

# Üstelleştirilmiş Genelleştirilmiş Gamma Dağılımında Yeni Bir Uygulama

## *A New Application in the Exponentiated Generalized Gamma Distribution*

Ceren ÜNAL<sup>1\*</sup> , Gamze ÖZEL<sup>1</sup> 

<sup>1</sup> Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Ankara, Türkiye.

### Öz

Dağılım sınıfları günümüzde uygulamalı bilimlerde sıkça kullanılmaktadır. Dağılım sınıflarının kendilerine özgü bazı özellikleri olduğundan her ailenin farklı kullanım alanları oluşmuştur. Örneğin dört parametrelili genelleştirilmiş gamma dağılımının bozulma oranı fonksiyonu monoton azalmayan ve modeli monoton olduğundan dağılım yaşam süresi verileri ve güvenilirlik analizinde oldukça yaygın kullanılmaktadır. Söz konusu kullanım alanların sürekli olarak genişlemesi ve yeni ihtiyaçların ortaya çıkmasıyla ayrıca mevcut dağılım aileleriyle çalışmanın bazı durumlarda çalışmayı zorlaştırdığı ve bazı dezavantajlara sahip olduğu için fazla problemle karşılaşmaktadır. Bu nedenle literature sürekli olarak farklı istatistiksel özelliklere sahip yeni dağılım sınıfları kazandırılmaktadır. Bu çalışmada da literature kazandırılan yeni dağılım sınıfından biri olarak üstelleştirilmiş genelleştirilmiş dağılım ailesi üzerinde durulacaktır. Özel olarak üstelleştirilmiş genelleştirilmiş gamma dağılımının özelliklerinden, kullanım alanlarından, diğer dağılımlara göre avantajlarından bahsedilecektir. Son olarak gerçek bir veri seti ile çalışmada kullanılacak dağılım test edilip dağılımın parametreleri en çok olabilirlik yöntemi aracılığıyla tahmin edilecektir.

**Anahtar Kelimeler:** Üstelleştirilmiş genelleştirilmiş dağılım, Hazard fonksiyonu, Dağılım fonksiyonu, En çok olabilirlik yöntemi

### Abstract

Distribution classes are frequently used in applied sciences today. Since distribution classes have some unique features, each family has different usage areas. For example, since the distortion rate function of the four-parameter generalized gamma distribution is not monotonously decreasing and the model is monotonous, the distribution is widely used in lifetime data and reliability analysis. With the continuous expansion of the said usage areas and the emergence of new needs, more problems are encountered because working with existing distribution families makes it difficult to work in some cases and has some disadvantages. For this reason, new distribution classes with different statistical properties are constantly being introduced to the literature. In this study, the exponentiated generalized distribution family, as one of the new distribution classes that are introduced to the literature, will be emphasized. In particular, the properties of the exponentiated generalized gamma distribution, its usage areas, and its advantages over other distributions are mentioned. Finally, the distribution is used in the study with a real data set is tested as well the parameters of the distribution are estimated using the maximum likelihood method.

**Keywords:** The exponential generalized distribution, Hazard function, Distribution function, The maximum likelihood method

## I. GİRİŞ

Üstelleştirilmiş genelleştirilmiş dağılım ailesi ilk olarak Gupta ve Kundu [1] tarafından üstelleştirilmiş üstel dağılım olarak adlandırılan ve dağılım fonksiyonu,

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

olan standart üstel dağılımın bir genellemesi biçiminde literature kazandırılmıştır. Bu aile günümüzde biyoloji ve mühendislik alanlarında sıkça kullanılmaktadır. Üstelleştirilmiş üstel dağılımın dağılım fonksiyonu aslında standart kümülatif üstel dağılımın  $\alpha$ . kuvvetidir. Söz konusu dağılım literature kazandırıldıktan sonra Gupta ve Kundu [2] tarafından bazı matematiksel özellikleri araştırılmıştır. Benzer şekilde Nadarajah ve Kotz [3] üstelleştirilmiş gamma, üstelleştirilmiş Fréchet ve üstelleştirilmiş Gumbel dağılımlarını literature kazandırmıştır. Bu çalışmada üstelleştirilmiş dağılımların bir genellemesi olan yani üstelleştirilmiş biçimdeki dağılımların bir genişlemesi ve bazı yapısal özellikleri üzerinde durulacaktır. Çalışmanın temelini oluşturan söz konusu dağılım sınıfının genel formu aşağıdaki gibidir.

$G(x)$  sürekli bir dağılımın kümülatif dağılım fonksiyonu olmak üzere üstelleştirilmiş genelleştirilmiş dağılım sınıfının kümülatif dağılım fonksiyonu,

$$F(x) = [1 - \{1 - G(x)\}^\alpha]^\beta, \beta > 0, \alpha > 0 \quad (1)$$

biçimindedir. Burada  $\beta$  ve  $\alpha$  ilave şekil parametreleridir. Fonksiyonda yer alan iki parametreye rağmen beta genelleştirilmiş dağılım ailesi Eugene vd., [4] aksine karmaşık yapıda bir fonksiyon değildir. Fakat Eşitlik (1) tamamlanmamış beta fonksiyonu içerir. Üstelleştirilmiş genelleştirilmiş dağılımın bu avantajına ek olarak (1) eşitliğinin işlem yapılabilirlik özelliğinden özellikle de simülasyon için dağılımdan sayı üretmede çalışmanın ikinci bölümünde verilecek olan kantil fonksiyonunun basit yapısından dolayı oldukça kullanışlıdır [11, 12]. Üstelleştirilmiş genelleştirilmiş dağılımın olasılık yoğunluk fonksiyonu,

$$f(x) = \alpha\beta\{1 - G(x)\}^{\alpha-1} [1 - \{1 - G(x)\}^\alpha]^{\beta-1} g(x), \quad (2)$$

biçimindedir. Üstelleştirilmiş genelleştirilmiş ailenin yoğunluğu (2) kuyrukların daha da esnek olmasına olanak sağlar.

Literatürde bulunan dağılımlar yukarıda bahsedildiği üzere bazı kısıtlara sahip olduğundan bu kısıtları mümkün olduğunca ortadan kaldırarak daha esnek hesaplamalar ve analizler yapılması sebebiyle böyle bir dağılım sınıfına odaklanılmıştır. Bu çalışmada üstelleştirilmiş türdeki dağılımları genişleterek elde edilen üstelleştirilmiş genelleştirilmiş dağılım sınıfından bahsedilecektir. Daha sonra bu dağılım ailesinin temel dağılımının gamma olması durumunda bazı istatistiksel ve matematiksel özellikleri üzerinde durulacaktır. Çalışmanın üçüncü bölümünde elde edilen dağılım için en çok olabilirlik yöntemi yardımıyla parametre tahmini üzerinde durulacaktır. Dördüncü bölümde elde edilen dağılımın uygulanabilir olduğu gerçek bir veri seti ile sınanacak ve literatürde yer alan bazı dağılımlar ile karşılaştırma yapılacaktır. Son bölümde ise, çalışmanın genel bir değerlendirilmesi yapılacaktır.

