The Complementary Exponentiated Exponential Geometric Lifetime Distribution

Francisco Louzada, Vitor Marchi, and James Carpenter

1 Department of Applied Mathematics and Statistics, ICMC, University of São Paulo, 13560-970 São Carlos, SP, Brazil
2 Department of Statistics, Federal University of São Paulo, 13565-905 São Carlos, SP, Brazil
3 London School of Hygiene and Tropical Medicine, University of London, Keppel Street, London WC1E 7HT, UK

Correspondence should be addressed to Francisco Louzada; louzada@icmc.usp.br

Received 8 August 2012; Revised 17 November 2012; Accepted 26 November 2012

1. Introduction

Several new classes of models have been introduced in recent years grounded in the simple exponential distribution. The main idea is to propose lifetime distributions which can accommodate practical applications where the underlying hazard functions are nonconstant, presenting monotone shapes, since the exponential distribution does not provide a reasonable fit in such situations. For instance, we can cite [1], which proposed a variation of the exponential distribution, the exponential geometric (EG) distribution, with decreasing hazard function, [2], which introduced the exponentiated exponential distribution as a generalization of the usual exponential distribution, which can accommodate data with increasing and decreasing hazard functions, [3], which proposed a generalized exponential distribution, which can accommodate data with increasing and decreasing hazard functions, [4], which proposed the exponentiated type distributions extending the Fréchet, gamma, Gumbel, and Weibull distributions, [5], which proposed another modification of the exponential distribution with decreasing hazard function, [6], which generalizes the distribution proposed by [5] by including a power parameter in this distribution, which can accommodate increasing, decreasing, and unimodal hazard functions, [7], which proposed the Poisson-exponential distribution, and [8], which proposed the complementary exponential geometric distribution, which is complementary to the exponential geometric distribution proposed by [1].

In this paper, following [8], we propose a new distribution family by extending the exponentiated exponential distribution [2] by compounding it with a geometric distribution, hereafter the complementary exponentiated exponential geometric distribution or simplistically the CE2G distribution. The new distribution genesis is stated on a complementary risk problem base [9] in presence of latent risks, in the sense that there is no information about which factor was responsible for the component failure and only the maximum lifetime value among all risks is observed. This family have one shape and two scale parameters accommodating increasing, decreasing, and bathtub failure rates.
The paper is organized as follows. In Section 2 we introduce the new CE2G distribution, derive the expressions for the probability density, survival, and hazard functions and the $p$th quantile, and present its genesis. In Section 3 we present some of its properties, such as its characteristic function, $r$th raw moment, mean and variance, order statistics, $r$th moment of the $r$th order statistic, mean residual lifetime, and modal value. In Section 8 we present the inferential procedure. In Section 10 the practical importance of the new distribution was demonstrated in three applications where our distribution outperforms several former lifetime distributions, such as the exponential, the exponential-geometric, the Weibull, the modified Weibull, and the generalized exponential Poisson distribution. Some final comments in Section 11 conclude the paper.

2. The CE2G Model

Let $Y$ be a nonnegative random variable denoting the lifetime of a component in some population. The random variable $Y$ is said to have a CE2G distribution with parameters $\lambda > 0$, $\alpha > 0$, and $0 < \theta < 1$ if its probability density function (pdf) is given by

$$f(y) = \frac{\alpha \lambda \theta e^{-\lambda y}(1 - e^{-\lambda y})^{\alpha - 1}}{[1 - (1 - \theta)(1 - e^{-\lambda y})^\alpha]^2}, \quad y > 0,$$  \hspace{1cm} (1)

where $\lambda$ is a scale parameter of the distribution, and $\alpha$ and $\theta$ are shape parameters. Figure I(a) shows the CE2G probability density function for $\lambda = 1$, $\theta = 0.05, 0.5, 0.95$, and $\alpha = 0.3, 1.0, 3$ and we can see that the function can be decreasing or unimodal.

The survival function of a CE2G distributed random variable is given by

$$S(y) = \frac{1 - (1 - e^{-\lambda y})^\alpha}{1 - (1 - \theta)(1 - e^{-\lambda y})^\alpha}, \quad y > 0,$$ \hspace{1cm} (2)

where $\alpha > 0$, $\theta \in (0, 1)$, and $\lambda > 0$.

From (2) and (1), the failure rate function, according to the relationship $h(y) = f(y)/S(y)$, is given by

$$h(y) = \frac{\alpha \lambda \theta e^{-\lambda y}(1 - e^{-\lambda y})^{\alpha - 1}}{[1 - (1 - \theta)(1 - e^{-\lambda y})^\alpha][1 - (1 - \theta)(1 - e^{-\lambda y})^\alpha]}. \hspace{1cm} \text{(3)}$$

The initial value is not finite if $\alpha < 1$ and otherwise is given by $h(0) = \lambda \theta$ if $\alpha = 1$ or $h(0) = 0$ if $\alpha > 1$ and the long-term hazard function value is $h(\infty) = \lambda$. The failure rate (3) can be increasing, decreasing, or bathtub as shown in Figure I(b), which shows some failure rate function shapes to $\lambda = 1, \theta = 0.05, 0.5, 0.95$, and $\alpha = 0.3, 1.0, 3$.

