

**KORELASYON YAPILARININ I.TİP HATA ÜZERİNE ETKİSİ****Benian TEKİNDAL<sup>1</sup>****ÖZET**

Bu çalışmada gözlemler arasındaki muhtelif korelasyon düzeylerinin, ANOVA F (F testi) ve Tekrarlanan ölçümlü varyans analizi tekniği (RF) bakımından gerçekleşen 1.Tip hata olasılığı üzerine etkisi araştırılmıştır. Bu amaçla, ( $\rho=0.0, 0.25, 0.50, 0.75$ ) olmak üzere 4 farklı korelasyon yapısı yanında farklı grup sayısı ( $k=3, 4, 5$ ) ve örnek hacmi de ( $n=3, 5, 8, 10, 20, 30$ ) dikkate alınmıştır. Her deneme koşulu 50 000 defa tekrarlanmıştır. Yapılan simülasyon denemeleri sonucunda,  $\rho=0.0$  iken her iki testin de deneme başında karşılaştırılan 1.Tip hata olasılığını deneme sonunda da aynı düzeyde korudukları görülmüştür. Ancak, gözlemler arasında korelasyonun bulunması halinde (gözlemlerin bağımlı olması) F testinin bundan oldukça etkilendiği görülmüştür.  $\rho=0.50$  ve daha fazla olduğunda  $n=5$  olması durumunda bile F testi bakımından gerçekleşen 1.Tip hata olasılığı %0.0 olarak gerçekleşmiştir. Buna karşın, RF testinin hem gözlemlerin bağımlı olması durumunda hem de bağımsız olması durumunda karşılaştırılan 1.Tip hata olasılığını deneme sonunda da hemen hemen aynı düzeyde koruduğu görülmüştür.

**Anahtar Kelimeler:** 1.Tip Hata, Tekrarlanan Ölçüm, Varyans Analizi Tekniği, Bağımsız ve Bağımlı Gözlemler

**THE EFFECT OF CORRELATION STRUCTURES ON TYPE I ERROR RATES****ABSTRACT**

We compared ANOVA F (F test) and repeated measurement analysis of variance techniques (RF test) with respect to Type I error rate under various experimental conditions (such as number of groups ( $k=3, 4, 5$ ) sample size (3, 5, 8, 10, 20, 30) and correlation structure among the observations ( $\rho=0.0, 0.25, 0.50, 0.75$ )). For  $\rho=0.0$  correlation coefficient, both F and RF test gave almost the same results in terms of protect the predetermined alpha level. At the end of 50,000 simulation trials, it was observed that the actual Type I errors of F test were affected by the correlation structure

among the observations, but not RF test. When  $\rho=0.50$  and more and  $n=5$ , the Type errors was 0. 0 % for F test. On the other hand, whether observations were independent or dependent the Type I error rates for RF test were almost equal to predetermined level.

