Journal of Economic and Social Research 2 (2) 2000, 1-17

Estimation and Hypothesis Testing for Exponential Lifetime Models with Double Censoring and Prior Information

Arturo J. Fernández[1]

Abstract. In this paper, on the basis of a doubly censored sample and in a Bayesian framework, the problem of estimating the mean lifetime, hazard rate, and survival function of the exponential lifetime model is addressed. Bayes estimators under squared-error loss functions are obtained in closed forms. Highest posterior density (*HPD*) estimators and credible intervals are computed using iterative methods. A Bayesian approach to hypothesis testing is also presented. Optimal answers to hypothesis testing problems are obtained in terms of Bayes factors. Both one- and two-sided tests are considered. Finally, an illustrative numerical example is included.

JEL Classification Codes: C120, C130.

Key Words: Bayesian Inference; The *HPD* Estimator and Interval; Double Censoring; Exponential Distribution; Survival Function.

1. Introduction

Statistical models and methods for survival data and other time-to-event data are extensively used in many fields, including the biomedical sciences, engineering, the environmental sciences, economics, actuarial sciences, management, and the social sciences. Examples of lifetime data include the times-to-failure of machine components in industrial reliability, the duration of strikes or periods of unemployment in economics, the times taken by subjects to complete specified tasks in psychological experimentation, the lengths of marriages in sociology, and the survival times of patients in a clinical trial. There are situations, however, where lifetime data are not measured in calendar time. For instance, the lifetime of an automobile is usually measured by its mileage, and the time-to-failure of a copier is

[1] Departamento de Estadística, I. O. y Computación, Universidad de La Laguna, SPAIN

measured by the number of copies made. Similarly, an insurance company uses the total amount of money paid to a client with a prolonged illness for measuring the time to death of such a client, and the reliability of a system designed for managing very large databases is frequently measured by the total amount of data stored in the database. Anyway, they all measure the *duration of life*.

The essential element in lifetime data analysis is the presence of a non-negative response, X, with appreciable dispersion and often with censoring. Due to sampling methods or the occurrence of some competing risk of removal from the study, several lifetimes of individuals may be censored. By censored data we mean that, in a potential sample of size n, a known number of observations is missing at either end (single censoring) or at both ends (double censoring). The type of censoring just described is often called Type II censoring. Doubly censored data has been considered, among other authors, by Harter and Moore (1968), Balakrishnan (1990), Raqab (1995), Balakrishnan and Chan (1995), Lalitha and Mishra (1996), and Kong and Fei (1996).

The Bayesian method of reasoning is currently riding a high tide of popularity in virtually all areas of statistical application. A distinctive feature of Bayesian inferences is that it takes explicit account of prior information in the analysis. Classical statistical inference, based on sampling theory, usually does not consider information beyond the sample data. The Bayesian use of relevant past experience, which is quantified by a prior distribution, produces more informative inferences in those cases where the prior distribution accurately reflects the variation in the unknown parameter. In addition, the Bayesian method usually requires less sample data to achieve the same quality of inferences than the method based on sampling theory. In many cases this is the practical motivation for using a Bayesian approach and represents the practical advantage in the use of prior information.

The exponential distribution $Exp(m)$ provides a population model which is useful in several areas of statistics. In lifetime data analysis, this distribution occupies an important role and inference procedures based on this model are widely used. References to the exponential model can be found, among others, in Bain (1978) and Lawless (1982). The probability density function *(pdf)* of a random variable $X \sim Exp(m)$, which represents the lifetime variable of interest, is given by

$$f(t|m) = m^{-1}\exp(-t/m), \quad t>0. \qquad (1)$$

The positive parameter m is the mean lifetime, while $l = m^{-1}$ is the hazard rate (instantaneous failure rate or force of mortality). The survival or reliability function is given by

$$R(t|m) = \Pr(X>t|m) = \exp(-t/m), \quad t>0. \qquad (2)$$

In this paper the important exponential lifetime model is considered and studied from a Bayesian perspective on the basis of a doubly censored sample. In Section 2, the likelihood function, prior distribution on the mean lifetime, and its corresponding posterior distribution are presented. Point and interval estimation procedures are derived in Section 3. Section 4 is concerned with hypothesis testing. Finally, an illustrative example is given in Section 5.