The $p$th quantile of the CE2G distribution is given by

$$Q(u) = F^{-1}(u) = \frac{-\ln \left[ \frac{1 - (1 - \theta)(1 - (u + 1))^{1/p}}{\lambda} \right]}{\alpha} \hspace{1cm} \text{for} \quad u \text{ has the uniform } U(0, 1) \text{ distribution and } F(y) = 1 - S(y) \text{ is the distribution function of } Y.$$  \hspace{1cm} (4)

Consider that in the study of reliability we can observe only the maximum component lifetime for each component among all risks. On many occasions, the information about what risk produces the dead of the component in analysis is not available or it is impossible that the true cause of failure is specified. Complementary risks (CR) problems arise in several areas and an extensive literature is available. Interested readers can see [10–12].

Then, in this context, our model can be derived as follows. Let $M$ be a random variable denoting the number of failure causes, $m = 1, 2, \ldots$ and considering $M$ with geometrical probability distribution given by

$$P(M = m) = \theta(1 - \theta)^{m - 1},$$ \hspace{1cm} (5)

where $0 < \theta < 1$ and $M = 1, 2, \ldots$

Also consider $t_i, i = 1, 2, 3, \ldots$ realizations of a random variable denoting the failure times, that is, the time-to-event due to the $i$th CR and, from [2], $t_i$ has an exponentiated exponential probability distribution with parameters $\lambda$ and $\alpha$, given by

$$f(t_i; \lambda, \alpha) = \alpha g(t; \lambda) G(t; \lambda) \hspace{1cm} \text{where } g(\cdot) \text{ and } G(\cdot) \text{ are the pdf and df, respectively, of the exponential distribution with parameter } \lambda.$$  \hspace{1cm} (6)

In the latent complementary risks scenario, the number of causes $M$ and the lifetime $t_i$ associated with a particular cause are not observable (latent variables), but only the maximum lifetime $Y$ among all causes is usually observed. So, we only observe the random variable given by

$$Y = \max\{T_1, T_2, \ldots, T_M\}.$$ \hspace{1cm} (7)

The following result shows that the random variable $Y$ has probability density function given by (1).

Proposition 1. If the random variable $Y$ is defined as (7), then, considering (5) and (6), $Y$ is distributed according to a CE2G distribution, with probability density function given by (1).

Proof. The conditional density function of (7) given $M = m$ is given by

$$f(y | M = m, \lambda, \alpha) = m \alpha \lambda e^{-\lambda y} (1 - e^{-\lambda y})^{\alpha - 1} \left( \frac{(1 - e^{-\lambda y})^{\alpha}}{m - 1} \right)^{m - 1} \hspace{1cm} \text{for } t > 0, \ m = 1, \ldots$$ \hspace{1cm} (8)
Then, the marginal probability density function of $Y$ is given by

$$f(y) = \sum_{m=1}^{\infty} m\alpha\lambda e^{-\lambda y} (1-e^{-\lambda y})^{\alpha-1} \left[ (1-e^{-\lambda y})^\alpha (1-\theta) \right]^{m-1}$$

This completes the proof. \hfill $\square$

3. Some Properties

Many of the most important features and characteristics of a distribution can be studied through its moments, such as mean and variance. A general expression for $r$th ordinary moment $\mu'_r = E(Y^r)$ of the CE2G distribution is hard to be obtained and we resume the mean and variance as follows.

The moment generating function of the $Y$ variable with density function given by (1) can be obtained analytically, if we consider the expression, given in [13, page 329, Equation (1.6)]:

$$\int_0^1 z^{p-1} (1-z)^{n-1} (1+bz^m)^{l-1} \, dz = \frac{\Gamma(p) \Gamma(n) \sum_{k=0}^{\infty} (\frac{1}{k})^\frac{b^k \Gamma(p+km)}{\Gamma(p+n+km)}}{\Gamma(n) \sum_{k=0}^{\infty} (\frac{1}{k})^\frac{b^k \Gamma(p+km)}{\Gamma(p+n+km)}}.$$ (10)

For any real number $t$, let $\Phi_Y(t)$ be the characteristic function of $Y$, that is, $\Phi_Y(t) = E[e^{itY}]$, where $i$ denotes the imaginary unit. With the preceding notations, we state the following.

