

# R Uygulamalı Yeniden Örnekleme Teknikleri

---

Zeynel CEBECİ



Prof. Dr. Zeynel CEBECİ

## R UYGULAMALI YENİDEN ÖRNEKLEME TEKNİKLERİ

ISBN 978-625-7052-60-3

DOI 10.14527/9786257052603

Kitap içeriğinin tüm sorumluluğu yazarına aittir.

© 2020, PEGEM AKADEMİ

Bu kitabın basım, yayım ve satış hakları Pegem Akademi Yay. Eğt. Dan. Hizm. Tic. A.Ş.'ye aittir. Anılan kuruluşun izni alınmadan kitabın tümü ya da bölümleri, kapak tasarımı; mekanik, elektronik, fotokopi, manyetik kayıt ya da başka yöntemlerle çoğaltılamaz, basılamaz ve dağıtılamaz. Bu kitap, T.C. Kültür ve Turizm Bakanlığı bandrolü ile satılmaktadır. Okuyucularımızın bandrolü olmayan kitaplar hakkında yayınevimize bilgi vermesini ve bandrolsüz yayınları satın almamasını diliyoruz.

Pegem Akademi Yayıncılık, 1998 yılından bugüne uluslararası düzeyde düzenli faaliyet yürüten **uluslararası akademik bir yayınev**idir. Yayımladığı kitaplar; Yükseköğretim Kurulunca tanınan yükseköğretim kurumlarının kataloglarında yer almaktadır. Dünyadaki en büyük çevrimiçi kamu erişim kataloğu olan **WorldCat** ve ayrıca Türkiye'de kurulan **Turcademy.com** tarafından yayınları taranmaktadır, indekslenmektedir. Aynı alanda farklı yazarlara ait 1000'in üzerinde yayını bulunmaktadır. Pegem Akademi Yayınları ile ilgili detaylı bilgilere <http://pegem.net> adresinden ulaşılabilir.

I. Baskı: Ağustos 2020, Ankara

Yayın-Proje: Şehriban Türüldür  
Dizgi-Grafik Tasarım: Müge Çetin  
Kapak Tasarımı: Pegem Akademi

Baskı: Ay-bay Kırtasiye İnşaat Gıda Pazarlama ve Ticaret Ltd. Şti.  
Çetin Emeç Bulvarı 1314. Cadde No: 37A-B Çankaya/ANKARA  
Tel: (0312) 472 58 55

Yayıncı Sertifika No: 36306  
Matbaa Sertifika No: 46661

### İletişim

Karanfil 2 Sokak No: 45 Kızılay/ANKARA  
Yayınevi: 0312 430 67 50 - 430 67 51  
Dağıtım: 0312 434 54 24 - 434 54 08  
Hazırlık Kursları: 0312 419 05 60  
İnternet: [www.pegem.net](http://www.pegem.net)  
E-ileti: [pegem@pegem.net](mailto:pegem@pegem.net)  
WhatsApp Hattı: 0538 594 92 40

## *Akademik hayatımızda büyük emekleri bulunan saygıdeğer hocam*

*Prof. Dr. Yüksel Bek'e*



Prof. Dr. Yüksel Bek, 1949 yılında Alaca (Çorum)'da doğdu. Atatürk Üniversitesi Ziraat Fakültesi'nden 1969 yılında mezun oldu. İki yıl asistanlık yaptıktan sonra, 1971 yılında devlet bursu ile lisansüstü öğrenimi için İngiltere'ye gitti. University of Reading, Applied Statistics (Department of Mathematics and Statistics) bölümünde lisans tamamlama eğitimini başarı ile bitirerek University of Reading, Applied Statistics Department'ta "A Comparison of Three Randomized Response Models for Quantitative Data" adındaki teziyle 1974 yılında yüksek lisans öğrenimini tamamladı. Tezi kendi alanında saygın istatistik dergilerinden JASA'da yayınlandı. Atatürk Üniversitesi Ziraat Fakültesi Zootehni Bölümü'ne dönerek o zamanki ismi ile İstatistik ve Panlama bilim dalında göreve başladı.

