SIMPLIFIED ESTIMATES FOR THE EXPONENTIAL DISTRIBUTION[1]

By A. E. Sarhan,[2] B. G. Greenberg, and Junjiro Ogawa[3]

University of North Carolina

1. Introduction and summary. In many practical applications involving statistical estimation, "inefficient" estimates may be the ones of choice for reasons of economy, in money, time, and effort. The usefulness of such estimates, which are generally based upon order statistics, was highlighted by Mosteller [6] who reasoned that observations in large samples could easily be arranged in order of magnitude by punch-card equipment. Moreover, there are many instances where the observations in the sample, by the nature of the measurement, occur naturally in order of magnitude. This happens, for example, in the case of fatigue and life-testing studies. This study deals with certain estimation problems of the one- and two-parameter exponential distributions.

The present study provides, for the parameters of the exponential distribution, estimators that are linear functions of specific subsets of the order statistics. These estimators are optimal in the sense that they provide the most efficient linear combinations of a given number of order statistics.

The use of only two observations from a sample, particularly a small one, represents a situation meriting special study. This paper will first consider which are the best two and which are the worst two order statistics to select for estimation purposes in samples up to size 20. Some consideration will also be given to the use of symmetrically placed order statistics.

For large samples, Ogawa [8], [9] derived optimum spacings to select subsets ranging in size from 1 to 15 for use as linear estimators in the one-parameter exponential distribution. The present paper will consider the same problem where a portion of the sample has been censored at the beginning of the distribution.

2. Estimation using two observations in small samples.

(a) Consider first the two-parameter exponential distribution given by the density

$$f(x) = \begin{cases} (1/\sigma)e^{-(x-\mu)/\sigma}, & \mu \leqq x, \\ 0, & \text{elsewhere}. \end{cases} \tag{2.1}$$

The purpose is to obtain the minimum variance unbiased linear estimates for μ and σ using only two order statistics from a sample of size n. Suppose that these

Received September 8, 1960; revised August 7, 1962.

[1] Sponsored by U. S. Army Research Office (Durham).

[2] Present address is Alexandria University, Alexandria, Egypt.

[3] Present address is Nihon University, Tokyo, Japan.

two order statistics are designated lth and mth in order of magnitude such that $x_{(1)} \leqq x_{(l)} < x_{(m)} \leqq x_{(n)}$.

The matrix of coefficients of (μ, σ) for the expected values (A) of these order statistics and the variance-covariance matrix (V) are as follows:

$$(2.2) \qquad A = \begin{bmatrix} 1 & \sum_1^l a \\ 1 & \sum_1^m b \end{bmatrix} \qquad V = \begin{bmatrix} \sum_1^l b & \sum_1^l b \\ \sum_1^l b & \sum_1^m b \end{bmatrix}$$

where $\sum_1^l a = \sum_{i=1}^l 1/(n - i + 1)$, and $\sum_1^l b = \sum_{i=1}^l 1/(n - i + 1)^2$.

To obtain the best linear unbiased estimates, we use

$$(2.3) \qquad (A'V^{-1}A)^{-1}A'V^{-1} = \left(1 \Big/ \sum_{l+1}^m a\right) \begin{bmatrix} \sum_1^m a & -\sum_1^l a \\ -1 & 1 \end{bmatrix}.$$

Therefore, the estimates are

$$(2.4) \qquad \mu^* = \left(1 \Big/ \sum_{l+1}^m a\right) \left[x_{(l)} \sum_1^m a - x_{(m)} \sum_1^l a\right],$$

$$(2.5) \qquad \sigma^* = \left(1 \Big/ \sum_{l+1}^m a\right) (x_{(m)} - x_{(l)}).$$

The variances for these estimates are taken from $(A'V^{-1}A)^{-1}$ and are

$$(2.6) \qquad V(\mu^*) = \left\{\sum_1^l b + \left[\left(\sum_1^l a\right)^2 \sum_{l+1}^m b \Big/ \left(\sum_{l+1}^m a\right)^2\right]\right\} \sigma^2,$$

$$(2.7) \qquad V(\sigma^*) = \left\{\sum_{l+1}^m b \Big/ \left(\sum_{l+1}^m a\right)^2\right\} \sigma^2.$$

The $(\text{mean})^* = \mu^* + \sigma^*$ and is given by

$$(2.8) \qquad (\text{mean})^* = \left(1 \Big/ \sum_{l+1}^m a\right) \left\{\left(\sum_1^m a - 1\right) x_{(l)} + \left(1 - \sum_1^l a\right) x_{(m)}\right\}$$

and its variance is

$$(2.9) \qquad V(\text{mean}^*) = \left\{\left[\sum_1^l b\right] + \left[\left(\sum_1^l a - 1\right)^2 \sum_{l+1}^m b \Big/ \left(\sum_{l+1}^m a\right)^2\right]\right\} \sigma^2.$$

