'ten az beklenen değere sahiptir. En küçük beklenen değer 21.36'dır

Ki kare tablosundaki (Tablo 6) anlamlılık değeri  $0.00 < 0,05$  olduğu için H<sub>0</sub> reddedilir. Yani değişkenler arasında anlamlı ilişki vardır.

‘Çalışan bir anne çocukları ile çalışmayan bir annede olduğu kadar sıcak ve güvenli bir ilişki kurabilir’ ifadesine kadınların %82,1’inin katıldığını, erkeklerin %59,3’ünün katıldığını görüyoruz. Ayrıca bu ifadeye kadınların %7,7’si, erkeklerin %13,1’i kararsızım cevabını, kadınların %10,3’ü, erkeklerin %27,6’sı katılmadığı cevabını vermiştir. Sonuç olarak bu ifadeye kadınlar daha fazla katılmıştır.

‘Genelde çalışan kişiler patronlarının kadın olmasından rahatsız olmazlar. (K12)’ ifadesinin, cinsiyet değişkeniyle Ki-Kare Testi;

H<sub>0</sub>: Cinsiyet ile K12 değişkeni arasında anlamlı ilişki yoktur.

H<sub>1</sub>: Cinsiyet ile K12 değişkeni arasında anlamlı ilişki vardır.

Tablo 7. Cinsiyet ve K12 için Ki-Kare Testi

			Genelde çalışan kişiler patronlarının kadın olmasından rahatsız olmazlar			Toplam
			kesinlikle katılmıyorum ve katılmıyorum	kararsızım	kesinlikle katılıyorum ve katılıyorum	
cinsiyet	Kadın	Miktar	108	65	206	379
		cinsiyete göre %	28.5%	17.2%	54.4%	100.0%
	Erkek	Miktar	51	37	133	221
		cinsiyete göre %	23.1%	16.7%	60.2%	100.0%
Toplam		Miktar	159	102	339	600
		cinsiyete göre %	26.5%	17.0%	56.5%	100.0%

**Ki-Kare Testi**

	Değer	serbestlik derecesi(df)	Asimptotik Anlamlılık(Sig.) (2-tarflı)
Pearson Ki-Kare	2.400 <sup>a</sup>	2	.301
Olabilirlik Oranı	2.424	2	.298
Geçerli N sayısı	600		

a. 0 hücre (.0%) 5'ten az beklenen değere sahiptir. En küçük beklenen değer 37.57'dir.

Ki kare tablosundaki (Tablo 7) anlamlılık değeri  $0.301 > 0,05$  olduğu için  $H_0$  kabul edilir. Yani değişkenler arasında anlamlı ilişki yoktur.

Normalde ilişki olmayan tablo çıktılarını dahil edilmemişti ancak cinsiyet için 12 ifadeden sadece bu ifade için değişkenler arası anlamlı ilişki çıkmadığı için ve erkeklerin sadece bu ifadeye kadınlardan daha fazla katıldığı için bu tablo da gösterildi. 'Genelde çalışan kişiler patronlarının kadın olmasından rahatsız olmazlar' ifadesine kadınlar erkeklerden daha fazla katılmamıştır.

K değişkenleriyle cinsiyet arasında yapılan ki-kare testlerinde K1-K11 arası değişkenlerle cinsiyet arasında yapılan ki-kare testlerinin hepsinde anlamlı bir ilişki vardır sonucuna ve kadınların bu K değişkenleri ifadelerine erkeklerden daha fazla katıldıkları sonucuna varılmıştır. Örnek olarak yukarıda K1 ile cinsiyet arasındaki ki-kare testi Tablo 6'da olduğu gibidir. Ancak sadece cinsiyet ve K12 değişkeni arasında anlamlı fark yoktur sonucuna varılmıştır ve erkekler bu ifadeye kadınlardan daha fazla katılmıştır. Bu durum kadınların erkeklere göre kadın patronlardan daha fazla rahatsız olduğunu gösteriyor.

**Uygulama**

Uygulamada güvenilirlik analizi, faktör analizi ve regresyon analizi yapılmıştır.

B1, B2, B3, B4, K1, K2, K3, K4, K5, K6, K7, K8, K9, K11, K12 değişkenlerine faktör analizi yapılmıştır. Faktör analizinde ortak varyanslar tablosunda K12 değişkeni çıkarılmış ortak varyans değeri 0,4'ten küçük olduğu için K12 değişkeninin çıkartılmasına karar verilmiştir ve analizlere tekrar başlanmıştır.