Doktora öğrenimini 1974-1976 yılları arasında "Faktör Analizi ile İlgili Bazı Problemler ve Psikiyatride Bir Uygulama" başlıklı tezi ile tamamladı. 1976-1981 yılları arasında Atatürk Üniversitesi Ziraat Fakültesi'nde Dr. Asistan olarak görev yaptı. 1981 yılında iklim verilerinin stokastik process ile modellenmesi konusunda hazırladığı doçentlik tezi ile Fen Bilimleri, Uygulamalı İstatistik (Biyostatistik) bilim dalında doçent oldu. YÖK yasası ile birlikte yeniden yapılanma sonrasında 1982 yılında Çukurova Üniversitesi Ziraat Fakültesi (ÇÜZF), Zootehni Bölümü, Biyometri ve Genetik Anabilim Dalı'nda doçent olarak göreve başladı. Aynı anabilim dalında 1989 yılında profesör oldu. 2000 yılında Ondokuz Mayıs Üniversitesi (OMÜ) Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı'nın kuruluşu ile birlikte Anabilim Dalı Başkanı olarak göreve başladı. Bu görevini 2016 yılına kadar sürdürerek aynı emekliye ayrıldı.

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



## Prof. Dr. Zeynel CEBECİ

Prof. Dr. Zeynel Cebeci, 1960 yılında Giresun'da doğdu. Lisans öğrenimini 1983 yılında Çukurova Üniversitesi Ziraat Fakültesi'nde; yüksek lisans ve doktora öğrenimlerini Çukurova Üniversitesi Fen Bilimleri Enstitüsü'nde sırasıyla 1985 ve 1990 yıllarında tamamladı. En İyi Doğrusal Yansız Tahmin için bilişim algoritmaları geliştirme konulu doktora çalışmasının bir bölümünü yapmak üzere 1987-1988 yıllarında Hohenheim Üniversitesi Haustiergenetik Bölümü'nde (Stuttgart, Almanya) araştırma çalışmalarında bulunmuş, ders ve kurslara katılmış; büyük boyutlu matrislerin tersinin alınması için Fortran dilinde dolaylı yöntemler üzerinde çalışmalar yürütmüştür.

1983-1985 yıllarında muhtelif özel sektör kuruluşlarında bilgisayar programcısı ve yönetici olarak çalışmıştır. 1986-1992 yıllarında Çukurova Üniversitesi Ziraat Fakültesi'nde araştırma görevlisi; 1992 yılı sonundan itibaren aynı bölümde yardımcı doçent olarak görev yapmıştır. 1993'te doçentliğe, 1999 yılı sonunda Çukurova Üniversitesi Ziraat Fakültesi Biyometri ve Genetik Anabilim Dalı'nda profesörlüğe atanmıştır. Halen aynı anabilim dalında öğretim üyesi olarak çalışmaktadır. Muhtelif yıllarda Çukurova Üniversitesi Bilgisayar Bilimleri Uygulama ve Araştırma Merkezi müdürü, Enformatik Bölüm başkanı ve Uzaktan Eğitim Uygulama ve Araştırma merkezi müdürü olarak görev yapmıştır.

