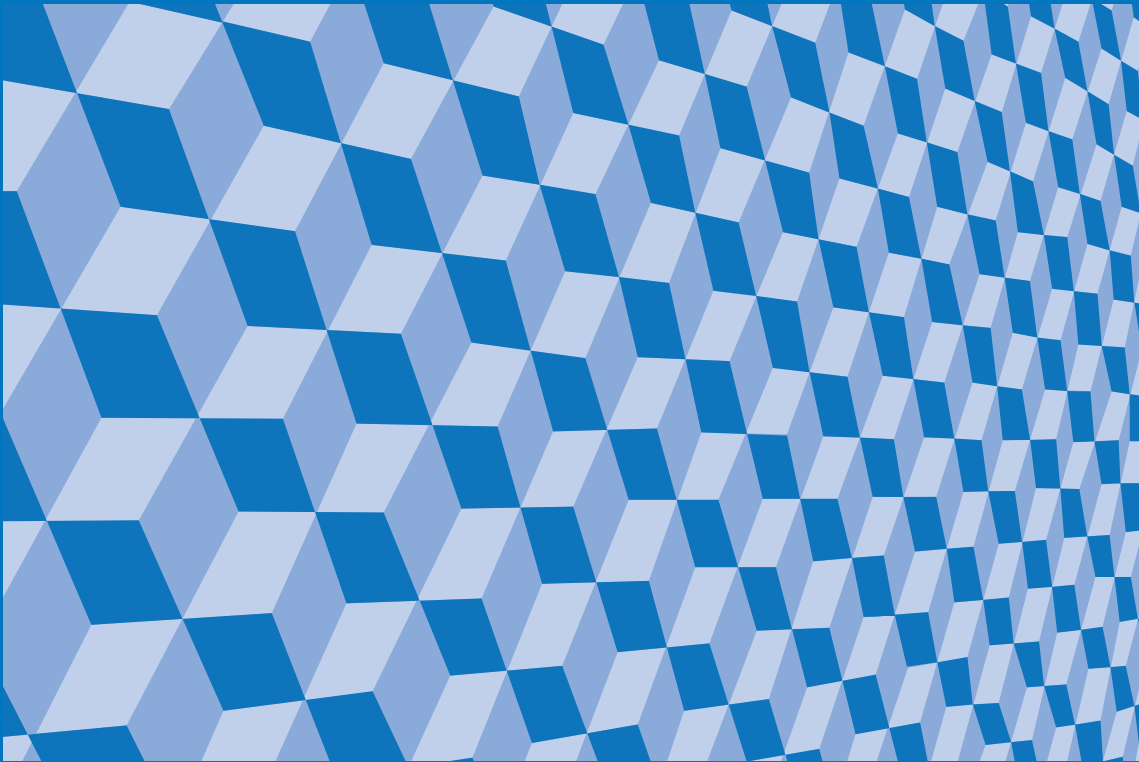




İSTATİSTİK ARAŞTIRMA DERGİSİ Journal of Statistical Research

**Cilt-Volume: 10 Sayı-Number: 01
Özel Sayı-Special Issue 2013**

ISSN 1303-6319



Prof. Dr. Yalçın TUNCER Anısına Düzenlenen İstatistik Sempozyumu, Konuşma ve Davetli Bildiriler, 20 Ekim 2011, ODTÜ Kültür ve Kongre Merkezi, Ankara
Statistical Symposium in Memory of Prof. Dr. Yalçın TUNCER, Speeches and Invited Papers, 20 October 2011, METU Cultural and Convention Center, Ankara

TÜRKİYE İSTATİSTİK KURUMU
Turkish Statistical Institute



İSTATİSTİK ARAŞTIRMA DERGİSİ

Journal of Statistical Research

Cilt-Volume: 10 Sayı-Number: 01
Özel Sayı-Special Issue 2013

Yayın istekleri için

For publication order

Döner Sermaye İşletmesi

Revolving Fund Management

Tel: + (312) 425 34 23 - 410 05 96 - 410 02 85**Faks-Fax:** + (312) 417 58 86**Yayın içeriğine yönelik sorularınız için**

For questions about contents of the publication

Dergi Editörlüğü

Journal Editorship

Tel: + (312) 410 03 67 - 284 45 00/171**Faks-Fax:** + (312) 425 34 05**İnternet**

Internet

http://www.tuik.gov.tr**http://www.turkstat.gov.tr****E-posta**

E-mail

dergi@tuik.gov.tr

journal@tuik.gov.tr

Yayın No

Publication Number

3935**ISSN****1303-6319****Türkiye İstatistik Kurumu**

Turkish Statistical Institute

Yücetepe Mah. Necatibey Cad. No: 114 06100 Çankaya-ANKARA / TÜRKİYE

Bu yayının 5846 Sayılı Fikir ve Sanat Eserleri Kanununa göre her hakkı Türkiye İstatistik Kurumu Başkanlığına aittir. Gerçek veya tüzel kişiler tarafından izinsiz çoğaltılamaz ve dağıtılamaz.

Turkish Statistical Institute reserves all the rights of this publication. Unauthorised duplication and distribution of this publication is prohibited under Law No: 5846.

Türkiye İstatistik Kurumu Matbaası, Ankara

Turkish Statistical Institute, Printing Division, Ankara

Tel: 0312 410 01 64 * Fax: 0312 418 50 82**Temmuz 2013**

July 2013

MTB: 2013-460 - 500 Adet-Copies

Değerli Okuyucular,

İstatistik Araştırma Dergisi'nin bu özel sayısını yakın bir zamanda yitirdiğimiz, başka bir ifadeyle ebedi hayata uğurladığımız sevgili hocamız ve çok değerli bilim insanımız Profesör Dr. Yalçın Tuncer'e ayırmış bulunuyoruz.

Yalçın hocamızı tüm yönleriyle anlatmak için, onu çok yakından ve iyi analiz ederek tanımış olmak gerekir. Bu önsözde, editör olarak ben naçizane, neden onun anısına bir özel sayı çıkarmış olduğumuz üzerinde özellikle durmak istiyorum. Bana göre, bir bilimsel derginin özel sayısı; o derginin ait olduğu bilim alanında, derginin basıldığı ülkede öncülük yapmış, birlikte olduğu bilim çevrelerinde, gerek o bilim alanında ve gerekse de yaşanan hayata dair unutulmayacak izler bırakmış, ve son olarak da, örnek alınabilecek insani ve ahlaki değerlere sahip olduğu geniş kitlelerce kabul görmüş bilim insanlarına ayrılmalıdır. Elbette Yalçın hocamızın Türkiye İstatistik Kurumu'na o dönemki adıyla Devlet İstatistik Enstitüsü'ne gerek Başkan Danışmanı gerekse Başkan Yardımcısı olarak sağladığı önemli katkılar nedeniyle de böyle bir durumu fazlasıyla haketmektedir.

Sevgili Yalçın hocamız hakkındaki bu özel sayımızda, özellikle onu en yakından tanıyan ve kendisi de iyi bir bilim insanı olan sevgili eşi Nilüfer hanımın yazdıkları beni daha çok etkilemiştir. Çünkü o yazdıklarının bir çoğunu benim de gözlemlediğimi ve hiç tereddütsüz onayladığımı belirtmek isterim. Yalçın hocamızın yaptıkları, davranışları, yaşam tarzı, hayata bakış açısı ve diğer yönlerinin, onu tanımış olan kişiler üzerinde bir biçimde etki bırakmış olabileceğine inanlardan biriyim.

İstatistik profesörü olarak Yalçın Tuncer, bilim ahlakına ve insani değerlere sahip, çok yönlü, çalışkan, alçak gönüllü (biraz duygulu) olma gibi özellikleriyle Ülkemizde istatistik biliminin gelişmesinde özel bir yeri olan bir bilim insanıdır.

Sevgili Yalçın Tuncer bıraktığın eserler ve örnek davranışların seni, hiçbir zaman unutturmayacak ve güzel hatırlatacaktır. Nurlar içinde kal...

Bu özel sayımızın basım aşamasına gelinceye kadar çok emeği geçen Sayın Prof. Dr. Öztaş AYHAN'a bilhassa teşekkür etmek istiyorum. Ayrıca, Derginin bu sayısının basım aşamasına gelmesinde desteğini esirgemeyen TÜİK Başkanı Sayın Birol AYDEMİR'e, derginin hemen her aşamasında emeği geçen Sayın Doç. Dr. Özlem İLK'e, Dergi Sekreteryasına ve emeği geçen tüm TÜİK çalışanlarına teşekkür ederim.

Prof. Dr. Fetih YILDIRIM
Dergi Editörü

TÜRKİYE İSTATİSTİK KURUMU TURKISH STATISTICAL INSTITUTE
İSTATİSTİK ARAŞTIRMA DERGİSİ JOURNAL OF STATISTICAL RESEARCH

Sahibi Owner

Türkiye İstatistik Kurumu Adına On Behalf of Turkish Statistical Institute
Birol AYDEMİR Birol AYDEMİR
Türkiye İstatistik Kurumu Başkanı President, Turkish Statistical Institute

Editör Editor

Prof. Dr. Fetih YILDIRIM Prof. Dr. Fetih YILDIRIM

Editör Yardımcısı Assistant Editor

Doç. Dr. Özlem İLK Assoc. Prof. Özlem İLK

Sekreteryä Secretariat

Buket AKGÜN
Z.Nur EMRE
Nurdan ELVER

İÇİNDEKİLER

CONTENTS

Sayfa
Page

ÖNSÖZ III FOREWORD

İÇİNDEKİLER VII CONTENTS

AMAÇ VE KAPSAM IX AIM AND SCOPE

Yalçın Tuncer'in Bilim Adamı Özellikleri 1 Yalçın Tuncer's Characteristics as a Scientist

*Nilüfer TUNCER**Nilüfer TUNCER*

Profesör Yalçın Tuncer'in Eğitim Anlayışı 3 Professor Yalçın Tuncer's Approach to Education

*Fetih YILDIRIM**Fetih YILDIRIM*

ODTÜ Uygulamalı İstatistik Bölümünün Kuruluş Yılları Üzerine Kısa Bir Not 7 A Short Note on the Years of Establishment of METU Applied Statistics Department

*M. Semih YÜCEMEN**M. Semih YÜCEMEN*

Karışık Varyans Analizi Modelinde Etki Faktörünün Rassallık Testi 9 Testing for Random Effect in the Mixed Analysis of Variance Model

*Bilgehan GÜVEN**Bilgehan GÜVEN*

Sihirli Kareler ve Latin Kareler 17 Magic Squares and Latin Squares

*Fikri AKDENİZ**Fikri AKDENİZ*

Bazı Üçgensel Dağılımlar Üzerine İstatistiksel Çıkarımlar: Sınır Değerleri Parametre Olan Durum 29 Statistical Inferences on Some Triangular Distributions: Case of Boundary Values Being Parameters

*İsmail ERDEM**İsmail ERDEM*

Sismik Tehlikenin Tahmini: Olasılık ve İstatistik Yöntemleri	42	Estimation of Seismic Hazard: Probabilistic and Statistical Methods
<i>M. Semih YÜCEMEN</i>		<i>M. Semih YÜCEMEN</i>
Prof. Dr. Yalçın Tuncer'in Resmi İstatistiklere Katkıları	58	Professor Yalçın Tuncer's Contribution to Official Statistics
<i>Orhan GÜVENEN</i>		<i>Orhan GÜVENEN</i>
Yeniden Yapılanma "Bologna Süreci"	75	Restructuring "Bologna Process"
<i>Reşat KASAP</i>		<i>Reşat KASAP</i>
Sansürlü ve Budanmış Veriler İçin Değişim Noktası Problemleri	85	Hazard Change Point Problems for Censored and Truncated Data
<i>Ülkü GÜRLER</i> <i>Deniz YENİGÜN</i>		<i>Ülkü GÜRLER</i> <i>Deniz YENİGÜN</i>
Eşbütünleşme İçin Tahmin ve Test: Spektral Bir Regresyon Yaklaşımı	95	Estimation and Testing for Cointegration: A Spectral Regression Approach
<i>Yılmaz AKDİ</i> <i>David A. DICKEY</i>		<i>Yılmaz AKDİ</i> <i>David A. DICKEY</i>

AMAÇ VE KAPSAM

“İstatistik Araştırma Dergisi (İAD)”, istatistik araştırmaların niteliğinin yükseltilmesi, istatistik yöntem ve uygulamalarının geliştirilmesi, literatürde yer alan çalışmaların tartışılması, istatistik uygulamalarıyla ilgili anket çalışmalarının ele alınması, kuramsal ve uygulama alanındaki araştırmacılar arasında iletişimin ortak çalışma ve yayınlarla güçlendirilmesi amacıyla, yayımlanan hakemli bir dergidir.

“İstatistik Araştırma Dergisi”nin kapsamında yer alan tematik konular aşağıda özet olarak verilmiştir:

- Bankacılık, Finans, Sigortacılık, Aktüerya ve Risk Yönetimi; Bayesci İstatistik; Benzetim Teknikleri; Bilgi Sistemleri; Biyoistatistik; Bulanık Teori; Demografi; Deney Tasarımı ve Varyans Analizi; Ekonometri; Genel Sayımlar ve Değerlendirmeleri; İstatistik Eğitimi; İstatistik Etiği; İstatistik Kuramı; İstatistiksel Kalite Kontrolü; Kamuoyu ve Piyasa Araştırmaları; Klinik Denemeler; Mühendislikte İstatistik Uygulamaları; Olasılık ve Stokastik Süreçler; Optimizasyon; Örnekleme ve Araştırma Tasarımları; Parametrik Olmayan İstatistiksel Yöntemler; Resmi İstatistikler; Toplum Bilimlerinde İstatistik; Veri Analizi ve Modelleme; Veri Madenciliği; Veri Yönetimi ve Karar Destek Sistemleri; Verimlilikte İstatistiksel Yaklaşımlar; Yönetişel Süreçlerde Performans Analizi; Yöneylem Araştırması; Zaman Serileri; Diğer İstatistiksel Yöntemler gibi istatistiğin her dalında yeni bilgi üretimine yönelik tüm araştırmalar.

Makale Dili ve Genel Kurallar

- Bu yayının 5846 Sayılı Fikir ve Sanat Eserleri Kanunu’na göre her hakkı Başbakanlık Türkiye İstatistik Kurumu Başkanlığı’na aittir. Gerçek veya tüzel kişiler tarafından izinsiz çoğaltılamaz ve dağıtılamaz.
- Makale taslakları WORD yazım dilinde, Times New Roman yazı tipinde, 12 punto büyüklükte, satırlar arasında bir satır boşluk bırakılarak yazılmalı, şekil ve grafikler JPG dosyaları olarak hazırlanmalıdır.
- A4 sayfa boyutunda; soldan 3,5 cm, sağdan, yukarıdan ve aşağıdan 2,5 cm boşluk bırakılmalıdır.
- Ana bölüm başlıklarının tümü büyük harf, 12 punto büyüklükte, koyu, ortalı ve Arap rakamları ile numaralandırılarak; alt bölüm başlıklarında ise sadece kelimelerin baş harfleri büyük diğerleri küçük harfle, 12 punto büyüklükte, koyu, sola dayalı ve ana bölüm başlığına endeksli olarak Arap rakamları ile numaralandırılarak yazılmalıdır.
- Makale taslağı yazımında, okuyucunun, çalışmanın her aşamasını anlama ve değerlendirmesine olanak verecek bir anlatım ve plana uyulmalıdır.
- Anlatım olabildiğince sade, anlaşılabilir, öz ve kısa olmalıdır. Gereksiz tekrarlardan, desteklenmemiş ifadelerden ve konu ile doğrudan ilişkisi olmayan açıklamalardan kaçınılmalıdır.
- Yazımda çok genel ifadeler kullanılmamalıdır. Yargı veya kesinlik içeren ifadeler mutlaka verilere/ referanslara dayandırılmalıdır.
- Araştırmacı/araştırmacılar tarafından probleme, hangi kuramsal/kavramsal açıdan yaklaşıldığı, gerekçeleri ile birlikte belirtilmelidir.
- Kullanılan araştırma yönteminin seçilme gerekçesi açıklanmalıdır. Bütün veri toplama araçlarının geçerliliği ve güvenilirliği belirtilmelidir.
- Araştırma sonucunda elde edilen veriler bir bütünlük içinde sunulmalıdır.
- Sadece elde edilen verilere dayanan sonuçlar sunulmalıdır.
- Sonuçların yorumları, varsa, literatürdeki diğer kaynaklarla desteklenerek, değerlendirilmelidir.
- Yararlanılan kaynaklar, çalışmanın kapsamını yansıtacak zenginlik ve yeterlikte olmalıdır.
- Türkçe ve İngilizce özetler; çalışmanın amacı, yöntemi, kapsamı ve temel bulgularını içermelidir.

Ayrıntılı bilgi için, <http://www.tuik.gov.tr> adresinden “İstatistik Araştırma Dergisi Kılavuzu”na bakınız.

AIM AND SCOPE

“*Journal of Statistical Research (JSR)*” is a refereed journal published with the aim to raise the quality of statistical researches, improve the statistical methodology and applications, discuss the studies included in literature, consider survey studies regarding the statistical application, and strengthen the communication between researchers in the field of theory and application by joint studies and publications.

The contents of the “*Journal of Statistical Research*” are summarized below:

- Researches aimed at producing new knowledge in every field of statistics such as Banking, Finance, Insurance Trade, Actuarial and Risk Management; Bayesian Statistics; Biostatistics; Clinic Tests; Data Analysis and Modeling; Data Management and Decision Support Systems; Data Mining; Demography; Econometrics; Experimental Design and Variance Analysis; Fuzzy Theory; General Census and Evaluation; Information Systems; Non-Parametric Statistical Methods; Official Statistics; Operational Research; Optimization; Sampling and Research Designs; Performance Analysis in Managerial Process; Probability and Stochastic Processes; Public Opinion and Market Researches; Statistical Applications in Engineering; Statistical Approaches in Efficiency; Statistical Ethics; Statistical Quality Control; Statistical Training; Statistics in Social Science; Statistics Theory; Simulation Techniques; Time Series; Other Statistical Methods.

Article Language and General Rules

- Prime Ministry, Turkish Statistical Institute reserves all the rights of this publication. Unauthorized duplication and distribution of this publication is prohibited under Law No: 5846.
- Article drafts should be prepared in WORD, using Times New Roman font, in 12 point size, with a blank line in between lines. Figures and tables should be prepared as JPG files.
- On A4 paper size; margins should be set as: left 3,5 cm; right, top and bottom 2,5 cm.
- Titles of the main sections should be all capitalized, in 12 point size, bold, centered and numbered with Arabic numerals; only the first letter of the words in the titles of the subsections should be capitalized, with 12 point size, bold, left justified and numbered with Arabic numerals indexed to the titles of the main sections.
- In article draft writing, writer should follow such a plan that reader should be able to understand and evaluate all the steps of the study.
- Narration should be as plain as possible, as well as comprehensible, compact and short. Unnecessary repetitions, unsupported declarations and explanations that are not in direct relation to the topic should be avoided.
- General statements should be avoided in writing. Statements that include judgment or facts must be supported by data/references.
- It should be stated, with justifications, from which theoretical/conceptual aspect the researcher/researchers have approached the problem.
- The reason of choosing the research methodology that is used should be explained. The validity and reliability of all the data collection tools should be presented.
- Data obtained as the result of the research should be presented in unity.
- Results that only rely on the obtained data should be presented.
- The interpretation of the results should be supported and evaluated by the other resources, if any, in the literature.
- Used resources should be in good wealth and proficiency that reflect the scope of the study.
- Turkish and English abstracts should include the goal, methodology, scope and main findings of the study.

For detailed information, please see “A Guide for Journal of Statistical Research” at <http://www.tuik.gov.tr>.

YALÇIN TUNCER'İN BİLİM ADAMI ÖZELLİKLERİ

Prof. Dr. Nilüfer TUNCER

Sayın Rektör, Değerli Konuklar,

Öncelikle, bu anlamlı Sempozyumu düzenleyerek, bir vefa örneği gösteren ODTÜ İstatistik Bölümü mensuplarına, ve burada hazır bulunan Rektör ve diğer yöneticilere teşekkürlerimi arz etmek isterim.

Bir insanın eşi hakkında, üstelik kaybedilmiş bir eş hakkında konuşması, takdir edersiniz ki oldukça zor. Ben olabildiğince objektif olmaya çalışarak, eşim Yalçın Tuncer'in "bilim adamı" yönüne değinmek istiyorum.

Her şeyden önce, bir bilim adamında olması gerekli "merak" unsuru, O'nda çok kuvvetli idi. Sadece kendi alanını değil, başka konuları, örneğin, sosyal ve fiziki antropoloji, tarih, felsefe, arkeoloji, doğa, linguistik, lehçeler, bitkiler, böcekler, eski eserler, fosiller hep ilgi alanı olmuştu. Ama, matematik ve istatistik, özellikle teorik istatistik tutkunuydu.

Sosyal antropolojide, Türkiye'deki göçler ve yerleşimler üzerine büyük bilgi sahibi idi. Anadolu lehçelerini, özellikle Ege lehçelerini iyi bilir; çok sevdiği Datça'da, köylülerle kendi lehçeleri ile konuşurdu. Kelimelerin etiyolojisine de meraklıydı. Hiç üşenmez, birçok sözlük karıştırır, bilgiye ulaşmadıkça rahat edemezdi.

Tarihi de iyi bilirdi. Özellikle antik dönemi, buna bağlı olarak arkeolojiyi anlatan yerli/yabancı her eseri merakla okurdu. Ege Bölgesi'ndeki antik dönem yerleşimlerine özel bir ilgisi vardı. "Karya"lı olmakla övünürdü.

Doğanın her yönünü severdi. Son durağı olan Tucson şehrini "Tabiat ve şehrin optimum bileşeni" olarak tanımlamıştı. Bitkiler, özellikle şifalı bitkiler, ve böcekler konusunda geniş bilgisi vardı.

Bütün bu konularda eline geçirdiği her şeyi okur ve belleğinde tutardı. Evdeki kütüphanesindeki eserlerin çeşitliliği, onun farklı ilgi ve meraklarının göstergesidir. Dolayısıyla, bir bilim adamında bulunması gerekli olan "araştırma" hevesi, yaşamının önemli bir parçası idi. Bir konuyu araştırmak istediğinde yurt içi kaynaklar yeterli olmazsa, yurt dışından kaynak getirir, gerekirse konuyla ilgisi olabilecek kişilerden bilgi alırdı.

Sebatkardı. Ele aldığı konuyu mutlaka sonuna kadar araştırır; ve eğer bir yayına dönüşmüşse, bir bilim adamı titizliği ile uğraşırdı. Yazılarında içerik ve şekilsel hiçbir hatanın olmaması için çok gayret ederdi.

Gene bir bilim adamı olarak, "farklı düşünebilme" yeteneğine sahipti. Konulara herkesten farklı bir görüş açısıyla bakar, analitik ve eleştirel yaklaşırdı. Bu yüzden bazen yadırganır, fakat zaman içinde haklılığı ortaya çıkardı.

O'nda gerçek bir bilim adamında görülen “tevazu” mevcuttu. Engin bilgisine rağmen, öne çıkmaktan hiç hoşlanmaz, ancak görüşlerini savunmaktan da çekinmezdi. Kimseyi kırmak istemezdi. Görüş ayrılıklarında o ortamdan çekilmeyi tercih ederdi.

Yöneticilik onun bilim adamı yapısına uygun değildi. Ama kurucusu olduğu ODTÜ ve Ankara Üniversitesi Fen Fakültesi İstatistik Bölümleri'nde uzun yıllar, Devlet İstatistik Kurumu'nda üç yıl, bir yıl da KKTC Lefke Avrupa Üniversitesi'nde mecburen yöneticilik görevini üstlendi. Bu süreçlerde canla başla çalışıp, bulunduğu kurumun çalışanları tarafından çok sevildi. O'nun her seviyeden insanla kısa sürede iletişim kurabilme gibi bir yeteneği de vardı.

Bir bilim adamında olması gereken özelliklerden biri de “bilgi paylaşımı”dır. Bu da O'nun için eğitim ve yayın yapmak demektir. Binlerce lisans öğrencisi, onlarca yüksek lisans ve doktora öğrencisi, bugün yurdun değişik yörelerindeki kurumlarda, üniversitelerde, bankalarda, TÜİK'de görev almış durumdadır. Hocalarının, istatistik derslerinde verdiği atipik örnekleri hala gülererek hatırlıyorlar.

Yayınlarının her birinin özgün olup, derin bir düşünce ve araştırma sonucu, büyük bir emekle ortaya çıktığını yakından biliyorum. Üzerinde son çalıştığı makalesi, nihayet Communications in Statistics dergisi tarafından basıldı. Basılı formunu göremedi ama internetteki edisyonunu Mayıs ayında gördüğünde çok mutlu oldu. Zira farklı fikirlerini derginin hakemlerine anlatabilmek için iki yıl uğraş vermişti.

Dürüstlük bir bilim adamının en önemli vasıflarındandır. Yalçın Tuncer'de bu fazlasıyla vardı. İnandığı davadan dönmez, sonuna kadar mücadele ederdi.

Son olarak, bazı bilim adamlarının O'nun için söylediklerini aktarmak istiyorum.

Alabama Üniversitesi'ndeki çalışmaları sırasında tanıştığı iki bilim adamından, Prof. Murat Tanık şöyle söylüyor;

“Arkadaşları, talebeleri ve tüm yazdıklarını okuyanlar, bilim tarihine bıraktığı imzasını görmektedir ve gelecek nesiller de bunu takdir edecektir.”

Aynı üniversitedeki diğer öğretim üyesi Prof. Murat Tanju ise duygularını “büyük bir hüznle dünyanın entelektüel ve insanlık açısından fakirleştiğini hissettim” diye dile getiriyor.

Yalçın Tuncer'in kişiler üzerinde derin etkisi kalmıştır. Bir istatistik toplantısına gidenler, mutlaka O'nun kendi hayatlarında iz bırakmış kişilere rastlarlar. Oğlumuz Dr.Tolga Tuncer de O'nu, kendisine hayatı boyunca örnek alacağı bir “model” olarak görmektedir.

Kuşkusuz ben değerli bir eş, oğlum mükemmel bir baba, öğrencileri iyi bir hoca kaybettik. Ancak eminim ki, özenle yetiştirdiği oğlu ve binlerce öğrencisi, O'nun mükemmeliyetçi anlayışını gelecek kuşaklara taşıyacaklardır.

Nur içinde yatsın.

PROFESÖR YALÇIN TUNCER'İN EĞİTİM ANLAYIŞI

Prof. Dr. Fetih YILDIRIM
Endüstri Mühendisliği Bölümü
Çankaya Üniversitesi

Bir Köy Çocuğu Yalçın Tuncer

Yalçın Hocayla ilk tanışmam, 1979 yılı sonbaharında ODTÜ Uygulamalı İstatistik Bölümüne kısmi görevli olarak ders vermek için davet edildiğim günlerde olmuştur. Hacettepe Üniversitesi Kütüphanesi'nde çalışan ve uzmanlık lisansüstü eğitimi yapan personeline istatistik dersi verdiğim sıralarda Kütüphanenin yöneticiliğini yapmakta olan ve zaman zaman derslere dinleyici olarak katılan eşi Nilüfer Hanımı da tanıyordum. Ancak Nilüfer Hanım eşinin istatistikçi olduğundan hiç bahsetmemişti. Daha sonraki sönestrde, Uygulamalı İstatistik Bölümü'nün tam zamanlı öğretim elemanı olarak göreve başlamıştım. Zamanla, Yalçın Hocayla daha yakınlaşmış ve birçok ortak yanımızın olduğunu fark etmiştik. O da ben de Anadolu'nun bir köyünde dünyaya gelmiştik. Nazilli'nin bir Türkmen köyünden olan hocamız, memleketini, yöresel ege şivesini çok sever ve iyi de kullanırdı. Çok yalın, alçak-gönüllü, güler yüzlü, sevecen, babacan tavırlı, içi-dışı bir kişilikli, kolay sinirlenebilen ancak kin tutmayan bir yapıda idi. Sahip olduklarıyla böbürlenmez, onlar yokmuş gibi davranırdı.

Eğitim: Öğretme-Öğrenme Süreci

Eğitim denilince akla bir öğrenme ve öğretme süreci veya etkinlikler kümesi gelir. Dolayısıyla eğitimin temel amaçlarından birincisi de eğitilenleri (öğrencileri veya öğrencileri) ve belki de eğitenleri de hayata hazırlamaktır. Bu anlamda eğer eğitim, tek yönlü bir etkinlik olursa yani bilgi alışverişi olmaz ise süreç eksik kalmış demektir. Çağdaş eğitimde amaç öğreten ile öğrenen ikilisinin karşılıklı bilgi alışverişi hedeflenmekte ve dolayısıyla, sonuçta memnuniyet karşılıklı olmaktadır. Elbette ki bu süreç boyunca eğitim etkisinin artırılmasını veya sürecin kısaltılmasını sağlayan bir takım araç ve gereçlere gereksinim vardır. Ancak, herşeye rağmen süreçte değişmeyen iki temel olmazsa olmaz unsüre; öğreten ve öğrenendir. Yalçın hocamız bunun farkında olan ve her zaman önemini de dile getiren bir kişiydi.

Öğrenen-Öğreten İlişkisi

Bir önceki paragrafta da değindiğim gibi öğrenen öğreten ilişkisi aslında hayatın ta kendisi değil midir? İnsan yaşadığı her an bilinçli ya da bilinçsiz olarak bir takım bilgileri edinmekte ve onları bir şekilde başkalarına aktarmaktadır. Doğal olarak, bu öğreten-öğrenen ilişkisi ne kadar sağlıklı işlerse ortaya o denli doğru ve etkili etkinlik çıkmaktadır. Yalçın Hocamız, öğretim etkinliğini bu temelde ele alır ve anılan alışverişte bir pazarlık olmaz derdi.

Katılımcı Eğitim

Katılımcı eğitim aslında birlikte öğrenme ve öğretme temelli bir eğitim sürecidir. Burada amaç, kitleleri eğitimde ortak paydada buluşturmak yani paylaşmak ve paylaşırma kültürünü yerleştirmek olmalıdır (yaygın veya paylaşımcı eğitim). Eğitilmiş toplumdan kastedilen olumlu anlamda toplum yararına eğitim sürecinden geçmiş insan öğeleri bütünüdür. Bu süreçte deneyimler paylaşılır ve eğitimin daha etkin olabilmesi için uygun ortam (atmosfer) oluşturulur. Bu süreç, bireylerin ve toplumların bilgi, beceri ve tutumlarını geliştirmeyi amaçlayan etkileşimli bir öğrenme sürecidir. Bu süreçte katılımcıların arzuları, konuya ilişkin görüşleri ve önerileri alınmalıdır. Ayrıca da, bu süreçte alınacak eğitimin anlam ve önemi iyi anlatılmalıdır. Sevgili Yalçın Hocamız, toplumun eğitiminde katılımcı eğitimin önemini ve kaçınılmaz olduğunu hep vurgulardı.

Teşvik Edici Olma

Çağdaş eğitim sürecinin diğer önemli ögesi, eğitenin köstekleyici değil destekleyici ve teşvik edici olmasıdır. Öğrenilecek konunun korkutarak değil sevdilerek veya kolaylıklar sağlanarak öğretilmesi gerekir. Bir internet sitesinde, **“Türkiye’de istatistik biliminin yaygınlaşması ve bu bilimin ‘soğuk yüzünü’ kırmak için bulduğu eğlenceli örneklerle ‘efsaneleşen’ Yalçın Tuncer, alanıyla ilgili çok sayıda kitap yazdı ve binlerce öğrenci yetiştirdi”** denilmektedir. Buradan da hocamızın eğitim yöntemine bakış açısı görülmektedir. Sorunlardan kaçmak yerine yılmadan üzerine gidilmesinden yana tavır almak onun eğitim düsturu olmuştur.

Kuşkucu- Sorgulayıcı Olma

Öğrenme süreci başlangıcı kuşkuyla başlar; akılda var olan bilinmeyene nedir, nasıldır, nerededir, kimdir, doğru mudur, vb. soruların yanıtlarını bulmak üzere yola çıkılır. Öğrenilen her bilginin de doğruluğu veya yanlışlığı sorgulanır. İşte bu kuşku duyma ve sorgulama süreci araştırıcı olmanın temel varsayımıdır. Öğrenilen hiç bir bilgiyi sorgulamadan kabullenmek, eğitimde dogmacılığa ve ezberlemeye götürür. Aslında önemli olan öğrenilmiş bilginin kalıcı olmasıdır ki bu da hazmedilmiş bilgi demektir. Yalçın Hocamız öğrencilerine araştırma ve öğrenme sürecinde kuşkunun ve sorgulamanın önemini ve yararını vurgulamaktaydı. Buna ilişkin bir anekdotumu anlatayım: *Bir lisansüstü dersinde hocamız, öğrencilerinden biriyle anlattığı bir konuda bir tartışma yaşar. Öyle ki öğrenci fikrinde ısrar ettikçe eder, bir türlü ikna olmaz ve dersten üzüntülü bir şekilde çıkar. Öğrenci bir iki gün sonra haksız olduğunu anlar ve hocamızdan özür diler. Sevgili hocamız o öğrenciye kırılmadan anlayış göstererek onu bağışlar.*

Rolleri Değiştirme: Kendine Güveni Arttırma

Çağdaş eğitimde ara sıra eğiten eğitilen rollerinin değiştirilmesi önerilmektedir. Şu sorunun sorulması “Siz olsaydınız bu konuyu nasıl öğrettirdiniz?” uygun ve yararlı olabilir. Böyle bir yöntemin, öğrencinin kendine güvenmesini ve önemsenmiş olma duygusunu arttıracak kuşkusuzdur. Böylece öğrenci, toplum önünde kendini anlatabilme, konuşabilme ve tartışabilme yeteneğini de elde etmiş olacaktır. İşte bu yöntem yine Yalçın Hocamızca sık sık kullanılan bir yöntemdi. Dolayısıyla, öğreten kişi, geleneksel otoriter rollerinden sıyrılarak yardım eden, destekleyen ve öğrenmeyi öğreten olma rolünü üstlenecektir.

Geçmişini Unutmama

Sahip olunan yetenek ve kapasitenin ne olduğunun farkında olma, iyi bir insan olmak için toplumca benimsenmiş aranan özelliklerden biridir. Bu konuda Yalçın Hocamız aşağıdaki hikayeyi anlatarak bu özelliğin önemini ve doğruluğunu dile getirip, “olduğu gibi ol, olamayacağına özenme” derdi. “Bir kurbağa su içmeye dereye gelir, bir de ne görsün! İnek salına salına yürüyor, eğilerek su içiyor. Kurbağa bir kendine bir de ineğe bakmış. Demiş ki “ne kadar alımlı ne kadar güçlü görünüyor, ben de onun gibi olmak istiyorum”. Kararını vermiş, “evet onun gibi olmalı”... Başlamış kendini şişirmeye; şişirmiş de şişirmiş... Kurbağa o kadar şişmiş ki, artık gözleri bile görmez olmuş, ama yetinmemiş, o alımlı inek gibi olmaya karar vermiş bir kere. Olmadığı bir forma girme çabası ne yazık ki kurbağanın sonu olmuş.

İnsan İlişkilerinde Önyargısız ve Kültürel Değerlere Saygılı Olma

Eğitim sürecinin hiçbir evresinde, insanları din, dil, ırk, renk ve yaşadığı veya doğduğu bölge gibi niteliklerine göre ayırmak doğru değildir. Diğer taraftan, eğitim alan insanların hakkında toplumda var olan herhangi bir önyargıya ödün vermeden eğitim verilmelidir. Aslında bu davranışı genel olarak her türlü insan ilişkilerinde göstermek gerekir. Bunlara ilaveten, eğitim sürecinde toplumda var olan kültürel değerlere de saygılı olmak gerekir. Çağdaş eğitimde bu ayrımların hiçbir şekilde yeri yoktur. Profesör Yalçın Hocamız bu konularda oldukça duyarlı idi. Özellikle böyle bir ayrıcalığı olan öğrencisi veya öğrencileri varsa bunlara o farklılıklarını hissettirmemek için özen gösterirdi.

Çağdaşlık, Bilgide Derinliğine Gelişme, Özünde Tutarlılık, Kalitede Sürdürülebilirlik

Bolonya sürecine göre, “Çağdaş eğitimde amaç, ezberleyen değil; bilgiyi öğrenmenin yanısıra öğrendiklerini anlamlandırabilen, araştıran, düşünen, yeni bilgiler üretebilen ve ürettiği bilgileri sorun çözmede kullanabilen bireylerin yetiştirilmesidir”. Bu, eğitimin her aşamasında yeniden yapılanma ve kaliteyi getirmektedir. Eğitim sürecinde asıl olan, uluslararası kuruluşlarca saptanmış ve bu alandaki bilim insanlarının da kabul görmüş standartlara göre eğitim vermektir. Aynı derecede önemli olan ise eğitimde tutturulmuş kaliteyi sürdürebilmektir, devamlılıktır. Ayrıca sahip olunan bilgide derinliğine geliştirme, eğitimin durağan olmaması gerektiği de aranan bir gerçektir. Bu arada, bu süreçte ahlak boyutunun unutulmaması ayrı bir önem taşımaktadır. Yalçın Hocamız, yaşamında hep sade, davranışlarının özünde tutarlı olmuş ve sahip olduğu bilgiyle yetinmemiştir. Bunun kanıtı da daha sonraları ölçüm kuramına ve soyut matematiğe merak duyması, öğrenmesi ve bunda da başarılı olmasıdır.

İstatistik: Deney, Gözlem, Ölçüm, Ölçülebilirlik ve Karar Verme

İstatistik bilimi, deney tasarlama ve elde edilen, ölçülebilen veya sayımlanabilen gözlem sonuçlarını anlamlandırmak için olasılık gibi bilimsel işlemleri kullanan süreçler bütünü olarak tanımlanabilir. İstatistiği, belirli özelliklere sahip deneyin matematiği olarak da tanımlamak mümkündür. Her ne kadar istatistik için yalan söylemek için kullanılan bir araç gibi yaygın kanı varsa da istatistik bir araç değil bir karar verme bilimidir. Buradan da istatistik biliminin öneminin ne kadar büyük olduğu

açık ve seçiktir. İstatistiksel yolla elde edilmiş sonuçların doğru yöntemlerle sunulması ve yorumlanması sürecinde ahlaki kurallara özen gösterilmesi çok önemli olup bu alanda çalışanlar ve kullananlar için gereklidir. Yalçın Hocamız için istatistikle uğraşmak neredeyse bir yaşam tarzı haline gelmiş ve kendisiyle yapılan bir röportajda, “Sizce istatistikçide aranan özellikler nelerdir?” diye sorulmuş ve bu soruya “Gerçek hayat sorunlarına eğildiği için araştırmacı ve analitik, bildiğini halkla paylaşacak derecede dürüst, liberal kafa yapılı bir kişilik olmalıdır.” diye cevap vermiştir.

Eğitimde Felsefenin Rolü

Felsefeyi, dünyada olup bitenleri, sebep sonuç ilişkilerini anlamaya, ideoloji geliştirmeye ve yargılamaya dayalı düşünce süreci şeklinde algıyorsak, eğitim de felsefe, mantıksal bir düşünme süreci geliştirmede önemli bir araçtır denilebilir. Felsefenin bu rolünün önemini kavramış olan Yalçın Hocamız, sık sık özellikle kendi uğraşı alanı olan istatistik eğitim müfredatında felsefe derslerinin gerekliliğini dile getirir ve kuramsal istatistik alanında derinlemesine ilerlemek için felsefenin olası katkılarını tartışmayı severdi.

İyi Bir Akademisyen Olma

İyi bir akademisyen olmak için alanında iyi bir eğitim almış olmanın yanında bence, şu özelliklere sahip olmak gerekir; iyi bir insan olmak, toplumdaki kopmamak, iyi bir dinleyici ve gözlemci olmak, gözlemlerden sonuç çıkarmak, gerçekçi olmak ve son olarak da duygularının esiri olmamaktır. Yalçın Hocamız bu özelliklere sahip iyi bir örnektir.

