Bayesian prediction for the log-normal model under Type II censoring

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 0405-17, Spring 2005

Center for Applied Mathematics and Statistics

NJIT

Bayesian prediction for the log-normal model under Type II censoring

Hafiz M. R. Khan[1], M. Safiul Haq[2] and Serge B. Provost[2]

1. Department of Mathematical Sciences, New Jersey Institute of Technology, University Heights, Newark, NJ 07102-1982, USA. E-mail: khan@adm.njit.edu

2. Department of Statistical and Actuarial Sciences, The University of Western Ontario, London, Ontario, CANADA N6A 5B7. E-mail: sbp@uwo.ca

Abstract

Given a Type II censored sample and a Type II 'median censored' sample from the two parameter log-normal distribution, we have derived the predictive distribution of future responses assuming a non-informative prior as well as an informative prior distribution for the parameters. We have also obtained various types of estimates of reliability functions. A numerical example illustrates the results.

Key words: Life testing model, Type II censored samples, Type II median censored sample, order statistics, log-normal distribution, predictive distribution.

1 Introduction

Numerous devices such as radios, televisions, personal computers, and various types of physical equipment have become an integral part of our daily lives and it is of interest to estimate their average life times. One may also wish to determine the effectiveness of a particular drug or serum by analyzing survival data in connection with some experiments. The failure times of a device or the survival times following a given treatment are recorded as life data. These data occur in an ordered manner; for example, in life tests the weakest 'unit' fails first, the second weakest fails next, and so on. These failure time data are modelled as life testing models. A widely used life testing model is the log-normal model. Under this model, the probability density function (pdf) of the life time Y is given by

$$f(y|\mu,\sigma) = \begin{cases} \frac{1}{(2\pi)^{1/2}\sigma\, y} \exp\left\{-\frac{(\log y-\mu)^2}{2\sigma^2}\right\},\ y \geq 0\ ; -\infty < \mu < \infty,\ \sigma \geq 0, \\ \\ 0 \qquad \text{elsewhere,} \end{cases} \tag{1}$$

where μ is the location parameter and σ is the scale parameter. The corresponding distribution function F(t) is given by

$$F(t|\mu,\sigma) = \int_{y=0}^{t} \frac{1}{(2\pi)^{1/2}\sigma\, y} \exp\left\{-\frac{(\log y-\mu)^2}{2\sigma^2}\right\} dy$$
$$= \Phi\left(\frac{\log t-\mu}{\sigma}\right).$$

The rth moments about origin of the log-normal distribution is given by $\mu_r' = \exp\left(r\mu + \frac{\sigma^2 r^2}{2}\right)$. The mean and variance of this pdf are respectively $\mu_1' = \exp\left(\mu + \frac{\sigma^2}{2}\right)$ and $\mu_2 = \exp\left(2\mu + \sigma^2\right)\left\{\exp\left(\sigma^2\right) - 1\right\}$. This distribution is positively skewed as its skewness is $\left(\exp\left\{\sigma^2\right\} + 2\right)\left(\exp\left\{\sigma^2\right\} - 1\right)^{1/2}$. For additional properties of this distribution, the reader is referred to Crow and Shimizu (1988) and Aitchison and Brown (1957), among others.

The log-normal model is appropriate when initially the failure rate is rather high and it subsequently decreases with time. Howard and Dodson (1961) and Adams (1962) showed that it is an applicable model for the life times of certain semiconductor devices and transistors. In life testing experiments, the failure times of all devices may be recorded, but in order to record the failures of all of the devices that are being tested, the experimenter would have to wait an inordinate amount of time. So in order to save both time and money, the experimenter may choose to stop the experiment after observing a certain number of failure times. In such an instance, we have what is known as a Type II censored sample. Type II censoring is extensively used in the fields of biological, industrial and medical sciences.

There has been numerous of studies on inference about the parameters of the log-normal distribution. For example, Zellner (1971) has discussed in detail Bayesian and non-Bayesian methods for estimating the parameters of the log-normal distribution and the log-normal regression processes based on a complete sample. He considered a diffuse prior for μ and σ to obtain minimal mean squared error estimators of the mean and the median of that distribution. Tiku (1968) has estimated the parameters of the log-normal distribution from censored data by considering the local approximation $g(y) \simeq \alpha + \beta y$, where $g(y)$ is the ratio of the ordinate and the probability integral of a standardized normal distribution. Rukhin (1986) has proposed an improved estimation of Finney's (1941) minimum variance unbiased estimator in terms of the mean square error in log-normal models. Sinha (1988) has derived the predictive density for a single future response from a log-normal distribution in a Bayesian framework making use of a non-informative prior.

In this paper, we discuss predictive inference for future responses given both Type II censored samples and Type II median censored samples from a log-normal life model. We also developed various types of estimates of the reliability function. Our derivations are based on the Bayesian framework. We utilized non-informative as well as informative prior distributions for the parameters. Our informative prior is general and convenient from an analytical point of view, and it is also easy to interpret. We illustrate our results by means of a numerical example.