**Proposition 2.** For the random variable $Y$ with CE2G distribution, we have that its characteristic function is given by

$$\Phi(t) = \alpha\theta \Gamma \left( 1 - \frac{it}{\lambda} \right) \sum_{k=0}^{\infty} \left( \frac{-2}{k} \right)^\frac{\Gamma(\alpha[k+1])(\theta-1)^k}{\Gamma(\alpha[k+1]+1-it/\lambda)},$$ (11)

where $i = \sqrt{-1}$. 

![Figure 1: (a) Probability density function of the CE2G distribution. (b) Failure rate function of the CE2G distribution. We fixed $\lambda = 1$.](image-url)
Proof. Consider the following:

\[
\Phi_Y(t) = \int_0^\infty e^{iy} f(y) \, dy = \int_0^\infty e^{iy} \frac{\alpha \theta e^{-\lambda y}(1 - e^{-\lambda y})^{\alpha-1}}{[1 - (1 - \theta)(1 - e^{-\lambda y})]^2} \, dy = \alpha \theta \int_0^1 \frac{z^{\alpha-1}(1-z)^{-it/\lambda}}{[1 - (1 - \theta)z]^2} \, dz,
\]

where the last equality follows from the change of variable \(z = 1 - e^{-\lambda y}\). Comparing the last integral with (10), obtaining \(b = \theta - 1\), and making the appropriate substitutions completed the proof. \(\square\)

**Proposition 3.** A random variable \(Y\) with density given by (1) has mean and variance given, respectively, by

\[
E(Y) = \frac{\theta}{\lambda^2} \sum_{k=0}^{\infty} \left(\frac{-2}{k} \right) \left[ \Psi(0, \alpha [k+1] + 1) - \Psi(0, 1) \right],
\]

\[
\text{Var}(Y) = \frac{\theta}{\lambda^2} \left\{ \sum_{k=0}^{\infty} \left(\frac{-2}{k} \right) \left[ \Psi(0, \alpha [k+1] + 1) - 2\Psi(0, 1) \right] \right. \]

\[
\left. - \Psi(1, \alpha [k+1] + 1) \right\} \quad \Xi \left[ \Psi(0, \alpha [k+1] + 1) - \Psi(0, 1) \right]\right\},
\]

where \(\Psi(n, z) = (a^{n+1}/dz^{n+1}) \ln(\Gamma(z))\) is known as PsiGamma function.

Proof. The first result follows from the relationship \(\Phi'_Y(t) \mid_{t=0} = E(Y)\). From the literature, \(\Phi''_Y(t) \mid_{t=0} = E(Y^2)\) and \(\text{Var}(Y) = E(Y^2) - [E(Y)]^2\), and with a little algebra follow the results. \(\square\)

Skewness is a measure of the asymmetry of the probability distribution. The skewness value can be positive or negative, or even undefined. Qualitatively, a negative skew indicates that the tail on the left side of the probability density function is longer than the right side and the bulk of the values lie to the right of the mean. A positive skew indicates that the tail on the right side is longer than the left side and the bulk of the values lie to the left of the mean. The skewness of a random variable \(Y\), say \(\gamma_1\), is given by the third standardized moment

\[
\gamma_1 = \frac{E[(Y - \mu)^3]}{(E[(Y - \mu)^2])^{3/2}}
\]

\[
= \frac{E(Y^3) - 3E(Y^2)E(Y) + 3E^2(Y)E(Y) - E^3(Y)}{[E(Y^2) - E^2(Y)]^{3/2}}
\]

(14)

Kurtosis is any measure of the “peakedness” of the probability distribution of a real-valued random variable. In a similar way to the concept of skewness, kurtosis is a descriptor of the shape of a probability distribution. It is common practice to use the kurtosis to provide a comparison of the shape of a given distribution to that of the normal distribution. One common measure of kurtosis, originating with Karl Pearson, say \(\gamma_2\), is based on a scaled version of the fourth moment, given by

\[
\gamma_2 = \frac{E[(Y - \mu)^4]}{(E[(Y - \mu)^2])^2}
\]

\[
= \frac{E(Y^4) - 4E(Y^3)E(Y) + 6E(Y^2)E^2(Y) - 3E^4(Y)}{[E(Y^2) - E^2(Y)]^2}
\]

(15)

Algebraic expressions of kurtosis and skewness are extensive to show, due to the fact that is necessary the algebraic moment expressions up order four. This moment can be obtained by algebraic manipulation to determine \(E(Y)\), \(E(Y^2)\), \(E(Y^3)\), and \(E(Y^4)\) in (14) and (15) through the Equation (11). Figure 2 shows the kurtosis \(\gamma_2\) and skewness \(\gamma_1\) of the CE2G distribution for \(\alpha = 0.3, 1.0, 3\) and for \(\theta = 1, 0.1, 0.5, 0.9\) and for \(\theta\) with \(\lambda = 1, \alpha = 0.3, 1.0, 3\).

