Brazilian Journal of Probability and Statistics
2015, Vol. 29, No. 3, 677–694
DOI: 10.1214/14-BJPS239

Estimation of parameters in Laplace distributions with interval censored data

V. L. D. Tomazella[a] and S. Nadarajah[b]

[a]*Universidade Federal de São Carlos*
[b]*University of Manchester*

Abstract. We derive maximum likelihood estimators for the parameters of the Laplace distribution for interval censored data. Existence and uniqueness of the estimators are proved. Simulations and real data applications show that the Laplace distribution can be a better model for interval censored data than competing models in spite of being simpler.

1 Introduction

In standard survival analysis, lifetimes are usually censored by the end of the period of study (see, e.g., Lawless (2003)). In this situation, we have right-censored lifetime data. If data are either observed exactly or right-censored (the most common situation in clinical trial settings), there are many parametric, semi-parametric, non-parametric methods available for estimating survival curves.

In some situations, however, events of interest are only known to have occurred within an interval of time, say $[L, U]$. This, can occur in a clinical trial, when patients are assessed only at pre-scheduled visits. If the event has not occurred at one visit, say at time L, but has by the following visit, say at time U, then the event must belong to the interval $[L, U]$. These data are known as interval censored data.

Lindsey (1998) suggests various parametric models for interval censored data. There are also non-parametric models for interval censored data. But Lindsey (1998) argues: "In the context of heavily interval censored data, the conclusions from parametric models are remarkably robust with changing distributional assumptions and generally more informative than the corresponding non-parametric models." Similar arguments are expressed by Sparling et al. (2006): "Parametric regression models in the presence of heavily interval censored data are robust and are generally more informative than the corresponding non-parametric models." A robust statistic is one that "is resistant to errors in the results, produced by deviations from assumptions (e.g., of normality). This means that if the assumptions are only approximately met, the estimators of the robust model will still have a reasonable efficiency, and reasonably small bias, as well as being asymptotically unbiased, meaning having a bias tending toward 0 as the sample size tends toward infinity" (Wikipedia). In Lindsey (1998) and Sparling et al. (2006), "robust"

Key words and phrases. Existence, maximum likelihood estimation, uniqueness.
Received September 2013; accepted February 2014.

might have meant that different parametric assumptions lead to similar estimates for quantities of interest like the probability of survival beyond 90 years.

Two prominent parametric models suggested by Lindsey (1998) are the log-logistic and log-Laplace distributions. Maximum likelihood (ML) estimation for interval censored data for the former has been considered by Zhou et al. (2007). But we are not aware of any work with respect to the log-Laplace distribution. In fact, we are not aware of any paper discussing any kind of estimation for any kind of censored data for the log-Laplace distribution.

However, estimation for interval censored data has been considered for many other models based on the Weibull distribution (Scallan, 1999; Tse et al., 2008; Ng and Wang, 2009; Ding et al., 2010), exponential distribution (Tse et al., 2002), lognormal distribution (Amin, 2008; Lin et al., 2009), gamma distribution (Lu and Tsai, 2009), generalized exponential distribution (Chen and Lio, 2010), exponentiated Weibull distribution (Hashimoto et al., 2010), geometric distribution (Patel and Gajjar, 2010), generalized Rayleigh distribution (Lio et al., 2011), and the Gompertz–Makeham distribution (Teimouri and Gupta, 2012). For a most excellent account of the statistical analysis of interval censored data, we refer the readers to Sun (2006).

The Laplace and log-Laplace distributions are mathematically simpler than the logistic and log-logistic distributions, in spite of the former being not differentiable at their location parameters. The Laplace and log-Laplace distributions have received wide ranging applications: low doses in dose response curves (Uppuluri, 1981); generalized nonlinear models for pharmacokinetic data (Lindsey et al., 2000); models of financial returns (Hürlimann, 2001); frequency of high microbial counts in commercial food products (Corradini et al., 2001; Peleg, 2002); stock price models and the approximation of currency exchange data (Kozubowski and Podgórski, 2003a); to mention just a few. For a comprehensive account of known applications, we refer the readers to Kozubowski and Podgórski (2003b).

The aim of this paper is to derive ML estimation of the parameters of the Laplace distribution for interval censored data. The contents are organized as follows. In Section 2, existence and uniqueness of the ML estimators are proved for the parameters of the Laplace distribution with interval censored data. In Section 3, a simulation study is conducted to assess the performance of the ML estimators for interval censored data. Comparisons are made to the ML estimators of the parameters of the logistic distribution (Zhou et al., 2007). These comparisons establish superior performance of the ML estimators of the parameters of the Laplace distribution. In Section 4, a real data application in Zhou et al. (2007) is revisited. The Laplace model is shown to provide a better fit to the data. Some possible future works are discussed in Section 5.

2 ML estimation

Supposed $X_1, X_2, \ldots, X_n$ is a random sample grouped into $k+1$ intervals, $[\tau_{i-1}, \tau_i]$, $i = 1, 2, \ldots, k+1$, where $-\infty = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_{k+1} = \infty$ are predetermined constants. Let $\boldsymbol{\tau} = \{\tau_0, \tau_1, \tau_2, \ldots, \tau_{k+1}\}$ denote the set of partitioned points. Let n_i denote the number of observations falling into the interval $[\tau_{i-1}, \tau_i]$, $i = 1, 2, \ldots, k+1$. Then the log-likelihood function is

$$\log L(\mu, \sigma) = \text{const} + \sum_{i=1}^{k+1} n_i \log\bigl[F(\tau_i; \mu, \sigma) - F(\tau_{i-1}; \mu, \sigma)\bigr], \tag{2.1}$$

where $F(\cdot\,; \mu, \sigma)$ denotes the common cumulative distribution function (c.d.f.) of $X_1, X_2, \ldots, X_n$ parameterized by μ ($-\infty < \mu < \infty$), a location parameter, and σ ($\sigma > 0$), a scale parameter.

Consider the re-parameterization $\theta_1 = \mu/\sigma$ and $\theta_2 = 1/\sigma$. There is a one-to-one correspondence between (μ, σ) and (θ_1, θ_2) for $-\infty < \theta_1 < \infty$ and $\theta_2 > 0$. The log-likelihood function in (2.1) can be rewritten as

$$\log L(\theta_1, \theta_2) = \text{const} + \sum_{i=1}^{k+1} n_i \log\bigl[F_0(\theta_2\tau_i - \theta_1) - F_0(\theta_2\tau_{i-1} - \theta_1)\bigr]. \tag{2.2}$$

Here, $F_0(x) = F(x; 0, 1)$ denotes the standard form of $F(\cdot\,; \mu, \sigma)$.

Now take $F_0(\cdot)$ to be the standard Laplace c.d.f., that is,

$$F_0(y) = \begin{cases} 1 - \dfrac{1}{2}\exp(-y), & y \geq 0, \\ \dfrac{1}{2}\exp(y), & y < 0. \end{cases}$$

The corresponding probability density function (p.d.f.) is

$$f_0(y) = \frac{1}{2}\exp\bigl(-|y|\bigr). \tag{2.3}$$

Theorem 2.1 shows that the ML estimators of (θ_1, θ_2), obtained by maximizing (2.2), exist and are unique.