**Key Words:** Type I Error Rate, Repeated Measurements, ANOVA F, Independent and Dependent Observations.

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

## 1.GİRİŞ

Uygulamada aynı deney ünitelerinden farklı zaman periyotlarında ya da farklı muamele koşulları altında verilerin elde edildiği durumlarla oldukça sık karşılaşmaktadır. Mesela, enerji içerikleri farklı üç rasyonun ivesi kuzularında canlı ağırlık artışına etkisinin araştırıldığı bir denemede deney ünitesi olan kuzuların canlı ağırlık artışlarının 10 hafta boyunca tespit edilmesi durumunda her kuzudan toplam 10 ölçüm değeri elde edilir. Söz konusu ölçümler aynı deney ünitelerine ait olduğu için ölçüm değerleri arasında bir korelasyon oluşacağından gözlemler bağımlı hale gelir (Winer, B.J., Brown, D.R., Michels, K.M., 1991; Keskin ve Mendeş, 2001). Bu gibi durumlarda üzerinde durulan özellik bakımından söz konusu muamele gruplarının karşılaştırılmasında klasik varyans analizi tekniğinin (ANOVA F testi) kullanılması uygun olmaz. Çünkü, varyans analizi tekniğinin kullanılabilmesi için söz konusu verilerin ayrı deney ünitelerinden elde edilmesi yani gözlemlerin bağımsız olması gerekir (Sokal ve Rohlf 1995; Zimmerman, 1998; Zar, 1999; Mendeş, 2002; Mendeş, 2003; Mendeş, 2004a). İşte bu gibi durumlarda uygulanan varyans analizi tekniği, tekrarlanan ölçümlü varyans analizi (repeated measurement design) olarak adlandırılır (Winer, B.J., Brown, D.R., Michels, K.M., 1991; Wilcox ve Keselman, 2003). Ancak, gözlemlerin bağımlı olması durumunda grup ortalamalarının varyans analizi tekniğinden yararlanılarak karşılaştırılması, veriler arasındaki korelasyon düzeyine bağlı olarak deneme sonunda gerçekleşen 1.Tip hata olasılığının, deneme başında kararlaştırılan 1.Tip hata olasılığından daha düşük olmasına neden olmaktadır. Yani negatif yönde bir sapmaya neden olur (negative biased) (Çamdeviren, 1995; Kurita, 1996; Wilcox ve ark., 2000). Halbuki, hipotez kontrolleri sonucunda varılan kararların güvenilirlikleri, deneme başında kararlaştırılan 1.Tip hata ( $\alpha$ ) ve II.Tip hata ( $\beta$ ) olasılıklarının (dolayısıyla testin gücünün), deneme sonunda da ne düzeyde korunup korunmadığı ile oldukça ilişkilidir. İstenen durum, deneme başında kararlaştırılan hata olasılıklarının deneme sonunda da aynı düzeyde korunmasıdır. Bu iki hata olasılıkları arasında  $\alpha=4\beta$  şeklinde ilişki olduğundan, 1.Tip ve II.Tip hata olasılığı genel olarak deneme başında 0.05 ve 0.20 olarak kararlaştırılmaktadır. Dolayısıyla deneme sonunda da bu iki hata olasılıklarının 0.05 ve 0.20 civarında gerçekleşmesi, elde edilen sonuçların güvenilirliklerini arttırması bakımından oldukça önemlidir (Wang, 1971; Murphy ve Myers 1998; Zar, 1999; Mendeş, 2004b). Diğer taraftan deneme sonunda gerçekleşen 1.Tip hata olasılığının  $0.045 \leq \alpha \leq 0.055$  ve  $0.025 \leq \alpha \leq 0.075$  şeklinde tarif edilen iki aralıktan herhangi birisine düşmesi (Bradley, 1978) ya da Cochran (1954) tarafından önerilen  $0.040 \leq \alpha \leq 0.060$  aralığa düşmesi durumunda da kabul edilebilir bir hata olasılığı olduğu bildirilmiştir. Buna karşın halen pek çok araştırmacı tarafından gözlemlerin bağımsızlığı ön şartının ihmal edilerek söz konusu grup ortalamalarının karşılaştırılmasında üzerinde tekrarlanan ölçümlerin alındığı faktörün (zaman periyodu, farklı muamele koşulları, vb) seviyelerinin birlikte değil de, her seferinde ayrı ayrı bir seviyesinin dikkate alınması ve söz konusu karşılaştırılmaların klasik varyans analizi tekniğinden yararlanılarak yapılması bir sorun olarak karşımıza çıkmaktadır. Bu çalışmanın amacı, hem gözlemlerin bağımlı olması durumunda söz konusu karşılaştırmaların klasik varyans analizi tekniği yerine tekrarlanan ölçümlü deneme düzenlerinde varyans analizi tekniğinden yararlanılarak yapılması gerektiğine dikkati çekmek, hem de ölçümler arasındaki muhtelif bağımlılık yapılarının yani korelasyonların kararlaştırılan 1.Tip hata olasılığını nasıl etkilediğini irdelemektir.

## 2.MATERYAL VE METOT

Bu çalışmanın materyalini Microsoft Fortran Developer Stüduo'nun IMSL kütüphanesinin RNMNV alt fonksiyonundan (subroutine function) yararlanılarak çok değişkenli standart normal dağılımdan üretilen tesadüf sayıları teşkil etmektedir (Anonymous, 1994). Çalışmada dikkate alınan deneme koşulları Çizelge 1'de topluca verilmiş olup, bu deneme koşullarının her birisi için 50 000 simülasyon denemesi yapılmıştır. I.Tip hata olasılığı ( $\alpha$ ) deneme başında %5 olarak kararlaştırılmıştır.

