Predictive Inference for Future Responses Given a Doubly Censored Sample from a Two Parameter Exponential Distribution

Hafiz M. R. Khan[(1)], **M. Safiul Haq**[(2)], **Serge B. Provost**[(2)]

[(1)] **Department of Mathematical Sciences and Center for Applied Mathematics and Statistics**

New Jersey Institute of Technology, Newark, NJ 07102

[(2)] **Department of Statistical and Actuarial Sciences The University of Western Ontario, London**

Ontario N6A 5B7, Canada

CAMS Report 0405-13, Spring 2005

Center for Applied Mathematics and Statistics

NJIT

ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Statistical Planning and Inference ■■■ (■■■) ■■■–■■■

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Predictive inference for future responses given a doubly censored sample from a two parameter exponential distribution

Hafiz M.R. Khan[a,*], M. Safiul Haq[b], Serge B. Provost[b]

[a]Department of Mathematical Sciences, New Jersey Institute of Technology, University Heights, Rm 211F, Cullimore Hall, Newark, NJ 07102-1982, USA

[b]Department of Statistical and Actuarial Sciences, The University of Western Ontario, London, Ont., Canada N6A 5B7

Received 23 January 2003; accepted 2 December 2004

Abstract

In this paper, we derive the predictive distributions of one and several future responses including their average, on the basis of a type II doubly censored sample from a two parameter exponential life testing model. We also determine the highest predictive density interval for a future response. A numerical example is provided to illustrate the results.

MSC: 62F15; 62N01; 62N05

Keywords: Order statistics; Two parameter exponential distribution; Type II doubly censored sample; Predictive distribution; Highest predictive density (HPD) region

* Corresponding author. Tel.: +1 973 596 5374; fax: +1 973 596 5591.
E-mail address: khan@adm.njit.edu (H.M.R. Khan).

0378-3758/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2004.12.002

1. Introduction

Consider the probability density function of a two parameter exponential distribution given by

$$f(x|\mu,\sigma)=\begin{cases}\dfrac{1}{\sigma}\exp\left(-\dfrac{x-\mu}{\sigma}\right), & x\geqslant\mu;\ \ \sigma>0,\\ 0 \quad \text{elsewhere},\end{cases} \tag{1}$$

where, μ is the location parameter and σ is the scale parameter. Data for survival and reliability analysis, as well as for biomedical and life testing studies have been modelled extensively by the two parameter exponential model.

In practice, usually a number of observations in a sample from some life testing model are below a certain level or while others are above a certain level. Observations below the lower level and above the upper level are often discarded. This is particularly true for data from biomedical studies. For example, suppose that in the study of a pesticide, $x_1,\dots,x_n$ are the survival times in minutes of n termites following their exposure to the pesticide, with $x_1\leqslant\cdots\leqslant x_n$; it may be the case that the first few observations and last few observations are influenced by some factors other than the effect of the pesticide. Hence to make inference about the effect of the pesticide from such a data set, it may be reasonable to remove these observations from the original data set. Such a sample where some data from both the lower and the upper ends are ignored is known as a type II doubly censored sample. Type II doubly censored samples have been considered for instance by Lalitha and Mishra (1996), Kambo (1978), Raqab (1995), Sarhan (1955), Fernandez (2000a, b), Hui (1980) and Ipsen (1949).

Most studies based on type II doubly censored data from the two parameter exponential model involve inference about the parameters. For instance, Kambo (1978) derived the maximum likelihood estimators of the location and scale parameters on the basis of a doubly censored sample, and comparisons were made with the minimum variance unbiased estimators. Fernandez (2000a) obtained maximum likelihood prediction estimator of the kth ordered future response, assuming first that only the scale parameter is unknown and then, that both the location and scale parameters are unknown. He also obtained the maximum likelihood estimators, the best linear unbiased estimators and the best linear invariant estimators of the location and scale parameters. Fernandez (2000b) derived the predictive density of the kth future ordered response from a Rayleigh distribution assuming a type II doubly censored sample in a Bayesian framework. Sarhan (1955) estimated the mean and standard deviation of some populations from singly and doubly censored samples. He compared the variances and efficiencies of the best linear estimators of the mean and standard deviation.

In Section 2, we derive the predictive distributions of one future response from the two parameter exponential life testing model, on the basis of a type II doubly censored sample. A Bayesian approach is used and the predictive distributions are derived by means of a conjugate prior for the scale parameter σ and a non-informative prior for the location parameter μ. We also obtain a 100 $(1-\alpha)$% highest predictive density (PD) region.

The cases of several future responses and their average are treated in Section 3. Finally, a numerical example is provided in Section 4.

2. Predictive inference for a single future response

The predictive density functions of a single future response and of the *i*th ordered response among *m* future responses are respectively derived in Sections 2.1 and 2.4. The HPD interval for a single future response is determined in Section 2.2 and the estimation of the hyperparameter of the prior distribution is discussed in Section 2.3.

2.1. Predictive distribution of a single future response

Let $x_1, \ldots, x_n$ be an ordered random sample of size n from model (1), $x_1 \leqslant \cdots \leqslant x_g$ be the g smallest ordered observations and $x_{\ell+1} \leqslant \cdots \leqslant x_n$ be the $n-\ell$ largest ordered observations. Only the remaining ordered observations, that is, $\mathbf{x} = x_{g+1}, \ldots, x_\ell$ are being used. It is assumed that the sample data are modelled by a two parameter exponential distribution. The likelihood function of μ and σ given the type II doubly censored sample $\mathbf{x} = x_{g+1}, \ldots, x_\ell$ is then

$$\mathrm{L}(\mu, \sigma|\mathbf{x}) \propto \left(F\left(\frac{x_{g+1}-\mu}{\sigma}\right)\right)^{g+1-1} \left\{\prod_{\alpha=g+1}^{\ell} f\left(\frac{x_\alpha - \mu}{\sigma}\right)\right\} \times \left(1 - F\left(\frac{x_\ell - \mu}{\sigma}\right)\right)^{n-\ell} \left(\frac{1}{\sigma}\right)^{\ell-g},$$

where for instance $F\left(\frac{x_{g+1}-\mu}{\sigma}\right) = 1 - \exp\left\{-\frac{x_{g+1}-\mu}{\sigma}\right\}$. Fernandez (2000a) and Raqab (1995) derived the predictive estimate of the kth failure time of the upper unobserved data from a maximum likelihood method. Fernandez (2000b) considered a predictive inference problem for an independent future sample from the Rayleigh distribution using the Bayesian approach with an inverted gamma prior. Following Fernandez' assumptions, we assign a natural conjugate prior to σ and a uniform prior to μ. Furthermore, μ and σ are assumed to be independently distributed.

