

Atf için / For Citation: M. Sandal, Z. Yıldız, "Yüksek Boyutlu Veriler İçin Behrens-Fisher Probleminin Çözümü ve Önerilen Test İstatistiklerinin Karşılaştırılması", *Süleyman Demirel Üniversitesi Fen Edebiyat Fakültesi Fen Dergisi*, 16(2), 397-415, 2021.



Yüksek Boyutlu Veriler İçin Behrens-Fisher Probleminin Çözümü ve Önerilen Test İstatistiklerinin Karşılaştırılması

Mehmet SANDAL*¹, Zeki YILDIZ²

¹Manisa Celal Bayar Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, 45000, Eskişehir, Türkiye

²Eskişehir Osmangazi Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, 26000, Eskişehir, Türkiye
*yazışılan yazar e-posta: mehmet.sandal@cbu.edu.tr

(Alınış: 05.03.2021, Kabul: 17.07.2021, Yayımlanma: 25.11.2021)

Özet: Çok değişkenli ikiden fazla grup ortalama vektörünün karşılaştırılması genellikle Çok Değişkenli Varyans Analizi (MANOVA) problemi olarak bilinmektedir. Ancak MANOVA problemlerinin çözümü için kullanılan klasik test istatistikleri, varsayım ihlallerinden oldukça fazla etkilenmektedir. Ayrıca bağımlı değişken sayısının gözlem sayısından daha büyük olduğu durumlarda çok değişkenli test istatistikleri yetersiz kalmaktadır. Bu çalışmanın amacı, yüksek boyutlu Behrens-Fisher problemlerinde ortalama vektörlerin eşit olup olmadığını belirlemek için önerilen bazı test istatistiklerini karşılaştırmaktır. Bu amaçla varyans-kovaryans matrislerinin heterojen olduğu durumlarda dört farklı test istatistiği için bir simülasyon analizi gerçekleştirilmiştir. Çalışmada üç farklı varyans-kovaryans modeli kullanılmıştır. Ayrıca bağımlı değişken sayısı ve gözlem sayısı için farklı deneysel koşullar dikkate alınmıştır. Çalışmanın sonuçları, test istatistiklerinin performanslarının genel olarak karşılaştırılabilir olduğunu göstermiştir. Ancak deneysel koşullara göre test istatistiklerinin performansının değiştiği gözlemlenmiştir.

Anahtar kelimeler: Çok değişkenli varyans analizi (MANOVA), Yüksek boyutlu veri, Behrens-Fisher problemi, Monte Carlo simülasyonu

Solution of Behrens-Fisher Problem For High-Dimensional Data and Comparison of Proposed Test Statistics

Abstract: Comparison of multivariate mean of more than two groups is generally known as multivariate variance analysis (MANOVA) problem. However, classical test statistics used to solve MANOVA problems are highly affected by assumption violations. In addition, multivariate test statistics are insufficient in cases where the number of dependent variables is greater than the number of observations. This study aims to compare some proposed test statistics to determine whether mean vectors are equal in high-dimensional Behrens-Fisher problems. For this purpose, a simulation analysis was carried out for four different test statistics in cases where variance-covariance matrices are heterogeneous. Three different variance-covariance models were used in the study. In addition, different experimental conditions were taken into account for the number of dependent variables and the number of observations. The results of the study showed that the performances of the test statistics were generally comparable. However, it has been observed that the performance of test statistics varies according to experimental conditions.

Key words: Multivariate variance analysis (MANOVA), High dimensional data, Behrens-Fisher problem, Monte Carlo simulation

Not: Bu makale doktora tez çalışmasından üretilmiştir.

Mehmet SANDAL, Mehmet.sandal@cbu.edu.tr, ORCID: <https://orcid.org/0000-0000-0000-0000>

Zeki YILDIZ, zyildiz@ogu.edu.tr, ORCID: <https://orcid.org/0000-0000-0000-0000>

1. Giriş

İstatistiksel araştırmalarda ilgilenilen temel konulardan bir tanesi ortalama karşılaştırma problemidir. İki ya da daha fazla grup ortalamasının eşit olup olmadığını belirlemek için çeşitli istatistiksel tekniklere başvurulmaktadır. Ancak araştırma probleminin çözümünde bağımlı ve bağımsız değişken sayısına göre farklı istatistiksel yöntemlerin kullanılması gerekmektedir. Tek bir bağımlı değişken mevcut iken ikiden fazla grup ortalamasını karşılaştırmak için “Varyans Analizi (ANOVA)” yöntemi kullanılmaktadır. Ancak bağımlı değişken sayısı birden fazla olduğunda, değişkenler arasındaki ilişkilerinde dikkate alınması gerekmektedir. Üstelik doğru olan yokluk hipotezinin reddedilmesi durumunda I. tip hata yapılmaktadır. Her bir değişken için ayrı ayrı ANOVA uygulanması, değişkenler arasındaki ilişkilerin göz ardı edilmesine ve I. tip hata yapma olasılığının artmasına neden olmaktadır. Bu durumda testin güven düzeyi azalmaktadır. Dolayısıyla birden fazla bağımlı değişken için çok değişkenli istatistiksel tekniklerin kullanılması gerekmektedir [1-2].

Çok değişkenli normal dağılıma sahip ikiden fazla grup ortalama vektörünün eşitlik sınaması genellikle “Çok Değişkenli Varyans Analizi (MANOVA)” problemi olarak tanımlanmaktadır. MANOVA’da, iki veya daha fazla bağımlı değişkenin aynı modelde değerlendirilmesi amaçlanmaktadır. Böylece değişkenler arasındaki ilişkiler göz ardı edilmemektedir. MANOVA problemlerinin çözümü için genellikle [3-6] tarafından önerilen test istatistikleri kullanılmaktadır. Bu test istatistikleri, gruplar arasındaki anlamlı farklılıkları belirlemek için son derece önemlidir. Ancak test istatistiklerinin uygulanabilmesi için çok değişkenli normal dağılım ve varyans-kovaryans matrislerinin homojenlik varsayımlarının sağlanması gerekmektedir. Bu koşullar sağlandığı zaman son derece güçlü ve anlamlı sonuçlar elde edilebilmektedir. Varyans-kovaryans matrislerinin homojen olmadığı durumlarda ise Behrens-Fisher problemi ile karşılaşılmaktadır. Bu problemin çözümünde [3-6] tarafından sunulan test istatistiklerinin kullanılması, bu istatistiklerin daha kötü bir performans ortaya koymasına neden olmaktadır.

MANOVA test istatistiklerinin hesaplanmasında bir başka problem ise varyans-kovaryans matrislerinin tersine ihtiyaç duyulmasıdır. Varyans-kovaryans matrislerinin tersinin alınabilmesi için bağımlı değişken sayısının gözlem sayısından daha küçük olması gerekmektedir. Ancak bağımlı değişken sayısının gözlem sayısından daha büyük olduğu yüksek boyutlu veriler söz konusu olduğunda varyans-kovaryans matrisi tekil olmakta ve tersi alınamamaktadır [7]. Dolayısıyla yüksek boyutlu veriler söz konusu olduğunda da bilinen test istatistiklerinin özellikleri ve kullanılabilirlikleri kısıtlanmaktadır.

Genetik bilimi başta olmak üzere tıp, biyoloji, endüstri, mühendislik vb. gibi birçok alanda yüksek boyutlu veriler ile sıklıkla karşılaşılmaktadır. Özellikle DNA mikrodizini üzerine yapılan araştırmalarda, daha küçük gözlemlerden elde edilen veriler için yüksek sayıdaki değişkene bağlı analizler gerçekleştirilmektedir. Son yıllarda hızlı bir gelişme gösteren teknoloji dünyası da yüksek boyutlu verilere ulaşılmasına ve bunların çözümlenmesine imkân sağlamaktadır. Bu durum yüksek boyutlu verilerin istatistiksel çalışmalarındaki önemini de giderek artırmaktadır. Ancak boyut sayısının gözlem sayısından daha fazla olması, çeşitli hesaplama ve istatistiksel çözümleme zorlukları ortaya çıkarmaktadır [8]. Bu nedenle yüksek boyutlu verilerin analiz edilmesi için uygun istatistiksel tekniklerin geliştirilmesine ihtiyaç duyulmaktadır.

Yüksek boyutlu iki örneklem ve k örneklem problemleri için varsayımların sağlandığı veya sağlanmadığı durumlara göre farklı çözümleme yaklaşımları benimsenmiştir. İki örneklem problemleri için ilk çalışmalar [9-12] tarafından gerçekleştirilmiştir. Yüksek

boyutlu MANOVA problemlerinde normallik ve homojenlik varsayımlarının farklı durumları için [13-19] tarafından çeşitli çözümlene yaklaşımları geliştirilmiştir. Ancak homojenlik varsayımının ihlal edildiği durumlarda yüksek boyutlu problemler için sınırlı sayıda test istatistiği önerilmiştir.

Bu çalışmanın amacı da yüksek boyutlu Behrens-Fisher problemlerinde ikiden fazla ($k > 2$) gruba ait ortalama vektörlerin eşitliğini sınamak için önerilen test istatistiklerini I. tip hata yapma olasılıkları ve güç değerleri bakımından karşılaştırmaktır. Bu amaçla çalışmanın üçüncü bölümünde, yüksek boyutlu veriler için önerilen test istatistikleri açıklanmıştır. Dördüncü bölümde, önerilen test istatistiklerinin farklı deneysel koşullar altında performanslarını belirlemek için bir simülasyon çalışması gerçekleştirilmiştir. Beşinci bölümde ise yüksek boyutlu k örneklem Behrens-Fisher problemlerinin çözümü için önerilen test istatistikleri gerçek bir veri örneği üzerinde uygulanmıştır.