2 Predictive distributions: non-informative prior

Inference about future responses from a given set of observed responses is known as predictive inference. In predictive inference, the observed data can be viewed as an informative experiment and the unobserved future data form the future experiment. The aim of predictive inference is to obtain probability statements about the future experiment given the informative experiment. Assume that n units of a device are being tested on a life test and that the experiment is terminated after r units have failed. Let y_1 be the first observed life time, y_2 be the next observed life time, and so on. In total, r ($\leq n$) failure times are being observed. Then $\mathbf{y} = y_1, \ldots, y_r$ is the observed sample of life times. We assume that the two parameter log-normal model (1) is appropriate for this experiment. Following Kim, Lee and Kang (2000), for a given set of data $\mathbf{y} = y_1, \ldots, y_r$, the likelihood function is given by

$$\begin{aligned} \mathrm{L}(\mu, \sigma|\mathbf{y}) &= \frac{n!}{(n-r)!}\left[\prod_{i=1}^{r} f(y_i)\right]\left[1 - F(y_r)\right]^{n-r} \\ &= \frac{n!}{(n-r)!}(2\pi)^{-r/2}\sigma^{-r}\left(\prod_{i=1}^{r} y_i^{-1}\right)\exp\left\{-\frac{\sum_{i=1}^{r}(\log y_i - \mu)^2}{2\sigma^2}\right\} \\ &\quad \times\left(1 - \Phi\left(\frac{\log y_r - \mu}{\sigma}\right)\right)^{n-r}. \end{aligned}$$

According to Fernandez (2000), the likelihood function tend to be fairly flat for incomplete data. For this reason, the ML estimator may be of limited value for inference about the future responses. It is therefore useful to assume a prior distribution for the parameters. In some situations, the experimenter may not have any information about the unknown parameters. In such instances, a non-informative prior or vague prior may be used in a Bayesian analysis. Let $p(\mu, \sigma)$ be the Jeffreys (1961) non-informative prior density of μ and σ; then the posterior density of μ and σ, given a set of data $\mathbf{y}$ is

$$\begin{aligned} p(\mu, \sigma|\mathbf{y}) &\propto p(\mu, \sigma)\mathrm{L}(\mu, \sigma|\mathbf{y}) \\ &= \Psi_1(\mathbf{y})\sigma^{-(r+1)}\exp\left\{-\frac{\sum_{i=1}^{r}(\log y_i - \mu)^2}{2\sigma^2}\right\} \\ &\quad \times\left(1 - \Phi\left(\frac{\log y_r - \mu}{\sigma}\right)\right)^{n-r}, \end{aligned}$$

where

$$\mathit{\Psi}_1^{-1}(\mathbf{y}) = \int_{\mu=-\infty}^{+\infty} \int_{\sigma=0}^{+\infty} \sigma^{-(r+1)} \exp\left\{-\frac{\sum_{i=1}^{r}(\log y_i - \mu)^2}{2\sigma^2}\right\} \times \left(1 - \Phi\left(\frac{\log y_r - \mu}{\sigma}\right)\right)^{n-r} d\sigma d\mu.$$

We will use this posterior density to derive the predictive distribution of future responses.

Consider a system composed of m identical units of a device, where each unit has the independent log-normal life distribution as described in (1). Let $z_1, \ldots, z_m$ be the m independent future responses. Then the predictive density function for m future responses is given by

$$\begin{aligned}
& p(\mathbf{z}|\mu, \sigma) \\
= & \int_{\mu=-\infty}^{+\infty} \int_{\sigma=0}^{\infty} \Big(\prod_{i=1}^{m} p(z_i|\mu, \sigma)\Big) \, p(\mu, \sigma|\mathbf{y}) d\sigma d\mu \\
= & \mathit{\Psi}_1(\mathbf{y}) \int_{\mu=-\infty}^{+\infty} \int_{\sigma=0}^{+\infty} \frac{1}{(2\pi)^{m/2}} \prod_{i=1}^{m} \left(\frac{1}{z_i}\right) \exp\left\{-\frac{\sum_{i=1}^{m}(\log z_i - \mu)^2}{2\sigma^2}\right\} \\
& \times \sigma^{-(m+r+1)} \exp\left\{-\frac{\sum_{i=1}^{r}(\log y_i - \mu)^2}{2\sigma^2}\right\} \left(1 - \Phi\left(\frac{\log y_r - \mu}{\sigma}\right)\right)^{n-r} d\sigma d\mu. \\
= & \frac{\mathit{\Psi}_1(\mathbf{y})}{(2\pi)^{m/2}} \prod_{i=1}^{m} \left(\frac{1}{z_i}\right) \int_{\mu=-\infty}^{+\infty} \int_{\sigma=0}^{+\infty} \sigma^{-(m+r+1)} \\
& \times \exp\left\{-\frac{\sum_{i=1}^{r}(\log y_i - \mu)^2 + \sum_{i=1}^{m}(\log z_i - \mu)^2}{2\sigma^2}\right\} \\
& \times \left(1 - \Phi\left(\frac{\log y_r - \mu}{\sigma}\right)\right)^{n-r} d\sigma d\mu, \quad z_i > 0, \; i = 1, \ldots, m.
\end{aligned}$$

There is no closed form representation of the above density function. Clearly, for an informative prior, the distributions becomes even more complicated. However, we may develop an approximate solution if the experiment is stopped after half of the experimental units have failed. We will refer to the sample resulting from such a censoring as a *Type II median censored sample.* Assuming that the sample median approximates the population median, for all $\mu \in (-\infty, \; +\infty)$ and $\sigma > 0$, y_r denoting the sample median, the predictive density becomes

$$p(\mathbf{z}|\mu, \sigma)$$

$$= \frac{\Psi_1(\mathbf{y})}{2^{(m+n-r)/2}\pi^{m/2}} \prod_{i=1}^{m} \left(\frac{1}{z_i}\right) \int_{\mu=-\infty}^{+\infty} \int_{\sigma=0}^{+\infty} \sigma^{-(m+r+1)}$$
$$\times \exp\left\{-\frac{\sum_{i=1}^{r}(\log y_i - \mu)^2 + \sum_{i=1}^{m}(\log z_i - \mu)^2}{2\sigma^2}\right\} d\sigma d\mu,$$
$$z_i > 0,\ i = 1, \ldots, m.$$