Mustafa Kemal Atatürk Der ki:

Son olarak, büyük Atatürk'ün, “Yüzyıllar süren derin bir umursamazlığın devlet yapısında açtığı yaraları sarmak için, gerekli olan çabaların en büyüğünü, hiç kuşkusuz eğitim alanında, esirgmeden göstermek gerekir” özdeyişini aklımızdan asla çıkarmamalıyız.

Sevgili Yalçın Hocamız, gelecek nesillerin seni örnek alması dileğiyle ruhun şad olsun.

ODTÜ UYGULAMALI İSTATİSTİK BÖLÜMÜNÜN KURULUŞ YILLARI ÜZERİNE KISA BİR NOT

Prof. Dr. M. Semih YÜCEMEN

İnşaat Mühendisliği Bölümü ve Deprem Çalışmaları EABD
Orta Doğu Teknik Üniversitesi

Bugün, Fen-Edebiyat Fakültesi'nde yer alan İstatistik Bölümü'nün kurulmasında Uygulamalı İstatistik Bölümü'nün önemli bir katkısı olmuştur. 1960-1975 yılları arasında ODTÜ'de şu İstatistik Bölümleri bulunmaktaydı.

- İdari İlimler Fakültesi'ne bağlı olarak 1961-62 akademik yılında kurulan Ekonomi ve İstatistik Bölümü
- 1969-70 döneminde Matematik Bölümü'nde kurulan İstatistik ve Yöneylem Araştırması Opsiyonu
- 1975 yılında Fen-Edebiyat Fakültesi'nde açılan Yöneylem Araştırması ve İstatistik Bölümü

1975-76 ders yılında İktisadi ve İdari Bilimler Fakültesi'nde de Uygulamalı İstatistik Bölümü kurulmuştur. Bu bölümün kurulması için aşağıda isimleri verilen öğretim üyeleri başvuruda bulunmuşlardır.

- Asst. Prof. Dr. Yalçın Tuncer
- Asst. Prof. Dr. M. Semih Yüccemen
- Asst. Prof. Dr. Dolun Öksoy
- Asst. Prof. Dr. Mustafa R. Yılmaz

Uygulamalı İstatistik Bölümünün kuruluş yıllarında aşağıda isimleri verilen öğretim üyeleri bölümün akademik kadrosunda yer almışlardır.

Yalçın Tuncer	Kurucu Öğretim Üyesi
M. Semih Yüccemen	Kurucu Öğretim Üyesi
Dolun Öksoy	Kurucu Öğretim Üyesi
Mustafa R. Yılmaz	Kurucu Öğretim Üyesi
Fetih Yıldırım	Öğretim Üyesi
İsmail Erdem	Öğretim Üyesi
Mehmet Şahinoğlu	Öğretim Görevlisi
İnsan Turnalı	Öğretim Görevlisi
Taşkın Atılğan	Öğretim Görevlisi
Taylan A. Ula	Öğretim Üyesi
Vural Akışık	Öğretim Üyesi

2547 sayılı Yüksek Öğretim Yasası (YÖK) kapsamında yapılan yeni düzenlemeyle 1982 yılında Yöneylem Araştırması ve İstatistik Bölümü, İstatistik Bölümü'ne

dönüştürülmüş ve aynı yıl Uygulamalı İstatistik Bölümü kapatılmış ve Bölümün öğretim üyeleri, 1983 yılında İstatistik Bölümü'ne katılmışlardır.

Prof. Dr. Yalçın Tuncer Uygulamalı İstatistik Bölümü'nün ilk Bölüm Başkanı olmuş ve bugün Üniversitemizde Fen-Edebiyat Fakültesi'nde yer alan İstatistik Bölümü'nün kurulmasında da önemli katkıları olmuştur.

TESTING FOR RANDOM EFFECT IN THE MIXED ANALYSIS OF VARIANCE MODEL

Bilgehan GÜVEN*

ABSTRACT

We consider the mixed analysis of variance (ANOVA) model. It is assumed that random effects in the model are from non-normal universes. Approximate test for random effect is established. The proposed approximate test is based on the asymptotic distribution of the F-ratio in this model and its robustness is given.

Keywords: The mixed ANOVA model, Asymptotic distribution, Approximate F test, Robustness of test.

1. INTRODUCTION

This paper establishes the approximate test for random effect in the mixed ANOVA model having the unbalanced one-fold nested error structure. To achieve this goal the asymptotic distribution of the F-statistic in the model is derived. For the asymptotic distribution of the F-ratio, the group sizes are assumed to be fixed.

The literature of the asymptotic distribution of the F-statistic for the random effects model in either the homoscedastic or heteroscedastic case is extensive: In studies [1, 2, 3, 5, 20], the asymptotic distribution of F-statistic can be reduced to the asymptotic distribution of the difference $MST - MSE$ since MSE converges in probability to constant. Here, MST is the mean square for treatment and MSE is the mean square for error.

In this study, the design matrix X of the model is considered an arbitrary design matrix and the different technique developed by Westfall [22] to obtain the asymptotic distribution of the F-statistic is used. The method of Westfall is based on finding the joint asymptotic distribution of (MST, MSE) and then applying delta method.

As it is explained in Sec. 2., the model is the one-way covariance model in which block effect is random or reduces to the one-way random-effects model or the split plot design with appropriate design matrix. Therefore, the results presented in this paper are applicable for the asymptotic distribution of F-statistic to these mixed ANOVA models.

*Prof. Dr., Mathematics Department, Çanakkale Onsekiz Mart University, 17100 Çanakkale, Turkey, e-mail: bguven@comu.edu.tr

Throughout the paper, we shall use the following notations. $tr(\mathbf{A})^2 = tr(\mathbf{A}\mathbf{A})$ if $\mathbf{A}$ is a symmetric matrix. If $\{d_n\}$ is a real sequence of n and r is a real number, then $d_n = o(n^r)$ if $n^{-r}d_n \rightarrow 0$ as $n \rightarrow \infty$ and $d_n = O(n^r)$ if $n^{-r}d_n$ has a finite nonzero limit as $n \rightarrow \infty$.

2. THE MODEL AND ASYMPTOTIC

The following linear model given in a matrix notation is generalization of some kind of the mixed ANOVA models.

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (1)$$

where $\mathbf{Y} = (\mathbf{Y}'_1, \mathbf{Y}'_2, \dots, \mathbf{Y}'_t)'$, $\mathbf{Y}'_i = (Y_{i1}, Y_{i2}, \dots, Y_{in_i})$ and we assume that the $n \times p$ matrix $\mathbf{X}$ is a matrix of regressors where $n = \sum_{i=1}^t n_i$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)'$. A random error $\boldsymbol{\epsilon}$ is defined as

$$\boldsymbol{\epsilon} = \mathbf{U}\boldsymbol{\xi} + \mathbf{e} \quad (2)$$

where the $n \times t$ matrix $\mathbf{U}$ matrix is:

$$\mathbf{U} = \text{diag}(1_{n_1}, 1_{n_2}, \dots, 1_{n_t}) \quad (3)$$

1_{n_i} is the $n_i \times 1$ vector of ones, $\boldsymbol{\xi}' = (\xi_1, \xi_2, \dots, \xi_t)$ and the vector $\mathbf{e}$ is defined similar to $\mathbf{Y}$. The random vectors $\boldsymbol{\xi}$ and $\mathbf{e}$ are independent and have zero-mean vector and covariance matrix $\sigma_1^2 \mathbf{I}_t$ and $\sigma_2^2 \mathbf{I}_n$ respectively where $\mathbf{I}_t$ and $\mathbf{I}_n$ are the $t \times t$ and $n \times n$ identity matrices.

The model is called either regression with one-fold nested error when $\mathbf{X}$ is a regression type design matrix (see Güven [10], Park and Burdick [16]) or the one-way random-effects model when $\mathbf{X} = \mathbf{1}_n$. Furthermore, the model arises in the split-plot experiment in which there are t -whole-plots and n_i -subplots in the i th whole plot. In the split-plot design the number of subplots in each whole-plot is generally the same.

Real data examples for the regression model with one-fold nested error are given in Park and Burdick [16], Vonesh and Carter [21].

For the random-effects model (1) the sum of squares for treatment SST and the sum of squares for error SSE are:

$$SST = \mathbf{Y}'\mathbf{Q}\mathbf{Y} \text{ and } SSE = \mathbf{Y}'(\mathbf{P} - \mathbf{Q})\mathbf{Y} \quad (4)$$

with the associated degrees of freedom $t - 1$ and $n - p - t + 1$. The $n \times n$ symmetric

idempotent matrices P and Q are defined as

$$P = I_n - H \quad (5)$$

and

$$Q = PU(U'PU)^{-1}U'P \quad (6)$$

respectively where $H = X(X'X)^{-1}X'$ and the superscript $'^{-1}$ denotes a generalized inverse of a matrix. (see [15], ch. 1). From equations (4), (5) and (6) one concludes that both SST and SSE differ when the design matrix X differs.

The F -ratio for the model is

$$F_n = MST/MSE \quad (7)$$

where $MST = SST/(t - 1)$ and $MSE = SSE/(n - p - t + 1)$. The subscript n is used to denote a finite sample size.

With the moment conditions that $E|\xi_i|^{4+\delta} < \infty$ and $E|e_{ij}|^{4+\delta} < \infty$ for some positive δ and the condition that p , the rank of a matrix X , is fixed, the following assumption ensures asymptotic normality of the F -ratio:

Assumption. Consider a sequence of models (1). The number of groups, t , tends to infinity in such a way that the group sizes, $n_1, n_2, \dots, n_t$ take values from a finite set of distinct positive integers $\{m_1, m_2, \dots, m_k\}$. Let $a_t(j)$ be the number of occurrences of m_j in $(n_1, n_2, \dots, n_t)$. Assume further that $(a_t(j)/t) \rightarrow p_j$ as $t \rightarrow \infty$.

Assumption is a modification of the asymptotic formulation of Westfall [22] and says that $(a_t(j)/t)$ has the limiting distribution as long as t tends to infinity. That is: $(a_t(j)/t) \rightarrow p_j$ as $t \rightarrow \infty$ where $0 \leq p_j \leq 1$ and $\sum_{j=1}^k p_j = 1$.

When assumption holds, it can be shown that

$$t = O(n) \text{ and } \sum_{i=1}^t n_i^l = O(n) \text{ for } l = 1, 2, 3, 4.$$

and we also have that

$$\sum_{i=1}^n q_{ii}^2 = O(n)$$

where q_{ii} is the i th diagonal entry of a matrix Q in (6), (see [14], p.872).

3. ASYMPTOTIC DISTRIBUTION OF F -RATIO

The asymptotic distribution of the F -ratio is obtained by using the delta method after establishing the joint asymptotic distribution of (MST, MSE) given by Westfall [22, 23]. We present theorem without giving its proof. The readers who are interested can find proof in [11, 12].

Main Theorem Suppose assumption holds. Then

$$\sqrt{n}(F_n - [1 + \rho c^{-1}]) \xrightarrow{d} N(0, \sigma_F^2)$$

as $n \rightarrow \infty$ where F_n is given in equation (7), the symbol $\xrightarrow{d}$ denotes convergency in distribution and

$$\begin{aligned} \sigma_F^2 = & \frac{2(\rho^2 \tau_2 + 2\rho + c)}{c^2} + \frac{2(1 + \rho c^{-1})^2}{(1 - c)^2} + k_1 \frac{\rho^2 \tau_2}{c^2} \\ & + k_2 \left(\frac{\gamma}{c^2} - \frac{2(1 + \rho c^{-1})(c - \gamma)}{c(1 - c)} + \frac{(1 + \rho)^2(1 - 2c - \gamma)}{(1 - c)^2} \right). \end{aligned} \quad (8)$$

where $\rho = \sigma_1^2 / \sigma_2^2$, the kurtosis parameters k_1 and k_2 for the underlying distributions of ξ_i and e_{ik} are defined as

$k_1 = E[\xi_i^4] / \sigma_1^4 - 3$ and $k_2 = E[e_{ik}^4] / \sigma_2^4 - 3$ respectively and

$$c = \lim_{n \rightarrow \infty} (t/n), \quad \tau_2 = \lim_{n \rightarrow \infty} (1/n) \sum_{i=1}^t n_i^2, \quad \gamma = \lim_{n \rightarrow \infty} (1/n) \sum_{i=1}^n q_{ii}^2.$$

It should be noted that the asymptotic variance σ_F^2 can be approximated by replacing the limit values c, τ_2 and γ with their finite sample size counterparts.

4. IMPLICATION FOR THE APPROXIMATE F-TEST

In this section, the approximate F -test for a null variance ratio is presented and robustness of the test is discussed.

When random effects ξ_i and e_{ij} are from non normal universes, the approximate F -test for $H_0 : \rho = 0$ vs. $H_0 : \rho > 0$ is described as follows. The test rejects H_0 , when

$$F_n > u_\alpha \quad (9)$$

where F_n is given in equation (7) and u_α is the upper $1 - \alpha$ quantile of the asymptotic null distribution of F_n .

From Theorem 3.2., the asymptotic null distribution of F_n is:

$$\sqrt{n}(F_n - 1) \xrightarrow{d} N(0, \sigma_{H_0}^2)$$

as $n \rightarrow \infty$ and

$$\sigma_{H_0}^2 = \frac{2}{c} + \frac{2}{(1-c)^2} + k_2 \left(\frac{\gamma}{c^2} - \frac{2(c-\gamma)}{c(1-c)} + \frac{1-2c+\gamma}{(1-c)^2} \right). \quad (10)$$

The test has an asymptotic level α , i.e., $P(F_n > u_\alpha \mid \rho = 0) = \alpha$ where the asymptotic null distribution of F_n is the normal distribution with the unit mean and variance $\sigma_{H_0}^2$ in equation (10). Then, u_α in (9) is determined by

$$u_\alpha = \frac{\sigma_{H_0}}{n} z_\alpha + 1 \quad (11)$$

where z_α is the upper $1 - \alpha$ quantile of the standard normal distribution.

The approximation to the power of the test for a finite sample, denoted by $K(\rho)$, is:

$$K(\rho) = P(F_n > u_\alpha \mid \rho > 0) = 1 - \Phi\left(\frac{u_\alpha - [1 + \rho c^{-1}]}{\sigma_F}\right), \quad (12)$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution function and σ_F^2 and u_α are given in equations (8) and (11) respectively.

5. ROBUSTNESS OF THE PROPOSED TEST

The size of the approximate F -test does not become asymptotically robust to non-normality when a model is unbalanced (see Westfall [22]). The model (1) is unbalanced. Therefore the size of the proposed test is not asymptotically robust to non-normality. In fact, The null variance $\sigma_{H_0}^2$ in (10) verifies that it is not asymptotically robust since $\sigma_{H_0}^2$ depends on k_2 .

Consider the model (1) with an unknown parameter β_1 and under the design matrix $\mathbf{X} = \mathbf{1}_n \in R^{n \times 1}$ where $n_i = m$ for all i . Then, the model is the balanced one-way random effect model. For this model, $n = tm$, the matrix $\mathbf{Q}$ is in (6) is

$$\mathbf{Q} = \text{diag}\left(\frac{1}{m}J_m, \frac{1}{m}J_m, \dots, \frac{1}{m}J_m\right) - \frac{1}{tm}J_{tm}$$

and the limit values c , τ_2 and γ defined in Main theorem are equal to $c = m^{-1}$, $\tau_2 = m$ and $\gamma = m^{-2}$.

When the limit values for the balanced one-way random effects model are substituted into equation (8), σ_F^2 is:

$$\sigma_F^2 = 2(1 + \rho m)^2 \frac{m^2}{m-1} + \rho^2 m^3 \left(k_1 + \frac{k_2}{m} \right).$$

Since $n = tm$, where m remains fixed, then from Theorem 3.2.

$$\sqrt{t}(F_n - [1 + \rho m]) \xrightarrow{d} N(0, \sigma_b^2)$$

as $t \rightarrow \infty$ where

$$\sigma_b^2 = 2(1 + \rho m)^2 \frac{m}{m-1} + \rho^2 m^2 \left(k_1 + \frac{k_2}{m} \right),$$

and the asymptotic variance σ_b^2 is before given by Akritas and Arnold ([1], p.221).

Then in the balanced one-way random effect model under H_0 , we have

$$\sqrt{t}(F_n - 1) \xrightarrow{d} n \left(0, \frac{2m}{m-1} \right),$$

as $t \rightarrow \infty$. As it is before pointed out in [1], ([2], p.369) and ([18], ch.10) the size of the approximate F -test in the balanced model is asymptotically robust to non-normality in errors.

As noted in [1, 5, 7, 13], the asymptotic null distributions are the same for both a fixed and random factor to be tested.

6. REFERENCES

- Akritas, M., Arnold, S., 2000. "Asymptotics for Analysis of Variance When the Number of Levels is Large", *Journal of the American Statistical Association*, 95, 212-226.
- Akritas, M. G., Papadatos, N., 2004. "Heteroscedastic One-Way ANOVA and Lack-of-Fit Tests", *Journal of the American Statistical Association*, 99, 368-382.
- Arnold, S. F., 1980. "Asymptotic Validity of F Tests for the Ordinary Linear Model and the Multiple Correlation Model", *Journal of the American Statistical Association*, 75, 890-894.
- Arnold, S. F., 1981. *The Theory of Linear Models and Multivariate Analysis*, New York: Wiley.

Bathke, A., 2004. "ANOVA F Test can Still be Used in Some Balanced Designs with Unequal Variances and Nonnormal Data", *Journal of Statistical Planning and Inference*, 126, 413-422.

Bathke, A., Harrar, S. W., Madden, L. W., 2008. "How to Compare Small Multivariate Samples Using Nonparametric Test", *Computational Statistics and Data Analysis*, 52, 4951-4965.

Boos, D. D., Brownie, C., 1995. "ANOVA and Rank Tests When the Number of Treatments is Large", *Statistics and Probability Letters*, 23, 183-191.

Brunner, E., Dette, H., Munk, A., 1997. "Box-Type Approximations in Nonparametric Factorial Design", *Journal of the American Statistical Association*, 92, 1494-1502.

Fuller, W. A., Battese, G. E., 1973. "Transformations for Estimation of Linear Models with Nested-Error Structure", *Journal of the American Statistical Association*, 68, 626-632.

Güven, B., 1995. "Maximum Likelihood Estimation in Simple Linear Regression with One Fold Nested Error", *Communications in Statistics-Theory and Methods*, 24, 121-130.

Güven, B., 2006. "The Limiting Distribution of the F-Statistic From Nonnormal Universes", *Statistics*, 40, 545-557.

Güven, B., 2012. "Testing for Random Effect in Fuller-Battesse Model", Accepted for publication in *Statistics*.

Harrar, S. W., Gupta, A. K., 2007. "Asymtotic Expansion for the Null Distribution of the F-Statistic in One-Way ANOVA Under Non-Normality", *Annals of the Institute of Statistical Mathematics*, 59, 531-556.

Horn, S. D., Horn, R. A., 1975. "Comparison of Estimators Heteroscedastic Variances in Linear Models", *Journal of the American Statistical Association*, 70, 872-879.

Khuri, A. I., Mathew, T. and Sinha, B. K., 1998. *Statistical Tests for Mixed Linear Models*, New York, Wiley.

Park, D. J., Burdick, R. K., 2005. "Performance of Confidence Intervals in Regression Models with Unbalanced One-Fold Nested Error Structure", *Communications in Statistics-Computation and Simulation*, 32, 717-732.

Rudin, W., 1964. *Principles of Mathematical Analysis*, New York, McGraw-Hill.

- Scheffé, H., 1959. The Analysis of Variance, New York: Wiley.
- Van der Vaart, A. W., 1998. Asymptotic Statistics, New York, Cambridge.
- Wang, L., Akritas M. G., 2006. "Two-Way Heteroscedastic ANOVA When the Number of Levels is Large", Statistica Sinica, 16, 1387-1408.
- Vonesh, F. E., Carter, R. L., 1987. "Efficient Inference for Random Coefficient Growth Models with Unbalanced Data", Biometrics, 43, 617-628.
- Westfall, P. H., 1986. "Asymptotic Normality of the Anova Estimates of Components of Variance in the Nonnormal Unbalanced Hierarchal Mixed Model", The Annals of Statistics, 14, 1572-1582.
- Westfall, P. H., 1987. "A Comparison of Variance Component Estimates for Arbitrary Underlying Distributions", Journal of the American Statistical Association, 82, 866-874.

KARIŞIK VARYANS ANALİZİ MODELİNDE ETKİ FAKTÖRÜNÜN RASSALLIK TESTİ

ÖZET

Bu çalışmada baz karışık varyans modelleri genelleştirildi. Bu modellerin rassal etki faktörleri ve hata terimleri normal olmayan bir kitleden geldiği varsayıldı. Etki faktörünün rassallığını test eden bir test elde edildi ve bu testin sağlamlığı çalışıldı.

Anahtar Kelimeler: Karışık ANOVA modeli, Asimptotik dağılım, Yaklaşık *F*-testi, Testin güçlülüğü.

SİHİRLİ KARELER VE LATİN KARELER

Fikri AKDENİZ*

ÖZET

Bu makalede 4000 yıldır bilinen sihirli kareler ele alınmıştır. Özellikle singüler matrisleri veren $2n \times 2n$ ($n \geq 2$) boyutlu sihirli matrislerin Moore-Penrose inversi ve özdeğerlerinin özellikleri incelenmiştir. Ayrıca, Latin kare tasarım modelleri pullar yardımıyla görsel olarak verilmiştir.

Anahtar Kelimeler: Latin kare, Moore-Penrose invers, Sihirli sabit, Sihirli kare, Yarı-sihirli kare.

1. GİRİŞ

Yazımıza ünlü düşünür, filozof ve matematikçi Bertrand Russel (1872-1970)'ın matematik hakkındaki düşüncesi ile başlamak istiyorum. B. Russel "Matematik, aynı şeyi değişik sözcüklerle söyleme sanatıdır. Doğru algılandığında yalnız gerçeği değil, bir heykeldeki gibi yüceltilmiş, donuk ve süssüz bir güzelliği içerir." demiştir. Ayrıca, M.S. 411 yılında İstanbul'da doğmuş ve 485 yılında Atina'da ölen Yunan filozofu Proclus Diadoclus "Bir yerde sayı varsa orada güzellikte vardır" özdeyişiyle matematiğin yalınlığını vurgulamıştır. Matematik evrensel bir dildir. Bunun anlamı eksilme ve bozulma olmaksızın daima var olacaktır. Kümeler kuramının kurucusu Alman matematikçi Georg Cantor (1845-1918)'in deyimiyle "Matematiğin özü özgürlüğünde aranmalıdır". Bu yazımızla sayılardaki yalınlığı ve matematiksel güzelliği paylaşmak istiyorum. Bilgilerin bir kısmını Akdeniz (2011)'de bulabilirsiniz. Bir soru ile konumuza gireceğim.

2. SİHİRLİ KARE NEDİR?

TANIM: $1, 2, 3, \dots, n^2$ sayılarının n satır ve n sütundan oluşan n^2 sayıda kutucuktan oluşan bir kare içine her satır, sütun ve köşegen üzerindeki sayılar toplamı aynı S sayısına eşit olacak şekildeki düzenlenmesine n -ci mertebeden **sihirli kare** (magic square) denir. 1 den n e kadar tamsayıların toplamının $1+2+3+\dots+n=n(n+1)/2$ olduğunu biliyoruz. Benzer biçimde $1^2+2^2+\dots+n^2$ toplamı için n yerine n^2 alınırsa n satırlı toplamı nS olur. O halde $nS=1+2+3+\dots+n^2=n^2((n^2+1)/2)$ olduğundan $S=n(n^2+1)/2$ bulunur. S sayısına "**sihirli sabit**" denir. Yalnız satır ve sütun toplamalarının tümü sabit bir S sayısına eşit olan kare matrisi **yarı-sihirli kare** (semi-magic square) denir. Buradaki gibi ardışık pozitif tam sayılarla oluşturulan sihirli kareye **doğal sihirli kare** denir. Bu yazıda öncelikli olarak doğal sihirli kareleri inceleyeceğiz.

Örneğin: $n=3$ için 1 den 9 a kadar sayılarla oluşturulan sihirli kare için sihirli sabit sayı: $S=3(3^2+1)/2=15$ tir. $n=3$ için sihirli kare örneği vereceğim. 3×3 tipindeki ilk sihirli kare aşağıdaki gibi sayılardan oluşmuştur. Bilinen en eski sihirli kare:

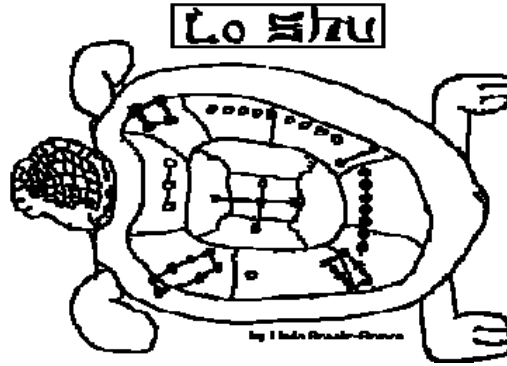
$$M = \begin{bmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{bmatrix}$$

* Prof. Dr., Çukurova Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü Emekli Öğretim Üyesi, e-posta: akdeniz@cu.edu.tr, fikriakdeniz@gmail.com

dir. Bu sihirli karedeki sayıların oluşturduğu yukarıdaki M gösterimine **sihirli matris** denir.

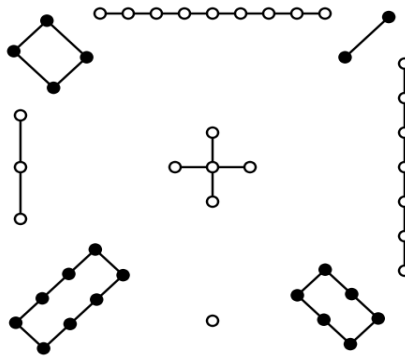
3. SİHİRLİ KARELERİN KISA TARİHİ

Sihirli kareler M.Ö. 21. asırdan beri bilinmektedir. Çin'in eski dönemlerinde çok sayıda taşkınlarla yol açan nehirler varmış. İnsanlar bu nehirlerden "Nehir Tanrısı" olarak bilinen ve su taşkınlarına neden olan "LUO" adlı nehrin kızgınlığını önlemek için nehir kıyısına gidip ona kurban sunmaya çalışılmış. Çin mitolojisine göre her bir tören anında nehirde bir kaplumbağa çıkıp kurbanın etrafında dolaşmış. Nehir tanrısı her defasında kurbanı kabul etmezmiş. Bu durum orada bulunan bir çocuğun, kaplumbağanın üzerindeki ilginç şekilleri işaret edene kadar böyle sürüp gitmiş. Kabuk üzerindeki işaretler 15 sayısını ifade ediyormuş. O günden sonra insanlar kurban sayısının kaplumbağa kabuğu üzerindeki miktar kadar, yani 15 olması gerektiği sonucuna varmışlar.



Şekil 1. Kutsal Kaplumbağa (The Divine Turtle)

Diğer bir Çin mitine göre, M.Ö. 21. asırda Antik Çin'de taşkınların kontrolü için hendekler yaptırıp ve kanallar açtırıp efsanevi kural koyucu olarak bilinen büyük Çin İmparatoru Yu bu ilginç sayısal şekle "LUO-SHU" adını verir. Çince "Luoshu" nun anlamı Luo nehri demektir. Çok daha önceleri bulunmasına rağmen bu sihirli kare ve Luo Shu hikayesi M.S. 1. yüzyılda Yih King başlıklı kitapta bulunmuştur.



Şekil 2. Luo Shu 3X3 Sihirli Kare

Sihirli kareler tarihsel süreç içinde Çin, Hindistan, Araplar, Antik Yunan ve Japonya'da düzenlenmişti. Hindistan'da 4x4 tipinde ilk sihirli kare ikinci asırda yaşamış Budist filozof Nagarjuna tarafından yazılmış Kaksaputta adlı eserde bulunmuştur. Orta Hindistan'da Khajuraho'da 10. asırdan kalma Parsvanath Hindu Tapınağındaki 4x4'lük sihirli karede her bir satır, sütun ve köşegenlerdeki sayılar toplamı 34 tür. Ayrıca 2x2'lik alt karelerin toplamı ve 3x3, 4x4 karenin köşelerindeki sayılar toplamı da 34 olur.



Şekil 3. Hint sayılarıyla 4X4 Sihirli Kare

7	12	1	14
2	13	8	11
16	3	10	5
9	6	15	4

Şekil 4. 4X4 (yukarıdaki) Sihirli Kare

Arap matematikçisi Ahmad Al-Buni 1250'li yıllarda sihirli karelerin özelliklerini incelemiştir. Avrupa'nın sihirli karelerle tanışması 13. asrın sonları 14. asırda İstanbul'da yaşamış olan Yunan matematikçi Manuel Moschopoulos ile gerçekleşmiştir. Manuel, sihirli kareler üzerine bilimsel inceleme yaptı. Daha sonraları 1450'lerde İtalyan matematikçi Fra Luca Bartolomeo de Pacioli (1446/7-1517) çok sayıda sihirli kare örneği toplamıştır. Avrupa'da sihirli karelerin ortaya çıkışı Alman sanatçısı Albrecht Dürer (1471-1528)'in kabartma resminde görülmektedir. Dürer 1514 yılında bakır plakette kabartma üzerine sağ üst köşesine sihirli kareyi yerleştirerek Melencolia'ı (*Melancholia I*) yapmıştır. Bir ressam olan Albrecht Dürer matematik ve sanat arasındaki ilişkiyi incelemiştir.

Şimdi Dürer'in bakır oymacılığı ile tabloyu yaptığı tarihi ve annesinin ölüm tarihi (1451-1514) 1514 yılını da içeren sihirli kareyi inceleyelim. 1 de 16 ya kadar tam sayılardan oluşan bir doğal sihirli karedir. Sihirli sabit= $S=34$ tür. En alt satırda 15 ve 14 sayıları yan yana gelerek sihirli kare özel olarak 1514 yılını vurgulamaktadır.

Bu güne kadar 880 adet farklı 4x4 sihirli kare hesaplanmış durumdadır. Sihirli kareler matematik tarihi içinde pek çok araştırmacının ilgisini çekti. Bunların arasında Leonhard Euler (1707-1783) ve 19. yüzyılda Ludwig Bieberbach (1886-1982). Ayrıca Henry Dudeney (1857-1930) ve Martin Gardner (1914-2010) gibi bulmaca çözerlerin çalışmalarını da eklemek gerekir.



Şekil 5. Melencolia I Albrecht Dürer (1514)



Şekil 6. Dürer'in 4×4 sihirli karesi

3.1 Benjamin Franklin 8x8 Sihirli Kareleri

Matematik dünyasında bir çoğu hala bilinmeyen bazı sihirli kareler 18. yüzyılda yaşamış olan Benjamin Franklin (1706-1790) tarafından oluşturulmuştur. Franklin sabun yapma işi ile uğraşan Josiah Franklin'in 17 çocuğundan biridir. Bir yazar, diplomat ve bilim adamı (Fizikçi ve Matematikçi) olarak Franklin kendi yaşadığı dönemde ve günümüzde övgüye değer biridir (Pasles, 2001). Örneğin, Londra ve Paris bilimler akademilerinin bilimsel üyeliğe seçtiği tek Amerikan'dı ve uzun yıllar da tek kalmıştı. Franklin'in matematiksel yaşamı karelerin kendileri ve almanaklardır. Franklin'in formal okul eğitimi 2 yıldan daha kısa bir sürede son bulmuş fakat onun eğitimi asla son bulmamıştır. Esas itibarıyla kendi kendini yetiştirmiştir. Başarılı bir insanın diğer insanlardan daha çok çalışması gerektiğine inanırdı. ABD 100 dolarında resmi olan kişidir. Franklin aşağıdaki gibi kareleri nasıl yazmıştı? El yazması dışında karşılaştırmak için elimizde bir belge de yoktur. Günümüze kadar bir çok araştırmacı Franklin karelerini incelediler. Bunlardan en ilginç 8x8 yarı sihirli karedir. Bu karedeki satır ve sütunların toplamı $S=8(8^2+1)/2=260$ tır. Köşegen toplamaları 260 değildir. Değişik renklerdeki simetrik yapıdaki sayılar toplamının da 260 olduğu görülmektedir.

52	61	4	13	20	29	36	45	52	61	4	13	20	29	36	45	52	61	4	13	20	29	36	45
14	3	62	51	46	35	30	19	14	3	62	51	46	35	30	19	14	3	62	51	46	35	30	19
53	60	5	12	21	28	37	44	53	60	5	12	21	28	37	44	53	60	5	12	21	28	37	44
11	6	59	54	43	38	27	22	11	6	59	54	43	38	27	22	11	6	59	54	43	38	27	22
55	58	7	10	23	26	39	42	55	58	7	10	23	26	39	42	55	58	7	10	23	26	39	42
9	8	57	56	41	40	25	24	9	8	57	56	41	40	25	24	9	8	57	56	41	40	25	24
50	63	2	15	18	31	34	47	50	63	2	15	18	31	34	47	50	63	2	15	18	31	34	47
16	1	64	49	48	33	32	17	16	1	64	49	48	33	32	17	16	1	64	49	48	33	32	17

Şekil 7. Yukarıda 8x8 lik yarı-sihirli Franklin kareleri görülmektedir. S=260

3.2 Pandiagonal Sihirli Kareler

Tanım. Bir matrisin, satır, sütun ve köşegen elemanlarının toplamlarının aynı sayıya eşit olması özelliğine ek olarak “kırılmış” köşegenlerin toplamı da aynı sihirli sayıya eşit ise bu sihirli kareye “pandiagonal” denir

$$P = \begin{pmatrix} 1 & 15 & 4 & 14 \\ 8 & 10 & 5 & 11 \\ 13 & 3 & 16 & 2 \\ 12 & 6 & 9 & 7 \end{pmatrix}$$

Yukarıdaki P matrisinde 6 adet kırılmış köşegen:

(1, 6, 16, 11); (8, 15, 9, 2); (13, 10, 4, 7); (14, 8, 3, 9); (4, 11, 13, 6); (15, 5, 2, 12) dir. Bu sayıların toplamı da sihirli sabite, S=34 e eşittir (Henrich, 1991).

3.3 Ramanujan’ın Sihirli Karesi

Cumhuriyet Bilim Teknik Dergisinin 12.01.2002 tarihli, 773. Sayısında yayınlanan “Hintli bir matematik dahisi” (Türkkaya Ataöv) başlıklı yazıda 22.12.1887 yılında doğmuş Hintli matematikçi Srinivasa Ramanujan’ın kısa yaşam öyküsü anlatılmıştı. Üniversite öğrenimi görmemiş olan bu *dâhi*, sayılarla oynamayı çok severdi. Ramanujan doğum tarihini de kullanarak aşağıdaki 4×4 biçimindeki kareyi oluşturmuştur.

$$\begin{bmatrix} 22 & 12 & 18 & 87 \\ 21 & 84 & 32 & 02 \\ 92 & 16 & 07 & 24 \\ 04 & 27 & 82 & 26 \end{bmatrix}$$

Soldan sağa, yukarıdan aşağıya tüm toplamlar 139 sayısını vermektedir. Ayrıca her bir köşegendeki 4 sayının toplamı da (22+84+07+26 ve 87+32+16+04) 139 u vermektedir. Köşeden köşeye toplamarda (22+04+26+87) 139 dur. Ayrıca 21+02+24+92=139, 12+18+82+27=139, 12+87+16+24=139, 21+32+82+04=139, 22+18+82+04=139 dur. Her bir köşedeki 2x2 matrislerin dört sayısının toplamı da 139 dur.



Şekil 8. Posta pulu üzerinde Ramanujan (1887-1920)

Satır ve sütun toplamlarının tümü 139 sayısına eşittir. Ayrıca her iki köşegendeki sayılar toplamı da 139 sayısına eşittir. Sihirli kareler üzerinde yoğun çalışmalar yapan Prof. Dr. George P.H. Styan (McGill Üniv. Kanada) Ocak 2012 içinde Ramanujan'ın doğumu ile ilgili aşağıdaki ilk satırları aynı olan başka bir sihirli kareyi göndermiştir.

$$\begin{bmatrix} 22 & 12 & 18 & 87 \\ 28 & 59 & 40 & 12 \\ 80 & 3 & 37 & 19 \\ 9 & 65 & 44 & 21 \end{bmatrix}$$

4. SİHIRLİ KARELER VE MOORE-PENROSE İNVERSİ

Booth ve Booth (1955)'da ifade edildiği gibi singüler olmayan bir sihirli karenin elemanlarıyla oluşturulan bir matrisin tersi de bir sihirli karedir. S “sihirli-sabit” olmak üzere bir matrisin tersinin satır ve sütunlarının toplamaları $1/S$ ye eşittir. Aşağıdaki 3×3 sihirli kareyi ele alalım. Singüler olmayan M matrisi için

$$M = [a_{ij}] = \begin{bmatrix} 2 & 9 & 4 \\ 7 & 5 & 3 \\ 6 & 1 & 8 \end{bmatrix}$$

$$\text{Det}(M) = \Delta = -360 \text{ tır.}$$

$$\begin{aligned} M^{-1} &= [a_{ij}]^{-1} = \left[\frac{A_{ji}}{\Delta} \right] = \\ &= -\frac{1}{360} \begin{bmatrix} 37 & -68 & 7 \\ -38 & -8 & 22 \\ -23 & 52 & -53 \end{bmatrix} \end{aligned}$$

Görülebileceği gibi M^{-1} matrisinin satır ve sütun toplamaları $1/15$ tir. Aynı özellik köşegen toplamaları için de geçerlidir. Birçok sihirli kare matris incelendikten sonra aşağıdaki sonuca varıyoruz.

Sonuç. $(2n+1) \times (2n+1)$ $n=1,2,3,\dots$ tipinde yarı-sihirli ya da sihirli kare matris singüler değil ve pozitif tam sayı olan en büyük öz değer sihirli sabit S ye eşittir. Ayrıca diğer öz değerler $\alpha \pm i\beta$ ya da $-\gamma \pm \delta, \pm \xi$ biçiminde toplamaları sıfırdır. Yani $\sum_{i=2}^{2n+1} \lambda_i = 0$

dır. Dürer'in sihirli karesine bakıldığında karşılık gelen kare matris singüler (tekil) ve rankı 3 tür. Bu nedenle Booth ve Booth (1955)'daki sonuç uygulanamaz. O halde matrislerin Moore-Penrose inverslerinin de sihirli kare özelliklerini taşıyıp taşımadığını öğrenmek istiyoruz.

Bir A matrisinin A^+ ile gösterilen Moore-Penrose inversi aşağıdaki dört koşulu sağlarsa A^+ tek olarak tanımlanır.

$$(i) AA^+A = A, (ii) A^+AA^+ = A^+, (iii)(AA^+)' = AA^+, (iv)(A^+A)' = A^+A$$

A bir kare matris olmasa bile bu invers daima vardır (Rao ve Mitra,1971)). Yalnız (i) koşulu sağlanırsa, yani $AA^-A=A$ ise, A^- ye A nın genelleştirilmiş-inversi denir. Her matris g-inverse sahiptir, fakat tek olmayabilir.

Şimdi Dürer tarafından yapılan "Melancholia I" resminde yer alan aşağıdaki meşhur sihirli kareyi düşünelim:

$$D = \begin{bmatrix} 16 & 03 & 02 & 13 \\ 05 & 10 & 11 & 08 \\ 09 & 06 & 07 & 12 \\ 04 & 15 & 14 & 01 \end{bmatrix}$$

Satır, sütun ve köşegen toplamaları 34, dört köşedeki 2×2 alt kare matrislerin elemanlarının toplamı da 34 tür. D singüler bir matristir ve rankı 3 tür. D , 4×4 tipinde sihirli kare matrisin MP-inversi de sihirli kare olan bir matrise örnektir. D nin özdeğerleri : $\lambda_1 = 34$, $\lambda_2 = 8.0$, $\lambda_3 = -8.0$, $\lambda_4 = 0$, $trace(D) = 34 = \sum_{i=1}^4 \lambda_i$ dir. Burada

D matrisinin en büyük özdeğeri sihirli sabit olan $S=34$ e eşittir. Görüldüğü gibi D nin MP-inversi:

$$D^+ = \frac{1}{34.80} \begin{bmatrix} 275 & -201 & -167 & 173 \\ 37 & -31 & -65 & 139 \\ -99 & 105 & 71 & 3 \\ -133 & 207 & 241 & -235 \end{bmatrix}$$

dir (Trenkler, 1994; Schmidt ve Trenkler, 2001). Görülebileceği gibi satır, sütun ve köşegen toplamaları, dört köşedeki alt kare matrislerdeki sayılar toplamı da $1/S = 1/34$ tür.