2. Materyal ve Metot

μ_i ortalama vektörüne ve Σ_i varyans-kovaryans matrisine sahip p boyutlu normal dağılım $N_p(\mu_i, \Sigma_i)$ olmak üzere $i = 1, 2, \dots, k$ için k tane bağımsız normal dağılıma sahip örneklem $X_{i1}, X_{i2}, \dots, X_{in_i} \sim N_p(\mu_i, \Sigma_i)$ olsun. Bu durumda p tane bağımlı değişkene sahip k anakütle ortalama vektörünün eşitliği

$$H_0: \mu_{1(p \times 1)} = \mu_{2(p \times 1)} = \dots = \mu_{k(p \times 1)} \quad (1)$$

$$H_1: \mu_i' \text{lerden en az biri farklıdır.}$$

biçimindeki hipotez ile ifade edilmektedir. Denklem 1'deki yokluk hipotezini sınamak için

$$W = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(X_{ij} - \bar{X}_i)' = \sum_{i=1}^k (n_i - 1) \hat{\Sigma}_i \quad (2)$$

$$B = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})' \quad (3)$$

biçimindeki grup içi ve gruplar arası kareler ve çarpımlar toplamı matrisleri dikkate alınmaktadır. Burada her bir gruba ait gözlem sayısı n_i olmak üzere i . örneklemin $p \times 1$ boyutlu ortalamalar vektörü \bar{X}_i ve genel ortalamalar vektörü ise \bar{X} ile gösterilmektedir. Ayrıca i . örneklemin varyans-kovaryans matrisi

$$\hat{\Sigma}_i = \frac{1}{(n_i - 1)} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(X_{ij} - \bar{X}_i)' \quad (4)$$

eşitliğindeki gibi hesaplanmaktadır. Böylece [3-6] tarafından önerilen test istatistikleri kullanılarak Denklem 1'de verilen yokluk hipotezi sınanmaktadır. MANOVA problemleri için gerekli olan varsayımlar sağlandığında ve bağımlı değişken sayısı gözlem sayısından daha küçük olduğu durumlarda [3-6] tarafından önerilen test istatistikleri oldukça güçlü testlerdir. Ancak yüksek boyutlu veriler söz konusu

olduğunda ise MANOVA test istatistikleri kullanılamamaktadır. Bu nedenle yeni yaklaşımların dikkate alınması gerekmektedir.

3. Alternatif Test İstatistikleri

3.1 Yamada ve Himeno (2015) testi

Yüksek boyutlu verilerde Denklem 1’de verilen yokluk hipotezini sınamak için [20] tarafından yeni bir test istatistiği önerilmiştir. [20], yüksek boyutlu verilerde homojenlik varsayımına ihtiyaç duymadan k ortalamalar vektörünün eşitliğini test etmek için [16] tarafından sunulan test istatistiğini

$$T_{YH} = tr(B) - \sum_{i=1}^k \left(1 - \frac{n_i}{N}\right) tr \hat{\Sigma}_i \quad (5)$$

biçiminde yeniden tanımlamıştır. Böylece [20], Denklem 5’teki test istatistiğinin

$$T_{YH}^* = \frac{T_{YH}}{\hat{\sigma}_{YH}} \xrightarrow{d} N(0,1) \quad (6)$$

biçiminde standart normal dağılıma yakınsadığını ifade etmektedir. Bu durumda $tr(\widehat{\Sigma}_i \widehat{\Sigma}_j) = tr(\hat{\Sigma}_i \hat{\Sigma}_j)$ ve

$$\theta_i = \frac{1}{(n_i - 1)} \sum_{j=1}^{n_i} \left((X_{ij} - \bar{X}_i)' (X_{ij} - \bar{X}_i) \right)^2 \quad (7)$$

için

$$\widehat{tr(\Sigma_i^2)} = \frac{(n_i - 1)}{n_i(n_i - 2)(n_i - 3)} \left[\{(n_i - 1)(n_i - 2)tr \hat{\Sigma}_i^2\} + (tr \hat{\Sigma}_i)^2 - n_i \theta_i \right] \quad (8)$$

olmak üzere σ_{YH}^2 ’in yansız bir tahmincisi

$$\hat{\sigma}_{YH}^2 = 2 \sum_{i=1}^k \left(1 - \frac{n_i}{N}\right)^2 \frac{n_i}{(n_i - 1)} \widehat{tr(\Sigma_i^2)} + 2 \sum_{i \neq j}^k \frac{n_i n_j}{N^2} tr(\widehat{\Sigma}_i \widehat{\Sigma}_j) \quad (9)$$

şeklindedir. Burada $N = \sum_{i=1}^k n_i$ olarak hesaplanmaktadır. Dolayısıyla α anlamlılık düzeyi için T_{YH}^* test istatistiği standart normal dağılımın kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir [20].

3.2 Zhou’ nun (2016) L^2 norm testi

Heterojenlik altında yüksek boyutlu MANOVA problemlerine ilişkin bir başka yaklaşım ise [21] tarafından sunulmuştur. [21], homojenlik varsayımı olmadan yüksek boyutlu ortalamalar vektörlerinin eşitliğini test etmek için L^2 -norm yaklaşımına dayalı

bir test istatistiği ortaya koymuştur. Öklid normu olarak da bilinen L^2 norm, iki nokta arasındaki en kısa uzaklığı ifade etmektedir. Bu durumda tek yönlü MANOVA probleminde normal dağılıma sahip k grup için L^2 norm uzaklığına dayalı bir test istatistiği

$$T_Z^* = \sum_{i=1}^k n_i \|\bar{X}_i - \bar{X}\|^2 \quad (10)$$

biçiminde tanımlanabilmektedir. Bu test istatistiği, ilgilenilen k örneklemin ortalamalar vektörü ile bütün örneklemelerin ortalamalar vektörü arasındaki L^2 uzaklığına bağlı bir test istatistiğidir. Denklem 10'daki test istatistiğinin olasılık dağılımına yakınsamak için Welch-Satterthwaite ki-kare yaklaşımı kullanılmıştır. Dolayısıyla $\hat{\beta}\chi_{\hat{d}}^2$ değeri, T_Z^* test istatistiği için kritik değer olarak belirlenmektedir. Bu durumda $i, j = 1, \dots, k$ için

$$\begin{cases} a_{ii} = 1 - \frac{n_i}{N} & i = j \\ a_{ij} = -\frac{\sqrt{n_i n_j}}{N} & i \neq j \end{cases} \quad (11)$$

olmak üzere $\hat{\beta}$ ve \hat{d} parametreleri sırasıyla

$$\hat{\beta} = \frac{\sum_{i=1}^k a_{ii}^2 \left[\frac{(n_i-1)^2}{(n_i-2)(n_i+1)} \left(\text{tr} \hat{\Sigma}_i^2 - \frac{1}{(n_i-1)} (\text{tr} \hat{\Sigma}_i)^2 \right) \right] + 2 \sum_{i \neq j} a_{ij}^2 \text{tr}(\hat{\Sigma}_i \hat{\Sigma}_j)}{\sum_{i=1}^k a_{ii} \text{tr} \hat{\Sigma}_i} \quad (12)$$

ve

$$\hat{d} = \frac{\sum_{i=1}^k a_{ii}^2 \left[\frac{n_i(n_i-1)}{(n_i-2)(n_i+1)} \left((\text{tr} \hat{\Sigma}_i)^2 - \frac{2}{n_i} \text{tr} \hat{\Sigma}_i^2 \right) \right] + 2 \sum_{i \neq j} a_{ii} a_{jj} (\text{tr} \hat{\Sigma}_i) (\text{tr} \hat{\Sigma}_j)}{\sum_{i=1}^k a_{ii}^2 \left[\frac{(n_i-1)^2}{(n_i-2)(n_i+1)} \left(\text{tr} \hat{\Sigma}_i^2 - \frac{1}{(n_i-1)} (\text{tr} \hat{\Sigma}_i)^2 \right) \right] + 2 \sum_{i \neq j} a_{ij}^2 \text{tr}(\hat{\Sigma}_i \hat{\Sigma}_j)} \quad (13)$$

şeklinde hesaplanmaktadır. Böylece hesaplanan T_Z^* test istatistiği $\hat{\beta}\chi_{\hat{d}}^2$ kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir [21].

