Optimal and suboptimal Bayes Procedures for selecting the best category in multinomial distribution

Saad Abed Madhi Kawther Fawzi Hamza
University of Babylon, Iraq
Department of Mathematics

Abstract

This paper deals with multinomial selection problem. Optimal Bayesian sequential and fixed sample size procedures are proposed for selecting the best (i.e largest probability) multinomial cell. These are constructed using Bayesian decision-theoretic approach in conjunction with the dynamic programming technique. Bayes risks are used for comparison between these procedures. Suboptimal Bayesian sequential methods are also considered and their performance is studies using Monte Carlo simulation. Performance characteristics such as the probability of correct selection expected sample size are evaluated assuming a maximum sample size. Single observation sequential rule as well as rule when group of observations are taken and fixed sample size rule are discussed. Some concluding remarks and suggestions for future work are also included.

1. Introduction

During the early fifties, it was pointed out by several researcher that testing the homogeneity of population means or variances is not satisfactory solution to a comparison of performance of several populations. One may wish to either rank them according to their performance or select one or more from among them for future use or future evaluation. These problems are known as ranking and selection problem.

Consider a multinomial distribution which is characterized by k events (cells) with probability vector $\underline{p} = (p_1, p_2, ..., p_k)$, where p_i is the probability of the event $E_i \quad (1 \le i \le k)$ with $\sum_{i=1}^{k} p_i = 1$. Let $n_1 n_2, ..., n_k$ be respective frequencies in k cells of the distribution with $\sum_{i=1}^{k} n_i = m$. Further, let $p_{[1]} \le p_{[2]} \le ... \le p_{[k]}$ denote the ordered values of the $p_i (1 \le i \le k)$. It is assumed that the values of p_i and of the $p_{[j]} (1 \le i, j \le k)$ is completely unknown. The goal of the experimenter is to select the most probable event, that, is the event associated with $p_{[k]}$, also called the best cell. A correct selection, denoted by *CS*, is defined as the selection of the best cell. According to this formulation we have a multinomial-decision selection problem.

The statistical formulation as stated above is typical of many practical problems arising in various fields of applications, that is, there are many situations where the multinomial distribution applies and the goal of practical as well as theoretical interest is to select the category that has the best probability (the best cell).Some applications are as follows.

- In social survey subjects might be or asked whether they agree, disagree or have no opinion about particular political the response most likely to be chosen by a randomly sampled individual.
- In marketing research, these procedures can be used to determine the most popular brand of a given product,
- A manufacturer would like to know which of three potential plant l maximize expected revenue,

- selection of the countries with the largest growth rate.
- selection of the largest elasticities.
- A medical research team conducts a clinical study comparing the success rates of five different drug regiments for a particular disease.

Considerable efforts have been expended to the development of multinomial selection procedures using different approaches. The most popular one is the indifference zone approach (IZ). According to this approach, the selection procedure should guarantee the following probability requirements:

$$p\{CS\} \geq p^* \text{ whenever } p_{[k]} \geq \delta^* p_{[k-1]} \quad \ldots (1.3.1)$$

Where $\{\delta^*, p^*\}$ with $1 < \delta^* < \infty$, $\frac{1}{k} < p^* < 1$ are specified by the experimenter prior to the start of experimentation. $P\{CS\}$ denotes the probability of a correct selection for a given certain selection rule. Using the indifference zone approach described above, the following selection procedures have been suggested in the literature. Bechhofer, et al. (1959) proposed a single-stage procedure for selecting the multinomial event associated with $p_{[k]}$. Calcoullos and Sobel (1966), Alam, et al. (1970), Alam (1971), Alam, et al. (1971), Gibbons, Olkin and Sobel (1977) and Ramey and Alam (1971) have given sequential procedures for the multinomial selection problem using sequential sampling with different stopping rules. A multistage design to select the best cell was studied by Hawang (1982).

Subset selection procedures, where the aim is to select a nonempty subset of cells which contains the best cell with a probability at least equal to a pre-assigned number p^*, are proposed by Gaupta and Nagel (1967). Minimize subset selection rule has been investigated by Berger (1980). Kulkarni (1981) and Bechhofer and Kalkarni (1984) studied procedures that denote use the probability condition required by the indifference zone approach.

Chen (1988) gave an expository survey of subset selection multinomial procedures. Goldsman (1984) first suggested the more general use of this type of procedure to find the simulated system mostly likely to produce the "most desirable" observation on a given trial, when "most desirable" can be almost any criterion of goodness.

The methods described so far do not take into account any prior information before the experimentation, therefore it is worth considering Bayesian approach to the selection problem. Jones and Madhi (1988) proposed some suboptimal sequential schemes for selecting the most probable event using stopping rules based on the difference between the largest and next-to-the largest posterior probabilities. Chick (1997) presented a Bayesian analysis of selecting the best simulated system Inone and Chick (1998) compare Bayesian and frequentist approaches for selecting the best system. Indeed, procedures based on the use of Monte Carlo simulation become desirable alternatives to overcome the difficulties associated with the use of optimal selection procedures.

Some related procedure using simulation, on the selection of events from multinomial distributions have been considered by Miller, Nelson and Reilly (1998) and Mausumi and Subir (1999). Kim and Nelson (2003) provides an advanced tutorial on the construction of ranking and selection procedures for selecting the best simulated system.

The remaining part of the paper is organized as follows.

In section two, optimal (Bayesian) selection schemes, namely the fully sequential and fixed sample size schemes, are derived using Bayesian approach with dynamic programming technique. They are investigated and compared using risks. In section three, Monte Carlo simulation studies have been carried out to evaluate the performance of some suboptimal Bayesian schemes in terms of other performance measures such as the probability of correct selection and the expected sample size. Finally, section four contains some concluding remarks and directions for future work.

2. Bayesian (Optimal) Procedures for Selecting the Best Multinomial Category (Cell)

2.1 Summary

In this section Bayesian (optimal) procedures for selecting the best cell in multinomial distribution are introduced. Bayesian decision- theoretic formulation to the problem is given in subsection 2.1. The stopping risks of linear losses are presented in subsection 2.2. In subsection 2.3, the fully Bayesian sequential selection procedure (BOS) is constructed using the dynamic programming technique in conjunction with Bayesian decision-theoretic formulation. Some numerical results are also presented in this subsection. Subsection 2.4 deals with Bayesian (optimal) fixed sample size procedure (BOF). Data comparisons between the schemes (BOS) and (BOF) are given in subsection 2.5.

2.1 Bayesian decision- theoretic formulation

Before we introduce the Bayesian procedures, we introduce some standards definitions and notations which are needed to construct the procedures. Let $\Omega_k : \{\underline{p} = (p_1, p_2, ..., p_k) : \sum_{i=1}^{k} p_i = 1 \ ; p_i \geq 0\}$ be the parameter space and $D = \{d_1, d_2, ..., d_k\}$ be the decision space where in the following terminal *k*-decision rule:

$d_i : p_i$ is the largest cell probability (i=1, 2, …, k).