## II. ÜSTELLEŞTİRİLMİŞ GENELLEŞTİRİLMİŞ GAMMA DAĞILIMININ ÖZELLİKLERİ

Literatürde üstelleştirilmiş genelleştirilmiş gamma dağılımı ilk olarak Cordeiro vd. [5, 6, 7] tarafından önerilmiştir. Bu dağılımın literatürde yer alan diğer dağılımlardan bazı avantajları bulunmaktadır. Üstelleştirilmiş genelleştirilmiş gamma dağılımının avantajları çeşitli yaşam süresi verilerinin modellenmesi problemlerinde hazard fonksiyonunun her formuyla uyumlu olmasıdır [9, 10]. Diğer avantajı ise özel olarak seçilen parametre değerlerinde Weibull, üstelleştirilmiş üstel, uyarlanmış Weibull ve genelleştirilmiş Rayleigh dağılımları gibi alt dağılımları içermesidir. Üstelleştirilmiş genelleştirilmiş gamma dağılımı aynı zamanda kuvvet tipi bozulma

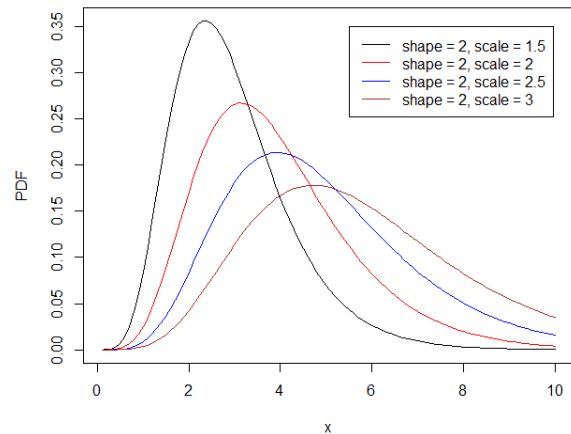
oranlarını modellemede uyumlu olmasının yanı sıra üstelleştirilmiş Weibull ve uyarlanmış Weibull dağılımları gibi özel alt dağılımlarla uyum iyiliği testi için uygundur [14]. Bu bölümde ilk olarak söz konusu dağılımın olasılık yoğunluğu verilecek ve farklı parametre değerlerine göre fonksiyonun grafikleri çizilerek dağılımın modu ve simetrisi hakkında yorumlar yapılacaktır. Ayrıca dağılımın özellikle simülasyon çalışmalarında sayı üretmede kullanılan kantil fonksiyonu, dağılımın momentleri, moment çıkaran fonksiyonu ve sıralı istatistikleri verilecektir. Sıralı istatistiklere ilişkin olarak dağılımın birinci, sonuncu ve herhangi bir sıralı istatistiğinin dağılım ve olasılık yoğunluk fonksiyonları verilecektir. Devamında ise, dağılımın olasılık yoğunluk fonksiyonuna bağlı olarak yaşam fonksiyonu ve hazard fonksiyonu verilecektir. Hazard fonksiyonunun dağılımın farklı parametre değerlerine göre grafikleri çizilerek yorumlama yapılacaktır.

### 2.1. Olasılık Yoğunluk Fonksiyonu

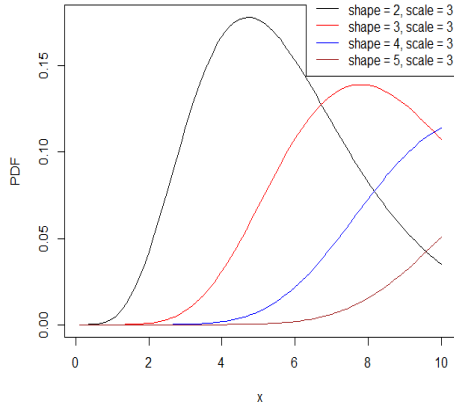
Üstelleştirilmiş genelleştirilmiş gamma dağılımında temel dağılımın parametreleri  $G \sim \text{Gamma}(a, b)$  ve dağılım fonksiyonu  $G(x) = \frac{\gamma(a, bx)}{\Gamma(a)}$   $x > 0, a > 0$  ölçek parametresi ve  $b > 0$  şekil parametresi olmak üzere olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{\alpha\beta b^\alpha x^{\alpha-1} e^{-bx}}{\Gamma(a)} \times \left\{1 - \frac{\gamma(a, bx)}{\Gamma(a)}\right\}^{\alpha-1} \times \left\{1 - \left(1 - \frac{\gamma(a, bx)}{\Gamma(a)}\right)^\alpha\right\}^{\beta-1}, \quad (3)$$

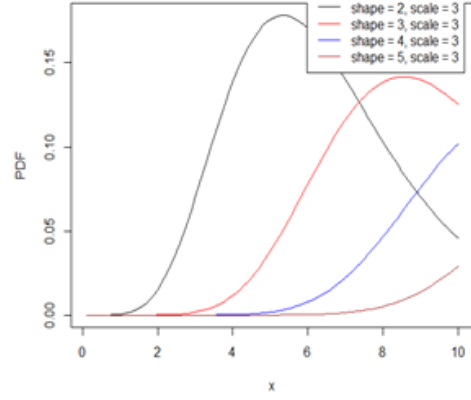
$x > 0$  şeklindedir. Burada  $\gamma(a, x) = \int_0^x y^{a-1} e^{-y} dy$  tamamlanmamış gamma fonksiyonu ve  $\Gamma(a) = \int_0^\infty y^{a-1} e^{-y} dy$  ifadesi ise gamma fonksiyonudur. Dağılımın farklı parametre değerleri için grafikleri sırasıyla Şekil 1 – Şekil 12 arasında yer almaktadır.



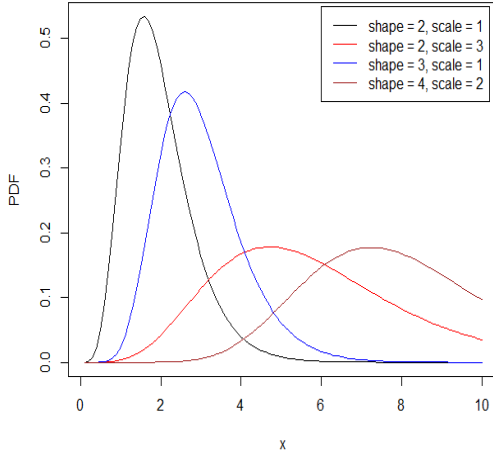
Şekil 1. Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 2, \beta = 3$ ) durumunda olasılık yoğunluk fonksiyonu grafiği



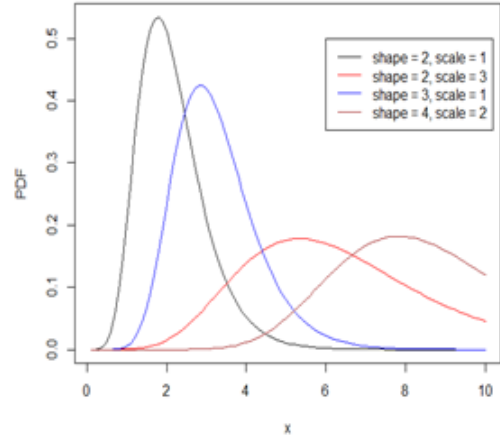
**Şekil 2.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 2, \beta = 3$ ) durumda olasılık yoğunluk fonksiyonu grafiği



