Mediterr. J. Math. 13 (2016), 2935–2949
DOI 10.1007/s00009-015-0665-5
1660-5446/16/052935-15
published online December 13, 2015

Mediterranean Journal of Mathematics

A Two-Parameter Distribution Obtained by Compounding the Generalized Exponential and Exponential Distributions

Božidar V. Popović, Miroslav M. Ristić and Gauss M. Cordeiro

Abstract. We introduce a new two-parameter lifetime distribution obtained by compounding the generalized exponential and exponential distributions. We assume that the shape parameter of the generalized exponential distribution is a random variable having the exponential distribution. The shapes of the density and hazard rate functions are derived. The model parameters are estimated by maximum likelihood, and an application of the proposed distribution is presented.

Mathematics Subject Classification. 62E15, 62P10, 62F10.

Keywords. Compounding model, Exponential distribution, Generalized exponential distribution, Identifiability.

1. Introduction

There are many practical situations where well-known distributions do not provide adequate fits to real data. Among the different methods for generating new distributions, the compounding of some discrete and important lifetime distributions has been very popular in lifetime modeling.

Adamidis and Loukas [1] introduced a two-parameter exponential–geometric (EG) distribution by compounding the exponential and geometric distributions. More general family of these distributions had been considered by Marshall and Olkin [9]. Similarly, the exponential Poisson (EP) and exponential logarithmic distributions were introduced by Kuş [6] and Tahmasbi and Rezaei [15], respectively. Barreto-Souza et al. [3] proposed the Weibull geometric (WG) model, while Lu and Shi [7] studied the Weibull–Poisson (WP) distribution as a natural extension of the EG and EP distributions. Further, Rodrigues et al. [13] defined the Weibull negative binomial distribution, which includes as sub-models the WG and WP distributions. Nadarajah et al. [11] built in a `Compounding` R script for computing continuous distributions obtained by compounding a continuous and a discrete distribution.

Gupta and Kundu [5] introduced the generalized exponential (GE) distribution with cumulative distribution function (cdf) given by

$$G(x|\alpha,\lambda) = (1-\mathrm{e}^{-\lambda x})^{\alpha}, \tag{1}$$

where $x > 0, \alpha > 0, \lambda > 0$. They showed that the corresponding probability density function (pdf) is log convex for $0 < \alpha < 1$ and log concave for $\alpha > 1$. Here, we propose a new distribution by assuming that the shape parameter $\alpha > 0$ of the GE distribution is a random variable itself. For a given α, let a random variable Y have the GE cdf with parameters $\alpha > 0$ and $\lambda > 0$ given by (1). Let α be a random variable having the exponential distribution with parameter $\theta > 0$. Then, we define a new random variable X by the cdf (for $x > 0$)

$$F(x) = F(x;\lambda,\theta) = \int_0^{\infty} G(x|\alpha,\lambda)\,\theta\,\mathrm{e}^{-\theta\alpha}d\alpha = \left[1-\frac{1}{\theta}\log(1-\mathrm{e}^{-\lambda x})\right]^{-1},\ \lambda,\theta > 0. \tag{2}$$

The pdf of X becomes

$$f(x) = f(x;\lambda,\theta) = \frac{\lambda\mathrm{e}^{-\lambda x}}{\theta\,(1-\mathrm{e}^{-\lambda x})\left[1-\frac{1}{\theta}\log(1-\mathrm{e}^{-\lambda x})\right]^2}. \tag{3}$$

Henceforth, we denote a random variable X having the density function (3) by $X \sim \mathrm{GEE}(\lambda,\theta)$, where "GEE" stands for the generalized exponential exponential distribution. The parameter λ is a scale parameter. It is easy to prove that if the random variable X has the pdf $f(x;\lambda,\theta)$, then the random variable $Z = \lambda X$ has the pdf $f(x;1,\theta)$. The density function $f(x;\lambda,\theta)$ has shapes given by Theorem 3.

We motivate the GEE distribution by the following facts:

1. Comparing with the GE and exponential distributions, one will see that the pdf (3) provides more flexibility than these distributions. This fact will make the GEE model more attract to the readers.
2. The next theorem links the GEE distribution with other well-known distributions.

Theorem 1. *For the random variable X having pdf (3), we have:*

(a) *The random variable $Y = -\log(1-\mathrm{e}^{-\lambda X})$ has the log-logistic distribution with shape parameter $\beta = 1$ and scale parameter $\alpha = \theta$.*
(b) *The random variable $Y = -\log(1-\mathrm{e}^{-\lambda X})$ has the generalized Pareto distribution with location parameter $\mu = 0$, scale parameter $\sigma = \theta$, and shape parameter $\xi = 1$.*
(c) *The random variable $Y = 1-\frac{1}{\theta}\log(1-\mathrm{e}^{-\lambda X})$ has the Pareto distribution with shape parameter one.*
(d) *The random variable $Y = \sqrt{-\frac{1}{\theta}\log(1-\mathrm{e}^{-\lambda X})}$ has the Burr distribution with shape parameters equal to two and one.*

(e) *The random variable* $Y = \log\left[1 - \frac{1}{\theta}\log(1-\mathrm{e}^{-\lambda X})\right]$ *has the exponential distribution with unity scale parameter.*

Proof. The proof is trivial and left as an exercise to the reader. □

Theorem 2. *If the random variable* V *has the log-logistic distribution with shape parameter* $\beta = 1$ *and scale parameter* θ*, then the random variable* $X = -\frac{1}{\lambda}(1 - e^{-V})$ *has the distribution with pdf* (3).

