REGRESSION ANALYSIS OF AN EXPONENTIAL RESPONSE WITH TYPE I CENSORING

Edsel A. Pena and Ann Inez N. Gironella[1]/

1. Introduction

Life-testing experiments are usually conducted to determine the life distribution of certain materials or components. The life distribution of components is important because it gives us the reliability of the components, that is, the probability that the component will survive a certain period of time. In ordinary experiments, a sample of test units are subjected to conditions in which they are intended to operate, and their failure times are observed. After all of the test units have failed, the life distribution is then estimated.

There are situations, however, in which failure of all the tests units takes an exceedingly longer time than the period in which a decision is needed; or a long experiment might be too costly. In these cases, there is a need to use a different experimental design and an appropriate method of estimation. For example, an electronic company might want to know within three months the life distribution of a transistor it produces. But the transistor might have a lifetime beyond three months, and therefore if an experiment with ten transistors is conducted, the company may need to wait for six months before all the transistors fail.

An intuitive solution to this dilemma is to censor the experiment; that is, terminate the experiment after a specified time and estimate the life distribution based on the test units that failed and did not fail. Of course, there is a certain loss of information with this method, but the loss may be offset by the gain in being able to make an early decision. In fact, as Epstein and Sobel (1953) have shown, the best estimator of the parameter of the exponential distribution based on r failures out of n test units had exactly the same accuracy and precision as the best estimator based on r failures out of r test units. This suggests that in order to compensate for the lost information, we can simply test more units to obtain more failures. There are many variations of censoring, and the scheme described above is customarily referred to as Type I censoring.

Another technique to obtain information in a shorter period is to accelerate the failure of the units by subjecting them to higher stresses such as by

1/ Instructor and Associate Professor of Statistics, respectively. Institute of Mathematical Science and Physics, UP at Los Banos, College, Laguna.

increasing the temperature and voltage, inference about the life distribution under usual operating conditions is then made on the basis of the accelerated failure times. Clearly, regression-type methods are needed to make such inferences.

2. Problem and Assumptions

Consider an accelerated life testing experiment. Prior to the experiment, s stress levels $x_1, \ldots, x_s$ and censoring values $c_1, \ldots, c_s$ are fixed on the basis of experimental and practical constraints.

The whole experiment can be viewed as composed of s independent experiments such that uniform conditions, except for the stress values, are maintained among these s experiments. In the ith experiment, that are n_i test units. if before time c_i all the n_i units have failed, the ith experiment yields a complete (all failed) sample; while if at time c_i there are still unfailed units, also called runouts, the ith experiment is terminated and yields a Type I censored sample. Thus the data from the whole experiment is composed of the failure times Y_{ij}, $j = 1, \ldots, r_i$, $i = 1, \ldots, s$, where r_i denotes the number of failed units in the ith experiment, and the censoring times c_i of the $(n_i - r_i)$ runouts, i=1, ..., s. It should be noted that $0 \; Y_{i1} \leq \ldots \leq Y_i \; \leq c_i$.
The problem is to develop the ML estimators of β_0 and β_1 in the model

$$\theta_i = E(Y \mid X_i) = \beta_0 + \beta_1 X_i > 0 \qquad (2.1)$$

and obtain their asymptotic variance-covariance matrix. In model (2.1), θ_i is the parameter of the exponential distribution which is the assumed distribution of the lifetimes of the test units in the ith experiment.

3. The Exponential Distribution and Type I Censoring

A random variable X is said to have an exponential distribution with parameter θ if its density function is given by

$$f(x) = (1/\theta)\exp(-x/\theta)\; I_{(0,\infty)}(x),\; \theta > 0 \qquad (3.1)$$

or if its distribution function is given by

$$F(x) = 1 - \exp(-x/\theta)\; I_{(0,\infty)}(x), \quad \theta > 0 \qquad (3.2)$$

It is well-known that the mean and variance of X are θ and θ^2, respectively.

The exponential distribution is a very important distribution in reliability theory, and it has been found to model the life distribution of electronic and mechanical components quite satisfactorily (Davis, 1952). Such components are characterized by the somewhat surprising property of "old as good as new", that is, an old component is as good as new component stochastically. Stated in another way, the probability of the component surviving x units of time given that it has survived t units of time is equal to the probability of surviving x units of time given that it is a new item. Symbolically, if X denotes the life of the component, then $P(X > x \mid X > t) = P(X > x)$, which is just the "memoryless" property of the exponential distribution. It is well-known that the exponential distribution is the only distribution which possesses such property. Another characterization is that it has a constant failure rate (Barlow and Proschan, 1981).

