جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة بابل/ كلية التربية

اجـراءات بيـزنية لمشكلة الاختيار في مجتمع متعدد الحدود

من قبل

كوثر فوزي حمزة السلطاني

رسالة مقدمة الى
قسم الرياضيات- كلية التربية- جامعة بابل
وهي جزء من متطلبات نيل درجة الماجستير في علوم الرياضيات

بإشراف

أ. م. د. سعد عبد ماضي

٢٠٠٦م ١٤٢٧هـ

REPUBLIC OF IRAQ

AND SCIENTIFIC RESEARCH

UNIVERSITY OF BABYLON

COLLEGE OF EDUCATION

Bayesian Procedures for Multinomial Selection Problem

By

Kawther Fawzi Hamza Al-Sultani

A THESIS

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS, COLLEGE

OF EDUCATION, UNIVERSITY OF BABYLON IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN

MATHEMATICS

Supervised by

Assist.Prof.Dr. Saad Abed Madhi

June ٢٠٠٦

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

﴿يَا أَيُّهَا الَّذِينَ إِمَنُواْ اذْكُرُوا اللَّهَ ذِكْرًا كَثِيرًا *
وَسَبِحُوهُ بُكْرَةً وَأَصِيلًا * هُوَ الَّذِي يُصَلِّي عَلَيْكُمْ
وَمَلَائِكَتُهُ لِيُخْرِجَكُم مِنَ الظُّلُمَاتِ إِلَى النُّورِ
وَكَانَ بِالْمُؤْمِنِينَ رَحِيمًا﴾

صدق الله العلي العظيم

سورة الاحزاب

اية ٤١-٤٣

الإهداء

الــى كل عين تتوضأ ببريق السماء وترجو نور المستوحشين في الظلم

الــى من استجاب الله دعائهم في صلاة الليل

أبـــي أمـــي

الــى ائلتي، التي ضحت براحتها من اجلي وقدمت
عائـــلــة عمـــي

الــى رفـاقــي في حيــاتـــــي

أخـوتـي عمـامـي

الــى كل من علمـني ولـو بشـطر كلمة

أسـاتذتي الافاضل

الخـلاصـة

في هذه الرسالة، اعطيت اجراءات بيزينية باختيار الخلية ذات الاحتمال الاكبر، التي تدعى (الخلية الاكثر احتمالاً) او الخلية الافضل، من توزيع متعدد الحدود الذي يحتوي على k من التصنيفات (الخلايا). لقد استعمل منهج القرار النظري لاشتقاق اجراءات مثلى تسلسلية وذات حجم ثابت لاختيار الخلية الافضل. هذه الاجراءات مثلى بمعنى انها تقلل معدل الخطورة نسبةً الى توزيعات قبلية (سابقة للمعاينة) على احتمالات الخلايا، انها تحديداً عائلة توزيعات درشلت. ايضاً تم دراسة بعض الاجراءات البيزينية المثلى جزئياً تحت قواعد معاينة مختلفة وقد قيمت خصائص انجازها باستخدام محاكاة مونتو كارلو. اخيراً تم عرض بعض الملاحظات الختامية والعمل المستقبلي.

ACNOWLEDGEMENT

First of all, I would like to thank for Allah who enabled me to reach this stage of education.

I'd like to express my gratitude to my supervisor Assist.Prof. Dr. Saad Abed Madhi for continuous help and advice during the research work.

I'd like to thank my father, for all his encouragement and help. His love and support is much appreciated. I would like to extend my love and deepest thanks, without him, I might have given up on this endeavor completely.

I'd like thank for my uncle's family to help, give me the strength and dedication to succeed, for which I will be forever grateful.

Also, I'm indebted to the dean of the college of education, Babylon University, Assist. Prof. Dr. Lua'y Abdul- Hani Al-Suwidi and to all the staff of the department of the Mathematics especial the head of department of mathematics, Dr. Iftichar Mudhar Talib, for their cooperation.

I would like to express my special thank to my brothers, my mother and my friends for their continuous help and encouragement.

