T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KANONİK KORELASYON VE FAKTÖR ANALİZİNDE BOYUT SEÇİMİ PROBLEMİNE PARALEL ANALİZ YAKLAŞIMLARI

SELAHATTİN AYDOĞDU

DOKTORA TEZİ
İSTATİSTİK ANABİLİM DALI
İSTATİSTİK PROGRAMI

DANIŞMAN
PROF. DR. GÜLHAYAT GÖLBAŞI ŞİMŞEK

İSTANBUL, 2019

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KANONİK KORELASYON VE FAKTÖR ANALİZİNDE BOYUT SEÇİMİ PROBLEMİNE PARALEL ANALİZ YAKLAŞIMLARI

Selahattin AYDOĞDU tarafından hazırlanan tez çalışması 17.05.2019 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı'nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Tez Danışmanı
Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK
Yıldız Teknik Üniversitesi

Jüri Üyeleri
Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK
Yıldız Teknik Üniversitesi _____

Prof. Dr. Fatma NOYAN TEKELİ
Yıldız Teknik Üniversitesi _____

Prof. Dr. İlknur Esen YILDIRIM
Marmara Üniversitesi _____

Doç. Dr. İbrahim DEMİR
Yıldız TeknikÜniversitesi _____

Doç. Dr. Ozan KOCADAĞLI
Mimar Sinan Güzel Sanatlar Üniversitesi _____

ÖNSÖZ

Çalışmalarım boyunca her türlü desteği sağlayan, beni yönlendiren, akademik hayatımda çok iyi bir rehber olan, çalışkanlığı ve azmini örnek aldığım değerli hocam Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK'e;
Çalışmamın şekillenmesinde değerli fikirleri ile beni yönlendiren hocalarım Prof. Dr. Fatma NOYAN TEKELİ ve Prof. Dr. İlknur Esen YILDIRIM'a;
Tez hazırlama sürecinde varlılığı ve destekleriyle her zaman yanımda olan kıymetli eşim Rukiye Nazan'a ve biricik kızım Büşram'a;
Teşekkürlerimi sunarım.

Mayıs, 2019

Selahattin AYDOĞDU

İÇİNDEKİLER

Sayfa
SİMGE LİSTESİ vi
KISALTMA LİSTESİ vii
ÇİZELGE LİSTESİ viii
ÖZET ix
ABSTRACT x
BÖLÜM 1
GİRİŞ 1
1.1 Literatür Özeti 1
1.2 Tezin Amacı 3
1.3 Orjinal Katkı 4
BÖLÜM 2
KANONİK KORELASYON ANALİZİNDE BOYUT SAYISININ BELİRLENMESİNDE PARALEL ANALİZ ve BOOTSTRAP MODİFİYELİ PARALEL ANALİZ YAKLAŞIMLARI 6
2.1 Kanonik Korelasyon Analizinde Boyut Sayısı Belirlenmesinde Kullanılan İstatistiksel Test Prosedürleri 8
2.2 Paralel Analizin İstatistiksel Temelleri 11
2.3 Paralel Analizin Kanonik Korelasyon Analizine Adaptasyonu 13
2.4 Paralel Analizin Bootstrap Modifikasyonu 13
2.5 Kanonik Korelasyon Analizinde Gerçek Veri İçin Paralel Analiz Yöntemi 14
2.6 Simülasyon Çalışması 16
2.6 Simülasyon Sonuçları 20
BÖLÜM 3
FAKTÖR ANALİZİNDE BOYUT SAYISININ BELİRLENMESİNDE PARALEL ANALİZ ve BOOTSTRAP MODİFİYELİ PARALEL ANALİZ YAKLAŞIMLARI 30
3.1 Faktör Analizi Modeli 31
3.2 Boyut Sayısı Belirleme Yöntemleri 33
3.2.1 G-K Kuralı 33
3.2.2 Paralel Analiz 33
3.3 Bootstrap Metodu 34
3.3.1 Bootstrap Persentillerine Dayalı Güven Aralıkları 35
3.2.3 Paralel Analiz Yönteminin Modifikasyonları 36

3.2.3.1 Bootstrap Modifiyeli Paralel Analiz 36
3.4 Gerçek Veri ile Paralel Analiz ve Bootstrap Modifiyeli Paralel Analiz Yaklaşımların Uygulamaları 37
3.5 Simülasyon Çalışması 40
3.5.1 Faktör Modelleri 40
3.6 Simülasyon Sonuçları 42
BÖLÜM 4
FAKTÖR ANALİZİNDE BOYUT SAYISININ BELİRLENMESİNDE PARALEL ANALİZE DAYANAN AMPİRİK OLABİLİRLİK GÜVEN ARALIĞI YAKLAŞIMI 50
4.1 Ampirik Olabilirlik Yöntemi 50
4.1.1 İki Örneklem Ampirik Olabilirlik Yöntemi 51
4.2 Paralel Analize Dayanan İki Örneklem Ampirik Olabilirlik Güven Aralığı Yöntemi 52
4.3 Gerçek Veri ile Önerilen Yaklaşımın Uygulamaları 53
4.4 Varsayımsal Faktör Modellerinde Önerilen Yaklaşımın Uygulamaları 54
4.5 Varsayımsal Faktör Modellerinde Önerilen Yaklaşımın Performans Değerlendirilmesi 57
BÖLÜM 5
SONUÇ ve ÖNERİLER 61
KAYNAKLAR 63
EK-A
ÖZGEÇMİŞ 96

SİMGE LİSTESİ

a	Eşik değer
$\sum$	Varyans-kovaryans matrisi
μ	Ortalama
σ^2	Varyans
d	Boyut sayısı
N	Örneklem sayısı
p	Değişken sayısı
q	Değişken sayısı
X	Bağımsız rasgele değişken
λ	Özdeğer
f	Faktör
Λ	Faktör yükleri matrisi
m	Varsayımsal modellerin faktör sayısı
S	Kovaryans matrisi
R	Korelasyon matrisi
ψ	Kalıntı matrisi
α	Anlam seviyesi
ρ	Anakütle korelasyon katsayısı
r	Örnek korelasyon katsayısı
θ	Parametre
Φ	Faktörlerin korelasyon matrisi
se	Standart hata
χ^2	Ki-kare dağılımı
w	Tracy-Widom dağılımının parametresi
l_1	İlk özdeğer
T_d^2	Lawley-Hotelling iz istatistiği
Λ_d	Wilks'in Λ_d olabilirlik oran istatistiği
V_d	Bartlett-Nanda-Pillai iz istatistiği
$S_B(T_d^2)$	T_d^2'ye dayalı Bartlett χ^2 yaklaşım prosedürü
$S_F(T_d^2)$	T_d^2'ye dayalı Fujikoshi χ^2 yaklaşım prosedürü
$S_B(\Lambda_d)$	Λ_d'ye dayalı Bartlett χ^2 yaklaşım prosedürü
$S_L(\Lambda_d)$	Λ_d'ye dayalı Lawley χ^2 yaklaşım prosedürü
$S_B(V_d)$	V_d'ye dayalı Bartlett χ^2 yaklaşım prosedürü
$S_F(V_d)$	V_d'ye dayalı Fujikoshi χ^2 yaklaşım prosedürü

KISALTMA LİSTESİ

AFA	Açıklayıcı Faktör Analizi
ADF	Ampirik Dağılım fonksiyonu
AIC	Akaike Bilgi Kriteri
BPA	Bootstrap Modifiyeli Paralel Analiz
$BPA_{ortalama}$	BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin ortalama değeri
BPA_{medyan}	BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin medyan değeri
BPA_{99}	BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 99. persentil eşik değeri
$BPA_{97.5}$	BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 97.5 persentil eşik değeri
BPA_{95}	BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 95. persentil eşik değeri
BPA_{90}	BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 90. persentil eşik değeri
CI	Güven Aralığı
EL	Ampirik Olabilirlik
ELCI	İki Örneklem Ampirik Olabilirlik Güven Aralığı
G-K	Guttman-Kaiser Kuralı
KKA	Kanonik Korelasyon Analizi
L	Tracy -Widom istatistiği
LB	Bootstrap Özdeğerleri (LB1, birinci bootstrap özdeğeri).
LR	Paralel Analiz Özdeğerleri (LR1, birinci paralel analiz özdeğeri).
LB-LR	Bootstrap özdeğerleri ile paralel özdeğerlerin ampirik olabilirlik güven aralıkları
MRFA	Minimum Rank Faktör Analizi
PELCI	Paralel Analize Dayanan İki örneklem Ampirik Olabilirlik Güven Aralığı Yöntemi
PA	Paralel Analiz
$PA_{ortalama}$	PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin ortalama değeri
PA_{medyan}	PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin medyan değeri
PA_{99}	PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 99. persentil eşik değeri.
$PA_{97.5}$	PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 97.5 persentil eşik değeri
PA_{95}	PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 95. persentil eşik değeri
PA_{90}	PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 90. persentil eşik değeri
TBA	Temel Bileşenler Analizi
TEF	Temel Eksen Faktörleşme Yöntemi
TW	Tracy-Widom Testi
YEM	Yapısal Eşitlik Modelleri

ÇİZELGE LİSTESİ

Sayfa

Çizelge 2. 1 Kanonik korelasyonların karesi (özdeğerler), sıralı özdeğerler için bootstrap %95 güven aralığının alt sınırına dayalı persentil değeri ve PA'ya dayanan sıralı özdeğerelerin ortalama, medyan, 95. ve 99. persentil değerleri 15

Çizelge 2. 2 Gerçek veride boyut sayısı için klasik test istatistikleri ve p değeri 15

Çizelge 2. 3 Simülasyon çalışmasında kullanılan kanonik korelasyonlar..... 16

Çizelge 2.4 Klasik test istatistikleri $(\alpha = 0.05)$ anlamlılık seviyesinde, Paralel Analiz (PA) ile %95 Bootstrap Güven Aralığı ile Modifiye Paralel Analiz (BPA) yöntemine göre boyut sayısı belirleme için göreli frekanslar 21

Çizelge 2. 5 Klasik test istatistikleri $(\alpha = 0.05)$ anlamlılık seviyesinde, Paralel Analiz (PA) ile %95 Bootstrap Güven Aralığı ile Modifiye Paralel Analize (BPA) göre kanonik korelasyon değerinin 0.5'ten büyük ve 0.5'ten küçük-eşit olmasına göre boyut belirleme için göreli frekanslar 26

Çizelge 3. 1 Üç farklı gerçek veri seti için boyut sayısını belirlemede G-K kuralı, $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} persentil değerleri..... 39

Çizelge 3. 2 Model 1 Tek faktörlü model için G-K kuralı, $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} yöntemlerine göre boyut sayısı belirleme için göreli frekanslar 44

Çizelge 3. 3 Model 2 korelasyonsuz üç faktörlü model için G-K kuralı, $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} yöntemlerine göre boyut sayısı belirleme için göreli frekanslar 46

Çizelge 3. 4 Model 3 korelasyonlu üç faktörlü model için G-K kuralı, $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} göre boyut sayısı belirleme için göreli frekanslar48

Çizelge 4. 1 Gerçek veri setleri için için boyut sayısını belirlemede G-K kuralı, PA_{99}, PA_{95} ve PELCI yöntemleri ile belirlenen faktör sayıları..... 54

Çizelge 4. 2 Faktör modellerinde G-K kuralı, PA_{99}, PA_{95} ve PELCI yöntemleri ile boyut sayısı belirleme 59

ÖZET

KANONİK KORELASYON VE FAKTÖR ANALİZİNDE BOYUT SEÇİMİ PROBLEMİNE PARALEL ANALİZ YAKLAŞIMLARI

Selahattin AYDOĞDU
İstatistik Anabilim Dalı
Doktora Tezi

Tez Danışmanı: Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK

Paralel analiz (PA), faktör analizinde boyut sayısı tespitinde kullanılan amprik bir yöntemdir. Bu tezde kanonik korelasyon analizi (KKA) ve faktör analizi yöntemlerinde boyut sayısını belirlemek için PA'ya dayanan yaklaşımlar önerilmektedir.

İlk olarak PA yöntemi KKA'ya adapte edilerek, boyut sayısının belirlenmesindeki kullanılabilirliği araştırılmıştır. Ayrıca bu bölümde KKA'da boyut sayısı belirlemede bootstrap modifiyeli paralel analiz (BPA) geliştirilmiştir. KKA'da boyut sayısı belirlemede PA ve BPA yöntemleri gerçek bir veri seti üzerinde uygulandıktan sonra önerilen yöntemler ile klasik test istatistiklerinin performanslarının karşılaştırılması amacıyla simülasyon çalışması yapılmıştır.

İkinci olarak ilk bölümde geliştirilmiş olan BPA yöntemi boyut sayısı belirlemek için faktör analizine adapte edilmiştir. Önerilen yöntem gerçek veri setleri üzerinde uygulandıktan sonra varsayımsal olarak oluşturulan faktör modellerindeki performansı simülasyon çalışmasıyla araştırılmıştır.

Üçüncü olarak faktör analizinde boyut sayısının belirlenmesinde, paralel ve bootstrap örneklere dayanan ampirik olabilirlik güven aralığı yaklaşımı önerilerek gerçek veride uygulaması yapılmış ve varsayımsal olarak oluşturulan faktör modellerinde performansı araştırılmıştır. Önerilen yaklaşımların KKA ve faktör analizinde boyut sayısı belirlemedeki performansları değerlendirilmiştir.

Anahtar Kelimeler: Paralel Analiz, Kanonik Korelasyon Analizi, Faktör Analizi, Bootstrap, Ampirik Olabilirlik Yöntemi

ABSTRACT

PARALLEL ANALYSİS APPROACHES FOR DİMENSİON SELECTİON İN CANONİCAL CORRELATİON AND FACTOR ANALYSİS

Selahattin AYDOĞDU
Department of Statistics
Phd. Thesis

Adviser: Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK

Parallel analysis (PA) is an empirical method used for determining the number of principal components or factors to retain in the factor analysis. In this thesis, PA based approaches were proposed to determine the number of dimensions in factor analysis and canonical correlation analysis (CCA).

Firstly, PA method was adapted to CCA and its usefulness was investigated in determining the number of dimensions. In this section, bootstrap modified parallel analysis (BPA) was also developed to determine the number of dimensions in CCA. After PA and BPA methods were applied to a real data set in the determination of the number of dimensions in CCA, a simulation study was performed to compare the performance of the proposed methods with the classical test statistics.

Secondly, the BPA method developed in the first section was adapted to factor analysis for determining the number of dimensions. After applying the proposed method on real data sets, performance in hypothetically generated factor models was investigated by simulation study.

Thirdly, empirical likelihood confidence interval approach based on parallel and bootstrap samples for determining number of dimensions in factor analysis was proposed. The proposed approach was applied to real data and its performance was investigated for hypothetical factor models. The performances of the proposed approaches were evaluated for determining the number of dimensions in CCA and factor analysis.

Keywords: Parallel Analysis, Canonical Correlation Analysis, Factor Analysis, Bootstrap, Empirical Likelihood

BÖLÜM 1

GİRİŞ

1.1 Literatür Özeti

Bilimlerin temelini yapılar ve sistemler oluşturur, yapılar alt sistemlerden veya kendisini oluşturan bileşenlerden oluşur ve genellikle çok boyutludur. Bu yapıları ortaya koyabilmek için bilim dallarında birçok teori geliştirilmiştir. Yapıların amacı olayları özetlemektir. Yapı ve yapı teorisini ortaya koymak için kullanılan en önemli istatistiksel teknik faktör analizidir. Faktör analizinde en büyük problemlerinden biri verinin boyut sayısının belirlenmesidir. Boyut sayısı, verilerin yorumlanmasını, düzenlenmesine ve özetlenmesine yardımcı olur. Değişkenler arasındaki bağımlılık yapısının ortadan kaldırılması ve boyut sayısının indirgenmesi istatistiksel yöntemlerin temel amacıdır.

Faktör analizinde ve temel bileşenler analizinde türetilecek maksimum faktör ve temel bileşen sayısı değişken sayısı ile sınırlıdır. Bununla birlikte faktör analizinde değişken sayısı kadar faktör (boyut) türetilmesi faktör analizinden beklenen yararı sağlamayacaktır; temel amaç toplam değişimi (varyansı) değişken sayısı kadar boyutla açıklamak değil, daha az sayıdaki boyutla açıklamaktır. Bu çerçevede, faktör analizi yöntemiyle belirlenecek önemli boyut sayısının ya da kaç adet boyutun dikkate alınacağının belirlenmesi önemlidir. Faktör sayısının az olduğu faktör uzaylarında boyut sayısının olması gerekenden az (underextraction) olması önemli bilgi kaybına yol açacaktır. Ayrıca uygun olmayan yöntemler kullanılarak boyut sayısının olması gerekenden fazla (overexraction) olması yanlış yorumlamalara yol açacaktır. Bu problemin ortadan kaldırılması amacıyla çeşitli yaklaşımlar geliştirilmiştir. Bu yaklaşımlara genellikle durma veya durdurma kuralları denir [1, 2, 3].

Literatürde temel bileşen veya faktör sayısına karar vermek için Guttman-Kaiser (G-K) diğer adıyla öz değer>1 kuralı, yamaç eğim testi (Scree plot), velicer'in MAP (Minimum Average Partial: En küçük ortalamalı kısmi korelasyon) testi, olabilirlik oran testi, Bartlett testi, paralel Analiz (PA) ve varyans yüzdesi ölçütü gibi bazı yöntemler kullanılmaktadır [4, 5, 6, 7, 8, 9, 10, 11, 12, 13].

Bu yöntemlerden en yaygın olarak kullanılan yöntemler Kaiser [14] ve Guttman [15] tarafından önerilen G-K kuralı ve Cattell tarafından önerilen yamaç eğim (scree plot) testidir. Bu yöntemler en çok yararlanılan yaklaşımlar olup hem temel bileşenler hem de faktör analiz yöntemlerinde kullanılabilmektedir. Bu iki yöntem SPSS ve SAS gibi birçok istatistiksel paket programlarında varsayılan olarak seçilmektedir. G-K kuralında öz değeri 1'den büyük olan faktör sayısı anlamlı boyut sayısı olarak alınır. G-K kuralı teorik olarak basit ve kolay kullanımı nedeniyle yaygın olarak kullanılmaktadır, fakat bu yaklaşımın uygulanmasında bazı problemler ortaya çıkmaktadır. Bunlardan birincisi bu kural örneklem büyüklüğünden etkilenmektedir [6, 10, 13, 17, 16]. Örneklem büyüklüğü arttıkça birden büyük özdeğer sayısının arttığı görülmektedir. İkincisi özdeğerin 1'den büyük olması keyfi seçimlere yol açmaktadır, örneğin özdeğeri 1.001 boyut olarak kabul edilirken, 0.999 özdeğeri boyut olarak kabul edilmemektedir [18]. Simülasyon çalışmalarında karşılaşılan bir diğer durum ise faktör yapısından bağımsız 1'den büyük özdeğerlerin sayısının değişken sayısının üçte biri kadar çıkmasıdır [19, 20].

Cattell'in Yamaç eğim (Scree plot) testi çok kullanılan tekniklerden biridir. Bu yöntem ise grafiksel bir sonuç ortaya koymaktadır, grafikte ilk keskin kırılmaya kadar olan faktörler anlamlı faktör olarak kabul edilir. Uygulamada bazen grafikte keskin düşüşlerin olmamasından veya birden fazla keskin düşüş olmasından dolayı kesin bir karar vermek zorlaşmaktadır. Araştırmacılara göre grafik ile boyut sayısına karar verilirken göz kararı ile grafik yorumlamadaki öznellik farklı sonuçlara neden olmaktadır. Birçok araştırmacı bu yöntemlerin faktör (boyut) sayısı belirlemede performansını yeterli görmemişlerdir [6, 10, 13].

Horn [7] tarafından önerilen paralel analiz (PA) faktör analizinde faktör sayısı belirlemede araştırmacılar tarafından en çok önerilen yöntemdir. Bu yöntem G-K kuralının örneklem tabanlı uyarlamasıdır. Yapılan çalışmalarda PA yönteminin faktör ve temel bileşen analizinde boyut sayısına karar verilmesinde yeterli bir teknik olduğu ve

diğer yöntemlere oranla daha iyi performans gösterdiği açıkça görülmektedir [6, 8, 11, 13, 16, 19, 21, 22, 23, 24, 25]. Paralel analiz yönteminin modifikasyonuyla ilgili birçok çalışma mevcuttur. PA yönteminde ortalama özdeğerin kullanımı gereğinden çok faktör belirleme (overexraction) problemi ortaya çıkarmıştır. Glorfeld [22], Horn'un PA'sını modifiye ederek, ortalama özdeğer kullanımı yerine özdeğerlerin örneğin 95. veya 99. persentil değerleri gibi $(1-\alpha)$ persentil değerlerinin kullanılmasını önermiştir [19]. Rasgele veriden elde edilen özdeğerlerin dağılımının %95 ve %99 gibi istenilen belli bir yüzdesinden veya kısaca 95. veya 99. persentilden büyük olduğunda ilgili temel bileşen veya faktörün analize katılmasına karar verilmesi ve eşik değerlerin yüksek persentiller formunda olması desteklenmiştir [12, 22, 25, 26]. PA'nın temel eksen faktörleşme yöntemi (PA-TEF) [27] ve PA'nın minimum rank faktör analizi (PA-MRFA) [28] modifikasyonları ise faktör çıkarma yöntemleri ve değişkenlerin seviyesinin ölçümü ile ilgilidir. PA'nın temel bileşen yöntemi (PA-TBA) ile kullanılması özellikle tek faktörlü ve düşük korelasyonlu modellerde PA-TEF yönteminden daha iyi performans göstermiştir [4]. PA'nın sıralı (ordinal) değişkenlerdeki [29] ve ikili (binary) değişkenlerdeki [25] boyut sayısını belirlenmedeki performansı ile ilgili çalışmalar bulunmaktadır. Bu çalışmalarda, pozitif çarpıklık koşulu altında, PA'nın pearson korelasyon ile kullanılması tetrakorik korelasyon ile kullanılmasının daha iyi olduğu görülmüştür [21]. Sıralı değişkenler için PA'nın polikorik korelasyon ile kullanılması önerilmektedir [30, 31]. PA sürekli verilerde dağılım şeklinden çok az etkilenmektedir [22, 32]. Buja ve Eyupoğlu [26], paralel örnekleri normal dağılım yerine permütasyon yöntemi ile elde ederek PA tekniğini uygulamış ve yapılan simülasyon çalışmasında permütasyon yöntemi kullanılarak uygulanan PA ile klasik PA yönteminin sonuçları benzer bulunmuştur. Lorenzo-Seva v.d. [9], boyut sayısını belirlemek için uyum analizine (correspondence analysis) PA'yı adapte etmiş ve PA'nın kullanılabilirliğini örnekler üzerinde göstererek simülasyon çalışmasıylada göstermiştir. PA tekniği, SPSS, SAS, Pyhton, Stata ve ViSta gibi istatistik paket programlarında sadece faktör analizinde faktör sayısına karar vermede bir seçenek olarak bulunmaktadır [5, 10, 18, 31].

1.2 Tezin Amacı

Çok değişkenli istatistiksel yöntemlerde, araştırılacak problemin içerisinde birbiriyle ilişkili çok sayıda değişken bulunabilir. Çok değişkenli istatistiksel yöntemlerin en

önemli amaçlarından biri de birbiriyle ilişkili değişken sayısını azaltmaktır. Bu ilişkili değişken sayısı yeteri kadar azaltıldığında problemin karmaşıklığı ortadan kalkar ve daha basit bir yapı elde edilmiş olur. Fakat değişken sayısı azaltıldığında problemle ilgili bilgi kaybı meydana gelecektir, geliştirilen istatistiksel yöntemlerle minimum bilgi kaybı ve daha az değişken sayısıyla problemin çözümü hedeflenmektedir. Çok değişkenli istatistiksel yöntemlerde boyut sayısı indirgeme problemi olarak adlandırılan bu problemin çözümü için çeşitli yöntemler önerilmektedir. Bu tez çalışmasında kanonik korelasyon analizi (KKA) ve faktör analizi yöntemlerinde boyut sayısı belirleme problemi için PA'ya dayanan alternatif yöntemler sunulmaktadır.

Bu tez çalışması beş ana bölümden oluşmaktadır.

İlk bölümde, literatür özeti, tezin amacı ve tezin içeriğiyle ile ilgili bilgiler verilmiştir.

İkinci bölümde, boyut sayısı belirlemek için KKA yöntemine Paralel analiz (PA) ve Bootstrap Modifiyeli Paralel Analiz (BPA)'nın adaptasyonu yapılarak PA ve BPA'nın KKA'da boyut sayısı problemine uygulanabilirliği araştırılmıştır. PA ve BPA'nın KKA'da boyut sayısının belirlenmesinde gerçek bir veri seti üzerinde uygulandıktan sonra boyut sayısı belirlemede önerilen yöntemler ile klasik test istatistiklerinin performansları simülasyon çalışmalarıyla karşılaştırılmıştır.

Üçüncü bölümde, ikinci bölümde geliştirilmiş olan BPA yöntemi boyut sayısı belirlemek için faktör analizine adapte edilmiştir. Önerilen yöntem gerçek veri setleri üzerinde uygulandıktan sonra varsayımsal olarak oluşturulan faktör modellerindeki performansı simülasyon çalışmasıyla araştırılmıştır.

Dördüncü bölümde, faktör analizinde boyut sayısının belirlenmesinde, Bootstrap ve paralel örnekler yardımıyla elde edilen özdeğerlerin iki örneklem ampirik olabilirlik güven aralığı (Empirical Likelihood Confidence Intervals) yaklaşımı (PELCI) önerilerek gerçek veride uygulaması yapılmış ve varsayımsal olarak oluşturulan faktör modellerinde performansı araştırılmıştır.

Son bölümde ise yöntemlerin sonuçları tartışılarak, çalışmalarla ilgili genel değerlendirme yapılmış ve önerilerde bulunulmuştur.

1.3 Orjinal Katkı

Bu tez çalışmasının KKA ve faktör analizi yöntemlerinde boyut belirleme problemine katkı sağlayacağı düşünülmektedir. KKA'da boyut sayısı belirlemede PA ve BPA

yöntemleri ilk defa kullanılmıştır. Faktör analizinde boyut sayısını belirlemede BPA yöntemi ilk defa kullanılmış ve performansı araştırılmıştır. Ayrıca ilk defa faktör analizinde boyut sayısı belirlenmesinde iki örneklem ampirik olabilirlik güven aralığı yaklaşımının kullanımı önerilmiştir.

BÖLÜM 2

KANONİK KORELASYON ANALİZİNDE BOYUT SAYISININ BELİRLENMESİNDE PARALEL ANALİZ ve BOOTSTRAP MODİFİYELİ PARALEL ANALİZ YAKLAŞIMLARI

"Şimşek, G. G., ve Aydoğdu S. (2016). Parallel analysis approach for determining dimensionality in canonical correlation analysis, Journal of Statistical Computation and Simulation. 86(17),3419-3431." makalesi [33] bu bölümden üretilmiştir.

Hotelling [34] tarafından geliştirilmiş olan kanonik korelasyon analizi (KKA) iki değişken seti arasındaki ilişkileri analiz etmek için kullanılan çok değişkenli istatistik yöntemlerden biridir. KKA, her bir setteki orjinal değişkenlerden, aynı set içinde korelasyonsuz olacak ve setler arasındaki korelasyonu maksimum yapacak şekilde, orjinal değişkenlerin lineer kombinasyonları olan yeni değişkenleri bulmaya çalışır. Bu yeni değişken çiftlerine kanonik değişkenler ve kanonik değişkenler arasındaki korelasyona da kanonik korelasyon denir [35, s.477]. Kanonik korelasyon sayısı boyut sayısı olarakta adlandırılır. Boyut sayısının belirlenmesi, KKA'nın en temel adımıdır ve aynı zamanda iki veri seti arasındaki ilişkinin kanonik değişken sayısının belirlenmesidir. Boyut sayısının olması gerekenden az olarak bulunması önemli kanonik korelasyonların kaybına yol açacaktır, boyut sayısının olması gerekenden fazla olması ise yanlış yorumlamalara yol açacaktır.

KKA'da boyut sayısı belirlenmesi ile ilgili literatürde birçok çalışma bulunmaktadır. Mendoza v.d. [36], KKA'da boyut sayısı belirlemek için kullanılan dört istatistiksel testi karşılaştırmıştır. Bu yöntemlerden üçü sıralı istatistiksel test, diğeri ise Hariss [37] tarafından önerilen testtir. Yapılan simülasyon çalışması sonucunda sıralı istatistiksel testlerden olan Lawley [38] χ^2 yaklaşımı diğer testlere oranla daha başarılı sonuçlara

ulaşmıştır. Fujikoshi ve Veitch [39], model seçimi için Akaike bilgi kriteri (AIC) ve regresyondaki Mallows Cp istatistiğine dayanan normallik varsayımı altında KKA'da boyut seçimi için iki yöntem geliştirmiştir. Gunderson ve Muirhead [40] bu iki yöntemi normallik varsayımı aranmadan KKA'da boyut sayısı seçimi için kullanmışlardır. Fan [41], Yapısal eşitlik modeli (YEM) seçim yaklaşımına dayanan önemli kanonik değişkenleri elde etmek için istatistiksel bir test prosedürü önermiştir. Bu yöntemde YEM parametrizasyonu kullanarak KKA modellerini kısıtlı en düşük boyut ile kısıtsız en yüksek boyut arasından χ^2 fark testlerini kullanarak seçer. Caliński ve Krzyśko [42], KKA'da gereksiz boyutları sıralayarak Lawley-Hotteling iz istatistiği ve Markus'un [43] oransal yöntemini kulanarak kapalı test prosedürünü (Closed test pocedure) önermişlerdir. Caliński v.d. [44], bu yöntemleri AIC ve 1) Lawley-Hoteling iz istatistiğne dayanan Bartlett ve Fujikoshi istatistiği, 2) Olabilirlik oran istatistiğine dayanan Bartlett ve Lawley istatistiği, 3) Bartlett-Nanda-Pillai iz istatistiğne dayanan Bartlett ve Fujikoshi χ^2 yaklaşımlarını simülasyon çalışması ile karşılaştırmıştır. Çalışmadaki simülasyon sonuçlarına göre AIC yöntemi kullanımı tavsiye edilmemiştir ve diğer üç yöntem ise özellikle küçük örnekler ve kanonik korelasyonların küçük olduğu anakütlelerde performansları tatmin edici bulunmamamıştır. Kanonik korelasyonların tümünün 0.5'ten büyük olduğu koşullarda Lawley-Hoteling iz istatistiğine dayanan yöntemler en iyi performansı göstermiştir. Kanonik korelasyonların en az birinin 0.5'e eşit veya küçük olduğunda ise en iyi performansı Fujikoshi tarafından önerilen Lawley-Hoteling iz istatistiğine dayanan Fujikoshi χ^2 yaklaşımı göstermiştir. Genel olarak bu yöntemler arasında en iyi performansı Lawley-Hotelling iz istatistiğine dayanan metodlar göstermiştir. Yamada ve Sugiyama [45], KKA'da boyut sayısı belirlemek için bir permütasyon testi geliştirmiştir. Bu yöntem özellikle örneklem sayısı fazla olduğu koşullarda iyi performans göstermiştir fakat örneklem sayısı az olduğunda performansı zayıf kalmıştır. Aoki ve Sato [46] kategorik verilerde KKA'da boyut sayısı tahmini için istatistiksel bir yöntem geliştirmiştir. Lee [47] KKA'da küçük örneklerde yanlılığı araştırmış ve Bootstrap örneklemesi kullanarak Jackknife yanlılık düzelticisi ile kanonik korelasyonlar için daha iyi güven aralığı oluşturmuştur.

Bu bölümde, Paralel analiz ve Bootstrap modifiye paralel analiz yöntemlerinin KKA'ya adaptasyonu yapılmıştır. KKA'da boyut sayısı belirlemede kullanılan istatistiksel test

prosedürleri ile beraber PA ve BPA yöntemlerinin performanslarının karşılaştırılması hem gerçek bir veri için yapılmış, hem de detaylı simülasyon çalışmalarıylada desteklenmiştir.

2.1 Kanonik Korelasyon Analizinde Boyut Sayısı Belirlenmesinde Kullanılan İstatistiksel Test Prosedürleri

$\mathrm{Y}' = \left[Y_1, Y_2, ..., Y_p \right]$ ve $\mathrm{X}' = \left[X_1, X_2, ..., X_q \right]$ iki rasgele veri seti matrisleri olsun. Burada $s = \min(p,q) = p \le q$, S ve R sırasıyla örnek kovaryans ve korelasyon matrisleri olmak üzere, $(\mathrm{Y}', \mathrm{X}')'$'nin örneklem sayısı $N = n+1$ ve $(p+q)$ değişkenli normal dağılımdan $\mathrm{N}_{p+q}(0, \Sigma)$ elde edilir. KKA'da amaç $U_i = \alpha_i' \mathrm{Y}$ ve $V_i = \beta_i' \mathrm{Y}$, $i = 1, ..., s$ kanonik değişkenlerinin bulunup U_i ve V_i arasındaki kanonik korelasyonu maksimum yapmaktır. $(\mathrm{Y}', \mathrm{X}')'$'nin, Σ, S ve R matrisleri,

$$\Sigma = \begin{bmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{bmatrix}, \ \mathrm{S} = \begin{bmatrix} \mathrm{S}_{YY} & \mathrm{S}_{YX} \\ \mathrm{S}_{XY} & \mathrm{S}_{XX} \end{bmatrix}, \text{ ve } \mathrm{R} = \begin{bmatrix} \mathrm{R}_{YY} & \mathrm{R}_{YX} \\ \mathrm{R}_{XY} & \mathrm{R}_{XX} \end{bmatrix} \tag{2.1}$$

burada $\rho_1^2 \ge \cdots \ge \rho_s^2 \ge 0$, $\Sigma_{yy}^{-1}\Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}$'nin özdeğerleridir.