### 4. Order Statistics

Order statistics are among the most fundamental tools in nonparametric statistics and inference. Let \(Y_1, \ldots, Y_n\) be a random sample taken from the CE2G distribution and \(Y_{1:n}, \ldots, Y_{n:n}\) denote the corresponding order statistics. Then, the pdf \(f_{i:n}(y)\) of the \(i\)th order statistics \(Y_{i:n}\) is given by

\[
f_{i:n}(x) = \frac{n!}{(k-1)! (n-k)!} F'(y)^{k-1} (1 - F(y))^{n-k} f(y).
\]

(16)

The \(r\)th moment of the \(i\)th order statistic \(Y_{i:n}\) can be obtained from the following result due to [14]:

\[
E \left[ Y_{i:n}^r \right] = r \sum_{p=n-r+1}^{n} \binom{n}{p+1} \left( \frac{p-1}{n-i} \right) \int_0^\infty y^{r-1} S(y)^p \, dy.
\]

(17)

Consider the binomial series expansion given by

\[
(1 - x)^{-r} = \sum_{k=0}^{\infty} \binom{r}{k} x^k,
\]

(18)
where \((r)_k\) is a Pochhammer symbol, given \((r)_k = r(r + 1) \cdots (r + k - 1)\) and if \(|x| < 1\) the series converge, and
\[
(-r)_k = (-1)^k(r - k + 1)_k. \tag{19}
\]

**Proposition 4.** For the random variable \(Y\) with CE2G distribution, we have that \(r\)th moment of the \(i\)th order statistic is given by
\[
E[Y^r_{\text{min}}] = \frac{\Gamma(r)}{\lambda^r} \sum_{p=n-i+1}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty (-1)^p (-i)_{(p-1)} \binom{n-i}{p} \binom{p}{n-i} \theta^j \alpha^{(j+l)+k} \ln^{r-1}(1-x) \ln^{(l)}(m+1)^r.
\tag{20}
\]

**Proof.** From (2) and (18), we have that
\[
\int_0^\infty y^{r-1}[S(y)]^p dy = \frac{(-1)^{-r} \lambda^{-1} \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty (1-\theta)^j (-p)_j (-p)_l j! l!}{\lambda^r \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{m=0}^\infty (1-\theta)^j (-p)_j (-p)_l j! l!} \ln^{r-1}(1-x) \ln^{(l)}(m+1)^r.
\tag{21}
\]

Using the change of variable \(\ln(1-x) = -u\) and the expansion (18) results in the kernel of the gamma distribution function as
\[
\int_0^\infty y^{r-1}[S(y)]^p dy = \frac{(-1)^{-r} \lambda^{-1} \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{m=0}^\infty (1-\theta)^j (-p)_j (-p)_l j! l!}{\lambda^r \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{n=0}^\infty (1-\theta)^j (-p)_j (-p)_l j! l!} \ln^{r-1}(1-x) \ln^{(l)}(m+1)^r.
\tag{22}
\]

Now considering (22) in (17) and the property (19), the result follows.

**5. Entropy**

An entropy of a random variable \(Y\) is a measure of variation of the uncertainty. A popular entropy measure is Rényi entropy [15].
If \( Y \) has the probability density function (1) then Rényi entropy is defined by

\[
y(\rho) = \frac{1}{1-\rho} \log \left( \int f^\rho(y) \, dy \right),
\]

where \( \rho > 0 \) and \( \rho \neq 1 \).

**Proposition 5.** If the random variable \( Y \) is defined as (7), then, the Rényi entropy, is given by

\[
y(\rho) = \frac{1}{1-\rho} \log \left( \alpha \lambda \theta \sum_{k=0}^{\infty} \left[ (1-\theta)^k (2\rho)_k \Gamma(\rho(\alpha-1)+k\alpha+1) \right] \right).
\]

**Proof.** From (23), we can calculate

\[
\int f^\rho(y) \, dy = \int_{0}^{\infty} (\alpha \lambda \theta)^\rho e^{-\lambda \rho y} (1-e^{-\lambda y})^{\rho(\alpha-1)} \, dy
\]

\[
= (\alpha \lambda \theta)^\rho \sum_{k=0}^{\infty} \left[ (1-e^{-\lambda y})^{\rho(\alpha-1)+k\alpha} \times (1-\theta)^k \frac{(2\rho)_k}{k!} \right] \, dy
\]

\[
= (\alpha \lambda \theta)^\rho \sum_{k=0}^{\infty} \left[ (1-e^{-\lambda y})^{\rho(\alpha-1)+k\alpha} \times \frac{(2\rho)_k}{k!} \lambda e^{-\lambda y} \right] \, dy
\]

\[
= \theta^2 \lambda \theta \sum_{k=0}^{\infty} \left[ (1-e^{-\lambda y})^{\rho(\alpha-1)+k\alpha} \right] \times \frac{(2\rho)_k}{k!} \frac{\lambda (1-\theta)^k}{(k+1)}
\]

6. **Reliability**

In the context of reliability, the stress-strength model describes the life of a component which has a random strength \( Y \) that is subjected to a random stress \( X \). The component fails at the instant that the stress applied to it exceeds the strength, and the component will function satisfactorily whenever \( Y > X \). So, \( R = \Pr(X < Y) \) is a measure of component reliability. In the area of stress-strength models there has been a large amount of work as regards estimation of the reliability \( R \) when \( Y \) and \( X \) are independent random variables belonging to the same univariate family of distributions.

