ANADOLU ÜNİVERSİTESİ

Bilim ve Teknoloji Dergisi A-Uygulamalı Bilimler ve Mühendislik
Cilt: 14 Sayı: 3 2013
Sayfa: 241-249

ARAŞTIRMA MAKALESİ /**RESEARCH ARTICLE**

A.Fırat ÖZDEMİR[1], Engin YILDIZTEPE[2]

BİR KONUM ÖLÇÜSÜ İLE k BAĞIMSIZ GRUBUN KARŞILAŞTIRILMASI İÇİN BİR TEST

ÖZ

k bağımsız grubu karşılaştırırken geleneksel olarak kullanılan tek-yönlü ANOVA F-testi'nin normal teori varsayımlarına bağlı olduğu bilinmektedir. Bu çalışmada budanmış ortalama ve bootstrap-t yöntemine dayalı yeni bir test önerilmiştir. Önerilen test, ANOVA F-testi ve onun parametrik olmayan karşılığı Kruskal-Wallis testi ile birinci tip hatanın korunması ölçütü kullanılarak orta büyüklükteki örneklem genişliklerinde Monte Carlo simülasyonları ile karşılaştırılmıştır. Önerilen test, normal teori varsayımlarının geçerli olduğu durumlarda diğer iki yöntem ile yakın sonuçlar verirken bu varsayımların bozulduğu deney düzenlerinde nominal anlam düzeyinin korunması bakımından daha başarılıdır.

Anahtar Kelimeler: Budanmış ortalama, Bootstrap-t, Gerçekleşen anlam düzeyi, k bağımsız grup

A TEST FOR COMPARING k INDEPENDENT GROUPS WITH A MEASURE OF LOCATION

ABSTRACT

It is well known that the conventional method for comparing k independent groups, one-way ANOVA F test, depends on normal theory assumptions. In this work, a new test is proposed which is based on trimmed means and a bootstrap t method. The proposed test, ANOVA F test and its rank-based nonparametric counterpart Kruskal-Wallis test were compared in terms of saving nominal Type I error via Monte Carlo simulations by using moderate sample sizes. It was found that the proposed test copes quite well under normality and homogeneity and it was much more preferable than those two traditional tests under non-normality and heterogeneity of variances.

Keywords: Trimmed mean, Bootstrap-t, Actual type I error, k independent groups

[1,] Dokuz Eylül Üniversitesi Fen Fakültesi İstatistik Bölümü, Tınaztepe Yerleşkesi, Buca, İzmir, Türkiye
Tel: 0 232 301 85 52, E-posta: firat.ozdemir@deu.edu.tr
[2,] Dokuz Eylül Üniversitesi Fen Fakültesi İstatistik Bölümü, Tınaztepe Yerleşkesi, Buca, İzmir, Türkiye
Tel: 0 232 301 86 04, E-posta: engin.yildiztepe@deu.edu.tr

Geliş: 28 Kasım 2012 **Düzeltme:** 11 Mart 2013 **Kabul:** 21 Haziran 2013

1. GİRİŞ

İkiden fazla sayıda grup karşılaştırılırken en çok kullanılan yöntem geleneksel ANOVA F-testidir. Ancak F-testinin diğer parametrik testler gibi normallik ve homojen varyanslılık varsayımlarının ihlal edilmesine karşı hassas bir test olduğu bilinmektedir. 0.05 nominal anlam düzeyi ve normallik varsayımı altında, her birinde 50 gözlem olan dört grup karşılaştırılırken homojen varyanslılığın ihlal edilmesi, gerçekleşen anlam düzeyinin 0.09 olmasına sebep olabilmiştir (Wilcox, vd., 1986). Tasarımın dengesiz olması, heterojen varyanslılığın etkisini daha da artırır. Özdemir ve Kurt, 2007 yılında yaptıkları çalışmada gerçekleşen anlam düzeyi olarak 0.24 gibi bir değer ile karşılaşmışlardır. Heterojen varyanslılık ve tasarımın dengesiz olmasına ek olarak, normallik varsayımından da uzaklaşmak, araştırmacının belirlediği nominal anlam düzeyinin çok büyük oranlarda bozulmasına ve testin gücünün azalmasına neden olur.