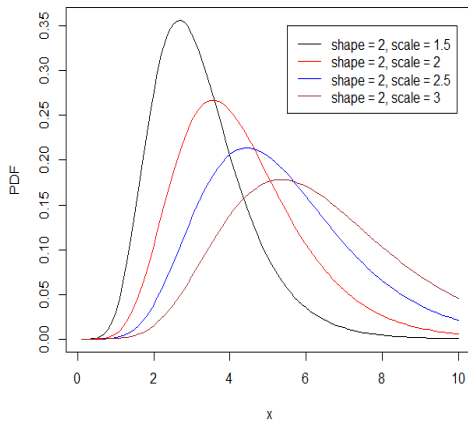
**Şekil 5.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 2, \beta = 4$ ) durumda olasılık yoğunluk fonksiyonu grafiği



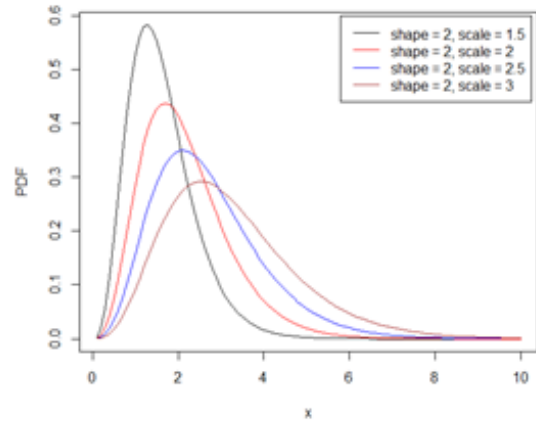
**Şekil 3.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 2, \beta = 3$ ) durumda olasılık yoğunluk fonksiyonu grafiği



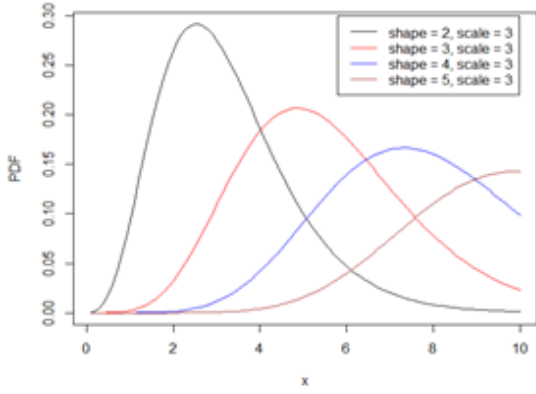
**Şekil 6.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 2, \beta = 4$ ) durumda olasılık yoğunluk fonksiyonu grafiği



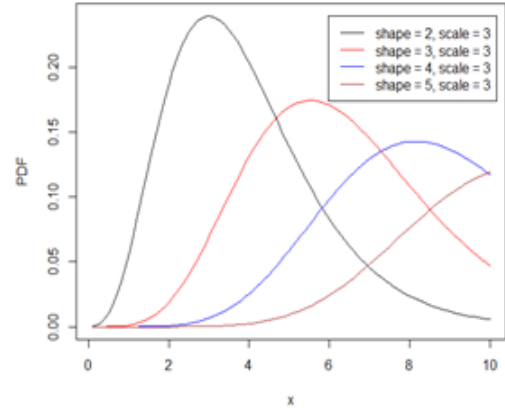
**Şekil 4.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 2, \beta = 4$ ) durumda olasılık yoğunluk fonksiyonu grafiği



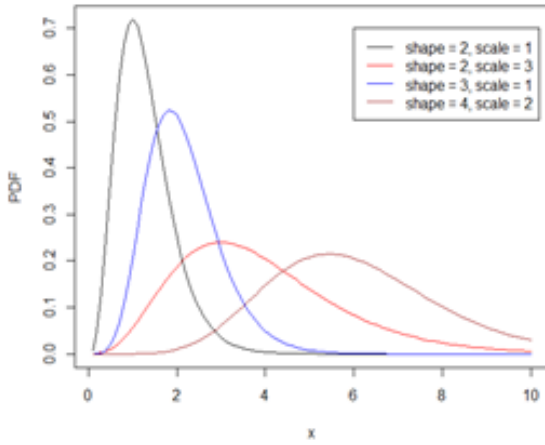
**Şekil 7.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 3, \beta = 2$ ) durumda olasılık yoğunluk fonksiyonu grafiği



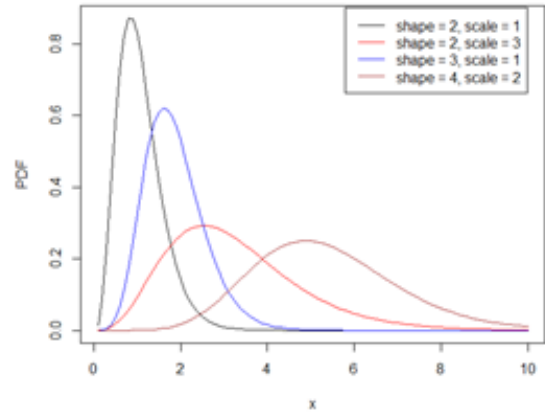
**Şekil 8.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 3, \beta = 2$ ) durumda olasılık yoğunluk fonksiyonu grafiği



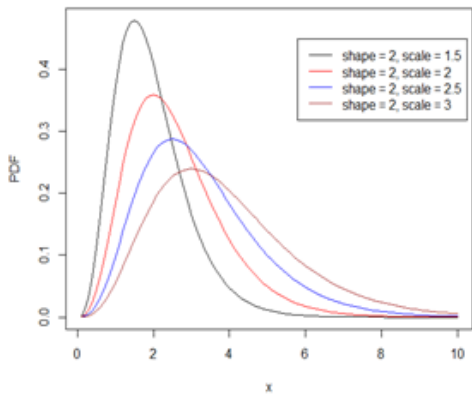
**Şekil 11.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 4, \beta = 2$ ) durumda olasılık yoğunluk fonksiyonu grafiği



**Şekil 9.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 3, \beta = 2$ ) durumda olasılık yoğunluk fonksiyonu grafiği



**Şekil 12.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 4, \beta = 2$ ) durumda olasılık yoğunluk fonksiyonu grafiği



**Şekil 10.** Üstelleştirilmiş genelleştirilmiş gamma dağılımında ( $\alpha = 4, \beta = 2$ ) durumda olasılık yoğunluk fonksiyonu grafiği

Grafiklerden görüldüğü üzere dağılımın tek modlu olduğu ve simetrik bir dağılım olmadığı, sağa çarpık bir dağılım olduğu açıktır. Üstelleştirilmiş genelleştirilmiş dağılımın  $\alpha$  ve  $\beta$  parametreleri dağılımın; yalnızca  $\beta$  parametresinin değişmesi durumunda dağılımın az da olsa konumunun değiştiği ve  $\alpha$  parametresinin değişmesi durumunda ise dağılımın basıklığının az da olsa değiştiği görülmektedir. Kısaca üstelleştirilmiş genelleştirilmiş dağılımın  $\alpha$  ve  $\beta$  parametrelerinin değişmesi durumunda dağılımın sağdan çarpık olduğu ve sadece  $\alpha$  parametresinin değişmesi durumunda ise basıklığının az miktarda değiştiği gözlemlenmiştir. Ayrıca temel dağılımın yani gamma dağılımının şekil ve ölçek parametreleri değiştiğinde ise şekil parametresine bağlı olarak dağılımın şeklinin, ölçek parametresine bağlı olarak dağılımın basıklığının değiştiği görülmektedir.

