Noname manuscript No.
(will be inserted by the editor)

On methods of estimation for the type II discrete Weibull distribution

Alessandro Barbiero

Received: date / Accepted: date

Abstract In this paper, we describe and analyze several methods of estimation for the type II discrete Weibull distribution, outlining their applicability and properties, assessing and comparing their performance via intensive Monte Carlo simulation experiments. We consider the standard maximum likelihood method, a method of proportion, and two variants of the least-squares method. The type II discrete Weibull distribution can be used in reliability engineering for modeling count data or discrete lifetimes and its use is theoretically motivated by its capability of modeling either bounded or unbounded support, and either increasing or decreasing failure rate. Statistical analyses of real datasets are presented to show the capability of the distribution in fitting reliability data and illustrate the application of the proposed inferential techniques.

Keywords Discrete lifetimes · Goodness-of-fit methods · Least-squares estimation · Maximum likelihood estimation · Monte Carlo simulation

Mathematics Subject Classification (2010) 62F10 · 62G05

1 Introduction

Almost all reliability studies assume that time is continuous and continuous probability distributions such as exponential, gamma, Weibull, normal and log-normal are commonly used to model the lifetime of a component or a structure. These distributions and the methods for estimating their parameters are well-known. In many practical situations, however, lifetime is not measured with the calendar time: for example, when a device works in cycles or on demands

A. Barbiero
Department of Economics, Management and Quantitative Methods
Università degli Studi di Milano
Tel.: +39-0250321533
Fax: +39-0250321505
E-mail: alessandro.barbiero@unimi.it

and the number of cycles or demands prior to failure is observed; or when the regular operation of a device is monitored once per period, and the number of time periods successfully completed is observed. Moreover, reliability data are often grouped into classes or truncated according to some censoring criterion. In all these situations, lifetime is modeled as a discrete random variable. Indeed, not so much work has been done on discrete stochastic models in reliability so far. Generally, most reliability concepts for continuous lifetimes have been adapted to the discrete context. Bracquemond and Gaudoin (2003) provided an exhaustive survey on discrete lifetime distributions; Bracquemond, Gaudoin, Roy, and Xie (2001); Bracquemond and Gaudoin (2002); Xie, Gaudoin, and Bracquemond (2002) discussed adjustments of some reliability concepts for continuous models to the discrete case.

Often, discrete distributions can be defined as counterparts of continuous distributions, see for example Khalique (1989); Noughabi, Roknabadi, and Borzadaran (2013). Among the continuous distributions used in the engineering field, the Weibull model is surely the most popular. As a discrete alternative, three main distributions have been introduced. The first one was presented in Nakagawa and Osaki (1975) and is referred to as type I discrete Weibull. The second one, which this article takes into consideration, was proposed and studied in Stein and Dattero (1984); the third in Padgett and Spurrier (1985). More recently, Noughabi, Borzadaran, and Roknabadi (2011) presented a discrete version of the continuous modified Weibull distribution proposed by Lai, Xie, and Murthy (2003).

The type II discrete Weibull (henceforth, if there is no ambiguity, simply discrete Weibull) random variable (rv) mimics the hazard rate function of its continuous analogue. The hazard rate function for a continuous rv is defined as

$$r(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{S(x)} \tag{1}$$

where $f(x)$ and $F(x)$ are the probability density function and cumulative distribution function, respectively; and $S(x) = P(X \geq x)$ is the survival function. For a discrete rv, the hazard rate function is usually defined as:

$$r(x) = r_x = \frac{P(X = x)}{S(x)}. \tag{2}$$

Note that for the discrete case $0 \leq r(x) \leq 1$. The discrete Weibull rv is built by imposing that the hazard rate function (2) is equal to its analogue in the continuous case:

$$r(x) = \begin{cases} cx^{\beta-1} & x = 1, 2, \ldots, m \\ 0 & x > m \end{cases} \tag{3}$$

with $0 < c < 1$, where m is defined as $\lfloor c^{-1/(\beta-1)} \rfloor$ if $\beta > 1$; if $0 \leq \beta \leq 1$, $m = +\infty$. The upper bound m of the support of the discrete Weibull rv is needed in order to assure that the hazard rate function in (3) always takes values smaller than 1 (restriction that is not needed in the continuous case). The chance of a finite support for a distribution intended to model count data may appear unusual and represent a theoretical or practical shortcoming; however, some authors indeed claimed that discrete distributions describing the lifetime of a system, device, etc, should be naturally defined on a finite support (see, for example, Jiang 2013).

The survival function S can be then expressed in terms of the hazard rates (3):

$$S(x) = (1 - r_1)(1 - r_2) \cdots (1 - r_{x-1}) = \prod_{j=1}^{x-1}(1 - r_j), \quad x = 2, \ldots, m \tag{4}$$

whereas obviously $S(1) = 1$. The corresponding probability mass function (pmf) has a complex expression and can be written for $\beta > 1$ as

$$P(X = x) = \begin{cases} c & x = 1 \\ S(x) - S(x+1) & x = 2, \ldots, m-1 \\ S(m) & x = m \end{cases} \tag{5}$$

whereas, if $0 \leq \beta \leq 1$,

$$P(X = x) = \begin{cases} c & x = 1 \\ S(x) - S(x+1) & x = 2, 3, \ldots \end{cases} \tag{6}$$

with $S(\cdot)$ given by (4). Note that if $\beta = 1$, the discrete Weibull distribution reduces to the geometric distribution with parameter c, with pmf $P(X = x) = c(1-c)^{x-1}$ for $x = 1, 2, \ldots$, and constant hazard rate function $r(x) = c$.

To the best of our knowledge, inferential techniques have not been directly studied and assessed for this discrete Weibull model. In the next Section, we present and discuss several estimation methods for the parameters of the discrete Weibull model, trying to outline possible issues and theoretical properties. In Section 3, we discuss a Monte Carlo simulation study aimed at comparing the performance of the methods of Section 2 under a number of artificial scenarios. Section 4 presents two applications of the discrete Weibull model to real datasets taken from the literature and applies the inferential techniques discussed in Sections 2 and 3. Section 5 concludes the paper with some final remarks and comments on future research directions.

2 Estimation

In this section, we consider three methods (with a further variant for the latter method) for the point estimation of the parameters c and β of the type II discrete Weibull distribution. Parameter estimation is provided under classical frequentist framework and it is thus based on a *i.i.d.* sample $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ of size n. We will assume that both parameters are unknown. The first method is the conventional maximum likelihood method; the second method (method of proportion) has been proposed for similar discrete models, characterized by at least two parameters, in the context of stochastic reliability; the third method is the least-squares method, which here exploits the specific expression of the hazard rate function for the model at study.

2.1 Maximum likelihood method

Point estimates of the parameters c and β can be derived by maximizing the likelihood function

$$L(c, \beta; x_1, \ldots, x_n) = \prod_{i=1}^{n} P(c, \beta; x_i) \tag{7}$$

or, equivalently, the log-likelihood function

$$\ell(c, \beta; x_1, \ldots, x_n) = \log L(c, \beta; x_1, \ldots, x_n) = \sum_{i=1}^{n} \log P(c, \beta; x_i) \tag{8}$$

Generally, it is not possible to obtain an analytical solution for the maximum likelihood estimates (MLEs) $(\hat{c}_{ML}, \hat{\beta}_{ML})$, admitted they exist: they can be only numerically derived by directly maximizing (7) or (8) (see Khan, Khalique, and Abouammoh (1989); Kulasekera (1993) for analogous issues for the type I discrete Weibull distribution).

For the model at study, particular care should be paid since once a sample $(x_1, x_2, \ldots, x_n)$ has been drawn, the parameter space for the vector (c, β) is subject to the constraint $x_{\max} = \max\{x_1, \ldots, x_n\} \leq \lfloor c^{-1/(\beta-1)} \rfloor$, that can be rewritten as $\beta \leq 1 - \log c / \log x_{\max}$.

