

**T.C.
KIRIKKALE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**MATEMATİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ**

DAĞILIM SEÇİMİ ÜZERİNE BİR ÇALIŞMA

REYHAN KARAKOÇ

Temmuz 2018

**T.C.
KIRIKKALE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**MATEMATİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ**

DAĞILIM SEÇİMİ ÜZERİNE BİR ÇALIŞMA

REYHAN KARAKOÇ

Temmuz 2018

Matematik Anabilim Dalında Reyhan KARAKOÇ tarafından hazırlanan
DAĞILIM SEÇİMİ ÜZERİNE BİR ÇALIŞMA adlı Yüksek Lisans Tezinin
Anabilim Dalı standartlarına uygun olduğunu onaylarım.

Prof. Dr. Kerim KOCA
Anabilim Dalı Başkanı

Bu tezi okuduğumu ve tezin **Yüksek Lisans Tezi** olarak bütün gereklilikleri yerine
getirdiğini onaylarım.

Dr. Öğr. Üyesi Cenker BİÇER
Danışman

Jüri Üyeleri

Başkan : (Unvanı, Adı ve Soyadı, İmzası) _____
Üye (Danışman) : (Unvanı, Adı ve Soyadı, İmzası) _____
Üye : (Unvanı, Adı ve Soyadı, İmzası) _____

...../...../.....

Bu tez ile Kırıkkale Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu Yüksek
Lisans derecesini onaylamıştır.

Prof. Dr. Mustafa YİĞİTOĞLU
Fen Bilimleri Enstitüsü Müdürü

ÖZET

DAĞILIM SEÇİMİ ÜZERİNE BİR ÇALIŞMA

KARAKOÇ, Reyhan

Kırıkkale Üniversitesi

Fen Bilimleri Enstitüsü

Matematik Anabilim Dalı, Yüksek Lisans tezi

Danışman: Dr. Öğr. Üyesi Cenker BİÇER

Temmuz2018,47sayfa

Gerçek hayat problemlerinde gözlenen veriler, aynı uygulama alanlarına sahip dağılımlar kullanılarak modellenabilir. Ancak verilere dayalı en uygun istatistiksel sonucun çıkarılabilmesi, verilerin en uygun dağılımla modellenmesine bağlıdır. Bu çalışmada, aynı uygulama alanlarına sahip genelleştirilmiş üstel ve Power Lindley dağılımları arasındaki ayırım problemi incelenmektedir. İlgili dağılımlar arasında bir ayırım yapabilmek için olabilirlik oranına dayalı bir ayırıcı istatistik ve ayırıcı istatistiğin asimptotik dağılımı elde edilmektedir. Elde edilen ayırıcı istatistiğin, doğru seçim performansı ve ulaşılan asimptotik sonuçların nasıl çalıştığı yapılan kapsamlı Monte-Carlo simülasyon çalışmaları ile değerlendirilmiştir. Simülasyon sonuçları, elde edilen ayırıcı istatistiğin, genelleştirilmiş üstel ve Power Lindley dağılımları arasındaki ayırıcı kullanılımasının uygun olduğunu göstermektedir.

Anahtar kelimeler: Power Lindley dağılımı, Genelleştirilmiş üstel dağılım, Asimptotik normallik, Olabilirlik oran yöntemi, Simülasyon.

ABSTRACT

A STUDY ON DISTRIBUTION SELECTION

KARAKOÇ, Reyhan

Kırıkkale University

Graduate School of Natural and Applied Sciences

Department of Mathematics, Master Thesis

Supervisor: Dr Öğr. Üyesi Cenker BİÇER

July 2018,47 pages

The data observed in real life problems can be modeled using distributions having the same application areas. However, to the optimal statistical inference based on the data depends on the modeling of the data with the optimal distribution. In this study, the problem of the discriminating between generalized exponential and Power Lindley distributions, which they have the same application fields, is examined. In order to make a discriminating between these distributions, a discriminative statistic based on the likelihood ratio is obtained and its asymptotic distribution is also obtained. Correct selection performance of the obtained discriminative statistic and how works the achieved asymptotic results are evaluated by comprehensive Monte-Carlo simulations. Simulation results show that the obtained discriminative statistic is appropriate using in the discriminating between the generalized exponential and Power Lindley distributions.

Key Words: Power Lindley distribution, Generalized exponential distribution, Asymptotic normality, Likelihood ratio method, Simulation.

TEŐEKKÜR

Tezimin hazırlanması esnasında hiçbir yardımcı esirgemeyen ve biz genç arařtırmacılara büyük destek olan, bilimsel deney imkanlarını sonuna kadar bizlerin hizmetine veren, tez yöneticisi hocam, Sayın Dr. Öğr. Üyesi Cenker BİÇER'e, tez çalışmalarım esnasında, bilimsel konularda daima yardımını gördüğüm hocam Sayın Dr. Öğr. Üyesi Hayrinisa BİÇER'e büyük fedakarlıklarla bana destek olan arkadaşım Muhammed ERGÜN'e, ve son olarak bana birçok konuda olduğu gibi, tezimi hazırlamam esnasında da yardımlarını esirgemeyen aileme teşekkür ederim.



İÇİNDEKİLER DİZİNİ

İÇİNDEKİLER	Sayfa
ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER DİZİNİ	iv
ŞEKİLLER DİZİNİ	v
ÇİZELGELER DİZİNİ	vi
1. GİRİŞ VE ÖNCEKİ ÇALIŞMALAR	1
2. BAZI DAĞILIMLAR	4
2.1. Üstel Dağılım	4
2.2. Weibull Dağılımı.....	6
2.3. Gamma Dağılımı.....	8
2.4. Log-Normal Dağılım.....	11
2.5. Genelleştirilmiş Üstel Dağılım.....	13
2.6. Power Lindley Dağılımı.....	16
3. OLABİLİRLİK ORANINA DAYALI AYIRICI İSTATİSTİK	19
4. BAZI DAĞILIM ÇİFTLERİ İÇİN RML İSTATİSTİĞİ	23
4.1. Weibull ve Log-Normal Dağılımları Arasında Ayrım.....	23
4.2. Log-Normal ve Gamma Dağılımları Arasındaki Ayrım.....	28
4.3. Genelleştirilmiş Üstel Dağılım ve Power Lindley Dağılımları Arasındaki Ayrım	32
5. SİMULASYON ÇALIŞMASI	40
6. SONUÇ	46
KAYNAKLAR	47

ŞEKİLLER DİZİNİ

<u>ŞEKİL</u>	<u>Sayfa</u>
Şekil 2. 1: Farklı parametre değerleri için üstel dağılımın olasılık yoğunluk fonksiyonunun grafiği.....	5
Şekil 2. 2: Farklı parametre değerleri için Weibull dağılımının olasılık yoğunluk fonksiyonunun grafiği.....	6
Şekil 2. 3: Farklı parametre değerleri için Weibull dağılımının bozulma fonksiyonunun grafiği.....	7
Şekil 2. 4: Farklı parametre değerleri için Gamma dağılımının olasılık yoğunluk fonksiyonunun grafiği.....	9
Şekil 2. 5: Farklı parametre değerleri için Gamma dağılımının bozulma fonksiyonunun grafiği.....	10
Şekil 2. 6: Farklı parametre değerleri için log-normal dağılımın olasılık yoğunluk fonksiyonunun grafiği.....	12
Şekil 2. 7: Farklı parametre değerleri için log-normal dağılımın bozulma fonksiyonunun grafiği.....	13
Şekil 2. 8: Farklı parametre değerleri için genelleştirilmiş üstel dağılımın olasılık yoğunluk fonksiyonunun grafiği.....	14
Şekil 2. 9: Farklı parametre değerleri için genelleştirilmiş üstel dağılımın bozulma fonksiyonunun fonksiyonun grafiği.....	15
Şekil 2. 10: Farklı parametre değerleri için Power Lindley dağılımının olasılık yoğunluk fonksiyonunun grafiği.....	17
Şekil 2. 11: Farklı parametre değerleri için Power Lindley dağılımının bozulma fonksiyonunun grafiği.....	18

ÇİZELGELER DİZİNİ

<u>ÇİZELGE</u>	<u>Sayfa</u>
Çizelge 5. 1: Veriler genelleştirilmiş üstel () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları.....	40
Çizelge 5. 2: Veriler genelleştirilmiş üstel () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları.....	41
Çizelge 5. 3: Veriler genelleştirilmiş üstel () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları.....	42
Çizelge 5. 4: Veriler Power Lindley () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları	43
Çizelge 5. 5: Veriler Power Lindley () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları	44
Çizelge 5. 6: Veriler Power Lindley () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları	45

1. GİRİŞ VE ÖNCEKİ ÇALIŞMALAR

Gözlenmiş bir verinin dağılımının belirlenmesi istatistikte oldukça eski bir problemdir. Her ne kadar Kolmogorov-Smirnov, Anderson Darling ve Ki-kare uyum iyiliği testleri bu problemin çözümü için yaygın olarak ve başarıyla kullanılan testler olsa da, çoğu zaman bu testler bir veri seti için birden fazla olası dağılım işaret edebilmektedir. Her ne kadar makul örneklem büyüklüğü için modeller benzer veri uyumuna sahip olsalar da doğru ya da doğruya yakın bir model seçimi arzu edilir. Özellikle, örneklem büyüklüğü küçük olduğunda da gözlemlere dayalı olarak mümkün en iyi kararı vermek oldukça önemlidir. Bir veri seti için uyum iyiliği testlerinin işaret ettiği olası dağılımlardan hangisi daha iyidir? Sorusunun cevabı Akaike bilgi kriteri (AIC), Bayes bilgi kriteri kullanılarak araştırılabilir. Ancak bu yöntemler mevcut olası dağılımlar arasından bir tanesini tavsiye niteliğinde öneren yöntemlerdir. Bu yöntemlerden başka olarak, bir veri setine model olabilecek olası iki dağılım arasından bir tanesinin seçilmesine yönelik en çok olabilirlik oran (RML) testi literatürde sıkça kullanılan bir yöntemdir.

Şu ana kadar, birçok araştırmacı karmaşık gerçek veri kümelerini optimal olarak modelleme amacıyla, veriler için uygun olduğu düşünülen alternatif dağılımlardan hangisinin seçileceğinin belirlenmesine yönelik bir çözüm arayışında birçok araştırmada bulunmuşlardır. Literatürde bir veri seti için olası iki dağılım arasından ayırım yapılmasına yönelik yapılmış çalışmalardan bazıları şöyledir.

Atkinson (1969, 1970), Cox (1961, 1962), Dyer (1973) verilen bir veri setinin herhangi iki olasılık dağılımından hangisi ile modelleneceği problemini ele almışlardır.

Dumonceaux ve Antle (1973) log-normal ve Weibull dağılımları arasından birisinin seçimi problemini RML yöntemine göre tartışmışlar ve doğru seçim olasılıklarını Monte Carlo simülasyonlarına dayalı olarak elde ederek kritik değerler için bir tablo sunmuşlardır.

Bain ve Engelhardt (1980), Weibull ve gama dağılımları arasında seçim yapmada RML yöntemini kullanmayı önermişlerdir. Araştırmacılar gamma ve Weibull dağılımları arasında doğru seçim olasılığının değerlerini Monte Carlo simülasyona göre elde etmişlerdir. Aynı problem Fearn ve Nebenzahl (1991) tarafından ele alınmış ve araştırmacılar gamma ve Weibull dağılımlarının ayırımı için kullanılan RML istatistiğinin asimptotik dağılımını elde etmişlerdir.

Gupta ve Kundu (2003), Weibull ile genelleştirilmiş üstel dağılımı, Gupta ve Kundu (2004), gamma ve genelleştirilmiş üstel dağılımı ayırmaya yönelik RML yöntemine dayalı kurallar vermişlerdir.

Kundu ve Manglick (2004), Kundu ve Manglick (2005) ve Kundu vd (2005) sırasıyla Weibull ile log-normal, log-normal ile gamma ve log-normal ile genelleştirilmiş üstel dağılımlarını ayırt etmeye yönelik RML istatistiklerini ve ilgili istatistiklere ait asimptotik dağılımları elde etmişlerdir.

İkinci tip sansürlü veriler için Weibull ve log-normal dağılımları arasında ayırım problemi Dey ve Kundu (2012) tarafından ele alınmış ve ilgili iki dağılım arasında ayırım yapabilmek için bir ayırıcı istatistik ve asimptotik dağılımı araştırmacılar tarafından verilmiştir.

Demirci Biçer ve Biçer (2017) ve Biçer ve Demirci Biçer (2018) geometrik süreçten gözlenen veri setleri için sırasıyla gamma ile Weibull ve gamma ile log-normal dağılımları arasından ayırım problemini incelemişlerdir.

Bu tez çalışmasının amacı, birbirinin alternatifi olan ve negatif değerli olmayan çarpık verileri modellemede kullanılan genelleştirilmiş üstel ve Power Lindley dağılımları arasındaki ayırım probleminin çözümünü RML yöntemine göre araştırmaktır.

Çalışmanın kalan bölümleri ise şu şekilde düzenlenmiştir. 2. Bölümde üstel, gamma, Weibull, log-normal, Power Lindley ve genelleştirilmiş üstel dağılımları ve ilgili dağılımların temel özellikleri hatırlatılacaktır.

Çalışmanın 3. bölümünde bir veri setini modelleyebilecek olası iki dağılım arasından ayırım yapmada kullanılan RML istatistiği genel anlamda verilmektedir.

Dördüncü bölümde, şu ana kadar yapılan çalışmalarla RML istatistiğine göre ortaya konmuş Weibull ile log-normal, gamma ile log-normal dağılım çiftleri arasındaki ayırım açıklama amaçlı olarak incelenmiştir. Genelleştirilmiş üstel ve Power Lindley dağılımları arasından ayırım probleminin çözümüne yönelik olarak RML istatistiği elde edilmektedir. Ayrıca elde edilen RML istatistiğinin asimptotik dağılımı yine bu bölümde verilmektedir.

Beşinci bölümde, genelleştirilmiş üstel ve Power Lindley dağılımları arasında ayırım yapabilmek için bu çalışmada elde edilen RML istatistiğinin doğru seçim performansını ortaya koymak için kapsamlı bir Monte-Carlo simülasyon çalışması verilmektedir.

Altıncı bölümde çalışmada elde edilen bulgular tartışılmaktadır.

2. BAZI DAĞILIMLAR

Bu bölümde aynı türden verilerin modellenmesinde kullanılacak üstel, Weibull, gamma, log-normal ve Power Lindley dağılımları ve bu dağılımların temel özellikleri hatırlatılacaktır.

