
GEOMETRİK SÜREÇ VERİLERİ İÇİN GAMMA VE WEİBULL DAĞILIMLARI ARASINDAKİ AYRIM

Çenker BİÇER¹

Hayrinisa DEMİRCİ BİÇER²

Öz

Uygulamalarda gözlemlenen verilerin en uygun biçimde istatistiksel analizini yapmak için veri kümesinin altında yatan dağılım en uygun biçimde belirlenmelidir. Çoğu zaman, bir veri kümesinin altında yatan dağılımı belirlemeye çalışırken kullanılan uyum iyiliği testleri, veri seti için birden fazla dağılım modelini işaret eder. Uyum iyiliği testlerinin sonuçlarına göre olası dağılım modelleri arasında, veri kümesi için optimal dağılım modelinin belirlenmesi problemi, istatistikte oldukça önemli bir problemdir. Bu çalışmada, geometric süreç verileri için Gamma ve Weibull dağılımları arasındaki ayırım problemi, en çok olabilirlik oran yöntemine göre araştırılmıştır. Ayrıcılık için kullanılan yöntemin doğru seçim performansını göstermek için, kapsamlı bir simülasyon çalışması yapılmış ve belirli bir güven düzeyinde ve test gücünde ayırım yapmak için gerekli minimum örneklem büyüklükleri elde edilmiştir. Buna ek olarak, açıklayıcı amaçlarla, gerçek bir veri seti kullanılarak bir uygulama yapılmıştır.

Anahtar Kelimeler: Geometrik Süreç, En Çok Olabilirlik Oranı, Gamma Dağılımı, Weibull Dağılımı
JEL Sınıflandırması: C630, C690, C020

DISCRIMINATING BETWEEN THE GAMMA AND WEIBULL DISTRIBUTIONS FOR GEOMETRIC PROCESS DATA

Abstract

To obtain an optimal statistical analysis of the data observed in the applications, the underlying distribution of the data set should be optimally determined. Most of times, the goodness-of-fit tests used when trying to determine the underlying distribution of a data set indicate more than one distribution model to data set. Among the possible distribution models according to the results of the goodness-of-fit tests, the problem of determination of the optimal distribution model for the data set is quite important problem in statistic. In this study, the problem of discriminating between the Gamma and Weibull distributions for geometric process data is investigated according to the ratio of maximum likelihood method. To show the correct selection performance of the method used for discrimination, a comprehensive simulation study is performed and in order to discriminating at fixed level of confidence and power of test, required minimum sample sizes are obtained. In addition, for illustrative purposes, an application is made by using a real data set.

Keywords: Geometric Process, Maximum Likelihood Ratio, Gamma Distribution, Weibull Distribution
JEL Classification: C630, C690, C020

¹ Yrd. Doç. Dr., Kırıkkale Üniversitesi, Fen Ed. Fak., İstatistik Bölümü, cbicer@kku.edu.tr

² Yrd. Doç. Dr., Kırıkkale Üniversitesi, Fen Ed. Fak., İstatistik Bölümü, hdbicer@hotmail.com

1. Giriş

Geometrik süreç yenileme sürecinin bir genellemesi olarak, trend içeren ardışık varış zamanlarını modellemek için Lam (1988) tarafından tanıtılmıştır. $\{N(t), t > 0\}$ stokastik süreci bir sayma süreci ve $\{X_n, n = 1, 2, \dots\}$ bu süreçten gelen bağımsız rasgele değişkenlerinin bir dizisi olmak üzere; $\{a^{n-1}X_n, n = 1, 2, \dots\}$ dizisi bağımsız ve aynı dağılımlı rasgele değişkenlerin herhangi bir dizisi olacak şekilde bir $a > 0$ sayısı var ise $\{N(t), t > 0\}$ sayma sürecinden gelen $\{X_n, n = 1, 2, \dots\}$ bağımsız rasgele değişkenlerinin dizisinin oluşturduğu sürece a oranlı bir geometrik süreç denir (Lam, 2007:38; Kara vd., 2015). X_1, X_2, \dots, X_n rasgele değişkenleri a oranlı bir geometrik süreçten bir örnek olmak üzere, X_1 rasgele değişkeni bu süreçte ilk olay gerçekleşinceye kadar geçen süreyi temsil etmektedir. Eğer X_1 rasgele değişkeninin dağılımı $f_{X_1}(x_1)$ ise bu durumda $X_i, i = 1, 2, \dots, n$ rasgele değişkenlerinin dağılımı

$$f_{X_i}(x_i) = a^{i-1}f_{X_1}(a^{i-1}x_i) \quad (1)$$

olmaktadır. Diğer taraftan

$$Y_i = a^{i-1}X_i, \quad i = 1, 2, \dots, n \quad (2)$$

dönüşümü ile tanımlanan Y_i rasgele değişkenleri birbirinden bağımsız ve X_1 rasgele değişkeni ile aynı dağılımlıdır. Ayrıca, μ ve σ^2 sırasıyla X_1 rasgele değişkeninin beklenen değer ve varyansı olmak üzere, eşitlik (2)'nin göz önüne alınmasıyla $X_i, i = 1, 2, \dots, n$ rasgele değişkeninin beklenen değer ve varyansı, sırasıyla

