Jaroslav Hustý

Subset selection of the largest location parameter based on L-estimates

Aplikace matematiky, Vol. 29 (1984), No. 6, 397–410

Persistent URL: http://dml.cz/dmlcz/104114

SUBSET SELECTION OF THE LARGEST LOCATION PARAMETER BASED ON L-ESTIMATES

Jaroslav Hustý

(Received December 30, 1982)

Consider the problem of selecting a subset containing the largest of several location parameters. A robust competitor to the classical Gupta's selection rule based on sample means (see [7]), namely a rule based on sample medians, was investigated by Gupta and Singh ([8]). As a certain generalization of both, a rule based on the L-estimates of location is proposed here. This rule is strongly monotone as well, and minimax in the class of all selection rules which satisfy the P^*-condition, the risk being the expected subset size, provided the underlying density has monotone likelihood ratio. The problem of fulfilling the P^*-condition is solved explicitly only asymptotically, under the asymptotic normality of the L-estimates used. However, after replacing their asymptotic variance by its estimate given in [12], this solution becomes fully distribution-free.

1. INTRODUCTION

Let $\pi_1, \ldots, \pi_k$ be $k \geqq 2$ populations with distribution functions $F(x - \theta_1), \ldots$ $\ldots, F(x - \theta_k)$, respectively, where $F(x)$ is assumed to be absolutely continuous with a density $f(x)$. The populations $\pi_1, \ldots, \pi_k$ thus differ only by the values $\theta_1, \ldots, \theta_k$ of a location parameter θ, which are unknown. Denoting the population with the largest location parameter as the best one, our aim is to select a subset of $\{\pi_1, \ldots \pi_k\}$ containing the best population. (If there are more than one population with the largest location parameter, then any one of them is tagged as the best one.) The decision (selection) is made on the basis of independent samples $X_{i1}, \ldots, X_{in}$ from π_i, $i = 1, \ldots, k$, all of the same sample size n. We denote the pooled sample space by $\mathscr{X}^{kn} = \{\boldsymbol{x}\} = \{(x_{11}, \ldots, x_{1n}, x_{21}, \ldots, x_{kn})\}$. The action space $\mathscr{A}$ consists of $2^k - 1$ nonempty subsets $\mathscr{S}$ of $\{\pi_1, \ldots, \pi_k\}$. A selection (decision) rule is a measurable function $\delta^{(n)} : \mathscr{X}^{kn} \times \mathscr{A} \to [0, 1]$ with the property

$$\sum_{\mathscr{S} \in \mathscr{A}} \delta^{(n)}(\boldsymbol{x}, \mathscr{S}) = 1 \quad \text{for all} \quad \boldsymbol{x} \in \mathscr{X}^{kn} .$$

$\delta^{(n)}(\boldsymbol{x}, \mathscr{S})$ is the probability of selecting the subset $\mathscr{S}$ when $\boldsymbol{x}$ was observed. The probability of including π_i in the selected subset when having observed $\boldsymbol{x}$,

$$\psi_i^{(n)}(\boldsymbol{x}) = \sum_{\mathscr{S}:\pi_i \in \mathscr{S}} \delta^{(n)}(\boldsymbol{x}, \mathscr{S}),$$

is the individual selection probability for π_i. As the risk, for the loss considered herein, depends on $\delta^{(n)}$ only through the individual selection probabilities, we may regard two rules with the same individual selection probabilities (for all $\boldsymbol{x}$) as equivalent and define the selection rule $\psi^{(n)}$ as a measurable mapping from $\mathscr{X}^{kn}$ into $\underbrace{[0, 1] \times \dots \times [0, 1]}_{\text{k---times}}$:

$$\psi^{(n)}(\boldsymbol{x}) = (\psi_1^{(n)}(\boldsymbol{x}), \dots, \psi_k^{(n)}(\boldsymbol{x})) .$$

Further, a nonrandomized selection rule may be equivalently defined by k (overlapping) subsets of $\mathscr{X}^{kn}$; if $(X_{11}, \dots, X_{kn})$ falls into the i-th subset, then π_i is selected into $\mathscr{S}$. These subsets are often defined by means of some statistics $Y_i = Y(X_{i1}, \dots$ $\dots, X_{in})$ with a distribution depending on θ_i $(i = 1, \dots, k)$.

We introduce the following usual notations. $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$ is the vector of the actual parameter values, $\Omega = \{\boldsymbol{\theta}\}$ the parameter space (k-dimensional Euclidean space in our case). $Y_{[1]} \leqq \dots \leqq Y_{[k]}$ are the ordered values of the used statistics $Y_1, \dots, Y_k$, $\theta_{[1]} \leqq \dots \leqq \theta_{[k]}$ the ordered values $\theta_1, \dots, \theta_k$, $\pi_{[1]}, \dots, \pi_{[k]}$ the corresponding populations and $Y_{(1)}, \dots, Y_{(k)}$ the corresponding statistics (so that $Y_{(i)}$ comes from $\pi_{[i]}$).

The selection of any subset $\mathscr{S}$ which contains the best population $\pi_{[k]}$ is called a correct selection (CS). For the probability of correct selection we write $\mathsf{P}\{\mathrm{CS}/\psi^{(n)}, \boldsymbol{\theta}\}$. We are interested only in rules with high probability of correct selection, namely in rules $\psi^{(n)}$ that satisfy the so called P^*-condition

$$\inf_{\boldsymbol{\theta} \in \Omega} \mathsf{P}\{\mathrm{CS}/\psi^{(n)}, \boldsymbol{\theta}\} \geqq P^* ,$$

where P^* is a preassigned fixed number, $1/k < P^* < 1$. The set of all selection rules that satisfy the P^*-condition is denoted by $\mathscr{R}_{P^*}^{(n)}$. Among them, it is further desirable to use only those selecting small subsets. Therefore, for the loss function (as in some papers quoted below) we take the number of the selected populations, $S = \operatorname{card} \mathscr{S}$. The risk of a selection rule $\psi^{(n)}$ is then given by the expected subset size $\mathsf{E}(S/\psi^{(n)}, \boldsymbol{\theta})$.

