

Üstel Dağılımlı Populasyonlardan Alınan Örneklerde Tek Yönlü Varyans Analizi Yöntemi ile Bazı Yaklaşım Testlerinin Güç Değeri Bakımından Karşılaştırılması

Sıddık KESKİN¹

Mehmet MENDEŞ²

Geliş Tarihi: 11.01.2002

Özet: Bu simülasyon çalışmasında; üstel (1.00) dağılımlı populasyondan alınmış örneklerde, varyanslar homojen olduğu ve olmadığı durumda, Varyans analizi yöntemi ve bazı yaklaşım testlerinin (Marascuilo, James'in ikinci -derece ve Alexander - Govern testleri) ampirik olarak gerçekleşen güç değerleri bakımından karşılaştırması yapılmıştır. Çalışmada grup sayısı 4 ve 5 olarak belirlenmiş ve gruplardaki gözlem sayıları ise her grupta eşit ve farklı olmak üzere 3 ile 100 arasında alınmıştır. Gruplardaki varyanslar ise eşitten başlayarak 17 kata kadar artırılmıştır. Grup ortalamaları arasındaki farklar (δ) 0.5, 1.0, 1.5 ve 2 standart sapma olarak alınmış ve her bir kombinasyon için ele alınan testlerin güç değerleri, 100,000 simülasyon denemesi sonunda elde edilmiştir. Sonuçta; varyanslar homojen olduğu durumda testlerin güç değerleri arasında belirgin bir farklılık olmadığı, ancak varyanslar heterojen olduğu durumda ise testlerin güç değerlerinin örnek genişliği, gruplardaki gözlem sayısının eşit olup olmaması ve grup ortalamaları arasındaki farka göre değiştiği görülmüştür.

Anahtar Kelimeler: üstel dağılım, varyans analizi yöntemi, Yaklaşım testleri, simülasyon, testin gücü

Comparison the Power of the Test of One - Way ANOVA Method and Some Approximation Tests for the Samples Drawn from the Exponential Distributed Population

Abstract: The present simulation study has been done to compare empirically the power of the ANOVA method, some other approximation tests (Marascuilo, Jame's second -order and Alexander Govern tests) for the samples drawn from exponential (1.00) distribution under the variances homogeneous or heterogeneous. In the study, the number of groups, including observation numbers varied from 3 to 100 observations was determined as 4 and 5. First, it was assumed that variances of the groups were homogen, then the ratio of variance heterogeneity was increased to seventeen times. Differences between group means (δ) were considered as 0.5, 1.0, 1.5 and 2.0 standart deviation. Power of the ANOVA method and the aproximation tests were obtained from 100,000 simulation trials for each combination. As a result, it appears that there is no considerable differences among all tests in terms of the power of the test under variance homogeneity. However, power of the test is affected by sample size, balanced or unbalanced design and differences between group means under variance heterogeneity.

Key Words: exponential distribution, ANOVA method, Approximation tests, simulation, power of test

Giriş

Varyans analizi tekniği başta biyoloji ve davranış bilimlri olmak üzere, bir çok alanda bağımsız iki ve daha fazla grup ortalamaları arasındaki farkın irdelenmesinde en yaygın olarak kullanılan istatistik tekniktir (Edgington 1974, Tomarken ve Serlin 1986). Ancak üzerinde dürulan özellik bakımından ele alınan grup ortalamaları arasındaki farkların irdelenmesinde bu tekniğin kullanılabilmesi için elde edilen verilerde bazı ön şartların sağlanması gerekir (Winer 1991, Sokal ve Rholf 1995). Pratikte çoğu zaman elde edilen verilerde varyans analizinin ön şartlarının sağlanamaması durumuyla karşılaşmaktadır. Bu gibi durumlarda, genellikle ya veriler uygun bir transformasyona tabi tutularak varyans analizi tekniğinin ön şartlarının sağlanması yoluna gidilir, ya da varyans analizinin parametrik olmayan karşılığı olan testler kullanılır. Ancak bir çok durumda, özellikle de varyansların heterojen olması durumunda bu çözüm yollarına gitmek de iyi sonuçlar vermemektedir. Zira, özellikle parametrik

olmayan testler, varyansların heterojen olmasından olumsuz yönde etkilenmekte, verilerin uygun bir transformasyona tabi tutulduktan sonra varyans analizi tekniğine başvurulması ise elde edilecek sonuçların yorumlanmasında bazı sakıncalar ortaya çıkarmaktadır (Oshima ve Algina 1992, Lix ve ark. 1996). Bu gibi durumlarda yukarıda bahsedilen iki çözüm yolu yerine, varyans analizi tekniğinin en önemli iki ön şartı olan normallik ve varyansların homojenliği ön şartlarının yerine gelmediği, ya da bu ön şartlardan sapmalar meydana geldiği durumlarda, varyans analizi tekniğine alternatif olarak geliştirilmiş yaklaşım testlerinin kullanılması önerilmektedir. Bu yaklaşım testleri arasında; Welch, Marascuilo, James'in ikinci-derece, Alexander-Govern, Brown – Forsythe , Wilcox ve Wilcox H_M vb. gibi bazı testler sayılabilir. Ancak bu testlerin gerek I. Tip hata ve gerekse de testin gücü bakımından incelenerek kullanılabilirliğinin tartışılması önemli bir konudur. Zira, bir

¹ Ankara Üniv. Ziraat Fak. Zootekni Bölümü–Ankara

² Ankara Üniv. Çankırı Orman Fak.–Çankırı

testin kullanılmasının önerilebilmesi için o testin sadece başlangıçta kararlaştırılan I. Tip hatayı test sonunda da koruyor olması yeterli olmayıp, aynı zamanda güç (power) değerinin de yüksek olması istenir. Bu amaca yönelik olarak çalışmada, yaklaşım testlerinden sadece Marascuilo, James'in ikinci-derece ve Alexander-Govern testleri ele alınarak, bazı değişik gözlem sayısı kombinasyonlarında ve varyans oranlarında bu testlerin ampirik olarak gerçekleşen güç değerlerinin hem kendi aralarında hem de varyans analizi yöntemi ile karşılaştırması yapılmıştır.