Benjamin Franklin tarafından oluşturulan 8×8 kare matrisin öz değerlerini inceleyelim: Öz değerlerin üçü $\lambda_1 = 260,00$, $\lambda_2 = -43,71281$, $\lambda_3 = 11,71281$ ve diğerleri sıfır

olmaktadır. Matris singülerdir. $trace(A) = 228 = \sum_{i=1}^8 \lambda_i$ dir. rank $(A)=3$ tür. En büyük özdeğer sihirli sabit $S=111$ e eşittir

5. PULLARLA OLUŞTURULAN 4X4 ÖZEL SİHİRLİ KARELER

1945 doğumlu iki öğretim üyemiz Prof. Dr. Fikri Akdeniz ve Prof. Dr. Öztaş Ayhan'ın yaşını (2011 yılında) veren özel sihirli kareyi örnek olarak vereceğim. Satır, sütun ve köşegen toplamaları 66 dır. Sihirli sabit $S=66$ dır.



Şekil 9. 4x4 lük özel sihirli kare. Sihirli sabit $S=66$ dır.

$$F = \begin{bmatrix} 22 & 9 & 20 & 15 \\ 19 & 16 & 21 & 10 \\ 13 & 18 & 11 & 24 \\ 12 & 23 & 14 & 17 \end{bmatrix}$$

Pulların üzerindeki değerler yukarıdaki F sihirli matrisindeki sayıların oluşturduğu sihirli kareyi vermektedir. Pullardan 5 tanesi 1958 yılında ülkemizde basılan kullanılmamış posta puludur. (Değerleri 20, 15, 10, 18, 12 kuruştur.) Diğerleri yabancı pullardan derlenmiştir (Chu ve ark. 2011).

6. LATİN KARE TASARIMI NEDİR?

$k \times k$ ($k \geq 2$) Latin kare tasarımı k satır ve k sütunu içeren bir karedir. Latin kareler orijinal olarak matematiksel merak olarak görünür, fakat istatistiksel uygulamalarda deneysel tasarımlar olarak 20. yüzyılın ilk zamanlarında kullanılmıştır. Örneğin 4 hastanede 4 farklı yöntem uygulayarak 4 deterjanı test etmek istiyoruz. Bu durumda 4×4 Latin kare tasarımı uygulanır. Her bir deterjanı herbir yöntemle herbir hastanede bir kez kullanacağız. İlk satır 1234 ve ilk sütun 1234 olmak üzere 4×4 latin kare matrisi:

$$L = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 2 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix}$$

dir. ANOVA incelemesinde Latin kare modelinde her bir y_{ijk} gözlem değeri aşağıdaki gibi yazılır:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}, i, j, k = 1, 2, 3, 4$$

Latin kareler SUDOKU'nun gerçek atalarıdır (Styan ve ark. 2009). Latin karelerle ilgili örneklere 700 yıldan daha eski Arap kaynaklarında raslanır. Bunlara yeni tip sihirli kare de denir. Sayılar, harfler ve sembollerle oluşturulur. Latin kare ve sihirli kare arasındaki fark kullanılan sembollerin sayısıdır. Örneğin 4x4'lük bir sihirli kare için 16 farklı sayı kullanılır. Oysa 4x4 lük bir latin kare için yalnız 4 farklı sayı gerekir.

7. POSTA PULLARI ÜZERİNDE LATİN KARELER

İlk posta pulu Mayıs 1840 yılında İngiltere'de basılmıştır. 1840'tan günümüze kadar binlerce pul basıldı, kataloglandı, koleksiyoncularca toplandı. Her bir harf bir satırda ya da bir sütunda yalnız bir kez kullanılacaktır.

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix}, \begin{bmatrix} a & b & c \\ b & c & a \\ c & a & b \end{bmatrix}, \begin{bmatrix} a & b & c & d \\ b & c & d & a \\ c & d & a & b \\ d & a & b & c \end{bmatrix}, \dots$$

2x2, 3x3, 4x4, 5x5 latin kare formatında posta pulları hazırlayan tek ülke Pakistandır. Bu yazımızda da Pakistan, (2x2), Malawi (4x4), Latin kareler seçildi (Chu ve ark. 2009).

Aşağıda 2x2'lik bir Latin kare verilmiştir. Türkiye Cumhuriyetinin kurucusu Mustafa Kemal Atatürk (1881--1938), ve Pakistan'ın kurucusu Quaid-e-Azam Muhammad Ali Jinnah (1876—1948). Türkiye Büyük Millet Meclisinin kuruluşunun 85. yıldönümünde Pakistan'da bastırılan $k=2$ için dördü bir arada pullar.

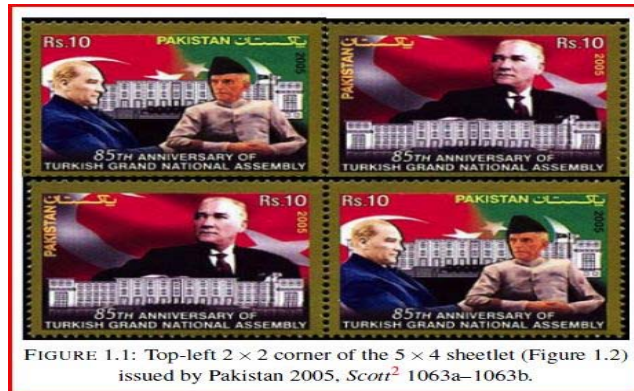


FIGURE 1.1: Top-left 2 x 2 corner of the 5 x 4 sheetlet (Figure 1.2) issued by Pakistan 2005, *Scott*² 1063a–1063b.

Şekil 10. 2x2 Latin Kare Tasarımı

Aşağıda 4×4'lük Latin kare görülmektedir.



Figure 4. Lilian's lovebird: Malawi 2009, PLS type a432.

Şekil 11. Muhabbet Kuşları: Malawi (Güneydoğu Afrika ülkesi)

Aynı pul (1 nolu) her bir hücrede asıl köşegende bulunmaktadır.

$$L_2 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}, \quad L_3 = \begin{bmatrix} A & B \\ B & A \end{bmatrix}$$

Sayılarla yazılan yukarıdaki 4×4 Latin kare, aynı zamanda 2×2 blok- Latin karedir (Chu ve ark. 2009).

5×5 Latin kare, 25 Ağustos-11 Eylül 1960 Roma yaz olimpiyatlarındaki **engelli koşu, futbol, engelli atlama (binicilik), basketbol ve güreş** yarışmalarını gösteren pullarla oluşturulan 5×5 Latin kare tasarımının sayılarla ifadesi:

$$L_5 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \\ 4 & 5 & 1 & 2 & 3 \\ 3 & 4 & 5 & 1 & 2 \\ 2 & 3 & 4 & 5 & 1 \end{bmatrix},$$

dir. 5×5 biçiminde pullarla elde edilen ilk Latin karedir (Loly and Styan, 2010). Türkiye bu olimpiyatlarda tümü güreşte olmak üzere 2 gümüş ve 7 altın madalya kazanmıştır. Aşağıda 1960 Roma Yaz Olimpiyatları anısına bastırılan 5×5 Latin Kare tasarımı ile uyumlu pulları görülmektedir.



Şekil 12. 1960 Roma Yaz Olimpiyatları anısına bastırılan 5×5 Latin Kare tasarımı ile uyumlu pullar

8. SONUÇ

Sihirli kareler 4000 yıldır sihirbazlık dünyasında kullanılmakta ve bilim dünyası için de ilgi odağı olmaktadır. Orijinal olarak bir matematiksel merak konusu olarak görünen fakat istatistik literatüründe 20. yüzyıl başından itibaren bir istatistiksel tasarım modeli olarak kullanılan LATİN KARELER deki matematiksel güzelliği, sihirli karelerdeki gizemi pullardan da yararlanarak sizlerle paylaşmak istedim. Yazımızı matematik eğitimcisi William Leonard Schaaf (1898-1992) 1978 yılında basılan “Matematik ve Bilim” adlı kitabındaki görüşü ile tamamlamak istiyorum. **“Dünyadaki posta pulları gerçekten uygarlığın aynasıdır.”**

9. KAYNAKLAR

- Akdeniz, F., 2011. Sihirli karelerin gizeminden Nye Gezinti İstatistik Dergisi, Yıl:1, Sayı: 3, 12-17.
- Ataöv, T., 2002. Hintli Bir Matematik Dahisi. Cumhuriyet Bilim Teknik, Sayı: 773.
- Booth, A. D., Booth, K. H. V. 1955. On Magic Squares. The Mathematical Gazette, 39, 132-133.
- Henrich, C. J., 1991. Magic squares and linear algebra. Amer. Math. Monthly 98, 481-488.

Ka Lok Chu, Puntanen, S., Styan, G. P. H., 2009. Some comments on philatelic Latin squares from Pakistan, Pak. J. Statist. 25(4), 427-471.

Ka Lok Chu, Drury, S. W., Styan, G. P. H., Trenkler, G., 2011. Magic Moore- Penrose inverses and philatelic magic squares with special emphasis on the Daniels-Zlobec magic square, Croatian Operational Research Review 2, 4-13.

Loly, P. D., Styan, G. P. H., 2010. Comments on 5x5 Philatelic Latin Squares CHANCE Vol.23(2), 1-10.

Pasles, P. C., 2001. The lost squares of Dr. Franklin: Ben Franklin's missing squares and the secret of the magic circle. The American Mathematical Monthly, 108, 489-511.

Rao, C. R., Mitra, S. K., 1971. Generalized Inverse of Matrices and its Applications, New York: John Wiley.

Schmidt, K. and Trenkler, G., 2001. The Moore-Penrose inverse of a semi- magic square is semi-magic. International Journal of Mathematical Education in Science and Technology, 32, 624-629.

Styan, G. H. P. Boyer, C., Ka Lok C., 2009. Some comments on Latin squares and on Graeco- Latin squares, illustrated with postage stamps and old playing cards. Stat. Papers 50: 917-941.

Trenkler, G., 1994. Singular magic squares. International International Journal of Mathematical Education in Science and Technology, 25, 595-597.

William L. S., 1978. Mathematics and Science: An Adventure in Postage Stamps National Council of Teachers of Mathematics, Reston, Virginia, 2.

MAGIC SQUARES AND LATIN SQUARES

ABSTRACT

In this paper, magic squares which have been known for 4000 years are presented. Especially, Moore-Penrose inverse and eigenvalue properties of $2n \times 2n$ ($n \geq 2$) singular magic squares are investigated. Furthermore, Latin square design models are illustrated visually by special stamps.

Keywords: Latin square, Moore-Penrose inverse, Magic constant, Magic square, Semi-magic square.

STATISTICAL INFERENCES ON SOME TRIANGULAR DISTRIBUTIONS: CASE OF BOUNDARY VALUES BEING PARAMETERS

İsmail ERDEM*

ABSTRACT

If continuous random variables X has a triangular distribution and if its boundary values θ_1 and/or θ_2 are unknown, then it may well be necessary to make some statistical inferences related to these parameters. For triangular distributions, with unknown boundary values, some estimators, as functions of ordered statistics, are proposed. The proposed estimators are compared based on their efficiencies. Based on efficiency criteria, the best estimator, among the proposed estimators, is determined. By the use of the best estimator, a confidence interval construction and the test of hypotheses procedures are developed. By means of a simulation process, matching accuracy between sampling results and theoretical findings is observed.

Keywords: Efficiencies of estimators, Estimators, Ordered statistics, Simulation, Triangular distribution.

1. INTRODUCTION

1.1 Type I Triangular Distribution

If a continuous random variable X has the following probability density function (pdf), its distribution will be named as Type I Triangular distribution.

$$f(x) = \frac{2}{(\theta - a)^2} (x - a), a < x < \theta, (a \text{ is a known constant}) \quad (1)$$

As it is shown below, all the moments of this distribution are the functions of the same parameter θ (upper bound value of X). If θ is unknown, then it must be estimated.

$$E(X^k) = \int_a^\theta x^k \frac{2}{(\theta - a)^2} (x - a) dx = \frac{2}{(\theta - a)^2} \left[\frac{x^{k+2}}{k+2} - \frac{ax^{k+1}}{k+1} \right] \Big|_{x=a}^\theta$$

$$E(X^k) = \frac{2}{(\theta - a)^2} \left[\frac{\theta^{k+2} - a^{k+2}}{k+2} - \frac{a\theta^{k+1} - a^{k+2}}{k+1} \right] \quad (2)$$

If we let $Y = (X - a)$, then the distribution of Y turns out to be a special Type I triangular distribution. Y assumes values in the interval $(0, \lambda = (\theta - a))$ and its pdf will be as given in (4).

*Prof. Dr., Başkent University, Faculty of Science and Letters, Department of Statistics and Computer Science, Ankara, e-mail: iserdem@baskent.edu.tr

$$F(y) = P(Y \leq y) = P(X \leq a + y) = \int_a^{a+y} \frac{2}{(\theta - a)^2} (x - a) dx = \frac{2}{(\theta - a)^2} \left[\frac{x^2}{2} - ax \right]_{x=a}^{x=a+y}$$

$$F(y) = \frac{y^2}{(\theta - a)^2}, 0 < y < (\theta - a) = \lambda \quad (3)$$

$$f(y) = \frac{2y}{\lambda^2}, 0 < y < \lambda \quad (4)$$

1.1.1. Estimation of the Parameter λ of the Random Variable $Y = (X - a)$ by the Largest Ordered Statistics $Y_{(n)}$

Let Ordered statistics obtained from a random sample of size n , taken from the pdf given in (4), be $Y_{(1)} < Y_{(2)} < \dots < Y_{(n)}$.

The pdf of $Y_{(n)}$ is as given below.

$$f_{Y_{(n)}}(y) = n[F(y)]^{n-1} f(y) = n \left[\frac{y^2}{\lambda^2} \right]^{n-1} \frac{2y}{\lambda^2} = \frac{2n}{\lambda^{2n}} y^{2n-1}, 0 < y < \lambda$$

The expected value and variance of $Y_{(n)}$ are as follows.

$$E(Y_{(n)}) = \frac{2n}{\lambda^{2n}} \int_0^\lambda y y^{2n-1} dy = \frac{2n}{\lambda^{2n}} \left[\frac{y^{2n+1}}{2n+1} \right]_0^\lambda = \frac{2n\lambda}{(2n+1)} \quad (5)$$

$$E(Y_{(n)}^2) = \frac{2n}{\lambda^{2n}} \int_0^\lambda y^2 y^{2n-1} dy = \frac{2n}{\lambda^{2n}} \left[\frac{y^{2n+2}}{2n+2} \right]_0^\lambda = \frac{2n\lambda^2}{(2n+2)} \quad (6)$$

$$Var(Y_{(n)}) = \frac{2n\lambda^2}{(2n+2)} - \frac{4n^2\lambda^2}{(2n+1)^2} = \frac{n\lambda^2}{(n+1)(2n+1)^2} \quad (7)$$

Since, $Y = (X - a)$ then

$$E(X_{(n)}) = E(Y_{(n)}) + a = \frac{2n\lambda}{(2n+1)} + a = \frac{2n\theta + a}{(2n+1)} \quad (8)$$

$$Var(X_{(n)}) = Var(Y_{(n)}) = \frac{n\lambda^2}{(n+1)(2n+1)^2} = \frac{n(\theta - a)^2}{(n+1)(2n+1)^2} \quad (9)$$

By the use of equation (8) we can obtain an unbiased estimator for the parameter θ of the random variable X . Hence,

$$T_1 = \frac{(2n+1)X_{(n)} - a}{2n} \quad (10)$$

$$Var(T_1) = \frac{(2n+1)^2}{4n^2} Var(X_{(n)}) = \frac{(\theta - a)^2}{4n(n+1)} \quad (11)$$

1.2 Estimation of the Parameter θ of the Random Variable X by the Sample Mean $\bar{X}$

We know that, the pdf of Y is $f(y) = \frac{2y}{\lambda^2}, 0 < y < \lambda$, then

$$E(Y) = \int_0^{\lambda} y \frac{2y}{\lambda^2} dy = \frac{2y^3}{3\lambda^2} \Big|_0^{\lambda} = \frac{2\lambda}{3} \quad (12)$$

$$E(Y^2) = \int_0^{\lambda} y^2 \frac{2y}{\lambda^2} dy = \frac{2y^4}{4\lambda^2} \Big|_0^{\lambda} = \frac{\lambda^2}{2} \quad (13)$$

$$Var(Y) = \frac{\lambda^2}{2} - \frac{4\lambda^2}{9} = \frac{\lambda^2}{18} \quad (14)$$

Since $Y = (X - a)$, $E(X) = E(Y) + a$, and $Var(X) = Var(Y)$.

$$E(X) = \frac{2\lambda}{3} + a = \frac{2(\theta - a)}{3} + a = \frac{2\theta + a}{3} \quad (15)$$

$$Var(X) = \frac{\lambda^2}{18} = \frac{(\theta - a)^2}{18} \quad (16)$$

For any random variable X, $E(\bar{X}) = E(X)$, and $Var(\bar{X}) = \frac{Var(X)}{n}$. Then for above

Type I Triangular distribution the following hold true.

$$E(\bar{X}) = \frac{2\theta + a}{3}, \text{ and } Var(\bar{X}) = \frac{(\theta - a)^2}{18n} \quad (17)$$

An unbiased estimator for θ as a function of $\bar{X}$ is as given below.

$$T_2 = \frac{3\bar{X} - a}{2} \text{ and its variance is } Var(T_2) = \frac{9}{4} Var(\bar{X}) = \frac{(\theta - a)^2}{8n}.$$

1.3 Comparisons of Unbiased Estimators on the Basis of Efficiency

$$T_1 = \frac{(2n+1)X_{(n)} - a}{2n}, \quad Var(T_1) = \frac{(\theta - a)^2}{4n(n+1)}$$

$$T_2 = \frac{3\bar{X} - a}{2}, \quad Var(T_2) = \frac{(\theta - a)^2}{8n}$$

For any integer $n > 1$, $Var(T_1) < Var(T_2)$. Hence $T_1 = \frac{(2n+1)X_{(n)} - a}{2n}$ is preferred over

$$T_2 = \frac{3\bar{X} - a}{2}.$$

1.4 Confidence Interval for θ of the Random Variable X

The better unbiased estimator for θ is $T_1 = \frac{(2n+1)X_{(n)} - a}{2n}$ and it is a linear function of $X_{(n)}$.

Hence, in the construction of a confidence interval it may be reasonable to use the pdf $X_{(n)}$.

From the pdf of X, $f(x) = \frac{2}{(\theta-a)^2}(x-a)$, $a < x < \theta$, the following are obtained.

$$F(x) = \int_a^x \frac{2}{(\theta-a)^2}(y-a)dy = \frac{2}{(\theta-a)^2} \left[\frac{y^2-a^2}{2} - ay + a^2 \right] = \frac{(x-a)^2}{(\theta-a)^2} \quad (18)$$

The pdf of $X_{(n)}$ is as given below.

$$f_{X_{(n)}}(x) = n \left[\frac{(x-a)^2}{(\theta-a)^2} \right]^{n-1} \frac{2}{(\theta-a)^2}(x-a) = \frac{2n}{(\theta-a)^{2n}}(x-a)^{2n-1}, a < x < \theta \quad (19)$$

From the probability statement, $P(x_{nL} < X_{(n)} < x_{nU}) = 1 - \alpha$, lower and upper confidence limits may be obtained.

$$P(X_{(n)} \leq x_{nL}) = \frac{2n}{(\theta-a)^{2n}} \int_a^{x_{nL}} (x-a)^{2n-1} dx = \alpha/2 \quad (20)$$

In (20) let $x-a = t$: then $dt = dx$; $x=a \rightarrow t=0$; $x=x_{nL} \rightarrow t=x_{nL}-a$.

$$P(X_{(n)} \leq x_{nL}) = \frac{2n}{(\theta-a)^{2n}} \int_a^{x_{nL}} (x-a)^{2n-1} dx = \frac{2n}{(\theta-a)^{2n}} \int_0^{x_{nL}-a} t^{2n-1} dt =$$

$$\frac{2n}{(\theta-a)^{2n}} \frac{t^{2n}}{2n} \Big|_{t=0}^{x_{nL}-a} = \left(\frac{x_{nL}-a}{\theta-a} \right)^{2n} = \frac{\alpha}{2}$$

x_{nL} is computed as given in (21).

$$x_{nL} = a + (\theta-a) \left(\frac{\alpha}{2} \right)^{1/2n} \quad (21)$$

Similarly, the upper confidence limit is obtained as in (22).

$$P(X \leq x_{nU}) = \frac{2n}{(\theta-a)^{2n}} \int_a^{x_{nU}} (x-a)^{2n-1} dx = \frac{2n}{(\theta-a)^{2n}} \int_0^{x_{nU}-a} t^{2n-1} dt =$$

$$\frac{2n}{(\theta-a)^{2n}} \frac{t^{2n}}{2n} \Big|_{t=0}^{x_{nU}-a} = \left(\frac{x_{nU}-a}{\theta-a} \right)^{2n} = 1 - \frac{\alpha}{2}$$

$$x_{nU} = a + (\theta - a) \left(1 - \frac{\alpha}{2}\right)^{1/2n} \quad (22)$$

If (21) and (22) are substituted in the probability statement, $P(x_{nL} < X_{(n)} < x_{nU}) = 1 - \alpha$, the following is obtained.

$$P\left(a + (\theta - a) \left(\frac{\alpha}{2}\right)^{1/2n} < X_{(n)} < a + (\theta - a) \left(1 - \frac{\alpha}{2}\right)^{1/2n}\right) = 1 - \alpha \quad (23)$$

If the inequalities in (23) are solved for θ simultaneously, we reach the following expression.

$$P\left(a + \frac{X_{(n)} - a}{(1 - \alpha/2)^{1/2n}} < \theta < a + \frac{X_{(n)} - a}{(\alpha/2)^{1/2n}}\right) = 1 - \alpha \quad (24)$$

Hence, a $100 * (1 - \alpha)\%$ confidence interval for θ can be given by the use of the following formula.

$$\left(a + \frac{X_{(n)} - a}{(1 - \alpha/2)^{1/2n}}, \quad a + \frac{X_{(n)} - a}{(\alpha/2)^{1/2n}}\right) \quad (25)$$

1.5 Tests of Hypotheses Related to the Parameter θ of Type I Triangular

In testing $H_0 : \theta = \theta_0$ against to any alternative hypothesis, $X_{(n)} = X_{Max}$ may be used as a proper test statistic. For the chosen level of significance α , the decision rules given in the following table are applicable.

$H_0 : \theta = \theta_0$ $H_1 : \theta \neq \theta_0$	$H_0 : \theta \leq \theta_0$ $H_1 : \theta > \theta_0$	$H_0 : \theta \geq \theta_0$ $H_1 : \theta < \theta_0$
If $x_{(n)} \geq x_{nU}$ or $x_{(n)} \leq x_{nL}$ H_0 is rejected	If $x_{(n)} \geq x_{nU}$ H_0 is rejected	If $x_{(n)} \leq x_{nL}$ H_0 is rejected
Don't reject H_0 otherwise	Don't reject H_0 otherwise	Don't reject H_0 otherwise
Where, $x_{nL} = a + (\theta_0 - a) \left(\frac{\alpha}{2}\right)^{1/2n}$ and $x_{nU} < a + (\theta_0 - a) \left(1 - \frac{\alpha}{2}\right)^{1/2n}$	Where, $x_{nU} < a + (\theta_0 - a) (1 - \alpha)^{1/2n}$	Where, $x_{nL} = a + (\theta_0 - a) (\alpha)^{1/2n}$

2. STATISTICAL INFERENCES RELATED TO TYPE II TRIANGULAR DISTRIBUTION

If a random variable X has the following pdf, it is said to have a Type II triangular distribution.

$$f(x) = \frac{2}{(a-\theta)^2} (a-x), \theta < x < a \quad (26)$$

All the moments of this distribution are functions of the parameter θ . For this reason, θ need to be estimated.

$$E(X^k) = \int_{\theta}^a x^k \frac{2}{(a-\theta)^2} (a-x) dx = \frac{2}{(a-\theta)^2} \left[\frac{ax^{k+1}}{k+1} - \frac{x^{k+2}}{k+2} \right]_{x=\theta}^a$$

$$E(X^k) = \frac{2}{(a-\theta)^2} \left[\frac{a^{k+2} - a\theta^{k+1}}{k+1} - \frac{a^{k+2} - \theta^{k+2}}{k+2} \right] \quad (27)$$

If we let $Y = (X - \theta)$, then the distribution of Y turns out to have a special Type II triangular distribution. Y assumes values in the interval $(0, \lambda = (a - \theta))$ and its pdf will be as given in (28).

Estimation of $\lambda = (a - \theta)$ will enable to estimate θ .

$$F(y) = P(Y \leq y) = P(X \leq \theta + y) = \int_{\theta}^{\theta+y} \frac{2}{(a-\theta)^2} (a-x) dx = \frac{2}{(a-\theta)^2} \left[ax - \frac{x^2}{2} \right]_{x=\theta}^{\theta+y}$$

$$F(y) = \frac{2(a-\theta)y - y^2}{(a-\theta)^2} = \frac{2\lambda y - y^2}{\lambda^2}, 0 < y < (a-\theta) = \lambda \quad (28)$$

$$f(y) = \frac{2(a-\theta) - 2y}{\lambda^2} = \frac{2(\lambda - y)}{\lambda^2}, 0 < y < \lambda \quad (29)$$

2.1 Estimation of the Parameter λ of the Random Variable by the First Ordered Statistics $Y_{(1)}$

Let Ordered statistics obtained from a random sample of size n , taken from the pdf given in (29) be $Y_{(1)} < Y_{(2)} < \dots < Y_{(n)}$.

The pdf of $Y_{(1)}$ is as given below.

$$f_{Y_{(1)}}(y) = n[1 - F(y)]^{n-1} f(y) = n \left[1 - \frac{2\lambda y - y^2}{\lambda^2} \right]^{n-1} \frac{2(\lambda - y)}{\lambda^2} =$$

$$\frac{2n}{\lambda^{2n}} (\lambda - y)^{2n-1}, 0 < y < \lambda \quad (30)$$

The expected value and variance of $Y_{(1)}$ are as follows.

$$E(Y_{(1)}) = \int_0^{\lambda} y \frac{2n(\lambda - y)^{2n-1}}{\lambda^{2n}} dy \quad (31)$$

In (32), if we let $(\lambda - y) = t$:

$$dt = -dy; y = 0 \rightarrow t = \lambda; y = \lambda \rightarrow t = 0$$

$$E(Y_{(1)}) = \int_0^{\lambda} y \frac{2n(\lambda - y)^{2n-1}}{\lambda^{2n}} dy = \frac{2n}{\lambda^{2n}} \int_0^{\lambda} (\lambda - t)t^{2n-1} dt =$$

$$\frac{2n}{\lambda^{2n}} \left[\frac{\lambda t^{2n}}{2n} - \frac{t^{2n+1}}{2n+1} \right]_{t=0}^{\lambda} = 2n\lambda \left[\frac{1}{2n} - \frac{1}{2n+1} \right] = \frac{\lambda}{2n+1}$$

$$E(Y_{(1)}) = \frac{\lambda}{2n+1} \quad (32)$$

$$E(Y_{(1)}^2) = \int_0^{\lambda} y^2 \frac{2n(\lambda - y)^{2n-1}}{\lambda^{2n}} dy = \frac{2n}{\lambda^{2n}} \int_0^{\lambda} (\lambda - t)^2 t^{2n-1} dt =$$

$$\frac{2n}{\lambda^{2n}} \left[\frac{\lambda t^{2n}}{2n} - \frac{2\lambda t^{2n+1}}{2n+1} + \frac{t^{2n+2}}{2n+2} \right]_{t=0}^{\lambda} = 2n\lambda^2 \left[\frac{1}{2n} - \frac{2}{2n+1} + \frac{1}{2n+2} \right] = \frac{\lambda^2}{(2n+1)(n+1)}$$

$$E(Y_{(1)}^2) = \frac{\lambda^2}{(2n+1)(n+1)} \quad (33)$$

By the use of the results in (32) and (33) we obtain

$$Var(Y_{(1)}) = \frac{\lambda^2}{(n+1)(2n+1)} - \frac{\lambda^2}{(2n+1)^2} = \frac{n\lambda^2}{(n+1)(2n+1)^2} \quad (34)$$

Since $Y = (X - \theta)$,

$$E(X_{(1)}) = E(Y_{(1)}) + \theta = \frac{\lambda}{(2n+1)} + \theta = \frac{2n\theta + a}{(2n+1)} \quad (35)$$

$$Var(X_{(1)}) = Var(Y_{(1)}) = \frac{n\lambda^2}{(n+1)(2n+1)^2} = \frac{n(a - \theta)^2}{(n+1)(2n+1)^2} \quad (36)$$

From the equation (35), we can obtain an unbiased estimator for θ as a function of $X_{(1)}$ as given in (37) and the variance of T_1 are obtained as in (38).

$$T_1 = \frac{(2n+1)X_{(1)} - a}{2n} \quad (37)$$

$$Var(T_1) = \frac{(2n+1)^2}{4n^2} Var(X_{(1)}) = \frac{(a - \theta)^2}{4n(n+1)} \quad (38)$$

2.2 Estimation of the Parameter θ of X by the Sample Mean $\bar{X}$

$$f(y) = \frac{2(a-\theta)-2y}{\lambda^2} = \frac{2(\lambda-y)}{\lambda^2}, 0 < y < \lambda$$

$$E(Y) = \int_0^\lambda y \frac{2(\lambda-y)}{\lambda^2} dy = \frac{2}{\lambda^2} \left[\frac{\lambda y^2}{2} - \frac{y^3}{3} \right]_{y=0}^\lambda = \frac{2}{\lambda^2} \left[\frac{\lambda^3}{2} - \frac{\lambda^3}{3} \right] = \frac{\lambda}{3} \quad (39)$$

$$E(Y^2) = \int_0^\lambda y^2 \frac{2(\lambda-y)}{\lambda^2} dy = \frac{2}{\lambda^2} \left[\frac{\lambda y^3}{3} - \frac{y^4}{4} \right]_{y=0}^\lambda = \frac{2}{\lambda^2} \left[\frac{\lambda^4}{3} - \frac{\lambda^4}{4} \right] = \frac{\lambda^2}{6} \quad (40)$$

$$Var(Y) = \frac{\lambda^2}{6} - \frac{\lambda^2}{9} = \frac{\lambda^2}{18} \quad (41)$$

Since $Y = (X - \theta)$, $E(X) = E(Y) + \theta$, and $Var(X) = Var(Y)$.

$$E(X) = \frac{\lambda}{3} + \theta = \frac{(a-\theta)}{3} + \theta = \frac{2\theta+a}{3} \quad (42)$$

$$Var(X) = \frac{\lambda^2}{18} = \frac{(a-\theta)^2}{18} \quad (43)$$

For any random variable X , $E(\bar{X}) = E(X)$, and $Var(\bar{X}) = \frac{Var(X)}{n}$. Then for the above Type II Triangular distribution the followings hold true.

$$E(\bar{X}) = E(X) = \frac{2\theta+a}{3}, \text{ and } Var(\bar{X}) = \frac{1}{n} Var(X) = \frac{(a-\theta)^2}{18n}.$$

An unbiased estimator for θ , as a function of $\bar{X}$, and its variance are as given below.

$$T_2 = \frac{3\bar{X} - a}{2} \quad (44)$$

$$Var(T_2) = \frac{9}{4} Var(\bar{X}) = \frac{(a-\theta)^2}{8n} \quad (45)$$

2.4 Comparisons of Unbiased Estimators on the Basis of Efficiency

$$T_1 = \frac{(2n+1)X_{(1)} - a}{2n}, \quad Var(T_1) = \frac{(\theta-a)^2}{4n(n+1)}$$

$$T_2 = \frac{3\bar{X} - a}{2}, \quad Var(T_2) = \frac{(\theta-a)^2}{8n}$$

For any $n > 1$ $Var(T_1) < Var(T_2)$. $T_1 = \frac{(2n+1)X_{(1)} - a}{2n}$ is a better unbiased estimator.

2.5 Confidence Interval for θ of the Type II Triangular Distribution

Since we let $Y = (X - \theta)$ and its pdf is obtained in (30) as given below

$$f_{Y_{(i)}}(y) = \frac{2n}{\lambda^{2n}} (\lambda - y)^{2n-1}, 0 < y < \lambda.$$

From the probability statement, $P(x_{1L} < X_{(1)} < x_{1U}) = 1 - \alpha$, lower and upper confidence limits may be obtained.

$$P(x_{1L} < X_{(1)} < x_{1U}) = P(x_{1L} < Y_{(1)} + \theta < x_{1U}) = P(x_{1L} - \theta < Y_{(1)} < x_{1U} - \theta) = 1 - \alpha$$

If we let $(x_{1L} - \theta) = y_{1L}$ and $(x_{1U} - \theta) = y_{1U}$

$$P(Y_{(1)} < y_{1L}) = \frac{2n}{\lambda^{2n}} \int_0^{y_{1L}} (\lambda - y)^{2n-1} dy = \alpha / 2 \quad (46)$$

$$P(Y_{(1)} < y_{1U}) = \frac{2n}{\lambda^{2n}} \int_0^{y_{1U}} (\lambda - y)^{2n-1} dy = 1 - \alpha / 2 \quad (47)$$

In (46) and (47), let $(\lambda - y) = t$:

$$dt = -dy; \quad y = 0 \rightarrow t = \lambda; \quad y = y_{1L} \rightarrow t = \lambda - y_{1L}$$

$$P(Y_{(1)} < y_{1L}) = \frac{2n}{\lambda^{2n}} \int_0^{y_{1L}} (\lambda - y)^{2n-1} dy = \frac{2n}{\lambda^{2n}} \int_{\lambda - y_{1L}}^{\lambda} t^{2n-1} dt = \left(\frac{\lambda}{\lambda - y_{1L}} \right)^{2n} = \alpha / 2 \quad (48)$$

Similarly,

$$P(Y_{(1)} < y_{1U}) = \frac{2n}{\lambda^{2n}} \int_0^{y_{1U}} (\lambda - y)^{2n-1} dy = \frac{2n}{\lambda^{2n}} \int_{\lambda - y_{1U}}^{\lambda} t^{2n-1} dt = \left(\frac{\lambda}{\lambda - y_{1U}} \right)^{2n} = 1 - \alpha / 2 \quad (49)$$

From the equations (48) and (49), y_{1L} and y_{1U} can be computed as given below.

$$y_{1L} = \lambda - (\lambda) \left(\frac{\alpha}{2} \right)^{-1/2n} = (a - \theta) - (a - \theta) \left(\frac{\alpha}{2} \right)^{-1/2n} \quad (50)$$

$$y_{1U} = \lambda - (\lambda) \left(1 - \frac{\alpha}{2} \right)^{-1/2n} = (a - \theta) - (a - \theta) \left(1 - \frac{\alpha}{2} \right)^{-1/2n} \quad (51)$$

Since we have $(x_{1L} - \theta) = y_{1L}$ and $(x_{1U} - \theta) = y_{1U}$, then

$$x_{1L} = y_{1L} + \theta \text{ and } x_{1U} = y_{1U} + \theta.$$

$$x_{1L} = (a) - (a - \theta) \left(\frac{\alpha}{2} \right)^{1/2n} \quad (52)$$

$$x_{1U} = (a) - (a - \theta) \left(1 - \frac{\alpha}{2} \right)^{-1/2n} \quad (53)$$

Substituting the results of (52) and (53) in the following probability expression yields (54).

$$P(x_{1L} < X_{(1)} < x_{1U}) = 1 - \alpha$$

$$P\left(a - (a - \theta)\left(\frac{\alpha}{2}\right)^{-1/2n} < X_{(1)} < a - (a - \theta)\left(1 - \frac{\alpha}{2}\right)^{-1/2n}\right) = 1 - \alpha \quad (54)$$

By solving inequalities in (54) for θ simultaneously we obtain the following probability statement.

$$P\left(a - (a - X_{(1)})(1 - \alpha/2)^{1/2n} < \theta < a - (a - X_{(1)})(\alpha/2)^{1/2n}\right) = 1 - \alpha \quad (55)$$

Hence, a $100 * (1 - \alpha)\%$ confidence interval for θ can be given by the use of the following formula.

$$\left(a - (a - X_{(1)})(1 - \alpha/2)^{1/2n}, \quad < a - (a - X_{(1)})(\alpha/2)^{1/2n}\right)$$

2.6 Tests of Hypotheses Related to the Parameter θ of Type II Triangular Distribution

In testing $H_0 : \theta = \theta_0$ against to any alternative hypothesis, $X_{(1)} = X_{Min}$ may be used as a proper test statistic. For the chosen level of significance α , the decision rules given in the following table are applicable.

$H_0 : \theta = \theta_0$ $H_1 : \theta \neq \theta_0$	$H_0 : \theta \leq \theta_0$ $H_1 : \theta > \theta_0$	$H_0 : \theta \geq \theta_0$ $H_1 : \theta < \theta_0$
If $x_{(1)} \geq x_{1U}$ or $x_{(1)} \leq x_{1L}$ H_0 is rejected.	If $x_{(1)} \geq x_{1U}$ H_0 is rejected.	If $x_{(1)} \leq x_{1L}$ H_0 is rejected.
Don't reject H_0 otherwise.	Don't reject H_0 otherwise.	Don't reject H_0 otherwise.
Where, $x_{1L} = a - (a - \theta_0)\left(\frac{\alpha}{2}\right)^{1/2n}$ and $x_{1U} < a - (a - \theta_0)\left(1 - \frac{\alpha}{2}\right)^{1/2n}$	Where, $x_{1U} < a - (a - \theta_0)(1 - \alpha)^{1/2n}$	Where, $x_{1L} = a - (a - \theta_0)(\alpha)^{1/2n}$

3. SIMULATION

For the purpose of simulations from both Type I, and Type II triangular distributions inverse transformations are used on the respective F(x) functions.

