



---

*Research article*

## The extended gamma distribution with regression model and applications

Emrah Altun<sup>1</sup>, Mustafa Ç. Korkmaz<sup>2</sup>, M. El-Morshedy<sup>3,4,\*</sup> and M. S. Eliwa<sup>4</sup>

<sup>1</sup> Department of Mathematics, Bartın University, Bartın, Turkey

<sup>2</sup> Department of Measurement and Evaluation, Artvin Çoruh University, Artvin, Turkey

<sup>3</sup> Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

<sup>4</sup> Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

\* **Correspondence:** Email: m.elmorshedy@psau.edu.sa.

**Abstract:** This paper introduces a new extension of the gamma distribution, named as a *new extended gamma* distribution, via mixture representation of  $x$ gamma and gamma distributions. The statistical properties of the proposed distribution are derived such as moment generating and characteristic functions, variance, skewness, and kurtosis measures, Lorenz curve, and mean residual life function. The maximum likelihood, parametric bootstrap, method of moments, least squares, and weighted least squares estimation methods are considered to obtain the unknown model parameters. The finite sample performance of estimation methods is discussed via a simulation study. Using the proposed distribution, we propose a new regression model for the right-skewed response variable as an alternative to the gamma regression model. Two real data sets are analyzed to convince the readers for the usefulness of the proposed model.

**Keywords:**  $x$ gamma distribution; gamma distribution; maximum likelihood; method of moments; bias-correction; randomized quantile residuals

**Mathematics Subject Classification:** 62E15

---

### 1. Introduction

In real-life problems, the analyzed data may consist of a mixed structure of more than one distribution. Especially, when data is heterogeneous structure model, using the mixture distribution will be an appropriate approach for data modeling. When the data has multimodality or unimodality, the mixture distributions are good choices to model the heterogeneous structure (Everitt and Hand [1]). However, mixture distributions have been used to model the heterogeneous data sets in many

areas such as survival, biomedical, engineering, and social sciences (see, Chen et al. [2], Erisoglu et al. [3, 4, 5]).

Mixture distributions can be obtained by both mixtures of two identical distributions and a mixture of two non-identical distributions. Recently, many studies about the mixture of two identical distributions were made such as Ahmad and Rahman [6], Jiang and Murthy [7], Sultan et al. [8], Zakerzadeh and Dolati [9], Ateya [10], Abouammoh et al. [11], El-Bassiouny et al. [12] and Karakoca et al. [13]. On the other hand, the Lindley distribution, proposed by Lindley [14], is the first distribution that comes to mind about a mixture of two non-identical distributions. Its probability density function (pdf) can be written as

$$f_L(x; \theta) = \frac{\theta^2}{1 + \theta} (1 + x) e^{-\theta x} = p f_E(x; \theta) + (1 - p) f_G(x; \theta), x > 0, \theta > 0 \quad (1.1)$$

where  $f_E(x; \theta) = \theta e^{-\theta x}$  (for  $x > 0$  and  $\theta > 0$ ) is the pdf of the exponential distribution with the scale parameter  $\theta$ ,  $f_G(x; \theta) = \theta^2 x e^{-\theta x}$  (for  $x > 0$  and  $\theta > 0$ ) is the pdf of the gamma distribution with the shape parameter 2 and scale parameter  $\theta$  and  $p = \theta / (\theta + 1)$  is the mixing proportion of distributions. Some mixture of two non-identical distributions has been also introduced in the literature, for example, the power Lindley distribution, denoted by  $PL(\alpha, \theta)$ , introduced by Ghitany et al. [15] with pdf:

$$f_{PL}(x; \alpha, \theta) = \frac{\alpha \theta^2}{\theta + 1} (1 + x^\alpha) x^{\alpha-1} e^{-\theta x^\alpha}, x, \alpha, \theta > 0,$$

the gamma Lindley distribution, denoted by  $GL(\alpha, \theta)$ , proposed by Nedjar and Zeghdoudi [16] with pdf:

$$f_{GL}(x; \alpha, \theta) = \frac{\theta^2}{\alpha(1 + \theta)} [(\alpha + \alpha\theta - \theta)x + 1] e^{-\theta x}, x, \alpha, \theta > 0,$$

and the xgamma distribution, denoted by  $XG(\theta)$ , proposed by Sen et al. [17] with pdf:

$$f_{x\text{gamma}}(x; \theta) = \frac{\theta^2}{1 + \theta} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x}, x, \theta > 0. \quad (1.2)$$

Although gamma distribution is widely used in lifetime data modeling, it is not able to model different characteristics of the data sets such as the bathtub and upside-down failure rates. To remove the drawbacks of the gamma distribution, we propose a new extended gamma (NEG) distribution as a two-component mixture of xgamma and gamma distributions with suitable mixing proportions. The proposed distribution has several advantages over the well-known distributions. For instance, it has closed-form expressions for its mean, variance and skewness, and kurtosis measures and provides very flexible hazard rate shapes for lifetime data modeling. The statistical properties of the NEG distribution are derived. The regression model defined under the NEG density is proposed to model the right-skewed dependent variable.

Other parts of the study are as follows. In Section 2, the main properties of the NEG distribution are derived. In Section 3, the parameter estimation problem of the NEG distribution is discussed comprehensively. In Section 4, a simulation study is given to compare the estimation methods for the parameters of the NEG distribution. The new regression model is defined and studied in Section 5. The real data applications of the NEG distribution are presented in Section 6. The study is ended with the concluding remarks, given in Section 7.

## 2. A new extended gamma distribution

The density of gamma distribution is

$$f_{\text{Gamma}}(x; \alpha, \theta) = \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x}, x > 0, \quad (2.1)$$

where  $\Gamma(\cdot)$  is

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx,$$

and  $\alpha > 0$  and  $\theta > 0$  are the shape and scale parameters, respectively. In the recent years, several generalization of gamma distribution are introduced such as new generalized gamma distribution by Bourguignon et al. [18], reflected shifted-truncated gamma distribution by Waymyers et al. [19], Kumaraswamy generalized gamma distribution by Pascoa et al. [20] and among others. Now, we introduce the NEG distribution with following proposition.

**Proposition 1.** *Let the random variable  $X$  follows a NEG distribution if its pdf is given by*

$$f(x; \alpha, \theta) = \frac{[\Gamma(\alpha)(2 + \theta x^2) + 4\theta^{\alpha-2} x^{\alpha-1}] \theta^2 e^{-\theta x}}{2(\theta + 3)\Gamma(\alpha)}, \quad (2.2)$$

where  $\alpha > 0$  and  $\theta > 0$  are shape and scale parameters, respectively. Hereafter, the density (2.2) is denoted as  $NEG(\alpha, \theta)$ .

*Proof.* The NEG distribution is defined as follows

$$f_{NEG}(x; \alpha, \theta) = p f_{x\text{gamma}}(x; \theta) + (1 - p) f_{\text{gamma}}(x; \alpha, \theta), \quad (2.3)$$

where  $p = (\theta + 1)/(\theta + 3)$ ,  $f_{x\text{gamma}}(x; \theta)$  and  $f_{\text{gamma}}(x; \alpha, \theta)$  are given in (1.2) and (2.1), respectively. Substituting (1.2) and (2.1) in (2.3), we have

$$\begin{aligned} f_{NEG}(x; \alpha, \theta) &= \left(\frac{\theta + 1}{\theta + 3}\right) \frac{\theta^2}{\theta + 1} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x} + \left(\frac{2}{\theta + 3}\right) \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x} \\ &= \frac{\Gamma(\alpha) \theta^2}{2(\theta + 3)\Gamma(\alpha)} (2 + \theta x^2) e^{-\theta x} + \frac{4\theta^\alpha}{2(\theta + 3)\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x} \\ &= \frac{\Gamma(\alpha) \theta^2 (2 + \theta x^2) e^{-\theta x} + 4\theta^\alpha x^{\alpha-1} e^{-\theta x}}{2(\theta + 3)\Gamma(\alpha)} \\ &= \frac{[\Gamma(\alpha)(2 + \theta x^2) + 4\theta^{\alpha-2} x^{\alpha-1}] \theta^2 e^{-\theta x}}{2(\theta + 3)\Gamma(\alpha)} \end{aligned} \quad (2.4)$$

The proof is completed. □

The NEG distribution is a mixture distribution of  $x\text{gamma}(\theta)$  and  $\text{Gamma}(\alpha, \theta)$  distributions with mixing proportion  $p = (\theta + 1)/(\theta + 3)$ . Therefore, statistical properties of NEG distribution can be

