BAYESIAN NONPARAMETRIC ESTIMATION OF FAILURE RATES WITH CENSORED DATA

J. S. RAO and R. C. TIWARI*

University of California, Santa Barbara

ABSTRACT: The failure time distribution is estimated in the nonparametric context when some of the observations are censored. The time interval is partitioned into fixed class intervals, and number of failures and number censored in each of these intervals are observed. Using a Dirichlet distribution as the prior, the resulting estimates of the survival function and the failure rate have a nice and simple form. If instead of the fixed time intervals, one uses the "natural" intervals formed by the observed failure times, this gives essentially the same results as in Ferguson and Phadia (1977), Susarla and Van Ryzin (1976), but in a much simpler way. Bayes estimation under the increasing and decreasing failure rates is also considered, and applications to accelerated life testing are discussed.

1. Introduction

Let T be a non-negative random variable representing the failure time, with distribution function F; it may be discrete or continuous or a mixture of both. The survival function $\bar{F}$ is given by

$$\bar{F}(t) = \text{Prob}\,(T \geqslant t) = 1 - F(t-0)$$

and the hazard or age-specific failure rate of F is given by

$$\lambda(t) = \lim_{h \to 0+} \frac{\text{Prob}\,(t \leqslant T < t+h \mid T \geqslant t)}{h}$$

if the limit exists. If F is continuous and λ is integrable, then

$$\bar{F}(t) = \exp\left\{ -\int_0^t \lambda(s)\, ds \right\}$$

Furthermore, if F has density f then

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)}$$

* Now in the Dept of Mathematics, I. I. T., Bombay.

Given a continuous failure time distribution F, consider a partition $\{(t_i, t_{i+1}]\}_{i=0}^{k}$ of $(0, \infty)$ with $t_0=0$ and $t_{k+1}=\infty$. Consider a set of N individuals. We take observations at the k time points $t_1 < t_2 < \ldots < t_k$ (say at the end of each hour, day or similar period, not necessarily of equal lengths) and observe how many individuals failed and how many were censored (e.g., left the study) in each of these intervals. Let n_i denote the number of failures and m_i the number censored in the interval $(t_i, t_{i+1}]$, $i=0,1,\ldots,k-1$. For definiteness, we shall assume that those censored in the interval $(t_i, t_{i+1}]$ survived past t_{i+1}. Without loss of generality, we can and shall assume that t_k is sufficiently large that no deaths or censorings occur beyond t_k so that $n_k=0$ and $m_k=0$. Let $n=\sum_{i=0}^{k-1} n_i$, $m=\sum_{i=0}^{k-1} m_i$ so that $N=n+m$. Define $n_{(i)}=\sum_{j=i}^{k-1} n_j$ and $m_{(i)}=\sum_{j=i}^{k-1} m_j$. Then $N_{(i)}=n_{(i)}+m_{(i)}$ is the number of individuals at risk at time t_i+0 ; that is, whose failure or censoring time is at least t_i, $i=0,1,\ldots,k-1$.

For $0 \leq i \leq k$ define

$$p_i=\int_{t_i}^{t_{i+1}} dF(t), \quad q_i=\frac{p_i}{(1-p_0-p_1-\cdots-p_{i-1})} \qquad \ldots \quad (1.1)$$

with $q_0=p_0$. Clearly, $q_k=1$. Holding $f(t)=p_i/(t_{i+1}-t_i)$ constant over the interval $(t_i, t_{i+1}]$, the survival function can be approximated by

$$\bar{F}(t)=\left\{(1-p_0-\cdots-p_{i-1})-\frac{(t-t_i)}{(t_{i+1}-t_i)}\, p_i\right\}$$

$$=\left\{1-\frac{(t-t_i)}{(t_{i+1}-t_i)}\, q_i\right\}\prod_{j=0}^{i-1}(1-q_j),\ t_i < t \leq t_{i+1}, \qquad \ldots \quad (1.2)$$

and the corresponding failure rate is

$$\lambda(t)=\frac{q_i}{\{(t_{i+1}-t_i)-(t-t_i)\,q_i\}},\ t_i < t \leq t_{i+1}, \quad i=0,1,\ldots,k. \qquad \ldots \quad (1.3)$$

Our problem is to estimate the survival function $\bar{F}(t)$ and the failure rate $\lambda(t)$ defined in (1.2) and (1.3) respectively using the data on $\mathbf{n}=\{n_i\}_{i=0}^{k-1}$ and $\mathbf{m}=\{m_i\}_{i=0}^{k-1}$. We will denote this data by $\mathbf{d}$.