Güvenirlilik analizi;

Tablo 8. Durum İşleme Özeti

		N	%
Durumlar	Geçerli	600	100.0
	Eksik Değer <sup>a</sup>	0	.0
Toplam		600	100.0

Analizde 600 geçerli gözlem vardır.

Tablo 9. Güvenirlilik Analizi

**Güvenirlilik Analizi**

Cronbach Alfa	Değişken sayısı
.906	14

Güvenirlilik analizinde Cronbach alfa değeri= 0,906 > 0,9 olduğu için güvenirlilik değeri oldukça yüksektir.

Tablo 10. Değişken Toplam İstatistikleri

**Değişken Toplam İstatistikleri**

	Değişken Silindiğinde Ölçek Ortalaması	Değişken Silindiğinde Ölçek Varyansı	Düzeltilmiş Korelasyon	Değişken Silindiğinde Cronbach Alfa
B1	53.33	93.756	.492	.904
B2	53.62	95.069	.354	.909
B3	52.85	95.301	.443	.906
B4	53.55	94.885	.356	.909
K1	53.34	87.738	.600	.901
K2	53.01	91.192	.588	.901
K3	53.48	86.844	.628	.899
K4	53.25	85.965	.730	.895
K5	53.31	85.938	.721	.895
K6	53.59	85.991	.678	.897
K7	53.00	87.411	.766	.894
K8	52.96	88.231	.724	.896
K9	52.98	88.125	.734	.895
K11	53.17	88.002	.692	.897

Tablo 10’da son sütun değişkenler silindiğinde cronbach alfanın alacağı değerleri gösteriyor. Herhangi bir değişken silinmesi cronbach alfayı yüksek değerde arttırsaydı, o değişken analizden çıkarılabilirdi. Burada böyle bir durum söz konusu olmadığı için analize devam edildi.

→ Güvenirlilik analizinden sonra faktör analizi yapıldı.

Faktör analizi;

Tablo 11. KMO ve Bartlett Testi

**KMO ve Bartlett Testi**

Kaiser-Meyer-Olkin Örnekleme Yeterliliğinin Ölçüsü		.920
Bartlett Küresellik Testi	Yaklaşık Ki-kare	4803.511
	df(serbestlik derecesi)	91
	Anlamlılık(Sig.)	.000

KMO testi, örneklem yeterliliğini gösterir. KMO oranının 0,5 ve üzerinde bir değer alması gerekir. Burada KMO değeri 0,92 olarak sonuçlanmıştır. Bu değer %90’ın üzerinde olduğu için örneklem yeterliliği çok iyi düzeydedir.

Bartlett testi, maddeler arası ilişkilerin olduğu gerçek korelasyon matrisiyle birim matris arasında anlamlı fark olup olmadığını sınar. Bartlett testi anlamlılığı 0,00 < 0,05 olduğu için H0 reddedilir. Yani anlamlı bir fark olduğu anlaşılır.



Tablo 12. Ters Görüntü Matrisi

Ters Görüntü Matrisi

	B1	B2	B3	B4	K11	K9	K8	K7	K6	K5	K4	K3	K2	K1	
<b>Ters Görüntü Korelasyonu</b>	<b>B1</b>	.847 <sup>a</sup>	-.369	-.303	-.241	.051	-.023	-.116	.018	-.009	-.080	-.002	.044	.062	-.085
	<b>B2</b>	-.369	.816 <sup>a</sup>	-.048	-.299	-.056	.071	.032	-.021	.052	.032	.027	-.110	.006	-.098
	<b>B3</b>	-.303	-.048	.873 <sup>a</sup>	-.240	-.012	-.048	.002	-.053	.051	.063	-.099	-.001	-.148	.109
	<b>B4</b>	-.241	-.299	-.240	.840 <sup>a</sup>	-.044	-.040	.068	-.052	-.062	.052	.021	-.001	.092	-.005
	<b>K11</b>	.051	-.056	-.012	-.044	.957 <sup>a</sup>	-.141	-.236	-.099	-.151	-.046	-.029	-.129	-.021	.040
	<b>K9</b>	-.023	.071	-.048	-.040	-.141	.932 <sup>a</sup>	-.328	-.316	.001	-.117	.014	-.027	-.022	-.054
	<b>K8</b>	-.116	.032	.002	.068	-.236	-.328	.929 <sup>a</sup>	-.245	-.006	.038	-.120	.011	-.056	-.016
	<b>K7</b>	.018	-.021	-.053	-.052	-.099	-.316	-.245	.942 <sup>a</sup>	-.062	-.202	-.017	-.003	-.049	-.058
	<b>K6</b>	-.009	.052	.051	-.062	-.151	.001	-.006	-.062	.925 <sup>a</sup>	-.175	-.188	-.108	.097	-.338
	<b>K5</b>	-.080	.032	.063	.052	-.046	-.117	.038	-.202	-.175	.932 <sup>a</sup>	-.362	-.092	-.065	.014
	<b>K4</b>	-.002	.027	-.099	.021	-.029	.014	-.120	-.017	-.188	-.362	.927 <sup>a</sup>	-.156	-.227	.016
	<b>K3</b>	.044	-.110	-.001	-.001	-.129	-.027	.011	-.003	-.108	-.092	-.156	.960 <sup>a</sup>	-.141	-.086
	<b>K2</b>	.062	.006	-.148	.092	-.021	-.022	-.056	-.049	.097	-.065	-.227	-.141	.928 <sup>a</sup>	-.233
	<b>K1</b>	-.085	-.098	.109	-.005	.040	-.054	-.016	-.058	-.338	.014	.016	-.086	-.233	.914 <sup>a</sup>

