⟨Reseach Paper⟩

Journal of the Korean
Statistcal Society
Vol. 12, No. 1, 1983

A Bayes Reliability Estimation from Life Test in a Stress-Strength Model**

Sung Sub Choi*
Jae Joo Kim*

ABSTRACT

A stress-strength model is formulated for s out of k system of identical components. We consider the estimation of system reliability from survival count data from a Bayesian viewpoint. We assume a quadratic loss and a Dirichlet prior distribution. It is shown that a Bayes sequential procedure can be established. The Bayes estimator is compared with the UMVUE obtained by Bhattacharyya and with an estimator based on Mann-Whitney statistic.

1. Introduction

Suppose a system consisting of k components is successful in its mission if at least $s(1\leq s\leq k)$ of these components survive a random stress. We assume that the component strengths $Y_1,\cdots,Y_k$ are independent with a common cumulative distribution function (cdf) $G\epsilon\mathscr{F}$ where $\mathscr{F}$ is the class of all continuous univariate cdf's. Further, the common stress experienced by each component has cdf $F\epsilon\mathscr{F}$ and is assumed to be independent of the Y's. The system reliability, the probability of s or more of $Y_1,\cdots,Y_k$ exceeding X, is then given by

$$R=R_{s,k}(F,G)=\sum_{i=s}^{k}\binom{k}{i}\int(1-G)^{i}G^{k-i}dF. \tag{1.1}$$

* Department of Computer Science & Statistics, Seoul National University
** This work was supported by the Ministry of Education, through the Research Institute of Basic Sciences, Seoul National University 1982—1983.

While many works for estimation of R have been performed assuming availability of independent measurements of component strengths and external stress, Bhattacharyya (1977) dealt with estimation of R from survival count data when groups of components are tested under common stresses. In each replication of the experiment, a string of components is exposed to a realization of the external stress and, rather than measuring the strengths and the stress, only the number of survivors is recorded.

Interest in this setting stems from the fact that numerical measurements of strengths and stresses often involve use of sophisticated instruments. As such these may be for more expensive than observations of survival and failure which may be obtained simply by visual inspection of the components. Of course, numerical measurements are expected to be more informative than counts. However, it may be economically advantageous to collect count data with a large number of samples than obtaining numerical measurements with fewer components, especially when the cost of numerical measuring is large enough relative to that of the component.

In this paper, we develope a Bayes sequential procedure (BSP) for estimation of R from survival count data under some assumptions. A Bayes fixed sample size estimator is compared with the uniformly minimum variance unbiased estimator(UMVUE) obtained by Bhattacharyya (1977) under the classical mean squared error (MSE) criterion. Simultaneously, we suggest a way of assessing prior informations that is essential to BSP. In addition, efficiencies of our result relative to that based on measurements of strengths and stresses when F and G are related by 'Lehmann alternative', is also evaluated.

2. Statement of the Problem

This section gives a decision theoretic formulation of the problem along with the introduction of some useful notations.

Suppose that the experiment consists of identical replicates in each of which a string of $m(m \geq k)$ components is tested under a common random stress. The component strengths are assumed to be independent and identically distributed (iid) with cdf $G \in \mathscr{F}$ and the stresses in different replications are iid with cdf $F \in \mathscr{F}$. Let Z_j denote the number of survivors in the j-th replication and define

$$N_{in}=\sum_{j=1}^{n} I(Z_j=i) \tag{2.1}$$

=number of replicates with i survivors in n replicates

for $i=0,1,\cdots,m$. Then the joint distribution of $N_n=(N_{0n},N_{1n},\cdots,N_{mn})'$ is a multinomial, $M(n,\boldsymbol{p})$, with n trials and with the vector of cell probabilities $\boldsymbol{p}=(p_0,p_1,\cdots,p_m)'$ given by

$$p_i=\binom{m}{i}\int(1-G)^iG^{m-i}\,dF,\quad i=0,1,\cdots,m. \tag{2.2}$$

Since the observable data N_n is a sufficient statistic for $\boldsymbol{p}$, it is necessary to express R in terms of $\boldsymbol{p}$. Note that

