ANADOLU ÜNİVERSİTESİ BİLİM VE TEKNOLOJİ DERGİSİ –B
Teorik Bilimler

ANADOLU UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY –B
Theoretical Sciences
Cilt/Vol.:1-Sayı/No: 1 : 57-71 (2011)

HOMOJEN OLMAYAN VARYANS VARSAYIMI ALTINDA ORTALAMALARIN EŞİTLİĞİ İÇİN BAZI TEST İSTATİSTİKLERİ VE KARŞILAŞTIRMALARI

Esra YİĞİT [1], Hamza GAMGAM [1]

ÖZ

İkiden fazla yığının ortalamalarının eşitliği hipotezinin testi amacıyla kullanılan klasik *F* testi, normallik ve yığın varyanslarının homojenlik varsayımına dayanır. Bu varsayımlar özellikle varyansların homojenlik varsayımı, sağlanmadığında klasik *F* testinin kullanılması uygun olmamaktadır. Bu durum özellikle örneklem hacmi büyük olmadığında, önemli bir sıkıntı doğurmaktadır. Literatürde bu konuyla ilgili bir çok test istatistiği geliştirilmiştir. Bu çalışmada Brown-Forsythe, genelleştirilmiş F, Scott-Smith, Welch, Xu-Wang testleri tanıtılmış ve testlerin farklı yığın parametreleri ve örnek hacimleri altında deneysel I.tip hata oranı ve testin gücü bakımından karşılaştırılması yapılmıştır.

Anahtar Kelimeler : Brown-Forsythe testi, Genelleştirilmiş F testi, Scott-Smith testi, Welch testi, Xu-Wang testi.

THE TESTS PROPOSED FOR THE ONE-WAY ANOVA UNDER UNEQUAL VARIANCES AND SIMULATION STUDY

ABSTRACT

Classical F-test to compare several means depends on the assumptions of homogenety of populations variances and normality. When these assumptions-especially the equality of variance-is dropped, classical F-test fails to reject the null hypothesis even the data actually provides strong evidence to do so. This can be considered as a serious problem in some applications especially when the sample size is not large. For this problem, a large number of tests are available in the literature. In this study Brown-Forsythe, generalized F, Scott-Smith, Welch, Xu-Wang are introduced. Also a simulation study is performed to compare these tests in different combination of variance, means, population number and sample size.

Keywords: The Brown-Forsythe test, The Genelleştirilmiş F test, The Scott-Smith test, The Welch test, The Xu-Wang test.

[1,] Gazi Üniversitesi Fen-Edebiyat Fakültesi İstatistik Bölümü Teknikokullar Ankara, eyigit@gazi.edu.tr. gamgam@gazi.edu.tr (Hamza GAMGAM)

Geliş: 20 Ekim 2009; **Düzeltme:** 24 Mart 2010; **Kabul:** 24 Eylül 2010

1. GİRİŞ

Klasik varyans analizi herbiri normal dağılıma sahip olan ikiden fazla yığının ortalamalarının eşitliği hipotezinin testi amacıyla kullanılır. Klasik tek faktörlü varyans analizi problemlerinde, herbiri normal dağılıma sahip yığın sayısı *k,* bu yığınlardan *i*. yığının bilinmeyen ortalaması μ_i ve genel ortalama μ olmak üzere, yokluk ve alternatif hipotezler

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_k = \mu \quad \text{ve} \quad H_1 : \exists \mu_i \neq \mu_j, \qquad 1 \leq i \leq j \leq k \tag{1.1}$$

biçiminde ifade edilir. *i.* yığından n_i hacimli bağımsız örneklemler $X_{i1}, \ldots, X_{in_i}$ ve $i=1,...k$ olsun. *i.* faktör düzeyinin bağımlı değişken üzerindeki etkisi α_i, *i.* yığından seçilen örneklemdeki *j.* birimin bağımlı değişken bakımından ölçüm değeri x_{ij}, hata terimi ε_{ij} olmak üzere, bilinen klasik tek faktörlü varyans analizi modeli

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \qquad i = 1, \ldots, k, \quad j = 1, \ldots, n_i$$

biçiminde ifade edilir (Scheffe,1959). Klasik varyans analizinde Eş. 1.1'deki H_0 hipotezini test etmek için gerekli olan varsayımlar;

1) n_i hacimli örneklemlerin her biri normal dağılımdan gelmiştir.

2) n_i hacimli örneklemlerin geldikleri yığınların varyansları homojendir.

3) örneklemler rastsal olarak seçilmişlerdir ve bağımsızdırlar.

şeklinde ifade edilebilir. Homojen varyans varsayımı sağlanmadığında yokluk hipotezinin reddini destekler nitelikte önemli kanıtlar olsa bile bazen klasik varyans analizi büyük hacimli örneklemler durumunda bile yokluk hipotezini reddedemeyebilir. Bir çok alanda, büyük hacimli örneklemlerin elde edilemeyeceği düşünülürse bu durum önemli bir sıkıntı doğurabilir. Büyük hacimli örneklemleri elde etmenin zor olduğu alanlardan biri biomedikal çalışmalardır. Böyle uygulamalarda her bir gözlem hayati öneme sahip olabilir veya bu gözlemi elde etmek çok pahalı olup zaman alabilir. Bu gibi durumlarda yeterli örneklem hacmine sahip olunamamaktadır. Böylece küçük hacimli örneklemlerle çalışma zorunluluğu ortaya çıkar. Böyle durumlarda klasik varyans analizi oldukça kötü sonuçlar vermesinden dolayı alternatif bir çok test geliştirilmiştir.

Bu test istatistiklerinin bazılarının dağılımı tam olarak bilinirken bazılarının dağılımı simülasyon yoluyla yaklaşık olarak bulunmaktadır (Weerahandi, 1995; Weerahandi, 2004). Homojen varyans varsayımının sağlanmadığı durumda ilk olarak normal dağılıma sahip iki yığın ortalamasının eşitliğine ait hipotezin testi incelenmiştir. Bu durum, Behrens-Fisher problemi olarak bilinmektedir. Behrens-Fisher problemi için önerilen ilk testlerden biri de Welch (1947)'in geliştirdiği testtir. Welch (1951), bu testi *k* yığının ortalamasının eşitliğine ait hipotezin testi için genelleştirmiştir. Scott ve Smith (1971) tarafından Scott-Smith test istatistiği, Brown-Forsythe (1974) tarafından klasik *F* testinin bir uyarlaması olan Brown-Forsythe testi önerilmiştir. Ayrıca Weerahandi (1995a), örneklemlerin ortak ölçüm dönüşümleri altında değişmeyen Genelleştirilmiş *F* testini geliştirmiştir. Genelleştirilmiş F testine dayalı bir diğer testte Xu ve Wang (2007) tarafından önerilmiştir. Ayrıca Gamage ve Weerahandi (1998) simülasyon yoluyla Genelleştirilmiş *F* testini, klasik *F*, ağırlıklandırılmış *F*, Welch ve Brown-Forsythe testleri ile karşılaştırmışlardır. Gerami ve Zahedian (2001) ise çalışmasında Welch testi, Scott-Smith testi, Weerahandi'nin Genelleştirilmiş F testi ve Chen ve Chen (1998) tarafından önerilen Tek Aşamalı testlerin karşılaştırmasını simülasyon yoluyla yapmışlardır.

