

## Classical and Bayesian inference approaches for the exponentiated discrete Weibull model with censored data and a cure fraction

Bruno Caparroz Lopes de Freitas<sup>1</sup>, Marcos Vinicius de Oliveira Peres<sup>2</sup>,  
Jorge Alberto Achcar<sup>3</sup>, Edson Zangiacomi Martinez<sup>4\*</sup>



\*Corresponding author

1. State University of Maringá, Master Program in Biostatistics, Maringá, Brazil, brunoclf19@gmail.com
2. Ribeirão Preto Medical School, University of São Paulo (USP), Ribeirão Preto, Brazil, mvperes1991@usp.br
3. Ribeirão Preto Medical School, University of São Paulo (USP), Ribeirão Preto, Brazil, achcar@fmrp.usp.br
4. Ribeirão Preto Medical School, University of São Paulo (USP), Ribeirão Preto, Brazil, edson@fmrp.usp.br

### Abstract

In this paper, we introduce maximum likelihood and Bayesian parameter estimation for the exponentiated discrete Weibull (EDW) distribution in the presence of randomly right-censored data. We also consider the inclusion of a cure fraction in the model. The performance of the maximum likelihood estimation approach is assessed by conducting an extensive simulation study with different sample sizes and different values for the parameters of the EDW distribution. The usefulness of the proposed model is illustrated with two examples considering real data sets.

**Key Words:** Survival analysis; Maximum likelihood estimation; Cure fraction; Bayesian inference; Discrete distributions; Censored data.

**Mathematical Subject Classification:** 62N01, 62N02, 62F15.

### 1. Background

Let  $T$  be a random variable denoting a survival time, and let  $t$  be an observation of  $T$ . The exponentiated discrete Weibull (EDW) distribution, introduced by Nekoukhou and Bidram(2015), has probability accumulated distribution function given by

$$F(t) = P(T \leq t) = \left[1 - \theta^{(t+1)^\alpha}\right]^\beta, \quad t \in \mathbb{N}^0,$$

where  $\alpha > 0$ ,  $\beta > 0$ ,  $0 < \theta < 1$ , and  $\mathbb{N}^0 = \{0, 1, 2, \dots\}$ . The correspondent probability mass function (pmf) is given by

$$f(t) = P(T = t) = \left[1 - \theta^{(t+1)^\alpha}\right]^\beta - \left[1 - \theta^{t^\alpha}\right]^\beta, \quad t \in \mathbb{N}^0, \quad (1)$$

and the corresponding hazard function is

$$h(t) = P(T = t | T \geq t) = \frac{P(T = t)}{P(T \geq t)} = \frac{f(t)}{S(t-1)} = \frac{\left[1 - \theta^{(t+1)^\alpha}\right]^\beta - \left[1 - \theta^{t^\alpha}\right]^\beta}{1 - \left[1 - \theta^{t^\alpha}\right]^\beta}, \quad t \in \mathbb{N}^0, \quad (2)$$

were  $S(t) = 1 - F(t) = P(T > t)$  is the survival function, defined in the context of survival analysis as the probability that an individual will be surviving at least until time  $t$ . The EDW distribution was studied by authors as Cardial et al.(2020), and a bivariate version was proposed by El-Morshedy et al.(2020). Some existing discrete distributions can be obtained as special cases of the EDW distribution (El-Morshedy et al.,2020), as listed below:

- (a) **Discrete Weibull (DW) distribution:** when  $\beta = 1$ , the expression (1) is the *pmf* of the two-parameter discrete Weibull distribution introduced by Nakagawa and Osaki(1975). The DW distribution has a decreasing failure rate when  $0 < \alpha \leq 1$  and an increasing failure rate when  $\alpha > 1$ . When  $\alpha = 1$ , the DW distribution becomes the discrete exponential (DE) distribution with a constant failure rate (Johnson and Kotz,1982).
- (b) **Discrete Rayleigh (DR) distribution:** when  $\beta = 1$  and  $\alpha = 2$ , the expression (1) is reduced to the *pmf* of a discrete Rayleigh distribution with an increasing failure rate (Roy,2004).
- (c) **Discrete generalized exponential (DGE) distribution:** this distribution is a special case of the EDW distribution with  $\alpha = 1$  (Nekoukhou et al.,2013).
- (d) **Discrete generalized Rayleigh (DGR) distribution:** when  $\alpha = 2$ , the expression (1) is the *pmf* of the two-parameter discrete generalized Rayleigh distribution introduced by Alamatsaz et al.(2016).

Let us consider a reparameterization of the EDW distribution such that  $\theta = e^{-\gamma}$ , where  $\gamma > 0$ , which simplifies obtaining estimators of the distribution parameters. Hereafter, we write EDW( $\alpha, \beta, \gamma$ ) to denote an EDW distribution with *pmf* given in (1), with  $\theta = e^{-\gamma}$ . For different values for the parameters ( $\alpha, \beta, \gamma$ ). Figure 1 presents graphs of the *pmf*, survival function, and hazard function of the EDW distribution considering different parameter values. The hazard function  $h(t)$  of the EDW distribution presents great flexibility of fit as increasing or decreasing shapes.

The main goal of this paper is to introduce the EDW distribution to model discrete lifetime data in the presence of right-censored time-to-event data. We also consider the inclusion of a cure fraction in the model. The article is organized as follows. Section 2 discusses the maximum likelihood and Bayesian estimates of the EDW distribution in the presence of censored data. Section 2 also presents the steps to obtain simulated samples from an EDW distribution. Section 3 considers a simulation study assuming the proposed model and compares the results from the frequentist and Bayesian approaches. In Section 3, we also conduct a simulation study to assess the performance of the maximum likelihood estimators in the presence of cure fraction. The usefulness of the proposed model is also shown considering two real data set applications. A brief revision of the major results and some concluding remarks are presented in Section 4.

## 2. Methods

### 2.1. Maximum likelihood estimation for complete data

Let  $T_1, \dots, T_n$  be a random sample of failure times from an EDW survival distribution with *pmf* given by equation (1). Considering  $\theta = e^{-\gamma}$ , we have

$$F(t) = \left[1 - e^{-\gamma(t_i+1)^\alpha}\right]^\beta, \quad t \in \mathbb{N}^0,$$

and the correspondent *pmf* is given by

$$f(t) = \left[1 - e^{-\gamma(t_i+1)^\alpha}\right]^\beta - \left(1 - e^{-\gamma t_i^\alpha}\right)^\beta.$$

Assuming the EDW model and a random sample of size  $n$ , the likelihood function for the parameters  $\alpha, \beta$  and  $\gamma$  is given by

$$L(\alpha, \beta, \gamma | \mathbf{t}) = \prod_{i=1}^n \left\{ \left[1 - e^{-\gamma(t_i+1)^\alpha}\right]^\beta - \left(1 - e^{-\gamma t_i^\alpha}\right)^\beta \right\},$$

where  $\mathbf{t} = (t_1, t_2, \dots, t_n)$ , and the corresponding log-likelihood function is given by

$$\ell(\alpha, \beta, \gamma | \mathbf{t}) = \sum_{i=1}^n \log \left\{ \left[1 - e^{-\gamma(t_i+1)^\alpha}\right]^\beta - \left(1 - e^{-\gamma t_i^\alpha}\right)^\beta \right\}.$$