Materyal ve Yöntem

Bu çalışmanın materyalini; IMSL LIBRARY desteği ile FORTRAN programlama dilinde yazılmış simülasyon programı kullanılarak üstel (1.00) (exponential) dağılımdan üretilen tesadüf sayıları oluşturmaktadır (Anonymous 1994). Çalışmada, grup sayısı uygulamada daha sık karşılaşılan 4 ve 5 grup ile sınırlı tutulmuştur. 4 gruplu denemeler için gruplardaki gözlem sayısı: 3:3:3:3, 4:4:4:4, 5:5:5:5, 10:10:10:10, 15:15:15:15, 20:20:20:20, 100:100:100:100, 3:5:7:9, 4:6:8:10, 10:14:18:22, 5:10:15:20, 10:20:30:40; varyans oranları: 1:1:1:1, 1:2:3:4, 1:4:7:10 olarak, 5 gruplu denemeler için gruplardaki gözlem sayısı: 3:3:3:3:3, 4:4:4:4:4, 5:5:5:5:5, 10:10:10:10:10, 15:15:15:15:15, 20:20:20:20:20, 100:100:100:100:100, 3:4:5:6:7, 5:7:9:11:13, 12:15:18:21:24, 5:10:15:20:25, 10:20:30:40:50 ve varyans oranları da: 1:1:1:1:1, 1:2:3:4:5, 1:5:9:13:17 olarak alınmıştır.

Üretilen tesadüf sayıları (X_{ij}); $Y_{ij} = \frac{X_{ij} - \mu_i}{\sigma_i}$; $i = 1, 2, \dots, k$

ve $j = 1, 2, \dots, n_i$ şeklinde standardize edilmiştir. Burada;

X_{ij} : i. populasyondaki j. gözlem değerini göstermekte olup, ortalaması μ_i ve varyansı σ_i^2 dir.

Y_{ij} : i. populasyondaki j. gözlemin standardize edilmiş değeri olup, ortalaması (μ) sıfır ve varyansı (σ^2) biridir.

μ_i : i. populasyonun ortalamasını,

σ_j : i. populasyonun standart sapmasını,

k : populasyon sayısını,

n_i : i. populasyondan alınan örneğin genişliğini göstermektedir.

Ele alınan testler için ampirik olarak gerçekleşen güç değerleri; 100,000 simülasyon denemesi sonunda belirlenmiştir. Güç değerlerinin elde edilmesi için belirlenen populasyonlardan (hepsi üstel dağılım) herhangi birisindeki standardize edilmiş gözlem değerlerine belirli sabit sayılar eklenmiş ve populasyon ortalamaları arasında standart sapma cinsinden farklar (δ) oluşturulmuştur. Standart sapma cinsinden oluşturulan bu farklar; $\delta_i = 0.5, 1, 1.5$ ve 2 olarak belirlenmiştir. Benzer bir yaklaşım da populasyon varyansları için yapılmıştır. Bunun için de belirlenen populasyon kombinasyonlarında, standardize edilmiş gözlem değerleri amaca uygun olarak belirli sabit sayılar ile çarpılmıştır.

Test İstatistikleri:

1. Marascuilo testi:

$$F_M = \sum_{i=1}^k W_i (\bar{X}_i - X'_{..})^2 / (k-1) \quad (1)$$

şeklinde tanımlanan F_M istatistiği, serbestlik dereceleri (k-1) ve $(1/\Lambda)$ olan merkezi bir F dağılımı gösterdiği varsayılmıştır (Marascuilo 1971). Λ_i hata serbestlik derecesinde bir düzeltme yapmada kullanılır ve bunun değeri;

$$\Lambda = \frac{3 \sum_{i=1}^k (1 - w_i / \sum_{i=1}^k w_i)^2 / (n_i - 1)}{(k^2 - 1)}$$

förmülü ile hesaplanır (Marascuilo 1971).

2. James'in ikinci derece testi:

Bu test istatistiğinin değeri;

$$U = \sum_{i=1}^k w_i (\bar{x}_i - x'_{..})^2 \quad (2)$$

förmülü ile bulunur (James 1951). H_0 hipotezinin doğru olduğu varsayımı altında, U istatistiğinin asimtotik olarak (k-1) serbestlik dereceli χ^2 dağılımı gösterdiği varsayılmıştır. Yukarıdaki eşitliklerde;

$$W_i = \frac{n_i}{s_i^2},$$

$$X'_{..} = \frac{\sum_{i=1}^k w_i \bar{x}_i}{W}$$

$$W = \sum_{i=1}^k w_i \text{ şeklinde bulunur.}$$

t ve s herhangi bir pozitif sayılar olmak üzere;

$$R_{st} = \sum_{k=1}^k \frac{1}{(n_i - 1)^s} \left(\frac{W_i}{\sum_{i=1}^k w_i} \right)^t \text{ olduğunda,}$$

$$X_{2s} = \frac{c^5}{[(k-1)(k+1)...(k+2s-3)]}$$

olarak tarif edilmektedir. Buradaki c; (k-1) serbestlik dereceli χ^2 tablo değeridir. U istatistiğinin test edilmesi için gerekli olan $h(\alpha)$ kritik değeri;

