Ann. Inst. Statist. Math.
39 (1987), Part A, 325-330

THE k-IN-A-ROW PROCEDURE IN SELECTION THEORY

PINYUEN CHEN

(Received Feb. 13, 1986)

Summary

A k-in-a-row procedure is proposed to select the most demanded element in a set of n elements. We show that the least favorable configuration of the proposed procedure which always selects the element when the same element has been demanded (or observed) k times in a row has a simple form similar to those of classical selection procedures. Moreover, numerical evidences are provided to illustrate the fact that k-in-a-row procedure is better than the usual inverse sampling procedure and fixed sample size procedure when the distance between the most demanded element and the other elements is large and when the number of elements is small.

1. Introduction

The k-in-a-row policy was considered by Kan and Ross [6] in studying the optimal list ordering problem. In the list ordering problem, we are given a set of n elements $e_1, e_2, \cdots, e_n$ which are to be arranged in some order. At each unit of time element e_i is demanded with probability p_i and replaced according to some replacement policy. The k-in-a-row policy makes a replacement when e_i has been demanded k times in a row. Kan and Ross [6] showed that for the move-to-the-front policy the average position of the next element demanded is a monotone decreasing function of k.

In this paper, we apply the k-in-a-row rule to the ranking and selection theory. Let $\pi_1, \pi_2, \cdots, \pi_n$ be n multinomial cells with cell probabilities $p_1, p_2, \cdots, p_n$ where $\sum_{i=1}^{n} p_i = 1$. The ordered values of $p_1, p_2, \cdots, p_n$ are denoted by $p_{[1]}, p_{[2]}, \cdots, p_{[n]}$. The experimenter is interested in selecting the cell with the largest cell probability $p_{[n]}$. The probability requirement of making a correct selection (hereinafter, referred

Key words and phrases: Fixed-sample sized procedure, inverse sampling procedure, k-in-a-row procedure, least favorable configuration, multinomial distribution, ranking and selection.

to as CS) under a proposed procedure R is given by

$$\mathrm{P}(\mathrm{CS}|R) \geqq P^* \tag{1.1}$$

whenever $p_{[n]}/p_{[n-1]} \geqq \delta^*$, where (P^*, δ^*) is preassigned.

Bechhofer, Elmaghraby and Morse [2] studied a fixed-sample size procedure (FSP) and Cacoullos and Sobel [3], motivated by Banach match-box problem (cf. Feller [5]), proposed a sequential type inverse sampling procedure (ISP) for selecting the largest multinomial cell probability. On selecting the multinomial cell with the smallest probability, Alam and Thompson [1] studied a fixed-sample size procedure and Chen [4] studied an inverse sampling procedure. In Section 2 of this paper, we propose the k-in-a-row procedure, which saves memory space and is sequential, and study its least favorable configuration (LFC). In Section 3, we provide the tables for the $\mathrm{P}(\mathrm{CS}|R)$ under LFC and their corresponding expected number of observations required. We also make remarks on the tables and the comparison of our procedure R with FSP and ISP.

2. k-in-a-row procedure R

Procedure R: Observations are taken one at a time and the sampling is terminated when one of the cells π_i has been observed k times in a row. Select the cell π_i as being the best (i.e. the cell with the largest probability $p_{[n]}$).

The least favorable configuration (LFC) is the parameter vector $(p_1, p_2, \cdots, p_n)$ in the parameter space $\{(p_1, p_2, \cdots, p_n) | p_{[n-1]} \leqq p_{[n]}/\delta^*\}$ that will give us the smallest $\mathrm{P}(\mathrm{CS}|R)$. Once we know the form of the LFC, we can compute N, the sample size, required to achieve the probability requirement P^*.

In order to find the LFC for the procedure R, we need the following proposition by Kan and Ross [6].