Anakütle kovaryans matrisi Σ pozitif tanımlı, $(p+q)$ değişkenli normal dağılımdan gelmek üzere, anakütle kanonik korelasyonları $1 \ge \rho_1 \ge \ldots \ge \rho_d \ge \rho_{d+1} \ge \ldots \ge \rho_s \ge 0$ olup, $\Sigma_{yy}^{-1}\Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}$'nin özdeğerlerinin pozitif kareköklerididir [48, s.324]. Örneklem kanonik korelasyonları $1 \ge r_1 \ge \ldots \ge r_s \ge 0$ olup, $S_{yy}^{-1}S_{yx}S_{xx}^{-1}S_{xy}$'nin özdeğerlerinin pozitif kareköklerididir.

KKA'da sıfır olmayan kanonik korelasyon sayısı d boyut sayısı olup, ayrıca Σ_{yx}'nin bilinmeyen rankı, aynı zamanda üzerinde çalışılacak kanonik değişken sayısıdır [39, 42]. Burada her bir hipotez birbirinden bağımsız olarak test edilir.

Boyut sayısına karar vermek için öncelikle $H_0 : \Sigma_{\mathrm{yx}} = 0$ sıfır hipotezi,

$$H_0 : \rho_1 = \rho_2 = \ldots = \rho_s = 0 \tag{2.2}$$

test edilir. H_0 kabul edilir ise $d=0$ olur. H_0 reddedilirse $\rho_1^2 \neq 0$, burada boyut sayısı $d=1$ olur ve takip edilen hipotez,

$H_0(1): \rho_1^2 \neq 0$ ve $\rho_2^2 = \rho_3^2 = \cdots = \rho_s^2 = 0$

(2.3)

benzer şekilde $H_0(1)$ kabul edilir ise $d=1$'dir. $H_0(1)$ red edilir ise $\rho_1^2 \neq 0$ ve $\rho_2^2 \neq 0$'dir yani boyut $d=2$'dir. Bunu genelleştirirsek, takip eden hipotez kabul edilene kadar devam eder

$$H_o(d): \rho_1^2 \neq 0,...,\rho_d^2 \text{ ve } \rho_{d+1}^2 = \rho_{d+2}^2 = ... = \rho_s^2 = 0,\ d = 1,2,...,s-1 \quad (2.4)$$

test edilir. H_d, $d = 0,1,2,...,s-1$ hipotezlerinin testinde aşağıda verilen test istatistikleri önerilmiştir. Aşağıdaki notasyonların benzeri Caliński vd. [44] çalışmasında kullanılmıştır.

Lawley-Hotelling iz istatistiği,

$$T_d^2 = \sum_{i=d+1}^{p} \frac{r_i^2}{1-r_i^2} \quad (2.5)$$

Wilks'in Λ_d olabilirlik oran istatistiği,

$$\Lambda_d = \prod_{i=d+1}^{p} \left(1-r_i^2\right) \quad (2.6)$$

Bartlett-Nanda-Pillai iz istatistiği,

$$V_d = \sum_{i=d+1}^{p} r_i^2 \quad (2.7)$$

H_d hipotezi altında, yaklaşımları Bartlett [49, 50, 51] tarafından önerilen,

$$S_B\left(T_d^2\right) = (n-p-q-2)T_d^2 \quad (2.8)$$

$$S_B\left(\Lambda_d\right) = -\left[n-1-\frac{1}{2}(p+q+1)\right]\ln \Lambda_d \quad (2.9)$$

$$S_B\left(V_d\right)=\left(n-1\right)V_d \tag{2.10}$$

Lawley [36] tarafından önerilen,

$$S_L\left(\Lambda_d\right)=-\left[n-1-\frac{1}{2}\left(p+q+1\right)+\sum_{i=1}^{d}r_i^{-2}\right]\ln\Lambda_d \tag{2.11}$$

Fujikoshi [52] tarafından önerilen,

$$S_F\left(T_d^2\right)=-\left(n-p-q-2+\sum_{i=1}^{d}r_i^{-2}\right)T_d^2 \tag{2.12}$$

$$S_F\left(V_d\right)=\left(n-1-2d+\sum_{i=1}^{d}r_i^{-2}\right)V_d \tag{2.13}$$

istatistikleri $(p-d)(q-d)$ serbestlik dereceli χ^2 dağılımına yaklaşmaktadırlar. İstatistik paket programlarında, Bartlett veya Lawley tarafından modifiye edilen Wilks'in Λ_{d} istatistiğinin kullanılması tercih edilmektedir. Mendoza vd. [36] tarafından yapılan simülasyon çalışmasında χ^2 yaklaşımlarına dayanan $S_B(\Lambda_{\mathrm{d}})$ ve $S_L(\Lambda_{\mathrm{d}})$ istatistikleri, Rao'nun F [53] ve Harris [37] tarafından önerilen yöntemlerden daha iyi şekilde boyut sayısını tespit etmişlerdir. Caliński v.d. [44] test istatistiklerinin performanslarını karşılaştırmak için Caliński vd. [44] tarafından yapılan simülasyon çalışmalarında, sıfır olmayan bütün kanonik korelasyonlar 0.5'in üstünde olduğunda T_d^2 'ye dayanan test istatistikleri T_d^2, $S_B\left(T_d^2\right)$ ve $S_F\left(T_d^2\right)$'nin daha iyi sonuçlar verdiği, anakütlede sıfır olmayan kanonik korelasyonlardan en az biri 0.5'in altında olduğunda ise $S_F\left(T_d^2\right)$'nin en iyi performansı sağladığı sonuçlarına varılmıştır. Gerçek veri setlerinde araştırmacılar genellikle Rao'nun F istatistiğini veya χ^2 yaklaşımlarına dayanan $S_B(\Lambda_{\mathrm{d}})$ ve $S_L(\Lambda_{\mathrm{d}})$ test istatisklerini kullanırlar.

Paralel Analiz (PA) Monte Carlo simülasyonuna dayalı bir tekniktir. Bu bölümde PA metodu KKA'ne adapte edilerek, KKA'deki boyut sayısına karar vermek için PA'nın kullanılabilirliği araştırılmaktadır.

2.2 Paralel Analizin İstatistiksel Temelleri

Horn [7] tarafından önerilen paralel analiz Monte Carlo simülasyonuna dayanan bir yöntemdir. Paralel analiz (PA) yöntemi G-K kuralının örneklem tabanlı uyarlamasıdır. Horn [7] özellikle küçük örneklerde faktörlerin açıklanan varyansı maksimum yapacak şekilde bulunmasına zorlanmasından dolayı, başlangıç faktörlerinin, örneklerde, anakütlede olduğundan daha yüksek özdeğerlere sahip olduklarını saptayarak, PA'yı geliştirmiştir. Horn tarafından önerilmiş olan bu teknikte n×p orijinal veri setiyle aynı sayıda değişken ve gözlem içerecek şekilde, kısaca veri setine paralel, birim kovaryans matrisli çok değişkenli normal dağılımdan n×p boyutlu rastgele veri setleri üretildikten sonra, orijinal veri setine ait korelasyon matrisinden elde edilen özdeğerler ile, rastgele veri setlerinden elde edilen korelasyon matrislerinin özdeğerlerinin ortalaması karşılaştırılarak, örneğin i. sıradaki özdeğerleri, i.ortalama özdeğerden büyük olduğunda karşılık gelen temel bileşen veya faktörün analizde kalmasına karar verilmektedir.

Paralel analiz rasgele veri matrisi teorisi bakış açısı ile incelendiğinde Tracy-Widom (TW) testi ile aynı temellere sahip olduğu görülmektedir. TW testi rasgele matris teorisi ve Tracy-Widom dağılımına dayanan bir yöntemdir. Bu yöntem, ilk özdeğer tespitinde paralel analiz ile aynı sonuçları vermektedir. Tracy-widom prosedürü Johnstone teoremine dayanmaktadır [54].

Johnstone teoremi, $\mathrm{X}_{(\mathrm{nxp})}$ birbirinden bağımsız ve çok değişkenli normal dağılıma sahip olsun $\mathrm{X} \sim \mathrm{N}(0, \mathrm{I})$. Örneklem kovaryans matrisi $S = X^{'}X$ ve l_1 ise S kovaryans matrisinin en büyük özdeğeridir. Merkezi ve konum ölçüleri,

$$\eta_{np} = (\sqrt{n - 1/2} + \sqrt{p - 1/2})^2 \quad , \tag{2.14}$$

$$\xi_{np} = (\sqrt{n - 1/2} + \sqrt{p - 1/2})(\frac{1}{\sqrt{n - 1/2}} + \frac{1}{\sqrt{p - 1/2}})^{\frac{1}{3}} \tag{2.15}$$

eğer,

$$\frac{n}{p} \to \gamma \geq 1 \tag{2.16}$$

ise

$$L_1 = \frac{l_1 - \mu_{np}}{\sigma_{np}} \sim F_1(w,1) \tag{2.17}$$

Burada $\mu_{np} = \eta_{np}$, $\xi_{np} = \sigma_{np}$ ve $F_1(w,1)$ yaklaşık Tracy-Widom dağılımıdır. L_1 ise l_1 özdeğeri için TW istatistiğidir. w Tracy-Widom dağılımının parametresidir.

Tracy Widom testi süreci,

nxp boyutlu X veri matrisi alınır ve pxp boyutlu S kovaryans matrisi hesaplanır.

S kovaryans matrisinden ilk özdeğere elde edilir.

Johnstone teoreminden l_1 özdeğeri için L_1 istatistiği elde edilir ve $\mu_{np} = r^2\eta_{np}$, $\sigma_{np} = r^2\xi_{np}$ parametreleri hesaplanır.

α anlam seviyesinde H_0 hipotezini test etmek için L_1 istatistiği ile Tracy-Widom persentil yaklaşımı karşılaştırılır.

Özdeğerlerin eşitlik testi,

$$\hat{\tau}^2 = \frac{1}{n(p-k+1)} \sum_{q=k}^{p} l_q \tag{2.18}$$

Burada α anlamlılık seviyesinde $H_0 : \lambda_k = \lambda_{k+1} = \ldots = \lambda_p$ test edilir. λ_k için yaklaşım dağılımının $\hat{x}_{100-\alpha}$ persentil değeri,

$$\hat{x}_{100-\alpha} = \hat{\tau}^2(\mu_{n,p-k+1} + x_{100-\alpha}\sigma_{n,p-k+1}) \tag{2.19}$$

şeklinde hesaplanır. Burada $x_{100-\alpha}$, Tracy-Widom dağılımının $100-\alpha$ persentil değeridir. $\hat{\tau}^2$ ise τ^2 ölçü parametresinin kestiricisidir.

ilk özdeğer için PA prosedürü,

- nxp boyutlu X veri matrisinden, pxp boyutlu örneklem kovaryans matrisi S hesaplanır. $S = X^{'}X$ matrisinden en büyük özdeğer olan l_1 hesaplanır.
- x^r rasgele matrisi çok değişkenli normal dağılımdan $N(0,\sum_r)$ türetilir. Kovaryans matrisi hesapalanır ve kovaryans matrisinde en büyük özdeğer olan l_1^r hesaplanır.
- 2. adım $N >> 1$ kez tekrarlanır ve N özdeğerler toplanır l_1^r'nin referans dağılımı oluşturulur.

- Referans dağılımdan $100-\alpha$ persentil değeri hesaplanır.
- l_1 ile $100-\alpha$ persentil değeri, α anlam seviyesinde H_0 hipotezi ile karşılaştırılır.

İlk özdeğer için TW testi ile PA aynı hipotezle test edilir. Burada TW testi, random matris asimptotik teorisine dayanır ve PA yeniden örnekleme kullanılarak elde edilen yeniden örnekleme dağılımından tahmin edilir. Bu benzerlikle birlikte PA'nın katı istatistiksel temeli ve tamamıyla çıkarımsal yaklaşım olduğu sonucuna varılır [54].

2.3 Paralel Analizin Kanonik Korelasyon Analizine Adaptasyonu

Kanonik korelasyon analizinde, bir anakütledeki sıfır olmayan her bir kanonik korelasyon bir boyut sayısıdır ve boyutla ilişkili her bir özdeğer (ρ_i^2), $V_i(i=1,...,d)$ tarafında açıklanan U_i'nin varyansının oranının bir ölçüsüdür. Geriye kalan özdeğerler $\rho_i^2(i=d+1,...,s)$, $s-d$ 'ye eşittir ve bunlar anakütlede sıfıra eşittir. Aksine örnekte, bu $\rho_i^2(i=d+1,...,s)$ sıfırdan büyük olabilir. Bu durumda önemsiz boyutların ortaya çıkmasına yol açabilir. Bu oranlama PA'nın KKA'ya adaptasyonu için kullanılır. Önemli boyut sayısıyla ilgili (kanonik korelasyonun karesi olan) özdeğerler, rasgele veriden elde edilen boyut sayısı ile ilgili özdeğerden büyük olmalıdır. PA'nın, KKA'ya adaptasyonu kısaca iki adımda özetlenebilir;

- Ampirik veriden, benzer sayıda (p,q) ve n ile büyük sayıda rasgele veri, bu veriler $\rho_1^2=\rho_2^2=...=\rho_s^2=0$ ile çok değişkenli normal dağılımdan Monte Carlo simülasyon sürecine benzer şekilde KKA için üretilir.
- Ampirik örnekten elde edilen i. özdeğer ile rasgele örneklerden elde edilen i. özdeğerin dağılımı karşılaştırılır. Burada ampirik özdeğer ve rasgele özdeğerler karşılaştırılabilir çünkü özdeğerler ölçekten bağımsız olup, ölçü birimi ve konumdaki değişimden etkilenmez.

2.4 Paralel Analizin Bootstrap Modifikasyonu

KKA'da, Anakütlede i. özdeğer ilgilenilen ρ_i^2 parametesine denktir. Bootstrap persentil $\%100(1-\alpha)$ güven aralığı, tahmini r_i^{2*} bootstrap dağılımına dayanır ve

$$\left[r_{(\alpha/2)}^{2*}, r_{(1-\alpha/2)}^{2*}\right] \tag{2.20}$$

şeklinde tanımlanır. Burada $r^{2*}_{(\alpha/2)}$ alt sınır ve $r^{2*}_{(1-\alpha/2)}$ üst sınır, sırasıyla $100(\alpha/2).$ ve $100(1-\alpha/2).$ bootstrap dağılımının persentil değerleridir [55].

Bu çalışmada, Paralel analiz yönteminin bootstrap ile modifikasyonu BPA olarak adlandırılmıştır. BPA'da KKA'da boyut sayısı problemi için önerilmektedir. BPA'nın KKA'da örneklem hatasından dolayı meydana gelen problemi çözeceği düşünülmektedir. BPA yaklaşımı, $\alpha = .05$ düzeyinde, ampirik veriden B kez iadeli olarak çekilmesine dayanır. i. özdeğer için bootstrap %95 persentil güven aralığının alt sınırı, rasgele veriden elde edilen i. özdeğerin dağılımı ile karşılaştırılır.

2.5 Kanonik Korelasyon Analizinde Gerçek Veri İçin Paralel Analiz Yöntemi

Bu bölümde KKA'da boyut sayısına karar vermek için bir kimyasal reaksiyon veri seti alınmıştır. Kimyasal reaksiyon deneyi N=19 kere tekrarlanmıştır. Burada bağımsız değişkenler X1: Sıcaklık, X2: Konsantrasyon (yoğunluk) ve X3: Zaman olarak alınmıştır. Bağımlı değişkenler ise Y1: Değişmeyen maddelerin yüzdesi, Y2: İstenen değişimin gerçekleştiği maddeler ve Y3: İstenmeyen yan ürünün yüzdesi olarak belirlenmiştir. Ayrıca modele X_1^2, X_2^2, X_3^2 ve X_1X_2, X_1X_3, X_2X_3'te eklenmiştir [56, s.390]. Bağımlı değişken sayısı $p = 3$ ve bağımsız değişken sayısı $q = 9$'dur. Bu model için kanonik korelasyon analizinde boyut sayısına karar vermek için klasik test yöntemleri ve PA yöntemi kullanılmıştır. Sonuçlar Çizelge 2.1 ve 2.2'de verilmiştir.

PA yaklaşımı kullanılarak $N(\mu = 0, \sum = I)$ olan normal dağılımdan N=19 büyüklüğünde ve korelasyonsuz olacak şekilde $p + q = 12$ değişkenli 1000 tane rasgele örnek üretilmiştir. İlk olarak, rasgele değişkenlerden üretilen özdeğerlerin dağılımından özdeğerlerin ortalama, medyan, 95. ve 99. persentil değerleri hesaplanır. Daha sonra orijinal veriden elde edilen sıralı özdeğerler (kanonik korelasyonların karesi) ile karşılaştırılır. Çizelge 2.1'de gösterildiği gibi, orijinal veriden elde edilen ilk iki özdeğerin, PA yaklaşımı ile elde edilen özdeğerlerin ortalama, medyan, 95. ve 99. persentil değerlerinden büyük olduğu görülmektedir. Buradan boyut sayısının iki olduğu görülmektedir. İkinci olarak BPA yaklaşımı için de, bootstrap yöntemiyle N=19 veri setinden 1000 bootstrap örneği iadeli olarak çekilerek, %95 güven aralığının alt sınırı ($\rho^{2*}_{(\alpha/2)}$) alınır. Burada ρ^{2*} bootstraplı özdeğerin $100(\alpha/2)$'inci değeri olup

bootstrap dağılımının $100(0.05/2)=2.5$'inci persentil değerine karşılık gelmektedir. Daha sonra bu özdeğerler PA'dan üretilen özdeğerlerin ortalama, medyan, 95., 99. persentil değerleri için %95 bootstrap güven aralığının alt sınırına dayanan persentil değerleri ile karşılaştırılır. Çizelge 2.1'de görüldüğü gibi özdeğerler için bootstrap %95 güven aralığı ilk iki alt sınırı (0.9841 ve 0.8887) değerleri, PA'dan elde edilen ortalama, medyan, 95. ve 99. persentil değerlerinden büyüktür. Bu durumda her iki PA yaklaşımıylada boyut sayısı $\mathrm{d}=2$ olarak bulunmuştur.

Çizelge 2. 1 Kanonik korelasyonların karesi (özdeğerler), sıralı özdeğerler için bootstrap %95 güven aralığının alt sınırına dayalı persentil değeri ve PA'ya dayanan sıralı özdeğerelerin ortalama, medyan, 95. ve 99. persentil değerleri

Boyut(d)	Kanonik korelasyon karesi (özdeğer)	Özdeğerler için bootstrap %95 CI alt sınırı	Paralel Analiz			
			$PA_{ortalama}$	PA_{medyan}	PA_{95}	PA_{99}
1	.9797	.9841	.7434	.7516	.8993	.9503
2	.9080	.8887	.4979	.4952	.6965	.7544
3	.2147	.2347	.2605	.2536	.4595	.5326

Denklemler 2.8-2.13'te verilen test istatistikleri gerçek veri setinde KKA'da boyut sayısı belirlemek için uygulanmıştır. Çizelge 2.2'de KKA'da boyut sayısının belirlenmesi için hesaplanan örnek özdeğerleri ve p (χ^2 dağılımı $(3-d)(9-d)$, $d=0,1,2$ serbetslik dereceli) olarak verilmiştir.

Çizelge 2. 2 Gerçek veride boyut sayısı için klasik test istatistikleri ve p değeri

	Bartlett χ^2			Lawley χ^2	Fujikoshi χ^2	
Boyut (d)	$S_B(T_d^2)$	$S_B(\Lambda_d)$	$S_B(V_d)$	$S_L(\Lambda_d)$	$S_F(T_d^2)$	$S_F(V_d)$
1	292.09 (*p*=.00)	75.04 (*p*=.00)	37.84 (*p*=.08)	75.04 (*p*=.00)	292.09 (*p*=.00)	37.84 (*p*=.08)
2	50.70 (*p*=.00)	30.22 (*p*=.02)	20.21 (*p*=.21)	30.27 (*p*=.02)	61.06 (*p*=.00)	19.11 (*p*=.26)
3	1.37 (*p*=.99)	2.78 (*p*=.90)	3.86 (*p*=.80)	2.81 (*p*=.90)	1.95 (*p*=.96)	3.46 (*p*=.84)

Barttlet'in $S_B(T_d^2)$ ve Fujikoshi'nin $S_F(T_d^2)$ testleri α=0.01 anlam seviyesinde boyut sayısını d=2 buldukları görülmektedir. Barttlet'in $S_B(\Lambda_d)$ ve Lawley'in $S_L(\Lambda_d)$ testleri α=0.05 anlam seviyesinde boyut sayısını $\mathrm{d}=2$ buldukları görülmektedir. Barttlet'in $S_B(V_d)$ ve Fujikoshi'nin $S_F(V_d)$ testleri α=0.01 anlam seviyesinde boyut sayısını d=1

bulmuşlardır. Barttlet'in $S_B(V_d)$ ve Fujikoshi'nin $S_F(V_d)$ testleri hariç diğer test istatistikleri boyut sayısını d=2 buldukları görülmektedir. PA yöntemi ve bazı klasik test istatistiklerini boyut sayısını d=2 buldukları görülmektedir.

2.6 Simülasyon Çalışması

PA, BPA ve klasik test prosedürlerinin performansları karşılaştırmak için simülasyon tasarımı yapılmıştır. Anakütle modelleri 4 koşul altında oluşturulmuştur;

- Varsayımsal anakütleden değişken sayıları $(p=4,q=4)$, $(p=4,q=8)$, $(p=6,q=7)$ ve $(p=6,q=8)$ olarak seçilmiştir Simülasyon çalışmalarında genellikle küçük veri setleri için değiken sayısı 4 ve 6 arasından seçilmektedir [7, 17, 38].
- Örneklem sayıları $N=25,50,75$ ve 100 olarak belirlenmiştir. $\mathrm{N}=25$ küçük örneklem sayısı olarak seçilmiş ve iki katı, üç katı ve dört katı alınarak örneklem büyüklükleri artırılmıştır.
- Boyut sayısı yöntemlere aşırı belirleme (overestimate) şansı verilebilmesi bakımından iki set içerisinde daha az değişkenli olan veri setindeki değişken sayısından az olarak alınmıştır. Boyut sayısı $d=0,1,2,3$ olarak belirlenmiştir.
- Boyut sayısı ve kanonik korelasyon değerleri göz önüne alınarak 7 adet kanonik korelasyon yapısı oluşturulmuştur. Çizelge 2.3'te görüldüğü gibi kanonik korelasyonlara 0 ile 0.8 arasında değişen değerler verilmiştir. İstenilen sayıda ve büyüklükte anakütle kanonik korelasyonlarına sahip örneklerin elde edilmesi için, Kanonik korelasyonların tümünün 0 eşit, 0,5'ten büyük olacak ve en az birinin 0. 5'ten küçük olacak şekilde sıfırdan farklı iki kanonik korelasyonlu modeller oluşturularak 7 adet kanonik korelasyon yapısı üretilmiştir.

Çizelge 2. 3 Simülasyon çalışmasında kullanılan kanonik korelasyonlar

$p=4$	$p=6$
Model 1: $\rho_1=\rho_2=\rho_3=\rho_4=0$	**Model 1:** $\rho_1=\rho_2=\rho_3=\rho_4=\rho_5=\rho_6=0$
Model 2: $\rho_1=.8,\ \rho_2=\rho_3=\rho_4=0$	**Model 2:** $\rho_1=.8,\ \rho_2=\rho_3=\rho_4=\rho_5=\rho_6=0$
Model 3: $\rho_1=.7,\ \rho_2=\ \rho_3=\rho_4=0$	**Model 3:** $\rho_1=.7,\ \rho_2=\ \rho_3=\rho_4=\rho_5=\rho_6=0$
Model 4: $\rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=0$	**Model 4:** $\rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=\rho_5=\rho_6=0$
Model 5: $\rho_1=.7,\ \rho_2=.5,\ \rho_3=\rho_4=0$	**Model 5:** $\rho_1=.7,\ \rho_2=.5,\ \rho_3=\rho_4=\rho_5=\rho_6=0$
Model6: $\rho_1=.8,\rho_2=.7,\rho_3=.5,\rho_4=0$	**Model6:** $\rho_1=.8,\ \rho_2=.7\ \rho_3=.5,\rho_4=\rho_5=\rho_6=0$
Model7: $\rho_1=.7,\rho_2=.5,\rho_3=.4,\rho_4=0$	**Model7:** $\rho_1=.7,\rho_2=.5,\ \rho_3=.4,\rho_4=\rho_5=\rho_6=0$

Bu 7 adet kanonik korelasyon yapısı,

- $d = 0$ modellerinde tüm kanonik korelasyonlar 0'dır. Yani boyut sıfırdır.
- $d = 1$ modellerinde ilk kanonik korelasyon 0.8'dir, diğer kanonik korelasyonlar 0'dır.
- $d = 1$ modellerinde ilk kanonik korelasyon 0.7'dir, diğer kanonik korelasyonlar 0'dır.
- $d = 2$ modellerinde ilk kanonik korelasyon 0.8'dir, ikincisi 0.7'dir ve diğer kanonik korelasyonlar 0'dır.
- $d = 2$ modellerinde ilk kanonik korelasyon 0.7'dir, ikincisi 0.5'dir ve diğer kanonik korelasyonlar 0'dır.
- $d = 3$ modellerinde ilk kanonik korelasyon 0.8'dir, ikincisi 0.7'dir, üçüncüsü 0.5 'dir ve diğer kanonik korelasyonlar 0'dır.
- $d = 3$ modellerinde ilk kanonik korelasyon 0.8'dir, ikincisi 0.7'dir, üçüncüsü 0.5 'dir ve diğer kanonik korelasyonlar 0'dır.
- $d = 3$ modellerinde ilk kanonik korelasyon 0.7'dir, ikincisi 0.5'dir, üçüncüsü 0.4 'dür ve diğer kanonik korelasyonlar 0'dır.

Yukarıdaki 7 adet simülasyon deseni Caliński vd. [44, p. 732, Tablo 1] tarafından da kullanılmıştır.

Kanonik korelasyonlar ve dolayısıyla bunların test istatistikleri tekil olmayan doğrusal dönüşümler altında değişmez olduğundan [36], $r_1^2,\ldots,r_p^2$'nin ortak dağılımın sadece $\rho_1^2,\ldots,\rho_p^2$ parametrelerine bağlıdır. Anakütle kovaryans matrisi Σ'nın kanonik formu;

$$\Sigma = \begin{bmatrix} I_p & \Delta \\ \Delta & Iq \end{bmatrix} \tag{2.21}$$

şeklinde seçilebilir. I_p ve I_q sırasıyla p ve q mertebesinden birim matrisleridir ve Δ ise korelasyonlarının *p*x*q* boyutlu köşegen matrisidir ve matrisin elemanları $(\rho_1,\ldots,\rho_s)$ 'dir.

Bu 7 yapı ve her bir (p,q) için dört adet kombinasyon ile toplamda 28 adet anakütle kanonik korelasyon modeli oluşturulmuştur. Mendoza v.d. [36] tarafından oluşturulan simülasyon deseni kullanılarak, 2.21'de verilen anakütle korelasyon matrisine sahip 28 modelin her biri için 4 farklı örnek büyüklüğünde 1000 örneklem seçilerek yapılan simülasyonla toplamda 112000 simülasyon verisi üretilmiştir.

Simülasyon çalışmasında aşağıda açıklaması verilen sekiz yöntemin performansı her bir örneklem için ayrı ayrı karşılaştırılmıştır. Bu sekiz yöntemden ikisi önerilecek olan PA ve BPA yöntemleridir. Diğer yöntemler ise daha önceki çalışmalarda önerilen χ^2 yaklaşım prosedürlerine dayanan klasik test istatistikleridir. Bu sekiz yöntem aşağıda tanımlanmıştır;

- PA yönteminde ampirik verilerin özdeğerleri ile rasgele veri özdeğerlerinin ortalama, medyan, 95. ve 99. persentil eşik değerleri karşılaştırılır ve sırasıyla $PA_{ortalama}$, PA_{medyan}, PA_{95} ve PA_{99} olarak tanımlanır.
- BPA yönteminde, ampirik verilerin özdeğerlerinin bootstrap %95 güven aralığının alt sınırı ile rasgele veri özdeğerlerin ortalama, medyan, 95. ve 99. persentil eşik değerleri karşılaştırılır ve sırasıyla $BPA_{ortalama}$, BPA_{medyan}, BPA_{95} ve BPA_{99} olarak tanımlanır.
- Altı adet χ^2 yaklaşım prosedürü ise T_d^2, Λ_d ve V_d 'e dayanmaktadır;
- Bu χ^2 yaklaşım prosedürlerinden iki tanesi T_d^2 dayanmakta; biri Barttlett'in χ^2 yaklaşımı $S_B\left(T_d^2\right)$ ve diğeri Fujikoshi χ^2 yaklaşımı $S_F\left(T_d^2\right)$'dir.
- Bu χ^2 yaklaşım prosedürlerinden iki tanesi Λ_{d} dayanmakta; biri Barttlett'in χ^2 yaklaşımı $S_B(\Lambda_d)$ ve diğeri Lawley χ^2 yaklaşımı $S_L(\Lambda_d)$'dir.
- Bu χ^2 yaklaşım prosedürlerinden iki tanesi Λ_{d} dayanmakta; biri Barttlett'in χ^2 yaklaşımı $S_B(V_d)$ ve diğeri Fujikoshi'nin χ^2 yaklaşımı $S_F(V_d)$'dır.

28 modelin her biri için aşağıdaki adımlar uygulanmıştır.

Simülasyon çalışması iki bölümden oluşmaktadır. Birinci bölümde PA yöntemi KKA'ya modifiye edilerek, KKA'daki boyut sayısına karar vermek için PA'nın kullanılabilirliği araştırılmıştır. İkinci bölümde BPA yani bootstrap örneklemesi kullanılarak, boyut sayısına karar verilecek olan örnekten elde edilen özdeğerlerin amprik örnekleme dağılımından anakütle özdeğerleri için persentile dayalı bootstrap güven aralığı hesaplanmış ve tek örneğin özdeğerleri yerine persentillere dayalı bootstrap güven aralıklarının alt sınırları kullanılmıştır.

Çalışmanın ilk bölümündeki klasik PA'nın kanonik korelasyon analizine adaptasyonu simülasyonu $N = 25,50,75$ ve 100 birimlik örnekler için aşağıdaki adımları izlemektedir.

- Birim kovaryans matrisli çok değişkenli normal dağılımdan $N(\mu = 0, \sum = I)$ pxq boyutlu bağımsız rastgele 1000 örnek çekilip, (kanonik korelasyonun karesi olan) özdeğerleri ve özdeğerlerin dağılımının 90. , 95., 99., 99.5. ve 99.9. persentil değerleri hesaplanır.
- Her bir model için varsayımsal anakütleden bir örnek çekilir. Çekilen örnekten elde edilen kanonik özdeğerler için denklem 2.8-2.13'te verilen altı test istatistiği hesaplanıp, hipotez testlerine dayanan boyut sayısı bulunur.
- Çekilen örneğin i. özdeğeri 1.adımda i. özdeğer için karşılık gelen persentil değerleri ile karşılaştırılır. Örneğin i. özdeğeri karşılık gelen i. persentil özdeğerden küçük oluncaya dek boyut sayısı bir arttırılarak PA'ya dayanan boyut sayısı bulunur.
- 1. adıma dönülür ve adımlar 1000 kez tekrarlanır.

Çalışmanın ikinci bölümünde PA yöntemi, bootstrap güven aralığı ile modifiye edilmiş ve yapılan simülasyon çalışması $N = 25,50,75$ ve 100 birimlik örnekler için aşağıdaki adımlardan oluşmuştur.

- Birim kovaryans matrisli çok değişkenli normal dağılımdan $N(\mu = 0, \sum = I)$ pxq boyutlu bağımsız rastgele 1000 örnek çekilip, (kanonik korelasyonun karesi olan) özdeğerleri ve özdeğerlerin dağılımının 90. , 95., 99., 99.5. ve 99.9 persentil değerleri hesaplanır.
- Her bir model için varsayımsal anakütleden bir örnek çekilir. Çekilen örnekten, örnekle aynı boyutta iadeli olarak 1000 örnek çekilip, persentile dayalı %95 bootstrap güven aralığının alt sınırı bulunur.
- Çekilen örnekten elde edilen kanonik özdeğerler için denklem 2.8-2.13'te verilen 6 test istatistiği hesaplanıp, hipotez testlerine dayanan boyut sayısı bulunur.
- i. özdeğerin 2. adımda bulunan alt sınırı 1.adımda i.özdeğer için karşılık gelen persentil değerleri ile karşılaştırılır. Örneğin i.özdeğer için bootstrap güven

aralığının alt sınırı karşılık gelen i. persentil özdeğerden küçük oluncaya dek boyut sayısı bir arttırılarak BPA'ya dayanan boyut sayısı bulunur.

- 1. adıma dönülür ve adımlar 1000 kez tekrarlanır.

2.6 Simülasyon Sonuçları

Her bir simülasyon örneği için χ^2 yaklaşım prosedürleri, PA ve BPA yöntemlerine dayalı boyut seçimi sonuçları kaydedilmiştir. Modellerden elde edilen tüm simülasyon sonuçları detaylı olarak Ek A'daki Tablo A-1'de verilmiştir. Simülasyon sonuçlarının ortalamaları alınarak özet halinde Çizelge 2.4 ve Çizelge 2.5'te verilmiştir. Elde edilen sonuçlara bakıldığında, PA ve BPA'nın boyut seçiminde küçük örneklerde karşılaştırılabilir ve hatta daha iyi sonuçlar verdiği halde büyük örneklerde test istatistiklerine göre performansı düşük olduğu görülmektedir. Simülasyon çalışmasında PA ve BPA yöntemlerinin ortalama, medyan, 99.9, 99.5, 99.0 ve 90. persentil değerlerinin sonuçları elde edilmiştir, fakat simülasyon sonuçları incelendiğinde genellikle 99. ve 95. persentil değerleri daha iyi sonuçlar vermişlerdir. Bu nedenle KKA boyut sayısı analizinde PA ve BPA'nın 99. ve 95. persentil yaklaşımlarının sonuçları değerlendirilmiştir. KKA'da boyut sayısı klasik test istatistikleri χ^2 yaklaşım prosedürleri $\alpha = 0.05$ anlamlılık seviyesinde test edilmiştir. Bu modellerden her bir yöntem için $N = 25, 50, 75$ ve 100 birimlik örnekler için boyut sayısı $d = 0$, 1, 2 ve 3 olacak şekilde elde edilen simülasyon sonuçları tablolar halinde oluşturulmuştur.

