UDC 519.2

M. FUSEK, J. MICHÁLEK

STATISTICAL INFERENCE FOR TYPE I MULTIPLY LEFT-CENSORED SAMPLES FROM WEIBULL DISTRIBUTION[1]

Abstract. Left-censored data with one or more detection limits occur frequently in many application areas. In this paper, the computational procedure for calculation of maximum likelihood estimates of the parameters for type I multiply left-censored data from underlying Weibull distribution is suggested and used considering various numbers of detection limits. The expected Fisher information matrix is analytically determined and its performance is compared with sample (observed) Fisher information matrix using simulations. Simulations are focused primarily on the properties of estimators for small sample sizes. Real data illustration is included.

Keywords: Fisher information matrix, maximum likelihood estimator, multiply left-censored sample, type I censoring, Weibull distribution.

INTRODUCTION

Left-censored data occur frequently in analyses of environmental or chemical data when for a given experimental unit the attribute being measured is not present above the detection limit (DL) d. In case the DL is fixed, we talk about type I (time) censoring. The number of experimental units under DL d, i.e. the number of left-censored experimental units, is a random variable. In case the number of censored units is fixed, we talk about type II (failure) censoring. This paper will be focused on the type I left censoring which can be described as follows. Assume that n experimental units are observed, and $X_1, \ldots, X_n$ are independent and identically distributed random variables. As a result of the experiment we get only measurements $X_i > d$ and the number of observations under the detection limit d. Most authors usually focus on right censoring. However, type I left-censored data are very frequent in real applications, and a brief overview of application areas can be found in [1]. In addition to that, multiply left-censored data often occur, mostly in environmental applications. They arise in case there are more, say k, detection limits $d_1 < d_2 < \ldots < d_k$, $k > 1$, and only observations above the highest detection limit d_k and the numbers of observations between detection limits d_{i-1} and d_i, $i = 1, \ldots, k$, $d_0 = 0$, are available [2–6]. In case $k = 2$, we talk about doubly left-censored data [7].

Various censoring techniques and statistical analyses of censored data are described in more details in many monographs [4, 8]. In order to estimate unknown parameters of particular distributions, the maximum likelihood (ML) approach is usually applied [9–11]. The ML method for estimating unknown parameters of the left-censored normal and log-normal distributions is well developed [5, 12, 13]. However, the interest in statistical analyses of censored data with asymmetric distributions which cannot be easily transformed into symmetric ones (e.g. normal distribution) has increased in many application areas recently. An overview of such distributions together with description of their properties can be found, for example, in [14]. Due to the monotonicity property of the hazard function and various shapes of the probability density function, Weibull distribution is very popular. Various approaches to inference on the parameters of the Weibull distribution can be found, for example, in [15–18]. A demonstration of parameters estimates properties based on numerical study for censored and uncensored Weibull

[1] The paper was written with the support of the research project PSVŘ 2 (University of Defence, Brno).

distribution can be found in [19]. Moreover, the generalized exponential (GE) distribution [20], which is similar to Weibull distribution, has been of great interest lately. ML estimates of unknown parameters in case of censored GE distribution [1, 21, 22] were studied as well. In addition to that, some authors deal with comparison of censored Weibull and GE distributions with each other ([23] and references inside) and also with gamma and log-normal distributions [1, 24].

An advantage of the ML approach lies (under certain regularity conditions) in good asymptotic properties of the obtained estimates. The speed of convergence of estimators distribution to normal distribution and the asymptotic bias can be quite easily assessed in practical situations using simulations. The properties of ML estimates in case of singly, doubly and triply left-censored exponential distributions (a special case of the GE distribution) based on simulations can be found in [6]. On the other hand, ML approach often requires special numerical algorithms for solving likelihood equations [25]. In case of Weibull distribution, a simple fixed point type algorithm for calculating the ML estimates of unknown parameters was proposed in [19]. A proper optimization procedures often based on EM algorithm [26, 27] can also be used. This approach is rather beneficial when calculating ML estimates of unknown parameters. However, estimation of parameters variances using the Fisher information matrix (FIM) is computationally demanding. This inconvenience can be overcome using the exact formulas for calculating the (expected) FIM proposed in this paper.

This contribution is organized as follows: in Section 1, a derivation of a computational procedure for determination of ML estimates of parameters of type I multiply left-censored Weibull distribution is described. In Section 2, the expected FIM is analytically determined. In Section 3, bias of the estimates of Weibull distribution parameters considering various DLs (censoring schemes) and censoring multiplicity $k = 1, 2, 3$ (single, double and triple censoring) is described. Furthermore, the sample (observed) FIM and the expected FIM are compared and the bias of the FIM estimates is analyzed using simulations. Simulations focus primarily on estimator properties for small sample sizes and also on comparing the estimates from singly ($k = 1$), doubly ($k = 2$) and triply ($k = 3$) left-censored samples. The paper was motivated by the need to process real environmental data described in [28]. The application of the derived method is presented in the last section.

1. ML ESTIMATION

Let $X_1, \ldots, X_n$ be a random sample from Weibull distribution with scale parameter $\lambda > 0$, shape parameter $\tau > 0$, cumulative distribution function (cdf)

$$F(x, \lambda, \tau) = \begin{cases} 1 - \exp\left[-\left(\dfrac{x}{\lambda} \right)^{\tau} \right] & \text{for } x \geq 0, \\ 0 & \text{for } x < 0, \end{cases} \tag{1}$$

probability density function (pdf)

$$f(x, \lambda, \tau) = \begin{cases} \dfrac{\tau}{\lambda^{\tau}} x^{\tau - 1} \exp\left[-\left(\dfrac{x}{\lambda} \right)^{\tau} \right] & \text{for } x \geq 0, \\ 0 & \text{for } x < 0, \end{cases} \tag{2}$$

and skewness

$$\gamma(\lambda, \tau) = \frac{2\Gamma^3\left(1 + \dfrac{1}{\tau}\right) - 3\Gamma\left(1 + \dfrac{1}{\tau}\right)\Gamma\left(1 + \dfrac{2}{\tau}\right) + \Gamma\left(1 + \dfrac{3}{\tau}\right)}{\left[\Gamma\left(1 + \dfrac{2}{\tau}\right) - \Gamma^2\left(1 + \dfrac{1}{\tau}\right)\right]^{3/2}}.$$

Furthermore, let $X_{(1)}, \ldots, X_{(n)}$ be the ordered sample of $X_1, \ldots, X_n$ which is type I multiply left-censored with detection limits $d_1, \ldots, d_k$ and we put $d_0 = 0$. Moreover, N_i is the number of observations in the interval $(d_{i-1}, d_i\rangle$ and N_0 is the number of uncensored observations $X_{(n-N_0+1)}, \ldots, X_{(n)}$. Considering the type I censoring, the random vector $(N_0, N_1, \ldots, N_k)$ has multinomial distribution $\mathrm{Mu}_{k+1}(n, \theta_0, \theta_1, \ldots, \theta_k)$, where $\theta_i = F(d_i, \lambda, \tau) - F(d_{i-1}, \lambda, \tau)$, $i = 1, \ldots, k$, $\theta_0 = 1 - F(d_k, \lambda, \tau)$ and $n = N_0 + N_1 + \ldots + N_k$. Particular marginal frequencies N_i have binomial distribution $\mathrm{Bi}(n, \theta_i)$, $i = 0, 1, \ldots, k$.

Using the results from [4], the log-likelihood function of the censored sample can be written in the form of

$$l(\lambda, \tau, N_0, \ldots, N_k, X_{(n-N_0+1)}, \ldots, X_{(n)}) = \log\left(\frac{n!}{N_1! \ldots N_k!}\right) + \\ + \sum_{i=n-N_0+1}^{n} \log[f(X_{(i)})] + \sum_{i=1}^{k} N_i \log[F(d_i, \lambda, \tau) - F(d_{i-1}, \lambda, \tau)], \qquad (3)$$

and we put $\sum_{i=n-N_0+1}^{n} \log[f(X_{(i)})] = 0$ for $N_0 = 0$.

The usual approach to estimation of parameters λ, τ is to derive likelihood equations and solve them numerically using, for example, the Newton–Raphson method or other methods for searching of extremes (see e.g. [29]) can be used. However, it was found out that in case the shape parameter value is low, there are numerical difficulties with obtaining the solution. For that very reason, ML estimates $\hat{\lambda}$ and $\hat{\tau}$ were obtained by maximizing the log-likelihood function (3) using the Nelder–Mead simplex algorithm [30]. When using this type of algorithm, it is necessary to select initial values of the parameters that need to be estimated. Starting values of the algorithm were selected using the moment estimator of parameters of the Weibull distribution based on samples in which the censored observations were replaced by a constant lying between the detection limits.

