

T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
İSTATİSTİK ANA BİLİM DALI



**ÇOK DEĞİŞKENLİ BEHRENS-FİŞER PROBLEMİ ALTINDA
İKİ BAĞIMSIZ GRUBUN ORTALAMA VEKTÖRLERİNİ
KARŞILAŞTIRMAK İÇİN ROBUST BİR YAKLAŞIM**

Yüksek Lisans Tezi

Gölnur KARAOSMAN

Danışman
Doç. Dr. Hasan BULUT

Bu çalışma, Ondokuz Mayıs Üniversitesi Proje Yönetim Ofisi tarafından
PYO.FEN.1904.21.015 proje numarası ile desteklenmektedir.

SAMSUN
2022

TEZ KABUL VE ONAYI

Glnur KARAOSMAN tarafından, **Do. Dr. Hasan BULUT** danışmanlığında hazırlanan **“OK DEĞİŖKENLİ BEHRENS-FİSHER PROBLEMİ ALTINDA İKİ BAĞIMSIZ GRUBUN ORTALAMA VEKTÖRLERİNİ KARŞILAŖTIRMAK İİN ROBUST BİR YAKLAŖIM”** bařlıklı bu alıřma, jrimiz tarafından 20/06/2022 tarihinde yapılan sınav sonucunda oy birlięi ile bařarılı bulunarak Yksek Lisans Tezi olarak kabul edilmiřtir.

	Unvanı Adı Soyadı niversitesi Ana Bilim/Ana Sanat Dalı	İmza	Sonuç
Bařkan	Prof. Dr. Yksel NER Ondokuz Mayıs niversitesi İstatistik Ana Bilim Dalı		☒
			Kabul
			☐ Ret
ye	Do. Dr. Hasan BULUT Ondokuz Mayıs niversitesi İstatistik Ana Bilim Dalı		☒
			Kabul
			☐ Ret
ye	Do. Dr. Tolga ZAMAN ankırı Karatekin niversitesi İstatistik Ana Bilim Dalı		☒
			Kabul
			☐ Ret

Bu tez, Enstit Ynetim Kurulunca belirlenen ve yukarıda adları yazılı jri yeleri tarafından uygun grlmřtir.

ONAY
.../.../2022
Prof. Dr. Ali BOLAT
Enstit Mdr

BİLİMSEL ETİKLİĞE UYGUNLUK BEYANI

Hazırladığım yüksek lisans tezinin bütün aşamalarında bilimsel etiğe ve akademik kurallara riayet ettiğimi, çalışmada doğrudan veya dolaylı olarak kullandığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin Kaynaklar'da gösterilenlerden oluştuğunu, her unsurun enstitü yazım kılavuzuna uygun yazıldığını ve TÜBİTAK Araştırma ve Yayın Etiği Kurulu Yönetmeliği'nin 3. bölüm 9. maddesinde belirtilen durumlara aykırı davranılmadığını taahhüt ve beyan ederim.

Etik Kurul Gerekli mi?

Evet ☐ (Gerekli ise ekler kısmına ekleyiniz)

Hayır ☒

15/05/2022

Gülnur KARAOSMAN

TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI

Tez Başlığı: ÇOK DEĞİŞKENLİ BEHRENS-FİŞER PROBLEMİ ALTINDA İKİ BAĞIMSIZ GRUBUN ORTALAMA VEKTÖRLERİNİ KARŞILAŞTIRMAK İÇİN ROBUST BİR YAKLAŞIM

Yukarıda başlığı belirtilen tez çalışması için şahsım tarafından 18/05/2022 tarihinde intihal tespit programından alınmış olan özgünlük raporu sonucunda;

Benzerlik oranı : % 10

Tek kaynak oranı : % 5 çıkmıştır.

18 /05 / 2022

Doç. Dr. Hasan BULUT

ÖZET

ÇOK DEĞİŞKENLİ BEHRENS-FİSHER PROBLEMİ ALTINDA İKİ BAĞIMSIZ GRUBUN ORTALAMA VEKTÖRLERİNİ KARŞILAŞTIRMAK İÇİN ROBUST BİR YAKLAŞIM

Gölnur KARAOSMAN
Ondokuz Mayıs Üniversitesi
Lisansüstü Eğitim Enstitüsü
İstatistik Ana Bilim Dalı
Yüksek Lisans, Mayıs/2022
Danışman: Doç. Dr. Hasan BULUT

Çok değişkenli istatistiksel çıkarımda, iki bağımsız grubun ortalama vektörlerinin eşitliğini test etmek için Hotelling T^2 istatistiği kullanılır. Bu istatistiğin kullanılabilmesi için, her iki gruba ait verilerin dağılımının çok değişkenli normal dağılımlı olması ve kovaryans matrislerinin homojenliği varsayımına ihtiyaç vardır. Ancak, homojen kovaryans matrislerinin homojenliği varsayımı gerçek verilerde sağlanamayabilir. Bu varsayımın sağlanamaması durumu literatürde Behrens-Fisher problemi olarak bilinir. Çok değişkenli Behrens-Fisher problemi altında bağımsız gruplar da iki ortalama vektörünün eşitliğini test etmek için birkaç çalışma vardır. Ancak önerilen bu testler veri setindeki aykırı değerlere karşı duyarlıdır. Bu çalışmada hem çok değişkenli Behrens-Fisher problemi altında çalışabilen hem de veri setindeki aykırı değerlerden etkilenmeyen bir test önerilmektedir. Önerilen yaklaşımın performansı hem temiz (aykırı değer eklenmemiş) hem de kirli (aykırı değer eklenmiş) veri setlerinde gerçekleştirilen simülasyon çalışmalarıyla incelenmiştir. Önerilen yaklaşımın temiz veri setlerinde diğer test istatistiklerine yakın sonuçlar verdiği, kirli veri setlerinde ise hem tip-1 hata hem de testin gücü bakımından çok daha başarılı olduğu görülmüştür. Böylece önerilen yaklaşımın çok değişkenli Behrens-Fisher problemi altında aykırı değerlerden etkilenmeden iki bağımsız grubun ortalama vektörlerinin karşılaştırılmasında kullanılabileceği gösterilmiştir. Ayrıca, gerçek veri uygulamalarında önerdiğimiz yaklaşımı kullanmak için MVTests paketinde bir R fonksiyonu oluşturulmuştur.

Anahtar Sözcükler: Hotelling T^2 , Hesaplamalı Yaklaşım testi, Çok değişkenli Behrens-Fisher problemi, Robust test yaklaşımı, RobCAT fonksiyonu.

ABSTRACT

A ROBUST TEST APPROACH FOR EQUALITY OF MEAN VECTORS OF TWO INDEPENDENT GROUPS UNDER THE MULTIVARIATE BEHRENS-FISHER PROBLEM

Gülnur KARAOSMAN
Ondokuz Mayıs University
Institute of Graduate Studies
Department of Statistics
Master, May/2022
Supervisor: Assoc. Prof. Dr. Hasan BULUT

In multivariate statistical inference, the Hotelling T^2 statistic is used to test the equality of mean vectors for two independent groups. This statistic needs the multivariate normality and homogeneous covariance matrices assumptions. However, homogeneous covariance matrices assumption may not be provided in real applications. Failure of this assumption is known in the literature as the Behrens-Fisher problem. There are several studies to test the equality of two mean vectors for the independent groups under the multivariate Behrens-Fisher problem. But these studies do not interest in outliers at data sets. In this study, we propose solving problems caused by multivariate Behrens-Fisher and outliers in the dataset. We compare our proposed approach with other approaches regarding empirical size and power at simulated data that are both clean (no outlier added) and contaminated (outlier added). Thus we show that our proposed approach can be used to test the equality of mean vectors for two independent groups under multivariate Behrens-Fisher problem without being affected by outliers in the data. Moreover, we construct an R function in the MVTTests package to use our proposed approach for real data applications.

Keywords: Hotelling T^2 , Computational approach test, Multivariate Behrens-Fisher Problem, Robust test approach, RobCat Function.

TEŞEKKÜR

Tez çalışmamın konusunun belirlenmesi, sonuçların değerlendirilmesi, yazımı aşamasında bilimsel katkılarını, yardımlarını, ilgisini ve sabrını esirgemeyen tez danışmanım Sayın Doç. Dr. Hasan Bulut’a sonsuz teşekkürlerimi sunarım.

Beni bugünlere getirip yetiştiren, hayatım boyunca bana ellerinden gelen en iyi olanakları sunan, eğitimim için her zaman yanımda olan ve beni her konuda her zaman yüreklendiren babama ve anneme minnettarım.

Gölnur KARAOSMAN



İÇİNDEKİLER

ÖZET	iii
ABSTRACT	iv
TEŞEKKÜR	v
SİMGELER VE KISALTMALAR	vii
TABLolar DİZİNİ	viii
1.GİRİŞ	1
2.TEST İSTATİSTİKLERİ	7
2.1. Bennett Testi	7
2.2. Yao Testi	8
2.3. Johansen Testi	8
2.4. Nel ve Van der Merwe Testi	9
2.5. Modifiye Edilmiş Nel ve Van der Merwe Testi	9
2.6. Hesaplamalı Yaklaşım Testi (Computational Approach Test-CAT)	9
3.ROBUST HESAPLAMALI YAKLAŞIM TESTİ (RobCAT).....	12
3.1. Minimum Kovaryans Determinant (MCD) Tahmin Edicileri.....	12
3.2. Robust Hesaplamalı Yaklaşım Testi Test Algoritması	12
3.3. RobCAT Algoritması Yazılım Kullanılabilirliği	13
4.SİMÜLASYON ÇALIŞMASI.....	14
4.1. Temiz Veri İçin Tip-1 Hata Oranları	15
4.2. Temiz Veri İçin Testin Gücü	18
4.3. Kirli Veri İçin Tip-1 Hata Oranları	24
4.4. Kirli Veri İçin Testin Gücü	31
5. SONUÇLAR	39
KAYNAKLAR	40
ÖZ GEÇMİŞ.....	43

SİMGELER VE KISALTMALAR

ANOVA	: Tek Yönlü Varyans Analizi
ARE	: Ortalama Bağlı Hata
CAT	: Hesaplamalı Yaklaşım Testi
MCD	: Minimum Kovaryans Determinant Tahmin Edicisi
RMLE	: Kısıtlı En Çok Olabilirlik Tahmin Edicisi
RobCAT	: Robust Hesaplamalı Yaklaşım Testi
S_{pl}	: Birleştirilmiş (Pooled) kovaryans matrisi
T_B	: Bennett test istatistiği
T_J	: Johansen Test istatistiği
T_{M-NV}	: Modifiye edilmiş Nel ve Van der Merwe Test istatistiği
T_{NV}	: Nel ve Van der Merwe Test istatistiği
T_Y	: Yao Test istatistiği

TABLÖLAR DİZİNİ

Tablo 4.1	Temiz veride Dizayn 1 için test istatistiklerinin deneysel tip-1 hata oranları..	15
Tablo 4.2	Temiz veride Dizayn 2 için test istatistiklerinin deneysel tip-1 hata oranları..	16
Tablo 4.3	Temiz veride Dizayn 3 için test istatistiklerinin deneysel tip-1 hata oranları..	17
Tablo 4.4	Temiz veride Dizayn 4 için test istatistiklerinin deneysel tip-1 hata oranları..	18
Tablo 4.5	Temiz veride Dizayn 1 için $\Delta = 2$ iken test istatistiklerinin gücü.....	19
Tablo 4.6	Temiz veride Dizayn 2 için $\Delta = 2$ iken test istatistiklerinin gücü.....	19
Tablo 4.7	Temiz veride Dizayn 3 için $\Delta = 2$ iken test istatistiklerinin gücü.....	20
Tablo 4.8	Temiz veride Dizayn 4 için $\Delta = 2$ iken test istatistiklerinin gücü.....	21
Tablo 4.9	Temiz veride Dizayn 1 için $\Delta = 5$ iken test istatistiklerinin gücü.....	22
Tablo 4.10	Temiz veride Dizayn 2 için $\Delta = 5$ iken test istatistiklerinin gücü.....	22
Tablo 4.11	Temiz veride Dizayn 3 için $\Delta = 5$ iken test istatistiklerinin gücü.....	23
Tablo 4.12	Temiz veride Dizayn 4 için $\Delta = 5$ iken test istatistiklerinin gücü.....	24
Tablo 4.13.	Kirli veride (kirlilik oranı %10 iken) Dizayn 1 için test istatistiklerinin deneysel tip-1 hata oranları.....	25
Tablo 4.14.	Kirli veride (kirlilik oranı %10 iken) Dizayn 2 için test istatistiklerinin deneysel tip-1 hata oranları.....	26
Tablo 4.15.	Kirli veride (kirlilik oranı %10 iken) Dizayn 3 için test istatistiklerinin deneysel tip-1 hata oranları.....	26
Tablo 4.16.	Kirli veride (kirlilik oranı %10 iken) Dizayn 4 için test istatistiklerinin deneysel tip-1 hata oranları.....	27
Tablo 4.17.	Kirli veride (kirlilik oranı %30 iken) Dizayn 1 için test istatistiklerinin deneysel tip-1 hata oranları.....	28
Tablo 4.18.	Kirli veride (kirlilik oranı %30 iken) Dizayn 2 için test istatistiklerinin deneysel tip-1 hata oranları.....	29
Tablo 4.19.	Kirli veride (kirlilik oranı %30 iken) Dizayn 3 için test istatistiklerinin deneysel tip-1 hata oranları.....	30
Tablo 4.20.	Kirli veride (kirlilik oranı %30 iken) Dizayn 4 için test istatistiklerinin deneysel tip-1 hata oranları.....	31
Tablo 4.21.	Kirli veride Dizayn 1 için $\Delta = 2$ iken test istatistiklerinin gücü.....	32
Tablo 4.22.	Kirli veride Dizayn 2 için $\Delta = 2$ iken test istatistiklerinin gücü.....	33
Tablo 4.23.	Kirli veride Dizayn 3 için $\Delta = 2$ iken test istatistiklerinin gücü.....	33

Tablo 4.24.	Kirli veride Dizayn 4 için $\Delta = 2$ iken test istatistiklerinin gücü.....	34
Tablo 4.25.	Kirli veride Dizayn 1 için $\Delta = 5$ iken test istatistiklerinin gücü.....	35
Tablo 4.26.	Kirli veride Dizayn 2 için $\Delta = 5$ iken test istatistiklerinin gücü.....	36
Tablo 4.27.	Kirli veride Dizayn 3 için $\Delta = 5$ iken test istatistiklerinin gücü.....	36
Tablo 4.28.	Kirli veride Dizayn 4 için $\Delta = 5$ iken test istatistiklerinin gücü.....	37



1.GİRİŞ

İstatistiksel arařtırmalarda, sıklıkla karřılařılan problemlerden biri farklı yığınlardan gelen iki veya daha fazla gruba ait ortalamaların karřılařtırılmasıdır. İki bağımsız grubun ortalamaları arasındaki farkın eřitliğini test etmek için örneklem sayısına göre Z ve t testlerinden yararlanırken, ikiden fazla grup söz konusu olduęunda bu testler yetersiz kalmaktadır. Tek bir bağımlı deęiřken söz konusu olduęunda anakütle varyansları eřit olan ve normal daęılıma sahip ikiden fazla grup ortalamasının eřitliğini ifade eden yokluk hipotezi “Tek Yönlü Varyans Analizi (ANOVA)” ile test edilmektedir (Özdamar, 2018). ANOVA; eğitim, psikoloji, saęlık, istatistik vb. birçok bilim dalında sıklıkla bařvurulan bir istatistiksel yöntemdir. Ancak bu yöntemin uygulanabilmesi için gözlemlerin bağımsız olması, bağımlı deęiřkenin normal daęılıma sahip olması ve tüm gruplardaki varyansların homojen olması gerekmektedir.

Birden fazla bağımlı deęiřkenin aynı modelde yer aldıęı problemler için her bir deęiřkene ayrı ayrı ANOVA uygulanması, deęiřkenler arasındaki iliřkilerin göz ardı edilmesine neden olmaktadır. Bu durum, tip-I hata yapma olasılığını arttırmakta ve elde edilen istatistiksel sonuçlar yanıltıcı olmaktadır. Dolayısıyla birden fazla bağımlı deęiřken söz konusu olduęunda, çok deęiřkenli istatistiksel analizlerin kullanılması gerekmektedir (Alpar, 2011; Finch ve French, 2013).