$$\sum_{i=s}^{k}\binom{k}{i}(1-G)^iG^{k-i}=\sum_{i=s}^{k}\sum_{j=0}^{m-k}\binom{k}{i}\binom{m-k}{j}(1-G)^{i+j}G^{m-i-j}$$

$$=\sum_{i=s}^{m}\left[\binom{m}{k}^{-1}\sum_{j=s}^{k}\binom{i}{j}\binom{m-i}{k-j}\right]\binom{m}{i}(1-G)^iG^{m-i} \tag{2.3}$$

where we make the convention that $\binom{a}{b}=0$ if $a<b$ or $b<0$. Integrating both sides of (2.3) with respect to F yields

$$R=\boldsymbol{h}'\boldsymbol{p} \tag{2.4}$$

where $\boldsymbol{h}=(h_0,h_1,\cdots,h_m)'$ and $h_i=\binom{m}{k}^{-1}\sum_{j=s}^{k}\binom{i}{j}\binom{m-i}{k-j}$ for $i=0,1,\cdots,m$.

The prior distribution of $\boldsymbol{p}$ is assumed to be a Dirichlet, $D(\nu)$, given by the probability density function (pdf),

$$g(\boldsymbol{p})=\Gamma(\nu)\prod_{i=0}^{m}[\Gamma(\nu_i)]^{-1}p_i^{\nu_i-1} \tag{2.5}$$

where $\boldsymbol{\nu}=(\nu_0,\nu_1,\cdots,\nu_m)'$, $\nu=\sum_{i=0}^{m}\nu_i$ and $\nu_i>0$ for $i=0,1,\cdots,m$. Then the expected value and the covariances of $\boldsymbol{p}$ are given in Johnson and Kotz (1972) as follows

$$E[\boldsymbol{p}|\boldsymbol{\nu}]=\nu^{-1}\boldsymbol{\nu} \tag{2.6}$$

and

$$\mathrm{Cov}[\boldsymbol{p}|\boldsymbol{\nu}]=\{\nu^2(\nu+1)\}^{-1}C(\boldsymbol{\nu})$$

where a $(m+1)\times(m+1)$ matrix $C(\boldsymbol{b})$, $\boldsymbol{b}=(b_0,b_1,\cdots,b_m)$, is defined by

$$\{C(\boldsymbol{b})\}_{(ii)}=b_i(b-b_i),\quad b=\sum_{i=0}^{m}b_i,$$

$$\{C(\boldsymbol{b})\}_{(ij)}=-b_ib_j \text{ for } i\neq j. \tag{2.7}$$

Note that the posterior distribution of $\boldsymbol{p}$ given N_n is $D(\boldsymbol{\nu}+N_n)$.

The loss function is assumed as follows; Let $c>0$ denote the relative cost of sampling per unit replicate as regards to the squared error of the decision rule d. If n replications have been taken, the loss is given by

$$L(n,d,\boldsymbol{p})=(R-d)^2+nc$$
$$=(\boldsymbol{h}'\boldsymbol{p}-d)^2+nc. \tag{2.8}$$

3. Bayes Sequential Estimation Procedure

Since the BSP consists of a stopping rule and a terminal decision rule, we first consider the latter. It follows from the loss function (2.8) that the Bayes estimator for fixed sample size n is the posterior expected value of R. That is, if we observed $\boldsymbol{N}_n$ after n replications, the terminal decision rule is

$$\hat{R}_n=\boldsymbol{h}'E[\boldsymbol{p}|\boldsymbol{\nu}+\boldsymbol{N}_n]$$
$$=(\nu+n)^{-1}\boldsymbol{h}'(\boldsymbol{\nu}+\boldsymbol{N}_n)$$
$$=(\nu+n)^{-1}\sum_{i=0}^{m}h_i(\nu_i+N_{in}) \tag{3.1}$$

with the posterior expected loss

$$U_n(\boldsymbol{N}_n)=E[L(n,R_n,\boldsymbol{p})|\boldsymbol{\nu}+\boldsymbol{N}_n]$$
$$=\{(\nu+n)^2(\nu+n+1)\}^{-1}\cdot\boldsymbol{h}'[C(\boldsymbol{\nu}+\boldsymbol{N}_n)]\boldsymbol{h}+\mathrm{nc} \tag{3.2}$$
$$=(\nu+n+1)^{-1}\{(\nu+n)^{-1}\sum_{i=0}^{m}h_i^2(\nu_i+N_{in})-\hat{R}^2{}_n\}+nc$$

Note that the Bayes estimator (3.1) coincides with the UMVUE obtained by Bhattacharyya (1977) in the case of vague prior $\boldsymbol{\nu}=\boldsymbol{0}$.