3 Inference about the parameters and estimates of the reliability and hazard functions under an informative prior

For predictive inference with an informative prior for μ and σ, we consider the following prior:

$$p(\mu, \sigma) \propto \sigma^{-2g-1} \exp\left\{-\frac{h}{2\sigma^2}\right\},\ -\infty < \mu < +\infty,\ \ \sigma,\ g,\ h \geq 0. \tag{2}$$

In this case, we assume that μ is uniform over the interval $(-\infty, +\infty)$ and the scale parameter σ has a square-root inverted gamma density (see Bernardo and Smith, 1994, p.119). Furthermore, we assume that μ and σ are independent. Here g and h are the hyperparameters, which reflect the experimenter's prior notion about the precision of σ. The density function given by (2) is suitable for deriving the posterior density. This prior reduces to Jeffreys (1961) non-informative prior for $g = 0$ and $h = 0$, and is asymptotically locally invariant for $h = 0,\ g = 1$; see Hartigan (1964). For the informative prior,

$E(\sigma) = (\frac{h}{2})^{1/2} \frac{\Gamma(g-1/2)}{\Gamma(g)}$,
and $Var(\sigma) = \frac{h}{2(g-1)} - \frac{h}{2}\left(\frac{\Gamma(g-1/2)}{\Gamma(g)}\right)^2;\ \sigma,\ g > 1.$

Fernandez (2000) considered this prior in connection with the prediction of future responses from a Rayleigh distribution using a Type II doubly censored sample. We use this prior to derive the predictive distribution of future responses from a log-normal distribution using a Type II median censored sample. Using the prior (2), the posterior density of μ and σ given a set of median censored data $\mathbf{y} = y_1, \ldots, y_r$, is

$$\begin{aligned} p(\mu, \sigma|\mathbf{y}) &\propto p(\mu, \sigma)\mathrm{L}(\mu, \sigma|\mathbf{y}) \\ &= \Psi_2(\mathbf{y})\sigma^{-(2g+r+1)} \exp\left\{-\frac{h + \sum_{i=1}^{r}(\log y_i - \mu)^2}{2\sigma^2}\right\}, \\ &\qquad -\infty < \mu < +\infty,\ \sigma > 0, \end{aligned}$$

where

$$\begin{aligned}
\Psi_2^{-1}(\mathbf{y}) &= \int_{\mu=-\infty}^{+\infty}\int_{\sigma=0}^{+\infty} \sigma^{-(2g+r+1)} \exp\left\{-\frac{h+\sum_{i=1}^{r}(\log y_i-\mu)^2}{2\sigma^2}\right\} d\sigma d\mu \\
&= \frac{2^{(2g+r-2)/2}\pi^{1/2}\Gamma(\frac{2g+r-1}{2})}{r^{1/2}\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)^{(2g+r-1)/2}}.
\end{aligned}$$

The marginal density of μ given a set of data $\mathbf{y}$ is then

$$\begin{aligned}
p(\mu|\mathbf{y}) &= r^{1/2}\frac{\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)^{(2g+r-1)/2}}{2^{(2g+r-2)/2}\pi^{1/2}\Gamma(\frac{2g+r-1}{2})} \\
&\quad\times\int_{\sigma=0}^{+\infty} \sigma^{-(2g+r+1)} \exp\left\{-\frac{h+\sum_{i=1}^{r}(\log y_i-\mu)^2}{2\sigma^2}\right\} d\sigma. \\
&= \frac{\sqrt{r}}{\sqrt{\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)}\beta(\frac{1}{2},\frac{2g+r-1}{2})} \\
&\quad\times\left[1+\frac{r\left(\mu-\frac{\sum_{i=1}^{r}\log y_i}{r}\right)^2}{\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)}\right]^{-\frac{2g+r}{2}}, \\
&\qquad\qquad -\infty<\mu<+\infty,
\end{aligned}$$

where $\beta(a_1,\ b_1)\ =\ \frac{\Gamma(a_1)\Gamma(b_1)}{\Gamma(a_1+b_1)}$. We note that $\sqrt{\frac{r(2g+r-1)}{\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)}}\times\left(\mu-\frac{\sum_{i=1}^{r}\log y_i}{r}\right)$ has a Student's -t distribution with $(2g+r-1)$ degrees of freedom. Thus a $(1-\alpha)100\%$ confidence interval for μ is

$$\left[\frac{\sum_{i=1}^{r}\log y_i}{r}\pm t_{(\frac{\alpha}{2},\ 2g+r-1)}\sqrt{\frac{\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)}{r(2g+r-1)}}\right],$$

where α is the significance level.