**Proposition 6.** If the random variable \( Y \) is defined as (7), then, the reliability \( R = \Pr(X, Y) \) for \( X \) and \( Y \) i.i.d is given by

\[
\theta^2 \sum_{k=0}^{\infty} \frac{(1-\theta)^k}{k! (k+2)}(3)
\]

**Proof.** For \( X \) and \( Y \) i.i.d CE2G r.v.'s where \( X \) is the stress and \( Y \) is the strength, the reliability \( R = \Pr(X < Y) \) is given by

\[
R = \int_{0}^{\infty} \int_{0}^{y} \frac{\alpha \lambda \theta e^{-\lambda x} (1-e^{-\lambda x})^{\alpha-1}}{[1 - (1-\theta)^2 (1-e^{-\lambda y})^\alpha]^2} \, dx \, dy
\]

\[
= \int_{0}^{\infty} \theta (1-e^{-\lambda y})^\alpha \frac{\alpha \lambda \theta e^{-\lambda y} (1-e^{-\lambda y})^{\alpha-1}}{[1 - (1-\theta)^2 (1-e^{-\lambda y})^\alpha]^2} \, dy
\]

\[
= \sum_{k=0}^{\infty} \theta^2 \lambda \theta \frac{(3)_k}{k!} (1-\theta)^k
\]

\[
= \sum_{k=0}^{\infty} \theta^2 \lambda \theta \frac{(3)_k}{k!} (1-\theta)^k
\]

So, using the (25) in \( y(\rho) \), the result follows. \[ \square \] This completes the proof. \[ \square \]
7. Residual Lifetime Distribution

Given that there was no failure prior to time $t$, the residual lifetime distribution of a random variable $X$, distributed as CE2G distribution, has the survival function given by

$$S_t(x) = \Pr [X > x + t | X > t] = \frac{1 - (1 - e^{-\lambda(x+t)})^\alpha}{1 - (1 - e^{-\lambda t})^\alpha} \times \frac{1 - (1 - \theta) (1 - e^{-\lambda t})^\alpha}{1 - (1 - \theta) (1 - e^{-\lambda(x+t)})^\alpha}.$$  

(28)

The mean residual lifetime of a continuous distribution with survival function $F(x)$ is given by

$$\mu(t) = E (X - t | X > t) = \frac{1}{S(t)} \int_t^\infty S(u) \, du.$$  

(29)

**Proposition 7.** For the random variable $Y$ with CE2G distribution, we have that the mean residual lifetime is given by

$$\mu(t) = \frac{1}{\lambda} \left( \frac{1 - (1 - \theta) (1 - e^{-\lambda t})^\alpha}{1 - (1 - e^{-\lambda t})^\alpha} \right) \times \sum_{i=0}^\alpha \sum_{j=0}^{\alpha-1} \frac{(1 - \theta)^j (-1)^i}{j!} \times \frac{1 - (1 - e^{-\lambda t})^{\alpha(i+j)+k+1}}{\alpha (i+j)+k+1}.$$  

(30)

**Proof.** From (29) and using $S(y)$ given by (2), we have that

$$\frac{1}{S(t)} \int_t^\infty S(u) \, du = \frac{1 - (1 - \theta) (1 - e^{-\lambda t})^\alpha}{1 - (1 - e^{-\lambda t})^\alpha} \times \int_t^\infty \frac{1 - (1 - e^{-\lambda u})^\alpha}{1 - (1 - \theta) (1 - e^{-\lambda u})^\alpha} \, du = \frac{1 - (1 - \theta) (1 - e^{-\lambda t})^\alpha}{\lambda} \times \int_1^X \frac{1 - X^\alpha}{1 - X^\alpha (1 - \theta) (1 - X)} \, dx.$$  

(31)

Now using (18) and making a binomial expansion in a similar way of the proof of Proposition 4 on (22), the result follows.