**2.2. Kantil Fonksiyonu**

Kantil fonksiyonu simülasyon çalışmalarında ilgilenilen dağılımdan sayı üretme konusunda oldukça önemlidir. Bu sebeple çalışmada araştırılan dağılımın kantil fonksiyonu,

$$x = Q_G \left( \left[ 1 - \left( 1 - u^{\frac{1}{\beta}} \right)^{\frac{1}{\alpha}} \right] \right) \tag{4}$$

biçimindedir. Eşitlikte yer alan  $Q_G(u)$ , gamma dağılımının kantil fonksiyonu ve  $u \sim U(0,1)$  dağılımından üretilen rasgele bir sayıdır. Yukarıda yer alan  $x$  eşitliği kullanılarak herhangi bir paket program yardımıyla üstelleştirilmiş genelleştirilmiş gamma dağılımından rasgele sayı üretilebilir.

**2.3. Dağılımın k. Momenti**

İlk olarak üstelleştirilmiş genelleştirilmiş dağılım sınıfına göre  $(k, j)$  inci  $X$  'in ağırlıklandırılmış momenti,

$$\tau_{k,j} = E[X^k G(X)^j] = \int_{-\infty}^{\infty} x^k G(x)^j g(x) dx \tag{5}$$

biçiminde tanımlanır. Eşitlik (5) kullanılarak  $k$  inci orjine göre momente şu şekilde geçilebilir.

$$E(X^k) = \alpha \beta \sum_{j=0}^{\infty} t_j \tau_{k,j} \tag{6}$$

ve  $t_j = t_j(\alpha, \beta)$  katsayıları,

$$t_j = \frac{(-1)^j \Gamma(\beta)}{j!} \sum_{k=0}^{\infty} \frac{(-1)^k \Gamma((k+1)\alpha)}{\Gamma(\beta-k)\Gamma((k+1)\alpha-j)k!}$$

biçiminde ifade edilir.  $X$  'in ağırlıklandırılmış momentinin ağırlıklarının toplamı sonsuz olabilir. Bu yüzden integralin hesaplanmasında problem çıkmaktadır. Bu problemi aşmak için ikinci bir formül verilir.

$Q_G(x) = G^{-1}(x)$  ve  $G(x) = u$  olmak üzere,

$$\tau_{k,j} = \int_0^1 Q_G(u)^k u^j du \tag{7}$$

biçimindedir. Burada integral  $(0,1)$  aralığı üzerinden rahatlıkla hesaplanabilir.

Üstelleştirilmiş genelleştirilmiş gamma dağılımının  $k$  inci momenti tamamlanmamış gamma fonksiyonunun güç serisi açılımı,

$$G_{a,b}(x) = \frac{(ba)^a}{\Gamma(a)} \sum_{m=0}^{\infty} \frac{(-bx)^m}{(a+m)m!},$$

dır. Buradan aşağıdaki seri açılımı,

$$f(x) = \frac{\alpha \beta b^a x^{a-1} e^{-bx}}{\Gamma(a)} \sum_{j=0}^{\infty} t_j \frac{(bx)^{aj}}{\Gamma(a)^j}$$

$$\sum_{m_1=0}^{\infty} \dots \sum_{m_j=0}^{\infty} \frac{(-bx)^{m_1+m_2+\dots+m_j}}{(a+m_1)\dots(a+m_j)m_1!\dots m_j!}$$

biçiminde elde edilir. Gerekli işlemler Cordeiro vd. (2011) tarafından yapılmış olup,

$$\tau_{k,j} = \frac{\Gamma(k+a(j+1))}{a^j b^k \Gamma(a)^{j+1}} F_A^{(j)}(r+a(j+1); a, \dots, a; a+1, \dots, a+1, -1, \dots, -1),$$

biçiminde elde edilir. Burada  $F_A^{(j)}$  Lauricella fonksiyonu olarak adlandırılır.

**2.4. Moment Çıkaran Fonksiyonu**

Çalışmada üzerinde durulan üstelleştirilmiş genelleştirilmiş gamma dağılımının üyesi olduğu sınıfın moment çıkarar fonksiyonunun üç farklı gösterimi vardır. Bu gösterimler sırasıyla aşağıda ifade edildiği gibidir. Çalışmada araştırılan dağılımın moment çıkarar fonksiyonuna ulaşmak için birinci gösterimi kullanmak yeterlidir. Birinci gösterim,

$$M(s) = E(e^{sX}) = \sum_{r=0}^{\infty} \frac{\mu_r'}{r!} s^r \tag{8}$$

biçimindedir. Burada  $\mu_r' = E(X^r)$  Eşitlik (6) aracılığıyla elde edilir.

İkinci gösterim ise,

$$f(x) = \sum_{j=0}^{\infty} t_j^* h_{j+1}(x), t_j^* = \frac{\alpha \beta t_j}{(k+1)} \text{ ve } h_{j+1}(x) = (j+1)g(x)G(x)^j \text{ ile } \text{Exp}^{j+1}(G) \text{ dağılımı olduğu Cordeiro vd. (2011) tarafından gösterilmiştir. Buradan,}$$

$$M(s) = E(e^{sX}) = \sum_{j=0}^{\infty} t_j^* M_{j+1}(s) \tag{9}$$

elde edilir. Bu eşitlikte  $M_{j+1}(s)$ ,  $\text{Exp}^{j+1}(G)$  dağılımının moment çıkarar fonksiyonudur. Bu nedenle  $M(s)$ , Exp-G dağılımının moment çıkarar fonksiyonu yardımıyla da belirlenebilir.

Son olarak,

$$M(s) = \sum_{j=0}^{\infty} t_j \rho_j(s) \tag{10}$$

ve

$$\rho_j(s) = \int_{-\infty}^{\infty} e^{sx} G(x)^j g(x) dx$$

dir. Temel dağılımın kantil fonksiyonunun  $Q_{G(u)} = G^{-1}(u)$  olduğunun bilinmesiyle

$$\rho_j(s) = \int_0^1 e^{sQ_G(u)} u^j du \tag{11}$$

olduğu açıktır. Buradan (10) ve (11) eşitliklerinden moment çıkarar fonksiyon direkt olarak elde edilebilir.