2. Modelling

Consider a random sample of size n from an $Exp(m)$ distribution, where $m \in$ Ù $=(0,\infty)$ is an unknown parameter, and let $x_r,\ldots,x_s$ be the ordered observations remaining when the $(r-1)$ smallest observations and the $(n-s)$ largest observations have been censored. The likelihood function for m, given the doubly censored sample $x=(x_r,\ldots,x_s)$, is then

$$L(m|x) = \frac{n!}{(r-1)!(n-s)!}\{1-R(x_r|m)\}^{r-1}\{R(x_s|m)\}^{n-s}\prod_{i=r}^{s} f(x_i|m).$$

According to (1) and (2), the likelihood becomes proportional to

$$L(m|x) \propto m^{-m}\exp\{-\mathbf{x}(x)/m\}\{1-\exp(-x_r/m)\}^{r-1}, \qquad (3)$$

where $m = s-r+1$ and $\mathbf{x}(x) = \sum_{i=r}^{s} x_i + (n-s)x_s$. Consequently, the sample evidence is contained in the sufficient statistic $U(x) = (x_r, \mathbf{x}(x))$.

Likelihood functions tend to be fairly flat for censored data. Moreover, asymptotic methods based on the maximum likelihood (ML) estimation are generally inadequate unless censoring levels are low and sample sizes are large. For these reasons the ML estimator may be of limited value. It is therefore especially important in our situation to assess a prior distribution for m. Frequently, prior information concerning an unknown lifetime parameter exists. Often such knowledge can be translated into a prior distribution. In this paper we consider prior densities of the form

$$g(m) \propto \exp(-a/m)\, m^{-(b+1)}, \quad m>0, \qquad (4)$$

where, to be a proper (inverted gamma) density, we must have $a>0$ and $b>0$. This prior distribution has advantages over many other distributions because of its analytical tractability, richness, and easy interpretability. The information summarized by this prior distribution may be either objective (i.e., based on test data from a comparable experiment) or subjective (i.e., based on an individual's experience, judgements, beliefs, and preferences) or both.

The hyperparameters a and b can be assessed to match the experimenter's notion of the location and precision of his/her prior distribution for the mean lifetime, m, since $E[m]=a/(b-1)$ and $Var[m]=a^2/\{(b-1)^2(b-2)\}$, provided that $b>2$. Similarly, the moment-matching method may be used to fit the prior distribution when the available prior knowledge is expressed in terms of the hazard rate, l, or the survival function at some fixed point $t>0$, $R_t=R(t|ì)$ since $E[l]=b/a$ and $Var[l]=b/a^2$, and, $E[R_t]=(1+t/a)^{-b}$ and $Var[R_t]=(1+2t/a)^{-b}-(1+t/a)^{-2b}$. Another suitable approach would be to subjectively estimate two percentiles of the prior distribution. For example, the experimenter's prior 10th and 90th percentiles for m yield two equations that can be solved for a and b.

If prior information about m is scanty, it may be appropriate to resort to the use of a diffuse prior distribution. Improper prior density for m, which can reasonably be accepted, are the Jeffreys' (1961) prior density, $(a=0, b=0)$, and an asymptotically locally invariant prior density proposed by Hartigan (1964), $(a=0,\ b=1)$. The results of this paper will hold for these non-informative prior densities.

Bayesian statistics have the advantage of being able to combine the subjective knowledge of the prior distribution with the knowledge contained in the data. Estimates can be obtained with relatively little data, which become extremely important in the case of expensive testing procedures. By combining the likelihood function (3) with the prior density (4), the posterior density of m is obtained to be

$$g(m \mid x) = \frac{\{a + \mathbf{x}(x)\}^{b+m} \exp\left[-\{a + \mathbf{x}(x)\}/m\right]\{1 - \exp(-x_r/m)\}^{r-1}}{m^{b+m+1}\Gamma(b+m)F_r\left[a+\mathbf{x}(x),\, b+m\right]}, \quad m > 0, \quad (5)$$

where

$$F_r\left[u,v\right] = \sum_{j=0}^{r-1}(-1)^j \binom{r-1}{j}\left(1 + j\frac{x_r}{u}\right)^{-v}, \quad u,v>0 \quad (F_1 \equiv 1).$$

All the information concerning the mean lifetime is now contained in the posterior distribution, which represents a modification of the subjective knowledge about m expressed by the prior distribution in light of the observed sample data. The Bayes theorem provides a mechanism for continually updating our knowledge about m as more sample data become available.

For large sample sizes the posterior distribution will be approximately numerically equal to the standardized likelihood function, and the difference between Bayesian inferences and inferences based on the likelihood function will be insignificant. Thus, at least in large samples, the choice of the prior distribution is not very crucial.