Proof. The proof is trivial and left as an exercise to the reader. □

Henceforth, we consider the following expansions:

$$(1-z)^q = \sum_{k=0}^{\infty} (-1)^k \binom{q}{k} z^k, \qquad \text{if } |z| < 1, \tag{4}$$

$$\log(1-z) = -z \sum_{k=0}^{\infty} \frac{z^k}{k+1}, \qquad \text{if } |z| < 1 \tag{5}$$

and

$$\binom{-x}{n} = \frac{(-1)^n \, x^{(n)}}{n!}, \tag{6}$$

where $x^{(k)} = x(x+1)\ldots(x+k-1)$ denotes the rising factorial.

We shall use the exponential integral defined as $E_i(x) = -\int_{-x}^{\infty} \frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t$. The exponential integral can be easily computed in MATHEMATICA using the function ExpIntegralEi[x].

In the sequel, we require the Whittaker's function defined by

$$W_{k,m}(z) = \frac{\mathrm{e}^{-z/2}\, z^k}{\Gamma\left(\frac{1}{2} - k + m\right)} \int_0^{+\infty} t^{-k-1/2+m} \left(1 + \frac{t}{z}\right)^{k-1/2+m} \mathrm{e}^{-t}\,\mathrm{d}t,$$

where $\Gamma(\cdot)$ is the gamma function. The values of the Whittaker function can be obtained using the MATHEMATICA function WhittakerW[k,m,z].

We use an equation of Gradshteyn and Ryzhik ([4], equation 0.314) for a power series raised to a positive integer power p

$$\left(\sum_{n\geq 0} a_n\, x^n\right)^p = \sum_{n\geq 0} c_{p,n}\, x^n, \tag{7}$$

where the coefficients $c_{p,n}$ can be determined using the following recurrence relation (for $m \geq 1$ and $c_{p,0} = a_0^p$)

$$c_{p,m} = (m\, a_0)^{-1} \sum_{k=1}^{m} [k(p+1) - m]\, a_k\, c_{p,m-k}.$$

The rest of the paper is organized as follows. We discuss the shapes of the pdf and cdf of the new distribution in Sect. 2. We provide in Sect. 3 several mathematical properties of the GEE model such as the ordinary and incomplete moments, probability-weighted moments (PWMs), mean deviations, Bonferroni and Lorenz curves, generating function and order statistics. The Rényi and Shannon entropies are derived in Sect. 4. The parameters of the GEE distribution are estimated by maximum likelihood in Sect. 5. Two applications to real data illustrate the potentiality of the new model in Sect. 6. Finally, concluding remarks are addressed in Sect. 7.

2. Shapes of pdf and hrf

Here, we will study some shapes of the pdf and hazard rate function (hrf) of the random variable X with the GEE distribution.

Theorem 3. *If* $0 < \theta < 1 - \log 2$*, then there exist two positive real numbers* $a < b$ *such that* $f(x)$ *is a decreasing function in* $(0, a) \cup [b, \infty)$ *and an increasing function in* $[a, b)$*. If* $\theta > 1 - \log 2$*, then the function* $f(x)$ *is a decreasing function. For both cases, we have that* $f(0+) = \infty$*.*

Proof. The first derivative of the logarithm of the function $f(x)$ can be expressed as

$$[\log f(x)]' = \frac{\lambda u(x)}{\theta(1 - \mathrm{e}^{-\lambda x})\left[1 - \frac{1}{\theta}\log(1 - \mathrm{e}^{-\lambda x})\right]},$$

where the function $u(x)$ is given by

$$u(x) = -\theta + 2\mathrm{e}^{-\lambda x} + \log(1 - \mathrm{e}^{-\lambda x}).$$

The function $u(x)$ is a unimodal function with $u(0+) = -\infty$ and $u(\infty) = -\theta$. The function $u(x)$ has a maximum at $x = \lambda^{-1}\log 2$, and it is equal to $1 - \theta - \log 2$. Thus, if $\theta > 1 - \log 2$, the function $u(x)$ is negative which implies that $f(x)$ is a decreasing function with $f(0+) = \infty$. In the second case, $0 < \theta < 1 - \log 2$, the function $u(x)$ has two roots, say $a < b$. This implies that $u(x)$ is negative for $x \in (0, a) \cup [b, \infty)$ and positive for $x \in [a, b)$. So, we obtain that $f(x)$ is a decreasing function for $x \in (0, a) \cup [b, \infty)$ and an increasing function for $x \in [a, b)$. Further, we have $f(0+) = \infty$. □

Based on Eqs. (2) and (3), the hrf $h(x)$ of X is given by

$$h(x) = \frac{\lambda \mathrm{e}^{-\lambda x}}{(1 - \mathrm{e}^{-\lambda x})\left[-\log(1 - \mathrm{e}^{-\lambda x})\right]\left[1 - \frac{1}{\theta}\log(1 - \mathrm{e}^{-\lambda x})\right]}.$$