The expression for the variance of σ^* (2.7) can be rewritten as

$$(2.10) \qquad \begin{aligned} V(\sigma^*) = {} & [(1/(n - l)^2) + \cdots + (1/(n - m + 1)^2)/ \\ & \{1/(n - l) + \cdots + 1/(n - m + 1)\}^2]\sigma^2. \end{aligned}$$

This expression attains a minimum for $l = 1$ and m according to the values of

TABLE A

Sets of two optimal order statistics to estimate parameters in the two-parameter exponential distribution

Order of Observations	n				
	2–6	7–10	11–15	16–20	21
l	1	1	1	1	1
m	n	$n-1$	$n-2$	$n-3$	$n-4$

n shown in Table A. It can be shown that the same results hold for estimating μ and therefore the mean as well.

Table 1 gives the coefficients to be used for estimating μ and σ with two order statistics, the variances of these estimates, and their efficiencies relative to the best linear unbiased estimate based upon the complete sample. It can be seen from this table that in estimating μ, the efficiency of two order statistics relative to the complete sample is high and, in fact, is never less than 93.7 %, this value occurring when $n = 6$. The efficiency increases with sample size after this point and at $n = 20$ it has attained a value of 97.55 %. The efficiency in estimating σ is not as high as that for μ, and decreases almost consistently, although slowly,

TABLE 1

Best two observations in estimating μ and σ, the coefficients in the BLE, the variances of the estimates and their relative efficiency (for the two-parameter exponential distribution).

n	l	m	Estimation of μ: Coefficient of $x_{(l)}$	Coefficient of $x_{(m)}$	$V(\mu)$*	R.E.	Estimation of σ: Coefficient of $x_{(l)}$	Coefficient of $x_{(m)}$	$V(\sigma)$*	R.E.
3	1	3	1.22222	−.22222	.17284	96.43	−.66667	+.66667	.55556	90.00
4	1	4	1.13636	−.13636	.08781	94.90	−.54545	+.54545	.40496	82.31
5	1	5	1.09600	−.09600	.05312	94.13	−.48000	+.48000	.32800	76.22
6	1	6	1.07299	−.07299	.03558	93.70	−.43796	+.43796	.28073	71.24
7	1	6	1.09852	−.09852	.02518	94.57	−.68966	+.68966	.23372	71.31
8	1	7	1.07848	−.07848	.01878	95.10	−.62780	+.62780	.20172	70.82
9	1	8	1.06468	−.06468	.01455	95.44	−.58212	+.58212	.17872	69.94
10	1	9	1.05468	−.05468	.01161	95.67	−.54676	+.54676	.16136	68.86
11	1	9	1.06362	−.06362	.00948	95.92	−.69981	+.69981	.14680	68.12
12	1	10	1.05483	−.05483	.00787	96.26	−.65795	+.65795	.13335	68.18
13	1	11	1.04798	−.04798	.00664	96.51	−.62375	+.62375	.12255	68.00
14	1	12	1.04251	−.04251	.00568	96.70	−.59519	+.59519	.11368	67.67
15	1	13	1.03806	−.03806	.00492	96.85	−.57092	+.57092	.10626	67.22
16	1	13	1.04209	−.04209	.00429	97.02	−.67345	+.67345	.09947	67.02
17	1	14	1.03801	−.03801	.00378	97.19	−.64625	+.64625	.09323	67.04
18	1	15	1.03459	−.03459	.00336	97.33	−.62258	+.62258	.08787	66.94
19	1	16	1.03167	−.03167	.00300	97.45	−.60177	+.60177	.08321	66.77
20	1	17	1.02916	−.02916	.00270	97.55	−.58329	+.58329	.07912	66.52
n	1	.7968n	1.00000	.00000	.00000	100.00	−.6275	+.6275		64.76

R.E. = Relative efficiency of the estimator is the efficiency relative to the minimum variance linear unbiased estimator based upon the complete sample.

as the sample size increases. It approaches an asymptotic limit of 64.76% relative efficiency.

The two largest order statistics give, in general, the lowest overall efficiency among all linear combinations of pairs of order statistics, and it is generally desirable to choose a pair as different as possible from the two largest order statistics, particularly in estimating μ. To estimate the value of μ, the farther l is from 1, the worse becomes the efficiency. The importance of the initial order statistic is not as great in estimating σ, however, and the loss is not so great, particularly as n increases.