SUPERVISOR CERTIFICATION

I certify that this thesis was prepared under my supervision at the Department of Mathematics/ College of Education/ University of Babylon, as a partial fulfillment of the requirements for the degree of Master of Science in Mathematics.

Signature:

Name: ***Dr. Saad Abd Madhi***

Date: / / ٢٠٠٦

RECOMMENDATION OF THE HEAD OF THE DEPARTMENT

In view of the available recommendations, I forward this research for debate by the Examining Committee.

Signature:

Name: Dr. Iftichar Mudhar Talib

Date: / / ٢٠٠٦

ABSTRACT

In this thesis, Bayesian procedures are given for selecting from a given multinomial distribution with k categories (cells), the cell with the largest probability, which is called the "most probable" or (the best cell). Decision-theoretic approach is used to derive optimal sequential and fixed sample size procedures to select the best cell, they are optimal in the sense that they minimize the average risk with respect to certain prior distributions on the cell probabilities, namely the family of Dirichlet distributions. Some suboptimal Bayes procedures under various sampling rules are also considered and their performance characteristics are evaluated using Monte Carlo simulation. Finally some concluding remarks and future work have been presented.

LIST OF CONTENTS

NOMENCLATURE

The following symbols are generally used through out the text. Others are defined as when used

Symbol ***Description***

R&S	Ranking and Selection.
n_i	The number of occurrences in category *i.*
n_i'	The number of prior frequency in the i^{th} cell.
n_i''	The number of posterior frequency in the i^{th} cell.
p_i	The parameter (the probability of cell i)
$\hat{p}_i$	The posterior mean of probability p_i
$k*$	The loss constant.
c	The sampling cost (constant cost of sampling one observation).
$S(n_1'', n_2'', \ldots; m'')$	The stopping risk.
$B(n_1'', n_2'', \ldots; m'')$	The continuation risk.
$D(n_1'', n_2'', \ldots; m'')$	The minimum risk (optimal risk).
BOS	Bayesian optimal sequential scheme.
BOF	Bayesian optimal fixed sample size.
RD%	Percent reduction in risk.
BSS	Bayesian suboptimal sequential scheme.
BSF	Bayesian suboptimal fixed sample size.
MC	Monte Carlo simulation.
P(*CS*)	Probability of correct selection.
E(*M*)	Expected of sample size.

M Number of observations

١.١ Prelude

In many practical situations the experimenter is confronted with the problem of choosing the best one, or the best few, of a number of populations or categories or ranking them according to their performance. For example different treatments (drugs) for a specific disease or several candidates for certain position.

The classical approach to such a problem is to test the homogeneity (null) hypothesis $H_{\circ}: \theta_1 = \theta_2 .... = \theta_k$, where $\theta_1, \theta_2 ...., \theta_k$ are the values of the parameters for these populations[٢٣]. But in many practical situations, when the hypothesis is rejected, one wants to know in which of a number of possible ways the actual situation differs from the one postulated by the null hypothesis. Therefore, there are methods to solve this problems, for example, the least significant differences based on the t- test has been used in the past to detect differences between the true unknown means of different varieties and thereby choosing the population which is the "best", say, the one with the largest mean. But this method is indirect, less efficient and does not provide an overall probability of a correct selection[٢٠],[١١]. Also, the multiple comparison procedures are techniques for constructing simultaneous confidence intervals for differences between population parameters[٥].

The ranking and selection (R&S) approach is different. It asks given the data on the distribution what is the probability that we can correctly rank them from worst to best? What is the probability choose the best population (perhaps the one with the largest population mean) or at least populations?[١٦]

R&S procedures are statistical methods specifically developed to select the best system or a subset that contains the best system design from a set of *k* competing alternatives [٣٧], [٢٥] .

Throughout this thesis we shall assume :

١- There is prior knowledge regarding the parameters of interest.

٢- The procedures are truncated, that is there is a maximum number of observations that can be taken during experimentation.

٣- The outcomes of the observations are independent and the probabilities for these outcomes remain constant from observation to observation.

١-٢ Statement of the problem and goal

Before discussing and stating the multinomial selection problem, we first develop the multinomial distribution model.