Bu çalışmanın ikinci bölümünde normal dağılıma sahip ancak homojen varyans varsayımını sağlamayan *k* yığın ortalamasının eşitliği hipotezinin testi için geliştirilen Welch Testi, Scott ve Smith Testi, Brown ve Forsythe Testi, Weerahandi'nin Genelleştirilmiş F Testi, Xu ve Wang'ın Genelleştirilmiş F testi tanıtılmıştır. Üçüncü bölümde ise simülasyon yoluyla bu test istatistiklerinin, farklı örneklem hacmi ve yığınların farklı varyansları altında deneysel I. tip hata bakımından karşılaştırmaları ve sonra da farklı örneklem hacmi, yığınların farklı varyansları ve ortalamaları altında testin gücü bakımından karşılaştırmaları yapılmıştır.

2. TEST İSTATİSTİKLERİ

Bu bölümde normal dağılıma sahip ancak homojen varyans varsayımını sağlamayan k yığın ortalamasının eşitliği hipotezinin testi için geliştirilen bazı test istatistiklerine yer verilmiştir. *i.* yığının varyansı σ_i^2 , i=1,…,k ve bunların homojen olmadığı varsayılsın, bu durumda *i.* örneklem için ortalaması $\overline{X}_i$ ve varyansı S_i^2 olmak üzere, Standartlaştırılmış Gruplar Arası Kareler Toplamı ve Standartlaştırılmış Hata Kareler Toplamı sırasıyla,

$$\tilde{S}_b = \tilde{S}_b\left(\sigma_1^2,\ldots,\sigma_k^2\right) = \sum_{i=1}^{k}\frac{n_i\overline{X}_i^2}{\sigma_i^2} - \frac{\left(\sum_{i=1}^{k}\frac{n_i\overline{X}_i}{\sigma_i^2}\right)^2}{\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}}, \tag{2.1}$$

$$\tilde{S}_e = \sum_{i=1}^{k}\frac{n_i S_i^2}{\sigma_i^2} \tag{2.2}$$

olarak ifade edilir. Homojen varyans varsayımı sağlanmadığı zaman klasik F test istatistiğinin dağılımı teorik olarak elde edilememektedir (Weerahandi, 1995). Bu yüzden ortalamaların eşitliği hipotezini test etmek için klasik F testine alternatif başka test istatistikleri kullanılmaktadır.

2.1 Welch Testi

Yığın varyansları homojen olmadığında k yığının ortalamasının eşitliği hipotezini test etmek için Welch (1951), Behren-Fisher probleminin çözümü için geliştirdiği testin genelleştirilmiş biçimini önermiştir. Bu test pratik olması bakımından uygulamalarda sıklıkla kullanılmaktadır. Buna göre

$w_i = \dfrac{n_i}{S_i^2}$ olmak üzere, Welch (1951) test istatistiği

$$W = \frac{\sum_{i=1}^{k} w_i\left[\left(\bar{X}_i - \bar{X}\right)^2 \Big/ (k-1)\right]}{1+\frac{2(k-2)}{k^2-1}\sum_{i=1}^{k}\frac{1}{n_i-1}\left(1-\frac{w_i}{\sum w_j}\right)^2} \tag{2.3}$$

şeklinde tanımlanır. Yokluk hipotezinin doğruluğu altında W test istatistiğinin dağılımı, k-1 ve f serbestlik dereceli F dağılımına yakınsamaktadır (Welch, 1951). Bu F dağılımının paydası için serbestlik derecesi

$$f = \frac{1}{\frac{3}{k^2-1}\sum_{i=1}^{k}\frac{1}{n_i-1}\left(1-\frac{w_i}{\sum w_j}\right)^2}$$

eşitliğinden elde edilmektedir. W test istatistiği için hesaplanan değer w olmak üzere,

$$P\left(F_{k-1,f} \rangle w\right) \langle \alpha$$

ise ortalamaların eşitliği hipotezi red edilir.

2.2 Scott-Smith Testi

Yığın varyansları homojen olmadığında k sayıda yığının ortalamalarının eşitliği hipotezinin testi için Scott-Smith (1971) tarafından önerilen test istatistiği aşağıda verilmiştir.

$$F_s = \sum_{i=1}^{k} \frac{n_i \left(\overline{X}_i - \overline{X}\right)^2}{S_i^{*2}} \tag{2.4}$$

Burada S_i^{*2}, $\quad S_i^{*2} = \dfrac{n_i - 1}{n_i - 3} S_i^2$ biçimindedir.

Yokluk hipotezinin doğruluğu altında F_s istatistiği k serbestlik dereceli χ^2 dağılımına yakınsamaktadır (Scott ve Smith,1971). Buna göre F_s istatistiği için hesaplanan değer f_s olmak üzere,

$$P\left(F_s \rangle f_s\right) \langle \alpha$$

ise ortalamaların eşitliği hipotezi red edilir.

2.3 Brown-Forsythe Testi

Brown ve Forsythe (1974) tarafından önerilen Brown-Forsythe testi klasik F testinin uyarlanmış bir biçimidir. Bu test istatistiği

$$B = \frac{\sum_{i=1}^{k} n_i \left(\overline{X}_i - \overline{X}\right)^2}{\sum_{i=1}^{k} \left(1 - \frac{n_i}{n}\right) S_i^2} \tag{2.5}$$

olarak önerilmiştir. Yokluk hipotezinin doğruluğu altında B istatistiği $k-1$ ve v serbestlik dereceli F dağılımına sahiptir (Brown ve Forsythe, 1974). Bu F dağılımının payda için serbestlik derecesi

$$v = \frac{\left[\sum_{i=1}^{k} \left(1 - \frac{n_i}{n}\right) S_i^2\right]^2}{\sum_{i=1}^{k} \frac{\left(1 - \frac{n_i}{n}\right)^2 S_i^4}{n_i - 1}}$$

eşitliği ile verilmiştir. Buna göre B istatistiği için hesaplanan değer b olmak üzere,

$$P\left(F_{k-1,v} \rangle b\right) \langle \alpha$$

olduğunda ortalamaların eşitliği hipotezi red edilir.

2.4 Weerahandi'nin Genelleştirilmiş *F* Testi

Klasik yaklaşımda test istatistiğinin dağılımının sağ kuyruk bölgesi kullanılırken genelleştirilmiş yaklaşımda ise test değişkeni için oluşturulan örnek uzayının sağ kuyruk bölgesi kullanılır. Genelleştirilmiş F testinin *p* değerinin hesaplanması için sağ kuyruk bölgedeki gözlenen örnek noktalarının sayısı dikkate alınır. Yokluk hipotezinin doğruluğu altında bu bölgenin olasılığı genelleştirilmiş *p* değerini vermektedir (Gamage ve Weerahandi, 1998). Weerahandi, bu yöntemde $S_i^2 = \frac{1}{n_i - 1}\sum_{i=1}^{n_i}\left(X_{ij} - \overline{X}_i\right)^2$ istatistiğinin yerine σ_i^2 parametresinin ençok olabilirlik tahmin edicisi olan $S_i^2 = \frac{1}{n_i}\sum_{i=1}^{n_i}\left(X_{ij} - \overline{X}_i\right)^2$ istatistiğinin kullanımını önermiştir (Weerahandi, 1995a). B_j değişkeni, $\frac{n_i S_i^2}{\sigma_i^2}$ istatistiğinin bir fonksiyonu olarak aşağıdaki gibi tanımlanmıştır.