That is, d_i denote the decision to select the event associated with the i^{th} cell as the most probable event, after the sampling is terminated.

Suppose the loss function in making decisions d_i, defined on $\Omega_k \times D$, is given as follows.

$$L(d_i, \underline{p}^*) = \begin{cases} k^*(p_{[k]} - p_i) & \text{if } (p_{[k]} \neq p_i) \\ 0 & \text{if } (p_{[k]} = p_i) \end{cases} \quad \ldots (2.1.1)$$

That is the loss if decision d_i is made when the true value of $\underline{p} = \underline{p}^*$. Where k^* is the loss constant, giving losses in terms of cost.

The Bayesian approach requires that we specify a prior probability density function $\pi(\underline{p})$, expressing our beliefs about $\underline{p}$ before we obtain the data. From a mathematical point of view, it would be convenient if $\underline{p}$ is assigned a prior distribution which is a member of a family of distributions closed under multinomial sampling or as a member of the conjugate family. The conjugate family in this case is the family of Dirichlet distribution. Accordingly, let $\underline{p}$ is assigned Dirichlet prior

distribution with parameters $m', n_1', n_2', ..., n_k'$. The normalized density function is given by

$$\pi(\underline{p}) = \frac{\Gamma\left(\sum_{i=1}^{k} n_i'\right)}{\prod_{i=1}^{k} \Gamma(n_i')} \prod_{i=1}^{k} p_i^{n_i'-1}, \text{ where } \quad m' = \sum_{i=1}^{k} n_i' \qquad \ldots (2.1.2)$$

and the marginal distribution for p_i is Beta density

$$f(p_i) = \frac{(m'-1)!}{(n_i'-1)!(m'-n_i'-1)!} p_i^{n_i'-1}(1-p_i)^{m'-n_i'-1}$$

Here $\underline{n'} = (n_1', n_2', ..., n_k')$, are regarded as hyperparameters specifying the prior distribution. They can be thought of "imaginary counts" from prior experience. If N_i be the number of times that category *i* is chosen in *m* independent trials, then

$\underline{N} = (N_1, ..., N_k)$ has a multinomial distribution with probability mass function

$$P_r(N_1 = n_1, N_2 = n_2, ..., N_k = n_k \mid p_1, ..., p_k) = P(\underline{n} \mid \underline{p})$$

$$= \frac{m!}{n_1! n_2! .. n_k!} \prod_{i=1}^{k} p_i^{n_i}, \text{ where } \sum_{i=1}^{k} n_i = m, \ \underline{n} = (n_1, ..., n_k).$$

Since

$$P(\underline{n} \mid \underline{p}) \propto p_1^{n_1} .... \ p_k^{n_k} \text{ and } \quad \pi(\underline{p}) \propto \ p_1^{n_1'-1} .... \ p_k^{n_k'-1},$$

then the posterior is $\pi(\underline{p} \mid \underline{n}) \propto \ p_1^{n_1+n_1'-1} .... \ p_k^{n_k+n_k'-1}$

This is a member of the Dirichlet family with parameters

$n_i'' = n_i' + n_i$ and $m'' = m' + m$ (i=1, …, k).

Hence, the posterior distribution has density function

$$\pi(\underline{p} \mid \underline{n}) = \frac{(m''-1)!}{(n_1''-1)!(n_2''-1)!....(n_k''-1)!} p_1^{n_1''-1} .... p_k^{n_k''-1} \qquad \ldots (2.1.3)$$

with posterior mean $\hat{p}_i = \frac{n_i''}{m''}$ (i= 1, 2, …, k), n_i'' will be termed the posterior frequency in the i^{th} cell. The marginal posterior distribution for p_i is the beta distribution with probability density function

$$f(p_i \mid n_i'') = \frac{\Gamma(m'')}{\Gamma(n_i'')\Gamma(m''-n_i'')} p_i^{n_i''-1}(1-p_i)^{m''-n_i''-1}.$$

Where Γ is gamma function.

2.2 The Stopping Risks

In this section, we derive the stopping risks (Bayes risk) of making decision d_i for linear loss function given in section 2.2. The stopping risk (the posterior expected loss) of the terminal decision d_i when the posterior distribution for $\underline{p}$ has parameters $(n_1'', n_2'', ..., n_k''; m'')$, that is when the sample path has reached $(n_1'', n_2'', ..., n_k''; m'')$ from the origin $(n_1', n_2', ..., n_k'; m')$, denoted by $S_i(n_1'', n_2'', ..., n_k''; m'')$, can be found as follows.

$$S_i(n_1'', n_2'', ..., n_k''; m'') = \underset{\pi(\underline{p}|\underline{n})}{E}[L(d_i, \underline{p}^*)]$$

$$= k^* \left[\underset{\pi(\underline{p}|\underline{n})}{E} (\text{þ}_{[k]}) - \frac{n_i''}{m''} \right] \qquad \ldots (2.2.1)$$

The value of $\underset{\pi(\underline{p}|\underline{n})}{E}[\text{þ}_{[k]}]$ is derived as follows.

$$\underset{\pi(\underline{p}|\underline{n})}{E}[\text{þ}_{[k]}] = \int_0^1 p_{[k]}.g(p_{[k]})dp_{[k]} ,$$

where $g(p_{[k]}) = k\, f(p_{[k]}).\left[F(p_{[k]})\right]^{k-1}$ be the probability density function of the largest order statistics $p_{[k]}$. Let the ordered values of $n_1'', n_2'', \ldots, n_k''$ is $n_{[1]}'' \le n_{[2]}'' \le, \ldots, \le n_{[k]}''$. The marginal posterior probability density function of p_i if $p_i = p_{[k]}$ is

$$f(p_{[k]}) = \frac{(m''-1)!}{(n_{[k]}''-1)!(m''-n_{[k]}''-1)!} p_{[k]}^{n_{[k]}''-1}(1-p_{[k]})^{m''-n_{[k]}''-1} \qquad \ldots (2.2.2)$$

and the cumulative density function is

$$F(p_{[k]}) = \sum_{j=n_{[k]}''}^{m''-1} \frac{(m''-1)!}{j!(m''-1-j)!}.{p_{[k]}}^{j}(1-p_{[k]})^{m''-1-j}$$

then,

$$\underset{\pi(\underline{p}|\underline{n})}{E}(\text{þ}_{[k]}) = \frac{k\left[(m''-1)!\right]^k}{(n_{[k]}''-1)!(m''-n_{[k]}''-1)!}\Bigg\{ \int_0^1 \left[\sum_{j=n_{[k]}''}^{m''-1} \frac{\left[\dfrac{p_{[k]}}{(1-p_{[k]})}\right]^j}{j!(m''-1-j)!} \right]^{k-1} .$$

