Statistical Methods & Applications 11: 187–200 (2002)

SMA

Computing maximum likelihood estimates from Type II doubly censored exponential data

Arturo J. Fernández, José I. Bravo and Íñigo De Fuentes

Departamento de Estadística, I. O. y Computación, Facultad de Matemáticas, Universidad de La Laguna, 38271 La Laguna, Tenerife, Spain (e-mail: ajfernan@ull.es)

Abstract. It is well-known that, under Type II double censoring, the maximum likelihood (*ML*) estimators of the location and scale parameters, θ and σ, of a two-parameter exponential distribution are linear functions of the order statistics. In contrast, when θ is known, the *ML* estimator of σ does not admit a closed form expression. It is shown, however, that the *ML* estimator of the scale parameter exists and is unique. Moreover, it has good large-sample properties. In addition, sharp lower and upper bounds for this estimator are provided, which can serve as starting points for iterative interpolation methods such as regula falsi. Explicit expressions for the expected Fisher information and Cramér-Rao lower bound are also derived. In the Bayesian context, assuming an inverted gamma prior on σ, the uniqueness, boundedness and asymptotics of the highest posterior density estimator of σ can be deduced in a similar way. Finally, an illustrative example is included.

Key words: Maximum likelihood estimation, Type II double censoring, exponential distributions, order statistics, Bayes estimators

1 Introduction

The two-parameter exponential distribution, $Exp(\theta, \sigma)$, provides a population model which is useful in several areas of statistics. Inference procedures based on this model are extensively used. In survival and reliability analysis, this distribution plays an important role and is often used to describe certain lifetime data in the biomedical area, as well as component failure observations or the reliability of an equipment in industrial applications. The probability density function of a random variable $T \sim Exp(\theta, \sigma)$ is given by

$$f(t; \theta, \sigma) = \sigma^{-1} \exp\{-(t - \theta)/\sigma\}, t > \theta, \sigma > 0, \quad (1)$$

where θ and σ are the corresponding location and scale parameters. Its mean, $100p$-th percentile, $0 < p < 1$, and cumulative distribution function (*cdf*) are respectively given by $\mu = \theta + \sigma$, $\xi_p = \theta - \sigma \log(1-p)$ and

$$F(t;\theta,\sigma) = F_t = 1 - \exp\left\{-\left(t-\theta\right)/\sigma\right\},\ t > \theta. \tag{2}$$

References about this model may be found, among many others, in Bain (1978) and Lawless (1982). In lifetime data analysis, the location parameter, θ, may be interpreted as an unknown point at which "life" begins or as a "guarantee time" during which failure cannot occur. In many situations one can reasonably assume that the probability of failure up to certain time θ is zero and, thereafter, the failure time follows the $Exp(0,\sigma)$ distribution. In addition, the $Exp(\theta,\sigma)$ model arises as the limiting form of the distribution of the minimum of random samples from some densities with support on (θ,∞). This property is often a justification for its use in reliability studies in which a complex mechanism fails when any one of its many components fails.

In reliability analysis, due to time limitations and/or other restrictions on data collection, several lifetimes of units put on test may not be observed. In addition, sometimes the lowest and/or highest few observations in a sample could be due to some negligence or some other extraordinary reasons. It is therefore convenient to remove those outlying observations. Type II censored samples are considered here, whereby, in an ordered sample of size n, a known number of observations is missing at either end (single censoring) or at both ends (double censoring). Doubly censored samples have been considered, among other authors, by Sarhan (1955), Tiku (1967), Harter and Moore (1968), Kambo (1978), Bhattacharyya (1985), Tiku et al. (1986), Leemis and Shih (1989), Raqab (1995), Balakrishnan and Chan (1995), Lalitha and Mishra (1996), Elfessi (1997) and Fernández (2000).

2 Maximum likelihood estimation

Consider a random sample of size n from an $Exp(\theta,\sigma)$ distribution (θ and σ being unknown), and let $t_r, ..., t_s$ be the ordered observations remaining when the $(r-1) = nq_1$ smallest observations and the $(n-s) = nq_2$ largest observations have been censored, where $r < s$, q_1 and q_2 are fixed, and $0 \leq q_1 + q_2 < 1$. The likelihood function of (θ,σ), given the Type II doubly censored sample $\mathbf{t} = (t_r, ..., t_s)$, can be written as

$$L(\theta,\sigma \mid \mathbf{t}) = \frac{n! \prod_{i=r}^{s} f(t_i;\theta,\sigma)}{(r-1)!(n-s)!} \left\{F(t_r;\theta,\sigma)\right\}^{r-1} \left\{1 - F(t_s;\theta,\sigma)\right\}^{n-s}.$$

We will first consider the case where both parameters, θ and σ, are unknown. Further, we will discuss the case where the parameter θ is known to be θ_0. Finally, as a numerical illustration, an example is presented in Section 3.

2.1 Unknown location parameter case

ML estimation will be considered here for the case in which both parameters in (1) are unknown. According to (1)–(2), the likelihood becomes proportional to

$$L(\theta,\sigma \mid \mathbf{t}) \propto \sigma^{-m} \exp\left[-\left\{V-(n-r+1)\,\theta\right\}/\sigma\right]\left[1-\exp\left\{-(t_r-\theta)/\sigma\right\}\right]^{r-1},$$

where $m = s - r + 1$ and $V = \sum_{i=r}^{s} t_i + (n-s)t_s$.

It is mentioned in Tiku (1967) that, unless $r = 1$, it is not possible to obtain a closed form expression for the *ML* estimates of θ and σ, $\widehat{\theta}$ and $\widehat{\sigma}$. For this reason, he considers a simplification of the likelihood function and proposes the following modified *ML* estimates:

$$\widehat{\theta}_T = \begin{cases} t_1 & \text{if } r = 1, \\ t_r - a_r U/m & \text{if } r > 1, \end{cases} \qquad \text{and} \qquad \widehat{\sigma}_T = U/m,$$

where $a_r \equiv a_{r:n} = \sum_{i=1}^{r} (n-i+1)^{-1}$ and $U = V - (n-r+1)\,t_r$. Nonetheless, the following explicit expressions for $\widehat{\theta}$ and $\widehat{\sigma}$ were derived in Kambo (1978):

$$\widehat{\theta} = t_r - c_r U/m \quad \text{and} \quad \widehat{\sigma} = U/m,$$

where $c_r = -\log\{1-(r-1)/n\} = -\log(1-q_1)$. Thus the *ML* estimates of θ and σ may be expressed as linear functions of the order statistics. Evidently, $\widehat{\mu} = \widehat{\theta} + \widehat{\sigma}$, $\widehat{\xi}_p = \widehat{\theta} - \widehat{\sigma}\log(1-p)$ and $\widehat{F}_{t_0} = 1 - \exp\{-(t_0 - \widehat{\theta})/\widehat{\sigma}\}$ are the respective *ML* estimates of the mean, $100p$-th percentile and *cdf* at $t_0 > \widehat{\theta}$. Since