**Çizelge 1. Çalışmada Dikkate Alınan Grup Sayısı, Örnek Hacmi ve Gözlemler Arasındaki İlişki Yapıları**

Grup sayısı (k)	Örnek hacmi (n)	Gözlemler Arasındaki İlişki düzeyleri (korelasyonlar)	
		Bağımsız	Bağımlı
3	3:3:3, 5:5:5, 8:8:8, 10:10:10, 20:20:20, 30:30:30	0.0	0.25, 0.50, 0.75
4	3:3:3:3, 5:5:5:5, 8:8:8:8, 10:10:10:10, 20:20:20:20, 30:30:30:30	0.0	0.25, 0.50, 0.75
5	3:3:3:3:3, 5:5:5:5:5, 8:8:8:8:8, 10:10:10:10:10, 20:20:20:20:20, 30:30:30:30:30	0.0	0.25, 0.50, 0.75

Simülasyon çalışmasında Çizelge 1'den de görüleceği üzere gözlem değerleri üretilirken yani korelasyon matrisi oluşturulurken;

a) Gruplardaki gözlemler birbirinden bağımsız olacak şekilde (yani gruplardaki gözlemler arasındaki korelasyon sıfır olduğunda)

b) Gruplardaki gözlemler arasında 0.25, 0.50 ve 0.75'lik korelasyonlar olacak şekilde

iki farklı korelasyon matrisi oluşturulmuştur. Bu korelasyonlardan 0.25 düşük korelasyon, 0.50 orta düzeyde korelasyon ve 0.75'de yüksek düzeyde korelasyon olarak kabul edilmiştir. Çalışmada dikkate alınan grup sayısı olarak uygulamada sık karşılaşılan 3, 4 ve 5 grubun olması durumları ve gruplardaki gözlem sayıları da yine uygulamada sık olarak karşılaşılan gözlem sayılarıyla bir uyum göstermesi açısından 3, 5, 8, 10, 20 ve 30 olarak kararlaştırılmıştır. Örnek hacminin böyle seçilmesinde etkili olan diğer bir faktör de, deney ünitesi kılığının çekildiği alanlarda RF testinin kullanılmasının elde edilecek sonuçların güvenilirliği bakımından bir avantaj sağlamasıdır. Çizelge 1'deki deneme koşulları için üretilen veriler önce ANOVA F testi ile değerlendirilmiş ve deneme sonunda gerçekleşen I.Tip hata olasılıkları tespit edilmiştir. Daha sonra da söz konusu verilerin değerlendirilmesinde tekrarlanan ölçümlü varyans analizi tekniği kullanılmış ve gerçekleşen I.Tip hata olasılıkları tespit edilmiştir. Elde edilen sonuçlar ise Çizelge 2'de topluca verilmiştir.

### 3.SİMÜLASYON BULGULARI VE TARTIŞMA

**Çizelge 2.** Farklı Korelasyon Yapıları, Örnek Hacmi ve Grup Sayısına Göre Gerçekleşen 1.Tip Hata Olasılıkları (%)

Grup sayısı (k)		3				4				5			
		Gözlemler arasındaki korelasyonlar ( $\rho$ )											
Örnek Hacmi	Test	0.00	0.25	0.50	0.75	0.00	0.25	0.50	0.75	0.00	0.25	0.50	0.75
3	F	4.9	3.7	0.9	0.3	4.8	3.5	0.8	0.2	5.1	3.4	0.7	0.1
	RF	5.1	4.8	4.9	4.8	5.0	5.0	4.9	4.7	4.8	4.9	5.1	5.0
5	F	5.1	3.3	0.7	0.2	4.9	3.0	0.6	0.1	5.0	2.5	0.4	0.0
	RF	5.2	5.1	4.9	4.9	5.2	5.1	5.0	5.0	4.9	5.0	5.0	4.9
8	F	5.0	2.7	0.5	0.1	5.1	2.7	0.4	0.0	4.9	2.1	0.3	0.0
	RF	5.1	5.0	5.1	4.9	5.2	5.1	5.0	4.8	5.2	5.1	5.1	5.0
10	F	4.9	2.5	0.2	0.0	5.0	2.4	0.4	0.0	5.2	2.0	0.1	0.0
	RF	4.9	5.2	4.9	4.9	4.9	4.9	5.1	5.1	4.8	4.9	4.9	4.8
20	F	5.2	2.4	0.2	0.0	5.1	2.2	0.1	0.0	5.1	1.8	0.0	0.0
	RF	5.1	5.1	4.9	5.1	5.0	5.2	5.1	4.8	4.9	5.1	5.1	5.0
30	F	5.1	2.2	0.1	0.0	5.1	2.0	0.0	0.0	5.2	1.7	0.0	0.0
	RF	5.0	5.1	5.2	5.1	4.9	4.8	4.9	4.9	5.1	5.1	4.9	5.0