2. FISHER INFORMATION MATRIX

In this section, the expected FIM for type I multiply left-censored samples will be derived. According to [9], the sample FIM can be defined (under the regularity conditions) using formula

$$\widetilde{\boldsymbol{J}}_n = \begin{bmatrix} \widetilde{J}_{11} & \widetilde{J}_{12} \\ \widetilde{J}_{21} & \widetilde{J}_{22} \end{bmatrix} = \begin{bmatrix} -\dfrac{\partial^2 l}{\partial \lambda^2} & -\dfrac{\partial^2 l}{\partial \lambda \partial \tau} \\ -\dfrac{\partial^2 l}{\partial \tau \partial \lambda} & -\dfrac{\partial^2 l}{\partial \tau^2} \end{bmatrix}. \qquad (4)$$

The sample FIM $\widetilde{\boldsymbol{J}}_n$ is an unbiased estimator of the expected FIM $\boldsymbol{J}_n$ and $\widetilde{\boldsymbol{J}}_n(\lambda, \tau) \to \boldsymbol{J}_n(\lambda, \tau)$ in probability for $n \to \infty$. On that account, in many applications when the exact determination of the expected FIM is complicated, the sample FIM is used instead [3, 31]. One major disadvantage of this approach is the rather extensive variability of the sample FIM. Many authors prefer another approach like bootstrap or Bayesian methods [19] for description of variability of parameters estimates. The expected FIM can be used for statistical inference and more precise description of the asymptotic variability of obtained estimates. On that account, our attention will now be paid to determination of the expected FIM. Similar problems were solved, for example, in [6] for exponential distribution, in [1] for GE distribution or in [22] and [23] using the hazard function according to [32].

The expected FIM can be calculated using formula

$$\boldsymbol{J}_n = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} = \mathrm{E}\widetilde{\boldsymbol{J}}_n = \begin{bmatrix} \mathrm{E}\widetilde{J}_{11} & \mathrm{E}\widetilde{J}_{12} \\ \mathrm{E}\widetilde{J}_{21} & \mathrm{E}\widetilde{J}_{22} \end{bmatrix}, \tag{5}$$

where

$$J_{11} = n\sum_{i=1}^{k} \frac{\tau^2 \left\{ d_{i-1}^{\tau} \exp\left[-\left(\frac{d_{i-1}}{\lambda} \right)^{\tau} \right] - d_i^{\tau} \exp\left[-\left(\frac{d_i}{\lambda} \right)^{\tau} \right] \right\}^2}{\lambda^{2\tau+2} \left\{ \exp\left[-\left(\frac{d_{i-1}}{\lambda} \right)^{\tau} \right] - \exp\left[-\left(\frac{d_i}{\lambda} \right)^{\tau} \right] \right\}} -$$

$$- n \frac{(d_k^{\tau} \lambda^{\tau} \tau^2 + d_k^{\tau} \lambda^{\tau} \tau - d_k^{2\tau} \tau^2) \exp\left[-\left(\frac{d_k}{\lambda} \right)^{\tau} \right]}{\lambda^{2\tau+2}} - n \frac{\tau}{\lambda^2} \exp\left[-\left(\frac{d_k}{\lambda} \right)^{\tau} \right] + \frac{\tau^2 + \tau}{\lambda^{\tau+2}} E_1,$$

$$J_{22} = n\sum_{i=1}^{k} \frac{\left\{ d_{i-1}^{\tau} \ln\left(\frac{d_{i-1}}{\lambda} \right) \exp\left[-\left(\frac{d_{i-1}}{\lambda} \right)^{\tau} \right] - d_i^{\tau} \ln\left(\frac{d_i}{\lambda} \right) \exp\left[-\left(\frac{d_i}{\lambda} \right)^{\tau} \right] \right\}^2}{\lambda^{2\tau} \left\{ \exp\left[-\left(\frac{d_{i-1}}{\lambda} \right)^{\tau} \right] - \exp\left[-\left(\frac{d_i}{\lambda} \right)^{\tau} \right] \right\}} -$$

$$- n \frac{(d_k^{\tau} \lambda^{\tau} - d_k^{2\tau}) \left[\ln\left(\frac{d_k}{\lambda} \right) \right]^2 \exp\left[-\left(\frac{d_k}{\lambda} \right)^{\tau} \right]}{\lambda^{2\tau}} + \frac{n}{\tau^2} \exp\left[-\left(\frac{d_k}{\lambda} \right)^{\tau} \right] +$$

$$+ \frac{(\ln \lambda)^2}{\lambda^{\tau}} E_1 - \frac{2 \ln \lambda}{\lambda^{\tau}} E_2 + \frac{1}{\lambda^{\tau}} E_3,$$

$$J_{12} = J_{21} = -n\sum_{i=1}^{k} \frac{\tau \left\{ d_{i-1}^{\tau} \exp\left[-\left(\frac{d_{i-1}}{\lambda} \right)^{\tau} \right] - d_i^{\tau} \exp\left[-\left(\frac{d_i}{\lambda} \right)^{\tau} \right] \right\}}{\lambda^{2\tau+1} \left\{ \exp\left[-\left(\frac{d_{i-1}}{\lambda} \right)^{\tau} \right] - \exp\left[-\left(\frac{d_i}{\lambda} \right)^{\tau} \right] \right\}} \times$$

$$\times \left\{ d_{i-1}^{\tau} \ln\left(\frac{d_{i-1}}{\lambda} \right) \exp\left[-\left(\frac{d_{i-1}}{\lambda} \right)^{\tau} \right] - d_i^{\tau} \ln\left(\frac{d_i}{\lambda} \right) \exp\left[-\left(\frac{d_i}{\lambda} \right)^{\tau} \right] \right\} +$$

$$+ n \frac{\left[d_k^{\tau} \tau \lambda^{\tau} \ln\left(\frac{d_k}{\lambda} \right) + d_k^{\tau} \lambda^{\tau} - d_k^{2\tau} \tau \ln\left(\frac{d_k}{\lambda} \right) \right] \exp\left[-\left(\frac{d_k}{\lambda} \right)^{\tau} \right]}{\lambda^{2\tau+1}} +$$

$$+ \frac{n}{\lambda} \exp\left[-\left(\frac{d_k}{\lambda} \right)^{\tau} \right] + \frac{\tau \ln \lambda - 1}{\lambda^{\tau+1}} E_1 - \frac{\tau}{\lambda^{\tau+1}} E_2.$$

The explicit expressions for E_1, E_2, E_3 and details of the derivation can be found in the Appendix 1.

The expected FIM can be used for statistical inference of parameters λ and τ. Considering the asymptotic properties of the ML estimator $\hat{\lambda}$ ($\hat{\tau}$ respectively), according to [10], $\sqrt{n}(\hat{\lambda}-\lambda)$ ($\sqrt{n}(\hat{\tau}-\tau)$ respectively) has asymptotically normal distribution $\mathrm{N}(0, \sigma_{\lambda}^2)$ ($\mathrm{N}(0, \sigma_{\tau}^2)$ respectively), where $\sigma_{\lambda}^2 = J_{11}^{-1}$ ($\sigma_{\tau}^2 = J_{22}^{-1}$ respectively). The properties of estimators $\hat{\lambda}$, $\hat{\tau}$ considering various sample sizes n, various number of detection limits k and various censoring schemes will be analyzed using simulations in the next section.

3. ESTIMATORS BEHAVIOR BASED ON SIMULATIONS

The estimates of parameters and parametric functions derived in the previous section will be analyzed and compared using simulations. Firstly, 20,000 Type I singly, doubly and triply left-censored random samples with size $n = 10, 20, 30, 50, 100$ from Weibull distribution were generated. Since λ is the scale parameter, and the maximum likelihood estimators are scale invariant, we take $\lambda = 1$ without loss of generality. In order to describe various shapes of Weibull distribution, we take $\tau = 0.5,\ 1.5,\ 3$, which corresponds to skewnesses $\gamma = 6.62,\ 1.07,\ 0.17$. Limits of detection d_i, $i = 1, \ldots, k$, $k = 1, 2, 3$, were chosen as quantiles of the Weibull distribution using equations $q_i = F(d_i, \lambda, \tau)$, where q_i are given in Table 1. Particular censoring schemes are denoted as $c_1, \ldots, c_9$, and correspond to various quantiles which determine detection limits for $k = 1, 2, 3$. Thus, for example, the q_1 given in column "Double" in Table 1 denotes the proportion of doubly censored observations, and describes the given censoring scheme. The censoring scheme c_1 represents the smallest proportion (10%) of censored data, and the censoring scheme c_9 represents the largest proportion (90%) of censored data in case of singly, doubly and triply censored samples. The large proportion of censored data corresponds to real data which will be analyzed later on.