Theorem 2.1. *Suppose that the grouped data* $n_1, n_2, \ldots, n_{k+1}$ *satisfy* $n_1 + n_{k+1} < n$, $n_{j-1} + n_j < n$, $2 \leq j \leq k+1$. *Then the ML estimators of* (θ_1, θ_2), *and consequently the ML estimators of* (μ, σ), *exist and are unique.*

Proof. In order to show existence and uniqueness of the ML estimators of (θ_1, θ_2), it is sufficient to:

- verify that the Hessian matrix is semi-negative definite at every point $(\theta_1, \theta_2) \in (-\infty, \infty) \times (0, \infty)$ and that it is negative definite at least at the points corresponding to the ML estimators. This is true from Lemma 2.1, Lemma 2.2 and Lemma 2.3;

- verify that for any given $\eta > 0$, there exists a compact subset $K \equiv K(\eta) \subset (-\infty, \infty) \times (0, \infty)$ such that $\{(\theta_1, \theta_2)\colon \log L(\theta_1, \theta_2) \geq -\eta\} \subset K$. This is shown in Lemma 2.4.

The proof is complete. □

Lemma 2.1. *Let H_1 and H_{k+1} be defined by*

$$H_1(\theta_1, \theta_2) = \begin{bmatrix} \dfrac{\partial^2}{\partial \theta_1^2} \log F_0(\theta_2 \tau_1 - \theta_1) & \dfrac{\partial^2}{\partial \theta_1 \, \partial \theta_2} \log F_0(\theta_2 \tau_1 - \theta_1) \\ \dfrac{\partial^2}{\partial \theta_1 \, \partial \theta_2} \log F_0(\theta_2 \tau_1 - \theta_1) & \dfrac{\partial^2}{\partial \theta_2^2} \log F_0(\theta_2 \tau_1 - \theta_1) \end{bmatrix}$$

and

$$H_{k+1}(\theta_1, \theta_2) = \begin{bmatrix} \dfrac{\partial^2}{\partial \theta_1^2} \log \overline{F}_0(\theta_2 \tau_k - \theta_1) & \dfrac{\partial^2}{\partial \theta_1 \, \partial \theta_2} \log \overline{F}_0(\theta_2 \tau_k - \theta_1) \\ \dfrac{\partial^2}{\partial \theta_1 \, \partial \theta_2} \log \overline{F}_0(\theta_2 \tau_k - \theta_1) & \dfrac{\partial^2}{\partial \theta_2^2} \log \overline{F}_0(\theta_2 \tau_k - \theta_1) \end{bmatrix},$$

respectively, where $\overline{F}_0(x) = 1 - F_0(x)$. Then, they are semi-negative definite.

Proof. Let $g_1(x) = \log F_0(x)$. If $x \geq 0$ then $\mathrm{d}^2 g_1(x)/\mathrm{d}x^2 = -2\exp(x)(2\exp(x) - 1)^{-2} < 0$ and

$$H_1(\theta_1, \theta_2) = \begin{bmatrix} -\dfrac{2\exp(\theta_2 \tau_1 - \theta_1)}{[2\exp(\theta_2 \tau_1 - \theta_1) - 1]^2} & \dfrac{2\tau_1 \exp(\theta_2 \tau_1 - \theta_1)}{[2\exp(\theta_2 \tau_1 - \theta_1) - 1]^2} \\ \dfrac{2\tau_1 \exp(\theta_2 \tau_1 - \theta_1)}{[2\exp(\theta_2 \tau_1 - \theta_1) - 1]^2} & -\dfrac{2\tau_1^2 \exp(\theta_2 \tau_1 - \theta_1)}{[2\exp(\theta_2 \tau_1 - \theta_1) - 1]^2} \end{bmatrix},$$

where we have

$$H_1(\theta_1, \theta_2) = \left[\frac{\mathrm{d}^2 g_1(x)}{\mathrm{d}x^2}\right]_{x = \theta_2 \tau_1 - \theta_1} \cdot \begin{pmatrix} 1 & -\tau_1 \\ -\tau_1 & \tau_1^2 \end{pmatrix}.$$

If $x < 0$, then $\mathrm{d}^2 g_1(x)/\mathrm{d}x^2 = 0$ and H_1 is the zero matrix. In both cases, $\det(H_1(\theta_1, \theta_2)) = 0$ and trace $(H_1) \leq 0$, so H_1 is semi-negative definite. Similarly, H_{k+1} is also semi-negative definite. □

Lemma 2.2. *Let $v < u$ and $g(u, v) = \log[F_0(u) - F_0(v)]$. Then the Hessian*

$$H(u, v) = \begin{bmatrix} \dfrac{\partial^2 g(u, v)}{\partial u^2} & \dfrac{\partial^2 g(u, v)}{\partial u \, \partial v} \\ \dfrac{\partial^2 g(u, v)}{\partial u \, \partial v} & \dfrac{\partial^2 g(u, v)}{\partial v^2} \end{bmatrix}$$

is semi-negative definite.

Proof. In order to prove that $H(u, v)$ is semi-negative definite, it is sufficient to prove that the following two conditions are satisfied:

(a) $\partial^2 g(u, v)/\partial u^2 < 0$ and $\partial^2 g(u, v)/\partial v^2 < 0$;
(b) the determinant of $H(u, v)$ is non-negative.

Note that

$$
\begin{aligned}
g(u, v) &= \log\bigl[F_0(u) - F_0(v)\bigr] \\
&= \begin{cases}
\log\Bigl[\frac{1}{2}\exp(-v) - \frac{1}{2}\exp(-u)\Bigr], & u \geq 0, v \geq 0, \\
\log\Bigl[1 - \frac{1}{2}\exp(-u) - \frac{1}{2}\exp(v)\Bigr], & u \geq 0, v < 0, \\
\log\Bigl[\frac{1}{2}\exp(u) - \frac{1}{2}\exp(v)\Bigr], & u < 0, v < 0.
\end{cases}
\end{aligned}
$$

In the case $u \geq 0, v \geq 0$,

$$
H(u, v) = \begin{bmatrix}
-\dfrac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} & \dfrac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} \\
\dfrac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} & -\dfrac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2}
\end{bmatrix} = H_2(u, v)
$$

say. We see that $\partial^2 g(u, v)/\partial u^2 < 0$, $\partial^2 g(u, v)/\partial v^2 < 0$ and $\det(H_2(u, v)) = 0$. In the case $u \geq 0, v < 0$,

$$
\begin{aligned}
H(u, v) &= \begin{bmatrix}
\dfrac{\exp(u)[-2 + \exp(v)]}{\{1 + \exp(u)[-2 + \exp(v)]\}^2} & \dfrac{\exp(u+v)}{\{1 + \exp(u)[-2 + \exp(v)]\}^2} \\
\dfrac{\exp(u+v)}{\{1 + \exp(u)[-2 + \exp(v)]\}^2} & \dfrac{\exp(u+v)[1 - 2\exp(u)]}{\{1 + \exp(u)[-2 + \exp(v)]\}^2}
\end{bmatrix} \\
&= H_3(u, v)
\end{aligned}
$$

say. We see that $\partial^2 g(u, v)/\partial u^2 < 0$, $\partial^2 g(u, v)/\partial v^2 < 0$ and $\det(H_3(u, v)) > 0$. In the case $u < 0, v < 0$,

$$
H(u, v) = \begin{bmatrix}
-\dfrac{\exp(u+v)}{[\exp(v) - \exp(u)]^2} & \dfrac{\exp(u+v)}{[\exp(v) - \exp(u)]^2} \\
\dfrac{\exp(u+v)}{[\exp(v) - \exp(u)]^2} & -\dfrac{\exp(u+v)}{[\exp(v) - \exp(u)]^2}
\end{bmatrix} = H_4(u, v)
$$

say. We see that $\partial^2 g(u, v)/\partial u^2 < 0$, $\partial^2 g(u, v)/\partial v^2 < 0$ and $\det(H_4(u, v)) = 0$. The proof is complete. □

It is not difficult to see that $g(u, v)$ for $u \geq 0$, $v < 0$ is greater than $g(u, v)$ for $u \geq 0$, $v \geq 0$ and $g(u, v)$ for $u < 0$, $v < 0$. So, the Hessian matrix corresponding to the ML estimators will be negative definite.