Thus, the joint prior density of μ and σ is

$$p(\mu, \sigma) \propto \frac{(\theta_0)^{\theta_0}}{\Gamma(\theta_0)} \sigma^{-(\theta_0+1)} \exp\left\{-\frac{\theta_0}{\sigma}\right\}; \quad \mu > 0; \sigma > 0; \theta_0 > 0, \tag{2}$$

where θ_0 is a hyperparameter, which is in fact the shape parameter of the prior density. It is therefore reasonable to assume that θ_0 is non-negative. This prior is analytically convenient and easily interpretable. It also increases the precision of the predictive density. Note that setting $\theta_0 = 0$ results in a non-informative prior density for the parameters. When the prior density is as specified in (2), the posterior density of μ and σ is given by

$$
\begin{aligned}
p(\mu,\sigma|\mathbf{x}) \propto\ & p(\mu,\sigma)\mathrm{L}(\mu,\sigma|\mathbf{x})\\
&= \Psi(\mathbf{x})\sum_{\gamma=0}^{g}(-1)^{\gamma}\binom{g}{\gamma}\sigma^{-(\ell-g+\theta_0+1)}\\
&\quad\times \exp\left\{-\frac{\sum_{\alpha=g+1}^{\ell}(x_\alpha-\mu)+(n-\ell)(x_\ell-\mu)+\gamma(x_{g+1}-\mu)+\theta_0}{\sigma}\right\},
\end{aligned}
$$

where

$$
\begin{aligned}
\Psi(\mathbf{x})^{-1} =\ & \int_{\mu=0}^{x_{g+1}}\int_{\sigma=0}^{+\infty}\sigma^{-(\ell-g+\theta_0+1)}\exp\left\{-\frac{\theta_0}{\sigma}\right\}\left(F\left(\frac{x_{g+1}-\mu}{\sigma}\right)\right)^{g}\\
&\times\left\{\prod_{\alpha=g+1}^{\ell} f\left(\frac{x_\alpha-\mu}{\sigma}\right)\right\}\left(1-F\left(\frac{x_\ell-\mu}{\sigma}\right)\right)^{n-\ell}\mathrm{d}\sigma\,\mathrm{d}\mu.\\
=\ & \sum_{\gamma=0}^{g}(-1)^{\gamma}\binom{g}{\gamma}\frac{\Gamma(\ell-g+\theta_0-1)}{(n+\gamma-g)}\\
&\times\left[\left((n-\ell)x_\ell+\theta_0-(n-g)x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0-1)}\right.\\
&\left.-\left((n-\ell)x_\ell+\gamma x_{g+1}+\theta_0+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0-1)}\right].
\end{aligned}
$$

Let y_f be a future response from the model specified by (1). Naturally, y_f is independent of the observed data. The predictive density of y_f for a given set of doubly censored data $\mathbf{x}=x_{g+1},\ldots,x_\ell$, denoted by $p(y_f|\mathbf{x})$, is as given below.

For $y_f<x_{g+1};\ \ 0<\mu<y_f;\ \ \sigma>0$,

$$
\begin{aligned}
p(y_f|\mathbf{x}) =\ & \int_{\mu=0}^{y_f}\int_{\sigma=0}^{+\infty} p(y_f|\mu,\sigma)\,p(\mu,\sigma|\mathbf{x})\,\mathrm{d}\sigma\,\mathrm{d}\mu\\
=\ & \Psi(\mathbf{x})\int_{\mu=0}^{y_f}\int_{\sigma=0}^{+\infty}\frac{1}{\sigma}\exp\left\{-\frac{(y_f-\mu)}{\sigma}\right\}\sigma^{-(\ell-g+\theta_0+1)}\exp\left\{-\frac{\theta_0}{\sigma}\right\}\\
&\times\left(F\left(\frac{x_{g+1}-\mu}{\sigma}\right)\right)^{g}\left\{\prod_{\alpha=g+1}^{\ell} f\left(\frac{x_\alpha-\mu}{\sigma}\right)\right\}\\
&\times\left(1-F\left(\frac{x_\ell-\mu}{\sigma}\right)\right)^{n-\ell}\mathrm{d}\sigma\,\mathrm{d}\mu
\end{aligned}
$$

$$
\begin{aligned}
&= \Psi(\mathbf{x}) \sum_{\gamma=0}^{g} (-1)^{\gamma} \binom{g}{\gamma} \frac{\Gamma(\ell - g + \theta_0)}{(n + \gamma - g + 1)} \\
&\times \left[\left(\theta_0 + (n-\ell)x_\ell + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha - (n+\gamma-g)y_f \right)^{-(\ell-g+\theta_0)} \right. \\
&\left. - \left(\theta_0 + y_f + (n-\ell)x_\ell + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha \right)^{-(\ell-g+\theta_0)} \right],
\end{aligned}
$$