2. Bayes estimation of survival function and failure rate

A Bayesian considers the parameter $\mathbf{q}=\{q_i\}_{i=0}^{k-1}$ as a random vector and represents his/her opinion about $\mathbf{q}$ by a probability distribution, called a prior distribution, or simply, a prior. After observing the failure data $\mathbf{d}=(\mathbf{n},\mathbf{m})$, the Bayesian may change the opinion about $\mathbf{q}$, and this is represented by the distribution of $\mathbf{q}$ given the data $\mathbf{d}$, called the posterior distribution. Let ξ be the prior density function of $\mathbf{q}$. The posterior density function of $\mathbf{q}$ given $\mathbf{d}$, namely $\xi(\mathbf{q}|\mathbf{d})$ is expressed by the relation

$$\xi(\mathbf{q}|\mathbf{d}) \propto \xi(\mathbf{q})L(\mathbf{q}|\mathbf{d}). \qquad \ldots (2.1)$$

Here, $L(\mathbf{q}|\mathbf{d})$ is the likelihood function of $\mathbf{q}$ at the point $\mathbf{d}$, and the proportionality symbol $\propto$ is used to indicate that the left side of (2.1) is equal to the right side of (2.1) divided by the factor $\int L(\mathbf{q}/\mathbf{d})\xi(\mathbf{q})\, d\mathbf{q}$, which does not depend on $\mathbf{q}$. It is clear from the relation (2.1) that the change of opinion about $\mathbf{q}$ after the data are obtained is effected through the likelihood function. A Bayesian regards the likelihood function as the sole reservoir of all the relevant information about the parameter that is contained in the data. This is usually stated as *the likelihood principle.* (See Basu (1975) for more on the likelihood principle.)

One usually employs a prior within a family $\mathcal{J}$ of distributions, which is large enough to accommodate various shades of opinion about the parameter $\mathbf{q}$. Further, if $\xi \in \mathcal{J}$ is a prior for $\mathbf{q}$, then the posterior $\xi(\mathbf{q}|\mathbf{d})$ ought to be in a simple computable form. If $\xi(\mathbf{q}|\mathbf{d}) \in \mathcal{J}$ for all $\xi \in \mathcal{J}$ and all data $\mathbf{d}$, then $\mathcal{J}$ is called a *conjugate family* of priors.

Observing that a prior for $\mathbf{q}=\{q_i\}_{i=0}^{k-1}$ corresponds to a prior for $\mathbf{p}=\{p_i\}_{i=0}^{k}$, we define a *conjugate family* of priors for $\mathbf{p}$ as follows. Let $\alpha=\{\alpha_i\}_{i=0}^{k}$ be a sequence of finite non-negative numbers. We say the random vector $\mathbf{p}$ has a k-dimensional Dirichlet distribution with parameter α and denote it by $\mathbf{p}\sim D(\alpha)$, if the distribution of $(p_0, p_1, \ldots, p_{k-1})$ is $D(\alpha_0, \alpha_1, \ldots, \alpha_{k-1}\,;\, \alpha_k)$ as defined in Wilks (1962), Section 7.7.

Under the prior $D(\alpha)$ for $\mathbf{p}$ the coordinates of random vector $\mathbf{q}$ are independent, with q_i having a Beta distribution with parameters α_i and $\alpha_{(i+1)}$ denoted by Beta $(\alpha_i, \alpha_{(i+1)})$, where $\alpha_{(i)}=\sum_{j=i}^{k} \alpha_j$, $i=0, 1, \ldots,$

$k-1$. The Bayes estimate under prior $D(\alpha)$ (with respect to the squared error loss function) of q_i based on no sample, called the prior Bayes estimate of q_i, is

$$\hat{q}_{i,\alpha}=E_{D(\alpha)}(q_i)=\frac{\alpha_i}{\alpha_{(i)}}, \; i=0, 1, \ldots, k-1. \quad \ldots \quad (2.2)$$