İstatistik ve tarımsal bilişim teknolojileri konulu lisans dersleri; Python, R ile istatistik analiz, biyoinformatik ve mikrodizi analizi konulu çeşitli lisansüstü dersleri vermektedir. Metadata, semantik, multilingual web, bulanık ve olabilirlikli kümeleme analizi, küme başlatma algoritmaları ve tarımda yapay sinir ağları konusunda bilimsel araştırmaları bulunmakta olup CRAN'de dağıtılan *kpeaks*, *inaparc* ve *ppclust* ile GitHub'ta yayınlanan *fcvalid* ve *VatAna* R paketlerinin geliştiricisidir. Avrupa Birliği tarafından ICT-PSP kapsamında desteklenen *Organic.Lingua* çokdilli web projesi son yıllarda katıldığı uluslararası projeler arasında yer almaktadır. *feedTrace* gıda izlenebilirliği sistemi, *Traglor: Türkiye Tarımsal Öğrenme Kaynakları Deposu* tarımsal bilişim alanında yapmış olduğu önemli projeler arasında bulunmaktadır. Hobi olarak doğada gezinmekte, kelebek ve kuş gözlemciliği ve makro fotoğrafçılığı yapmaktadır. *AdaMerOs: Türkiye Kelebek Gözlemcileri Topluluğu* (*adameros.org*) kurucusu ve sistem geliştiricisidir.



## ÖN SÖZ

Hemen her araştırma alanında bilgi keşfi için istatistik yöntemler ve veri bilimi teknikleri kullanılarak veri analizi yapılır. Günümüzde bilgi ve iletişim teknolojisindeki ilerlemeler sayesinde bilgisayar ve istatistik yazılımların kullanılması veri analizini oldukça kolaylaştırmış ve yaygınlaştırmıştır. Ancak önemli bir sorun hala karşımızda durmaktadır. Yazılımların çoğu parametrik istatistik ve daha az da olsa parametrik olmayan istatistik yöntemleri kapsamaktadır. Diğer yandan parametrik yöntemlerle analiz sonuçlarının doğru ve güvenilir olması için bir dizi varsayımın geçerli olup olmadığının test edilmesi de gerekmektedir. Bu, araştırmacı veya veri analizcisinin bu varsayımlar hakkında bilgili olması, gerekli kontrolleri yapması ve verisine uygun analiz yöntemi ve testi seçmesi anlamına gelmektedir. Dahası bu gereksinimlerin çoğunun sağlandığı hallerde bile bazı varsayımların geçerli olmadığı birçok durumla karşılaşmak bir sürpriz değildir. Böylesi sorunların da bir çaresi olmalıdır. Varsayımlara dayanmayan yeniden örneklemeli istatistik bu çarelerden biridir. Yeniden örnekleme hesaplama yoğunluklu tekniklerden oluşmaktadır. Ancak sevindiricidir ki günümüzde bilgisayarların işlem hızları ve bellek kapasitelerinin artışı hesaplama güçlüklerini önemli düzeyde sorun olmaktan çıkarmıştır. Böylece yeniden örneklemeyle dayanan istatistik yöntemleri kullanma olanağı doğmuş ve önümüzdeki yıllarda yaygınlaşacağı beklenmektedir.

Yeniden örneklemeyle dayanan bu yeni istatistik analiz döneminde klasik parametrik yöntemlerin dayandığı varsayımlar ve matematiksel temelleri anlamaya çalışmak yerine araştırma sorusu ve sonuçlar üzerine daha fazla yoğunlaşarak daha verimli çalışma olanağı da sağlanabilecektir. Bu, bir araştırmacının karmaşık formül ve tabloları değil analiz sonuçlarını anlamaya daha fazla zaman ayırması demektir. Yeniden örnekleme teknikleri işte bu nedenle istatistiği daha kolay öğrenilen bir araç haline getirmeyi vaat eden alternatif yeni bir yol sunmuş gibi gözükmektedir.

Yeniden örneklemeyle dayanan veri analizinde araştırma hipotezlerini daha az veri ile deneysel olarak test edebilmek mümkün olabilmektedir. Parametrik yöntemler ortalama, parametrik olmayan yöntemler ortancalara dayalı karşılaştırmalara, ancak her halde makul ölçüde büyük örneklere ihtiyaç duyarken yeniden örnekleme tekniklerinden biri olarak bootstrap küçük veriyi sanal olarak büyütürken bu soruna çözüm sunulabilmektedir. Bootstrap ile güven aralıkları hesaplamak için herhangi bir formüle ihtiyaç duyulmaması da önemli bir avantajdır. Ancak daha da önemlisi istenilen herhangi bir istatistik için formüle gereksinim olmaksızın güven aralığı hesaplanmasının mümkün olmasıdır. Parametrik ve parametrik olmayan istatistikte örneğin çeyreklik ölçüleri için güven aralığı hesaplanamazken bootstrap ile bunu yapabilmek mümkündür.

