

INADMISSIBILITY RESULTS FOR THE SELECTED SCALE PARAMETERS

BY P. VELLAISAMY

Indian Institute of Technology

Let X_1, X_2, \dots, X_k be k independent gamma random variables with different scale parameters but with a common known shape parameter. Suppose the population corresponding to the largest $X_{(1)}$ [or the smallest $X_{(k)}$] observation is selected. The problem of estimating the scale parameter $\theta_{(1)}$ [or $\theta_{(k)}$] of the selected population is considered. We derive, using the method of differential inequalities, explicit estimators that dominate the natural or the existing estimators. The improved estimators of $\theta_{(1)}$ are similar to that of DasGupta estimators for the usual simultaneous estimation problem. An implication of this result for the simultaneous estimation of the selected subset is also considered.

1. Introduction. Estimation of a characteristic of the selected population arises in various practical situations. Suppose, as an example, a doctor has experimented k types of drugs for a particular disease and chooses the best one. He might naturally be interested in obtaining an estimate of the effectiveness of the selected drug for the successful diagnosis of the disease. These types of problems of estimation after selection have been, of late, studied for various probability models. Some of the recent references in this area are Cohen and Sackrowitz (1982, 1989), Sackrowitz and Samuel-Cahn (1984, 1987) and Venter (1988).

Let X_1, X_2, \dots, X_k be k independent random variables, where the density of the X_i is

$$(1.1) \quad f_i(x|\theta_i, p) = \theta_i^{-p} e^{-(x/\theta_i)} x^{p-1} / \Gamma(p).$$

Assume the scale parameters θ_i 's, $0 < \theta_i < \infty$, are unknown and the shape parameter p is known. Let $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(k)}$ denote the order statistics of the X_i 's. We shall call the population with the largest (or the smallest) scale parameter the best population.

Suppose we employ the natural rule, to select the best population, according to which the population corresponding to $X_{(1)}$ [or $X_{(k)}$] is selected. Let $\theta_{(1)}$ [or $\theta_{(k)}$] denote the scale parameter associated with the selected population. Observe that, for example, $\theta_{(1)}$ is a random quantity, and is given by

$$(1.2) \quad \theta_{(1)} = \begin{cases} \theta_i, & \text{if } X_i = X_{(1)}, \\ 0, & \text{otherwise.} \end{cases}$$

Received June 1991; revised March 1992.

AMS 1980 subject classifications. Primary 62C15; secondary 62F10, 62F07.

Key words and phrases. Estimation after selection, gamma scale parameters, inadmissible estimators, differential inequalities, simultaneous estimation after selection.

In this paper we consider the problem of estimating $\theta_{(1)}$ and $\theta_{(k)}$ under the squared error loss (unless stated otherwise) defined, for instance, by

$$(1.3) \quad L(t, \theta_{(1)}) = (t - \theta_{(1)})^2.$$

We first consider the problem of estimating $\theta_{(k)}$. A natural estimator of $\theta_{(k)}$ is $X_{(k)}/p$. Also a natural competitor of $X_{(k)}/p$ is $X_{(k)}/(p+1)$, the analog of the best scale invariant estimator (of θ_1) $X_1/(p+1)$ of the component problem (i.e., the case $k=1$). We have not been able to compare these two estimators analytically. However, we derive a class of estimators for $p > (k-1)^{-1}$ that dominate the natural estimator $X_{(k)}/p$. For the estimation of $\theta_{(1)}$, it is known that the estimator $X_{(1)}/(p+1)$ is better than the natural estimator $X_{(1)}/p$ [Vellaisamy and Sharma (1989)]. In this paper we obtain a class of estimators that are better than $X_{(1)}/(p+1)$, and are similar to DasGupta [(1986) Theorem 4] estimators for the usual simultaneous estimation of gamma scale parameters.