In fact, given a sample $(x_1, \ldots, x_n)$ not each pair (c, β) belonging to the natural parameter space $(0, 1) \times [0, \infty)$ is feasible: (c, β) must ensure that the corresponding value of the upper bound m is greater than or equal to each x_i, i.e., greater than or equal to $x_{\max}$ (see Figure 1). As seen before, since a value of β smaller than or equal to 1 leads to a support equal to $\mathbb{N}^+$, the subset $(0, 1) \times [0, 1]$ continues to belong to the "feasible parameter region" in any case. As a function of c and β, given $(x_1, \ldots, x_n)$, the log-likelihood function presents a discontinuity on the points of the parameter space $(0, 1) \times [0, +\infty)$ belonging to the curve of equation $\beta = 1 - \log c / \log(x_{\max} + 1)$. If we suppose that $m > x_{\max}$, the likelihood function can be written as

$$L(c, \beta, x_1, \ldots, x_n) = c^n \prod_{k=2}^{x_{\max}} k^{(\beta-1)\sum_{i=1}^{n} 1_{x_i=k}} [1 - c(k-1)^{\beta-1}]^{\sum_{i=1}^{n} 1_{x_i \geq k}} \tag{9}$$

and then the log-likelihood becomes

$$\ell(c, \beta; x_1, \ldots, x_n) = n \log c + (\beta - 1) \sum_{k=2}^{x_{\max}} \sum_{i=1}^{n} 1_{x_i=k} \log k + \sum_{k=2}^{x_{\max}} \sum_{i=1}^{n} 1_{x_i \geq k} \log[1 - c(k-1)^{\beta-1}] \tag{10}$$

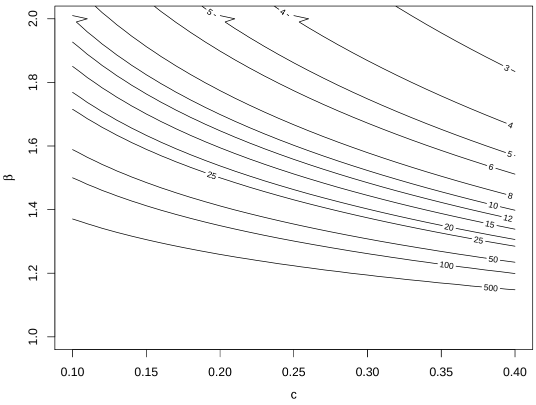


Fig. 1: Contour plot of $c^{-1/(\beta-1)}$. Once a sample is drawn and the maximum value $x_{\max} = \max\{x_i; i = 1, \ldots, n\}$ is observed, the parameter space of (c, β), initially equal to $(0, 1) \times \mathbb{R}^+$, is restricted to the area below the contour line of level $K = x_{\max}$.

Letting $n_k = \sum_{i=1}^n 1_{x_i=k}$ and $N'_k = \sum_{i=1}^n 1_{x_i \geq k}$, we get

$$\ell(c, \beta; x_1, \ldots, x_n) = n \log c + (\beta - 1) \sum_{k=2}^{x_{\max}} n_k \log k + \sum_{k=2}^{x_{\max}} N'_k \log[1 - c(k-1)^{\beta-1}]. \tag{11}$$

On the contrary, if we suppose that $m = x_{\max}$ (i.e., the sample comprises at least one value equal to m) the log-likelihood becomes:

$$\ell(c, \beta; x_1, \ldots, x_n) = (n - n_{x_{\max}}) \log c + (\beta - 1) \sum_{k=2}^{x_{\max}-1} n_k \log k + \sum_{k=2}^{x_{\max}} N'_k \log[1 - c(k-1)^{\beta-1}] \tag{12}$$

Note that since the support of the distribution depends upon the value of the parameters c and β, not much can be claimed about the asymptotic distribution of the MLEs (see Sen, Singer, and Pedroso De Lima 2010; LeCam 1970). If regularity conditions hold, large sample approximate $(1 - \alpha)$ confidence intervals can be separately built for c and β as

$$\hat{\theta}_h \mp z_{1-\alpha/2} \sqrt{\hat{I}_{hh}^{-1}} \tag{13}$$

with $\theta_h = c$, $\hat{\theta}_h = \hat{c}_{ML}$ if $h = 1$; $\theta_h = \beta$, $\hat{\theta}_h = \hat{\beta}_{ML}$ if $h = 2$, and $\hat{I}_{hh}$ the observed Fisher Information matrix for the sample at hand,

$$\hat{I}_{hh} = -\frac{1}{n} \sum_{i=1}^n \left. \frac{\partial^2 \log P(X = x_i; \theta_1, \theta_2)}{\partial \theta_h^2} \right|_{\hat{c}_{ML}, \hat{\beta}_{ML}}. \tag{14}$$

We will now consider two cases where the maximum likelihood method poses some problems, due to the dependence of the parameter space on the observed sample values.

Example 1: non-existence of MLE, existence of an upper bound of ℓ Let us suppose we want to derive the MLEs of c and β based on the *i.i.d.* sample $(x_1, x_2, x_3) = (1, 2, 3)$, or one of its 6 possible permutations, drawn from the discrete Weibull model at study. The log-likelihood function can assume one of the two following expressions:

$$\ell_1(c, \beta; 1, 2, 3) = 3 \log c + (\beta - 1)(\log 2 + \log 3) + 2 \log(1 - c) + \log(1 - c2^{\beta-1}) \quad (15)$$

if $m > 3$, or

$$\ell_2(c, \beta; 1, 2, 3) = 2 \log c + (\beta - 1) \log 2 + 2 \log(1 - c) + \log(1 - c2^{\beta-1}) \quad (16)$$

if $m = 3$.

It can be shown that the function ℓ admits an upper limit ($\ell_{\sup} = -3.312405$) for (c, β) tending to $(\hat{c}_{ML*}, \hat{\beta}_{ML*}) = (0.3058, 1.8546)$ from the region corresponding to ℓ_2, i.e., $\ell_{\sup} = \lim_{c \to \hat{c}_{ML*}, \beta \to \hat{\beta}_{ML*}} \ell_2$, with c and β such that $m(c, \beta) = 3$. The point $(\hat{c}_{ML*}, \hat{\beta}_{ML*})$ lies on the dashed curve of Figure 2a, which is the set of discontinuity points for ℓ described before. $(\hat{c}_{ML*}, \hat{\beta}_{ML*})$ can be considered an extension of the maximum likelihood estimates, since they satisfy the condition

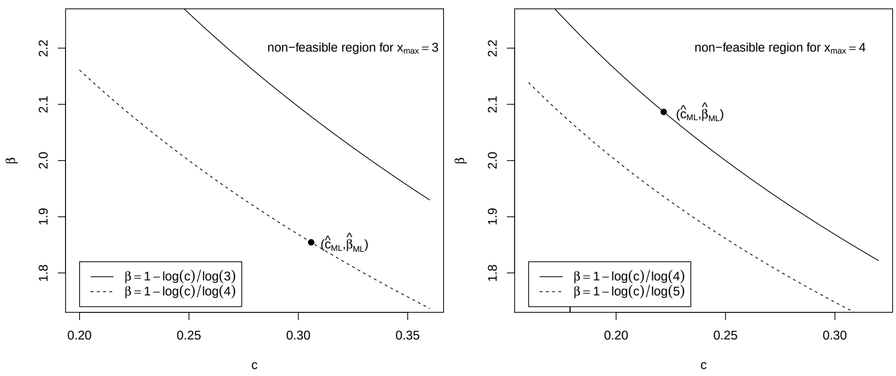
$$\lim_{(c,\beta) \to (\hat{c}_{ML*}, \hat{\beta}_{ML*}) | m(c,\beta) = x_{\max}} L(c, \beta; x_1, \ldots, x_n) = \sup_{(c,\beta)} L(c, \beta; x_1, \ldots, x_n)$$

Large-sample confidence intervals (13) for c and β cannot be built because the (observed) Fisher information matrix (14) cannot be computed in $(\hat{c}_{ML*}, \hat{\beta}_{ML*})$ (the log-likelihood is not continuous here).

Example 2: MLE lying on the boundary Let us suppose we want to derive the MLEs of c and β based on the *i.i.d.* sample of size $n = 10$ $\boldsymbol{x} = (1, 1, 2, 2, 2, 2, 3, 3, 3, 4)$ (or any of its possible permutations) drawn from the discrete Weibull model at study. The MLEs are $\hat{c}_{ML} = 0.2216039$ and $\hat{\beta}_{ML} = 2.0869723$ and the corresponding value of the log-likelihood is -12.82967. The corresponding value of $m = \lfloor \hat{c}_{ML}^{-1/(\hat{\beta}_{ML}-1)} \rfloor$ is equal to $4 = x_{\max}$. Thus, the point whose coordinates are the MLEs lies on the boundary between the feasible and the non-feasible region, see Figure (2b). For this reason, as in the previous example, large-sample confidence intervals (13) for c and β cannot be built because the (observed) Fisher information matrix (14) cannot be computed in $(\hat{c}_{ML}, \hat{\beta}_{ML})$.

