# İki Örneklem Behrens Fisher Problemi

Esin Budak

**YÜKSEK LİSANS TEZİ**

İstatistik Anabilim Dalı

Temmuz 2021

Two Samples Behrens Fisher Problem

Esin Budak

**MASTER OF SCIENCE THESIS**

Department of Statistics

July 2021

# İki Örneklem Behrens Fisher Problemi

Esin Budak

Eskişehir Osmangazi Üniversitesi
Fen Bilimleri Enstitüsü
Lisansüstü Yönetmeliği Uyarınca
İstatistik Anabilim Dalı
Uygulamalı İstatistik Bilim Dalında
YÜKSEK LİSANS TEZİ
Olarak Hazırlanmıştır

Danışman: Prof. Dr. Zeki Yıldız

Temmuz 2021

**ETİK BEYAN**

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Prof. Dr. Zeki Yıldız danışmanlığında hazırlamış olduğum “İki Örneklem Behrens Fisher Problemi” başlıklı tezimin özgün bir çalışma olduğunu; tez çalışmamın tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; tezimde verdiğim bilgileri, verileri akademik ve bilimsel etik ilke ve kurallara uygun olarak elde ettiğimi; tez çalışmamda yararlandığım eserlerin tümüne atıf yaptığımı ve kaynak gösterdiğimi ve bilgi, belge ve sonuçları bilimsel etik ilke ve kurallara göre sunduğumu beyan ederim. 01/07/2021

Esin BUDAK

# ÖZET

Çok değişkenli istatistiksel çalışmalarda, iki ortalama vektörünün eşitliğini test edilmesini araştırmak en eski konulardan biridir. Gerçekten de, birçok çıkarım yönteminin yanı sıra pratik sorunlara çeşitli analitik yaklaşımlar için Behrens-Fisher problemi bir test alanı olmuştur. Bu çalışmada, iki gruplu tasarımlarda çok değişkenli ortalama eşitlik hipotezlerini test etmek için istatistiksel prosedürleri incelenmiştir. Çok değişkenli ortalamalar için geleneksel test olan Hotelling'in $T^2$'si, verilerin dağılımı, anakütle varyansları ve kovaryansları hakkında belirli varsayımlara dayanır. Hotelling $T^2$'si, çok değişkenli normal dağılım varsayımı altında iki ortalama vektörünü karşılaştırmak için kullanılan bir testtir. Ancak, bu test varyans-kovaryans matrisinin homojenliği varsayımına dayanmaktadır. Bu varsayım sağlanmadığı zaman, bu durum literatürde çok değişkenli Behrens-Fisher problemi olarak bilinir. Çok değişkenli Behrens-Fisher problemi söz konusu olduğunda, homojen varyans-kovaryans matrisine dayalı test istatistiklerinin performansı etkilenmektedir. Çok değişkenli normal dağılım varsayımı altında iki ortalama vektörün eşitliğini test etmek için literatürde birçok çalışma bulunmaktadır. En önemli çalışmalar, Bennett (1951), Yao (1965), Johansen (1980), Nel ve Van Der Merwe (1986), Modifiye Edilmiş Nel ve Van der Merwe Testi, Yanagihara ve Yuan (2005) F yaklaşımı, Düzeltilmiş Bartlet (MB) testi, Yanlılık düzeltme (BC) ve ikinci sıra (S) yöntemleridir. Yukarıda bahsedilen testler için, çok değişkenli Behrens-Fisher Problemleri Simülasyon Çalışmaları R-studio'da çalıştırılıp tablo değerleri elde edilmiştir. Ayrıca, Mısır'da beş farklı çağdan kalma erkek kafatası örnekleri üzerine yapılan ölçümler alınarak örnek bir uygulamaya yer verilmiştir. Araştırmanın önemi, iki değişkenli örneklem büyüklüklerinde daha iyi performans ve sonuçların elde edilebilir olduğunu göstermektedir.

**Anahtar Kelimeler:** Çok Değişkenli Behrens-Fisher Problemi, Varyans-Kovaryans Matrisi, Hotelling $T^2$

## SUMMARY

In multivariate statistical studies, investigating the equality of two mean vectors is one of the oldest topics. Indeed, the Behrens-Fisher problem has been a testing ground for many inference methods as well as for various analytical approaches to practical problems. In this study, statistical procedures for testing multivariate mean equality hypotheses in two-group designs were examined. Hotelling's $T^2$, the traditional test for multivariate means, relies on certain assumptions about the distribution of data, population variances, and covariances. Hotelling's $T^2$ is a test used to compare two mean vectors under the assumption of a multivariate normal distribution. However, this test is based on the assumption of homogeneity of the variance-covariance matrix. When this assumption is not met, it is known in the literature as the multivariate Behrens-Fisher problem. In the case of the multivariate Behrens-Fisher problem, the performance of test statistics based on homogeneous variance-covariance matrix is affected. There are many studies in the literature to test the equality of two mean vectors under the assumption of a multivariate normal distribution. The most important studies are Bennett (1951), Yao (1965), Johansen (1980), Nel and Van Der Merwe (1986), Modified Nel and Van der Merwe Test, Yanagihara and Yuan (2005) F Approach, Modified Bartlet (MB), Bias Correction Presedure (BC) and Second Order Presedure (S). For the tests mentioned above, multivariate Behrens-Fisher Problems Simulation Studies were run in R-studio and table values were obtained. In addition, measurements made on male skull samples from five different eras in Egypt were taken and a sample application was included. The importance of the research shows that better performance and results can be obtained in bivariate sample sizes.

**Key Words:** Multivariate Behrens-Fisher Problem, Variance Co-variance Matrix, Hotelling $T^2$

# İÇİNDEKİLER

**Sayfa**

**İÇİNDEKİLER(devam)**

<u>**Sayfa**</u>

# ÇİZELGELER DİZİNİ

**Çizelge** **Sayfa**

# SİMGELER VE KISALTMALAR DİZİNİ

| **Simgeler** | **Açıklama** |
|---|---|
| A | Anlamlılık Düzeyi |
| $H_0$ | Yokluk (sıfır) hipotezi |
| $\boldsymbol{H_1}$ | Alternatif hipotez |
| $\boldsymbol{\mu}$ | Ortalamalar Vektörü |
| $\Sigma$ | Kovaryans matrisi |
| $P$ | Bağımlı değişken sayısı |
| $\boldsymbol{\rho}$ | Korelasyon katsayısı |
| $F$ | F dağılımı |
| $k$ | Grup sayısı |
| $\boldsymbol{T_{Hot}}$ | Hotelling testi |
| $\boldsymbol{T_{Joh}}$ | Johansen testi |
| $\boldsymbol{T_{Yao}}$ | Yao testi |
| $\boldsymbol{T_{MNV}}$ | Modifiye edilmiş Nel ve Van der Merwe testi |
| $\boldsymbol{T_{BC}}$ | Yanlılık düzeltme testi |
| $\boldsymbol{T_{S}}$ | İkinci sıra (second) testi |
| $\boldsymbol{T_{YY}}$ | Yanagihara ve Yuan testi |
| $\boldsymbol{T_{NV}}$ | Nel ve Van der Merwe testi |
| $\boldsymbol{T_{MB}}$ | Modifiye edilmiş Bartlet testi |
| **Kısaltmalar** | **Açıklama** |
| AAD | Ortalama mutlak tutarsızlık |
| BC | Yanlılık düzeltme |
| BF | Behrens-Fisher |
| CAT | Hesaplamalı yaklaşım testi |
| MB | Modifiye edilmiş Bartlet |
| MNV | Modifiye edilmiş Nel ve Van der Merwe |
| NV | Nel ve Van der Merwe |
| S | İkinci sıra |
| YY | Yanagihara ve Yua |

# 1. GİRİŞ VE AMAÇ

İstatistiksel çözümlemede kullanılacak doğru tekniğin seçimi önemlidir. İstatistiksel çözümleme teknikleri, özellikle parametrik teknikler çeşitli önemli varsayımları gerektirirler. Bu tekniklerle yapılacak çözümlemelerin sonuçlarındaki anlamlılık söz konusu varsayımların sağlanmasına bağlıdır.

İki ve daha fazla gruba ait ortalamaların karşılaştırılması genelde istatistiksel çalışmalarda karşılaşılan bir problemdir. (Sandal, 2020) Çok değişkenli istatistiksel analiz tekniği ise, iki veya daha fazla değişkeni birlikte incelemeyi sağlayan istatistiksel yöntemleri belirtmek için kullanılmaktadır. (Erdoğan, 2018). İki anakütle arasındaki farkın test edilmesi, araştırmacılar ve uygulamalı istatistikçiler tarafından sıkça karşılaşılan sorunlardandır. Temel alınan birim ve değişken sayısı çok fazla olup, bu birim ve değişkenlerin karşılıklı etkileşimlerinde söz konusudur. Bundan dolayı bu etkileşimleri dikkate alan çok değişkenli istatistiksel analiz tekniklerine ihtiyaç duyulmaktadır.

Birden fazla bağımlı değişken söz konusu olduğunda da, çok değişkenli istatistiksel tekniklerin kullanılması gerekmektedir (Alpar, 2011; Finch ve French, 2013). Yöntemlerin kullanılabilmesi için gereken şartlar; gözlemlerin bağımsız, normal dağılıma sahip ve varyansların homojen olmasıdır. İki bağımsız örneklem kullanılmasının sebebi, dağılımların ilişkileri karşılaştırılırken, analistin veriler hakkında yapacağı varsayımlara bağlı olarak birçok farklı yaklaşımın olmasıdır.

t-testi, birbirinden bağımsız iki grubun ortalamalarını karşılaştırmak için kullanılan testtir; erkeklere karşı kadınlar, sporculara karşı sporcu olmayanlar, gençlere karşı yaşlı olanlar durumunda ya da bir testte ölçülen ortalamaları karşılaştırma durumlarında incelemek istediğimizde kullanılması tercih edilen testtir. t-testi, iki ortalama arasındaki farkın istatistiksel anlamlılığını test etmek için kullanılan parametrik bir tekniktir. t-testinde sadece iki ortalama karşılaştırılmaktadır. Ayrıca t-testi; normallik, eşit varyanslılık ve bağımsızlık varsayımları sağlandığında kullanılabilmektedir. Tek örneklem durumunda seçilen örneklemin o evrenden gelip gelmediği incelenirken ilişkisiz örneklem durumunda ise iki farklı örneklem ortalaması arasında farkın anlamlı olup olmadığı araştırılmaktadır.

Ayrıca ilişkili örneklemlerde ise aynı bireylere yönelik iki farklı ölçüm ortalamaları arasında anlamlı bir farklılık olup olmadığı test edilmektedir.

Ancak, varyansların homojenliği varsayımı ihlal edilirse, normal t-testi eşit olmayan örneklem büyüklükleri için artık yeterli değildir. Tip I hata olasılığı ciddi şekilde etkilenir. Bilinmeyen varyanslara sahip normal anakütlelerden iki örneklem eşitliğini test etme problemi Behrens-Fisher (BF) problemi olarak bilinir. İstatistikte, Walter Behrens ve Ronald Fisher'ın adını taşıyan Behrens-Fisher problemi, iki anakütlenin varyanslarının eşit olmadığı varsayıldığında, normal olarak dağılan iki anakütle ortalamaları arasındaki farkla ilgili aralık tahmini ve hipotez testi problemi, iki bağımsız örneğe dayanmaktadır. Homojenlik varsayımı sağlanmadığı zaman bağımsız grup ortalamaları arasındaki farkın belirlenmesine yönelik ilk çalışmalar Behrens ve Fisher tarafından gerçekleştirilmiştir. Behrens (1929) , varyansları eşit olmayan tek değişkenli iki bağımsız grup ortalamasının eşitliğini test etmek içinse bir çözüm ortaya koymuştur. Fisher (1935) ise Behrens'in (1929) çözümü olan, fiducial aralık yöntemini kullanarak doğruluğunu yine kanıtlamıştır. Homojenlik varsayımı sağlanmadığında iki veya daha fazla bağımsız grubun ortalamalarını karşılaştırma problemi "Behrens-Fisher Problemi (BFP)" olarak adlandırılmıştır. Bu problem oldukça zorlu ve karmaşık yapısı nedeniyle dikkat çekip ve ilgi duyulan önemli bir istatistiksel araştırma problemidir. BF problemin zorluğunun sebebi ise nuisance parametrelerin varlığında tam olasılıklı çözümlerin elde edilememesindendir. Temeli tek değişkenli durumlara dayanan Behrens-Fisher problemi, 1950'li yıllarından itibaren çok değişkenli problemler için de genişletilmiştir. Bu nedenle, bağımlı değişken sayısının tek olduğu durumlarda bu problem tek değişkenli Behrens- Fisher problemi olarak adlandırılırken, bağımlı değişken sayısının iki veya daha fazla olduğu durumlarda ise çok değişkenli Behrens-Fisher problemi adını almıştır.