Next, ML estimates of parameters λ, τ from 20,000 samples were calculated by maximization of (3), and their mean values $\bar{\hat{\lambda}}$, $\bar{\hat{\tau}}$ were determined. It was found out (see Fig. 1) that the estimate of parameter λ has lower bias for higher values of parameter τ (i.e. for lower skewness) considering various censoring schemes from Table 1. In all the following tables, only results for censoring schemes c_1, c_3, c_5, c_7, c_9 are presented. Moreover, estimates $\hat{\tau}$ are similar bias-wise for various values of τ (not shown in figures). It can be seen from Table 2 (only results for single and double censoring

Table 1

Censoring	Quantiles for determination of DL values considering single, double and triple censoring and various censoring schemes					
	Single	Double		Triple		
	q_1	q_1	q_2	q_1	q_2	q_3
c_1	0.10	0.05	0.10	0.03	0.07	0.10
c_2	0.20	0.10	0.20	0.07	0.13	0.20
c_3	0.30	0.15	0.30	0.10	0.20	0.30
c_4	0.40	0.20	0.40	0.13	0.27	0.40
c_5	0.50	0.25	0.50	0.17	0.33	0.50
c_6	0.60	0.30	0.60	0.20	0.40	0.60
c_7	0.70	0.35	0.70	0.23	0.47	0.70
c_8	0.80	0.40	0.80	0.27	0.53	0.80
c_9	0.90	0.45	0.90	0.30	0.60	0.90

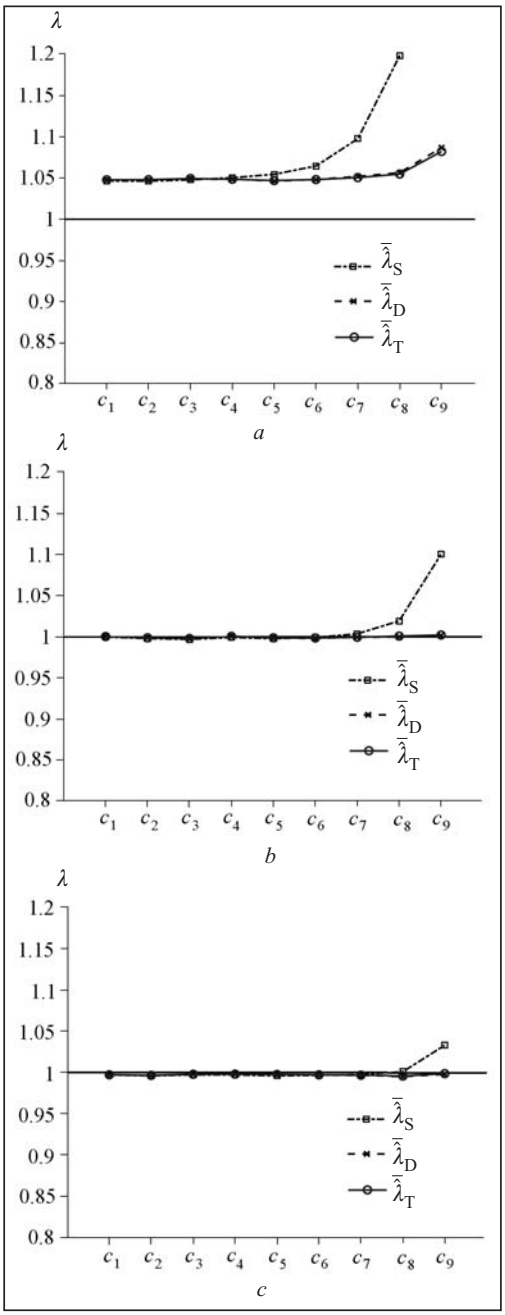

Fig. 1. The average ML estimates of parameter λ considering various values of τ and single (index S), double (index D), triple (index T) censoring; $\tau = 0.5$ (*a*), $\tau = 1.5$ (*b*), $\tau = 3$ (*c*); sample size $n = 30$

are shown), where the average ML estimates $\bar{\hat{\lambda}}$, $\bar{\hat{\tau}}$ and their average mean square errors (MSE) can be found, that the ML estimate $\bar{\hat{\lambda}}$ is rather satisfying until the censoring scheme c_7 even when the sample size is small ($n = 10$). For $n > 10$, the bias of $\bar{\hat{\lambda}}$ is very small, and from the practical point of view negligible for the censoring scheme c_7 and lower. The effect of multiplicity of the censoring on the estimation of parameter λ is noticeable only for higher detection limits depending on the sample size. For $n \geq 30$, the differences among the single, double and triple censoring are almost negligible until scheme c_8 when, in accordance with expectations, the highest bias of estimate is present in case of single censoring. The ML estimate $\bar{\hat{\tau}}$ is significantly biased even when the censoring is low, and sample size $n = 100$ (see Table 2).

Variances of the ML estimators $\hat{\lambda}$, $\hat{\tau}$ of parameters λ, τ were estimated by means of sample variances $s^2(\hat{\lambda})$, $s^2(\hat{\tau})$ of simulated values. Furthermore, the asymptotic variances $\sigma^2(\lambda) = J_{11}^{-1}(\lambda, \tau)$, $\sigma^2(\tau) = J_{22}^{-1}(\lambda, \tau)$ were compared with their estimates $\sigma^2(\hat{\lambda}) = J_{11}^{-1}(\hat{\lambda}, \hat{\tau})$, $\sigma^2(\hat{\tau}) = J_{22}^{-1}(\hat{\lambda}, \hat{\tau})$ based on the expected FIM (5) and $\tilde{\sigma}^2(\hat{\lambda}) = \tilde{J}_{11}^{-1}(\hat{\lambda}, \hat{\tau})$, $\tilde{\sigma}^2(\hat{\tau}) = \tilde{J}_{22}^{-1}(\hat{\lambda}, \hat{\tau})$ based on the sample FIM (4) using simulations. The estimates $\sigma^2(\hat{\lambda})$, $\sigma^2(\hat{\tau})$ and $\tilde{\sigma}^2(\hat{\lambda})$, $\tilde{\sigma}^2(\hat{\tau})$ were averaged over 20,000 repetitions, and from now on when speaking about the estimates $\sigma^2(\hat{\lambda})$, $\sigma^2(\hat{\tau})$, $\tilde{\sigma}^2(\hat{\lambda})$, $\tilde{\sigma}^2(\hat{\tau})$ of the asymptotic variances $\sigma^2(\lambda)$, $\sigma^2(\tau)$, we will have in mind the estimates averaged over 20,000 repetitions. These estimates together with the corresponding empirical sample variances $S^2(\hat{\lambda}) = ns^2(\hat{\lambda})$, $S^2(\hat{\tau}) = ns^2(\hat{\tau})$ will be further compared with the asymptotic variances considering various sample sizes *n.* Due to a rather large number of samples, the estimators $S^2(\hat{\lambda})$, $S^2(\hat{\tau})$ allow us to assess the bias of estimators $\sigma^2(\hat{\lambda})$, $\sigma^2(\hat{\tau})$, $\tilde{\sigma}^2(\hat{\lambda})$, $\tilde{\sigma}^2(\hat{\tau})$, and the bias of asymptotic variances $\sigma^2(\lambda)$, $\sigma^2(\tau)$ from the true (simulation-based) variances $S^2(\hat{\lambda})$, $S^2(\hat{\tau})$ of the estimates.

Table 2

Censoring	The average ML estimates $\bar{\hat{\lambda}}$, $\bar{\hat{\tau}}$ and their average MSE (in the parenthesis) considering single (S) and double (D) censoring and sample size n					
	$\lambda = 1,\ n = 10$		$\lambda = 1,\ n = 30$		$\lambda = 1,\ n = 100$	
	S	D	S	D	S	D
c_1	1.00 (0.050)	1.00 (0.050)	1.00 (0.017)	1.00 (0.017)	1.00 (0.005)	1.00 (0.005)
c_3	0.99 (0.052)	1.00 (0.050)	1.00 (0.017)	1.00 (0.017)	1.00 (0.005)	1.00 (0.005)
c_5	0.99 (0.058)	0.99 (0.051)	1.00 (0.020)	1.00 (0.017)	1.00 (0.006)	1.00 (0.005)
c_7	1.03 (0.071)	1.00 (0.053)	1.00 (0.029)	1.00 (0.018)	1.00 (0.009)	1.00 (0.005)
c_9	1.33 (0.214)	1.07 (0.066)	1.10 (0.112)	1.00 (0.020)	1.02 (0.038)	1.00 (0.006)
Censoring	$\tau = 1.5,\ n = 10$		$\tau = 1.5,\ n = 30$		$\tau = 1.5,\ n = 100$	
	S	D	S	D	S	D
c_1	1.75 (0.355)	1.75 (0.350)	1.57 (0.067)	1.57 (0.065)	1.52 (0.016)	1.52 (0.016)
c_3	1.78 (0.446)	1.76 (0.382)	1.58 (0.081)	1.57 (0.069)	1.52 (0.019)	1.52 (0.016)
c_5	1.96 (58.58)	1.77 (0.408)	1.60 (0.119)	1.58 (0.076)	1.53 (0.028)	1.52 (0.019)
c_7	2.76 (161.9)	1.77 (0.442)	1.67 (0.273)	1.59 (0.091)	1.54 (0.051)	1.52 (0.022)
c_9	6.59 (558.9)	1.62 (0.418)	2.84 (80.11)	1.58 (0.095)	1.66 (0.278)	1.53 (0.026)

