Volume: 07 Number: 02 Category: 02 Page: 45-57 ISSN No: 1303-6319 Cilt: 07 Sayı: 02 Kategori: 02 Sayfa: 45-57

SAĞA ÇARPIK DAĞILIM ORTALAMALARI İÇİN BAZI TESTLERİN KULLANIMI VE KARŞILAŞTIRMALARI

Emre Erçin SARISOY* **Hamza GAMGAM****

ÖZET

Bu makalenin amacı, sağa çarpık dağılımlardan çekilen örneklemler ile kitle ortalamasına ilişkin yokluk hipotezini test etmek için kullanılan Johnson'ın düzeltilmiş karesel t testini, Sutton'ın bileşik testini ve Chen'in testini tanıtmak, ayrıca bunları 1. tip hata oranları ve testin gücü bakımından karşılaştırmaktır. Bu amaçla bazı sağa çarpık dağılımlardan üretilen veri kümeleri üzerinden simülasyon çalışması yapılmış ve Chen'in testinin daha güçlü olduğu tespit edilmiştir.

Anahtar Kelimeler: Chen'in testi, Çarpıklık, Johnson'ın düzeltilmiş karesel t testi, Sutton'ın bileşik testi, Testin gücü.

1. GİRİŞ

Kitle ortalamasına ilişkin yokluk hipotezinin test edilmesinde kullanılan klasik yöntemler Z ve Student t testleridir. Bu test istatistiklerini kullanmak için örneklemin normal dağılıma sahip bir kitleden geldiği varsayımı geçerli olmalıdır.

Merkezi Limit Teoremi'ne göre normal dağılıma sahip olmayan bir kitleden rasgele örneklem seçildiğinde, örneklem büyüklüğü (n) yeteri kadar büyük iken örneklem ortalamasının dağılımı μ ortalama ve σ^2/n varyans ile normal dağılıma yakınsar. Dolayısıyla, dağılımı ne olursa olsun herhangi bir kitleden seçilen örneklem büyüklüğü yeterince büyük olan örneklem ile kitle ortalamasına ilişkin yokluk hipotezinin testinde $Z \sim N(0,1)$ istatistiği kullanılır. Eğer örneklemin geldiği kitle normal dağılıma sahip ve kitle varyansı (σ^2) bilinmiyor ise, bu istatistik yerine Student (1908)'ın önermiş olduğu $t \sim t_{n-1}$ istatistiği kullanılır. Örnek hacminin küçük ve örneklemin çekildiği kitlenin dağılımı çarpık iken, kitle ortalamasına ilişkin yokluk hipotezinin testinde bu test istatistiklerini kullanmak doğru olmaz.

Örnek hacmi yeteri kadar büyük değilken, normal dağılıma sahip olmayan kitleler için ortalamaya ilişkin hipotez testleri ve güven aralıkları hep tartışma konusu olmuştur. Sophister (1928), Neyman ve Pearson (1928), Nair (1941) çarpıklığın basıklığa göre t istatistiğini daha çok etkilediğini göstermişlerdir. Rider (1929), Geary (1936), Laderman (1939), Perlo (1933) ve Anscombe (1950)'un çarpık dağılımlardan çekilen örneklemler ile kitle ortalamasına ilişkin yokluk hipotezini test etmeye yönelik çalışmaları, t istatistiğinde bir düzeltme yapmaya yöneliktir (Johnson, 1978).

Box ve Andersen (1955) parametrik olmayan yöntemlerin kullanımına dayanan çalışmalar yapmışlardır. Ayrıca, Tukey (1964), Andrews vd. (1972) ve Yuen (1974)'in budama yöntemine dayanan çalışmaları vardır (Johnson, 1978).

* TÜİK Uzmanı, Türkiye İstatistik Kurumu, Ankara, e-posta: emresarisoy@tuik.gov.tr
**Prof. Dr., Gazi Üniversitesi, GÜFEF, İstatistik Bölümü, Ankara, e-posta: gamgam@gazi.edu.tr

Kitle normal dağılıma sahip değil ve örnek hacmi küçük iken Bartlett (1935), Chung (1946) ve Gayen (1949) Edgeworth ve Gram-Charlier açılımlarını kullanarak çeşitli düzeltilmiş t istatistikleri tanımlamışlardır.

Bu çalışmanın ikinci bölümünde sağa çarpık dağılımların ortalaması ile ilgili yokluk hipotezlerinin testi için Johnson'ın düzeltilmiş karesel t testi, Sutton'ın bileşik testi ve Chen'in testi tanıtılacak ve bu testlerin "deneysel 1. tip hata oranları" ve "testin gücü değerleri" bakımından karşılaştırılması için yapılan simülasyon çalışması hakkında bilgi verilecektir. Üçüncü bölümde ise simülasyon çalışması sonucunda elde edilen bulgulara yer verilecektir. Dördüncü bölümde, çalışmanın genel sonuçları değerlendirilecek ve bazı önerilerde bulunulacaktır.

