

ANOVA F ve K Testlerinin III. Tip Hata Olasılıkları Bakımından Karşılaştırılması

Mehmet MENDEŞ¹

Geliş Tarihi: 02.12.2002

Özet: Çift taraflı hipotez kontrollerinde, I. Tip ve II. Tip hata olasılıkları dışında, III. Tip hata (γ) olarak adlandırılan başka bir hata olasılığı da söz konusudur. III. Tip hata olasılığı, birbirleriyle karşılaştırılacak grup ortalamaları arasındaki farkın yönünün dikkate alınmamasından kaynaklanır. III. Tip hata olasılığı özellikle küçük hacimli örneklerle çalışılması durumunda testin gücünü etkilemektedir. Bu çalışmada, Anova F ve K testlerinin III. Tip hata olasılıkları bakımından karşılaştırılması amacıyla bir simülasyon çalışması yapılmıştır. Yapılan 100000 simülasyon denemesi sonunda her iki test bakımından gerçekleşen III. Tip hata olasılıklarının örnek genişliği ve ortalamalar arası farkın artmasına paralel olarak giderek küçüldüğü ve bunun Anova F testinde daha belirgin olduğu görülmüştür. III. Tip hata olasılığının grup sayısı ve varyansların heterojenlik derecesinden olumsuz etkilendiği ve bunun K testinde çok daha belirgin olduğu görülmüştür. Varyans oranları 1:3:5 olan üç χ^2 (3) dağılımlı populasyonun olması durumunda F ve K testi bakımından gerçekleşen III. Tip hata olasılıklarının genel olarak %0.54-1.57 ve %0.92-4.37 arasında değiştiği görülmüştür. Varyans oranlarının 1:2:4:6 olması durumunda III. Tip hata olasılıkları %1.63-5.01 ve %2.15-14.6 olarak bulunmuştur. Aynı koşullarda dağılımların üstel ($\exp(0.5)$) olması özellikle K-testi bakımından gerçekleşen III. Tip hata olasılıklarının belirgin bir şekilde artmasına neden olmuştur.

Anahtar Kelimeler: III. Tip hata, güven aralığı, testin gücü, varyansların homojenliği, normal olmayan dağılım

Comparison of ANOVA F and K Tests in Terms of Type III Error Rates

Abstract: There exist another type of error, called Type III error, in addition to Type I and Type II error in two-tailed hypothesis testing. Type III error (γ) is the probability of getting the direction wrong. The effect of Type III error on test power is more pronounced especially when sample sizes are small. In this study a simulation study was conducted to compare Anova F and K tests with respect to Type III error rates. At the end of 100000 simulation trial it was seen that the probability of a rejection in the wrong direction (Type III error) decreased as sample size and population mean differences increased. This was more obvious for ANOVA F test. Also it was seen that probability of the Type III error rate increased as number of groups and heterogeneity of variance increased. It is more pronounced with respect to K test. When samples were taken from three Chi-square with 3 df populations under variance heterogeneity (1:3:5), the probability of Type III error rate vary 0.54-1.54% for F test and 0.92-4.37% for K test. When number of group increased to four and variance ratio was 1:2:4:6, the probability of Type III error rate vary 1.63-5.01% for F test and 2.15-14.6% for K test. When samples were taken from three exponential (0.50) populations under the same experimental conditions, Type III error rate increased especially in terms of K-statistics.

Key Words: Type III error, confidence interval, power of test, homogeneity of variance, nonnormal distribution

Giriş

İstatistik analizlerinin temel amacı, üzerinde durulan özellik bakımından örnek değerlerinden yararlanılarak populasyon parametreleri hakkında bilgi edinmektir (Zar 1999). Uygulamada bu bilgi, genellikle populasyon ortalamaları ile ilgilidir. Üzerinde durulan özellik bakımından iki ve daha fazla bağımsız grup ortalaması arasındaki farkın irdelenmesinde en yaygın kullanılan istatistik teknik, varyans analizi tekniğidir (ANOVA F). Ancak, bu teknikten beklenen yararların sağlanabilmesi için özellikle karşılaştırılacak muamele gruplarının alınmış oldukları varsayılan populasyonların varyanslarının homojen ve dağılımlarının normal olması gerekir (Welch 1951, Brown-Forsythe 1974, Edgington 1974, Tomärken ve Serlin 1986, Wilcox 1989, Winer ve ark. 1991, Sokal ve Rohlf 1995, Ott 1998, Mendes 2002; Mendes ve Başpınar 2003). Bu iki varsayımın sağlanamadığı durumlar için ANOVA F testi yerine kullanılabilecek değişik yaklaşım testleri geliştirilmiştir. Bunlardan birisi de K istatistiğidir

(Krutchkoff 1988). Birbirinden bağımsız k adet ortalama arasındaki farkların irdelenebilmesi için;