2.1. Üstel Dağılım

X pozitif reel değerli bir rasgele değişken olmak üzere; eğer X rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x > 0, \theta > 0 \\ 0, & \text{diğer yerlerde (d.y.)} \end{cases} \quad (2.1.1)$$

ve dağılım fonksiyonu

$$F(x) = 1 - e^{-\frac{x}{\theta}}, \quad x > 0 \quad (2.1.2)$$

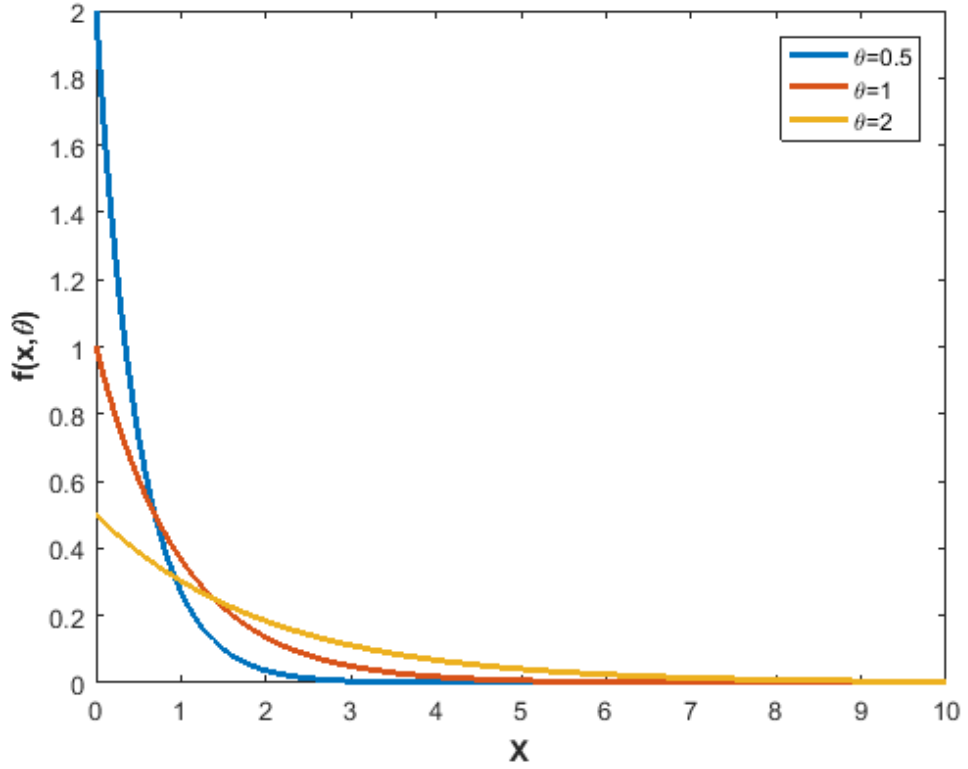
biçiminde ise X rasgele değişkenine üstel dağılıma sahiptir denir ve $X \sim \text{üstel}(\theta)$ biçiminde gösterilir. Burada θ dağılımın ölçek parametresidir. Üstel dağılımın beklenen değeri, varyansı ve moment çıkarıcı fonksiyonu sırasıyla

$$E(X) = \theta \quad (2.1.3)$$

$$\text{Var}(X) = \theta^2 \quad (2.1.4)$$

$$M_x(t) = \frac{1}{1 - \theta t}, \quad t < \frac{1}{\theta} \quad (2.1.5)$$

biçimindedir. Üstel dağılım üstel aileye mensup tek tepeli bir dağılım olup farklı parametre değerleri için dağılımın olasılık yoğunluk fonksiyonunun grafiği Şekil 2.1.1 de verilmiştir.



Şekil 2. 1: Farklı parametre değerleri için üstel dağılımın olasılık yoğunluk fonksiyonunun grafiği

Ayrıca üstel dağılımın bozulma (hazard) fonksiyonu;

$$h(x, \theta) = \frac{\theta e^{-\frac{x}{\theta}}}{1 - \left(1 - e^{-\frac{x}{\theta}}\right)} = \theta \quad (2.1.6)$$

dır. Eşitlik (2.1.6)'dan açıkça görüleceği gibi dağılım sabit değerli bozulma fonksiyonuna sahiptir.

Üstel dağılım yaşam sürelerinin modellenmesinde kullanılan gamma, Weibull ve Rayleigh dağılımları ile ilişkili bir dağılımdır.

2.2. Weibull Dağılımı

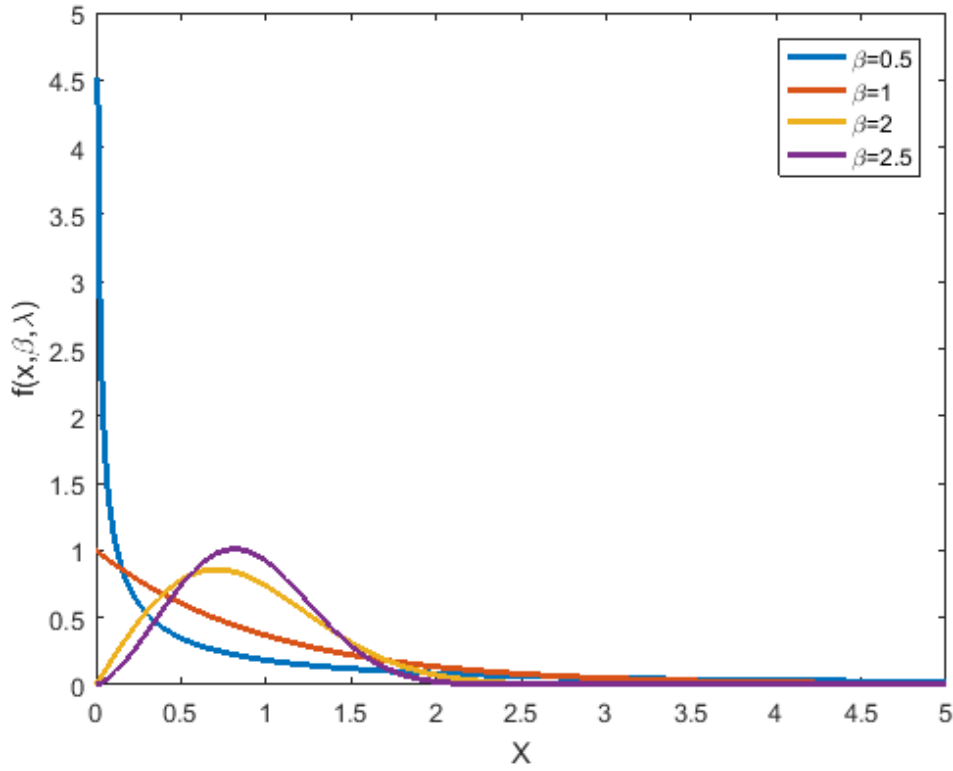
Bir X rasgele değişkenin olasılık yoğunluk fonksiyonu

$$f_{WE}(x, \beta, \lambda) = \beta \lambda^\beta x^{\beta-1} e^{-(x\lambda)^\beta}, x > 0 \quad (2.2.1)$$

biçiminde ise X rasgele değişkenine Weibull dağılımına sahiptir denir. Burada $\beta > 0$ dağılımın ölçek parametresi $\lambda > 0$ ise dağılımın şekil parametresidir. Ayrıca Weibull dağılımın dağılım fonksiyonu

$$F_{WE}(x, \beta, \lambda) = 1 - e^{-(x\lambda)^\beta}, x > 0 \quad (2.2.2)$$

ile verilir. Farklı parametre değerleri için Weibull dağılımının olasılık yoğunluk fonksiyonunun grafiği Şekil 2.2.1 de verilmiştir.



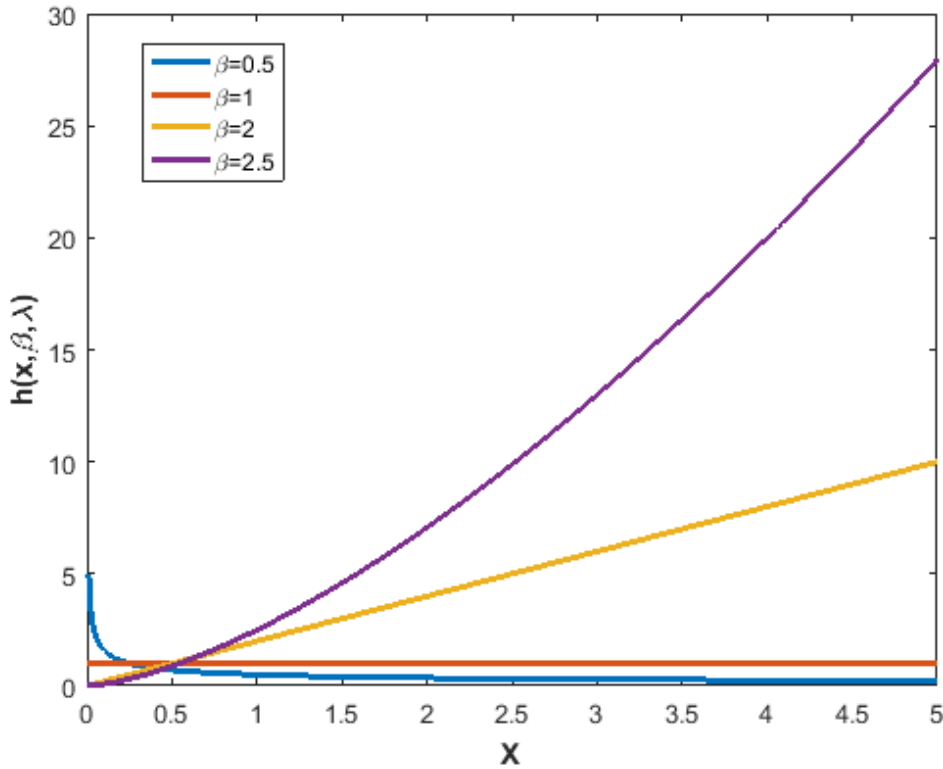
Şekil 2.2: Farklı parametre değerleri için Weibull dağılımının olasılık yoğunluk fonksiyonunun grafiği

Güvenilirlik ve yaşam analizi uygulamaları gibi birçok alandan elde edilen farklı tiplerden pozitif verinin modellenmesinde Weibull dağılımı başarı ile kullanılan

esnek bir dağılımdır. Güvenilirlik analizi uygulamalarında sıkça kullanılan Weibull dağılımı için bozulma fonksiyonu

$$h(x, \lambda, \beta) = \frac{\beta \lambda^\beta x^{\beta-1} e^{-(x\lambda)^\beta}}{1 - (1 - e^{-(x\lambda)^\beta})} = \beta \lambda^\beta x^{\beta-1} \quad (2.2.3)$$

olup, parametrelerin farklı değerleri için (2.2.3) ile verilen bozulma fonksiyonunun grafiği Şekil 2.2.2 de verildiği gibi olmaktadır.



Şekil 2.3: Farklı parametre değerleri için Weibull dağılımının bozulma fonksiyonunun grafiği

Weibull dağılımının beklenen değeri, varyansı ve moment çıkarıcı fonksiyonu sırasıyla

$$E(X) = \frac{1}{\lambda} \Gamma\left(1 + \frac{1}{\beta}\right) \quad (2.2.4)$$

$$Var(X) = \frac{1}{\lambda^2} \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \left(\Gamma\left(1 + \frac{1}{\beta}\right) \right)^2 \right] \quad (2.2.5)$$

ve

$$M_x(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{1}{\lambda^n} \Gamma\left(1 + \frac{n}{\beta}\right), \beta \geq 1 \quad (2.2.6)$$

ile verilir. Üstel dağılımda olduğu gibi Weibull dağılımı da üstel aileye mensup tek tepeli bir dağılımdır.

2.3. Gamma Dağılımı

Bir X rasgele değişkenin olasılık yoğunluk fonksiyonu

$$f(x; \alpha, \beta) = x^{\alpha-1} \frac{e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}, \quad x > 0 \quad (2.3.1)$$

biçiminde ise X rasgele değişkenine gamma dağılımına sahiptir denir. Burada $\beta > 0$ dağılımın ölçek parametresi $\alpha > 0$ ise dağılımın şekil parametresidir. Ayrıca gamma dağılımın dağılım fonksiyonu

$$F(x, \alpha, \beta) = \frac{\gamma(\alpha, x/\beta)}{\Gamma(\alpha)}, \quad x > 0 \quad (2.3.2)$$

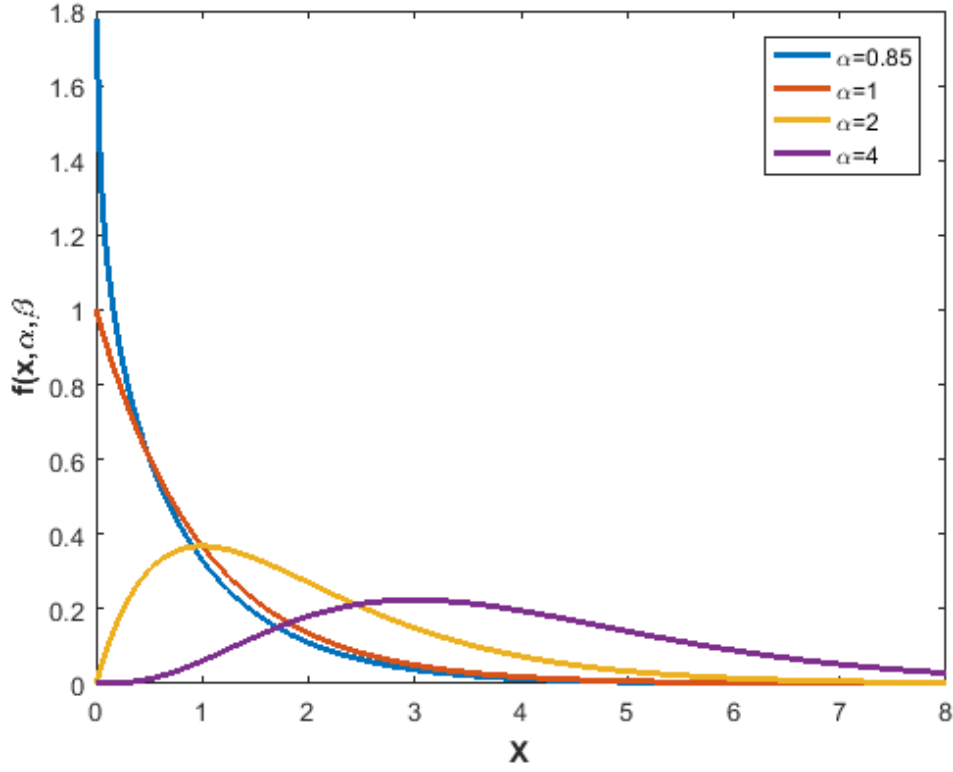
ile verilir, burada $\gamma(\alpha, x/\beta)$

$$\gamma(\alpha, x/\beta) = \int_0^{x/\beta} t^{\alpha-1} e^{-t} dt \quad (2.3.3)$$

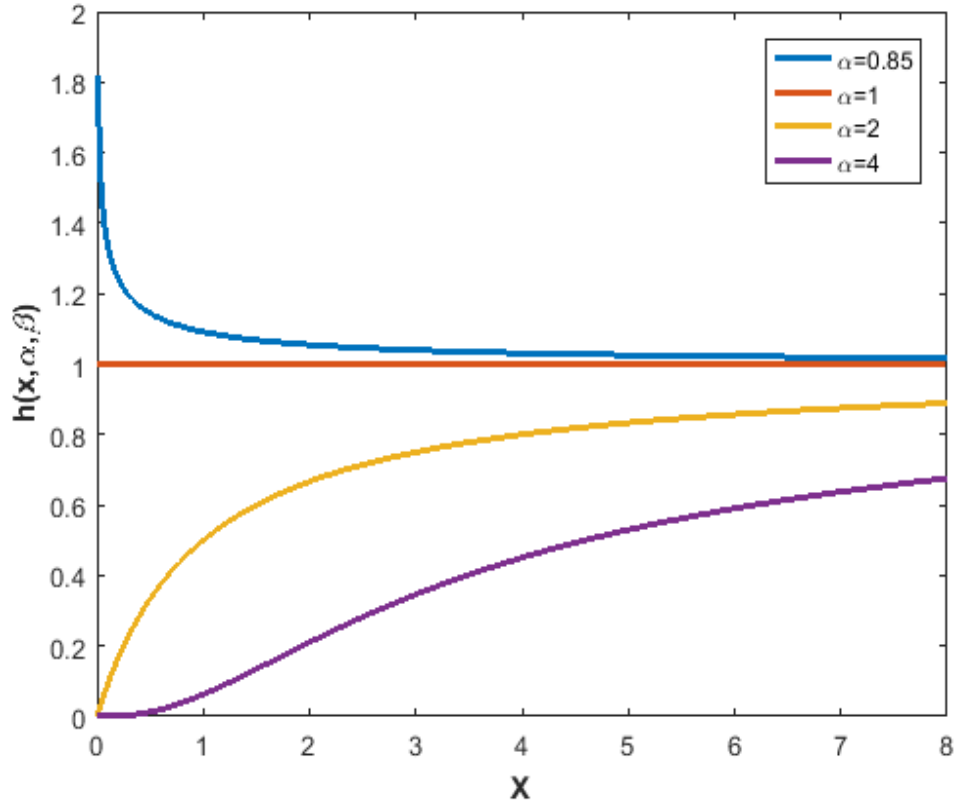
biçiminde tanımlı tam olmayan gamma fonksiyonunu ifade etmektedir . Ayrıca (2.3.1) eşitliği ile verilen olasılık yoğunluk fonksiyonunun ve (2.3.2) de verilen dağılım fonksiyonunun kullanılmasıyla, gamma dağılımının bozulma fonksiyonu

$$h(x, \alpha, \beta) = \frac{1/(\Gamma(\alpha) \beta^\alpha) x^{\alpha-1} e^{-(x/\beta)}}{1 - \gamma(\alpha, x/\beta)/\Gamma(\alpha)}, \quad x > 0 \quad (2.3.4)$$

biçiminde elde edilir. Weibull dağılımında olduğu gibi gamma dağılımı da güvenilirlik ve yaşam analizi uygulamalarında oldukça yoğun olarak kullanılan önemli bir dağılımdır. α ve β parametrelerin farklı değerleri için gamma dağılımının olasılık yoğunluk fonksiyonunun ve bozulma fonksiyonunun grafikleri Şekil 2.3.1 ve Şekil 2.3.2 de verildiği gibi olmaktadır.