For the type I triangular distribution over the interval (5,10) [$\theta = 5$ and $a = 10$] the pdf is

$$f(x) = \frac{2(x-5)}{25}, 5 < x < 10. \text{ Hence we obtain}$$

$$F(x) = \frac{(x-5)^2}{25}, 5 < x < 10.$$

If a random sample is taken from the *Uniform*[0,1] we obtain the corresponding X value in the

type I triangular distribution by the use of the following formula

$$x = 5(1 + \sqrt{F(x)}).$$

The median, m, for this Type I triangular distribution is:

$$F(m) = \frac{(m-5)^2}{25} = \frac{1}{2} \rightarrow \frac{(m-5)}{5} = \frac{1}{\sqrt{2}} \rightarrow m = 5 + \frac{5}{\sqrt{2}} \cong 8.535$$

Similarly, for the type II triangular distribution over the interval (5, 10) [$a = 5$ and $\theta = 10$] the pdf is

$$f(x) = \frac{2(10-x)}{25}, 5 < x < 10, \text{ and the respective distribution function is}$$

$$F(x) = 1 - \frac{(10-x)^2}{25}, 5 < x < 10.$$

If a random sample is taken from the *Uniform*[0,1] we obtain the corresponding X value in the

type II triangular distribution by the use of the following formula

$$x = 10 - 5\sqrt{1 - F(x)}.$$

The median, m, for this Type II triangular distribution is:

$$F(m) = 1 - \frac{(10-m)^2}{25} = \frac{1}{2} \rightarrow \frac{(10-m)}{5} = \frac{1}{\sqrt{2}} \rightarrow m = 10 - \frac{5}{\sqrt{2}} \cong 6.464.$$

Table 1. 50 Random Samples of Sizes n=30, From Uniform (0, 1)

u_1	u_2	u_3	u_4	u_5	...	u_{48}	u_{49}	u_{50}
0,517788	0,498149	0,73755	0,231313	0,906975	...	0,357554	0,955998	0,259729
0,746209	0,260626	0,239369	0,473726	0,815237		0,171827	0,454524	0,764033
0,377066	0,449645	0,178527	0,453634	0,773392		0,483062	0,988522	0,711061
0,692188	0,387486	0,778406	0,122221	0,626263		0,261265	0,929536	0,805118
0,606739	0,406041	0,60322	0,43981	0,234233		0,759478	0,567909	0,919013
0,421063	0,729669	0,367489	0,722403	0,2267		0,750751	0,153411	0,369019
0,80357	0,124422	0,543059	0,808808	0,504451		0,324161	0,27868	0,516978

Table 2. 50 Random Samples of Sizes n=30, From TYPE I Triangular Distribution and Estimators:

$$x_i = 5(1 + \sqrt{u_i})$$

	X1	X2	X3	X4	X5		X49	X50		
	8,597876	8,528984	9,294036	7,404752	9,761761		9,888757	7,548182		
	9,31917	7,552579	7,446267	8,441389	9,514525		8,370919	9,37045		
	8,070286	8,352779	7,112625	8,367617	9,397135		9,971222	9,216223		
	8,24447	9,271034	8,031045	9,249713	7,380651		6,958389	8,037347		
	9,482104	6,763672	8,684627	9,496689	8,551237		7,639509	8,595061	Mean	Var
Mean	8,36463	8,312875	8,433408	8,721029	8,75695		8,426896	8,771766	8,4317	0,0460
Var	1,296027	1,174865	0,980048	1,171179	0,798496		1,739837	0,767044	1,2601	0,0940
Xmax	9,94276	9,89737	9,945686	9,860279	9,930852		9,971222	9,864761	9,921	0,0044
Xmin	5,166093	6,20382	6,297411	5,459905	6,89126		5,809413	6,676654	5,8748	0,2427
T1	10,02514	9,978993	10,02811	9,941284	10,01303		10,05408	9,94584	10,003	0,0045
T2	10,04695	9,969312	10,15011	10,58154	10,63542		10,14034	10,65765	10,147	0,1037
Med	8,672594	8,436987	8,651681	9,223838	8,953614		8,842737	8,854906	8,6634	0,0948
Tmed	10,19383	9,860634	10,16426	10,97341	10,59125		10,43445	10,45166	10,180	0,1897

The above simulation results are in accordance with the theoretical findings, for that the estimator T_1 is seen to estimate the true parameter θ of the Type I triangular distribution very accurately.

Table 3. 50 Random samples of sizes n=30, from TYPE II Triangular Distribution and Estimators:

$$x_i = 10 - 5\sqrt{1 - u_i}$$

	X1	X2	X3	X4		X48	X49	X50		
	6,527927	6,457928	7,438506	5,616261		5,992363	8,951165	5,69805		
	7,481117	5,700658	5,639291	6,372764		5,449799	6,307182	7,571179		
	6,053693	6,290705	5,468244	6,30417		6,405081	9,464323	7,312349		
	6,195606	7,400334	6,023473	7,365624		7,50376	5,399488	6,028285		
	7,783981	5,321382	6,62013	7,813727		5,88953	5,753473	6,525013	Mean	Var
Mean	6,651137	6,596583	6,683158	7,051849		6,528217	6,903038	7,029369	6,767	0,0464
Var	1,088856	1,187989	1,158918	1,211741		1,605636	1,807143	1,11921	1,335	0,0732
Xmax	9,245597	8,992148	9,265025	8,82625		9,072844	9,464323	8,844968	9,194	0,1313
Xmin	5,002759	5,147082	5,171261	5,021196		5,017774	5,06595	5,289498	5,102	0,0105
T1	5,002805	5,149533	5,174115	5,02155		5,018071	5,067049	5,294323	5,104	0,0109
T2	4,976705	4,894875	5,024737	5,577773		4,792326	5,354557	5,544054	5,151	0,1045
Median	6,607878	6,368598	6,584901	7,32477		6,198722	6,803171	6,815711	6,630	0,1153
Tmed	5,202815	4,864422	5,170321	6,216653		4,624181	5,479002	5,496736	5,234	0,2306

The above simulation results are in accordance with the theoretical findings, for that the estimator T_1 is seen to estimate the true parameter θ of the Type II triangular distribution very accurately.

4. REFERENCES

- Balakrishnan, N., Cohen, A. C., 1991. Order Statistics and Inference. New York: Academic Press.
- Balakrishnan, N., Rao, C. R. (Eds.), 1998. Handbook of Statistics, Vol. 16: Order Statistics: Theory and Methods. Amsterdam, Netherlands: Elsevier.
- Balakrishnan, N., Rao, C. R. (Eds.), 1998. Order Statistics: Applications. Amsterdam, Netherlands: Elsevier.
- David, H. A., Order Statistics, 2nd ed., 1981. New York: Wiley.
- Gibbons, J. D., Chakraborti, S. (Eds.), 1992. Nonparametric Statistic Inference, 3rd ed. exp. rev. New York: Dekker.
- Hogg, R. V., Craig, A. T., 1970. Introduction to Mathematical Statistics, 3rd ed. New York: Macmillan.
- Rose, C., Smith, M. D., 2002. Order Statistics. §9.4 in Mathematical Statistics with Mathematica. New York: Springer-Verlag, pp. 311-322.
- Rose, C., Smith, M. D., 2005. Computational Order Statistics. Mathematica J. 9, 790-802.

BAZI ÜÇGENSEL DAĞILIMLAR ÜZERİNE İSTATİSTİKSEL ÇIKARIMLAR: SINIR DEĞERLERİ PARAMETRE OLAN DURUM

ÖZET

Eğer sürekli rastgele değişkenler üçgensel dağılıma sahipse ve eğer θ_1 ve/veya θ_2 sınır değerleri bilinmiyorsa, o zaman bu parametrelerle ilgili bazı istatistiksel çıkarımlar yapmak gerekli olabilir. Bilinmeyen sınır değerleri olan üçgensel dağılımlar için, sıralı istatistiklerin bir fonksiyonu olarak bazı tahmin ediciler önerilmiştir. Önerilen tahmin ediciler etkinliklerine dayalı olarak karşılaştırılmıştır. Önerilen tahmin ediciler arasındaki en iyi tahmin edici, etkinlik kriterine göre belirlenmiştir. En iyi tahmin edicinin kullanılmasıyla, bir güven aralığı oluşturma ve hipotez testi izlekleri geliştirilmiştir. Simülasyon prosesi yoluyla, örneklem sonuçları ile teorik bulgular arasında eşleşen bir doğruluk gözlenmiştir.

Anahtar Kelimeler: Tahmin edicilerin etkinlikleri, Tahmin ediciler, Sıralı istatistikler, Simülasyon, Üçgensel dağılım.

SİSMİK TEHLİKENİN TAHMİNİ: OLASILIK VE İSTATİSTİK YÖNTEMLERİ

M. Semih YÜCEMEN*

ÖZET

Deprem oluşumlarının zaman, yer ve şiddet bakımından gösterdikleri rassallık ve çeşitli belirsizlikler nedeni ile sismik tehlikenin tahmininde olasılık ve istatistik yöntemlerine dayanan bir yaklaşım gereklidir. Ancak bu yöntemler, çoğunlukla, eldeki verilerin ve fiziksel olayların uygulanan stokastik modellere uyumları kontrol edilmeden kullanılmaktadır. Böyle bir yaklaşım hatalı sonuçlara yol açabilmekte, çoğu kez de uygulayıcılar bu durumun farkında olmamaktadır. Deterministik yaklaşımlara karşın, olasılık ve istatistik kuramları çerçevesinde geliştirilen bir yöntemin katkısı, yer hareketi değişkenleri için tek bir değer yerine bir değerler kümesi ile bu küme üzerinde tanımlanmış bir olasılık dağılımının belirlenmesi şeklinde olmaktadır. Bu bildiride sismik tehlike analizinin temelini oluşturan modeller özetlenerek, özellikle aktif faylara ağırlık verilerek geliştirilecek olan sismik tehlike haritalarının oluşturulmasında kullanılacak stokastik modeller üzerinde durulmuş ve yukarıda sözü edilen muhtemel hatalı uygulamalara dikkat çekilmiştir.

Anahtar Kelimeler: Olasılıksal sismik tehlike analizi, Ortogonal regresyon, Yinelenme modeli, Magnitüd-tekrerrür ilişkisi.

1. GİRİŞ

Sismik tehlike analizindeki aşamaların her birinin içerdiği rassal değişkenliklerden (aleatory) ve bilgi eksikliğinden (epistemic) kaynaklanan belirsizliklerin göz önünde bulundurulabilmesi için mutlaka olasılık ve istatistik yöntemlerinin kullanılması gerekir. Ancak bu yöntemler, çoğunlukla, eldeki verilerin ve fiziksel olayların, uygulanan stokastik modellere uyumları kontrol edilmeden kullanılmaktadır. Böyle bir yaklaşım hatalı sonuçlara yol açabilmekte, çoğu kez de uygulayıcılar bu durumun farkında olmamaktadırlar.

Deterministik yaklaşımlara karşın, olasılık ve istatistik çerçevesinde geliştirilecek bir yöntemin katkısı, zemin hareketi değişkenleri için tek bir değer yerine bir değerler kümesi ile bu küme üzerinde tanımlanmış bir olasılık dağılımının belirlenmesi şeklinde olmaktadır.

Son yıllarda aktif faylardan kaynaklanan sismik tehlikenin tahmini daha da önem kazanmıştır. Faylar üzerinde gerçekleştirilen kapsamlı çalışmalar neticesinde fayların ana özelliklerine ait elde edilen bilgiler ve verilerin de bunda önemli bir katkısı olmuştur. Yeni nesil sismik tehlike haritalarının hazırlanmasında aktif faylardan kaynaklanan sismik tehlike özellikle dikkate alınmalı ve faylarla ilgili mevcut bilgilerle uyumlu stokastik modeller kullanılmalıdır.

Belirli bir inşaat sahasındaki sismik tehlikenin belirlenmesi için uygulanacak olan olasılıksal sismik tehlike analizi (OSTA) disiplinlerarası bir çalışmayı gerektirmektedir.

*Prof. Dr., Orta Doğu Teknik Üniversitesi, İnşaat Mühendisliği Bölümü ve Deprem Çalışmaları EABD, e-posta: yucemen@metu.edu.tr

ve bu amaçla yürütülecek olan çalışmanın başlıca aşamaları aşağıda sıralanmıştır.

- (i) Etkilenme alanının tespiti ve bu alanda geçmişte meydana gelmiş depremlerle ilgili bilgileri içeren deprem kataloglarından yararlanarak bir sismik veri tabanının hazırlanması ve gerekli değişiklikleri ve ayarlamaları yaparak bu veri tabanının yeknesak ve yansız olmasının sağlanması.
- (ii) İncelenen bölgede alansal ve çizgisel sismik kaynakların tanımlanması. Aktif fay haritasının hazırlanması ve tanımlanan fayların özelliklerini ifade eden parametrelerin değerlerinin belirlenmesi.
- (iii) Deprem merkez-üstlerinin konumlarını ve bunların belirlenen aktif faylar ile ilişkilerini incelemek üzere bir sismotektonik haritanın çizilmesi. Sismik veri tabanında yer alan depremlerin merkez-üstlerinin konumuna göre sismik kaynaklara dağıtılması ve buna bağlı olarak da her sismik kaynak için elde edilecek magnitüd-tekerrür ilişkisine göre magnitüd için bir olasılık dağılımının çıkartılması ve diğer sismisite parametrelerinin değerlerinin saptanması. Belirlenen sismik kaynakların hiçbirisi ile ilişkilendirilemeyen depremlerin katkısını da dikkate almak üzere alansal geri plan sismik kaynakların tanımlanması.
- (iv) Depremlerin zaman içinde oluşumları için uygun bir stokastik modelin seçilmesi.
- (v) Yerel verilere dayanan bir yer hareketi tahmin (azalım) ilişkisinin çıkartılması ya da mevcutlar arasından uygun birinin seçilmesi.
- (vi) Göz önünde tutulan tüm sismik kaynakların inşaat sahasındaki sismik tehlikeye katkılarını birleştirecek bir hesaplama algoritmasına göre seçilen deprem şiddeti ölçüsü ya da zemin hareketi parametresi için olasılık dağılımının elde edilmesi. Sayısal hesaplamalar, bu amaçla hazırlanmış olan bilgisayar yazılımları kullanılarak yapılacağından uygun bir yazılım paketinin seçilmesi.
- (vii) Değişik türdeki belirsizliklerin değerlendirilmesinin ya doğrudan yapılması ya da duyarlılık analizleri ve mantık ağacı ya da benzeri yöntemler yolu ile bilgi eksikliğinden kaynaklanan (epistemik) belirsizliklerin etkilerinin sonuçlara yansıtılması.
- (viii) Belirli aşılma olasılıklarına (ya da tekerrür sürelerine) karşı gelen deprem yer hareketi değerlerinin elde edilmesi ve bu değerlere göre deprem tehlikesi haritalarının çizilmesi.

Burada bu aşamaların içinden sadece deprem veri tabanının hazırlanması, magnitüd-tekerrür ilişkisi ve depremlerin zaman içinde oluşumları için uygun bir stokastik modelin seçilmesi konuları ele alınmıştır. Söz konusu bu aşamaların her biri aşağıdaki bölümlerde daha ayrıntılı olarak ortaya konulmuştur.

2. DEPREM VERİ TABANININ OLUŞTURULMASI

Sismik veri tabanının oluşturulmasında deprem katalogları en önemli veri kaynağıdır. Ancak deprem kataloglarındaki veriler doğrudan kullanıma uygun olmayabilir. Genellikle deprem kataloglarında deprem magnitüdüleri değişik ölçeklerde verilmektedir. Bunların tek bir ölçeğe çevrilerek, magnitüd açısından yeknesak bir veri tabanının oluşturulması tavsiye edilmektedir. Poisson modelinin içerdiği bağımsızlık varsayımı nedeni ile öncü ve artçı depremlerin ayıklanması da gerekebilir. Diğer bir problem de, deprem kataloglarında yer alan küçük magnitüd değerli depremler ile çok uzun tekerrür

sürekli büyük magnitüdü depremlerin sayılarının eksik olmasının yarattığı yanlılıktır. Aşağıdaki alt-bölümlerde bu problemlerin çözümüne yönelik işlemler kısaca özetlenmiştir.

2.1 Deprem Veri Tabanının Tek Bir Magnitüd Ölçeğine Göre Oluşturulması

Kataloglarda yer alan değişik magnitüd ölçeklerinin tek bir magnitüd ölçeğine çevrilmesinde moment magnitüdünün (M_w) esas alınması uygun olacaktır. Farklı büyüklük ölçeklerine göre (cisim dalga magnitüdü – M_b , süre magnitüdü – M_d , yerel magnitüd – M_L ve yüzey magnitüdü – M_s) raporlanan deprem kayıtlarının M_w ölçeğine çevrilmesi oldukça önemli bir sorun teşkil etmektedir. Bu amaçla değişik ampirik dönüşüm ilişkileri geliştirilmiştir (örneğin Boore ve Joyner, 1982, Ulusay, v.d. 2004). Bu ilişkilerin elde edilmesinde yaygın olarak standart en küçük kareler regresyonu kullanılmıştır. Bu yöntem, aralarında bağıntı kurulacak değişkenlerden yalnızca bağımlı değişkende (M_w) hata (depremin rassal oluşumundan ileri gelen) olması durumunu göz önüne almaktadır. Ancak deprem büyüklüklerinin çeşitli nedenlerden kaynaklanan belirsizlikler yüzünden hatasız olarak belirlenmesi mümkün değildir. Dolayısıyla aralarında bağıntı kurulacak olan bağımlı ve bağımsız değişkenlerin her ikisinin de hata içermesi kaçınılmazdır. Böyle bir durumda regresyon analizi yapılabilmesi için ortogonal regresyon yönteminin kullanılması uygun olacaktır. Castellaro, v.d. (2006) standart regresyon yoluyla elde edilen dönüşüm denklemlerindeki yanlılığı incelemiş ve geliştirdikleri “Unified Italian Catalogue”da 0.4 magnitüd değerlerine erişen hataların olabileceğini belirtmişlerdir.

Deniz ve Yüçemen (2010), ortogonal regresyon yöntemini ve son yüzyıl içerisinde ülke çapında meydana gelmiş bütün depremlerden oluşan bir veri tabanını kullanarak bir dizi dönüşüm ilişkileri elde etmiştir. Bu ilişkiler Denklem 1’de gösterilmiştir.

$$M_w = 2.25 \times M_b - 6.14 \quad (1a)$$

$$M_w = 1.27 \times M_d - 1.12 \quad (1b)$$

$$M_w = 1.57 \times M_L - 2.66 \quad (1c)$$

$$M_w = 0.54 \times M_s + 2.81 \quad (1d)$$

Ortogonal regresyon, çevirim ilişkilerinin eğimlerini standart en küçük kareler yöntemine göre her zaman daha büyük tahmin etmektedir. Bu nedenle büyük depremlerin magnitüdlerini geleneksel yöntemlere göre daha büyük vermektedir. Küçük depremler için bunun tersi geçerli olmakla birlikte, bu depremlerin sismik tehlikeye katkısı zaten oldukça küçük seviyelerde kalmaktadır. Dolayısıyla deprem tehlikesinin tahmininde ortogonal regresyonun kullanılması durumunda geleneksel yöntemin sonuçlarına göre daha emniyetli tarafta değerler elde edilecektir.

2.2 Deprem Veri Tabanının İkincil Depremlerden Arındırılması

Olasılıksal sismik tehlike analizinde yaygın bir şekilde kullanılan Poisson modeli depremlerin gerek mekan, gerekse zaman açısından birbirlerinden bağımsız bir şekilde meydana geldikleri varsayımına dayanır. Poisson modelinin gerektirdiği bağımsızlık koşulunu sağlamak için deprem öbekleşmelerinin belirlenerek öncü ve artçı depremlerin (ikincil depremler) sismik veri tabanından çıkartılması gerekir.

Öncü ve artçı şokların zaman ve mekana göre belirlenmesi için birçok çalışmalar yapılmıştır (örneğin: Omori, 1894, Gardner ve Knopoff, 1974, Prozorov ve Dziewonski, 1982, Van Dyck, 1985, Utsu, v.d., 1995, Savage ve Rupp, 2000 ve Kagan, 2002). Öncü

ve artçı depremler zamansal ve mekansal olarak ana şok etrafında benzer dağılımlar göstermektedirler. Bu nedenle, ikincil depremlerin tayini öncü ve artçı depremler için farklılık göstermemektedir. Sözü geçen çalışmalar mühendislik uygulamaları için belirli bir büyüklük seviyesindeki depremlerin, deprem bölgesi, sismik kaynak, ilgili fayın uzunluğu ve çeşidi gibi ayrımlar gözetilmeksizin aynı ikincil deprem aktivitesine yol açtığını kabul eden çalışmalardır. Burada da her bir deprem büyüklüğü seviyesi için, bu seviyede bulunan bir ana şoka belirli bir zaman ve uzaklık penceresi içinde kalan bütün depremlerin ilgili ana şokun artçı depremleri olduğu kabul edilmiştir. Bir depremin öncü deprem sayılabilmesi için ise, kendi büyüklük seviyesi için belirlenmiş olan zaman ve uzaklık pencerelerinin içerisinde, kendisinden daha büyük bir deprem bulunması gerekmektedir. Böyle durumlarda magnitudü daha büyük olan ikinci depremin ana şok olduğu varsayılmıştır. Bu varsayımlara istisna olarak, yalnızca magnitudü 6.0'dan büyük olan bütün depremlerin ana şok olduğu kabul edilmiştir.

Deniz (2006), yukarıda belirtilen varsayımlar çerçevesinde ve Van Dyck (1985), Utsu, v.d. (1995), Savage ve Rupp (2000) ve Kagan (2002) tarafından yapılan çalışmalara dayanarak zaman ve mekan pencerelerinin boyutlarını belirlemiştir. Bu dört çalışmada verilen değerlere dayanarak artçı şoklar için uzaklık penceresi boyutlarını, Gardner ve Knopoff (1974) ve Savage ve Rupp (2000) tarafından verilen değerlerin ortalamasını alarak da zaman pencerelerinin boyutlarını tespit etmiştir. Elde edilen uzaklık ve zaman pencerelerinin boyutları Tablo 1'de verilmiştir. Tabloda yer almayan ara değerlerin hesabında, zaman için doğrusal, uzaklık için de log-doğrusal enterpolasyon kullanılması önerilmiştir. Tablo 1'de verilen değerler kısıtlı bir araştırmaya dayanarak elde edilmiş öneri niteliğindeki değerlerdir ve bu durum sözü edilen değerlerin kullanımında göz önünde tutulmalıdır.

Tablo 1. Öncü ve artçı depremlerin ayırt edilmesinde kullanılacak olan uzaklık ve zaman pencerelerinin boyutları (Deniz, 2006)

Magnitud	Uzaklık (km)	Zaman (gün)
4.5	35.5	42
5.0	44.5	83
5.5	52.5	155
6.0	63.0	290
6.5	79.4	510
7.0	100.0	790
7.5	125.9	1326
8.0	151.4	2471

2.3 Deprem Veri Tabanındaki Eksikliklerin (Yanlılığın) Giderilmesi

Üstel olasılık yoğunluk işlevinin parametrelerinin tahmini için kullanılacak deprem katalog verilerinin her magnitud düzeyinde eksiksiz olması gerekmektedir. Zaman içinde geriye doğru gidildikçe kataloglardaki deprem kayıtlarının hem kalitesi düşmekte hem de sayısı azalmaktadır. Yakın zaman içinde küçük, büyük tüm depremler kaydedilirken, çok eski kayıtlar sadece büyük depremleri içermektedir. Ayrıca kaydedilmiş depremler daima iskan edilmiş bölgelerde olup, insan yaşamından çok uzak yerde olan büyükçe depremlerin bile kayda geçmeme olasılığı vardır. Dolayısı ile deprem kataloglarındaki bu eksiklikler veri tabanında hem zamanda hem de mekanda yanlılıklara sebep olmaktadır. Buna bağlı olarak da bulunan tekerrür ilişkileri uzun süreli oluş sıklıklarını gerçekçi bir biçimde vermeyebilir. Bu nedenle, belirli bir

magnitüd aralığına düşen depremlerin eksiksiz olarak kayda geçirildiği zaman dilimini belirlemek gerekmektedir. Bu zaman dilimi belirlendikten sonra da o magnitüd aralığındaki depremlerin oluş sıklığı, sadece o zaman diliminde oluşan depremler göz önünde tutularak yapılacaktır.

Katalogda yer alan deprem sayılarının suni olarak, gözlemlerde mevcut eksikliklerden arındırılması için Stepp (1973) tarafından geliştirilmiş olan bir yöntem yaygın bir biçimde kullanılmaktadır. Bu yöntemde, depremler belirlenen magnitüd aralıklarına göre gruplandırılmakta ve her gruptaki depremlerin oluşumu noktasal bir süreç olarak alınmaktadır. Örneklem ortalamasının varyansının, örneklem sayısı ile ters orantılı olduğunu belirten temel istatistik kuralı bu yönteme esas teşkil etmektedir. Bu kurala göre gözlem sayısını çoğaltarak varyansı istenildiği kadar küçültmek mümkündür, yeter ki deprem kayıtları zaman içinde eksiksiz ve süreç de durağan olsun. Eğer deprem oluşumu durağan bir süreç ise ortalama değer, varyans ve diğer istatistiksel momentler sabit kalacaktır. Birim zamana isabet eden deprem sayıları $k_1, k_2, \dots, k_n$ ile gösterilecek olursa, bu örneklem grubu için yansız ortalama deprem sayısı:

$$\lambda = \frac{1}{n} \sum_{i=1}^n k_i \quad (2)$$

varyans da

$$\sigma_\lambda^2 = \frac{\lambda}{n} \quad (3)$$

olur. Denklem 3’de, n , birim zaman aralıklarının sayısıdır. Birim zaman aralığı bir yıl olarak alınırsa

$$\sigma_\lambda = \frac{\sqrt{\lambda}}{\sqrt{T}} = \sqrt{\frac{\lambda}{T}} \quad (4)$$

elde edilir. Burada, σ_λ ortalama değerin standart sapması, T ise örneklemin yıl cinsinden zaman aralığıdır. Eğer durağanlık varsayımı geçerli ise, göz önünde tutulan örnekleme, λ belirlenen bir magnitüd aralığında sabit kalacak ve σ_λ , $1/\sqrt{T}$ şeklinde değişecektir. Eğer λ sabit olursa, o zaman durağanlığın gerçekleşeceği zaman aralığı, ortalama değer için iyi bir tahmin oluşturacak kadar uzun, ama kayıtların eksik olduğu aralıkları içermeyecek kadar da kısa olmalıdır. Bu zaman aralığı belirlendikten sonra, seçilen magnitüd grubundaki depremlerin ortalama yıllık sayısı ise sadece o zaman aralığında oluşan depremler göz önünde tutularak yapılacaktır.

3. DEPREM MAGNİTÜD-TEKERRÜR MODELLERİ

3.1 Üstel Dağılım Modeli

Deprem magnitüdlerinin olasılık dağılımı, magnitüdlere ile bunların oluş sıklıkları arasındaki ilişkiyi gösteren tekerrür bağıntılarından çıkartılır. En yaygın kullanılan ilişki Richter (1958) tarafından önerilen aşağıdaki doğrusal magnitüd-sıklık ilişkisidir:

$$\log_{10} N(m) = a - bm \quad (5)$$

burada, $N(m)$ = birim zaman içinde magnitüd değeri m 'ye eşit ya da m 'den büyük ortalama deprem sayısı; a ve b = ilgili bölge için saptanan katsayılar; m = Richter magnitüdü ve $\log_{10} = 10$ tabanına göre logaritmadır. Eğer $\alpha = a(\ln 10)$ ve $\beta = b(\ln 10)$ olarak tanımlanırsa, Denklem 5 şu şekilde yazılabilir.

$$N(m) = e^{\alpha - \beta m} \quad (6)$$

Denklem 6'da α , bir bölgede olabilecek depremlerin toplam sayısına ilişkin bilgiyi yansıtmaktadır. α değerleri, genellikle göz önünde tutulan bölgenin büyüklüğü ve incelenen zaman süresi ile doğrudan doğruya ilgilidir. Buna karşılık β değerleri daha çok bölgenin tektonik yapısı ile ilişkilidir ve büyük magnitüdlü depremlerin küçükler olan göreceli oranını gösterir. Bu bakımdan, β değerleri bölgenin tektonik açıdan sismik etkinliğinin bir göstergesi olarak kabul edilir. β 'nın değeri sismik bölgeler arasında farklılık göstermektedir.

Genellikle, sismik tehlike analizlerinde magnitüd için m_0 gibi bir alt sınır saptanır. Alt sınır değerinden daha küçük magnitüdlü depremler mühendislik yapılarında bir hasar yaratamayacaklarından, bunlar sismik tehlike analizine katılmazlar. Ayrıca, m_0 'dan daha küçük depremler için istatistiksel veriler çoğu kez güvenilir değildir. Geçmiş deprem kayıtları, sonsuz enerjinin açığa çıkmasının olanaksız olduğunu göstermektedir. Diğer bir deyimle, magnitüd için bir üst sınır vardır. Deprem magnitüdlerinin üst sınırı, o bölgede beklenebilecek en büyük deprem magnitüdü m_1 ile belirlenecektir. Magnitüd için bir alt ve üst sınırın olduğu varsayılırsa, Denklem 6 kullanılarak magnitüd için aşağıda verilen birikimli dağılım işlevi, $F_M(m)$, elde edilir:

$$F_M(m) = \Pr(M < m | m_0 \leq m \leq m_1) = k \left(1 - e^{-\beta(m - m_0)} \right) \quad (7)$$

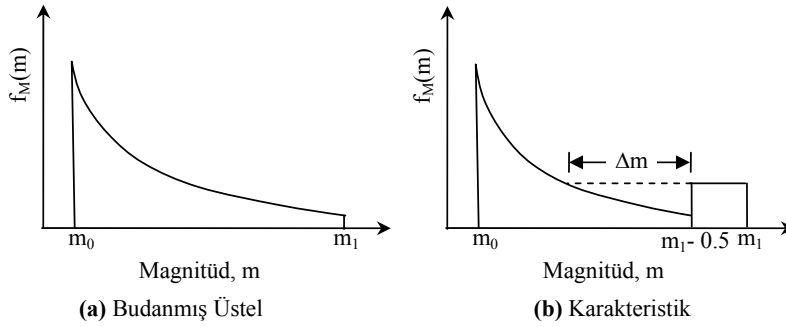
Burada,

$$k = \left[1 - e^{-\beta(m_1 - m_0)} \right]^{-1} \quad (8)$$

olup, birikimli dağılım işlevinin $m = m_1$ değerinde 1.0 olmasını sağlayan standartlaştırma katsayısıdır. Birikimli dağılım işlevinin magnitüde göre türevinin alınması ile aşağıda verilen olasılık yoğunluk işlevi, $f_M(m)$, bulunur:

$$\begin{aligned} f_M(m) &= k\beta e^{-\beta(m - m_0)} & m_0 \leq m \leq m_1 \\ &= 0 & \text{diğer yerlerde} \end{aligned} \quad (9)$$

Bu şekilde elde edilen budanmış üstel olasılık yoğunluk işlevi, Şekil 1(a)'da gösterilmiştir.

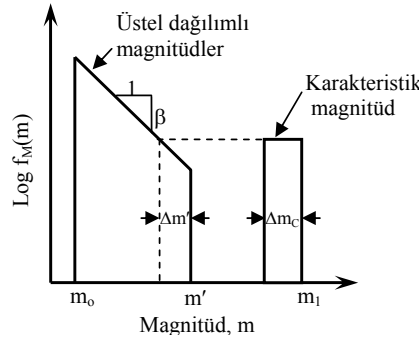


Şekil 1. Budanmış üstel ve karakteristik deprem modelleri için magnitüd olasılık yoğunluk işlevleri

Gözlem verilerine dayanarak üstel dağılımın parametrelerinin tahmini için değişik istatistiksel yöntemler kullanılabilir. Doğrusal regresyon ve en büyük olabilirlik istatistiksel tahmin yöntemleri en fazla tercih edilenlerdir. Standart en küçük kareler yöntemi, gözlenen ve tahmin edilen değerler arasındaki farkların karelerinin toplamalarının en küçüklenmesine dayanmaktadır. En küçük kareler regresyon yönteminin değişik uygulamaları mümkündür. Örneğin: frekans ya da birikimli frekans verilerinin kullanılması, regresyonun her bir magnitüd düzeyindeki frekanslara o magnitüd düzeyindeki gözlem sayısına göre verilen ağırlıklara göre yapılması (McGuire, 2004). En büyük olabilirlik yöntemi ise üstel dağılımın parametrelerini gözlenen magnitüd verilerinin olabilirliğini en büyükleyecek şekilde tahmin etmektedir. Değiştirilmiş en büyük olabilirlik yöntemi ise parametrelerin sağlam (robust) bir biçimde tahminine olanak sağlamaktadır (Yücemen ve Akkaya, 2012).

3.2 Karakteristik Deprem Modeli

Geçmiş deprem verilerine göre hesaplanan deprem tekerrür tahminleri ile sismolojik ve jeolojik incelemelere göre yapılanlar arasındaki çelişkiler, araştırmacıları bu çelişkileri giderecek yeni tekerrür modellerinin geliştirilmesine teşvik etmiştir. Bunlar arasında Schwartz ve Coppersmith (1984) tarafından önerilen karakteristik deprem modeli en fazla kabul gören olmuştur. Schwartz ve Coppersmith (1984), üstel dağılım modelinin büyük alanlardaki magnitüd dağılımını yeterli bir biçimde tanımladığını, ama fay segmentlerinde oluşan büyük magnitüdü depremlerin oluş sıklığını eksik tahmin ettiğini belirterek, karakteristik deprem modelini önermişlerdir. Youngs ve Coppersmith (1985), karakteristik deprem modeli için geçerli olacak bir olasılık yoğunluk işlevini çıkartmışlardır. Bu modelde deprem magnitüdüleri m' değerine kadar üstel dağılımlı olarak alınmışlardır. Magnitüdü m' den büyük depremler karakteristik deprem olarak tanımlanmışlar ve bunların $m_1 - \Delta m_C$ ve m_1 arasında bir biçimli dağılım gösterdikleri varsayılmıştır (Şekil 2).



Şekil 2. Schwartz ve Coppersmith (1984) tarafından önerilen karakteristik deprem modeli

Bu modelin kullanımı için Youngs ve Coppersmith (1985) bazı basitleştirici varsayımlar yapmışlardır. Δm_C , 0.5 ve $m' = m_1 - \Delta m_C$ olarak alınmıştır. Karakteristik depremin frekansının da üstel dağılımın $(m' - 1.0)$ değerindeki frekansa eşit olduğu varsayılmıştır. Bu varsayımların uygulanması ve olasılık yoğunluk işlevinin altındaki toplam alanın 1 olmasının sağlanması için gerekli işlemin yapılması ile, karakteristik deprem modeli için aşağıda verilen olasılık yoğunluk işlevi elde edilmiştir:

$$f_M(m) = \begin{cases} k\beta e^{-\beta(m-m_0)} & m_0 \leq m \leq m_1 - 0.5 \\ k\beta e^{-\beta(m_1 - \frac{3}{2} - m_0)} & m_1 - 0.5 \leq m \leq m_1 \end{cases} \quad (10)$$

burada, k olasılık yoğunluk işlevinin altındaki toplam alanın 1 olmasını sağlayan katsayıdır ve şu şekilde ifade edilmiştir:

$$k = \left[1 - e^{-\beta(m_1 - 0.5 - m_0)} + \beta e^{-\beta(m_1 - \frac{3}{2} - m_0)} 0.5 \right]^{-1} \quad (11)$$

Bu varsayımlara göre ortaya çıkan olasılık yoğunluk işlevinin biçimi, Şekil 1(b)'de gösterilmiştir. Bu şekilde, $\Delta m = 1$ dir.

Sismik tehlike analizinde en önemli parametrelerden biri de sismik kaynaklar için belirlenecek olan deprem magnitüdü üst sınırıdır. Bu değerin tahmini için değişik yöntemler vardır. Gözlenmiş en büyük deprem magnitüdünün bir miktar artırılarak kullanılması bu yöntemlerden bir tanesidir. Diğer yöntemler kırılma boyu-magnitüd ve atım-magnitüd korelasyonlarına dayanmaktadır (Wells ve Coppersmith, 1994). Mevcut tüm yerel veriler ve belirtilen bu yöntemler kullanılarak m_1 için en iyi tahmin yapılmalıdır.

4. DEPREMLERİN ZAMAN İÇİNDE OLUŞUM MODELLERİ

Depremlerin zaman içinde gösterdikleri rassal dağılımın modellenmesi için değişik stokastik modeller geliştirilmiştir. Olasılıksal sismik tehlike analizi çalışmalarının çoğunluğunda depremlerin zaman içindeki oluşumları homojen Poisson süreci ile modellenmektedir. Poisson modelinde deprem olaylarının birbirlerinden bağımsız oldukları varsayılmaktadır. Bir sismik kaynak içerisinde ve belirli bir t zaman aralığında en az bir deprem olma olasılığı şöyledir:

$$\Pr(N \geq 1) = 1 - e^{-v \cdot t} \quad (12)$$

burada, v , incelenen bölgede, birim zaman süresinde (genellikle bir yıl) meydana gelen magnitudü m_0 'a eşit veya m_0 'dan büyük depremlerin ortalama sayısı olup, $1/v$ ise yıl cinsinden ortalama tekerrür süresine eşittir. Poisson modelinde, v zaman içinde değişmeyen sabit bir değere eşittir.

Depremlerin zamana olan bağımlılığını modellemek üzere yinelenme modellerini kullanmak mümkündür. Yinelenme sürecine dayanan modellerde, depremlerin oluşumu, aynı dağılıma sahip tekerrür süreleri olan bir olaylar dizisi şeklinde alınmaktadır. Diğer bir deyimle gelecekte olacak depremin beklenen zamanı sadece bir önceki depremin olduğu tarihe bağımlıdır. Fay segmentinde meydana gelen bir deprem sonraki deprem için yinelenme sürecini başlangıç durumuna getirterek tekrar başlatmaktadır. Zamana bağımlılık, tekerrür süreleri için varsayılan dağılımlara bağlı olan tehlike oranı yolu ile modellenmektedir. Tehlike oranı aşağıda denklemleri verilen tehlike fonksiyonuna bağlıdır:

$$h(t) = \frac{f_T(t)}{1 - F_T(t)} \quad (13)$$

burada, $f_T(t)$ ve $F_T(t)$, sırası ile olaylar arası zamanın olasılık yoğunluk ve birikimli dağılım işlevleridir. Örneğin, Poisson süreci, tehlike oranının sabit ve tekerrür sürelerinin üstel dağılıma sahip olduğu bir yinelenme sürecidir.

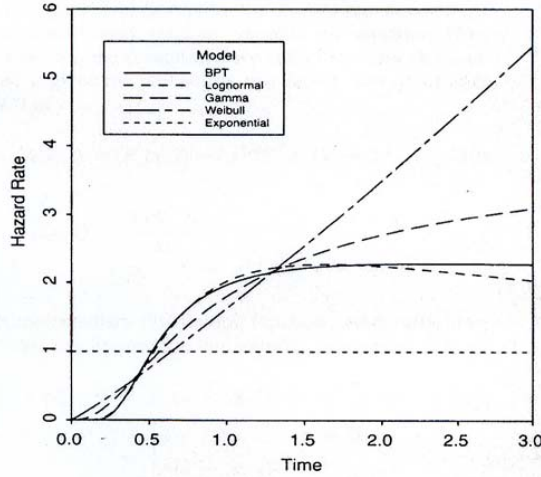
Esteva (1970) tekerrür süreleri için gamma dağılımını önermiştir. Weibull dağılımı, elastik geri tepme kuramı (Reid, 1910) ile uyumlu olarak, en son deprem olayından sonra geçen süre ile artan bir tehlike oranına sahip olması nedeni ile tekerrür zamanları için sıkça kullanılmıştır (örneğin, Kameda ve Ozaki, 1979; Hagiwara, 1974; Brillinger, 1982). Yakın bir zaman önce Brownian Aşma Zamanı (Brownian Passage Time) modeli, karakteristik depremlerin tekerrür sürelerinin olasılık dağılımı için önerilmiştir (Matthews, v.d., 2002). Brownian Aşma Zamanı (BAZ) modeli, San Francisco için 2002 yılında yapılan sismik tehlike analizinde kullanılmıştır. BAZ modeli için geçerli olan olasılık yoğunluk işlevi şöyledir (Matthews, v.d. 2002):

$$f_T(t) = \left(\frac{\mu}{2\pi\alpha^2 t^3} \right)^{1/2} e^{-\left(\frac{(t-\mu)^2}{2\alpha^2 \mu t} \right)} \quad (14)$$

burada, μ ortalama tekerrür süresi ve α aperiodyklik parametresi olup aynı zamanda standart sapmanın ortalama değere oranı olan değişkenlik katsayısına eşittir. Yinelenme modelinde, tekerrür süresi için kullanılan değişik olasılık dağılımları için geçerli olan tehlike oranı işlevlerinin değişimi Şekil 3'de gösterilmiştir. Bu şekilde de görüleceği üzere lognormal ve BAZ modelleri birbirlerine çok yakın değerler vermekte olup, depremlerin zamana bağımlı oluşum süreçlerini en iyi şekilde tasvir etmektedirler.