$$E(X_i) = \frac{\mu}{a^{i-1}}, \quad i = 1, 2, \dots, n \quad (3)$$

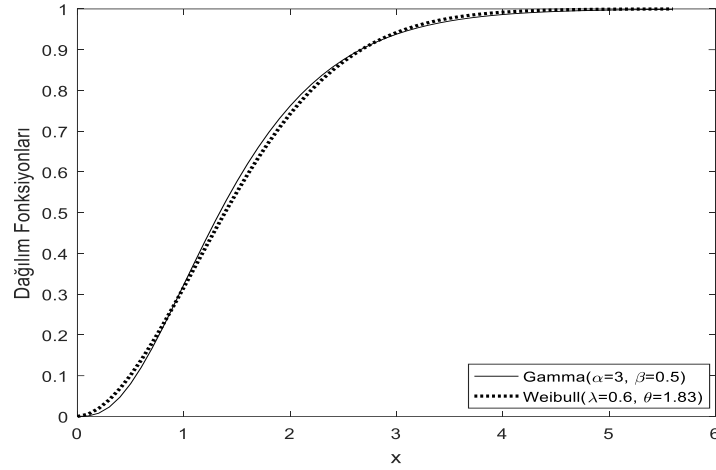
ve

$$Var(X_i) = \frac{\sigma^2}{a^{2i-2}} \quad (4)$$

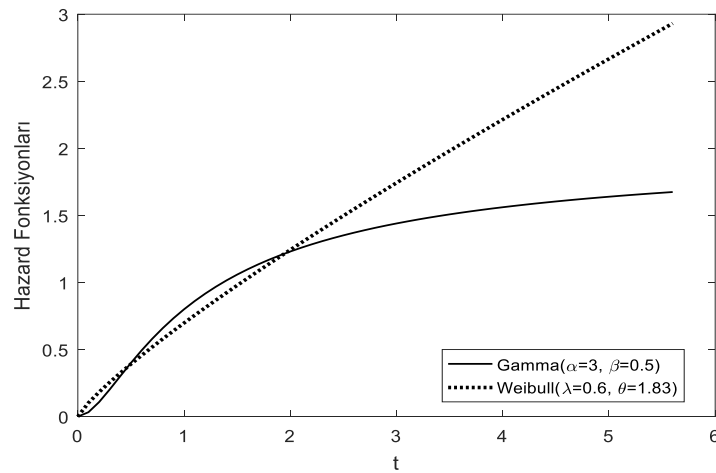
olarak kolayca hesaplanır. Eşitlik (3) ve eşitlik (4)'den X_2, X_3, \dots, X_n rasgele değişkenlerinin beklenen değer ve varyansları yalnızca X_1 rasgele değişkeninin beklenen değer ve varyansı ile ifade edilebilmektedir. Bununla birlikte, geometrik süreçlerde parametrik sonuç çıkarım X_1 rasgele değişkeninin dağılımına dayalı olarak yapılabilir. Dolayısıyla Geometrik süreç ile uyumlu bir veri setinin istatistiksel analizi için, X_1 rasgele değişkeninin dağılımın doğru veya doğruya en yakın biçimde belirlenmesi oldukça önemlidir.

İstatistikte, gözlenmiş bir veri setinin dağılımının belirlenmesinde Ki-Kare, Kolmogorov-Smirnov, Anderson-Darling vb. bazı uyum iyiliği testleri en sık başvurulan yöntemlerdendir. Ancak bu testler genellikle bir veri seti için birden fazla dağılım modeline işaret eder. Her ne kadar veri seti için olası dağılımlar, veri setini modellerken benzer veri uyumu gösterebilirler, model varsayımından dolayı kuyruk olasılıkları ve hazard fonksiyonları gibi bazı karakteristikler önemli derecede etkilenebilmektedir. Bu durumu açıklayabilmek için Gamma(3, 0.5) ve Weibull(0.6, 1.83) dağılımlarını göz önüne alalım. Şekil 1 ve Şekil 2 sırasıyla ilgili dağılımlara ait dağılım fonksiyonlarının ve hazard fonksiyonlarının grafiklerini göstermektedir.

Şekil 1: Gamma(3, 0.5) ve Weibull(0.6, 1.83) Dağılımlarının Dağılım Fonksiyonları



Şekil 2: Gamma(3, 0.5) ve Weibull(0.6, 1.83) Dağılımlarının Hazard Fonksiyonları



Şekil 1'den görülebileceği gibi her iki dağılım fonksiyonu neredeyse aynı veri uyumu sağlamaktadır, ancak Şekil 2'de hazard fonksiyonlarının birbirinden tamamen farklı oldukları görülmektedir. Dolayısıyla, bir veri setinden optimal sonuç çıkarım için en uygun dağılım modeli belirlenmelidir. En çok olasılık oran yöntemi (RML), veri setini modellemek için kullanılabilecek olası dağılımlar arasından optimal dağılımın belirlenmesi probleminin çözümü için kullanılan yöntemlerden bir tanesidir. Literatürde RML yöntemi kullanılarak klasik yaklaşım altında, Gamma ve Weibull (Fearn ve Nebenzahl, 1991), Log-Normal ve Gamma (Kundu ve Manglick, 2005), Weibull ve Log-Normal (Kundu ve Manglick, 2004), Weibull ve Log-Logistic (Elsherpiny vd, 2013), Log-Normal ve Genelleştirilmiş Üstel (Kundu vd., 2005), Genelleştirilmiş Rayleigh ve Log-Normal (Kundu ve Ragab, 2007), Genelleştirilmiş Rayleigh ve Weibull (Ragab ve Kundu, 2013) gibi bazı örtüşen dağılım çiftleri arasındaki seçim problemleri incelenmiştir. Verilerin ikinci tip sansürlü olması durumunda ise Weibull ve Log-Normal dağılımları arasındaki ayırım Dey ve Kundu (2012) tarafından ele alınmıştır. Bununla birlikte, bu problem geometrik bir süreçten elde edilen veri seti için açıktır.

Bu çalışma, ilk varış zamanının dağılımının Gama veya Weibull olduğu ve geometrik süreç ile uyumlu olan bir veri seti için Gamma ve Weibull dağılımlarını ayırt etme problemini RML yöntemine göre göz önüne almaktadır. Bu çalışmanın kalan kısımları ise şu şekilde düzenlenmiştir: 2. bölümde, geometrik süreç verileri için RML yöntemine dayalı bir ayırıcı istatistik verilmektedir. Araştırmacı tarafından istenilen anlam düzeyinde ve güçte bir ayırım işlemi gerçekleştirebilmek için gerekli

minimum örneklem büyüklüğü, Fearn ve Nebenzahl (1991) tarafından verilen asimptotik sonuçlara dayalı olarak, yine bu bölümde elde edilmektedir. Bölüm 3’de, çalışmada kullanılan ayırıcı istatistiğin ayırım performansı gerçekleştirilen bazı simülasyon çalışmaları ile incelenmektedir. Gerçek bir veri seti kullanılarak yapılan uygulamanın verildiği dördüncü bölümden sonra beşinci ve son bölüm sonuç bölümüdür.

2. Geometrik Süreç Verileri için RML

Varsayalım (X_1, X_2, \dots, X_n) a oranlı bir geometrik süreçten alınan rasgele bir örneklem ve X_1 , $\alpha > 0$ ve $\beta > 0$ parametreleriyle Gamma dağılımına sahip olsun. Bu durumda X_1 rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_{GA}(x_1; \alpha, \beta) = \frac{1}{\beta \alpha \Gamma(\alpha)} x_1^{\alpha-1} e^{-x_1/\beta}, x_1 > 0; \alpha > 0, \beta > 0 \quad (5)$$

dır. Burada β ve α sırasıyla dağılımın ölçek ve şekil parametresidir. Ayrıca $X_i, i = 1, 2, \dots, n$, rasgele değişkenleri için olabilirlik fonksiyonu, (1) eşitliğinin göz önüne alınmasıyla

$$\begin{aligned} L_{GA}(a_{GA}, \alpha, \beta; x_1, x_2, \dots, x_n) &= \prod_{i=1}^n a_{GA}^{i-1} f_{x_1; GA}(a_{GA}^{i-1} x_i; \alpha, \beta) \\ &= \frac{a_{GA}^{n(n-1)}}{\beta^{n\alpha} (\Gamma(\alpha))^n} \prod_{i=1}^n x_i^{\alpha-1} e^{-a_{GA}^{i-1} x_i/\beta} \end{aligned} \quad (6)$$

