Communications for Statistical Applications and Methods
2017, Vol. 24, No. 4, 325–337

https://doi.org/10.5351/CSAM.2017.24.4.325
Print ISSN 2287-7843 / Online ISSN 2383-4757

# Application of the Weibull-Poisson long-term survival model

Valdemiro Piedade Vigas[a], Josmar Mazucheli[1,b], Francisco Louzada[c]

[a]Instituto de Matemática, Universidade Federal do Mato Grosso do Sul, Brazil;
[b]Departamento de Estatística, Universidade Estadual de Maringá, Brazil;
[c]Institute of Mathematical Science and Computing, Universidade de São Paulo, Brazil

**Abstract**

In this paper, we proposed a new long-term lifetime distribution with four parameters inserted in a risk competitive scenario with decreasing, increasing and unimodal hazard rate functions, namely the Weibull-Poisson long-term distribution. This new distribution arises from a scenario of competitive latent risk, in which the lifetime associated to the particular risk is not observable, and where only the minimum lifetime value among all risks is noticed in a long-term context. However, it can also be used in any other situation as long as it fits the data well. The Weibull-Poisson long-term distribution is presented as a particular case for the new exponential-Poisson long-term distribution and Weibull long-term distribution. The properties of the proposed distribution were discussed, including its probability density, survival and hazard functions and explicit algebraic formulas for its order statistics. Assuming censored data, we considered the maximum likelihood approach for parameter estimation. For different parameter settings, sample sizes, and censoring percentages various simulation studies were performed to study the mean square error of the maximum likelihood estimative, and compare the performance of the model proposed with the particular cases. The selection criteria Akaike information criterion, Bayesian information criterion, and likelihood ratio test were used for the model selection. The relevance of the approach was illustrated on two real datasets of where the new model was compared with its particular cases observing its potential and competitiveness.

Keywords: competing risks, likelihood, long-term, Weibull-Poisson distribution, survival analysis

## 1. Introduction

Survival data in presence of competing risks arise in several areas, such as public health, actuarial science, biomedical studies, demography, and industrial reliability. A scenario classical of the competing risks occurs when there is no information about which risk was responsible for the component failure (or individual death) and only the minimum lifetime value among all risks is observed. This information is not accessible in many situations or it is impossible to specify the true cause of failure. This sometimes occurs when there is an interest to observe the lifetime of a system in series. In this case, the lifetime duration depends on a set of components.

In recent years, several authors proposed probability distributions which properly accommodate survival data in the presence of latent competing risks. For example, Adamidis and Loukas (1998) proposed a compounding distribution, denoted by exponential geometric (EG) distribution, which properly accommodate survival data in the presence of latent competing risks. Kuş (2007) proposed

[1] Corresponding author: Departamento de Estatística, Universidade Estadual de Maringá, Ave. Colombo, 5.790 Jd. Universitário Maringá, Paraná, CEP 87020-900, Brazil. E-mail: jmazucheli@gmail.com

Published 31 July 2017 / journal homepage: http://csam.or.kr

another compounding distribution that properly accommodates survival data in the presence of latent competing risks and exponential-Poisson distribution (EP). Tahmasbi and Rezaei (2008) introduced the logarithmic exponential distributions. Chahkandi and Ganjali (2009) introduced exponential power series that contains the distributions cited EG, EP, and logarithmic exponential as special cases. Louzada *et al.* (2011) proposed a complementary EG distribution that properly accommodates survival data in the presence of latent complementary risks. More details can be found in: Marshall and Olkin (1997), Louzada-Neto (1999), Crowder (2001), Lu and Tsiatis (2005), Pintilie (2006), and Khan and King (2016).

Other characteristic in survival data arises when part of the population is not susceptible to the event of interest and is considered as immune or cured. Models which consider that one part of the population is cured have been widely developed and are usually called long-term survival models. For instance, a population can respond favorably to a treatment in clinical studies and therefore be considered cured. Perhaps the most popular type of long-term model was introduced by Boag (1949) and Berkson and Gage (1952). In this model, it is assumed that a certain proportion of the patients, say $p$, are cured in the sense that they do not present the event of interest over a long time and can be seen as immune to the cause of failure under study. Maller and Zhou (1995), Yakovlev *et al.* (1996), Chen and Ibrahim (2001), Pons and Lemdani (2003), Cooner *et al.* (2006), Perperoglou *et al.* (2007), Perdoná and Louzada-Neto (2011), and Ortega *et al.* (2017) later considered long-term mixture modeling. Mazucheli *et al.* (2012) proposed the polysurvival model with long-term survivors when there are situations in which a fraction of systems is subject to failure from independent competing causes of failure, while the remaining proportion $p$ is cured or has not presented the event of interest during the duration of the study.

In this paper, we proposed a Weibull-Poisson (WP) long-term model. This new model is based on the WP distribution, Bereta *et al.* (2011), in a structure of modeling long-term. The WP distribution arises on the competitive risk scenario, in which the lifetime associated with a particular risk is not observable, rather only the minimum lifetime value among all risks is noticed. However, this model can be used in any other situation as long as it fits the data well. We now call this situation as WP long-term distribution or long-term Weibull-Poisson (LWP) distribution.

The new model, due to its flexibility in accommodating various forms of the risk function, seems to be an important model that can be used in a variety of problems for survival data modeling in a latent competing risk scenario with long-term implications. In addition to that, the LWP model is also suitable to test goodness-of-fit of some special sub-models, such as new exponential-Poisson long-term distribution (LEP) and Weibull long-term distribution (LW). We demonstrate, by means of an application to real data, that the LWP model can produce better fits than some other known models. It therefore represents a good alternative for lifetime data analysis. We hope this generalization may attract wider applications in survival analysis. The inferential part of this model is carried out using the asymptotic distribution of maximum likelihood estimators. Considering that the LWP model is embedded in LEP and LW, the likelihood ratio test (LR) can be used to discriminate such models. Studies were conducted via Monte Carlo simulation to evaluate the performance of LWP distribution through means, mean squared error (MSE) for the maximum likelihood estimates (MLEs), power of the test, Akaike information criterion (AIC), and Bayesian information criterion (BIC) for model selection.