Theorem 4. *The hrf is bathtub-shaped for* $0 < \theta < 2$ *and decreasing for* $\theta > 2$*. For all parameter values, we have that* $h(\infty) = \lambda$*.*

Proof. The logarithm of $h(x)$ is given by

$$[\log h(x)]' = \frac{s(x)}{\theta(1 - \mathrm{e}^{-\lambda x})\left[-\log(1 - \mathrm{e}^{-\lambda x})\right]\left[1 - \frac{1}{\theta}\log(1 - \mathrm{e}^{-\lambda x})\right]},$$

where $s(x) = \lambda\theta\mathrm{e}^{-\lambda x} + \lambda(\theta - 2\mathrm{e}^{-\lambda x})\log(1 - \mathrm{e}^{-\lambda x}) - \lambda[\log(1 - \mathrm{e}^{-\lambda x})]^2$. If $0 < \theta < 2$, then there exists a positive real number a such that $s(x)$ is increasing in $(0, a)$ and decreasing in (b, ∞) with $s(0+) = -\infty$ and $s(\infty) = 0$. Thus, there exists a positive real number b such that, for $x < b$, the function $s(x)$ is negative and, for $x > b$, the function $s(x)$ is positive. This fact implies that $h(x)$ is bathtub-shaped.

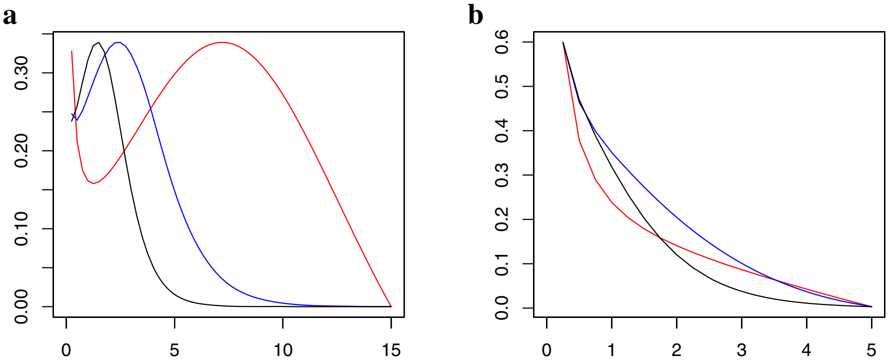

Figure 1. GEE pdf for $\lambda = 0.25$ (*red*), $\lambda = 0.75$ (*blue*), $\lambda = 1.25$ (*black*) and: **a** $\theta = 0.15$; **b** $\theta = 0.75$ (color figure online)

If $\theta > 2$, then $s(x)$ is an increasing function with $s(0+) = -\infty$ and $s(\infty) = 0$. Thus, $s(x)$ is a negative function, which implies that the hrf is a decreasing function. $\square$

Remark 1. *The hazard rate function $h(\cdot)$ has a finite positive limit λ, which implies according to Marshall and Olkin* ([8], Proposition B.3) *that residual life function converges to the exponential distribution with parameter λ.*

Figures 1 and 2 display, respectively, shapes of the pdf and hrf for different parameter values.

The finite limit property of the hazard rate function allows to use to prove the identifiability of the distribution with respect to parameters λ and θ.

Theorem 5. *The distribution function F given by* (2) *is identifiable with respect to parameters θ and λ.*

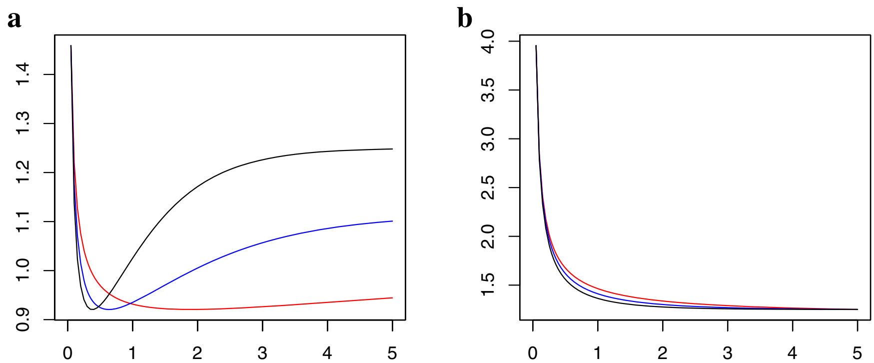

Figure 2. GEE hrf for $\lambda = 0.25$ (*red*), $\lambda = 0.75$ (*blue*), $\lambda = 1.25$ (*black*) and: **a** $\theta = 0.75$; **b** $\theta = 3.75$ (color figure online)

Proof. Let us suppose that $F(x;\lambda_1,\theta_1) = F(x;\lambda_2,\theta_2)$ for all $x > 0$. We will show that this condition implies that $\lambda_1 = \lambda_2$ and $\theta_1 = \theta_2$. First, we note that this condition implies that $h(x;\lambda_1,\theta_1) = h(x;\lambda_2,\theta_2)$ for all $x > 0$. Now, letting x tends to ∞ on both sides, and using the result from Theorem 4 that $h(\infty;\lambda,\theta) = \lambda$, we obtain that $\lambda_1 = \lambda_2$. Next, condition $F(x;\lambda_1,\theta_1) = F(x;\lambda_2,\theta_2)$ implies that $\theta_1(1-e^{-\lambda_1 x}) = \theta_2(1-e^{-\lambda_2 x})$ and since $\lambda_1 = \lambda_2$, we obtain finally that $\theta_1 = \theta_2$. Thus, we have proved the identifiability of the distribution function F. □

3. Mathematical Properties

Without loss of generality in order to simplify final expressions given in the Sects. 3 and 4, we will set $\lambda = 1$.