Şekil 2. 4: Farklı parametre değerleri için Gamma dağılımının olasılık yoğunluk fonksiyonunun grafiği



Şekil 2. 5: Farklı parametre değerleri için Gamma dağılımının bozulma fonksiyonunun grafiği

Gamma dağılımının beklenen değeri, varyansı ve moment çıkaran fonksiyonu sırasıyla

$$E(X) = \alpha\beta \quad (2.3.5)$$

$$Var(X) = \alpha\beta^2 \quad (2.3.6)$$

ve

$$M_x(t) = (1 - \beta t)^{-\alpha}, \quad t < 1/\beta \quad (2.3.7)$$

dır.

2.4. Log-Normal Dağılım

Bir X rasgele değişkenin olasılık yoğunluk fonksiyonu

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} \quad (2.4.1)$$

biçiminde ise X rasgele değişkenine Log-Normal dağılımına sahiptir denir. Burada μ dağılımın reel değerli konum parametresi $\sigma > 0$ ise dağılımın ölçek parametresidir. Log-normal dağılımın dağılım fonksiyonu

$$F(x; \mu, \sigma) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[\frac{\ln(x) - \mu}{\sigma\sqrt{2}} \right] \quad (2.4.2)$$

ile verilir. Burada $\operatorname{erf}(k)$

$$\operatorname{erf}(k) = \frac{2}{\sqrt{\pi}} \int_0^k e^{-t^2} dt \quad (2.4.3)$$

biçiminde tanımlı hata fonksiyonudur (Abramowitz ve Stegun,1964). Log-normal dağılımın beklenen değeri ve varyansı sırasıyla

$$E(X) = e^{\mu + \sigma^2/2} \quad (2.4.4)$$

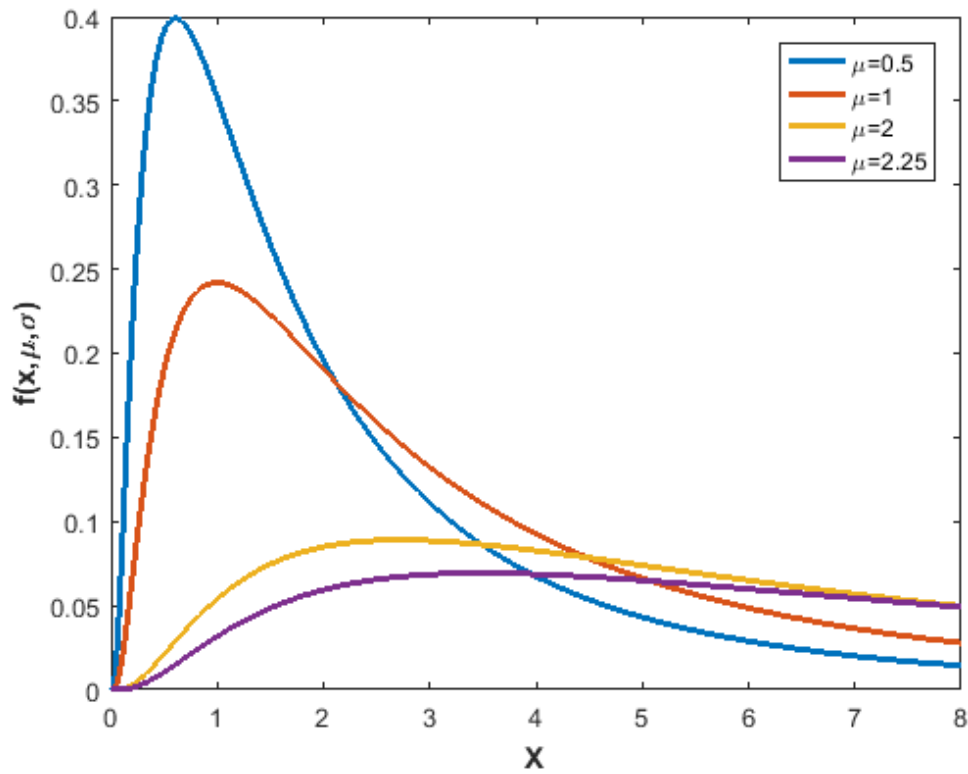
Ve

$$\operatorname{Var}(X) = (e^{\sigma^2} - 1) e^{2\mu + \sigma^2} \quad (2.4.5)$$

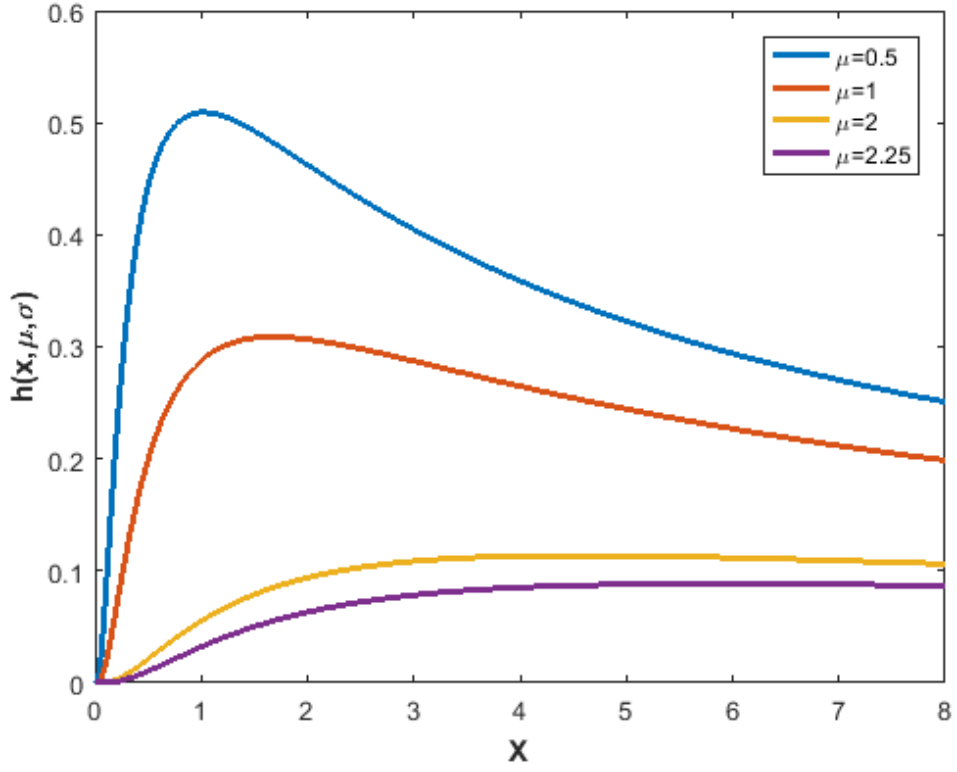
dır. Log-normal dağılmış bir X rasgele değişken için bozulma fonksiyonu (2.4.1) ve (2.4.2) eşitliklerinde verilen olasılık yoğunluk ve dağılım fonksiyonlarının kullanılmasıyla

$$h(x, \alpha, \beta) = \frac{\sqrt{\frac{2}{\pi}} \frac{1}{x\sigma} \exp\left(-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right)}{1 + \operatorname{erf}\left[\frac{\ln(x) - \mu}{\sigma\sqrt{2}}\right]}, \quad x > 0 \quad (2.4.6)$$

biçiminde yazılır. Log-normal dağılımın olasılık yoğunluk ve bozulma fonksiyonlarının farklı parametrelere göre grafikleri Şekil 2.4.1- 2.4.2 de verilmektedir.



Şekil 2. 6: Farklı parametre değerleri için log-normal dağılımın olasılık yoğunluk fonksiyonunun grafiği



Şekil 2. 7: Farklı parametre değerleri için log-normal dağılımın bozulma fonksiyonunun grafiği

2.5. Genelleştirilmiş Üstel Dağılım

Bir X rasgele değişkenin olasılık yoğunluk fonksiyonu

$$f(x; \alpha, \lambda) = \alpha \lambda (1 - e^{-\lambda x})^{\alpha-1} e^{-\lambda x}, \quad x > 0 \quad (2.5.1)$$

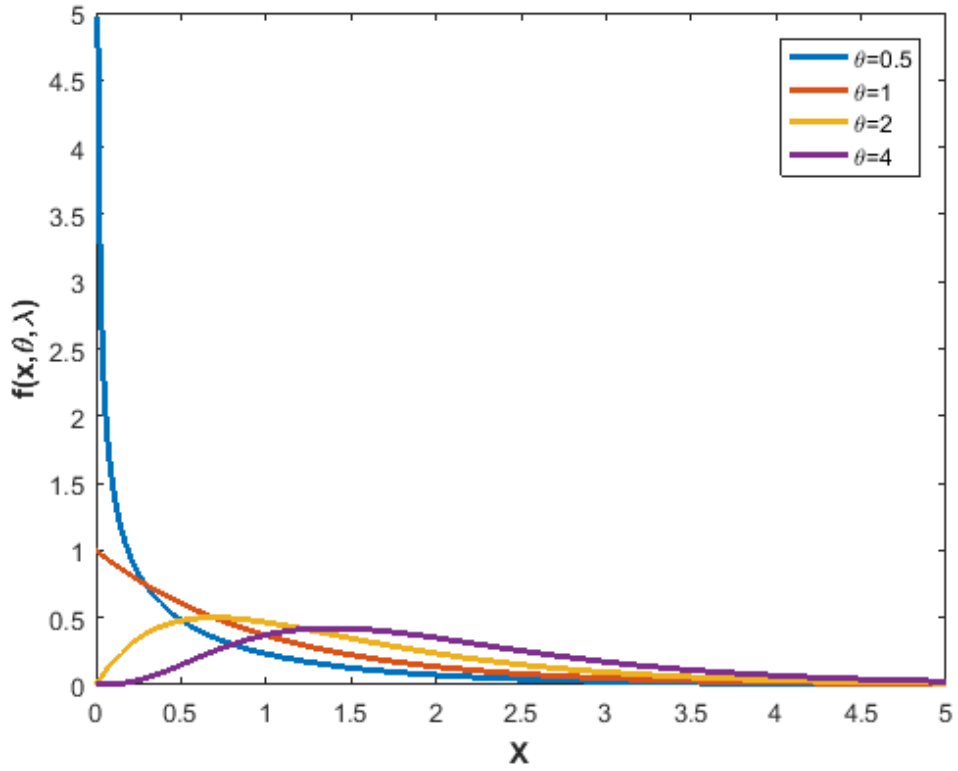
biçiminde oluyorsa X rasgele değişkenine genelleştirilmiş üstel dağılıma sahiptir denir. Burada $\alpha > 0$ dağılımın şekil parametresi $\beta > 0$ ise dağılımın ölçek parametresidir. Genelleştirilmiş üstel dağılım için dağılım fonksiyonu,

$$F(x; \alpha, \lambda) = (1 - e^{-\lambda x})^\alpha, \quad x > 0 \quad (2.5.2)$$

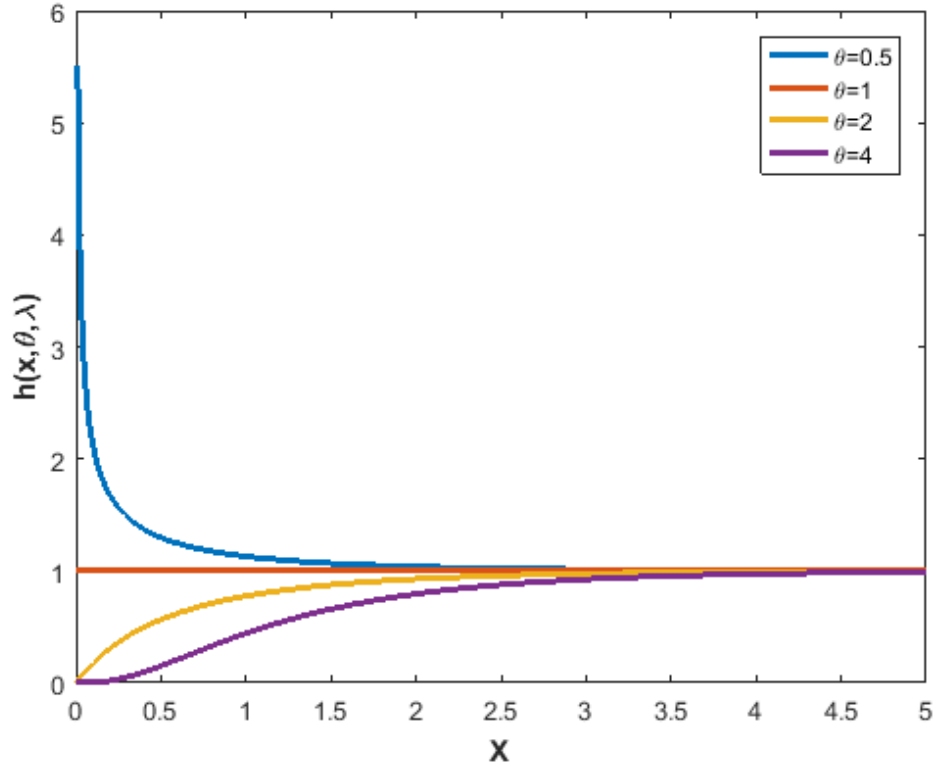
olup dağılımın bozulma fonksiyonu eşitlik (2.5.1) ve (2.5.2) den kolayca

$$h(x, \alpha, \lambda) = \frac{\alpha \lambda (1 - e^{-\lambda x})^{\alpha-1} e^{-\lambda x}}{1 - (1 - e^{-\lambda x})^\alpha} \quad (2.5.3)$$

olarak elde edilir (Gupta ve Kundu,1999). $\lambda = 1$ alınarak α (şekil) parametresinin değişik değerleri için genelleştirilmiş üstel dağılımın olasılık yoğunluk ve bozulma fonksiyonu grafikleri Şekil 2.5.1 ve Şekil 2.5.2' de verilmiştir.



Şekil 2. 8: Farklı parametre değerleri için genelleştirilmiş üstel dağılımın olasılık yoğunluk fonksiyonunun grafiği



Şekil 2. 9: Farklı parametre değerleri için genelleştirilmiş üstel dağılımın bozulma fonksiyonunun fonksiyonun grafiği

Genelleştirilmiş üstel dağılımın beklenen değer, varyans ve moment çıkaran fonksiyonu, sırasıyla;

$$E(X) = \frac{1}{\lambda} (\Psi(\alpha+1) - \Psi(1)) \quad (2.5.4)$$

$$Var(X) = \frac{1}{\lambda^2} (\Psi'(1) - \Psi'(\alpha+1)) \quad (2.5.6)$$

ve

$$M_x(t) = \alpha \lambda \Gamma(\alpha) \Gamma\left(1 - \frac{t}{\lambda}\right) / \Gamma\left(\alpha + 1 - \frac{t}{\lambda}\right) \quad (2.5.7)$$

dır. Burada $\Psi(\alpha) = \frac{d \ln(\alpha)}{d\alpha}$ biçiminde tanımlı Psi fonksiyonunu ifade etmektedir.

