

UYGULAMALI ÇOK DEĞİŞKENLİ İSTATİSTİKSEL ANALİZ

Prof. Dr. Hüseyin TATLIDİL
Hacettepe Üniversitesi
Fen Fakültesi İstatistik Bölümü
Öğretim Üyesi

ANKARA

İÇİNDEKİLER

Sayfa

GÖSTERİM.....	IX
1. GİRİŞ.....	1
1.1. Çok Değişkenli İstatistiksel Analiz ve Kullanım Alanları	1
1.2. Çok Değişkenli İstatistiksel Analizde Kullanılan Matris Kuramı	3
1.2.1. Gösterim ve Tanımlar	3
1.2.2. Örüntü ya da Model Matrisleri	4
1.2.3. Determinantlar	7
1.2.4. Matris Tersi	9
1.2.5. Genelleştirilmiş Matris Tersi	11
1.2.6. Matris Parçalanması	11
1.2.7. Matris Rankı	13
1.2.8. Özdeğerler ve Özvektörler	14
1.2.9. Bir Matrisin İzi	15
1.2.10. Doğrudan Çarpım	15
1.2.11. Karesel Formlar ve Tanımlılık	17
1.2.12. Matris Faktörleştirilmesi	18
1.2.13. İdempotent Matrisler	19
1.2.14. Türev ve İntegral	20
1.2.15. Matris Dönüşümlerinin Jakobiyenleri	22
Çözülmüş Örnekler	24
Sorular	29
2. SÜREKLİ ÇOK DEĞİŞKENLİ DAĞILIMLAR.....	31
2.1. Çok Değişkenli İstatistiksel Analizde Kullanılan Matris ve Vektörler	31

2.2.	Çok Değişkenli İstatistiksel Analizde Kullanılan	
	Özetleyici Bilgiler	32
2.3.	Çok Değişkenli Dağılımlar.....	37
	Çözülmüş Örnekler	44
	Sorular	50
3.	ÇOK DEĞİŞKENLİ NORMAL DAĞILIM (Ç.D.N.D.).....	53
3.1.	Karakteristik Fonksiyonlar ve Momentler	57
3.2.	Ç.D.N.D.'ın Moment Türeten Fonksiyonu ve	
	Parametreleri	60
3.3.	Kitle Parametrelerinin En Çok Olabilirlik Kestiricileri	63
3.4.	Marjinal Normal Dağılım	65
3.5.	Koşullu Normal Dağılım	70
3.6.	Çoklu Korelasyon	76
3.7.	Normal Dağılıma Sahip Değişkenlerin Doğrusal	
	Bağıntılarının Dağılımı.....	79
3.8.	Alt Rastlantı Vektör Değişkenlerinin Bağımsızlığı.....	83
3.9.	Yoğunluk Fonksiyonu Verilmiş İken Parametrelerin	
	Bulunması.....	85
3.10.	Karesel Formların Dağılımı.....	87
	Çözülmüş Örnekler	90
	Sorular	101
4.	ÇOK DEĞİŞKENLİ HİPOTEZ TESTLERİ.....	105
4.1.	Giriş.....	105
4.2.	Çok Değişkenli Test Yöntemleri.....	107
4.3.	Tek Örneklem İlişkin Testler	108
4.4.	İki Örneklem İlişkin Testler	117
4.5.	Çok Örneklem İlişkin Testler	120
4.6.	Behrens-Fisher Problemi	126

	<u>Sayfa</u>
Çözülmüş Örnekler	128
Sorular	134
5. TEMEL BİLEŞENLER ANALİZİ.....	138
5.1. Temel Bileşenlerin Elde Edilmesi	139
5.2. Temel Bileşenlerin Özellikleri ve Sağladığı Yararlar	144
5.3. Temel Bileşenler Analizinin Gerekliliği.....	145
5.4. Temel Bileşen Sayısının Belirlenmesi	146
Çözülmüş Örnekler	155
Sorular	162
6. FAKTÖR ANALİZİ	167
6.1. Faktör Analizinin Amacı.....	167
6.1.1. Faktör Analizi ile Temel Bileşenler Analizi Arasındaki Benzerlikler	171
6.1.2. Faktör Döndürmesi ve Kavramsal Anlamlılık	173
6.2. Döndürme Türleri	178
6.2.1. Dik Döndürme Türleri	179
6.2.2. Eğik Döndürme Türleri	181
6.3. Faktör Bulma Yöntemleri.....	185
6.3.1. Çoklu Gruplandırma Yöntemi.....	185
6.3.2. Ardışık Çoklu Gruplandırma Yöntemi	191
Çözülmüş Örnekler	201
Sorular	213
7. KANONİK KORELASYON ANALİZİ.....	216
7.1. İlişki Kavramı ve Kanonik Korelasyon Analizinin Amacı.....	216
7.2. Kanonik Değişkenler, Kanonik Korelasyonların Tanımı ve Elde Edilmesi.....	218
7.3. Kanonik Değişkenlerle Orijinal Değişkenler Arasındaki Korelasyonlar ve Yorumları	223