Çok deęiřkenli istatistiksel analiz, iki veya daha fazla deęiřkeni aynı anda incelemeye yarayan istatistiksel yöntemleri tanımlamakta kullanılan genel bir terimdir. Çok deęiřkenli istatistiksel analizde birbiriyle iliřki halinde çok sayıda deęiřken söz konusudur. Ayrıca, gerçek hayatta ve modern bilimsel çalıřmalarda temel alınan birim ve deęiřken sayısı birden fazla olup, bu birim ve deęiřkenlerin de karřılıklı etkileřimleri söz konusudur. Bu nedenle çok deęiřkenli istatistiksel analiz tekniklerine ihtiyaç duyulur (Erdoğan,2018). Çok deęiřkenli normal daęılıma sahip iki veya daha fazla bağımlı deęiřken bakımından, ikiden fazla grup ortalamasının eřitliğini test etmek istedięimizde “Çok Deęiřkenli Varyans Analizi” teknięinden yararlanılmaktadır.

Çok deęiřkenli varyans analizinin özel bir durumu da iki grubun ortalama vektörlerinin karřılařtırılmasıdır. Hem çok deęiřkenli varyans analizinde hem de iki ortalama vektörünün karřılařtırılmasında ortaya çıkan problemlerden biri de

kovaryans matrislerinin homojenlik varsayımının sağlanmamasıdır. Homojenlik varsayımının sağlanmadığı durumlarda grup ortalamaları arasındaki farkın belirlenmesine yönelik ilk çalışmalar Behrens (1929) ve Fisher (1935) tarafından gerçekleştirilmiştir. Bu nedenle homojenlik varsayımının sağlanmadığı durumlarda iki veya daha fazla grubun ortalama vektörlerini karşılaştırma problemi literatürde “Behrens-Fisher Problemi” olarak adlandırılmaktadır. Behrens-Fisher problemi tek değişkenli duruma dayanmakla birlikte 1951’li yıllarda çok değişkenli problemler için de genişletilmiştir. Bağımlı değişken sayısının tek olduğu durumlarda tek değişkenli Behrens-Fisher problemi olarak adlandırılırken, bağımlı değişken sayısının iki veya daha fazla olduğu durumlarda ise bu durum çok değişkenli Behrens-Fisher problemi olarak adlandırılır. Bu tez çalışmasında da iki ortalama vektörünün eşitliğinin test edilmesinde karşılaşılan çok değişkenli Behrens-Fisher problemi üzerinde durulacaktır.

Çok değişkenli Behrens-Fisher probleminde iki grup ortalama vektörü arasındaki farkın çözümüne ilişkin ilk çalışmalar Bennett (1951), James (1954), Ito (1964), Yao (1965), Subrahmaniam ve Subrahmaniam (1973), Johansen (1980), Nel ve Van der Merwe (1986), Hwang ve Paulson (1986), Kim (1992), Christensen ve Rencher (1997), Krishnamoorthy ve Yu (2004) ve Yanagira ve Yuan (2005) tarafından ortaya koyulmuştur. Çok değişkenli Behrens-Fisher problemine ilk yaklaşık çözüm 1951 yılında Bennett tarafından geliştirilmiştir. Bennett, Scheffé (1943)’nin sonuçlarını çok değişkenli durum için genişleterek genelleştirmiş ve Behrens-Fisher problemine bir çözüm önermiştir. Ancak Bennett’in geliştirdiği istatistik, gözlemlerin sırasına bağlı olduğundan, çok büyük örnek çaplarında kullanımı oldukça zordur (Yanagihara ve Yuan, 2005). James (1954), test istatistiğinin kritik değerini belirlemek için tek değişkenli Behrens-Fisher probleminde önerdiği birinci ve ikinci derece seri açılımı yaklaşımını çok değişkenli Behrens-Fisher problemi için genelleştirmiştir. Ancak bu genelleme iki veya daha fazla faktör söz konusu olduğunda oldukça karmaşık bir hale gelmiştir. Bu sebeple çok değişkenli Behrens-Fisher problemi için Yao (1965), Johansen (1980) ve Kim (1992) başta olmak üzere birçok araştırmacı tarafından Welch (1947)’in yaklaşık serbestlik derecesi yöntemi kullanılarak test istatistikleri önerilmiştir (Sandal,2020).

Yao (1965), Hotelling’in T^2 istatistiğini değiştirerek yaklaşık F dağılımına sahip önerdiği test istatistiğini James’in birinci derece seri açılımı yaklaşımı ile

karşılaştırmış ve yaklaşık bir çözüm elde etmiştir. Subrahmaniam ve Subrahmaniam (1973), yaptıkları simülasyon çalışmaları ile Bennett'in önerdiği test istatistiğinin örnek çaplarının eşit olmadığı durumlarda deneysel gücün kötü olduğunu göstermiştir. Johansen (1980), Behrens-Fisher problemi için Welch'in tek değişkenli yaklaşık serbestlik derecesi çözümlemesini genelleştirmiştir. Nel ve Van der Merwe (1986), Wishart matrislerinin doğrusal birleşimlerini kullanarak Welch'in yaklaşık serbestlik derecesi çözümlemesine farklı bir genelleme sunmuş ve önerdikleri test istatistiğini tip-1 hata oranı ve testin gücü bakımından karşılaştırmak amacıyla bir simülasyon çalışması yapmıştır. Hwang ve Paulson (1986) da benzer bir çalışma yapmıştır. Kim (1992) ise iki ortalama vektörü için güven elipsoidlerinin geometrisini kullanarak bir test istatistiği önermiş ve Yao'nun testi ile karşılaştırmıştır.

Christensen ve Rencher (1997), çok değişkenli Behrens-Fisher problemi söz konusu olduğu durumda, Hotelling T^2 ve Bennett (1951), Yao (1965), Johansen (1980), Nel ve Van der Merwe (1986), Hwang ve Paulson (1986), Kim (1992) tarafından önerilen yaklaşık metotları, tip-1 oranları ve testin güç değerleri bakımından karşılaştırmak amacıyla bir simülasyon çalışması yapmış ve Van der Merwe ile Kim testlerinin daha yüksek güce sahip olduğunu göstermiştir. Krishnamoorthy ve Yu (2004) ise, Nel ve Van der Merwe (1986) tarafından sunulan test istatistiğine bir düzeltme uygulayarak modifiye edilmiş Nel ve Van der Merwe testi önermiştir.

Bahsedilen çalışmaların hemen hemen hepsi iki ortalama vektörünün eşitliği için önerilen asimptotik çözümlerdir. Bilindiği üzere asimptotik bir dağılıma sahip testlerde özellikle küçük örnek çaplarında tip-1 hata bakımından nominal değerden uzak sonuçlar vermektedir. Bu soruna çözüm olarak yeniden örneklemeye dayalı Parametrik Bootstrap yönteminin özel bir hali olarak bilinen Hesaplamalı Yaklaşım Testi (Computational Approach Test- CAT-2018) önerilmiştir. Yeniden örneklemeye dayalı bu yöntemde elde edilen örnekten faydalanarak yapay örnekler üretilmiş ve test istatistiğinin yapay bir dağılımı oluşturulmuştur.

CAT yöntemi son yıllarda birçok istatistiksel araştırmada sıklıkla kullanılan bir yöntemdir. Bu konuyla ilgili başlıca çalışmalar Pal ve diğerleri (2007), Chang ve Pal (2008), Chang ve diğerleri (2010, 2011, 2016), Gökpınar ve Gökpınar (2012, 2014),

Gökpınar ve diğerleri (2013), Mutlu ve diğerleri (2017), Jafari ve Kazemi (2017) ve Erdoğan (2018)'dir.

Pal ve diğerleri (2007), Parametrik Bootstrap Yönteminin özel bir hali olan Hesaplamalı Yaklaşım Yöntemini (Computational Approach Test-CAT) önermişlerdir. Bu yöntem yokluk hipotezinin doğruluğu altında kısıtlı en çok olabilirlik (Restricted Maximum Likelihood-RML) tahmin edicilerine dayanmaktadır. Ayrıca, test istatistiğinin dağılımını teorik olarak bulmak gerekmediğinden ve p değerini doğrudan elde ettiğinden dolayı kullanımı kolay bir yöntemdir.

Chang ve Pal (2008), Behrens-Fisher problemi için CAT yöntemini önermişler ve bu yöntemi Welch–Satterthwaite testi (WST), Cochran–Cox testi (CCT), Genelleştirilmiş p-değeri testi (GPT) ve Singh–Saxena–Srivastava testine (SSST) göre karşılaştırıp, CAT yönteminin performansının daha iyi olduğunu göstermişlerdir.

Chang ve arkadaşları (2010), varyansların eşitliği varsayımı altında ikiden fazla normal yığın ortalamasının eşitliğini test etmek için CAT yöntemine dayalı alternatif bir test önermişler ve bu testi, tek yönlü ANOVA ve ANOM (Analysis of Means-Ortalamaların Analizi) yöntemiyle karşılaştırmışlardır. Chang ve arkadaşları 2011 yılında ise ikiden fazla gama dağılımının ortalamasının eşitliğini test etmek için CAT yöntemini uygulamışlardır.

Gökpınar ve Gökpınar (2012), Chang ve arkadaşlarının testini geliştirerek varyansların eşit olmadığı durumda k yığın ortalamasının eşitliğini test etmek için CAT yöntemine dayalı alternatif bir test önermişlerdir. Ayrıca Brown-Forsythe, Weerahandi'nin Genelleştirilmiş p, Parametrik Bootstrap ve Welch testleri ile karşılaştırmışlardır. Gökpınar ve diğerleri (2013), ölçek parametrelerinin eşitsizliği durumu altında ikiden fazla ters Gauss dağılımının ortalamalarının eşitliği için hesaplamalı yaklaşım testine dayanan bir test prosedürü önermişlerdir. Bu yöntemi literatürde bulunan testler ile karşılaştırmışlardır. Gökpınar ve Gökpınar(2014), ikiden fazla log-normal yığının ortalamasının eşitliğinin testi için CAT yöntemine dayalı yeni bir test önermişlerdir. Farklı örnek çapları ve grup sayıları için Monte-Carlo simülasyonunu kullanarak bazı mevcut metotlarla karşılaştırmışlardır (Aksoy ve Gökpınar, 2020).

Chang ve diğerkleri (2016), yaptıkları çalışmada varyans homojenliğinin testinde en çok olabilirlik metoduna dayalı olarak CAT yöntemini önermişler ve yaptıkları simülasyon çalışmasında bu testin karşılaştırılan testlere göre daha iyi bir test olduklarını göstermişlerdir.

Mutlu ve diğerkleri (2017), heterojen varyans varsayımı altında Gökpınar ve Gökpınar (2012)'ın yaptığı prosedürü geliştirerek, tek yönlü varyans analizi (ANOVA) için CAT yöntemine dayanan ve yaptıkları diğerk popüler testlerle karşılaştırmaları sonucu daha yüksek güç değerlerine sahip olduğu bilinen test istatistiğini geliştirmişlerdir.

Jafari ve Kazemi (2017), iki log-normal dağılımın ortalamalarının eşitliğini test etmek için CAT yöntemine dayalı alternatif bir test önermişlerdir. Daha sonra Jafari ve arkadaşları tarafından ikiden fazla log-normal yığının ortalamalarını karşılaştırmak için, CAT yöntemine dayalı alternatif bir test önermişlerdir. Bu testi; F-testi, en çok olabilirlik oranı testi, genelleştirilmiş p-değeri yaklaşımı ile karşılaştırmışlardır.

Erdoğan (2018), çok değişkenli normal dağılım varsayımı altında iki ortalama vektörün eşitliği test etmek için Hesaplamalı Yaklaşım Testine dayalı yeni bir test istatistiği ileri sürmüştür. Sunulan yeni test istatistiği, tip-1 hata yapma olasılığı ve testin gücü bakımından Bennett (1951), Yao (1965), Johansen (1980), Nel ve Van Der Merwe (1986), Krishnamoorthy ve Yu (2004) testleri ile karşılaştırmıştır. Çalışmanın sonuçları, yeni testin oldukça iyi performans gösterdiğini ortaya koymuştur.

Literatürde, kovaryans matrislerinin heterojenliği altında, iki ortalama vektörünün eşitliği için yaygın olarak kullanılan testlerin özellikle küçük ve farklı örnek çaplarında ve yüksek değişken sayısı varlığı altında, I.tip hata bakımından çok iyi sonuç vermediği yapılan çalışmalarda gösterilmiştir (Christensen ve Rencher, 1997). Uygulamalarda yüksek örnek çaplarıyla çalışmak çok da gerçekçi bir durum değildir. Ayrıca birçok problemde değişken sayısı oldukça fazla ve veri setinde aykırı değerler olabilmektedir. Dolayısıyla, özellikle küçük örnek çaplarında, yüksek değişken sayısında ve aykırı değerlerin bulunduğu veri setlerinde, tip-1 hata oranları nominal değerden uzaklaşmakta ve testin gücü azalmaktadır. Verilerin analizinde daha iyi tip-1 hata ve daha yüksek testin gücü elde etmek sonuçların güvenilirliğini

artırmaktadır. Ayrıca tanıtılan bu yöntemler veri setindeki aykırı değerlere karşı dayanıklı değildir.

Bu tez çalışmasının amacı, çok değişkenli Behrens-Fisher problemi altında iki bağımsız grubun ortalama vektörlerini karşılaştırmak amacıyla aykırı değerlerden etkilenmeyen bir test yaklaşımı geliştirmektir. Önerilen yaklaşım CAT algoritmasına dayanmaktadır.

Çalışmanın ikinci bölümünde, çok değişkenli normal dağılıma sahip ancak kovaryans matrislerinin homojen olmadığı durumlarda iki bağımsız grubun ortalama vektörlerinin eşitliği testi için geliştirilen bazı test istatistiklerine yer verilmiştir. Bu tezde önerilen robust yaklaşımla karşılaştırmak üzere iki bağımsız grup ortalamasının eşitliği hipotezi için önerilmiş olan Bennett (1951), Yao (1965), Johansen (1980), Nel ve Van der Merwe (1986), Modifiye edilmiş Nel ve Van der Merwe (2004) ve Hesaplamalı Yaklaşım (CAT-2018) testleri detaylı olarak tanıtılmıştır.

Üçüncü bölümde, Parametrik Bootstrap yöntemi olan CAT yönteminde klasik tahmin ediciler yerine minimum kovaryans determinant (MCD) tahmin edicilerinin kullanılmasına dayanarak robust bir yaklaşımı önerilmiş ve CAT yöntemine dayanan robust bir test prosedürü geliştirilmiştir.

Dördüncü bölümde, Monte Carlo simülasyonu yoluyla, daha önceden önerilen test istatistikleriyle önerilen CAT yöntemine dayalı robust yaklaşımın (RobCAT) farklı değişken sayısı, örnek çapı, kovaryans matrisleri ve ortalama vektörleri kombinasyonları kullanarak testlerin tip-1 hata oranları ve testin gücü bakımından karşılaştırılmış ve sonuçlar tablo haline getirilerek özetlenmiştir.

Beşinci bölümde ise, yapılan simülasyon sonuçlarına yer verilmiş ve önerilerde bulunulmuştur.

2.TEST İSTATİSTİKLERİ

$X_{11}, X_{12}, \dots, X_{1n_1}$ örneği $N_p(\mu_1, \Sigma_1)$ ve $X_{21}, X_{22}, \dots, X_{2n_2}$ örneği $N_p(\mu_2, \Sigma_2)$ çekilen bir örnek olsun. Çekilen tüm örneklerin bağımsız olduğu varsayılmaktadır.

$\Sigma_1 = \Sigma_2 = \Sigma$ olduğunda, iki örneklem için Hotelling T^2 test istatistiği $H_0: \mu_1 = \mu_2$ yokluk hipotezini test etmek için kullanılır ve bu istatistik aşağıdaki gibi tanımlanır:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2)' S_{pl}^{-1} (\bar{X}_1 - \bar{X}_2) \quad (2.1)$$

Burada $(.)^T$ herhangi bir matrisin transpozu, $\bar{X}_k$ ve n_k k. grup için sırasıyla örnek ortalama vektörü ve örnek boyutudur ($k=1,2$). Ayrıca S_{pl} birleştirilmiş kovaryans matrisidir ve aşağıdaki gibi hesaplanır:

$$S_{pl} = \frac{1}{n_1 + n_2 - 2} [(n_1 - 1)S_1 + (n_2 - 1)S_2] \quad (2.2)$$

Burada S_1 ve S_2 sırasıyla birinci ve ikinci gruplar için örnek varyans-kovaryans matrisleridir.

Ancak Hotelling T^2 test istatistiğini uygulayabilmemiz için çok değişkenli normallik varsayımına ve kovaryans matrislerinin eşitliği ($\Sigma_1 = \Sigma_2 = \Sigma$) durumuna ihtiyaç duyarız. Kovaryans matrislerinin birbirinden farklı olması durumu çok değişkenli Behrens-Fisher problemi olarak adlandırılır ve bu problem altında Hotelling T^2 istatistiğinin tip-1 hata oranı nominal değerlere yakın değildir. Bu yüzden bu istatistik çok değişkenli Behrens-Fisher problemi altında uygulanamaz. Bu problem üzerine birçok çözüm önerilmiştir. Bunlardan bazıları: Bennett(1951), Yao(1965), James(1954), Johansen (1980), Nel ve Van der Merwe (1986), Modifiye edilmiş Nel ve Van der Merwe (2004) ve Hesaplamalı Yaklaşım (CAT-2018) testleridir.