biçiminde yazılır. Benzer düşünce ile eğer X_1 rasgele değişkeni $\theta > 0$ ve $\lambda > 0$ parametreleriyle Weibull dağılımına sahip olsun. Bu durumda da X_1 rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_{WE}(x; \theta, \lambda) = \theta \lambda^\theta x^{\theta-1} e^{-(x\lambda)^\theta}, x > 0, \theta > 0, \lambda > 0 \quad (7)$$

olup burada θ dağılımın şekil parametresi λ ise dağılımın ölçek parametresidir. Gamma dağılımı durumunda olduğu gibi $X_i, i = 1, 2, \dots, n$, rasgele değişkenlerine ait olabilirlik fonksiyonu

$$\begin{aligned} L_{WE}(a_{WE}, \lambda, \theta; x_1, x_2, \dots, x_n) &= \prod_{i=1}^n a_{WE}^{i-1} f_{x_1; WE}(a_{WE}^{i-1} x_i; \lambda, \theta) \\ &= a_{WE}^{\frac{n(n-1)}{2}} \theta^n \lambda^{n\theta} \prod_{i=1}^n (a_{WE}^{i-1} x_i)^{\theta-1} e^{-\lambda^\theta \sum_{i=1}^n a_{WE}^{i-1} x_i} \end{aligned} \quad (8)$$

biçiminde yazılır.

a oran parametrelili geometrik süreçle uyumlu (X_1, X_2, \dots, X_n) rasgele örneklemde X_1 rasgele değişkeni için Gamma ve Weibull dağılımları arasındaki ayırmada ilk olarak

$$\begin{aligned} H_0: X_1 \sim GA(\alpha, \beta) \text{ veya denk olarak } Y_i \sim GA(\alpha, \beta) \\ H_1: X_1 \sim WE(\theta, \lambda) \text{ veya denk olarak } Y_i \sim WE(\lambda, \theta) \end{aligned} \quad (9)$$

hipotezi göz önüne alınsın. (9) ile verilen hipotezi test edebilmek için RML yöntemine dayalı bir ayırıcı istatistik, eşitlik (6) ile verilen olabilirlik fonksiyonunun (8)’da verilen olabilirlik fonksiyonuna oranlanması ve logaritmasının alınmasıyla

$$T = \ln \left(\frac{L_{GA}(a_{GA}, \alpha, \beta; x_1, x_2, \dots, x_n)}{L_{WE}(a_{WE}, \lambda, \theta; x_1, x_2, \dots, x_n)} \right) \quad (10)$$

olarak yazılabilir. (10) ile verilen istatistik, parametrelerin tahminlerine bağlı olarak

$$T = \ln \left(\frac{L_{GA}(\hat{a}_{GA}, \hat{\alpha}, \hat{\beta}; x_1, x_2, \dots, x_n)}{L_{WE}(\hat{a}_{WE}, \hat{\lambda}, \hat{\theta}; x_1, x_2, \dots, x_n)} \right) = \ln \left[\prod_{i=1}^n \frac{\hat{a}_{GA}^{i-1} f_{GA}(\hat{a}_{GA}^{i-1} x_i; \hat{\alpha}, \hat{\beta})}{\hat{a}_{WE}^{i-1} f_{WE}(\hat{a}_{WE}^{i-1} x_i; \hat{\lambda}, \hat{\theta})} \right] \quad (11)$$

biçiminde yazılır. Burada $\hat{\alpha}, \hat{\beta}, \hat{\theta}$ ve $\hat{\lambda}$ ilgili parametrelere ait en çok olabilirlik tahmin edicileri, \hat{a}_{WE} ve \hat{a}_{GA} ise sırasıyla Weibull ve Gamma dağılımları durumu için geometrik sürecin oran parametresinin en çok olabilirlik tahmin edicisidir.

Eşitlik (11) ile verilen T istatistiğinde \hat{a}_{WE} ve \hat{a}_{GA} yerine Lam ve Chan (1998) tarafından önerilen

$$\hat{\alpha} = \exp\left(\frac{6}{(n-1)n(n+1)} \sum_{i=1}^n (n-2i+1) \ln X_i\right) \quad (12)$$

parametrik olmayan tahmin edicisi kullanılırsa

$$T = \ln \left[\prod_{i=1}^n \frac{f_{GA}(\hat{\alpha}^{i-1} x_i; \hat{\alpha}, \hat{\beta})}{f_{WE}(\hat{\alpha}^{i-1} x_i; \hat{\theta}, \hat{\lambda})} \right] \quad (13)$$

olur. Ayrıca, (12) eşitliği ile verilen $\hat{\alpha}$ tahmin edicisinin eşitlik (6) ve (8) da kullanılmasıyla $\hat{\alpha}$ ve $\hat{\beta}$ tahmin edicileri arasındaki ilişki

$$\hat{\alpha} = \frac{\sum_{i=1}^n \hat{\alpha}^{i-1} x_i}{n \hat{\beta}} \quad (14)$$

olarak elde edilir. $\hat{\theta}$ ve $\hat{\lambda}$ tahmin edicileri arasındaki ilişki ise

$$\hat{\lambda} = \left(\frac{n}{\sum_{i=1}^n (\hat{\alpha}^{i-1} x_i)^{\hat{\theta}}} \right)^{\frac{1}{\hat{\theta}}} \quad (15)$$

biçiminde elde edilir. Eşitlik (14)-(15) ile verilen ilişkiler çerçevesinde eşitlik (12) ile verilen T istatistiği

$$T = n \left[\hat{\alpha} \ln(\hat{\alpha} \bar{X}_a / \bar{X}_a) - \hat{\theta} \ln(\hat{\lambda} \bar{X}_a) - \ln(\hat{\theta} \Gamma(\hat{\alpha})) - \hat{\alpha} + 1 \right] \quad (16)$$

olarak elde edilir. Burada $\bar{X}_a = (\prod_{i=1}^n \hat{\alpha}^{i-1} X_i)^{1/n}$ ve $\bar{X}_a = \frac{1}{n} \sum_{i=1}^n \hat{\alpha}^{i-1} X_i$ dir. (16) eşitliği ile elde edilen T istatistiğinin değerine bağlı olarak (9) ile verilen hipotez için

$$\begin{cases} T > 0, \text{ ise } H_0 \text{ hipotezi kabul edilir} \\ T < 0, \text{ ise } H_0 \text{ hipotezi reddedilir} \end{cases} \quad (17)$$