**2.5. Sıralı İstatistikler**

i.  $f_{i:n}(x), i = 1, 2, \dots, n$   $i$  . sıra istatistiğinin olasılık yoğunluk fonksiyonu ve  $X_1, X_2, \dots, X_n$  rasgele değişkenleri birbirinden bağımsız aynı dağılımlı olmak üzere,

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} F(x)^{i-1} \{1 - F(x)\}^{n-i}$$

dır. Burada,  $f(x)$  ve  $F(x)$  yerine üstelleştirilmiş genelleştirilmiş dağılımın sırasıyla olasılık yoğunluk fonksiyonu ve dağılım fonksiyonu yerine yazılırsa;

$$f_{i:n}(x) = \frac{\alpha\beta}{B(i, n-i+1)} g(x) \{1 - G(x)\}^{\alpha-1} [1 - \{1 - G(x)\}^\alpha]^{\beta i-1} \times \{1 - [1 - \{1 - G(x)\}^\alpha]^\beta\}^{n-i}$$

olduğu açıktır. Eşitlikte binom açılımı kullanılarak  $f_{i:n}(x)$  daha açık hale getirilebilir.

$$f_{i:n}(x) = \frac{\alpha\beta}{B(i, n-i+1)} f(x) \sum_{k=0}^{n-i} \frac{(-1)^k}{(i+k)} \binom{n-i}{k} f_{a,b,\beta(i+k),\alpha}(x) \tag{12}$$

Eşitlik (12),  $i$  inci sıra istatistiğinin üstelleştirilmiş genelleştirilmiş gamma dağılımının olasılık yoğunluk fonksiyonunu ifade eder. Eşitliğe ilişkin olarak dağılımın 1. ve n. sıra istatistiği sırasıyla aşağıdaki gibi elde edilir.

$$f_{1:n}(x) = \frac{\alpha\beta}{B(1,n)} f(x) \sum_{k=0}^{n-1} \frac{(-1)^k}{(1+k)} \binom{n-1}{k} f_{a,b,\beta(1+k),\alpha}(x),$$

ve

$$f_{n:n}(x) = \frac{\alpha\beta}{nB(n,1)} f(x) f_{a,b,\beta(n),\alpha}(x),$$

biçimindedir. Burada,

$$f(x) = \frac{\alpha\beta b^\alpha x^{\alpha-1} e^{-bx}}{\Gamma(a)} \left\{1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right\}^{\alpha-1} \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^{\beta-1}$$

ve  $x > 0$  olarak tanımlıdır.

ii.  $F_{i:n}(x), i = 1, 2, \dots, n$   $i$  . sıra istatistiğinin dağılım fonksiyonu  $X_1, X_2, \dots, X_n$  rasgele değişkenleri birbirinden bağımsız aynı dağılımlı olmak üzere,

$$F_{i:n}(x) = \sum_{k=i}^n \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k} \tag{13}$$

Burada,  $F(x) = \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^\beta$  ve

$\gamma(a, x) = \int_0^x y^{\alpha-1} e^{-y} dy$  biçimindedir. Burada Eşitlik (13) yeniden düzenlenirse,

$$F_{i:n}(x) = \sum_{k=i}^n \binom{n}{k} \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^{\beta k} [1 - \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^\beta]^{n-k}$$

ile üstelleştirilmiş genelleştirilmiş gamma dağılımının  $i$  inci sıra istatistiğinin dağılım fonksiyonu elde edilir. Eşitliğe ilişkin olarak dağılımın 1. ve n. sıra istatistiği sırasıyla aşağıdaki gibi elde edilir.

$$F_{1:n}(x) = \sum_{k=1}^n \binom{n}{k} \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^{\beta k} [1 - \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^\beta]^{n-k}$$

ve

$$F_{n:n}(x) = \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^{\beta n}.$$

Ayrıca, söz konusu dağılımın sıra istatistiğinin  $k$  inci momentini

$$\begin{aligned} \mu'_{k(i:n)} &= \frac{\beta}{b^k B(i, n-i+1)} \\ &\times \sum_{l=0}^{n-i} \sum_{j=0}^{\infty} \sum_{m=0}^j \frac{(-1)^{l+j+m}}{\Gamma(a)^{\beta(i+l)}} \binom{n-i}{l} \binom{\beta(i+l)-1}{j} \binom{j}{m} I\left(a, \frac{k}{a}, m\right) \end{aligned}$$

biçimde elde edilir. Burada,

$$\begin{aligned} I\left(a, \frac{k}{a}, m\right) &= a^{-m} \Gamma\left(\frac{a}{\alpha} + a(m+1)\right) \\ &\times F_A^{(m)}(r + a(m+1); a, \dots, a; a+1, \dots, a+1, -1, \dots, -1) \end{aligned}$$

ile gösterilmektedir. Verilen bu denklem sistemleri herhangi bir paket programı yardımıyla çözümlenerek üstelleştirilmiş genelleştirilmiş gamma dağılımının sıralı istatistiğinin  $k$  inci momentine ulaşılmış olur.

**2.6. Yaşam ve Hazard Fonksiyonları**

Üstelleştirilmiş genelleştirilmiş gamma dağılımının olasılık yoğunluk fonksiyonu aşağıdaki gibi tanımlı olduğu Eşitlik (3)'de verilmişti. Olasılık yoğunluk fonksiyonuna bağlı olarak üstelleştirilmiş genelleştirilmiş gamma dağılımının yaşam fonksiyonu ve hazard fonksiyonu sırasıyla Eşitlik (14) – (15)'de verilmektedir.

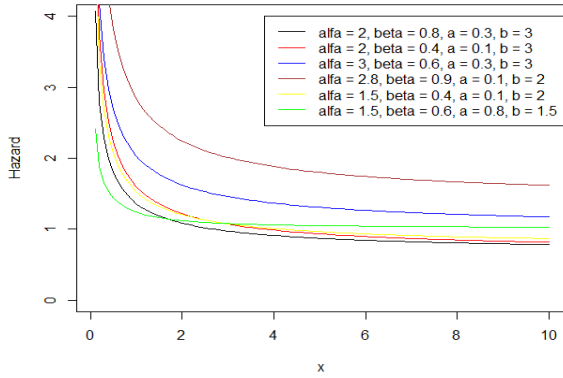
$$S(x) = 1 - F(x) = 1 - \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^\beta,$$

ve

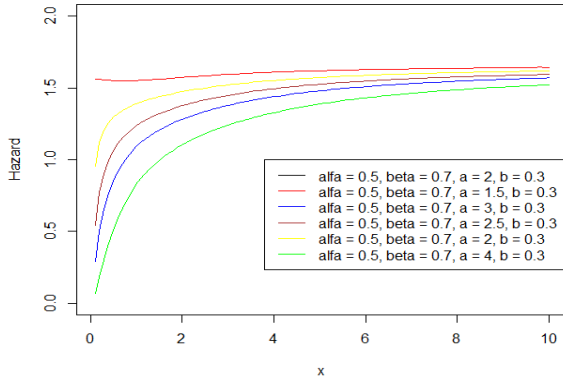
$h(x) = \frac{f(x)}{S(x)}$  formülünden yararlanılarak,

$$h(x) = \frac{\alpha\beta b^\alpha x^{\alpha-1} e^{-bx} \left\{1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right\}^{\alpha-1} \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^{\beta-1}}{1 - \left\{1 - \left(1 - \frac{\gamma(a,bx)}{\Gamma(a)}\right)^\alpha\right\}^\beta} \tag{15}$$

olarak elde edilir. Eşitlik (15)'de verilen hazard fonksiyonu R paket programı aracılığıyla kullanılarak dağılımın farklı parametre değerlerinde grafikleri Şekil 13 – 14'de elde edilmektedir.



