

EXPONENTIATED DISCRETE WEIBULL DISTRIBUTION FOR CENSORED DATA

Marcílio Ramos Pereira CARDIAL¹
Juliana Betini FACHINI-GOMES¹
Eduardo Yoshio NAKANO¹

- **ABSTRACT:** This paper further develops the statistical inference procedure of the exponentiated discrete Weibull distribution (EDW) for data with the presence of censoring. This generalization of the discrete Weibull distribution has the advantage of being suitable to model non-monotone failure rates, such as those with bathtub and unimodal distributions. Inferences about EDW distribution are presented using both frequentist and bayesian approaches. In addition, the classical Likelihood Ratio Test and a Full Bayesian Significance Test (FBST) were performed to test the parameters of EDW distribution. The method presented is applied to simulated data and illustrated with a real dataset regarding patients diagnosed with head and neck cancer.
- **KEYWORDS:** Survival analysis; discrete distributions; exponentiated distributions; full bayesian significance test (FBST)

1 Introduction

The Weibull distribution is widely used to model survival data. This popularity is due to its versatility and relative simplicity (RINNE, 2008). The Weibull distribution and its generalizations have been applied over the years in survival analysis when the data are continuous. However, in some cases the data are discrete. This can arise, for example, when the survival time is measured in months, cycles or counts. It is not always acceptable to use a continuous model to analyze discrete data. To overcome this problem, Nakagawa and Osaki (1975) introduced the discrete Weibull distribution (DW), which models discrete data with monotone (increasing or decreasing) failure rates (more details on the discretization of Weibull

¹Universidade de Brasília - UnB, Departamento de Estatística, CEP: 70910-900, Brasília, DF, Brasil. E-mail: *marciliocardial@gmail.com; jfachini@unb.br; eynakano@gmail.com*

distribution can be seen in Vila et al. (2019)). Nevertheless, the Nakagawa and Osaki's DW distribution does not have a reasonable parametric fit for phenomena that present non-monotone failure rates, such as those with bathtub or unimodal distribution.

In recent years, new families of distributions for continuous data have been proposed based on modifications of the Weibull distribution to deal with bathtub-shaped or unimodal failure rates. An example is the exponentiated Weibull distribution introduced by Mudholkar and Srivastava (1993), which is part of the family of exponentiated distributions introduced by Gupta and Kundu (2001). An analog for discrete data of the exponentiated Weibull distribution was introduced by Nekoukhou and Bidram (2015). However, inferences about the parameters of the exponentiated discrete Weibull distribution (EDW) have only been presented for uncensored data.

This paper formulates the EDW model in a context of survival analysis, allowing the presence of censored data in the inference process. The inferences are made according to the frequentist and bayesian approaches. The methods developed are illustrated with simulated data and a real dataset on survival time of patients suffering from head and neck cancer. All the simulations and estimates were carried out using the free R software (R CORE TEAM, 2019).

2 Exponentiated discrete Weibull model

Let X be a continuous random variable. The corresponding discrete variable is denoted by $T = [X]$, where $[X]$ represents "the integer part of X ".

By considering that, $X \sim Weibull(\eta, \beta)$ the cumulative distribution function of the discrete Weibull distribution (DW) according to Nakagawa and Osaki (1975) is given by:

$$F_{dw}(t) = 1 - q^{(t+1)^\beta}, \quad t = 0, 1, 2, \dots,$$

where $q = \exp\left\{\frac{-1}{\eta^\beta}\right\}$, $0 < q < 1$ and $\beta > 0$ is the shape parameter.

The cumulative distribution function of the exponentiated discrete Weibull distribution (EDW) can be expressed by:

$$\begin{aligned} F_{edw}(t) &= [F_{dw}(t)]^\gamma \\ &= \left[1 - q^{(t+1)^\beta}\right]^\gamma, \quad t = 0, 1, 2, \dots \end{aligned} \quad (1)$$

Here, $\gamma > 0$ is the shape parameter, which can also be called the resilience parameter (NEKOUKHO; BIDRAM, 2015).

Consequently, the EDW has probability distribution given by:

$$p_{edw}(t) = \left[1 - q^{(t+1)^\beta}\right]^\gamma - \left[1 - q^{t^\beta}\right]^\gamma, \quad t = 0, 1, 2, \dots \quad (2)$$

where $\gamma > 0$, $\beta > 0$ and $0 < q < 1$.

The discrete Weibull distribution (DW), Exponentiated geometric distribution (EG), geometric distribution (G), Exponentiated discrete Rayleigh distribution (EDR) and discrete Rayleigh distribution (DR) are special cases of the EDW distribution. The EDW distribution reduces to:

- DW distribution, if $\gamma = 1$;
- EG distribution, if $\beta = 1$;
- G distribution, if $\gamma = 1$ and $\beta = 1$;
- EDR distribution, if $\beta = 2$;
- DR distribution, if $\gamma = 1$ and $\beta = 2$ (NEKOUKHOU; BIDRAM, 2015).

For illustration of the special cases of EDW distribution, Figure 1 shows the probability distribution for different values of q , β and γ .

The survival function of the EDW distribution is expressed by:

$$S_{edw}(t) = 1 - \left[1 - q^{(t+1)^\beta}\right]^\gamma, \quad t = 0, 1, 2, \dots \quad (3)$$

Through equations (2) and (3), the failure rate function of the EDW distribution is defined by:

$$\begin{aligned} h_{edw}(t) &= \frac{p_{edw}(t)}{S_{edw}(t) + p_{edw}(t)} \\ &= \frac{\left[1 - q^{(t+1)^\beta}\right]^\gamma - \left[1 - q^{t^\beta}\right]^\gamma}{1 - \left[1 - q^{t^\beta}\right]^\gamma}, \quad t = 0, 1, 2, \dots \end{aligned} \quad (4)$$

Figure 2 presents the failure rate graph of the EDW distribution for different values of q , β and γ .

One of the advantages of working with this generalization of the DW distribution is the fact it provides a parametric fit that allows working with various types of failure rates described in the literature.

Figure 2 shows that depending on the combination of parameters, the failure rate function assumes increasing, decreasing, bathtub or unimodal behavior. Thus, the EDW distribution is more flexible to fit data than the DW distribution, which accommodates only monotone and constant failure rates.

An alternative way to find the EDW is first to exponentiated the continuous Weibull distribution, obtaining the exponentiated Weibull distribution proposed by Mudholkar and Srivastava (1993), and then to discretize the exponentiated continuous Weibull to obtain the exponentiated discrete Weibull distribution.

