ESTIMATION OF THE PARAMETERS OF THE EXPONENTIAL DISTRIBUTION BASED ON OPTIMUM ORDER STATISTICS IN CENSORED SAMPLES

By A. K. Md. Ehsanes Saleh[1]

University of Western Ontario

1. Introduction and summary. In the theory of estimation it is well known that when all the observations in a sample are available, it is sometimes possible to find estimators that are the most efficient linear combinations of a given number of order statistics. In many practical situations, we encounter censored samples, that is, samples where values of some of the observations are not available, and it is desired to obtain linear estimators based on a few optimal order statistics from such a sample.

The present study concerns the determination of the optimal set of order statistics for a given integer k (where k is much less than the number of observations in the censored sample), in estimating the parameters of the exponential distribution when the sample is censored. The study is based on the asymptotic theory and under Type II censoring scheme.

The problem of estimation of the parameters of the exponential distribution in censored samples has been considered by Sarhan and Greenberg (1957). The choice of optimal set of order statistics for the scale parameter alone in a left censored sample has been studied numerically by Sarhan, Greenberg and Ogawa (1963).

For the estimation of the parameters of exponential distribution based on optimal set of order statistics, we present in Section 2 the asymptotically best linear unbiased estimates (BLUE's) of the parameters based on k sample quantiles of given orders when all the sample values are available and define the censored samples considered. In Sections 3–5, the detailed treatment for the determination of the k optimum order statistics in singly and doubly censored sample is presented. In Section 6 some extremal properties of a related function are given. The results are always referred to in the text of Sections 2 to 5 to establish uniqueness of the optimal order statistics so determined.

Further, for $k = 2(1)4$ and proportion of censoring the right $1 - \beta = .05(.05).40$, Table I has been prepared for the estimation of the scale parameter (assuming the location parameter known) furnishing the coefficients of the BLUE and the spacings corresponding to the optimum order statistics. For $k = 2(1)4$ and equal proportions of censoring on both sides from .05 to .25 at steps of .05, Table II has been prepared for the simultaneous estimation of the location and the scale parameters.

Received 13 January 1965; revised 4 June 1966.

[1] Now at Carleton University, Ottawa.

TABLE I

The optimum values of $p_1, \cdots, p_k$ for the linear estimate of σ based on k order statistics when the sample is censored on the right for $k = 2(1)4$ and for $1 - \beta = .05(.05)40$ with coefficients $b_1, \cdots, b_k$

Per Cent Censored	k	t_1	t_2	t_3	t_4	Q_k^*	p_1	p_2	p_3	p_4	b_1	b_2	b_3	b_4
5	2	1.0177	2.6613			.8203	.6386	.9500			.5232	.1790		
	3	.7033	1.6294	3.0000		.8885	.5051	.8080	.9500		.4373	.2354	.1029	
	4	.5132	1.1341	1.9208	3.0000	.9151	.4014	.6783	.8536	.9500	.3643	.2407	.1423	.8894
10	2	.9278	2.3025			.8159	.6046	.9000			.5181	.2255		
	3	.5877	1.3217	2.3025		.8621	.4444	.7334	.9000		.4156	.2546	.1821	
	4	.4304	.9335	1.5395	2.3025	.8785	.3498	.6068	.7856	.9000	.3413	.2439	.1628	.1627
15	2	.7973	1.8971			.7932	.5495	.8500			.5124	.3118		
	3	.5084	1.1219	1.8971		.8245	.3985	.6712	.8500		.4021	.2669	.2614	
	4	.3733	.8004	1.2982	1.8971	.8356	.3115	.5509	.7272	.8500	.3266	.2457	.1762	.2386
20	2	.6961	1.6094			.7603	.5015	.8000			.5089	.4012		
	3	.4462	.9712	1.6094		.7822	.3599	.6214	.8000		.3921	.2763	.3459	
	4	.3285	.6975	1.1187	1.6094	.7899	.2800	.5022	.6733	.8000	.3157	.2469	.1864	.3204
25	2	.6127	1.3863			.7218	.4582	.7500			.5067	.4975		
	3	.3945	.8493	1.3863		.7374	.3260	.5723	.7500		.3843	.2837	.4382	
	4	.2910	.6134	.9749	1.3863	.7429	.2525	.4585	.6228	.7500	.3070	.2476	.1945	.4105
30	2	.5413	1.2039			.6798	.4180	.7000			.5050	.6036		
	3	.3499	.7465	1.2039		.6909	.2952	.5260	.7000		.3777	.2899	.5410	
	4	.2586	.5418	.8545	1.2039	.6949	.2279	.4183	.5747	.7000	.2997	.2481	.2015	.5115
35	2	.4788	1.0499			.6356	.3805	.6500			.5038	.7227		
	3	.3107	.6575	1.0498		.6436	.2671	.4819	.6500		.3721	.2953	.6575	
	4	.2299	.4791	.7509	1.0498	.6463	.2054	.3807	.5281	.6500	.2935	.2486	.2074	.6265
40	2	.4230	.9162			.5898	.3449	.6000			.5029	.8593		
	3	.2754	.5786	.9162		.5954	.2409	.4494	.6000		.3672	.3000	.7916	
	4	.2041	.4232	.6595	.9162	.5973	.1847	.3451	.4829	.6000	.2881	.2489	.2126	.7593

2a. Asymptotically best linear unbiased estimates (BLUE's) of the parameters based on k sample quantiles. Consider the one and two-parameter exponential distributions given by

$$f(x) = \sigma^{-1}e^{-x/\sigma}, \qquad 0 \leqq x \leqq \infty; \sigma > 0; \tag{2.1}$$

$$f(x) = \sigma^{-1}e^{-(x-\mu)/\sigma}; \qquad \mu \leqq x \leqq \infty; \sigma > 0; \tag{2.2}$$

where μ and σ are, respectively, the location and the scale parameters of the distribution. Assume that the sample size n is large and $k \leqq n$. Let the ordered observations in a random sample of size n be $x_{(1)} < \cdots < x_{(n)}$ and consider the k sample quantiles $x_{(n_1)}, \cdots, x_{(n_k)}$ where $n_1, \cdots, n_k$ are the respective ranks which satisfy the inequality $1 \leqq n_1 < \cdots < n_k \leqq n$. The integers $n_1, \cdots, n_k$ are determined by k fixed real numbers $p_1, \cdots, p_k$ which satisfy the order relation $0 < p_1 < \cdots < p_k < 1$, and $n_i = [np_i] + 1, i = 1, \cdots, k$. $[np_i]$ is the Euler notation denoting the largest integer not exceeding np_i. Define $p_0 = 0$ and $p_{k+1} = 1$. Ogawa (1960) and Kulldorff (1963a) have shown that the asymptotically best linear unbiased estimates (BLUE) of μ and σ based on k sample quantiles $x_{(n_1)}, \cdots, x_{(n_k)}$ are

$$\hat{\sigma} = \sum_{i=1}^{k} c_i x_{(n_i)}; \tag{2.3}$$

$$\hat{\mu} = x_{(n_1)} - \hat{\sigma}u_1, \tag{2.4}$$

where

$$c_1 = -\{(u_2 - u_1)/(e^{u_2} - e^{u_1})L\}; \tag{2.5a}$$

$$c_i = L^{-1}\{(u_i - u_{i-1})/(e^{u_i} - e^{u_{i-1}}) - (u_{i+1} - u_i)/(e^{u_{i+1}} - e^{u_i})\}, \quad i = 2, \cdots, (k-1); \tag{2.5b}$$

$$c_k = L^{-1}\{(u_k - u_{k-1})/(e^{u_k} - e^{u_{i-1}})\}; \tag{2.5c}$$

$$L = \sum_{i=2}^{k} (u_i - u_{i-1})^2/(e^{u_i} - e^{u_{i-1}}), \tag{2.6}$$

where $u_i = \ln(1 - p_i)^{-1}, i = 1, \cdots, k$, are quantiles of the standard exponential distribution corresponding to $p_1, \cdots, p_k$.

