Cilt: 06 Sayı: 01 Kategori: 02 Sayfa: 65-74 ISSN No: 1303-6319 Volume: 06 Sayı: 01 Category: 02 Page: 65-74

GRUPLANDIRILMIŞ VERİLERİN ÜSTEL DAĞILIMA UYUMUNDA AĞIRLIKLANDIRILMIŞ KOLMOGROV-SMIRNOV TESTLERİ İLE OLABİLİRLİK ORANI VE Kİ-KARE TESTLERİNİN KARŞILAŞTIRILMASI

Hamza GAMGAM* **Esra YİĞİT****

ÖZET

Bu çalışmada, gruplandırılmış verilerin üstel dağılıma uyumu için Gulati ve Neus (2003) tarafından önerilen Ağırlıklandırılmış Kolmogrov-Smirnov test istatistikleri tanıtılmıştır. Bu istatistiklerle, olabilirlik oranı ve ki-kare Uyum İyiliği test istatistikleri, farklı alternatif dağılımı, grup sayısı ve örnek çapı için güç bakımından karşılaştırılması yapılmıştır. Karşılaştırmalar sonucunda özellikle sağa çarpık dağılımlarda, Ağırlıklandırılmış Kolmogrov-Smirnov test istatistiklerinin güç bakımından performansının, Pearson χ^2 ve olabilirlik oranı test istatistiğine göre daha iyi olduğu görülmüştür.

Anahtar Kelimeler: Ağırlıklandırılmış Kolmogrov-Smirnov, Bootstrap, Gruplandırılmış veri, Pearson Ki-Kare, Olabilirlik oranı, Uyum iyiliği.

1. GİRİŞ

Üstel dağılım, başarısızlık gerçekleşene kadar geçen sürenin dağılımıdır. Bu sebepten birçok endüstriyel ve biyolojik uygulamalarda sıklıkla kullanılır. Hafızasızlık özelliğinden dolayı bir sonraki başarısızlığın ne zaman gerçekleşebileceğinin tahmini için oldukça faydalıdır. Gruplandırılmış verileri, özellikle bazı uygulama alanında kullanmak gerekli olabilir. Örneğin gözlem birimlerinin tam olarak ölçülemediği durumlarda gruplandırılmış veri kullanmak daha sağlıklı olur. Birçok deneyde birimleri sürekli değişken olarak gözlemlemek çok zor ya da imkânsızdır. Bunun yerine önceden, belirlenmiş belirli aralıklarda ölçüm yapmak hem daha kolay hem de daha ucuzdur. Örneğin belediyenin yeni aldığı bir grup otobüsün bozulma zamanları ile ilgili çalışma yapmak istenmesi durumunda, bunların her birinin bozulma zamanlarının kaydını ayrı ayrı tutmak yerine, belirli zaman aralıklarında kaç tanesinin bozulduğunun ölçülmesi daha kolay ve ucuz bir yoldur.

* Gazi Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü Teknikokullar, Ankara, e-posta: gamgam@gazi.edu.tr
** Gazi Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü Teknikokullar, Ankara, e-posta: eyigit@gazi.edu.tr

Gruplandırılmış verilerde uyum iyiliğini ölçmek için ilk test istatistiği Pearson (1900) tarafından geliştirilmiştir. Daha sonra Kolmogrov Smirnov ve Cramer Von Mises test istatistikleri geliştirilmiştir. Kolmogrov Smirnov test istatistiğini Schmid (1958) ve Conover (1972) gruplandırılmış veri için düzenlemişlerdir. Kuldorf (1961), gruplandırılmış veriler için üstel dağılımın parametre tahminini en çok olabilirlik yöntemi ile elde etmiştir. Ayrıca gruplandırılmış veriler için Seo ve Yum (1993) tarafından üstel dağılımın parametresinin tahmini için bir yöntem geliştirilmiştir. Gulati ve Neus (2003) gruplandırılmış veriler için Kolmogrov-Smirnov test istatistiğini ağırlıklandırarak, üstel dağılım için iki uyum iyiliği test istatistiği önermiştir. Daha sonra Baklizi (2006), Gulati ve Neus (2003) tarafından üstel dağılım için geliştirilen uyum iyiliği test istatistiğini Rayleigh dağılımına uygulamıştır.

