Volume: 09 Number: 01 Category: 01 Page: 11-20 ISSN No: 1303-6319 Cilt: 09 Sayı: 01 Kategori: 01 Sayfa: 11-20

GAMMA, WEİBULL VE LOG-NORMAL DAĞILIMLARININ DOĞRU SEÇİM OLASILIKLARINA GÖRE AYRIŞTIRILMASI

Hayrinisa DEMİRCİ BİÇER* **Cemal ATAKAN****

ÖZET

Gamma, Weibull ve Log-Normal dağılımları, çarpık verilerin analizi için sık kullanılan dağılımlardır. Bu çalışmada, verilen bir veri setinin, Gamma, Weibull ya da Log-Normal dağılımlardan hangisi ile modelleneceği problemi üzerinde durulmuştur. Verilen bir veri setinin dağılımının Gamma, Weibull ya da Log-Normal dağılımından hangisinden geldiğine, veri setinin dağılımının sırasıyla ilgili dağılımlara göre oluşturulan yokluk hipotezi altında Monte-Carlo simülasyonları ve asimptotik sonuçlardan hesaplanan doğru seçim olasılıkları ile karar verilmiştir.

Anahtar Kelimeler: Doğru seçim olasılığı, Gamma dağılımı, Log-Normal dağılım, Olabilirlik oranı, Weibull dağılımı.

1. GİRİŞ

Gamma, Weibull ve Log-Normal dağılımları, özellikle, mühendislik, sağlık ve fen alanlarında çarpık verilerin analizi için sık kullanılan dağılımlardır ve oldukça yaygın bir uygulama alanlarına sahiptirler.

Literatürde, verilen bir veri setinin iki olasılık dağılımından hangisi ile modelleneceği ile ilgili birçok çalışma mevcuttur. Atkinson (1969, 1970), Cox (1961, 1962), Dyer (1973) kitleden alınan örneklemin, ele aldıkları iki olasılık dağılımından hangisi ile modelleneceği problemini ele almışlardır. Dumonceaux ve Antle (1973) Log-Normal ve Weibull dağılımlarını ayırt etmek için en çok olabilirlik oranına ilişkin kritik değerler elde etmişlerdir. Bain ve Englehardt (1980) simülasyon sonuçlarına göre Weibull ve Gamma dağlımlarına ilişkin doğru seçim olasılıklarını elde etmişlerdir. Wiens (1999) bir veri setinin dağılımının Log-Normal veya Gamma dağılımı olması durumlarını incelemiştir.

Veri setinin, ortak uygulama alanlarına sahip olan bu dağılımlardan hangisi ile modelleneceği oldukça önemlidir. Bu çalışmada, örneklemin dağılımının Gamma, Weibull ya da Log-Normal dağılımlarının hangisi ile modelleneceği problemi üzerinde durulmuştur. Elde edilen doğru seçim olasılığına göre, veri setinin dağılımının Gamma, Weibull ya da Log-Normal dağılımlarından hangisine uyduğuna karar verilmektedir. Doğru seçim olasılığı, en çok olabilirlik oranının dağılımına göre elde edilebilir. Ancak, bu oranın dağılımına ilişkin analitik ifadeler henüz elde edilemediği için, doğru seçim olasılığı olabilirlik oranının logaritmasının asimptotik dağılımına göre elde edilebilir. Kitleden alınan örneklemin dağılımının sırasıyla ilgili dağılımlara göre oluşturulan yokluk hipotezi altında elde edilen asimptotik dağılımlarına ve Monte-Carlo simülasyonlarında elde edilen sonuçlara göre doğru seçim olasılıkları hesaplanmıştır.

---

*Arş. Gör. Dr., Kırıkkale Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, Kırıkkale, e-posta:hdbicer@hotmail.com

**Doç. Dr., Ankara Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 06100 Ankara, e-posta: atakan@science.ankara.edu.tr

2. EN ÇOK OLABİLİRLİK ORANLARI

Verilen bir veri setinin dağılımının, Gamma, Weibull ya da Log-Normal dağılımlarından, hangisine uyduğuna karar verebilmek için bu dağılımlar ile ilgili hipotezler altında en çok olabilirlik fonksiyonlarının oranına bağlı bir karar kriteri elde edilebilir.

