<**Research Paper**>

Journal of the Korean
Statistical Society
Vol. 11, No. 2, 1982

Optimal Life Testing Procedure for a System with Exponentially Distributed Failure Times

Sang Un Yun*

ABSTRACT

The choice of constants that define a life testing procedure is considered in terms of the test termination time (censoring time) and the number of items to be tested subject to a given range of variance of the expected life time, where the failure time of life testing is exponentially distributed.

1. Introduction

The life-testing problem has received a lot of attention in statistical literature. Generally the life-time distribution is assumed to be exponential and a censoring procedure is used to estimate the parameters of this distribution. Blight (1972) has studied the parameter estimation problem when the testing facility and the total replacement are limited. He obtains the optimal numbers of items to start the test and to be replaced during the test at a pre-determined total number of failures. In this paper we assume that the total testing cost depends linearly on the number of tested items n and the time period of the testing operation, T or the censoring time. The accuracy of an estimator is measured by its variance. The optimal values of n and T for a required accuracy are obtained by using the exact variance, the Cramer-Rao lower bound, and Boardman and Kendall's approximation (1970), which is the same as Mendenhall and Lehman's (1960). A practical example is used to demonstrate the cost saving of these optimal results.

* Korea Institute for Defense Analyses

The optimal procedure to estimate the survival function is studied in Section 4. The important extention for the grouped observations is considered in Section 5.

2. Optimal Number of Testing Items and Censoring Time

Let the life-time x of an item have an exponential distribution with density function

$$f(x)=\frac{1}{\theta}e^{-x/\theta},\ x>0,\ \theta>0,$$

where θ is an unknown parameter. It is well known (e.g. Moeshberger and David (1971)), that the maximum likelihood estimator (MLE) $\hat{\theta}$ of θ is

$$\hat{\theta}=\frac{\sum_{j=1}^{r} x_j+(n-r)T}{r}, \tag{1}$$

where n is the total number of testing items, r is the number of items failed before the censoring time T, and x_j is the failure time of the jth item which failed before T.

The exact mean and variance of $\hat{\theta}$ is not easy to obtain. Boardman and Kendall (1970), and Mendenhall and Lehman (1960) have shown that

$$E(\hat{\theta})=\theta-T/(1-e^{-T/\theta})+nTE(1/r) \tag{2}$$

and

$$\mathrm{Var}(\hat{\theta})=\left\{\theta^2-\frac{T^2e^{-T/\theta}}{(1-e^{-T/\theta})^2}\right\}E\left[\frac{1}{r}\right]+n^2T^2E\left[\frac{1}{r^2}\right]-n^2T^2E^2\left[\frac{1}{r}\right] \tag{3}$$

where r has a binomial distribution with parameters n and $p=1-e^{-T/\theta}$ and $E\left[\frac{1}{r^i}\right]$ is conditioned on $r>0$.

Exact values of the expectations $E\left[\frac{1}{r}\right]$ and $E\left[\frac{1}{r^2}\right]$ can only be obtained numerically for small n. Boardman and Kendall (1970) modified Johnson's (1960) result to obtain an approximate to them, i.e.,

$$E\left[\frac{1}{r}\right]=\frac{n-2}{n\{(n-1)p-1\}} \tag{4}$$

and

$$E\left[\frac{1}{r^2}\right]=\frac{(n-2)(n-3)}{n^2\{(n-1)p-1\}\{(n-1)p-2\}} \tag{5}$$

where $p=1-e^{-\frac{T}{\theta}}$. Mendenhall and Lehman (1960) also obtained the same result by using a Beta approximation to a binomial distribution. Figure 1 compares the differences of the variance expressions, such as CRLB, Boardman and Kendall's (1970) approximation (B) and the exact variance (E) when $\theta=10.0$ and $n=10$ (10) 50, 70, 100 and $T=4.0$ (2) 12.0, 15.0, 20.0, 50.0, 100.0. Figure 1 shows that the variance