Şekil 13. Üstelleştirilmiş genelleştirilmiş gamma dağılımı durumunda farklı parametre değerlerinde olasılık hazard fonksiyonu grafiği



Şekil 14. Üstelleştirilmiş genelleştirilmiş gamma dağılımı durumunda farklı parametre değerlerinde olasılık hazard fonksiyonu grafiği  
Grafikleri yorumlamadan önce üstelleştirilmiş genelleştirilmiş gamma dağılımının hazard fonksiyonunun Cordeiro vd. (2011) tarafından,

- i.  $\alpha a > 1, \beta > 1$  ve  $\forall x > 0, h(x)' > 0$  ise  $h(x)$  artan,
- ii.  $0 < \alpha a < 1, 0 < \beta < 1$  ve  $\forall x > 0, h(x)' < 0$  ise  $h(x)$  azalan,
- iii.  $\alpha < 1$  ve  $\alpha\beta \geq 1$  ise  $h(x)$  tek modlu,
- iv.  $\alpha > 1$  ve  $\alpha\beta < 1$  ise  $h(x)$  fonksiyonun grafiği küvet tipi şeklinde

özellikleri olduğunu elde etmişlerdir.

Bu bilgi altında Şekil 13 ve Şekil 14’de hazard fonksiyonunun grafik yorumlandığında Şekil 13’te parametrelerin (iv) önermede yer alan koşulları sağladığı yani grafiğin şeklinin küvet tipi biçiminde olduğu görülür. Şekil 14’te üstelleştirilmiş genelleştirilmiş dağılımın parametreleri sabit tutulduğunda ve diğer parametrelerin (ii) koşulu altında değiştirildiği durumda grafiğin şeklinin azalan olduğu görülebilir.

### III. PARAMETRE TAHMİNİ

Çalışmanın bu bölümünde üstelleştirilmiş genelleştirilmiş dağılım sınıfına ilişkin olarak parametre tahmini üzerinde durulacaktır. Parametre tahminini elde etmek için en çok olabirlik tahmin yöntemi kullanılacaktır.

$x_1, x_2, \dots, x_n$  ÜGG( $\alpha, \beta, \lambda$ ) dağılımından alınan  $n$  tane örneklem olsun. Burada  $\lambda$  gamma dağılımına ( $G(x; \lambda)$ ) ilişkin parametre vektörüdür. Üstelleştirilmiş genelleştirilmiş dağılım sınıfının logaritmik olabirlik fonksiyonu  $\theta = (\alpha, \beta, \lambda)$  parametre vektörü olmak üzere,

$$\begin{aligned} \ell(\theta) = & n \log(\alpha) + n \log(\beta) \\ & + \sum_{i=1}^n \log(g(x_i; \lambda)) \\ & + (\alpha - 1) \sum_{i=1}^n \log(1 - G(x_i; \lambda)) \\ & + (\beta - 1) \sum_{i=1}^n \log[1 - (1 - G(x_i; \lambda))^\alpha] \end{aligned} \quad (16)$$

biçiminde elde edilir. Eşitlik (16), herhangi bir paket program kullanılarak çözülebilir.  $\theta$  parametre vektörüne ilişkin parametrelerin skor bileşenleri ise,

$$U_\alpha(\theta) = \frac{n}{\alpha} + \sum_{i=1}^n \log(1 - G(x_i; \gamma)) \left[ 1 - \frac{(\beta - 1)[1 - G(x_i; \gamma)]^\alpha}{1 - (1 - G(x_i; \gamma))^\alpha} \right],$$

$$U_\beta(\theta) = \frac{n}{\beta} + \sum_{i=1}^n \log(1 - (1 - G(x_i; \gamma))^\alpha),$$

$$\begin{aligned} U_{\gamma_j}(\theta) = & \sum_{i=1}^n \left\{ \frac{[g(x_i; \gamma)]_{\gamma_j}}{g(x_i; \gamma)} - \frac{(\alpha - 1)[G(x_i; \gamma)]_{\gamma_j}}{1 - G(x_i; \gamma)} \right. \\ & \left. + \frac{\alpha(\beta - 1)[1 - G(x_i; \gamma)]^{\alpha-1} [G(x_i; \gamma)]_{\gamma_j}}{1 - [1 - G(x_i; \gamma)]^\alpha} \right\}, \end{aligned}$$

biçimindedir. Burada,  $[g(x_i; \gamma)]_{\gamma_j} = \frac{\partial g(x_i; \gamma)}{\partial \gamma_j}$  ve  $[G(x_i; \gamma)]_{\gamma_j} = \frac{\partial G(x_i; \gamma)}{\partial \gamma_j}, j = 1, 2, \dots, p$  olarak ifade edilir. Son olarak elde edilen skor bileşenleri,

$$U_\alpha(\theta) = U_\beta(\theta) = U_{\gamma_j}(\theta) = 0,$$

lineer olmayan eşitliklerden herhangi bir paket program yardımıyla  $\theta$ 'nın en çok olabirlik tahminleri ( $\hat{\theta}$ ) elde edilmektedir.

Dağılımın parametreleri, hipotez testleri ve aralık tahmini için  $(p + 2) \times (p + 2)$  boyutunda gözlemlenmiş bilgi matrisine,  $J = J(\theta)$ , gerek duyulur. Bu matrisin birimleri,

$$J_{\alpha,\alpha} = -\frac{n}{\alpha^2} + (\beta - 1) \sum_{i=1}^n \frac{[1 - G(x_i; \gamma)]^\alpha \{\log(1 - G(x_i; \gamma))\}^2}{\{1 - [1 - G(x_i; \gamma)]^\alpha\}^2},$$

$$J_{\alpha,\beta} = -\sum_{i=1}^n \frac{[1 - G(x_i; \gamma)]^\alpha \log(1 - G(x_i; \gamma))}{1 - [1 - G(x_i; \gamma)]^\alpha},$$

$$J_{\alpha,\gamma_j} = -\sum_{i=1}^n \frac{[\hat{G}(x_i; \gamma)]_{\gamma_j}}{1 - G(x_i; \gamma)} + (\beta - 1) \sum_{i=1}^n \frac{[\hat{G}(x_i; \gamma)]_{\gamma_j} [1 - G(x_i; \gamma)]^{\alpha-1}}{\{1 - [1 - G(x_i; \gamma)]^\alpha\}^2} \times \{1 - [1 - G(x_i; \gamma)]^\alpha\} + \alpha \log(1 - G(x_i; \gamma)) \times \{1 - [1 - G(x_i; \gamma)]^\alpha\} + \alpha [1 - G(x_i; \gamma)]^\alpha\},$$

$$J_{\beta,\beta} = -\frac{n}{\beta^2},$$

$$J_{\beta,\gamma_j} = \alpha \sum_{i=1}^n \frac{[\hat{G}(x_i; \gamma)]_{\gamma_j} [1 - G(x_i; \gamma)]^{\alpha-1}}{\{1 - [1 - G(x_i; \gamma)]^\alpha\}^2},$$