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

H_1 : En az iki ortalama arasındaki fark önemlidir, hipotez takımının test edilmesi gerekir. Çift taraflı hipotez kontrollerinde H_0 hipotezinin reddedilmesi, grup ortalamalarından en az ikisinin üzerinde durulan özelliğe etkilerinin farklı olduğu anlamına gelir. Ancak, bu farkların yönü hakkında bir bilgiye $H_0: \mu_1 - \mu_2 = 0$ ulaşılamaz. şeklinde kurulan kontrol hipotezinin reddedilmesi durumunda; $H_1: \mu_1 - \mu_2 < 0$ ya da $H_2: \mu_1 - \mu_2 > 0$ şeklinde iki alternatif hipotezden herhangi birisinin kabul edilmesi söz konusudur. Ancak, uygulamada $H_0: \mu_1 - \mu_2 = 0$ hipotezinin reddedilmesi durumunda sadece üzerinde durulan özellik bakımından bu ki muamele grubunun ortalaması arasındaki farkın istatistik olarak önemli olduğu

¹ Çanakkale Onsekiz Mart Üniv. Ziraat Fak. Zooteknik Bölümü - Çanakkale

söylenir. Ancak, bu iki grup ortalaması arasındaki farkın yönü dikkate alınmamaktadır. Halbuki, söz konusu muamele gruplarından hangisinin diğerinden büyük / küçük olduğu yani farkın yönü testin gücünü etkilemektedir. Bu etki ise özellikle küçük hacimli örneklerle çalışılması durumunda daha da belirginleşmektedir (Leventhal ve Huynh 1996).

Aynı denemenin benzer koşullarda tekrarlanması halinde ret edilen hipotez sayısının nispi frekansı testin gücünü temsil eder (Tiku 1971, Murphy ve Myers 1998). Ancak bu durum her zaman geçerli değildir. Çünkü, bazı durumlarda araştırmacı ele alınan populasyon ortalamaları arasındaki farkın gerçek yönü hakkında bir yanılgıya düşebilir. Böyle hallerde I. Tip ve II. Tip hata olasılıklarından başka bir hata yapma olasılığı (III. Tip hatadan (γ)) ile de karşılaşılır (Kimmel 1957, Kimball 1957, Kaiser 1960, Shaffer 1972, Hopkins 1973, Leventhal ve Huynh 1996, McDonald 1999, Sharon ve Carpenter 1999). III. Tip hata, gerçekte yanlış olan H_0 hipotezinin doğru bir şekilde reddedilmesi ancak, etkilerin yönü hakkında yanlış bir sonuca ulaşılmasıdır. Mesela A ve B populasyon ortalamaları arasında gerçekte $\mu_A - \mu_B = 0.5$ gibi ($\mu_A > \mu_B$ şeklinde) bir fark olsun. Bu durumda araştırmacı iki tip hata ile karşılaşabilir: Bunlar;

1) Gerçekte bu iki populasyon ortalaması arasında fark varken, yapılan test sonunda fark yoktur şeklinde bir sonuca ulaşmak yani II. Tip hata yapma olasılığı (β),

2) Gerçekte bu iki populasyon ortalaması arasında $\mu_A > \mu_B$ şeklinde fark varken, yapılan test sonunda bu iki populasyon ortalaması arasında $\mu_A < \mu_B$ şeklinde bir farkın olduğu sonucuna ulaşmakla yapılan hata yani, III. Tip hatadır (γ) (Kimball 1957, McDonald 1999, Sharon ve Carpenter 1999, Sansgiry ve Akman 2000).

Bu iki hata olasılığı da doğrudan testin gücü üzerinde etkilidir. Çünkü III. Tip hata ile testin gücü arasında,

Güç=1- β - γ şeklinde bir ilişki vardır. Kaiser (1960) ve Leventhal ve Huynh (1996) testin gücünü "reddedilmesi gereken kontrol hipotezinin ortalamalar arasındaki farkın yönünün de dikkate alınarak reddedilme olasılığı" olarak yeniden tanımlamışlardır. Dolayısıyla istatistik testlerinin güç bakımından karşılaştırılmasında ve uygun örnek hacminin belirlenmesinde III. Tip hatanın dikkate alınması gerekir. Testin önemlilik yönünün bilinmesi, ampirik olarak gerçekleşen testin gücünden bunun payının giderilmesine ve dolayısıyla daha güvenilir sonuçların elde edilmesine imkan sağlar. III. Tip hatanın testin gücüne olan etkisi özellikle küçük hacimli örneklerle çalışılması durumunda ortaya çıkmaktadır. Büyük hacimli örneklerle çalışılması durumunda ise III. Tip hatanın etkisi oldukça düşük seviyelerde kalmaktadır. Uygulamada daha ziyade küçük hacimli örneklerle çalışılmasından dolayı III. Tip hata olasılığının dikkate alınması elde edilecek sonuçların güvenilirliği açısından oldukça önemlidir (Kaiser 1960, McDonald 1999).