Çizelge 2.4'te χ^2 yaklaşım prosedürleri, paralel analiz (PA) ve %95 bootstrap güven aralığı ile modifiye paralel analiz (BPA) yöntemine göre boyut belirleme sonuçlarının ortalamaları verilmiştir.

Çizelge 2.4 Klasik test istatistikleri $(\alpha = 0.05)$ anlamlılık seviyesinde, Paralel Analiz (PA) ile %95 Bootstrap Güven Aralığı ile Modifiye Paralel Analiz (BPA) yöntemine göre boyut sayısı belirleme için göreli frekanslar

			[a,b] χ^2 Yaklaşımı prosedürleri					
			T_d^2 dayalı yöntemler		Λ_d dayalı yöntemler		V_d dayalı yöntemler	
N	*d*	Belirlenen *d(%)*	[1] $S_B\left(T_d^2\right)\partial$	[2] $S_F\left(T_d^2\right)$	[3] $S_B\left(\Lambda_d\right)$	[4] $S_L\left(\Lambda_d\right)$	[5] $S_B\left(V_d\right)$	[6] $S_F\left(V_d\right)$
25	0	Doğru	88.13	88.13	94.28	94.28	**98.85**	**98.85**
		Fazla	11.88	11.88	5.73	5.73	1.15	1.15
		Eksik	30.08	30.08	46.05	46.05	82.50	82.50
	1	Doğru	**66.49**	63.36	50.81	50.50	16.03	16.58
		Fazla	3.44	6.56	3.14	3.45	1.48	0.93
		Eksik	73.83	65.56	77.05	75.71	86.44	89.09
	2	Doğru	25.36	**31.53**	21.70	22.45	12.16	10.01
		Fazla	0.81	2.91	1.25	1.84	1.40	0.90
		Eksik	94.61	88.13	93.26	91.71	93.15	94.99
	3	Doğru	5.20	**10.91**	6.30	7.58	6.15	4.46
		Fazla	0.19	0.96	0.44	0.71	0.70	0.55
50	0	Doğru	92.33	92.33	94.73	94.73	**97.48**	**97.48**
		Fazla	7.68	7.68	5.28	5.28	2.53	2.53
		Eksik	3.34	3.34	7.50	7.50	22.05	22.05
	1	Doğru	**91.68**	89.96	88.45	87.89	74.91	75.18
		Fazla	4.99	6.70	4.05	4.61	3.04	2.78
		Eksik	30.11	26.35	34.30	32.93	42.98	44.71
	2	Doğru	67.33	**69.09**	63.00	63.61	54.43	52.83
		Fazla	2.56	4.56	2.70	3.46	2.60	2.46
		Eksik	66.38	58.61	66.84	63.60	68.50	69.91
	3	Doğru	32.36	**38.63**	31.55	33.99	29.66	28.09
		Fazla	1.26	2.76	1.61	2.41	1.84	2.00
75	0	Doğru	93.80	93.80	95.18	95.18	**97.03**	**97.03**
		Fazla	6.20	6.20	4.83	4.83	2.98	2.98
		Eksik	0.15	0.15	0.59	0.59	2.15	2.15
	1	Doğru	94.33	93.31	94.30	94.03	**93.44**	93.68
		Fazla	5.53	6.54	5.11	5.39	4.41	4.18
		Eksik	12.84	11.45	15.34	14.56	19.16	19.50
	2	Doğru	**83.58**	83.10	81.09	80.83	77.29	76.89
		Fazla	3.59	5.45	3.58	4.61	3.55	3.61
		Eksik	39.19	33.26	40.29	37.16	41.99	42.01
	3	Doğru	58.65	**62.91**	57.18	59.43	55.24	54.90
		Fazla	2.16	3.83	2.54	3.41	2.78	3.09
100	0	Doğru	93.85	93.85	95.13	95.13	**96.03**	**96.03**
		Fazla	6.15	6.15	4.88	4.88	3.98	3.98
		Eksik	0.01	0.01	0.01	0.01	0.16	0.16
	1	Doğru	94.83	94.04	95.26	94.99	95.73	**95.88**
		Fazla	5.16	5.95	4.73	5.00	4.11	3.96
		Eksik	5.25	4.79	6.59	6.14	8.68	8.75
	2	Doğru	**90.74**	89.86	89.39	89.11	87.35	87.13
		Fazla	4.01	5.35	4.03	4.75	3.98	4.13
		Eksik	22.68	18.91	23.76	21.29	25.45	24.96
	3	Doğru	74.61	**76.84**	73.36	74.90	71.44	71.56
		Fazla	2.71	4.25	2.88	3.81	3.11	3.48

Çizelge 2. 4 Klasik test istatistikleri $(\alpha = 0.05)$ anlamlılık seviyesinde, Paralel Analiz (PA) ile %95 Bootstrap Güven Aralığı ile Modifiye Paralel Analiz (BPA) yöntemine göre boyut sayısı belirleme için göreli frekanslar (Devamı)

			[b]PA dayalı yöntemler			
			Paralel Analiz		Bootstrap Modifiyeli Paralel Analiz	
N	*d*	Belirlenen *d(%)*	[7] PA_{95}	[8] PA_{99}	[9] BPA_{95}	[10] BPA_{99}
25	0	Doğru	94.75	**98.88**	83.70	97.50
		Fazla	5.25	1.20	16.30	2.50
		Eksik	47.46	69.88	30.48	67.25
	1	Doğru	**37.18**	25.75	31.93	23.11
		Fazla	15.36	4.38	37.59	9.64
		Eksik	54.95	76.48	34.93	67.95
	2	Doğru	24.76	16.71	**34.53**	22.04
		Fazla	20.29	6.81	30.55	10.01
		Eksik	62.51	82.78	56.75	82.06
	3	Doğru	20.15	12.28	**31.60**	15.24
		Fazla	17.34	4.95	11.65	2.70
50	0	Doğru	95.13	98.85	94.18	**99.50**
		Fazla	4.88	1.15	5.83	0.50
		Eksik	3.38	8.98	4.83	16.53
	1	Doğru	67.53	**77.93**	63.09	73.04
		Fazla	29.10	13.10	32.09	10.44
		Eksik	9.70	20.06	11.63	26.10
	2	Doğru	46.20	55.19	57.09	**62.29**
		Fazla	44.10	24.75	31.29	11.61
		Eksik	15.56	30.55	32.38	54.25
	3	Doğru	41.99	46.09	**54.74**	41.98
		Fazla	42.45	23.36	12.89	3.78
75	0	Doğru	94.93	98.90	95.30	**99.58**
		Fazla	5.08	1.10	4.70	0.43
		Eksik	0.13	0.48	0.88	2.34
	1	Doğru	68.45	84.16	66.69	**86.66**
		Fazla	31.43	15.36	32.44	11.00
		Eksik	1.94	5.06	3.63	11.59
	2	Doğru	48.75	64.36	61.86	**73.88**
		Fazla	49.31	30.58	34.51	14.54
		Eksik	4.09	9.55	14.03	29.30
	3	Doğru	42.90	56.24	**67.01**	64.04
		Fazla	53.01	34.21	18.96	6.66

Çizelge 2. 4 Klasik test istatistikleri $(\alpha = 0.05)$ anlamlılık seviyesinde, Paralel Analiz (PA) ile %95 Bootstrap Güven Aralığı ile Modifiye Paralel Analiz (BPA) yöntemine göre boyut sayısı belirleme için göreli frekanslar (Devamı)

			[b]PA dayalı yöntemler			
			Paralel Analiz		Bootstrap Modifiyeli Paralel Analiz	
N	*d*	Belirlenen *d(%)*	[7] PA_{95}	[8] PA_{99}	[9] BPA_{95}	[10] BPA_{99}
100	0	Doğru	94.90	98.65	95.60	**99.58**
		Fazla	5.10	1.35	4.40	0.43
		Eksik	0.01	0.03	0.60	0.73
	1	Doğru	68.83	84.80	67.05	**88.45**
		Fazla	31.16	15.18	32.35	10.83
		Eksik	0.46	1.24	1.39	4.10
	2	Doğru	49.25	67.05	62.08	**80.16**
		Fazla	50.29	31.71	36.54	15.74
		Eksik	1.14	3.40	5.54	14.43
	3	Doğru	43.88	58.84	72.86	**77.36**
		Fazla	54.99	37.76	21.60	8.21

Not: *N*: örnek sayısı, *d*: boyut sayısı, [a] χ^2 Yaklaşım prosedürleri $\alpha = .05$ güven seviyesinde hesaplanmıştır, [b]En iyi performansı gösterenler bolt olarak gösterilmiştir,

[1] $S_B\left(T_d^2\right)$: T_d^2'ye dayalı Bartlett χ^2 yaklaşım prosedürü,

[2] $S_F\left(T_d^2\right)$: T_d^2'ye dayalı Fujikoshi χ^2 yaklaşım prosedürü,

[3] $S_B(\Lambda_d)$: Λ_d'ye dayalı Bartlett χ^2 yaklaşım prosedürü,

[4] $S_L(\Lambda_d)$: Λ_d'ye dayalı Lawley χ^2 yaklaşım prosedürü,

[5] $S_B(V_d)$: V_d'ye dayalı Bartlett χ^2 yaklaşım prosedürü,

[6] $S_F(V_d)$: V_d'ye dayalı Fujikoshi χ^2 yaklaşım prosedürü,

[7] PA_{95} : PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 95. persentil eşik değeri,

[8] PA_{99} : PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 99. persentil eşik değeri,

[9] BPA_{95} : BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 95. persentil eşik değeri,

[10] BPA_{99} : BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 99. persentil eşik değeri

Örneklem büyüklüğü $N = 25$ olduğunda sırasıyla boyut d =0, 1, 2 ve 3 için boyut $d = 0$ yani (tüm kanonik korelasyonlar sıfır) olduğunda bütün yöntemler benzer sonuçlar vermiştir. Boyut sayısı $d = 0$ olduğunda yöntemlerin doğru boyut tespit etme yüzdeliklerinin ortalamasına göre, en yüksek performansa sahip olan PA_{99}, 98.88 değeri ile, test istatistiklerinden $S_B(V_d)$ ve $S_F(V_d)$ 98.85 değeri ile iyi performans göstermektedirler. Boyut sayısı $d = 1$ olduğu durumda $S_B\left(T_d^2\right)$ testinin, 66.48 değeri ile en yüksek değere sahip olduğu, PA ve BPA yöntemlerinin performansı test istatistiklerine göre düşük, fakat test istatistikleri ile karşılaştırabilir olduğu görülmektedir. Boyut sayısı $d = 2$ olduğu duruma bakıldığında PA prosedürleri içinde

34.52 en yüksek değeri ile BPA_{95}'nın performansı ile karşılaştırıldığında, test istatistiklerinden 31.52 değeri ile $S_F\left(T_d^2\right)$'nın performansının kısmen daha düşük olduğu görülmektedir. Boyut sayısı $d=3$ olduğu duruma bakıldığında $S_F\left(T_d^2\right)$'nın perfomansının 10.91 olduğu; 31.6 değeri ile BPA_{95}'nın test istatistikleri içinde en yüksek değere sahip olan test istatistiğinden yaklaşık 3 kat daha doğru sonuç bulduğu görülmektedir. Küçük örnekler için sonuçlara bakıldığında PA ve BPA'nın 95. persentil değerinin kullanımı önerilebilir. Altı test istatistiği arasındaki performansa bakıldığında $S_F\left(T_d^2\right)$ ve $S_B\left(T_d^2\right)$ testlerinin başarı oranlarının geriye kalan test istatistiklerine göre daha iyi olduğu görülmektedir.

Örneklem büyüklüğü $N=50$ için sırasıyla boyut sayısının $d=$0, 1, 2 ve 3 olma koşuları incelendiğinde, doğru boyut sayısı $d=0$ olduğunda, tüm yöntemlerin 92 ve üzeri sonuçlar ile çok iyi performans gösterdikleri, BPA_{99}'nın 99.3 değeri ile en yüksek yüzdeye ulaştığı görülmektedir. $d=1$ boyutlu olma koşulunda bakıldığında test istatistiklerinden $S_B\left(T_d^2\right)$, 91.67 değeri ile en iyi sonucu verdiği görülmektedir. Doğru boyut sayısı $d=2$ olduğu koşulda ise test istatistiklerinde $S_F\left(T_d^2\right)$, 69.08 değeri ile en yüksek olup, BPA_{99}'nın 60.28 ile PA'ya oranla daha iyi sonuç verdiği görülmektedir. $d=3$ boyutlu duruma bakıldığında $S_F\left(T_d^2\right)$, BPA_{99} ve PA_{99} yöntemleri birbirine yakın sonuçlar bulmuştur. PA_{99}, 43.86 değeri ile tüm yöntemler arasında en iyi sonucu vermiştir. $N=25$ birimlik örneklere göre $N=50$ birimlik örneklerde genel olarak tüm yöntemlerin performanslarında iyileşme görülmektedir. $N=50$ birimlik örneklere bakıldığında genel anlamda hem PA hem de BPA yöntemlerinin iyi sonuçlar verdiği görülmektedir. PA hem de BPA yöntemlerinin 99. persentil değerinin kullanımı önerilebilir.

Örneklem büyüklüğü $N=75$ için sonuçlar sırasıyla $d=$0, 1, 2 ve 3 boyutlu olma koşuları için incelendiğinde, nispeten tüm yöntemlerin performanslarının yükseldiği görülmektedir. Doğru boyut sayısı $d=0$ olma koşulu altında en yüksek performansa sahip olan yöntem BPA_{99} (99.575) olduğu görülmektedir. Doğru boyut sayısının $d=1$ bakıldığında en yüksek $S_B\left(T_d^2\right)$ (94.32) iken, BPA_{99} 'nın (86.66) başarı değeri ile

özellikle $d=1$ boyutlu olma koşulu altında diğer örnek büyüklüklerine göre performansının artığı görülmektedir. Doğru boyut sayısının $d=2$ olma koşulunda ise test istatistiklerinde $S_F\left(T_d^2\right)$ (83.57) en yüksek değere sahipken, BPA_{99} (73.87) kabul edilebilir bir performans sergilemektedir. $d=3$ boyutlu duruma bakıldığında tüm yöntemler arasında BPA_{99}, 64.03 değeri ile en iyi performansı göstermektedir. Ayrıca $S_F\left(T_d^2\right)$ ve $S_B\left(T_d^2\right)$ test istatistiklerin yanısıra BPA_{99} ve PA_{99} yöntemlerinin kullanımı önerilebilir.

Örneklem büyüklüğü $N=100$ için sırasıyla $d=$0, 1, 2 ve 3 boyutlu olma koşuları incelendiğinde, doğru boyut sayısının $d=0$ olma koşulu altında en yüksek performansa sahip olan yöntemin BPA_{99} (99.57) olduğu görülmektedir. $d=1$ boyutlu duruma bakıldığında $S_B\left(T_d^2\right)$ (95.77) değeri ile en iyi sonucu vermiştir. $d=2$ boyutluda ise $S_F\left(T_d^2\right)$ 90.73 değeri en iyi sonucu vermiştir. $d=3$ boyutlu duruma bakıldığında BPA_{99} 77.36 değeri ile en uygun sonuca ulaşmıştır. Buna göre bakıldığında hem PA_{99} hem de BPA_{99} kullanımı önerilebilir.

Çizelge 2. 5 Klasik test istatistikleri $(\alpha = 0.05)$ anlamlılık seviyesinde, Paralel Analiz (PA) ile %95 Bootstrap Güven Aralığı ile Modifiye Paralel Analize (BPA) göre kanonik korelasyon değerinin 0.5'ten büyük ve 0.5'ten küçük-eşit olmasına göre boyut belirleme için göreli frekanslar

			[a,b] χ^2 Yaklaşım Prosedürleri					
			T_d^2 dayalı yöntemler		Λ_d dayalı yöntemler		V_d dayalı yöntemler	
	d	Belirlenen *d* (%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	[c] $\rho_i > .5$	Eksik	39.49	36.13	52.06	51.56	81.23	82.53
		Doğru	57.82	**58.03**	45.20	45.21	17.06	16.39
		Fazla	2.69	5.84	2.74	3.23	1.72	1.08
	[d] $\rho_i \leq .5$	Eksik	92.85	86.38	92.18	90.76	93.50	95.19
		Doğru	6.88	**12.51**	7.34	8.48	5.83	4.31
		Fazla	0.27	1.12	0.48	0.77	0.67	0.50
50	[c] $\rho_i > .5$	Eksik	4.33	3.89	8.19	8.00	21.25	21.95
		Doğru	**91.07**	89.61	87.79	87.34	75.48	75.08
		Fazla	4.61	6.50	4.02	4.66	3.27	2.98
	[d] $\rho_i \leq .5$	Eksik	62.23	54.98	64.23	61.35	67.77	69.17
		Doğru	36.51	**42.18**	34.21	36.32	30.52	28.98
		Fazla	1.27	2.85	1.56	2.33	1.72	1.85
75	[c] $\rho_i > .5$	Eksik	0.19	0.17	0.58	0.56	1.97	2.03
		Doğru	**94.68**	93.44	94.60	94.10	93.66	93.88
		Fazla	5.13	6.39	4.82	5.34	4.38	4.10
	[d] $\rho_i \leq .5$	Eksik	34.59	29.74	36.89	34.32	40.23	40.42
		Doğru	63.03	**66.11**	60.44	62.08	56.98	56.43
		Fazla	2.38	4.15	2.67	3.60	2.78	3.15
100	[c] $\rho_i > .5$	Eksik	0.01	0.01	0.02	0.02	0.13	0.13
		Doğru	95.12	94.21	95.43	95.05	95.72	**95.87**
		Fazla	4.88	5.78	4.56	4.93	4.15	4.00
	[d] $\rho_i \leq .5$	Eksik	18.62	15.80	20.23	18.28	22.73	22.45
		Doğru	78.33	**79.62**	76.58	77.62	73.96	73.84
		Fazla	3.05	4.58	3.19	4.11	3.32	3.71

Çizelge 2. 5 Klasik test istatistikleri $(\alpha = 0.05)$ anlamlılık seviyesinde, Paralel Analiz (PA) ile %95 Bootstrap Güven Aralığı ile Modifiye Paralel Analize (BPA) göre kanonik korelasyon değerinin 0.5'ten büyük ve 0.5'ten küçük-eşit olmasına göre boyut belirleme için göreli frekanslar (Devam)

			[b] PA dayalı yöntemler			
			Paralel Analiz		Bootstrap Modifiyeli Paralel Anali	
N	*d*	Belirlenen *d* (%)	[7] PA_{95}	[8] PA_{99}	[9] BPA_{95}	[10] BPA_{99}
25	[c] $\rho_i > .5$	Eksik	32.96	67.10	27.51	61.65
		Doğru	**36.46**	26.17	35.65	26.91
		Fazla	30.58	6.73	36.84	11.44
	[d] $\rho_i \leq .5$	Eksik	66.99	85.65	55.27	83.19
		Doğru	18.27	10.33	**29.72**	13.35
		Fazla	14.74	4.03	15.02	3.46
50	[c] $\rho_i > .5$	Eksik	2.43	6.59	3.69	12.48
		Doğru	61.35	74.20	61.88	**74.93**
		Fazla	36.22	19.21	34.43	12.59
	[d] $\rho_i \leq .5$	Eksik	16.66	33.13	28.86	53.43
		Doğru	42.46	45.27	**54.73**	41.94
		Fazla	40.88	21.60	16.41	4.63
75	[c] $\rho_i > .5$	Eksik	0.08	0.12	0.81	1.83
		Doğru	61.78	78.19	64.08	**84.55**
		Fazla	38.14	21.69	35.12	13.62
	[d] $\rho_i \leq .5$	Eksik	4.02	9.74	11.54	26.98
		Doğru	44.96	58.32	**66.30**	65.17
		Fazla	51.03	31.94	22.16	7.85
100	[c] $\rho_i > .5$	Eksik	0.01	0.02	0.01	0.03
		Doğru	61.98	78.62	65.01	**85.89**
		Fazla	38.02	21.37	34.98	14.08
	[d] $\rho_i \leq .5$	Eksik	1.07	15.62	4.44	12.18
		Doğru	45.99	78.33	70.22	**78.09**
		Fazla	6.05	52.94	25.34	9.73

Not: [a, b, 1,2,3,4,5,6,7,8,9,10] İçin açıklamalar Çizelge 2. 4'te verilmiştir

[c]Tüm anakütle modellerinde kanonik korelasyonların 0.5 üstünde olduğu ($\rho_i > 0.5\,(i = 1,\ldots,d)$).

[d]Tüm anakütle modellerinde kanonik korelasyonların 0.5 eşit veya en az birinin küçük olduğu ($\rho_i \leq .5\,(i = 1,\ldots,d)$).

Çizelge 2.5'te tüm modellerde kanonik korelasyon değerlerinin tümünün 0.5'ten büyük, kanonik korelasyon değerlerinin tümünün 0.5'ten küçük veya en az birinin eşit olması koşulları altında simülasyon sonuçlarının ortalamaları alınarak verilmiştir. Simülasyon sonuçlarının klasik test istatistikleri, paralel analiz (PA) ile %95 bootstrap güven aralığı ile modifiye paralel analiz (BPA) yöntemine göre boyut seçimi için elde edilen sonuçlar incelendiğinde, örneklem sayısı $N = 25$ olduğunda birinci koşul altında (tüm $\rho > 0.5$) klasik test istatistiklerinden $S_F\left(T_d^2\right)$, 41.5 ile en iyi performansı göstermiştir. İkinci koşul altında (en az bir tane $\rho \leq 0.5$) için sonuçlara bakıldığında en iyi sonucu BPA_{99}, 23.02 değeri ile göstermiştir. Burada özellikle ikinci koşul altında (en az bir tane $\rho \leq 0.5$) BPA yöntemi test istatistiklerine göre yaklaşık 5 katı daha iyi doğru boyut sayısını hesapladığı görülmektedir.

Örneklem büyüklüğü $N = 50$ için bakıldığında birinci koşul altında (tüm $\rho > 0.5$) klasik test istatistiklerinde en yüksek performans sahip test 74.81 değeri ile $S_F\left(T_d^2\right)$'dir. İkinci koşul altında (en az bir tane $\rho \leq 0.5$) için sonuçlara incelendiğinde ise BPA_{95}'nın 42.42 değeri ile en iyi performansa sahip olduğu görülmektedir.

Örneklem büyüklüğü $N = 75$ için incelendiğinde birinci koşul altında (tüm $\rho > 0.5$) klasik test istatistiklerinde en yüksek performansa sahip olan test 86.74 değeri ile $S_B\left(T_d^2\right)$ olup BPA_{99} 84.85 değeri ile önemli bir sonuca ulaşmıştır. İkinci koşul altında (en az bir tane $\rho \leq 0.5$) için sonuçlara bakıldığında BPA_{95}'nın 61.6 değeri ile en iyi sonuca sahip olduğu görülmektedir.

Örneklem büyüklüğü $N = 100$ için bakıldığında birinci koşul altında (tüm $\rho > 0.5$) klasik test istatistiklerinde en iyi performansa sahip olan test $S_B\left(T_d^2\right)$ (92.22) olduğu görülmektedir. İkinci koşul altında (en az bir tane $\rho \leq 0.5$) için sonuçlara bakıldığında ise BPA_{95}'nın 73.27 değeri ile en iyi performansa sahip olduğu görülmektedir. Kanonik korelasyon değerlerinin en az birinin 0,5'in altında veya eşit olduğunda boyut sayısına karar verilmesinde, özellikle küçük örneklemlerde test istatistiklerine nazaran klasik PA ve Bootstrap modifiyeli PA'nın çok daha iyi olduğu sonucu çıkarılabilir.

Özetlenecek olursa boyut seçimi için özellikle kanonik korelasyonların en az birinin 0.5 ten küçük veya eşit olduğunda PA ve BPA'nın klasik test istatistiklerinden daha iyi sonuç verdiği görülmektedir. Kanonik korelasyonların tümünün 0.5 ten büyük

olduğunda PA ve BPA'nın klasik test istatistikleri ile karşılaştırılabilir olduğu görülmektedir. Bununla beraber, kanonik korelasyonda boyut seçimi için modifiye edilen PA'nın özellikle küçük örneklerdeki performansı dikkat çekmektedir. Yukarıda verilen anakütle modelleri için yapılan simülasyon çalışmasının sonuçlarına göre, PA yaklaşımı küçük örneklerde, anakütledeki boyut sayısını, klasik test yöntemlerine göre %15'den fazla doğrulukta belirlemektedir. Bootstrap güven aralığı ile modifiye edilen Paralel Analizde küçük örneklerde 95. persentil değerlerinin, diğer örneklerde ise 99. persentil değerlerinin eşik değer olarak kullanılması önerilmektedir. Ayrıca bütün kanonik korelasyonların sıfır olduğu anakütle için hem klasik PA ile hem de modifiye PA ile bütün örnek büyüklüklerinde ve özellikle 99. ve 95. persentil değerlerinde boyut sayısı en yüksek doğrulukta belirlenmektedir. Bu çalışmada simülasyon sonuçları incelendiğinde, modifikasyonun PA'nın performansını arttırdığı görülmektedir. Simülasyon sonuçlarına göre test istatistiklerinin performansına bakıldığında küçük örneklemlerde gerçek boyut sayısı $d = 0$ olduğu durum dışında düşük olduğu, büyük ve orta büyüklükteki örneklemlerde ise performanslarının örnek büyüklüğü artıkça iyileştiği görülmektedir. Gerçek boyut sayısı $d = 0{,}1$ ve 2 olduğunda test istatistiklerinin performansları orta büyüklükteki örneklemlerde yeterli olup, $d = 3$ boyutlu olduğunda ise örneklem büyüklüğü artıkça daha iyi sonuçlar vermişlerdir. Ayrıca, özellikle kanonik korelasyonların en az birinin 0,5 ten küçük olduğunda koşulda test istatistiklerinin performanslarının düştüğü görülmüştür. Genel olarak altı test istatistiği kendi aralarındaki performansa bakıldığında Fujikoshi'nin $S_F\left(T_d^2\right)$ ve Bartlett'in $S_B\left(T_d^2\right)$ testlerinin başarı oranı diğer test istatistiklerine göre daha iyi olduğu söylenilebilir.

Sonuç olarak PA ve BPA yöntemleri KKA'da boyut sayısına karar verilmesinde alternatif yöntemler olarak önerilebilir. PA yönteminin faktör ve temel bileşen analizi yöntemlerindeki başarılı performansı önceki çalışmalarla ortaya konulmuştu, ayrıca bu çalışmada kanonik korelasyon analizinde boyut sayısı belirlemede performansının da iyi olduğu sonucu elde edilmiştir. Bu çalışma ile PA'nın ve bootstrap modifiyeli PA'nın KKA'daki performansı kabul edilebilir seviyede bulunuş olup, faktör ve temel bileşen analizi haricindeki diğer çok değişkenli yöntemlerdeki boyut sayısı probleminin çözümüne de katkısının olabileceği düşünülmektedir.

BÖLÜM 3

FAKTÖR ANALİZİNDE BOYUT SAYISININ BELİRLENMESİNDE PARALEL ANALİZ ve BOOTSTRAP MODİFİYELİ PARALEL ANALİZ YAKLAŞIMLARI

Bu bölümde, faktör analizinde boyut (faktör) sayısını belirlemede paralel analiz (PA), bootstrap modifiyeli paralel analiz (BPA) ve klasik yöntemlerden G-K Kuralı üç faklı gerçek veride incelenmiş ve simülasyon çalışmalarıylada sonuçlar desteklenmiştir. Faktör Analizinde PA faktör sayısına karar verilmesinde kullanılan yöntemlerden biridir. Araştırmalarda faktör analizinde boyut sayısı belirlenmesinde PA yöntemi diğer kullanılan yöntemlerden daha iyi performans göstermektedir. Bu bölümde faktör analizinde boyut sayısı belirlemede paralel analizin ortalama, medyan, 90., 95., 97.5 ve 99. persentil değerleri kullanılmıştır. Ayrıca bootstrap modifiyeli paralel analiz (BPA) analiz yöntemi oluşturulmuş ve ilk defa faktör analizinde boyut sayısını belirlemede kullanılmıştır. Boyut sayısı belirlemede BPA'nın da ortalama, medyan, 90., 95., 97.5 ve 99. persentil değeri kullanılmıştır. Faktör analizinde boyut sayısını belirlemede yukarıda belirtilen yöntemlerin performansları üç gerçek veri ve simülasyon çalışmalarıyla ayrıntılı olarak incelenmiştir.

Faktör analizi, veri matrisinin temelini oluşturan yapıyı tanımlamayı amaç edinen ve temel işlevi dışında birçok çok değişkenli istatistiksel yöntemin uygulamasında önemli roller üstlenen çok değişkenli istatistiksel bir yöntemdir. Aralarında ilişki bulunan p sayıdaki değişkenle açıklanan bir yapıyı, kendi içlerinde ilişkili; ancak aralarında ilişkisiz daha az sayıdaki $(m < p)$ yeni değişkenlerle açıklamaya yarayan yöntemdir. Faktör analizinin iki temel amacı vardır. Birincisi değişken sayısını azaltmak (boyut indirgemek), ikincisi ise değişkenleri sınıflamak yani değişkenler arasındaki yapıyı kovaryans matrisini bulmaktır. Faktör analizi ile ilgili ilk çalışmaların Charles Spearman

tarafından yapıldığı bilinir. Spearman (1994) çalışmasında iki faktör teorisini ortaya atmıştır. Daha sonra Charles Spearman, Karl Pearsen ve Louis Thurstone faktör analizinin gelişmesinde, gizil (soyut) kavramların ölçülmesi ile ilgili çalışmalarla öncülük etmişlerdir [11, 23, 57].

3.1 Faktör Analizi Modeli

$Y'_{1xp} = [y_1, y_2, \ldots, y_p]$ rasgele değişkenler ortalama vektörü μ ve varyans-kovaryans matrisi Σ olsun. Her bir p değişken vektörü için,

$$Y_{1xp} - \mu_{px1} = \Lambda_{pxm}\, f_{mx1} + \varepsilon_{px1} \tag{3.1}$$

faktör modeli için $m < p$ olmak üzere, temel bileşenler analizinde $m \leq p$'dir. $y_1, y_2, \ldots, y_p$ değişkenlerinin $f_1, f_2, \ldots, f_m$ faktörlerinin lineer kombinasyonu olduğu varsayılır.

Faktör modeli 3.1' in açık hali,

$$\begin{aligned} y_1 - \mu_1 &= \Lambda_{11} f_1 + \Lambda_{12} f_2 + \cdots + \Lambda_{1m} f_m + \varepsilon_1 \\ y_2 - \mu_2 &= \Lambda_{21} f_1 + \Lambda_{22} f_2 + \cdots + \Lambda_{2m} f_m + \varepsilon_2 \\ &\vdots \\ y_p - \mu_p &= \Lambda_{p1} f_1 + \Lambda_{21} f_2 + \cdots + \Lambda_{pm} f_m + \varepsilon_p \end{aligned} \tag{3.2}$$

olup Λ_{ij}: faktör yükleri ve f_j: faktörleri göstermektedir. (3.2)'deki modelin matris formunda $y - \mu = \Lambda f + \varepsilon$ biçiminde yazılabilir.

Burada $Z = y - \mu$ alınırsa,

$$Z = \Lambda f + \varepsilon \tag{3.3}$$

formunda yazılabilir. Varyans-kovaryans matrisi $COV(Z) = COV(\Lambda f + \varepsilon) = \Sigma$ elde edilir. Burada $\mathrm{E}(Z) = \mathrm{E}(\Lambda f + \varepsilon) = \mathrm{E}(f) = \mathrm{E}(\varepsilon) = 0$'dır. Ayrıca $\mathrm{E}(\varepsilon) = 0$ ve $Z = \Lambda f + \varepsilon$ regresyon modeline benzemekte fakat f faktörlerin doğrudan gözlemlenememesi açısından regresyondan ayrılmaktır.

Faktör analizinin varsayımları,

- $rank(\Lambda) = r < p$ (faktör sayısı değişken sayısından küçüktür.)

- $E(Z / f) = \Lambda f$
- $E(ff') = \Omega$, $\psi = E(\varepsilon\varepsilon')$ yani $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$'dır. $E(ff') = I$ başlangıçta faktörler ortonormal vektörlerdir.
- $E(f\varepsilon') = 0$ yani kalıntılar ve faktörler korelasyonsuz, $\text{cov}(f, \varepsilon) = 0$'dır. Burada

$$\begin{aligned} E(ZZ') &= E((\Lambda f + e)(\Lambda f + e)') \\ &= \Lambda\Omega\Lambda' + E(\varepsilon\varepsilon') \\ &= \tau + \psi \end{aligned} \tag{3.4}$$

faktörler $\tau = \Lambda\Omega\Lambda'$ ile açıklanan varyansı ve $\psi = E(\varepsilon\varepsilon')$ kalıntıların varyans-kovaryans matrisi, değişkenlerin faktörleri ile açıklanmayan varyansıdır. Faktörler korelasyonsuz olduğunda $\text{COV}(f) = I$ birim matrise eşit olur. Eğer faktörler birbirine dik ise, değişkenler ve faktörler arasındaki korelasyonlar Λ faktör yüklerine eşit olur.
Değişkenlerin varyansı,

$$var(y_i) = \Lambda_{i1}^2 + \Lambda_{i2}^2 + \cdots + \Lambda_{im}^2 + \varepsilon_i \tag{3.5}$$

olur. Faktör analizinin temeli, $E(ZZ') = \sum = \tau + \psi$ formülüne dayanmaktadır. Varyans-kovaryans matrisinin köşegen elamanlarının karekökü standart sapmalar matrisi oluşturulur. S standart sapmalar matrisini göstermek üzere korelasyon matrisi,

$$R = S^{-1} * \Sigma * S^{-1} \tag{3.6}$$

ile elde edilebilir. Faktör analizinde, temel problem kovaryans veya korelasyon matrisinden faktör yükleri matrisinin elde edilmesidir [35].
Faktör analizinde verilerin oran veya aralık ölçeğinde olması istenir. Örneklem büyüklüğünün makul bir seviyede olması istenir ve küçük örnekler ilişki katsayılarının güvenilirliklerini olumsuz yönde etkilemektedir. Araştırmacılar, örnek büyüklüğünün en az gözlenen değişken sayısının beş katı olmasını tavsiye etmişlerdir. Eğer az sayıda belirgin faktör varsa örneklem büyüklüğü N=50 olarak belirlenebilir. Faktör modelinde değişkenlerin ve bu değişkenlerin doğrusal kombinasyonlarının çok değişkenli normal dağılım gösteren bir anakütleden çekilmesi olmasını varsayar. Ayrıca normallik varsayımı değişken çiftleri arasındaki ilişkinin doğrusal olduğunu gösterir. Korelasyon matrisinde en az birkaç değişken arasınadaki korelasyon belirli bir büyüklükte, örneğin

en az 0.30 olmalıdır. Eğer bu korelasyon düşük ise uygun faktör sayısına ulaşmak mümkün olmamaktadır [23, 58].