3.3 Hu vd.nin (2017) testi

Yüksek boyutlu verilerde ve homojenlik varsayımının sağlanmadığı durumlarda [12] tarafından önerilen test istatistiği, ikiden fazla grup ortalama vektörü için [22] tarafından genellenmiştir. İki den fazla ortalama vektörün karşılaştırılması için Öklid uzaklık ölçüsüne dayalı olarak oluşturulan test istatistiği

$$T_{HBWW} = \sum_{i < j}^k (\bar{X}_i - \bar{X}_j)' (\bar{X}_i - \bar{X}_j) - (k-1) \sum_{i=1}^k \frac{\text{tr} \hat{\Sigma}_i}{n_i} \quad (14)$$

biçiminde tanımlanmaktadır. [22]'e göre Denklem 14'deki T_{HBWW} test istatistiği

$$\frac{T_{HBWW} - E(T_{HBWW})}{\sqrt{Var(T_{HBWW})}} \xrightarrow{d} N(0,1) \quad (15)$$

olacak şekilde normal dağılıma yakınsamaktadır. Burada $E(T_{HBWW})$ değeri

$$E(T_{HBWW}) = \sum_{i < j}^k \|\mu_i - \mu_j\|^2 \quad (16)$$

şekindedir. [12]'nin test istatistiğine paralel olarak T_{HBWW} istatistiğinin tanımlanması için $Var(T_{HBWW})$ 'ın tahmin edilmesi gerekmektedir. Bu durumda $p \rightarrow \infty$ ve $N \rightarrow \infty$ için

$$\widehat{tr\Sigma_i^2} = \frac{(n_i - 1)^2}{(n_i + 1)(n_i - 2)} \left(tr\Sigma_i^2 - \frac{1}{n_i - 1} (tr\Sigma_i)^2 \right) \quad (17)$$

$$tr(\widehat{\Sigma_i\Sigma_j}) = tr\hat{\Sigma}_i\hat{\Sigma}_j \quad (18)$$

olmak üzere minimum varyanslı yansız tahmincisi (UMVUE)

$$(\hat{\sigma}_{HBWW})^2 = \sum_{i=1}^k \frac{2(k-1)^2}{n_i(n_i-1)} \widehat{tr\Sigma_i^2} + \sum_{i < j}^k \frac{4}{n_i n_j} tr(\widehat{\Sigma_i\Sigma_j}), \quad i \neq j \in \{1, \dots, k\} \quad (19)$$

şekindedir. Dolayısıyla yokluk hipotezi ve bütün varsayımlar altında $p \rightarrow \infty$ ve $N \rightarrow \infty$ için test istatistiği

$$T_{HBWW}^* = \frac{T_{HBWW}}{\hat{\sigma}_{HBWW}} \rightarrow N(0,1) \quad (20)$$

olmaktadır. Denklem 20'deki T_{HBWW}^* test istatistiği, α anlamlılık düzeyine göre standart normal dağılımın kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir. Ayrıca $k = 2$ durumunda T_{HBWW} test istatistiği [12]'nin iki değişkenli test istatistiğine dönüşmektedir [22].

3.4 Cao vd.nin (2019) testi

[12], varyans- kovaryans matrislerinin heterojenliği altında iki ortalama vektör arasındaki farkın eşitliğini test etmek için [11]'in sunduğu test istatistiğinden $\sum_{j=1}^{n_i} X'_{ij}X_{ij}$ terimini çıkartarak yeni bir test istatistiği ortaya koymuştur. [23]'ün test istatistiği ise [12]'nin bir terim çıkarma düşüncesinin [16] tarafından sunulan test istatistiğine uygulanmasına dayanmaktadır.

[16]'nın test istatistiğine [12]'nin $X'_{it}X_{it}$ terimini çıkarma düşüncesi uygulandığında, [16]'nın test istatistiği

$$T_{CPH} = \sum_{i=1}^k \frac{N - n_i}{N(n_i - 1)} \sum_{t \neq s}^{n_i} X'_{it} X_{is} - \sum_{i \neq j}^k \frac{n_i n_j}{N} \bar{X}'_i \bar{X}_j \quad (21)$$

şeklinde yeniden yazılabilmektedir. Bu durumda test istatistiğinin asimptotik dağılımına yakınsamak için standart sapmasının tutarlı bir tahmincisinin belirlenmesi gerekmektedir. [23] tarafından [24] ile [22]'nin çalışmalarındaki tahmincilerle göre iki farklı test istatistiği tanımlamıştır. [22]'nin çalışmasındaki tahmin edici dikkate alındığında standart sapmanın tutarlı bir tahmincisi

$$\widehat{tr(\Sigma_i^2)} = \frac{(n_i - 1)^2}{(n_i + 1)(n_i - 2)} \left[tr \hat{\Sigma}_i^2 - \frac{1}{(n_i - 1)} (tr \hat{\Sigma}_i)^2 \right] \quad (22)$$

olmak üzere

$$\tilde{\sigma}_{CPH}^2 = \frac{2}{N^2} \left[\sum_{i=1}^k \frac{n_i(N - n_i)^2}{(n_i - 1)} tr(\widehat{\Sigma_i^2}) + \sum_{i \neq j}^k n_i n_j tr(\widehat{\Sigma_i \Sigma_j}) \right] \quad (23)$$

biçiminde elde edilmektedir. Böylece test istatistiğinin asimptotik dağılımı

$$T_{CPH}^* = \frac{T_{CPH}}{\tilde{\sigma}_{CPH}} \quad (24)$$

olacak şekilde standart normal dağılıma yakınsamaktadır [23]. Bu durumda α anlamlılık düzeyinde Denklem 24'teki T_{CPH}^* test istatistiği standart normal dağılımın kritik değerinden daha büyük ise yokluk hipotezi reddedilmektedir.

4. Bulgular

Bu çalışmanın amacı; yüksek boyutlu verilerde ve homojenlik varsayımının sağlanmadığı durumlarda, ortalama vektörlerin eşitliğini sınamak için önerilen test istatistiklerinin performanslarını farklı varyans-kovaryans modellerine göre karşılaştırmaktır. Bu amaçla yüksek boyutlu verilerde MANOVA problemleri için [20], [21], [22] ve [23] tarafından önerilen test istatistikleri bir simülasyon çalışması ile karşılaştırılmıştır. Test istatistiklerinin performanslarını karşılaştırmak için ilk olarak çok değişkenli normal dağılıma sahip rassal gözlemler türetilmiştir. Homojen olmayan varyans-kovaryans matrisleri altında test istatistiklerinin ortaya koyduğu I. tip hata yapma olasılıklarını belirlemek için üç farklı model kullanılmıştır:

1. **Model:** $\Sigma_i = i * I_p \quad i = 1, \dots, k$

2. **Model:** $\Sigma_i = \sigma_{js} = \begin{cases} u_{js} & j = s \\ 0 & j \neq s \end{cases}$

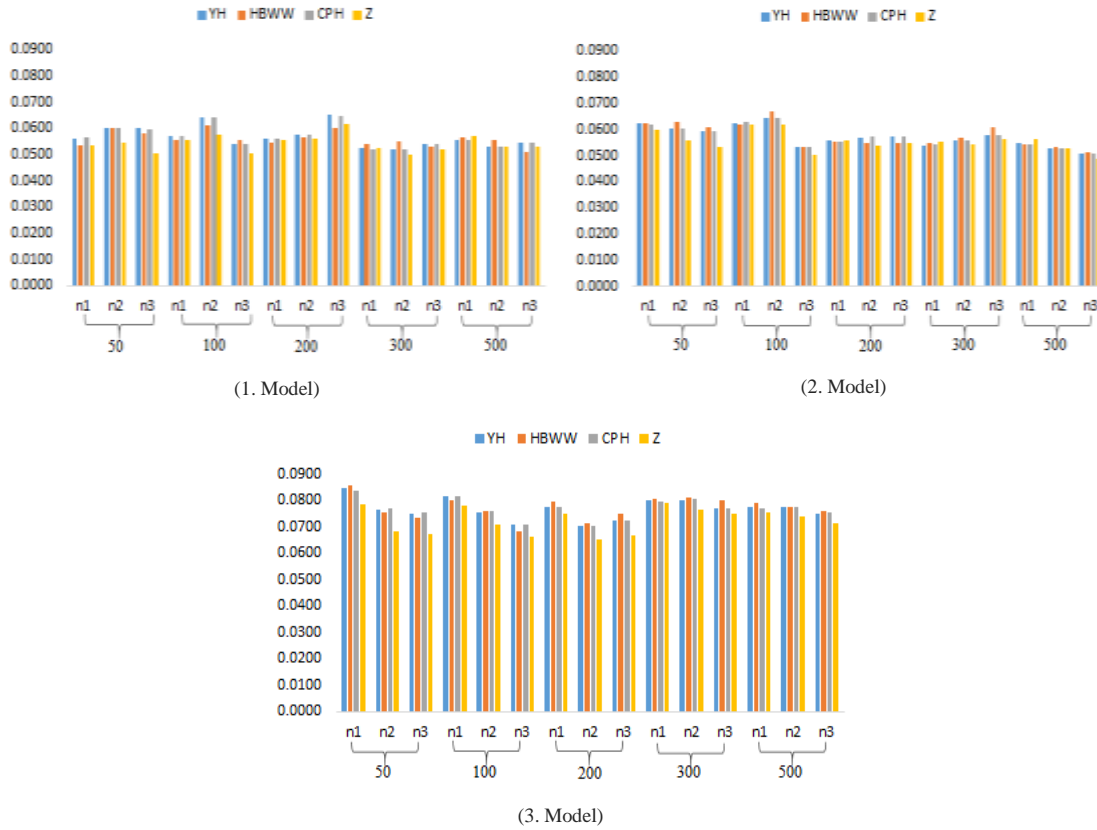
$$u_{js} \sim U(l, m) \quad \begin{array}{ll} i = 1 \Rightarrow U(0,2) & i = 3 \Rightarrow U(0.5,2.5) \\ i = 2 \Rightarrow U(1,4) & i = 4 \Rightarrow U(3,6) \end{array}$$

3. **Model:** $\Sigma_i = \sigma_{js} = \begin{cases} 1 & j = s \\ (-1)^{j+s} \frac{i * \rho}{s} & j < s \end{cases} \quad i = 1, \dots, k; \quad \rho = 0,05$

Burada $U(l, m)$; l ve m parametrelili Uniform dağılımı, J_p ; p boyutlu 1'lerden oluşan vektörü ve I_p ise p boyutlu birim matrisi temsil etmektedir. Böylece hem heterojenlik hem de farklı kovaryans yapıları altında test istatistiklerinin performansları araştırılmıştır. Çok değişkenli normal dağılım varsayımı altında her bir test istatistiğinin deneysel olasılık (p) değerleri hesaplanmıştır. Elde edilen olasılık değerleri α anlamlılık düzeyi ile karşılaştırılmış ve olasılık değerinden daha büyük olan α değerleri için yokluk hipotezi reddedilmiştir. Yapılan tekrar (iterasyon) sayısı N^* olmak üzere bu işlemler N^* kez tekrar edilmiştir. Böylece test istatistiklerinin her biri için elde edilen red sayıları toplamının yapılan iterasyon sayısına oranlanması ile I. tip hata yapma olasılıkları belirlenmiştir.