This paper is organized as follows. Section 2 introduces the LWP distribution and presents some forms of its hazard rate function. We also described its $k^{th}$ order statistics. Section 3 presents an inferential procedure based on a maximum likelihood approach; in addition, the selection criteria of AIC, BIC, and the LR test help select the best model. Section 4 presents the results of a simulation

study conducted to assess the performance of the new model. Section 5 compares the new distribution with its particular cases on two real datasets, observing its potential and competitiveness. The last section presents the conclusion and some final comments.

## 2. Formulation of the long-term Weibull-Poisson distribution

Let $T$, a non-negative random variable, represent the lifetime of an individual in a population that has non cured individuals (susceptible) with probability $1-p$ and cured individuals (not susceptible) with probability $p$. Thus, given population survival function, $S_{\text{pop}}(t)$, we have that $\lim_{t\to\infty} S_{\text{pop}}(t) = p$, wherein $p$ is the proportion of individuals not susceptible to the event of interest. The model Berkson and Gage (1952) is characterized by the survival function:

$$S_{\text{pop}}(t) = p + (1-p)S_0(t), \tag{2.1}$$

where $S_0(t)$ is the survival function for individuals who are under risk, such that for $\lim_{t\to\infty} S_0(t) = 0$, then $\lim_{t\to\infty} S_{\text{pop}}(t) = p$ (improper survival function). The function $S_0(t)$ can be specified by the usual survival functions such as a Weibull, and exponential, among others. In this work we used the survival function of the WP distribution (Bereta *et al.*, 2011), given by:

$$S_0(t) = \frac{\exp\{\alpha \exp[-(\beta t)^\gamma]\} - 1}{\exp(\alpha) - 1}. \tag{2.2}$$

WP distribution results from latent competing risk scenarios, in which the lifetime associated with a particular risk is not observable, and only the minimum lifetime value among all risks is noticed. Application of WP distribution in survival studies were investigated by Bereta *et al.* (2011). Replacing (2.2) of (2.1) the improper survival function, $S_{\text{pop}}(t)$, can be written as:

$$S_{\text{pop}}(t) = \frac{p\exp(\alpha) - p + (\exp\{\alpha \exp[-(\beta t)^\gamma]\}) - 1 - p(\exp\{\alpha \exp[-(\beta t)^\gamma]\}) + p}{\exp(\alpha) - 1}, \tag{2.3}$$

where $t > 0$, $\alpha > 0$, $\beta > 0$, $\gamma > 0$, and $0 < p < 1$. The probability density function (pdf), $f_{\text{pop}}(t)$, is obtained directly considering that $f_{\text{pop}}(t) = -dS_{\text{pop}}(t)/dt$ and is given by

$$f_{\text{pop}}(t) = \frac{(1-p)\alpha \exp\{\alpha \exp[-(\beta t)^\gamma] - (\beta t)^\gamma\}\beta^\gamma t^{\gamma-1}\gamma}{\exp(\alpha) - 1}. \tag{2.4}$$

The pdf (2.4) is defined as the LWP distribution, LWP($\boldsymbol{\theta}$), where $\boldsymbol{\theta} = (\alpha, \beta, \gamma, p)$. The hazard function, $h_{\text{pop}}(t)$, is obtained by the relation $h_{\text{pop}}(t) = S_{\text{pop}}(t)/f_{\text{pop}}(t)$, and is given by

$$h_{\text{pop}}(t) = \frac{(1-p)\alpha \exp\{\alpha \exp[-(\beta t)^\gamma] - (\beta t)^\gamma\}\beta^\gamma t^{\gamma-1}\gamma}{p\exp(\alpha) - p + (\exp\{\alpha \exp[-(\beta t)^\gamma]\}) - 1 - p(\exp\{\alpha \exp[-(\beta t)^\gamma]\}) + p}. \tag{2.5}$$

Figure 1 illustrates some of the possible shapes of the hazard function $h_{\text{pop}}(t)$ for the selected parameter values. This figure indicates that the hazard function is flexible and can accommodate various forms, such as increasing, decreasing, and unimodal. For example, the risk function is decreasing for gamma values less than 1. For values greater than 6, the risk function is increasing.

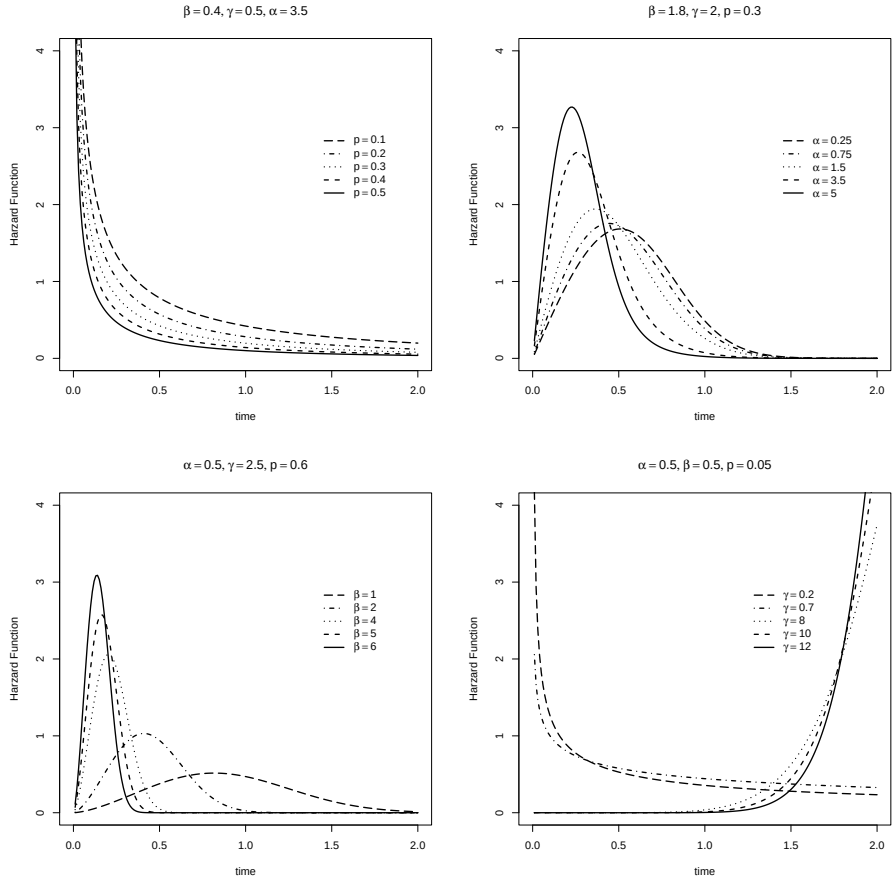


Figure 1: *Plots of the hazard function for the Weibull-Poisson long-term distribution.*

## 2.1. Characteristics of the long-term Weibull-Poisson distribution

LWP distribution opens new possibilities for several types of fitted data. It is observed that when $\alpha \to 0$ in equation (2.3) the LWP distribution is reduced to the LW. For $\gamma = 1$ in equation (2.3) the LWP distribution is reduced to new the LEP.