2. GUPTA-TYPE RULE BASED ON L-ESTIMATES

A Gupta-type selection rule for the location parameter case using some statistics $Y_1, \dots, Y_k$ is defined as follows:

$$\text{(1)} \qquad \text{select } \pi_i \quad \text{iff} \quad Y_i \geqq Y_{[k]} - d \quad (i = 1, \dots, k)$$

where $d = d(n) > 0$ is a constant chosen as small as possible and so that the P^*-condition is satisfied. Gupta ([7]) proposed this rule for the normal means case with sample means $\bar{X}_i = Y_i$. As the sample mean is too sensitive to deviations from normality, other, more robust rules have also been studied. Lately, Gupta and Singh ([8]) have investigated a rule based on sample medians. A rule $\psi_L^{(n)}$ based on L-estimates (linear combinations of order statistics) is suggested here. These statistics are easy to compute and a special case of them, the trimmed mean, may be regarded as a compromise between the sample mean, optimal for the normal distribution, and the sample median, robust against deviations from normality. We obtain the rule $\psi_L^{(n)}$ if we put

$$Y_i = L_i = \sum_{j=1}^{n} \lambda_j X_{i[j]}$$

into (1), where $X_{i[1]} < \ldots < X_{i[n]}$ is the ordered i-th sample. The coefficients $\lambda_1, \ldots, \lambda_n$ are assumed to be generated by a weight function $J(u) \geqq 0, u \in (0, 1), \int_0^1 J(u) \, \mathrm{d}u = 1$, through the relation

$$\lambda_j = \frac{1}{n} J\left(\frac{j}{n+1}\right).$$

The statistic L_i has the distribution function $G^{(n)}(y - \lambda^{(n)} \theta_i)$, where $\lambda^{(n)} = \sum_{i=1}^{n} \lambda_j$ is supposed to be positive and

$$G^{(n)}(y) = n! \int \cdots \int_{\substack{x_1 < \ldots < x_n \\ \Sigma \lambda_j x_j \leqq y}} f(x_1) \ldots f(x_n) \, \mathrm{d}x_1 \ldots \mathrm{d}x_n ,$$

and the corresponding density is $g^{(n)}(y - \lambda^{(n)} \theta_i)$. Similarly as in (1.11) and (1.12) of [7] we get

$$\begin{aligned} (2) \qquad \mathsf{P}\{CS/\psi_L^{(n)}, \boldsymbol{\theta}\} &= \mathsf{P}\{L_{(k)} \geqq L_{[k]} - d \mid \boldsymbol{\theta}\} = \\ &= \mathsf{P}\{L_{(i)} \leqq L_{(k)} + d \,, \quad i = 1, \ldots, k-1 \mid \boldsymbol{\theta}\} = \\ &= \int_{-\infty}^{\infty} \left[\prod_{i=1}^{k-1} G^{(n)}(y + d + \lambda^{(n)}(\theta_{[k]} - \theta_{[i]}))\right] g^{(n)}(y) \, \mathrm{d}y \,. \end{aligned}$$

It is clear from this expression that the minimum probability of CS is achieved at any parameter point $\boldsymbol{\theta}$ with $\theta_{[1]} = \theta_{[k]}$ and that the minimal value of the constant d ensuring the P^*-condition is determined (at least theoretically) from the equation

$$(3) \qquad \mathsf{P}\{CS/\psi_L^{(r)}, \theta_{[1]} = \theta_{[k]}\} = \int_{-\infty}^{\infty} [G^{(n)}(x + d)]^{k-1} \, g^{(n)}(y) \, \mathrm{d}y = P^* \,.$$

As mentioned above, CS means the selection of $\pi_{[k]}$. Generally, for any i $(1 \leqq i \leqq k)$

$$(4) \qquad \mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \boldsymbol{\theta}\} = \int_{-\infty}^{\infty} \left[\prod_{\substack{r=1 \\ r \neq i}}^{k} G^{(n)}(y + d + \lambda^{(n)}(\theta_{[i]} - \theta_{[r]}))\right] g^{(n)}(y) \, \mathrm{d}y$$

is obtained in the same way as (2). It is now readily seen that $\psi_L^{(n)}$ is strongly monotone, i.e. $\mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \boldsymbol{\theta}\}$ is nondecreasing in $\theta_{[i]}$ (when all other components of $\boldsymbol{\theta}$ are fixed) and nonincreasing in $\theta_{[r]}$, $r \neq i$ (when all other components of $\boldsymbol{\theta}$ are fixed), for all $i = 1, \ldots, k$. Thus, according to Remark 4.2 in [11], $\psi_L^{(n)}$ is also monotone, i.e.

$$\mathsf{P}\{\text{select } \pi_{[r]}/\psi_L^{(n)}, \boldsymbol{\theta}\} \geqq \mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \boldsymbol{\theta}\}$$

for all $1 \leqq i < r \leqq k$ and $\boldsymbol{\theta} \in \Omega$. This, of course, entails also the unbiasedness of $\psi_L^{(n)}$:

$$\mathsf{P}\{\text{select } \pi_{[k]}/\psi_L^{(n)}, \boldsymbol{\theta}\} \geqq \mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \boldsymbol{\theta}\}$$

for all $1 \leqq i < k$ and $\boldsymbol{\theta} \in \Omega$.

