BUDANMIŞ ORTALAMALAR İÇİN KARŞILAŞTIRMA TESTLERİ

COMPARISON TESTS FOR TRIMMED MEANS

GÖZDE İŞLER

PROF. DR. SEVİL BACANLI

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı için Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

2021

ÖZET

BUDANMIŞ ORTALAMALAR İÇİN KARŞILAŞTIRMA TESTLERİ

GÖZDE İŞLER

Yüksek Lisans, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. Sevil BACANLI

Haziran 2021, 62 sayfa

Student-t, ANOVA-F testi gibi istatistiksel anlamlılık testleri tıp, psikoloji, biyoloji gibi birçok disiplinde yaygın olarak kullanılmaktadır. Testlerin 1. tip hata oranı üzerinde tam bir kontrol sağlaması için normallik ve varyasların homojenliği varsayımlarının sağlanması gerekmektedir. Varsayımlar sağlanmadığında testlerin 1.tip hata ve güç oranları olumsuz etkilenmektedir. Bu problem dayanıklı konum ölçüsü budanmış ortalamanın kullanılmasıyla büyük ölçüde giderilmektedir. Ayrıca budanmış ortalamaya dayalı testlerin bootstrap yöntemlerle birlikte kullanımı da elde edilen sonuçları geliştirebilmektedir.

Bu çalışmada, bağımlı ve bağımsız grup durumunda ortalama ve budanmış ortalama karşılaştırılmasında kullanılan testler 1.tip hata oranı bakımından karşılaştırılmıştır. Bağımsız grupların karşılaştırılmasında Student-t testi, Welch testi, Yuen testi, bootstrap-t ile Yuen testi, ANOVA-F testi, budanmış ortalama ile Welch testi, budanmış ortalama ve bootstrap-t ile Welch testi kullanılmıştır. Bağımlı grupların karşılaştırılmasında ise bağımlı örneklem t testi, Yuen testi, Huynh-Feldt($\tilde{\epsilon}$)-düzeltmeli ANOVA-F testi, budanmış ortalama ile $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi, bootstrap-t ile ANOVA-F testi ve budanmış ortalama ve bootstrap-t ile ANOVA-F testi kullanılmıştır. Simülasyon çalışması ile elde edilen sonuçlara göre bağımsız grup düzeninde varsayım bozulmalarının sebep olduğu problemlerden kaçınmak için bağımsız iki grup durumunda Yuen testi ve bootstrap-t ile Yuen testi; bağımsız

k grup durumunda ise budanmış ortalama ile Welch testi önerilmektedir. Bağımlı grup düzeninde ise iki bağımlı grup durumunda çarpık-aşırı kuyruklu dağılım varlığında Yuen testi, k bağımlı grup durumunda ise budanmış ortalama ve bootstrap-t ile ANOVA-F testi önerilmektedir.

Bu çalışmanın bağımsız grup karşılaştırmasının yanı sıra bağımlı grup karşılaştırmasını da içermesi bakımından literatüre katkı sağlayacağı düşünülmektedir.

Anahtar Kelimeler: Budanmış ortalama, ANOVA-F testi, Yuen testi, Welch testi, Bootstrap-t

ABSTRACT

COMPARISON TESTS FOR TRIMMED MEANS

GÖZDE İŞLER

Master of Science, Department of Statistics

Supervisor: Prof. Dr. Sevil BACANLI

June 2021, 62 pages

Statistical significance tests, such as Student's t test, ANOVA-F test, are widely used in many disciplines, including medicine, psychology, and biology. For these tests to exact control over the rate of a Type I error, normality and homoscedasticity must be satisfied. Type I error and power rates are adversely affected when these assumptions are violated. These problems are greatly reduced when using trimmed mean that is a robust measure of location. Also, the results could be improved by combining bootstrap methods with the tests based on trimmed means.

In this study, the tests for mean and trimmed mean equality were compared in terms of Type I Error probability in the case of dependent and independent groups. In comparison of independent groups Student's t test, Welch test, Yuen test, Yuen test with bootstrap-t, ANOVA-F test, Welch test with trimmed mean, Welch test with trimmed mean and bootstrap-t were used. In comparison of dependent groups paired sample t test, Yuen test, Huynh-Feldt($\tilde{\epsilon}$)-adjusted ANOVA-F test, $\tilde{\epsilon}$-adjusted ANOVA-F test with trimmed mean, ANOVA-F test with bootstrap-t, ANOVA-F test with trimmed mean and bootstrap-t were used. According to results of simulation study, within the context of independent groups designs, to avoid the problems of assumption violation, in the case of two independent groups, the recommended tests are Yuen test and Yuen test with bootstrap-t, while in the case of k independent groups Welch test with trimmed mean is recommended. Within the context of dependent groups designs, in the case of two dependent groups, the suggested test is Yuen test for skew-heavy tailed distribution, on the other hand, in the case of k dependent groups, ANOVA-F test with trimmed mean and bootstrap-t is recommended.

It is considered that this study will contribute to the literature in terms of including independent group comparison as well as dependent group comparison.

Keywords: Trimmed mean, ANOVA-F test, Yuen test, Welch test, Bootsrap-t

TEŞEKKÜR

Tez çalışma sürecimin her aşamasında değerli yorum ve katkılarıyla bana yol gösteren, sonsuz sabırla beni çalışmaya teşvik eden çok kıymetli hocam Sayın Prof. Dr. Sevil BACANLI'ya, lisans ve lisansüstü ders döneminde ve tezimin ilk safhasında bana yol gösteren, destek ve yardımlarını esirgemeyen değerli hocam Sayın Prof. Dr. Tülay SARAÇBAŞI'na, hayatım boyunca her zaman yanımda olan, bu süreçte bana her zamankinden daha çok anlayış gösteren başta annem Gülay İŞLER'e babam Mehmet İŞLER'e, abim Gökhan İŞLER'e ve tüm aileme, ayrıca bu süreçte desteklerini arkamda hissettiğim değerli arkadaşlarıma teşekkürü bir borç bilirim.

İÇİNDEKİLER

ÖZET i
ABSTRACT iii
TEŞEKKÜR v
İÇİNDEKİLER vi
ÇİZELGELER DİZİNİ viii
SİMGELER VE KISALTMALAR ix
1. GİRİŞ 1
2. GENEL BİLGİLER 4
2.1. Budanmış Ortalama 4
2.2. Winsorized Ortalama 6
2.3. Winsorized Varyans 6
2.4. Budanmış Ortalamanın Standart Hatası 6
2.5. Bootstrap Yöntemler 8
2.5.1. Yüzdelik Bootstrap Yöntemi 8
2.5.2. Bootstrap-t Yöntemi 9
3. BAĞIMSIZ GRUPLARIN KARŞILAŞTIRILMASINDA KULLANILAN YÖNTEMLER 11
3.1. Bağımsız İki Grubun Karşılaştırılmasında Kullanılan Yöntemler 11
3.1.1. Student-t Testi 12
3.1.2. Welch Testi 13
3.1.3. Yuen Testi 13
3.1.4. Bootstrap-t ile Yuen Testi 15
3.1.5. Budanmış Ortalama ile Bootstrap Yüzdelik Yöntemi 16
3.2. Bağımsız k Grubun Karşılaştırılmasında Kullanılan Yöntemler 16
3.2.1. ANOVA-F Testi 18
3.2.2. Welch Testi 19
3.2.3. Budanmış Ortalama ile Welch Testi 20
3.2.4. Budanmış Ortalama ve Bootstrap-t ile Welch Testi 21
3.2.5. Budanmış Ortalama ile Bootstrap Yüzdelik Yöntemi 22
3.2.6. Box Testi 23
4. BAĞIMLI GRUPLARIN KARŞILAŞTIRILMASINDA KULLANILAN YÖNTEMLER 25
4.1. Bağımlı İki Grubun Karşılaştırılmasında Kullanılan Yöntemler 25

4.1.1. Bağımlı Örneklem t Testi 26
4.1.2. Yuen Testi 26
4.2. Bağımlı k Grubun Karşılaştırılmasında Kullanılan Yöntemler 27
4.2.1. Tekrarlı Ölçümlü ANOVA-F Testi 29
4.2.2. Budanmış Ortalama ile Huynh-Feldt(ϵ) Düzeltmeli ANOVA-F Testi 30
4.2.3. Budanmış Ortalama ve Bootstrap-t ile ANOVA- F Testi 32
5. SİMÜLASYON ÇALIŞMASI 33
5.1. Bağımsız Grup Durumu İçin Simülasyon Çalışması 33
5.1.1. Bağımsız Grup Durumu İçin Simülasyon Düzeni 33
5.1.2. Bağımsız Grup Durumu İçin Simülasyon Sonuçları 36
5.2. Bağımlı Grup Durumu İçin Simülasyon Çalışması 45
5.2.1. Bağımlı Grup Durumu İçin Simülasyon Düzeni 46
5.2.2. Bağımlı Grup Durumu İçin Simülasyon Sonuçları 47
6. SONUÇ VE TARTIŞMA 53
7. KAYNAKLAR 56
EKLER 60
EK 1 – Simülasyonda Kullanılan R Paketleri ve Fonksiyonları 60
EK 2 - Tez Çalışması Orjinallik Raporu 61
ÖZGEÇMİŞ 62

ÇİZELGELER DİZİNİ

Çizelge 1. k=2 için deney düzenleri 34
Çizelge 2. k=4 için deney düzenleri 34
Çizelge 3. k=6 için deney düzenleri 35
Çizelge 4. Bağımsız grup simülasyon çalışmasında kullanılan g ve h parametre değerleri 36
Çizelge 5. Bağımsız k=2 durumunda normal dağılım için gerçekleşen 1.tip hata oranları 38
Çizelge 6. Bağımsız k=2 durumunda g&h(g=0,5 h=0,2) dağılım için gerçekleşen 1.tip hata oranları . 38
Çizelge 7. Bağımsız k=2 durumunda g&h(g=0,5 h=0,5) dağılım için gerçekleşen 1.tip hata oranları . 39
Çizelge 8. Bağımsız k=2 durumunda g&h(g=1 h=0) dağılım için gerçekleşen 1.tip hata oranları 39
Çizelge 9. Bağımsız k=2 durumunda g&h(g=1 h=0,5) dağılım için gerçekleşen 1.tip hata oranları 40
Çizelge 10. Bağımsız k=4 durumunda normal dağılım için gerçekleşen 1.tip hata oranları 40
Çizelge 11. Bağımsız k=6 durumunda normal dağılım için gerçekleşen 1.tip hata oranları 41
Çizelge 12. Bağımsız k=4 durumunda g&h(g=0,5 h=0,2) dağılım için gerçekleşen 1.tip hata oranları41
Çizelge 13. Bağımsız k=6 durumunda g&h(g=0,5 h=0,2) dağılım için gerçekleşen 1.tip hata oranları42
Çizelge 14. Bağımsız k=4 durumunda g&h(g=0,5 h=0,5) dağılım için gerçekleşen 1.tip hata oranları42
Çizelge 15. Bağımsız k=6 durumunda g&h(g=0,5 h=0,5) dağılım için gerçekleşen 1.tip hata oranları43
Çizelge 16. Bağımsız k=4 durumunda g&h(g=1 h=0) dağılım için gerçekleşen 1.tip hata oranları 43
Çizelge 17. Bağımsız k=6 durumunda g&h(g=1 h=0) dağılım için gerçekleşen 1.tip hata oranları 44
Çizelge 18. Bağımsız k=4 durumunda g&h(g=1 h=0,5) dağılım için gerçekleşen 1.tip hata oranları .. 44
Çizelge 19. Bağımsız k=6 durumunda g&h(g=1 h=0,5) dağılım için gerçekleşen 1.tip hata oranları .. 45
Çizelge 20.Bağımlı k grup simülasyon çalışmasında kullanılan g ve h parametre değerleri 46
Çizelge 21. Bağımlı k=2 durumunda normal dağılım için gerçekleşen 1.tip hata oranları 48
Çizelge 22. Bağımlı k=2 durumunda g&h(g=0 h=0,5) dağılım için gerçekleşen 1.tip hata oranları 48
Çizelge 23. Bağımlı k=2 durumunda g&h(g=0,5 h=0) dağılım için gerçekleşen 1.tip hata oranları 49
Çizelge 24. Bağımlı k=2 durumunda g&h(g=0,5 h=0,5) dağılım için gerçekleşen 1.tip hata oranları . 49
Çizelge 25. Bağımlı k=4 durumunda normal dağılım için gerçekleşen 1.tip hata oranları 50
Çizelge 26. Bağımlı k=4 durumunda g&h(g=0 h=0,5) dağılım için gerçekleşen 1.tip hata oranları 50
Çizelge 27. Bağımlı k=4 durumunda g&h(g=0,5 h=0) dağılım için gerçekleşen 1.tip hata oranları 51
Çizelge 28. Bağımlı k=4 durumunda g&h(g=0,5 h=0,5) dağılım için gerçekleşen 1.tip hata oranları . 52

SİMGELER VE KISALTMALAR

Simgeler

μ	Kitle ortalaması
μ_b	Kitle budanmış ortalaması
$\bar{y}_b$	Örneklem budanmış ortalaması
x_i	Örneklem Winsorized değer
$\bar{\boldsymbol{x}}$	Örneklem winsorized ortalaması
s_w^2	Örneklem winsorized varyans
Σ	Varyans-kovaryans matrisi
$\boldsymbol{\sigma^2}$	Kitle varyansı

Kısaltmalar

ANOVA	Tek yönlü varyans analizi
W	Welch testi
Y	Yuen testi
YB	Bootstrap-t ile Yuen testi
WB	Budanmış ortalama ile Welch testi
WBB	Budanmış ortalama ve bootstrap-t ile Welch testi
$FBud.$	Budanmış ortalama ile $\tilde{\varepsilon}$- düzeltmeli ANOVA-F testi
BF	Bootstrap-t ile ANOVA-F testi
$BFBud.$	Budanmış ortalama ve bootstrap-t ile ANOVA-F testi

1. GİRİŞ

İki ve daha fazla grubun ortalamalarının karşılaştırıldığı hipotez testleri tıp, biyoloji, psikoloji vb. alanlarda yaygınlıkla kullanılmaktadır. Bağımsız iki grubun ortalamaları arasında anlamlı bir farklılık olup olmadığının incelenmesinde en çok kullanılan test t testi iken grup sayısının ikiden fazla olduğu durumda en çok kullanılan test ANOVA-F testidir. Testlerin geçerli sonuçlar verebilmesi için varsayımların sağlanması gerekmektedir. Varsayımlardan ilki karşılaştırılan grupların normal dağılıma sahip kitleden alınmasıdır. İkinci varsayım ise grupların alındığı kitle varyanslarının homojen olmasıdır. Ancak uygulamada varsayımlar çoğunlukla sağlanamamaktadır. Varsayımlardan biri veya her ikisinin de sağlanmadığı durumlar için çeşitli testler önerilmiştir.

Varyansların homojenliği varsayımının sağlanmadığı durumda grup ortalamalarının karşılaştırılabilmesi için parametrik olmayan testlerin kullanılması, veri dönüşümü yapılması, dayanıklı istatistiksel yöntemlerin kullanılması gibi yaklaşımlar geliştirilmiş olup bu yaklaşımlardan biri de yaklaşık testlerdir. Testlerde yokluk hipotezinin doğru olduğu varsayımı altında test istatistiklerinin örneklem dağılımları, yaklaşık olarak F, t gibi örneklem dağılımları iyi bilinen test istatistiklerinin dağılımı gibidir (Özdemir, 2006). Literatürde Welch (1951), Box (1954), Brown-Forsythe (1974), Alexander-Govern (1994) tarafından yaklaşık testler önerilmiştir. Ancak yapılan çalışmalar varyansların homojenliği varsayımının yanı sıra normallik varsayımı da sağlanmadığında özellikle de örneklem büyüklüklerinin eşit olmadığı durumda, bu testlerin 1.tip hata oranı kontrolünü sağlayamadığını göstermiştir (Lix ve Keselman, 1998).

Kitle ortalaması ve varyansının tahmin edicileri olan $\bar{y}$ ve s^2'nin aykırı değerlerden oldukça etkilendiği bilinmektedir. Özellikle aşırı kuyruklu dağılım altında ortalamanın standart hatası oldukça büyümektedir (Lix ve Keselman, 1998). Bu tahmin edicilerin yerine budanmış ortalama ve winsorized varyans gibi dayanıklı tahmin edicilerin kullanımı Welch testini, normallik ve varyansların homojenliği varsayımının sağlanmamasının yarattığı etkiye karşı dayanıklı hale getirmektedir (Yılmaz ve Özdemir, 2016).

Budanmış ortalama veri setindeki en büyük ve en küçük değerlerin belli bir oranının çıkarılmasıyla kalan gözlemlerin ortalamasının hesaplanmasıyla elde edilmektedir. Budanacak gözlemlerin oranı araştırmacılar tarafından belirlenmektedir. Budanmış ortalamaya dayalı hipotez testine ilk olarak Tukey ve McLaughlin (1963)'in çalışmasında yer verilmiş ve Yuen (1974) tarafından iki grubun budanmış ortalamalarının karşılaştırılması incelenmiştir. Yuen'in önerdiği yöntem ise ikiden fazla grubun karşılaştırıldığı durumlar için geliştirilmiştir (Wu, 2007).

Kitlenin parametreleri tahmin edilmek istendiğinde ya da kitle parametreleri karşılaştırılırken kitle dağılımının bilinmediği durumlarla karşılaşılabilir. Bootstrap yöntemler ile kitle dağılımının bilinmesine gerek olmadan ya da kitle dağılımına ilişkin herhangi bir varsayımda bulunmadan çıkarsama yapılabilmektedir (Özdemir ve Navruz, 2016). Efron (1979) tarafından önerilen bootstrap yöntemi bir tekrarlı örnekleme yöntemidir. Bootstrap yöntemde örneklem kitle olarak düşünülür ve esas örneklemden yerine konularak her biri aynı büyüklükte M tane yeni örneklem üretilir (Yıldıztepe ve Özdemir, 2013). İlgilenilen istatistiğin M tane örneklemin her biri için hesaplanmasıyla birlikte elde edilen örnekleme dağılımı güven aralıklarının elde edilmesinde ve hipotez testlerinde kullanılır.

Grupların budanmış ortalamalarının karşılaştırıldığı testlerin bootstrap yöntemlerle birlikte kullanılmasının 1.tip hata oranı kontrolüne etkisinin araştırıldığı Wilcox ve arkadaşlarının çalışmasında (Wilcox ve ark., 1998), budanmış ortalama ve bootstrap yöntemlerin birlikte kullanıldığı testlerin 1. tip hata oranı kontrolünde daha iyi olduğu sonuçlar elde edilmiştir (Keselman ve ark., 2003).

Bağımsız grupların karşılaştırılmasında olduğu gibi bağımlı grup karşılaştırmalarında kullanılan testlerin de geçerli sonuçlar verebilmesi için varsayımların sağlanması gerekmektedir. Bağımlı iki grubun karşılaştırılmasında en yaygın kullanılan yöntem olan bağımlı örneklem t testinin uygulanabilmesi için normallik varsayımının sağlanması; bağımlı ikiden fazla grubun karşılaştırılmasında ise en sık kullanılan yöntem olan tekrarlı ölçümlü ANOVA-F testi için normallik ve varyansların homojenliği varsayımının yanı sıra küresellik varsayımının sağlanması gerekmektedir. Küresellik varsayımı sağlanmadığı durumlar için serbestlik derecesi düzeltmelerine dayalı bazı yöntemler, çok değişkenli test istatistikleri gibi yaklaşımlar önerilmekle birlikte küresellik varsayımının yanı sıra normalliğin de sağlanmadığı durumlarda budanmış ortalama ve winsorized varyans-kovaryansın kullanılması önerilen yöntemlerden biridir.