Lemma 2.3. *Let* $g_i(\theta_1, \theta_2) = \log[F_0(\theta_2\tau_i - \theta_1) - F_0(\theta_2\tau_{i-1} - \theta_1)]$ *and let* $\mathcal{H}_i$ *be its associated Hessian matrix*, $2 \leq i \leq k$. *Then*, $\mathcal{H}_i$ *is semi-negative definite.*

Proof. Let $u = \theta_2\tau_i - \theta_1$ and $v = \theta_2\tau_{i-1} - \theta_1$. It is easy to see that $\mathcal{H}_i(u, v) \equiv A' H(u, v) A$, where

$$A = \begin{pmatrix} -1 & \tau_i \\ -1 & \tau_{i-1} \end{pmatrix}.$$

By Lemma 2.2, $\mathcal{H}_i$ is semi-negative definite. □

Lemma 2.4. *Assume that* $n_1 + n_{k+1} < n$, $n_{j-1} + n_j < n$, $2 \le j \le k+1$. *For any given* $\eta > 0$, *there exists a compact subset* $K \equiv K(\eta) \subset (-\infty, \infty) \times (0, \infty)$ *such that*

$$\{(\theta_1, \theta_2)\colon \log L(\theta_1, \theta_2) \ge -\eta\} \subset K.$$

Proof. This follows from Lemmas 2.5 and 2.6. □

Lemma 2.5. *Assume that* $n_1 + n_{k+1} < n$. *Then*

$$\lim_{\theta_2 \to 0^+} \sup_{-\infty < \theta_1 < \infty} \log L(\theta_1, \theta_2) = -\infty. \tag{2.4}$$

Proof. From the assumption $n_1 + n_{k+1} < n$, there exists an index $2 \le i \le k+1$ such that $n_i > 0$. From (2.3), we have that $f_0(t) \le 1/2 < 1$ for all $t \in (-\infty, \infty)$. So, for $\theta_2 \to 0^+$ and $\theta_1 \to \infty$, we have

$$\log\left[\int_{\theta_2\tau_{i-1}-\theta_1}^{\theta_2\tau_i-\theta_1} f_0(t)\,\mathrm{d}t\right] = \log\left[\frac{1}{2}\exp(\theta_2\tau_{i-1} - \theta_1) - \frac{1}{2}\exp(\theta_2\tau_i - \theta_1)\right]$$

and

$$\begin{aligned} \log L(\theta_1, \theta_2) &\le n_i \log \int_{\theta_2\tau_{i-1}-\theta_1}^{\theta_2\tau_i-\theta_1} f_0(t)\,\mathrm{d}t \\ &\le n_i \log\left[\frac{1}{2}\exp(\theta_2\tau_{i-1} - \theta_1) - \frac{1}{2}\exp(\theta_2\tau_i - \theta_1)\right]. \end{aligned}$$

So,

$$\begin{aligned} \sup_{\theta_1 \to \infty} \log L(\theta_1, \theta_2) &\le \sup_{\theta_1 \to \infty} n_i \log\left[\frac{1}{2}\exp(\theta_2\tau_{i-1} - \theta_1) - \frac{1}{2}\exp(\theta_2\tau_i - \theta_1)\right] \\ &\le \sup_{\theta_1 \to \infty} (-n_i\theta_1) + n_i \log\left[\frac{1}{2}\exp(\theta_2\tau_{i-1}) - \frac{1}{2}\exp(\theta_2\tau_i)\right] \end{aligned}$$

and similarly, for $\theta_2 \to 0^+$ and $\theta_1 \to -\infty$. Hence, (2.4) holds. □

Lemma 2.6. *Assume that* $n_1 + n_{k+1} < n$ *and* $n_{j-1} + n_j < n$ *for all* $2 \le j \le k+1$. *Then*

$$\lim_{\theta_2 \to \infty} \sup_{-\infty < \theta_1 < \infty} \log L(\theta_1, \theta_2) = -\infty \tag{2.5}$$

and

$$\lim_{|\theta_1|\to\infty} \sup_{\theta_2>0} \log L(\theta_1, \theta_2) = -\infty. \tag{2.6}$$

Proof. Let $I = \{1 \le j \le k+1\colon\ n_j > 0\}$. For every fixed $\theta_2 > 0$, it is evident that

$$\log L(\theta_1, \theta_2) \le n_i \log \int_{\theta_2\tau_{i-1}-\theta_1}^{\theta_2\tau_i-\theta_1} f_0(t)\,\mathrm{d}t \equiv M(\theta_1).$$

So, $\sup_{-\infty<\theta_1<\infty} \log L(\theta_1, \theta_2) = \sup_{-\infty<\theta_1<\infty} M(\theta_1) = -\infty$ by Lemma 2.5. As a result, for every fixed θ_2, there is a $\theta_1^* = \theta_1^*(\theta_2)$ such that

$$\sup_{-\infty<\theta_1<\infty} \log L(\theta_1, \theta_2) = \log L\bigl(\theta_1^*, \theta_2\bigr) = \sum_{i\in I} n_i \log \int_{\theta_2\tau_{i-1}-\theta_1^*}^{\theta_2\tau_i-\theta_1^*} f_0(t)\,\mathrm{d}t.$$

Let $2C = \min_{1\le j\le k+1}(\tau_j - \tau_{j-1}) > 0$ and let $A > 0$ be given. For any $\theta_2 A$, there exists at least one index i in I such that

(a) $\theta_2\tau_{i-1} - \theta_1^* = \theta_2(\tau_{i-1} - \theta_1^*/\theta_2)$ and $\theta_2\tau_i - \theta_1^* = \theta_2(\tau_i - \theta_1^*/\theta_2)$ have the same sign;
(b) $|\tau_{i-1} - \theta_1^*/\theta_2| \ge C$ and $|\tau_i - \theta_1^*/\theta_2| \ge C$.