and for $y_f > x_{g+1}$; $0 < \mu < x_{g+1}$; $\sigma > 0$,

$$
\begin{aligned}
p(y_f|\mathbf{x}) &= \int_{\mu=0}^{x_{g+1}} \int_{\sigma=0}^{+\infty} p(y_f|\mu,\sigma)\, p(\mu,\sigma|\mathbf{x})\, \mathrm{d}\sigma\, \mathrm{d}\mu \\
&= \Psi(\mathbf{x}) \int_{\mu=0}^{x_{g+1}} \int_{\sigma=0}^{+\infty} \frac{1}{\sigma} \exp\left\{ -\frac{(y_f-\mu)}{\sigma} \right\} \sigma^{-(\ell-g+\theta_0+1)} \exp\left\{ -\frac{\theta_0}{\sigma} \right\} \\
&\times \left(F\left(\frac{x_{g+1}-\mu}{\sigma} \right) \right)^{g} \left\{ \prod_{\alpha=g+1}^{\ell} f\left(\frac{x_\alpha - \mu}{\sigma} \right) \right\} \\
&\times \left(1 - F\left(\frac{x_\ell - \mu}{\sigma} \right) \right)^{n-\ell} \mathrm{d}\sigma\, \mathrm{d}\mu \\
&= \Psi(\mathbf{x}) \sum_{\gamma=0}^{g} (-1)^{\gamma} \binom{g}{\gamma} \frac{\Gamma(\ell - g + \theta_0)}{(n + \gamma - g + 1)} \\
&\times \left[\left(y_f + \theta_0 + (n-\ell)x_\ell + \sum_{\alpha=g+1}^{\ell} x_\alpha - (n-g+1)x_{g+1} \right)^{-(\ell-g+\theta_0)} \right. \\
&\left. - \left(y_f + \theta_0 + (n-\ell)x_\ell + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha \right)^{-(\ell-g+\theta_0)} \right],
\end{aligned}
$$

$p(y_f|\mathbf{x})$ being equal to zero elsewhere.

It was verified that $\Pr(\mathrm{Y}_f \leqslant x_{g+1}) + \Pr(\mathrm{Y}_f \geqslant x_{g+1}) = 1$. When $\theta_0 = 0$, the above results coincide with those obtained from a non-informative prior. For $\theta_0 = 0$, $g = 0$, and $\ell = r$, this distribution is same as the predictive distribution of a right censored sample. Assuming an unrestricted location parameter, that is, $-\infty < \mu < \infty$, Dunsmore (1974) derived the predictive distribution of an ordered future response given a type II censored sample.

2.2. Highest predictive density (HPD) interval

For inference purposes, one may be interested in determining the HPD interval of a future response y_f. An HPD interval $\mathscr{S}$ is of the form $\mathscr{S} = \{y_f : p(y_f|\mathbf{x}) \geqslant s_\alpha\}$, where s_α is the largest constant such that $\Pr(y_f \in \mathscr{S}|\mathbf{x}) = 1 - \alpha$ and α denotes the significance level, see for instance Box and Tiao (1973). We note that the predictive density derived in Section 2 is composed of two parts: one that increases from zero to x_{g+1} and the other which decreases from x_{g+1} to plus infinity. Thus, the HPD interval $[s_1, s_2]$ for y_f must simultaneously satisfy

$$\Pr(s_1 \leqslant y_f \leqslant s_2) = 1 - \alpha \quad \text{and} \quad p(s_1|\mathbf{x}) = p(s_2|\mathbf{x}).$$

In our case, we have $\Pr(s_1 < y_f < x_{g+1}) + \Pr(x_{g+1} < y_f < s_2) = 1 - \alpha$, where s_1 and s_2 are to be chosen so that $p(s_1|\mathbf{x}) = p(s_2|\mathbf{x})$. For arbitrary s_1 and s_2:

$$\begin{aligned}
&\Pr(s_1 < y_f < x_{g+1}) \\
&= \int_{y_f=s_1}^{x_{g+1}} \Psi(\mathbf{x}) \sum_{\gamma=0}^{g} (-1)^\gamma \binom{g}{\gamma} \frac{\Gamma(\ell - g + \theta_0)}{(n + \gamma - g + 1)} \\
&\quad \times \left[\left(\theta_0 + (n-\ell)x_\ell + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha - (n + \gamma - g) y_f \right)^{-(\ell - g + \theta_0)} \right. \\
&\quad \left. - \left(y_f + \theta_0 + (n-\ell)x_\ell + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha \right)^{-(\ell - g + \theta_0)} \right] \mathrm{d}y_f
\end{aligned}$$

and

$$\begin{aligned}
&\Pr(x_{g+1} < y_f < s_2) \\
&= \int_{y_f=x_{g+1}}^{s_2} \Psi(\mathbf{x}) \sum_{\gamma=0}^{g} (-1)^\gamma \binom{g}{\gamma} \frac{\Gamma(\ell - g + \theta_0)}{(n + \gamma - g + 1)} \\
&\quad \times \left[\left(y_f + \theta_0 + (n-\ell)x_\ell + \sum_{\alpha=g+1}^{\ell} x_\alpha - (n - g + 1) x_{g+1} \right)^{-(\ell - g + \theta_0)} \right. \\
&\quad \left. - \left(y_f + \theta_0 + (n-\ell)x_\ell + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha \right)^{-(\ell - g + \theta_0)} \right] \mathrm{d}y_f.
\end{aligned}$$

A closed form solution for s_1 and s_2 is not tractable. However, for a specific set of data, a solution can be determined as shown in Section 4.

2.3. Estimation of the hyperparameter of the prior distribution

The hyperparameter of the prior distributions is assumed to be unknown. The hyperparameter may be estimated from the data set by applying the method of moments or the maximum likelihood method, see Geisser (1993, p. 65). From (1) and (2), the likelihood function $\mathrm{L}(\theta_0|\mathbf{x})$ is obtained as

