

**VARYANSLARIN HOMOJENLİĞİNDE KULLANILAN BAZI YÖNTEMLERİN
I.TİP HATA VE TESTİN GÜCÜ BAKIMINDAN KARŞILAŞTIRILMASI: BİR
MONTE CARLO SİMÜLASYON ÇALIŞMASI****Benian TEKİNDAL¹****ÖZET**

Bu çalışmada, değişik deneme koşullarında (dağılım şekli, varyans oranları, gözlem sayısı, grup sayısı vb) Brown-Forsythe, Cochran, Fmax ve Satterwaite gibi bazı varyans homojenliğini test yöntemlerinin ampirik olarak gerçekleşen 1.Tip hata olasılıkları ve testin gücü bakımından karşılaştırılması yapılmıştır. Bu amaçla 100.000 denemeden oluşan Monte Carlo simülasyonları yapılmıştır. Yapılan simülasyon denemeleri sonucunda, genel olarak bütün deneme koşulları için geçerli olabilecek hiçbir testin önerilemeyeceği, her testin farklı deneme koşullarında kullanılmasının uygun olabileceği sonucuna varılmıştır. Populasyon dağılımları normal olduğunda, örnek hacmi ne olursa olsun karşılaştırılan 1.Tip hata olasılığının deneme sonunda da korunması bakımından en güvenilir sonuçları Cochran ve Fmax testlerinin; normal olmayan dağılımlarda ise en güvenilir sonuçları Satterwaite testinin gerçekleştirdiği görülmüştür. Diğer yandan söz konusu testlerin güç değerlerinin, dağılım şekline göre çok varyans oranları ve örnek hacminden etkilendiği görülmüştür.

Anahtar Kelimeler: İstatistik Test, Varyansların Homojenliği, 1.Tip Hata, Testin Gücü, Simetrik Dağılım, Varyans Analizi

**TYPE I ERROR RATES AND TEST POWER OF SOME VARIANCE
HOMOGENEITY TESTS UNDER DIFFERENT EXPERIMENTAL CONDITIONS:
A MONTE CARLO SIMULATION STUDY****ABSTRACT**

In this study, some homogeneity tests such as Brown-Forsythe, Cochran, Fmax and Satterwaite tests were compared for Type I error rate and test power under various experimental conditions. Data used in the study were obtained from Monte Carlo simulation which were formed 100.000 trials. At the end of 100.000 simulation trials, it was observed that none of the tests studied were the best for every kind of experimental conditions. Also, it was seen that if the distributions were normal, Cochran and Fmax tests gave the closest value to the predetermined alpha level regardless of the sample size. In contrary for the non-normal distributions, Satterwaite test gave the closest value to the predetermined alpha level regardless of the sample size. On the other hand, the power levels of the tests were affected by the variance ratios and sample sizes rather than being affected by the distribution.

Keywords: Statistical Test, Homogeneity of Variance, Type I Error Rate, Power of Test, Symmetrical Distribution, Variance Analysis

¹Gazi Üniversitesi, Endüstriyel Sanatlar Eğitim Fakültesi, Bilgisayar Eğitimi Bölümü, 06500 Beşevler/Ankara. benian@gazi.edu.tr

1. GİRİŞ

Ziraat, tıp, biyoloji ve sosyal bilimler gibi bir çok alanda bağımsız iki ve daha fazla muamele grubu ortalaması arasındaki farkın irdelenmesinde yaygın olarak kullanılan varyans analizi tekniğinden (ANOVA F testi) beklenen yararların sağlanabilmesi, çalışılan veri setinde normallik, varyansların homojenliği, gözlemlerin bağımsızlığı ve etkilerin eklenebilir olması gibi bazı varsayımlarının sağlanmasına bağlıdır (Conover ve ark., 1981; Sokal ve Rohlf, 1995; Zar, 1999; Mendeş, 2002). Bunlardan deneme sonuçlarının hassasiyetini en fazla etkileyen varsayımlar ise normallik ve varyansların homojenliğidir (Keskin, 2002; Mendeş, 2003). Çünkü, bu iki varsayım çalışılan popülasyona bağlı oldukları için deneme başında yerine getirilmeleri mümkün olamamaktadır. Diğer yandan, elde edilen verilerde normallikten sapmalar fazla olmadıkça ya da örnek yeterince büyük ve eşit hacimli örneklerle çalışılıyorsa normallik varsayımlarının sağlanıp sağlanmaması elde edilen sonuçların hassasiyetini pek etkilememektedir. Ancak, aynı şeylerin varyansların homojenliği varsayımı için söylenmesi pek mümkün değildir. Çünkü, bu varsayımın yerine gelmemesi, analiz sonuçlarının önemlilik seviyesi ve duyarlılığını yani 1.Tip hata ve testin gücünü çok büyük ölçüde etkilemektedir (Cochran ve Cox 1957; Alexander ve Govern, 1994; Lim ve Loh, 1996; Mendeş, 2002). Buradan hareketle “varyansların homojenliği” pek çok araştırmacı tarafından varyans analizinin en önemli varsayımı olduğu bildirilmektedir (Bishop ve Dudewicz 1978; Wilcox, 1989; Steel ve ark.,1997; Ott 1998).