Order statistics play an important role in quality control testing and reliability, where a practitioner needs to predict the failure of future items based on how many times develop early failure. These predictors are often based on moments of order statistics (Louzada *et al.*, 2014). Considering LWP distribution, the pdf of $k^{th}$ order statistic is given as follows.

**Proposition 1.** *Let $T_1, T_2, \ldots, T_n$ be the iid random variable such that $T_k \sim LWP(\boldsymbol{\theta})$ for $k = 1, 2, \ldots, n$. The pdf of the $k^{th}$ order statistic, say $T_{k:n}$, is given by (for $t > 0$)*

$$f_{k:n}(t) = g_{k:n}(t)(1-p)^k \left[ \frac{p\exp(\alpha) + \exp\left\{\alpha \exp\left\{-(\beta t)^\gamma\right\}\right\}(1-p) - 1}{\exp\left\{\alpha \exp\left\{-(\beta t)^\gamma\right\}\right\} - 1} \right]^{n-k},$$

*where $g_{k:n}(t)$ is order statistic of WP with parameters $\alpha$, $\beta$ and $\gamma$.*

**Proof**: We now derive an explicit expression for the density of the $i^{th}$ order statistic $T_{k:n}$, say $f_{k:n}(t)$, in a random sample of size $n$ from the LWP distribution. It is known that

$$f_{k:n}(t) = \frac{1}{B(k, n-k+1)} f(t)(F(t))^{k-1}(S(t))^{n-k},$$

where $B(k, n-k+1) = (n-k)!(k-1)!/n!$. Using the definition, we have,

$$\begin{aligned}
f_{k:n}(t) &= \frac{1}{B(k, n-k+1)} f_{\text{pop}}(t)(F_{\text{pop}}(t))^{k-1}(S_{\text{pop}}(t))^{n-k} \\
&= \frac{1}{B(k, n-k+1)} \frac{(1-p)\alpha \exp\{\alpha \exp[-(\beta t)^\gamma] - (\beta t)^\gamma\}\beta^\gamma t^{\gamma-1}\gamma}{\exp(\alpha)-1} \left[\frac{(z(t)-\exp(\alpha))(-1+p)}{\exp(\alpha)-1}\right]^{k-1} \\
&\quad \times \left[\frac{p\exp(\alpha) - p + (z(t)) - 1 - p(z(t)) + p}{\exp(\alpha)-1}\right]^{n-k} \\
&= \frac{\alpha \exp\{\alpha \exp[-(\beta t)^\gamma] - (\beta t)^\gamma\}\beta^\gamma t^{\gamma-1}\gamma(1-p)^k}{B(k, n-k+1)(\exp(\alpha)-1)^n} [z(t) - \exp(\alpha)]^{k-1} \\
&\quad \times [p\exp(\alpha) + (z(t))(1-p) - 1]^{n-k} \\
&= \alpha \exp\{\alpha \exp[-(\beta t)^\gamma] - (\beta t)^\gamma\}\beta^\gamma t^{\gamma-1}\gamma [\exp(\alpha) - z(t)]^{k-1} \\
&\quad \times \frac{[z(t)-1]^{n-k}}{B(k, n-k+1)[\exp(\alpha)-1]^n}(1-p)^k \left[\frac{p\exp(\alpha) + (z(t))(1-p) - 1}{z(t)-1}\right]^{n-k}.
\end{aligned}$$

But the order statistic of a WP distribution is given by,

$$g_{k:n}(t) = \alpha \exp\{\alpha \exp[-(\beta t)^\gamma] - (\beta t)^\gamma\}\beta^\gamma t^{\gamma-1}\gamma [\exp(\alpha) - z(t)]^{k-1} \times \frac{[z(t)-1]^{n-k}}{B(k, n-k+1)[\exp(\alpha)-1]^n},$$

where $z(t) = \exp\{\alpha \exp[-(\beta t)^\gamma]\}$. Thus, we have the result. □

## 3. Inference

### 3.1. Estimation by maximum likelihood

Let $t_1, t_2, \ldots, t_n$ a random sample of size $n$ of the LWP($\boldsymbol{\theta}$) distribution, where $\boldsymbol{\theta} = (\alpha, \beta, \gamma, p)$. Assuming that the lifetimes $t_i$ and censoring are independent, the lifetimes are identically distributed and that censorship is not informative, the log-likelihood function of $\boldsymbol{\theta}$, $\ell(\boldsymbol{\theta})$ can be written as

$$\begin{aligned}
\ell(\boldsymbol{\theta}) \propto &\sum_{i=1}^{n} \delta_i \log\left\{\alpha\gamma\beta^\gamma t_i^{\gamma-1}(1-p)\right\} + \sum_{i=1}^{n} \delta_i [\alpha \exp\{-(\beta t_i)^\gamma\} - (\beta t_i)^\gamma] - \sum_{i=1}^{n} \delta_i \log\{\exp(\alpha) - 1\} \\
&+ \sum_{i=1}^{n} (1-\delta_i) \log[\exp(\alpha)(z(t))(1-p) - 1] - \sum_{i=1}^{n} (1-\delta_i) \log\{\exp(\alpha) - 1\}, \qquad (3.1)
\end{aligned}$$

where $z(t) = \exp\{\alpha \exp[-(\beta t_i)^\gamma]\}$. $\delta_i$ is a censoring indicator variable, which is equal to 0 or 1, respectively, if the data is censured or observed.