Bu kitapta bootstrap ve permütasyon testleri başta olmak üzere yeniden örneklemeyle dayalı istatistik analiz teknikleri hakkında temel bilgiler verilmekte, performans karşılaştırmaları yapılmakta ve çok sayıda örnek ile uygulamalı ola-

rak açıklanmaktadır. Kitapta ayrıca entropi ve karşılıklı bilgi tabanlı bilgi teorisi hakkında giriş düzeyinde bilgi verilmektedir. Tek, iki ve çok değişkenli veri benzetiminde Monte Carlo teknikleri yanında bir araştırmacının ister doğrusal ister doğrusal olmayan modellere dayanan verilerini analiz etmesi için gerekli bilgi ve uygulamalar da sunulmaktadır. Örneğin çok az kaynakta incelenen uzaklık ve bilgi teorisi tabanlı doğrusal olmayan korelasyon ve regresyon analizi yöntemleri kitapta geniş şekilde incelenmiştir. Henüz yeterli düzeyde tanınmayan ancak esnek modelleme olanağı sunan GAM analizi ve splaynlar da kitapta ayrıntılı şekilde anlatılmaktadır.

Kitaptaki uygulamalar R ortamında gerçekleştirilmiş olup okuyucuların veri bilimi ve grafik analizin lideri durumunda olan özgür ve açık kaynak R yazılımı ile tanışmaları ve çalışmalarını amaçlanmıştır. R, hemen her bilim dalı ve endüstriyel uygulama için özel veya genel amaçlı ondokuzbin civarında paketi ile yeniden örnekleme istatistikte de lider durumdadır. R'de sunulan bootstrap işlemlerinin birçoğu henüz diğer istatistik yazılımlarında bulunmamaktadır. Bu önemli avantaj, R'nin paketler ekosistemi ve kendi programlama diliyle yeni yöntem ve algoritmaların kodlanması ve uygulanması için kolay, esnek ve hızlı kodlama olanağı sunmasından ileri gelmektedir. Sonuç olarak bu kitap, çok sayıda uygulamayla yeniden örnekleme istatistik analizi kısa yoldan kapsamlı şekilde açıklayan bir kitap olarak yazılmıştır. Bir başka deyişle bu kitap veri biliminde hızlı analizi en kestirme yoldan eşsiz bir yazılım ekosistemi olan R'nin gücü ile keşfetmenizi sağlayacak bir öğrenme ve başvuru kaynağıdır.

Yararlı olması dileğiyle,

Prof. Dr. Zeynel Cebeci

Çukurova Üniversitesi, Ziraat Fakültesi

Biyometri ve Genetik Anabilim Dalı

Balcalı, Sarıçam / Adana

zebeci@cukurova.edu.tr / cebeciz@gmail.com



## İÇİNDEKİLER

|                                    |     |
|------------------------------------|-----|
| Öz Geçmiş .....                    | v   |
| Ön Söz.....                        | vii |
| Kitaptaki Örneklerle Çalışma ..... | xv  |

### 1. BÖLÜM GİRİŞ

|   |   |
|---|---|
| 1.1. Klasik Parametrik Testler ve Yeniden Örnekleme ..... | 4 |
| 1.1.1. Dağılım Varsayımı .....                            | 4 |
| 1.1.2. Rastlantısal Örnekleme Varsayımı .....             | 4 |
| 1.1.3. Örneklem Büyüklüğü .....                           | 5 |
| 1.1.4. Doğruluk ve Doğrulama .....                        | 6 |
| 1.1.5. Hesaplama Güçlükleri .....                         | 6 |
| 1.1.6. Anlaşılabilirlik .....                             | 6 |
| 1.2. Yeniden Örnekleme Tekniklerinin Geleceği .....       | 7 |