Definition. A random variable X is said to be distributed as a truncated exponential with truncation $c > 0$ and parameter θ if its density function is given by

$$f^T(x) = \frac{\frac{1}{\theta}\exp(-\frac{x}{\theta})}{1 - \exp(-c/\theta)} I_{(0,c)}(x), \theta > 0 \tag{3.3}$$

$$= \frac{f(x)}{F(c)} I_{(0,c)}(x)$$

or if its distribution function is given by

$$F^T(x) = \frac{F(x)}{F(c)} I_{(0,c)}(x) + I_{(c,\infty)}(x) \tag{3.4}$$

If X has pdf (3.3) or df (3.4) then it has mean

$$E(X) = \theta - c\,\overline{F}(c)/F(c) \text{ where } \overline{F}(c) = 1 - F(c) \tag{3.5}$$

and variance

$$V(X) = \theta^2 - c^2\,\overline{F}(c)/F^2(c) \tag{3.6}$$

Let $X_1, \ldots, X_n$ be a random sample from the exponential distribution. Denote by $Y_1, Y_2, \ldots, Y_n$ the order statistics of $X_1, \ldots, X_n$. Then the joint density of $Y_1, \ldots, Y_2$ is given by:

$$f_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n) = n!\,(\tfrac{1}{\theta})\exp(-\tfrac{1}{\theta}\sum_{i=1}^{n} y_i), \tag{3.7}$$

$$0 < y_1 \leq \cdots \leq y_n < \infty$$

We also have the following results.

Theorem 1. The random variable R denoting the number of X_i, $i = 1, \ldots, n$ less than or equal to $c > 0$ has probability mass function

$$f_R(r) = {}_nC_r \, [F(c)]^r \, [\bar{F}(c)]^{n-r}, \quad r = 0, 1, \ldots, n$$

Proof: For any i, i = 1, ..., n, $P(X_i \leq c) = F(c)$. Since the X_i's are independent, then $f_R(r) = P(R=r) = P(\text{exactly } r \; X_i\text{'s} \leq c) = \binom{n}{r} [F(c)]^r [\bar{F}(c)]^{n-r}$, Q.E.D.

Since R is binomial with parameters n and probability of success F(c), it follows that:

Corollary: The r.v. R has expected value

$$E(R) = n\,F(c) \tag{3.8}$$

and variance

$$V(R) = n\,F(c)\,\bar{F}(c) \tag{3.9}$$

Theorem 2. The joint density of $Y_1, \ldots, Y_n$ given R = r is

$$f_{y_1, \ldots, y_{n/2}}(y_1, \ldots, y_n | r) = \frac{1}{f_p(r)}\, n!\, \left(\frac{1}{\theta}\right)^n \exp\left(-\frac{1}{\theta} \sum_{i=1}^{n} y_i\right), \tag{3.10}$$

for $0 < y_1 \leq \cdots \leq y_r \leq c < y_{r+1} \leq \cdots \leq y_n < \infty$, and r = 0, 1, ..., n.

Proof: The joint density of $Y_1, \ldots, Y_n$ given R = r is equivalent to the joint density of $Y_1, \ldots, Y_n$ conditional on the event $Y_1 \leq \cdots \leq Y_r \leq c < Y_{r+1} \leq \cdots \leq Y_n$. The probability of the conditioning event is given by (3.7), hence the theorem follows from the definitions of the conditional density function, Q.E.D.

The following lemma is a useful result and can be established by mathematical induction.

Lema: Given real numbers $0 < a \leq b < \infty$, then

$$I(a, b, n) \quad \int_a^b \int_{y_1}^b \cdots \int_{y_{n-2}}^b \int_{y_{n-1}}^b \exp\{-\sum_{j=1}^{n} y_j\} \, dy_n \cdots dy_1$$

$$= (-1)^n \frac{1}{n!} [\exp(-b) - \exp(-1)]^n \tag{3.11}$$

for any positive integer n.

Theorem 3. The joint density of $Y_1, \ldots, Y_R$ given R=r is

$$f_{y_1, \ldots, Y_{R/R}}(y_1, \ldots, y_r/r) = \frac{1}{f_R(r)}\, {}_nP_r \left(\frac{1}{\theta}\right)^r \exp\left(-\frac{S}{e}\right) \tag{3.12}$$

where $S = \sum_{j=1}^{r} y_j + (n-r)c$

and $P_r = \frac{n!}{(n-r)!}$

Proof:

$$f_{y_1,\ldots,y_{R/R}}(y_1,\ldots,y_r/r) = \int_c^\infty \int_{y_{r+1}}^\infty \cdots \int_{y_{n-1}}^\infty f_{y_1,\ldots,y_{n/R}}(y_1,\ldots y_{n/r})\, dy_n \ldots dy_{r+1}$$

$$= \frac{1}{f_R(r)}\, n!\, \left(\frac{1}{\theta}\right)^r \exp\left(-\frac{1}{\theta}\sum_{j=1}^{r} y_j\right) \times$$

$$\int_c^\infty \int_{y_{r+1}}^\infty \cdots \int_{y_{n-1}}^\infty \left(\frac{1}{\theta}\right)^{n-r} \exp\left(-\frac{1}{\theta}\sum_{j=r+1}^{n} y_j\right) dy_n \ldots dy_{r+1}$$

$$= \frac{1}{f_R(r)}\, n!\, \left(\frac{1}{\theta}\right)^r \exp\left(-\frac{1}{\theta}\sum_{j=1}^{r} y_j\right) \times \lim_{b\to\infty} I\left(\frac{c}{\theta}, \frac{b}{\theta}, n-r\right)$$

$$= \frac{1}{f_R(r)}\, n!\, \left(\frac{1}{\theta}\right)^r \exp\left(-\frac{1}{\theta}\sum_{j=1}^{r} y_j\right) \times \frac{1}{(n-r)!}\left[-\frac{(n-r)c}{\theta}\right]$$

$$= \frac{1}{f_R(r)}\ \frac{n!}{(n-r)!}\left(\frac{1}{\theta}\right)^r \exp - \frac{\sum_{j=1}^{r} y_j + (n-r)c}{\theta}$$