biçiminde oluşturulan bir karar kuralına göre karar verilir (Bain ve Engelhardt, 1980). Ancak T istatistiği kullanılarak gerçekleştirilen hipotez testinde, T istatistiğinin dağılımı bilinmediğinden dolayı testin gücü veya anlamlılık düzeyi hakkında bir şey söylenememektedir. Bir diğer deyişle yapılan seçimin doğru olması olasılığının hesaplanabilmesi için T 'nin dağılımının belirlenmesi gerekmektedir. T istatistiğinin dağılımı henüz açık olarak elde edilememekle birlikte doğru seçim olasılığının yaklaşık olarak hesaplanabilmesi ve istenen bir anlam düzeyinde test gerçekleştirilebilmek için gerekli örneklem çapının yaklaşık olarak hesaplanabilmesi için asimptotik dağılımının belirlenmesi oldukça önemlidir. Böylece gerçek bir veri setinin içerdiği gözlem sayısına göre testin anlamlılığı ve gücü hakkında yaklaşık bir sonuca varılabilir. T istatistiğinin asimptotik dağılımının hesaplamak için ilk olarak (2) eşitliği ile verilen dönüşüm göz önüne alınsın. Ayrıca, X_1, X_2, \dots, X_n a oran parametrelili geometrik süreçten bir örneklem ve $X_1 \sim GA(\alpha, \beta)$ veya $X_1 \sim WE(\lambda, \theta)$ olsun. $n \rightarrow \infty$ ve $\hat{\alpha} \xrightarrow{h.h.h.y} a$ için (10) eşitliği ile verilen T istatistiği Y_1, Y_2, \dots, Y_n dönüşüm değişkenleri ile

$$T_Y = \ln \left[\prod_{i=1}^n \frac{f_{GA}(y_i; \alpha, \beta)}{f_{WE}(y_i; \theta, \lambda)} \right] \quad (18)$$

biçiminde ifade edilebilir. Y_1, Y_2, \dots, Y_n rasgele değişkenleri birbirinden bağımsız ve aynı dağılımlı olmalarının yanısıra X_1 rasgele değişkeniyle de aynı dağılımlıdır. (18) eşitliğinde verilen T_Y istatistiği ve (10) eşitliği ile verilen T istatistiğinin aynı dağılımlı olduğu aşikardır. (18) eşitliğinde verilen T_Y istatistiği ise klasik yaklaşımda Gamma ve Weibull dağılımlarının ayırımı için kullanılan RML yöntemine dayalı bir ayırıcı istatistiktir. T_Y istatistiğinin asimptotik dağılımı

- veriler Gamma dağılımından
- veriler Weibull dağılımından

olması durumları göz önünde bulundurularak Fearn ve Nebenzahl (1991) tarafından yapılan çalışma ile elde edilmiştir. Fearn ve Nebenzahl (1991) tarafından elde edilen asimptotik sonuçlar Tablo 1'de verilmiştir.

Tablo 1' de verilen $\psi(\cdot)$ digamma fonksiyonudur. Ayrıca, $\tilde{\theta}$ veri Gamma dağılımından geldiğinde Weibull dağılımının hatalı parametresi, $\tilde{\alpha}$ ise veri Weibull dağılımından geldiğinde Gamma dağılımının hatalı parametresi olarak adlandırılır. $\tilde{\theta}$ ve $\tilde{\alpha}$ 'nın hesabı için Fearn ve Nebenzahl (1991)'e bakınız.

Tablo 1: T_Y ve T İstatistiği için Asimptotik Sonuçlar

Durum 1: X_1 rasgele değişkeni Gamma dağılımlı yani, $Y_i \sim GA(\alpha, \beta) \equiv X_1 \sim GA(\alpha, \beta)$ ise $T_Y, T \sim$ Asimptotik Normal ($n \times AE_{GA}, n \times AVar_{GA}$)	
Beklenen değer (AE_{GA}):	$1 - \alpha - 2\ln\Gamma(\alpha) - \ln\tilde{\theta} + \ln\Gamma(\tilde{\theta} + \alpha) + (\alpha - \tilde{\theta})\psi(\alpha)$
Varyans ($AVar_{GA}$):	$\left\{ \frac{(\alpha - \tilde{\theta})^2 \psi'(\alpha) + 2(\alpha - \tilde{\theta})[\psi(\tilde{\theta} + \alpha) - \psi(\alpha)] + \Gamma(\alpha)\Gamma(2\tilde{\theta} + \alpha)}{[\Gamma(\tilde{\theta} + \alpha)]^2} - \alpha - 1 \right\}$
Durum 2: X_1 rasgele değişkeni Weibull dağılımlı yani, $Y_i \sim WE(\lambda, \theta) \equiv X_1 \sim WE(\lambda, \theta)$ ise $T_Y, T \sim$ Asimptotik Normal ($n \times AE_{WE}, n \times AVar_{WE}$)	
Beklenen değer (AE_{WE}):	$-\tilde{\alpha} \left(\ln\tilde{\alpha} + \frac{\psi(1)}{\theta} - \ln\Gamma\left(1 + \frac{1}{\theta}\right) - 1 \right) + \ln\theta + \ln\Gamma(\tilde{\alpha}) + \psi(1) - 1$
Varyans ($AVar_{WE}$):	$\left\{ (\theta - \tilde{\alpha})^2 \frac{\pi^2}{6\theta^2} + \tilde{\alpha}^2 \frac{\Gamma\left(1 + \frac{2}{\theta}\right)}{\left(\Gamma\left(1 + \frac{1}{\theta}\right)\right)^2} - \tilde{\alpha}^2 + 2(\theta - \tilde{\alpha}) \left[\frac{\tilde{\alpha}}{\theta} \psi\left(\frac{1}{\theta}\right) - \frac{\tilde{\alpha}}{\theta} \psi(1) + \tilde{\alpha} \right] - 1 \right\}$

T istatistiği için Durum 1'de verilen asimptotik dağılımın göz önüne alınmasıyla, X_1 rasgele değişkeni gerçekte Gamma dağılımlı olduğunda (17) ile verilen karar kuralına göre doğru seçim yapılmış olması olasılığı, diğer bir deyişle H_0 hipotezi doğru iken kabul edilmesi olasılığı

$$P(T > 0) \approx 1 - \Phi\left(-\frac{n \times AE_{GA}}{\sqrt{n \times AVar_{GA}}}\right) \quad (19)$$

olarak elde edilir. Burada Φ standart normal dağılımın dağılım fonksiyonunu ifade etmektedir. Ayrıca testin anlam düzeyi ξ ile gösterilmek üzere; (19) ifadesinden

$$1 - \xi = 1 - \Phi\left(-\frac{n \times AE_{GA}}{\sqrt{n \times AVar_{GA}}}\right) \quad (20)$$

yazılabilir. Böylece (20) eşitliğinden testin anlam düzeyi

$$\xi = \Phi\left(-\frac{n \times AE_{GA}}{\sqrt{n \times AVar_{GA}}}\right) \quad (21)$$

olarak elde edilir. Eşitlik (21) in kullanılmasıyla istenilen bir ξ^* anlam düzeyinde test işlemi gerçekleştirilebilir için gerekli örneklem büyüklüğü ise

$$n \approx \frac{(Z_{\xi^*})^2 AVar_{GA}}{(AE_{GA})^2} \quad (22)$$

olarak bulunur. Burada Z_{ξ^*} , ξ^* olasılığı için standart normal dağılımın dağılım fonksiyonunun tersinin değeridir.