$$\begin{aligned}
\mathrm{L}(\theta_0|\mathbf{x}) &= \int_{\mu=0}^{x_{g+1}}\int_{\sigma=0}^{+\infty} p(\mathbf{x}|\mu,\sigma)p(\mu,\sigma)\,\mathrm{d}\sigma\,\mathrm{d}\mu \\
&= \int_{\mu=0}^{x_{g+1}}\int_{\sigma=0}^{+\infty} \frac{n!}{g!(n-\ell)!}\left(F\left(\frac{x_{g+1}-\mu}{\sigma}\right)\right)^{g+1-1} \\
&\quad\times\left\{\prod_{\alpha=g+1}^{\ell} f\left(\frac{x_\alpha-\mu}{\sigma}\right)\right\} \\
&\quad\times\left(1-F\left(\frac{x_\ell-\mu}{\sigma}\right)\right)^{n-\ell}\left(\frac{1}{\sigma}\right)^{\ell-g}\frac{(\theta_0)^{\theta_0}}{\Gamma(\theta_0)}\sigma^{-(\theta_0+1)}\exp\left\{-\frac{\theta_0}{\sigma}\right\}\mathrm{d}\sigma\,\mathrm{d}\mu \\
&= \frac{n!}{g!(n-\ell)!}\sum_{\gamma=0}^{g}(-1)^\gamma\binom{g}{\gamma}\frac{\Gamma(\ell-g+\theta_0-1)(\theta_0^{\theta_0})}{(n+\gamma-g)\Gamma(\theta_0)} \\
&\quad\times\left[\left((n-\ell)x_\ell+\theta_0-(n-g)x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0-1)}\right. \\
&\quad\left.-\left((n-\ell)x_\ell+\gamma x_{g+1}+\theta_0+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0-1)}\right].
\end{aligned}$$

The maximum likelihood estimate of θ_0 is obtained by equating the derivative of $\mathrm{L}(\theta_0|\mathbf{x})$ with respect to θ_0 to zero. There is no closed form solution for $\hat{\theta}_0$; however, a numerical method such as the Newton–Raphson iterative algorithm may be used to solve for $\hat{\theta}_0$. The iterative value of $\hat{\theta}_0$ may be substituted in $p(y_f|\mathbf{x})$ as obtained in Section 2.1 to determine the predictive density of a single future response.

2.4. Predictive distribution for the ith ordered future response

Consider the life testing model (1) and let y_i, $i=1,2,\ldots,m$ be the ith ordered future response in a set of m future responses. The density function of the ith ordered future response is then

$$p(y_i|\mu,\sigma)=\frac{m!}{(i-1)!(m-i)!}[F(y_i)]^{i-1}[1-F(y_i)]^{m-i}f(y_i)$$
$$=\frac{1}{\sigma}\sum_{j=0}^{i-1}\frac{(-1)^j\binom{i-1}{j}}{\beta(m-i+1,i)}\exp\left\{-\frac{(m-i+j+1)(y_i-\mu)}{\sigma}\right\},$$

where $\beta(\alpha,\delta)=\frac{\Gamma(\alpha)\Gamma(\delta)}{\Gamma(\alpha+\delta)}$, and the predictive density of y_i given a type II doubly censored sample is given by

$$p(y_i|\mathbf{x})=\int_{\mu=0}^{\min(x_{g+1},y_i)}\int_{\sigma=0}^{+\infty}p(y_i|\mu,\sigma)p(\mu,\sigma|\mathbf{x})\,\mathrm{d}\sigma\,\mathrm{d}\mu$$
$$=\int_{\mu=0}^{\min(x_{g+1},y_i)}\int_{\sigma=0}^{+\infty}\frac{1}{\sigma}\sum_{j=0}^{i-1}\frac{(-1)^j\binom{i-1}{j}}{\beta(m-i+1,i)}$$
$$\times\exp\left\{-\frac{(m-i+j+1)(y_i-\mu)}{\sigma}\right\}$$
$$\times\Psi(\mathbf{x})\sigma^{-(\ell-g+\theta_0+1)}\exp\left\{-\frac{\theta_0}{\sigma}\right\}\left(F\left(\frac{x_{g+1}-\mu}{\sigma}\right)\right)^g$$
$$\times\prod_{\alpha=g+1}^{\ell}f\left(\frac{x_\alpha-\mu}{\sigma}\right)\left(1-F\left(\frac{x_\ell-\mu}{\sigma}\right)\right)^{n-\ell}\mathrm{d}\sigma\,\mathrm{d}\mu$$
$$=\Psi(\mathbf{x})\Gamma(\ell-g+\theta_0)\frac{\sum_{\gamma=0}^{g}\sum_{j=0}^{i-1}(-1)^{j+\gamma}\binom{i-1}{j}\binom{g}{\gamma}}{\beta(m-i+1,i)(m+n+\gamma+j-i-g+1)}$$
$$\times\left[\left((n-\ell)x_\ell+(m-i+j+1)y_i+\theta_0+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha\right.\right.$$
$$\left.-(m+n+\gamma+j-i-g+1)\min(x_{g+1},y_i)\right)^{-(\ell-g+\theta_0)}$$
$$-\left((n-\ell)x_\ell+(m-i+j+1)y_i+\theta_0+\gamma x_{g+1}\right.$$
$$\left.\left.+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0)}\right].$$

The predictive density for the ith ordered future response among m future responses given a set of doubly censored sample $\mathbf{x}$ is thus given by

$$p(y_i|\mathbf{x}) = \begin{cases} \Psi(\mathbf{x})\Gamma(\ell-g+\theta_0)\displaystyle\sum_{\gamma=0}^{g}\sum_{j=0}^{i-1}\frac{(-1)^{j+\gamma}\binom{i-1}{j}\binom{g}{\gamma}}{(m+n+\gamma+j-i-g+1)\beta(m-i+1,i)} \\ \times\left[\left((n-\ell)x_\ell+(m-i+j+1)y_i+\theta_0+\displaystyle\sum_{\alpha=g+1}^{\ell}x_\alpha\right.\right. \\ \left.-(m+n+j-i-g+1)x_{g+1}\right)^{-(\ell-g+\theta_0)} \\ -\left((n-\ell)x_\ell+(m-i+j+1)y_i+\theta_0+\gamma x_{g+1}\right. \\ \left.\left.+\displaystyle\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0)}\right] \\ \text{for } y_i > x_{g+1}, \\ \Psi(\mathbf{x})\Gamma(\ell-g+\theta_0)\displaystyle\sum_{\gamma=0}^{g}\sum_{j=0}^{i-1}\frac{(-1)^{j+\gamma}\binom{i-1}{j}\binom{g}{\gamma}}{(m+n+\gamma+j-i-g+1)\beta(m-i+1,i)} \\ \times\left[\left((n-\ell)x_\ell+\theta_0+\gamma x_{g+1}+\displaystyle\sum_{\alpha=g+1}^{\ell}x_\alpha-(n+\gamma-g)y_i\right)^{-(\ell-g+\theta_0)}\right. \\ -\left((n-\ell)x_\ell+(m-i+j+1)y_i+\theta_0+\gamma x_{g+1}\right. \\ \left.\left.+\displaystyle\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0)}\right] \\ \text{for } y_i \leqslant x_{g+1}, \\ 0 \quad \text{elsewhere.} \end{cases} \tag{3}$$