MLEs for parameter vector $\boldsymbol{\theta} = (\alpha, \sigma, \gamma, p)$ can be obtained by maximizing the log-likelihood function (3.1) and solving the system of equations given by:

$$U(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0.$$

The components of the score vector $U(\boldsymbol{\theta})$ are given by

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \frac{\sum_{i=1}^{n} \delta_i}{\alpha} + \sum_{i=1}^{n} \delta_i z(t_i) - \frac{\sum_{i=1}^{n} \delta_i \exp(\alpha)}{\sum_{i=1}^{n}(1-\delta_i)\exp(\alpha) - 1}$$
$$+ \frac{\exp\{\alpha v(t_i) - (\beta t_i)^{\gamma}\}\beta^{\gamma} t_i^{\gamma-1}(1-p)p\exp(\alpha) + v(t_i)(1-p)p\exp(\alpha)}{z(t_i)(1-p)p(\exp(\alpha)-1)} + \frac{\sum_{i=1}^{n}(1-\delta_i)}{\exp(\alpha)-1},$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta} = \frac{\sum_{i=1}^{n} \delta_i \gamma}{\beta} + \sum_{i=1}^{n} \delta_i \left( \frac{-\alpha(\beta t_i)^{\gamma}\gamma v(t_i)}{\beta} - \frac{(\beta t_i)^{\gamma}\gamma}{\beta} \right) \frac{-\sum_{i=1}^{n}(1-\delta_i)\alpha(\beta t_i)^{\gamma}\gamma v(t_i) z(t_i)(1-p)p\exp(\alpha)}{\beta[z(t_i)(1-p)p(\exp(\alpha)-1)]},$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \gamma} = \frac{\sum_{i=1}^{n} \delta_i [\alpha\beta^{\gamma}(t_i)^{\gamma-1}(1-p) + \alpha\gamma\beta^{\gamma}\log(\beta) t_i^{\gamma-1}(1-p) + \alpha\gamma\beta^{\gamma} t_i^{\gamma-1}\log(t_i)(1-p)]}{\alpha\gamma\beta^{\gamma} t_i^{\gamma-1}(1-p)} +$$
$$+ \sum_{i=1}^{n} \delta_i \left[-\alpha(\beta t_i)^{\gamma}\log(\beta t_i) v(t_i) - (\beta t_i)^{\gamma}\log(\beta t_i)\right]$$
$$\times \frac{-\sum_{i=1}^{n}(1-\delta_i)\alpha(\beta t_i)^{\gamma}\log(\beta t_i) v(t_i) z(t_i)(1-p)p\exp(\alpha)}{z(t_i)(1-p)p(\exp(\alpha)-1)},$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial p} = \frac{\sum_{i=1}^{n} \delta_i}{1-p} + \frac{\sum_{i=1}^{n}(1-\delta_i)\left(-\exp\{\alpha v(t_i)\}\exp(\alpha)p + \exp\{\alpha\exp v(t_i)(1-p)\exp(\alpha)\}\right)}{z(t_i)(1-p)p(\exp(\alpha)-1)},$$

where $z(t_i) = \exp\{\alpha\exp[-(\beta t_i)^{\gamma}]\}$; $v(t_i) = \exp[-(\beta t_i)^{\gamma}]$.

We can use numerical methods to solve the system of equations because there is no closed analytical method to find the estimators. The estimates of these parameters were therefore obtained by numerical methods, using an iterative process. We used the command *optim* in software R through of the method BFGS (Broyden-Fletcher-Goldfarb-Shanno). The inference procedure for $\boldsymbol{\theta}$ can be based on standard asymptotic likelihood techniques. Approximate confidence intervals and hypothesis tests on model parameters can be conducted using multivariate normal distribution with covariance matrix $\mathbf{I}(\boldsymbol{\theta})^{-1}$, where $\mathbf{I}(\boldsymbol{\theta})$ is the expected information matrix. The approximation $\hat{\boldsymbol{\theta}} \sim N_k(\boldsymbol{\theta}, -\ddot{\mathbf{L}}_{\theta\theta})$, where $k$ is the number of parameters of the model. $\ddot{\mathbf{L}}_{\theta\theta} = \partial^2 l(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T$ is the matrix of the second derivatives of the log likelihood (assumed evaluated at $\boldsymbol{\theta}$) and $\mathbf{I}(\boldsymbol{\theta}) = -E(\ddot{\mathbf{L}}_{\theta\theta})$, can be used instead of the Fisher information matrix when we have censored (censoring is random and non-informative) observations.

## 3.2. Model selection

For the selection of the model that best fits the data were used the AIC model selection criteria, BIC, and the LR. The AIC and BIC are defined by

$$\text{AIC} = -2\log(L) + 2k; \qquad \text{BIC} = -2\log(L) + k\log(n),$$

where $L$ the MLEs, $k$ is the number of parameters in the model and $n$ is the size sample. The preferred model is the one with the smallest value on each criterion.

As LWP distribution is reduced in the LEP and LW distributions may be used the LR. LR can be used to discriminate when testing nested models. We can compute the maximum values of the unrestricted and restricted log-likelihoods to construct LR statistics for testing some sub-models. The statistic ($\omega_n$) converges to a chi-square distribution with degrees of freedom equal to the difference between the number of parameters in the two models. However, for comparison of a nonnested survival model, under certain conditions of regularity, the distribution of the statistical likelihood ratio

under $H_0$ is a mixture with a weights (0.5 and 0.5) of distribution $\chi^2$ with a degree of freedom with a discrete distribution with the mass concentrated in the value 0, this is, $P(\omega_n \leq w) = 1/2 + 1/2P(\chi^2_1 \leq w)$. More details are found in Maller and Zhou (1995) and Cancho *et al.* (2011).