$$h(\alpha) = c + (1/2)(3X_4 + X_2) \sum_{i=1}^k (1 - W_i/W)^2 / v_i + [(1/16)(3X_4 + X_2)^2 \times (1 - (i-3)/c) \left[\sum_{i=1}^k (1 - W_i/W)^2 / v_i \right]^2 + (1/2)(3X_4 + X_2) [(8R_{23} - 10R_{22} + 4R_{21} - 6R_{12}^2 + 8R_{12}R_{11} - 4R_{11}^2) + (2R_{23} - 4R_{22} + 2R_{21} - 2R_{12}^2 + 4R_{12}R_{11} - 2R_{11}^2) \times (X_2 - 1) + (1/4)(-R_{12}^2 + 4R_{12}R_{11} - 2R_{12}R_{10} - 4R_{11}^2 + 4R_{11}R_{10} - R_{10}^2) \times (3X_4 - 2X_2 - 1)] + (R_{23} - 3R_{22} + 3R_{21} - R_{20}) \times (5X_6 + 2X_4 + X_2) + (3/16)(R_{12}^2 - 4R_{23} + 6R_{22} - 4R_{21} + R_{20}) (35X_8 + 15X_6 + 9X_4 + 5X_2) + (1/16) \times (-2R_{22} + 4R_{21} - R_{20} + 2R_{12}R_{10} - 4R_{11}R_{10} + R_{10}^2) \times (9X_8 - 3X_6 - 5X_4 - X_2) + (1/4)(-R_{22} + R_{11}^2) \times (27X_8 + 3X_6 + X_4 + X_2) + (1/4)(R_{23} - R_{12}R_{11}) \times (45X_8 + 9X_6 + 7X_4 + 3X_2)]$$

eşitliği ile bulunmuştur (James 1951)

Daha sonra hesaplanan U istatistiği $h(\alpha)$ kritik değeri ile karşılaştırılmış ve $U \geq h(\alpha)$ olduğu durumda ilgili H_0 hipotezi ret edilmiştir Wilcox 1989).

3. Alexander-Govern testi:

k tane bağımsız gruptan her birisi için

$$t_i = \frac{\bar{X}_i - X^+}{S_{\bar{X}_i}} \text{ şeklinde hesaplanan tek örnek t-test}$$

istatistiklerinin hesaplanması ve buradan hareketle Alexander-Govern testi için test istatistiği; hesaplanan t_i değerlerinin doğrudan bir normallik transformasyonuna tabi tutulması ve elde edilen standart normal değişkenlerin karelerinin alınıp toplanması sonucunda elde edilmiştir. Bu t_i değerleri için kullanılan normallik transformasyonu;

$$Z_i = c + \frac{(c^3 + 3c)}{b} - \frac{(4c^7 + 33c^5 + 240c^3 + 855)}{(10b^2 + 8bc^4 + 1000b)} \quad (3)$$

eşitliğine göre yapılmıştır.

(3) no 'lu eşitlikte; $a = v_i - 0.5$, $b = 48 a^2$

$$c = \sqrt{a \times \ln\left(1 + \frac{t_i^2}{v_i}\right)} \quad \text{ve} \quad v_i = n_i - 1 \text{ şeklinde tarif}$$

edilmektedir.

Bu eşitlikten elde edilen Z_i değerleri, birbirlerinden bağımsız standart normal değişkenler haline dönüşmüş olup, bu değişkenlerin karelerinin alınması ve toplanması sonucunda Alexander-Govern (AG) testi için test istatistiği;

$$AG = \sum_{i=1}^k Z_i^2 \text{ eşitliğinden elde edilmiştir. Bu test istatistiği;}$$

yaklaşık olarak (k-1) serbestlik dereceli χ^2 dağılımı göstermektedir. Buradan elde edile AG değerleri, (k-1)

serbestlik dereceli χ^2 dağılımına göre test edilmiştir (Alexander ve Govern 1994).

Bulgular ve Tartışma

Üstel dağılımdan alınan 4 v 5 gruplu örneklerde; değişik örnek genişliği, varyans oranları ve δ değerleri kombinasyonlarında 100.000 simülasyon denemesi sonucunda ele alınan testlere ait ampirik olarak gerçekleşen güç değerleri Çizelge 1, Çizelge 2 ve Çizelge 3 'de verilmiştir.

Gruplardaki varyans oranlarının eşit olduğu diğer bir ifade ile varyansların homojen olduğu durum için elde edilen sonuçların verilmiş olduğu Çizelge 1 incelendiğinde; 4 ve 5 gruplu denemeler için elde edilen sonuçların genel olarak birbirine benzer olduğu görülür. Bunun yanısıra; gruplardaki gözlem sayılarının eşit olup olmamasının da, diğer bir ifade ile grupların dengeli veya dengesiz oluşunun da ele alınan testlerin güç değerleri üzerine belirgin bir etki yapmadığı söylenebilir. Gruplardan birisinin ortalamasının diğer gruplardan 0.5 ve 1 standart sapma farklı olduğu durumda; gruplardaki gözlem sayısı 100 olduğunda bile testlerin güç değerleri %50' nin altındadır. Her hangi bir istatistik test için % 80'in üzerinde olan güç değerinin yeterli olabileceği kabul edilirse, δ 'nın 1.5 ve 2 değerleri için gruplardaki gözlem sayısının ancak 100 olması durumunda yeterli bir güç değeri elde edilebileceği söylenebilir. Çizelge 1'den; ele alınan kombinasyonlar için bu 4 testin güç değerlerinin bir birine yakın olduğu, yani varyanslar homojen olduğu durumda üstel dağılımdan alınan örneklerde Varyans analizi yöntemi ile diğer testlerin güç bakımından bir birine benzer sonuçlar verdiği görülür.

Gruplardaki varyans oranlarının, diğer bir ifade ile grup varyanslarının farklı olduğu durum için ampirik olarak gerçekleşen güç değerlerinin verilmiş olduğu Çizelge 2 ve Çizelge 3 birlikte incelendiğinde; grupların dengeli ve gruplardaki gözlem sayısının 3 ile 15 arasında olması durumunda, gerek 4 grup için gerekse de 5 grup için elde edilen güç değerlerinin çok fazla değişmediği görülür. Ancak gruplardaki gözlem sayısının 20 ve 100 olması durumunda; grup sayısının ve δ 'nın farklı oluşunun, gerçekleşen güç değeri üzerine belirgin bir etkisinin olduğu görülmektedir.