Corollary 2. The order statistics $Y_1, \ldots Y_r$ given that $R = r$ are the order statistics of a sample of size r from the truncated exponential distribu-distribution. Consequently,

$$\sum_{j=1}^{r} Y_j = \sum_{j=1}^{r} X_j ,$$

Proof From (3.12), we obtain

$$f_{y_1,\ldots,Y_{R/R}}(y_1,\ldots,y_{r/r}) = \frac{\frac{n!}{(n-r)!}\left(\frac{1}{\theta}\right)^r \exp - \frac{\sum_{j=1}^{r} Y_j + (n-r)c}{\theta}}{\frac{n!}{r!(n-r)!}\left[\bar{F}(c)\right]^r \left[F(c)\right]^{n-r}}$$

$$= \frac{r!\left(\frac{1}{\theta}\right)^r \exp\left(-\frac{1}{\theta}\sum_{j=1}^{r} Y_j\right)}{\left[F(c)\right]^r}$$

$$= r! \sum_{i=1}^{r} f^T(y_i), \quad 0 < y_1 \le \ldots \le y_r \le c$$

Therefore the conditional joint density of $Y_1,\ldots, Y_r$ given $R = r$ is equal to the joint density of the order statistics of size r from the truncated exponential distribution. Clearly, $\sum_{j=1}^{r} Y_j = \sum_{j=1}^{r} X_j$ Q.E.D.

From an experimental point of view, $S = \sum_{j=1}^{r} Y_j + (n-r)c$

is the total time on test of n test units at the time of censoring.

Theorem 4. The joint density function of $Y_1, \ldots, Y_R$, R is

$$f_{Y_1, \ldots, Y_R, R}(y_1, \ldots, y_r, r) = {}_nP_r \left(\frac{1}{\theta}\right)^r \exp\left(-\frac{s}{\theta}\right) \quad (3.13)$$

for $r = 0, 1, \ldots, n$; $0 < y_1 \leq \ldots \leq y_r \leq c$.

Proof: Since $f_{Y_1,\ldots,Y_{R,R}}(y_1,\ldots,y_r,r) = f_R(r) f_{Y_1,\ldots,Y_R|R}(y_1,\ldots,y_{r/r})$.
The theorem follows from theorems 1 and 3. Q.E.D.

By the factorization theorem, the statistic $S = \sum_{j=1}^{R} Y_j + (n-R)c$ is not a sufficient statistic for θ, but (S, R) is jointly sufficient for θ.

Theorem 5. The statistic $S = \sum_{j=1}^{R} Y_j + (n-R)c$ has expected value

$$E(S) = \theta n F(c) = \theta E(R) \quad (3.14)$$

and

$$V(S) = \theta n F(c) \{\theta (1+\bar{F}(c) - 2c \frac{\bar{F}(c)}{F(c)}\} \quad (3.15)$$

Proof: First note that

$$E(S|R) = E(\sum_{j=1}^{R} Y_j/R) + (n-R)c$$

By corollary 2, it follows that

$$E(\sum_{j=1}^{R} Y_j/R) = E(\sum_{j=1}^{R} Y_j|R) = \sum_{j=1}^{R} E(X_j|R)$$

$$= \sum_{j=1}^{R} \{\theta - \frac{c\bar{F}(c)}{F(c)}\} = R\{\theta - c\frac{\bar{F}(c)}{F(c)}\}.$$

Therefore, $E(S|R) = R(\theta - \frac{c\bar{F}(c)}{F(c)}) + (n-R)c = R(\theta - \frac{S}{F(c)}) + nc$,

Since $E(S) = E[E(S|R)]$ then
$E(S) = nc + (\theta - \frac{S}{F(c)}) E(R)$

$$= nc + \left(\theta - \frac{S}{F(c)} \right) nF(c)$$

$$= \theta nF(c) = \theta E(R)$$

For the variance, recall that

$$V(S) = E(V(S|R)) + V(E(S|R)).$$

We have $V(E(S|R)) = (\theta - \frac{c}{F(c)})^2 V(R) = (\theta - \frac{c}{F(c)})^2 nF(c) \bar{F}(c)$

On the other hand,

$$V(S/R) = V\left(\sum_{j=1}^{R} Y_j/R \right) = V\left(\sum_{j=1}^{R} X_j/R \right) = \sum_{j=1}^{R} V(X_j/R)$$

$$= R \left(\theta^2 - c^2 \frac{\bar{F}(c)}{F^2(c)} \right)$$

Therefore, $E(V(S|R) = \left| \theta^2 - c^2 \frac{\bar{F}(c)}{F^2(c)} \right| nF(c)$

hence

$$V(S) = nF(c) \{ \theta^2 - c^2 \frac{\bar{F}(c)}{F^2(c)} + (\theta - \frac{c}{F(c)})^2 \bar{F}(c) \}$$

$$= nF(c) \{ \theta^2 (1 + \bar{F}(c) - 2c \frac{\bar{F}(c)}{F(c)} \}$$

$$= \theta nF(c) \{ \theta^2 (1 + \bar{F}(c) - 2c \frac{\bar{F}(c)}{F(c)} \}$$

4. The ML Estimators

The likelihood function of the sample described in Section 2 is

$$L = L(\beta_0, \beta_1) = \prod_{i=1}^{s} L_i$$

where L_i, $i = 1, \ldots, s$ is the likelihood function of the ith sample. By theorem 4, the log- likelihood function

$$\ell = \sum_{i=1}^{s} \log (n_i P_{r_i}) - \sum_{i=1}^{s} r_i \log \theta_i - \sum_{i=1}^{s} \frac{S_i}{\theta_i} \qquad (4.1)$$

where $S_i = \sum_{j=1}^{r} Y_{ij} + (n_i - r_i) c_i$, and $\theta_i = \beta_0 + \beta_1 X_i$.