$$U = \sum_{i=r+1}^{s} (n-i+1)\,(t_i - t_{i-1}) = \sum_{i=r+1}^{s} z_i$$

and $z_{r+1}, .., z_s$ are independent $Exp(0,\sigma)$ variates (see David, 1970, p. 17), it follows that $2U/\sigma = 2m\widehat{\sigma}/\sigma$ is distributed as χ^2 with $2(m-1)$ degrees of freedom. Moreover, it can readily be shown that the distribution of $\exp\{-(t_r-\theta)/\sigma\}$ is $beta$ with parameters $n-r+1$ and r, which implies that $(t_r-\theta)\,/\,(2m\widehat{\sigma})$ is pivotal. The above results can be used immediately to construct confidence intervals and tests of significance for σ and θ. As $2m\widehat{\sigma}/\sigma \sim \chi^2_{2(m-1)}$, it turns out that $E(\widehat{\sigma}) = \sigma - \sigma/m$ and $Var(\widehat{\sigma}) = (m-1)\,\sigma^2/m^2$. It is well-known (David, 1970, p. 39) that $E(t_r) = \theta + a_r\sigma$ and $Var\,(t_r) = Cov(t_r, t_k) = b_r\sigma^2$ for $r \le k \le n$, where $b_r \equiv b_{r:n} = \sum_{i=1}^{r} (n-i+1)^{-2}$. Notice that $(t_r - \theta)$ coincides with $\sum_{i=1}^{r} z_i/\,(n-i+1)$. Thus, $E(\widehat{\theta}) = \theta + \sigma\,\{a_r - (m-1)\,c_r/m\}$, $Var(\widehat{\theta}) = \sigma^2\left\{b_r + (m-1)\,c_r^2/m^2\right\}$ and $Cov(\widehat{\theta},\widehat{\sigma}) = -c_r\,(m-1)\,\sigma^2/m^2$.

By using an extended version of the strong law of large numbers, it is clear that $\widehat{\sigma} = \sum_{i=r+1}^{s} z_i/m$ converges almost surely ($\overset{a.s.}{\to}$) to σ (all limits in this paper are taken as $n \to \infty$). Moreover, $t_r \overset{a.s.}{\to} \xi_{q_1} = \theta - \sigma\log(1-q_1)$. Therefore, $\widehat{\theta}$ and $\widehat{\sigma}$ are strongly consistent estimators of θ and σ, respectively. In addition, from the central limit theorem, $n^{1/2}(\widehat{\sigma} - \sigma)$ converges in distribution ($\overset{d}{\to}$) to a normal $N\left(0, \sigma^2/\,(1-q_1-q_2)\right)$ variable. Likewise, provided that $q_1 > 0$, it can be shown that $n^{1/2}(t_r - \xi_{q_1}) \xrightarrow{d} N\left(0, q_1\sigma^2/\left(1-q_1\right)\right)$; if $q_1 = 0$, $\widehat{\theta} \sim Exp(\theta, \sigma/n)$, that is, $2n(\widehat{\theta}-\theta)/\sigma \sim \chi^2_2$. Consequently, by noting that t_r and $\widehat{\sigma}$ are independent, and $(\widehat{\theta}-\theta, \widehat{\sigma}-\sigma)$ coincides with $(t_r - \xi_{q_1}, 0) + (\log(1-q_1), 1)(\widehat{\sigma}-\sigma)$, it follows that, when $q_1 > 0$, the asymptotic distribution of $n^{1/2}(\widehat{\theta}-\theta, \widehat{\sigma}-\sigma)$ is bivariate normal with mean zero vector and variance-covariance matrix

$$\Lambda = \sigma^2 \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \sigma^2 \begin{pmatrix} \frac{q_1}{1-q_1} + \frac{\{\log(1-q_1)\}^2}{1-q_1-q_2} & \frac{\log(1-q_1)}{1-q_1-q_2} \\ \frac{\log(1-q_1)}{1-q_1-q_2} & \frac{1}{1-q_1-q_2} \end{pmatrix}.$$

Thus, the asymptotic variances of $n^{1/2}\widehat{\mu}$ and $n^{1/2}\widehat{\xi}_p$ are given respectively by

$$Avar(n^{1/2}\widehat{\mu}) = \sigma^2 \left[\frac{q_1}{1-q_1} + \frac{\{1+\log(1-q_1)\}^2}{1-q_1-q_2}\right]$$

and

$$Avar(n^{1/2}\widehat{\xi}_p) = \sigma^2 \left[\frac{q_1}{1-q_1} + \frac{\{\log(1-q_1)-\log(1-p)\}^2}{1-q_1-q_2}\right].$$

Furthermore, by using the delta method (Rao, 1973),

$$Avar(n^{1/2}\widehat{F}_{t_0}) = (1-F_{t_0})^2 \left[\lambda_{11} - 2\lambda_{12}\log(1-F_{t_0}) + \lambda_{22}\left\{\log(1-F_{t_0})\right\}^2\right].$$

In spite of the fact that $\widehat{\theta}$ and $\widehat{\sigma}$ are both biased, their biases (which are both of order m^{-1}) may easily be corrected. In this way, the estimators $\sigma^* = m\widehat{\sigma}/(m-1)$ and $\theta^* = t_r - a_r\sigma^*$ are unbiased for σ and θ, respectively, and $Var(\theta^*) = \sigma^2\left\{b_r + a_r^2/\left(m-1\right)\right\}$, $Var(\sigma^*) = \sigma^2/\left(m-1\right)$ and $Cov(\theta^*, \sigma^*) = -a_r\sigma^2/\left(m-1\right)$. The above estimators, θ^* and σ^*, are the same as the best linear unbiased (*BLU*) estimators, which were obtained in Sarhan (1955). It is readily seen that $\widehat{\theta}$ and $\widehat{\sigma}$ are jointly sufficient. They are also complete for θ and σ, which formally establishes (by the Lehmann-Scheffé theorem) that θ^* and σ^* are the best ones in the class of all estimators, not merely among linear estimators. Furthermore, it is obvious that $\widehat{\theta}$ and $\widehat{\sigma}$ are asymptotically efficient (unbiased and of minimum variance). Notice, however, that the mean square error (mse) of $\widehat{\sigma}$ is strictly less than the mse of σ^*. Also observe that, in terms of mse, $\widehat{\theta}$ is neither uniformly better nor uniformly worse than θ^*. Nevertheless, among linear estimators of σ and θ, the corresponding best linear invariant (*BLI*) estimators, $\sigma^{**} = \widehat{\sigma}$ and $\theta^{**} = t_r - a_r\sigma^{**}$ (see Nelson, 1982, p. 307, for construction), provide the smallest mse's. It is directly deduced that $E(\theta^{**}) = \theta + a_r\sigma/m$, $E(\sigma^{**}) = \sigma - \sigma/m$, $mse(\theta^{**}) = \sigma^2\left(b_r + a_r^2/m\right)$, $mse(\sigma^{**}) = \sigma^2/m$ and $E\left[\left(\theta^{**}-\theta\right)\left(\sigma^{**}-\sigma\right)\right] = -a_r\sigma^2/m$.