### 2. BÖLÜM R'İN KURULMASI VE R'DE ÇALIŞMA

|   |    |
|---|----|
| 2.1. R Nedir? .....                                   | 9  |
| 2.2. R'nin İndirilmesi ve Kurulması .....             | 10 |
| 2.3. R'nin Güncellenmesi .....                        | 12 |
| 2.4. R'de Çalışma .....                               | 12 |
| 2.5. R'de Veri Türleri .....                          | 16 |
| 2.5.1. Atomik Veriler .....                           | 17 |
| 2.5.2. Bileşik Veriler .....                          | 19 |
| 2.5.3. Faktörler .....                                | 32 |
| 2.5.4. Tablolar .....                                 | 34 |
| 2.6. R'de İşlemler .....                              | 35 |
| 2.6.1. Atama İşlemleri .....                          | 35 |
| 2.6.2. Aritmetik İşlemler .....                       | 36 |
| 2.6.3. Karşılaştırma İşlemleri .....                  | 36 |
| 2.6.4. Mantıksal Bağlaçlar .....                      | 36 |
| 2.6.5. Diğer İşlemler .....                           | 37 |
| 2.7. R'de Fonksiyonlar .....                          | 37 |
| 2.7.1. Temel Fonksiyonlar .....                       | 38 |
| 2.7.2. Fonksiyon Tanımlama .....                      | 39 |
| 2.8. R'de Program Akışı, Kontroller ve Döngüler ..... | 41 |
| 2.8.1. Akış Kontrolü .....                            | 41 |

|  |    |
|--|----|
| 2.8.2. Döngüler .....  | 44 |
| 2.9. R Paketleri .....   | 45 |
| 2.9.1. Paketlerin İndirilmesi ve Kurulması .....                     | 46 |
| 2.9.2. Paketlerin Çalışma Alanına Yüklenmesi ve Çalıştırılması ..... | 49 |
| 2.10. R'de Dosya Okuma ve Yazma .....                                | 51 |
| 2.10.1. Veri Kümeleri .....  | 51 |
| 2.10.2. Veri Dosyaları.....  | 53 |
| 2.10.3. R'de Grafik Dosyaları.....                                   | 59 |
| 2.11. Alıştırmalar .....   | 61 |

### **3. BÖLÜM**

#### **İSTATİSTİK YÖNTEMLERE KISA BAKIŞ**

|  |    |
|--|----|
| 3.1. Açıklayıcı İstatistikler .....                        | 65 |
| 3.1.1. Açıklayıcı İstatistiklerin Hesaplanması.....        | 65 |
| 3.1.2. Açıklayıcı İstatistikler için Grafikler .....       | 69 |
| 3.2. Çıkarımsal İstatistik .....                           | 72 |
| 3.2.1. Ortalama İçin Nokta ve Güven Aralığı Kestirimi..... | 73 |
| 3.2.2. Ortalama İçin Nokta ve Güven Aralığı Hesaplama..... | 76 |
| 3.3. Parametrik ve Parametrik Olmayan İstatistik .....     | 77 |
| 3.3.1. Varsayımların Kontrol Edilmesi .....                | 78 |
| 3.3.2. R'de Parametrik Varsayımların Kontrolü.....         | 79 |
| 3.4. Yeniden Örnekleme İstatistik.....                     | 80 |
| 3.5. Alıştırmalar .....                                    | 85 |

