



Tek yönlü VA için seçenек bir test yöntemi ve geliştirilen bilgisayar yazılımı

Engin Yıldıztepe

Dokuz Eylül Üniversitesi
Fen-Edebiyat Fakültesi İstatistik Bölümü
Kaynaklar Yerleşkesi 35160, Buca,
İzmir, Türkiye
engin.yildiztepe@deu.edu.tr

A. Fırat Özdemir

Dokuz Eylül Üniversitesi
Fen-Edebiyat Fakültesi İstatistik Bölümü
Kaynaklar Yerleşkesi 35160, Buca,
İzmir, Türkiye
firat.ozdemir@deu.edu.tr

Özet

Oldukça geniş uygulama alanlarına sahip olan tek yönlü varyans çözümlemesi yöntemi, özellikle dengesiz deney düzenlerinde sabit varyanslık varsayımının sağlanmamasından çok etkilenir. Bu etkilenme, deneydeki tekrar sayısı ve kitle varyansı eşleşmesine göre farklı yönlerde olmak üzere, araştırmacı tarafından belirlenen nominal anlam düzeyinden sapma biçiminde gerçekleşir (Harwell, M.R. ve diğerleri, 1992; Lix, L.M. ve diğerleri, 1996). Bu çalışmada Özdemir ve Kurt (2006) tarafından etken düzeyleri özel seçilmiş tek yönlü varyans çözümlemesinde değişen varyanslık durumunda kullanılmak üzere önerilen B^2 test yöntemine kullanım kolaylığı sağlamak amacıyla geliştirilen bir özel yazılım tanıtılacaktır. Bu yazılım <http://kisi.deu.edu.tr/userweb/engin.yildiztepe/BKare.rar> adresinden temin edilebilir.

Anahtar sözcükler: ANOVA; Değişen varyanslık; B^2 testi, İstatistik yazılımı.

Abstract

An Alternative Test for One Way ANOVA and Special Software

One way analysis of variance which has a wide application area is very sensitive to inequality of variance problem especially when the design is unbalanced. The impact of this assumption violation give rise to an increase or decrease in nominal significance level which is determined by the researcher depending on the pairing of number of replication and population variances (Harwell, M.R. et al., 1992; Lix, L.M. et al., 1996). In this study, a special software has been presented which was developed in order to ease the application of B^2 test which was proposed by Özdemir and Kurt (2006) to create an alternative when there is a inequality of variance problem in one way ANOVA. This software is available at <http://kisi.deu.edu.tr/userweb/engin.yildiztepe/BKare.rar>.

Keywords: ANOVA; Heteroscedasticity; B^2 test; Statistical software.

1. Giriş

Varyans çözümlemesi genellikle ikiden fazla kitle ortalamasının eşitliği test edilirken yararlanılan bir yöntemdir. Burada test edilen hipotez,

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

$$H_1 : \mu_i \neq \mu_{i^*} \text{ en az bir } i, i^* \text{ çifti için}$$

biçiminde kurulur. Bu hipotez, cevap değişkeninde (Y) gözlenen toplam değişim varyans çözümlemesi yardımıyla,

$$GKT = HKT + DKT$$

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 + \sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y}_{..})^2 \quad (1)$$

olarak iki kısma ayrıldıktan sonra tüm $i=1,2,\dots,k$ düzeyleri için $Y_{ij} \sim NID(\mu_i, \sigma^2)$ $j=1,2,\dots,n_i$ olmak üzere

$$\frac{\sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y}_{..})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 / (N-k)} = \frac{DKO}{HKO} \sim F_{k-1, N-k} \quad (2)$$

biçiminde tanımlanan test istatistiğine sahip F testi ile test edilir. F testi'nin uygulanabilmesi için