Note that if $i \in I$ with both $\theta_2\tau_{i-1} - \theta_1^* > 0$ and $\theta_2\tau_i - \theta_1^* > 0$ then one has

$$\int_{\theta_2\tau_{i-1}-\theta_1^*}^{\theta_2\tau_i-\theta_1^*} f_0(t)\,\mathrm{d}t = \left[\frac{1}{2}\exp(\theta_2\tau_i) - \theta_1^* - \frac{1}{2}\exp(\theta_2\tau_{i-1}) - \theta_1^*\right]$$

and

$$\int_{\theta_2\tau_{i-1}-\theta_1^*}^{\theta_2\tau_i-\theta_1^*} \exp(-t)\,\mathrm{d}t = \exp\bigl[-\bigl(\theta_2\tau_{i-1} - \theta_1^*\bigr)\bigr] - \exp\bigl[-\bigl(\theta_2\tau_i - \theta_1^*\bigr)\bigr].$$

So,

$$\int_{\theta_2\tau_{i-1}-\theta_1^*}^{\theta_2\tau_i-\theta_1^*} f_0(t)\,\mathrm{d}t < \int_{\theta_2\tau_{i-1}-\theta_1^*}^{\theta_2\tau_i-\theta_1^*} \exp(-t)\,\mathrm{d}t < \exp\bigl[-\bigl(\theta_2\tau_{i-1} - \theta_1^*\bigr)\bigr].$$

For any given $A > 0$ and for all $\theta_2 > A/C$, we have $-(\theta_2\tau_{i-1} - \theta_1^*) < -CA/C = -A$ and $n_i \log \int_{\theta_2\tau_{i-1}-\theta_1^*}^{\theta_2\tau_i-\theta_1^*} f_0(t)\,\mathrm{d}t < -n_i A \le A$. So,

$$\log L\bigl(\theta_1^*, \theta_2\bigr) = n_i \log \int_{\theta_2\tau_{i-1}-\theta_1^*}^{\theta_2\tau_i-\theta_1^*} f_0(t)\,\mathrm{d}t < -A \tag{2.7}$$

for all $\theta_2 > A/C$. The inequality (2.7) can be shown in a similar way in the cases of both $\theta_2\tau_{i-1} - \theta_1^* < -C$ and $\theta_2\tau_i - \theta_1^* < -C$. So, we see that for every given large number $A > 0$ and for all $\theta_2 > A/C$, $\sup_{-\infty<\theta_1<\infty} \log L(\theta_1, \theta_2) = \log L(\theta_1^*, \theta_2) < -A$. Due to the arbitrariness of A, we conclude that (2.5) holds.

By a similar approach, one can show that $\sup_{\theta_2>0} \log L(\theta_1, \theta_2) < -A$ for large $A > 0$ and large $|\theta_1|$. Again the arbitrariness of A implies (2.6) holds. $\square$

3 Simulation study

Here, we perform a simulation study to see how the ML estimators of μ and σ: (1) vary with respect to n; (2) vary with respect to different values of $\boldsymbol{\tau}$; (3) compare between the Laplace and logistic models. Our simulation study is more informative than the one in Zhou et al. (2007). We use the following scheme:

- simulate ten thousand random samples of size n from the Laplace distribution with $\mu = 0$ and $\sigma = 1$;
- for each of the ten thousand samples, compute the ML estimates of μ and σ for the Laplace model and the associated 95 percent confidence intervals based on normal approximation;
- compute the biases, the mean squared errors, the skewness, the kurtosis, the coverage probabilities and the coverage lengths for $\widehat{\mu}$ and $\widehat{\sigma}$ over the ten thousand replications;
- simulate ten thousand random samples of size n from the logistic distribution with $\mu = 0$ and $\sigma = \sqrt{6}/\pi$;
- for each of the ten thousand samples, compute the ML estimates of μ and σ for the logistic model and the associated 95 percent confidence intervals based on normal approximation;
- compute the biases, the mean squared errors, the skewness, the kurtosis, the coverage probabilities and the coverage lengths for $\widehat{\mu}$ and $\widehat{\sigma}$ over the ten thousand replications.

We took n to vary from $20, 21, \ldots, 100$. We took $\boldsymbol{\tau}$ to be one of $(-\infty, 0, 0.5, 1, 1.5, 2, \infty)$, $(-\infty, -0.5, 0.25, 1, 1.75, 2.5, \infty)$, $(-\infty, -0.25, 0, 1, 2, 2.5, \infty)$, or $(-\infty, -2, -1, 0.5, 1, 3, \infty)$. The parameters of the Laplace and logistic distributions were taken in such a way that they have equal means and equal variances. The coverage probabilities were computed as the proportion of the ten thousand confidence intervals containing the true parameter value. The coverage lengths were computed as the mean length of the ten thousand confidence intervals.

The ML estimates were obtained numerically by maximizing (2.2). The numerical maximization was performed by using optimize in the R statistical package (R Development Core Team, 2014). Our numerical calculations showed that the surface of (2.2) was smooth. The optimize was executed for a wide range of starting values. The solution was unique all the time, as suggested by the theory.

Plots of the biases, the mean squared errors, the skewness, the kurtosis, the coverage probabilities and the coverage lengths versus n for $\widehat{\mu}$ and $\widehat{\sigma}$ are shown in Figures 1–6. The actual values plotted are the lowess (Cleveland, 1979, 1981) smoothed versions versus n for $n = 20, 21, \ldots, 100$. While lowess smoothing, we used the default options. These are: a smoothing span of 2/3, three 'robustifying' iterations and the speed of computations determined by 0.01th of the range of the n values.

We can observe the following from Figures 1–6:

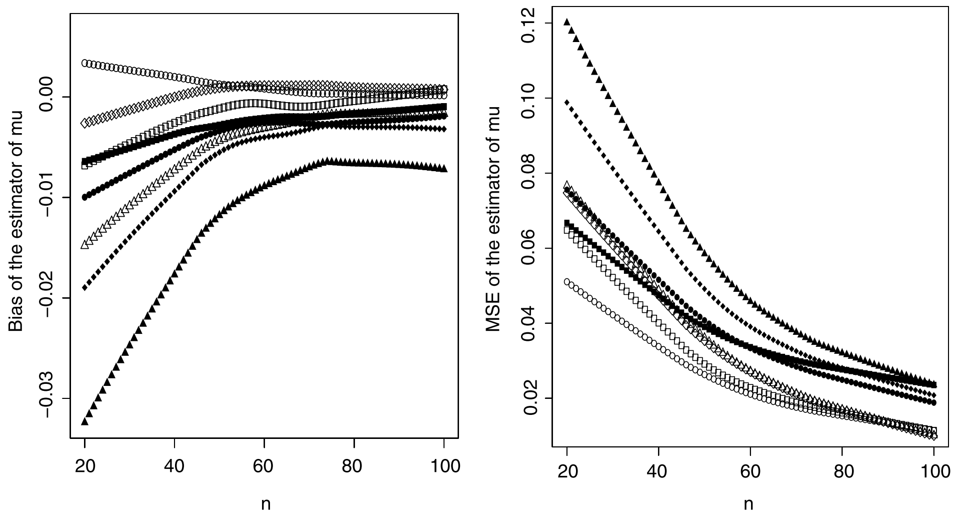

Figure 1 *Biases* (*left*) *and mean squared errors* (*right*) *of* $\widehat{\mu}$ *versus* n *for* $\boldsymbol{\tau} = (-\infty, 0, 0.5, 1, 1.5, 2, \infty)$ (*square*), $\boldsymbol{\tau} = (-\infty, -0.5, 0.25, 1, 1.75, 2.5, \infty)$ (*circle*), $\boldsymbol{\tau} = (-\infty, -0.25, 0, 1, 2, 2.5, \infty)$ (*triangle*), $\boldsymbol{\tau} = (-\infty, -2, -1, 0.5, 1, 3, \infty)$ (*diamond*), *the Laplace model* (*open symbol*) *and the logistic model* (*filled symbol*).