### **4. BÖLÜM**

#### **TEK DEĞİŞKENLİ BOOTSTRAP**

|   |     |
|---|-----|
| 34.1. Bootstrap Yöntemleri.....                                       | 88  |
| 4.1.1. Parametrik Olmayan Bootstrap .....                             | 88  |
| 4.1.2. Parametrik Olmayan Bootstrap Uygulaması .....                  | 99  |
| 4.1.3. R'de Bootstrap Paketleri İle Çalışma.....                      | 127 |
| 4.1.4. Bootstrap Başarım Analizi.....                                 | 150 |
| 4.1.5. G&H Dağılımları ve Dayanıklı Kestiriciler İçin Bootstrap ..... | 169 |
| 4.2. Parametrik Bootstrap.....  | 184 |
| 4.3. Parametrik Bootstrap Uygulaması.....                             | 184 |
| 4.4. boot Paketi ile Parametrik Bootstrap .....                       | 193 |
| 4.5. Yarı Parametrik Bootstrap.....                                   | 194 |
| 4.6. Yarı Parametrik Bootstrap Uygulaması.....                        | 195 |
| 4.7. Alıştırmalar .....   | 198 |

## 5. BÖLÜM JACKKNİFE İLE YENİDEN ÖRNEKLEME

|  |     |
|--|-----|
| 5.1. Jackknife Yöntemi .....                               | 201 |
| 5.2. Jackknife Uygulaması .....                            | 203 |
| 5.3. Çoklu İstatistikler İçin Jackknife Kestirimleri ..... | 210 |
| 5.4. Alıştırmalar .....                                    | 213 |

## 6. BÖLÜM İKİ ÖRNEKLEMİN KARŞILAŞTIRILMASI

|   |     |
|---|-----|
| 6.1. Bağımsız Örneklerin Karşılaştırılması .....      | 215 |
| 6.1.1. Permütasyon Testi .....                        | 227 |
| 6.1.2. Bootstrap Uygulaması .....                     | 241 |
| 6.2. Eşleştirilmiş Örneklerin Karşılaştırılması ..... | 249 |
| 6.2.1. Permütasyon Testi .....                        | 251 |
| 6.2.2. Bootstrap Uygulaması .....                     | 255 |
| 6.3. Alıştırmalar .....                               | 259 |

## 7. BÖLÜM KORELASYON ANALİZİNDE BOOTSTRAP

|  |     |
|--|-----|
| 7.1. İki Değişkenli Veri Benzetimi .....                       | 261 |
| 7.1.1. Fonksiyonel İlişki Benzetimleri .....                   | 261 |
| 7.1.2. Korelasyonlu Veri Benzetimi .....                       | 274 |
| 7.1.3. HHG Paketi İle Veri Benzetimi .....                     | 284 |
| 7.1.4. mlbench Paketi İle Veri Benzetimi .....                 | 287 |
| 7.2. Korelasyon Hesaplama Yöntemleri .....                     | 288 |
| 7.2.1. Pearson Korelasyonu .....                               | 288 |
| 7.2.2. Spearman Korelasyonu .....                              | 292 |
| 7.2.3. Kendall tau Korelasyonu .....                           | 294 |
| 7.2.4. Korelasyon Analizindeki Sorunlar .....                  | 298 |
| 7.2.5. Uzaklık Korelasyonu .....                               | 304 |
| 7.2.6. Hoeffding'in D Ölçüsü .....                             | 310 |
| 7.2.7. Maksimum Bilgi Katsayısı .....                          | 313 |
| 7.2.8. Rastlantısal Bağımlılık Katsayısı .....                 | 338 |
| 7.2.9. Diğer Doğrusal Olmayan Yöntemler .....                  | 341 |
| 7.2.10. Korelasyon Yöntemlerinin Karşılaştırılması .....       | 346 |
| 7.2.11. Korelasyonlar Arasındaki İlişkilerin İncelenmesi ..... | 348 |
| 7.2.12. Korelasyon Matrisleri .....                            | 352 |