- Gözlemlerin birbirinden bağımsız elde edilmesi
- Gözlemlerin elde edildiği kitle dağılımlarının normal dağılıma uyması
- Gözlemlerin elde edildiği kitle dağılımlarına ait varyansların homojen olması

varsayımlarının sağlanması gerekir. Her bir deneme ile elde edilen gözlemlerin çekildiği kitle varyanslarının homojen olmaması durumu farklı varyanslılık (heteroscedasticity) olarak adlandırılır. Farklı varyanslılık durumunda, diğer iki varsayım sağlanıyorsa deneydeki cevap değişkeninin dağılımı tüm $i=1,2,\dots,k$ düzeyleri için $Y_{ij} \sim NID(\mu_i, \sigma_i^2)$ $j=1,2,\dots,n_i$ biçiminde olur. Farklı varyanslılık durumunda karşılaşılan iki temel sorundan ilki F testine ait test istatistiğinin dağılımının $F_{k-1, N-k}$ dağılımına uymaması, ikincisi ise araştırmacı tarafından belirlenen anlam düzeyinin test sonunda korunmamasıdır.

1.1 Farklı varyanslılık durumunda F oranının dağılımında gözlenen değişim

Farklı varyanslılık durumunda $\frac{DKO}{HKO}$ oranının dağılımı konusunda yapılan en yetkin çalışma (Box, 1954) dir [9]. Box, farklı varyanslılık durumunda, deneme ve hata karelerine ait karesel biçimlerin dağılımında gözlenen değişimi tek yönlü varyans çözümlemesine uyarlamış ve $\frac{DKO}{HKO}$ oranının

$$b = \frac{N-k}{N(k-1)} \frac{\sum_{i=1}^k (N-n_i) \sigma_i^2}{\sum_{i=1}^k (n_i-1) \sigma_i^2} \quad h' = \frac{\left\{ \sum_{i=1}^k (N-n_i) \sigma_i^2 \right\}^2}{\left\{ \sum_{i=1}^k n_i \sigma_i^2 \right\}^2 + N \sum_{i=1}^k (N-2n_i) \sigma_i^4} \quad h = \frac{\left\{ \sum_{i=1}^k (n_i-1) \sigma_i^2 \right\}^2}{\left\{ \sum_{i=1}^k (n_i-1) \sigma_i^4 \right\}} \quad (3)$$

değerleri ile $bF_{h',h}$ dağılımına yaklaştığını belirtmiştir.

1.2 Araştırmacı tarafından belirlenen anlam düzeyi (nominal) ile gerçekleşen anlam düzeyi (actual) arasındaki fark

Düzenlenen deney sonunda elde edilen gözlem değerleri, varyans çözümlemesi ve F testi ile çözümlendiğinde farklı varyanslılık durumu ile karşılaşılırsa, gerçekleşen anlam düzeyi bazı durumlarda belirlenen anlam düzeyinin 3 ya da 4 katına kadar yükselebilirken bazı durumlarda ise bu düzeyin çok altında kalabilir [11]. İlk durumda araştırmacı, göze aldığı I. Tip hata yapma riskinin 3-4 katı daha fazla bir risk ile karşı karşıya olurken gerçekte doğru olan ve reddedilmemesi gereken $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ hipotezi, farklı varyanslılık problemi nedeniyle reddedilebilir. İkinci durumda ise II. Tip hata yapma olasılığı artacağından gerçekte denemeler arasında var olan farklar yakalanamayabilir.

Belirlenen anlam düzeyi ile gerçekleşen anlam düzeyi arasındaki fark kitle varyansları ve her deneme ile yapılan tekrar sayılarına bağlıdır. Farklı varyanslılığın, F testi sonuçları üzerindeki etkilerini görmek için büyük örneklem genişliği durumuna yani F oranını oluşturan DKO ve HKO değerlerinin beklenen değerlerine bakmak gerekir [9].