8. Inference

Assuming the lifetimes are independently distributed and are independent from the censoring mechanism, the maximum likelihood estimates (MLEs) of the parameters are obtained by direct maximization of the log-likelihood function given by

$$\ell(\theta, \lambda, \alpha) = \ln \left( \alpha \theta \lambda \right) \sum_{i=1}^n \beta_i \gamma_i + (\alpha - 1) \sum_{i=1}^n \beta_i \ln \left( 1 - e^{-\lambda y_i} \right) + \sum_{i=1}^n (1 - \beta_i) \ln \left( 1 - (1 - \theta) (1 - e^{-\lambda y_i})^\alpha \right) - \sum_{i=1}^n (1 + \beta_i) \ln \left( 1 - (1 - \theta) (1 - e^{-\lambda y_i})^\alpha \right),$$  

(32)

where $\beta_i$ is a censoring indicator, which is equal to 0 or 1, respectively, if the data is censored or observed. The advantage of this procedure is that it runs immediately using existing statistical packages. We have considered the optim routine of the R [16].

Large-sample inference for the parameters is based on the MLEs and their estimated standard errors. For $(\alpha, \theta, \lambda)$, we consider the observed Fisher information matrix given by

$$I_F(\alpha, \theta, \lambda) = \left( I_{aa} I_{a\theta} I_{a\lambda} I_{\theta\theta} I_{\theta\lambda} I_{\lambda\lambda} \right)_{(\alpha, \theta, \lambda) = (\hat{\alpha}, \hat{\theta}, \hat{\lambda})}.$$  

(33)

where the elements of the matrix $I_F(\alpha, \theta, \lambda)$ are given in the appendix.

Under conditions that are fulfilled for the parameters $\alpha, \theta$, and $\lambda$ in the interior of the parameter space, the asymptotic distribution of $(\hat{\alpha}, \hat{\theta}, \hat{\lambda})$, as $n \to \infty$, is a normal 3-variate with zero mean and variance covariance matrix $I_F^{-1}(\alpha, \theta, \lambda)$.

In order to compare different distributions, we relied upon several authors in the literature, for example, [6, 17–19], which use the Akaike information criterion (AIC) and Bayesian information criterion (BIC) values, which are defined, respectively, by $-2\ell(\cdot) + 2q$ and $-2\ell(\cdot) + q \log(n)$, where $\ell(\cdot)$ is the LogLikelihood evaluated in the MLE vector on respective distribution, $q$ is the number of parameters estimated, and $n$ is the sample size. The best distribution corresponds to a lower AIC and BIC values.

9. Simulation Study

Regarding the performance of the MLEs in the process of estimation, a study was performed based on one hundred generated dataset from the CE2G with six different sets of parameters for $n = 20, 50, 100, 200, 500$, and 1000. In order to have unbounded parameters, we consider the following restrictions on the parameters in estimation process.
the parameter $\theta$, we considered the transformation $\theta = e^{\theta^*} / (1 + e^{\theta^*})$, where $\theta^* \in \mathbb{R}$, and for $\alpha$ and $\lambda$ consider an exponential transformation. Based on the literature of the MLEs, we can return on the original parameters thought of the transformations. For the calculus of their variances, we use the delta method. The values $(\alpha, \lambda, \theta) = (1, 1, 0.5)$ were used as the initial values for all numerics simulations since $\lambda > 0$, $\alpha > 0$, and $0 < \theta < 1$.

The results are condensed in Table 1, which shows the averages of the MLEs, $\bar{\bar{\alpha}}, \bar{\bar{\lambda}}, \bar{\bar{\theta}}$, together with coverage probability of the 95% confidence intervals for parameters of the CE2G, $C(\alpha, \lambda, \theta)$, the bias, the mean squared error, MSE, and their deviance, $Sd(\bar{\bar{\alpha}}, \bar{\bar{\lambda}}, \bar{\bar{\theta}})$. These results suggest that the MLEs estimates have performed adequately. The deviance of the MLEs decrease when sample size increases. The empirical coverage probabilities are close to the nominal coverage level, particularly, as sample size increases.

10. Applications

In this section, we compare the CE2G distribution fit with several usual lifetime distributions on three datasets extracted

Figure 3: (a) Empirical TTT plot for the dataset $T_1$, $T_2$, and $T_3$, respectively. (b) Models fitting for the dataset $T_1$, $T_2$, and $T_3$, respectively.
<table>
<thead>
<tr>
<th>$n$</th>
<th>$\text{Av}(\hat{\alpha}, \hat{\lambda}, \hat{\theta})$</th>
<th>$\text{Sd}(\hat{\alpha}, \hat{\lambda}, \hat{\theta})$</th>
<th>Bias</th>
<th>MSE</th>
<th>$C(\alpha, \lambda, \theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>(0.510, 1.000, 0.600)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>(0.250, 0.500, 0.250)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>(0.125, 0.250, 0.125)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>(0.062, 0.125, 0.062)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>(0.031, 0.062, 0.031)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>(0.016, 0.031, 0.016)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.000, 0.000, 0.000)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>$\alpha = (1.48, 3.10, 0.75)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>(2.16, 1.16, 0.16)</td>
<td>(1.04, 0.94, 0.94)</td>
<td>(0.96, 0.96, 0.96)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>(1.85, 1.05, 0.10)</td>
<td>(0.94, 0.94, 0.94)</td>
<td>(0.96, 0.96, 0.96)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>(1.75, 1.00, 0.10)</td>
<td>(0.90, 0.90, 0.90)</td>
<td>(0.96, 0.96, 0.96)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>(1.65, 0.95, 0.09)</td>
<td>(0.85, 0.85, 0.85)</td>
<td>(0.96, 0.96, 0.96)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>(1.55, 0.90, 0.09)</td>
<td>(0.80, 0.80, 0.80)</td>
<td>(0.96, 0.96, 0.96)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>(1.45, 0.85, 0.09)</td>
<td>(0.75, 0.75, 0.75)</td>
<td>(0.96, 0.96, 0.96)</td>
<td>(0.99, 0.99, 0.99)</td>
<td></td>
</tr>
</tbody>
</table>