α ve β parametreli Gamma dağılımına sahip X rasgele değişkeninin olasılık yoğunluk fonksiyonu,

$$f_{GA}(x;\alpha,\beta)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}\ ,x>0,\alpha>0,\beta>0 \tag{1}$$

olmak üzere, α ve β parametrelerinin en çok olabilirlik tahmin edicileri arasındaki ilişki

$$\hat{\beta}=\frac{\overline{X}}{\hat{\alpha}} \tag{2}$$

olarak elde edilir.

η ve σ parametreli Log-Normal dağılıma sahip X rasgele değişkenin olasılık yoğunluk fonksiyonu,

$$f_{LN}(x;\sigma,\eta)=\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\ln\eta)^2}{2\sigma^2}}\ ,x>0,\ \sigma>0,\ \eta>0 \tag{3}$$

dır. η ve σ parametrelerinin en çok olabilirlik tahmin edicileri,

$$\hat{\eta}=\left(\prod_{i=1}^{n}X_i\right)^{\frac{1}{n}}\quad \text{ve}\quad \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(\ln X_i-\ln\hat{\eta})^2 \tag{4}$$

biçiminde elde edilir.

θ ve λ parametreli Weibull dağılımına sahip X rasgele değişkeninin olasılık yoğunluk fonksiyonu,

$$f_{WE}(x;\theta,\lambda)=\theta\lambda^{\theta}x^{\theta-1}e^{-(x\lambda)^{\theta}}\ ,x>0,\ \theta>0,\ \lambda>0 \tag{5}$$

olmak üzere, θ ve λ parametrelerinin en çok olabilirlik tahmin edicileri arasındaki ilişki

$$\hat{\lambda}=\left(n\Big/\sum_{i=1}^{n}X_i^{\hat{\theta}}\right)^{1/\hat{\theta}} \tag{6}$$

olarak elde edilir.

En çok olabilirlik oranının logaritması ile kitleden alınan örneklemin dağılımının Gamma ya da Log-Normal dağıldığına karar verilebilir. Yokluk ve alternatif hipotezler sırasıyla,

H_0 : Örneklem Gamma dağılımından alınmıştır.
H_1 : Örneklem Log-Normal dağılımından alınmıştır.

olmak üzere, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}$ ve $\hat{\eta}$ en çok olabilirlik tahmin edicilerine bağlı olarak en çok olabilirlik oranının logaritması,

$$T_1 = \ln\left[\frac{L_{GA}\left(\hat{\alpha},\hat{\beta}\right)}{L_{LN}\left(\hat{\sigma},\hat{\eta}\right)}\right] = n\left[\ln\left(\Gamma\left(\hat{\alpha}\right)\right) - \ln\left(\hat{\sigma}\right) - \hat{\alpha}\ln\left(\frac{\tilde{X}}{\hat{\beta}}\right) + \hat{\alpha} - \frac{1}{2}\left(1+\ln\left(2\pi\right)\right)\right] \tag{7}$$

olarak elde edilir. Burada $\tilde{X} = \left(\prod_{i=1}^{n} X_i\right)^{1/n}$ örneklemin geometrik ortalamasıdır. Eğer $T_1 > 0$ ise veri setinin dağılımının Gamma dağılımından, aksi halde Log-Normal dağılımından olduğuna karar verilir. H_0'ın doğru olduğu koşulu altında (7) eşitliği ile verilen T_1'in dağılımı β'dan bağımsızdır ve sadece α parametresine bağlıdır. Benzer olarak, H_1'in doğru olduğu koşulu altında ise, T_1'in dağılımı sadece σ'ya bağlıdır.

Benzer biçimde, veri setinin dağılımının Log-Normal ya da Weibull dağıldığına karar verilebilir. Bu durumda, yokluk ve alternatif hipotezler sırasıyla,

H_0 : Örneklem Log-Normal dağılımından alınmıştır.
H_1 : Örneklem Weibull dağılımından alınmıştır.

olmak üzere, $\hat{\sigma}$, $\hat{\eta}$, $\hat{\theta}$ ve $\hat{\lambda}$ en çok olabilirlik tahmin edicilerine bağlı olarak en çok olabilirlik oranının logaritması,

$$T_2 = \ln\left[\frac{L_{LN}\left(\hat{\sigma},\hat{\eta}\right)}{L_{WE}\left(\hat{\theta},\hat{\lambda}\right)}\right] = n\left[\frac{1}{2} - \ln\left(\hat{\sigma}\hat{\theta}\left(\hat{\lambda}\hat{\eta}\right)^{\hat{\theta}}\sqrt{2\pi}\right)\right] \tag{8}$$

olarak elde edilir. Eğer $T_2 > 0$ ise örneklemin Log-Normal dağılımdan, aksi halde Weibull dağılımından geldiğine karar verilir. Eğer veri Log-Normal dağılımından geliyor ise, (8) eşitliği ile verilen T_2'nin dağılımı (σ,η)'dan bağımsızdır. Benzer olarak, eğer veri Weibull dağılımından geliyor ise, T_2'nin dağılımı (θ,λ)'dan bağımsızdır.