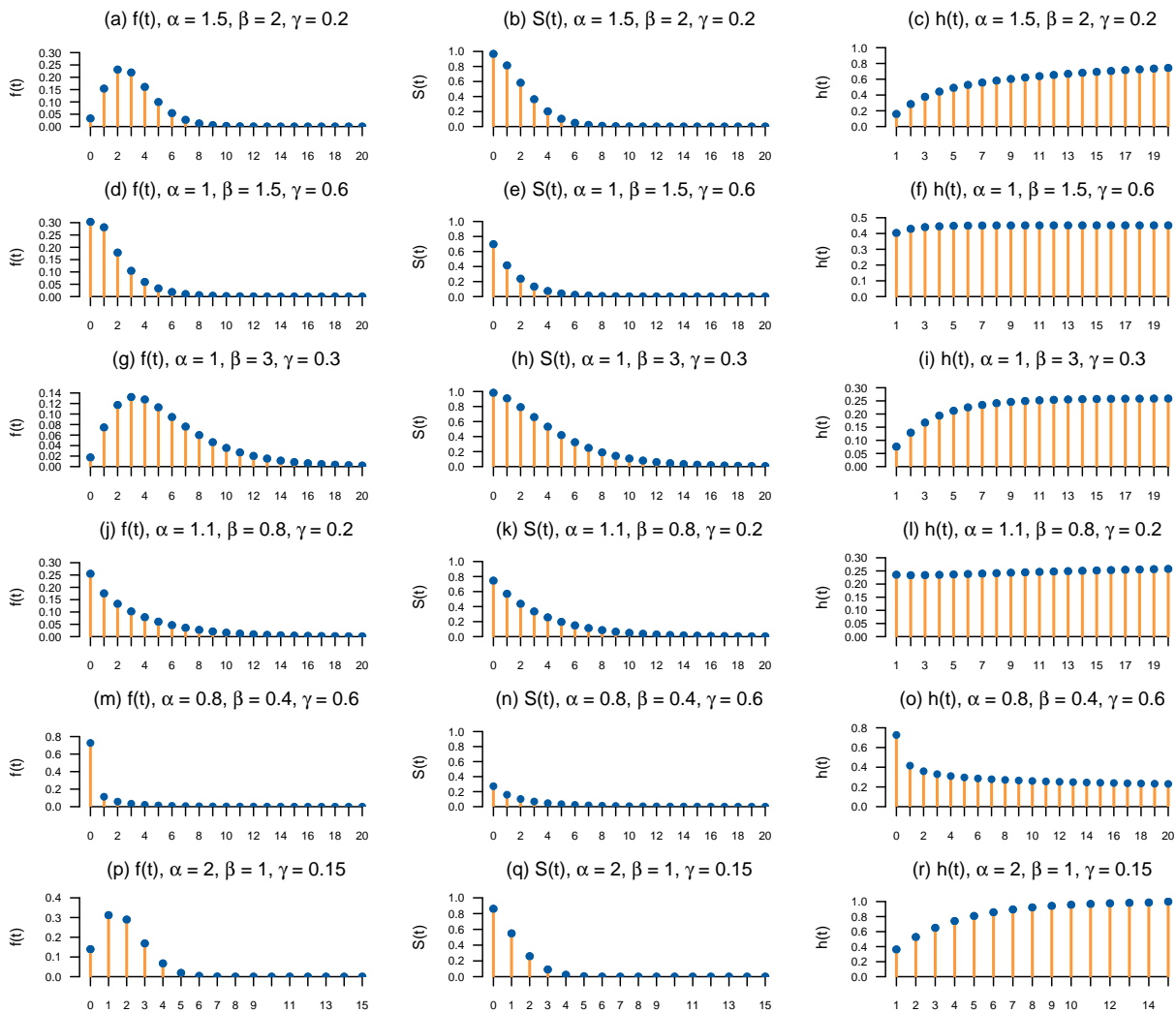


Figure 1: The pmf, survival function and hazard function of EDW( $\alpha, \beta, \gamma$ ) for different values of  $\alpha, \beta$  and  $\gamma$ .

By deriving the log-likelihood function with respect to  $\alpha, \beta$  and  $\gamma$ , we have the following equations:

$$\frac{\partial \ell}{\partial \alpha} = \beta \gamma \sum_{i=1}^n \frac{u(\alpha, \beta, \gamma, t_i + 1) \log(t_i + 1) - u(\alpha, \beta, \gamma, t_i) \log t_i}{(1 - e^{-\gamma(t_i+1)^\alpha})^\beta - (1 - e^{-\gamma t_i^\alpha})^\beta},$$

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^n \frac{(1 - e^{-\gamma(t_i+1)^\alpha})^\beta \log(1 - e^{-\gamma(t_i+1)^\alpha}) - (1 - e^{-\gamma t_i^\alpha})^\beta \log(1 - e^{-\gamma t_i^\alpha})}{(1 - e^{-\gamma(t_i+1)^\alpha})^\beta - (1 - e^{-\gamma t_i^\alpha})^\beta}$$

and

$$\frac{\partial \ell}{\partial \gamma} = \beta \sum_{i=1}^n \frac{u(\alpha, \beta, \gamma, t_i + 1) - u(\alpha, \beta, \gamma, t_i)}{(1 - e^{-\gamma(t_i+1)^\alpha})^\beta - (1 - e^{-\gamma t_i^\alpha})^\beta},$$

where

$$u(\alpha, \beta, \gamma, t_i) = t_i^\alpha e^{-\gamma t_i^\alpha} (1 - e^{-\gamma t_i^\alpha})^{\beta-1}. \quad (3)$$

The maximum likelihood (ML) estimators can be obtained numerically solving the equations  $\frac{\partial \ell}{\partial \alpha} = 0, \frac{\partial \ell}{\partial \beta} = 0$  and  $\frac{\partial \ell}{\partial \gamma} = 0$  (Millar, 2011). Let the ML estimates be  $(\hat{\alpha}_{ML}, \hat{\beta}_{ML}, \hat{\gamma}_{ML})^T$ , which can be obtained by a numerical method such as the Newton-Raphson procedure and its variants (Broyden, 1970; Fletcher, 1970; Nelder and Mead, 1965). In

this article, we use the maxLik package in the statistical software R, which numerically maximises the likelihood function (Henningsen and Toomet, 2011). Inferences of interest as hypothesis tests or confidence intervals (CI) for the parameters of the EDW distribution are usually obtained from the asymptotical normality of the ML estimators. The asymptotic multivariate normal distribution of the ML estimators  $\hat{\alpha}_{ML}$ ,  $\hat{\beta}_{ML}$  and  $\hat{\gamma}_{ML}$  is denoted by

$$\begin{bmatrix} \hat{\alpha}_{ML} \\ \hat{\beta}_{ML} \\ \hat{\gamma}_{ML} \end{bmatrix} \sim N \left( \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}, \begin{bmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{bmatrix} \right),$$

where  $V_{11}$ ,  $V_{22}$  and  $V_{33}$  are the variances of  $\hat{\alpha}_{ML}$ ,  $\hat{\beta}_{ML}$  and  $\hat{\gamma}_{ML}$ , respectively,  $V_{12} = V_{21}$  is the covariance between  $\hat{\alpha}_{ML}$  and  $\hat{\beta}_{ML}$ ,  $V_{13} = V_{31}$  is the covariance between  $\hat{\alpha}_{ML}$  and  $\hat{\gamma}_{ML}$ , and  $V_{32} = V_{23}$  is the covariance between  $\hat{\beta}_{ML}$  and  $\hat{\gamma}_{ML}$ . Approximate  $100(1 - \nu)\%$  Wald-type CIs for  $\alpha$ ,  $\beta$  and  $\gamma$  are, respectively, given by

$$\hat{\alpha}_{ML} \mp z_{\nu/2} \sqrt{\hat{V}_{11}}, \quad \hat{\beta}_{ML} \mp z_{\nu/2} \sqrt{\hat{V}_{22}} \quad \text{and} \quad \hat{\gamma}_{ML} \mp z_{\nu/2} \sqrt{\hat{V}_{33}}, \quad (4)$$

where  $z_{\nu/2}$  denotes the upper  $\nu$ -th percentile of the standard normal distribution. The asymptotic variances of the ML estimators are given by the elements of the inverse of the Fisher's information matrix. The expected information matrix is given by

$$I(\alpha, \beta, \gamma) = \begin{bmatrix} -E \left( \frac{\partial^2 \ell}{\partial \alpha^2} \right) & -E \left( \frac{\partial^2 \ell}{\partial \alpha \partial \beta} \right) & -E \left( \frac{\partial^2 \ell}{\partial \alpha \partial \gamma} \right) \\ -E \left( \frac{\partial^2 \ell}{\partial \beta \partial \alpha} \right) & -E \left( \frac{\partial^2 \ell}{\partial \beta^2} \right) & -E \left( \frac{\partial^2 \ell}{\partial \beta \partial \gamma} \right) \\ -E \left( \frac{\partial^2 \ell}{\partial \gamma \partial \alpha} \right) & -E \left( \frac{\partial^2 \ell}{\partial \gamma \partial \beta} \right) & -E \left( \frac{\partial^2 \ell}{\partial \gamma^2} \right) \end{bmatrix},$$

whose elements are the expected values of the second partial derivatives of the log-likelihood function with respect to the parameters. Since the above confidence intervals given by (4) may lead to negative lower bounds sometimes, the logarithmic transformation and the delta method can be used to approximate the distribution of  $\log \hat{\alpha}_{ML}$ ,  $\log \hat{\beta}_{ML}$  and  $\log \hat{\gamma}_{ML}$ . That is, we have that for a given estimator  $\hat{\lambda}_{ML}$  and a correspondent  $\lambda$  parameter,

$$\sqrt{n} \left( \log(\hat{\lambda}_{ML}) - \log(\lambda) \right) \xrightarrow{D} N \left( 0, \frac{n \text{Var}(\hat{\lambda}_{ML})}{\lambda^2} \right),$$

where  $\hat{\lambda}_{ML} \in (\hat{\alpha}_{ML}, \hat{\beta}_{ML}, \hat{\gamma}_{ML})$ ,  $\lambda \in (\alpha, \beta, \gamma)$  and  $\xrightarrow{D}$  denotes convergence in distribution. Asymptotic  $100(1 - \nu)\%$  confidence intervals for  $\log(\alpha)$ ,  $\log(\beta)$  and  $\log(\gamma)$  can then be obtaining by

$$\log(\hat{\lambda}_{ML}) \mp z_{\nu/2} \frac{\sqrt{\text{Var}(\hat{\lambda}_{ML})}}{\hat{\lambda}_{ML}} \equiv (L_{\lambda}, U_{\lambda}),$$

and using the inverse logarithmic transformation, the approximate confidence intervals for  $\alpha$ ,  $\beta$  and  $\gamma$  are respectively obtained as  $(e^{L_{\alpha}}, e^{U_{\alpha}})$ ,  $(e^{L_{\beta}}, e^{U_{\beta}})$  and  $(e^{L_{\gamma}}, e^{U_{\gamma}})$ .

In this article, the Akaike and Bayesian information criteria (AIC and BIC) are used to compare the models based on the EDW distribution and its special cases described in the previous section. For a given model fit to the data, the AIC value is given by  $AIC = -2 \log(L) + 2k$ , where  $L$  is the log-likelihood of the model and  $k$  is the number of model parameters (Akaike, 1974). The BIC value is given by  $BIC = -2 \log(L) + k \log(n)$ , where  $n$  is the sample size (Schwarz, 1978). When the sample size is small, many authors suggested the use of the corrected AIC (AICC) as an alternative to AIC. The AICC value is given by  $AICC = AIC + 2k(k+1)/(n-k-1)$  (Hurvich and Tsai, 1989). The AIC, BIC, and AICC values directly compare models, with lower values indicating a better model.