Dayanıklı (robust) istatistik, varsayım ihlalleri sonucunda sıklıkla karşılaşılan aykırı değerlere karşı hassas olmayan dayanıklı konum ve değişim ölçülerinin kullanımı yolu ile nominal anlam düzeyinden sapma ve düşük güç sorununa çözüm getirmeye çalışır. Literatürde en çok karşılaşılan dayanıklı konum ölçülerinden birisi standart hatası aykırı değerlerden daha az etkilendiği bilinen budanmış ortalamadır (Staudte ve Sheather, 1990).

Özdemir ve Kurt, 2007 yılında yaptıkları çalışmada k bağımsız grubu karşılaştırmak için normal dağılım varsayımı altında kullanılan ve ki-kare dağılan bir test istatistiğine sahip bir yöntem önermişlerdir. Bu çalışmada ise 2007 yılındaki çalışmayı temel alan ancak kitle budanmış ortalamalarını karşılaştıran ve bootstrap örnekleme yaklaşımıyla belirlenen bir karar kuralına sahip yeni bir yöntem önerilmiştir. Yöntemin uygulanabilmesi için normallik varsayımına ihtiyaç yoktur. İkinci bölümde önerilen yöntem anlatılmıştır. Üçüncü bölümde yapılan simulasyon çalışması ve sonuçları, son bölümde ise bulunan sonuçlara ait yorumlar yer almaktadır.

2. BUDANMIŞ ORTALAMA ve BOOTSTRAP–T İLE BİR TEST (B_{tk}^2)

Önerilen yöntem, kitle budanmış ortalamalarının eşitliği $, H_0 : \mu_{t1} = \mu_{t2} = ... = \mu_{tk} = \mu_t$, hipotezini test etmektedir. $i = 1,...,n_j$; $j = 1,2,...,k$ olmak üzere $X_{1j},...,X_{n_j j}$ rassal örneklemler olsun. $\overline{X}_{tj}$, $\overline{Y}_{wj}$ ve s_{wj}^2 sırasıyla örneklem budanmış ve winsorized ortalamaları ile winsorized varyansını gösterir ve bu istatistikler aşağıda verildiği gibi hesaplanır.

$X_{(1)j} \le X_{(2)j} \le ... \le X_{(n_j)j}$ sıralanmış örneklem, γ budama yüzdesi (çalışmada %20 alınmıştır, (Wilcox, 2012)), g_j=[γn_j] her iki kuyruktan simetrik olarak budanacak gözlem sayısı ve $h_j = n_j - 2g_j$ olmak üzere örneklem budanmış ortalaması,

$$\overline{X}_{tj} = \frac{1}{h_j} \sum_{i=g_j+1}^{n_j-g_j} X_{(i)j} \tag{1}$$

winsorized ortalama,

$$\overline{Y}_{wj} = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_{ij} \tag{2}$$

ve winsorized varyans ise,

$$s_{wj}^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \left(Y_{ij} - \overline{Y}_{wj}\right)^2 \tag{3}$$

eşitlikleri ile hesaplanırlar. Son iki formülde

$$\begin{aligned} Y_{ij} &= X_{(g_j+1)j} \quad if \quad X_{ij} \le X_{(g_j+1)j} \\ &= X_{ij} \qquad if \quad X_{(g_j+1)j} < X_{ij} < X_{(n_j-g_j)j} \\ &= X_{(n_j-g_j)j} \quad if \quad X_{ij} \ge X_{(n_j-g_j)j} \end{aligned}$$

olarak kullanılır. Örneklem budanmış ortalamasının standart hatasının karesi,

$$s_{\overline{X}_{tj}}^2 = \frac{s_{wj}}{(1-2\gamma)\sqrt{n_j}} \tag{4}$$

biçimindedir (Wilcox, 2012). Burada her bir $(j = 1,2,...,k)$ grup için

$$\omega_j = \frac{\dfrac{1}{s_{\overline{X}_{tj}}^2}}{\dfrac{1}{s_{\overline{X}_{t1}}^2} + \dfrac{1}{s_{\overline{X}_{t2}}^2} + \ldots + \dfrac{1}{s_{\overline{X}_{tk}}^2}} \tag{5}$$