The ML estimators of β_0 and β_1, denoted by b_0 and b_1, respectively, are values of β_0 and β_1 maximizing (4.1) and satisfying $\beta_0 + \beta_1 X_i > 0$ for $i = 1, \ldots, s$. The first order partials of (4.1) are

$$\ell_0 = \frac{\partial \ell}{\partial \beta_0} \quad - \sum_{i=1}^{s} \frac{r_i}{\theta_i} + \sum_{i=1}^{s} \frac{S_i}{\theta_i^2}$$

$$\ell_1 = \frac{\partial \ell}{\partial \beta_1} = - \sum_{i=1}^{s} \frac{r_i x_i}{\theta_i} + \sum_{i=1}^{s} \frac{S_i x_i}{\theta_i^2} \qquad (4.2)$$

Both ℓ_0 and ℓ_1 are nonlinear functions of β_0 and β_1, hence it is practically impossible to obtain close forms of the ML estimators. Estimates are therefore obtained iteratively.

Using the Newton-Raphson procedure, which is based on first-order Taylor approximations of ℓ_0 and ℓ_1, the ML estimates are obtained by solving iteratively the simultaneous equations below in $b_0^{(k+1)}$ and $b_1^{(k+1)}$ and where $b_0^{(k+1)}$ and $b_1^{(k+1)}$ are the (k+1)th iterates. The equations are

$$A_{11}^{(k)} b_0^{(k+1)} + A_{12}^{(k)} b_1^{(k+1)} = B_1^{(k)}$$

$$A_{12}^{(k)} b_0^{(k+1)} + A_{22}^{(k)} b_1^{(k+1)} = B_2^{(k)}$$

where $$A_{11}^{(k)} = \sum_{i=1}^{s} \frac{r_i}{{}^{(k)}\theta_i^2} - 2 \sum_{i=1}^{s} \frac{S_i}{{}^{(k)}\theta_i^3}$$

$$A_{12}^{(k)} = \sum_{i=1}^{s} \frac{r_i x_i^2}{{}^{(k)}\theta_i^2} - 2 \sum_{i=1}^{s} \frac{S_i x_i^2}{{}^{(k)}\theta_i^3}$$

$$A_{22}^{(k)} = \sum_{i=1}^{s} \frac{r_i x_i}{{}^{(k)}\theta_i^2} - 2 \sum_{i=1}^{s} \frac{S_i x_i^2}{{}^{(k)}\theta_i^3}$$

$$B_1^{(k)} = \sum_{i=1}^{s} \frac{r_i}{{}^{(k)}\theta_i} - \sum_{i=1}^{s} \frac{S_i}{{}^{(k)}\theta_i^2}$$

$$B_2^{(k)} = \sum_{i=1}^{s} \frac{r_i x_i}{{}^{(k)}\theta_i} - \sum_{i=1}^{s} \frac{S_i x_i}{{}^{(k)}\theta_i^2}$$

and $${}^{(k)}\theta_i = b_0^{(k)} + b_1^{(k)} x_i$$

When using this method, provision must be made to ensure that the initial iterates $(b_0^{(0)}, b_1^{(0)})$, which the user should supply, is "near" the solution (b_0, b_1). Furthermore, at each iteration, the iterates must satisfy $b_0^{(k)} + b_1^{(k)}x_i > 0$ for i = 1, ..., s. The authors believe that this algorithm is not effective when the true values of β_0 and β_1 are rather small, say in the range from 0 to 10. Indeed, this method is theoretically inefficient because so much information has been lost by using a linear approximation to the nonlinear functions ℓ_0 and ℓ_1.

A second-order method can be employed and an algorithm for this method is described below:

Let $\underline{b}^T = (b_0, b_1)$, $\underline{g}(\underline{b})^T = (\ell_0(\underline{b}), \ell_1(b))$

and $\underline{H}(\underline{b}) = \begin{bmatrix} \ell_{00}(b) & \ell_{01}(b) \\ \ell_{00}(b) & \ell_{11}(b) \end{bmatrix}$ be the vector of iterates,

the gradient vector, and the Hessian matrix, respectively. The second order partials are:

$$\ell_{00} = \frac{\partial^2 \ell}{\partial \ell_0^2} = \sum_{i=1}^{s} \frac{r_i}{\theta_i^2} - 2\sum_{i=1}^{s} \frac{s_i}{\theta_i^3}$$

$$\ell_{01} = \frac{\partial^2 \ell}{\partial \ell_0 \partial \ell_1} = \sum_{i=1}^{s} \frac{r_i x_i}{\theta_i^2} - 2\sum_{i=1}^{s} \frac{s_i x_i}{\theta_i^3} \qquad (4.4)$$

$$\ell_{11} = \frac{\partial^2 \ell}{\partial \beta_1^2} = \sum_{i=1}^{s} \frac{r_i x_i^2}{\theta_i^2} - 2\sum_{i=1}^{s} \frac{s_i x_i^2}{\theta_i^3}$$

A superscript of T and -1 will denote matrix transpose and inverse, respectively.