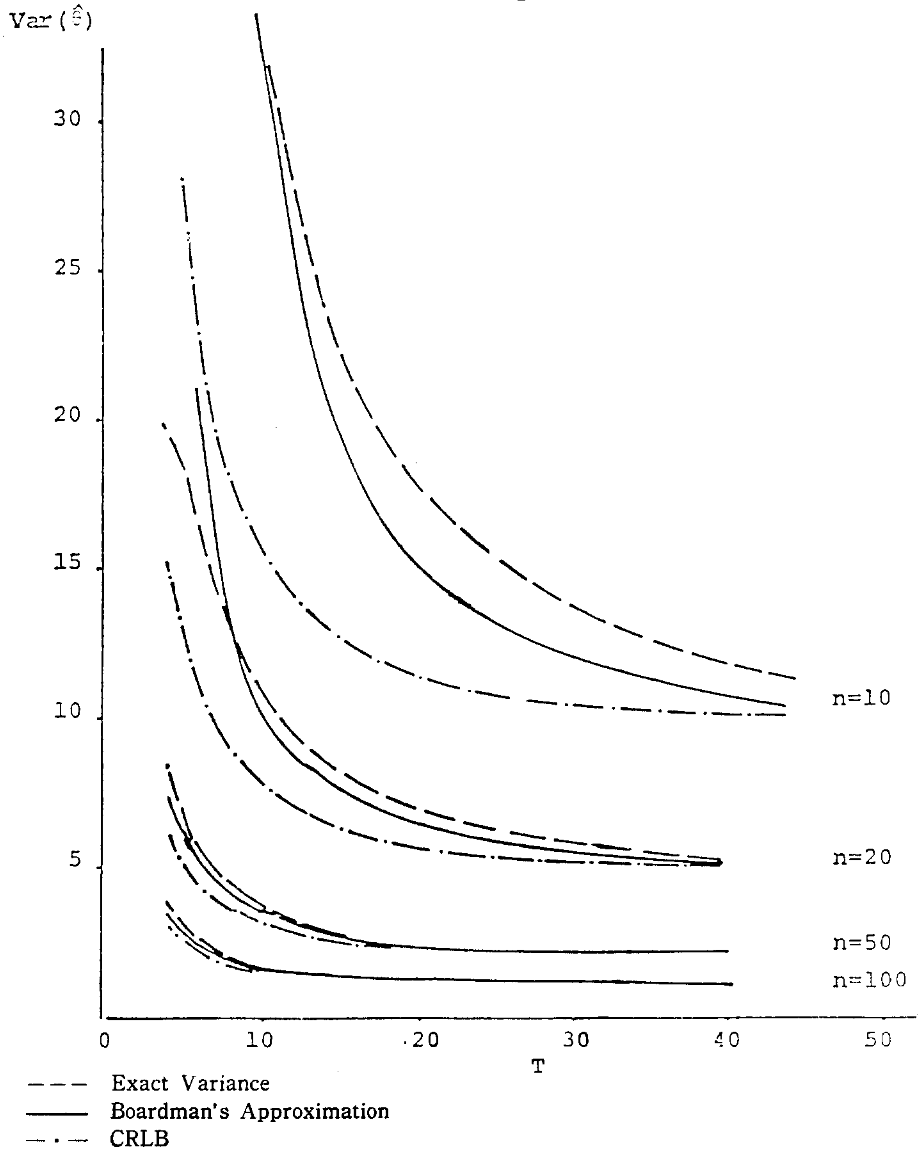

Figure 1. Comparison of the approximate variances and the exact variances for $\hat{E}(x)=\hat{\theta}$ when $\theta=10$.

is more sensitive to the changes in number of testing items than that in censoring time. And the CRLB seems to be nearly attained when n and (or) T are reasonably large.