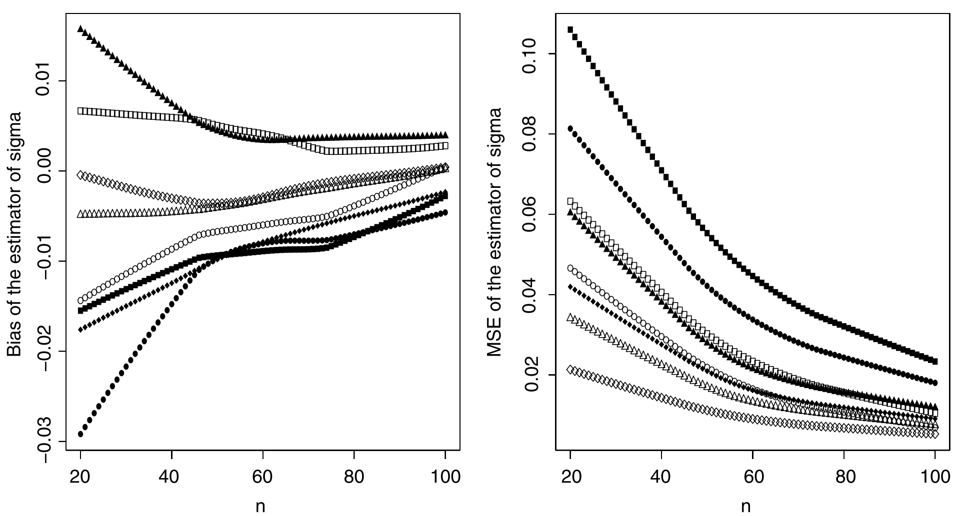

Figure 2 *Biases* (*left*) *and mean squared errors* (*right*) *of* $\widehat{\sigma}$ *versus* n *for* $\boldsymbol{\tau} = (-\infty, 0, 0.5, 1, 1.5, 2, \infty)$ (*square*), $\boldsymbol{\tau} = (-\infty, -0.5, 0.25, 1, 1.75, 2.5, \infty)$ (*circle*), $\boldsymbol{\tau} = (-\infty, -0.25, 0, 1, 2, 2.5, \infty)$ (*triangle*), $\boldsymbol{\tau} = (-\infty, -2, -1, 0.5, 1, 3, \infty)$ (*diamond*), *the Laplace model* (*open symbol*) *and the logistic model* (*filled symbol*).

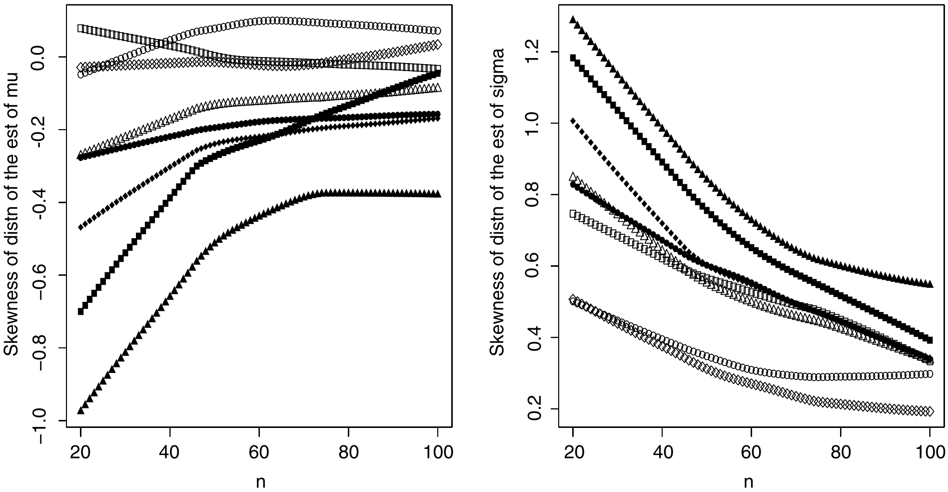

Figure 3 *Skewness of the distributions of $\widehat{\mu}$ (left) and $\widehat{\sigma}$ (right) versus n for $\boldsymbol{\tau} = (-\infty, 0, 0.5, 1, 1.5, 2, \infty)$ (square), $\boldsymbol{\tau} = (-\infty, -0.5, 0.25, 1, 1.75, 2.5, \infty)$ (circle), $\boldsymbol{\tau} = (-\infty, -0.25, 0, 1, 2, 2.5, \infty)$ (triangle), $\boldsymbol{\tau} = (-\infty, -2, -1, 0.5, 1, 3, \infty)$ (diamond), the Laplace model (open symbol) and the logistic model (filled symbol).*

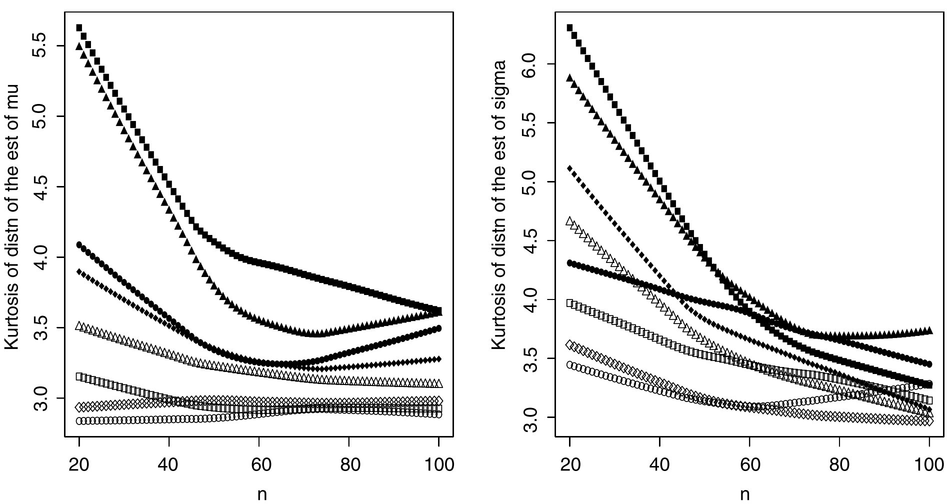

Figure 4 *Kurtosis of the distributions of $\widehat{\mu}$ (left) and $\widehat{\sigma}$ (right) versus n for $\boldsymbol{\tau} = (-\infty, 0, 0.5, 1, 1.5, 2, \infty)$ (square), $\boldsymbol{\tau} = (-\infty, -0.5, 0.25, 1, 1.75, 2.5, \infty)$ (circle), $\boldsymbol{\tau} = (-\infty, -0.25, 0, 1, 2, 2.5, \infty)$ (triangle), $\boldsymbol{\tau} = (-\infty, -2, -1, 0.5, 1, 3, \infty)$ (diamond), the Laplace model (open symbol) and the logistic model (filled symbol).*