$$p_{[k]}^{n_{[k]}''}(1-p_{[k]})^{km''-n_{[k]}''-k} \Bigg\}.dp_{[k]}$$

$$= \frac{k[(m''-1)!]^k}{(n_{[k]}''-1)!(m''-n_{[k]}''-1)!} \sum_{j_1=n_{[k]}''}^{m''-1} \sum_{j_2=n_{[k]}''}^{m''-1} \ldots\ldots \sum_{j_{k-1}=n_{[k]}''}^{m''-1} \int_0^1 \left(\frac{p_{[k]}}{1-p_{[k]}} \right)^{j_1+j_2+\ldots+j_{k-1}} .$$

$$\frac{p_{[k]}^{n_{[k]}''}(1-p_{[k]})^{m''-n_{[k]}''-k}}{j_1!(m''-j_1-1)!\ldots j_{k-1}(m''-j_{k-1}-1)!} dp_{[k]}$$

$$= \frac{k[(m''-1)!]^k}{(n_{[k]}''-1)!(m''-n_{[k]}''-1)!} \sum_{j_1=n_{[k]}''}^{m''-1} \sum_{j_2=n_{[k]}''}^{m''-1} \ldots\ldots \sum_{j_{k-1}=n_{[k]}''}^{m''-1} .$$

$$\frac{(j_1+j_2+\ldots+j_{k-1}+n_{[k]}'')!(km''-n_{[k]}''-k-j_1-j_2-\ldots-j_{k-1})!}{j_1!\,j_2!\ldots j_k!(m''-j_1-1)!(m''-j_2-1)!\ldots.(m''-j_{k-1}-1)!}$$

Hence $S_i(n_1, n_2, \ldots, n_k; m) = k^*\Big\{ \dfrac{k[(m''-1)!]^k}{(n_{[k]}''-1)!(m''-n_{[k]}''-1)!} \sum_{j_1=n_{[k]}''}^{m''-1} \sum_{j_2=n_{[k]}''}^{m''-1} \ldots\ldots \sum_{j_{k-1}=n_{[k]}''}^{m''-1}$

$$\frac{(j_1+j_2+\ldots+j_{k-1}+n_{[k]}'')!(km''-n_{[k]}''-k-j_1-j_2-\ldots-j_{k-1})!}{j_1!\,j_2!\ldots j_{k-1}!(m''-j_1-1)!(m''-j_2-1)!\ldots.(m''-j_{k-1}-1)!} -$$

$$\frac{n_i''}{m''} \Big\}. \qquad \ldots (2.2.3)$$

2.3 Fully Bayesian (Optimal) Sequential Procedure

2.3.1 Construction of the Procedure BOS

Fully Bayesian sequential schemes, denoted by BOS, to select the best cell is presented. The Bayesian decision formulation, given in subsection 2.1, in conjunction

with the dynamic programming technique is used to construct this procedure. Before a decision is made, random variables $\underline{x}_1, \underline{x}_2, \ldots$ are observed sequentially where for a given value $\underline{p} \in \Omega_k$ the $\underline{x}_i$'s are independent and identically distributed with a common density function given by

$f(\underline{\varepsilon}_i | \underline{p}) = p_i$ where $\underline{\varepsilon}_i = (y_1, y_2, \ldots, y_k)$ with $y_i = 1$ and $y_j = 0$ if $i \neq j$. The procedure is truncated where a maximum sample size of *N* is given.

At each point $(n_1'', n_2'', \ldots, n_k''; m'')$ in k-dimensional integer space, the optimal decision to stop or continue is made by comparing the stopping risk with the risk of taking one more observation.

At the point $(n_1'', n_2'', \ldots, n_k''; m'')$, let

$S_i(n_1'', n_2'', \ldots, n_k''; m'')$ be the Bayes (stopping) risk of making the terminal decision d_i

$B(n_1'', n_2'', \ldots, n_k''; m'')$ be the risk of taking one further observation and proceeding optimality thereafter, termed the continuation risk.

$D(n_1'', n_2'', \ldots, n_k''; m'')$ be the minimum risk (optimal risk) giving the optimal policy.

At each point, there are k possible transitions $(\underline{n}'' + \underline{\varepsilon}_i; m'' + 1)$ with probability $\hat{\underline{p}}_i$.

Then the dynamic programming equations for the procedure are

$B(n_1'', n_2'', \ldots, n_k''; m'') = c + \sum_{i=1}^{k} \hat{p}_i D(\underline{n}'' + \underline{\varepsilon}_i; m'' + 1)$, where c is constant cost of sampling one observation.

Knowing $S_i(n_1'', n_2'', \ldots, n_k''; m'')$, ($i$= 1, …, k) and $B(n_1'', n_2'', \ldots, n_k''; m'')$, the equation for $D(n_1'', n_2'', \ldots, n_k''; m'')$ is given by

$$D(n_1'', n_2'', \ldots, n_k''; m'') = \min\left[S(n_1'', n_2'', \ldots, n_k''; m''),\ B(n_1'', n_2'', \ldots, n_k''; m'')\right],$$

where

$$S(n_1'', n_2'', \ldots, n_k''; m'') = \min_{i=1,\ldots,k} S_i(n_1'', n_2'', \ldots, n_k''; m'')$$

Suppose that the procedure is truncated at *N* observations, then the dynamic programming equations above are used successively from this end point to the origin to partition the *k* dimensional integer space into stopping and continuation points. Due to the dynamic programming technique of computation it is not known which points are reachable by any simple path starting at $(n_1', n_2', \ldots, n_k'; m')$ until this origin is reached.

The stopping rule of the optimal sequential scheme can be described as follows.

At the point $(n_1'', n_2'', \ldots, n_k''; m'')$

(i) stop sampling and make that terminal decision with smaller risk as soon as

$D(n_1'', n_2'', \ldots, n_k''; m'') = S(n_1'', n_2'', \ldots, n_k''; m'') \leq B(n_1'', n_2'', \ldots, n_k''; m'')$

(ii) If no terminal decision has been reached before *N*, then terminate sampling and take that terminal decision with smaller risk.

(iii) If $D(n_1'', n_2'', \ldots, n_k''; m'') = B(n_1'', n_2'', \ldots, n_k''; m'') < S(n_1'', n_2'', \ldots, n_k''; m'')$

then continue sampling with the population which has smaller continuation risk. The terminal decision is as follows.