The *ML* estimators of μ and ξ_p are given by $\widehat{\mu} = \widehat{\theta} - \widehat{\sigma}\log(1-p)$ and $\widehat{\xi}_p = \widehat{\theta} - \widehat{\sigma}\log(1-p)$. In addition, $E(\widehat{\mu}) = \mu + \sigma[a_r - c_r + (c_r-1)/m]$, $Var(\widehat{\mu}) = \sigma^2\left\{b_r + (m-1)(c_r-1)^2/m^2\right\}$, $E(\widehat{\xi}_p) = \xi_p + \sigma[a_r - c_r + \left\{c_r + \log(1-p)\right\}/m]$ and $Var(\widehat{\xi}_p) = \sigma^2[b_r + (m-1)\left\{c_r + \log(1-p)\right\}^2/m^2]$.

Similarly, the *BLU* estimators of μ and ξ_p can be obtained as $\mu^* = \theta^* + \sigma^*$ and $\xi_p^* = \theta^* - \sigma^* \log(1-p)$. Clearly, $Var(\mu^*) = \sigma^2 \{b_r + (a_r - 1)^2/(m-1)\}$ and $Var(\xi_p^*) = \sigma^2[b_r + \{a_r + \log(1-p)\}^2/(m-1)]$. In the same manner, the *BLI* estimators of μ and ξ_p are given by $\mu^{**} = \theta^{**} + \sigma^{**}$ and $\xi_p^{**} = \theta^{**} - \sigma^{**}\log(1-p)$. It is also derived that $E(\mu^{**}) = \mu + (a_r - 1)\sigma/m$, $E(\xi_p^{**}) = \xi_p + \{a_r + \log(1-p)\}\sigma/m$, $mse(\mu^{**}) = \sigma^2\{b_r + (a_r-1)^2/m\}$ and $mse(\xi_p^{**}) = \sigma^2[b_r + \{a_r + \log(1-p)\}^2/m]$.

It is clear that the *ML*, *BLU* and *BLI* estimators differ little for practical purposes (unless sample sizes are small). The choice of either estimator is mostly a matter of taste.

2.2 *Known location parameter case*

There are many situations where the location parameter θ may be known. In such cases, a better estimation for σ may be obtained. If θ is known to be θ_0, the likelihood function of σ is proportional to

$$L(\sigma \mid \mathbf{t},\, \theta_0) \propto \sigma^{-m} \exp(-W/\sigma) \{1 - \exp(-w_r/\sigma)\}^{r-1},$$

where $m = s - r + 1$ and $W = \sum_{i=r}^{s} w_i + (n-s)w_s$, in which $w_i = t_i - \theta_0$.

The *ML* estimate of σ, now denoted by $\widehat{\sigma}_0$, can be found by solving

$$G(\sigma) = \frac{\partial}{\partial\sigma} \log L(\sigma \mid \mathbf{t},\, \theta_0) = \frac{1}{\sigma^2}\left\{W - m\sigma - \frac{(r-1)w_r}{\exp(w_r/\sigma) - 1}\right\} = 0. \quad (3)$$

This equation is not solvable analytically (unless $r = 1$, in which case $\widehat{\sigma}_0 = W/m$). One approach is to find a modified *ML* estimate by linearizing the likelihood equation appropriately. For details about this methodology, see Tiku et al. (1986). Equation (3) provides, however, a unique *ML* estimate, $\widehat{\sigma}_0$ (theorem 1), which though not expressible explicitly, can be determined by numerical methods. The solution could be obtained by using the Newton-Raphson method or the EM algorithm (see Dempster et al. 1977). Nevertheless, interpolation methods such as regula falsi are preferable since sharp bounds on the value of $\widehat{\sigma}_0$ can be derived (theorem 2).

Theorem 1 *The likelihood function $L(\sigma \mid \mathbf{t},\, \theta_0)$ has a unique global maximum in $(0, +\infty)$, i. e., there is a unique ML estimate $\widehat{\sigma}_0$.*

Proof. The *ML* equation can also be written as $h(\sigma) = g(\sigma)$, where $h(\sigma) = W - m\sigma$ and $g(\sigma) = (r-1)w_r/\{\exp(w_r/\sigma) - 1\}$. It can easily be seen that h and g are decreasing and non-decreasing in $(0, \infty)$, respectively, $G(0^+) = W > 0$ and $G(W/m) \le 0$. Therefore, likelihood equation (3) admits a unique solution, $\widehat{\sigma}_0$. In addition, $0 < \widehat{\sigma}_0 \le W/m$. Since $h'(\widehat{\sigma}_0) < 0$ and $g'(\widehat{\sigma}_0) \ge 0$, then $G'(\widehat{\sigma}_0) = \{h'(\widehat{\sigma}_0) - g'(\widehat{\sigma}_0)\}/\widehat{\sigma}_0^2 < 0$ and hence $\widehat{\sigma}_0$ is the global maximum point. □

Lemma 1 *For $\sigma > 0$,*

$$\max\left\{0,\, (r-1)\left(\sigma - \frac{3\sigma w_r}{6\sigma + w_r}\right)\right\} \le g(\sigma) \le \frac{6(r-1)\sigma^3}{6\sigma^2 + 3\sigma w_r + w_r^2}.$$

Proof. Notice that for $x > 0$, $\exp(x) = \sum_{k=0}^{\infty} x^k/k! > 1+x+x^2/2+x^3/6 > 1$, and for $0 < x < 3$, $\exp(x) = 1 + x + (x^2/2)\sum_{k=2}^{\infty} 2x^{k-2}/k! < 1 + x + (x^2/2)/(1 - x/3)$. Therefore,

$$\max\left(0, 1 - \frac{3x}{6+x}\right) < \frac{x}{\exp(x) - 1} < \frac{1}{1 + x/2 + x^2/6}, \qquad x > 0,$$

and hence the stated result follows. □

Theorem 2 *The ML estimate $\widehat{\sigma}_0$ verifies that $\widehat{\sigma}_L \leq \widehat{\sigma}_0 \leq \widehat{\sigma}_U$, where*

$$\widehat{\sigma}_U = \min\left(\frac{W}{m}, \frac{\gamma}{2} + \frac{1}{2}\left\{\gamma^2 + (2/3)Ww_r/s\right\}^{1/2}\right)$$

and

$$\widehat{\sigma}_L = \left(R + D^{1/2}\right)^{1/3} + \left(R - D^{1/2}\right)^{1/3} - A/3,$$

in which

$$\gamma = \{W - (m - 2r + 2)w_r/6\}/s, \; R = \frac{9AB - 27C - 2A^3}{54},$$

$$D = \frac{1}{108}\left\{\left(4AC - B^2\right)\left(A^2 - 4B\right) - 2ABC + 27C^2\right\} > 0,$$

$$A = \frac{mw_r - 2W}{2s}, \quad B = \frac{mw_r^2 - 3Ww_r}{6s} \quad \text{and} \quad C = \frac{Ww_r^2}{6s}.$$