Consider an arbitrary experiment or process where each outcome is classified into one of *k* possible mutually exclusive possibilities, which we call categories (or cells). Let $p_i, o \le p_i \le 1$, denote the probability of an outcome in the i^{th} category for *i*=١, ٢, …, *k*. Since the *k* categories are mutually exclusive and exhaustive, the probabilities must sum to one, that $\sum_{i=1}^{k} p_i = 1$. If an experiment of this type is repeated a fixed number of times, say *m* then we can count the number of times an outcome is classified in each category. Let n_i denote the number of occurrences for category *i*. Since each outcome must be

in one of the *k* designated categories, we have $\sum_{i=1}^{k} n_i = m$. The joint probability distribution of this set of outcomes is given by the multinomial distribution, specifically,

$$f(n_1, n_2, ..., n_k) = \frac{m!}{n_1! n_2! .... n_k!} \prod_{i=1}^{k} p_i^{n_i} \qquad \ldots (١.٢.١)$$

For any specified integer values of the components of the vector $(n_1, n_2, ... n_k)$ and the parameters $p_1, p_2, ... p_k$, which sum to ١, equation (١.٢.١) gives the probability of exactly n_i outcomes in category *i*, jointly for all i ,(*i*= ١, ٢, ..., *k*)[١٦].

A single observation in a multinomial model is regarded as a vector with *k* components, of which exactly one is equal to ١ and all the others are equal to ٠. The position of the components with the ١ indicates which category describes the outcome of that observation. For example, if *k*=٤ and the observation vector (٠, ١, ٠, ٠) the outcome of this single observation is category ٢. In general, if the i^{th} component is ١, this indicates that the outcome is in the category labeled *i*. The value of n_i is then the sum over all *m* vectors of the number of ١'s in the i^{th} component.

Now, our problem can be formulated as follows:

Consider a multinomial distribution which is characterized by *k* events (cells) with probability p_i associated with the i^{th} cell (*i*=١, ٢,...,*k*). Let $p_{[1]} \le p_{[2]} \le ... \le p_{[k]}$ denotes the ordered values of $p_1, ..., p_k$, the "most probable" event, also called the best cell, is defined to be that cell which has

the probability $p_{[k]}$ associated with it. Nothing is known about the value of $p_i's$, that is we do not know which cell associated with $p_{[k]}$. According to this formulation we have a multiple-decision problem.

Our goal in this thesis is to solve such problem. For this purpose we propose a Bayesian sequential and fixed sample size designs using decision-theoretic approach with dynamic programming or performing some Monte Carlo simulation studies.

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). For examples [٣٩]. Sample of blood is taken from each of large number of students, and each student is classified as type A, type B, type AB or type O. The goal is to identify the blood type that occurs most frequently among these students or, five candidates are running for an office in a certain high populated city. A random sample of registered voters is taken and each person is asked to name the one candidate he prefers. The goal is to determine which candidate has the greatest support among the potential voters. An other application to the identification of the most popular television program has been noted in Dedwicz [١٢].

١.٣ Literature Survey

Considerable efforts have been expended to the development of multinomial selection procedures, that is to solve the problem of selecting the best cell in multinomial distribution. Some of these used indifference zone

approach. According to this approach, the experiment has to define a measure of distance δ as the ratio of $p_{[k]}$ to $p_{[k-1]}$, that is

$$\delta = p_{[k]} / p_{[k-1]} \quad \text{where } k \geq ١ \qquad \ldots (١.٣.١)$$

and specified two predetermined constants p^*, δ^* such that

$$\frac{1}{k} < p^* < 1 \quad \text{and} \quad 1 < \delta^* < \infty$$

the numbers δ^* is the smallest ration δ that the experimenter feels worthy of consideration, and p^* is the smallest probability of correctly choosing the best cell that the experimenter allows for any values of ($p_1, p_2, \ldots p_k$) that satisfy $\delta \geq \delta^*$ for some specified $\delta^* > 1$.