Benzer biçimde X_1 rasgele değişkeni gerçekte Weibull dağılımlı olduğunda T istatistiği $n \times AE_{WE}$ ortalama ve $n \times AVar_{WE}$ varyansı ile asimptotik olarak Normal dağılıma sahip olduğundan, (17) kuralına göre verilen kararın doğru olması olasılığı diğer bir deyişle H_0 hipotezi yanlış iken reddedilmesi olasılığı

$$P(T < 0) \approx \Phi\left(-\frac{n \times AE_{WE}}{\sqrt{n \times AVar_{WE}}}\right) \quad (23)$$

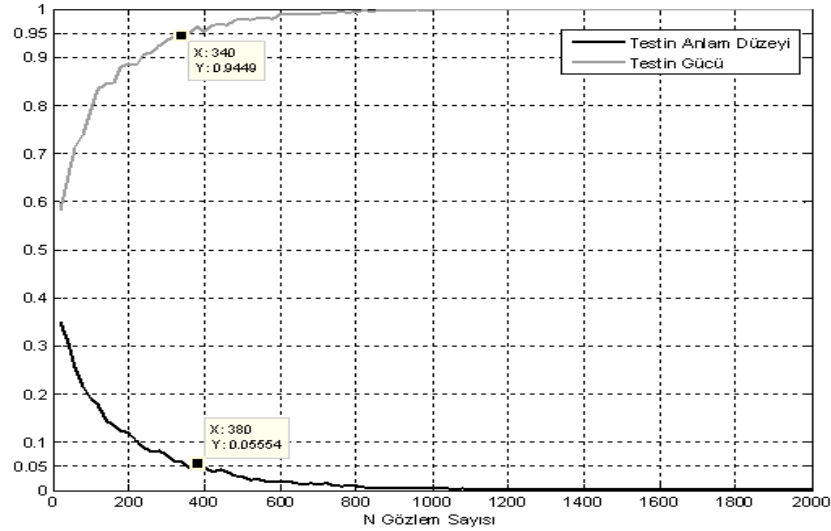
olarak elde edilir. (23) ifadesi ayrıca testin gücünü göstermektedir. İstenen güç düzeyinde bir test işlemi gerçekleştirebilmek için gerekli örneklem büyüklüğü ise eşitlik (23) ün kullanılmasıyla

$$n \approx \frac{(Z_{\varphi^*})^2 AVar_{WE}}{(AE_{WE})^2} \quad (24)$$

olarak elde edilir. Burada φ^* istenen güç düzeyini göstermektedir.

İstenen bir anlam düzeyi ve güçte test gerçekleştirebilmek için gerekli örneklem sayısını belirlemek amacıyla $n = 20,40,60, \dots, 2000$, $a = 0.95$ seçilerek ve (19)-(24) eşitlikleri ile verilen asimptotik sonuçları kullanarak, gözlem sayısı n 'e göre testin güç ve anlamlılık düzeyleri 1000 tekrarlı simülasyon çalışmasına dayalı olarak hesaplanmış ve Şekil 3'de verilmiştir.

Şekil 3: Gözlem Sayısı N'e Göre Testin Gücü ve Anlamlılık Düzeyi



Şekil 3'den $\xi \cong 0.05$ anlam düzeyinde bir test işlemi gerçekleştirmek için yaklaşık 380 ve $\varphi \cong 0.95$ güçte bir test için ise yaklaşık 340 birimlik örneklem kullanılması gerektiği sonucuna varılabilir. Ayrıca analiz edilecek veri setindeki gözlem sayısına göre testin gücü ve anlam düzeyi hakkında yaklaşık olarak bir sonuca varılabilir.

3. Simülasyon Çalışması

Bu bölümde, T istatistiğinin ayırım performansını ortaya koymak için yapılan simülasyon çalışmasına yer verilmektedir. Simülasyon çalışmasında sırasıyla $X_1 \sim GA(\alpha, \beta)$ ve $X_1 \sim WE(\lambda, \theta)$ olmak üzere iki farklı durum düşünülmüştür. Her iki durumda da T istatistiğinin doğru seçim olasılığı hem simülasyon hem de T istatistiğinin asimptotik dağılımına dayalı olarak hesaplanmıştır. Birinci durumda $a = 0.90, 0.95, 1.05, 1.10$, $\alpha = 0.5, 2$, $\beta = 0.5, 2, 4, 6, 8, 10$ ve $n = 30, 50, 100, 200$ olarak seçilmiş ve simülasyon gerçekleştirilmiştir. 1000 tekrara dayanarak elde edilen sonuçlar ilk satırlarda simülasyona dayalı sonuçlar ve ikinci satırlarda asimptotik sonuçlar olmak üzere Tablo 2-3 de verilmiştir.

Tablo 2: $X_1 \sim GA(\alpha, \beta)$ ve $\alpha = 0.5$ Durumu için Doğru Seçim Olasılıkları

β	a	n				β	a	n			
		30	50	100	200			30	50	100	200
0.5	0.9	0.587	0.624	0.741	0.804	6	0.9	0.580	0.643	0.702	0.837
		0.573	0.607	0.723	0.804			0.558	0.629	0.699	0.828
	0.95	0.571	0.642	0.735	0.853		0.95	0.580	0.621	0.704	0.840
		0.546	0.622	0.739	0.840			0.557	0.611	0.693	0.833
	1.05	0.571	0.646	0.732	0.836		1.05	0.614	0.684	0.753	0.834
		0.566	0.619	0.720	0.824			0.582	0.637	0.737	0.832
	1.1	0.569	0.651	0.735	0.815		1.1	0.569	0.648	0.737	0.845
		0.553	0.639	0.725	0.807			0.555	0.629	0.725	0.834
2	0.9	0.594	0.618	0.725	0.821	8	0.9	0.591	0.633	0.745	0.843
		0.566	0.609	0.715	0.809			0.566	0.609	0.737	0.839
	0.95	0.603	0.654	0.719	0.825		0.95	0.608	0.640	0.722	0.833
		0.587	0.633	0.713	0.812			0.570	0.611	0.715	0.817
	1.05	0.597	0.633	0.725	0.809		1.05	0.578	0.655	0.705	0.832
		0.580	0.618	0.708	0.820			0.551	0.633	0.687	0.825
	1.1	0.567	0.644	0.734	0.815		1.1	0.613	0.667	0.715	0.844
		0.566	0.624	0.723	0.806			0.584	0.643	0.724	0.832
4	0.9	0.597	0.642	0.730	0.827	10	0.9	0.598	0.678	0.730	0.827
		0.583	0.628	0.716	0.819			0.580	0.647	0.737	0.821
	0.95	0.558	0.651	0.739	0.810		0.95	0.599	0.668	0.744	0.832
		0.538	0.630	0.737	0.805			0.568	0.629	0.722	0.814
	1.05	0.607	0.653	0.731	0.832		1.05	0.613	0.621	0.751	0.839
		0.581	0.627	0.728	0.831			0.581	0.610	0.731	0.835
	1.1	0.587	0.633	0.704	0.852		1.1	0.596	0.652	0.716	0.829
		0.557	0.602	0.714	0.835			0.577	0.629	0.711	0.831