Çizelge 2, 4 gruplu denemeler için incelendiğinde; gruplardan birisinin varyansı diğer grupların varyanslarının 4 katı bile olsa, Varyans analizi yöntemini gerek dengeli gruplarda gerekse de dengesiz gruplarda diğer testlerle yakın bir güç değerine sahip olduğu dikkati çekmektedir. Bunun yanı sıra; grupların dengeli ve gruplardaki gözlem sayısının da 20 ve 100 olması durumunda en yüksek güç değerinin $\delta = 0.5$ olduğu durum için gerçekleştiği görülmektedir. Grup sayısının 5 olması durumunda; grupların dengesiz oluşu, VA yöntemi ile diğer yaklaşım testlerinden elde edilen güç değerlerini belirgin bir şekilde farklılaştırmaktadır. Misal olarak; 25 gözlemin bulunduğu grubun ortalaması diğer grupların ortalamalarından 0.5 standart sapma farklı olduğu durumda, VA yönteminin gerçekleşen güç değeri % 4.8 olurken, AG testinin gerçekleşen güç değeri % 35.8 olmuştur.

Çizelge 1. Dört ve beş gruplu denemelerde varyans oranları sırası ile 1:1:1:1 ve 1:1:1:1:1 olduğu durumda ele alınan testlerin güç değerleri

Exp (1.00)	Grup sayısı (k) = 4				Grup sayısı (k) = 5				
Gözlemler	VA	Maras- cuilo	James	AG	Gözlemler	VA	Maras- cuilo	James	AG
Grup ortalamaları: 0: 0: 0: 0.5 ($\delta = 0.5$)					Grup ortalamaları: 0: 0: 0: 0: 0.5 ($\delta = 0.5$)				
3:3:3:3	4.5	6.9	3.2	4.0	3:3:3:3:3	4.5	10.0	3.4	4.8
4:4:4:4	4.4	6.6	2.8	4.8	4:4:4:4:4	4.3	9.0	2.8	5.6
5:5:5:5	4.5	6.5	2.9	5.3	5:5:5:5:5	4.5	8.9	3.2	6.4
10:10:10:10	5.1	7.8	5.8	7.0	10:10:10:10:10	5.2	9.2	6.3	7.8
15:15:15:15	5.6	7.8	6.5	7.1	15:15:15:15:15	5.5	9.2	7.3	8.1
20:20:20:20	5.8	8.1	7.2	7.5	20:20:20:20:20	6.0	8.9	7.6	8.0
100:100:100:100	12.5	13.3	13.2	13.1	100:100:100:100:100	12.0	13.1	12.8	12.8
3:5:7:9	5.3	7.8	3.3	8.4	3:4:5:6:7	5.0	9.4	3.0	7.6
4:6:8:10	5.0	6.9	3.6	-7.1	5:7:9:11:13	5.2	9.1	5.1	8.6
10:14:18:22	5.7	7.0	5.7	7.0	12:15:18:21:24	5.6	8.0	6.5	7.5
5:10:15:20	5.7	8.0	5.6	10.1	5:10:15:20:25	5.5	9.9	6.9	12.4
10:20:30:40	6.0	7.0	6.0	8.9	10:20:30:40:50	6.0	8.0	6.7	10.2
Grup ortalamaları: 0: 0: 0: 1 ($\delta = 1.0$)					Grup ortalamaları: 0: 0: 0: 0: 1 ($\delta = 1.0$)				
3:3:3:3	5.0	7.9	3.6	4.9	3:3:3:3:3	5.0	11.1	4.0	5.6
4:4:4:4	4.9	7.6	3.2	5.5	4:4:4:4:4	5.3	10.5	3.5	6.7
5:5:5:5	5.5	8.3	4.0	6.7	5:5:5:5:5	5.5	10.6	4.0	7.7
10:10:10:10	7.3	10.8	8.2	9.5	10:10:10:10:10	6.8	12.0	8.5	10.2
15:15:15:15	9.1	12.1	10.5	11.0	15:15:15:15:15	8.6	12.8	10.5	11.2
20:20:20:20	11.1	13.8	12.6	12.8	20:20:20:20:20	9.9	13.8	12.0	12.4
100:100:100:100	41.8	44.2	43.9	43.7	100:100:100:100:100	39.1	42.5	41.8	41.6
3:5:7:9	6.1	7.6	3.4	8.1	3:4:5:6:7	5.8	10.2	3.6	8.0
4:6:8:10	6.3	7.3	3.8	7.3	5:7:9:11:13	6.6	9.6	5.4	8.7
10:14:18:22	8.4	8.9	7.4	8.6	12:15:18:21:24	8.4	10.4	8.5	9.5
5:10:15:20	7.2	7.6	5.1	9.3	5:10:15:20:25	7.0	9.3	6.0	11.2
10:20:30:40	8.9	7.9	6.7	9.9	10:20:30:40:50	8.5	8.5	7.1	10.5
Grup ortalamaları: 0: 0: 0: 1.5 ($\delta = 1.5$)					Grup ortalamaları: 0: 0: 0: 0: 1.5 ($\delta = 1.5$)				
3:3:3:3	5.9	9.1	4.1	5.7	3:3:3:3:3	5.8	13.1	4.7	6.7
4:4:4:4	6.4	9.8	4.5	7.3	4:4:4:4:4	6.3	12.7	4.4	8.3
5:5:5:5	7.3	11.2	5.6	8.7	5:5:5:5:5	7.0	13.4	5.6	9.6
10:10:10:10	11.3	15.9	12.5	13.7	10:10:10:10:10	10.7	17.2	12.6	14.3
15:15:15:15	15.8	20.2	17.9	18.2	15:15:15:15:15	14.5	20.3	17.0	17.6
20:20:20:20	20.2	24.0	22.3	22.2	20:20:20:20:20	18.5	24.2	21.5	21.6
100:100:100:100	78.9	83.5	83.2	83.1	100:100:100:100:100	76.7	82.6	82.2	82.0
3:5:7:9	7.7	8.8	4.2	9.0	3:4:5:6:7	6.8	11.8	4.5	9.1
4:6:8:10	8.3	9.0	5.0	8.6	5:7:9:11:13	8.3	11.6	6.8	10.0
10:14:18:22	13.0	13.6	11.5	13.1	12:15:18:21:24	13.3	15.8	13.2	14.2
5:10:15:20	9.5	9.1	6.0	10.8	5:10:15:20:25	9.1	10.6	6.7	12.2
10:20:30:40	14.1	11.9	10.2	15.1	10:20:30:40:50	13.4	12.2	10.2	15.0
Grup ortalamaları: 0: 0: 0: 2 ($\delta = 2.0$)					Grup ortalamaları: 0: 0: 0: 0: 2 ($\delta = 2.0$)				
3:3:3:3	7.4	11.2	5.3	7.1	3:3:3:3:3	7.0	15.1	5.3	7.7
4:4:4:4	8.6	12.8	6.0	9.3	4:4:4:4:4	8.1	16.3	5.8	10.3
5:5:5:5	10.2	15.1	8.0	11.5	5:5:5:5:5	9.4	17.8	7.8	12.4
10:10:10:10	17.9	23.9	19.7	20.5	10:10:10:10:10	16.7	24.9	19.1	20.4
15:15:15:15	25.7	31.6	28.7	28.5	15:15:15:15:15	24.0	31.8	27.6	27.5
20:20:20:20	33.6	39.7	37.3	36.8	20:20:20:20:20	31.8	39.8	36.4	35.8
100:100:100:100	96.6	98.4	98.3	98.3	100:100:100:100:100	96.5	98.6	98.6	98.5
3:5:7:9	9.5	10.9	5.8	10.8	3:4:5:6:7	8.4	14.4	5.9	10.6
4:6:8:10	10.8	12.0	7.0	11.0	5:7:9:11:13	11.2	15.2	9.3	12.9
10:14:18:22	20.2	21.7	18.7	20.6	12:15:18:21:24	21.5	25.4	21.7	22.5
5:10:15:20	13.1	12.4	8.5	14.6	5:10:15:20:25	12.2	13.7	8.7	15.5
10:20:30:40	21.1	19.8	17.2	24.5	10:20:30:40:50	19.2	19.2	16.1	23.8