Wu, v.d. (1995), karma (hybrid) yinelenme modelini geliştirmişlerdir. Bu modele göre, büyük magnitudlü karakteristik depremlerin zamana olan bağımlılığı yinelenme süreci ile modellenmiş ve daha küçük depremler için kabul edilen üstel dağılım ile birleştirilmiştir. Karma yinelenme modeli, büyük magnitudlü karakteristik depremlerin periyodik olarak meydana geldiği varsayımını içerebilmek için hafızasız Poisson

modelini değiştirmektedir. Bu değişiklik karakteristik depremler için kabul edilen tek adımlik hafızadır. Diğer bir deyimle, bu modelde büyük magnitüdlü karakteristik depremlerin oluşumunun bir önceki karakteristik depremden sonra geçen zamana bağımlı olduğu varsayılmaktadır. Buna ek olarak daha önceki bölümde de açıklandığı gibi, büyük magnitüdlü karakteristik depremlerin olasılık dağılımı için üstel dağılım yerine, karakteristik depremin yer aldığı dar aralıkta yoğunlaşmış bir biçimli (uniform) bir olasılık dağılımı varsayılmaktadır (bakınız Şekil 1 (b) ve 2).



Şekil 3. Tekerrür süresi için kullanılan değişik olasılık dağılımları için geçerli olan tehlike oranı işlevleri. Üstel dağılım haricindeki tüm dağılımlar için ortalama değer 1, standart sapma ise 0.5' dir. (Matthews, v.d. 2002)

5. YER HAREKETİ TAHMİN DENKLEMİ

Sismik tehlikenin tahmini için depremin etkinliğini yansıtan bir zemin hareketi parametresi seçilmelidir. Bundan sonra da o parametrenin azalımına ilişkin bir model geliştirmelidir. Zemin hareketi tahmin ya da azalım modelleri zemin hareketi parametrelerinin özelliklerinin odak noktasından ya da sismik kaynağın seçilen bir noktasından uzaklaştıkça nasıl değişeceğini gösteren ve çoğunlukla gözlemsel yollarla elde edilen denklemlerdir. Bu denklemler genellikle m magnitüdündeki bir depremin, r uzaklığındaki inşaat sahasında yaratacağı en büyük zemin hareketi parametresinin değerini veren bir fonksiyon şeklindedir. Uzaklık olarak, merkez-üstü, odak ya da sismik kaynak üzerindeki bir noktadan ölçülen mesafeler alınmaktadır. Ayrıca inşaat sahasının zemin özelliklerini yansıtan bir parametre de bu ilişkilerde yer alabilmektedir. Bazı azalım ilişkileri fayın türünü de göz önünde tutmaktadır. Bu azalım ilişkilerinin genel yapısı aşağıda verilen denklem ile tanımlanabilir (Araya ve Der Kiureghian, 1988):

$$Y = N_y f(M, R, SP_i) \quad (15)$$

Burada, Y =tahmin edilecek olan kuvvetli yer hareketi parametresi (bağımlı değişken); N_y =azalım ilişkisindeki (ortalama tahmin eğrisi) belirsizlik (saçılım) için rassal düzeltme katsayısı; R =depremden inşaat sahasına olan “tanımlanmış” uzaklık ölçüsü; M =deprem büyüklüğünü gösteren herhangi bir ölçekteki magnitüd değeri; SP_i = deprem kaynağı, dalga yayılma hattı, yerel zemin koşulları ile ilgili parametreler.

Zemin hareketi tahmin ilişkisi genellikle en küçük kareler yönteminin gözlemsel verilere uygulanması ile elde edilen bir eğri şeklindedir. Bu eğrinin etrafındaki saçılımdan doğan belirsizliğin analize yansıtılması, bundan sonraki bölümde anlatılacağı üzere rassal düzeltme katsayısı N_y ile sağlanmaktadır.

6. BELİRSİZLİKLERİN ANALİZİ

Olasılıksal sismik tehlike analizinde temelde birbirinden farklı iki tür belirsizlik vardır. Şu an kullanılan terminolojiye göre bunlar rassal (aleatory) ve bilgiye dayalı (epistemic) belirsizliklerdir. Rassallıktan kaynaklanan belirsizlikler sismik tehlikeyi etkileyen fiziksel olayların doğasında mevcut olan rassallık ve değişkenlikten kaynaklanmaktadır ve daha fazla veri ve bilgi elde edilerek azaltılmaları mümkün değildir. Gelecekte olacak bir depremin yeri, büyüklüğü, fay kırılmasının boyutları ve yönü bu tür belirsizliğin örnekleridir. Bilgi/veri eksikliğinden kaynaklanan belirsizliği ise elde edilecek yeni bilgiler ve veriler ile azaltmak mümkündür. Sismotektonik bölgelerin konumu, sismisite parametrelerinin dağılımlarındaki istatistiksel parametrelerdeki belirsizlikler bu tür belirsizliğe örneklerdir.

6.1 Zemin Hareketi Tahmin Modelindeki Belirsizlik

Sismik tehlike analizinde en önemli belirsizlik kaynaklarından biri kullanılan azalım ilişkisidir. Bu “aleatorik” belirsizliğin modele doğrudan yansıtılması Denklem 15’de gösterildiği gibi azalım ilişkisinin N_y ile simgelenen rassal düzeltme katsayısı ile çarpılması ve birinci mertebe belirsizlik analizi yönteminin kullanılması ile sağlanmıştır. Genellikle, N_y , beklenen değeri (ortalaması) 1 olan (yani ortalama tahmin eğrisinde bir yanlışlığın olmadığı) lognormal dağılımlı bir değişken olarak alınmaktadır. Azalım ilişkileri çoğunlukla $\ln Y$ cinsinden ifade edildiğinden, düzeltme katsayısı da toplam ve $\ln N_y$ şeklinde olacaktır. Bu durumda $\ln N_y$ ortalaması, $\mu=0$ ve standart sapması, σ_a , olan normal dağılımlı bir rassal değişkene dönüşür. Azalım ilişkisindeki belirsizliğin (saçılımın) ölçüsü olan standart sapma, σ_a , vasıtası ile azalım ilişkisindeki belirsizlik sismik tehlike sonuçlarına yansıtılmaktadır. Genellikle literatürde yer alan azalım ilişkileri, σ_a değeri ile birlikte verilmektedir. Yerel verilere dayanmayan “ithal” azalım ilişkilerinin kullanıldığı durumda σ_a değerini arttırmak gerekebilir.

6.2 Sismik Kaynakların Konumundaki Belirsizlik

Diğer bir belirsizlik sebebi de sismik kaynak bölgelerinin belirlenmesi ile ilgilidir. Bu belirsizliğin sonuçlara yansıtılması için klasik sismik tehlike modellerindeki deterministik bölge sınırları yerine, bu sınırların konumunun rassal olduğu varsayılmıştır. Bender (1986) tarafından önerilen modelde depremlerin beklenen konumu birbiçimli dağılım olarak alınmış, ancak ortalama konum etrafındaki belirsizlik iki değişkenli Gauss (normal) dağılımı ile modellenmiştir. Ortalama vektör, $(\bar{x}, \bar{y})$ en olası konumu, standart sapma σ ise konumdaki belirsizliği göstermektedir. En olası konumdan, δ_x ve δ_y miktarlarında bir sapma olasılığı aşağıdaki yoğunluk işlevi ile orantılı olacaktır:

$$f(\delta_x, \delta_y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\delta_x^2 + \delta_y^2}{2\sigma^2}\right) \quad (16)$$

Burada, x ve y birbirine dik iki eksen simgelemektedir ve her iki eksen boyunca konumdaki belirsizliğin eşit olduğu varsayılmıştır.

Yücemen ve Gülkan (1994), Bender (1986) tarafından önerilen modeli, konumdaki yöne bağlı belirsizliği içerecek şekilde geliştirmişlerdir. Eğer konumdaki belirsizlik birbirine dik iki eksen boyunca aynı değilse, o zaman Denklem 16 şu şekilde yazılacaktır (Yücemen ve Gülkan, 1994):

$$f(\delta_x, \delta_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\left(\frac{\delta_x^2}{2\sigma_x^2} + \frac{\delta_y^2}{2\sigma_y^2}\right)\right] \quad (17)$$

Denklem 16 ve 17 ile bir sismik kaynağın sınırlarındaki belirsizlik inşaat sahasında oluşacak zemin hareketi parametresine yansıtılmış olacaktır; zira bir sismik kaynaktaki oluşacak depremlerin beklenen konumları (δ_x, δ_y) miktarında kaydırılırsa, bu o sismik kaynağın sınırlarının da aynı miktarda yer değiştireceği anlamına gelir. Burada dikkat çekilmesi gereken husus, bir kaynak bölgesinde meydana gelen depremin koordinatlarındaki σ_x ve σ_y ile gösterilen belirsizliğin, o depremi içine alan bölgenin sınırlarında aynı değerlere sahip bir belirsizliğe eşdeğer olduğu gerçeğidir. Sismik kaynakların konumundaki belirsizliği yukarıda anlatılan model ile sonuçlara doğrudan yansıtmak SEISRISK-III programı (Bender ve Perkins, 1987) ile mümkün olmaktadır.

6.3 Sismik Kaynak Parametrelerinin Tahmin Edilen Değerlerindeki Belirsizlikler ve Mantık Ağacı Yöntemi

Olasılıksal sismik tehlike analizi modelinde bulunan yıllık sismik tehlike değeri, $P(Y > y)$, sismik parametreler v , β ve m_1 'in verilen değerlerine, sismik kaynak modellemesine ve diğer varsayımlara bağlıdır. Dolayısı ile hesaplanan sismik tehlike değerleri koşullu olasılıklardır, ve $P(Y > y / v, \beta, m_1, A_j)$ şeklinde yazılmalıdır. Burada A_j değişik varsayımları simgelemektedir. Sismik parametrelerin tahminindeki ve değişik varsayımlardaki epistemik belirsizliklerin sonuçlara yansıtılması, aşağıda özetlenen ve genellikle mantık ağacı (logic-tree) olarak adlandırılan şu basit yöntemle, dolaylı olarak yapılabilir (Yücemen, 1982):

- (i) Sismik parametrelerin değerlerine, sismik kaynakların konumuna, azalım ilişkisine, sismik veri tabanına ve diğer hususlara ilişkin her varsayıma, o varsayımın diğerlerine göre doğru olma olasılığını yansıtan öznel olasılık değerleri verilir.
- (ii) Her bir varsayım grubu için (mesela, bir v değeri, bir β değeri, bir m_1 değeri, bir azalım ilişkisi ve kaynakların konumuna ilişkin bir varsayım), o grubu oluşturan varsayımların öznel olasılıklarının çarpımına eşit olan birleşik olasılık değeri bulunur. Bu şekilde hesaplanan birleşik olasılıkların toplamının bire eşit olması gereklidir. Ayrıca en iyi tahminlerden oluşan grup için bulunacak birleşik olasılık değerinin de en büyük olması beklenir.
- (iii) Her bir varsayım için sismik tehlike hesabı yapılır. Bulunan aşılma olasılığı o varsayım grubu için belirlenen birleşik olasılık değeri ile çarpılır. Birleşik olasılıklarla çarpılmış aşılma olasılıklarının toplamı aranılan sismik tehlike değerini verecektir.

Toplam olasılık kuramına göre hesaplanan bu ağırlıklı ortalama sismik tehlikeye “Bayes” tahmini denilecektir. Matematiksel olarak ifade edilirse:

$$P(Y > y) = \sum_{j=1}^n P(Y > y/G_j) w_j \quad (18)$$

burada, G_j =j sayılı varsayım grubu; $w_j=P(G_j)$, j sayılı varsayım grubunun diğerlerine göre doğru olma olasılığını yansıtan birleşik öznel olasılık; n =göz önünde tutulan varsayım takımlarının toplam sayısıdır.

7. SONUÇLAR

Bir “deprem laboratuvarı” olarak nitelendirebileceğimiz ülkemizde geçmiş deprem verilerinden en iyi bir şekilde yararlanmak için, istatistik, sismoloji, jeoloji ve deprem mühendisliği dallarında yetişmiş uzmanlardan oluşan grup çalışmalarına gereksinim vardır. Deprem, istatistik uzmanlarının katılımının kaçınılmaz olduğu ve disiplinlerarası bir ortamda araştırılması gereken bir konudur. Burada elde edilen başlıca sonuçlar aşağıda özetlenmiştir.

(i) Yeni nesil sismik tehlike haritalarının hazırlanmasında aktif faylardan kaynaklanan sismik tehlikenin özellikle dikkate alınması ve faylarla ilgili mevcut bilgilerle uyumlu stokastik modellerin kullanılması gereklidir. Aktif faylardan kaynaklanan tehlikenin modellenmesine uygun olan yinelenme rassal süreci bu tür çalışmalarda göz önünde tutulmalıdır.

(ii) Rassal (aleatory) ya da bilgilerin yetersizliğinden (epistemic) kaynaklanan tüm belirsizliklerin etkilerinin sismik tehlike analizine yansıtılması çok önemlidir. Bunun gerçekleştirilmesi ise olasılık ve istatistik kuram ve yöntemlerinin kullanılması ile mümkündür. Ancak bu kuram ve yöntemlerin uygulamasında öngörülen koşulların sağlanıp sağlanmadığının gerektiği gibi kontrol edilmesi şarttır.

(iii) Kataloglarda yer alan değişik magnitüd ölçeklerinin tek bir magnitüd ölçeğine, örneğin moment magnitüdüne (M_w), dönüştürülmesi gereklidir. Dönüşüm ilişkilerinin elde edilmesinde EKK regresyonu yerine ortogonal regresyon yönteminin kullanılması daha uygun olacaktır. Ortogonal regresyon, çevirim ilişkilerinin eğimlerini standart en küçük kareler yöntemine göre her zaman daha büyük tahmin etmekte ve bu nedenle de büyük depremlerin magnitüdlerini geleneksel yöntemle göre daha büyük vermektedir. Dolayısıyla deprem tehlikesinin tahmininde ortogonal regresyonun kullanılması durumunda geleneksel yöntemin sonuçlarına göre daha emniyetli tarafta değerler elde edilecektir.

(iv) Depremlerin zaman uzayında oluşumları bir Poisson süreci olarak alınabilir. Poisson dağılımı, rassal olarak oluşan büyük magnitüdü ana şokların oluşumu için geçerli bir model olup, bu basit modelden elde edilen sonuçlar genellikle daha karmaşık bazı diğer stokastik modellerden (örneğin Markov) bulunan sonuçlarla uyum içindedir. Ancak Kuzey Anadolu ve San Andreas gibi fay zonlarında yapılan paleo-sismik çalışmalar bu faylarda belirli büyüklükteki depremlerin periyodik olarak ortaya çıktığını göstermiştir. Diğer bir deyimle, bu gibi fayların yaratabileceği ve karakteristik deprem olarak adlandırılan büyük depremlerin tekerrür süreleri daha önce meydana gelmiş

büyük magnitüdlü sismik etkinlikle bağımlıdır. Bu durumda yinelenme modeli, Poisson modeline kıyasla karakteristik depremlerin oluşumu için daha uygun bir stokastik model olmaktadır.

8. KAYNAKLAR

Araya, R., Der Kiureghian, A., 1988. Seismic Hazard Analysis: Improved Models, Uncertainties and Sensitivities, EERC Report No. UCB/EERC-90/11, College of Engineering, University of California, Berkeley.

Bender, B., 1986. Modeling Source Zone Boundary in Seismic Hazard Analysis, Bull. Seism. Soc. Am., 76(2): 329-341.

Bender, B., Perkins, D., 1987. SEISRISK III: A Computer Program for Seismic Hazard Estimation, U.S.G.S. Bulletin 1772.

Boore, D. M., Joyner, W. B., 1982. The Empirical Prediction of Ground Motion, Bull. Seism. Soc. Am., 72(6): 43-60.

Brillenger, D. R., 1982. Some Bounds for Seismic Risk, Bull. Seism. Soc. Am., 72(4): 1403-1410.

Castellaro, S., Mulargia, F., Kagan, Y. Y., 2006. Regression Problems for Magnitudes, Geophys J Int, 165: 913-930.

Deniz, A. (2006) Estimation of Earthquake Insurance Premium Rates for Turkey, M.Sc. Thesis, Dept. of Civil Engineering, METU.

Deniz, A., Yüçemen, M. S., 2010. Magnitude Conversion Problem for the Turkish Earthquake Data, Natural Hazards, 55(2): 333-352.

Esteva, L., 1970 Seismic Risk and Seismic Design Decisions, in Seismic Design for Nuclear Power Plants, Ed. Hansen, R. J., MIT Press, Cambridge, Mass.

Gardner, J. K., Knopoff, L., 1974. Is the Sequence of Earthquakes in Southern California, with Aftershocks Removed, Poissonian?, Bull. Seism. Soc. Am., 64: 1363-1367.

Hagiwara, Y., 1974. Probability of Earthquake Occurrence as Obtained from a Weibull Distribution Analysis of Crustal Strain, Tectonophysics, 23(3): 313-318.

Kagan, Y. Y., 2002. Aftershock Zone Scaling, Bull. Seism. Soc. Am., 92(2): 641-655.

Kameda, H., Ozaki, Y., 1979. A Renewal Process Model for Use In Seismic Risk Analysis, Memoirs of the Faculty of Engineering, Kyoto University, 41: 11-35.

Matthews, M. V., Ellsworth, W. L., Reasenberg, P. A., 2002. A Brownian Model for Recurrent Earthquakes, Bull. Seism. Soc. Am., 92(6): 2233-2250.

- McGuire, R. K., 2004. Seismic Hazard and Risk Analysis, EERI, MNO-10, Oakland, CA.
- Omori, F., 1894. On the Aftershocks of Earthquakes, Journal of College of Science, Imperial University, Tokyo, 7: 111-200.
- Prozorov, A. G., Dziewonski, A. M., 1982. A Method of Studying Variations in the Clustering Property of Earthquakes: Application to the Analysis of Global Seismicity, Journal of Geophysical Research, 87(B4): 2829-2839.
- Reid, H. F., 1910. The Mechanics of the Earthquake, The California Earthquake of April 18, 1906, Report of the State Investigation Commission, Carnegie Institution of Washington, Washington, D.C., 2: 16-28.
- Richter, C. F., 1958. Elementary Seismology, W.H. Freeman and Company, San Francisco.
- Savage, M. K., Rupp, S. H., 2000. Foreshock Probabilities in New Zealand, New Zealand Journal of Geology and Geophysics, 43: 461-469.
- Schwartz, D. P., Coppersmith K. J., 1984. Fault Behavior and Characteristic Earthquakes: Examples from the Wasatch and San Andreas Fault Zones, J. Geophys. Res. 89: 5681-5698.
- Stepp, J.C., 1973. Analysis of the Completeness of the Earthquake Sample in the Puget Sound Area, in Contributions to Seismic Zoning, S. T. Harding (Ed.), National Oceanic and Atmospheric Technical Report EERL 267-ESL 30.
- Ulusay, R., Tuncay, E., Sönmez, H., Gökçeoğlu, C., 2004. An Attenuation Relationship Based on Turkish Strong Motion Data and Iso-Acceleration Map of Turkey, Engineering Geology, 74: 265-291.
- Utsu, T., Ogata, Y., Matsu'ura, R. S., 1995. The Centenary of the Omori Formula for a Decay Law of Aftershock Activity, Journal of Physics of the Earth, 43: 1-33.
- Van Dyck, J. F. M., 1985. Statistical Analysis of Earthquake Catalogs, PhD. Thesis, Civil Engineering Department, Massachusetts Institute of Technology, Cambridge.
- Wells, D. L., Coppersmith, K. J., 1994. New Empirical Relationships Among Magnitude, Rupture Length, Rupture Width, Rupture Area and Surface Displacement, Bull. Seism. Soc. Am., 84(4): 974-1002.
- Wu, S. C., Cornell, C. A., Winterstein, S. R., 1995. A Hybrid Model and its Implication on Seismic Hazard Results, Bull. Seism. Soc. Am., 85: 1-16.
- Youngs, R. R., Coppersmith, K. J., 1985. Implications of Fault Slip Rates and Earthquake Recurrence Models to Probabilistic Seismic Hazard Estimates, Bull. Seism. Soc. Am., 75: 939-964.

Yüçemen, M. S., 1982. Sismik Risk Analizi, Orta Doğu Teknik Üniversitesi, Ankara, 160 s.

Yüçemen, M. S., Gülkan, P., 1994. Seismic Hazard Analysis with Randomly Located Sources, Natural Hazards, Kluwer Academic Publishers, 9: 215-233.

Yüçemen, M. S., Akkaya, A. D., 2012. Robust Estimation of Magnitude-Frequency Relationship Parameters, Structural Safety, 38: 32-39.

ESTIMATION OF SEISMIC HAZARD: PROBABILISTIC AND STATISTICAL METHODS

ABSTRACT

Considering the aleatory uncertainties related to earthquake occurrences with respect to time, space, magnitude and the additional epistemic uncertainties, probabilistic methods appear to be more appropriate. However, in implementing the probabilistic and statistical methods, engineers very seldom check the validity of the underlying assumptions with respect to the available data. This may lead to serious errors and most often those who apply these methods are unaware of the resulting errors. In this paper, the basic steps for the development of seismic hazard maps are stated together with the necessary background and supporting information for the implementation of these steps. Also attention is drawn to the possible errors committed in utilizing the statistical methods for the assessment of seismic hazard.

Keywords: Probabilistic seismic hazard analysis, Orthogonal regression, Renewal model, Magnitude-recurrence relationship.

PROF. DR. YALÇIN TUNCER'İN RESMİ İSTATİSTİKLERE KATKILARI

Orhan GÜVENEN*

ÖZET

Türkiye Cumhuriyeti İstatistik Bilimi tarihinde, çok değerli Prof. Dr. Yalçın Tuncer Hocamızın (1936-2011) en onurlu en üst düzeyde katkıları vardır. Bu çalışmada Prof. Dr. Yalçın Tuncer'in, Türkiye Cumhuriyeti Başbakanlık Devlet İstatistik Enstitüsü'nde yapmış olduğu çok önemli katkıların tümünü kapsamayan örnekler sunulmuştur. Enstitü Başkan Yardımcısı olarak, Araştırma (ATİT) Dairesi çalışmaları, sosyal istatistikler, 1990 Genel Nüfus Sayımı Komitesi Başkanı olarak yapmış olduğu önemli çalışmalar belirtilmiştir. Ayrıca, 1991 Genel Tarım Sayımı, 1992 Sanayi ve İş Yerleri Sayımı, 1994 Hanehalkı Gelir ve Tüketim Harcamaları Araştırması (Anketi) Komite Başkanı, DİE Araştırma Sempozyumu, İstatistik Eğitim Merkezi, Tarih İstatistikleri, Dünya ve Türkler'de Matematik ve İstatistik Çalıştayları ve bu çalışmada tümü belirtilmeyen kapsamda, büyük katkılarını en derin minnet ve saygıyla anıyoruz.

Anahtar Kelimeler: İstatistik metodolojisi, Sosyal istatistikler, Genel nüfus sayımı, İstatistik eğitim merkezi, Dünyada ve Türkler'de matematik ve istatistik.

1. GİRİŞ

Türkiye Cumhuriyeti İstatistik Bilimi tarihinde, çok değerli Prof. Dr. Yalçın Tuncer Hocamızın, en onurlu en üst düzeyde katkıları vardır. Türkiye Cumhuriyeti Başbakanlık Devlet İstatistik Enstitüsü (DİE)'nde, Hocamızla çalışmış olmak benim için bir onurdur. Ben, Hocamızın maiyetinde çalıştım. 1988-1994 yıllarında görevde olduğum dönemde, şu anda aramızda olan ve olmayan çok değerli hocalarımızla, Türkiye resmi istatistiklerine "Türkiye Optimali" ve uluslararası düzeyde, etkin olma hasretimizin hedefinde, çalışmalarımızı yönlendirdik. Bu hasretle, Prof. Dr. Yalçın Tuncer Hocamızın, o kendine özgü, soylu bilim insanı, bilge kişiliğiyle müstesna katkıları olmuştur. Kendilerini en derin saygıyla, rahmetle anıyorum. Bu mütevazı çalışmada, çok değerli katkılarını esirgemeyen Prof. Dr. Öztaş Ayhan Hocamıza ve TÜİK Daire Başkanı Sayın Şennur Onur'a minnet duygularımı iletmek isterim. Bu sempozyumu düzenleyen ODTÜ İstatistik Bölüm Başkanı Prof. Dr. Öztaş Ayhan ve tüm emeği geçenlere sade vatandaş kimliğimle en derin minnet duygularımı sunarım. Bu çalışma, Prof. Dr. Yalçın Tuncer'in Türkiye Cumhuriyeti Başbakanlık Devlet İstatistik Enstitüsü'nde yapmış olduğu çok önemli katkıların tümünü kapsamayan örnekler sunmaktadır.

*Prof. Dr., Bilkent Üniversitesi, Stratejik Karar Sistemleri, Ekonometri ve Uluslararası Ekonomi Profesörü. Bilkent Üniversitesi Dünya Sistemleri, Ekonomileri ve Stratejik Araştırmalar Enstitüsü Direktörü ve Muhasebe Bilgi Sistemleri Bölümü Başkanı, e-posta: gorhan@bilkent.edu.tr

2. PROF. DR. YALÇIN TUNCER'İN YÖNETİM VE DANIŞMANLIK GÖREVLERİ

Prof. Dr. Yalçın Tuncer, 1990 yılı başlangıcında kısa bir dönem Enstitü Başkanlık Danışmanı görevini yapmış ve hemen ardından, Enstitü Başkan Yardımcılığı görevini büyük bir sorumluluk, etkenlik ve başarı ile sürdürmüştür. 1990 yılının ikinci yarısından itibaren Enstitü Başkan Yardımcısı olarak görev yapmıştır. İdari ve metodolojik görevlerine ek olarak belirli Komite Çalışmaları'na başkanlık etmiştir. İdari görevi kapsamında iki daire başkanlıkları kendisine bağlı olarak çalışmaktaydı. Prof. Tuncer'e idari olarak bağlı olan iki teknik daire bulunmaktaydı. Bunlar, *Araştırma Tetkik ve İstatistiksel Teknikler* ve *Sosyal İstatistikler* daire başkanlıklarıydı.

Araştırma Tetkik ve İstatistiksel Teknikler (ATİT) Dairesi Başkanlığı, Devlet İstatistik Enstitüsü'nün istatistik metodolojilerini belirleyen ve uygulayan çok önemli bir daire si olma niteliğindedir. Tüm teknik dairelerle işbirliği halinde çalışmaktaydı. Enstitüde yapılan ve örneklemeye dayanan tüm araştırmaların metodolojisini ve uygulama planlarını ATİT düzenlemekteydi. Ayrıca araştırma verilerinin analizinden sonra, araştırmaların örneklem ve örneklem dışı hataları yine ATİT tarafından yapılmaktaydı.

Sosyal İstatistikler Dairesi Başkanlığı, Enstitü'nün sosyal istatistiklerinden (Demografi, Genel Nüfus Sayımı, Doğum İstatistikleri, Ölüm İstatistikleri, Göç İstatistikleri, Evlenme İstatistikleri, Boşanma İstatistikleri, İntihar İstatistikleri) sorumlu olan bu dairenin çalışma kapsamı içerisinde öncelikle 1990 Genel Nüfus Sayımının yapılması gelmekteydi.

3. ÇALIŞMA KOMİTESİ BAŞKANLIKLARI

Prof. Dr. Yalçın Tuncer, 1990 Genel Nüfus Sayımı Komitesi Başkanı olarak sayım ile ilgili dairelerin yakın işbirliği ve bir küme kapsamında çalışması konusunda çok başarılı olmuştur. Sayımda uygulanan bilimsel metodolojiler ve "Nüfus Sayımının Cevap Güvenilirliği Araştırması"na öncülük etmiştir. Belirtilen cevap güvenilirliği araştırması DİE'de ilk olarak yapılmıştır. Araştırma sonuçları ise iki ayrı dilde DİE yayını bir kitap olarak kullanıcılara sunulmuştur.

Prof. Dr. Öztaş Ayhan, Prof. Dr. Haluk Kasnakoğlu ile birlikte Prof. Dr. Yalçın Tuncer ve DİE mensupları, çok değerli katkıları ile 1991 Genel Tarım Sayımını gerçekleştirmişlerdir. Bu tarım sayımında ilk kez, tüm köylerden tam sayım olarak, tarımın yapısı ile ilgili bilgi alınmıştır. Buna ilaveten temsili olan örneklemle seçilmiş köylerden, yine temsili olarak seçilen tarımsal işletme (tarımsal hanehalkı) düzeyinde ankete dayanan kapsamlı tarımsal işletme bilgileri elde edilmiştir.

1992 Genel Sanayi ve İş Yerleri Sayımı'da ilk kez tüm soru kağıtlarında sektörlerin yapılarını daha iyi yorumlamaya ve daha etkin analiz yapılmasına olanak sağlamak için, madde ayrıntısında, bilgi derlenmesinde öncü rol oynamış, bu bilgilerin girdi çıktı

(Input – Output) hesaplamalarında ve Milli Gelir Çalışmalarında kullanılabilecek düzeyde teminine büyük katkıları olmuştur.

Prof. Dr. Yalçın Tuncer, 1994 Hanehalkı Gelir ve Tüketim Harcamaları Araştırması Komitesi Başkanı niteliğiyle, bu araştırmanın planlama aşamasında metodolojiye önemli katkıda bulunmuştur. Bu katkılar TÜFE endeks yapısına önemli katma değer sağlamıştır. Tüketici Fiyatları Endeksi (TÜFE), hanehalklarının tüketimine yönelik mal ve hizmet fiyatlarının zaman içindeki değişimini ölçmektedir. 1994 temel yıllık TÜFE'nin temel amacı, piyasada tüketime konu olan mal ve hizmetlerin fiyatlarındaki değişimi ölçerek enflasyon oranını hesaplamaktır. Bu kavram, hane halklarının nihai parasal harcamalarını, kendi tüketimlerine yönelik tüketimleri ve hane halkları için geçerli izafi kiralari ve aynı gelirleri de kapsamaktadır. Ağırlıkların tespitinde ve endeks hesaplamasında amaca göre bireysel tüketim sınıflaması kullanılmış ve bu harcamalar 10 ana grup, 33 alt grup altında toplanmıştır. Endekste 410 madde kapsama alınmıştır. EUROSTAT içinde belirtilen ve ilgili uluslararası ve bölgesel standartlarla, Türkiye'nin metodolojisi arasında önemli bir fark yoktur. 1994 temel yıllık kentsel yerler TÜFE'de 410 madde, 35 yerleşim yerinden, 6390 iş yerinden endeks kapsamında takip edilmiştir. Endeksin coğrafi kapsamı, nüfusu 20.001 ve daha fazla olan kentsel yerleşim yerleri içindir. Endekste 35 yerleşim yerinden fiyat derlenmiştir. 1994 = 100 Temel Yıllı Kentsel Yerler Tüketici Fiyatları Endeksi Laspeyres formulu $L = w * (P_i / P_0)$ kullanılarak hesaplanmıştır.

4. ARAŞTIRMA SEMPOZYUMU'NUN BAŞLATILMASI

1990 yılında başlatılan (Ulusal İstatistik, Matematik, Ekonometri Sempozyumu) günümüzde halen Ulusal İstatistik Araştırma Sempozyumu olarak devam etmektedir. 1990'da başlatılan bu sempozyumlarda Prof. Dr. Yalçın Tuncer'in katkıları çok önemli olmuştur.

Sempozyumda sunulan bildirilerin önemli bir hakemlik sürecinde değerlendirilerek basılması ve istatistikçilerin kullanımına sunulması çalışmalarında, Hocamızın katkıları çok etkin olmuştur. Bu kapsamdaki yayınlar, sonraki yıllarda da Araştırma Sempozyumu bildiriler kitabı olarak bilim dünyasına katkılarını sürdürmüştür.

5. İSTATİSTİK EĞİTİM MERKEZİ'NİN GELİŞTİRİLMESİ

İstatistik Eğitim Merkezi'ni geliştirmesi ve hizmet içi en üst bilimsel düzeyde merkezin kurulması çalışmasına katkıda bulunmuşlardır. Bu kapsamda, çok nitelikli bir lisansüstü eğitim programının geliştirilmesini sağlamışlardır. Eğitim İngilizce yapılmaktaydı, bu lisansüstü eğitimin, Hocamız tarafından düzenlenen program kitapçığı Ek 1'de aynen sunulmuştur. Bu eğitim programlarına, Devlet İstatistik Enstitüsü ve DİE dışından adayların müracaat ederek sınava girme olanakları sağlanmış ve başarılı olanlar o dönemde istatistik ve resmi istatistikler kapsamında uluslararası düzeyde bir eğitim almışlardır.

6. YENİ ARAŞTIRMA VE UYGULAMA MERKEZLERİN KURULMASINA KATKILARI

T.C. Başbakanlık Devlet İstatistik Enstitüsü Başkanlığı bünyesinde 1989–1990 yıllarında kurulan, ülkemiz için çok önemli bu 3 araştırma ve uygulama merkezinin kuruluş ve gelişmesinde Prof. Dr. Yalçın Tuncer Hocamızın çok değerli katkıları olmuştur. Bu merkezler; *Harzemli Bilgisayar Merkezi*, *Uluğ Bey Bilgi Sistemleri Merkezi*, ve *Piri Reis Uzaktan Algulama ve Coğrafi Bilgi Sistemleri Merkezi*'dir.

7. TARİHİ İSTATİSTİKLER DİZİSİ'NE KATKILARI

DİE kapsamında “Tarihi İstatistikler” yayınlamak amacıyla olduğumuzu, çok değerli Hocamız Prof. Dr. Halil İnalcık’a arzettiğim ve bizleri yönlendirmesinin büyük şeref olacağını belirttiğimde, çok etkilendi, ve cevabı, “Dünya imparatorlukları içinde bu kapsamda en büyük yazılı kaynağa sahip 3 imparatorluktan biridir, Türk–Osmanlı İmparatorluğu”. Lütfettiler, çalışmalara başladık. Projenin uygulanmasını, Prof. Dr. Şevket Pamuk, değerli meslektaşlarla etkin şekilde başlattı. Bu kapsamda da Prof. Dr. Yalçın Tuncer Hocamız, kıymetli desteğini hiç eksik etmedi. Tarihi istatistikler projesi yayınları bu çalışmalarla başlatıldı ve sürdürüldü. “Tarihi İstatistikler” dizisinin devamı Türk İstatistik Bilimi açısından büyük önem taşımaktadır. Bu hususu, saygıyla, yetkililere arz ediyoruz.

Prof. Dr. Yalçın Tuncer hocamızın, Tarihi İstatistikler Dizisi’nin başlatılmasında ve geliştirilmesinde önemli katkıları olmuştur. Bu dizide yer alan kitaplar ülkemizde konusunda uzman olan bilim insanlarının katkılarıyla gerçekleştirilmiştir. Dizide yer alan kitapların hazırlanması belirli bir zaman süreci içinde gerçekleşmiş ve kitapların basımı, DİE matbaasında yapılmıştır. Ülkemizde resmi istatistiklerin gelişmesinin tarihsel sürecini kapsayan bu dizide yer alan kitaplar aşağıda liste olarak sunulmuştur:

19. Yüzyılda Osmanlı Dış Ticareti (Ottoman Foreign Trade in the 19th Century)

Osmanlı İmparatorluğu’nun ve Türkiye’nin Nüfusu 1500–1927 (The Population of Ottoman Empire and Turkey 1500 – 1927)

Osmanlı Dönemi Tarım İstatistikleri 1909, 1913 and 1914 (Agricultural Statistics of Turkey During the Ottoman Period 1909, 1913 and 1914)

Osmanlı Sanayii 1913, 1915 Yılları Sanayi İstatistiki (Ottoman Industry Industrial Census of 1913, 1915)

Osmanlı Devleti’nin İlk İstatistik Yıllığı 1897 (The First Statistical Yearbook of Ottoman Empire 1897)

Tanzimat’tan Cumhuriyet’e Modernleşme Sürecinde Eğitim İstatistikleri 1839–1924 (Education Statistics in Modernization From the Tanzimat to the Republic 1839–1924)

Osmanlı Mali İstatistikleri Bütçeler 1841–1918 (Ottoman Financial Statistics Budgets 1841–1918)

Osmanlı Devleti'nde Bilgi ve İstatistik (Data and Statistics in the Ottoman Empire)

İstanbul ve Diğer Kentlerde 500 Yıllık Fiyatlar ve Ücretler 1469 – 1998 (500 Years of Prices and Wages in İstanbul and Other Cities 1469–1998)

Osmanlı İmparatorluğu'nda İlk Nüfus Sayımı 1831 (First Population Census in the Ottoman Empire 1831)

8. DÜNYA'DA VE TÜRKLER'DE MATEMATİK VE İSTATİSTİK ÇALIŞTAYLARI

DİE kapsamında başlattığımız bu çalıştaylardan ülkemize, matematik ve istatistik'te değerli bilim insanları çok önemli çalışmalarda bulundular. Bir Türk deyişi, “Bazı insanların gölgesi yeter” der. Prof. Dr. Yalçın Tuncer Hocamızın, bu kapsamda da varlığı çok etkiliydi. Bu dönemde, 1958 yılı “Fields Medal” sahibi ünlü Fransız matematikçi Prof. Dr. René Thom (1923–2002) ve Prof. Dr. Edmond Malinvaud, DİE'de verdikleri seminerlerde olağanüstü kişilikleriyle, matematik ve istatistikte değerli ufuklar açtılar.

9. DİE UZMANLIK VE UZMAN YARDIMCILIĞI KOMİSYONU BAŞKANLIĞI

Devlet İstatistik Enstitüsü çalışanlarının teknik ve metodolojik altyapılarını güçlendirmek amacı ile belirli, çok iyi düzenlenmiş ve hakkaniyetli sınavlara gereksinim duyulmaktaydı. Uzmanlık sınavları yapan komitenin koordinasyon görevini değerli Hocamız gerçekleştirmişlerdir. Devlet İstatistik Enstitüsü uzman yardımcılarının belirlenmesi için yapılan sınavların koordinasyonunda, Prof. Dr. Yalçın Tuncer Hocamız başkanlığında yapılmıştır.

10. GENEL DEĞERLENDİRME

Prof. Dr. Yalçın Tuncer Hocamızın, müstesna katkıları, “Türkiye Optimali” ve uluslararası boyutta olmuştur. Minnet, şükran ve en derin saygı duygularımızla, büyük özlemle anıyoruz çok değerli Prof. Dr. Yalçın Tuncer Hocamızı. Eşi, Prof. Dr. Nilüfer Tuncer hanımefendiye en derin saygılarımızı arzediyoruz. Büyük soluklu bir bilim insanının yanında, eş, anne, bilim insanı boyutuyla Hanımefendinin, tüm bu gerçekleştirilen çalışmalarda görünmeyen, derin ve sessiz katkıları olmuştur, saygılarımızı arzediyoruz. Evlatları, Tıp Doktoru, Onkoloji Uzmanı Tolga Tuncer'in çok değerli baba ve annesinin soluğunda ülkemize ve insanlığa katkılarını sürdüreceğini ve bunun ailesine ve ülkemize bir armağan olacağını düşünüyoruz.

PROFESSOR YALÇIN TUNCER'S CONTRIBUTION TO OFFICIAL STATISTICS

ABSTRACT

Professor Yalçın Tuncer has honorable upper level contribution to Statistical Science in the Republic of Turkey. This study presents some of the important reflections of the contributions of Professor Yalçın Tuncer to the State Institute of Statistics of the Prime Ministry of Turkish Republic. As the Vice President of the State Institute of Statistics, his contributions to the works of the Methodology Department, Social Statistics Department, and as the Chairman of the 1990 General Population Census of Turkey was highly appreciated. In addition to these, his contribution to the 1991 Turkish Census of Agriculture.

Keywords: Statistical methodology, Social statistics, General population census, Historical statistics series, Mathematics and statistics in Turkey and the world.