2.1. Bennett Testi

Çok değişkenli Behrens-Fisher problemi altında, iki ortalama vektörünün eşitliğini test etmek için Bennett (1951) tarafından önerilen test istatistiği kullanılabilir. Bu istatistik $n_2 \geq n_1$ olmak üzere aşağıdaki gibi tanımlanır:

$$T_B = n_1 \bar{Z}' S_Z^{-1} \bar{Z} \quad (2.3)$$

Burada: $z_i = X_{1i} - \sqrt{\frac{n_1}{n_2}} X_{2i} + \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} X_{1i} - \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i}$ ($i = 1, 2, \dots, n_1$), $\bar{Z}$ ve

S_z için sırasıyla ortalama vektörü ve kovaryans matrisidir. Bu istatistik F istatistiğine dönüştürülerek aşağıdaki gibi hesaplanır:

$$F_B = \frac{n_1 - p}{p(n_1 - 1)} T_B \sim F_{p, n_1 - p} \quad (2.4)$$

2.2. Yao Testi

Yao (1965), Tukey (1959) tarafından önerilen Welch testinin bir uzantısını kullanarak bir test istatistiği önerdi. Yao'nun testi, T_*^2 istatistiğine dayalı olup, bu istatistik Eşitlik (2.5) ile verilmiştir.

$$T_*^2 = (\bar{X}_1 - \bar{X}_2)' S_e^{-1} (\bar{X}_1 - \bar{X}_2) \quad (2.5)$$

Burada:

$$S_e = \frac{1}{n_1} S_1 + \frac{1}{n_2} S_2 \quad (2.6)$$

ve $\bar{X}_k$ ve S_k , $k = (1, 2)$ için k . örneğin sırasıyla ortalama vektörü ve kovaryans matrisidir. Yao'nun test istatistiği ise aşağıdaki gibi tanımlanır:

$$T_Y = \frac{w - p + 1}{p w} T_*^2 \sim F_{p, w - p + 1} \quad (2.7)$$

Burada $\frac{1}{w} = \frac{1}{(T_*^2)^2} \sum_{k=1}^2 \frac{1}{n_k - 1} \left[(\bar{X}_1 - \bar{X}_2)' S_e^{-1} \frac{S_1}{n_1} S_e^{-1} (\bar{X}_1 - \bar{X}_2) \right]^2$ dir.

2.3. Johansen Testi

Johansen(1980) tarafından önerilen test istatistiği Eşitlik (2.8)'de verilmiştir.

$$T_j = \frac{T_*^2}{C} \sim F_{p, v} \quad (2.8)$$

Burada T_*^2 Eşitlik (2.5)'de verildiği gibi, $v = p(p + 2)/(3D)$ şeklinde hesaplanır. Ayrıca C ve D sabitleri aşağıdaki gibi hesaplanır:

$$C = p - 2D - \frac{6D}{p(p-1)+2} \quad (2.9)$$

$$D = \sum_{k=1}^2 \frac{1}{2(n_k - 1)} \{iz[(I - V^{-1}V_i)^2] + [iz(I - V^{-1}V_i)]^2\} \quad (2.10)$$

Ayrıca $k = 1, 2$ için $V_k = (S_k / n_k)^{-1}$ ve $V = V_1 + V_2$ olarak hesaplanır.

2.4. Nel ve Van der Merwe Testi

Nel ve Van der Merwe (1986) tarafından önerilen test istatistiği, Eşitlik (2.5)'teki T_*^2 istatistiği kullanılarak hesaplanır. Bu test istatistiği aşağıdaki gibi tanımlanır:

$$T_{NV} = \frac{V_{NV-p+1}}{p V_{NV}} T_*^2 \sim F_{p, V_{NV-p+1}} \quad (2.11)$$

Burada $S_e = \frac{1}{n_1} S_1 + \frac{1}{n_2} S_2$ olarak hesaplanır ve V_{NV} şu şekilde tanımlanır:

$$V_{NV} = \frac{iz(S_e)^2 + [iz(S_e)]^2}{\frac{1}{n_1-1} \left\{ iz\left(\frac{S_1}{n_1}\right)^2 + \left[iz\left(\frac{S_1}{n_1}\right) \right]^2 \right\} + \frac{1}{n_2-1} \left\{ iz\left(\frac{S_2}{n_2}\right)^2 + \left[iz\left(\frac{S_2}{n_2}\right) \right]^2 \right\}} \quad (2.12)$$

2.5. Modifiye Edilmiş Nel ve Van der Merwe Testi

Krishnamoorthy ve Yu (2004), Nel ve Van der Merwe'nin test istatistiğini modifiye ederek yeniden oluşturmuştur. Modifiye edilmiş bu test istatistiği aşağıdaki gibi tanımlanır:

$$T_{M-NV} = \frac{(V_{M-NV-p+1}) T_*^2}{p V_{M-NV}} \sim F_{p, V_{M-NV-p+1}} \quad (2.13)$$

Burada T_*^2 Eşitlik (2.5)'de verildiği gibi ve V_{M-NV} aşağıdaki gibi tanımlanır:

$$V_{M-NV} = \frac{p(p+1)(n-2)}{\hat{\phi}_1 + \hat{\phi}_2} \quad (2.14)$$

ve burada $\hat{\phi}_1$ ve $\hat{\phi}_2$ sırası ile

$$\hat{\phi}_1 = \frac{n_2^2 (n-2)}{n^2 (n_1-1)} \{iz(S_1 \bar{S}^{-1})\}^2 + \frac{n_1^2 (n-2)}{n^2 (n_2-1)} \{iz(S_2 \bar{S}^{-1})\}^2 \quad (2.15)$$

$$\hat{\phi}_2 = \frac{n_2^2 (n-2)}{n^2 (n_1-1)} iz(S_1 \bar{S}^{-1} S_1 \bar{S}^{-1}) + \frac{n_1^2 (n-2)}{n^2 (n_2-1)} iz(S_2 \bar{S}^{-1} S_2 \bar{S}^{-1}) \quad (2.16)$$

eşitlikleri ile ifade edilir. Ayrıca;

$$\bar{S} = \frac{n_2}{n} S_1 + \frac{n_1}{n} S_2 \quad (2.17)$$

olarak bulunur.

2.6. Hesaplamalı Yaklaşım Testi (Computational Approach Test-CAT)

İstatistiksel hipotez testlerinde kritik problemlerden biri de test istatistiklerinin teorik dağılımının kesin olarak bilinmemesidir. Ayrıca birçok test istatistiğinin teorik dağılımı bulunamamaktadır. Bu durum söz konusu olduğunda, asimptotik yöntemler kullanılarak test istatistiğinin yaklaşık bir dağılımı elde edilebilir. Ancak bu

yöntemle, büyük örnek çaplarında iyi sonuçlar verebilirken özellikle küçük örnek çaplarında tip-1 hata bakımından kabul edilebilen sınırın çok üstünde sonuçlar verebilmektedirler (Erdoğan, 2018).

Tip-1 hata oranının nominal değerlerden çok uzak olması problemini çözmek için yeniden örnekleme yöntemlerine dayalı birçok yöntem geliştirilmiştir. Bootstrap yöntemi bunlardan en önemlileridir. Bootstrap yöntemleri parametrik ve parametrik olmayan bootstrap yöntemleri olarak ikiye ayrılır. Lee (1994), eğer örneğin geldiği dağılım biliniyorsa, parametrik bootstrap yönteminin parametrik olmayan bootstrap yönteminden daha iyi olduğunu gösterdi. Parametrik bootstrap yönteminin özel bir hali olan CAT, test istatistiklerinin dağılımının teorik olarak bulunamadığı durumlarda simülasyonlar ile test istatistiklerinin dağılımını oluşturmaya dayanmaktadır. Bu yöntem, p-değerini tahmin etmenin en etkili yoludur. Pal ve diğerleri, sıfır hipotezi altında kısıtlı en çok olabilirlik tahmin edicilerini (RMLE) kullanarak CAT yöntemini önerdi. Bu yöntem, örnekleme dağılımı hakkında herhangi bir bilgiye ihtiyaç duymaz ve doğrudan bir p değeri elde ettiği için kullanımı kolaydır (Pal ve diğerleri, 2007).

Gökpinar ve diğerleri (2020), farklı dağılım parametreleri ile çok değişkenli normal dağılımlardan iki bağımsız grubun ortalama vektörlerin eşitliğini test etmek için CAT'e dayalı bir algoritma önerdi. CAT yöntemi, $H_0: \mu_1 = \mu_2$ altında aşağıda verilen kısıtlı en çok olabilirlik tahmin edicilerine dayanmaktadır. Sıfır hipotezi altındaki log-olabilirlik fonksiyonu aşağıdadır:

$$\ln L = \frac{-p(n_1+n_2)}{2} \ln(2\pi) - \frac{n_1}{2} \ln(|\Sigma_1|) - \frac{n_2}{2} \ln(|\Sigma_2|) - \frac{1}{2} i_z \left(\Sigma_1^{-1} \sum_{i=1}^{n_1} [X_{1i} - \mu]' [X_{1i} - \mu] \right) - \frac{1}{2} i_z \left(\Sigma_2^{-1} \sum_{i=1}^{n_2} [X_{2i} - \mu]' [X_{2i} - \mu] \right) \quad (2.18)$$

Burada μ , H_0 altında bilinmeyen ortak ortalama vektör parametresidir. Denklem (2.18)'de Σ_1, Σ_2 ve μ 'nün kısmi türevlerini alarak kısıtlı en çok olabilirlik tahmin edicileri aşağıdaki gibi elde edilir:

$$\hat{\Sigma}_{1(RML)} = \frac{\sum_{i=1}^{n_1} [X_{1i} - \hat{\mu}_{(RML)}][X_{1i} - \hat{\mu}_{(RML)}]'}{n_1} \quad (2.19)$$

$$\hat{\Sigma}_{2(RML)} = \frac{\sum_{i=1}^{n_2} [X_{2i} - \hat{\mu}_{(RML)}][X_{2i} - \hat{\mu}_{(RML)}]'}{n_2} \quad (2.20)$$

$$\hat{\mu}'_{(RML)} = [n_1 \hat{\Sigma}_{1(RML)}^{-1} + n_2 \hat{\Sigma}_{2(RML)}^{-1}]^{-1} (n_1 \hat{\Sigma}_{1(RML)}^{-1} \bar{X}_1 + n_2 \hat{\Sigma}_{2(RML)}^{-1} \bar{X}_2) \quad (2.21)$$

Bu denklemlerin yakın formları yoktur. Bu nedenle onları $(\hat{\Sigma}_1^{(k)}, \hat{\Sigma}_2^{(k)}, \hat{\mu}^k)$ tahminleri olarak, k-adım tahminlerinden yinelemeli olarak aşağıdaki gibi elde edebiliriz.

$$\hat{\Sigma}_1^{(k+1)} = \frac{\sum_{i=1}^{n_1} [X_{1i} - \hat{\mu}^k][X_{1i} - \hat{\mu}^k]'}{n_1} \quad (2.22)$$

$$\hat{\Sigma}_2^{(k+1)} = \frac{\sum_{i=1}^{n_2} [X_{2i} - \hat{\mu}^k][X_{2i} - \hat{\mu}^k]'}{n_2} \quad (2.23)$$

$$\hat{\mu}^{(k+1)} = \left[n_1 (\hat{\Sigma}_1^{(k+1)})^{-1} + n_2 (\hat{\Sigma}_2^{(k+1)})^{-1} \right]^{-1} \left(n_1 (\hat{\Sigma}_1^{(k+1)})^{-1} \bar{X}_1 + n_2 (\hat{\Sigma}_2^{(k+1)})^{-1} \bar{X}_2 \right) \quad (2.24)$$

Biz burada $\hat{\mu}^{(0)} = [n_1 S_1^{-1} + n_2 S_2^{-1}]^{-1} (n_1 S_1^{-1} \bar{X}_1 + n_2 S_2^{-1} \bar{X}_2)$ olarak kullandık. $\hat{\Sigma}_1^{(k)}$, $\hat{\Sigma}_2^{(k)}$ ve $\hat{\mu}^{(k)}$; $\hat{\Sigma}_{1(RML)}$, $\hat{\Sigma}_{2(RML)}$ ve $\hat{\mu}_{(RML)}$ olarak belirtilen sıfır hipotezi altında kısıtlı en çok olabilirlik tahmin edicilerine yakınsanır.

Gökpınar ve diğerleri (2020), bu kısıtlı en çok olabilirlik tahmin edicilerini kullanarak çok değişkenli Behrens-Fisher problemi altında $H_0: \mu_1 = \mu_2$ hipotezini test etmek için aşağıdaki CAT algoritmasını önermiştir:

Adım 1: Test istatistiği, örneklem kullanılarak hesaplanır.

$$\widehat{\eta}_{ML} = (\bar{X}_1 - \bar{X}_2)' \left(\frac{\hat{\Sigma}_1}{n_1} + \frac{\hat{\Sigma}_2}{n_2} \right)^{-1} (\bar{X}_1 - \bar{X}_2) \quad (2.25)$$

Burada, $\hat{\Sigma}_1 = \frac{\sum_{i=1}^{n_1} [X_{1i} - \hat{\mu}][X_{1i} - \hat{\mu}]'}{n_1}$, $\hat{\Sigma}_2 = \frac{\sum_{i=1}^{n_2} [X_{2i} - \hat{\mu}][X_{2i} - \hat{\mu}]'}{n_2}$ ve $\hat{\mu} = \hat{\mu}^0$ 'nin başlangıç değerinden daha öncesinde bahsedilmiştir.

Adım 2: H_0 hipotezi altında, μ, Σ_1 ve Σ_2 parametrelerinin kısıtlı en çok olabilirlik tahmin edicileri, Eşitlik (2.22-2.24) yinelemeli olarak kullanılarak $\hat{\Sigma}_{1(RML)}$, $\hat{\Sigma}_{2(RML)}$ ve $\hat{\mu}_{(RML)}$ hesaplanır.

Adım 3: $N_p(\hat{\mu}_{(RML)}, \hat{\Sigma}_{i(RML)}), k = 1, 2$ dağılımından rastgele üretilmiş bir M yapay örnekleri oluşturulur. Türetilen her bir örnek için $\hat{\eta}_{ML}^{(j)}$ ($j = 1, 2, \dots, M$) değeri tekrar hesaplanır.

Adım 4: $\hat{p} = \# \left(\hat{\eta}_{ML}^{(j)} > \hat{\eta}_{ML} \right) / M$ yardımıyla p değeri elde edilir. $\hat{p} < \alpha$ olduğunda H_0 hipotezi reddedilir.

3. ROBUST HESAPLAMALI YAKLAŞIM TESTİ (RobCAT)

Çok değişkenli Behrens-Fisher problemi altında şimdiye kadar önerilen test istatistikleri incelendiğinde bu test istatistiklerinin verideki aykırı değerlere duyarlı olduğu görülmektedir. Bu çalışma, verilerdeki aykırı değerlerden etkilenmeyen bir test süreci önermektedir. Bu amaçla Gökpınar ve diğerleri (2020) tarafından önerilen CAT algoritmasını ve Willems ve diğerleri (2002), Todorov ve Filzmoser (2010) tarafından robust bir test elde etmek için kullanılan MCD tahmin edicilerini birleştirdik.

Burada klasik tahminler yerine Minimum Kovaryans Determinant tahmin edicilerini kullanmayı öneriyoruz.

3.1. Minimum Kovaryans Determinant (MCD) Tahmin Edicileri

Rousseeuw (1985), çok değişkenli verilerin konum parametresinin MCD tahmin edicisini, minimum kovaryans matrisi determinantına sahip X 'in h noktalarının ortalama vektörü olarak tanımladı. Dağılım parametresinin MCD tahmincisi de bu alt kümenin kovaryans matrisidir. MCD tahmin edicileri α kırpma oranı ve $h = (1 - \alpha) * n$ olarak temel alınır. Ayrıca MCD tahmin edicilerinin kırılma noktası, kırpma oranına eşittir (Bulut ve Öner, 2017).

MCD tahminlerini elde etmek için, R programlama dilindeki *robustbase* paketindeki *covMcd()* fonksiyonu kullanılmıştır (Bulut, 2019).