3.1. Moments

By means of the Theorem 2, the nth moment of X can be expressed as

$$\mathbb{E}(X^n) = \mathbb{E}\left(-\log(1-e^{-V})\right)^n = (-1)^n\,\frac{1}{\theta}\int_0^{+\infty}\frac{\left[\log(1-\mathrm{e}^{-v})\right]^n}{\left(1+\frac{1}{\theta}v\right)^2}\,\mathrm{d}v\,.$$

By setting $1+\frac{1}{\theta}v = u$, using both (5) and (7), the last equation reduces to

$$\mathbb{E}(X^n) = \sum_{k\geq 0} c_{n,k}\int_0^{+\infty}\frac{\mathrm{e}^{-\theta(n+k)v}}{(v+1)^2}\,\mathrm{d}v, \tag{8}$$

where (for $i \geq 1$)

$$c_{n,i} = \frac{1}{i}\sum_{j=1}^{i}\frac{[j(n+1)-i]}{(j+1)}\,c_{n,k-j}, \tag{9}$$

and $c_{n,\,0} = 1$. From the result (3.353.2) in [4], Eq. (8) can be expressed as

$$\mathbb{E}(X^n) = \sum_{k\geq 0} c_{n,k}\left\{1-\mathrm{e}^{-\theta(n+k)}\,E_i\left[-\theta(n+k)\right]\right\}. \tag{10}$$

Incomplete moments play important role for inequality measurement, for example mean deviations and the Lorenz curve, Pietra and Gini measures. Using a similar approach to the one for ordinary moments and equation (3.353.1) in [4], the nth incomplete moment of X is given by

$$\begin{aligned} m_n(z) = \sum_{k\geq 0} c_{n,k}\Bigg\{&\frac{(1-\mathrm{e}^{-z})^{n+k}}{1-\frac{1}{\theta}\log(1-\mathrm{e}^{-x})} + \frac{\theta(n+k)}{(1-\mathrm{e}^{-z})^{-1/\theta}} \\ &E_i\left[(\log(1-\mathrm{e}^{-z})-\theta)(n+k)\right]\Bigg\}, \end{aligned} \tag{11}$$

where $c_{n,k}$ is defined by (9).

Probability-weighted moments (PWMs) cover the summarization and description of theoretical distributions. The primary use of these moments is in the estimation of parameters for a probability distribution whose inverse cannot be expressed explicitly. One can use PWMs when maximum likelihood estimates are unavailable or difficult to compute. The $(s,r)th$ PWM of X is formally defined as

$$\tau_{s,r} = \mathbb{E}\left\{X^s\, F(X)^r\right\} = \int_0^{\infty} x^s\, F(x)^r\, f(x)\, \mathrm{d}x.$$

By equation (3.353.1) in [4] and using similar arguments as above, we obtain

$$\tau_{s,r} = \sum_{k\geq 0} c_{n,k} \left\{ \sum_{l=1}^{r+1} (l-1)!\, [-\theta(s+k)]^{r+1-l} \right. \\ \left. - \, [-\theta(s+k)]^{r+1}\, \mathrm{e}^{\theta(s+k)} E_i\,[-\theta(s+k)] \right\}.$$

3.2. Mean Deviations

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. The mean deviations of X about the mean and the median are defined by $\delta_1 = \int_0^{+\infty} |x-\mu|\, f(x)\, \mathrm{d}x$ and $\delta_2 = \int_0^{+\infty} |x - M|\, f(x)\, \mathrm{d}x$, where $\mu = E(X)$ and M is the median. Hence, the quantile function (qf) of X is $F^{-1}(p) = -\log\left[1-\mathrm{e}^{\theta(1-\frac{1}{p})}\right]$, where $p \in [0,1)$. So, the median equals to $M = F^{-1}(1/2) = -\log\left(1-\mathrm{e}^{-\theta}\right)$.

These measures can be calculated using the relationships $\delta_1 = 2[\mu F(\mu) - m_1(\mu)]$ and $\delta_2 = \mathbb{E}(X) - 2m_1(M)$. Clearly, $F(M)$ and $F(\mu)$ are easily calculated from Eq. (2), whereas the first ordinary moment and the first incomplete moment follow from (10) and (11), respectively.

Equation (11) with $n = 1$ is useful to derive the Bonferroni and Lorenz curves and mean deviations, that is, the Bonferroni and Lorenz curves are defined (for a given probability π) by $B(\pi) = m_1(q)/(\pi\, \mu_1')$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $q = F^{-1}(\pi)$ can be determined from the qf above.