2.2 Method of proportion

Through this method (see, for example, Khan, Khalique, and Abouammoh 1989), the unknown parameters c and β of the discrete Weibull model are estimated by equating the first two probabilities $P(X = 1)$ and $P(X = 2)$ to the corresponding sample rates of 1s and 2s. Suppose that the sample contains at least a 1 and at least a 2. Letting $\hat{p}_1 = \sum_{i=1}^{n} \mathbb{1}_{x_i=1}/n$ and $\hat{p}_2 =$



(a) Discontinuity of the loglikelihood function (b) Boundary of the parameter space

Fig. 2: MLEs and feasible/non-feasible parameter regions for two samples from a discrete Weibull rv: $\boldsymbol{x} = (1, 2, 3)$ (left) and $\boldsymbol{x} = (1, 1, 2, 2, 2, 2, 3, 3, 3, 4)$ (right)

$\sum_{i=1}^{n} \mathbb{1}_{x_i=2}/n$ be the proportions of 1s and 2s in the sample, respectively, we can derive the expression of the estimates for c and β by the method of proportion as

$$\hat{c}_P = \hat{p}_1 \tag{17}$$

$$\hat{\beta}_P = 1 + \log\left[\frac{\hat{p}_2}{\hat{p}_1(1-\hat{p}_1)}\right] / \log(2) \tag{18}$$

being $P(X = 2) = c2^{\beta-1}(1 - c)$. Though being very straightforward and providing an analytical expression for the estimates, the method of proportion exhibits however an intrinsic weakness since it automatically discards plenty of information contained in the sample, making use only of the sample rates of 1s and 2s. Moreover, it can yield 'inconsistent' estimates for β, i.e., values smaller than zero, whereas its natural parameter space is $\mathbb{R}_0^+$. Inconsistent estimates thus occur if and only if

$$1 + \log\left[\frac{\hat{p}_2}{\hat{p}_1(1-\hat{p}_1)}\right] / \log(2) < 0 \Leftrightarrow \hat{p}_2 < \frac{\hat{p}_1(1-\hat{p}_1)}{2}.$$

2.3 Least-squares method

This method exploits the expression of the failure rate function for the discrete Weibull distribution given by (3). Making the logarithm of the two sides in the expression (3) of the hazard rate function we get

$$\log r = \log c + (\beta - 1)\log x \tag{19}$$

that can be rewritten as a linear model

$$z = a + bw \tag{20}$$

letting $z = \log r$, $w = \log x$, $a = \log c$, $b = \beta - 1$. Now, for each sample observation x_i, $i = 1 \ldots, n$, we can estimate the corresponding hazard rate value $r_i = r(x_i)$ as the ratio between the sample frequency of values equal to x_i and the sample frequency of values equal to or greater than x_i, $\hat{r}_i = \hat{r}(x_i) = \sum_{j=1}^{n} \mathbb{1}_{x_j = x_i} / \sum_{j=1}^{n} \mathbb{1}_{x_j \geq x_i}$, which is a number between 0 and 1 (equal to zero for all the integer values x not present in the sample; equal to 1 for $x_{\max} = \max\{x_i, i = 1, \ldots, n\}$, see also Bracquemond and Gaudoin (2002)). The estimates of the parameters a and b in the linear model (19), with z replaced by $\hat{z} = \log \hat{r}$, can be derived by the least-squares method and are denoted as $\hat{a}$ and $\hat{b}$,

$$\hat{b} = \text{cov}(w, \hat{z})/\text{var}(w) \quad \hat{a} = \bar{\hat{z}} - \hat{b}\bar{w} \tag{21}$$

and then the parameter estimates $\hat{c}_{LS}$ and $\hat{\beta}_{LS}$ of the discrete Weibull rv can be obtained as

$$\hat{c}_{LS} = \exp(\hat{a}) \quad \hat{\beta}_{LS} = \hat{b} + 1 \tag{22}$$

A preliminary simulation study has shown that this method often produces highly biased estimators; a possible remedy is to compute the estimates of the linear model parameters a and b ignoring the sample values x_i equal to $x_{\max}$, to which corresponds a value of the sample hazard rate equal to 1. This modification seems to empirically diminish the absolute value of the bias of both estimates $\hat{c}_{LS}$ and $\hat{\beta}_{LS}$. We will denote these new estimators as $\hat{c}_{MLS}$ and $\hat{\beta}_{MLS}$, MLS standing for “Modified Least-Squares”.

Example The following sample is drawn from a discrete Weibull rv with $c = 0.2$ and $\beta = 0.5$: $\boldsymbol{x} = (1, 1, 1, 1, 2, 2, 2, 2, 5, 6, 7, 8, 9, 12, 15, 16, 30)$. The least-squares method provides the estimates $\hat{c}_{LS} = 0.1594$, $\hat{\beta}_{LS} = 1.2187$; the estimate of β is far greater than the true value 0.5. This can explained, apart from sampling variability considerations, as follows: when β is between 0 and 1, the support of the discrete Weibull rv is the whole $\mathbb{N}^+$, and the closer is β to 0, the larger is the probability of large integers; for the given choice of parameters, there is a probability about 20% of values greater than 20. For such large values, the theoretical hazard rate function r is however very small (see Eq.(3)), but the corresponding value of the sample hazard rate function $\hat{r}$ may be moderately high (for the sample at hand, $\hat{r}(16) = 0.5$); thus, the largest values in a sample have a sort of “leverage effect” in the linear model (20), leading to a positive value for $\hat{b}$ and thus a value larger than 1 for $\hat{\beta}$, as for the sample here considered (over-estimate). This bad behaviour seems liable to be mitigated only by increasing the sample size. More accurate details will be provided in Section 3, devoted to a simulation study.

2.4 An illustrative example

Let consider the sample of size $n = 20$ drawn from the discrete Weibull rv with parameters $c = 0.1$ and $\beta = 1.5$, whose frequency distribution is given by

x_i	1	2	3	4	5	6	7	9
O_i	1	2	2	8	2	3	1	1

where O_i denotes the observed frequency of value x_i. The method of proportion can be applied to this sample and yields the estimates $\hat{c}_P = 1/20 = 0.05$ and $\hat{\beta}_P = 1 + \log\left[\frac{1/10}{1/20(19/20)}\right] / \log(2) = 2.074$. As for the least-squares method, the values of the sample hazard rates for each distinct observed value and the transformed values w_i and z_i are reported below.

x_i	1	2	3	4	5	6	7	9
$\hat{r}_i$	0.05	0.1053	0.1176	0.5333	0.2857	0.6	0.5	1
w_i	0	0.6931	1.0986	1.3863	1.6094	1.7918	1.9459	2.1972
$\hat{z}_i$	-2.9957	-2.2509	-2.1405	-0.6287	-1.2528	-0.5108	-0.6931	0

It is then easy to calculate the intermediate estimates

$$\hat{b} = \text{cov}(w, \hat{z})/\text{var}(w) = 1.4322 \quad \hat{a} = \bar{\hat{z}} - \hat{b}\bar{w} = -3.0398 \tag{23}$$

and finally $\hat{c}_{LS} = 0.0478$ and $\hat{\beta}_{LS} = 2.4322$. The modification of the least-squares method provides similar results: $\hat{c}_{MLS} = 0.0466$ and $\hat{\beta}_{MLS} = 2.4562$. The maximum likelihood method yields the estimates $\hat{c}_{ML} = 0.0497$ and $\hat{\beta}_{ML} = 2.3034$, but here raises some difficulty since it is not possible to build the large sample confidence intervals for c and β, as described in subsection 2.1. Note that all the point estimates for c are smaller than the true value of c, and all the point estimates for β are greater than the true value of β. However, the four methods provide very similar estimates for each of the two parameters.

3 Simulation study

Here, we describe a Monte Carlo simulation study we performed to assess the properties of the estimators presented in the previous section under different experimental conditions. A similar comparative study was carried out in Barbiero (2013) for comparing estimators for the type III discrete Weibull distribution.