As one can see from the previous formula, the test procedure is completely determined by the choice of n and T. If we let C_t be the cost of one unit test running time and C_i be the cost of one test item, then the optimal test for a given variance of $\hat{\theta}$ is to minimize

$$C=C_tT+C_in$$

subject to

$$\text{Var}\ (\hat{\theta})=V, \tag{6}$$

where V is a given value. The optimal n and T can be found by iteration method(e.g. Newton's method). The initial values of n and T can be found by letting Var($\hat{\theta}$) equal to the CRLB of (1), i.e.,

$$V=\frac{\theta^2}{n\left[1-e^{-\frac{T}{\theta}}\right]} \tag{7}$$

It can be shown by the method of Lagrange multipliers that the optimal solution for (6) and (7) is

$$\hat{n}=\frac{C_ik+\sqrt{C_i^2k^2+4C_tC_i\theta_k}}{2C_i}$$
$$\hat{T}=\theta\left\{\ln\left[C_t+\frac{C_in}{\theta}\right]-\ln C_t\right\}, \tag{8}$$

where

$$k=\frac{\theta^2}{V}.$$

In practice θ is unknown. We may use the previous knowledge of θ to solve for n and T. Table 1 lists various optimal T and n by the criteria of the CRLB, Boardman and Kendall's (1970) approximation (B) for the variance, and the exact variance (E). These values are recorded for five different ratios of cost, C_t and C_i; (9 : 1), (7 : 3), (5 : 5), (3 : 7), (1 : 9), and for the variance of the estimation, V in (6), equal to 1.0, 2.0(2) 16.0.

Since the CRLB is nearly attained for large n and (or) T, the solution using (8) appears to be quite good even with a relatively large variance limit, V, in (6).

3. An Example

Mendenhall and Hader (1958) analyzed some data on the failure times of radio transmitter receivers; the failures were classified into two types: those confirmed on arrival at the maintenance center, and those unconfirmed. Since they assumed that an item is preordained to fall by only one cause according to a binomial mixture and an associated conditional failure density function, Boardman and Kendall (1970) considered other

Table 1. Comparison of optimal solutions to estimate the life time by various criteria such that CRLB, Boardman's approximation (B), and exact variance (E).

	$C_t=1.0$, $C_i=9.0$		$C_t=3.0$, $C_i=7.0$		$C_t=5.0$, $C_i=5.0$		$C_t=7.0$, $C_i=3.0$		$C_t=9.0$, $C_i=1.0$	
VAR	T	n	T	n	T	n	T	n	T	n
1.0 CRLB	45.21	101	32.29	104	24.77	109	18.08	119	10.10	157
B	45.50	102	34.46	105	25.62	111	18.63	122	10.22	162
E	52.16	102	37.15	106	25.39	112	18.33	123	10.38	161
2.0 CRLB	38.48	51	25.93	53	19.17	58	13,54	67	7.26	96
B	41.41	52	29.43	55	20.23	61	14.40	70	7.60	101
E	47.10	51	27.63	55	20.18	60	14.83	68	7.71	100
4.0 CRLB	31.95	26	20.19	28	14.35	32	9.83	29	5.17	51
B	38.66	27	22.59	31	15.46	36	10.94	43	5.62	67
E	32.96	27	22.15	30	16.28	34	11.04	42	5.83	66
6.0 CRLB	27.91	17	17.35	20	11.94	23	8.08	29	4.20	47
B	35.06	19	19.44	23	13.65	27	9.11	34	4.76	54
E	36.12	18	21.21	21	14.63	25	9.36	33	4.94	54
8.0 CRLB	25.42	13	15.04	15	10.65	19	7.07	24	3.68	40
B	33.24	15	17.49	19	11.92	23	8.15	29	4.24	47
E	33.29	14	19.25	17	13.02	21	8.53	28	4.50	47
10.0 CRLB	23.89	11	13.95	13	9.56	16	6.42	21	3.29	35
B	37.24	12	17.43	16	11.28	20	7.44	26	3.83	43
E	28.35	12	17.05	15	12.57	18	7.96	25	4.17	43
12.0 CRLB	22.08	9	12.72	11	8.75	14	5.72	18	2.96	31
B	31.47	11	17.38	14	10.76	18	6.90	24	3.67	39
E	30.36	10	17.05	13	11.21	17	7.54	23	3.96	40
14.0 CRLB	21.04	8	12.04	10	7.88	12	5.22	16	2.79	29
B	29.88	10	16.16	13	9.89	17	6.70	22	3.43	37
E	28.33	9	16.02	12	11.49	15	7.46	21	3.79	38
16.0 CRLB	19.88	7	11.31	9	7.42	11	4.96	15	2.62	27
B	30.68	9	15.84	12	9.45	16	6.31	21	3.30	35
E	28.91	8	15.82	11	11.20	14	7.18	20	3.71	36