Similarly, the marginal density of σ given the data $\mathbf{y}$ is given by

$$p(\sigma|\mathbf{y}) = r^{1/2}\frac{\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)^{(2g+r-1)/2}}{2^{(2g+r-2)/2}\pi^{1/2}\Gamma(\frac{2g+r-1}{2})}$$

$$\times \int_{\mu=-\infty}^{+\infty} \sigma^{-(2g+r+1)} \exp\left\{-\frac{h+\sum_{i=1}^{r}(\log y_i-\mu)^2}{2\sigma^2}\right\} d\mu.$$

$$= \frac{\left(h+\sum_{i=1}^{r}(\log y_i)^2 - \frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)^{\frac{2g+r-1}{2}}}{2^{\frac{2g+r-3}{2}}\Gamma(\frac{2g+r-1}{2})}\sigma^{-(2g+r)}$$

$$\times \exp\left\{-\frac{\left(h+\sum_{i=1}^{r}(\log y_i)^2 - \frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)}{2\sigma^2}\right\}.$$

Noting that $\frac{(h+\sum_{i=1}^{r}(\log y_i)^2 - \frac{(\sum_{i=1}^{r}\log y_i)^2}{r})}{\sigma^2}$ is distributed as a χ^2 with $(2g+r-1)$ degrees of freedom, a $(1-\alpha)100\%$ confidence interval for σ^2 is
$\left[\frac{(h+\sum_{i=1}^{r}(\log y_i)^2 - \frac{(\sum_{i=1}^{r}\log y_i)^2}{r})}{\chi^2_{(\frac{\alpha}{2},\ 2g+r-1)}}, \frac{(h+\sum_{i=1}^{r}(\log y_i)^2 - \frac{(\sum_{i=1}^{r}\log y_i)^2}{r})}{\chi^2_{(1-\frac{\alpha}{2},\ 2g+r-1)}}\right]$.

These marginal distributions agree with those given by Sinha (1988) for a vague prior when the hyperparameters are such that $g = h = 0$.

Reliability Function:

The reliability of a device is defined as $R(t) = \Pr\{$ the device functions at least t units of time$\} = 1-F(t)$. This concept of 'reliability' is frequently used in the medical and engineering fields, e.g., reliable treatment, reliable service station, reliable products, etc. Under the squared-error loss function, Bayes estimators of $R(t|\mathbf{y})$ is obtained as

$$R(t|\mathbf{y})$$

$$= \int_{y=t}^{\infty}\int_{\mu=-\infty}^{\infty}\int_{\sigma=0}^{\infty} p(y|\mu,\sigma)p(\mu,\sigma|\mathbf{y})d\sigma d\mu dy$$

$$= \int_{y=t}^{\infty}\int_{\mu=-\infty}^{\infty} r^{1/2}\frac{\left(h+\sum_{i=1}^{r}(\log y_i)^2 - \frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)^{(2g+r-1)/2}\Gamma(\frac{2g+r+1}{2})}{\pi\Gamma(\frac{2g+r-1}{2})y}$$

$$\times\left[h+\sum_{i=1}^{r}(\log y_i-\mu)^2+(\log y-\mu)^2\right]^{-\frac{(2g+r+1)}{2}} d\mu dy.$$

Thus for a fixed value of t and a given set of data, one can evaluate $R(t|\mathbf{y})$.

Using the maximum likelihood estimates (MLE) of μ and σ, the estimate of the reliability function is obtained as

$$
\begin{aligned}
R(t|\hat{\mu},\hat{\sigma}) &= \Pr\{y \geq t | \mu = \hat{\mu}, \text{and } \sigma = \hat{\sigma}\} \\
&= \int_{y=t}^{\infty} \frac{1}{(2\pi)^{1/2}\hat{\sigma}\ y} \exp\left\{-\frac{(\log y - \hat{\mu})^2}{2\hat{\sigma}^2}\right\} dy.
\end{aligned}
$$

Following Sinha (1986), the uniformly minimum variance unbiased estimator (UMVUE) of $R(t)$ is $R(t) = 1 - \Phi\left(\frac{logt-\hat{\mu}}{\hat{\sigma}\sqrt{\frac{n-1}{n}}}\right)$.

Hazard Function:

The hazard function has been extensively used in the reliability modelling of a device. It is defined as the ratio of the probability of failure at time t to the probability that failure has not yet occurred by time t, e.g.,

$$
\begin{aligned}
H(t|\hat{\mu},\hat{\sigma}) &= f(t|\hat{\mu},\hat{\sigma})/(1 - F(t|\hat{\mu},\hat{\sigma})) \\
&= \frac{\frac{1}{(2\pi)^{1/2}\hat{\sigma}t} \exp\left\{-\frac{(\log t - \hat{\mu})^2}{2\hat{\sigma}^2}\right\}}{1 - \int_{y=0}^{t} \frac{1}{(2\pi)^{1/2}\hat{\sigma}\ y} \exp\left\{-\frac{(\log y - \hat{\mu})^2}{2\hat{\sigma}^2}\right\} dy}.
\end{aligned}
$$

This function can be evaluated numerically.

4 Predictive inference for a single future response under an informative prior

Using the posterior density given in Section 3, the predictive density function of a single future response z, for a given set of data $\mathbf{y}$ can be expressed as

$$
\begin{aligned}
& p(z|\mathbf{y}) \\
= \; & \int\int p(z|\mu,\sigma)p(\mu,\sigma|\mathbf{y})d\sigma d\mu \\
= \; & \frac{\Gamma(\frac{2g+r}{2})\sqrt{\frac{r}{(r+1)}}}{\Gamma(\frac{2g+r-1}{2})\pi^{\frac{1}{2}}\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)^{1/2} z} \\
& \times\left[1+\frac{r\left(\log z-\frac{\sum_{i=1}^{r}\log y_i}{r}\right)^2}{(r+1)\left(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r}\right)}\right]^{-(2g+r)/2}, \\
& 0<z<\infty.
\end{aligned}
$$