(b) Consider now the one-parameter exponential distribution (so that the value of μ is known to be zero). In this case, the comparable equations are

$$\sigma^* = \left\{\left(\sum_{l+1}^{m} b \sum_{1}^{l} a - \sum_{1}^{l} b \sum_{l+1}^{m} a\right) x_{(l)} + \left(\sum_{1}^{l} b \sum_{l+1}^{m} a\right) x_{(m)}\right\} \Big/ \left[\left(\sum_{1}^{l} a\right)^2 \sum_{l+1}^{m} b + \sum_{1}^{l} b \left(\sum_{l+1}^{m} a\right)^2\right], \tag{2.11}$$

$$V(\sigma^*) = \left\{\left(\sum_{1}^{l} b \sum_{l+1}^{m} b\right) \Big/ \left[\left(\sum_{1}^{l} a\right)^2 \sum_{l+1}^{m} b + \sum_{1}^{l} b \left(\sum_{l+1}^{m} a\right)^2\right]\right\} \sigma^2. \tag{2.12}$$

The variance for estimating σ with two order statistics attains a minimum at different points than previously, and the results are summarized in Table B.

TABLE B

Sets of two optimal order statistics to estimate σ in the one-parameter exponential distribution

Order of Observations	n					
	2–4	5–7	8–11	12–15	16–18	19–21
l	$n-1$	$n-2$	$n-3$	$n-4$	$n-6$	$n-7$
m	n	n	n	n	$n-1$	$n-1$

It is clear from Table B that the first ordered observation is no longer crucial as in the estimation procedure for the two-parameter exponential distribution. In fact, the value of l is almost the opposite in Table B since it deviates increasingly from the small order statistics as the sample size increases.

Table 2 presents the coefficients for $x_{(l)}$ and $x_{(m)}$ to be used in estimating σ from the one-parameter exponential distribution, the variances of the estimates, and their efficiency relative to the best linear estimate based upon complete sample. Note that the efficiency decreases gradually to a level of 84.63% for $n = 20$. Also, it can be observed that the estimate based upon the highest two order statistics has a variance equal to the coefficient to be used with $x_{(m)}$.

For large values of n, the asymptotic efficiency for estimating σ with only two order statistics is 82.03%, and the two order statistics to be used are $.6386n$

TABLE 2

Best two observations, the coefficients in the best estimate, the variance of the estimate and its relative efficiency (for the one-parameter exponential distribution).

	n	l	m		Coefficient of		Variance	R.E.
					$x_{(l)}$	$x_{(m)}$		
	2	1	2		.50000	.50000	.50000	100.00
$l = n - 1$	3	2	3		.44737	.34211	.34211	97.44
	4	3	4		.41304	.26522	.26522	94.26
	5	3	5		.52800	.25682	.21402	93.45
$l = n - 2$	6	4	6		.49389	.21665	.18055	92.31
	7	5	7		.46753	.18862	.15718	90.89
	8	5	8	$m = n$	.55363	.18776	.13940	89.67
$l = n - 3$	9	6	9		.52706	.16799	.12472	89.09
	10	7	10		.50503	.15250	.11322	88.32
	11	8	11		.48639	.14003	.10396	87.44
	12	8	12		.55263	.14062	.09609	86.72
$l = n - 4$	13	9	13		.53345	.13047	.08915	86.28
	14	10	14		.51670	.12190	.08330	85.75
	15	11	15		.50192	.11458	.07829	85.15
	16	10	15		.51933	.21701	.07354	84.98
$l = n - 6$	17	11	16		.50658	.20443	.06928	84.91
	18	12	17	$m = n - 1$	.49505	.19343	.06555	84.75
$l = n - 7$	19	12	18		.53132	.19337	.06213	84.71
	20	13	19		.51982	.18387	.05908	84.63
	large	*	**		.5232	.1790		82.03

R.E. = Relative efficiency of the estimator is the efficiency relative to the minimum variance linear unbiased estimator based upon the complete sample.

* .6386n

** .9266n

and .9266n. If we check these values against $n = 20$, $l = 12.77$ and $m = 18.53$. By adjusting them to the nearest integral values, order statistics 13 and 19 result, and these are the same as those in Table 2. Similarly, the asymptotic expressions for the coefficients, .5232 and .1790 compare favorably with the exact values given in Table 2 for $n = 20$, viz. .5198 and .1839. The closeness of the results for $n = 20$ as well as neighboring sample sizes suggests that the asymptotic approximations might be used instead of the exact values for $n > 20$.

Inasmuch as the optimal values of l and m appear to come from the upper end of the distribution, it suggests that an estimator based upon the smallest order statistics will have maximum loss in efficiency. In fact, the worst set of two order statistics is the first and second. The variance of an estimate based upon the first two order statistics is .5, and the relative efficiency to the complete sample is $1/(2n) \to 0$ as $n \to \infty$.

3. Estimation using two symmetric observations in small samples. Although the use of two symmetric order statistics may not always be the most efficient pair to use in estimating μ and σ, it sometimes has the advantage of being simpler to select the required observations. This convenience afforded by symmetry suggests a special investigation into the nature of efficiency with two symmetrical order statistics. In this section the discussion of efficiency in using two order statistics as estimators is intended to be relative to the variance of the best linear estimates based upon all order statistics. This will permit us to judge not only which two symmetric order statistics are best in the group of all two-fold symmetric sets but also the efficiency of using two particular order statistics instead of the best linear estimate based upon all observations.