## 4. Simulation study

To examine the performance of the LWP model compared with LEP and LW models and assess the performance of MLEs for the parameters of the new model, a simulation study was done for sample sizes (40, 100, 400, 600, and 800) with 40% and 60% censored observations in each sample that generated 1,000 random samples. In this study, the survival time of $Y$ follows LWP distribution with a density function given by (2.4), where the values of the $Y$ were generated from inverse transformation method. The times of censoring $C$ is a random variable exponential with parameter $\lambda = 0.2$ where $\lambda$ was adjusted until censoring percentages, 0.40 or 0.60, were reached. Besides, the survival time $t_i$ and of censure in the simulation were considered through $t_i = \min(y_i, c_i)$, $i = 1, 2, \ldots, n$. Samples of different sizes were generated and the parameter values were fixed in $\alpha = 3.5, \beta = 1.5, \gamma = 1.2$. The values of parameter $p$ was fixed in 0.20 and 0.30, respectively. Below is the process of this simulation:

1. Generate $u_i \sim U(0, 1)$
2. Generate $y_i$, $i = 1, 2, \ldots, n$, such that:

$$y_i = \begin{cases} \infty, & \text{if } u_i \leq 1 - p, \\ F^{-1}(u_i) = \dfrac{1}{\beta}\left[\ln(\alpha) - \ln\left(\ln(\exp(\alpha)(1 - p - u_i) + u_i) - \ln(1 - p)\right)\right]^{\frac{1}{\gamma}}, & \text{if } u_i > 1 - p, \end{cases}$$

   where $F(\cdot)$ is the cumulative distribution function of the LWP distribution
3. Generate variable of censure $c_i$; $c_i \sim \exp(0.2)$
4. Find $t_i = \min(t_i, c_i)$
5. If $y_i > c_i$ or $y_i = \infty$, then $\delta_i = 1$, otherwise, $\delta_i = 0$, to $i = 1, \ldots, n$.

For each combination of $n$, parameter values and censoring percentages, 1,000 samples were generated in order to obtain the MLEs for the LWP model. In addition, the likelihood-ratio test was performed for each fit, through the hypotheses $H_0 : \gamma = 1$ versus $H_1 : \gamma \neq 1$, which is equivalent in comparing the LWP model with the LEP model. All statistics were compared with the $\chi^2_1$ critical value at a significance level of 5%. It also realized the procedure for the analysis of the LWP and LW models, but in this case the hypotheses for the test were $H_0 : \alpha \to 0$ versus $H_1 : \alpha > 0$. The $95^{th}$ percentile of this distribution represented by $w_{0.95}$, is such that $1/2 + 1/2P(\chi^2_1 \leq w_{0.95}) = 0.95$, so that $w_{0.95} = 2.705543$. Therefore, it rejects $H_0$ at a significance level of 5% to $\omega_n > 2.7055$.

Table 1 shows the proportion of times that the AIC and BIC of the LWP model was less than the LW and LEP models. Furthermore, it was also calculated the power of the test with 5% of the significance for different sample sizes, percentages of censored observations, and different parameter values of the new model. Table 1 indicates that the AIC proportion of the LWP model was lower than the LW model as well as between 35.2% and 82.5% to $n$ bigger than 100. In addition, higher proportions of the AIC are in the scenario in which parameter $p$ is 0.20, in particular, when $n$ increased and/or % censoring decreased.

In relation to BIC, the proportion varied between 0.15% and 46.60%. For $n > 100$ the proportion of BIC increased to around 19% and reduced its variability. However, lower values of the AIC

Table 1: Proportion of the AIC, BIC, and power of the test of LWP, LW, and LEP models for different values of $n$ and censored observations

| $p$ | $n$ | % of censorship | LW model | | | LEP model | | |
|---|---|---|---|---|---|---|---|---|
| | | | AIC | BIC | Power of the test | AIC | BIC | Power of the test |
| 0.20 | 40 | 40% | 0.089 | 0.015 | 0.048 | 0.217 | 0.002 | 0.142 |
| | | 60% | 0.052 | 0.010 | 0.070 | 0.270 | 0.023 | 0.168 |
| | 100 | 40% | 0.298 | 0.045 | 0.213 | 0.317 | 0.009 | 0.197 |
| | | 60% | 0.019 | 0.026 | 0.132 | 0.233 | 0.027 | 0.298 |
| | 400 | 40% | 0.670 | 0.216 | 0.636 | 0.687 | 0.071 | 0.531 |
| | | 60% | 0.409 | 0.205 | 0.372 | 0.549 | 0.102 | 0.404 |
| | 600 | 40% | 0.766 | 0.257 | 0.745 | 0.612 | 0.178 | 0.508 |
| | | 60% | 0.465 | 0.212 | 0.448 | 0.532 | 0.270 | 0.470 |
| | 800 | 40% | 0.825 | 0.258 | 0.758 | 0.849 | 0.400 | 0.783 |
| | | 60% | 0.748 | 0.232 | 0.449 | 0.803 | 0.303 | 0.709 |
| 0.30 | 40 | 40% | 0.118 | 0.001 | 0.006 | 0.230 | 0.005 | 0.083 |
| | | 60% | 0.109 | 0.005 | 0.005 | 0.165 | 0.071 | 0.170 |
| | 100 | 40% | 0.296 | 0.002 | 0.225 | 0.263 | 0.005 | 0.108 |
| | | 60% | 0.209 | 0.011 | 0.171 | 0.144 | 0.018 | 0.100 |
| | 400 | 40% | 0.750 | 0.199 | 0.632 | 0.622 | 0.050 | 0.485 |
| | | 60% | 0.352 | 0.190 | 0.403 | 0.532 | 0.069 | 0.470 |
| | 600 | 40% | 0.759 | 0.349 | 0.701 | 0.667 | 0.227 | 0.487 |
| | | 60% | 0.355 | 0.204 | 0.409 | 0.618 | 0.213 | 0.462 |
| | 800 | 40% | 0.819 | 0.466 | 0.733 | 0.609 | 0.231 | 0.768 |
| | | 60% | 0.550 | 0.346 | 0.549 | 0.575 | 0.470 | 0.596 |

AIC = Akaike's information criterion; BIC = Bayesian information criterion; LWP = Weibull-Poisson long-term, LW = Weibull long-term; LEP = exponential-Poisson long-term.

occurred when the size of $n$ was smaller than 400 and independent of the scenarios. In addition, the highest BIC proportions were in the scenario in which the parameter $p = 0.20$, in particular, when $n$ increased and/or % censoring decreased. It was noted that the AIC selection criteria was better than the BIC for and LWP model independent of the size of $n$ and % censored observations. The power of the test indicated that the values vary between 40.9% and 75.8% to $n$ bigger than 400. The analysis of the LWP and LEP models showed that the AIC proportion of the LWP model was lower than the LEP model in values between 14.40% and 84.9%.