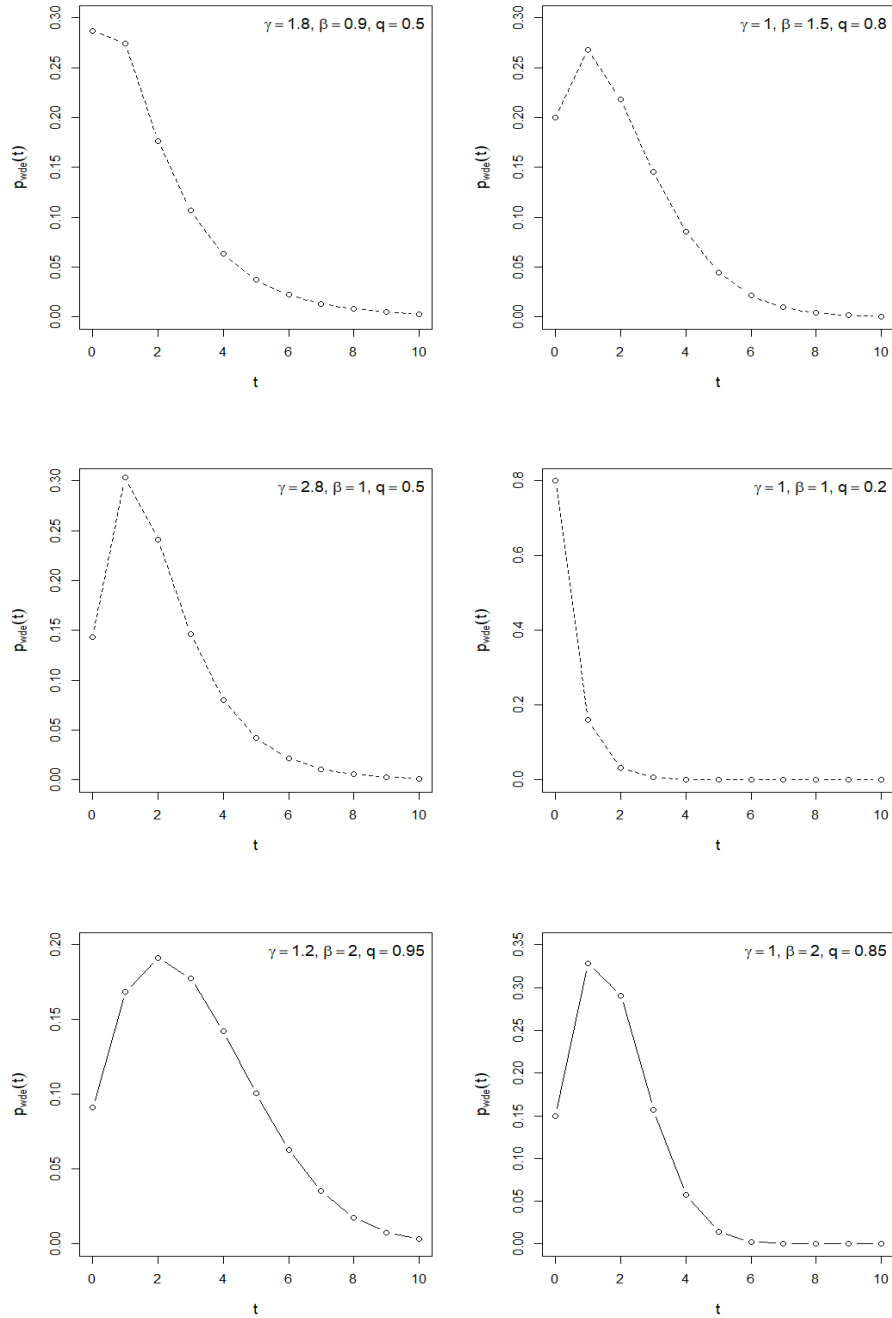
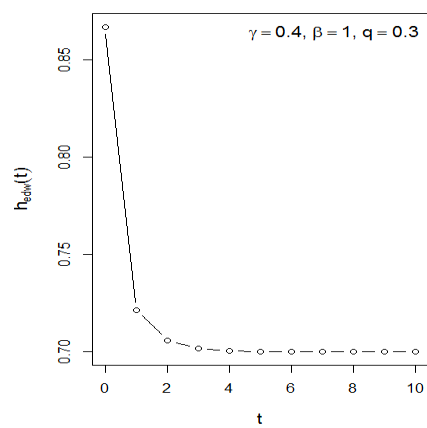
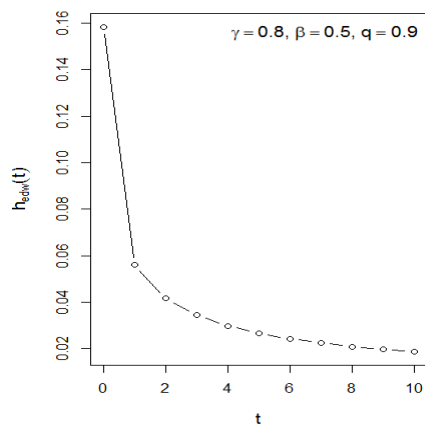
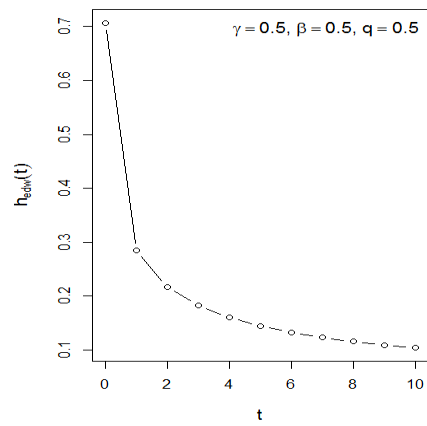
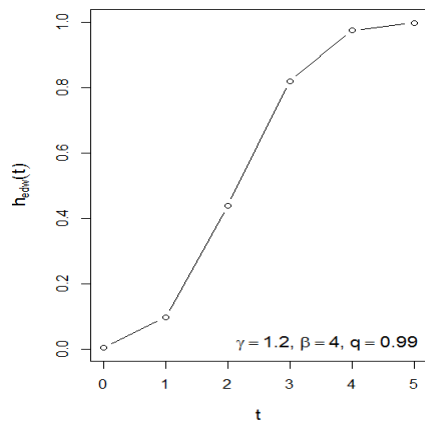
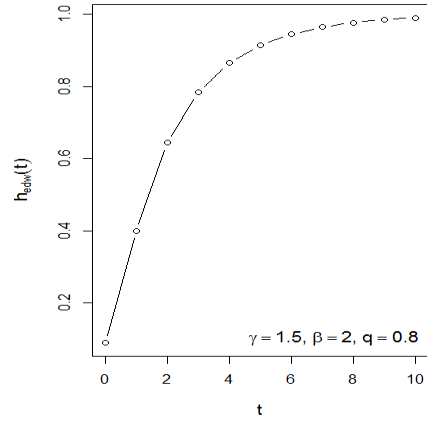
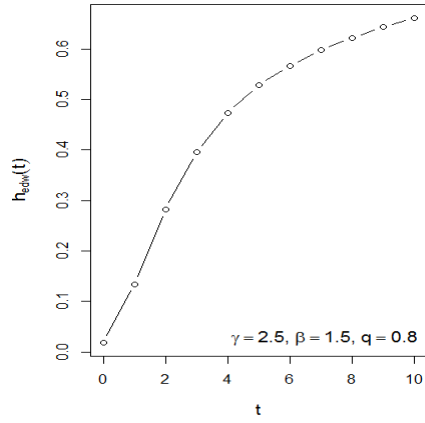


Figure 1 - Probability distribution of the exponentiated discrete Weibull distribution for different values of q , β and γ .