The variances, and the covariance of the estimates of $\hat{\mu}$ and $\hat{\sigma}$ are

$$V(\hat{\mu}) = \{u_1^2/L + (e^{u_1} - 1)\}\sigma^2/n; \tag{2.7a}$$

$$V(\hat{\sigma}) = L^{-1}\sigma^2/n; \tag{2.7b}$$

$$\operatorname{cov}(\hat{\mu}, \hat{\sigma}) = (u_1/L)(\sigma^2/n); \tag{2.7c}$$

and the generalized variance of $\hat{\mu}$ and $\hat{\sigma}$ defined by $V(\hat{\mu})\cdot V(\hat{\sigma}) - \operatorname{Cov}^2(\hat{\mu}, \hat{\sigma})$ is

$$\Lambda = [(e^{u_1} - 1)/L]\sigma^4/n^2. \tag{2.8}$$

If $\mu = 0$, then the corresponding estimate of σ based on the k quantiles is

$$\hat{\sigma} = \sum_{i=1}^{k} b_i x(n_i) \tag{2.9}$$

with variance

$$V(\hat{\sigma}) = Q_k^{-1}\sigma^2/n \tag{2.10}$$

where

$$b_i = Q_k^{-1}\{(u_i - u_{i-1})/(e^{u_i} - e^{u_{i-1}}) - (u_{i+1} - u_i)/(e^{u_{i+1}} - e^{u_i})\}, \quad i = 1, \cdots, (k-1); \tag{2.11a}$$

$$b_k = Q_k^{-1}\{(u_k - u_{k-1})/(e^{u_k} - e^{u_{k-1}})\}; \tag{2.11b}$$

$$Q_k = \textstyle\sum_{i=1}^{k} (u_i - u_{i-1})^2/(e^{u_i} - e^{u_{i-1}}). \tag{2.12}$$

The subscript of Q denotes its dimension. It may be noted that the above results are based on the asymptotic distribution of the k-quantiles [Mosteller (1946); Ogawa (1951)] and on the application of the Gauss-Markoff theorem [Kulldorff (1963b)].

2b. Censored samples considered. Consider $x_1, x_2, \cdots, x_n$ to be a random sample of size n from the one-parameter distribution (2.1) or the two-parameter distribution (2.2). Let $x_{(r_1)} < x_{(r_1+1)} < \cdots < x_{(n)}$ be the uncensored portion of the sample of size $(n - r_1 + 1)$ in which $x_{(r_1)}$ is the smallest observation. The integer r_1 is the rank of the observation $x_{(r_1)}$ in the complete sample and is determined by a prefixed real number α such that $0 < \alpha < 1$, using the relation $r_1 = [n\alpha] + 1$. Then, α is the proportion of censoring on the left. Alternatively, let $x_{(1)} < x_{(2)} < \cdots < X_{(r_2)}$ be the uncensored portion of the sample of size r_2 in which $x_{(r_2)}$ be the largest available observation. The integer r_2 is determined as above by a real number β such that $0 < \beta < 1$, using the relation $r_2 = [n\beta] + 1$. Then, $1 - \beta$ is the proportion of censoring on the right. Finally, let $x_{(r_1)} < x_{(r_1+1)} < \cdots < x_{(r_2)}$ be the uncensored portion of the sample of size $(r_2 - r_1 + 1)$ in which $x_{(r_1)}$ is the smallest and $x_{(r_2)}$ is the largest available observations in the censored sample. The integers r_1 and r_2 are determined by two real numbers α and β such that $0 < \alpha < \beta < 1$, using the above mentioned relations. Then, the quantities α and $1 - \beta$ are respectively the proportions of censoring on the left and on the right and we have a doubly censored sample. The next subsequent sections deal with the determination of optimum order statistics when we have singly and doubly censored samples as discussed above.

3. Determination of the optimum order statistics for the estimation of the parameters from a left censored sample. For a given $k(<n - r_1 + 1)$, consider the order statistics $x_{(n_1)}, \cdots, x_{(n_k)}$ where $n_1, \cdots, n_k$ are the respective ranks of the order statistics determined by k fixed real numbers $p_1, \cdots, p_k$ having the order relations $\alpha \leqq p_1 < \cdots < p_k < 1$ and $n_i = [np_i] + 1$, $i = 1, \cdots, k$. Set $p_0 = 0$ and $p_{k+1} = 1$. Define $u_i = \ln(1 - p_i)^{-1}$, $i = 1, \cdots, k$, corresponding to $p_1, \cdots, p_k$. Then $u_1, \cdots, u_k$ satisfy the inequality $\ln(1 - \alpha)^{-1} \leqq u_1 < \cdots < u_k < \infty$. Subject to these restrictions on u's, the BLUE's of μ and σ based on k order statistics and their variances, covariance and the generalized variance are given by (2.3) through (2.8) respectively.

3.1. *Two-parameter problem.* In order to determine the optimum order statistics for this problem, we obtain the k optimum spacings $p_1^0, \cdots, p_k^0$ by maximizing $L(e^{u_1} - 1)^{-1}$ with respect to $u_1, \cdots, u_k$ over the domain $\ln(1-\alpha)^{-1} \leqq u_1 < \cdots < u_k < \infty$. It is known [Saleh and Ali, (1966)] that $L(e^{u_1} - 1)^{-1}$ is a monotonically decreasing function of u_1 and attains its maximum when u_1 assumes its smallest value $u_1^0 = \ln(1-\alpha)^{-1}$ in the domain. Therefore, the optimum spacing is $p_1^0 = \alpha$ and the corresponding optimum rank of the order statistics is $r_1 = [n\alpha] + 1$. The remaining $k-1$ optimum spacings are determined by maximizing $e^{-u_1^0}(e^{u_1^0} - 1)^{-1}Q_{k-1}(t_1, \cdots, t_{k-1})$ with respect to $t_1, \cdots, t_{k-1}$ where $t_{i-1} = u_i - u_1^0$, $i = 2, \cdots, k$, and $t_0 = 0$. By Theorem 6.1 (Section 6), this function $Q_{k-1}(t_1, \cdots, t_{k-1})$ has a unique maximum over the domain $0 < t_1 < \cdots < t_{k-1} < \infty$. Let $(t_1^0, \cdots, t_{k-1}^0)$ be the point at which this maximum occurs with maximum value Q_{k-1}^0. Then setting $\lambda_i^0 = 1 - e^{-t_i^0}$, $i = 1, \cdots, k-1$, and using the relation $t_{i-1} = u_i - u_1^0$, $i = 2, \cdots, k$ the optimum spacings are

$$p_{i+1}^0 = \alpha + (1-\alpha)\,\lambda_i^0, \qquad i = 1, \cdots, (k-1),$$

and the optimum ranks of the order statistics are

$$n_i^0 = [np_i^0] + 1, \qquad i = 2, \cdots, k.$$

The BLUE's of μ and σ based on $x_{(r_1)}, x_{(n_2^0)}, \cdots, x_{(n_k^0)}$ are:

$$\hat{\mu} = x_{(r_1)} - \hat{\sigma}\ln(1-\alpha)^{-1};$$
$$\hat{\sigma} = b_0^0 x_{(r_1)} + \textstyle\sum_{i=1}^{k-1} b_i^0 x_{(n_{i+1}^0)},$$

where $b_0^0 = -\sum_{i=1}^{k-1} b_i^0$ and $b_1^0, \cdots, b_{k-1}^0$ are given in Table 3 [Sarhan, Greenberg, and Ogawa (1963)]. The joint asymptotic efficiency (JAE) and asymptotic relative efficiency (ARE) of the BLUE's compared to the BLUE's using all the observations in the censored sample are:

$$\begin{aligned} \text{JAE}(\hat{\mu}, \hat{\sigma}) &= Q_{k-1}^0; \\ \text{ARE}(\hat{\sigma}) &= Q_{k-1}^0; \\ \text{ARE}(\hat{\mu}) &= Q_{k-1}^0\{\alpha + \ln^2(1-\alpha)^{-1}\}/\alpha Q_{k-1}^0 + \ln^2(1-\alpha)^{-1}. \end{aligned}$$

For $k = 5$, $n = 70$ and $\alpha = .41$, it is easily verified that the optimum order statistics are $x_{(29)}, x_{(47)}, x_{(60)}, x_{(66)}$ and $x_{(70)}$ and the BLUE's are

$$\hat{\mu} = x_{(29)} - \hat{\sigma}\ln(.59)^{-1},$$

$$\hat{\sigma} = -.7871x_{(29)} + .3907x_{(47)} + .2361x_{(60)} + .1195x_{(66)} + .0409x_{(70)}.$$

The coefficients are taken from "Table 3, (1963)"[2] for $k = 4$. The efficiencies are JAE $(\hat{\mu}, \hat{\sigma}) = 92.69\%$, ARE $(\hat{\sigma}) = 92.69\%$ and ARE $(\hat{\mu}) = 97.03\%$.

3.2. *One-parameter problem.* In order to determine the optimum order statistics. for the single parameter σ, we determine the k optimum spacings by maximizing the expression Q_k in (2.12) over the domain $\ln(1-\alpha)^{-1} \leqq u_1 < \cdots < u_k < \infty$.

[2] Table 3 (1963) refers to Sarhan, Greenberg, and Ogawa (1963).