Tez çalışmasındaki amaç varsayımların sağlanmadığı durumlarda iki ve daha fazla grubun ortalama ve budanmış ortalamalarının karşılaştırmasında kullanılan testlerin tanıtılması ve bu testlerin simülasyon çalışması ile oluşturulan farklı deney düzenlerinde elde edilen 1. tip hata oranları yönünden performanslarının incelenmesidir.

Genel Bilgiler başlığı altında grupların karşılaştırılmasında kullanılan test istatistiklerinde, bilinen ortalama ve varyans yerine kullanılan dayanıklı tahmin edicilerden budanmış ortalama ve winsorized varyans ile test istatistiklerinin dağılımının kestirilmesinde kullanılan bootstrap yöntemler hakkında bilgi verilmiştir.

Üçüncü ve dördüncü bölümde sırasıyla bağımsız ve bağımlı grupların ortalamalarının ve budanmış ortalamalarının karşılaştırılmasında kullanılan testler tanıtılmıştır.

Beşinci bölümde simülasyon çalışmasıyla ilgili bilgiler verilmiş ve grupların ortalamalarının ve budanmış ortalamalarının karşılaştırıldığı testler 1. tip hata oranı kontrolü açısından karşılaştırılmıştır. Son bölümde ise simülasyon çalışması ile elde edilen sonuçlar yorumlanmıştır.

2. GENEL BİLGİLER

2.1. Budanmış Ortalama

Budanmış ortalama, veri setindeki en küçük ve en büyük gözlemlerin belli bir oranda budanması sonucunda geri kalan gözlemlerin ortalamasıdır. Budanmış ortalamanın kullanılmasının amacı normal ve normal olmayan dağılımlar için daha küçük standart hata elde etmektir (Di ve ark., 2014).

Kitle budanmış ortalaması μ_b'nin tahmin edilmesinde kullanılan örneklem budanmış ortalaması $\bar{y}_b$ aşağıdaki adımlar takip edilerek elde edilir (Wilcox, 2012).

1) $\{y_1, y_2, \dots, y_n\}$ gözlem değerlerinden oluşan n büyüklüğündeki rasgele örneklemin gözlem değerleri küçükten büyüğe sıralanır:

$$y_{(1)} \leq y_{(2)} \leq \cdots \leq y_{(n)}.$$

2) Küçükten büyüğe sıralanmış gözlem değerlerinin her iki kuyruğundan budanacak gözlemlerin oranı γ ile gösterilmek üzere budanacak gözlem sayısı g ve etkin örneklem büyüklüğü h

$$g = [\gamma n] \,, \tag{2.1}$$
$$h = n - 2g \,, \tag{2.2}$$

olarak hesaplanır.

Elde edilen g'nin tamsayı olmaması halinde g en yakın alt tamsayıya yuvarlanır.

3) Örneklem budanmış ortalaması $\bar{y}_b$,

$$\bar{y}_b = \frac{1}{h} \sum_{i=g+1}^{n-g} y_i \tag{2.3}$$

biçiminde tanımlanır.

Örneklem budanmış ortalamasının hesaplanışı aşağıda verilen örnek veri seti üzerinde gösterilmiştir.

ÖRNEK 1: Gözlem değerleri $\{14, 5, 26, 32, 9, 45, 78, 60, 51, 56, 37, 92, 18, 68\}$ olan rasgele bir örneklem olsun.

Budama oranı γ'nın %20 olduğu varsayımı altında örneklem %20 budanmış ortalaması aşağıdaki adımlar takip edilerek hesaplanır.

1) Gözlem değerleri küçükten büyüğe sıralanır:

5, 9, 14, 18, 26, 32, 37, 45, 51, 56, 60, 68, 78, 92.

2) Budama oranı ile örneklem büyüklüğü değerleri Eşitlik (2.1)'de yerine konarak budanacak gözlem sayısı $g = 0{,}20 \times 14 = 2{,}8$ hesaplanır. Bulunan 2,8 değeri en yakın alt tamsayı olan 2'ye yuvarlanır.

3) Küçükten büyüğe sıralanmış veri setinin sol ve sağ kuyruğundan 2'şer tane olmak üzere 4 gözlem değerinin çıkarılmasıyla etkin örneklem büyüklüğü $h = 14 - 2(2) = 10$ olarak elde edilir.

4) Etkin örneklem büyüklüğü ve budama yapıldıktan sonra kalan gözlem değerleri Eşitlik (2.3)'te yerine konularak örneklem budanmış ortalaması,

$$\bar{y}_b = \frac{1}{10}\sum_{i=3}^{12} y_i$$

$$= \frac{14+18+\cdots+60+68}{10} = 40{,}7$$

biçiminde elde edilir.

Budanmış ortalamanın varyans tahmini için matematiksel teknikler geliştirilmiştir. Buna göre kitle budanmış ortalamasının varyansının tahmini için winsorized ortalama ve winsorized varyans tahminlerinin elde edilmesi gerekmektedir (Wilcox, 2003).

2.2. Winsorized Ortalama

Örneklem gözlem değeri y_i'ye karşılık gelen x_i winsorized değeri,

$$x_i = \begin{cases} y_{g+1} & , \; y_i \leq y_{g+1} \\ y_i & , \; y_{g+1} < y_i < y_{n-g} \\ y_{n-g} & , \; y_i \geq y_{n-g} \end{cases} \tag{2.4}$$

biçiminde tanımlanır. Eşitlikte, $y_{(g+1)}$ örneklemin budanmamış en küçük gözlem değerini, $y_{(n-g)}$ ise örneklemin budanmamış en büyük gözlem değerini göstermektedir. Örneklem winsorized ortalaması $\bar{x}$ ise,

$$\boldsymbol{\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i} \tag{2.5}$$

biçiminde tanımlanır.

2.3. Winsorized Varyans

Örneklem winsorized varyansı s_w^2 ,

$$s_w^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{2.6}$$

biçiminde tanımlanır.

2.4. Budanmış Ortalamanın Standart Hatası

Örneklem budanmış ortalamasının standart hatası,

$$s_b = \frac{s_w}{(1-2\gamma)\sqrt{n}} \tag{2.7}$$

biçiminde tanımlanır.

Örnek 1'de verilen veri seti için winsorized ortalama, winsorized varyans ve budanmış ortalamanın standart hatasının hesaplanışı örnek 2 ile verilmiştir.

ÖRNEK 2:

1) $g = 2$ değerine göre budanması gereken en küçük 2 gözlem değeri(5 ve 9) budanmayıp yerlerine budanmamış en küçük gözlem değeri olan 14'ün alınmasıyla, aynı şekilde $g = 2$ değerine göre budanması gereken en büyük 2 gözlem değeri(78 ve 92) budanmayıp yerlerine budanmamış en büyük gözlem değeri olan 68'in alınmasıyla, 14 ve 68 arasında kalan gözlem değerlerinin ise aynen kalmasıyla winsorized değerler elde edilir.

$$x_i = \begin{cases} 14 \ , \ y_i \leq 14 \\ y_i \ \ , \ 14 < y_{ij} < 68 \\ 68 \ , \ y_i \geq 68 \end{cases}$$

2) 14, **14**, 26, 32, **14**, 45, 68, 60, 51, 56, 37, **68**, 18, **68** olarak elde edilen winsorized değerlerin Eşitlik (2.5)'te yerine koyulmasıyla örneklemin winsorized ortalaması aşağıdaki gibi elde edilir.

$$\bar{x} = \frac{1}{14}\sum_{i=1}^{14} x_i$$

$$= \frac{14+14+\cdots+18+68}{14} = 40{,}79.$$

3) Winsorized değerlerin ve winsorized ortalamanın Eşitlik (2.6)'da yerine koyulmasıyla örneklemin winsorized varyansı aşağıdaki gibi elde edilir.

$$s_w^2 = \frac{1}{14-1}\sum_{i=1}^{14}(x_i - 40{,}79)^2$$

$$= \frac{(14-40{,}79)^2+(14-40{,}79)^2+\cdots+(18-40{,}79)^2+(68-40{,}79)^2}{13}$$

$$= 21{,}35.$$

4) Winsorized varyansın, budama oranının(γ) ve örneklem büyüklüğünün Eşitlik (2.7)'de yerine koyulmasıyla örneklem budanmış ortalamasının standart hatası ise,

$$s_b = \frac{\sqrt{21{,}35}}{(1 - 2 * 0{,}2)\sqrt{14}} = 2{,}06$$

olarak bulunur.

2.5. Bootstrap Yöntemler

Efron (1979) tarafından önerilen bootstrap, bir yeniden örnekleme yöntemidir. Bu yöntemlerde aşağıda detayına yer verildiği üzere kitleden çekilen örneklemden yeni örneklemler üretilir (Yıldıztepe ve Özdemir, 2013). Bootstrap yöntemlerin arkasındaki temel fikir, örneklemi kullanarak yaklaşık bir örnekleme dağılımı elde etmek ve elde edilen dağılımı güven aralığı hesabında ve hipotez testinde kullanmaktır (Wilcox, 2010).

$\{y_1, y_2, \ldots, y_n\}$ olasılık dağılımı bilinmeyen bir kitleden çekilen n büyüklüğünde rasgele bir örneklem olsun. Örneklemden yerine koyarak rasgele çekilen ve $\{y_1^*, y_2^*, \ldots, y_n^*\}$ olarak gösterilen n tane gözlemin oluşturduğu örneklem bootstrap örneklem olarak adlandırılmaktadır. Bootstrap örneklemler kullanılarak esas örneklemin çekildiği kitlenin, ilgilenilen parametresi θ için güven aralığı elde edilmesi ve $H_0: \theta = \theta_0$ gibi hipotezlerin test edilmesi mümkündür (Wilcox, 2010).

Literatürde yüzdelik bootstrap, yanlılığı düzeltilmiş yüzdelik bootstrap, bootstrap-t gibi önerilen birçok bootstrap yöntem bulunmakla birlikte tez çalışmasında temel bootstrap yöntemlerden yüzdelik bootstrap ve bootstrap-t yöntemine yer verilmiştir.

2.5.1. Yüzdelik Bootstrap Yöntemi

$\{y_1, y_2, \ldots, y_n\}$ ortalaması μ olan bir kitleden rasgele çekilen n büyüklüğünde örneklem olsun ve bu örneklemin ortalaması $\bar{y}$ ile gösterilsin. Bu örneklemden $\{y_1^*, y_2^*, \ldots, y_n^*\}$ ile gösterilen n büyüklüğünde bootstrap örneklem çekilsin. Elde edilen bootstrap örneklemin ortalaması $\bar{y}^*$ ile gösterilmek üzere bu süreç M kez tekrarlandığında ortalaması sırasıyla $\bar{y}_1^*, \bar{y}_2^*, \ldots, \bar{y}_M^*$ olan M tane bootstrap örneklem elde edilecektir. Bu durumda kitle ortalaması için güven aralığı $\left(\bar{y}_{(l+1)}^*, \bar{y}_{(u)}^*\right)$ olacaktır. M bootstrap örneklemin küçükten büyüğe sıralanmış hali $\bar{y}_{(1)}^* \leq \cdots \leq \bar{y}_{(M)}^*$ olmak üzere $l = \frac{M\alpha}{2}$ ve $u = M - l$ hesaplanır. Burada α anlamlılık düzeyini, l ve u ise güven aralığının sırasıyla alt ve üst sınırına karşılık gelen değerlerin bulunmasında kullanılan sıra numaralarını ifade etmektedir (Wilcox, 2003).

2.5.2. Bootstrap-t Yöntemi

Kitle ortalaması μ'ye ait güven aralığının elde edilmesi veya $H_0: \mu = \mu_0$ hipotezinin test edilmesi t test istatistiğinin,

$$t = \frac{\bar{y} - \mu_0}{s/\sqrt{n}} \tag{2.8}$$

Student-t dağılımına sahip olduğu varsayımıyla mümkündür. Ancak normallik varsayımının sağlanmadığı bazı durumlarda t test istatistiğinin dağılımına ilişkin bu yaklaşım ile 1.tip hata kontrolü ve testin gücü açısından istenmeyen sonuçlarla karşılaşılabilmektedir. Normallik varsayımı olmadan t test istatistiğinin dağılımının kestirilmesi bu probleme bir çözüm olarak karşımıza çıkmaktadır (Wilcox, 2010).

Yüzdelik-t olarak da bilinen bootstrap-t yöntemi t'nin dağılımının kestirilmesinde kullanılabilmektedir. Bootstrap-t yöntemi ile $H_0: \mu = \mu_0$ hipotezinin test edilmesi için aşağıdaki adımlar izlenir:

1) Gözlem değerleri $\{y_1, y_2, \ldots, y_n\}$ ve ortalaması $\bar{y}$ olan n büyüklüğündeki rasgele örneklem için Eşitlik (2.8)'de verilen t test istatistiği hesaplanır.
2) İlk adımda tanımlanan örneklemden n büyüklüğünde bootstrap örneklem çekilir.
3) Gözlem değerleri $\{y_1^*, y_2^*, \ldots, y_n^*\}$ olan bootstrap örneklemin ortalaması $\bar{y}^*$ ve standart sapması s^* hesaplanır.
4) Elde edilen bootstrap örneklem için t^* değeri hesaplanır.

$$t^* = \frac{\bar{y}^* - \bar{y}}{s^*/\sqrt{n}} \tag{2.9}$$

5) Bu süreç M kere tekrarlandığında M tane t^* değeri buna bağlı olarak da t'nin örneklem dağılımının kestirimi elde edilecektir.
6) M tane t^* değeri küçükten büyüğe sıralanır:

$$t^*_{(1)} \leq t^*_{(2)} \leq \cdots \leq t^*_{(M)}.$$

7) Ardından H_0 hipotezinin testi için alt kritik değerin sıra sayısı $l = M\alpha/2$ eşitliği ile hesaplanır ve elde edilen değer en yakın tamsayıya yuvarlanır. Üst kritik değerin sıra sayısı ise $u = M - l$ eşitliği ile hesaplanır.
8) Eğer $t \leq t^*_{(l+1)}$ veya $t \geq t^*_{(u)}$ ise H_0 hipotezi reddedilir.
9) Kitle ortalaması μ için $(1-\alpha\ )$ güven düzeyine karşılık gelen güven aralığı

$$\left(\bar{y} - t^*_{(u)}\frac{s}{\sqrt{n}}\ , \bar{y} - t^*_{(l+1)}\frac{s}{\sqrt{n}}\right)$$

biçiminde tanımlanır (Wilcox, 2003).

3. BAĞIMSIZ GRUPLARIN KARŞILAŞTIRILMASINDA KULLANILAN YÖNTEMLER

Bu bölümde bağımsız iki ve ikiden fazla grubun konum parametrelerinden ortalama ve budanmış ortalama kullanılarak karşılaştırılması incelenmiştir.

3.1. Bağımsız İki Grubun Karşılaştırılmasında Kullanılan Yöntemler

Kitle ortalaması μ_j $(j = 1,2)$ ile gösterilmek üzere bağımsız iki grubun kitle ortalamaları arasında anlamlı bir farklılık olup olmadığı Eşitlik (3.1)'de verilen H_0 hipotezi ile test edilir:

$$H_0: \mu_1 = \mu_2 \, . \tag{3.1}$$

Student-t testi, Eşitlik (3.1)'deki H_0 hipotezinin testinde en yaygın kullanılan yöntemlerden biri olup 1. tip hata olasılığı üzerinde tam bir kontrol sağlanması açısından aşağıdaki varsayımların sağlanması gerekmektedir (Wilcox, 2003):

1. Örneklemin rasgele çekilmesi (Rasgelelik varsayımı),
2. Örneklemin çekildiği kitlenin normal dağılması (Normallik varsayımı),
3. Kitle varyanslarının eşit olması (Varyansların homojenliği varsayımı).

Normallik varsayımının sağlandığı ancak varyansların homojenliği varsayımının sağlanmadığı durumda iki grubun konum ölçüleri karşılaştırılırken ortaya çıkan probleme Behrens-Fisher problemi denir (Yılmaz ve Özdemir, 2016). Welch (1938) tarafından önerilen test, bu probleme çözüm sağlamakla birlikte normallik varsayımı sağlanmadığında 1. tip hata oranının kontrolünde başarılı sonuçlar verememektedir (Guo ve Luh, 2000). Bu iki varsayımın sağlanmadığı durumlara karşı dayanıklı tahmin ve test yöntemleri önerilmiş olup bunlardan biri de budanmış ortalama ve winsorized varyansı kullanan Yuen testidir. Burada kitle budanmış ortalaması μ_{bj} $(j = 1,2)$ ile gösterilmek üzere bağımsız iki grubun kitle budanmış ortalamaları arasında anlamlı bir farklılık olup olmadığı

$$H_0: \mu_{b1} = \mu_{b2} \tag{3.2}$$

ile test edilir.

3.1.1. Student-t Testi

Eşitlik (3.1) ile verilen H_0 hipotezi Student's t testi ile aşağıdaki adımlar takip edilerek test edilir:

1) j. gruptan gözlem değerleri y_{ij} $\left(i = 1, \dots, n_j\, ; j = 1{,}2\right)$ ile gösterilmek üzere, rasgele çekilen n_j büyüklüğündeki örneklemin ortalaması $\bar{y}_{.j}$ ve varyansı s_j^2 hesaplanır:

$$\bar{y}_{\cdot j} = \frac{\sum_{i=1}^{n_j} y_{ij}}{n_j}, \tag{3.3}$$

$$s_j^2 = \frac{\sum_{i=1}^{n_j} \left(y_{ij} - \bar{y}_{\cdot j}\right)^2}{n_j - 1}. \tag{3.4}$$

2) Eşitlik(3.5)'de verildiği üzere toplam varyans σ_t^2'nin tahmini s_t^2,

$$s_t^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \tag{3.5}$$

şeklinde tanımlanır.

3) Test istatistiği t,

$$t = \frac{(\bar{y}_{.1} - \bar{y}_{.2}) - (\mu_1 - \mu_2)}{\sqrt{s_t^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{3.6}$$

biçiminde tanımlanır.

4) $|t| \geq t_{(v,1-\alpha/2)}$ ise H_0 hipotezi reddedilir. Burada serbestlik derecesi v ,

$$v = n_1 + n_2 - 2 \tag{3.7}$$

biçiminde tanımlanır.

İki grubun kitle ortalamaları arasındaki fark $(\mu_1 - \mu_2)$ için $(1 - \alpha\ )$ güven düzeyine karşılık gelen güven aralığı ise

$$(\bar{y}_{.1} - \bar{y}_{.2}) \pm t_{(v,1-\alpha/2)}\sqrt{s_t^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)},$$

biçiminde tanımlanır (Wilcox, 2003).

3.1.2. Welch Testi

Eşitlik (3.1) ile verilen H_0 hipotezi Welch testi ile aşağıdaki adımlar izlenerek test edilir.

1) j. gruptan gözlem değerleri y_{ij} $\left(i = 1, \ldots, n_j \, ; j = 1{,}2\right)$ ile gösterilmek üzere rasgele çekilen n_j büyüklüğündeki örneklemin ortalaması $\bar{y}_{.j}$ ve varyansı s_j^2 hesaplanır.

2) $\bar{y}_{.j}$'nin standart hatasının karesi $s_{\bar{y}_{.j}}^2$ Eşitlik (3.8)'deki gibi elde edilir.

$$s_{\bar{y}_{.j}}^2 = \frac{s_j^2}{n_j}. \tag{3.8}$$

3) Welch'in test istatistiği W,

$$W = \frac{(\bar{y}_{.1} - \bar{y}_{.2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{3.9}$$

biçiminde tanımlanır.