## 2.2. Maximum-likelihood estimation including censored data

Considering the presence of right censored observations, the contribution of the  $i$ th individual for the likelihood function based on a random sample  $(t_i, d_i)$  of size  $n$  is given by

$$L_i = [f(t_i)]^{d_i} [S(t_i)]^{1-d_i},$$

where  $d_i$  is a censoring indicator variable, that is,  $d_i = 1$  for an observed lifetime and  $d_i = 0$  for a censored lifetime ( $i = 1, \dots, n$ ). Assuming the EDW model, the likelihood function for  $\alpha$ ,  $\beta$  and  $\gamma$  is given by

$$L(\alpha, \beta, \gamma | \mathbf{t}, \mathbf{d}) = \prod_{i=1}^n \left\{ \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta - \left( 1 - e^{-\gamma t_i^\alpha} \right)^\beta \right\}^{d_i} \left\{ 1 - \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta \right\}^{1-d_i},$$

where  $\mathbf{d} = (d_1, d_2, \dots, d_n)$ , and the corresponding log-likelihood function is

$$\begin{aligned} \ell(\alpha, \beta, \gamma | \mathbf{t}, \mathbf{d}) &= \sum_{i=1}^n d_i \log \left\{ \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta - \left( 1 - e^{-\gamma t_i^\alpha} \right)^\beta \right\} \\ &\quad + \sum_{i=1}^n (1 - d_i) \log \left\{ 1 - \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta \right\}. \end{aligned}$$

By deriving the log-likelihood function with respect to  $\alpha$ ,  $\beta$  and  $\gamma$ , we have the following expressions:

$$\begin{aligned} \frac{\partial \ell}{\partial \alpha} &= \beta \gamma \sum_{i=1}^n d_i \frac{u(\alpha, \beta, \gamma, t_i + 1) \log(t_i + 1) - u(\alpha, \beta, \gamma, t_i) \log t_i}{\left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta - \left( 1 - e^{-\gamma t_i^\alpha} \right)^\beta} \\ &\quad - \beta \gamma \sum_{i=1}^n \frac{(1 - d_i) \log(t_i + 1) u(\alpha, \beta, \gamma, t_i + 1)}{1 - \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta}, \end{aligned}$$

$$\begin{aligned} \frac{\partial \ell}{\partial \beta} &= \sum_{i=1}^n d_i \frac{\left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta \log \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right] - \left( 1 - e^{-\gamma t_i^\alpha} \right)^\beta \log \left( 1 - e^{-\gamma t_i^\alpha} \right)}{\left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta - \left( 1 - e^{-\gamma t_i^\alpha} \right)^\beta} \\ &\quad - \sum_{i=1}^n \frac{(1 - d_i) \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta \log \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]}{1 - \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta} \end{aligned}$$

and

$$\frac{\partial \ell}{\partial \gamma} = \beta \sum_{i=1}^n d_i \frac{u(\alpha, \beta, \gamma, t_i + 1) - u(\alpha, \beta, \gamma, t_i)}{\left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta - \left( 1 - e^{-\gamma t_i^\alpha} \right)^\beta} - \beta \sum_{i=1}^n \frac{(1 - d_i) u(\alpha, \beta, \gamma, t_i + 1)}{1 - \left[ 1 - e^{-\gamma(t_i+1)^\alpha} \right]^\beta},$$

where  $u(\alpha, \beta, \gamma, t_i)$  is given by (3). The ML estimates of the three parameters are again found equating these expressions to zero and solving them simultaneously. Due to the complexity of these expressions, it is appropriate to use the optimization algorithms such as a Newton-Raphson algorithm for maximizing the likelihood function numerically.

### 2.3. A cure fraction model based on the EDW distribution

In several medical studies related to survival data analysis, it is observed a fraction of individuals who appear not to be subject to the occurrence of the event of interest. This can occur, for example, in clinical trials where some patients who respond to treatment may be free of signs of the disease for an extended period and are possibly cured (Morbiducci et al., 2003). This situation is usually observed when the graph of the usual non-parametric Kaplan-Meier estimator of the survival function for the lifetimes in the presence of censored data shows a long and stable plateau with heavy censoring at the extreme right of the plot, suggesting the presence of immune individuals. However, this plot is not sufficient to state that a model with a cure fraction is suitable for a given data set, and experts' arguments are also important. For example, when analysing data from a clinical trial, it is essential to use medical statements to justify whether a cure of the disease under study is possible or whether immunity to the event of interest can occur. Some specific models commonly used to handle data with such characteristics are the mixture and non-mixture models described by Achcar et al. (2012) and Martinez et al. (2013) and the models based on defective distributions (Rocha et al., 2017; Scudilio et al., 2019; Hamdeni and Gasmi, 2020).

The survival function of the standard mixture model (Farewell, 1982) is given by

$$S(t) = \eta + (1 - \eta)S_0(t),$$

where  $\eta \in (0, 1)$  is the cure fraction or the proportion of immunes, and  $S_0(t)$  is a baseline survival function for the susceptible individuals. We can note that if  $S_0(t)$  is a function that converges to zero as  $t$  tends to infinity, then  $S(t)$  converges to  $\eta$  as  $t$  tends to infinity. Considering a random sample  $(t_i, d_i)$  of size  $n$ ,  $i = 1, \dots, n$ , the contribution of the  $i$ th subject for the likelihood function is given by

$$L_i = [f(t_i)]^{d_i} [S(t_i)]^{1-d_i} = [(1-\eta)f_0(t)]^{d_i} [\eta + (1-\eta)S_0(t)]^{1-d_i},$$

where  $f_0(t)$  is the baseline probability density function for the susceptible individuals and  $d_i$  is a censoring indicator variable. Assuming the mixture model based on the EDW distribution, the log-likelihood function for  $\alpha, \beta, \gamma$  and  $\eta$  is given by

$$\begin{aligned} L(\alpha, \beta, \gamma, \eta | \mathbf{t}, \mathbf{d}) &= \prod_{i=1}^n \left\{ (1-\eta) \left[ \left( 1 - e^{-\gamma(t_i+1)^\alpha} \right)^\beta - \left( 1 - e^{-\gamma t_i^\alpha} \right)^\beta \right] \right\}^{d_i} \\ &\quad \times \left\{ \eta + (1-\eta) \left[ 1 - \left( 1 - e^{-\gamma(t_i+1)^\alpha} \right)^\beta \right] \right\}^{1-d_i}, \end{aligned}$$

and the correspondent log-likelihood function for  $\alpha, \beta, \gamma$ , and  $\eta$  is given by

$$\begin{aligned} \ell(\alpha, \beta, \gamma, \eta | \mathbf{t}, \mathbf{d}) &= \sum_{i=1}^n d_i \log [(1-\eta)f_0(t_i)] + \sum_{i=1}^n (1-d_i) \log [\eta + (1-\eta)S_0(t_i)] \\ &= \sum_{i=1}^n d_i \log(1-\eta) + \sum_{i=1}^n d_i \log \left[ \left( 1 - e^{-\gamma(t_i+1)^\alpha} \right)^\beta - \left( 1 - e^{-\gamma t_i^\alpha} \right)^\beta \right] \\ &\quad + \sum_{i=1}^n (1-d_i) \log \left\{ \eta + (1-\eta) \left[ 1 - \left( 1 - e^{-\gamma(t_i+1)^\alpha} \right)^\beta \right] \right\}. \end{aligned}$$

The ML estimators are obtained by solving the corresponding equations after differentiating the log-likelihood function  $\ell(\alpha, \beta, \gamma, \eta | \mathbf{t}, \mathbf{d})$  with respect to  $\alpha, \beta, \gamma$  and  $\eta$  and setting the resultant derivative to be equal to zero.

## 2.4. Bayesian analysis

Under a Bayesian approach, a prior distribution must be assigned to each parameter of the EDW model. We thus assume gamma prior distributions for the parameters  $\alpha, \beta$  and  $\gamma$  considering that these parameters are real and positive numbers. Thus, we assume gamma prior distributions  $\alpha \sim \text{Gamma}(u_\alpha, v_\alpha)$ ,  $\beta \sim \text{Gamma}(u_\beta, v_\beta)$  and  $\gamma \sim \text{Gamma}(u_\gamma, v_\gamma)$ , where  $u_\alpha, v_\alpha, u_\beta, v_\beta, u_\gamma$  and  $v_\gamma$  are known hyperparameters, and  $\text{Gamma}(u, v)$  denotes a gamma distribution with mean  $u/v$  and variance  $u/v^2$ . When considering data with a cure fraction, we assume a beta prior distribution for the proportion  $\eta$  of the long-term survivors denoted by  $\eta \sim \text{Beta}(a, b)$  where  $a$  and  $b$  are known hyperparameters. In all cases, we assume prior independence among the parameters included in the model. The joint posterior distribution for the model's parameters is thus obtained via Bayes theorem by combining the joint prior distribution with the correspondent likelihood function (Carlin and Louis, 2001). To get information on the joint posterior distributions for the parameters of the model, we used Markov chain Monte Carlo (MCMC) methods as the Gibbs and Metropolis-Hastings algorithms (Gelman et al., 2014). The simulation algorithm generated 1,005,000 samples of the joint posterior distribution of interest with a burn-in phase of 5,000 simulated samples to eliminate the effect of the initial values in the iterative procedure and considering a thinning interval of size 100 to have approximately independent samples. The Bayes Monte Carlo estimates of the parameters based on the simulated 10,000 Gibbs samples were obtained as the sample means (use of a squared error loss function) of the Gibbs samples drawn from the joint posterior distribution, and the uncertainty in the estimates was indicated by the 95% highest probability density (HPD) intervals (Chen and Shao, 1999).