**F:** ANOVA F, **RF:** Tekrarlanan Ölçümlü Varyans Analizi

Simülasyon sonuçları incelendiğinde gözlemlerin bağımsız olması durumunda ( $\rho=0.0$ ), grup sayısı ve örnek hacmi ne olursa olsun hem Anova F testi (F) hem tekrarlanan ölçümlü varyans analizi (RF) bakımından gerçekleşen 1.Tip hata olasılıklarının hep %5 civarında olduğu dikkati çekmektedir (Çizelge 2). Yani gözlemler bağımsız iken deneme başında karşılaştırılan 1.Tip hata olasılığı hem F testinde hem de RF testinde deneme sonunda da aynı düzeyde korunmaktadır. Ancak, F testi bakımından gerçekleşen 1.Tip hata olasılıkları gözlemler arasında düşük korelasyonların bulunması durumunda bile giderek %5'ten sapmaktadır. Söz konusu sapmanın da negatif yönde olduğu (%5'ten daha düşük) olduğu görülmektedir. Yani gözlemler arasında korelasyonların bulunması, deneme başında %5 olarak karşılaştırılan 1.Tip hata olasılığının deneme sonunda %5'ten daha düşük olarak

(negative biased) gerçekleşmesine neden olmaktadır. Bu çalışmada elde edilen bulgular Kurita (1996), Wilcox ve ark. (2000) ve Wilcox ve Keselman (2003) yaptığı simülasyon çalışmalarını destekleyen bulgulardır. Bu açıdan bakıldığında her üç çalışma arasında bir paralelliğin olduğu görülür. Mesela,  $k=3$  ve gözlemler arasında  $\rho=0.25$ 'lik (düşük korelasyon) bir korelasyonun bulunması durumunda RF testi bakımından gerçekleşen en düşük 1.Tip hata olasılığı %4.8 iken, F testi bakımından gerçekleşen en büyük 1.Tip hata olasılığı %3.7 dir. Gözlemler arasındaki korelasyonun artması durumunda F testi bakımından gerçekleşen 1. Tip hata yapma olasılıklarında belirgin bir düşme olmaktadır. Bu düşme, grup sayısı ve gruplardaki gözlem sayısının artmasına paralel olarak daha da belirginleşmektedir. Dolayısıyla aralarında korelasyon bulunan gözlemler içeren grupların karşılaştırılmasında F testinden yararlanması, gerçekleşen 1.Tip hata olasılığının daima kararlaştırılandan daha düşük olmasına neden olmaktadır. Bu durum, örnek hacmi arttıkça gözlemler arasındaki korelasyonun daha belirgin olarak ortaya çıkmasından dolayı çok daha belirgin olmaktadır. Büyük hacimli örneklerle çalışılması durumunda ise 1.Tip hata olasılığının 0.0 olarak gerçekleşmesi gibi bir durumla karşılaşılmaktadır. Yani gerçekleşen 1.Tip hata olasılığı için  $0.025 \leq \alpha^* \leq 0.075$  gibi geniş sayılabilecek bir aralık veren Bradley kriterine bile uymamaktadır. Mesela, gerçekleşen 1.Tip hata olasılıkları;  $k=3$ ,  $\rho=0.75$  ve  $n=3$  iken 0.3 olarak gerçekleşmekte,  $n=8$  iken 0.1 olarak ve  $n=20$  iken 0.0 olarak gerçekleşmektedir. Aynı koşullarda grup sayısının 4 olması durumunda gerçekleşen 1.Tip hata olasılıklarının 0.2, 0.0 ve 0.0 olarak gerçekleştiği, grup sayısının 4 olması durumunda ise 1.Tip hata olasılıklarının 0.1, 0.0 ve 0.0 olarak gerçekleştiği görülmektedir. Dikkat edileceği üzere grup sayısı, gruplardaki gözlem sayısı ve gözlemler arasındaki korelasyon arttıkça, 50 000 deneme sonucunda reddedilen  $H_0$  hipotezi sayısı sıfır olmaktadır. Buradan hareketle, farklı gruplarda yer alan gözlemler arasında bir bağımlılığın yani korelasyonun mevcut olmasının gerçekleşen 1. Tip hata olasılıklarında önemli düzeyde düşmeye sebep olduğu söylenebilir. Bu düşmeye grup sayısı ve gruplardaki gözlem sayılarının da etkisi vardır. Tekrarlanan ölçümlü varyans analizinde, bütün deneme koşullarında gerçekleşen 1.Tip hata olasılıklarının hep deneme başında kararlaştırılan seviyede (%5) kalmıştır. Wang (1971), Hsiung and Olejnik (1996), Schneider ve Penfield (1997), Ware (1997), Cliff (1997), Wilcox (1997), Gorham (1998), Zimmerman (1998), Myers (1998), Zar (1999), Ferron ve Sentovich (2002), Mendeş (2003), Mendeş ve Başpınar (2003) yaptıkları simülasyon çalışmalarında gerçekleşen 1.Tip hata olasılığı için kabul edilebilir olasılıkların %4.5-5.9 arasında değişen olasılıklar olduklarını bildirmişlerdir. Bu açıdan düşünüldüğünde RF testi bakımından gerçekleşen 1.Tip hata olasılıkları, hangi deneme koşulu olursa olsun hep %5 civarındadır ve bu olasılıkları da bu literatür çalışmalarını destekler niteliktedir.