$$B_j = \frac{\left(\sum_{i=1}^{j}\frac{n_i S_i^2}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{j+1}\frac{n_i S_i^2}{\sigma_i^2}\right)}, \qquad j = 1,\ldots,k-1 \tag{2.6}$$

Buna göre B_j istatistiğinin dağılımı

$$B_j \sim beta\left[\sum_{i=1}^{j}\frac{(n_i - 1)}{2}, \frac{(n_{j+1} - 1)}{2}\right]$$

olur (Weerahandi, 1995a). *i*. örnek için $\frac{n_i S_i^2}{\sigma_i^2}$ rasgele değişkeni Eş. 2.2'de verilen $\tilde{S}_e$ ve Eş. 2.6'da verilen B_j değişkenine bağlı olarak aşağıdaki gibi elde edilir.

$$\frac{n_1 S_1^2}{\sigma_1^2} = \tilde{S}_e B_1 B_2 \ldots B_{k-1},$$

$$\frac{n_i S_i^2}{\sigma_i^2} = \tilde{S}_e\left(1 - B_{i-1}\right)B_i \ldots B_{k-1}, \qquad i = 2,\ldots,k-1$$

$$\frac{n_k S_k^2}{\sigma_k^2} = \tilde{S}_e\left(1 - B_{k-1}\right).$$

$\frac{n_i S_i^2}{\sigma_i^2}$ istatistiği için yapılan ayrıştırma, bilinmeyen bir parametreye bağlı değildir. Bu yüzden H_0 hipotezinin kabul edilip edilmemesi bu ifadeyi etkilememektedir.

Bu durumda Eş. 2.1'de verilen $\tilde{S}_b$ fonksiyonun gözlem değerlerinden elde edilen fonksiyonu $\tilde{s}_b$ olmak üzere genelleştirilmiş *p* değeri

$$=1-E\left(H_{k-1,n-k}\left\{\frac{n-k}{k-1}\tilde{s}_b\left[\frac{n_1s_1^2}{B_1B_2\ldots B_{k-1}},\frac{n_2s_2^2}{(1-B_1)B_2\ldots B_{k-1}},\frac{n_3s_3^2}{(1-B_2)B_3\ldots B_{k-1}},\ldots,\frac{n_ks_k^2}{(1-B_{k-1})}\right]\right\}\right)$$

şeklinde tanımlanır (Weerahandi, 1995a). $H_{k-1,n-k}$, *k-1 ve n-k* serbestlik derecesine sahip F dağılımının kümülatif dağılım fonksiyonudur ve Genelleştirilmiş p, bu ifadenin beklenen değeri alınarak bulunur. Buna göre $p<\alpha$ olduğunda yokluk hipotezi red edilir.

2.5 Xu ve Wang'ın Genelleştirilmiş *F* Testi

Homojen olmayan varyans durumunda yığınların ortalamalarının eşitliği hipotezinin testi için önerilen diğer bir test istatistiği de Xu ve Wang (2007) tarafından geliştirilen genelleştirilmiş F testidir. Eş. 1.1'deki yokluk hipotezinin

H_0: $\mu_1=\mu_k$; $\mu_2=\mu_k$;... ;$\mu_{k-1}=\mu_k$

biçiminde gösterilebileceği Xu ve Wang (2007) tarafından ifade edilmiştir. v_a *ile* v_b

$$v_a=\left(\mu_1,\mu_2,\ldots,\mu_{k-1}\right)', \qquad v_b = {}_{k-1}\mu_k$$

ile gösterilirse, yokluk hipotezi ve karşıt hipotez

$$H_0: v_a=v_b\ , \qquad H_1: v_a\neq v_b \tag{2.7}$$

şeklinde yeniden düzenlenebilir. Burada ${}_{k-1}$, $\left(k-1\right)\times 1$ boyutlu tüm elemanları 1 olan vektördür. Xu ve Wang (2007) tarafından geliştirilen genelleştirilmiş p için gerekli tanımlar aşağıdaki gibidir. Genelleştirilmiş F testine dayalı bir test olduğundan σ_i^2 yerine en çok olabilirlik tahmin edicisi olan $S_i^2=\frac{1}{n_i}\sum_{j=1}^{n_i}\left(X_{ij}-\bar{X}_i\right)^2$ kullanılmıştır. $\bar{Y}_a$, $\bar{Y}_b$, S_a *ve* S_b ifadeleri

$$\bar{Y}_a=\left(\bar{X}_1,\ldots,\bar{X}_{k-1}\right)' \qquad \bar{Y}_b=\mathbf{1}_{k-1}\bar{X}_k$$

$$S_a=k\ddot{o}\c{s}\left(\frac{S_1^2}{n_1-1},\ldots,\frac{S_{k-1}^2}{n_{k-1}-1}\right), \qquad S_b=\frac{S_k^2}{n_k-1}S_k^2\mathbf{1}_{k-1}\mathbf{1}'_{k-1}$$

şeklinde tanımlanmıştır. Genelleştirilmiş test değişkeninin değeri olan t

$$t=\left(\bar{y}_a-\bar{y}_b\right)'\left(s_a+s_b\right)^{-1}\left(\bar{y}_a-\bar{y}_b\right)$$

olarak bulunur. Burada $\bar{y}_a$, $\bar{y}_b$, s_a *ve* s_b sırasıyla $\bar{Y}_a$, $\bar{Y}_b$, S_a *ve* S_b'nin gözlem değerleridir. Test değişkeni T ise,

$$T=Y'\left(\left(s_a+s_b\right)^{-1/2}\left(k\ddot{o}\c{s}\left(\frac{s_1^2}{U_1},\ldots,\frac{s_{k-1}^2}{U_{k-1}}\right)+\frac{s_k^2}{U_k}\mathbf{1}_{k-1}\mathbf{1}'_{k-1}\right)\left(s_a+s_b\right)^{-1/2}\right)Y$$

şeklinde tanımlanır (Xu ve Wang, 2007). Burada Y ve U_i değerleri

$$Y \sim N\left(0, I_{k-1}\right) \qquad U_i \sim \chi^2_{n_i-1}, \qquad i = 1, \ldots, k$$

şeklindedir. Böylece H_o hipotezinin doğruluğu altında genelleştirilmiş p değeri

$$p = P\left(T \geq t \mid H_0\right)$$

olarak ifade edilir.

3. SİMÜLASYON ÇALIŞMASI

Bu bölümde, Bölüm 2'de tanıtılan test istatistiklerinin deneysel I.tip hata oranları ve güçleri bakımından karşılaştırılması simülasyon yoluyla yapılarak sonuçlar tablolarla ve grafiklerle gösterilmiştir.

3.1 Test İstatistiklerinin Deneysel I.Tip Hata Oranları Bakımından Karşılaştırılması

Simülasyon çalışmasında, test istatistiklerinin deneysel I.tip hata oranları bakımından karşılaştırılmasında ortalamaları eşit olarak belirlenen *k=3*, *k=5* ve *k*=7 sayıdaki yığınlar için, örnek hacimleri ve yığın varyanslarının çeşitli kombinasyonları kullanılmıştır. Bu kombinasyonlar örnek hacimlerinin eşit ve farklı, yığın varyanslarının homojen ($\sigma_1^2 = \ldots = \sigma_k^2$) ve heterojen ($\sigma_1^2 < \ldots < \sigma_k^2$ veya $\sigma_1^2 > \ldots > \sigma_k^2$), örnek hacimleri ile yığın varyanslarının doğru ve ters orantılı olduğu durumlar dikkate alınarak oluşturulmuştur. Ayrıca yığın varyanslarının heterojenlikleri artan bir şekilde oluşturularak heterojenlik seviyeleri de dikkate alınmıştır.