3.2 Boyut Sayısı Belirleme Yöntemleri

Bu çalışmada, Faktör analizinde boyut sayısının belirlenmesinde G-K kuralı, Paralel analiz ve Bootstrap modifiyeli paralel analiz yöntemleri kullanılmıştır.

3.2.1 G-K Kuralı

Kaiser [14] ve Guttman [15] tarafından önerilen G-K (öz değer >1) kuralı faktör analizinde yaygın olarak kullanılmaktadır. Bu yöntem, bir faktörün açıklayıcılığının en azından bir değişkenin açıklayıcılığı kadar olması mantığına dayanır. Bu ölçütte, özdeğeri 1'den büyük olan faktörler önemli faktörler olarak dikkate alınırken, özdeğeri 1'den küçük olan faktörler ise önemsiz faktörler sayılarak dikkate alınmazlar. G-K kuralı, Cronbach α katsayısını esas alan bir yöntemdir. Cronbach α katsayısı değişkenlerin güvenilirliğini hesaplamak için kullanılan bir yöntemdir.

$$r_{yy'} = \frac{p}{(p-1)}.\{(S_y^2 - \sum_i S_i^2)/S_y^2\} \tag{3.7}$$

burada y'nin transpozu y' ile arasındaki korelasyon, değişken sayısı p, y'nin varyansı S_y^2 ve toplam varyans $\sum_i S_i^2$'ye bağlıdır. Kaiser [14] bu formül ile faktör sayısının belirlenebileceği belirtilmiştir. Bir temel bileşeninin güvenilirliği,

$$r_{xx'} = (\frac{p}{(p-1)}).(\frac{\lambda_x - 1}{\lambda_x}) \tag{3.8}$$

formülü ile hesaplanabilir. λ_x bileşenin özdeğeri olmak üzere, $\lambda_x < 1$ olduğunda güvenilirlik negatif çıkacaktır dolayısıyla bileşen olarak belirlenemeyecektir [17, 59].

3.2.2 Paralel Analiz

Horn [7] tarafından önerilen bu yöntem faktör sayısını belirlemek için rasgele sayı üretilmesine dayalı bir yöntemdir. PA Monte Carlo simülasyonu kullanılarak orjinal veri setine paralel rasgele veri seti üretilmesine dayanmaktadır. G-K kuralında bir biri ile

ilişkisiz değişkenlerin korelasyon matrisinden elde edilen anakütle özdeğerleri ile örneklemden elde edilen özdeğerler karşılaştırlır. Guttman [15], anakütle modeli tam olarak bilindiğinde, anakütle korelasyon matrisinin rankı için alt sınır, sadece 1'den büyük özdeğerlerden oluştuğunu göstermiştir. Horn [7], bir biri ile ilişkisiz değişkenlerin örnek korelasyon matrisinin başlangıçta 1'den büyük olan özdeğerlerin, örneklem hatasından dolayı finalde elde edilen özdeğerlerin 1'den küçük elde edildiğini göstermiştir. Horn [7] özellikle küçük örneklerde faktörlerin açıklanan varyansı maksimum yapacak şekilde bulunmasına zorlanmasından dolayı, başlangıç faktörlerinin, örneklerde, anakütlede olduğundan daha yüksek özdeğerlere sahip olduklarını saptayarak, PA'yı geliştirmiştir. Horn tarafından önerilmiş olan bu teknikte n×p orijinal veri setiyle aynı sayıda değişken ve gözlem içerecek şekilde, kısaca veri setine paralel, birim kovaryans matrisli çok değişkenli normal dağılımdan n×p boyutlu rastgele veri setleri üretildikten sonra, orijinal veri setine ait korelasyon matrisinden elde edilen özdeğerler ile, rastgele veri setlerinden elde edilen korelasyon matrislerinin özdeğerlerinin ortalaması karşılaştırılarak, örneğin i. sıradaki özdeğerleri, i.ortalama özdeğerden büyük olduğunda karşılık gelen temel bileşen veya faktörün analizde kalmasına karar verilmektedir. PA'da özdeğerleri hesaplamak için tekrar (iterasyon) sayısının ne kadar olacağı ile ilgili belirli bir kural yoktur. Horn [7] özdeğereleri hesaplamak için yapılan tekrar sayısının makul seviyede olmasını önermiştir. Çoğu araştırmacı tekrar sayısının 500 ile 1000 arasında olmasını önermişlerdir [4, 6, 8].

3.3 Bootstrap Metodu

Temel olarak bootstrap, $X_1, X_2, \cdots, X_n$ birbirinden bağımsız aynı dağılımına sahip F fonksiyonundan yeniden örnekleme ile elde edilen $X_1^*, X_2^*, \cdots X_n^*$'nin F_n gözleme dayalı dağılım fonksiyonunun oluşturulmasıdır.

Ampirik dağılım fonksiyonu (ADF) ve Plug-in prensibi

$x_1, x_2, \cdots, x_n$ olasılık dağılımı olan F'den çekilmiş olsun. Burada her bir x kesikli tesadüfi değişken olsun ve aralığı $\{x_i, 1 \le i \le n\}$ olduğu varsayılsın. Bu durumda ADF, bu tesadüfi değişkenin kümülatif dağılım fonksiyonudur ve

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(X_i \le x)} \tag{3.9}$$

şeklinde ifade edilir. ADF, $x_i, i = 1, 2, \cdots, n$ her biri için $\frac{1}{n}$ olasılık değerine sahip keşikli bir dağılımdır. Burada $1_{(X_i \le x)}$ notasyonu aşağıdaki şekilde açılanabilir. R'nin herhangi alt kümesinden biri A olsun. Bu durumda

$$1_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \tag{3.10}$$

olur. Böylece A kümesinin en büyük değeri 1, en küçük değeri ise 0'dır [55].

3.3.1 Bootstrap Persentillerine Dayalı Güven Aralıkları

θ parametresine ilişkin plug-in kestiricisi $\hat{\theta}$ ve standart hatası $s\hat{e}$ ile gösterildiğinde, standart güven aralığı,

$$\left[\hat{\theta} - z^{(1-\alpha)}.s\hat{e}, \hat{\theta} - z^{(\alpha)}.s\hat{e}\right] \tag{3.11}$$

aralığın uç noktaları bootstrapa dayanan, $\hat{\theta}^*$, $N(\hat{\theta}, s\hat{e}^2)$ dağılımından çekilen bir rasgele değişkeni temsil eder. Bu durumda $\theta_{alt} = \hat{\theta}^{*(\alpha)} = \hat{\theta} - z^{(1-\alpha)}.s\hat{e}$ ve $\theta_{üst} = \hat{\theta}^{*(1-\alpha)} = \hat{\theta} - z^{(\alpha)}.s\hat{e}$ değerlerine sırasıyla $\hat{\theta}^*$'nın $100\alpha.$ ve $100(1-\alpha).$ persentillerine karşılık gelir. x^* bootstrap veri kümesi $\hat{F}$ dağılımından oluşturulduğu ve bootstrap tekrar sayısı olan $\hat{\theta}^* = s(x^*)$ bilindiği varsayılsın ayrıca $\hat{\theta}^*$'ın kümülatif dağılım fonksiyonu $\hat{G}$ olsun. Buradan haraketle $(1-2\alpha)$ yüzde aralığı, $\hat{G}$'nin α ve $(1-\alpha)$ yüzdeleri aşağıdaki,

$$\left[\hat{\theta}_{\%alt}, \hat{\theta}_{\%üst}\right] = \left[\hat{G}^{-1}(\alpha), \hat{G}^{-1}(1-\alpha)\right] \tag{3.12}$$

gibi tanımlanır. Bu durumu genelleştirecek olursak $x^{*1}, x^{*2}, \cdots, x^{*B}$ olmak üzere B tane bağımsız veri kümesi oluşturulduğunu varsayalım. Buradan hareketle, $\hat{\theta}^*(b) = s(x^{*b})$ $b = 1, 2, \cdots, B$ olmak üzere, $(1-2\alpha)$ yüzdelik aralığı

$$\left[\hat{\theta}_{\%alt}, \hat{\theta}_{\%üst}\right] = \left[\hat{G}_B^{-1}(\alpha), \hat{G}_B^{-1}(1-\alpha)\right] \tag{3.13}$$

gibi olur [55].

3.2.3 Paralel Analiz Yönteminin Modifikasyonları

Paralel analiz yönteminin modifikasyonuyla ilgili birçok çalışma mevcuttur. PA yönteminde kullanılan ortalama özdeğerin kullanımı gereğinden çok faktör belirleme (overexraction) problemi ortaya çıkarmıştır. Rasgele veri özdeğerlerinin dağılımının (%95 gibi) belirli bir yüzdesinin kullanılması önerilmektedir. Glorfeld [22], Horn'un PA'sını modifiye ederek ortalama özdeğer kullanımı yerine, büyük sayıda rasgele veri üretmiş ve $(1-\alpha)$ persentil değerlerini örneğin 95. veya 99. persentil değerlerinin kullanılmasını önermiştir. Rastgele veriden elde edilen özdeğerlerin dağılımının %95 ve %99 gibi istenilen belli bir yüzdesinden veya kısaca 95. veya 99. persentilden büyük olduğunda ilgili temel bileşen veya faktörün analize katılmasına karar verilmesi ve eşik değerlerin yüksek persentiller formunda olması desteklenmiştir [12, 22, 26]. Burada amaç boyut sayısının belirlenmesinde I. tip hatanın kontrol altına alınmak istenmesidir [20].

Bu çalışmada, PA'nın, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, ve PA_{90} persentil değerleri kullanılacaktır. Ayrıca yeni tanıştırılacak olan bootstrap modifiyeli paralel analiz modifikasyonunda (BPA)'nın $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, ve BPA_{90} persentil değerleri kullanılacaktır.

3.2.3.1 Bootstrap Modifiyeli Paralel Analiz

Bootstrap yönteminin faktör analizinde boyut sayısının belirlenmesinde kullanımı ile ilgili çalışmalar bulunmaktadır [55]. Bootstrap tekniği ile G-K kuralı birlikte kullanılarak temel bileşen analizi (TBA) ve faktör analizi yöntemlerinde temel bilşen ve faktör sayısına karar vermek için kullanılmıştır. Özdeğerler, bootstrap yöntemi ile çekilerek ve bootstrap %95 güven aralığı oluşturularak, elde edilen bootstrap güven aralığının 1 ile karşılaştırılmasıyla boyut sayısına karar verilmiştir. Bootstrap tahmini ortalamasının standart hatasının 1'e yakın olup olmamasına göre ve güven aralığının üst üste binmiş (overlap) olup olmadığı da incelenmiştir [24, 57, 60, 61, 62].

Bootstrap tahminleme yöntemleri arasında en temel yöntem parametrik olamayan bootstrap tekniğidir. Bu teknikte orjinal veri setinden iadeli olarak B kez çok büyük sayıda örneklem çekilir ve çekilen her bir örneklem sayısı orjinal veri setindeki sayısına eşit olur. İlgilenilen parametre için elde edilen ampirik bootstrap dağılımı ile her bir

örneklemin parametreleri ile karşılaştırılır. Bu tekniğin en büyük avantajı herhangi bir dağılım ve parametrik varsayımların bulunmamasıdır.

Faktör analizinde BPA uygulanırken; paralel analiz bootstrap örneklemesi kullanılarak, boyut sayısına karar verilecek olan örnekten elde edilen özdeğerlerin amprik örnekleme dağılımından anakütle özdeğerleri için persentile dayalı bootstrap güven aralığı hesaplanmış ve tek örneğin özdeğerleri yerine persentillere dayalı bootstrap güven aralıklarının medyanları kullanılmıştır. Faktör analizinde BPA'nın $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, ve BPA_{90} persentilleri kullanılmıştır.

3.4 Gerçek Veri ile Paralel Analiz ve Bootstrap Modifiyeli Paralel Analiz Yaklaşımların Uygulamaları

Faktör analizinde boyut sayısı belirlemede G-K kuralı, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} persentil yöntemlerinin üç gerçek veri setinde uygulanmıştır. Analizler SAS 9.4 programında yapılmıştır. Çizelge 3.1'de gerçek üç veri seti için faktör analizinde boyut sayısı hesaplamada kullanılan yöntemlerin sonuçları verilmiştir. Gerçek veri setleri için aşağıda ayrıntılı bilgiler verilmiştir.

İlk olarak analizde Araba markaları verisi kullanılmıştır. Veri seti, 23 farklı araba modeli ile ilgili 40 kişiye anket yapılmış, veriler elde edilmiştir. Ayrıca bağımsız 8 değişken her bir değişken için 1 (çok iyi)'den 6 (çok kötü) olarak kodlanmıştır. Bağımsız değişkenler, X1: A Ekonomi, X2: B Servis, X3: C Yıpranma payı değeri, X4: D Fiyat, X5: E Dizayn, X6: F Sportif, X7: G Güvenlik ve X8: H Kolay elde edilebilirlik, değişkenlerinden oluşmaktadır [63, s185,344].

İkinci veri seti WISCSEM (Çocuklar için Weschler zekâ ölçeği alt skorları) verisidir. Bu veri seti 11 değişeken ve 177 gözlemden oluşmaktadır. Bağımsız değişkenler, X1: Bilgi, X2: Kapsam, X3: Aritmetik, X4: Benzerlikler, X5: Kelimeler, X6: Sayı dizisi, X7: Resim tamalama, X8: Resim düzenleme, X9: Blok tasarımı, X10: Nesne derleme, X11: Kodlama, değişkenlerinden oluşmaktadır [64].

Son veri seti Boston Housing verileri, veriler 14 değişken ve 506 gözlemden oluşmaktadır. Bağımsız değişkenler, X1: Şehir merkezinde suç oranı, X2: Büyük arazilerde imarlı arsa oranı, X3: Parakende de olmayan iş alanı oranı, X4: Charles River

(eğer nehrin sınırındaysa 1, diğer bölge 0), X5: Nitrik osit konsantrasyonu, X6: Konut başına ortalama oda sayısı, X7: 1940 öncesi inşa edilen mülk sahibi birimleri oranı, X8: Boston'daki beş istihdam merkezinin ağırlıklı mesafeleri, X9: Radyal otoyolara erişilebilirlik indeksi, X10: 10,000 Dolar başına emlak vergisi oranı, X11: Öğrenci/öğretmen oranı, X12: $1000.(\mathrm{B} - 0.63)^2\mathrm{I}(\mathrm{B} < 0.63)$ Burada B Afrikalı Amerikanların oranı, X13: Düşük gelirli yüzdesi, X14: 1,000 Dolar evlerin sahiplerinin medyanı, değişkenlerinden oluşmaktadır [63, s346].

Her bir gerçek veri seti için ayrı ayrı G-K kuralı ile faktör analizinde boyut sayısı belirlenmiştir. Her bir gerçek veri setinde PA ve BPA süreci;

PA analizi yapabilmek için $\mathrm{N}(\mu = 0, \Sigma = \mathrm{I})$ olan normal dağılımdan değişken sayısına ve orijinal örnek büyüklüğüne paralel veri türetilmiştir. Rasgele verilerin özdeğerleri ve özdeğerlerin dağılımının ortalama, medyan ile 90., 95., 97.5 ve 99. persentil değerleri hesaplanmıştır. Her bir orjinal veri setinin özdeğerleri hesaplanmış ve ragele veriden elde edilen özdeğerlerin ortalama, medyan , 90., 95., 97.5 ve 99. persentil değerleri ile karşılaştırılmıştır. Boyut sayısı, gerçek veriden hesaplanan özdeğerlerin rasgele veriden hesaplanan özdeğerlerden büyük olduğu noktadır. BPA uygulanan süreçte ise, her bir gerçek veri setinde ayrı ayrı BPA analizi yapabilmek için $N(\mu = 0, \Sigma = \mathrm{I})$ olan normal dağılımdan değişken sayısına ve orijinal örnek büyüklüğüne paralel veri üretilmiştir. Rasgele verilerin özdeğerleri ve özdeğerlerin dağılımının ortalama, medyan, 90., 95., 97.5 ve 99. persentil değerleri hesaplanmıştır. Her bir orjinal veri setinin özdeğerlerinin persentileleri dayalı bootstrap güven aralıklarının medyanları hesaplanmış ve ragele veriden elde edilen özdeğerlerin ortalama, medyan ile 90., 95., 97.5 ve 99. persentil değerleri ile karşılaştırılmıştır. Boyut sayısı, gerçek veriden hesaplanan özdeğerlerin medyanı rasgele veriden hesaplanan özdeğerlerden büyük olduğu noktadır. Üç gerçek veri seti için tüm yöntemlerin faktör analizinde buldukları boyut sayıları Çizelge 3.1'de verilmiştir.

Çizelge 3. 1 Üç farklı gerçek veri seti için boyut sayısını belirlemede G-K kuralı, $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} persentil değerleri

	Klasik Metot	Paralel Analiz					
	[1]G-K kuralı	Persentiler					
Veriler	d	[2]$PA_{ortalama}$	[3]PA_{medyan}	[4]PA_{99}	[5]$PA_{97.5}$	[6]PA_{95}	[7]PA_{90}
Araba markaları verisi	2	2	2	2	2	2	2
WIHCSEM verisi	4	2	2	2	2	2	2
Boston Housing verisi	3	3	3	3	3	3	3
		Bootstrap Modifyeli Paralel Analiz					
		Persentiler					
Veriler	d	[8]$BPA_{ortalama}$	[9]BPA_{medyan}	[10]BPA_{99}	[11]$BPA_{97.5}$	[12]BPA_{95}	[13]BPA_{90}
Araba markaları verisi	2	2	2	2	2	2	2
WIHCSEM verisi	4	2	2	2	2	2	2
Boston Housing verisi	3	3	3	3	3	3	3

Not: *N*: örnek sayısı, *d*: boyut sayısı
[1]G-K kuralı: özdeğer 1'den büyük kuralı
[2]$PA_{ortalama}$: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin ortalama değeri
[3]PA_{medyan}: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin medyan değeri
[4]PA_{99}: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 99. persentil eşik değeri
[5]$PA_{97.5}$: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 97.5. persentil eşik değeri
[6]PA_{95}: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 95. persentil eşik değeri
[7]PA_{90}: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 90. persentil eşik değeri
[8]$BPA_{ortalama}$: BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin ortalama değeri
[9]BPA_{medyan}: BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin medyan değeri
[10]BPA_{99}: BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 99. persentil eşik değeri
[11]$BPA_{97.5}$: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 97.5. persentil eşik değeri
[12]BPA_{95}: BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 95. persentil eşik değeri
[13]BPA_{90}: BPA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 90. persentil eşik değeri

Analiz sonuçlarına bakıldığında, ilk gerçek veri olan Araba markaları için bütün yöntemlerin aynı sonucu verdiği ve hepsininde boyut sayısını $d = 2$ olarak bulduğu görülmektedir. İkinci veri WIHCSEM için ise klasik metodlardan G-K kuralını boyut sayısını $d = 4$, PA ve BPA yöntemlerinin tüm modifikasyonlarının boyut sayısını $d = 2$ olarak bulduğu görülmektedir. Son gerçek veri Boston Housing için ise boyut sayısını, G-K kuralının $d = 3$, PA ve BPA'nın tüm yöntemlerinin $d = 3$ bulduğu görülmektedir.

Genel olarak gerçek üç veri seti sonuçları incelendiğinde gerçek veride PA ve BPA'nın başarılı performans gösterdiği görülmektedir. Her üç veride de PA ve BPA'nın tüm modifikasyonları aynı sonuçları vermiştir. G-K kuralı ise WIHCSEM verisinde boyut sayısını olması gerekenden fazla bulmuş, diğer veriler için ise PA ve BPA ile aynı sonuçlar vermiştir.

3.5 Simülasyon Çalışması

Analiz için kullanılacak veriler simülasyon çalışmasıyla elde edilmiştir. Araştırmada değişkenlerin örneklem büyüklüğü, faktör sayısı, faktör yükü, faktör başına düşen madde sayısı faktör analizi literatür taraması ve refarans alınan çalışmalara paralel olarak seçilmiştir.

Çalışmada üç farklı faktör modeli oluşturulmuştur. Modeller: tek faktörlü, üç faktörlü korelasyonlu ve üç faktörlü korelasyonsuz olarak alınmıştır.

Faktör sayısı (m): 1 ve 3 olarak alınmıştır.

Faktör yükleri: 0.9, 0.7 ve 0.5'tir.

Örneklem büyüklükleri: 50, 100, 200 ve 400 olarak seçilmiştir.

Değişken sayısı: 18 olarak alınmıştır.

Her koşul için çok değişkenli normal dağılımdan 1000 adet veri üretilmiştir. Her bir örneklem büyüklüğü için simülasyon çalışmasında PA, BPA ve klasik metotlardan G-K kuralı yöntemleri üç farklı faktör modeli için ayrı ayrı uygulanmıştır. Bu modeller için veriler korelasyon matrislerinden üretilmiştir.

3.5.1 Faktör Modelleri

Her üç faktör modelinde de, değişken sayısı 18 olarak alınmıştır. Modeller tek faktörlü, üç faktörlü korelasyonsuz ve üç faktörlü korelasyonlu olmak üzere belirlenmiştir. Her

bir model için faktör sayısı belirleme süreci aşağıda verilmiştir. m faktör sayısını göstermektedir.

- Model 1: Tek faktörlü model (m=1) için korelasyon matrisinin oluşturulması süreci;

  Faktör yüklerinin belirlenmesi, $\Lambda_{1,1} = ... = \Lambda_{1,6} = .9, \Lambda_{1,7} = ... = \Lambda_{1,12} = .7,$ $\Lambda_{1,13} = ... = \Lambda_{1,18} = .5$ faktör yükleri büyük (0.9), orta (0.7) ve küçük (0.5) karma olarak belirlenir. Kovaryans matrisi, $\mathrm{COV}(f) = [1.00]$ korelasyon matrisi hesaplanır. Hesaplanan korelasyon matrisi model 1 için elde edilmiştir.

- Model 2: Korelasyonsuz üç faktörlü model (m=3) için korelasyon matrisinin elde edilmesi süreci

  Faktör yüklerinin belirlenmesi, $\Lambda_{1,1} = ... = \Lambda_{1,6} = .9,$ $\Lambda_{2,7} = ... = \Lambda_{2,12} = .7,$ $\Lambda_{3,13} = ... = \Lambda_{3,18} = .5$ faktör yükleri 1. faktör için büyük (0.9), 2. faktör için orta (0.7) ve 3. faktör için küçük (0.5) olarak belirlenir. Faktörler korelasyonsuz olduğunda $\mathrm{COV}(f) = I$ birim matrise eşit olur,

  $$\mathrm{cov}(f) = \begin{pmatrix} 1.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

  korelasyonsuz üç faktörlü model (m=3) için korelasyon matrisi elde edilir.

- Model 3: Korelasyonlu üç faktörlü model (m=3) için korelasyon matrisinin elde edilmesi;

  Faktör yüklerinin belirlenmesi, $\Lambda_{1,1} = ... = \Lambda_{1,6} = .9,$ $\Lambda_{2,7} = ... = \Lambda_{2,12} = .7,$ $\Lambda_{3,13} = ... = \Lambda_{3,18} = .5$ faktör yükleri 1. faktör için büyük (0.9), 2. faktör için orta (0.7) ve 3. faktör için küçük (0.5) olarak belirlenir.

  Faktörler korelasyonlu $\Phi = .5$ olması koşulu altında; kovaryans matrisi,

  $$\mathrm{cov}(f) = \begin{pmatrix} 1.00 & 0.50 & 0.50 \\ 0.50 & 1.00 & 0.50 \\ 0.50 & 0.50 & 1.00 \end{pmatrix}$$

olarak belirlenir. Korelasyonlu üç faktörlü model için korelasyon matrisi elde edilir.

Korelasyon matrisleri belirlendikten sonra model 1, model 2 ve model 3 için ayrı ayrı aşağıdaki simülasyon süreci uygulanır. Çalışma faktör analizinde boyut sayısı belirlemede kullanılan klasik metot G-K kuralı, $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95},

PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} yöntemlerinin faktör analizine adaptasyonu simülasyonunda gerçek verimizin yaklaşık 3 katı, 5 katı 10 katı ve 20 katı alınarak $N = 50, 100, 200, 400$ olarak belirlenir ve modeldeki değişken başına örnek birim sayısı için aşağıdaki adımlar oluşturulur.

- Birim kovaryans matrisli çok değişkenli normal dağılımdan nx18 bağımsız rastgele 1000 örnek çekilip, özdeğerleri ve özdeğerlerin dağılımının ortalama, medyan ile 90., 95., 97.5 ve 99. persentil değerleri hesaplanır.
- Her bir model için varsayımsal anakütleden bir örnek çekilir. Çekilen örnekten, örnekle aynı boyutta iadeli olarak 1000 örnek çekilip, özdeğerlerin bootstrap dağılımının medyanı bulunur.
- Her bir model için varsayımsal anakütleden bir örnek çekilir. Çekilen örnekten elde edilen özdeğerler için G-K (özdeğer>1) kuralı ile boyut sayısı bulunur.
- Çekilen örneğin i. özdeğeri 1.adımda i.özdeğer için karşılık gelen persentil değerleri ile karşılaştırılır. Örneğin i.özdeğeri karşılık gelen i.persentil özdeğerden küçük oluncaya dek boyut sayısı bir arttırılarak PA'ya dayanan boyut sayısı bulunur.
- i. özdeğerin 2. adımda bulunan medyanı 1.adımda i. özdeğer için karşılık gelen persentil değerleri ile karşılaştırılır. Örneğin i.özdeğer için bootstrap dağılımının medyanı karşılık gelen i. persentil özdeğerden küçük oluncaya dek boyut sayısı bir arttırılarak BPA'ya dayanan boyut sayısı bulunur.
- 1. adıma dönülür ve adımlar 1000 kez tekrarlanır.

3.6 Simülasyon Sonuçları

Simülasyon sonuçları model 1 için Çizelge 3.2, model 2 için Çizelge 3.3 ve model 3 içinde Çizelge 3.4 verilmiştir. Çizelge 3.2'de model 1 için klasik metot G-K kuralı, $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} persentil sonuçları bulunmuştur.

Model 1 faktör sayısı m=1 için sonuçlar Çizelge 3.2'de verilmiştir. Çizelge 3.2 incelendiğinde, örneklem büyüklüğü $N = 50$ olduğunda yöntemlerin performansı incelendiğinde PA ve BPA'nın yüksek performansa sahip olduğu görülmektedir. $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95} ve PA_{90}'nın doğru boyut sayısını %100 bulduğu

görülmektedir. BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95} ve BPA_{90}'nın ise %98.60 değeri doğru boyut sayısını tespit ettiği görülmektedir. G-K kuralının %1.60 değeri ile performansının düşük olduğu görülmektedir. Örnek büyüklüğü küçük olduğu koşullarda G-K kuralının performansının zayıf olduğu görülmektedir.

Örneklem büyüklüğü N=100 olduğunda ise PA ve BPA'nın yüksek performansa sahip olduğu görülmektedir. PA'nın en yüksek performansa sahip olduğu doğru boyut sayısını %100 olarak bulduğu görülmektedir. BPA'nın %99.70 değeri ile performansının iyi olduğu görülmektedir. G-K kuralının %13.0 değeri ile düşük performans göstermektedir. En iyi performansı PA göstermiştir, BPA'da tatmin edici performans göstermiştir.

Örneklem büyüklüğü N=200 olduğunda ise PA ve BPA'nın yüksek performansa sahip olduğu görülmektedir. G-K kuralı performansını %67.50 yükseltiği görülmektedir.

Örneklem büyüklüğü N=400 olduğunda ise yine PA'nın en yüksek performansa sahip olduğu doğru boyut sayısını %100 tespit ettiği görülmektedir. BPA'nın %99.90 değeri ie tespit ettiği görülmektedir. G-K kuralının %99.10 değeri ile başarılı performans göstermiştir.

Genel aritmetik ortalamaları alındığında bütün örneklem sayıları için ortalamalara bakıldığında PA'nın ortalaması %100 ve BPA'nın %99.60 iken, G-K kuralının ortalaması %45.0 bulmuştur. Bir boyutlu modelde PA ve BPA yöntemlerinin çok iyi performans gösterdiği sonucuna ulaşılmıştır. Uygulamada en çok kullanılan G-K kuralının ise performansının örneklem sayısından etkilendiği görülmektedir.

Çizelge 3. 2 Model 1 Tek faktörlü model için G-K kuralı, $PA_{ortalama}$, PA_{medyan} , PA_{99} , $PA_{97.5}$, PA_{95} , PA_{90} ve $BPA_{ortalama}$, BPA_{medyan} , BPA_{99} , $BPA_{97.5}$, BPA_{95} , BPA_{90} yöntemlerine göre boyut sayısı belirleme için göreli frekanslar

	Klasik Metot		Paralel Analiz						
	[1]G-K kuralı		Persentiler						
N	d		d	[2]$PA_{ortalama}$	[3]PA_{medyan}	[4]PA_{99}	[5]$PA_{97.5}$	[6]PA_{95}	[7]PA_{90}
50	1	**1.6**	**1**	**100**	**100**	**100**	**100**	**100**	**100**
	2	19.9							
	3	53.5							
	4	25.0							
100	**1**	**13.0**	**1**	**100**	**100**	**100**	**100**	**100**	**100**
	2	58.0							
	3	28.0							
	4	1.0							
200	**1**	**67.5**	**1**	**100**	**100**	**100**	**100**	**100**	**100**
	2	31.9							
	3	0.6							
400	**1**	**99.1**	**1**	**100**	**100**	**100**	**100**	**100**	**100**
	2	0.9							
Gen	**1**	**45.3**	**1**	**100**	**100**	**100**	**100**	**100**	**100**
	2	27.68							
	3	20.53							
	4	6.5							

	Bootstrap Modifyeli Paralel Analiz						
	Persentiler						
N	d	[8]$BPA_{ortalama}$	[9]BPA_{medyan}	[10]BPA_{99}	[11]$BPA_{97.5}$	[12]BPA_{95}	[13]BPA_{90}
50	0	1.10	1.10	1.10	1.10	1.10	1.10
	1	98.40	98.4	**98.9**	**98.9**	**98.9**	**98.9**
	2	0.50	0.50				
100	0	0.30	0.30	0.30	0.30	0.30	0.30
	1	**99.7**	**99.7**	**99.7**	**99.7**	**99.7**	**99.7**
200	0	0.10	0.10	0.10	0.10	0.10	0.10
	1	**99.9**	**99.9**	**99.9**	**99.9**	**99.9**	**99.9**
400	0	0.10	0.10	0.10	0.10	0.10	0.10
	1	**99.9**	**99.9**	**99.9**	**99.9**	**99.9**	**99.9**
Gen	0	0.40	0.40	0.40	0.40	0.40	0.40
	1	99.48	99.48	**99.6**	**99.6**	**99.6**	**99.6**
	2	0.13	0.13				

Not: [1,2,3,4,5,6,7,8,9,10,11,12,13] açıklamaları Çizelge 3.1'de verilmiştir. Gen: Her bir örneklem için yöntemlerin performanslarının ortalaması

Simülasyon sonuçları Model 2 faktör sayısı m=3 için Çizelge 3.3'te verilmiştir. Çizelge 3.3 incelendiğinde, örneklem büyüklüğü N=50 olduğunda PA ve BPA'nın yüksek performans gösterdiği görülmektedir. BPA'nın $BPA_{ortalama}$ ve BPA_{medyan} %98.50 değeri ile en iyi performansı göstermiştir. BPA'nın geriye kalan diğer persentilerin doğru boyut sayısını tespit etmede % 90 değeri üzerinde performans göstermiştir. PA'nın $PA_{ortalama}$ ve PA_{medyan} persentil değerleri sırasıyla doğru boyut sayısını % 94.5 ve %94.6 tespit etmiştir. G-K kuralının ise doğru boyut sayısını %3.70 ile düşük seviyede tespit etmiştir. Örneklem büyüklüğü N=100 olduğunda ise doğru boyut sayısı tespit etmede $PA_{ortalama}$, PA_{medyan}, PA_{90} ve PA_{95}'in %99.80 en yüksek performansa sahip oldukları görülmektedir. BPA'nın yaklaşık %99 değeri ile iyi perfomans gösterdiği görülmektedir. G-K kuralının ise %30.6 değeri düşük performans göstermektedir.

Örneklem büyüklüğü N=200 olduğunda ise PA'nın doğru boyut sayısını en iyi performansla %100 ile bulduğu görülmektedir. BPA'nın %99.70 ile başarılı bir performans göstermiştir. G-K kuralının %85.1 ile performansının iyileştiği görülmektedir.

Örneklem büyüklüğü N=400 olduğunda ise G-K kuralı, PA ve BPA'nın doğru boyut sayısı bulmada %99 üstü başarı göstermişlerdir.