Recalling the material presented in Section 2, the two order statistics to estimate both parameters in the two-parameter exponential distribution which have highest efficiency for samples of size $n \leqq 6$ are the first and last order statistics. Inasmuch as these are already symmetric, it is clear that the best two symmetric order statistics must be based upon $x_{(1)}$ and $x_{(n)}$ where $n \leqq 6$.

The graphs shown in Figures 1 and 2 present in more detail the efficiencies of the estimates, $\mu^{*\prime}$ and $\sigma^{*\prime}$, where the primes represent estimates using two symmetric order statistics for sample sizes of $n \leqq 20$. The relationships in Figure 1 show clearly that the first and last order statistics are always the symmetric ones to be preferred for maximum efficiency in estimating μ. Moreover,

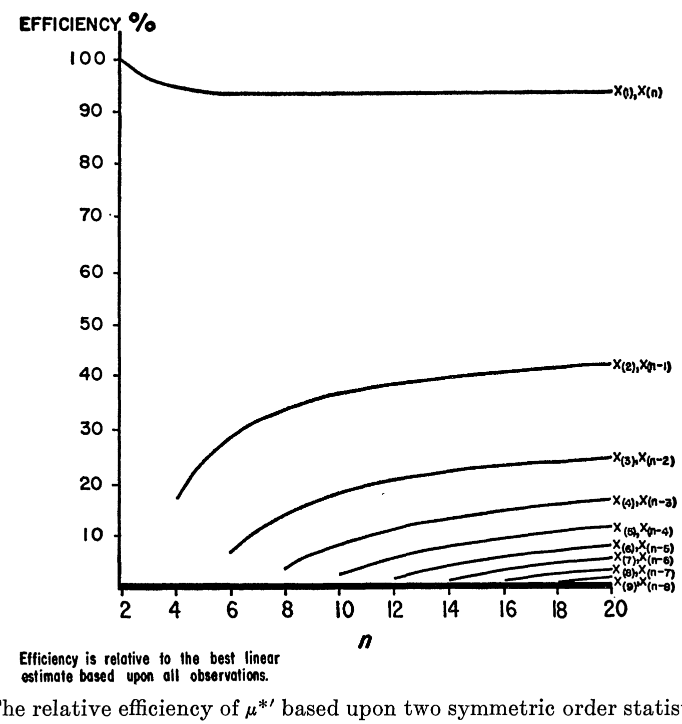
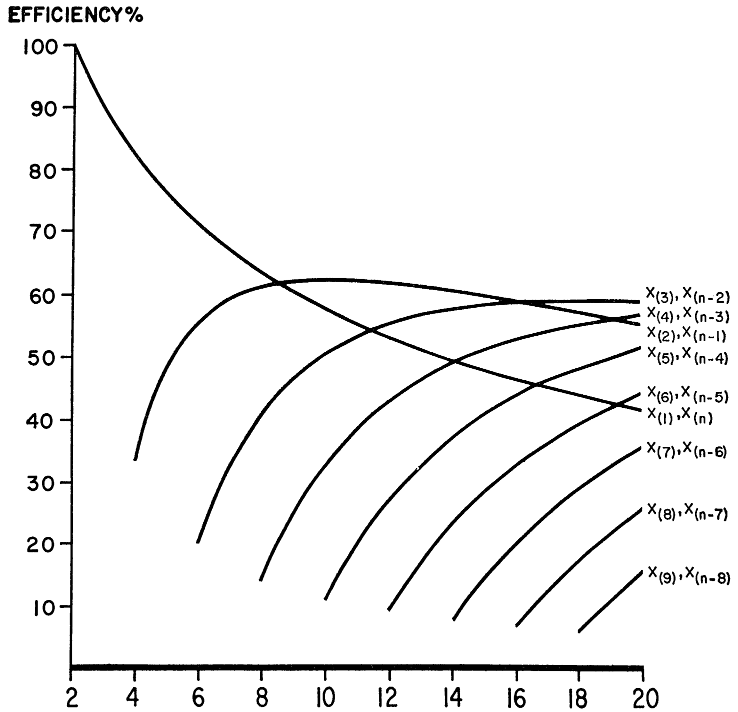


FIG. 1. The relative efficiency of $\mu^{*\prime}$ based upon two symmetric order statistics from the two-parameter exponential distribution.



Efficiency is relative to the best linear estimate based upon all observations.

FIG. 2. The relative efficiency of $\sigma^{*\prime}$ based upon two symmetric order statistics from the two-parameter exponential distribution.

the efficiency of the first and last order statistic is quite high relative to the best linear estimate using the entire sample, because this pair includes the first order statistic which is known to be most influential in determining μ. Consequently, as the two symmetric order statistics diverge from the extremes of the distribution, the efficiency drops rapidly.