Since the marginal distribution of $\boldsymbol{N}_n$ is a multivariate binomial-beta, $BB(n,\boldsymbol{\nu})$ [Johnson and Kotz (1969)], the conditional probability of $\boldsymbol{N}_n$ given $\boldsymbol{N}_{n-1}$ is

$$P_r[\boldsymbol{N}_n=\boldsymbol{N}_{n-1}+\boldsymbol{e}_i|\boldsymbol{N}_{n-1}]=(\nu_i+N_{i(n-1)})/(\nu+n-1),\quad i=0,1,\cdots,\ m, \tag{3.3}$$

where $\boldsymbol{e}_i$ is the $(m+1)$ dimensional unit vector of which the i-th element is 1. By taking conditional expectation of each element of
$C(\boldsymbol{\nu}+\boldsymbol{N}_n)$ given $\boldsymbol{N}_{n-1}$, we have

$$E[C(\boldsymbol{\nu}+\boldsymbol{N}_n)|\boldsymbol{N}_{n-1}]=(\nu+\mathrm{n}-1)/(\nu+\mathrm{n})\cdot C(\boldsymbol{\nu}+\boldsymbol{N}_n). \tag{3.4}$$

Hence

$$E[U_{n-1}(\boldsymbol{N}_{n-1})-U_n(\boldsymbol{N}_n)|\boldsymbol{N}_{n-1}]=[(\nu+n)(\nu+n-1)]^{-2}\boldsymbol{h}'[C(\boldsymbol{\nu}+\boldsymbol{N}_{n-1})]\boldsymbol{h}-c$$
$$\leq[2(\nu+n)]^{-2}-c\leq 0 \tag{3.5}$$

for all $n\geq J=\langle 1/2\,c^{-1/2}-\nu\rangle$ where $\langle\iota\rangle$ denotes the integer part of ι.

It follows from (2.8) and (3.5) that BSP can be truncated at J and the stopping rule can be completely determined by the backward induction. Define recursively as

$$V_J(N_J)=U_J(N_J)$$
$$\text{and } V_n(N_n)=\min\{U_n(N_n),\ E[V_{n+1}(N_{n+1})\,|\,N_n]\} \tag{3.6}$$
$$=\min\{U_n(N_n),\ {}_n\Sigma^{i=0}(\nu_i+N_{in})/(\nu+n)\cdot V_{n+1}(N_n+e_i)\}$$
for $n=J-1,\ J-2,\cdots,0.$

Then, by DeGroot (1970), the following is estabilished.

Theorem 1. The Bayes sequential estimation procedure for the general problem specified by (1.1), (2.5) and (2.8) is given as follows;

(a) Stopping rule: Stop sampling after taking $n(0\leq n\leq J-1)$ replications if and only if

$$U_n(N_n)\leq V_n(N_n)$$

where $U_n(N_n)$ and $V_n(N_n)$ are given by (3.2) and (3.6), respectively. Stop sampling at $n=J$.

(b) Terminal decision rule: After the sampling is terminated with observations N_n, estimate R as shown in (3.1).

Examples of the procedure for the estimation of 2 out of 3 system reliability is illustrated in Table 3.1, 3.2 and 3.3 for $m=3,4$ and 5, respectively. The prior distributions are assumed to be D(0.6, 0.6, ⋯, 0.6) in all cases. The value of n in the table represents the number of replications or the stage of sampling. The set in each cell represents the sample point N_n where the sampling should be continued. The complement of the continuation set represents the sample point where the sampling should be terminated. The expected Bayes risk of the procedure, $V_0(N_0)$, is compared with that of the Bayes fixed sample size procedure (BFSSP). The expected Bayes risk of BFSSP, V_F, is computed by