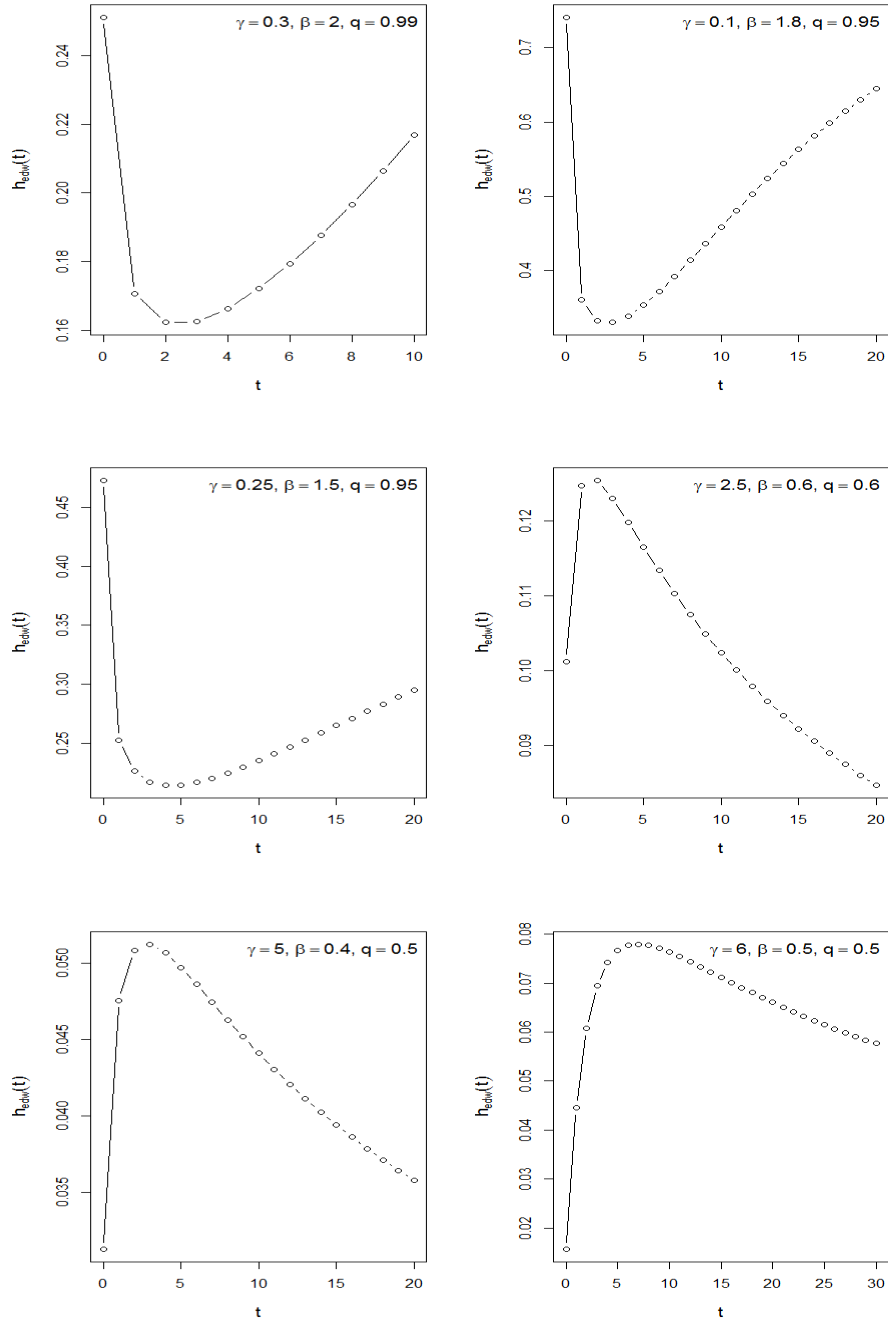


Figure 2 - Failure rate of the exponentiated discrete Weibull distribution for different values of q, β and γ .

3 Maximum likelihood estimators

Let $\mathbf{t} = (t_1, t_2, \dots, t_n)$, be a random sample observed from a population with the probability distribution and survival function of the EDW defined in equations (2) and (3). The vector of parameters $\boldsymbol{\theta}$ is defined by $\boldsymbol{\theta} = (q, \beta, \gamma)$ and $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_n)$, where δ_i denotes the censoring indicator variable, which is equal to 1 if time t_i is failure and 0 for right censoring according to Colosimo and Giolo (2006). The likelihood function of the EDW is given by:

$$L(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta}) \propto \prod_{i=1}^n \left\{ \left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma \right\}^{\delta_i} \left\{ 1 - \left[1 - q^{(t_i+1)^\beta} \right]^\gamma \right\}^{(1-\delta_i)} \quad (5)$$

where $\gamma > 0$, $\beta > 0$ and $0 < q < 1$.

The log likelihood function is given by:

$$\begin{aligned} \ell(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta}) &= \sum_{i=1}^n \left\{ \delta_i \log \left\{ \left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma \right\} \right\} + \\ &\quad \sum_{i=1}^n \left\{ (1 - \delta_i) \log \left\{ 1 - \left[1 - q^{(t_i+1)^\beta} \right]^\gamma \right\} \right\} + c, \end{aligned} \quad (6)$$

where c is a constant that does not depend on $\boldsymbol{\theta}$.

Deriving (6) in relation to q , β and γ we obtain the likelihood equations (Appendix A):

$$\frac{\partial \ell(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta})}{\partial q}, \quad \frac{\partial \ell(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta})}{\partial \beta} \quad \text{and} \quad \frac{\partial \ell(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta})}{\partial \gamma}. \quad (7)$$

The values of \hat{q} , $\hat{\beta}$ and $\hat{\gamma}$ that satisfy equations (7) are the maximum likelihood estimates (MLE) of the EDW distribution. Note that the solutions of equation (7) cannot be obtained analytically. However, the MLE can be obtained numerically through Newton-Raphson algorithm.

Since the parameters of the EDW distribution are limited in the parametric spaces, i.e., $\gamma > 0$, $\beta > 0$ and $0 < q < 1$, it is necessary to perform a transformation to make them unconstrained. Thus, for the parameter q we considered the log-log transformation and for the parameters β and γ we considered the logarithmic transformation. This causes the confidence intervals of q , β and γ to be given, respectively, by:

$$\left[(\hat{q})^{e^{Z_{1-\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{u})}}}; (\hat{q})^{e^{-Z_{1-\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{u})}}} \right], \quad (8)$$

where $\hat{u} = \log[-\log \hat{q}]$,

$$\left[\hat{\beta}^{e^{-Z_{1-\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{v})}}}; \hat{\beta}^{e^{Z_{1-\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{v})}}} \right], \quad (9)$$

where $\hat{v} = \log \hat{\beta}$, and

$$\left[\hat{\gamma} e^{-Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{w})}}; \hat{\gamma} e^{Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{w})}} \right], \quad (10)$$

where $\hat{w} = \log \hat{\gamma}$ and $Z_{1-\alpha/2}$ is the quantile $(1 - \alpha/2)$ of a standard normal distribution.

The estimates of the variances, $\widehat{Var}(\hat{u})$, $\widehat{Var}(\hat{v})$ and $\widehat{Var}(\hat{w})$, can be obtained numerically by the delta method (CASELLA; BERGER, 2002).