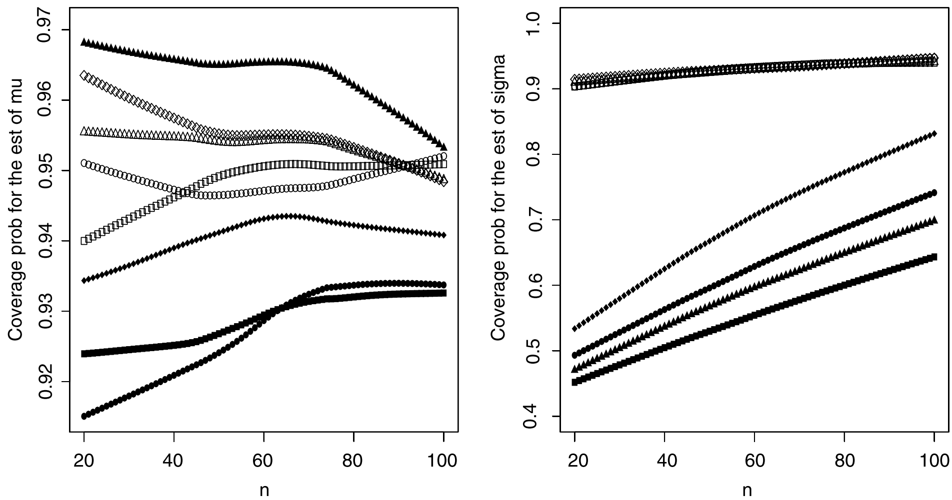

Figure 5 *Coverage probabilities for $\widehat{\mu}$ (left) and $\widehat{\sigma}$ (right) versus n for $\boldsymbol{\tau} = (-\infty, 0, 0.5, 1, 1.5, 2, \infty)$ (square), $\boldsymbol{\tau} = (-\infty, -0.5, 0.25, 1, 1.75, 2.5, \infty)$ (circle), $\boldsymbol{\tau} = (-\infty, -0.25, 0, 1, 2, 2.5, \infty)$ (triangle), $\boldsymbol{\tau} = (-\infty, -2, -1, 0.5, 1, 3, \infty)$ (diamond), the Laplace model (open symbol) and the logistic model (filled symbol).*

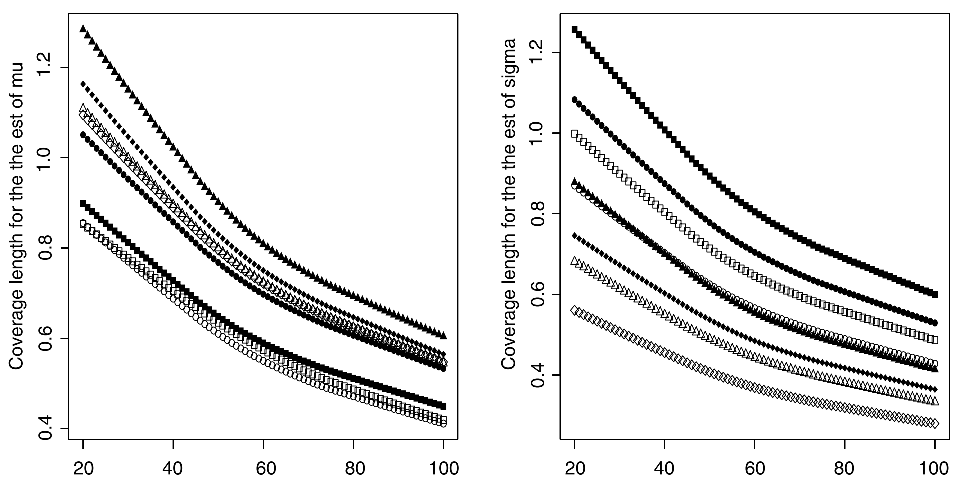

Figure 6 *Coverage lengths for $\widehat{\mu}$ (left) and $\widehat{\sigma}$ (right) versus n for $\boldsymbol{\tau} = (-\infty, 0, 0.5, 1, 1.5, 2, \infty)$ (square), $\boldsymbol{\tau} = (-\infty, -0.5, 0.25, 1, 1.75, 2.5, \infty)$ (circle), $\boldsymbol{\tau} = (-\infty, -0.25, 0, 1, 2, 2.5, \infty)$ (triangle), $\boldsymbol{\tau} = (-\infty, -2, -1, 0.5, 1, 3, \infty)$ (diamond), the Laplace model (open symbol) and the logistic model (filled symbol).*

- biases and mean squared errors of $\widehat{\mu}$ and $\widehat{\sigma}$ are generally smaller for the Laplace model;
- variation of the biases and mean squared errors of $\widehat{\mu}$ and $\widehat{\sigma}$ with respect to $\boldsymbol{\tau}$ is generally smaller for the Laplace model;
- biases are generally negative for both $\widehat{\mu}$ and $\widehat{\sigma}$ and for both the models;
- the magnitude of the biases generally decreases as n increases;
- mean squared errors generally decrease as n increases;
- skewness of the distributions of $\widehat{\mu}$ and $\widehat{\sigma}$ are generally closer to zero, the skewness for the normal distribution, for the Laplace model;
- skewness is generally negative for $\widehat{\mu}$;
- skewness is generally positive for $\widehat{\sigma}$;
- skewness generally approaches zero, the skewness for the normal distribution, as n increases;
- kurtosis of the distributions of $\widehat{\mu}$ and $\widehat{\sigma}$ are generally closer to three, the kurtosis for the normal distribution, for the Laplace model;
- kurtosis generally decreases to three, the kurtosis for the normal distribution, as n increases;
- coverage probabilities for $\widehat{\mu}$ and $\widehat{\sigma}$ are generally closer to the nominal level for the Laplace model;
- variation of the coverage probabilities for $\widehat{\mu}$ and $\widehat{\sigma}$ with respect to $\boldsymbol{\tau}$ is generally smaller for the Laplace model;
- coverage probabilities generally approach the nominal level as n increases;
- coverage lengths for $\widehat{\mu}$ and $\widehat{\sigma}$ are generally smaller for the Laplace model;
- coverage lengths generally decrease as n increases.

These observations suggest that the fitted distribution is closer to the *true* distribution of interval censored observations (as measured by biases, mean squared errors, coverage probabilities and coverage lengths) when the uncensored observations are from the Laplace distribution. The observations also suggest that the distribution of the ML estimators is closer to the normal distribution (as measured by skewness and kurtosis) when the uncensored observations are from the Laplace distribution.

Obviously, biases, mean squared errors, coverage probabilities and coverage lengths are not the ultimate measures of closeness. More objective measures are the chi-square statistics, Kolmogorov–Smirnov statistics and the Cramér–von Mises statistics. When the computations were repeated for these measures, the results were similar to those reported. The corresponding figures are not shown here for space concerns and to avoid repetitive discussion.

The figures presented in Zhou et al. (2007) only showed how the estimates, $\widehat{\mu}$ and $\widehat{\sigma}$, varied with respect to n and $\boldsymbol{\tau}$. Zhou et al. (2007) did not consider biases, mean squared errors, coverage probabilities, coverage lengths, skewness, kurtosis, chi-square statistics, Kolmogorov–Smirnov statistics or the Cramér–von

Mises statistics. Furthermore, Zhou et al. (2007) drew different figures for different $\boldsymbol{\tau}$. This did not allow for a fair comparison of the estimators with respect to different $\boldsymbol{\tau}$.