2. YÖNTEM

2.1 Johnson'ın Düzeltilmiş Karesel t Testi

Bir X rastgele değişkeni için Cornish ve Fisher (1937) Açılımı, Eşitlik 1'de verilen biçimde ifade edilir:

$$CF(X) = \mu + \sigma\zeta + (\mu_3 / \sigma^2)(\mu_3\zeta^2 - 1) + ... \tag{1}$$

Burada ζ standart normal rastgele değişkeni, μ parametresi X rastgele değişkeninin ortalamasını, σ^2 bu değişkenin varyansını ve μ_3 ise bu değişkenin üçüncü merkezsel momentini ifade etmektedir. Johnson'ın düzeltilmiş karesel t istatistiğin türetilmesi Wallace (1958)'ın CF yaklaşımı ile ilişkilidir (Johnson, 1978).

Johnson (1978)'ın çalışmasında Eşitlik 1'deki tüm momentler ele alınmamıştır ve seri belirli bir terimden sonra devam ettirilmemiştir. Eşitlik 1 ifadesi $\bar{X}$'ya göre düzenlenirse;

$$CF(\bar{X}) = \mu + {}^{\sigma}\!/\!{}_{\sqrt{n}}\,\zeta + ({}^{\mu_3}\!/\!{}_{6n\sigma^2})(\zeta^2 - 1) + o(n^{-\frac{3}{2}}) \tag{2}$$

eşitliği elde edilir (Johnson, 1978). Burada $o(n^{-\frac{3}{2}})$ terimi, ilk iki terimden sonraki terimlerin $n^{\frac{3}{2}}$ terimine bölümü ile elde edilen serinin sıfıra yakınsadığını göstermektedir. Eşitlik 2'deki $CF(\bar{X})$ ifadesi, Cornish-Fisher açılımının iki terimli bir gösterimidir.

Johnson (1978), $\bar{X}$'nın momentlerinden faydalanarak yeni bir test istatistiği türetmiştir. Bunun için ilk adım olarak Eşitlik 1'den Eşitlik 2 elde edilmiştir. t istatistiğine ilişkin düzeltmeyi türetmek için bir başka ifade ile, düzeltilmiş t istatistiğini elde etmek için bu açılımdan $\zeta^2 - 1$ terimi çıkartılmıştır. Daha sonraki terimlerin çarpıklığı azaltmadaki etkisi çok az olduğundan, bu terimler etkisiz olarak kabul edilmiştir. Bu şekilde elde edilen t_1 değişkeni Eşitlik 3'teki gibi ifade edilebilir (Johnson, 1978).

$$t_1 = \left\{ (\bar{X}-\mu) + \lambda + \gamma \left[(\bar{X}-\mu)^2 - (\sigma^2/n) \right] \right\} (\sigma^2/n)^{-1/2} \tag{3}$$

Burada λ, n'nin bir fonksiyonu ve γ momentsel bir orandır ve $\gamma = \mu_3 / 3\sigma^4$, $\lambda = \mu_3 / 2n\sigma^2$ olarak tanımlanır. Buradan

$$t_1 = \left[(\bar{X}-\mu) + \frac{\mu_3}{6\sigma^2 n} + \frac{\mu_3}{3\sigma^4}(\bar{X}-\mu)^2 \right] \left[\frac{\sigma^2}{n} \right]^{-1/2} \tag{4}$$

olarak elde edilir. Bu istatistik n–1 serbestlik dereceli t dağılımına sahiptir.

t_1 istatistiğine ilişkin yan, payındaki $\mu_3 / 2\sigma^2 n$ ve 3. merkezsel momenti olan $4\mu_3 / n$ terimine bağlıdır. t_1 istatistiğinin payı ile paydası arasındaki kovaryans $4\mu_3 / n^2$'den daha küçüktür. Aynı zamanda bu istatistiğin γ ve λ terimlerini içermesi, çarpıklıktan kaynaklanan yüksek yanlılığın etkisini azaltmaktadır. Uygulamada μ_3 ve σ^2 bilinmediğinden bu parametreler örnekten tahmin edilir. Buna göre $H_0 : \mu = \mu_0$ yokluk hipotezinin, $H_1 : \mu > \mu_0$ seçenek hipotezine karşı testinde

$$\left[(\bar{X}-\mu_0) + \frac{\hat{\mu}_3}{6S^2 n} + \frac{\hat{\mu}_3}{3S^4}(\bar{X}-\mu_0)^2 \right] > t_{n-1,\alpha} \frac{S}{\sqrt{n}} \tag{5}$$

ise 1-α güven düzeyinde H_0 red edilir. Burada, $t_{n-1,\alpha}$ değeri n-1 serbestlik dereceli t dağılımından elde edilen kritik değerdir.

t_1 istatistiğinin payı doğrusal olmadığından güven aralığı için basit bir tanımlama yapılamaz. Ama γ ve λ yan miktarını ve çarpıklığın etkisini azalttığından dolayı karesel yapıyı içermeyen yeni bir istatistik tanımlanabilir. Johnson (1978), kitle ortalamasına ilişkin güven aralığını elde edebilmek için $(\bar{X}-\mu_0)^2$ terimini seriden çıkarmıştır. Buna göre, 1-α güven düzeyinde kitle ortalamasına ilişkin güven aralığı

$$\left[(\bar{X} + \frac{\hat{\mu}_3}{6S^2 n}) \right] \pm t_{n-1,\alpha/2} S / \sqrt{n} \tag{6}$$

ile bulunabilir (Johnson, 1978).