Varyansların homojenliğinin varyans analizi tekniğinin en önemli varsayımı olarak kabul edilmesi, bu varsayımın sağlanıp sağlanmadığının doğru olarak belirlenebilmesi gereğini ortaya çıkartmış ve bu amaç doğrultusunda çok sayıda test yöntemi geliştirilmiştir (Conover ve ark., 1981; Çamdeviren ve Mendeş, 2005). Bu testlerin birçoğu birbirinin değiştirilmiş şekli olup söz konusu testlerinin gerek 1.Tip hata ve gerekse de testin gücü bakımından birbirleri ile karşılaştırıldığı birçok çalışma yapılmıştır (Conover ve ark., 1981; Piepho, 1996; Keskin, 2002; Mendeş, 2003). Ancak, söz konusu çalışmalar oldukça dar bir çerçevede yapılmış olduklarından hangi testin hangi deneme koşulunda kullanılması halinde, elde edilecek sonuçların güvenilirliği açısından daha iyi olabileceği hakkında bazı çelişiklere düşülebilmekte ve dolayısıyla hangi test yönteminin hangi durumlarda kullanılabileceğine ilişkin karar vermek zorlaşmaktadır. Söz konusu test yöntemlerinden yararlanılarak varyansların homojen olup olmadığına yönelik verilecek kararlar, dağılım şekli, örnek hacmi, varyans oranları, örnek hacmi ile varyans oranları arasındaki ilişkiler, grup sayısı gibi çeşitli özelliklerden etkilenir. Bu çalışmada, varyans homojenlik testlerinden Brown-Forsythe, Cochran, Hartley Fmax ve Satterwaite testleri, deneme sonunda ampirik olarak gerçekleşen 1.Tip hata olasılıkları ve testin gücü bakımından birbirleriyle bir çok deneme koşulunda karşılaştırılmıştır. Bu testlerin normallikten küçük sapmalardan bile etkilenmeleri göz önüne alınarak normal dağılım dışında şekil olarak simetrik, ancak normal olmayan 10 serbestlik dereceli t dağılımı (t(10)) ve (10, 10) parametrelerine sahip Beta (Beta (10,10)) dağılımları da dikkate alınmıştır.

2. ÇALIŞMADA ELE ALINAN TESTLER

2.1. Brown-Forsythe (BF) Testi

Bu testin esası; her gruptaki gözlem değerlerinin kendi grup ortanca değerinden (median) olan sapmalarının mutlak değerlerine varyans analizi uygulanmasıdır. Y_{ij} ; j. gruptaki i. gözlem değerini ve \tilde{Y}_j 'da j. grubun ortanca değerini göstermek üzere, her gruptaki gözlem değerlerinin ortanca değerden olan sapmalarının mutlak değerleri;

$$X_{ij} = |Y_{ij} - \tilde{Y}_j| \quad (1)$$

olarak ifade edilir. Daha sonra X_{ij} değerlerine varyans analizi yapılması sonucu BF istatistiği;

$$BF = \frac{\sum_{j=1}^k n_j (\bar{X}_j - \bar{X}_{..})^2 / (k-1)}{\sum_{j=1}^k \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2 / (N-k)} \quad \left(N = \sum_{j=1}^k n_j \right) \quad (2)$$

olarak hesaplanır. Burada; N toplam gözlem sayısını, \bar{X}_j j.grubun ortalamasını, $\bar{X}_{..}$ genel ortalamayı, n_j ise j.grubun gözlem sayısını göstermektedir. Yapılan hesaplamalar sonucunda; BF istatistiği, $F_{\alpha; (k-1), (N-k)}$ tablo değerinden büyükse “grup varyansları homojendir” şeklinde kurulan H_0 hipotezi reddedilir (Brown ve Forsythe, 1974).

2.2. Cochran (COH) Testi

Ele alınan gruplarda eşit sayıda gözlemin bulunmasını gerektiren Cochran testi için test

$$\text{istatistiği; } COH = \frac{S_{\max}^2}{\sum_{j=1}^k S_j^2} \quad (3)$$

şeklinde hesaplanır. Bu şekilde hesaplanan test istatistiğinin k ve n-1'e göre düzenlenmiş Cochran tablo değerinden büyük olması durumunda H_0 hipotezi reddedilir. Normal dağılımdan olan küçük sapmalara karşı bile son derece duyarlı olan Cochran testi, şartlar sağlandığında karşılaştırılan 1.Tip hatayı koruma bakımından oldukça güvenilir bir testtir (Conover ve ark., 1981; Johnson, 1991; Zar, 1999).

2.3. Hartley veya F_{\max} (F_{\max}) Testi

Bu yöntemde de Cochran testinde olduğu gibi varyansları karşılaştırılacak muamele gruplarının eşit sayıda gözlem içermesi gerektirmektedir. Bu test her ne kadar söz konusu muamele gruplarında eşit sayıda gözlem içermesi gerektirmesi bakımından Cochran testine bir benzerlik gösteriyorsa da, test istatistiğinin, sadece en büyük varyans ile en küçük

varyansı dikkate alınarak hesaplanması bakımından bir farklılık göstermektedir. Bu farklılıkta bu testin dezavantajı olarak kabul edilmektedir. Bu test için test istatistiği;

$$F_{\max} = \frac{S_{\max}^2}{S_{\min}^2} \quad (4)$$

olarak hesaplanır. Yapılan hesaplamalar sonucunda elde edilen test istatistiği; α yanılma seviyesinde, k ve n-1'e göre düzenlenmiş Hartley tablo değeri ile karşılaştırılır ve hesaplanan değer tablo değerinden büyük ise H_0 hipotezi reddedilir (Johnson, 1991).

2.4. Satterwaite (SATT) Testi

Bu test, test istatistiğinin hesaplanması bakımından Brown-Forsythe testi ile aynı, fakat elde edilen test istatistiğinin karşılaştırılacağı tablo değeri bakımından farklıdır. Çünkü ST testinde, test istatistiğinin F tablo değeri ile karşılaştırılmasında dikkate alınan serbestlik derecesinde bir düzeltme yapılır ve söz konusu test istatistiği k ve v serbestlik dereceli F tablo değeri ile karşılaştırılır. Düzeltilmiş serbestlik derecesi olan v değeri;

$$v = \frac{\left(\sum_{j=1}^k X_j \right)^2}{\sum X_j^2 / v_j} \quad X_j = \sum_{i=1}^k (Y_{ij} - \bar{Y}_j)^2 \quad \text{ve} \quad v_j = n_j - 1 \text{ dir} \quad (5)$$

şeklinde hesaplanır (Loh, 1987).