Tablo 3: $X_1 \sim GA(\alpha, \beta)$ ve $\alpha = 2$ Durumu için Doğru Seçim Olasılıkları

β	a	n				β	a	n			
		30	50	100	200			30	50	100	200
0.5	0.9	0.598	0.645	0.703	0.768	6	0.9	0.618	0.679	0.699	0.799
		0.599	0.658	0.701	0.760			0.614	0.660	0.693	0.790
	0.95	0.627	0.659	0.724	0.790		0.95	0.609	0.647	0.688	0.814
		0.605	0.640	0.721	0.782			0.615	0.638	0.697	0.801
	1.05	0.636	0.655	0.714	0.804		1.05	0.631	0.657	0.704	0.801
		0.611	0.650	0.717	0.806			0.610	0.643	0.715	0.794
1.1	0.630	0.641	0.739	0.778	1.1	0.650	0.672	0.693	0.786		
	0.620	0.639	0.722	0.772		0.633	0.672	0.691	0.780		
2	0.9	0.606	0.656	0.729	0.802	8	0.9	0.652	0.661	0.741	0.807
		0.613	0.645	0.715	0.788			0.630	0.650	0.733	0.805
	0.95	0.596	0.647	0.698	0.825		0.95	0.638	0.652	0.711	0.803
		0.595	0.637	0.700	0.810			0.631	0.649	0.704	0.798
	1.05	0.638	0.663	0.711	0.797		1.05	0.614	0.658	0.757	0.785
		0.607	0.645	0.710	0.775			0.603	0.646	0.744	0.783
1.1	0.638	0.655	0.712	0.795	1.1	0.651	0.638	0.747	0.797		
	0.614	0.654	0.710	0.802		0.615	0.630	0.727	0.784		
4	0.9	0.641	0.667	0.722	0.808	10	0.9	0.633	0.670	0.700	0.793
		0.624	0.654	0.712	0.794			0.626	0.665	0.707	0.785
	0.95	0.631	0.645	0.711	0.780		0.95	0.613	0.651	0.716	0.802
		0.616	0.639	0.695	0.777			0.614	0.661	0.706	0.791
	1.05	0.622	0.662	0.731	0.813		1.05	0.632	0.649	0.712	0.792
		0.610	0.654	0.723	0.813			0.625	0.639	0.716	0.795
1.1	0.607	0.650	0.718	0.780	1.1	0.640	0.638	0.731	0.785		
	0.605	0.657	0.721	0.792		0.626	0.638	0.715	0.781		

Tablo 2 ve Tablo 3'den anlaşılmaktadır ki gözlem sayısı n artıkça hem simülasyona hemde asimptotik sonuçlara dayalı doğru seçim olasılıkları artmaktadır. Bu durum geometrik sürecin oran parametresi a 'nın her değeri için geçerlidir. Bunun yanısıra α parametresinin değeri büyüdüğü zaman doğru seçim olasılıklarının da arttığı görülmektedir.

İkinci durumu göz önüne alan simülasyon çalışmasında ise ilk durumdakine benzer biçimde $a = 0.90, 0.95, 1.05, 1.10, \lambda = 0.5, 2, \theta = 0.5, 2, 4, 6, 8, 10$ ve $n = 30, 50, 100, 200$ olarak seçilmiştir. 1000 tekrarlı simülasyon çalışması ile elde edilen sonuçlar Tablo 4-5 de verildiği gibidir.

Tablo 4: $X_1 \sim WE(\lambda, \theta)$ ve $\lambda = 0.5$ Durumu için Doğru Seçim Olasılıkları

θ	a	n				θ	a	n			
		30	50	100	200			30	50	100	200
0.5	0.9	0.709	0.778	0.858	0.941	6	0.9	0.777	0.836	0.923	0.992
		0.708	0.768	0.836	0.917			0.764	0.823	0.916	0.985
	0.95	0.703	0.784	0.849	0.946	0.95	0.736	0.843	0.941	0.985	
		0.713	0.776	0.838	0.930		0.742	0.824	0.923	0.983	
	1.05	0.712	0.772	0.856	0.916	1.05	0.767	0.831	0.934	0.983	
		0.710	0.764	0.846	0.909		0.763	0.824	0.924	0.982	
	1.1	0.730	0.759	0.844	0.937	1.1	0.776	0.826	0.933	0.982	
		0.734	0.760	0.836	0.930		0.769	0.829	0.923	0.981	
2	0.9	0.633	0.690	0.810	0.897	8	0.9	0.756	0.861	0.940	0.992
		0.623	0.675	0.795	0.891			0.764	0.843	0.927	0.982
	0.95	0.627	0.694	0.802	0.876	0.95	0.777	0.866	0.952	0.994	
		0.628	0.694	0.798	0.884		0.774	0.845	0.934	0.988	
	1.05	0.640	0.680	0.784	0.885	1.05	0.762	0.854	0.934	0.984	
		0.640	0.686	0.766	0.879		0.769	0.844	0.933	0.983	
	1.1	0.646	0.717	0.818	0.894	1.1	0.750	0.844	0.946	0.994	
		0.644	0.699	0.800	0.896		0.755	0.846	0.944	0.986	
4	0.9	0.729	0.796	0.896	0.975	10	0.9	0.776	0.825	0.950	0.993
		0.732	0.809	0.891	0.971			0.772	0.826	0.936	0.985
	0.95	0.710	0.803	0.909	0.969	0.95	0.769	0.860	0.947	0.991	
		0.730	0.797	0.900	0.974		0.769	0.861	0.936	0.988	
	1.05	0.733	0.827	0.899	0.983	1.05	0.802	0.842	0.960	0.992	
		0.721	0.822	0.908	0.973		0.786	0.839	0.945	0.988	
	1.1	0.726	0.824	0.900	0.976	1.1	0.770	0.846	0.948	0.987	
		0.724	0.806	0.899	0.970		0.764	0.845	0.941	0.985	

Tablo 5: $X_1 \sim WE(\lambda, \theta)$ ve $\lambda = 2$ Durumu için Doğru Seçim Olasılıkları