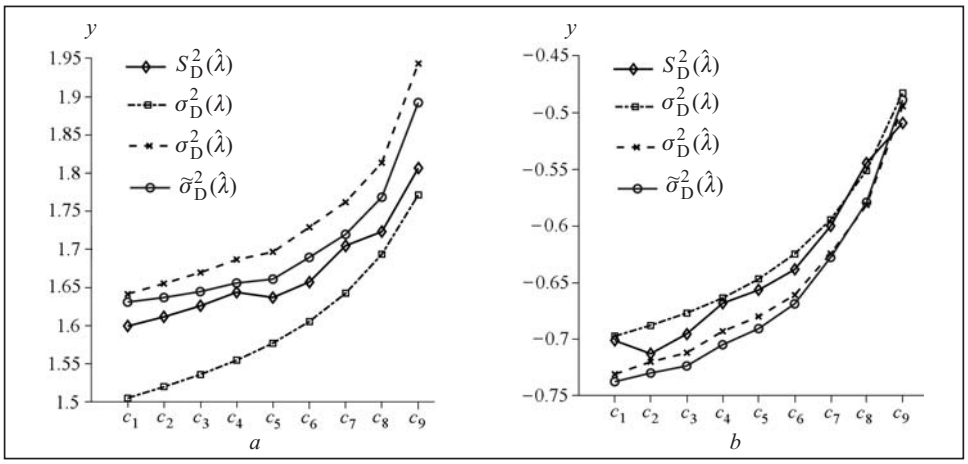

Fig. 2. Comparison of the estimates of variance $S^2(\hat{\lambda})$, $\sigma^2(\hat{\lambda})$, $\tilde{\sigma}^2(\hat{\lambda})$ and the asymptotic variance $\sigma^2(\lambda)$ considering double censoring and sample size $n = 30$; $\tau = 0.5$ (*a*), $\tau = 1.5$ (*b*); logarithmic scale on the *y*-axis

It can be seen from Fig. 2 that behavior of estimates $\sigma^2(\hat{\lambda})$, $\tilde{\sigma}^2(\hat{\lambda})$ is significantly influenced by the value of parameter τ. In case $\tau < 1$, i.e. the skewness of the sample is high (> 2), both estimates are higher than the asymptotic variance $\sigma^2(\lambda)$. When $\tau > 1$, i.e. the skewness of the sample is low (< 2), both estimates are lower than the asymptotic variance $\sigma^2(\lambda)$. The comparison of above mentioned characteristics of variance (see Table 3 in case of double censoring) shows that the anticipated bias of estimator $\sigma^2(\hat{\lambda})$ is substantial for small sample sizes. Furthermore, the estimator $S^2(\hat{\lambda})$ is of lower (higher respectively) values than asymptotic variance $\sigma^2(\lambda)$ for $\tau > 1$ ($\tau < 1$ respectively). All of the estimators of variance almost coincide for $\tau > 1$ and the sample size $n \geq 100$. Furthermore, the asymptotic variance $\sigma^2(\lambda)$ (obtained from the expected FIM) was analyzed considering various sample sizes and censoring schemes. With the exception of schemes $c_1 - c_3$, variability of the estimators is, as expected, the lowest for triple censoring for an arbitrary sample size.

Table 3

Censoring	Comparison of the estimates of variance $S^2(\hat{\lambda})$, $\sigma^2(\hat{\lambda})$, $\tilde{\sigma}^2(\hat{\lambda})$ and the asymptotic variance $\sigma^2(\lambda)$ considering double censoring and sample size n							
	$\tau = 0.5,\ n = 10$				$\tau = 0.5,\ n = 100$			
	$S^2(\hat{\lambda})$	$\sigma^2(\lambda)$	$\sigma^2(\hat{\lambda})$	$\tilde{\sigma}^2(\hat{\lambda})$	$S^2(\hat{\lambda})$	$\sigma^2(\lambda)$	$\sigma^2(\hat{\lambda})$	$\tilde{\sigma}^2(\hat{\lambda})$
c_1	6.2696	4.6015	6.8042	6.6231	4.4943	4.4680	4.6574	4.6416
c_3	6.0831	4.9152	6.9288	6.4813	4.5989	4.5557	4.7488	4.7115
c_5	6.0369	5.2786	7.2980	6.5860	4.8258	4.7072	4.9366	4.8834
c_7	6.3032	5.7804	8.2039	7.1779	5.1093	4.9976	5.1648	5.1014
c_9	8.7636	6.6003	13.0720	11.0980	5.7891	5.6527	5.8591	5.7752
Censoring	$\tau = 1.5,\ n = 10$				$\tau = 1.5,\ n = 100$			
	$S^2(\hat{\lambda})$	$\sigma^2(\lambda)$	$\sigma^2(\hat{\lambda})$	$\tilde{\sigma}^2(\hat{\lambda})$	$S^2(\hat{\lambda})$	$\sigma^2(\lambda)$	$\sigma^2(\hat{\lambda})$	$\tilde{\sigma}^2(\hat{\lambda})$
c_1	0.4954	0.5036	0.4560	0.4474	0.5004	0.4957	0.4904	0.4894
c_3	0.4982	0.5205	0.4707	0.4551	0.4955	0.5039	0.4988	0.4971
c_5	0.5093	0.5391	0.4884	0.4723	0.5191	0.5189	0.5139	0.5122
c_7	0.5257	0.5664	0.5235	0.5144	0.5484	0.5489	0.5434	0.5430
c_9	0.6064	0.6271	0.7248	0.7230	0.6156	0.6176	0.6104	0.6120

Table 4

Censoring	Comparison of the estimates of variance $S^2(\hat{\tau})$, $\sigma^2(\hat{\tau})$, $\tilde{\sigma}^2(\hat{\tau})$ and the asymptotic variance $\sigma^2(\tau)$ considering double censoring and sample size n; $\tau = 1.5$							
	$n = 10$				$n = 100$			
	$S^2(\hat{\tau})$	$\sigma^2(\tau)$	$\sigma^2(\hat{\tau})$	$\tilde{\sigma}^2(\hat{\tau})$	$S^2(\hat{\tau})$	$\sigma^2(\tau)$	$\sigma^2(\hat{\tau})$	$\tilde{\sigma}^2(\hat{\tau})$
c_1	2.8844	1.4203	2.0839	2.1990	1.5242	1.4187	1.4668	1.4764
c_3	3.1440	1.5105	2.2399	2.4523	1.6002	1.5485	1.5978	1.6169
c_5	3.3591	1.6513	2.4820	2.8280	1.8402	1.7315	1.7901	1.8214
c_7	3.6670	1.9000	2.9202	3.2968	2.1143	1.9827	2.0569	2.0976
c_9	4.0300	2.4558	3.1862	3.6107	2.5537	2.3383	2.4623	2.4876

The behavior of $\hat{\tau}$ variance estimators is similar for various values of τ. The comparison of the characteristics of variance (see Table 4 in case of double censoring) shows that the anticipated bias of estimator $\sigma^2(\hat{\tau})$ is substantial for small sample sizes. Furthermore, the estimator $S^2(\hat{\tau})$ is of higher values than asymptotic variance $\sigma^2(\tau)$ for all sample sizes and censoring schemes. Furthermore, the asymptotic variance $\sigma^2(\tau)$ (obtained from the expected FIM) was analyzed considering various sample sizes and censoring schemes. With the exception of schemes $c_1 - c_2$, the variability of the estimators is, as expected, the lowest for triple censoring for arbitrary sample size.