Bu farklılığın δ 'nin 1, 1.5 ve 2 olması durumunda azalma eğilimine girdiği dikkat çekmektedir (Çizelge 2). Bu durum, üstel dağılımda gruplardaki gözlem sayısının artması ile birlikte, birbirinden büyük farkı olan gözlem değerlerinin örneğe girmiş olmasından kaynaklanmış olabilir.

Çizelge 3 incelendiğinde; 4 gruplu denemelerde grup varyanslarındaki farklılığın artması ile dengeli gruplarda testlerin güç değerlerinde belirgin bir farklılık

görülmektedir, dengesiz gruplarda Marascuilo ve AG testlerinin güç değerlerinin diğer iki testin güç değerlerinden daha yüksek olduğu görülmüştür. Misal olarak; dengesiz gruplarda 40 adet gözlem içeren grubun ortalaması diğer grupların ortalamasından 0.5 standart sapma farklı olduğunda Varyans analizi yönteminin güç değeri % 16.2 olurken, AG testinin güç değeri % 71.7 olmuştur.

Çizelge 2. Dört ve beş gruplu denemelerde varyans oranları sırası ile 1:1:1:4 ve 1:2:3:4: 5 olduğu durumda ele alınan testlerin güç değerleri

Exp (1.00)	Grup sayısı (k) = 4				Grup sayısı (k) = 5				
Gözlemler	VA	Maras-cuilo	James	AG	Gözlemler	VA	Maras-cuilo	James	AG
Grup ortalamaları : 0: 0: 0: 0.5 ($\delta = 0.5$)					Grup ortalamaları : 0: 0: 0: 0: 0.5 ($\delta = 0.5$)				
3:3:3:3	5.1	7.3	3.3	4.4	3:3:3:3:3	5.6	10.4	3.7	5.0
4:4:4:4	5.1	6.5	2.7	4.6	4:4:4:4:4	5.9	9.3	3.0	5.7
5:5:5:5	5.6	6.7	3.1	5.3	5:5:5:5:5	6.5	9.5	3.4	6.7
10:10:10:10	8.1	9.7	6.9	8.5	10:10:10:10:10	10.0	13.4	8.8	11.5
15:15:15:15	11.3	12.9	10.8	11.8	15:15:15:15:15	14.3	18.8	15.3	17.0
20:20:20:20	14.3	16.3	14.7	15.3	20:20:20:20:20	19.1	24.5	21.5	22.7
100:100:100:100	71.6	76.4	76.0	76.0	100:100:100:100:100	88.6	92.8	92.6	92.7
3:5:7:9	3.0	11.0	4.0	11.6	3:4:5:6:7	3.6	11.9	3.4	9.9
4:6:8:10	3.5	11.2	5.9	11.6	5:7:9:11:13	4.8	18.5	12.1	18.3
10:14:18:22	7.2	18.8	16.7	19.4	12:15:18:21:24	10.7	27.5	24.3	27.5
5:10:15:20	3.6	19.6	15.6	23.2	5:10:15:20:25	4.8	29.6	24.3	35.8
10:20:30:40	7.5	28.3	26.6	34.0	10:20:30:40:50	11.4	42.7	40.0	51.2
Grup ortalamaları: 0: 0: 0: 0: 1 ($\delta = 1.0$)					Grup ortalamaları : 0: 0: 0: 0: 0: 1 ($\delta = 1.0$)				
3:3:3:3	5.1	7.3	3.4	4.5	3:3:3:3:3	5.4	11.0	4.0	5.6
4:4:4:4	5.0	7.0	3.1	5.1	4:4:4:4:4	5.6	9.7	3.3	6.2
5:5:5:5	5.1	6.8	3.2	5.4	5:5:5:5:5	6.2	9.3	3.6	6.7
10:10:10:10	7.2	8.2	5.7	7.2	10:10:10:10:10	8.7	11.2	7.3	9.3
15:15:15:15	9.0	9.7	8.0	8.8	15:15:15:15:15	11.8	14.2	12.3	12.7
20:20:20:20	11.4	11.7	10.3	10.8	20:20:20:20:20	15.7	17.8	15.4	16.3
100:100:100:100	52.2	52.8	52.4	52.4	100:100:100:100:100	75.8	79.6	79.2	79.3
3:5:7:9	2.7	8.9	3.2	9.2	3:4:5:6:7	3.3	10.4	3.4	8.3
4:6:8:10	3.3	9.3	4.7	9.2	5:7:9:11:13	4.3	15.4	9.3	14.9