After observing the data **d** the likelihood function of **q** at the point **d** is

$$L(\mathbf{q}|\mathbf{d})=\prod_{i=0}^{k-1} p_i^{n_i}(1-p_0-p_1-\cdots-p_i)^{m_i}$$

$$=\prod_{i=0}^{k-1} q_i^{n_i}(1-q_i)^{n_{(i+1)}+m_{(i)}} \quad \ldots \quad (2.3)$$

Using (2.1) and (2.3) the posterior distribution of **q** given the data **d** is given by the relation

$$\xi(\mathbf{q}|\mathbf{d})\propto\prod_{i=0}^{k-1} q_i^{\alpha_i+n_i-1}(1-q_i)^{\alpha_{(i+1)}+n_{(i+1)}+m_{(i)}-1}. \quad \ldots \quad (2.4)$$

The Bayes estimate of q_i based on the data **d** is

$$q^N_{i,\alpha}=E_{D(\alpha)}(q_i|\mathbf{d})=\frac{\alpha_i+n_i}{\alpha_{(i)}+n_{(i)}+m_{(i)}}$$

$$=\frac{\alpha_i+n_i}{\alpha_{(i)}+N_{(i)}}, \; i=0, 1, \ldots, k-1. \quad \ldots \quad (2.5)$$

Clearly, from (2.2) and (2.5) we have

$$\hat{q}^N_{i,\alpha}=\zeta_i\hat{q}_{i,\alpha}+(1-\zeta_i)\hat{\lambda}_i, \quad \ldots \quad (2.6)$$

where $\zeta_i=\alpha_{(i)}/(\alpha_{(i)}+N_{(i)})$. Thus $q^N_{i,\alpha}$ is a weighted mean of the prior Bayes estimate $\hat{q}_{i,\alpha}$ and the empirical estimate $\hat{\lambda}_i=n_i/N_{(i)}$ with weights ζ_i and $(1-\zeta_i)$ respectively. Substituting $\hat{q}_{i,\alpha}$ and $\hat{q}^N_{i,\alpha}$ in equations (1.2) and (1.3), the Bayes estimates of $\bar{F}$ and λ for no sample size and with the data **d** are

$$\hat{\bar{F}}_\alpha(t)=\left\{1-\frac{(t-t_i)\,\alpha_i}{(t_{i+1}-t_i)\,\alpha^{(i)}}\right\}\prod_{j=0}^{i-1}\left(1-\frac{\alpha_j}{\alpha_{(j)}}\right)$$

$$\hat{\bar{F}}_\alpha(t)=\left\{1-\frac{(t-t_i)\,(\alpha_i+n_i)}{(t_{i+1}-t_i)(\alpha_{(i)}+N_{(i)})}\right\}\prod_{j=0}^{i-1}\left(1-\frac{\alpha_j+n_j}{(\alpha_{(j)}+N_{(j)})}\right), \; t_i<t\leqslant t_{i+1},$$

$$\ldots \quad (2.7)$$

and

$$\hat{\lambda}_\alpha(t)=\frac{\alpha_i}{\{(t_{i+1}-t_i)\alpha_{(i)}-(t-t_i)\alpha_i\}}$$

$$\hat{\lambda}^N_\alpha(t)=\frac{\alpha_i+n_i}{\{(t_{i+1}-t_i)(\alpha_{(i)}+N_{(i)})-(t-t_i)(\alpha_i+n_i)\}},\ t_i<t\leqslant t_{i+1}, \quad \dots (2.8)$$

$i=0, 1,\dots, k-1$.

It appears that $\alpha_{(i)}$, the sum of the parameters α_i and $\alpha_{(i+1)}$ of the distribution Beta $(\alpha_i, \alpha_{(i+1)})$ of q_i, enters into the expressions (2.2) and (2.5) as the "prior sample size". This has given rise to the general feeling that allowing $\alpha_{(i)}$ to become small not only makes the "prior sample size" small but also it corresponds to no prior information. By investigating the limit of the Bayes estimate of q_i when $\alpha_{(i)}$ is allowed to converge to zero, we show below that it is misleading to think of $\alpha_{(i)}$ as the prior sample size and the smallness of $\alpha_{(i)}$ as having no information. (This result is discussed in Sethuraman and Tiwari (1981) in the context of Dirichlet measures.)