σ_i^2 $i=1,2,\dots,k$ kitle varyansları olmak üzere hata kareler ortalamasının beklenen değeri

$$E(HKO) = E\left[\frac{1}{N-k} \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2\right] = \frac{1}{N-k} \sum_{i=1}^k (n_i - 1) \sigma_i^2 \quad (4)$$

biçimindedir. $N-k = \sum_{i=1}^k (n_i - 1)$ olduğu için bu değer $(n_i - 1)$ ağırlıkları ile σ_i^2 değerlerinin bir ağırlıklı ortalamasıdır. H_0 'ın doğruluğu altında deneme kareler ortalamasının beklenen değeri ise

$$\begin{aligned} E(DKO) &= E\left[\frac{1}{k-1} \sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y}_{..})^2\right] \\ &= \frac{1}{k-1} \left[\sum_{i=1}^k n_i E(\bar{Y}_i - \mu)^2 - N E(\bar{Y}_{..} - \mu)^2 \right] \\ &= \frac{1}{k-1} \left[\sum_{i=1}^k n_i \frac{\sigma_i^2}{n_i} - N \frac{\sum_{i=1}^k n_i \sigma_i^2}{N^2} \right] \\ &= \frac{1}{N(k-1)} \sum_{i=1}^k (N - n_i) \sigma_i^2 \end{aligned} \quad (5)$$

biçimindedir. Bu değer de $(N - n_i)$ ağırlıkları ile σ_i^2 değerlerinin ağırlıklı ortalamasıdır.

Her deneme ile eşit sayıda tekrarın yapıldığı dengeli deney düzenlerinde H_0 hipotezi doğru ise $E(HKO) = E(DKO)$ olacağı için farklı varyanslılığın etkisi $\text{Var}(DKO)$ değerinin incelenmesi ile gözlenebilir. Bu değer

$$Var(DKO) = \frac{2\bar{\sigma}^4}{k-1} \left[1 + \frac{(k-2)}{k(k-1)} \frac{\sum_{i=1}^k (\sigma_i^2 - \bar{\sigma}^2)^2}{\bar{\sigma}^4} \right] \quad (6)$$

biçimindedir. Kitle varyanslarının homojen olması durumunda parantez içindeki değer 1'e eşit olurken farklı varyanslılık durumunda 1'den büyük çıkar. Farklı varyanslılık, Var(DKO)'nun alacağı değer, homojen varyanslılık durumuna göre daha büyük olmasına neden olur. F dağılımı tek kuyruklu bir dağılım olduğu için Var(DKO)'nun büyümesi E(DKO)'nun büyümesine bu da F oranının büyümesine neden olacaktır [9]. Dengeli deney düzenleri için farklı varyanslılık durumunda F oranının payı paydasından daha büyük değer alma eğilimi gösterir. Bunun sonucu olarak da gerçekleşen anlam düzeyi, belirlenen anlam düzeyinden daha büyük çıkar ve test, liberal bir eğilim gösterir. Ancak farklı varyanslılık nedeniyle belirlenen anlam düzeyinin korunamaması sorunu, dengesiz deney düzenlerinde olduğu kadar ciddi boyutta değildir [2, 10].