Örneklemin dağılımının Gamma ya da Weibull dağıldığına karar verilebilmesi için, yokluk ve alternatif hipotezler sırasıyla,

H_0 : Örneklem Gamma dağılımından alınmıştır.
H_1 : Örneklem Weibull dağılımından alınmıştır.

olmak üzere, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\theta}$ ve $\hat{\lambda}$ en çok olabilirlik tahmin edicilerine bağlı olarak en çok olabilirlik oranının logaritması

$$T_3 = \ln\left[\frac{L_{GA}\left(\hat{\alpha},\hat{\beta}\right)}{L_{WE}\left(\hat{\theta},\hat{\lambda}\right)}\right] = n\left[\hat{\alpha}\ln\left(\hat{\alpha}\tilde{X}/\bar{X}\right) - \hat{\theta}\ln\left(\hat{\lambda}\tilde{X}\right) - \ln\left(\hat{\theta}\Gamma\left(\hat{\alpha}\right)\right) - \hat{\alpha} + 1\right] \tag{9}$$

olarak elde edilir. Burada $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ örneklemin aritmetik ortalamasıdır.

Eğer $T_3 > 0$ ise örneklemin dağılımının Gamma dağılımı, aksi halde Weibull dağılımından olduğuna karar verilir.

3. EN ÇOK OLABİLİRLİK ORANLARININ ASİMPTOTİK DAĞILIMLARI

Bu bölümde ilgili veri setinin iki dağılımdan hangisinden geldiğine karar vermek için asimptotik sonuçlara bağlı karar kuralı elde edilmeye çalışılacaktır. Bu amaçla, karşılaştırmaya ilişkin hipotezler

A) A1. $H_0 : X \sim f_{GA}\left(x;\alpha,\beta\right)$ vs $H_1 : X \sim f_{LN}\left(x;\sigma,\eta\right)$
A2. $H_0 : X \sim f_{LN}\left(x;\sigma,\eta\right)$ vs $H_1 : X \sim f_{GA}\left(x;\alpha,\beta\right)$
B) B1. $H_0 : X \sim f_{LN}\left(x;\sigma,\eta\right)$ vs $H_1 : X \sim f_{WE}\left(x;\theta,\lambda\right)$
B2. $H_0 : X \sim f_{WE}\left(x;\theta,\lambda\right)$ vs $H_1 : X \sim f_{LN}\left(x;\sigma,\eta\right)$
C) C1. $H_0 : X \sim f_{GA}\left(x;\alpha,\beta\right)$ vs $H_1 : X \sim f_{WE}\left(x;\theta,\lambda\right)$
C2. $H_0 : X \sim f_{WE}\left(x;\theta,\lambda\right)$ vs $H_1 : X \sim f_{GA}\left(x;\alpha,\beta\right)$

olarak göz önüne alınsın. H_0 yokluk hipotezleri altında en çok olabilirlik oranının logaritmasının asimptotik dağılımlarının elde edilmesinde kullanılacak olan Lemma ve Teorem, A1 durumu göz önüne alınarak, verinin Gamma dağılımına sahip olduğu varsayımı altında verilmiştir. Diğer durumlar için de benzer sonuçlar aynı şekilde elde edilir.