Ayrıca Best ve Rayner (2007) χ^2'nin bileşenlerinden faydalanarak, gruplandırılmış veriler için üstel dağılıma uyum iyiliği test istatistiği geliştirmiştir.

Çalışmanın ikinci bölümünde, Gulati ve Neus (2003) tarafından önerilen SW1 ve SW2 test istatistikleri tanıtılmıştır. Üçüncü bölümde ise bu istatistikler ile olabilirlik oranı ve ki-kare test istatistiğinin güç bakımından karşılaştırılmasında kullanılacak yöntemin algoritması verilmiştir. Ayrıca bu bölümde farklı parametreli bazı dağılımlar için, farklı grup sayıları, örnek çaplarına göre SW1, SW2 olabilirlik oranı ile ki-kare testlerinin güçleri hesaplanmış ve bunlara ait güç grafikleri çizilmiştir.

2. YÖNTEM

2.1 Gruplandırılmış Veriler için Kullanılan SW1 ve SW2 Test İstatistikleri

Olasılık yoğunluk fonksiyonu f(x) olan bir dağılımdan n tane gözlem alınsın ve bunlar, $x_1, x_2, \ldots, x_{k-1}$ kesim noktaları olmak üzere, $(0, x_1)$, $(x_1, x_2), \ldots, (x_{k-1}, \infty)$ k sayıda gruba ayrılsın. Gruplara düşen gözlem sayıları sırasıyla $n_1, n_2, \ldots, n_k$ olsun. Gruplandırılmış verilerin üstel dağılıma uyumu için yokluk hipotezi aşağıdaki gibidir (Gulati ve Neus (2003)).

H_0: $f(x)=\theta\exp(-\theta x)$, ($\theta>0$ bilinmiyor , $x>0$)

θ'nın en çok olabilirlik tahmin edicisi, $n_1<n$ ve $n_k<n$ koşulları altında, Kulldorf (1961) tarafından,

$$\sum_{i=1}^{k-1} \frac{n_i\left(x_i - x_{i-1}\right)}{e^{\theta(x-x_{i-1})}-1} - \sum_{i=2}^{k} n_i x_{i-1} = 0 \tag{1}$$

eşitliği her aralığın eşit olduğu varsayımı altında çözülerek;

$$\hat{\theta} = \frac{1}{x_1}\ln\left(1+\frac{n-n_k}{\sum_{i=2}^{k}(i-1)n_i}\right) \tag{2}$$

şeklinde elde edilmiştir (Gulati ve Neus, 2003). Eğer bu varsayım kullanılmaz ise $\hat{\theta}$ için tekrarlı çözüm gerekir.

1≤i≤k için n_i i. gruptaki gözlem sayısı ve x_i de i. grubun üst sınırı olmak üzere, üstel dağılım için x_i değerinden daha küçük olma olasılığı,

$$F(x_i,\hat{\theta}) = 1 - \exp(-x_i\theta)$$

biçiminde ifade edilir. Bu olasılığın deneysel sonucu ise,

$$F_n(x_i) = \sum_{j=1}^{i} n_j / n$$

şeklinde tanımlanır. Bu iki olasılığın farkı,

$$S_i = F_n(x_i) - F(x_i,\hat{\theta})$$

olarak ifade edilir. SW1 ve SW2 test istatistikleri aşağıdaki gibi önerilmiştir.