From (5), it follows that the posterior survival function of m given x is

$$S(c \mid x) = \int_c^\infty g(m \mid x)\,dm = \frac{G_r\left[a+\mathbf{x}(x), b+m; c\right]}{F_r\left[a+\mathbf{x}(x), b+m\right]}, \qquad c>0, \qquad (6)$$

where

$$G_r\left[u,v;c\right] = \sum_{j=0}^{r-1}(-1)^j \binom{r-1}{j} H\left(v; \frac{u+jx_r}{c}\right)\left(1 + j\frac{x_r}{u}\right)^{-v}, \qquad u,v>0,$$

in which $H(\bullet,\bullet)$ is the incomplete gamma function defined by

$$H(v;g) = \frac{1}{\Gamma(v)}\int_0^g z^{v-1}\exp(-z)\,dz, \qquad v,g>0.$$

Using the transformations $m = l^{-1}$ and $m = t(-\log R_t)^{-1}$, the respective posterior densities of l and R_t, implied by the prior density (4) on m, can be readily derived. Specifically,

$$h_1(l|x) = \frac{1}{l^2} g(l^{-1}|x), \quad l > 0,$$

and

$$h_2(R_t|x) = \frac{t}{R_t(\log R_t)^2} g(t(-\log R_t)^{-1}|x), \quad 0 < R_t < 1.$$

3. Point and Interval Estimation

From a decision-theoretic viewpoint, in order to select a single value as representing our "best" estimator of m, one must first specify a loss function, $\mathcal{L}(m, \tilde{m})$, which represents the cost involved in using the estimate $\tilde{m}$ when the true value is m. A commonly used loss function for estimating m is the squared-error loss, $\mathcal{L}(m, \tilde{m}) = (m - \tilde{m})^2$. Under this loss function, the Bayes estimator of m is the mean of the posterior density (5) given by

$$\tilde{m} = E[m|x] = \frac{F_r[a+\boldsymbol{x}(x), b+m-1]}{F_r[a+\boldsymbol{x}(x), b+m]} \left\{ \frac{a+\boldsymbol{x}(x)}{b+m-1} \right\}, \quad (b+m>1). \tag{7}$$

Other problems of interest are those of estimating l and R_t. For squared-error loss functions, the Bayes estimators of l and R_t are given by

$$\tilde{l} = E[m^{-1}|x] = \frac{F_r[a+\boldsymbol{x}(x), b+m-1]}{F_r[a+\boldsymbol{x}(x), b+m]} \left\{ \frac{b+m}{a+\boldsymbol{x}(x)} \right\}, \tag{8}$$

and

$$\tilde{R}_t = E[\exp(-t/m)|x] = \frac{F_r[a+\boldsymbol{x}(x)+t, b+m]}{F_r[a+\boldsymbol{x}(x), b+m]} \left\{ 1 + \frac{t}{a+\boldsymbol{x}(x)} \right\}^{-(b+m)} \tag{9}$$

The posterior risks (minimum posterior expected losses) of $\tilde{m}$, $\tilde{l}$, and $\tilde{R}_t$ are given by the posterior variances $Var[m|x]$, $Var[l|x]$, and $Var[R_t|x]$, respectively. The estimators $\tilde{l}$, $\tilde{m}$, and $\tilde{R}_t$ are not unbiased in sample theory sense. However, they are Bayes unbiased, i.e., their means coincide with the corresponding prior means.

Another loss function in popular use is the absolute-error loss function, $\mathcal{L}(m,\tilde{m})=|m-\tilde{m}|$. If this loss is deemed suitable, the posterior medians of m, l, and R_t represent the appropriate Bayes estimators. If there is no compelling reason to accept some specific loss function, we will base parameter estimation on the *ML* principle. From the Bayesian perspective, that leads to the mode of posterior density or the *HPD* estimator. Since the posterior density (5) is unimodal, the *HPD* estimator of m, denoted by $\hat{m}$, can be derived by solving the modal equation

$$(b+m+1)\hat{m}-\{a+\boldsymbol{x}(x)\}+\frac{(r-1)x_r}{\exp(x_r/\hat{m})-1}=0.$$

This equation cannot be solved explicitly (unless r=1). However, since

$$\max(0,\hat{m}-x_r/2)\le\frac{x_r}{\exp(x_r/\hat{m})-1}\le\hat{m},$$

it is obtained that $\hat{a}\le\hat{m}\le\hat{b}$, where

$$\hat{a}=\frac{a+\boldsymbol{x}(x)}{b+m+r}\qquad\text{and}\qquad\hat{b}=\min\left\{\frac{a+\boldsymbol{x}(x)}{b+m+1},\ \frac{a+\boldsymbol{x}(x)+(r-1)x_r/2}{b+m+r}\right\}.$$

Thus, the rule of false position (iterative linear interpolation) can be used to determine $\hat{m}$. Experience has shown that the rate of convergence is very rapid. In fact, it converges even faster than in Newton's method.