Such a function has been studied in Section 6. By Theorem 6.1, Q_k has a unique maximum inside the domain $0 < u_1 < \cdots < u_k < \infty$. Let this maximum be attained at $(u_1^0, \cdots, u_k^0)$. Now, for the maximum of Q_k over the domain $\ln(1-\alpha)^{-1} \leqq u_1 < \cdots < u_k < \infty$, two situations arise according to the proportion of censoring: (i) α is such that $\ln(1-\alpha)^{-1} \leqq u_1^0$ and (ii) α is such that $\ln(1-\alpha)^{-1} > u_1^0$. In the first case, we take the maximum of Q_k corresponding to $(u_1^0, \cdots, u_k^0)$ and the optimum order statistics are the same as those in the uncensored sample. In the second case, by Corollary 6.2.1, Q_k as a function of u_1 alone is decreasing and attains its maximum when u_1 attains its smallest value $\ln(1-\alpha)^{-1}$. Therefore, the optimum value of p_1 is α and the rank of the corresponding order statistics is $r_1 = [n\alpha] + 1$. The remaining $(k-1)$ are determined as outlined in Section 3.1 with the optimum spacings $p_{i+1}^0 = \alpha + (1-\alpha)\lambda_i^0$, $i = 1, \cdots, (k-1)$, where $\lambda_1^0, \cdots, \lambda_{k-1}^0$ are the optimum spacings for the BLUE of σ when $\mu = 0$ when selecting $k-1$ order statistics in an uncensored sample. The BLUE of σ is thus given by

$$\hat{\sigma} = b_0^0 x_{(r_1)} + \sum_{i=1}^{k-1} b_i x_{(n_{i+1}^0)}$$

where $b_0^0 = (Q_{k-1}^0)^{-1}\{\ln(1-\alpha)^{-1}/\alpha - t_1^0/(e^{t_1^0} - 1)\}$ with $t_1^0 = \ln(1-\lambda_1^0)^{-1}$. The asymptotic relative efficiency (ARE) of this BLUE compared to the BLUE using all the observations in the censored sample is

$$\text{ARE}(\hat{\sigma}) = [\alpha Q_{k-1}^0 + \ln(1-\alpha)^{-1}]/[\alpha + \ln^2(1-\alpha)^{-1}].$$

The quantity Q_{k-1}^0 is the efficiency of the BLUE of σ based on $(k-1)$ order statistics in uncensored sample. For $k = 6$, $n = 70$ and $\alpha = .41$, it is easily verified that the optimum order statistics are $x_{(29)}$, $x_{(45)}$, $x_{(57)}$, $x_{(64)}$, $x_{(68)}$, and $x_{(70)}$. The BLUE of σ is $\hat{\sigma} = .4082x_{(29)} + .3463x_{(45)} + .2320x_{(57)} + .1402x_{(64)} + .0709x_{(68)} + .0243x_{(70)}$, and ARE $(\hat{\sigma}) = 96.20\%$. The coefficients of $x_{(45)}$, $x_{(57)}$, $x_{(64)}$, $x_{(68)}$ and $x_{(70)}$, are taken from "Table 3 (1963)".

4. Determination of the optimum order statistics for the estimation of the parameters from a right censored sample. For a given k $(<r_2)$, consider the order statistics $x_{(n_1)}, \cdots, x_{(n_k)}$ with ranks $n_1, \cdots, n_k$ determined by k fixed real numbers $p_1, \cdots, p_k$ satisfying the order relation $0 < p_1 < \cdots < p_k \leqq \beta$, and $n_i = [np_i] + 1$, $i = 1, \cdots, k$. Define $p_0 = 0$ and $p_{k+1} = 1$.

4.1. *One-parameter problem.* In order to estimate the scale parameter σ, we define $t_i = \ln(1-p_i)^{-1}$, $i = 1, \cdots, k$, corresponding to $p_1, \cdots, p_k$. Then $t_1, \cdots, t_k$ satisfy the inequality $0 < t_1 < \cdots < t_k \leqq \ln(1-\beta)^{-1}$. Based on the k order statistics $x_{(n_1)}, \cdots, x_{(n_k)}$, the asymptotic BLUE of σ and its variance are given by (2.9) and (2.10) respectively replacing u's by t's with the above restriction on the t's.

For the determination of the optimum order statistics, we have to minimize the variance of the estimate, or equivalently, we maximize $Q_k(t_1, \cdots, t_k)$ given by

$$Q_k(t_1, \cdots, t_k) = \sum_{i=1}^{k} (t_i - t_{i-1})^2/(e^{t_i} - e^{t_{i-1}}), \qquad t_0 = 0$$

with respect to $t_1, \cdots, t_k$ over the domain $0 < t_1 < \cdots < t_k \leqq \ln(1-\beta)^{-1}$. By Theorem 6.1, the above function Q_k has a unique maximum inside the domain $0 \leqq t_1 \leqq \cdots \leqq t_k \leqq \infty$. Let this maximum correspond to the point $(t_1^0, \cdots, t_k^0)$. For the maximum of Q_k over the domain $0 < t_1 < \cdots < t_k \leqq \ln(1-\beta)^{-1}$, we have to consider the following two cases: (i) the proportion of censoring $1-\beta$ is such that $t_k^0 \leqq \ln(1-\beta)^{-1}$; (ii) $1-\beta$ is such that $t_k^0 > \ln(1-\beta)^{-1}$.

In the first case, we take the maximum of Q_k corresponding to $(t_1^0, \cdots, t_k^0)$, and the k optimum order statistics are the same as those in the uncensored sample.

In the second case, by Theorem 6.3, the maximum of Q_k occurs on the boundary $t_k = \ln(1-\beta)^{-1}$. It follows that the optimum value of p_k is β and the rank of the corresponding optimum order statistic is $r_2 = [n\beta] + 1$. Thus, we include the largest available observation $x_{(r_2)}$ in the relevant sample.

To determine the remaining $(k-1)$ optimum order statistics, we maximize $Q_k(t_1, \cdots, t_k)$ keeping $t_k = \ln(1-\beta)^{-1}$ fixed. By Theorem 6.3, this maximum is unique and corresponds to the solution of the system of equations

$$\tau_{i+1} + \tau_i - 2t_i = 0, \qquad i = 1, \cdots, k,$$

with $t_k = \ln(1-\beta)^{-1}$. The equation above is discussed in Section 6.

Suppose the quantities $t_1^*, \cdots, t_{k-1}^*$ correspond to the solution of the above system. Then the optimum values of $p_1, \cdots, p_{k-1}$ are obtained by setting

$$p_i^* = 1 - e^{-t_i^*}, \qquad i = 1, \cdots, (k-1).$$

Thus the ranks of the remaining $(k-1)$ optimum order statistics are given by

$$n_i^* = [np_i^*] + 1, \qquad i = 1, \cdots, (k-1).$$

Therefore, the relevant sample consists of the order statistics $x_{(n_1^*)}, \cdots, x_{(n_{k-1}^*)}$, $x_{(r_2^*)}$ and are uniquely determined. The estimate of σ based on $x_{(n_1^*)}, \cdots, x_{(n_{k-1}^*)}$, $x_{(r_2)}$ and the coefficients depending on $t_1^*, \cdots, t_k^*$ are given by

$$\hat{\sigma} = \textstyle\sum_{i=1}^{k-1} b_i^* x_{(n_i^*)} + b_k^* x_{(r_2)},$$

where

$$b_i^* = (Q_k^*)^{-1}\{(t_i^* - t_{i-1}^*)/(e^{t_i^*} - e^{t_{i-1}^*}) - (t_{i+1}^* - t_i^*)/e^{t_{i+1}^*} - e^{t_i^*}\},$$
$$i = 1, \cdots, (k-1);$$

$$b_k^* = (Q_k^*)^{-1}(1-\beta)\{(\ln(1-\beta)^{-1} - t_{k-1}^*)/(1-(1-\beta)e^{t_{k-1}})\},$$

where Q_k^* is the maximum value of Q_k over the domain $0 < t_1 < \cdots < t_k \leqq \ln(1-\beta)^{-1}$. The asymptotic relative efficiency (ARE) of the estimate compared to the best linear estimate using all the observations in the censored sample is given by ARE $(\hat{\sigma}) = Q_k^*/\beta$. The values of $t_1^*, \cdots, t_k^*, p_1^*, \cdots, p_k^*, b_1^*, \cdots, b_k^*$ and Q_k^* have been computed for $k = 2(1)4$ and $1-\beta = .05(.05).40$ and are presented in Table I. Using this table, we illustrate the estimation procedure by the following example:

EXAMPLE. Assume $k = 4$, $n = 72$, $1 - \beta = .20$ and $r_2 = 58$. From "Table 3 (1963)" it is easily checked that this situation belongs to case (ii). First we choose $x_{(58)}$. For the remaining three optimum order statistics, we use Table I, for $k = 4$ and obtain

$$p_1^* = .2800, \qquad p_2^* = .5022, \qquad p_3^* = .6733.$$

The optimum order statistics are thus $x_{(21)}$, $x_{(37)}$, $x_{(49)}$, and $x_{(58)}$. The BLUE is

$$\hat{\sigma} = .3157x_{(21)} + .2469x_{(37)} + .1864x_{(49)} + .3204x_{(48)};$$

$$\text{ARE}(\hat{\sigma}) = 98.74\%.$$

4.2. *Two-parameter problem.* In order to estimate the location and scale parameters μ and σ consider the order statistics $x_{(m)}, x_{(n_2)}, \cdots, x_{(n_k)}$ with ranks $m, n_2, \cdots, n_k$ such that $1 \leqq m < n_2 < \cdots < n_k \leqq r_2$. The ranks $n_2, \cdots, n_k$ are determined by $(k - 1)$ fixed real numbers $p_2, \cdots, p_k$ satisfying the order relation $0 < p_2 < \cdots < p_k \leqq \beta$ and the rank m is such that $1 \leqq m < [np_2] + 1$. Define $p_0 = 0$ and $p_{k+1} = 1$. Let $u_i = \ln(1 - p_i)^{-1}$ corresponding to $p_2, \cdots, p_k$ and obtain $u_1 = \ln(1 - m/(n + \frac{1}{2}))^{-1}$ by using the Euler-Maclaurin summation formula as is done in Section 3 [Saleh and Ali (1966)].

Then

$$\ln(1 - (n + \tfrac{1}{2})^{-1})^{-1} \leqq u_1 < \ln(1 - p_2)^{-1}$$

and $u_1, u_2, \cdots, u_k$ satisfy the inequality

$$\ln(1 - (n + \tfrac{1}{2})^{-1})^{-1} \leqq u_1 < \cdots < u_k \leqq \ln(1 - \beta)^{-1},$$

Based on the k order statistics, $x_{(m)}, x_{(n_2)}, \cdots, x_{(n_k)}$, the asymptotic BLUE's of μ and σ and their variances, covariance and generalized variance have the same functional form as obtained in (2.3) through (2.8) but now with the above restrictions on the u's.

For the determination of the optimum order statistics we maximize $L(e^{u_1} - 1)^{-1}$ with respect to $u_1, \cdots, u_k$ over the domain $\ln(1 - (n + \frac{1}{2})^{-1})^{-1} \leqq u_1 < \cdots < u_k \leqq \ln(1 - \beta)^{-1}$. Treating $L(e^{u_1} - 1)^{-1}$ as a continuous function of u_1, we note that the maximum of the function is attained when u_1 assumes its smallest value $\ln(1 - (n + \frac{1}{2})^{-1})^{-1}$. Setting $p_1 = (n + \frac{1}{2})^{-1}$, the order statistic $x_{(1)}$ corresponds to p_1 and must be included as one of the k optimum order statistics.

To determine the remaining $(k - 1)$ optimum order statistics, we maximize

$$e^{-u_1^0}(e^{u_1^0} - 1)^{-1}Q_{k-1}(t_1, \cdots, t_{k-1})$$

where $u_1^0 = \ln(1 - (n + \frac{1}{2})^{-1})^{-1}$ and $Q_{k-1}(t_1, \cdots, t_{k-1})$ is defined over $0 < t_1 < \cdots < t_{k-1} \leqq T$ and T is given by

$$T = \ln\{(2n - 1)/(1 - \beta)(2n + 1)\}.$$

Thus, our problem reduces to maximizing $Q_{k-1}(t_1, t_2, \cdots, t_{k-1})$ with respect to $t_1, \cdots, t_{k-1}$ over the given domain. By Theorem 6.1, the above function Q_{k-1} has a unique maximum inside the domain $0 < t_1 < \cdots < t_{k-1} < \infty$. Let this maximum correspond to the point $(t_1^0, \cdots, t_{k-1}^0)$. For the maximum of Q_{k-1} over the domain $0 < t_1 < \cdots < t_{k-1} \leqq T$, we have to consider the following two cases: (i) The proportion of censoring $1 - \beta$ is such that $t_{k-1}^0 \leqq T$; (ii) $1 - \beta$ is such that $t_{k-1}^0 > T$.

In the first case, we take the maximum corresponding to $(t_1^0, \cdots, t_{k-1}^0)$, and the $(k - 1)$ optimum order statistics are the same as those in the uncensored sample. Thus, the results of Saleh and Ali (1966) should be used.

In the second case, by Theorem 6.3, the function $Q_{k-1}(t_1, \cdots, t_{k-1})$ has its maximum on the boundary $t_{k-1} = T$. The remaining $(k - 2)$ coordinates are uniquely determined by the solution of the system of equations

$$\tau_{i+1} + \tau_i - 2t_i = 0, \qquad i = 1, \cdots, (k-1),$$

with $t_{k-1} = T$. Let $t_1^*, \cdots, t_{k-2}^*$ be the solution of the above system of equations· Then setting

$$\lambda_{i*} = 1 - e^{t_i^*}, \qquad i = 1, \cdots, (k-2),$$

and $\lambda_{k-1}^* = 1 - e^{-T}$, we obtain the optimum values of $p_2, \cdots, p_k$ by the relations

$$p_{i+1}^* = (2 + (2\text{n} - 1)\lambda_i^*)/(2n + 1), \qquad i = 1, \cdots, (k-2),$$

and $p_k^* = \beta$. The ranks of the optimum order statistics are uniquely determined by using the relations

$$n_i^* = [np_i^*] + 1, \qquad \text{i} = 2, \cdots, (k-1),$$

and

$$n_k^* = [n\beta] + 1 = r_2 .$$

Thus, the relevant sample for the estimation of μ and σ consists of the optimum order statistics

$$x_{(1)}, x_{(n_2^*)}, \cdots, x_{(n_{k-1}^*)} \quad \text{and} \quad x_{(r_2)} .$$

The estimates of μ and σ based on the optimum order statistics and the coefficients based on $t_1^*, \cdots, t_{k-1}^*$ and T are given by

$$\hat{\mu} = x_{(1)} - \hat{\sigma} \ln [(2n + 1)/(2n - 1)];$$

$$\hat{\sigma} = b_1^* x_{(1)} + \textstyle\sum_{i=2}^{k-1} b_i^* x_{(n_i^*)} + b_k^* x_{(r_2)} ,$$

where

$$b_1^* = -(Q_{k-1}^*)^{-1}(t_1^*/(e^{t_1^*} - 1));$$

$$b_i^* = (Q_{k-1}^*)^{-1}\{(t_{i-1}^* - t_{i-2}^*)/(e^{t_{i-1}^*} - e^{t_{i-2}^*}) - (t_i^* - t_{i-1}^*)/(e^{t_i^*} - e^{t_{i-1}^*})\}, \quad i = 2, \cdots, (k-1);$$

$$b_k^* = (Q_{k-1}^*)^{-1}((T - t_{k-1}^*)/(e^T - e^{t_{k-1}^*}));$$

and Q^*_{k-1} is the maximum value of $Q_{k-1}(t_1, \cdots, t_{k-1})$ over the domain $0 < t_1 < \cdots < t_{k-1} \leqq T$.