$$\text{SW1}=\sqrt{n}\sum_{j=1}^{k-1}\frac{|S_j|}{\Psi_1(j)} \qquad (3)$$

ve

$$\text{SW2} = \sqrt{n}\sum_{j=1}^{k-1}\Psi_2(j)|S_j| \qquad (4)$$

Burada, $\Psi_1(j)$ ve $\Psi_2(j)$ ağırlık fonksiyonları olmak üzere,

$$\Psi_1(j) = \sqrt{F\left(x_i,\hat{\theta}\right)\left(1 - F(x_i,\hat{\theta}\right)},$$

$$\Psi_2(j) = (k/2 - j)^2 +$$

olarak tanımlanır (Gulati ve Neus (2003)). ψ_2'deki (+) ifadesi;

$$x+ = \begin{cases} x & x \neq 0 \\ 1 & x = 0 \end{cases}$$

şeklinde tanımlanır. SW1 test istatistiği dağılımın uçlarına ağırlık verirken, SW2 test istatistiği dağılımın merkezine daha fazla ağırlık verir. SW1 ve SW2 test istatistiklerinin tam dağılımı bilinmediği için Bootstrap yöntemi ile p değeri bulunarak, anlamlılık düzeyine göre test sonucu yorumlanır.

3. BULGULAR

3.1 Simülasyon Çalışması ve Güçlerinin Karşılaştırılması

SW1, SW2 olabilirlik oranı ve χ^2 testlerinin güç karşılaştırılması için kullanılan algoritma aşağıda verilmiştir.

1) Çeşitli dağılımlardan ((Ki-kare (1), Ki-kare (4), Weibull (0.8, 1), Lognormal (0, 1), Normal (2, 2), Tekdüze (0, 2.5), Beta (2, 2) ve Lojistik (1.2, 0.35)) n sayıda rassal sayı üretilerek, k gruba bölünür ve (2) eşitliği kullanılarak her biri için $\hat{\theta}$ değeri hesaplanır.
2) Bunların her biri için SW1, SW2, olabilirlik oranı ve χ^2 test istatistiklerinin değerleri elde edilir.

3) 1. adımda elde edilen her bir $\hat{\theta}$ değeri ile üstel dağılımdan, n sayıda rassal sayı üretilerek k gruba bölünür ve bu şekilde 10000 tane Bootstrap örnek oluşturularak, SW1 ve SW2 istatistiklerinin Bootstrap dağılımları elde edilir.
4) Olabilirlik oranı ve χ^2 testi için $\chi^2_{(k-2)}$ dağılımı kullanılarak, *p* değerleri hesaplanır.
5) 2. adımda elde edilen SW1 ve SW2 test istatistiklerinin değerleri için SW1 ve SW2 test istatistiklerinin 3. adımda oluşturulan dağılımları kullanılarak, Bootstrap *p* değerleri hesaplanır.
6) 4. ve 5. adımdaki *p* değerleri seçilen α=0.05 ile karşılaştırılıp, testin sonucu bulunur.
7) Bu işlem 10000 kez tekrar edilerek, 4. ve 5. adımdaki red sayıları bulunur ve bu red sayıları 10000'e bölünerek, her bir test istatistiği için testin gücü hesaplanır.

Simülasyon çalışmasında, yukarıda belirtilen dağılımlardan, *n* çaplı (*n*= 20: 20: 200) örnekler seçilerek farklı grup sayılarına (*k*= 4, 6, 10) ayrılmıştır. Ki-kare (1), Ki-kare (4), Weibull (0.8, 1), Lognormal (0, 1), Normal (2, 2), Tekdüze (0, 2.5) ve Lojistik (1.2, 0.35) için herhangi bir k değeri için i. grubun aralıkları;

$$[x_{i-1}, x_i) = \begin{cases} \left[\dfrac{2(i-1)}{k-1}, \dfrac{2i}{k-1}\right) & i < k \\ \left[\dfrac{2(i-1)}{k-1}, \infty\right) & i = k \end{cases} \tag{5}$$