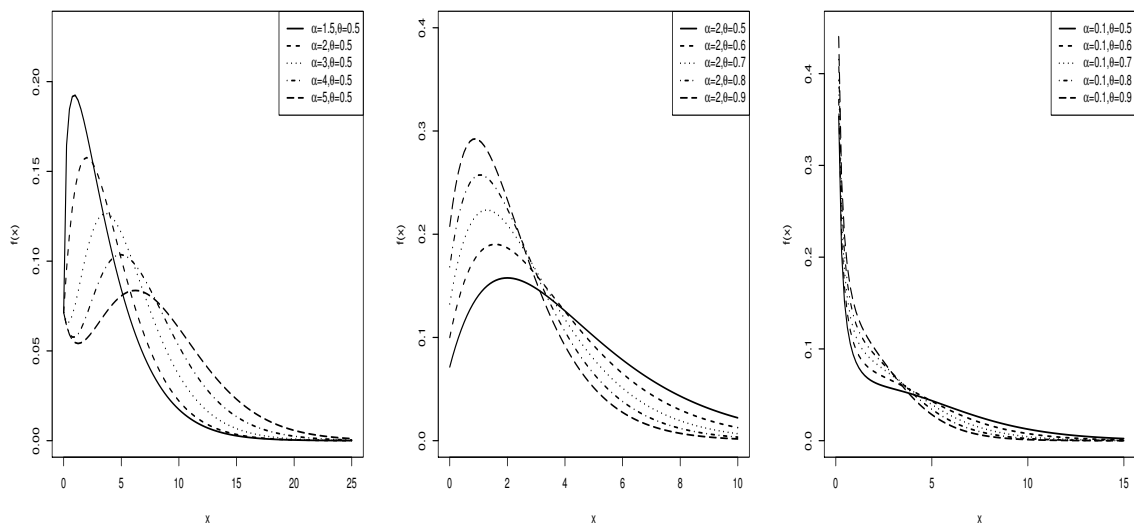
derived using the properties of mixture distributions. The corresponding cumulative distribution function (cdf) to (2.2) is

$$F(x) = \frac{\Gamma(\alpha) \left[ 2 - e^{-\theta x} (2 + 2\theta x + \theta^2 x^2) \right] + 2\theta \Gamma(\alpha) (1 - e^{-\theta x}) + 4 [\Gamma(\alpha) - \gamma(\alpha, \theta x)]}{2\Gamma(\alpha) (\theta + 3)}, \quad (2.5)$$

where  $\gamma(\alpha, \theta x)$  is

$$\gamma(\alpha, \theta x) = \int_{\theta x}^{\infty} t^{\alpha-1} e^{-\theta t} dt.$$

which is called as incomplete gamma function. The possible pdf shapes of NEG distribution are displayed in Figure 1. These figures reveal the flexibility of NEG distribution which can be used to model right-skewed and bi-modal data sets.



**Figure 1.** The pdf shapes of the NEG distribution.

### 2.1. Survival and mean residual life functions

The main reliability characteristics of NEG distribution are discussed via survival and mean residual life (MRL) functions. The survival function (sf) represents the probability that a patient, device, or any object of interest survive after a specified time point. Therefore, sf has important application fields in reliability and survival analysis. The sf is defined as  $S(t) = 1 - F(t)$  which is given below for the NEG distribution

$$S(t) = 1 - \frac{\Gamma(\alpha) \left[ 2 - e^{-\theta t} (2 + 2\theta t + \theta^2 t^2) \right] + 2\theta \Gamma(\alpha) (1 - e^{-\theta t}) + 4 [\Gamma(\alpha) - \gamma(\alpha, \theta t)]}{2\Gamma(\alpha) (\theta + 3)}. \quad (2.6)$$

The mean residual life function of NEG distribution is

$$MRL(t) = \frac{1}{S(t)} \int_t^{\infty} xf(x) dx - t, \quad (2.7)$$

where  $S(t)$  is given in (2.6) and

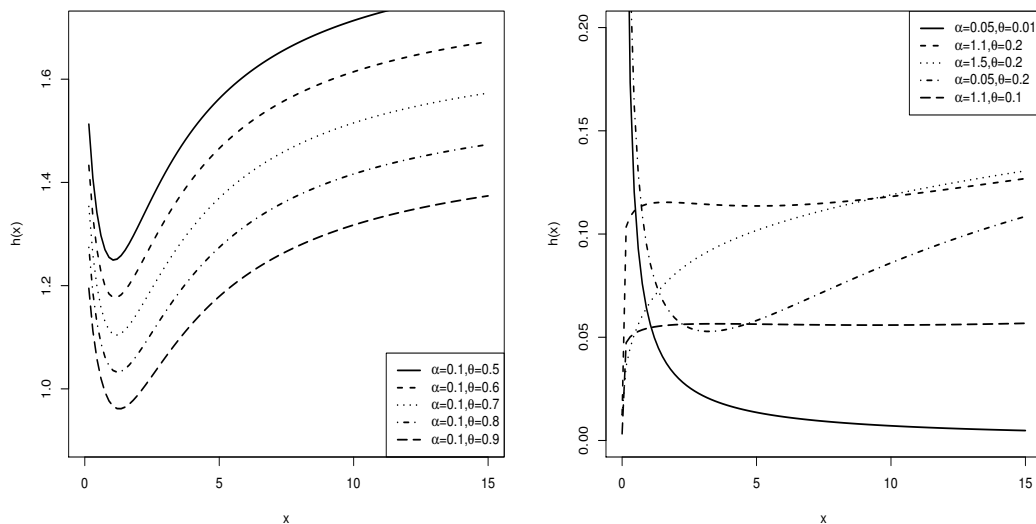
$$\int_t^{\infty} xf(x) dx = \frac{e^{-t\theta}(t\theta + 1)}{\theta + 3} + \frac{2\gamma(\alpha + 1, t\theta)}{\theta\Gamma(\alpha)(\theta + 3)} + \frac{e^{-t\theta}(t^3\theta^3 + 3t^2\theta^2 + 6t\theta + 6)}{2\theta(\theta + 3)} \quad (2.8)$$

## 2.2. Hazard rate function

The hazard rate function (hrf) is defined as  $h(t) = f(t)/S(t)$  which is given below for the NEG distribution

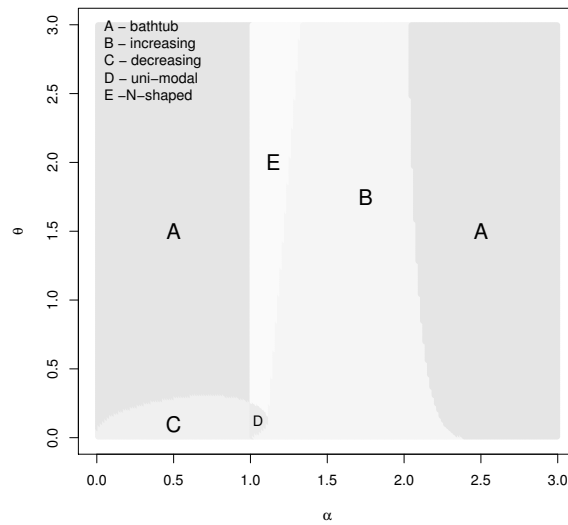
$$h(t) = \frac{4\theta^\alpha t^\alpha + 2\theta^2 t\Gamma(\alpha) + \theta^3 t^3 \Gamma(\alpha)}{t(2\Gamma(\alpha) + 4e^{\theta t}\gamma(\alpha, \theta t) + 2\theta\Gamma(\alpha) + 2\theta t\Gamma(\alpha) + \theta^2 t^2 \Gamma(\alpha))}. \quad (2.9)$$

The possible hrf shapes of NEG distribution are displayed in Figure 2. These plots reveal that NEG distribution is a very attractive distribution to model the different characteristics of the lifetime data sets with increasing, decreasing, bathtub, and upside-down shapes.



**Figure 2.** The hrf shapes of the NEG distribution.

Further, Figure 3 indicates the possible shape regions of hrf of NEG distribution. They are bathtub, increasing, decreasing, uni-modal (upside-down) and N-shaped (increasing-decreasing-increasing) shapes.



**Figure 3.** The possible hrf shapes regions of NEG distribution for parameters values.

### 2.3. Lorenz curve

The Lorenz curve (LC), introduced by Lorenz [21], is widely used in economics to represent the distribution of income and inequality of wealth distribution. The LC of  $X$  random variable is given by

$$L(t) = \frac{\int_0^t x f(x) dx}{\int_0^{\infty} t f(t) dt} \quad (2.10)$$

which  $f(\cdot)$  is the pdf of  $X$  random variable. So, substituting pdf of NEG distribution, (2.2), in (2.10), the LC of NEG distribution is

$$L(t) = \frac{\Gamma(\alpha) \left[ 6 - e^{-t\theta} (6 + 6\theta t + 3\theta^2 t^2 + \theta^3 t^3) + 2\theta - 2e^{-t\theta} \theta (1 + \theta t) \right] + 4(\Gamma(\alpha + 1) - \gamma(\alpha + 1, \theta t))}{2\Gamma(\alpha) (2\alpha + \theta + 3)} \quad (2.11)$$

It is obvious that (2.11) can be used in modeling the distribution of income by economists.