The higher proportions of the AIC was in the scenario in which the parameter $p = 0.20$ and $n > 100$, with a minimum proportion of 57.5%. However, lower AIC values occurred when the size of $n$ was smaller than 100 and independent of the scenarios. However, the proportion varied between 0.2% and 47.00% in relation to BIC. For $n > 400$ the proportion increased to around 21.3% and reduced its variability. In addition, the highest BIC proportions were when $p = 0.20$, in particular, when $n$ increased. It was also noted that the AIC selection criteria was better than the BIC for a LWP model independent of the size of $n$ and % censored observations.

In relation to the power of the test, the values varied between 8.3% and 78.3%. However, the power of the minimum test was 40.4% when $n > 100$. Lower values of AIC, BIC, and the power of the test occurred when the size of $n$ was smaller than 400 and independent of the scenarios. The MSE of MLE's were also calculated for simulated samples under conditions similar to those used in previous simulations.

A simulation study was then done with times of censure $C$ as a random variable exponentially distributed with parameter $\lambda = 2$. Parameter values were fixed in $\alpha = 3.5$, $\beta = 0.5$, $\gamma = 2$, $p = 0.20$, and 0.30. For each combination of parameter values, with censoring percentages, 1,000 samples were generated in order to obtain the MLEs of LWP distribution. The simulation results showed that the

Table 2: Mean and mean squared error (whithin parenthesis) of the estimates of the parameters of Weibull-Poisson long-term model with $\alpha = 3.5, \beta = 0.5$, and $\gamma = 2$

| $n$ | Parameters | $p = 0.20$ | | $p = 0.30$ | |
|---|---|---|---|---|---|
| | | 40% | 60% | 40% | 60% |
| 40 | $\alpha$ | 0.776 (9.061) | 0.798 (8.817) | 0.612 (9.686) | 0.573 (9.837) |
| | $\beta$ | 0.839 (0.179) | 0.859 (0.206) | 0.900 (0.242) | 0.989 (0.313) |
| | $\gamma$ | 1.610 (0.259) | 1.731 (0.271) | 1.479 (0.373) | 1.653 (0.262) |
| | $p$ | 0.128 (0.009) | 0.281 (0.015) | 0.175 (0.020) | 0.380 (0.009) |
| 100 | $\alpha$ | 0.614 (9.856) | 0.813 (9.098) | 2.112 (7.107) | 0.420 (10.673) |
| | $\beta$ | 0.845 (0.159) | 0.854 (0.178) | 1.883 (0.631) | 0.991 (0.282) |
| | $\gamma$ | 1.511 (0.273) | 1.621 (0.202) | 1.549 (0.234) | 1.530 (0.271) |
| | $p$ | 0.135 (0.005) | 0.295 (0.011) | 0.265 (0.001) | 0.386 (0.008) |
| 400 | $\alpha$ | 2.329 (9.122) | 1.561 (10.125) | 1.512 (10.343) | 0.608 (14.399) |
| | $\beta$ | 0.799 (0.138) | 0.755 (0.125) | 0.873 (0.201) | 0.953 (0.244) |
| | $\gamma$ | 1.494 (0.268) | 1.604 (0.177) | 1.453 (0.532) | 1.493 (0.274) |
| | $p$ | 0.137 (0.004) | 0.300 (0.010) | 0.143 (0.001) | 0.388 (0.007) |
| 600 | $\alpha$ | 1.455 (6.859) | 2.074 (7.477) | 0.563 (8.688) | 1.563 (10.226) |
| | $\beta$ | 0.737 (0.117) | 0.697 (0.101) | 0.898 (0.200) | 0.795 (0.165) |
| | $\gamma$ | 1.508 (0.251) | 1.612 (0.162) | 1.367 (0.407) | 1.475 (0.290) |
| | $p$ | 0.137 (0.004) | 0.302 (0.010) | 0.193 (0.001) | 0.377 (0.006) |
| 800 | $\alpha$ | 5.596 (6.045) | 1.845 (6.905) | 2.834 (4.214) | 6.408 (9.554) |
| | $\beta$ | 1.103 (0.201) | 0.712 (0.111) | 1.599 (0.007) | 1.119 (0.206) |
| | $\gamma$ | 1.771(0.058) | 1.584(0.194) | 1.718(0.172) | 1.683 (0.111) |
| | $p$ | 0.217(0.004) | 0.292(0.011) | 0.231(0.110) | 0.265 (0.001) |

data in Table 2 indicated that the estimates of the parameters of LWP model are close to the true value when we increase the sample size, except for parameters $\alpha$ and $\beta$. The MSE in most cases increase as the censoring percentage increased. However, in most cases, the MSE decreases as the values of $n$ increased. It is also observed in most cases that the MSE is smaller for the $p = 0.2$ with respect to $p = 0.3$. Higher values of MSE occurred when the size of $n$ was smaller than 400.

## 5. Applications

In this section, we compared the LWP distribution with LW and LEP distributions on two data sets to identify an appropriate survival time model. The first data set was related to the time (in days) until the recurrence to the crime of 477 individuals who are in a semi-open regime, in which 60% of the observations are censored. The second data set was extracted from Kalbfleisch and Prentice (2002). The data referred to the survival time (in days) of 195 patients with carcinoma of the oropharynx, in which 30% of the observations are censored. Data was verified before we fit a model to the risk function of the times observed using a graphical method based on the total time test (TTT) that is also known as a TTT plot. This method is useful when there is information about the risk function of the studied variable. According to Aarset (1987), the empirical version of the TTT plot is given by $G(r/n) = [(\sum_{i=1}^{r} Y_{i:n}) - (n-r)Y_{r:n}]/(\sum_{i=1}^{r} Y_{i:n})$, where $r = 1, \ldots, n$ and $Y_{i:n}$ represents the order statistics of the sample. Aarset (1987) showed that the hazard function is constant if the TTT plot is graphically presented as a straight diagonal, the hazard function is increasing (or decreasing) if the TTT plot is concave (or convex). The hazard function is U-shaped if the TTT plot is convex and then concave, if not, the hazard function is unimodal. The TTT plot for two datasets in Figure 2 (a), (c), indicates a unimodal shaped failure rate function. We can therefore try using the LWP distribution for the modeling of data. Table 3 shows the estimates and standard error (in parentheses) from the MLE for the two data sets using the command *optim* of the software R.