Önceden belirlenen yığınların sayısına göre yukarıda bahsedilen durumlar dikkate alınarak aynı ortalamalı normal dağılımlı yığınlardan n_i hacminde örnekler üretilmiştir *(i=1,..,k, j=1,...,n_i).* Bu örneklerin herbiri için ortalamaların eşitliği hipotezinin testi amacıyla 2. bölümde bahsedilen klasik varyans analizi (KF), Welch (W), Scott-Smith (SS), Brown-Forsythe (BF), Weerahandi'nin Genelleştirilmiş F (GF), ve Xu ile Wang'ın Genelleştirilmiş F (XW) testlerinin algoritmaları kullanılarak test istatistiklerinin herbirinin değeri hesaplanmış ve bu testlerin herbiri için p değerleri bulunmuştur.

p değerleri seçilen nominal α (anlam seviyesi) değeri ile karşılaştırılarak "p değeri < nominal α" ise ortalamaların eşitliği hipotezi red edilmiştir. Bu işlemler 5000 kez tekrarlanarak nominal α=0.05 durumu için her bir test istatistiğinin ortalamaların eşitliği hipotezini red etme sayıları saptanmış ve bunlar 5000 tekrar sayısına bölünerek her bir test istatistiği için deneysel I.tip hata oranları hesaplanmıştır. Bu testlerden genelleştirilmiş p değeri gerektiren GF ve XW testlerinin p değerini bulmak amacıyla herbiri ayrıca 5000 tekrara dayalı simülasyon ile olasılık dağılımları türetilerek bu dağılımlardan genelleştirilmiş p değerleri hesaplanmıştır.

Simülasyon çalışması Matlab R2008a program dili kullanılarak yapılmıştır ve nominal α=0.05 için elde edilen testlerin deneysel I.tip hata oranları *k=3*, *k=5* ve *k*=7 için sırasıyla Tablo 3.1, 3.2, 3.3'de verilmiştir. Bu sonuçlarla ilgili yorumlar aşağıda verilmiştir.

Tablo 3.1 Nominal α =0.05, *k=3,* farklı örneklem hacimleri ve farklı yığın varyansları için test istatistiklerinin deneysel I.tip hata oranları

n_i	σ_i	KF	W	SS	BF	GF	XW
(4,4,4)	(1,1,1)	0.0494	0.0422	0.0360	0.0348	0.0324	0.0150
	(4,4,4)	0.0498	0.0432	0.0366	0.0374	0.0342	0.0158
	(9,9,9)	0.0492	0.0388	0.0366	0.0338	0.0332	0.0132
	(1,1.25,1.5)	0.0565	0.0443	0.0386	0.0417	0.0363	0.0166
	(1,2,4)	0.0798	0.0521	0.0975	0.0571	0.0431	0.0368
	(1,4,9)	0.0932	0.0674	0.2996	0.0588	0.0652	0.0446
(10,10,10)	(1,1,1)	0.0498	0.0492	0.0314	0.0490	0.0460	0.0370
	(4,4,4)	0.0496	0.0526	0.0322	0.0501	0.0480	0.0380
	(9,9,9)	0.0499	0.0532	0.0292	0.0502	0.0470	0.0394
	(1,1.25,1.5)	0.0548	0.0522	0.0376	0.0506	0.0482	0.0426
	(1,2,4)	0.0730	0.0502	0.1446	0.0576	0.0494	0.0610
	(1,4,9)	0.0760	0.0448	0.4202	0.0606	0.0478	0.0680
(30,30,30)	(1,1,1)	0.0474	0.0490	0.0224	0.0474	0.0466	0.0532
	(4,4,4)	0.0476	0.0479	0.0210	0.0476	0.0466	0.0480
	(9,9,9)	0.0474	0.0490	0.0224	0.0474	0.0466	0.0504
	(1,1.25,1.5)	0.0490	0.0476	0.0260	0.0480	0.0464	0.0640
	(1,2,4)	0.0694	0.0499	0.1481	0.0645	0.0493	0.0708
	(1,4,9)	0.0724	0.0510	0.4504	0.0658	0.0510	0.0756
(3,5,7)	(1,1,1)	0.0496	0.0508	0.0172	0.0446	0.0390	0.0166
	(4,4,4)	0.0476	0.0510	0.0155	0.0452	0.0396	0.0160
	(9,9,9)	0.0485	0.0504	0.0144	0.0396	0.0342	0.0178
	(1,1.25,1.5)	0.0336	0.0500	0.0174	0.0410	0.0322	0.0166
	(1,2,4)	0.0292	0.0512	0.0264	0.0574	0.0406	0.0248
	(1,4,9)	0.0332	0.0566	0.0350	0.0638	0.0486	0.0452
	(1.5,1.25,1)	0.0792	0.0586	0.0214	0.0500	0.0448	0.0266
	(4,2,1)	0.1852	0.0680	0.0908	0.0712	0.0568	0.0368
	(9,4,1)	0.2332	0.0766	0.3554	0.0728	0.0734	0.0432
	(1,1,1)	0.0486	0.0492	0.0302	0.0454	0.0448	0.0158
	(4,4,4)	0.0496	0.0504	0.0298	0.0408	0.0452	0.0358
(7,10,13)	(9,9,9)	0.0490	0.0490	0.0296	0.0470	0.0454	0.0396
	(1,1.25,1.5)	0.0388	0.0460	0.0344	0.0476	0.0426	0.0142
	(1,2,4)	0.0300	0.0500	0.1284	0.0570	0.0446	0.0560
	(1,4,9)	0.0314	0.0534	0.3746	0.0632	0.0524	0.0640
	(1.5,1.25,1)	0.0816	0.0560	0.0364	0.0580	0.0388	0.0390
	(4,2,1)	0.1462	0.0540	0.1360	0.0632	0.0526	0.0470
	(9,4,1)	0.1688	0.0548	0.4136	0.0666	0.0580	0.0534
(20,25,30)	(1,1,1)	0.0494	0.0534	0.0254	0.0540	0.0506	0.0500
	(4,4,4)	0.0505	0.0460	0.0226	0.0474	0.0442	0.0506
	(9,9,9)	0.0496	0.0478	0.0224	0.0468	0.0462	0.0536
	(1,1.25,1.5)	0.0418	0.0528	0.0314	0.0540	0.0502	0.0570
	(1,2,4)	0.0386	0.0470	0.1370	0.0598	0.0458	0.0672
	(1,4,9)	0.0394	0.0470	0.4254	0.0634	0.0492	0.0684
	(1.5,1.25,1)	0.0694	0.0521	0.0305	0.0559	0.0509	0.0478
	(4,2,1)	0.1110	0.0482	0.1446	0.0652	0.0486	0.0454
	(9,4,1)	0.1248	0.0464	0.4456	0.0678	0.0484	0.0470