3.2. Robust Hesaplamalı Yaklaşım Testi Algoritması

Bu yaklaşımda n_1 ve n_2 yerine örneklem boyutu olarak h_1 ve h_2 uzunluklarını kullandık. Çünkü Bölüm 3.1.'deki tanıma göre, h gözlemlerinden elde edilen en çok olabilirlik tahminlerinin MCD tahmin edicileri diyebiliriz.

Farklı dağılım parametrelerine sahip çok değişkenli normal dağılımlardan gelen ortalama vektörlerinin eşitliğini test etmek için, aşağıdaki gibi CAT yöntemine ve MCD tahminlerine dayalı robust bir algoritma öneriyoruz:

Adım 1: Test istatistiği, örneklem kullanılarak hesaplanır.

$$\widehat{\eta}_{ML} = (\bar{X}_1 - \bar{X}_2)' \left(\frac{\hat{\Sigma}_1}{h_1} + \frac{\hat{\Sigma}_2}{h_2} \right)^{-1} (\bar{X}_1 - \bar{X}_2) \quad (3.1)$$

Burada, $\hat{\Sigma}_1 = \frac{\sum_{i=1}^{h_1} |X_{1i} - \hat{\mu}| |X_{1i} - \hat{\mu}|'}{h_1}$, $\hat{\Sigma}_2 = \frac{\sum_{i=1}^{h_2} |X_{2i} - \hat{\mu}| |X_{2i} - \hat{\mu}|'}{h_2}$ ve $\hat{\mu} = [h_1 S_1^{-1} + h_2 S_2^{-1}]^{-1} (h_1 S_1^{-1} \bar{X}_1 + h_2 S_2^{-1} \bar{X}_2)$, h_1 ve h_2 sırasıyla birinci ve ikinci gruplardan en iyi alt kümedeki gözlemlerin sayısıdır. Ayrıca, $(\bar{X}_1, S_1)$ ve $(\bar{X}_2, S_2)$ sırasıyla birinci ve ikinci örneklerdeki MCD tahminleridir.

Adım 2: H_0 hipotezi altında, μ , Σ_1 ve Σ_2 parametrelerinin kısıtlı en çok olabilirlik tahmin edicileri (RMLE), sırasıyla h_1 ve h_2 örnek boyutlarına sahip en iyi alt kümeleri kullanarak Eşitlik (2.22-2.24)'den yinelemeli olarak $\hat{\Sigma}_{1(RML)}$, $\hat{\Sigma}_{2(RML)}$ ve $\hat{\mu}_{(RML)}$ hesaplanır. Bu, $k=1$ ya da 2 için eşitlikte n_k yerine h_k değerlerinin kullanıldığı anlamına gelir.

Adım 3: $N_p(\hat{\mu}_{(RML)}, \hat{\Sigma}_{k(RML)})$, $k = 1, 2$ dağılımından rastgele üretilmiş bir M yapay örnekleri oluşturulur. Türetilen her bir örnek için $\hat{\eta}_{ML}^{(j)}$ ($j = 1, 2, \dots, M$) değeri tekrar hesaplanır.

Adım 4: $\hat{p} = \# \left(\hat{\eta}_{ML}^{(j)} > \hat{\eta}_{ML} \right) / M$ yardımıyla p değeri elde edilir. $\hat{p} < \alpha$ olduğunda H_0 hipotezi reddedilir. (Burada $\#$ işareti ile $\hat{\eta}_{ML}^{(j)}$ değerlerinden kaç tanesinin $\hat{\eta}_{ML}$ değerinden büyük olduğu belirlenir ve M değerine bölünerek $\hat{p}$ değeri elde edilir.)

3.3. RobCAT Algoritması Yazılım Kullanılabilirliği

RobCAT algoritması, (Bulut, 2019) CRAN'ında "MVTtests" R paketinde "RobCAT" fonksiyonu kullanılarak uygulanır. "RobCAT" fonksiyonu çok değişkenli Behrens-Fisher problemi altında iki ortalama vektörünü karşılaştırmak için robust CAT algoritmasına dayalı olarak p değerini hesaplar. Bu fonksiyon dört argümandan oluşur. X ve Y bağımsız değişkenleri, sırasıyla birinci ve ikinci örnek için *matrix* ya da *data.frame* 'dir. M bir yineleme sayısıdır ve varsayılan değer $M=1000$ 'dir. Son olarak, son argüman *alpha*'dır. (0,1) aralığında değerler alır ve aykırı değer eklenmemiş verinin yüzdesini ifade eder.

4.SİMÜLASYON ÇALIŞMASI

Çok değişkenli normal dağılım varsayımı ve heterojen dağılım parametreleri altında temiz (aykırı değer eklenmemiş) ve kirli (aykırı değer eklenmiş) veri setlerinde önerdiğimiz RobCAT yaklaşımımızı diğer test istatistikleri ve CAT ile karşılaştırmak için bir simülasyon çalışması yapıyoruz. Bu çalışmada test edilecek hipotezler:

$$H_0: \mu_1 = \mu_2 = \mu$$

$$H_1: \mu_1 \neq \mu_2$$

olarak tanımlanır. Bu karşılaştırma için Bölüm 2’de tanıtılan Bennett (T_B), Yao (T_Y), Johansen (T_J), Nel ve Van der Merwe (T_{NV}), Modifiye edilmiş Nel ve Van der Merwe (T_{M-NV}), Hesaplamalı Yaklaşım testi (CAT) test istatistikleriyle CAT algoritmasına dayalı önerdiğimiz RobCAT yaklaşımının; farklı değişken sayısı, örnek çapı, kovaryans matrisleri ve ortalama vektörleri kombinasyonlarını kullanarak testlerin deneysel tip-1 hata oranı ve güçleri bakımından Monte Carlo simülasyonu yoluyla karşılaştırıyoruz.

Sıfır hipotezi (H_0) altında tip-1 hata oranı için çok değişkenli normal dağılımda aynı ortalama vektörleri için örnekler üretiyoruz. Burada hem aykırı değerlerin eklenmediği (temiz) hem de aykırı değerlerin eklendiği (kirli) veri için ortalama vektörleri:

$$\mu_1 = \mu_2 = [0 \ 0 \ 0 \ \dots \ 0]$$

olarak alınmıştır. Sıfır hipotezi (H_0) altında testin gücü için çok değişkenli normal dağılımda farklı ortalama vektörleri için örnekler üretiyoruz. Hem aykırı değerlerin eklenmediği (temiz) hem de aykırı değerlerin eklendiği(kirli) veri için ortalama vektörleri:

$$\mu_1 = [0 \ 0 \ 0 \ \dots \ 0]$$

$$\mu_2 = [\Delta \ \Delta \ \Delta \ \dots \ \Delta]$$

olarak alınmıştır. Buradaki işlemlerde $\Delta=2$ ve $\Delta=5$ için kombinasyonlar yapılmıştır. Hem tip-1 hata oranları hem de güç için kovaryans matrisleri popülasyonları farklıdır ve (Gökpınar ve diğerleri, 2020)’e benzer dört farklı kovaryans yapısı aşağıdaki gibidir:

$$\text{Dizayn 1: } \Sigma_1 = I_p + 0.5(J_p - I_p) \quad ; \quad \Sigma_2 = 3I_p + 0.5(J_p - I_p)$$

$$\text{Dizayn 2: } \Sigma_1 = [\sigma_{rs}] = (0.6)^{|r-s|} \quad ; \quad \Sigma_2 = (0.6)^{|r-s|} + 2I_p$$

$$\text{Dizayn 3: } \Sigma_1 = I_p + 0.5(J_p - I_p) \quad ; \quad \Sigma_2 = 9I_p + 0.5(J_p - I_p)$$

$$\text{Dizayn 4: } \Sigma_1 = [\sigma_{rs}] = (0.6)^{|r-s|} \quad ; \quad \Sigma_2 = (0.6)^{|r-s|} + 8I_p$$

Burada $I_{p,p}$ boyutlu bir birim matris ve J_p , 1'lerden oluşan $p \times p$ matrisidir.

Farklı örneklem boyutları ve değişken sayıları kombinasyonları ile her bir durum için rastgele 1000 kez örneklem üretildi. Nominal anlamlılık düzeyini $\alpha = 0.05$ alındı. Ayrıca nominal boyutun (0.05) korunmasındaki genel performansı ölçmek için, aşağıdaki gibi, test istatistiklerinin ortalama bağıl hata (ARE) değerlerini kullanarak gerekli hesaplamalar yapıldı.

$$ARE = \frac{100}{M\alpha} \sum_{j=1}^M |\hat{\alpha}_i - \alpha| \quad (4.1)$$

Burada M kombinasyonların sayısı, $\hat{\alpha}_i$ ise test istatistiğinin i . Kombinasyona ait tip-1 hata değeridir. En küçük ARE değerine sahip test istatistiğinin nominal boyutun korunması açısından diğer istatistiklerden daha iyi performans gösterdiği, söylenebilir.

4.1. Temiz Veri İçin Tip-1 Hata Oranları

Bu alt bölümde daha önce de belirttiğimiz gibi tip-1 hata oranı için ortalama vektörleri eşit olarak alınmış ve bu grupta veriye aykırı bir değer eklenmemiştir (kirlenmemiştir). Diğer bir deyişle yokluk hipotezinin doğru olduğu durumda test işlemleri gerçekleştirilmelidir. Dolayısıyla söz konusu hipotezi nominal boyut olan %5'e yakın bir oranda reddedilmesi beklenmektedir. Farklı kombinasyonlar için elde edilen deneysel tip-1 hata oranları tabloda verilmiştir.

Tablo 4.1. Temiz veride Dizayn 1 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T_B	T_Y	T_J	T_{NV}	T_{M-NV}	CAT	RobCAT
2	7	7	0,055	0,054	0,245	0,049	0,051	0,046	0,112
	7	15	0,041	0,045	0,112	0,044	0,044	0,043	0,096
	10	20	0,048	0,042	0,085	0,042	0,042	0,038	0,085
	15	15	0,047	0,044	0,093	0,042	0,042	0,040	0,094
	15	30	0,061	0,049	0,071	0,049	0,049	0,049	0,079
	20	20	0,055	0,046	0,087	0,048	0,047	0,045	0,081
	20	40	0,057	0,062	0,078	0,062	0,062	0,059	0,094
	30	50	0,051	0,053	0,063	0,053	0,053	0,052	0,084

Tablo 4.1. (Devam)

3	7	7	0,050	0,064	0,336	0,050	0,062	0,043	0,103
	7	15	0,047	0,055	0,153	0,047	0,056	0,044	0,090
	10	20	0,044	0,042	0,105	0,041	0,043	0,033	0,073
	15	15	0,041	0,041	0,116	0,040	0,041	0,042	0,095
	15	30	0,040	0,044	0,077	0,042	0,044	0,038	0,084
	20	20	0,051	0,058	0,122	0,052	0,052	0,046	0,096
	20	40	0,050	0,054	0,080	0,053	0,054	0,049	0,071
	30	50	0,055	0,064	0,081	0,064	0,064	0,057	0,084
4	7	7	0,049	0,058	0,472	0,040	0,053	0,025	0,060
	7	15	0,065	0,045	0,185	0,038	0,043	0,038	0,081
	10	20	0,057	0,055	0,154	0,052	0,054	0,042	0,079
	15	15	0,046	0,049	0,162	0,048	0,049	0,041	0,086
	15	30	0,056	0,047	0,105	0,048	0,049	0,049	0,082
	20	20	0,046	0,041	0,114	0,042	0,041	0,037	0,068
	20	40	0,046	0,046	0,08	0,045	0,047	0,038	0,057
	30	50	0,052	0,056	0,081	0,056	0,056	0,053	0,073
ARE Değerleri			0.24	0.304	4.114	0.274	0.272	0.39	1.614

Tablo 4.1'e göre Dizayn 1 için, minimum ARE değerine sahip olan Bennett testi tip-1 hata oranları bakımından temiz veri setlerinde daha önce önerilen testlerden ve önerdiğimiz RobCAT yaklaşımından daha başarılıdır.

Tablo 4.2. Temiz veride Dizayn 2 için test istatistiklerinin deneysel Tip-1 hata oranları

p	n ₁	n ₂	T_B	T_Y	T_J	T_{NV}	T_{M-NV}	CAT	RobCAT
2	7	7	0,042	0,040	0,239	0,043	0,043	0,035	0,095
	7	15	0,056	0,052	0,123	0,051	0,053	0,050	0,112
	10	20	0,04	0,05	0,095	0,051	0,051	0,049	0,094
	15	15	0,060	0,053	0,117	0,051	0,051	0,048	0,102
	15	30	0,058	0,067	0,098	0,066	0,067	0,067	0,100
	20	20	0,054	0,056	0,105	0,057	0,057	0,051	0,094
	20	40	0,067	0,065	0,087	0,065	0,065	0,063	0,107
	30	50	0,043	0,052	0,065	0,052	0,052	0,048	0,084
3	7	7	0,037	0,051	0,339	0,041	0,047	0,034	0,099
	7	15	0,052	0,040	0,131	0,040	0,040	0,032	0,084
	10	20	0,067	0,061	0,116	0,060	0,063	0,056	0,093
	15	15	0,062	0,063	0,143	0,055	0,054	0,048	0,104
	15	30	0,040	0,046	0,080	0,047	0,046	0,041	0,079
	20	20	0,053	0,056	0,112	0,055	0,053	0,044	0,090
	20	40	0,046	0,053	0,075	0,054	0,054	0,054	0,080
	30	50	0,058	0,047	0,068	0,048	0,047	0,046	0,075

Tablo 4.2. (Devam)

	7	7	0,051	0,074	0,499	0,048	0,057	0,034	0,061
	7	15	0,048	0,049	0,206	0,047	0,052	0,038	0,065
	10	20	0,052	0,046	0,129	0,044	0,046	0,038	0,074
	15	15	0,046	0,050	0,171	0,048	0,049	0,039	0,087
4	15	30	0,059	0,053	0,105	0,052	0,053	0,044	0,078
	20	20	0,048	0,047	0,141	0,049	0,047	0,042	0,076
	20	40	0,047	0,04	0,071	0,04	0,04	0,036	0,071
	30	50	0,054	0,059	0,089	0,058	0,058	0,054	0,089
ARE									
Değerleri			0.332	0.32	4.408	0.284	0.27	0.398	1.786

Tablo 4.2.'ye göre Dizayn 2 için, minimum ARE değerine sahip olan Modifiye edilmiş Nel ve Van der Merwe testi tip-1 hata oranları bakımından temiz veri setlerinde daha önce önerilen testlerden ve önerdiğimiz RobCAT yaklaşımından daha başarılıdır.

Tablo 4.3. Temiz veride Dizayn 3 için test istatistiklerinin deneysel Tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,052	0,074	0,346	0,062	0,071	0,049	0,13
	7	15	0,053	0,048	0,118	0,047	0,047	0,045	0,10
	10	20	0,039	0,043	0,088	0,043	0,044	0,044	0,08
	15	15	0,051	0,045	0,108	0,044	0,043	0,038	0,10
	15	30	0,052	0,045	0,068	0,043	0,043	0,038	0,07
	20	20	0,058	0,058	0,111	0,057	0,057	0,05	0,09
	20	40	0,05	0,043	0,066	0,043	0,042	0,045	0,07
	30	50	0,043	0,044	0,063	0,043	0,043	0,044	0,08
3	7	7	0,049	0,070	0,465	0,033	0,045	0,039	0,13
	7	15	0,044	0,040	0,130	0,037	0,037	0,033	0,07
	10	20	0,057	0,047	0,110	0,045	0,046	0,046	0,08
	15	15	0,056	0,055	0,155	0,049	0,050	0,043	0,09
	15	30	0,050	0,043	0,084	0,045	0,044	0,036	0,07
	20	20	0,039	0,038	0,107	0,037	0,037	0,032	0,06
	20	40	0,056	0,054	0,072	0,053	0,053	0,055	0,06
	30	50	0,048	0,045	0,061	0,046	0,045	0,043	0,07
4	7	7	0,058	0,089	0,667	0,037	0,067	0,053	0,09
	7	15	0,056	0,036	0,217	0,034	0,038	0,025	0,07
	10	20	0,037	0,033	0,127	0,031	0,031	0,029	0,06
	15	15	0,053	0,057	0,208	0,044	0,047	0,039	0,09
	15	30	0,045	0,049	0,099	0,047	0,047	0,043	0,07
	20	20	0,055	0,056	0,152	0,052	0,055	0,043	0,07
	20	40	0,046	0,046	0,081	0,046	0,046	0,038	0,06
	30	50	0,054	0,058	0,085	0,058	0,057	0,049	0,06
ARE									
Değerleri			0.242	0.452	4.932	0.37	0.372	0.434	1.60

Tablo 4.3.'e göre Dizayn 3 için, minimum ARE değerine sahip olan Bennett testi tip-1 hata oranları bakımından temiz veri setlerinde daha önce önerilen testlerden ve önerdiğimiz RobCAT yaklaşımından daha başarılıdır.

