Sankhyā : The Indian Journal of Statistics
1999, Volume 61, Series A, Pt. 1, pp. 139-147

SELECTING THE t BEST CELLS OF A MULTINOMIAL DISTRIBUTION

By MAUSUMI BOSE
and
SUBIR KUMAR BHANDARI
Indian Statistical Institute, Calcutta

SUMMARY. The problem of selecting a subset of $s(\geq t)$ cells which contains the t best cells of a multinomial distribution is considered for the fixed-sample-size selection procedure for arbitrary s and t, $s \geq t$. The least favourable configuration over the difference zone is obtained for large sample sizes. This settles two interesting conjectures about the least favourable configuration by Chen and Hwang (*Commun. Statist. - Theor. Meth.* 13(10), 1289-98, 1984) in the affirmative.

1. **Introduction**

Consider a multinomial distribution with k cells and unknown probability vector p. In a general multinomial selection problem, the objective is to select a subset of $s(\geq t)$ cells which contains the t "best" cells, that is, the t cells with the t largest probabilities, $1 \leq t \leq k$, $k \geq 3$. For selecting this subset, a widely used selection procedure is the fixed-sample-size procedure, where a single sample of a fixed size n is drawn and the s cells having the largest frequencies are selected, with ties broken by randomization.

A usual requirement for the selection is that the probability of correct selection (PCS) must not be lower than a prespecified level p^* if the true configuration p lies in some preference zone. Two preference zones have been extensively used in the literature. One is the "difference zone", which is defined by $D(t,k,b) = \{p | p_t \geq p_{t+1} + b\}$ where b is a constant in the interval $(0, \frac{1}{t})$ and $p_1 \geq p_2 \geq \ldots \geq p_k$ are the ranked multinomial cell probabilities in p. The other is the "ratio zone" which is defined by $R(t,k,\theta) = \{p | p_t \geq \theta p_{t+1}\}$ where θ is a constant greater than one and the p_i's are as defined earlier.

Paper received. April 1997; revised October 1997.
AMS (1991) *subject classification.* 62F07.
Key words and phrases. Least favourable configuration, probability of correct selection, difference zone.

Definition 1.1. For a selection procedure, the configuration p in $D(t,k,b)$ which minimizes the PCS over $D(t,k,b)$ is called a least favourable configuration (LFC) over $D(t,k,b)$. The LFC over $R(t,k,\theta)$ is defined similarly.

For the case $s=t=1$, the selection problem reduces to that of selecting the best multinomial cell and Bechofer, Elmagraby and Morse (1959) proposed a fixed-sample-size procedure for this problem. Kesten and Morse (1959) showed that in this case, the slippage configuration $\{p : p_1 = \frac{\theta}{k-1+\theta},\ p_2 = \ldots = p_k = \frac{1}{k-1+\theta}\}$ is the LFC over $R(1,k,\theta)$.

Later, Alam and Thompson (1972) pointed out that the formulation of the problem of selecting the best cell as given by Bechofer, Elmagraby and Morse(1959), is not satisfactory when $2 \leq t \leq k-1$. For studying the case $s=t=k-1$, i.e. for the selection of the least probable cell, Alam and Thompson (1972) proposed to use the preference zone $D(k-1,k,b)$ instead of $R(k-1,k,\theta)$. For detailed discussion on why the zone $R(t,k,\theta)$ is not a suitable preference zone for general $t(>1)$, we refer to Cheng and Hwang (1984).

So, in this paper, for the general selection problem with arbitrary $s, t(s \geq t)$ using the fixed-sample-size procedure, we consider $D(t,k,b)$ as the preference zone. For the particular case $s=t=1$, Bhandari and Bose (1987) obtained the LFC over $D(t,k,b)$ and for the case $s=t=k-1$, Alam and Thompson (1972) showed that the usual slippage configuration is the LFC over $D(k-1,k,b)$.

For the case $s=t$, for arbitrary s and t, Chen and Hwang(1984) studied this problem and gave some interesting conjectures about the LFC over $D(t,k,b)$. They showed why the truth of their main conjecture would be very useful in multinomial selection theory. Their main conjecture is as follows:

Conjecture 1. (Chen and Hwang, 1984). For any sample size n and any preference zone $D(t,k,b)$, a LFC over $D(t,k,b)$ is of the form :

$$p : p_1 = \ldots = p_t = \delta + b, p_{t+1} = \ldots = p_{t+l} = \delta; p_{t+l+1} = \ldots = p_k = 0. \quad \ldots (1.1)$$

for some l, $1 \leq l \leq k-t$.