4 Bayesian inference

In considering the likelihood function (5) and $\boldsymbol{\theta} = (q, \beta, \gamma)$, since $\beta > 0$, $\gamma > 0$ and $0 < q < 1$, we considered the following prior distributions for the parameters: $\beta \sim \text{Gamma}(a_1, b_1)$, $\gamma \sim \text{Gamma}(a_2, b_2)$ and $q \sim \text{Beta}(a_3, b_3)$, where a_1 , a_2 , a_3 , b_1 , b_2 and b_3 are known positive hyperparameters. Assuming independence of the prior parameters, the joint posterior distribution of $\boldsymbol{\theta} = (q, \beta, \gamma)$ is proportional to:

$$\begin{aligned} \pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}) &\propto \pi(\boldsymbol{\theta}) L(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta}) \\ &\propto \beta^{a_1-1} e^{-b_1 \beta} \gamma^{a_2-1} e^{-b_2 \gamma} q^{a_3-1} (1-q)^{b_3-1} \\ &\quad \prod_{i=1}^n \left\{ \left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma \right\}^{\delta_i} \left\{ 1 - \left[1 - q^{(t_i+1)^\beta} \right]^\gamma \right\}^{(1-\delta_i)} \end{aligned} \quad (11)$$

and the posterior conditional distributions are given by:

$$\pi(\beta | \gamma, q, \mathbf{t}, \boldsymbol{\delta}) \propto \beta^{a_1-1} e^{-b_1 \beta} \prod_{i=1}^n \left\{ \left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma \right\}^{\delta_i} \left\{ 1 - \left[1 - q^{(t_i+1)^\beta} \right]^\gamma \right\}^{(1-\delta_i)},$$

$$\pi(\gamma | \beta, q, \mathbf{t}, \boldsymbol{\delta}) \propto \gamma^{a_2-1} e^{-b_2 \gamma} \prod_{i=1}^n \left\{ \left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma \right\}^{\delta_i} \left\{ 1 - \left[1 - q^{(t_i+1)^\beta} \right]^\gamma \right\}^{(1-\delta_i)}$$

and

$$\pi(q | \gamma, \beta, \mathbf{t}, \boldsymbol{\delta}) \propto q^{a_3-1} (1-q)^{b_3-1} \prod_{i=1}^n \left\{ \left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma \right\}^{\delta_i} \left\{ 1 - \left[1 - q^{(t_i+1)^\beta} \right]^\gamma \right\}^{(1-\delta_i)}.$$

The posterior distribution (11) cannot be obtained analytically, but samples from it can be obtained numerically through MCMC (Markov Chain Monte Carlo) methods by applying steps of the Metropolis-Hastings algorithm (METROPOLIS et al., 1953; HASTINGS, 1970).

The point and interval estimates of the parameters were obtained from the posterior mean and its respective highest posterior density (HPD) interval. The

predictive values of the survival function of the EDW distribution are given by:

$$\begin{aligned}
 \widehat{S}(t) &= \int_0^\infty \int_0^\infty \int_0^1 P[T > t, \boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}] dq d\beta d\gamma \\
 &= \int_0^\infty \int_0^\infty \int_0^1 \left(1 - [1 - q^{(t+1)^\beta}]^\gamma\right) \pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}) dq d\beta d\gamma \\
 &= E_{\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}} \left[1 - [1 - q^{(t+1)^\beta}]^\gamma\right], \quad t = 0, 1, 2, \dots,
 \end{aligned} \tag{12}$$

where $\boldsymbol{\theta} = (q, \beta, \gamma)$ and $\pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta})$ is the posterior distribution given by (11).

Note that the expectancy in (12) can be easily obtained from the values generated by the posterior distribution (11).

We performed hypothesis testing of the parameters of the EDW model by the full bayesian significance test (FBST), which is based on the e -value (evidence of the null hypothesis). To describe the FBST, the interest is to test a null hypothesis $H_0 : \boldsymbol{\theta} \in \Theta_0$, where $\Theta_0 \subset \Theta$. Let Θ be the parametric space and $\boldsymbol{\theta} = (q, \beta, \gamma)$ the vector of parameters of the EDW distribution. The set tangent to H_0 is:

$$T_{H_0} = \{\boldsymbol{\theta} \in \Theta : \pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}) > \pi(\boldsymbol{\theta}_{H_0}^* | \mathbf{t}, \boldsymbol{\delta})\},$$

where $\boldsymbol{\theta}_{H_0}^*$ is the vector of parameters that maximizes the posterior density, $\pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta})$, under the null hypothesis H_0 . Hence, the proposed evidence measure (the e -value) is defined by:

$$e\text{-value}(H_0) = 1 - P(\boldsymbol{\theta} \in T_{H_0} | \mathbf{t}, \boldsymbol{\delta}).$$

The e -value is the volume (probability) of the posterior density in the set of the parametric spaces that consists of the points with posterior density smaller than the maximum point of the density under the hypothesis (De BRAGANÇA PEREIRA and STERN, 1999). More details about the FBST can be found in De Bragança Pereira and Stern (1999) and Madruga et al. (2001).

5 Simulation study

This section presents the computational simulations and their results, which were obtained with the R software. The objective of the simulations is to generate survival times of the EDW distribution by the inversion method, considering a random right censoring mechanism. In the same way as Brunello and Nakano (2015), the censoring was incorporated in the samples independently of the survival time by the censoring indicator variable generated by a Bernoulli distribution, with the censoring percentages specified below.

We considered different scenarios for the simulation, considering the behavior of the failure rate in the distribution under analysis, as depicted in Figure 2. Table 1 describes the scenarios chosen for the simulations.

Table 1 - Scenarios used in the simulations

Scenario	q	β	γ	Failure rate
1	0.80	1.5	2.5	Increasing
2	0.90	0.5	0.8	Decreasing
3	0.99	2.0	0.3	Bathtub
4	0.50	0.5	6.0	Unimodal

The bias and mean square error (MSE) of the estimators of the exponentiated discrete Weibull model were calculated from 10,000 Monte Carlo replicates considering sample sizes of $n = 30, 80, 100, 200$ and 500 and censoring percentages of $0\%, 10\%$ and 20% . Table 2 presents the estimated bias and MSE values for each parameter according to Scenario 1, for different sample sizes and censoring percentages.

Table 2 - Estimates of bias and MSE for Scenario 1

		Scenario 1					
		$q = 0.80, \beta = 1.5, \gamma = 2.5$					
n	Parameters	Censoring Percentages					
		0%		10%		20%	
		Bias	MSE	Bias	MSE	Bias	MSE
30	q	-0.0041	0.0017	-0.0004	0.0022	0.0035	0.0028
	β	0.0206	0.0284	-0.0218	0.0338	-0.0643	0.0441
	γ	0.1303	0.0877	0.1419	0.0782	0.1529	0.0695
80	q	-0.0048	0.0007	-0.0043	0.0012	-0.0021	0.0016
	β	0.0003	0.0104	-0.0501	0.0157	-0.0971	0.0272
	γ	0.0959	0.0545	0.1354	0.0739	0.1647	0.0599
100	q	-0.0055	0.0007	-0.0049	0.0010	-0.0031	0.0014
	β	-0.0037	0.0806	-0.0531	0.0141	-0.1009	0.0252
	γ	0.0925	0.0573	0.1317	0.0648	0.1659	0.0538
200	q	-0.0049	0.0004	-0.0064	0.0008	-0.0046	0.0011
	β	-0.0060	0.0051	-0.0616	0.0106	-0.1076	0.0221
	γ	0.0692	0.0461	0.1309	0.0780	0.1770	0.0655
500	q	-0.0046	0.0004	-0.0071	0.0006	-0.0058	0.0008
	β	-0.0076	0.0032	-0.0666	0.0083	-0.1143	0.0197
	γ	0.0541	0.0428	0.1246	0.0729	0.1751	0.0627

The results reported in Table 2 show the accuracy of the proposed model, since the estimates of bias and MSE are small, even for relatively small samples ($n = 30$), and the bias estimates tend to values near zero when the sample size increases.