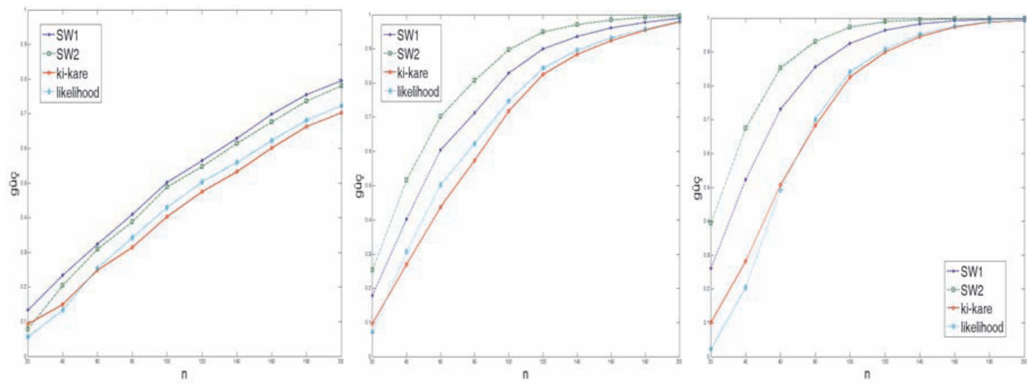
şeklinde alınmıştır. Burada $2i/(k-1)$ kesim noktaları olarak adlandırılır. Beta (2, 2) için;

herhangi bir k değeri için i. grubun aralıkları;

$$[x_{i-1}, x_i) = \begin{cases} \left[\dfrac{(i-1)}{k-1}, \dfrac{i}{k-1}\right) & i < k \\ \left[\dfrac{(i-1)}{k-1}, \infty\right) & i = k \end{cases} \tag{6}$$

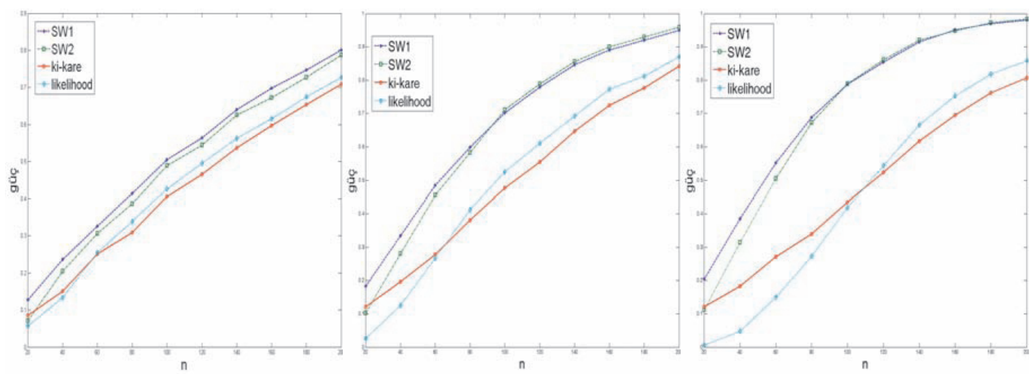
şeklinde alınmıştır. Burada $i/(k-1)$ kesim noktaları olarak adlandırılır.

Her bir durum için SW1, SW2 olabilirlik oranı ve χ^2 test istatistiklerinin değeri hesaplanmıştır. Daha sonra üstel dağılımdan örnek çapı *n* için, 10000 tane Bootstrap örnek üretilerek, SW1 ve SW2 istatistiklerinin Bootstrap dağılımları elde edilmiştir. Oluşturulan bu dağılımlar kullanılarak, SW1 ve SW2 testleri için Bootstrap *p* değeri hesaplanmıştır. Olabilirlik oranı ve χ^2 testi için $\chi^2_{(k-2)}$ dağılımı kullanılarak, *p* değeri elde edilmiştir. Bu *p* değerleri α=0.05 ile karşılaştırılarak, test sonucu bulunmuştur. Bu işlem 10 000 kez tekrarlanarak her bir test istatistiği için red sayıları saptanmış ve bu red sayıları 10 000'e bölünerek, her bir test istatistiği için testin gücü hesaplanmıştır. Bu durum *n*= 20: 20: 200 örnek çapları için yapılarak, Şekil 1-8'deki grafikler elde edilmiştir.