Model 2'de yöntemler dört örneklem için genel ortalamaları alındığında ise G-K kuralı doğru boyut sayısı tespitinde %54.83, PA'nın en yüksek %98.58 ve BPA'nın en yüksek olanın ise %99.18 ile performans göstermektedir.

Genel olarak bakıldığında BPA ve PA'nın tatmin edici performans gösterdiği, G-K kuralının örneklem büyüklüğünden etkilendiği söylenebilir.

Çizelge 3. 3 Model 2 korelasyonsuz üç faktörlü model için G-K kuralı, $PA_{ortalama}$, PA_{medyan} , PA_{99} , $PA_{97.5}$, PA_{95} , PA_{90} ve $BPA_{ortalama}$, BPA_{medyan} , BPA_{99} , $BPA_{97.5}$, BPA_{95} , BPA_{90} yöntemlerine göre boyut sayısı belirleme için göreli frekanslar

		Klasik Metot	Paralel Analiz					
		[1]G-K kuralı	Persentiler					
N	d		d	[2]$PA_{ortalama}$	[3]PA_{medyan}	[4]PA_{99}	[5]$PA_{97.5}$	[6]PA_{95}	[7]PA_{90}
50			2	5.40	5.30	20.90	17.00	13.90	11.40
	3	**3.70**	**3**	**94.50**	**94.60**	79.10	83.00	86.10	88.60
	4	54.3	4	0.10	0.10				
	5	39.8							
	6	2.20							
100			2	0.20	0.20	0.40	0.30	0.20	0.20
	3	**30.6**	**3**	**99.80**	**99.80**	99.60	99.70	**99.80**	**99.80**
	4	62.50							
	5	6.90							
200	**3**	**85.10**	**3**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
	4	14.90							
400	**3**	**99.90**	**3**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
	4	0.10							
Gen			2	1.40	1.40	5.33	4.33	3.53	2.90
	3	**54.83**	**3**	**98.58**	**98.58**	94.68	95.68	96.48	97.10
	4	32.95	4	0.03	0.03				
	5	11.68							
	6	0.55							

		Bootstrap Modifyeli Paralel Analiz					
		Persentiler					
N	d	[8]$BPA_{ortalama}$	[9]BPA_{medyan}	[10]BPA_{99}	[11]$BPA_{97.5}$	[12]BPA_{95}	[13]BPA_{90}
50	0	0.80	0.80	0.80	0.80	0.80	0.80
	2	0.40	0.40	12.70	8.60	6.20	3.90
	3	**98.50**	**98.50**	86.50	90.6	93.0	95.3
	4	0.30	0.30				
100	0	1.10	1.10	1.10	1.10	1.10	1.10
	2			0.20	0.20	0.10	0.10
	3	**98.90**	**98.90**	98.70	98.70	98.80	98.80
200	0	0.30	0.30	0.30	0.30	0.30	0.30
	3	**99.70**	**99.70**	**99.70**	**99.70**	**99.70**	**99.70**
400	0	0.40	0.40	0.40	0.40	0.40	0.40
	3	**99.60**	**99.60**	**99.60**	**99.60**	**99.60**	**99.60**
Gen	0	0.65	0.65	0.65	0.65	0.65	0.65
	2	0.10	0.10	3.23	2.20	1.58	1.00
	3	**99.18**	**99.18**	96.13	97.15	97.78	98.35
	4	0.08	0.08				

Not: [1,2,3,4,5,6,7,8,9,10,11,12,13] açıklamaları Çizelge 3.1'de verilmiştir. Gen: Her bir örneklem için yöntemlerin performanslarının ortalaması

Simülasyon sonuçları Model 3 faktör sayısı m=3 için Çizelge 3.4'te verilmiştir. Çizelge 3.4 incelendiğinde örneklem büyüklüğü N=50 olduğunda yöntemlerin doğru boyut sayısı bulmada performansları incelendiğinde $BPA_{ortalama}$ ve BPA_{medyan} sırasıyla %54.3 ve %55 değeri ile en yüksek performansa sahip oldukları görülmektedir. $PA_{ortalama}$ ve PA_{medyan} ise sırasıyla %35.4 ve %35.7 ile doğru boyut sayısı tespitinde bulunmuşlardır. G-K kuralının %5.9 ile düşük bir performans göstermiştir.

Örneklem büyüklüğü N=100 olduğunda ise yine doğru boyut sayısı tespitinde en iyi performansı $BPA_{ortalama}$ ve BPA_{medyan} yaklaşık %83 değeri ile en yüksek performansa sahiptir. $PA_{ortalama}$ ve PA_{medyan} %72 değeri iyi performans gösterdiği görülmektedir. G-K kuralının %36.6 değeri ile düşük performans göstermektedir.

Örneklem büyüklüğü N=200 olduğunda ise doğru boyut sayısı tespitinde en iyi performansı $BPA_{ortalama}$ ve BPA_{medyan} yaklaşık %99 değeri ile sahiptir. $PA_{ortalama}$ ve PA_{medyan} %97.5 değeri yüksek performans gösterdiği görülmektedir. G-K kuralının %87 değeri ile iyi performans göstermektedir.

Örneklem büyüklüğü N=400 olduğunda ise doğru boyut sayısı belirlemede G-K kuralı, PA ve BPA yöntemleri yaklaşık %100 değeri ile yüksek performans göstermiştir.

Genel ortalama alındığında ise BPA'nın en yüksek performans %84.38 değerine sahip olduğu, PA'nın ise en yüksek performansının %76.3 olduğu ve G-K kuralının ise %57.45 değerine sahip olduğu görülmektedir.

Çizelge 3. 4 Model 3 korelasyonlu üç faktörlü model için G-K kuralı, $PA_{ortalama}$, PA_{medyan} , PA_{99} , $PA_{97.5}$, PA_{95} , PA_{90} ve $BPA_{ortalama}$, BPA_{medyan} , BPA_{99} , $BPA_{97.5}$, BPA_{95} , BPA_{90} göre boyut sayısı belirleme için göreli frekanslar

	Klasik Metot		Paralel Analiz						
	[1]G-K kuralı		Persentiler						
N	d		d	[2]$PA_{ortalama}$	[3]PA_{medyan}	[4]PA_{99}	[5]$PA_{97.5}$	[6]PA_{95}	[7]PA_{90}
50			1	7.5	7.20	25.80	21.70	18.90	15.00
			2	57.1	57.10	63.50	65.20	64.20	63.50
	3	**5.9**	**3**	**35.4**	**35.70**	10.7	13.1	16.9	21.5
	4	55.1							
	5	37.3							
	6	1.7							
100			1	0.1	0.1	2.20	1.80	1.20	0.90
			2	28.3	27.8	58.60	53.60	48.40	43.20
	3	**36.6**	**3**	**71.6**	**72.1**	39.20	44.60	50.40	55.90
	4	58.4							
	5	5.0							
200			2	2.50	2.50	12.4	9.7	7.9	6.3
	3	**87.4**	**3**	**97.5**	**97.5**	87.6	90.3	92.1	93.7
	4	12.5							
	5	0.1							
400			2		0.10	0.1	0.1		
	3	**99.9**	**3**	**100**	99.9	99.9	99.9	**100**	**100**
	4	0.1							
Gen			1	1.9	1.83	7.00	5.88	5.03	3.98
			2	21.98	21.85	33.65	32.15	30.15	28.25
	3	**57.45**	**3**	**76.13**	**76.33**	59.35	61.98	64.83	67.78
	4	31.53							
	5	10.6							
	6	0.43							

	Bootstrap Modifyeli Paralel Analiz					
	Persentiler					
N	d	[8]$BPA_{ortalama}$	[9]BPA_{medyan}	[10]BPA_{99}	[11]$BPA_{97.5}$	[12]BPA_{95}	[13]BPA_{90}
50	0	0.80	0.80	0.80	0.80	0.80	0.80
	1	0.20	0.10	6.70	5.00	3.30	2.40
	2	44.70	44.1	79.90	77.40	74.30	69.30
	3	**54.3**	**55**	12.60	16.80	21.60	27.50
100	0	0.50	0.50	0.50	0.50	0.50	0.50
	2	16.20	15.80	52.20	45.50	38.70	34.20
	3	**83.30**	**83.70**	47.3	54.00	60.8	65.3
200	0	0.10	0.10	0.10	0.10	0.10	0.10
	2	1.00	1.00	8.30	5.90	4.70	3.70
	3	**98.9**	**98.90**	91.6	94.00	95.2	96.2
400	0	0.10	0.10	0.10	0.10	0.10	0.10
	3	**99.9**	**99.9**	**99.9**	**99.9**	**99.9**	**99.9**
Gen	0	0.38	0.38	0.38	0.38	0.38	0.38
	1	0.05	0.03	1.70	1.25	0.83	0.60
	2	15.48	15.23	35.1	32.20	29.43	26.8
	3	**84.10**	**84.38**	62.83	66.18	69.38	72.23

Not: [1,2,3,4,5,6,7,8,9,10,11,12,13] açıklamaları Çizelge 3.1'de verilmiştir. Gen: Her bir örneklem için yöntemlerin sonuçlarının ortalaması

Tek faktörlü model (Model1, m=1), üç faktörlü korelasyonsuz model (Model2, m=3) ve üç faktörlü korelasyonlu model (Model3, m=3) için genel olarak, faktör analizinde boyut belilemede PA ve BPA'nın klasik yöntemlerden tüm modellerde daha iyi sonuçlar verdiği görülmektedir. PA ve BPA'nın korelasyonuz 3 faktörlü ve tek faktörlü modelde bütün örnek büyüklüklerinde başarı oranları %94 üzeri olduğu, koreasyonlu üç faktörlü modelde ise örnek sayısı artığında performanslarının artığı görülmektedir. Klasik metotlardan G-K kuralı küçük örneklerde düşük performans gösterirken, örneklem sayısı artıkça daha iyi performans gösterdiği görülmektedir. Yöntemlerin korelasyonlu ve korelasyonsuz modellerdeki performanslarına bakıldığında PA ve BPA'nın performansı diğer iki yönteme göre çok büyük oranda başarılı oldukları görülmektedir. PA ve BPA karşılaştırıldığında ise özellikle küçük örneklerde BPA'nın PA'dan daha hassas olduğu sonucu ortaya çıkmaktadır. Model 3'te PA ve BPA faktör analizinde boyut sayısına bakılırken modifikasyonlardan ortalama ve medyanın, 99., 97.5., 95. ve 90. persentillere göre daha iyi sonuçlar verdiği görülmektedir. Ayrıca model 3'te yani korelasyonlu modelde BPA'nın performansının daha iyi olduğu görülmektedir. Model 1 ve model 2'de PA ve BPA'nın özellikle küçük örneklerde klasik yöntemlere göre çok büyük farkla doğru boyut sayısını tespit ettikleri görülmektedir. Bu fark G-K kuralı için örnek sayısı artıkça düştüğü görülmektedir. PA ve BPA'nın korelasyonsuz modellerde kullanımı önerilmektedir. Model 3 için bakıldığında örnek birim sayısı N=50 olduğunda PA ve BPA'nın performansının düşük olduğu görülmektedir fakat klasik yöntemlere göre çok daha iyi doğru boyut sayısı belirledikleri görülmektedir.

Üç farkılı faktör modellerinde boyut sayısı problemi için yapılan simülasyon çalışmasının sonuçlarına göre, PA ve BPA'nın diğer yöntemlerden daha iyi sonuçlar verdiği görülmektedir. Özellikle faktörler korelasyonlu olduğu modellerde PA ve BPA'nın persentil değerlerinin kullanılması daha uygun olacağı görülmektedir. Bununla beraber, faktör analizinde boyut belirleme için BPA'nın özellikle korelasyonlu modeldeki performansı dikkat çekmektedir.

BÖLÜM 4

FAKTÖR ANALİZİNDE BOYUT SAYISININ BELİRLENMESİNDE PARALEL ANALİZE DAYANAN AMPİRİK OLABİLİRLİK GÜVEN ARALIĞI YAKLAŞIMI

4.1 Ampirik Olabilirlik Yöntemi

Owen [65] tarafından önerilen Ampirik olabilirlik (EL) yöntemi, olabilirlik oran prensiplerine dayanır. EL, log-olabilirlik fonksiyonuna dayanan çıkarımsal parametrik olmayan bir yöntemdir, güven aralığı oluşturmak ve hipotez testi yapmak için kullanılır. Ampirik olabilirlik, herhangi bir dağılım ailesine bağlı değildir ve asimptotik olarak güçlü özelliklere sahiptir. Ayrıca ampirik olabilirlik (EL), herhengi bir dağılım ailesine ihtiyaç olmadan olabilirlik koşullarını sağlayan parametrik olamayan yöntemdir [65].

Ampirik olabilirlik yöntemi birçok araştırmacı tarafından araştırılmış ve birçok farklı alanda uygulanmıştır. Örneğin tahmin denklemlerinde, genel lineer modeller, kısmi lineer modeller, multiresponse regresyonda yarı parametrik ve parametrik modeller, zaman serilerinde, sensörlü verilerde ve lineer regresyonda birçok alanda uygulanabilirliği gösterilmiştir [65, 66, 67, 68, 69].

Birbirinden bağımsız $X_1,\ldots,X_n$ rasgele değişkenler bilinmeyen F dağılımına sahip olsun. Bağımsız $X_1,\ldots,X_n$ rasgele değişkenlerin ampirik kümülatif dağılım fonksiyonu,

$$F_n(x)=\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}(\mathrm{X}_i\le x) \tag{4.1}$$

ile gösterilsin. Burada $I(X)$ veri kümesi fonksiyonunun indikatörüdür. Ayrıca kümalatif F dağılımının olabilirlik fonksiyonu,

$$L(F)=\prod_{i=1}^{n}(F(X_i)-F(X_i-)) \tag{4.2}$$

şeklindedir [65].

4.1.1 İki Örneklem Ampirik Olabilirlik Yöntemi

$\{X_1,\ldots,X_n\}$ ve $\{Y_1,\ldots,Y_k\}$, $X \in R^d$ ve $Y \in R^d$ olmak üzere birbirinden bağımsız rasgele vektörlerdir. X ve Y'nin ortalaması sırasıyla μ_x ve μ_y olmak üzere, bilinmeyen θ parametresi ise X ve Y'nin ortalamaları arasındaki farka, $\theta = \mu_x - \mu_y \in R^d$ eşittir. Olasılıksal vektörler $p=(p_1,\ldots,p_n)$ ve $q=(q_1,\ldots,q_k)$, $p_i \geq 0$, $q_j \geq 0$ olsun,

$$\sum_{i=1}^{n} p_i = 1 \text{ ve } \sum_{i=1}^{k} q_i = 1 \tag{4.3}$$

bu durumda $\mu_x(p) = \sum_{i=1}^{n} p_i X_i$ ve $\mu_y(q) = \sum_{j=1}^{k} q_j Y_j$ olur.

$$\theta(p,q) = \mu_y(q) - \mu_x(p), \tag{4.4}$$

$\theta \in \mathrm{R}^d$ için iki örnek ampirik olabilirlik fonksiyonu,

$$L(\theta) = \max_{(p,q):\theta(p,q)=\theta} \left(\prod_{i=1}^{n} p_i \right) \left(\prod_{j=1}^{k} q_j \right) \tag{4.5}$$

şeklindedir. θ için iki örnek ampirik log-olabilirlik fonksiyonu,

$$l(\theta) = -2 \max_{(p,q):\theta(p,q)=\theta} \left(\sum_{i=1}^{n} \log(np_i) + \sum_{j=1}^{k} \log(kq_j) \right) \tag{4.6}$$

biçiminde gösterilebilir. Burada $N = n+k$, $f_n = N/n$ ve $f_k = N/k$ ve $n \geq k > d$ olsun. Lagrange çarpanları metodu ile

$$l(\theta) = 2\left[\sum_{i=1}^{n} \log\left\{1 - f_n \lambda^T (X_i - \mu_x)\right\} + \sum_{j=1}^{k} \log\left\{1 + f_k \lambda^T (Y_j - \mu_y)\right\} \right] \tag{4.7}$$

olarak yazılabilir. Burada lagrange çarpanı ile $\lambda = \lambda(\theta)$

$$\sum_{i=1}^{n} \frac{X_i - \mu_x}{1 - f_n \lambda^T (X_i - \mu_x)} = 0 \text{ ve } \sum_{j=1}^{k} \frac{Y_j - \mu_y}{1 + f_k \lambda^T (Y_j - \mu_y)} = 0 \tag{4.8}$$

$$\sum_{j=1}^{k}\frac{Y_j}{1+f_k\lambda^T(Y_j-\mu_y)}-\sum_{i=1}^{n}\frac{X_i}{1-f_n\lambda^T(X_i-\mu_x)}=\theta \tag{4.9}$$

(4.8) eşitliklerini sağlar.

Jing [70] çalışmasıyla X ve Y 'nin varayans-kovaryans matrisleri sırasıyla $\sum_x$ ve $\sum_Y$ 'nin tam ranklı sonlu kovaryans matrisleri olmaları koşulu altında,

$$l(\theta)\xrightarrow{d}\chi^2 \tag{4.10}$$

olduğunu göstererek $(1-\alpha)$- seviyesinde θ'nın güven aralığını,

$$\{\theta\,|\,-2\log R(\theta)<\chi^2_{1,1-\alpha}\} \tag{4.11}$$

olarak elde etmiştir [64, 66, 67, 68, 69, 70, 71, 72].

4.2 Paralel Analize Dayanan İki Örneklem Ampirik Olabilirlik Güven Aralığı Yöntemi

Bu bölümde önerilen PELCI (empirical likelihood confidence interval basen on paralel analysis), özdeğerlerin iki örneklem ortalaması farkının ampirik olabilirlik güven aralığını kullanılarak faktör sayısı bulmayı amaçlayan bir yöntemdir. PELCI yöntemi sürecinde;

$\{X_1,\ldots,X_n\}$ ve $\{Y_1,\ldots,Y_k\}$ sırasıyla paralel özdeğer örnekleri ve bootstrap özdeğer örnekleri olsun. Burada μ_x paralel özdeğerlerin ortalaması ve μ_y bootstrap özdeğerlerin ortalaması olsun. θ ise bootstrap özdeğerlerin ortalaması ile paralel özdeğerlerin ortalaması arasındaki fark olarak tanımlansın [72]. Burada iki örneklem ampirik olabilirlik güven aralığı ile θ'nın güven aralıkları elde edilmektedir.

PELCI yöntemi,

- Birinci bootstrap özdeğeri ile birinci paralel özdeğer ortalamaları arasındaki farkın (θ'nın) güven aralığı iki örneklem ampirik olabilirlik güven aralığı yöntemi ile elde edilir.
- Ampirik olabilirlik yöntemi ile bulunan güven aralığı pozitif ise, bu birinci faktör olarak alınır. Aynı prosedür ikinci özdeğerler içinde uygulanır.
- Bu adımlar negatif güven aralığı bulununcaya dek devam edilir. Hesaplanan pozitif güven aralığı sayısı boyut sayısı olarak alınır.

PELCI yöntemi üç farklı gerçek veri seti ve dokuz farklı faktör modelinde boyut sayısı problemine çözüm için kullanılmıştır.

4.3 Gerçek Veri ile Önerilen Yaklaşımın Uygulamaları

Bu bölümde daha önceki çalışmalarda faktör analizi yapılmış üç adet gerçek veri seti için faktör sayısı belirlemek için G-K kuralı, PA ve PELCI yöntemleri yardımıyla faktör sayıları elde edilecektir. Analizler için SAS 9.4 ve R programı kullanılmıştır. G-K kuralı ve PA yöntemini analiz sonuçaları SAS 9.4'te hesaplanmıştır. PELCI için ise veriler SAS programında üretildikten sonra R programında analizleri yapılmıştır. Gerçek üç veri seti ile ilgili ayrıntılı bilgiler Bölüm 3.4'te verilmiştir. Her bir veri seti için faktör analizinde G-K kuralı ile boyut sayısı hesaplamıştır. PA analizi yapabilmek için $N\left(\mu = 0, \Sigma = \mathrm{I}\right)$ olan normal dağılımdan değişken sayısına ve orijinal örnek büyüklüğüne paralel veri türetilmiştir. Rasgele verilerin özdeğerleri ve özdeğerlerin dağılımının 95. ve 99. persentil değerleri hesaplanmıştır. Her bir orjinal veri setinin özdeğerleri hesaplanmış ve ragele veriden elde edilen özdeğerlerin 95. ve 99. persentil değerleri ile karşılaştırılmıştır. Boyut sayısı, gerçek veriden hesaplanan özdeğerlerin rasgele veriden hesaplanan özdeğerlerden büyük olduğu noktadır. Her bir orjinal veri setinin özdeğerleri hesaplanır ve ragele veriden elde edilen özdeğerlerin 95. ve 99. persentil değerleri ile karşılaştırılır. Boyut sayısı için gerçek veriden hesaplanan özdeğerlerin rasgele veriden hesaplanan özdeğerlerden büyük olduğu noktadır. Boyut sayısı için gerçek veriden hesaplanan özdeğerlerin medyanı rasgele veriden hesaplanan özdeğerlerden büyük olduğu noktadır. Üç gerçek veri seti için tüm yöntemlerin buldukları boyut sayıları Çizelge 4.1'de verilmiştir.

Üç gerçek veri seti için PELCI süreci;

- Her bir veri seti ayrı ayrı Σ varyans-kovaryans matrisinden bir örnek çekilir.
- 1000 adet iadeli olarak bootstrap örneği bu örnekten elde edilir ve paralel analiz kullanılarak 1000 adet de paralel örnek elde edilir.
- Bootstrap ve paralel örnekler için özdeğerler hesaplanır.
- Birinci bootstrap özdeğeri ile birinci paralel özdeğer ortalamaları arasındaki farkın güven aralığı iki örneklem ampirik olabilirlik yöntemi ile elde edilir.

- Ampirik olabilirlik yöntemi ile bulunan güven aralığı pozitif ise, bu birinci faktör olarak alınır. Aynı prosedür ikinci özdeğerler içinde uygulanır.
- Bu adımlar negatif güven aralığı bulununcaya dek devam edilirek bulunan pozitif güven aralığı sayısı boyut sayısı olarak alınır.

Süreç her üç veri seti için ayrı ayrı uygulanmıştır ve sonuçları Çizelge 4.1'de verilmiştir.

Çizelge 4. 1 Gerçek veri setleri için için boyut sayısını belirlemede G-K kuralı, PA_{99}, PA_{95} ve PELCI yöntemleri ile belirlenen faktör sayıları

	[1]G-K kuralı	[2]PA_{99}	[3]PA_{95}	[4]PELCI(%95 Güven aralığı)				
Veriler	d	d	d	d	LB1-LR1	LB2-LR2	LB3-LR3	LB4-LR4
Araba Markaları verisi	2	2	2	2	(3.406; 3.42)	(0.277; 0.289)	(-0.844; -0.837)	
WIHCSEM verisi	2	2	2	2	(2.388; 2.418)	(0.27; 0.29)	(-0.005; -0.004)	(-0.091; -0.082)
Boston Housing verisi	3	3	3	3	(5.265; 5.286)	(0.437; 0.450)	(0.180; 0.188)	(-0.205; -0.199)

Not; d: Boyut sayısı
[1]G-K kuralı: özdeğerin 1'de büyük olma kuralı
[2]PA_{99}: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 99. persentil eşik değeri
[3]PA_{95}: PA'ya dayanan rasgele verilerin sıralı özdeğerlerinin 95. persentil eşik değeri
[4]PELCI: İki Örneklem Ampirik Olabilirlik Güven Aralığı
LB: Bootstrap özdeğerleri (LB1, birinci bootstrap özdeğeri)
LR: Paralel analiz özdeğerleri (LR1, birinci paralel analiz özdeğeri)
LB-LR: Bootstrap özdeğerleri ile paralel özdeğerlerin ampirik olabilirlik güven aralıkları

Çizelge 4.1'de sonuçlara bakıldığında dört yöntemde benzer sonuçlar vermiştir. Araba markaları ve WIHCSEM verilerinde tüm yöntemler aynı sonucu belirlemiştir. Boston Housing verisinde ise PA_{95}, PA_{99}, PELCI ve G-K kuralı aynı sonucu belirlemiştir. Bu veri setleri için PELCI yöntemi tatmin edici sonuç ortaya koymuştur.

4.4 Varsayımsal Faktör Modellerinde Önerilen Yaklaşımın Uygulamaları

Bu bölümde, varyans-kovaryans matrisinden manipüle edilerek üretilen verilerin özellikleri yer almaktadır. Değişken sayısı, gözlem sayısı, faktör sayısı, faktör yükü ve faktörler arasında arasındaki korelasyon değerlerine göre varsayımsal modeller oluşturulmuştur. Veri üretimi ve analizinde SAS 9.4 ve R programı kullanılmıştır.

Her bir varsayımsal model için,

Verilerin örneklem birim sayısı: N= 50, 100, 200 ve 400 olarak belirlenmiştir.

Değişken sayısı: Değişken sayısı p=12 ve 18 olarak belirlenmiştir.

Faktör sayısı: Faktör modellerine göre faktör sayısı m=1, 2 ve 3 olarak belirlenmiştir.

Faktör yükleri: Tek faktörlü model için $\Lambda_1 = \ldots = \Lambda_6 = .9,\ \Lambda_7 = \ldots = \Lambda_{12} = .7,$ $\Lambda_{13} = \ldots = \Lambda_{18} = .5$, iki faktörlü model için $\Lambda_{1,1} = \ldots = \Lambda_{1,9} = .7, \Lambda_{2,10} = \ldots = \Lambda_{2,18} = .4,$ üç faktörlü model için $\Lambda_{1,1} = \ldots = \Lambda_{1,6} = .9, \Lambda_{2,7} = \ldots = \Lambda_{2,12} = .7, \Lambda_{3,13} = \ldots = \Lambda_{3,18} = .5$ olarak belirlenmiştir.

Faktör korelasyonları: İki ve üç faktörlü modeller için faktörler arasındaki korelasyon $\Phi = .5$ olarak alınmıştır.

Varsayımsal faktör modelleri,

Model 1: İlk modelde değişken sayısı p=18, faktör sayısı m=1 ve faktör yükleri $\Lambda_1 = \ldots = \Lambda_6 = .9,\ \Lambda_7 = \ldots = \Lambda_{12} = .7,\ \Lambda_{13} = \ldots = \Lambda_{18} = .5$ olarak alınmıştır.

Model 2: ikinci faktör modelinde değişken sayısı p=18, faktör sayısı m=2 ve faktör yükleri $\Lambda_{1,1} = \ldots = \Lambda_{1,9} = .7,\ \Lambda_{2,10} = \ldots = \Lambda_{2,18} = .4,$ olarak alınmıştır.

Model 3: Üçüncü faktör modelinde ise değişken sayısı p=18, faktör sayısı m=2, faktörler arasındaki korelasyon $\Phi = .5$ ve faktör yükleri $\Lambda_{1,1} = \ldots = \Lambda_{1,6} = .7,\ \Lambda_{2,7} = \ldots = \Lambda_{2,12} = .4$ olarak alınmıştır.

Model 4: Dördüncü faktör modelinde ise değişken sayısı p=18, faktör sayısı m=3, faktörler arasındaki korelasyon $\Phi = .5$ ve faktör yükleride $\Lambda_{1,1} = \ldots = \Lambda_{1,6} = .9,$ $\Lambda_{2,7} = \ldots = \Lambda_{2,12} = .7,\ \Lambda_{3,13} = \ldots = \Lambda_{3,18} = .5$ olarak alınmıştır.

Model 5: Beşinci faktör modelinde ise değişken sayısı p=18, faktör sayısı m=3 ve faktör yükleride $\Lambda_{1,1} = \ldots = \Lambda_{1,6} = .9,\ \Lambda_{2,7} = \ldots = \Lambda_{2,12} = .7,\ \Lambda_{3,13} = \ldots = \Lambda_{3.18} = .5$ olarak alınmıştır.

Model 6: Altıncı faktör modelinde ise değişken sayısı P=12, faktör sayısı m=1 ve faktör yükleride $\Lambda_1 = \ldots = \Lambda_4 = .3,\ \Lambda_5 = \ldots = \Lambda_8 = .5,\ \Lambda_9 = \ldots = \Lambda_{12} = .7$ olarak alınmıştır.

Model 7: Yedinci faktör modelinde ise değişken sayısı p=12, faktör sayısı m=2 ve faktör yükleride $\Lambda_{1,1} = \ldots = \Lambda_{1,6} = .7,\ \Lambda_{2,7} = \ldots = \Lambda_{2,12} = .4$ olarak alınmıştır.

Model 8: Sekizinci faktör modelinde ise değişken sayısı p=12, faktör sayısı m=3 ve faktör yükleride $\Lambda_{1,1} = \ldots = \Lambda_{1,4} = .3,\ \Lambda_{2,5} = \ldots = \Lambda_{2,8} = .5,\ \Lambda_{3.9} = \ldots = \Lambda_{3,12} = .7$ olarak alınmıştır.

Model 9: Dokuzuncu faktör modelinde ise değişken sayısı p=12, faktör sayısı m=3, faktörler arasındaki korelasyon $\Phi = .5$ ve faktör yükleride $\Lambda_{1,1} = \ldots = \Lambda_{1,4} = .7$, $\Lambda_{2,5} = \ldots = \Lambda_{2,8} = .5$, $\Lambda_{3,9} = \ldots = \Lambda_{3,12} = .4$ olarak alınmıştır.

Anakütle kovaryans matrisi aşağıdaki eşitlikten elde edilmiştir,

$$\Sigma = \Lambda\Omega\Lambda' + \psi \qquad (4.12)$$

Burada Σ değişkenler arasındaki anakütle kovaryans matrisi, Λ faktör yükleri matrisi, Ω faktörlerin kovaryans matrisi ve ψ ise hatalar arasındaki anakütle kovaryans köşegen matrisidir. Veri üretimi için aşağıdaki adımsal süreç uygulanır,

Her bir faktör modeli için faktör analizinde G-K kuralı ile boyut sayısı hesaplamıştır. PA yöntemi için süreç; her bir faktör modeli için ayrı ayrı için $N(\mu = 0, \Sigma = I)$ olan normal dağılımdan bağımsız değişken sayısına paralel rasgele değişkenlerden kadar veri türetilecektir. Rasgele verilerin özdeğerleri ve özdeğerlerin dağılımının 95. ve 99. persentil değerleri hesaplanır. Her bir orjinal veri setinin özdeğerleri hesaplanır ve rasgele veriden elde edilen özdeğerlerin 95. ve 99. persentil değerleri ile karşılaştırılır. Boyut sayısı için gerçek veriden hesaplanan özdeğerlerin rasgele veriden hesaplanan özdeğerlerden büyük olduğu noktadır. Boyut sayısı için gerçek veriden hesaplanan özdeğerlerin medyanı rasgele veriden hesaplanan özdeğerlerden büyük olduğu noktadır.

Her bir model için veri seti için PELCI süreci;

- Her bir faktör modeli için ayrı ayrı Σ varyans-kovaryans matrisinden bir örnek çekilir.
- 1000 adet iadeli olarak bootstrap örneği bu örnekten elde edilir ve paralel analiz kullanılarak 1000 adet paralel örnek elde edilir.
- Bootstrap ve paralel örneklerden için özdeğerler hesaplanır.
- Birinci bootstrap özdeğeri ile birinci paralel özdeğer ortalamaları arasındaki farkın güven aralığı iki örneklem ampirik olabilirlik yöntemi ile elde edilir.
- Ampirik olabilirlik yöntemi ile bulunan güven aralığı pozitif ise, bu birinci faktör olarak alınır. Aynı prosedür ikinci özdeğerler içinde tekrarlanır.
- Bu adımlar negatif güven aralığı bulununcaya dek devam edilir. Hesaplanan pozitif güven aralığı sayısı boyut sayısı olarak alınır.

4.5 Varsayımsal Faktör Modellerinde Önerilen Yaklaşımın Performans Değerlendirilmesi

PELCI yöntemi her bir faktör modeli için ayrı ayrı uygulanmıştır. Modeller için tüm yöntemlerin faktör analizinde buldukları boyut sayıları Çizelge 4.2'de verilmiştir. Çizelge 4.2'de G-K kuralı, PA_{99}, PA_{95} ve PELCI yöntemlerinin 9 farklı faktör analizi modelinde buldukları faktör sayısı verilmektedir. Çizelge 4.2 incelendiğinde, model 1 faktör sayısı m=1 için sonuçlara bakıldığında faktör analizi için boyut sayısına karar verilmesinde bütün örnek büyüklüklerine göre PELCI ve PA yöntemlerinin doğru boyut sayısını tespit ettikleri görülmektedir. Fakat G-K kuralının örnek birimi N=400 dışında faktör sayısını olması gerekenden fazla sayıda hasapladığı görülmektedir. Model 2 m=2 için bakıldığında doğru boyut sayısı tespitinde PELCI ve PA yöntemlerinin başarılı olduğu görülmektedir. G-K kuralının tüm örneklem sayılarında, boyut sayısını olması gerekenden fazla tespit ettiği görülmektedir. Model 3 m=2 için bakıldığında PA'nın tüm örneklemlerde boyut sayısını tespit ettiği, PELCI'nın örneklem sayısı N=50 için olması gerekenden fazla sayıda boyut bulduğu diğer örneklem sayılarında doğru sayıda boyut hesapladığı görülmektedir. G-K kuralı ise tüm örnek büyüklüklerinde, boyut sayısını olması gerekenden fazla tespit ettiği görülmektedir. Model 4 m=3 için bakıldığında PELCI yöntemi tüm örnek büyüklüklerinde doğru boyut sayısını çok başarılı olarak hesaplamıştır. Örneklem sayısı N=50 ve 100 olduğunda PA yönteminin boyut sayısını olması gerekenden az bulduğu, N=200 ve 400 için doğru tespit ettiği görülmektedir. G-K kuralının örneklem sayısı N=50 ve 100 olduğunda boyut sayısını olması gerekenden fazla hesapladığı, N=200 ve 400 için doğru tespit ettiği görülmektedir. Model 5 m=3 için ise bütün örnek büyüklükleri için PELCI ve PA yöntemleri doğru boyut sayısını tespit ettikleri görülmektedir. G-K kuralının 200 birimden küçük örneklerde (N=50 ve 100 için) boyut sayısını olması gerekenden fazla bulduğu, N=200 ve 400 için ise doğru hesapladığı görülmektedir. Model 6 m=1 için ise PA'nın tüm örnek büyüklüklerinde doğru boyut sayısını tespit ettiği, PELCI'nın N=50 birimlik örneklerde olması gerekenden fazla sayıda boyut bulduğu diğer örnek büyüklüklerinde sayılarında doğru sayıda boyut hesapladığı görülmektedir. G-K kuralının ise tüm örnek büyüklüklerinde, boyut sayısını olması gerekenden fazla tespit ettiği görülmektedir. Model 7 m=2 için ise tüm örnek büyüklüklerinde PELCI ve PA yöntemlerin doğru boyut sayısını tespit

ettikleri görülmektedir. Fakat G-K kuralının örneklem sayısı N=400 dışında faktör sayısını olması gerekenden fazla sayıda hasapladığı görülmektedir. Model 8 m=3 için tüm örnek büyüklüklerinde PELCI ve PA yöntemleri doğru boyut sayısını tespit ettikleri görülmektedir. G-K kuralının örnek büyüklüğü N=50 için olması gerekenden fazla bulduğu, diğer örneklem sayılarında ise boyut sayısını doğru hesapladığı görülmektedir. Model 9 m=3 için PELCI yönteminin doğru boyut sayısı tespitinde diğer yöntemlere göre başarılı olduğu görülmektedir. PELCI yöntemi örnek büyüklüğü N=50 birimde boyut sayısını olmasın gerekenden düşük sayıda bulmuş, diğer örnek büyüklüklerinde doğru hesaplamıştır. PA yöntemi ise tüm örnek büyüklüklerinde boyut sayısını olması gerekenden düşük bulmuş, G-K kuralı N=50 ve 100 birimlik örnekler için boyut sayısını olması gerekenden fazla hesaplamış, diğer örnek büyüklüklerinde boyut sayısını doğru hesaplamıştır.