The relationships in Figure 2 show that the efficiencies of estimating σ with two symmetric order statistics are not as high as those observed in estimating μ. The maximum efficiencies for estimating σ with two symmetric order statistics are achieved when the order statistics are chosen as shown in Table C.

TABLE C

Most efficient two symmetric order statistics to estimate σ in two-parameter exponential distribution

	Sample size		
	$2 \leqq n \leqq 8$	$9 \leqq n \leqq 16$	$17 \leqq n \leqq 20$
Order of observations	$x_{(1)}$, $x_{(n)}$	$x_{(2)}$, $x_{(n-1)}$	$x_{(3)}$, $x_{(n-2)}$

Thus, the efficiency of using the first and last order statistic to estimate σ drops very rapidly for $n > 2$, and by the time that $n = 9$, the penultimate order statistics exceed it in efficiency among the class of symmetric pairs.

For the one-parameter exponential distribution, the maximum efficiencies for estimating σ with two symmetric order statistics can be determined from Figure 3, and are attained when the order statistics are selected as shown in Table D.

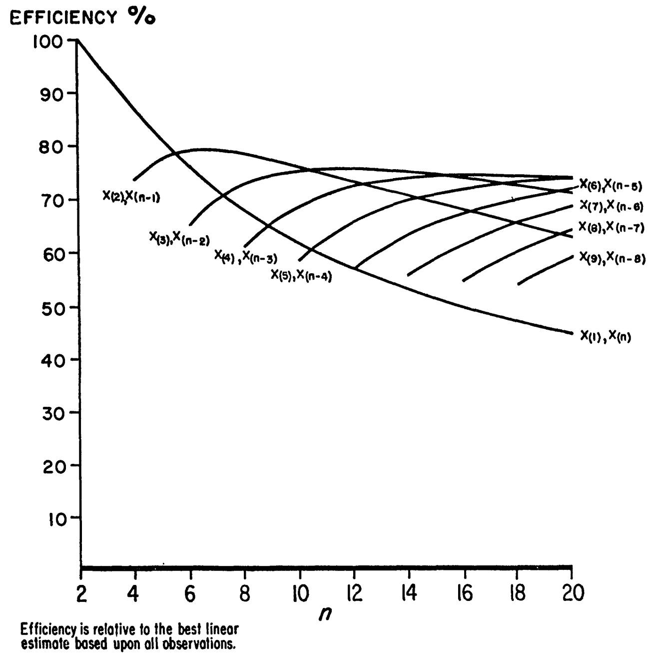


FIG. 3. The relative efficiency of $\sigma^{*\prime}$ based upon two symmetric order statistics from the one-parameter exponential distribution.

TABLE D

Most efficient two symmetric order statistics to estimate σ in one-parameter exponential distribution

	Sample size			
	$2 \leqq n \leqq 5$	$6 \leqq n \leqq 10$	$11 \leqq n \leqq 15$	$16 \leqq n \leqq 20$
Order of Observations	$x_{(1)}$, $x_{(n)}$	$x_{(2)}$, $x_{(n-1)}$	$x_{(3)}$, $x_{(n-2)}$	$x_{(4)}$, $x_{(n-3)}$

Since σ is most efficiently estimated in the one-parameter exponential distribution by $\bar{x}$, then any estimate based upon the difference or the weighted sum of two order statistics will be of relatively lower efficiency. For example, the range and quasi-range would be expected to produce an estimate of σ with a variance considerably higher than that of $\bar{x}$. Harter [3] found that the efficiency of quasi-ranges in this case ranged from 50.00 to 61.73 per cent for samples of size 2 to 100. The quasi-range estimators would have lower efficiency than the use of the two symmetric order statistics here because the former uses equal and opposite weights whereas the latter estimates have positive and differing weights. These weights are obtained as follows:

The coefficient or weight for the first ordered observation used in the estimate, say $x_{(r)}$, is

$$C\left[\sum_{r+1}^{n-r+1} b \sum_{1}^{r} a - \sum_{1}^{r} b \sum_{r+1}^{n-r+1} a\right],$$

where a and b are defined in Section 2 and where

$$C = \left[\left(\sum_{1}^{r} a\right)^{2} \sum_{r+1}^{n-r+1} b + \sum_{1}^{r} b \left(\sum_{r+1}^{n-r+1} a\right)^{2}\right]^{-1}.$$

The coefficient for its symmetric mate, say $x_{(n-r+1)}$ is

$$C\left(\sum_{1}^{r} b \sum_{r+1}^{n-r+1} a\right).$$

4. Optimal spacings in large samples. Ogawa [7] developed the general large sample theory for estimating location and scale parameters based upon sample quantiles. In particular, he determined the asymptotically optimal spacings for the normal distribution. Although several of the tables in that publication had regrettable errors in computation, the numerical results have been confirmed by Cox [1], Kulldorff [4], [5] and others.