Simülasyon çalışmasında bütün test istatistikleri için 5000 iterasyon gerçekleştirilmiştir. Her bir test istatistiği için $k = 3$ ve $k = 4$ olmak üzere iki farklı grup düzeyi ve her bir grup için $p = 50, 100, 200, 300, 500$ olacak şekilde beş farklı değişken yapısı dikkate alınmıştır. Ayrıca anlamlılık düzeyi $\alpha = 0,05$ olarak belirlenirken, test istatistiklerinin algoritmalarını oluşturmak için RStudio programı kullanılmıştır.

Üç farklı varyans-kovaryans modeli için $k = 3$ durumunda $p = 50, 100, 200, 300, 500$ değerlerine ve n_i gözlem sayılarına göre test istatistiklerine ait deneysel hata yapma olasılıklarının grafikleri Şekil 1'de yer almaktadır.



Şekil 1: $k = 3$ durumunda üç farklı varyans-kovaryans modeline göre test istatistiklerinin I. tip hata yapma olasılıkları

Ayrıca her bir varyans-kovaryans modeli için $k = 3$ grup sayısı ile $p = 50, 100, 200, 300, 500$ değerlerine göre test istatistiklerine ait deneysel hata yapma olasılıkları Tablo 1'de verilmiştir.

Tablo 1. $k = 3$ durumunda üç farklı varyans-kovaryans modeline göre test istatistiklerinin I. tip hata yapma olasılıkları

Σ_i	p	(n_i)	T_{YH}^*	T_{HBWW}^*	T_{CPH}^*	T_Z^*
1. Model	50	n_1	0.0562	0.0534	0.0566	0.0536
		n_2	0.0602	0.0602	0.0600	0.0546
		n_3	0.0600	0.0580	0.0596	0.0504
	100	n_1	0.0570	0.0556	0.0572	0.0556
		n_2	0.0640	0.0610	0.0640	0.0574
		n_3	0.0540	0.0554	0.0542	0.0504
	200	n_1	0.0562	0.0546	0.0560	0.0558
		n_2	0.0578	0.0564	0.0578	0.0560
		n_3	0.0650	0.0600	0.0648	0.0614
	300	n_1	0.0526	0.0540	0.0522	0.0524
		n_2	0.0518	0.0548	0.0520	0.0498
		n_3	0.0538	0.0530	0.0540	0.0520
	500	n_1	0.0558	0.0568	0.0556	0.0572
		n_2	0.0532	0.0556	0.0530	0.0528
		n_3	0.0544	0.0512	0.0546	0.0528
2. Model	50	n_1	0.0622	0.0626	0.0620	0.0596
		n_2	0.0602	0.0630	0.0602	0.0558
		n_3	0.0594	0.0606	0.0594	0.0532
	100	n_1	0.0622	0.0618	0.0630	0.0620
		n_2	0.0644	0.0670	0.0644	0.0618
		n_3	0.0534	0.0534	0.0534	0.0504
	200	n_1	0.0556	0.0552	0.0552	0.0558
		n_2	0.0570	0.0548	0.0572	0.0538
		n_3	0.0572	0.0548	0.0574	0.0548
	300	n_1	0.0540	0.0546	0.0542	0.0554
		n_2	0.0556	0.0568	0.0558	0.0542
		n_3	0.0578	0.0610	0.0580	0.0562
	500	n_1	0.0546	0.0542	0.0544	0.0564
		n_2	0.0528	0.0532	0.0528	0.0526
		n_3	0.0506	0.0514	0.0506	0.0488
3. Model	50	n_1	0.0846	0.0858	0.0838	0.0784
		n_2	0.0764	0.0754	0.0768	0.0684
		n_3	0.0750	0.0734	0.0754	0.0672
	100	n_1	0.0816	0.0802	0.0814	0.0778
		n_2	0.0756	0.0758	0.0760	0.0706
		n_3	0.0708	0.0682	0.0710	0.0660
	200	n_1	0.0776	0.0794	0.0774	0.0748
		n_2	0.0702	0.0714	0.0702	0.0650
		n_3	0.0726	0.0750	0.0724	0.0668
	300	n_1	0.0800	0.0804	0.0796	0.0788
		n_2	0.0802	0.0812	0.0806	0.0766
		n_3	0.0772	0.0802	0.0772	0.0750
	500	n_1	0.0774	0.0792	0.0770	0.0756
		n_2	0.0776	0.0776	0.0774	0.0740
		n_3	0.0750	0.0760	0.0756	0.0714
		$n_1 = (10,15,25)$	$n_2 = (20,30,50)$	$n_3 = (30,45,75)$		

Şekil 1 ve Tablo 1 incelendiğinde $k = 3$ için birinci ve ikinci modele göre test istatistiklerinin I. tip hata oranları %5 ile %6 civarında elde edilirken, üçüncü modelde ise %7'den daha fazla olduğu gözlemlenmiştir. Bağımlı değişken sayısı arttıkça test istatistikleri için elde edilen deneysel hata olasılıklarının da nominal anlamlılık düzeyine yaklaştığı gözlemlenmiştir. Bütün deneysel koşullar dikkate alındığında T_Z^* test istatistiği, diğer test istatistiklerine göre daha düşük I. tip hata yapma olasılığı ortaya koymaktadır. T_{CPH}^* ve T_{YH}^* test istatistikleri genel olarak birbirlerine yakın değerler

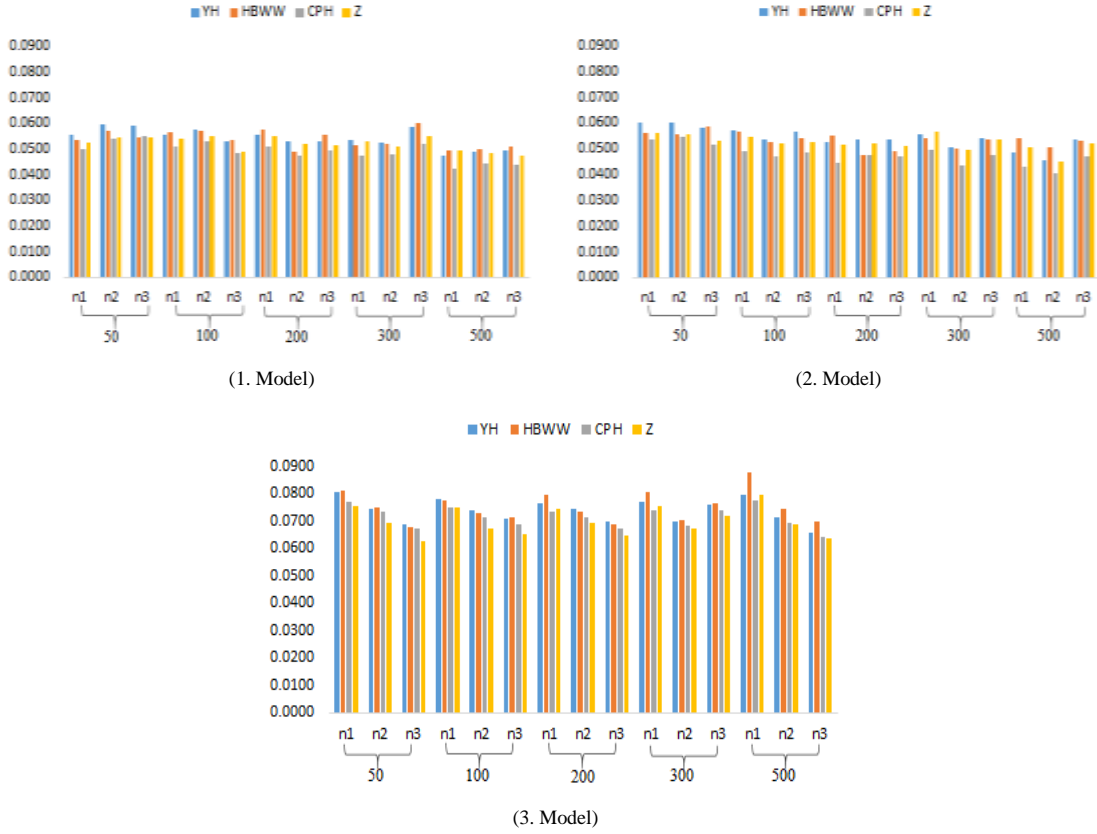
ortaya koyarken, T_{HBWW}^* test istatistiği ise varyans-kovaryans modeline göre farklılık göstermektedir. Birinci model için T_{HBWW}^* test istatistiği genel olarak T_{CPH}^* ve T_{YH}^* 'a göre daha düşük değerler ortaya koymuştur. Ancak ikinci ve üçüncü modelde ise en yüksek deneysel hata yapma olasılıkları T_{HBWW}^* test istatistiğinden elde edilmiştir.