Behrens-Fisher problemlerinin çözümünde başta Fisher (1936), Welch (1947) ve Wald (1955) olmak üzere birçok yöntem geliştirilmiştir. Ancak Wald (1955) yöntemi sadece eşit gözlem sayıları söz konusu olduğunda kullanılabilmektedir. Bu yöntemler içerisinde asimptotik olarak en güçlü test istatistiği ise Welch 'in (1947) geliştirdiği Welch-t testidir. (Pfanzagl, 1974)

Eğer çalışmalarda $\sigma_1 \neq \sigma_2$ ve ikisi de bilinmiyorsa Behrens-Fisher problemi ile karşı karşıyayız. Bu problem için evrensel olarak kabul edilmiş bir test prosedürü yoktur, ancak bir dizi test literatür taramasında incelenmiştir.

Bu tez çalışmasının amacı da; varyans-kovaryans matrislerinin homojenlik varsayımı sağlanmadığında iki ortalama vektörün eşitliğini test etmek için alternatif çözümleme yaklaşımlarını araştırmak ve çeşitli deneysel koşullar altında önerilen farklı metotlar için test istatistiklerinin performanslarını karşılaştırmaktır. Bu metotların I. tip hata olasılıkları için bir Monte Carlo simülasyon çalışması yapılmıştır.

Bu çalışmanın ikinci bölümünde literatür taramasına yer verilmiştir. Üçüncü bölümde normal dağılıma sahip ancak varyans-kovaryans matrislerinin heterojenliği altında iki anakütle ortalama vektörlerinin eşitliği hipotezinin testi için geliştirilen bazı test istatistiklerine yer verilmiştir. İki ana kütle ortalamasının eşitliği hipotezinin testi için Yao (1965) testi, Johansen (1980) testi, Nel ve Van der Merwe (1986) testi, Krishnamoorthy ve Yu (2004)'nun modifiye edilmiş Nel ve Van der Merve testi, Yanagihara ve Yuan testi, Kawasaki ve Seo tarafından önerilen Yanlılık Düzeltme (BC) testi ve İkinci Sıra (S) Yöntemleri testi ayrıntılı olarak tanıtılmıştır. Dördüncü bölümde, simülasyon yoluyla bu test istatistiklerinin, farklı örnek hacmi, farklı değişken sayısı ve farklı varyans-kovaryans matrisleri altında deneysel I.tip hata yapma olasılıkları bakımından karşılaştırmaları yapılmıştır ve gerçek veri uygulamasına yer verilmiştir. Beşinci ve son bölümde ise sonuçlara yer verilmiştir.

# 2. LİTERATÜR ARAŞTIRMASI

Homojen varyans varsayımı sağlanmadığında yokluk hipotezinin reddini sağlayan önemli çalışmalar olsa da, bazen klasik varyans analizi büyük hacimli örneklemlerde bile $H_0$ yokluk hipotezini reddedemeyebilir. (Sandal, 2020) Çeşitli alanlarda, büyük hacimli örneklemlerin elde edilemeyeceği de düşünülürse bu durum analizlerde önemli bir sıkıntı doğurabilir. Bu sebeple küçük hacimli örneklemlerle çalışma zorunluluğu ortaya çıkar. Böyle durumlarda klasik varyans analizi oldukça kötü sonuçlar vermesinden dolayı alternatif testler geliştirilmiştir. Bu testlerden ilki Behrens-Fisher problemi için önerilen Welch (1947)'in örnek çalışmasında uyguladığı bir test olarak literatürde bulunmaktadır. Welch (1951), bu testi k kitlesinin ortalamasının eşitliğine ait hipotez testi için genelleştirmiştir.

İki ortalama vektörün eşitliği testlerini dikkate almak temel bir problemdir. Eşit olmayan kovaryanslarla ortalama karşılaştırma, yapısı gereği zordur ve bu durum istatistiksel olarak Behrens-Fisher problemi olarak bilinir. Örneklem ortalama vektörü ve örnek kovaryans matrisi kullanılarak $T^2$'ye dayalı yaklaşık çözümler için Welch (1938) ve Scheffe (1943) tarafından, tek değişkenli durumlara yaklaşık çözümler önerilmiştir. Çok değişkenli Behrens-Fisher problemini çözmek için en eski yöntemlerden biri, Bennett (1951) tarafından Scheffe'nin (1943) tek değişkenli çözümüne dayanarak türetilmiştir. Bennett (1951), James (1954), Yao (1965), Subrahmaniam ve Subrahmaniam (1973), Johansen (1980), Nel ve Van der Merwe (1986), Nel vd. (1990) tarafından T'nin sıfır hipotezi altında iki veri kümesinin dağılımını elde eden bazı yaklaşık çözümler önerilmiştir. Bu konuda yapılan bazı diğer araştırmalar da Hwang ve Paulson (1986), Kim (1992), Christensen ve Rencher (1997), Krishnamoorthy ve Yu (2004)'nun testleridir. Son zamanlarda literatürde Yanagıhara ve Yuan (2005), Zezula ve diğerleri (2009)'nin çalışmaları da mevcuttur. Krishnamoorthy ve Yu (2012), daha önceki çalışmalarında değiştirilmiş Nel ve Van der Merwe'nin test prosedürünü monoton bir modele sahip eksik veri durumuna genişleten bir çözüm olarak önermektedir. Behrens-Fisher problemi, Seko ve diğerlerinin önerdiği gibi, bir monoton eksik model durumu ve $\Sigma_1 = \Sigma_2$ olduğunda, iki normal anakütlenin ortalama vektörleri arasındaki farkla ilgilidir. (2011). Gir´on ve del Castillo (2010), iki bağımsız çok değişkenli Student t dağılımının geliştirilmesi olarak tanımlanan çok değişkenli Behrens-Fisher dağılımını incelemiştir. U momentleri Yanagihara ve Yuan (2005) tarafından elde edilmiştir. Yanagihara ve Yuan (2005), çok değişkenli Behrens-Fisher problemine, yaklaşık serbestlik derecelerine sahip iki F yaklaşımı (BC ve Second) ve değiştirilmiş Bartlett istatistiği olan üç yaklaşık çözüm sağlamıştır. Bununla birlikte, kovaryans matrisleri arasındaki fark büyük olduğunda bu çözümlerin iyi tahminler olmadığı görülmektedir. Kawasaki ve Seo (2014) U istatistiğinin birinci ve ikinci momentleri için $N^{-2}$ terimine kadar asimtotik açılımlar türeterek yeni bir yaklaşık çözüm önermiştir. Fisher, Welch, Aspin, Cochran ve Cox, Qin ve Jing'in hepsi farklı çözümler önermektedir. Tsui ve Weerahandi tarafından genelleştirilmiş p-değeri kavramı kullanılarak önerilen fikir, aykırı parametrelerin varlığında Behrens-Fisher problemini test etmek için olan çözümün bir yaklaşımıdır. Tsui ve Tang, BF problemi için genelleştirilmiş p-değerinin dağılım özelliğini çoklu teste uygulamıştır. Kim ve Cohen, Behrens-Fisher problemini güvene dayalı, Bayesçi ve sıklıkçı yaklaşımlar altında ele almak için kullanılan temel kavramların ve uygulamaların gözden geçirilmesini önermektedir. Singh vd. Jackknife metodolojisini kullanarak yeni bir test önermiştir. Dong, BF problemi için ampirik olasılık yaklaşımını ele almıştır. Chang ve Pal, BF problemini yeniden gözden geçirmiş ve yeni geliştirilen bir "Hesaplamalı Yaklaşım Testi" ni (CAT) uygulamıştır. Erdoğan (2018) ise, çok değişkenli normal dağılım varsayımı altında iki ortalama vektörün eşitliğini test etmek için hesaplamalı yaklaşım testine dayalı yeni bir test istatistiği ileri sürmüştür. Sunulan yeni test istatistiği, I.tip 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.

# 3. YÖNTEM

Bu bölümde Çok değişkenli Behrens-Fisher Problemi ayrıntılı olarak ele alınacaktır.

## 3.1 Çok Değişkenli Behrens Fisher Problemi ve Çözümleme Yaklaşımları

İkiden fazla anakütlenin ortalamalarının eşitliğinin test edilmesi için gerekli varsayımlar sağlandığında Klasik F-testi yeterince güçlüdür. Ancak anakütle varyansları homojen olmadığında Klasik F-testi bu gücünü kaybetmektedir. Anakütle varyansları homojen olmadığında veya bilinmediğinde ikiden fazla anakütlenin ortalamasının eşitliğinin test edilmesi problemi ‘’Genelleştirilmiş Behrens-Fisher Problemi’’ olarak adlandırılır. İstatistiksel olarak, Walter Behrens ve Ronald Fisher'ın adını alan Behrens-Fisher sorunu, iki anakütlenin varyanslarının eşit olmadığı varsayıldığında normal dağılmış iki anakütle arasındaki farka ilişkin aralık tahmini ve hipotez testi problemidir. İki normal dağılımın ortalamaları test edilirken klasik iki örneklem t testine başvurulur. Araştırmak istenen anakütle varyansları eşit değilse ve bilinmiyorsa ya Welch (1947)’in T testi ya da Satterhwaite (1946)’ın yaklaşık F testi önerilir. Ancak Welch in T testi normal dağılım durumunda güçlüdür.

İki anakütlenin ortalamalarının eşitliğinin test edilmesi için gerekli varsayımlar sağlandığında Klasik t-testi yeterince güçlüdür. Ancak anakütle varyansları homojen olmadığında Klasik t-testi bu gücünü kaybetmektedir. Anakütle varyansları homojen olmadığında veya bilinmediğinde ikiden fazla anakütlenin ortalamasının eşitliğinin test edilmesi problemi ya da başka bir deyişle hata değişkenleri farklı ve bilinmeyen varyanslara sahip olduğunda iki ortalamanın eşitliğini test eden ‘’Genelleştirilmiş Behrens-Fisher Problemi’’ olarak adlandırılır.

İki anakütle dağılımının konum, ölçek ve şekil bakımından anakütle ortalamalarının arasındaki farklılığın testini etkileyebilir. Bu sebeple iki anakütle ortalamalrı arasındaki farklılık araştırılırken, öncelikli olarak anakütlelerin dağılımları arasında fark olup olmadığı test edilir. Bu testin sonucunda da dağılımlar arasında fark olup olmadığı test edilir. Bu testin sonucunda dağılımlar arasında fark olmadığı ve anakütlenin $N(\mu_1, \sigma_1^2)$ , $N(\mu_2, \sigma_2^2)$ şeklinde normal dağıldığı varsayılsın. Bu durumda ele alınan anakütlelerin varyansları bilinmiyorsa ve eşit değilse ($\sigma_1^2 \neq \sigma_2^2$) buradaki durum literatürde iki anakütle ortalamalarının testi olan Behrens Fisher problemi olarak bilinir.

İstatistiksel olarak, Walter Behrens ve Ronald Fisher'ın adını alan Behrens-Fisher sorunu, iki anakütle varyanslarının eşit olmadığı varsayıldığında normal dağılmış iki anakütle arasındaki farka ilişkin aralık tahmini ve hipotez testi problemidir.

Bağımlı değişken sayısının tek olduğu durumlarda bu problem tek değişkenli Behrens- Fisher problemi olarak adlandırılırken, bağımlı değişken sayısının iki veya daha fazla olduğu durumlarda ise çok değişkenli Behrens-Fisher problemi adını almaktadır.