Lemma : $X_1, X_2, \ldots, X_n$ rasgele örneklemi Gamma dağıldığında $n \to \infty$ için aşağıdaki asimptotik sonuçlar elde edilir. Burada a.s. ile hemen hemen her yerde yakınsama ifade edilmektedir.

a. $E_{GA}\left(\ln f_{GA}\left(X;\alpha,\beta\right)\right) = \max_{\bar{\alpha},\bar{\beta}} E_{GA}\left(\ln f_{GA}\left(X;\bar{\alpha},\bar{\beta}\right)\right)$

olduğunda $\hat{\alpha} \to \alpha$ a.s. ve $\hat{\beta} \to \beta$ a.s. dır.

b. $E_{GA}\left(\ln f_{LN}\left(X;\tilde{\sigma},\tilde{\eta}\right)\right)=\max_{\sigma,\eta} E_{GA}\left(\ln f_{LN}\left(X;\sigma,\eta\right)\right)$
olduğunda $\hat{\sigma}\to\tilde{\sigma}$ a.s. ve $\hat{\eta}\to\tilde{\eta}$ a.s. dır.

c. $T_1^*=\ln\left[\frac{L_{GA}\left(\alpha,\beta\right)}{L_{LN}\left(\tilde{\sigma},\tilde{\eta}\right)}\right]$ olarak tanımlansın. $\frac{1}{\sqrt{n}}\left[T_1-E_{GA}\left(T_1\right)\right]$ ve $\frac{1}{\sqrt{n}}\left[T_1^*-E_{GA}\left(T_1^*\right)\right]$ aynı asimptotik dağılıma sahiptirler.

TEOREM : Eğer veri Gamma dağılımına sahip ise, T_1, $E_{GA}\left(T_1\right)$ ortalama ve $Var_{GA}\left(T_1\right)$ varyansı ile asimptotik normal dağılır.

İspat: Merkezi Limit Teoremi kullanılarak $\frac{1}{\sqrt{n}}\left[T_1-E_{GA}\left(T_1\right)\right]$'nin asimptotik normal dağıldığı gösterilebilir. Böylece, Merkezi Limit Teoremi ve Lemma c. den ispat tamamlanmış olur (White, 1982).

Lemma b'den, $\tilde{\sigma}$, $\tilde{\eta}$, $E_{GA}\left(T_1\right)$ ve $Var_{GA}\left(T_1\right)$ ifadelerini elde etmek için

$$\begin{aligned} g\left(\sigma,\eta\right) &= E_{GA}\left(\ln\left(f_{LN}\left(X;\sigma,\eta\right)\right)\right)\\ &= E_{GA}\left[-\frac{1}{2}\ln\left(2\pi\right)-\ln\sigma-\psi\left(\alpha\right)-\ln X-\frac{1}{2\sigma^2}\left(\ln X-\ln\eta\right)^2\right]\\ &= -\frac{1}{2}\ln\left(2\pi\right)-\ln\left(\sigma\beta\right)-\psi\left(\alpha\right)-\frac{1}{2\sigma^2}\left[\psi'\left(\alpha\right)+\left(\psi\left(\alpha\right)+\ln\left(\beta/\eta\right)\right)^2\right] \qquad (10)\end{aligned}$$

fonksiyonunu tanımlayalım. $g\left(\sigma,\eta\right)$'nın sırasıyla σ ve η'ya göre türevlerinin sıfıra eşitlenmesi ile

$$\tilde{\sigma}=\left(\psi'\left(\alpha\right)\right)^{1/2} \text{ ve } \tilde{\eta}=\beta e^{\psi(\alpha)} \qquad (11)$$

olarak elde edilir. Burada $\psi\left(\alpha\right)=\frac{d}{d\alpha}\ln\left(\Gamma\left(\alpha\right)\right)$ ve $\psi'\left(\alpha\right)=\frac{d}{d\alpha}\psi\left(\alpha\right)$ dır [4].

T_1 'nin asimptotik beklenen değeri ve varyansı;

$$\begin{aligned} \frac{E_{GA}\left(T_1\right)}{n} \approx AE_{GA} &= E_{GA}\left[\ln f_{GA}\left(X;\alpha,1\right)-\ln f_{LN}\left(X;\tilde{\sigma},\tilde{\eta}\right)\right]\\ &= \frac{1}{2}\ln\left(2\pi\right)+\frac{1}{2}+\ln\tilde{\sigma}-\ln\left(\Gamma\left(\alpha\right)\right)+\alpha\left(\psi\left(\alpha\right)-1\right)\end{aligned}$$

ve

$$\frac{Var_{GA}(T_1)}{n} \approx AVar_{GA} = Var_{GA}\left[\ln f_{GA}(X;\alpha,1) - \ln f_{LN}(X;\tilde{\sigma},\tilde{\eta})\right]$$