### 2.4. Moments and related measures

Here, the mean, variance and related measures of the NEG distribution are derived.

**Proposition 2.** The  $r$ th raw moments about the origin of  $X$  are

$$E(X^r) = \frac{2\theta\Gamma(\alpha)\Gamma(r+1) + \Gamma(\alpha)\Gamma(r+3) + 4\Gamma(r+\alpha)}{2(\theta+3)\theta^r\Gamma(\alpha)}. \quad (2.12)$$

*Proof.*

$$\begin{aligned}
 E(X^r) &= \int_0^{\infty} x^r f(x) dx \\
 &= \int_0^{\infty} x^r \frac{\theta^2}{2(\theta+3)} (2 + \theta x^2) e^{-\theta x} dx + \int_0^{\infty} x^r \left(\frac{2}{\theta+3}\right) \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x} dx \\
 &= \int_0^{\infty} \frac{\theta^2}{2(\theta+3)} (2x^r + \theta x^{r+2}) e^{-\theta x} dx + \int_0^{\infty} \left(\frac{2}{\theta+3}\right) \frac{\theta^\alpha}{\Gamma(\alpha)} x^{r+\alpha-1} e^{-\theta x} dx \\
 &= \frac{2\theta\Gamma(r+1) + \Gamma(r+3)}{2\theta^r(\theta+3)} + \frac{2\Gamma(\alpha+r)}{\Gamma(\alpha)\theta^r(\theta+3)} \\
 &= \frac{2\theta\Gamma(\alpha)\Gamma(r+1) + \Gamma(\alpha)\Gamma(r+3) + 4\Gamma(\alpha+r)}{2(\theta+3)\theta^r\Gamma(\alpha)} \tag{2.13}
 \end{aligned}$$

The proof is completed.  $\square$

Using (2.12), the first (mean) and second raw moments of  $X$  are given, respectively, by

$$E(X) = \mu = \frac{2\alpha + \theta + 3}{\theta(\theta + 3)}, \tag{2.14}$$

$$E(X^2) = \frac{2(\alpha^2 + \alpha + \theta + 6)}{\theta^2(\theta + 3)}. \tag{2.15}$$

The variance of  $X$  is

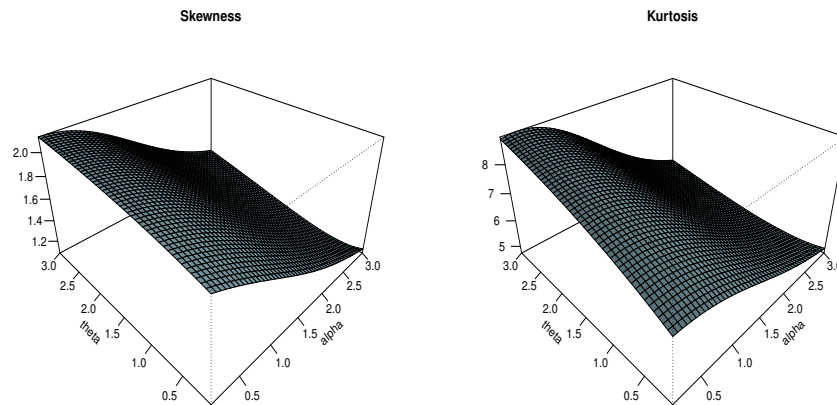
$$\text{Var}(X) = \sigma^2 = \frac{2\alpha^2(\theta + 1) - 2\alpha(\theta + 3) + (\theta + 3)(\theta + 9)}{\theta^2(\theta + 3)^2}. \tag{2.16}$$

Using the well-known relations, the skewness and kurtosis measures of NEG distribution are given, respectively, by

$$S = \frac{2(\alpha^3\theta^2 - \alpha^3 + 6\alpha^2\theta + 18\alpha^2 - \alpha\theta^2 - 24\alpha\theta - 63\alpha + \theta^3 + 21\theta^2 + 99\theta + 135)}{\theta^3(\theta + 3)^3\sigma^3} \tag{2.17}$$

$$K = \left( \begin{array}{l} \frac{2\alpha^4 + 12\alpha^3 + 22\alpha^2 + 12\alpha + 24\theta + 360}{\theta^4(\theta+3)} \\ - \frac{3(2\alpha+\theta+3)^4}{\theta^4(\theta+3)^4} - \frac{4(2\alpha+\theta+3)(2\alpha^3+6\alpha^2+4\alpha+6\theta+60)}{\theta^4(\theta+3)^2} \\ + \frac{6(2\alpha+\theta+3)^2(2\alpha^2+2\alpha+2\theta+12)}{\theta^4(\theta+3)^3} \end{array} \right) \sigma^{-4} \tag{2.18}$$

The skewness and kurtosis of the NEG distribution are displayed in Figure 4 for different values of the parameters  $\alpha$  and  $\theta$ . From these figures, we conclude the following results: (i) when  $\alpha$  increases, the skewness, and kurtosis both decrease; (ii) when  $\theta$  increases, the skewness, and kurtosis both increase.



**Figure 4.** The skewness and kurtosis measures of the NEG distribution.

**Proposition 3.** The moment generating function of  $X$  is

$$M(t) = \frac{\theta^2}{(\theta - t)(\theta + 3)} + \frac{\theta^3}{(\theta - t)^3(\theta + 3)} + \frac{2\theta^\alpha}{(\theta - t)^\alpha(\theta + 3)}. \quad (2.19)$$

*Proof.*

$$\begin{aligned} M(t) &= \int_0^{\infty} e^{tx} f(x) dx. \\ &= \int_0^{\infty} e^{tx} \frac{\theta^2}{(\theta + 3)} e^{-\theta x} dx + \int_0^{\infty} e^{tx} \frac{1}{2(\theta + 3)} \theta^3 x^2 e^{-\theta x} dx + \int_0^{\infty} e^{tx} \frac{2}{(\theta + 3)\Gamma(\alpha)} \theta^\alpha x^{\alpha-1} e^{-\theta x} dx. \\ &= \frac{\theta^2}{(\theta - t)(\theta + 3)} + \frac{\theta^3}{(\theta - t)^3(\theta + 3)} + \frac{2\theta^\alpha}{(\theta - t)^\alpha(\theta + 3)}. \end{aligned}$$

The proof is completed. □

Substituting  $t$  by  $it$  in (2.19), the characteristic function of NEG distribution is

$$\varphi(t) = \frac{\theta^2}{(\theta - it)(\theta + 3)} + \frac{\theta^3}{(\theta - it)^3(\theta + 3)} + \frac{2\theta^\alpha}{(\theta - it)^\alpha(\theta + 3)}. \quad (2.20)$$

### 2.5. Generating random variables

Since the NEG distribution is a mixture distribution of  $x\text{gamma}(\theta)$  and  $\text{Gamma}(\alpha, \theta)$  distributions with mixing proportion  $p = (\theta + 1)/(\theta + 3)$ , the below algorithm can be used for generating data from the NEG distribution.

1. Set the sample size  $n$  and the parameter values  $\alpha$  and  $\theta$ .
2. Generate  $U_i \sim \text{uniform}(0, 1), i = 1, 2, \dots, n$ .



3. Generate  $V_i \sim \text{xgamma}(\theta), i = 1, 2, \dots, n$ .
4. Generate  $W_i \sim \text{Gamma}(\alpha, \theta), i = 1, 2, \dots, n$ .
5. If  $U_i \leq (\theta + 1)/(\theta + 3)$ ,  $X_i = V_i$ , otherwise,  $X_i = W_i$ .

### 3. Estimation

In this section, three estimation methods are used to obtain unknown parameters of NEG distribution. These are maximum likelihood (ML), parametric bootstrap, and method of moments (MM) estimation methods. The rest of this section is devoted to inference on these estimation methods.