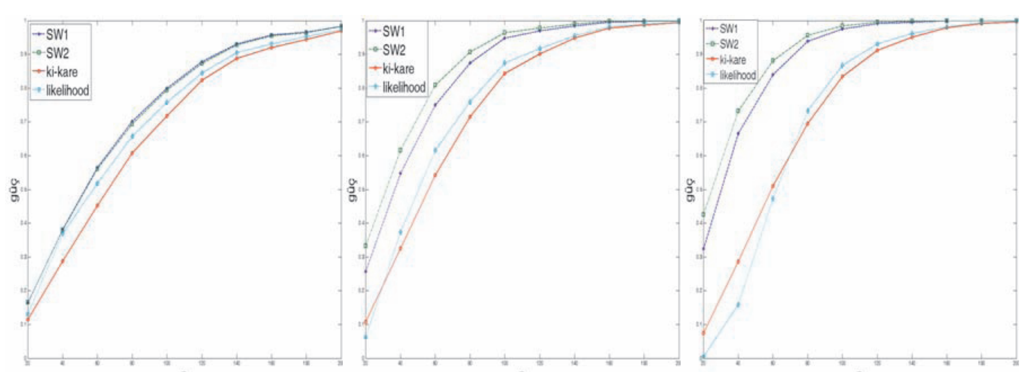
(a) (b) (c)

Şekil 1. χ_1^2 **dağılımından üretilen verilerin, grup sayısı a) k=4; b) k=6; c) k=10 iken, üstel dağılıma uygunluk testinin güç**



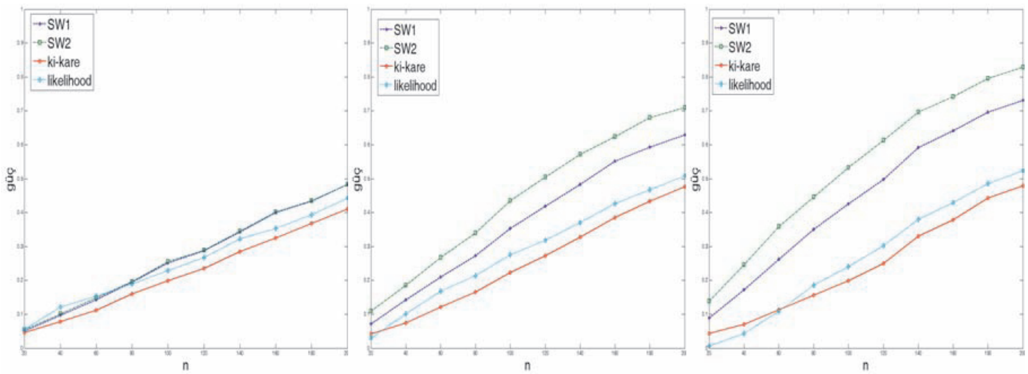
(a) (b) (c)

Şekil 2. χ_4^2 **dağılımından üretilen verilerin, grup sayısı a) k=4; b) k=6; c) k=10 iken, üstel dağılıma uygunluk testinin güç grafikleri**



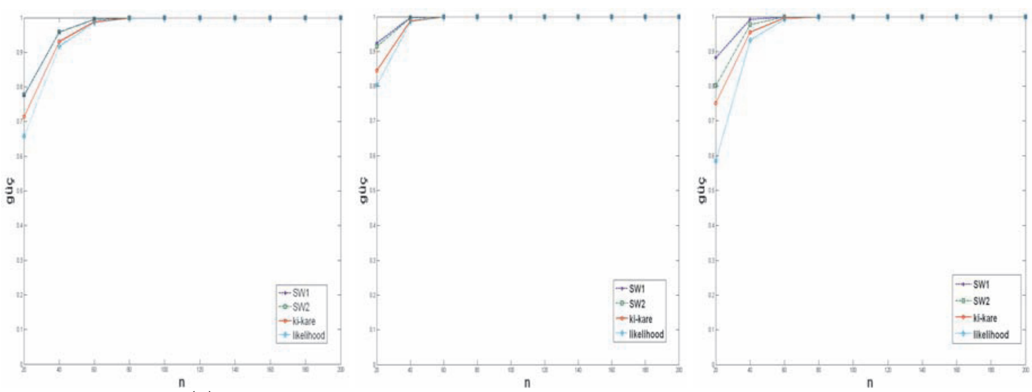
(a) (b) (c)