Her bir varyans-kovaryans modeli için $k = 4$ ve $p = 50,100,200,300,500$ değerlerine göre test istatistiklerinin deneysel hata yapma olasılıkları Tablo 2'de verilmiştir.

Tablo 2. $k = 4$ durumunda üç farklı varyans-kovaryans modeline göre test istatistiklerinin I. tip hata yapma olasılıkları

Σ_i	p	(n_i)	T_{YH}^*	T_{HBWW}^*	T_{CPH}^*	T_Z^*
1. Model	50	n_1	0.0558	0.0534	0.0498	0.0524
		n_2	0.0594	0.0572	0.0540	0.0544
		n_3	0.0592	0.0546	0.0550	0.0546
	100	n_1	0.0558	0.0564	0.0512	0.0542
		n_2	0.0578	0.0572	0.0530	0.0550
		n_3	0.0530	0.0536	0.0484	0.0490
	200	n_1	0.0556	0.0576	0.0510	0.0552
		n_2	0.0532	0.0490	0.0476	0.0518
		n_3	0.0532	0.0554	0.0494	0.0514
	300	n_1	0.0536	0.0514	0.0472	0.0532
		n_2	0.0526	0.0520	0.0480	0.0512
		n_3	0.0584	0.0602	0.0520	0.0550
	500	n_1	0.0476	0.0496	0.0422	0.0494
		n_2	0.0488	0.0502	0.0442	0.0486
		n_3	0.0496	0.0512	0.0440	0.0476
2. Model	50	n_1	0.0602	0.0562	0.0534	0.0560
		n_2	0.0604	0.0554	0.0544	0.0558
		n_3	0.0584	0.0588	0.0518	0.0530
	100	n_1	0.0570	0.0566	0.0492	0.0544
		n_2	0.0538	0.0526	0.0470	0.0520
		n_3	0.0568	0.0540	0.0486	0.0524
	200	n_1	0.0526	0.0550	0.0444	0.0518
		n_2	0.0538	0.0476	0.0474	0.0522
		n_3	0.0534	0.0490	0.0472	0.0512
	300	n_1	0.0556	0.0540	0.0496	0.0566
		n_2	0.0508	0.0500	0.0434	0.0494
		n_3	0.0540	0.0536	0.0478	0.0536
	500	n_1	0.0488	0.0540	0.0432	0.0506
		n_2	0.0454	0.0506	0.0404	0.0450
		n_3	0.0536	0.0530	0.0472	0.0520
3. Model	50	n_1	0.0806	0.0812	0.0772	0.0756
		n_2	0.0746	0.0750	0.0732	0.0694
		n_3	0.0690	0.0678	0.0670	0.0628
	100	n_1	0.0778	0.0774	0.0750	0.0750
		n_2	0.0738	0.0728	0.0714	0.0670
		n_3	0.0708	0.0714	0.0686	0.0650
	200	n_1	0.0764	0.0796	0.0732	0.0742
		n_2	0.0742	0.0732	0.0714	0.0692
		n_3	0.0698	0.0688	0.0670	0.0646
	300	n_1	0.0770	0.0806	0.0740	0.0756
		n_2	0.0698	0.0702	0.0682	0.0670
		n_3	0.0760	0.0764	0.0738	0.0716
	500	n_1	0.0794	0.0878	0.0774	0.0794
		n_2	0.0712	0.0746	0.0694	0.0686
		n_3	0.0656	0.0698	0.0640	0.0636
$n_1 = (10,15,25,30)$		$n_2 = (20,30,50,60)$		$n_3 = (30,45,75,90)$		

Ayrıca üç farklı varyans-kovaryans modeli için $k = 4$ durumunda $p = 50, 100, 200, 300, 500$ değerlerine ve n_i gözlem sayılarına göre test istatistiklerine ait deneysel hata yapma olasılıkları Şekil 2’de yer almaktadır.



Şekil 2: $k = 4$ durumunda üç farklı varyans-kovaryans modeline göre test istatistiklerinin I. tip hata yapma olasılıkları

$k = 4$ için simülasyon sonuçlarını gösteren Tablo 2 incelendiğinde ise, $k = 3$ durumuna göre I. tip hata oranlarının nominal anlamlılık düzeyine daha yakın olduğu gözlemlenmiştir. Üç model için elde edilen sonuçlara göre genel olarak T_{CPH}^* ve T_Z^* test istatistikleri daha düşük deneysel hata yapma olasılıkları ortaya koymuştur. Ayrıca birinci ve ikinci modelde T_{HBWW}^* test istatistiğinin performansı T_{YH}^* ’dan daha iyi olduğu görülürken, üçüncü modelde ise en kötü performans T_{HBWW}^* test istatistiğinden elde edilmiştir. Bu durumda $k = 4$ için tek bir varyans-kovaryans modeline göre T_{HBWW}^* test istatistiğinin performansının T_{YH}^* ’a göre daha yüksek olduğu gözlemlenmiştir.

Simülasyon çalışması sonucunda elde edilen deneysel I. tip hata yapma olasılıklarına ($\hat{\alpha}$) göre test istatistiklerinin performanslarını belirlemek için [25], [26] ve [27] tarafından kullanılan “Ortalama Mutlak Tutarsızlık (AAD)” değeri hesaplanmıştır. Farklı deneysel koşullara göre elde edilen I. tip hata yapma olasılıkları ($\hat{\alpha}$) için AAD değeri

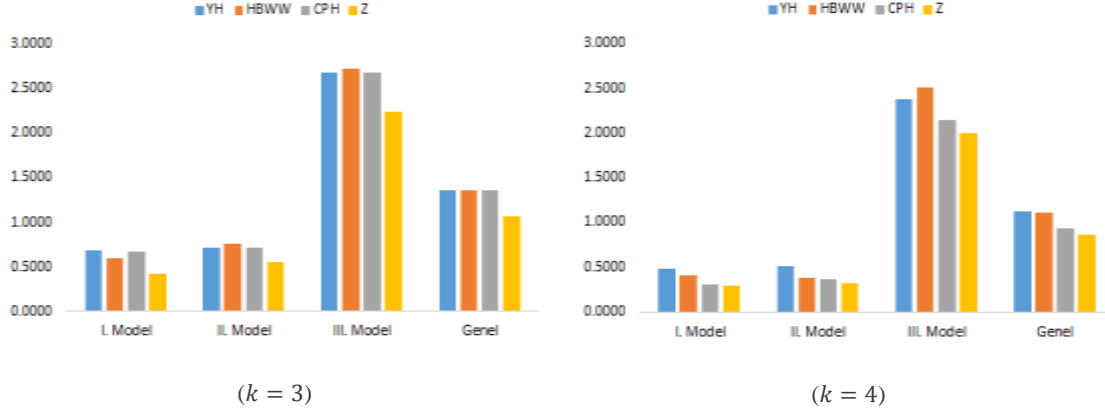
$$AAD = \frac{1}{\tilde{N}} \sum_{t=1}^{\tilde{N}} (\hat{\alpha}_t - \alpha) \quad (25)$$

biçiminde hesaplanmaktadır. AAD değeri düşük olan test istatistiğinin, daha iyi bir performans gösterdiği ifade edilmektedir. Tüm durumlar için elde edilen AAD değerleri Tablo 3'te yer almaktadır.

Tablo 3. Test istatistiklerine ait AAD değerleri

Grup Sayısı	Kovaryans Modeli	T_{YH}^*	T_{HBWW}^*	T_{CPH}^*	T_Z^*
$k = 3$	1. Model	0.6800	0.6000	0.6773	0.4173
	2. Model	0.7133	0.7627	0.7200	0.5547
	3. Model	2.6787	2.7280	2.6787	2.2427
	Genel:	1.3573	1.3636	1.3587	1.0716
$k = 4$	1. Model	0.4773	0.4120	0.3027	0.2920
	2. Model	0.5080	0.3813	0.3613	0.3147
	3. Model	2.3733	2.5107	2.1387	1.9907
	Genel:	1.1196	1.1013	0.9342	0.8658

Ayrıca $k = 3$ ve $k = 4$ durumunda üç farklı varyans-kovaryans modeli için hesaplanan AAD değerlerinin grafiği Şekil 3'te gösterilmiştir.



Şekil 3: $k = 3$ ve $k = 4$ durumunda üç farklı varyans-kovaryans modeline göre elde edilen AAD değerleri

Tablo 3 ve Şekil 3'e göre $k = 3$ ve $k = 4$ durumlarında bütün modellere göre en düşük AAD değeri T_Z^* test istatistiğinden elde edilmiştir. $k = 3$ için T_{YH}^* ve T_{CPH}^* test istatistiklerinin AAD değerleri oldukça yakın çıkmaktadır. Ancak $k = 4$ olduğunda ise T_{CPH}^* test istatistiğinin performansının T_{YH}^* 'a göre daha iyi olduğu söylenebilir. Ayrıca bütün modellerin I. tip hata yapma olasılıkları toplamına göre hesaplanan genel AAD değerleri incelendiğinde, $k = 3$ için T_{HBWW}^* test istatistiğinin AAD değeri T_{YH}^* 'a göre daha yüksektir. Dolayısıyla $k = 3$ için T_{HBWW}^* test istatistiğinin performansı daha düşüktür. Ancak $k = 4$ durumunda T_{HBWW}^* test istatistiğinin performansının T_{YH}^* 'a göre daha yüksek olduğu görülmektedir. Bu durum grup sayısı arttıkça T_{HBWW}^* test istatistiğinin performansının varyans-kovaryans modelinden daha az etkilendiğini göstermektedir.