3.3. Generating Function

The moment generating function (mgf) $M(t) = \mathbb{E}(\mathrm{e}^{tX})$ of X can be obtained using the same approach for the moments. We have

$$M(t) = \int_0^{+\infty} \frac{(1-\mathrm{e}^{-\theta v})^{-t}}{(v+1)^2}\, \mathrm{d}v\,.$$

Based on (4), equation (3.353.2) in [4] and (6), the last equation reduces to

$$M(t) = \sum_{k\geq 0} \frac{(t)^{(k)}}{k!}\, \left\{1 + \theta\, k\, \mathrm{e}^{\theta k}\, E_i\,(-\theta k)\right\}.$$

3.4. Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1, \ldots, X_n$ is a random sample from the GEE distribution.

Let $X_{i:n}$ denote the ith order statistic. From Eqs. (2) and (3), the power series expansion of the pdf of $X_{i:n}$ can be expressed as

$$f_{i:n}(x) = \frac{K\,\mathrm{e}^{-x}}{\theta(1-\mathrm{e}^{-x})} \sum_{j=0}^{n-i} \frac{(-1)^j}{\left[1-\frac{1}{\theta}\log(1-\mathrm{e}^{-x})\right]^{i+j+1}} \binom{n-i}{j},$$

where $K = \frac{n!}{(i-1)!(n-i)!}$.

We obtain the moments of the order statistics using a similar approach of Sect. 2. We have

$$\mathbb{E}(X_{i:n}^n) = K \sum_{j=0}^{n-i} \sum_{k\geq 0} \frac{(-1)^i\, c_{n,k}}{(i+j)!} \binom{n-1}{j} \times \left\{ \sum_{l=1}^{i+j} \frac{(l-1)!}{\left[-\theta(n+k)\right]^l} - \left[\theta(n+k)\right]^{i+j} \mathrm{e}^{\theta(n+k)}\, E_i\left[-\theta(n+k)\right] \right\}. \quad (12)$$

The importance of the moments of order statistics arises in the applied statistics such as quality control testing and reliability. If the reliability of an item is high, the duration of "all items fail" life test can be too expensive in both time and money. This fact prevents a practitioner from knowing enough about the product in a relatively short time. Therefore, a practitioner needs to predict the failure of future items based on the times of a few early failures. These predictions are often based on moments of order statistics.

The L-moments are summary statistics for distributions and data samples. They are analogous to ordinary moments but are computed from linear functions of the ordered data values. The rth L-moment of X is defined by

$$\lambda_r = \sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j} \binom{r-1+j}{j} \beta_j,$$

where $\beta_j = E\{XF(X)^j\}$. In particular, $\lambda_1 = \beta_0$, $\lambda_2 = 2\beta_1 - \beta_0$, $\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$, and $\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$. In general, $\beta_r = (r+1)^{-1}E(X_{r+1:r+1})$, so it can be computed from (12). The L-moments have several advantages over ordinary moments; for example, they apply for any distribution having finite mean; no higher-order moments need be finite.

4. Entropy

An entropy is a measure of variation or uncertainty of a random variable. Two popular entropy measures are the Rényi and Shannon entropies [12,14]. The Rényi entropy of a random variable with pdf $f(\cdot)$ is defined as

$$I_R(\gamma) = \frac{1}{1-\gamma} \log\left(\int_0^\infty f^\gamma(x)dx\right)$$

for $\gamma > 0$ and $\gamma \neq 1$. The Shannon entropy of a random variable is defined by $E\{-\log[f(X)]\}$. It is the special case of the Rényi entropy when $\gamma \uparrow 1$.

Here, we derive explicit expressions for the Rényi and Shannon entropies of X. We have

$$I_R(\gamma) = \frac{1}{1-\gamma} \log\left\{ \left(\frac{1}{\theta}\right)^{\gamma} \int_0^{\infty} \frac{e^{-\gamma x}}{(1-e^{-x})^{\gamma}\left[1-\frac{1}{\theta}\log(1-e^{-x})\right]^{2\gamma}}\, dx \right\}$$
$$= \frac{1}{1-\gamma} \log\left\{ \left(\frac{1}{\theta}\right)^{\gamma} \int_0^{\infty} \frac{e^{-(\gamma-1)x}}{(1-e^{-x})^{\gamma-1}} \frac{e^{-x}}{(1-e^{-x})\left[1-\frac{1}{\theta}\log(1-e^{-x})\right]^{2\gamma}}\, dx \right\}.$$

Then, using the binomial expansion of $(1-e^{-x})^{1-\gamma}$ and the same algebra which leads to (10), the last equation becomes

$$I_R(\gamma) = \tfrac{1}{1-\gamma} \log\left\{ \left(\tfrac{1}{\theta}\right)^{\gamma-1} \sum_{i,j\geq 0} (-1)^{i+j} \binom{1-\gamma}{i}\binom{\gamma-1+i}{j} \int_0^{+\infty} \frac{e^{-\theta j v}}{(1+v)^{2\gamma}} dx \right\}. \tag{13}$$

By equation (3.382.3) from [4], Eq. (13) turns out to be

$$I_R(\gamma) = \frac{1}{1-\gamma} \log\left\{ \left(\frac{1}{\theta}\right)^{\gamma-1} \sum_{i,j\geq 0} (-1)^{i+j} (j\theta)^{\gamma-1} \binom{1-\gamma}{i}\binom{\gamma-1+i}{j} e^{\frac{j\theta}{2}}\, W_{-\gamma;\frac{1-2\gamma}{2}}(j\theta) \right\}. \tag{14}$$