Bu çalışmada, varyans analizi tekniği ile bu tekniğin normallik ve varyansların homojenliği ön şartlarının

gerçekleşmediği durumlarda yerine kullanılması önerilen testlerden birisi olan K-istatistiği ile III. Tip hata bakımından karşılaştırılması yapılmıştır.

Materyal ve Yöntem

Çalışmanın materyalini, Microsoft Fortran Developer Studio'nun IMSL kütüphanesinden yararlanılarak üretilen tesadüf sayıları oluşturmaktadır (Anonymous, 1994). Çalışmanın amacı için IMSL'in RNCHI, ve RNEXP fonksiyonları kullanılarak χ^2 (3) ve Üstel (0.50) parametrelerine sahip populasyonlardan uygulamada sık karşılaşılan k=3,4 grup için çeşitli örnek hacimlerine göre tesadüf örnekleri üretilmiştir. Üretilen örneklerin alınmış oldukları populasyonların ortalamaları arasında fark oluşturmak için, bunlardan birincisindeki tesadüf sayılarına standart sapma cinsinden belirli sabit sayılar ($\delta=0.25, 0.50, 0.75, 1.0$) ilave edilmiştir. Populasyonların varyansları arasında heterojenlik oluşturmak için de örneklerdeki tesadüf sayıları sabit sayılarla ($\sigma=1, \sqrt{2}, \sqrt{3}, \sqrt{4}, \sqrt{5}, \sqrt{6}, \sqrt{9}, \sqrt{12}, \sqrt{15}$) çarpılmıştır. Böylece, değişik deneme koşullarında Anova F ve K testlerinin gerçekleşen III. Tip hata olasılıkları bakımından karşılaştırılması sağlanmıştır. Bu örneklerin özellikleri; k=3 için, $n_1:n_2:n_3 = 3:3:3, 5:5:5, 8:8:8, 10:10:10, 12:12:12, 15:15:15, 20:20:20, 25:25:25, 30:30:30$ ve $\sigma_1^2:\sigma_2^2:\sigma_3^2 = 1:1:1, 1:3:5, 1:5:15$ olarak ve k=4 için, $n_1:n_2:n_3:n_4 = 3:3:3:3, 5:5:5:5, 8:8:8:8, 10:10:10:10, 12:12:12:12, 15:15:15:15, 20:20:20:20, 25:25:25:25, 30:30:30:30$ ve $\sigma_1^2:\sigma_2^2:\sigma_3^2:\sigma_4^2 = 1:1:1:1, 1:2:4:6, 1:5:9:12$ olarak belirlenmiştir. Sonra önce bu testlerin ampirik olarak gerçekleşen gücü (1- β), her ortalamalar arasındaki fark (δ)-varyans oranları-gözlem sayıları-grup sayısı kombinasyonuna göre 100000 simülasyon denemesi sonunda ret edilen hipotez sayılarının (r) nispi frekanslarının hesaplanması ($f = r/100000$) ile elde edilmiştir. Daha sonra bundan yararlanarak III. Tip hata olasılıkları, ret edilen hipotezlerin toplamı (r) içerisinde, gerçekte büyük ortalamalı populasyon ortalamasının diğer populasyon ortalamalarından kaç kez (r_1) küçük olduğunun sayılması ve bu sayının nispi frekansa dönüştürülmesi ($\gamma = r_1/100000$) ile elde edilmiştir. Çalışmada, I. Tip hata olasılığı (α), deneme başında %5 olarak kararlaştırılmıştır. Varyansların homojenliği ön şartının sağlanmadığı durumlar için önerilen testlerden biri olan K-testi için test istatistiği aşağıdaki şekilde elde edilmiştir. Her bir grubun ortalaması;

$$\bar{X}_i = \frac{\sum X_{ij}}{n_i} \quad 1$$

şeklinde tarif edilir. Bu şekilde elde edilen grup ortalamalarından yararlanılarak genel ortalama ise ;

$$\bar{X} = \frac{\sum_{i=1}^k \bar{X}_i n_i}{\sum_{i=1}^k n_i} \quad 2$$

şeklinde. Burada; k : grup sayısını,

X_{ij} : i . Gruptaki j . gözlem değerini, n_i : i . gruptaki

gözlem adetini ve $\hat{\sigma}_i^2$: σ_i^2 'nin tahminini göstermektedir.