### 3.1.1 Yaklaşık serbestlik dereceleriyle $\mathbf{T^2}$'yi kullanan yaklaşımlar

Çok değişkenli iki grup ortalama vektörü arasındaki farkın anlamlı olup olmadığını belirlemek için,

$$\begin{aligned} & H_0: \boldsymbol{\mu_1} = \boldsymbol{\mu_2} \\ & H_1: \boldsymbol{\mu_1} \neq \boldsymbol{\mu_2} \end{aligned} \tag{3.1}$$

eşitliğindeki yokluk hipotezi sınanmaktadır. Anakütle varyans-kovaryans matrisleri homojen olduğu durumlarda bu hipotezi sınamak için Hotelling $T^2$ test istatistiğinden yararlanılmaktadır. $i = 1,2$ ve $j = 1,2, \ldots, n_i$ olmak üzere $N_p(\mu_i, \Sigma_i)$ dağılımına sahip p boyutlu $X_{ij}$ rasgele vektörler için Hotelling $T^2$ test istatistiği,

$$T_H^2 = (\bar{x}_1 - \bar{x}_2)' \left[ \left( \frac{1}{n_1} + \frac{1}{n_2} \right) S_p \right]^{-1} (\bar{x}_1 - \bar{x}_2) \sim \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{sd1, sd2, (1-\alpha)} \tag{3.2}$$

biçimindedir. (Hotelling, 1931; Krishnamoorthy ve Xia, 2010). Burada i. örneklemin $\mathrm{p} \times 1$ boyutlu ortalamalar vektörü,

$$\bar{\mathrm{x}}_{\mathrm{i}} = \frac{1}{\mathrm{n}_{\mathrm{i}}} \sum_{\mathrm{j}=1}^{\mathrm{n}_{\mathrm{i}}} \mathrm{x}_{\mathrm{ij}} \tag{3.3}$$

ve ortak varyans-kovaryans matrisi ise,

$$\mathrm{S}_{\mathrm{p}} = \frac{(\mathrm{n}_1 - 1)\mathrm{S}_1 + (\mathrm{n}_2 - 1)\mathrm{S}_2}{\mathrm{n}_1 + \mathrm{n}_2 - 2} \tag{3.4}$$

olarak ifade edilmektedir. Ayrıca i. örneklemin varyans-kovaryans matrisi,

$$\mathrm{S}_{\mathrm{i}} = \frac{1}{\mathrm{n}_{\mathrm{i}} - 1} \sum_{\mathrm{j}=1}^{\mathrm{n}_{\mathrm{i}}} \left(\mathrm{x}_{\mathrm{ij}} - \bar{\mathrm{x}}_{\mathrm{i}}\right)\left(\mathrm{x}_{\mathrm{ij}} - \bar{\mathrm{x}}_{\mathrm{i}}\right)' \tag{3.5}$$

eşitliğindeki gibi hesaplanmaktadır. Eşitlik 3.2’de verilen $\mathrm{T}_{\mathrm{H}}^2$ istatistiği $sd_1 = \mathrm{p}$ ve $sd_2 = \mathrm{n}_1 + \mathrm{n}_2 - \mathrm{p} - 1$ serbestlik dereceli F dağılımı ile karşılaştırılmaktadır. Eğer test istatistiği $\mathrm{F}_{\mathrm{sd1,sd2,}(1-\alpha)}$ kritik değerinden daha büyük ise eşitlik 3.1’deki yokluk hipotezi reddedilmekte ve gruplar arasındaki farkın anlamlı olduğuna karar verilmektedir (Hotelling, 1931). Ancak varyans-kovaryans matrislerinin homojen olmadığı durumlarda ise Hotelling $\mathrm{T}^2$ istatistiği için yaklaşık bir dağılım belirlenememektedir. Bu durumda alternatif yöntemler için,

$$\mathrm{T}^2 = (\bar{\mathrm{x}}_1 - \bar{\mathrm{x}}_2)' \left(\frac{\mathrm{S}_1}{\mathrm{n}_1} + \frac{\mathrm{S}_2}{\mathrm{n}_2}\right)^{-1} (\bar{\mathrm{x}}_1 - \bar{\mathrm{x}}_2) \tag{3.6}$$

eşitliğindeki test istatistiğine dayalı olarak çözümleme yaklaşımları geliştirilmiştir. Behrens-Fisher problemi söz konusu olduğunda bu eşitlikte test istatistiği için yaklaşık çözümlerin belirlenmesi ve uygulanması gerekmektedir.

### 3.1.2 Alternatif test istatistikleri

#### 3.1.2.1 Yao testi

Welch (1947), tek değişkenli Behrens-Fisher problemleri için t dağılımına dayanan yeni bir test istatistiği önermiş ve bu istatistiğin kritik değerini belirlemek için seri açılımı ve yaklaşık serbestlik derecesi olmak üzere iki farklı yöntem geliştirmiştir. Yao (1965) ise Welch (1947)'in yaklaşık serbestlik derecesi kavramını çok değişkenli iki örneklem Behrens-Fisher problemleri için genellemiştir. Bu durumda $W_i = {}^{S_i}/_{n_i}$ ve $W = \sum_{i=1}^{2} W_i$ olmak üzere,

$$v_{Yao} = \left[\sum_{i=1}^{2} \frac{1}{n_i - 1}\left(\frac{(\bar{x}_1 - \bar{x}_2)'(W)^{-1}(W_i)(W)^{-1}(\bar{x}_1 - \bar{x}_2)}{(\bar{x}_1 - \bar{x}_2)'(W)^{-1}(\bar{x}_1 - \bar{x}_2)}\right)^2\right]^{-1} \tag{3.7}$$

olmak üzere Yao'nun (1965) test istatistiği, Eşitlik 3.6'da verilen $T^2$ istatistiğine dayalı olarak

$$T_{Yao} = \frac{v_{Yao} - p + 1}{v_{Yao} * p} T^2 \sim F_{p, v_{Yao} - p + 1} \tag{3.8}$$

eşitliği yardımıyla hesaplanmaktadır. Eğer $T_{Yao}$ test istatistiği $F_{p, v_{Yao} - p + 1}$ kritik değerinden daha büyük ise Eşitlik 3.1'deki yokluk hipotezi reddedilmektedir (Yao, 1965).

### 3.1.2.2 Johansen testi

Johansen (1980) çok değişkenli Behrens-Fisher problemleri için yaklaşık serbestlik derecesi kavramını kullanarak test istatistiğinin olasılık dağılımına yakınsamayı amaçlamıştır. Eşitlik 3.1'deki yokluk hipotezini sınamak için Johansen (1980) testinde,

$$c = p + 2D - \frac{6D}{p+2} \tag{3.9}$$

olmak üzere $T^2$ test istatistiği c sabitine oranlamakta ve

$$T_{Joh} = \frac{T^2}{c} \tag{3.10}$$

eşitliği $sd_1 = p$ ve $sd_2 = \frac{p(p+2)}{3D}$ serbestlik dereceli F dağılımına yakınsamaktadır. Burada D değeri,

$$D = \frac{1}{2}\sum_{i=1}^{2}\frac{1}{n_i - 1}\{tr[(I - (W_1^{-1} + W_2^{-1})^{-1}W_i^{-1})]^2 + [tr(I - (W_1^{-1} + W_2^{-1})^{-1}W_i^{-1})]^2\} \tag{3.11}$$

eşitliği yardımıyla hesaplanmaktadır. Eğer Johansen test istatistiği $T_{Joh} > F_{sd_1,sd_2,1-\alpha}$ olacak şekilde $F_{sd_1,sd_2,1-\alpha}$ kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir.

### 3.1.2.3 Nel ve Van der Merwe testi

Nel ve Van der Merwe (1986) ise Welch (1947)'in yaklaşık serbestlik derecesi yöntemine dayanan farklı bir test istatistiği geliştirmişlerdir. Nel ve Van Der Merwe (1986), Eşitlik 3.6'daki test istatistiğinin dağılımına yakınsamak için,

$$v_{NV}=\frac{iz((W)^2)+\left(iz(W)\right)^2}{\sum_{i=1}^{2}\frac{1}{n_i-1}\left[iz((W_i)^2)+\left(iz(W_i)\right)^2\right]} \tag{3.12}$$

olmak üzere yeni bir test istatistiği olarak,

$$T_{NV}=\frac{v_{NV}-p+1}{v_{NV}*p}T^2\sim F_{p,v_{NV}-p+1} \tag{3.13}$$

eşitliğini önermişlerdir. Eğer $T_{NV}$ test istatistiği $F_{p,v_{NV}-p+1}$ kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir (Nel ve Van Der Merwe, 1986).

### 3.1.2.4 Modifiye Edilmiş Nel ve Van der Merwe testi

Krishnamoorthy ve Yu (2004), farklı bir serbestlik derecesi dikkate alarak Nel ve Van Der Merwe'nin (1986) test istatistiğini yeniden düzenlemiştir. Böylece,

$$v_{MNV}=\frac{p+p^2}{\sum_{i=1}^{2}\frac{1}{n_i-1}\left[iz((W_iW^{-1})^2)+\left(iz(W_iW^{-1})\right)^2\right]} \tag{3.14}$$

olmak üzere Krishnamoorthy ve Yu’nun (2004) test istatistiği, Eşitlik 3.6’daki $T^2$ istatistiğine dayalı olarak,

$$T_{MNV}=\frac{v_{MNV}-p+1}{v_{MNV}*p}T^2\sim F_{p,v_{MNV}-p+1} \tag{3.15}$$

eşitliği yardımıyla hesaplanmaktadır. Eğer $T_{MNV}$ test istatistiği $F_{p,v_{MNV}-p+1}$ kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir (Krishnamoorthy ve Yu, 2004).

### 3.1.2.5 Yanagihara ve Yuan (2005) F yaklaşımı

Yanagihara ve Yuan (2005), çok değişkenli Behrens- Fisher problemlerinde $T^2$'nin olasılık dağılımına yakınsamak için bir F dağılımından yararlanmışlar ve Welch'in (1938) test yaklaşımını çok değişkenli örneklemler için genellemişlerdir. Bu durumda,

$$\hat{\theta}_1 = \frac{p\hat{\eta}_1 + (p-2)\hat{\eta}_2}{p(p+2)}$$

$$\hat{\theta}_2 = \frac{\hat{\eta}_1 + 2\hat{\eta}_2}{p(p+2)} \tag{3.16}$$

ve

$$v_{YY} = \frac{\left(n_1 + n_2 - 2 - \hat{\theta}_1\right)^2}{(n_1 + n_2 - 2)\hat{\theta}_2 - \hat{\theta}_1} \tag{3.17}$$

olmak üzere Yanagira ve Yuan'ın (2005) F dağılımına bağlı test istatistiği,

$$T_{YY} = \frac{\left(n_1 + n_2 - 2 - \hat{\theta}_1\right)}{(n_1 + n_2 - 2) * p} T^2 \sim F_{p, v_{YY}} \tag{3.18}$$

şeklinde ifade edilmektedir. Burada $S_W = \frac{n_2}{n_1+n_2} S_1 + \frac{n_1}{n_1+n_2} S_2$ olmak üzere $\hat{\eta}_1$ ve $\hat{\eta}_2$ parametrelerinin tahmincileri sırasıyla,

$$\hat{\eta}_1 = \frac{n_2^2(n_1 + n_2 - 2)}{n^2(n_1 - 1)} \left(iz\left(S_1 S_W{}^{-1}\right)\right)^2 + \frac{n_1^2(n_1 + n_2 - 2)}{n^2(n_2 - 1)} \left(iz\left(S_2 S_W{}^{-1}\right)\right)^2 \tag{3.19}$$

$$\begin{aligned} \hat{\eta}_2 = {} & \frac{n_2^2(n_1 + n_2 - 2)}{n^2(n_1 - 1)} \left(iz\left(S_1 S_W{}^{-1} S_1 S_W{}^{-1}\right)\right) \\ & + \frac{n_1^2(n_1 + n_2 - 2)}{n^2(n_2 - 1)} \left(iz\left(S_2 S_W{}^{-1} S_2 S_W{}^{-1}\right)\right) \end{aligned} \tag{3.20}$$

eşitlikleri yardımıyla hesaplanmaktadır. Eğer $T_{YY}$ test istatistiği $F_{p,v_{YY}}$ kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir (Yanagira ve Yuan, 2005; Kawasaki ve Seo, 2015).

### 3.1.2.6 Düzeltilmiş Bartlet (MB) testi

Yanagira ve Yuan (2005) çalışmalarında, Fujiskoshi'nin (2000) Düzeltilmiş Bartlett (MB) yaklaşımını çok değişkenli durumlara genelleyerek çok değişkenli Behrens-Fisher problemleri için farklı bir test istatistiği daha önermişlerdir. Yanagira ve Yuan (2005) MB test istatistiğini

$$T_{MB} = \left((n_1 + n_2 - 2)\hat{\beta}_1 + \hat{\beta}_2\right) \log\left(1 + \frac{T^2}{(n_1 + n_2 - 2)\hat{\beta}_1}\right) \sim X_p^2 \tag{3.21}$$

biçiminde tanımlamışlardır. Burada,

$$\hat{\xi}_1 = \frac{\hat{\eta}_1 + \hat{\eta}_2}{p} \tag{3.22}$$

ve

$$\hat{\xi}_2 = \frac{2(p+3)\hat{\eta}_1 + 2(p+4)\hat{\eta}_2}{p(p+2)} \tag{3.23}$$

olmak üzere Eşitlik 3.21'deki test istatistiğinde yer alan $\hat{\beta}_1$ ve $\hat{\beta}_2$ parametrelerinin tahminleri ise sırasıyla,

$$\hat{\beta}_1 = \frac{2}{\hat{\xi}_2 - 2\hat{\xi}_1} \tag{3.24}$$

$$\hat{\beta}_2 = \frac{(p+2)\hat{\xi}_2 - 2(p+4)\hat{\xi}_1}{2\left(\hat{\xi}_2 - 2\hat{\xi}_1\right)} \tag{3.25}$$

şeklinde hesaplanmaktadır. Eşitlik 3.21'de verilen $T_{MB}$ test istatistiği $p$ serbestlik dereceli $X_p^2$ kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir (Yanagira ve Yuan, 2005).