Consider the convergent sequences of non-negative numbers $\boldsymbol{\alpha}^\gamma=\{\alpha^\gamma_i\}^k_{i=0}$, $\gamma=0, 1, 2,\dots$. If $\alpha^\gamma_j\to\alpha^0_j$ as $\gamma\to\infty$, then $\mathbf{q}^\gamma\xrightarrow{D}\mathbf{q}^0$, where for each γ, $\mathbf{q}^\gamma=\{q^\gamma_i\}$ is as defined in (1.1) and $\xrightarrow{D}$ represents convergence in distribution, and

$$\hat{q}_{i,\alpha^\gamma}\to\hat{q}_{i,\alpha^0},\quad \hat{q}^N_{i,\alpha^\gamma}\to\hat{q}^N_{i,\alpha^0},\quad i=0, 1,\dots, k-1, \quad \dots (2.9)$$

as $\gamma\to\infty$. Further, if we let $\alpha^\gamma_{(i)}$ converge to zero such that $\alpha^\gamma_i/\alpha^\gamma_{(i)}$ converges to a constant θ_i, $0<\theta_i<1$, then in the limit as γ tends to infinity q^γ_i is the Binomial $(1, \theta_i)$ random variable; and

$$\hat{q}^N_{i,\alpha^\gamma}\to\frac{n_i}{N_{(i)}}, \quad \dots (2.10)$$

which is the empirical estimate $\hat{\lambda}_i$, $i=0, 1,\dots, k-1$. Also, $\hat{\bar{F}}^N_{\alpha^\gamma}$ (which is as defined in (2.7) with α replaced by α^γ) in the limit is the empirical estimate

$$\hat{\bar{F}}(t)=\left\{1-\frac{(t-t_i)\;n_i}{(t_{i+1}-t_i)\;N_{(i)}}\right\}\prod_{j=0}^{i-1}\left(1-\frac{n_j}{N_{(j)}}\right),\; t_i<t\leqslant t_{i+1},\; i=0,1,\ldots,k-1. \quad \ldots \quad (2.11)$$

In the remainder of this section, we explore the relationship between the Bayes estimates (2.7) and (2.8) with the well known Kaplan-Meier estimates (Kaplan and Meier (1958)). Instead of starting with a fixed partition $\{(t_i, t_{i+1}]\}_{i=0}^{k}$ of $(0, \infty)$, let $\{t_i\}_{i=1}^{k}$ now denote the distinct observed failure times, again with $t_0=0$ and $t_{k+1}=\infty$. Let n_i denote the number of failures at t_{i+1}, $i=0, 1,\ldots, k-1$, and let m_i, $n_{(i)}$, $m_{(i)}$ and $N_{(i)}$ remain as before. We wish to partition the time-interval using t_i's, $i=1, 2,\ldots, k$ as before. This procedure is clearly justified when T is a discrete random variable and $\{t_i\}_{i=1}^{k}$ is its support so that failures occur only at these points. On the other hand, one may look at the case of continuous failure times as being approximated by a discrete situation like this where, because of observational restrictions, one observes the process at t_{i+1} and makes the approximation that the n_i failures in the interval actually occured at t_{i+1} instead of over the period $(t_i, t_{i+1}]$, $i=0, 1,\ldots, k-1$. This is analogous to the assumptions one makes in computing statistics like the mean and variance from continuous data that has been grouped into class intervals. It should be remarked that the procedure used by Susarla and Van Ryzin (1976) amounts implicity to such a partitioning of the time-interval using the distinct observed failure times and using it to find the censoring numbers (cf. their Section 3). In this case, the survival function and the failure rate are given by (cf. equations (1.2) and (1.3))

$$\bar{F}(t)=\prod_{j=0}^{i}(1-q_j),\quad t_i<t\leqslant t_{i+1}, \quad \ldots \quad (2.1\;)$$

and

$$\lambda(t)=\begin{cases} q_i \text{ at } t=t_{i+1} \\ 0,\; t_i<t\leqslant t_{i+1},\; i=0,1,\ldots,k. \end{cases} \quad \ldots \quad (2.13)$$