#### 3.1. Maximum likelihood

The log-likelihood function of NEG  $(\alpha, \theta)$  is

$$\ell(\Phi) = 2n \ln(\theta) - n\theta\bar{x} + \sum_{i=1}^n \ln \left[ \Gamma(\alpha) (2 + \theta x_i^2) + 4\theta^{\alpha-2} x_i^{\alpha-1} \right] - n \ln [2(\theta + 3)\Gamma(\alpha)], \quad (3.1)$$

Let  $x_1, x_2, \dots, x_n$  be a random sample following the distribution, NEG  $(\alpha, \theta)$  and where  $\Phi = (\alpha, \theta)$  is the unknown parameter vector. Taking the partial derivatives of (3.1) with respect to  $\alpha$  and  $\theta$ , the score vector components are obtained as

$$\frac{\partial \ell}{\partial \theta} = \frac{2n}{\theta} - \frac{n}{\theta + 3} + \sum_{i=1}^n \frac{4(\alpha - 2)\theta^\alpha x_i^\alpha + \theta^3 \Gamma(\alpha)}{\theta [4\theta^\alpha x_i^\alpha + \theta^2 x_i \Gamma(\alpha) (\theta x_i^2 + 2)]}, \quad (3.2)$$

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^n \frac{\Gamma(\alpha) \psi(\alpha) (\theta x_i^2 + 2) + 4\theta^{\alpha-2} x_i^{\alpha-1} \log(\theta) + 4\theta^{\alpha-2} x_i^{\alpha-1} \log(x_i)}{\Gamma(\alpha) (\theta x_i^2 + 2) + 4\theta^{\alpha-2} x_i^{\alpha-1}} - n\psi(\alpha), \quad (3.3)$$

where  $\psi(\alpha)$  is the digamma function which is defined as  $\psi(\alpha) = \partial \ln(\Gamma(\alpha))/\partial \alpha$ . The ML estimators of  $(\alpha, \theta)$ , say  $(\hat{\alpha}, \hat{\theta})$ , can be obtained by simultaneous solutions of (3.2) and (3.3) against zero. However, it is not possible to obtain explicit forms of ML estimators of  $\alpha$  and  $\theta$  because of the non-linear functions in (3.2) and (3.3). Therefore, the direct maximization of (3.1) is needed to obtain the ML estimators of NEG distribution. Here, the **optim** function of R software is used to minimize minus of (3.1) which is equivalent to maximization of (3.1). The observed information matrix is given by

$$I_F(\Phi) = - \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\theta} \\ I_{\theta\alpha} & I_{\theta\theta} \end{bmatrix}_{(\hat{\alpha}, \hat{\theta})}$$

where  $I_{\alpha\alpha} = \partial^2 \ell / \partial \alpha^2$ ,  $I_{\alpha\theta} = I_{\theta\alpha} = \partial^2 \ell / \partial \alpha \partial \theta$ ,  $I_{\theta\theta} = \partial^2 \ell / \partial \theta^2$ . The squared values of the diagonal elements of the inverse of above matrix give the standard errors of the estimated parameters. So, we have the following quantities for the confidence intervals of the parameters,  $\alpha$  and  $\theta$ , respectively,

$$\hat{\alpha} \pm z_{p/2} \sqrt{\text{Var}(\hat{\alpha})}, \quad \hat{\theta} \pm z_{p/2} \sqrt{\text{Var}(\hat{\theta})}$$

where  $\text{Var}(\hat{\alpha})$  and  $\text{Var}(\hat{\theta})$  are the square root of the diagonal elements of  $I_F^{-1}(\Phi)$  and  $z_{p/2}$  is easily obtained by the quantile function of  $N(0, 1)$ .

### 3.2. Parametric bootstrap

Parametric bootstrap method, introduced by Efron [22] can be used to obtain bias-corrected ML estimators of NEG distribution. The estimated bias of  $\hat{\Phi}$  is

$$\hat{\mathbf{B}}(\hat{\Phi}) = \frac{1}{B} \sum_{i=1}^B \hat{\Phi}_{(j)} - \hat{\Phi}, \quad (3.4)$$

where  $\hat{\Phi}_j$  is the MLE of  $\hat{\Phi}$  obtained from the  $j$ th bootstrap sample generated by assuming  $\hat{\Phi}$  is true and  $B$  is the bootstrap replications (see, Mazucheli et al. [23, 24]). So, the bootstrap bias-corrected (BBC) estimator of  $\hat{\Phi}$  is

$$\hat{\Phi}_{BBC} = 2\hat{\Phi} - \frac{1}{B} \sum_{i=1}^B \hat{\Phi}_{(j)}. \quad (3.5)$$

### 3.3. Method of moments

The MM is the simple and effective estimation method generally for large sample sizes. The idea of MM estimators is based on equating the theoretical moments to empirical ones. So, equating the first two theoretical moments of NEG distribution to their empirical counterparts, we have

$$\frac{2\alpha + \theta + 3}{\theta(\theta + 3)} = m_1, \quad (3.6)$$

$$\frac{2(\alpha^2 + \alpha + \theta + 6)}{\theta^2(\theta + 3)} = m_2, \quad (3.7)$$

where  $m_1 = \frac{1}{n} \sum_{i=1}^n x_i$  and  $m_2 = \frac{1}{n} \sum_{i=1}^n x_i^2$  are the first and second sample moments. The MM estimators of  $\alpha$  and  $\theta$ , say  $\hat{\alpha}_{MM}$  and  $\hat{\theta}_{MM}$  are simultaneous solutions of (3.6) and (3.7).

### 3.4. Least squares

Let  $x_{1:n}, x_{2:n}, x_{3:n}, \dots, x_{n:n}$  be an ordered sample of size  $n$  from a probability distribution with cdf  $F(\cdot)$ . The mean and variance of  $F(X_{j:n})$  are, respectively,

$$E[F(X_{j:n})] = \frac{j}{n+1}, \quad (3.8)$$

$$\text{Var}[F(X_{j:n})] = \frac{j(n-j+1)}{(n+1)^2(n+2)}. \quad (3.9)$$

The least squares estimators (LSEs) of  $\alpha$  and  $\theta$  are procured by means of the minimization of the following equation.

$$\sum_{j=1}^n \left( F(X_{j:n}) - \frac{j}{n+1} \right)^2, \quad (3.10)$$

where  $F(Y_{j:n})$  is in (2.5) and replacing it in (3.10), we have

$$\sum_{j=1}^n \left[ \frac{\Gamma(\alpha) \left[ 2 - e^{-\theta x_{j:n}} (2 + 2\theta x_{j:n} + \theta^2 x_{j:n}^2) \right] + 2\theta \Gamma(\alpha) (1 - e^{-\theta x_{j:n}}) + 4 \left[ \Gamma(\alpha) - \gamma(\alpha, \theta x_{j:n}) \right]}{2\Gamma(\alpha) (\theta + 3)} - \frac{j}{n+1} \right]^2. \quad (3.11)$$

The detailed information on the LSE can be found in Ding [25] and Ding et al. [26].

### 3.5. Weighted least squares

The weighted LSEs (WLSEs) of  $\alpha$  and  $\theta$  are, respectively,

$$\sum_{j=1}^n w_j \left[ F(X_{j:n}) - \frac{j}{n+1} \right]^2, \quad (3.12)$$

where

$$w_j = \frac{1}{\text{Var}[F(X_{j:n})]} = \frac{(n+1)^2 (n+2)}{j(n-j+1)}. \quad (3.13)$$

Replacing eqrefwj and (2.5) with  $w_j$  and  $F(X_{j:n})$ , we have

$$\sum_{j=1}^n \frac{(n+1)^2 (n+2)}{j(n-j+1)} \left[ \frac{\Gamma(\alpha) \left[ 2 - e^{-\theta x_{j:n}} (2 + 2\theta x_{j:n} + \theta^2 x_{j:n}^2) \right] + 2\theta \Gamma(\alpha) (1 - e^{-\theta x_{j:n}}) + 4 \left[ \Gamma(\alpha) - \gamma(\alpha, \theta x_{j:n}) \right]}{2\Gamma(\alpha) (\theta + 3)} - \frac{j}{n+1} \right]^2. \quad (3.14)$$

The minimization of (3.14) give the WLSEs of the parameters,  $\alpha$  and  $\theta$ .