These results have also been applied to the one-parameter exponential distribution by Ogawa [8], [9] and his notation is used herewith for convenience.

Consider the single-parameter exponential distribution with density function,

$$g(x) = \begin{cases} (1/\sigma)e^{-x/\sigma}, & 0 \leqq x \\ 0, & \text{elsewhere}, \end{cases} \tag{4.1}$$

where we have an ordered random sample of size n.

Let there be k fixed real numbers, $\lambda_1, \lambda_2, \cdots, \lambda_k$, such that $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k < 1$, and suppose that we select the k sample λ_i quantiles ($i = 1, 2, \cdots, k$) to estimate σ. The k order statistics are $x_{(n_1)}, x_{(n_2)}, \cdots, x_{(n_k)}$, where $n_i = [n\lambda_i] + 1$, and the symbol $[n\lambda_i]$ stands for the greatest integer not exceeding $n\lambda_i$. (This definition of the sample quantile has been chosen for convenience along the lines outlined by Cramér ([2], 367–68), but any reasonable alternative definition would lead to the same asymptotic results.)

The standardized exponential distribution has density

$$f(x) = \begin{cases} \exp(-x), & 0 \leqq x, \\ 0, & \text{elsewhere}. \end{cases} \tag{4.2}$$

If the λ_i-quantile of the standardized distribution is u_i, so that

$$\lambda_i = 1 - \exp(-u_i),$$

and that of the original distribution is x_i, then

$$x_i = u_i\sigma, \qquad i = 1, 2, \cdots, k. \tag{4.3}$$

Ogawa [8], [9] has shown that the asymptotically best linear unbiased estimate (σ^*) of σ is

$$\sigma^* = Y/K_2, \tag{4.4}$$

where

$$Y = \sum_{i=1}^{k+1} [u_i \exp(-u_i) - u_{i-1} \exp(-u_{i-1})][x_{(n_i)} \exp(-u_i) - x_{(n_{i-1})} \exp(-u_{i-1})]/[\exp(-u_{i-1}) - \exp(-u_i)]$$

and where

$$K_2 = \sum_{i=1}^{k+1} [u_i \exp(-u_i) - u_{i-1} \exp(-u_{i-1})]^2/[\exp(-u_{i-1}) - \exp(-u_i)],$$

and the latter is also the asymptotic relative efficiency of the linear systematic statistic based upon the sample λ_i-quantiles. Another method of expressing Y is

$$Y = \sum_{i=1}^{k} a_i x_{(n_i)} \tag{4.5}$$

where

$$a_i = \exp(-u_i)\{[u_i \exp(-u_i) - u_{i-1} \exp(-u_{i-1})]/[\exp(-u_{i-1}) - \exp(-u_i)] - [u_{i+1} \exp(-u_{i+1}) - u_i \exp(-u_i)]/[\exp(-u_i) - \exp(-u_{i+1})]\}.$$

The coefficients a_i for optimal spacings have been tabulated in the references by Ogawa for $k = 1(1)15$. Inasmuch as the resultant value of Y has to be divided by the particular K_2 for the given value of k, in order to derive σ^*, it is more convenient to tabulate b_i such that

$$\sigma^* = \sum_{i=1}^{k} b_i x_{(n_i)}. \tag{4.6}$$

The values of b_i ($i = 1, 2, \cdots, 15$), u_i and λ_i are given in Table 3 such that the asymptotic relative efficiency, K_2 (also tabulated) is a maximum. This table is repeated here because the earlier ones by Ogawa had values of a_i rather than b_i as well as several computational errors in the third and fourth decimal places.[4]

[4] A recent report by Kulldorff [5a] improved the precision of Ogawa's and our own tabulations (in the fourth decimal place) by use of an electronic computer.

TABLE 3

Asymptotically optimum spacings for estimates, asymptotic relative efficiencies and the coefficients of best estimates for the one-parameter exponential distribution