Also, the Bayes estimates of $\bar{F}$ and λ are given by (cf. equations (2.7) and (2.8))

$$\hat{\bar{F}}{}^{N}_{\alpha}(t)=\prod_{j=0}^{i}\left(1-\frac{\alpha_j+n_j}{\alpha_{(j)}+N_{(j)}}\right),\quad t_i<t\leqslant t_{i+1}, \quad \ldots \quad (2.14)$$

and

$$\hat{\lambda}_{\alpha}^{N}(t)=\begin{cases}\dfrac{\alpha_i+n_i}{\alpha_{(i)}+N_{(i)}} & \text{at } t=t_{i+1}\\[2ex] 0, & t_i < t < t_{i+1},\ i=0,1,\ \dots\ ,k-1.\end{cases} \quad \dots \quad (2.15)$$

Now an equation similar to (2.6) holds with $\hat{\lambda}_i$ replaced by the Kaplan-Meier estimate $\hat{\lambda}_{KM}(t)=\dfrac{n_i}{N_{(i)}}$ at $t=t_{i+1}$. Again, if $\alpha_{(i)}\to 0$ and $\dfrac{\alpha_i}{\alpha_{(i)}}\to\theta_i$ $0<\theta_i<1$, in the sense discussed above, then the estimates (2.14) and (2.15) converge to the Kaplan-Meier estimates of $\bar{F}$ and λ, namely

$$\bar{F}_{KM}(t)=\prod_{j=0}^{i}\left(1-\frac{n_j}{N_{(j)}}\right),\quad t_i < t \leqslant t_{i+1}, \quad \dots \quad (2.16)$$

and

$$\hat{\lambda}_{KM}(t)=\begin{cases}\dfrac{n_i}{N_{(i)}} & \text{at } t=t_{i+1}\\[2ex] 0, & t_i < t < t_{i+1},\ i=0,1,\ \dots\ ,k-1.\end{cases} \quad \dots \quad (2.17)$$

3. Estimation of $\bar{F}$ and λ when F has increasing (decreasing) failure rate

We say F has an increasing (decreasing) failure rate, i.e., F has IFR (DFR) if $\lambda(t)$ is a monotone increasing (decreasing) function of t. We see from (1.2) that F has IFR (DFR) if the coordinates of $\mathbf{q}$ satisfy the relation $q_{i+1} \geqslant (\leqslant)\, q_i$, for each i. In the Bayesian framework, the prior ξ for $\mathbf{q}$ should incorporate the assumption

$$q_{i+1} \geqslant (\leqslant)\, q_i,\ i=0,\ 1,\ \dots\ ,\ k-2. \quad \dots \quad (3.1)$$

We now consider the cases when F has IFR and DFR separately, since two different priors are appropriate in these two cases. When F has IFR a logical assumption on the prior $D(\alpha)$ for $\mathbf{p}$ is

$$\alpha_{i+1} \geqslant \alpha_i,\ i=0,\ 1,\ \dots\ ,\ k-1. \quad \dots \quad (3.2)$$

Let X and Y be any two random variables. The notation $X \geqslant^{St} Y$ denotes the fact that X is stochastically larger than Y; that is

$$\text{Prob}\,(X \geqslant x) \geqslant \text{Prob}\,(Y > x) \text{ for all real } x.$$

Let $Y_{\alpha,\beta}$ denote a random variable having distribution Beta (α,β). It is well known that $Y_{\alpha',\beta'} \geqslant^{St} Y_{\alpha,\beta}$ if $\alpha' \geqslant \alpha$ and $\beta' \leqslant \beta$. Then under assumption (3.2) we get

$$p_{i+1} \geqslant^{St} p_i, \; i=0, 1, \ldots, k-2. \qquad \ldots \quad (3.3)$$

Clearly, (3.3) implies that

$$q_{i+1} \geqslant^{St} q_i, \; i=0, 1, \ldots, k-2. \qquad \ldots \quad (3.4)$$

Also, under the assumption (3.2) equation (2.2) gives

$$\hat{q}_{i+1,\alpha} \geqslant \hat{q}_{i,\alpha}, \; i=0, 1, \ldots, k-2. \qquad \ldots \quad (3.5)$$