#### 3.1.2.7 Yanlılık düzeltme (BC) ve ikinci sıra (S) yöntemleri

Kawasaki ve Seo (2015), F dağılımının serbestlik derecelerini düzelterek çok değişkenli Behrens-Fisher problemi için "Yanlılık Düzeltme Yöntemi (BC)" ve "İkinci Sıra Yöntemi (S)" olmak üzere iki farklı yaklaşık çözüm önermiştir. $\mathrm{N} = \mathrm{n}_1 + \mathrm{n}_2 - 2$ olmak üzere,

$$\hat{\mathrm{v}}_{\mathrm{BC}} = \frac{2(\mathrm{N}^2 - \mathrm{N}\theta_1 + \theta_2 - \theta_3 - \theta_1^*)^2}{\mathrm{N}^2(\mathrm{N}^2 - 2\mathrm{N}\theta_1 + 2\mathrm{N}\theta_4 + 2\theta_5 - \theta_6 - 2\theta_1^* + 2\theta_4^*) - (\mathrm{N}^2 - \mathrm{N}\theta_1 + \theta_2 - \theta_3 - \theta_1^*)^2} \tag{3.26}$$

ve

$$\hat{\phi}_{\mathrm{BC}} = \frac{\mathrm{N}^2 \mathrm{v}_{\mathrm{BC}}}{\mathrm{N}^2 - \mathrm{N}\theta_1 + \theta_2 - \theta_3 - \theta_1^*} \tag{3.27}$$

için BC yöntemine ait test istatistiği,

$$\mathrm{T}_{\mathrm{BC}} = \frac{\hat{\mathrm{v}}_{\mathrm{BC}}}{\mathrm{p}\hat{\phi}_{\mathrm{BC}}} \mathrm{T} \sim \mathrm{F}_{\mathrm{p}, \hat{\mathrm{v}}_{\mathrm{BC}}, (1-\alpha)} \tag{3.28}$$

biçimindedir. Benzer şekilde S yöntemine ait test istatistiği ise,

$$\hat{\mathrm{v}}_{\mathrm{s}} = \frac{2(\mathrm{N}^2 - \mathrm{N}\theta_1 + \theta_2 - \theta_3)^2}{\mathrm{N}^2(\mathrm{N}^2 - 2\mathrm{N}\theta_1 + 2\mathrm{N}\theta_4 + 2\theta_5 - \theta_6) - (\mathrm{N}^2 - \mathrm{N}\theta_1 + \theta_2 - \theta_3)^2} \tag{3.29}$$

ve

$$\hat{\phi}_s = \frac{N^2 v_s}{N^2 - N\theta_1 + \theta_2 - \theta_3} \tag{3.30}$$

olmak üzere

$$T_s = \frac{\hat{v}_s}{p\hat{\phi}_s} T \sim F_{p,\hat{v}_s,(1-\alpha)} \tag{3.31}$$

biçiminde hesaplanmaktadır. Eğer $T_{BC}$ ve $T_S$ test istatistikleri sırasıyla $F_{p,\hat{v}_{BC},(1-\alpha)}$ ve $F_{p,\hat{v}_s,(1-\alpha)}$ kritik değerlerinden daha büyük ise yokluk hipotezi reddedilmektedir. Ayrıca $l = 1, \ldots, 6$ için $\theta_l$ ve $\theta_l^*$ değerleri Kawasaki ve Seo (2015)'da görüldüğü gibi hesaplanmaktadır.

# 4. BULGULAR VE TARTIŞMA

Bu bölümde çok değişkenli iki örneklem Behrens-Fisher probleminin çözümü için önerilen test istatistiklerinin bir simülasyon çalışması ile karşılaştırılması amaçlanmıştır. Bu amaçla RStudio programından yararlanılarak bütün test istatistiklerinin algoritmaları oluşturulmuştur. Daha sonra uygulama bölümünde çok değişkenli iki örneklem Behrens-Fisher problemlerinin çözümü için önerilen test istatistiklerinin gerçek bir veri örneği üzerinde karşılaştırılması yapılmıştır.

## 4.1 Simülasyon Çalışması

Simülasyon çalışmasının ilk aşamasında çok değişkenli normal dağılıma sahip rassal örneklemler türetilmiştir. Varyans-kovaryans matrislerinin heterojen olmasını sağlamak için $i = 1,2$ ve $j = 1, \ldots, p$ olmak üzere;

$$\Sigma_i = \begin{cases} 1 & i = j \\ (\rho_i)^{|i-j|} & i \neq j \end{cases} \tag{4.1}$$

eşitliğindeki AR(1) kovaryans modeli kullanılmıştır. Çok değişkenli normal dağılıma sahip ancak varyans-kovaryans matrisleri homojen olmayan veriler için önerilen test istatistiklerinin kritik değerleri elde edilmiştir. Böylece test istatistiğinin değeri ile kritik değerler karşılaştırılmış ve yokluk hipotezinin reddedilip reddedilemediğine karar verilmiştir. Bu işlem çok sayıda tekrar edilmiş ve her bir test istatistiği için elde edilen red sayıları, yapılan tekrar sayısına oranlanarak I. tip hata yapma olasılıkları elde edilmiştir.

Simülasyon çalışmasında test istatistiklerine ait I. tip hata yapma olasılıklarını belirlemek için 20000 tekrar (iterasyon) gerçekleştirilmiştir. İki örneklem durumunda bağımlı değişken sayısının test istatistikleri üzerindeki etkisini görmek için $p = 3,5,7,9$ olmak üzere dört farklı durum ele alınmıştır. Ayrıca gözlem büyüklükleri ile varyans-kovaryans matrislerinin heterojenlik durumlarının etkisini belirlemek amacıyla farklı deneysel koşullar altında I. tip hata yapma olasılıkları elde edilmiştir.

Simülasyon sonucunda elde edilen deneysel I. tip hata yapma olasılıklarına ($\hat{\alpha}$ ) göre test istatistiklerinin performanslarını belirlemek için Yanagira ve Yuan (2005) tarafından kullanılan “Ortalama Mutlak Tutarsızlık (AAD)” değeri hesaplanmıştır. AAD değerinin hesaplanması için

$$AAD = \frac{1}{T}\sum_{t=1}^{T}(\hat{\alpha}_t - \alpha) \quad (4.2)$$

eşitliğinden yararlanılmaktadır. Daha düşük AAD değerine sahip test istatistiklerinin anlamlılık düzeyine daha yakın değerler ortaya koyduğu söylenebilmektedir.

### 4.1.1 $p = 3$ durumunda I. tip hata yapma olasılıkları

Belirlenen deneysel koşullara göre $p = 3$ için %5 ($\alpha = 0.05$) anlamlılık düzeyinde elde edilen 1. Tip hata yapma olasılıkları Çizelge 4.1’de gösterilmiştir.

Çizelge 4.1 incelendiğinde $p = 3$ durumunda $T_{Hot}$ istatistiği için anlamlılık düzeyine en yakın değer $\rho = (0.4{,}0.6)$ ve $n_1$ gözlem büyüklüğünde .0501 olarak elde edilmiştir. Ancak tüm koşullar altında en kötü performans $\rho = (0.1{,}0.9)$ ve $n_6$ durumunda .1381 değeri ile $T_{Hot}$ istatistiğinden elde edilmiştir. Üstelik $\rho = (0.1{,}0.9)$ için $n_4$ gözlem büyüklüğünde I. tip hata yapma olasılığı %9.34 iken $n_6$ gözlem büyüklüğünde %13.81’e yükseldiği görülmektedir. Dolayısıyla $T_{Hot}$ istatistiği eşit olmayan gözlem büyüklüklerinden ve gözlem sayıları arasındaki farkın artışından çok fazla etkilenmektedir. Korelasyon katsayıları birbirine yaklaştıkça I. tip hata yapma olasılıkları anlamlılık düzeyine yaklaşmaktadır. Ancak korelasyon katsayıları arasındaki fark arttıkça, $T_{Hot}$ istatistiğinin anlamlılık düzeyinden uzaklaştığı gözlemlenmiştir.

**Çizelge 4.1 $\boldsymbol{p=3}$ için elde edilen I. tip hata yapma olasılıkları**

| $n_i$ | $(\rho_1,\rho_2)$ | $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_S$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | (0.1,0.9) | .0498 | .0537 | .0527 | .0497 | .0604 | .0491 | .0497 | .0484 | .0497 |
| | (0.2,0.8) | .0484 | .0518 | .0497 | .0487 | .0554 | .0473 | .0486 | .0479 | .0489 |
| | (0.3,0.7) | .0470 | .0496 | .0470 | .0470 | .0525 | .0446 | .0466 | .0461 | .0470 |
| | (0.4,0.6) | .0452 | .0491 | .0464 | .0463 | **.0501** | .0437 | .0461 | .0459 | .0466 |
| $n_2$ | (0.1,0.9) | .0508 | .0510 | .0530 | .0510 | .0553 | .0509 | .0509 | .0507 | .0509 |
| | (0.2,0.8) | .0507 | .0507 | .0520 | .0508 | .0527 | .0507 | .0508 | .0507 | .0508 |
| | (0.3,0.7) | .0508 | .0511 | .0518 | .0512 | .0526 | .0506 | .0511 | .0511 | .0512 |
| | (0.4,0.6) | .0510 | .0512 | .0513 | .0514 | .0519 | .0509 | .0513 | .0514 | .0514 |
| $n_3$ | (0.1,0.9) | .0501 | .0501 | .0516 | .0504 | .0527 | .0504 | .0504 | .0502 | .0504 |
| | (0.2,0.8) | .0496 | .0495 | .0502 | .0498 | .0507 | .0497 | .0498 | .0498 | .0498 |
| | (0.3,0.7) | .0489 | .0487 | .0492 | .0489 | .0496 | .0488 | .0488 | .0488 | .0488 |
| | (0.4,0.6) | .0479 | .0478 | .0480 | .0480 | .0481 | .0479 | .0479 | .0480 | .0480 |
| $n_4$ | (0.1,0.9) | .0512 | .0511 | .0561 | .0508 | **.0934** | .0512 | .0509 | .0499 | .0508 |
| | (0.2,0.8) | .0503 | .0500 | .0534 | .0499 | .0784 | .0503 | .0500 | .0497 | .0499 |
| | (0.3,0.7) | .0500 | .0493 | .0517 | .0493 | .0662 | .0496 | .0495 | .0490 | .0493 |
| | (0.4,0.6) | .0498 | .0493 | .0498 | .0493 | .0570 | .0494 | .0494 | .0490 | .0493 |
| $n_5$ | (0.1,0.9) | .0488 | .0495 | .0457 | .0498 | .0367 | .0491 | .0494 | .0493 | .0494 |
| | (0.2,0.8) | .0484 | .0485 | .0460 | .0486 | .0388 | .0483 | .0485 | .0486 | .0486 |
| | (0.3,0.7) | .0494 | .0492 | .0471 | .0492 | .0416 | .0489 | .0492 | .0492 | .0493 |
| | (0.4,0.6) | **.0502** | .0497 | .0480 | .0501 | .0450 | .0496 | **.0499** | .0497 | .0501 |
| $n_6$ | (0.1,0.9) | .0499 | .0507 | .0566 | .0503 | **.1381** | .0510 | .0506 | .0494 | .0504 |
| | (0.2,0.8) | .0505 | .0501 | .0545 | .0500 | .1031 | .0511 | .0504 | .0494 | .0501 |
| | (0.3,0.7) | .0503 | .0497 | .0522 | .0496 | .0793 | .0502 | .0499 | .0490 | .0497 |
| | (0.4,0.6) | .0501 | .0492 | .0498 | .0492 | .0617 | .0500 | .0495 | .0486 | .0492 |
| $n_7$ | (0.1,0.9) | .0513 | .0524 | .0438 | .0527 | .0341 | .0523 | .0524 | .0523 | .0525 |
| | (0.2,0.8) | .0524 | .0525 | .0461 | .0526 | .0363 | .0526 | .0527 | .0524 | .0525 |
| | (0.3,0.7) | .0527 | .0525 | .0477 | .0527 | .0389 | .0526 | .0527 | .0525 | .0527 |
| | (0.4,0.6) | **.0531** | .0521 | .0500 | .0521 | .0439 | .0525 | **.0526** | .0521 | .0522 |
| | ***AAD*** | ***.1157*** | ***.1139*** | ***.2600*** | ***.1171*** | ***1.4232*** | ***.1475*** | ***.1214*** | ***.1346*** | ***.1161*** |
| $n_1=(10,10)$ | | $n_3=(50,50)$ | | | $n_5=(30,20)$ | | $n_7=(40,20)$ | | | |
| $n_2=(25,25)$ | | $n_4=(20,30)$ | | | $n_6=(20,40)$ | | | | | |

$T_{Joh}$, $T_{Yao}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$ ve $T_{MB}$ test istatistikleri için elde edilen olasılık değerlerinin genel olarak birbirine yakın değerler ortaya koyduğu gözlemlenmiştir. Gözlem sayılarının eşit olduğu durumlarda anlamlılık düzeyine oldukça yakın değerler elde edilmiştir. Ancak en küçük ve en büyük gözlem büyüklüğü arasındaki fark arttıkça I. tip hata yapma olasılıklarının anlamlılık düzeyinden uzaklaştığı görülmektedir. $n_5$ ve $\rho=(0.4{,}0.6)$ için $T_{Yao}$ ve $T_S$ istatistikleri için sırasıyla .0502 ve .0499 değerleri elde edilmiştir. Ancak $n_7$ ve $\rho=(0.4{,}0.6)$ durumunda $T_{Yao}$ ve $T_S$ istatistiklerinin I. tip hata yapma olasılıklarının sırasıyla .0531 ve .0526 değerlerine yükseldiği görülmüştür. $T_{NV}$ istatistiği, $T_{Hot}$ istatistiğine göre daha iyi performans göstermiştir. Ancak $T_{Yao}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$ ve $T_{MB}$ test istatistiklerine göre $T_{NV}$ istatistiğinin I. tip hata yapma olasılıklarının anlamlılık düzeyinden daha uzak olduğu belirlenmiştir.