$$\begin{aligned} &= -\frac{1}{\tilde{\sigma}^2}\Big[\alpha(\alpha+1)\big(\psi'(\alpha+2)+\psi(\alpha+2)\big)^2 \\ &\quad -\psi(\alpha)\big(1-\psi(\alpha)(\alpha+2)\big)-\alpha\psi''(\alpha) \\ &\quad -\frac{1}{4\tilde{\sigma}^2}\big(\psi'''(\alpha)-3\psi(\alpha)\psi''(\alpha)\big)\Big] \\ &\quad +\alpha+\alpha^2\tilde{\sigma}^2-\psi(\alpha)(\alpha-2)-\frac{3}{2} \end{aligned}$$

dir. Bu durumda, T_1, $E_{GA}(T_1)$ oratalama ve $Var_{GA}(T_1)$ varyanslı asimptotik normal dağıldığından, doğru seçim olasılığı; yani A1 de verilen H_0 hipotezinin red edilememe olasılığı

$$P(T_1 > 0) \approx \Phi\left(-\frac{E_{GA}(T_1)}{\sqrt{Var_{GA}(T_1)}}\right) = \Phi\left(-\frac{n \times AE_{GA}}{\sqrt{n \times AVar_{GA}}}\right) \qquad (12)$$

dir. Burada Φ standart normal rasgele değişkenin kümülatif dağılım fonksiyonudur.

4. SİMÜLASYON ÇALIŞMASI

Bu bölümde, farklı örneklem büyüklükleri için Bölüm 3 de elde edilen asimptotik sonuçların simülasyon çalışması ile Monte-Carlo simülasyon çalışmalarına bağlı olarak doğru seçim olasılıkları hesaplanmıştır.

A: H_0 hipotezi Gamma olduğu ve alternatif hipotezi Log-Normal olduğu durum ele alınmıştır. $L_{GA}(\alpha,2)$'den örneklem büyüklüğü $n = 20$, 40, 60, 80 ve 100 birimlik rasgele örneklemler 10000 tekrarlı üretilerek (7) eşitliği ile verilen T_1'in sıfırdan büyük olup olmadığı kontrol edilmiştir. Doğru seçim olasılığının bir tahmin değerini elde etmek için, T_1'nin sıfırdan büyük olduğu durumların sayısı hesaplanmıştır. Asimptotik sonuçlar kullanılarak (12) eşitliği ile verilen doğru seçim olasılıkları hesaplanarak, sonuçlar Tablo 1'de verilmiştir. H_0 hipotezinin Log-Normal olduğu durum için de $L_{LN}(\sigma,2)$'den örneklem büyüklüğü $n = 20$, 40, 60, 80 ve 100 birimlik rasgele örneklemler üretilerek, $T_1 < 0$ olup olmadığı kontrol edilmiştir. Sonuçlar Tablo 2'de verilmiştir. Her bir tabloda, ilk eleman Monte-Carlo simülasyonlarının sonucunu ve parantez içindeki değerler asimptotik sonuçlardan elde edilen doğru seçim olasılıklarını göstermektedir.

Tablo 1. Veri Gamma dağıldığında Monte-Carlo simülasyonları ve asimptotik sonuçlara bağlı doğru seçim olasılıkları

$\alpha \downarrow n \rightarrow$	20	40	60	80	100
	MC AS	MC AS	MC AS	MC AS	MC AS
2	0.7245(0.7216)	0.8023(0.7970)	0.8512(0.8456)	0.8797(0.8800)	0.9058(0.9055)
4	0.6549(0.6675)	0.7394(0.7299)	0.7844(0.7734)	0.8098(0.8068)	0.8354(0.8336)
6	0.6303(0.6419)	0.7083(0.6965)	0.7419(0.7356)	0.7652(0.7665)	0.7953(0.7919)
8	0.6399(0.6285)	0.6851(0.6786)	0.7251(0.7150)	0.7478(0.7440)	0.7673(0.7683)
10	0.6249(0.6154)	0.6718(0.6609)	0.7059(0.6943)	0.7245(0.7213)	0.7442(0.7441)

Tablo 2. Veri Log-Normal dağıldığında Monte-Carlo simülasyonları ve asimptotik sonuçlara bağlı doğru seçim olasılıkları