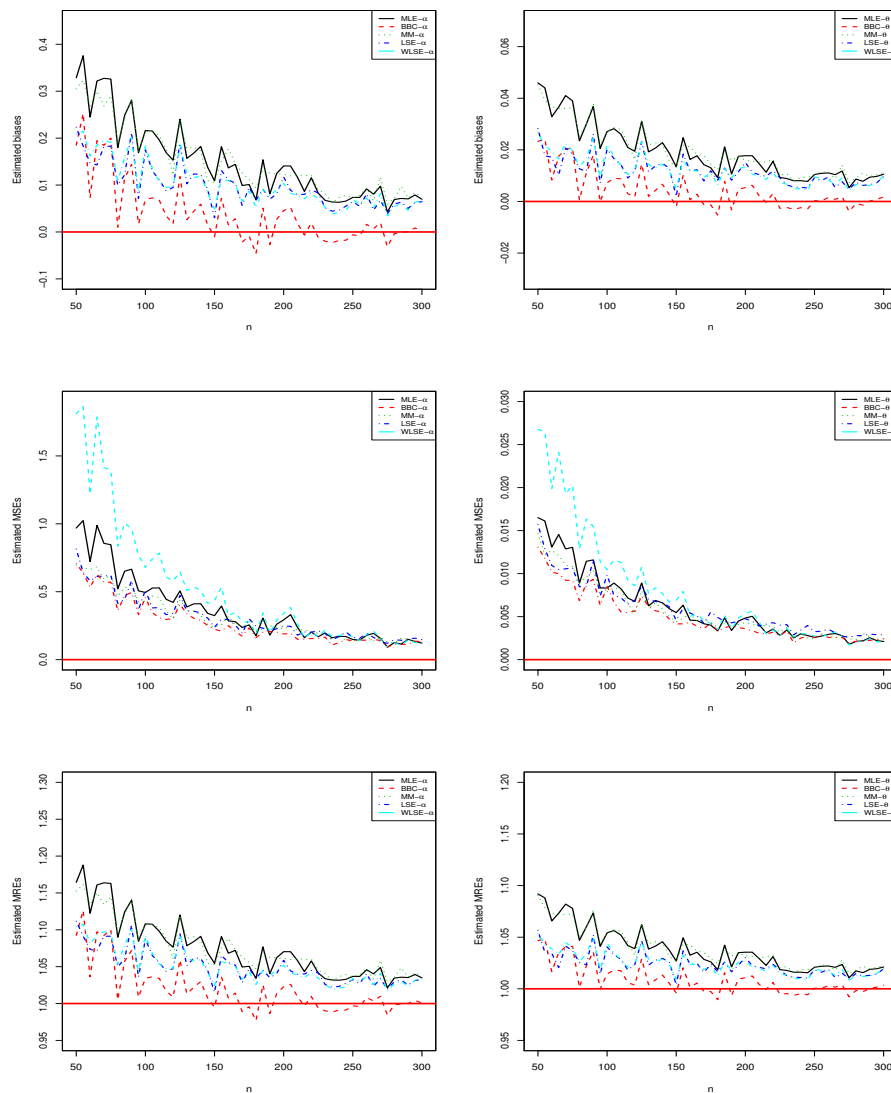
## 4. Simulation studies

The simulation studies are widely used to compare the finite sample performance of estimation methods (see, Zaka et al. [27] and Zaidi et al. [28]). Now, we give a simulation study to see the efficiencies of the presented estimation methods, in Section 3. The bootstrap replication  $B$  is chosen 1,000. The below steps are implemented.

1. Set the parameters  $\alpha$ ,  $\theta$  and the number of simulation replications,  $N$ ,
2. Generate random variables from  $\text{NEG}(\alpha, \theta)$ ,
3. Using the generated samples in step 2, estimated the parameters of  $\text{NEG}(\alpha, \theta)$  by means of ML, BBC, LS and WLS estimation method
4. Estimate the biases, mean square errors (MSEs) and mean relative errors (MREs) for each parameters and estimation methods.
5. Repeat steps 2–4  $N$  times.

#### 4.1. Scenario I

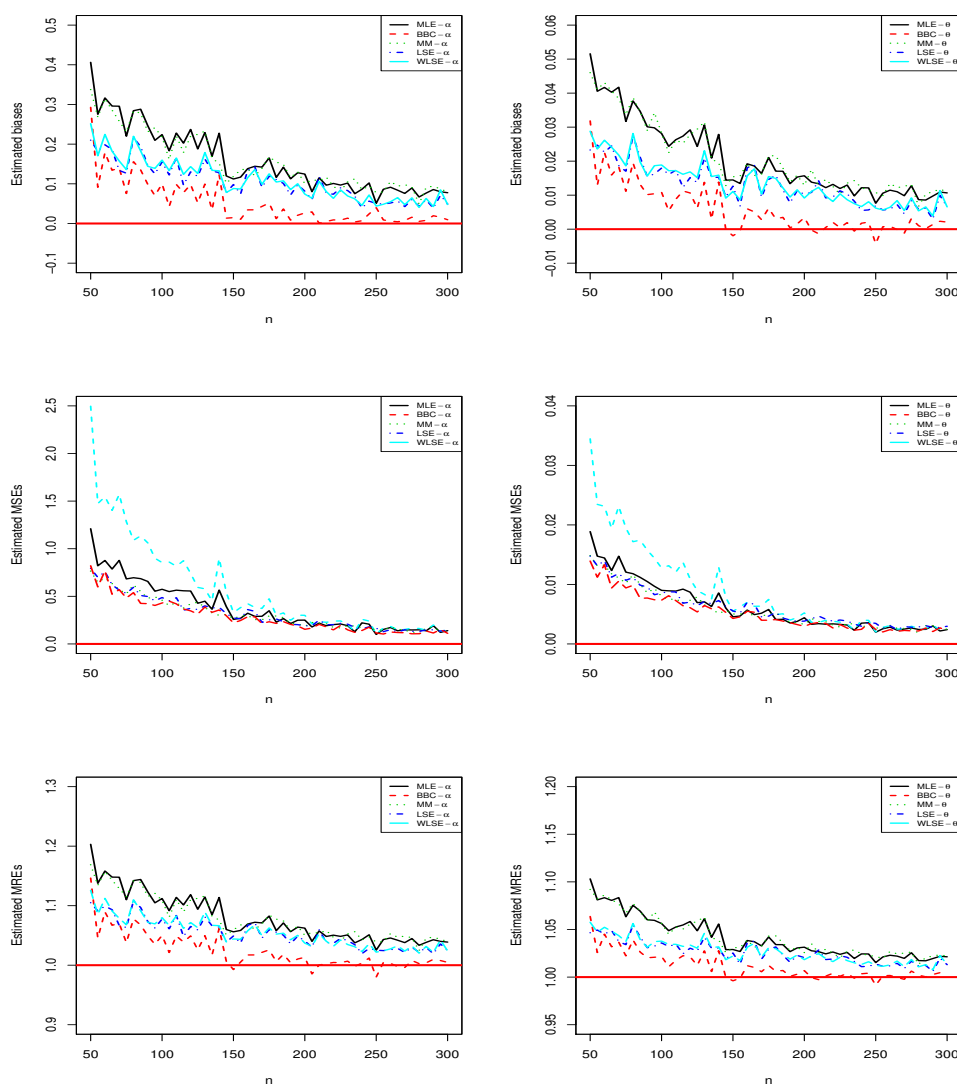
In the first simulation scenario, the following settings are used:  $N = 10,000$ ,  $\alpha = 2$  and  $\theta = 0.5$ . The required formulations of biases, MSEs, and MREs can be found in Altun et al. [29]. We know that when  $n \rightarrow \infty$ , the estimated biases, and MSEs are near the zero and MREs are near the one. The statistical software, R, is used. The simulation results are graphically summarized in Figure 5. As seen from these plots, when the sample size increases, the estimated biases, and MSEs are near zero for all estimation methods. As expected, the estimated MSEs are also near the one for all estimation methods. However, we suggest the use of BBC estimation method for small samples since it approaches nominal values of the bias, MSE, and MRE faster than the other four estimation methods. One can obtain similar results for different simulation design.



**Figure 5.** The simulation results of the NEG distribution for the first scenario.

## 4.2. Scenario II

In the second scenario, the parameter values of the NEG distribution is changed. The used parameter values are  $\alpha = 0.5$  and  $\theta = 3$ . The simulation results are summarized in Figure 6. The results are very similar to the results of the first scenario. All estimation methods work well based on the estimated values of the biases, MSEs, and MREs. However, the BBC method is better for small sample sizes.



**Figure 6.** The simulation results of the NEG distribution for the second scenario.

## 5. NEG regression model

The gamma regression model, proposed by Cepeda-Cuervo [30] and Cepeda and Gamerman [31], is widely used to model the right-skewed response variable (see, Bateson [32]). Recently, Altun [33, 34] studied several models such as Lomax and weighted-exponential regression models to provide more accurate modeling of the skewed response variables. Let  $Y \sim \text{Gamma}(\alpha, \theta)$ . Using the transformation

$\theta = \alpha/\mu$  in (2.1), we have

$$f(y; \alpha, \mu) = \frac{1}{y\Gamma(\alpha)} \left(\frac{\alpha y}{\mu}\right)^\alpha e^{-\alpha y/\mu}, \quad y > 0, \quad (5.1)$$

where  $\mu > 0$ ,  $\alpha > 0$  and  $E(Y) = \mu$ . Now, using the appropriate link function, we can link the covariates to mean of the random variable  $Y$ . In classical gamma regression model (with constant shape parameter), the mean structure is given by

$$g(\mu_i) = \mathbf{x}_i^T \boldsymbol{\beta}, \quad (5.2)$$

where  $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^T$  is the vector of unknown regression parameters ( $k < n$ ) and  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ik})^T$  is the vector of covariates. The link function,  $g(\cdot) : (0, \infty) \mapsto \mathfrak{R}$  is the increasing and twice differentiable function (see, McCullagh and Nelder [35]).