We exploit the technique of constructing improved estimators by solving certain differential inequalities on the sample space. This technique is well known and has been successfully employed by many statisticians, prominently by Berger (1980), Hwang (1982) and DasGupta (1984, 1986), among others, for the simultaneous estimation problems where no selection is involved. Our method of proof is along the lines similar to that of Berger (1980) and that of DasGupta (1986). We also apply the UV method of Robbins (1988) to obtain the unbiased estimators of the risk difference.

2. The main results. We first state a lemma which is essentially Corollary 2 of Berger (1980), and is useful in finding unbiased estimators of certain functions of $\mathbf{X} = (X_1, \dots, X_k)$ and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$.

LEMMA 2.1 [Berger (1980)]. *Suppose X_1, X_2, \dots, X_k are k independent random variables with densities defined in (1.1). Let $u(\mathbf{x})$ be any real valued function defined on R^k such that (i) $E_{\boldsymbol{\theta}}|u(\mathbf{X})| < \infty$ and (ii) the indefinite integral*

$$h_i(\mathbf{x}) = \int_0^{x_i} u(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_k) t^{p-1} dt$$

exists for all $x_i \in R^1$. Then

$$(2.1) \quad v(\mathbf{x}) = x_i^{1-p} h_i(\mathbf{x})$$

satisfies the condition

$$E_{\boldsymbol{\theta}}(v(\mathbf{X})) = \theta_i E_{\boldsymbol{\theta}}(u(\mathbf{X})) \quad \text{for all } \boldsymbol{\theta}.$$

We are now ready to state and prove the main results of the paper.

THEOREM 2.1. *Consider the estimation of $\theta_{(k)}$ under squared error loss. Let $p > (k-1)^{-1}$ and $\varphi: R^{(k-1)} \rightarrow R^1$ be any real valued function satisfying the*

following conditions:

- (i) $0 < \varphi \leq 2(kp - p - 1)/p(p + 1)^2 = M$, say.
- (ii) φ is nonincreasing in each of the $(k - 1)$ arguments.

Then any estimator

$$\delta_\varphi(\mathbf{X}) = \frac{X_{(k)}}{p} \left[1 + p(p + 1 - X_{(k)})\varphi(X_{(1)}, \dots, X_{(k-1)})e^{-X_{(k)}} \right]$$

dominates $\delta(\mathbf{X}) = X_{(k)}/p$.

PROOF. Let $\delta_1(\mathbf{X}) = \delta_1(X_{(1)}, X_{(2)}, \dots, X_{(k)})$ be any estimator of $\theta_{(k)}$. Then the risk of δ_1 under squared error loss is

$$(2.2) \quad R(\delta_1, \theta) = \sum_{i=1}^k E_\theta \left[\delta_1^2(X_{(1)i}, \dots, X_{(k-1)i}, X_i) - 2\delta_1(X_{(1)i}, \dots, X_{(k-1)i}, X_i)\theta_i + \theta_i^2 \right] I(X_i < X_{(k-1)i}),$$

where $X_{(1)i} \geq X_{(2)i} \geq \dots \geq X_{(k-1)i}$ denote the ordered values of $X_1, X_2, \dots, X_{i-1}, X_{i+1}, \dots, X_k$. Applying Lemma 2.1, an unbiased estimator of

$$D_1(\theta) = E_\theta \left[\theta_i \delta_1(X_{(1)i}, \dots, X_{(k-1)i}, X_i) I(X_i < X_{(k-1)i}) \right]$$

can be seen to be

$$(2.3) \quad \eta_1(\mathbf{X}) = X_i^{1-p} \left[g_1(X_{(1)i}, \dots, X_{(k-1)i}, X_i) I(X_i < X_{(k-1)i}) + g_1(X_{(1)i}, \dots, X_{(k-1)i}, X_{(k-1)i}) I(X_{(k-1)i} < X_i) \right],$$

where

$$(2.4) \quad g_1(x_1, x_2, \dots, x_k) = \int_0^{x_k} \delta_1(x_1, x_2, \dots, x_{k-1}, t) t^{p-1} dt.$$