Later, Chen (1986) considered the case $s \geq t$. He proved half of Conjecture 1 for the special case $s=t$ and posed the other half as an open problem.

To solve this problem, in this paper, we study the general case $s \geq t$ for arbitrary s and t and derive the LFC over $D(t,k,b)$ for the fixed-sample-size procedure, for large values of n. Our result settles Conjecture 1 of Chen and Hwang(1984) in the affirmative, for large n. Another conjecture of Chen and Hwang (1984) is also shown to be true.

Chen and Hwang (1984) showed that the slippage configuration

$$p : p_1 = \ldots = p_t = b + \frac{1-bt}{k},\ p_{t+1} = \ldots = p_k = \frac{1-bt}{k} \quad \ldots (1.2)$$

is always a LFC over $D(t,k,b)$ for $t = 1,2,\ldots,k-1; k = 3,4,\ldots$ and b in $(0,\frac{1}{t})$ if the sample size $n = 1,2,3$ or 5. They also gave counterexamples to show that when $n = 4, k = 3$ and $t = 1$, p as in (1.2) is sometimes a LFC and sometimes not, depending on the value of b. Thus, for a fixed sample size n, there may be no general solution to the LFC over general $D(t, k, b)$, which works for all n. So, in this case, the large sample LFC remains interesting.

In Section 2, we obtain a large sample expansion for the probability of correct selection. In Section 3, this expansion is used to obtain the large sample LFC over $D(t, k, b)$ for arbitrary s and t, $s \geq t$, $t = 1, 2, \dots, k$, $k \geq 3$, b in $(0, \frac{1}{t})$. A number of remarks on the theorems proved in this paper are also given in Section 3. The proofs of the theorems make extensive use of the 'rich-to-poor transfer' technique and some standard results in analysis.

2. The Probability of Correct Selection

Consider a sample of size n drawn from a multinomial distribution with probability vector p. Let the k elements of p be ordered as

$$p_1 \geq p_2 \geq \dots \geq p_t \geq \dots \geq p_s \geq \dots \geq p_k, \sum_{i=1}^{k} p_i = 1.$$

Let x be the vector of ordered cell frequencies and

$$\Omega_x = \left\{ (x_1, x_2, \dots, x_k) : x_1 \geq x_2 \geq \dots \geq x_t \geq \dots \geq x_s \geq \dots \geq x_k; \sum_{i=1}^{k} x_i = n \right\}.$$

Let $PCS(p)$ denote the probability of correct selection at p, when the aim is to select $s(\geq t)$ cells which contain the t best cells, using the fixed sample-size procedure with a sample of size n.

In Theorem 2.1, we obtain a large sample expansion for $PCS(p)$. The proof of the theorem consists of two main parts; first, the expression for $[1 - PCS(p)]^{\frac{1}{n}}$ is reduced to a maximum of a certain function + $o(1)$. Then, this function is maximised using the 'rich-to-poor transfer' technique and the theorem follows.

THEOREM 2.1. *$PCS(p)$ admits the expansion*

$$\log(1 - PCS(p)) = n \log[1 - (s - j_0 + 2)(A - G)] + o(n), \quad \text{as } n \to \infty,$$

where j_0 is the largest integer among $t, t+1, \dots, s$ such that $p_{j_0} > G$, and

$$A = (p_t + \sum_{i=j_0+1}^{s+1} p_i) \frac{1}{s - j_0 + 2}, \quad G = (p_t . p_{j_0+1} \dots p_s p_{s+1})^{\frac{1}{s-j_0+2}}. \qquad \dots (2.1)$$

PROOF. The probability of incorrect selection, or $1 - PCS(p)$, may be expressed in terms of its dominating term as follows :

$$1 - PCS(p) \simeq c \sum_{\Omega_x} \left\{ \frac{n!}{\prod_{i=1}^{k} x_i!} (\prod_{i=1}^{t-1} p_i^{x_i}) p_t^{x_{s+1}} (\prod_{i=t}^{s} p_{i+1}^{x_i}) (\prod_{i=s+2}^{k} p_i^{x_i}) \right\},$$

where c is a constant, $0 < c < \infty$.