Let $\underline{b}^{(k)}$ be the vector of iterates after the kth iteration. To determine $\underline{b}^{(k+1)}$, a direction vector is first obtained as follows:

$$\underline{d}^{(k)} = S' \; \underline{H}(\underline{b}^{(k)})^{-1} \; \underline{g}(\underline{b})^{(k)}) + S^2\underline{e}^k$$

where

$$S^1 = \text{sign}\,[-\{g(b^{(k)})\}^T \{\underline{H}(\underline{b}^{(k)})\}^{-1} \{g(\underline{b}^{(k)})\}],$$

$$S^2 = \text{sign}\,[\{g(\underline{b}^{(k)})\}\;\underline{e}^*],$$

$$e^* + \max\;(\lambda_1, 0)\;\underline{e}_1,$$

$$\lambda_1 = \text{largest eigenvalue of } \underline{H}(\underline{b}^{(k)}), \text{ and}$$

$$\underline{e}_1 = \text{eigenvector associated with } \lambda_1.$$

Then $\underline{b}^{(k+1)}$ is obtained by maximizing the log-likelihood function in (4.1) in the direction of $\underline{d}^{(k)}$, that is, $\underline{b}^{(k+1)} = \underline{b}^{(k)} + \underline{\tau}^* \underline{d}^{(k)}$

where $\tau^* \in T = \{\tau : \max (\underline{b}^{(k)} + \underline{d}^{(k)}),\ 0 < \alpha < \infty\}$

The Fibonacci Search Algorithm is used to obtain τ^*. This algorithm is a variant of the Modified Newton Procedure and the Second-Order Optimum Technique (see Zangwill (1969)).

By theorem 5. $E(\ell_0) = E(\ell_1) = 0$, and $E(\ell_{00})$, $E\ \ell_{01})$ and $E(\ell_{11})$ exist and are nonzero. Therefore, for large samples (see Kendall and Stuart, 1961), the ML estimators (b_0, b_1) is distributed as bivariate normal with mean (β_0, β_1) and asymptotic variances and covariance given by:

$$\mathrm{Var}(b_0) = \Delta^{-1} \sum_{i=1}^{s} n_i F_i(c_i) X_i^2 / \theta_i^2$$

$$\mathrm{Var}(b_1) = \Delta^{-1} \sum_{i=1}^{s} n_i F_i(c_i) / \theta_i^2 \qquad (4.5)$$

$$\mathrm{Cov}(b_0, b_1) = -\Delta^{-1} \sum_{i=1}^{s} n_i F_i(c_i) X_i^2 / \theta_i^2$$

where $$\Delta = \{\sum_{i=1}^{s} n_i F_i(c_i) / \theta_i^2\}\{\sum_{i=1}^{s} n_i F_i(c_i) X_i^2 / \theta_i^2\} - \{\sum_{i=1}^{s} n_i F_i(c_i) X_i / \theta_i^2\}^2$$

$$\theta_i = \beta_0 + \beta_1 x_i$$

and $F_i(c_i) = 1 - \exp(-c_i/\theta_i)$.

In order to obtain estimates of the asymptotic variances and covariances, $\hat{\theta}_i = b_0 + b_1 x_i$ is substituted in (4.5) for θ_i.

<u>Numerical Example</u>. Table 4.1 contains the resulting data from a computer simulated experiment, the experiment being of the type described in Section 2. The model is that at each stress level X_i, the response variable is exponentially distributed with mean $\theta_i = 2 - x_i$.

This data was analyzed using the ML method developed in the preceding section, and a summary of the iterations is presented in Table 4.2. The initial iterates were $b_o^{(0)} = 2.0$ and $b_1^{(0)} = -1.0$, the true values of the model,

Table 4.1. Simulated Data for Numerical Example

Stress Level (X_i)	Units Tested (n_i)	Uncensored Units (r_i)	Censoring Time (c_i)	Observed Values (Y_{ij})
0.0	1	0	1.39	-
0.25	1	1	1.39	0.14
0.50	1	0	1.39	-
0.75	1	0	1.39	-
1.00	1	1	1.39	0.24

Table 4.2. Summary of Iterations for ML Method

Iteration Number (k)	b_0	b_1	Value of log-likelihood function - Constant
0	2.000000	-1.000000	-3.603589
1	3.158213	-1.658252	-3.582346
2	2.876690	-2.304753	-3.242906
4	3.286557	-2.842378	-3.086101
7	4.495989	-3.944261	-2.887740
10	5.495789	-5.158663	-2.648423
13	8.105762	-7.819685	-2.531618
16	8.635826	-8.388651	-2.518728
18	8.872850	-8.626531	-2.517985
19	8.967540	-8.719568	-2.517907
20	8.965623	-8.719618	-2.517878

It is interesting to note that the ML estimates are quite far from the true values of the model of 2.00 and -1.00. However, this was one of those samples in the simulation experiment which gave estimates which were far from the true values. A true example of how randomness can disappoint us!