4) Eğer $|W| \geq t_{(\hat{v},1-\alpha/2)}$ ise H_0 hipotezi reddedilir. Burada serbestlik derecesi $\hat{v}$, Eşitlik (3.10)'da verildiği gibi elde edilir.

$$\hat{v} = \frac{\left(s_{\bar{y}_{.1}}^2 + s_{\bar{y}_{.2}}^2\right)^2}{\frac{\left(s_{\bar{y}_{.1}}^2\right)^2}{n_1 - 1} + \frac{\left(s_{\bar{y}_{.2}}^2\right)^2}{n_2 - 1}}\ . \tag{3.10}$$

5) İki grubun kitle ortalamaları arasındaki fark $(\mu_1 - \mu_2)$ için $(1 - \alpha\,)$ güven düzeyine karşılık gelen güven aralığı ise,

$$(\bar{\boldsymbol{y}}_{.\mathbf{1}} - \bar{\boldsymbol{y}}_{.\mathbf{2}}) \pm \boldsymbol{t}_{(\hat{\boldsymbol{v}},\mathbf{1}-\boldsymbol{\alpha}/\mathbf{2})}\sqrt{\frac{\boldsymbol{s}_{\mathbf{1}}^{\mathbf{2}}}{\boldsymbol{n}_{\mathbf{1}}} + \frac{\boldsymbol{s}_{\mathbf{2}}^{\mathbf{2}}}{\boldsymbol{n}_{\mathbf{2}}}}$$

biçiminde tanımlanır (Algina ve ark., 1994).

3.1.3. Yuen Testi

Kitle budanmış ortalamalarının karşılaştırılması için Yuen (1974) tarafından önerilen bu test winsorized varyansların heterojen olmasına izin vermektedir. Budama oranının 0 olduğu durumda Yuen testi, varyanslar heterojen olduğunda kullanılabilen Welch testine dönüşmektedir (Wilcox, 2012).

Eşitlik (3.2) ile verilen H_0 hipotezi Yuen testi ile aşağıdaki adımlar izlenerek test edilir:

1) j. gruptan gözlem değerleri y_{ij} $\left(i = 1, \dots, n_j \,; j = 1,2\right)$ ile gösterilmek üzere rasgele çekilen n_j büyüklüğündeki örneklem için budanacak gözlem sayısı g_j , etkin örneklem büyüklüğü h_j , budanmış ortalama $\bar{y}_{bj}$ ve winsorized varyans s_{wj}^2 sırasıyla Eşitlik (2.1), Eşitlik (2.2), Eşitlik (2.3) ve Eşitlik (2.6) kullanılarak elde edilir.

2) $\bar{y}_{bj}$'nin standart hatasının karesi için Yuen tarafından önerilen d_j,

$$d_j = \frac{(n_j-1)s_{wj}^2}{h_j(h_j-1)} \qquad (3.14)$$

şeklinde hesaplanır.[1]

3) Yuen'in test istatistiği t_y,

$$t_y = \frac{(\bar{y}_{b1}-\bar{y}_{b2})-(\mu_{b1}-\mu_{b2})}{\sqrt{d_1+d_2}} \qquad (3.15)$$

biçiminde elde edilir.

4) Eğer $\left|t_y\right| \geq t_{(1-\frac{\alpha}{2}\,;\,\hat{v}_y)}$ ise H_0 hipotezi reddedilir. Burada serbestlik derecesi $\hat{v}_y$,

$$\hat{v}_y = \frac{(d_1+d_2)^2}{\frac{{d_1}^2}{h_1-1}+\frac{{d_2}^2}{h_2-1}} \qquad (3.16)$$

biçiminde tanımlanır (Wilcox ve Keselman, 2001).

İki grubun kitle budanmış ortalamaları arasındaki fark $(\mu_{b1} - \mu_{b2})$ için $(1 - \alpha\,)$ güven düzeyine karşılık gelen güven aralığı ise

$$(\bar{y}_{b1} - \bar{y}_{b2}) \pm t_{\left(1-\frac{\alpha}{2}\,;\,\hat{v}_y\right)}\sqrt{d_1 + d_2}$$

[1] d_j için bundan sonra "Yuen'in standart hata kare tahmini" ifadesi kullanılacaktır.

olarak tanımlanır (Wilcox, 2003).

3.1.4. Bootstrap-t ile Yuen Testi

Eşitlik (3.15)'te tanımlanan t_y test istatistiğinin dağılımının bootstrap-t yöntemi ile kestirilmesi ve Eşitlik (3.2)'deki H_0 hipotezinin Yuen testi ile test edilmesi için aşağıdaki adımlar takip edilir.

1) j. gruptan gözlem değerleri y_{ij} $\left(i = 1, \dots, n_j\, ; j = 1,2\right)$ ile gösterilmek üzere rasgele çekilen n_j büyüklüğündeki örneklem için budanmış ortalama $\bar{y}_{bj}$, Yuen'in standart hata kare tahmini d_j ve Yuen'in test istatistiği t_y sırasıyla Eşitlik (2.3), Eşitlik (3.14) ve Eşitlik (3.15) ile hesaplanır.
2) 1. adımda elde edilen n_{j} büyüklüğündeki örneklemden gözlem değerleri $y^*_{1\mathrm{j}}, y^*_{2\mathrm{j}}, \dots, y^*_{n_{\mathrm{j}}\mathrm{j}}$ şeklinde gösterilmek üzere bootstrap örneklem çekilir. Elde edilen bootstrap örneklemin budanmış ortalaması $\bar{y}^*_{bj}$ ile Yuen'in standart hata kare tahmini d^*_j hesaplanır.
3) Elde edilen bootstrap örneklem için Yuen'in test istatistiği t^*_y,

$$t^*_y = \frac{(\bar{y}^*_{b1} - \bar{y}^*_{b2}) - (\bar{y}_{b1} - \bar{y}_{b2})}{\sqrt{d^*_1 + d^*_2}} \tag{3.17}$$

biçiminde tanımlanır.

4) 2. ve 3. adımın M kez tekrarlanmasıyla elde edilen $t^*_{y1}\,, t^*_{y2}\,, \dots, t^*_{yM}$ değerleri küçükten büyüğe $t^*_{y(1)} \leq \cdots \leq t^*_{y(M)}$ şeklinde sıralanır.
5) H_0 hipotezinin testi için alt kritik değerin sıra sayısı $l = M\alpha/2$ eşitliği ile hesaplanır ve elde edilen değer en yakın tamsayıya yuvarlanır. Üst kritik değerin sıra sayısı ise $u = M - l$ eşitliği ile hesaplanır.

6) 1. adımda elde edilen t_y değeri $t^*_{y(l+1)}$ ve $t^*_{y(u)}$ kritik değerleriyle karşılaştırılır. Eğer, $t_y \leq t^*_{y(l+1)}$ veya $t_y \geq t^*_{y(u)}$ ise Eşitlik (3.2)'deki H_0 hipotezi reddedilir.

İki grubun kitle budanmış ortalamaları arasındaki fark $(\mu_{b1} - \mu_{b2})$ için $(1 - \alpha\,)$ güven düzeyine karşılık gelen güven aralığı ise

$$\left((\bar{y}_{b1} - \bar{y}_{b2}) - t^*_{y(u)}\sqrt{d_1 + d_2}\,, (\bar{y}_{b1} - \bar{y}_{b2}) - t^*_{y(l+1)}\sqrt{d_1 + d_2}\ \right)$$

olarak tanımlanır (Wilcox, 2012).

3.1.5. Budanmış Ortalama ile Bootstrap Yüzdelik Yöntemi

Eşitlik (3.2)'deki $\boldsymbol{H_0}$ hipotezi bootstrap yüzdelik yöntemi ile aşağıdaki adımlar takip edilerek test edilir.

1) j. gruptan gözlem değerleri y_{ij} $\left(i = 1, \dots, n_j \,; j = 1,2\right)$ ile gösterilmek üzere, rasgele çekilen n_j büyüklüğündeki örneklemden n_j büyüklüğünde bootstrap örneklem çekilir ve elde edilen bootstrap örneklemin budanmış ortalaması $\bar{y}^*_{bj}$ hesaplanır.
2) $j = 1,2$ için elde edilen $\bar{y}^*_{b1}$ ve $\bar{y}^*_{b2}$ değerleri arasındaki fark bulunur:

$D^* = \bar{y}^*_{b1} - \bar{y}^*_{b2}$.

3) Bu süreç M kere tekrarlanarak $D^*_1, \dots, D^*_M$ değerleri hesaplanır.
4) $l = M\alpha/2$ eşitliği hesaplanır ve elde edilen değer en yakın tamsayıya yuvarlanır, ardından $u = M - l$ eşitliği hesaplanır.
5) İki grubun kitle budanmış ortalamaları arasındaki fark $(\mu_{b1} - \mu_{b2})$ için $(1 - \alpha\,)$ güven düzeyine karşılık gelen güven aralığı $\left(D^*_{(l+1)}, D^*_{(u)}\right)$ olarak elde edilir. Eğer bu aralık 0'ı içermiyorsa H_0 hipotezi reddedilir (Wilcox, 2003).

3.2. Bağımsız $\boldsymbol{k}$ Grubun Karşılaştırılmasında Kullanılan Yöntemler

Bağımsız iki grubun kitle ortalamalarının eşitliğine ilişkin hipotez, bağımsız k grubun kitle ortalamalarının eşitliği araştırılmak istendiğinde Eşitlik (3.18)'de verilen hipoteze dönüşür:

$$H_0\colon \mu_1 = \mu_2 = \cdots = \mu_k \,. \qquad (3.18)$$

Eşitlik (3.18) ile verilen H_0 hipotezinin testine ilişkin yöntem, varyans analizi (Analysis of Variance, ANOVA)[2] olarak adlandırılmaktadır. 1920'lerde, Sir Ronald Fisher tarafından önerilen ANOVA-F testinin geçerli sonuçlar verebilmesi için aşağıda sıralanan varsayımların sağlanması gerekmektedir:

[2] ANOVA, varyanslar arasında kıyaslama yaparak F test istatistiğine göre kitle ortalamaları arasında anlamlı bir farklılık olup olmadığına karar verilmesini sağladığından bundan sonra bu yöntem için ANOVA-F testi ifadesi kullanılacaktır.

1. Rasgelelik varsayımı,
2. Gözlemlerin bağımsız olması,
3. Normallik varsayımı,
4. Varyansların homojenliği varsayımı.

Uygulamada normallik ve varyansların homojenliği varsayımlarının ayrı ayrı veya aynı anda sağlanmadığı durumlarla karşılaşılabilmektedir. Normallik varsayımının sağlandığı, ancak varyansların homojenliği varsayımının sağlanmadığı durumlarda ANOVA-F testinin gücü düşebilmektedir. Varyansların homojenliğinin sağlanmadığı durumlarda H_0 hipotezinin test edilebilmesi için birçok yöntem önerilmiştir. Bu yöntemlerden biri de Welch (1951) tarafından önerilmiştir. Ancak bu testler normallik varsayımının sağlanmadığı durumlarda kullanıldığında 1. tip hata oranının kontrolünün sağlanamaması ve testin gücünün düşmesi problemleri ortaya çıkabilmektedir (Wilcox, 2003). Bu testlerin, bilinen ortalama ve varyans yerine budanmış ortalama ve winsorized varyans gibi dayanıklı konum ve değişkenlik parametreleriyle kullanılması varsayım bozulmasının söz konusu etkilerine karşı duyarsız test istatistikleri elde edilmesini mümkün kılmaktadır (Lix ve ark., 1998). Bilinen ortalama yerine budanmış ortalama kullanıldığında grupların kitle budanmış ortalamalarının karşılaştırılması için H_0 hipotezi,

$$H_0: \mu_{b1} = \mu_{b2} = \cdots = \mu_{bk} \tag{3.19}$$

biçiminde ifade edilir.

Normallik varsayımı bozulumuna karşı %20 budanmış ortalamanın kullanılmasının etkili bir yöntem olduğu hem teorik çalışmalarla hem de simülasyon çalışmalarıyla gösterilmiştir (Wilcox ve Keselman, 2001). Ayrıca budanmış ortalamayla birlikte bootstrap yöntemlerin kullanımının elde edilen sonuçlara etkisi de birçok çalışmada (örneğin; Wilcox ve ark. (1998), Wilcox ve Keselman (2001), Yılmaz ve Özdemir (2016), Özdemir ve ark. (2018)) araştırılmıştır.

3.2.1. ANOVA-F Testi

ANOVA-F testi için $\boldsymbol{H_0}$ hipotezi eşitlik (3.18)' de verildiği gibi kurulur ve aşağıda verilen adımlar izlenir:

1) $\boldsymbol{j}$. gruptan gözlem değerleri $\boldsymbol{y_{ij}}$ $(\boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{n_j}\,; \boldsymbol{j} = \mathbf{1}, \dots, \mathbf{k})$ ile gösterilmek üzere rasgele çekilen $\boldsymbol{n_j}$ büyüklüğündeki örneklemin ortalaması $\bar{\boldsymbol{y}}_{\cdot j}$ ve genel ortalamayı gösteren $\bar{\boldsymbol{y}}_{\cdot\cdot}$ sırasıyla Eşitlik (3.3) ve Eşitlik (3.20) ile hesaplanır:

$$\bar{\boldsymbol{y}}_{\cdot\cdot} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} y_{ij}}{N}. \tag{3.20}$$

Burada $\boldsymbol{N} = \sum_{j=1}^{k} \boldsymbol{n_j}$'dir.

2) $\boldsymbol{y_{ij}}$ gözlem değerlerinin genel ortalama $\bar{\boldsymbol{y}}_{\cdot\cdot}$'den sapmalarının kareleri toplamı $\boldsymbol{KT_T}$, gruplar arasındaki sapmaların kareleri toplamı $\boldsymbol{KT_{GA}}$ ve grup içindeki sapmaların kareleri toplamı $\boldsymbol{KT_{Gİ}}$,

$$\boldsymbol{KT_T} = \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(\boldsymbol{y_{ij}} - \bar{\boldsymbol{y}}_{\cdot\cdot}\right)^2 \tag{3.21}$$

$$\boldsymbol{KT_{GA}} = \sum_{j=1}^{k}\boldsymbol{n_j}\left(\bar{\boldsymbol{y}}_{\cdot j} - \bar{\boldsymbol{y}}_{\cdot\cdot}\right)^2 \tag{3.22}$$

$$\boldsymbol{KT_{Gİ}} = \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(\boldsymbol{y_{ij}} - \bar{\boldsymbol{y}}_{\cdot j}\right)^2 \tag{3.23}$$

olarak hesaplanır.

3) 2. adımda verilen eşitliklere ilişkin serbestlik dereceleri sırasıyla aşağıdaki gibi elde edilir:

$$\boldsymbol{v_T} = (\boldsymbol{N} - \mathbf{1}), \tag{3.24}$$

$$\boldsymbol{v_{GA}} = (\boldsymbol{k} - \mathbf{1}), \tag{3.25}$$

$$\boldsymbol{v_{Gİ}} = (\boldsymbol{N} - \boldsymbol{k}). \tag{3.26}$$

4) ANOVA-F test istatistiği ise,

$$\boldsymbol{F_h} = \frac{KT_{GA}/v_{GA}}{KT_{Gİ}/v_{Gİ}} \tag{3.27}$$

biçiminde tanımlanır. Eğer $\boldsymbol{F_h} \geq \boldsymbol{F}_{(v_{GA};v_{Gİ};1-\alpha)}$ ise $\boldsymbol{H_0}$ hipotezi reddedilir (Wilcox, 2003).

3.2.2. Welch Testi

Welch testi için $\boldsymbol{H_0}$ hipotezi eşitlik (3.18)'de verildiği gibi kurulur ve aşağıda verilen adımlar izlenir.

1) $\boldsymbol{j}$. gruptan gözlem değerleri $\boldsymbol{y_{ij}}$ $\left(\boldsymbol{i = 1, \dots, n_j \,; j = 1, \dots, \mathrm{k}}\right)$ ile gösterilmek üzere, rasgele çekilen $\boldsymbol{n_j}$ büyüklüğündeki örneklemin ortalaması $\boldsymbol{\bar{y}_{.j}}$ ve varyansı $\boldsymbol{s_j^2}$ sırasıyla Eşitlik (3.3) ve Eşitlik (3.4) ile hesaplanır.
2) Her bir grup için w_j ağırlıkları

$$w_j = \frac{n_j}{s_j^2} \tag{3.28}$$

olarak elde edilir.

3) $\tilde{y}$ ile gösterilmek üzere grup ortalamalarının ağırlıklı ortalaması aşağıdaki gibi hesaplanır:

$$\tilde{y} = \frac{1}{U} \sum_{j=1}^{k} w_j \, \bar{y}_{.j} \,. \tag{3.29}$$

Burada $U = \sum_{j=1}^{k} w_j$'dir.

4) Ardından Welch'in F test istatistiği F_w'nun hesabında kullanılmak üzere Eşitlik (3.30) ve Eşitlik (3.31) ile verilen A ve B değerleri hesaplanır:

$$A = \frac{1}{k-1} \sum_{j=1}^{k} w_j \left(\bar{y}_{.j} - \tilde{y}\right)^2, \tag{3.30}$$

$$B = \frac{2(k-2)}{k^2-1} \sum_{j=1}^{k} \frac{\left(1-\frac{w_j}{U}\right)^2}{n_j-1} \,. \tag{3.31}$$

5) Welch'in F test istatistiği F_w,

$$F_w = \frac{A}{B+1} \tag{3.32}$$

biçiminde tanımlanır.

6) Eşitlik (3.18)'deki H_0 hipotezi doğru iken F_w dağılımı yaklaşık olarak v_1 ve v_2 serbestlik dereceli F dağılımına sahiptir. Burada v_1 ve v_2

$$v_1 = k - 1, \tag{3.33}$$

$$v_2 = \left[\frac{3}{k^2-1}\sum_{j=1}^{k}\frac{\left(1-\frac{w_j}{U}\right)^2}{n_j-1}\right]^{-1} \tag{3.34}$$

şeklinde hesaplanır.

7) Eğer $F_w \geq F_{(v_1;v_2;1-\alpha)}$ ise H_0 hipotezi reddedilir (Wilcox, 2003).

3.2.3. Budanmış Ortalama ile Welch Testi

Varyansların heterojenliğine karşı dayanıklı olan Welch testini, bilinen ortalama ve varyans yerine dayanıklı konum ve değişkenlik parametrelerini kullanarak normallik varsayımı bozulmalarına karşı da dayanıklı hale getirebilmek mümkündür.

Budanmış ortalama ile Welch testi için $\boldsymbol{H_0}$ hipotezi Eşitlik (3.19)' da verildiği gibi kurulur ve aşağıda verilen adımlar izlenir.

1) j. gruptan gözlem değerleri y_{ij} $(i = 1, \dots, n_j\,;\, j = 1, \dots, \mathrm{k})$ ile gösterilmek üzere rasgele çekilen n_j büyüklüğündeki örneklemin budanmış ortalaması $\bar{y}_{bj}$ ile Yuen'in standart hata kare tahmini d_j sırasıyla Eşitlik (2.3) ve Eşitlik (3.14) kullanılarak hesaplanır.
2) Elde edilen d_j'nin tersinin alınmasıyla w_j ağırlık değeri

$$w_j = \frac{1}{d_j} \tag{3.35}$$

olarak hesaplanır.