Bunlardan yararlanılarak bağımsız grup ortalamalarını karşılaştırmak amacıyla kullanılacak K istatistiği;

$$K = \frac{\sum_i \left[(\bar{X}_i - \bar{X})^2 * n_i / \hat{\sigma}_i^2 \right]}{(k-1)} \quad 3$$

şeklinde elde edilir ve bu istatistik yaklaşık olarak $F(k-1, N-k-1)$ dağılıma sahiptir. Eşitlikte, k karşılaştırılacak bağımsız sayısını göstermektedir. K test istatistiğinin

hesaplanmasında her bir $\hat{\sigma}_i^2$ yerine toplanmış varyans

tahmininin kullanılması halinde K testi, F testine benzer olur. Dolayısıyla bu iki test aynı sonucu verir (Krutchkoff 1988, Mendeş 2002). Bu iki testin bir birleriyle karşılaştırılması amacıyla K test istatistiği hesaplanırken

$\hat{\sigma}_i^2$ yerine ortak bir hata varyansı (toplanmış varyans) değilse, her grubun kendi varyansı kullanılmıştır.

III. Tip hata olasılığı reddedilmesi gereken bir hipotezin farkların yönü dikkate alınmadan reddedilme olasılığıdır. Çift taraflı hipotez kontrollerinde kontrol (H_0) ve alternatif (H_1) hipotezleri aşağıdaki gibi tarif edilebilir.

- 1) $H_1 : \mu_1 - \mu_2 < 0$
- 2) $H_2 : \mu_1 - \mu_2 = 0$ ($= H_0 : \mu_1 - \mu_2 = 0$)
- 3) $H_3 : \mu_1 - \mu_2 > 0$

Bilindiği üzere testin gücü H_2 hipotezinin (klasik olarak H_0 hipotezi) reddedilmesi durumunda tanımlanabilir. Dolayısıyla bu durumda, testin gücü H_1 veya H_3 ile ifade edilen alternatif hipotezlerinin kabul edilmesi durumunda tanımlanabilir. Bu durumda, H_1 ve H_3 hipotezlerine bağlı olarak iki farklı III. tip hata olasılığı söz konusudur. Bunlar ise, gerçekte H_1 hipotezi geçerli iken yani $H_1 : \mu_1 - \mu_2 < 0$ hipotezi geçerli iken

$H_3 : \mu_1 - \mu_2 > 0$ şeklinde ifade edilen H_3 hipotezinin

Gerçek Durum

		H_1 doğru	H_2 doğru	H_3 doğru
Verilen karar	H_1 Kabul	Doğru karar $1 - \beta_{21} - \gamma_{31}$ Testin gücü ₁₁	I. Tip hata α_{12}	III. Tip hata γ_{13} Testin gücü ₁₃
	H_2 Kabul	II. Tip hata β_{21}	Doğru karar $1 - \alpha_{12} - \alpha_3$	II. Tip hata β_{23}
	H_3 Kabul	III. Tip hata γ_{31} Testin gücü ₃₁	I. Tip hata α_{32}	Doğru karar $1 - \beta_{23} - \gamma_{13}$ Testin gücü ₃₃

kabul edilmesi durumunda yapılan III. Tip hata ile (γ_{31}),

gerçekte $H_3 : \mu_1 - \mu_2 > 0$ şeklinde ifade edilen H_3 hipotezi

geçerli iken $H_1 : \mu_1 - \mu_2 < 0$ şeklinde ifade edilen H_1

hipotezinin kabul edilmesiyle yapılan III. Tip hatadır (γ_{13}).

H_1 veya H_3 ile ifade edilen alternatif hipotezlerinin kabul edilmesine göre testin gücü 4 farklı biçimde ifade edilebilmektedir (Testin gücü₁₁, Testin gücü₁₃, Testin gücü₃₁, Testin gücü₃₃). Dikkat edileceği üzere testin gücü, III. Tip hata olasılığının etkilenmektedir. Dolayısıyla III. Tip hata olasılığının dikkate alınmadığı çalışmalarda testin gücü gerçek değerinden daima yüksek olacaktır. Ancak, III. Tip hata olasılığının testin gücü üzerine olan etkisi, örnek hacmi ve ortalamalar arasındaki farkın artmasına paralel olarak azalır (Leventhal ve Huynh 1996, McDonald 1999).