Düzeltilmiş güven aralığı çok kullanışlıdır. Aralık bu düzeltmeye göre elde ediliyorsa, aralığın genişliği çok artmaz. Ayrıca, çarpık dağılımlarda medyan ve ortalama arasında belirli bir farkın olmasından dolayı CF açılımlarının kullanılarak bu farkın minimuma indirilmesi ve bu şekilde güven aralığının oluşturulması daha iyi sonuçlar elde edilmesini sağlar (Johnson, 1978).

2.2 Sutton'ın Bileşik Testi

Kitle ortalamasına ilişkin yokluk hipotezini test etmek için varyansı ve üçüncü merkezsel momenti bilinmeyen bir kitleden, örneklem büyüklüğü n olan rastgele örneklemin seçildiğini varsayalım. Bu örneklemden, örneklem ortalaması ve örneklem varyansı hesaplansın. Johnson (1978) üçüncü merkezsel momenti sıfırdan farklı bir değer varsayarak n-1 serbestlik dereceli Student t dağılımına sahip t_1 istatistiğini tanımlamıştır. Çarpıklık çok büyük olmadığında, bu test istatistiğinin Student t istatistiğine göre daha güçlü olduğu, Sutton (1993)'ın yaptığı simülasyon çalışmaları ile gösterilmiştir. Ayrıca t_1 testinin kitlenin dağılımı sağa çarpık iken daha etkili olduğu Sutton (1993) tarafından vurgulanmıştır.

Sutton (1993) Üstel, Gamma, Ki-Kare, Weibull ve Lognormal gibi dağılımların ortalamalarına ilişkin yokluk hipotezini test etmek amacıyla hem Student t testini, hem de Johnson'ın Düzeltilmiş Karesel t testini kullanmıştır (Sutton, 1993). Eğer çarpıklık Johnson (1978)'ın incelediği duruma göre daha fazla ve örnek hacmi yeterince büyük değil ise deneysel 1. tip hata oranının nominal değerden farklı olacağını yine Sutton (1993) çalışmasında göstermiştir. Böylece her örneklem büyüklüğü için en az Johnson'ın Düzeltilmiş Karesel Testi kadar güçlü birleşik bir test elde etmiştir. Bu test yine Johnson'ın düzeltilmiş karesel t istatistiğine dayanmaktadır.

$t_1 > t_{n-1,\alpha}$ veya $t_1 / s_{t_1 sb} > z_\alpha$ veya $\left\{(n-1)/(n-3)\right\}^{\frac{1}{2}} t_1/s_{t_1 sb} > t_{n-1,\alpha}$ ise $H_0 : \mu = \mu_0$ yokluk hipotezi $H_1 : \mu > \mu_0$ hipotezine karşı 1-α güven düzeyinde red edilir. Burada, $s_{t_1 sb}$ değeri t_1 istatistiğinin bootstrap standart hatasıdır. Bu bileşik test için red bölgesi α=0.01 ve α=0.05 için oluşturulmuş ve Monte Carlo çalışmaları ile bileşik testin güç bakımından iyi özelliklere sahip olduğu sonucuna varılmıştır (Sutton, 1993).

2.3 Chen'in Testi

Sağa çarpık dağılımların ortalamalarına ilişkin yokluk hipotezinin testi için önerilen bu istatistik, herhangi bir X rastgele değişkeni için aşağıda ifade edilen Edgeworth Açılımları kullanılarak elde edilmiştir (Chen, 1995).

$$P\left\{\frac{\sqrt{n}(\bar{X}-\mu)}{S} \le x - \frac{1}{6\sqrt{n}}\hat{\beta}_1(1+2x^2)\right\} = \Phi(x) + o(n^{-\frac{1}{2}}) \tag{7}$$

Burada $\Phi(x)$ standart normal olasılık yoğunluk fonksiyonudur ve

$$\hat{\beta}_1 = \frac{\sum\left(X_i - \bar{X}\right)^3 / n}{S^3}$$

olarak tanımlanmıştır (Hall, 1983). $o(n^{-\frac{1}{2}})$ ise ilk iki terimden sonraki terimlerin $n^{\frac{1}{2}}$' ye bölümü ile elde edilen serinin sıfıra yakınsadığını göstermektedir. $H_0 : \mu = \mu_0$ hipotezinin $H_1 : \mu > \mu_0$ hipotezine karşı testinde

$$\frac{\sqrt{n}(\bar{X}-\mu_0)}{S} > Z_\alpha - \frac{1}{6\sqrt{n}}\hat{\beta}_1(2Z_\alpha^2+1) \tag{8}$$

ise yokluk hipotezi 1-α güven düzeyinde red edilir (Chen, 1995). Burada Z_α standart normal dağılımın tablo değeridir. Karar kuralı aşağıdaki yaklaşım kullanılarak elde edilir.

$$\frac{\sqrt{n}(\bar{X}-\mu_0)}{S} > x - \frac{1}{6\sqrt{n}}\hat{\beta}_1(2x^2+1)\ ,\ (\forall\, x) \tag{9}$$

Burada $a=\hat{\beta}_1/6\sqrt{n}$ ve t=$\sqrt{n}(\bar{X}-\mu_0)/S$ olmak üzere, $n\to\infty$ iken (1-8a(t+a))≥ 0 varsayımı altında Eşitlik 8 kullanılarak

$$P\left\{\frac{1-\sqrt{(1\text{-}8\text{a}(\text{t}+\text{a}))}}{4\text{a}} > Z_\alpha\right\} \leq \alpha \tag{10}$$

ifadesi yazılabilir. Taylor açılımı kullanılarak

$$\frac{1-\sqrt{(1\text{-}8\text{a}(\text{t}+\text{a}))}}{4\text{a}} = t+a+2at^2+4a^2(t+2t^3)+o(n^{-1}) \tag{11}$$

elde edilir [Chen, 1995].