Çalışmanın son aşamasında ise üç farklı varyans-kovaryans modeli için test istatistiklerinin güç değerleri hesaplanmıştır. Güç değerlerinin belirlenmesi için alternatif hipotez altında $\mu_1 = 0$, $\mu_2 = (-1)^i * (b_1, \dots, b_p)'$ ve $b_1 \sim U(0, t)$, $\mu_3 = -\mu_2$, $\mu_4 = 2\mu_2$ seçilmiştir. $k = 3$ için birinci ve ikinci modelde $t = 0.15$, üçüncü modelde

$t = 0.20$ olarak alınmıştır. $k = 4$ durumunda ise birinci ve ikinci modelde $t = 0.10$, üçüncü modelde $t = 0.13$ olarak alınmıştır. Böylece alternatif hipotez altında test istatistiklerinin güç değerleri elde edilmiştir. $k = 3$ ve $k = 4$ için üç farklı modele göre elde edilen güç değerleri sırasıyla Tablo 4 ve Tablo 5'te yer almaktadır.

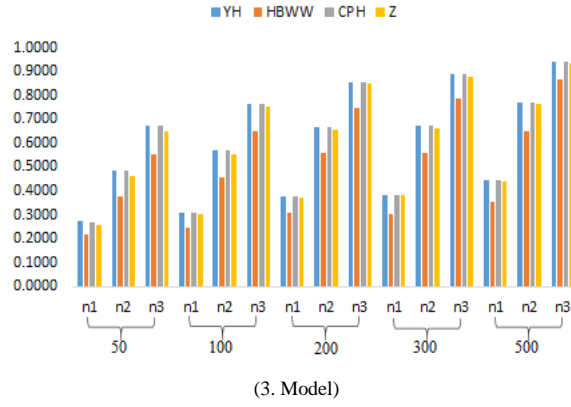
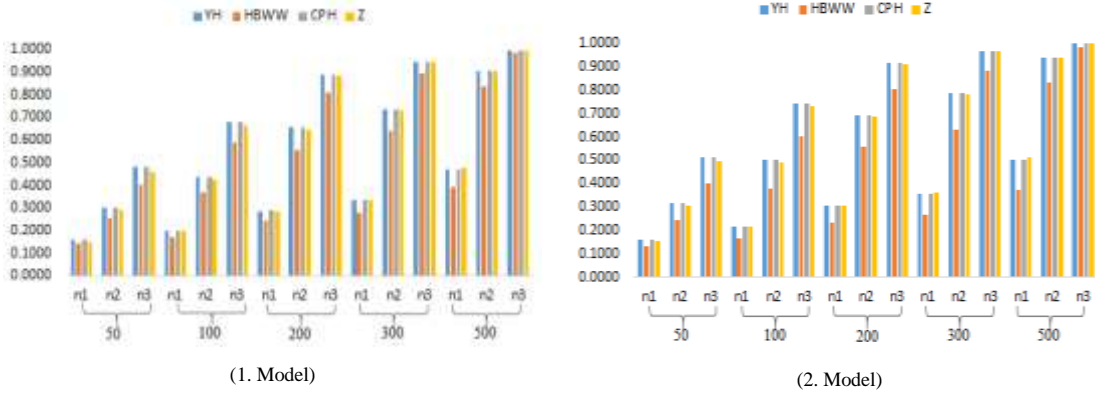
Tablo 4. $k = 3$ durumunda üç farklı varyans-kovaryans modeline göre test istatistiklerinin güç değerleri

Σ_i	p	(n_i)	T_{YH}^*	T_{HBWW}^*	T_{CPH}^*	T_Z^*
1. Model	50	n_1	0.1606	0.1408	0.1616	0.1498
		n_2	0.3034	0.2586	0.3038	0.2904
		n_3	0.4812	0.4042	0.4810	0.4588
	100	n_1	0.2004	0.1702	0.2016	0.1970
		n_2	0.4384	0.3688	0.4386	0.4252
		n_3	0.6792	0.5908	0.6794	0.6626
	200	n_1	0.2868	0.2456	0.2876	0.2864
		n_2	0.6542	0.5576	0.6542	0.6456
		n_3	0.8862	0.8104	0.8858	0.8802
	300	n_1	0.3358	0.2784	0.3364	0.3370
		n_2	0.7376	0.6410	0.7382	0.7318
		n_3	0.9468	0.8924	0.9466	0.9454
	500	n_1	0.4728	0.3924	0.4732	0.4766
		n_2	0.9074	0.8356	0.9074	0.9062
		n_3	0.9966	0.9854	0.9966	0.9962
2. Model	50	n_1	0.1616	0.1352	0.1610	0.1562
		n_2	0.3198	0.2458	0.3194	0.3054
		n_3	0.5150	0.3996	0.5148	0.4960
	100	n_1	0.2192	0.1686	0.2176	0.2156
		n_2	0.5010	0.3808	0.5016	0.4904
		n_3	0.7416	0.6026	0.7408	0.7302
	200	n_1	0.3078	0.2348	0.3062	0.3068
		n_2	0.6928	0.5584	0.6938	0.6882
		n_3	0.9172	0.8076	0.9170	0.9138
	300	n_1	0.3592	0.2670	0.3588	0.3626
		n_2	0.7874	0.6324	0.7858	0.7834
		n_3	0.9666	0.8842	0.9666	0.9650
	500	n_1	0.5018	0.3736	0.5050	0.5116
		n_2	0.9404	0.8352	0.9404	0.9396
		n_3	0.9992	0.9850	0.9992	0.9990
3. Model	50	n_1	0.2724	0.2180	0.2710	0.2602
		n_2	0.4872	0.3772	0.4880	0.4656
		n_3	0.6728	0.5551	0.6726	0.6516
	100	n_1	0.3100	0.2480	0.3100	0.3046
		n_2	0.5726	0.4566	0.5728	0.5552
		n_3	0.7656	0.6478	0.7658	0.7506
	200	n_1	0.3782	0.3086	0.3780	0.3730
		n_2	0.6676	0.5572	0.6676	0.6564
		n_3	0.8574	0.7466	0.8576	0.8472
	300	n_1	0.3832	0.3036	0.3838	0.3804
		n_2	0.6722	0.5570	0.6724	0.6604
		n_3	0.8866	0.7850	0.8870	0.8760
	500	n_1	0.4446	0.3528	0.4440	0.4414
		n_2	0.7720	0.6490	0.7722	0.7630
		n_3	0.9410	0.8648	0.9414	0.9366
		$n_1 = (10,15,25)$	$n_2 = (20,30,50)$		$n_3 = (30,45,75)$	

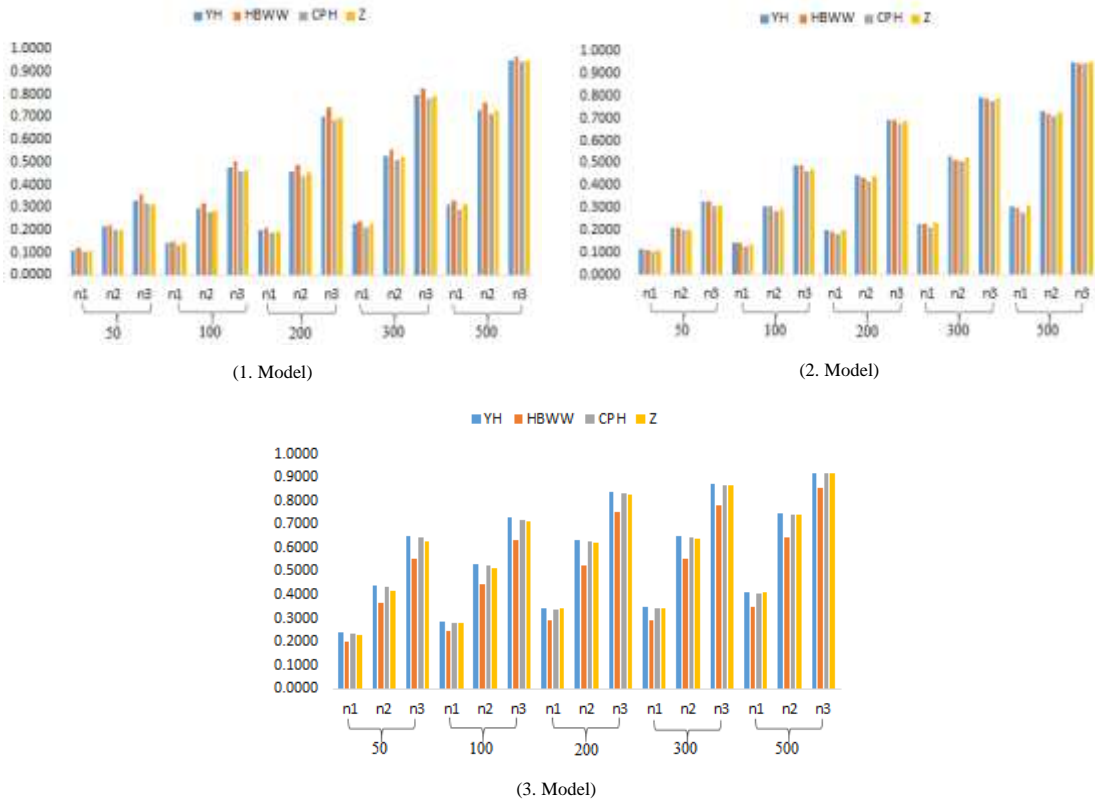
Tablo 5. $k = 4$ durumunda üç farklı varyans-kovaryans modeline göre test istatistiklerinin güç değerleri