3. MAXIMUM VALUE OF THE RISK

Using (4), we obtain for the risk of the rule $\psi_L^{(n)}$:

$$\mathsf{E}(S/\psi_L^{(n)}, \boldsymbol{\theta}) = \sum_{i=1}^{k} \mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \boldsymbol{\theta}\} = \tag{5}$$

$$= \sum_{i=1}^{k} \int_{-\infty}^{\infty} \Big[\prod_{\substack{r=1 \\ r \neq i}}^{k} G^{(n)}(y + d + \lambda^{(n)}(\theta_{[i]} - \theta_{[r]}))\Big] \, g^{(n)}(y) \, \mathrm{d}y \, .$$

To investigate its behaviour, it is desirable that the density $g^{(n)}(y - \lambda^{(n)}\theta)$ have monotone likelihood ratio (MLR) in θ. This follows from the MLR-property of the underlying density $f(x - \theta)$, provided the coefficients $\lambda_1, \ldots, \lambda_n$ are such that all coefficients between any two positive ones are again positive (see [10]). Then we come to the following conclusions:

(i) The risk $\mathsf{E}(S/\psi_L^{(n)}, \boldsymbol{\theta})$ attains its maximum at any point $\boldsymbol{\theta}$ with $\theta_{[1]} = \theta_{[k]}$, and

$$\max_{\boldsymbol{\theta} \in \Omega} \mathsf{E}(S/\psi_L^{(n)}, \boldsymbol{\theta}) = \mathsf{E}(S/\psi_L^{(n)}, \theta_{[1]} = \theta_{[k]}) = kP^* \, .$$

This is proved in the same way as Theorem 1 (and Corrolary 1) in [7].

(ii) The rule $\psi_L^{(n)}$ is minimax in the class $\mathscr{R}_{P^*}^{(n)}$ with respect to the risk $\mathsf{E}(S/\psi^{(n)}, \boldsymbol{\theta})$. This is a direct consequence of Theorem 3.2 of [4].

(iii) It is clear from (i) that the maximum value of $\mathsf{E}(S/\psi_L^{(n)}, \boldsymbol{\theta})$ can be diminished only in some subset of Ω. If we choose the subset

$$C(\gamma^*) = \{\boldsymbol{\theta} \in \Omega : \theta_{[k-1]} \leqq \theta_{[k]} - \gamma^*\} \tag{6}$$

with some $\gamma^* < 0$, we get (see Theorem 1 and Corrolary 2 in [7])

$$\max_{\boldsymbol{\theta} \in C(\gamma^*)} \mathsf{E}(S/\psi_L^{(n)}, \boldsymbol{\theta}) = \mathsf{E}(S/\psi_L^{(n)}, \theta_{[1]} = \theta_{[k-1]} = \theta_{[k]} - \gamma^*) = \tag{7}$$

$$= \int_{-\infty}^{\infty} [G^{(n)}(y + d + \lambda^{(n)}\gamma^*)]^{k-1} g^{(n)}(y)\,\mathrm{d}y +$$

$$+ (k - 1) \int_{-\infty}^{\infty} [G^{(n)}(y + d)]^{k-2} G^{(n)}(y + d - \lambda^{(n)}\gamma^*) g^{(n)}(y)\,\mathrm{d}y\,.$$

The condition

$$\max_{\theta \in C(\gamma^*)} \mathsf{E}(S/\psi_L^{(n)}, \theta) \leqq 1 + \varepsilon \tag{8}$$

with given $\gamma^* > 0$ and $\varepsilon > 0$ can then be used for determining the sample size as the smallest n for which (8) is fulfilled. $C(\gamma^*)$ is sometimes called the preference zone for the risk.

4. ASYMPTOTIC SOLUTION FOR KNOWN DISTRIBUTION

Once one can assume the knowledge of $F(x)$ and has chosen some appropriate $J(u)$, it should be possible to find $G^{(n)}(y)$ (at least numerically for sufficiently many y) and to evaluate the integrals in (3) and (7) (see the algorithm given in [5]). But this seems to be rather laborious as one would have to do it for various n and d (and k). Therefore we give an asymptotic solution of the problem of finding d and n according to the requirements (3) and (8), respectively, analogous to that given in [2]. We shall also investigate some asymptotic properties of the sequences $\{\psi_L^{(n)}\}_{n=1}^{\infty}$ of selection rules of the same type. To be able to do this, we shall use the results of [13], especially the asymptotic normality of L-estimates. We first list conditions on the functions $J(u)$ and $F(x)$, some of which will have to be fulfilled:

(A) $J(u)$ is bounded on $(0, 1)$.

(B) $J(u)$ is continuous a.e. F^{-1}.

(C) $J(u) = 0$ for $u \in (0, \alpha)$ and $u \in (1 - \alpha, 1)$, where $0 < \alpha < \frac{1}{2}$.

(D) $\int_{-\infty}^{\infty} x^2\,\mathrm{d}F(x) < \infty$.

(E) $\lim_{x \to \infty} x^{\beta}[1 - F(x) + F(-x)] = 0$ for some $\beta > 0$.

The following theorem yields the asymptotic solution for $d = d(n)$.

Theorem 1. *Let $P^* \in (1/k, 1)$ be given and for every $n \geqq 1$, let $d = d(n)$ be determined so that (3) holds. Let the conditions (A), (B), (D) or the conditions (A), (B), (C), (E) be satisfied. Then*

$$d(n) = \delta\,\sigma(J, F)\,n^{-1/2} + o(n^{-1/2}) \tag{9}$$

as $n \to \infty$, where δ is the solution of

$$P^* = Q_{k-1}(2^{-1/2}\delta, \ldots, 2^{-1/2}\delta)\,, \tag{10}$$

Q_{k-1} is the distribution function of the $(k-1)$-dimensional normal distribution with zero means, unit variances and covariances $\frac{1}{2}$, and

$$(11) \qquad \sigma^2(J, F) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J[F(x)]\, J[F(y)]\, [F(\min(x, y)) - F(x)\, F(y)]\, \mathrm{d}x\, \mathrm{d}y\,.$$