Let $P(\mathrm{CS})$ denotes the probability of a correct selection. for a selection rule SR, the experimenter then requires

$$P(\mathrm{CS}/\mathrm{SR}) \geq p^* \text{ whenever } \delta \geq \delta^* \qquad \ldots. (١.٣.٢)$$

Define $\Omega_{\delta^*} = \left\{ \underline{p} \mid \delta \geq \delta^* \right\}$. Ω_{δ^*} is called as the preference zone and its complement as the indifference zone. The experimenter wishes to minimize the expected sample size needed to satisfy (١.٣.١). Since the probability of a correct selection and the expected sample size is dependant on the true value of the vector $\underline{p} = (p_1, p_2, \ldots, p_k)$, the experimenter needs to find the vector $\underline{p}^* \in \Omega_{\delta^*}$ that satisfies the formula below

$$P(\mathrm{CS} \mid \underline{p} = \underline{p}^*) = \min_{\underline{p} \in \Omega_{\delta^*}} P(\mathrm{CS} \mid \underline{p}) \qquad \ldots (١.٣.٣)$$

this vector $\underline{p}^*$ is called the least favorable configuration.

Using the formulation described above, the following sampling procedures have been suggested:

Bechhofer, Elmagharby and Morse (١٩٥٩) first posed the multinomial selection problem. They have considered, in this paper, a single stage selection procedure in which a sample of size *m*, where *m* is a predetermined integer, is drawn from multinomial population. The best cell is taken to be the cell with the largest sample count with ties broken by randomization.

Calcoullos and Sobel (١٩٦٦), Alam (١٩٧١) and Ramey and Alam (١٩٧٩) proposed and investigated sequential schemes where observations are taken one at a time. In the sequential sampling of Calcoullos and Sobel, called inverse sampling, sampling is terminated when the count in any cell is N^*, when N^* is a predetermined integer. The cell with count N^* is chosen as the best cell. In the second paper, sampling is terminated when the difference between the largest cell count and the second largest cell count is r^*, where r^* is a predetermined integer. The cell which has the highest count when sampling is terminated is selected as the best cell.

In the third paper, sampling is terminated when the count in any cell is N^* or when the difference between the largest cell count and the second largest cell count is r^*, where r^* and N^* are predetermined positive integers and $r^* \leq N^*$.

Alam, Kenze and James (١٩٧٠) suggested a sequential sampling rule, where a random number of observations is taken from the multinomial distribution at each stage of sampling, and this number is distributed according to a Poisson distribution with mean λ. Sampling is stopped when the count in any cell is greater than or equal to a given positive integer N^*. Single-stage

scheme for selecting the worst cell (the cell with the smallest probability), using the indifference zone approach, were studied by Alam and Thompson (١٩٧٢) and Gibbons et al. (١٩٧٧).

Hawang (١٩٨٢) suggested a multistage scheme for the problem.

Subset selection schemes, where the aim is to select a nonempty subset of cells which contains the largest or smallest cell probability with a probability at least equal to a pre-assigned number, p^*, have been proposed by some authors such as Gupta and Nagel (١٩٦٧) and Berger (١٩٨٠) [١١].

Kulkarni (١٩٨١) proposed a closed sequential scheme without referring to the probability requirements given above. Bechhofer and Kulkarni (١٩٨٤) studied this procedure using numerical comparisons.

Goldsman (١٩٨٤)[٣٤] 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, where "most desirable" can be almost any criterion of goodness.

Some related procedure on the selection of events from multinomial distributions have been considered by Miller, Nelson and Reilly (١٩٩٨) and Mausumi and Subir (١٩٩٩).

However, in many situations, we may have some information about the unknown parameters prior to the experimentation; this knowledge may come from past experiences. Therefore, it is worth considering the Bayesian approach to the problem. Jones and Madhi (١٩٨٨) proposed some suboptimal sequential schemes for selecting the most probable event using stopping rule based on the difference between the largest and next - to- the largest posterior probabilities.

In this thesis we suggested an optimal Bayesian sequential selection scheme for *k*- variate multinomial distribution and compared it with an optimal fixed sample size scheme.