It should be noted that the difference between $\hat{m}$ and the *ML* estimator of m, m^*, is numerically small when $\boldsymbol{x}(x)$ and m are large in relation to a and b, respectively. Indeed, the difference $|\hat{m}-m^*|$ converges in probability to zero as $m\to\infty$.

The *HPD* estimators of l and R_t are given by

$$\hat{\boldsymbol{l}} = \frac{1}{\hat{m}_1} \quad \text{and} \quad \hat{R}_t = \begin{cases} \exp(-t/\hat{m}_2) & \text{if} \quad t < a + \boldsymbol{x}(x), \\ 0 & \text{if} \quad t \geq a + \boldsymbol{x}(x), \end{cases}$$

where $\hat{m}_1$ and $\hat{m}_2$ are the *HPD* estimators of m when the pair of hyperparameters is $(a, b - 2)$ and $(a-t, b-2)$, respectively.

It is also noted that $\hat{m}$, $\hat{\boldsymbol{l}}$, and $\hat{R}_t$ are asymptotically of minimum variance, unbiased, and normal (see Halperin, 1952, and Bhattacharyya, 1985). When $n \rightarrow \infty$ (with $q_1 = (r-1)/n$ and $q_2 = (n-s)/n$ fixed), x_r and x_s converge in probability to $\ddot{a}_1$ and $\ddot{a}_2$, respectively, where $R(\ddot{a}_1 | m) = 1 - q_1$ and $R(\ddot{a}_2 | m) = q_2$, and

$$E\left[\frac{1}{n}\sum_{i=r}^{s} x_i\right] \xrightarrow[n\rightarrow\infty]{} \int_{d_1}^{d_2} t f(t|m)\,dt = m\left[(1-q_1)\{1-\log(1-q_1)\} - q_2(1-\log q_2)\right],$$

in which the term involving $\log q_2$ drops out when $q_2 = 0$. Hence,

$$-\frac{1}{n} E\left[\frac{\partial^2 \log L(m|\mathrm{x})}{\partial m^2}\right] \xrightarrow[n\rightarrow\infty]{} \frac{c}{m^2},$$

where

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

Therefore, the asymptotic variances of $\sqrt{n}\hat{m}$, $\sqrt{n}\,\hat{\boldsymbol{l}}$, and $\sqrt{n}\hat{R}_t$, are given by m^2/c, $\boldsymbol{l}^2/c$, and $(R_t \log R_t)^2/c$, respectively.

It is well-known that *ML* estimators are modes of posterior densities corresponding to uniform priors. Therefore, the *ML* estimators of m, l, and R_t are the *HPD* estimators when the pair of hyperparameters (a, b) coincide with (0,-1), (0,1), and (t,1), respectively.

Expressions of Bayes estimators (7), (8), and (9) are considerably simplified when $r=1$ since all of the F_r can be deleted. In this case, the *HPD* estimators are also in closed forms.

Another common Bayesian approach to inference is to present credible sets (or intervals) for m. A set C on Ù such that $\Pr(m \in C \mid x)=1-a$ is called a $100(1-a)\%$ credible set (or credible interval if C is an interval) for m. Moreover, if $C=[c_1, c_2]$ and $\Pr(m \le c_1 \mid x)=\Pr(m \ge c_2 \mid x)=a/2$, then C is called a symmetric $100(1-a)\%$ credible interval for m. The interval $[c_1, c_2]$ satisfying simultaneously $S(c_1 \mid x)=1-a/2$ and $S(c_2 \mid x)=a/2$, where $S(c \mid x)$ is defined by (6), is the symmetric $100(1-a)\%$ credible interval for m. Therefore, $[1/c_2, 1/c_1]$ and $[\exp(-t/c_1), \exp(-t/c_2)]$ are the symmetric $100(1-a)\%$ credible intervals for l and R_t, respectively.

In choosing a credible set for m, it is usually desirable to minimize its size. A set $C=\{m \in$ Ù$: g(m \mid x) \ge c_á\}$, where $c_á$ is the largest constant such that $\Pr(m \in C \mid x)=1-a$, is called a $100(1-a)\%$ *HPD* credible set for m. Such a credible set is very appealing intuitively since it groups together the "most likely" values of m and is always computable from the posterior density.