Proof. We put $\tilde{L}_i = \sum_{j=1}^{n} \lambda_j \tilde{X}_{i[j]}$ where $\tilde{X}_{ij} = X_{ij} - \theta_i$, so that the variables $\tilde{X}_{ij}$ $(i = 1, \ldots, k; j = 1, \ldots, n)$ are i.i.d. with the distribution function $F(x)$ and also $\tilde{L}_1, \ldots, \tilde{L}_k$ are i.i.d. From (3), we have

$$P^* = \mathrm{P}\{n^{1/2}[2\,\sigma^2(J, F)]^{-1/2}\, (\tilde{L}_{(i)} - \tilde{L}_{(k)}) \leqq$$
$$\leqq d(n)\, n^{1/2}[2\,\sigma^2(J, F)]^{-1/2},\ i = 1, \ldots, k-1\} =$$
$$= H_{k-1}^{(n)}(d(n)\, n^{1/2}[2\,\sigma^2(J, F)]^{-1/2}, \ldots, d(n)\, n^{1/2}[2\,\sigma^2(J, F)]^{-1/2}) = H_{k-1}^{(n)}(\boldsymbol{a}(n))\,,$$

say. According to the assumptions, the distribution functions $H_{k-1}^{(n)}$ converge to Q_{k-1} as $n \to \infty$ (see Theorems 1 and 2 or Theorem 5 of [13]) and this convergence is uniform in the argument since Q_{k-1} is continuous. Hence for any $\eta > 0$,

$$\eta \geqq |Q_{k-1}(\boldsymbol{a}(n)) - H_{k-1}^{(n)}(\boldsymbol{a}(n))| = |Q_{k-1}(\boldsymbol{a}(n)) - P^*|$$

for n sufficiently large. The existence of the limit

$$\lim_{n\to\infty} Q_{k-1}(\boldsymbol{a}(n)) = P^*$$

and (10) entail

$$(12) \qquad \lim_{n\to\infty} d(n)\, n^{1/2}[2\,\sigma^2(J, F)]^{-1/2} = 2^{-1/2}\delta\,,$$

which is equivalent to (9).

Using (9), we get the following selection rule for large n:

$$(13) \qquad \text{select } \pi_i \quad \text{iff} \quad L_i \geqq L_{[k]} - \delta[\sigma^2(J, F)/n]^{1/2}\,.$$

The values δ are tabulated in [3], Table I, for $k = 2(1)\,10$ (the entry t of the table must be $t = 1$) and various P^*. $\sigma^2(J, F)$ is tabulated in [9] for the case of trimmed means and F normal, double exponential, logistic and uniform.

If we wish the condition (8) with any prescribed $\varepsilon > 0$ to be satisfied for large n, the sequence $\{\psi_L^{(n)}\}$ should possess the property of the so called risk consistency with respect to the preference zones $C(\gamma^*)$, $\gamma^* > 0$, i.e.

$$(14) \qquad \lim_{n\to\infty} \max_{\boldsymbol{\theta}\in C(\gamma^*)} \mathrm{E}(S/\psi_L^{(n)}, \boldsymbol{\theta}) = 1$$

for every $\gamma^* > 0$. This is ensured by the next theorem.

Theorem 2. *Let the assumptions of Theorem 1 be fulfilled. Then the sequence $\{\psi_L^{(n)}\}$ is risk consistent with respect to the system of preference zones $\{C(\gamma^*), \gamma^* < 0\}$ with $C(\gamma^*)$ given by (6).*

Proof. Let $\tilde{L}_i$ have the same meaning as in the proof of Theorem 1. With regard to $d(n) \to 0$, $\lambda^{(n)} \to 1$, we get for any $\boldsymbol{\theta} \in C(\gamma^*)$, $i = 1, \ldots, k-1$ and n large enough

$$\mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \boldsymbol{\theta}\} \leqq \mathsf{P}\{L_{(i)} \geqq L_{(k)} - d(n)/\boldsymbol{\theta}\} \leqq$$
$$\leqq \mathsf{P}\{\tilde{L}_{(i)} \geqq \tilde{L}_{(k)} + \lambda^{(n)}\gamma^* - d(n)\} \leqq \mathsf{P}\{|\tilde{L}_{(i)} - \mathsf{E}\tilde{L}_{(i)}| \geqq (\lambda^{(n)}\gamma^* - d(n))/2 \text{ or}$$
$$|\tilde{L}_{(k)} - \mathsf{E}\tilde{L}_{(k)}| \geqq (\lambda^{(n)}\gamma^* - d(n))/2\} \leqq 8 \operatorname{Var} \tilde{L}_1/(\lambda^{(n)}\gamma^* - d(n))^2 .$$

So we have also

$$\max_{\boldsymbol{\theta} \in C(\gamma^*)} \mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \boldsymbol{\theta}\} \leqq 8 \operatorname{Var} \tilde{L}_1/(\lambda^{(n)}\gamma^* - d(n))^2$$

and from $\lim_{n\to\infty} \operatorname{Var} \tilde{L}_1 = 0$ (see Theorem 1 or 5 of [13])

$$\lim_{n\to\infty} \max_{\boldsymbol{\theta} \in C(\gamma^*)} \mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \boldsymbol{\theta}\} = 0 , \quad i = 1, \ldots, k-1 .$$

The rules are such that $S \geqq 1$; hence

$$1 \leqq \limsup_{n\to\infty} \max_{\boldsymbol{\theta} \in C(\gamma^*)} \mathsf{E}(S/\psi_L^{(n)}, \boldsymbol{\theta}) \leqq$$
$$\leqq \sum_{i=1}^{k-1} \lim_{n\to\infty} \max_{\boldsymbol{\theta} \in C(\gamma^*)} \mathsf{P}\{\text{select } \pi_{[i]}/\psi_L^{(n)}, \theta\} +$$
$$+ \limsup_{n\to\infty} \max_{\boldsymbol{\theta} \in C(\gamma^*)} \mathsf{P}\{\text{select } \pi_{[k]}/\psi_L^{(n)}, \boldsymbol{\theta}\} \leqq 1$$

and (14) is proved.

Let us mention that we do not need the MLR-property of $g^{(n)}(y - \lambda^{(n)}\theta)$ for the risk consistency of $\{\psi_L^{(n)}\}$.

Under (14), for each $\varepsilon > 0$ and each $\gamma^* > 0$ we can find a natural n^* such that for every $n \geqq n^*$ (8) is fulfilled. We shall denote by $n^*(\gamma^*, \varepsilon) = n^*(\gamma^*, \varepsilon; \psi_L)$ the minimal n^* with this property; here ψ_L stands for the whole sequence $\{\psi_L^{(n)}\}$. Asymptotical considerations about $\max_{\boldsymbol{\theta} \in C(\gamma^*)} \mathsf{E}(S/\psi^{(n)}, \boldsymbol{\theta})$ can be connected either with $\varepsilon \to 0$ and γ^* fixed or $\gamma^* \to 0$ and ε fixed. We shall treat the latter case first.