Tablo 3.2 Nominal α =0.05, *k=5,* farklı örneklem hacimleri ve farklı yığın varyansları için test istatistiklerinin deneysel I.tip hata oranları

n_i	σ_i	K	W	SS	BF	GF	XW
(4,4,4,4,4)	(1,1,1,1,1)	0.0486	0.0472	0.0494	0.0334	0.0618	0.0162
	(4,4,4,4,4)	0.0480	0.0472	0.0472	0.0344	0.0568	0.0120
	(9,9,9,9,9)	0.0462	0.0522	0.0510	0.0318	0.0640	0.0108
	(18,18,18,18,18)	0.0478	0.0568	0.0574	0.0360	0.0702	0.0114
	(1,1.25,1.5,1.75,2)	0.0638	0.0634	0.0640	0.0442	0.0740	0.0226
	(1,2,4,6,8)	0.0920	0.0804	0.1899	0.0571	0.0936	0.0342
	(1,4,9,13,18)	0.0978	0.0856	0.4198	0.0556	0.1092	0.0358
(10,10,10,10,10)	(1,1,1,1,1)	0.0486	0.0518	0.0420	0.0468	0.0586	0.0522
	(4,4,4,4,4)	0.0488	0.0520	0.0404	0.0470	0.0580	0.0484
	(9,9,9,9)	0.0498	0.0548	0.0436	0.0488	0.0616	0.0508
	(18,18,18,18,18)	0.0474	0.0474	0.0434	0.0464	0.0568	0.0524
	(1,1.25,1.5,1.75,2)	0.0626	0.0518	0.0494	0.0570	0.0610	0.0586
	(1,2,4,6,8)	0.0852	0.0514	0.1844	0.0566	0.0644	0.0760
	(1,4,9,13,18)	0.0880	0.0516	0.5476	0.0664	0.0684	0.0768
(30,30,30,30,30)	(1,1,1,1,1)	0.0494	0.0484	0.0300	0.0490	0.0510	0.0666
	(4,4,4,4,4)	0.0494	0.0500	0.0310	0.0490	0.0538	0.0736
	(9,9,9,9,9)	0.0504	0.0550	0.0318	0.0496	0.0570	0.0688
	(18,18,18,18,18)	0.0472	0.0486	0.0346	0.0512	0.0514	0.0758
	(1,1.25,1.5,1.75,2)	0.0582	0.0484	0.0376	0.0562	0.0516	0.0832
	(1,2,4,6,8)	0.0830	0.0494	0.2822	0.0790	0.0542	0.1044
	(1,4,9,13,18)	0.0854	0.0502	0.5812	0.0814	0.0502	0.1000
(3,4,5,6,7)	(1,1,1,1,1)	0.0466	0.0592	0.0326	0.0378	0.0656	0.0132
	(4,4,4,4,4)	0.0500	0.0630	0.0334	0.0394	0.0686	0.0164
	(9,9,9,9,9)	0.0504	0.0636	0.0336	0.0410	0.0666	0.0108
	(18,18,18,18,18)	0.0498	0.0678	0.0356	0.0438	0.0782	0.0150
	(1,1.25,1.5,1.75,2)	0.0356	0.0548	0.0330	0.0454	0.0650	0.0210
	(1,2,4,6,8)	0.0380	0.0606	0.0616	0.0650	0.0688	0.0486
	(1,4,9,13,18)	0.0326	0.0682	0.0720	0.0620	0.0790	0.0536
	(2,1.75,1.5,1.25,1)	0.1088	0.0804	0.0426	0.0554	0.0884	0.0168
	(8,6,4,2,1)	0.2114	0.0862	0.2180	0.0580	0.1000	0.0404
	(18,13,9,4,1)	0.2236	0.0920	0.5140	0.0618	0.1124	0.0492
(7,9,11,13,15)	(1,1,1,1,1)	0.0506	0.0521	0.0420	0.0462	0.0608	0.0470
	(4,4,4,4,4)	0.0508	0.0530	0.0430	0.0494	0.0622	0.0470
	(9,9,9,9,9)	0.0504	0.0535	0.0466	0.0496	0.0672	0.0474
	(18,18,18,18,18)	0.0502	0.0538	0.0436	0.0484	0.0632	0.0448
	(1,1.25,1.5,1.75,2)	0.0378	0.0522	0.0518	0.0596	0.0584	0.0588
	(1,2,4,6,8)	0.0426	0.0528	0.2398	0.0770	0.0602	0.0860
	(1,4,9,13,18)	0.0448	0.0580	0.4958	0.0804	0.0702	0.0822
	(2,1.75,1.5,1.25,1)	0.0926	0.0526	0.0488	0.0522	0.0638	0.0442
	(8,6,4,2,1)	0.1744	0.0612	0.2820	0.0650	0.0724	0.0460
	(18,13,9,4,1)	0.1832	0.0588	0.5740	0.0672	0.0732	0.0526
(20,23,26,29,32)	(1,1,1,1,1)	0.0486	0.0564	0.0304	0.0494	0.0590	0.0570
	(4,4,4,4,4)	0.0474	0.0478	0.0310	0.0490	0.0510	0.0622
	(9,9,9,9,9)	0.0510	0.0542	0.0324	0.0516	0.0564	0.0668
	(18,18,18,18,18)	0.0498	0.0522	0.0328	0.0514	0.0536	0.0610
	(1,1.25,1.5,1.75,2)	0.0438	0.0506	0.0390	0.0568	0.0524	0.0834
	(1,2,4,6,8)	0.0466	0.0436	0.2486	0.0726	0.0472	0.0982
	(1,4,9,13,18)	0.0496	0.0502	0.5488	0.0790	0.0556	0.0918
	(2,1.75,1.5,1.25,1)	0.0778	0.0478	0.0448	0.0558	0.0582	0.0482
	(8,6,4,2,1)	0.1360	0.0466	0.2870	0.0736	0.0518	0.0504
	(18,13,9,4,1)	0.1310	0.0494	0.5864	0.0778	0.0638	0.0504

Tablo 3.3. Nominal α =0.05, k=7, farklı örneklem hacimleri ve farklı yığın varyansları için test istatistiklerinin deneysel I.tip hata oranları