3. MATERYAL VE YÖNTEM

Brown-Forsythe, Cochran, F_{\max} ve Satterwaite testlerinin 1.Tip hata ve Testin gücü bakımından karşılaştırılması amacıyla Monte Carlo simülasyon tekniğinden yararlanılarak, normal (0,1), Beta (10,10) ve t (10) dağılımı gösteren popülasyonlardan tesadüf sayıları üretilmiştir. Tesadüf sayıları üretilirken eşit sayıda gözlem içeren ancak farklı varyans oranlarına sahip üç ve beş popülasyonun olduğu durumlar dikkate alınmıştır. Bu çalışmada, Normal (0,1), β (10, 10) ve t (10) dağılımı gösteren popülasyonlardan ele alınan deneme koşullarına göre tesadüf sayılarının üretilmesinde Fortran Power Station Develeoper IMSL (1994) kütüphanesinden yararlanılmıştır. Söz konusu popülasyonlardan tesadüf sayılarının üretilmesinde IMSL (1994) kütüphanesinin RNNOA, RNBET ve RNSTT alt grupları kullanılmıştır.

Yukarıda parametreleri verilen popülasyonlar doğal olarak farklı ortalama ve varyansa sahiptirler. Bu durum dikkate alınarak, her popülasyondaki gözlem değerleri olarak kabul

edilen tesadüf sayıları $X_{ij} = \frac{Y_{ij} - \mu_j}{\sigma_j}$ şeklinde standardize edilmiştir. Böylece,

dağılımların şekilleri değiştirilmeden, $\mu = 0$ ve $\sigma^2 = 1$ olan dağılımlar elde edilmiştir. Bu ifade de; Y_{ij} , j. popülasyondaki i. gözlem değerini, μ_j j. popülasyonun ortalamasını, σ_j j. popülasyonun standart sapmasını ve X_{ij} de j. popülasyondaki i. gözlemin standardize edilmiş değerini göstermektedir. Çalışmada, uygulamada sık olarak karşılaşılan 3 ve 5

grubun olması durumları dikkate alınmıştır. Söz konusu grupların her birinde eşit sayıda gözlem ile çalışılmış ve gruplardaki gözlem sayıları 4,6,8,...,30 olarak alınmıştır. Çalışılan testlerin varyansları oranları 1:2:3, 1:5:10, 1:2:3:4:5 ve 1:3:6:9:12 şeklinde alınmış ve heterojenliğin oluşturulabilmesi için gözlem değerleri olarak kabul edilen standardize edilmiş tesadüf sayıları $1, \sqrt{2}, \sqrt{3}, \sqrt{4}, \sqrt{5}, \sqrt{6}, \sqrt{9}, \sqrt{10}, \sqrt{12}$ gibi sabit sayılarla çarpılmışlardır. Daha sonra çalışılan testler için ampirik olarak gerçekleşen I. tip hata olasılıkları, 100.000 simülasyon denemesi sonunda reddedilen H_0 hipotezlerinin sayılması ve yüzdeye dönüştürülmesiyle; bu testlerin güç değerleri ise grupların varyans oranlarından yararlanılmak suretiyle yine 100.000 simülasyon denemesi sonucunda reddedilen H_0 hipotezlerinin sayılarak yüzdeye dönüştürülmesiyle elde edilmiştir. Genel olarak her ne kadar %80'lik bir güç değeri yeterli güç değeri olarak kabul ediliyorsa da bazı yazarlar tarafından %50 ve daha fazla güç değeri gerçekleştiren testlerin güçlü testler olarak kabul edilebilmesinden hareketle bu çalışmada testin gücü için %70 ve daha yüksek güç değeri yeterli güç değeri olarak kabul edilmiştir (Cohen, 1988; Murphy ve Myers, 1998).

4. ARAŞTIRMA BULGULARI VE TARTIŞMA

Çalışmada, dikkate alınan deneme koşullarında söz konusu testler bakımından ampirik olarak gerçekleşen I.Tip hata olasılıkları ve testin güç değerleri Çizelge 1-3'te topluca verilmiştir.