Böylece test istatistiği

$$t_2 = t+a+2at^2+4a^2(t+2t^3) \tag{12}$$

olur. Bu istatistik, standart normal dağılıma sahiptir. Bu açılım, orijinal değişkenler ile ifade edilirse,

$$\begin{aligned} t_2 = {} & \frac{\sqrt{n}(\bar{X}-\mu_0)}{S} + \frac{1}{6\sqrt{n}}\hat{\beta}_1\left[2\left(\frac{\sqrt{n}(\bar{X}-\mu_0)}{S}\right)^2+1\right]+ \\ & \frac{1}{9n}\hat{\beta}_1^2\left[\frac{\sqrt{n}(\bar{X}-\mu_0)}{S}+2\left(\frac{\sqrt{n}(\bar{X}-\mu_0)}{S}\right)^2\right] \end{aligned} \tag{13}$$

elde edilir (Chen, 1995). Bu test istatistiği örnek hacmi ne kadar küçük olursa olsun kullanılabilir (Chen, 1995). Eğer $t_2 > Z_\alpha$ ise yokluk hipotezi α anlamlılık düzeyinde red edilir. Bu istatistik Student t istatistiği ve Johnson'ın düzeltilmiş karesel t istatistiğinin geliştirilmiş bir halidir. Chen (1995), yaptığı simülasyon çalışmaları sonucunda, test için Z_α kritik değerinin kullanılmasının küçük örneklem büyüklüğüne sahip örneklemlerde çok iyi sonuçlar verdiğini tespit etmiştir. Diğer taraftan, bu kritik değer kullanıldığında örneklem büyüklüğü 20'den fazla iken deneysel 1. tip hata oranlarının, nominal 1. tip hata değerini aştığı gözlemlenmiştir. Ayrıca, Chen (1995) çalışmasında küçük örnek hacimlerinde bu test için $t_{n-1,\alpha}$ kritik değerinin kullanılabileceğini belirtmiştir. Dolayısıyla, bu çalışmada da Chen'in testi için $t_{n-1,\alpha}$ kritik değeri kullanılmıştır.

2.4 Simülasyon Çalışması

Simülasyon çalışmasında, λ=0.25 parametreli üstel dağılım ($\mu = 4$, $\sigma^2 = 16$ ve *Çarpıklık* = 2), k = 0,50, $\lambda = 1,00$ parametrelerine sahip Weibull dağılımı ($\mu = 2.00$, $\sigma^2 = 20.00$ ve *Çarpıklık* = 6.62) ve $\alpha = 1$, $\beta = 2$ parametrelerine sahip Beta dağılımı ($\mu = 0.33$, $\sigma^2 = 20.00$ ve *Çarpıklık* $= 0.57$) kullanılmıştır. Belirtilen parametreler için ele alınan üç dağılım sağa çarpıktır.

Eğer X rastgele değişkeni Üstel dağılıma sahipse, bu değişkenin olasılık yoğunluk fonksiyonu

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases} \tag{14}$$

olarak ifade edilir ve X : $Üstel(\lambda)$ ile gösterilir. Burada λ dağılımın tek parametresi olmak üzere pozitif bir sayıdır.

Weibull dağılımı bir sürekli olasılık dağılımı olup, olasılık yoğunluk fonksiyonu

$$f(x;k,\lambda) = \begin{cases} \dfrac{k}{\lambda}\left(\dfrac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & , \quad x \geq 0; k > 0; \lambda > 0 \\ 0 & , \quad d.h. \end{cases} \tag{15}$$

olarak ifade edilir ve X : $Weibull(k,\lambda)$ ile gösterilir. Burada k şekil parametresi ve λ ölçek parametresidir.

Beta dağılımı [0,1] aralığında α ve β gibi iki tane pozitif şekil parametresi ile normalize edilmiş bir sürekli olasılık dağılımı ailesidir. Beta dağılımının olasılık yoğunluk fonksiyonu

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} & , \quad \alpha > 0, \beta > 0, x \in [0,1] \\ 0 & , \quad d.h. \end{cases} \tag{16}$$

olarak ifade edilir ve $X : Beta(\alpha,\beta)$ ile gösterilir. Olasılık yoğunluk fonksiyonundaki $B(\alpha,\beta)$ ifadesi Gamma fonksiyonlarından faydalanarak, aşağıdaki gibi elde edilir.

$$B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{17}$$

Simülasyon çalışmasında algoritmalar Matlab R2006a paket programı ile yazılmıştır. Simülasyon programları hem deneysel 1. tip hata oranlarını, hem de testin gücü değerlerini elde etmeye yöneliktir. Bu oranları elde etmek için kullanılan kayma miktarı (kym), kitle ortalamasında meydana gelen kaymayı göstermektedir. Simülasyon çalışmasında; aşağıda verilen örneklem büyüklükleri, nominal 1. tip hata oranları ve kayma miktarları kullanılmıştır.