Therefore the risk is equal to

$$(2.5) \quad \begin{aligned} R(\delta_1, \theta) &= \sum_{i=1}^k E_\theta \left[\left\{ \delta_1^2(X_{(1)i}, \dots, X_{(k-1)i}, X_i) - 2X_i^{1-p} g_1(X_{(1)i}, \dots, X_{(k-1)i}, X_i) + \theta_i^2 \right\} I(X_i < X_{(k-1)i}) \right. \\ &\quad \left. - 2X_i^{1-p} g_1(X_{(1)i}, \dots, X_{(k-1)i}, X_{(k-1)i}) I(X_i > X_{(k-1)i}) \right] \\ &= E_\theta \left[\delta_1^2(X_{(1)}, \dots, X_{(k)}) - 2X_{(k)}^{1-p} g_1(X_{(1)}, \dots, X_{(k)}) \right. \\ &\quad \left. - 2G_1(X_{(1)}, \dots, X_{(k)}) + \theta_{(k)}^2 \right], \end{aligned}$$

where

$$G_1(X_{(1)}, \dots, X_{(k)}) = X_{(1)}^{1-p} g_1(X_{(2)}, \dots, X_{(k)}, X_{(k)}) + \sum_{i=2}^{k-1} X_{(i)}^{1-p} g_1(X_{(1)}, \dots, X_{(i-1)}, X_{(i+1)}, \dots, X_{(k)}, X_{(k)}).$$

Using the fact

$$g_1^{k(1)}(x_1, \dots, x_k) = \frac{\partial}{\partial x_k} g_1(x_1, \dots, x_k) = \delta_1(x_1, \dots, x_k) x_k^{p-1},$$

we can write

$$(2.6) \quad R(\delta_1, \theta) = E_{\theta} \left[X_{(k)}^{2(1-p)} \left(g_1^{k(1)}(X_{(1)}, \dots, X_{(k)}) \right)^2 - 2X_{(k)}^{1-p} g_1(X_{(1)}, \dots, X_{(k)}) - 2G_1(X_{(1)}, \dots, X_{(k)}) + \theta_{(k)}^2 \right].$$

Let $\delta_2(\mathbf{X}) = \delta_2(X_{(1)}, \dots, X_{(k)})$ be any other competing estimator of $\theta_{(k)}$ such that $h(x_1, \dots, x_k) = g_2(x_1, \dots, x_k) - g_1(x_1, \dots, x_k)$ is a nonincreasing function in each of first $(k - 1)$ arguments, and $g_2(x_1, \dots, x_k)$ be defined as in (2.4) with δ_1 replaced by δ_2 . Then an unbiased estimator of the risk difference $R(\delta_2, \theta) - R(\delta_1, \theta)$ is given by

$$(2.7) \quad \Delta_1(\mathbf{X}) = X_{(k)}^{2(1-p)} \left\{ \left(g_2^{k(1)}(X_{(1)}, \dots, X_{(k)}) \right)^2 - \left(g_1^{k(1)}(X_{(1)}, \dots, X_{(k)}) \right)^2 \right\} - 2X_{(k)}^{1-p} h(X_{(1)}, \dots, X_{(k)}) - 2\{G_2(X_{(1)}, \dots, X_{(k)}) - G_1(X_{(1)}, \dots, X_{(k)})\}.$$

Note that

$$\begin{aligned} &G_2(X_{(1)}, \dots, X_{(k)}) - G_1(X_{(1)}, \dots, X_{(k)}) \\ &= X_{(1)}^{1-p} h(X_{(2)}, \dots, X_{(k)}, X_{(k)}) \\ &\quad + \sum_{i=2}^{k-1} X_{(i)}^{1-p} h(X_{(1)}, \dots, X_{(i-1)}, X_{(i+1)}, \dots, X_{(k)}, X_{(k)}) \\ &\geq (k - 1) X_{(k)}^{1-p} h(X_{(1)}, \dots, X_{(k-2)}, X_{(k-1)}, X_{(k)}), \end{aligned}$$