3) $\tilde{y}$ ile gösterilmek üzere grupların budanmış ortalamalarının ağırlıklı ortalaması aşağıdaki gibi elde edilir:

$$\tilde{y} = \frac{1}{U}\sum_{j=1}^{k} w_j\, \bar{y}_{bj}\,. \tag{3.36}$$

4) Ardından Welch'in F test istatistiği F_{wb}'nin hesabında kullanılmak üzere Eşitlik (3.37) ve Eşitlik (3.38) ile verilen A ve B değerleri hesaplanır:

$$A = \frac{1}{k-1}\sum_{j=1}^{k} w_j\left(\bar{y}_{bj} - \tilde{y}\right)^2, \tag{3.37}$$

$$B = \frac{2(k-2)}{k^2-1}\sum_{j=1}^{k}\frac{\left(1-\frac{w_j}{U}\right)^2}{h_j-1}\,. \tag{3.38}$$

5) Welch'in F test istatistiği F_{wb} ,

$$F_{wb} = \frac{A}{B+1} \tag{3.39}$$

biçiminde tanımlanır.

6) Eşitlik (3.19)'daki H_0 hipotezi doğru iken F_{wb} yaklaşık olarak v_1 ve v_2 serbestlik dereceli F dağılımına sahiptir. Burada v_1 ve v_2 sırasıyla Eşitlik (3.40) ve (3.41) ile

$$v_1 = k - 1\,, \tag{3.40}$$

$$v_2 = \left[\frac{3}{k^2-1}\sum_{j=1}^{k}\frac{\left(1-\frac{w_j}{U}\right)^2}{h_j-1}\right]^{-1} \tag{3.41}$$

şeklinde hesaplanır.

7) Eğer $F_{wb} \geq F_{(v_1;v_2;1-\alpha)}$ ise H_0 hipotezi reddedilir (Wilcox, 2012).

3.2.4. Budanmış Ortalama ve Bootstrap-t ile Welch Testi

1.tip hata olasılığının kontrolü bakımından %20 budanmış ortalama ve bootstrap-t yönteminin Welch testi ile kullanımı iyi sonuçlar vermektedir (Özdemir ve Yılmaz, 2016).

Bootstrap-t yöntemi ile Eşitlik (3.39)'da tanımlanan $\boldsymbol{F_{wb}}$ test istatistiğinin dağılımının kestirilmesi ve Eşitlik (3.19)'daki $\boldsymbol{H_0}$ hipotezinin Welch testi için aşağıdaki adımlar izlenir.

1) j. gruptan gözlem değerleri y_{ij} $(i = 1, \dots, n_j\,; j = 1, \dots, \mathrm{k})$ ile gösterilmek üzere, rasgele çekilen n_j büyüklüğündeki örneklemden gözlem değerleri $y_{1\mathrm{j}}^*, y_{2\mathrm{j}}^*, \dots, y_{n_\mathrm{j}\mathrm{j}}^*$ şeklinde gösterilmek üzere bootstrap örneklem çekilir ve Eşitlik (3.42) hesaplanır:

$$D_{ij}^* = y_{\mathrm{ij}}^* - \bar{y}_{b\mathrm{j}}\,. \tag{3.42}$$

2) Eşitlik (3.39)'daki F_{wb} test istatistiği, D_{ij}^* değerlerine dayalı olarak hesaplanarak F_{wb}^* ile gösterilen test istatistiği elde edilir.
3) 1. ve 2. adımın M kez tekrarlanmasıyla F_{wb1}^* , F_{wb2}^* , ... , F_{wbM}^* değerleri elde edilir. Ardından bu değerler küçükten büyüğe sıralanır:

$F^{*}_{wb(1)} \leq F^{*}_{wb(2)} \leq \cdots \leq F^{*}_{wb(M)}$.

4) H_0 hipotezinin testi için kritik değerin sıra sayısı $a = (1-\alpha)M$ eşitliği ile hesaplanır ve elde edilen a değeri en yakın tamsayıya yuvarlanır.
5) Eğer $F_{wb} \geq F^{*}_{wb(a)}$ ise H_0 hipotezi reddedilir (Özdemir ve ark., 2018).

3.2.5. Budanmış Ortalama ile Bootstrap Yüzdelik Yöntemi

Schrader ve Hettmansperger (1980) ile He ve arkadaşları (1990) tarafından önerilen yüzdelik bootstrap yöntemi için H_0 hipotezi Eşitlik (3.19)'da verildiği gibidir ve test için aşağıdaki adımlar izlenir:

1) j. gruptan gözlem değerleri y_{ij} $\left(i = 1, \ldots, n_j\, ; j = 1, \ldots, \mathrm{k}\right)$ ile gösterilmek üzere rasgele çekilen n_j büyüklüğündeki örneklemin budanmış ortalaması $\bar{y}_{bj}$ hesaplanır ve elde edilen $\bar{y}_{bj}$ değerlerinin ortalaması,

$$\bar{y}_{b.} = \frac{1}{k}\sum_{j=1}^{k} \bar{y}_{b\mathrm{j}} \tag{3.43}$$

olarak elde edilir.

2) Test istatistiği H,

$$H = \frac{1}{N}\sum_{j=1}^{k} n_j\left(\bar{y}_{b\mathrm{j}} - \bar{y}_{b.}\right)^2 \tag{3.44}$$

biçiminde tanımlanır. Burada $\boldsymbol{N} = \sum_{j=1}^{k} \boldsymbol{n_j}$'dir.

3) j. grubun her bir gözlem değerinden grubun budanmış ortalaması $\bar{y}_{b\mathrm{j}}$ çıkarılarak D_{ij} değerleri elde edilir:

$D_{ij} = y_{ij} - \bar{y}_{b\mathrm{j}}$.

4) j. grup için 3. adım ile elde edilen $D_{1j}, D_{2j}, \ldots, D_{n_j j}$ değerlerinden bootstrap örneklem çekilir.
5) 3. ve 4. adımın k grup için uygulanmasıyla elde edilen k tane bootstrap örneklem kullanılarak H^{*} test istatistiği hesaplanır.
6) 4.ve 5. adımın M kere tekrarlanmasıya elde edilen $H_1^{*}, H_2^{*}, \ldots, H_M^{*}$ değerleri küçükten büyüğe sıralanır:

$H^*_{(1)} \leq H^*_{(2)} \leq \cdots \leq H^*_{(M)}$.

6) H_0 hipotezinin testi için kritik değerin sıra sayısı $a = (1-\alpha)M$ eşitliği ile hesaplanır ve elde edilen a değeri en yakın tamsayıya yuvarlanır.
7) Eğer $H \geq H^*_{(a)}$ ise H_0 hipotezi reddedilir (Wilcox, 2003).

3.2.6. Box Testi

Box testi için H_0 hipotezi Eşitlik (3.19)'da verildiği gibidir ve test için adımlar aşağıda verilmiştir:

1) j. gruptan gözlem değerleri y_{ij} $(i = 1, \dots, n_j\, ; j = 1, \dots, \mathrm{k})$ ile gösterilmek üzere rasgele çekilen n_j büyüklüğündeki örneklem için budanacak gözlem sayısı g_j , etkin örneklem büyüklüğü h_j , budanmış ortalama $\bar{y}_{bj}$ ve winsorized varyans s^2_{wj} sırasıyla Eşitlik (2.1), Eşitlik (2.2), Eşitlik (2.3) ve Eşitlik (2.6) kullanılarak elde edilir.

2) Eşitlik (3.45) ve Eşitlik (3.46) ile sırasıyla s^2_{jbox} ve $\bar{y}_{bbox}$ değerleri

$$s^2_{jbox} = \frac{(n_j-1)s^2_{wj}}{(h_j-1)}, \tag{3.45}$$

$$\bar{y}_{bbox} = \frac{\sum_{j=1}^{k} h_j \bar{y}_{\mathrm{bj}}}{H} \tag{3.46}$$

biçiminde hesaplanır. Burada $H = \sum_{j=1}^{k} h_j$'dir.

3) Test istatistiği F_{box} ,

$$F_{box} = \frac{\sum_{j=1}^{k} h_j \left(\bar{y}_{\mathrm{bj}} - \bar{y}_{bbox}\right)^2}{\sum_{j=1}^{k} 1-(h_j/H)s^2_{jbox}} \tag{3.47}$$

biçiminde tanımlanır.

4) Eşitlik (3.19)'daki H_0 hipotezi doğru iken F_{box} yaklaşık olarak

$$\hat{v}_1 = \frac{\left[\sum_{j=1}^{k}(1-f_j)s^2_{jbox}\right]^2}{\left(\sum_{j=1}^{k} s^2_{jbox} f_j\right)^2 + \sum_{j=1}^{k} s^4_{jbox}(1-2f_j)} \tag{3.48}$$

$$\hat{v}_2 = \frac{\left[\sum_{j=1}^{k}(1-f_j)s_{jbox}^2\right]^2}{\sum_{j=1}^{k} s_{jbox}^4 (1-f_j)^2/(h_j-1)} \tag{3.49}$$

serbestlik dereceli F dağılımına sahiptir. Burada $f_j = h_j/H$'dir.

5) Eğer $\boldsymbol{F_{box}} \geq \boldsymbol{F}_{(\boldsymbol{\alpha};\hat{\boldsymbol{v}}_1;\hat{\boldsymbol{v}}_2)}$ ise $\boldsymbol{H_0}$ hipotezi reddedilir (Wilcox, 2012).

4. BAĞIMLI GRUPLARIN KARŞILAŞTIRILMASINDA KULLANILAN YÖNTEMLER

Bu bölümde bağımlı iki ve ikiden fazla grubun konum parametrelerinden ortalama ve budanmış ortalama kullanılarak karşılaştırılması incelenmiştir.

4.1. Bağımlı İki Grubun Karşılaştırılmasında Kullanılan Yöntemler

Uygulamada bağımsız grupların karşılaştırılmasının yanı sıra bağımlı grupların karşılaştırılması gereken durumlar da mevcuttur. Bağımlı gruplar, aynı deneklerin farklı zamanlarda ya da farklı durumlarda uygulanan deneme sonucu elde edilen ölçüm sonuçlarını ifade etmektedir.

$(y_{11}, y_{12}), \ldots., (y_{n1}, y_{n2})$ şeklinde gösterilmek üzere n çift gözlemden oluşan bağımlı iki grubun ortalamaları arasında anlamlı bir farklılık olup olmadığı Eşitlik (4.1)'de verilen H_0 hipotezi ile test edilir:

$$H_0: \mu_D = 0 \,. \tag{4.1}$$

Bu hipotezin testinde en yaygın kullanılan yöntemlerden biri bağımlı örneklem t testidir. Testin uygulanabilmesi için normallik varsayımının sağlanması gerekmektedir. Test, gruplar özdeş dağılıma sahip olduğunda 1. tip hata bakımından iyi sonuçlar verirken, gruplar çarpıklık yönünden farklılık gösterdiğinde yanlı sonuçlar verebilmektedir. Bu problemin çözümüne ilişkin genel yaklaşım konum ölçüsü için budanmış ortalama gibi dayanıklı ölçümlerin kullanılmasıdır (Wilcox, 2003).

Bağımlı iki grubun budanmış ortalamaları arasında anlamlı bir farklılık olup olmadığı Eşitlik (4.2)'de verilen H_0 hipotezi

$$H_0: \mu_{b1} = \mu_{b2} \tag{4.2}$$

ile test edilir.

Eşitlik (4.2) ile verilen H_0 hipotezi, bağımsız iki grubun budanmış ortalamalarının karşılaştırılmasında kullanılan Yuen testi ile aşağıda yer verilen adımlar izlenerek test edilebilmektedir.

4.1.1. Bağımlı Örneklem t Testi

Bağımlı iki grup karşılaştırılmasında H_0 hipotezi Eşitlik (4.1)'de verildiği gibidir ve t testi için aşağıdaki adımlar izlenir:

1) y_{ij} $(i = 1, \dots, n\,; j = 1,2)$, j. grubun i. deneğine ait gözlem değeri olmak üzere (y_{i1}, y_{i2}) şeklindeki n tane gözlem çifti için D_i değerleri hesaplanır:

$$D_i = y_{i1} - y_{i2}\,. \tag{4.3}$$

2) D_i değerlerinin ortalaması $\bar{D}$ ve varyansı s_D^2 sırasıyla,

$$\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n}\,, \tag{4.4}$$

$$s_D^2 = \frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1} \tag{4.5}$$

şeklinde hesaplanır.

3) Test istatistiği t_D ,

$$t_D = \frac{\bar{D}}{s_D/\sqrt{n}} \tag{4.6}$$

biçiminde tanımlanır.

4) Eğer $|t_D| \geq t_{(v,1-\alpha/2)}$ ise H_0 hipotezi reddedilir. Burada $v = (n-1)$'dir.
5) Bağımlı iki grubun kitle ortalamaları arasındaki fark $(\mu_1 - \mu_2)$ için $(1-\alpha\,)$ güven düzeyine karşılık gelen güven aralığı ise,

$$\left(\boldsymbol{\bar{D} - t_{(v,1-\alpha/2)}\frac{s_D}{\sqrt{n}}\,, \bar{D} + t_{(v,1-\alpha/2)}\frac{s_D}{\sqrt{n}}}\right)$$

biçiminde tanımlanır (Wilcox, 2003).

4.1.2. Yuen Testi

Yuen testi için H_0 hipotezi Eşitlik (4.2)'de verildiği gibidir ve test için aşağıdaki adımlar izlenir:

1) Etkin örneklem büyüklüğü h, j. grubun x_{ij} ile gösterilen winsorized değerleri ve winsorized değerlerin ortalaması $\bar{x}_{.j}$ sırasıyla Eşitlik (2.2), Eşitlik (2.4) ve Eşitlik (2.5) ile elde edilir.
2) İki grubun budanmış ortalamaları arasındaki farkın standart hatasının karesinin Yuen yaklaşımıyla tahmini için Eşitlik (4.7) ve Eşitlik (4.8) ile verilen d_j ve d_{12} değerleri sırasıyla aşağıdaki gibi hesaplanır:

$$d_j = \frac{\sum_{i=1}^{n}\left(x_{ij}-\bar{x}_{.j}\right)^2}{h(h-1)}, \tag{4.7}$$

$$d_{12} = \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_{.1})(x_{i2}-\bar{x}_{.2})}{h(h-1)}. \tag{4.8}$$

3) Yuen'in test istatistiği t_y ,

$$t_y = \frac{\bar{y}_{b1}-\bar{y}_{b2}}{\sqrt{d_1+d_2-2d_{12}}} \tag{4.9}$$

biçiminde tanımlanır.

4) Eğer $|t_y| \geq t_{(v,1-\alpha/2)}$ ise H_0 hipotezi reddedilir. Burada $v = (h-1)$'dir.

Bağımlı iki grubun kitle budanmış ortalamaları arasındaki fark $(\mu_{b1} - \mu_{b2})$ için $(1-\alpha)$ güven düzeyine karşılık gelen güven aralığı ise,

$$(\bar{y}_{b1} - \bar{y}_{b2}) \pm t_{(v,1-\alpha/2)}\sqrt{d_1 + d_2 - 2d_{12}}$$

biçiminde tanımlanır (Wilcox, 2012).

4.2. Bağımlı k Grubun Karşılaştırılmasında Kullanılan Yöntemler

İkiden fazla bağımlı grubun karşılaştırıldığı deney düzenleri tekrarlı ölçümlü deney düzeni olarak da adlandırılmaktadır. k bağımlı grubun ortalamalarının eşitliğinin araştırılmasında Eşitlik (4.10)'da verilen H_0 hipotezi

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_k \tag{4.10}$$

şeklinde kurulur.

Bu deney düzeninde H_0 hipotezinin testinde genel olarak ANOVA kullanılmaktadır. Tekrarlı ölçümlü deney düzeninde ANOVA'nın uygulanabilmesi için aşağıdaki varsayımların sağlanması gerekmektedir (Alpar, 2006):

1. Normallik varsayımı,
2. Varyansların homojenliği varsayımı,
3. Küresellik varsayımı.

Küresellik, tekrarlı ölçüm deney düzenindeki ardışık seviyeler arasında farklara ilişkin varyansların homojen olması olarak ifade edilmektedir (Keskin ve Mendeş, 2001). Küresellik varsayımının sağlanıp sağlanmadığı küresellik katsayısı $\boldsymbol{\epsilon}$ değeri ile belirlenir. Bu varsayımın sağlanmadığı durumlar için serbestlik derecesi düzeltmelerine dayalı bazı yöntemler geliştirilmiştir. Huyn-Feldt düzeltmesi bu yöntemlere örnek olarak verilebilir. Huyn-Feldt için küresellik varsayımının bozulumu açısından $\boldsymbol{\epsilon = 0}$ en yüksek seviyede bozulmayı, $\boldsymbol{\epsilon = 1}$ ise en düşük seviyede bozulmayı ifade eder. Genellikle $\boldsymbol{\epsilon \geq 0,75}$ istenir (Alpar, 2006). Bunun dışında çok değişkenli test istatistikleri, çok değişkenli küreselliğin varsayımlarına bağlı olmayan istatistikler ile tek değişkenli ve çok değişkenli yaklaşımların kombinasyonlarını içeren karma modeller de küresellik varsayımının sağlanmadığı durumlarda kullanılmaktadır (Keselman ve ark., 2001).

Çok değişkenli analizler için küresellik gerekmemesine rağmen normalliğin sağlanması gerekmektedir ve yapılan çalışmalar bu analiz yöntemlerinin aşırı çarpıklığa karşı oldukça hassas olduğunu göstermektedir. Tek değişkenli yaklaşımlarda küresellik varsayımının yanı sıra normalliğin de sağlanmadığı durumlarda kullanılan çeşitli yaklaşımlar bulunmaktadır. Bunlardan biri de bilinen ortalama ve varyans-kovaryans yerine budanmış ortalama ve winsorized varyans-kovaryans kullanılmasıdır (Berkovits ve ark., 2000). Bu durumda $\boldsymbol{k}$ bağımlı grubun budanmış ortalamaları Eşitlik (4.11)'de verilen $\boldsymbol{H_0}$ hipotezi

$$H_0: \mu_{b1} = \mu_{b2} = \cdots = \mu_{bk} \tag{4.11}$$

ile test edilir.

Bunun yanı sıra bağımsız grupların karşılaştırılmasında olduğu gibi bağımlı grupların karşılaştırılmasında da testler bootstrap yöntemlerle birlikte kullanılabilmektedir (Örneğin; Berkovits ve ark. (2000), Wilcox ve ark. (2000)).

4.2.1. Tekrarlı Ölçümlü ANOVA-F Testi

Tekrarlı ölçümlü ANOVA-F testi için H_0 hipotezi Eşitlik (4.10)'da verildiği gibidir ve test için aşağıdaki adımlar izlenir:

1) $\boldsymbol{y_{ij}}$ $(\boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{n}\, ; \boldsymbol{j} = \mathbf{1}, \dots, \mathbf{k})$, $\boldsymbol{j}$. grubun $\boldsymbol{i}$. deneğine ait gözlem değeri olmak üzere $\boldsymbol{j}$. grubun gözlemleri toplamı $\boldsymbol{y_{.j}}$, $\boldsymbol{i}$. deneğe ait gözlem değerlerinin toplamı $\boldsymbol{y_{i.}}$ ve tüm gözlemlerin toplamı $\boldsymbol{y_{..}}$ aşağıdaki gibi hesaplanır:

$$\boldsymbol{y_{.j}} = \sum_{i=1}^{n} \boldsymbol{y_{ij}}\,, \tag{4.12}$$

$$\boldsymbol{y_{i.}} = \sum_{j=1}^{k} \boldsymbol{y_{ij}}\,, \tag{4.13}$$

$$\boldsymbol{y_{..}} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \boldsymbol{y_{ij}}\,. \tag{4.14}$$

2) $\boldsymbol{y_{ij}}$ gözlem değerlerinin genel ortalamadan sapmalarının kareleri toplamı $\boldsymbol{KT_{Toplam}}$, denekler arasındaki sapmaların kareleri toplamı $\boldsymbol{KT_{denek}}$, denemeler arasındaki sapmaların kareleri toplamı $\boldsymbol{KT_{deneme}}$ ve $\boldsymbol{KT_{hata}}$ sırasıyla,

$$\boldsymbol{KT_{Toplam}} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} y_{ij}^2 - \frac{(y_{..})^2}{N}, \tag{4.15}$$

$$\boldsymbol{KT_{deneme}} = \frac{\sum_{j=1}^{k} (y_{.j})^2}{n} - \frac{(y_{..})^2}{N}, \tag{4.16}$$

$$\boldsymbol{KT_{denek}} = \frac{\sum_{i=1}^{n} (y_{i.})^2}{k} - \frac{(y_{..})^2}{N}, \tag{4.17}$$

$$\boldsymbol{KT_{Hata}} = \boldsymbol{KT_{T}} - \boldsymbol{KT_{Denek}} - \boldsymbol{KT_{Deneme}} \tag{4.18}$$

şeklinde hesaplanır. Burada $\boldsymbol{N} = \sum_{j=1}^{k} \boldsymbol{n}$'dir.