Örneklem Büyüklükleri (*n*)= 7, 13, 20, 30, 50
Nominal 1. tip hata oranları (α) = 0.01, 0.05 ve 0.10
Kayma Miktarları (*kym*)= 0.50; 1.00; 1.50

Kayma miktarları, Beta dağılımı için 0.05, 0.10, 0.15 olarak alınmıştır. Bunun nedeni bu dağılımın 0 ile 1 arasında değer almasıdır. Çalışmada 10000 iterasyon kullanılmıştır.

Test isimleri uzun olduğu için tabloları elde etmede kolaylık olması açısından bazı kısaltmalar yapılmıştır. Bu kısaltmalar; Student t testi için ***ST***, Johnson'ın düzeltilmiş karesel t testi için ***JDKT***, Sutton'ın bileşik testi için ***SBT*** ve Chen'in testi için ***CT*** şeklindedir.

3. BULGULAR

Tablo 1'de, Üstel, Beta ve Weibull dağılımlarından 10000 iterasyona dayalı 7, 13, 20, 30 ve 50 örneklem büyüklükleri ile $H_1 : \mu > \mu_0$ hipotezine karşı $H_0 : \mu = \mu_0$ hipotezinin test edilmesinde her bir test istatistiği için elde edilen deneysel 1. tip hata oranları verilmiştir.

Tablo 1'de Üstel dağılım incelendiğinde, nominal α=0.01 ve örneklem büyüklüğü 7 iken Chen'in testine ilişkin deneysel 1. tip hata oranının 0.009 olduğu görülmektedir. Aynı koşullarda Sutton'ın bileşik testi için deneysel 1. tip hata oranı 0.004 olarak gerçekleşmiştir. Diğer taraftan Weibull dağılımı için yine aynı koşullar incelendiğinde ise, Chen'in testine ilişkin deneysel 1. tip hata oranının 0.005 ve Sutton'ın bileşik testine ilişkin deneysel 1. tip hata oranının ise 0.003 olduğu görülmektedir. Diğer testler için Üstel dağılımdan ve Weibull dağılımından çekilen örneklemlerden hesaplanan deneysel 1. tip hata oranları ise nominal değerlere yaklaşamamıştır.

İlgili parametreler için Beta dağılımı, diğer dağılımlara göre daha düşük çarpıklık değerine sahiptir. Bu durumda, Chen'in testi ve Sutton'ın bileşik testi için deneysel 1. tip hata oranları, nominal α'ya, diğer testler için elde edilen deneysel 1. tip hata oranlarına göre daha çok yaklaşmıştır. Beta dağılımı için hem Sutton'ın bileşik testi, hem de Chen'in testi deneysel 1. tip hata oranları bakımından hemen hemen aynı sonuçları vermiştir. Diğer testler için deneysel 1. tip hata oranları çarpıklıkların düşük olmasından dolayı belirli bir artış göstermişlerdir. Fakat, bu oranlar çoğu durumda nominal α'ya yaklaşamamıştır.

Üstel dağılım ve Weibull dağılımı için örneklem büyüklüğünün 20, 30 veya 50 olduğu durumlarda, Chen'in testi deneysel 1. tip hata oranları bakımından nominal α'ya en yakın sonuçlara sahiptir. Aynı koşularda Beta dağılımı incelendiğinde ise Sutton'ın bileşik testi ve Chen'in testi nominal α'ya en yakın sonuçlara sahiptir. Genel olarak düşünüldüğünde, Chen'in testi ele alınan dağılımlardan çekilen örneklem ile kitle ortalamasına ilişkin yokluk hipotezinin test edilmesinde diğer testlerden daha iyi sonuçlara sahiptir.

Doğal bir sonuç olarak, örneklem büyüklüğü arttıkça, bütün testler için deneysel 1. tip hata oranlarının nominal α'ya yaklaştığı gözlemlenmiştir.

Tablo 1. Beta, Üstel ve Weibull dağılımları için deneysel 1. tip hata oranları

Örneklemin Geldiği Kitlenin Dağılımı	Nominal α	Test İstatistiği	n=7	n=13	n=20	n=30	n=50
Beta Dağılımı α=1 β=2 (0.57)	0.01	ST	0.005	0.005	0.006	0.006	0.007
		JDKT	0.006	0.006	0.007	0.007	0.008
		SBT	0.008	0.008	0.008	0.009	0.009
		CT	0.008	0.008	0.008	0.009	0.009
	0.05	ST	0.031	0.033	0.035	0.040	0.041
		JDKT	0.031	0.033	0.037	0.043	0.045
		SBT	0.033	0.037	0.041	0.047	0.048
		CT	0.032	0.040	0.048	0.049	0.050
	0.10	ST	0.071	0.077	0.080	0.087	0.092
		JDKT	0.072	0.079	0.084	0.094	0.098
		SBT	0.074	0.084	0.088	0.097	0.101
		CT	0.074	0.084	0.089	0.098	0.101
Üstel Dağılım λ=0.25 (2.00)	0.01	ST	0.000	0.000	0.000	0.002	0.003
		JDKT	0.001	0.002	0.004	0.005	0.008
		SBT	0.004	0.005	0.005	0.007	0.009
		CT	0.009	0.010	0.010	0.011	0.011
	0.05	ST	0.012	0.015	0.018	0.023	0.029
		JDKT	0.014	0.031	0.033	0.041	0.048
		SBT	0.022	0.034	0.036	0.044	0.050
		CT	0.040	0.048	0.052	0.053	0.055
	0.10	ST	0.040	0.053	0.054	0.065	0.071
		JDKT	0.053	0.079	0.080	0.093	0.097
		SBT	0.062	0.081	0.083	0.096	0.101
		CT	0.090	0.095	0.102	0.104	0.106
Weibull Dağılımı k=0.50 λ=1.00 (6.62)	0.01	ST	0.000	0.000	0.000	0.000	0.000
		JDKT	0.001	0.002	0.002	0.003	0.007
		SBT	0.003	0.004	0.004	0.006	0.008
		CT	0.005	0.005	0.005	0.008	0.010
	0.05	ST	0.002	0.002	0.004	0.006	0.011
		JDKT	0.013	0.014	0.017	0.018	0.022
		SBT	0.014	0.020	0.024	0.035	0.041
		CT	0.022	0.031	0.041	0.044	0.050
	0.10	ST	0.013	0.019	0.023	0.030	0.042
		JDKT	0.032	0.053	0.077	0.078	0.085
		SBT	0.041	0.072	0.080	0.086	0.087
		CT	0.055	0.082	0.088	0.090	0.100