Simülasyon sonucunda elde edilen AAD değerleri incelendiğinde ise $p = 3$ için en küçük AAD değerinin .1139 ile $T_{Joh}$ istatistiğinden elde edildiği görülmektedir. Bu test istatistiğini sırasıyla $T_{Yao}$, $T_{MB}$, $T_{MNV}$, $T_S$ ve $T_{YY}$ istatistikleri takip etmektedir. $T_{BC}$ ve $T_{NV}$ istatistikleri daha yüksek AAD değerleri ortaya koymuştur. En kötü performans ise 1.4232 ile $T_{Hot}$ test istatistiğinden elde edilmiştir.

### 4.1.2 $p = 5$ durumunda I. tip hata yapma olasılıkları

Belirlenen deneysel koşullara göre $p = 5$ için %5 $(\alpha = 0.05)$ anlamlılık düzeyinde elde edilen 1. Tip hata yapma olasılıkları Çizelge 4.2'de gösterilmiştir.

**Çizelge 4.2 $p = 5$** için elde edilen I. tip hata yapma olasılıkları

| $n_i$ | $(\rho_1,\rho_2)$ | $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_S$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | (0.1,0.9) | .0542 | .0695 | .0548 | .0505 | .0723 | .0519 | .0517 | .0458 | .0531 |
| | (0.2,0.8) | .0480 | .0628 | .0483 | .0467 | .0596 | .0466 | .0480 | .0442 | .0492 |
| | (0.3,0.7) | .0444 | .0602 | .0449 | .0449 | .0547 | .0432 | .0457 | .0427 | .0473 |
| | (0.4,0.6) | .0417 | .0576 | .0423 | .0434 | **.0499** | .0417 | .0439 | .0419 | .0452 |
| $n_2$ | (0.1,0.9) | .0482 | .0491 | .0529 | .0481 | .0570 | .0485 | .0482 | .0472 | .0483 |
| | (0.2,0.8) | .0491 | **.0501** | .0512 | .0494 | .0542 | **.0495** | .0496 | .0489 | .0496 |
| | (0.3,0.7) | .0486 | .0495 | .0496 | .0492 | .0516 | .0487 | .0492 | .0488 | .0494 |
| | (0.4,0.6) | .0473 | .0482 | .0477 | .0476 | .0492 | .0469 | .0477 | .0475 | .0478 |
| $n_3$ | (0.1,0.9) | .0488 | .0489 | .0518 | .0494 | .0537 | .0494 | .0493 | .0491 | .0493 |
| | (0.2,0.8) | .0485 | .0488 | .0504 | .0493 | .0512 | .0493 | .0493 | .0491 | .0493 |
| | (0.3,0.7) | .0489 | .0488 | .0502 | .0493 | .0509 | .0493 | .0493 | .0493 | .0493 |
| | (0.4,0.6) | .0497 | .0491 | .0498 | .0496 | .0503 | .0494 | .0496 | .0495 | .0496 |
| $n_4$ | (0.1,0.9) | .0502 | .0514 | .0638 | .0491 | **.1240** | .0505 | .0493 | .0471 | .0493 |
| | (0.2,0.8) | .0498 | .0498 | .0563 | .0475 | .0909 | .0496 | .0484 | .0462 | .0479 |
| | (0.3,0.7) | .0450 | .0483 | .0515 | .0469 | .0711 | .0481 | .0476 | .0457 | .0471 |
| | (0.4,0.6) | .0492 | .0481 | .0482 | .0471 | .0578 | .0479 | .0477 | .0467 | .0473 |
| $n_5$ | (0.1,0.9) | .0477 | .0497 | .0408 | .0492 | .0303 | .0489 | .0491 | .0488 | .0491 |
| | (0.2,0.8) | .0468 | .0483 | .0398 | .0474 | .0338 | .0468 | .0473 | .0471 | .0475 |
| | (0.3,0.7) | .0462 | .0467 | .0400 | .0461 | .0369 | .0455 | .0462 | .0457 | .0463 |
| | (0.4,0.6) | .0467 | .0475 | .0427 | .0467 | .0416 | .0467 | .0471 | .0463 | .0468 |
| $n_6$ | (0.1,0.9) | .0524 | .0539 | .0648 | .0503 | **.1955** | .0529 | .0512 | .0479 | .0508 |
| | (0.2,0.8) | .0530 | **.0520** | .0587 | .0500 | .1357 | **.0529** | .0507 | .0481 | .0503 |
| | (0.3,0.7) | .0535 | .0507 | .0537 | .0488 | .0962 | .0522 | .0503 | .0473 | .0492 |
| | (0.4,0.6) | .0530 | .0495 | .0488 | .0480 | .0669 | .0510 | .0493 | .0468 | .0481 |
| $n_7$ | (0.1,0.9) | .0510 | .0523 | .0334 | .0523 | .0263 | .0514 | .0519 | .0514 | .0521 |
| | (0.2,0.8) | .0503 | .0508 | .0368 | .0507 | .0290 | .0504 | .0508 | .0502 | .0507 |
| | (0.3,0.7) | .0517 | .0510 | .0340 | .0503 | .0331 | .0512 | .0510 | .0498 | .0504 |
| | (0.4,0.6) | .0536 | .0519 | .0450 | .0514 | .0403 | .0529 | .0520 | .0507 | .0515 |
| | ***AAD*** | ***.2438*** | ***.3002*** | ***.6000*** | ***.1854*** | ***2.2257*** | ***.2191*** | ***.1707*** | ***.2673*** | ***.1643*** |
| $n_1 = (10,10)$ <br> $n_2 = (25,25)$ | | $n_3 = (50,50)$ <br> $n_4 = (20,30)$ | | | $n_5 = (30,20)$ <br> $n_6 = (20,40)$ | | $n_7 = (40,20)$ | | | |

Çizelge 4.2 incelendiğinde $p = 5$ durumunda $T_{Hot}$ istatistiği için anlamlılık düzeyine en yakın değer $\rho = (0.4,0.6)$ ve $n_1$ gözlem büyüklüğünde .0499 olarak elde edilmiştir. Ancak tüm koşullar altında en kötü performans $\rho = (0.1,0.9)$ ve $n_6$ durumunda .1955 değeri ile $T_{Hot}$ istatistiğinden elde edilmiştir. Üstelik $\rho = (0.1,0.9)$ için $n_4$ gözlem büyüklüğünde I. tip hata yapma olasılığı %12.40 iken $n_6$ gözlem büyüklüğünde %19.55'e yükseldiği görülmektedir. Dolayısıyla $T_{Hot}$ istatistiği eşit olmayan gözlem büyüklüklerinden ve gözlem sayıları arasındaki farkın artışından çok fazla etkilenmektedir. Korelasyon katsayıları birbirine yaklaştıkça I. tip hata yapma olasılıkları anlamlılık düzeyine yaklaşmaktadır. Ancak korelasyon katsayıları arasındaki fark arttıkça, $T_{Hot}$ istatistiğinin anlamlılık düzeyinden uzaklaştığı gözlemlenmiştir. Çizelge 4.2 'de de bağımlı değişken sayısı arttıkça en kötü performansı $T_{Hot}$ test istatistiği göstermektedir.

$T_{Joh}$, $T_{Yao}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$, $T_{NV}$ ve $T_{MB}$ test istatistikleri için elde edilen olasılık değerlerinin genel olarak birbirine yakın değerler ortaya koyduğu gözlemlenmiştir. Gözlem sayılarının eşit olduğu durumlarda anlamlılık düzeyine oldukça yakın değerler elde edilmiştir. Ancak en küçük ve en büyük gözlem büyüklüğü arasındaki fark arttıkça I. tip hata yapma olasılıklarının anlamlılık düzeyinden uzaklaştığı görülmektedir. $n_2$ ve $\rho = (0.2,0.8)$ için $T_{Joh}$ ve $T_{BC}$ istatistikleri için sırasıyla .0501 ve .0495 değerleri elde edilmiştir. Ancak $n_6$ ve $\rho = (0.2,0.8)$ durumunda $T_{Joh}$ ve $T_{BC}$ istatistiklerinin I. tip hata yapma olasılıklarının sırasıyla .0520 ve .0529 değerlerine yükseldiği görülmüştür. $T_{NV}$ istatistiği, $T_{Hot}$ istatistiğine göre daha iyi performans göstermiştir. Ancak $T_{Yao}$, $T_{Joh}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$ ve $T_{MB}$ test istatistiklerine göre $T_{NV}$ istatistiğinin I. tip hata yapma olasılıklarının anlamlılık düzeyinden daha uzak olduğu belirlenmiştir.

Simülasyon sonucunda elde edilen AAD değerleri incelendiğinde ise $p = 5$ için en küçük AAD değerinin .1643 ile $T_{MB}$ istatistiğinden elde edildiği görülmektedir. Bu test istatistiğini sırasıyla $T_S$, $T_{MNV}$, $T_{BC}$, $T_{Yao}$ ve $T_{YY}$ istatistikleri takip etmektedir. $T_{Joh}$ ve $T_{NV}$ istatistikleri daha yüksek AAD değerleri ortaya koymuştur. En kötü performans ise 2.2257 ile $T_{Hot}$ test istatistiğinden elde edilmiştir.

### 4.1.3 $p = 7$ durumunda I. tip hata yapma olasılıkları

Belirlenen deneysel koşullara göre $p = 7$ için %5 $(\alpha = 0.05)$ anlamlılık düzeyinde elde edilen 1. Tip hata yapma olasılıkları Çizelge 4.3'te gösterilmiştir.