$$J_{\gamma_j, \gamma_s} = \sum_{i=1}^n \frac{[\hat{g}(x_i; \gamma)]_{\gamma_j \gamma_s} g(x_i; \gamma) - [\hat{g}(x_i; \gamma)]_{\gamma_j} [\hat{g}(x_i; \gamma)]_{\gamma_s}}{[g(x_i; \gamma)]^2} + \sum_{i=1}^n \frac{[\hat{G}(x_i; \gamma)]_{\gamma_j \gamma_s} [1 - G(x_i; \gamma)] + [\hat{G}(x_i; \gamma)]_{\gamma_j} [\hat{G}(x_i; \gamma)]_{\gamma_s}}{[1 - G(x_i; \gamma)]^2} - \sum_{i=1}^n \frac{[1 - G(x_i; \gamma)]^{\alpha-2} \{(\alpha - 1) [\hat{G}(x_i; \gamma)]_{\gamma_j} [\hat{G}(x_i; \gamma)]_{\gamma_s} - [1 - G(x_i; \gamma)] [\hat{G}(x_i; \gamma)]_{\gamma_j \gamma_s}\}}{[1 - [1 - G(x_i; \gamma)]^\alpha]^\alpha} + \alpha \sum_{i=1}^n \frac{[\hat{G}(x_i; \gamma)]_{\gamma_j} [\hat{G}(x_i; \gamma)]_{\gamma_s} [1 - G(x_i; \gamma)]^{2(\alpha-1)}}{\{1 - [1 - G(x_i; \gamma)]^\alpha\}^2},$$

biçimindedir. Burada  $[\hat{g}(x_i; \gamma)]_{\gamma_j \gamma_s} = \frac{\partial^2 g(x_i; \gamma)}{\partial \gamma_j \partial \gamma_s}$ ,  $[\hat{G}(x_i; \gamma)]_{\gamma_j \gamma_s} = \frac{\partial^2 G(x_i; \gamma)}{\partial \gamma_j \partial \gamma_s}$ ,  $j, s = 1, 2, \dots, p$  olup,  $\sqrt{n}(\hat{\theta} - \theta)$  biçiminde tanımlı ifadenin asimptotik dağılımı  $N_{(p+2)}(0, I(\theta)^{-1})$ 'dir. Burada  $I(\theta)$  beklenen bilgi matrisini ifade etmektedir. Genelde  $I(\theta)$  'nın yerine  $\hat{\theta}$  'nın hesaplanan gözlenmiş bilgi matrisi  $(J(\hat{\theta}))$  alınır. Sonuç olarak, parametrelerin yaklaşık güven bölgeleri ve güven aralıkları  $N_{(p+2)}(0, J(\hat{\theta})^{-1})$  çok değişkenli normal dağılıma dayalı olarak oluşturulur.

### IV. UYGULAMA

Çalışmanın bu bölümünde gerçek bir veri setinin dağılımının üstelleştirilmiş genelleştirilmiş gamma dağılımı olması durumunda parametre tahmini yapılacak ve veri seti için literatürde yer alan bazı dağılımlar seçilmesi durumunda karşılaştırma yapılacaktır. Tahmin ve analiz yapılırken R paket programından yararlanılmıştır. Bu veri seti Maguire vd. [8] tarafından yayınlanan Büyük Britanya'da meydana gelen art arda gerçekleşen 109 adet kömür madeni kazaları arasındaki geçen günlere ilişkindir. Söz konusu veri setinin sıralanmış biçimi Çizelge 1'de detaylı olarak verilmiştir.

İlk olarak, söz konusu veri setinin betimleyici istatistikleri Çizelge 2'de detaylı olarak verilmektedir. Veri setine uygulanan en çok olabilirlik yöntemine ilişkin tahminler ise Çizelge 3'de yer almaktadır.

**Tablo 1.** Art arda gerçekleşen kömür madeni kazaları arasında geçen günler

1	4	7	11	13	15	15	17	18	19	19	20	20	22	23	28	29	31	32	36	37	47	48	49	50	54	54	55	59	59	61	61	66	72	72	75	78	78	81	93	96	99	108	113	114	120	120	120	123	124	129	131	137	145	151	156	171	176	182	188	189	195	203	208	215	217	217	217	224	228	233	255	271	275	275	275	286	291	312	312	312	315	326	326	329	330	336	338	345	348	354	361	364	369	378	390	457	467	498	517	566	644	745	871	1312	1357	1613	1630.
---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	-------

**Tablo 2.** Maden kazası veri setine ilişkin betimleyici istatistikler

Minimum	Aralık (Range)	Toplam	Medyan	Ortalama	Varyans	Standart Sapma	Değişim Katsayısı	Maksimum
1	1629	25432	145	233.321	87873.33	296.434	127.050	1.630

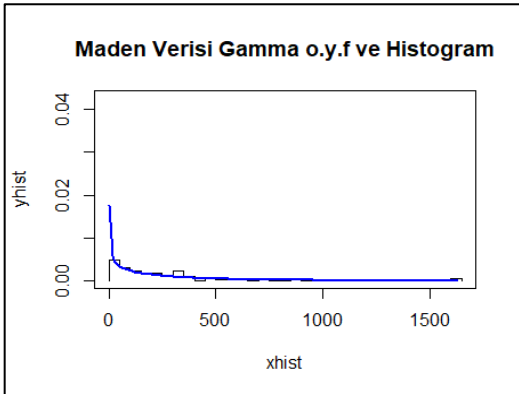
**Tablo 3.** Maden kazası veri setine ilişkin farklı dağılım durumlarında yapılan uyum iyiliği testleri ve en çok olabilirlik tahmin edicilerinin tahmini

Model	Uyum İyiliği Testi	Bilgi Kriterleri			Tahmin Değerleri (Hata)			
		K-S Test (P-Değeri)	AIC	BIC	CAIC	$\alpha$	$\beta$	$a$
EG-Gamma	0.3762	1409,8	1420,98	1410,6	0,2018 (0,099)	0,1413 (0,024)	6,218 (0,0987)	0,0183 (0,0057)
EG-Weibull	0,5466	1409,7	1420,4	1410,1	0,1765 (0,0287)	2,0268 (0,3275)	6,0035 (0,0232)	0,6186 (0,03169)
E-Gamma	0,3296	1410,9	1419,02	1411,1	1,021 (0,5031)	-	0,80517 (0,3614)	0,003407 (0,5136)
E-LogNormal	0,4511	1409,99	1418,07	1410,2	0,0934	-	6,6138 (0,252247)	0,56607 (0,1250)