3.1 Simulation design

The Monte Carlo simulation study comprises a number of artificial settings obtained by combining different values of the two distribution parameters c and β and sample size n. In greater detail, we considered the following values for c: $0.05, 0.1, 0.2, 0.3$; for β: $0.5, 0.75, 1, 1.25, 1.5$; and for the sample size n: $20, 50, 100$. With regard to c, values larger than 0.3 were not taken into consideration since they would imply a 1-probability too much high for practical cases. As for β, we considered values either smaller than or equal to 1, which provide a decreasing or constant hazard rate function and an unbounded support ($\beta = 1$ is a particular case, since for this value the discrete Weibull distribution degenerates into a geometric); or greater than 1, which provide an increasing hazard rate function and a bounded support. With regard to the sample size, we did not consider values smaller than 20, as they would have produced problems with some estimators and made the comparison among different estimators more difficult and the presentation of overall results less readable.

$N = 5,000$ samples of size n ($n = 20, 50, 100$) were simulated for each of the $5 \times 4 = 20$ combinations of parameter values. For each sample, the parameter estimates were computed according to each of the methods presented in the previous section (abbreviated as ML, P, LS, MLS) and asymptotic 95% confidence intervals based on large-sample theory were built for both parameters. The MC bias, i.e. the value $bias = \frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta) = \bar{\hat{\theta}} - \theta$, and the standard error, i.e., the value $SE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \bar{\hat{\theta}})^2}$ –with $\theta = c, \beta$– of each point estimator $\hat{\theta}_i$ were computed over all the $5,000$ random samples, as well as the actual coverage rate and the average length for the above mentioned confidence intervals.

The simulation study has been implemented in the programming environment `R` (R Development Core Team 2016) and the routines related to the discrete Weibull model (calculating the probability mass function, the cumulative distribution function, the hazard rate function, the quantile function, and implementing pseudo-random generation and sample estimation) will be made freely available through the package `DiscreteWeibull` (Barbiero 2015), which already implements types I and III discrete Weibull distributions.

3.2 Simulation results

The results of the simulation study are reported in Tables 1–3 (estimators of the first parameter c) and 4–6 (estimators of the second parameter β).

Please note that, as revealed in advance, the method of proportion for the point estimation of c and β cannot be applied for each possible sample, since it requires the presence in the sample of at least a 1 and a 2. Similarly, the large-sample confidence intervals for the two parameters based on the maximum likelihood estimates and Fisher's information matrix are not always derivable (recall the examples of Section 2). The rate of "non-feasible" samples for the two inferential techniques clearly depend upon the combination of parameter values and the sample size. The summary indexes (bias, standard error, coverage rate and average length) are clearly computed and averaged only on the samples where they can be applied.

Looking at the results one can note the following points:

- the method of proportion produces an estimator $\hat{c}_P$ whose bias is non-negligible for small sample size and low values of c; whereas for $n = 50, 100$, the bias is never greater than .004 under each scenario. As to $\hat{\beta}_P$, its bias

is relevant especially for low values of c and β (and $n = 20$); whereas for larger values of c and β, it becomes negligible even for small sample size. The maximum likelihood method produce a nearly unbiased estimator of c (a negative bias appears only for large values of c), independently from the value of β; the bias of $\hat{\beta}_{ML}$ is non-negligible for small sample size and increases in absolute value increasing c (keeping fixed β) or increasing β (keeping fixed c). The least-squares methods yield negatively biased estimators of c (the bias in absolute value decreases increasing β for a fixed c); the modified version significantly reduces the absolute bias, especially for small values of n. It should be noted that for some combinations of c and β the behavior of $\hat{c}_{LS}$ (and $\hat{c}_{MLS}$) is quite odd, since its bias in absolute value tends to increase with n. For example, for $c = 0.1$ and $\beta = 0.5$, the bias — moving n from 20 to 50 and 100 — passes from -0.028 (-0.019) to -0.047 (-0.042) and -0.050 (-0.046). The bias of the estimator $\hat{\beta}_{LS}$ is dramatically high, especially for small values β, and makes it practically unusable; however, again, the modified version of the method seem to yield some benefit, lowering the magnitude of the bias. Such an unexpected behavior can be justified recalling the arguments of section 2.3.
- as to the standard error, among the estimators of c, the one yielded by the method of proportion is the most variable, under each scenario and for each sample size examined. The gap does not seem to reduce moving to the largest sample size. The standard error of all the estimators of c increases in absolute value with c and is almost independent from the value of β. As to the estimators of β, again, the method of proportion shows the largest standard error under each possible scenario and for each sample size examined. The least-squares method and its modification overcome the maximum likelihood method for most of the cases, but this advantage is paid in terms of a larger absolute bias and vanishes for $n = 100$ and for small values of c and β. For a given sample size n, the standard error of the ML and LS (MLS) estimators increases with β keeping c fixed, and slightly increases with c keeping β fixed. The behavior of the standard error of $\hat{\beta}_P$ is quite singular and opposite to that of its competitors: for the largest sample size examined here, it decreases with β keeping c fixed, and it decreases with c keeping β fixed. It can be easily explained if we recall that the method of proportion just exploits the information contained in the sample units equal to 1 or 2, and the fraction of these units increases with both c and β, thus making the estimator more precise.
- the modification to the least-squares method on the one hand seems to reduce the high value of absolute bias that characterizes the original couple of estimators, but, on the other hand, increases the standard error. Overall, the least-squares method and its modification, despite the fact they are easily applicable to every possible sample, both need to be properly adjusted in order to make the bias less significant.
- overall, the maximum likelihood method is the most reliable method. In particular, it is highly reliable in estimating β, whereas the other three

methods often produce a non-negligible bias for it. In some scenarios, however, the performances of the ML and LS (MLS) estimators is not so different from each other. In this sense, figure 3, which displays the Monte Carlo distributions of the estimators for c and β when $c = 0.2$, $\beta = 1.25$ and $n = 50$, is quite suggestive. At first glance, it shows that in this scenario the estimators derived by the method of proportion (especially the estimator for β) are more variable than those derived by the other three methods, whose performance is instead quite similar.

- the confidence interval for c hardly achieves the nominal coverage even when the sample size is high ($n = 100$). As a general trend, its performance gets better for larger values of c, as one could expect, since the distribution of $\hat{c}_{ML}$ resembles normality better when c approaches 0.5. In fact, the real coverage of the confidence interval for c is quite smaller than the nominal level if c itself is small (say, 0.05, 0.1), whereas for larger values (0.2, 0.3), the actual coverage tends to get closer. When $c = 0.05, 0.1$, increasing the value of β reduces the coverage rate; whereas, when $c = 0.2, 0.3$, increasing the value of β increases the coverage rate.
- the confidence interval for β, on the contrary, always has a coverage rate close to the nominal one, even for smaller sample size, but is also little precise. Some variability in the coverage rate is observed for larger values of c, varying the value of β: increasing β there is a tendency to over-coverage, decreasing β there is a tendency to under-coverage. On the contrary, for small values of c, the coverage rate is very close to 95% independently from the value of β.

Despite complexity of computation, MLEs (of c and β) are overall the best estimators in terms of statistical performance; LS, MLS, and P estimators have the advantage of having an analytical expression, but the latter often suffer from large variability (especially $\hat{\beta}_P$), whereas LS and MLS are often significantly biased. Least-squares estimators have however a much closer performance to MLE and may be susceptible to possible enhancement. Results are somewhat not surprising, especially if compared to analogous studies of other discrete models (see, for example, Barbiero 2013; Barbiero 2016).

4 Application

In this section, we apply the inferential techniques described so far to two real datasets taken from the literature.