point estimates of the parameters by assuming that an item on test can fail by either one of the other two subsystems. They obtained the point estimates of two parameters, θ_1 and θ_2 as $\hat{\theta}_1=450.61$, $\hat{\theta}_2=918.07$, where $n=369$ and $T=630$.

Since θ_1 and θ_2 are the parameters from two independent exponential distributions one can combine them into one exponential distribution with parameter $\theta=\theta_1\theta_2/(\theta_1+\theta_2)$. Accordingly,

$$\hat{E}(t)=\theta=\frac{\hat{\theta}_1\hat{\theta}_2}{\hat{\theta}_1+\hat{\theta}_2}=302.3$$

One can find the exact variance (E) and the approximate variance (B) by Boardman and Kendell (1970) and the CRLB when $n=369$ and $T=630$, which are shown to be

$$\begin{aligned} \mathrm{Var}(E)&=312.19 \\ \mathrm{Var}(B)&=285.29 \\ \mathrm{CRLB}\ &=282.75 \end{aligned} \tag{9}$$

Table 2. Optimal solution for n and T in Mendenhall and Hader's example with different ratios of costs, C_i and C_t.

	(C_t,C_i)	(1 : 9)			(3 : 7)			(5 : 5)			(7 : 3)			(9 : 1)		
	VAR	T	n	C	T	n	C	T	n	C	T	n	C	T	n	C
CRLB	282.75	738.6	353	3915.6	437.9	422	4267.8	299.9	513	4064.6	200.8	665	3400.3	103.7	1113	2046.0
B	285.29	735.8	354	3921.8	440.5	422	4275.6	298.9	516	4074.7	201.2	667	3409.4	104.4	1110	2049.9
M		630.0	369	3951.0	630.0	369	4473.0	630.0	369	4464.5	630.0	369	5517.0	630.0	369	6039.0

Table 2 summarizes the optimal solutions for different ratios of the costs, C_t and C_i under the restrictions that they have the same variances as (9). In Table 2, each row for the CRLB, B, shows the optimal T and n, and the total cost, respectively for different ratios of the costs C_t and C_i; (1 : 9), (3 : 7), (5 : 5), (7 : 3), (9 : 1), with respect to each criteria as constraints. The last row (M) shows the total cost comparison for the case when the test is conducted with the T and n in Mendenhall and Hader's (1958) paper. The exact variance E is not computed since it requires a large amount of computing time.

As one can see from Table 2, there are large cost differences between the optimal design and the design in Mendenhall and Hader's (1958) example when the cost of unit testing time, C_t, is more than the cost of unit item, C_i.

4. Optimal Procedure to Estimate the Survival Function

In this section, optimal n and T are derived to estimate the survival function. The survival function for the exponential distribution is

$$\begin{aligned} H(t_0)&=p_r(X>t_0)=1-F(t_0) \\ &=e^{-\frac{t_0}{\theta}} \end{aligned}$$

where t_0 is a given time and X is a random variable of survival time.

From (1), maximum likelihood estimator of $H(t_0)$,

$$\hat{H}(t_0)=\exp\left[-\frac{rt_0}{\sum_{j=1}^{r} x_j+(n-r)T}\right] \tag{9}$$