For the unimodal posterior density (5), the $100(1-a)\%$ *HPD* (and shortest, of course) credible interval $[c_1, c_2]$ for m must simultaneously satisfy $S(c_1 \mid x) - S(c_2 \mid x)=1-a$ and $g(c_1 \mid x)=g(c_2 \mid x)$, which lead to the system

$$\begin{cases} G_r\left[a+\boldsymbol{x}(x), b+m; c_1\right]-G_r\left[a+\boldsymbol{x}(x), b+m; c_2\right]=(1-\boldsymbol{a})F_r\left[a+\boldsymbol{x}(x), b+m\right] \\ \dfrac{(c_2-c_1)\{a+\boldsymbol{x}(x)\}}{c_1 c_2}=(b+m+1)\log\left(\dfrac{c_2}{c_1}\right)+(r-1)\log\left\{\dfrac{1-\exp(-x_r/c_1)}{1-\exp(-x_r/c_2)}\right\} \end{cases}$$

Similarly, the $100(1-a)\%$ *HPD* credible interval $[d_1, d_2]$ for l must verify

$$\begin{cases} G_r\left[a+\boldsymbol{x}(x), b+m; d_2^{-1}\right]-G_r\left[a+\boldsymbol{x}(x), b+m; d_1^{-1}\right]=(1-\boldsymbol{a})F_r\left[a+\boldsymbol{x}(x), b+m\right], \\ (d_2-d_1)\{a+\boldsymbol{x}(x)\}=(b+m-1)\log\left(\dfrac{d_2}{d_1}\right)+(r-1)\log\left\{\dfrac{1-\exp(-d_2x_r)}{1-\exp(-d_1x_r)}\right\}, \end{cases}$$

and the 100(1–a)% *HPD* credible interval [exp(–t/w_1),exp(–t/w_2)] for R_t is given by the solution of

$$\begin{cases} G_r\left[a+\boldsymbol{x}(x), b+m; w_1\right]-G_r\left[a+\boldsymbol{x}(x), b+m; w_2\right]=(1-\boldsymbol{a})F_r\left[a+\boldsymbol{x}(x), b+m\right], \\ \dfrac{(w_2-w_1)\{a+\boldsymbol{x}(x)-t\}}{w_1w_2}=(b+m-1)\log\left(\dfrac{w_2}{w_1}\right)+(r-1)\log\left\{\dfrac{1-\exp(-x_r/w_1)}{1-\exp(-x_r/w_2}\right\}. \end{cases}$$

4. Hypothesis Testing

In general, the Bayesian approach to hypothesis testing, due primarily to Jeffreys (1961), is simpler and more sensible than the traditional approach, which is based on the ideas of Fisher, Neyman and Pearson, and also avoids several fairly substantial criticisms of the classical hypothesis testing (see, for example, Bernardo and Smith, 1994: 474-475; and Carlin and Louis, 1996: 45-47).

Suppose that one wishes to decide whether the unknown parameter *m* lies in $Ù_0$ or in $Ù_1$, where $Ù_0$ and $Ù_1$ are two disjoint subsets of the parameter space Ù. In Bayesian analysis, the task of deciding between the null hypothesis H_0: $\boldsymbol{m} \in Ù_0$ and the alternative hypothesis H_1: $\boldsymbol{m} \in Ù_1$ is conceptually straightforward. One merely calculates the posterior probabilities $p_0(x)=\Pr(\boldsymbol{m} \in Ù_0|x)$ and $p_1(x)=\Pr(\boldsymbol{m} \in Ù_1|x)$ and decides between H_0 and H_1 accordingly. Obviously, there is no need to distinguish formally between the null (working) hypothesis and the alternative. In addition, there is no limit on the number of hypotheses that may be simultaneously considered.

When the loss is "0$-k_i$" loss (i.e., $\mathcal{L}(m, a_i)=0$ if $m \in$ Ù$_i$ and $\mathcal{L}(m,a_i)=k_i>0$ if $m \in$ Ù$_{1-i}$, where a_i represents the action of accepting H_i, i=0,1), the posterior expected losses of a_0 and a_1 are $k_0 p_1(x)$ and $k_1 p_0(x)$, respectively. The Bayes decision is that corresponding to the smallest posterior expected loss. Thus, action a_i is taken if and only if *(iff)* $B_{10}(x)>(k_1 q_0)/(k_0 q_1)$, where $q_0(q_1)$ denotes the positive prior probability of Ù$_0$ (Ù$_1$), and the quantity $B_{10}(x)=\{p_1(x)/p_0(x)\}/(q_1/q_0)$, which is the ratio of the posterior odds of H_1 to the prior odds of H_1, is called the Bayes factor in favor of Ù$_1$. Usually Ù=Ù$_0 \cup$ Ù$_1$, in which case the null hypothesis H_0: $m \in$ Ù$_0$ is rejected when $p_1(x)>k_1/(k_0+k_1)$. When both hypotheses are simple (i.e., Ù$_i=\{m_i\}$, i=0,1), $B_{10}(x)$ is then just the likelihood ratio of H_1: $m=m_1$ to H_0: $m=m_0$, i.e., $B_{10}(x)=L(m_1|x)/L(m_0|x)$, and H_0 is rejected *iff*

$$\exp\{x(x)\,((1/m_0)-(1/m_1))\} > \frac{k_1 q_0}{k_0 q_1}\left(\frac{m_1}{m_0}\right)^m \left\{\frac{1-\exp(-x_r/m_0)}{1-\exp(-x_r/m_1)}\right\}^{r-1}.$$