$\sigma \downarrow n \rightarrow$	20	40	60	80	100
	MC AS	MC AS	MC AS	MC AS	MC AS
0.5	0.6645(0.6827)	0.7493(0.7414)	0.7912(0.7848)	0.8197(0.8242)	0.8408(0.9583)
0.7	0.7149(0.7207)	0.7994(0.7959)	0.8494(0.8444)	0.8947(0.8789)	0.9054(0.9045)
0.9	0.7603(0.7509)	0.8453(0.8309)	0.8819(0.8796)	0.9152(0.9122)	0.9357(0.9350)
1.1	0.7699(0.7723)	0.8651(0.8545)	0.9051(0.9020)	0.9417(0.9323)	0.9510(0.9525)
1.3	0.7849(0.7876)	0.8798(0.8705)	0.9199(0.9166)	0.9495(0.9448)	0.9632(0.9628)

B: H_0 hipotezi Log-Normal olduğu ve alternatif hipotezi Weibull olduğu durum ele alınmıştır. $L_{LN}(2,2)$'den örneklem büyüklüğü $n = 20,\ 40,\ 60,\ 80$ ve 100 birimlik rasgele örneklemler üretilerek (8) eşitliği ile verilen $T_2 > 0$ olup olmadığı kontrol edilmiştir. Sonuçlar Tablo 3'de verilmiştir. H_0 hipotezi Weibull ve alternatif hipotez Log-Normal olduğu durum ele alınarak $L_{WE}(2,2)$'den n birimlik rasgele örneklemler üretilerek, $T_2 < 0$ olup olmadığı kontrol edilmiştir. Sonuçlar Tablo 4'de verilmiştir.

Tablo 3. Veri Log-Normal dağıldığında Monte-Carlo simülasyonları ve asimptotik sonuçlara bağlı doğru seçim olasılıkları

n	20	40	60	80	100
MC	0.7838	0.8668	0.9162	0.9421	0.9595
AS	0.7808	0.8632	0.9110	0.9394	0.9588

Tablo 4. Veri Weibull dağıldığında Monte-Carlo simülasyonları ve asimptotik sonuçlara bağlı doğru seçim olasılıkları

n	20	40	60	80	100
MC	0.7782	0.8660	0.9090	0.9389	0.9581
AS	0.7766	0.8590	0.9062	0.9360	0.9556

C: H_0 hipotezi Gamma olduğu ve alternatif hipotezi Weibull olduğu durum ele alınmıştır. Örneklem büyüklüğü $n = 20,\ 40,\ 60,\ 80$ ve 100 ve şekil parametresi $\alpha = 2,\ 4,\ 6,\ 8,\ 10$ ve 12 olarak seçilmiştir. $L_{GA}(\alpha,2)$'den n birimlik rasgele örneklem üretildi ve $T_3 > 0$ olup olmadığı kontrol edilmiştir. Sonuçlar Tablo 5'de verilmiştir. Benzer olarak, H_0 hipotezi Weibull ve alternatif hipotez Gamma olduğu durum ele alınarak, örneklem büyüklüğü $n = 20,\ 40,\ 60,\ 80\ \text{ve} 100$ ve $\theta = 2,\ 4,\ 6,\ 8$ ve 10 olduğu durumlar için $T_3 < 0$ olup olmadığı kontrol edildi. Sonuçlar Tablo 6'da verilmiştir.

Tablo 5. Veri Gamma dağıldığında Monte-Carlo simülasyonları ve asimptotik sonuçlara bağlı doğru seçim olasılıkları

$\alpha \downarrow n \rightarrow$	20	40	60	80	100
	MC AS	MC AS	MC AS	MC AS	MC AS
2	0.8914(0.8907)	0.9014(0.9088)	0.9203(0.9103)	0.9367(0.9329)	0.9601(0.9592)
4	0.8370(0.8384)	0.8420(0.8479)	0.8627(0.8595)	0.8024(0.8797)	0.8923(0.8924)
6	0.7185(0.7068)	0.7476(0.7465)	0.7568(0.7530)	0.7739(0.7681)	0.7839(0.7833)
8	0.5794(0.5812)	0.5750(0.5627)	0.5938(0.5912)	0.6163(0.6088)	0.5962(0.5965)
10	0.5468(0.5340)	0.5369(0.5217)	0.5456(0.5438)	0.5390(0.5343)	0.5320(0.5342)
12	0.5278(0.5115)	0.5073(0.5078)	0.5171(0.5102)	0.5115(0.5097)	0.5095(0.5096)

Tablo 6. Veri Weibull dağıldığında Monte-Carlo simülasyonları ve asimptotik sonuçlara bağlı doğru seçim olasılıkları