Σ_i	p	(n_i)	T_{YH}^*	T_{HBWW}^*	T_{CPH}^*	T_Z^*
1. Model	50	n_1	0.1116	0.1196	0.1042	0.1072
		n_2	0.2140	0.2222	0.2004	0.2006
		n_3	0.3308	0.3586	0.3164	0.3152
	100	n_1	0.1446	0.1512	0.1334	0.1410
		n_2	0.2938	0.3156	0.2790	0.2850
		n_3	0.4756	0.5058	0.4584	0.4628
	200	n_1	0.1988	0.2106	0.1858	0.1966
		n_2	0.4594	0.4860	0.4376	0.4520
		n_3	0.7026	0.7388	0.6866	0.6958
	300	n_1	0.2252	0.2400	0.2114	0.2266
		n_2	0.5262	0.5556	0.5080	0.5224
		n_3	0.7974	0.8250	0.7800	0.7896
	500	n_1	0.3104	0.3312	0.2924	0.3128
		n_2	0.7326	0.7650	0.7112	0.7294
		n_3	0.9524	0.9660	0.9460	0.9506
2. Model	50	n_1	0.1172	0.1130	0.1066	0.1118
		n_2	0.2134	0.2100	0.1976	0.2022
		n_3	0.3272	0.3306	0.3056	0.3096
	100	n_1	0.1432	0.1438	0.1286	0.1400
		n_2	0.3088	0.3070	0.2844	0.2976
		n_3	0.4890	0.4888	0.4628	0.4756
	200	n_1	0.1998	0.1966	0.1826	0.1996
		n_2	0.4460	0.4364	0.4208	0.4386
		n_3	0.6938	0.6944	0.6752	0.6870
	300	n_1	0.2308	0.2270	0.2116	0.2326
		n_2	0.5286	0.5164	0.5074	0.5244
		n_3	0.7954	0.7876	0.7768	0.7890
	500	n_1	0.3076	0.2994	0.2802	0.3114
		n_2	0.7304	0.7222	0.7118	0.7288
		n_3	0.9506	0.9470	0.9458	0.9504
3. Model	50	n_1	0.2400	0.2020	0.2332	0.2284
		n_2	0.4406	0.3634	0.4334	0.4194
		n_3	0.6502	0.5552	0.6438	0.6274
	100	n_1	0.2878	0.2484	0.2836	0.2828
		n_2	0.5316	0.4458	0.5250	0.5154
		n_3	0.7286	0.6336	0.7210	0.7104
	200	n_1	0.3462	0.2926	0.3402	0.3418
		n_2	0.6316	0.5272	0.6266	0.6208
		n_3	0.8356	0.7552	0.8326	0.8280
	300	n_1	0.3470	0.2938	0.3408	0.3422
		n_2	0.6480	0.5528	0.6432	0.6408
		n_3	0.8722	0.7826	0.8682	0.8652
	500	n_1	0.4126	0.3506	0.4084	0.4136
		n_2	0.7444	0.6428	0.7412	0.7392
		n_3	0.9204	0.8540	0.9178	0.9164
		$n_1 = (10,15,25,30)$	$n_2 = (20,30,50,60)$	$n_3 = (30,45,75,90)$		

Ayrıca üç farklı varyans-kovaryans modeli için $k = 3$ ve $k = 4$ durumunda $p = 50,100,200,300,500$ değerlerine ve n_i gözlem sayılarına göre test istatistiklerine ait güç değerleri sırasıyla Şekil 4 ve Şekil 5'te yer almaktadır.



Şekil 4: $k = 3$ durumunda üç farklı varyans-kovaryans modeline göre test istatistiklerinin güç değerleri



Şekil 5: $k = 4$ durumunda üç farklı varyans-kovaryans modeline göre test istatistiklerinin güç değerleri

Üç farklı varyans-kovaryans modeli için $k = 3$ ve $k = 4$ durumunda elde edilen güç istatistikleri incelendiğinde bağımlı değişken sayısının küçük olduğu durumlarda test istatistiklerinin performansı daha düşük iken, boyut sayısı arttıkça test istatistiklerinin performanslarının da yükseldiği gözlemlenmiştir. Her iki tabloya göre en küçük güç değerleri T_{HBWW}^* test istatistiğinden elde edilmiştir. T_Z^* , T_{CPH}^* ve T_{YH}^* test istatistiklerinin güç değerleri ise daha yüksek bulunmuştur. T_Z^* test istatistiğinin performansı T_{CPH}^* ve T_{YH}^* 'a göre daha iyi olmasına rağmen güç değerleri daha küçük elde edilmiştir. Bu durum, T_Z^* test istatistiği için elde edilen I. tip hata yapma olasılıklarının T_{CPH}^* ve T_{YH}^* test istatistiklerine göre çok daha küçük olmasından kaynaklanmaktadır.

5. Gerçek Veri Uygulaması

Bu bölümde yüksek boyutlu k örneklem Behrens-Fisher problemlerinin çözümü için önerilen test istatistikleri gerçek bir veri örneği üzerinde uygulanmıştır. Bu amaçla [28] tarafından sunulan ses rehabilitasyon veri seti kullanılmıştır. Bu veri seti; konuşma sinyallerinden klinik bilgilerin çıkarılması amacıyla 126 konuşma sinyalini karakterize etmek için kullanılan 309 farklı özellik içermektedir. Üstelik [29], yüksek boyutlu kovaryans matrislerinin homojenlik testi için ses rehabilitasyon veri setinden yararlanmıştır.

Çalışmada konuşma sinyallerine ait veriler rassal olarak üç farklı gruba ayrılmıştır. Ayrıca dengeli ve dengesiz gözlem büyüklüklerinde test istatistiklerinin performanslarını karşılaştırmak amacıyla iki farklı gözlem durumu kullanılmıştır. Böylece yüksek boyutlu k örneklem Behrens-Fisher problemleri için önerilen test istatistikleri kullanılarak iki farklı gözlem durumuna göre üç grubun ortalamalar vektörü arasında farklılık olup olmadığı test edilmiştir. %5 anlamlılık düzeyinde test istatistiklerine ait olasılık değerleri Tablo 6'da yer almaktadır.

Tablo 6. Ses rehabilitasyon veri seti için test istatistiklerine ait olasılık değerleri

(n_1, n_2, n_3)	T_{YH}^*	T_{HBWW}^*	T_{CPH}^*	T_Z^*
(20,30,40)	.1957	.2295	.1996	.1498
(30,30,30)	.4437	.4453	.4453	.3096

Tablo 6 incelendiğinde dengeli ve dengesiz gözlem büyüklüklerine göre elde edilen olasılık değerleri $\alpha = .05$ 'ten daha büyük olduğu için yokluk hipotezi kabul edilmektedir. Bu durumda %5 anlamlılık düzeyinde üç grubun ortalamalar vektörü arasında anlamlı bir farklılık olmadığı söylenebilmektedir.

6. Sonuç ve Yorum

Son yıllarda bilgisayar teknolojisinde meydana gelen gelişmeler, büyük boyutlarda verilerin elde edilmesini ve bu verilerin kaydedilmesini kolaylaştırmaktadır. Bu durum değişken sayısının çok büyük olduğu veriler için istatistiksel analizlerin gerçekleştirilmesine imkân sağlamaktadır. Ancak günümüzde kullanılan istatistiksel yöntemler genellikle gözlem sayısının değişken sayısından daha fazla olduğu durumlar ile ilgilenmektedir. Bu nedenle klasik yöntemlerin birçoğu yüksek boyutlu verilerin analiz edilmesinde kullanılamamaktadır. Dolayısıyla yüksek boyutlu veri kavramının giderek yaygınlaştığı ve gelecekte daha da önemli hale geleceği düşünüldüğünde yüksek boyutlu veriler için uygun çözümlere yaklaşımlarının geliştirilmesi ve bu yöntemlerin kullanılması oldukça önemlidir.

Bu çalışmanın amacı, yüksek boyutlu Behrens-Fisher problemlerinde ortalama vektörlerin eşitliğini sınamak için uygun test istatistiklerini araştırmaktır. Bu amaçla varyans-kovaryans matrislerinin homojen olmadığı durumlar için literatürde önerilen dört test istatistiği bir simülasyon çalışması ile karşılaştırılmıştır. Test istatistiklerinin

performansını etkileyen faktörleri gözlemek için üç farklı varyans-kovaryans modeli dikkate alınmıştır. Ayrıca her bir test istatistiğinin farklı deneysel koşullar altındaki performansları belirlenmiştir.

Yüksek boyutlu verilerde MANOVA problemlerinin çözümü için yeni bir test istatistiği öneren [21], önerilen bu test istatistiğinin T_{YH}^* 'dan daha iyi bir performans ortaya koyduğunu belirtmektedir. Bizim simülasyon sonuçlarımız da T_Z^* test istatistiğinin performansının T_{YH}^* 'dan daha iyi olduğunu göstermektedir. Ayrıca varyans-kovaryans matrislerinin heterojen olduğu durumda [21]'in test istatistiğinin [22]'nin test istatistiğinden de daha iyi bir performans ortaya koyduğu görülmüştür.