since $h(x_1, \dots, x_k)$ is nonincreasing in the first $(k - 1)$ arguments. Therefore,

$$\Delta_1(\mathbf{X}) \leq \Delta_2(\mathbf{X}) = X_{(k)}^{2(1-p)} \left\{ \left(g_2^{k(1)}(X_{(1)}, \dots, X_{(k)}) \right)^2 - \left(g_1^{k(1)}(X_{(1)}, \dots, X_{(k)}) \right)^2 \right\} - 2kX_{(k)}^{(1-p)} h(X_{(1)}, \dots, X_{(k)}).$$

Now let $\delta_1(X_{(1)}, \dots, X_{(k)}) = X_{(k)}/p$. Then

$$h^{k(1)} = g_2^{k(1)} - g_1^{k(1)} = g_2^{k(1)} - x_k^p/p$$

and

$$(g_2^{k(1)})^2 - (g_1^{k(1)})^2 = (h^{k(1)})^2 + 2x_k^p h^{k(1)}/p.$$

Hence,

$$(2.8) \quad \Delta_2(\mathbf{X}) = X_{(k)}^{2(1-p)}(h^{k(1)})^2 + \frac{2}{p}X_{(k)}^{2-p}h^{k(1)} - 2kX_{(k)}^{1-p}h.$$

Therefore, we have to solve the differential inequality

$$(2.9) \quad \Delta_3(\mathbf{x}) < 0 \quad \text{for } x_1 \geq x_2 \geq \dots \geq x_k > 0,$$

where

$$\Delta_3(\mathbf{x}) = x_k^{(1-p)}(h^{k(1)})^2 + \frac{2}{p}x_k h^{k(1)} - 2kh.$$

Now let

$$h(x_1, \dots, x_k) = \varphi(x_1, \dots, x_{k-1})\psi(x_k),$$

where φ satisfies the assumptions of the theorem and $\psi(t)$ is positive and is differentiable at all $t \in (0, \infty)$. Then our problem reduces to that of solving the inequality

$$(2.10) \quad \Delta_4(x_k) = x_k^{1-p}(\psi'(x_k))^2 + a_1x_k\psi'(x_k) - a_2\psi(x_k) < 0,$$

where $a_1 = 2/pM$, $a_2 = 2k/M$, $\psi'(x) = (d/dx)\psi(x)$ and $x_k > 0$. It can be shown that $\psi(x_k) = x_k^{p+1}e^{-x_k}$ is a solution of the above inequality. Hence

$$(2.11) \quad h(x_1, \dots, x_k) = \varphi(x_1, \dots, x_{k-1})x_k^{p+1}e^{-x_k}$$

is a solution of the differential inequality given in (2.9).

Let us now take

$$(2.12) \quad \delta_2(x_1, \dots, x_k) = x_k s(x_1, \dots, x_k)/p,$$

where $s(x_1, \dots, x_k)$ is a suitable function to be determined later. Then, from (2.11) and (2.12),

$$(2.13) \quad \begin{aligned} h^{k(1)}(x_1, \dots, x_k) &= x_k^p [s(x_1, \dots, x_k) - 1]/p. \\ &= \varphi(x_1, \dots, x_{k-1})(p + 1 - x_k)x_k^p e^{-x_k} \end{aligned}$$

and hence,

$$s(x_1, \dots, x_k) = 1 + p\varphi(x_1, \dots, x_{k-1})(p + 1 - x_k)e^{-x_k}.$$

Therefore,

$$\delta_2(\mathbf{X}) = \frac{X_{(k)}}{p} [1 + p(p + 1 - X_{(k)})\varphi(X_{(1)}, \dots, X_{(k-1)})e^{-X_{(k)}}]$$

improves upon $\delta_1(\mathbf{X})$, which proves the theorem. \square

REMARK 2.1. We have been able to find improved estimators when $p > (k - 1)^{-1}$, by solving the differential inequality in a rather crude way. Al-

though our result covers most of the practical cases, it will be interesting to obtain improved estimators for the case $p < (k - 1)^{-1}$ also.