Bulgular ve Tartışma

Yapılan 100000'er simülasyon denemeleri sonunda, ortalamalar arasındaki fark (δ)- varyans oranları-gözlem sayıları- grup sayısı kombinasyonlarında göre gerçekleşen III. Tip hata olasılıkları (γ) Çizelge 1 ve Çizelge 2'de topluca verilmiştir. Çizelge 1 incelendiğinde, örneklerin homojen varyanslı (1:1:1) üç serbestlik dereceli Ki-kare dağılımı gösteren populasyonlardan alınmaları durumunda ($\chi^2(3)$), ortalamalar arasındaki fark 0.25 standart sapma iken ($\delta=0.25$) Anova F testi bakımından gerçekleşen III. Tip hata olasılıklarının genel olarak %0.7-1.19 arasında, $\delta=0.50$ iken % 0.35-0.97 arasında, $\delta=0.75$ iken % 0.08-0.85 arasında ve $\delta=1.0$ iken %0.0-0.71 arasında değiştiği görülmektedir. Aynı koşullarda K-istatistiği bakımından gerçekleşen III. Tip hata olasılıklarının ise sırasıyla %1.02-2.56, %0.50-2.34, %0.09-2.14 ve %0.0-1.97 arasında değiştiği görülmektedir. Dikkat edileceği üzere genel olarak K-istatistiği bakımından gerçekleşen III. Tip hata olasılıkları Anova F testi bakımından gerçekleşen III. Tip hata olasılıklarından daha yüksektir. Ancak, ortalamalar arasındaki fark ve örnek hacminin artmasına paralel olarak ele alınan iki testin de giderek birbirine yakın sonuçlar gerçekleştirme eğilimine girdikleri görülmektedir. Aynı koşullarda populasyon varyanslarının 1:3:5 şeklinde orta derecede heterojenleştirilmesi durumunda her iki testinde bu deneme koşulundan olumsuz yönde etkilenmektedir. Bu koşullarda populasyon varyanslarının 1:5:15 şeklinde daha da heterojenleştirilmesi durumunda, gerçekleşen III. Tip hata olasılıklarının çok daha belirgin bir şekilde arttığı görülmektedir. Dikkat edileceği üzere, populasyon varyanslarının heterojenleştirilmesine paralel olarak her iki test bakımından gerçekleşen III. Tip hata olasılıkları da belirgin bir şekilde artmakta ve bu artış K-istatistiğinde çok daha belirgin olmaktadır. Dolayısıyla bu koşullarda Anova F testi bakımından elde edilen sonuçların daha güvenilir olduğu söylenebilir. Zira III. Tip hatanın, testin gücü üzerine olan etkisi ANOVA F testi bakımından daha düşüktür.

(Güç=1- β - γ). Buradan hareketle bu iki testin III. Tip hata olasılığının dikkate alınmadan testin gücü bakımından karşılaştırmak amacıyla yapılan simülasyon çalışmalarında, Anova F testi bakımından gerçekleşen güç

değerlerinin, K-istatistiği bakımından gerçekleşen III. Tip hata olasılıklarına tercih edilebileceği rahatlıkla ileri sürülebilir.

Söz konusu populasyonların dağılımlarının üstel (0.50) olması durumunda elde edilen sonuçların, aynı koşullarda dağılımların $\chi^2(3)$ olması durumunda elde edilen sonuçlara oldukça benzer oldukları söylenebilir. Çünkü, dağılım şeklinin $\chi^2(3)$ ya da üstel (0.50) olması, her iki test bakımında da gerçekleşen III. Tip hata olasılıklarını dikkate değer bir şekilde etkilemediği göze çarpmaktadır. Genel olarak ele alınan bütün deneme koşulları birlikte değerlendirildiğinde, her iki test bakımından da gerçekleşen III. Tip hata olasılıklarının artmasına neden olmaktadır. Söz konusu bu artış özellikle ortalamalar arasındaki farkın 0.25-0.75 ve örnek hacminin küçük olduğu durumlarda daha da belirgin olduğu görülmektedir. Diğer yandan örnek hacmi ve ortalamalar arasındaki fark arttıkça III. Tip hata olasılığı azalmaktadır. Grup sayısının 3'ten 4'e çıkartılması halinde, her iki testinde bundan olumsuz yönde etkilendiği ve söz konusu bu olumsuz etkinin K-istatistiği bakımından çok daha belirgin olduğu görülmektedir. Yani her iki testte grup sayısının artmasından olumsuz yönde etkilenmektedir. Halbuki varyans analizi ile bu teste alternatif olabilecek pek çok testin ampirik olarak gerçekleşen testin güç değerleri

bakımından karşılaştırılması amacıyla yapılan pek çok simülasyon çalışmasında söz konusu testlerin III. Tip hata olasılıklarının dikkate alınmamasından dolayı başta Anova F testi olmak üzere pek çok testin grup sayısına karşı güçlü testler oldukları bildirilmiştir (Welch 1951, Brown-Forsythe 1974, Krutchkoff 1988, Wilcox 1989, Alexander-Govern 1994).

Sonuç

Çalışmadan elde edilen ampirik sonuçlara göre;

1. Testin gücünün hesaplanmasında ve dolayısıyla da uygun örnek genişliğinin belirlenmesinde, mutlaka III. Tip hata olasılığı göz önüne alınmalıdır.
2. Ortalamalar arasındaki fark ve örnek hacimleri arttıkça, III. Tip hata olasılığı azalmaktadır.
3. Varyansların heterojenlik dereceleri arttıkça, III. Tip hata olasılığı da artmaktadır.
4. Grup sayısının artmasına paralel olarak, III. Tip hata olasılığı da artmaktadır.
5. Ortalamalar arası farka bağlı olarak büyük hacimli örneklerle çalışılması durumunda, III. Tip hatanın göz ardı edilebileceği sonuçlarına varmak mümkündür.