Consider now the prior density

$$g^*(m)=\begin{cases} q_0 g_0(m) & if \quad m \in \Omega_0, \\ q_1 g_1(m) & if \quad m \in \Omega_1, \end{cases}$$

where g_i is a proper density over Ù$_i$ (e.g., g_i may be an inverted gamma density truncated in Ù$_i$) for i=0,1. With this prior density

$$B_{10}(x)=\frac{\int_{\Omega_1} L(m|x)g_1(m)dm}{\int_{\Omega_0} L(m|x)g_0(m)dm},$$

i.e., $B_{10}(x)$ is the ratio of "weighted" (by g_1 and g_0) likelihoods of Ù$_1$ to Ù$_0$.

In practical situations, one may need to assess whether some statements about m lying in a particular subset Ù$_0$ of Ù are reasonable in the absence of an alternative hypothesis. If $p_0(x)$ is sufficiently small, one will want to reject

H_0: $\mu \in \Omega_0$, and this can be expressed in terms of "significance levels." Thus, given $\alpha \in (0,1)$, a 100α % Bayesian test of significance will reject H_0: $\mu \in \Omega_0$ *iff* $p_0(x) \le \alpha$. Bernardo (1980) considers the case of a non-informative prior density and states that the posterior probability $p_0(x)$ gives a meaningful measure of the appropriateness of H_0 only if H_0 and H_1 are both simple or both composite (and of the same dimensionality).

In classical statistics, it is very common to test a point null hypothesis H_0: $\mu = \mu_0$, against the two-sided alternative H_1: $\mu \neq \mu_0$. The Bayesian viewpoint to this problem differs radically from the classical one. A typical approach (see Jeffreys, 1961, Ch. 5) to conducting a Bayesian test is to give μ_0 a positive probability q_0, while giving $\mu \in \Omega_1 = \Omega - \{\mu_0\}$ the density $q_1 g_1(\mu)$, where $q_1 = 1 - q_0$ and g_1 is a proper density of the form (4). Thus, under "$0 - k_i$" loss, the null hypothesis H_0: $\mu = \mu_0$ is rejected when

$$B_{10}(x) = \frac{\int_{\Omega_1} L(\mu|x) g_1(\mu) d\mu}{L(\mu_0|x)} > \frac{k_1 q_0}{k_0 q_1}$$

i.e., when

$$\frac{F_r\left[a+\xi(x), b+m\right] \exp\{\xi(x)/\mu_0\}}{\{a+\xi(x)\}^{b+m} \{1-\exp(-x_r/\mu_0)\}^{r-1}} > \frac{k_1 q_0}{k_0 q_1} \left\{ \frac{\Gamma(b)}{\Gamma(b+m) a^b \mu_0^m} \right\}$$

Another method for testing H_0: $\mu = \mu_0$ is described by Lindley (1965: 58-62). Intuitively, for α close to zero, the $100(1-\alpha)$% *HPD* credible interval for μ, $C_\alpha(x)$, contains the "most plausible" values of μ given the data x. Taking this into account, the null hypothesis H_0: $\mu = \mu_0$ can reasonably be accepted *iff* $C_\alpha(x)$ contains μ_0. This method is limited to cases where the prior information on μ is vague; in particular, where $q_0 = 0$.

In order to test a one-sided null hypothesis H_0: $\mu \le \mu_0$ against H_1: $\mu > \mu_0$ a reasonable loss function is "$0 - K_i(\mu)$" loss, i.e., $\mathcal{L}(\mu, a_i) = K_i(\mu) \mathrm{I}(\mu \in \Omega_{1-i})$, $i = 0,1$, where $K_0(\mu)$ and $K_1(\mu)$ are non-decreasing positive functions of $(\mu - \mu_0)$ and $(\mu_0 - \mu)$, respectively. The optimal answer to the hypothesis testing problem is to reject H_0 *iff*