Çizelge 4. 2 Faktör modellerinde G-K kuralı, PA_{99}, PA_{95} ve PELCI yöntemleri ile boyut sayısı belirleme

Faktör modelleri		[1]G-K Kuralı	[2] PA_{99}	[3] PA_{95}		[4]PELCI(%95 Güven aralıkları)			
	N	d	d	d	d	LB1-LR1	LB2-LR2	LB3-LR3	LB4-LR4
Model 1 m=1	50	3	1	1	1	(7.635; 7.949)	(-0.427; -0.342)		
	100	3	1	1	1	(7.843; 8.086)	(-0.435; -0.379)		
	200	2	1	1	1	(8.023; 8.188)	(-0.380; -0.346)		
	400	1	1	1	1	(8.684; 8.789)	(-0.388; -0.364)		
Model 2 m=2	50	6	2	2	3	(3.052; 3.247)	(0.437; 0.556)	(0.02; 0.1)	(-0.120; -0.071)
	100	4	2	2	2	(3.606; 3.764)	(0.892; 0.984)	(-0.072; 0.019)	
	200	4	2	2	2	(3.502; 3.603)	(0.752; 0.826)	(-0.036; 0.002)	
	400	3	2	2	2	(3.444; 3.514)	(0.949; 1.008)	(-0.121; -0.101)	
Model 3 m=2	50	5	2	2	3	(3.409; 3.583)	(0.429; 0.557)	(0.0009; 0.071)	(-0.141; -0.087)
	100	5	2	2	2	(3.574; 3.727)	(0.355; 0.428)	(-0.019; 0.032)	(-0.138; -0.101)
	200	5	2	2	2	(3.827; 3.944)	(0.269; 0.335)	(-0.07; -0.038)	
	400	3	2	2	2	(3.847; 3.944)	(0.538; 0.586)	(-0.12; -0.098)	
Model 4 m=3	50	4	2	2	3	(3.379; 3.593)	(1.039; 1.182)	(0.194; 0.295)	(-0.143; -0.076)
	100	5	2	2	3	(4.665; 4.857)	(1.022; 1.127)	(0.172; 0.259)	(-0.293; -0.253)
	200	3	3	3	3	(5.177; 5.341)	(0.951; 1.041)	(0.243; 0.3)	(-0.259; -0.229)
	400	3	3	3	3	(5.269; 5.363)	(1.219; 1.287)	(0.406; 0.44)	(-0.244; -0.223)
Model 5 m=3	50	4	3	3	3	(3.284; 3.424)	(1.354; 1.473)	(0.525; 0.625)	(-0.237; -0.171)
	100	4	3	3	3	(3.552; 3.64)	(1.459; 1.555)	(0.546; 0.62)	(-0.221; -0.179)
	200	3	3	3	3	(3.705; 3.765)	(1.813; 1.886)	(0.978; 1.043)	(-0.315; -0.286)
	400	3	3	3	3	(3.809; 3.851)	(1.998; 2.048)	(1.089; 1.133)	(-0.286; -0.267)
Model 6 m=1	50	4	1	1	2	(2.141; 2.335)	(0.034; 0.116)	(-0.126; -0.068)	
	100	3	1	1	1	(2.706; 2.843)	(-0.086; -0.042)		
	200	3	1	1	1	(2.929; 3.021)	(-0.111; -0.078)		
	400	2	1	1	1	(2.573; 2.649)	(-0.107; -0.087)		

Çizelge 4. 2 Faktör modellerinde G-K kuralı, PA_{99} , PA_{95} ve PELCI yöntemleri ile boyut sayısı belirleme (devamı)

Faktör modelleri		[1]G-K Kuralı	[2] PA_{99}	[3] PA_{95}	[4]PELCI(%95 Güven aralıkları)			
	N	d	d	d	d	LB1-LR1	LB2-LR2	LB3-LR3	LB4-LR4
Model 7 m=2	50	3	2	2	2	(1.942; 2.096)	(0.30; 0.416)	(-0.066; -0.005)	
	100	4	2	2	2	(1.865; 1.945)	(0.559; 0.634)	(-0.084; -0.036)	
	200	3	2	2	2	(2.185; 2.251)	(0.867; 0.917)	(-0.202; -0.174)	
	400	2	2	2	2	(2.24; 2.287)	(0.794; 0.835)	(-0.183; -0.164)	
Model 8 m=3	50	4	3	3	3	(1.431; 1.56)	(0.838; 0.943)	(0.086; 0.156)	(-0.101; -0.046)
	100	3	3	3	3	(1.248; 1.333)	(0.805; 0.88)	(0.203; 0.257)	(-0.096; -0.058)
	200	3	3	3	3	(1.295; 1.351)	(1.068; 1.122)	(0.148; 0.192)	(-0.099; -0.069)
	400	3	3	3	3	(1.282; 1.32)	(1.087; 1.122)	(0.335; 0.371)	(-0.152; -0.132)
Model 9 m=3	50	4	1	1	2	(1.674; 1.835)	(0.121; 0.221)	(-0.054; -0.0004)	
	100	4	1	1	3	(1.424; 1.543)	(0.182; 0.246)	(0.006; 0.054)	(-0.072; -0.04)
	200	3	2	2	3	(1.672; 1.754)	(0.382; 0.447)	(0.055; 0.097)	(-0.141; -0.112)
	400	3	2	2	3	(2.04; 2.107)	(0.124; 0.155)	(0.038; 0.065)	(-0.129; -0.11)

Not: [1,2,3,4] açıklamaları Çizelge 4.1'de verilmiştir, d: Boyut sayısı. N: örneklem sayısı
LB: Bootstrap özdeğerleri (örn: LB1, birinci bootstrap özdeğeri)
LR: Paralel analiz özdeğerleri (örn: LR1, birinci paralel analiz özdeğeri)
LB-LR: Bootstrap özdeğerleri ile paralel özdeğerlerin ampirik olabilirlik güven aralıkları.
m: Varsayımsal modellerin faktör sayısı

Genel olarak analiz sonuçlar incelendiğinde önerilen PELCI yöntemi tüm modellerde tatmin edici bir performansı göstermiştir. Diğer iki yöntemle karşılaştırıldığında ise, genel olarak diğer yöntemlerden daha iyi performans göstermiştir. Sadece dokuz modelin üçünde örnek büyüklüğü N=50 olduğunda doğru boyut sayısını hatalı tespit ettiği görülmektedir. Araştırmacılar tarafından önerilen PA_{99} ve PA_{95} yöntemlerinin korelasyonlu modellerde performanslarının düştüğü görülmektedir. PELCI yönteminin genel olarak araştırmacılar tarafından önerilen PA_{99} ve PA_{95} daha iyi sonuçlar hesapladığı söylenilebilir. G-K kuralının örnek büyüklüğünden oldukça etkilendiği görülmektedir.

BÖLÜM 5

SONUÇ ve ÖNERİLER

Araştırmada ilk olarak çok değişkenli istatistiksel yöntemlerden biri olan kanonik korelasyon analizinde boyut sayısının belirlenmesinde paralel analiz (PA) ve bootstrap modifiyeli paralel analiz (BPA) yöntemlerinin kullanılması önerilmiştir. KKA'da boyut sayısı hesaplamada daha önceki çalışmalarda en çok önerilen yöntemlerden altı adet χ^2 yaklaşım prosedürü; Barttlett'in χ^2 yaklaşımı $S_B\left(T_d^2\right)$, Fujikoshi χ^2 yaklaşımı $S_F\left(T_d^2\right)$, Barttlett'in χ^2 yaklaşımı $S_B\left(\Lambda_d\right)$, Lawley χ^2 yaklaşımı $S_L\left(\Lambda_d\right)$, Barttlett'in χ^2 yaklaşımı $S_B\left(V_d\right)$, Fujikoshi χ^2 yaklaşımı $S_F\left(V_d\right)$ ile PA ve BPA yöntemlerinin performansları gerçek bir veride ve geniş simülasyon çalışmalarıyla karşılaştırılmıştır. Bu çalışmada bu altı hipotez testi yaklaşımı arasında en iyi performansı Fujikoshi'nin $S_F\left(T_d^2\right)$ ve Bartlett'in $S_B\left(T_d^2\right)$ testleri göstermiştir. Tez çalışmamızda PA ve BPA yöntemlerinin KKA yöntemine adaptasyonları yapılarak, KKA'da boyut sayısı tespitinde uygulanabilirlikleri araştırılmıştır. KKA'da boyut sayısı belirlemede PA ve BPA'nın sırasıyla $PA_{ortalama}$, PA_{medyan}, PA_{95}, PA_{99} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{95} ve BPA_{99} kullanılarak, PA ve BPA'nın yüksek persentillerinin PA_{95}, PA_{99} ve BPA_{95}, BPA_{99} kullanımı önerilmiştir. Gerçek veri uygulamasında ve simülasyon çalışmaları sonucunda PA ve BPA yöntemlerinin sonuçları tatmin edici şekilde bulunmuştur. Boyut sayısı tespitinde PA ve BPA'nın özellikle kanonik korelasyonların (en az bir tane $\rho \leq 0.5$) olduğu koşullarda performansı dikkat çekmektedir. KKA'da boyut sayısı belirlenmesinde PA ve BPA yöntemleri, klasik test istatistiklerine (Hipotez testleri) altenatif yöntem olarak kullanılabileceği önerilmektedir. KKA'da boyut sayısı belirleme probleminin çözümüne PA ve BPA'nın katkı sunacağı düşünülmektedir.

İkinci olarak faktör analizinde faktör sayısının tespitinde BPA yöntemi geliştirilmiş ve PA ve BPA yöntemlerinin $PA_{ortalama}$, PA_{medyan}, PA_{99}, $PA_{97.5}$, PA_{95}, PA_{90} ve $BPA_{ortalama}$, BPA_{medyan}, BPA_{99}, $BPA_{97.5}$, BPA_{95}, BPA_{90} persentil değerleri kullanılmıştır. Faktör analizinde boyut sayısı belirlemede klasik yöntemlerden G-K kuralı yöntemi ile PA ve BPA yöntemleri üç gerçek veri seti için uygulanmış ve simülasyon çalışmalarıyla performansları karşılaştırılmıştır. PA ve BPA yöntemlerinin klasik yöntemlerden çok daha iyi sonuçlar bulduğu görülmüştür. BPA yönteminin özellikle korelasyonlu faktör modellerindeki başarısı dikkat çekmektedir.

Son bölümde ise faktör analizinde faktör sayısının belirlenmesinde paralel analizden ve bootstrap örneklerinden elde edilen veri ile oluşturulan iki örneklem ampirik olabilirlik güven aralığına dayanan yönteminin kullanılabilirliği araştırılmış, önerilen bu yöntem PELCI olarak adlandırılmıştır. Faktör analizinde boyut belirlenmesi için PELCI, G-K kuralı, PA_{99} ve PA_{95} yöntemleri üç gerçek veri seti için uygulanmış ve varsayımsal 9 farklı faktör modelinde performansları incelenmiştir. PELCI yöntemi gerçek veri setlerinde ve varsayımsal faktör modellerinde iyi performans göstermiştir. PELCI yönteminin korelasyonlu modellerde doğru boyut sayısı tespitinde başarısı dikkat çekicidir.

KKA'da kanonik korelasyon sayısı belirlemek için hipotez testlerinin yanı sıra PA ve BPA yöntemlerinin kullanılması önerilmiştir.

Faktör analizinde boyut sayısı belirlemek için BPA ve PELCI yöntemleri oluşturulmuştur ve faktör sayısı belirleme için kullanılmaları önerilmiştir.

Bundan sonra yapılacak araştırmalarda, PA ve BPA yöntemlerinin ayırma analizi (Diskriminant analizi) yöntemine adapte edilip ayırma fonksiyonu sayısının belirlenmesinde kullanılması önerilebilir. Ayrıca BPA yönteminin uyum analizi yöntemine adapte edilip boyut sayısı belirlemede uygulanması önerilebilir.

KAYNAKLAR

[1] Reddon, J. R., Marceau, R. ve Jackson, D. N. (1982). "An application of singular value decomposition to the factor analysis of MMPI items", Applied Psychological Measurement 6, pp. 275-283.

[2] Yavuz, G. & Doğan, N. (2015). "Boyut sayısı belirlemede Velicer'in map testi ve Horn'un paralel analizinin kullanılması", Hacettepe Üniversitesi Eğitim Fakültesi Dergisi [Hacettepe University Journal of Education], 30(3), 176-188.

[3] Alpar, C. R. (2011) Uygulamalı Çok Değişkenli İstatistiksel Yöntemler, Detay Yayıncılık, 3. Baskı, Ankara.

[4] Crawford, A. V., Green, S. B., Levy, R., Lo, W., Scott, L., Svetina, D. ve Thompson, M. S., (2010). "Evaluation of parallel analysis methods for determining the number of factors", Educational and Psychological Measurement, 70, pp. 885–901.

[5] Dinno, A., (2009). "Implementing Horn's parallel analysis for principal component analysis and factor analysis", Stata Journal, 9, pp. 291–298.

[6] Hayton, J. C., Allen, D. G. ve Scarpello, V. (2004). "Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis", Organizational Research Methods 7, pp. 191–205.

[7] Horn, J. L. (1965), "A rationale and test for the number of factors in factor analysis", Psychometrika 30, pp. 179-185.

[8] Humphreys, L. G. ve Montanelli, R. G. (1975). "An investigation of the parallel analysis criterion for determining the number of common factors", Multivariate Behavioral Research 10, pp. 193-206.

[9] Lorenzo-Seva, U., Timmerman, M. E. ve Kiers, A. L. (2011). "The hull method for selecting the number of common factors", Multivariate Behavioral Research 46, pp. 340-364.

[10] O'Connor, B. P. (2000). "SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test", Behavior Research Methods, Instruments, & Computers 32, pp. 396–402.

[11] Parmet, Y., Schechtman, E. ve Sherman, M. (2010). "Factor analysis revisited-How many factors are there?", Communication in Statistics-Simulation and Computation 39, 1893-1908.

[12] Turner, N. E. (1998). “The effect of common variance and structure pattern on random data eigenvalues: Implications for the accuracy of parallel analysis”, Educational and Psychological Measurement 58, pp. 541-568.

[13] Velicer, W. F., Eaton, C. A. ve Fava, J. L. (2000). “Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components”, In R. D. Goffin & E. Helmes (Eds.), Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy (pp. 41– 71). New York, NY: Kluwer Academic/Plenum Publishers.

[14] Kaiser, H. F. (1960). “The application of electronic computers to factor analysis”, Educational and Psychological Measurement 20, 141-151.

[15] Guttman, L. (1954). “Some necessary conditions for common factor analysis”, Psychometrika 19, pp. 149-161.

[16] Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. ve. Strahan, E. J. (1999). “Evaluating the use of exploratory factor analysis in psychological research”, Psychological Methods 4, pp. 272–299.

[17] Cliff, N. (1988). “The eigenvalues-greater-than-one rule and the reliability of components”, Psychological Bulletin, 103(2), 276-279.

[18] Ledesma R. D. ve Valero-Mora, P. (2007). “Determining the number of factors to retain in EFA: an easy-to-use computer program for carrying out parallel analysis”, Practical Assessment Research & Evaluation 12(2), pp. 1-10.

[19] Zwick, W. R. ve Velicer, W. F. (1986). “Factor influencing five rules for determining the number of components to retain”, Psychological Bulletin 99, pp. 432-442.

[20] Piconne, A. V. (2009). A Comprasion of Three Computational Procedures For Solving the Number of Factor Problem in Exploratory Factor Analysis, Doktora Tezi, University of Northern Colorado, Colorado.

[21] Cho, S. Li, F. ve Bandalos, D., (2009). “Accuracy of the parallel analysis procedure with polychoric correlations”, Educational and Psychological Measurement 69, pp. 748–759.

[22] Glorfeld, L. W. (1995). “An improvement on Horn’s parallel analysis methodology for selecting the correct number of factors to retain”, Educational and Psychological Measurement 55, pp. 377-393.

[23] Henson, R. K. ve Roberts, J. K. (2006). “Use of exploratory factor analysis in published research: Common errors and some comment on improved practice”, Educational and Psychological Measurement 66, pp. 393–416.

[24] Peres-Neto, P., Jackson, D. ve Somers, K. (2005). “How many principal components? Stopping rules for determining the number of non-trivial axes revisited”, Computational Statistics & Data Analysis 49, pp. 974–997.

[25] Weng, L. ve Cheng, C. (2005). “Parallel analysis with unidimensional binary data”, Educational and Psychological Measurement 65, pp. 697–716.

[26] Buja A. ve Eyuboglu, N. (1992). “Remarks on parallel analysis”, Multivariate Behavioral Research 27, pp. 509-540.

[27] Humphreys, L. G. ve Ilgen, D. R. (1969). “Note on a criterion for the number of common factors”, Educational and Psychological Measurement 29, pp. 571–578.

[28] Timmerman M. E. ve Lorenzo-Seva, U. (2011). “Dimensionality assessment of ordered polytomous items with parallel analysis”, Psychological Methods 16, pp. 209–220.

[29] Tran U. S. ve Formann, A. K. (2009). “Performance of parallel analysis in retrieving unidimensionality in the presence of binary data”, Educational and Psychological Measurement 69, pp. 50 – 61.

[30] Garrido, L. E., Abad, F. J. ve Ponsoda, V. (2013). “A new look at Horn’s parallel analysis with ordinal variables”, Psychological Methods 18(4), pp. 454-474.

[31] Liu O. L. ve Rijmen, F. (2008). “A modified procedure for parallel analysis of ordered categorical data”, Behavior Research Methods 40(2), pp. 556-562.

[32] Dinno, A., (2009). “Exploring the sensitivity of Horn’s parallel analysis to the distributional form of random data”, Multivariate Behavioral Research 44(3), pp. 362-388.

[33] Şimşek, G. G., ve Aydoğdu S. (2016). “Parallel analysis approach for determining dimensionality in canonical correlation analysis”, Journal of Statistical Computation and Simulation. 86(17),3419-3431.

[34] Hotelling, H. (1936). “Relations between two sets of variates”, Biometrika 28, pp. 321-377.

[35] Timm, N. H. (2002). Applied Multivariate Analysis, New York: Springer-Verlag.

[36] Mendoza, J. L., Markos, V. H. ve Gonter, R. (1978). “A new perspective on sequential testing procedures in canonical analysis”, A Monte Carlo evaluation. Multivariate Behavioral Research 13, pp. 371-382.

[37] Harris, R. J. (1976), “The invalidity of partitioned-U tests in canonical correlation and multivariate analysis of variance”, Multivariate Behavioral Research 11, pp. 353-365.

[38] Lawley, D. N. (1959). “Tests of significance in canonical analysis”, Biometrika 28, pp. 321–377.

[39] Fujikoshi, Y. ve Veitch, L. G. (1979). “Estimation of dimensionality in canonical correlation analysis”, Biometrika 66, pp. 345-351.

[40] Gunderson, B. K. ve Muirhead, R. J. (1997). “On estimating the dimensionality in canonical correlation analysis” Journal of Multivariate Analysis 62, pp. 121-136.

[41] Fan, X. (1997). “Canonical correlation analysis and structural equation modeling: What do they have in common?”, Structural Equation Modeling. 4(1), pp. 65-79.

[42] Caliński T. ve Krzyśko, M. (2005). “A closed testing procedure for canonical correlations” Communications in Statistics-Theory and Methods 34, pp. 1105-1116.

[43] Marcus, R., Peritz, E. ve Gabriel, K. R. (1976). “On closed testing procedures with special reference to ordered analysis of variance”, Biometrika 63, pp. 655–660.

[44] Caliński, T., Krzyśko, M. ve Wołyński, W.A., (2006). “A comparison of some tests for determining the number of nonzero canonical correlations”, Communication in Statistics-Simulation and Computation 35, pp. 727-749.

[45] Yamada, T. ve Sugiyama, T. (2006). “On the permutation test in canonical correlation analysis”, Computational Statistics & Data Analysis 50, pp. 2111-2123.

[46] Aoki, K. ve Sato, Y. (2007). “A test statistic in canonical correlation analysis for categorical variables”, Behaviormetrika, 34(1), pp. 59-74.

[47] Lee, H. S. (2007). “Canonical correlation analysis using small number of samples”, Communications in Statistics-Simulation and Computation 36(5), pp. 973-985.

[48] Fujikoshi, Y., Ulyanov, V.V. ve Shimizu, R. (2010). “Multivariate Statistics: High-Dimensional and Large-Sample Approximations”, Hoboken: John Wiley & Sons, Inc.

[49] Bartlett, M. S. (1938). “Further aspects of the theory of multiple regression”, Proceedings of the Cambridge philosophical society. 34, pp. 33-40.

[50] Bartlett, M. S. (1941). “The statistical significance of canonical correlations”, Biometrika 32, pp. 29-37.

[51] Bartlett, M. S. (1947). “Multivariate analysis (with discussion)”, Journal of the Royal Statistical Society Supplement. 9(B), pp. 176-197.

[52] Fujikoshi, Y. (1977). “Asymptotic expansion of the distributions of the latent roots in MANOVA and the canonical correlations”, Journal of Multivariate Analysis 7, pp.386–396.

[53] Rao, C. R. (1951). “An asymptotic expansion of the distribution of Wilks’ criterion”, Bulletin of the Institute of International Statistics 33(2), pp. 177-180.

[54] Saccenti, E. ve Timmerman M. E (2016). “Considiring Horn’s Parallel Analysis from a Random Matrix Theory of Point Wiev”, Psychometrika, doi: 10.1007/s11336-016-9515-z.

[55] Efron, B. ve Tibshirani, R. J. (1994). An Introduction to the Bootstrap, Boca Raton: Chapman & Hall/CRC.

[56] Rencher, A., ve Christensen, W. F. (2012). Methods of multivariate analysis, 3rd edition, Wiley, p. 390

[57] Larsen, R. ve Warne, R. T. (2010).“ Estimating confidence intervals for eigenvalues in exploratory factor analysis”, Behavior Research Methods. 42 (3), pp. 871-876.

[58] Büyüköztürk, Ş. (2002). “Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı”, Kuram ve Uygulamada Eğitim Yönetimi, 32, 470-483.

[59] Uyar, S. (2012). Açımlayıcı Faktör Analizinde Boyut Sayısını Belirlemede Kullanılan Yöntemlerin Karşılaştırılması (Yayınlanmamış yüksek lisans tezi.), Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü. Ankara.

[60] Jackson, D. A. (1993). “Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches”, Ecology 74, pp. 2204-2214.

[61] Warne, R. T. ve Larsen, R. (2014).“ Evaluating a proposed modification of the Guttman rule for determining the number of factors in an exploratory factor analysis”, Psychological Test and Assessment Modeling 56, pp. 104-123.

[62] Zientek, L. R. ve Thompson, B. (2007). “Applying the bootstrap to the multivariate case: Bootstrap component/factor analysis”, Behavior Research Methods 39(2), pp. 318-325.

[63] Härdle, W., & Hlávka, Z. (2007). Multivariate statistics, Springer, Berlin.

[64] Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate analysis, California State University Northridge: Harper Collins College Publishers.

[65] Owen, A. (1988). “Empirical likelihood ratio confidence intervals for a single functional”, Biometrika, 75 (2), 237–249.

[66] Wu F. ve Tsao M. (2014). “Two-sample extended empirical likelihood”, Statistics & Probability Letters, 84; 81-87.

[67] Qin, J. (1994). “Semi-empirical likelihood ratio confidence intervals for the difference of two sample means”, Annals of the Institute of Statistical Mathematics, 46(1), 117-126.

[68] Wu, C., & Yan, Y. (2012). “Empirical likelihood inference for two-sample problems”, Stat. Interface, 5, 345-354.

[69] Claeskens, G., Jing, B. Y., Peng, L., & Zhou, W. (2003). “Empirical likelihood confidence regions for comparison distributions and ROC curves”, Canadian Journal of Statistics”, 31(2), 173-190.

[70] Jing, B. Y., (1995). “Two-sample empirical likelihood method”, Statistics and Probability Letters, 24, 315–319.

[71] Valeinis J. ve Cers E. (2010). “Extending the two-sample empirical likelihood. Mathematical Modeling and Analysis”, 15(1):137-151

[72] Tsao, M. (2013). “Extending the empirical likelihood by domain expansion” The Canadian Journal of Statistics, 41 (2), 257-274.

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri

		Model 1: $p = 4,\ q = 4,\ \rho_1 = \rho_2 = \rho_3 = \rho_4 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik						
	Doğru	91	91	95.1	95.1	97.6	97.6
	Fazla	9	9	4.9	4.9	2.4	2.4
50	Eksik						
	Doğru	93.3	93.3	94.7	94.7	96.8	96.8
	Fazla	6.7	6.7	5.3	5.3	3.2	3.2
75	Eksik						
	Doğru	95.1	95.1	96.1	96.1	96.7	96.7
	Fazla	4.9	4.9	3.9	3.9	3.3	3.3
100	Eksik						
	Doğru	95	95	96.2	96.2	96.8	96.8
	Fazla	5	5	3.8	3.8	3.2	3.2
Gen	Eksik						
	Doğru	93.6	93.6	95.53	95.53	96.98	96.98
	Fazla	6.4	6.4	4.48	4.48	3.03	3.03

		Model 1: $p = 4,\ q = 4,\ \rho_1 = \rho_2 = \rho_3 = \rho_4 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik								
	Doğru	51.6	49.7	94	98.5	55.1	53	98.3	99.8
	Fazla	48.4	50.3	6	1.5	44.9	47	1.7	0.2
50	Eksik								
	Doğru	55.4	51.3	94.2	98.7	65.5	59.9	97.8	99.1
	Fazla	44.6	48.7	5.8	1.3	34.5	40.1	2.2	0.9
75	Eksik								
	Doğru	53.6	49.2	95.7	98.9	67.7	61.1	99.6	100
	Fazla	46.4	50.8	4.3	1.1	32.3	38.9	0.4	
100	Eksik								
	Doğru	58	52.4	95	99.1	71.8	63.4	99.8	100
	Fazla	42	47.6	5	0.9	28.2	36.6	0.2	
Gen	Eksik								
	Doğru	54.65	50.65	94.73	98.8	65.03	59.35	99.2	99.93
	Fazla	45.35	49.35	5.28	1.2	34.98	40.65	0.8	0.08

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 2 :, $p = 4,\ q = 4,\ \rho_1 = .8,\ \rho_2 = \rho_3 = \rho_4 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B\left(T_d^2\right)$	[2] $S_F\left(T_d^2\right)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	5.5	5.5	12.8	12.8	49	49
	Doğru	90.4	88	82.8	82.4	48.1	49
	Fazla	4.1	6.5	4.4	4.8	2.9	2
50	Eksik					0.2	
	Doğru	93.7	92.6	94.2	94	94.8	0.2
	Fazla	6.3	7.4	5.8	6	5	95
75	Eksik						
	Doğru	94.2	93.9	94.6	94.5	94.9	95.2
	Fazla	5.8	6.1	5.4	5.5	5.1	4.8
100	Eksik						
	Doğru	94.1	93.4	94.5	94.2	94.8	94.9
	Fazla	5.9	6.6	5.5	5.8	5.2	5.1
Gen	Eksik	1.38	1.38	3.2	3.2	12.3	12.3
	Doğru	93.1	91.98	91.53	91.28	83.15	83.53
	Fazla	5.53	6.65	5.28	5.53	4.55	4.18

		Model 2 $p = 4,\ q = 4,\ \rho_1 = .8,\ \rho_2 = \rho_3 = \rho_4 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	0.1	0.1	9.3	26.2	2	2	57.2	67.9
	Doğru	25.3	22.9	66.8	64	43.6	38.9	42	31.6
	Fazla	74.6	77	23.9	9.8	54.4	59.1	0.8	0.5
50	Eksik					0.6	0.6	0.7	2.5
	Doğru	25.4	22.7	69.1	86.5	47.4	39.5	94.7	96.4
	Fazla	74.6	77.3	30.9	13.5	52	59.9	4.6	1.1
75	Eksik					0.9	0.9	0.9	0.9
	Doğru	25.8	21.5	70.2	84.3	50.2	41.7	94.8	98.3
	Fazla	74.2	78.5	29.8	15.7	48.9	57.4	4.3	0.8
100	Eksik					0.5	0.5	0.5	0.5
	Doğru	25	21.3	71.7	85.9	50	41.7	95.8	98.9
	Fazla	75	78.7	28.3	14.1	49.5	57.8	3.7	0.6
Gen	Eksik	0.03	0.03	2.33	6.55	1	1	7.45	15.28
	Doğru	25.38	22.1	69.45	80.18	47.8	40.45	88.28	83.9
	Fazla	74.6	77.8	28.23	13.28	51.2	58.55	4.28	0.83

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 3 : $p = 4, q = 4, \rho_1 = .7,\ \rho_2 = \rho_3 = \rho_4 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	23.4	23.4	34.9	34.9	66.8	66.8
	Doğru	72.8	70.5	61.8	61.3	31	31.5
	Fazla	3.8	6.1	3.3	3.8	2.2	1.7
50	Eksik	1.2	1.2	2	2	5	5
	Doğru	93.9	93.1	93.5	93.1	91.1	91.2
	Fazla	4.9	5.7	4.5	4.9	3.9	3.8
75	Eksik						
	Doğru	94.6	94	94.8	94.5	95.1	95.2
	Fazla	5.4	6	5.2	5.5	4.9	4.8
100	Eksik						
	Doğru	96.1	95.7	96.3	96	96.7	96.7
	Fazla	3.9	4.3	3.7	4	3.3	3.3
Gen	Eksik	6.15	6.15	9.23	9.23	17.95	17.95
	Doğru	89.35	88.33	86.6	86.23	78.48	78.65
	Fazla	4.5	5.53	4.18	4.55	3.58	3.4

		Model 3 : $p = 4, q = 4, \rho_1 = .7,\ \rho_2 = \rho_3 = \rho_4 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	2.9	2.8	31.7	58.5	5.2	4.8	56.6	83.5
	Doğru	25	21.3	51.3	35.4	43.2	37.7	39.4	16
	Fazla	72.1	75.9	17	6.1	51.6	57.5	4	0.5
50	Eksik			1.4	4.1	1.3	1.2	8.4	26.9
	Doğru	24.8	22.1	70.5	83.2	51.6	44.4	88.5	72.6
	Fazla	75.2	77.9	28.1	12.7	47.1	54.4	3.1	0.5
75	Eksik				0.1	0.9	0.9	1	3.6
	Doğru	22.7	19	71.1	86.2	48.8	41.8	96.1	95.9
	Fazla	77.3	81	28.9	13.7	50.3	57.3	2.9	0.5
100	Eksik					0.5	0.5	0.5	0.6
	Doğru	23.8	19.7	70.7	87.3	49.2	41.4	96.9	98.9
	Fazla	76.2	80.3	29.3	12.7	50.3	58.1	2.6	0.5
Gen	Eksik	0.73	0.7	8.28	15.68	1.98	1.85	16.63	28.65
	Doğru	24.08	20.5	65.9	73.03	48.2	41.33	80.23	70.85
	Fazla	75.2	78.7	25.83	11.3	49.83	56.83	3.15	0.5

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 4 : $p=4, q=4, \rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B\left(T_d^2\right)$	[2] $S_F\left(T_d^2\right)$	[3] $S_B\left(\Lambda_d\right)$	[4] $S_L\left(\Lambda_d\right)$	[5] $S_B\left(V_d\right)$	[6] $S_F\left(V_d\right)$
25	Eksik	29.50	24.20	35.40	33.70	50.20	55.50
	Doğru	68.10	70.80	61.50	62.10	46.30	41.00
	Fazla	2.40	5.00	3.10	4.20	3.50	3.50
50	Eksik	0.50	0.50	0.50	0.50	0.70	0.80
	Doğru	95.10	93.50	94.60	94.40	94.40	94.40
	Fazla	4.40	6.00	4.90	5.10	4.90	4.80
75	Eksik						
	Doğru	95.00	93.90	94.80	94.30	94.80	94.80
	Fazla	5.00	6.10	5.20	5.70	5.20	5.20
100	Eksik						
	Doğru	95.90	94.80	95.80	95.20	95.70	95.50
	Fazla	4.10	5.20	4.20	4.80	4.30	4.50
Gen	Eksik	7.50	6.18	8.98	8.55	12.73	14.08
	Doğru	88.53	88.25	86.68	86.50	82.80	81.43
	Fazla	3.98	5.58	4.35	4.95	4.48	4.50