The Geweke convergence diagnostic method was employed to assess convergence of the chains (Geweke, 1992; Cowles and Carlin, 1996; El Adlouni et al., 2006). This method takes two non-overlapping parts of a simulated Markov chain and compares the means of both parts. The MCMCpack package of the R software (Martin and Quinn, 2006) considers by default the first 10% and the last 50%. Supposing that  $\phi$  is a parameter of interest and  $\phi^{(1)}, \dots, \phi^{(B)}$  are the related

simulated values, we calculate a  $Z$ -score given by

$$Z = \frac{\bar{\phi}^B - \bar{\phi}^A}{\hat{\sigma}_{\bar{\phi}^B - \bar{\phi}^A}},$$

where  $\bar{\phi}^A$  and  $\bar{\phi}^B$  are respectively the means of the initial and the final parts, and  $\hat{\sigma}_{\bar{\phi}^B - \bar{\phi}^A}$  is an estimated standard deviation of the difference  $\bar{\phi}^B - \bar{\phi}^A$ . Considering that  $Z$  asymptotically follows a standardized normal distribution, we obtain convergence for a chain if its correspondent absolute  $Z$  value is less than 1.96. The convergence of the chains was also assessed by visual inspection using traceplots. Plots of the autocorrelation function (ACF) were used to check for the correlation between successive MCMC samples.

In this article, the Bayesian model selection is based on the Logarithm of the Pseudo-Marginal Likelihood (LPML; Ibrahim et al., 2014). The LPML can be obtained through the Conditional Predictive Ordinate (CPO) values (Geisser and Eddy, 1979). An estimate of the LPML can be calculated as

$$\widehat{LPML} = \sum_{i=1}^n \log(\widehat{CPO}_i),$$

where  $\widehat{CPO}_i$  is the Monte Carlo estimate of the CPO for the  $i$ th observation. A larger LPML indicates a better fit of the proposed model to the data.

## 2.5. Simulating samples from a EDW distribution

Nekoukhou and Bidram (2015) showed that if  $X$  is a random variable that follows a continuous exponentiated Weibull distribution with cumulative function given by

$$F(x) = \left[1 - e^{-\gamma x^\alpha}\right]^\beta,$$

where  $x > 0$ ,  $\alpha > 0$ ,  $\beta > 0$  and  $\gamma > 0$ , then  $[X]$  has an  $EDW(\alpha, \beta, \gamma)$  distribution. In this case,  $g(x) = [x]$  denotes the floor function, i.e. the greatest integer less than or equal to  $x$ . Based on this remark and considering the Algorithms 6 and 7 presented in the article from Ramos et al. (2020), we can simulate a sample of size  $n$  from the EDW distribution with right-censored data following the steps:

**Step 1.** Fix values of  $\alpha$ ,  $\beta$  and  $\gamma$ . If a cure fraction is present, also fix a value for  $\eta$ .

**Step 2.** Generate  $n$  random samples from  $u_i \sim Uniform(0, 1)$ .

**Step 3.** Consider  $t'_{0i} = F^{-1}(u_i; \alpha, \beta, \gamma)$ , where

$$F^{-1}(u_i; \alpha, \beta, \gamma) = \left[ -\frac{\log(1 - u_i^{1/\beta})}{\gamma} \right]^{1/\alpha}.$$

**Step 4.** If a cure fraction is not present, then  $t'_i = t'_{0i}$  and skip to Step 7.

**Step 5.** Generate  $n$  random samples from a Bernoulli distribution with parameter  $\eta$ , denoted by  $b_i \sim Bernoulli(\eta)$ .

**Step 6.** If  $b_i = 1$ , then  $t'_i = \infty$ . If  $b_i = 0$ , then  $t'_i = t'_{0i}$ .

**Step 7.** Generate  $n$  random samples from  $c_i \sim Uniform(0, \max(t'_{0i}))$ . This is the distribution that controls the censorship mechanism.

**Step 8.** Pairs of values  $(t_1, d_1), (t_2, d_2), \dots, (t_n, d_n)$  are thus obtained. If  $t'_i \leq c_i$ , then  $t_i = [t'_i]$  and  $d_i = 1, i = 1, \dots, n$ . Otherwise,  $t_i = [c_i]$  and  $d_i = 0, i = 1, \dots, n$ .

### 3. Results

#### 3.1. Applications to simulated data

In order to exemplify the application of the EDW model and to compare the results from the frequentist and Bayesian approaches, we simulate samples of size  $n = 50, 100$ , and  $300$  from the EDW distribution for  $(\alpha, \beta, \gamma)$  equal to (a)  $(1.5, 2, 0.2)$ , (b)  $(1, 1.5, 0.6)$ , (c)  $(1, 3, 0.3)$ , (d)  $(1.1, 0.8, 0.2)$ , (e)  $(0.8, 0.4, 0.6)$ , and (f)  $(2, 1, 0.15)$ , according to the steps described in the subsection 2.5. The seed value of the pseudo random number generator was arbitrarily fixed to a value of 34. For each sample size, we evaluate the ML estimations of the parameters using the maxLik package of the R software (Henningsen and Toomet, 2011) and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimization method (Broyden, 1970; Fletcher, 1970). Confidence intervals were obtained using the delta method and the logarithmic transformation, as described in the previous section. The MCMCpack package (Martin and Quinn, 2006) was used to obtain the Bayesian estimates, the 95% HPD intervals, and the Geweke's  $Z$ -score. Under the Bayesian estimation framework, prior distributions of  $\alpha$ ,  $\beta$  and  $\gamma$  were given by  $\alpha \sim \text{Gamma}(1, 1)$ ,  $\beta \sim \text{Gamma}(1, 1)$  and  $\gamma \sim \text{Gamma}(1, 1)$ . Tables 1 and 2 give all simulation results, where we can note that the estimates of the parameters tend to be closer to the nominal values when  $n$  increases. ML and Bayesian point estimates are relatively close to each other. However, it is observed that under the frequentist framework, we have fairly wide confidence intervals, especially when considering smaller samples. All absolute values of Geweke's  $Z$ -score in Tables 1 and 2 are less than 1.96, suggesting satisfactory convergence of the chains to a stable distribution.

#### 3.2. Applications to simulated data with a cure fraction

In this subsection, a simulation study is presented in order to exemplify the application of the EDW model under right censoring mechanism in presence of a cure fraction, and to compare the obtained results from the frequentist and Bayesian approaches. We simulate samples of size  $n = 50, 100, 300, 1000$  and  $5000$  from the EDW distribution with a cure fraction, according to the steps described in the subsection 2.5. The chosen values to perform this study were  $(\alpha, \beta, \gamma) = (1.1, 0.8, 0.2)$  and the  $\eta$  parameter was set at 0.2 and 0.5. Confidence intervals for  $\alpha$ ,  $\beta$  and  $\gamma$  were obtained using the delta method and the logarithmic transformation, as described in the subsection 2.1. However, confidence intervals for  $\eta$  were obtained using the traditional Wald method, since all their bounds were in the interval  $(0, 1)$ . Under the Bayesian estimation framework, prior distributions of  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\eta$  were assumed as  $\alpha \sim \text{Gamma}(1, 1)$ ,  $\beta \sim \text{Gamma}(1, 1)$ ,  $\gamma \sim \text{Gamma}(1, 1)$  and  $\eta \sim \text{Beta}(0.5, 0.5)$ . Table 3 shows the obtained results, from where it is noted that the ML and Bayesian point estimates tend to be closer as the sample size increases. The frequentist standard errors tend to decrease as the sample sizes increase, as it is expected. However, we have fairly wide confidence intervals when considering smaller samples. The Bayesian 95% HPD intervals are relatively narrow. Geweke's  $Z$ -scores together with visual inspection of the simulated chains suggest that convergence was achieved for each parameter.

#### 3.3. A simulation study

In this subsection, we perform a simulation study to examine the performance of the maximum likelihood estimation method. The simulation study was conducted from random samples of sizes 25, 30, 40, 45, ..., 400 of the EDW distribution for different values of the parameters. All simulation results were based on  $B = 1,000$  replications for the different sample sizes considered for each parameter setting. We computed the coverage probability of the confidence intervals, bias, and MSE of parameter estimates based on these  $B$  values, and the obtained results are presented in Figure 2. The coverage probability is the proportion of the time that the interval contains the nominal value of the respective parameter. In this study, the nominal confidence level is taken as 95%. The bias in the estimation of a parameter  $\lambda$  and the MSE are respectively estimated by

$$\widehat{Bias}(\hat{\lambda}_{ML}) = \frac{1}{B} \sum_{b=1}^B \left( \hat{\lambda}_{ML}^{(b)} - \lambda_N \right) \quad \text{and} \quad \widehat{MSE}(\hat{\lambda}_{ML}) = \frac{1}{B} \sum_{b=1}^B \left( \hat{\lambda}_{ML}^{(b)} - \lambda_N \right)^2,$$

where  $\hat{\lambda}_{ML} \in (\hat{\alpha}_{ML}, \hat{\beta}_{ML}, \hat{\gamma}_{ML})$  is the ML estimate for a given parameter,  $\hat{\lambda}_{ML}^{(b)}$  is the ML estimate obtained for  $\lambda$  considering the  $b$ -th simulated sample,  $\lambda_N$  is the corresponding nominal value for  $\lambda$ ,  $\lambda \in (\alpha, \beta, \gamma)$ , and  $B$  is the number of simulated samples. This simulation study was carried out using the software R.