Dengesiz deney düzenlerinde farklı varyanslılığın etkisini incelemek için iki uç durum ele alınmalıdır. Bunlardan birincisi literatürde aynı yönde eşleşme (AYE) olarak adlandırılan varyansı büyük olan kitleye ait deneme ile yapılan tekrar sayısının büyük, varyansı küçük olan kitleye ait deneme ile yapılan tekrar sayısının küçük olması durumudur. Bu durumda Eşitlik (4)'te verilen E(HKO)'da büyük varyansın ağırlığı büyük olurken Eşitlik (5)'te verilen E(DKO)'da küçük varyansın ağırlığı büyük olmaktadır. Bunun sonucunda $\frac{DKO}{HKO}$ olarak tanımlanan F oranının payı, paydasından daha küçük değerler alma eğilimi gösterecek ve gerçekleşen anlam düzeyi, belirlenen anlam düzeyinden daha küçük çıkacaktır. F testi, farklı varyanslılığın gözleendiği ve aynı yönde eşleşmenin olduğu dengesiz deney düzenlerinde tutucu eğilim gösterir ve bunun sonucu olarak gerçekte denemeler arasında olan anlamlı farklar bulunamayabilir [3, 5]. Dengesiz deney düzenlerinde incelenmesi gereken ikinci uç durum ise literatürde ters yönde eşleşme (TYE) olarak adlandırılan, varyansı büyük olan kitleye ait deneme ile yapılan tekrar sayısının küçük, varyansı küçük olan kitleye ait deneme ile yapılan tekrar sayısının büyük olması durumudur. Bu durumda Eşitlik (4)'te verilen E(HKO)'da büyük varyansın ağırlığı küçük olurken Eşitlik (5)'te verilen E(DKO)'da büyük varyansın ağırlığı büyük olmaktadır. Bunun sonucunda $\frac{DKO}{HKO}$ olarak tanımlanan F oranının payı, paydasından daha büyük değerler alma eğilimi gösterecek ve gerçekleşen anlam düzeyi, belirlenen anlam düzeyinden daha büyük çıkacaktır. F testi, farklı varyanslılığın gözleendiği ve ters yönde eşleşmenin olduğu dengesiz deney düzenlerinde liberal eğilim gösterir ve bunun sonucu olarak denemeler arasında gerçekte anlamlı olmayan farklar bulunabilir [3, 5].

Uygulamada, hangi denemeye ait kitle varyansının büyük, hangi denemeye ait kitle varyansının küçük olduğunun bilinmesi elbette mümkün değildir. Burada, F testinin farklı varyanslılıktan en çok etkilendiği iki uç durum üzerinde durulmuştur. Literatürde, farklı varyanslılık sorununa çözüm olarak önerilen yaklaşımlar değerlendirilirken, özellikle bu iki uç durumdaki performanslar göz önüne alınmaktadır.

2. Yeni ve basit bir çözüm önerisi

Özel seçimli deneyler için, tek yönlü varyans çözümlemesinde farklı varyanslılık sorununa çözüm olarak geliştirilen yaklaşımları,

- Veri dönüşümleri
- Yaklaşık (approximate) testler
- Tam (exact) testler
- Parametrik olmayan testler
- Ağırlıklı en küçük kareler tahmin yöntemi

- Güçlü (robust) istatistiksel yöntemler

biçiminde 6 ana başlıkta gruplandırmak mümkündür. Bu bölümde anlatılacak yaklaşım yaklaşık testler grubuna dâhildir.

k düzeyinin özel olarak seçildiği ve bu düzeylere (denemelere) ait kitle dağılımlarının normal olduğu tek etkenli bir deneyden elde edilen Y_{ij} ($i=1,2,\dots,k$; $j=1,2,\dots,n_i$) gözlemleri ile hesaplanan deneme ortalamaları \bar{Y}_i ve bu ortalamalara ait standart hatalar da

$$S_{\bar{Y}_i} = \left[\frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}{n_i(n_i - 1)} \right]^{1/2} \quad i = 1, 2, \dots, k \quad (7)$$

olsun. Her $i=1,2,\dots,k$ denemesi için $\sum_{i=1}^k \omega_i = 1$ olacak şekilde,

$$\omega_i = \frac{1/S_{\bar{Y}_i}^2}{\sum_{i=1}^k \left(1/S_{\bar{Y}_i}^2 \right)} \quad (8)$$

ağırlıkları hesaplınsın. Hesaplanan ω_i ağırlıkları ve \bar{Y}_i deneme ortalamaları kullanılarak

$$Y^+ = \sum_{i=1}^k \omega_i \bar{Y}_i \quad (9)$$

biçiminde tanımlanan bir ağırlıklı ortalama hesaplınsın. Y^+ ağırlıklı ortalama değeri, μ genel ortalamanın varyans ağırlıklı bir tahminidir. Bu değerler kullanılarak hesaplanan