2.6. Power Lindley Dağılımı

Power Lindley dağılımı ilk kez Githany vd (2013) tarafından pozitif değerli verilerin modellenmesi için önerilmiş Gamma, Weibul, Log-Normal, Genelleştirilmiş üstel dağılım gibi popüler dağılımlara önemli bir alternatif olan bir dağılımdır. Power Lindley dağılımının olasılık yoğunluk fonksiyonu

$$f(x, \theta, \beta) = \frac{\theta\beta^2}{\beta+1} (1+x^\theta) x^{\theta-1} e^{-\beta x^\theta} \quad (2.6.1)$$

ve dağılım fonksiyonu

$$F(x, \theta, \beta) = 1 - \left(1 + \frac{\beta}{\beta+1} x^\theta\right) e^{-\beta x^\theta}, \quad x > 0 \quad (2.6.2)$$

biçimindedir. Burada $\beta > 0$ ölçek parametresi, $\theta > 0$ ise şekil (ters şekil) parametresidir. Esasında, Power Lindley dağılımı, (θ, β) parametrelili bir Weibull ve $(2, \theta, \beta)$ parametrelili genelleştirilmiş gamma dağılımlarının iki bileşenli bir karmasıdır. Dağılımın beklenen değeri, varyansı ve moment çıkaran fonksiyonu, sırasıyla;

$$E(X) = \frac{1}{\lambda} (\Psi(\alpha+1) - \Psi(1)) \quad (2.6.3)$$

$$Var(X) = \frac{1}{\lambda^2} (\Psi'(1) - \Psi'(\alpha+1)) \quad (2.6.4)$$

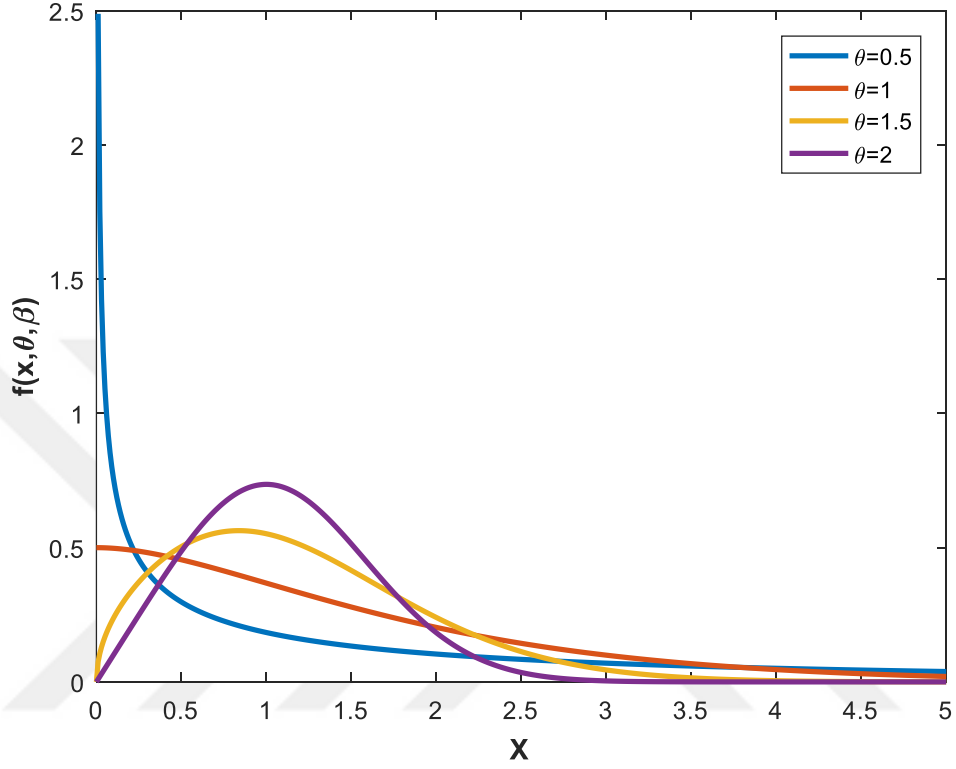
ve

$$M_x(t) = \alpha\lambda\Gamma(\alpha)\Gamma\left(1-\frac{t}{\lambda}\right) / \Gamma\left(\alpha+1-\frac{t}{\lambda}\right) \quad (2.6.5)$$

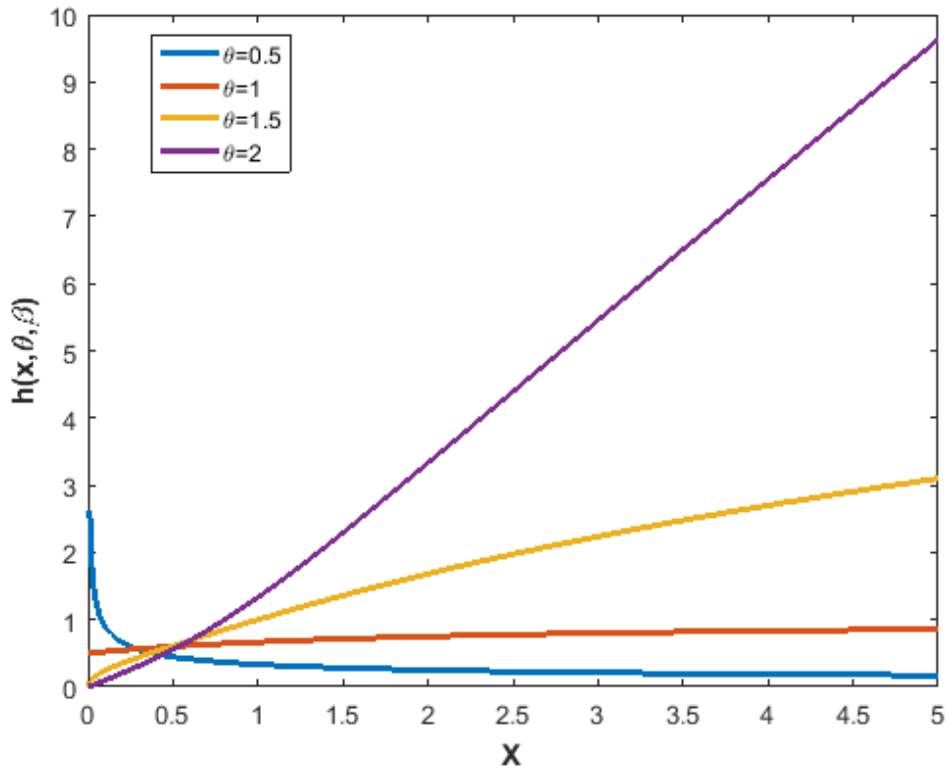
dır (Githany vd, 2013). Dağılımın bozulma fonksiyonu ise

$$H(x) = \theta\beta^2 \frac{(1+x^\theta)x^{\theta-1}}{\beta+1+\beta x^\theta} \quad (2.6.6)$$

biçiminde olup farklı parametre değerleri için Power Lindley dağılımının olasılık yoğunluk fonksiyonu ve bozulma fonksiyonlarına ait grafikler Şekil 2.6.1- 2.6.2 de verilmiştir.



Şekil 2. 10: Farklı parametre değerleri için Power Lindley dağılımının olasılık yoğunluk fonksiyonunun grafiği



Şekil 2. 11: Farklı parametre değerleri için Power Lindley dağılımının bozulma fonksiyonunun grafiği

3. OLABİLİRLİK ORANINA DAYALI AYIRICI İSTATİSTİK

Bu kısımda aralarında ayırım yapılması düşünülen ve birbirinin alternatifi olan iki dağılım için en büyüklenmiş olabilirlik fonksiyonlarının oranına dayalı olarak tanımlanan ayırıcı istatistik verilecektir.

X_1, X_2, \dots, X_n rasgele örnekleme $f(x, \theta_1, \theta_2)$ veya $g(x, \beta_1, \beta_2)$ olasılık yoğunluk fonksiyonlarına sahip dağılımların herhangi birinden alınmış bir rasgele örneklem olmak üzere; X_1, X_2, \dots, X_n , $f(x, \theta_1, \theta_2)$ olasılık yoğunluk fonksiyonlu dağılımdan alınmış olduğunda X_1, X_2, \dots, X_n rasgele değişkenlerinin olabilirlik fonksiyonu

$$L_f = \prod_{i=1}^n f(x_i, \theta_1, \theta_2) \quad (3.1)$$

ve (3.1) ile verilen olabilirlik fonksiyonun logaritması

$$\ln L_f = \sum_{i=1}^n \ln f(x_i, \theta_1, \theta_2) \quad (3.2)$$

biçiminde ve benzer olarak, X_1, X_2, \dots, X_n , $g(x, \beta_1, \beta_2)$ olasılık yoğunluk fonksiyonlu dağılımdan alınmış olduğunda X_1, X_2, \dots, X_n rasgele değişkenlerinin olabilirlik fonksiyonu

$$L_g = \prod_{i=1}^n g(x_i, \beta_1, \beta_2) \quad (3.3)$$

ve (3.1) ile verilen olabilirlik fonksiyonun logaritması

$$\ln L_g = \sum_{i=1}^n \ln g(x_i, \beta_1, \beta_2) \quad (3.4)$$

biçiminde gösterilsin. Bu gösterimler altında f ve g dağılımlarının ayırımı için kullanılacak bir test istatistiği en büyüklenmiş olabilirlik fonksiyonlarının oranına dayalı olarak

$$T = \ln \left[\frac{L_f}{L_g} \right] \quad (3.5)$$

biçiminde tanımlanır. Ayrıca her iki dağılımın parametrelerinin tahmin edilmesi ve elde edilen tahmin değerlerinin f ve g dağılımlarında yerlerinde kullanılmasıyla (3.5) ile tanımlanan T istatistiğinin örnekleme bağlı değeri

$$T = \ln \left[\frac{L_{\hat{f}}}{L_{\hat{g}}} \right] \quad (3.6)$$

biçiminde verilir. Böylece (3.6) ile hesaplanan T istatistiğinin değeri göz önünde bulundurularak; $T > 0$ ise X_1, X_2, \dots, X_n rastgele örnekleme f olasılık yoğunluk fonksiyonlu dağılımdan alınmıştır, aksi halde X_1, X_2, \dots, X_n örnekleme g olasılık yoğunluk fonksiyonlu bir dağılımın bir örneklemdir, biçiminde karar verilir.

(3.5) de verilen T istatistiği kullanılarak yapılan test için testin gücü ve anlam düzeyi hakkında bir şey söylenemez. Yapılan testin gücü ve anlam düzeyi hakkında bir şey söylenebilmesi için T istatistiğinin dağılımının belirlenmesi gerekmektedir. Ancak henüz hiçbir durumda T istatistiğinin dağılımı açık olarak belirlenememiştir. Dolayısıyla Testin gücü ve anlam düzeyi hakkında yaklaşık da olsa bir değer verilebilmesi için kullanılan istatistiğin asimptotik dağılımı oldukça önem kazanmaktadır. T istatistiğinin asimptotik dağılımı örneklemin f olasılık yoğunluk fonksiyonlu dağılımdan veya g olasılık yoğunluk fonksiyonlu dağılımdan alınmış olması durumları göz önünde bulundurularak belirlenebilmektedir.

Lemma 3.1: X_1, X_2, \dots, X_n rasgele değişkenleri $f(x, \theta_1, \theta_2)$ olasılık yoğunluk fonksiyonlu dağılımın bir örnekleme olsun. Bu durumda $n \rightarrow \infty$ için

$$i) \hat{\theta}_1 \xrightarrow{h.h.h.y} \theta_1 \text{ ve } \hat{\theta}_2 \xrightarrow{h.h.h.y} \theta_2$$

$$E_f \left(\ln \left(f(x; \theta_1, \theta_2) \right) \right) = \max_{\hat{\theta}_1, \hat{\theta}_2} E_f \left(\ln \left(f(x; \hat{\theta}_1, \hat{\theta}_2) \right) \right)$$

$$ii) \hat{\beta}_1 \xrightarrow{h.h.h.y} \tilde{\beta}_1, \hat{\beta}_2 \xrightarrow{h.h.h.y} \tilde{\beta}_2$$

$$E_f \left(\ln \left(g(x; \tilde{\beta}_1, \tilde{\beta}_2) \right) \right) = \max_{\beta_1, \beta_2} E_f \left(\ln \left(g(x; \beta_1, \beta_2) \right) \right)$$

iii) $n^{-\frac{1}{2}}[T - E_f(T)]$ ve $n^{-\frac{1}{2}}[T^* - E_f(T^*)]$ asimptotik olarak denktir. Burada

$$T^* = \ln \left[\frac{L_f(\hat{\theta}_1, \hat{\theta}_2)}{L_g(\tilde{\beta}_1, \tilde{\beta}_2)} \right] \text{ dir.}$$

varsayımları altında T istatistiği $E_f(T)$ beklenen değeri ve $V_f(T)$ varyansı ile asimptotik normal dağılır. Burada

$$\frac{E_f(T)}{n} = E_f \left(\ln f(x, \hat{\theta}_1, \hat{\theta}_2) \right) - E_f \left(\ln g(x, \tilde{\beta}_1, \tilde{\beta}_2) \right)$$

ve

$$\frac{V_f(T)}{n} = V_f \left(\ln f(x, \hat{\theta}_1, \hat{\theta}_2) - \ln g(x, \tilde{\beta}_1, \tilde{\beta}_2) \right)$$

olarak hesaplanır. (Gupta ve Kundu,2004)

Lemma 3.2: X_1, X_2, \dots, X_n rasgele değişkenleri $g(x, \beta_1, \beta_2)$ olasılık yoğunluk fonksiyonlu dağılımın bir örnekleme olsun. Bu durumda $n \rightarrow \infty$ için

$$i) \hat{\theta}_1 \xrightarrow{h.h.h.y} \theta_1 \text{ ve } \hat{\theta}_2 \xrightarrow{h.h.h.y} \theta_2$$

$$E_f \left(\ln \left(f(x; \theta_1, \theta_2) \right) \right) = \max_{\hat{\theta}_1, \hat{\theta}_2} E_f \left(\ln \left(f(x; \hat{\theta}_1, \hat{\theta}_2) \right) \right)$$

$$ii) \hat{\beta}_1 \xrightarrow{h.h.h.y} \tilde{\beta}_1, \hat{\beta}_2 \xrightarrow{h.h.h.y} \tilde{\beta}_2$$

$$E_f \left(\ln \left(g(x; \tilde{\beta}_1, \tilde{\beta}_2) \right) \right) = \max_{\beta_1, \beta_2} E_f \left(\ln \left(g(x; \beta_1, \beta_2) \right) \right)$$

iii) $n^{-\frac{1}{2}}[T - E_f(T)]$ ve $n^{-\frac{1}{2}}[T^* - E_f(T^*)]$ asimptotik olarak denktir. Burada

$$T^* = \ln \left[\frac{L_f(\hat{\theta}_1, \hat{\theta}_2)}{L_g(\tilde{\beta}_1, \tilde{\beta}_2)} \right] \text{ dir.}$$

varsayımları altında T istatistiğinin dağılımı $E_g(T)$ beklenen değeri ve $V_g(T)$ varyansı ile asimptotik normal olmaktadır. Burada

$$\frac{E_g(T)}{n} = E_g(\ln f(x, \tilde{\theta}_1, \tilde{\theta}_2)) - E_g(\ln g(x, \hat{\beta}_1, \hat{\beta}_2))$$

ve

$$\frac{V_g(T)}{n} = V_g(\ln f(x, \tilde{\theta}_1, \tilde{\theta}_2) - \ln g(x, \hat{\beta}_1, \hat{\beta}_2))$$

olarak hesaplanır.