Now, employing the notation of Section 4, we obtain the expectation of quantities leading to the CRLB. Since the joint density function of ordered failure time, $x_1, \cdots, x_r$ and the number of failures, r, given censoring time, T, and the number of items to be tested, n, is

$$f(x_1, \cdots, x_r, r|n, x_j \leq T, \forall_j)=\left[\begin{matrix} n \\ r \end{matrix}\right]\left[\prod_{j=1}^{r}\frac{1}{\theta}e^{-\frac{x_j}{\theta}}\right]\left[e^{-\frac{T}{\theta}}\right]^{n-r}$$

$$=\left[\begin{matrix} n \\ r \end{matrix}\right]\left[\prod_{j=1}^{r}\left(-\frac{\ln H_0}{t_0}\, e^{\frac{x_j \ln H_0}{t_0}}\right)\right]\left(e^{\frac{T\ln H_0}{t_0}}\right]^{n-r}$$

where, as in Section 4,

$$H_0=H(t_0)=e^{-\frac{t_0}{\theta}}$$

and the logarithm of the likelihood function is

$$\ln L \propto \sum_{j=1}^{r}\left\{\ln(-\ln H_0)-\ln t_0+\frac{\ln H_0}{t_0}x_j\right\}+(n-r)\frac{\ln H_0}{t_0}\ T.$$

Now, the second derivative is found to be

$$\frac{\partial^2 \ln L}{\partial H_0^2}= -r\left\{-\frac{1}{H_0^2\ln H_0}-\frac{1}{H_0^2(\ln H_0)^2}\right\}+\sum_{j=1}^{r}\left(-\frac{x_j}{t_0H_0^2}\right)-\frac{(n-r)T}{t_0H_0^2}$$

$$=\frac{r}{H_0^2(\ln H_0)^2}(\ln H_0+1)-\frac{1}{t_0H_0^2}\sum_{j=1}^{r}x_j-\frac{T}{t_0H_0^2}(n-r)$$

Furthermore, the facts that

$$E[H(r, x)]=E(E[H(r, x)\,|\,r])$$

and that the density of r is

$$g(r)=\binom{n}{r}\left(1-e^{-\frac{T}{\theta}}\right)^{r}\left(e^{-\frac{T}{\theta}}\right)^{n-r}$$

lead to

$$E\left[\frac{\partial^2 \ln L}{\partial H_0^2}\right]=\frac{n\left(1-e^{\frac{T\ln H_0}{t_0}}\right)(2\ln H_0+1)}{H_0^2(\ln H_0)^2}.$$

Thus,

$$CRLB(H_0)=-\frac{H_0^2\ln^2 H_0}{n\left(1-e^{\frac{T\ln H_0}{t_0}}\right)(2\ln H_0+1)} \tag{10}$$

Now the optimal test for a given accuracy, as in (6) is to minimize

Table 3. Optimal procedures to estimate the survival function for the various combinations of t_0's and cost ratios when the prior estimate of θ is 10.

		$C_t=1$, $C_i=100$		$C_t=1$, $C_i=9$		$C_t=3$, $C_i=7$		$C_t=5$, $C_i=5$		$C_t=7$, $C_i=3$		$C_t=9$, $C_i=1$		$C_t=100$, $C_i=1$	
t_0	CRLB	T	n	T	n	T	n	T	n	T	n	T	n	T	n
6.1	0.0025	70.27	199	51.99	200	38.83	203	30.88	209	23.48	220	13.78	267	4.43	557
	0.0100	62.16	50	38.48	51	26.08	53	19.24	58	13.57	67	7.29	96	2.23	249
8.0	0.0025	67.08	86	43.75	87	30.94	90	23.53	95	17.07	105	9.43	141	2.93	339
	0.0100	53.82	21	30.60	22	19.29	25	13.60	28	9.27	35	4.84	56	1.47	157
10.0	0.0025	62.90	54	39.26	55	26.78	58	19.85	62	14.05	71	7.57	101	2.32	261
	0.0100	49.22	13	26.47	14	16.01	16	11.06	20	7.44	25	3.85	42	1.16	123
12.0	0.0025	59.33	37	35.71	38	23.62	41	17.14	45	11.93	53	6.33	79	1.93	212
	0.0100	45.57	9	23.32	10	13.67	12	9.32	15	6.22	20	3.21	34	0.97	101
15.0	0.0025	54.22	22	30.96	23	19.58	26	13.84	29	9.44	36	4.94	57	1.50	161
	0.0100	40.60	5	19.31	6	10.89	8	7.32	10	4.85	14	2.49	25	0.75	77
18.0	0.0025	49.30	13	26.52	14	16.05	17	11.09	20	7.47	25	3.87	42	1.17	123
	0.0100	35.84	3	15.81	4	8.64	5	5.76	7	3.80	10	1.94	19	0.58	60
20.0	0.0025	46.03	9	23.70	10	13.94	12	9.52	15	6.36	20	3.28	34	0.99	103
	0.0100	32.73	2	13.72	3	7.38	4	4.89	6	3.22	8	1.65	16	0.49	50
25.0	0.0025	37.86	4	17.25	5	9.55	6	6.38	8	4.22	12	2.16	21	0.65	67
	0.0100	25.22	1	9.39	1	4.91	2	3.23	3	2.12	5	1.08	10	0.32	32
30.0	0.0025	29.86	1	11.95	2	6.35	3	4.19	5	2.76	7	1.41	13	0.42	43
	0.0100	18.41	0	6.24	0	3.21	1	2.11	2	1.38	3	0.70	6	0.21	21
50.0	0.0025	6.96	0	2.13	0	1.08	0	0.71	0	0.47	1	0.24	2	0.07	7
	0.0100	3.53	0	1.07	0	0.54	0	0.36	0	0.23	0	0.12	1	0.04	3