4.1 Number of accesses between disk failures

In the first dataset, considered in Bebbington and Lai (1998), the observed number of trials between errors in read-write error testing of a certain com-

Table 1: Simulation results for the estimators of parameter c, $n = 20$. Legenda: SE=standard error, P=method of proportion, ML=maximum likelihood method, LS=least-squares method, MLS=modified least-squares method, CI=confidence interval based on ML estimates and Fisher information matrix, rate=actual coverage rate, ave l=average length

c	β	bias($\hat{c}$)				SE($\hat{c}$)				CI(c)	
		P	ML	LS	MLS	P	ML	LS	MLS	rate	ave l
0.05	0.5	0.028	0.000	-0.010	-0.004	0.038	0.026	0.020	0.022	0.877	0.094
0.05	0.75	0.028	0.000	-0.008	-0.001	0.039	0.027	0.022	0.024	0.871	0.101
0.05	1	0.028	0.000	-0.004	0.003	0.039	0.028	0.024	0.027	0.867	0.100
0.05	1.25	0.028	0.000	0.000	0.008	0.039	0.029	0.026	0.030	0.863	0.102
0.05	1.5	0.028	0.000	0.005	0.013	0.038	0.029	0.029	0.032	0.862	0.103
0.1	0.5	0.013	-0.002	-0.028	-0.019	0.058	0.043	0.035	0.038	0.892	0.161
0.1	0.75	0.013	-0.001	-0.021	-0.011	0.059	0.045	0.038	0.042	0.888	0.167
0.1	1	0.014	-0.001	-0.013	-0.003	0.059	0.046	0.041	0.044	0.887	0.171
0.1	1.25	0.014	-0.001	-0.005	0.005	0.059	0.047	0.043	0.047	0.884	0.173
0.1	1.5	0.014	-0.001	0.001	0.011	0.060	0.048	0.045	0.050	0.884	0.175
0.2	0.5	0.000	-0.005	-0.049	-0.034	0.086	0.069	0.062	0.066	0.905	0.261
0.2	0.75	0.002	-0.004	-0.035	-0.020	0.087	0.070	0.064	0.068	0.905	0.267
0.2	1	0.002	-0.004	-0.021	-0.007	0.087	0.072	0.066	0.070	0.906	0.271
0.2	1.25	0.003	-0.005	-0.010	0.002	0.088	0.073	0.068	0.072	0.907	0.275
0.2	1.5	0.003	-0.006	-0.002	0.008	0.087	0.073	0.070	0.074	0.931	0.284
0.3	0.5	0.000	-0.007	-0.055	-0.035	0.101	0.087	0.083	0.087	0.910	0.332
0.3	0.75	0.001	-0.007	-0.037	-0.019	0.101	0.089	0.084	0.088	0.909	0.337
0.3	1	0.001	-0.007	-0.022	-0.006	0.101	0.090	0.085	0.089	0.913	0.341
0.3	1.25	0.002	-0.008	-0.011	0.001	0.102	0.090	0.086	0.090	0.927	0.347
0.3	1.5	0.002	-0.008	-0.003	0.005	0.102	0.089	0.087	0.091	0.948	0.358

Table 2: Simulation results for the estimators of parameter c, $n = 50$

c	β	bias($\hat{c}$)				SE($\hat{c}$)				CI(c)	
		P	ML	LS	MLS	P	ML	LS	MLS	rate	ave l
0.05	0.5	0.004	0.000	-0.027	-0.025	0.028	0.016	0.009	0.010	0.918	0.061
0.05	0.75	0.004	0.000	-0.022	-0.019	0.028	0.017	0.011	0.012	0.917	0.064
0.05	1	0.004	0.000	-0.016	-0.012	0.028	0.017	0.013	0.014	0.917	0.067
0.05	1.25	0.004	0.000	-0.009	-0.005	0.028	0.018	0.015	0.016	0.914	0.068
0.05	1.5	0.004	0.000	-0.004	0.000	0.028	0.018	0.016	0.017	0.916	0.069
0.1	0.5	0.000	-0.001	-0.047	-0.042	0.042	0.027	0.019	0.020	0.923	0.103
0.1	0.75	0.000	-0.001	-0.034	-0.028	0.042	0.028	0.022	0.023	0.923	0.107
0.1	1	0.000	-0.001	-0.021	-0.015	0.042	0.028	0.024	0.025	0.924	0.110
0.1	1.25	0.000	-0.001	-0.012	-0.006	0.042	0.029	0.026	0.027	0.922	0.112
0.1	1.5	0.000	-0.001	-0.005	0.000	0.042	0.029	0.027	0.029	0.921	0.113
0.2	0.5	0.000	-0.003	-0.063	-0.052	0.056	0.043	0.036	0.038	0.930	0.167
0.2	0.75	0.000	-0.003	-0.039	-0.029	0.056	0.044	0.039	0.041	0.930	0.171
0.2	1	0.000	-0.003	-0.022	-0.013	0.056	0.045	0.041	0.043	0.928	0.174
0.2	1.25	0.000	-0.003	-0.011	-0.004	0.056	0.045	0.043	0.044	0.930	0.176
0.2	1.5	0.000	-0.003	-0.004	0.001	0.056	0.046	0.044	0.046	0.940	0.177
0.3	0.5	0.000	-0.004	-0.062	-0.047	0.063	0.054	0.050	0.052	0.937	0.212
0.3	0.75	0.000	-0.004	-0.036	-0.024	0.063	0.055	0.051	0.053	0.935	0.216
0.3	1	0.000	-0.003	-0.019	-0.009	0.063	0.056	0.052	0.054	0.937	0.218
0.3	1.25	0.000	-0.004	-0.009	-0.002	0.063	0.056	0.054	0.055	0.938	0.219
0.3	1.5	0.000	-0.005	-0.003	0.001	0.063	0.056	0.055	0.057	0.953	0.223

Table 3: Simulation results for the estimators of parameter c, $n = 100$

c	β	bias($\hat{c}$)				SE($\hat{c}$)				CI(c)	
		P	ML	LS	MLS	P	ML	LS	MLS	rate	ave l
0.05	0.5	0.001	0.000	-0.032	-0.030	0.021	0.011	0.006	0.006	0.933	0.044
0.05	0.75	0.001	0.000	-0.024	-0.022	0.021	0.012	0.008	0.008	0.929	0.046
0.05	1	0.001	0.000	-0.016	-0.013	0.021	0.012	0.009	0.010	0.928	0.048
0.05	1.25	0.001	0.000	-0.009	-0.006	0.021	0.013	0.011	0.011	0.928	0.049
0.05	1.5	0.001	0.000	-0.004	-0.002	0.021	0.013	0.012	0.012	0.928	0.050
0.1	0.5	0.000	0.000	-0.050	-0.046	0.030	0.019	0.013	0.014	0.935	0.073
0.1	0.75	0.000	0.000	-0.032	-0.028	0.030	0.020	0.016	0.016	0.934	0.077
0.1	1	0.000	0.000	-0.018	-0.014	0.030	0.020	0.018	0.018	0.935	0.079
0.1	1.25	0.000	0.000	-0.009	-0.006	0.030	0.021	0.019	0.020	0.9336	0.080
0.1	1.5	0.000	0.000	-0.004	-0.001	0.030	0.021	0.020	0.021	0.932	0.081
0.2	0.5	0.000	-0.001	-0.059	-0.051	0.040	0.031	0.026	0.027	0.937	0.119
0.2	0.75	0.000	-0.001	-0.033	-0.025	0.040	0.032	0.028	0.029	0.935	0.122
0.2	1	0.000	-0.001	-0.017	-0.011	0.040	0.032	0.030	0.031	0.936	0.124
0.2	1.25	0.000	-0.001	-0.007	-0.003	0.040	0.032	0.031	0.032	0.938	0.125
0.2	1.5	0.000	-0.001	-0.002	0.000	0.040	0.033	0.032	0.033	0.938	0.125
0.3	0.5	0.001	-0.001	-0.055	-0.044	0.047	0.039	0.036	0.037	0.941	0.150
0.3	0.75	0.001	-0.001	-0.029	-0.020	0.047	0.040	0.037	0.038	0.941	0.153
0.3	1	0.001	-0.001	-0.014	-0.007	0.047	0.040	0.039	0.039	0.942	0.155
0.3	1.25	0.001	-0.001	-0.005	-0.001	0.047	0.040	0.039	0.040	0.943	0.155
0.3	1.5	0.001	-0.002	-0.001	0.001	0.047	0.041	0.040	0.041	0.954	0.157