$$\simeq c \sum_{\Omega_x} \left\{ \frac{n^{n+\frac{1}{2}}}{\prod_{i=1}^{k} x_i^{x_i+\frac{1}{2}}} (\prod_{i=1}^{t-1} p_i^{x_i}) p_t^{x_{s+1}} (\prod_{i=t}^{s} p_{i+1}^{x_i}) (\prod_{i=s+2}^{k} p_i^{x_i}) \right\}. \quad \ldots (2.2a)$$

Let

$$\Omega_q = \left\{ q = (q_1, \ldots, q_k) : q_i = \frac{x_i}{n} \; \forall i = 1, \ldots, k, x = (x_1, \ldots, x_k) \in \Omega_x \right\}.$$

Then, from (2.2a), after some algebra, we have

$$\begin{aligned} & 1 - PCS(p) \\ \simeq \; & \frac{c}{n^{\frac{k-1}{2}}} \sum_{\Omega_q} \left\{ \frac{1}{(\prod_{i=1}^{k} q_i^{q_i})} (\prod_{i=1}^{t-1} p_i^{q_i}) p_t^{q_{s+1}} (\prod_{i=t}^{s} p_{i+1}^{q_i}) (\prod_{i=s+2}^{k} p_i^{q_i}) \right\}^n \frac{1}{\prod_{i=1}^{k} q_i^{\frac{1}{2}}}. \\ = \; & \frac{1}{\prod_{i=1}^{k} q_i^{\frac{1}{2}}} \cdot \frac{c}{n^{\frac{k-1}{2}}} \cdot \sum_{q \in \Omega_q} \{f(q)\}^n \, g(q) \quad \text{(say)}, \end{aligned} \quad \ldots (2.2b)$$

Now, since $g(q) > 1$

$$\in \left\{ \max_{\Omega} f(q) - \eta \right\}^n \leq \sum_{\Omega_q} \{f(q)\}^n g(q) \;\; \text{for large } n,$$

where

$$\Omega = \left\{ q = (q_1, \ldots, q_k) : \sum_{i=1}^{k} q_i = 1; 0 \leq q_i \leq 1; \; q_1 \geq q_2 \geq \ldots \geq q_k \right\}$$

and $\in$ is the Lebesgue measure of $\{q : f(q) > \max f(q) - \eta\}$ for fixed η, η small.

Let $\Omega_Q = \{q = (q_1, \ldots, q_k) : q_i = \frac{h_i}{n}$, where h_i 's are integers, $0 \leq h_i \leq n$, $i = 1, \ldots, k\}$. Then,

$$\begin{aligned} \sum_{\Omega_q} \{f(q)\}^n g(q) \;\; & \leq \;\; \left\{ \max_{\Omega} f(q) \right\}^n \sum_{\Omega_q} g(q) \\ & \leq \;\; \left\{ \max_{\Omega} f(q) \right\}^n \sum_{\Omega_Q} g(q), \\ & \leq \;\; \left\{ \max_{\Omega} f(q) \right\}^n n^k c_1 \;\; \text{for some constant } c_1, \end{aligned}$$

since $\frac{1}{n^k} \sum_{\Omega_Q} g(q)$ tends to a Riemann integral which is finite.

Hence for large n,

$$\left[\sum_{\Omega_q} \{f(q)\}^n \, g(q)\right]^{\frac{1}{n}} \to \max_{\Omega} f(q),$$

and so from (2.2b),

$$[1 - PCS(p)]^{\frac{1}{n}} = \max_{\Omega} f(q) + o(1). \qquad \ldots (2.3)$$

Define

$$F(q) = \log f(q) + \lambda(\sum_{i=1}^{k} q_i - 1) \qquad \ldots (2.4)$$

where λ is the Lagrangian multiplier. Then, to find the maximiser point of f in Ω, it is enough to find $q^* = (q_1^*, \ldots, q_k^*)$ such that q^* maximizes F and $q_1^* \geq q_2^* \geq \cdots \geq q_k^*$.