Tablo 2, Tablo 3 ve Tablo 4'te Beta, Üstel ve Weibull dağılımlarından 10000 tekrar ile seçilen 7, 13 ve 30 büyüklüğündeki örnekler ile $H_1 : \mu > \mu_0$ hipotezine karşı $H_0 : \mu = \mu_0$ hipotezinin test edilmesinde her bir test istatistiği için testin gücü değerleri verilmiştir.

Tablo 2'de yer alan Beta dağılımından çekilen örneklemlerle elde edilen testin gücü değerleri çarpıklığının düşük olması nedeniyle bütün testler için birbirine yakın çıkmıştır. Fakat Chen'in testi ve Sutton'ın bileşik testi, diğer testlere oranla daha yüksek testin gücü değerlerine sahiptir.

Tablo 2. Beta dağılımı için testin gücü değerleri

Örneklemin Geldiği Kitlenin Dağılımı	Nominal α	Test İstatistiği	*n*=7			*n*=13			*n*=30		
			kym = **0.05**	*kym* = **0.10**	*kym* = **0.15**	*kym* = **0.05**	*kym* = **0.10**	*kym* = **0.15**	*kym* = **0.05**	*kym* = **0.10**	*kym* = **0.15**
Beta Dağılımı	0.01	ST	0.014	0.037	0.084	0.029	0.110	0.325	0.083	0.427	0.881
α=1		JDKT	0.015	0.038	0.085	0.029	0.126	0.401	0.100	0.502	0.931
β=2		SBT	0.025	0.060	0.135	0.041	0.168	0.480	0.118	0.534	0.938
(0.57)		CT	0.025	0.060	0.140	0.041	0.169	0.480	0.119	0.538	0.940
	0.05	ST	0.086	0.192	0.367	0.134	0.372	0.711	0.276	0.745	0.980
		JDKT	0.088	0.194	0.394	0.146	0.415	0.773	0.303	0.789	0.987
		SBT	0.098	0.220	0.446	0.160	0.437	0.789	0.315	0.797	0.989
		CT	0.100	0.230	0.448	0.161	0.437	0.791	0.317	0.798	0.989
	0.10	ST	0.182	0.347	0.590	0.250	0.559	0.855	0.423	0.864	0.994
		JDKT	0.183	0.366	0.636	0.270	0.597	0.885	0.451	0.883	0.996
		SBT	0.200	0.390	0.660	0.281	0.608	0.889	0.461	0.888	0.997
		CT	0.200	0.391	0.660	0.285	0.608	0.893	0.465	0.889	0.997

Tablo 3'te, üstel dağılım incelendiğinde, testin gücü değerlerinin düşük gerçekleştiği görülmektedir. Bütün örneklem büyüklüklerinde, kayma miktarı (kym) 0,50 iken en yüksek testin gücü değerlerine Chen'in testi sahiptir.

Üstel dağılımdan seçilen örneklem ile kitle ortalamasına ilişkin yokluk hipotezinin test edilmesinde, Chen'in testinin daha güçlü bir test olduğu aşikardır.

Tablo 3. Üstel dağılım için testin gücü değerleri

Örneklemin Geldiği Kitlenin Dağılımı	Nominal α	Test İstatistiği	*n*=7			*n*=13			*n*=30		
			kym = **0.50**	*kym* = **1.00**	*kym* = **1.50**	*kym* = **0.50**	*kym* = **1.00**	*kym* = **1.50**	*kym* = **0.50**	*kym* = **1.00**	*kym* = **1.50**
Üstel Dağılım	0.01	ST	0.002	0.003	0.006	0.003	0.006	0.018	0.011	0.042	0.101
λ=0.25		JDKT	0.002	0.005	0.008	0.008	0.019	0.044	0.032	0.100	0.220
(2.00)		SBT	0.009	0.013	0.025	0.015	0.035	0.072	0.041	0.118	0.249
		CT	0.019	0.035	0.062	0.029	0.066	0.125	0.059	0.159	0.314
	0.05	ST	0.022	0.041	0.066	0.041	0.086	0.153	0.093	0.220	0.395
		JDKT	0.032	0.058	0.095	0.073	0.142	0.232	0.142	0.310	0.512
		SBT	0.045	0.078	0.124	0.082	0.160	0.253	0.149	0.321	0.524
		CT	0.077	0.129	0.193	0.111	0.207	0.319	0.174	0.356	0.562
	0.10	ST	0.072	0.121	0.177	0.115	0.206	0.311	0.197	0.385	0.589
		JDKT	0.098	0.158	0.227	0.159	0.274	0.396	0.249	0.466	0.659
		SBT	0.111	0.177	0.247	0.167	0.283	0.408	0.255	0.473	0.667
		CT	0.155	0.235	0.315	0.196	0.328	0.457	0.270	0.495	0.688