Finally, using the variance estimators $\sigma^2(\hat{\lambda})$, $\sigma^2(\hat{\tau})$, $\tilde{\sigma}^2(\hat{\lambda})$, $\tilde{\sigma}^2(\hat{\tau})$, the lower and the upper confidence limits of the estimate of parameters λ and τ can be obtained. The coverage probability of 95% confidence interval, computed as the proportion of the number of times, out of 20,000 replications, the estimated 95% confidence interval contains the true parameter value, is calculated. In general, the coverage probability of λ is better with higher values of τ, because the estimator of parameter λ performs better for higher values of parameter τ. When $\tau < 1$, i.e. the skewness of the sample is high (> 2), the estimator based on expected FIM (5) performs better than the estimator based on sample FIM (4) for all censoring schemes, especially for small sample sizes

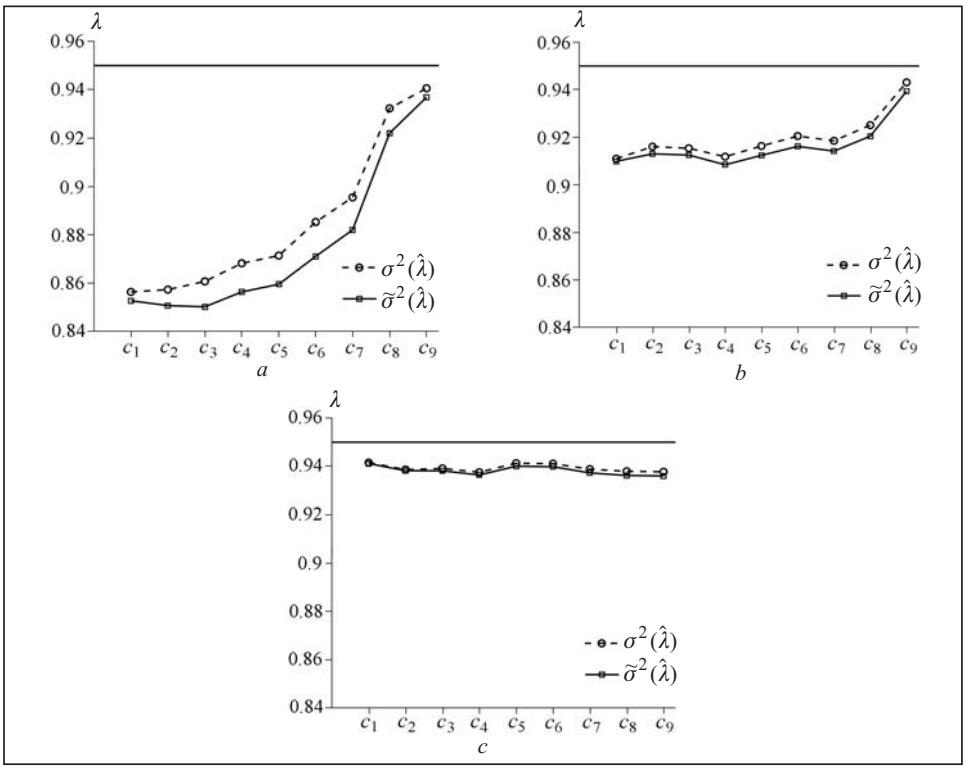

Fig. 3. Coverage probabilities for parameter $\hat{\lambda}$ considering various estimates of variance and double censoring; $n = 10$ (*a*), $n = 30$ (*b*), $n = 100$ (*c*); $\tau = 0.5$

Table 5

Censoring	Coverage probabilities (CP) for parameter $\hat{\lambda}$ based on expected FIM (5) considering single (S), double (D) and triple (T) censoring and sample size n; $\tau = 1.5$								
	$n = 10$, S			$n = 10$, D			$n = 10$, T		
	LCL	UCL	CP	LCL	UCL	CP	LCL	UCL	CP
c_1	0.593	1.398	0.888	0.596	1.398	0.887	0.592	1.403	0.890
c_3	0.580	1.406	0.896	0.589	1.402	0.894	0.588	1.406	0.898
c_5	0.562	1.428	0.909	0.580	1.408	0.898	0.584	1.407	0.901
c_7	0.549	1.515	0.912	0.572	1.432	0.918	0.581	1.425	0.923
c_9	0.729	1.926	0.599	0.554	1.588	0.942	0.576	1.565	0.955
Censoring	$n = 30$, S			$n = 30$, D			$n = 30$, T		
	LCL	UCL	CP	LCL	UCL	CP	LCL	UCL	CP
c_1	0.754	1.245	0.930	0.755	1.245	0.931	0.752	1.249	0.935
c_3	0.745	1.248	0.935	0.750	1.245	0.932	0.750	1.246	0.934
c_5	0.730	1.266	0.939	0.747	1.250	0.933	0.751	1.246	0.933
c_7	0.685	1.322	0.940	0.741	1.257	0.935	0.749	1.250	0.934
c_9	0.570	1.631	0.791	0.725	1.278	0.949	0.739	1.266	0.947
Censoring	$n = 100$, S			$n = 100$, D			$n = 100$, T		
	LCL	UCL	CP	LCL	UCL	CP	LCL	UCL	CP
c_1	0.863	1.137	0.944	0.863	1.137	0.944	0.861	1.139	0.947
c_3	0.859	1.140	0.947	0.862	1.137	0.945	0.861	1.138	0.947
c_5	0.850	1.149	0.947	0.860	1.140	0.943	0.862	1.137	0.942
c_7	0.819	1.182	0.948	0.856	1.143	0.946	0.860	1.139	0.944
c_9	0.671	1.371	0.899	0.846	1.151	0.944	0.854	1.145	0.949

Table 6

Censoring	Coverage probabilities (CP) for parameter $\hat{\tau}$ based on expected FIM (5) considering single (S), double (D) and triple (T) censoring and sample size n; $\tau = 1.5$								
	$n = 10$, S			$n = 10$, D			$n = 10$, T		
	LCL	UCL	CP	LCL	UCL	CP	LCL	UCL	CP
c_1	0.876	2.620	0.942	0.889	2.605	0.942	0.861	2.637	0.949
c_3	0.813	2.751	0.937	0.870	2.650	0.939	0.855	2.660	0.957
c_5	0.688	3.233	0.933	0.831	2.706	0.943	0.835	2.676	0.960
c_7	0.384	5.137	0.937	0.760	2.789	0.950	0.772	2.719	0.967
c_9	0.000	13.689	0.967	0.605	2.642	0.961	0.632	2.577	0.976
Censoring	$n = 30$, S			$n = 30$, D			$n = 30$, T		
	LCL	UCL	CP	LCL	UCL	CP	LCL	UCL	CP
c_1	1.119	2.029	0.946	1.126	2.020	0.946	1.111	2.036	0.954
c_3	1.075	2.085	0.946	1.109	2.038	0.947	1.110	2.036	0.954
c_5	1.003	2.200	0.944	1.087	2.069	0.948	1.100	2.049	0.953
c_7	0.864	2.483	0.938	1.058	2.114	0.945	1.071	2.085	0.954
c_9	0.391	5.290	0.953	0.996	2.166	0.958	1.014	2.121	0.971
Censoring	$n = 100$, S			$n = 100$, D			$n = 100$, T		
	LCL	UCL	CP	LCL	UCL	CP	LCL	UCL	CP
c_1	1.280	1.763	0.950	1.285	1.758	0.951	1.277	1.766	0.958
c_3	1.254	1.790	0.949	1.274	1.768	0.951	1.275	1.765	0.956
c_5	1.212	1.846	0.948	1.261	1.784	0.950	1.270	1.773	0.954
c_7	1.127	1.963	0.947	1.244	1.805	0.948	1.253	1.792	0.957
c_9	0.861	2.467	0.942	1.223	1.836	0.947	1.231	1.820	0.964

(see Fig. 3 for double censoring and $\tau = 0.5$). For $\tau > 1$, coverage probabilities of both estimators almost coincide. The results showed that the coverage probability of λ are very similar considering double and triple censoring for all sample sizes and various values of τ (see Table 5 for $\tau = 1.5$; the average lower (LCL) and upper (UCL) confidence limits are included). The coverage probability gets higher with a higher censoring scheme for small sample sizes ($n < 50$). In case of single censoring, the behavior is similar until censoring scheme c_7.

The coverage probability of τ is similar considering double and triple censoring for all values of τ (see Table 6 for $\tau = 1.5$; the average lower (LCL) and upper (UCL) confidence limits are included), because all the estimates of τ are similar bias-wise. The coverage probabilities are quite close to the prescribed significance level for both estimators (based on the expected and the sample FIM) and practically coincide for $n > 50$.

4. REAL DATA EXAMPLE

Statistical methods derived in previous sections were used in the analysis of the worldwide commonly used synthetic musk compounds (e.g. galaxolide, musk ketone, musk xylene, etc.) present in fish muscle (see [28] for more details). Here we show one example, specifically modeling of galaxolide concentration using doubly left-censored Weibull distribution (see Fig. 4). The real sample consists of 30 fish from the carp family *(Leuciscus cephalus*). Fish tissue samples were analyzed, and, among others, polycyclic musk compound called galaxolide was explored with $N_0 = 4$, $N_1 = 3$, $N_2 = 23$, $d_1 = 8.9488\ \mu\text{g/kg}$, $d_2 = 29.8294\ \mu\text{g/kg}$ and $X_{(27)} = 30.3630\ \mu\text{g/kg}$, $X_{(28)} = 39.2597\ \mu\text{g/kg}$, $X_{(29)} = 48.9161\ \mu\text{g/kg}$, $X_{(30)} = 79.7756\ \mu\text{g/kg}$. The level of censoring is high, and approximately corre-

sponds to censoring scheme c_9. The unknown parameters of the Weibull distribution were estimated, and particular 95% confidence intervals were calculated using estimators of variance based on the expected and the sample FIM. Specifically, the ML estimate $\hat{\lambda} = 23.98$ ($\hat{\tau} = 1.61$ respectively) with confidence interval $\lambda \in \langle 17.91, 30.05 \rangle$ ($\tau \in \langle 1.10, 2.12 \rangle$ respectively) with estimate of variance based on the expected FIM, and $\lambda \in \langle 17.82, 30.13 \rangle$ ($\tau \in \langle 1.17, 2.06 \rangle$ respectively) with estimate of variance based on the sample FIM.