5. Results and Discussion

Due to the difficulty of obtaining a closed form of the ML estimators, a computer simulation experiment, was conducted to compare the ML estimators and the ILS estimators (see Scheme and Hohn (1979) for discussion of ILS estimators), over varying sample factors.

The assumptions of the simulation experiment were identical to those set forth in Section 2. The parameters β_0 and β_1 have values 2.00 and -1.00, respectively. Testing was conducted uniformly over the stress interval $[0, 1]$, that is, ranging from the minimum stress of $X_1 = 0$ (with $\theta_1 = 2$) to the maximum stress of $X_s = 1$ (with $\theta_s = 1$).

The sample factors considered in the experiment were the following:

1. The censoring probability at the minimum stress level. This determined the censoring time at each stress level, since it was assumed that the censoring times were identical. For a given p, the censoring time is c = -2 log p.

2. The number of stress level s. This include the minimum and maximum stresses of $X_1 = 0$ and $X_s = 1$. The stress levels were equally spaced over the stress interval /0, 1/.

3. The overall sample size $n = \sum_{i=1}^{s} n_i$, where n_i is the number of units on the ith stress level.

Due to limited resources, the simulation was conducted on only 3 combinations of the factors. These combinations are enumerated in Table 5.1

Table 5.1 Combinations for Simulation Experiment

Combination	p	s	n_i	n
1	0.50	5	1 each	5
2	0.75	5	1 each	5
3	0.50	20	1 each	20

Table 5.2 presents a summary of the features of the simulation experiment. The number of samples analyzed in each of the combinations vary due to the setting of upper limits for the number of iterations and computer running time. Under the heading "Successfully analyzed" are the number of samples in which estimates were obtained. In both the ML and ILS methoas, there were inherently unanalyzable samples, which are samples which have at most one uncensored value. Notice that combination 2 has the highest rate of unanalyzable samples (43%). This is so because the censoring probability was 0.75 hence more values were censored. Aside from these unanalyzable samples, there were others where the ML method did not produce estimates. These are the ones in which the ML method did not converge, and their numbers can be found under the heading "nonconvergent". The column "all uncensored" represents the number of samples where all the values are uncensored.

The estimates of the parameters of the sampling uncensored (s.d.'s) of the estimators of β_0 are presented in Table 5.3. It should be pointed out at this stage that all the numbers in Tables 5.3 - 5.6 are just estimates of the parameters of the sampling distributions, hence are subject to random variation themselves. The root-mean-square error is just the square root of the mean-square-error. Under the heading "quantiles", q_1 and q_3 denotes the estimates of the 0.25th and 0.75th quantiles of the s.d.'s of estimators of β_1.

Table 5.2 Summary of Features of the Simulation Experiment

Combination	Factors				Samples Analyzed	Successfully Analyzed	
	p	s	n_i	n		ML	ILS
1	0.50	5	1 each	5	130	99	118
2	0.75	5	1 each	5	200	105	114
3	0.50	20	1 each	20	1[illegible]0	167	180

Table 5.2 continued

Unanalyzable Samples[a]	Nonconvergent[b] (for ML)	All Uncensored[c]	Mean No. of Uncensored Values	
			All Samples	Convergent Only
12	9	10	2.98	3.15
86	9	0	1.66	2.43
0	13	0	12.41	12.52

[a]Samples which has at most one uncensored value.
[b]Convergence not achieved at the upper limit of number of iterations
[c]No censored values (complete sample).

Table 5.3 Statistics of Estimators of β_0 (β_0 = 2.00)

Combination	Method	Mean(t-value)[a]	Standard Deviation	RMSE	Quantiles				
					Min	q_1	Med	q_3	Max
1	ML	2.55(2.15**)	2.58	2.64	0.08	0.68	1.67	3.21	10.20
	ILS	1.31(-11.2***)	0.67	0.96	-0.08	0.78	1.27	1.82	3.15
2	ML	1.52(-3.58**)	1.36	1.44	0.01	0.39	1.09	2.92	4.64
	ILS	0.69(48***)	0.29	1.36	0.05	0.49	0.73	0.91	1.22
3	ML	2.10(1.23)	1.09	1.09	0.17	1.43	1.92	2.63	4.75
	ILS	1.30(-2.65***)	0.35	0.78	0.49	1.09	1.29	1.49	2.43

[a]t-values with (*) are significant at = 0.10
t-values with (**) are significant at = 0.05
t-values with (***) are significant at = 0.01

Tables 5.4.1-5.4.3 presents the estimators of the parameter of the s.d.'s but classified according to the number of censored values in the samples.

The remaining tables and figure presents a similar information as the preceding ones except that they pertain to the estimators of β_1.

The conclusions which can be deduced from the results of the simulation experiment are limited to the assumptions of the experiment, in particular, that $\beta_0 = 2.0$ and $\beta_1 = -1.0$. In essence, this is one of the limitations of the experiment, and an extensive comparison of the ML and ILS estimators requires further study.

Examining Table 5.3, there can be no doubt that the ILS estimator of β_0 is negatively biased. The tests of significance, although not very appropriate under the circumstances because of an observed assymmetry of the sampling distributions, showed highly significant results. Furthermore, the true values of β_0 is not contained in the interquantile ranges for all the three combinations.