ağırlıkları hesaplanır. Bu ağırlıklar kullanılarak $X^+ = \omega_1 \overline{X}_{t1} + \omega_2 \overline{X}_{t2} + \ldots + \omega_k \overline{X}_{tk}$ ağırlıklı ortalaması ve her bir grup için

$$T_{tj} = \frac{\overline{X}_{tj} - X^+}{s_{\overline{X}_{tj}}} \tag{6}$$

istatistikleri hesaplanır. Her T_{tj} değerine

$$z_{tj} = \pm \frac{4\nu_j^2 + \dfrac{5(2z_c^2 + 3)}{24}}{4\nu_j^2 + \nu_j + \dfrac{(4z_c^2 + 9)}{12}} \nu_j^{1/2} \left\{ \ln\left(1 + \frac{T_{tj}^2}{\nu_j}\right) \right\}^{1/2} \sim N(0,1) \tag{7}$$

biçiminde verilen Bailey'in normallik dönüşümü uygulanır (Bailey, 1980). Burada $\nu_j = h_j - 1$ ve $h_j = n_j - 2g_j$ dir. Önerilen yönteme ait test istatistiği,

$$B_{tk}^2 = \sum_{j=1}^{2} z_{tj}^2 = \sum_{j=1}^{2} \left(\frac{4\nu_j^2 + \frac{5(2z_c^2+3)}{24}}{4\nu_j^2 + \nu_j + \frac{(4z_c^2+9)}{12}} \nu_j^{1/2} \left\{ \ln\left(1 + \frac{\left(\frac{\overline{X}_{tj} - X^+}{S_{\overline{X}_{tj}}}\right)^2}{\nu_j} \right) \right\}^{1/2} \right)^2 \tag{8}$$

biçimindedir.

Bu aşamada H_0 ile ilgili karar verebilmek için bootstrap-t yaklaşımından yararlanılmıştır;

1. n_j büyüklüğündeki $X_{1j}^*,...,X_{n_j j}^*$ bootstrap örneklem, orijinal $X_{1j},...,X_{n_j j}$ örnekleminden tekrarlı olarak seçilir ve $C_{ij}^* = X_{ij}^* - \overline{X}_{tj}$ $i = 1,...,n_j$; $j = 1,2,...,k$ hesaplanır.

2. Hesaplanan $C_{ij}^* = X_{ij}^* - \overline{X}_{tj}$'ler ile $\overline{X}_{tj}^*$, $\overline{Y}_{wj}^*$, s_{wj}^{*2}, $s_{\overline{X}_{tj}}^{*2}$, ω_j^*, X^{*+}, T_{tj}^* ve $B_{tk}^{*\,2}$ değerleri hesaplanır.

3. İlk iki basamak B=599 kez tekrar edilir ve böylece bootstrap örneklemlere dayanan B adet test istatistiği $B_{tk1}^{*\,2},...,B_{tkB}^{*\,2}$ oluşur.

4. 599 adet $B_{tk1}^{*\,2},...,B_{tkB}^{*\,2}$ değeri $B_{tk(1)}^{*2} \leq ... \leq B_{tk(B)}^{*2}$ biçiminde küçükten büyüğe doğru sıralanır.

5. En yakın tamsayıya yuvarlanacak şekilde $a = (1-\alpha)B$ değeri hesaplanır ve $B_{tk(a)}^{*2}$ bulunur.

6. Orijinal $X_{1j},...,X_{n_j j}$ örneklemi ile hesaplanan $B_{tk}^2 > B_{tk(a)}^{*2}$ ise H_0 reddedilir.

3. SİMÜLASYON ÇALIŞMASININ TASARIMI

Bu bölümde önerilen test, ANOVA-F testi ve Kruskal-Wallis testi ile dengeli ve dengesiz deney düzenlerinde gerçekleşen I.tip hata yapma olasılığı açısından karşılaştırılmıştır. R programlama dilinde yazılan kodlarla, k=3 ve 6 grup için dört teorik dağılımdan üretilen veriler ve üç gerçek veri setinden yararlanılarak, her düzen ile 10000 tekrar olacak biçimde gerçekleşen anlam düzeyleri hesaplanmıştır. Çalışmada kullanılan teorik dağılımlar normal dağılım, farklı parametre değerleri ile iki g&h dağılımı (Hoaglin, 1985) ve bozulmuş normal dağılım, gerçek veriler ise http://www.freewebs.com/tedstats adresinden alınan farklı karakterdeki üç veri setidir.