Equation (14) is very complicated for limiting, and then, we can derive an explicit expression for the Shannon entropy based on its definition. We can write

$$\mathbb{E}\{-\log[f(X)]\} = \log\theta - \mathbb{E}(X) + \mathbb{E}\left\{\log(1-e^{-X})\right\} + 2\mathbb{E}\left\{\log\left[1-\frac{1}{\theta}\log(1-e^{-X})\right]\right\}. \tag{15}$$

The first expectation in (15) follows easily from (10) for $n=1$. Using Eq. (5) and the same approach of Sect. 2, we obtain

$$\mathbb{E}\left\{\log(1-e^{-X})\right\} = -\sum_{i,j\geq 0} \frac{(-1)^j}{i}\binom{i}{j}\,[1+j\theta\, E_i(-j\theta)]. \tag{16}$$

Setting $1-\frac{1}{\theta}\log(1-e^{-x}) = u$, we easily obtain

$$\mathbb{E}\left\{\log\left[1-\frac{1}{\theta}\log(1-e^{-X})\right]\right\} = 1. \tag{17}$$

By inserting (10) (for $n=1$), (16), and (17) into (15), we obtain the Shannon entropy.

5. Estimation

In this section, we consider the maximum likelihood estimation (MLE) of the unknown parameters λ and θ. Suppose the observed sample $x_1,\ldots,x_n$ of size

n from the distribution (3). The log-likelihood function for (λ, θ) is given by

$$l(\lambda,\theta) = n\log\lambda - n\log\theta - \lambda\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\log(1-\mathrm{e}^{-\lambda x_i}) - 2\sum_{i=1}^{n}\log\left[1-\frac{1}{\theta}\log(1-\mathrm{e}^{-\lambda x_i})\right].$$

So, the components of the score function satisfy equations

$$\frac{\partial l(\lambda,\theta)}{\partial\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}\frac{x_i}{1-\mathrm{e}^{-\lambda x_i}} + \frac{2}{\theta}\sum_{i=1}^{n}\frac{x_i\,\mathrm{e}^{-\lambda x_i}}{(1-\mathrm{e}^{-\lambda x_i})\left[1-\frac{1}{\theta}\log(1-\mathrm{e}^{-\lambda x_i})\right]} = 0, \tag{18}$$

$$\frac{\partial l(\lambda,\theta)}{\partial\theta} = -\frac{n}{\theta} - \frac{2}{\theta^2}\sum_{i=1}^{n}\frac{\log(1-\mathrm{e}^{-\lambda x_i})}{1-\frac{1}{\theta}\log(1-\mathrm{e}^{-\lambda x_i})} = 0. \tag{19}$$

Now, we will study the existence and uniqueness of the MLE estimates when the other parameter is known (or given).

Theorem 6. *If the parameter θ is known, then the equation (18) has at least one root on the interval $(0,+\infty)$.*

Proof. One can readily verify that $\lim_{\lambda\to 0}\frac{\partial l(\lambda,\theta)}{\partial\lambda} = +\infty$ and $\lim_{\lambda\to+\infty}\frac{\partial l(\lambda,\theta)}{\partial\lambda} = -\sum_{i=1}^{n} x_i$. So, there exists at least one solution on the interval $(0,+\infty)$.

This completes the proof. □

Theorem 7. *Let us suppose that the parameter λ is known. Then, the root of equation (19) lies in the interval $\left(-\log\left(1-e^{-\lambda x_{(n)}}\right), -\log\left(1-e^{-\lambda x_{(1)}}\right)\right)$ and is unique for all $\lambda > 0$, where $x_{(1)} = \min(x_1, x_2, \ldots, x_n)$ and $x_{(n)} = \max(x_1, x_2, \ldots, x_n)$.*

Proof. Let us define the function $\psi(\theta) = -\frac{n}{2} + \sum_{i=1}^{n}\frac{y_i}{\theta+y_i}$, where $y_i = -\log(1-e^{-\lambda x_i})$. Then, the function $\frac{\partial l(\lambda,\theta)}{\partial\theta}$ can be represented as $\frac{\partial l(\lambda,\theta)}{\partial\theta} = \frac{2}{\theta}\psi(\theta)$, which implies that we can derive its behavior through the behavior of the function $\psi(\theta)$. We have that $\psi(\theta)$ is almost surely decreasing on $(0,+\infty)$, and it holds $\lim\limits_{\theta\downarrow 0}\psi(\theta) = \frac{n}{2}$ and $\lim\limits_{\theta\uparrow 0}\psi(\theta) = -\frac{n}{2}$. Thus, the function $\psi(\theta)$ has the unique root, which implies that the Eq. (19) has the unique solution.