It was verified that $\Pr(Y_i \leqslant x_{g+1}) + \Pr(Y_i \geqslant x_{g+1}) = 1$. This result coincides with that obtained by Evans and Nigm (1980) for a non-negative location parameter when $\theta_0=0$, $g=0$, $\ell=r=G+1$, $n=C$, $x_1=B$, $(n-r)x_r+\sum_{i=1}^{r}x_i-nx_1=H$, $i=K$ and $m=N$.

For example, in a life testing experiment involving m computers, one may be interested in making inference about the lifetime of the ith computer to fail, given the data (type II doubly censored). Letting $i=1$ in (3) will yield the distribution of the earliest failure time while letting $i=m$ will provide that of the latest.

3. Predictive inference for several future responses

First we give the joint density function of m future responses. The predictive distribution of their average is then derived.

3.1. Predictive distribution for m future responses

Let $y_1,\ldots,y_m$ be m ordered future responses from model (1). Then, the density function of $\mathbf{y}=y_1,\ldots,y_m$ given μ and σ, is

$$p(\mathbf{y}|\mu,\sigma)=m!\prod_{i=1}^{m}p(y_i|\mu,\sigma)$$

$$=m!\frac{1}{\sigma^m}\exp\left(-\frac{\sum_{i=1}^{m}(y_i-\mu)}{\sigma}\right);\qquad \mu\leqslant y_1\leqslant\cdots\leqslant y_m\leqslant\infty.$$

The predictive density of $y_1,\ldots,y_m$ is defined as

$$p(y_1,\ldots,y_m|\mathbf{x})$$

$$=\int_{\mu=0}^{\min(x_{g+1},y_1)}\int_{\sigma=0}^{+\infty}p(y_1,\ldots,y_m|\mu,\sigma)p(\mu,\sigma|\mathbf{x})\,\mathrm{d}\sigma\,\mathrm{d}\mu$$

$$=\Psi(\mathbf{x})m!\sum_{\gamma=0}^{g}(-1)^{\gamma}\binom{g}{\gamma}\int_{\mu=0}^{\min(x_{g+1},y_1)}\int_{\sigma=0}^{+\infty}\sigma^{-(m+\ell-g+\theta_0+1)}$$

$$\times\exp\left(-\frac{\sum_{i=1}^{m}(y_i-\mu)}{\sigma}\right)$$

$$\times\exp\left\{-\frac{\sum_{\alpha=g+1}^{\ell}(x_\alpha-\mu)+(n-\ell)(x_\ell-\mu)+\gamma(x_{g+1}-\mu)+\theta_0}{\sigma}\right\}\mathrm{d}\sigma\,\mathrm{d}\mu,$$

$$=\Psi(\mathbf{x})m!\Gamma(m+\ell-g+\theta_0)\sum_{\gamma=0}^{g}(-1)^{\gamma}\binom{g}{\gamma}\int_{\mu=0}^{\min(x_{g+1},y_1)}\left(\sum_{i=1}^{m}y_i+(n-\ell)x_\ell\right.$$

$$\left.+\gamma x_{g+1}+\theta_0+\sum_{\alpha=g+1}^{\ell}x_\alpha-(m+n+\gamma-g)\mu\right)^{-(m+\ell-g+\theta_0)}\mathrm{d}\mu.$$

Thus the predictive density of $y_1, \ldots, y_m$ given $\mathbf{x} = x_{g+1}, \ldots, x_\ell$, is

$$p(y_1, \ldots, y_m|\mathbf{x}) = \begin{cases} \Psi(\mathbf{x}) m! \Gamma(m+\ell-g+\theta_0-1) \sum_{\gamma=0}^{g} \dfrac{(-1)^\gamma \binom{g}{\gamma}}{(m+n+\gamma-g)} \\ \times \left[\left((n-\ell)x_\ell + \sum_{i=1}^{m} y_i + \theta_0 + \sum_{\alpha=g+1}^{\ell} x_\alpha \right.\right. \\ \left. -(m+n-g)x_{g+1} \right)^{-(m+\ell-g+\theta_0-1)} \\ \left. -\left((n-\ell)x_\ell + \sum_{i=1}^{m} y_i + \theta_0 + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha \right)^{-(m+\ell-g+\theta_0-1)} \right] \\ \text{for } y_m \geqslant \cdots \geqslant y_1 \geqslant x_{g+1}, \\ \Psi(\mathbf{x}) m! \Gamma(m+\ell-g+\theta_0-1) \sum_{\gamma=0}^{g} \dfrac{(-1)^\gamma \binom{g}{\gamma}}{(m+n+\gamma-g)} \\ \times \left[\left((n-\ell)x_\ell + \sum_{i=1}^{m} y_i + \theta_0 + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha \right.\right. \\ \left. -(m+n+\gamma-g)y_1 \right)^{-(m+\ell-g+\theta_0-1)} \\ \left. -\left((n-\ell)x_\ell + \sum_{i=1}^{m} y_i + \theta_0 + \gamma x_{g+1} + \sum_{\alpha=g+1}^{\ell} x_\alpha \right)^{-(m+\ell-g+\theta_0-1)} \right] \\ \text{for } y_1 \leqslant \cdots \leqslant y_m \leqslant x_{g+1}, \\ 0 \quad \text{elsewhere.} \end{cases} \tag{4}$$

3.2. *Predictive distribution for the average of the future responses*

One may be interested in making inference about the average of the future responses from several devices. Hui (1980) employed the structural approach for deriving the average predictive density of future failure times given a type II doubly censored sample. We make use of a Bayesian approach to obtain the predictive distribution of the average of a sample of future responses.