4.1. I.Tip Hata Olasılıkları Bakımından Elde Edilen Sonuçlar

Çizelge 1 incelendiğinde, özellikle söz konusu örneklerin alınmış oldukları popülasyonların dağılımlarının normal olması ve gruplarda eşit sayıda gözlem bulunması durumunda en güvenilir sonuçları gerçekleştiren Cochran ve Fmax testlerinin, normalliğe karşı oldukça duyarlı oldukları ve normallikten olan küçük sapmalardan bile olumsuz yönde etkilendikleri görülmektedir. Bu durum, söz konusu testlerin popülasyon dağılımlarının normal olması durumunda hep %5 veya %5 civarında gerçekleşen I.Tip hata olasılıklarının, dağılımların diklik katsayılarının artmasına paralel olarak giderek %5'ten uzaklaşma eğilimine girmelerinden de açıkça görülmektedir. Mesela, normal dağılımlı popülasyonlardan alınan üç örnekte de 4'er gözlem gibi oldukça düşük sayıda gözlemin bulunması durumunda Cochran ve Fmax testleri bakımından sırasıyla % 5.0 ve %5.3 olarak gerçekleşen I.Tip hata olasılıklarının, dağılımların t (10) olması ve örneklerde 30'ar gözlemin bulunması durumunda bile sırası ile %10.9 ve %11 olarak gerçekleştiği görülmektedir. Bu bulgular Gartside (1972), Conover ve ark. (1981), Keskin (2002) ve Mendeş (2002) yapmış oldukları çalışma bulgularıyla benzerlik göstermekle birlikte, bu bulguların özellikle Gartside (1972), Conover ve ark. (1981) yaptıkları çalışmalarla simülasyon sayılarının farklılıklarından dolayı bazı küçük farklılıklar da göstermektedir. Diğer yandan teorik temelleri aynı olan, ancak serbestlik dereceleri bakımından aralarında bazı farklılıklar bulunan Brown-Forsythe ve Satterwaite testlerinden Brown-Forsythe testi dağılım şekli ne olursa olsun gruplarda 4'er gözlemin bulunması hariç, genel olarak %3.0-4.1 arasında I.Tip hata olasılıkları gerçekleştirirken, Satterwaite testinin aynı deneme koşullarında, genel olarak %3.8-4.7 arasında I.Tip hata olasılıkları gerçekleştirmektedir. Ele alınan bütün deneme koşullarında örnek hacimlerinin 4:4:4 veya 4:4:4:4 olması hariç, genel olarak Brown-Forsythe testi bakımından gerçekleşen I.tip hata olasılıklarının diğer üç test bakımından gerçekleşen I.tip hata olasılıklarından daha düşük olduğu görülmektedir (Çizelge 1). Loh (1987) yapmış olduğu çalışmada, normal dağılımlı popülasyonlardan alınan gruplarda (4 grup) 4'er adet gözlemin bulunması durumunda (1000 deneme) Brown-

Forsythe ve Satterwaite testleri bakımından gerçekleşen I. tip hata olasılıklarını % 8.0 ve % 5.3 olarak bulmuştur. Bu çalışmada ise bu değerler 5 grubun olması durumunda sırasıyla % 6.5 ve % 3.6 olarak bulunmuştur. Lim ve Loh (1996) ise yaptıkları çalışmada, (1000 deneme) gruplarda 10'ar adet gözlemin olması durumunda Brown-Forsythe testi bakımından gerçekleşen I.tip hata olasılığını %2.9 olarak bulmuştur. Bunun yanı sıra, Sharma (1991) yaptığı çalışmada (10.000 deneme), gruplarda 10'ar ve 20'şer gözlem bulunduğunda; Brown-Forsythe testine ait gerçekleşen I. tip hata olasılıklarının sırasıyla; % 3.30 ve % 3.38 olduğu belirtilmiştir. Bu çalışmada ise söz konusu I.Tip hata olasılıkları; % 2.9 ve % 3.5 olarak bulunmuştur. Bu çalışma bulguları genel olarak daha önce yapılmış çalışma bulgularını desteklemekle birlikte, söz konusu çalışmalarda dikkate alınan deneme koşullarındaki farklılıklardan dolayı (simulasyon sayısı, dağılım şekli, örnek hacmi vb gibi) bazı küçük farklılıklarda göstermektedir. Aynı zamanda bu çalışmada çalışılan testlerden hiçbirisinin grup sayısından etkilenmediği Çizelge 1'de dikkati çeken başka bir husus olarak göze çarpmaktadır. Bu söylenenlerden hareketle, ele alınan testlerden hiç birisinin dikkate alınan bütün deneme koşulları için önerilemeyeceği sonucuna varılabilir. Çünkü, çalışılan testlerden hiç birisi dikkate alınan bütün deneme koşullarında deneme başında karşılaştırılan I.tip hata olasılığını (0.05), deneme sonunda da aynı düzeyde koruyamamaktadır. Buna karşın herhangi bir öneride bulunmak gerekirse, normal dağılımlı popülasyonlarla çalışılması durumunda Cochran veya Fmax testlerinden herhangi birisini, β (10, 10) veya t (10) dağılımlı popülasyonlarla çalışılması durumunda ise, genel olarak %4 civarında I.Tip hata olasılıkları gerçekleştiren Satterwaite testinin kullanılmasının uygun olabileceği söylenebilir.

4.2. Testin Gücü Bakımından Elde Edilen Sonuçlar

Bilindiği üzere, bir testin önerilebilmesi yani sağlam (robust) bir test olarak kabul edilmesine karar verilmesinde, I.Tip hata olasılığı dışında bu testlerin güç değeri bakımından da karşılaştırılması ve bu iki durumun birlikte dikkate alınarak bir öneride bulunulmaması, elde edilecek sonuçların güvenilirliği bakımından daha doğru olacaktır. Bu durum dikkate alınarak, bu çalışmada ele alınan testler testin gücü bakımından da birbirleri ile karşılaştırılmış ve elde edilen bulgular Çizelge 2 ve Çizelge 3'te topluca sunulmuştur. Çizelge 2 incelendiğinde, örneklerin alınmış oldukları popülasyonların dağılımları ne olursa olsun, gerçekleşen güç değerleri bakımından en güvenilir sonuçları Fmax testinin verdiği görülmektedir. Çünkü bu test, varyans oranları 1:2:3 olan üç popülasyonun olması durumunda, her üç dağılım koşulunda da örneklerde 28 ve daha fazla gözlemin bulunması durumunda, popülasyon varyanslarının 1:5:10 olması dağılımların normal olması durumunda örneklerde 8'er gözlemin, dağılımların β (10, 10) ve t (10) olması durumlarında ise örneklerde 10'ar gözlemin bulunması durumunda yeterli güç değerine ulaşmaktadır. Diğer yandan, ele alınan diğer üç test popülasyon varyans oranlarının 1:2:3 olması durumunda yeterli güç değerine ulaşamamıştır. Buna karşın popülasyon varyanslarının 1:5:10 olması durumunda Brown-Forsythe testi dağılımlar normal ve t (10) iken örneklerde 14'er gözlemin bulunması durumunda, β (10, 10) iken 12'şer gözlemin bulunması durumunda yeterli güç değerine ulaşmıştır. Cochran testinin aynı koşullarda dağılım şeklinden etkilenmediği ve her üç durumda da 16'şar gözlem ile çalışılması durumunda yeterli güç değerine ulaştığı görülmektedir. Satterwaite testi ise dağılımların normal ve β (10, 10) olması durumunda 12'şer gözlem ile dağılımların t (10) olması durumunda ise 14'er gözlem ile çalışılması durumunda yeterli güç değerine ulaştığı görülmektedir. Grup sayısının 5'e çıkması durumunda da yine benzer sonuçların elde edildiği dikkati