a. Örnekleme Yeterliliği Ölçümleri (MSA)

Ters görüntü matrisi tablosunda (Tablo 12) madde eliminasyonu için .50 sınır değer olarak kabul edilir. Düşük değerler, maddenin diğer maddelerle yeterli korelasyonunun olmadığını gösterir.

Burada değerler (matris köşegen değerleri) her maddenin diğer maddelerle korelasyonunun yeteri kadar iyi olduğunu (>0,50) gösteriyor.

Tablo 13. Ortak Faktör Varyansları

Ortak Faktör Varyansları

	Başlangıç	Çıkarılmış
<b>B1</b>	1.000	.725
<b>B2</b>	1.000	.657
<b>B3</b>	1.000	.556
<b>B4</b>	1.000	.686
<b>K11</b>	1.000	.601
<b>K9</b>	1.000	.673
<b>K8</b>	1.000	.661
<b>K7</b>	1.000	.708
<b>K6</b>	1.000	.591
<b>K5</b>	1.000	.689
<b>K4</b>	1.000	.682
<b>K3</b>	1.000	.503
<b>K2</b>	1.000	.487
<b>K1</b>	1.000	.447

Extraction Method: Principal Component Analysis.

Tablo 13’de sol sütunda başlangıç maddelerinin ortak varyanslarının 1 olmasının sebebi;

Faktörlerin çözümü sürecine her maddenin bir faktör varsayılarak başlanması nedeniyle, başlangıçta madde sayısı kadar faktör vardır ve her maddenin tüm maddelerle ilişkisiyle zaten varyansın tümü açıklanmış olur.

Tablo 13’de sağ sütun(çıkarılmış), maddelerin ilişkili oldukları tüm faktörlerle aralarındaki korelasyon katsayılarının karelerinin toplamını gösterir. Bu değer 0,4’ün altındaysa, o değişkenin çıkarılması gerekir. Burada tüm değerler 0.4’ün üzerindedir.

Tablo 14. Açıklanan Toplam Varyans

Açıklanan Toplam Varyans

Bileşen	Başlangıç Özdeğerleri			Faktörleştirme Sonrası Değerler			Rotasyon Sonrası Değerler		
	Toplam	Varyans Yüzdesi	Birikimli Yüzde	Toplam	Varyans Yüzdesi	Birikimli Yüzde	Toplam	Varyans Yüzdesi	Birikimli Yüzde
1	6.542	46.731	46.731	6.542	46.731	46.731	5.934	42.385	42.385
2	2.125	15.178	61.909	2.125	15.178	61.909	2.733	19.524	61.909
3	.846	6.039	67.949						
4	.708	5.057	73.006						
5	.578	4.126	77.132						
6	.534	3.817	80.949						
7	.469	3.353	84.302						
8	.421	3.005	87.307						
9	.395	2.819	90.126						
10	.337	2.407	92.533						
11	.297	2.124	94.657						
12	.287	2.050	96.707						
13	.239	1.710	98.417						
14	.222	1.583	100.000						