Thus the prior Bayes estimate of λ is increasing in t and, therefore, it follows that the prior Bayes estimate of F has IFR. However, under the assumption (3.2) the Bayes estimates $\{\hat{q}^N_{i,\alpha}\}$ based on the data **d** need not satisfy a relation similar to (3.5) unless the data is such that $n_{i+1} \geqslant n_i$, $i=0, 1, \ldots, k-2$. But the data need not satisfy these conditions. If we have strong reasons to believe that the Bayes estimates $\{\hat{q}^N_{i,\alpha}\}$ should satisfy the relation

$$\hat{q}^N_{i+1,\alpha} \geqslant \hat{q}^N_{i,\alpha}, \; i=0, 1, \ldots, k-2. \qquad \ldots \quad (3.6)$$

then we require

$$n_{i+1} \geqslant n_i, \; i=0, 1, \ldots, k-2. \qquad \ldots \quad (3.7)$$

If a reversal occurs in (3.7), that is, if $n_{j+1} < n_j$ for some j, then we pool the adjacent violators and replace both n_{j+1} and n_j by $\frac{1}{2}(n_j + n_{j+1})$. We now check if the new sequence is properly ordered, i.e., the new sequence of n_i's satisfies the relation (3.7). If a reversal exists, then we replace again by appropriate averages. We continue this procedure until all reversals are eliminated. This procedure of pooling adjacent violators is commonly used in isotonic regression (cf. Proschan and Singapurwalla (1980)).

When F has DFR then $D(\alpha)$ does not serve as a reasonable prior for **p** (or **q**). In this case an appropriate prior for **q** is to consider the q_i's to be independent with q_i having a Beta (a_i, b_i) distribution with $a_i > 0$, $b_i > 0$, $i=0, 1, \dots, k-1$. The prior for **p** is then well-defined and it is called the generalized Dirichlet distribution—of which the Dirichlet distribution is a special case. Under the assumption

$$a_i \geqslant a_{i+1}, \; i=0, 1, \dots, k-2 \qquad \dots \quad (3.8)$$

and

$$b_i \leqslant b_{i+1}, \; i=0, 1, \dots, k-2, \qquad \dots \quad (3.9)$$

we get

$$q_i \geqslant^{st} q_{i+1}, \; i=0, 1, \dots, k-2. \qquad \dots \quad (3.10)$$

The Bayes estimate of q_i (under the generalized Dirichlet prior for **p**) for no sample is

$$\hat{q}_{i,\text{GD}} = \frac{a_i}{a_i+b_i}, \; i=0, 1, \dots, k-1 \qquad \dots \quad (3.11)$$

and based on data **d** is

$$\hat{q}^N_{i,\text{GD}} = \frac{a_i+n_i}{a_i+b_i+N_{(i)}}, \; i=0, 1, \dots, k-1. \qquad \dots \quad (3.12)$$

Clearly under the assumptions (3.8) and (3.9) the prior Bayes estimates $\{q_{i,GD}\}$ satisfy

$$\hat{q}_{i,\text{GD}} \geqslant \hat{q}_{i+1,\text{GD}}, \; i=0, 1, \dots, k-2. \qquad \dots \quad (3.13)$$

showing that the prior Bayes estimate of $\lambda(t)$ and $\bar{F}(t)$ are decreasing with t. Again, as in IFR, under the assumptions (3.8) and (3.9) the Bayes estimates $\{\hat{q}^N_{i,GD}\}$ need not satisfy a relation similar to (3.13). If we strongly believe that $\{\hat{q}^N_{i,\text{GD}}\}$ should satisfy the relation

$$\hat{q}^N_{i,\text{GD}} \geqslant \hat{q}^N_{i+1,\text{GD}}, \; i=0, 1, \dots, k-2. \qquad \dots \quad (3.14)$$

then we require

$$a_i+n_i \geqslant a_{i+1}+n_{i+1}$$

$$a_i+b_i+n_i+m_i \leqslant a_{i+1}+b_{i+1}, \; i=0, 1, \dots, k-2. \qquad \dots \quad (3.15)$$

Again, if a reversal in either of the inequalities in (3.15) occurs then we replace it by pooling adjacent violators (through the method described in the IFR case) until all reversals are eliminated.

4. Accelerated life tests

In accelerated life tests the individuals are tested under different stress environments. Consider l accelerated stress environments $E_1, E_2, \ldots, E_l$ and let E_u denote the usual or use-condition stress. Let $E_i \supset E_j$ denote the fact that E_i is more severe than E_j. We assume the following cases.