çekmektedir. Dolayısıyla Fmax testinin diğer testlere göre birçok deneme koşulunda daha güvenilir sonuçlar gerçekleştirdiği ileri sürülebilir.

Çizelge 1. 100.000 Simülasyon Denemesi Sonucunda Ampirik Olarak Gerçekleşen 1.Tip Hata Olasılıkları (%)

Dağ. Şekli	Gözlemler	k=3				Gözlemler	k=5			
		1.Tip Hata Olasılığı (%)					1.Tip Hata Olasılığı (%)			
		BF	COCH.	Fmax	SATT		BF	COCH.	Fmax	SATT
Normal (0,1)	4:4:4	6,5	5,0	5,3	1,9	4:4:4:4:4	6,2	4,7	5,1	3,6
	6:6:6	4,0	5,0	5,2	4,1	6:6:6:6:6	3,5	4,7	4,7	4,0
	8:8:8	3,3	4,7	5,0	4,0	8:8:8:8:8:8	3,0	5,0	5,1	3,7
	10:10:10	3,2	4,8	5,0	4,1	10:10:10:10:10	2,9	4,9	4,8	3,5
	12:12:12	3,6	5,0	5,3	4,3	12:12:12:12:12	3,1	4,8	5,0	3,8
	14:14:14	3,6	5,0	5,2	4,2	14:14:14:14:14	3,1	4,8	4,9	3,7
	16:16:16	3,5	5,1	4,8	4,1	16:16:16:16:16	3,2	5,1	4,9	3,7
	18:18:18	3,6	5,0	5,2	4,2	18:18:18:18:18	3,2	5,1	4,8	3,9
	20:20:20	3,9	5,0	4,9	4,3	20:20:20:20:20	3,5	5,1	5,1	4,0
	22:22:22	3,8	5,0	5,1	4,3	22:22:22:22:22	3,8	5,1	5,5	4,3
	24:24:24	3,9	5,3	5,3	4,6	24:24:24:24:25	3,6	5,0	5,0	4,1
	26:26:26	3,6	4,7	4,8	4,1	26:26:26:26:26	3,6	5,0	4,9	4,1
	28:28:28	3,9	5,0	4,9	4,4	28:28:28:28:28	3,7	5,2	4,9	4,2
	30:30:30	4,0	4,9	4,8	4,4	30:30:30:30:30	3,7	4,8	4,9	4,1

Çizelge 1'in Devamı

Beta (10,10)	4:4:4	6,8	4,4	4,9	2,1	4:4:4:4:4	6,4	3,8	4,4	3,6
	6:6:6	3,8	3,8	4,1	4,0	6:6:6:6:6	3,8	3,7	4,0	4,2
	8:8:8	3,4	4,1	4,0	4,1	8:8:8:8:8:8	2,8	3,3	3,9	3,6
	10:10:10	3,2	3,7	3,7	3,8	10:10:10:10:10	3,0	3,3	3,6	3,6
	12:12:12	3,0	3,6	3,9	3,8	12:12:12:12:12	3,1	3,1	3,5	3,6
	14:14:14	3,3	3,6	3,9	4,0	14:14:14:14:14	3,0	3,2	3,5	3,6
	16:16:16	3,5	3,5	3,7	4,2	16:16:16:16:16	3,3	3,1	3,6	3,8
	18:18:18	3,8	3,6	3,6	4,3	18:18:18:18:18	3,2	3,2	3,1	3,6
	20:20:20	3,6	3,2	3,5	4,1	20:20:20:20:20	3,2	2,8	3,3	3,8
	22:22:22	3,6	3,5	3,6	4,1	22:22:22:22:22	3,4	3,1	3,3	3,9
	24:24:24	3,9	3,5	3,4	4,3	24:24:24:24:25	3,5	3,1	3,2	4,0
	26:26:26	4,3	3,5	3,8	4,7	26:26:26:26:26	3,5	3,2	3,2	4,0
	28:28:28	4,0	3,4	3,4	4,4	28:28:28:28:28	3,6	3,1	3,3	3,9
30:30:30	3,9	3,5	3,5	4,4	30:30:30:30:30	3,6	3,0	3,1	4,0	
t (10)	4:4:4	6,6	7,1	6,4	3,2	4:4:4:4:4	6,4	8,6	6,5	5,0
	6:6:6	3,7	8,1	7,2	4,4	6:6:6:6:6	3,5	10,2	8,2	4,2
	8:8:8	3,3	8,8	8,1	4,2	8:8:8:8:8:8	3,0	10,8	9,1	3,9
	10:10:10	3,3	9,6	8,9	4,3	10:10:10:10:10	3,1	11,9	10,1	4,0
	12:12:12	3,4	9,6	9,1	4,2	12:12:12:12:12	3,4	12,3	10,8	4,2
	14:14:14	3,3	10,0	9,6	4,0	14:14:14:14:14	2,9	12,0	11,1	3,5
	16:16:16	3,9	10,6	10,3	4,6	16:16:16:16:16	3,3	12,8	12,1	4,1
	18:18:18	3,5	10,5	10,5	4,2	18:18:18:18:18	3,7	13,2	12,2	4,3
	20:20:20	4,1	10,5	10,5	4,7	20:20:20:20:20	3,8	13,4	13,1	4,4
	22:22:22	4,1	10,9	10,8	4,6	22:22:22:22:22	3,9	13,4	12,7	4,4
	24:24:24	4,0	10,8	11,2	4,5	24:24:24:24:25	3,8	13,2	12,8	4,3
	26:26:26	4,0	10,6	10,8	4,4	26:26:26:26:26	3,9	13,3	13,3	4,4
	28:28:28	3,9	10,8	10,7	4,4	28:28:28:28:28	3,8	13,6	13,3	4,2
30:30:30	4,1	10,9	11,0	4,5	30:30:30:30:30	4,0	13,6	13,6	4,4	