Also, since $y_{(1)} + y_i \le 2y_i$ for all $i = 1, 2, \ldots, n$ then

$$\psi\left(y_{(1)}\right) = -\frac{n}{2} + \sum_{i=1}^{n}\frac{y_i}{y_{(1)}+y_i} \ge -\frac{n}{2} + \sum_{i=1}^{n}\frac{1}{2} = 0\,,$$

and analogously

$$\psi\left(y_{(n)}\right) = -\frac{n}{2} + \sum_{i=1}^{n}\frac{y_i}{y_{(n)}+y_i} \le -\frac{n}{2} + \sum_{i=1}^{n}\frac{1}{2} = 0\,.$$

From the last two equations, we found interval which is the solution of the Eq. (19). □

Table 1. Descriptive statistics for the Aarset data set

Mean	Median	SD	Skewness	Kurtosis	Min	Max
45.69	48.5	32.8352	−0.1378	1.4138	0.1	86

Table 2. Estimated parameters, AIC, KS, and p value for Aarset data

Model	$\widehat{\lambda}$	$\widehat{\alpha}$	$\widehat{\theta}$	$\widehat{\beta}$	AIC	KS	p
GE	0.0187 (0.0036)	0.7798 (0.1352)	–	–	483.99	0.2042	0.0309
GEE	0.0383 (0.0058)	–	0.2158 (0.0883)	–	473.57	0.1491	0.2162
MOE	0.0326 (0.0070)	2.6214 (1.0927)	–	–	483.10	0.1617	0.1464
Weibull	0.9490 (0.1095)	–	–	44.9125 (6.9465)	486	0.1928	0.0486

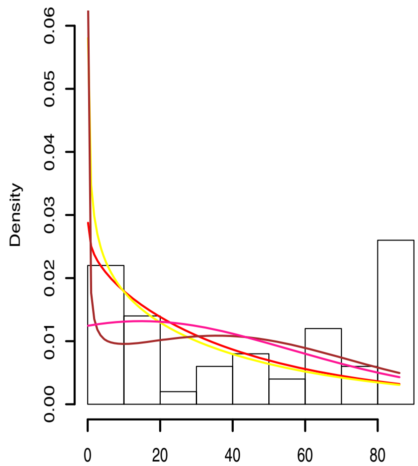

Figure 3. Histogram of the Aarset data and fitted density functions of the Weibull (*red*), GE (*yellow*), GEE (*brown*), and MOE (*pink*) distributions (color figure online)

6. Application

Here, we use two real data sets to compare the fits of the GEE distribution with other fits from the Weibull, generalized exponential (GE), and Marshal–Olkin exponential (MOE) distributions. The parameters are estimated using maximum likelihood and reported jointly with standard errors in parentheses in Tables 2 and 4.

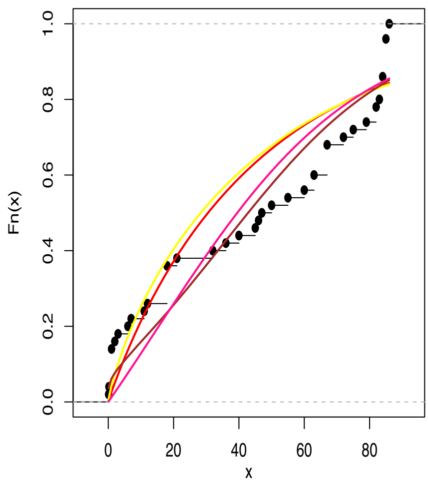

Figure 4. Empirical distribution function of the current data and fitted cdfs of the Weibull (*red*), GE (*yellow*), GEE (*brown*), and MOE (*pink*) distributions (color figure online)

Table 3. Descriptive statistics for the test stop data

Mean	Median	SD	Skewness	Kurtosis	Min	Max
0.0014	5.8460	3.4375	−0.2294	1.7753	0.0014	10.76

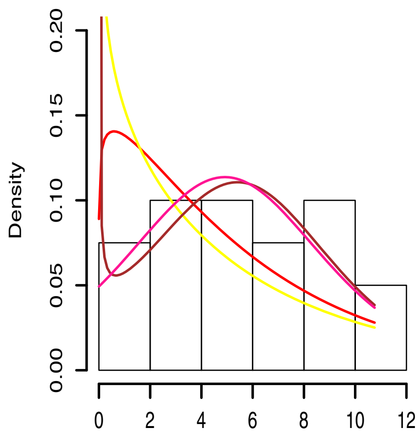

Figure 5. Histogram of the test stopped data and fitted pdfs of the Weibull (*red*), GE (*yellow*), GEE (*brown*), and MOE (*pink*) distributions (color figure online)

6.1. Application 1: Aarset Data

This data set appears in [2]. The data represent the lifetimes of 50 devices. Table 1 gives a data set descriptive summary.

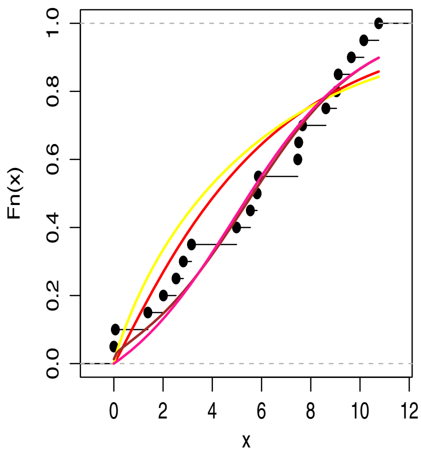

Figure 6. Empirical distribution function of the current data and fitted cdfs of the Weibull (*red*), GE (*yellow*), GEE (*brown*), and MOE (*pink*) distributions (color figure online)

Table 4. Estimated parameter, AIC, KS, and p, values for the test stopped data

Model	$\widehat{\lambda}$	$\widehat{\alpha}$	$\widehat{\theta}$	$\widehat{\beta}$	AIC	KS	p
GE	0.1570	0.8377	–	–	113.24	0.2493	0.1397
	(0.0467)	(0.2300)					
GEE	0.4202	–	0.0974	–	100.41	0.1375	0.7958
	(0.0939)		(0.0707)				
MOE	0.3984	8.0859	–	–	107.78	0.1472	0.7254
	(0.1046)	(5.7289)					
Weibull	1.0892	–	–	5.8163	113.50	0.2205	0.2465
	(0.2210)			(1.2218)			

Figure 3 displays the histogram of the current data and the fitted pdfs. Figure 4 displays the empirical cdf for the current data superimposed with the fitted cdfs. These two figures reinforce that the GEE distribution provides the best fit to the blood cancer data.