Tablo 4.4. Temiz veride Dizayn 4 için test istatistiklerinin deneysel Tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,053	0,058	0,349	0,050	0,054	0,042	0,133
	7	15	0,056	0,055	0,124	0,056	0,056	0,045	0,094
	10	20	0,063	0,05	0,097	0,05	0,05	0,051	0,09
	15	15	0,050	0,052	0,124	0,052	0,052	0,045	0,094
	15	30	0,052	0,059	0,085	0,058	0,058	0,054	0,086
	20	20	0,046	0,046	0,1	0,045	0,046	0,045	0,09
	20	40	0,057	0,054	0,074	0,054	0,054	0,051	0,077
	30	50	0,045	0,053	0,067	0,054	0,054	0,052	0,071
3	7	7	0,053	0,063	0,441	0,030	0,044	0,032	0,098
	7	15	0,05	0,045	0,130	0,045	0,046	0,039	0,089
	10	20	0,059	0,061	0,131	0,060	0,060	0,050	0,106
	15	15	0,047	0,049	0,181	0,045	0,048	0,034	0,110
	15	30	0,046	0,047	0,079	0,046	0,046	0,043	0,069
	20	20	0,042	0,043	0,119	0,041	0,043	0,039	0,09
	20	40	0,056	0,056	0,078	0,056	0,055	0,052	0,082
	30	50	0,038	0,045	0,075	0,046	0,045	0,043	0,075
4	7	7	0,043	0,091	0,673	0,028	0,065	0,046	0,094
	7	15	0,049	0,054	0,187	0,051	0,05	0,038	0,074
	10	20	0,048	0,048	0,15	0,048	0,047	0,039	0,075
	15	15	0,054	0,056	0,217	0,047	0,049	0,042	0,099
	15	30	0,055	0,061	0,122	0,057	0,057	0,049	0,075
	20	20	0,047	0,056	0,149	0,054	0,056	0,047	0,085
	20	40	0,061	0,06	0,099	0,061	0,061	0,058	0,075
	30	50	0,05	0,046	0,076	0,047	0,047	0,041	0,065
ARE									
Değerleri			0.236	0.44	5.454	0.29	0.242	0.318	1.772

Tablo 4.4.'e göre Dizayn 4 için, minimum ARE değerine sahip olan Bennett testi tip-1 hata oranları bakımından temiz veri setlerinde daha önce önerilen testlerden ve önerdiğimiz RobCAT yaklaşımından daha başarılıdır.

4.2. Temiz Veri İçin Testin Gücü

Bu alt bölümde veri setlerine aykırı değer eklenmemiş ve daha önce belirtildiği gibi ortalama vektörleri farklı olarak oluşturulmuştur ($\Delta=2$, $\Delta=5$ alınarak). Bu yüzden yokluk hipotezi gerçekte yanlıştır. Dolayısıyla yapılacak hipotez testlerinde yokluk hipotezinin reddedilmesi beklenmektedir. Bu nedenle yokluk hipotezinin reddedilme oranları farklı kombinasyonlar için deneysel güç olarak aşağıdaki tablolarda verilmiştir.

Tablo 4.5. Temiz veride Dizayn 1 için $\Delta=2$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,564	0,695	0,946	0,682	0,691	0,629	0,624
	7	15	0,786	0,939	0,981	0,938	0,942	0,934	0,896
	10	20	0,96	0,992	0,998	0,992	0,993	0,989	0,984
	15	15	0,984	0,992	0,998	0,992	0,992	0,992	0,989
	15	30	0,999	1	1	1	1	1	1
	20	20	0,997	0,998	0,999	0,998	0,998	0,999	0,999
	20	40	0,999	1	1	1	1	1	0,999
	30	50	1	1	1	1	1	1	1
3	7	7	0,489	0,700	0,975	0,655	0,673	0,579	0,546
	7	15	0,684	0,924	0,992	0,936	0,945	0,900	0,851
	10	20	0,943	0,998	0,999	0,998	0,998	0,992	0,988
	15	15	0,982	0,994	1	0,994	0,994	0,994	0,989
	15	30	1	1	1	1	1	1	1
	20	20	0,999	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
4	7	7	0,338	0,660	0,985	0,574	0,619	0,430	0,325
	7	15	0,496	0,895	0,990	0,914	0,933	0,845	0,782
	10	20	0,905	0,988	0,999	0,988	0,991	0,985	0,971
	15	15	0,975	0,992	1	0,990	0,990	0,989	0,984
	15	30	0,998	1	1	1	1	1	1
	20	20	0,998	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tablo 4.5. incelendiğinde Dizayn 1 için $\Delta=2$ iken, daha önceden önerilen testlerin gücü çok az da olsa önerilen ROBCAT algoritmasından daha yüksektir. Veri temiz olmasına karşın bazı durumlarda aynı güce sahip olduğu bile görülmektedir. Dolayısıyla tabloya göre önerilen algoritmanın güç bakımından diğer testlere yakın bir performansa sahip olduğu söylenebilir. Ancak tüm test istatistikleri için küçük örneklem değerlerinde deneysel güç değerinin 1'den uzaklaştığını görmekteyiz.

Tablo 4.6. Temiz veride Dizayn 2 için $\Delta=2$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,593	0,708	0,947	0,698	0,693	0,628	0,643
	7	15	0,762	0,923	0,975	0,926	0,932	0,909	0,879
	10	20	0,952	0,989	0,992	0,99	0,991	0,986	0,985
	15	15	0,978	0,985	0,993	0,985	0,985	0,984	0,983
	15	30	1	1	1	1	1	1	1
	20	20	0,999	1	1	1	1	1	0,997
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tablo 4.6. (Devam)

3	7	7	0,462	0,676	0,974	0,626	0,643	0,520	0,533
	7	15	0,640	0,919	0,983	0,929	0,938	0,890	0,856
	10	20	0,946	0,991	0,999	0,991	0,993	0,986	0,981
	15	15	0,982	0,993	1	0,993	0,992	0,990	0,986
	15	30	0,998	1	1	1	1	1	0,999
	20	20	0,998	1	1	1	1	0,998	0,998
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
4	7	7	0,375	0,670	0,990	0,590	0,632	0,467	0,355
	7	15	0,497	0,913	0,996	0,923	0,942	0,872	0,808
	10	20	0,917	0,994	0,999	0,994	0,994	0,989	0,981
	15	15	0,988	0,996	1	0,996	0,996	0,995	0,995
	15	30	1	1	1	1	1	1	1
	20	20	0,998	0,999	0,999	0,999	0,999	0,999	0,999
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tablo 4.6. incelendiğinde Dizayn 2 için $\Delta=2$ iken, daha önceden önerilen testlerin gücü çok az da olsa önerilen ROBCAT algoritmasından daha yüksektir. Veri temiz olmasına karşın bazı durumlarda aynı güce sahip olduğu bile görülmektedir. Dolayısıyla tabloya göre önerilen algoritmanın güç bakımından diğer testlere yakın bir performansa sahip olduğu söylenebilir. Ancak tüm test istatistikleri için küçük örneklem değerlerinde deneysel güç değerinin 1'den uzaklaştığını görmekteyiz.

Tablo 4.7. Temiz veride Dizayn 3 için $\Delta=2$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T_B	T_Y	T_J	T_{NV}	T_{M-NV}	CAT	RobCAT
2	7	7	0,309	0,373	0,834	0,337	0,349	0,289	0,366
	7	15	0,506	0,722	0,843	0,719	0,720	0,702	0,724
	10	20	0,762	0,876	0,932	0,872	0,873	0,858	0,871
	15	15	0,774	0,808	0,909	0,797	0,798	0,780	0,809
	15	30	0,941	0,968	0,979	0,968	0,967	0,966	0,963
	20	20	0,89	0,903	0,954	0,899	0,9	0,894	0,901
	20	40	0,991	0,996	0,998	0,996	0,996	0,995	0,995
	30	50	1	1	1	1	1	1	1
3	7	7	0,266	0,418	0,922	0,293	0,350	0,293	0,382
	7	15	0,440	0,764	0,918	0,764	0,760	0,716	0,712
	10	20	0,768	0,901	0,960	0,901	0,899	0,886	0,884
	15	15	0,798	0,843	0,955	0,830	0,832	0,797	0,817
	15	30	0,959	0,992	0,995	0,992	0,992	0,989	0,986
	20	20	0,919	0,928	0,980	0,924	0,923	0,910	0,935
	20	40	0,997	0,999	0,999	0,999	0,999	0,999	0,999
	30	50	1	1	1	1	1	1	1

Tablo 4.7. (Devam)

3	7	7	0,202	0,411	0,969	0,189	0,298	0,249	0,248
	7	15	0,340	0,764	0,954	0,768	0,768	0,703	0,673
	10	20	0,704	0,939	0,981	0,931	0,93	0,917	0,898
	15	15	0,801	0,870	0,975	0,839	0,842	0,801	0,833
	15	30	0,959	0,992	0,999	0,991	0,991	0,993	0,992
	20	20	0,942	0,959	0,992	0,954	0,954	0,945	0,956
	20	40	0,996	0,998	1	0,998	0,998	0,999	0,999
	30	50	1	1	1	1	1	1	1

Tablo 4.7. incelendiğinde Dizayn 3 için $\Delta = 2$ iken, daha önceden önerilen testlerin gücü genel olarak 1'e yakındır. Önerilen ROBCAT algoritmasında da aynı durum geçerlidir. Veri temiz olmasına karşın bazı durumlarda aynı güce sahip olduğu bile görülmektedir. Dolayısıyla tabloya göre önerilen algoritmanın güç bakımından diğer testlere yakın bir performansa sahip olduğu söylenebilir. Ancak tüm test istatistikleri için küçük örneklem değerlerinde deneysel güç değerinin 1'den uzaklaştığını görmekteyiz.

Tablo 4.8. Temiz veride Dizayn 4 için $\Delta = 2$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,290	0,361	0,817	0,333	0,341	0,271	0,365
	7	15	0,508	0,687	0,836	0,683	0,680	0,675	0,722
	10	20	0,753	0,858	0,914	0,854	0,853	0,843	0,859
	15	15	0,710	0,743	0,877	0,735	0,736	0,718	0,760
	15	30	0,944	0,977	0,984	0,976	0,976	0,970	0,973
	20	20	0,871	0,891	0,938	0,889	0,889	0,876	0,878
	20	40	0,987	0,99	0,993	0,99	0,99	0,989	0,989
	30	50	0,999	1	1	1	1	0,999	0,999
3	7	7	0,261	0,388	0,912	0,292	0,332	0,259	0,348
	7	15	0,424	0,739	0,908	0,737	0,736	0,706	0,692
	10	20	0,740	0,906	0,952	0,905	0,900	0,889	0,880
	15	15	0,803	0,847	0,954	0,830	0,831	0,807	0,811
	15	30	0,961	0,990	0,996	0,988	0,988	0,986	0,988
	20	20	0,935	0,945	0,977	0,940	0,941	0,937	0,943
	20	40	0,995	0,998	0,999	0,998	0,998	0,998	0,997
	30	50	1	1	1	1	1	1	1
4	7	7	0,186	0,436	0,976	0,197	0,305	0,239	0,246
	7	15	0,338	0,764	0,942	0,769	0,771	0,698	0,666
	10	20	0,712	0,92	0,976	0,912	0,914	0,899	0,894
	15	15	0,808	0,896	0,978	0,857	0,864	0,818	0,822
	15	30	0,960	0,998	0,999	0,997	0,997	0,995	0,996
	20	20	0,951	0,968	0,991	0,96	0,96	0,954	0,96
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tablo 4.8. incelendiğinde Dizayn 4 için $\Delta = 2$ iken, daha önceden önerilen testlerin gücü çok az da olsa önerilen ROBCAT algoritmasından daha yüksektir. Veri

temiz olmasına karşın bazı durumlarda aynı güce sahip olduğu bile görülmektedir. Bu kovaryans matrisi dizaynında testin deneysel güç değerlerinin 1'den oldukça uzaklaştığını görmekteyiz.

Tablo 4.9. Temiz veride Dizayn 1 için $\Delta=5$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	1	1	1	1	1	0,999	0,978
	7	15	1	1	1	1	1	1	1
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
3	7	7	0,994	1	1	1	1	0,981	0,946
	7	15	1	1	1	1	1	1	0,999
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
4	7	7	0,946	0,999	1	0,999	1	0,952	0,881
	7	15	0,992	1	1	1	1	1	0,994
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tablo 4.9. incelendiğinde Dizayn 1 için $\Delta=5$ iken, daha önceden önerilen testlerin gücü çok az da olsa önerilen ROBCAT algoritmasından daha yüksektir. Veri temiz olmasına karşın genel olarak tüm durumlarda aynı güce sahip olduğu ve 1 olduğu bile görülmektedir.

Tablo 4.10. Temiz veride Dizayn 2 için $\Delta=5$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,998	1	1	1	1	0,999	0,977
	7	15	1	1	1	1	1	1	0,999
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tablo 4.10. (Devam)

3	7	7	0,996	1	1	1	1	0,983	0,941
	7	15	1	1	1	1	1	1	0,999
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
4	7	7	0,964	1	1	1	1	0,973	0,886
	7	15	0,996	1	1	1	1	1	0,986
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tabloya göre, veri temiz olmasına karşın genel olarak tüm durumlarda RobCAT algoritmasının diğer testlerle aynı güce sahip olduğu ve 1 olduğu görülmektedir.

Tablo 4.11. Temiz veride Dizayn 3 için $\Delta=5$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,950	0,971	1	0,968	0,968	0,890	0,831
	7	15	0,999	1	1	1	1	1	1
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
3	7	7	0,914	0,979	1	0,963	0,973	0,806	0,787
	7	15	0,991	1	1	1	1	1	1
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
4	7	7	0,798	0,968	1	0,936	0,968	0,778	0,767
	7	15	0,946	1	1	1	1	1	0,991
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tablo 4.11. incelendiğinde Dizayn 3 için $\Delta=5$ iken, daha önceden önerilen testlerin gücü çok az da olsa önerilen ROBCAT algoritmasından daha yüksektir. Veri temiz olmasına karşın genel olarak tüm durumlarda aynı güce sahip olduğu ve 1 olduğu bile görülmektedir. Dolayısıyla tabloya göre önerilen algoritmanın güç bakımından diğer testlere yakın bir performansa sahip olduğu söylenebilir.

Tablo 4.12. Temiz veride Dizayn 4 için $\Delta=5$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,954	0,972	1	0,970	0,971	0,899	0,843
	7	15	0,999	1	1	1	1	1	1
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
3	7	7	0,913	0,966	1	0,959	0,965	0,794	0,772
	7	15	0,994	1	1	1	1	1	0,999
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1
4	7	7	0,804	0,975	1	0,934	0,965	0,752	0,761
	7	15	0,959	1	1	1	1	1	0,994
	10	20	1	1	1	1	1	1	1
	15	15	1	1	1	1	1	1	1
	15	30	1	1	1	1	1	1	1
	20	20	1	1	1	1	1	1	1
	20	40	1	1	1	1	1	1	1
	30	50	1	1	1	1	1	1	1

Tablo 4.12 incelendiğinde Dizayn 4 için $\Delta=5$ iken, daha önceden önerilen testlerin gücü çok az da olsa önerilen ROBCAT algoritmasından daha yüksektir. Veri temiz olmasına karşın genel olarak tüm durumlarda aynı güce sahip olduğu ve 1 olduğu bile görülmektedir. Dolayısıyla tabloya göre önerilen algoritmanın güç bakımından diğer testlere yakın bir performansa sahip olduğu söylenebilir.

4.3. Kirli Veri İçin Tip-1 Hata Oranları

Bu alt bölümde Bölüm 4.1.'den farklı olarak örneklere aykırı değerler eklenmiştir. Bu aykırı değerleri eklemek için 1. Gruptaki son gözlem satırı 10 ile, 2. Grupta bulunan son gözlem satırı ise -10 ile çarpılmıştır. Buradaki kirlilik oranı alpha argümanı ile belirtilmiştir. Kirlilik oranı %10 ve %30 alınarak, sırasıyla gerekli kombinasyonlar yapılmıştır. Veriler kirlenmeden önce aynı ortalama vektörüne

sahip iki örnek olduğundan yokluk hipotezi doğrudur. Bu nedenle söz konusu hipotezin %5'e yakın bir oranda reddedilmesi beklenmektedir. Farklı kombinasyonlar için elde edilen deneysel tip-1 hata oranları aşağıdaki tablolarda verilmiştir.