The asymptotic joint efficiency (JAE) and the asymptotic relative efficiencies compared to the best linear estimate using all the observations in the censored sample are given by

$$\begin{aligned} \text{JAE}\ (\hat{\mu}, \hat{\sigma}) &= [(2n-1)^2/2(2n+1)(n-1)]\cdot Q^*_{k-1}/\beta; \\ \text{ARE}\ (\hat{\sigma}) &= [(2n-1)/(2n+1)]\cdot Q^*_{k-1}/\beta; \\ \text{ARE}\ (\hat{\mu}) &= (Q^*_{k-1}/\beta)\{[(2n+1)\ln^2((2n+1)/(2n-1)) + 2\beta] \\ &\qquad [(2n+1)\ln^2((2n+1)/(2n-1)) + 2Q^*_{k-1}]^{-1}\}. \end{aligned}$$

The following example illustrates the above estimation procedure:

EXAMPLE. Assume $k = 4, n = 72, 1 - \beta = .1479$ and $r_2 = 62$. First we choose $x_{(1)}$. To choose the remaining three order statistics, we compute $T = \ln (1 - \beta)^{-1} - \ln (2n + 1)/(2n - 1) = 1.8971$. From "Table 3 (1963)" we observe that this situation belongs to case (ii) and we determine t_1^* and t_2^* by solving the system of equations discussed above with $t_3^* = T$. The solutions are (using Table I)

$$t_1^* = .5084 \quad \text{and} \quad t_2^* = 1.1219$$

whence

$$\lambda_1^* = .4217 \quad \text{and} \quad \lambda_2^* = .7334,$$

and we obtain $p_2^* = .4292$, $p_3^* = .7302$ and $p_4^* = .8521$. The corresponding ranks of the optimum order statistics are $n_2^* = 31$, $n_3^* = 53$ and $n_4^* = 62$. Thus, the relevant sample for the estimation of μ and σ consists of the optimum order statistics $x_{(1)}$, $x_{(31)}$, $x_{(53)}$, and $x_{(62)}$. The BLUE's are given by

$$\begin{aligned} \hat{\mu} &= x_{(1)} - \hat{\sigma} \ln (145/143), \\ \hat{\sigma} &= -.9287x_{(1)} + .4021x_{(31)} + .3669x_{(53)} + .2614x_{(62)}. \end{aligned}$$

The asymptotic efficiencies are JAE $(\hat{\mu}, \hat{\sigma}) = 97.6\%$, ARE $(\hat{\sigma}) = 98.0\%$ and ARE $(\hat{\mu}) = 99.3\%$.

5. Determination of the optimum order statistics for the estimation of the parameters from doubly censored samples. For a given $k(< r_2 - r_1 + 1)$ consider the order statistics $x_{(n_1)}, \cdots, x_{(n_k)}$ with ranks $n_1, \cdots, n_k$ determined by k fixed real numbers $p_1, \cdots, p_k$ satisfying the order relation $\alpha \leqq p_1 < \cdots < p_k \leqq \beta$, and $n_i = [np_i] + 1$, $i = 1, \cdots, k$. Define $p_0 = 0$ and $p_{k+1} = 1$.

5.1. *Two-parameter problem.* For the estimation of the location and scale parameters μ and σ, we define $u_i = \ln (1 - p_i)^{-1}$, $i = 1 \cdots, k$, corresponding to $p_1, \cdots, p_k$. Then $u_1, \cdots, u_k$ satisfy the inequality $\ln (1 - \alpha)^{-1} \leqq u_1 < \cdots < u_k \leqq \ln (1 - \beta)^{-1}$. The asymptotic BLUE's of μ and σ based on the k order statistics $x_{(n_1)}, \cdots, x_{(n_k)}$ and their variances, covariance and the generalized variance are given by (2.3) through (2.8) with the above restriction on the u's.

In order to determine the optimum order statistics, we have to minimize the generalized variance given in (2.8), or equivalently, maximize $L(e^{u_1} - 1)^{-1}$ with respect to $u_1, \cdots, u_k$ over the domain $\ln(1-\alpha)^{-1} \leqq u_1 < \cdots < u_k \leqq \ln(1-\beta)^{-1}$. It is known (Section 3) that the maximum of the function $L(e^{u_1} - 1)^{-1}$ is attained when u_1 assumes its smallest value $\ln(1-\alpha)^{-1}$. Therefore, the optimum value of p_1 is α and the corresponding optimum order statistic is $x_{(r_1)}$ since $r_1 = [n\alpha] + 1$. Thus, $x_{(r_1)}$ is one of the optimum order statistics in the relevant sample.

To determine the remaining $(k-1)$ optimum order statistics, we use the transformation

$$t_{i-1} = u_i - u_1^0 \qquad i = 2, \cdots, k,$$

where $u_1^0 = \ln(1-\alpha)^{-1}$. Then the function $L(e^{u_1} - 1)^{-1}$ becomes $e^{-u_1^0}(e^{u_1^0} - 1)^{-1}Q_{k-1}(t_1 \cdots t_{k-1})$, where $Q_{k-1}(t_1, \cdots, t_{k-1})$ is defined, over the domain $0 < t_1 < \cdots < t_{k-1} \leqq T$ and T is given by $T = \ln[(1-\alpha)/(1-\beta)]$. Thus, our problem reduces to maximizing $Q_{k-1}(t_1, \cdots, t_{k-1})$ with respect to $t_1, \cdots, t_{k-1}$ over the given domain. By Theorem 6.1, the above function Q_{k-1} has a unique maximum inside the domain $0 < t_1 < \cdots < t_{k-1} < \infty$. Let this maximum correspond to the point $(t_1^0, \cdots, t_{k-1}^0)$. For the maximum of Q_{k-1} over the domain $0 < t_1 < \cdots < t_{k-1} \leqq T$, we have to consider the following two cases: (i) The quantity T is such that $t_{k-1}^0 \leqq T$; (ii) T is such that $t_{k-1}^0 > T$. In the first case, we take the maximum corresponding to $(t_1^0, \cdots, t_{k-1}^0)$, and the $(k-1)$ optimum order statistics are the same as those obtained in the left-censored sample with fixed α. We may thus use the results of Section 3.

In the second case, by Theorem 6.3, the function $Q_{k-1}(t_1, \cdots, t_{k-1})$ has its maximum on the boundary $t_{k-1} = T$. The remaining $(k-2)$ coordinates are uniquely determined by the solution of the system of equations

$$\tau_{i+1} + \tau_i - 2t_i = 0, \qquad i = 1, \cdots, (k-1),$$

with $t_{k-1} = T$. Let $(t_1^*, \cdots, t_{k-2}^*)$ be the solution of the above system of equations. Then setting $\lambda_i^* = 1 - e^{-t_i^*}$, $i = 1, \cdots, (k-2)$, and $\lambda_{k-1}^* = 1 - e^{-T}$, we obtain the optimum values for $p_2, \cdots, p_k$ by the relations

$$p_{i+1}^* = \alpha + (1-\alpha)\lambda_i^*, \qquad i = 1, \cdots, (k-2);$$
$$p_k^* = \beta.$$

The ranks of the optimum order statistics are obtained by using the relations $n_i^* = [np_i^*] + 1$, $i = 2, \cdots, (k-1)$, and $n_k^* = [n\beta] + 1 = r_2$. Thus relevant sample for the estimation of μ and σ consists of the order statistics $x_{(r_1)}, x_{(n_2^*)}, \cdots, x_{(n_{k-1}^*)}$ and $x_{(r_2)}$. The estimates of μ and σ based on the optimum order statistics and the coefficients based on $t_1^*, \cdots, t_{k-1}^*$ and T are given by

$$\hat{\mu} = x_{(r_1)} - \hat{\sigma}\ln(1-\alpha)^{-1};$$
$$\hat{\sigma} = b_1^* x_{(r_1)} + \textstyle\sum_{i=2}^{k-1} b_i^* x_{(n_i^*)} + b_k^* x_{(r_2)},$$

where

$b_1^* = -(Q_{k-1}^*)^{-1}(t_1^*/(e^{t_1^*} - 1));$

$b_i^* = (Q_{k-1}^*)^{-1}\{(t_{i-1}^* - t_{i-2}^*)/(e^{t_{i-1}^*} - e^{t_{i-2}^*})$

$\qquad - (t_i^* - t_{i-1}^*)/(e^{t_i^*} - e^{t_{i-1}^*})\}, \quad i = 2, \cdots, (k-1);$

$b_k^* = (Q_{k-1}^*)^{-1}\{(T - t_{k-1}^*)/(e^T - e^{t_{k-1}^*})\};$

and Q_{k-1}^* is the maximum value of $Q_{k-1}(t_1, \cdots, t_{k-1})$ over the domain $0 < t_1 < \cdots < t_{k-1} \leqq T$. The joint asymptotic efficiency (JAE) and the asymptotic relative efficiencies (ARE) compared to the best linear estimate using all the observations in the censored sample are given by

$$\begin{aligned} \text{JAE}(\hat{\mu}, \hat{\sigma}) &= [(1-\alpha)/(\beta-\alpha)]\cdot Q_{k-1}^*; \\ \text{ARE}(\hat{\sigma}) &= [(1-\alpha)/(\beta-\alpha)]Q_{k-1}^*; \\ \text{ARE}(\hat{\mu}) &= \{[\alpha(\beta-\alpha) + (1-\alpha)\ln^2(1-\alpha)^{-1}]/ \\ &\qquad (\beta-\alpha)[\alpha Q_{k-1}^* + \ln^2(1-\alpha)^{-1}]\}Q_{k-1}^*. \end{aligned}$$

The values of $t_1^*, \cdots, t_{k-1}^*, \lambda_1^*, \cdots, \lambda_{k-1}^*, b_1^*, \cdots, b_k^*$ and Q_{k-1}^* have been computed for $k = 2(1)4$ and $\alpha = 1 - \beta = .05(.05).25$, and are presented in Table II. Using this table, we illustrate the above estimation procedure by the following example:

EXAMPLE. Assume $k = 5, n = 62, \alpha = 1 - \beta = .10, r_1 = 7$ and $r_2 = 56$. First we choose $x_{(7)}$. To determine the remaining four optimum order statistics, we first compute $T = \ln[(1-\alpha)/(1-\beta)] = 1.7746$. From Table I, we obtain for $k = 4$,

$$t_1^0 = .6003, \qquad t_2^0 = 1.3544, \qquad t_3^0 = 2.3721, \qquad t_4^0 = 3.9657$$

and observe $T = 1.7746 < t_4^0 = 3.9657$. Hence, we are concerned with case (ii) discussed above. We take $x_{(56)}$ and determine the remaining three order statistics by consulting Table II whence $\lambda_1^* = .2987$, $\lambda_2^* = .5307$, $\lambda_3^* = .7056$ and we obtain $p_2^* = .3488$, $p_3^* = .5779$, $p_4^* = .7350$. The corresponding ranks of the order statistics are $n_2^* = 22$, $n_3^* = 36$, $n_4^* = 46$. Thus the optimum order statistics are $x_{(7)}, x_{(22)}, x_{(36)}, x_{(46)}$ and $x_{(56)}$. The BLUE's are

$$\hat{\mu} = x_{(7)} - \hat{\sigma}\ln(.90)^{-1},$$

$$\hat{\sigma} = -1.0185x_{(7)} + .3221x_{(22)} + .2462x_{(36)} + .1805x_{(46)} + .2697x_{(56)}.$$

The asymptotic efficiencies are: JAE $(\hat{\mu}, \hat{\sigma}) = 91.88\%$, ARE $(\hat{\sigma}) = 91.88\%$ and ARE $(\hat{\mu}) = 91.11\%$.

5.2. *One-parameter problem.* For the estimation of the scale parameter σ, we define

$$u_i = \ln(1-p_i)^{-1}, \qquad i = 1, \cdots, k,$$

corresponding to $p_1, \cdots, p_k$. Then $u_1, \cdots, u_k$ satisfy the inequality $\ln(1-\alpha)^{-1} \leqq u_1 < \cdots < u_k \leqq \ln(1-\beta)^{-1}$. The asymptotic BLUE of σ based

TABLE II

The optimum values of $\lambda_1, \cdots, \lambda_{k-1}$ for the linear estimates of μ and σ based on k optimum order statistics with their coefficients in doubly censored sample for $k = 2(1)4$ and for $\alpha = .10(.05).25 = 1 - \beta$

Per Cent Censored	10			20			30			40			50		
K	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4
t_1	.8952	.5680	.4162	.7551	.4825	.3547	.7409	.4738	.3484	.6127	.3946	.2911	.4907	.3182	.2354
t_2	2.1972	1.2713	.9001	1.7746	1.0586	.7572	1.7342	1.0375	.7429	1.3863	.8494	.6135	1.0786	.6743	.4910
t_3		2.1972	1.4782		1.7746	1.2227		1.7342	1.1975		1.3863	.9750		1.0786	.7704
t_4			2.1972			1.7746			1.7342			1.3863			1.0786
Q_k^0	.8122	.8543	.8693	.7812	.8083	.8179	.7766	.8025	.8116	.7218	.7374	.7429	.6445	.6530	.6561
λ_1	.5817	.4333	.3405	.5301	.3828	.2987	.5234	.3774	.2942	.4582	.3261	.2526	.3819	.2725	.2097
λ_2	.8889	.7196	.5935	.8304	.6531	.5307	.8235	.6457	.5243	.7501	.5723	.4585	.6599	.4905	.3880
λ_3		.8889	.7720		.8340	.7056		.8235	.6981		.7501	.6228		.6599	.5372
λ_4			.8889			.8304			.8235			.7501			.6599
b_1	−.7612	−.8693	−.7482	−.8571	−.9625	−1.0184	−.8690	−.9140	−1.0298	−1.0040	−1.1061	−1.1601	−1.2018	−1.3004	−1.3518
b_2	.5166	.4122	.3376	.5109	.3979	.3221	.5104	.3965	.3205	.5063	.3843	.3074	.5040	.3732	.2947
b_3	.2446	.2577	.2445	.3462	.2709	.2462	.3586	.2721	.2464	.4974	.2837	.2476	.6978	.2942	.2485
b_4		.1994	.1661		.2937	.1804		.3054	.1819		.4381	.1946		.6330	.2063
b_5			.1792			.2697			.2810			.4105			.6023

on the k order statistics $x_{(n_1)}, \cdots, x_{(n_k)}$ and its variance are given by (2.9) and (2.10), respectively with the above restrictions on the u's.

In order to determine the optimum order statistics, we have to minimize the variance of the estimate given in (2.10), or equivalently, maximize $Q_k(u_1, \cdots, u_k)$ with respect to $u_1, \cdots, u_k$ where

$$Q_k(u_1, \cdots, u_k) = \sum_{i=1}^{k} (u_i - u_{i-1})^2/(e^{u_i} - e^{u_{i-1}})$$

with $u_0 = 0$ and $c_1 = \ln(1-\alpha)^{-1} \leqq u_1 < \cdots < u_k \leqq \ln(1-\beta)^{-1} = c_2$.

By Theorem 6.1, the above function Q_k has a unique maximum inside the domain $0 < u_1 < \cdots < u_k < \infty$. Let this maximum correspond to the point $(u_1^0, \cdots, u_k^0)$. If c_1 and c_2 are such that $c_1 \leqq u_1^0$ and $c_2 \geqq u_k^0$ hold simultaneously, then we take the point $(u_1^0, \cdots, u_k^0)$ to which corresponds the maximum Q_k over the domain $c_1 \leqq u_1 < \cdots < u_k \leqq c_2$, and the optimum order statistics are the same as those in the uncensored sample.

If c_1 and c_2 are such that $c_1 \leqq u_1^0$ and $c_2 \geqq u_k^0$ are not satisfied simultaneously, then the maximum of Q_k occurs on the boundary of the domain $c_1 \leqq u_1 < \cdots < u_k \leqq c_2$ in the following three possible combinations:

(i) $u_1 = c_1$ and $u_k = c_2$;

(ii) $u_1 = c_1$ and $u_k < c_2$;

(iii) $u_1 > c_1$ and $u_k = c_2$.

In the first case, in order to determine the remaining $(k-2)$ coordinates, we proceed as follows: We make the transformation

$$t_{i-1} = u_i - u_1, \qquad i = 1, \cdots, k, \tag{5.2.1}$$

with $u_1 = c_1$ whence $Q_k(u_1, \cdots, u_k)$ becomes

$$c_1^2/(e^{c_1} - 1) + e^{-c_1} Q_{k-1}(t_1, \cdots, t_{k-1}),$$

where

$$Q_{k-1}(t_1, \cdots, t_{k-1}) = \sum_{i=1}^{k-1} (t_i - t_{i-1})^2/(e^{t_i} - e^{t_{i-1}}) \tag{5.2.2}$$

with $t_0 = 0$ and $0 < t_1 < \cdots < t_{k-1} \leqq c_2 - c_1 = T$ (say). Now taking $t_{k-1} = T$, we maximize $Q_{k-1}(t_1, \cdots, t_{k-1})$. By Theorem 6.3, the maximum corresponds to the solution of the system of equations

$$\tau_{i+1} + \tau_i - 2t_i = 0, \qquad i = 1, \cdots, (k-1),$$

with $t_{k-1} = T$, and is unique. The determination of the optimum order statistics is the same as in case (ii) of Section 5.1. The optimum order statistics are $x_{(r_1)}, x_{(n_1^*)}, \cdots, x_{(n_{k-1}^*)}, x_{(r_2)}$. The estimate of σ is given by

$$\hat{\sigma} = b_1^* x_{(r_1)} + \sum_{i=2}^{k-1} b_i^* x_{(n_i)} + b_k^* x_{(r_2)},$$

where

$$b_1^* = (Q_{k-1}^*)^{-1}\{\ln(1-\alpha)^{-1}/\alpha - t_1^*/(e^{t_1^*}-1)\};$$

$$b_i^* = (Q_{k-1}^*)^{-1}\{(t_{i-1}^* - t_{i-2}^*)/(e^{t_{i-1}^*} - e^{t_{i-2}^*}) - (t_i^* - t_{i-1}^*)/(e^{t_i^*} - e^{t_{i-1}^*})\}, \qquad i = 2, \cdots, (k-1);$$

$$b_k^* = (Q_{k-1}^*)^{-1}\{(T - t_{k-2}^*)/(e^T - e^{t_{k-2}^*})\};$$

and Q_{k-1}^* is the maximum value of $Q_{k-1}(t_1, \cdots, t_{k-1})$ with $T = t_{k-1} = \ln((1-\alpha)/(1-\beta))$. The asymptotic relative efficiency (ARE) of the estimate is

$$\text{ARE}(\hat{\sigma}) = (1-\alpha)\{\alpha Q_{k-1}^* + \ln(1-\alpha)^{-1}\} \{(\beta-\alpha)\alpha + (1-\alpha)\ln^2(1-\alpha)^{-1}\}^{-1}.$$