Şekil 3. Log-normal (0, 2) dağılımından üretilen verilerin, grup sayısı a) k=4; b) k=6; c) k=10 iken, üstel dağılıma uygunluk testinin güç grafikleri



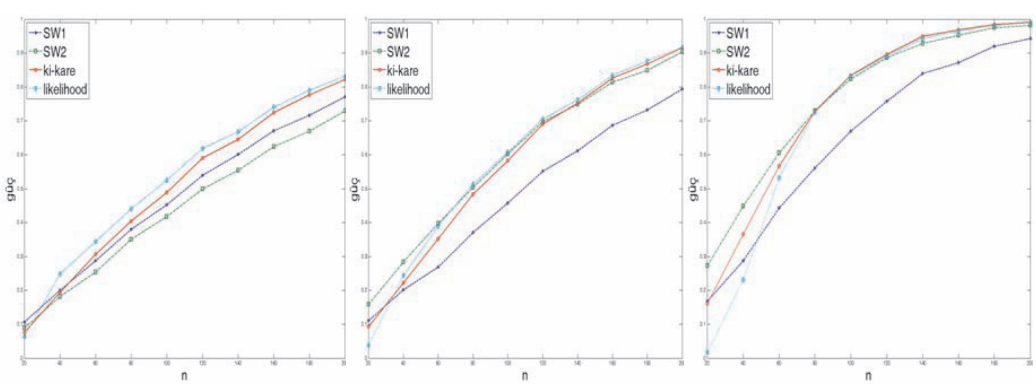
(a) (b) (c)

Şekil 4. Weibull (1, 0.8) dağılımından üretilen verilerin, grup sayısı a) k=4; b) k=6; c) k=10 iken, üstel dağılıma uygunluk testinin güç grafikleri



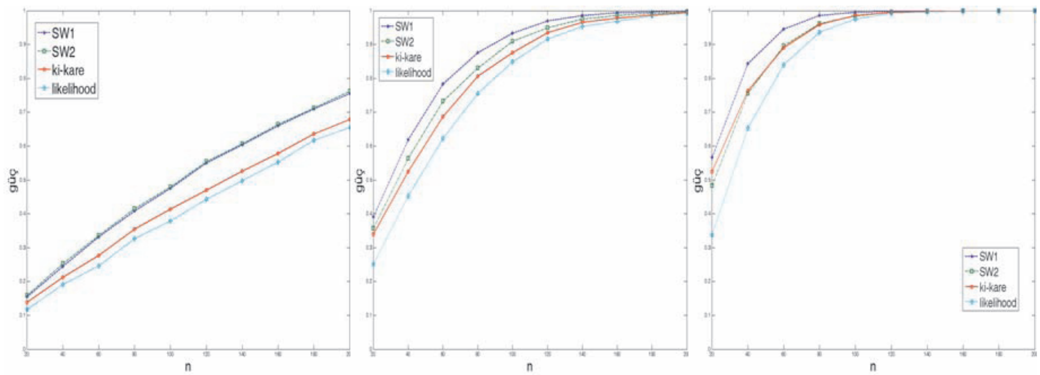
(a) (b) (c)

Şekil 5. Lojistik (1.2, 0.35) dağılımından üretilen verilerin, grup sayısı a) k=4; b) k=6; c) k=10 iken, üstel dağılıma uygunluk testinin güç grafikleri