θ	a	n				θ	a	n					
		30	50	100	200			30	50	100	200		
0.5	0.9	0.706	0.749	0.862	0.938	6	0.9	0.737	0.829	0.924	0.983		
		0.705	0.754	0.840	0.920			0.743	0.824	0.926	0.978		
	0.95	0.695	0.789	0.870	0.944		0.95	0.748	0.837	0.928	0.990		
		0.706	0.768	0.848	0.922			0.740	0.820	0.910	0.984		
	1.05	0.711	0.764	0.835	0.931		1.05	0.777	0.815	0.934	0.987		
		0.721	0.766	0.826	0.917			0.778	0.815	0.926	0.983		
	1.1	0.711	0.753	0.852	0.934		1.1	0.729	0.858	0.936	0.990		
		0.712	0.752	0.830	0.925			0.743	0.834	0.931	0.981		
	2	0.9	0.615	0.706	0.774		0.890	8	0.9	0.781	0.842	0.938	0.995
			0.609	0.707	0.794		0.887			0.774	0.844	0.933	0.987
0.95		0.627	0.689	0.774	0.890	0.95	0.779		0.865	0.948	0.993		
		0.633	0.697	0.778	0.887		0.764		0.846	0.942	0.988		
1.05		0.635	0.719	0.781	0.892	1.05	0.761		0.834	0.943	0.984		
		0.631	0.710	0.783	0.885		0.761		0.830	0.935	0.984		
1.1		0.622	0.694	0.807	0.886	1.1	0.730		0.855	0.943	0.990		
		0.630	0.696	0.798	0.890		0.741		0.839	0.937	0.984		
4		0.9	0.714	0.823	0.922	0.972	10		0.9	0.780	0.829	0.950	0.991
			0.713	0.811	0.910	0.968				0.779	0.835	0.934	0.988
	0.95	0.730	0.784	0.911	0.976	0.95		0.762	0.859	0.955	0.990		
		0.728	0.785	0.907	0.976			0.769	0.863	0.934	0.987		
	1.05	0.697	0.814	0.903	0.971	1.05		0.775	0.867	0.941	0.992		
		0.707	0.804	0.897	0.972			0.778	0.848	0.929	0.987		
	1.1	0.718	0.809	0.908	0.978	1.1		0.774	0.855	0.950	0.988		
		0.709	0.810	0.901	0.970			0.776	0.842	0.936	0.986		

Tablo 4 ve Tablo 5 ile verilen simülasyon çalışması sonuçları, tamamen birinci durumda elde edilen sonuçlara benzer olarak yorumlanabilir. a parametresinin her değeri için n arttıkça hem simülasyona hem de asimptotik sonuçlara dayalı doğru seçim olasılıkları artmaktadır. Ayrıca Tablo 2-4 incelendiğinde, ikinci durum için gerçekleştirilen simülasyon çalışması ile elde edilen doğru seçim olasılıkları parametrelerin ve gözlem sayısının her değerinde birinci durumdan daha yüksek çıkmaktadır.

4. Uygulama

Çalışmada yapılanları açıklama amaçlı olarak, bu kısımda, geometrik süreçle uyumlu bir veri seti olan “No. 3 planlanmamış bakım veri seti” kullanılarak bir uygulamaya yer verilmiştir.

71 gözlem içeren No. 3 planlanmamış bakım veri seti, U.S.S. Halfbeak denizaltılarının ana motorları için planlanmamış ardışık bakım zamanlarını içermektedir (Ascher ve Feingold, 1984). Bu veri setinin, oran parametresi 1'den küçük olan bir geometrik süreç ile uyumlu olduğu Chan vd. (2004) tarafından gösterilmiş ve araştırmacılar tarafından bu veri setinde ilk varış zamanının dağılımı Gamma olarak seçilmiş ve veri seti geometrik süreç kullanılarak modellenmiştir. Aydoğdu vd. (2010) tarafından aynı veri seti için ilk varış zamanının dağılımı Weibull seçilerek geometrik süreçle modellenmiştir. Acaba bu veri setinde ilk varış zamanının dağılımı için düşünülen Gamma

ve Weibull dağılımlarından hangisi daha uygundur? Uygun dağılımı belirleyebilmek için hesaplanan parametre tahminleri ve T istatistiğinin değeri Tablo 6'da verildiği gibidir.

Tablo 6: No. 3 Planlanmamış Bakım Veri Seti için Parametre Tahmini ve T İstatistiğinin Değeri

\hat{a}	Model	Parametre	Tahmin	Log-olabilirlik	T
1.0401	Gamma	α	0.6581	-546.8785	-0.0419
		β	0.001479		
	Weibull	λ	0.0012	-546.8366	
		θ	0.7616		

Tablo 6' dan görülebileceği gibi $T < 0$ 'dır. Bu durumda (17) ile verilen karar kuralına göre ilk varış zamanının dağılımı için Weibull dağılımı daha uygundur denilebilir. Diğer taraftan bir geometrik süreçte X_1, X_2, \dots, X_n gözlemleri

$$\hat{X}_k = E(X_1)\hat{a}^{1-k}, \quad k = 1, 2, \dots, n \quad (25)$$

biçiminde tahmin edilebilir (Kara vd, 2015). (25) eşitliği ile verilen tahminler kullanılarak hata kareler ortalaması (MSE_X) ise

$$MSE_X = \frac{1}{n} \sum_{k=1}^n (X_k - \hat{X}_k)^2 \quad (26)$$

biçiminde hesaplanabilir. İlk varış zamanının dağılımının Gamma ve Weibull olması durumuna göre No. 3 planlanmamış bakım veri seti için ilk varış zamanının beklenen değeri ve MSE_X değerleri Tablo 7 de verilmiştir.

Tablo 7: No. 3 Planlanmamış Bakım Veri Seti için Beklenen Değer ve MSE_X Değerleri

Model Dağılım	Beklenen Değer $E(X_1)$	MSE_X
Gamma	973.8308	201801.27689
Weibull	977.177	201595.27778

Tablo 7'den görüleceği gibi X_1 rasgele değişkeni için model olarak Weibull dağılımı kullanıldığında MSE_X değeri daha küçük olmaktadır. Bu ise T istatistiğine göre verilen kararı desteklemektedir.

5. Sonuç

Bu çalışmada geometrik süreçle uyumlu bir veri setinde ilk varış zamanının dağılımı için Gamma ve Weibull dağılımları arasındaki ayırım probleminin çözümü RML yöntemine göre araştırılmıştır. İki dağılım arasında ayırım yapmak için kullanılan istatistiğe ait doğru seçim performansı yapılan bir dizi simülasyon çalışmasıyla gösterilmiştir. Elde edilen sonuçlardan, doğru seçim olasılığı için asimptotik sonuçlar ile simülasyona dayalı sonuçların örtüşükleri görülmüştür. Bununla birlikte asimptotik sonuçlar kullanılarak, gözlem sayısına göre testin güç ve anlamlılık düzeyi yaklaşık olarak hesaplanmıştır. Daha önceki yapılan çalışmalarda her iki dağılım da kullanılarak modellenmiş gerçek bir veri seti, açıklayıcı amaçlar için analiz edilmiştir. Hem simülasyon çalışması sonuçlarından hemde veri analizi sonucundan, RML yönteminin geometrik süreç verileri için Gamma ve Weibull dağılımları arasındaki ayırımda kullanımının uygun olduğu söylenebilir.