[23] ise önerdiği test istatistiğini [20] ile [22]'nin test istatistikleri ile karşılaştırmış ve özellikle dengesiz gözlem büyüklüğünde T_{CPH}^* test istatistiğinin daha iyi bir performans ortaya koyduğunu göstermiştir. Çalışmamızda incelenen deneysel koşullara göre $k = 3$ durumunda T_{CPH}^* ve T_{YH}^* test istatistikleri için yakın sonuçlar elde edilirken, grup sayısı arttıkça T_{CPH}^* test istatistiğinin performansının daha iyi olduğu görülmüştür. T_{Hu} test istatistiği ise birinci varyans-kovaryans modeline göre T_{CPH}^* ve T_{YH}^* test istatistiklerinden daha iyi performans göstermektedir. Ancak diğer modellerde T_{HBWW}^* test istatistiğinin deneysel hata yapma oranlarının en yüksek olduğu gözlemlenmiştir. Bu durum T_{HBWW}^* test istatistiğinin varyans-kovaryans modelinden fazlasıyla etkilendiğini göstermektedir. [23]'te, test istatistiklerinin performansının varyans-kovaryans modelinden etkilendiğini ifade etmektedir.

Bütün sonuçlar dikkate alındığında genel olarak T_Z^* ve T_{CPH}^* test istatistiklerinin daha iyi performans ortaya koyduğu gözlemlenirken, T_{YH}^* ve T_{HBWW}^* test istatistiklerinin performansının ise daha düşük kaldığı görülmüştür. Grup sayısı arttıkça T_Z^* ve T_{CPH}^* test istatistiklerinin performansları yaklaşmaktadır. Ayrıca grup sayısı veya bağımlı değişken sayısı arttıkça, bütün test istatistiklerinin deneysel hata yapma olasılıklarının nominal anlamlılık düzeyine yaklaştığı görülmüştür. Bu durum, bağımlı değişken sayısı ile gözlem sayısı arasındaki oran yükseldikçe önerilen test istatistiklerinin daha iyi sonuçlar ortaya koyacağını göstermektedir.

Bu çalışma ile yüksek boyutlu ve varyans-kovaryans matrislerinin homojen olmadığı problemlerde, ikiden fazla grup ortalamasını karşılaştırmak için uygun yaklaşımların belirlenmesi hedeflenmiştir. Ayrıca önerilen test istatistiklerinin bir arada değerlendirilmesi ve farklı kovaryans matrisleri altında performanslarının karşılaştırılması bakımından literatüre katkı sağlamaktadır. Böylece varsayımların sağlanmadığı durumlarda uygun istatistiksel çözümleme yaklaşımına karar verilebilmektedir. Ancak test istatistiklerinin performansları; varyans-kovaryans matrislerinin yapısından, bağımlı değişken sayısından ve grup sayısından çok fazla etkilenmektedir. Bu durum test istatistiklerini performanslarının farklı deneysel koşullar altında değişmesine neden olmaktadır. Bu nedenle veri setinin özellikleri dikkate alınarak test istatistiklerinin güçleri belirlenmeli ve analizler için en yüksek performansa sahip test istatistiği seçilmelidir. Böylece elde edilen sonuçların doğru ve güvenilir olması beklenmektedir.

Araştırmacıların Katkı Oranı Beyanı

Mehmet Sandal: Metodoloji, Yazılım, Araştırma, İnceleme ve Düzenleme

Zeki Yıldız: Metodoloji, Araştırma, İnceleme ve Düzenleme, Denetim/Gözlem/Tavsiye

Destek ve Teşekkür Beyanı

Bu çalışmanın yazarları olarak herhangi bir destek ve teşekkür beyanımız bulunmadığını bildiririz.

Çatışma Beyanı

Bu çalışmanın yazarları olarak herhangi bir çatışma beyanımız bulunmadığını bildiririz.

Etik Kurul Onayı ve/veya Aydınlatılmış Onam Bilgileri

Bu çalışmanın yazarları olarak herhangi bir etik kurul onayı ve/veya aydınlatılmış onam bilgileri beyanımız bulunmadığını bildiririz.

Kaynakça

- [1] R. Alpar, *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*, Detay Yayıncılık, Ankara, 2011.
- [2] H. Finch, and B. French, “A monte carlo comparison of robust MANOVA test statistics,” *J. Mod. Appl. Stat. Methods*, 12 (2), 35-81, 2013.
- [3] S.S. Wilks, “Certain generalizations in the analysis of variance,” *Biometrika*, 24 (3-4), 471-494, 1932.
- [4] H. Hotelling, “A generalized T test and measure of multivariate dispersion,” *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 23-41, 1951.
- [5] K.C.S. Pillai, “Some new test criteria in multivariate analysis,” *Ann. Math. Stat.*, 26 (1), 117–121, 1955.
- [6] S.N. Roy, *Some Aspects of Multivariate Analysis*, Wiley, New York, 1957.
- [7] K. Jiamwattanapong, and S. Chongcharoen, “A new test for the mean vector in high-dimensional data,” *Songklanakarin Journal of Science and Technology*, 37 (4), 477-484, 2015.
- [8] Y. Shen, Z. Lin, and J. Zhu, “Shrinkage-based regularization tests for high-dimensional data with application to gene set analysis,” *Comput. Stat. Data Anal.*, 55 (7), 2221–2233, 2011.
- [9] A.P. Dempster, “A high dimensional two sample significance test,” *Ann. Math. Stat.*, 29 (4), 995–1010, 1958,
- [10] A.P. Dempster, “A significant test for the separation of two highly multivariate small samples,” *Biometrics*, 16 (1), 41–50, 1960.
- [11] Z. Bai, and H. Saranadasa, “Effect of high dimension: By an example of a two sample problem,” *Statistica Sinica*, 6 (2), 311–329, 1996.
- [12] S. Chen, and Y. Qin, “A two-sample test for high-dimensional data with applications to gene-set testing,” *Ann. Stat.*, 38 (2), 808–835, 2010.
- [13] Y. Fujikoshi, T. Himeno, and H. Wakaki, “Asymptotic results of a high dimensional MANOVA test and power comparisons when the dimension is large compared to the sample size,” *Journal of Japan Statistical Society*, 34, 19-26, 2004.
- [14] M.S. Srivastava, and Y. Fujikoshi, “Multivariate analysis of variance with fewer observations than the dimension,” *J. Multivar. Anal.*, 97 (9), 1927–1940, 2006.
- [15] M.S. Srivastava, “Multivariate theory for analyzing high-dimensional data,” *Journal of the Japan Statistical Society*, 37 (1), 53–86, 2007.
- [16] J.R. Schott, “Some high-dimensional tests for a one-way MANOVA,” *J. Multivar. Anal.*, 98 (9), 1825-1839, 2007.
- [17] J.T. Zhang, and J. Xu, “On the k-sample Behrens-Fisher problem for high-dimensional data,” *Science in China Series A:Mathematics*, 52 (6), 1285-1304, 2009.
- [18] T. Yamada, and M.S. Srivastava, “A test for the multivariate analysis of variance in high-dimension,” *Commun. Stat. Theory Methods*, 41, 2602–2612, 2012.
- [19] M.S. Srivastava, and T. Kubokawa, “Tests for multivariate analysis of variance in high dimension under non-normality,” *J. Multivar. Anal.*, 115, 204-216, 2013.
- [20] T. Yamada, and T. Himeno, “Testing homogeneity of mean vectors under heteroscedasticity in high-dimension,” *J. Multivar. Anal.*, 139, 7-27, 2015.
- [21] B. Zhou, “Linear hypothesis testing for high-dimensional data under heteroscedasticity,” Phd Dissertation, National University of Singapore, Singapore, 2016.
- [22] J. Hu, Z. Bai, C. Wang, and W. Wang, “On testing the equality of high dimensional mean vectors with unequal covariance matrices,” *Ann. Inst. Stat. Math.*, 69 (2), 365-387, 2017.

- [23] M. X. Cao, J. Park, and D. J. He, "A test for the k sample Behrens–Fisher problem in high dimensional data," *J. Stat. Plan. Inference*, 201, 86–102, 2019.
- [24] M. Aoshima, and K. Yata, "Two-stage procedures for high-dimensional data," *Seq. Anal.s*, 30 (4), 356-399, 2011.
- [25] H. Yanagira, and K.H. Yuan, "Three approximate solutions to the multivariate Behrens–Fisher problem," *Commun. Stat. Simulation Comput.*, 34 (4), 975-988, 2005.
- [26] T. Kawasaki, and T. Seo, "A Two Sample Test for Mean Vectors with Unequal Covariance Matrices," *Commun. Stat. Simulation Comput.*, 44 (7), 1850-1866, 2015.
- [27] P. Sukcharoen, and S. Chongcharoen, "A test on the multivariate Behrens–Fisher Problem in high-dimensional data by block covariance estimation," *J. Math. Stat.*, 15 (1), 44-54, 2019.
- [28] A. Tsanas, M.A. Little, C. Fox, and L.O. Ramig, "Objective automatic assessment of rehabilitative speech treatment in Parkinson's disease," *IEEE Trans. Neural Syst. Rehabilitation Eng.*, 22 (1), 181-190, 2014.
- [29] A. Qayed, and D. Han, "Homogeneity test of several high-dimensional covariance matrices for stationary processes under non-normality," *arXiv preprint arXiv:2008.09259*, 2020.