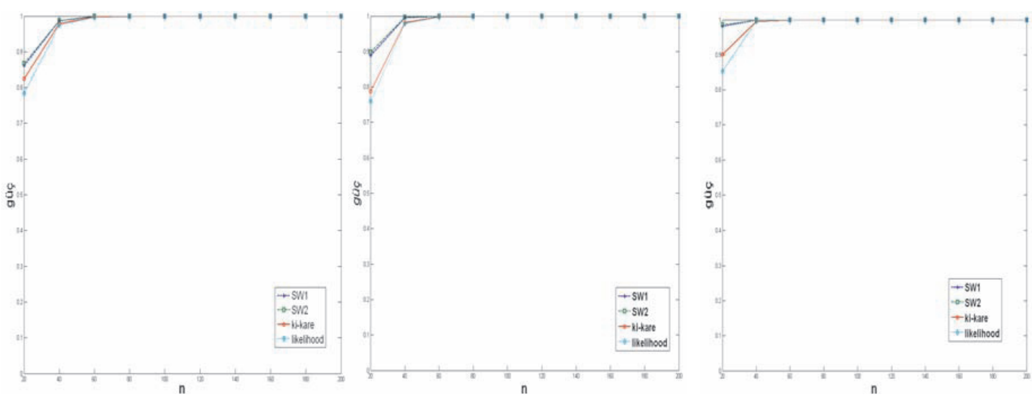
(a) (b) (c)

Şekil 6. Normal (2, 2) dağılımından üretilen verilerin, grup sayısı a) k=4; b) k=6; c) k=10 iken, üstel dağılıma uygunluk testinin güç grafikleri



(a) (b) (c)

Şekil 7. Tekdüze (0, 2.5) dağılımından üretilen verilerin, grup sayısı a) k=4; b) k=6; c) k=10 iken, üstel dağılıma uygunluk testinin güç grafikleri



(a) (b) (c)

Şekil 8. Beta (2, 2) dağılımından üretilen verilerin, grup sayısı a) k=4; b) k=6; c) k=10 iken, üstel dağılıma uygunluk testinin güç grafikleri

Şekil 1, 2, 3, 4, 5'te SW1, SW2, Ki-kare ve olabilirlik oran test istatistikleri, testlerin gücü bakımından alternatif dağılım biçiminin sağa çarpık olduğu durum incelenmiştir. Bu yüzden sağa çarpık dağılımlardan; χ_1^2, χ_4^2, Log-normal (0, 2), Weibull (1, 0.8) ve Lojistik (1.2, 0.35) dağılımları ele alınmıştır. Genel olarak, SW2 test istatistiğinin güç bakımından SW1, χ^2 ve olabilirlik oran test istatistiklerine göre daha iyi olduğu görülmektedir. Ayrıca SW1, SW2 üstel dağılıma benzeyen diğer dağılımları ayırt etmekte daha iyi sonuç vermektedir. Aynı zamanda tüm test istatistikleri artmasına rağmen, bu artış miktarı SW1 ve SW2'de ki-kare ve olabilirlik oranına göre daha fazladır. Grup sayısı küçükken (k=4) tüm test istatistikleri birbirine yakın sonuçlar vermekte, grup sayısı arttıkça (k= 6, 10) SW1 ve SW2 ile χ^2 ve olabilirlik oran test istatistikleri arasındaki fark artmakta ve χ^2 en kötü sonucu vermektedir. Bu durum hücrelerin beklenen değerinin 5'ten küçük olmasından kaynaklanmaktadır. Örneğin, n=20, k=10 durumunda gözlenen grup frekansları 2 olacaktır. Bu durumda ki-kare ve olabilirlik oranı, yapıları gereği iyi sonuç vermeyecektir.