Table 3.1; Continuation sets (N_0+N_1, N_2+N_3) for $m=3$

n	$c=0.003$	$c=0.0025$
0	All	All
1	All	All
2	All	All
3	All	All
4	(3, 1), (2, 2), (1, 3)	All
5	(3, 2), (2, 3)	(4, 1), (3, 2), (2, 3), (1, 4)
6	None	(4, 2), (3, 3), (2, 4)
7		None
$V_0(N_0)$	0.03862	0.03573
V_F	0.03885	0.03601

$$V_F=\min_n\{E[U_n(N_n)]\}$$
$$=\min\{[(\nu+n^*)(\nu+1)]^{-1}h'[C(\nu)]h+n^*c,\ [(\nu+n^*+1)(\nu+1)]^{-1}\cdot h'[C(\nu)]h+(n^*+1)c\} \quad (3.7)$$

where $n^*=<\{[c(\nu+1)]^{-1}h'[C(\nu)]h\}^{1/2}-\nu>$.

Table 3.2; Continuation sets $(N_0+N_1,\ N_2, N_3+N_4)$ for $m=4$

n	$c=0.003$	$c=0.0025$
0	All	All
1	All	All
2	All	All
3	(3, 0, 0), (2, 1, 0), (2, 0, 1), (1, 1, 1) (1, 0, 2), (0, 1, 2), (0, 0, 3)	(3, 0, 0), (2, 1, 0), (2, 0, 1), (1, 2, 0) (1, 1, 1), (1, 0, 2), (0, 1, 2), (0, 0, 3)
4	(3, 0, 1), (2, 0, 2), (1, 0, 3)	(3, 0, 1), (2, 1, 1), (2, 0, 2), (1, 1, 2) (1, 0, 3)
5	None	(3, 0, 2), (2, 0, 3)
6		None
$V_0(N_0)$	0.03324	0.03100
V_F	0.03343	0.03125

Table 3.3; Continuation sets $(N_0+N_1, N_2, N_3, N_4+N_5)$ for $m=5$

n	$c=0.003$	$c=0.00025$
0	All	All
1	All	All
2	(2, 0, 0, 0), (1, 1, 0, 0), (1, 0, 1, 0) (1, 0, 0, 1), (0, 1, 1, 0), (0, 1, 0, 1) (0, 0, 1, 1), (0, 0, 0, 2)	All
3	(2, 0, 0, 1), (1, 1, 0, 1), (1, 0, 1, 1) (1, 0, 0, 2)	(3, 0, 0, 0), (2, 1, 0, 0), (2, 0, 1, 0) (2, 0, 0, 1), (1, 1, 0, 1), (1, 0, 2, 0) (1, 0, 1, 1), (1, 0, 0, 2), (0, 2, 0, 1) (0, 1, 0, 2), (0, 0, 1, 2), (0, 0, 0, 3)
4	None	(3, 0, 0, 1), (2, 0, 1, 1), (2, 0, 0, 2) (1, 1, 0, 2), (1, 0, 0, 3)
5		None
$V_0(N_0)$	0.03018	0.02840
V_F	0.03034	0.02853

4. Comparative Studies for Moderate Samples

In this section, the Bayes estimator (3.1) is compared with Bhattacharyya's UMVUE based on survival counts and with the Mann-Whitney statistic based on numerical measurements under the classical MSE criterion.

If N_n is observed after (fixed) n samples are obtained, the UMVUE is given in Bhattacharyya(1977) as

$$\tilde{R}_n = n^{-1} \cdot h' N_n \tag{4.1}$$

where h is the same as that defined in (2.4). And the MSE of (4.1) is

$$\begin{aligned} \mathrm{MSE}(\tilde{R}_n) &= \mathrm{Var}[\tilde{R}_n | p] \\ &= n^{-1} \cdot h' [C(p)] h \\ &= n^{-1} \left\{ \sum_{i=0}^{m} h_i^2 p_i - \left(\sum_{i=0}^{m} h_i p_i \right)^2 \right\} \end{aligned} \tag{4.2}$$

where $C(p)$ is defined as (2.7).