BF: Brown-Forsythe testi **COCH:** Cochran testi **Fmax :** Hartley F max testi **SATT:** Satterwaite testi

Çizelge 2. 100.000 simülasyon denemesi sonucunda ampirik olarak gerçekleşen güç değerleri (%), $k=3$

Dağ. Şekli	Var. Oranı	Gözlemler	Testin gücü (%)				Var. Oranı	Testin gücü (%)			
			BF	COCH.	Fmax	SATT		BF	COCH.	Fmax	SATT
Normal (0,1)	1:2:3	4:4:4	9,6	9,6	8,7	8,8	1:5:10	17,9	23,5	27,0	23,8
		6:6:6	8,5	13,9	13,2	11,1		22,7	34,5	52,6	31,0
		8:8:8	11,3	18,2	19,4	14,4		36,2	44,4	73,1	44,4
		10:10:10	14,3	21,8	25,3	17,6		52,6	53,8	85,6	60,0
		12:12:12	18,6	25,3	31,0	21,9		67,8	61,6	92,9	73,2
		14:14:14	23,4	28,7	37,9	26,5		79,8	68,0	96,7	83,4
		16:16:16	28,0	32,3	42,4	31,0		87,9	74,3	98,2	90,1
		18:18:18	33,3	36,5	48,6	36,3		93,3	79,2	99,3	94,5
		20:20:20	38,2	39,5	53,7	41,0		96,2	83,3	99,8	96,9
		22:22:22	42,8	42,9	58,2	45,4		98,0	87,0	99,8	98,3
		24:24:24	47,8	46,0	63,6	50,2		98,9	89,5	100,0	99,1
		26:26:26	52,3	49,4	67,2	54,6		99,5	92,1	100,0	99,6
		28:28:28	57,2	52,2	71,4	59,2		99,8	93,4	100,0	99,8
30:30:30	60,8	55,2	75,1	62,6	99,9	95,2	100,0	99,9			
eta (10,10)	1:2:3	4:4:4	10,1	9,0	7,9	9,5	1:5:10	18,2	22,5	25,4	24,2
		6:6:6	9,0	12,7	12,4	11,4		23,8	32,8	51,3	31,7
		8:8:8	10,7	16,1	17,0	13,6		37,6	43,8	73,5	46,3
		10:10:10	15,0	20,3	23,2	18,4		54,5	53,3	86,4	62,2
		12:12:12	19,5	23,7	29,3	22,8		70,8	61,2	94,0	75,8
		14:14:14	24,3	27,3	35,2	27,6		81,9	68,9	97,2	85,1
		16:16:16	29,7	31,1	41,3	32,8		90,2	75,1	98,7	92,0
		18:18:18	34,7	34,8	47,3	37,4		94,6	79,9	99,6	95,6
		20:20:20	40,3	38,6	53,1	43,3		97,3	84,3	99,8	97,8
		22:22:22	45,2	41,7	58,3	48,0		98,7	87,7	99,9	98,9
		24:24:24	50,0	45,2	62,9	52,2		99,3	90,3	100,0	99,4
		26:26:26	55,1	49,0	67,3	57,0		99,6	92,6	100,0	99,7
		28:28:28	59,2	52,0	71,3	61,2		99,9	94,5	100,0	99,9
30:30:30	63,2	55,1	75,3	65,0	99,9	95,9	100,0	99,9			
t (10)	1:2:3	4:4:4	9,1	12,1	10,3	9,6	1:5:10	16,6	24,8	28,8	22,9
		6:6:6	8,1	17,2	17,0	10,7		20,8	36,3	54,1	28,8
		8:8:8	9,6	21,7	23,7	12,7		32,3	46,2	72,5	40,5
		10:10:10	13,1	25,8	30,1	16,2		47,1	54,4	84,0	54,6
		12:12:12	16,6	29,4	35,7	19,6		61,5	61,1	90,7	67,6
		14:14:14	20,6	32,5	41,4	23,6		73,5	67,1	94,7	77,7
		16:16:16	25,3	35,8	46,1	28,2		83,2	73,0	97,2	86,1
		18:18:18	29,0	39,8	51,3	31,7		89,3	77,4	98,5	91,0
		20:20:20	33,3	41,8	55,8	36,0		93,5	81,1	99,0	94,7
		22:22:22	38,5	45,6	60,8	40,9		96,1	84,7	99,5	96,8
		24:24:24	41,5	48,0	63,8	44,1		97,8	87,4	99,8	98,2
		26:26:26	45,8	50,8	67,0	48,3		98,7	89,8	99,9	98,9
		28:28:28	50,3	53,4	70,7	52,5		99,3	92,0	99,9	99,4
30:30:30	53,2	56,0	73,2	55,1	99,7	93,7	99,9	99,7			