Now, from (2.3) and (2.4),

$$\frac{\partial F(q)}{\partial q_{s+1}} \geq \frac{\partial F(q)}{\partial q_s} \quad \text{since } p_t \geq p_{s+1} \text{ and } q_s \geq q_{s+1}.$$

So, F may be increased by increasing q_{s+1} as much as possible. So, from the form of (2.3) and (2.4), after equating the appropriate derivatives of F with respect to the q_i's, to zero, it can be shown that, for some constant a, the maximum of F occurs at

$$q_i^* = ap_i \quad \text{for } i = 1, \ldots, t-1, \qquad \ldots (2.5)$$

$$\begin{aligned} &= ap_{i+1} \quad \text{for } i = t, t+1, \ldots, s-1, \\ &= a(p_t p_{s+1})^{\frac{1}{2}} \quad \text{for } i = s, s+1, \end{aligned}$$

$$= ap_i \quad \text{for } i = s+2, \ldots, k, \qquad \ldots (2.6)$$

if $q_{s-1}^* > q_s^*$ or equivalently, if

$$p_s > (p_t p_{s+1})^{\frac{1}{2}}. \qquad \ldots (2.7)$$

Again, if condition (2.7) does not hold, then it can be shown that $\frac{\partial F}{\partial q_s} \geq \frac{\partial F}{\partial q_{s-1}}$ and then, q^* is given by (2.5), (2.6) and

$$\begin{aligned} q_i^* &= ap_{i+1} \quad \text{for } i = t, t+1, \ldots, s-2 \\ &= a(p_t p_{s+1} p_s)^{\frac{1}{3}} \quad \text{for } i = s+1, s, s-1 \end{aligned}$$

if

$$q_{s-2}^* > q_{s-1}^*, \; i.e., \; if \; p_{s-1} > (p_t p_{s+1} p_s)^{\frac{1}{3}}. \qquad \ldots (2.8)$$

We continue to check successive conditions like (2.7), (2.8) until for some $j_0, t+1 \le j_0 \le s$,

$$p_{j_0+1} < (p_t p_{s+1} \dots p_{j_0+2})^{\frac{1}{s-j_0+1}} \quad \text{and } p_{j_0} > (p_t p_{s+1} p_s \dots p_{j_0+1})^{\frac{1}{s-j_0+2}}.$$

Then, q^* is given by (2.5), (2.6) and

$$\left.\begin{array}{rcl} q_i^* & = & a p_{i+1} \quad \text{for } i = t, t+1, \dots, j_0 - 1 \\ & = & a(p_t p_{s+1} p_s \dots p_{j_0+1})^{\frac{1}{s-j_0+2}} \quad \text{for } i = s+1, s, \dots, j_0 \end{array}\right\} \qquad \dots (2.9)$$

Now, from (2.4), $\frac{dF}{d\lambda} = 0$ implies, on simplification,

$$\begin{array}{rcl} a & = & \left[1 - (p_t + p_{s+1} + \dots + p_{j_0+1}) + (s - j_0 + 2)(p_t p_{s+1} \dots p_{j_0+1})^{\frac{1}{s-j_0+2}}\right]^{-1} \\ & = & \left[1 - (s - j_0 + 2)(A - G)\right]^{-1}, \end{array} \qquad \dots (2.10)$$

where A and G are as in (2.1).

Using (2.3), (2.5), (2.6) and (2.9), after some algebra, it follows that

$$\max_{\Omega} \log f(q) = -\log\ a$$

where a is given by (2.10).

Hence from (2.3), the theorem follows.

3. **Derivation of the LFC**

The following Lemma will be used in the proof of the subsequent theorem. Their proofs are given after some remarks on the results.

LEMMA 3.1. *In a group of m elements, $e_1, e_2, \dots, e_m$, not all equal, if e_i is increased to $e_i + h$, for some h, for all $i = 1, 2, \dots, m$, then the geometric mean of the m elements increases by an amount $\ge h$, for small h.*

THEOREM 3.1. *For the fixed-sample-size selection procedure, with preference zone $D(t, k, b)$, as $n \to \infty$, the limiting form of the LFC is given by the configuration*

$$p : p_1 = p_2 = \dots = p_t = \delta + b > p_{t+1} = \dots = p_{s+1} = \delta > p_{s+2} = \dots = p_k = 0,$$

where $\delta = \frac{1-bt}{s+1}$.

REMARK 3.1. For $s = t$, Theorem 3.1 settles Conjecture 1 of Chen and Hwang (1984) stated in (1.1), in the affirmative, for large samples.

REMARK 3.2. Theorem 3.1 supports the main result in Chen (1986), where half of Conjecture 1 was proved.