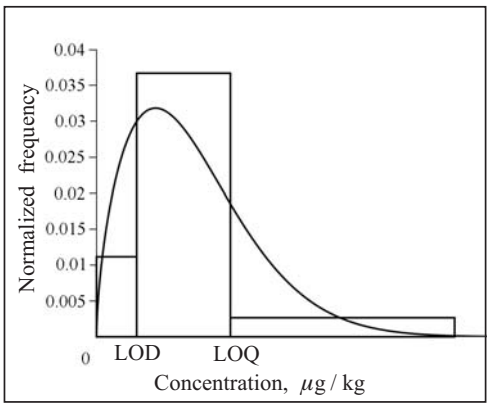

Fig. 4. Histogram of galaxolide concentration with Weibull density

From the practical point of view, it would be more interesting to have confidence limits for the mean concentration. In order to obtain them, results from the previous section can be utilized, and the delta method can be used [28, 33].

CONCLUSIONS

This paper dealt with type I multiply left-censored Weibull distribution. It was described how to estimate parameters of censored Weibull distribution using the method of maximum likelihood, and the expected FIM was analytically determined. Moreover, simulation results showed what bias of estimators $\hat{\lambda}$, $\hat{\tau}$ and $\sigma^2(\hat{\lambda})$, $\sigma^2(\hat{\tau})$ can be expected considering various sample sizes, various censoring schemes, and degree of censoring. Furthermore, it is shown what change in the variance estimate can be expected when the sample FIM is used instead of the expected FIM. Using different variance estimators, the lower and the upper confidence limits of the parameters estimates were obtained and the coverage probabilities of 95% confidence intervals were calculated. It was shown that the coverage probability of λ is better with higher values of τ, and when $\tau < 1$, i.e. the skewness of the sample is high (> 2), the estimator based on the expected FIM performs better than the estimator based on the sample FIM for all censoring schemes, especially for small sample sizes.

The problem of measured values found below the detection limits is common in many application areas. The methods derived in this paper can be favorably used instead of various ad hoc methods (e.g. replacing censored values with a constant) whenever dealing with type I multiply left-censored data from Weibull distribution. All the procedures used were implemented in the Matlab environment (version R2015a), and are available from the first author upon request.

APPENDIX 1. DERIVATION OF THE EXPECTED FIM

The expected FIM is calculated using formula

$$\boldsymbol{J}_n = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} = \begin{bmatrix} -\mathrm{E}\dfrac{\partial^2 l}{\partial \lambda^2} & -\mathrm{E}\dfrac{\partial^2 l}{\partial \lambda \partial \tau} \\ -\mathrm{E}\dfrac{\partial^2 l}{\partial \tau \partial \lambda} & -\mathrm{E}\dfrac{\partial^2 l}{\partial \tau^2} \end{bmatrix},$$

where

$$J_{11} = -\sum_{i=1}^{k} H_i^{\lambda\lambda}(\lambda,\tau)\mathrm{E}(N_i) - \frac{\mathrm{E}(N_0)\tau}{\lambda^2} + \frac{\tau^2+\tau}{\lambda^{\tau+2}}\mathrm{E}\left(\sum_{i=n-N_0+1}^{n} X_{(i)}^{\tau}\right),$$

$$J_{22} = -\sum_{i=1}^{k} H_i^{\tau\tau}(\lambda,\tau)\mathrm{E}(N_i) + \frac{\mathrm{E}(N_0)}{\tau^2} + \frac{(\ln\lambda)^2}{\lambda^\tau}\mathrm{E}\left(\sum_{i=n-N_0+1}^{n} X_{(i)}^\tau\right) -$$

$$-\frac{2\ln\lambda}{\lambda^\tau}\mathrm{E}\left(\sum_{i=n-N_0+1}^{n} X_{(i)}^\tau \ln X_{(i)}\right) + \frac{1}{\lambda^\tau}\mathrm{E}\left[\sum_{i=n-N_0+1}^{n} X_{(i)}^\tau (\ln X_{(i)})^2\right], \qquad (6)$$

$$J_{12} = J_{21} = -\sum_{i=1}^{k} H_i^{\lambda\tau}(\lambda,\tau)\mathrm{E}(N_i) + \frac{\mathrm{E}(N_0)}{\lambda} + \frac{\tau\ln\lambda - 1}{\lambda^{\tau+1}}\mathrm{E}\left(\sum_{i=n-N_0+1}^{n} X_{(i)}^\tau\right) -$$

$$-\frac{\tau}{\lambda^{\tau+1}}\mathrm{E}\left(\sum_{i=n-N_0+1}^{n} X_{(i)}^\tau \ln X_{(i)}\right),$$

and

$$H_i^{\lambda\lambda}(\lambda,\tau) = \frac{(d_i^\tau\lambda^\tau\tau^2 + d_i^\tau\lambda^\tau\tau - d_i^{2\tau}\tau^2)\exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]}{\lambda^{2\tau+2}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - \exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}} -$$

$$\frac{(d_{i-1}^\tau\lambda^\tau\tau^2 + d_{i-1}^\tau\lambda^\tau\tau - d_{i-1}^{2\tau}\tau^2)\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right]}{\lambda^{2\tau+2}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - \exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}} -$$

$$-\frac{\tau^2\left\{d_{i-1}^\tau\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - d_i^\tau\exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}^2}{\lambda^{2\tau+2}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - \exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}^2},$$

$$H_i^{\tau\tau}(\lambda,\tau) = \frac{(d_i^\tau\lambda^\tau - d_i^{2\tau})\left[\ln\left(\frac{d_i}{\lambda}\right)\right]^2\exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]}{\lambda^{2\tau}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - \exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}} - \frac{(d_{i-1}^\tau\lambda^\tau - d_{i-1}^{2\tau})\left[\ln\left(\frac{d_{i-1}}{\lambda}\right)\right]^2\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right]}{\lambda^{2\tau}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - \exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}} -$$

$$-\frac{\left\{d_{i-1}^\tau\ln\left(\frac{d_{i-1}}{\lambda}\right)\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - d_i^\tau\ln\left(\frac{d_i}{\lambda}\right)\exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}^2}{\lambda^{2\tau}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - \exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}^2},$$

$$H_i^{\lambda\tau}(\lambda,\tau) = \frac{\left[d_i^{2\tau}\tau\ln\left(\frac{d_i}{\lambda}\right) - d_i^\tau\lambda^\tau\tau\ln\left(\frac{d_i}{\lambda}\right) - d_i^\tau\lambda^\tau\right]\exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]}{\lambda^{2\tau+1}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - \exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}} -$$

$$-\frac{\left[d_{i-1}^{2\tau}\tau\ln\left(\frac{d_{i-1}}{\lambda}\right) - d_{i-1}^\tau\lambda^\tau\tau\ln\left(\frac{d_{i-1}}{\lambda}\right) - d_{i-1}^\tau\lambda^\tau\right]\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right]}{\lambda^{2\tau+1}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^\tau\right] - \exp\left[-\left(\frac{d_i}{\lambda}\right)^\tau\right]\right\}} +$$

$$+\frac{\tau\left\{d_{i-1}^{\tau}\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^{\tau}\right]-d_i^{\tau}\exp\left[-\left(\frac{d_i}{\lambda}\right)^{\tau}\right]\right\}}{\lambda^{2\tau+1}\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^{\tau}\right]-\exp\left[-\left(\frac{d_i}{\lambda}\right)^{\tau}\right]\right\}^2}\times$$

$$\times\left\{d_{i-1}^{\tau}\ln\left(\frac{d_{i-1}}{\lambda}\right)\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^{\tau}\right]-d_i^{\tau}\ln\left(\frac{d_i}{\lambda}\right)\exp\left[-\left(\frac{d_i}{\lambda}\right)^{\tau}\right]\right\},$$

$i=1,\dots,k$. Since the frequencies N_i, $i=0,\dots,k$, have binomial distribution, their expectations are

$$\mathrm{E}(N_i)=n\theta_i=\begin{cases}n\left\{\exp\left[-\left(\frac{d_{i-1}}{\lambda}\right)^{\tau}\right]-\exp\left[-\left(\frac{d_i}{\lambda}\right)^{\tau}\right]\right\} & \text{for } i=1,\dots,k,\\ n\exp\left[-\left(\frac{d_k}{\lambda}\right)^{\tau}\right] & \text{for } i=0.\end{cases} \tag{7}$$

Furthermore, it remains to derive the expectations

$$E_1=\mathrm{E}\left(\sum_{i=n-N_0+1}^{n}X_{(i)}^{\tau}\right),\ E_2=\mathrm{E}\left(\sum_{i=n-N_0+1}^{n}X_{(i)}^{\tau}\ln X_{(i)}\right),\ E_3=\mathrm{E}\left[\sum_{i=n-N_0+1}^{n}X_{(i)}^{\tau}(\ln X_{(i)})^2\right], \tag{8}$$

where $X_{(i)}$ is the i-th ordered statistic.