At the point $(n_1'', n_2'', \ldots, n_k''; m'')$, we choose decision d_i and select cell i to be the best cell if

$$S_i(n_1'',n_2'',...,n_k'';m'') = \min_{j\neq i} S_j(n_1'',n_2'',...,n_k'';m'')$$

2.3.2 Numerical Results and discussion

Since the optimal overall risk depends on the sample size N, the loss constants k^*, constant cost of sampling one observation (sampling cost) c, and the prior parameters $(n_1',n_2',...,n_k';m')$, therefore it varies as they vary.

In this subsection we present some numerical work to study the effect of these factors, on the optimal overall risk of the optimal scheme BOS under linear loss function.

The numerical results where k=3 and 4 are given in table (2-1). From these tables we note that as the prior increases the optimal overall risk decreases. The tables also show that the optimal overall risk decreases as N increases, for N=2, 3, under different values of c and k^*.

Table (2-2) shows that, the optimal overall risk increases when c and k^* increases for different values of prior for k= 4.

Table (2-1)

The influence of prior information and the sample size on the overall risk of the optimal procedure (BOS), under Linear Loss function, various priors and various N for k=3,4.

K	c, k^*	sample size / prior	2	3
3	c=100 k^*=2000	(1, 1, 1; 3)	196.7361	162.8710
		(1, 1, 2; 4)	175.8961	147.8613
		(0, 0, 2; 2)	237.9121	231.747
		(0, 0, 3; 3)	227.0475	222.6183
		(0, 0, 2; 2)	237.9121	231.747
		(1, 0, 2; 4)	191.4361	127.7342
	c=100 k^*=100	(1, 0, 1; 2)	152.2544	136.9046
		(2, 1, 1; 4)	151.9065	109.5610
		(1,0, 1; 2)	153.2544	136.9046
		(1,0, 2, 3)	152.5242	124.1075
		(1, 0, 1; 2)	153.2544	136.9046
		(1, 2, 1; 4)	151.4222	105.0155
4	c=100 k^*=2000	(1, 1, 0, 0; 2)	283.3326	208.5714
		(1, 2, 0, 0; 3)	185.7144	172.2223
		(0, 0, 1, 1; 2)	216.1564	180.0779
		(0, 1, 1, 1; 3)	161.6808	132.5372
		(0, 0, 1, 0; 1)	316.6667	185.3335
		(0, 1, 1, 1; 3)	161.6808	132.5372
	c=10 k^*=100	(1, 0, 1, 0; 2)	21.66663	19.42857
		(1, 1, 1, 0; 3)	20.35715	19.40476
		(0, 1, 0, 0; 1)	23.33333	18.333340
		(0, 1, 1, 1; 3)	18.08404	12.95852
		(0, 1, 0, 0; 1)	23.33333	18.333340
		(1, 2, 0, 0; 3)	19.52378	17.380940

Table (2-2)

The influence of the loss constant and sampling cost on the optimal overall risk in (BOS), for different prior information, when N=2, 3 and k=4.

k	N=2	$c,\ k^*$	
4	Prior	100, 1000	100, 2000
	(1, 1, 0, 0; 2)	216.6663	283.3326
	(0, 1, 1, 1; 3)	130.8404	161.6808
	(0, 0, 1, 0; 1)	233.3333	316.6667
	N=3	100, 1000	100, 2000
	(1, 1, 0, 0; 2)	194.2857	208.5714
	(0, 1, 1, 1; 3)	129.5852	132.5372
	(0, 0, 1, 0; 1)	183.3334	185.3335
	N=3	50, 1000	100, 2000
	(1, 0, 1, 0; 2)	104.2857	208.5714
	(0, 1, 1, 1; 3)	72.76047	132.5372
	(1, 0, 0, 0; 1)	91.66674	183.3335
	N=3	10, 1000	50, 1000
	(1, 0, 1, 0; 2)	32.28572	104.2857
	(0, 1, 0, 1; 3)	34.03896	90.03896
	(1, 0, 0, 0; 1)	18.333341	91.66674

2.4 The Bayesian (Optimal) fixed Sample Size Procedure (BOF)

In this subsection we present Bayesian optimal fixed (BOF) sample size scheme for selecting the best cell in multinomial population using Dirchelet priors and linear loss function. Fixed sample size means that exactly N observations are taken.

2.4.1 Construction of the Procedure BOF

At the point $(n_1'', n_2'', ..., n_k''; m'')$, where $m'' = m' + m$ and $\sum_{i=1}^{k} n_i'' = m''$, $S_i, (i = 1,...,k)$ denote the stopping risk of taking decision $d_i, (i = 1,...,k)$. The terminal decision rule for BOF is as follows.

Let $S = \min(S_1, S_2, ..., S_k)$, called optimal risk using this procedure.

Take decision d_i if $S_i = S$, $(i = 1,...,k)$

where the stopping risk $S_i, (i = 1,...,k)$ for decision d_i is given by

$$S_i(n_1'', n_2'', ..., n_k''; m'') = \underset{\pi(\underline{p}|\underline{n})}{E} \{L(d_i, \underline{p}^*)\}$$

$$= mc + \underset{\pi(\underline{p}|\underline{n})}{E} \{L(p_{[k]} - p_i)\}. \qquad \ldots (2.4.1.1)$$

If m observations have been taken.

2.4.2 Numerical Results and discussion

Some numerical work has been carried out to investigate the optimal fixed sample scheme (BOF).

The numerical results given in tables (2-3), when k= 4 and 5, show that the Bayes risk using this procedure increases as N increases under different priors. From the same tables we also note that as the prior increases the Bayes risk decreases.

In table (2-4) we observed that the Bayes risk using this procedure (BOF) increases as the sampling cost and loss constant increase for different values of prior and N, when k=3, 5.

Table (2-3)

The influence of prior information and N on the Bayes risk of the optimal procedure (BOF), under Linear Loss function, different priors and various N for k=4,5

c=100, k^*=2000				
K	N	Prior		
		(0, 0, 0, 1; 1)	(1, 0, 0, 1; 2)	(2, 0, 0, 1; 3)
4	2	346.0317	268.1319	235.5957
	3	731.8682	664.4044	319.481
	4	764.4044	713.8524	410.5958
	5	813.8524	775.1185	505.3295
	N	Prior		
		(1, 0, 0, 1; 2)	(1, 0, 1, 1; 3)	(1, 0, 2, 1; 4)
	2	268.1319	235.5957	219.4810
	3	664.4044	613.8524	310.5958
	4	713.8524	675.1185	405.3295
5	N	Prior		
		(0, 1, 0, 0, 0; 1)	(0, 1, 0, 1, 0; 2)	(0, 1, 1, 1, 0; 3)
	2	1351.515	1221.479	1085.959
	3	1675.000	1544.405	1407.918
	4	1904.762	1783.385	165.4.399
	N	Prior		
		(0, 1, 0, 0, 0; 1)	(0, 2, 0, 0, 0; 2)	(0, 3, 0, 0, 0; 3)
	2	1351.515	721.4786	285.9586
	3	1675.000	1144.405	741.251
	4	1904.762	1450.051	1082.97
	N	Prior		
		(0, 1, 0, 0, 0; 1)	(0, 1, 0, 1, 0; 2)	(0, 2, 0, 1, 0; 3)
	2	1351.515	1221.479	685.9586
	3	1675.000	1544.405	1074.584
	4	1904.762	1783.385	1368.685

Table (2-4)

The influence of the loss constant and sampling cost on the Bayes risk of the optimal procedure (BOF) for different prior information when N=2, 3, 4 and *k*= 5.