Table I: Mean of the MLEs, their deviances, coverages, bias and MSE.
from the literature. The first dataset, \( T1 \), refers to the serum-reversal time (days) of 143 children contaminated with HIV from vertical transmission at the university hospital of the Ribeirão Preto Scholl of Medicine (Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto) from 1986 to 2001 [20]. Serum reversal can occur in children born from mothers infected with HIV.

The second dataset, \( T2 \), is lifetimes in hours of 417 forty-watt, 110-volt internally frosted incandescent lamps taken from 42 weekly quality control [21]. Survival times, in days, are given for the control group of lamps on original dataset.

The third dataset, \( T3 \), gives the survival times for laboratory mice, which were exposed to a fixed dose of radiation at an age of 5 to 6 weeks. The cause of death for each mouse was determined after autopsy to be one of three possibilities: thymic lymphoma (\( C1 \)), reticulum cell sarcoma (\( C2 \)), or other causes (\( C3 \)) [22]. Consider here the minces of \( C3 \) in the control group.

Firstly, in order to identify the shape of a lifetime data failure rate function, we will consider a graphical method based on the TTT plot [23]. In its empirical version, the TTT plot is given by \( G(r/n) = [(\sum_{i=1}^{r} Y_{i:n}) + (n-r)Y_{r:n}]/(\sum_{i=1}^{n} Y_{i:n}) \), where \( r = 1, \ldots, n \) and \( Y_{i:n}, i = 1, \ldots, n \) represent the order statistics of the sample. It has been shown that the failure rate function is increasing (decreasing) if the TTT plot is concave (convex). Figure 3(a) shows concave TTT plots for the \( T1 \), \( T2 \), and \( T3 \) datasets, indicating increasing failure rate functions.

We compare the CE2G distribution fits with the exponential distribution with probability density function given by \( f(x) = \lambda e^{-\lambda x} \), the exponentiated exponential distribution, EE, with probability density function given by \( f(x) = \alpha \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{\alpha-1} \), the EG distribution [1] with probability density function given by \( f(x) = \lambda(1 - \theta + \theta e^{-\lambda x})^{-\alpha} \), the Weibull distribution with probability density function given by \( f(x) = (\theta/\lambda)(x/\lambda)^{\alpha-1} e^{-(x/\lambda)^{\alpha}} \), where the shape parameter is \( \alpha \) and scale parameter is \( \lambda \), the gamma distribution with probability density function given by \( f(x) = (1/\Gamma(\theta)) x^{\theta-1} e^{-x/\lambda} \), with shape parameter \( \theta \) and scale parameter \( \lambda \), the modified Weibull (MW) distribution with probability density function given by \( f(x) = \alpha \lambda e^{\beta x} e^{-\lambda e^{-\beta x}} \), where \( \alpha, \beta \geq 0 \) and \( \lambda > 0 \), the generalized exponential Poisson (GEP) distribution [6] with probability density function given by \( f(x) = (\alpha \beta / \lambda(1 - e^{-\beta x})^\alpha) (1 - e^{-\lambda x} e^{-\beta x})^{-\alpha} \), the generalized Birnbaum-Saunders (BS-G) distribution [24] with probability density function given by \( f(x) = (\sqrt{(y - \mu)/\beta + \sqrt{\beta/(x - \mu)})/2\alpha(\sqrt{(y - \mu)/\beta - \sqrt{\beta/(x - \mu)})}) \), where \( \phi(\cdot) \) is the probability density function of the standard normal distribution, and the Birnbaum-Saunders (BS) distribution. The BS distribution is obtained considering \( \mu = 0 \) in the BS-G probability density function.