4. BAZI DAĞILIM ÇİFTLERİ İÇİN RML İSTATİSTİĞİ

Bu kısımda aralarında ayırım yapılması düşünülen ve birbirinin alternatifi olan iki dağılım için en büyüklenmiş olabilirlik fonksiyonlarının oranına dayalı olarak tanımlanan ayırıcı istatistik verilecektir.

4.1. Weibull ve Log-Normal Dağılımları Arasında Ayırım

X_1, X_2, \dots, X_n rasgele örnekleme $f_{LN}(x, \sigma, \theta)$ veya $f_{WE}(x, \beta, \lambda)$ olasılık yoğunluk fonksiyonlarına sahip dağılımların herhangi birinden alınmış bir rasgele örneklem olmak üzere; $X_1, X_2, \dots, X_n, f_{LN}(x, \sigma, \theta)$ olasılık yoğunluk fonksiyonlu dağılımdan alınmış olduğunda X_1, X_2, \dots, X_n rasgele değişkenlerinin olabilirlik fonksiyonu

$$L_{LN} = \prod_{i=1}^n f_{LN}(x_i, \sigma, \theta) \quad (4.1.1)$$

ve (4.1.1) ile verilen olabilirlik fonksiyonun logaritması

$$\ln L_{LN} = \sum_{i=1}^n \ln f_{LN}(x_i, \sigma, \theta) \quad (4.1.2)$$

biçiminde ve benzer olarak, $X_1, X_2, \dots, X_n, f_{WE}(x, \beta, \lambda)$ olasılık yoğunluk fonksiyonlu dağılımdan alınmış olduğunda X_1, X_2, \dots, X_n rasgele değişkenlerinin olabilirlik fonksiyonları

$$L_{WE} = \prod_{i=1}^n f_{WE}(x_i, \beta, \lambda) \quad (4.1.3)$$

ve (4.1.1) ile verilen olabilirlik fonksiyonun logaritması

$$\ln L_{WE} = \sum_{i=1}^n \ln f_{WE}(x_i, \beta, \lambda) \quad (4.1.4)$$

biçiminde gösterilsin. Bu gösterimler altında log-Normal ve Weibull dağılımlarının ayırımı için kullanılacak bir test istatistiği (3.5) ve (3.6) eşitlikleri göz önünde bulundurulması ve f_{LN} ve f_{WE} fonksiyonlarının kullanılmasıyla

$$\begin{aligned}
T &= \ln \left(\frac{L_{LN}(\hat{\sigma}, \hat{\theta})}{L_{WE}(\hat{\beta}, \hat{\lambda})} \right) \\
&= \ln \left(\prod_{i=1}^n f_{LN}(x_i, \hat{\sigma}, \hat{\theta}) \right) - \ln \left(\prod_{i=1}^n f_{WE}(x_i, \hat{\beta}, \hat{\lambda}) \right) \\
&= \sum_{i=1}^n \ln \left(f_{LN}(x_i, \hat{\sigma}, \hat{\theta}) \right) - \sum_{i=1}^n \ln \left(f_{WE}(x_i, \hat{\beta}, \hat{\lambda}) \right) \\
&= \sum_{i=1}^n \ln \left(\underbrace{\frac{1}{\sqrt{2\pi} x_i \hat{\sigma}} e^{-\left[\frac{(\ln x_i - \ln \hat{\theta})^2}{2\hat{\sigma}^2} \right]}}_A \right) - \sum_{i=1}^n \ln \left(\underbrace{\hat{\beta} \hat{\lambda}^{\hat{\beta}} x_i^{\hat{\beta}-1} e^{-x_i \hat{\lambda}^{\hat{\beta}}}}_B \right) \\
&= A + B
\end{aligned} \tag{4.1.5}$$

olur. Eşitlik (4.1.5) de

$$\begin{aligned}
A &= \sum_{i=1}^n \ln 2\pi - \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln \hat{\sigma} - \sum_{i=1}^n \left[\ln x_i - \ln \hat{\theta} \right]^2 / 2\hat{\sigma}^2 \\
&= -\frac{n}{2} \ln 2\pi - n \ln \hat{\theta} - n \ln \hat{\sigma} - \frac{n}{2}
\end{aligned} \tag{4.1.6}$$

ve

$$\begin{aligned}
B &= \sum_{i=1}^n \ln \left(\hat{\beta} \hat{\lambda}^{\hat{\beta}} x_i^{\hat{\beta}-1} e^{-x_i \hat{\lambda}^{\hat{\beta}}} \right) \\
&= \sum_{i=1}^n \ln \hat{\beta} + \hat{\beta} \sum_{i=1}^n \ln \hat{\lambda} + (\hat{\beta} - 1) \sum_{i=1}^n \ln x_i - \hat{\lambda}^{\hat{\beta}} \sum_{i=1}^n x_i^{\hat{\beta}}
\end{aligned} \tag{4.1.7}$$

dir. (4.1.6) ve (4.1.7) eşitliklerinin (4.1.5) eşitliğinde yerlerinde kullanılmasıyla

$$T = n \left(\frac{1}{2} - \ln \left(\hat{\sigma} \hat{\beta} (\hat{\lambda} \hat{\theta})^{\hat{\beta}} \sqrt{2\pi} \right) \right) \tag{4.1.8}$$

elde edilir.

Böylece (4.18) ile elde edilen T istatistiğinin değeri 0'dan büyükse veri seti için Log-Normal dağılımın daha uygun olduğu söylenebilir aksi halde Weibull dağılımının veri setinin modellenmesinde daha uygun olduğu sonucuna varılır.

(4.18) ile elde edilen T istatistiğinin asimptotik dağılımı 3. Bölümde verilen Lemma 1 ve Lemma 2 göz önüne alınarak Kundu ve Manglick (2004) tarafından elde edilmiştir.

Verilerin log-normal dağılımdan geldiği varsayımı altında üçüncü bölümde verilen Lemma 1'in göz önüne alınmasıyla (4.1.8) eşitliğinde verilen istatistiğinin dağılımı $E_{LN}(T)$ beklenen değeri ve $V_{LN}(T)$ varyansı ile asimptotik normaldir. Burada

$$\begin{aligned}
\frac{E_{LN}(T)}{n} &= \int_0^{\infty} T f_{LN}(x; \sigma, \theta) dx \\
&= \int_0^{\infty} \ln \left[\frac{L_{LN}(x, \sigma, \theta)}{L_{WE}(x, \tilde{\beta}, \tilde{\lambda})} \right] f_{LN}(x; \sigma, \theta) dx \\
&= \underbrace{\int_0^{\infty} \ln(f_{LN}(x, \sigma, \theta)) f_{LN}(x; \sigma, \theta) dx}_A - \underbrace{\int_0^{\infty} \ln(f_{WE}(x, \tilde{\beta}, \tilde{\lambda})) f_{LN}(x; \sigma, \theta) dx}_B
\end{aligned} \tag{4.1.9}$$

olup

$$\begin{aligned}
A &= \int_0^{\infty} \ln \left[\frac{1}{\sqrt{2\pi} x \sigma} e^{-[(\ln x - \ln \theta)^2 / 2\sigma^2]} \right] f_{LN}(x; \sigma, \theta) dx \\
&= \int_0^{\infty} \left[\underbrace{-\frac{1}{2} \ln 2\pi}_C - \underbrace{\ln x}_D - \underbrace{\ln \sigma}_E - \underbrace{[(\ln x - \ln \theta)^2 / 2\sigma^2]}_F \right] f_{LN}(x; \sigma, \theta) dx
\end{aligned}$$

dır. Burada

$$C = \int_0^{\infty} \left[\frac{1}{2} \ln 2\pi \right] f_{LN}(x; \sigma, \theta) dx = \frac{1}{2} \ln 2\pi,$$

$$D = \int_0^{\infty} \underbrace{[\ln x]}_{E(\ln x)} f_{LN}(x; \sigma, \theta) dx,$$

$$E = \ln \sigma,$$

$$D = \theta e^{\frac{1}{2}\sigma^2},$$

$$F = e^{2\sigma^2} \theta^2.$$

Ayrıca (4.1.9) eşitliğinde

$$B = \ln \tilde{\beta} + \tilde{\beta} \ln \tilde{\lambda} + (\tilde{\beta} - 1) \ln \theta - (\tilde{\lambda})^{\tilde{\beta}} e^{\tilde{\beta} \ln \theta + \frac{\tilde{\beta}}{2} \sigma^2}$$

$$= \ln \tilde{\beta} + \tilde{\beta} \ln \tilde{\lambda} + (\tilde{\beta} - 1) \ln \theta - (\tilde{\lambda} \theta)^{\tilde{\beta}} e^{\tilde{\beta} \sigma^2 / 2}$$

olarak hesaplanır. Böylece $E_{LN}(T)$ beklenen değeri

$$E_{LN}(T) = -\frac{1}{2} \ln 2\pi - \ln \sigma + \ln \theta - 2(\ln \theta)^2 - \sigma^2$$

$$- \ln \tilde{\beta} - \tilde{\beta} \ln \tilde{\lambda} - (\tilde{\beta} - 1) \ln \theta + (\tilde{\lambda} \theta)^{\tilde{\beta}} e^{\tilde{\beta} \sigma^2 / 2}$$

$$= -\frac{1}{2} \ln 2\pi - \frac{1}{2} - \ln \tilde{\beta} - \tilde{\beta} \ln \tilde{\lambda} + (\tilde{\lambda})^{\tilde{\beta}} e^{\tilde{\beta} / 2}$$

olarak hesaplanır.

$V_{LN}(T)$ asimptotik varyansı ise (Kundu ve Manglick, 2004) tarafından

$$\frac{V_{LN}(T)}{n} \approx AV_{LN}(\sigma) = V_{LN} \left(\ln f_{LN}(x; 1, 1) - \ln f_{WE}(x; \tilde{\beta}, \tilde{\lambda}) \right)$$

$$= V \left(\tilde{\beta} \ln X + \frac{1}{2} (\ln X)^2 - (\tilde{\lambda} X)^{\tilde{\beta}} \right)$$

$$= \tilde{\beta}^2 + \frac{1}{2} + (\tilde{\lambda})^{2\tilde{\beta}} (e^{2\tilde{\beta}^2} - e^{\tilde{\beta}^2}) - 3\tilde{\beta}^2 (\tilde{\lambda})^{\tilde{\beta}} e^{(1/2)\tilde{\beta}^2}$$

$$= e - \frac{5}{2} = 0.2182818$$

biçiminde elde edilmiştir.

Tamamen benzer adımlar izlenerek, verilerin Weibull dağılımından geldiği durum da (4.1.8) eşitliği ile verilen T istatistiğinin asimptotik dağılımı $E_{WE}(T)$ beklenen değeri ve $V_{WE}(T)$ varyansı ile asimptotik normaldir. Burada

$$\begin{aligned}\frac{E_{WE}(T)}{n} &\approx AM_{WE} = E_{WE} \left[\ln \left(f_{LN}(x; \tilde{\sigma}, \tilde{\theta}) \right) - \ln \left(f_{WE}(x; 1, 1) \right) \right] \\ &= -\frac{1}{2} \ln(2\pi) - \ln \tilde{\sigma} - \frac{1}{2\tilde{\sigma}^2} \left[\frac{\pi^2}{6} + (\psi(1) - \ln \tilde{\theta})^2 \right] + 1 - \psi(1) \\ &= \frac{1}{2} - \frac{3}{2} \ln \pi + \frac{1}{2} \ln 3 - \psi(1) \\ &= -0.905730.\end{aligned}$$

ve

$$\begin{aligned}\frac{V_{WE}(T)}{n} &\approx AV_{WE}(\beta) = V_{WE} \left[\ln \left(f_{LN}(x; \tilde{\sigma}, \tilde{\theta}) \right) - \ln \left(f_{WE}(x; 1, 1) \right) \right] \\ &= V_{WE} \left[\ln X + \frac{(\ln X - \ln \tilde{\theta})^2}{2\tilde{\sigma}^2} - X \right] \\ &= \left(1 - \frac{\ln \tilde{\theta}}{\tilde{\sigma}^2} \right)^2 \left(\Gamma''(1) - (\Gamma'(1))^2 \right) + \frac{1}{4\tilde{\sigma}^4} \left(\Gamma^{(4)}(1) - (\Gamma''(1))^2 \right) \\ &+ 1 - \frac{1}{2\tilde{\sigma}^2} \Gamma'(1) + \frac{1}{\tilde{\sigma}^2} \left(1 - \frac{\ln \tilde{\theta}}{\tilde{\sigma}^2} \right) \left(\Gamma'''(1) - \Gamma'(1)\Gamma''(1) \right) - 2 \left(1 - \frac{\ln \tilde{\theta}}{\tilde{\sigma}^2} \right) \\ &= \left(1 - \frac{\psi(1)}{\psi'(1)} \right)^2 \left(\Gamma''(1) - (\Gamma'(1))^2 \right) + \frac{1}{4(\psi'(1))^4} \left(\Gamma^{(4)}(1) - (\Gamma''(1))^2 \right) \\ &+ 1 + \frac{1}{\psi'(1)} \left(1 - \frac{\psi(1)}{\psi'(1)} \right) \left(\Gamma'''(1) - \Gamma'(1)\Gamma''(1) \right) - 2 \left(1 - \frac{\psi(1)}{\psi'(1)} \right) - 2 \frac{\psi(1)}{\psi'(1)} \\ &= \psi'(1) \left(1 - \frac{\psi(1)}{\psi'(1)} \right)^2 + \frac{1}{4(\psi'(1))^2} \left(\Gamma^{(4)}(1) - (\Gamma''(1))^2 \right) - 1 \\ &+ \frac{1}{\psi'(1)} \left(1 - \frac{\psi(1)}{\psi'(1)} \right) \left(\Gamma'''(1) - \Gamma'(1)\Gamma''(1) \right) \\ &= 0.2834081.\end{aligned}$$

dir (Kundu ve Manglick, 2004)