Table 3 presents the estimated bias and MSE values for each parameter in Scenario 2 for different sample sizes and censoring percentages

Table 3 - Estimates of bias and MSE for Scenario 2

		Scenario 2					
		$q = 0.90, \beta = 0.5, \gamma = 0.8$					
n	Parameters	Censoring Percentages					
		0%		10%		20%	
		Bias	MSE	Bias	MSE	Bias	MSE
30	q	-0.0042	0.0011	0.0052	0.0009	0.0149	0.0011
	β	0.0092	0.0035	0.0079	0.0039	0.0072	0.0047
	γ	0.0289	0.0157	0.0394	0.0107	0.0461	0.0094
80	q	-0.0043	0.0005	0.0037	0.0005	0.0129	0.0007
	β	0.0001	0.0013	-0.0032	0.0015	-0.0052	0.0019
	γ	0.0246	0.0096	0.0389	0.0095	0.0479	0.0079
100	q	-0.0045	0.0005	0.0030	0.0005	0.0127	0.0006
	β	-0.0014	0.0011	-0.0046	0.0014	-0.0065	0.0017
	γ	0.0241	0.0095	0.0388	0.0102	0.0472	0.0083
200	q	-0.0041	0.0004	0.0034	0.0005	0.0129	0.0005
	β	-0.0025	0.0008	-0.0059	0.0010	-0.0082	0.0011
	γ	0.0184	0.0070	0.0363	0.0095	0.0483	0.0085
500	q	-0.0035	0.0002	0.0041	0.0003	0.0134	0.0004
	β	-0.0031	0.0005	-0.0061	0.0007	-0.0084	0.0007
	γ	0.0141	0.0042	0.0331	0.0059	0.0472	0.0053

It can be seen in Table 3 that the bias and MSE estimates are larger when the sample is small and gradually decline as the sample increases, attesting to the accuracy of the EDW model.

Table 4 refers to the estimated bias and MSE values for each parameter in Scenario 3 for different sample sizes and censoring percentages.

Scenario 3 considers a bathtub-shaped failure rate. The estimated bias and MSE values presented in Table 4 are very near zero irrespective of the sample size, indicating the accuracy of the EDW model for this scenario.

Table 4 - Estimates of bias and MSE for Scenario 3

		Scenario 3					
		$q = 0.99, \beta = 2.0, \gamma = 0.3$					
n	Parameters	Censoring Percentages					
		0%		10%		20%	
		Bias	MSE	Bias	MSE	Bias	MSE
30	q	0.0002	1×10^{-5}	0.0010	1×10^{-5}	0.0019	1×10^{-5}
	β	0.0571	0.0079	0.0357	0.0047	0.0195	0.0029
	γ	0.0157	0.0059	0.0377	0.0087	0.0615	0.0127
80	q	0.0006	2×10^{-6}	0.0014	5×10^{-6}	0.0026	1×10^{-5}
	β	0.0378	0.0041	0.0184	0.0015	0.0126	0.0011
	γ	0.0029	0.0019	0.0205	0.0029	0.0397	0.0047
100	q	0.0006	2×10^{-6}	0.0014	4×10^{-6}	0.0027	1×10^{-5}
	β	0.0337	0.0034	0.01482	0.0011	0.0111	0.0008
	γ	0.0015	0.0016	0.0187	0.0023	0.0366	0.0038
200	q	0.0005	1×10^{-6}	0.0014	3×10^{-6}	0.0028	9×10^{-6}
	β	0.0244	0.0017	0.0071	0.0004	0.0089	0.0006
	γ	-0.0007	0.0008	0.0169	0.0013	0.0317	0.0023
500	q	0.0004	7×10^{-7}	0.0015	2×10^{-6}	0.0029	9×10^{-6}
	β	0.0161	0.0008	0.0039	0.0001	0.0088	0.0005
	γ	-0.0018	0.0004	0.0152	0.0006	0.02787	0.0013

Table 5 contains the estimated bias and MSE values for each parameter in Scenario 4 for different sample sizes and censoring percentages.

The estimated bias and MSE values presented in Table 5 are relatively small and tend to decrease as the sample size increases, again confirming the accuracy of the model.

Table 5 - Estimates of bias and MSE for Scenario 4

		Scenario 4					
		$q = 0.50, \beta = 0.5, \gamma = 6.0$					
n	Parameters	Censoring Percentages					
		0%		10%		20%	
		Bias	MSE	Bias	MSE	Bias	MSE
30	q	0.0019	0.0042	0.0068	0.0046	0.0102	0.0051
	β	0.0134	0.0053	-0.0002	0.0055	-0.0159	0.0061
	γ	0.2590	0.1357	0.3222	0.1746	0.3836	0.2173
80	q	-0.0049	0.0015	-0.0033	0.0016	-0.0017	0.0019
	β	0.0016	0.0018	-0.0139	0.0021	-0.0302	0.0029
	γ	0.2339	0.1020	0.3188	0.1562	0.4172	0.2273
100	q	-0.0063	0.0012	-0.0051	0.0013	-0.0039	0.0015
	β	-0.0002	0.0014	-0.0156	0.0018	-0.0324	0.0027
	γ	0.2314	0.1014	0.3179	0.1557	0.4257	0.2300
200	q	-0.0074	0.0007	-0.0082	0.0007	-0.0082	0.0009
	β	-0.0023	0.0007	-0.0193	0.0011	-0.0365	0.0022
	γ	0.2021	0.0937	0.3304	0.1667	0.4711	0.2594
500	q	-0.0075	0.0004	-0.0105	0.0005	-0.0119	0.0006
	β	-0.0033	0.0003	-0.0216	0.0008	-0.0394	0.0019
	γ	0.1608	0.0889	0.3589	0.1926	0.5293	0.3113

In general, the estimated bias and MSE values decline with increasing sample size in the situation without censoring, since the likelihood function only relies on the probability distribution. With censoring, bias will always exist, because the “true” values of the parameters do not consider censoring, so it is expected for bias to appear. It is natural for the estimates of bias and MSE to increase when the censoring percentage is higher since these estimates are naturally biased by the fact the fact that the likelihood function in the presence of censoring counts on the contribution of the survival function. The accuracy of the EDW was demonstrated in all the scenarios. Therefore, the model can be used when the failure rate is increasing (Scenario 1) and decreasing (Scenario 2). Besides this, another advantage is that the model can be used when the risk function has a bathtub (Scenario 3) or unimodal (Scenario 4) distribution.