$\theta \downarrow n \rightarrow$	20	40	60	80	100
	MC AS	MC AS	MC AS	MC AS	MC AS
2	0.5495(0.5271)	0.5513(0.5386)	0.5578(0.5470)	0.5588(0.5544)	0.5615(0.5609)
4	0.5449(0.5224)	0.5501(0.5344)	0.5494(0.5435)	0.5509(0.5492)	0.5541(0.5530)
6	0.5273(0.5143)	0.5309(0.5218)	0.5369(0.5258)	0.5316(0.5300)	0.5339(0.5321)
8	0.5228(0.5103)	0.5198(0.5143)	0.5243(0.5176)	0.5204(0.5193)	0.5237(0.5229)
10	0.5184(0.5073)	0.5201(0.5099)	0.5248(0.5131)	0.5153(0.5148)	0.5165(0.5156)

5. TARTIŞMA VE SONUÇ

Bu çalışmada, verilen bir veri setinin, ortak uygulama alanlarına sahip olan Gamma, Weibull ya da Log-Normal dağılımlarından hangisine uyduğuna karar vermek için ilgili dağılımlara göre oluşturulan yokluk hipotezi altında Monte-Carlo simülasyonlarından ve asimptotik sonuçlardan doğru seçim olasılıkları elde edilmiştir.

Elde edilen tablolardan, örneklem büyüklüğü arttıkça doğru seçim olasılıklarının arttığı ve Monte-Carlo simülasyonu ile elde edilen doğru seçim olasılığı ile asimptotik sonuçlardan elde edilen doğru seçim olasılığı arasındaki farkın küçüldüğü görülmektedir.

Asimptotik sonuçlara göre elde edilen doğru seçim olasılıklarına bakıldığında büyük örneklemlerde olduğu gibi küçük örneklemler için de asimptotik sonuçların Monte-Carlo kadar iyi sonuçlar verdiği görülmektedir.

6. KAYNAKLAR

Atkinson, A., 1969. A Test of Discriminating Between Modes, Biometrika, 56, 337–341.

Atkinson, A., 1970. A Method for Discriminating Between Models (with discussions), Journal of the Royal Statistical society. Ser. B, 32, 323–353.

Bain, L. J., Englehardt, M., 1980. Probability of Correct Selection of Weibull Versus Gamma Based on Likelihood Ratio. Communications in Statistics. Series A, 9, 375–381.

Bernardo, J. M., 1976. Algorithm As 103: Psi (Digamma) Function, J. Roy. Statist. Soc. Ser. C, 25, 315-317.

Cox, D. R., 1961. Tests of Separate Families of Hypotheses, Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability, Berkeley, University of California Press, 105–123.

Cox, D. R., 1962. Further Results on Tests of Separate Families of Hypotheses, Journal of the Royal Statistical Society, Ser. B, 24, 406–424.

Dumonceaux, R., Antle, C.E., 1973. Discriminating Between the Log-normal and Weibull Distribution, Technometrics, 15, 923–926.

Dyer, A. R., 1973. Discrimination Procedure for Separate Families of Hypotheses, Journal of the American Statistical Association, 68, 970–974.

Johnson, N. L., Kotz, S., Balakrishnan, N., 1994. Continuous Univariate Distributions, Volume 1, John Wiley & Sons, Inc. USA.

Wiens, B.L., 1999. When Log-Normal and Gamma Models Give Different Results: A Case Study, American Statistician, 53, 89–93.

White, H., 1982. Reqularity Conditions for Cox’s Test of Non-nested Hypotheses, J. Econometrics, 19, 301–318.

DISCRIMINATION OF THE GAMMA, WEIBULL AND LOG-NORMAL DISTRIBUTIONS ACCORDING TO PROBABILITY OF CORRECT SELECTION

ABSTRACT

Gamma, Weibull and Log-Normal distributions are often used distributions for modelling skewed data. In this study, for a given data set the problem of selecting Gamma, Weibull or Log-Normal distributions is focused on for the modelling. Distribution of a given set of data, Gamma, Weibull or Log-Normal distribution has been decided by the probability of correct selection that is calculated by constructing Monte-Carlo simulations and asymptotic results under the null hypothesis.

Keywords: Probability of correct selection, Gamma distribution, Log-Normal distribution, Likelihood ratio tests, Weibull distribution.