10:14:18:22	5.7	13.8	12.1	14.1	12:15:18:21:24	9.3	21.3	18.4	21.0
5:10:15:20	3.5	15.4	11.9	18.0	5:10:15:20:25	4.3	24.7	19.6	29.4
10:20:30:40	6.2	21.9	20.2	26.2	10:20:30:40:50	10.0	35.5	32.9	43.1
Grup ortalamaları: 0: 0: 0: 0: 1.5 ($\delta = 1.5$)					Grup ortalamaları: 0: 0: 0: 0: 0: 1.5 ($\delta = 1.5$)				
3:3:3:3	4.9	7.9	3.6	4.8	3:3:3:3:3	5.3	11.5	4.0	5.9
4:4:4:4	4.9	7.5	3.2	5.5	4:4:4:4:4	5.4	10.0	3.5	6.5
5:5:5:5	5.1	7.7	3.7	6.2	5:5:5:5:5	5.9	10.0	4.1	7.3
10:10:10:10	6.4	8.0	6.1	7.2	10:10:10:10:10	8.3	10.7	7.3	9.1
15:15:15:15	8.1	9.1	7.7	8.3	15:15:15:15:15	10.8	12.2	9.8	10.8
20:20:20:20	9.6	9.9	8.8	9.2	20:20:20:20:20	13.6	14.3	12.3	13.0
100:100:100:100	39.0	34.8	34.4	34.3	100:100:100:100:100	64.0	63.2	62.7	62.7
3:5:7:9	2.9	8.1	3.2	8.3	3:4:5:6:7	3.3	10.4	3.4	8.2
4:6:8:10	3.2	8.2	4.3	7.9	5:7:9:11:13	3.9	13.4	8.0	12.6
10:14:18:22	5.2	11.5	9.8	11.4	12:15:18:21:24	8.2	17.2	14.7	16.6
5:10:15:20	3.3	12.8	9.3	14.6	5:10:15:20:25	4.0	20.9	16.0	24.4
10:20:30:40	5.9	17.6	16.1	20.4	10:20:30:40:50	8.9	29.7	27.2	35.5
Grup ortalamaları: 0: 0: 0: 0: 2 ($\delta = 2.0$)					Grup ortalamaları: 0: 0: 0: 0: 0: 2 ($\delta = 2.0$)				
3:3:3:3	5.3	9.0	4.2	5.8	3:3:3:3:3	5.3	12.5	4.4	6.5
4:4:4:4	5.1	9.0	4.1	6.7	4:4:4:4:4	5.6	11.7	4.1	7.7
5:5:5:5	5.4	9.1	4.8	7.3	5:5:5:5:5	5.8	11.5	4.7	8.4
10:10:10:10	6.1	9.5	7.3	8.5	10:10:10:10:10	7.7	11.6	8.2	9.8
15:15:15:15	7.6	9.9	8.6	9.2	15:15:15:15:15	9.9	12.5	10.2	11.2
20:20:20:20	9.0	10.8	9.7	10.0	20:20:20:20:20	12.4	13.8	12.0	12.6
100:100:100:100	33.9	30.7	30.4	30.3	100:100:100:100:100	55.4	51.9	51.4	51.4
3:5:7:9	2.8	8.2	3.5	8.0	3:4:5:6:7	3.2	10.7	3.8	8.3
4:6:8:10	3.3	8.5	4.6	8.0	5:7:9:11:13	3.9	12.9	7.8	11.8
10:14:18:22	5.0	10.8	9.4	10.5	12:15:18:21:24	7.4	15.5	13.3	14.7
5:10:15:20	3.2	11.2	8.0	12.5	5:10:15:20:25	4.0	18.1	13.4	20.4
10:20:30:40	5.3	14.7	13.3	16.5	10:20:30:40:50	8.1	25.4	23.0	29.2

Beş gruplu denemelerde grupların dengesiz olması durumunda Varyans analizi yönteminin güç değeri diğer yaklaşım testlerinden belirgin bir şekilde düşüktür. Gözlem sayısı 25 olan grubun ortalaması, diğer grupların ortalamasından 0.5 standart sapma farklı olduğunda, Varyans analizi yönteminin güç değeri % 8.4 olarak gerçekleşirken, AG testinin güç değeri % 70.1 olarak gerçekleşmiştir.

Bu değerler, gruplardaki gözlem sayısının 10:20:30:40:50 olması durumunda sırası ile %27.4 ve % 93.0 olarak gerçekleşmiştir. Çizelge 3 'te, 5 gruplu denemelerde gruplardan birinin ortalamasının (büyük varyanslı grubun) diğer gruplardan olan farklılığının (Standart sapma cinsinden) 0.5 'ten 2' ye çıkması da testlerin güç değerlerinde önemli bir değişikliğe sebep olmaktadır.