$$t_i = \frac{\bar{Y}_i - Y^+}{S_{\bar{Y}_i}} \quad (10)$$

istatistiğinin dağılımı $\nu_i = n_i - 1$ serbestlik dereceli Student t dağılımına uyar. Bu aşamada, hesaplanan her bir t_i değerine, Bailey tarafından Student t dağılan rassal değişkenler üzerinde kullanılmak üzere önerilen

$$z_i = \pm \frac{4v_i^2 + \frac{5(2z_c^2 + 3)}{24}}{4v_i^2 + v_i + \frac{(4z_c^2 + 9)}{12}} v_i^{1/2} \left\{ \ln \left(1 + \frac{t_i^2}{v_i} \right) \right\}^{1/2} \sim N(0,1) \quad (11)$$

biçimindeki lokal tam normallik dönüşümü uygulanır [1]. Burada z_c değeri, standart normal dağılımda araştırmacının belirlediği anlam düzeyinin yarısına ($\alpha/2$) karşılık gelen kritik değere eşittir ($\alpha = 0.05$ ise $z_c = 1.96$). Dönüşümden sonra elde edilen z_i değerlerinin kareleri alınarak

$$B^2 = \sum_{i=1}^k z_i^2 = \sum_{i=1}^k \left(\frac{4v_i^2 + \frac{5(2z_c^2 + 3)}{24}}{4v_i^2 + v_i + \frac{(4z_c^2 + 9)}{12}} v_i^{1/2} \left\{ \ln \left(1 + \frac{\left(\frac{\bar{Y}_i - Y^+}{S_{\bar{Y}_i}} \right)^2}{v_i} \right) \right\}^{1/2} \right)^2 \quad (12)$$

biçiminde tanımlanan, dağılımı $k-1$ serbestlik dereceli χ^2 dağılımına yaklaşan ve tarafımızdan B^2 olarak isimlendirilen istatistik hesaplanır. Hesaplanan bu değer, $\chi_{k-1,1-\alpha}^2$ tablo değerini geçerse $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ hipotezi reddedilir.

3. B^2 testi yazılımı

Tek yönlü sabit etkili deney düzenlerinden elde edilen sonuçların çözümlenmesinde farklı varyanslılık sorununun giderilmesi amacıyla geliştirilmiş B^2 Testi'ne [8] ait test istatistiği hesaplanırken Bailey tarafından 1980 yılında önerilen lokal tam normallik dönüşümünden yararlanılır. İşlem yükünün hafifletilmesi için bazı tablolar üretilmişse de hesaplamaların bir bilgisayar yazılımı ile yapılmasının uygulamacılar açısından kolaylık sağlayacağı düşünülmüştür ve bu amaçla B^2 Testi yazılımı geliştirilmiştir. B^2 Testi yazılımı Delphi 7.0 uygulama geliştirme aracı ile geliştirilmiştir.

3.1. B^2 testi yazılımı veri girişi

Kullanıcılar grup ve tekrar sayılarını belirterek veri girişi yapabilecekleri gibi metin dosyalarında bulunan verileri de kullanabilirler. Yazılım çalıştırıldığında ilk olarak kullanıcıdan veri kaynağının seçilmesini ister. Eğer veriler bir dosyada kayıtlı ise "Dosyadan" seçeneği seçilerek "Tamam" butonuna basılır, böylece kullanıcılar verileri içeren csv uzantılı (veriler virgülle ayrılmış) dosyalarını, açılan form ile seçebilirler. Grup ve tekrar sayılarını belirterek veri girişi için "Veri Girişi" seçeneği işaretlenmelidir. Daha sonra aktif hale gelen "Grup Sayısı" ve "Tekrar Sayısı" alanlarına deneydeki grup sayısı ve tekrar sayısı yazılabilir. "Veri Girişi" seçeneğini işaretleyen kullanıcı "Grup Sayısı" ve "Tekrar Sayısı" alanlarını boş geçemez ve bu alanlara ikiden küçük bir sayı giremez. Bu alanlarda belirttiği grup ve tekrar sayılarına göre, verileri açılan ana form üzerindeki tabloya girebilir.