**Çizelge 4.3** $\boldsymbol{p = 7}$ için elde edilen I. tip hata yapma olasılıkları

| $n_i$ | $(\rho_1, \rho_2)$ | $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_S$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | (0.1,0.9) | .0653 | .1067 | .0560 | .0526 | .0875 | .0536 | .0537 | .0442 | .0610 |
| | (0.2,0.8) | .0494 | .0891 | .0442 | .0445 | .0625 | .0456 | .0464 | .0394 | .0510 |
| | (0.3,0.7) | .0407 | .0794 | .0396 | .0400 | .0529 | .0410 | .0430 | .0363 | .0460 |
| | (0.4,0.6) | .0378 | .0760 | .0376 | .0387 | .0491 | .0394 | .0414 | .0354 | .0445 |
| $n_2$ | (0.1,0.9) | .0528 | .0559 | .0591 | .0505 | .0669 | .0516 | .0508 | .0493 | .0511 |
| | (0.2,0.8) | .0508 | .0533 | .0527 | .0495 | .0575 | .0504 | .0500 | .0486 | .0501 |
| | (0.3,0.7) | .0493 | .0522 | .0504 | .0494 | .0530 | .0494 | .0497 | .0488 | .0498 |
| | (0.4,0.6) | .0488 | .0519 | .0497 | .0494 | .0513 | .0490 | .0496 | .0491 | .0497 |
| $n_3$ | (0.1,0.9) | .0516 | .0514 | .0564 | .0515 | .0599 | .0518 | .0514 | .0511 | .0515 |
| | (0.2,0.8) | .0514 | .0512 | .0539 | .0515 | .0557 | .0518 | .0515 | .0513 | .0515 |
| | (0.3,0.7) | **.0501** | **.0499** | .0512 | .0502 | .0517 | **.0502** | .0503 | .0501 | .0503 |
| | (0.4,0.6) | .0491 | .0488 | .0496 | .0494 | **.0501** | .0491 | .0495 | .0492 | .0496 |
| $n_4$ | (0.1,0.9) | .0588 | .0642 | .0762 | .0545 | **.1581** | .0578 | .0557 | .0510 | .0564 |
| | (0.2,0.8) | .0587 | .0592 | .0647 | .0525 | .1093 | .0557 | .0535 | .0500 | .0532 |
| | (0.3,0.7) | .0562 | .0566 | .0566 | .0510 | .0813 | .0540 | .0523 | .0495 | .0518 |
| | (0.4,0.6) | .0541 | .0547 | .0507 | .0495 | .0626 | .0522 | .0510 | .0487 | .0503 |
| $n_5$ | (0.1,0.9) | .0502 | .0541 | .0378 | .0511 | .0299 | .0508 | .0513 | .0502 | .0516 |
| | (0.2,0.8) | .0499 | .0532 | .0388 | .0504 | .0337 | .0499 | .0506 | .0450 | .0508 |
| | (0.3,0.7) | .0494 | .0537 | .0413 | .0512 | .0381 | .0507 | .0516 | .0505 | .0516 |
| | (0.4,0.6) | .0518 | .0540 | .0446 | .0515 | .0440 | .0517 | .0518 | .0507 | .0518 |
| $n_6$ | (0.1,0.9) | .0617 | .0655 | .0793 | .0553 | **.2571** | .0596 | .0562 | .0502 | .0571 |
| | (0.2,0.8) | .0651 | .0626 | .0680 | .0544 | .1664 | .0603 | .0561 | .0507 | .0557 |
| | (0.3,0.7) | **.0645** | **.0609** | .0602 | .0544 | .1111 | **.0603** | .0570 | .0520 | .0561 |
| | (0.4,0.6) | .0639 | .0592 | .0541 | .0546 | .0744 | .0596 | .0566 | .0527 | .0552 |
| $n_7$ | (0.1,0.9) | .0504 | .0526 | .0250 | .0511 | .0210 | .0507 | .0511 | .0505 | .0513 |
| | (0.2,0.8) | .0508 | .0525 | .0295 | .0503 | .0254 | .0507 | .0508 | .0496 | .0507 |
| | (0.3,0.7) | .0533 | .0533 | .0342 | .0512 | .0306 | .0525 | .0519 | .0505 | .0515 |
| | (0.4,0.6) | .0568 | .0546 | .0402 | .0512 | .0387 | .0545 | .0529 | .0504 | .0519 |
| | ***AAD*** | ***.5139*** | ***.9975*** | ***.9907*** | ***.2521*** | ***3.0671*** | ***.3825*** | ***.2804*** | ***.2457*** | ***.2639*** |
| $n_1 = (10,10)$ | | $n_3 = (50,50)$ | | | $n_5 = (30,20)$ | | $n_7 = (40,20)$ | | | |
| $n_2 = (25,25)$ | | $n_4 = (20,30)$ | | | $n_6 = (20,40)$ | | | | | |

Çizelge 4.3 incelendiğinde $p = 7$ durumunda $T_{Hot}$ istatistiği için anlamlılık düzeyine en yakın değer $\rho = (0.4,0.6)$ ve $n_3$ gözlem büyüklüğünde .0501 olarak elde edilmiştir. Ancak tüm koşullar altında en kötü performans $\rho = (0.1,0.9)$ ve $n_6$ durumunda .2571 değeri ile $T_{Hot}$ istatistiğinden elde edilmiştir. Üstelik $\rho = (0.1,0.9)$ için $n_4$ gözlem büyüklüğünde I. tip hata yapma olasılığı %15.81 iken $n_6$ gözlem büyüklüğünde %25.71'e yükseldiği görülmektedir. Dolayısıyla $T_{Hot}$ istatistiği eşit olmayan gözlem

büyüklüklerinden ve gözlem sayıları arasındaki farkın artışından çok fazla etkilenmektedir. Korelasyon katsayıları birbirine yaklaştıkça I. tip hata yapma olasılıkları anlamlılık düzeyine yaklaşmaktadır. Ancak korelasyon katsayıları arasındaki fark arttıkça, $T_{Hot}$ istatistiğinin anlamlılık düzeyinden uzaklaştığı gözlemlenmiştir. Diğerlerinde olduğu gibi Çizelge 4.3 ‘de bağımlı değişken sayısı arttıkça en kötü performansı $T_{Hot}$ test istatistiği göstermektedir.

$T_{Joh}$, $T_{Yao}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$, $T_{NV}$ ve $T_{MB}$ test istatistikleri için elde edilen olasılık değerlerinin genel olarak birbirine yakın değerler ortaya koyduğu gözlemlenmiştir. Gözlem sayılarının eşit olduğu durumlarda anlamlılık düzeyine oldukça yakın değerler elde edilmiştir. Ancak en küçük ve en büyük gözlem büyüklüğü arasındaki fark arttıkça I. tip hata yapma olasılıklarının anlamlılık düzeyinden uzaklaştığı görülmektedir. $n_3$ ve $\rho = (0.3,0.7)$ için $T_{Yao}$, $T_{Joh}$ ve $T_{BC}$ istatistikleri için sırasıyla .0501, 0.499 ve .0502 değerleri elde edilmiştir. Ancak $n_6$ ve $\rho = (0.3,0.7)$ durumunda $T_{Yao}$, $T_{Joh}$ ve $T_{BC}$ istatistiklerinin I. tip hata yapma olasılıklarının sırasıyla .0645, .0609 ve .0603 değerlerine yükseldiği görülmüştür. $T_{Joh}$ istatistiği, $T_{Hot}$ istatistiğine göre daha iyi performans göstermiştir. Ancak $T_{Yao}$, $T_{NV}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$ ve $T_{MB}$ test istatistiklerine göre $T_{Joh}$ istatistiğinin I. tip hata yapma olasılıklarının anlamlılık düzeyinden daha uzak olduğu belirlenmiştir.

Simülasyon sonucunda elde edilen AAD değerleri incelendiğinde ise $p = 7$ için en küçük AAD değerinin .2457 ile $T_{YY}$ istatistiğinden elde edildiği görülmektedir. Bu test istatistiğini sırasıyla $T_{MNV}$, $T_{MB}$, $T_S$, $T_{BC}$ ve $T_{Yao}$ istatistikleri takip etmektedir. $T_{NV}$ ve $T_{Joh}$ istatistikleri daha yüksek AAD değerleri ortaya koymuştur. En kötü performans ise 3.0671 ile $T_{Hot}$ test istatistiğinden elde edilmiştir.

### 4.1.4 $p = 9$ durumunda I. tip hata yapma olasılıkları

Belirlenen deneysel koşullara göre $p = 9$ için %5 $(\alpha = 0.05)$ anlamlılık düzeyinde elde edilen 1. Tip hata yapma olasılıkları Çizelge 4.4'de gösterilmiştir.

**Çizelge 4.4** $p = 9$ için elde edilen I. tip hata yapma olasılıkları

| $n_i$ | $(\rho_1, \rho_2)$ | $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_S$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | (0.1,0.9) | .0742 | .1858 | .0570 | .0574 | .1069 | .0531 | .0582 | .0423 | .0798 |
| | (0.2,0.8) | .0480 | .1433 | .0418 | .0450 | .0715 | .0460 | .0479 | .0362 | .0631 |
| | (0.3,0.7) | .0376 | .1275 | .0385 | .0401 | .0559 | .0429 | .0436 | .0335 | .0540 |
| | (0.4,0.6) | .0346 | .1206 | .0368 | .0376 | **.0510** | .0420 | .0422 | .0325 | .0511 |
| $n_2$ | (0.1,0.9) | .0557 | .0623 | .0619 | .0513 | .0736 | .0529 | .0519 | .0491 | .0528 |
| | (0.2,0.8) | .0525 | .0587 | .0554 | .0507 | .0606 | .0521 | .0519 | .0497 | .0521 |
| | (0.3,0.7) | .0511 | .0575 | .0519 | .0505 | .0557 | .0514 | .0518 | .0498 | .0518 |
| | (0.4,0.6) | .0482 | .0559 | .0488 | .0486 | .0512 | .0487 | .0496 | .0482 | .0495 |
| $n_3$ | (0.1,0.9) | .0518 | .0531 | .0575 | .0519 | .0617 | .0528 | .0520 | .0512 | .0522 |
| | (0.2,0.8) | .0514 | .0522 | .0549 | .0516 | .0569 | .0523 | .0519 | .0512 | .0519 |
| | (0.3,0.7) | **.0508** | **.0509** | .0518 | .0508 | .0534 | **.0509** | .0511 | .0506 | .0510 |
| | (0.4,0.6) | .0511 | .0511 | .0512 | .0511 | .0519 | .0510 | .0511 | .0509 | .0511 |
| $n_4$ | (0.1,0.9) | .0709 | .0815 | .0940 | .0575 | **.1897** | .0603 | .0578 | .0525 | .0608 |
| | (0.2,0.8) | .0677 | .0729 | .0738 | .0556 | .1244 | .0605 | .0577 | .0530 | .0587 |
| | (0.3,0.7) | .0640 | .0669 | .0612 | .0543 | .0870 | .0583 | .0561 | .0516 | .0562 |
| | (0.4,0.6) | .0588 | .0619 | .0522 | .0522 | .0652 | .0551 | .0541 | .0507 | .0535 |
| $n_5$ | (0.1,0.9) | .0482 | .0566 | .0311 | .0490 | .0266 | .0492 | .0494 | .0482 | .0500 |
| | (0.2,0.8) | .0494 | .0564 | .0342 | .0503 | .0326 | .0503 | .0511 | .0496 | .0512 |
| | (0.3,0.7) | .0495 | .0566 | .0370 | .0504 | .0380 | .0506 | .0514 | .0497 | .0513 |
| | (0.4,0.6) | .0510 | .0566 | .0399 | .0499 | .0425 | .0508 | .0512 | .0490 | .0511 |
| $n_6$ | (0.1,0.9) | .0772 | .0826 | .0937 | .0585 | **.3028** | .0615 | .0590 | .0493 | .0620 |
| | (0.2,0.8) | .0769 | .0739 | .0730 | .0561 | .1886 | .0619 | .0580 | .0506 | .0589 |
| | (0.3,0.7) | **.0750** | **.0680** | .0585 | .0535 | .1182 | **.0605** | .0562 | .0499 | .0558 |
| | (0.4,0.6) | .0685 | .0630 | .0489 | .0527 | .0753 | .0585 | .0549 | .0496 | .0543 |
| $n_7$ | (0.1,0.9) | .0489 | .0544 | .0180 | .0501 | .0170 | .0500 | .0506 | .0494 | .0508 |
| | (0.2,0.8) | .0509 | .0550 | .0211 | .0501 | .0226 | .0504 | .0506 | .0489 | .0506 |
| | (0.3,0.7) | .0531 | .0563 | .0270 | .0504 | .0275 | .0526 | .0518 | .0493 | .0509 |
| | (0.4,0.6) | .0589 | .0579 | .0328 | .0512 | .0355 | .0546 | .0529 | .0496 | .0522 |
| | ***AAD*** | ***.8825*** | ***2.2836*** | ***1.4004*** | ***.3143*** | ***3.7829*** | ***.4414*** | ***.3593*** | ***.2804*** | ***.4632*** |

$n_1 = (10,10)$ $n_3 = (50,50)$ $n_5 = (30,20)$ $n_7 = (40,20)$
$n_2 = (25,25)$ $n_4 = (20,30)$ $n_6 = (20,40)$

Çizelge 4.4 incelendiğinde $p = 9$ durumunda $T_{Hot}$ istatistiği için anlamlılık düzeyine en yakın değer $\rho = (0.4,0.6)$ ve $n_1$ gözlem büyüklüğünde .0510 olarak elde edilmiştir. Ancak tüm koşullar altında en kötü performans $\rho = (0.1,0.9)$ ve $n_6$ durumunda .3028 değeri ile $T_{Hot}$ istatistiğinden elde edilmiştir. Üstelik $\rho = (0.1,0.9)$ için $n_4$ gözlem büyüklüğünde I. tip hata yapma olasılığı %18.97 iken $n_6$ gözlem büyüklüğünde %30.28'e yükseldiği görülmektedir. Dolayısıyla $T_{Hot}$ istatistiği eşit olmayan gözlem büyüklüklerinden ve gözlem sayıları arasındaki farkın artışından çok fazla etkilenmektedir. Korelasyon katsayıları birbirine yaklaştıkça I. tip hata yapma olasılıkları anlamlılık düzeyine yaklaşmaktadır. Ancak korelasyon katsayıları arasındaki fark arttıkça, $T_{Hot}$ istatistiğinin anlamlılık düzeyinden uzaklaştığı gözlemlenmiştir. Çizelge 4.4 'de de bağımlı değişken sayısı arttıkça en kötü performansı $T_{Hot}$ test istatistiği göstermektedir.