We next consider the estimation of $\theta_{(1)}$. The following theorem, which is similar to Theorem 4 of DasGupta (1986), presents improved estimators over the existing estimator $X_{(1)}/(p + 1)$ [Vellaisamy and Sharma (1989)].

THEOREM 2.2. *Let $Y = (\prod_{i=1}^k X_{(i)})^{1/k}$, and for $t > 0$, $s(t)$ be a real valued nondecreasing function such that (i) $s(t)/t$ is nonincreasing, and (ii) $0 < s(t)/t \leq 2k(k - 1)/(p + k)^2$. Then any estimator of the form*

$$(2.14) \quad \delta_s(X_{(1)}, \dots, X_{(k)}) = \frac{X_{(1)}}{p + 1} + \frac{1}{(p + 1)} \left[ps(Y) + \frac{1}{k} Ys'(Y) \right]$$

dominates $X_{(1)}/(p + 1)$.

PROOF. We present here only the important steps as most part of the proof is similar to that of Theorem 2.1.

Let $\nu_1(\mathbf{X}) = \nu_1(X_{(1)}, \dots, X_{(k)})$ be an arbitrary estimator of $\theta_{(1)}$. Then an unbiased estimator of

$$D_2(\theta) = E_{\theta}[\theta_i \nu_1(X_i, X_{(1)i}, \dots, X_{(k-1)i})]$$

can be seen to be

$$(2.15) \quad \eta_2(\mathbf{X}) = X_i^{(1-p)} q_1(X_i, X_{(1)i}, \dots, X_{(k-1)i}) I(X_i > X_{(1)i}),$$

where

$$q_1(x_1, x_2, \dots, x_k) = \int_{x_2}^{x_1} \nu_1(t, x_2, \dots, x_k) t^{p-1} dt.$$

Let $\nu_2(\mathbf{X}) = \nu_2(X_{(1)}, \dots, X_{(k)})$ be any other estimator of $\theta_{(1)}$, and $q_2(x_1, \dots, x_k)$ be the corresponding function associated with ν_2 . Then an unbiased estimator of the risk difference $R(\nu_2, \theta) - R(\nu_1, \theta)$ is,

$$(2.16) \quad H(X_{(1)}, \dots, X_{(k)}) = X_{(1)}^{2(1-p)} \left[(q_2^{1(1)})^2 - (q_1^{1(1)})^2 \right] - 2X_{(1)}^{1-p} (q_2 - q_1),$$

where for example

$$q_1^{1(1)}(x_1, \dots, x_k) = \frac{\partial}{\partial x_1} q(x_1, \dots, x_k) = x_1^{p-1} \nu_1(x_1, \dots, x_k).$$

Let now $\nu_1(X_{(1)}, \dots, X_{(k)}) = X_{(1)}/(p + 1)$, and

$$\nu_2(X_{(1)}, \dots, X_{(k)}) = \frac{X_{(1)}}{(p + 1)} w(X_{(1)}, \dots, X_{(k)}),$$

where w is some suitable function.

Following Berger (1980), let

$$F(x_1, \dots, x_k) = (p + 1)(q_2(x_1, \dots, x_k) - q_1(x_1, \dots, x_k)),$$

is better than the natural estimator

$$U = \frac{1}{p} \sum_{r=1}^k X^{(r)} I(X_{(r+1)} < cX_{(1)} \leq X_{(r)}),$$

for the estimation of

$$Q = \sum_{r=1}^k \theta^{(r)} I(X_{(r+1)} < cX_{(1)} \leq X_{(r)})$$

under the squared error loss function defined, for example, by

$$L(Q, U) = \sum_{r=1}^k \left[\sum_{j=1}^r \left(\frac{X_{(j)}}{p} - \theta_{(j)} \right)^2 \right] I(X_{(r+1)} < cX_{(1)} \leq X_{(r)}),$$

where $X^{(r)} = (X_{(1)}, \dots, X_{(r)})$, $\theta^{(r)} = (\theta_{(1)}, \dots, \theta_{(r)})$ and $\theta_{(i)}$ is the parameter associated with $X_{(i)}$, $i = 1, \dots, k$.