However, some computational and analytical complexities are associated with the implementation of such optimal schemes, particularly for large number of cells and large sample sizes. Therefore from practical points of view, it is preferable to find schemes which easy and simple to implement and which has some good characteristics. Therefore Bayesian suboptimal selection schemes for the multinomial selection problem are investigated using Monte Carlo simulation with stopping rule based on the ratio of the largest and next - to- largest posterior probabilities.

١.٤ Some Definitions and Concepts

This section presents some definitions and concepts which are useful in constructing the proposed selection procedures.

١.٤.١ Bayesian Decision approach

١. Prior and Posterior Distributions [٤١]

Bayesian statistics, (Named for the Revd Thomas Bayes (١٧٠٢- ١٧٦١), and a mateur ١٨th century mathematician), represents a different approach to statistical inference. Data are still assumed to come from a distribution belonging to a known parametric family. However, whereas classical statistics considers the parameters to be fixed but unknown, the Bayesian approach

treats them as random variables in their own right. Prior beliefs about ϑ are represented by the prior distribution, with a prior probability density (or mass) function, $\pi(\theta)$. The posterior distribution has posterior density (or mass) function, $\pi(\theta \mid x_1, x_2, ..., x_n)$, and captures our beliefs about ϑ after they have been modified in the light of the observed data.

By Baye's formula, for the density function

$$\pi(\theta \mid x_1, x_2, ..., x_n) = \frac{f(x_1, x_2, ...x_n \mid \theta)\pi(\theta)}{\int_{\Omega} f(x_1, x_2, ...x_n \mid \theta)\pi(\theta)d\theta}$$

The denominator of the above equation does not involve ϑ and so in practice is usually not calculated. Bayes rule is written,

$$\pi(\theta \mid x_1, ..., x_n) \propto f(x_1, ..., x_n \mid \theta)\, \pi(\theta)$$

In fact the actual choice of $\pi(\theta)$ depends upon the experimenter and the information and experience available to him at the time of doing the experiment. A prior which contains no information about ϑ is called non-informative prior or (vague prior). Mathematical and computational difficulties may arise from using some prior distributions.

A reasonable method of overcoming these difficulties is to use a particular class of prior distributions. This class of prior has been termed as natural conjugate priors in Raffa and Schlaifer [٣٤], for example the Beta distribution is a natural conjugate prior to Binomial distribution.

٢. Elements of a Bayesian Decision Problem

In this point we present elements of Bayesian decision theoretic that we needed in this thesis.

i The basic elements of statistical decision theory are:[٣٠],[١٧]

١. A parameter space $\Omega = \{\underline{\theta}\}$, which may be vector- valued, of possible states of nature,

٢. An action space $A = \{a\}$ of the possible courses of action,

٣. A loss function $L(\underline{\theta}, a)$ representing the loss incurred when an action *a* is taken and the state of nature is $\underline{\theta}$,

٤. An observable r.v. $\underline{X}$, which may be vector- valued, defined on a sample space $\chi = \{\underline{x}\}$ such that when $\underline{\theta}$ is the true state of nature, $\underline{X}$ has probability distribution $f(\underline{x}, \underline{\theta})$,

٥. A decision space $D = \{d(\underline{x})\}$ of possible decision functions defined on χ that maps χ into the action space A,

٦. A prior probability distribution π is defined on the space Ω.

ii Risk [٣٠],[٢٢]

For given decision function *d*, the loss function may be written as

$L\{\underline{\theta}, d(\underline{x})\}$,

Since our action *a* depends on the particular sample data $\underline{x}$ that we observe. Thus, we see that the loss is a r.v. and depends on the sample outcome. Therefore, let us define the *risk* to be the expected value of the loss

function. That is, the risk $R(\underline{\theta},d)$ is a function of $\underline{\theta},d$, and the loss function L such that

$$R(\underline{\theta},d) = E\left[L\{\underline{\theta},d(\underline{x})\}\,|\,\underline{\theta}\right]$$

$$= \int_{\chi} L\{\underline{\theta},d(\underline{x})\}\, f(\underline{x}\,|\,\underline{\theta})d\underline{x}.$$

iii Bayes Risk [٣٠],[٢٢]