n_i	σ_i	KF	W	SS	BF	GF	XW
(4,4,4,4,4,4,4)	(1,1,1,1,1,1,1)	0.0505	0.0690	0.0620	0.0345	0.1080	0.0145
	(4,4,4,4,4,4,4)	0.0605	0.0750	0.0715	0.0495	0.1120	0.0100
	(9,9,9,9,9,9,9)	0.0445	0.0700	0.0620	0.0315	0.1005	0.0085
	(18,18,18,18,18,18,18)	0.0570	0.0815	0.0665	0.0385	0.1135	0.0145
	(1,1.25,1.25,1.5,1.75,1.75,2)	0.0670	0.0795	0.0690	0.0500	0.1155	0.0165
	(1,2,2,4,6,6,8)	0.1080	0.0940	0.1545	0.0550	0.1265	0.0400
	(1,4,4,9,13,13,18)	0.1045	0.1105	0.3215	0.0555	0.1500	0.0440
(10,10,10,10,10, 10,10)	(1,1,1,1,1,1,1)	0.0560	0.0565	0.0455	0.0550	0.0770	0.0640
	(4,4,4,4,4,4,4)	0.0465	0.0545	0.0475	0.0445	0.0645	0.0505
	(9,9,9,9,9,9,9)	0.0550	0.0515	0.0425	0.0510	0.0665	0.0630
	(18,18,18,18,18,18,18)	0.0560	0.0580	0.0475	0.0550	0.0695	0.0600
	(1,1.25,1.25,1.5,1.75,1.75,2)	0.0565	0.0555	0.0540	0.0510	0.0690	0.0725
	(1,2,2,4,6,6,8)	0.0855	0.0560	0.2155	0.0715	0.0740	0.0800
	(1,4,4,9,13,13,18)	0.0835	0.0595	0.4475	0.0675	0.0880	0.0845
(30,30,30,30,30, 30,30)	(1,1,1,1,1,1,1)	0.0485	0.0555	0.0360	0.0485	0.0595	0.0790
	(4,4,4,4,4,4,4)	0.0495	0.0555	0.0350	0.0495	0.0605	0.0815
	(9,9,9,9,9,9,9)	0.0480	0.0495	0.0290	0.0475	0.0555	0.0690
	(18,18,18,18,18,18,18)	0.0530	0.0535	0.0390	0.0525	0.0595	0.0935
	(1,1.25,1.25,1.5,1.75,1.75,2)	0.0595	0.0475	0.0520	0.0570	0.0535	0.0920
	(1,2,2,4,6,6,8)	0.0840	0.0540	0.2090	0.0785	0.0585	0.1100
	(1,4,4,9,13,13,18)	0.0890	0.0475	0.4885	0.0835	0.0525	0.1125
(3,4,4,5,6,6,7)	(1,1,1,1,1,1,1)	0.0525	0.0780	0.0500	0.0445	0.1075	0.0135
	(4,4,4,4,4,4,4)	0.0540	0.0760	0.044	0.0425	0.1075	0.0100
	(9,9,9,9,9,9,9)	0.0550	0.0770	0.0480	0.0470	0.1005	0.0135
	(18,18,18,18,18,18,18)	0.0465	0.0815	0.0430	0.0350	0.1075	0.0135
	(1,1.25,1.25,1.5,1.75,1.75,2)	0.0335	0.0655	0.410	0.0440	0.0930	0.0240
	(1,2,2,4,6,6,8)	0.0395	0.0815	0.0835	0.0750	0.1025	0.0550
	(1,4,4,9,13,13,18)	0.0425	0.0825	0.0840	0.0735	0.1115	0.0600
	(2,1.75,1.75,1.5,1.25,1.25,1)	0.0985	0.0840	0.0435	0.0405	0.1120	0.0140
	(8,6,6,4,2,2,1)	0.2120	0.1105	0.1650	0.0570	0.1245	0.0355
	(18,13,13,9,4,4,1)	0.2255	0.1070	0.4520	0.0560	0.1510	0.0490
(7,9,9,11,13,13,15)	(1,1,1,1,1,1,1)	0.0450	0.0505	0.0455	0.0450	0.0665	0.0565
	(4,4,4,4,4,4,4)	0.0420	0.0510	0.0420	0.0405	0.0695	0.0495
	(9,9,9,9,9,9,9)	0.0525	0.0575	0.0515	0.0515	0.0705	0.0605
	(18,18,18,18,18,18,18)	0.0475	0.0535	0.0440	0.0430	0.0705	0.0465
	(1,1.25,1.25,1.5,1.75,1.75,2)	0.0370	0.0560	0.0560	0.0560	0.0680	0.0680
	(1,2,2,4,6,6,8)	0.0415	0.0485	0.1850	0.0860	0.0625	0.1020
	(1,4,4,9,13,13,18)	0.0480	0.0600	0.4195	0.0885	0.0795	0.1005
	(2,1.75,1.75,1.5,1.25,1.25,1)	0.0965	0.0575	0.0550	0.0535	0.0745	0.0430
	(8,6,6,4,2,2,1)	0.1770	0.0580	0.2115	0.0710	0.0765	0.0430
	(18,13,13,9,4,4,1)	0.1965	0.0560	0.4605	0.0705	0.0760	0.0445
(20,23,23,26,29,29,32)	(1,1,1,1,1,1,1)	0.0470	0.0465	0.0305	0.0445	0.0515	0.0620
	(4,4,4,4,4,4,4)	0.0505	0.0530	0.0355	0.0515	0.0590	0.0805
	(9,9,9,9,9,9,9)	0.0480	0.0545	0.0350	0.0490	0.0600	0.0745
	(18,18,18,18,18,18,18)	0.0455	0.0470	0.0335	0.0440	0.0535	0.0795
	(1,1.25,1.25,1.5,1.75,1.75,2)	0.0485	0.0485	0.0385	0.0660	0.0555	0.0970
	(1,2,2,4,6,6,8)	0.0545	0.0505	0.2035	0.0805	0.0580	0.1090
	(1,4,4,9,13,13,18)	0.0575	0.0425	0.4585	0.0865	0.0535	0.1075
	(2,1.75,1.75,1.5,1.25,1.25,1)	0.0810	0.0510	0.0425	0.0585	0.0565	0.0610
	(8,6,6,4,2,2,1)	0.1325	0.0575	0.2230	0.0795	0.0670	0.0530
	(18,13,13,9,4,4,1)	0.1365	0.0560	0.4715	0.0885	0.0660	0.0540

Örneklem hacimleri eşit ve yığın varyansları heterojen olduğu durum altında testler incelendiğinde, W, BF ve GF testlerinin deneysel I.tip hata oranlarının nominal α =0.05 değerine yakın sonuçlar verdiği gözlenmektedir. XW testi için bu oranın, örnek hacimleri arttıkça nominal α =0.05 değerinden uzaklaştığı görülmektedir. Örneğin yığın varyansları $\sigma_1^2 = 1, \sigma_2^2 = 4,\ \sigma_3^2 = 9$ için örnek hacmi n_1=n_2=n_3=4 olduğunda XW testinin deneysel I.tip hata oranı 0.0446 iken n_1=n_2=n_3=30 olduğunda 0.0756 olduğu gözlenmektedir. SS testinin ise diğer testlere göre heterojenlikten çok fazla olumsuz yönde etkilendiği görülmektedir. Yığın sayısı arttığında ise örneklem hacimleri küçükken BF testinin; örneklem hacimleri arttığında da W ve GF testlerinin deneysel I.tip hata oranlarının nominal α =0.05 değerine oldukça yakın sonuçlar verdiği görülmektedir.

Örneklem hacimleri ile yığın varyanslarının doğru orantılı olduğu durum altında testler incelendiğinde, W, BF ve GF testlerinin deneysel I.tip hata oranlarının nominal α =0.05 değerine oldukça yaklaştığı görülmektedir. Ayrıca XW testi için bu oran, heterojenlik az olduğunda nominal α =0.05 değerinden uzaklaşırken heterojenlik arttığında bu değere yaklaşmıştır. Yığın sayısı arttığında ise KF testi olumlu yönde etkilenirken BF ve XW testleri bu durumdan olumsuz yönde etkilenmiş olduğu gözlenmektedir.

Örneklem hacimleri ile yığın varyanslarının ters orantılı olduğu durum altında testler incelendiğinde örneklem hacimleri küçükken XW testinin, örneklem hacmi arttıkça bu teste ek olarak G*F* ve W testlerinin deneysel I.tip hata oranlarının nominal α =0.05 değerine oldukça yakın sonuçlar verdiği görülmektedir. Yığın sayısı arttığında da benzer sonuçların elde edildiği gözlenmiştir.