3) 2.adımda verilen eşitliklere ilişkin serbestlik dereceleri,

$$v_{Toplam} = N - 1, \tag{4.19}$$

$$v_{denek} = n - 1, \tag{4.20}$$

$$v_{deneme} = k - 1, \tag{4.21}$$

$$v_{hata} = (n - 1)(k - 1), \tag{4.22}$$

olarak tanımlanır.

4) Tekrarlı ölçümlü ANOVA-F test istatistiği,

$$\boldsymbol{F_h} = \frac{\boldsymbol{KT_{deneme}/v_{deneme}}}{\boldsymbol{KT_{hata}/v_{hata}}} \tag{4.23}$$

biçiminde tanımlanır. Eğer $F_h \geq F_{(v_{deneme};v_{hata};1-\alpha)}$ ise H_0 hipotezi reddedilir (Alpar, 2006).

4.2.2. Budanmış Ortalama ile Huynh-Feldt($\tilde{\epsilon}$) Düzeltmeli ANOVA-F Testi

Budanmış ortalama ile Huynh-Feldt($\tilde{\epsilon}$)-düzeltmeli ANOVA-F testi**,** tekrarlı ölçümlü ANOVA-F testinde küresellik varsayımının sağlanmadığı durumda ortalamaların karşılaştırılmasında önerilen Huynh-Feldt yönteminin budanmış ortalamalar için genelleştirilmiş halidir (Wilcox, 2012).

Budama oranı $\gamma = 0$ olduğunda bu yöntem ortalamaların karşılaştırılmasında kullanılan $\tilde{\epsilon}$-düzeltmeli ANOVA-F testine indirgenmektedir (Wilcox ve ark.,2000).

Budanmış ortalama ile $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi için H_0 hipotezi Eşitlik (4.11)'de verildiği gibidir ve test için aşağıdaki adımlar izlenir:

1) j. grubun budanmış ortalaması $\bar{y}_{bj}$ ve budanmış ortalamaların ortalaması $\bar{y}_{b.}$ sırasıyla Eşitlik (2.3) ve Eşitlik (3.43) ile hesaplanır.
2) j. grubun x_{ij} ile gösterilen winsorized değerleri elde edilir. Elde edilen x_{ij} değerleri üzerinden j. grubun ortalaması $\bar{x}_{.j}$, i. deneğe ait gözlem değerlerinin ortalaması $\bar{x}_{i.}$ ve genel ortalamayı gösteren $\bar{x}_{..}$ sırasıyla Eşitlik (2.5), Eşitlik (4.24) ve Eşitlik (4.25) ile hesaplanır:

$$\bar{x}_{i.} = \frac{1}{k}\sum_{j=1}^{k} x_{ij}, \tag{4.24}$$

$$\bar{x}_{..} = \frac{1}{N}\sum_{j=1}^{k}\sum_{i=1}^{n} x_{ij}\ . \tag{4.25}$$

Burada $\boldsymbol{N = nk}$'dır.

3) $j = 1, \dots k\ ; l = 1, \dots k$ olmak üzere winsorized kovaryans matrisi $V = (v_{jl})$,

$$v_{jl} = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_{ij} - \bar{x}_{.j}\right)(x_{il} - \bar{x}_{.l}) \tag{4.26}$$

şeklinde elde edilir.

Burada v_{jl}, $j = l$ iken s_{wj}^2 , $j \neq l$ iken v_{jl}'dir.

4) Eşitlik (4.27) ve Eşitlik (4.28) hesaplanır:

$$Q_c = (n - 2g) \sum_{j=1}^{k} \left(\bar{y}_{bj} - \bar{y}_{b.}\right)^2, \tag{4.27}$$

$$Q_e = \sum_{j=1}^{k} \sum_{i=1}^{n} \left(x_{ij} - \bar{x}_{.j} - \bar{x}_{i.} + \bar{x}_{..}\right)^2. \tag{4.28}$$

5) Test istatistiği F_h,

$$F_h = \frac{R_c}{R_e} \tag{4.29}$$

biçiminde tanımlanır. Burada $\boldsymbol{R_c}$ ve $\boldsymbol{R_e}$ aşağıdaki gibi hesaplanır:

$$R_c = \frac{Q_c}{k-1}, \tag{4.30}$$

$$\boldsymbol{R_e} = \frac{\boldsymbol{Q_e}}{(\boldsymbol{n}-\boldsymbol{2g}-\boldsymbol{1})(\boldsymbol{k}-\boldsymbol{1})}. \tag{4.31}$$

6) Serbestlik derecesi v_1 ve v_2'nin elde edilebilmesi için v_{jl} değerleri kullanılarak aşağıdaki eşitlikler hesaplanır:

$$\bar{v}_{..} = \frac{1}{k^2} \sum_{j=1}^{k} \sum_{l=1}^{k} v_{jl} , \tag{4.32}$$

$$\bar{v}_d = \frac{1}{k} \sum_{j=1}^{k} v_{jj} , \tag{4.33}$$

$$\bar{v}_{j.} = \frac{1}{k} \sum_{l=1}^{k} v_{jl} , \tag{4.34}$$

$$A = \frac{k^2 (\bar{v}_d - \bar{v}_{..})^2}{k-1}, \tag{4.35}$$

$$B = \sum_{j=1}^{k} \sum_{l=1}^{k} v_{jl}^2 - 2k \sum_{j=1}^{k} \bar{v}_{j.}^2 + k^2 \bar{v}_{..}^2 , \tag{4.36}$$

$$\hat{\epsilon} = \frac{A}{B}, \tag{4.37}$$

$$\tilde{\epsilon} = \frac{n(k-1)\hat{\epsilon} - 2}{(k-1)(n-1-(k-1)\hat{\epsilon})} . \tag{4.38}$$

7) Serbestlik derecesi v_1 ve v_2 olmak üzere,

$v_1 = (k-1)\tilde{\epsilon}\,,$ (4.39)

$v_2 = (k-1)(n-2g-1)\tilde{\epsilon}$ (4.40)

biçiminde hesaplanır.

8) Eğer $F_h \geq F_{(v_1;v_2;1-\alpha)}$ ise H_0 hipotezi reddedilir (Wilcox, 2000).

4.2.3. Budanmış Ortalama ve Bootstrap-t ile ANOVA- F Testi

Eşitlik (4.11)'deki H_0 hipotezinin Eşitlik (4.29)'da verilen F test istatistiği ile test edilmesinde kullanılacak kritik değer ve testin adımları aşağıda verilmiştir.

Budama oranı $\gamma = 0$ olduğunda bu yöntem ortalamaların karşılaştırılmasında kullanılan bootstrap-t ile ANOVA-F testine dönüşmektedir (Wilcox ve ark.,2000).

1) $\boldsymbol{j}$. grubun $\boldsymbol{i}$. deneğine ait gözlem değeri $\boldsymbol{y_{ij}}$'den grubun budanmış ortalaması $\boldsymbol{\bar{y}_{b\mathrm{j}}}$ çıkarılır:

$D_{ij} = y_{ij} - \bar{y}_{b\mathrm{j}}\,.$

2) j. grup için elde edilen $D_{1j},\dots,D_{nj}$ değerlerinden gözlem değerleri $D^*_{1j},\dots,D^*_{nj}$ ile gösterilmek üzere n büyüklüğünde bootstrap örneklem çekilir.
3) Elde edilen D^*_{ij} değerleri üzerinden F^* ile gösterilmek üzere Eşitlik (4.29)'da verilen F_h test istatistiği hesaplanır.
4) 2. ve 3. adımın M kez tekrar edilmesiyle F^*_1 , F^*_2 , … , F^*_M değerleri elde edilir.
5) Elde edilen F^* değerleri küçükten büyüğe sıralanır:

$F^*_{(1)} \leq F^*_{(2)} \leq \cdots \leq F^*_{(M)}\,.$

6) H_0 hipotezinin testi için kritik değerin sıra sayısı $a = (1-\alpha)M$ eşitliği ile hesaplanır ve elde edilen a değeri en yakın tamsayıya yuvarlanır.
7) Eğer $F_h \geq F^*_{(a)}$ ise H_0 hipotezi reddedilir. Burada F_h Eşitlik (4.29) ile y_{ij} değerlerine dayalı olarak hesaplanmış değerdir (Wilcox, 2012).

5. SİMÜLASYON ÇALIŞMASI

Bu bölümde, bağımsız ve bağımlı grup durumunda ortalama ve budanmış ortalamaların hipotez testinde kullanılan yöntemlerin karşılaştırılması için farklı deney düzenleri oluşturulmuş ve incelenen deney düzenlerinde yöntemlere ait gerçekleşen 1. tip hata oranları elde edilerek yöntemlerin varsayım bozulmalarına karşı dayanıklılıkları incelenmiştir.

5.1. Bağımsız Grup Durumu İçin Simülasyon Çalışması

Bağımsız iki grubun karşılaştırılmasında kullanılan yöntemlerden

- Student-t testi (t),
- Welch testi (W),
- Yuen testi (Y),
- Bootstrap-t ile Yuen testi (YB);

ve bağımsız k grubun karşılaştırılmasında kullanılan yöntemlerden

- ANOVA-F testi (F),
- Welch testi (W),
- Budanmış ortalama ile Welch testi (WB) ve
- Budanmış ortalama ve bootstrap-t ile Welch testi (WBB)

farklı deney düzenleri altında gerçekleşen 1. tip hata oranları bakımından karşılaştırılmıştır. R(4.0.2) programlama dili kullanılarak yapılan simülasyon çalışmasında nominal anlamlılık düzeyi $\boldsymbol{\alpha = 0,05}$, budama oranı $\boldsymbol{\gamma = \%20}$ olarak seçilmiştir. Her bir deney düzeninde 10.000 tekrar yapılmış ve bootstrap-t yönteminin kullanıldığı testler için $\boldsymbol{M}$ değeri 599 alınmıştır. Simülasyon çalışmasında kullanılan R paketleri ve fonksiyonları Ek-1'de verilmiştir.

5.1.1. Bağımsız Grup Durumu İçin Simülasyon Düzeni

Simülasyon çalışmasında kullanılan deney düzenleri; $\boldsymbol{k = 2}$, $\boldsymbol{k = 4}$ ve $\boldsymbol{k = 6}$ için örneklem büyüklüklerinin eşit olup olmadığı, normallik ve varyansların homojenliği varsayımlarının sağlanıp sağlanmadığı durumlar ve örneklem büyüklükleri ile varyansların pozitif ve negatif eşleşmesi durumları[3] dikkate alınarak oluşturulmuştur.

Örneklem büyüklükleri ile varyansların pozitif eşleşmesi örneklem büyüklüğü en büyük olan gruba karşılık gelen kitlenin varyansının da en büyük; örneklem büyüklüğü en küçük olan gruba karşılık gelen kitlenin varyansının da en küçük olması anlamına gelmektedir. Örneklem büyüklükleri ile varyansların negatif eşleşmesi ise örneklem büyüklüğü en büyük olan gruba karşılık gelen kitlenin varyansının en küçük; örneklem büyüklüğü en küçük olan gruba karşılık gelen kitlenin varyansının ise en büyük olması anlamına gelmektedir.

Kullanılan örneklem büyüklükleri ve varyans değerleri, grup sayıları $\boldsymbol{k = 2}$, $\boldsymbol{k = 4}$ ve $\boldsymbol{k = 6}$ için sırasıyla Çizelge 1, Çizelge 2 ve Çizelge 3'te gösterilmektedir.

Çizelge 1. $k = 2$ için deney düzenleri

Örneklem Büyüklüğü	Varyans	
	Homojen Varyans	Heterojen Varyans
20-20	1-1	1-16
		1-36
10-20	1-1	1-16 ; 16-1
		1-36 ; 36-1

Çizelge 2. $k = 4$ için deney düzenleri

Örneklem Büyüklüğü	Varyans	
	Homojen Varyans	Heterojen Varyans
20-20-20-20	1-1-1-1	1-1-1-36
		1-4-9-16
10-15-20-25	1-1-1-1	1-1-1-36 ; 36-1-1-1
		1-4-9-16 ; 16-9-4-1
15-20-25-30	1-1-1-1	1-1-1-36 ; 36-1-1-1
		1-4-9-16 ; 16-9-4-1

[3] Örneklem büyüklükleri ile varyansların pozitif ve negatif eşleşmesi durumları için bundan sonra kısaca "pozitif eşleşme" ve "negatif eşleşme" ifadesi kullanılacaktır.

Çizelge 3. $k = 6$ için deney düzenleri

Örneklem Büyüklüğü	Varyans	
	Homojen Varyans	Heterojen Varyans
20-20-20-20-20-20	1-1-1-1-1-1	1-1-1-1-1-36
		1-1-4-9-9-16
10-15-15-20-20-25	1-1-1-1-1-1	1-1-1-1-1-36; 36-1-1-1-1-1
		1-1-4-9-9-16; 16-9-9-4-1-1
15-20-20-25-25-30	1-1-1-1-1-1	1-1-1-1-1-36; 36-1-1-1-1-1
		1-1-4-9-9-16; 16-9-9-4-1-1

Normallik varsayımının sağlanmadığı deney düzenleri oluşturmak için $g\&h$ dağılımı (Hoaglin, 1985) kullanılmıştır. $g\&h$ dağılımında g parametresi çarpıklığın miktarını ve yönünü, h parametresi ise basıklığı kontrol etmektedir. $g\&h$ dağılımı $g = h = 0$ iken standart normal dağılıma karşılık gelmektedir. g parametresinin değeri arttıkça dağılımın çarpıklığı artmakta, h parametresinin değeri arttıkça da dağılım daha kuyruklu hale gelmektedir.

$g\&h$ dağılımından veri üretebilmek için öncelikle, standart normal dağılıma sahip Z raslantı değişkeni üretilmektedir. Ardından, Z raslantı değişkeninin aldığı değerler Z_{ij} $(i = 1, \dots, n_j\, ; j = 1, \dots, k)$ ile gösterilmek üzere Z_{ij} ile g ve h parametre değerlerinin, Eşitlik (5.1)'de yerine konmasıyla $g\&h$ dağılımına sahip ε_{ij} elde edilmektedir:

$$\varepsilon_{ij} = \frac{\exp(gZ_{ij})-1}{g} exp\left(\frac{hZ_{ij}^2}{2}\right), \tag{5.1}$$

$g = 0$ olduğunda ε_{ij} ,

$$\varepsilon_{ij} = Z_{ij} * exp\left(\frac{hZ_{ij}^2}{2}\right), \tag{5.2}$$

biçiminde elde edilmektedir.

$g = 0$ iken $g\&h$ dağılımı 0 ortalamaya sahip olduğundan varyansların homojen olduğu ve homojen olmadığı düzenler oluşturulurken her bir ε_{ij}'nin σ_j ile çarpılması Eşitlik (3.18) ve Eşitlik (3.19)'da yer alan hipotezleri etkilemeyecektir. Ancak $g > 0$ iken $g\&h$ dağılımı, Eşitlik (5.3)'te yer verilen ortalamaya sahiptir.

$$\mu_{gh} = \frac{exp\left(\frac{g^2}{2(1-h)}\right)-1}{g(1-h)^{1/2}} \tag{5.3}$$

Varyansların homojen olduğu ve homojen olmadığı düzenler oluşturulurken ε_{ij}'den μ_{gh}'nin çıkarılması ile elde edilen sonucun σ_j ile çarpılması sonucunda yeni gözlem değerleri aşağıdaki gibi elde edilir:

$$y_{ij} = \mu_j + \sigma_j(\varepsilon_{ij} - \mu_{gh})\,. \tag{5.4}$$

Budanmış ortalamaların karşılaştırıldığı Eşitlik (3.19) durumunda ise Eşitlik (5.5)'te gösterildiği üzere yeni gözlem değerleri,

$$y_{ij} = \mu_j + \sigma_j(\varepsilon_{ij} - \mu_{bgh}) \tag{5.5}$$

ile elde edilmektedir (Wilcox, 1994).

Wilcox (1994)'ın çalışmasında, çeşitli g ve h parametre değerlerine karşılık gelen μ_{gh}, μ_{bgh}, çarpıklık ve basıklık katsayısının yer aldığı tablodan yararlanarak simülasyon çalışmasında kullanılan g ve h parametre değerleri ile bu parametre değerlerine karşılık gelen μ_{gh}, μ_{bgh} ile dağılımın şekline ilişkin tanımlamaya Çizelge 4'te yer verilmiştir.

Çizelge 4. Bağımsız grup simülasyon çalışmasında kullanılan g ve h parametre değerleri

g	h	μ_{gh}	μ_{bgh} (%20)	Dağılım tanımlaması
0	0	0	0	Standart normal
0,5	0,2	0,3782	0,0564	Kuyruklu çarpık
0,5	0,5	0,8033	0,0600	Aşırı kuyruklu çarpık
1,0	0,0	0,6487	0,1110	Hafif kuyruklu aşırı çarpık
1,0	0,5	2,4300	0,1231	Aşırı kuyruklu aşırı çarpık

5.1.2. Bağımsız Grup Durumu İçin Simülasyon Sonuçları

Simülasyon çalışmasındaki her bir tekrar; belirlenen deney düzeni çerçevesinde veri üretimini ve üretilen veri üzerinden Eşitlik (3.1), Eşitlik (3.2), Eşitlik (3.18) ve Eşitlik (3.19)'da yer alan hipotezlerin testinde kullanılan yöntemlerin p değerlerinin elde edilmesi sürecini içermektedir.

Tekrar sayısı 10.000 seçilen simülasyon çalışmasında karşılaştırılan testlere ait gerçekleşen 1.tip hata oranı olarak, 10.000 tekrar içinde 0,05 anlamlılık düzeyinde reddedilen hipotezlerin oranı kullanılmıştır (Wilcox,1994).