$$C = C_t T + C_i n$$

subject to

$$CRLB(H(t_0)) = V \tag{11}$$

where V is a given value. Again, by Lagrange multipliers, the optimal solution for (10) and (11) is

$$\hat{T} = -\theta \ln\left[\frac{(2a-b)-\sqrt{b(b-4a)}}{2a}\right]$$
$$\hat{n} = -\frac{t_0^2 e^{-\frac{2t_0}{\theta}}}{\theta^2 v\left(1-e^{-\frac{T}{\theta}}\right)\left(-\frac{2t}{\theta}+1\right)} \tag{12}$$

where

$$a = C_1 \theta^3 V\left(-\frac{2t_0}{\theta}+1\right)$$

and

$$b=C_2t_0^2e^{\frac{2t_0}{\theta}}.$$

In practice θ is unknown. But we may use the previous knowledge of θ to solve for n and T. Table 3 shows some examples of optimal procedure to estimate the Survival Function for the various cases such that t_0=6.1, 8.0, 10.0, 12.0, 15.0, 18.0, 20.0, 25.0, 30.0, 50.0, and $(C_i\,;\,C_t)$=(100 : 1), (9 : 1), (7 : 3), (5 : 5), (3 : 7), (1 : 9), (1 : 100), when θ is assumed to be 10. It is quite intuitive that when t_0 is much larger than the expected survival time θ, we do not need too many items for testing.

5. Optimal Procedure Based on Grouped Observations

Another important extension of the general method is to group the observations into intervals. Suppose that the range of variation of the lifetime is partitioned into h time intervals of length g such that $T=hg$. Let r_i, $i=1,\cdots,h$, denote the number of individuals failing in the interval $((i-1)g,\ ig)$. Then $r=\sum_{i=1}^{h}r_i$ is the total number of failures in the test.From Moeschberger and David (1971), the MLE of θ, $\hat{\theta}$, is

$$\hat{\theta}=\frac{rg}{r\ln\left\{1+(r-r_k)/\sum_{i=2}^{h}(i-1)r_i\right\}}$$

and

$$CRLB(\hat{\theta})=\frac{\theta^2}{n}(1+r)$$

where

$$r=\frac{\theta^2\left(e^{-\frac{g}{\theta}}-1\right)^2}{g^2e^{-\frac{g}{\theta}}\left[1-\exp\left\{-\frac{(h-1)g}{\theta}\right\}\right]}-1$$

Here, again, n is the number of items to be tested.

Now, the optimal design is to minimize

$$C=C_tgh+C_in$$

subject to

$$v=CRLB(\hat{\theta})$$

Here, again, C_t and C_i are given as the cost of one unit of testing time and the cost of one item, respectively, and g and v are given.