Tablo 4.13. Kirli veride (kirlilik oranı %10 iken) Dizayn 1 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,023	0,008	0,118	0,005	0,007	0,007	0,036
	7	15	0,252	0,015	0,069	0,011	0,013	0,013	0,058
	10	20	0,27	0,012	0,036	0,014	0,014	0,012	0,05
	15	15	0,030	0,023	0,055	0,020	0,022	0,025	0,052
	15	30	0,242	0,033	0,049	0,028	0,032	0,036	0,063
	20	20	0,044	0,029	0,053	0,023	0,025	0,026	0,07
	20	40	0,175	0,03	0,043	0,028	0,029	0,029	0,06
	30	50	0,195	0,028	0,036	0,026	0,027	0,028	0,054
3	7	7	0,035	0,012	0,238	0,007	0,013	0,005	0,032
	7	15	0,294	0,019	0,086	0,01	0,017	0,017	0,036
	10	20	0,312	0,024	0,056	0,018	0,022	0,021	0,044
	15	15	0,032	0,024	0,09	0,018	0,022	0,021	0,043
	15	30	0,297	0,037	0,063	0,031	0,033	0,039	0,062
	20	20	0,027	0,033	0,067	0,029	0,033	0,032	0,037
	20	40	0,237	0,022	0,038	0,015	0,019	0,021	0,057
	30	50	0,365	0,040	0,053	0,036	0,039	0,038	0,063
4	7	7	0,036	0,041	0,397	0,016	0,034	0,017	0,033
	7	15	0,290	0,026	0,142	0,011	0,021	0,023	0,033
	10	20	0,364	0,026	0,093	0,013	0,022	0,019	0,043
	15	15	0,046	0,039	0,126	0,032	0,041	0,031	0,056
	15	30	0,352	0,032	0,079	0,024	0,031	0,031	0,049
	20	20	0,051	0,035	0,088	0,029	0,035	0,027	0,052
	20	40	0,319	0,035	0,061	0,03	0,033	0,033	0,057
	30	50	0,294	0,031	0,055	0,028	0,029	0,033	0,056
ARE									
Değerleri			7.282	1.092	2.174	1.396	1.118	1.232	0.448

Tablo 4.13. incelendiğinde Dizayn 1 için kirlilik oranı %10 iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği görülmektedir. Öte yandan RobCAT algoritması veri setindeki aykırı değerden etkilenmeden söz konusu yokluk hipotezini genel olarak nominal değer olan %5'e yakın oranlarda reddetmiştir. Bu durum önerilen RobCAT algoritmasının verideki aykırı değerlerden etkilenmediğini göstermektedir. Tabloya göre minimum ARE değerine sahip olan RobCAT algoritması tip-1 hata bakımından kirli veri setlerinde diğer algoritmalarından daha başarılıdır.

Tablo 4.14. Kirli veride (kirlilik oranı %10 iken) Dizayn 2 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,025	0,016	0,144	0,012	0,016	0,012	0,054
	7	15	0,270	0,018	0,066	0,012	0,016	0,017	0,051
	10	20	0,237	0,027	0,056	0,025	0,027	0,025	0,074
	15	15	0,020	0,019	0,057	0,018	0,018	0,02	0,04
	15	30	0,214	0,018	0,04	0,017	0,017	0,016	0,054
	20	20	0,025	0,022	0,037	0,02	0,021	0,019	0,046
	20	40	0,207	0,024	0,036	0,023	0,024	0,021	0,057
	30	50	0,206	0,03	0,039	0,03	0,03	0,031	0,066
3	7	7	0,039	0,024	0,220	0,019	0,023	0,019	0,034
	7	15	0,294	0,023	0,093	0,018	0,02	0,018	0,039
	10	20	0,320	0,027	0,067	0,016	0,024	0,02	0,041
	15	15	0,045	0,033	0,089	0,028	0,03	0,026	0,054
	15	30	0,309	0,028	0,049	0,024	0,027	0,025	0,049
	20	20	0,041	0,034	0,064	0,025	0,03	0,028	0,051
	20	40	0,256	0,03	0,051	0,026	0,031	0,029	0,049
	30	50	0,236	0,037	0,049	0,034	0,037	0,033	0,064
4	7	7	0,044	0,032	0,374	0,017	0,032	0,015	0,025
	7	15	0,266	0,022	0,139	0,013	0,019	0,011	0,029
	10	20	0,35	0,035	0,109	0,018	0,033	0,031	0,045
	15	15	0,039	0,03	0,110	0,025	0,026	0,023	0,048
	15	30	0,352	0,029	0,073	0,021	0,028	0,028	0,046
	20	20	0,035	0,026	0,075	0,023	0,028	0,022	0,046
	20	40	0,303	0,03	0,065	0,023	0,029	0,029	0,049
	30	50	0,328	0,043	0,065	0,04	0,043	0,039	0,061
ARE									
Değerleri			7.07	1.066	2.134	1.342	1.146	1.286	0.4

Tablo 4.14. incelendiğinde, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği RobCAT algoritmasının ise aykırı değerlerden etkilenmediği ve 0.05 anlamlılık düzeyine yakın oranlarda reddettiği görülmektedir. Tabloya göre minimum ARE değerine sahip olan RobCAT algoritması tip-1 hata bakımından kirli veri setlerinde diğer algoritmalarından daha başarılıdır.

Tablo 4.15. Kirli veride (kirlilik oranı %10 iken) Dizayn 3 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,038	0,021	0,206	0,016	0,019	0,02	0,049
	7	15	0,316	0,013	0,066	0,012	0,014	0,012	0,047
	10	20	0,365	0,026	0,065	0,024	0,027	0,027	0,055
	15	15	0,035	0,033	0,08	0,03	0,032	0,029	0,044
	15	30	0,303	0,018	0,044	0,019	0,019	0,022	0,06
	20	20	0,025	0,019	0,047	0,019	0,019	0,017	0,05
	20	40	0,265	0,024	0,038	0,025	0,025	0,021	0,048
	30	50	0,266	0,037	0,049	0,037	0,037	0,04	0,056

Tablo 4.15. (Devam)

3	7	7	0,035	0,028	0,283	0,012	0,024	0,015	0,046
	7	15	0,386	0,022	0,089	0,015	0,019	0,015	0,035
	10	20	0,420	0,025	0,065	0,022	0,024	0,021	0,052
	15	15	0,032	0,029	0,098	0,023	0,028	0,025	0,036
	15	30	0,383	0,032	0,059	0,03	0,032	0,031	0,057
	20	20	0,042	0,032	0,087	0,027	0,03	0,028	0,056
	20	40	0,350	0,042	0,064	0,040	0,040	0,040	0,074
	30	50	0,342	0,034	0,048	0,032	0,033	0,032	0,045
4	7	7	0,043	0,029	0,294	0,009	0,026	0,016	0,057
	7	15	0,382	0,024	0,106	0,015	0,02	0,017	0,041
	10	20	0,465	0,035	0,131	0,031	0,034	0,028	0,045
	15	15	0,039	0,028	0,141	0,014	0,023	0,016	0,038
	15	30	0,440	0,024	0,058	0,023	0,025	0,022	0,035
	20	20	0,037	0,03	0,112	0,026	0,034	0,024	0,037
	20	40	0,362	0,033	0,061	0,03	0,034	0,032	0,06
	30	50	0,351	0,031	0,05	0,028	0,03	0,029	0,047
ARE									
Değerleri			9.536	1.06	2.256	1.282	1.104	1.262	0.402

Tablo 4.15. incelendiğinde Dizayn 3 için kirlilik oranı %10 iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği RobCAT algoritmasının ise aykırı değerlerden etkilenmediği ve 0.05 anlamlılık düzeyine yakın oranlarda reddettiği görülmektedir. Tabloya göre minimum ARE değerine sahip olan RobCAT algoritması tip-1 hata bakımından kirli veri setlerinde diğer algoritmalarından daha başarılıdır.

Tablo 4.16. Kirli veride (kirlilik oranı %10 iken) Dizayn 4 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,031	0,02	0,202	0,014	0,017	0,016	0,053
	7	15	0,328	0,021	0,056	0,018	0,02	0,02	0,055
	10	20	0,359	0,024	0,05	0,021	0,022	0,024	0,051
	15	15	0,042	0,036	0,085	0,027	0,033	0,03	0,047
	15	30	0,313	0,021	0,042	0,021	0,021	0,024	0,052
	20	20	0,037	0,036	0,067	0,033	0,033	0,029	0,05
	20	40	0,257	0,033	0,046	0,033	0,033	0,032	0,052
	30	50	0,286	0,036	0,048	0,035	0,037	0,032	0,059
3	7	7	0,040	0,027	0,295	0,010	0,02	0,018	0,04
	7	15	0,370	0,028	0,095	0,027	0,028	0,02	0,043
	10	20	0,445	0,033	0,085	0,023	0,03	0,024	0,047
	15	15	0,038	0,038	0,103	0,029	0,037	0,036	0,044
	15	30	0,370	0,025	0,054	0,024	0,026	0,026	0,051
	20	20	0,044	0,039	0,085	0,032	0,039	0,032	0,045
	20	40	0,347	0,045	0,064	0,042	0,045	0,041	0,06
	30	50	0,307	0,033	0,054	0,031	0,033	0,032	0,07

Tablo 4.16. (Devam)

4	7	7	0,033	0,037	0,501	0,011	0,031	0,01	0,041
	7	15	0,349	0,036	0,141	0,020	0,031	0,021	0,03
	10	20	0,413	0,027	0,086	0,018	0,024	0,021	0,039
	15	15	0,042	0,032	0,149	0,023	0,035	0,028	0,047
	15	30	0,419	0,033	0,072	0,029	0,033	0,028	0,047
	20	20	0,043	0,031	0,095	0,029	0,035	0,028	0,029
	20	40	0,365	0,032	0,056	0,026	0,031	0,026	0,059
	30	50	0,389	0,028	0,053	0,028	0,028	0,029	0,043
ARE									
Değerleri			9.336	0.898	2.83	1.212	0.958	1.146	0.34

Tablo 4.16. incelendiğinde Dizayn 4 için kirlilik oranı %10 iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği RobCAT algoritmasının ise aykırı değerlerden etkilenmediği ve 0.05 anlamlılık düzeyine yakın oranlarda reddettiği görülmektedir. Tabloya göre minimum ARE değerine sahip olan RobCAT algoritması tip-1 hata bakımından kirli veri setlerinde diğer algoritmalarından daha başarılıdır.

Tablo 4.17. Kirli veride (kirlilik oranı %30 iken) Dizayn 1 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,016	0,010	0,125	0,009	0,009	0,008	0,041
	7	15	0,251	0,014	0,060	0,010	0,012	0,011	0,048
	10	20	0,256	0,023	0,044	0,018	0,020	0,019	0,048
	15	15	0,022	0,022	0,047	0,021	0,021	0,021	0,050
	15	30	0,230	0,013	0,040	0,012	0,014	0,014	0,047
	20	20	0,027	0,020	0,038	0,017	0,020	0,017	0,054
	20	40	0,192	0,033	0,048	0,032	0,033	0,038	0,073
	30	50	0,196	0,028	0,041	0,027	0,028	0,028	0,070
3	7	7	0,040	0,024	0,232	0,009	0,020	0,013	0,028
	7	15	0,298	0,015	0,081	0,009	0,015	0,014	0,043
	10	20	0,331	0,022	0,067	0,015	0,021	0,023	0,047
	15	15	0,044	0,025	0,070	0,020	0,026	0,025	0,055
	15	30	0,289	0,026	0,045	0,022	0,026	0,026	0,043
	20	20	0,038	0,027	0,058	0,025	0,026	0,026	0,044
	20	40	0,268	0,026	0,045	0,026	0,026	0,025	0,068
	30	50	0,247	0,040	0,049	0,038	0,040	0,035	0,069
4	7	7	0,052	0,031	0,368	0,013	0,029	0,013	0,027
	7	15	0,282	0,026	0,142	0,012	0,021	0,016	0,026
	10	20	0,375	0,021	0,092	0,016	0,019	0,017	0,026
	15	15	0,040	0,030	0,101	0,023	0,030	0,024	0,036
	15	30	0,325	0,019	0,057	0,014	0,017	0,016	0,044
	20	20	0,050	0,030	0,096	0,026	0,031	0,026	0,055
	20	40	0,311	0,028	0,056	0,022	0,028	0,026	0,054
	30	50	0,311	0,032	0,059	0,028	0,031	0,028	0,066

ARE								
Değerleri	7.074	1.064	1.934	1.466	1.352	1.382	0.532	

Tablo 4.17. incelendiğinde Dizayn 1 için kirlilik oranı %30 iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği ve nominal değerden uzaklaştığı görülmektedir. Öte yandan RobCAT algoritması veri setindeki aykırı değerden genel olarak etkilenmeden söz konusu yokluk hipotezini nominal değer olan %5'e yakın oranlarda reddetmiştir. Ancak söz konusu önerdiğimiz algoritmanın değişken değeri arttıkça ve örneklem değeri azaldıkça nominal değerden ulaştığını görmekteyiz. Tabloya göre minimum ARE değerine sahip olan RobCAT algoritması tip-1 hata bakımından kirli veri setlerinde diğer algoritmalarından daha başarılıdır.

Tablo 4.18. Kirli veride (kirlilik oranı %30 iken) Dizayn 2 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,027	0,012	0,130	0,009	0,011	0,012	0,041
	7	15	0,262	0,028	0,069	0,020	0,023	0,027	0,047
	10	20	0,263	0,018	0,042	0,013	0,015	0,017	0,051
	15	15	0,033	0,020	0,049	0,017	0,020	0,018	0,040
	15	30	0,184	0,016	0,029	0,014	0,016	0,014	0,052
	20	20	0,026	0,024	0,043	0,023	0,023	0,021	0,050
	20	40	0,221	0,032	0,041	0,029	0,030	0,033	0,073
	30	50	0,183	0,029	0,039	0,026	0,028	0,026	0,060
3	7	7	0,034	0,024	0,230	0,013	0,021	0,012	0,033
	7	15	0,278	0,027	0,106	0,016	0,023	0,022	0,048
	10	20	0,329	0,019	0,064	0,013	0,015	0,018	0,039
	15	15	0,031	0,030	0,080	0,025	0,032	0,026	0,045
	15	30	0,292	0,025	0,055	0,023	0,024	0,024	0,050
	20	20	0,040	0,023	0,053	0,019	0,024	0,017	0,045
	20	40	0,238	0,026	0,044	0,023	0,025	0,025	0,046
	30	50	0,237	0,029	0,047	0,029	0,029	0,033	0,060
4	7	7	0,048	0,024	0,375	0,011	0,025	0,011	0,014
	7	15	0,284	0,015	0,132	0,005	0,013	0,009	0,020
	10	20	0,368	0,030	0,099	0,015	0,027	0,022	0,047
	15	15	0,045	0,038	0,114	0,028	0,032	0,028	0,048
	15	30	0,343	0,039	0,078	0,028	0,035	0,038	0,051
	20	20	0,031	0,025	0,082	0,022	0,027	0,022	0,033
	20	40	0,286	0,029	0,060	0,027	0,029	0,031	0,071
	30	50	0,283	0,029	0,048	0,028	0,029	0,030	0,061
ARE									
Değerleri	6.872	1.118	1.946	1.448	1.248	1.328	0.466		

Tablo 4.18. incelendiğinde Dizayn 2 için kirlilik oranı %30 iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği ve nominal değerden uzaklaştığı görülmektedir. Öte yandan RobCAT algoritması veri setindeki aykırı değerden genel olarak etkilenmeden söz konusu yokluk hipotezini nominal değer olan %5'e yakın oranlarda reddetmiştir. Ancak söz konusu önerdiğimiz

algoritmanın örneklem değeri azaldıkça nominal değerden ulaştığını görmekteyiz. Tabloya göre minimum ARE değerine sahip olan RobCAT algoritması tip-1 hata bakımından kirli veri setlerinde diğer algoritmalarından daha başarılıdır.