The Basyes risk of a decision d is the expected value of the risk $R(\underline{\theta},d)$ with respect to the prior distribution π on *Ω;*

Namely,

$$r(\pi,d) = E\left[R(\Theta,d)\right]$$

$$= \int_{\Omega}\int_{\chi} L\{\underline{\theta},d(\underline{x})\}\, f(\underline{x}\,|\,\underline{\theta})\pi(\underline{\theta})d\underline{x}d\underline{\theta}.$$

The Bayes risk is of value as it sets up a linear ordering on the decision space D. (i.e. the decision maker prefers decision function $d_١$ to $d_٢$ if it has smaller Bayes risk).

iv Bayes Decision Function [٣٠]

Since the Bayes risk sets up a linear ordering on the decision space D, the search begins for the decision function d which minimizes the Bayes risk for a specified prior distribution π, if such a decision function say $d_{\pi}(\underline{x})$, exists it is known as the *Bayes decision function* and its associated Bayes risk

$$r(\pi) = r(\pi, d_\pi) = \min_{d \in D} r(\pi, d),$$

is known as the *minimum Bayes risk.*

٣. Construction of Bayes Decision Functions[٣٠]

Provided that the order of integration can be reserved, which we shall assume here, the Bayes risk becomes

$$r(\pi, d) = \int_{\chi} f(\underline{x}) \left\{ \int_{\Omega} \{(\underline{\theta}, d\,\underline{x})\} \pi(\underline{\theta} \mid \underline{x}) d\underline{\theta} \right\} d\underline{x} \ .$$

Since $\pi(\underline{\theta} \mid \underline{x}) = \dfrac{\left[\prod_{i=1}^{n} f(x_i \mid \underline{\theta}) \right] \pi(\underline{\theta})}{f(\underline{x})}$

$$= \frac{f(\underline{x} \mid \underline{\theta}) \pi(\underline{\theta})}{f(\underline{x})} \ ,$$

is the joint posterior probability distribution of Θ given $\underline{x}$, therefore, to minimize the Bayes risk, a decision function *d*($\underline{x}$) should be chosen such that the inner integral is a minimum. The inner integral is simply the conditional expectation $E\left[L\{\underline{\theta}, d(\underline{x})\} \mid \underline{X}\right]$ of the loss with respect to the joint posterior distribution of Θ given $\underline{x}$, or simply, the *posterior risk*. Thus, the Bayes decision function with respect to a given π can be found without computing the value of the minimum Bayes risk. If we let $\phi(d, \underline{x})$ denote the posterior risk and Let

$$\phi(\underline{x}) = \phi(d_\pi, \underline{x}) = \min\ \phi(d, \underline{x}) = \min\ E\left[\{\Theta, d(\underline{x})\} \mid \underline{X}\right],$$

Then the minimum Bayes risk becomes

$$r(\pi) = \int_{\chi} \phi(\underline{x})\, f(\underline{x}) d\,\underline{x} = E[\phi(\underline{x})],$$

Which is the expectation of $\phi(\underline{x})$ using the joint marginal distribution of $\underline{x}$.

١.٤.٢ Sequential Decision Procedure [٣٣]

We describe now formally the sequential *decisions* procedure.

When D be the set of possible decisions.

Definition: sequential decision (selection) procedure

Let $x_1, x_2, \ldots$ be random variables observed sequentially. A sequentially decision procedure consists in:

١- Sampling rule: The sampling rule prescribes how observations are taken sequentially, such as one at a time or a group at a time.

٢- A stopping rule σ_n which specifies whether a decision must be taken without taking any further observation.

If at least one observation is taken, this rule specifies for every set of observed values $(x_1, \ldots, x_n)$, with $n \geq 1$, whether to stop sampling and take a decision out of D or to take another observation X_{n+1}.

٣- A decision rule δ_n which specifies the decision to be taken if $n \geq 1$ observations have been taken, then one takes an action $\delta_n(x_1, \ldots, x_n) \in D$.

Once a decision has been taken, the sampling process is stopped.