$$\int_{\Omega_1} K_0(m)g(m|x)dm > \int_{\Omega_0} K_1(m)g(m|x)dm,$$

where Ù$_0=(0,m_0]$ and Ù$_1=(m_0,\infty]$. In the case of $K_0(m) = k_0(m - m_0)$ and $K_1(m) = k_1(m_0 - m)$ and prior density (4), $H_0\!: m \le m_0$ is rejected *iff*

$$\tilde{m} > m_0 + (1-(k_0/k_1))\{J(m_0;x) - m_0 p_1(x)\},$$

where $\tilde{m} = E[m|x]$, $p_1(x) = S(m_0|x)$ and

$$J(m_0;x) = \int_{m_0}^{\infty} mg(m|x)dm = \frac{G_r[a+\boldsymbol{x}(x), b+m-1; m_0]}{F_r[a+\boldsymbol{x}(x), b+m]}\left\{\frac{a+\boldsymbol{x}(x)}{b+m-1}\right\}.$$

In the symmetric case, $k_0 = k_1$, this reduces to rejecting, $H_0\!: m \le m_0$ *iff* $\tilde{m} \ge m_0$, as one might intuitively have expected.

On the other hand, under "$0 - k_i$" loss ($K_i(m)$□$k_i > 0$, $i = 0,1$) and prior density (4), the Bayes test rejects $H_0\!: m \le m_0$ *iff* $S(m_0|x) > k_1/(k_0+k_1)$, i.e., *iff*

$$G_r[a+\boldsymbol{x}(x), b+m; m_0] > \frac{k_1}{k_0+k_1} F_r[a+\boldsymbol{x}(x), b+m]$$

Reliability demonstration tests (see Martz and Waller, 1982, Ch. 10) are usually performed on a newly designed device in order to demonstrate its reliability, mean lifetime, or whether it has achieved a required level of performance. One may be interested in testing whether $R_t \ge r_0$ against the alternative $R_t < r_0$ for a specified t and r_0. Using the fact that $R_t = \exp(-t/m)$ this problem is equivalent to testing $m \ge m_0$ against the alternative $m < m_0$, where $m_0 = t(-\log r_0)^{-1}$. One may also specify some particular positive value r_1 less than r_0, which can be regarded as a clearly unacceptable reliability, and consider now the alternative $R_t < r_1$. The region $r_1 \le R_t < r_0$ is sometimes referred to as the indifference region. Problems of this nature occur quite frequently in quality control and in writing up warranties for products.

5. An Illustrative Example

As a numerical illustration, an example is presented in this section. The following data represent failure times, in minutes, for a specific type of electrical insulation (Lawless, 1982: 138 and 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 9-th failure. Hence $n=12$, $r=3$, $s=9$ and $m=7$. Assuming that the data came from an exponential distribution, three cases are considered. In Case I the Jeffreys' prior density on m is selected, i.e. $(a,b)=(0,0)$ in (4). In Cases II and III it is believed that the prior variance $Var[m]$ is 64 and 16, respectively; in both the prior mean $E[m]$ is assumed to be 80. Thus, the available prior information indicates $a=8080$ and $b=102$ in Case II, while $a=32080$ and $b=402$ in Case III.

Table I contains the Bayes (under squared-error loss) and the *HPD* estimates for m, l, and R_t (at $t=50{,}100$). The corresponding *ML* estimates are included in Table II.

Consider that it is desirable to test $H_0{:}R_{t=50} \geq r_0 = 0.50$ against $H_1{:}R_t < r_0$ under "$0-k_i$" loss, where $k_1(k_0)$ represents the positive loss incurred by committing a type I (type II) error. The null hypothesis is rejected whenever $S(-t/\log r_0|x) < k_0/(k_0+k_1)$. Therefore, $H_0{:}R_{t=50} \geq 0.50$ is rejected *iff* the ratio k_0/k_1 is greater than 1.65314, 5.39653, and 60.61412 in Cases I, II, and III, respectively.

Suppose now, for instance, that one wishes to decide between the hypotheses $H_0{:}\ m \geq m_0 = 75$ and $H_1{:}\ m < m_1 = 60$, where m_0 represents the contract specified or desired mean lifetime, while m_1 represents the minimum mean lifetime that a consumer is willing to accept. Assuming "$0-K_i(i)$" loss function, where $K_0(m)=k_0(m_1 - m)$ and $K_1(m)=k_1(m - m_0)$, and prior density (4), H_0 is rejected *iff*

$$\tilde{m} < m_1 p_1(x) + J(m_1;x) - (k_1/k_0)\{J(m_0;x) - m_0 p_0(x)\},$$

where $\tilde{m}=E[m\ |x]$, $p_0(x)=S(m_0|x)$, $p_1(x)=1-S(m_1|x)$, and $J(m_i;x)=\int_{m_i}^{\infty} mg(m|x)\,dm,\ i=0,1.$

After some algebra, this reduces to rejecting the null hypothesis whenever $\tilde{m}<c$, where c is 88.3492–17.8886(k_1/k_0), 79.7687–5.9868(k_1/k_0) and 79.9372–5.1354(k_1/k_0) in Cases I, II, and III, respectively. In particular, when $k_0=k_1$, H_0 is accepted in the three cases.