PROPOSITION 2.1. *Given a sequence of independent multinomial trials—each resulting in outcome i with probability p_i, $\sum_{i=1}^{n} p_i = 1$. Then the probability that a run of k_1 successive trials all resulting in outcome number 1 occurs before any run of k_i successive i outcomes, $i=2, \cdots, n$ equals*

$$\frac{p_1^{k_1}(1-p_1)/(1-p_1^{k_1})}{\sum_{i=1}^{n} p_i^{k_i}(1-p_i)/(1-p_i^{k_i})}. \tag{2.1}$$

From (2.1), the $\mathrm{P}(\mathrm{CS}|R)$ can be written as

$$\text{(2.2)} \qquad P(CS\,|\,R)=\frac{p_{[n]}^k(1-p_{[n]})/(1-p_{[n]}^k)}{\sum_{i=1}^{n} p_{[i]}^k(1-p_{[i]})/(1-p_{[i]}^k)}\,.$$

Before we state and prove the main theorem on LFC, we need the following lemma.

LEMMA 2.1. *The real-valued function* $f(x)=\dfrac{x^k(1-x)}{1-x^k}$ *has a non-decreasing derivative when* $\dfrac{k-1}{k+1}\geqq\dfrac{x-x^k}{1-x^{k+1}}$ *for positive* x *and* $k\geqq1$.

PROOF. The second derivative of $f(x)$ can be easily obtained as

$$\text{(2.3)} \qquad f''(x)=\begin{cases} kx^{k-2}\left[\dfrac{(k-1)(1-x^{k+1})+(k+1)(x^k-x)}{(1-x^k)^3}\right] & \text{for } k\geqq2 \\ 0 & \text{for } k=1\,. \end{cases}$$

Since $x>0$, thus $f''(x)\geqq0$ when $\dfrac{k-1}{k+1}\geqq\dfrac{x-x^k}{1-x^{k+1}}$.

THEOREM 2.1. *The LFC under the procedure* R *is in the following form:*

$$\text{(2.4)} \qquad 0=p_{[1]}=p_{[2]}=\cdots=p_{[h]}\leqq p_{[h+1]}\leqq p_{[h+2]}=\cdots=p_{[n-1]}<p_{[n]}$$

where h *is an integer satisfying* $0\leqq h\leqq k-2$ *and* $p_{[n]}/p_{[n-1]}=\delta^*$.

PROOF. We start with the general configuration

$$\text{(2.5)} \qquad p_{[1]}\leqq p_{[2]}\leqq\cdots\leqq p_{[n-1]}<p_{[n]}$$

where $p_{[n]}/p_{[n-1]}\geqq\delta^*$.

Fix all the $p_{[i]}$'s except $p_{[1]}$ and $p_{[2]}$. Let $x=p_{[2]}$, then $p_{[1]}=1-x-\sum_{i=3}^{n} p_{[i]}$ and $P(CS|R)$ becomes a function of only one variable x. It can be written as

$$P(CS\,|\,R)=\frac{f(p_{[n]})}{f(x)+f\left(1-x-\sum_{i=3}^{n} p_{[i]}\right)+\sum_{i=3}^{n} f(p_{[i]})}\,.$$

Since the derivative $\left[f(x)+f\left(1-x-\sum_{i=3}^{n} p_{[i]}\right)+\sum_{i=3}^{n} f(p_{[i]})\right]'=f'(x)-f'\left(1-x-\sum_{i=3}^{n} p_{[i]}\right)=f'(p_{[2]})-f'(p_{[1]})\geqq0$ by Lemma 2.1, thus $P(CS|R)$ is non-decreasing in x for $\dfrac{k-1}{k+1}\geqq\dfrac{x-x^k}{1-x^{k+1}}$. Consider the following cases:

Case 1. When $k=1$, it is clear that $P(CS|R)$ is non-decreasing in x for all x.

Case 2. When $k=2$, $\frac{k-1}{k+1}=\frac{1}{3}\geqq\frac{x-x^3}{1-x^4}$ for all $x\in(0,1)$.

Case 3. When $k\geqq 3$, $\frac{k-1}{k+1}\geqq\frac{1}{2}\geqq x\geqq\frac{x-x^k}{1-x^{k+1}}$ since $0\leqq x\leqq p_{[n]}$ and $p_{[n]}+x\leqq 1$.