Ek: T.C. Başbakanlık Devlet İstatistik Enstitüsü 1990 Yılı Lisansüstü Eğitim Programı

POST-GRADUATE STUDY PROGRAMS IN APPLIED STATISTICS

**STATE INSTITUTE OF STATISTICS POST-
GRADUATE STUDIES, TRAINING AND
RESEARCH CENTER**

TURKEY

STATE INSTITUTE OF STATISTICS POST-GRADUATE STUDY PROGRAMS IN APPLIED STATISTICS

I. INTRODUCTION

The State Institute of Statistics (SIS) conducts special post-graduate study programs in applied statistics. The programs are designed to train personnel to collect, process and analyze statistical data for use in public and private agencies. The medium of instruction of these programs is English.

All programs are full-year academic engagements consisting of classroom and laboratory sessions and field applications. A special emphasis is placed upon computer applications in each area of concern. The programs are offered in the following areas of concentration:

- Sampling and Survey Designs
- Computer Data Systems and Analysis
- Econometrics
- Demography

The post-graduate study programs for these areas are so designed as to provide basic skills and knowledge in each specific area.

Participants can choose any one of the above areas or combine any two of these as Major and Minor areas at a ratio of three to two.

The aim of all the programs is a firm training in both Statistical Theory and Applications. All programs require a mathematical background at an introductory level in standard Calculus and Linear Algebra. Those who have a deficiency in this area can take mathematical courses subject to the approval of the academic advisor in charge. Each participant is assigned an academic advisor who is responsible for arranging a curriculum of courses designed for individual needs.

2. OBJECTIVES AND CURRICULUM

The objectives of all programs are to provide theoretical and practical training for persons in statistical organizations or statistical units of public and private agencies responsible for data collection and processing operations. The programs offered are also beneficial for those who are engaged in research activities and statistical analyses.

The training is carried out through classroom and laboratory sessions, workshops, field exercises and group projects. To provide a general framework for the main practical applications, both theory and principles are presented. However, the major emphasis is on practice. A typical curriculum for each program is designed as in Table-1. Course descriptions and further explanations are given in Sections 7 and 9 below.

3. GENERAL REGULATIONS

All programs require two-semesters of formal training. A semester consists of approximately 12-14 academic weeks, an academic week being made up of five daily training periods. A daily training period in turn is composed of a total of six hours of classroom and laboratory sessions, workshops, field exercises and presentations of projects. Courses are evaluated in terms of credit hours, which are defined as the sum of the weekly theoretical sessions and one half of the weekly application hours (i.e., laboratory sessions, workshops, field exercises and presentations of group projects). Hence, in terms of credit hours, formal training must consist of at least a total of 50 credit hours per year.

A specific calendar of activities may vary slightly each academic year. However, the formal training generally takes place within a period that ranges from October of one calendar year to June of the next.

4. GRADUATION REQUIREMENTS

Following the successful completion of at least a 50 credit-hour course load, each participant is required to prepare and defend a thesis. Upon the successful defense of a thesis, the participant is awarded a certificate.

Successful completion of the prescribed course load means that the participant receives at least a (C) average on the whole and a minimum of (D) in each course. Letter grading corresponds to the following numerical grading:

<u>Points Grade</u>	<u>Letter</u>
90-100	A (Excellent)
80-89	B (Good)
70-79	C (Average)
60-69	D (Poor)
00-59	F (Failure)

Within the first month of the second semester of the formal training, each participant is required to choose a thesis topic with a thesis supervisor under whom the thesis will be prepared. The thesis topic is to be approved by an *ad hoc graduate committee* composed of the thesis supervisor and at least two more experts in the area concerned.

The participant has to complete the preparation and defense of the thesis within one year following enrollment for the training program. The thesis defense and the thesis itself have to be evaluated by a thesis examination committee composed of 2 - 4 experts in addition to the supervisor.

5. PROGRAM FINANCE

Training program fees are charged to participants or their sponsoring agencies. Participation costs other than the program fees consist of travel and subsistence expenses. These latter expenses are also expected to be covered by participants themselves or their sponsors. Program fees include SIS course fees and other fees paid to agencies outside SIS.

SIS does not offer fellowship funds. Therefore, participants and/or their sponsors are expected to meet all related expenses through their own resources. However, there is a limited number of tuition grants to be given to foreign participants from developing countries. The grants are given on a competitive basis.

6. QUALIFICATIONS FOR PARTICIPATION

The SIS Post-Graduate Study Programs in Applied Statistics are primarily designed for statisticians or research workers who have some working experience in statistical organizations or for the future personnel of such organizations. The educational background of these participants may vary, but all are required to have university or other higher-education degrees. Formal university education in Mathematics, Statistics, Economics or other Social and Administrative Sciences is desirable. Participants who want to enroll directly for the basic statistical program required for all SIS Applied Statistical Training Programs must have at least one- year standard calculus and one semester linear algebra background. Participants without this background will be required to make up for their deficiencies.

The minimum required English knowledge of participants must be equivalent to a TOEFL score of 450 or 75 points to be received in the SIS proficiency exam. SIS does not offer deficiency courses in English.

7. COURSE DESCRIPTIONS

7.1. STATISTICS

Probability: Random experiments/Outcome space and event space/Probability as a normalized measure on an event space/Independence and conditional probabilities/Measurable functions and random variables/Convergence modes in probability theory/Probability distributions and their characterizations.

Theory of Statistics: Sampling/Sampling distributions/Basic limit theorems/Fundamentals of estimation: Uniformly minimum risk estimators, uniformly minimum variance estimators/Bayes estimators, minimax estimators, maximum likelihood and least squares estimators/Fundamentals of hypotheses testing: Neyman-Pearson Theorem, monotone likelihood ratios, similar tests, likelihood ratio tests, simultaneous inference techniques such as union-intersection principle and Bonferroni inequality.

Statistical Design of Experiments: Fundamentals of statistical design and analysis of comparative experiments/Randomization, replication, experimental error, blocking

and confounding/Analysis of variance, response levels, orthogonal composite design, rotatable designs, fractional factorial experiments, incomplete block design and mixed models.

Time Series Analysis: Autocovariance and autocorrelation functions/Stationarity/General linear process/Stationary models: AR, MA, ARMA/Nonstationary models: ARIMA/Forecasting/Model identification/Estimation/Diagnostic checks/Seasonal models/Multivariate time series.

Data Analysis: Exploratory data analysis techniques: stem and leaf displays, boxplots, letter-value displays, x-y plotting, resistant line, smoothing data, coded tables, median polish, computer graphics, roctograms, etc.

Nonparametrics: Ranks and order statistics/Randomness, independence and symmetry hypotheses/Nonparametric tests of such hypotheses/Tests of location and scale: sign, median/Wilcoxon tests/Nonparametric measures of association and related tests: Kendall's tau and Spearman's rho/Nonparametric goodness-of-fit tests: Kolmogorov-Smirnov and Chi-square tests.

Regression Analysis: General linear model of full-rank and less-than-full rank/Concepts of product-moment, multiple and partial correlations/General theorems governing simple and multiple regression models/Related estimation and hypotheses testing discussions.

7.2. SAMPLING AND SURVEY DESIGNS

Survey Sampling Techniques: Methods of survey sampling: Probability sampling methods, simple random sampling, stratified element sampling, systematic sampling, equal-sized cluster sampling, unequal-sized cluster sampling, ratio estimation, probability proportional-to-size selection, and multistage sampling.

Survey Research Methods: Survey research methods used in social research/Basic methods in survey research/Methods of data collection/Questionnaire design and construction/ Questionnaire wording and order effects/Data scaling techniques/Organization and administration of fieldwork/Survey errors.

Design of Sample Surveys: Basic methods/Planning of sample surveys/Basic survey designs/Design of field application/Specific survey designs.

Nonsampling Errors and Quality Control in Censuses and Surveys: Types and components of nonsampling errors. Types of nonobservation errors (noncoverage, nonresponse) and observation errors (memory, response, processing). Statistical modeling of nonsampling errors/ Statistical quality control/ Methodologies and indices for response reliability.

Variance Estimation in Sample Surveys: Variance estimation techniques for linear statistics for sample designs such as simple random, stratified element, systematic and equal size clusters/Variance estimation for such nonlinear statistics as ratios and correlation coefficients and for complex sample designs/Methods of variance estimation for random groups, half sample replication, jackknife and bootstrap.

Workshop in Sampling Techniques: Stratification techniques/Controlled selection/Segmenting and listing/Small domains and rare items/Fertility surveys/Sampling establishments/Retaining units/Observational units of variable sizes/Computation of sampling errors.

Longitudinal Survey Design: Repeated cross-section designs/Types of survey designs/Panel designs/Sample designs over time/Repeated studies/Designs for periodic samples/Panel effects/Split-panel designs/Weighting procedures/Combining overlapping samples.

7.3. COMPUTER DATA SYSTEMS AND ANALYSIS

Computer Organization: Number representation/Binary arithmetic units/Memory devices/Basic computer hardware organization/Microprocessors/Operating Cycles: Introduction and their use/Input and output devices and I/O operations.

Introduction to Computers and Information Processing I: Computers and information technology/Evaluation of computers/Basic components of a computer: Computer hardware organization, central processing unit, memory, input/output, secondary storage, software concepts, compilers/interpreters, operating systems, editors, programming languages, system software, auxiliary memory devices, file organization and processing/Flowcharting and structured programming in BASIC, variables, simple BASIC statements, looping, subprogramming in BASIC.

Introduction to Computers and Information Processing II: An overview of PC software packages and Disk Operating System/Word processing and graphics/Spreadsheet concepts/Database management/Query processing and report generation/Use of PC's in data communications/Distributed processing and networking.

Structured Analysis and Design of Information Systems: Fundamental concepts/Information systems life cycles/Classical approach/Structured approach/Structured tools: Hierarchy charts and HIPO, data flow diagrams, SADT, structure charts, Warnier/Orr diagrams, Jackson's notation, entity-relationship model, non-graphical tools. Information systems development methodologies, functional decomposition methodologies, data-oriented methodologies, prescriptive methodologies.

Fundamentals of Structured Programming: Machine, assembly and high level languages/Data representation/Integers, reals, characters, problem solving and algorithm development/Program structures/Introduction to PASCAL.

7.4. ECONOMETRICS

Microeconomic Theory: Theory of the firm/Theory of the consumer/Market equilibrium/Market imperfections/Welfare economics/Introduction to general equilibrium theory/Introduction to game theory/Introduction to uncertainty/Applications.

Macroeconomic Theory: Basic nonstochastic macro models: Classical, Keynesian, Tobin's dynamic aggregative models/Inflation and unemployment: Philips curve, expectations and Okun law, monetarist inflation models/Open economy macro

analysis: Balance of payments and Keynesian analysis, money, capital flows and monetary approach.

Monetary Theory and Policy: Concepts and their measurement/National income/Financial system/Money supply analysis/Demand for money/Interest rate determination/Money and income determination/Money in macro models/Inflation.

Econometrics: Simple regression/Multiple regression/Multicollinearity/Regression on dummy variables/Heteroscedasticity/Autocorrelation/Simultaneous equation models. .

International Trade: The Law of comparative advantage/International equilibrium/Hecksher-Ohlin model/Growth and trade/Theory of tariffs/Domestic distortions and noneconomic objectives/Customs unions.

Economic Statistics: National income accounting/Input-Output/Price and quantity indices/Business, households and agricultural establishments/Time series/Collection of economic data.

7.5. DEMOGRAPHY

Population Dynamics: Determinants and consequences of population trends/History of world population development/Factors of the rate of population growth/Population growth controls/Theories of human population/Demographic transition/Migration.

Techniques of Demographic Analysis: Methods of fertility and mortality analysis/Types and sources of demographic data/Techniques of data standardization/Precision and errors/Life table construction/Use of abridged and model life tables/Estimation of fertility and mortality.

Inference from Incomplete Data: Incomplete data/Indirect methods for estimation of fertility and mortality/Pregnancy history techniques/Brass techniques/Inference from incomplete data.

Demographic Data Collection Methods: Methods to measure population change/Data collection basis for single round surveys, multiround surveys, dual record system/Demographic surveys/Family health surveys/Contraceptive prevalence surveys.

Package Programs and Computer Applications: Computer programs for demographic analysis. Purposes of the subroutines. Installation and operation/Special programs such as MORTPAC, MCPDA, BRASF and BRAMS. Brass and Brass type estimates (i.e., Sullivan, Trussel and Feeney), P/F ratios and other statistics.

Population and Housing Census Methodologies: Concepts in population and its evaluation/Knowledge of map design, map construction and map reproduction methods/Methodology in use of maps for censuses Use of area methods in population and housing censuses/ Enumeration techniques/Blocking and segmenting methodologies/Methods of fieldwork allocation and administration/Methods of tabulation and analysis.

8. FACULTY

Ziya AKTAŞ,	Professor of Computer Science, METU Ph. D.: Lehigh University, USA M.S.: METU B.S.: METU
Öztaş AYHAN,	Associate Professor of Statistics, METU Ph. D.: University of Wales, UK M.S.: University of Aberdeen, UK B.S.: University of Istanbul
Haluk ERLAT,	Professor of Econometrics, METU Ph. D.: University of Pennsylvania, USA M.A.: University of Pennsylvania, USA B.A.: University of Ankara
Orhan GÜVENEN,	Professor of Econometrics, Bilkent University, on leave from University of Paris IX President, State Institute of Statistics Ph. D., M.S.: University of Paris I, France B.S.: University of Istanbul
Haluk KASNAKOĞLU,	Professor of Economics, METU Ph. D.: University of Wisconsin, USA M.S.: METU B.S.: METU
Zehra KASNAKOĞLU,	Associate Professor of Economics, METU Ph. D.: University of Wisconsin, USA M.S.: METU B.S.: METU

Tosun TERZİOĞLU,	Professor of Mathematics, METU Dr. Phil.nat.: Goethe University, Germany B.S.: University of Newcastle upon Tyne, UK
Mehmet R. TOLUN,	Assistant Professor of Computer Science, METU Ph. D.: University of Kent at Canterbury, UK M.S.: University of Kent at Canterbury, UK B.S.: University of Kent at Canterbury, UK
Yalçın TUNCER,	Professor of Statistics, METU Ph. D.: University of Pittsburgh, USA M.S.: University of Pittsburgh, USA B.S.: University of Ankara
Ercan UYGUR,	Professor of Econometrics, University of Ankara Ph. D.: University of East Anglia, UK M.S.: University of Warwick, UK B.S.: METU

9. FURTHER INFORMATION

Further information can be obtained by writing to:

Director of Training Programs
The President's Office
State Institute of Statistics
Necatibey Cad. No: 114
06100 Ankara
TURKEY

TABLE I: C U R R I C U L A

FIRST SEMESTER
(Common for all programs)

A. Statistics	Credit points ⁽¹⁾	
Core Courses:		
	<u>a</u>	<u>b</u>
- Probability	(4.2)	5
- Theory of statistics	(4.2)	5
- Data analysis	(4.2)	5
Electives: (two of the following)		
- Design of experiments	(4.2)	5
- Time series analysis	(4.2)	5
- Nonparametrics	(4.2)	5
- Regression analysis	(4.2)	5

SECOND SEMESTER
(Any of B, C, D and E below ⁽²⁾)

B. Sampling and Survey Designs		
Core Courses:		
- Design of sample surveys	(4.2)	5
- Survey sampling techniques	(4.2)	5
- Survey research methods	(4.2)	5
Electives: (two of the following)		
- Nonsampling Errors and Quality Control in Censuses and Surveys	(4.2)	5
- Variance estimation in sample survey.	(4.2)	5
- Workshop in sampling techniques	(4.2)	5
- Longitudinal survey design	(4.2)	5
C. Computer Data Systems and Analysis		
Core Courses:		
- Computer organization	(4.2)	5
- Introduction to computers and information processing I	(4.2)	5
- Introduction to computers and information processing II	(4.2)	5

Electives: (two of the following)

Fundamentals of structured programming	(4.2)	5
Structured analysis and design of information systems	(4.2)	5
Statistical package programs	(4.2)	5
Databases	(4.2)	5

D. Econometrics

Core Courses:

- Regression analysis	(4.2)	5
- Time series analysis	(4.2)	5
- Econometrics	(4.2)	5
- Economic statistics	(4.2)	5

Electives: (three of the following)

- Microeconomic theory	(4.2)	5
- Macroeconomic theory	(4.2)	5
- Monetary theory and policy	(4.2)	5
- International trade	(4.2)	5

E. Demography Core Courses:

- Population dynamics	(4.2)	5
- Techniques of demographic analysis	(4.2)	5
- Demographic data collection methods	(4.2)	5

Electives: (two of the following)

- Inference from incomplete data	(4.2)	5
- Package programs and computer applications	(4.2)	5
- Population and housing census methodologies	(4.2)	5

NOTES: (1) The meanings of a, b and c are as follows:

- a: Theoretical hours
- b: Recitation hours
- c: Credit hours $a+(b/2)$

(2) Alternatively, three courses chosen from any one of B, C, D, and E and two courses from any one of the remaining three areas will provide major and minor specialization areas. The selected courses in this latter case must be all core courses.

YENİDEN YAPILANMA “BOLOGNA SÜRECİ”

Reşat KASAP*

ÖZET

Bu çalışmada, üniversitelerin yeniden yapılanması olarak bilinen Bologna Süreci ile ilgili genel ve pratik bilgiler özetle verilmiştir.

Anahtar Kelimeler: Bologna Süreci, Yeterlilik, Kalite güvencesi, Hayat boyu öğrenme

1. LİZBON VE BOLOGNA SÜREÇLERİ

Kısaca “Bologna Süreci” olarak adlandırılan Üniversitelerin yeniden yapılanma süreci olarak bilinen bu süreç ile ilgili genel ve pratik bilgiler içeren bu yazı, bu alanda yapılan sunumlardan bir derleme çalışması olup, yararlanılan kaynaklar metin sonunda belirtilmiştir.

Avrupa, Yükseköğretim Alanını yeniden yapılandırmaktadır.

1.1 Lizbon Süreci:

Amaç: Dünyanın en güçlü, bilgiye dayalı rekabet edebilir ekonomisini oluşturmak. Bunun için Avrupa Bilgi Toplumunun ve bunu gerçekleştirecek Avrupa Yükseköğretim Sisteminin oluşturulması.

İtici Güç: Ekonomik Rekabet

Araçlar: Daha fazla rekabet, yüksek kalitede araştırma, mükemmellik, gelişmiş teknoloji, yenilik

1.2 Bologna Süreci:

Amaç: Avrupa Toplumunun (46 Ülke) ihtiyaçlarına uygun Avrupa Yükseköğretim Sisteminin oluşturulması

İtici Güç: Ekonomik ve sosyal ortak hedefler

Araçlar: Daha fazla işbirliği, farklılık, esneklik, keşfedicilik, ve ortak referans noktalar

1.3 Bologna Sürecinin Hedefleri

*Yükseköğretim sistemlerini birbiri ile uyumlu, kolay anlaşılır ve ülke ve uluslararası bir çerçevede tanınır duruma getirmek

*Derece ve diplomaların tanınırlığını sağlamak

*Yükseköğretim kurumları arasında işbirliğini hareketliliği ve mezunların istihdam edilebilirliğini arttırmak

*Uluslararası rekabette yer almak

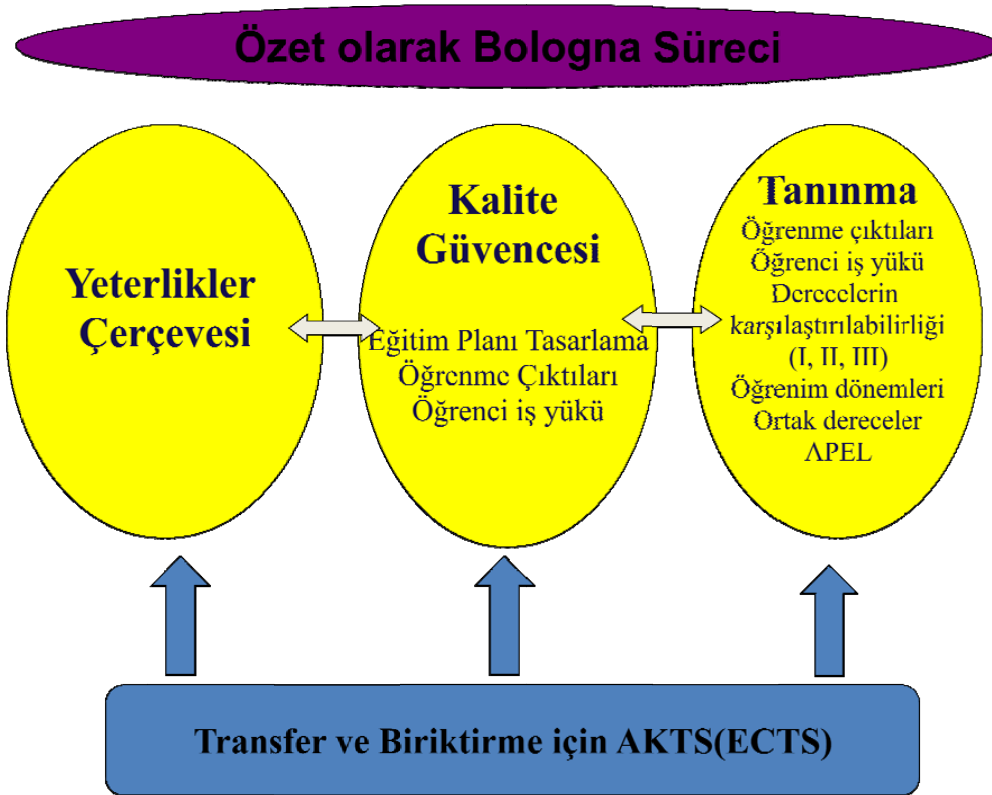
*Öğrenci merkezli uygulamaları ve ders programlarını yeniden oluşturmak

*Yükseköğretimimizi yeterlilikler çerçevesinde yeniden yapılandırarak kalite ve şeffaflığı artırmak

*Prof. Dr., Gazi Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Ankara, e-posta: rkasap@gazi.edu.tr

*Kalite güvencesi sisteminin uygulanmasını sağlamak

*Öğrencilerin, öğretim elemanlarının, mezunların dolaşımını/hareketliliğini
özendirmek, kolaylaştırmak, kalitesini artırmak

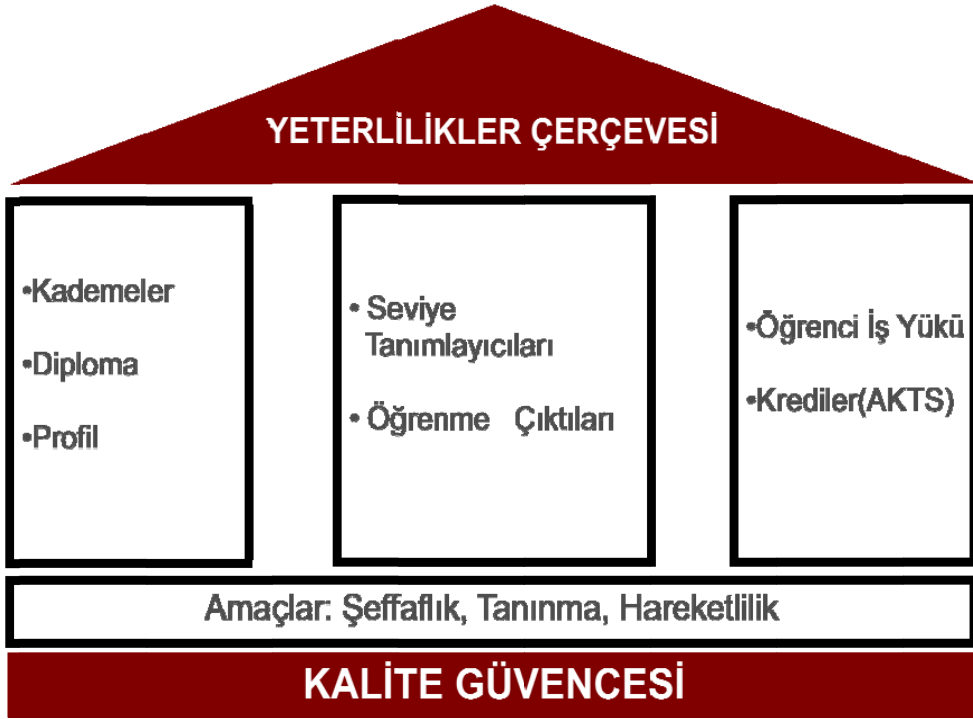


Şekil 1: Özet olarak Bologna Süreci

2. YETERLİLİKLER ÇERÇEVESİ

Yeterlilikler Çerçevesi Ana Unsurları:

- Avrupa Yeterlilikler Çerçevesi
- Türkiye Yeterlilikler Çerçevesi
- Temel Alan Yeterlilikleri
- Program Yeterlilikleri
- Öğrenme Çıktıları ve Yetkinlikler



Şekil 2: Yeterlilik Çerçevesi ana unsurları

2.1 Temel Kavramlar

Yeterlilik: Geçerliliği kabul edilen bir öğretim programının başarıyla tamamlanması sonucu o program için öngörülen öğrenme çıktıların (kazanımların) kazanıldığını onaylayan ve yetkili bir otorite tarafından basılı olarak verilen derece, diploma veya sertifika türü belgedir.

Öğrenme Çıktıları (kazanımları): Bir öğrenme sürecinin tamamlanmasının ardından öğrenenin neleri bileceğinin, neleri yapabileceğinin ve nelere yetkin olacağının ifade edilmesidir.

Yeterlilikler Çerçevesi: Yeterlilikleri, öğrenme çıktılarına göre sınıflandıran ve organize eden yapıdır.

Türkiye Yeterlilikler Çerçevesi (TYÇ): Türkiye düzeyde yükseköğretim yeterlilikleri arasındaki ilişkiyi açıklayan, ülke ve uluslararası paydaşlarca tanınan ve ilişkilendirilebilen, yeterliliklerin belirli bir düzen içerisinde yapılandırıldığı bir sistemdir. Bu sistem aracılığıyla, yükseköğretimde tüm yeterlilikler ve diğer öğrenme kazanımları açıklanabilir ve tutarlı bir şekilde birbiri ile ilişkilendirilebilir.

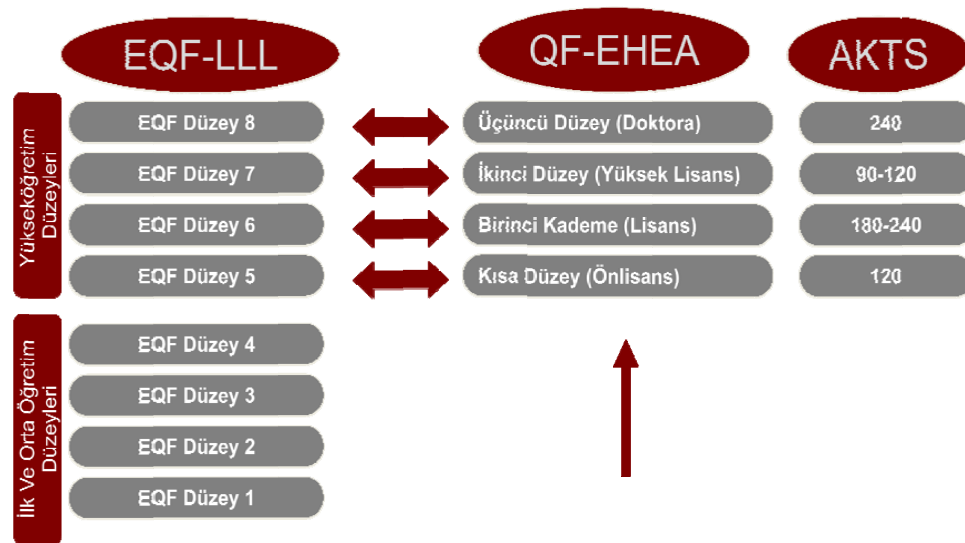
Avrupa Yeterlilikler Çerçevesi: Türkiye Yeterlilikler Çerçevelerinin ilgi tutulabileceği ve bu sayede farklı ülkelerin yeterliliklerinin de birbirleriyle ilişkilendirilebileceği şemsiye (üst) çerçevedir.

2.2 Avrupa Yeterlilikler Çerçevesi

Avrupa Yükseköğretim Alanı için Yeterlilikler Çerçevesi (The overarching framework for qualifications of the EHEA) “QF - EHEA”: Bologna Süreci-Bergen Kararı, 19-20 Mayıs 2005

Yaşam Boyu Öğrenim için Avrupa Yeterlilikler Çerçevesi (European Qualifications Framework for Lifelong Learning) “EQF-LLL”: Avrupa Parlamentosu ve Konseyi tarafından 23 Nisan 2008’de kabul edilmiştir.

İki ayrı çerçevenin yaklaşımları farklı, amaçları aynıdır.



Şekil 3: Avrupa Yeterlilik Çerçevesi

2.3 Türkiye Yükseköğretim Yeterlilikler Çerçevesi

Türkiye Yeterlilikler Çerçevesi (TYÇ) ülkeler ve yükseköğretim için neden önemlidir?

- Düzeyler arasında öğrenme çıktılarına dayalı açık tanımlamalar ve ilişkiler ile yeterliliklerin amaç ve çıktılarını açık hale getirir,
- Farklı yeterlilikler arasındaki bütünlüğü ve örtüşmeyi ortaya çıkararak, yeterliliklerin bir sistem içerisinde birbirleri ile ilişkilendirilebilmelerini ve bu sayede düzeyler arasında ilerlemeyi ve geçişi sağlar,
- Paydaşların katılımı ile kabul edilmiş, eğitime rehberlik eden Türkiye için bir çerçeve sunar;
- Mevcut yeterliliklerin anlaşılması, düzenlenmesi ve geliştirilmesine olanak sağlayan bir yapı oluşturur,
- Yeni yeterliliklerin tasarlanabilmesi için bir içerik sunar.

- Yeterliliklerin reformunu kolaylaştırarak, değişen toplumsal ihtiyaçlara uygun yeni yeterliliklerin geliştirilmesine yardımcı olur
- Öğrenci ve diğer paydaşlara karşı sorumluluklarını yerine getiren (özerk) yükseköğretim kurumlarını destekler,
- Yeterliliklerin vatandaşlar, işverenler ve toplumun tüm kesimleri için rollerini ve yararlarını göstererek bunlara ulaşılmasını teşvik eder;
- Ülke dışında yükseköğretimin tanınmasını ve cazibesini artırır.

TYÇ öğretim elemanları için neden önemlidir?

- Öğrencilerine, öğrenim gördükleri programın ve program içerisindeki derslerin amaçlarını ve bunları başarı ile tamamladıklarında kazanacakları yeterlikleri açıklamalarına yardımcı olur,
- Başka bir eğitim kurumunda alınmış olan derslerin tanınmasını, intibakını ve kredilendirilmesini kolaylaştırır,
- Ders programı değişimini ve içerik geliştirilmesini kolaylaştırır,
- Derslerin ülke ve uluslararası düzeyde daha kolay anlaşılmasını ve tanınmasını sağlar,
- Ölçme ve değerlendirme süreçlerinin öğrenme çıktıları ile ilişkilendirilebilmesini ve bu sayede derslerin başarı ve etkinliğinin geliştirilebilmesine yardımcı olur.

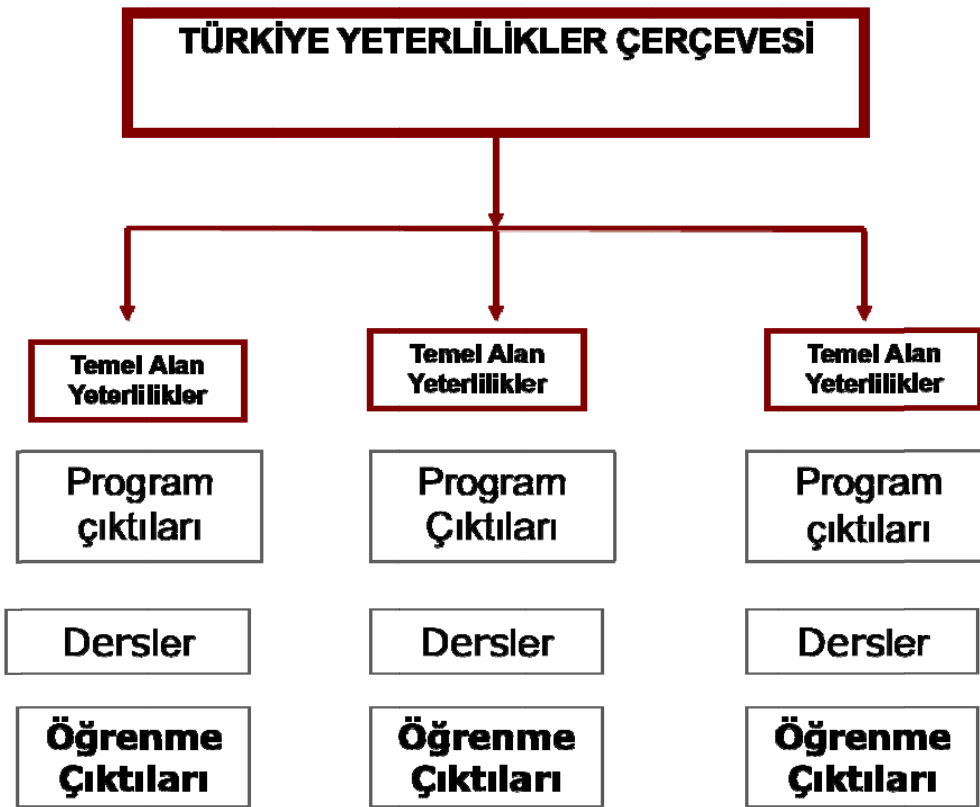
TYÇ öğrenciler için neden önemlidir?

- Öğrencilerin eğitim programlarını ve derslerini bilinçli seçmelerine ve bunları başarı ile tamamladıklarında hangi yeterliliklere (neleri bileceklerine, uygulayabileceklerine ve hangi sosyal ve iletişim yetkinliklerine) sahip olacaklarını önceden bilmelerine yardımcı olur,
- Öğrencilerin öğrenim programları dışındaki aktivitelerinde öğreneceklerinin anlaşılır olmasına yardım eder,
- Öğrenci hareketliliği için engelleri azaltır ve hayat boyu öğrenmeyi teşvik eder,
- Eğitim-Öğretim düzeyleri arasında yatay ve dikey geçişleri anlaşılabilir hale getirir ve kolaylaştır,
- Yeterliliklere giriş ve çıkış noktalarının belirlenmesi ile yeterliliklerin kazanılmasında çeşitli alternatif yollar yaratarak yeterliliklere ulaşımı ve sosyal katılımı geliştirir,
- Ara yeterlilikler dahil bütün yükseköğretim yeterliliklerinin ve kredi aralıklarının kapsamlı bir listesini sağlayarak öğrenenlere yardım ederler ve onlara yönelik bütün eğitim fırsatlarını açıklar.

TYÇ işverenler için neden önemlidir?

- İstihdam edecekleri mezunların neleri bildiğini, bunları hangi ölçüde uygulamaya aktarabileceğini, bilgi ve becerilerinden ne bekleyeceklerini anlamalarına yardım eder,
- Eğitim düzeyleri arasındaki yeterliliklere dayalı farkı ve bunların, ihtiyaçlara göre istihdamına yönelik daha bilinçli tercih yapmalarına yardımcı olur,
- İlgili oldukları alanlarda eğitim-öğretim programlarının geliştirilmesine paydaş olarak katılmalarını ve beklentilerini aktarmalarını kolaylaştırır.

Türkiye Yeterlilikler Çerçevesi belirlendi, sonra ne yapılacak?



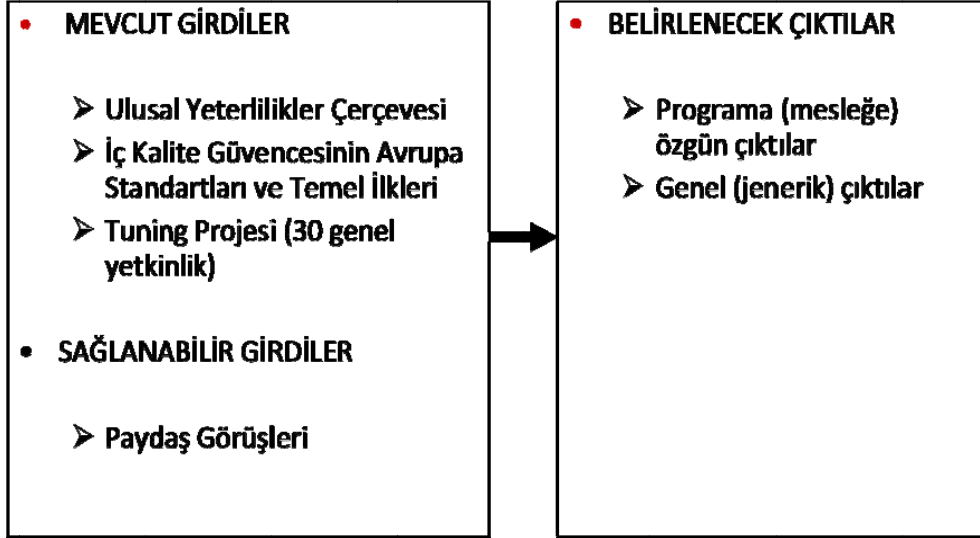
Şekil 4: Türkiye Yeterlilik Çerçevesi

3. PROGRAM YETERLİLİKLERİ

- * Türkiye Yeterliliklerin genel çerçevesi yasalastı.
- * Temel Alan Yeterlilikler çalışmaları tamamlandı.
- * Bölüm/Program yeterlilikleri (çıktıları) belirlemek veya güncellemek için ne yapabilir?
- * Program çıktıları, programı oluşturan modüllerin/derslerin çıktıları ile nasıl ilişkilendirebilir?
- * Fakültede/Bölümde/Programda ne yapılacak?

3.1 Program Çıktılarının Belirlenmesi

Tablo 1. Program Çıktılarının Belirlenmesi



Amaç: Programdan mezun olanların sahip olması beklenen yeterliklerin Türkiye ve Temel Alan yeterlilikler doğrultusunda belirlenmesi (*Kurum misyonuna ve akreditasyon ölçütlerine uygun, sürekli iyileştirme için ölçülüp değerlendirilebilen*)

- Nasıl bir mezun?
- Hangi Yetkinlikler ve Beceriler (Competences and Skills) in kazandırılmış olması bekleniyor?
- Program Eğitim Amaçlarının Belirlenmesi
- Dış Paydaş Görüşleri
(Mezunlar, işverenler, Meslek Odası Temsilcileri, Danışma Kurulu)
- İç Paydaş Görüşleri
(Öğretim Üye/Elemanları, Bölüm Personeli, Öğrenciler)
- Bölüm/Programın misyonu ve vizyonu
(Üniversite misyonu ve vizyonu dikkate alınarak)
- Program Çıktılarının Belirlenmesi

Girdiler

- Türkiye Yeterlilikler Çerçevesi
- Temel Alan Yeterlilikler
- Eğitim Amaçları

Program Çıktıları

1. Programa Özgün Çıktılar
2. Genel (Jenerik) Program Çıktıları

Çıktı (Outcome): Yetkinlik ve beceri seti

3.2 Öğretim Programının Hazırlanması / Gözden Geçirilmesi

* Program çıktılarını sağlayacak eğitim planının hazırlanması veya mevcut planın gözden geçirilerek yeniden şekillendirilmesi

* Derslerin 4 yıla (8 yarıyla) yerleştirilmesi

30 AKTS (ECTS)/yarıyıl

Dikkat: AKTS yerel kredi değildir! Öğrenci iş yüküdür.
60 AKTS(ECTS)/yıl
Toplam 240 AKTS(ECTS) 4 yıl için

Bireysel olarak öğretim üyesi dersle ilgili ne yapacak?

Programdaki herhangi bir ders:

Ders Amaç ve Hedefleri:

Öğrenme Sürecinde öğrencinin öğrenme çıktılarına dayalı öğreneceklerini ölçülebilir biçimde ifade eden öğretmenin geniş ve genel beyanıdır.

Amaç: Genellikle öğretmenin bakış açısından dersin genel içeriğini ve yönünü belirtmek üzere yazılır.

Hedef: Öğretme niyetinin özel bir ifadesidir ve öğretmenin bir derste işleyeceği alanlardan bir tanesini gösterir.