		Model 4 : $p=4, q=4, \rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	0.30	0.20	7.20	24.00	3.60	3.00	34.20	62.50
	Doğru	21.20	16.30	55.10	56.70	72.90	65.10	63.60	36.90
	Fazla	78.50	83.50	37.70	19.30	23.50	31.90	2.20	0.60
50	Eksik				0.20	1.10	1.10	1.70	5.40
	Doğru	22.10	17.60	58.50	74.80	74.60	65.10	96.10	94.40
	Fazla	77.90	82.40	41.50	25.00	24.30	33.80	2.20	0.20
75	Eksik					0.90	0.90	0.90	1.10
	Doğru	19.90	15.40	57.40	74.80	74.20	65.10	96.20	98.00
	Fazla	80.10	84.60	42.60	25.20	24.90	34.00	2.90	0.90
100	Eksik					0.80	0.80	0.80	0.80
	Doğru	20.00	15.50	58.00	73.70	74.90	67.60	96.60	98.40
	Fazla	80.00	84.50	42.00	26.30	24.30	31.60	2.60	0.80
Gen	Eksik	0.08	0.05	1.80	6.05	1.60	1.45	9.40	17.45
	Doğru	20.80	16.20	57.25	70.00	74.15	65.73	88.13	81.93
	Fazla	79.13	83.75	40.95	23.95	24.25	32.83	2.48	0.63

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 5 : $p=4, q=4,\ \rho_1=.7,\ \rho_2=.5\ \rho_3=\rho_4=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	79.50	73.30	81.70	79.80	87.90	89.60
	Doğru	19.90	24.80	17.50	18.80	11.30	9.30
	Fazla	0.60	1.90	0.80	1.40	0.80	1.10
50	Eksik	30.30	26.60	32.90	31.60	37.20	37.80
	Doğru	67.50	68.90	64.60	64.40	59.90	58.70
	Fazla	2.20	4.50	2.50	4.00	2.90	3.50
75	Eksik	9.60	8.50	10.40	10.00	11.70	11.80
	Doğru	87.10	86.40	86.30	85.40	85.00	84.00
	Fazla	3.30	5.10	3.30	4.60	3.30	4.20
100	Eksik	2.00	1.80	2.40	2.20	2.60	2.50
	Doğru	94.00	93.20	93.60	93.00	93.30	92.80
	Fazla	4.00	5.00	4.00	4.80	4.10	4.70
Gen	Eksik	30.35	27.55	31.85	30.90	34.85	35.43
	Doğru	67.13	68.33	65.50	65.40	62.38	61.20
	Fazla	2.53	4.13	2.65	3.70	2.78	3.38

		Model 5 : $p=4, q=4,\ \rho_1=.7,\ \rho_2=.5\ \rho_3=\rho_4=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	6.10	5.60	53.70	78.40	18.30	14.20	81.80	95.50
	Doğru	25.20	20.00	31.30	17.00	71.30	70.60	17.90	4.50
	Fazla	68.70	74.40	15.00	4.60	10.40	15.20	0.30	
50	Eksik	0.40	0.40	6.30	14.20	3.20	2.30	36.20	63.20
	Doğru	21.90	15.90	56.30	64.00	78.60	72.10	63.20	36.70
	Fazla	77.70	83.70	37.40	21.80	18.20	25.60	0.60	0.10
75	Eksik	0.10	0.10	1.00	3.30	1.80	1.50	14.70	32.50
	Doğru	19.80	14.10	58.20	72.00	79.20	69.70	83.80	67.30
	Fazla	80.10	85.80	40.80	24.70	19.00	28.80	1.50	0.20
100	Eksik				0.40	0.80	0.80	4.00	9.90
	Doğru	20.60	14.90	59.20	77.20	79.20	71.20	94.50	89.90
	Fazla	79.40	85.10	40.80	22.40	20.00	28.00	1.50	0.20
Gen	Eksik	2.00	1.53	15.25	24.08	6.03	4.70	34.18	50.28
	Doğru	94.90	16.23	51.25	57.55	77.08	70.90	64.85	49.60
	Fazla	3.10	82.25	33.50	18.38	16.90	24.40	0.98	0.13

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 6 : $p=4, q=4, \rho_1=.8,\ \rho_2=.7,\ \rho_3=.5, \rho_4=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	74.50	61.60	71.20	67.60	70.60	73.50
	Doğru	24.40	35.30	26.80	29.40	26.50	23.50
	Fazla	1.10	3.10	2.00	3.00	2.90	3.00
50	Eksik	20.90	16.80	20.40	18.50	20.30	20.80
	Doğru	76.00	78.80	76.20	77.10	75.60	74.90
	Fazla	3.10	4.40	3.40	4.40	4.10	4.30
75	Eksik	3.20	2.30	3.20	3.10	3.10	3.30
	Doğru	93.10	93.10	92.70	92.40	92.40	92.20
	Fazla	3.70	4.60	4.10	4.50	4.50	4.50
100	Eksik	1.10	0.80	1.10	1.00	1.10	1.10
	Doğru	94.70	94.20	94.40	94.00	94.10	93.90
	Fazla	4.20	5.00	4.50	5.00	4.80	5.00
Gen	Eksik	24.93	20.38	23.98	22.55	23.78	24.68
	Doğru	72.05	75.35	72.53	73.23	72.15	71.13
	Fazla	3.03	4.28	3.50	4.23	4.08	4.20

		Model 6 : $p=4, q=4, \rho_1=.8,\ \rho_2=.7,\ \rho_3=.5, \rho_4=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	3.00	2.20	20.10	38.80	34.40	26.10	81.40	93.70
	Doğru	38.70	25.00	51.80	47.00	65.60	73.60	18.60	6.30
	Fazla	58.30	72.80	28.10	14.20		0.30		
50	Eksik	0.20	0.20	1.30	3.90	6.60	4.40	30.50	49.80
	Doğru	37.50	25.20	64.50	76.80	92.60	93.90	69.20	50.10
	Fazla	62.30	74.60	34.20	19.30	0.80	1.70	0.30	0.10
75	Eksik				0.20	0.90	0.70	7.00	15.20
	Doğru	33.80	23.90	62.30	76.70	97.60	96.70	92.50	84.60
	Fazla	66.20	76.10	37.70	23.10	1.50	2.60	0.50	0.20
100	Eksik				0.10	0.60	0.50	1.90	3.90
	Doğru	33.80	22.70	62.50	76.10	97.80	96.80	97.80	96.00
	Fazla	66.20	77.30	37.50	23.80	1.60	2.70	0.30	0.10
Gen	Eksik	0.80	0.60	5.35	10.75	10.63	7.93	30.20	40.65
	Doğru	35.95	24.20	60.28	69.15	88.40	90.25	69.53	59.25
	Fazla	63.25	75.20	34.38	20.10	0.98	1.83	0.28	0.10

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

z		Model 7 : $p=4, q=4,\ \rho_1=.7,\ \rho_2=.5,\ \rho_3=.4,\ \rho_4=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	95.30	89.50	94.60	91.80	93.30	93.80
	Doğru	4.50	9.50	5.00	7.30	5.80	5.30
	Fazla	0.20	1.00	0.40	0.90	0.90	0.90
50	Eksik	63.90	55.50	62.80	57.90	62.20	60.80
	Doğru	34.60	41.50	35.60	39.10	36.00	36.40
	Fazla	1.50	3.00	1.60	3.00	1.80	2.80
75	Eksik	28.80	23.60	28.40	24.70	28.00	23.60
	Doğru	67.70	70.90	67.10	69.80	67.20	70.90
	Fazla	3.50	5.50	4.50	5.50	4.80	5.50
100	Eksik	10.60	9.30	10.60	9.60	10.60	10.00
	Doğru	85.70	85.90	85.50	85.60	85.10	85.20
	Fazla	3.70	4.80	3.90	4.80	4.30	4.80
Gen	Eksik	49.65	44.48	49.10	46.00	48.53	47.80
	Doğru	48.13	51.95	48.30	50.45	48.53	48.73
	Fazla	2.23	3.58	2.60	3.55	2.95	3.48

		Model 7 : $p=4, q=4,\ \rho_1=.7,\ \rho_2=.5,\ \rho_3=.4,\ \rho_4=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	13.50	10.00	59.90	83.20	70.90	61.50	98.00	99.60
	Doğru	36.70	26.20	27.10	13.10	29.10	38.50	2.00	0.40
	Fazla	49.80	63.80	13.00	3.70				
50	Eksik	2.40	1.90	14.30	29.80	34.60	25.90	80.90	92.60
	Doğru	38.80	26.40	57.60	56.40	65.10	73.50	19.10	7.40
	Fazla	58.80	71.70	28.10	13.80	0.30	0.60		
75	Eksik	0.30	0.20	2.50	5.40	9.10	5.90	47.40	67.80
	Doğru	34.30	22.50	60.50	72.30	90.10	92.30	52.50	32.20
	Fazla	65.40	77.30	37.00	22.30	0.80	1.80	0.10	
100	Eksik	0.10	0.10	0.60	1.70	3.00	2.30	21.80	37.80
	Doğru	34.30	23.20	61.60	76.00	96.00	95.50	78.10	62.20
	Fazla	65.60	76.70	37.80	22.30	1.00	2.20	0.10	
Gen	Eksik	4.08	3.05	19.33	30.03	29.40	23.90	62.03	74.45
	Doğru	36.03	24.58	51.70	54.45	70.08	74.95	37.93	25.55
	Fazla	59.90	72.38	28.98	15.53	0.53	1.15	0.05	

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 8 : $p=4, q=8, \rho_1 = \rho_2 = \rho_3 = \rho_4 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik						
	Doğru	88.20	88.20	94.60	94.60	94.60	99.30
	Fazla	11.80	11.80	5.40	5.40	5.40	0.70
50	Eksik						
	Doğru	91.40	91.40	93.50	93.50	93.50	96.50
	Fazla	8.60	8.60	6.50	6.50	6.50	3.50
75	Eksik						
	Doğru	93.40	93.40	94.80	94.80	94.80	97.00
	Fazla	6.60	6.60	5.20	5.20	5.20	3.00
100	Eksik						
	Doğru	94.20	94.20	95.40	95.40	95.40	96.00
	Fazla	5.80	5.80	4.60	4.60	4.60	4.00
Gen	Eksik						
	Doğru	91.80	91.80	94.58	94.58	94.58	97.20
	Fazla	8.20	8.20	5.43	5.43	5.43	2.80

		Model 8 : $p=4, q=8, \rho_1 = \rho_2 = \rho_3 = \rho_4 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik								
	Doğru	53.00	53.60	95.30	99.00	7.10	7.40	90.40	98.70
	Fazla	47.00	46.40	4.70	1.00	92.90	92.60	9.60	1.30
50	Eksik								
	Doğru	54.80	52.80	94.70	98.70	19.50	15.00	95.60	99.40
	Fazla	45.20	47.20	5.30	1.30	80.50	85.00	4.40	0.60
75	Eksik								
	Doğru	54.30	50.50	94.60	98.60	23.10	18.90	96.60	99.70
	Fazla	45.70	49.50	5.40	1.40	76.90	81.10	3.40	0.30
100	Eksik								
	Doğru	55.80	51.40	94.60	98.60	24.20	19.00	97.30	99.90
	Fazla	44.20	48.60	5.40	1.40	75.80	81.00	2.70	0.10
Gen	Eksik								
	Doğru	54.48	52.08	94.98	98.75	18.48	15.08	94.98	99.43
	Fazla	45.53	47.93	5.03	1.25	81.53	84.93	5.03	0.58

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 9 : $p=4, q=8,\ \rho_1=.8,\ \rho_2=\ \rho_3=\rho_4=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	15.80	15.80	33.50	33.50	33.50	83.60
	Doğru	80.00	75.80	63.10	62.50	63.10	15.60
	Fazla	4.20	8.40	3.40	4.00	3.40	0.80
50	Eksik			0.40	0.40	0.40	8.10
	Doğru	95.00	92.40	95.70	94.90	95.70	89.00
	Fazla	5.00	7.60	3.90	4.70	3.90	2.90
75	Eksik						
	Doğru	95.70	94.80	96.00	95.80	96.00	97.20
	Fazla	4.30	5.20	4.00	4.20	4.00	2.80
100	Eksik						
	Doğru	94.10	93.20	94.50	94.40	94.50	95.10
	Fazla	5.90	6.80	5.50	5.60	5.50	4.90
Gen	Eksik	3.95	3.95	8.48	8.48	8.48	22.93
	Doğru	91.20	89.05	87.33	86.90	87.33	74.23
	Fazla	4.85	7.00	4.20	4.63	4.20	2.85

		Model 9 : $p=4, q=8,\ \rho_1=.8,\ \rho_2=\ \rho_3=\rho_4=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	1.70	1.70	31.90	58.60	1.40	1.40	21.80	55.90
	Doğru	20.00	19.10	45.80	34.90	2.30	1.60	44.40	34.10
	Fazla	78.30	79.20	22.30	6.50	96.30	97.00	33.80	10.00
50	Eksik			0.30	0.60	1.10	1.10	1.40	4.20
	Doğru	21.70	19.70	69.70	85.20	5.70	3.90	72.40	88.10
	Fazla	78.30	80.30	30.00	14.20	93.20	95.00	26.20	7.70
75	Eksik					0.70	0.70	0.70	0.70
	Doğru	19.90	17.50	68.00	85.20	4.00	3.10	74.40	93.10
	Fazla	80.10	82.50	32.00	14.80	95.30	96.20	24.90	6.20
100	Eksik					0.70	0.70	0.70	0.70
	Doğru	21.20	19.30	69.50	84.40	5.00	3.80	75.10	91.60
	Fazla	78.80	80.70	30.50	15.60	94.30	95.50	24.20	7.70
Gen	Eksik	0.43	0.43	8.05	14.80	0.98	0.98	6.15	15.38
	Doğru	20.70	18.90	63.25	72.43	4.25	3.10	66.58	76.73
	Fazla	78.88	80.68	28.70	12.78	94.78	95.93	27.28	7.90

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 10 : $p=4, q=8, \rho_1=.7,\ \rho_2=\ \rho_3=\rho_4=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	40.90	40.90	58.70	58.70	91.30	91.30
	Doğru	55.60	52.80	38.30	38.10	7.20	7.80
	Fazla	3.50	6.30	3.00	3.20	1.50	0.90
50	Eksik	5.00	5.00	11.10	11.10	30.90	30.90
	Doğru	90.90	88.80	86.10	85.00	67.10	67.10
	Fazla	4.10	6.20	2.80	3.90	2.00	2.00
75	Eksik	0.10	0.10	0.50	0.50	1.80	1.80
	Doğru	94.40	93.10	94.40	93.90	93.80	93.80
	Fazla	5.50	6.80	5.10	5.60	4.40	4.40
100	Eksik						
	Doğru	95.50	94.70	95.90	95.50	96.10	96.30
	Fazla	4.50	5.30	4.10	4.50	3.90	3.70
Gen	Eksik	11.50	11.50	17.58	17.58	31.00	31.00
	Doğru	84.10	82.35	78.68	78.13	66.05	66.25
	Fazla	4.40	6.15	3.75	4.30	2.95	2.75

		Model 10 : $p=4, q=8, \rho_1=.7,\ \rho_2=\ \rho_3=\rho_4=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				PA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	9.80	9.70	61.70	83.90	2.80	2.80	46.70	80.90
	Doğru	19.90	19.10	25.80	12.90	2.40	1.90	29.10	14.30
	Fazla	70.30	71.20	12.50	3.20	94.80	95.30	24.20	4.80
50	Eksik	0.20	0.10	15.00	20.20	1.30	1.30	9.40	29.40
	Doğru	24.20	22.20	73.90	71.50	5.30	3.60	69.40	65.40
	Fazla	75.60	77.70	11.10	8.30	93.40	95.10	21.20	5.20
75	Eksik				0.60	0.70	0.70	0.90	3.90
	Doğru	20.80	18.90	72.00	86.10	6.00	3.90	77.70	89.80
	Fazla	79.20	81.10	28.00	13.30	93.30	95.40	21.40	6.30
100	Eksik					0.60	0.60	0.60	0.90
	Doğru	21.40	20.10	70.90	87.00	5.50	3.50	76.30	94.40
	Fazla	78.60	79.90	29.10	13.00	93.90	95.90	23.10	4.70
Gen	Eksik	2.50	2.45	16.70	24.88	1.35	1.35	14.40	28.78
	Doğru	21.58	20.08	59.58	64.98	4.80	3.23	63.13	65.98
	Fazla	75.93	77.48	23.73	10.15	93.85	95.43	22.48	5.25

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 11 : $p=4, q=8, \rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	59.30	48.70	67.20	65.20	84.60	87.90
	Doğru	39.70	46.70	31.30	32.30	13.90	11.40
	Fazla	1.00	4.60	1.50	2.50	1.50	0.70
50	Eksik	4.80	4.30	7.10	6.90	14.40	15.80
	Doğru	91.50	90.10	89.10	88.80	82.30	81.20
	Fazla	3.70	5.60	3.80	4.30	3.30	3.00
75	Eksik	0.10	0.10	0.20	0.10	0.50	0.50
	Doğru	94.60	92.90	94.80	93.80	94.50	95.00
	Fazla	5.30	7.00	5.00	6.10	5.00	4.50
100	Eksik						
	Doğru	95.30	94.10	95.40	94.80	95.50	95.70
	Fazla	4.70	5.90	4.60	5.20	4.50	4.30
Gen	Eksik	16.05	13.28	18.63	18.05	24.88	26.05
	Doğru	80.28	80.95	77.65	77.43	71.55	70.83
	Fazla	3.68	5.78	3.73	4.53	3.58	3.13

		Model 11 : $p=4, q=8, \rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	1.50	1.60	32.60	65.50	2.40	2.40	19.30	55.40
	Doğru	12.70	11.00	34.50	22.90	6.90	5.30	54.50	37.60
	Fazla	85.80	87.40	32.90	11.60	90.70	92.30	26.20	7.00
50	Eksik			0.40	1.70	0.70	0.70	1.50	4.90
	Doğru	12.30	11.20	50.40	67.10	9.10	6.20	73.90	87.20
	Fazla	87.70	88.80	49.20	31.20	90.20	93.10	24.60	7.90
75	Eksik					1.00	1.00	1.00	1.20
	Doğru	12.60	10.20	48.00	65.50	7.40	5.00	72.20	88.20
	Fazla	87.40	89.80	52.00	34.50	91.60	94.00	26.80	10.60
100	Eksik					0.30	0.30	0.30	0.30
	Doğru	11.50	9.50	47.40	66.70	7.40	4.50	71.90	89.00
	Fazla	88.50	90.50	52.60	33.30	92.30	95.20	27.80	10.70
Gen	Eksik	0.38	0.40	8.25	16.80	1.10	1.10	5.53	15.45
	Doğru	12.28	10.48	45.08	55.55	7.70	5.25	68.13	75.50
	Fazla	87.35	89.13	46.68	27.65	91.20	93.65	26.35	9.05

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 12 : $p = 4, q = 8, \rho_1 = .7,\ \rho_2 = .5,\ \rho_3 = \rho_4 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	90.60	82.80	92.00	90.30	95.40	96.50
	Doğru	8.90	15.50	7.50	8.90	4.10	3.20
	Fazla	0.50	1.70	0.50	0.80	0.50	0.30
50	Eksik	55.80	49.00	61.50	59.50	70.60	71.70
	Doğru	43.10	47.90	37.30	39.00	28.30	27.20
	Fazla	1.10	3.10	1.20	1.50	1.10	1.10
75	Eksik	24.50	21.70	29.20	27.60	35.90	36.40
	Doğru	72.20	72.80	67.30	67.90	61.00	59.90
	Fazla	3.30	5.50	3.50	4.50	3.10	3.70
100	Eksik	7.30	6.30	9.30	8.80	13.80	13.80
	Doğru	88.70	88.50	86.80	86.50	82.40	82.20
	Fazla	4.00	5.20	3.90	4.70	3.80	4.00
Gen	Eksik	44.55	39.95	48.00	46.55	53.93	54.60
	Doğru	53.23	56.18	49.73	50.58	43.95	43.13
	Fazla	2.23	3.88	2.28	2.88	2.13	2.28

		Model 12 : $p = 4, q = 8, \rho_1 = .7,\ \rho_2 = .5,\ \rho_3 = \rho_4 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	13.00	12.40	78.00	93.00	2.90	2.80	59.50	90.20
	Doğru	15.00	13.80	13.60	5.20	15.70	12.30	30.70	8.70
	Fazla	72.00	73.80	8.40	1.80	81.40	84.90	9.80	1.10
50	Eksik	1.30	1.10	17.60	37.00	1.40	1.30	23.40	56.50
	Doğru	15.40	14.30	46.10	46.70	16.20	11.90	65.20	41.50
	Fazla	83.30	84.60	36.30	16.30	82.40	86.80	11.40	2.00
75	Eksik			3.30	9.00	0.50	0.50	5.80	24.80
	Doğru	12.80	11.40	49.60	63.70	12.40	8.10	77.50	70.70
	Fazla	87.20	88.60	47.10	27.30	87.10	91.40	16.70	4.50
100	Eksik			0.50	1.10	1.10	1.10	2.40	8.60
	Doğru	12.60	10.70	50.80	68.50	10.30	6.60	78.40	85.60
	Fazla	87.40	89.30	48.70	30.40	88.60	92.30	19.20	5.80
Gen	Eksik	3.58	3.38	24.85	35.03	1.48	1.43	22.78	45.03
	Doğru	13.95	12.55	40.03	46.03	13.65	9.73	62.95	51.63
	Fazla	82.48	84.08	35.13	18.95	84.88	88.85	14.28	3.35

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 13 : $p = 4, q = 8, \rho_1 = .8,\ \rho_2 = .7,\ \rho_3 = .5,\ \rho_4 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik		84.10	92.30	90.50	94.00	97.10
	Doğru	93.50	14.80	7.30	8.90	5.50	2.80
	Fazla	6.50	1.10	0.40	0.60	0.50	0.10
50	Eksik	48.80	40.20	49.90	46.50	54.30	58.00
	Doğru	49.60	56.30	48.10	50.20	43.60	39.70
	Fazla	1.60	3.50	2.00	3.30	2.10	2.30
75	Eksik	16.00	13.50	16.80	16.10	18.20	18.50
	Doğru	81.30	82.40	80.40	80.60	78.90	78.40
	Fazla	2.70	4.10	2.80	3.30	2.90	3.10
100	Eksik	4.40	3.10	5.10	4.40	6.10	6.20
	Doğru	92.90	92.90	92.20	92.20	90.90	90.70
	Fazla	2.70	4.00	2.70	3.40	3.00	3.10
Gen	Eksik	40.68	35.23	41.03	39.38	43.15	44.95
	Doğru	57.58	61.60	57.00	57.98	54.73	52.90
	Fazla	1.75	3.18	1.98	2.65	2.13	2.15

		Model 13 : $p = 4, q = 8, \rho_1 = .8,\ \rho_2 = .7,\ \rho_3 = .5,\ \rho_4 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	3.80	2.90	43.20	73.50	3.80	3.00	50.20	77.50
	Doğru	18.00	14.80	31.10	18.40	64.70	57.20	48.80	22.40
	Fazla	78.20	82.30	25.70	8.10	31.50	39.80	1.00	0.10
50	Eksik	0.40	0.40	3.70	10.90	1.20	1.10	15.10	36.20
	Doğru	12.80	11.30	46.10	58.70	61.40	52.20	82.00	63.20
	Fazla	86.80	88.30	50.20	30.40	37.40	46.70	2.90	0.60
75	Eksik			0.50	2.10	0.90	0.90	3.90	11.70
	Doğru	11.50	8.90	44.20	61.60	56.00	46.00	92.00	87.00
	Fazla	88.50	91.10	55.30	36.30	43.10	53.10	4.10	1.30
100	Eksik			0.10	0.30	0.60	0.60	1.10	2.60
	Doğru	13.70	12.20	45.20	61.70	54.60	45.50	94.00	95.80
	Fazla	86.30	87.80	54.70	38.00	44.80	53.90	4.90	1.60
Gen	Eksik	1.05	0.83	11.88	21.70	1.63	1.40	17.58	32.00
	Doğru	14.00	11.80	41.65	50.10	59.18	50.23	79.20	67.10
	Fazla	84.95	87.38	46.48	28.20	39.20	48.38	3.23	0.90

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 14 : $p=4, q=8, \rho_1=.7,\ \rho_2=.5,\ \rho_3=.4,\ \rho_4=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B\left(T_d^2\right)$	[2] $S_F\left(T_d^2\right)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	99.40	96.30	98.40	97.70	98.40	99.10
	Doğru	0.50	3.10	1.40	2.00	1.40	0.80
	Fazla	0.10	0.60	0.20	0.30	0.20	0.10
50	Eksik	87.90	80.70	87.90	85.00	89.10	89.10
	Doğru	11.50	18.20	11.30	13.90	10.10	10.10
	Fazla	0.60	1.10	0.80	1.10	0.80	0.80
75	Eksik	59.90	52.40	60.80	56.50	62.40	61.50
	Doğru	39.30	44.90	38.00	41.10	36.00	36.60
	Fazla	0.80	2.70	1.20	2.40	1.60	1.90
100	Eksik	37.40	31.80	38.10	34.60	39.40	38.30
	Doğru	59.40	63.70	58.60	61.40	57.30	57.80
	Fazla	3.20	4.50	3.30	4.00	3.30	3.90
Gen	Eksik	71.15	65.30	71.30	68.45	72.33	72.00
	Doğru	27.68	32.48	27.33	29.60	26.20	26.33
	Fazla	1.18	2.23	1.38	1.95	1.48	1.68

		Model 14 : $p=4, q=8, \rho_1=.7,\ \rho_2=.5,\ \rho_3=.4,\ \rho_4=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	3.80	2.90	43.20	73.50	3.80	3.00	50.20	77.50
	Doğru	18.00	14.80	31.10	18.40	64.70	57.20	48.80	22.40
	Fazla	78.20	82.30	25.70	8.10	31.50	39.80	1.00	0.10
50	Eksik	0.40	0.40	3.70	10.90	1.20	1.10	15.10	36.20
	Doğru	12.80	11.30	46.10	58.70	61.40	52.20	82.00	63.20
	Fazla	86.80	88.30	50.20	30.40	37.40	46.70	2.90	0.60
75	Eksik			0.50	2.10	0.90	0.90	3.90	11.70
	Doğru	11.50	8.90	44.20	61.60	56.00	46.00	92.00	87.00
	Fazla	88.50	91.10	55.30	36.30	43.10	53.10	4.10	1.30
100	Eksik			0.10	0.30	0.60	0.60	1.10	2.60
	Doğru	13.70	12.20	45.20	61.70	54.60	45.50	94.00	95.80
	Fazla	86.30	87.80	54.70	38.00	44.80	53.90	4.90	1.60
Gen	Eksik	1.05	0.83	11.88	21.70	1.63	1.40	17.58	32.00
	Doğru	14.00	11.80	41.65	50.10	59.18	50.23	79.20	67.10
	Fazla	84.95	87.38	46.48	28.20	39.20	48.38	3.23	0.90

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 15 : $p=6, q=7, \rho_1=\rho_2=\rho_3=\rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik						
	Doğru	85.90	85.90	93.90	93.90	99.30	99.30
	Fazla	14.10	14.10	6.10	6.10	0.70	0.70
50	Eksik						
	Doğru	92.20	92.20	95.40	95.40	98.30	98.30
	Fazla	7.80	7.80	4.60	4.60	1.70	1.70
75	Eksik						
	Doğru	93.40	93.40	94.90	94.90	97.00	97.00
	Fazla	6.60	6.60	5.10	5.10	3.00	3.00
100	Eksik						
	Doğru	93.00	93.00	94.10	94.10	95.30	95.30
	Fazla	7.00	7.00	5.90	5.90	4.70	4.70
Gen	Eksik						
	Doğru	91.13	91.13	94.58	94.58	97.48	97.48
	Fazla	8.88	8.88	5.43	5.43	2.53	2.53

		Model 15 : $p=6, q=7, \rho_1=\rho_2=\rho_3=\rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik								
	Doğru	46.40	47.60	94.00	99.00	3.20	3.30	79.40	96.80
	Fazla	53.60	52.40	6.00	1.00	96.80	96.70	20.60	3.20
50	Eksik								
	Doğru	50.10	48.00	95.40	99.00	7.50	6.40	92.70	99.40
	Fazla	49.90	52.00	4.60	1.00	92.50	93.60	7.30	0.60
75	Eksik								
	Doğru	54.90	52.50	94.70	98.80	11.10	8.60	94.40	99.50
	Fazla	45.10	47.50	5.30	1.20	88.90	91.40	5.60	0.50
100	Eksik								
	Doğru	53.90	50.20	94.20	98.20	9.10	6.30	93.50	99.00
	Fazla	46.10	49.80	5.80	1.80	90.90	93.70	6.50	1.00
Gen	Eksik								
	Doğru	51.33	49.58	94.58	98.75	7.73	6.15	90.00	98.68
	Fazla	48.68	50.43	5.43	1.25	92.28	93.85	10.00	1.33

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 16 : $p = 6,\ q = 7, \rho_1 = .8,\ \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	24.50	24.50	44.50	44.50	88.30	88.30
	Doğru	71.60	67.60	51.90	51.80	10.50	11.20
	Fazla	3.90	7.90	3.60	3.70	1.20	0.50
50	Eksik	0.20	0.20	1.70	1.70	14.90	14.90
	Doğru	93.20	91.40	93.30	92.90	82.10	82.60
	Fazla	6.60	8.40	5.00	5.40	3.00	2.50
75	Eksik						
	Doğru	93.00	91.10	93.30	93.30	94.20	94.30
	Fazla	7.00	8.90	6.70	6.70	5.80	5.70
100	Eksik						
	Doğru	94.30	93.40	95.00	94.50	95.80	96.20
	Fazla	5.70	6.60	5.00	5.50	4.20	3.80
Gen	Eksik	6.18	6.18	11.55	11.55	25.80	25.80
	Doğru	88.03	85.88	83.38	83.13	70.65	71.08
	Fazla	5.80	7.95	5.08	5.33	3.55	3.13

		Model 16 : $p = 6,\ q = 7, \rho_1 = .8,\ \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	3.30	3.40	45.60	71.40	1.90	1.90	19.60	58.00
	Doğru	18.80	18.40	37.70	24.70	0.10	0.10	25.70	28.10
	Fazla	77.90	78.20	16.70	3.90	98.00	98.00	54.70	13.90
50	Eksik				0.90	1.20	1.20	1.30	4.30
	Doğru	19.80	17.10	66.50	80.20	0.30	0.10	49.50	77.30
	Fazla	80.20	82.90	33.50	18.90	98.50	98.70	49.20	18.40
75	Eksik					0.80	0.80	0.80	0.80
	Doğru	18.30	16.20	64.10	80.60	0.50	0.10	50.20	80.80
	Fazla	81.70	83.80	35.90	19.40	98.70	99.10	49.00	18.40
100	Eksik					1.20	1.20	1.20	1.20
	Doğru	18.40	16.40	66.80	83.50	0.40	0.30	53.00	82.30
	Fazla	81.60	83.60	33.20	16.50	98.40	98.50	45.80	16.50
Gen	Eksik	0.83	0.85	11.40	18.08	1.28	1.28	5.73	16.08
	Doğru	18.83	17.03	58.78	67.25	0.33	0.15	44.60	67.13
	Fazla	80.35	82.13	29.83	14.68	98.40	98.58	49.68	16.80

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 17 : $p = 6, q = 7,\ \rho_1 = .7,\ \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	49.90	49.90	67.90	67.90	93.50	93.50
	Doğru	47.60	45.00	30.00	29.80	5.50	6.00
	Fazla	2.50	5.10	2.10	2.30	1.00	0.50
50	Eksik	8.90	8.90	18.90	18.90	43.80	43.80
	Doğru	86.60	85.10	77.60	77.20	54.20	54.40
	Fazla	4.50	6.00	3.50	3.90	2.00	1.80
75	Eksik	0.20	0.20	1.50	1.50	5.80	5.80
	Doğru	95.30	94.30	94.20	93.90	90.70	90.90
	Fazla	4.50	5.50	4.30	4.60	3.50	3.30
100	Eksik					0.30	0.30
	Doğru	94.70	93.40	95.40	95.00	95.80	96.00
	Fazla	5.30	6.60	4.60	5.00	3.90	3.70
Gen	Eksik	14.75	14.75	22.08	22.08	35.85	35.85
	Doğru	81.05	79.45	74.30	73.98	61.55	61.83
	Fazla	4.20	5.80	3.63	3.95	2.60	2.33

		Model 17 : $p = 6, q = 7,\ \rho_1 = .7,\ \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	10.60	11.30	71.20	90.90	1.80	1.80	41.20	80.40
	Doğru	20.30	19.30	21.00	7.70	0.30	0.30	20.20	12.60
	Fazla	69.10	69.40	7.80	1.40	97.90	97.90	38.60	7.00
50	Eksik	0.30	0.30	8.70	23.60	1.20	1.20	8.20	32.40
	Doğru	20.30	19.20	63.60	65.80	0.90	0.60	49.70	56.90
	Fazla	79.40	80.50	27.70	10.60	97.90	98.20	42.10	10.70
75	Eksik			0.40	1.20	0.70	0.70	1.00	4.30
	Doğru	18.30	16.70	67.40	84.30	0.90	0.50	56.80	83.00
	Fazla	81.70	83.30	32.20	14.50	98.40	98.80	42.20	12.70
100	Eksik					0.60	0.60	0.60	0.90
	Doğru	18.90	16.70	67.60	82.90	0.60	0.40	56.60	85.70
	Fazla	81.10	83.30	32.40	17.10	98.80	99.00	42.80	13.40
Gen	Eksik	2.73	2.90	20.08	28.93	1.08	1.08	12.75	29.50
	Doğru	19.45	17.98	54.90	60.18	0.68	0.45	45.83	59.55
	Fazla	77.83	79.13	25.03	10.90	98.25	98.48	41.43	10.95