Proof. Since $h(\widehat{\sigma}_0) = g(\widehat{\sigma}_0)$, it is derived from lemma 1 that

$$m\widehat{\sigma}_0 - W \leq 0 \text{ and } m\widehat{\sigma}_0 - W + (r-1)\left(\widehat{\sigma}_0 - \frac{3\widehat{\sigma}_0 w_r}{6\widehat{\sigma}_0 + w_r}\right) \leq 0.$$

Therefore, $\widehat{\sigma}_0 \leq W/m$ and $\widehat{\sigma}_0 \leq \gamma/2 + \left\{\gamma^2/4 + Ww_r/(6s)\right\}^{1/2}$. In a similar way,

$$m\widehat{\sigma}_0 - W + \frac{6(r-1)\widehat{\sigma}_0^3}{6\widehat{\sigma}_0^2 + 3\widehat{\sigma}_0 w_r + w_r^2} \geq 0,$$

i.e., $\widehat{\sigma}_0^3 + A\widehat{\sigma}_0^2 + B\widehat{\sigma}_0 + C \geq 0$. Upon using the Cardanus formula, it is deduced that $\widehat{\sigma}_0 \geq \widehat{\sigma}_L$. □

When $r = 1$, $\widehat{\sigma}_0$ reduces to W/m and the inequality $\widehat{\sigma}_L \leq \widehat{\sigma}_0 \leq \widehat{\sigma}_U$ becomes an equality; otherwise, the rule of false position or regula falsi (iterative linear interpolation) is an easy and convenient procedure to determine $\widehat{\sigma}_0$. Firstly, theorem 2 provides sharp bounds on the value of $\widehat{\sigma}_0$. Secondly, this method always works well (Newton's method, for example, may fail to converge if the initial value is not close to the solution) and experience has shown that the rate of convergence is very rapid (e.g., Stoer and Bulirsch, 1983). Additionally, computational costs are reduced (there is no need to calculate derivatives) and, above all, estimation errors may be accurately bounded.

There is no known analytic method of determining the mean and variance of $\widehat{\sigma}_0$. Large-sample properties of *ML* estimation are considered in theorem 3 instead. Explicit expressions for the expected Fisher information and Cramér-Rao lower bound are derived in theorem 4.

Theorem 3 *The ML estimators $\widehat{\sigma}_0$ and $\widehat{F}_{t_0} = 1 - \exp\{-(t_0 - \theta_0)/\widehat{\sigma}_0\}$, with $t_0 > \theta_0$, are strongly consistent, and asymptotically efficient and normal. The asymptotic variances of $n^{1/2}\widehat{\sigma}_0$ and $n^{1/2}\widehat{F}_{t_0}$ are given respectively by σ^2/α and $\{(1 - F_{t_0})\log(1 - F_{t_0})\}^2/\alpha$, where*

$$\alpha = \begin{cases} 1 - q_2 & \text{if } q_1 = 0, \\ 1 - q_1 - q_2 + (1 - q_1)\{\log(1 - q_1)\}^2/q_1 & \text{if } q_1 > 0. \end{cases}$$

Proof. Since the exponential $Exp(\theta_0, \sigma)$ model is regular, the asymptotic efficiency and normality, and the strong consistency of $\widehat{\sigma}_0$ can be verified following the lines of Section 3 in Bhattacharyya (1985) with appropriate extensions of his theorems 1 and 2 to the double censoring case (as he indicated in Section 2, p. 400), and it is therefore omitted. Upon using the delta method, it is clear that $\widehat{F}_{t_0}$ satisfies the same asymptotic properties.

In order to derive the asymptotic variance of $n^{1/2}\widehat{\sigma}_0$, observe that, when $n \to \infty$ (with q_1 and q_2 fixed), t_r and t_s converge $a.s.$ to the corresponding population percentiles $\xi_{q_1} = \theta_0 - \sigma\log(1 - q_1)$ and $\xi_{1-q_2} = \theta_0 - \sigma\log q_2$. Moreover, by virtue of an extended version of the strong law of large numbers,

$$\begin{aligned} \frac{1}{n}\sum_{i=r}^{s} w_i &\xrightarrow{a.s.} \int_{\xi_{q_1}}^{\xi_{1-q_2}} (t - \theta_0) f(t; \theta_0, \sigma)\, dt \\ &= \sigma\left[(1 - q_1)\{1 - \log(1 - q_1)\} - q_2(1 - \log q_2)\right], \end{aligned}$$

where the term involving $\log q_2$ drops out when $q_2 = 0$, which implies that

$$\frac{W}{n} = \frac{\sum_{i=r}^{s} w_i + (n - s)w_s}{n} \xrightarrow{a.s.} \sigma\{1 - q_1 - q_2 - (1 - q_1)\log(1 - q_1)\}.$$

Hence,

$$-\frac{1}{n}\frac{\partial^2}{\partial\sigma^2}\log L(\sigma \mid \mathbf{t}, \theta_0) \xrightarrow{a.s.} \frac{\alpha}{\sigma^2},$$

and the asymptotic variance of $n^{1/2}\widehat{\sigma}_0$ follows. The asymptotic variance of $n^{1/2}\widehat{F}_{t_0}$ can be found by using the delta method. □

Theorem 4 *The Cramér-Rao lower bound for the variance of any unbiased estimator of σ is given by $CRLB(\sigma) = \sigma^2/(m + \beta_{r:n})$, $r = 1, \ldots, n - 1$, where*

$$\beta_{1:n} = 0, \quad \beta_{2:n} = 2n(n - 1)\sum_{i=0}^{\infty}(n + i)^{-3}$$

and

$$\beta_{r:n} = n(n - r + 1)\left(b_{r-2:n-1} + a_{r-2:n-1}^2\right)/(r - 2), \; r = 3, \ldots, n - 1.$$

Moreover, for $t_0 > \theta_0$, $CRLB(F_{t_0}) = \{(1 - F_{t_0})\log(1 - F_{t_0})\}^2/(m + \beta_{r:n})$.