In case (ii), by applying the transformation (5.2.1), the problem reduces to maximizing Q_{k-1} given in (5.2.2) over the domain $0 \leqq t_1 \leqq \cdots \leqq t_{k-1} \leqq T$. By Theorem 6.1, Q_{k-1} has a unique maximum inside the domain $0 \leqq t_1 \leqq \cdots \leqq t_{k-1} \leqq \infty$. Let this maximum correspond to the point $(t_1^0, \cdots, t_{k-1}^0)$. Now if $t_{k-1}^0 < T$, we take the maximum corresponding to the point $(t_1^0, \cdots, t_{k-1}^0)$ and the desired solution is obtained. The determination of the optimum order statistics is the same as in the case of left censored sample dealt with in Section 3.2. If $t_{k-1}^0 = T$, again we take the maximum corresponding to $(t_1^0, \cdots, t_{k-1}^0)$ and proceed as in Section 3.2. If $t_{k-1}^0 \leqq T$, the problem reduces to case (ii) in Section (5.1).

In case (iii), one possible way is to proceed as follows: We maximize $Q_k(u_1, \cdots, u_k)$ for fixed $u_k = c_2$ over the domain $0 < u_1 < \cdots < u_k = c_2$ by solving the system of equations

$$\tau_{i+1} + \tau_i - 2u_i = 0, \qquad i = 1, \cdots, k,$$

with $u_k = c_2$. Let this solution be $(u_1^*, \cdots, u_{k-1}^*)$. Now if, $u_1^* > c_1$, the desired solution is obtained. The determination of the optimum order statistics is the same as in the right censored problem dealt with in Section 4.1. If $u_1^* = c_1$, again we proceed as in Section 4.1. If $u_1^* < c_1^*$, we reduce the problem to case (i) above, i.e., we preassign $u_1 = c_1$ and $u_k = c_k$. It is expected that in doing so, the loss of efficiency will be negligible.

6. Some extremal problems. In the previous sections, we have observed that for a given k, the determination of the unique set of optimum order statistics for the estimation of the parameters of the exponential distribution depends on the uniqueness of the maximum of the function Q_{k-1} which has been studied by Saleh and Ali (1966). Here we present an alternative proof for the uniqueness of the maximum of Q_{k-1} given by

$$(6.1) \qquad Q_{k-1} = \textstyle\sum_{i=1}^k (t_i e^{-t_i} - t_{i-1}e^{-t_{i-1}})^2/(e^{-t_{i-1}} - e^{-t_i})$$

with $t_0 = 0$ and $t_k e^{-t_k} = 0$. In order to obtain the maximum of Q_{k-1} we have to show that the following system of equations

(6.2) $$\tau_{i+1} + \tau_i - 2t_i = 0, \qquad i = 1, \cdots, (k-1),$$

where

(6.3) $$1 - \tau_i = (t_i e^{-t_i} - t_{i-1}e^{-t_{i-1}})/(e^{-t_i} - e^{-t_{i-1}})$$

has a unique solution. That (6.2) corresponds to the maximum of Q_{k-1} has been shown in Saleh and Ali (1966). To establish the uniqueness of the solution we need the following lemma.

LEMMA 6.1. *Equal probability spacings,* i.e., t_{i-1}, t_i *and* t'_{i+1} *are such that* $F(t_i) - F(t_{i-1}) = F(t'_{i+1}) - F(t_i)$, $i = 1, \cdots, (k-1)$, *where* $F(t) = 1 - e^{-t}$ *do not satisfy the equations in* (6.2).

This implies for fixed t_{i-1} and t_i the value computed according to (6.2) must be less than t'_{i+1}. For proof the readers are referred to Saleh (1964).

THEOREM 6.1. *The system of equations*

$$\tau_{i+1} + \tau_i - 2t_i = 0, \qquad i = 1, \cdots, (k-1),$$

has one and only one solution and this solution corresponds to the maximum of Q_{k-1}.

PROOF. In order to prove the theorem, we study the effect of displacement in any t on the succeeding ones satisfying the given equations:

Consider a fixed value of $t_1 > 0$. This determines $k - 2$ succeeding t's, t_2, $t_3, \cdots, t_{k-1}$ computed in accordance with $2t_i = \tau_i + \tau_{i+1}$. Taking differentials on both sides, we obtain

(6.4) $$[e^{-t_{i-1}}(t_{i+1} - \tau_{i+1})/(e^{-t_i} - e^{-t_{i+1}})]\, d(t_{i+1} - t_i)$$
$$= [e^{-t_{i+1}}(\tau_i - t_{i-1})/(e^{-t_{i-1}} - e^{-t_i})]\, d(t_i - t_{i-1}),$$

where d is the differential operator. From this differential equation it follows that if $d(t_i - t_{i-1}) > 0$, then $d(t_{i+1} - t_i) > 0$ since all other terms are positive. Now, if we can prove that for a positive increment in t_1, $d(t_2 - t_1) > 0$, then by induction, we can prove that an infinitesimal increment in t_1 increases the values of all the succeeding t's.

First, we show that $dt_1 > 0$ implies $d(t_2 - t_1) > 0$. Consider $2t_1 = \tau_1 + \tau_2$. Taking differentials, we obtain

$$[e^{-t_2}(t_2 - \tau_2)/(e^{-t_1} - e^{-t_2})]\, d(t_2 - t_1) = [\tau_1/(1 - e^{-t_1})^{-1}]\, dt_1$$

or

$$dt_2/dt_1 = 1 + \tau_1(e^{-t_1} - e^{-t_2})/(1 - e^{-t_1})e^{-t_2}.$$

Since the second term on the right is positive, it follows that t_2 increases with t_1.

Now we show that an increase in t_1 increases all the succeeding t's and at the same time satisfies the equation $2t_i = \tau_i + \tau_{i+1}$. From the differential equation (6.4), we obtain

$$d(t_{i+1} - t_i)/d(t_i - t_{i-1}) = e^{-t_{i-1}}(\tau_i - t_{i-1})(e^{-t_i} - e^{-t_{i+1}})$$
$$[e^{-t_{i+1}}(t_{i+1} - \tau_{i+1})(e^{-t_{i-1}} - e^{-t_i})]^{-1} > 0, \qquad i = 1, \cdots, (k-1),$$

which implies

$$d(t_{i+1} - t_i)/d(t_2 - t_1) = A_i \text{ (say)} > 0$$

so that $d(t_{i+1} - t_i) = A_i\, d(t_2 - t_1)$ and we obtain

$$dt_{i+1}/dt_1 > 0, \qquad i = 1, \cdots, (k-1).$$

This proves the statement.

Therefore, assigning a small positive value to t_1 and assuming that $k - 2$ points $t_2, \cdots, t_{k-1}$ can be found satisfying $2t_i = \tau_i + \tau_{i+1}$, we increase t_1 continuously until t_{k-1} satisfies the condition $\tau_k = 1 + t_{k-1}$. By monotonicity of the t's, the solution will be unique.

Now, we prove the existence of the solution in the following manner:

We divide the interval [0, 1] into k equal parts each that

$$1 - e^{-t_1^0} = e^{-t_1^0} - e^{-t_2^0} = \cdots = e^{-t_{k-3}^0} - e^{-t_{k-2}^0} = e^{-t_{k-1}^0}.$$

First choose t_1^0 as a trial value of t_1 and generate t_2^* from the relation $2t_1 = \tau_1 + \tau_2$. Similarly generate $t_3^*, \cdots, t_{k-1}^*$. By Lemma 6.1 $t_2^* < t_2^0; \cdots; t_{k-1}^* < t_{k-1}^0$. Therefore, increasing t_1^0 continuously whereby $t_2^*, \cdots, t_{k-1}^*$ increase, we finally obtain $\tau_k^* = 1 + t_{k-1}^*$ and the solution is unique. This proves that the given system of equations has one and only one solution. By Lemma 4.6 [Saleh and Ali, (1966)], this solution corresponds to the maximum of Q_{k-1}. Hence Q_{k-1} has a unique maximum inside the domain $0 \leqq t_1 \leqq \cdots \leqq t_{k-1} \leqq \infty$. This completes the proof of the theorem. For a geometrical interpretation of the nature of solution we refer the readers to Saleh (1964).