Lemma 1. *Let the assumptions of Theorem 1 be satisfied, let $\varepsilon \in (0, kP^* - 1)$ be fixed and let $g^{(n)}(y - \lambda^{(n)}\theta)$ have* MLR *in θ (for all n) (cf. the remark at the beginning of Section 3). Then $n^*(\gamma^*, \varepsilon; \psi_L)$ is a nonincreasing function of γ^* and*

(15) $$\lim_{\gamma^*\to 0} n^*(\gamma^*, \varepsilon; \psi_L) = \infty ,$$

so that the numbers $\{n^*(\gamma^*, \varepsilon; \psi_L), \gamma^* > 0\}$ can be arranged in an increasing *sequence* $\{n_m\}$. *If we choose for every* n_m *a* $\gamma^{(m)}$ *such that* $n^*(\gamma^{(m)}, \varepsilon; \psi_L) = n_m$, *then*

$$\lim_{m\to\infty} \gamma^{(m)} = 0\,. \tag{16}$$

Proof. The obvious monotonicity of $n^*(\gamma^*, \varepsilon)$ entails the existence of $\lim\limits_{\gamma^*\to 0} n^*(\gamma^*, \varepsilon)$. Assume this limit to be equal to $n_0 < \infty$. Then there is $\gamma_0^* > 0$ such that for every $\gamma^* \leqq \gamma_0^*$ we have $n^*(\gamma^*, \varepsilon) = n_0$, i.e. $1 + \varepsilon \geqq \max\limits_{\theta\in C(\gamma^*)} \mathsf{E}(S/\psi_L^{(n_0)}, \boldsymbol{\theta})$. This maximum is a function nondecreasing with $\gamma^* \searrow 0$ and continuous according to (7). Hence we get

$$1 + \varepsilon \geqq \lim_{\gamma^*\to 0} \max_{\theta\in C(\gamma^*)} \mathsf{E}(S/\psi_L^{(n_0)}, \boldsymbol{\theta}) =$$

$$= k \int [G^{(n_0)}(y + d(n_0))]^{k-1}\, g^{(n_0)}(y)\, \mathrm{d}y = kP^*\,,$$

a contradiction with the assumptions, and (15) is proved. Similarly, $\{\gamma^{(m)}\}$ is a decreasing sequence and $\lim\limits_{m\to\infty} \gamma^{(m)} > 0$ would lead to a contradiction with the risk consistency of $\{\psi_L^{(n)}\}$.

Theorem 3. *Let the assumptions of Lemma* 1 *be satisfied. Then*

$$n^*(\gamma^*, \varepsilon; \psi_L) = (\gamma/\gamma^*)^2\, \sigma^2(J, F) + o((\gamma^*)^{-2}) \tag{17}$$

as $\gamma^* \to 0$, *where* γ *is the solution of*

$$\begin{aligned} 1 + \varepsilon = {} & Q_{k-1}((\delta + \gamma)\, 2^{-1/2}, \ldots, (\delta + \gamma)\, 2^{-1/2}) + \\ & + (k - 1)\, Q_{k-1}(\delta\, 2^{-1/2}, \ldots, \delta\, 2^{-1/2}, (\delta - \gamma)\, 2^{-1/2}) \end{aligned} \tag{18}$$

and δ, Q_{k-1} *are the same as in Theorem* 1.

Proof. We choose a sequence $\{\gamma^{(m)}\}$ as in Lemma 1. According to (5) and (7) and with the same notation as in the proof of Theorem 1 we have

$$\max_{\theta\in C(\gamma^{(m)})} \mathsf{E}(S/\psi_L^{(n_m)}, \boldsymbol{\theta}) =$$

$$= \sum_{i=1}^{k} \mathsf{P}\{L_{(i)} \geqq L_{(r)} - d(n_m),\, r = 1, \ldots, k;\, r \neq i/\theta_{[1]} = \theta_{[k-1]} = \theta_{[k]} - \gamma^{(m)}\} =$$

$$= \sum_{i=1}^{k-1} \mathsf{P}\{\tilde{L}_{(r)} - \tilde{L}_{(i)} \leqq d(n_m),\, r = 1, \ldots, k - 1;\, r \neq i,\, \tilde{L}_{(k)} - \tilde{L}_{(i)} \leqq$$

$$\leqq d(n_m) - \lambda^{(n_m)}\gamma^{(m)}\} + \mathsf{P}\{\tilde{L}_{(r)} - \tilde{L}_{(k)} \leqq d(n_m) + \lambda^{(n_m)}\gamma^{(m)},\, r = 1, \ldots, k - 1\}\,.$$

Each vector $(n_m^{1/2}[2\,\sigma^2(J, F)]^{-1/2}\, (\tilde{L}_{(r)} - \tilde{L}_{(i)}),\ r = 1, \ldots, k;\, r \neq i)$, $i = 1, \ldots, k$, has the same distribution function $H_{k-1}^{(n_m)}$, which tends to Q_{k-1} (uniformly) for $n_m \to \infty$.

Simplifying the notation by putting

$$\begin{aligned} a(n_m) &= n_m^{1/2}[2\,\sigma^2(J, F)]^{-1/2}\, d(n_m)\,, \\ b_m &= n_m^{1/2}[2\,\sigma^2(J, F)]^{-1/2}\, \lambda^{(n_m)}\, \gamma^{(m)}\,, \\ c_m &= (n_m - 1)^{1/2}\,[2\,\sigma^2(J, F)]^{-1/2}\, \lambda^{(n_m - 1)}\, \gamma^{(m)}\,, \end{aligned}$$

we can write with regard to the choice of $\gamma^{(m)}$

$$\begin{gathered} (k-1)\, H_{k-1}^{(n_m)}(a(n_m), \ldots, a(n_m), a(n_m) - b_m) + H_{k-1}^{(n_m)}(a(n_m) + b_m, \ldots, a(n_m) + b_m) \leqq \\ \leqq 1 + \varepsilon < (k-1)\, H_{k-1}^{(n_m - 1)}(a(n_m - 1), \ldots, a(n_m - 1), a(n_m - 1) - c_m) + \\ + H_{k-1}^{(n_m - 1)}(a(n_m - 1) + c_m, \ldots, a(n_m - 1) + c_m)\,. \end{gathered}$$