Letting $\bar{y}=\frac{1}{r}\sum_{i=1}^{r}\log y_i$ and $S_1^2=\frac{1}{2g+r-1}\frac{r+1}{r}(h+\sum_{i=1}^{r}(\log y_i)^2-\frac{(\sum_{i=1}^{r}\log y_i)^2}{r})$, one has

$$
\begin{aligned}
& p(z|\mathbf{y}) \\
= \; & \frac{\Gamma(\frac{2g+r}{2})}{\Gamma(\frac{2g+r-1}{2})\sqrt{\pi(2g+r-1)}\;S_1} \\
& \times\left[1+\frac{1}{2g+r-1}\frac{\left(\log z-\frac{\sum_{i=1}^{r}\log y_i}{r}\right)^2}{S_1^2}\right]^{-(2g+r)/2}, \; 0<z<\infty.
\end{aligned}
$$

It is observed that $z|\mathbf{y}$ has a log-t distribution with location parameter $\bar{y}$, scale parameter S_1^2, and $(2g+r-1)$ degrees of freedom. Therefore, the mean of $\log z$ is $\bar{y}$ and the variance of $\log z$ is $\frac{(2g+r-1)}{(2g+r-3)}S_1^2$. Thus a $(1-\alpha)100\%$ prediction interval for z will be $\exp\left\{\bar{y}\pm t_{(2g+r-1,\frac{\alpha}{2})}\sqrt{\frac{(2g+r-1)}{(2g+r-3)}S_1^2}\right\}$, where $t_{(2g+r-1,\frac{\alpha}{2})}$ is the upper $(100\;\alpha/2)th$ percentile of the t-distribution with $(2g+r-1)$ degrees of freedom. This predictive density is similar to Sinha's (1988) for a non-informative prior with $g=0,\; h=0$ and n = r. With the same conditions, the result for $\log z$ is similar to that of Thabane (1998) who studied the prediction of a single future response based on a Box-Cox transformed observed response, the transformed response following a univariate normal distribution. The above result also agrees with that of Johnson, Kotz and Balakrishnan

(1995, p.400) for a non-informative prior when $(r-1) = 2q$.

4.1 Predictive distribution for two future responses

From an application point of view, one may be interested in making inference about more than one future response. The predictive density of two future responses z_1 and z_2, for the given set of data $\mathbf{y}$ is

$$\begin{aligned}
& p(z_1, z_2|\mathbf{y}) \\
= & \int_{\mu=-\infty}^{\infty}\int_{\sigma=0}^{\infty}\prod_{j=1}^{2} p(z_j|\mu,\sigma)p(\mu,\sigma|\mathbf{y})d\sigma d\mu \\
= & \frac{\Gamma(\frac{2g+r+1}{2})|\mathbf{D}^{-1}|^{1/2}}{(\pi(2g+r-1))^{2/2}\Gamma(\frac{2g+r-1}{2})S_2^2}\left(\prod_{j=1}^{2}\frac{1}{z_j}\right) \\
& \times\left[1+\frac{1}{(2g+r-1)S_2^2}\left\{(\log\mathbf{z}-\bar{y}\mathbf{1}_2)^{'}\mathbf{D}^{-1}(\log\mathbf{z}-\bar{y}\mathbf{1}_2)\right\}\right]^{-\frac{2g+r+1}{2}}, \\
& \qquad z_1, z_2 > 0,
\end{aligned}$$

where $\mathbf{D}^{-1} = I_2 - \frac{1}{2+r}\mathbf{1}_2\mathbf{1}_2^{'}$, $\mathbf{1}_2^{'} = (1,\ 1)$, and $S_2^2 = \frac{1}{2g+r-1}(h + \sum_{i=1}^{r}(\log y_i)^2 - \frac{(\sum_{i=1}^{r}\log y_i)^2}{r})$, i.e., $\log\mathbf{z}|\mathbf{y} \sim t_2(\bar{y}\mathbf{1}_2, S_2^2\mathbf{D}, 2g+r-1)$, where $t_2()$ denote a bivariate t distribution. It is to be noted that $E(\log\mathbf{z}) = \bar{y}\mathbf{1}_2$ and $Cov(\log\mathbf{z}) = \frac{(2g+r-1)}{(2g+r-3)}S_2^2\mathbf{D}$.

Following a similar procedure, the predictive density function for a set of m future responses $\mathbf{z} = (z_1, z_2, \ldots, z_m)$ is obtained as $p(\mathbf{z}|\mathbf{y})$

$$= \begin{cases} \frac{\Gamma(\frac{2g+r+m-1}{2})|\mathbf{G}^{-1}|^{1/2}}{(\pi(2g+r-1)S_m^2)^{m/2}\Gamma(\frac{2g+r-1}{2})}\left(\prod_{j=1}^{m}\frac{1}{z_j}\right) \\ \times\left[1+\frac{1}{(2g+r-1)S_m^2}\left\{(\log\mathbf{z}-\bar{y}\mathbf{1}_m)^{'}\mathbf{G}^{-1}(\log\mathbf{z}-\bar{y}\mathbf{1}_m)\right\}\right]^{-\frac{2g+r+m-1}{2}}, \\ \qquad z_i > 0, \\ 0, \qquad \text{elsewhere} \end{cases}$$

where $\mathbf{G}^{-1} = I_m - \frac{1}{m+r}\mathbf{1}_m\mathbf{1}_m^{'}$, $\mathbf{1}_m^{'} = (1,\ldots,1)$, and $S_m^2 = \frac{1}{2g+r-1}(h + \sum_{i=1}^{r}(\log y_i)^2 - \frac{(\sum_{i=1}^{r}\log y_i)^2}{r})$.
It is noted that $\mathbf{z}|\mathbf{y}$ is an m-dimensional multivariate log-t random vector with $(2g+r-1)$ degrees of freedom, i.e., $\log\mathbf{z}|\mathbf{y} \sim t_m(\bar{y}\mathbf{1}_m, S_m^2\mathbf{G}, 2g+r-1)$.