Using properties of the conditional expectation $\mathrm{E}(\mathrm{E}(X\mid Y))=\mathrm{E}(X)$ (X,Y are random variables), and due to the fact that $N_0\sim\mathrm{Bi}(n,\theta_0)$, one gets

$$E_1=\mathrm{E}\left(\sum_{i=n-n_0+1}^{n}X_{(i)}^{\tau}\right)=\mathrm{E}\left[\mathrm{E}\left(\sum_{i=n-N_0+1}^{n}X_{(i)}^{\tau}|N_0\right)\right]=$$

$$=\sum_{n_0=0}^{n}\mathrm{E}\left(\sum_{i=n-n_0+1}^{n}X_{(i)}^{\tau}|n_0\right)\mathrm{P}(N_0=n_0)=\sum_{n_0=0}^{n}\left[\sum_{i=n-n_0+1}^{n}\mathrm{E}(X_{(i)}^{\tau})\right]\binom{n}{n_0}\theta_0^{n_0}(1-\theta_0)^{n-n_0}=$$

$$=\sum_{n_0=0}^{n}\left[\sum_{i=n-n_0+1}^{n}\mathrm{E}(X_{(i)}^{\tau})\right]\binom{n}{n_0}\exp\left[-n_0\left(\frac{d_k}{\lambda}\right)^{\tau}\right]\left\{1-\exp\left[-\left(\frac{d_k}{\lambda}\right)^{\tau}\right]\right\}^{n-n_0}, \tag{9}$$

and analogically

$$E_2=\mathrm{E}\left(\sum_{i=n-n_0+1}^{n}X_{(i)}^{\tau}\ln X_{(i)}\right)=\sum_{n_0=0}^{n}\left[\sum_{i=n-n_0+1}^{n}\mathrm{E}(X_{(i)}^{\tau}\ln X_{(i)})\right]\times$$

$$\times\binom{n}{n_0}\exp\left[-n_0\left(\frac{d_k}{\lambda}\right)^{\tau}\right]\left\{1-\exp\left[-\left(\frac{d_k}{\lambda}\right)^{\tau}\right]\right\}^{n-n_0}, \tag{10}$$

$$E_3=\mathrm{E}\left[\sum_{i=n-n_0+1}^{n}X_{(i)}^{\tau}(\ln X_{(i)})^2\right]=\sum_{n_0=0}^{n}\left\{\sum_{i=n-n_0+1}^{n}\mathrm{E}[X_{(i)}^{\tau}(\ln X_{(i)})^2]\right\}\times$$

$$\times\binom{n}{n_0}\exp\left[-n_0\left(\frac{d_k}{\lambda}\right)^{\tau}\right]\left\{1-\exp\left[-\left(\frac{d_k}{\lambda}\right)^{\tau}\right]\right\}^{n-n_0}. \tag{11}$$

The pdf of the variable $X_{(i)}$ is in the form of (see [34])

$$f_{(i)}(x) = n\binom{n-1}{i-1} f(x)[F(x)]^{i-1}[1-F(x)]^{n-i},\ \ i = 1, 2, \dots, n,$$

where F is cdf (1) and f is pdf (2) of the uncensored Weibull distribution. Gradually, one gets pdf of $f_{(i)}(x)$ in the form of

$$f_{(i)}(x) = n\binom{n-1}{i-1}\frac{\tau}{\lambda^\tau}x^{\tau-1}\sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}\exp\left[-\left(\frac{x}{\lambda}\right)^\tau(n-i+j+1)\right].$$

Using $\int\limits_0^\infty t\mathrm{e}^{-t}dt = \Gamma(2) = 1$, $\int\limits_0^\infty t\ln t\mathrm{e}^{-t}dt = \Gamma'(2) = 1-\gamma_e$ and $\int\limits_0^\infty t(\ln t)^2\mathrm{e}^{-t}dt = \Gamma''(2) = \frac{\pi^2}{6} - 2\gamma_e + \gamma_e^2$ (see [35]), the expectations (8) are obtained in the form of

$$\mathrm{E}(X_{(i)}^\tau) = n\binom{n-1}{i-1}\lambda^\tau\sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}(n-i+j+1)^{-2}, \tag{12}$$

$$\mathrm{E}(X_{(i)}^\tau\ln X_{(i)}) = n\binom{n-1}{i-1}\frac{\lambda^\tau}{\tau}\sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}(n-i+j+1)^{-2}\left[\ln\left(\frac{\lambda^\tau}{n-i+j+1}\right)+1-\gamma_e\right], \tag{13}$$

$$\mathrm{E}[X_{(i)}^\tau(\ln X_{(i)})^2] = n\binom{n-1}{i-1}\frac{\lambda^\tau}{\tau^2}\sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}(n-i+j+1)^{-2}\times$$

$$\times\left\{\left[\ln\left(\frac{\lambda^\tau}{n-i+j+1}\right)\right]^2 + 2\ln\left(\frac{\lambda^\tau}{n-i+j+1}\right)(1-\gamma_e) + \frac{\pi^2}{6} - 2\gamma_e + \gamma_e^2\right\}, \tag{14}$$

where $\gamma_e \doteq 0.57722$ is Euler's constant. Substituting (12)–(14) into (9)–(11), one gets

$$E_1 = n\lambda^\tau\sum_{n_0=0}^{n}\sum_{i=n-n_0+1}^{n}\binom{n-1}{i-1}\sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}(n-i+j+1)^{-2}\binom{n}{n_0}\times$$

$$\times\exp\left[-n_0\left(\frac{d_k}{\lambda}\right)^\tau\right]\left\{1-\exp\left[-\left(\frac{d_k}{\lambda}\right)^\tau\right]\right\}^{n-n_0}, \tag{15}$$

$$E_2 = n\frac{\lambda^\tau}{\tau}\sum_{n_0=0}^{n}\sum_{i=n-n_0+1}^{n}\binom{n-1}{i-1}\sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}(n-i+j+1)^{-2}\binom{n}{n_0}\exp\left[-n_0\left(\frac{d_k}{\lambda}\right)^\tau\right]\times$$

$$\times\left\{1-\exp\left[-\left(\frac{d_k}{\lambda}\right)^\tau\right]\right\}^{n-n_0}\left[\ln\left(\frac{\lambda^\tau}{n-i+j+1}\right)+1-\gamma_e\right], \tag{16}$$

$$E_3 = n\frac{\lambda^\tau}{\tau^2}\sum_{n_0=0}^{n}\sum_{i=n-n_0+1}^{n}\binom{n-1}{i-1}\sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}(n-i+j+1)^{-2}\binom{n}{n_0}\times$$

$$\times\exp\left[-n_0\left(\frac{d_k}{\lambda}\right)^\tau\right]\left\{1-\exp\left[-\left(\frac{d_k}{\lambda}\right)^\tau\right]\right\}^{n-n_0}\times$$

$$\times\left\{\left[\ln\left(\frac{\lambda^\tau}{n-i+j+1}\right)\right]^2 + 2\ln\left(\frac{\lambda^\tau}{n-i+j+1}\right)(1-\gamma_e) + \frac{\pi^2}{6} - 2\gamma_e + \gamma_e^2\right\}. \tag{17}$$

Moreover, substituting (7) and (15)–(17) into (6), the elements of the FIM (5) are obtained.