4.2. Log-Normal ve Gamma Dağılımları Arasındaki Ayrım

X_1, X_2, \dots, X_n rasgele örnekleme $f_{LN}(x, \sigma, \theta)$ olasılık yoğunluk fonksiyonlu log-normal veya $f_{GA}(x, \alpha, \beta)$ olasılık yoğunluk fonksiyonlu gamma dağılımından alınmış bir rasgele örneklem olsun. X_1, X_2, \dots, X_n , rasgele değişkenlerinin sırasıyla log-normal ve gamma dağılımlarından alınmış olması göz önünde bulundurularak X_1, X_2, \dots, X_n rasgele değişkenlerine ait olabilirlik fonksiyonu sırasıyla

$$L_{LN}(x, \theta, \sigma) = \prod_{i=1}^n f_{LN}(x_i, \theta, \sigma) \quad (4.2.1)$$

ve

$$L_{GA}(x, \alpha, \lambda) = \prod_{i=1}^n f_{GA}(x_i, \alpha, \beta) \quad (4.2.2)$$

biçiminde gösterilsin

log-normal ve gamma dağılımları arasında seçim yapabilmek için RML yöntemine dayalı bir ayırıcı istatistik (3.6) eşitliğinin kullanılmasıyla

$$T = \ln \left(\frac{L_{LN}(x_i, \sigma, \theta)}{L_{GA}(x_i, \alpha, \beta)} \right) \quad (4.2.3)$$

biçiminde kolayca yazılabilir . Burada $L_{LN}(\sigma, \theta)$ Log- Normal dağılımın olabilirlik fonksiyonu nu ve $L_{GA}(\alpha, \beta)$ Gamma dağılımın olabilirlik fonksiyonunu göstermektedir. Log-Normal dağılımın parametrelerine ait en çok olabilirlik tahmin edicileri $\hat{\sigma}, \hat{\theta}$ ve Gamma dağılımının parametrelerinin en çok olabilirlik tahmin edicileri $\hat{\alpha}, \hat{\beta}$ biçiminde gösterilmek üzere, parametre tahmin ediciler kullanılarak T istatistiği

$$T = \ln L_{LN}(x_i, \hat{\sigma}, \hat{\theta}) - \ln L_{GA}(x_i, \hat{\alpha}, \hat{\beta}) \quad (4.2.4)$$

olarak yazılır. Sırasıyla (2.4.1) ve (2.3.1) eşitlikleri ile verilen log-normal dağılımın olasılık yoğunluk fonksiyonu ve gamma dağılımının olasılık yoğunluk fonksiyonu göz önüne alınmasıyla (4.2.4) eşitliği ile verilen T istatistiği açık olarak;

$$\ln L_{LN} = -\frac{n}{2} \ln(2\pi) - \sum_{i=1}^n \ln(x_i) - n \ln \hat{\sigma} - \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n \left(\ln \frac{x_i}{\hat{\theta}} \right)^2 \quad (4.2.5)$$

ve

$$\ln L_{GA} = -n\hat{\alpha} \ln \hat{\beta} - n \ln \Gamma(\hat{\alpha}) + (\hat{\alpha} - 1) \sum_{i=1}^n \ln x_i - \frac{1}{\hat{\beta}} \sum_{i=1}^n x_i \quad (4.2.6)$$

olmak üzere (4.2.5) ve (4.2.6) eşitliklerinin (4.1.4) eşitliğinde yerinde kullanılmasıyla

$$\begin{aligned} T &= -\frac{n}{2} \ln(2\pi) - \sum_{i=1}^n \ln(x_i) - n \ln \hat{\sigma} - \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n \left(\ln \frac{x_i}{\hat{\theta}} \right)^2 - \\ &\quad \left(-n(\hat{\alpha} \ln \hat{\beta} + \ln \Gamma(\hat{\alpha})) + (\hat{\alpha} - 1) \sum_{i=1}^n (\ln(x_i)) - \frac{1}{\hat{\beta}} \sum_{i=1}^n x_i \right) \\ &= -\frac{n}{2} \ln(2\pi) - \hat{\alpha} \sum_{i=1}^n \ln(x_i) - n \ln \hat{\sigma} - \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n \left(\ln \frac{x_i}{\hat{\theta}} \right)^2 + \\ &\quad n(\hat{\alpha} \ln \hat{\beta} + \ln \Gamma(\hat{\alpha})) + \frac{1}{\hat{\beta}} \sum_{i=1}^n x_i \end{aligned} \quad (4.2.7)$$

olarak elde edilir. $\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \Rightarrow n\bar{X} = \sum_{i=1}^n x_i$ ve $\sum_{i=1}^n \ln x_i = \ln \left(\prod_{i=1}^n x_i \right)$ olduğu göz önüne alınarak (4.2.7) ile verilen T istatistiği

$$\begin{aligned} T &= n \left(\left(\hat{\alpha} \ln \hat{\beta} + \ln \Gamma(\hat{\alpha}) \right) - \frac{1}{2} \ln(2\pi) - \ln \hat{\sigma} + \frac{\bar{X}}{\hat{\beta}} \right) - \\ &\quad \hat{\alpha} \sum_{i=1}^n \ln(x_i) - \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n \left(\ln \frac{x_i}{\hat{\theta}} \right)^2 \end{aligned} \quad (4.2.8)$$

olarak yazılır. Ayrıca $\left(\prod_{i=1}^n x_i \right)^{1/n} = \tilde{X}$ ve $\alpha \ln \prod_{i=1}^n x_i = \alpha n \ln(\tilde{X})$ gösterimlerinin kullanılmasıyla

$$T = n \left(\hat{\alpha} \ln \hat{\beta} + \ln \Gamma(\hat{\alpha}) - \frac{1}{2} \ln(2\pi) - \ln \hat{\sigma} + \frac{\bar{X}}{\hat{\beta}} \right) - \hat{\alpha} \prod_{i=1}^n \ln(x_i) - \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n \left(\ln \frac{x_i}{\hat{\theta}} \right)^2 \quad (4.2.9)$$

yazılabilir. (4.2.9) eşitliğinin düzenlenmesiyle

$$T = n \left(\ln \frac{\Gamma(\hat{\alpha})}{\hat{\sigma}} - \hat{\alpha} \ln \frac{\tilde{X}}{\hat{\beta}} + \frac{\bar{X}}{\hat{\beta}} - \frac{1}{2n\hat{\sigma}^2} \sum_{i=1}^n \left(\ln \frac{x_i}{\theta} \right)^2 - \frac{1}{2} \ln(2\pi) \right) \quad (4.2.10)$$

elde edilir. Unutulmamalıdır ki, log-normal dağılım için

$$\hat{\theta} = \tilde{X} \quad \text{ve} \quad \hat{\sigma} = \left(\frac{1}{n} \sum_{i=1}^n \left(\ln \frac{x_i}{\hat{\theta}} \right)^2 \right)^{1/2} \quad \text{ve} \quad \text{gamma dağılımı için} \quad \hat{\alpha} = \frac{\bar{X}}{\hat{\beta}} \quad \text{ilişkileri}$$

geçerlidir. (4.2.10) eşitliği ile elde edilen T istatistiğinin değeri 0'dan büyükse veri seti için log-normal dağılımın daha uygun olduğu söylenebilir aksi halde gamma dağılımının veri setinin modellenmesinde daha uygun olduğu kararı verilir.

Kundu ve Manglick (2005) tarafından (4.2.10) ile verilen T istatistiğinin asimptotik dağılımı Çizelge 4'de verildiği şekilde elde edilmiştir.

Çizelge 4. 1: Log-normal ve gamma dağılımlarının ayrımı için RML yöntemine dayalı ayırıcı istatistik için asimptotik sonuçlar.

Veriler log-normal dağılımdan geldiğinde
<p>$E_{LN}(T)$ asimptotik beklenen değer:</p> $\frac{E_{LN}(T)}{n} = AM_{LN}(\sigma) = \left(-\ln 2\pi - \ln \sigma - \frac{1}{2} \right) - \left(-\tilde{\alpha} \ln \tilde{\beta} - \ln \Gamma(\tilde{\alpha}) - \frac{1}{\tilde{\beta}} e^{\frac{1}{2}\sigma^2} \right)$ $= -\left(\ln 2\pi + \ln \sigma + \frac{1}{2} \right) + \tilde{\alpha} \ln \tilde{\beta} + \ln \Gamma(\tilde{\alpha}) + \frac{1}{\tilde{\beta}} e^{\frac{1}{2}\sigma^2}$ <p>$V_{LN}(T)$ asimptotik Varyans:</p> $\frac{V_{LN}(T)}{n} = AV_{LN}(\sigma) = V_{GA} \left[\ln \left(f_{LN}(x; \sigma^2, 1) \right) - \ln \left(f_{GA}(x; \tilde{\alpha}, \tilde{\beta}) \right) \right]$ $= V_{LN} \left[-\tilde{\alpha} \ln X + \frac{X}{\tilde{\beta}} - \frac{1}{2\tilde{\sigma}^2} (\ln X)^2 \right]$ $= \tilde{\alpha} + \sigma^2 + \frac{1}{\tilde{\beta}^2} e^{\tilde{\sigma}^2} (e^{\tilde{\sigma}^2} - 1) + \frac{1}{2} - \frac{\tilde{\alpha}}{\tilde{\beta}} Cov(\ln X, X) - \frac{X}{\tilde{\beta}\tilde{\sigma}^2} Cov((\ln X)^2, X)$
Veriler Gamma dağılımdan geldiğinde
<p>$E_{GA}(T)$ asimptotik beklenen değer:</p> $\frac{E_{GA}(T)}{n} = AM_{GA}(\alpha) = E_{GA} \left[\ln \left(f_{LN}(X; \tilde{\sigma}, \tilde{\theta}) \right) - \ln \left(f_{GA}(X; \alpha, 1) \right) \right]$ $= -\frac{1}{2} \ln(2\pi) - \ln \tilde{\sigma} - \frac{1}{2\tilde{\sigma}^2} \left(\psi'(\alpha) + (\psi(\alpha) - \ln \tilde{\theta})^2 \right)$ $+ \ln(\Gamma(\alpha)) + \alpha(1 - \psi(\alpha))$ <p>$V_{GA}(T)$ asimptotik varyans:</p> $\frac{V_{GA}(T)}{n} = AV_{GA}(\alpha) = V_{GA} \left[\ln \left(f_{LN}(X; \tilde{\sigma}, \tilde{\theta}) \right) - \ln \left(f_{GA}(X; \alpha, 1) \right) \right]$ $= \alpha + \alpha^2 \psi'(\alpha) - 2\alpha(\psi(\alpha+1) + \psi(\alpha)) -$ $\frac{1}{\tilde{\sigma}^2} \left[\alpha(\alpha+1)(\psi'(\alpha+2) - \psi(\alpha+2))^2 - (\ln \tilde{\theta}) \alpha \psi'(\alpha+1) + \alpha(\ln \tilde{\theta})^2 - \alpha(\psi'(\alpha) - \psi(\alpha))^2 \right]$ $- 2(\ln \tilde{\theta}) \psi(\alpha) - 2\alpha \psi'(\alpha) \psi(\alpha) - \tilde{\alpha} \psi''(\alpha) + 2(\ln \theta) \psi'(\alpha)$ $+ \frac{1}{4\tilde{\sigma}^4} \left[\psi'''(\alpha) + 4\psi(\alpha) \psi''(\alpha) + 4\psi'(\alpha)(\psi(\alpha))^2 + 2(\psi'(\alpha))^2 \right]$ $- 4(\ln \tilde{\theta}) \psi''(\alpha) - 8\psi(\alpha) \psi'(\alpha) \ln \tilde{\theta} + 4\psi'(\alpha)(\ln \tilde{\theta})^2$

4.3. Genelleştirilmiş Üstel Dağılım ve Power Lindley Dağılımları Arasındaki Ayrım

Bu kısımda genelleştirilmiş üstel dağılım ve Power Lindley dağılımlarıyla uyumlu bir veri seti için ilgili iki dağılımdan hangisinin optimal olduğunu belirleyebilmek için RML yöntemine dayalı ayırıcı bir istatistik ve bu ayırıcı istatistiğin asimptotik dağılımı elde edilecektir.

Varsayalım ki; X_1, X_2, \dots, X_n rasgele örnekleme $f_{GE}(x, \alpha, \lambda)$ olasılık yoğunluk fonksiyonuna sahip genelleştirilmiş üstel veya $f_{PL}(x, \theta, \beta)$ olasılık yoğunluk fonksiyonuna sahip Power Lindley dağılımlarının herhangi birinden alınmış bir rasgele örneklem olsun; X_1, X_2, \dots, X_n , rasgele değişkenlerinin sırasıyla genelleştirilmiş üstel ve Power Lindley dağılımlarından alınmış olduğunda X_1, X_2, \dots, X_n rasgele değişkenlerine ait logaritmik olabilirlik fonksiyonu sırasıyla

$$\ln(L_{GE}(x, \alpha, \lambda)) = n \ln(\alpha) + n \ln(\lambda) + (\alpha - 1) \sum_{i=1}^n \ln(1 - e^{-\lambda x_i}) - \lambda \sum_{i=1}^n x_i \quad (4.3.1)$$

ve

$$\begin{aligned} \ln(L_{PL}(x, \theta, \beta)) = & n \ln(\theta) + 2n \ln(\beta) - n \ln(\beta + 1) + \\ & \sum_{i=1}^n \ln(1 + x_i^\theta) + (\theta - 1) \sum_{i=1}^n \ln(x_i) - \beta \sum_{i=1}^n (x_i^\theta) \end{aligned} \quad (4.3.2)$$

dır. (4.3.1) ve (4.3.2) eşitlikleri yardımıyla ve (3.5) ile verilen tanımlama kullanılarak ilgili iki dağılım arasında ayırım yapabilecek bir istatistik

$$\begin{aligned} T = \ln \left(\frac{L_{GE}(x, \alpha, \lambda)}{L_{PL}(x, \theta, \beta)} \right) &= \ln(L_{GE}(x, \alpha, \lambda)) - \ln(L_{PL}(x, \theta, \beta)) \\ &= n \ln(\alpha) + n \ln(\lambda) + (\alpha - 1) \sum_{i=1}^n \ln(1 - e^{-\lambda x_i}) - \lambda \sum_{i=1}^n x_i \\ &= - \left\{ n \ln(\theta) + 2n \ln(\beta) - n \ln(\beta + 1) + \right. \\ & \quad \left. \sum_{i=1}^n \ln(1 + x_i^\theta) + (\theta - 1) \sum_{i=1}^n \ln(x_i) - \beta \sum_{i=1}^n (x_i^\theta) \right\} \end{aligned} \quad (4.3.3)$$

olarak kolayca yazılabilir. (4.3.3) eşitliğinde gerekli sadeleştirmeler yapılarak

$$T = \left[n \left[\ln \left(\frac{\alpha \lambda (\beta + 1)}{2\theta \beta^2} \right) \right] \right] + \sum_{i=1}^n \left[\ln \left(\frac{(1 - e^{\lambda x_i})^{\alpha-1}}{x_i^{\theta-1} + x_i^{2\theta-1}} \right) - \beta x_i^\theta \right] \quad (4.3.4)$$

elde edilir. α , λ , θ ve β parametrelerinin sırasıyla $\hat{\alpha}$, $\hat{\lambda}$, $\hat{\theta}$ ve $\hat{\beta}$ en çok olabilirlik tahmin edicileri (4.3.4) eşitliğinde yerlerinde kullanılmasıyla

$$T = \left[n \left[\ln \left(\frac{\hat{\alpha} \hat{\lambda} (\hat{\beta} + 1)}{2\hat{\theta} \hat{\beta}^2} \right) \right] \right] + \sum_{i=1}^n \left[\ln \left(\frac{(1 - e^{\hat{\lambda} x_i})^{\hat{\alpha}-1}}{x_i^{\hat{\theta}-1} + x_i^{2\hat{\theta}-1}} \right) - \hat{\beta} x_i^{\hat{\theta}} \right] \quad (4.3.5)$$

biçiminde yazılır. Burada

$$\sum_{i=1}^n (x_i^{\hat{\theta}}) = \frac{2n}{\hat{\beta}} - \frac{n}{\hat{\beta} + 1} \quad (4.3.6)$$

dır (Ghithany vd, 2013). (4.3.6) eşitliğinin (4.3.5) eşitliğinde yerinde kullanılmasıyla

$$\begin{aligned} T &= \left[n \left[\ln \left(\frac{\hat{\alpha} \hat{\lambda} (\hat{\beta} + 1)}{2\hat{\theta} \hat{\beta}^2} \right) \right] \right] - \frac{(\hat{\alpha} - 1)n}{\hat{\alpha}} + \sum_{i=1}^n \left[\ln \left(\frac{1}{(1 + x_i^{\hat{\theta}}) x_i^{\hat{\theta}-1}} \right) - \hat{\beta} x_i^{\hat{\theta}} \right] \\ &= \left[n \left[\ln \left(\frac{\hat{\alpha} \hat{\lambda} (\hat{\beta} + 1)}{2\hat{\theta} \hat{\beta}^2} \right) \right] \right] - \left(n - \frac{n}{\hat{\alpha}} \right) + \sum_{i=1}^n \left[\ln \left(\frac{1}{(1 + x_i^{\hat{\theta}}) x_i^{\hat{\theta}-1}} \right) - \hat{\beta} \left(\frac{2n}{\hat{\beta}} - \frac{n}{\hat{\beta} + 1} \right) \right] \end{aligned} \quad (4.3.7)$$

elde edilir. (4.3.7) eşitliği yeniden düzenlenirse

$$\begin{aligned} T &= \left[n \left[\ln \left(\frac{\hat{\alpha} \hat{\lambda} (\hat{\beta} + 1)}{2\hat{\theta} \hat{\beta}^2} \right) \right] \right] - \left(n - \frac{n}{\hat{\alpha}} \right) + \sum_{i=1}^n \left[\ln \left(\frac{1}{(1 + x_i^{\hat{\theta}}) x_i^{\hat{\theta}-1}} \right) \right] - \sum_{i=1}^n \left[\frac{n(\hat{\beta} + 2)}{\hat{\beta} + 1} \right] \\ T &= \left[n \left[\ln \left(\frac{\hat{\alpha} \hat{\lambda} (\hat{\beta} + 1)}{2\hat{\theta} \hat{\beta}^2} \right) - \frac{n(\hat{\beta} + 2)}{\hat{\beta} + 1} \right] + \frac{1}{\hat{\alpha}} - 1 \right] + \sum_{i=1}^n \left[\ln \left(\frac{1}{(1 + x_i^{\hat{\theta}}) x_i^{\hat{\theta}-1}} \right) \right] \end{aligned} \quad (4.3.8)$$

elde edilir. Böylece (4.3.8) ile elde edilen T istatistiğinin değeri 0'dan büyükse veri setini genelleştirilmiş üstel dağılımla modellenmelidir aksi halde Power Lindley dağılımının veri setini modellemede daha uygun olduğu söylenir.