Çizelge 1. Ortalamalar arasındaki fark (δ)-varyans oranları-gözlem sayıları-grup sayısı kombinasyonlarında göre gerçekleşen ampirik III. Tip hata olasılığı (γ) (%), $k=3$

Varyans oranları		$\chi^2(3)$						Exp (0.50)					
Dağılımlar		1:1:1		1:3:5		1:5:15		1:1:1		1:3:5		1:5:15	
δ_i	$n_1:n_2:n_3$	F	K	F	K	F	K	F	K	F	K	F	K
0.25	3:3:3	1.19	2.56	1.57	4.37	2.04	5.66	1.20	2.85	1.51	4.79	1.73	6.21
	5:5:5	0.95	1.90	1.48	3.78	1.87	5.35	1.04	2.11	1.42	4.13	1.36	6.00
	8:8:8	0.96	1.64	1.45	3.26	1.80	5.10	1.01	1.81	1.38	4.04	1.34	5.57
	10:10:10	0.87	1.47	1.40	3.04	1.68	5.08	0.99	1.77	1.35	3.89	1.27	5.38
	15:15:15	0.84	1.24	1.33	2.57	1.56	4.90	0.96	1.58	1.29	3.08	1.23	5.13
	20:20:20	0.78	1.07	1.24	2.60	1.45	4.53	0.85	1.27	1.24	3.01	1.17	4.81
	25:25:25	0.73	1.05	1.15	2.41	1.40	4.31	0.79	1.25	1.18	2.70	1.11	4.59
30:30:30	0.70	1.02	1.10	2.32	1.33	4.18	0.70	1.09	1.11	2.63	1.04	4.33	
0.50	3:3:3	0.97	2.34	1.46	4.25	1.82	5.49	1.04	2.87	1.31	4.55	1.60	5.77
	5:5:5	0.87	1.87	1.25	3.67	1.47	4.72	0.77	2.07	1.29	3.97	1.39	5.44
	8:8:8	0.76	1.56	1.20	3.12	1.35	4.61	0.70	1.59	1.26	3.60	1.27	5.19
	10:10:10	0.69	1.35	1.25	2.88	1.27	4.52	0.63	1.55	1.20	3.34	1.16	4.97
	15:15:15	0.62	1.08	1.22	2.57	1.20	4.40	0.61	1.18	1.12	3.07	1.11	4.92
	20:20:20	0.52	0.78	1.18	2.24	1.15	4.06	0.51	0.85	1.04	2.84	1.07	4.65
	25:25:25	0.39	0.57	1.16	2.08	1.11	3.92	0.46	0.74	1.00	2.48	1.00	4.32
30:30:30	0.35	0.50	1.06	1.71	1.09	3.36	0.34	0.53	0.93	2.27	0.93	4.03	
0.75	3:3:3	0.85	2.14	1.38	4.08	1.80	5.10	0.73	2.57	1.23	4.24	1.48	5.65
	5:5:5	0.72	1.79	1.16	3.31	1.37	4.16	0.60	1.96	1.14	3.81	1.32	4.87
	8:8:8	0.55	1.13	1.11	2.95	1.33	3.99	0.57	1.55	1.07	3.47	1.19	4.57
	10:10:10	0.48	0.94	1.08	2.76	1.13	3.92	0.51	1.07	1.04	3.08	1.15	4.35
	15:15:15	0.29	0.48	1.06	2.34	1.10	3.80	0.32	0.59	1.00	2.87	1.06	4.26
	20:20:20	0.16	0.24	1.03	2.06	1.03	3.69	0.23	0.35	0.99	2.61	1.00	4.21
	25:25:25	0.14	0.17	1.00	1.93	0.96	3.54	0.14	0.19	0.90	2.26	0.94	4.12
30:30:30	0.08	0.09	0.87	1.59	0.77	3.01	0.06	0.08	0.74	1.89	0.82	3.92	
1.0	3:3:3	0.71	1.97	1.15	3.66	1.56	4.79	0.62	2.39	1.10	4.05	1.36	5.25
	5:5:5	0.52	1.42	1.08	3.11	1.23	3.99	0.51	1.60	1.03	3.57	1.28	4.17
	8:8:8	0.36	0.74	1.05	2.48	1.15	3.48	0.34	0.84	0.99	3.19	1.17	3.82
	10:10:10	0.27	0.55	1.03	2.29	1.00	2.99	0.26	0.64	0.95	2.63	1.09	3.38
	15:15:15	0.09	0.15	0.99	2.07	0.85	2.85	0.11	0.20	0.81	2.36	1.00	3.20
	20:20:20	0.06	0.07	0.89	1.69	0.77	2.79	0.04	0.05	0.73	1.67	0.92	3.06
	25:25:25	0.01	0.01	0.66	1.21	0.65	2.75	0.01	0.01	0.72	1.42	0.84	2.99
30:30:30	0.00	0.00	0.54	0.92	0.48	2.54	0.02	0.02	0.55	0.99	0.72	2.87	