We can note in Figure 2 that in all simulations, the biases and the MSE reduces as the sample sizes increase for



**Table 1: Maximum likelihood and Bayesian estimates for the parameters of the model based on EDW distribution.**

Nominal values ( $\alpha, \beta, \gamma$ )			Maximum likelihood estimates			Bayesian estimates			
	$n$	Parameter	Estimate	Std. error	95% CI	Mean	Std. dev.	95% HPD interval	Geweke Z-score
(1.5,2,0.2)	50	$\alpha$	1.7202	1.4276	(0.3382, 8.7497)	1.5565	0.4354	(0.7910, 2.3817)	-1.77
		$\beta$	1.7024	2.4696	(0.0991, 29.234)	2.2147	1.0833	(0.5723, 4.3979)	0.70
		$\gamma$	0.1036	0.3211	(0.0002, 45.107)	0.1985	0.1527	(0.0025, 0.5169)	0.82
	100	$\alpha$	1.4928	0.6480	(0.6375, 3.4955)	1.4368	0.3443	(0.8356, 2.1384)	-0.39
		$\beta$	1.7545	1.3825	(0.3745, 8.2198)	2.1134	0.9515	(0.7139, 4.0161)	-0.34
		$\gamma$	0.1675	0.2406	(0.0100, 2.7989)	0.2384	0.1578	(0.0063, 0.5397)	-0.19
	300	$\alpha$	1.3502	0.3733	(0.7853, 2.3214)	1.3597	0.2672	(0.8798, 1.8795)	-0.99
		$\beta$	2.0280	1.0535	(0.7326, 5.6139)	2.2047	0.8226	(0.8973, 3.8911)	0.28
		$\gamma$	0.2120	0.1825	(0.0391, 1.1463)	0.2441	0.1342	(0.0146, 0.5040)	0.38
(1,1.5,0.6)	50	$\alpha$	0.9873	0.8699	(0.1755, 5.5521)	1.0892	0.3354	(0.5395, 1.7679)	0.01
		$\beta$	1.5273	2.6830	(0.0488, 47.775)	1.5351	0.9059	(0.3531, 3.3144)	-0.64
		$\gamma$	0.5332	1.0633	(0.0106, 26.570)	0.5126	0.3085	(0.0420, 1.1159)	-0.02
	100	$\alpha$	0.8824	0.5417	(0.2649, 2.9391)	1.0503	0.2935	(0.5560, 1.6301)	-0.28
		$\beta$	1.8358	2.4335	(0.1366, 24.669)	1.5505	0.8160	(0.3592, 3.1978)	-0.92
		$\gamma$	0.6965	0.8995	(0.0554, 8.7534)	0.5576	0.3030	(0.0565, 1.1436)	-0.62
	300	$\alpha$	1.1397	0.3454	(0.6293, 2.0641)	1.1000	0.2500	(0.6227, 1.5625)	-1.24
		$\beta$	1.0844	0.6100	(0.3600, 3.2662)	1.3452	0.6073	(0.4757, 2.6272)	0.93
		$\gamma$	0.3538	0.2788	(0.0755, 1.6581)	0.4509	0.2316	(0.0741, 0.9167)	1.02
(1,3,0.3)	50	$\alpha$	1.2353	0.8137	(0.3397, 4.4922)	1.2420	0.3291	(0.6576, 1.8785)	0.94
		$\beta$	2.0059	2.4786	(0.1780, 22.599)	2.1710	1.0704	(0.6236, 4.2785)	-0.99
		$\gamma$	0.1363	0.3197	(0.0014, 13.523)	0.1810	0.1408	(0.0014, 0.4673)	-0.83
	100	$\alpha$	1.1067	0.4314	(0.5154, 2.3760)	1.1320	0.2524	(0.6901, 1.6137)	-1.15
		$\beta$	2.1190	1.5987	(0.4830, 9.2966)	2.2444	0.9486	(0.8190, 4.1923)	0.60
		$\gamma$	0.1918	0.2438	(0.0159, 2.3160)	0.2193	0.1385	(0.0073, 0.4958)	0.66
	300	$\alpha$	0.9382	0.2503	(0.5562, 1.5828)	1.0097	0.1719	(0.6917, 1.3490)	-0.14
		$\beta$	2.8260	1.5776	(0.9462, 8.4401)	2.6338	0.9259	(1.1688, 4.5379)	-0.83
		$\gamma$	0.2965	0.2308	(0.0645, 1.3638)	0.2685	0.1311	(0.0536, 0.5237)	-0.57
(1.1,0.8,0.2)	50	$\alpha$	0.8733	2.0361	(0.0091, 84.264)	0.8841	0.3061	(0.4011, 1.5184)	1.70
		$\beta$	1.1065	4.5555	(0.0003, 3535.4)	1.3908	0.9101	(0.2301, 3.0955)	-1.92
		$\gamma$	0.3090	1.9087	(0.0001, 56050.8)	0.4095	0.2997	(0.0030, 0.9910)	-1.89
	100	$\alpha$	0.8614	1.1402	(0.0643, 11.531)	0.9355	0.2893	(0.4594, 1.5027)	0.96
		$\beta$	1.1449	2.7800	(0.0098, 133.55)	1.2381	0.7192	(0.2585, 2.7107)	-1.31
		$\gamma$	0.3427	1.1826	(0.0004, 296.55)	0.3676	0.2567	(0.0063, 0.8884)	-0.99
	300	$\alpha$	0.9177	0.4312	(0.3654, 2.3048)	1.0321	0.2469	(0.5959, 1.5424)	-0.20
		$\beta$	1.0417	0.8752	(0.2007, 5.4060)	0.9839	0.4582	(0.3228, 1.8787)	-0.75
		$\gamma$	0.2824	0.3683	(0.0219, 3.6383)	0.2556	0.1745	(0.0105, 0.6089)	-0.58
(0.8,0.4,0.6)	50	$\alpha$	0.5640	1.0525	(0.0145, 21.871)	0.6393	0.2731	(0.2332, 1.1999)	-0.43
		$\beta$	0.6678	2.6145	(0.0003, 1435.6)	0.7810	0.6224	(0.0798, 2.0711)	0.21
		$\gamma$	0.8177	2.8632	(0.0001, 782.13)	0.8180	0.5077	(0.0163, 1.7767)	0.17
	100	$\alpha$	0.6051	0.6336	(0.0777, 4.7110)	0.7045	0.3047	(0.2833, 1.2595)	0.14
		$\beta$	0.7541	1.7710	(0.0076, 75.252)	0.8645	0.7311	(0.0575, 2.2341)	-1.55
		$\gamma$	0.9164	1.7867	(0.0201, 41.846)	0.8857	0.5286	(0.0049, 1.8716)	-1.25
	300	$\alpha$	0.6252	0.2538	(0.2822, 1.3852)	0.6312	0.1759	(0.3377, 0.9874)	-0.46
		$\beta$	0.8545	0.7831	(0.1418, 5.1501)	1.1498	0.8357	(0.2049, 2.6883)	0.14
		$\gamma$	0.9018	0.6922	(0.2003, 4.0598)	1.0138	0.4764	(0.1891, 1.8991)	0.21

**Table 2: Maximum likelihood and Bayesian estimates for the parameters of the model based on EDW distribution.**

Nominal values ( $\alpha, \beta, \gamma$ )			Maximum likelihood estimates			Bayesian estimates			
$\alpha, \beta, \gamma$	$n$	Parameter	Estimate	Std. error	95% CI	Mean	Std. dev.	95% HPD interval	Geweke Z-score
(2,1,0.15)	50	$\alpha$	1.6231	2.3455	(0.0956, 27.565)	1.7005	0.5616	(0.7877, 2.8636)	1.56
		$\beta$	1.5783	3.9419	(0.0118, 210.94)	1.7157	0.9562	(0.4185, 3.5768)	-1.94
		$\gamma$	0.2458	1.0112	(0.0001, 780.01)	0.2921	0.2214	(0.0012, 0.7237)	-1.75
	100	$\alpha$	1.4716	1.6024	(0.1742, 12.435)	1.6032	0.4628	(0.8354, 2.5040)	0.39
		$\beta$	1.4762	2.8402	(0.0340, 64.099)	1.5109	0.7769	(0.4260, 3.0059)	-1.31
		$\gamma$	0.2886	0.8556	(0.0001, 96.267)	0.2983	0.2063	(0.0113, 0.7144)	-1.17
	300	$\alpha$	1.5224	1.0345	(0.4019, 5.7669)	1.7087	0.4148	(1.0041, 2.5029)	1.32
		$\beta$	1.3427	1.5675	(0.1362, 13.234)	1.2496	0.5034	(0.4868, 2.3626)	-1.16
		$\gamma$	0.2395	0.4664	(0.0053, 10.885)	0.2157	0.1408	(0.0131, 0.5178)	-1.15

each parameter, as expected under asymptotic theory. However, we observe that when considering the nominal values  $(\alpha, \beta, \gamma) = (0.8, 0.4, 0.6)$ , the curves for the bias and MSE approach zero more slowly, and the estimates of the coverage probabilities do not reach the 95% nominal value even for larger sample sizes. Panel (m) of Figure 1 suggests that the *pmf* of the EDW distribution given these parameter values has a flat shape after  $t = 0$ .