Let $y_{f1}, \ldots, y_{fm}$ be the m independent future responses from model (1). Then the probability density function of the average future failure time denoted by $\bar{y}_f$, is given by

$$p(\bar{y}_f|\mu,\sigma)=\begin{cases}\dfrac{(m)^m(\bar{y}_f-\mu)^{m-1}}{\Gamma(m)\sigma^m}\exp\left\{-\dfrac{m(\bar{y}_f-\mu)}{\sigma}\right\}, & \bar{y}_f\geqslant\mu;\ \mu\geqslant 0,\ \sigma>0,\\ 0\quad\text{elsewhere.}\end{cases}$$

The predictive density function for the average future response $\bar{y}_f$ given $\mathbf{x}$, is then

$$\begin{aligned}
&p(\bar{y}_f|\mathbf{x})\\
&=\int_{\mu=0}^{\min(x_{g+1},\bar{y}_f)}\int_{\sigma=0}^{+\infty}p(\bar{y}_f|\mu,\sigma)p(\mu,\sigma|\mathbf{x})\,\mathrm{d}\sigma\,\mathrm{d}\mu\\
&=\Psi(\mathbf{x})\int_{\mu=0}^{\min(x_{g+1},\bar{y}_f)}\int_{\sigma=0}^{+\infty}\frac{(m)^m(\bar{y}_f-\mu)^{m-1}}{\Gamma(m)\sigma^m}\\
&\quad\times\exp\left\{-\frac{m(\bar{y}_f-\mu)}{\sigma}\right\}\sigma^{-(\ell-g+\theta_0+1)}\exp\left\{-\frac{\theta_0}{\sigma}\right\}\left(F\left(\frac{x_{g+1}-\mu}{\sigma}\right)\right)^g\\
&\quad\times\prod_{\alpha=g+1}^{\ell}f\left(\frac{x_\alpha-\mu}{\sigma}\right)\left(1-F\left(\frac{x_\ell-\mu}{\sigma}\right)\right)^{n-\ell}\mathrm{d}\sigma\,\mathrm{d}\mu\\
&=\Psi(\mathbf{x})\frac{m^m\Gamma(m+\ell-g+\theta_0)}{\Gamma(m)}\sum_{i=0}^{m-1}\sum_{\gamma=0}^{g}(-1)^{\gamma+i}\binom{g}{\gamma}\binom{m-1}{i}(\bar{y}_f)^{m-i-1}\\
&\quad\times\left(m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(m+\ell-g+\theta_0)}\int_{\mu=0}^{\min(x_{g+1},\bar{y}_f)}\mu^i\\
&\quad\times\left(1-\frac{(m+n+\gamma-g)\mu}{m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha}\right)^{-(m+\ell-g+\theta_0)}\mathrm{d}\mu.
\end{aligned}$$

To evaluate this integral, we consider the following cases:
Case 1: $\bar{y}_f\geqslant x_{g+1}$
Let

$$z=\frac{(m+n+\gamma-g)\mu}{m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha}\quad\text{and}$$

$$q=\frac{(m+n+\gamma-g)x_{g+1}}{m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha}.$$

Then,

$$
\begin{aligned}
&p(\bar{y}_f|\mathbf{x})\\
&\quad=\Psi(\mathbf{x})\frac{m^m\Gamma(m+\ell-g+\theta_0)}{\Gamma(m)}\sum_{i=0}^{m-1}\sum_{\gamma=0}^{g}(-1)^{\gamma+i}\binom{g}{\gamma}\binom{m-1}{i}(\bar{y}_f)^{m-i-1}\\
&\qquad\times\left(m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(m+\ell-g+\theta_0-i-1)}\\
&\qquad\times(m+n+\gamma-g)^{-(i+1)}\int_{z=0}^{q}z^i(1-z)^{-(m+\ell-g+\theta_0)}\,\mathrm{d}z\\
&\quad=\Psi(\mathbf{x})\frac{m^m\Gamma(m+\ell-g+\theta_0)}{\Gamma(m)}\sum_{i=0}^{m-1}\sum_{\gamma=0}^{g}(-1)^{\gamma+i}\binom{g}{\gamma}\binom{m-1}{i}(\bar{y}_f)^{m-i-1}\\
&\qquad\times\left(m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(m+\ell-g+\theta_0-i-1)}\\
&\qquad\times\frac{(m+n+\gamma-g)^{-(i+1)}}{(1+i)}\left({}_2F_1\left[1+i,\ell-g+m+\theta_0,2+i,\right.\right.\\
&\qquad\left.\frac{(-g+m+n+\gamma)\,x_{g+1}}{m\bar{y}_f+\theta_0+\gamma x_{1+g}+(n-\ell)\,x_\ell+\sum_{\alpha=g+1}^{\ell}x_\alpha}\right]\frac{1}{(g-m-n-\gamma)}\\
&\qquad\left.\times\left(\frac{m\bar{y}_f+\theta_0+\gamma x_{1+g}+(n-\ell)x_\ell+\sum_{\alpha=g+1}^{\ell}x_\alpha}{(-g+m+n+\gamma)x_{g+1}}\right)^{-(1+i)}\right),\quad \bar{y}_f\geqslant x_{g+1},
\end{aligned}
$$

where ${}_2F_1(\eta,\xi;\zeta;z)=\sum_{k=0}^{\infty}z^k\left((\eta)_k(\xi)_k/(\zeta)_k\right)/k!$ and $(\eta)_k=\Gamma(\eta+k)/\Gamma(\eta)$.