Testlere ait gerçekleşen 1. tip hata oranlarına göre testin varsayım bozulmalarına karşı dayanıklı olup olmadığını belirlemek için Bradley (1978) tarafından, liberal ölçüt olarak adlandırılan dayanıklılık ölçütü önerilmiştir. Buna göre teste ait gerçekleşen 1. tip hata oranı $\hat{\alpha}$'nın $0,5\alpha \leq \hat{\alpha} \leq 1,5\alpha$ aralığında kalması durumunda testin dayanıklı olduğu şeklinde yorum yapılabilmektedir. Bu çalışmada, dayanıklılık ölçütü olarak birçok çalışmada (örneğin; Najdi ve Ahad (2019), Oberfeld ve Franke (2013), Cribbie ve ark. (2012), Wilcox ve ark. (2000)) kullanıldığı görülen $0,5\alpha \leq \hat{\alpha} \leq 1,5\alpha$ aralığı tercih edilmiştir. Buna göre, nominal anlamlılık düzeyi α'nın 0,05 olarak alındığı bu çalışmada testin dayanıklılığına ilişkin olarak, gerçekleşen 1. tip hata oranı $0,025 \leq \hat{\alpha} \leq 0,075$ aralığında olduğunda dayanıklı, $0,025$ 'den küçük olduğunda tutucu, $0,075$'ten büyük olduğunda ise liberal sonuç verdiği şeklinde yorum yapılmıştır.

Normal dağılım ve farklı parametre değerleriyle $g\&h$ dağılımı altında $k = 2$ durumunda Student-t testi (t), Welch testi (W), Yuen testi (Y) ve bootstrap- t ile Yuen testi (YB); $k = 4$ ve $k = 6$ durumunda ANOVA-F testi (F), Welch testi (W), budanmış ortalama ile Welch testi (WB) ve budanmış ortalama ve bootstrap- t ile Welch testi (WBB) için farklı deney düzenlerinde elde edilen gerçekleşen 1.tip hata oranlarına Çizelge 5-19'da yer verilmiştir. Çizelgelerde altı çizili değerler Bradley'in dayanıklılık ölçütüne göre tutucu sonuçları, koyu yazılmış değerler ise liberal sonuçları göstermektedir.

$k = 2$ durumunda normal dağılım için elde edilen sonuçların yer aldığı Çizelge 5 incelendiğinde t testi hariç diğer testlerin dayanıklı sonuçlar verdiği, bir başka deyişle 1. tip hata oranı kontrolünü sağladığı; t testinin ise örneklem büyüklüklerinin farklı ve varyansların homojen olmadığı durumda dayanıklı olmadığı görülmektedir.

Çizelge 5. Bağımsız $k = 2$ durumunda normal dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	t	W	Y	YB
20-20	1-1	0,0481	0,0480	0,0494	0,0483
	1-16	0,0512	0,0469	0,0547	0,0488
	1-36	0,0539	0,0481	0,0525	0,0469
10-20	1-1	0,0493	0,0471	0,0544	0,0532
	1-16	<u>0,0112</u>	0,0473	0,0540	0,0490
	16-1	**0,1599**	0,0500	0,0588	0,0519
	1-36	<u>0,0101</u>	0,0481	0,0528	0,0492
	36-1	**0,1820**	0,0528	0,0591	0,0481

Kuyruklu çarpık olarak tanımlanan $g\&h(g = 0{,}5\ h = 0{,}2)$ dağılım için gerçekleşen 1. tip hata oranlarına Çizelge 6'da yer verilmiştir. Sonuçlara göre budanmış ortalamaya dayalı Y ve YB testi 1. tip hata oranı kontrolünü her durumda sağlamıştır. t ve W testi ise durumların çoğunda liberal sonuç vermiştir.

Çizelge 6. Bağımsız $k = 2$ durumunda $g\&h\ (g = 0{,}5\ h = 0{,}2)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	t	W	Y	YB
20-20	1-1	0,0418	0,0390	0,0412	0,0428
	1-16	**0,0852**	**0,0807**	0,0544	0,0528
	1-36	**0,0948**	**0,0887**	0,0527	0,0501
10-20	1-1	0,0408	0,0407	0,0428	0,0459
	1-16	0,0341	0,0703	0,0499	0,0521
	16-1	**0,1890**	**0,0938**	0,0621	0,0589
	1-36	0,0342	**0,0863**	0,0496	0,0496
	36-1	**0,2164**	**0,0979**	0,0596	0,0526

Aşırı kuyruklu çarpık dağılıma karşılık gelen $g\&h\ (g = 0{,}5\ h = 0{,}5)$ dağılım için gerçekleşen 1. tip hata oranlarının yer aldığı Çizelge 7 incelendiğinde sonuçların Çizelge 6'daki sonuçlarla benzer olduğu, budanmış ortalamaya dayalı Y ve YB testinin 1. tip hata oranı kontrolünü her durumda sağladığı, t ve W testinin ise varyansların homojen olduğu durumlar hariç durumların tamamında liberal sonuç verdiği görülmüştür.

Çizelge 7. Bağımsız $k = 2$ durumunda $g\&h$ $(g = 0,5\ h = 0,5)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	t	W	Y	YB
20-20	1-1	0,0313	0,0290	0,0371	0,0427
	1-16	**0,1628**	**0,1571**	0,0498	0,0539
	1-36	**0,1878**	**0,1804**	0,0491	0,0516
10-20	1-1	0,0347	0,0309	0,0379	0,0422
	1-16	**0,0805**	**0,1299**	0,0436	0,0496
	16-1	**0,2606**	**0,1683**	0,0487	0,0491
	1-36	**0,0960**	**0,1691**	0,0483	0,0525
	36-1	**0,2996**	**0,1788**	0,0461	0,0435

Hafif kuyruklu aşırı çarpık olarak tanımlanan $g\&h$ $(g = 1\ h = 0)$ dağılım için gerçekleşen 1. tip hata oranlarına Çizelge 8'de yer verilmiştir. Sonuçlara göre t ve W testi çoğunlukla liberal sonuçlar verirken budanmış ortalamaya dayalı Y testi negatif eşleşmenin olduğu durumlar hariç 1. tip hata oranı kontrolünü her durumda sağlamıştır. Budanmış ortalama ve bootstrap-t yöntemine dayalı YB testi ise negatif eşleşmenin olduğu 1 durum dışında 1. tip hata oranı kontrolünü sağlamıştır.

Çizelge 8. Bağımsız $k = 2$ durumunda $g\&h$ $(g = 1\ h = 0)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	t	W	Y	YB
20-20	1-1	0,0359	0,0326	0,0355	0,0356
	1-16	**0,1244**	**0,1207**	0,0697	0,0638
	1-36	**0,1323**	**0,1258**	0,0664	0,0593
10-20	1-1	0,0399	0,0486	0,0470	0,0472
	1-16	0,0654	**0,1015**	0,0601	0,0594
	16-1	**0,2214**	**0,1450**	**0,0832**	**0,0760**
	1-36	0,0673	**0,1139**	0,0626	0,0569
	36-1	**0,2534**	**0,1551**	**0,0820**	0,0704

$k = 2$ için yapılan simülasyon çalışmasıyla ilgili son olarak Çizelge 9'da, aşırı kuyruklu aşırı çarpık olarak tanımlanan $g\&h$ $(g = 1\ h = 0,5)$ dağılım için gerçekleşen 1. tip hata oranlarına yer verilmiştir. Sonuçlara bakıldığında t ve W testinin varyansların heterojen olduğu durumların tamamında liberal sonuç verdiği, budanmış ortalamaya dayalı Y testinin tüm durumlarda 1. tip hata oranı kontrolünü sağladığı, budanmış ortalamaya ve bootstrap-t yöntemine dayalı YB testinin ise negatif eşleşmenin olduğu 1 durum hariç 1. tip hata oranı kontrolünü sağladığı görülmektedir.

Çizelge 9. Bağımsız $k = 2$ durumunda $g\&h$ $(g = 1\ h = 0{,}5)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	t	W	Y	YB
20-20	1-1	0,0218	0,0192	0,0340	0,0374
	1-16	**0,3848**	**0,3786**	0,0569	0,0622
	1-36	**0,4151**	**0,4093**	0,0613	0,0648
10-20	1-1	0,0271	0,0253	0,0305	0,0353
	1-16	**0,2985**	**0,3500**	0,0484	0,0550
	16-1	**0,4595**	**0,4097**	0,0750	**0,0797**
	1-36	**0,3241**	**0,3935**	0,0521	0,0573
	36-1	**0,5175**	**0,4470**	0,0701	0,0697

$k = 4$ ve $k = 6$ durumunda normal dağılım için elde edilen sonuçların yer aldığı Çizelge 10 ve Çizelge 11 incelendiğinde F testi hariç diğer testlerin dayanıklı sonuçlar verdiği görülmektedir. F testinin ise varyanslar homojenken her durumda 1. tip hata oranı kontrolünü sağladığı, ancak varyanslar heterojenken, özellikle de negatif eşleşme durumlarında liberal sonuç verme eğiliminde olduğu görülmektedir.

Çizelge 10. Bağımsız $k = 4$ durumunda normal dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	F	W	WB	WBB
20-20-20-20	1-1-1-1	0,0503	0,0473	0,0554	0,0427
	1-1-1-36	**0,1088**	0,0529	0,0573	0,0440
	1-4-9-16	0,0692	0,0496	0,0607	0,0450
10-15-20-25	1-1-1-1	0,0465	0,0473	0,0594	0,0464
	1-1-1-36	0,0398	0,0509	0,0583	0,0496
	36-1-1-1	**0,2469**	0,0499	0,0650	0,0368
	1-4-9-16	0,0281	0,0494	0,0583	0,0504
	16-9-4-1	**0,1480**	0,0518	0,0640	0,0396
15-20-25-30	1-1-1-1	0,0542	0,0539	0,0622	0,0508
	1-1-1-36	0,0472	0,0508	0,0551	0,0461
	36-1-1-1	**0,2006**	0,0502	0,0560	0,0397
	1-4-9-16	0,0319	0,0496	0,0522	0,0449
	16-9-4-1	**0,1294**	0,0518	0,0593	0,0407

Çizelge 11. Bağımsız $k = 6$ durumunda normal dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	*F*	*W*	*WB*	*WBB*
20-20-20-20-20-20	1-1-1-1-1-1	0,0490	0,0493	0,0595	0,0363
	1-1-1-1-1-36	**0,1289**	0,0485	0,0598	0,0363
	1-1-4-9-9-16	**0,0766**	0,0509	0,0637	0,0378
10-15-15-20-20-25	1-1-1-1-1-1	0,0486	0,0545	0,0670	0,0389
	1-1-1-1-1-36	0,0592	0,0482	0,0646	0,0385
	36-1-1-1-1-1	**0,2572**	0,0488	0,0682	0,0290
	1-1-4-9-9-16	0,0356	0,0539	0,0670	0,0451
	16-9-9-4-1-1	**0,1580**	0,0539	0,0722	0,0305
15-20-20-25-25-30	1-1-1-1-1-1	0,0507	0,0493	0,0580	0,0390
	1-1-1-1-1-36	**0,0753**	0,0536	0,0611	0,0393
	36-1-1-1-1-1	**0,2260**	0,0510	0,0611	0,0352
	1-1-4-9-9-16	0,0408	0,0549	0,0607	0,0434
	16-9-9-4-1-1	**0,1382**	0,0518	0,0618	0,0333

$g\&h$ $(g = 0{,}5\ h = 0{,}2)$ dağılım için gerçekleşen 1. tip hata oranlarına Çizelge 12 ve Çizelge 13'te yer verilmiştir. Sonuçlara göre budanmış ortalamaya dayalı WB ve WBB testi 1. tip hata oranı kontrolünü her durumda sağlamıştır. F testi, normal dağılım altındaki sonuçlara benzer olarak varyanslar homojenken her durumda 1. tip hata kontrolünü sağlamış, ancak varyanslar heterojenken liberal sonuç verme eğiliminde olmuştur. W testi ise normal dağılım altında elde edilen sonuçların aksine dağılımın çarpık ve kuyruklu hale gelmesiyle birlikte varyanslar heterojenken çoğunlukla liberal sonuç vermiştir.

Çizelge 12. Bağımsız $k = 4$ durumunda $g\&h$ $(g = 0{,}5\ h = 0{,}2)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	*F*	*W*	*WB*	*WBB*
20-20-20-20	1-1-1-1	0,0371	0,0424	0,0430	0,0354
	1-1-1-36	**0,1442**	0,0708	0,0529	0,0426
	1-4-9-16	0,0737	**0,0874**	0,0564	0,0450
10-15-20-25	1-1-1-1	0,0411	0,0477	0,0443	0,0387
	1-1-1-36	0,0735	0,0677	0,0479	0,0424
	36-1-1-1	**0,2780**	**0,0823**	0,0598	0,0401
	1-4-9-16	0,0350	0,0716	0,0451	0,0433
	16-9-4-1	**0,1376**	**0,0951**	0,0628	0,0396
15-20-25-30	1-1-1-1	0,0431	0,0501	0,0499	0,0399
	1-1-1-36	**0,0827**	0,0654	0,0490	0,0427
	36-1-1-1	**0,2304**	**0,0817**	0,0550	0,0418
	1-4-9-16	0,0357	**0,0768**	0,0502	0,0465
	16-9-4-1	**0,1217**	**0,0977**	0,0586	0,0396

Çizelge 13. Bağımsız $k = 6$ durumunda $g\&h$ $(g = 0{,}5\ h = 0{,}2)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	F	W	WB	WBB
20-20-20-20-20-20	1-1-1-1-1-1	0,0379	0,0494	0,0485	0,0297
	1-1-1-1-1-36	**0,1620**	0,0660	0,0515	0,0336
	1-1-4-9-9-16	0,0713	**0,0952**	0,0571	0,0357
10-15-15-20-20-25	1-1-1-1-1-1	0,0445	0,0544	0,0510	0,0300
	1-1-1-1-1-36	**0,0988**	0,0678	0,0573	0,0337
	36-1-1-1-1-1	**0,2860**	**0,0761**	0,0615	0,0275
	1-1-4-9-9-16	0,0349	**0,0790**	0,0521	0,0339
	16-9-9-4-1-1	**0,1455**	**0,1132**	0,0699	0,0332
15-20-20-25-25-30	1-1-1-1-1-1	0,0424	0,0541	0,0478	0,0324
	1-1-1-1-1-36	**0,1043**	0,0678	0,0521	0,0338
	36-1-1-1-1-1	**0,2491**	0,0743	0,0549	0,0353
	1-1-4-9-9-16	0,0409	**0,0818**	0,0495	0,0359
	16-9-9-4-1-1	**0,1183**	**0,1021**	0,0594	0,0369

$g\&h$ $(g = 0{,}5\ h = 0{,}5)$ dağılım için gerçekleşen 1. tip hata oranlarının yer aldığı Çizelge 14 ve Çizelge 15 incelendiğinde $k = 4$ durumu için budanmış ortalamaya dayalı WB ve WBB testlerinin 1. tip hata kontrolünü her durumda sağladığı; F ve W testinin ise varyanslar heterojen iken liberal sonuç verme eğiliminde olduğu görülmektedir. $k = 6$ durumunda ise budanmış ortalamaya dayalı testlerden WB testi 1. tip hata kontrolünü her durumda sağlamaya devam ederken WBB testinin bazı durumlarda tutucu sonuç verdiği; F ve W testinin ise varyanslar heterojenken liberal sonuç verme eğilimini koruduğu görülmektedir.

Çizelge 14. Bağımsız $k = 4$ durumunda $g\&h$ $(g = 0{,}5\ h = 0{,}5)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	F	W	WB	WBB
20-20-20-20	1-1-1-1	<u>0,0237</u>	0,0283	0,0376	0,0359
	1-1-1-36	**0,2365**	**0,1382**	0,0413	0,0370
	1-4-9-16	**0,0851**	**0,1525**	0,0437	0,0383
10-15-20-25	1-1-1-1	0,0347	0,0305	0,0365	0,0359
	1-1-1-36	**0,1504**	**0,1147**	0,0379	0,0389
	36-1-1-1	**0,3240**	**0,1318**	0,0418	0,0303
	1-4-9-16	0,0428	**0,1187**	0,0371	0,0400
	16-9-4-1	**0,1447**	**0,1763**	0,0462	0,0328
15-20-25-30	1-1-1-1	0,0304	0,0304	0,0378	0,0340
	1-1-1-36	**0,1685**	**0,1179**	0,0389	0,0387
	36-1-1-1	**0,2982**	**0,1376**	0,0443	0,0357
	1-4-9-16	0,0460	**0,1312**	0,0438	0,0450
	16-9-4-1	**0,1286**	**0,1750**	0,0468	0,0389

Çizelge 15. Bağımsız $k = 6$ durumunda $g\&h$ $(g = 0{,}5\ h = 0{,}5)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	F	W	WB	WBB
20-20-20-20-20-20	1-1-1-1-1-1	0,0290	0,0334	0,0346	0,0244
	1-1-1-1-1-36	**0,2275**	**0,1131**	0,0380	0,0261
	1-1-4-9-9-16	**0,0785**	**0,1814**	0,0419	0,0282
10-15-15-20-20-25	1-1-1-1-1-1	0,0381	0,0343	0,0384	0,0236
	1-1-1-1-1-36	**0,1782**	**0,1021**	0,0411	0,0288
	36-1-1-1-1-1	**0,3157**	**0,1188**	0,0449	0,0249
	1-1-4-9-9-16	0,0406	**0,1524**	0,0401	0,0292
	16-9-9-4-1-1	**0,1450**	**0,2057**	0,0471	0,0240
15-20-20-25-25-30	1-1-1-1-1-1	0,0325	0,0313	0,0368	0,0261
	1-1-1-1-1-36	**0,1896**	**0,1109**	0,0406	0,0291
	36-1-1-1-1-1	**0,2865**	**0,1199**	0,0405	0,0281
	1-1-4-9-9-16	0,0485	**0,1638**	0,0408	0,0320
	16-9-9-4-1-1	**0,1319**	**0,2113**	0,0509	0,0310

$g\&h$ $(g = 1\ h = 0)$ dağılım için gerçekleşen 1. tip hata oranlarına Çizelge 16 ve Çizelge 17'de yer verilmiştir. Sonuçlara göre F ve W testi çoğunlukla liberal sonuçlar verirken budanmış ortalamaya dayalı WBB testi pozitif eşleşmenin olduğu 1 durum hariç 1. tip hata oranı kontrolünü her durumda sağlamıştır. Budanmış ortalamaya dayalı WB testi ise $k = 4$ durumunda negatif eşleşmenin olduğu 2 durum, $k = 6$ durumunda ise 4'ü negatif eşleşmenin olduğu durum olmak üzere 5 durum dışında 1. tip hata oranı kontrolünü sağlamıştır.