The following theorem helps in finding the set of solutions $(t_1^0, \cdots, t_k^0)$ for which the k-dimensional function $Q_k(t_1, \cdots, t_k)$ is maximum from a knowledge of the set of solution $(t_1^*, \cdots, t_{k-1}^*)$ for which the $(k-1)$-dimensional function $Q_{k-1}(t_1, \cdots, t_{k-1})$ is maximum.

THEOREM 6.2. *If the $(k-1)$-dimensional function $Q_{k-1}(t_1', \cdots, t_{k-1}')$ defined by*

$$Q_{k-1}(t_1', \cdots, t_{k-1}') = \textstyle\sum_{i=1}^{k-1} (t_i' - t_{i-1}')^2/(e^{t_i'} - e^{t_{i-1}'})$$

with $t_0' = 0$ and $0 < t_1' < \cdots < t_{k-1}' < \infty$ attains its unique maximum at $(t_1^*, \cdots, t_{k-1}^*)$, then the k-dimensional function $Q_k(t_1, \cdots, t_k)$ defined by

$$Q_k(t_1, \cdots, t_k) = \textstyle\sum_{i=1}^{k} (t_i - t_{i-1})^2/(e^{t_i} - e^{-t_{i-1}})$$

with $t_0 = 0$ and $0 < t_1 < \cdots < t_k < \infty$ has its unique maximum at $(t_1^0, \cdots, t_k^0)$ where

$$t_{i+1}^0 = t_i^* + t_i^0, \qquad i = 1, \cdots, (k-1),$$

and t_1^0 is given by the root of the equation

$$t_1/(1 - e^{-t_1}) = 1 + (1 - Q_{k-1}^0)^{\frac{1}{2}}, \tag{6.5}$$

Q_{k-1}^0 denoting the maximum value of $Q_{k-1}(t_1', \cdots, t_{k-1}')$ at $(t_1^*, \cdots, t_{k-1}^*)$.

For proof see Saleh (1964).

COROLLARY 6.2.1. *If $t_1^0 < T \leqq t_1$, then $Q_k(t_1)$ is a monotonically decreasing function of t_1 and attains its maximum at $t_1 = T$.*

PROOF. Consider the derivative

$$dQ_k(t_1)/dt_1 = e^{-t_1}\{(1 - Q_{k-1}^0) - [t_1/(1 - e^{-t_1})^{-1} - 1]^2\}.$$

Now, the derivative is zero at $t_1 = t_1^0$ by (6.5). But since $t_1/(1 - e^{-t_1})$ is an increasing function of t_1, the derivative is negative when t_1^0 for $t_1 \geqq T$. This proves the corollary.

THEOREM 6.3. (i) *The function Q_k defined over the domain $0 \leqq t_1 \leqq \cdots \leqq t_k \leqq T < t_k^0$ attains its maximum on the boundary $t_k = T$.*

(ii) *The maximum of Q_k is unique and corresponds to the solution of the system of equations*

$$\tau_{i+1} + \tau_i - 2t_i = 0, \qquad i = 1, \cdots, (k - 1), \tag{6.6}$$

with $t_k = T$ where τ_i is defined by (6.3).

PROOF. By Theorem 6.1, the function Q_k defined over the domain $0 \leqq t_1 \leqq \cdots \leqq t_k \leqq \infty$ has unique maximum inside the domain, and let this maximum be attained at the point $(t_1^0, \cdots, t_k^0)$ which satisfies the inequality $0 < t_1^0 < \cdots < t_k^0 < \infty$. If T is such that $T < t_k^0$, then Q_k cannot have a maximum inside the domain $0 \leqq t_1 \leqq \cdots \leqq t_k \leqq T$. Therefore, the maximum of Q_k occurs on the boundary. Proceeding as in Lemma 4.3 [Saleh and Ali, (1966)], we see that if $t_1 \to 0$ or $t_i \to t_{i+1}$, the maximum does not occur. Therefore the maximum of Q_k occurs when $t_k = T$.

Now, for the maximum of $Q_k(t_1, \cdots, t_{k-1}, t_k)$, we solve the following system of equations:

$$(\partial/\partial t_i)Q_k(t_1, \cdots, t_{k-1}, t_k) = 0, \quad i = 1, \cdots, (k - 1),$$

with $t_k = T$. Proceeding along similar lines as in the proof of Lemma 4.6 [Saleh and Ali (1966)], we see that the maximum corresponds to the solution of the system of equations

$$\tau_{i+1} + \tau_i - 2t_i = 0, \qquad i = 1, \cdots, k,$$

with $t_k = T$. We show now that this system of equations has unique solution.

Consider a small fixed value of $t_1 > 0$. This determines the $k - 1$ succeeding t's, $t_2 < \cdots < t_{k-1} < t_k < T$, according to the relation

$$2t_i = \tau_i + \tau_{i+1}, \qquad i = 1, \cdots, k.$$

Proceeding along similar lines as in the proof of Theorem 6.1, we obtain the same relation as in (6.4) and $dt_{i+1}/dt_1 > 0$, that is to say, if t_1 is increased, all the succeeding t's increase, satisfying the relation

$$2t_i = \tau_i + \tau_{i+1}, \qquad i = 1, \cdots, k.$$

Therefore, increasing t_1 continuously, we can find $(k - 1)$ points $t_2, \cdots, t_k$ satisfying the above relation until $t_k = T$. By monotonicity, this will be true, and the solution of the system of equations (6.6) is unique.

We prove the existence of the solution as follows: We divide the interval $[0, 1 - e^{-T}]$ into k equal parts such that

$$1 - e^{-t_1'} = e^{-t_1'} - e^{-t_2'} = \cdots = e^{-t_{k-1}'} - e^{-T}.$$

Keeping T fixed and choosing t_1' as a trial value, generate t_2^* from the relation $2t_1 = \tau_1 + \tau_2$. Similarly, generate $t_3^*, \cdots, t_k^*$. By Lemma (6.1) $t_2^* < t_2' \cdots t_k^* < T$. Therefore, increasing t_1' continuously whereby $t_2^*, \cdots, t_k^*$ increase, we finally obtain $t_k^* = T$ and the solution is unique. Hence, the system of equations (6.6) has one and only one solution, and the solution corresponds to the unique maximum of Q_k over the domain $0 < t_1 < \cdots < t_k \leqq T < t_k^0$. This completes the proof. For a geometrical interpretation we refer the readers to Saleh (1964).

It is interesting to note that Kulldorff (1958) obtained the system of equations (6.6) in connection with grouping problem in truncated exponential distribution and conjectured that the system has one and only one solution.

Acknowledgment. I wish to express my gratitude to Professor Mir M. Ali for suggesting the problem, his keen interest during the research period and helpful suggestions. I am also grateful to the referee for his comments and suggestions, which were helpful in condensing the original manuscript.

REFERENCES

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

[2] Kulldorff, G. (1963a). On the optimum spacings of the sample quantiles from an exponential distribution. Department of Statistics, University of Lund, Sweden.

[3] Kulldorff, G. (1963b). Estimation of one or two parameters of the exponential distribution on the basis of suitably chosen order statistics. *Ann. Math. Statist.* **34** 1419–1431.

[4] Mosteller, F. (1946). On some useful inefficient statistics. *Ann. Math. Statist.* **17** 377–408.

[5] Ogawa, J. (1951). Contributions to the theory of systematic statistics. *Osaka Math. J.* **3** 175-213.

[6] Ogawa, J. (1960). Determination of optimum spacings for the estimation of the scale parameters of an exponential distribution based on sample quantile. *Ann. Inst. Statist. Math.* **2** 135–141.

[7] Saleh, A. K. Md. Ehsanes (1964). On the estimation of the parameters of the exponential distribution based on optimum order statistics in censored samples. Ph.D. Dissertation. University of Western Ontario, London, Canada.

[8] Saleh, A. K. Md. Ehsanes and Ali, M. M. (1966). Asymptotic optimum quantiles for the estimation of the parameters of the negative exponential distribution. *Ann. Math. Statist.* **37** 143–151.

[9] Sarhan, A. E. and Greenberg, B. (1957). Tables for best linear estimates by order statistics of the parameters of the single exponential distribution from singly and doubly censored samples. *Amer. Statist. Assoc.* **57** 58–87.

[10] Sarhan, A. E., Greenberg, B. and Ogawa (1963). Simplified estimates for the exponential distribution. *Ann. Math. Statist.* **34** 117–121.