Case 2: $\bar{y}_f\leqslant x_{g+1}$

Let

$$
z=\frac{(m+n+\gamma-g)\mu}{m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha}\quad\text{and}
$$

$$
s=\frac{(m+n+\gamma-g)\bar{y}_f}{m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha}.
$$

Then

$$
\begin{aligned}
&p(\bar{y}_f|\mathbf{x})\\
&= \Psi(\mathbf{x})\frac{m^m\Gamma(m+\ell-g+\theta_0)}{\Gamma(m)}\sum_{i=0}^{m-1}\sum_{\gamma=0}^{g}(-1)^{\gamma+i}\binom{g}{\gamma}\binom{m-1}{i}(\bar{y}_f)^{m-i-1}\\
&\quad\times\left(m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0+m-i-1)}\\
&\quad\times(m+n+\gamma-g)^{-(i+1)}\int_{z=0}^{s}z^i(1-z)^{-(\ell-g+\theta_0+m)}\,\mathrm{d}z\\
&= \Psi(\mathbf{x})\frac{m^m\Gamma(m+\ell-g+\theta_0)}{\Gamma(m)}\sum_{i=0}^{m-1}\sum_{\gamma=0}^{g}(-1)^{\gamma+i}\binom{g}{\gamma}\binom{m-1}{i}(\bar{y}_f)^{m-i-1}\\
&\quad\times\left(m\bar{y}_f+\theta_0+(n-\ell)x_\ell+\gamma x_{g+1}+\sum_{\alpha=g+1}^{\ell}x_\alpha\right)^{-(\ell-g+\theta_0+m-i-1)}\\
&\quad\times\frac{(m+n+\gamma-g)^{-(i+1)}}{(1+i)}\left({}_2F_1\left[1+i,\ell-g+m+\theta_0,2+i,\right.\right.\\
&\quad\left.\frac{(-g+m+n+\gamma)\bar{y}_f}{m\bar{y}_f+\theta_0+\gamma x_{1+g}+(n-\ell)x_\ell+\sum_{\alpha=g+1}^{\ell}x_\alpha}\right]\frac{1}{(g-m-n-\gamma)}\\
&\quad\times\left.\left(\frac{m\bar{y}_f+\theta_0+\gamma x_{1+g}+(n-\ell)x_\ell+\sum_{\alpha=g+1}^{\ell}x_\alpha}{(-g+m+n+\gamma)\bar{y}_f}\right)^{-(1+i)}\right),\quad \bar{y}_f\leqslant x_{g+1}.
\end{aligned}
$$

When $\theta_0=0$, the above result agrees with that of Hui (1980), who followed the structural approach of Fraser (1968).

4. A numerical example

We consider a data set which was analyzed by Lawless (1982, p. 138). The data represents the failure times, in minutes, of an electrical insulation device and was modelled as a two parameter exponential distribution:

-, -, 24.4, 28.6, 43.2, 46.9, 70.7, 75.3, 95.5, -, -, -.

The first two observations are censored due to the fact that experimenter failed to observe these failure times and the experiment was stopped before the failure times of the last three devices were observed. Making use of these data, Fernandez (2000a) and Raqab (1995) have estimated the maximum likelihood predictors of the 11th and 12th failure times of the upper unobserved data. We used these data to derive the predictive distribution of the first future response y_f. We also considered the predictive distributions of the average of the future responses as well as that of type II doubly censored future responses.

In this case, $n=12$, $\ell=9$, $g=2$, $x_\ell=95.5$, $x_{g+1}=24.4$, $\sum_{\alpha=3}^{9} x_\alpha=384.60$, $\theta_0=0.2328$ and $\Psi(\mathbf{x}) = 1.5882 \times 10^{15}$. One may be interested in the distribution of the first future response y_f. The density function of y_f is given by

$$p(y_f|\mathbf{x}) = \begin{cases} \Psi(\mathbf{x}) \displaystyle\sum_{\gamma=0}^{g} (-1)^\gamma \binom{g}{\gamma} \frac{\Gamma(9-2+0.2328)}{(12+\gamma-2+1)} [(0.2328+(12-9)95.5+24.4\gamma \\ \quad +384.60-(12+\gamma-2)y_f)^{-(9-2+0.2328)} \\ \quad -(0.2328+y_f+(12-9)95.5+24.4\gamma \\ \quad +384.60)^{-(9-2+0.2328)}] \quad \text{for } y_f \leqslant 24.4, \\ \Psi(\mathbf{x}) \displaystyle\sum_{\gamma=0}^{g} (-1)^\gamma \binom{g}{\gamma} \frac{\Gamma(9-2+0.2328)}{(12+\gamma-2+1)} \\ \quad \times[(y_f+0.2328+(12-9)95.5+384.60 \\ \quad -(12-2+1)24.4)^{-(9-2+0.2328)} \\ \quad -(y_f+0.2328+(12-9)95.5 \\ \quad +24.4\gamma+384.60)^{-(9-2+0.2328)}] \quad \text{for } y_f > 24.4, \\ 0 \quad \text{elsewhere.} \end{cases}$$

By setting $p(s_1|\mathbf{x}) = p(s_2|\mathbf{x})$, where $p(.|\mathbf{x})$ is given in Section 2.1 and aiming for $\Pr(s_1 \leqslant y_f \leqslant s_2) = 0.95$, we determined that in this case the 95% HPD (as well as the shortest credible) interval for a single future response is specified by $s_1 = 5.78263$ and $s_2 = 273.738$. The density of y_f is plotted in Fig. 1 where the 95% HPD interval corresponds to the unshaded region. The effect of the hyperparameter θ_0 on the mean and variance

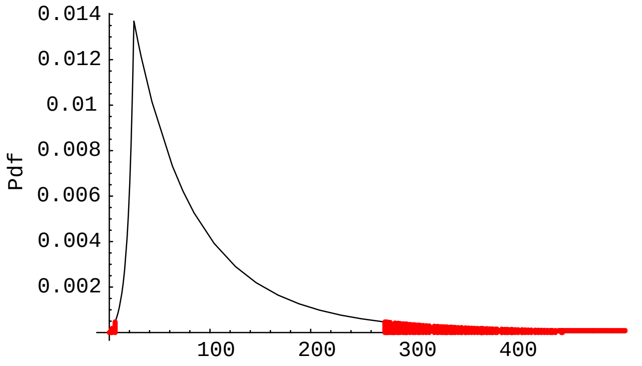

Fig. 1. Predictive density for the first future response with HPD interval, where 95% HPD interval corresponds to the unshaded region.