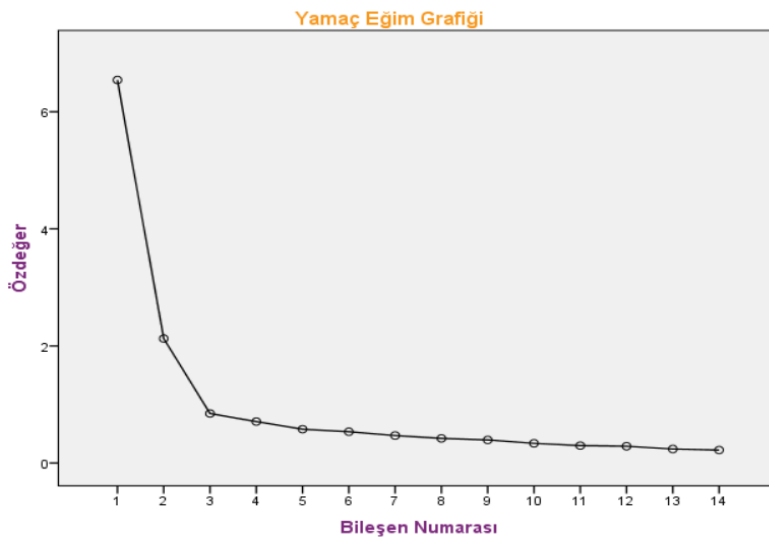
Extraction Method: Principal Component Analysis.

İyi bir faktöriyel çözümlemede en az sayıda faktörle en yüksek varyansın açıklanması beklenir. Toplam varyansın %50-75'ini açıklayan bir analiz mutluluk vericidir.

Burada (Tablo 14) özdeğeri 1'in üzerinde bulunan 2 adet faktör sayılabilecek değişken bulunmuştur.

İlk faktör toplam varyansın %46'sını, ikinci faktör tüm varyansın %15'ini ve 2 faktör toplam tüm varyansın %61,9'unu açıklamaktadır.

Rotasyon faktörlerin görelî önemini eşitlemeyi amaçlar. Rotasyondan sonra; İlk faktör toplam varyansın %42'sini, ikinci faktör tüm varyansın %19'unu açıklamaktadır. Sonuçta yine 2 faktör toplam tüm varyansın %61,9'unu açıklamaktadır.



Şekil 12 Yamaç Eğim Grafiği

Yamaç eğim grafiğine (Şekil 12) bakarak da faktör sayısına ulaşabiliriz. Grafik ya 2 ya da 3 faktör olabileceğini gösteriyor. 3 olmasına karar verilirse analizin yeniden yapılması lazım. 2 faktörle devam edildi.

Tablo 15. Rotasyonsuz/Rotasyonlu Bileşen Matrisi

	Rotasyonsuz Bileşen Matrisi <sup>a</sup>		Rotasyonlu Bileşen Matrisi <sup>a</sup>	
	Bileşen		Bileşen	
	1	2	1	2
K7	.834	-.113	.826	
K9	.810	-.129	.819	.109
K8	.802	-.134	.817	.204
K4	.801	-.202	.800	.180
K5	.797	-.230	.795	.173
K11	.766	-.122	.761	.108
K6	.747	-.182	.756	.171
K3	.697	-.130	.696	.138
K2	.667	-.205	.695	
K1	.664		.646	.174
B4	.374	.739		.825
B2	.367	.723	.216	.824
B1	.507	.684		.807
B3	.471	.577	.223	.711

Faktör yükleri tablosundan (Tablo 15) hangi değişkenin hangi faktör üzerinden en yüksek yüke sahipse o değişkeni o faktöre atılır. Varimax dönüşümü yapıldıktan sonra değişkenlerin faktör yüklerinde değişiklik olsa da değişkenler aynı faktöre atanmıştır. K değişkenleri 1. faktöre, B değişkenleri 2. faktöre atanmıştır.

Regresyon analizi;

Faktör analizi sonucu 2 faktör oluşturuldu. Daha sonraki analizler için her faktörle ilgili skorlar regresyon yöntemiyle kaydedildi. Birinci faktör K değişkenlerini, ikinci faktör B değişkenlerini içeriyor. Birinci faktör soruların içeriğinden dolayı deneyim-performans, ikinci faktör motivasyon olarak isimlendirildi.

‘K10’ değişkeni (Çalışan kadının aile hayatı düzensiz değildir.) ‘Düzen’ olarak isimlendirildi.

Bağımlı değişken ‘düzen’, bağımsız değişkenler ‘deneyim-performans’ ve ‘motivasyon’ değişkenleri olarak seçilerek regresyon analizi yapıldı.