$c,\ k^*$	N	Prior	
		(0, 1, 0, 1, 0; 2)	(0, 0, 2, 0, 1; 3)
400, 500	2	1050.001	805.8107
	3	1300.000	1208.138
	4	1600.001	1608.957
400, 900	2	1250.001	810.4592
	3	1380.00	1214.648
	4	1600.002	1616.122
$c,\ k^*$	N	Prior	
		(0, 1, 0, 1, 0; 2)	(0, 0, 2, 0, 1; 3)
100, 1000	2	250.0002	201.1621
	3	320.000	301.6276
	4	400.0002	401.7914
400, 1000	2	850.0002	801.1621
	3	1220.000	1201.628
	4	1600.000	1601.791
700, 1000	2	1450.00	1401.162
	3	2120.00	2101.628
	4	2800.00	2801.791

2.5 Data comparisons between the Selection Procedures BOS and BOF

In this subsection, data comparisons are made between the procedures BOS and BOF. We discuss the efficiency of the fully sequential scheme in term of the percent reduction gained due to the use of BOS scheme risk using the formula

$$RD = \frac{Risk\ (BOF) - Risk\ (BOS)}{Risk(BOF)}\% \quad . \qquad \ldots (2.5.1)$$

Table (2-5) shows that for $k = 3, 4$ for different N, different priors and various k^*, c that the percent reduction in risk increases as N increases.

The comparisons above assume that observation cost in all cases remain the same whatever sampling is being used, in practice, an adjustment in cost may be necessary to reflect the ease of use of same of the sampling methods. In practice the factors such as the cost of sampling, ethical considerations, delayed and instantaneous responses etc. may play important roles in choosing the sampling method.

Under the assumption of equal sampling cost and other related factors the fully sequential scheme with a maximum total sample size of N will have a smaller risk for BOS than for BOF. Suppose that there is a fixed cost associated with each sample, in addition to a cost per unit sampled, where the fixed cost is the same irrespective of the sample size. In the case of fixed sample scheme of N observations, the fixed cost is incurred once. For BOS each observation is considered as separate sample, the fixed cost may be incurred up to N times.

Generally speaking, if the observations are very costly and no fixed cost associated with the sampling so that the cost of sampling is a function of the observation only, then BOS is preferable. On the other hand, if the fixed cost associated with sampling stage is the most important and not the cost of observations than the optimal fixed

sample scheme may be preferred. The very slight loss of efficiency of BOF will usually be more compensated for by its greater simplicity of use comparable with BOS in terms of time and computer storage required to output sampling scheme.

Table (2-5)

The effect N on the percentage reduction in risk for different priors for k =3, 4.

K=3	$c=100$, $k^*=100$			
Prior	N	BOF	BOS	RD%
(1, 0, 1; 3)	2	207.8571	153.2544	26.267345
	3	314.359	136.9046	56.4496
(2, 1, 1; 4)	2	204.3561	151.9065	25.665786
	3	303.528	109.5610	63.904153
(1, 0, 2; 3)	2	217.5025	152.5242	29.874737
	3	301.0157	124.1075	41.229577
(1, 2, 1; 4)	2	212.3106	151.4222	28.678926
	3	310.7577	105.0155	66.206629
K=4	$c=100$, $k^*=2000$			
Prior	N	BOF	BOS	RD%
(1, 1, 0, 0; 2)	2	533.3309	283.3326	46.874895
	3	633.33090	208.5714	67.067547
(1, 2, 0, 0; 3)	2	485.7133	185.7144	61.764604
	3	692.9507	172.2223	75.146529
(0, 1, 0, 0; 1)	2	866.6667	316.6667	63.461536
	3	966.6642	183.3335	81.034417
(0, 1, 1, 1; 3)	2	564.4044	161.6808	71.353731
	3	613.8524	132.5375	78.408946

3. Bayesian (Suboptimal) Procedures for Selecting the Best Multinomial Cell Using Monte Carlo Simulation

3.1 Summary

In this section, we consider some Bayesian (suboptimal) sequential schemes for selecting the best cell in multinomial population and their performance is studied using Monte Carlo Simulation methods.

3.2 Description of the MC Studies

In this subsection, we shall illustrate the method of MC simulation as it is applied to our procedures. MC studies have been carried out to investigate the performance characteristics of the proposed procedure such as the probability of correctly selecting the best cell $P(CS)$ and the expected number of observations $E(M)$. Computer programs, which simulate the operations of these procedures were written in Fortran power station. The simulation program performs a large number of runs (t=5000), which are assumed to be independent, in order to obtain MC estimates with high precision. At each run mutually independent multinomial observations are generated by using the assumed probability model under $\underline{p}$ and $\underline{n}$ specified in advance and then the selection procedure is applied. After generating observations from multinomial distribution, we calculated the posterior estimate of p_i as:

$\hat{p}_i = \dfrac{n'_i + n_i}{m' + m}$,$(i = 1,...,k)$ Such that $m = \sum_{i=1}^{k} n_i$, $m' = \sum_{i=1}^{k} n'_i$; … (3.2.1)

$n''_i = n'_i + n_i$ will be termed the posterior frequencies in the i^{th} cell.

In our work, we assumed values of the following quantities;

$\underline{p} = (p_1, p_2,..., p_k)$, $\underline{n'} = (n'_1, n'_2,..., n'_k)$ (prior frequencies) such that $\sum_{i=1}^{k} p_i = 1$ and N. The observed values of performance characteristics are accumulated. At the end of all runs, these accumulated values are divided by number of runs, t, to obtain the MC estimates of the performance characteristics of interest. The simulation computer program was written in FORTRAN power station and run on pentum III. The computer generates the necessary random numbers as input data for the simulation model and analysis the behaviours of the schemes. Listings of the programs are available on request from the author.

3.3 Bayesian (suboptimal) sequential procedures (BSS)

3.3.1 Construction of the procedure (BSS)

These procedures are constructed using the following sampling rules

1- R_1: Fully sequential where observations are taken sequentially one at a time until a terminal decision is reached.