Now, as in gamma regression model, we propose an alternative regression model to the gamma regression model by means of following proposition.

**Proposition 4.** Let  $\theta = \left\{ \sqrt{8\alpha\mu + 9\mu^2 + 6\mu + 1} - 3\mu + 1 \right\} (2\mu)^{-1}$  be in (2.2). Under this re-parametrization, the pdf of NEG is

$$f(y; \alpha, \theta) = \left( \frac{\sqrt{8\alpha\mu + 9\mu^2 + 6\mu + 1} - 3\mu + 1}{2\mu} \right)^2 \left[ \begin{array}{l} \Gamma(\alpha) \left( 2 + \frac{\sqrt{8\alpha\mu + 9\mu^2 + 6\mu + 1} - 3\mu + 1}{2\mu} y^2 \right) \\ + 4 \left( \frac{\sqrt{8\alpha\mu + 9\mu^2 + 6\mu + 1} - 3\mu + 1}{2\mu} \right)^{\alpha-2} y^{\alpha-1} \end{array} \right] \times \left[ 2 \left( \frac{\sqrt{8\alpha\mu + 9\mu^2 + 6\mu + 1} - 3\mu + 1}{2\mu} + 3 \right) \Gamma(\alpha) \right]^{-1} \exp \left[ \left( \sqrt{8\alpha\mu + 9\mu^2 + 6\mu + 1} - 3\mu + 1 \right) (2\mu)^{-1} y \right] \quad (5.3)$$

where  $\mu > 0$ ,  $\alpha > 0$  and  $E(Y) = \mu$ .

The reason of this re-parametrization is to make the mean equation of NEG simple for linking the covariates to mean of NEG random variable. The log-link function is used to link the covariates to mean of the random variable  $Y$ , as follows

$$\ln(\mu_i) = \mathbf{x}_i^T \boldsymbol{\beta}. \quad (5.4)$$

### 5.1. Estimation of the parameters of NEG regression model

We use the ML estimation method to obtain the unknown parameter of NEG regression model. The log-likelihood function of NEG regression model is

$$\begin{aligned} \ell(\boldsymbol{\tau}) = & 2 \sum_{i=1}^n \ln \left( \frac{\sqrt{8\alpha\mu_i + 9\mu_i^2 + 6\mu_i + 1} - 3\mu_i + 1}{2\mu_i} \right) \\ & + \sum_{i=1}^n \ln \left[ \Gamma(\alpha) \left( 2 + \frac{\sqrt{8\alpha\mu_i + 9\mu_i^2 + 6\mu_i + 1} - 3\mu_i + 1}{2\mu_i} y_i^2 \right) + 4 \left( \frac{\sqrt{8\alpha\mu_i + 9\mu_i^2 + 6\mu_i + 1} - 3\mu_i + 1}{2\mu_i} \right)^{\alpha-2} y_i^{\alpha-1} \right], \\ & - \sum_{i=1}^n \ln \left[ 2 \left( \frac{\sqrt{8\alpha\mu_i + 9\mu_i^2 + 6\mu_i + 1} - 3\mu_i + 1}{2\mu_i} + 3 \right) \Gamma(\alpha) \right] \\ & + \sum_{i=1}^n \left[ \left( \sqrt{8\alpha\mu_i + 9\mu_i^2 + 6\mu_i + 1} - 3\mu_i + 1 \right) (2\mu_i)^{-1} y_i \right] \end{aligned} \quad (5.5)$$

where  $\boldsymbol{\tau} = (\boldsymbol{\beta}, \alpha)$ ,  $\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$ . The ML estimators of  $\boldsymbol{\tau}$ , say  $\hat{\boldsymbol{\tau}}$  is obtained by direct maximization of 5.5. To do this, **optim** function of R software is used. The well-known property of the MLE method

is that the asymptotic distribution of  $\widehat{\boldsymbol{\tau}} - \boldsymbol{\tau}$  is  $N_{k+2}(0, K(\boldsymbol{\tau})^{-1})$ . Here,  $K(\boldsymbol{\tau})$  represents the expected information matrix which is approximated by observed information matrix. We calculate the observed information matrix evaluated at  $\widehat{\boldsymbol{\tau}}$  with the dimension  $(k + 2) \times (k + 2)$  to obtain the standard errors of the estimated parameters. It can be obtained by **hessian** function of **R** software.

## 5.2. Residuals analysis

The randomized quantile residuals (rqrs), introduced by Dunn and Smyth [36] is used to check the model accuracy on the fitted data set. The rqrs are defined by

$$\hat{r}_i = \Phi^{-1}(\hat{u}_i), \quad (5.6)$$

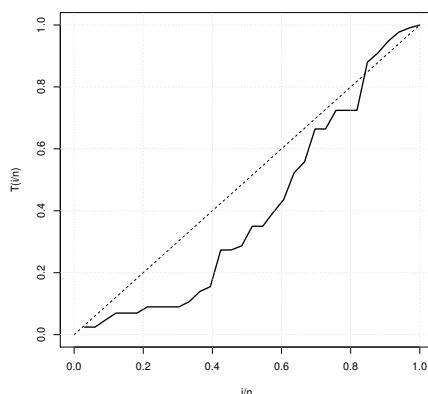
where  $\hat{u}_i = F(y_i | \hat{\boldsymbol{\beta}}, \alpha)$ . The rqrs are distributed  $N(0, 1)$  if the fitted model is statistically valid.

## 6. Empirical studies

In this section, NEG distribution is compared with competitive models to demonstrate its performance in real data modeling. Two real data sets are analyzed. The first application deals with the univariate data modeling. The second one is done for regression modeling.

### 6.1. Myelogeneous leukaemia data

We compare the NEG distribution with several competitive models such as Lindley, xgamma, gamma, power Lindley and generalized Lindley distribution. The data represents the survival times of the patients having the acute Myelogeneous Leukaemia (see, Feigl and Zelen [37]). The data set was recently studied by Mead [38]. The selection of the best model for the data can be done with model selection criteria. In this study, we use three goodness-of-fit tests. These are Kolmogorov-Smirnov ( $KS$ ), Cramer-von-Mises, ( $W^*$ ) and Anderson-Darling ( $A^*$ ). Additionally, two information criteria, Akaike Information Criteria (AIC) and Bayesian information criterion (BIC), is used to select the best model for the models passing the first step. The first step is the goodness-of-fit test. If the p-value of K-S test is larger than the critical value, 0.05, it means that the distribution provides sufficient fit. In the second step, we use AIC, BIC,  $A^*$  and  $W^*$  statistics to select the model. In this case, the distribution with the lowest values of the AIC and BIC statistics is the best model for the data. Before starting the data analysis, it is very useful to have information about the empirical behavior of the hrf. The total time test (TTT) plot is used for this goal (see, Aarset [39]). According to the shape of the TTT plot, we decide the empirical shape of the hrf. If the TTT plot has a concave shape, the hrf is increasing, otherwise, the hrf is decreasing. If the TTT plot has a convex-concave shape, the hrf is a bathtub. The TTT plot of the used data is displayed in Figure 7 which indicates that the empirical hrf is the bathtub.



**Figure 7.** TTT plot for the data set.

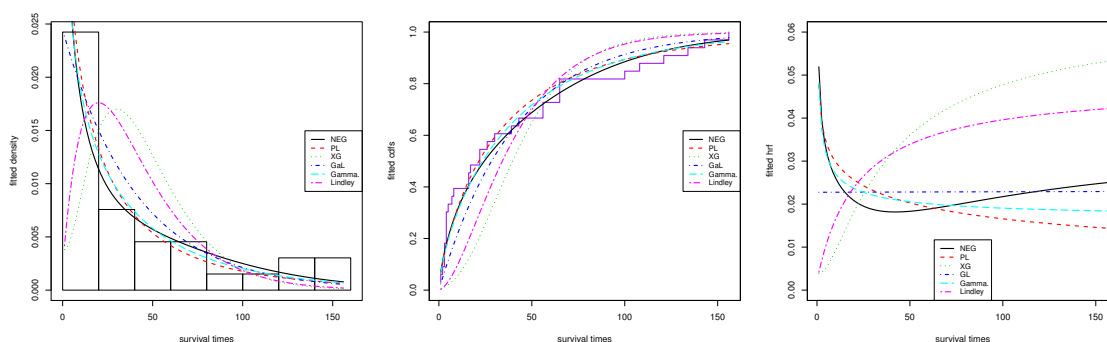
Table 1 contains the results for the fitted distributions. As mentioned above, we decide the best-fitted model in two steps. Based on the p-values of the KS test, NEG, power Lindley, generalized Lindley, and gamma distribution provide adequate fits. So, these distributions are passed the first step. Now, the distribution with the lowest values of the AIC, BIC,  $A^*$  and  $W^*$  statistics indicates the best model. According to Table 1, the NEG distribution has the lowest values of these statistics. Therefore, the proposed distribution is the best choice for the data used.