|   |     |
|---|-----|
| 7.3. Korelasyonlar İçin Bootstrap Uygulaması .....                    | 356 |
| 7.3.1. Bootstrap Sonuçlarının İncelenmesi ve Görselleştirilmesi ..... | 358 |
| 7.3.2. Güven Aralıklarının Hesaplanması .....                         | 360 |
| 7.3.3. Çok Sayıda Korelasyon için Bootstrap .....                     | 365 |
| 7.4. Korelasyon Yöntemlerinin Bootstrap Başarımları .....             | 369 |
| 7.5. Korelasyonların Önemlilik Testi .....                            | 378 |
| 7.5.1. Korelasyonlar İçin Permütasyon Testi .....                     | 379 |
| 7.5.2. Jackknife ile Korelasyon Analizi .....                         | 386 |
| 7.6. Alıştırmalar .....   | 390 |

## 8. BÖLÜM

### REGRESYON ANALİZİNDE BOOTSTRAP

|   |     |
|---|-----|
| 8.1. Doğrusal Regresyon Modelleri İçin Bootstrap .....      | 396 |
| 8.1.1. Basit Doğrusal Regresyon .....                       | 396 |
| 8.1.2. Çoklu Doğrusal Regresyon İçin Bootstrap .....        | 407 |
| 8.1.3. Kantil Regresyonu .....                              | 422 |
| 8.2. Doğrusal Olmayan Modeller İçin Bootstrap .....         | 428 |
| 8.2.1. Polinomiyal Regresyon .....                          | 431 |
| 8.2.2. Basamak Fonksiyonları .....                          | 433 |
| 8.2.3. Çekirdek Regresyonu .....                            | 435 |
| 8.2.4. Lokal Regresyon .....                                | 440 |
| 8.2.5. Regresyon Splaynları .....                           | 446 |
| 8.2.6. Süper Yumuşatıcılar .....                            | 453 |
| 8.2.7. Lojistik Regresyon .....                             | 455 |
| 8.2.8. Büyüme Modelleri .....                               | 461 |
| 8.2.9. Genelleştirilmiş Eklemeli Modeller .....             | 479 |
| 8.2.10. Çokdeğişkenli Uyarlanabilir Regresyon Splaynı ..... | 502 |
| 8.3. Alıştırmalar .....                                     | 515 |

## 9. BÖLÜM

### ÇAPRAZ DOĞRULAMA VE MODEL SEÇME

|   |     |
|---|-----|
| 9.1. Model Karşılaştırma Ölçütleri .....          | 520 |
| 9.2. Çapraz Doğrulama Teknikleri .....            | 528 |
| 9.2.1. Bölünmüş Örnekleme Çapraz Doğrulama .....  | 528 |
| 9.2.2. Monte Carlo Çapraz Doğrulaması .....       | 535 |
| 9.2.3. K-katlı Çapraz Doğrulama .....             | 540 |
| 9.2.4. Tekrarlanan K-katlı Çapraz Doğrulama ..... | 547 |
| 9.2.5. Bootstrap İle Doğrulama .....              | 551 |

|   |     |
|---|-----|
| 9.3. Model Karşılaştırma ve Değişken Seçme .....        | 555 |
| 9.3.1. Çapraz Doğrulama İle Model Seçme .....           | 556 |
| 9.3.2. Rde rminer Paketi İle Model Seçme.....           | 559 |
| 9.3.3. Rde caret ve caretEnsemble ile Model Seçme ..... | 563 |
| 9.4. Alıştırmalar .....                                 | 575 |