#### 4.SONUÇ VE ÖNERİLER

50000 simülasyon denemesi sonucunda elde edilen bulgular doğrultusunda;

- Gözlemlerin bağımsız olması durumunda hem F, hem de RF testinin deneme başında kararlaştırılan 1.Tip hata olasılığını deneme sonunda da aynı düzeyde korudukları,
- Gözlemlerin bağımlı olması (gözlemler arasında korelasyonun bulunması) durumunda, F testinin kullanılması durumunda başlangıçta kararlaştırılan 1.Tip hatanın deneme sonunda negatif sapmalı olarak tahmin edilmesine neden olmaktadır. Bu durum ise, gözlemler arasındaki korelasyonun artması durumunda çok daha belirgin olmaktadır. Yani gözlemler arasındaki korelasyonun artması, ilgili  $H_0$  hipotezinin gereğinden fazla sayıda kabul

edilmesine, hatta orta ve yüksek korelasyon yapısında  $H_0$  hipotezlerinden hiç birinin reddedilmemesine neden olmaktadır. Buna karşın RF testi bakımından gerçekleşen 1.Tip hata olasılıklarının gözlemler arasındaki korelasyon yapısındaki değişimlerden etkilenmediği, dolayısıyla bu gibi durumlarda RF testinin kullanılması gerektiği sonuçlarına varmak mümkündür.

#### KAYNAKLAR

- Anonymous, (1994). FORTRAN subroutines for Mathematical Applications. IMSL MATH/LIBRARY. Vol.1-2. *Visual Numerics*, Inc., Houston, USA.
- Bradley, J.C. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31, 144-152.
- Cliff, N.(1997). Robustness and Power of D Statistics in Comparison to t For Paired Data. DAI-B 57/07, 4774.
- Cochran, W.G (1954).Some methods for strengthening the common  $\chi^2$ -tests.*Biometrics*, 10, 417-451.
- Çamdeviren, H. (1995). Tekrarlanan Ölçümlü Deneme Düzenleri. (Yayınlanmamış Yüksek Lisans Tezi). A.Ü.Fen Bilimleri Enstitüsü
- Ferron, J., Sentovich, C. (2002). Statistical power of randomization tests used with multiple- baseline designs. *Journal of Experimental Education*, 70(2), 165-178.
- Gorham, J.L., (1998). The Effects on Type I Error Rate and Power of Selected Competitors to the ANOVA F Test in Randomized Block Design Under Non-Normality and Variance Heterogeneity. DAI-B 58/12, 6650.
- Hsiung, T.H. , Olejnik, S. (1996). Type I Error Rates and Statistical Power for the James 2<sup>nd</sup>-Order Test and the Univariate F-Test in 2-Way Fixed Effects ANOVA Models Under Heteroscedasticity and/or Nonnormality. *Jour. of Experimental Education*, 65(1) 57-71.
- Keskin S., Mendeş, M. (2001). Faktörlerden Birinin Seviyelerinde Tekrarlanan Ölçüm İki Faktörlü Deneme Düzenleri. Selçuk Üniv.Ziraat Fak. Dergisi. C:15, S: 25, 42-53.
- Kurita, K. (1996). The Biasing Effects of Violating the Independence Assumption Upon the Power of t-Test. Jap.*Jour. of Educ. Psyc.* 44(2), 234-242.
- Mendes, M. (2002). Normallik ve Varyansların Homojenliği Ön Şartlarının Sağlanması Durumlarında Varyans Analizi T Yekniği Yerine Kullanılabilecek Bazı Parametrik Testlerin 1.Tip Hata ve Testin Gücü Bakımından İrdelenmesi. (Yayınlanmamış Doktora Tezi. )Ankara Üniv. Fen Bilimleri Ens.
- Mendeş, M., (2002). Varyansların heterojen olması durumunda K istatistiği (KANOVA) ile ANOVA F Testinin Gerçekleşen 1.Tip Hata Olasılıkları Bakımından Karşılaştırılması. *A.Ü.Z.F. Tarım Bilimleri Dergisi*, 8(3), 238-241.
- Mendes, M. (2002). Levene, Bartlett, Neyman-Pearson ve Bartlett 2 testlerinin 1.Tip hata olasılıkları bakımından karşılaştırılması. *Tarım Bilimleri Dergisi*, 9(2), 143-146.