We also performed a sensitivity study of the choice of hyperparameters of the prior distribution of γ , where $\gamma \sim Gamma(a_2, b_2)$. We chose γ because the main inferences in this work are focused on this parameter. We kept the values of the hyperparameters of the prior of β (a_1 and b_1) constant and equal to 0.001 and the values of the hyperparameters of the prior of q (a_3 and b_3) constant and equal to 1. In other words, we considered non-informative (diffuse) priors for the parameters β and q . This sensitivity study considered Scenario 1 ($q = 0.8, \beta = 1.5, \gamma = 2.5$). The estimates were based on a chain with size 10,000, considering a burn-in of 1,000. The convergence of marginal posteriors was verified by Geweke Diagnostic. Table

6 presents the results of a sensitivity study of the choice of hyperparameters of the prior distribution of γ with the point estimates of the (posterior mean) of the EDW model, the respective 95% HPD credible intervals and the e -value of the FBST for hypothesis $H_{01} : \gamma = 1$ of Scenario 1, considering $n = 200$ and 0% censoring. Here we decided to use a sample without censoring to allow better control of the true value of the parameters in the samples.

Table 6 - Influence of the choice of hyperparameters a_2 and b_2 of the prior of γ on the inference of the parameters of the EDW distribution for Scenario 1

a_2	b_2	Parameters	Estimates	HPD Interval 95%	e -value $H_{01} : \gamma = 1$
0.001	0.001	q	0.7271	[0.6291; 0.8176]	0.0274
		β	1.3971	[1.1965; 1.5808]	
		γ	3.1249	[2.1831; 4.1952]	
0.01	0.01	q	0.7303	[0.6352; 0.8177]	0.0314
		β	1.4036	[1.2242; 1.6048]	
		γ	3.0869	[2.2014; 4.1546]	
1	1	q	0.7435	[0.6535; 0.8282]	0.0677
		β	1.4303	[1.2360; 1.6050]	
		γ	2.9467	[2.1469; 3.9197]	
1	2	q	0.7595	[0.6668; 0.8385]	0.1442
		β	1.4649	[1.2583; 1.6461]	
		γ	2.7866	[1.9944; 3.7073]	
1	3	q	0.7752	[0.6952; 0.8495]	0.2115
		β	1.5002	[1.3026; 1.6839]	
		γ	2.6279	[1.9229; 3.3796]	
3	1	q	0.7296	[0.6363; 0.8209]	0.0099
		β	1.4020	[1.2080; 1.5929]	
		γ	3.0963	[2.1644; 4.1045]	
5	1	q	0.7156	[0.6200; 0.8089]	0.0002
		β	1.3739	[1.1869; 1.5578]	
		γ	3.2418	[2.2769; 4.3217]	

Table 6 shows that the inference is robust when the prior variance is not small. Note that the estimates of the parameters and their HPD intervals do not vary greatly when the prior variance of γ is larger than 1, even when varying the values of the prior mean. Recall that if $\gamma \sim \text{Gamma}(a_2, b_2)$, then the mean and variance of γ are, respectively, a_2/b_2 and a_2/b_2^2 (according to the parameterization adopted here). However, in cases when the prior variance is small (informative priors), and hence the prior mean is smaller than 1, the values of the posterior mean of γ are smaller. In these cases it is possible to note the influence of the priors on the posterior estimates of the parameters.

We performed the FBST for each pair of hyperparameters in Table 6, considering the limit of 0.05 for the e -value. Note that in Scenario 1, the true value of

the parameter adopted in the simulations ($\gamma=2.5$) is different than 1. Therefore, for diffuse priors (priors with large variance), the inference is dominated by the sample, and consequently hypothesis $H_{01} : \gamma = 1$ is rejected (e -value smaller than 0.05). Besides this, as a_2 increases (the prior mean of γ increases), the posterior mean of γ tends to rise. Consequently, the e -value gradually declines, providing further evidence to reject H_{01} . In counterpart, when the prior distribution is informative (small variance) and is centered in values smaller than 1, the estimate of γ tends to decrease (more closely approaches H_{01}), strengthening the evidence that hypothesis H_{01} can be true.

6 Application to real data

Here we illustrate application of the EDW model with data from Efron (1988), a study of 51 patients with head and neck cancer conducted by the Northern California Oncology Group. The survival times were initially recorded in days, but like in Efron (1988), we transformed the data in days into months, considering 1 month = 30.438 days. Therefore, the response variable T is the time in complete months from the start of the study until the death of the patient or censoring. Here $T = 0$ indicates the patient died within a month of the positive diagnosis. The sample is composed of 51 observations, of which 9 (18%) are censored (Table 7).

Table 7 - Survival times in months of patients with head and neck cancer

Survival times in months	
0, 1, 1, 2, 2, 2+, 2, 2, 2, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6+, 7, 7, 7, 8, 8, 9, 9+, 9, 10+, 13, 13, 13, 14, 17, 17+, 19, 19, 36, 36+, 37, 40+, 44+, 46+, 46	

Note: “+” indicates a censored observation.

Source: data from Efron (1988)

Table 8 presents the point and interval estimates of the parameters of the EDW model for the data in Table 7 according to the classic and bayesian approaches. In the bayesian analysis, we adopted diffuse priors for the parameters, i.e., $q \sim Beta(1, 1)$, $\gamma \sim Gamma(0.001, 0.001)$ and $\beta \sim Gamma(0.001, 0.001)$. The bayesian estimates were obtained through the empirical posterior distribution obtained via MCMC, considering a chain size of 10,000 with burn-in of 1,000.

Table 8 - Estimates of the parameters of the EDW distribution for the data in Table 7

Parameters	Frequentist		Bayesian	
	MLE	CI 95%	Estimate	HPD 95%
q	0.1647	(0.1375; 0.1942)	0.1935	(0.1579; 0.2452)
β	0.2893	(0.2519; 0.3321)	0.3485	(0.2888; 0.3542)
γ	18.813	(14.229; 24.873)	17.425	(13.142; 19.997)

The point estimates (posterior means) of the parameters according to the bayesian approach are near the MLE. However, the interval estimates obtained by the bayesian approach (HPD interval) for the parameters β and γ have smaller amplitude than the classic confidence intervals.

We calculated the estimates of the survival function according to the EDW, DW, exponentiated geometric (EG) and geometric (G) distributions. These functions were estimated by two approaches, frequentist and bayesian. Figure 3 shows the bayesian fit of these distributions to the data in Table 7. The results according to the frequentist approach were similar.