Çizelge 3. Dört ve beş gruplu denemelerde varyans oranları sırası ile 1: 4 : 7 :10 ve 1: 5: 9: 13 : 17 olduğu durumda ele alınan testlerin güç değerleri

Exp (1..00)	Grup sayısı (k) = 4				Grup sayısı (k) = 5				
	Gözlemler	VA	Maras-cuilo	James	AG	Gözlemler	VA	Maras-cuilo	James
Grup ortalamaları: 0: 0: 0: 0.5 ($\delta = 0.5$)					Grup ortalamaları: 0: 0: 0: 0: 0.5 ($\delta = 0.5$)				
3:3:3:3	6.4	8.2	3.8	4.9	3:3:3:3:3	7.5	12.7	4.7	6.2
4:4:4:4	6.9	7.7	3.3	5.4	4:4:4:4:4	8.8	12.2	4.2	7.6
5:5:5:5	8.0	8.8	3.9	6.8	5:5:5:5:5	9.9	14.0	4.8	9.9
10:10:10:10	15.3	21.3	15.7	19.5	10:10:10:10:10	20.3	37.4	27.0	34.2
15:15:15:15	24.7	37.3	33.0	35.7	15:15:15:15:15	34.1	62.1	55.9	60.0
20:20:20:20	34.8	51.9	48.9	50.6	20:20:20:20:20	49.6	79.2	76.0	77.9
100:100:100:100	99.8	99.9	99.9	99.9	100:100:100:100:100	100.0	100.0	100.0	100.0
3:5:7:9	3.3	16.7	5.9	17.1	3:4:5:6:7	4.4	18.1	5.0	15.2
4:6:8:10	4.3	18.9	10.0	18.7	5:7:9:11:13	7.3	4.0	28.8	40.0
10:14:18:22	14.0	45.8	42.4	47.0	12:15:18:21:24	26.0	74.4	71.0	74.8
5:10:15:20	5.6	38.9	33.1	43.7	5:10:15:20:25	8.4	62.8	56.9	70.1
10:20:30:40	16.2	64.6	62.7	71.7	10:20:30:40:50	27.4	88.6	87.3	93.0
Grup ortalamaları : 0: 0: 0: 1 ($\delta = 1.0$)					Grup ortalamaları : 0: 0: 0: 0: 1 ($\delta = 1.0$)				
3:3:3:3	6.3	8.1	3.8	4.9	3:3:3:3:3	7.2	12.4	4.6	6.0
4:4:4:4	6.6	7.5	3.3	5.4	4:4:4:4:4	8.2	11.3	3.9	6.9
5:5:5:5	7.6	8.0	3.7	6.1	5:5:5:5:5	9.3	12.4	4.4	8.5
10:10:10:10	13.1	16.2	11.6	14.5	10:10:10:10:10	18.1	30.2	20.8	27.0
15:15:15:15	20.7	28.1	24.3	26.5	15:15:15:15:15	29.9	52.5	46.0	50.0
20:20:20:20	29.0	40.8	37.7	39.4	20:20:20:20:20	43.5	70.7	66.9	69.1
100:100:100:100	99.0	99.8	99.8	99.8	100:100:100:100:100	99.9	100.0	100.0	100.0
3:5:7:9	3.0	13.2	4.2	13.6	3:4:5:6:7	4.2	15.7	4.3	12.7
4:6:8:10	4.1	15.2	7.6	14.9	5:7:9:11:13	6.8	34.2	23.2	33.8
10:14:18:22	11.9	37.5	34.0	38.2	12:15:18:21:24	23.1	67.5	63.6	67.7
5:10:15:20	5.1	32.8	27.5	37.3	5:10:15:20:25	7.5	57.5	51.5	64.9
10:20:30:40	13.9	56.6	54.4	63.6	10:20:30:40:50	24.5	84.6	83.0	90.2
Grup ortalamaları: 0: 0: 0: 1.5 ($\delta = 1.5$)					Grup ortalamaları: 0: 0: 0: 0: 1.5 ($\delta = 1.5$)				
3:3:3:3	6.1	8.2	3.8	5.1	3:3:3:3:3	7.0	12.5	4.7	6.3
4:4:4:4	6.2	7.6	3.4	5.5	4:4:4:4:4	7.7	11.0	3.8	6.8
5:5:5:5	6.7	7.5	3.6	5.8	5:5:5:5:5	9.0	11.8	4.5	8.1
10:10:10:10	11.4	12.5	8.8	10.9	10:10:10:10:10	16.8	24.6	16.2	21.4
15:15:15:15	17.4	20.7	17.4	19.3	15:15:15:15:15	26.5	43.0	36.6	40.5
20:20:20:20	24.6	30.6	27.7	29.1	20:20:20:20:20	37.9	60.7	56.4	58.9
100:100:100:100	96.4	98.8	98.8	98.8	100:100:100:100:100	99.9	100.0	100.0	100.0
3:5:7:9	3.0	11.1	3.7	11.0	3:4:5:6:7	4.0	13.8	3.8	11.0
4:6:8:10	3.9	12.7	6.2	12.2	5:7:9:11:13	6.5	29.4	19.1	28.8
10:14:18:22	10.4	29.8	26.7	30.3	12:15:18:21:24	20.3	59.7	55.3	59.8
5:10:15:20	4.7	27.3	21.9	30.8	5:10:15:20:25	7.2	52.4	45.9	59.3
10:20:30:40	11.9	48.7	46.6	55.1	10:20:30:40:50	22.0	80.6	78.8	86.9
Grup ortalamaları: 0: 0: 0: 2 ($\delta = 2.0$)					Grup ortalamaları: 0: 0: 0: 0: 2 ($\delta = 2.0$)				
3:3:3:3	5.9	8.4	3.9	5.2	3:3:3:3:3	6.9	12.5	4.6	6.2
4:4:4:4	6.2	8.0	3.7	6.0	4:4:4:4:4	7.3	10.8	3.9	6.7
5:5:5:5	6.6	7.7	3.9	6.0	5:5:5:5:5	8.4	11.0	4.3	7.6
10:10:10:10	10.4	10.4	7.2	8.9	10:10:10:10:10	15.3	20.0	12.9	17.0
15:15:15:15	15.2	15.9	13.1	14.5	15:15:15:15:15	24.1	34.8	28.6	32.3
20:20:20:20	20.7	22.5	20.1	21.1	20:20:20:20:20	33.8	50.8	46.2	48.7
100:100:100:100	90.5	95.2	95.1	95.2	100:100:100:100:100	99.8	99.9	99.9	99.9
3:5:7:9	3.0	9.9	3.5	9.7	3:4:5:6:7	3.9	13.0	3.9	10.0
4:6:8:10	3.7	11.0	5.3	10.4	5:7:9:11:13	6.0	25.3	15.7	24.3
10:14:18:22	9.3	23.7	20.8	23.8	12:15:18:21:24	18.4	51.4	46.7	51.2
5:10:15:20	4.4	22.8	17.5	25.5	5:10:15:20:25	6.7	47.0	40.3	53.3
10:20:30:40	10.7	41.0	38.9	46.4	10:20:30:40:50	20.0	76.2	74.0	83.0