Tablo 16. Model Özeti

Model Özeti				
Model	R	R Kare	Düzeltilmiş R Kare	Tahminin Standart Hatası
1	.812 <sup>a</sup>	.659	.658	.613

a. Tahminler: (Sabit), Motivasyon, Deneyim ve Performans

Model özeti tablosu (Tablo 16) yorumları;

R → Bağımsız değişkenler olan ‘motivasyon’ ve ‘deneyim-performans’ birlikte ele alındığında, bağımlı değişken olan ‘düzen’ ile aralarındaki çoklu derecenin ilişkisi  $R=0,812$ ’dir.

R Kare → Motivasyon ve deneyim-performans bağımsız değişkenleri, düzen durumundaki değişimin %65,9’unu açıklamaktadır. Değişimin kalan kısmı başka nedenlerden kaynaklanmış olabilir.

Düzeltilmiş R Kare → Modelin tahmin gücü %65,8’dir. Bu değer yüksek olmaması gözlem sayısı yetersizliğinden kaynaklanmış olabilir.

Tablo 17. Varyans Analizi

Varyans Analizi (ANOVA)<sup>a</sup>

Model	Kareler Toplamı	serbestlik derecesi(df)	Kareler Ortalaması	F	Anlamlılık (Sig.)
1 Regresyon	434.576	2	217.288	578.020	.000 <sup>b</sup>
Kalıntı	224.423	597	.376		
Toplam	658.998	599			

a. Bağımlı Değişken: Düzen

b. Tahminler: (Sabit), Motivasyon, Deneyim ve Performans

Varyans analizi tablosu (Tablo 17) yorumları;

Kareler Toplamı (regresyon) → Bağımlı değişkendeki değişimle ilgili kareler toplamı 434,576'dır.

Kareler Toplamı (kalıntı) → Analizde açıklanamayan değişimle ilgili kareler toplamı 224,423'tür.

Kareler Ortalaması(regresyon) → Regresyon için kareler toplamının serbestlik derecesine bölümüdür. Yani  $434,576/2 = 217,288$ 'dir.

Kareler Ortalaması(kalıntı) → Kalıntı için kareler toplamının serbestlik derecesine bölümüdür. Yani  $224,423/597 = 0.376$ 'dir.

F → Kareler ortalamalarının birbirine oranı; regresyon/kalıntı =  $217,288/0,376 = 578,02$

Anlamlılık → (Sig = 0,0) <0,05 olduğu için model anlamlıdır.

Tablo 18. Katsayılar

Katsayılar<sup>a</sup>

Model	Standartlaştırılmamış Katsayılar		Standartlaştırılmış Katsayılar	t	Anlamlılık (Sig.)
	B	Std. Hata	Beta		
1 (Sabit)	4.252	.025		169.859	.000
Deneyim ve Performans	.815	.025	.777	32.544	.000
Motivasyon	.247	.025	.235	9.846	.000

a. Bağımlı Değişken: Düzen

Katsayılar tablosu (Tablo 18) yorumları;

Standartlaştırılmamış Katsayılar(sabit) → Bağımsız değişkenlerin hepsi 0 olduğunda 'düzen' değişkeninin alacağı değerdir. Yani sabit terim 4,252'dir.

Standartlaştırılmamış Katsayılar (Deneyim ve Performans) → 0,815'dir. Deneyim ve performansın artması düzeni artırır.

Standartlaştırılmamış Katsayılar (Motivasyon) → 0,247'dir. Çalışma motivasyonunun artması düzeni artırır.

Anlamlılık →  $H_0: \beta_0 = 0$        $H_0: \beta_1 = 0$        $H_0: \beta_2 = 0$

$H_1: \beta_0 \neq 0$        $H_1: \beta_1 \neq 0$        $H_1: \beta_2 \neq 0$

Tüm anlamlılık değerleri 0,05'ten küçük olduğu 3 H<sub>0</sub> hipotezi de reddedilmiştir. Yani tüm katsayılar anlamlı katkı sağlamıştır.

Regresyon denklemi bu şekildedir;

Düzen =  $4,252 + 0,815 * \text{Deneyim ve Performans} + 0,247 * \text{Motivasyon}$

Regresyon denkleminde göre; kadının iş hayatında olmasıyla ilgili olan deneyim ve performans değişkeninin değeri arttıkça aile hayatı düzeni artış gösterir. Ayrıca kadınların çalışmasının nedenleri olan motivasyon değişkeninin değeri arttıkça yine aile hayatı düzeni artış gösterir.