We also performed a simulation study comparing the estimators for the log-Laplace and log-logistic distributions. As in Zhou et al. (2007), we considered estimators of the mode of the p.d.f. and the mode of the failure rate function for each distribution. The results were similar to those reported.

4 Data application

Here, we consider the data set used in Section 4 of Zhou et al. (2007), a data set from a life test study of locomotive controls. The data set and its original reference can be found in Lawless (2003). According to Zhou et al. (2007), the data results from a study that "recorded the mileage at which 96 different locomotive controls failed. The test was terminated after 135,000 miles, by which time 37 failures had occurred. In addition, there are 59 censoring times, all equal to 135,000 miles".

Zhou et al. (2007) partitioned the data into twelve groups and recorded the number of observations falling within each group. The partitioned intervals and the observed frequencies are given in Table 1. In order to make the data suitable for models defined over the entire real line, a log-transformation was applied to the intervals. The logged intervals are shown in the third column of Table 1.

We fitted the Laplace and logistic models to the data in Table 1 by the method of ML. For the Laplace model, we obtained the estimates $\widehat{\mu} = 5.031$, $\widehat{\sigma} = 0.482$, $\mathrm{Var}(\widehat{\mu}) = 0.004$, $\mathrm{Var}(\widehat{\sigma}) = 0.007$, $\mathrm{Cov}(\widehat{\mu}, \widehat{\sigma}) = 0.002$ with $\log L = 151.5$. For the logistic model, we obtained the estimates $\widehat{\mu} = 5.080$, $\widehat{\sigma} = 0.380$, $\mathrm{Var}(\widehat{\mu}) = 0.008$,

Table 1 *Estimates of parameters and goodness of fit statistics*

No	Interval	log (interval)	Observed frequency	Expected frequency	
				Logistic model	Laplace model
1	<35	$(-\infty, 3.555]$	1	2.25453	1.713705
2	35–45	$(3.555, 3.807]$	1	1.546345	1.55515
3	45–55	$(3.807, 4.007]$	5	1.952273	2.133774
4	55–65	$(4.007, 4.174]$	1	2.379348	2.725425
5	65–75	$(4.174, 4.317]$	3	2.805545	3.268483
6	75–85	$(4.317, 4.443]$	8	3.264085	3.767965
7	85–95	$(4.443, 4.554]$	2	3.673676	4.10645
8	95–105	$(4.554, 4.654]$	1	4.116794	4.367169
9	105–115	$(4.654, 4.745]$	4	4.56486	4.514625
10	115–125	$(4.745, 4.828]$	7	4.985005	4.527251
11	125–135	$(4.828, 4.905]$	4	5.457734	4.492645
12	>135	$(4.905, \infty)$	59	58.9998	58.82736

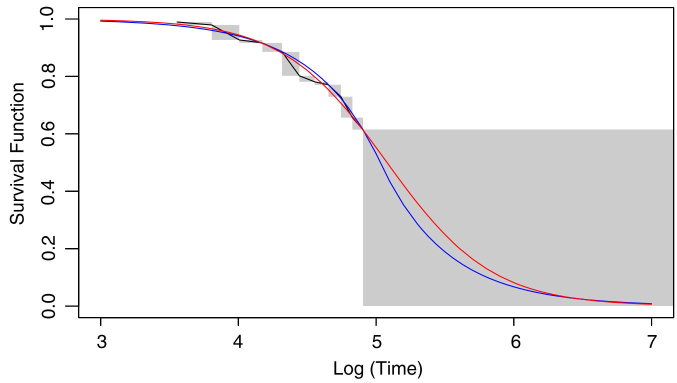

Figure 7 *Non-parametric estimate of the survival function as well as those obtained from the fitted Laplace (curve of dashes) and fitted logistic (curve of dots) models.*

$\mathrm{Var}(\widehat{\sigma}) = 0.003$, $\mathrm{Cov}(\widehat{\mu}, \widehat{\sigma}) = 0.003$ with $\log L = 150.9$. The expected frequencies computed using these estimates are shown in the last two columns of Table 1. The chi-squared statistics computed using the observed and expected frequencies are 6.413×10^{-10} and 5.066×10^{-4} for the Laplace and logistic models, respectively. The Kolmogorov–Smirnov statistics computed are 1.429×10^{-11} and 4.945×10^{-3} for the Laplace and logistic models, respectively. The Cramér–von Mises statistics computed are 7.246×10^{-9} and 1.083×10^{-5} for the Laplace and logistic models, respectively. The Kolmogorov–Smirnov and Cramér–von Mises statistics were computed by comparing the empirical and fitted c.d.f.s.

Hence, we see that the Laplace model is better in that it has smaller $\mathrm{Var}(\widehat{\mu})$, smaller $\mathrm{Cov}(\widehat{\mu}, \widehat{\sigma})$, larger log-likelihood, much smaller chi-square statistic, much smaller Kolmogorov–Smirnov statistic and much smaller Cramér–von Mises statistic. The parameter estimates are not too different between the two models.

Although the Laplace and logistic models are not nested, their likelihood values can be compared since they have the same number of parameters. In fact, criteria like the Akaike information criterion (Akaike, 1974), a commonly used criterion for comparing non-nested models, reduce to the standard likelihood ratio statistic for models having the same number of parameters.

As a final check, we have plotted the non-parametric ML estimate of the interval censored survival function in Figure 7. This estimate was computed using the function icfit in the R contributed package interval (Fay and Shaw, 2010; R Development Core Team, 2014). Superimposed in the figure are the estimates of the survival function from the fitted Laplace and logistic models. We can see that the Laplace model provides a better fit. As a quantitative evidence, Table 2 gives non-parametric estimates of the cell probabilities as well as those obtained from the fitted Laplace and logistic models. We can see that the sum of the absolute deviations between the second and third columns of Table 2 is 0.208. The sum of the

Table 2 *Estimates of cell probabilities*

	Cell probabilities		
log (interval)	Non-parametric	Logistic	Laplace
$(-\infty, 3.555)$	0.01041667	0.02348469	0.01785109
$(3.555, 3.807)$	0.01041667	0.01610777	0.01619948
$(3.807, 4.007)$	0.05208333	0.02033617	0.02222681
$(4.007, 4.174)$	0.01041667	0.02478488	0.02838984
$(4.174, 4.317)$	0.03125	0.02922442	0.0340467
$(4.317, 4.443)$	0.08333333	0.03400088	0.03924964
$(4.443, 4.554)$	0.02083333	0.03826746	0.04277552
$(4.554, 4.654)$	0.01041667	0.04288327	0.04549134
$(4.654, 4.745)$	0.04166667	0.04755063	0.04702734
$(4.745, 4.828)$	0.07291667	0.05192713	0.04715886
$(4.828, 4.905)$	0.04166667	0.0568514	0.04679839
$(4.905, \infty)$	0.6145833	0.6145813	0.612785

absolute deviations between the second and fourth columns of Table 2 is 0.203. This is the final evidence that the Laplace model provides a better fit.

For this particular data set, we see that the Laplace model gives a better fit. This does not mean that the Laplace model will always give a better fit for every real data set. Of course, there will be many data sets where the logistic model will provide a better fit. The message of Section 3 is that the parameter estimates are likely to have smaller biases, smaller mean squared errors, better coverage properties, skewness closer to zero and kurtosis closer to three if the best fitting model was the Laplace distribution. In other words, the parameter estimates are likely to have larger biases, larger mean squared errors, worse coverage properties, skewness more distant from zero and kurtosis more distant from three if the best fitting model was the logistic distribution.