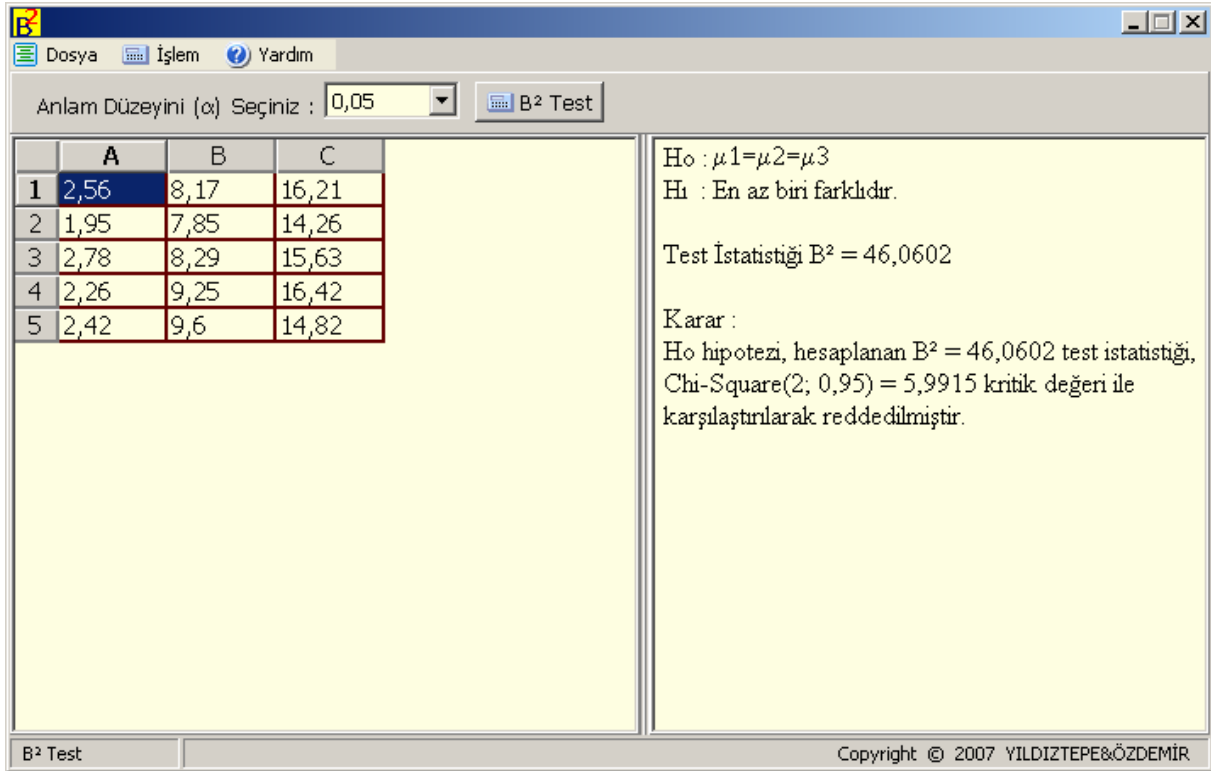
Yazılımdan, dengesiz deneylerden elde edilen verilerin çözümlenmesinde de yararlanılabilir. Bunun için veri girişinde tekrar sayısı olarak maksimum tekrar sayısı belirtilmeli ve tekrar sayısı eksik olan hücreler boş bırakılmalıdır.