3.2 Test İstatistiklerinin Güç Bakımından Karşılaştırılması

Simülasyon çalışmasında testlerin güç bakımından karşılaştırılması amacıyla farklı ortalamalı normal dağılımlı yığınlardan n_i hacimli örneklemler üretilmiştir $(i=1...k,\ j=1,...,n_i)$. $k=3$, $k=5$ ve k=7 sayıdaki yığınlar için farklı örneklem hacimleri ve farklı yığın varyanslarının çeşitli kombinasyonları kullanılmıştır. Bu örneklemlerin her biri için gerçekte yanlış olan ortalamaların eşitliği hipotezinin testi amacıyla 2. bölümde bahsedilen test istatistiklerinin değerleri hesaplanmış ve bu testlerin her biri için p değerleri bulunmuştur. Bu işlemler 5000 kez tekrarlanarak nominal α =0.05 için her bir test istatistiğinin ortalamaların eşitliği hipotezini red etme sayıları saptanmış ve bunlar tekrar sayısına yani 5000'e bölünerek her bir test istatistiğinin deneysel güç değerleri (oranları) hesaplanmıştır. Bu sonuçlar Tablo 3.4, 3.5 ve 3.6'da özetlenmiştir.

Tablo 3.4. Nominal α =0.05, *k=3,* farklı örneklem hacimleri ve farklı yığın varyansları için test istatistikleri için güç değerleri

n_i	σ_i	(μ_1, μ_2, μ_3)	KF	W	SS	BF	GF	XW
(4,4,4)	(1,4,9)	(-1.5,0,1.5)	0.120	0.105	0.431	0.072	0.112	0.049
		(-4.5,0,4.5)	0.430	0.481	0.910	0.274	0.541	0.182
		(-7.5,0,7.5)	0.752	0.843	1.000	0.581	0.924	0.461
		(-10.5,0,10.5)	0.952	0.994	1.000	0.830	1.000	0.712
(30,30,30)	(1,4,9)	(-1.5,0,1.5)	0.368	0.615	0.949	0.347	0.615	0.317
		(-4.5,0,4.5)	1.000	1.000	1.000	1.000	1.000	0.997
		(-7.5,0,7.5)	1.000	1.000	1.000	1.000	1.000	1.000
		(-10.5,0,10.5)	1.000	1.000	1.000	1.000	1.000	1.000
(3,5,7)	(1,4,9)	(-1.5,0,1.5)	0.053	0.132	0.032	0.101	0.112	0.062
		(-4.5,0,4.5)	0.281	0.724	0.083	0.439	0.695	0.355
		(-7.5,0,7.5)	0.731	0.991	0.190	0.874	0.988	0.797
		(-10.5,0,10.5)	0.966	1.000	0.392	0.994	1.000	0.981
	(9,4,1)	(-1.5,0,1.5)	0.278	0.099	0.522	0.083	0.116	0.049
		(-4.5,0,4.5)	0.673	0.458	0.956	0.248	0.592	0.232
		(-7.5,0,7.5)	0.946	0.866	1.000	0.466	0.946	0.600
		(-10.5,0,10.5)	0.996	0.985	1.000	0.635	0.999	0.886
(20,25,30)	(1,4,9)	(-1.5,0,1.5)	0.256	0.577	0.948	0.332	0.577	0.341
		(-4.5,0,4.5)	0.998	1.000	1.000	0.999	1.000	0.997
		(-7.5,0,7.5)	1.000	1.000	1.000	1.000	1.000	1.000
		(-10.5,0,10.5)	1.000	1.000	1.000	1.000	1.000	1.000
	(9,4,1)	(-1.5,0,1.5)	0.402	0.495	0.925	0.271	0.509	0.474
		(-4.5,0,4.5)	0.999	1.000	1.000	0.998	1.000	1.000
		(-7.5,0,7.5)	1.000	1.000	1.000	1.000	1.000	1.000
		(-10.5,0,10.5)	1.000	1.000	1.000	1.000	1.000	1.000

Tablo 3.5. Nominal α =0.05, *k=5,* farklı örneklem hacimleri ve farklı yığın varyansları için test istatistikleri için güç değerleri

n_i	σ_i	(μ_1, μ_2, μ_3, μ_4, μ_5)	**KF**	**W**	**SS**	**BF**	**GF**	**XW**
(4,4,4,4,4)	(1,4,9,13,18)	(-3,-1.5,0,1.5,3)	0.115	0.129	0.638	0.067	0.178	0.051
		(-9,-4.5,0,4.5,9)	0.426	0.665	0.995	0.283	0.835	0.168
		(-15,-7.5,0,7.5,15)	0.839	0.985	1.000	0.676	0.997	0.393
		(-21,-10.5,0,10.5,21)	0.982	1.000	1.000	0.912	1.000	0.636
(30,30,30,30, 30)	(1,4,9,13,18)	(-3,-1.5,0,1.5,3)	0.413	0.835	0.998	0.395	0.849	0.334
		(-9,-4.5,0,4.5,9)	1.000	1.000	1.000	1.000	1.000	0.993
		(-15,-7.5,0,7.5,15)	1.000	1.000	1.000	1.000	1.000	1.000
		(-21,-	1.000	1.000	1.000	1.000	1.000	1.000
(3,,4,5,6,7)	(1,4,9,13,18)	(-3,-1.5,0,1.5,3)	0.059	0.150	0.097	0.089	0.177	0.060
		(-9,-4.5,0,4.5,9)	0.319	0.857	0.360	0.452	0.913	0.338
		(-15,-7.5,0,7.5,15)	0.814	0.999	0.733	0.906	1.000	0.761
		(-21,-10.5,0,10.5,21)	0.986	1.000	0.948	0.995	1.000	0.964
	(18,13,,9,4,1)	(-3,-1.5,0,1.5,3)	0.298	0.156	0.757	0.074	0.226	0.038
		(-9,-4.5,0,4.5,9)	0.705	0.716	0.999	0.264	0.900	0.089
		(-15,-7.5,0,7.5,15)	0.972	0.991	1.000	0.602	1.000	0.158
		(-21,-10.5,0,10.5,21)	0.998	1.000	1.000	0.815	1.000	0.260
(20,23,26,29,32)	(1,4,9,13,18)	(-3,-1.5,0,1.5,3)	0.270	0.801	0.994	0.352	0.793	0.349
		(-9,-4.5,0,4.5,9)	1.000	1.000	1.000	1.000	1.000	0.997
		(-15,-7.5,0,7.5,15)	1.000	1.000	1.000	1.000	1.000	1.000
		(-21,-10.5,0,10.5,21)	1.000	1.000	1.000	1.000	1.000	1.000
	(18,13,,9,4,1)	(-3,-1.5,0,1.5,3)	0.451	0.723	0.992	0.309	0.753	0.703
		(-9,-4.5,0,4.5,9)	1.000	1.000	1.000	1.000	1.000	1.000
		(-15,-7.5,0,7.5,15)	1.000	1.000	1.000	1.000	1.000	1.000
		(-21,-10.5,0,10.5,21)	1.000	1.000	1.000	1.000	1.000	1.000