Sonuç

Üstel dağılımdan alınan 4 ve 5 gruplu örneklerde varyanslar eşit olduğu durumda, gruplar dengeli olsun veya olmasın varyans analizi yönteminin güç değeri, diğer yaklaşım testlerinin güç değerlerinden önemli ölçüde farklı olmamaktadır.

Marascuilo testinin örneklerde küçük gözlem kombinasyonlarının bulunması durumunda, grup sayısının farklı oluşundan olumsuz yönde etkilendiği ve bu olumsuz etkinin ise ortalamalar arasındaki farkın artmasına paralel olarak daha da belirginleştiği görülmüştür.

Varyans analizi yönteminin, özellikle populasyon varyansları ile bu populasyonlardan alınan örneklerdeki gözlem sayılarının doğru eşleştirilmesi durumunda, diğer testlere göre çok daha düşük güç değeri gerçekleştirdiği görülmüştür. Özellikle, populasyon ortalamaları arasındaki farkın 0 ve 1 olması durumunda, bu düşük güç değerinin daha da belirginleştiği dikkat çekmektedir. Dolayısıyla, bu gibi durumlarda alternatif testlerin VA yöntemine tercih edilmesi gerektiği söylenebilir. Varyanslar homojen iken, bütün testlerin ele alınan bütün gözlem kombinasyonlarında birbirlerine oldukça yakın güç değeri verdikleri ve bu koşullarda güç bakımından bu testlerin her birisinin diğeri yerine kullanılabileceği söylenebilir.

Buna karşın, gruplardaki varyans oranları farklı olduğu zaman, özellikle 5 gruplu denemelerde grupların dengesizliği, ele alınan testlerin güç değerlerini etkilemektedir. Beş gruplu denemelerde, gruplardaki gözlem sayısının 15'in üzerine çıkması durumunda, yaklaşım testlerinden Marascuilo ve AG testlerinin güç değerlerinin VA yöntemine göre daha yüksek olduğu söylenebilir. Ancak, bilindiği üzere bir testin kullanılmasının önerilebilmesi için o testin aynı zamanda başlangıçta kararlaştırılan I. Tip hatayı da koruyor olması gerekir. Bir test için sadece güç değerine veya sadece I. Tip hata değerine göre öneride bulunmak yanlış yorumlamalara sebep olabilir. Çalışmada belirlenen şartlar için güç değerleri bakımından her ne kadar Marascuilo ve AG testlerinin daha iyi sonuçlar verdiği düşünülür ise de, bu testlerin I. Tip hata bakımından karşılaştırması yapılmadan kullanılıp kullanılmayacağı konusunda öneride bulunmak sakıncalı olacaktır. Dolayısıyla, bu testlerin I. Tip hata bakımından karşılaştırılması başka bir çalışmaya bırakılmıştır.

Kaynaklar

- Alexander, R. A. and D. M. Govern, 1994. A new and simpler approximation for Anova under variance heterogeneity. *J. of Edu. Stat.*, 19 (2) 91-101.
- Anonymous, 1994. IMSL MATH / LIBRARY FORTRAN subroutines for mathematical applications. Vol: 1, 2, Visual Numerics Inc. USA.
- Edgington, E. S. 1974. A new tabulated of statistical procedures used in APA journals. *Am. Psychologist*, 29, 25-26.
- James, G. S. 1951. The comparison of several groups of observation when the ratios of the population variances are known. *Biometrika*, 38, 324-329.
- Lix, L. M, J. C. Keselman and H. J. Keselman, 1996. Consequences of assumption violations revised: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational. Research*, 66 (4) 579-619
- Marascuilo, L. A. 1971. *Statistical Methods for Beevioral Science Research*. McGraw-Hill 578p. New York.
- Oshima, T. C. and J. Algina, 1992. Type I error for James's second-order test and Wilcox' H_M test under heteroscedasticity and non-normality. *British J. of Mathematichal and Statistical Psychology*, 45, 255-263.

Sokal, R. R. and F. J. Rohlf, 1995. *Bimetry*. W. H. Freeman and Company. 887p., New York.

Tomarken A. J. and R. C. Serlin, 1986. Comparison of anova alternatives under variance heterogeneity and specific noncentrality structures. *Psycholoical Bulletin* 99 (1) 90-99.

Wilcox, R. R. 1989. Adjusting for unequal variances when the comparing mean in one- way and two-way fixed effect anova models. *J. of Educational Statistics*, 14 (3) 269-278.

Winer, B. J. 1991. *Statistical Principles in Experimental Design*. McGraw-Hill Company. 907p. New York.