Let us now define

$$L = \sum_{r=1}^k L^{(r)} I(X_{(r+1)} < cX_{(1)} \leq X_{(r)})$$

and

$$M = \sum_{r=1}^k M^{(r)} I(X_{(r+1)} < cX_{(1)} \leq X_{(r)})$$

to be two estimators of Q , where for $1 \leq r \leq k$,

$$L^{(r)} = (L_1, L_2, \dots, L_r), \quad M^{(r)} = (M_1, L_2, \dots, L_r)$$

with $L_1 = X_{(1)}/(p+1)$, $M_1 = \delta_s(X_{(1)}, \dots, X_{(k)})$ defined in (2.14), and L_2, \dots, L_r are arbitrary estimators. An immediate consequence of Theorem 2.2 is that the estimator M dominates L .

It is now clear that the estimator

$$U_s = \delta_s(X_{(1)}, \dots, X_{(k)}) I(X_{(2)} < cX_{(1)}) + \frac{1}{p} \sum_{r=2}^k X^{(r)} I(X_{(r+1)} < cX_{(1)} \leq X_{(r)})$$

dominates the existing estimator U_1 of Q .

Acknowledgments. The author is grateful to an Associate Editor for pointing out an error in the original manuscript and to the referee for some useful suggestions.

REFERENCES

- BERGER, J. O. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma scale parameters. *Ann. Statist.* **8** 545–571.
- COHEN, A. and SACKROWITZ, H. (1982). Estimating the mean of the selected population. In *Statistical Decision Theory and Related Topics III* (S. S. Gupta and J. O. Berger, eds.) 1 243–270. Academic, New York.

- COHEN, A. and SACKROWITZ, H. (1989). Two stage conditionally unbiased estimators of the selected mean. *Statist. Probab. Lett.* **18** 273–278.
- DASGUPTA, A. (1984). Admissibility in the gamma distribution: Two examples. *Sankhyā Ser. A.* **46** 395–407.
- DASGUPTA, A. (1986). Simultaneous estimation in the multiparameter gamma distribution under weighted quadratic losses. *Ann. Statist.* **14** 206–219.
- HWANG, J. T. (1982). Improving upon standard estimators in discrete exponential families with applications to Poisson and negative binomial cases. *Ann. Statist.* **10** 857–867.
- GUPTA, S. S. (1963). On a selection and ranking procedure for gamma populations. *Ann. Inst. Statist. Math.* **14** 119–216.
- ROBBINS, H. E. (1988). The UV method of estimation. In *Statistical Decision Theory and Related Topics IV* (S. S. Gupta and J. O. Berger, eds.) **1** 265–270. Springer, New York.
- SACKROWITZ, H. and SAMUEL-CAHN, E. (1984). Estimation of the mean of a selected negative exponential population. *J. Roy. Statist. Soc. Ser. B* **46** 242–249.
- SACKROWITZ, H. and SAMUEL-CAHN, E. (1987). Evaluating the chosen population: A Bayes and minimax approach. In *Adaptive Statistical Procedures and Related Topics* (J. Van Ryzin, ed.) 386–399. IMS, Hayward, Calif.
- VELLAISAMY, P. and SHARMA, D. (1989). A note on the estimation of the selected gamma population. *Comm. Statist. Theory Methods* **18** 555–560.
- VELLAISAMY, P. (1990). Average worth and simultaneous estimation of the selected subset. *Ann. Inst. Statist. Math.* To appear.
- VENTER, J. H. (1988). Estimation of the mean of the selected population. *Comm. Statist. Theory Methods* **17** 797–805.

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY
POWAI, BOMBAY 400 076
INDIA