The fitted densities, empirical cdfs, and hrf plots of all model are sketched in Figure 8. Hence, we observe that the NEG fitting and successfully captures the empirical shape of the data set. Further, all estimated hrfs have not fitted as bathtub shaped except the NEG model. Hence, only the NEG model deals with Figure 7 for the data set.

**Table 1.** The results of the fitted distributions.

Model	$\hat{\alpha}$	$\hat{\theta}$	$-\hat{\ell}$	AIC	BIC	$A^*$	$W^*$	KS
NEG	0.6911 (0.1416)	0.0351 (0.0064)	152.7853	309.5706	312.5637	0.5783	0.0848	0.1277 [0.6545]
Power Lindley	0.5809 (0.0676)	0.2477 (0.0674)	153.9968	311.9935	314.9865	0.6578	0.0963	0.1344 [0.5895]
Xgamma		0.0674 (0.00069)	176.3484	354.6967	356.1932	13.5500	1.2762	0.3545 [0.0005]
Generalized Lindley	0.0239 (0.0165)	0.0244 (0.0093)	155.4502	314.9003	317.8934	2.3067	0.3250	0.2182 [0.0863]
Gamma	0.6874 (0.1440)	0.0168 (0.0049)	153.6737	311.3473	314.3404	0.6803	0.1013	0.1390 [0.5468]
Lindley		0.0478 (0.0059)	168.8337	339.6674	341.1639	9.5929	0.8975	0.3250 [0.0018]





**Figure 8.** Estimated densities (left), cdfs (middle) and hrfs (right).

## 6.2. Homicide data

Now, we demonstrate the importance of the NEG regression model over the gamma regression model. The data source is the Better Life Index (BLI). The BLI is calculated for OECD countries as well as Brazil Russia and South Africa. These data set can be download from <https://stats.oecd.org/index.aspx?DataSetCode=BLI2016>. It consists of 11 indicators. These indicators have 24 variables. Here, we relate the homicide rate ( $y_i$ ) with long-term unemployment rate ( $x_{i1}$ ) and labour market insecurity ( $x_{i2}$ ). The NEG and gamma regression models are used to model below regression structure.

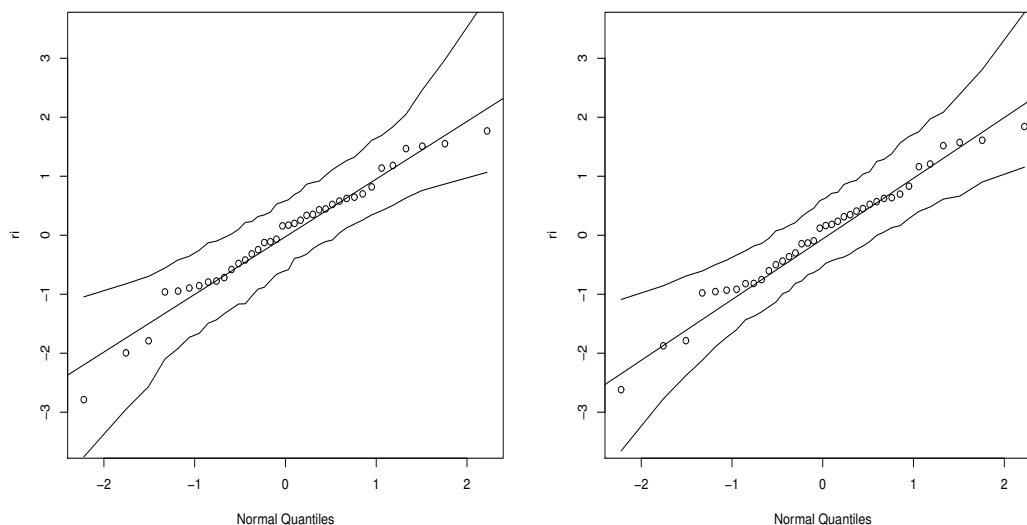
$$g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} \quad (6.1)$$

Table 2 lists the estimated parameters of gamma and NEG regression models and corresponding standard errors and p-values. As seen from these results, all estimated parameters are found statistically significant for both regression models. According to the estimated regression coefficients, we conclude that when the long-term unemployment rate increases, the homicide rate decreases surprisingly. On the other hand, when the labour market insecurity increases, the homicide rate also increases, as expected.

**Table 2.** The results of the NEG and gamma regression models.

Parameters	Gamma				NEG			
	Estimates	Std. Errors	t-value	p-value	Estimates	Std. Errors	t-value	p-value
$\beta_0$	2.017	0.173	11.639	<0.001	2.015	0.180	11.190	<0.001
$\beta_1$	-0.199	0.070	-2.839	0.007	-0.203	0.067	-3.048	0.002
$\beta_2$	0.126	0.044	2.844	0.007	0.128	0.045	2.849	0.004
$\alpha$	1.566	-	-	-	1.841	-	-	-
$-\ell$	117.537				115.278			
AIC	243.075				238.557			
BIC	249.625				245.107			

After fitting the regression models, we check the model accuracy with residual analysis. The rqs are obtained and plotted in Figure 9. Both regression model does not contain any possible outlier observation.



**Figure 9.** The rqs of gamma (left) and NEG (right) regression models.

## 7. Conclusions and future work

In this study, the two-parameter distribution is defined and studied comprehensively. The different estimation methods are investigated to estimate the unknown parameters of NEG distribution. Statistical properties of NEG distribution are derived. More importantly, a new regression model for the right-skewed response variable is introduced and is compared with the gamma regression model via an application to the actuarial data set. Empirical results show that NEG distribution could be a competitive model and can provide better modeling ability than its counterparts. The advanced residual analysis and influential diagnostics of the NEG regression model is a planned future work of this study. We believe that NEG distribution gains much attention from practitioners and increases its popularity soon.

## Appendix

The derivations of the cdf and Lorenz curve are given below.

$$\begin{aligned}
 F(x) &= \int_0^x f(t) dt. \\
 &= \int_0^x \frac{\theta^2}{2(\theta+3)} (2 + \theta t^2) e^{-\theta t} dt + \int_0^\infty \left( \frac{2}{\theta+3} \right) \frac{\theta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\theta t} dt. \\
 &= \frac{2 - e^{-x\theta} (2 + 2\theta x + \theta^2 x^2) + 2\theta (1 - e^{-x\theta})}{2(\theta+3)} + \frac{2\Gamma(\alpha) - 2\gamma(\alpha, \theta x)}{\Gamma(\alpha)(\theta+3)}
 \end{aligned}$$

$$= \frac{\Gamma(\alpha) \left[ 2 - e^{-x\theta} (2 + 2\theta x + \theta^2 x^2) \right] + 2\Gamma(\alpha)\theta(1 - e^{-x\theta}) + 4[\Gamma(\alpha) - \gamma(\alpha, \theta x)]}{2\Gamma(\alpha)(\theta + 3)}.$$

$$L(t) = \frac{\overbrace{\int_0^t xf(x) dx}^A}{\underbrace{\int_0^\infty tf(t) dt}_B}$$

$$B = \int_0^\infty tf(t) dt = E(X) = \frac{2\alpha + \theta + 3}{\theta(\theta + 3)}$$

$$A = \int_0^t xf(x) dx$$

$$= \int_0^t x \frac{\theta^2}{2(\theta + 3)} (2 + \theta x^2) e^{-\theta x} dx + \int_0^t x \left( \frac{2}{\theta + 3} \right) \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x} dx$$

$$= \int_0^\infty \frac{\theta^2}{2(\theta + 3)} (2x + \theta x^3) e^{-\theta x} + \int_0^\infty \left( \frac{2}{\theta + 3} \right) \frac{\theta^\alpha}{\Gamma(\alpha)} x^\alpha e^{-\theta x} dx$$

$$= \frac{6 - e^{-t\theta} (6 + 6\theta t + 3\theta^2 t^2 + \theta^3 t^3) + 2\theta - 2e^{-t\theta}\theta(1 + \theta t)}{2\theta(\theta + 3)} + \frac{2(\Gamma(\alpha + 1) - \gamma(\alpha + 1, \theta t))}{\Gamma(\alpha)\theta(\theta + 3)}$$

$$= \frac{\Gamma(\alpha) \left[ 6 - e^{-t\theta} (6 + 6\theta t + 3\theta^2 t^2 + \theta^3 t^3) + 2\theta - 2e^{-t\theta}\theta(1 + \theta t) \right] + 4(\Gamma(\alpha + 1) - \gamma(\alpha + 1, \theta t))}{2\Gamma(\alpha)\theta(\theta + 3)}$$

$$L(t) = \frac{A}{B}$$

$$= \frac{\Gamma(\alpha) \left[ 6 - e^{-t\theta} (6 + 6\theta t + 3\theta^2 t^2 + \theta^3 t^3) + 2\theta - 2e^{-t\theta}\theta(1 + \theta t) \right] + 4(\Gamma(\alpha + 1) - \gamma(\alpha + 1, \theta t))}{2\Gamma(\alpha)(2\alpha + \theta + 3)}$$