Örneğin, bir dersin amacı

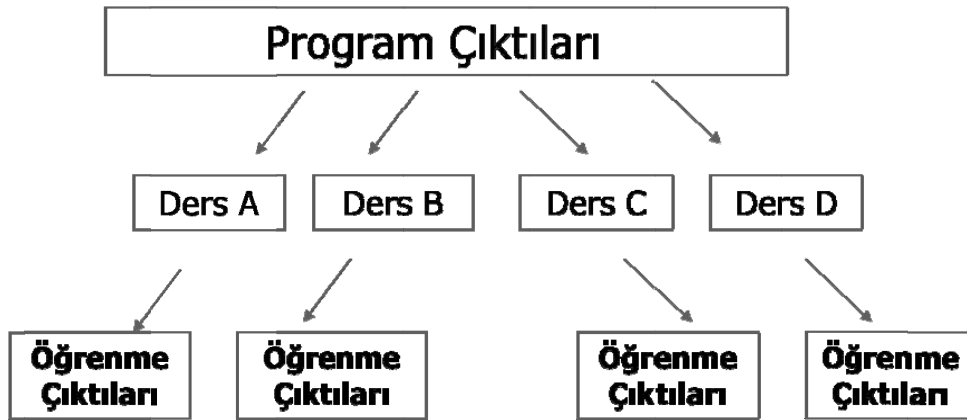
“Öğrencilere kapsamlı bir akışkanlar mekaniği bilgisini geniş bir bakış açısı ile kazandırmak”,

Aynı dersin hedeflerinden birisi; “Öğrencilere bir süreç vererek borulama sistemi tasarımı yapma becerisi kazandırmak” olabilir.

Dolayısıyla bir dersin amacı öğretme niyetinin genel bir ifadesi iken; hedefi derste öğretim etkinliğinin ne başarmayı beklediğidir.

Ders öğrenme çıktıları:

(Öğretmenin ne anlattığına değil, öğrenenin ne kazandığına dayanan öğrenme)



Şekil 5. Ders öğrenme çıktıları

İş yükü tahmini:

Anket: Öğretim Üyeleri

1. Dersinizde hangi program çıktıları kazandırmayı düşünüyorsunuz?
2. Dersinizin öğrenme çıktıları nelerdir?
3. Dersinizde hangi öğrenim aktiviteleri mevcut?

4. Öğrenme çıktıları nasıl ölçüp değerlendiriyorsunuz?
5. Bu öğrenme çıktıları kazandırabilmek için tahmin edilen çalışma süresi nedir?

AKTS/ECTS kredilerinin belirlenmesi:

Tablo2. AKTS/ECTS Kredilerinin belirlenmesinde örnek

Örnek

2.Yıl, 2.Yarıyıl

Akademik Yıl: 34-40 hafta

Haftalık Saat: 40-42 saat

Yıllık Çalışma Saati:1360-1680 saat/yıl

Ortalama Yıllık Çalışma Saati:(1360+1680)/2= 1520 saat/yıl

AKTS kredisi başına düşen iş yükü: $\frac{1520\text{saat/yıl}}{60\text{AKTS/yıl}} = 25-26 \text{ saat/AKTS}$

Ders	Saat/hafta	Tahmin Edilen İş Yükü (saat/yarıyıl)	AKTS Kredisi
A	2	105	4
B	4	160	6
C	4	155	6
D	3	120	5
E	5	203	7
F	2	55	2
Toplam	20	798	30

Sonuç olarak sürdürülebilir gelişme ve kalite kültürü:

Tablo 3. Özetle sürdürülebilir gelişme ve kalite kültürü

- Öğrenci Anketleri Öğrenme Tanınma
- İşveren Anketleri Çıktıları AKTS DE Hareketlilik
- Mezun Anketleri İş Yükü İstihdam Edilebilirlik

4. KAYNAKLAR

<http://www.ehea.info>

<http://www.tyyc.yok.gov.tr>

<http://www.ua.gov.tr>

RESTRUCTURING “BOLOGNA PROCESS”

ABSTRACT

In this study, known as the restructuring of universities in general and practical information related the Bologna Process was summarized.

Keywords: Bologna Process, Qualification, Quality assurance, Lifelong learning

HAZARD CHANGE POINT PROBLEMS FOR CENSORED AND TRUNCATED DATA

Ülkü GÜRLER*

Deniz YENİGÜN**

ABSTRACT

This paper is a summary of the work by Gürler and Yenigün (2011) prepared for the conference in memory of the late Professor Yalçın Tuncer. We consider the estimation of the hazard change point when the observations are subject to random censoring and truncation. The hazard function is assumed to have single jump and otherwise piecewise constant. Full and conditional likelihood approaches are considered and the conditions under which they perform better are discussed.

Keywords: Censoring, Estimation, Hazard Function, Truncation.

1. HAZARD FUNCTION

Hazard function is one of the important functions in reliability, survival, actuarial and other studies, which quantifies the instantaneous risk of failure of an item at a given time point. From statistical point of view, main problems of interest are the estimation of the hazard function, and the estimation of the effects of the covariates. There has been a huge literature on the characteristics and the estimation of hazard functions which are well known, hence the related literature is omitted here. Various functional forms of survival functions that are observed in applications are depicted in Figure 1 below.

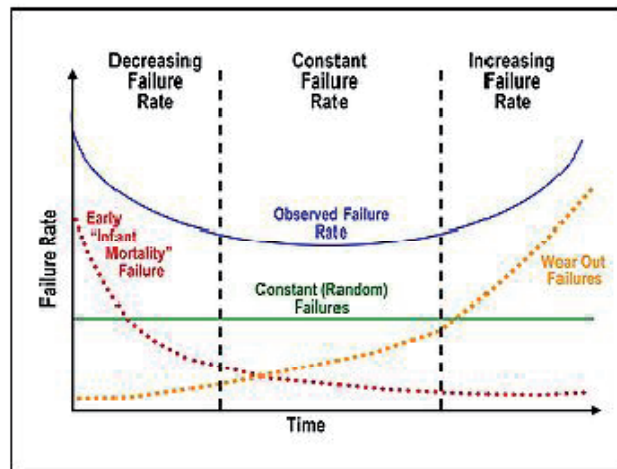


Figure 1: Various failure rate functions

*Prof. Dr., Bilkent University, Department of Industrial Engineering, Ankara, e-mail: ulku@bilkent.edu.tr

**Asst. Prof. Dr., Bilkent University, Faculty of Business Administration, Ankara e-mail: yenigun@bilkent.edu.tr

2. CENSORING AND TRUNCATION

In medical follow-up or reliability studies the random variable of interest (lifetime) may not be fully observable. Common forms of incomplete data are censoring and truncation. There is a vast literature on the analysis of censored and truncated data, including Lynden-Bell (1971), Woodroffe (1985), Kalbeisch and Lawless (1991), Gürler and Wang (1993), Gürler and Prewitt (2000), Frobish and Ebrahimi (2009), Dupuy (2009).

Right Censoring: Right Censoring occurs when event is observed only if it occurs prior to some predetermined time. Let X denote the random variable of interest (lifetime), let C denote the censoring variable. Define $T = \min(X, C)$ and let $\delta = I(T = X)$. Here X and C are independent and nonnegative. Then in the presence of right censoring, one observes the pairs (T, δ) . The observed data is denoted by (t_i, δ_i) for $i = 1, \dots, n$.

Right Truncation: Right Truncation occurs if time origin of lifetime is after the time origin of the study. Let X denote the random variable of interest (lifetime), let Y denote the truncation variable. Here X and Y are independent and nonnegative. The pair (X, Y) is observable only if $X \leq Y$. In other words, observations come from conditional distribution of (X, Y) , given that $X \leq Y$. The observed data is denoted by (x_i, y_i) for $i = 1, \dots, n$.

A well-known example of right truncation is the Transfusion Related (TR) AIDS Data. Here X is the time from transfusion to diagnosis of AIDS, Y is the time from transfusion to the end of study. Various realizations of X and Y are illustrated in Figure 2.

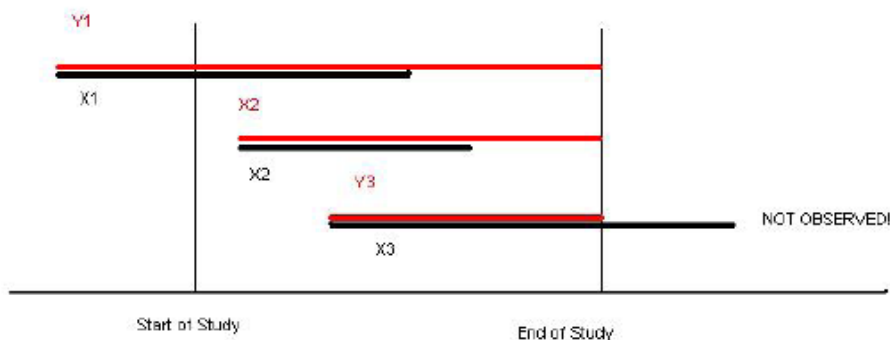


Figure 2: TR-AIDS Data

Left Truncation Right Censoring: Consider a random variable of interest X , representing the time until an event occurs, which may correspond to the survival time of a patient after a treatment or the time until failure of a component. Let Y and C be the truncation and censoring variables respectively, which prevent the complete observation of the variable X . We assume that X , Y , C are independent and

nonnegative. Let $T = \min(X, C)$, and $\delta = I(T = X)$, where I is the indicator function. In the presence of left truncation and right censoring, instead of observing independent and identically distributed (i.i.d.) samples of X , we observe triplets (T, Y, δ) only if $Y \leq T$, otherwise nothing is observed. Thus the observations come from the conditional distribution of (T, Y, δ) , given that $Y \leq T$. The observed data are given by a set of i.i.d. observations (t_i, y_i, δ_i) for $i = 1, \dots, n$.

3. HAZARD CHANGE-POINT MODELS

In some cases, abrupt changes in the hazard function may be observed. For example, lag for effectiveness of a treatment may change over time. It is of interest to detect the location and the size of the change. One of the earliest works that consider changes in the hazard function is by Matthews and Farewell (1982) which studied a piecewise constant hazard model with a single change-point given by

$$\lambda(t) = \begin{cases} \beta & 0 \leq t < \tau \\ \beta + \theta & t \geq \tau \end{cases}, \quad (1)$$

where β and $\beta + \theta > 0$. Here β represents the initial constant hazard rate, θ represents the size of the change in the hazard rate, and τ is the location of the change-point, all of which are unknown.

4. PRELIMINARIES

Suppose X has the hazard function λ as given in (1). Then, the p.d.f. f , the c.d.f. F , the survival function S , and the cumulative hazard function Λ of X are as given below, which are all piecewise functions.

$$f(x) = \begin{cases} \beta e^{-\beta x} \equiv f_1(x) & 0 \leq x < \tau \\ (\beta + \theta) e^{-\beta x - \theta(x-\tau)} \equiv f_2(x) & x \geq \tau \end{cases}, \quad (2)$$

$$F(x) = \begin{cases} 1 - e^{-\beta x} \equiv F_1(x) & 0 \leq x < \tau \\ 1 - e^{-\beta x - \theta(x-\tau)} \equiv F_2(x) & x \geq \tau \end{cases}, \quad (3)$$

$$S(x) = \begin{cases} e^{-\beta x} \equiv S_1(x) & 0 \leq x < \tau \\ e^{-\beta x - \theta(x-\tau)} \equiv S_2(x) & x \geq \tau \end{cases} \quad (4)$$

and

$$\Lambda(x) = \begin{cases} \beta x & 0 \leq x < \tau \\ \beta x + \theta(x - \tau) & x \geq \tau \end{cases}. \quad (5)$$

5. CONSTRUCTING THE LIKELIHOOD FUNCTION WITH THE FULL LIKELIHOOD APPROACH

As described above, in the left truncation and right censoring model one observes triplets (T, Y, δ) only if $Y \leq T$, otherwise nothing is observed. Hence the observed variables belong to the following conditional distribution:

$F_1 \equiv F_1(t, y, \delta | Y \leq T) = P(T \leq t, Y \leq y, \delta | Y \leq T)$. Let $\alpha = P(Y \leq T)$ be the probability that a (Y, T) pair is observed without truncation. We can write α more explicitly as follows:

$$\begin{aligned} \alpha &= P(Y \leq T) = P(Y \leq \min(X, C)) = \int_0^\infty \int_y^\infty \int_y^\infty f(x)h(c)g(y)dc dx dy \\ &= \int_0^\infty \bar{F}(y)\bar{H}(y)dG(y). \end{aligned} \quad (6)$$

We decompose F_1 into two parts, the sub-distribution function of uncensored observations, F_u , and the sub-distribution function of censored observations, F_c . These distributions can be expressed as follows:

$$\begin{aligned} F_u &\equiv F_u(t, y, \delta = 1 | Y \leq T) = P(T \leq t, Y \leq y, \delta = 1 | Y \leq T) \\ &= P(T \leq t, Y \leq y, \delta = 1, Y \leq T)\alpha^{-1} \\ &= \alpha^{-1} \int_0^t \int_u^y \bar{H}(x)dF(x)dG(u). \end{aligned} \quad (7)$$

The corresponding sub-density of censored observations is

$$f_u(t, y) = \frac{\partial F_u}{\partial y \partial t} = \alpha^{-1} g(y)\bar{H}(t)f(t). \quad (8)$$

Similarly, the sub-distribution function of censored observations is

$$\begin{aligned} F_c &\equiv F_c(t, y, \delta = 0 | Y \leq T) = P(T \leq t, Y \leq y, \delta = 0 | Y \leq T) \\ &= P(T \leq t, Y \leq y, \delta = 0, Y \leq T)\alpha^{-1} \\ &= \alpha^{-1} \int_0^y \int_u^t \bar{F}(c)dH(c)dG(u), \end{aligned} \quad (9)$$

and the corresponding sub-density function is

$$f_c(t, y) = \frac{\partial F_c}{\partial y \partial t} = \alpha^{-1} g(y)\bar{F}(t)h(t). \quad (10)$$

Now consider the observed sample (t_i, y_i, δ_i) for $i=1, \dots, n$. The likelihood contribution of an observed uncensored triplet (t_j, y_j, δ_j) for some $j \in \{1, \dots, n\}$ is $f_u(t_j, y_j)$, and the likelihood contribution for an observed censored triplet (t_k, y_k, δ_k) for some $k \in \{1, \dots, n\}$, $k \neq j$, is $f_c(t_k, y_k)$. Then the likelihood function can be written as follows:

$$L = \prod_{i=1}^n \alpha^{-1} g(y_i) [\bar{H}(t_i) f(t_i)]^{\delta_i} [\bar{F}(t_i) h(t_i)]^{1-\delta_i}. \quad (11)$$

6. CONSTRUCTING THE LIKELIHOOD FUNCTION WITH THE CONDITIONAL LIKELIHOOD APPROACH

Klein and Moeschberger (2003) summarized the likelihood construction techniques frequently used in survival analysis literature. According to this construction, various types of censoring and truncation schemes have different contributions to the likelihood function. For example, if X is a random variable of interest with p.d.f. f and survival function S , and if X is subject to right censoring, then the contribution of an observed exact lifetime x to the likelihood function is given by $f(x)$, and the contribution of an observed censoring time c to the likelihood function is given by $S(c)$. When we generalize this approach to the left truncation and right censoring model, we have the following.

Recall that in the left truncation and right censoring model, one observes the triplets (T, Y, δ) only if $Y \leq T$, otherwise nothing is observed. Consider an observed random sample (t_i, y_i, δ_i) for $i=1, \dots, n$. In this case, the contribution of an exact lifetime $(t_i = x_i)$ to the likelihood function is $f(x_i)/S(y_i)$, and the contribution of an observed censoring time $(t_i = c_i)$ to the likelihood function is $S(c_i)/S(y_i)$. Putting together all the components, one may write the conditional likelihood function as

$$L \propto \prod_{i \in D} \frac{f(x_i)}{S(y_i)} \prod_{i \in R} \frac{S(c_i)}{S(y_i)}, \quad (12)$$

where D is the set of observations where the real lifetimes are observed and R is the set observations where the censoring times are observed only.

7. MAXIMUM LIKELIHOOD ESTIMATION

When we construct the likelihood function for the piecewise constant hazard model (1), is not differentiable with respect to τ , hence it is not possible to find the M.L.E.'s for Ψ using standard methods. Therefore, we take the same approach as in Section 3, where we first fix the value of τ and find the M.L.E.'s for the remaining parameters as a

function of τ . Then we search for the value of τ as our estimator, which maximizes the likelihood function over a number of grid points on a specific interval $[\tau_0, \tau_1]$.

8. FULL LIKELIHOOD APPROACH

The censoring variable C and the truncation variable Y are both assumed to have exponential distributions with rate parameters γ and ν respectively. Note that the hazard function of the censoring variable is $\lambda_c(t) = \gamma$, and the hazard function of the truncation variable is $\lambda_y(t) = \nu$.

Consider an observed random sample (t_i, y_i, δ_i) for $i = 1, \dots, n$. For a fixed τ , let A and B denote the set of observations such that $t_i \leq \tau$ and $t_i > \tau$, respectively. Formally, $A = \{i : t_i \leq \tau\}$, $B = \{i : t_i > \tau\}$. Let n_{1A} denote the number uncensored observations that are less than or equal to τ , n_{1B} denote the number of uncensored observations that are larger than τ , and n_B denote the number of observations that are larger than τ . Let $\bar{t}$, $\bar{y}$ and $\bar{\delta}$ denote the sample means. Then for a fixed τ , after some steps the log-likelihood function (12) can be written as

$$\begin{aligned} \log L = n \log \left(\frac{\gamma \nu}{\alpha} \right) - (\gamma + \beta) n \bar{t} - \nu n \bar{y} - \theta \left(\sum_B t_i - n_B \tau \right) \\ + n_{1A} \log \beta + n_{1B} \log (\beta + \theta) - n \bar{\delta} \log \gamma, \end{aligned} \quad (13)$$

where $w = \beta + \gamma + \nu$

$$\text{and } \alpha = P(Y \leq T) = \frac{\nu}{w} - \frac{\nu \theta e^{-w\tau}}{w(w + \theta)}, \quad (14)$$

Taking the derivative of $\log L$ with respect to the unknown parameters, the score vector (13) is computed as

$$U(\Psi_\tau) = \begin{bmatrix} \frac{\partial \log L}{\partial \beta} \\ \frac{\partial \log L}{\partial \theta} \\ \frac{\partial \log L}{\partial \gamma} \\ \frac{\partial \log L}{\partial \nu} \end{bmatrix} = \begin{bmatrix} \frac{nE_1}{\alpha} - \sum_{i=1}^n t_i + \sum_A \delta_i \frac{1}{\beta} + \sum_B \delta_i \frac{1}{\beta + \theta} \\ \frac{nE_2}{\alpha} - \sum_B (t_i - \tau) + \sum_B \delta_i \frac{1}{\beta + \theta} \\ \frac{n}{\alpha} E_1 + \frac{n}{\gamma} - \sum_{i=1}^n t_i - \sum_{i=1}^n \delta_i \frac{1}{\gamma} \\ -n \left(\frac{1}{\nu} - \frac{E_1}{\alpha} \right) + \sum_{i=1}^n \left(\frac{1}{\nu} - y_i \right) \end{bmatrix}.$$

Here the quantities E_1 and E_2 are given by

$$E_1 = \int_0^\infty y \bar{F}(y) \bar{H}(y) dG(y)$$

$$= \frac{\nu}{w^2} + \nu \exp - w \tau \left[\frac{\tau(w + \theta) + 1}{(w + \theta)^2} - \frac{\tau w + 1}{w^2} \right]$$

and

$$E_2 = \int_\tau^\infty (y - \tau) \bar{F}(y) \bar{H}(y) dG(y) = \frac{\nu \exp - w \tau}{(w + \theta)^2}.$$

For the fixed τ , the M.L.E. $\hat{\Psi}_\tau = (\hat{\beta}, \hat{\theta}, \hat{\gamma}, \hat{\nu})_\tau$ for Ψ_τ is obtained as the solution to the system of equations $U(\Psi_\tau) = 0$. This solution is obtained by numerical methods since closed form expressions cannot be obtained.

9. CONDITIONAL LIKELIHOOD APPROACH

Let us start with the problem of finding the M.L.E.'s of β and θ for a fixed τ . Consider an observed random sample (t_i, y_i, δ_i) for $i = 1, \dots, n$. Note that the p.d.f. and the survival function of X are piecewise functions as described in (2) and (4). Then for a fixed τ , there are six possible types of observations that have different contributions to the likelihood function. Let

$$A = \{(t_i, y_i, \delta_i) : \delta_i = 1, y_i < x_i \leq \tau\}, B = \{(t_i, y_i, \delta_i) : \delta_i = 1, y_i \leq \tau < x_i\},$$

$$C = \{(t_i, y_i, \delta_i) : \delta_i = 1, \tau < y_i < x_i\}, D = \{(t_i, y_i, \delta_i) : \delta_i = 0, y_i < c_i \leq \tau\},$$

$$E = \{(t_i, y_i, \delta_i) : \delta_i = 0, y_i \leq \tau < c_i\}, F = \{(t_i, y_i, \delta_i) : \delta_i = 0, \tau < y_i < c_i\}.$$

For example, A denotes the set of observed triplets for which t is an actual lifetime x , and both x and the observed truncation variable y are less than τ . The contribution of such (x, y) pairs to the likelihood function is $f_1(x)/S_1(y)$, where f_1 and S_1 are as described in (2) and (4). We define the sets B , C , D , E , F , and their likelihood contributions similarly. Let n_A denote the number of observed triplets in set A , let $\prod_A$ denote the product over set A , let $n_{B,C}$ denote the total number of observed triplets in sets B and C , and let $\sum_{B,C}$ denote the sum over sets B and C . Define all the other related subscripts similarly. Then for a fixed τ , we can write the likelihood function as follows:

$$L(\beta, \theta | y, t, \delta) = \prod_A \frac{f_1(x_i)}{S_1(y_i)} \prod_B \frac{f_2(x_i)}{S_1(y_i)} \prod_C \frac{f_2(x_i)}{S_2(y_i)} \prod_D \frac{S_1(c_i)}{S_1(y_i)} \prod_E \frac{S_2(c_i)}{S_1(y_i)} \prod_F \frac{S_2(c_i)}{S_2(y_i)}$$

$$= \beta^{n_A} (\beta + \theta)^{n_{B,C}} \exp \left\{ n\beta(\bar{y} - \bar{t}) + \theta \left[\sum_{C,F} y_i - \sum_{B,C} x_i - \sum_{E,F} c_i + n_{B,E}\tau \right] \right\}.$$

Following the notation above, for a fixed τ , let $\Psi_\tau = \{\beta, \theta\}_\tau$ be the parameter set to be estimated, and let $U(\Psi_\tau)$ be the corresponding score vector composed of the first derivatives of the log-likelihood function, which is given by

$$U(\Psi_\tau) = \begin{bmatrix} \frac{\partial \ln L}{\partial \beta} \\ \frac{\partial \ln L}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{n_A}{\beta} + \frac{n_{B,C}}{\beta + \theta} + n(\bar{y} - \bar{t}) \\ \frac{n_{B,C}}{\beta + \theta} + \sum_{C,F} y_i - \sum_{B,C} x_i - \sum_{E,F} c_i + n_{B,E}\tau \end{bmatrix}. \quad (15)$$

Then, the M.L.E. $\hat{\Psi}_\tau = (\hat{\beta}, \hat{\theta})_\tau$ for Ψ_τ is obtained as the solution of the system of equations $U(\Psi_\tau) = 0$, which results in the estimators:

$$\hat{\beta} = \frac{n_A}{\sum_{C,F} y_i - \sum_{B,C} x_i - \sum_{E,F} c_i + n_{B,E}\tau - n(\bar{y} - \bar{t})} \quad (16)$$

$$\text{and } \hat{\theta} = \frac{-n_{B,C}}{\sum_{C,F} y_i - \sum_{B,C} x_i - \sum_{E,F} c_i + n_{B,E}\tau} - \hat{\beta}. \quad (17)$$

Let $\tau_i \in [\tau_0, \tau_1]$, $i = 1, \dots, m$ denote the fixed grid points in the search interval and let L_{τ_i} denote the maximum of the likelihood function for $\tau = \tau_i$. That is

$$L_{\tau_i} = L(\{\hat{\beta}, \hat{\theta}\}_{\tau_i}, \tau_i).$$

Then the proposed estimators for the change-point τ and the rest of the parameters are given by

$$\hat{\tau} = \argmax_{\tau_i} L_{\tau_i} \quad (18)$$

$$\text{and } \hat{\Psi} = \{(\hat{\beta}, \hat{\theta})_{\hat{\tau}}, \hat{\tau}\}. \quad (19)$$

10. COMPARISON OF FULL AND CONDITIONAL LIKELIHOOD METHODS

The full likelihood model specifies parametric families of distributions for the censoring and truncation variables and it is expected to give more accurate results when the model

specification is correct. Its drawbacks are the risk of model misspecification, the large number of parameters to be estimated, and the lack of closed form estimators which forces one to use numerical methods for estimation. Sometimes the numerical methods may not lead to the maximum likelihood estimators, especially for small sample sizes. The conditional likelihood model on the other hand, does not assume any parametric families of distributions for censoring and truncation variables, and focuses only on estimating the model parameters of the hazard function. This simpler approach provides closed form estimators for the model parameters and does not have the risk of model misspecification. The conditional likelihood approach, however, emphasizes more the observed values of the censoring and truncation variables and somewhat overlooks their random nature. This results in increased bias and variance especially for small samples, which disappears as the sample size increases.

For a large set of parameter settings, numerical experiments are done in Gürler and Yenigün (2001) and the following conclusions are drawn: (i) Both methods can easily be implemented and their performances are comparable; (ii) When distributional assumptions of full likelihood method are correct, two methods are close for estimating location of change-point and initial hazard rate; (iii) Full likelihood performs better for estimating size of change and the difference tends to vanish as the sample size increases; (iv) Full likelihood is not robust to model misspecification and in some cases it is outperformed by conditional likelihood model.

11. REFERENCES

- Dupuy, J. F., 2009. Detecting Change in a Hazard Regression Model with Right-Censoring. *Journal of Statistical Planning and Inference* 139, 1578 -1586.
- Frobish, D., Ebrahimi, N., 2009. Parametric Estimation of Change-Points for Actual Event Data in Recurrent Events Models. *Computational Statistics and Data Analysis* 53, 671-682.
- Gürler, Ü., Prewitt, C., 2000. Bivariate Density Estimation with Randomly Truncated Data. *Journal of Multivariate Analysis*, Vol. 74, 88-115.
- Gürler, Ü., Wang, J. L., 1993. Nonparametric Estimation of Hazard Functions and Their Derivatives Under Truncation Model. *The Annals of the Institute of Statistical Mathematics*, Vol. 45, 249-264.
- Gürler, Ü., Yenigün, D., 2001. Full and Conditional Likelihood Approaches for Hazard Change-Point Estimation with Truncated and Censored Data. *Computational Statistics & Data Analysis*, Vol. 55, 2856-2870.
- Kalbfleisch, J. D., Lawless, J. F., 1989. Inference Based on Retrospective Ascertainment: An Analysis of The Data on Transfusion-Related AIDS. *Journal of the American Statistical Association* 84, 360-372.

Klein, J. P., Moeschberger, M.L., 2003. Survival Analysis, Second Ed., Springer, New York.

Lynden-Bell, D., 1971. A Method of Allowing for Known Observational Selection in Small Samples Applied to 3CR Quasars. Monthly Notices of the Royal Astronomical Society, Vol. 155, p. 95.

Matthews, D. E., Farewell, V. T., 1982. On Testing for a Constant Hazard Against a Change-Point Alternative. Biometrics 38, 463-468.

Woodroffe, M., 1986. Estimating a Distribution Function with Truncated Data. The Annals of Statistics. Vol. 13, No. 1., 163-177.

SANSÜRLÜ VE BUDANMIŞ VERİLER İÇİN DEĞİŞİM NOKTASI PROBLEMLERİ

ÖZET

Bu makale, Gürler ve Yenigün'ün (2011) merhum Profesör Yalçın TUNCER'in anısına düzenlenmiş olan konferans için hazırlanmış çalışmasının bir özetidir. Gözlemler rastgele sansür ve budanmaya tabi olduğunda tehlike değişim noktasının tahmini dikkate alınmıştır. Tehlike fonksiyonunun tek bir sıçrama yaptığı ve bunun dışında parçalı sabit olduğu varsayılmıştır. Tam ve koşullu olabilirlik yaklaşımları düşünülmüş ve daha iyi performans gösterdikleri koşullar tartışılmıştır.

Anahtar Kelimeler: Budanmış veri, Hazard fonksiyonu, Kesikli veri, tahmin.

ESTIMATION AND TESTING FOR COINTEGRATION: A SPECTRAL REGRESSION APPROACH

Yılmaz AKDİ*

David A. DICKEY**

ABSTRACT

A popular topic in the econometrics and time series area is the cointegrating relationship among the components of a vector autoregressive time series. The problem became important after the work of Engle and Granger (1987) and has been addressed by many authors: Johansen (1988), Stock and Watson among many others. Engle and Granger's least squares method and Johansen's conditional maximum likelihood method have received the most attention. These tests are routinely applied to economic time series because the notion of cointegration has a natural interpretation. Our method uses low frequency components of the cross periodogram to estimate the cointegration relationship between cointegrated time series. The method improves the results of ordinary least squares method proposed by Engle and Granger in some cases.

Keywords: Time series, Cointegration, Periodogram ordinate, Spectral regression.

1. INTRODUCTION

Unit root tests comprise a standard diagnostic tool in applied time series analysis. There are several procedures to test for a unit root (e.g. Dickey and Fuller, 1979). Test procedures have also been developed to test for seasonal unit roots (Dickey, Hasza and Fuller (1984), Hylleberg, Engle, Granger and Yoo, 1990). Dickey and Pantula (1987) propose a procedure to test for multiple unit roots. Series with unit roots described as integrated. Akdi and Dickey (1998) developed a procedure to test for a unit root using the periodogram ordinates of a univariate time series.

Time series variables with a common (joint) stochastic trend form a cointegrated system. That is, if all of the individual time series are integrated, say of order one, it is sometimes possible that some linear combination of the series will be integrated of order zero (that is, stationary). Thus, the multiple time series $\underline{Y}_t$ is a vector of nonstationary time series, but there exists a vector (or a matrix) $\underline{\beta}$ such that $\underline{\beta}' \underline{Y}_t$ is a stationary system. This notion is known as cointegration and $\underline{\beta}$ is called the cointegrating vector (or matrix).

Engle and Granger (1987) have proposed an estimation procedure for the cointegrating vector. They used a regression approach to estimate the cointegrating vector $\underline{\beta}$. Johansen (1988) gave an estimation procedure that has become very popular. Levy (2002) take advantage of a squared coherency, phase and gain to study the cointegrating relationship for a bivariate cointegrated system. He derives some restrictions by

*Prof. Dr., Department of Statistics, Faculty of Science, Ankara University, Ankara, e-mail: akdi@science.ankara.edu.tr

**Department of Statistics, Nort Carolina State Universty, Raleigh, USA

studying cross-spectral properties of a cointegrated bivariate system. Boswijk and Lucas (2003) considers a semi-nonparametric cointegration test by using LM-testing principles. Breitung (2002) variance ratio testing procedure to test for a unit root and he suggests a generalization of the variance ratio test for cointegration. Chen and Hurvich (2003) study the asymptotic distribution of a tapered narrow-band least squares estimator of the cointegrating vector $\underline{\beta}$ in the framework of fractional cointegration.

Deo and Hurvich (2001) study the estimators based on the log periodogram regression and they obtain the asymptotic bias and variance. They suggest to use low frequencies in the context of the long memory stochastic volatility model. Finite sample properties of spectral regression estimators have been studied by Chambers (2001) with simulation. Marunacci (2000) deals with a somewhat related problem. He considers spectral regression for cointegrated time series with long memory innovations. He provides a functional central limit theorem as a quadratic forms in nonstationary fractionally integrated processes. We investigate an estimation procedure for the cointegrating vector based on the periodogram. For simplicity, bivariate series are considered in detail and the extensions to higher dimensional autoregressive processes are discussed.

In section 2, some notation and definitions are introduced. Section 3 deals with the estimation procedures and consistency results. Several estimation strategies are given. These are regression using the frequency components of the series and different number of frequencies used. Simulation results indicate that using all frequencies in the regression gives the worst result. Even in this case, better results are obtained than those of least squares. Section 4 discusses the extension of the method to higher dimensional processes and finally Section 5 include real data example and a Monte-Carlo simulation study.

2. NOTATION AND MOTIVATION

Consider a first order vector autoregressive (VAR(1)) time series model

$$\underline{Y}_t = A\underline{Y}_{t-1} + \underline{e}_t \quad (1)$$

where $\underline{e}_t$ is a sequence of independent normally distributed random variables with mean vector $\underline{0}$ and variance covariance matrix V . Note that the process is stationary if all eigenvalues of A are less than 1 in absolute value and nonstationary otherwise.

If the coefficient matrix A has distinct eigenvalues, then there are matrices Q and M such that $AQ = QM$ where M is a diagonal matrix of the eigenvalues of A . The transformation $\underline{Z}_t = Q^{-1}\underline{Y}_t$ gives the canonical form of the series. For example, a bivariate series with

$$A = \begin{bmatrix} 1.8 & -0.8 \\ 1.2 & -0.2 \end{bmatrix}, \quad M = \begin{bmatrix} 1 & 0 \\ 0 & 0.6 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$$

the canonical form of the series can be obtained by setting $A = QMQ^{-1}$ as $\underline{Z}_t = M\underline{Z}_{t-1} + \underline{\xi}_t$ where $\underline{Z}_t = (U_t, S_t)'$ and $U_t = U_{t-1} + \xi_{1,t}$ and $S_t = 0.6S_{t-1} + \xi_{2,t}$. Subtracting $\underline{Y}_{t-1}$ from both sides of (1) we have $\underline{Y}_t - \underline{Y}_{t-1} = (A - I)\underline{Y}_{t-1} + \underline{e}_t = \Pi\underline{Y}_{t-1} + \underline{e}_t$. That is, $\nabla \underline{Y}_t = \Pi\underline{Y}_{t-1} + \underline{e}_t$. In our example, notice that the matrix Π can be written as

$$\Pi = A - I = \begin{bmatrix} 1.8 & -0.8 \\ 1.2 & -0.2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.8 & -0.8 \\ 1.2 & -1.2 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 1.2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} = \underline{\alpha}\underline{\beta}'$$

such that $W_t = 0.6W_{t-1} + \xi_t$ is stationary. Here, $W_t = \underline{\beta}' \underline{Y}_t$.

The inverse transformation displays the cointegration relationship. One way to estimate this cointegrating vector in a bivariate case is to regress $Y_{2,t}$ on $Y_{1,t}$. If the residual series is stationary, then the bivariate series is cointegrated. This has been studied by Engle and Granger (1987) and consistency properties have been discussed by Stock (1987).

In a dimension p process, $\underline{Y}_t = A\underline{Y}_{t-1} + \underline{e}_t$ subtract $\underline{Y}_{t-1}$ from both sides to obtain

$$\underline{Y}_t - \underline{Y}_{t-1} = (A - I)\underline{Y}_{t-1} + \underline{e}_t = \Pi\underline{Y}_{t-1} + \underline{e}_t$$

where $\Pi = A - I = \underline{\alpha}\underline{\beta}'$. If Π is of rank r , ($0 < r < p$) and $\underline{\alpha}'\underline{\beta}$ is nonsingular, the series is a linear combination of r stationary and $p - r$ unit root canonical series, as used by Johansen (1988). We have shown Π for our bivariate example.

When a bivariate series $\underline{Y}_t$ is given, each component of the series can be considered as a sum of a stationary and a nonstationary series:

$$\begin{aligned} Y_{1,t} &= q_{11}U_t + q_{12}S_t \\ Y_{2,t} &= q_{21}U_t + q_{22}S_t \end{aligned} \quad (2)$$

where U_t and S_t represent nonstationary (unit root series) and stationary series, respectively. From this representation, it can be seen that $Y_{2,t} - (q_{21}/q_{11})Y_{1,t} = (q_{22} - (q_{21}q_{12}/q_{11}))S_t = cS_t$ is stationary. That is, $\underline{\beta}'\underline{Y}_t$ is stationary when $\underline{\beta}' = (-q_{21}/q_{11}, 1)$. That is, $\underline{\beta}'$ is a cointegrating vector. Of course, the coefficient matrix A is unknown and thus has to be estimated.

Periodograms: Assume that a bivariate series $\underline{Y}_t$ with components $Y_{1,t}$ and $Y_{2,t}$ is given and assume that the representation in (2) is available. For any univariate time series $X_t, t = 1, 2, 3, \dots, n$, the periodogram ordinate at the frequency w_k is defined by

$$I_X(w_k) = \frac{n}{2} (a_k^2 + b_k^2)$$

where $a_k = \frac{2}{n} \sum_{t=1}^n (X_t - \mu) \cos(w_k t)$, $b_k = \frac{2}{n} \sum_{t=1}^n (X_t - \mu) \sin(w_k t)$. Here, μ is the mean of the series and when $w_k = 2\pi k / n$, since $\sum_{t=1}^n \cos(w_k t) = \sum_{t=1}^n \sin(w_k t) = 0$ the periodogram is invariant to the mean whether it is known or estimated.

Given a set of observations $\{Y_1, Y_2, \dots, Y_n\}$ the periodograms of U_t and S_t are defined as in (3) below:

$$I_{u,n}(w_k) = \frac{n}{2} (a_{u,k}^2 + b_{u,k}^2), \quad I_{s,n}(w_k) = \frac{n}{2} (a_{s,k}^2 + b_{s,k}^2) \quad (3)$$

and the real part of the cross periodogram ordinate between U_t and S_t is defined as

$$\text{Re} \{ I_{u,s,n}(w_k) \} = \frac{n}{2} (a_{u,k} a_{s,k} + b_{u,k} b_{s,k}) \quad (4)$$

where $w_k = 2\pi k / n, k = 0, 1, 2, \dots, [n/2]$ (here, $[x]$ denotes the integer part of x). The Fourier coefficients for the series U_t and S_t are

$$\begin{aligned} a_{u,k} &= \frac{2}{n} \sum_{t=1}^n U_t \cos(w_k t) \quad , \quad a_{s,k} = \frac{2}{n} \sum_{t=1}^n S_t \cos(w_k t). \\ b_{u,k} &= \frac{2}{n} \sum_{t=1}^n U_t \sin(w_k t) \quad , \quad b_{s,k} = \frac{2}{n} \sum_{t=1}^n S_t \sin(w_k t). \end{aligned} \quad (5)$$

For any stationary time series X_t , the normalized periodogram ordinate is asymptotically distributed as chi-square with two degrees of freedom.