Çizelge 16. Bağımsız $k = 4$ durumunda $g\&h$ $(g = 1\ h = 0)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	F	W	WB	WBB
20-20-20-20	1-1-1-1	0,0362	0,0594	0,0459	0,0312
	1-1-1-36	**0,1670**	**0,1053**	0,0596	0,0424
	1-4-9-16	**0,0797**	**0,1395**	0,0691	0,0459
10-15-20-25	1-1-1-1	0,0407	0,0666	0,0527	0,0359
	1-1-1-36	**0,1018**	**0,0984**	0,0532	0,0420
	36-1-1-1	**0,3060**	**0,1372**	0,0744	0,0476
	1-4-9-16	0,0422	**0,1069**	0,0582	0,0438
	16-9-4-1	**0,1549**	**0,1876**	**0,1005**	0,0569
15-20-25-30	1-1-1-1	0,0382	0,0614	0,0478	0,0342
	1-1-1-36	**0,1050**	**0,0896**	0,0576	0,0409
	36-1-1-1	**0,2590**	**0,1203**	0,0681	0,0441
	1-4-9-16	0,0449	**0,1100**	0,0542	0,0421
	16-9-4-1	**0,1305**	**0,1601**	**0,0821**	0,0556

Çizelge 17. Bağımsız $k = 6$ durumunda $g\&h\ (g = 1\ h = 0)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	F	W	WB	WBB
20-20-20-20-20-20	1-1-1-1-1-1	0,0381	**0,0754**	0,0576	0,0261
	1-1-1-1-1-36	**0,1859**	**0,1087**	0,0608	0,0274
	1-1-4-9-9-16	**0,0777**	**0,1682**	**0,0778**	0,0365
10-15-15-20-20-25	1-1-1-1-1-1	0,0408	**0,0869**	0,0650	0,0272
	1-1-1-1-1-36	**0,1266**	**0,1105**	0,0637	0,0237
	36-1-1-1-1-1	**0,3213**	**0,1505**	**0,0844**	0,0364
	1-1-4-9-9-16	0,0438	**0,1396**	0,0736	0,0346
	16-9-9-4-1-1	**0,1544**	**0,2092**	**0,1057**	0,0420
15-20-20-25-25-30	1-1-1-1-1-1	0,0408	**0,0851**	0,0613	0,0301
	1-1-1-1-1-36	**0,1283**	**0,1019**	0,0590	0,0292
	36-1-1-1-1-1	**0,2773**	**0,1263**	**0,0754**	0,0384
	1-1-4-9-9-16	0,0466	**0,1418**	0,0724	0,0402
	16-9-9-4-1-1	**0,1302**	**0,1888**	**0,0864**	0,0415

$k = 4$ ve $k = 6$ için simülasyon çalışmasıyla ilgili son olarak Çizelge 18 ve Çizelge 19'da, $g\&h\ (g = 1\ h = 0{,}5)$ dağılım için gerçekleşen 1. tip hata oranlarına yer verilmiştir. Sonuçlara bakıldığında F ve W testinin varyansların heterojen olduğu durumların tamamında liberal sonuç verdiği, budanmış ortalamaya dayalı WB testinin tüm durumlarda 1. tip hata oranı kontrolünü sağladığı, budanmış ortalamaya dayalı WBB testinin ise $k = 4$ iken 1 durumda, $k = 6$ durumunda ise 13 durumun 7'sinde tutucu sonuç verdiği görülmektedir.

Çizelge 18. Bağımsız $k = 4$ durumunda $g\&h\ (g = 1\ h = 0{,}5)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	F	W	WB	WBB
20-20-20-20	1-1-1-1	0,0211	0,0346	0,0298	0,0240
	1-1-1-36	**0,4365**	**0,3667**	0,0433	0,0363
	1-4-9-16	**0,2094**	**0,5306**	0,0552	0,0427
10-15-20-25	1-1-1-1	0,0296	0,0345	0,0337	0,0266
	1-1-1-36	**0,3778**	**0,3440**	0,0389	0,0348
	36-1-1-1	**0,4995**	**0,4028**	0,0547	0,0372
	1-4-9-16	**0,1532**	**0,4692**	0,0358	0,0315
	16-9-4-1	**0,2830**	**0,5842**	0,0672	0,0441
15-20-25-30	1-1-1-1	0,0319	0,0375	0,0356	0,0267
	1-1-1-36	**0,3990**	**0,3546**	0,0417	0,0354
	36-1-1-1	**0,4764**	**0,3922**	0,0517	0,0421
	1-4-9-16	**0,1686**	**0,4832**	0,0452	0,0369
	16-9-4-1	**0,2658**	**0,5641**	0,0665	0,0508

Çizelge 19. Bağımsız $k = 6$ durumunda $g\&h$ $(g = 1\ h = 0{,}5)$ dağılım için gerçekleşen 1.tip hata oranları

n_j	σ_j^2	F	W	WB	WBB
20-20-20-20-20-20	1-1-1-1-1-1	0,0255	0,0439	0,0325	0,0146
	1-1-1-1-1-36	**0,4275**	**0,3546**	0,0458	0,0255
	1-1-4-9-9-16	**0,1889**	**0,6555**	0,0524	0,0304
10-15-15-20-20-25	1-1-1-1-1-1	0,0302	0,0501	0,0370	0,0166
	1-1-1-1-1-36	**0,3821**	**0,3267**	0,0380	0,0188
	36-1-1-1-1-1	**0,4673**	**0,3832**	0,0544	0,0246
	1-1-4-9-9-16	**0,1473**	**0,6153**	0,0420	0,0245
	16-9-9-4-1-1	**0,2576**	**0,7042**	0,0701	0,0328
15-20-20-25-25-30	1-1-1-1-1-1	0,0297	0,0527	0,0346	0,0177
	1-1-1-1-1-36	**0,3886**	**0,3307**	0,0384	0,0208
	36-1-1-1-1-1	**0,4532**	**0,3765**	0,0468	0,0279
	1-1-4-9-9-16	**0,1529**	**0,6276**	0,0488	0,0287
	16-9-9-4-1-1	**0,2333**	**0,6817**	0,0650	0,0348

5.2. Bağımlı Grup Durumu İçin Simülasyon Çalışması

Bağımlı iki grubun karşılaştırılmasında kullanılan yöntemlerden

- Bağımlı örneklem t testi (t),
- Yuen testi (Y);

ve bağımlı k grubun karşılaştırılmasında kullanılan yöntemlerden

- $\tilde{\varepsilon}$-düzeltmeli ANOVA- F testi (F) ,
- Budanmış ortalama ile $\tilde{\varepsilon}$-düzeltmeli ANOVA- F testi ($FBud.$),
- Bootstrap- t ile ANOVA- F testi (BF),
- Budanmış ortalama ve bootstrap- t ile ANOVA- F testi ($BFBud.$)

farklı deney düzenleri altında gerçekleşen 1. tip hata oranları bakımından karşılaştırılmıştır. R(4.0.2) programlama dili kullanılarak yapılan simülasyon çalışmasında nominal anlamlılık düzeyi düzeyi $\boldsymbol{\alpha = 0,05}$ ve budama oranı $\boldsymbol{\gamma = \%20}$ olarak seçilmiştir. Bootstrap-t yönteminin kullanılmadığı deney düzenlerinde 10.000 tekrar yapılırken bootstrap-t yönteminin kullanıldığı deney düzenlerinde Wilcox ve arkadaşlarının (2000) çalışmasına dayanılarak 1.000 tekrar yapılmıştır. Bootstrap-t yöntemine dayalı BF ve $BFBud.$ testi için M değeri 599 alınmıştır.

5.2.1. Bağımlı Grup Durumu İçin Simülasyon Düzeni

Simülasyon çalışmasında kullanılan deney düzenleri; $\boldsymbol{k = 2}$ ve $\boldsymbol{k = 4}$ için örneklem büyüklüğünün 10 ve 30 olduğu durumlarda, normallik varsayımının sağlanıp sağlanmadığı durumlar ve farklı varyans-kovaryans matrisleri dikkate alınarak oluşturulmuştur.

Normallik varsayımının sağlanmadığı deney düzenleri oluşturmak için bağımsız grup simülasyon çalışmasında olduğu gibi $\boldsymbol{g}\&\boldsymbol{h}$ dağılımı kullanılmıştır. Kullanılan g ve h parametre değerleri ile bu parametre değerlerine karşılık gelen μ_{gh}, μ_{bgh} ile dağılımın şekline ilişkin tanımlama Çizelge 20'de verilmiştir.

Çizelge 20. Bağımlı k grup simülasyon çalışmasında kullanılan g ve h parametre değerleri

g	h	μ_{gh}	μ_{bgh} (%20)	Dağılım tanımlaması
0,0	0,0	0	0	Standart normal
0,0	0,5	0	0	Aşırı kuyruklu
0,5	0,0	0,2263	0,0541	Hafif kuyruklu çarpık
0,5	0,5	0,8033	0,0600	Aşırı kuyruklu çarpık

$k = 2$ iken varyansların homojen olduğu durum için $[\sigma_1^2 \quad \sigma_2^2] = [1 \quad 1]$; varyansların heterojen olduğu durum için $[\sigma_1^2 \quad \sigma_2^2] = [1 \quad 16]$ kullanılırken, $k = 4$ iken varyansların homojen olduğu durum için $[\sigma_1^2 \quad \sigma_2^2 \quad \sigma_3^2 \quad \sigma_4^2] = [1 \quad 1 \quad 1 \quad 1]$; heterojen olduğu durum için $[\sigma_1^2 \quad \sigma_2^2 \quad \sigma_3^2 \quad \sigma_4^2] = [1 \quad 4 \quad 9 \quad 16]$ kullanılmıştır.

$k = 2$ durumunda normallik varsayımının sağlandığı deney düzenlerine ilişkin veri seti üretilirken öncelikle, ortalaması $\mu = [0 \quad 0]$ kovaryans matrisi ise Σ olan iki değişkenli normal dağılımdan gözlem değerleri Z_{ij} $(i = 1, \dots, n; j = 1,2)$ ile gösterilmek üzere $[Z_1 \quad Z_2]$ raslantı değişkeni vektörü üretilmiştir. Kovaryans matrisi (Değişkenlerin varyansı 1 olduğundan korelasyon matrisi de denilebilir.) olarak Σ_1 ve Σ_2 ile gösterilen 2 farklı matris kullanılmıştır.

$$\Sigma_1 = \begin{bmatrix} 1 & 0,1 \\ & 1 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} 1 & 0,8 \\ & 1 \end{bmatrix}$$

Varyansların homojen olduğu durum için $[\sigma_1 \quad \sigma_2] = [1 \quad 1]$ olarak belirlendiğinden Z_{ij} aynen alınırken varyansların heterojen olduğu durum için Z_{ij} değerleri j. gruba karşılık gelen σ_j yani $[\sigma_1 \quad \sigma_2] = [1 \quad 4]$ değerleriyle çarpılmış, böylece Eşitlik (4.1) ve Eşitlik (4.2)'de yer verilen hipotezlerin test edilmesinde kullanılacak veri seti elde edilmiştir.

Normallik varsayımının sağlanmadığı deney düzenlerine ilişkin veri seti üretilirken de öncelikle, ortalaması $\mu = [0 \quad 0]$ kovaryans matrisi ise Σ olan iki değişkenli normal dağılımdan gözlem değerleri Z_{ij} $(i = 1, \dots, n; j = 1{,}2)$ ile gösterilmek üzere $[Z_1 \quad Z_2]$ raslantı değişkeni vektörü üretilmiştir. Kovaryans matrisi olarak yukarıda yer verilen Σ_1 ve Σ_2 ile gösterilen 2 farklı matris kullanılmıştır. Ardından Z_{ij} ile g ve h parametre değerleri Eşitlik (5.1)'de; $g = 0$ olması durumunda ise Eşitlik (5.2)'de yerine konarak ε_{ij} elde edilmiştir. Sonrasında g ve h parametre değerlerine karşılık gelen μ_{gh} değeri ile ε_{ij} ve σ_j'nin Eşitlik (5.4)'te yerine konmasıyla Eşitlik (4.1)'de yer verilen hipotezin test edilmesinde kullanılacak y_{ij} elde edilmiştir. Eşitlik (4.2)'de yer alan budanmış ortalamaların karşılaştırıldığı hipotezin testinde ise y_{ij}'nin elde edilmesinde Eşitlik (5.5) kullanılmıştır.

$k = 4$ durumunda da normallik varsayımının sağlandığı ve sağlanmadığı deney düzenlerinde Eşitlik (4.10) ve Eşitlik (4.11)'de yer verilen hipotezlerin test edilmesine ilişkin veri seti üretilirken $k = 2$ durumu için belirtilen adımlar izlenmiştir. Kovaryans matrisi olarak Wilcox ve arkadaşları (2000)'nın çalışmasında da kullanılan ve aşağıda Σ_3, Σ_4, Σ_5 ve Σ_6 ile gösterilen 4 farklı matris kullanılmıştır.

$$\Sigma_3 = \begin{bmatrix} 1 & 0{,}1 & 0{,}1 & 0{,}1 \\ & 1 & 0{,}1 & 0{,}1 \\ & & 1 & 0{,}1 \\ & & & 1 \end{bmatrix}, \Sigma_4 = \begin{bmatrix} 1 & 0{,}5 & 0{,}5 & 0{,}5 \\ & 1 & 0{,}5 & 0{,}5 \\ & & 1 & 0{,}5 \\ & & & 1 \end{bmatrix}, \Sigma_5 = \begin{bmatrix} 1 & 0{,}8 & 0{,}8 & 0{,}8 \\ & 1 & 0{,}8 & 0{,}8 \\ & & 1 & 0{,}8 \\ & & & 1 \end{bmatrix}, \Sigma_6 = \begin{bmatrix} 1 & 0{,}8 & 0{,}5 & 0{,}2 \\ & 1 & 0{,}5 & 0{,}2 \\ & & 1 & 0{,}2 \\ & & & 1 \end{bmatrix}$$

5.2.2. Bağımlı Grup Durumu İçin Simülasyon Sonuçları

Bağımsız grup için yapılan simülasyon çalışmasında olduğu gibi bağımlı grup simülasyon çalışmasında da, karşılaştırılan testlere ilişkin elde edilen gerçekleşen 1. tip hata oranları Bradley (1978)'in önerdiği dayanıklılık ölçütü kullanılarak yorumlanmıştır.

Normal dağılım, aşırı kuyruklu ($g = 0\ h = 0{,}5$) dağılım ve hafif kuyruklu çarpık ($g = 0{,}5\ h = 0$) dağılım için elde edilen sonuçların gösterildiği Çizelge 21, Çizelge 22 ve Çizelge 23'e bakıldığında iki testin de her durumda 1. tip hata oranı kontrolünü sağladığı görülmektedir.

Çizelge 21. Bağımlı $k = 2$ durumunda normal dağılım için gerçekleşen 1.tip hata oranları

n	Kovaryans Matrisi	σ^2	t	Y
10	Σ_1	1-1	0,0471	0,0392
		1-16	0,0503	0,0541
	Σ_2	1-1	0,0502	0,0349
		1-16	0,0475	0,0593
30	Σ_1	1-1	0,0473	0,0486
		1-16	0,0480	0,0532
	Σ_2	1-1	0,0542	0,0483
		1-16	0,0489	0,0520

Çizelge 22. Bağımlı $k = 2$ durumunda $g\&h$ ($g = 0\ h = 0{,}5$) dağılım için gerçekleşen 1.tip hata oranları

n	Kovaryans Matrisi	σ^2	t	Y
10	Σ_1	1-1	0,0311	0,0309
		1-16	0,0315	0,0338
	Σ_2	1-1	0,0302	0,0296
		1-16	0,0284	0,0410
30	Σ_1	1-1	0,0362	0,0437
		1-16	0,0325	0,0468
	Σ_2	1-1	0,0317	0,0370
		1-16	0,0324	0,0442

Çizelge 21 ve Çizelge 22 karşılaştırıldığında t testinin normal dağılım altında nominal anlamlılık seviyesine daha yakın sonuçlar verdiği görülmektedir.

Çizelge 23. Bağımlı $k = 2$ durumunda $g\&h\ (g = 0{,}5\ h = 0)$ dağılım için gerçekleşen 1.tip hata oranları

n	Kovaryans Matrisi	σ^2	t	Y
10	Σ_1	1-1	0,0434	0,0371
		1-16	0,0701	0,0557
	Σ_2	1-1	0,0397	0,0310
		1-16	0,0686	0,0663
30	Σ_1	1-1	0,0452	0,0427
		1-16	0,0535	0,0528
	Σ_2	1-1	0,0455	0,0414
		1-16	0,0540	0,0575

Çizelge 23'te yer alan hafif kuyruklu çarpık dağılım için elde edilen 1. tip hata oranlarına göre t ve Y testi birbirine yakın sonuçlar vermiştir. Ayrıca her iki test için de $n = 30$ durumunda nominal anlamlılık seviyesine daha yakın sonuçlar elde edilmiştir.

Aşırı kuyruklu çarpık olarak tanımlanan $g\&h\ (g = 0{,}5\ h = 0{,}5)$ dağılımı için elde edilen sonuçlar Çizelge 24'te verilmiştir. Sonuçlara bakıldığında dağılımın hem çarpık hem kuyruklu olmasıyla birlikte t testinin liberal sonuç verme eğiliminde olduğu, Y testinin ise 1 durum hariç 1. tip hata oranı kontrolünü sağladığı görülmektedir.

Çizelge 24. Bağımlı $k = 2$ durumunda $g\&h\ (g = 0{,}5\ h = 0{,}5)$ dağılım için gerçekleşen 1.tip hata oranları

n	Kovaryans Matrisi	σ^2	t	Y
10	Σ_1	1-1	0,0270	0,0277
		1-16	**0,1376**	0,0331
	Σ_2	1-1	0,0248	0,0241
		1-16	**0,2090**	0,0476
30	Σ_1	1-1	0,0276	0,0420
		1-16	**0,1562**	0,0484
	Σ_2	1-1	0,0278	0,0381
		1-16	**0,2174**	0,0530

$k = 4$ durumunda normal dağılım için elde edilen sonuçların yer aldığı Çizelge 25'e bakıldığında budanmış ortalamaya dayalı $FBud.$ testinin $n = 10$ durumunda kovaryans matrisi Σ_5 ve varyanslar homojenken verdiği tutucu sonuç hariç tüm testlerin her durumda 1. tip hata oranı kontrolünü sağladığı görülmektedir.

Çizelge 25. Bağımlı $k = 4$ durumunda normal dağılım için gerçekleşen 1.tip hata oranları

n	Kovaryans Matrisi	σ^2	F	$FBud.$	BF	$BFBud.$
10	Σ_3	1-1-1-1	0,0447	0,0397	0,035	0,038
		1-4-9-16	0,0543	0,0553	0,047	0,040
	Σ_4	1-1-1-1	0,0473	0,0317	0,041	0,039
		1-4-9-16	0,0582	0,0525	0,041	0,055
	Σ_5	1-1-1-1	0,0460	0,0208	0,037	0,030
		1-4-9-16	0,0521	0,0462	0,043	0,043
	Σ_6	1-1-1-1	0,0536	0,0418	0,032	0,035
		1-4-9-16	0,0572	0,0538	0,037	0,035
30	Σ_3	1-1-1-1	0,0506	0,0462	0,053	0,058
		1-4-9-16	0,0514	0,0507	0,063	0,070
	Σ_4	1-1-1-1	0,0477	0,0434	0,051	0,050
		1-4-9-16	0,0446	0,0467	0,055	0,060
	Σ_5	1-1-1-1	0,0459	0,0374	0,041	0,045
		1-4-9-16	0,0489	0,0508	0,064	0,049
	Σ_6	1-1-1-1	0,0489	0,0450	0,041	0,054
		1-4-9-16	0,0511	0,0464	0,052	0,061

Aşırı kuyruklu simetrik olarak tanımlanan $g\&h$ $(g = 0\ h = 0{,}5)$ dağılım için gerçekleşen 1. tip hata oranlarına Çizelge 26'da yer verilmiştir. F testi durumların çoğunda tutucu sonuç verirken bootstrap-t yöntemine dayalı F testi durumların tamamında tutucu sonuç vermiştir. Budanmış ortalamaya dayalı $FBud.$ testi ise 1 durum hariç 1. tip hata kontrolünü sağlarken, bu testin bootstrap-t yöntemiyle birlikte kullanıldığı $BFBud.$ testi 1. tip hata kontrolünü her durumda sağlamıştır.