## 10. BÖLÜM

### VARYANS ANALİZİNDE YENİDEN ÖRNEKLEME

|  |     |
|--|-----|
| 10.1. Parametrik Tek Yönlü ANOVA .....                         | 579 |
| 10.1.1. Tek Yönlü Varyans Analizi Uygulaması.....              | 583 |
| 10.1.2. Varyans Analizi İle İlgili Varsayımların Kontrolü..... | 588 |
| 10.2. ANOVA İçin Bootstrap İşlemi .....                        | 593 |
| 10.2.1. Model Tabanlı Bootstrap .....                          | 593 |
| 10.2.2. Wild Bootstrap .....                                   | 601 |
| 10.3. Doğrusal Modeller için Permütasyon Testi .....           | 607 |
| 10.4. İkili Karşılaştırma İçin Permütasyon Testi .....         | 610 |
| 10.5. Alıştırmalar .....                                       | 615 |
| Kaynaklar.....   | 617 |
| Ekler .....  | 625 |
| İngilizce-Türkçe Terim Karşılıkları Klavuzu .....              | 625 |
| Dizin.....   | 629 |

## TEŐEKKÜR

*Ondokuz Mayıs Üniversitesi Tıp Fakóltesi Biyoistatistik Anabilim Dalı başkanı iken emekliye ayrılan ve kendisine atfettiđim bu kitabın bilimsel edisyonunda katkı sađlayan deđerli hocam Sn. Prof. Dr. Yüksel Beke,*

*Bu kitap dâhil R ile istatistik ve veri bilimi serisi kitaplarımı yazma isteđimi güçlendiren Çukurova Üniversitesi Ziraat Fakóltesi öğretim üyelerinden, deđerli öğrencim ve meslektaşım Doç.Dr. Uđur Serbester'e,*

*Kitabın yazım sürecinde akademik hayatımın dışındaki birçok ödevimi özveriyle üstlenen sevgili eşim Nalan Cebeci'ye,*  
*Teşekkürlerimle...*

## Kitaptaki Örneklerle Çalışma

Kitaptaki R fonksiyonları ve örnekler aşağıda görüldüğü gibi gri zeminli kutu içinde Courier yazı tipinde verilmiştir.

```
# Fonksiyon 6.7: Eşleştirilmiş örneklerde ortalamaların farkı
> calc.meandif <- function(x, indices){
+   n <- nrow(x)
+   if(missing(indices)) indices <- 1:n
+   x <- x[indices,]
+   delta <- mean(x[c(1:(n/2)),1]-x[-c(1:(n/2)),1])
+   names(delta) <- "dif.Ort"
+   return(delta)
+ }
```

```
# Örnek 4.20a: boot paketi ile çoklu istatistikler için bootstrap
# Bağımlılık - Örnek: 4.1a; Fonksiyon: 4.6, Paket: boot
# boot paketini kur, kurulu ise çalışma alanına yükle
> if(!require(boot, quietly=TRUE)) {
+   install.packages("boot", repo="https://cloud.r-project.org/")
+   require(boot, quietly=TRUE)}
# boot ile x verisinde çoklu istatistikler için bootstrap
> bootmult <- boot(data=x, statistic=calc.meandif,
+   sim="ordinary", R=2000)
> bootmult$t0[1]
[1] 4.34542
```

Kitapta yukarıda örnekleri sunulan fonksiyon tanımlamaları ve örneklerde:

- Satır başlarındaki > simgesi R tarafından otomatik olarak görüntülenen komut satırı göstergesidir. Örnekler elle yazılarak giriliyorsa yazılmaması gerekir.
- Satır başlarındaki + simgesi R tarafından otomatik olarak görüntülenen komut satırı devam işaretidir. Komutlar tek bir satıra sığmadığında veya kullanıcı komutu tamamlamadan <Enter> tuşuna bastığında R izleyen satıra geçer ve satırın başına bunun öncekinin devamı olduğunu gösteren + simgesi koyar. Kitaptaki örnekler elle giriliyorsa bu karakterin yazılmaması gerekir.
- Komutlar çalıştırdıktan sonra sonuçları komuttan sonraki satırlarda görüntülenir. Bazı sonuçlar [1], [8] gibi köşeli parantezler içinde sayılarla başlayan şekilde gösterilir. Bu sayılar hesaplama sonucu bir vektör