(4.3.8) eşitliği ile elde edilen T istatistiğinin asimptotik dağılımı üçüncü bölümde verilen iki durum göz önüne alınarak; verilerin genelleştirilmiş üstel dağılımdan veya verilerin Power Lindley dağılımından geldiği varsayımları ile elde edilebilir.

DURUM 4.3.1 Verilerin genelleştirilmiş üstel dağılımdan geldiğini ve üçüncü bölümde verilen Durum 1'i göz önünde bulundurarak eşitlik (4.3.8) ile verilen T istatistiğinin asimptotik dağılımı $E_{GE}(T)$ beklenen değeri ve $V_{GE}(T)$ varyansı ile Asimptotik Normal dir.

$E_{GE}(T)$ beklenen değeri, Lemma 3.1 in göz önüne alınmasıyla

$$\begin{aligned} \frac{E_{GE}(T)}{n} &\approx AM_{GE} = E_{GE} \left[\ln(f_{GE}(x; \alpha, \lambda)) - \ln(f_{PL}(x; \tilde{\theta}, \tilde{\beta})) \right] \\ &= E_{GE} \left[\ln(f_{GE}(x; \alpha, \lambda)) \right] - E_{GE} \left[\ln(f_{PL}(x; \tilde{\theta}, \tilde{\beta})) \right] \end{aligned} \quad (4.3.9)$$

biçiminde yazılabilir. $E_{GE}(T)$ asimptotik beklenen değerinin açık bir ifadesini elde edebilmek için ilk olarak (4.3.9) eşitliğinde mevcut $E_{GE} \left[\ln(f_{GE}(x; \alpha, \lambda)) \right]$ terimini hesaplayalım. Beklenen değer tanımından

$$\begin{aligned} E_{GE} \left(\ln(f_{GE}(x; \alpha, \lambda)) \right) &= \int_0^{\infty} \ln(f_{GE}(x)) f_{GE}(x) dx \\ &= \int_0^{\infty} \left(\ln \alpha + \ln \lambda + (\alpha - 1) \ln(1 - e^{-\lambda x}) - \lambda x \right) f_{GE}(x) dx \\ &= \int_0^{\infty} (\ln \alpha) f_{GE}(x) dx + \int_0^{\infty} (\ln \lambda) f_{GE}(x) dx + \\ &\quad (\alpha - 1) \int_0^{\infty} \ln(1 - e^{-\lambda x}) f_{GE}(x) dx - \lambda \int_0^{\infty} x f_{GE}(x) dx \quad (4.3.10) \\ &= \ln \alpha + \ln \lambda - \lambda E_{GE}(x) + (\alpha - 1) E_{GE} \left(\ln(1 - e^{-\lambda x}) \right) \end{aligned}$$

olarak elde edilir. Burada $E_{GE}(x)$ genelleştirilmiş üstel dağılımın beklenen değeri olup

$$E_{GE}(x) = \left(\frac{1}{\lambda} (\psi(\alpha+1) - \psi(1)) \right) \quad (4.3.11)$$

dir (Gupta ve Kundu, 2001) Ayrıca (4.3.10) denklemindeki $E_{GE}(\ln(1-e^{-\lambda x}))$ terimi,

$$E_{GE}(\ln(1-e^{-\lambda x})) = \int_0^{\infty} \ln(1-e^{-\lambda x}) \alpha \lambda (1-e^{-\lambda x})^{\alpha-1} e^{-\lambda x} dx \quad (4.3.12)$$

olup, $u = 1 - e^{-\lambda x}$ dönüşümü yardımıyla

$$E_{GE}(\ln(1-e^{-\lambda x})) = \alpha \underbrace{\int_0^1 \ln u (u)^{\alpha-1} du}_{-\frac{1}{\alpha^2}} = -\frac{1}{\alpha} \quad (4.3.13)$$

olarak kolayca hesaplanır.

Şimdi de (4.3.9) beklenen değerini hesaplayabilmek için $E_{GE}[\ln(f_{PL}(x; \tilde{\theta}, \tilde{\beta}))]$

beklenen değerini hesaplayalım. Power Lindley dağılımının olasılık yoğunluk

fonksiyonu göz önünde bulundurularak $E_{GE}[\ln(f_{PL}(x; \tilde{\theta}, \tilde{\beta}))]$ terimi

$$E_{GE}(\ln(f_{PL}(x; \tilde{\theta}, \tilde{\beta}))) = \underbrace{\int_0^{\infty} \ln(\tilde{\theta}) f_{GE} dx}_A + 2 \underbrace{\int_0^{\infty} \ln(\tilde{\beta}) f_{GE} dx}_B - \underbrace{\int_0^{\infty} \ln(\tilde{\beta} + 1) f_{GE} dx}_C + \underbrace{\int_0^{\infty} \ln(1 + x^{\tilde{\theta}}) f_{GE} dx}_D - \underbrace{\tilde{\beta} \int_0^{\infty} x^{\tilde{\theta}} f_{GE} dx}_E \quad (4.3.14)$$

biçiminde yazılabilir. (4.3.14) beklenen değerinde $A = \ln(\tilde{\theta})$, $B = 2 \ln(\tilde{\beta})$ ve

$C = -\ln(\tilde{\beta} + 1)$ oldukları aşikardır. (4.3.14) eşitliğindeki D terimi ise genelleştirilmiş

üstel dağılımın olasılık yoğunluk fonksiyonu göz önüne alınarak

$$\begin{aligned}
D &= \int_0^{\infty} \ln(1+x^{\tilde{\theta}}) f_{GE} dx \\
&= \int_0^{\infty} \ln(1+x^{\tilde{\theta}}) \alpha \lambda (1-e^{-\lambda x})^{\alpha-1} e^{-\lambda x} dx \\
&= D_x
\end{aligned}$$

(4.3.15)

biçiminde hesaplanır. (4.3.15) ile verilen integral analitik olarak alınamamakla birlikte sayısal olarak kolayca hesaplanabilir. Son olarak (4.3.14) eşitliğindeki E terimi genelleştirilmiş üstel dağılımın θ . momentini olup

$$E = -\tilde{\beta} \int_0^{\infty} x^{\tilde{\theta}} f_{GE} dx = -\tilde{\beta} \mu_{\tilde{\theta}}$$

(4.3.16)

dır, burada $\mu_{\theta_k} = \alpha k! \sum_{j=0}^{\infty} (-1)^j \binom{\alpha-1}{j} \frac{1}{(j+1)^{k+1}}$ dir (Gupta ve Kundu, 2001).

Elde edilen A , B , C , D ve E terimleri göz önünde bulundurularak (4.3.14) ile verilen beklenen değer

$$E_{GE} \left[\ln \left(f_{PL} (x; \tilde{\theta}, \tilde{\beta}) \right) \right] = \ln(\tilde{\theta}) + 2 \ln(\tilde{\beta}) - \ln(\tilde{\beta} + 1) + D_x - \beta \mu_{\tilde{\theta}} \quad (4.3.17)$$

biçiminde elde edilir.

Böylece (4.3.9) ile hesaplanmak istenen $\frac{E_{GE}(T)}{n}$ asimptotik beklenen değeri

(4.3.10), (4.3.11), (4.3.13) ve (4.3.17) eşitliklerinin göz önüne alınmasıyla

$$\begin{aligned}
\frac{E_{GE}(T)}{n} \approx AM_{GE} &= \left[\ln \alpha + \ln \lambda - \lambda \left(\frac{1}{\lambda} (\psi(\alpha+1) - \psi(1)) \right) - \frac{(\alpha-1)}{\alpha} \right] - \\
&\quad \left[\ln(\tilde{\theta}) + 2 \ln(\tilde{\beta}) - \ln(\tilde{\beta} + 1) + D_x - \tilde{\beta} \sum_{i=1}^n X_i Y_i \mu_{\tilde{\theta}} \right]
\end{aligned}$$

(4.3.18)

olarak elde edilir.

Durum 4.3.1 için asimptotik varyans $V_{GE}(T)$ ise

$$\begin{aligned} \frac{V_{GE}(T)}{n} &\approx AV_{GE}(T) = V_{GE} \left[\ln(f_{GE}(x; \alpha, \lambda)) - \ln(f_{PL}(x; \tilde{\theta}, \tilde{\beta})) \right] \\ &= \int_0^{\infty} \left\{ \left[\ln(f_{GE}(x; \alpha, \lambda)) - \ln(f_{PL}(x; \tilde{\theta}, \tilde{\beta})) \right] - E \left[\ln(f_{GE}(x; \alpha, \lambda)) - \ln(f_{PL}(x; \tilde{\theta}, \tilde{\beta})) \right] \right\}^2 f_{GE}(x; \alpha, \lambda) dx \end{aligned}$$

(4.3.19)

biçiminde hesaplanır. Ancak (4.3.19) ile verilen integral analitik olarak elde edilememektedir ve sayısal olarak hesaplanmalıdır.

DURUM 4.3.2: Varsayalım X_1, X_2, \dots, X_n gözlemleri θ, β parametreleriyle Power Lindley dağılımından alınmış olsun. Bu durumda (4.3.8) ile verilen T istatistiğinin asimptotik dağılımı $E_{PL}(T)$ beklenen değeri ve $V_{PL}(T)$ varyansı ile Asimptotik Normal dir.

İspat: (4.3.8) ile verilen T istatistiğinin asimptotik beklenen değeri 3. bölümde verilen Lemma 2'nin göz önüne alınmasıyla

$$\begin{aligned} \frac{E_{PL}(T)}{n} &\approx AM_{PL} = E_{PL} \left[\ln(f_{GE}(x; \tilde{\alpha}, \tilde{\lambda})) - \ln(f_{PL}(x; \theta, \beta)) \right] \\ &= E_{PL} \left[\ln(f_{GE}(x; \tilde{\alpha}, \tilde{\lambda})) \right] - E_{PL} \left[\ln(f_{PL}(x; \theta, \beta)) \right] \end{aligned}$$

(4.3.20)

yazılabilir. $E_{PL}(T)$ asimptotik beklenen değeri için açık bir ifade elde edebilmek için (4.3.20) eşitliğinde verilen $E_{PL} \left[\ln(f_{GE}(x; \alpha, \lambda)) \right]$ ve $E_{PL} \left[\ln(f_{PL}(x; \theta, \beta)) \right]$ beklenen değerleri hesaplanmalıdır. Sırasıyla (GE pdf) ve (PL pdf) ile verilen genelleştirilmiş üstel dağılımın ve Power Lindley dağılımının olasılık yoğunluk fonksiyonları göz önüne alınırsa

$$\begin{aligned}
E_{PL} \left(\ln \left(f_{GE} (x; \tilde{\alpha}, \tilde{\lambda}) \right) \right) &= \int_0^{\infty} \ln \left(f_{GE} (x; \tilde{\alpha}, \tilde{\lambda}) \right) f_{PL} (x; \theta, \beta) dx \\
&= \int_0^{\infty} \ln \tilde{\lambda} f_{PL} (x; \theta, \beta) dx + \int_0^{\infty} \ln \tilde{\alpha} f_{PL} (x; \theta, \beta) dx + \\
&\quad \int_0^{\infty} \left((\tilde{\alpha} - 1) \ln (1 - e^{-\tilde{\lambda}x}) \right) f_{PL} (x; \theta, \beta) dx - \int_0^{\infty} \tilde{\lambda} x f_{PL} (x; \theta, \beta) dx \\
&= \ln \tilde{\lambda} + \ln \tilde{\alpha} + (\tilde{\alpha} - 1) \zeta_{PL} (\tilde{\lambda}, x) - \tilde{\lambda} \frac{\Gamma \left(\frac{1}{\theta} \right) (\theta (\beta + 1) + \theta)}{\theta^2 \beta^{1/\theta} (\beta + 1)}
\end{aligned}$$

(4.3.21)

biçiminde elde edilir. Burada $\zeta_{PL} (\tilde{\lambda}, x) = \int_0^{\infty} \ln (1 - e^{-\tilde{\lambda}x}) \frac{\theta \beta^2}{\beta + 1} (1 + x^\theta) x^{\theta-1} e^{-\beta x^\theta} dx$ dir.

Benzer şekilde $E_{PL} \left[\ln \left(f_{PL} (x; \theta, \beta) \right) \right]$ beklenen değeri

$$\begin{aligned}
E_{PL} \left(\ln \left(f_{PL} (x; \theta, \beta) \right) \right) &= \int_0^{\infty} \ln (\theta) f_{PL} (x; \theta, \beta) dx + \\
&\quad 2 \int_0^{\infty} \ln (\beta) f_{PL} (x; \theta, \beta) dx - \int_0^{\infty} \ln (\beta + 1) f_{PL} (x; \theta, \beta) dx + \\
&\quad \int_0^{\infty} \ln (1 + x^\theta) f_{PL} (x; \theta, \beta) dx - \beta \int_0^{\infty} x^\theta f_{PL} (x; \theta, \beta) dx \\
&= \ln \theta + 2 \ln \beta - \ln (\beta + 1) + \frac{(-e^\beta E_i (-\beta) - 1)}{(\beta + 1)} + \frac{\beta + 2}{(\beta + 1)}
\end{aligned} \tag{4.3.22}$$

olarak kolayca hesaplanır. Burada $E_i (-\beta) = \int_{-\beta}^{\infty} \frac{e^{-x}}{x} dx$ biçiminde tanımlanmış üstel

integrali (Abramowitz ve Stegun, 1964) ifade etmektedir .

Böylece (4.3.20) eşitliğinde (4.3.21) ve (4.3.22)'ün yerlerinde kullanılmasıyla

$$\frac{E_{PL}(T)}{n} \approx AM_{PL} = \left[\ln \tilde{\lambda} + \ln \tilde{\alpha} + (\tilde{\alpha} - 1) \zeta_{PL}(\tilde{\lambda}, x) - \tilde{\lambda} \frac{\Gamma\left(\frac{1}{\theta}\right)(\theta(\beta+1) + \theta)}{\theta^2 \beta^{1/\theta} (\beta+1)} \right] - \left[\ln \theta + 2 \ln \beta - \ln(\beta+1) + \frac{(-e^\beta E_i(-\beta) - 1)}{(\beta+1)} + \frac{\beta+2}{(\beta+1)} \right]$$

(4.3.23)

olarak elde edilir.

Durum 4.3.2 için asimptotik varyans $V_{PL}(T)$,

$$\frac{V_{PL}(T)}{n} \approx AV_{PL}(T) = V_{PL} \left[\ln(f_{GE}(x; \tilde{\alpha}, \tilde{\lambda})) - \ln(f_{PL}(x; \theta, \beta)) \right] = \int_0^\infty \left\{ \left[\ln(f_{GE}(x; \tilde{\alpha}, \tilde{\lambda})) - \ln(f_{PL}(x; \theta, \beta)) \right] - E \left[\ln(f_{GE}(x; \tilde{\alpha}, \tilde{\lambda})) - \ln(f_{PL}(x; \theta, \beta)) \right] \right\}^2 f_{PL}(x; \hat{\theta}, \hat{\beta}) dx$$

(4.3.24)

integralinin sayısal çözümünden elde edilir.