2- $R_2(h)$: Group sequential where observations are taken sequentially in group of (h) observations at each stage N will be a multiple of h. both sampling rules the stopping rule that based on the ratio distance, that is stop sampling at the sample size m when

$\left(\hat{p}_{[k]} / \hat{p}_{[k-1]}\right) \geq \delta_\circ$, … (3.3.1)

where $\delta_\circ (1 < \delta_\circ < \infty)$ is preassigned and $\hat{p}_{[i]}$ is the i^{th} ordered posterior mean, (i=1, …, k). In these cases the best cell is chosen to be that with the largest posterior probability with ties being broken by randomization.

3.3.2 Results and Discussion

The criteria used to judge the performance of the rules are:

$P(CS)$ and $E(M)$. The program generates the necessary random numbers as input data for the simulated model and analyses the behaviour of the scheme. A listing of the programs is given in the appendix (B). The above criteria are calculated in each case from the results of a Monte Carlo Simulation of (5000) runs.

Tables (3-1, 3-2) show that the values of $P(CS)$ and $E(M)$ increases as the value of the parameter $\delta_\circ$ increases, for fully and group sequential schemes. These values are calculated for 4-cells and 5-cells.

From tables (3-3, 3-4) we see that as N increases the values of $P(CS)$ and $E(M)$ are also increases for (fully and group) sequential. However, the fully sequential method R_1, has slightly smaller $P(CS)$ and smaller expected sample sizes compared with the group sequential R_2 for different k-cell and N.

The effect of group size in R_2 is also investigated, some results are presented in tables (3-5, 3-6) for N=100, 120, 200 and group size (1, 2, 4, 5, 10, 20, 25). These tables show that $P(CS)$ and $E(M)$ increases as N increases. Also, we noted that in the same tables the $P(CS)$ and $E(M)$ increase as $\delta_\circ$ increase.

Table (3-1)

The effect of $\delta_\circ$ on the performance characteristics of the schemes using sampling methods (R_1 and R_2), for different N and prior frequencies, when k=4, h=2 and fixed values of $\underline{p} = (p_1,...,p_4)$.

N=100	$\underline{n'} = (4,7,5,3)$; $\underline{p} = (.03,.31,.19,.47)$		
	Performance characteristics		
Sampling Rule	$\delta_\circ$	$P(CS)$	$E(M)$
R_1	1.7	5.644000E-01	50.8046
	1.8	6.012000E-01	57.2916
	1.9	7.466000E-01	75.2932
	2.1	8.404000 E-01	89.266
	2.2	8.462000E-01	91.5954
	2.3	8.748000E-01	95.3192
	2.5	8.986000E-01	98.3284
R_2	1.7	5.996000E-01	54.0848
	1.8	6.124000E-01	59.2932
	1.9	7.946000E-01	80.2772
	2.1	8.420000E-01	89.6812
	2.2	8.546000E-01	92.2388
	2.3	8.854000E-01	96.1596
	2.5	9.010000E-01	98.7028

Table (3-2)

The effect of $\delta_\circ$ on the performance characteristics of the schemes using sampling methods (R_1 and R_2), for different N and prior frequencies, when k=5, h=2 and fixed values of $\underline{p} = (p_1,...,p_5)$.

N=100	$\underline{n'} = (6,9,8,3,7)$; $\underline{p} = (.11,.15,.18,.20,.36)$		
	Performance characteristics		
Sampling Rule	$\delta_\circ$	$P(CS)$	$E(M)$
R_1	1.3	7.504000E-01	25.5298
	1.4	8.988000E-01	42.8194
	1.5	9.358000E-01	54.5398
	1.6	9.696000E-01	68.1960
	1.7	9.822000E-01	77.7882
	1.9	9.832000E-01	90.1912
	2	9.854000E-01	92.7066
R_2	1.3	7.768000E-01	26.9728
	1.4	9.178000E-01	44.1800
	1.5	9.452000E-01	55.6216
	1.6	9.734000E-01	69.2428
	1.7	9.826000E-01	78.7580
	1.9	9.836000E-01	90.1980
	2	9.844000E-01	93.3772

Table (3-3)

The effect of N on the performance characteristics using sampling methods (R_1 and R_2), for different of $\delta_\circ$, prior frequencies and fixed values of $\underline{p} = (p_1,...,p_4)$, when k=4, h=2 and N=50(10)100.

$\delta_\circ = 1.5$	$\underline{n'} = (14,11,16,17)$; $\underline{P} = (.3,.2,.1,.4)$			
Sampling Rule	N	Performance characteristics		
		$P(CS)$	$P(NCS)$	$E(M)$
R_1	50	9.010000E-01	9.900000E-02	40.3086
	60	9.076000E-01	9.240000E-02	46.2154
	70	9.116000E-01	8.840000E-02	51.7206
	80	9.254000E-01	7.460000E-02	57.0542
	90	9.280000E-01	7.200000E-02	62.8632
	100	9.386000E-01	6.140000E-02	67.9200
R_2	50	9.016000E-01	9.840000E-02	40.496800
	60	9.098000E-01	9.020001E-02	46.568000
	70	9.150000E-01	8.500000E-02	52.228000
	80	9.242000E-01	7.580000E-02	58.366400
	90	9.280000E-01	7.200000E-02	64.005600
	100	9.342000E-01	6.580000E-02	68.560000

Table (3-4)

The effect of N on the performance characteristics using sampling methods (R_1 and R_2), for different of $\delta_\circ$, prior frequencies and fixed values of $\underline{p} = (p_1,...,p_5)$, when k=5 and N=50(10)100.

$\delta_\circ = 2.2$	$\underline{n'} = (4,5,6,2,3)$; $\underline{P} = (.11,.18,.15,.20,.36)$			
Sampling Rule	N	Performance Characteristics		
		$P(CS)$	$P(NCS)$	$E(M)$
R_1	50	8.370000E-01	1.630000E-01	49.190800
	60	8.852000E-01	1.148000E-01	58.554800
	70	9.156000E-01	8.440000E-02	67.823200
	80	9.396000E-01	6.040000E-02	77.001400
	90	9.544000E-01	6.040000E-02	86.333200
	100	9.784000E-01	2.860000E-02	95.341600
R_2	50	8.512000E-01	1.488000E-01	49.24000
	60	8.928000 E-01	1.072000E-01	58.665600
	70	9.182000 E-01	8.180000E-02	68.071600
	80	9.478000 E-01	5.220000E-02	77.361600
	90	9.596000E-01	5.220000E-02	86.547200
	100	9.756000E-01	3.120000E-02	95.754800

Table (3-5)

Performance characteristics of the group sequential schemes for k=4, using R_2(h) sampling method for different group and various $\delta_\circ$, and N.