3. ESTIMATION AND CONSISTENCY RESULTS

In this section, an estimation procedure for the cointegrating vector based on the periodogram ordinates is discussed. Note that, if a vector process is stationary, then each component is marginally stationary. Akdi and Dickey (1998) show that the periodogram of the unit root process U_t satisfies

$$\frac{4k^2 \pi^2}{n^2} I_{u,n}(w_k) \xrightarrow{D} \sigma^2 (Z_1^2 + Z_2^2) \text{ as } n \rightarrow \infty \quad (6)$$

Here, Z_1 and Z_2 are independent standard normal random variables. This result can be used to test for a unit root. The critical values of the distribution under the null hypothesis of a unit root is tabulated in Akdi and Dickey (1998). For a stationary time series $\{S_t, t = 1, 2, 3, \dots, n\}$, $I_{s,n}(w_k) = (n/2)(a_k^2 + b_k^2) = O_p(1)$ as shown in Fuller (1996). This means that the unit root dominates all the asymptotic properties. The cited references show that $a_{u,k} = O_p(n^{1/2})$, $a_{s,k} = O_p(n^{-1/2})$ and $b_{u,k} = O_p(n^{1/2})$, $b_{s,k} = O_p(n^{-1/2})$. Thus we have

$$\begin{aligned} \text{Re} \{ I_{u,s,n}(w_k) \} &= \frac{n}{2} (a_{u,k} a_{s,k} + b_{u,k} b_{s,k}) \\ &= O(n) [O_p(n^{1/2}) O_p(n^{-1/2}) + O_p(n^{1/2}) O_p(n^{-1/2})] = O_p(n). \end{aligned}$$

Theorem 1. Consider the transformation in (2). For ease of notation and to emphasize the analogy to regression, write X_t for $Y_{1,t}$ and Y_t for $Y_{2,t}$. For each fixed k , the ratio

$y_k = \text{Re} \{I_{XY}(w_k)\}$ to $x_k = I_X(w_k)$ is a consistent estimator of the ratio q_{21}/q_{11} . That is, for each fixed k

$$C_n(w_k) = \frac{\text{Re} \{I_{XY}(w_k)\}}{I_X(w_k)} \xrightarrow{P} \frac{q_{21}}{q_{11}}, \text{ as } n \rightarrow \infty.$$

Proof: Notice that

$$b_{X,k}(w_k) = \frac{2}{n} \sum_{t=1}^n X_t \sin(w_k t) = \frac{2}{n} \sum_{t=1}^n (q_{11}U_t + q_{12}S_t) \sin(w_k t) = q_{11}b_{u,k} + q_{12}b_{s,k}$$

and similarly $a_{X,k}(w_k) = q_{11}a_{u,k} + q_{12}a_{s,k}$ and thus

$$\begin{aligned} n^{-2}I_X(w_k) &= (2n)^{-1}(a_{X,k}^2 + b_{X,k}^2) = (2n)^{-1}[(q_{11}a_{u,k} + q_{12}a_{s,k})^2 + (q_{11}b_{u,k} + q_{12}b_{s,k})^2] \\ &= (2n)^{-1}[q_{11}^2(a_{u,k}^2 + b_{u,k}^2)] + O_p(n^{-1}) = q_{11}^2(n^{-2}I_u(w_k)) + O_p(n^{-1}). \end{aligned}$$

Likewise the cross periodogram ordinate of X_t and Y_t can be written as follows:

$$\text{Re} \{n^{-2}I_{XY}(w_k)\} = q_{11}q_{21}(n^{-2}I_u(w_k)) + O_p(n^{-1})$$

and thus,

$$C_n(w_k) = \frac{\text{Re} \{I_{XY}(w_k)\}}{I_X(w_k)} = \frac{\text{Re} \{n^{-2}I_{XY}(w_k)\}}{n^{-2}I_X(w_k)} = \frac{q_{11}q_{21}}{q_{11}^2} + O_p(n^{-1})$$

which completes the proof.

As a result of Theorem 1, when we fix the number of frequencies in the regression, we still have the consistency. That is, consider the following regression model

$$y_k = \beta x_k + \xi_k, \quad k = 1, 2, 3, \dots, m \quad (7)$$

where y_k is the real part of the cross periodogram ordinate of the series X_t and Y_t as defined in (4) at frequency k , and x_k is the periodogram ordinate of the series X_t .

Then the ordinary least squares estimate $\hat{\beta}_{1,n}$ of β in model (7) is a consistent estimator of the ratio q_{21}/q_{11} . That is,

$$\hat{\beta}_{1,n} = \frac{\sum_{k=1}^m x_k y_k}{\sum_{k=1}^m x_k^2} = \frac{\frac{1}{m} \sum_{k=1}^m x_k y_k}{\frac{1}{m} \sum_{k=1}^m x_k^2} \xrightarrow{P} \frac{q_{21}}{q_{11}} \text{ as } n \rightarrow \infty \quad (8)$$

Moreover, using the intercept term in the regression,

$$y_k = \alpha + \beta x_k + \xi_k, \quad k = 1, 2, 3, \dots, m \quad (9)$$

the ordinary least squares estimate of β is still a consistent estimator of the ratio q_{21}/q_{11} . That is,

$$\hat{\beta}_{1,n} = \frac{\sum_{k=1}^m (x_k - \bar{x})(y_k - \bar{y})}{\sum_{k=1}^m (x_k - \bar{x})^2} \xrightarrow{P} \frac{q_{21}}{q_{11}} \text{ as } n \rightarrow \infty. \quad (10)$$

4. HIGHER ORDER SERIES

In this section, higher order and higher dimension vector autoregressive time series are considered. Consider the following time series model

$$\underline{Y}_t = A_1 \underline{Y}_{t-1} + A_2 \underline{Y}_{t-2} + \dots + A_p \underline{Y}_{t-p} + \underline{e}_t \quad (11)$$

where $\underline{Y}_t$ is a k -variate random vector, A_i 's are appropriate matrices and $\underline{e}_t$ is a sequence of *i.i.d.* random vectors with mean-vector $\underline{0}$ and variance-covariance matrix Σ . Subtracting $\underline{Y}_{t-1}$ from both sides, the model becomes $\nabla \underline{Y}_t = \Pi \underline{Y}_{t-1} + B_1 \nabla \underline{Y}_{t-1} + \dots + B_p \nabla \underline{Y}_{t-p} + \underline{e}_t$ where $\Pi = I - A_1 - A_2 - \dots - A_p$, $B_i = -(A_{i+1} + A_{i+2} + \dots + A_p)$. The number, $p - r$, of unit roots of the characteristic equation $|m^p I - m^{p-1} A_1 - \dots - A_p| = 0$ is the rank deficiency of Π . In this case, there exist rank r matrices α and β such that $\Pi = \alpha \beta'$. We assume that each element of the response vector $\underline{Y}_t$ has a unit root and stationary first difference. That is, $\underline{Y}_t$ is integrated of order 1. Thus, it is clear that

$$\nabla \underline{Y}_t = A_1 \underline{Y}_{t-1} + B \underline{W}_t \quad (12)$$

where $\underline{W}_t$, being a linear combination of stationary series, is stationary. Thus, the problem reduces to estimating the coefficient matrix Π .

Using a similar argument to that in section 2 and partitioning $\underline{Y}_t$, the following representation is available for the higher dimensional processes, too.

$$\begin{aligned} \underline{Y}_{1,t} &= Q_{11} \underline{U}_t + Q_{12} \underline{S}_t \\ \underline{Y}_{2,t} &= Q_{21} \underline{U}_t + Q_{22} \underline{S}_t \end{aligned} \quad (13)$$

where $\underline{U}_t$ represents the components with unit roots and $\underline{S}_t$ represents stationary components of the series. Therefore, the problem reduces to estimating $Q_{21} Q_{11}^{-1}$. We are assuming Q_{11} is $(p-r) \times (p-r)$ and of full rank, by which assumption we have identified as a set of series to construct $\underline{Y}_{1,t}$ involves all of the $p-r$ nonstationary common trends. That is, $Q_{11}^{-1} \underline{Y}_{1,t}$ is $\underline{U}_t$ plus stationary components. Only in this way, $\underline{Y}_{1,t}$ can be used to remove all the nonstationary components from the other series. Analogous to the bivariate case $Q_{21} Q_{11}^{-1}$ is to be estimated. Exactly as in the bivariate case, the order of the unit root parts of the Fourier transforms of the data dominate. Specifically, the Fourier coefficient matrices for $\underline{Y}_t$ are:

$$\begin{aligned} \underline{A}_{1,k} &= \frac{2}{n} \sum_{t=1}^n \underline{Y}_{1,t} \cos(w_k t) \quad , \quad \underline{B}_{1,k} = \frac{2}{n} \sum_{t=1}^n \underline{Y}_{1,t} \sin(w_k t) \\ \underline{A}_{2,k} &= \frac{2}{n} \sum_{t=1}^n \underline{Y}_{2,t} \cos(w_k t) \quad , \quad \underline{B}_{2,k} = \frac{2}{n} \sum_{t=1}^n \underline{Y}_{2,t} \sin(w_k t) \end{aligned}$$

and those for the series $\underline{U}_t$ and $\underline{S}_t$ are similar and denoted by lower case letters correspondingly; $\underline{a}_{u,k}$, $\underline{a}_{s,k}$, $\underline{b}_{u,k}$, $\underline{b}_{s,k}$. The relationship between the Fourier coefficients

$\underline{A}_{i,k}$, $\underline{B}_{i,k}$, $i=1,2$ and $\underline{a}_{u,k}$, $\underline{a}_{s,k}$, $\underline{b}_{u,k}$, $\underline{b}_{s,k}$ is as follows:

$$\begin{aligned}\underline{A}_{1,k} &= Q_{11}\underline{a}_{u,k} + Q_{12}\underline{a}_{s,k}, \quad \underline{B}_{1,k} = Q_{11}\underline{b}_{u,k} + Q_{12}\underline{b}_{s,k} \\ \underline{A}_{2,k} &= Q_{21}\underline{a}_{u,k} + Q_{22}\underline{a}_{s,k}, \quad \underline{B}_{2,k} = Q_{21}\underline{b}_{u,k} + Q_{22}\underline{b}_{s,k}\end{aligned}$$

The periodogram ordinate of $\underline{Y}_{1,t}$ can be calculated as follows:

$$P_{11}(w_k) = \frac{n}{2} \left(\underline{A}_{1,k} \underline{A}_{1,k}' + \underline{B}_{1,k} \underline{B}_{1,k}' \right)$$

and the real part of the cross periodogram ordinate will be

$$R_{12}(w_k) = \frac{n}{2} \left(\underline{A}_{1,k} \underline{A}_{2,k}' + \underline{B}_{1,k} \underline{B}_{2,k}' \right).$$

Now define the *cointegrating spectrum* as

$$C_n(w_k) = R_{12}'(w_k) [P_{11}(w_k)]^{-1}. \quad (14)$$

The following are the multivariate versions of Theorem 1.

Theorem 2. The estimator in (14) is a consistent estimator of $Q_{21}Q_{11}^{-1}$ for each fixed k .

As before, a fixed number of frequencies can be combined in a multivariate regression to give a consistent estimate.

Theorem 3. Using the transformations in (13), the ordinary least squares estimator $\hat{\beta}_n$ of $\underline{\beta}$ is a consistent estimator for $Q_{21}Q_{11}^{-1}$ using the regression model

$$Y_k = \beta X_k + E_k, \quad k=1,2,\dots,m$$

where Y_k and X_k are as before and the estimator is defined as

$$\hat{\beta}_n = \left(\sum_{k=1}^m Y_k' X_k \right) \left(\sum_{k=1}^m X_k' X_k \right)^{-1}.$$

5. AN EXAMPLE AND SIMULATIONS

A. Demonstration with Real Data: As an example, we use quarterly US consumption and income data set of Beaulieu and Miron (1993). It covers the period 1946:1 through 1985:4. Data are log transformed and seasonally adjusted using X12 adjustment program. First of all, we check if the series are of $I(1)$. The time series plots and their identification plots are given in Figure 1 below.

Both series are modelled with first order autoregressive model, AR(1) as suggested by the values of AIC and SBC statistics obtained from PROC ARIMA in SAS. The

autocorrelations decay very slowly. Here, X and Y denote income and consumption under the logarithmic transformation, respectively. First differenced series and their identification plots are given in Figure 2.

The stationarity of the series have been checked with standard Dickey-Fuller test and the periodogram based unit root test proposed by Akdi and Dickey (1998). The models for X and Y are

$$X_t = \alpha_{X,0} + \alpha_{X,1}X_{t-1} + e_{X,t}, t = 1,2,3,\dots,160$$

$$Y_t = \alpha_{Y,0} + \alpha_{Y,1}Y_{t-1} + e_{Y,t}, t = 1,2,3,\dots,160$$

and the results have been summarized below:

Table 1. Summary of the Results

	$\hat{\tau}_\mu$	Critical Value	Periodogram	$T_n(w_1)$	Critical Value	Conclusion
X_t	3.201	-2.89	66.9539	688.254	0.178	Unit Root
Y_t	2.415	-2.89	50.6841	300.582	0.178	Unit Root
∇X_t	-13.442	-2.89	0.0040204	0.03953	0.178	Stationary
∇Y_t	-14.811	-2.89	0.0041481	0.024176	0.178	Stationary

The periodogram test is left tailed. We fail to reject the null hypothesis of a unit root for each series. That is, both series are integrated of order 1 according to the results of Dickey-Fuller and periodogram tests. The periodogram based method has certain advantages over conventional tests. Firstly, conventional tests require the estimation of too many AR parameters to account for the dynamics/seasonality of the series. Secondly, test results change with the sample size in conventional tests, while the periodogram based method requires no parameter estimation except for variance. Thirdly, the critical values of the test statistics are free of sample size constraints. Thus, these might have considerable advantages, especially for small samples.

In order to check whether these series are cointegrated, we calculate the periodograms and the real part of the cross periodograms of the series and regress the real part of the cross periodograms on the periodogram of one of the series (income). The estimate of the regression coefficient with spectral regression is $0.86814 \approx 0.87$. The time series and its identification plots of $Z_t = Y_t - 0.87X_t$ given in Figure 3, seem to show a stationary series but this requires a statistical check. If this series is stationary, then the estimated cointegration vector is $\underline{\beta} = (-0.87, 1)'$. The values of AIC statistic and the identification plots imply an AR(1) model for Z_t , $Z_t = \alpha Z_{t-1} + v_t$. To test $H_0: \alpha = 1$ against $H_a: \alpha < 1$ one can regress $D_t = Z_t - Z_{t-1}$ on Z_{t-1} and calculate the standard t -statistic as: $\hat{\tau}_a = \hat{\alpha} / s(\hat{\alpha}) \approx -3.325$ which is smaller than 10% critical value -3.134 (obtained by simulation- see Table 2) and we reject H_0 at the 10% level. The parameter estimates and their standard errors are given below:

$$\begin{array}{rcl} \nabla Z_t & = & \hat{\alpha}_0^* + \hat{\alpha}_1^* Z_{t-1} \\ \text{Est.} & & -0.0192 \quad -0.104 \\ \text{Std.Err.} & & 0.0056 \quad 0.0313 \\ t-stat & & -3.428 \quad -3.325 \end{array}$$

The power is approximated as 0.804 which is obtained in a similar way given in Table 3 in the Annex.

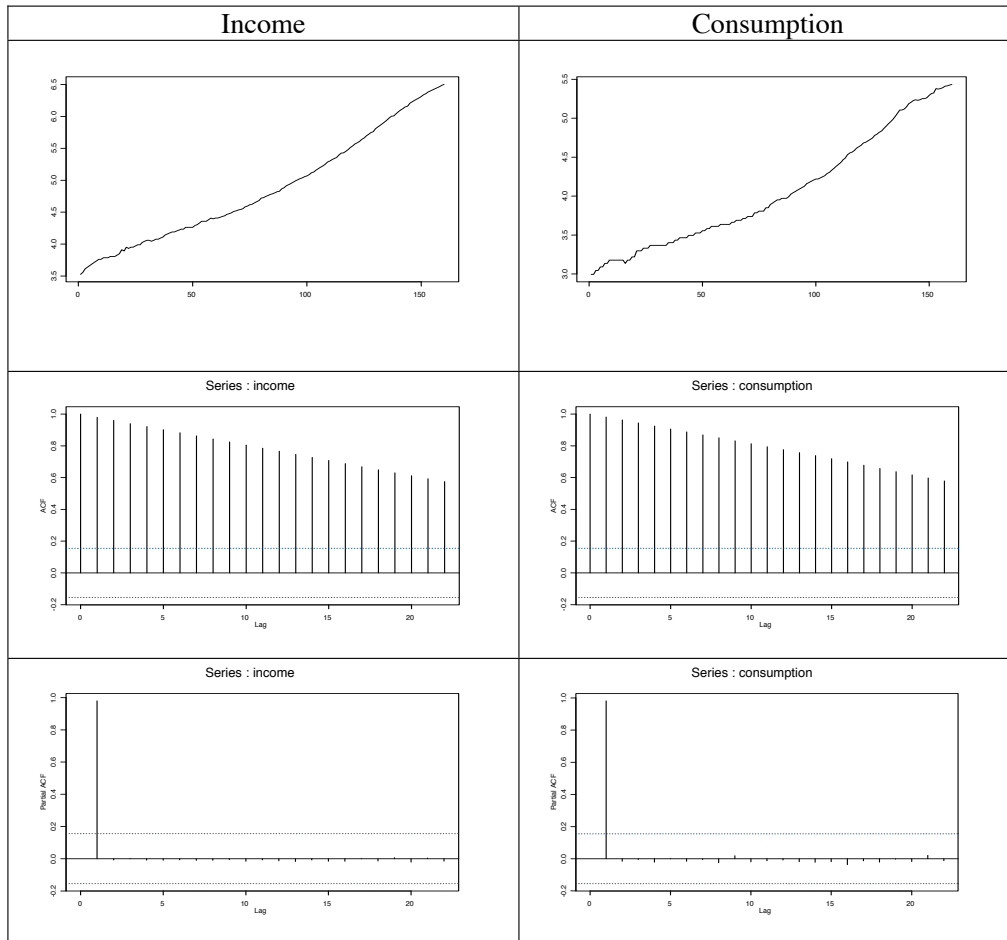


Figure 1. Original Series and Their Identification Plots

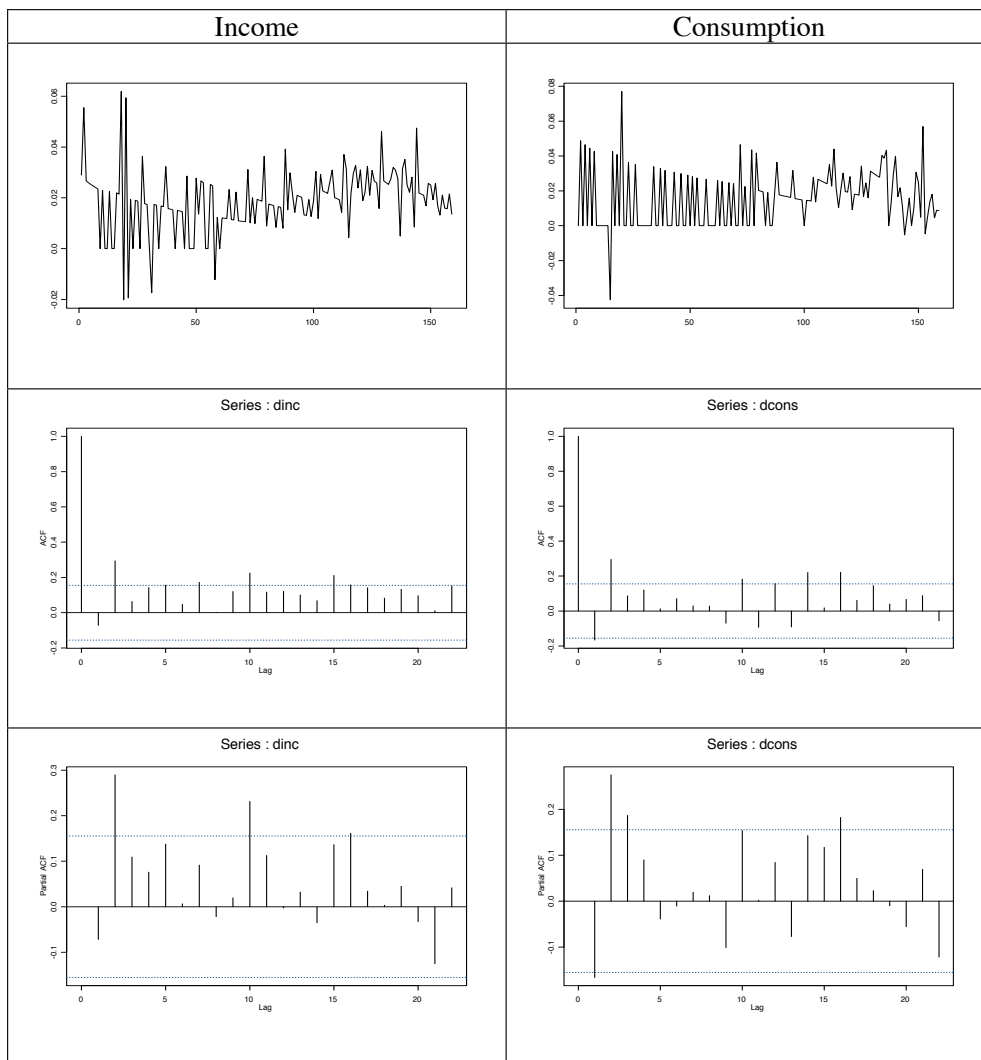


Figure 2. First Differences and Their Identification Plots

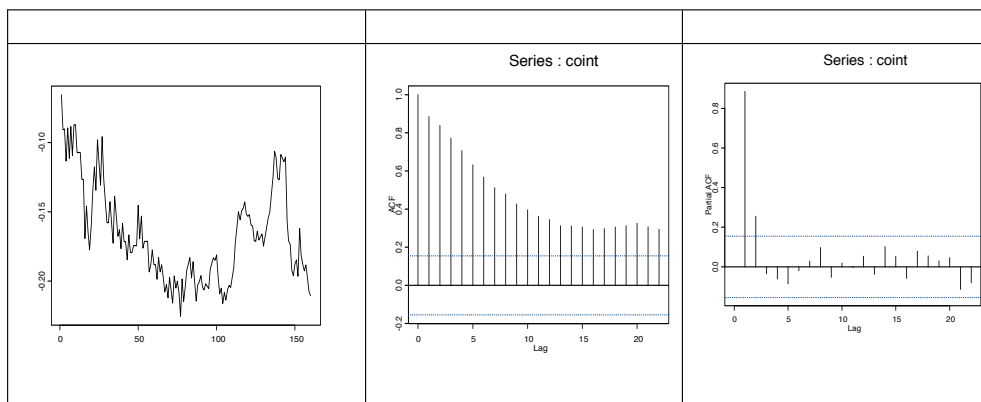


Figure 3. Cointegrated Series

This critical value comes from a simulation reported in Table 2. Bivariate series with roots 1 and ρ were generated, Z was computed by the periodogram method shown here and then the differenced Z was regressed on its first lag Z and empirical percentiles computed. The percentiles of τ_a can be used as the critical values for testing the null hypothesis of no cointegration. The table is similar to the table of Engle and Granger but here the periodogram method is used rather than ordinary regression to construct Z . In other words, adjusted consumption and income are cointegrated at 10% significance level. In a similar way, the performance of the test is tabulated for different ρ 's as listed in Table 3. Here, 1000 replications are used and for different ρ 's the number of rejections were counted.

In addition, the cointegrating relationship between consumption and income is analyzed with the Engle and Granger approach. The estimated cointegrating vector ($0.856996 \approx 0.86$) is very close to the one estimated through the spectral regression approach. When the first differenced residual series R_t (obtained from the regression of Y_t on X_t) is regressed on R_{t-1} , we get the following estimation results:

$$\begin{array}{rcl} \nabla \hat{R}_t & = & -0.000684 - 0.10886 \hat{R}_{t-1} \\ Std.Err & & (0.0011) \quad (0.0327) \\ t-stat & & -0.620 \quad -3.326 \end{array}$$

Since, $\hat{\tau}_\mu(E - G) = -3.326 < -3.033$, we reject the null hypothesis of no cointegration at the 10% level. The critical value -3.033 is obtained by simulation with 5000 replications. Note that both tests fail to reject the null hypothesis of no cointegration at 5% significance level.

Moreover, the Johansen (1988) method is also applied for investigating the cointegration relation between these variables. The corresponding squared canonical correlations are 0.099310 and 0.050393. Then, the value of Johansen's trace statistics is calculated as $\hat{\lambda}_{trace} = -160[\ln(1 - 0.099310) + \ln(1 - 0.050393)] = 25.008$ which is greater than 10% critical value (10.3) and we reject the null hypothesis of no cointegration. Thus, we find that the consumption and income series are cointegrated at 10% significance level.

B. Simulation Study: We consider the series generated from a normal distribution with mean zero and variance 1 according to (15). That is, $e_{1,t} \sim N(0,1)$, $e_{2,t} \sim N(0,1)$ with $U_t = \rho_1 U_{t-1} + e_{1,t}$ and $S_t = \rho_2 S_{t-1} + e_{2,t}$ for $\rho_1 = 1$, $|\rho_2| < 1$. The equations in (2) allow us to write

$$\begin{aligned} Z_{1,t} &= U_t + 2S_t \\ Z_{2,t} &= 3U_t + 4S_t \end{aligned} \tag{15}$$

where U_t represent a unit root time series and S_t represents a stationary time series. Notice that both series $Z_{1,t}$ and $Z_{2,t}$ are nonstationary because both include U_t . That is, the nonstationary bivariate series can be written as

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} U_t \\ S_t \end{bmatrix}. \quad (16)$$

But $Z_{2,t} - 3Z_{1,t} = -2S_t$ which is stationary. Based on this cointegration relationship, in the following simulations we expect to get estimates close to 3 for different numbers of observations. We run 10,000 regressions for each case (different ρ_2 and n) and average these spectral regression estimates of the cointegrating vectors. We also run 10,000 regressions of $Z_{2,t}$ on $Z_{1,t}$ for each case and average these 10,000 ordinary least squares estimates of the cointegrating coefficient as in Engle and Granger (1987), labeled OLS in Table 1a. All frequency regressions does not have fixed k as n gets large so our asymptotic theory does not apply to it. Different number of frequencies were considered in the regressions and we observed that using all frequencies in the regression gives the worst result. We also analyze the bivariate series with Johansen's method. Standard deviations and mean squared errors of estimation are reported for comparison. For the bivariate case, the sample sizes used as $n = 50, 100, 200$. The values of the parameters $\rho_1 = 1$ and $\rho_2 = 0.7, 0.8, 0.9, 0.95, 0.99$ are considered and the results are tabulated in Table 1b. Notice that there is no cointegration when $\rho_2 = 1$. In our simulations, we observe that Johansen's method produces some bizarre outliers which cause large mean squared error and standard deviations and therefore the bias of Johansen's method is large.

In Table 1a, we take the average of 10,000 regression estimates obtained from both ordinary least squares and spectral regression for different values of ρ_2 . We are looking for the values close to 3. Using all frequencies in the regression yields the worse results. Even in the worst case spectral regression gives a better result than that obtained from ordinary least squares. As it is seen from the table, the estimates gets further away from 3 as ρ_2 approaches 1. In Table 1b, SRE is our periodogram estimator using all frequencies in the regression. When ρ_2 takes the value 1, there is no cointegration because the matrix Π has rank of zero.

6. CONCLUSION

In this study, periodogram based cointegration method have been proposed. The method improves the OLS method proposed by Engle and Granger (1987) in certain cases.

7. REFERENCES

Akdi, Y., Dickey, D. A., 1998. Periodograms of Unit Root Time Series: Distributions and Tests, Communications in Statistics: Theory and Methods, 27, 69-87.

- Beaulieu, J., Miron, J. A., 1993. Seasonal Unit Roots in Aggregate US Data, *Journal of Econometrics*, 55, 305-328.
- Bloomfield, P., 1976. *Fourier Analysis of Time Series: An Introduction*, Wiley, New York.
- Boswijk, H. P., Lucas, A., 2002. Semi-nonparametric Cointegration Testing, *Journal of Econometrics*, 108, 253-280.
- Breitung, J., 2002. Nonparametric Tests for Unit Roots and Cointegration, *Journal of Econometrics*, 108, 343-368.
- Chambers, M. J., 2001. Temporal Aggregation and the Finite Sample Performance of Spectral Regression Estimators in the Cointegrated System: A Simulation Study, *Econometric Theory*, Vol. 17, Number 3, 591-607.
- Chen, W. W., Hurvich, C. M., 2003. Estimating Fractional Cointegration in the Presence of Polynomial Trends, *Journal of Econometrics*, 117, 95-121.
- Choi, I., Phillips, P. B. C., 1993. Testing for a Unit Root by Frequency Domain Regression, *Journal of Economics*, 59, 263-286.
- Deo, R. S., Hurvich, C. M., 2001. On the log Periodogram Regression Estimators of the memory Parameters in the Long Memory Volatility Models, *Econometric Theory*, Vol. 17, Number 4, 686-710.
- Dickey, D. A., Fuller, W. A., 1979. Distributions of Estimators for Autoregressive Times Series with a Unit Root, *Journal of American Statistical Association*, 74, 427-431.
- Dickey, D. A., Pantula, S. G., 1987. Determining the order of Differencing in Autoregressive Processes, *Journal of Business and Economics Statistics*, 5, 455-461.
- Dickey D. A., Hasza, D. P., Fuller, W. A., 1984. Testing For Unit Roots in Seasonal Time Series, *Journal of American Statistical Association*, 79, 355-367.
- Engle, R.F., C.W.J. Granger., 1987. Cointegration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55, 251-276.
- Fuller, W.A., 1996. *Introduction to Statistical Time Series*, Wiley, New York.
- Hylleberg, S., Engle, R. F., Granger, C. W. J., Yoo, B. S., 1990. Seasonal Integration and Cointegration, *Journal of Econometrics*, 99, 215-238.
- Johansen, S., 1988. Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.
- Levy, D., 2002. Cointegration in Frequency Domain, *Journal of Time Series Analysis*, Vol.23, No.3, 333-339.

Marunicci, D., 2000. Spectral regression For Cointegrated Time Series With Long-Memory Innovations, Journal of Time Series Analysis, Vol. 21, Issue 6, 685-705.

Stock, J. H., 1987. Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors, Econometrica, 55, 1035-1056.

Stock, J. H., Watson, M. W., 1993. A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems, Econometrica, 61, 783-820.

EŞBÜTÜNLEŞME İÇİN TAHMİN VE TEST: SPEKTRAL BİR REGRESYON YAKLAŞIMI

ÖZET

Ekonometri ve zaman serileri alanındaki popüler bir konu, vektör otoregressif zaman serilerinin bileşenleri arasındaki eşbütünleşme ilişkisidir. Engle ve Granger (1987)'in çalışmalarından sonra problem önemli hale gelmiş ve Johansen (1988), Stock ve Watson gibi başka pek çok yazar tarafından da bu probleme işaret edilmiştir. Engle ve Granger'in en küçük kareler metodu ile Johansen'in koşullu maksimum olabilirlik metodu en çok dikkati çekenlerdendir. Bu testler ekonomik zaman serilerine rutin olarak uygulanmıştır çünkü eşbütünleşme nosyonu doğal bir yoruma sahiptir. Bizim metodumuz, eşbütünleşmiş zaman serileri arasındaki eşbütünleşme ilişkisini tahmin etmek için çapraz periyodogramın düşük frekanslı bileşenlerini kullanır. Bazı durumlarda bu metod, Engle ve Granger tarafından önerilen sıradan en küçük kareler metodunun sonuçlarını geliştirir.

Anahtar Kelimeler: Zaman serileri, Eşbütünleşme, Periyodogram ordinat, Spektral regresyon.

Annex. Tables

Table 1a. Simulations (* marks the closest average to 3 in each row)

	5 Freq	10 Freq.	All Freq.	OLS
n=50				
$\rho_2 = 0.7$	2.641*	2.619	2.600	2.482
$\rho_2 = 0.8$	2.524*	2.511	2.498	2.421
$\rho_2 = 0.9$	2.273	2.367*	2.362	2.329
$\rho_2 = 0.95$	2.282*	2.280	2.279	2.268
$\rho_2 = 0.99$	2.221*	2.221	2.220	2.217
n=100				
$\rho_2 = 0.7$	2.843*	2.822	2.786	2.617
$\rho_2 = 0.8$	2.736*	2.715	2.684	2.552
$\rho_2 = 0.9$	2.536*	2.523	2.507	2.436
$\rho_2 = 0.95$	2.380*	2.373	2.367	2.337
$\rho_2 = 0.99$	2.235*	2.234	2.233	2.231
n=200				
$\rho_2 = 0.7$	2.957*	2.948	2.916	2.746
$\rho_2 = 0.8$	2.905*	2.890	2.852	2.687
$\rho_2 = 0.9$	2.746*	2.727	2.693	2.569
$\rho_2 = 0.95$	2.546*	2.533	2.514	2.446
$\rho_2 = 0.99$	2.267*	2.267	2.265	2.260

Table 1b. Simulations (Comparisons for SRE with OLS and Johansen with respect to MSE and STD, used all frequencies in SRE)

N=50	$\rho_2 = 0.7$	$\rho_2 = 0.8$	$\rho_2 = 0.9$	$\rho_2 = 0.95$	$\rho_2 = 0.99$
SRE	2.600	2.498	2.362	2.279	2.220
MSE(SRE)	0.245	0.344	0.510	0.626	0.715
STD(SRE)	0.292	0.305	0.321	0.325	0.327
OLS(EG)	2.482	2.421	2.329	2.268	2.217
MSE(EG)	0.311	0.381	0.310	0.238	0.247
STD(EG)	0.205	0.215	0.310	0.238	0.247
JOH	3.632	2.408	3.286	8.48	2.449
MSE(JOH)	3482.5	7081.5	833.22	36382835	2427.5
STD(JOH)	59.01	84.15	28.87	603.187	49.268
N=100					
SRE	2.786	2.684	2.507	2.367	2.233
MSE(SRE)	0.097	0.169	0.333	0.500	0.693
STD(SRE)	0.227	0.264	0.300	0.315	0.326
OLS(EG)	2.617	2.552	2.436	2.337	2.231
MSE(EG)	0.178	0.238	0.364	0.491	0.652
STD(EG)	0.179	0.194	0.215	0.228	0.246
JOH	2.968	2.936	-70.28	1.0248	2.178
MSE(JOH)	193.68	257.95	52078513	17247	1610.06
STD(JOH)	13.91	16.06	7216.53	131.319	40.119
N=200					
SRE	2.916	2.852	2.693	2.514	2.265
MSE(SRE)	0.026	0.057	0.161	0.324	0.649
STD(SRE)	0.168	0.188	0.258	0.297	0.332
OLS(EG)	2.746	2.687	2.569	2.446	2.260
MSE(EG)	0.084	0.124	0.223	0.353	0.608
STD(EG)	0.141	0.162	0.193	0.216	0.248
JOH	3.014	3.051	3.527	9.018	2.789
MSE(JOH)	0.021	8.733	3028.01	35054277	14941.76
STD(JOH)	0.144	2.95	55.03	592.06	122.24

Table 2. Critical values for τ_a (5,000 replicates) The values for $\rho = 1$ are used for testing the null hypothesis of no cointegration.

	0.01	0.05	0.10	0.90	0.95	0.99
N=50						
$\rho = 1$	-4.121	-3.464	-3.173	-0.885	-0.488	0.290
$\rho = 0.95$	-4.225	-3.613	-3.264	-1.203	-0.880	-0.300
$\rho = 0.90$	-4.318	-3.685	-3.336	-1.412	-1.148	-0.593
$\rho = 0.80$	-4.498	-3.906	-3.610	-1.774	-1.530	-1.079
$\rho = 0.70$	-4.788	-4.164	-3.873	-2.065	-1.851	-1.419
N=100						
$\rho = 1$	-4.125	-3.434	-3.114	-0.967	-0.602	0.188
$\rho = 0.95$	-4.183	-3.659	-3.370	-1.523	-1.255	-0.677
$\rho = 0.90$	-4.420	-3.938	-3.655	-1.936	-1.714	-1.226
$\rho = 0.80$	-5.146	-4.513	-4.233	-2.563	-2.368	-1.982
$\rho = 0.70$	-5.650	-5.032	-4.737	-3.031	-2.835	-2.489
N=160						
$\rho = 1$	-4.050	-3.415	-3.134	-1.005	-0.588	0.159
$\rho = 0.95$	-4.330	-3.810	-3.544	-1.820	-1.605	-1.133
$\rho = 0.90$	-4.815	-4.305	-4.043	-2.426	-2.231	-1.823
$\rho = 0.80$	-5.726	-5.183	-4.924	-3.318	-3.132	-2.735
$\rho = 0.70$	-6.432	-5.880	-5.629	-3.969	-3.796	-3.403
N=200						
$\rho = 1$	-3.994	-4.424	-3.121	-1.047	-0.646	0.105
$\rho = 0.95$	-4.446	-3.925	-6.643	-1.995	-1.781	-1.317
$\rho = 0.90$	-5.022	-4.544	-4.278	-2.703	-2.502	-2.107
$\rho = 0.80$	-6.043	-5.538	-5.307	-3.734	-3.526	-3.099
$\rho = 0.70$	-6.807	-6.361	-6.103	-4.494	-4.261	-3.848

Table 3. Power of τ_a (number of rejections out of 1,000 replications)

$\alpha = 0.05$				$\alpha = 0.10$			
$\rho \downarrow N \rightarrow$	100	160	200	$\rho \downarrow N \rightarrow$	100	160	200
1.00	0.041	0.045	0.047	1.00	0.107	0.093	0.103
0.95	0.087	0.133	0.169	0.95	0.160	0.213	0.290
0.90	0.167	0.329	0.516	0.90	0.287	0.533	0.705
0.80	0.452	0.866	0.960	0.80	0.649	0.950	0.978
0.70	0.734	0.984	1.000	0.70	0.868	0.996	1.000
0.60	0.886	1.000	1.000	0.60	0.964	1.000	1.000
0.50	0.965	1.000	1.000	0.50	0.992	1.000	1.000

DANIřMA KURULU ÜYELERİ - ADVISORY BOARD MEMBERS

Ali YAZICI
Alper GÜVEL
Asaf Savaş AKAT
Aşır GENÇ
Aydın ÖZTÜRK
Ayşe GÜNDÜZ HOŞGÖR
Bedriye SARAÇOĞLU
Coşkun Can AKTAN
Deniz GÖKÇE
Ekrem ERDEM
Ercan UYGUR
Erdem BAŞCI
Erinç YELDAN
Erol TAYMAZ
Eser KARAKAŞ
Fatih ÖZATAY
Fatin SEZGİN
Fikri AKDENİZ
Fikri ÖZTÜRK
Gülay BAŞARIR KIROĞLU
Güven SAK
Haluk LEVENT
Hamza EROL
İlhan TEKELİ
İmdat KARA
İnsan TUNALI
Levent KANDİLLER
Mehmet KAYTAZ
Meltem DAYIOĞLU TAYFUR
Metin TOPRAK
Mustafa ACAR
Mustafa AYTAÇ
Nihat BOZDAĞ
Onur BASKAN
Orhan GÜVENEN
Ömer Faruk ÇOLAK
Ömer L. GEBİZLİOĞLU
Özkan ÜNVER
Öztaş AYHAN
Reşat KASAP
Savaş ALPAY
Seyfettin GÜRSOY
Süleyman GÜNAY
Turan EROL
Ümit OKTAY FIRAT
Yasin AKTAY
Yılmaz AKDİ

Atılım Üniversitesi
Çukurova Üniversitesi
Bilgi Üniversitesi
Selçuk Üniversitesi
Ege Üniversitesi
Orta Doğu Teknik Üniversitesi
Gazi Üniversitesi
Dokuz Eylül Üniversitesi
Bahçeşehir Üniversitesi
Erciyes Üniversitesi
Türkiye Ekonomi Kurumu
T.C. Merkez Bankası
Bilkent Üniversitesi
Orta Doğu Teknik Üniversitesi
Bahçeşehir Üniversitesi
TOBB Ekonomi ve Teknoloji Üniversitesi
Bilkent Üniversitesi
Çukurova Üniversitesi
Ankara Üniversitesi
Mimar Sinan Güzel Sanatlar Üniversitesi
TOBB Ekonomi ve Teknoloji Üniversitesi
Galatasaray Üniversitesi
Çukurova Üniversitesi
Orta Doğu Teknik Üniversitesi
Başkent Üniversitesi
Koç Üniversitesi
Yaşar Üniversitesi
Işık Üniversitesi
Orta Doğu Teknik Üniversitesi
İstanbul Üniversitesi
Aksaray Üniversitesi
Uludağ Üniversitesi
Gazi Üniversitesi
Ege Üniversitesi
Bilkent Üniversitesi
Gazi Üniversitesi
Kadir Has Üniversitesi
Ufuk Üniversitesi
Orta Doğu Teknik Üniversitesi
Gazi Üniversitesi
SESRTCIC
Bahçeşehir Üniversitesi
Hacettepe Üniversitesi
Ankara Strateji Enstitüsü
Marmara Üniversitesi
Selçuk Üniversitesi
Ankara Üniversitesi