Table 4: Simulation results for the estimators of parameter β, $n = 20$

β	c	bias($\hat{\beta}$)				SE($\hat{\beta}$)				CI(β)	
		P	ML	LS	MLS	P	ML	LS	MLS	rate	ave l
0.5	0.05	0.445	0.044	0.836	0.782	0.791	0.131	0.088	0.088	0.950	0.486
0.5	0.1	0.329	0.051	0.815	0.749	0.961	0.149	0.110	0.114	0.950	0.554
0.5	0.2	0.109	0.066	0.742	0.650	1.033	0.179	0.125	0.137	0.946	0.664
0.5	0.3	-0.030	0.081	0.672	0.556	0.928	0.207	0.130	0.153	0.941	0.759
0.75	0.05	0.243	0.059	0.693	0.619	0.813	0.181	0.127	0.126	0.951	0.668
0.75	0.1	0.189	0.066	0.652	0.568	0.992	0.198	0.147	0.153	0.950	0.733
0.75	0.2	0.039	0.079	0.565	0.463	1.045	0.224	0.161	0.179	0.945	0.835
0.75	0.3	-0.050	0.091	0.494	0.375	0.905	0.249	0.169	0.200	0.940	0.921
1	0.05	0.075	0.074	0.547	0.459	0.836	0.228	0.164	0.165	0.950	0.842
1	0.1	0.075	0.080	0.496	0.403	1.023	0.243	0.181	0.191	0.948	0.898
1	0.2	0.016	0.090	0.410	0.307	1.036	0.265	0.199	0.224	0.945	0.983
1	0.3	-0.035	0.099	0.347	0.234	0.859	0.285	0.208	0.250	0.950	1.049
1.25	0.05	-0.088	0.089	0.407	0.309	0.868	0.274	0.200	0.205	0.950	1.006
1.25	0.1	-0.010	0.093	0.357	0.261	1.045	0.284	0.216	0.232	0.947	1.050
1.25	0.2	0.006	0.101	0.284	0.187	1.027	0.301	0.235	0.268	0.953	1.109
1.25	0.3	-0.030	0.111	0.233	0.133	0.821	0.313	0.249	0.302	0.974	1.144
1.5	0.05	-0.220	0.103	0.276	0.175	0.903	0.317	0.236	0.248	0.949	1.161
1.5	0.1	-0.073	0.107	0.242	0.149	1.052	0.324	0.253	0.278	0.951	1.184
1.5	0.2	0.010	0.114	0.189	0.106	0.994	0.327	0.272	0.315	0.974	1.199
1.5	0.3	-0.001	0.107	0.150	0.071	0.761	0.301	0.286	0.351	0.983	1.225

Table 5: Simulation results for the estimators of parameter β, $n = 50$

β	c	bias($\hat{\beta}$)				SE($\hat{\beta}$)				CI(β)	
		P	ML	LS	MLS	P	ML	LS	MLS	rate	ave l
0.5	0.05	0.200	0.018	0.786	0.748	1.068	0.077	0.074	0.074	0.949	0.294
0.5	0.1	0.024	0.022	0.702	0.653	1.049	0.088	0.079	0.080	0.948	0.335
0.5	0.2	-0.064	0.028	0.571	0.505	0.791	0.106	0.080	0.083	0.945	0.402
0.5	0.3	-0.061	0.034	0.487	0.407	0.629	0.123	0.085	0.091	0.940	0.462
0.75	0.05	0.094	0.025	0.605	0.555	1.095	0.106	0.097	0.097	0.950	0.405
0.75	0.1	0.005	0.028	0.503	0.446	1.038	0.117	0.101	0.102	0.949	0.444
0.75	0.2	-0.047	0.033	0.383	0.316	0.742	0.134	0.105	0.110	0.947	0.506
0.75	0.3	-0.043	0.038	0.316	0.242	0.571	0.148	0.111	0.121	0.944	0.559
1	0.05	0.017	0.031	0.436	0.380	1.108	0.134	0.119	0.119	0.951	0.510
1	0.1	0.002	0.033	0.342	0.284	1.017	0.143	0.122	0.125	0.949	0.545
1	0.2	-0.033	0.038	0.249	0.187	0.685	0.158	0.131	0.139	0.947	0.597
1	0.3	-0.021	0.041	0.197	0.133	0.519	0.169	0.137	0.151	0.943	0.638
1.25	0.05	-0.036	0.037	0.296	0.239	1.110	0.160	0.140	0.142	0.951	0.610
1.25	0.1	0.010	0.039	0.225	0.170	0.985	0.167	0.146	0.152	0.950	0.637
1.25	0.2	-0.011	0.042	0.156	0.103	0.638	0.178	0.155	0.167	0.947	0.674
1.25	0.3	-0.008	0.046	0.118	0.067	0.476	0.187	0.162	0.179	0.954	0.697
1.5	0.05	-0.060	0.042	0.189	0.135	1.109	0.185	0.162	0.167	0.950	0.703
1.5	0.1	0.030	0.045	0.140	0.092	0.945	0.189	0.170	0.178	0.949	0.719
1.5	0.2	0.003	0.049	0.094	0.053	0.598	0.197	0.178	0.191	0.965	0.733
1.5	0.3	0.000	0.053	0.069	0.031	0.440	0.192	0.184	0.206	0.971	0.735

Table 6: Simulation results for the estimators of parameter β, $n = 100$

β	c	bias($\hat{\beta}$)				SE($\hat{\beta}$)				CI(β)	
		P	ML	LS	MLS	P	ML	LS	MLS	rate	ave l
0.5	0.05	-0.014	0.009	0.711	0.683	1.076	0.053	0.058	0.058	0.949	0.205
0.5	0.1	-0.052	0.010	0.586	0.551	0.796	0.060	0.059	0.059	0.949	0.233
0.5	0.2	-0.031	0.013	0.439	0.393	0.526	0.073	0.058	0.060	0.947	0.279
0.5	0.3	-0.028	0.016	0.361	0.307	0.420	0.084	0.062	0.065	0.947	0.320
0.75	0.05	-0.027	0.012	0.497	0.462	1.064	0.073	0.073	0.073	0.949	0.282
0.75	0.1	-0.032	0.013	0.375	0.336	0.749	0.080	0.073	0.073	0.950	0.309
0.75	0.2	-0.020	0.016	0.265	0.221	0.484	0.092	0.077	0.079	0.945	0.352
0.75	0.3	-0.021	0.018	0.212	0.165	0.385	0.101	0.082	0.086	0.945	0.388
1	0.05	-0.022	0.015	0.322	0.285	1.027	0.092	0.088	0.088	0.949	0.355
1	0.1	-0.012	0.016	0.232	0.195	0.701	0.098	0.090	0.091	0.950	0.379
1	0.2	-0.015	0.018	0.158	0.120	0.455	0.108	0.096	0.100	0.947	0.415
1	0.3	-0.013	0.020	0.122	0.084	0.358	0.116	0.102	0.107	0.943	0.443
1.25	0.05	-0.004	0.017	0.201	0.165	0.978	0.110	0.103	0.104	0.950	0.424
1.25	0.1	0.007	0.018	0.140	0.107	0.658	0.115	0.106	0.108	0.949	0.443
1.25	0.2	-0.006	0.020	0.092	0.061	0.429	0.122	0.113	0.118	0.946	0.469
1.25	0.3	-0.006	0.021	0.067	0.039	0.335	0.127	0.118	0.124	0.943	0.485
1.5	0.05	0.008	0.020	0.120	0.089	0.943	0.126	0.118	0.120	0.949	0.489
1.5	0.1	0.020	0.021	0.083	0.055	0.625	0.129	0.123	0.126	0.949	0.500
1.5	0.2	-0.001	0.022	0.051	0.029	0.405	0.134	0.130	0.135	0.945	0.509
1.5	0.3	-0.001	0.026	0.035	0.016	0.310	0.136	0.133	0.141	0.962	0.509

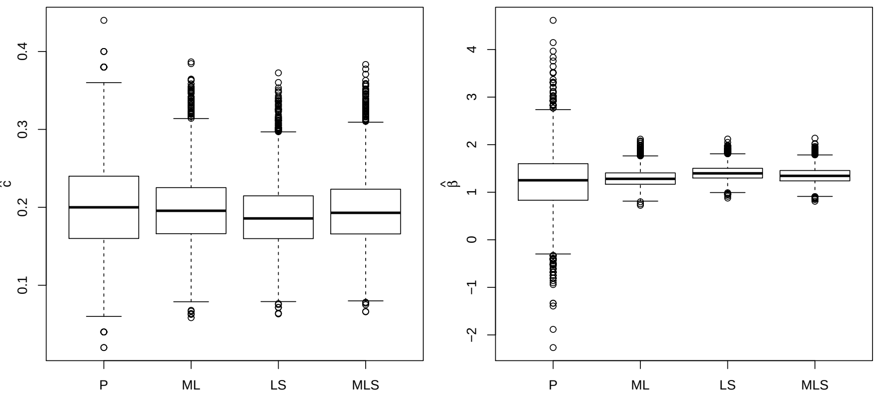


Fig. 3: Monte Carlo distribution of estimators $\hat{c}$ (left) and $\hat{\beta}$ (right), when $c = 0.2$, $\beta = 1.25$, $n = 50$. Please note that the y-scale in the two boxplots is not the same.