The uniform convergence of $H_{k-1}^{(n_m)}$ to Q_{k-1} yields

$$\begin{gathered} (k-1)\, Q_{k-1}(a(n_m), \ldots, a(n_m), a(n_m) - b_m) + \\ + Q_{k-1}(a(n_m) + b_m, \ldots, a(n_m) + b_m) \leqq 1 + \varepsilon + k\eta\,, \\ (k-1)\, Q_{k-1}(a(n_m - 1), \ldots, a(n_m - 1), \\ a(n_m - 1) - c_m) + Q_{k-1}(a(n_m - 1) + c_m, \ldots, a(n_m - 1) + c_m) > 1 + \varepsilon - k\eta \end{gathered}$$

for any $\eta > 0$ and n_m sufficiently large (i.e. for all m sufficiently large). Denoting the function of two variables on the left hand sides of these inequalities by Q, we get

$$\begin{aligned} \limsup_{m \to \infty} Q(a(n_m), b_m) &\leqq 1 + \varepsilon\,, \\ \liminf_{m \to \infty} Q(a(n_m - 1), c_m) &\geqq 1 + \varepsilon\,. \end{aligned}$$

Q_{k-1} as a continuous distribution function is uniformly continuous, hence so is also Q. Further Q is increasing in the first variable, and decreasing in the second one when both have the same sign (this can be easily proved by using formula (1.3) of [6] for Q_{k-1} and differentiating Q with respect to the second variable). As $\{a(n_m)\}$ is convergent (see (12)), we may write

$$\begin{gathered} 1 + \varepsilon \leqq Q(\liminf a(n_m - 1), \limsup c_m) = \\ = Q(\lim a(n_m), \limsup b_m) \leqq Q(\limsup a(n_m), \liminf b_m) \leqq 1 + \varepsilon\,, \end{gathered}$$

which entails the existence of $\lim_{m \to \infty} b_m$ and

$$Q(\lim a(n_m), \lim b_m) = 1 + \varepsilon\,.$$

Comparing this with (12) and (18) and recalling that $\lambda^{(n_m)} \to 1$ we obtain

$$\lim_{m \to \infty} n_m^{1/2} \gamma^{(m)} [2\,\sigma^2(J, F)]^{-1/2} = \gamma\, 2^{-1/2}\,.$$

The arbitrariness of the choice of the sequence $\{\gamma^{(m)}\}$ implies

(19) $$\lim_{\gamma^* \to 0} n^*(\gamma^*, \varepsilon; \psi_L) \cdot (\gamma^*)^2 = \gamma^2 \sigma^2(J, F),$$

which is equivalent to (17).

Theorem 3 yields the asymptotic solution for $n^*(\gamma^*, \varepsilon; \psi_L)$. A combination of (19) with (12) gives

$$\lim_{\gamma^* \to 0} d(n^*(\gamma^*, \varepsilon; \psi_L))/\gamma^* = \delta/\gamma$$

or

$$d(n^*(\gamma^*, \varepsilon; \psi_L)) = \delta\gamma^*/\gamma + o(\gamma^*).$$

That is, if we use the selection rule

(20) $$\text{select } \pi_i \quad \text{iff} \quad L_i \geqq L_{[k]} - \delta\gamma^*/\gamma$$

with the sample size approximately determined by

(21) $$n \approx (\gamma/\gamma^*)^2 \sigma^2(J, F),$$

the P^*-condition and the condition (8) are (under the assumptions of Theorem 3) approximately satisfied for small γ^*.

Theorem 3 enables us also to determine the asymptotic relative efficiency (ARE) of the rule ψ_L with respect to Gupta's means rule for $\gamma^* \to 0$ and ε fixed, the ARE of a rule ψ_2 with respect to a rule ψ_1 being in this case defined as

$$\lim_{\gamma^* \to 0} \frac{n^*(\gamma^*, \varepsilon; \psi_1)}{n^*(\gamma^*, \varepsilon; \psi_2)}.$$

Theorem 4. *Let the conditions* (A), (B), (D) *be satisfied,* $\varepsilon \in (0, kP^* - 1)$ *and for all* n let $g^{(n)}(y - \lambda^{(n)}\theta)$ *as well as the density of the sample means have* MLR *in* θ *(cf. the remark at the beginning of Section* 3*). Then the* ARE *of the rule* ψ_L *relative to Gupta's rule for* $\gamma^* \to 0$ *and* ε *fixed is*

(22) $$e_{L,G}(F) = \frac{\sigma^2(F)}{\sigma^2(J, F)}$$

(and hence does not depend on ε*), where* $\sigma^2(F) = \sigma^2(1, F)$ *is the variance of the distribution* F.

For some numerical values of (22) see Table in [9].

The case $\varepsilon \to 0$, γ^* fixed, is similar, though we do not get so detailed results. Let us suppose that the assumptions of Theorem 1 are satisfied and the functions $g^{(n)}(y-\lambda^{(n)}\theta)$ have MLR in θ. $n^*(\gamma^*, \varepsilon; \psi_L)$ is obviously a monotone function of ε, but it may be bounded. For instance, if F and hence also $G^{(n)}$ have a finite support $[a, b]$ and $\lambda^{(n)}\gamma^* > b - a + d(n)$, it follows from (7) that $\max_{\theta \in C(\gamma^*)} \mathsf{E}(S/\psi_L^{(n)}, \theta) = 1$. Of course,