### Conflict of interest

The authors declare there is no conflict of interest.

---

## References

1. B. S. Everitt, D. J. Hand, *Finite Mixture Distributions*, Chapman and Hall, London, 1981.
2. W. C. Chen, B. M. Hill, J. B. Greenhouse, J. V. Fayos, Bayesian analysis of survival curves for cancer patients following treatment, *Bayesian Stat.*, **2** (1985), 299–328.
3. M. Erisoglu, N. Calis, T. Servi, U. Erisoglu, M. Topaksu, The mixture distribution models for interoccurrence times of earthquakes, *Russian Geol. Geophys.*, **52** (2011a), 685–692.
4. U. Erisoglu, M. Erisoglu, H. Erol, A mixture model of two different distributions approach to the analysis of heterogeneous survival data, *Int. J. Comput. Math. Sci.* **5** (2011b), 75–79.
5. U. Erisoglu, M. Erisoglu, H. Erol, Mixture model approach to the analysis of heterogeneous survival data, *Pakistan J. Stat.*, **5** (2012), 115–130.
6. K. E. Ahmad, A. M. Abd-El Rahman, Updating a nonlinear discriminant function estimated from a mixture of two Weibull distributions, *Math. Comput. Model.*, **19** (1994), 41–51.
7. R. Jiang, D. N. P. Murthy, Two sectional models involving three Weibull distributions, *Qual. Reliab. Eng. Int.*, **13** (1997), 83–96.
8. K. S. Sultan, M. A. Ismail, A. S. Al-Moisheer, Mixture of two inverse Weibull distributions: Properties and estimation, *Comput. Stat. Data Anal.*, **51** (2007), 5377–5387.
9. H. Zakerzadeh, A. Dolati, The generalized Lindley distribution, *J. Math. Ext.*, **3** (2009), 1–17.
10. S. F. Ateya, Maximum likelihood estimation under a finite mixture of generalized exponential distributions based on censored data, *Stat. Papers*, **55** (2014), 311–325.
11. A. M. Abouammoh, A. M. Alshangiti, I. E. Ragab, A new generalized Lindley distribution, *J. Stat. Comput. Simul.*, **85** (2015), 3662–3678.
12. A. H. El-Bassiouny, E. D. Medhat, M. Abdelfattah, M. S. Eliwa, Mixture of exponentiated generalized Weibull-Gompertz distribution and its applications in reliability, *J. Stat. Appl. Probab.*, **5** (2016), 455–468.
13. A. Karakoca, U. Erisoglu, M. Erisoglu, A comparison of the parameter estimation methods for bimodal mixture Weibull distribution with complete data, *J. Appl. Stat.*, **42** (2015), 1472–1489.
14. D. V. Lindley, Fiducial distributions and Bayes' theorem, *J. Royal Stat. Society. Series B (Methodological)*, **20** (1958), 102–107.
15. M. E. Ghitany, D. K. Al-Mutairi, N. Balakrishnan, L. J. Al-Enezi, Power Lindley distribution and associated inference, *Comput. Stat. Data Anal.*, **64** (2013), 20–33.
16. S. Nedjar, H. Zeghdoudi, On gamma Lindley distribution: Properties and simulations, *J. Comput. Appl. Math.*, **298** (2016), 167–174.
17. S. Sen, S. S. Maiti, N. Chandra, The xgamma distribution: Statistical properties and application, *J. Modern Appl. Stat. Methods*, **15** (2016), 774–788.
18. M. Bourguignon, M. D. C. S. Lima, J. Leao, A. D. Nascimento, L. G. B. Pinho, G. M. Cordeiro, A new generalized gamma distribution with applications, *Am. J. Math. Manage. Sci.*, **34** (2015), 309–342.

19. S. D. Waymyers, S. Dey, H. Chakraborty, A new generalization of the gamma distribution with application to negatively skewed survival data, *Commun. Stat.-Simul. Comput.*, **47** (2018), 2083–2101.
20. M. A. de Pascoa, E. M. Ortega, G. M. Cordeiro, The Kumaraswamy generalized gamma distribution with application in survival analysis, *Stat. Methodol.*, **8** (2011), 411–433.
21. M. O. Lorenz, Methods of measuring the concentration of wealth, *Publ. Am. Stat. Assoc.*, **9** (1905), 209–219.
22. B. Efron, *The jackknife, the bootstrap, and other resampling plans*, Vol. 38. Philadelphia, PA, USA: SIAM, 1982.
23. J. Mazucheli, A. F. B. Menezes, S. Dey, Bias-corrected maximum likelihood estimators of the parameters of the inverse Weibull distribution, *Comm. Stat.-Simul. Comput.*, **48** (2019), 2046–2055.
24. J. Mazucheli, A. F. B. Menezes, S. Dey, Improved maximum-likelihood estimators for the parameters of the unit-gamma distribution, *Comm. Stat.-Theory Methods*, **47** (2018), 3767–3778.
25. F. Ding, Decomposition based fast least squares algorithm for output error systems, *Signal Process.*, **93** (2013), 1235–1242.
26. F. Ding, P. X. Liu, G. Liue, Gradient based and least-squares based iterative identification methods for OE and OEMA systems, *Signal Process.*, **20** (2010), 664–677.
27. A. Zaka, A. S. Akhter, R. Jabeen, The new reflected power function distribution: Theory, simulation application., *Aims Math.*, **5** (2020), 5031–5054.
28. S. M. Zaidi, M. M. A. Sobhi, M. El-Morshedy, A. Z. Afify, A new generalized family of distributions: Properties and applications, *Aims Math.*, **6** (2021), 456–476.
29. E. Altun, H. M. Yousof, G. G. Hamedani, A new log-location regression model with influence diagnostics and residual analysis, *Facta Universitatis, Series: Math. Inf.*, **33** (2018), 417–449.
30. E. C. Cuervo, *Modelagem da variabilidade em modelos lineares generalizados*, (Doctoral dissertation, Tese de D. Sc., IM-UFRJ, Rio de Janeiro, RJ, Brasil, 2001.
31. E. Cepeda, D. Gamerman, Bayesian methodology for modeling parameters in the two parameter exponential family, *Revista Estadística*, **57** (2005), 93–105.
32. T. F. Bateson, Gamma regression of interevent waiting times versus poisson regression of daily event counts: Inside the epidemiologist’s toolbox-selecting the best modeling tools for the job, *Epidemiology*, **20** (2009), 202–204.
33. E. Altun, Weighted-exponential regression model: An alternative to the gamma regression model, *Int. J. Model., Simul., Sci. Comput.*, **10** (2019), 1–15.
34. E. Altun, The Lomax regression model with residual analysis, *J. Appl. Stat.*, **Forthcoming** (2020), 1–10.
35. J. McCullagh, J. Nelder, *Generalized Linear Models*, Second Edition, Chapman and Hall, London, 1989.
36. P. K. Dunn, G. K. Smyth, Randomized quantile residuals, *J. Comput. Graphical Stat.*, **5** (1996), 236–244.

- 
37. P. Feigl, M. Zelen, Estimation of exponential survival probabilities with concomitant information, *Biometrics*, **21** (1965), 826–838.
38. M. E. Mead, The beta exponentiated Burr XII distribution, *J. Stat.: Advances Theory Appl.* **12** (2014), 53–73.
39. M. V. Aarset, How to identify a bathtub hazard rate, *IEEE Trans. Reliab.*, **36** (1987), 106–108.



©2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)