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 18 : $p=6,\ q=7, \rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	67.80	54.10	73.70	72.30	87.70	91.60
	Doğru	31.40	42.00	24.90	25.60	10.40	7.20
	Fazla	0.80	3.90	1.40	2.10	1.90	1.20
50	Eksik	9.00	6.30	13.40	11.90	28.00	31.40
	Doğru	87.30	87.50	83.10	83.50	68.80	65.90
	Fazla	3.70	6.20	3.50	4.60	3.20	2.70
75	Eksik	0.60	0.50	1.00	1.00	2.10	2.70
	Doğru	95.90	93.80	95.70	94.50	94.50	94.30
	Fazla	3.50	5.70	3.30	4.50	3.40	3.00
100	Eksik						
	Doğru	95.70	94.60	95.90	95.30	96.00	96.20
	Fazla	4.30	5.40	4.10	4.70	4.00	3.80
Gen	Eksik	19.35	15.23	22.03	21.30	29.45	31.43
	Doğru	77.58	79.48	74.90	74.73	67.43	65.90
	Fazla	3.08	5.30	3.08	3.98	3.13	2.68

		Model 18 : $p=6,\ q=7, \rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	3.10	2.90	43.00	72.90	1.70	1.70	9.10	44.20
	Doğru	7.60	7.10	26.50	17.40	0.30		31.80	32.70
	Fazla	89.30	90.00	30.50	9.70	98.00	98.30	59.10	23.10
50	Eksik			0.50	2.30	0.90	0.90	1.40	3.70
	Doğru	9.00	8.00	44.40	64.30	0.30	0.10	41.50	72.30
	Fazla	91.00	92.00	55.10	33.40	98.80	99.00	57.10	24.00
75	Eksik					0.40	0.40	0.40	0.60
	Doğru	9.40	8.20	45.10	62.50	0.40	0.10	41.00	72.30
	Fazla	90.60	91.80	54.90	37.50	99.20	99.50	58.60	27.10
100	Eksik					0.50	0.50	0.50	0.50
	Doğru	8.90	7.50	43.50	62.50	0.30	0.20	41.70	73.60
	Fazla	91.10	92.50	56.50	37.50	99.20	99.30	57.80	25.90
Gen	Eksik	0.78	0.73	10.88	18.80	0.88	0.88	2.85	12.25
	Doğru	8.73	7.70	39.88	51.68	0.33	0.10	39.00	62.73
	Fazla	90.50	91.58	49.25	29.53	98.80	99.03	58.15	25.03

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 19 : $p=6, q=7, \rho_1=.7,\ \rho_2=.5,\ \rho_3=\rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	93.20	87.80	93.20	92.70	96.30	97.90
	Doğru	6.50	10.80	6.30	6.50	3.20	2.00
	Fazla	0.30	1.40	0.50	0.80	0.50	0.10
50	Eksik	62.70	55.10	69.30	66.20	77.90	79.80
	Doğru	36.50	42.60	29.70	32.00	21.10	19.40
	Fazla	0.80	2.30	1.00	1.80	1.00	0.80
75	Eksik	32.40	29.10	38.40	36.60	47.10	47.80
	Doğru	65.40	66.80	59.20	60.20	50.30	49.60
	Fazla	2.20	4.10	2.40	3.20	2.60	2.60
100	Eksik	15.40	14.20	19.00	17.50	25.40	25.50
	Doğru	80.60	80.20	76.50	77.40	70.20	69.60
	Fazla	4.00	5.60	4.50	5.10	4.40	4.90
Gen	Eksik	50.93	46.55	54.98	53.25	61.68	62.75
	Doğru	47.25	50.10	42.93	44.03	36.20	35.15
	Fazla	1.83	3.35	2.10	2.73	2.13	2.10

		Model 19 : $p=6, q=7, \rho_1=.7,\ \rho_2=.5,\ \rho_3=\rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	15.90	16.20	85.70	96.80	1.40	1.40	43.10	83.70
	Doğru	12.50	11.30	6.90	2.00	1.20	0.90	28.00	11.00
	Fazla	71.60	72.50	7.40	1.20	97.40	97.70	28.90	5.30
50	Eksik	2.30	2.00	25.20	48.00	1.20	1.20	16.20	46.00
	Doğru	11.10	10.60	35.50	34.20	1.10	0.60	47.90	45.90
	Fazla	86.60	87.40	39.30	17.80	97.70	98.20	35.90	8.10
75	Eksik	0.50	0.50	5.30	12.50	0.30	0.30	3.30	17.20
	Doğru	10.50	8.80	44.60	58.20	0.50	0.30	56.60	70.20
	Fazla	89.00	90.70	50.10	29.30	99.20	99.40	40.10	12.60
100	Eksik	0.20	0.20	1.60	4.10	0.40	0.40	1.90	6.90
	Doğru	9.70	8.20	44.50	62.90	0.50	0.10	49.70	75.40
	Fazla	90.10	91.60	53.90	33.00	99.10	99.50	48.40	17.70
Gen	Eksik	4.73	4.73	29.45	40.35	0.83	0.83	16.13	38.45
	Doğru	10.95	9.73	32.88	39.33	0.83	0.48	45.55	50.63
	Fazla	84.33	85.55	37.68	20.33	98.35	98.70	38.33	10.93

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 20 : $p=6, q=7, \rho_1=.8.\ \rho_2=.7,\ \rho_3=.5,\ \rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	96.90	87.90	94.80	93.50	95.10	98.30
	Doğru	3.00	11.30	5.00	6.20	4.60	1.70
	Fazla	0.10	0.80	0.20	0.30	0.30	
50	Eksik	58.60	50.00	61.70	57.90	66.00	69.60
	Doğru	39.40	45.80	35.50	38.50	30.80	27.30
	Fazla	2.00	4.20	2.80	3.60	3.20	3.10
75	Eksik	25.90	20.80	28.70	26.50	32.40	33.40
	Doğru	71.80	75.60	68.60	70.10	64.80	63.80
	Fazla	2.30	3.60	2.70	3.40	2.80	2.80
100	Eksik	9.10	7.40	10.00	9.10	11.80	12.40
	Doğru	87.90	87.50	86.70	86.50	84.60	83.80
	Fazla	3.00	5.10	3.30	4.40	3.60	3.80
Gen	Eksik	47.63	41.53	48.80	46.75	51.33	53.43
	Doğru	50.53	55.05	48.95	50.33	46.20	44.15
	Fazla	1.85	3.43	2.25	2.93	2.48	2.43

		Model 20 : $p=6, q=7, \rho_1=.8.\ \rho_2=.7,\ \rho_3=.5,\ \rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	4.40	4.50	54.00	82.60	2.00	2.00	27.10	63.10
	Doğru	11.50	9.80	18.30	10.70	4.00	2.60	50.00	31.80
	Fazla	84.10	85.70	27.70	6.70	94.00	95.40	22.90	5.10
50	Eksik	0.30	0.30	6.10	14.20	1.30	1.30	7.00	24.70
	Doğru	7.20	5.50	35.30	48.30	2.80	1.60	65.30	64.30
	Fazla	92.50	94.20	58.60	37.50	95.90	97.10	27.70	11.00
75	Eksik			0.50	2.40	0.30	0.30	1.00	5.00
	Doğru	5.60	5.20	32.80	51.60	1.70	0.60	60.80	81.30
	Fazla	94.40	94.80	66.70	46.00	98.00	99.10	38.20	13.70
100	Eksik			0.10	0.40	0.50	0.50	0.50	1.10
	Doğru	5.80	4.70	33.40	50.40	0.80	0.50	56.90	82.20
	Fazla	94.20	95.30	66.50	49.20	98.70	99.00	42.60	16.70
Gen	Eksik	1.18	1.20	15.18	24.90	1.03	1.03	8.90	23.48
	Doğru	7.53	6.30	29.95	40.25	2.33	1.33	58.25	64.90
	Fazla	91.30	92.50	54.88	34.85	96.65	97.65	32.85	11.63

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 21 : $p=6, q=7, \rho_1=.7,\ \rho_2=.5,\ \rho_3=.4,\ \rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	99.80	98.20	99.40	99.00	99.10	99.70
	Doğru	0.20	1.60	0.50	0.80	0.70	0.20
	Fazla		0.20	0.10	0.20	0.20	0.10
50	Eksik	91.10	85.00	90.90	88.30	90.80	92.20
	Doğru	8.70	13.70	8.70	11.10	8.80	7.40
	Fazla	0.20	1.30	0.40	0.60	0.40	0.40
75	Eksik	71.60	62.60	73.00	67.90	74.50	73.80
	Doğru	27.20	34.10	25.40	29.30	23.60	23.70
	Fazla	1.20	3.30	1.60	2.80	1.90	2.50
100	Eksik	49.90	40.90	52.10	46.10	55.10	53.10
	Doğru	48.60	56.30	46.30	51.80	43.30	45.20
	Fazla	1.50	2.80	1.60	2.10	1.60	1.70
Gen	Eksik	78.10	75.33	78.85	75.33	79.88	79.70
	Doğru	21.18	23.25	20.23	23.25	19.10	19.13
	Fazla	0.73	1.43	0.93	1.43	1.03	1.18

		Model 21 : $p=6, q=7, \rho_1=.7,\ \rho_2=.5,\ \rho_3=.4,\ \rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	18.70	18.30	89.10	98.20	2.40	2.20	59.00	91.40
	Doğru	15.50	13.60	5.20	1.30	15.10	11.30	33.80	7.50
	Fazla	65.80	68.10	5.70	0.50	82.50	86.50	7.20	1.10
50	Eksik	2.30	2.10	31.10	56.40	1.40	1.40	36.90	69.20
	Doğru	11.10	9.50	30.30	26.60	10.20	7.90	53.20	29.50
	Fazla	86.60	88.40	38.60	17.00	88.40	90.70	9.90	1.30
75	Eksik	0.80	0.80	10.20	22.50	0.80	0.80	15.10	41.40
	Doğru	8.60	7.70	33.90	40.10	6.80	4.10	66.10	53.80
	Fazla	90.60	91.50	55.90	37.40	92.40	95.10	18.80	4.80
100	Eksik	0.10	0.10	2.70	7.80	0.30	0.30	4.30	19.50
	Doğru	6.90	5.90	35.00	49.60	3.50	2.00	70.40	74.80
	Fazla	93.00	94.00	62.30	42.60	96.20	97.70	25.30	5.70
Gen	Eksik	5.48	5.33	33.28	46.23	1.23	1.18	28.83	55.38
	Doğru	10.53	9.18	26.10	29.40	8.90	6.33	55.88	41.40
	Fazla	84.00	85.50	40.63	24.38	89.88	92.50	15.30	3.23

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 22 : $p = 6, q = 8, \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = 0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik						
	Doğru	87.40	87.40	93.50	93.50	99.20	99.20
	Fazla	12.60	12.60	6.50	6.50	0.80	0.80
50	Eksik						
	Doğru	92.40	92.40	95.30	95.30	98.30	98.30
	Fazla	7.60	7.60	4.70	4.70	1.70	1.70
75	Eksik						
	Doğru	93.30	93.30	94.90	94.90	97.40	97.40
	Fazla	6.70	6.70	5.10	5.10	2.60	2.60
100	Eksik						
	Doğru	93.20	93.20	94.80	94.80	96.00	96.00
	Fazla	6.80	6.80	5.20	5.20	4.00	4.00
Gen	Eksik						
	Doğru	91.58	91.58	94.63	94.63	97.73	97.73
	Fazla	8.43	8.43	5.38	5.38	2.28	2.28

		Model 22 : $p = 6, q = 8, \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = 0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik								
	Doğru	47.60	49.20	95.70	98.70	2.00	2.10	66.70	94.70
	Fazla	52.40	50.80	4.30	1.30	98.00	97.90	33.30	5.30
50	Eksik								
	Doğru	52.10	50.00	96.20	99.00	4.00	3.10	89.30	99.30
	Fazla	47.90	50.00	3.80	1.00	96.00	96.90	10.70	0.70
75	Eksik								
	Doğru	51.80	48.70	94.70	99.30	3.80	3.10	90.60	99.10
	Fazla	48.20	51.30	5.30	0.70	96.20	96.90	9.40	0.90
100	Eksik								
	Doğru	56.90	53.90	94.50	98.60	5.60	3.60	91.80	99.40
	Fazla	43.10	46.10	5.50	1.40	94.40	96.40	8.20	0.60
Gen	Eksik								
	Doğru	52.10	50.45	95.28	98.90	3.85	2.98	84.60	98.13
	Fazla	47.90	49.55	4.73	1.10	96.15	97.03	15.40	1.88

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 23 : $p=6, q=8, \rho_1=.8,\ \rho_2=\rho_3=\rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	27.30	27.30	47.70	47.70	91.30	91.30
	Doğru	68.70	64.50	48.70	48.30	7.50	8.10
	Fazla	4.00	8.20	3.60	4.00	1.20	0.60
50	Eksik	1.00	1.00	3.00	3.00	23.50	23.50
	Doğru	94.40	92.20	93.20	92.90	73.70	74.00
	Fazla	4.60	6.80	3.80	4.10	2.80	2.50
75	Eksik					0.20	0.20
	Doğru	93.30	92.20	93.80	93.70	94.80	95.10
	Fazla	6.70	7.80	6.20	6.30	5.00	4.70
100	Eksik						
	Doğru	95.00	94.30	95.20	95.10	95.90	96.10
	Fazla	5.00	5.70	4.80	4.90	4.10	3.90
Gen	Eksik	7.08	7.08	12.68	12.68	28.75	28.75
	Doğru	87.85	85.80	82.73	82.50	67.98	68.33
	Fazla	5.08	7.13	4.60	4.83	3.28	2.93

		Model 23 : $p=6, q=8, \rho_1=.8,\ \rho_2=\rho_3=\rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	4.80	5.20	53.90	78.40	2.00	2.00	13.30	49.30
	Doğru	17.90	17.50	31.90	18.60			15.10	24.20
	Fazla	77.30	77.30	14.20	3.00	98.00	98.00	71.60	26.50
50	Eksik			0.60	2.80	0.60	0.60	0.90	4.30
	Doğru	19.50	18.30	69.00	84.10	0.20	0.20	40.50	73.30
	Fazla	80.50	81.70	30.40	13.10	99.20	99.20	58.60	22.40
75	Eksik					0.70	0.70	0.70	0.70
	Doğru	17.70	16.10	67.00	83.60			39.20	76.00
	Fazla	82.30	83.90	33.00	16.40	99.30	99.30	60.10	23.30
100	Eksik					0.40	0.40	0.40	0.40
	Doğru	17.70	15.90	65.30	81.80	0.10		39.70	76.20
	Fazla	82.30	84.10	34.70	18.20	99.50	99.60	59.90	23.40
Gen	Eksik	1.20	1.30	13.63	20.30	0.93	0.93	3.83	13.68
	Doğru	18.20	16.95	58.30	67.03	0.08	0.05	33.63	62.43
	Fazla	80.60	81.75	28.08	12.68	99.00	99.03	62.55	23.90

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 24 : $p=6, q=8, \rho_1=.7,\ \rho_2=\rho_3=\rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	53.30	53.30	68.40	68.40	96.20	96.20
	Doğru	45.20	42.70	29.90	29.80	3.10	3.40
	Fazla	1.50	4.00	1.70	1.80	0.70	0.40
50	Eksik	10.40	10.40	22.90	22.90	50.00	50.00
	Doğru	85.70	84.10	74.00	73.10	47.60	48.10
	Fazla	3.90	5.50	3.10	4.00	2.40	1.90
75	Eksik	0.90	0.90	2.70	2.70	9.40	9.40
	Doğru	94.10	93.10	93.30	92.60	87.30	87.70
	Fazla	5.00	6.00	4.00	4.70	3.30	2.90
100	Eksik	0.10	0.10	0.10	0.10	1.00	1.00
	Doğru	94.80	94.20	95.30	95.20	95.60	95.70
	Fazla	5.10	5.70	4.60	4.70	3.40	3.30
Gen	Eksik	16.18	16.18	23.53	23.53	39.15	39.15
	Doğru	79.95	78.53	73.13	72.68	58.40	58.73
	Fazla	3.88	5.30	3.35	3.80	2.45	2.13

		Model 24 : $p=6, q=8, \rho_1=.7,\ \rho_2=\rho_3=\rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	16.30	17.40	74.40	91.10	1.60	1.60	33.00	72.80
	Doğru	18.70	18.00	17.10	7.80			13.70	13.60
	Fazla	65.00	64.60	8.50	1.10	98.40	98.40	53.30	13.60
50	Eksik	0.40	0.40	10.90	24.80	1.40	1.40	8.30	28.20
	Doğru	21.60	20.40	62.20	64.50	0.10		40.00	54.30
	Fazla	78.00	79.20	26.90	10.70	98.50	98.60	51.70	17.50
75	Eksik			0.60	1.90	0.70	0.70	1.00	3.80
	Doğru	20.10	18.10	67.80	83.00	0.30	0.20	44.30	76.40
	Fazla	79.90	81.90	31.60	15.10	99.00	99.10	54.70	19.80
100	Eksik			0.10	0.20	0.20	0.20	0.30	0.60
	Doğru	18.00	16.50	68.10	85.60	0.10		43.00	79.60
	Fazla	82.00	83.50	31.80	14.20	99.70	99.80	56.70	19.80
Gen	Eksik	1.20	1.30	13.63	20.30	0.93	0.93	3.83	13.68
	Doğru	18.20	16.95	58.30	67.03	0.08	0.05	33.63	62.43
	Fazla	80.60	81.75	28.08	12.68	99.00	99.03	62.55	23.90

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 25 : $p=6, q=8, \rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B\left(T_d^2\right)$	[2] $S_F\left(T_d^2\right)$	[3] $S_B\left(\Lambda_d\right)$	[4] $S_L\left(\Lambda_d\right)$	[5] $S_B\left(V_d\right)$	[6] $S_F\left(V_d\right)$
25	Eksik	76.70	66.00	80.00	79.10	92.20	95.30
	Doğru	22.70	29.90	18.20	18.50	5.90	4.50
	Fazla	0.60	4.10	1.80	2.40	1.90	0.20
50	Eksik	10.90	8.90	17.30	16.70	35.50	39.00
	Doğru	85.50	84.50	79.10	78.30	61.00	58.00
	Fazla	3.60	6.60	3.60	5.00	3.50	3.00
75	Eksik	0.40	0.20	1.10	0.90	3.80	3.90
	Doğru	96.00	94.20	95.50	94.40	92.60	93.00
	Fazla	3.60	5.60	3.40	4.70	3.60	3.10
100	Eksik			0.10	0.10	0.30	0.30
	Doğru	95.90	94.70	95.90	95.40	95.60	96.00
	Fazla	4.10	5.30	4.00	4.50	4.10	3.70
Gen	Eksik	22.00	18.78	24.63	24.20	32.95	34.63
	Doğru	75.03	75.83	72.18	71.65	63.78	62.88
	Fazla	2.98	5.40	3.20	4.15	3.28	2.50

		Model 25 : $p=6, q=8, \rho_1=.8,\ \rho_2=.7,\ \rho_3=\rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	3.60	3.60	53.00	83.80	2.30	2.30	7.60	39.70
	Doğru	10.70	10.40	24.00	11.00			22.50	30.80
	Fazla	85.70	86.00	23.00	5.20	97.70	97.70	69.90	29.50
50	Eksik	0.10	0.10	1.30	3.10	0.90	0.90	1.10	3.60
	Doğru	9.60	8.70	42.70	60.80	0.10	0.10	26.30	60.90
	Fazla	90.30	91.20	56.00	36.10	99.00	99.00	72.60	35.50
75	Eksik					0.40	0.40	0.40	0.40
	Doğru	7.30	6.70	43.20	62.20			26.00	62.80
	Fazla	92.70	93.30	56.80	37.80	99.60	99.60	73.60	36.80
100	Eksik					0.50	0.50	0.50	0.50
	Doğru	9.00	8.30	44.20	62.10			26.70	62.10
	Fazla	91.00	91.70	55.80	37.90	99.50	99.50	72.80	37.40
Gen	Eksik	0.93	0.93	13.58	21.73	1.03	1.03	2.40	11.05
	Doğru	9.15	8.53	38.53	49.03	0.03	0.03	25.38	54.15
	Fazla	89.93	90.55	47.90	29.25	98.95	98.95	72.23	34.80

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 26 : $p=6, q=8, \rho_1=.7,\ \rho_2=.5,\ \rho_3=\rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenend(%)	[1] $S_B(T_d^2)$	[2] $S_F(T_d^2)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	94.00	87.60	93.20	92.60	97.20	98.40
	Doğru	5.70	11.70	6.40	6.90	2.20	1.50
	Fazla	0.30	0.70	0.40	0.50	0.60	0.10
50	Eksik	66.90	60.10	72.40	70.10	79.50	81.40
	Doğru	32.10	37.70	26.50	28.50	19.60	17.80
	Fazla	1.00	2.20	1.10	1.40	0.90	0.80
75	Eksik	35.10	31.50	42.40	40.30	52.20	52.90
	Doğru	62.40	64.00	55.10	56.10	45.60	44.50
	Fazla	2.50	4.50	2.50	3.60	2.20	2.60
100	Eksik	17.30	16.00	21.90	20.50	27.30	27.90
	Doğru	79.80	78.80	75.20	75.30	70.10	69.00
	Fazla	2.90	5.20	2.90	4.20	2.60	3.10
Gen	Eksik	53.33	48.80	57.48	55.88	64.05	65.15
	Doğru	45.00	48.05	40.80	41.70	34.38	33.20
	Fazla	1.68	3.15	1.73	2.43	1.58	1.65

		Model 26 : $p=6, q=8, \rho_1=.7,\ \rho_2=.5,\ \rho_3=\rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenend(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	18.40	19.50	86.40	97.40	2.00	2.00	24.80	72.40
	Doğru	11.40	11.00	6.20	1.50	0.60	0.40	27.20	14.10
	Fazla	70.20	69.50	7.40	1.10	97.40	97.60	48.00	13.50
50	Eksik	2.50	2.30	26.30	54.00	0.50	0.50	11.50	41.50
	Doğru	10.90	9.80	35.70	29.60	0.30	0.30	42.60	43.40
	Fazla	86.60	87.90	38.00	16.40	99.20	99.20	45.90	15.10
75	Eksik	0.20	0.20	5.90	15.70	1.10	1.10	2.50	14.90
	Doğru	8.90	7.40	43.90	56.00			41.60	61.50
	Fazla	90.90	92.40	50.20	28.30	98.90	98.90	55.90	23.60
100	Eksik			1.60	4.30	0.50	0.50	0.70	5.30
	Doğru	10.00	8.70	46.40	62.80			37.10	67.30
	Fazla	90.00	91.30	52.00	32.90	99.50	99.50	62.20	27.40
Gen	Eksik	5.28	5.50	30.05	42.85	1.03	1.03	9.88	33.53
	Doğru	10.30	9.23	33.05	37.48	0.23	0.13	37.13	46.58
	Fazla	84.43	85.28	36.90	19.68	98.75	98.85	53.00	19.90

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 27 : $p=6, q=8, \rho_1=.8,\ \rho_2=.7,\ \rho_3=.5,\ \rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B\left(T_d^2\right)$	[2] $S_F\left(T_d^2\right)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	97.60	89.70	95.80	94.70	95.60	98.60
	Doğru	2.40	9.50	4.00	5.00	3.90	1.20
	Fazla		0.80	0.20	0.30	0.50	0.20
50	Eksik	65.30	53.90	68.00	64.20	72.40	75.30
	Doğru	33.70	42.90	30.60	33.50	25.90	23.10
	Fazla	1.00	3.20	1.40	2.30	1.70	1.60
75	Eksik	31.90	24.30	34.80	31.20	39.40	41.90
	Doğru	65.40	70.90	62.30	64.90	57.70	55.00
	Fazla	2.70	4.80	2.90	3.90	2.90	3.10
100	Eksik	13.60	10.90	15.50	13.60	18.90	19.30
	Doğru	84.40	85.10	82.30	82.90	78.60	78.10
	Fazla	2.00	4.00	2.20	3.50	2.50	2.60
Gen	Eksik	52.10	44.70	53.53	50.93	56.58	58.78
	Doğru	46.48	52.10	44.80	46.58	41.53	39.35
	Fazla	1.43	3.20	1.68	2.50	1.90	1.88

		Model 27 : $p=6, q=8, \rho_1=.8,\ \rho_2=.7,\ \rho_3=.5,\ \rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	5.90	6.0	61.80	90.70	2.10	2.10	14.90	52.90
	Doğru	8.80	8.20	13.70	4.50	1.40	1.00	43.30	34.70
	Fazla	85.30	85.80	24.50	4.80	96.50	96.90	41.80	12.40
50	Eksik	0.40	0.40	65.30	3.80	1.10	1.10	4.20	17.20
	Doğru	6.80	5.30	33.70	27.20	0.60	0.20	51.70	68.50
	Fazla	92.80	94.30	1.00	69.00	98.30	98.70	44.10	14.30
75	Eksik			0.70	2.20	0.40	0.40	0.80	3.30
	Doğru	6.80	5.90	30.00	48.10	0.20	0.10	44.40	72.50
	Fazla	93.20	94.10	69.30	49.70	99.40	99.50	54.80	24.20
100	Eksik			0.10	0.70	0.70	0.70	0.70	1.30
	Doğru	4.90	4.00	30.70	47.80	0.10	0.10	41.10	71.40
	Fazla	95.10	96.00	69.20	51.50	99.20	99.20	58.20	27.30
Gen	Eksik	1.58	1.60	17.08	27.30	1.08	1.08	5.15	18.68
	Doğru	6.83	5.85	27.43	36.35	0.58	0.35	45.13	61.78
	Fazla	91.60	955	55.50	36.35	98.35	98.58	49.73	19.55

Tablo A.1 28 Kanonik korelasyon modeli için yöntemlerin boyut sayısı belirlenmesini eksik, doğru ve fazla tespit yüzdelikleri **(**Devam)

		Model 28 : $p=6, q=8, \rho_1=.7, \rho_2=.5,\ \rho_3=.4,\ \rho_4=\rho_5=\rho_6=0$					
		[a] χ^2 – Yaklaşım prosedürleri					
		T_d^2 Dayalı		Λ_d Dayalı		V_d Dayalı	
N	Belirlenen d(%)	[1] $S_B\left(T_d^2\right)$	[2] $S_F\left(T_d^2\right)$	[3] $S_B(\Lambda_d)$	[4] $S_L(\Lambda_d)$	[5] $S_B(V_d)$	[6] $S_F(V_d)$
25	Eksik	99.90	97.7	99.60	98.90	99.10	99.80
	Doğru	0.10	2.20	0.40	1.00	0.80	0.20
	Fazla		0.10		0.10	0.10	
50	Eksik	94.50	86.8	93.10	90.50	92.90	93.50
	Doğru	5.40	11.8	6.40	8.50	6.50	5.80
	Fazla	0.10	1.40	0.50	1.00	0.60	0.70
75	Eksik	76.20	66.6	76.60	71.30	77.90	77.10
	Doğru	23.40	31.4	22.90	27.20	21.30	21.50
	Fazla	0.40	2.00	0.50	1.50	0.80	1.40
100	Eksik	55.30	47.1	57.60	51.90	60.60	59.30
	Doğru	43.30	49.1	40.90	44.80	37.60	37.80
	Fazla	1.40	3.80	1.50	3.30	1.80	2.90
Gen	Eksik	81.48	74.5	81.73	78.15	82.63	82.43
	Doğru	18.05	23.6	17.65	20.38	16.55	16.33
	Fazla	0.48	1.83	0.63	1.48	0.83	1.25

		Model 28 : $p=6, q=8, \rho_1=.7, \rho_2=.5,\ \rho_3=.4,\ \rho_4=\rho_5=\rho_6=0$							
		Önerilen PA'ya dayalı yöntemler							
		PA				BPA			
N	Belirlenen d(%)	[11] PA_{mean}	[12] PA_{median}	[7] PA_{95}	[8] PA_{99}	[13] BPA_{mean}	[14] BPA_{median}	[9] BPA_{95}	[10] BPA_{99}
25	Eksik	22.00	21.8	88.70	98.20	2.40	2.40	40.30	81.60
	Doğru	11.80	10.8	5.00	1.20	4.70	2.60	39.90	15.50
	Fazla	66.20	67.4	6.30	0.60	92.90	95.00	19.80	2.90
50	Eksik	2.90	2.70	36.50	63.80	1.00	1.00	26.00	61.80
	Doğru	9.70	8.30	27.10	22.00	3.70	2.00	56.30	35.50
	Fazla	87.40	89.0	36.40	14.20	95.30	97.00	17.70	2.70
75	Eksik	0.90	0.90	10.50	24.30			8.70	32.50
	Doğru	7.80	6.10	31.90	41.30	2.30	1.30	57.20	58.80
	Fazla	91.30	93.0	57.60	34.40	97.70	98.70	34.10	8.70
100	Eksik	0.20	0.10	3.90	11.20	0.60	0.60	3.00	16.30
	Doğru	5.50	4.50	32.30	45.30	0.80	0.50	59.20	70.30
	Fazla	94.30	95.4	63.80	43.50	98.60	98.90	37.80	13.40
Gen	Eksik	6.50	6.38	34.90	49.38	1.00	1.00	19.50	48.05
	Doğru	8.70	7.43	24.08	27.45	2.88	1.60	53.15	45.03
	Fazla	84.80	86.2	41.03	23.18	96.13	97.40	27.35	6.93

Not: [a, 1,2,3,4,5,6,7,8,9,10] açıklamaları Çizelge 2.4'te verilmiştir *N*: Örneklem sayısı, *d*: Boyut, , [11] PA_{mean} = $\mathrm{PA}_{\text{ortalama}}$: PA'ya dayalı rasgele veri özdeğerlerinin ortalama eşik değeri, [12] PA_{median} = $\mathrm{PA}_{\text{medyan}}$: PA'ya dayalı rasgele veri özdeğerlerinin medyan eşik değeri, [13] BPA_{mean} = $BPA_{ortalama}$:BPA'ya dayalı rasgele veri özdeğerlerinin ortalama eşik değeri, [14] BPA_{median} = BPA_{medyan} : BPA'ya dayalı rasgele veri özdeğerlerinin medyan eşik değeri.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı :Selahattin AYDOĞDU

Doğum Tarihi ve Yeri :Arıcak/1988

Yabancı Dili :İngilizce

E-posta :saydogd@gmail.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Y. Lisans	İstatistik	Fırat Üniversitesi	2011
Lisans	İstatistik	Fırat Üniversitesi	2009
Lise	Fen Bilimleri	İbrahim Uçkunkaya lisesi	2004

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2010-2012	Tunceli Üniversitesi/İİBF/Ekonometri Bölümü	Araştırma Görevlisi
2012-2013	Niğde Üniversitesi/Fen Edebiyat Fakültesi/İstatistik Bölümü	Araştırma Görevlisi
2013-2016	Yıldız Teknik Üniversitesi/Fen Edebiyat Fakültesi/İstatistik Bölümü	Araştırma Görevlisi

YAYINLARI
Makale

1. Gölbaşı Şimşek Gülhayat, Aydoğdu Selahattin (2016). “Parallel analysis approach for determining dimensionality in canonical correlation analysis.” Journal of Statistical Computation and Simulation, 86(17), 3419-3431, Doi: 10.1080/00949655.2016.1161044 (Yayın No: 2804790).

Bildiri

1. Aydoğdu Selahattin, Gölbaşı Şimşek Gülhayat (2016). “A hybrid approach for determining the number of factors using empirical likelihood confidence intervals, bootstrap and parallel analysis.” 2nd International Conference on Advances in Statistics (2nd ICAS'2016) (Sözlü Bildiri)(Yayın No:2804835).

2. Çene Erhan, Aydoğdu Selahattin (2015). “What determines mathematics achievement: A bayesian network approach.” 9th International Statistics Congress. (Sözlü Bildiri)(Yayın No:1709813).

3. Aydoğdu Selahattin, Gölbaşı Şimşek Gülhayat(2015). “Application of empirical likelihood regression to a real data set with outliers.” AIOC(2015) series-International Conference on Advances in Statistics (Sözlü Bildiri)(Yayın No:1709789).

4. Aydoğdu Selahattin, Gölbaşı Şimşek Gülhayat (2014). “Bootstrap approach in paralel analysis to determine non-zero canonical correlations.” The International Conference on Trends and Perspectives in Linear Statistical Inference (LinStat2014). İsveç (Sözlü Bildiri)(Yayın No:1288201).

5. Gölbaşı Şimşek Gülhayat, Aydoğdu Selahattin (2014). “Parallel Analysis with Bootstrap Modification in Factor Analysis.” 9th International Statistics Day Symposium – ISDS. Antalya (Sözlü Bildiri)(Yayın No:1288199).

6. Gölbaşı Şimşek Gülhayat, NOYAN Fatma, Aydoğdu Selahattin (2013). “Kanonik Korelasyon Analizinde Boyut Sayısına Karar Verilmesinde Paralel Analiz Yaklaşımı.” 8.Uluslararası İstatistik Kongresi (Sözlü Bildiri)(Yayın No:781665).

7. Aydoğdu Selahattin, Gürcan Mehmet, Kılıç Muhammet Burak (2012). “Genelleştirilmiş Beta Dağılımı ve Parametrelerinin Bulunması.” 13.Uluslararası Ekonometri, Yöneylem Araştırması ve İstatistik Konferansı, EYI, (Sözlü Bildiri)(Yayın No:781369).

8. Kılıç Muhammet Burak, Gürcan Mehmet, Aydoğdu Selahattin (2010). “Bernstein Polinomlarinin Bazi Ozellikleri ve Istatistiksel Çıkarsamalari.” 7. Istatistik Gunleri Sempozyumu, (Yayın No:781099).