Tablo 4.19. Kirli veride (kirlilik oranı %30 iken) Dizayn 3 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,030	0,021	0,199	0,012	0,016	0,017	0,050
	7	15	0,325	0,019	0,067	0,018	0,017	0,016	0,043
	10	20	0,309	0,013	0,036	0,015	0,013	0,015	0,049
	15	15	0,030	0,024	0,079	0,019	0,023	0,023	0,051
	15	30	0,286	0,029	0,055	0,031	0,031	0,027	0,056
	20	20	0,029	0,026	0,050	0,025	0,026	0,025	0,047
	20	40	0,260	0,034	0,047	0,032	0,033	0,031	0,067
	30	50	0,247	0,026	0,040	0,026	0,026	0,027	0,063
3	7	7	0,038	0,033	0,294	0,017	0,024	0,018	0,047
	7	15	0,368	0,023	0,098	0,020	0,022	0,019	0,042
	10	20	0,391	0,025	0,065	0,023	0,025	0,022	0,062
	15	15	0,041	0,033	0,095	0,025	0,032	0,030	0,044
	15	30	0,385	0,019	0,046	0,018	0,019	0,018	0,042
	20	20	0,034	0,030	0,079	0,029	0,030	0,029	0,051
	20	40	0,332	0,020	0,044	0,020	0,020	0,018	0,062
	30	50	0,309	0,037	0,055	0,034	0,036	0,034	0,071
4	7	7	0,046	0,037	0,496	0,009	0,039	0,016	0,051
	7	15	0,328	0,026	0,149	0,015	0,024	0,016	0,034
	10	20	0,438	0,030	0,098	0,024	0,029	0,021	0,045
	15	15	0,055	0,043	0,149	0,026	0,035	0,031	0,050
	15	30	0,424	0,024	0,055	0,023	0,024	0,024	0,051
	20	20	0,039	0,033	0,106	0,029	0,033	0,025	0,046
	20	40	0,377	0,037	0,071	0,037	0,041	0,036	0,063
	30	50	0,375	0,037	0,055	0,033	0,036	0,033	0,064
ARE Değerleri			9.044	1.038	2.804	1.28	1.112	1.298	0.346

Tablo 4.19. incelendiğinde Dizayn 3 için kirlilik oranı %30 iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği ve nominal değerden uzaklaştığı görülmektedir. Öte yandan RobCAT algoritması veri setindeki aykırı değerden genel olarak etkilenmeden söz konusu yokluk hipotezini nominal değer olan %5'e yakın oranlarda reddetmiştir. Tabloya göre minimum ARE değerine sahip olan RobCAT algoritması tip-1 hata bakımından kirli veri setlerinde diğer algoritmalarından daha başarılıdır.

Tablo 4.20. Kirli veride (kirlilik oranı %30 iken) Dizayn 4 için test istatistiklerinin deneysel tip-1 hata oranları

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,034	0,018	0,202	0,011	0,014	0,013	0,050
	7	15	0,348	0,025	0,074	0,024	0,024	0,025	0,046
	10	20	0,338	0,025	0,058	0,022	0,024	0,023	0,053
	15	15	0,033	0,022	0,069	0,020	0,022	0,019	0,042
	15	30	0,308	0,028	0,044	0,024	0,028	0,028	0,060
	20	20	0,029	0,025	0,054	0,022	0,023	0,023	0,044
	20	40	0,291	0,028	0,043	0,028	0,028	0,030	0,083
	30	50	0,250	0,035	0,043	0,033	0,033	0,036	0,083
3	7	7	0,036	0,016	0,294	0,008	0,015	0,008	0,049
	7	15	0,370	0,027	0,095	0,024	0,026	0,025	0,046
	10	20	0,395	0,022	0,055	0,018	0,023	0,020	0,047
	15	15	0,032	0,031	0,101	0,023	0,027	0,026	0,039
	15	30	0,387	0,022	0,049	0,023	0,025	0,022	0,043
	20	20	0,047	0,040	0,092	0,031	0,035	0,032	0,049
	20	40	0,339	0,045	0,064	0,042	0,043	0,042	0,073
	30	50	0,321	0,028	0,044	0,029	0,029	0,029	0,066
4	7	7	0,034	0,037	0,465	0,009	0,030	0,014	0,035
	7	15	0,335	0,034	0,137	0,019	0,029	0,019	0,029
	10	20	0,452	0,031	0,085	0,024	0,029	0,023	0,034
	15	15	0,048	0,031	0,152	0,021	0,030	0,025	0,035
	15	30	0,412	0,033	0,065	0,028	0,033	0,031	0,042
	20	20	0,034	0,029	0,103	0,026	0,030	0,021	0,035
	20	40	0,383	0,033	0,056	0,032	0,033	0,030	0,066
	30	50	0,357	0,029	0,054	0,027	0,028	0,030	0,054
ARE									
Değerleri			9.318	0.992	2.704	1.264	1.078	1.212	0.546

Tablo 4.20. incelendiğinde Dizayn 4 için kirlilik oranı %30 iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği ve nominal değerden uzaklaştığı görülmektedir. Öte yandan RobCAT algoritması veri setindeki aykırı değerden genel olarak etkilenmeden söz konusu yokluk hipotezini nominal değer olan %5'e yakın oranlarda reddetmiştir. Ancak söz konusu önerdiğimiz algoritmanın p değeri 4 alındığında küçük örneklem çaplarında nominal değerden ulaştığını görmekteyiz. Tabloya göre minimum ARE değerine sahip olan RobCAT algoritması tip-1 hata bakımından kirli veri setlerinde diğer algoritmalarından daha başarılıdır.

4.4. Kirli Veri İçin Testin Gücü

Bu alt bölümde Bölüm 4.2.'den farklı olarak örneklere aykırı değerler eklenmiştir. Bu aykırı değerleri eklemek için 1. Gruptaki son gözlem satırı 10 ile, 2. Grupta bulunan son gözlem satırı ise -10 ile çarpılmıştır ve daha önce belirtildiği gibi ortalama vektörleri farklı olarak oluşturulmuştur ($\Delta=2$, $\Delta=5$ alınarak). Burada veriler

kirletilmeden önce farklı ortalama vektörüne sahip iki örnek olduğundan yokluk hipotezi gerçekte yanlıştır. Bu nedenle yokluk hipotezinin reddedilme oranları farklı kombinasyonlar için deneysel güç olarak aşağıdaki tablolarda verilmiştir.

Tablo 4.21. Kirli veride Dizayn 1 için $\Delta = 2$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,189	0,093	0,300	0,066	0,081	0,075	0,495
	7	15	0,252	0,141	0,275	0,144	0,152	0,136	0,851
	10	20	0,381	0,262	0,370	0,269	0,274	0,257	0,975
	15	15	0,495	0,323	0,411	0,315	0,321	0,315	0,976
	15	30	0,616	0,442	0,504	0,449	0,452	0,437	1
	20	20	0,514	0,386	0,442	0,367	0,383	0,376	0,997
	20	40	0,791	0,592	0,625	0,588	0,592	0,581	1
	30	50	0,947	0,76	0,782	0,758	0,759	0,754	1
3	7	7	0,209	0,149	0,504	0,071	0,141	0,104	0,380
	7	15	0,310	0,264	0,471	0,233	0,257	0,247	0,788
	10	20	0,443	0,386	0,496	0,370	0,382	0,370	0,972
	15	15	0,578	0,441	0,569	0,404	0,439	0,427	0,983
	15	30	0,717	0,610	0,683	0,605	0,613	0,601	1
	20	20	0,691	0,556	0,628	0,541	0,554	0,539	0,999
	20	40	0,860	0,718	0,771	0,717	0,718	0,715	1
	30	50	0,964	0,863	0,883	0,864	0,866	0,857	1
4	7	7	0,170	0,180	0,683	0,065	0,175	0,097	0,210
	7	15	0,311	0,357	0,605	0,302	0,351	0,306	0,660
	10	20	0,504	0,498	0,671	0,476	0,503	0,475	0,954
	15	15	0,641	0,509	0,676	0,441	0,501	0,468	0,983
	15	30	0,748	0,726	0,794	0,712	0,725	0,711	0,998
	20	20	0,761	0,663	0,74	0,626	0,657	0,636	1
	20	40	0,901	0,841	0,871	0,837	0,845	0,83	1
	30	50	0,981	0,922	0,937	0,921	0,921	0,918	1

Tablo 4.21. incelendiğinde Dizayn 1 için $\Delta = 2$ iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği önerdiğimiz RobCAT algoritmasının ise genel olarak etkilenmediği ve 1'e yakın sonuçlar verdiği görülmektedir. Tablo 4.21'e göre RobCAT algoritmasının gücü diğer algoritmaların gücünden oldukça yüksektir. Bu durum önerilen RobCAT algoritmasının verideki aykırı değerlerden etkilenmediğini göstermektedir.

Tablo 4.22. Kirli veride Dizayn 2 için $\Delta=2$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,181	0,085	0,313	0,069	0,082	0,079	0,470
	7	15	0,253	0,134	0,265	0,141	0,145	0,130	0,833
	10	20	0,358	0,271	0,352	0,269	0,278	0,261	0,977
	15	15	0,434	0,275	0,363	0,263	0,269	0,267	0,967
	15	30	0,633	0,461	0,503	0,458	0,460	0,447	0,999
	20	20	0,556	0,406	0,471	0,401	0,407	0,403	0,994
	20	40	0,792	0,58	0,635	0,58	0,582	0,577	1
	30	50	0,924	0,776	0,803	0,775	0,777	0,771	1
3	7	7	0,203	0,145	0,469	0,079	0,134	0,114	0,339
	7	15	0,305	0,252	0,446	0,234	0,254	0,225	0,755
	10	20	0,438	0,403	0,511	0,399	0,409	0,390	0,971
	15	15	0,560	0,406	0,536	0,369	0,402	0,385	0,976
	15	30	0,723	0,603	0,674	0,593	0,605	0,586	0,999
	20	20	0,666	0,514	0,611	0,494	0,511	0,498	0,998
	20	40	0,872	0,723	0,761	0,723	0,727	0,718	1
	30	50	0,967	0,848	0,867	0,846	0,849	0,846	1
4	7	7	0,176	0,199	0,690	0,069	0,192	0,104	0,229
	7	15	0,288	0,357	0,608	0,301	0,345	0,305	0,702
	10	20	0,505	0,532	0,663	0,514	0,534	0,503	0,969
	15	15	0,639	0,524	0,695	0,447	0,521	0,481	0,985
	15	30	0,779	0,739	0,807	0,726	0,740	0,725	1
	20	20	0,758	0,656	0,755	0,617	0,647	0,633	0,999
	20	40	0,894	0,813	0,857	0,811	0,821	0,808	1
	30	50	0,986	0,914	0,931	0,911	0,914	0,909	1

Tablo 4.22. incelendiğinde Dizayn 2 için $\Delta=2$ iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği önerdiğimiz RobCAT algoritmasının ise genel olarak etkilenmediği ve 1'e yakın sonuçlar verdiği görülmektedir.

Tablo 4.23. Kirli veride Dizayn 3 için $\Delta=2$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,138	0,092	0,345	0,057	0,079	0,074	0,212
	7	15	0,333	0,157	0,308	0,160	0,163	0,142	0,683
	10	20	0,451	0,294	0,404	0,295	0,3	0,29	0,83
	15	15	0,367	0,285	0,407	0,267	0,277	0,272	0,725
	15	30	0,525	0,427	0,481	0,428	0,434	0,417	0,953
	20	20	0,455	0,392	0,477	0,376	0,384	0,374	0,88
	20	40	0,685	0,563	0,593	0,558	0,561	0,554	0,995
	30	50	0,844	0,697	0,723	0,693	0,699	0,696	0,997

Tablo 4.23. (Devam)

3	7	7	0,143	0,128	0,501	0,061	0,117	0,087	0,237
	7	15	0,354	0,220	0,417	0,209	0,223	0,205	0,640
	10	20	0,462	0,367	0,489	0,357	0,367	0,356	0,865
	15	15	0,464	0,372	0,537	0,309	0,355	0,330	0,770
	15	30	0,594	0,459	0,577	0,434	0,454	0,436	0,909
	20	20	0,593	0,495	0,600	0,466	0,489	0,475	0,921
	20	40	0,581	0,498	0,602	0,470	0,487	0,461	0,903
	30	50	0,893	0,808	0,843	0,804	0,807	0,804	1
4	7	7	0,132	0,134	0,692	0,031	0,105	0,062	0,145
	7	15	0,349	0,305	0,554	0,280	0,305	0,265	0,542
	10	20	0,549	0,455	0,624	0,437	0,453	0,424	0,869
	15	15	0,648	0,525	0,655	0,498	0,520	0,492	0,922
	15	30	0,677	0,541	0,682	0,490	0,524	0,493	0,931
	20	20	0,645	0,578	0,716	0,539	0,567	0,549	0,934
	20	40	0,842	0,803	0,852	0,797	0,803	0,793	1
	30	50	0,942	0,88	0,907	0,875	0,878	0,875	1

Tablo 4.23. incelendiğinde Dizayn 3 için $\Delta = 2$ iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği önerdiğimiz RobCAT algoritmasının ise genel olarak etkilenmediği ve 1'e yakın sonuçlar verdiği görülmektedir.

Tablo 4.24. Kirlili veride Dizayn 4 için $\Delta = 2$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _V	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,122	0,093	0,376	0,056	0,084	0,079	0,236
	7	15	0,333	0,162	0,295	0,162	0,166	0,144	0,653
	10	20	0,38	0,252	0,362	0,25	0,253	0,247	0,826
	15	15	0,513	0,365	0,451	0,347	0,358	0,356	0,882
	15	30	0,505	0,337	0,423	0,331	0,335	0,338	0,865
	20	20	0,481	0,398	0,477	0,38	0,384	0,379	0,894
	20	40	0,67	0,549	0,585	0,548	0,548	0,545	0,992
	30	50	0,825	0,678	0,705	0,679	0,679	0,676	0,998
3	7	7	0,150	0,111	0,548	0,054	0,102	0,073	0,217
	7	15	0,392	0,226	0,422	0,206	0,229	0,192	0,602
	10	20	0,495	0,382	0,512	0,371	0,379	0,366	0,863
	15	15	0,590	0,458	0,560	0,430	0,451	0,440	0,912
	15	30	0,583	0,438	0,549	0,411	0,434	0,408	0,920
	20	20	0,574	0,481	0,599	0,446	0,467	0,461	0,931
	20	40	0,757	0,664	0,710	0,660	0,663	0,656	1
	30	50	0,889	0,797	0,819	0,793	0,796	0,795	1

Tablo4.24. (Devam)

	7	7	0,141	0,150	0,717	0,035	0,136	0,084	0,146
	7	15	0,356	0,301	0,553	0,261	0,291	0,239	0,538
	10	20	0,516	0,459	0,636	0,437	0,459	0,418	0,876
	15	15	0,645	0,520	0,675	0,485	0,504	0,488	0,938
4	15	30	0,640	0,542	0,678	0,499	0,533	0,507	0,927
	20	20	0,671	0,607	0,744	0,552	0,598	0,565	0,955
	20	40	0,807	0,77	0,815	0,766	0,772	0,757	0,999
	30	50	0,926	0,866	0,89	0,865	0,867	0,859	1

Tablo 4.24. incelendiğinde Dizayn 4 için $\Delta = 2$ iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği önerdiğimiz RobCAT algoritmasının ise genel olarak etkilenmediği ve 1'e yakın sonuçlar verdiği görülmektedir.

Tablo 4.25. Kirlili veride Dizayn 1 için $\Delta = 5$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
	7	7	0,243	0,126	0,313	0,073	0,115	0,105	0,967
	7	15	0,392	0,153	0,276	0,171	0,169	0,136	1
	10	20	0,752	0,287	0,377	0,295	0,295	0,283	1
2	15	15	0,470	0,324	0,392	0,298	0,315	0,312	1
	15	30	0,962	0,527	0,594	0,529	0,526	0,523	1
	20	20	0,551	0,428	0,482	0,408	0,421	0,419	1
	20	40	0,997	0,796	0,835	0,793	0,796	0,775	1
	30	50	1	0,953	0,962	0,952	0,953	0,949	1
	7	7	0,326	0,229	0,559	0,105	0,219	0,160	0,922
	7	15	0,418	0,363	0,522	0,353	0,356	0,337	0,983
	10	20	0,825	0,531	0,642	0,528	0,535	0,515	1
3	15	15	0,709	0,552	0,662	0,505	0,550	0,534	1
	15	30	0,988	0,731	0,789	0,724	0,732	0,705	1
	20	20	0,786	0,646	0,711	0,608	0,641	0,630	1
	20	40	0,998	0,866	0,904	0,858	0,863	0,857	1
	30	50	1	0,976	0,982	0,974	0,976	0,973	1
	7	7	0,288	0,316	0,787	0,081	0,315	0,205	0,835
	7	15	0,395	0,560	0,764	0,536	0,549	0,519	0,978
	10	20	0,819	0,672	0,783	0,654	0,671	0,655	1
	15	15	0,765	0,673	0,786	0,593	0,674	0,640	1
4	15	30	0,995	0,856	0,900	0,842	0,853	0,845	1
	20	20	0,886	0,801	0,863	0,764	0,795	0,784	1
	20	40	0,999	0,938	0,958	0,931	0,935	0,937	1
	30	50	1	0,986	0,990	0,984	0,987	0,983	1

Tablo 4.25. incelendiğinde Dizayn 1 için $\Delta = 5$ iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği önerdiğimiz RobCAT algoritmasının ise etkilenmediği, 1 ve 1'e yakın sonuçlar verdiği görülmektedir. Tablo 4.25.'e göre RobCAT algoritmasının gücü diğer algoritmaların gücünden

oldukça yüksektir. Bu durum önerilen RobCAT algoritmasının verideki aykırı değerlerden etkilenmediğini göstermektedir.