Following Box and Tiao (1973, p. 117), the quadratic form

$q = \frac{1}{mS_m^2}(\log\mathbf{z} - \bar{y}\mathbf{1}_m)^{'}\mathbf{G}^{-1}(\log\mathbf{z} - \bar{y}\mathbf{1}_m)$ has an F-distribution with m and $(2g+r-1)$ degrees of freedom. Then a $(1-\alpha)100\%$ predictive ellipsoid for $\log\mathbf{z}$ is given by

$$(\log\mathbf{z} - \bar{y}\mathbf{1}_m)^{'}\mathbf{G}^{-1}(\log\mathbf{z} - \bar{y}\mathbf{1}_m) \leq mS_m^2 F_{\alpha,(m,\ 2g+r-1)}.$$

5 A numerical example

We consider the data analyzed by Tiku (1967), where he determined the maximum likelihood estimates of the parameters μ and σ assuming a normal distribution for the base 10 log days. The data are the days on which the first 7 of a sample of 10 tested mice died after being inoculated with a uniform culture of human tuberculosis in a life testing experiment. Obviously, the data constitutes a Type II censored sample. We have obtained the maximum likelihood estimates $\hat{\mu} = 3.9928$, and $\hat{\sigma} = 0.1046$ for the parameters of an assumed log-normal distribution for the days.

Table 1: Life data of mice after being inoculated

Days after inoculation	log days after inoculation
41	3.714
44	3.784
46	3.829
54	3.989
55	4.007
58	4.060
60	4.094

The predictive density for a single future response involves the hyperparameters g and h. We have attempted to estimate g and h from the data. However, the likelihood equations do not yield a closed form solution, and even a numerical solution of the likelihood equation using the Newton-Raphson algorithm does not converge to a solution. For this reason we will consider some specific values of g and h.

In Table 2, with $g = h$, we provide certain values of $R(t)$, the reliability function based on the MLE, a Bayes estimate, a uniform minimum variance unbiased estimate, and on the predictive density obtained from an informative prior. It is noted that the values of $R(t)$ differ very slightly between UMVUE and MLE. With Bayes estimates, the values of $R(t)$ are higher than those obtained when using UMVUE and MLE. It is interesting to note that the reliability function using the predictive density function is higher than that of other estimates.

Table 2: Comparison of reliability estimates with $g = h$

$Days$	$g = h$	$\hat{R}_{MLE}(t)$	$\hat{R}_{UMVUE}(t)$	$\hat{R}_{Bayes}(t)$	$\hat{R}_{PredInf}(t)$
61	05	0.1295	0.1114	0.1328	0.3816
63	10	0.0753	0.0603	0.1026	0.3734
65	15	0.0413	0.0304	0.0844	0.3607
67	20	0.0214	0.0143	0.0720	0.3473
69	25	0.0105	0.0064	0.0628	0.3340
71	30	0.0049	0.0027	0.0557	0.3210
73	35	0.0022	0.0011	0.0498	0.3085
75	40	0.0010	0.0004	0.0451	0.2963
77	45	0.0004	0.0002	0.0409	0.2847
100	60	0.0000	0.0000	0.0245	0.1786

In Table 3, for increasing values of g and h with $0 < h < 1$, we have determined the values of $R(t)$ from the Bayes estimates, the predictive density with an informative prior and the predictive density with a non-informative prior. It is observed that for $0 < h < 1$, the values of $R(t)$ obtained with Bayes estimates are consistently smaller than those obtained with a predictive density with an informative prior which in turn happen to be much smaller than those corresponding to the non-informative prior case.

Table 4 and Table 5 show the variance, standard deviation, 95% predictive intervals and the relative precision for a single future response with informative prior for specific values of the hyperparameters. In Table 4, we have considered the range, $0 < h < 1$, for h while in Table 5, $h > 1$. We note that for the given set of data, the mean and variance for a single future response with non-informative prior are respectively 51.6449 and 235.0810. With a non-informative prior, the 95% predictive interval for a single future response is (34.3918, 74.6627). With an informative prior, the lower limit is larger than that obtained with the non-informative prior, but the upper limit is smaller and the lengths of the predictive intervals are always smaller. The relative precision of the estimate of a future response, as defined by Cochran (1977), were calculated for $0 < h < 1$ and $h > 1$. It is observed that for $0 < h < 1$, the predictive estimate of z is significantly more efficient than it is for $h > 1$. In Table 4, it is observed that the increase in precision of a future response confirms that an informative prior can provide better predictive estimates.