7.4.	Kanonik Korelasyon Katsayılarının Önem Kontrolleri	225
7.4.1.	Bartlett Testi.....	225
7.4.2.	Roy'un En Büyük Özdeğer Yaklaşımı.....	227
7.5.	Kısmi Kanonik Korelasyonlar ve Yorumları	228
	Çözülmüş Örnekler	234
	Sorular	252
8.	DİSKRİMİNANT (AYIRMA) ANALİZİ	256
8.1.	Giriş.....	256
8.2.	İki Grup Olması Durumunda Diskriminant Analizi.....	258
8.2.1.	Hatalı Sınıflandırma Olasılıklarının Bulunması.....	260
8.2.2.	Diskriminant Fonksiyonunun Önem Kontrolü	261
8.2.3.	Bireylerin Sınıflara Ayrılmasına İlişkin Olasılıkların Tahmini	262
8.3.	İkiden Çok Grup Olması Halinde Diskriminant Analizi	265
8.3.1.	Diskriminant Fonksiyonunun Önem Kontrolü	268
8.3.2.	Normal Dağılımlı Kitlelerde Diskriminant Analizi.....	269
8.4.	Özel Durumlarda Kullanılan Diskriminant Fonksiyonları.....	273
	Çözülmüş Örnekler	279
	Sorular	285
9.	LOJİSTİK REGRESYON ANALİZİ.....	289
9.1.	Giriş.....	289
9.2.	Doğrusal Olasılık Modeli	289
9.3.	İki Grup Lojistik Modeller	291
9.3.1.	Lojistik Ayrımsama.....	294
9.3.2.	İki Grup Lojistik Modelde Kestirim Yöntemleri	295
9.3.3.	Uyum İyiiliği ve Sapma Ölçütleri	297

9.4.	Çoklu Grup Lojistik Modeller	304
9.4.1.	Çoklu Grup Lojistik Modellerde Kestirim Yöntemleri	307
9.4.2.	Uyum İyiliği Testleri, Değişken Seçimi ve Temel Sınıfın Belirlenmesi	311
9.5.	Lojistik Regresyon ile Diskriminant Analizinin Karşılaştırılması	313
	Çözülmüş Örnekler	315
	Sorular	324
10.	KÜMELEME ANALİZİ	329
10.1.	Giriş	329
10.2.	Kümelemede Kullanılan Benzerlik ve Uzaklık Ölçütleri	330
10.3.	Kümeleme Yöntemleri	334
10.3.1.	Hiyerarşik Kümeleme Yöntemleri	334
10.3.2.	Hiyerarşik Olmayan Kümeleme Yöntemleri	338
10.3.3.	Hacim İlişkisine Dayalı Kümeleme Yöntemi	339
10.4.	Küme Sayısının Belirlenmesi	341
	Çözülmüş Örnekler	343
	Sorular	350
11.	ÇOK BOYUTLU ÖLÇEKLEME (Ç.B.Ö.)	353
11.1.	Giriş	353
11.2.	Ç.B.Ö. Yöntemleri	354
11.2.1.	Metrik Ölçekleme Yöntemi	355
11.2.2.	Metrik Olmayan Ölçekleme Yöntemi	361
11.3.	Ç.B.Ö. Yöntemlerinin Karşılaştırılması	366
11.4.	Ç.B.Ö. ile Temel Bileşenlerin Karşılaştırılması	367
	Çözülmüş Örnekler	368
	Sorular	378

12. ÇOK DEĞİŞKENLİ REGRESYON ANALİZİ	381
12.1. Giriş.....	381
12.2. Parametrelerin En Çok Olabilirlik Kestiricileri.....	382
12.3. Hipotez Testleri	386
12.4. Korelasyon Katsayıları	393
12.4.1. Çoklu Korelasyon Katsayısı	394
12.4.2. Kısmi Korelasyon Katsayısı	395
12.4.3. Vektörler Arası Korelasyon	396
12.5. Model Uygunluğunun Testi	396
Çözülmüş Örnekler	399
Sorular	403
EK-A İSTATİSTİK TABLOLARI	408
Tablo 1. Standart Normal Dağılım Tablosu	409
Tablo 2. t Dağılımı Tablosu	410
Tablo 3. χ^2 Dağılımı Tablosu	411
Tablo 4. F Dağılımı Tablosu	412
Tablo 5. Heck Grafikleri.....	413
Tablo 6. Beta Dağılımı Tablosu	417
EK-B İNGİLİZCE-TÜRKÇE İSTATİSTİK TERİMLERİ LİSTESİ	419
KAYNAKÇA	421
DİZİN	423

GÖSTERİM

Diğer bilim dallarında olduğu gibi, istatistik konusunda yayınlanan kitapların bir çoğunda da farklı gösterim biçimleriyle karşılaşılmalıdır.

Gösterim yönünden genel bir kuralın benimsenmesi, kaynağın izlenmesini kolaylaştıracağı gibi, yanlış anlamayı da ortadan kaldıracaktır. Bu nedenle, kitapta özel olarak tanımlanmamışsa küçük harfler (karakter) gözlemleri ya da değerleri, koyu (bold) büyük harfler ise matrisleri göstermede kullanılmıştır. Bazı bölümlerde θ , Λ , α , γ gibi sembollerin teknik olanaksızlıklar nedeniyle bold basılamaması dışında vektörler için (gerekli açıklamalar yapılarak) indisli büyük ya da küçük harfler koyu olarak kullanılmıştır. Kitapta özellikle belirtilmemişse vektör denince sütun vektöründen söz edilmektedir ve \sim sembolü dağılım anlamında kullanılmaktadır.

Diğer istatistik kitaplarında olduğu gibi, μ , Σ gibi semboller kitle parametreleri \bar{x} , S gibi semboller ise örneklem değerleri için kullanılmıştır.