REMARK 3.3. Chen and Hwang (1984) made another conjecture as follows:

CONJECTURE IV. (Chen and Hwang (1984)): For each $t = 1, 2, \ldots$, each $k \geq t+2$ and any b in $(0, \frac{1}{t})$, there exists an N_0 such that the configuration in (1.2) is not a LFC over $D(t,k,b)$ if the sample size $n \geq N_0$. Theorem 3.1 proves this conjecture to be also true.

REMARK 3.4. Chen and Hwang (1984) has shown that for some values of n, t, k and b, the slippage configuration is a LFC over $D(t,k,b)$ whereas it is not a LFC for some other combination of values of n, t, k and b. Thus, it seems that there is no general solution to the LFC over general $D(t,k,b)$, which works for all n. So, the large sample LFC remains interesting.

REMARK 3.5. The main theorems of Bhandari and Bose (1987 and 1989) follow from Theorem 3.1 as particular cases.

REMARK 3.6. Chen (1986) has also considered the problem when $s \leq t$ and the goal is to select a subset of size s which contains any s of the t best cells. Results corresponding to theorems 2.1 and 3.1 can be proved along similar lines.

REMARK 3.7 : The expansion in Theorem 2.1 can be used to get the limiting form of the LFC for configurations in preference zones other than the difference zone considered here.

PROOF OF LEMMA 3.1. The geometric mean of $e_i + h, i = 1, \ldots, m$

$$\begin{aligned} &= [(e_1+h)(e_2+h)\ldots(e_m+h)]^{\frac{1}{m}} \\ &= (e_1e_2\ldots e_m)^{\frac{1}{m}}\left[(1+\frac{h}{e_1})(1+\frac{h}{e_2})\ldots(1+\frac{h}{e_m})\right]^{\frac{1}{m}} \\ &\geq (e_1\ldots e_m)^{\frac{1}{m}}\left[1+h\left(\frac{1}{e_1}+\frac{1}{e_2}+\ldots+\frac{1}{e_m})\right)\right]^{\frac{1}{m}} \\ &= (e_1\ldots e_m)^{\frac{1}{m}}\left[1+\frac{h}{H.M}\right]+o(h^2) \end{aligned}$$

where $H.M.$ is the harmonic mean of $e_1, \ldots, e_m$.

$$\begin{aligned} &= (e_1\ldots e_m)^{\frac{1}{m}} + \frac{(e_1\ldots e_m)^{\frac{1}{m}}}{H.M.}.h + o(h^2) \\ &> (e_1\ldots e_m)^{\frac{1}{m}} + h + o(h^2) \end{aligned}$$

since all e_i's are not equal, $i = 1, \ldots, m$, $(e_i, \ldots, e_m)^{\frac{1}{m}} > H.M.$ Hence the Lemma follows.

PROOF OF THEOREM 3.1. From Definition 1.1 and Theorem 2.1, it is clear that to find the LFC in $D(t,k,b)$, it is enough to find $p \in D(t,k,b)$ which maximizes $G - A$, where G and A are as in (2.1) and involves only $p_t, p_{j_0+1}, \ldots, p_{s+1}$.

From (2.1) it follows that

$$\begin{aligned}\frac{\partial}{\partial p_t}(G-A) &= \frac{1}{s-j_0+2}\left[\left(\frac{p_{j_0+1}}{p_t}.\frac{p_{j_0+2}}{p_t}\dots\frac{p_{s+1}}{p_t}\right)^{\frac{1}{s-j_0+2}}-1\right]\\ &< 0, \text{ since } p_t > p_{j+1}\geq\dots\geq p_{s+1} \text{ for all } p \text{ in } D(t,k,b).\end{aligned}$$

Hence $G-A$ is a decreasing function of p_t and so the LFC must be in the following sub-class of $D(t,k,b)$;

$$P=\left\{p: p=(p_1,\dots,p_k), p_1\geq\dots\geq p_k,\ \sum_{i=1}^{k} p_i=1,\ p_t=p_{t+1}+b\right\}.$$

Consider any $p\in P$. Now, by appropriate transfer among the elements of P, we try to maximize $G-A$.

First, by rich-to-poor transfer among $p_{j_0+1},\dots,p_{s+1}$ we make $p_{j_0+1}=\dots=p_{s+1}$. This transfer leaves A unchanged but increases G, thereby increasing $G-A$.

Next, by transferring appropriate amounts to p_1 from each of $p_i,\ i=t+1,\dots,j_0,\ s+2,\dots,k$ we make

$$p_{t+1}=\dots=p_{j_0}=p_{j_0+1} \text{ and } p_{s+2}=\dots=p_k=0,$$

and then by rich-to-poor transfer among $p_1,\dots,p_{t-1}$ we make $p_1=\dots p_{t-1}$. These transfers leave $G-A$ unchanged.