Table 3: Comparison of reliability estimates with informative & non-informative prior

$Days$	g	h	$\hat{R}_{Bayes}(t)$	$\hat{R}_{PredInf}(t)$	$\hat{R}_{PredNon}(t)$
61	1	0.050	0.0685	0.1414	0.1429
63	5	0.100	0.0154	0.0549	0.1091
65	10	0.150	0.0035	0.0170	0.0835
67	15	0.200	0.0012	0.0050	0.0641
69	20	0.250	0.0003	0.0014	0.0496
71	25	0.350	0.0001	0.0006	0.0386
73	30	0.400	0.0000	0.0002	0.0303
75	35	0.450	0.0000	0.0000	0.0240
77	40	0.500	0.0000	0.0000	0.0192
100	55	0.95	0.0000	0.0000	0.0026

Table 4: Comparison of variances for predictive densities with non-informative prior & informative prior for increasing values of g and $0 < h < 1$

g	h	$Var(\log z)$	$SD(z)$	95% P. I. of z	Relative precision
2	0.010	0.0204	7.3640	(36.9042, 69.5797)	4.3349
4	0.030	0.0154	6.3789	(38.8285, 66.1315)	5.7774
6	0.050	0.0129	5.8389	(39.885, 64.3798)	6.8954
8	0.070	0.0115	5.4935	(40.5567, 63.3136)	7.7896
10	0.090	0.0106	5.2523	(41.0223, 62.5949)	8.5215
12	0.110	0.0099	5.0738	(41.7081, 61.5656)	9.1318
14	0.130	0.0093	4.9360	(41.9252, 61.2469)	9.6487
16	0.150	0.0089	4.8264	(42.0991, 60.9939)	10.0920
18	0.170	0.0086	4.7370	(42.2416, 60.7882)	10.4764
20	0.190	0.0084	4.6627	(42.3605, 60.6175)	10.8129
32	0.310	0.0074	4.3765	(42.7989, 59.9965)	12.2736
50	0.490	0.0068	4.2037	(43.0941, 59.5856)	13.3031
64	0.630	0.0065	4.1316	(43.2166, 59.4167)	13.7717
85	0.840	0.0063	4.0653	(43.3285, 59.2632)	14.2241
100	0.999	0.0062	4.0342	(43.3540, 59.2283)	14.4443

Table 5: Comparison of variances for predictive densities with non-informative prior & informative prior for increasing values of g and $h > 1$

g	h	$Var(\log z)$	$SD(z)$	95% P. I. of z	Relative precision
2	1.50	0.2331	32.3303	(17.2824, 148.5780)	0.2249
4	2.50	0.2506	32.5554	(17.3140, 148.3070)	0.2218
6	3.50	0.2594	32.7545	(17.3798, 147.7460)	0.2191
8	4.50	0.2647	32.9006	(17.4337, 147.289)	0.2172
10	5.50	0.2682	33.0074	(17.4734, 146.954)	0.2158
12	6.50	0.2707	33.0876	(18.2774, 140.490)	0.2147
14	7.50	0.2726	33.1495	(18.2129, 140.987)	0.2139
16	8.50	0.2740	33.1986	(18.1631, 141.3740)	0.2133
18	9.50	0.2752	33.2385	(18.1234, 141.6830)	0.2128
20	10.50	0.2762	33.2715	(18.0911, 141.936)	0.2124
32	16.50	0.2795	33.3901	(17.9779, 142.830)	0.2109
50	25.50	0.2817	33.4668	(17.9068, 143.397)	0.2099
64	32.50	0.2825	33.4979	(17.8785, 143.625)	0.2095
85	43.00	0.2833	33.5260	(17.8531, 143.828)	0.2091
100	50.50	0.2837	33.5389	(17.8415, 143.923)	0.2090

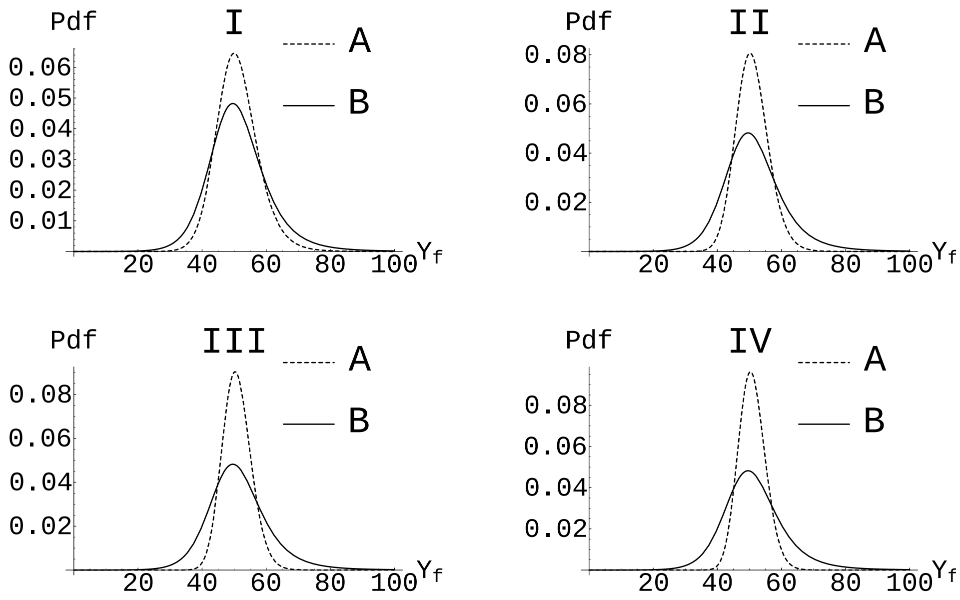

Figure 1: Predictive densities for a single future response $Z = Y_f$, under informative prior:A, non-informative prior:B, for some values of the hyperparameters I:[g=3, h=0.02], II:[g=11, h=0.10], III:[g=25, h=0.24] & IV:[g=50, h=0.49].