$n^*(\gamma^*, \varepsilon; \psi_L)$ is unbounded when (7) is a strictly monotone function of γ^*. As the boundedness of $n^*(\gamma^*, \varepsilon; \psi_L)$ apparently implies that γ^* was chosen unnecessarily large, this case is of no practical interest and we may assume

$$\lim_{\varepsilon \to 0} n^*(\gamma^*, \varepsilon; \psi_L) = \infty ,$$

though we do not have a general boundary for γ^* like for ε in Lemma 1. We do not get an analogue of Theorem 3, either. We may only argue that the approximate equality (for small ε)

$$\max_{\boldsymbol{\theta} \in C(\gamma^*)} \mathsf{E}(S/\psi_L^{(n^*(\gamma^*,\varepsilon))}, \boldsymbol{\theta}) \approx 1 + \varepsilon$$

implies (see the proof of Theorem 3)

$$Q(2^{-1/2}\delta, [n^*(\gamma^*, \varepsilon)/2\, \sigma^2(J, F)]^{1/2}\, \lambda^{(n^*(\gamma^*,\varepsilon))}\gamma^*) \approx 1 + \varepsilon ,$$

hence

$$n^*(\gamma^*, \varepsilon; \psi_L) \approx (\gamma(\varepsilon)/\gamma^*)^2\, \sigma^2(J, F)$$

where $\gamma(\varepsilon)$ is the solution of (18) for γ. So we obtain practically the same solution as in the case ε fixed, $\gamma^* \to 0$ (see (20) and (21)). Also the ARE (defined here as

$$\lim_{\varepsilon \to 0} \frac{n^*(\gamma^*, \varepsilon; \psi_1)}{n^*(\gamma^*, \varepsilon; \psi_2)},$$

of the rule ψ_L relative to the means rule does not depend on γ^* and is again equal to (22). This is not true generally, for any two rules; cf. [1], formula (4.22).

5. ASYMPTOTIC SOLUTION FOR UNKNOWN DISTRIBUTION

As we have seen, even the asymptotic solution for $d(n)$ and $n^*(\gamma^*, \varepsilon; \psi_L)$ depends on the underlying distribution, namely, through $\sigma^2(J, F)$. This disadvantage can be avoided by replacing $\sigma^2(J, F)$ by its estimate proposed by Sen in [12] (a suggestion of J. Jurečková in a personal communication). Then of course the question arises whether and in what sense the asymptotic results remain valid.

According to Sen, one gets an estimate of $\sigma^2(J, F)$ from the i-th sample $X_{i1}, \ldots, X_{in}$ by putting the corresponding empirical distribution function $F_i^{(n)}$ into (11) instead of F:

$$\hat{\sigma}_i^2 = \sigma^2(J, F_i^{(n)}) =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J[F_i^{(n)}(x)]\, J[F_i^{(n)}(y)]\, [F_i^{(n)}(\min(x, y)) - F_i^{(n)}(x)\, F_i^{(n)}(y)]\, \mathrm{d}x\, \mathrm{d}y =$$

$$= \frac{1}{n^2} \sum_{j=1}^{n-1} \sum_{l=1}^{n-1} J\left(\frac{j}{n}\right) J\left(\frac{l}{n}\right) [n \,.\, \min(j, l) - jl]\, (X_{i[j+1]} - X_{i[j]})\, (X_{i[l+1]} - X_{i[l]}) \,.$$

Being translation invariant, the variables $\hat{\sigma}_1^2, \ldots, \hat{\sigma}_k^2$ are identically and independently distributed so that we may take

$$\hat{\sigma}^2(J, F) = \hat{\sigma}_{(n)}^2 = \frac{1}{k} \sum_{i=1}^{k} \hat{\sigma}_i^2$$

for an estimate of $\sigma^2(J, F)$. For its use to our purpose, the following property is of great importance:

(F) $\hat{\sigma}_i^2$ converges to $\sigma^2(J, F)$ in probability as $n \to \infty$.

Sen in [12], Theorem 4.3, gives (rather unpleasant) conditions for the almost sure convergence of his estimate to $\sigma^2(J, F)$. In case that $J(u)$ trims the extremes (condition (C)) the almost sure convergence and hence also (F) follows immediately from the Glivenko theorem. (F) obviously implies the convergence of $\hat{\sigma}_{(n)}^2$ to $\sigma^2(J, F)$ in probability.

Let the sample size n be given. If we replace $\sigma^2(J, F)$ by $\hat{\sigma}_{(n)}^2$ in (13) we get the following selection rule $\hat{\psi}_L^{(n)}$:

$$\text{select } \pi_i \text{ iff } \quad L_i \geqq L_{[k]} - \delta(\hat{\sigma}_{(n)}^2/n)^{1/2} \,. \tag{23}$$

The next theorem shows that it is really possible to use this rule (for large n).

Theorem 5. *Let $P^* \in (1/k, 1)$ be given and let δ be the solution of (10). Let the conditions (A), (B), (D), (F) or the conditions (A), (B), (C), (E) be fulfilled. Then the sequence of selection rules $\{\hat{\psi}_L^{(n)}\}$ defined by (23) satisfies asymptotically the P^*-condition, i.e.*

$$\lim_{n \to \infty} \inf_{\boldsymbol{\theta} \in \Omega} \mathsf{P}\{CS/\hat{\psi}_L^{(n)}, \boldsymbol{\theta}\} = P^* \,.$$

Proof. Using the notation of the proof of Theorem 1, we have

$$\mathsf{P}\{CS/\hat{\psi}_L^{(n)}, \boldsymbol{\theta}\} = \mathsf{P}\{\tilde{L}_{(i)} - \tilde{L}_{(k)} \leqq \delta(\hat{\sigma}_{(n)}^2/n) + \lambda^{(n)}(\theta_{[k]} - \theta_{[i]}) \,, \quad i = 1, \ldots, k - 1 \,/\, \boldsymbol{\theta}\} \,.$$

As neither $\tilde{L}_i$ nor $\hat{\sigma}_{(n)}^2$ depend on $\boldsymbol{\theta}$, the infimum of $\mathsf{P}\{CS/\hat{\psi}_L^{(n)}, \boldsymbol{\theta}\}$ over Ω is achieved for $\theta_{[1]} = \theta_{[k]}$. Thus we may write

$$\inf_{\boldsymbol{\theta} \in \Omega} \mathsf{P}\{CS/\hat{\psi}_L^{(n)}, \boldsymbol{\theta}\} = \\ = \mathsf{P}\{[\sigma^2(J, F)/\sigma_{(n)}^2]^{1/2} \, n^{1/2}[2\,\sigma^2(J, F)]^{-1/2} \, (\tilde{L}_{(i)} - \tilde{L}_{(k)}) \leqq 2^{-1/2}\delta \,, \quad i = 1, \ldots, k - 1\} \,.$$

The proof is completed with help of Cramér's lemma recalling that $\sigma^2(J, F)/\hat{\sigma}_{(n)}$ tends to 1 in probability.