Çizelge 3. 100.000 simülasyon denemesi sonucunda ampirik olarak gerçekleşen güç değerleri (%), k=5

Dağ. Şekli	Var. Oranı	Gözlemler	Testin gücü (%)				Var. Oranı	Testin gücü (%)			
			BF	COCH.	Fmax	SATT		BF	COCH.	Fmax	SATT
Normal (0,1)	1:2:3:4:5	4:4:4:4:4	12,4	13,9	11,4	15,3	1:3:6:9:12	18,9	22,0	23,0	25,7
		6:6:6:6:6	12,0	18,8	20,4	16,2		24,6	31,5	52,4	32,3
		8:8:8:8:8:8	16,9	24,0	32,8	21,3		38,8	40,0	75,5	46,5
		10:10:10:10:10	24,2	29,3	44,3	28,9		57,0	49,8	89,5	63,5
		12:12:12:12:12	33,3	34,0	55,9	37,7		72,3	57,0	95,7	77,3
		14:14:14:14:14	42,9	38,8	65,9	46,9		84,6	64,8	98,4	87,4
		16:16:16:16:16	51,9	43,6	74,3	55,4		92,0	70,9	99,5	93,5
		18:18:18:18:18	60,5	48,2	80,6	63,5		96,2	76,6	99,9	96,9
		20:20:20:20:20	68,4	52,5	86,2	71,0		98,2	81,4	99,9	98,6
		22:22:22:22:22	74,6	56,9	90,3	76,8		99,3	85,1	100,0	99,5
		24:24:24:24:25	80,8	61,2	92,8	82,4		99,7	89,2	100,0	99,8
		26:26:26:26:26	85,2	64,9	95,2	86,6		99,9	91,1	100,0	99,9
		28:28:28:28:28	89,2	68,5	96,8	90,1		100,0	93,4	100,0	100,0
30:30:30:30:30	91,7	72,1	97,7	92,4	0	95,2	100,0	100,0			
						100,0					
eta (10,10)	1:2:3:4:5	4:4:4:4:4	12,4	12,2	10,0	15,5	1:3:6:9:12	19,0	20,1	22,2	26,1
		6:6:6:6:6	12,6	16,3	18,6	16,9		25,8	29,6	50,8	33,4
		8:8:8:8:8:8	17,9	21,9	29,8	22,4		40,9	38,3	75,1	49,0
		10:10:10:10:10	25,5	26,6	42,3	30,2		59,6	47,5	89,8	65,5
		12:12:12:12:12	35,1	31,8	53,7	39,6		75,4	55,6	96,1	79,3
		14:14:14:14:14	44,5	36,4	64,9	48,7		86,7	63,5	98,7	89,2
		16:16:16:16:16	55,1	41,7	74,1	58,7		93,5	70,3	99,6	94,8
		18:18:18:18:18	63,4	46,3	80,7	66,6		96,9	76,0	99,9	97,5
		20:20:20:20:20	71,4	51,2	86,5	73,9		98,6	81,2	100,0	98,9
		22:22:22:22:22	77,7	54,9	90,3	79,8		99,5	85,6	100,0	99,6
		24:24:24:24:25	83,9	59,4	93,9	85,4		99,8	89,2	100,0	99,9
		26:26:26:26:26	87,2	64,0	95,6	88,5		99,9	91,4	100,0	99,9
		28:28:28:28:28	91,0	67,0	97,2	91,8		100,0	93,8	100,0	100,0
30:30:30:30:30	93,4	71,0	98,2	94,0	0	95,3	100,0	100,0			
						100,0					
T (10)	1:2:3:4:5	4:4:4:4:4	11,5	17,1	13,5	15,1	1:3:6:9:12	17,2	25,7	27,3	24,1
		6:6:6:6:6	11,1	24,1	26,0	15,3		22,0	35,3	55,6	29,5
		8:8:8:8:8:8	15,3	29,7	38,3	19,8		34,5	44,5	76,2	41,9
		10:10:10:10:10	21,8	34,7	50,3	26,0		50,6	53,2	88,4	57,5
		12:12:12:12:12	29,0	39,7	59,9	33,3		65,1	59,7	94,3	70,6
		14:14:14:14:14	37,0	44,0	68,2	41,3		77,8	66,8	97,2	81,3
		16:16:16:16:16	45,1	48,8	75,6	48,7		86,7	72,5	98,8	89,0
		18:18:18:18:18	52,9	52,8	80,8	56,1		92,5	77,1	99,4	93,9
		20:20:20:20:20	61,1	56,8	85,6	63,8		96,2	81,8	99,8	96,9
		22:22:22:22:22	67,8	60,5	88,9	70,2		98,0	84,9	99,9	98,3
		24:24:24:24:25	73,2	64,0	91,4	75,2		99,1	88,2	100,0	99,3
		26:26:26:26:26	78,5	67,2	93,9	80,2		99,6	90,5	100,0	99,6
		28:28:28:28:28	83,0	70,3	95,4	84,4		99,7	92,5	100,0	99,8
30:30:30:30:30	86,1	72,8	96,4	87,1	99,9	94,2	100,0	99,9			

5. SONUÇ VE ÖNERİLER

Elde edilen ampirik sonuçlara göre;