$T_{Yao}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$, $T_{NV}$ ve $T_{MB}$ test istatistikleri için elde edilen olasılık değerlerinin genel olarak birbirine yakın değerler ortaya koyduğu gözlemlenmiştir. Gözlem sayılarının eşit olduğu durumlarda anlamlılık düzeyine oldukça yakın değerler elde edilmiştir. Ancak en küçük ve en büyük gözlem büyüklüğü arasındaki fark arttıkça I. tip hata yapma olasılıklarının anlamlılık düzeyinden uzaklaştığı görülmektedir. $n_3$ ve $\rho = (0.3,0.7)$ için $T_{Yao}$, $T_{Joh}$ ve $T_{BC}$ istatistikleri için sırasıyla .0508, .0509 ve .0509 değerleri elde edilmiştir. Ancak $n_6$ ve $\rho = (0.3,0.7)$ durumunda $T_{Yao}$, $T_{Joh}$ ve $T_{BC}$ istatistiklerinin I. tip hata yapma olasılıklarının sırasıyla .0750, .0680 ve .0605 değerlerine yükseldiği görülmüştür. $T_{Joh}$ istatistiği, $T_{Hot}$ istatistiğine göre daha iyi performans göstermiştir. Ancak $T_{Yao}$, $T_{NV}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$ ve $T_{MB}$ test istatistiklerine göre $T_{Joh}$ istatistiğinin I. tip hata yapma olasılıklarının anlamlılık düzeyinden daha uzak olduğu belirlenmiştir.

Simülasyon sonucunda elde edilen AAD değerleri incelendiğinde ise $p = 9$ için en küçük AAD değerinin .2804 ile $T_{YY}$ istatistiğinden elde edildiği görülmektedir. Bu test istatistiğini sırasıyla $T_{MNV}$ , $T_S$, $T_{BC}$, $T_{MB}$ ve $T_{Yao}$ istatistikleri takip etmektedir. $T_{NV}$ ve $T_{Joh}$ istatistikleri daha yüksek AAD değerleri ortaya koymuştur. En kötü performans ise 3.7829 ile $T_{Hot}$ test istatistiğinden elde edilmiştir.

### 4.2 Uygulama

Bu bölümde çok değişkenli iki örneklem Behrens-Fisher problemlerinin çözümü için önerilen test istatistiklerinin gerçek bir veri örneği üzerinde karşılaştırılması amaçlanmıştır. Bu amaçla Mısır'da beş farklı çağdan kalma erkek kafatası örnekleri üzerine yapılan ölçümler kullanılmıştır. Bu veri seti; hanedanlık (firavunlar) öncesi ilk dönemler (yaklaşık MÖ 4000); hanedanlık öncesi son dönemler (yaklaşık MÖ 3300); 12. ve 13. hanedanlar (yaklaşık MÖ 1850); Ptolemaios hanedanlık dönemi (yaklaşık MÖ 200) ve Roma dönemi (yaklaşık MS 150) olmak üzere beş farklı döneme ait kafatası örneklerine ilişkin gözlem verileri içermektedir. Kafatası örneklerinin her biri için "$x =$ Kafatasının maksimum genişliği (mm)", "$y =$ Kafatasının bazibregmatik yüksekliği (mm)", "$z =$ Kafatasının bazialveolar uzunluğu (mm)" ve "$t =$ Kafatasının burun yüksekliği (mm)" olmak üzere dört farklı ölçüm bilinmektedir. Ayrıca her bir dönemden 30 kafatası örneği olmak üzere toplam 150 gözlemden oluşmaktadır. Bu veri seti Thomson ve Randall-Maciver (1905) tarafından sunulmuş ve Krishnamoorthy ve Lu (2010), Zhang ve Liu (2011) ile Zhang (2012) tarafından $k$ örneklem Behrens-Fisher problemleri için uygulanmıştır. Ayrıca "skulls" veri seti, R.Studio paket programında yer alan "HSAUR2::skulls" paketi yardımıyla elde edilmiştir.

Eşit ve eşit olmayan gözlem büyüklüklerinde test istatistiklerinin performansı göstermek amacıyla her bir örneklem için farklı sayıda gözlem kullanılmıştır. Ayrıca MÖ4000-MÖ3300, MÖ3300-MÖ1850, MÖ1850-MÖ200 ve MÖ200-MS150 olmak üzere dört farklı karşılaştırma yapılmıştır. Böylece çok değişkenli iki örneklem Behrens-Fisher problemleri için önerilen test istatistikleri kullanılarak grup ortalamaları arasındaki farkların zamanla değişip değişmediğinin tespit edilmesi amaçlanmıştır.

Kullanılan veri setine ait gözlem sayıları, ortalama vektörler ve varyans-kovaryans matrisleri sırasıyla,

$$n_1 = 21 \qquad n_2 = 21 \qquad n_3 = 10 \qquad n_4 = 6 \qquad n_5 = 15$$

$$\hat{\mu}_1 = \begin{pmatrix} 131.71 \\ 134.05 \\ 98.00 \\ 50.52 \end{pmatrix} \qquad \hat{\mu}_2 = \begin{pmatrix} 132.81 \\ 133.00 \\ 99.86 \\ 49.90 \end{pmatrix} \qquad \hat{\mu}_3 = \begin{pmatrix} 133.80 \\ 136.30 \\ 95.70 \\ 49.60 \end{pmatrix} \qquad \hat{\mu}_4 = \begin{pmatrix} 136.33 \\ 129.67 \\ 94.83 \\ 50.50 \end{pmatrix}$$

$$\hat{\mu}_5 = \begin{pmatrix} 135.07 \\ 131.33 \\ 94.40 \\ 52.33 \end{pmatrix}$$

$$\hat{\Sigma}_1 = \begin{pmatrix} 29.914 & 6.614 & 3.300 & 7.407 \\ & 20.848 & -4.200 & 1.624 \\ & & 28.800 & -1.600 \\ & & & 6.762 \end{pmatrix}$$

$$\hat{\Sigma}_2 = \begin{pmatrix} 29.262 & -1.800 & 4.721 & 1.931 \\ & 24.700 & 3.200 & 8.250 \\ & & 15.629 & 2.836 \\ & & & 10.790 \end{pmatrix}$$

$$\hat{\Sigma}_3 = \begin{pmatrix} 19.733 & 3.622 & 0.711 & -3.756 \\ & 18.233 & 1.211 & 9.467 \\ & & 24.456 & 5.978 \\ & & & 14.044 \end{pmatrix}$$

$$\hat{\Sigma}_4 = \begin{pmatrix} 17.067 & -0.267 & 1.067 & 4.800 \\ & 29.067 & 13.333 & 10.400 \\ & & 52.967 & 15.300 \\ & & & 8.300 \end{pmatrix}$$

$$\hat{\Sigma}_5 = \begin{pmatrix} 30.638 & -2.167 & -0.100 & 1.190 \\ & 31.381 & 12.357 & 9.810 \\ & & 16.543 & 6.071 \\ & & & 12.952 \end{pmatrix}$$

şeklinde elde edilmiştir. Böylece iki örneklem ortalama vektörleri arasındaki farkı sınamak için ilgili örneklemlerin gözlem sayıları, ortalama vektörleri ve varyans-kovaryans matrisleri yardımıyla T test istatistiği,

$$T = (\overline{x}_A - \overline{x}_B)' \left( \frac{\hat{\Sigma}_A}{n_A} + \frac{\hat{\Sigma}_B}{n_B} \right)^{-1} (\overline{x}_A - \overline{x}_B) \tag{5.1}$$

eşitliği ile hesaplanmaktadır.

MÖ4000-MÖ3300 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından bir farklılık olup olmadığını test etmek için,

$$H_0{}^1: \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{14} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \\ \mu_{23} \\ \mu_{24} \end{pmatrix} \tag{5.2}$$

şeklindeki yokluk hipotezi oluşturulmuştur. Bu durumda MÖ4000-MÖ3300 dönemine ait kafatası örnekleri için %5 anlamlılık düzeyinde test istatistiklerine ait olasılık değerleri Çizelge 4.5'de yer almaktadır.

**Çizelge 4.5** MÖ 4000-MÖ 3300 dönemine ait gözlemler için olasılık değerleri

| $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_S$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|
| .62278 | .61620 | .62333 | .62358 | .62256 | .62418 | .62333 | .62325 | .62299 |

Çizelge 4.5 incelendiğinde bütün test istatistikleri için elde edilen olasılık değerleri (p-value) $\alpha = .05$'ten daha yüksek olduğu için yokluk hipotezi kabul edilmektedir. Bu durumda %5 anlamlılık düzeyinde MÖ4000-MÖ3300 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından anlamlı bir farklılık olmadığı söylenebilmektedir.

MÖ3300-MÖ1850 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından bir farklılık olup olmadığını test etmek için,

$$H_0{}^2: \begin{pmatrix} \mu_{21} \\ \mu_{22} \\ \mu_{23} \\ \mu_{24} \end{pmatrix} = \begin{pmatrix} \mu_{31} \\ \mu_{32} \\ \mu_{33} \\ \mu_{34} \end{pmatrix} \tag{5.3}$$

şeklindeki yokluk hipotezi oluşturulmuştur. Bu durumda MÖ3300-MÖ1850 dönemine ait kafatası örnekleri için %5 anlamlılık düzeyinde test istatistiklerine ait olasılık değerleri Çizelge 4.6'da yer almaktadır.

**Çizelge 4.6** MÖ 3300-MÖ 1850 dönemine ait gözlemler için olasılık değerleri

| $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_{S}$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|
| .13011 | .09789 | .11270 | .10773 | ***.04708*** | .09719 | .10301 | .11091 | .10651 |

Çizelge 4.6 incelendiğinde Hotelling test istatistiği haricinde bütün test istatistikleri sonucunda elde edilen olasılık değerleri (p-value) $\alpha = .05$'ten daha büyük olduğu için yokluk hipotezi kabul edilmektedir. Ancak Hotelling test istatistiği için $p = .04708 < \alpha = .05$ olduğundan yokluk hipotezi reddedilmektedir. Bu durumda Hotelling test istatistiğine göre %5 anlamlılık düzeyinde MÖ3300-MÖ1850 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından anlamlı bir farklılık olduğu söylenebilmektedir. Ancak diğer bütün test istatistiklerine göre göre %5 anlamlılık düzeyinde MÖ3300-MÖ1850 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından anlamlı bir farklılık olmadığı görülmektedir.

MÖ1850-MÖ200 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından bir farklılık olup olmadığını test etmek için,

$$H_0{}^3 : \begin{pmatrix} \mu_{31} \\ \mu_{32} \\ \mu_{33} \\ \mu_{34} \end{pmatrix} = \begin{pmatrix} \mu_{41} \\ \mu_{42} \\ \mu_{43} \\ \mu_{44} \end{pmatrix} \tag{5.4}$$

şeklindeki yokluk hipotezi oluşturulmuştur. Bu durumda MÖ1850-MÖ200 dönemine ait kafatası örnekleri için %5 anlamlılık düzeyinde test istatistiklerine ait olasılık değerleri Çizelge 4.7'de yer almaktadır.

**Çizelge 4.7** MÖ 1850-MÖ 200 dönemine ait gözlemler için olasılık değerleri

| $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_{S}$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|
| .13106 | .06520 | .13442 | .09217 | ***.03526*** | .08011 | .08507 | .09964 | .08786 |

Çizelge 4.7 incelendiğinde Hotelling test istatistiği haricinde bütün test istatistikleri sonucunda elde edilen olasılık değerleri (p-value) $\alpha = .05$'ten daha büyük olduğu için yokluk hipotezi kabul edilmektedir. Ancak Hotelling test istatistiği için $p = .03526 < \alpha = .05$ olduğundan yokluk hipotezi reddedilmektedir. Bu durumda Hotelling test istatistiğine göre %5 anlamlılık düzeyinde MÖ1850-MÖ200 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından anlamlı bir farklılık olduğu söylenebilmektedir. Ancak diğer bütün test istatistiklerine göre göre %5 anlamlılık düzeyinde MÖ1850-MÖ200 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından anlamlı bir farklılık olmadığı görülmektedir.

MÖ200-MS150 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından bir farklılık olup olmadığını test etmek için,

$$H_0{}^4 : \begin{pmatrix} \mu_{41} \\ \mu_{42} \\ \mu_{43} \\ \mu_{44} \end{pmatrix} = \begin{pmatrix} \mu_{51} \\ \mu_{52} \\ \mu_{53} \\ \mu_{54} \end{pmatrix} \tag{5.5}$$

şeklindeki yokluk hipotezi oluşturulmuştur. Bu durumda MÖ200-MS150 dönemine ait kafatası örnekleri için %5 anlamlılık düzeyinde test istatistiklerine ait olasılık değerleri Çizelge 4.8'de yer almaktadır.