Thus $\frac{k-1}{k+1}\geqq\frac{x-x^k}{1-x^{k+1}}$ for all the possible values of k and thus $P(CS|R)$ is non-increasing in x for all $x\in(0,1)$. Therefore, we can move $p_{[2]}$ toward $p_{[k-1]}$ (simultaneously move $p_{[1]}$ to 0) and won't increase the value of $P(CS|R)$. Repeated the above argument to all the $p_{[i]}$'s. We will finally change the configuration (2.5) to (2.4) and obtain the same or smaller $P(CS|R)$. Thus the LFC must be in the form (2.4).

With the help of the above theorem, it is now easy to find the LFC by considering $P(CS|R)$ as a real-valued function of a single variable $y=p_{[n]}$. The tables of the $P(CS|R)$ under LFC for various k, n and δ^* values are provided in the next section.

3. Tables and remarks

In Table A, we provide the $P(CS|R)$ values under LFC for $n=2$, 3 and 4, $k=2$ to 6, and $\delta^*=2$, 3, 5 and 10. We also calculate the expected number of total observations required under LFC and they are the values in the parenthesis ( ) right next to the corresponding P(CS).

Remark 1. The expected number of total observations required using the proposed procedure R, $E(T)$ under any configuration $(p_1, p_2, \cdots, p_n)$ can be calculated by the formula (cf. Ross [7]):

$$E(T)=\frac{1}{\sum_{i=1}^{n}\frac{p_i^{k_i}(1-p_i)}{1-p_i^k}}. \tag{3.1}$$

By Lemma 2.1, we can prove that $E(T)$ attains its maximum at the configuration

$$p_{[1]}=p_{[2]}=\cdots=p_{[n-1]}=p_{[n]}/\delta^*. \tag{3.2}$$

Although we have not been able to prove that (3.2) is the LFC under procedure R, the numerical evidences show that the LFC for all the cases we considered in Table A is indeed (3.2). Since the calculation of $P(CS|R)$ under the simple formula (2.1) by computer is an inexpensive job, the identification of the exact LFC is only of theoretical interest.

Table A. Exact P(CS|R) and E(T) under LFC for $n=2$, 3 and 4, $k=2$ to 6, and $\delta^*=2$, 3, 5 and 10.

$n=2$

δ^* \ k	2		3		4		5		6	
2	.7619	(2.85)	.8455	(6.02)	.9078	(11.06)	.9483	(18.75)	.9722	(30.30)
3	.8653	(2.69)	.9387	(5.14)	.9752	(8.42)	.9907	(12.73)	.9966	(18.41)
5	.9409	(2.48)	.9834	(4.29)	.9958	(6.41)	.9990	(8.92)	.9998	(11.91)
10	.9828	(2.27)	.9975	(3.63)	.9996	(5.10)	.9999	(6.71)	.9999	(8.48)

$n=3$

δ^* \ k	2		3		4		5		6	
2	.6257	(3.75)	.7623	(10.33)	.8610	(24.73)	.9248	(54.26)	.9612	(113.22)
3	.7714	(3.42)	.8952	(8.12)	.9587	(16.09)	.9850	(29.20)	.9948	(50.81)
5	.9002	(2.96)	.9734	(5.77)	.9939	(9.53)	.9987	(14.62)	.9997	(21.61)
10	.9696	(2.53)	.9958	(4.28)	.9995	(6.31)	.9999	(8.72)	.9999	(11.59)

$n=4$

δ^* \ k	2		3		4		5		6	
2	.5333	(4.66)	.6794	(16.56)	.8038	(50.99)	.8898	(143.35)	.9414	(381.48)
3	.7000	(4.19)	.8600	(12.04)	.9452	(28.36)	.9804	(60.78)	.9932	(125.15)
5	.8607	(3.49)	.9629	(7.68)	.9917	(14.04)	.9983	(23.86)	.9996	(39.24)
10	.9548	(2.83)	.9938	(5.09)	.9992	(7.91)	.9999	(11.53)	.9996	(16.21)