### 3.4. Applications to real data sets

This section presents applications of the methods discussed in this article to real data sets from the literature.

#### 3.4.1. Survival of HIV-infected patients with high-grade non-Hodgkin's lymphomas

Let us consider as a first example the statistical analysis of a data set where a cure fraction is not assumed. Manyau et al.(2020) published a retrospective review of medical records of adult HIV-infected patients treated in Zimbabwe for high-grade large cell non-Hodgkin's lymphoma between 2015–2017. We consider a sub-sample of  $n = 92$  patients who received three or more cycles of treatment. The variable of interest is the overall survival (OS), defined as the time in months since the diagnosis of lymphoma until the date of death. According to the study planning, 26 patients were right-censored at 20 months after the diagnosis. Table 4 shows the ML and Bayesian estimates considering the EDW distribution and their special cases previously described in section 1: discrete Weibull (DW), discrete exponential (DE), discrete Rayleigh (DR), discrete generalized exponential (DGE) and discrete generalized Rayleigh (DGR). Considering the Bayesian approach introduced in the subsection 2.4, prior distributions of  $\alpha, \beta$  and  $\gamma$  were assumed as  $\alpha \sim \text{Gamma}(1, 1)$ ,  $\beta \sim \text{Gamma}(1, 1)$  and  $\gamma \sim \text{Gamma}(1, 1)$ . MCMC methods were used to get the posterior summaries of interest. Convergence of the simulation algorithm was achieved for all parameters based on the Geweke's statistics.

We can note that the Bayesian and maximum likelihood estimates showed in Table 4 are very similar. Figure 3 compares the survival function estimated by the Kaplan-Meier method and fitted by parametric models using the ML estimation method (left panel) and the Bayesian approach (right panel). Figure 3 also provides the AIC, BIC, and AICC values for the frequentist models and LPML values for the Bayesian models. The models based on the EDW and DGE distributions have the lowest AIC and AICC values, and under the Bayesian approach, these models have the highest LPML values, suggesting that they have the best fit to data among the candidate models.

#### 3.4.2. Survival of patients with non-Hodgkin's lymphoma and low ADMA concentration

In this second example, we consider a medical data set with the presence of a cure fraction. Let us consider a sub-sample of  $n = 14$  patients with non-Hodgkin's lymphoma and low asymmetric dimethylarginines (ADMA) plasma concentrations of the Hematology Clinic of Wroclaw Medical University in Poland. This data set was introduced by Chachaj et al.(2018) and referred to the time-to-death of the patients. These times are: 1, 1, 1, 1, 5, 7, 13, 16, 40+, 40+, 40+, 40+, 40+ and 40+ months, where + denotes a censored observation. We have six censored observations at

**Table 3: Maximum likelihood and Bayesian estimates for the parameters of the model based on EDW distribution including a cure fraction given by  $\eta$ . Nominal values for the parameters  $\alpha$ ,  $\beta$  and  $\gamma$  are 1.1, 0.8 and 0.2, respectively.**

$\eta$	$n$	Parameter	Maximum likelihood estimates			Bayesian estimates			
			Estimate	Std. error	95% CI	Mean	Std. dev.	95% HPD interval	Geweke Z-score
0.2	50	$\alpha$	1.4319	1.1571	(0.2938, 6.9780)	0.8772	0.3776	(0.2966, 1.6375)	1.59
		$\beta$	0.9368	1.2822	(0.0641, 13.6976)	1.4885	0.9285	(0.2367, 3.3811)	-0.85
		$\gamma$	0.1603	0.4262	(0.0001, 29.3792)	0.4210	0.3097	(0.0013, 1.0356)	-1.11
		$\eta$	0.2751	0.0759	(0.1264, 0.4238)	0.1763	0.1149	(0.0001, 0.3731)	1.40
	100	$\alpha$	2.9363	1.9790	(0.7834, 11.0058)	1.0407	0.3944	(0.4048, 1.8305)	0.02
		$\beta$	0.2064	0.1729	(0.0400, 1.0657)	1.0276	0.5533	(0.2097, 2.1292)	-0.13
		$\gamma$	0.0018	0.0091	(0.0001, 32.2225)	0.2967	0.2168	(0.0017, 0.7306)	-0.15
		$\eta$	0.2870	0.0514	(0.1862, 0.3878)	0.2638	0.0667	(0.1307, 0.3929)	0.22
	300	$\alpha$	2.6031	1.3381	(0.9505, 7.1291)	1.3115	0.3786	(0.7367, 2.1097)	-0.07
		$\beta$	0.2451	0.1527	(0.0723, 0.8309)	0.6580	0.2333	(0.2445, 1.0951)	-0.18
		$\gamma$	0.0027	0.0096	(0.0001, 2.6496)	0.1129	0.0805	(0.0007, 0.2643)	-0.22
		$\eta$	0.2123	0.0302	(0.1531, 0.2715)	0.1954	0.0388	(0.1230, 0.2696)	0.68
	1000	$\alpha$	0.8260	0.2044	(0.5086, 1.3416)	0.8139	0.1599	(0.5328, 1.1445)	-0.95
		$\beta$	1.3483	0.6018	(0.5622, 3.2337)	1.5359	0.5996	(0.6224, 2.6806)	1.14
		$\gamma$	0.4132	0.2378	(0.1338, 1.2765)	0.4698	0.2026	(0.1127, 0.8498)	0.95
		$\eta$	0.1855	0.0180	(0.1503, 0.2207)	0.1828	0.0181	(0.1471, 0.2183)	-0.29
	5000	$\alpha$	1.0817	0.1007	(0.9013, 1.2983)	1.0577	0.0969	(0.8649, 1.2437)	1.49
		$\beta$	0.8314	0.1289	(0.6134, 1.1267)	0.8813	0.1412	(0.6167, 1.1489)	-1.60
		$\gamma$	0.2166	0.0585	(0.1275, 0.3678)	0.2390	0.0628	(0.1177, 0.3553)	-1.61
		$\eta$	0.2089	0.0071	(0.1951, 0.2227)	0.2084	0.0071	(0.1943, 0.2219)	1.32
0.5	50	$\alpha$	1.1762	1.4896	(0.0983, 14.0750)	0.9987	0.3746	(0.3732, 1.7860)	0.04
		$\beta$	1.0596	2.3561	(0.0136, 82.7680)	1.5492	0.9102	(0.2981, 3.3883)	0.42
		$\gamma$	0.2541	0.9192	(0.0002, 305.189)	0.4379	0.2918	(0.0094, 1.0094)	0.30
		$\eta$	0.4923	0.0806	(0.3344, 0.6502)	0.4670	0.1033	(0.2687, 0.6569)	0.54
	100	$\alpha$	3.7726	3.5717	(0.5899, 24.1272)	0.9180	0.3904	(0.2642, 1.6764)	0.03
		$\beta$	0.1594	0.1791	(0.0176, 1.441)	1.2199	0.7285	(0.2214, 2.7116)	0.44
		$\gamma$	0.0003	0.0023	(0.0001, 32.2225)	0.3518	0.2600	(0.0002, 0.8824)	0.36
		$\eta$	0.5773	0.0544	(0.4707, 0.6830)	0.5243	0.1192	(0.2479, 0.7065)	1.17
	300	$\alpha$	1.8915	1.1633	(0.5667, 6.3138)	0.8904	0.3354	(0.3500, 1.5599)	0.77
		$\beta$	0.3667	0.2918	(0.0771, 1.7442)	1.3036	0.7165	(0.2980, 2.7406)	-0.90
		$\gamma$	0.0174	0.0538	(0.0001, 7.3739)	0.3340	0.2304	(0.0016, 0.7918)	-0.85
		$\eta$	0.4747	0.0353	(0.4054, 0.5440)	0.4373	0.0608	(0.3219, 0.5377)	0.18
	1000	$\alpha$	0.9901	0.2750	(0.5744, 1.7065)	0.9084	0.2011	(0.5348, 1.2932)	-0.47
		$\beta$	0.9799	0.4627	(0.3884, 2.4723)	1.2971	0.5744	(0.4857, 2.4198)	0.83
		$\gamma$	0.2727	0.2030	(0.0634, 1.1730)	0.3923	0.2089	(0.0807, 0.8165)	0.58
		$\eta$	0.5071	0.0185	(0.4707, 0.5434)	0.5043	0.0191	(0.4672, 0.5423)	-0.13
	5000	$\alpha$	1.0612	0.1260	(0.8408, 1.3393)	1.0226	0.1210	(0.7853, 1.2588)	0.51
		$\beta$	0.8374	0.1665	(0.5671, 1.2365)	0.9245	0.2034	(0.6066, 1.3423)	-0.65
		$\gamma$	0.2291	0.0776	(0.1179, 0.4451)	0.2685	0.0905	(0.1149, 0.4463)	-0.69
		$\eta$	0.5168	0.0080	(0.5012, 0.5325)	0.5162	0.0080	(0.5005, 0.5315)	-0.41

$t = 40$  months. Table 5 shows the ML and Bayesian estimates for the model based on the EDW distribution. For a Bayesian analysis of the model, we assumed the same priors for  $\alpha$ ,  $\beta$  and  $\gamma$  considered in the example of subsection 3.4.1 and a beta  $Beta(0.5, 0.5)$  prior for the parameter  $\eta$ . MCMC methods were used to get the posterior summaries of interest. We note that ML and Bayesian point estimates are relatively close to each other. The large frequentist confidence intervals for the parameters can be attributed to the small sample size.