	k														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
u_1	1.5936	1.0177	.7541	.6003	.4994	.4274	.3741	.3323	.2988	.2719	.2494	.2303	.2137	.1994	.1875
λ_1	.7968	.6386	.5296	.4514	.3931	.3478	.3121	.2827	.2583	.2381	.2207	.2057	.1924	.1808	.1710
b_1	.6275	.5232	.4477	.3907	.3463	.3108	.2819	.2579	.2375	.2201	.2053	.1923	.1807	.1703	.1613
u_2		2.6113	1.7718	1.3544	1.0997	.9268	.8015	.7064	.6311	.5707	.5213	.4797	.4440	.4131	.3869
λ_2		.9266	.8300	.7419	.6670	.6042	.5513	.5066	.4680	.4349	.4063	.3810	.3585	.3384	.3208
b_2		.1790	.2266	.2361	.2320	.2228	.2119	.2009	.1905	.1804	.1711	.1627	.1551	.1479	.1412
u_3			3.3654	2.3721	1.8538	1.5271	1.3009	1.1338	1.0052	.9030	.8201	.7516	.6934	.6434	.6006
λ_3			.9655	.9067	.8434	.7828	.7277	.6782	.6340	.5946	.5596	.5284	.5001	.4745	.4515
b_3			.0775	.1195	.1402	.1492	.1519	.1511	.1483	.1446	.1402	.1356	.1312	.1268	.1224
u_4				3.9657	2.8715	2.2812	1.9012	1.6332	1.4326	1.2771	1.1524	1.0504	.9653	.8928	.8309
λ_4				.9810	.9434	.8978	.8506	.8047	.7613	.7212	.6841	.6502	.6191	.5905	.5643
b_4				.0409	.0709	.0902	.1017	.1082	.1116	.1127	.1124	.1111	.1093	.1073	.1051
u_5					4.4651	3.2989	2.6553	2.2335	1.9320	1.7045	1.5265	1.3827	1.2641	1.1647	1.0803
λ_5					.9885	.9631	.9297	.8928	.8551	.8181	.7827	.7491	.7175	.6880	.6605
b_5					.0243	.0456	.0615	.0725	.0799	.0847	.0875	.0891	.0896	.0895	.0889
u_6						4.8925	3.6730	2.9876	2.5323	2.2039	1.9539	1.7568	1.5964	1.4635	1.3522
λ_6						.9925	.9746	.9496	.9205	.8896	.8583	.8274	.7974	.7686	.7413
b_6						.0156	.0311	.0438	.0536	.0607	.0659	.0694	.0718	.0733	.0741
u_7							5.2666	4.0053	3.2864	2.8042	2.4533	2.1842	1.9705	1.7958	1.6510
λ_7							.9948	.9818	.9626	.9394	.9140	.8874	.8606	.8340	.8081
b_7							.0107	.0222	.0324	.0407	.0472	.0522	.0560	.0587	.0607

u_8								5.5989	4.3041	3.5583	3.0536	2.6836	2.3979	2.1699	1.9833
λ_8								.9963	.9865	.9715	.9528	.9317	.9091	.8858	.8624
b_8								.0076	.0164	.0246	.0316	.0374	.0421	.0458	.0486
u_9									5.8977	4.5760	3.8077	3.2839	2.8973	2.5973	2.3574
λ_9									.9973	.9897	.9778	.9625	.9448	.9255	.9053
b_9									.0056	.0124	.0191	.0250	.0301	.0344	.0380
u_{10}										6.1696	4.8254	4.0380	3.4976	3.0967	2.7848
λ_{10}										.9979	.9920	.9824	.9697	.9548	.9383
b_{10}										.0042	.0097	.0151	.0202	.0247	.0285
u_{11}											6.4190	5.0557	4.2517	3.6970	3.2842
λ_{11}											.9984	.9936	.9858	.9752	.9625
b_{11}											.0033	.0077	.0122	.0165	.0204
u_{12}												6.6493	5.2694	4.4511	3.8845
λ_{12}												.9987	.9949	.9883	.9794
b_{12}												.0026	.0062	.0100	.0137
u_{13}													6.8630	5.4688	4.6386
λ_{13}													.9990	.9958	.9903
b_{13}													.0021	.0051	.0083
u_{14}														7.0624	5.6563
λ_{14}														.9991	.9965
b_{14}														.0017	.0042
u_{15}															7.2499
λ_{15}															.9993
b_{15}															.0014
K_2	.6476	.8203	.8910	.9269	.9476	.9606	.9693	.9754	.9798	.9832	.9857	.9878	.9894	.9907	.9918

The values given in Table 3 have been especially helpful in dealing with estimates of the parameter in samples containing a large series of cases. For example, the distribution of age at death among those dying under one year of age appears to follow closely the one-parameter exponential distribution. In the United States, such infant deaths are over 100,000 per year and the use of order statistics is well justified in estimation problems relating thereto.

Situations occur, however, where the optimum spacings given in Table 3 cannot be applied because the magnitude of the initial ordered observations in the sample have been censored. Thus, many countries do not classify age at death in the same detail as done in this country in which the basis of classification is under 1 hour, 1.0 to 23.9 hours, 24.0 to 47.9 hours, etc. Where the vital statistics do not specify the exact age at death under 1 day, for instance, about 40% of the deaths under one year will have had the initial ordered observations censored. This means that if the magnitude of the first 40% of the ordered observations were unavailable, one would be restricted to the use of four or less spacings since $\lambda_1 < .40$ for all $k \geqq 5$.