Çizelge 2. Ortalamalar arasındaki fark (δ)-varyans oranları-gözlem sayıları-grup sayısı kombinasyonlarında göre gerçekleşen ampirik III. Tip hata olasılığı (γ) (%), $k=4$

Varyans oranları		$\chi^2 (3)$						Exp (0.50)					
Dağılımlar		1:1:1:1		1:2:4:6		1:5:9:12		1:1:1:1		1:2:4:6		1:5:9:12	
δ_i	$n_1:n_2:n_3$	F	K	F	K	F	K	F	K	F	K	F	K
0.25	3:3:3:3	2.94	9.76	5.01	14.60	5.11	15.78	3.01	11.26	4.96	18.27	5.58	22.46
	5:5:5:5	2.47	6.70	4.86	12.23	4.88	13.16	2.70	7.92	4.91	15.04	5.53	18.73
	8:8:8:8	2.76	5.72	4.83	11.06	4.84	11.87	2.63	6.48	4.88	12.69	5.40	16.36
	10:10:10:10	2.42	4.93	4.79	10.19	4.81	10.95	2.55	5.72	4.84	11.67	5.37	15.11
	15:15:15:15	2.62	4.39	4.74	9.83	4.79	10.63	2.41	4.93	4.82	9.99	5.28	13.24
	20:20:20:20	2.72	4.16	4.71	8.88	4.76	9.08	2.36	4.33	4.66	9.37	5.14	11.88
	25:25:25:25	2.53	3.86	4.69	8.15	4.66	8.41	2.32	3.90	4.50	8.85	5.00	10.74
	30:30:30:30	2.47	3.43	4.65	7.56	4.51	8.05	2.30	3.73	4.38	8.36	4.75	10.01
0.50	3:3:3:3	2.53	9.37	4.59	13.17	4.76	14.63	2.51	10.53	4.55	18.10	5.39	21.96
	5:5:5:5	2.42	6.99	4.57	10.09	4.63	11.88	2.36	7.86	4.46	14.55	5.34	18.67
	8:8:8:8	2.17	5.30	4.50	9.89	4.60	10.0	2.24	5.89	4.44	11.96	5.09	15.97
	10:10:10:10	2.29	4.90	4.43	8.24	4.59	9.17	2.21	5.65	4.34	11.58	5.08	14.81
	15:15:15:15	2.11	3.69	4.40	7.69	4.51	8.98	2.11	4.44	4.33	9.93	5.02	12.55
	20:20:20:20	1.75	2.88	4.35	7.27	4.41	8.18	2.05	3.45	4.30	8.60	4.92	10.85
	25:25:25:25	1.78	2.46	4.29	6.77	4.37	7.12	1.74	2.77	4.20	8.29	4.79	10.21
	30:30:30:30	1.45	1.97	4.19	5.62	4.21	6.82	1.64	2.37	4.09	7.56	4.60	9.74
0.75	3:3:3:3	2.40	9.48	4.47	12.06	4.37	13.41	2.48	9.98	4.15	17.72	5.05	21.82
	5:5:5:5	2.20	6.49	4.41	9.91	4.28	10.42	2.01	7.14	4.02	14.31	4.96	18.11
	8:8:8:8	2.11	4.82	4.26	9.63	4.23	9.36	2.00	5.70	4.01	11.04	4.95	15.59
	10:10:10:10	1.87	3.82	4.01	8.02	4.19	8.57	1.77	4.60	3.95	10.46	4.81	14.30
	15:15:15:15	1.30	2.13	3.63	6.67	4.02	7.84	1.24	2.44	3.83	9.44	4.70	12.17
	20:20:20:20	0.77	1.10	3.34	5.93	3.88	7.37	0.91	1.43	3.62	8.16	4.58	10.14
	25:25:25:25	0.46	0.62	3.04	5.28	3.74	7.01	0.50	0.73	3.50	7.78	4.50	8.84
	30:30:30:30	0.29	0.40	2.99	4.83	3.33	6.15	0.31	0.40	3.43	7.09	4.11	7.56
1.0	3:3:3:3	2.27	8.60	4.01	10.26	3.95	12.00	2.11	9.71	3.83	16.68	4.66	20.79
	5:5:5:5	1.94	5.62	3.85	9.32	3.81	9.92	1.88	6.90	3.54	12.62	4.49	16.44
	8:8:8:8	1.45	3.10	3.74	8.40	3.78	9.16	1.39	3.50	3.44	10.08	4.41	13.60
	10:10:10:10	0.97	1.95	3.58	7.52	3.65	9.00	1.02	2.34	3.29	9.58	4.39	12.56
	15:15:15:15	0.48	0.67	3.01	5.11	3.57	7.01	0.49	0.85	2.89	7.59	4.34	10.74
	20:20:20:20	0.16	0.22	2.33	4.36	3.07	5.36	0.17	0.27	2.48	5.77	4.05	9.47
	25:25:25:25	0.08	0.09	2.12	2.90	2.54	3.86	0.07	0.08	2.10	4.18	3.98	8.10
	30:30:30:30	0.04	0.04	1.63	2.15	2.01	2.91	0.03	0.03	1.70	3.08	3.58	6.39