Figure 4 compares the estimated values of the survival function obtained by the Kaplan-Meier method and by the parametric models using the ML and Bayesian approaches (left and right panels, respectively). The graphs visually suggest that the EDW distribution and its special cases fit the data well, except the model based on the DR distribution,

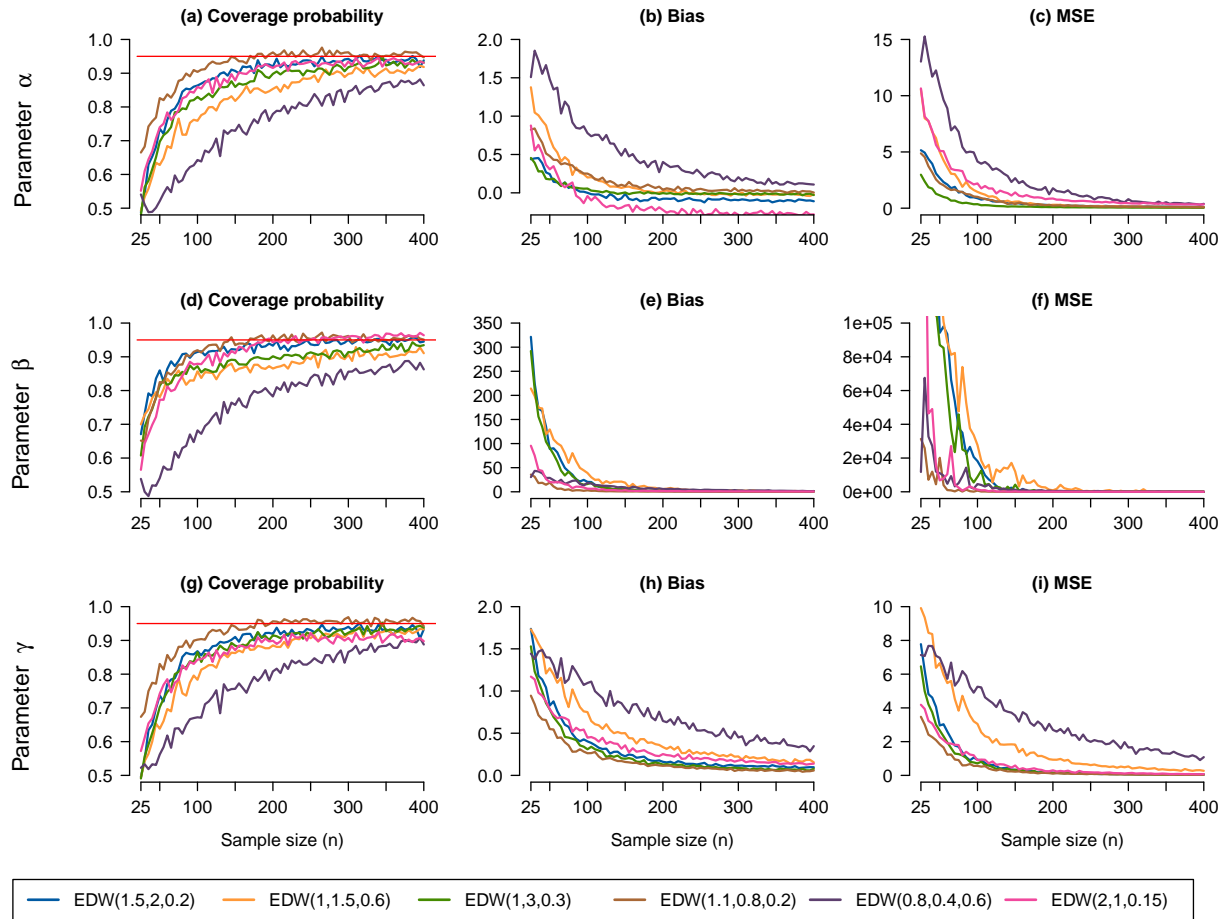


Figure 2: Results from the simulation study. The graphs show the estimates of the coverage probabilities of the 95% confidence intervals for the parameters  $\alpha$ ,  $\beta$  and  $\gamma$  based on the delta method, the average bias, and the mean squared error ( $MSE$ ).

**Table 4: Maximum likelihood and Bayesian estimates for the models based on the EDW distribution and its special cases, applied to the HIV-infected patients data introduced by Manyau et al.(2020).**

Model	Parameter	Maximum likelihood estimates			Bayesian estimates			
		Estimate	Std. error	95% CI	Mean	Std. dev.	95% HPD interval	Geweke Z-score
EDW	$\alpha$	0.8058	0.2955	(0.3927, 1.6535)	0.7812	0.1915	(0.4944, 1.2088)	0.41
	$\beta$	3.1594	2.2408	(0.7868, 12.6857)	3.1371	1.5006	(1.0989, 6.3456)	-0.58
	$\gamma$	0.1794	0.2145	(0.0172, 1.8682)	0.1904	0.1390	(0.0106, 0.4831)	-0.51
DW	$\alpha$	1.5558	0.1879	(1.2278, 1.9714)	1.4490	0.1757	(1.1081, 1.7955)	-0.32
	$\gamma$	0.0094	0.0051	(0.0032, 0.0272)	0.0144	0.0074	(0.0029, 0.0287)	0.33
DE	$\gamma$	0.0430	0.0061	(0.0326, 0.0567)	0.0435	0.0061	(0.0320, 0.0555)	-0.12
DR	$\gamma$	0.0027	0.0004	(0.0020, 0.0036)	0.0027	0.0004	(0.0020, 0.0035)	0.21
DGE	$\beta$	2.2303	0.4531	(1.4978, 3.3212)	2.1597	0.4241	(1.3935, 3.0114)	0.06
	$\gamma$	0.0854	0.0147	(0.0609, 0.1198)	0.0833	0.0140	(0.0566, 0.1112)	0.02
DGR	$\beta$	0.7537	0.1225	(0.5481, 1.0365)	0.7721	0.1214	(0.5423, 1.0102)	1.22
	$\gamma$	0.0019	0.0005	(0.0012, 0.0033)	0.0021	0.0005	(0.0011, 0.0030)	1.50

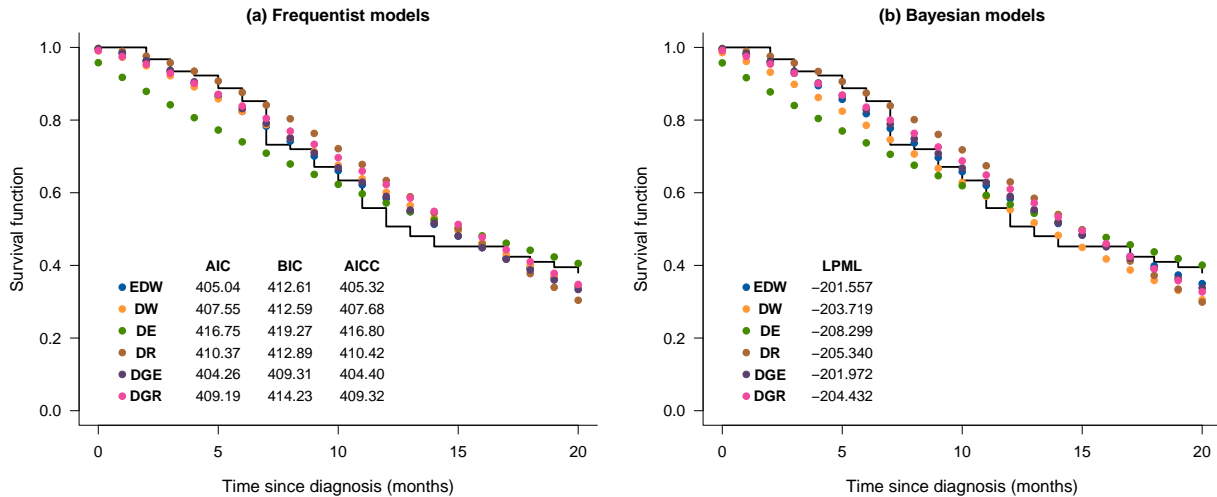


Figure 3: Survival function for the HIV-infected patients data estimated by the Kaplan- Meier method and by using the models based on the EDW distribution and its special cases. ML and Bayesian estimates are shown in panels (a) and (b), respectively.

which seems to be further away from the Kaplan-Meier curve. However, the model based on the DE distribution is the one with the lowest AICC value under the frequentist approach and with the highest LPML value under the Bayesian framework, which suggests that a model based on a baseline distribution with only one parameter may be adequate for these data.