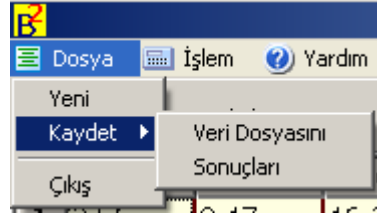
Şekil 1. Veri kaynağı seçimi

3.2. B^2 testi yazılımı kullanım

Veri girişi tamamlandıktan sonra kullanıcı test istatistiğini hesaplatırabilir. Testin anlam düzeyi formun üst kısmındaki liste kutusundan seçilebilir. Varsayılan değer 0,05 dir. Test sonuçlarını elde etmek için formun üst kısmındaki “ B^2 Test” butonu veya menüden “İşlem → Test değerini hesapla” seçeneği kullanılabilir.

Şekil 2. B^2 testi ana form

Formun sağ tarafında test istatistiği ve $\chi^2_{k-1,1-\alpha}$ değerine göre karşılaştırılarak verilen karar gösterilir. Testin sonuçları formun sağ tarafındaki editörden kopyalanabilir, metin dosyası olarak kaydedilebilir. Kullanıcı sadece sonuçları değil tabloya girdiği verileri de csv uzantılı dosya olarak kaydedebilir. B^2 Testi yazılımı <http://kisi.deu.edu.tr/userweb/engin.yildiztepe/BKare.rar> adresinden temin edilebilir.



Şekil 3. Kaydetme seçenekleri

4. Sonuç

B^2 Testi, kitle dağılımı normal olduğunda tek yönlü sabit etkili deney düzenlerinde, değişen varyanslılık nedeniyle gerçekleşen anlam düzeyinin nominal anlam düzeyinden sapsmasını kontrol edebilen bir testtir. Değişen varyanslılık problemi olmadığı durumlarda ise verdiği sonuçlar, F Testi sonuçları ile uyum göstermektedir. Burada hazırlanan yazılım, eşit ya da farklı deneme ve tekrar sayıları ile B^2 Testi'ni değişen varyanslılık durumunda seçenek bir test olarak kullanmak isteyen uygulamacılar için büyük kullanım kolaylığı sağlamıştır.

Kaynaklar

- [1] Bailey, B.J.R. (1980), Accurate normalizing transformations of Student's t Variate, *Applied Statistics*, 29, 304-306.
- [2] Box, G.E.P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, *Annals of Mathematical Statistics*, 25, 290-302.
- [3] Brown, M.B., & Forsythe, A.B. (1974). The small sample behavior of some statistics which test equality of several means, *Technometrics*, 16, 129-132.
- [4] Cantu, M. (2003). *Mastering Delphi 7*. CA:Sybex.
- [5] Clinch, J. J., & Keselman, H. J. (1982). Parametric alternatives to the analysis of variance, *Journal of Educational Statistics*, 7, 207-214.
- [6] Harwell, M.R., Rubinstein, E.N., Hayes, W.S., Olds, C.C. (1992). Summarizing Monte Carlo results in methodological research: The one and two factor fixed effects ANOVA cases, *Journal of Educational Statistics*, 17, 315-339.
- [7] Lix, L.M., Keselman, J.C., Keselman, H.J. (1996). Consequences of Assumption Violations Revisited: A Quantitative Review of Alternatives to the One-Way Analysis of Variance F Test, *Review of Educational Research*, 4, 579-619.
- [8] Özdemir, A.F. & Kurt, S. (2006). One Way Fixed Effect Analysis of Variance Under Variance Heterogeneity And A Solution Proposal, *Selçuk Journal Of Applied Mathematics*, 7, 81-91.
- [9] Rupert, G., & Miller, J.R.(1986). *Beyond ANOVA, basics of applied statistics*. Newyork: John Wiley & Sons. Inc.
- [10] Scheffé, H. (1959). *The Analysis of Variance*. Newyork: John Wiley & Sons.Inc.
- [11] Wilcox, R.R., Charlin, V.L., & Thompson, K.L. (1986). New Monte Carlo results on the robutness of the ANOVA F, W and F* statistics, *Communications in Statistics: Simulation and Computation*, 15, 933-943.