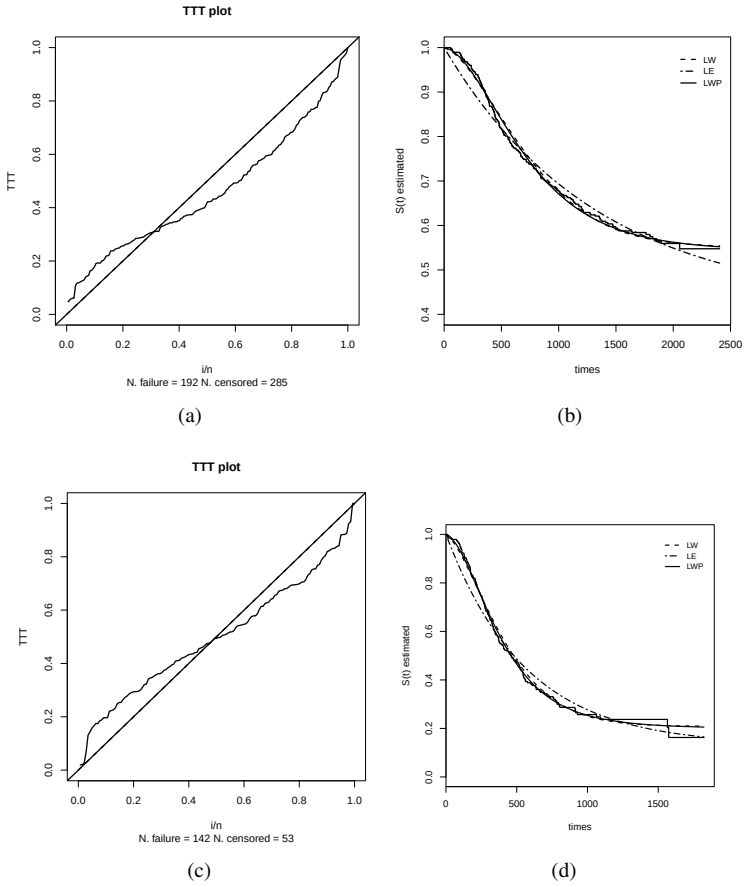


Figure 2: *(a, c) TTT plot; (b, d) Kaplan-Meier curve with estimated survival function via LEP, LW, and LWP distributions, for the recurrence to the crime data (a, b) and for the carcinoma of the oropharynx data (c, d). TTT = total time test; LW = Weibull long-term distribution; LEP = exponential-Poisson long-term distribution; LWP = long-term Weibull-Poisson.*

We also compared LWP distribution with its particular cases in considering AIC and BIC. AIC and BIC values are presented in Table 4 with the log-likelihood $\ell(\cdot)$ of the three models for two datasets. It can be observed that LWP distribution had the lowest AIC values in relation to LEP and LW distributions, indicating that this model is more appropriate to the data. The results are corroborated by the plots in Figure 2 (b), (d), which show the fitted survival functions for all fitted distributions superimposed to the empirical survival function.

A likelihood ratio test was used to select models that best fit the data because LWP distribution is reduced in the LEP and LW distributions. The test procedure is shown in Section 3.2. The hypothesis for LEP and LWP distribution were: $H_0 : \gamma = 1$, i.e., the LEP model is adequate versus $H_1 : \gamma \neq 1$, i.e., the LWP model is adequate. Data that refers to the recurrence of the crime dedicated that the test statistic was 37.362 and the critical value given by the $95^{th}$ percentile of distribution $\chi^2$ with one degree of freedom, is $\chi^2_1 = 3.84$ with the $p$-value ($< 0.001$) was less than a significance level of the 5%, which leads us to reject the null hypothesis. Data that refers to carcinoma of the oropharynx

Table 3: MLEs (and corresponding standard errors in parentheses) for the parameters of the fitted distributions for two datasets

| Datasets | Distributions | $\hat{\beta}$ | $\hat{\gamma}$ | $\hat{\alpha}$ | $\hat{p}$ |
|---|---|---|---|---|---|
| Recurrence to the crime | LEP | 0.007e−01 | - | 0.001 | 0.420 |
| | | (0.001e−01) | - | (0.028) | (0.062) |
| | LW | 0.001 | 1.563 | - | 0.551 |
| | | (7.589e−04) | (0.104) | - | (0.026) |
| | LWP | 0.007e-01 | 1.735 | 2.228 | 0.544 |
| | | (0.002e−01) | (0.125) | (1.459) | (0.028) |
| Carcinoma of the oropharynx | LEP | 0.001 | - | 0.007e−01 | 0.131 |
| | | (0.001e−01) | - | (0.021) | (0.040) |
| | LW | 0.002 | 1.453 | - | 0.208 |
| | | (0.001e−01) | (0.103) | - | (0.034) |
| | LWP | 0.001 | 1.656 | 3.458 | 0.201 |
| | | (0.003e−01) | (0.132) | (2.013) | (0.036) |

MLE = maximum likelihood estimate; LEP = exponential-Poisson long-term distribution; LW = Weibull long-term distribution; LWP = long-term Weibull-Poisson.