By Lagrange multipliers,

$$h=-\frac{\theta}{g}\ln E$$

and

$$n=\frac{\theta^2}{v}(r+1)$$

where

$$E=\frac{A+B-\sqrt{B^2+2AC_i\theta^3(e-1)^2}}{2Ae}$$

$$e=e^{\frac{g}{\theta}}$$

$$A=-2vC_t g^2 e$$

$$B=C_i\theta^3(e-1)^2$$

Table 4 lists some examples of optimal procedure for various cases such that the ratios of C_t, C_i are (1 : 100), (1 : 9), (3 : 7), (5 : 5), (7 : 3), (9 : 1), (100 : 1) and $CRLB$'s are 1.0, 4.0, 8.0, 12.0, and g's are 0.4, 1.0, 2.0 when θ is assumed to be 10. It shows that the censoring time, $T=hg$, remains the same despite of the changes of g, and the number of testing item, n, is slightly increased as g becomes larger. And in comparison with Table 1 for the case of continuous observations, n and T in Table1 are, of course, smaller than those in Table 4.

Table 4. Optimal procedures to estimate the life time for the grouped data when θ is assumed to be 10.

		$C_t=1$		$C_t=1$		$C_t=3$		$C_t=5$		$C_t=7$		$C_t=9$		$C_t=100$	
		$C_i=100$		$C_i=9$		$C_i=7$		$C_i=5$		$C_i=3$		$C_i=1$		$C_i=1$	
CRLB	g	h	n	h	n	h	n	h	n	h	n	h	n	h	n
1.0	0.4	173	208	114	207	81	203	62	198	46	190	26	168	8	126
	1.0	68	222	46	220	33	217	25	211	19	202	11	177	4	131
	2.0	34	249	23	247	17	243	13	236	10	224	6	194	2	127
4.0	0.4	139	51	80	50	52	48	37	45	25	41	14	35	4	28
	1.0	56	55	32	54	21	51	15	48	10	43	6	37	2	27
	2.0	28	62	17	60	11	57	8	53	5	45	3	37	1	25
8.0	0.4	122	25	65	24	39	23	27	21	18	19	10	16	3	13
	1.0	49	27	26	26	16	24	11	22	8	20	4	16	2	13
	2.0	25	31	13	29	8	26	6	24	4	20	2	15	1	12
12.0	0.4	112	17	57	16	33	14	23	13	15	12	8	10	3	9
	1.0	45	18	23	17	14	15	9	13	6	12	4	10	1	8
	2.0	23	20	12	19	7	17	5	15	3	12	2	10	1	8

6. Conclusion

The optimal procedures given by (8) and (11), which are obtained by the criteria of *CRLB* instead of the exact variance, are satisfactory as shown in the Table 1 except for the extreme cases. Even in the extreme cases one can get some general idea by the criteria of *CRLB* about that how many items and how much time for test are economically required to achieve certain accuracy of the estimations.

REFERENCES

Blight, B. (1972) On the Most Economical Choice of a Life Testing Procedure for Exponentially Distributed Data, *Technometrics*, Vol. 14, 613-618.

Boardman, T.J., and Kendell, P. (1970) Estimation in Compound Exponential Failure Models, *Technometrics*, Vol. 12, 891-900.

Johnson, N. (1960) An Approximation to the Multinomial Distribution: Some Properties and Applications, *Biometrika*, Vol. 47, 93-102.

Mendenhall, W., and Lehman, Jr. E. (1960) An Approximation of the Negative Moments of the Positive Binomal Useful in Life Testing, *Technometrics*, Vol. 2, No. 2, 227-242.

Mendenhall, W., and Hader, R. (1958) Estimation of Parameters of Mixed Exponentially Distributed Failure Time Distributions from Censored Life Test Data, *Biometrika*, Vol. 45, 504-520.

Moeschberger, M., and David, H.A. (1971) Life Tests under Competing Causes of Failure and the Theory of Competing Risks, *Biometrics*, Vol. 27, 909-933.