### Sonuç

Bu çalışmada, 600 kişiyle yapılan anketten kadının iş hayatındaki rolünü nasıl gördüklerini öğrenmek amaçlanmıştır. Yapılan anket ve analizlerden birçok farklı sonuç elde edilmiştir. Ankete katılan katılımcılara genel olarak bakıldığında kadınların iş hayatında yeri olduğu düşüncesinde oldukları görülmüştür. Buna karşın farklı düşünce ve yorumlarda verilmiştir. Yüz-yüze anket yöntemi uygulanırken bazı düşüncelerle karşılaşılmıştır. Katılımcılar buradaki düşüncelere kendilerinin farklı düşündüğünü söyleyip, ama artık böyle kabul edilmiş değişmez diyerek kendi düşünceleri dışında cevaplar verme çabasına girmişlerdir. Sizin düşünceniz ne denmesine rağmen benim düşüncem bu durumu değiştirmiyor diyenler olmuştur.

Regresyon denkleminde göre, kadının iş hayatında olmasıyla ilgili olan deneyim ve performans değişkeninin değeri arttıkça ve kadınların çalışmasının nedenleri olan motivasyon değişkeninin değeri arttıkça aile hayatı düzeni artar.

### Kaynakça

- Akkaş, İ., 2019. Cinsiyet ve Toplumsal Cinsiyet Kavramları Çerçevesinde Ortaya Çıkan Toplumsal Cinsiyet Ayrımcılığı, Ekev Akademi Dergisi
- Anonim, 2013. Ev eksenli çalışan kadınlar ve sosyal haklar derneği, Diyarbakır'da kadın istihdamı, sorunlar ve stratejik çözüm önerileri projesi araştırma raporu, Diyarbakır
- Berkman, Y., 2005. Attitudes Towards Women Managers: Development of a New Measure, Koç Üniversitesi
- Büyüköztürk, Ş., 2002. 'Faktör Analizi; Temel Kavramlar ve Ölçek Geliştirmede Kullanımı', Ankara Üniversitesi, Eğitim bilimleri Fakültesi, Eğitim Bilimleri Bölümü
- Dünya Ekonomi Forumu (World Economic Forum), 2018, Global Gender Gap Report
- Ekonomi Analiz, 2013. Güvenilirlik Analizi (Reliability Analysis)
- Euronews, 2018, Cinsiyet Eşitliği Raporu
- Euronews, 2019, 'Eşit Ödeme Günü'
- Gültekin, F., 2020. Regresyon Analizi, <http://w3.balikesir.edu.tr/~bsentuna/wp-content/uploads/2013/03/Regresyon-Analizi.pdf> (Erişim Tarihi: 29.10.2020)
- Kaçar, A., 2020. Kadınlar Halk Fırkasından Türk Kadınlar Birliğine Kadın Hareketi ve Basına Yansımaları, Yüksek Lisans Tezi, Bursa
- Kuzgun, Y., Sevim, S., 2004. Kadınların Çalışmasına Karşı Tutum ve Dini Yönelim Arasındaki İlişki, Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi
- Küçük, M., 2015. Çalışma Hayatında Kadınlar ve Karşılaştıkları Sorunlar: Bir İşverene Bağlı Olarak Çalışan Emekçi Kadınlara İlişkin Bir Araştırma, Ekonomi Bilimleri Dergisi, 7 (1), 1-17.
- Soysal, A., 2010. 'Kadın Girişimcilerin Özellikleri, Karşılaştıkları Sorunlar ve İş Kuracak Kadınlara Öneriler: Kahramanmaraş İlinde Bir Araştırma', Kilis 7 Aralık Üniversitesi, İİBF İşletme Bölümü
- Şentürk, N., 2019. Toplumsal Cinsiyet Bağlamında Kadının Medyada Temsili, Yüksek Lisans Tezi, Erzurum
- TUİK, 2019a. İşgücü İstatistikleri.
- TUİK, 2019b. İşgücü Girdi, İşgücü Maliyeti ve Kazanç İstatistikleri
- Vatandaş, C., 2011. Toplumsal Cinsiyet ve Cinsiyet Rollerinin Algılanışı, Istanbul Journal of Sociological Studies, 0 (35), 29-56.
- Yıldırım, H., 2016. Ki Kare Bağımsızlık Testi Uygulanması ve Yorumlanması/SPSS ile Veri Analizi. <https://www.youtube.com/watch?v=uTmmdTxxreo>. (Erişim Tarihi:29.10.2020) 7