Case 1 : $E_u \supset E_l \supset \ldots \supset E_2 \supset E_1$.

Case 2 : $E_1 \supset E_2 \supset \ldots \supset E_l \supset E_u$.

Case 1 is referred to as accelerated life testing with continuously increasing stress starting with stress level E_1, whereas Case 2 is the more usual situation for accelerated life tests. See for instance Mann *et al.* (1974).

Let F_j be the failure time distribution under stress environment E_j. Suppose N_j individuals are observed under stress environment $E_j, j=1, 2, \ldots, l$. Let $n_{i,j}$ be the number of failures and $m_{i,j}$ the number censored in the interval $(t_i, t_{i+1}]$, $i=0, 1, \ldots, k-1$. Let $n_j = \sum_{i=0}^{k-1} n_{i,j}$ and $m_j = \sum_{i=0}^{k-1} m_{i,j}$ so that $N_j = n_j + m_j$. Define $n_{(i),j} = \sum_{r=i}^{k-1} n_{r,j}$, $m_{(i),j} = \sum_{r=i}^{k-1} m_{r,j}$. Then $N_{(i),j} = n_{(i),j} + m_{(i),j}$ is the number of individuals at risk at time t_i+0 under stress environment E_j. Let $\mathbf{n}_j = \{n_{i,j}\}$ and $\mathbf{m}_j = \{m_{i,j}\}$ and denote the data $\{\mathbf{n}_j, \mathbf{m}_j\}$ by $\mathbf{d}_j$. For each j in Case 1 we can obtain the Bayes estimates of the survival function F_j and the failure rate λ_j under the prior $D(\boldsymbol{\alpha}_j)$ for $\mathbf{p}$, where $\boldsymbol{\alpha}_j = \{\alpha_{i,j}\}_{i=0}^{k}$. We assume that this j^{th} sequence is related to a given sequence $\alpha = \{\alpha_0, \alpha_1, \ldots, \alpha_{k+l-1}\}$ by $\alpha_{i,j} = \alpha_{i+j-1}$ and α satisfies

$$\alpha_{i+1} \geqslant \alpha_i, \; i=0, 1, \ldots, k+l-1. \qquad \ldots \quad (4.1)$$

We then have from (3.4) that

$$q_{i+1,j} \geqslant^{St} q_{i,j}, \; i=0, 1, \ldots, k-2 \qquad \ldots \quad (4.2)$$

and

$$q_{i,j+1} \geqslant^{St} q_{i,j},\ j=1, 2,\ldots, l. \qquad \ldots (4.3)$$

Also, under the assumption (4.1) we get

$$\hat{q}_{i+1,j,\alpha} \geqslant \hat{q}_{i,j,\alpha},\ i=0, 1,\ldots, k-2 \qquad \ldots (4.4)$$

and

$$\hat{q}_{i,j+1,\alpha} \geqslant \hat{q}_{i,j,\alpha},\ j=1, 2,\ldots, l. \qquad \ldots (4.5)$$

Thus the prior Bayes estimates of $\bar{F}_j$'s have I$\hat{\text{F}}$R, and from (4.3) and (4.5) we get

$$\lambda_l(t) \geqslant^{St} \lambda_{l-1}(t) \geqslant^{St} \ldots \geqslant^{St} \lambda_1(t) \qquad \ldots (4.6)$$

and

$$\hat{\lambda}_{l,\alpha}(t) \geqslant \hat{\lambda}_{l-1,\alpha}(t) \geqslant \ldots \geqslant \hat{\lambda}_{1,\alpha}(t). \qquad \ldots (4.7)$$