The MSE of the estimator (3.1) given p is represented as

$$\begin{aligned} \mathrm{MSE}(\hat{R}_n) &= \mathrm{Var}[\hat{R}_n | p] + (E[\hat{R}_n | p] - R)^2, \quad R = h'p \\ &= (\nu+n)^{-2} [n h' [C(p)] h + \nu^2 (R_0 - R)^2\}, \quad R_0 = (h'\nu)/\nu. \end{aligned} \tag{4.3}$$

Then the relative efficiency of the Bayes estimator to the UMVUE is given by

$$\begin{aligned} e(\hat{R}_n;\ \tilde{R}_n) &= \mathrm{MSE}(\tilde{R}_n) / \mathrm{MSE}(\hat{R}_n) \\ &= [(n+\nu)/n]^2 (1 + \nu^2 \rho / n)^{-1} \end{aligned} \tag{4.4}$$

where $\rho = (R_0 - R)^2 / h' [C(p)] h$.

Note that $e(\hat{R}_n;\ \tilde{R}_n)$ is determined by ν and ρ, for given n and p and that the smaller value of ρ (*i.e.*, the more accurate value of R_0) gives the higher efficiency. In fact, if $\rho \leq n^{-1}$, then $e(\hat{R}_n;\ \tilde{R}_n)$ is always greater than 1 irrespective of ν. It is also noticeable that, when ρ is fixed, the efficiency is greater than 1 for $0 < \nu < 2\rho^{-1}$ and that $e(\hat{R}_n;\ \tilde{R}_n)$ is maximized at $\nu = \rho^{-1}$ with maximum value $1 + (n\rho)^{-1}$.

Hence the following suggestion may be possible to determine the parameter vector $\boldsymbol{\nu}$ of a prior distribution; Suppose that we have some prior information about the structure of R, that is, we have some prior feeling that a $(m+1)$ vector $f = (f_0, f_1, \cdots, f_m)'$, $f_i > 0$ for all i and $\sum_{i=0}^{m} f_i = 1$, would be probable for p and the variance of R would be $\sigma > 0$. Then, if we determine $\boldsymbol{\nu} = (\nu_0, \nu_1, \cdots, \nu_m)'$ as

$$\nu = h' [C(f)] h / \sigma,$$

and $$\nu_i = \nu f_i, \quad i = 0, 1, \cdots, m, \tag{4.5}$$

the relative efficiency $e(\hat{R}_n;\ \tilde{R}_n)$ in (4.4) could be expected to be close to its maximum value unless f and σ are much different from their true values.

Relative efficiencies for 1 out of 2 system with $m=5$ are listed in Table (4.1). Prior probability guesses $f_i=\binom{5}{i}(0.5)^5$, $i=0, 1, \cdots, 5$ with prior guesses $\sigma=0.005$, 0.01, 0.05, 0.1 are applied to real probabilities $p_i=\binom{5}{i}\theta^i(1-\theta)^{5-i}$, $i=0\ 1, \cdots, 5$ where θ is varied from 0.3 to 0.7 for moderate sample sizes $n=20$. It shows that the efficiencies are relatively good in general except in considerably 1) pessimistic cases. Hence the Bayes estimator (3.1) could be expected to be better than UMVUE (4.1) by considerate choice of f and 2) conservative evaluation of σ.

Table 4.1 : Efficiencies of Bayes estimator relative to UMVUE

θ \ σ	0.005	0.01	0.05	0.1
0.3	0.469	0.800	1.071	1.047
0.4	1.207	1.307	1.104	1.054
0.45	1.932	1.541	1.113	1.056
0.5	2.44	1.642	1.116	1.057
0.55	1.864	1.524	1.112	1.056
0.6	1.015	1.215	1.100	1.053
0.7	0.262	0.534	1.030	1.035

Another interesting comparison results from a consideration of the UMVUE of R for a single component $(s=k=1)$ system under the sampling of numerical measurement. Consider a sampling scheme such that; A random sample of n stress measurements are observed in addition to observing an independent random sample of mn strength measurements. With the general nonparametric model $F\in\mathcal{F}$, $G\in\mathcal{F}$, let $(X_1, \cdots, X_m)$ and $(Y_1, \cdots, Y_{mn})$ denote the stress and the strength measurements respectively. As noted in Birnbaum (1956), the UMVUE of R is given by $R_n{}^*=W/(mn^2)$ where $W=\sum_{i=1}^{n}\sum_{j=1}^{mn} I(Y_j>X_i)$ is the Mann-Whitney statistic.