## Statistical Inference for Doubly Geometric Process with Lognormal Interarrival Times

Gültaç EROĞLU İNAN<sup>1</sup>, M. Hilmi PEKALP<sup>1</sup>, Halil AYDOĞDU<sup>1</sup>

<sup>1</sup>Ankara University, Ankara, Turkey

### Abstract

The geometric process is an important stochastic model and widely used in the many areas of the applied statistics. However, it has some difficulties because of model limitations. In situation that it cannot be used, the doubly geometric process is proposed as an alternative model. Recently, this process has become popular since it overcomes some limitations of geometric process. In applications related to doubly geometric process, the parameter estimation problem will naturally arise like in the geometric process. In this study, this problem is handled when the distribution of the first occurrence time is assumed to be lognormal. The parameters of the lognormal distribution and the model parameters of the doubly geometric process are estimated by maximum likelihood method. A simulation study is presented to evaluate the small sample performance of the estimators with different parameter values and sample sizes.

**Key words:** Asymptotic normality, doubly geometric process, geometric process, Lognormal distribution, maximum likelihood estimation

## **POSTER PRESENTATIONS**

## POSTER PRESENTATIONS CONTENT

Classification of Mushrooms of Agaricus and Lepiota Mushroom Families by its Edibility with Decision Tree Algorithms .....	300
Investigation of Factors Affecting Diamond Prices with Data Mining.....	301
Use of Multivariate Adaptive Regression Splines Algorithm (MARS) in Binary Logistic Regression Analysis.....	302
Model Building, Assessment and Optimization of Multivariate Adaptive Regression Splines Algorithm (MARS) for Regression Type Problems.....	303



**Classification of Mushrooms of Agaricus and Lepiota Mushroom Families by its  
Edibility with Decision Tree Algorithms**

Burcu DURMUŞ<sup>1</sup>, Öznur İŞÇİ GÜNERİ<sup>1</sup>

<sup>1</sup>Muğla Sıtkı Koçman University, Faculty of Science, Department of Statistics, Kötekli Campus,  
Muğla

**Abstract**

In this study, data mining algorithms that can classify mushrooms according to their edibility status were determined and classification analysis was made with decision trees. 23 different characteristics belonging to Agaricus and Lepiota mushroom families were discussed. One of these qualities is the 'class' category, which indicates that the fungus is edible or non-edible. This study, it is based on the statistical significance of the decision tree results. For this purpose, data were analyzed with different algorithms. The results of the study showed that decision tree algorithms predicted the edible status of mushrooms almost 100%. On the other hand, this study revealed that 7 out of 22 attributes do not play any role in the classification of mushroom except.

**Key words:** Mushroom, Agaricus and Lepiota, classification, edibility, Decision Tree.

## Investigation of Factors Affecting Diamond Prices with Data Mining

Burcu DURMUŞ<sup>1</sup>, Öznur İŞÇİ GÜNERİ<sup>1</sup>

<sup>1</sup>Muğla Sıtkı Koçman University, Faculty of Science, Department of Statistics, Kötekli Campus,  
Muğla

### Abstract

There are several factors that affect diamond, like other commercial products on the market. Many factors such as shape, cut, color, clarity play a big role in diamond pricing. In this study, 9 different variables (carat, cut, color, clarity, depth, table, length, width, depth) affecting the price of diamonds were discussed and the effects of these variables on price were examined with data mining. The data were analyzed by classification method with the help of R program. A successful result was obtained with 93.92% correct classification rate and 0.923 Kappa statistics. In addition, the 'accuracy' value was found to be 0.89 and results in classification performance were supported. Descriptive statistics and data mining results regarding variables were presented to the reader throughout the study with the help of graphics.