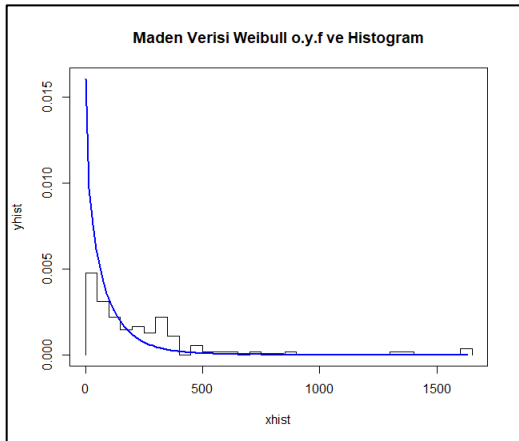


Farklı modellere yapılan uyum iyiliği testleri sonucunda veri setinin tabloda yer alan dört farklı modele uyum sağladığı görülmektedir. Uyum iyiliği testinin ardından parametrik bir tahmin yöntemi olan en çok olabilirlik tahmin sonuçları elde edilmiştir. Bu sonuçlara göre, çalışmanın temelini oluşturan üstelleştirilmiş genelleştirilmiş dağılım sınıfına ilişkin iki modelin diğer modellere göre parametreleri çok daha iyi tahmin ettiği gözlenmiştir. Fakat modelin uygunluğunun göreceli bir kriter olan bilgi kriterlerinden literatürde en sık kullanılan Akaike Bilgi Kriteri (AIC) açısından değerlendirildiğinde en iyi performansın EG-Weibull (Üstelleştirilmiş Genelleştirilmiş Weibull) dağılımı ve EG-Gamma (Üstelleştirilmiş Genelleştirilmiş Gamma) dağılımı ile verildiği gözlemlenmiştir. AIC açısından en iyi model olarak yorumlanabilen EG-Weibull dağılımı ve EG-Gamma dağılımı parametreleri tahmin etme noktasında kıyaslandığında  $\alpha$  ve  $a$  parametreleri için EG-Weibull dağılımının  $\beta$  ve  $b$  parametreleri açısından EG-Gamma dağılımının daha iyi tahminler elde ettiği görülmüştür.

Aynı zamanda, veri setinin EG-Weibull ve EG-Gamma dağılımına uyumunu görsel olarak test etmede elde edilen grafikler sırasıyla Şekil 15 ve Şekil 16'da gösterilmektedir.



Şekil 15. Veri setine ilişkin histogram ve EG-gamma olasılık yoğunluk fonksiyonu grafiği



Şekil 16. Veri setine ilişkin histogram ve eg-weibull olasılık yoğunluk fonksiyonu grafiği

Büyük Britanya'da meydana gelen art arda gerçekleşen 109 adet kömür madeni kazaları arasındaki geçen günlere ilişkin veri setinin yapılan analizler sonucunda model belirleme kriterlerine ve parametre tahminlerine göre EG-Gamma ve EG-Weibull modellerinin en uyumlu olduğu gözlemlenmiştir. Bu iki modele ilişkin olarak olasılık yoğunluk fonksiyonun çizitleri söz konusu veri setinin histogram grafiği ile birlikte çizilerek görsel anlamda da veri setinin bu dağılımlara uyumu Şekil 15 ve Şekil 16 'da verilmiştir. Fakat bu grafikler incelendiğinde EG-Gamma dağılımının veri setini daha iyi modellediği söylenebilir.

Özet olarak, veri setine uyumu sınanan dört modelden yapılan ilk analiz sonucunda bilgi kriterlerine göre EG-Gamma ve EG-Weibull dağılımının veri setini daha iyi açıkladığı görülmüştür. Daha sonra söz konusu iki model için görsel olarak veri setine uyumu test edilmiştir. Grafikler değerlendirildiğinde veri setini EG-Gamma dağılımının EG-Weibull dağılımına göre daha iyi modellediği yorumuna ulaşılmıştır.

## V. TARTIŞMA VE SONUÇ

Çalışmada üstelleştirilmiş genelleştirilmiş gamma dağılımı üzerinde durulmuştur. Üstelleştirilmiş genelleştirilmiş gamma dağılımının diğer dağılımlara kıyasla avantajları çeşitli yaşam süresi verilerinin modellenmesi problemlerinde hazard fonksiyonunun her formuyla uyumlu olmasıdır. Bir diğer avantajı ise özel olarak seçilen parametre değerlerinde Weibull, üstelleştirilmiş üstel, uyarlanmış Weibull ve genelleştirilmiş Rayleigh dağılımları gibi alt dağılımları içermesidir. Ayrıca üstelleştirilmiş genelleştirilmiş gamma dağılımı kuvvet tipi bozulma oranlarını modellemede uyumlu olmasının yanı sıra üstelleştirilmiş Weibull ve uyarlanmış Weibull dağılımları gibi özel alt dağılımlarla uyum iyiliği testi için uygundur.

Çalışmanın ikinci bölümünde, üstelleştirilmiş genelleştirilmiş gamma dağılımının bazı istatistiksel ve matematiksel özellikleri verilmiştir. Üçüncü bölümde bu dağılıma ilişkin parametrik parametre tahmini yöntemi olan en çok olabilirlik tahmin yöntemi verilmiştir. Dördüncü bölümde ise, ikisi üstelleştirilmiş genelleştirilmiş dağılım sınıfının alt dağılımları diğer iki dağılım ise üstelleştirilmiş dağılım ailesinin alt dağılımları olmak üzere toplam dört dağılımın Büyük Britanya 'da meydana gelen maden kazalarına ilişkin veri setine uyumunun test edilmesi, üçüncü bölümde verilen en çok olabilirlik tahmin edicilerinin bulunması ve en iyi modelin belirlenmesi üzerinde durulmuştur. Veri setine uygulanan aşamalardan sonra en iyi modelin çalışma kapsamında test edilen dört dağılım arasında üstelleştirilmiş genelleştirilmiş gamma dağılımı olduğu yorumlanması yapılmıştır.

Sonuç olarak çalışmada üstelleştirilmiş genelleştirilmiş gamma dağılımının veri setini en iyi modelleyebildiği ancak başka çalışmalarda üstelleştirilmiş

genelleştirilmiş dağılım sınıfına ait bazı özel dağılımların ya da farklı bir dağılım sınıfının veri setini daha iyi modelleyebilme durumu mevcuttur. Söz konusu durum araştırmacıya bırakılmıştır.

## KAYNAKLAR

- [1] Gupta, R.D. & Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, 41(2), 173-188.
- [2] Gupta, R.D. & Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and Weibull. *Biometrical Journal*, 43(1), 117-130.
- [3] Nadarajah, S. & Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematicae*, 92(2), 97-111.
- [4] Eugene, N., Lee, C. & Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*, 31(4), 497-512.
- [5] Cordeiro, G.M. & Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81(7), 883– 898.
- [6] Cordeiro, G.M. & Nadarajah, S. (2011). Closed-form expressions for moments of a class of beta generalized distributions. *Brazilian Journal of Probability and Statistics*, 25(1), 14-33.
- [7] Cordeiro, G.M., Ortega, E.M. & Silva, G.O. (2011). The exponentiated generalized gamma distribution with application to lifetime data. *Journal of Statistical Computation and Simulation*, 81(7), 827-842.
- [8] Maguire, B.A. Pearson, E. & Wynn, A. (1952). The time intervals between industrial accidents. *Biometrika*, pages 168–180.
- [9] Dubey, S.D. (1970). Compound gamma, beta and F distributions. *Metrika*, 16(1), 27–31.
- [10] Zografos, K. & Balakrishnan, N. (2009). On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6(4), 344–362.
- [11] Alzaatreh, A., Lee, C. & Famoye, F. (2013). A new method for generating families of continuous distributions. *METRON*, 71(1), 63–79.
- [12] Alzaghal, A., Famoye, F. & Lee, C. (2013). Exponentiated T-X family of distributions with some applications. *International Journal of Statistics and Probability*, 2(3), 31–49.
- [13] Marshall, A. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. *Biometrika*, 84(3), 641–652.
- [14] Chahkandi, M. & Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics & Data Analysis*, 53(12), 4433–4440.
- [15] Lindley, D.V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20, 102–107.