Considering once again the Lehmann subfamily $\bar{F}=\bar{G}^\lambda$, we have

$$\begin{aligned} \mathrm{MSE}(R_n{}^*) &= \mathrm{Var}[R_n{}^* \mid \lambda] \\ &= [\lambda/(\lambda+1)+(mn-1)\lambda/(\lambda+2)+(n-1)\{1-2/(\lambda+1)+1/(2\lambda+1)\} \\ &\quad -(n+mn-1)\lambda^2/(\lambda+1)^2]/(mn^2). \end{aligned} \tag{4.6}$$

On the other hand, MSE of the Bayes estimator from survival counts with the same number of samples under $\bar{F}=\bar{G}^\lambda$ is given by

$$\mathrm{MSE}^*(\hat{R}_n)=(\nu+n)^{-2}\{n\lambda(\lambda+m+1)/[m(\lambda+1)^2(\lambda+2)]+\nu^2(R_0-\lambda/(\lambda+1))^2\} \tag{4.7}$$

1) the cases in which we give too much confidence to inaccurate f.

2) It is safer to take a moderate value of σ than to assess the value of σ too small.

where R_0 is defined as (4.3) with $h_i=i/m$, $i=0, 1, \cdots, m$. Then the efficiency of Bayes estimator based on counts relative to UMVUE based on measurements for the case $s=k=1$ and $\bar{F}=\bar{G}^{\lambda}$ is given by

$$e^*(\hat{R}_n\ ;\ R_n^*)=\mathrm{MSE}(R_n^*)/\mathrm{MSE}^*(\hat{R}_n) \qquad (4.8)$$

Table (4.2) provides numerical values of $e^*(\hat{R}_n\ ;\ R_n^*)$ for $m=5$. Even prior guesses $f_i=1/6$, $i=0, 1, \cdots, 5$ with prior variance guesses $\sigma=0.005, 0.01, 0.05, 0.1$ are applied to the cases $\lambda=1.0, 1.2, 1.4, 1.6, 1.8$ for moderate sample size $n=20$. ν is calculated as (4.5). The efficiencies of $\tilde{R}$ relative to R_n^* are also listed for comparison.

Table 4.2 : Efficiencies of $\hat{R}_n$ relative to R_n^* for the case $s=k=1$ and $\bar{F}=\bar{G}^{\lambda}$

λ / σ	0.005	0.01	0.05	0.1	$e^*(\tilde{R}_n;R_n^*)$
1.0	4.057	2.167	1.078	0.968	0.864
1.2	2.682	1.912	1.065	0.961	0.859
1.4	1.439	1.478	1.044	0.950	0.852
1.6	0.861	1.114	1.017	0.938	0.845
1.8	0.572	0.815	0.990	0.928	0.840

※ $e(\tilde{R}_n\ ;\ R_n^*)=\mathrm{MSE}(R_n^*)/\mathrm{MSE}^*(\hat{R}_n)$ is the efficiency of $\tilde{R}_n$ relative to R_n^* where $\mathrm{MSE}^*(\tilde{R}_n)=n^{-1}\lambda(\lambda+m-1)/[m(\lambda+1)^2(\lambda+2)]$.

Acknowledgment

The authors wish to thank the referee for his careful reading of the manuscript and for his valuable comments.

REFERENCES

(1) Bhattacharyya, G.K. (1977). Reliability Estimation from Survivor Count Data in a Stress Strength Setting, *Journal of the Indian Assoc. for Productivity, Quality and Reliability*, 2, 1—16.

(2) Birnbaum, Z.W. (1956). On a Use of the Mann-Whitney Statistic, *Proceeding of the third Berkeley Symposium on Math. Statistics and Prob.*, 1, 13—17.

(3) DeGroot, M.H. (1970). *Optimal Statistical Decisions*, McGrow-Hill, Inc., New York.

(4) Johnson, N.L. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*, John Wiley & Sons, Inc., New York.

(5) Johnson, N.L. and Kotz, S. (1969). *Distribution in Statistics: Discrete Distributions*, Houghton Mifflin Company, Boston.