Proof. Notice that, after some calculations, the expected Fisher information, is obtained to be

$$
\begin{aligned}
I_{(r,n,m)}(\sigma) &= -E\left[\frac{\partial^2}{\partial\sigma^2}\log L(\sigma \mid \mathbf{t},\, \theta_0)\right] \\
&= \frac{m}{\sigma^2} + \frac{(r-1)}{\sigma^4} E\left[\frac{w_r^2 \exp(w_r/\sigma)}{\{\exp(w_r/\sigma)-1\}^2}\right] = \frac{m+\beta_{r:n}}{\sigma^2},
\end{aligned}
$$

where $\beta_{1:n} = 0$ and

$$
\beta_{r:n} = \frac{n!}{(r-2)!\,(n-r)!}\int_0^\infty x^2 \exp\{-(n-r+2)\,x\}\{1-\exp(-x)\}^{r-3}\,dx
$$

for $r = 2, ..., n-1$. It can be shown that

$$
\beta_{2:n} = n\,(n-1)\sum_{i=0}^{\infty}\int_0^\infty x^2 \exp\{-(n+i)\,x\}\,dx = 2n(n-1)\sum_{i=0}^{\infty}(n+i)^{-3},
$$

and, for $r = 3, ..., n-1$,

$$
\beta_{r:n} = \frac{n(n-r+1)}{r-2}\left[Var\left(X_{r-2:n-1}\right) + \{E\left[X_{r-2:n-1}\right]\}^2\right],
$$

where $X_{r-2:n-1}$ is the $(r-2)$-th order statistics in a sample of size $(n-1)$ from an $Exp(0,1)$ distribution, which yields that

$$
\beta_{r:n} = \frac{n(n-r+1)}{r-2}\left[\sum_{i=1}^{r-2}(n-i)^{-2} + \left\{\sum_{i=1}^{r-2}(n-i)^{-1}\right\}^2\right].
$$

Hence, the expressions of $CRLB(\sigma) = 1/I_{(r,n,m)}(\sigma)$ and $CRLB(F_{t_0}) = \{(\partial/\partial\sigma)\,F_{t_0}\}^2/I_{(r,n,m)}(\sigma)$ follow directly. □

Remark 1 Approximate value of $\beta_{r:n}$ for $r = 2, ..., n-1$.

Since

$$
S_n = \sum_{i=0}^{\infty}(n+i)^{-3} = \frac{1}{2}\int_0^\infty \frac{x^2\exp(-nx)}{1-\exp(-x)}dx = \frac{1}{2n^3}\int_0^\infty \frac{t^2\exp(-t)}{1-\exp(-t/n)}dt,
$$

it turns out that

$$
S_n \simeq \frac{1}{2n^2}\int_0^\infty \frac{t\exp(-t)}{1-\frac{t}{2n}+\frac{t^2}{6n^2}}dt \simeq \frac{1}{2n^2}\int_0^\infty t\exp(-t)\left(1+\frac{t}{2n}+\frac{t^2}{12n^2}\right)dt
$$

when n is large. Thus

$$
\beta_{2:n} = 2n(n-1)S_n \simeq 2n(n-1)\left(\frac{1}{2n^2}+\frac{1}{2n^3}+\frac{1}{4n^4}\right) = 1-\frac{1}{2n^2}-\frac{1}{2n^3}.
$$

This approximation is very accurate even for $n = 5$. In addition, it is well-known that $1 + 1/2 + .. + 1/n \simeq \log(n + 1/2) + \gamma$ and $1 + 1/2^2 + .. + 1/n^2 \simeq \pi^2/6 - (n + 1/2)^{-1}$ for n large (say, $n \geq 10$), where $\gamma = 0.577215 \cdot\cdot$ is Euler's constant. Hence, $a_{j:n} \simeq -\log\{1 - j/(n + 1/2)\}$ and $b_{j:n} \simeq j/\{(n + 1/2)(n + 1/2 - j)\}$ provided that j and $n - j$ are large. As a consequence, we can derive an approximate value of $\beta_{r:n}$ when r and $(n - r)$ are large. It is also clear that $\widehat{\sigma}_0$ is asymptotically efficient since $n\alpha/\left(m + \beta_{r:n}\right) \to 1$. Observe that $\beta_{r:n}/n$ converges to $(1 - q_1)\{\log(1 - q_1)\}^2/q_1$ when $q_1 > 0$. Evidently, for complete samples, $(r = 1, s = n)$, $I_{(1,n,n)}(\sigma)$ is simply n/σ^2; otherwise, there is loss of information due to censoring. For instance, notice that $I_{(2,n,n-1)}(\sigma) = (n - 1 + \beta_{2:n})/\sigma^2$ is approximately equal to $\left\{n - 1/(2n^2) - 1/(2n^3)\right\}/\sigma^2$ for $n \geq 5$. In general, when the sample size grows to infinity, $I_{(r,n,m)}(\sigma)/I_{(1,n,n)}(\sigma)$ converges to α.

It is an empirical fact that transforming an estimator to remove the dependency of the variance on the unknown parameter tends to improve the convergence to normality, due to the reduction of the skewness. In this way, by using the delta method, it can be seen that $n^{1/2}(\log \widehat{\sigma}_0 - \log \sigma)$ is asymptotically $(\overset{a}{\sim})$ $N(0, 1/\alpha)$. Similarly, it is easy to show that $n^{1/2}[\log\{-\log(1 - \widehat{F}_{t_0})\} - \log\{(t_0 - \theta_0)/\widehat{\sigma}_0\}] \overset{a}{\sim} N(0, 1/\alpha)$. These results can be used for test and confidence interval construction. However, it is preferable to use that $2W/\sigma \sim \chi^2_{2s}$ when $r = 1$, and $2Y/\sigma \sim \chi^2_{2(m-1)}$ if $r > 1$, where $Y = W - (n - r + 1)\,w_r$. Note that $W = \sum_{i=1}^{s} z_i$ if $r = 1$ and $Y = \sum_{i=r+1}^{s} z_i$ if $r > 1$, where $z_i = (n - i + 1)\,(w_i - w_{i-1})$, w_0 being 0, and also that the z_i's are independent $Exp(0, \sigma)$ variates.

On the other hand, in this case, $\sigma_0^* = (a_r w_r/b_r + m\widehat{\sigma})/d_r$ is the *BLU* estimator of σ, where $d_r \equiv d_{r:n} = a_r^2/b_r + m - 1$ (see Leemis and Shih, 1989), and its variance is $Var(\sigma_0^*) = \sigma^2/d_r$, whereas the *BLI* estimator of σ is $\sigma_0^{**} = \sigma_0^* d_r/(1 + d_r)$ and its mse is $\sigma^2/(1 + d_r)$. Therefore, from theorem 4, the efficiency of σ_0^*, $eff\,(\sigma_0^*) = CRLB\,(\sigma)/Var(\sigma_0^*)$, is given by $d_{r:n}/(m + \beta_{r:n})$. An approximate value of $eff\,(\sigma_0^*)$ can readily be found using the results in remark 1. In particular, if $r = 2$ and $n \geq 5$, $eff\,(\sigma_0^*) \simeq \left[s - 1/\left\{n^2 + (n-1)^2\right\}\right] / \left[s - (n + 1)/\left(2n^3\right)\right]$.