The Bayes estimate of $q_{i,j}$ based on the data $\mathbf{d}_j$ is

$$\hat{q}^N_{i,j,\alpha} = \frac{\alpha_{i,j}+n_{i,j}}{\alpha_{(i),j}+N_{(i),j}},\ i=0, 1,\ldots, k-1,$$

$$j=1, 2,\ldots, l. \qquad \ldots (4.8)$$

Clearly, under the assumptions (4.1) the Bayes estimates $\{\hat{q}^N_{i,j,\alpha}\}$ need not satisfy relations similar to (4.4) and (4.5) and hence $\hat{\lambda}^N_{j,\alpha}(t)$'s need not satisfy relations similar to (4.7). However, if data are such that

$$n_{i+1,j} \geqslant n_{i,j},\ i=0, 1,\ldots, k-2$$

$$n_{i,j+1} \geqslant n_{i,j},\ j=1, 2,\ldots, l, \qquad \ldots (4.9)$$

and

$$n_{i,j+1}+m_{i,j+1} \leqslant n_{i,j}+m_{i,j},\ j=1, 2,\ldots, l, \qquad \ldots (4.10)$$

then the Bayes estimates $\{\hat{q}^N_{i,j,\alpha}\}$ will satisfy

$$\hat{q}^N_{i+1,j,\alpha} \geqslant \hat{q}^N_{i,j,\alpha},\quad i=0, 1, \ldots, k-2$$

$$\hat{q}^N_{i,j+1,\alpha} \geqslant \hat{q}^N_{i,j,\alpha},\quad j=1, 2,\ldots, l. \qquad \ldots (4.11)$$

Also, $\hat{\lambda}^N_{j,\alpha}(t)$'s will satisfy

$$\hat{\lambda}^N_{l,\alpha}(t) \geqslant \hat{\lambda}^N_{l-1,\alpha}(t) \geqslant \cdots \geqslant \hat{\lambda}^N_{1,\alpha}(t). \qquad \ldots \quad (4.12)$$

Case 2 is treated the same way using the generalized Dirichlet prior for p and, therefore, it is omitted here. (Also, see Proschan and Singapurwalla (1980) for this case.)

The Bayes estimates $\{\hat{q}^N_{i,j,\alpha}\}$ are used to estimate the survival function $\bar{F}_u$ and λ_u under E_u. For this we assume the model (cf. Proschan and Singapurwalla (1980)

$$\hat{q}^N_{i,l,\alpha} = w_0 + w_1 \hat{q}^N_{i,l-1,\alpha} + w_2 \hat{q}^N_{i,l-2,\alpha} + \cdots + w_{l-1} \hat{q}^N_{i,1,\alpha}, \quad i=0,1,\ldots,k-1. \qquad \ldots \quad (4.13)$$

We can use the method of least squares to obtain the estimates $\hat{w}_0, \hat{w}_1, \ldots, \hat{w}_{l-1}$. Using these estimates, an estimate of $q_{i,u}$ under E_u is given by

$$\hat{q}_{i,u,\alpha} = \hat{w}_0 + \hat{w}_1 \hat{q}^N_{i,l,\alpha} + \hat{w}_2 \hat{q}^N_{i,l-1,\alpha} + \cdots + \hat{w}_{l-1} \hat{q}^N_{i,2,\alpha}, \quad i=0, 1, \ldots, k-1. \qquad \ldots \quad (4.14)$$

Substituting the $q^*_{i\,u,\alpha}$'s from (4.14) in the expressions (1.2) and (1.3), we get estimates of $\bar{F}_u$ and λ_u.

REFERENCES

Basu, D. (1975) : Statistical information and likelihood. *Sankhyā*. A, 37, 1-71.

Ferguson, T. S. and Phadia, E. G. (1977) : Bayesian nonparametric estimation based on censored data. *Ann. Statist.* 3, 736-740.

Kaplan, E. L. and Meier, P. (1958) : Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* 53, 457-481.

Mann, N. R., Schaffer, R. E. and Singapurwalla, N. D. (1974) : *Methods for Statistical Analysis of Reliability and Life Data.* John Wiley, N. Y.

Proschan, F. and Singapurwalla, N. D. (1980) : A new approach to inference from accelerated life tests. *IEEE Transactions on Reliability.* 29, 2, 98-102.

Sethuraman, J. and Tiwari, Ram C. (1981) : Convergence of Dirichlet measures and the interpretation of their parameter. F. S. U. Tech. Report No 583. (To appear in 3rd Purdue Symposium).

Susarla, V. and Van Ryzin, J. (1976) : Nonparametric Bayesian estimation of survival curves from incomplete observations. *J. Amer. Statist. Assoc.* 71, 897-902.

Wilks, S. S. (1962) : *Mathematical Statistics.* John Wiley, N. Y.