Remark 2. In Table A, we use Theorem 2.1 and formula (2.2) to find the LFC and use formula (2.2) and (3.1) to calculate the P(CS|R) and E(T) under LFC. Suppose that $x=p_{[n]}$, then the upper and lower limits of x are $U=1/(1+(1/\delta^*))$ and $L=1/(1+(n-1)/\delta^*)$ respectively. When x is fixed, we can find the integer h by the inequality

$$1-\frac{p_{[n]}}{\delta^*}\leqq(n-h-2)\frac{p_{[n]}}{\delta^*}+p_{[n]}\,, \tag{3.3}$$

since there is at most one $p_{[h+1]}$ value such that $0\leqq p_{[h+1]}\leqq p_{[n-1]}=\frac{p_{[n]}}{\delta^*}$. Thus we can compute P(CS|R) using (2.2) under the configuration

$$\left(0,\ 0,\cdots,\ p_{[h+1]},\ \frac{x}{\delta^*},\ \frac{x}{\delta^*},\cdots,\frac{x}{\delta^*},\ x\right). \tag{3.4}$$

When we compute Table A, we let the x value run from L to U with each increment 1×10^{-3}. We find that the minimum P(CS|R) always happens when the configuration is at (3.2).

Remark 3. A reasonable criteria for comparing two sequential procedures is the size of E(T) under the same p^*, δ^* and n values. Comparing our Table A with the Table II in Cacoullos and Sobel [3], we can see the procedure R will do a better job for large δ^* and small n. For example, with $p^*=.75$, $\delta^*=2$ and $n=2$, $E(T|R)=2.8571 <3.96=E(T|\text{ISP})$. Suppose one is to compare procedure R with FSP, he can use E(T) and N the sample size respectively. Procedure R also gives better result for large δ^* and small n. For example, with $p^*=.95$, $\delta^*=3$ and $n=3$, $E(T|R)=16.0971<17=N_{\text{FSP}}$ (cf. Table A-3 of Bechhofer, Elmaghraby and Morse [2]). There are also other advantages of Procedure R. It saves the memory space (cf. Kan and Ross [6]). It is also suitable for additive model (that if we require $p_{[n]}-p_{[n-1]} \geqq \delta^*$ instead of $p_{[n]}/p_{[n-1]} \geqq \delta^*$) which neither FSP nor ISP can do. The result of the LFC for additive model under procedure R can be similarly proved to be

$$p_{[1]}=p_{[2]}=\cdots=p_{[n-1]}=p_{[n]}-\delta^* . \tag{3.5}$$

Remark 4. Other goals in ranking and selection (e.g. comparison with a control or a standard, subset selection, partitioning, . . .) can also be achieved by k-in-a-row procedure. With the technique we use in Theorem 2.1, we should be able to obtain the similar result for LFC under these goals.

SYRACUSE UNIVERSITY

REFERENCES

[1] Alam, K. and Thompson, J. R. (1971). On selecting the least probable multinomial event, *Ann. Math. Statist.*, **43**, 1983-1990.

[2] Bechhofer, R. E., Elmaghraby, S. and Morse, N. (1959). A single-sample multiple-decision procedure for selecting the multinomial event which has the highest probability, *Ann. Math. Statist.*, **30**, 102-119.

[3] Cacoullos, T. and Sobel, M. (1966). An inverse-sampling procedure for selecting the most probable event in a multinomial distribution, *Proc. 1st Internat. Symp. Multivariate Anal.* (ed. P. R. Krishnaiah), Academic Press, New York, 423-455.

[4] Chen, P. (1985). Subset selection for the least probable multinomial cell, *Ann. Inst. Statist. Math.*, **37**, 303-314.

[5] Feller, W. (1957). *An Introduction to Probability Theory and Its Applications*, Wiley, New York.

[6] Kan, Y. C. and Ross, S. M. (1980). Optimal list order under partial memory constraints, *J. Appl. Prob.*, **17**, 1004-1015.

[7] Ross, S. M. (1985). *Introduction to Probability Models*, 3rd ed., Academic Press, Orlando, Florida.