Obviously, the *ML* estimators of μ and ξ_p are $\widehat{\mu} = \theta_0 + \widehat{\sigma}_0$ and $\widehat{\xi}_p = \theta_0 - \widehat{\sigma}_0 \log(1 - p)$. Their asymptotic variances are given by $(\mu - \theta_0)^2/\alpha$ and $(\xi_p - \theta_0)^2/\alpha$, respectively. The *BLU* estimators of μ and ξ_p can be obtained as $\mu^* = \theta_0 + \sigma_0^*$ and $\xi_p^* = \theta_0 - \sigma_0^* \log(1 - p)$, and their mse's as $mse(\mu^*) = \sigma^2/d_r$ and $mse(\xi_p^*) = \{\log(1 - p)\}^2 \sigma^2/d_r$. In the same manner, the *BLI* estimators of μ and ξ_p are given by $\mu^{**} = \theta_0 + \sigma_0^{**}$ and $\xi_p^{**} = \theta_0 - \sigma_0^{**} \log(1 - p)$. Their mse's are $mse(\mu^{**}) = \sigma^2/(1 + d_r)$ and $mse(\xi_p^{**}) = \{\log(1 - p)\}^2 \sigma^2/(1 + d_r)$. Moreover, $E(\mu^{**}) = \mu - \sigma/(1 + d_r)$ and $E(\xi_p^{**}) = \xi_p + \sigma \log(1 - p)/(1 + d_r)$.

It is important to note here that, since $d_{r:n}/n$ converges to α as $n \to \infty$, the *ML, BLU* and *BLI* estimators and their mse's are asymptotically equal.

Remark 2 A Bayesian approach.

There are two important practical benefits of a Bayesian analysis. One is the increased quality of the inferences, provided that prior knowledge accurately reflects the true variation in the parameter. The other is the reduction in testing requirement (test time and/or sample size) that often occurs in Bayesian experiments. For the $Exp\,(\theta_0, \sigma)$ model, the natural conjugate prior density on σ under many sampling schemes is the inverted gamma density

$$\phi(\sigma \mid \theta_0) = \frac{a^b}{\Gamma(b)} \sigma^{-(b+1)} \exp(-a/\sigma), \quad \sigma > 0, \quad (a > 0, b > 0).$$

This density is widely used in Bayesian analysis both for its flexibility in representing prior information and for its mathematical tractability. Moments- and percentiles-matching methods may be used to fit this prior density. However, if prior knowledge about σ is vague, or the researcher is unwilling to involve his prior beliefs concerning σ or, simply, a posterior based mainly on the data at hand is desired, an alternative is to use a noninformative (default or diffuse) prior. The most commonly used default prior is Jeffreys' (1961) prior, $\phi_J(\sigma \mid \theta_0)$, which is proportional to the square root of the expected Fisher information, i.e., $\phi_J(\sigma \mid \theta_0) \propto \left\{I_{(r,n,m)}\,(\sigma)\right\}^{1/2}, \sigma > 0$. According to theorem 4, $I_{(r,n,m)}\,(\sigma) = (m + \beta_{r:n})\,/\sigma^2$ and therefore $\phi_J(\sigma \mid \theta_0) \propto \sigma^{-1}, \sigma > 0$. Thus, diffuseness in our prior for σ can be represented by the (improper) choice $a = b = 0$. This density is not proper in the sense that it has infinite mass. However, it is the natural noninformative prior for the scale parameter σ due to the fact that all prior densities for σ that are invariant under scale transformations are proportional to σ^{-1}.

In light of the observed sample data, $\mathbf{t} = (t_r, ..., t_s)$, our actual opinion about σ is summarized by the posterior density of σ, which is given by Bayes' theorem as

$$\phi(\sigma \mid \mathbf{t}, \theta_0) = \frac{(a + W)^{b+m} \left\{1 - \exp\left(w_r/\sigma\right)\right\}^{r-1} \exp\left\{-(a + W)/\sigma\right\}}{\Gamma(b + m) \Lambda_r\,(a + W, b + m)\,\sigma^{b+m+1}}, \quad \sigma > 0,$$

where

$$\Lambda_r\,(p, q) = \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \left(1 + k\frac{w_r}{p}\right)^{-q}, \quad p, q > 0 \quad \left(\Lambda_1\,(\cdot,\cdot) \equiv 1\right).$$

Using the transformation $\sigma = -\,(t_0 - \theta_0)\,/\log\left(1 - F_{t_0}\right), t_0 > \theta_0$, it follows that the posterior density of F_{t_0} is given by

$$\Phi(F_{t_0} \mid \mathbf{t}, \theta_0) = \frac{(t_0 - \theta_0)\,\phi\left(-\,(t_0 - \theta_0)\,/\log\left(1 - F_{t_0}\right) \mid \mathbf{t}, \theta_0\right)}{\left(1 - F_{t_0}\right)\left\{\log\left(1 - F_{t_0}\right)\right\}^2}, \quad 0 < F_{t_0} < 1.$$

From a decision-theoretic viewpoint, in order to select a single value as representing our "best" estimator of σ, one must first specify a loss function, $\mathcal{L}(\sigma, \widetilde{\sigma})$, which represents the cost involved in using the estimate $\widetilde{\sigma}$ when the true value is σ. Under zero-one loss ($\mathcal{L}(\sigma, \widetilde{\sigma}) = 0$ if $\widetilde{\sigma} = \sigma$, and equals 1 otherwise), the Bayes estimator of σ (i.e., the value of $\widetilde{\sigma}$ that minimizes the posterior expected loss), denoted by $\widetilde{\sigma}_{(a,b)}$, coincides with the posterior mode or highest posterior density (*HPD*) estimator, which can be found by solving the modal equation

$$a + W - (b + m + 1)\,\sigma - \frac{(r-1)w_r}{\exp\left(w_r/\sigma\right) - 1} = 0.$$

Obviously, *HPD* estimators are in closed forms when $r = 1$. Otherwise, the existence and uniqueness of $\widetilde{\sigma}_{(a,b)}$ are evident following the reasoning in theorem 1. Moreover, sharp bounds on $\widetilde{\sigma}_{(a,b)}$ can be derived from theorem 2. In addition, for $t_0 > \theta_0$, the *HPD* estimator of F_{t_0}, $\widetilde{F}_{t_0;(a,b)}$, turns out to be

$$\widetilde{F}_{t_0;(a,b)} = \begin{cases} 1 - \exp\left\{-\left(t_0 - \theta_0\right)/\widetilde{\sigma}_{(a-t_0+\theta_0,b-2)}\right\} & \text{if } t_0 < \theta_0 + a + W, \\ 1 & \text{if } t_0 \geq \theta_0 + a + W. \end{cases}$$

Two commonly used loss functions for estimating σ are squared-error loss, $\mathcal{L}(\sigma, \widetilde{\sigma}) = (\sigma - \widetilde{\sigma})^2$, and absolute-error loss, $\mathcal{L}(\sigma, \widetilde{\sigma}) = |\sigma - \widetilde{\sigma}|$. For squared-error loss functions, the Bayes estimator of σ and F_{t_0}, $t_0 > \theta_0$, are simply the respective posterior means, $E\left[\sigma \mid \mathbf{t}, \theta_0\right]$ and $E\left[F_{t_0} \mid \mathbf{t}, \theta_0\right]$, which are given by

$$\int_0^{\infty} \sigma\phi(\sigma \mid \mathbf{t}, \theta_0) d\sigma = \frac{(a+W)\,\Lambda_r\,(a+W, b+m-1)}{(b+m-1)\,\Lambda_r\,(a+W, b+m)}, \quad (b+m > 1),$$

and

$$\int_0^{1} F_{t_0}\Phi(F_{t_0} \mid \mathbf{t}, \theta_0) dF_{t_0} = \frac{\Lambda_r\,(a+W+t_0, b+m)}{\Lambda_r\,(a+W, b+m)\left\{1 + t_0/\left(a+W\right)\right\}^{b+m}},$$

whereas, if absolute-error loss functions are deemed suitable, the posterior medians of σ and F_{t_0} represent the appropriate Bayes estimators. In the right censoring case, as $\Lambda_1\left(\cdot,\cdot\right) \equiv 1$, the above estimators are considerably simplified.