$\underline{n}' = (n_1', .., n_4')$ $\underline{p} = (p_1, .., p_4)$	N=100 $\delta_\circ$	h	Performance characteristics		
			$P(CS)$	$P(NCS)$	$E(M)$
$\underline{n}$ =(4,7,5,3) $\underline{p}$ =(.03,.31,.19,.47)	1.8	1	6.012000E-01	3.988000E-01	57.2916
		2	6.214000E-01	3.786000E-01	59.2932
		4	6.570000E-01	3.343000E-01	63.480
		5	7.672000E-01	2.328000E-01	75.072
		10	8.582000E-01	1.418000E-01	85.152
		20	9.128000E-01	8.720000E-02	89.908
		25	9.274000E-01	7.260001E-02	91.475
	2.3	1	8.748000E-01	1.252000E-01	95.3192
		2	8.854000E-01	1.146000E-01	96.1596
		4	8.966000E-01	1.034000E-01	97.2096
		5	9.046000E-01	9.540001E-01	97.864
		10	9.064000E-01	9.360000E-02	97.654
		20	9.388000E-01	6.120000E-02	99.220
		25	9.394000E-01	6.060000E-02	99.250
	2.5	1	8.986000E-01	1.014000E-01	98.3284
		2	9.010000E-01	9.900000E-02	98.7028
		4	9.048000E-01	9.520000E-01	98.7576
		5	9.140000E-01	8.600000E-02	99.436
		10	9.142000E-01	8.580000E-02	99.456
		20	9.360000E-01	6.400001E-02	99.640
		25	9.422000E-01	5.780000E-02	99.715

Table (3-6)

The effect of $\delta_\circ$ on the performance characteristics using $R_2(h)$ (group sequential schemes) sampling method with various group sizes, $\delta_\circ$ and prior information, k=5.

$\underline{n}' = (n_1', .., n_5')$ $\underline{p} = (p_1, .., p_5)$	N=200 $\delta_\circ$	h	Performance Characteristics		
			$P(CS)$	$P(NCS)$	$E(M)$
$\underline{n}'$=(6,9,8,3,7) $\underline{p}$ = (.11,.15,.18,.20,.36)	1.5	1	9.488000E-01	5.120000E-02	69.3442
		2	9.566000E-01	4.340000E-02	71.0680
		4	9.756000E-01	2.440000E-02	75.2928
		5	9.804000E-01	1.960000E-02	76.3400
		10	9.894000E-01	1.060000E-02	79.7520
		20	9.936000E-01	6.400000E-03	80.2000
		25	9.978000E-01	2.200000E-03	81.5550
	1.6	1	9.820000E-01	1.800000E-02	95.1896
		2	9.866000E-01	1.340000E-02	96.4788
		4	9.904000E-01	9.600000E-03	99.3944
		5	9.910000E-01	9.000000E-03	100.203
		10	9.966000E-01	3.400000E-03	103.586
		20	9.720000E-01	2.800000E-03	104.720
		25	9.992000E-01	6.000000E-04	105.770

3.5 Bayesian (Suboptimal) Fixed Procedures (BSF)

3.5.1 Construction of the procedure (BSF)

In this subsection, we propose Bayesian (suboptimal) fixed scheme (BFS) for selecting the best cell in multinomial population.

In this procedure, a fixed sampling rule (denoted by R_3), where a sample size is taken from the multinomial population; that is we don't need any stopping rule since all *N* observation are taken. We calculate $\hat{p}_i, (i = 1,2,...,k)$ for the procedure and then select the largest cell probability. Let the ordered values of $\hat{p}_1, \hat{p}_2, ..., \hat{p}_k$ is $\hat{p}_{[1]} \leq \hat{p}_{[2]} \leq, ..., \leq \hat{p}_{[k]}$, that is, select the cell associated with the $\hat{p}_{[k]}$. Some numerical results are given is the next subsection.

3.5.2 Numerical Results

In tables (3-8, 3-9), we calculated the probability of correct selection and we see that it increases as *N* increases, so we compare the performance characteristics of this procedure (BFS) with (BSS). These results are obtained for different number of cells, different *N* and prior information.

Table (3-7)

The effect of N on the performance characteristics using Sampling methods (R_1 and R_3) for k=4, fixed prior $\underline{p} = (p_1, ..., p_4)$ and different prior frequencies.

$\underline{n}' = (n_1', .., n_4')$ $\underline{p} = (p_1, .., p_4)$	Sampling Rule	$\delta_\circ$ =1.5	Performance characteristics	
		N	*P*(*CS*)	$E(M)$
$\underline{n}'(5,21,17,27)$ $\underline{p} = (.31,.03,.47,.19)$	R_1 R_3	50	7.658000E-01 8.086000E-01	46.7856 50.00
	R_1 R_3	60	8.418000E-01 8.944000E-01	55.1958 60.00
	R_1 R_3	70	8.810000E-01 9.486000E-01	63.8098 70.000
	R_1 R_3	80	8.958000E-01 9.722000E-01	71.2794 80.000
	R_1 R_3	90	9.012000E-01 9.858000E-01	77.7618 90.000
	R_1 R_3	100	9.066000E-01 9.916000E-01	84.0434 100.000

Table (3-8)

The effect of N on the performance characteristics using sampling methods (R_1 and R_3) for k=5, fixed prior $\underline{p} = (p_1,...,p_5)$ and different prior frequencies.

$\underline{n'} = (n'_1,..,n'_5)$ $\underline{p} = (p_1,...,p_5)$	Sampling Rule	$\delta_\circ$ =1.1	Performance characteristics	
		N	$P(CS)$	$E(M)$
$\underline{n}'(13,25,29,27,31)$ $\underline{p} = (.31,.17,.16,.20,.34)$	R_1 R_3	100	9.878000E-01 9.896000E-01	11.383 100
	R_1 R_3	110	9.882000E-01 9.912000E-01	11.450 110
	R_1 R_3	120	9.888000E-01 9.936000E-01	11.4836 120
	R_1 R_3	150	9.888000E-01 9.976000E-01	11.5474 150
	R_1 R_3	170	9.892000E-01 9.982000E-01	11.574 170

4. Conclusion and directions for future work

4.1 Conclusions

Ranking and selection procedures provides excellent tools for selecting the best of k competing alternatives. In this paper we attempt to apply Bayesian statistical decision theory which leads to a quite different approach to the selection problem as the concepts of loss of taking a certain decision when particular values of the parameters of interest are true, the cost of sampling and some prior information about the parameters of the underlying distributions are involved. Furthermore, since reaching a decision as quickly as possible is desirable it seems sensible to employ sequential technique to achieve the aim. The main property of a sequential procedure are that the sample size required to terminate the procedure is a random variable since it depends on the results of observations and they are economical in that a decision may be reached earlier by sequential procedure than by that using a fixed sample size.

However, optimal sequential sampling needs to use a computer with high speed and large capacity to do the calculations. The large computer storage is necessary to use the recursive formula when the number of stages are large makes computations particularly only for small value of k and N. Therefore, Bayesian suboptimal schemes, which are simple and easy to apply, are proposed.