**Key words:** Classification, datamining, decision tree, diamond, diamond features

## Use of Multivariate Adaptive Regression Splines Algorithm (MARS) in Binary Logistic Regression Analysis

Ecevit EYDURAN<sup>1</sup>, Daniel ZABORSKI<sup>2</sup>, Hakan DUMAN<sup>3</sup>

<sup>1</sup>Iğdır University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Quantitative Methods, Iğdır, Turkey

<sup>2</sup>Laboratory of Biostatistics, Department of Ruminants Science, West Pomeranian University of Technology, Klemensa Janickiego 29, 71-270 Szczecin, Poland

<sup>3</sup>Iğdır University, Iğdır Vocational School, Department of Transportation Services, Iğdır, Turkey

### Abstract

Binary logistic regression is a generalized linear model used for solving classification problems in maximum likelihood estimation and is a popular analysis technique employed with a binary response variable and predictor variables i.e. categorical and continuous variables. In binary logistic regression as a logit model, the log odds of the response variable are analyzed as a linear combination of predictors. The aim of this study is to indicate how to use multivariate adaptive regression splines as a data mining algorithm, which enables ones to describe the high dimensional relationship between sets of variables as a modified form of Classification and Regression Tree (CART) algorithm, for binary logistic regression analysis in R environment in order to gain a different perspective and make easier interpretation with the use of predictors with hinges. For this aim, an R script file was formed with some R packages i.e. caret, earth, resample, e1071, ggplots and dplyr. With train function of caret R package, the optimum model that produces the greatest accuracy was determined among the candidate models with tuning parameters i.e. degree=1:3 and nprune=2:40, on the basis of method= "earth", Enhanced Adaptive Regression Through Hinges, an R package required for multivariate adaptive regression splines modeling. Resampling methods such as cross-validation, repeated cross-validation and bootstrap were specified for model building, assessment and optimization procedure in statistical analysis of multivariate adaptive regression splines for binary logistic regression. It could be emphasized that use of multivariate adaptive regression splines in binary logistic regression may be a good alternative for solving classification type problems and making easier interpretations about the obtained analysis results. Consequently, it is considered that the R script file formed here for a binary classification problem will be a valuable preference for researchers who desire solving similar problems and making a proper decision in the future.

**Keywords:** MARS, binary logistic regression, generalized linear models, classification

**Model Building, Assessment and Optimization of Multivariate Adaptive Regression  
Splines Algorithm (MARS) for Regression Type Problems**

Ecevit EYDURAN<sup>1</sup>, Daniel ZABORSKI<sup>2</sup>, Meleksen AKIN<sup>3</sup>, Cem TIRINK<sup>4</sup>

<sup>1</sup>Iğdır University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Quantitative Methods, Iğdır, Turkey

<sup>2</sup>Laboratory of Biostatistics, Department of Ruminants Science, West Pomeranian University of Technology, Klemensa Janickiego 29, 71-270 Szczecin, Poland

<sup>3</sup>Iğdır University, Agricultural Faculty, Department of Landscape Architecture, Iğdir, Turkey

<sup>4</sup>Iğdır University, Agricultural Faculty, Department of Animal Science, Biometry Genetics Unit, Iğdır, Turkey

**Abstract**

The main purposes of the present work were to specify how to optimize applied predictive model of multivariate adaptive regression splines with the objective to predict a continuous variable from categorical and continuous predictor variables and how to create an R script file providing the optimum multivariate adaptive regression splines modeling. Earth R package allows applied predictive modeling of multivariate adaptive regression splines on predictors with hinges for solving regression type problems. The valuable functions of the caret (Classification And Regression Training) R package permit model building, evaluation and optimization process for a more sophisticated predictive MARS model in method= "earth", as an abbreviation of Enhanced Adaptive Regression Through Hinges. For this purpose, a train function of the caret R package as a perfect statistical tool was implemented with the specification of two determinative tuning parameters i.e. nprune= 2:40 and degree= 1:3 and a useful R script was created here for finding the best nprune and degree values based on the smallest RMSE value in method= "earth", which allows ones to perform applied predictive modeling of multivariate adaptive regression splines for regression type problems. The influence of resampling methods; namely, cross-validation, repeated cross-validation, bootstrap and bootstrap-632 on determining the best tuning parameters were comparatively evaluated based on method= "earth" in caret R package. Predictive performances of training and test sets with the scope of efficiently assessing overfitting problem i.e. model generalization ability of MARS predictive model ascertained as the best in training set were measured on ehaGoF R package, estimating 15 goodness of fit criteria. Consequently, it is hoped that the R script file created here will be a pioneer reference for further applied predictive modeling studies of multivariate adaptive regression splines.

**Key words:** Earth package, Caret package, regression, resampling, Bootstrap

### NATIONALITY OF PARTICIPATIONS

<b>Country</b>	<b>n</b>	<b>%</b>
Turkey	79	79
Albania	9	9
Algeria	2	2
Pakistan	5	5
South Africa	3	3
Poland	1	1
Iraq	1	1
Total	100	100.00