Tablo 3.6. Nominal α =0.05, k=7, farklı örneklem hacimleri ve farklı yığın varyansları için test istatistikleri için güç değerleri

n_i	σ_i	(μ_1, μ_2, μ_3, μ_4, μ_5, μ_6, μ_7)	**KF**	**W**	**SS**	**BF**	**GF**	**XW**
(4,4,4,4,4,4,4)	(1,4,4,9,13,13,18)	(-3,-1.5,-1.5,0,1.5,1.5,3)	0.121	0.164	0.619	0.062	0.247	0.047
		(-9,-4.5,-4.5,0,4.5,4.5,9)	0.467	0.803	0.998	0.304	0.932	0.162
		(-15,-7.5,-7.5,0,7.5,7.5,15)	0.874	0.999	1.000	0.723	1.000	0.373
		(-21,-10.5,-10.5,0,10.5,10.5,21)	0.994	1.000	1.000	0.947	1.000	0.603
(30,30,30,30,30,3	(1,4,4,9,13,13,18)	(-3,-1.5,-1.5,0,1.5,1.5,3)	0.456	0.927	0.998	0.441	0.937	0.359
		(-9,-4.5,-4.5,0,4.5,4.5,9)	1.000	1.000	1.000	1.000	1.000	0.996
		(-15,-7.5,-7.5,0,7.5,7.5,15)	1.000	1.000	1.000	1.000	1.000	1.000
		(-21,-10.5,-10.5,0,10.5,10.5,21)	1.000	1.000	1.000	1.000	1.000	1.000
(3,4,4,5,6,6,7)	(1,4,4,9,13,13,18)	(-3,-1.5,-1.5,0,1.5,1.5,3)	0.065	0.177	0.132	0.101	0.236	0.097
		(-9,-4.5,-4.5,0,4.5,4.5,9)	0.356	0.937	0.528	0.469	0.964	0.403
		(-15,-7.5,-7.5,0,7.5,7.5,15)	0.871	1.000	0.913	0.937	1.000	0.791
		(-21,-10.5,-10.5,0,10.5,10.5,21)	0.997	1.000	0.997	0.999	1.000	0.969
	(18,13,13,9,4,4,1)	(-3,-1.5,-1.5,0,1.5,1.5,3)	0.304	0.208	0.781	0.0803	0.302	0.083
		(-9,-4.5,-4.5,0,4.5,4.5,9)	0.739	0.882	1.000	0.324	0.982	0.641
		(-15,-7.5,-7.5,0,7.5,7.5,15)	0.987	1.000	1.000	0.686	1.000	0.996
		(-21,-10.5,-10.5,0,10.5,10.5,21)	0.999	1.000	1.000	0.908	1.000	1.000
(20,23,23,26,29,29,32)	(1,4,4,9,13,13,18)	(-3,-1.5,-1.5,0,1.5,1.5,3)	0.289	0.865	0.998	0.371	0.884	0.375
		(-9,-4.5,-4.5,0,4.5,4.5,9)	1.000	1.000	1.000	1.000	1.000	0.996
		(-15,-7.5,-7.5,0,7.5,7.5,15)	1.000	1.000	1.000	1.000	1.000	1.000
		(-21,-10.5,-10.5,0,10.5,10.5,21)	1.000	1.000	1.000	1.000	1.000	1.000
	(18,13,13,9,4,4,1)	(-3,-1.5,-1.5,0,1.5,1.5,3)	0.512	0.860	0.997	0.379	0.886	0.875
		(-9,-4.5,-4.5,0,4.5,4.5,9)	1.000	1.000	1.000	1.000	1.000	1.000
		(-15,-7.5,-7.5,0,7.5,7.5,15)	1.000	1.000	1.000	1.000	1.000	1.000
		(-21,-10.5,-10.5,0,10.5,10.5,21)	1.000	1.000	1.000	1.000	1.000	1.000

Testlerin deneysel I.tip hata oranlarının nominal α =0.05 değerinden oldukça uzaklaştığı durumlarda ilgili testin güç bakımından değerlendirmesi dikkate alınmamıştır. Buna göre testleri güç bakımından genel olarak değerlendirildiğinde yığın sayısı k=3 iken GF ve W testleri ile örnek hacimleri ve yığın varyanslarının ters orantılı olduğu durumda da XW testinin yüksek güç değerlerine sahip olduğu görülmektedir. Hem yığın sayısının artışının hem de örneklem hacminin küçük olmasının testlerin güç değerlerini oldukça etkileyerek birbirinden farklılık gösterdiği gözlenmiştir. Örneklem hacmi küçükken BF ve özellikle XW testlerinin diğer testlere göre yüksek güç değerlerine sahip olduğu görülmektedir. Örneklem hacmi arttığında ise W ve GF testleri ile örneklem hacimleri ile yığın varyanslarının ters orantılı olduğu durumda da XW testinin yüksek güç değerlerine sahip olduğu görülmektedir. Testin gücü örneklem hacminin artan bir fonksiyonu olduğundan örneklem hacmi büyüdükçe testlerin güç değerlerinin 1 değerine yaklaşması beklenen bir sonuçtur.

4. SONUÇ

Testler karşılaştırılırken hangi testin daha iyi olduğunu anlamak için sadece deneysel I.tip hata oranlarına değil aynı zamanda güç değerlerine de bakılması gerekmektedir. Yığın varyansları heterojen olduğunda örneklem hacimleri küçükken KF testinin deneysel I.tip hata oranları nominal α =0.05 değerinden oldukça uzaklaştığından yorumlarda bu testin güç bakımından karşılaştırılması yapılmamıştır. Örneklem hacimleri arttığında KF testinin deneysel I.tip hata oranları nominal α =0.05 değerine yakın sonuçlar vermesine rağmen, diğer testlere göre güç değerleri daha düşüktür.

Genel olarak durumları değerlendirdiğimizde örneklem hacimlerinin küçük olması tüm testlerin güç değerlerini düşürmüştür. Örneklem hacimleri arttığında beklenildiği gibi testlerin güç değerleri artmaktadır. Diğer bir ifadeyle yığından daha büyük hacimli örneklem almanın heterojenliğin etkisini azalttığı söylenebilir. Testleri güç bakımından tüm durumlar altında incelediğimizde, genellikle SS ve KF testlerinin deneysel I.tip hata oranları nominal α değerinden oldukça uzak olduğu için değerlendirmeye alınmamıştır. W ve özellikle GF testleri genel olarak diğer testlere göre daha yüksek güç değerlerine sahip olduğu gözlenmiştir. Ayrıca XW testi örneklem hacimleri ile yığın varyanslarının ters ilişkili olduğu durumda özellikle örneklem hacmi arttığında oldukça yüksek güç değerlerine sahip olduğunu görülmektedir.

KAYNAKLAR

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

Chen, S. ve Chen, J.H. (1998). Single-Stage analysis of variance under heteroscedasticity. *Communications in Statistics and Simulations* 27(3), 641-666.

Gamage, J. ve Weerahandi, S. (1998). Size performance of some tests in one-way ANOVA *Communications in Statistics and Simulations* 27(3), 625-640.

Gerami, A. ve Zahedian, A. (2001). Comparing the means of normal populations with unequal variances. roceedings of the 53rd Session of Intenational Statistical Institute**,** Seoul, Korea.

Scheffe, H. (1959). *The analysis of variance*, Wiley, New York, 30-80.

Scott, A.J. ve Smith, T.M.F. (1971). Interval estimates for linear combinations of means. *Applied Statistics* 20(3), 276-285.

Weerahandi, S. (1995a). ANOVA under unequal error variances. *Biometrica* 38, 330-336.

Weerahandi, S. (1995). *Exact statistical method for data analysis*. Springer-Verlag, NewYork, 2-50.

Weerahandi, S. (2004). *Generalized inference in repeated measures: Exact methods in MANOVA and mixed models*. Wiley, New York, 1-60.

Welch, B.L. (1947). The generalization of student's problem when several different population variances are involved. *Biometrika* 34, 28-35.

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

Xu, L. ve Wang, S. (2007). A new generalized p-value and its upper bound for ANOVA under unequal erros variances. *Communications in Statistics Theory and Methods* 37, 1002-1010.