puter hard disk are reported. The observed sample values are

$$5,1,1,1,3,2,4,3,2,3,1,1,1,3,1,3,1,6,4,1,9,2,6,2,1,3,1,3,$$
$$1,1,10,2,7,1,7,1,1,2,1,1,6,1,2,1,4,1,1,1,3,5,1,1,1,1,5,2,$$
$$4,5,1,2,2,1,3,1,1,1,3,1,2,1,1,1,1,1,1,5,2,2,4,6,1,3,1,1,1$$

and their empirical distribution is

x_i	1	2	3	4	5	6	7	9	10
O_i	43	13	11	5	5	4	2	1	1

On this sample, we can apply the inferential techniques described in Section 2. The maximum likelihood estimates are $\hat{c}_{ML} = 0.4725$ and $\hat{\beta}_{ML} = 0.8053$. The estimates derived by the method of proportion are $\hat{c}_P = 0.5059$ and $\hat{\beta}_P = 0.2913$. The least-squares method provides $\hat{c}_{LS} = 0.4639$ and $\hat{\beta}_{LS} = 0.8846$ and its modified version $\hat{c}_{MLS} = 0.4708$ and $\hat{\beta}_{MLS} = 0.8384$. The estimate of the parameter c of the (hypothesized) underlying discrete Weibull distribution are quite close to each other. The estimates for the other parameter β are quite similar for the maximum likelihood and least-squares methods, but they are both quite different from the estimate by the method of proportion. However, the four estimates for β being all smaller than 1, it descends that the corresponding (estimated) support of the discrete model is unbounded and the failure rate is estimated to be decreasing. Approximate 95% CIs for c and β can be built separately as $(0.3697, 0.5754)$ and $(0.5416, 1.0691)$, respectively.

Figure 4 displays the relative frequency of the observed sample values along with their theoretical probability.

In figure 5, the sample failure rate function is plotted (filled circles), for each observed value x. Note that since the value 8 is not observed, the corresponding sample failure rate is zero; the same happens for each integer value greater than

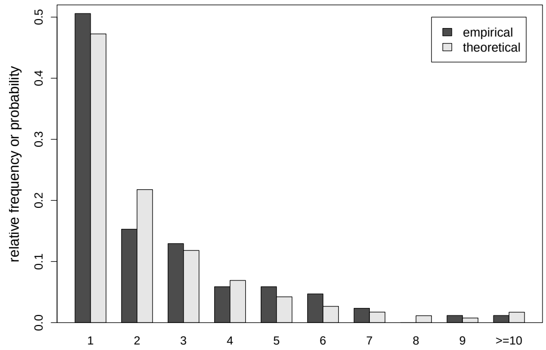


Fig. 4: Empirical distribution of sample values for the first dataset and corresponding theoretical probabilities under the type II discrete Weibull model with the two parameters set equal to their MLE.

$x_{\max} = 9$. The estimated failure rate function (3) (computed according to the Weibull model with parameters set equal to their least-squares estimates) is superimposed (empty squares).

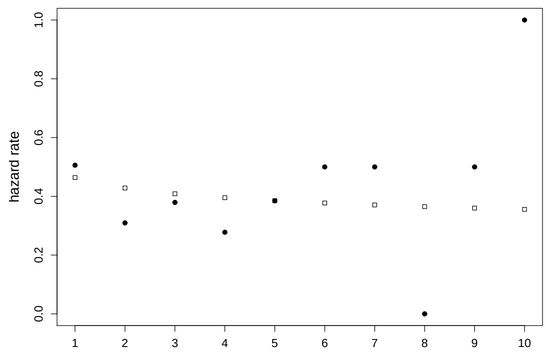


Fig. 5: Sample and theoretical failure rate functions (represented by filled circles and empty squares, respectively) for the first dataset under the type II discrete Weibull model with the two parameters set to their least-squares estimates.

We can then try to check if the discrete Weibull model (with its two parameters estimated through the maximum likelihood method) fits the data adequately. For this aim, we resort to the chi-squared goodness-of-fit test, after having properly grouped some values in order to ensure that each value or group of values has an expected frequency at least equal to 5. For this case, we need to collapse into a unique class all the values equal to or greater than 5. The observed and expected frequencies (the latter denoted by E_i) are then those re-

Table 7: Observed vs. expected frequencies for the first dataset

x_i	1	2	3	4	≥ 5
observed frequency O_i	43	13	11	5	13
expected frequency E_i	40.17	18.51	10.04	5.87	10.41

ported in Table 7. The chi-squared statistic, given by $\chi^2 = \sum_{i=1}^{h=5}(O_i - E_i)^2/E_i$, takes value 2.707. Now, under the hypothesized discrete model, the chi-square statistic is asymptotically distributed as a chi square with $h-2-1=2$ degrees of freedom. Then, the associated p-value is $P(\chi^2_2 > 2.707) = 0.258$. We can accept the hypothesis the observed sample comes from a discrete Weibull type II distribution with parameters c and β equal to the corresponding MLEs.

Now we can compare the type II discrete Weibull model with other alternative discrete models suggested in the literature for fitting reliability data. The comparison can be carried out considering the log-likelihood function computed on its MLEs, and then calculating, for example, the AIC (Akaike Information Criterion) index, given by $AIC = 2k - 2\ell$, where k is the number of parameters of the model and ℓ denotes the log-likelihood function computed plugging-in the MLEs of the parameters. Here, we will consider the geometric (G) distribution as a competing stochastic model, whose pmf is

$$P(X = x) = p(1-p)^{x-1} \quad x = 1, 2, \dots \tag{24}$$

which is - as already pointed out - a particular case of the type II discrete Weibull model; the type I discrete Weibull (W1) distribution (Nakagawa and Osaki 1975), whose pmf is

$$P(X = x) = q^{(x-1)^\beta} - q^{x^\beta} \quad x = 1, 2, \dots \tag{25}$$

with $0 < q < 1$ and $\beta > 0$; the discrete inverse Weibull (IW) distribution (Jazi, Lai, and Alamatsaz 2010), with pmf

$$P(X = x) = \begin{cases} q & x = 1 \\ q^{x^{-\beta}} - q^{(x-1)^{-\beta}} & x = 2, 3, \dots \end{cases} \tag{26}$$

with $0 < q < 1$ and $\beta > 0$; the zero-truncated Poisson-Lindley (ZTPL) distribution (Ghitany, Al-Mutairi, and Nadarajah 2008):

$$P(X = x) = \frac{\theta^2}{\theta^2 + 3\theta + 1} \frac{x + \theta + 2}{(\theta + 1)^x} \quad x = 1, 2, \dots \tag{27}$$

with $\theta > 0$; and, finally, the generalized binomial (GNB) distribution, with pmf

$$P(X = x) = \begin{cases} 1 - b & x = 1 \\ ab(1-a)^{x-2} & x = 2, 3, \dots, \end{cases} \tag{28}$$

The MLE of the parameter of the geometric distribution is $\hat{p}_G = n/\sum_{i=1}^n x_i = 0.417$. The MLEs of the parameters of the type I discrete Weibull distribution are $\hat{q}_{W1} = 0.516$ and $\hat{\beta}_{W1} = 0.823$. The MLEs of the parameters of

Table 8: Values of the AIC index for the type II discrete Weibull model (W2) and alternative discrete models for the first dataset

model	W2	G	W1	IW	ZTPL	GNB
AIC index	278.936	279.1109	278.1796	289.9503	280.7802	276.38

the discrete inverse Weibull distribution are $\hat{q}_{IW} = 0.490$ and $\hat{\beta}_{IW} = 1.193$. The MLE of the parameter of the zero-truncated Poisson-Lindley distribution is $\hat{p}_{ZTPL} = 1.005$. Finally, for the generalized negative binomial model, the MLEs are $\hat{a}_{GNB} = 0.355$ and $\hat{b}_{GNB} = 0.5$. The values of the AIC index for the alternative models listed above are reported in Table 8 and show that actually the generalized negative binomial provides the best fit, followed by the type I and type II discrete Weibull distributions.