It is worth to remark that, since *ML* estimators are modes of posterior densities corresponding to flat priors, the *ML* estimators of σ and F_{t_0} coincide with the *HPD* ones when the pairs of hyperparameters (a, b) are respectively $(0, -1)$ and $(t_0 - \theta_0, 1)$. It should be noted that the difference between *HPD* and *ML* estimators is numerically small when W and m are large in relation to a and b, respectively. Indeed, its difference converges $a.s.$ to 0 as $n \to \infty$. By virtue of theorem 3, it is clear that $\widetilde{\sigma}_{(a,b)} \xrightarrow{a.s.} \sigma$ and $n^{1/2}\{\widetilde{\sigma}_{(a,b)} - \sigma\} \xrightarrow{d} N(0, \sigma^2/\alpha)$. As well, because the influence of the prior distribution diminishes with increasing sample size, both the posterior mean and median approach the *ML* estimator and have the same asymptotic properties. Moreover, for large sample sizes, the posterior density $\phi(\sigma \mid \mathbf{t}, \theta_0)$ is approximately normal with mean the *HPD* estimator, $\widetilde{\sigma}_{(a,b)}$, and variance the inverse of the expected Fisher information evaluated at $\widetilde{\sigma}_{(a,b)}$, $\{\widetilde{\sigma}_{(a,b)}\}^2/\left(m + \beta_{r:n}\right)$.

3 An illustrative example

The following data represent failure times, in minutes, for a specific type of electrical insulation (Lawless, 1982, p. 138; Raqab, 1995): -, -, 24.4, 28.6, 43.2, 46.9, 70.7, 75.3, 95.5, -, -, -. In this example, the experimenter failed to observe the two smallest failure times and the experiment was terminated at the time of the ninth failure. Hence $n = 12$, $r = 3$, $s = 9$ and $m = 7$. Raqab (1995) assumed that the data came from an $Exp(\theta, \sigma)$ distribution.

The *ML* estimates of θ, σ and the expected time to failure, μ, are obtained to be $\widehat{\theta} = 13.27578$, $\widehat{\sigma} = 61.01429$ and $\widehat{\mu} = \widehat{\theta} + \widehat{\sigma} = 74.29007$. Their corresponding root mean square errors ($rmse$'s) are given by 12.68091, 23.06123 and 20.02325.

Table 1 shows the *ML*, *BLU* and *BLI* estimates of θ, σ, μ and ξ_p, $p = 0.1, 0.5, 0.9$, and their $rmse$ values. Among them, the *BLI* estimates give the smallest $rmse$ values, whereas the *BLU* estimates give the largest. The *ML* and *BLI* estimates provide similar $rmse$ values.

Table 1. ML, BLU and BLI estimates and their $rmse$'s

	ML		BLU		BLI	
	estimate	$rmse$	estimate	$rmse$	estimate	$rmse$
θ	13.27578	12.68091	4.878510	13.82932	7.667294	11.56909
σ	61.01429	23.06123	71.18333	29.06047	61.01429	23.06123
μ	74.29007	20.02325	76.06184	23.92822	68.68158	19.33828
$\xi_{0.1}$	19.70378	11.66087	12.37842	12.32159	14.09579	10.44100
$\xi_{0.5}$	55.56716	14.63326	54.21903	16.61117	49.95917	13.68104
$\xi_{0.9}$	153.7659	48.05008	168.7842	60.01834	148.1579	47.76870

On the other hand, let us assume that the location parameter θ is known to be θ_0. For instance, suppose that $\theta_0 = 0$, which implies that $W = 671.1$. In this case, it turns out that $\widehat{\sigma}_L = 77.11759$ and $\widehat{\sigma}_U = 77.14201$. Hence the estimation error of the midpoint, $\widehat{\sigma}_M = (\widehat{\sigma}_L + \widehat{\sigma}_U)/2 = 77.12980$, is less than $(\widehat{\sigma}_L - \widehat{\sigma}_U)/2 = 0.01221$. Since the *ML* estimate $\widehat{\sigma}_0$ is known to lie in the interval $[\widehat{\sigma}_L, \widehat{\sigma}_U]$, it is convenient to employ the rule of the false position to determine $\widehat{\sigma}_0$. The first iteration yields:

$$\widehat{\sigma}_L + \frac{(\widehat{\sigma}_U - \widehat{\sigma}_L)G(\widehat{\sigma}_L)}{G(\widehat{\sigma}_L) - G(\widehat{\sigma}_U)} = 77.13508525,$$

whereas the exact value of the *ML* estimate, $\widehat{\sigma}_0 = 77.13508211$, is obtained in the second iteration. Generally, $\widehat{\sigma}_M$ is a good approximation to $\widehat{\sigma}_0$ (in this case, the exact estimation error of $\widehat{\sigma}_M$ is 0.00529).

The *ML* estimate $\widehat{\sigma}_0$ can also be found by using the *EM* algorithm or the Newton-Raphson method. In such cases, given the current estimate $\widehat{\sigma}_{(j)}$ of σ, the new estimate $\widehat{\sigma}_{(j+1)}$ is defined as

$$\widehat{\sigma}_{(j+1)} = \frac{W}{n} - \frac{q_1 w_r}{\exp\left\{w_r/\widehat{\sigma}_{(j)}\right\} - 1} + (q_1 + q_2)\widehat{\sigma}_{(j)} \text{ or } \widehat{\sigma}_{(j+1)} = \widehat{\sigma}_{(j)} - \frac{G(\widehat{\sigma}_{(j)})}{G'(\widehat{\sigma}_{(j)})},$$

respectively. Under fairly general conditions, the derived sequences converge to $\widehat{\sigma}_0$. Clearly, the midpoint, $\widehat{\sigma}_M$, can be used as an initial estimate in finding $\widehat{\sigma}_0$. In that case, due to the sharpness of our bounds, the convergence is reached in a few iterations.

If, instead, it is assumed an $Exp(\theta_0, \sigma)$ distribution, with $\theta_0 = 5, 10$ and 15, for the data at hand, it follows from theorem 2 that $[71.05799, 71.07287]$, $[64.99105, 64.99858]$ and $[58.91474, 58.91739]$ contain the *ML* estimates of σ, respectively. The corresponding *ML* estimates of σ are found to be 71.06872, 64.99652 and 58.91668. Observe that the *ML* estimate and the lower and upper bounds are very close.