Table I: The Bayes and The *HPD* Estimates for m, l, $R_{t=50}$ and $R_{t=100}$

Case	$\tilde{m}$	$\hat{m}$	$\tilde{l}$	$\hat{l}$	$\tilde{R}_{t=50}$	$\hat{R}_{t=50}$	$\tilde{R}_{t=100}$	$\hat{R}_{t=100}$
I	86.7770	69.4073	0.01297	0.01152	0.53472	0.53748	0.29783	0.26017
II	79.7659	78.3414	0.01265	0.01254	0.53220	0.53236	0.28425	0.28136
III	79.9372	79.5492	0.01254	0.01251	0.53444	0.53449	0.28590	0.28513

Table II: *ML* Estimates for m, l, $R_{t=50}$ and $R_{t=100}$

m^*	l^*	$R^*_{t=50}$	$R^*_{t=100}$
77.1351	0.01296	0.52298	0.27361

Acknowledgment

I would like to thank the editor and referees for their careful reading of the manuscript and for their helpful suggestions.

References

Bain, L. J. (1978) *Statistical Analysis of Reliability and Life-Testing Models.* Dekker, New York.

Balakrishnan, N. (1990) “On the Maximum Likelihood Estimation of the Location and Scale Parameters of Exponential Distribution Based on

Multiply Type II Censored Samples." *Journal of Applied Statistics* 17: 55-61.

Balakrishnan, N. & Chan, P. S. (1995) "Maximum Likelihood Estimation for the Log-Gamma Distribution Under Type II Censored Samples and Associated Inference." In N. Balakrishnan (Ed.), *Recent Advances in Life-Testing and Reliability,* pp. 409-421 CRC Press.

Bernardo, J. M. (1980) "A Bayesian Analysis of Classical Hypothesis Testing." In J. M. Bernardo, M. H. De Groot, D. V. Lindley, & A. F. M. Smith (Eds.), *Bayesian Statistics: Proceedings of the First International Meeting* Valencia, Spain: Valencia University Press.

Bernardo, J. M. & Smith, A. F. M. (1994) *Bayesian Theory*. John Wiley: New York.

Bhattacharyya, G. K. (1985) "On Asymptotics of Maximum Likelihood and Related Estimators Based on Type II Censored Data." *Journal of the American Statistical Association* 80: 398-404.

Carlin, B. & Louis, T. (1996) *Bayes and Empirical Bayes Methods for Data Analysis.* Chapman and Hall: London.

Halperin, M. (1952) "Maximum Likelihood Estimation in Truncated Samples." *Annals of Mathematical Statistics* 23: 226-238.

Harter, H. L. & Moore, A. H. (1968) "Maximum Likelihood Estimation, from Doubly Censored Samples, of the Parameters of the First Asymptotic Distribution of Extreme Values." *Journal of the American Statistical Association* 63: 889-901.

Hartigan, J. A. (1964) "Invariant Prior Distributions." *Annals of Mathematical Statistics* 35: 836-845.

Jeffreys, H. (1961) *Theory of Probability*. Clarendom Press: Oxford.

Kong, F. & Fei, H. (1996) "Limits Theorems for the Maximum Likelihood Estimate Under General Multiply Type II Censoring." *Annals of the Institute of Statistical Mathematics* 48: 731-755.

Lalitha, S. & Mishra, A. (1996) “Modified Maximum Likelihood Estimation for Rayleigh Distribution.” *Communications in Statistics - Theory and Methods* 25: 389-401.

Lawless, J. F. (1982) *Statistical Models and Methods for Lifetime Data*. John Wiley: New York.

Lindley, D. V. (1965) *Introduction to Probability and Statistics from a Bayesian Viewpoint. Part 2: Inference*, C.U.P.: Cambridge.

Martz, H. F. & Waller, R. A. (1982) *Bayesian Reliability Analysis*. John Wiley: New York.

Raqab, M. Z. (1995) “On the Maximum Likelihood Prediction of the Exponential Distribution Based on Doubly Type II Censored Samples.” *Pakistan Journal of Statistics* 11: 1-10.