Table 2 provides the AIC and BIC criterion values for each distribution. They provide evidence in favor of our CE2G distribution for the datasets \( T1 \) and \( T2 \) in all of the three comparison criterion. For the dataset \( T3 \), the CE2G distribution provides similar fitting to the Weibull and MW distributions, implying that the CE2G distribution is a competitor to the usual survival distributions. These results are corroborated by the empirical Kaplan-Meier survival functions and the fitted survival functions shown in Figure 3(b). The MLEs (and their corresponding standard errors in parentheses) of the parameters \( \alpha, \theta \times 1000 \), and \( \lambda \times 10000 \) of the CE2G distribution are given, respectively, by 3.7469 (0.5688), 41.4860 (9.7659), and 17536.46 (7.1814) for \( T1 \), by 5.1765 (19.4159), 0.2625 (0.9915), and 94.6676 (7.1870) for \( T2 \), and by 0.0018180 (0.9818), 0.0698 (0.3770), and 78.7704 (11.5084) for \( T3 \).

### 11. Concluding Remarks

In this paper, a new lifetime distribution is provided and discussed. The CE2G distribution accommodates increasing, decreasing, and bathtub failure rate functions and arises in a latent complementary risks scenario, where the lifetime associated with a particular risk is not observable but only the maximum lifetime value among all risks. The properties of the proposed distribution are discussed, including a formal proof of its probability density function and explicit algebraic formulas for its survival and hazard functions, moments, rth moment of the rth order statistic, mean residual lifetime, modal value, and the observed Fisher information matrix. Maximum likelihood inference is implemented straightforwardly. The practical importance of the new distribution was demonstrated in three applications where the CE2G distribution provided the best fit in comparison with several other former lifetime distributions.
Appendix

In this appendix, we show the values of the elements of the observed Fisher information matrix in (33). From (32), we obtain

\[
I_{\alpha \alpha} = \sum_{i=1}^{n} \left( \frac{c_i}{\alpha^2} + \frac{(1-c_i)L_i^a \ln^2(L_i)}{R_i} + \frac{(1-c_i)L_i^{2a} \ln^2(L_i)}{R_i^2} \right) + \frac{(1+c_i)(1-\theta)L_i^a \ln^2(L_i)}{T_i} + \frac{(1+c_i)(1-\theta)^2L_i^{2a} \ln^2(L_i)}{T_i^2},
\]

\[
I_{\alpha \beta} = I_{\beta \alpha} = \sum_{i=1}^{n} \left( \frac{(1+c_i)L_i^a \ln(L_i)}{T_i} + (1+c_i)(1-\theta)L_i^{2a} \ln(L_i) \right),
\]

\[
I_{\alpha \lambda} = I_{\lambda \alpha} = \sum_{i=1}^{n} \left( \frac{-c_i X_i}{L_i} + \frac{\alpha(1-c_i)L_i^a \ln(L_i) X_i}{L_i R_i} + \frac{(1-c_i)L_i^a X_i}{L_i R_i} - \frac{\alpha(1+c_i)(1-\theta)L_i^a \ln(L_i) X_i}{L_i T_i} - \frac{(1+c_i)(1-\theta)L_i^{2a} X_i}{L_i T_i} - \frac{\alpha(1+c_i)(1-\theta)^2L_i^{2a} \ln(L_i) X_i}{L_i T_i^2} \right),
\]

\[
I_{\beta \beta} = I_{\lambda \lambda} = \sum_{i=1}^{n} \left( \frac{c_i}{\beta^2} + \frac{(1+c_i)L_i^{2a}}{T_i^2} \right),
\]

\[
I_{\beta \lambda} = I_{\lambda \beta} = \sum_{i=1}^{n} \left( \frac{\alpha(1+c_i)L_i^a X_i}{L_i T_i} + \frac{\alpha(1+c_i)(1-\theta)L_i^{2a} X_i}{L_i T_i^2} \right),
\]

\[
I_{\lambda \lambda} = \sum_{i=1}^{n} \left( \frac{c_i}{\lambda^2} + \frac{(\alpha-1)c_i y_i X_i}{L_i} + \frac{(\alpha-1)c_i X_i^2}{L_i^2} - \frac{\alpha(1-c_i)L_i^a y_i X_i}{L_i R_i} - \frac{\alpha(1-c_i)L_i^a X_i^2 (1-\alpha)}{L_i^2 R_i} + \frac{\alpha^2(1-c_i)L_i^{2a} X_i^2}{L_i^2 R_i^2} + \frac{(1+c_i)(1-\theta)L_i^a y_i X_i}{L_i T_i} + \frac{(1+c_i)(1-\theta)L_i^{2a} X_i^2}{L_i T_i^2} \right),
\]

(A.1)

where \( L_i = 1 - e^{-\lambda y_i} \), \( R_i = 1 - L_i^a \), \( T_i = 1 - (1-\theta)L_i^a \), and \( X_i = y_i e^{-\lambda y_i} \).

Acknowledgments

V. Marchi and F. Louzada are supported by the Brazilian organizations CAPES and CNPq, respectively. The authors are grateful to Dr. Gauss Cordeiro, Editor of this special issue in, as well as to the anonymous Referees for their comments, criticisms, and suggestions, which lead to important improvements.

References