Table 1
Sensitivity of the mean and standard deviation of the predictive distribution with respect to various values of the hyperparameter θ_0

Values of θ_0	Mean (y_f)	SD (y_f)
0.2328	93.100	78.450
0.5000	17.270	48.660
0.7500	3.570	22.880
1.0000	0.740	10.370
1.5000	0.032	2.096
2.0000	0.001	0.424

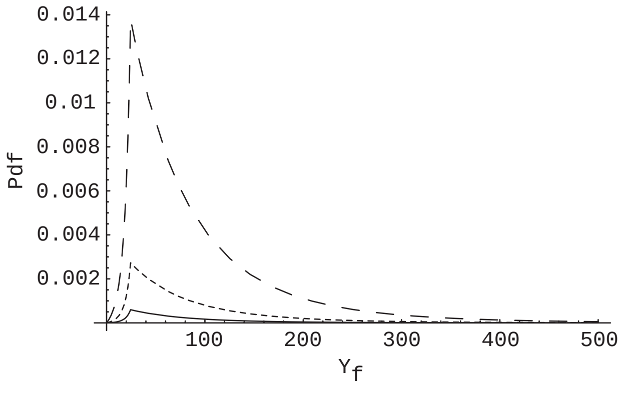

Fig. 2. Predictive densities for certain values of the hyperparameter, namely, $\theta_0=0.2328$ (long dashes), $\theta_0=0.5000$ (short dashes) and $\theta_0=0.7500$ (solid line).

of the predictive distribution is illustrated in Table 1. The predictive densities for certain values of hyperparameter are plotted in Fig. 2. In practice, the hyperparameter could be estimated from information available from previous experiments.

Let $m=5$ in the representation of the predictive density of the average of m future responses derived in Section 3.2; then, the predictive density of $\bar{y}_f$, given a doubly censored sample $\mathbf{x}$, is obtained as

$$p(\bar{y}_f|\mathbf{x}) \propto \begin{cases} \sum_{i=0}^{5-1}\sum_{\gamma=0}^{2}(-1)^{\gamma+i}\binom{2}{\gamma}\binom{5-1}{i}(\bar{y}_f)^{5-i-1} \\ \quad \times\big[(5\bar{y}_f+0.2328+(12-9)95.5+24.4\gamma \\ \quad +384.60)^{-(5+9-2+0.2328-i-1)} \\ \quad \times(5+12+\gamma-2)^{-(i+1)}\int_{z=0}^{\frac{(5+12+\gamma-2)\min(24.4,\bar{y}_f)}{5\bar{y}_f+0.2328+(12-9)95.5+24.4\gamma+384.60}} \\ \quad \times z^i(1-z)^{-(5+9-2+0.2328)}\,\mathrm{d}z\big], \quad \text{for } 0\leqslant \min(24.4,\bar{y}_f)\leqslant\infty, \\ 0 \quad \text{elsewhere.} \end{cases}$$

This predictive density is plotted in Fig. 3 with the appropriate normalizing constant.

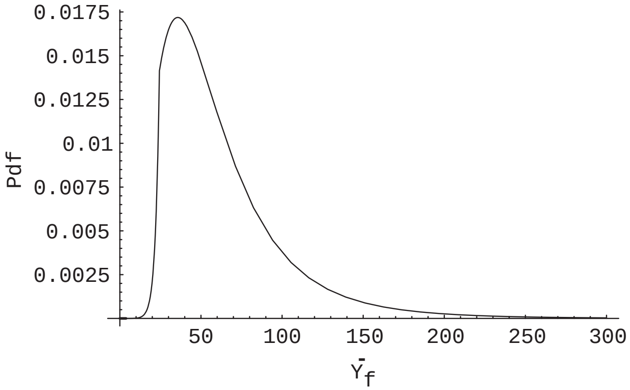

Fig. 3. Predictive density for the average of five future responses.

Acknowledgements

The first author thanks the University of Western Ontario for awarding him a Research Assistantship and a Special University Scholarship. The second and third authors gratefully acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada. The authors are also grateful to an Associate Editor and a referee for helpful comments.

References

Box, G.E.P., Tiao, G.C., 1973. Bayesian Inference in Statistical Analysis. Addison-Wesley Publishing Company Inc., Ontario.

Dunsmore, I.R., 1974. The Bayesian predictive distribution in life testing models. Technometrics 16, 455–460.

Evans, I.G., Nigm, A.M., 1980. Bayesian prediction for the left truncated exponential distribution. Technometrics 22, 201–204.

Fernandez, A.J., 2000a. On maximum likelihood prediction based on type II doubly censored exponential data. Metrika 50, 211–220.

Fernandez, A.J., 2000b. Bayesian inference from type II doubly censored Rayleigh data. Statist. Probab. Lett. 48, 393–399.

Fraser, D.A.S., 1968. The Structure of Inference. Wiley, New York.

Geisser, S., 1993. Predictive Inference: An Introduction. Chapman & Hall, New York.

Hui, G.E., 1980. Prediction Distribution for Some Life Testing Models. Ph.D. Dissertation, The University of Western Ontario.

Ipsen Jr., J., 1949. A Practical Method of estimating the mean and standard deviation of truncated normal distributions. Human Biol. 21, 1–16.

Kambo, N.S., 1978. Maximum likelihood estimators of the location and scale parameters of the exponential distribution from a censored sample. Comm. Statist. Theory Methods A 12, 1129–1132.

Lalitha, S., Mishra, A., 1996. Maximum likelihood estimation for Rayleigh distribution. Comm. Statist. Theory Methods 25 (2), 389–401.

Lawless, J.F., 1982. Statistical Models and Methods for Lifetime Data. Wiley, New York.

Raqab, M.Z., 1995. On the maximum likelihood prediction of the exponential distribution based on doubly type II censored samples. Pakistan J. Statist. 11 (1), 1–10.

Sarhan, A.E., 1955. Estimation of the mean and standard deviation by order statistics, Part III. Ann. Math. Statist. 26, 576–592.