4.2 Immunogold assay data

Cullen, Walsh, Nicholson, and Harris (1990) gave counts of sites with 1, 2, 3, 4 and 5 particles from immunogold assay data. The counts were 122, 50, 18, 4, 4. The MLEs of the parameters of the type II discrete Weibull model are $\hat{c}_{ML} = 0.615$ and $\hat{\beta}_{ML} = 1.094$. Note that $\hat{\beta}_{ML} > 1$, thus the distribution is estimated to have an upper-bounded support ($m = 173$), although the estimated cumulative probability of the maximum observed value (5) is very close to 1 (≈ 0.996). The parameter estimates derived through the method of proportion are $\hat{c}_P = 0.616$ and $\hat{\beta}_P = 1.095$, those derived through the least-squares method are $\hat{c}_{LS} = 0.615$ and $\hat{\beta}_{LS} = 1.105$, and finally those provided by the modified least-squares method are $\hat{c}_{MLS} = 0.620$ and $\hat{\beta}_{MLS} = 1.058$. It is very interesting to note that the four estimates of c are almost identical; similarly, the four estimates of β are very close to each other and slightly larger than 1. Thus in this example, the four point estimators seem to converge much more than for the first dataset analyzed in the previous subsection. Approximate 95% CIs for c and β can be built separately as $(0.5496, 0.6814)$ and $(0.9149, 1.2732)$, respectively.

We are interested in testing the null hypothesis H_0: "Number of attached particles is the type II dicrete Weibull random variable" versus the alternative hypothesis H_1: "Number of attached particles is not the type II dicrete Weibull random variable". Plugging in the MLEs for c and β, we can compute the expected theoretical frequencies and then, after merging the last two observed values in order to exceed the 5 threshold, the chi-square statistic. The chi-square statistic takes the value $\chi^2 = 0.0123$ and the p-value of the test is given by $P(\chi^2_{4-2-1} \geq 0.0123) = 0.912$. It follows that the null hypothesis H_0 cannot be rejected; indeed, the large p-value and the close agreement between the observed and expected frequencies (Table 9) suggest that the type II discrete Weibull distribution provides very good fit.

Supposing the data has been generated by a type I discrete Weibull model (with support given by the set of positive integers), we can derive the MLEs,

Table 9: Observed vs. expected frequencies for the immunogold assay data

x_i	1	2	3	4	($\geq$)5
observed frequency O_i	122	50	18	4	4
expected frequency E_i	121.9	50.0	17.8	5.8	2.5

Table 10: Sample and theoretical hazard function for the immunogold assay data

x_i	1	2	3	4	5
sample hazard rate $\hat{r}_i$	0.616	0.658	0.692	0.5	1
expected hazard rate r_i^*	0.615	0.657	0.682	0.701	0.716

$\hat{q} = 0.384$ and $\hat{\beta} = 1.093$, and again compute the value of the corresponding χ^2 statistic, which is 0.0264, with a p-value 0.871. The zero-truncated Poisson Lindley distribution (Ghitany, Al-Mutairi, and Nadarajah 2008) (with the MLE of its unique parameter p equal to 2.183) provides a p-value for the χ^2 test equal to 0.511. In relative terms, computing the AIC, we have that the best model is the zero-truncated Poisson Lindley distribution ($AIC = 411.2441$), followed by the type II (412.6335) and type I (412.6813) discrete Weibull.

In Table 10, the values of the sample hazard rate function are reported for each distinct observed value; the theoretical values, computed recalling the formula (3) plugging in the MLEs into c and β for the type II discrete Weibull model, are reported as well.

5 Final remarks

In this paper we considered the type II discrete Weibull distribution, built by mimicking the hazard rate function of the analogous continuous model, and presented four types of point estimators for its two parameters, derived using the maximum likelihood method, the method of proportion, the least-squares method, and a modification of the latter. The construction of large-sample confidence intervals based on the maximum likelihood estimates and Fisher information matrix was also considered. We first outlined the applicability and properties of such estimators and secondly assessed and compared their performance (in terms of bias and standard error; coverage rate and average length) via intensive simulation experiments; both aspects were missing in the literature.

The simulations led to identify the features of each estimator for different combinations of values of the two parameters and different sample sizes. Although there is not an overall 'uniformly most efficient' estimator for all the experimental settings examined, however, as a general result, despite its non-analytical form, the maximum likelihood estimator is the most reasonable choice; the two types of least-squares-based estimators need some further modification, which has to be explored, in order to reduce their bias; the method of proportion, though providing estimates with an analytical expression, however discards most of the information contained in the sample and thus reveals nothing much reliable. Attention should be paid when dealing with small samples, where the method of proportion may fail to produce estimates and large-sample confidence intervals may be not computable.

Two applications to real data prove that the discrete Weibull model can fit data even better than more popular distributions; the implementation of the model and the related inferential techniques have been easily worked out in the `R` programming environment and will be made available through a new version of the `DiscreteWeibull` package (Barbiero 2015).

Thus, despite the criticality highlighted in this paper, the discrete Weibull model has been shown to be a competitive stochastic model for count data and this may encourage its use and the further development of *ad-hoc* inferential techniques. In this work, we focused on estimation under the classical frequentist framework, but one can also think of exploring Bayesian estimation, whose formulation would require a specification of prior distributions for the two parameters and whose solution would possibly require the use of typical Bayesian tools, such as MCMC. Since frequentist estimation has already risen some issues, we expect that Bayesian estimation, besides requiring a totally different approach, would be even more challenging, and thus the subject would probably deserve a work of its own.

References

Barbiero A (2013) Parameter Estimation for Type III Discrete Weibull Distribution: A Comparative Study. J Prob Stat `http://dx.doi.org/10.1155/2013/946562`.

Barbiero A (2015) Discrete Weibull Distributions (Type 1 and 3). R package version 1.0.1.

Barbiero A (2016) A comparison of methods for estimating parameters of the type I discrete Weibull distribution. Stat Interface 9(2):203-212.

Bebbington MS, Lai C-D (1998) A generalized negative binomial and applications. Comm Stat - Theory Methods 27(10):2515–2533.

Bracquemond C, Gaudoin O, Roy D, Xie M (2001) On some discrete notions of aging. In: System and Bayesian Reliability, a volume in honor of Richard E. Barlow (Hayakawa Y, Irony T, Xie M, editors):185–197. World Scientific, Singapore.

Bracquemond C, Gaudoin O (2002) The empirical failure rate of discrete reliability data. 3rd International Conference on Mathematical Methods in Reliability, MMR 2002, Trondheim, Norway, June 2002.

Bracquemond C, Gaudoin O (2003) A survey on discrete lifetime distributions. Int J. Reliab Qual Safety Eng 10(1):69–98.

Cullen MJ, Walsh J, Nicholson LV, Harris JB (1990) Ultrastructural localization of dystrophin in human muscle by using gold immunolabelling. Proc Royal Society London, Series B 20:197–210.

Ghitany ME, Al-Mutairi DK, Nadarajah S (2008) Zero-truncated Poisson-Lindley distribution and its application. Math Comp Simul 79:279–287.

Khalique A (1989) On discrete failure-time distributions. Reliab Eng Syst Saf 25(2):99–107.

Khan MSA, Khalique A, Abouammoh AM (1989) On Estimating Parameters in a Discrete Weibull Distribution. IEEE Trans Reliab 38(3):348–350.

Kulasekera KB (1993) Approximate MLE's of the parameters of a discrete Weibull distribution with type I censored data. Microel Reliab 34(7):1185–1188.

Jazi MA, Lai C-D, Alamatsaz MH (2010) A discrete inverse Weibull distribution and estimation of its parameters. Stat Methodol 7(2):121–132.

Jiang R (2013) A new bathtub curve model with a finite support. Reliab Eng Syst Saf 119:44–51.

Lai C-D, Xie M, Murthy DNP (2003) A Modified Weibull Distribution. IEEE Trans Reliab 25:33–37.

LeCam L (1970) On the Assumptions Used to Prove Asymptotic Normality of Maximum Likelihood Estimates. Annals Math Stat 41(3):802–828.

Nakagawa T, Osaki S (1975) The discrete Weibull distribution. IEEE Trans Reliab 24:300–301.

Noughabi MS, Borzadaran GRM, Roknabadi AHR (2011) Discrete modified Weibull distribution. Metron 49(2):207–222.

Noughabi MS, Roknabadi AHR, Borzadaran GRM (2013) Some discrete lifetime distributions with bathtub-shaped hazard rate functions. Qual Eng 25:225-236.

Padgett WJ, Spurrier JD (1985) Discrete failure models. IEEE Trans Reliab 34:253–256.

R Development Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.

Sen PK, Singer JM, Pedroso De Lima AC (2010) From Finite Sample to Asymptotic Methods in Statistics. Cambridge: Cambridge University Press.

Stein WE, Dattero R (1984) A new discrete Weibull distribution. IEEE Trans Reliab 33:196–197

Xie M, Gaudoin O, Bracquemond C(2002) Redefining failure rate function for discrete distributions Int J Reliab Qual Saf Eng, 9(3):275–286.