A general solution to this problem would be to consider the asymptotically optimal spacings when the lowest available sample quantile is preassigned. Under the assumption that λ_1 is fixed in advance, it is of interest to know the optimal spacings for the best linear unbiased estimate and their efficiency relative to the case when λ_1 is not predetermined.

Following the principle used by Ogawa, the asymptotically optimal spacings, for a fixed u_1, are those which maximize K_2. Proceeding along the lines of development used in the unrestricted case, it is soon evident that for this specific distribution the optimal spacings of the sample quantiles, when λ_1 is fixed, have the same relative positioning as when λ_1 is not preassigned.

Table 4 presents as an illustration optimal spacings for k up to 9 when $\lambda_1 = .4$. An examination of the values of K_2 in this table reveals that the loss of efficiency in estimating σ is rather trivial even in those cases not anticipated. For example, consider $k = 9$, in Table 3, where the asymptotic efficiency is 97.98 per cent when $\lambda_1 = 25.83$ percentile, and the subsequent quantiles are chosen accordingly. If, however, the 25th percentile is not available or, in fact, even if 40% of the data are censored such that λ_1 must be .400, then the results in Table 4 show the asymptotic efficiency is affected only slightly from 97.98 to 97.66.

Addendum. Since submission of the original manuscript upon which the present revision is based, Harter has published a paper closely related to the work contained herein. His tabulations confirm the values given in Tables 1 and 2 and extend the cases up to $n = 100$. His coefficients have been carried out to six decimal places. The reference is Harter, H. Leon (1961). Estimating the parameters of negative exponential populations from one or two order statistics. *Ann. Math. Statist.* **32** 1078–1090.

TABLE 4

Optimum spacings, asymptotic relative efficiencies and coefficients in the BLE of σ, given $\lambda_1 = .4$

	k				
	5	6	7	8	9
u_1	.5108	.5108	.5108	.5108	.5108
λ_1	.4000	.4000	.4000	.4000	.4000
b_1	.3466	.3163	.2949	.2790	.2666
u_2	1.1111	1.0102	.9382	.8849	.8431
λ_2	.6708	.6359	.6087	.5872	.5696
b_2	.2293	.2051	.1852	.1685	.1546
u_3	1.8652	1.6105	1.4376	1.3123	1.2172
λ_3	.8451	.8002	.7625	.7308	.7039
b_3	.1386	.1374	.1327	.1267	.1204
u_4	2.8829	2.3646	2.0379	1.8117	1.6446
λ_4	.9440	.9060	.8697	.8366	.8069
b_4	.0701	.0830	.0889	.0908	.0907
u_5	4.4765	3.3823	2.7920	2.4120	2.1440
λ_5	.9886	.9660	.9387	.9104	.8828
b_5	.0240	.0420	.0537	.0608	.0649
u_6		4.9759	3.8097	3.1661	2.7443
λ_6		.9931	.9778	.9578	.9357
b_6		.0144	.0272	.0368	.0435
u_7			5.4033	4.1838	3.4984
λ_7			.9955	.9848	.9698
b_7			.0093	.0186	.0263
u_8				5.7774	4.5161
λ_8				.9969	.9891
b_8				.0064	.0133
u_9					6.1097
λ_9					.9978
b_9					.0045
K_2	.9476	.9600	.9678	.9730	.9766

REFERENCES

[1] Cox, D. R. (1957). Note on grouping. *J. Amer. Statist. Assoc.* **30** 543–547.

[2] Cramér, Harald (1946). *Mathematical Methods of Statistics*, Princeton University Press, Princeton.

[3] Harter, H. Leon (1959). The use of sample quasi-ranges in estimating population standard deviation. *Ann. Math. Statist.* **30** 980–999.

[4] KULLDORFF, GUNNAR (1958). Maximum likelihood estimation for the exponential distribution when the sample is grouped. Department of Statistics, University of Lund, Sweden.

[5] KULLDORFF, GUNNAR (1961). *Contributions to the Theory of Estimation from Grouped and Partially Grouped Samples*. Almqvist and Wiksell, Stockholm.

[5a] KULLDORFF, GUNNAR (1962). On the asymptotically optimum spacings for the exponential distribution. Preliminary Report, Department of Statistics, University of Lund.

[6] MOSTELLER, FREDERICK (1946). On some useful inefficient statistics. *Ann. Math. Statist.* **17** 377–408.

[7] OGAWA, JUNJIRO (1951). Contribution to the theory of systematic statistics, I. *Osaka Math. J.* **3** No. 2 175–213.

[8] OGAWA, JUNJIRO (1957). A further contribution to the theory of systematic statistics, *University of North Carolina Institute of Statistics, Mimeograph Series*, No. 168, 1–32.

[9] OGAWA, JUNJIRO (1960). Determination of optimum spacings for the estimation of the scale parameter of an exponential distribution based on sample quantiles. *Ann. Inst. Statist. Math.* **12** 135–141.