g&h dağılımı, çarpıklığın kontrol edilebileceği g ve basıklığın kontrol edilebileceği h parametreleri ile normal dağılımdan uzaklaşmayı derecelendirerek gözleme imkanı veren ve g=h=0 için standart normal dağılıma eşdeğer olan bir dağılımdır. Standart normal dağılımdan üretilen bir Z değeri ile

g≠0 için, $$X = \frac{(\exp(gZ)-1)\exp\left(\frac{hZ^2}{2}\right)}{g} \tag{9}$$

g=0 için, $$X = Z\exp\left(\frac{hZ^2}{2}\right) \tag{10}$$

dönüşümleri ile bu dağılımdan veri üretilir. Çalışmada kullanılan g ve h değerleri ile elde edilen dağılımlara ait çarpıklık ve basıklık değerleri Tablo I'de verilmiştir.

Tablo 1. Kullanılan g&h dağılımlarına ait çarpıklık ve basıklık değerleri

g	**h**	$\hat{\kappa}_1$	$\hat{\kappa}_2$
0	0	0	3
0.5	0.2	10.554	200.736
0.5	0.4	76.877	14091.46

Bozulmuş normal dağılım, ortalamaları aynı standart sapmaları farklı iki normal dağılımdan elde edilir. Bozulmuş normal dağılımdan veri üretmek için $0 < \varepsilon < 1$ olmak üzere veriler, $1-\varepsilon$ olasılıkla örneğin standart normal dağılımdan ve ε olasılıkla ortalaması 0, standart sapması K olan bir normal dağılımdan türetilir. Dağılım fonksiyonları kullanılarak

$$H(x) = (1-\varepsilon)\Phi(x) + \varepsilon\,\Phi\left(x/K\right) \tag{11}$$

biçiminde ifade edilen bozulmuş normal dağılım simetrik ve ağır kuyruklu bir dağılımdır (Wilcox, 2012). Bu çalışmada ε= 0.1 ve K=10 olarak alınmıştır.

Simülasyon çalışmasında kullanılan gerçek verilere ilişkin bazı istatistikler Tablo II'de verildiği gibidir.

Tablo 2. Kullanılan gerçek verilere ait istatistikler

	N	Ortalama	Medyan	Standart Sapma	IQR	Çarpıklık
Carnegie Verisi						
Lisans FTE	4607	2858.95	1145.0	4617.55	2828	3.35
Standart Testler						
SAT Sayısal	15168	545.33	540.0	89.49	130	0.026
Enstitü Verisi						
Toplam Süre	20154	52.83	46.0	34.12	45	0.94

Bradley, 1978'de önerdiği tutucu dayanıklılık ölçütünde, varsayım ihlali durumunda nominal anlam düzeyi α=0.05 için gerçekleşen anlam düzeyi 0.045 ile 0.055 arasında kalırsa, araştırmacı tarafından belirlenen anlam düzeyinin korunduğunun söylenebileceğini belirtmiştir. Aşağıda verilen tablolarda bu aralığın dışında kalan gerçekleşen anlam düzeyleri kalın ve vurgulu yazılmıştır.

4. SİMÜLASYON ÇALIŞMASININ SONUÇLARI

Tablo 3. Kitle dağılımları Normal iken gerçekleşen anlam düzeyleri.