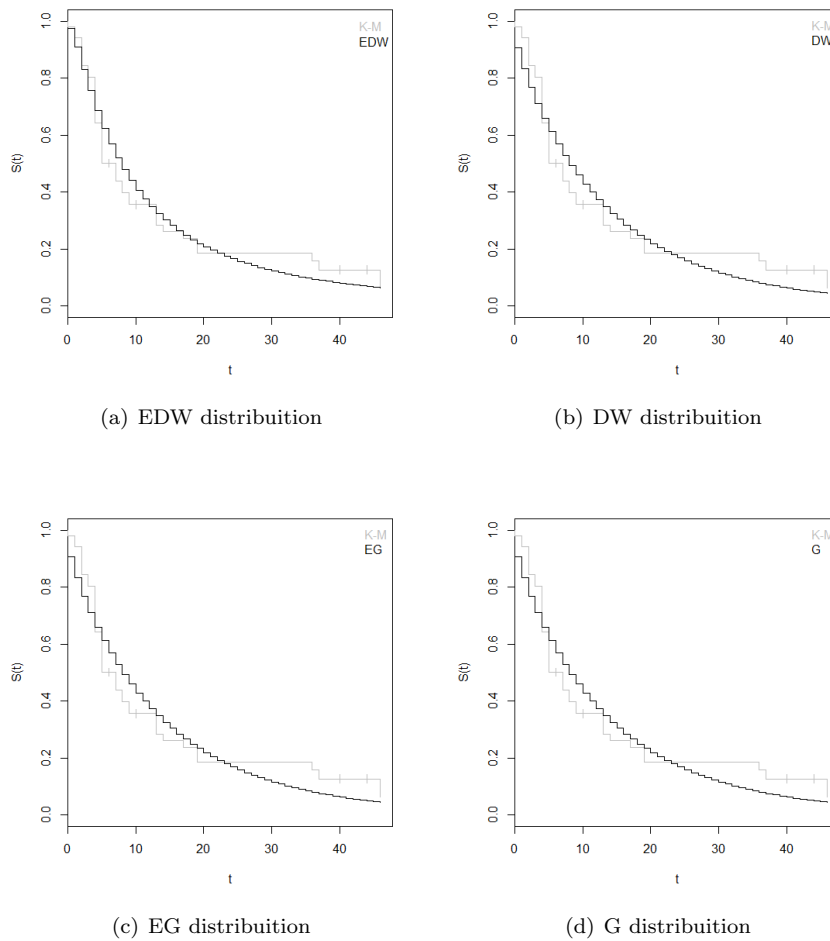


Figure 3 - Bayesian fit of the EDW, DW, EG and G distributions for the survival times (in months) of patients with head and neck cancer.

It can be seen that the EDW distribution produces estimates nearest to the Kaplan-Meier (K-M) estimates, indicating this model best fits the dataset. The estimates obtained by the DW, EG and G distributions present larger deviations in relation to the K-M estimate, suggesting they do not adequately fit this dataset (Figure 3).

To check whether a simpler distribution (with fewer parameters) than the EDW can be used, we tested the following hypotheses:

- H_{01} : $\gamma = 1$: The DW distribution is adequate;
- H_{02} : $\beta = 1$: The EG distribution is adequate; and
- H_{03} : $\gamma = 1$ and $\beta = 1$: The G distribution is adequate.

We tested these hypotheses by the likelihood ratio test (LRT), considering 5% significance, and by the FBST, considering a cutoff point of 5% for the e -value. The results are reported in Table 9.

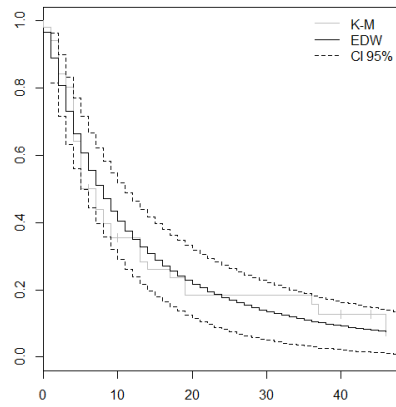
Table 9 - Hypothesis tests of the parameters of the EDW distribution

Hypothesis	LRT		FBST
	$\chi^2(\text{df})$	p -value	e -value
$H_{01} : \gamma = 1$	10.395 (1)	0.0013	0.0208
$H_{02} : \beta = 1$	10.997(1)	0.0009	0.0178
$H_{03} : \gamma = 1$ and $\beta = 1$	10.575(2)	0.0051	0.0178

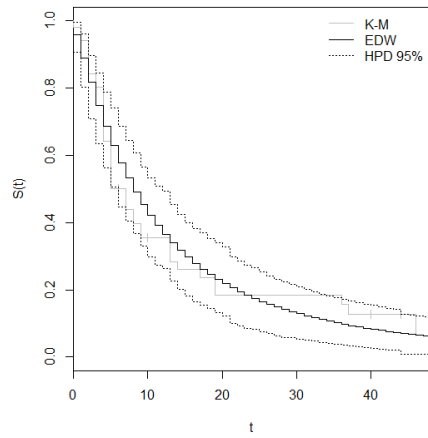
Table 9 shows that hypotheses H_{01} , H_{02} e H_{03} are rejected in the two approaches (frequentist and bayesian). This indicates that the simpler distributions, such as the DW and EG, which accommodate monotone failures, and the G distribution, which accommodates constant failures, are not adequate for this dataset.

Figure 4 presents the K-M estimate of the survival function and the point and interval estimates of the survival function of the EDW distribution for the data in Table 7. The classical confidence interval of the EDW distribution's survival function was obtained by the delta method and the bayesian credible interval was obtained by the predictive distribution (12).

Figure 5 presents the estimate of the failure rate according to the EDW model for the data on patients with head and neck cancer. Note that the failure rate has unimodal behavior, which is in accordance with rejection of the adjustment of the GE, DW and G distributions, which only accommodate monotone failure rates.

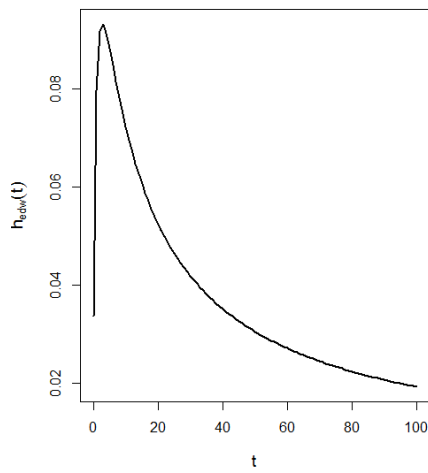


(a) Frequentist

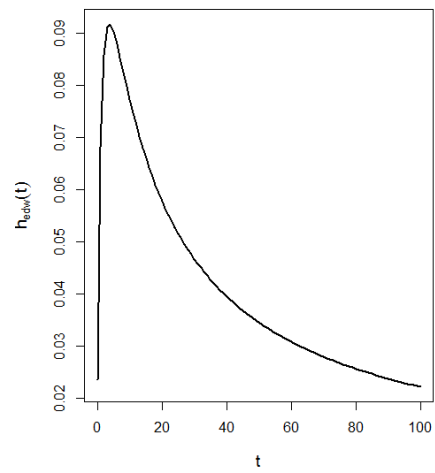


(b) Bayesian

Figure 4 - Point and interval estimates of the survival function of the EDW distribution for the data in Table 7, according to the frequentist and bayesian approaches.



(a) Frequentist



(b) Bayesian

Figure 5 - Estimates of the failure rate function of the EDW distribution for the data in Table 7, according to the frequentist and bayesian approaches.

Conclusions

The EDW distribution is very versatile, since it accommodates constant, monotone (increasing and decreasing) and non-monotone (bathtub and unimodal) failure rates. In this paper, the EDW distribution was formulated from exponentialization (GUPTA; KUNDU, 2001) of the DW distribution of Nakagawa and Osaki (1975). However, this distribution can also be characterized by discretization of the exponentiated continuous Weibull distribution, presented by Mudholkar and Srivastava (1993). Both procedures result in the same probability distribution.