Table 5.4.1 Statistics of Estimators of $\beta_0(\beta_0 = 2.0)$ Classified According to the Number of Censored Values for Combination 1

Number of Censored	Method	Number of Samples	Mean	Standard Deviation	Minimum Value	Maximum Value
0	ML	8	0.80	0.37	0.26	1.36
	ILS	10	0.73	0.25	0.22	1.22
1	ML	26	1.03	0.57	0.13	2.25
	ILS	33	0.86	0.44	-0.08	1.55
2	ML	38	2.29	1.28	0.18	4.31
	ILS	46	1.43	0.50	0.39	2.34
3	ML	27	4.91	3.64	0.36	10.20
	ILS	29	1.84	0.74	0.05	3.15

Table 5.4.2 Statistics of Estimates of $\beta_0(\beta_0 = 2.0)$ Classified According to the Number of Censored Values for Combination 2

Number of Censored	Method	Number of Samples	Mean	Standard Deviation	Minimum Value	Maximum Value
1	ML	10	0.52	0.36	0.06	1.23
	ILS	10	0.40	0.21	0.11	1.78
2	ML	25	0.81	0.52	0.01	1.65
	ILS	29	0.61	0.23	0.20	1.90
3	ML	70	1.92	1.48	0.01	9.64
	ILS	75	0.75	0.30	0.05	1.22

Table 5.4.3 Statistics of Estimates of $\beta_0(\beta_0 = 2.0)$ Classified According to the Number of Censored Values for Combination 3

Number of Censored	Method	Number of Samples	Mean	Standard Deviation	Minimum Value	Maximum Value
4	ML	10	1.05	0.33	0.50	1.65
	ILS	10	0.93	0.28	0.57	1.50
5	ML	15	1.53	0.42	0.74	2.00
	ILS	15	1.12	0.22	0.67	1.41
6	ML	28	1.55	0.90	0.77	3.18
	ILS	30	1.08	1.22	0.72	1.53
7	ML	30	1.89	0.66	0.29	3.13
	ILS	30	1.26	0.29	0.69	1.91
8	ML	33	2.19	0.77	0.17	3.23
	ILS	38	1.38	0.29	0.64	1.95
9	ML	19	2.26	0.93	0.75	3.91
	ILS	20	1.32	0.26	0.79	1.86
10	ML	17	3.03	1.02	1.63	4.59
	ILS	18	1.56	0.27	1.16	2.13
11	ML	5	3.56	1.17	1.91	4.75
	ILS	7	1.69	0.22	1.30	2.01

On the other hand, the ML estimator of β_0 is biased for combinations 1 and 2. In combination 3, however, there is no reason to conclude that it is biased. In fact, the mean values is 2.10 with standard error of 0.08. By inspection of Table 5.2, we notice that combination 3 has no unanalyzable samples, while the other two have. It seems logical to attribute the significant boas of combination 2 to the 86 unanalyzable samples, because these are the samples which had large values, and hence were supposed to yield higher estimates.

Looking at Table 5.4.1, we notice that the cause of the positive bias for combination 1 were the large estimates obtained from samples with 3 censored values. This leads to the conjecture that the ML estimator gives large estimates when there are many censored values. But in Table 5.4.2, the maximum value of the estimates obtained from highly censored samples is only 4.64, and in Table 5.4.3, there are not many large estimates. It therefore seems that the sampling distribution of the ML estimator of β_0 is positively skewed, making the probability of large values small, but still possible. Note also that the true value of β_0 is contained in the interquantile ranges of the ML estimator.

Examining the root-mean-square error (RMSE's) of the ML and ILS estimators of β_0, we notice that the RMSE of the ILS estimator is smaller than the RMSE of the ML estimator in all three combinations. The reason is that the ILS estimator has very small variance relative to that of the ML estimator. This precludes the immediate conclusion that the ML estimator is preferable over the ILS estimator.

Inspecting Table 5.5, we notice that the ILS estimtor of β_1 is positively biased for all three combinations. On the other hand, the ML estimator is unbiased for combinations 1 and 3, and positively biased for combination 2. Again, we explain this as due to the unanalyzable samples. It is worthwhile noting that for β_0 we had negative bias while for β_1 we had positive bias. This is so because the estimators of β_0 and β_1 are always negatively correlated.

Table 5.5.1 Statistics of Estimate of $\beta_1 (\beta_1 = -1.0)$

Combination	Method	Mean(t-value)[a]	Standard Deviation	RMSE	Quantiles Min	q_1	Med	q_3	Max
1	ML	-0.98(0.06)	3.78	3.78	-9.81	-2.63	-0.75	0.99	8.79
	ILS	-0.38(6.14***)	1.09	1.25	-2.70	-1.10	-0.50	0.36	2.92
2	ML	-0.34(2.81***)	2.39	2.48	-4.41	-2.42	-0.49	-0.89	4.65
	ILS	-0.19(15.5***)	0.56	0.99	-1.17	-0.65	-0.30	0.21	1.60
3	ML	-1.03(-0.25)	1.46	1.46	-6.52	-1.81	-1.01	-0.15	2.84
	ILS	-0.46(13.1***)	0.55	0.78	1.98	0.77	0.48	0.16	0.79

[a]t-values with (*) are significant at = 0.10
t-values with (**) are significant at = 0.05
t-values with (***) are significant at = 0.01