Notice that for the case of given n we have a one-sample asymptotically distribution-free selection rule. This corresponds in a certain sense to Gupta's one-sample normal means rule under common unknown variances (see [7], Section 4).

If we also wish to control the expected size of the selected subset, we must use a two-sample rule as we have first to estimate the needed sample size $n^*(\gamma^*, \varepsilon)$ with help of an estimate of $\sigma^2(J, F)$. The procedure is as follows: We take samples $X_{i1}, \ldots, X_{im}$ of a sample size m from π_i, $i = 1, \ldots, k$, and use them to estimate $\sigma^2(J, F)$: $\hat{\sigma}^2(J, F) = \hat{\sigma}^2_{(m)}$. We put this estimate into the approximate formula (21) to get the sample size that ensures (asymptotically) (8):

$$\hat{n} \approx (\gamma/\gamma^*)^2 \, \hat{\sigma}^2_{(m)} \,. \tag{24}$$

We put $n = \max(m, \hat{n})$ and, in the case $n > m$, take further samples $X_{im+1}, \ldots, X_{in}$ from π_i, $i = 1, \ldots, k$. The decision is made according to (20), where the statistics are computed from the pooled samples $X_{i1}, \ldots, X_{im}, \ldots, X_{in}$.

Since the decision rule is the same as in the case of known $\sigma^2(J, F)$, the only problem that arises here is the question of the sense of the approximate equality in (24).

Theorem 6. *Let the assumptions of Theorem 5 be satisfied and let $g^{(n)}(y - \lambda^{(n)}\theta)$ have* MLR *in θ for all n. Then*

$$n^*(\gamma^*, \varepsilon; \psi_L) = (\gamma/\gamma^*)^2 \, \hat{\sigma}^2_{(m)} + o_P((\gamma^*)^{-2}) \quad \textit{for} \quad \gamma^* \to 0 \quad \textit{and} \quad m \to \infty \,. \tag{25}$$

Proof. We may write

$$[n^*(\gamma^*, \varepsilon; \hat{\psi}_L) - (\gamma/\gamma^*)^2 \, \hat{\sigma}^2_{(m)}] \, (\gamma^*)^2 =$$
$$= n^*(\gamma^*, \varepsilon; \psi_L) \, . \, (\gamma^*)^2 - \gamma^2 \, \sigma^2(J, F) + [\sigma^2(J, F) - \hat{\sigma}^2_{(m)}] \, \gamma^2 \,,$$

where the first difference tends to zero for $\gamma^* \to 0$ and the second one tends to zero in probability for $m \to \infty$.

References

[1] *R. E. Balow, S. S. Gupta:* Selection procedures for restricted families of probability distributions. Ann. Math. Statist. 40 (1969), 905—917.

[2] *N. S. Bartlett, Z. Govindarajulu:* Some distribution-free statistics and their application to the selection problem. Ann. Inst. Statist. Math. 20 (1968), 79—97.

[3] *R. E. Bechhofer:* A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. 25 (1954), 16—39.

[4] *R. L. Berger:* Minimax subset selection for loss measured by subset size. Ann. Statist. 7(1979), 1333—1338.

[5] *R. J. Carrol, S. S. Gupta:* On the probabilities of rankings of k populations. J. Statist. Comput. Simul. 5 (1977), 145—157.

[6] *R. N. Curnow, C. W. Dunnet:* The numerical evaluation of certain multivariate normal integrals. Ann. Math. Statist. 33 (1962), 571—579.

[7] *S. S. Gupta:* On some multiple decision (selection and ranking) rules. Technometrics 7 (1965), 225—245.

[8] *S. S. Gupta, A. K. Singh:* On rules based on sample medians for selection of the largest location parameter. Commun. Statist. — Theor. Meth. A 9 (1980), 1277—1298.

[9] *J. Hustý:* Ranking and selection procedures for location parameter case based on L-estimates. Apl. mat. 26 (1981), 377—388.

[10] *J. Hustý:* Total positivity of the density of a linear combination of order statistics. To appear in Čas. pěst. mat.

[11] *T. J. Santner:* A restricted subset selection approach to ranking and selection problems. Ann. Statist. 3 (1975), 334—349.

[12] *P. K. Sen:* An invariance principle for linear combinations of order statistics. Z. Wahrscheinlichkeitstheorie verw. Gebiete 42 (1978), 327—340.

[13] *S. M. Stigler:* Linear functions of order statistics with smooth weight functions. Ann. Statist. 2 (1974), 676—693.

Souhrn

SELEKCE PODMNOŽINY S NEJVĚTŠÍM PARAMETREM POLOHY ZALOŽENÁ NA L-ODHADECH

JAROSLAV HUSTÝ

Uvažuje se problém selekce podmnožiny populací obsahující populaci s největším parametrem polohy. Jakožto zobecnění selekčních pravidel založených na výběrovém průměru a na výběrovém mediánu se navrhuje pravidlo založené na L-odhadu polohy. Toto pravidlo je silně monotonní a minimaxové, vezmeme-li za rizikovou funkci očekávaný rozsah podmnožiny, a jestliže základní hustota má monotónní věrohodnostní poměr. Problém splnění P^*-podmínky je explicitně vyřešen pouze asymptoticky, jestliže užité L-odhady jsou asymptoticky normální. Nicméně nahradíme-li jejich asymptotický rozptyl odhadem, řešení již nebude záviset na distribuci.

Author's address: RNDr. *Jaroslav Hustý,* CSc., Matematický ústav ČSAV, Žitná 25, 115 67 Praha 1.