Gruplar	n	σ	F	KW	B_{tk}^2
	25,25,25	1,1,1	**0.0398**	0.0494	0.0488
	25,25,25	1,2,4	**0.0838**	**0.0676**	0.0510
k=3	25,30,35	1,1,1	**0.0387**	0.0475	0.0485
	25,30,35	1,2,4	**0.0687**	**0.0564**	0.0520
	25,30,35	4,2,1	**0.1109**	**0.0830**	0.0510
	25,25,25,25,25,25	1,1,1,1,1,1	**0.0383**	0.0457	0.0476
	25,25,25,25,25,25	1,2,4,1,2,4	**0.0855**	**0.0662**	0.0459
k=6	25,30,35,25,30,35	1,1,1,1,1,1	**0.0443**	0.0496	0.0477
	25,30,35,25,30,35	1,2,4,1,2,4	**0.0583**	0.0522	0.0492
	25,30,35,25,30,35	4,2,1,4,2,1	**0.1089**	**0.0842**	0.0457

Tablo 3'de Normal dağılan kitlelerden alınan sonuçlar verilmektedir. Bu tabloda görülen nominal değerlerden sapmaların sebebi heterojen varyanslılıktır. Kruskal-Wallis testinin gerçekleşen anlam düzeyi toplam 10 deney düzeninin 5'inde Bradley'in belirlediği sınırlar dışında kalmıştır. Bu sayı F-testi için 4, B_{tk}^2 testi için 0 dır.

Tablo 4. Kitle dağılımları g=0.5, h=0.2 iken gerçekleşen anlam düzeyleri.

Gruplar	n	σ	F	KW	B_{tk}^2
	25,25,25	1,1,1	**0.0398**	0.0494	0.0488
	25,25,25	1,2,4	**0.0838**	**0.0676**	0.0510
k=3	25,30,35	1,1,1	**0.0387**	0.0475	0.0485
	25,30,35	1,2,4	**0.0687**	**0.0564**	0.0520
	25,30,35	4,2,1	**0.1109**	**0.0830**	0.0510
	25,25,25,25,25,25	1,1,1,1,1,1	**0.0383**	0.0457	0.0476
	25,25,25,25,25,25	1,2,4,1,2,4	**0.0855**	**0.0662**	0.0459
k=6	25,30,35,25,30,35	1,1,1,1,1,1	**0.0443**	0.0496	0.0477
	25,30,35,25,30,35	1,2,4,1,2,4	**0.0583**	0.0522	0.0492
	25,30,35,25,30,35	4,2,1,4,2,1	**0.1089**	**0.0842**	0.0457

Tablo IV, kitle dağılımları g=0.5, h=0.2 iken elde edilen sonuçları vermektedir. Heterojen varyanslılık ile beraber normal dağılım varsayımı bozulduğunda F-testi tüm deney düzenlerinde gerçekleşen anlam düzeylerini Bradley'in limitleri dışında vermiştir. Kruskal-Wallis testine ait 5 sonucun limit dışı olduğu görülürken bu sayı B_{tk}^2 testi için sıfırdır.

Tablo 5. Kitle dağılımları g=0.5, h=0.4 iken gerçekleşen anlam düzeyleri.

Gruplar	n	σ	F	KW	B_{tk}^2
	25,25,25	1,1,1	**0.0332**	0.0493	0.0479
	25,25,25	1,2,4	**0.0874**	**0.0608**	0.0460
k=3	25,30,35	1,1,1	**0.0360**	0.0501	0.0497
	25,30,35	1,2,4	**0.0634**	0.0512	0.0492
	25,30,35	4,2,1	**0.1153**	**0.0743**	0.0458
	25,25,25,25,25,25	1,1,1,1,1,1	**0.0332**	0.0489	**0.0441**
	25,25,25,25,25,25	1,2,4,1,2,4	**0.0775**	**0.0655**	**0.0405**
k=6	25,30,35,25,30,35	1,1,1,1,1,1	**0.0383**	0.0522	0.0462
	25,30,35,25,30,35	1,2,4,1,2,4	**0.0620**	0.0529	0.0465
	25,30,35,25,30,35	4,2,1,4,2,1	**0.1071**	**0.0779**	0.0453

Tablo 5, kitle dağılımları g=0.5, h=0.4 iken elde edilen sonuçları vermektedir. Dağılımın çarpıklık ve basıklığının daha da arttığı, normallikten uzaklaşmanın daha etkili olduğu bu tabloda F-testine ait tüm sonuçlar Bradley'in limitleri dışında yer alırken Kruskal-Wallis testine ait 4, B_{tk}^2 testine ait 2 sonuç limitler dışında kalmıştır.