- 1) Ele alınan testlerden özellikle Cochran ve Fmax testlerinin dağılımların diklik katsayısından (yani normallikten sapmalardan) oldukça etkilendikleri;
- 2)100.000 simülasyon denemesi sonucunda ampirik olarak gerçekleşen 1.Tip hata olasılıkları bakımından normal dağılışlarda en güvenilir sonuçları Cochran ve Fmax testleri gerçekleştirdiğinden bu testlerden herhangi birinin önerilebileceği,
- 3) Populasyon dağılımları t(10) olması durumunda en sapmalı sonuçları Cochran ve Fmax testinde olduğu,
- 4) Çalışılan bütün testlerin (nispeten dağılımlar t (10) iken Cochran ve Fmax testleri hariç) gerçekleşen 1.Tip hata olasılıkları bakımından genel olarak grup sayısından pek etkilenmedikleri,
- 5) Varyans oranları (etki büyüklüğü) arttıkça bütün testlerin giderek daha yüksek güç değeri gerçekleştirme eğilimine girdikleri, dolayısıyla daha az gözlemle çalışılması durumunda bile yeterli güç değerine ulaşabildikleri,
- 6) Varyans oranları arttıkça çalışılan testlerin birbirlerine oldukça yakın güç değerleri gerçekleştirdikleri,
- 7) 1.Tip hata ve testin gücü birlikte değerlendirildiğinde en güvenilir sonuçları dağılımların normal olması kaydıyla Fmax testinin sağladığı sonuçlarına varmak mümkündür.

KAYNAKLAR

- Alexander, R.A., and Govern, D.M. (1994). A New and Simpler Approximation for ANOVA Under Variance Heterogeneity. *Journal of Educational Statistics*, 19, 91-101.
- VISUAL NUMERICS, (1994). FORTRAN subroutines for Mathematical Applications. IMSL MATH/LIBRARY. Vol.1-2. *Visual Numerics*, Inc., Houston, USA.
- Bishop, T.A, and Dudewicz, E.J.(1978). Exact analysis of variance with unequal variances: Test procedures and tables. *Technometrics*, 20, 419-430.
- Brown, M.B., and Forsythe, A.B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, 69, 364-367.
- Çamdeviren, H., ve Mendes, M. (2005). A Simulation Study for Type III Error Rates of Some Variance Homogeneity Tests. *Pakistan Journal of Statistics*, 21 (2), 223-234.
- Cochran, W.G. and Cox, G.M. (1957). *Experimental design*. John Willey and Sons Inc. New York, USA. 455.
- Cohen, J.(1988).Statistical Power Analysis for the Behavioral Sciences.Second Ed.,*Lawrence Erlbaum Associates*, Hillsdale, 567, New Jersey, USA.
- Conover, W.J. , Johnson, M.E. and Johnson, M.M. (1981). A Comparative study of tests for homogeneity of variance, with application to the outhor continental shelf bidding data. *Technometrics*, 23(4), 351-361.

- Gardside, P.S. (1972). A study of methods for comparing several variances. *JASA*, 67 (338); 342-346.
- Johnson, J.D.(1991). *Applied Multivariate Data Analysis*. Springer-Verlag New York, Inc.USA, 621.
- Keskin, S. (2002). Varyansların Homojenliğini Test Etmede Kullanılan Bazı Yöntemlerin 1.Tip Hata Ve Testin Gücü Bakımından İrdelenmesi. (Yayınlanmamış Doktora Tezi), Ankara Üniversitesi Fen ilimleri Enstitüsü.
- Lim, T.S. and Loh, W.Y. (1996). A Comparison of tests of equality of variances. *Computational Statistics & Data Analysis*. Vol. 22, 287-301.
- Loh, W.Y. (1987). Some modifications of Levene's test of variance homogeneity. *Journal of Statistical Computation and Simulation*, 28, 213-226.
- Mendeş, M.(2002). Normal Dağılım ve Varyansların Homojenliği Ön Şartlarının Gerçekleşmediği Durumlarda Varyans Analizi Tekniğinin Yerine Kullanılabilecek Bazı Parametrik Testlerin 1.Tip Hata ve Testin Gücü Bakımından İrdelenmesi. (Yayınlanmamış Doktora Tezi), Ankara Üniversitesi Fen Bilimleri Enstitüsü.
- Mendeş, M. (2003). Levene, Bartlett, Neyman-Pearson ve Bartlett 2 testlerinin 1.tip hata olasılıkları bakımından karşılaştırılması, *A.Ü.Z.F. Tarım Bilimleri Dergisi*, 9(2), 143-146.
- Murphy, K.R., and Myers, B.(1998) *Statistical Power Analysis. A simple and General Model for Traditional and Modern Hypothesis Tests*. Lawrence Erlbaum Assoc., Publishers,120, London, UK.
- Ott, L. (1998). *An Introduction to Statistical Methods and Data Analysis*. Third Edition. PWS-Kent Publishing Company. USA, 835 .
- Piepho., H.,P.(1996). A Monte Carlo Test for Variance Homogeneity in Linear Models. *Biometrical Journal*. Vol. 38, 461-473.
- Sharma, S.C. (1991). *A new Jackknife test for homogeneity of variances*. Commun.Statist.-Simula, 20 (263); 479-495.
- Sokal, R.R., and Rohlf, F.J, (1995). *Biometry.The Prenciples and Practice of Statistics in Biological Research*. Third Ed.W.H.Freeman and Co. New York, 887 .
- Steel, R.G.D., Torrie, J.H. and Dickey, D.A. (1997). *Principles and procedures of statistics. A biometrical asroach*. McGraw-Hill, 666, New York. USA.
- Wilcox, R.R (1989) Adjusting for Unequal Variances When Comparing Means in One-Way and Two-Way Effects ANOVA models. *Journal of Educational Statistics*. 14, .69-278.
- Zar, J.H. (1999). *Biostatistical Analysis*. Fourth Ed.,Prentice-Hall, Inc. USA, 683.