We compute the MLEs of the model parameters and adopt the Akaike information criteria (AIC) and p values corresponding to the Kolmogorov–Smirnov (KS) test for comparing the fitted models. The results in Table 2 indicate that the GEE distribution has the lowest AIC and the largest p value among all fitted distributions.

6.2. Application 2: Test Stopped Data

Here, we use data points representing failure times. The data are taken from [10]. The descriptive statistics of the failure times are given in Table 3.

From the figures in Table 3, it is obvious that the GEE distribution provides a better fit to the failure time data than the other models.

The histogram of the failure time data and the fitted pdfs is displayed in Fig. 5, whereas the empirical cdf of these data and the fitted cdfs is displayed in Fig. 6. These two figures reinforce our earlier observation that the GEE distribution provides the best fit.

7. Concluding Remarks

We introduce a new two-parameter model, called the generalized exponential exponential (GEE) distribution, and study some of its structural properties. We provide explicit expressions for the density function, moments and incomplete moments, probability-weighted moments, generating function, mean deviations, Bonferroni and Lorenz curves, and two measures of entropy. Our formulas related with the GEE model are manageable and, with the use of modern computer resources with analytic and numerical capabilities, may turn into adequate tools comprising the arsenal of applied statisticians. The model parameters are estimated by maximum likelihood, and the existence of the ML estimates is proved. This distribution is a very competitive model to other lifetime distributions. In fact, we prove that this can be superior to some widely known lifetime distributions by means of two examples with real data.

Acknowledgements

We are grateful to the Editor and anonymous referees for comments that greatly improved the quality of the paper. The authors are very grateful to one of the referee for providing them the proof of Theorem 7.

References

[1] Adamidis, K., Loukas, S.: A lifetime distribution with decreasing failure rate. Stat. Probab. Lett. **39**, 35–42 (1998)

[2] Aarset, M.V.: How to identify bath tub hazard rate. IEEE Trans. Reliab. **36**(1), 106–108 (1987)

[3] Barreto-Souza, W., Morais, A.L., Cordeiro, G.M.: The Weibull-geometric distribution. J. Stat. Simul. Comput. **81**, 645–657 (2010)

[4] Gradshteyn, I.S., Ryzhik, I.M.: Table of Integrals, Series, and Products. 7th edn. Academic Press, San Diego (2007)

[5] Gupta, R.D., Kundu, D.: Generalized exponential distribution. Aust. N. Z. J. Stat. **41**, 173–188 (1999)

[6] Kuş, C.: A new lifetime distribution. Comput. Stat. Data Anal. **51**, 4497–4509 (2007)

[7] Lu, W., Shi, D.: A new compounding life distribution: the Weibull-Poisson distribution. J. Appl. Stat. **39**, 21–38 (2012)

[8] Marshall, A.W., Olkin, I.: Life Distributions. Springer, New York (2007)

[9] Marshall, A.W., Olkin, I.: A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika. **84**, 641–652 (1997)

[10] Murthy, D.N.P., Xie, M., Jiang, R.: Weibull Models. Wiley, New York (2005)

[11] Nadarajah, S., Popović, B.V., Ristić, M.M.: Compounding: an R package for computing continuous distributions obtained by compounding a continuous and a discrete distribution. Comput. Stat. **28**, 977–992 (2013)

[12] Rényi, A.: On measures of entropy and information. In: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Volume I, PP. 547–561. University of California Press, Berkeley (1961)

[13] Rodrigues, C., Cordeiro, G.M., Demetrio, C.G.B., Ortega, E.M.M.: The Weibull negative binomial distribution. Adv. Appl. Stat. **22**, 25–55 (2011)

[14] Shannon, C.E.: Prediction and entropy of printed English. The Bell System Tech. J. **30**, 50–64 (1951)

[15] Tahmasbi, R., Rezaei, S.: A two-parameter lifetime distribution with decreasing failure rate. Comput. Stat. Data Anal. **52**, 3889–3901 (2008)

Božidar V. Popović
Faculty of Philosophy
University of Montenegro
Nikšić, Montenegro
e-mail: `bozidarpopovic@gmail.com`

Miroslav M. Ristić
Faculty of Sciences and Mathematics
University of Niš
Niš, Serbia
e-mail: `miristic72@gmail.com`

Gauss M. Cordeiro
Departamento de Estatística
Universidade Federal de Pernambuco
Recife, PE, Brazil
e-mail: `gausscordeiro@uol.com.br`

Received: January 28, 2015.
Revised: November 9, 2015.
Accepted: November 26, 2015.