4.2 Directions for Future Work

Some directions for future work are given as follows:

1- Group sequential sampling can be tried where observations are taken in groups to build Bayesian sequential scheme for the selection problem.
2- The problem of selecting the least probable cell can be attempted.
3- To simplify the formula (2.3.3) we can use stirling's approximation for large factorials.
4- An upper bound for risks may be found using functional analysis.

5- General loss functions may be tried, where linear loss is considered as a special case.

6- In some problems the experimenter might be interested in selecting a subset of the cells including the best cell. In this problem a correct selection is the selection of any subset including the cell with i^{th} largest probability. Bayesian approach can be used to solve such as a problem.

References

1. Alam, K. (1971). On selecting the most probable category Technometrics, vol.13, No.4, pp. 843-850.
2. Alam, K. Seo, K. and Thompson, J. R. (1971). A sequential sampling rule for selecting the most Probable Multinomial event. Annals of the INST. Of statist. Math. 33, No.3, pp. 365- 374.
3. Alam, K. Kenzo, S. and James, R. (1970). A sequential sampling rule for selecting the most probable Multinomial event.
4. Alam, K. and Thampson, J. R. (1972). On selecting the least probable Multinomial Event. Ann. Math. Statist. 43. pp. 1981-1990.
5. Bechhofer, R. E., Elmaghraby, S. and Morse, N. (1959). A single- sample multiple-decision procedure for selecting the Multinomial event which has the largest probability. Ann. Math. Statist., 30, pp. 102-119.
6. Bechhofer, R. E. and Kulkarni, R.V. (1984). Closed sequential which have the largest probabilities. Communications in statistics-Theory and Methods. A13 (24). pp. 2997- 3031.
7. Bellman, R.E. (1957). Dynamic programming. Rand Corporation Research Studies, Princetion, N. J.
8. Berger, R.L. (1980). Minimax subset selection for the Multinomial distribution. Journal of statistical planning and inference 4, pp. 391-402.
9. Calcoullos, T. and Sobel M. (1966). An inverse- sampling procedure for selecting the most probable event in a Multinomial distribution. In P.R. Krishnaiah (ed.), Multivariate Analysis. NewYork. Academic Press. N. Y. pp. 423- 455.
10. Chen, p. (1988) On selecting the best of K systems. An expository Survey of subset selection Multinomial procedures. In proceeding of the winter Simulation conference.
11. Chick (1997) .Selection the best system: A decision- theoretic approach. proceeding of the 1997 Winter simulation conferece,326-333.
12. Dudewicz, E. J. (1976). Introduction to statistics and probability. Holt, Ririehart and Winston.
13. El-Sayyad, G. M. and Freeman, P. R. (1973). Bayesian sequential estimation of a Poisson process rate. Biometrika, 60, pp. 289- 296.
14. Freeman, P.R. (1970). Optimal Bayesian sequential estimation of the median effective dose. Biometrika, 57, 1, pp. 79- 89.
15. Freeman, P.R. (1972). Sequential estimation of the size of a population Biometrika, 59, 1, pp. 9-17.
16. Gibbons, J. D., Olkin, I. and Sobel, M. (1977). Selecting and ordering populations: A new Statistical methology. John Wiley and Sons, N.Y.
17. Goldsman, D. (1984). On selecting the best of K systems: An expository Survey of indifference- zone Multinomial procedures. In proceedings of the winter simulation conference.

18. Gupta, S.S. and Nagel, K. (1967). On selecting and ranking procedures and order statistics from the Multinomial distribution. Sankhya, ser. B, vol. 29, parts 1 and 2, pp. 1-34.
19. Gupta, S. S. and Panchapkesan, S. (1979). Multiple decision procedures: Theory and methodology of selecting and ranking populations. John Wiley and Sons, New York.
20. Hawang, F. K. (1982). A multistage selection scheme for the most probable event. Americal Journal of Mathematical and Management Sciences, vol. 2, No.1, pp. 3-12.
21. Horrace, W. C. (2002). Selection procedures for order statistics in Empirical Economic studies, 1-26.
22. Inone,kand s.e.chick (1988).comparison of Bayesian and frequentist assessments of uncertainty for selecting the best system .proceeding of the 1998 Winter simulation conference,727-734.
23. Jones, p.w. and Madhi, S.A (1988). Bayesan sequential methods for choosing the best Multinomial cell: Some simulation results. Statistical paper 29, 125-132.
24. Kim, S.H., and Nelson, B.L, (2003). Selecting the best system. Proceedings (2003). Winter simulation conference, Georgia Institute of Technology, North Western Universi
25. Kulkarni, R. V. (1981). Closed adaptive sequential procedures for selecting the best of $k \geq 2$ Bernoulli populations ph. D. Dissertation, Cornell University, Ithaca, N.Y.
26. Mausumi, B. and Subir, K.B. (1999). Selecting the *t* Best cells of a Multinomial distribution. The Indian Journal of statistics, volume 61, series A, pt 1, pp. 139-147, Calcutta.
http://202.54.54.1471/search/61a1/61a16926.pdf
27. Miller, J.O., B.L. Nelson. And C.H. Reilly. (1998). Efficient multinomial selection in simulation. Naval Research Logistics 45: 459-489.
28. Ramey, T.J. and Alam, K. (1979). A Sequential procedure for selecting the most probable Multinomial event. Biometrika, 66, 1, pp. 171-173.
29. Ramey,Jr,J.T (1974). Selecting the most probable event . Dept.of mathematical sciences,Clemson university.
30. Ranking and selection: a comprehensive survey on the theory, Advanced Simulation.
http://www.bilkent.edu.tr/~goren/simulation.htmL

الملخص

يتناول هذا البحث مشكلة الاختيار في مجتمع متعدد الحدود. اقترحت اجراءات مثلى بيزنية متسلسة وذات عينة ثابتة لاختيار افضل خلية متعدد حدود (اي لها اكبر احتمال). تم بناء هذه الاجراءات باستخدام منهج نظرية القرار البيزني مع طريقة البرمجة الديناميكية. استخدمت الخطورة البيزنية للمقارنة بين هذه الاجراءات. درست ايضاً طرق متسلسلية بيزنية مثلى جزئياً واستخدمت محاكاة مونتو كارلو لتقييم كفاءة هذه الطرق المثلى جزئياً. تم تقييم خصائص الانجاز كاحتمال الاختيار الصحيح والحجم المتوقع لعدد المشهدات على اقتراح حجم العينة الاعلى. نوقشت بعض طرق المعاينة مثل قاعدة اخذ المشاهدات واحدة بعد الاخرى او على شكل مجاميع او على شكل حجم عينة ثابت وتضمن البحث ايضاً بعض الملاحظات الختامية والمقترحات لاعمال مستقبلية.