Table 5.5.2 Statistics of Estimates of $\beta_1 (\beta_1 = 1.0)$ Classified According to the Number of Censored Values for Combination 1

Number of Censored	Method	Number of Samples	Mean	Standard Deviation	Minimum Value	Maximum Value
0	ML	8	-0.52	0.51	-1.32	0.11
	ILS	10	-0.39	0.48	-0.97	0.52
1	ML	26	-0.17	1.01	-1.741	1.47
	ILS	33	-0.17	0.81	-1.54	1.15
2	ML	38	-0.92	2.46	-4.11	3.55
	ILS	46	-0.55	1.08	-2.44	1.73
3	ML	27	-1.98	6.51	-9.81	8.79
	ILS	29	-0.37	1.48	-2.70	2.92

Table 5.5.3 Statistics of Estimates of $\beta_1(\beta_1 = 1.0)$ Classified According to the Number of Censored Values for Combination 2

Number of Censored	Method	Number of Samples	Mean	Standard Deviation	Minimum Value	Maximum Value
1	ML	10	-0.20	0.56	-1.19	0.41
	ILS	10	-0.11	0.36	-0.66	0.26
2	ML	25	-0.05	.112	-1.53	1.89
	ILS	33	-0.24	0.50	-0.17	0.62
3	ML	70	-0.47	2.84	-4.41	4.65
	ILS	75	-0.17	0.61	-1.18	1.60

Table 5.5.4 Statistics of Estimates of $\beta_1(\beta_1 = -1.0)$ Classified According to the Number of Censored Values for Combination 3

Number of Censored	Method	Number of Samples	Mean	Standard Deviation	Minimum Value	Maximum Value
4	ML	10	-0.17	0.68	-1.33	1.16
	ILS	10	-0.22	0.55	-1.37	0.62
5	ML	15	-0.82	0.74	-1.69	0.47
	ILS	15	-0.45	0.41	-1.01	0.20
6	ML	28	-0.62	0.98	-3.08	0.94
	ILS	30	-0.29	0.43	-1.30	0.64
7	ML	20	-0.97	1.22	-2.85	2.56
	ILS	30	-0.50	0.57	-1.85	0.53
8	ML	33	-1.11	1.43	-2.87	2.84
	ILS	38	-0.55	0.57	-1.53	0.76
9	ML	19	-0.93	1.66	-3.68	2.27
	ILS	20	-0.33	0.53	-1.26	0.78
10	ML	17	-1.78	1.62	-4.25	0.74
	ILS	18	-0.58	0.54	-1.67	0.15
11	ML	5	-2.01	1.84	-3.80	0.83
	ILS	7	-0.65	0.45	-1.35	0.04

Notice that all the interquartile ranges of the ML estimator contain -1.0, the true value of β_1, while the interquartile ranges of the ILS estimator do not contain -1.0, except in combination 1. But, just like in the case of β_0, the ILS estimator has smaller RMSE than the ML estimator.

On the basis of these observations, we can conclude that the ML estimators tend to be unbiased when there are not, any unanalyzable samples, while the ILS estimators are biased. However, the ML estimators are very variant relative to the ILS estimators, making its RMSE greater than the RMSE of the ILS estimators.

The question is which is preferable between the two methods? As was pointed out in the beginning of this subsection, we can only make conclusions with respect to the given parameter values of $\beta_0 = 2.0$ and $\beta_1 = -1.0$. In order to make a general choice, it would entail examination of the performances of the estimators for other parameter values. In particular, if we change the values of β_0 and β_1, does it increase the bias of the ILS estimators, or decrease the variance of the ML estimators? What happens to their RMSE's at different values of β_0 and β_1? This we cannot answer at this stage. Going back to the particular case of $\beta_0 = 2.00$ and $\beta_1 = -1.00$, there is reason to prefer the ML estimator over the ILS estimator because of its unbiasedness. In the context of accelerated life tests, we would prefer unbiased estimators so that we could also unbiasedly estimate the parameter values at the usual operating conditions, which are estimated by extrapolation. Another reason for preferring the ML estimators is that we do not really know the magnitude of the bias of the ILS estimator. That bias might depend on θ, n, s, and p. If we know the magnitude of bias, then the ILS would be preferable because then we can just correct for the bias.

Thus, it is not advisable to use the methods developed for the normal case because it might lead to very biased results. Even for large samples (combination 3), the ILS estimators are biased, hence it seems that it would also hold true in the asymptotic sense.

7. References

Aitkin, M. 1981. A note on the regression analysis of censored data. Technometrics 23, 161-163.

Barlow, R. and Froschan, F. 1981. Statistical Theory of Reliability and Life Testing. Second Printing Publisher. To Begin With, 1137 Homell Drive, Silver Spring, MD. 20904

Davis, D. 1952. An Analysis of Some Failure Data. JASA. 47:113-150.

Epstein, B. and Sobel, M. 1955. Life Testing JASA, 48:485-502.

Kendall, M.G. and Stuart, A. 1961. The Advanced Theory of Statistics. Vol. 2, Inference and Relationship. NY Hofner Publishing Co. Inc.

Schemee and Hahn. 1979. A Simple Method for Regression Analysis with Censored Data. Technometrics. 21 417-432.

Zangwill. 1969. Nonlinear Programming: A Unified Approach. NJ: Prentice-Hall, Englewood Cliffs.