- Mendeş, M. ve A.Pala, (2003). "Type I Error Rate and Power of Three Normality Tests," *Pakistan Journal of Information and Technology*, 2 (2), 135–139
- Mendes, M. (2004a). ANOVA F ve K-istatistiğinin III.Tip Hata Olasılıkları Bakımından Karşılaştırılması. . *A.Ü.Z.F. Tarım Bilimleri Dergisi*, 10 (2), 121-126.
- Mendes, M. (2004b). Ortanca Değerin Kullanılmasının ANOVA F testi bakımından gerçekleşen I.Tip Hata Olasılığı Üzerine Etkisi *A.Ü.Z.F. Tarım Bilimleri Dergisi*, 10 (1), 20-23.
- Murphy, K.R., Myers, B. (1998). *Statistical Power Analysis*. A simple and General Model for Traditional and Modern Hypothesis Tests. Lawrence Erlbaum Assoc., Publishers, s.120, London, UK.
- Myers, L. (1998). Comparability of the James' Second-Order Asroximation Test and the Alexander and Govern a Statistic for Non-Normal Heterosdecastic Data. *Journal of Statistical Computation&Simulation*, 60, 207-223.
- Schanieder, P.J., Penfield, D.A. (1997). Alexander and Govern's Approximations: Providing an Alternative to ANOVA Under Variance Heterogeneity, *The Journal of Experimental Education*, 65(3), 271-286.
- Sokal, R.R., Rohlf, F.J. (1995). *Biometry*. The Principles and Practice of Statistics in Biological Research. Third Ed. W.H.Freeman and Co. New York, 887 .
- Wang, Y.Y. (1971). Probabilities of the Type I Errors of the Welch Test for the Behrens-Fisher Problem. *Jour. of the American Statist. Assoc.* 66(335), 605-608.
- Ware, W.B., (1997). Detecting Departures From Normality: A Monte Carlo Simulation of a New Omnibus Test Based on Moments. DAI-A 58/04, 1196.
- Wilcox, R.R. (1997). A Bootstrap modification of the Alexander-Govern ANOVA method, plus comments on compairing trimmed means. *Educational and Psychological Measurement*, 57, 655-665.
- Wilcox, R.R., Keselman, H.J., Muska, J., Cribbie, R. (2000). Repeated measures ANOVA: Some new results on comparing trimmed means and means. *British Journal of Mathematical and Statistical Psychology*, 53, 69-82.
- Wilcox, R.R., Keselman, H.J. (2003). Repeated measures one-way ANOVA based on a modified one-step M-estimator. *British Journal of Mathematical and Statistical Psychology*, 56, 15-25.
- Winer, B.J., Brown, D.R., Michels, K.M. (1991).Statistical Princeptles in Experimental Design. Third Ed., McGraw-Hill, Inc., s.1057, USA.
- Zar, J.H.(1999). *Biostatistical Analysis*. Fourth Ed. Prentice Hall Inc., Upper Saddle River, New Jersey, USA. 663 S.
- Zimmerman, D.W. (1998). Invalidation of Parametric and Nonparametric Statistical Tests by Concurrent violation of two Assumptions. *Journal of Experimental Education*, 67, 55-69.