Tablo 4 incelendiğinde, Weibull dağılımı için Sutton'ın bileşik ve Chen'in testlerinin örneklem büyüklüğü 7, nominal α=0.01 ve kayma miktarı 0,50 olduğunda aynı güce sahip olduğu görülmektedir. Diğer durumlarda Chen'in testine ilişkin testin gücü değerleri daha yüksek gerçekleşmiştir.

Örneklem büyüklüğünün 13 ve 30 olduğu durumda Weibull dağılımı için testlere ilişkin testin gücü değerleri incelendiğinde, Chen'in testinin daha güçlü bir test olduğu anlaşılmaktadır. Örneklem büyüklüğü arttıkça, Weibull dağılımdan çekilen örneklem ile kitle ortalamasına ilişkin yokluk hipotezinin test edilmesinde tüm testler için testin gücü değerleri artmıştır. Kayma miktarı arttıkça, tüm testler için testin gücü değerlerinde yine artış meydana gelmiştir.

Tablo 4. Weibull dağılımı için testin gücü değerleri

Örneklemin Geldiği Kitlenin Dağılımı	Nominal α	Test İstatistiği	*n*=7			*n*=13			*n*=30		
			kym= 0.50	*kym*= 1.00	*kym*= 1.50	*kym*= 0.50	*kym*= 1.00	*kym*= 1.50	*kym*= 0.50	*kym*= 1.00	*kym*= 1.50
Weibull	0.01	ST	0.000	0.001	0.034	0.000	0.002	0.179	0.008	0.026	0.674
Dağılımı		JDKT	0.005	0.007	0.049	0.008	0.017	0.345	0.022	0.179	0.726
k=0.50		SBT	0.024	0.030	0.125	0.028	0.066	0.416	0.042	0.269	0.729
λ=1.00		CT	0.024	0.031	0.258	0.030	0.079	0.506	0.050	0.306	0.733
(6.62)	0.05	ST	0.006	0.020	0.231	0.013	0.064	0.492	0.050	0.271	0.733
		JDKT	0.035	0.058	0.318	0.058	0.199	0.567	0.147	0.523	0.737
		SBT	0.054	0.109	0.363	0.075	0.237	0.574	0.160	0.545	0.737
		CT	0.060	0.139	0.456	0.090	0.291	0.595	0.194	0.595	0.738
	0.10	ST	0.035	0.101	0.408	0.067	0.226	0.587	0.157	0.524	0.737
		JDKT	0.096	0.202	0.480	0.157	0.396	0.610	0.272	0.685	0.738
		SBT	0.113	0.230	0.489	0.167	0.409	0.612	0.279	0.692	0.738
		CT	0.115	0.268	0.532	0.197	0.457	0.619	0.308	0.721	0.749

Beta dağılımı, Üstel dağılım ve Weibull dağılımından çekilen örneklemler ile kitle ortalamasına ilişkin yokluk hipotezini test etmede testin gücü değerleri bakımından Chen'in testinin daha iyi bir test olduğu söylenebilir.

4. TARTIŞMA VE SONUÇ

Bu çalışmada, sağa çarpık kitlelerden seçilen örneklemler ile kitle ortalamasının testi için kullanılan Johnson'ın düzeltilmiş karesel t testi, Sutton'ın bileşik testi, Chen'in testi deneysel 1. tip hata oranları ve testin gücü değerleri bakımından karşılaştırılmıştır. Bu amaçla, Üstel (0.25), Weibull (0.50, 1) ve Beta (1, 2) dağılımlarından örneklemler çekilmiş ve her bir test için deneysel 1. tip hata oranları ve testin gücü değerleri hesaplanmıştır.

Gerçekte doğru olan yokluk hipotezini reddetme (1. tip hata) oranları için Tablo 1'e bakıldığında, Johnson'ın düzeltilmiş karesel testine ilişkin deneysel 1. tip hata oranları nominal α 'ya yaklaşmışsa da, bu genelde çarpıklığın düşük ve örneklem büyüklüğünün büyük olduğu durumlarda gerçekleşmiştir. Sutton'ın bileşik testi ve Chen'in testi özellikle küçük örnekler için deneysel 1. tip hata oranları bakımından hemen hemen yakın sonuçlara sahiptir. Fakat küçük örneklerde başarılı gibi görünen Sutton'ın bileşik testi, aynı başarıyı büyük örneklemlerde gösterememiştir. Chen'in testi ise çoğu durumda, deneysel 1. tip hata oranları bakımından nominal α 'ya yaklaşmıştır. Bu sonuçlar, Chen'in testinin daha iyi bir test olduğunun göstergesidir.