**Çizelge 4.8** MÖ 200-MS 150 dönemine ait gözlemler için olasılık değerleri

| $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_S$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|
| .47032 | .49850 | .61782 | .54799 | .73223 | .51283 | .53441 | .54982 | .54108 |

Çizelge 4.8 incelendiğinde bütün test istatistikleri sonucunda elde edilen olasılık değerleri (p-value) $\alpha = .05$'ten daha büyük olduğu için yokluk hipotezi kabul edilmektedir. Bu durumda %5 anlamlılık düzeyinde MÖ200-MS150 dönemine ait kafatası örnekleri arasında dört farklı ölçüm bakımından anlamlı bir farklılık olmadığı söylenebilmektedir.

# 5. SONUÇ VE ÖNERİLER

Bu çalışmada, varyans-kovaryans matrisinin heterojenliği altında, iki grup ortalama vektörlerinin eşitliği için kullanılan test istatistikleri deneysel olarak I.tip hata yapma olasılıkları bakımından karşılaştırılmıştır.

Simülasyon sonuçları incelendiğinde Çizelge 4.1, Çizelge 4.2, Çizelge 4.3 ve Çizelge 4.4'de $T_{Hot}$ dayalı test istatistiklerinin deneysel I.tip hata değerlerinin, ilgili anlamlılık düzeyinde (α=0,05) en uzak değerler aldığı görülmektedir. Ayrıca, $T_{Hot}$'e dayalı test istatistikleri özellikle bağımlı değişken sayısı arttıkça ve varyans-kovaryans matrisleri farklılaştıkça, kötü bir istatistiksel sonuç elde edildiği görülmektedir.

Genel olarak tüm testlerin deneysel I.tip hata oranlarını incelendiğinde, örnek çapı eşit ve küçükken, heterojenlik büyükken testlerinin deneysel I.tip hata yapma oranları α'dan büyük çıkmıştır. Heterojenlik azaldığında ise, deneysel I.tip hata yapma oranları α'ya daha yakın çıkmıştır. Değişken sayısı arttığında ise, bu testlerin bundan oldukça olumsuz etkilenerek α'dan daha da uzaklaştığı görülmektedir. $T_{Yao}$, $T_{NV}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$ ve $T_{MB}$ testlerinin ise tüm bu durumlarda deneysel I.tip hata yapma oranları α'ya yakın çıkmıştır. $T_{Hot}$ istatistiği eşit olmayan gözlem büyüklüklerinden ve gözlem sayıları arasındaki farkın artışından çok fazla etkilenmiş ve en kötü performansı göstermiştir. Örnek çapı eşit ve arttığında ise tüm testlerin deneysel I.tip hata oranları α'ya yaklaşmıştır. Korelasyon katsayıları birbirine yaklaştıkça I. tip hata yapma olasılıkları anlamlılık düzeyine yaklaşmaktadır. Ancak korelasyon katsayıları arasındaki fark arttıkça genel olarak, $T_{Hot}$ istatistiğinin nominal (α=0,05) uzaklaştığı gözlemlenmiştir. Korelasyon değerinin artışı testleri etkilerken, aynı zamanda özellikle değişken sayısının artışının da testleri etkilediği görülmektedir.

Bütün sonuçlar göz önüne alındığında, heterojenlik altında simülasyon sonucunda elde edilen deneysel I. tip hata yapma olasılıklarına ($\hat{\alpha}$) göre, test istatistiklerinin performanslarını belirlemek için kullanılan "Ortalama Mutlak Tutarsızlık (AAD)" değeri, yukarıdaki her bir çizelgenin son satırında hesaplanmıştır. Çizelge 4.1, Çizelge 4.2, Çizelge 4.3 ve Çizelge 4.4 için hesaplanan AAD değerleri Çizelge 5.1'de topluca gösterilmiştir.

**Çizelge 5.1** Tüm olasılık değerlerindeki ortalama mutlak tutarsızlık (AAD) değerleri

| **AAD /** | $T_{Yao}$ | $T_{Joh}$ | $T_{NV}$ | $T_{MNV}$ | $T_{Hot}$ | $T_{BC}$ | $T_{S}$ | $T_{YY}$ | $T_{MB}$ |
|---|---|---|---|---|---|---|---|---|---|
| **p=3** | .1157 | .1139 | .2600 | .1171 | 1.4232 | .1475 | .1214 | .1346 | .1161 |
| **p=5** | .2438 | .3002 | .6000 | .1854 | 2.2257 | .2191 | .1707 | .2673 | .1643 |
| **p=7** | .5139 | .9975 | .9907 | .2521 | 3.0671 | .3825 | .2804 | .2457 | .2639 |
| **p=9** | .8825 | 2.2836 | 1.4004 | .3143 | 3.7829 | .4414 | .3593 | .2804 | .4632 |

Çizelge 5.1 incelendiğinde Yao, MNV, BC, S, YY test istatistikleri sonucunda elde edilen 4 olasılık değerinde de (p-value), $\alpha = 0{,}05$ durumunda AAD değerlerinin aykırı bir değere sahip olmadığı görülmektedir.

P=9 durumunda Johansen, Nel Van Der Merwe ve Hotelling test istatistiklerinin daha kötü sonuçlar verdiği görülmektedir. Literatürde bulunan çalışmalardaki gibi küçük örneklem değerlerinde Johansen testinin iyi bir test istatistiği vermediği bu çalışmada da görülmüştür. Ayrıca elde edilen sonuçlardan NV test istatistiği için, Modifiye Edilmiş NV test istatistiğinin gerekli olduğu bir kez daha doğrulamıştır.

Tüm olasılık değerleri göz önüne alındığında da heterojenlik durumunda Hotelling test istatistiğinin çok iyi performansa sahip olmadığı görülmüştür.

Sonuç olarak AAD değerleri için;

Olasılık değerleri p=3 den p=9 değerine doğru gittikçe, düşük deneysel hata oranlarının $\alpha = 0{,}05$'e yaklaştığı, yüksek deneysel hata oranlarının da $\alpha = 0{,}05$'den uzaklaştığı görülmektedir. Dolayısıyla p(olasılık değeri) arttıkça kötü sonuç veren testlerin oranı artmaktadır.

Sonuç olarak, $\boldsymbol{T_{Hot}}$'e dayalı test istatistiği, özellikle küçük örnek çaplarında birçok durumda iyi sonuçlar veren $T_{Yao}$, $T_{NV}$, $T_{MNV}$, $T_{BC}$, $T_{S}$, $T_{YY}$ ve $T_{MB}$ testlerine alternatif olarak kullanılmamalıdır.

Bu çalışmada incelenen $T_{Yao}$, $T_{NV}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$ ve $T_{MB}$ testlerinde birbirine yakın değerler elde edildiği görülmüştür. Bu tip analizlerde $\boldsymbol{T_{Hot}}$ test istatistiği kullanımı yerine, diğer test istatistiklerinin ($T_{Yao}$, $T_{NV}$, $T_{MNV}$, $T_{BC}$, $T_S$, $T_{YY}$ ve $T_{MB}$ ) kullanılması önerilmektedir.

# KAYNAKLAR DİZİNİ

Bartlett, M.S., 1937, Properties of sufficiency and statistical tests. Proceedings of Royal Society of London. Series A, Mathematical and Physical Sciences,160(901): 268-282.

Bartlett, M.S., 1939, A note on tests of significance in multivariate analysis, Mathematical Proceedings of the Cambridge Philosophical Society, 35(2): 180–185.

Behrens, W.V., 1929, Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtungen (A contribution to error estimation with few observations), Landwirtschaftliches Jahrbuch, 68: 807-837.

Bennett, B.M., 1951, Note on a solution of the generalized Behrens–Fisher problem, Annals of the Institute Statistical Mathematics, 2, 87–90.

Çavuş, M., 2016, Aykırı değer durumunda genelleştirilmiş Behrens-Fisher problemi için düzeltilmiş testler, Yüksek Lisans Tezi, Anadolu Üniversitesi, 129 s.

Erdoğan, S., 2018, Heterojenlik altında iki grup ortalama vektörlerinin karşılaştırılması için önerilen yeni bir hesaplamalı yaklaşım testi, Yüksek Lisans Tezi, Gazi Üniversitesi, 78 s.

Fisher, R.A., 1935, The fiducial argument in statistical inference, Annals of Eugenics, 6(4): 391-398.

Gamage, J., Mathew, T., Weerahandi, S., 2004, Generalized p-values and generalized confidence regions for the multivariate Behrens-Fisher problem and MANOVA. Journal of Multivariate Analysis, 88(1): 177-189.

Gokpınar E .,Karanfıl S.,Ebegıl M. ,Ozdemir Y. And Gokpınar F(2015). A Computational Approach Test for the Equality ofTwo Multivariate Normal Mean Vectors under Heterogeneity of Covariance Matrices, Hacettepe Journal ofMathematics and Statistics,

Hotelling, H., 1951, A generalized T test and measure of multivariate dispersion, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 23-41.

Johansen, S., 1980, The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression, Biometrika, 67(1): 85-92.

**KAYNAKLAR DİZİNİ (devam)**


Kanık, E.A., 1999, Çok değişkenli varyans analizinde kovaryans matrislerinin homojenliği ön şartı, Doktora Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, 126 s.

Kawasaki, T., Seo, T., 2014, A two sample test for mean vectors with unequal covariance matrices, Communications in Statistics: Simulation and Computation, 44(7): 1850-1866.

Kim, S., 1992, A practical solution to the multivariate Behrens- Fisher problem, Biometrika, 79(1): 171-176

Krishnamoorthy, K., Yu, J., 2004, Modified Nel and Van der Merwe test for the multivariate Behrens–Fisher problem, Statistics and Probability Letters, 66: 161–169.

Krishnamoorthy, K., Lu, F., Mathew, T., 2007, A parametric bootstrap approach for ANOVA with unequal variances: Fixed and random models, Computational Statistics and Data Analysis, 51(12): 5731-5742.

Krishnamoorthy, K., Lu, F., 2010, A parametric bootstrap solution to the MANOVA under heteroscedasticity, Journal of Statistical Computation and Simulation, 80(8): 873-887.

Nel, D.G., Van der Merwe, C.A., 1986, A Solution to the Multivariate Behrens-Fisher Problem, Communication Statistics-Theory and Methods, 15(12): 3719-3735.

Oyeyemi, G.M., Adebayo, P.O., Adeleke, B.L., 2018, Heteroscedasticity in one way multivariate analysis of variance, Journal of Özdamar, K., 2018, Eğitim, Sağlık ve Sosyal Bilimler İçin SPSS Uygulamalı Temel İstatistik, Nisan Kitabevi, s.116. Physical Mathematics, 9(2): 1-5.

Özkip, E., Yazıcı, B., Sezer, A., 2014, A simulation study on tests for the Behrens-Fisher problem, Türkiye Klinikleri Journal of Biostatistics, 6(2): 59-66.

Satterthwaite, F. E. (1946), "An approximate distribution of estimates of variance components.", biometrics Bulletin, 2 (6): 110–114, doi:10.2307/3002019, JSTOR 3002019, PMID 20287815

## KAYNAKLAR DİZİNİ (devam)

Taysı,R., Çelik, Ş., 2018,Homojen olmayan varyans varsayımı altında ortalamaların eşitliği için Brown-Forstye ve Welch istatistiklerinin mısır verimi örneğine uygulanması, Fırat Üniversitesi Fen Bilimleri Dergisi 30(1),23-27,2018 30(1), 23-327,2018

Thomson, A. and Randall-Maciver, R. (1905) Ancient Races of the Thebaid, Oxford: Oxford University Press.

Tsui, K.W., Weerahandi, S., 1989, Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters, Journal of the American Statistical Association, 84(406): 602-607.

Yao, Y., 1965, An approximate degrees of freedom solution to the Multivariate Behrens-Fisher problem, Biometrika, 52(1-2): 139-147.

Yiğit, E., 2009, Homojen olmayan varyans varsayımı altında ortalamaların eşitliği için test istatistikleri, Gazi ÜniversitesZhang, J.T., Xu, J., 2009, On the k-sample Behrens-Fisher problem for high-dimensional data, Science in China Series A:Mathematics, 52 (6): 1285-1304. i Fen Bilimleri Enstitüsü, 109 s.

Zhang, J.T., Liu, X., 2011, A modified Bartlett test for heteroscedastic one-way MANOVA, Metrika, 76: 135–152.

Zhang, J.T., 2012, An approximate Hotelling $T^2$-test for heteroscedastic one-way MANOVA, Open Journal of Statistics, 2(1): 1-11.