Çizelge 26. Bağımlı $k = 4$ durumunda $g\&h$ $(g = 0\ h = 0{,}5)$ dağılım için gerçekleşen 1.tip hata oranları

n	Kovaryans Matrisi	σ^2	F	$FBud.$	BF	$BFBud.$
10	Σ_3	1-1-1-1	0,0177	0,0292	0,005	0,032
		1-4-9-16	0,0233	0,0355	0,014	0,036
	Σ_4	1-1-1-1	0,0195	0,0265	0,006	0,031
		1-4-9-16	0,0228	0,0314	0,011	0,039
	Σ_5	1-1-1-1	0,0207	0,0176	0,008	0,025
		1-4-9-16	0,0223	0,0292	0,012	0,031
	Σ_6	1-1-1-1	0,0285	0,0301	0,021	0,046
		1-4-9-16	0,0256	0,0362	0,021	0,051
30	Σ_3	1-1-1-1	0,0230	0,0370	0,007	0,045
		1-4-9-16	0,0262	0,0424	0,018	0,051
	Σ_4	1-1-1-1	0,0209	0,0340	0,012	0,045
		1-4-9-16	0,0243	0,0409	0,017	0,061
	Σ_5	1-1-1-1	0,0182	0,0299	0,014	0,051
		1-4-9-16	0,0204	0,0370	0,016	0,047
	Σ_6	1-1-1-1	0,0263	0,0418	0,013	0,037
		1-4-9-16	0,0304	0,0431	0,016	0,033

Hafif kuyruklu çarpık olarak tanımlanan $g\&h\ (g = 0{,}5\ h = 0)$ dağılım için gerçekleşen 1. tip hata oranlarının yer aldığı Çizelge 27 incelendiğinde aşırı kuyruklu simetrik olarak tanımlanan $g\&h\ (g = 0\ h = 0{,}5)$ dağılım için elde edilen sonuçların aksine F testinin her durumda 1. tip hata kontrolünü sağladığı, bootstrap-t yöntemine dayalı BF testinin ise 3 durum hariç 1. tip hata kontrolünü sağladığı görülmektedir. Budanmış ortalamaya dayalı $FBud.$ testi ise 1 durum hariç 1. tip hata kontrolünü sağlarken, bu testin bootstrap-t yöntemiyle birlikte kullanıldığı $BFBud.$ testi 1. tip hata kontrolünü her durumda sağlamıştır.

Çizelge 27. Bağımlı $k = 4$ durumunda $g\&h\ (g = 0{,}5\ h = 0)$ dağılım için gerçekleşen 1.tip hata oranları

n	Kovaryans Matrisi	σ^2	F	$FBud.$	BF	$BFBud.$
10	Σ_3	1-1-1-1	0,0317	0,0351	<u>0,024</u>	0,046
		1-4-9-16	0,0502	0,0492	0,044	0,047
	Σ_4	1-1-1-1	0,0366	0,0309	<u>0,018</u>	0,028
		1-4-9-16	0,0609	0,0521	0,041	0,052
	Σ_5	1-1-1-1	0,0343	<u>0,0171</u>	<u>0,023</u>	0,028
		1-4-9-16	0,0709	0,0572	0,054	0,057
	Σ_6	1-1-1-1	0,0399	0,0356	0,027	0,043
		1-4-9-16	0,0519	0,0499	0,039	0,050
30	Σ_3	1-1-1-1	0,0410	0,0383	0,041	0,036
		1-4-9-16	0,0454	0,0454	0,035	0,043
	Σ_4	1-1-1-1	0,0401	0,0367	0,040	0,047
		1-4-9-16	0,0473	0,0490	0,046	0,054
	Σ_5	1-1-1-1	0,0427	0,0329	0,031	0,051
		1-4-9-16	0,0556	0,0542	0,051	0,049
	Σ_6	1-1-1-1	0,0468	0,0447	0,048	0,058
		1-4-9-16	0,0547	0,0513	0,051	0,068

Bağımlı k grup simülasyon çalışmasında son olarak; çarpık aşırı kuyruklu olarak tanımlanan $g\&h\ (g = 0{,}5\ h = 0{,}5)$ dağılım durumu incelenmiş ve bu dağılım altında gerçekleşen 1. tip hata oranlarına Çizelge 28'de yer verilmiştir. Sonuçlara bakıldığında F testinin ve bootstrap-t yöntemine dayalı BF testinin çoğunlukla tutucu sonuç verdiği, liberal sonuç verdikleri durumların da olduğu, dolayısıyla bu 2 testin çoğunlukla 1. tip hata kontrolünü sağlayamadığı görülmektedir. Budanmış ortalamaya dayalı $FBud.$ testi ise tutucu sonuç verdiği 2 durum hariç 1. tip hata kontrolünü sağlarken, testin bootstrap-t yöntemiyle birlikte kullanıldığı $BFBud.$ testi tutucu sonuç verdiği 1 durum hariç her durumda 1. tip hata kontrolünü sağlamıştır.

Çizelge 28. Bağımlı $k = 4$ durumunda $g\&h$ $(g = 0{,}5\ h = 0{,}5)$ dağılım için gerçekleşen 1.tip hata oranları

n	Kovaryans Matrisi	σ^2	F	$FBud.$	BF	$BFBud.$
10	Σ_3	1-1-1-1	0,0146	0,0254	0,010	0,032
		1-4-9-16	0,0582	0,0364	0,039	0,046
	Σ_4	1-1-1-1	0,0166	0,0230	0,00**6**	0,020
		1-4-9-16	**0,0877**	0,0346	0,056	0,037
	Σ_5	1-1-1-1	0,0152	0,0147	0,011	0,026
		1-4-9-16	**0,1692**	0,0348	**0,132**	0,050
	Σ_6	1-1-1-1	0,0227	0,0272	0,005	0,034
		1-4-9-16	0,0708	0,0354	0,038	0,045
30	Σ_3	1-1-1-1	0,0173	0,0346	0,009	0,052
		1-4-9-16	0,0619	0,0420	0,045	0,045
	Σ_4	1-1-1-1	0,0143	0,0363	0,006	0,029
		1-4-9-16	**0,0891**	0,0412	0,066	0,033
	Σ_5	1-1-1-1	0,0143	0,0289	0,008	0,036
		1-4-9-16	**0,1500**	0,0434	**0,125**	0,052
	Σ_6	1-1-1-1	0,0187	0,0326	0,010	0,048
		1-4-9-16	0,0737	0,0412	0,051	0,055

6. SONUÇ VE TARTIŞMA

Çalışmada grupların karşılaştırılmasına en yaygın kullanılan parametrik testler ile bu testlerin uygulanabilmesi için sağlanması gereken normallik ve varyansların homojenliği varsayımlarının ayrı ayrı veya aynı anda sağlanmadığı durumda uygulanabilecek testler tanıtılmış, ardından testler, simülasyon çalışması ile elde edilen testlere ait gerçekleşen 1. tip hata oranları bakımından karşılaştırılmıştır. Testlerin 1. tip hata oranları bağımlı ve bağımsız grup ayrımında iki ve ikiden fazla grubun ortalama ve budanmış ortalama bakımından karşılaştırılması için yapılan simülasyon çalışması ile elde edilmiş ve sonuçlar Bradley (1978) tarafından önerilen liberal dayanıklılık ölçütü dikkate alınarak $0{,}025 \leq \hat{\alpha} \leq 0{,}075$ aralığına göre yorumlanmıştır.

Bağımsız iki grubun karşılaştırılması için yapılan simülasyon çalışmasında Student-t testi, Welch testi, Yuen testi ve bootstrap-t ile Yuen testi kullanılmıştır. Normal dağılım altında Student-t testi hariç diğer testler her durumda 1. tip hata oranının kontrolünü sağlarken Student-t testi örneklem büyüklüklerinin farklı ve varyansların homojen olmadığı durumda dayanıklı sonuçlar verememiştir. Varyansların homojenliği varsayımı bozulumuna karşı dayanıklı Welch testi, farklı g ve h parametre değerleriyle elde edilen normal olmayan dağılımlar altında beklendiği üzere liberal sonuç verme eğiliminde olmuştur. Budanmış ortalamaya dayalı Yuen testi ve bootstrap-t ile Yuen testi ise g parametresinin 1 olduğu bir başka deyişle dağılımın aşırı çarpık olduğu durumdaki birkaç durum hariç varsayım bozulmalarına karşı dayanıklı sonuçlar vermiştir.

ANOVA-F testi, Welch testi, budanmış ortalama ile Welch testi ve budanmış ortalama ve bootstrap-t ile Welch testi kullanılarak bağımsız $\boldsymbol{k}$ grubun karşılaştırıldığı simülasyon çalışmasında $\boldsymbol{k = 2}$ için elde edilen sonuçlara çok benzer sonuçlar elde edilmiştir. Normallik varsayımının sağlandığı ancak varyansların heterojen olduğu özellikle negatif eşleşme durumunda ANOVA-F testinin liberal sonuç verme eğiliminde olduğu, diğer testlerin ise normallik varsayımının sağlandığı tüm durumlarda dayanıklı sonuç verdiği görülmüştür. Welch testi normal olmayan dağılımlarda 1.tip hata oranı kontrolünü çoğunlukla sağlayamamıştır. Kuyruklu ve aşırı kuyruklu çarpık dağılımlarda budanmış ortalama ile Welch testi her durumda dayanıklı sonuç verirken budanmış ortalama ve bootstrap-t ile Welch testi $\boldsymbol{k = 6}$ iken 4 durum için verdiği tutucu sonuçlar hariç 1.tip hata oranı kontrolünü sağlamıştır. Aşırı çarpık dağılımlara gelindiğinde ise budanmış ortalama ile Welch testinin liberal, budanmış ortalama ve bootstrap-t ile Welch testinin ise tutucu sonuçlar verdiği durumlar olmakla birlikte bu iki testin 1.tip hata oranı kontrolünün sağlandığı durum sayısı bakımından ANOVA-F testi ve Welch testine nazaran daha iyi olduğu görülmüştür.

Bağımlı grupların karşılaştırılmasına ilişkin simülasyon çalışmasında ilk olarak bağımlı örneklem t testi ve Yuen testi ile karşılaştırılan iki bağımlı grup durumuna yer verilmiştir. Sonuçlara göre dağılımın normal olduğu durum ile yalnızca çarpık ya da yalnızca kuyruklu olduğu durumlarda her iki test de 1.tip hata oranı kontrolünü sağlamıştır. Ancak dağılımın hem çarpık hem de aşırı kuyruklu olduğu durumda bağımlı örneklem t testi varyansların heterojen olduğu tüm durumlarda liberal sonuç vermiştir. Yuen testi ise $\boldsymbol{n = 10}$ durumunda gruplar arası korelasyon yüksekken verdiği bir tutucu sonuç hariç her durumda 1.tip hata oranı kontrolünü sağlamıştır.

Bağımlı $\boldsymbol{k}$ grubun karşılaştırıldığı simülasyon çalışmasında ise $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi, budanmış ortalama ile $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi, bootstrap-t ile ANOVA-F testi ile budanmış ortalama ve bootstrap-t ile ANOVA-F testi incelenmiştir. Normal dağılım altında tüm testler 1.tip hata oranı kontrolü açısından başarılı olmuştur. Ancak aşırı kuyruklu dağılım ($\boldsymbol{h = 0,5}$) altında $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi ile bootstrap-t ile ANOVA-F testi tutucu sonuçlar vermiştir. Dağılım hem çarpık hem de aşırı kuyruklu ($\boldsymbol{g = 0,5\ h = 0,5}$) olduğunda da aynı şekilde $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi ve bootstrap-t ile ANOVA-F testi 1.tip hata oranı kontrolünde başarılı olamamıştır. Budanmış ortalama ile $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi ve budanmış ortalama ve bootstrap-t ile ANOVA-F testi ise hem aşırı kuyruklu dağılım ($\boldsymbol{h = 0,5}$) altında hem de çarpık ve aşırı kuyruklu ($\boldsymbol{g = 0,5\ h = 0,5}$) dağılım altında birkaç durum dışında dayanıklı sonuçlar vermiştir.

Sonuç olarak; bağımsız grupların karşılaştırılmasında testlerden Student-t testi, F testi ve Welch testi varsayımların sağlanmadığı durumlarda 1. tip hata kontrolünde başarılı olamazken, budanmış ortalamaya dayalı Yuen testi, bootstrap-t ile Yuen testi ve budanmış ortalama ile Welch testi varsayımların sağlanmadığı durumlara karşı dayanıklı sonuç vermektedir. Bağımlı grupların karşılaştırılmasında ise bağımlı örneklem t testi çarpık-aşırı kuyruklu dağılım altında, $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi ile bootstrap-t ile ANOVA-F testi ise aşırı kuyruklu dağılım ile çarpık-aşırı kuyruklu dağılımda 1. tip hata kontrolünü sağlayamazken budanmış ortalama ile $\tilde{\epsilon}$-düzeltmeli ANOVA-F testi ile budanmış ortalama ve bootstrap-t ile ANOVA-F testi çarpık-aşırı kuyruklu dağılım da dahil incelenen tüm dağılımlarda birkaç durum hariç 1. tip hata kontrolünü sağlamaktadır.

7. KAYNAKLAR

Alexander, R. A., Govern, D. M., A new and simpler approximation for ANOVA under variance heterogeneity, Journal of Educational Statistics, 19(2), 91-101, 1994.

Algina, J., Oshima, T. C., Lin, W. Y., Type I error rates for Welch's test and James's second-order test under nonnormality and inequality of variance when there are two groups, Journal of Educational Statistics, 19(3), 275-291, 1994.

Alpar, R., Spor Bilimlerinde Uygulamalı İstatistik, Nobel Akademik Yayıncılık, 2006.

Berkovits, I., Hancock, G. R., Nevitt, J., Bootstrap resampling approaches for repeated measure designs: relative robustness to sphericity and normality violations, Educational and Psychological Measurement, 60(6), 877-892, 2000.

Box, G. E., Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification, The annals of mathematical statistics, 25(2), 290-302, 1954.

Bradley, J. V., Robustness? British Journal of Mathematical and Statistical Psychology, 31, 144–152, 1978.

Brown, M. B., Forsythe, A. B., The small sample behavior of some statistics which test the equality of several means, Technometrics, 16(1), 129-132, 1974.

Cribbie, R. A., Fiksenbaum, L., Keselman, H. J., Wilcox, R. R., Effect of non-normality on test statistics for one-way independent groups designs, British Journal of Mathematical and Statistical Psychology, 65(1), 56-73, 2012.

Di, N. F. M., Yahaya, S. S. S., Abdullah, S., Comparing Groups Using Robust H Statistic With Adaptive Trimmed Mean, Sains Malaysiana, Vol: 43, No:4, Page:643-648, 2014.

Efron, B., Bootstrap methods: another look at the jackknife, The Annals of Statistics, 7(1), 1-26, 1979.

Guo, J. H., Luh, W. M., An invertible transformation two-sample trimmed t-statistic under heterogeneity and nonnormality. Statistics & Probability Letters, 49(1), 1-7, 2000.

He, X., Simpson, D. G., Portnoy, S. L., Breakdown robustness of tests, Journal of the American Statistical Association, 85, 446–452, 1990.

Hoaglin, D. C., Mosteller, F., Tukey, J. W., Exploring data tables, trends, and shapes, New York: Wiley, 1985.

Keselman, H. J., Algina, J., Kowalchuk, R. K., The analysis of repeated measures designs: a review, British Journal of Mathematical and Statistical Psychology, 54(1), 1-20, 2001.

Keselman, H. J., Wilcox, R. R., Lix, L. M., A generally robust approach to hypothesis testing in independent and correlated groups designs, Psychophysiology, 40(4), 586-596, 2003.

Keskin, S., Mendes, M., Faktörlerden birinin seviyelerinde tekrarlanan ölçüm bulunan iki faktörlü deneme düzenleri. S.Ü. Ziraat Fakültesi Dergisi, 15(25): 42-53, 2001.

Lix, L. M., Keselman, H. J., To trim or not to trim: Tests of location equality under heteroscedasticityand nonnormality, Educational and Psychological Measurement, 58(3), 409-429, 1998.

Najdi, N. F. N., Ahad, N. A., A Modification of ANOVA with Trimmed Mean, Malaysian Journal of Social Sciences and Humanities (MJSSH), 4(4), 109-118, 2019.

Oberfeld, D., Franke, T., Evaluating the robustness of repeated measures analyses: The case of small sample sizes and nonnormal data, Behavior research methods, 45(3), 792-812, 2013.

Schrader, R. M., Hettmansperger, T. P., Robust analysis of variance, Biometrika, 67, 93–101, 1980.

Özdemir, A. F., Anova Methods For The Group Means With Unknown Variances, Doktora Tezi, Dokuz Eylül Üniversitesi Fen Bilimleri Enstitüsü, İzmir, 2006.

Özdemir, A. F., Navruz, G., Bootstrap-t ve Yüzdelik Bootstrap Yöntemlerinde Tekrar Sayısı, Budama Yüzdesi ve Dağılımın Sonuçlara Etkisi, Nevşehir Bilim ve Teknoloji Dergisi, 5(2), 74-85, 2016.

Özdemir, A. F., Wilcox, R. R., Yildiztepe, E., Comparing J independent groups with a method based on trimmed means, Communications in Statistics-Simulation and Computation, 47(3), 852-863, 2018.

Tukey, J. W., McLaughlin, D. H., Less vulnerable confidence and significance procedures for location based on a single sample: Trimming/Winsorization 1. Sankhyā: The Indian Journal of Statistics, Series A, 331-352, 1963.

Welch, B. L., The significance of the difference between two means when the population variances are unequal, Biometrika, 29, 350–362, 1938.

Welch, B. L., On the comparison of several mean values: An alternative approach. Biometrika, 38, 330-336, 1951.

Wilcox, R. R., A one-way random effects model for trimmed means, Psychometrika 59, 289–306, 1994.

Wilcox, R. R., Applying Contemporary Statistical Techniques, Academic Press, 2003.

Wilcox, R. R., Keselman, H. J., Kowalchuk, R. K., Can tests for treatment group equality be improved?: The bootstrap and trimmed means conjecture, British Journal of Mathematical and Statistical Psychology, 51(1), 123-134. 1998.

Wilcox, R. R., Keselman, H. J., Muska, J., Cribbie, R., Repeated measures ANOVA: Some new results on comparing trimmed means and means, British Journal of Mathematical and Statistical Psychology, 53(1), 69-82, 2000.

Wilcox, R. R., Keselman, H. J., Using trimmed means to compare K measures corresponding to two independent groups, Multivariate Behavioral Research, 36(3), 421-444, 2001.

Wilcox, R. R., Fundamentals of modern statistical methods: Substantially improving power and accuracy, Springer, 2010.

Wilcox, R. R., Introduction to robust estimation and hypothesis testing, Academic press, 2012.

Wu, P. C., Modern One-Way ANOVA F Methods: Trimmed Means, One Step M-estimators and Bootstrap Methods, J. Quantitative Res., 155- 172, 2007.

Yıldıztepe, E., Özdemir, A. F., Asimetrik ve Ağır Kuyruklu Dağılımların Konum Parametresinin Bootstrap Güven Aralıkları İçin Bir Benzetim Çalışması, Anadolu University of Sciences & Technology-A: Applied Sciences & Engineering, 14(3), 2013.

Yılmaz, İ., Özdemir, A. F., Robust Methods for Comparing K Independent Groups, Türkiye Klinikleri Journal of Biostatistics, 8(2), 143-151, 2016.

Yuen, K. K., The two-sample trimmed t for unequal population variances, Biometrika, 61(1), 165-170, 1974.

EKLER

EK 1 – Simülasyonda Kullanılan R Paketleri ve Fonksiyonları

1. MASS: Support Functions and Datasets for Venables and Ripley's MASS
    - mvrnorm
2. metaSEM: Meta-Analysis using Structural Equation Modeling
    - vec2symMat
3. onewaytests: One-Way Tests in Independent Groups Designs
    - welch.test
4. reshape2: Flexibly Reshape Data: A Reboot of the Reshape Package
    - melt
5. stats: The R Stats Package
    - rnorm
    - t.test
    - oneway.test
6. WRS2: A Collection of Robust Statistical Methods
    - t1waybt
    - yuen
    - yuend
    - yuenbt
    - rmanova
    - rmanovab