Kaynaklar

- Alexander, R. A. and D. M. Govern, 1994. A new and simpler approximation for ANOVA under variance heterogeneity. *J. of Educational Statistics*, 19, 91-101.
- Anonymous, 1994. FORTRAN Subroutines for Mathematical Applications. IMSL MATH/LIBRARY. Vol. 1-2. Visual Numerics, Inc., Houston, USA.
- Brown, M. B. and A. B. Forsythe, A. B. 1974. The small sample behavior of some statistics which test the equality of several means. *Technometrics*, 16, 129-132.
- Edgington, E. S. 1974. A new tabulation of statistical procedures used in APA journals. *American Psychologist*, 29, 25-26.
- Hopkins, B. 1973. Educational research and Type III errors. *The Journal of Experimental Education*, 41, 31-32.
- Kaiser, H. F. 1960. Directional statistical decisions. *Psychological Review*, 67, 160-167.
- Kimball, A. W. 1957. Errors of the third kind in statistical consulting. *J. Am. Stat. Assoc.*, 52, 133-142.
- Kimmel, H. D. 1957. Three criteria for the use of one-tailed tests. *Psychological Bulletin*, 54, 351-353.
- Krutchoff, R. G. 1988. One-way fixed effects analysis of variance when the error variances may be Unequal. *J. Statist. Comput. Simul.* 30, 259-71.
- Leventhal, L. and C. L. Huynh, 1996. Directional decisions for two-tailed tests: power, error rates and sample size. *psychological Methods*, 278-292.
- Mcdonald, P. 1999. Power, Type I and Type III error rates of parametric and nonparametric statistical tests. *The J. of Experimental Education*, 67, 367-379.
- Mendeş, M. 2002. Varyansların heterojen olması durumunda K-istatistiği (KANOVA) ile ANOVA F testinin gerçekleşen 1.Tip hata olasılıkları bakımından karşılaştırılması. *Ankara Üniv. Ziraat Fak. Tarım Bilimleri Dergisi*, 9 (1) 23-28, Ankara.
- Mendeş, M. ve E. Başpınar, 2003. Normal olmayan dağılımlı populasyonlardan alınan örneklerde hesaplanan çeşitli test istatistiklerinin 1.Tip hata olasılıkları bakımından karşılaştırılması. *Ankara Üniv. Ziraat Fak. Tarım Bilimleri Dergisi*, 8 (3), 238-241, Ankara.
- Murphy, K. R. and B. Myors, 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 .
- Sansgiry, P. and O. Akman, 2000. Transformations of the Lognormal Distribution as a Selection Model. The American Statistician, 307-309.
- Sharon, S. and K. M. Carpenter, 1999. The right answer for the wrong question: consequences of type III Error for public health research. American J. of Public Health, 1175-1181.
- Shaffer, J. P. 1972. Directional statistical hypothesis and comparisons among means. Psychological Bulletin, 77, 195-197.
- Sokal, R. R. and F. J. Rohlf, 1995. Biometry. The Principles and Practice of Statistics in Biological Research. Third Ed. W. H. Freeman and Co. New York, 887 .
- Tomarken, A. J. and R. C. Serlin, 1986. Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. Psychological Bulletin, 99, 90-99.
- Tiku, M. L. 1971. Power function of F-test under non-normal situations. J. Amer. Statist. Assoc. 66, 913-916.
- Welch, B. L. 1951. On the comparison of several mean values: an alternative approach. Biometrika, 38, 330-336.
- Wilcox, R. R. 1989 Adjusting for unequal variances when comparing means in one-way and two-way effects ANOVA models. J. of Educational Statistics. 14, 69-278.
- Winer, B. J., D. R. Brown and K. M. Michels, 1991. Statistical Principles in Experimental Design. Third Ed., Mc Graw-Hill, Inc. USA, 1057 .
- Zar, J. H. 1999. Biostatistical Analysis. Fourth Ed., Prentice-Hall, Inc. USA, 683.

İletişim adresi:**Mehmet MENDEŞ**Çanakkale Onsekiz Mart Üniversitesi Ziraat Fakültesi
Zootekni Bölümü-Çanakkale