**Table 5: Maximum likelihood and Bayesian estimates for the model based on the EDW distribution with a cure fraction, applied to the non-Hodgkin's lymphoma patients data introduced by Chachaj et al.(2018).**

Model	Parameter	Maximum likelihood estimates			Bayesian estimates			
		Estimate	Std. error	95% CI	Mean	Std. dev.	95% HPD interval	Geweke Z-score
EDW	$\alpha$	0.8158	0.8676	(0.1014, 6.5601)	0.7000	0.3006	(0.2484, 1.3548)	0.55
	$\beta$	1.6529	3.3579	(0.0308, 88.620)	1.4819	1.0146	(0.2914, 3.7454)	-0.54
	$\gamma$	0.3252	0.9493	(0.0011, 99.319)	0.3397	0.2789	(0.0010, 0.9336)	-0.90
	$\eta$	0.4281	0.1327	(0.1680, 0.6882)	0.3943	0.1579	(0.0001, 0.6269)	0.66

#### 4. Concluding Remarks

In this article, the three-parameter EDW distribution introduced by Nekoukhou and Bidram(2015) was studied, considering the inclusion of right-censored data and a cure fraction. The maximum likelihood and the Bayesian estimation methods were employed to estimate the unknown model parameters. Applications to simulated data showed that both the ML and Bayesian approaches are computationally feasible to estimate the parameters of EDW distribution. In a simulation study, we observed that both bias and MSE tend to zero when there is an increase in the sample size  $n$ , suggesting that the ML estimators are asymptotically unbiased for the parameters. Applications with real data have shown that the EDW probability distribution and its special cases can fit the data very well, under both the ML and Bayesian frameworks. The model can be easily implemented in existing free computational programs like R, and the EDW distribution can attract more comprehensive applications in diverse fields.

#### Acknowledgement

The authors want to thank the anonymous referees for their helpful comments on revising the manuscript.

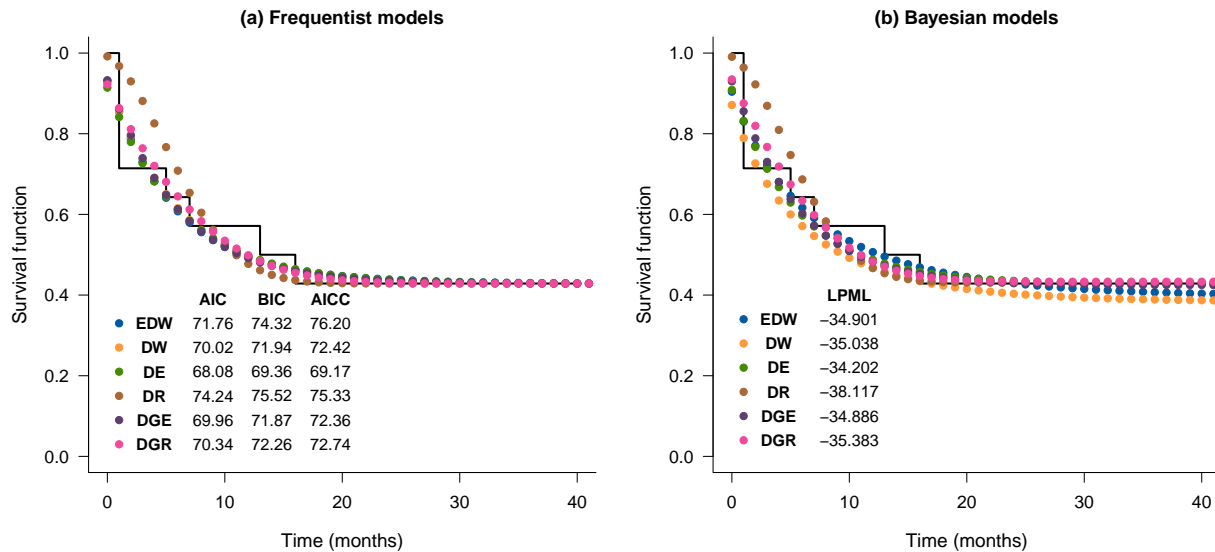


Figure 4: Survival function for the non-Hodgkin's lymphoma patients data estimated by the Kaplan- Meier method and by using the models based on the EDW distribution and its special cases. ML and Bayesian estimates are shown in panels (a) and (b), respectively.

## References

1. Achcar, J. A., Coelho-Barros, E. A., and Mazucheli, J. (2012). Cure fraction models using mixture and non-mixture models. *Tatra Mountains Mathematical Publications*, 51:1–9.
2. Akaike, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723.
3. Alamatsaz, M., Dey, S., Dey, T., and Harandi, S. (2016). Discrete generalized Rayleigh distribution. *Pakistan Journal of Statistics*, 32:1–20.
4. Broyden, C. (1970). The convergence of a class of double-rank minimization algorithms. *Journal of the Institute of Mathematics and Its Applications*, 6:76–90.
5. Cardial, M., Fachini-Gomes, J., and Nakano, E. (2020). Exponentiated discrete Weibull distribution for censored data. *Brazilian Journal of Biometrics*, 38:35–56.
6. Carlin, B. and Louis, T. (2001). *Bayes and Empirical Bayes Methods for Data Analysis*, 2nd ed. Chapman and Hall, Boca Raton.
7. Chachaj, A., Wiśniewski, J., Rybka, J., Butrym, A., Biedroń, M., Krzystek-Korpaczka, M., Fleszar, M. G., Karczewski, M., Wróbel, T., Mazur, G., Gamian, A., and Szuba, A. (2018). Asymmetric and symmetric dimethylarginines and mortality in patients with hematological malignancies — a prospective study. *PLoS One*, 13:e0197148.
8. Chen, M. H. and Shao, Q. M. (1999). Monte carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics*, 8:69–92.
9. Cowles, M. and Carlin, B. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, 91(434):883–904.
10. El Adlouni, S., Favre, A., and Bobée, B. (2006). Comparison of methodologies to assess the convergence of Markov chain Monte Carlo methods. *Computational Statistics & Data Analysis*, 50(10):2685–2701.
11. El-Morshedy, M., Eliwa, M., El-Gohary, A., and Khalil, A. (2020). Bivariate exponentiated discrete Weibull distribution: statistical properties, estimation, simulation and applications. *Mathematical Sciences*, 14:29–42.
12. Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, 38:1041–1046.
13. Fletcher, R. (1970). A new approach to variable metric algorithms. *Computer Journal*, 13:317–322.
14. Geisser, S. and Eddy, W. F. (1979). A predictive approach to model selection. *Journal of the American Statistical Association*, 74:153–160.
15. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian Data*

- Analysis, volume 2.* CRC press Boca Raton, FL.
16. Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bernardo, J. M., Berger, J., Dawid, A. P., Smith, A. F. M. (Eds.), Bayesian Statistics*, volume 4, pages 169–193. Oxford University Press.
  17. Hamdeni, T. and Gasmi, S. (2020). The Marshall–Olkin generalized defective Gompertz distribution for surviving fraction modeling. *Communications in Statistics – Simulation and Computation*, pages 1–14.
  18. Henningsen, A. and Toomet, O. (2011). maxlik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26:443–458.
  19. Hurvich, C. M. and Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76:297–307.
  20. Ibrahim, J. G., Chen, M. H., and Sinha, D. (2014). *Bayesian Survival Analysis*. Wiley Online Library.
  21. Johnson, N. L. and Kotz, S. (1982). Developments in discrete distributions, 1969-1980, correspondent paper. *International Statistical Review/Revue Internationale de Statistique*, 50:71–101.
  22. Manyau, M. C. P., Mudzviti, T., Rusakaniko, S., Mberi, E. T., Maponga, C. C., and Morse, G. D. (2020). Survival of HIV-infected patients with high-grade non-Hodgkin’s lymphomas: A retrospective study of experiences in Zimbabwe. *PLoS ONE*, 15:e0239344.
  23. Martin, A. D. and Quinn, K. M. (2006). Applied Bayesian inference in R using MCMCpack. *R News*, 6:2–7.
  24. Martinez, E. Z., Achcar, J. A., Jácome, A. A., and Santos, J. S. (2013). Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application to gastric cancer data. *Computer Methods and Programs in Biomedicine*, 112:343–355.
  25. Millar, R. B. (2011). *Maximum likelihood estimation and inference: with examples in R, SAS and ADMB*. John Wiley & Sons.
  26. Morbiducci, M., Nardi, A., and Rossi, C. (2003). Classification of cured individuals in survival analysis: the mixture approach to the diagnostic-prognostic problem. *Computational Statistics & Data Analysis*, 41:515–529.
  27. Nakagawa, T. and Osaki, S. (1975). The discrete Weibull distribution. *IEEE Transactions on Reliability*, 24:300–301.
  28. Nekoukhrou, V., Alamatsaz, M. H., and Bidram, H. (2013). Discrete generalized exponential distribution of a second type. *Statistics*, 47:876–887.
  29. Nekoukhrou, V. and Bidram, H. (2015). The exponentiated discrete Weibull distribution. *SORT Statistics and Operations Research Transactions*, 39:127–146.
  30. Nelder, J. and Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7:308–313.
  31. Ramos, P., Guzman, D., Mota, A., Rodrigues, F., and Louzada, F. (2020). Sampling with censored data: a practical guide. *arXiv preprint arXiv:2011.08417*.
  32. Rocha, R., Nadarajah, S., Tomazella, V., and Louzada, F. (2017). A new class of defective models based on the Marshall–Olkin family of distributions for cure rate modeling. *Computational Statistics & Data Analysis*, 107:48–63.
  33. Roy, D. (2004). Discrete Rayleigh distribution. *IEEE Transactions on Reliability*, 53:255–260.
  34. Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6:461–464.
  35. Scudilio, J., Calsavara, V. F., Rocha, R., Louzada, F., Tomazella, V., and Rodrigues, A. S. (2019). Defective models induced by gamma frailty term for survival data with cured fraction. *Journal of Applied Statistics*, 46:484–507.