The EDW distribution reduces to the exponentiated geometric, discrete Weibull, geometric, exponentiated discrete Rayleigh and discrete Rayleigh distributions for determined fixed shape parameter values. This fact permitted applying hypothesis testing of the shape parameters of the EDW distribution as a model selection procedure. We considered the likelihood ratio test (LRT) and the full bayesian significance test (FBST). Both these tests were useful to decide on the simplest model to be used to fit the data.

The application of the EDW distribution was illustrated with a real dataset on survival of patients with head and neck cancer. The distribution proved to be a good option in this case, where the risk function presented non-monotone behavior.

Acknowledgments

The first author would like to thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (Brazil) for the financial support. Grant number 1484321. Authors thank reviewers and editors for their comments and suggestions.

CARDIAL, M. R. P.; FACHINI-GOMES, J. B.; NAKANO, E. Y. Distribuição Weibull discreta exponenciada para dados censurados. *Rev. Bras. Biom.*, Lavras, v.38, n.1, p.35-56, 2020.

- *RESUMO: Este trabalho faz um estudo inferencial da distribuição Weibull Discreta Exponenciada (WDE) para dados com presença de censura. Esta generalização da distribuição Weibull Discreta tem a vantagem de ser adequada para modelar taxa de falhas não-monótona como taxas de falha em forma de banheira e unimodal. Inferências sobre a distribuição WDE foram realizadas segundo as abordagens frequentista e bayesiana. Ademais, o Teste de Razão de Verossimilhanças e um Teste de Significância Genuinamente Bayesiano (FBST) foram realizados para testar os parâmetros da distribuição WDE. A metodologia apresentada foi aplicada em dados simulados e ilustrada com um conjunto de dados reais de pacientes diagnosticados com câncer de pescoço e cabeça.*
- *PALAVRAS-CHAVE: Análise de sobrevivência; distribuições discretas; distribuições exponenciadas; teste de significância genuinamente bayesiano.*

References

- BRUNELLO, G. H.; NAKANO, E. Y. Inferência bayesiana no modelo Weibull discreto em dados com presença de censura. *TEMA*, v.16, n.2, p.97-110, 2015.
- CASELLA, G.; Berger, R. L. *Statistical inference*. 2.ed. Kentucky: Duxbury, 2002.
- COLOSIMO, E. A.; GIOLO, S. R. *Análise de sobrevivência aplicada*. 1.ed. São Paulo: Edgard Blücher, 2006.
- De BRAGANÇA PEREIRA, C. A.; STERN, J. Evidence and credibility: full Bayesian significance test for precise hypotheses. *Entropy*, v.1, n.4, p.99-110, 1999.
- EFRON, B. Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve. *Journal of The American Statistical Association*, v.83, n.402, p.414-425, 1988.
- GUPTA, R. D.; KUNDU, D. Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, v.43, n.1, p.117-130, 2001.
- HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, v.57, n.1, p.97-109, 1970.
- MADRUGA, M. R.; ESTEVES, L. G.; WECHSLER, S. On the bayesianity of pereira-stern tests. *Test*, v.10, n.2, p.291-299, 2001.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH M. N., TELLER, A. H., TELLER, E. Equation of state calculations by fast computing machines. *The journal of chemical physics*, v.21, n.6, p.1087-1092, 1953.
- MUDHOLKAR, G. S.; SRIVASTAVA, D. K. Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, v.42, n.2, p.299-302, 1993.
- NAKAGAWA, T.; OSAKI, S. The discrete Weibull distribution. *IEEE Transactions on Reliability*, v.24, n.5, p.300-301, 1975.
- NEKOUKHOV, V.; BIDRAM, H. The exponentiated discrete Weibull distribution. *SORT*, v.39, n.1, p.127-146, 2015.

R CORE TEAM (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Available from: <https://www.R-project.org/>.

RINNE, H. *The Weibull distribution: a handbook*. 1.ed. New York: Chapman and Hall/CRC, 2008.

VILA, R.; NAKANO, E. Y.; SAULO, H. Theoretical results on the discrete Weibull distribution of Nakagawa and Osaki. *Statistics*, v.53, n.2, p.339-363, 2019.

Received on 10.04.2019.

Approved after revised on 09.08.2019.

Appendix

A Likelihood Equations of the EDW

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta})}{\partial q} &= \sum_{i=1}^n \left\{ \delta_i \left\{ \frac{\left[\gamma \left(1 - q^{(t_i+1)^\beta} \right)^{\gamma-1} (t_i+1)^\beta q^{(t_i+1)^\beta - 1} \right]}{\left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma} \right\} \right\} + \\ &\sum_{i=1}^n \left\{ \delta_i \left\{ \frac{- \left[\gamma \left(1 - q^{t_i^\beta} \right)^{\gamma-1} t_i^\beta q^{t_i^\beta - 1} \right]}{\left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma} \right\} \right\} + \\ &\sum_{i=1}^n \left\{ (1 - \delta_i) \left\{ \frac{\left[-\gamma \left(1 - q^{(t_i+1)^\beta} \right)^{\gamma-1} (t_i+1)^\beta q^{(t_i+1)^\beta - 1} \right]}{1 - \left[1 - q^{(t_i+1)^\beta} \right]^\gamma} \right\} \right\}; \end{aligned}$$

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta})}{\partial \beta} &= \sum_{i=1}^n \left\{ \delta_i \left\{ \frac{\left[\gamma \left(1 - q^{(t_i+1)^\beta} \right)^{\gamma-1} (t_i+1)^\beta \log(t_i+1) q^{(t_i+1)^\beta} \log(q) \right]}{\left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma} \right\} \right\} - \\ &\sum_{i=1}^n \left\{ \delta_i \left\{ \frac{\left[\gamma \left(1 - q^{t_i^\beta} \right)^{\gamma-1} t_i^\beta \log(t_i) q^{t_i^\beta} \log(q) \right]}{\left[1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[1 - q^{t_i^\beta} \right]^\gamma} \right\} \right\} + \sum_{i=1}^n \left\{ (1 - \delta_i) \right. \\ &\left. \left\{ \frac{\left[-\gamma \left(1 - q^{(t_i+1)^\beta} \right)^{\gamma-1} (t_i+1)^\beta \log(t_i+1) q^{(t_i+1)^\beta} \log(q) \right]}{1 - \left[1 - q^{(t_i+1)^\beta} \right]^\gamma} \right\} \right\} \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta})}{\partial \gamma} &= \sum_{i=1}^n \left\{ \delta_i \left[\frac{\left(1 - q^{(t_i+1)^\beta} \right)^\gamma \log \left(1 - q^{(t_i+1)^\beta} \right) \left(1 - q^{t_i^\beta} \right)^\gamma \log \left(1 - q^{t_i^\beta} \right)}{\left(1 - q^{(t_i+1)^\beta} \right)^\gamma - \left(1 - q^{t_i^\beta} \right)^\gamma} \right] \right\} + \\ &\sum_{i=1}^n \left\{ (1 - \delta_i) \left[\frac{- \left(1 - q^{(t_i+1)^\beta} \right)^\gamma \log \left(1 - q^{(t_i+1)^\beta} \right)}{1 - \left(1 - q^{(t_i+1)^\beta} \right)^\gamma} \right] \right\} \end{aligned}$$