Table 2 contains the *ML* estimates of σ, μ and ξ_p, $p = 0.1, 0.5, 0.9$, and their asymptotic $rmse$ values based on theorem 3 when $\theta_0 = 0, 5, 10$ and 15. The corresponding *BLU* and *BLI* estimates and their $rmse$ values are also included. Notice that *ML* and *BLU* estimation give very close results. Obviously, the *BLI* estimates provide the smallest $rmse$ values.

Table 2. ML, BLU and BLI estimates ($rmse's$) when $\theta = \theta_0$

θ_0	Estimate ($rmse$)	σ	μ	$\xi_{0.1}$	$\xi_{0.5}$	$\xi_{0.9}$
0	ML	77.13508	77.13508	8.126995	53.46598	177.6102
		(25.7196)	(25.7196)	(2.70983)	(17.8275)	(59.2216)
	BLU	77.09113	77.09113	8.122362	53.43550	177.5089
		(25.7207)	(25.7207)	(2.70995)	(17.8283)	(59.2242)
	BLI	69.36922	69.36922	7.308777	48.08308	159.7285
		(21.9547)	(21.9547)	(2.31316)	(15.2178)	(50.5525)
5	ML	71.06872	76.06872	12.48784	54.26108	168.6418
		(23.6969)	(23.6969)	(2.49671)	(16.4254)	(54.5640)
	BLU	71.03621	76.03621	12.48441	54.23855	168.5669
		(23.7006)	(23.7006)	(2.49710)	(16.4280)	(54.5726)
	BLI	63.92080	68.92080	11.73473	49.30652	152.1831
		(20.2303)	(20.2303)	(2.13148)	(14.0226)	(46.5820)
10	ML	64.99652	74.99652	16.84807	55.05216	159.6600
		(21.6722)	(21.6722)	(2.28339)	(15.0220)	(49.9020)
	BLU	64.98129	74.98129	16.84646	55.04160	159.6249
		(21.6804)	(21.6804)	(2.28426)	(15.0277)	(49.9210)
	BLI	58.47237	68.47237	16.16068	50.52996	144.6376
		(18.5059)	(18.5059)	(1.94979)	(12.8273)	(42.6115)
15	ML	58.91668	73.91668	21.20749	55.83793	150.6607
		(19.6449)	(19.6449)	(2.06980)	(13.6168)	(45.2341)
	BLU	58.92636	73.92634	21.20851	55.84464	150.6830
		(19.6602)	(19.6602)	(2.07141)	(13.6274)	(45.2694)
	BLI	53.02395	68.02395	20.58663	51.75340	137.0921
		(16.7816)	(16.7816)	(1.76811)	(11.6321)	(38.6410)

Finally, suppose now that one wants to adopt the Bayesian perspective, and assume that prior knowledge about σ does not exist. As mentioned earlier, Jeffreys' prior is deemed appropriate in this case due to our prior state of ignorance about σ. Assuming this default prior (i.e., $a = b = 0$) and also that $\theta_0 = 0$, the *HPD* estimate of σ is found to be $\widetilde{\sigma}_{(0,0)} = 69.4073$. Evidently, the *HPD* estimates of μ and ξ_p coincide with $\theta_0 + \widetilde{\sigma}_{(0,0)}$ and $\theta_0 + \widetilde{\sigma}_{(0,0)} \log(1-p)$, respectively. In particular, the *HPD* estimates of $\xi_{0.1}$, $\xi_{0.5}$ and $\xi_{0.9}$ are 7.31279, 48.10947 and

159.81621, respectively. Moreover, the *HPD* estimate of F_{t_0} is given by $\widetilde{F}_{t_0;(0,0)} = 1 - \exp\left\{-\left(t_0 - \theta_0\right)/\widetilde{\sigma}_{(\theta_0 - t_0, -2)}\right\}$ if $\theta_0 < t_0 < \theta_0 + 671.1$, and equals 1 if $t_0 \geq \theta_0 + 671.1$. For instance, the corresponding *HPD* estimates of F_{t_0} for $t_0 = 50$ and 100 are obtained to be 0.46252 and 0.73983; the respective *ML* estimates are given by 0.47702 and 0.72639.

Acknowledgements. The authors would like to thank the Editor, Associate Editor and referees for valuable comments.

References

Bain LJ (1978) Statistical Analysis of Reliability and Life-Testing Models. Marcel Dekker, New York

Balakrishnan N, Chan PS (1995) Maximum likelihood estimation for the log-gamma distribution under Type II censored samples and associated inference. In: Balakrishnan N (ed) Recent Advances in Life-Testing and Reliability. CRC Press, Inc., pp. 409–421

Bhattacharyya GK (1985) On asymptotics of maximum likelihood and related estimators based on Type II censored data. J. Amer. Statist. Assoc. 80:398–404

David HA (1970) Order Statistics. John Wiley, New York

Dempster AP, Laird NM, Rubin DB (1977) Maximun likelihood from incomplete data via the EM algorithm (with discussion). J. Roy. Statist. Soc. B39:1–38

Elfessi A (1997) Estimation of a linear function of the parameters of an exponential distribution from doubly censored data. Statist. and Probab. Letters 36:251–259

Fernández AJ (2000) Bayesian inference from type II doubly censored Rayleigh data. Statist. and Probab. Letters 48:393–399

Harter HL, Moore AH (1968) Maximum likelihood estimation, from doubly censored samples, of the parameters of the first asymptotic distribution of extreme values. J. Amer. Statist. Assoc. 63:889–901

Jeffrey H. (1961) Theory of Probability. Clarendom press, Oxford

Kambo NS (1978) Maximum likelihood estimators of the location and scale parameters of the exponential distribution from a censored sample. Commun. Statist. A12:1129–1132

Lalitha S, Mishra A (1996) Modified maximum likelihood estimation for Rayleigh distribution. Commun. Statist. A25:389–401

Lawless JF (1982) Statistical Models and Methods for Lifetime Data. John Wiley, New York

Leemis LM, Shih LH (1989) Exponential parameter estimation for data sets containing left and right censored observations. Commun. Statist. B18:1077–1085

Nelson W (1982) Applied Life Data Analysis. John Wiley, New York

Rao CR (1973) Linear Statistical Inference and Its Applications. John Wiley, New York

Raqab MZ (1995) On the maximum likelihood prediction of the exponential distribution based on doubly Type II censored samples. Pak. J. Statist. 11:1–10

Sarhan AE (1955) Estimation of the mean and standard deviation by order statistics, Part. III. Ann. Math. Statist. 26:576–592

Stoer J, Bulirsch R (1983) Introduction to Numerical analysis. Springer, Berlin Heidelberg New York

Tiku ML (1967) A note on estimating the location and scale parameters of the exponential distribution from a censored sample. Aust. J. Statist. 9:48–54

Tiku ML, Tan WY, Balakrishnan N (1986) Robust Inference. Marcel Dekker, New York