Tablo 6. Kitle dağılımları CN(0.1,10) iken gerçekleşen anlam düzeyleri.

Gruplar	n	F	KW	B_{tk}^2
k=3	25,25,25	**0.0340**	0.0483	0.0514
	25,30,35	**0.0389**	**0.0442**	0.0503
k=6	25,25,25,25,25,25	**0.0403**	0.0463	0.0455
	25,30,35,25,30,35	**0.0427**	**0.0445**	0.0473

Tablo 6, kitle dağılımları simetrik ağır kuyruklu bir dağılım olan CN(0.1,10) iken elde edilen sonuçları vermektedir. F-testinin gerçekleşen anlam düzeyleri tüm deney düzenleri için Bradley'in limitleri dışındadır. Kruskal-Wallis testi için iki defa limit aşımı gözlenirken, B_{tk}^2 testi için hiç limit aşımı gözlenmemiştir.

Tablo 7. Kitle dağılımları Lisans FTE isimli gerçek veri seti iken gerçekleşen anlam düzeyleri

Gruplar	n	F	KW	B_{tk}^2
k=3	25,25,25	**0.0409**	0.0495	**0.0325**
	25,30,35	**0.0427**	0.0476	**0.0354**
k=6	25,25,25,25,25,25	**0.0411**	**0.0443**	**0.0225**
	25,30,35,25,30,35	**0.0402**	**0.0440**	**0.0285**

Tablo 7, kitle dağılımları sağdan çarpık ve değişimin çok yüksek olduğu Lisans FTE isimli gerçek veri seti iken elde edilen sonuçları vermektedir. F-testi ve önerilen B_{tk}^2 testi tüm deney düzenleri için Bradley limitlerinin dışında sonuçlar verirken Kruskal-Wallis testine ait sonuçlar 2 düzen için limitlerin dışında kalmıştır.

Tablo 8. Kitle dağılımları SAT sayısal isimli gerçek veri seti iken gerçekleşen anlam düzeyleri.

Gruplar	n	F	KW	B_{tk}^2
k=3	25,25,25	0.0465	0.0463	0.0476
	25,30,35	0.0462	0.0470	0.0497
k=6	25,25,25,25,25,25	0.0490	0.0455	0.0452
	25,30,35,25,30,35	0.0515	0.0495	0.0465

Tablo 8, kitle SAT Sayısal isimli gerçek veri seti iken elde edilen sonuçları vermektedir. Tüm testlere ait sonuçlar Bradley'in limitleri içindedir.

Tablo 9. Kitle Toplam Saatler isimli gerçek veri seti iken gerçekleşen anlam düzeyleri.

Gruplar	n	F	KW	B_{tk}^2
k=3	25,25,25	0.0511	0.0484	0.0455
	25,30,35	0.0509	0.0510	0.0485
k=6	25,25,25,25,25,25	0.0515	0.0508	0.0453
	25,30,35,25,30,35	0.0506	0.0497	0.0467

Tablo 9, kitle dağılımları Toplam Saatler isimli gerçek veri seti iken elde edilen sonuçları vermektedir. Tüm testlere ait sonuçlar Bradley'in limitleri içindedir.

5. SONUÇLAR

k bağımsız grubu karşılaştırmak için yeni önerilen test ve geleneksel olarak kullanılan iki test I. tip hatanın korunması ölçütü ile farklı kitlelerden çekilen örneklemlerle karşılaştırılmıştır. Normal dağılım ve homojen varyanslılık varsayımlarının sağlandığı durumlarda şüphesiz ki F-testi en uygun seçenektir. Ancak bu varsayımların sağlanmadığı durumlarla uygulamada sıklıkla karşılaşılmaktadır. Bu durumlarda F testinin kullanılması, göze alınanın çok ötesinde hatalı karar verme olasılığı ile karşı karşıya kalmak anlamına gelmektedir. Tasarımın dengesiz olması varsayım ihlallerinin F-testi üzerindeki etkisini daha da artırmaktadır.