5. SİMULASYON ÇALIŞMASI

Çalışmanın bu bölümünde, geliştirilmiş üstel dağılım ve Power Lindley dağılımları arasında ayırım yapmak için RML yöntemine dayalı olarak elde edilen ayırıcı istatistiğin doğru ayırım performansını değerlendirebilmek ve yine kısım 4.3 de elde edilen asimptotik sonuçların nasıl çalıştığını ortaya koymak için yapılan simülasyon çalışmalarına yer verilmektedir. Simülasyon çalışmasında iki farklı senaryo üzerinde durulmuştur. İlk senaryoda, verilerin geliştirilmiş üstel dağılımlı olduğu varsayılarak, geliştirilmiş üstel dağılımın α ve λ parametreleri sırasıyla $\alpha = 0.5, 1, 2$ ve $\lambda = 0.25, 0.5, 1, 1.5, 2, 4, 10$ olarak seçilmiştir. Ayrıca, yöntemin farklı örneklem büyüklüklerindeki performansını ve büyük örneklerdeki davranışını ortaya koyabilmek için örnek sayısı n , 30, 50, 80, 100 ve 200 olarak alınmıştır. İlk senaryo için 1000 tekrar ile elde edilen simülasyona dayalı doğru seçim olasılıkları (SDSO) ve asimptotik doğru seçim olasılıkları (ADSO) Çizelge 5.1-5.3 ile çizelgelenmiştir.

Çizelge 5. 1: Veriler geliştirilmiş üstel () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları

λ	Yöntem	n				
		30	50	80	100	200
0.25	SDSO	0.7000	0.7240	0.7680	0.8200	0.8440
	ADSO	0.6676	0.6829	0.7457	0.7760	0.8401
0.50	SDSO	0.6880	0.7080	0.7440	0.7280	0.8600
	ADSO	0.6678	0.7018	0.7428	0.7195	0.8429
1.00	SDSO	0.6960	0.7160	0.7880	0.7640	0.8360
	ADSO	0.6664	0.7202	0.7655	0.7757	0.8255
1.50	SDSO	0.6640	0.7080	0.7520	0.7800	0.8880
	ADSO	0.6476	0.7092	0.7581	0.7487	0.8660
2.00	SDSO	0.7080	0.7320	0.7840	0.8000	0.8760
	ADSO	0.6786	0.7230	0.7637	0.7913	0.8741
4.00	SDSO	0.6960	0.7760	0.7760	0.7640	0.8600
	ADSO	0.6711	0.7346	0.7665	0.7761	0.8585
10.00	SDSO	0.7040	0.7360	0.8120	0.7720	0.8720

ADSO	0.6927	0.7399	0.7843	0.7773	0.8545
------	--------	--------	--------	--------	--------

Çizelge 5. 2: Veriler genelleştirilmiş üstel () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları.

λ	Yöntem	n				
		30	50	80	100	200
0.25	SDSO	0.6640	0.6960	0.7120	0.8040	0.8640
	ADSO	0.6572	0.6876	0.7121	0.7630	0.8497
0.50	SDSO	0.6760	0.7400	0.7760	0.8440	0.8560
	ADSO	0.6453	0.7056	0.7472	0.7866	0.8632
1.00	SDSO	0.6920	0.6920	0.7960	0.8120	0.8720
	ADSO	0.6272	0.6915	0.7916	0.8167	0.8842
1.50	SDSO	0.6720	0.6480	0.8160	0.8160	0.9200
	ADSO	0.6660	0.6462	0.8143	0.8130	0.9073
2.00	SDSO	0.5880	0.7120	0.7760	0.8040	0.8720
	ADSO	0.5932	0.7021	0.7643	0.7827	0.8541
4.00	SDSO	0.5720	0.6960	0.7040	0.7560	0.8320
	ADSO	0.5648	0.6819	0.7024	0.7387	0.8250
10.00	SDSO	0.5720	0.6520	0.7040	0.7360	0.8600
	ADSO	0.5468	0.6274	0.7124	0.7320	0.8529

Çizelge 5. 3: Veriler genelleştirilmiş üstel () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları

λ	Yöntem	n				
		30	50	80	100	200
0.25	SDSO	0.7000	0.7720	0.8160	0.8440	0.8840
	ADSO	0.6812	0.7256	0.8236	0.8204	0.8966
0.50	SDSO	0.7000	0.7840	0.8040	0.8720	0.9440
	ADSO	0.6718	0.7959	0.8028	0.8737	0.9485
1.00	SDSO	0.7120	0.8000	0.8760	0.8920	0.9560
	ADSO	0.6848	0.7710	0.8837	0.8813	0.9523
1.50	SDSO	0.7200	0.8120	0.8400	0.9280	0.9560
	ADSO	0.6956	0.7947	0.8408	0.9246	0.9504
2.00	SDSO	0.6880	0.7760	0.8200	0.8960	0.9360
	ADSO	0.6615	0.7834	0.8277	0.8873	0.9448
4.00	SDSO	0.6760	0.7680	0.8040	0.8280	0.9120
	ADSO	0.6718	0.7510	0.7950	0.8162	0.9200
10.00	SDSO	0.6280	0.7040	0.7960	0.8760	0.9240
	ADSO	0.6287	0.7228	0.8044	0.8634	0.9228

Çizelge 5.1-5.3 'den açık olarak görülebileceği gibi, tüm durumlarda simülasyona dayalı doğru seçim olasılıkları SDSO ve asimptotik sonuçlara dayalı doğru seçim olasılıkları ASDO birbirlerini oldukça yakından takip etmektedir. Bunun yanı sıra, yine tüm durumlarda örneklem çapı n arttıkça doğru seçim olasılıklarının arttığı görülmektedir.

İkinci senaryo olarak, verilerin Power Lindley dağılımından alındığı varsayılmıştır. İlk senaryodakine benzer olarak Power Lindley dağılımın θ ve β parametrelerinin değerleri sırasıyla $\theta = 0.5, 1, 2$ ve $\beta = 0.25, 0.5, 1, 1.5, 2, 4, 10$ ve $n, 30, 50, 80, 100, 200$ olarak alınmıştır. 1000 tekrarlı simülasyon çalışmaları ile elde edilen SDSO ve ADSO değerleri Çizelge 5.4-5.6 ile çizelgelenmiştir.

Çizelge 5. 4: Veriler Power Lindley () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları

β	Yöntem	n				
		30	50	80	100	200
0.25	SDSO	0.5080	0.6320	0.5840	0.6280	0.7360
	ADSO	0.5231	0.6393	0.6291	0.6550	0.7570
0.50	SDSO	0.5560	0.5880	0.6040	0.6560	0.7280
	ADSO	0.5908	0.6167	0.6534	0.6763	0.7465
1.00	SDSO	0.5560	0.6400	0.5800	0.6560	0.7240
	ADSO	0.5575	0.5869	0.5794	0.6609	0.6770
1.50	SDSO	0.5600	0.5240	0.5800	0.6360	0.6720
	ADSO	0.5601	0.5393	0.6115	0.6369	0.6495
2.00	SDSO	0.4720	0.5560	0.5480	0.5240	0.6120
	ADSO	0.4978	0.5662	0.5549	0.5782	0.6389
4.00	SDSO	0.5360	0.5240	0.5280	0.5480	0.5800
	ADSO	0.5108	0.5041	0.5348	0.5266	0.5784
10.00	SDSO	0.4880	0.4880	0.5280	0.4800	0.6160
	ADSO	0.4825	0.4435	0.4968	0.5066	0.5436

Çizelge 5. 5: Veriler Power Lindley () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları

β	Yöntem	n				
		30	50	80	100	200
0.25	SDSO	0.7960	0.8560	0.9320	0.9360	0.9920
	ADSO	0.7868	0.8404	0.9086	0.9257	0.9853
0.50	SDSO	0.7840	0.8600	0.9440	0.9280	0.9760
	ADSO	0.8075	0.8494	0.9193	0.9358	0.9815
1.00	SDSO	0.7640	0.8840	0.9120	0.9200	0.9920
	ADSO	0.7221	0.8876	0.9106	0.9287	0.9911
1.50	SDSO	0.8280	0.8360	0.9520	0.9360	0.9760
	ADSO	0.7883	0.8408	0.9437	0.9378	0.9797
2.00	SDSO	0.7920	0.8320	0.8960	0.9160	0.9840
	ADSO	0.7997	0.8477	0.9010	0.9118	0.9777
4.00	SDSO	0.7560	0.8200	0.8880	0.9320	0.9720
	ADSO	0.7130	0.8172	0.8773	0.9288	0.9718
10.00	SDSO	0.7800	0.8160	0.8680	0.9040	0.9760
	ADSO	0.7299	0.8070	0.8674	0.9044	0.9647

Çizelge 5. 6: Veriler Power Lindley () dağılımlı olduğunda, parametresinin farklı değerleri için doğru seçim olasılıkları

β	Yöntem	n				
		30	50	80	100	200
0.25	SDSO	0.9160	0.9600	1.0000	0.9960	1.0000
	ADSO	0.8602	0.9468	0.9766	0.9900	0.9991
0.50	SDSO	0.9040	0.9680	1.0000	0.9960	1.0000
	ADSO	0.8911	0.9600	0.9855	0.9866	0.9997
1.00	SDSO	0.9280	0.9800	0.9840	0.9960	1.0000
	ADSO	0.9134	0.9667	0.9867	0.9914	0.9998
1.50	SDSO	0.9200	0.9320	0.9920	0.9960	1.0000
	ADSO	0.8926	0.9327	0.9871	0.9920	0.9994
2.00	SDSO	0.9000	0.9400	0.9880	0.9960	1.0000
	ADSO	0.9018	0.9384	0.9862	0.9853	0.9996
4.00	SDSO	0.8760	0.9520	0.9840	0.9880	0.9960
	ADSO	0.8723	0.9450	0.9800	0.9868	0.9991
10.00	SDSO	0.8720	0.9320	0.9880	0.9880	1.0000
	ADSO	0.8738	0.9314	0.9852	0.9894	0.9981

İlk senaryoda elde edilen sonuçlara benzer olarak, Çizelge 5.4 – 5.6 ile verilen ikinci senaryoya ait simülasyon çalışması sonuçlarından da tüm durumlarda SDSO ve ASDO değerlerinin birbirini yakından takip ettiği ve örneklem çapı n arttıkça doğru seçim olasılıklarının arttığı görülmektedir. Bununla birlikte Power Lindley dağılımının θ parametresinin değeri büyüdükçe doğru seçim olasılıklarının arttığı görülmektedir.

6. SONUÇ

Power Lindley dağılımı ve genelleştirilmiş üstel dağılım, sağlık, fen ve mühendislik gibi farklı alanlardan elde edilen pozitif değerli ve çarpık verilerin modellenmesinde başarı ile kullanılan dağılımlardır. Power Lindley dağılımı ilk olarak Githany vd.(2013)'nin yaptığı çalışmada, Lindley dağılımına daha esnek bir veri modelleme yapısı kazandırmak amacıyla tanımlanmıştır. Genelleştirilmiş üstel dağılım ise Gupta ve Kundu (1999) tarafından üstel dağılıma dayalı olarak, Weibull, Log-Normal, Lindley ve Gamma dağılımı gibi popüler dağılımlara alternatif olarak önerilmiştir. Gözlenmiş bir veri seti için, ortak uygulama alanlarına sahip olan Power Lindley ve genelleştirilmiş üstel dağılımdan hangisinin optimal bir model olacağı belirlenmesi istatistiksel sonuç çıkarım açısından önemlidir.

Bu çalışmada genelleştirilmiş üstel dağılım ve Power Lindley dağılımları arasındaki ayırım problemi göz önüne alınmıştır. İlgilenilen problemin çözümüne yönelik olarak, çalışmada RML yöntemine dayalı bir ayırıcı istatistik elde edilmiştir. Ayrıca elde edilen ayırıcı istatistiğin asimptotik dağılımı belirlenmiştir.

Yapılan simülasyon çalışmaları ile çalışmada elde edilen ayırıcı istatistiğin doğru seçim performansı değerlendirilmiştir. Simülasyon çalışmalarından elde edilen sonuçlara göre tüm durumlarda gözlem sayısı n arttıkça yöntemin doğru seçim olasılığı artmaktadır. Ayrıca küçük örneklem çaplarında gerçekleştirilen simülasyon çalışmalarına göre hem simülasyona dayalı ayırım performansı hem de asimptotik sonuçlara dayalı ayırım performansı oldukça tatminkardır.

Sonuç olarak RML yöntemine dayalı olarak elde edilen ayırıcı istatistiğin genelleştirilmiş üstel ve Power Lindley dağılımlarının ayırımında kullanılabileceği söylenebilir.

KAYNAKLAR

- Abramowitz, M., & Stegun, I. A.. Handbook of mathematical functions: with formulas, graphs, and mathematical tables. Courier Corporation.(1964), (Vol. 55)
- Atkinson, A. C, A. test for discriminating between models. *Biometrika*, (1969), 56(2), 337-347.
- Atkinson, A. C, A method for discriminating between models. *Journal of the Royal Statistical Society. Series B (Methodological)*, . (1970). , 323-353.
- Biçer, C., & Biçer, H. D., Geometrik süreç verileri için gamma ve weibull dağılımları arasındaki ayırım. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, Eyi , Özel Sayı, (2018), 18, 239-252.
- Cox, D. R.. Tests of separate families of hypotheses. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*.(1961, June), (Vol. 1, pp. 105-123).
- Cox, D. R. Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society. Series B (Methodological)*, (1962), 406-424.
- Dey, A. K., & Kundu, D. Discriminating between the Weibull and log-normal distributions for Type-II censored data. *Statistics*, (2012), 46(2), 197-214.
- Dumonceaux, R., & Antle, C. E. Discrimination between the log-normal and the Weibull distributions. *Technometrics*, (1973),15(4), 923-926.
- Dyer, A. R. Discrimination procedures for separate families of hypotheses. *Journal of the American Statistical Association*, . (1973),68(344), 970-974.

- Demirci Biçer, H., Biçer C. Discrimination between gamma and lognormal distributions for geometric process data, *Researches on Science and Art in 21st Century Turkey*, Gece Publishing, (2017), 2830 -2836.
- Bain, L. J., & Engelhardt, M.. Probability of correct selection of Weibull versus gamma based on likelihood ratio. *Communications in statistics-theory and methods*, (1980), 9(4), 375-381.
- Fearn, D. H., & Nebenzahl, E.. On the maximum likelihood ratio method of deciding between the Weibull and Gamma distributions. *Communications in Statistics-Theory and Methods*, (1991),20(2), 579-593.
- Ghitany, M. E., Al-Mutairi, D. K., Balakrishnan, N., & Al-Enezi, L. J. Power Lindley distribution and associated inference. *Computational Statistics & Data Analysis*, . (2013), 64, 20-33.
- Gupta, R. D., & Kundu, D.. Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, (1999), 41(2), 173-188.
- Gupta, R. D., & Kundu, D.. Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation*, (2001), 69(4), 315-337.
- Gupta, R. D., & Kundu, D. Discriminating between Weibull and generalized exponential distributions. *Computational statistics & data analysis*, . (2003), 43(2), 179-196.
- Gupta, R. D., & Kundu, D.. Discriminating between gamma and generalized exponential distributions. *Journal of Statistical Computation & Simulation*, (2004), 74(2), 107-121.
- Kundu, D., & Manglick, A.. Discriminating between the Weibull and log normal distributions. *Naval Research Logistics (NRL)*, (2004), 51(6), 893-905.

Kundu, D., & Manglick, A. (2005). Discriminating between the log-normal and gamma distributions. *Journal of the Applied Statistical Sciences*, 14, 175-187.

Kundu, D., Gupta, R. D., & Manglick, A. (2005). Discriminating between the log-normal and generalized exponential distributions. *Journal of Statistical Planning and Inference*, 127(1-2), 213-227.