Table 4: The log-likelihood $\ell(\cdot)$, AIC, and BIC of the fitted distributions LEP, LW, and LWP for two datasets

| Distributions | Recurrence to the crime | | | Carcinoma of the oropharynx | | |
|---|---|---|---|---|---|---|
| | $\ell(\cdot)$ | AIC | BIC | $\ell(\cdot)$ | AIC | BIC |
| LEP | −1727.523 | 3461.036 | 3473.539 | −1082.608 | 2171.216 | 2181.035 |
| LW | −1712.804 | 3431.608 | 3444.111 | −1082.608 | 2151.398 | 2161.217 |
| LWP | −1710.762 | 3429.524 | 3446.194 | −1082.608 | 2150.508 | 2162.232 |

AIC = Akaike's information criterion; BIC = Bayesian information criterion; LEP = exponential-Poisson long-term distribution; LW = Weibull long-term distribution; LWP = long-term Weibull-Poisson.

indicated a test statistic of 24.076, with a $p$-value ($< 0.001$) less than the significance level of 5%, which leads us to reject the null hypothesis. In relation the LW and LWP distributions, hypotheses were: $H_0 : \alpha \to 0$, i.e., the LW model is adequate versus $H_1 : \alpha > 0$, i.e., the LWP model is adequate. Data that refers to the recurrence to the crime indicated that the test statistic was 4.084 and larger than $1/2 + 1/2\, P(\chi^2_1 \leq c) = 2.705$, at a significance level of the 5%, which leads us to reject the null hypothesis. Data that refers to carcinoma of the oropharynx indicated that the test statistic was 4.258 and large than 2.705, at a significance level of the 5%, which leads us to reject the null hypothesis. Both cases indicated evidence in favor of the LWP model for the two datasets.

## 6. Concluding remarks

This paper proposed a LWP model. This model is an extension of the WP distribution proposed by Bereta *et al.* (2011) on a long term mixture structure. That is, we proposed a new model inside of an competing risk scenario with long-term survival. This distribution was proposed as an alternative to model this type of data and presented as particular cases of the new LEP and the LW. Properties of the proposed model, including its probability density, survival, and hazard functions, were discussed. Its order statistics were also provided. We used a maximum likelihood method for the parameter estimation. This article also compared the performance of the proposed model based on MSE, AIC, BIC, and the test power through a simulation study. The simulation study indicated that more favorable results were obtained with the proportions of AIC, BIC, test power, and MSE for sample sizes larger than 100. The practical importance of the LWP distribution was demonstrated in two dataset where LWP distribution provided a competitive fitting in comparison to some usual long-term distributions such as the LE and LW. This model therefore is expected to be used in other datasets.

## References


Aarset MV (1987). How to identify a bathtub hazard rate, *IEEE Transactions on Reliability*, **36**, 106–108.

Adamidis K and Loukas S (1998). A lifetime distribution with decreasing failure rate, *Statistics and Probability Letters*, **39**, 35–42.

Bereta EM, Louzanda F, and Franco MA (2011). The Poisson-Weibull distribution, *Advances and Applications in Statistics*, **22**, 107–118.

Berkson J and Gage RP (1952). Survival curve for cancer patients following treatment, *Journal of the American Statistical Association*, **47**, 501–515.

Boag JW (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy, *Journal of the Royal Statistical Society Series B (Methodological)*, **11**, 15–53.

Cancho VG, Louzada-Neto F, and Barriga GDC (2011). The Poisson-exponential lifetime distribution, *Computational Statistics & Data Analysis*, **55**, 677–686.

Chahkandi M and Ganjali M (2009). On some lifetime distributions with decreasing failure rate, *Computational Statistics & Data Analysis*, **53**, 4433–4440.

Chen HM and Ibrahim JG (2001). Maximum likelihood methods for cure rate models with missing covariates, *Biometrics*, **57**, 43–52.

Cooner F, Banerjee S, and McBean AM (2006). Modelling geographically referenced survival data with a cure fraction, *Statistical Methods in Medical Research*, **15**, 307–324.

Crowder MJ (2001). *Classical Competing Risks*, Chapman & Hall, New York.

Kalbfleisch JD and Prentice RL (2002). *The Statistical Analysis of Failure Time Data* (2nd ed), John Wiley & Sons, Hoboken.

Khan MS and King R (2016). New generalized inverse Weibull distribution for lifetime modeling, *Communications for Statistical Applications and Methods*, **23**, 147–161.

Kuş C (2007). A new lifetime distribution, *Computational Statistics & Data Analysis*, **51**, 4497–4509.

Louzada F, Roman M, and Cancho VG (2011) The complementary exponential geometric distribution: model, properties, and a comparison with its counterpart, *Computational Statistics & Data Analysis*, **55**, 2516–2524.

Louzada F, Yamachi CY, Marchi VA, and Franco MA (2014). The long-term exponentiated complementary exponential geometric distribution under a latent complementary causes framework, *TEMA (São Carlos)*, **15**, 19–35.

Louzada-Neto F (1999). Polyhazard models for lifetime data, *Biometrics*, **55**, 1281-1285.

Lu K and Tsiatis AA (2005). Comparison between two partial likelihood approaches for the competing risks model with missing cause of failure, *Lifetime Data Analysis*, **11**, 29–40.

Maller RA and Zhou S (1995). Testing for the presence of immune or cured individuals in censored survival data, *Biometrics*, **51**, 1197–1205.

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

Mazucheli J, Louzada F, and Achcar JA (2012). The polysurvival model with long-term survivors, *Brazilian Journal of Probability and Statistics*, **26**, 313–324.

Ortega EM, Cordeiro GM, Hashimoto EM, and Suzuki AK (2017). Regression models generated by gamma random variables with long-term survivors, *Communications for Statistical Applications and Methods*, **24**, 43–65.

Perdoná GC and Louzada-Neto F (2011). A general hazard model for lifetime data in the presence of cure rate, *Journal of Applied Statistics*, **38**, 1395–1405.

Perperoglou A, Keramopoullos A, and van Houwelingen HC (2007). Approaches in modelling long-term survival: an application to breast cancer, *Statistics in Medicine*, **26**, 2666–2685.

Pintilie M (2006). *Competing Risks: A Practical Perspective*, John Wiley & Sons, Southern Gate, Chichester.

Pons O and Lemdani M (2003). Estimation and test in long-term survival mixture models, *Computational Statistics & Data Analysis*, **41**, 465–479.

Tahmasbi R and Rezaei S (2008). A two-parameter lifetime distribution with decreasing failure rate, *Computational Statistics & Data Analysis*, **52**, 3889–3901.

Yakovlev AY, Tsodikov AD, and Asselain B (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*, World Scientific Pub Co Inc., Hackensack.

*Received October 18, 2016; Revised March 18, 2017; Accepted June 29, 2017*