Sadece heterojen varyanslılık altında F-testi ve Kruskal-Wallis testlerindeki limit dışı anlam düzeyi sayıları benzer iken heterojen varyanslılığın yanına g=0.5, h=0.2 ile normalliğin asimetri ve aykırı değerler ile bozulması problemi de eklenince F-testi tüm düzenler için limit dışı değerler vermiş, Kruskal-Wallis testi'nde F-testi kadar olmasa da limit dışı değerler görülmüş, B_{tk}^2 testi için ise tüm gerçekleşen anlam düzeyleri Bradley'in limitleri dahilinde çıkmıştır.

g=0.5 ve h=0.4 ile normalliğin daha ileri derecede bozulması B_{tk}^2 testinde de iki adet limit dışı anlam düzeyi ile karşılaşılmasına sebep olmuştur.

Simetrik ve ağır kuyruklu dağılımların etkisinin görülmesi amacıyla çalışmaya dahil edilen bozulmuş normal dağılımdan veri çekilmesi durumunda Kruskal-Wallis testi kısmen etkilenmiş ancak F-testine ait tüm sonuçlar yine limit dışı olmuştur. Önerilen B_{tk}^2 testi bu durumda nominal anlam düzeyini koruyabilmiştir.

Güçlü bir asimetrik yapı gösteren Lisans FTE isimli veri setinde B_{tk}^2 testi F-testi gibi nominal anlam düzeyini koruyamamıştır. Bu veri setinin kullanıldığı deney düzenleri için her iki teste ait tüm gerçekleşen anlam düzeyleri nominal değerin altında kalmıştır. Başka bir ifade ile her iki testte tutucu eğilim göstermiştir. Diğer iki simetrik gerçek veri setlerinde ise tüm testler araştırmacının belirlediği nominal anlam düzeyini koruyabilmişlerdir.

F-testi ve Kruskal-Wallis testi herhangi bir varsayım ihlali durumunda I.tip hata yapma olasılığını koruyamadıkları için testler arasında güç karşılaştırması yapılmamıştır.

Bu çalışmada, F-testinin normal teori varsayımlarına karşı hassaslığı bir kez daha gözlenmiştir. Literatürde ilk alternatif olarak düşünülen Kruskal-Wallis testinin de güvenilir bir seçenek olmayacağına dair sonuçlar bulunmuştur. Önerilen B_{tk}^2 testi teorik ve gerçek verilerin kullanıldığı toplam 46 deney düzeninin 40'nda, Kruskal-Wallis testi 28'inde, F-testi ise 14'ünde araştırmacı tarafından belirlenen nominal anlam düzeyini Bradley'in belirlediği limitler içinde korumayı başarmıştır. Çalışmaya konu olan test, normal teori varsayımlarının bozulduğu ve orta büyüklükteki örneklem genişlikleri ile çalışıldığı durumlarda Anova F-testine seçenek olarak düşünülebilir.

KAYNAKLAR

Bailey, B.J.R. (1980). Accurate Normalizing Transformations of Student's t Variate. *Applied Statistics,* 29*,* 304-306.

Bradley, J.V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31, 144-152.

Hoaglin, D.C. (1985). Summerazing Shape Numerically: The g and h Distributions. In D. Hoaglin, F. Mosteller, & J. Tukey (Eds.), *Exploring Data Tables, Trends, and Shapes*. 461-513. New York: Wiley.

Özdemir A.F., Kurt S. (2007). Tek Yönlü Özel Seçimli Varyans Çözümlemesinde Farklı Varyanslılık Sorunu ve Bir Çözüm Önerisi. *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, 8 (1), 93-102.

Staudte, R.G., and Sheater, S.J. (1990). *Robust Estimation and Testing*. Newyork: Wiley.

Ted' Micceri's Web Site.
(http://www.freewebs.com/tedstats. Alıntı tarihi: 01.10.2012.)

Wilcox, R.R. (2012). *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier Academic Press, Third Edition.

Wilcox, R.R., Charlin, V., Thompson, K.L. (1986). New Monte Carlo Results on The Robustness of The ANOVA F, W and F* Statistics. *Communications in Statistics: Simulations and Computations*, 15, 933-944.