Şekil 6, 7, 8'de kullanılan test istatistikleri, testlerin gücü bakımından dağılım biçiminin simetrik olduğu durumlarda; Normal (2, 2), Tekdüze (0, 2.5), Beta (2, 2) incelenmiştir. Test istatistikleri güç bakımından Normal (2, 2) ve Tekdüze (0, 2.5) dağılımlarında birbirine yakın sonuçlar vermiştir. Normal dağılımda χ^2, olabilirlik oranı ve SW2 en iyi sonucu verirken, SW1 en kötü sonucu vermektedir. Tekdüze dağılımında ise, bunun tersi olarak SW1 en iyi sonucu verirken, olabilirlik oranı en kötü sonucu vermektedir. Beta dağılımında ise, test istatistikleri bütün grup düzeyleri için yüksek güç değerlerine sahip olup, SW1 ve SW2 diğer test istatistiklerine göre daha iyi sonuç vermektedir.

Sola çarpık dağılımlardan; Weibull (1.5, 8) ve Gumbel (2) dağılımları incelendiğinde biçim olarak üstel dağılıma benzemediğinden en düşük örnek çapında (n=20) bile tüm test istatistiklerinin, gücü 1'e yakın çıkmaktadır.

4. TARTIŞMA VE SONUÇ

Grup sayısı ve örneklem çapı arttıkça, tüm test istatistiklerinin özellikle de SW1 ve SW2 test istatistiklerin güç değerleri artmaktadır. Sağa çarpık alternatif dağılımlardan üretilen verilerin üstel dağılıma uygunluğunun testinde, SW1 ve SW2 test istatistikleri daha iyi sonuç verirken, simetrik alternatif dağılımlardan üretilen verilerin üstel dağılıma uygunluğunun testinde SW1 ve SW2'nin, Ki-kare ve olabilirlik oranı test istatistiklerinden üstün olmadığı görülmüştür. Sola çarpık dağılımlardan üretilen veriler için ise, tüm test istatistikleri çok iyi sonuçlar vermiştir. Bu çalışma, farklı dağılımlar için test istatistiği geliştirilerek, genişletilebilir.

5. KAYNAKLAR

Baklizi, A., (2006). Weighted Kolmogrov-Smirnov type tests for grouped Rayleigh data. Applied Mathematical Modelling 30:437-445.

Best, D. J., Rayner, J.C.W. (2007). Chi-Squared components for tests of fit and improved models fort he grouped exponential distribution. Computational Statistics and Data Analysis 51:3946-3954.

Conover, W. J. (1972). A Kolmogrov goodness-of-fit test for discontinuous distributions. Journal of the American Statistical Association 67:591-596.

Gulati, S., Neus, J., (2003). Goodness of fit statistics for the exponential distribution when the data grouped. Comm. Statist. Theory Methods 32, 681-700.

Kulldorf, G., (1961). Estimation from grouped and partially grouped samples. New York: John Wiley & Sons.

Pearson, K., (1900). On a criterion that a given system of deviations from the probable in the case of correlated system of variable is such that it can be reasonably supposed to have arisen from random sampling, Philos. Mag., 5th series 50 (1900) 157-175.

Schmid, P., (1958). On the Kolmogrov and Smirnov limit theorems for discontinuous distribution functions. Annals of Mathematical Statistics 29:1011-1027.

Seo, S. K., Yum, B. J. (1993). Estimation methods for the mean of the exponential distribution based on grouped and censored data. IEEE Transactions on Reliability, 42 (1):87-96.

COMPARISONS OF WEIGHTED KOLMOGROV-SMIRNOV, LIKELIHOOD RATIO AND CHI-SQUARE GOODNESS OF FIT TESTS FOR THE EXPONENTIAL DISTRIBUTION BASED ON THE GROUPED DATA

ABSTRACT

In this paper, Weighted Kolmogrov-Simirnov test statistics which are used to test whether the grouped data fits to exponential distrubution proposed by Gulati and Neus (2003) are defined. These test statistics are compared with Likelihood ratio and Chi-square goodness of test statistic in terms of power under diffrent alternative distribution, group size and sample size. The simulation results showed that specially for right skewed distributuions Weighted Kolmogrov-Smirnov test statistics outperformed Pearson Chi-Square test statistics in terms of power.

Keywords: Bootstrap, Goodness of fit, Grouped data, Likelihood ratio, Pearson Chi-Square, Weighted Kolmogrov-Smirnov.