5 Conclusions and future work

We have derived ML estimators of the parameters of the Laplace distribution by assuming an interval censored data. We have proved existence and uniqueness of the ML estimators. Simulations and real data applications have shown that the Laplace distribution can be a better model for interval censored data than the logistic distribution in terms of: biases of the ML estimates, variation of the biases of the ML estimates, mean squared errors of the ML estimates, variation of the mean squared errors of the ML estimates, skewness of the distribution of the ML estimates, kurtosis of the distribution of the ML estimates, coverage probabilities for the ML estimates, variation of the coverage probabilities for the ML estimates, coverage lengths for the ML estimates, variance of the ML estimates, covariance of the

ML estimates, log-likelihood values, chi-square statistics, Kolmogorov–Smirnov statistics, Cramér–von Mises statistics and the deviation between the fitted and non-parametric estimates of the interval censored survival function.

In Sections 3 and 4, we have assumed asymptotic normality of the ML estimators of μ and σ. For asymptotic normality, certain regularity conditions must be satisfied; see, for example, Ferguson (1996), page 121. However, these conditions are hardly checked in published papers (even in theoretical journals) and in practical data analysis. But it could be a possible future work.

Another future work is to extend the results of this paper for asymmetric Laplace distributions as well as for generalized Laplace distributions. A comprehensive account of known generalizations of the Laplace distribution is given in Kozubowski and Nadarajah (2010).

Acknowledgments

The authors would like to thank the Editor, the Associate Editor and the two referees for careful reading and for their comments which greatly improved the paper.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716–723. MR0423716

Amin, Z. H. (2008). A note on the parameter estimation for the lognormal distribution based on progressively type I interval censored samples. *Model Assisted Statistics and Applications* **3**, 169–176. MR2528638

Chen, D. G. and Lio, Y. L. (2010). Parameter estimations for generalized exponential distribution under progressive type-I interval censoring. *Computational Statistics and Data Analysis* **54**, 1581–1591. MR2600858

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74**, 829–836. MR0556476

Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *The American Statistician* **35**, 54.

Corradini, M. G., Normand, M. D., Nussinovitch, A., Horowitz, J. and Peleg, M. (2001). Estimating the frequency of high microbial counts in commercial food products using various distribution functions. *Journal of Food Protection* **64**, 674–681.

Ding, C., Yang, C. and Tse, S.-K. (2010). Accelerated life test sampling plans for the Weibull distribution under type I progressive interval censoring with random removals. *Journal of Statistical Computation and Simulation* **80**, 903–914. MR2723138

Fay, M. P. and Shaw, P. A. (2010). Exact and asymptotic weighted logrank tests for interval censored data: The interval R package. *Journal of Statistical Software* **36**, 1–34.

Ferguson, T. S. (1996). *A Course in Large Sample Theory*. London: Chapman and Hall. MR1699953

Hashimoto, E. M., Ortega, E. M. M., Cancho, V. G. and Cordeiro, G. M. (2010). The log-exponential Weibull regression model for interval censored data. *Computational Statistics and Data Analysis* **54**, 1017–1035. MR2580935

Hürlimann, W. (2001). Financial data analysis with two symmetric distributions. *Astin Bulletin* **31**, 187–211. MR1945639

Kozubowski, T. J. and Nadarajah, S. (2010). Multitude of Laplace distributions. *Statistical Papers* **51**, 127–148. MR2556591

Kozubowski, T. and Podgórski, K. (2003a). A log-Laplace growth rate model. *The Mathematical Scientist* **28**, 49–60. MR1995165

Kozubowski, T. and Podgórski, K. (2003b). Log-Laplace distributions. *International Mathematical Journal* **3**, 467–495. MR1966732

Lawless, F. J. (2003). *Statistical Models and Methods for Lifetime Data*. New York: Wiley. MR1940115

Lin, C.-T., Wu, S. J. S. and Balakrishnan, N. (2009). Planning life tests with progressively type-I interval censored data from the lognormal distribution. *Journal of Statistical Planning and Inference* **139**, 54–61. MR2460551

Lindsey, J. K. (1998). A study of interval censoring in parametric regression models. *Lifetime Data Analysis* **4**, 329–354.

Lindsey, J. K., Byrom, W. D., Wang, J., Jarvis, P. and Jones, B. (2000). Generalized nonlinear models for pharmacokinetic data. *Biometrics* **56**, 81–88.

Lio, Y. L., Chen, D.-G. and Tsai, T.-R. (2011). Parameter estimations for generalized Rayleigh distribution under progressively type-I interval censored data. *Open Journal of Statistics* **1**, 46–57. MR2922304

Lu, W. and Tsai, T.-R. (2009). Interval censored sampling plans for the gamma lifetime model. *European Journal of Operational Research* **192**, 116–124. MR2458459

Ng, H. K. T. and Wang, Z. (2009). Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample. *Journal of Statistical Computation and Simulation* **79**, 145–159. MR2655680

Patel, M. N. and Gajjar, K. A. (2010). Inference under progressive interval censoring for geometric distribution. *International Journal of Statistics and Economics* **5**, 21–36. MR2727604

Peleg, M. (2002). Interpretation of the irregularly fluctuating microbial counts in commercial dairy products. *International Dairy Journal* **12**, 255–262.

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at www.r-project.org.

Scallan, A. J. (1999). Regression modelling of interval censored failure time data using the Weibull distribution. *Journal of Applied Statistics* **26**, 613–618. MR1704554

Sparling, Y. H., Younes, N., Lachin, J. M. and Bautista, O. M. (2006). Parametric survival models for interval censored data with time-dependent covariates. *Biostatistics* **7**, 599–614.

Sun, J. (2006). *The Statistical Analysis of Interval Censored Failure Time Data*. New York: Springer. MR2287318

Teimouri, M. and Gupta, A. K. (2012). Estimation methods for the Gompertz–Makeham distribution under progressively type-I interval censoring scheme. *National Academy Science Letters—India* **35**, 227–235.

Tse, S.-K., Ding, C. and Yang, C. (2008). Optimal accelerated life tests under interval censoring with random removals: The case of Weibull failure distribution. *Statistics* **42**, 435–451. MR2457647

Tse, S.-K., Yuen, H.-K. and Yang, C. (2002). Statistical analysis of exponential lifetimes under an integrated type-II interval censoring scheme. *Journal of Statistical Computation and Simulation* **72**, 461–471. MR1927520

Uppuluri, V. R. R. (1981). Some properties of the log-Laplace distribution. In *Statistical Distributions in Scientific Work, Volume 4*. *NATO Advanced Study Institute Series C: Mathematical and Physical Sciences* **79**, 105–110. Dordrecht: Reidel. MR0656148

Zhou, Y. Y., Mi, J. and Guo, S. (2007). Estimation of parameters in logistic and log-logistic distribution with grouped data. *Lifetime Data Analysis* **13**, 421–429. MR2409958

Departamento de Estatística
Universidade Federal de São Carlos
13565-905, São Carlos-SP
Brazil

School of Mathematics
University of Manchester
Manchester M13 9PL
United Kingdom