Gerçekte yanlış olan yokluk hipotezini reddetme (testin gücü) değerleri için Tablo 2 ve sonraki tablolara bakıldığında, Johnson'ın düzeltilmiş karesel testi bazı durumlarda güçlü bir test olarak görünse de, bu sadece çarpıklığın düşük ve örneklemin büyük olduğu durumlarda söz konusudur. Testin gücü değerleri için elde edilen çoğu tabloda, Chen'in testinin gücü en yüksektir. Sutton'ın bileşik testi bazı durumlarda testin gücü açısından, Chen'in testine yaklaşmış olsa da, bu durum sadece küçük örneklemlerde veya yüksek kayma miktarlarında elde edilmiştir. Beta dağılımının çarpıklığı düşük olduğundan, Sutton'ın bileşik testi testin gücü değerleri bakımından, Chen'in testine ilişkin değerler ile yakın sonuçlara sahiptir.

Bazı durumlarda, Sutton'ın bileşik testi ile Chen'in testi için testin gücü değerleri ve deneysel 1. tip hata oranları bakımından yakın sonuçlar elde edilse de, Sutton'ın bileşik testi Bootstrap örneği çekmek gerektirdiğinden ve üç aşamalı bir karar mekanizmasına sahip olduğundan, uygulamada kullanılması zor bir yöntemdir. Chen'in testi ise tek bir karar kuralına sahip olduğundan daha kullanışlıdır.

Chen'in testi, ele alınan dağılımlar için gerçekte doğru olan yokluk hipotezini reddetme oranları ve gerçekte yanlış olan yokluk hipotezini reddetme değerleri bakımından diğer testlere göre daha iyi bir testtir.

5. KAYNAKLAR

Andrews, D. F., Bickel, P. J., Hampel, F. R., Huber, P. J., Rogers, W. H., Tukey, J. W., 1972. Robust estimates of location: Survey and advances. Princeton University Press, Princeton.

Anscombe, F. J., 1950. Tables of hyberbolic transformation sinh. $\sqrt{X}$. Journal of the Royal Statistical Society, A/113, 228-229.

Bartlett, M. S., 1935. The effect of non-normality on the t distribution. Proceedings of the Cambridge Philosophical Society. 31, 223-231.

Box, G. E. P., Andersen, S. L., 1955. Permutation theory in the derivation of robust criteria and the study of departures from assumption. Journal of the Royal Statistical Society, B/17, 1-26.

Chen, L., 1995. Testing the mean of skewed distributions. Journal of The American Statistical Association. Vol. 90, No. 430, 767-772.

Chung, K., 1946. The approximate distribution of Student's statistics. Annals of Mathematical Statistics, 17, 447-465.

Cornish, E. A., Fisher, R. A., 1937. Moments and cumulants in the specification of distributions. Revue of the International Statistics Institute, 5, 307-327.

Gayen, A. K., 1949. The distribution of 'Student's' t in random samples of any size drawn from non-normal universes. Biometrika, 36, 353-369.

Geary, R. C., 1936. The distribution of Student's ratio for non-normal samples. Journal of the Royal Statistical Society, 3/2, 178-184.

Hall, P., 1983. Inverting an Edgeworth expansion. The Annals of Statistics, 11: 569-576.

Johnson, N. J., 1978. Modified t tests and confidence intervals for asymmetrical populations. Journal of the American Statistical Association, Vol. 73, No. 363, 536-544.

Laderman, J., 1939. The distribution of Student's ratio for samples of two items drawn from non-normal universes. Annals of Mathematical Statistics, 10, 376-379.

Nair, A. K. N., 1941. Distribution of Student's t in the correlation coefficient in sample from non-normal population. Sankhya, 5, 383-400.

Neyman, J., Pearson, E. S., 1928. On the use and interpretation of certain test criteria for purposes of statistical inference part I. Biometrika, 20A, 175-240.

Perlo, V., 1933. On the distribution of Student's ratio for samples of three drawn from a rectangular distribution. Biometrika, 25, 203-204.

Rider, P. R., 1929. On the distribution of the ratio of mean to standard deviation in small samples from non-normal universes. Biometrika, 21, 124-143.

Sophister, 1928. Discussion of small samples drawn from an infinite skew population. Biometrika, 20A, 389-423.

Student, 1908. The probable error of a mean. Biometrika, 6, 1-25.

Sutton, C. D., 1993. Computer-intensive methods for tests about the mean of an asymmetrical distribution. Journal of the American Statistical Association, 88, 802-810.

Tukey, J. W., 1964. Data analysis and behavioral sciences. Unpublished Manuscript.

Wallace, D. L., 1958. Asymptotic approximations to distributions. Annals of Mathematical Statistics, 29, 635-654.

Yuen, K. K., 1974. The two sample trimmed t for unequal population variances. Biometrika, 61, 165-170.

USE OF SOME TESTS FOR MEANS OF POSITIVELY SKEWED DISTRIBUTION AND COMPARISONS

ABSTRACT

The purpose of this article is to introduce Johnson's modified square t test, Sutton's composite test and Chen' s test applied to test the null hypotheses of population mean with samples drawn from positively skewed distributions, also compare them on the account of type 1 error rates and power of the test. In this context, a simulation study is carried out via data group which is produced from some positively skewed distributions and it is determined that Chen's test is more powerful than other tests.

Keywords: Chen's test, Johnson's modified square t test, Power of the test, Skewness, Sutton's composite test.