Kaynakça

- Ascher H. ve Feingold, H. (1984). *Repairable Systems Reliability*. New York: Marcel Dekker.
- Aydođdu, H., Őenođlu, B. ve Kara, M. (2010). Parameter Estimation in Geometric Process with Weibull Distribution. *Applied Mathematics and Computation*, 217(6), 2657-2665.
- Bain, L. J. ve Engelhardt, M. (1980). Probability of Correct Selection of Weibull Versus Gamma Based on Livelihood Ratio. *Communications in Statistics-Theory and Methods*, 9(4), 375-381.
- Chan, J. S., Lam, Y. ve Leung, D. Y. (2004). Statistical İnference for Geometric Processes with Gamma Distributions. *Computational Statistics & Data Analysis*, 47(3), 565-581.
- Dey, A. K. ve Kundu, D. (2012). Discriminating Between the Weibull and Log-normal Distributions for Type-II Censored Data. *Statistics*, 46(2), 197-214.
- Elsherpieny, E. A., Ibrahim, N. S. ve Radwan, N. U. (2013). Discriminating Between Weibull and Log-logistic Distributions. *International Journal of Innovative Research in Science, Engineering and Technology*, 2(8), 3358-3371.
- Fearn, D. H. ve Nebenzahl, E. (1991). On the Maximum Likelihood Ratio Method of Deciding Between the Weibull and Gamma Distributions. *Communications in Statistics-Theory and Methods*, 20(2), 579-593.
- Kara, M., Aydođdu, H. ve Tőrksen, Ö. (2015). Statistical Inference for Geometric Process with the Inverse Gaussian Distribution. *Journal of Statistical Computation and Simulation*, 85(16), 3206-3215.
- Kundu, D. ve Manglick, A. (2005). Discriminating Between the Log-Normal and Gamma Distributions. *Journal of the Applied Statistical Sciences*, 14, 175-187.
- Kundu, D. ve Manglick, A. (2004). Discriminating Between the Weibull and Log-Normal Distributions *Naval Research Logistics (NRL)*, 51(6), 893-905.
- Kundu, D., Gupta, R. D., ve Manglick, A. (2005). Discriminating Between the Log-normal and Generalized Exponential Distributions. *Journal of Statistical Planning and Inference*, 127(1), 213-227.
- Kundu, D. ve Raqab, M. Z. (2007). Discriminating Between the Generalized Rayleigh and Log-normal Distribution. *Statistics*, 41(6), 505-515.
- Lam, Y. (1988). Geometric Processes and Replacement Problem. *Acta Mathematicae Applicatae Sinica*, 4, 366-377.
- Lam, Y. (2007). *The Geometric Process and Its Applications*. Singapore: World Scientific.
- Lam, Y. ve Chan, S. K. (1998). Statistical Inference for Geometric Processes with Lognormal Distribution. *Computational Statistics & Data Analysis*, 27(1), 99-112.
- Raqab, M. Z. (2013). Discriminating Between the Generalized Rayleigh and Weibull Distributions. *Journal of Applied Statistics*, 40(7), 1480-1493.

DISCRIMINATING BETWEEN THE GAMMA AND WEIBULL DISTRIBUTIONS FOR GEOMETRIC PROCESS DATA

Extended Abstract

Aim: Let stochastic process $\{N(t), t > 0\}$ is a count process and $\{X_n, n = 1, 2, \dots\}$ be a set of random variables from count process $\{N(t), t > 0\}$. The process $\{X_n, n = 1, 2, \dots\}$ is called a geometric process with ratio parameter a if there exists any real constant $a > 0$ such that $\{a^{n-1}X_n, n = 1, 2, \dots\}$ are set of independently and identically distributed random variables with distribution F (Lam 2007). X_1, X_2, \dots, X_n be a random sample from geometric process with ratio parameter a then the random variable X_1 is indicate the first occurrence time. The parametric inference in the geometric process is made by assuming the distribution of random variable X_1 is known. Also, let μ and σ^2 are expected value and variance of the random variable X_1 , respectively, then expected value and variance of the other random variables can be fully expressed by μ and σ^2 . Hence, for analysis of a set of data consistent with geometric process, determining the underlying distribution of a set of data as true or close to true is quite important. The main goal of this study, in the stage of modelling of a set of data consistent with geometric process, it is investigate the solution of the problem of discriminating between the Gamma and Weibull distributions for the random variable X_1 .

Method(s): The ratio of maximum likelihood (RML) method is used for discrimination between the Gama and Weibull distributions for geometric process data. The RML is a commonly used method of discriminating between two overlapping distribution families. The most important advantage of the RML method is that it is an easily applicable method. However, distribution of the discriminative statistic obtained by RML cannot be expressed explicitly and its asymptotic distribution becomes quite important. By using the asyptotic distribution of the RML, requiried minimum sample sizes for the fixed confidence level and power of the test can be obtained.

Findings: In order to discriminating between the Gamma and Weibull distribution for geometric process data, a discriminative statistic based on random sample (X_1, X_2, \dots, X_n) from geometric process with ratio parameter a has been obtained as follow

$$T = n \left[\hat{\alpha} \ln(\hat{\alpha} \bar{X}_a / \bar{X}_a) - \hat{\theta} \ln(\hat{\lambda} \bar{X}_a) - \ln(\hat{\theta} \Gamma(\hat{\alpha})) - \hat{\alpha} + 1 \right] \quad (1)$$

where $\hat{\alpha}$ and $\hat{\theta}$ are maximum likelihood estimators of parameters α and θ , respectively and $\bar{X}_a = (\prod_{i=1}^n \hat{\alpha}^{i-1} X_i)^{1/n}$, $\bar{X}_a = \frac{1}{n} \sum_{i=1}^n \hat{\alpha}^{i-1} X_i$ and $\hat{\alpha} = \exp\left(\frac{6}{(n-1)n(n+1)} \sum_{i=1}^n (n-2i+1) \ln X_i\right)$. Besides, the asymptotic distribution of the statistic given by Equation (1) has been evaluated according to the results obtained by Fearn and Nebenzahl (1991). The correct selection performance of the discriminative statistic used in this study has been examined through a comprehensive simulation study. The results of the simulation study has shown that the correct selection performance of the used statistic is quite satisfactory even at small sample sizes. In addition, according to sample size n and for fixed value of the ratio parameter a , the confidence level and power of the test has been obtained based on asymptotic distribution of the statistic T

Conclusion: In this study, in order to solution of the problem of the discriminating between the Gamma and Weibull distributions for geometric process data, a discriminative statistic has been obtained based on the RML method. The correct selection performance and power of the test for the statistics used to discriminating between the two distributions has been shown with a series of simulation studies. From obtained result by the simulation study, it is seen that asymptotic results and simulation-based results are overlap in terms of the probability of correct selection. In addition, a real data set which is modeled using both Gamma and Weibull distributions in the previous studies has been analyzed for explanatory purposes. As a result of both data analysis and the simulation study, it can be said that the use of the RML method for the discriminating between the Gamma and Weibull distributions for geometric process data is appropriate.