So, the LFC will be in the sub-class

$$\begin{aligned}P_1 &= \left\{p: p=(p_1,\dots,p_k), \sum_{i=1}^{k} p_i=1,\ p_1=\dots=p_{t-1}\geq p_t=p_{t+1}+b>p_{t+1}=p_{t+2}\right.\\ &= \dots=p_{j_0+1}=\dots=p_{s+1}\geq p_{s+2}=\dots=p_k=0\}\subset P.\end{aligned}$$

Finally, for any $p\in P_1$, we may take away an amount y from each of $p_1,\dots,p_{t-1}$ and distribute this amount equally among $p_t,p_{t+1},\dots,p_{j_0+1},\dots,p_{s+1}$, each receving $\frac{(t-1)y}{s-t+2}$, where y is such that after the transfer, $p_{t-1}=p_t$. By Lemma 3.1, this transfer leads to an increase in $G-A$.

Hence the LFC will be in the class

$$\begin{aligned}P_2 &= \left\{p: p=(p_1,\dots,p_k), \sum_{i=1}^{k} p_i=1, p_1=\dots=p_t=p_{t+1}+b>p_{t+1}=\dots=p_{s+1}\right.\\ &\geq p_{s+2}=\dots=p_k=0\}\subset P_1.\end{aligned}$$

From the form of P_2 it is clear that $G-A$ cannot be increased any further.

Thus, the LFC over $D(t,k,b)$ is at $p=(p_1,\dots,p_k)$, given by,

$$\begin{aligned}p_i &= \delta+b \text{ for } i=1,\dots,t\\ &= \delta \text{ for } i=t+1,\dots,s+1\\ &= 0 \text{ for } i=s+2,\dots,k.\end{aligned}$$

where $\delta = \frac{1-bt}{s+1}$. Hence the theorem follows.

Another selection procedure which is widely used in multinomial selection problems is the inverse sampling procedure. This procedure has been considered by Cacoullos and Sobel(1966) for $s = t = 1$, by Chen and Sobel(1984) for $s = t$, and others. An extension of the results of this paper to the case of inverse sampling is currently under investigation.

Acknowledgement. The authors would like to thank the referee for constructive comments which have lead to an improved paper.

References

ALAM, K., AND THOMPSON, J.R. (1972). On selecting the least probable multinomial event. *Ann. Math. Statist.* **43**, 1981-1990.

BECHHOFER, R.E., ELMAGHRABY, S.A. AND MORSE, N. (1959). A single-sample multiple decision procedure for selecting the multinomial event which has the largest probability. *Ann. Math. Statist.*, **30**, 102-119.

BHANDARI, S.K. AND BOSE, A. (1987). On selecting the most likely event. *J. Statist. Plann. Inf.*, **17**, 227-240.

– – –– (1989). Selecting the t-best cells in a multinomial distribution. *Commun. Statist. - Theor. Meth.*, **18 (9)**, 3313-3326.

CACOULLOS, T. AND SOBEL, M. (1966). An inverse sampling procedure for selecting the most probable event in a multinomial distribution. *Proceedings of the 1st International Symposium on Multivariate Analysis,* ed. by P.R. Krishnaish, Academic Press, New York, 423-455.

CHEN, P. (1986). On the least favourable configuration in multinomial selection problems, *Commun. Statist. - Theor. Meth.*, **15(2)**, 367-385.

CHEN, R.W. AND HWANG, F.K. (1984). Some theorems, counterexamples, and conjectures in multinomial selection theory, *Commun. Statist. Theor. Meth.*, **13(10)**, 1289-1298.

CHEN, P. AND SOBEL, M. (1984). Selecting the t best cells of a multinomial using inverse sampling inequalities in Statistics and Probability, *IMS Lecture Notes - Monograph Series* Vol. **5**, 206-210.

KESTEN, H., AND MORSE, N. (1959). A property of the multinomial distribution. *Ann. Math. Statist.*, **30**, 120-127.

MAUSUMI BOSE AND SUBIR KUMAR BHANDARI
INDIAN STATISTICAL INSTITUTE
203 B.T. ROAD
CALCUTTA 700 035
INDIA
e-mail : mausumi@isical.ac.in