For different values of g and h with $0 < h < 1$, the predictive densities $p(z|\mathbf{y})$ have been determined. The graphs of these densities are plotted in Figure 1. The predictive densities appear to be in conformity with the given data. These densities appear to be slightly skewed with a mode approximately at 50. We also note that with an informative prior, the variance of z decreases. Figure 2 shows the predictive density function of two future responses.

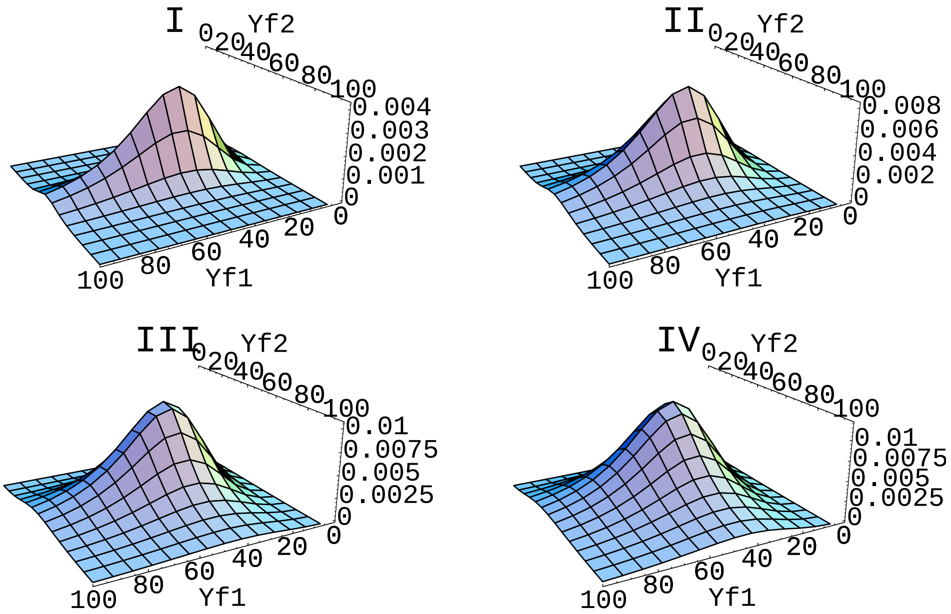

Figure 2: Predictive densities for the future responses $Z_1 = Y_{f1}$ and $Z_2 = Y_{f2}$, for some values of the hyperparameters I:[g=2 & h=0.01], II:[g=18 & h=0.17], III:[g=51 & h=0.50] & IV:[g=100 & h=0.99]. The vertical axes represent the density function.

6 Conclusion

We have used the Bayesian framework to derive the predictive distributions of future responses given both Type II censored samples and Type II median censored samples, assuming non-informative and informative prior distributions for the parameters. It is noted that based on a median censored sample and certain values of the hyperparameters, the predictive densities turn out to be more efficient than those obtained with a non-informative prior.

References

Adams JD (1962) Failure time distribution estimation. Semi-conductor Rel. **2**, 41-52

Aitchison J, Brown JAC (1957) The Log-normal Distribution. Cambridge University Press, Cambridge

Bernardo JM, Smith AFM (1994) Bayesian Theory. Wiley, New York

Box JEP, Tiao GC (1973) Bayesian Inference in Statistical Analysis. Adison Wesley, London

Cochran WG (1977) Sampling Techniques, 109-113. John Wiley & Sons, New York

Crow EL, Shimizu K (1988) Log-normal Distributions: Theory and Applications. Marcel Dekker, Inc., New York

Fernandez AJ (2000) Bayesian inference from Type II doubly censored Rayleigh data. Statistics & Probability Letters **48**, 393-399

Finney DJ (1941) On the distribution of a variate whose logarithm is normally distributed. Supplement to the Journal of the Royal Statistical Society **7**, 155-161

Hartigan JA (1964) Invariant prior distributions. Ann. Math. Statist. **35**, 836-845

Howard BT, Dodson GA (1961) High stress aging to failure of semiconductor devices. Proc. 7th Nat. Symp. Rel. Quality Control. Ins. Electrical & Electronic Engineers, New York

Jeffreys H (1961) Theory of Probability. 3rd ed. Oxford University Press, Oxford

Johnson LN, Kotz S, and Balakrishnan N (1995) Continuous Univariate Distributions 2, Second Edition, John Willey & Sons, Inc., New York

Kim DH, Lee WD and Kang SG (2000) Bayesian model selection for life time data under Type II censoring. Commun. Statis. - Theory Meth. **29**(12), 2865-2878

Rukhin AL (1986) Improved estimation in log-normal models. Journal of the American Statistical Association **81**, Theory and Methods, 1046-1049.

Sinha SK (1986) Reliability and Life Testing. John Wiley & Sons Inc., New York

Sinha SK (1988) Bayesian inference about the prediction/credible intervals and reliability function for log-normal distribution. Journal of the Indian Statistical Association **27**, 73-78

Thabane L (1998) Contributions to Bayesian Statistical Inference. PhD Thesis. Department of Statistical and Actuarial Sciences. The University of Western Ontario

Tiku ML (1967) Estimating the mean and standard deviation from a censored normal sample. Biometrika **54**, 155-165

Tiku ML (1968) Estimating the parameters of lognormal distribution from censored samples. American Statistical Association Journal **63**, 134-140

Zellner A (1971) Bayesian and non-Bayesian analysis of the log-normal regression. Journal of the American Statistical Association **66**(334), Theory and Method Section, 326-330