REFERENCES

1. Mitra S., Kundu D. Analysis of left censored data from the generalized exponential distribution. *J. Stat. Comput. Simul.* 2008. Vol. 78. P. 669–679. DOI: https://doi.org/10.1080/00949650701344158.
2. Aboueissa A.E.-M.A., Stoline M.R. Maximum likelihood estimators of population parameters from doubly left-censored samples. *Environmetrics.* 2006. Vol. 17. P. 811–826. DOI: https://doi.org/10.1002/env.795.
3. Aboueissa A.E.-M.A. Maximum likelihood estimators of population parameters from multiply censored samples. *Environmetrics*. 2009. Vol. 20. P. 312–330. DOI: https://doi.org/10.1002/env.931.
4. Cohen A.C. Truncated and censored samples. New York: Marcel Dekker, 1991. 328 p.
5. El-Shaarawi A.H., Naderi A. Statistical inference from multiply censored environmental data. *Environ. Monit. Assess*. 1991. Vol. 17. P. 339–347. DOI: https://doi.org/10.1007/BF00399313.
6. Fusek M., Michálek J. Statistical analysis of type I multiply left-censored samples from exponential distribution. *J. Stat. Comput. Simu*l. 2015. Vol. 85. P. 2148–2163. DOI: https://doi.org/10.1080/00949655.2014.923886.
7. Fusek M., Michálek J. Statistical methods for analyzing musk compounds concentration based on doubly left-censored samples. *Int. J. Math. Model. Method. Appl. Sci*. 2013. Vol. 7. P. 755–763.
8. Cox D.R., Oakes D. Analysis of survival data. New York: Chapman and Hall/CRC, 1984. 201 p.
9. Barndorff-Nielsen O.E., Cox D.R. Inference and asymptotics. London: Chapman and Hall/CRC, 1994. 360 p.
10. Lehmann E.L., Casella G. Theory of point estimation, second ed. New York: Springer-Verlag, 1998. 617 p.
11. Samosonok O.S. Analysis of empirical estimates obtained for Gibbs distribution parameters by the maximum likelihood method. *Cybernetics and Systems Analysis*. 2013. Vol. 49, N 2. P. 316–324. DOI: https://doi.org/10.1007/s10559-013-9514-3.
12. El-Shaarawi A.H., Dolan D.M. Maximum likelihood estimation of water concentrations from censored data. Can. J. Fish. Aquat. Sci. 1989. Vol. 46. P. 1033–1039. DOI: https://doi.org/10.1139/f89-134.
13. El-Shaarawi, A.H., Esterby, S.R. Replacement of censored observations by a constant: an evaluation. *Water Res.* 1992. Vol. 26. P. 835–844. DOI: https://doi.org/10.1016/0043-1354(92)90015-V.
14. Johnson N.L., Kotz S., Balakrishnan N. Continuous univariate distributions, Vol. 1, 2nd ed. New York: John Wiley and Sons, 1994. 784 p.
15. Banerjee A., Kundu D. Inference based on type-II hybrid censored data from a Weibull distribution. *IEEE Trans. Relia*b. 2008. Vol. 57. P. 369–378. DOI: https://doi.org/10.1109/TR.2008.916890
16. Kundu D. On hybrid censored Weibull distribution. *J. Stat. Plan. Inference.* 2007. Vol. 137. P. 2127–2142. DOI: https://doi.org/10.1016/j.jspi.2006.06.043.
17. Kundu D. Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. *Technometrics*. 2008. Vol. 50. P. 144–154. DOI: https://doi.org/10.1198/004017008000000217.
18. Shoari N., Dubé J.S., Chenouri S. Estimating the mean and standard deviation of environmental data with below detection limit observations: Considering highly skewed data and model misspecification. *Chemosphere*. 2015. Vol. 138. P. 599–608. DOI: https://doi.org/10.1016/ j.chemosphere.2015.07.009.
19. Joarder A., Krishna H., Kundu D. Inferences on Weibull parameters with conventional type-I censoring. *Comput. Stat. Data Anal*. 2011. Vol. 55. P. 1–11. DOI: https://doi.org/10.1016/j.csda.2010.04.006.
20. Gupta R.D., Kundu D. Generalized exponential distribution. *Aust. N. Z. J. Stat*. 1999. Vol. 41. P. 173–188. DOI: https://doi.org/10.1111/1467-842X.00072.
21. Gupta R.D., Kundu D. Generalized exponential distribution: Existing results and some recent developments. *J. Stat. Plan. Inference*. 2007. Vol. 137. P. 3537–3547. DOI: https://doi.org/10.1016/j.jspi.2007.03.030.
22. Zheng G. On the Fisher information matrix in type II censored data from the exponentiated exponential family. *Biom. J.* 2002. Vol. 44, Iss. 3. P. 353–357. DOI: https://doi.org/10.1002/1521-4036(200204)44:3<353::AID-BIMJ353>3.0.CO;2-7.
23. Gupta R.D., Kundu D. On the comparison of Fisher information of the Weibull and GE distributions. *J. Stat. Plan. Inference.* 2006. Vol. 136. DOI: P. 3130–3144. https://doi.org/10.1016/j.jspi. 2004.11.013.
24. Dube S., Pradhan B., Kundu D. Parameter estimation of the hybrid censored log-normal distribution. *J. Stat. Comput. Simul*. 2011. Vol. 81. DOI: P. 275–287. https://doi.org/10.1080/ 00949650903292650.
25. Chen D., Lu J.C., Li C.S., Park J. Parameter estimation for bivariate shock models with singular distribution for censored data with concomitant order statistics. *Aust. N. Z. J. Stat.* 2000. Vol. 42. P. 323–336. DOI: https://doi.org/10.1111/1467-842X.00129.
26. Helsel D.R. Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere*. 2006. Vol. 65. P. 2434–2439. DOI: https://doi.org/10.1016/j.chemosphere.2006.04.051.

27. Helsel D.R. Statistics for censored environmental data using Minitab® and R. New York: John Wiley and Sons, 2012. 344 p.
28. Fusek M., Michálek J., Vávrová M. Evaluation of contamination data with non-detects using censored distributions. *Fresenius Environ. Bull*. 2015. Vol. 24. P. 4165–4172.
29. Voronin A.N. A compromise method in constrained optimization problems. C*ybernetics and Systems Analysis*. 2013. Vol. 49, N 1. P. 77–80. DOI: https://doi.org/10.1007/s10559-013-9487-2.
30. Lagarias J.C., Reeds J.A., Wright, M.H., Wright, P.E. Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM J. Optim.* 1998. Vol. 9. P. 112–147. DOI: https://doi.org/10.1137/S1052623496303470.
31. Fahrmeier L., Tutz G. Multivariate statistical modelling based on generalized linear models, second ed. New York: Springer-Verlag, 2001. 518 p.
32. Efron B., Johnstone I.M. Fisher's information in terms of the hazard rate. *Ann. Stat*. 1990. Vol. 18. P. 38–62. DOI: https://doi.org/10.1214/aos/1176347492.
33. Fusek M., Michálek J. On the confidence intervals for mean of left-censored Weibull distribution. 22nd International Conference on Soft Computing MENDEL 2016, Brno, Czech Republic. 2016. P. 249–254.
34. Hogg R.V., McKean, J.W., Craig, A. Introduction to Mathematical Statistics, sixth ed. Upper Saddle River (NJ): Pearson Education, 2005. 694 p.
35. Fichtengoltz G.M. Course on the differential and integral calculus, Vol. 2 [in Russian]. Moscow: Fiz.-Math.-Giz., 1959.

Надійшла до редакції 08.11.2018

М. Фусек, Я. Міхалек

СТАТИСТИЧНЕ ВИВЕДЕННЯ ДЛЯ БАГАТОРАЗОВО ЦЕНЗУРОВАНОЇ ЗЛІВА ВИБІРКИ ТИПУ I ДЛЯ РОЗПОДІЛУ ВЕЙБУЛЛА

Анотація. У багатьох галузях науки часто зустрічаються задачі з цензурованими зліва даними з однією або кількома межами виявлення. У цій роботі запропоновано процедуру для обчислення оцінок максимальної правдоподібності параметрів багаторазового цензурування зліва типу I з розподілу Вейбулла з урахуванням різної кількості меж виявлення. Очікувану інформаційну матрицю Фішера визначено аналітично та її вигляд порівняно з вибірковою (спостережуваною) інформаційною матрицею Фішера. Моделювання здебільшого ґрунтується на властивостях оцінок вибірок малих розмірів. Приклади проілюстровано на реальних даних.

Ключові слова: багаторазово цензурована зліва вибірка, оцінка максимальної правдоподібності, розподіл Вейбулла, інформаційна матриця Фішера, цензурування типу I.

М. Фусек, Я. Михалек

СТАТИСТИЧЕСКИЙ ВЫВОД ДЛЯ МНОГОКРАТНО ЦЕНЗУРИРОВАННОЙ СЛЕВА ВЫБОРКИ ТИПА I ДЛЯ РАСПРЕДЕЛЕНИЯ ВЕЙБУЛЛА

Аннотация. Во многих областях науки часто встречаются задачи с цензурированными слева данными с одной или несколькими границами обнаружения. В данной работе предложена процедура для вычисления оценок максимальной правдоподобности параметров многократного цензурирования слева типа I для распределения Вейбулла с учетом разного числа границ обнаружения. Ожидаемая информационная матрица Фишера определена аналитически и ее вид сравнен с выборочной (наблюдаемой) информационной матрицей Фишера. Моделирование основано, главным образом, на свойствах оценок выборок малых размеров. Примеры проиллюстрированы на реальных данных.

Ключевые слова: многократно цензурированная слева выборка, оценка максимальной правдоподобности, распределение Вейбулла, информационная матрица Фишера, цензурирование типа I.

Michal Fusek,
Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic, e-mail: fusekmi@feec.vutbr.cz.

Jaroslav Michálek,
Department of Quantitative Methods, Faculty of Military Leadership, University of Defence, Brno, Czech Republic, e-mail michalek@fme.vutbr.cz.