Tablo 4.26. Kirli veride Dizayn 2 için $\Delta = 5$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,229	0,118	0,312	0,083	0,112	0,099	0,966
	7	15	0,556	0,212	0,328	0,234	0,227	0,201	0,998
	10	20	0,713	0,293	0,396	0,308	0,308	0,283	1
	15	15	0,854	0,431	0,500	0,418	0,428	0,416	1
	15	30	0,850	0,394	0,461	0,384	0,394	0,383	1
	20	20	0,568	0,454	0,512	0,429	0,444	0,443	1
	20	40	0,992	0,777	0,817	0,775	0,778	0,775	1
	30	50	1	0,942	0,950	0,940	0,941	0,938	1
3	7	7	0,323	0,226	0,577	0,106	0,225	0,163	0,934
	7	15	0,455	0,390	0,575	0,391	0,393	0,366	0,993
	10	20	0,811	0,519	0,642	0,523	0,532	0,519	1
	15	15	0,673	0,539	0,642	0,492	0,543	0,531	1
	15	30	0,983	0,711	0,767	0,707	0,713	0,702	1
	20	20	0,768	0,654	0,731	0,619	0,654	0,648	1
	20	40	0,998	0,904	0,932	0,901	0,906	0,898	1
	30	50	1	0,981	0,988	0,981	0,981	0,981	1
4	7	7	0,265	0,293	0,779	0,081	0,298	0,185	0,820
	7	15	0,407	0,535	0,740	0,516	0,531	0,494	0,977
	10	20	0,855	0,716	0,830	0,707	0,719	0,698	1
	15	15	0,794	0,697	0,814	0,610	0,688	0,669	1
	15	30	0,988	0,852	0,902	0,849	0,854	0,841	1
	20	20	0,859	0,785	0,849	0,742	0,784	0,765	1
	20	40	0,999	0,935	0,955	0,934	0,935	0,931	1
	30	50	1	0,988	0,995	0,988	0,988	0,987	1

Tablo 4.26 incelendiğinde Dizayn 2 için $\Delta = 5$ iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği önerdiğimiz RobCAT algoritmasının ise etkilenmediği, 1 ve 1'e yakın sonuçlar verdiği görülmektedir.

Tablo 4.27. Kirli veride Dizayn 3 için $\Delta = 5$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,487	0,314	0,434	0,331	0,332	0,288	0,997
	7	15	0,518	0,337	0,459	0,351	0,351	0,319	0,999
	10	20	0,635	0,410	0,473	0,409	0,414	0,399	1
	15	15	0,743	0,515	0,577	0,505	0,510	0,509	1
	15	30	0,743	0,451	0,521	0,444	0,452	0,439	1
	20	20	0,550	0,486	0,536	0,473	0,480	0,477	1
	20	40	0,989	0,777	0,811	0,777	0,778	0,770	1
	30	50	1	0,912	0,929	0,912	0,913	0,912	1

Tablo 4.27. (Devam)

3	7	7	0,544	0,539	0,662	0,551	0,547	0,515	0,996
	7	15	0,555	0,541	0,679	0,550	0,550	0,523	0,994
	10	20	0,698	0,597	0,696	0,589	0,606	0,581	1
	15	15	0,845	0,683	0,762	0,669	0,685	0,667	1
	15	30	0,822	0,669	0,737	0,650	0,666	0,651	1
	20	20	0,743	0,694	0,761	0,683	0,689	0,682	1
	20	40	0,994	0,881	0,908	0,880	0,883	0,878	1
	30	50	1	0,961	0,971	0,961	0,962	0,960	1
4	7	7	0,441	0,661	0,801	0,665	0,665	0,648	0,988
	7	15	0,453	0,681	0,810	0,675	0,687	0,656	0,990
	10	20	0,737	0,739	0,845	0,739	0,752	0,707	1
	15	15	0,889	0,806	0,872	0,774	0,798	0,777	1
	15	30	0,864	0,772	0,866	0,753	0,776	0,755	1
	20	20	0,852	0,816	0,889	0,786	0,808	0,793	1
	20	40	0,999	0,953	0,968	0,951	0,955	0,947	1
	30	50	1	0,980	0,985	0,978	0,981	0,977	1

Tabloya göre, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği önerdiğimiz algoritmanın ise etkilenmediği görülmektedir.

Tablo 4.28. Kirli veride Dizayn 4 için $\Delta = 5$ iken test istatistiklerinin gücü

p	n ₁	n ₂	T _B	T _Y	T _J	T _{NV}	T _{M-NV}	CAT	RobCAT
2	7	7	0,244	0,177	0,437	0,130	0,162	0,155	0,690
	7	15	0,396	0,280	0,398	0,291	0,290	0,266	1
	10	20	0,626	0,421	0,512	0,431	0,433	0,411	1
	15	15	0,470	0,407	0,482	0,389	0,403	0,402	1
	15	30	0,932	0,637	0,691	0,636	0,640	0,626	1
	20	20	0,576	0,524	0,577	0,512	0,521	0,514	1
	20	40	0,571	0,517	0,585	0,500	0,512	0,506	1
	30	50	1	0,906	0,926	0,906	0,907	0,901	1
3	7	7	0,518	0,542	0,667	0,556	0,559	0,517	0,995
	7	15	0,511	0,484	0,622	0,499	0,499	0,465	0,996
	10	20	0,705	0,607	0,697	0,598	0,613	0,578	1
	15	15	0,845	0,677	0,756	0,659	0,682	0,657	1
	15	30	0,819	0,679	0,749	0,656	0,675	0,662	1
	20	20	0,732	0,696	0,762	0,676	0,693	0,682	1
	20	40	0,994	0,901	0,923	0,899	0,901	0,895	1
	30	50	1	0,956	0,964	0,953	0,955	0,951	1
4	7	7	0,470	0,679	0,815	0,681	0,682	0,667	0,986
	7	15	0,478	0,662	0,800	0,655	0,665	0,633	0,984
	10	20	0,765	0,736	0,841	0,730	0,746	0,709	1
	15	15	0,884	0,801	0,873	0,778	0,802	0,781	1
	15	30	0,884	0,777	0,871	0,752	0,778	0,763	1
	20	20	0,838	0,811	0,881	0,781	0,799	0,784	1
	20	40	0,995	0,948	0,966	0,948	0,948	0,939	1
	30	50	1	0,982	0,989	0,980	0,983	0,978	1

Tablo 4.28 incelendiğinde Dizayn 4 için $\Delta = 5$ iken, daha önceden önerilen testlerin veri setindeki aykırı değerlerden oldukça etkilendiği önerdiğimiz RobCAT algoritmasının ise etkilenmediği, 1 ve 1'e yakın sonuçlar verdiği görülmektedir.

Genel olarak $\Delta = 2$ ve $\Delta = 5$ iken deneysel güç değerlerini inceleyecek olursak, $\Delta = 5$ iken n_1, n_2 örneklem kombinasyonlarında testin gücü değerlerinin 1 ve 1'e yakın sonuç verdiği ve daha önceden önerilen test istatistiklerinin de 1'e yaklaştığı görülmektedir.



5. SONUÇLAR

Bu çalışmada, çok değişkenli Behrens-Fisher problemi altında iki bağımsız grubun ortalama vektörlerini karşılaştırmak için kullanılabilecek robust bir test yaklaşımı önerilmektedir. RobCAT adı verilen bu yaklaşım CAT algoritmasına ve MCD tahmin edicilerine dayanmaktadır. Bu problem altında daha önceden kullanılan testler ve RobCAT algoritmasına dayanan test istatistiği deneysel olarak tip-1 hata oranları ile güç değerleri bakımından karşılaştırılmıştır.

Simülasyon sonuçları incelendiğinde, aykırı değerlerin eklenmediği temiz veri setlerinde önerilen RobCAT algoritmasının daha önceden kullanılan testlere yakın tip-1 hata oranlarına ve güç değerlerine sahip olduğu görülmektedir. Öte yandan, veriye aykırı değerler eklendiğinde RobCAT algoritmasının tip-1 hata oranı ve güç değerleri bakımından diğer testlere göre daha başarılı olduğu görülmektedir. Daha önceden önerilen testlerin ise aykırı değerlerden oldukça etkilendiği görülmektedir.

Sonuç olarak, çok değişkenli Behrens-Fisher problemi altında, RobCAT algoritması daha önceden önerilen testler gibi herhangi bir asimptotik dağılıma dayanmadığından küçük örneklem boyutları için kullanılabilir. Ayrıca önerilen bu yaklaşım daha önce önerilmiş olanların aksine kirlenmiş veri setlerinde de kullanılabilmektedir.

Ayrıca, RobCAT algoritmasını gerçek uygulama çalışmalarında kullanabilmek için MVTTests paketinde (Bulut, 2019:132-138) bir R fonksiyonu oluşturulmuştur. Böylece, araştırmacılar bu R fonksiyonunu kullanarak gerçek veri setlerinde aykırı değerlerden etkilenmeden RobCat algoritmasını uygulayabileceklerdir. Bu özellikleri göz önüne alındığında hem çok değişkenli Behrens-Fisher hem de aykırı değer problemlerine aynı anda çözüm bulan bu çalışmanın, çok değişkenli istatistiksel çıkarsamaya katkı sağlayacağına inanılmaktadır.

KAYNAKLAR

- Alpar, R. (2011). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*, Detay Yayıncılık. Baskı. Ankara
- Aksoy, S. ve Gökpınar, F. (2020). Homojen Olmayan Varyans Varsayımı Altında Ortalamaların Eşitliği için Skor ve Wald İstatistiğine Dayalı Alternatif Testler. *Gazi Üniversitesi Fen Fakültesi Dergisi*, 1(1-2), 78-100.
- Behrens, W. (1929). Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtungen (A contribution to error estimation with few observations), *Landwirtschaftliches Jahrbuch*.
- Bennett, B.M. (1951). Note on a solution of the generalized Behrens–Fisher problem, *Annals of the Institute Statistical Mathematics*.
- Bulut, H. and Öner, Y. (2017) The evaluation of socio-economic development of development agency regions in Turkey using classical and robust principal component analyses. *Journal of Applied Statistics*.
- Bulut, H. (2019). An R package for multivariate hypothesis tests: MVTests. *Technological Applied Sciences*, 14(4), 132-138.
- Chang, C. H. and Pal, N. (2008). A revisit to the Behrens–Fisher problem: comparison of five test methods. *Communications in Statistics—Simulation and Computation®*, 37(6), 1064-1085.
- Chang, C. H. et al. (2010). A note on comparing several poisson means. *Communications in Statistics-Simulation and Computation*, 39(8), 1605-1627.
- Chang, C. H. et al. (2011). Testing the equality of several gamma means: a parametric bootstrap method with applications. *Computational Statistics*, 26(1), 55-76.
- Chang, C. H. et al. (2016). A revisit to test the equality of variances of several populations. *Communications in Statistics: Simulation and Computation*, 46, 6360-6384.
- Christensen, W. F. and Rencher, A. C. (1997). A comparison of Type I error rates and power levels for seven solutions to the multivariate Behrens-Fisher problem. *Communications in Statistics-Simulation and Computation*, 26(4), 1251-1273.
- Erdoğan, S. (2018). *Heterojenlik altında iki grup ortalama vektörlerinin karşılaştırılması için önerilen yeni bir hesaplamalı yaklaşım testi*. Basılmamış Yüksek Lisans Tezi. Gazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, Ankara.
- Finch, H. and French, B. (2013). A Monte Carlo comparison of robust MANOVA test statistics. *Journal of Modern Applied Statistical Methods*, 12(2), 4.
- Fisher, R. A. (1935). The fiducial argument in statistical inference. *Annals of eugenics*, 6(4), 391-398.
- Gökpınar E.Y. and Gökpınar F. (2012). A test based on computational approach for equality of means under unequal variance assumption. *Hacettepe Journal of Mathematics and Statistics*, 41(4), 605-613.
- Gökpınar E. et al. (2013). A new computational approach for testing equality of inverse Gauss means under heterogeneity. *Hacettepe Journal of Mathematics and Statistics*, 42(5), 585-590.
- Gökpınar F. and Gökpınar E. (2014). A computational approach for testing equality of coefficients of variation in k normal populations. *Hacettepe Journal Of Mathematics And Statistics*, 44 (5), 1197-1213.

- Gökpınar, E. et al. (2020). A Computational Approach Test for the Equality of Two Multivariate Normal Mean Vectors under Heterogeneity of Covariance Matrices. *REVSTAT Statistical Journal*.
- Hwang, H.L. and Paulson, A.S. (1986). Some methods for the multivariate two-sample location problem. Unpublished manuscript, *Oak Ridge National Laboratory*, Oak Ridge, TN.
- Ito, K. and Schull, W.J. (1964). On the robustness of the T^2 test in multivariate analysis of variance when variance-covariance matrices are not equal. *Biometrika*, 51(1-2): 71-82.
- Jafari, A.A. and Kazemi, M.R. (2017). Computational approach test for inference about several correlation coefficients: Equality and common. *Communications in Statistics-Simulation and Computation*, 0,1-14.
- James, G.S. (1954). Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown. *Biometrika*, 41, 19-43.
- Johansen, S. (1980). The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression. *Biometrika*, 67(1).
- Kim, S. (1992). A practical solution to the multivariate Behrens- Fisher problem. *Biometrika*, 79(1): 171-176.
- Krishnamoorthy, K. and Yu J. (2004). Modified Nel and Van der Merwe test for the multivariate Behrens-Fisher problem. *Statistics & probability letters*, 66(2).
- Lee, S.M.S. (1994). Optimal choice between parametric and non-parametric bootstrap estimates. in *Mathematical Proceedings of the Cambridge Philosophical Society*.
- Mutlu, H.T. et al. (2017). A new computational approach test for one-way ANOVA under heteroscedasticity. *Communications in Statistics-Theory and Methods*, 46(16).
- Nel, D. and Van der Merwe C. (1986). A solution to the multivariate Behrens-Fisher problem. *Communications in Statistics-Theory and Methods*, 15(12): p.3719-3735.
- Özdamar, K. (2018). *Eğitim, Sağlık ve Sosyal Bilimler İçin SPSS Uygulamalı Temel İstatistik*. Nisan Kitabevi.
- Pal, N. et al. (2007). A computational approach to statistical inferences. *Journal of Applied Probability and Statistics*. 2(1): p. 13-35.
- Rencher, A. (2002). *Methods of multivariate analysis*. Canada: John Willey & Sons. Inc. Publications.
- Rousseeuw, P.J. (1985). Multivariate estimation with high breakdown point. *Mathematical statistics and applications*. 8(37): p. 283-297.
- Sandal, M. (2020). *Kovaryans Matrislerinin Homojenliği Varsayımı Sağlanmadığında İstatistiksel Çözümleme Yaklaşımları*. Basılmamış Doktora Tezi. Osmangazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, Eskişehir.
- Scheffé, H. (1943). On solutions of the Behrens-Fisher problem, based on the t-distribution. *Annals of Mathematical Statistics*, 14(1): 35-44.
- Subrahmaniam, K. and Subrahmaniam, K. (1973). On the multivariate Behrens-Fisher problem. *Biometrika*, 60, 107-111.
- Todorov, V. and P. Filzmoser. (2010a). Robust Statistic for the One-way MANOVA. *Computational Statistics & Data Analysis*. 54(1): p. 37-48.
- Todorov, V. and P. Filzmoser. (2010b). An object-oriented framework for robust multivariate analysis. *Journal of Statistical Software*. 32: p. 1-47.

- Tukey, J.W. (1959). A quick compact two sample test to Duckworth's specifications. *Technometrics*. 1(1): p. 31-48.
- Yanagihara, H. and Yuan K.H. (2005). Three approximate solutions to the multivariate Behrens–Fisher problem. *Communications in Statistics Simulation and Computation*. 34(4): p. 975-988.
- Yao, Y. (1965). An approximate degrees of freedom solution to the multivariate Behrens Fisher problem. *Biometrika*. 52(1/2): p. 139-147.
- Welch, B. L. (1947). The generalization of student's problem when several different population variances are involved. *Biometrika*, 34, 28-35.
- Willems, G. et al. (2002). A robust Hotelling test. *Metrika*. 55(1-2): p. 125-138.



ÖZ GEÇMİŞ

Gölnur KARAOSMAN, Samsun Milli Piyango Anadolu Lisesi'ni bitirdikten sonra Ondokuz Mayıs Üniversitesi Fen Edebiyat Faköltesi İstatistik Anabilim dalından 2020 yılında mezun oldu. 2020 yılında OMÜ İstatistik Yüksek Lisans programına girdi.

İletişim Bilgileri

ORCID ID: 0000-0001-7455-2404