١.٤.٣ Dynamic Programming

Backward induction has been used in the literature as computational technique for finding the optimal sequential procedures for different statistical decision problems. Bellman (١٩٥٧) introduced alternative term to backward induction, called it dynamic programming and showed how it could be used to solve multistage decision processes. The general ideas of dynamic programming and its applications can be found in Simpson (١٩٦١). The linkage between dynamic programming and decision theory was given in Lindley (١٩٦١). Wetherill (١٩٦١), Lindley and Barnett (١٩٦٥), Freeman (١٩٧٠, ١٩٧٢), El-Sayyad and Freeman (١٩٧٣) and Madhi (١٩٨٦) among others employed dynamic programming allied with the concept of truncation, that is the maximum number of observations is fixed in advance, to construct some optimal sequential procedures.

The dynamic programming technique is particularly important in multistage processes where decision are taken sequentially and where they are not independent, in that decisions taken earlier will affect decisions made later. Consequently in order to find the best decision at the present time, it is necessary to know the best decision in the future and the only way to know that is to work backwards from the optimal future decisions back to the origin. This technique is used in chapter ٢ to develop the optimal Bayesian selection procedure.

١.٥ The Plan of the Thesis

The thesis can be summarized as follows:

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

٢.١ Summary

In this chapter Bayesian optimal procedures for selecting the best cell in multinomial distribution are introduced. Bayesian decision- theoretic formulation to the problem is given in section ٢.٢. The stopping risks of linear losses are presented in section ٢.٣. In section ٢.٤, the fully optimal 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 section. Section ٢.٥ deals with Bayesian optimal fixed sample size procedure (BOF). Comparisons between the schemes (BOS) and (BOF) are given in section ٢.٦.

٢.٢ Bayesian decision- theoretic formulation

For the multinomial distribution with k cells, let the unknown probability of an observation in the i^{th} cell be p_i, (*i*= ١, ٢, …, *k*). 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*=١, ٢, ..., *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:

$$L(d_i, \underline{p}^*) = \begin{cases} k^*(p_{[k]} - p_i) & \text{if } (p_{[k]} > p_i) \\ 0 & \text{if } (p_{[k]} = p_i) \end{cases} \quad \dots (٢.٢.١)$$

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

, where $m' = \sum_{i=1}^{k} n_i'$ ... (٢.٢.٢) $\pi(\underline{p}) = \dfrac{\Gamma m'}{\prod_{i=1}^{k} \Gamma(n_i')} \prod_{i=1}^{k} p_i^{n_i'-1}$

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. The equivalent sample size is $(n_1' + n_2' + ... + n_k' = m')$. In addition to the prior information, we obtain some sample information from the multinomial population. If there are *k* categories, p_i is the probability that the i^{th} category is selected in a single trail, and the trails are independent, then the number of times each category is selected has a multinomial distribution. More precisely, let N_i be the number of times that category *i* is chosen in *m* independent trials.

Then

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

$$= \frac{m!}{n_1! n_2! ... n_k!} \prod_{i=1}^{k} p_i^{n_i}$$, where $\sum_{i=1}^{k} n_i = m$ and $\sum_{i=1}^{k} p_i = 1$.

The posterior distribution is derived from the prior probability function

and the multinomial distribution by means of Bayes theorem as follows.

$$\pi(\underline{p} \mid \underline{n}) = \frac{P(\underline{n} \mid \underline{p})\pi(\underline{p})}{\int P(\underline{n} \mid \underline{p})\pi(\underline{p}) d\,\underline{p}} = \frac{P(\underline{n} \mid \underline{p})\pi(\underline{p})}{P(\underline{n})}.$$

Now,

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

Then the posterior distribution is

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

This is a member of the Dirichlet family with parameters

and $\quad m'' = m' + m \ \ (i=1, \ldots, k). n_i'' = n_i' + n_i$

Hence, the posterior distribution has density function

$$\pi(\underline{p} \mid \underline{n}) = \frac{(m''-1)!}{(n_1''-1)!(n_2''-1)!\ldots.(n_k''-1)!} p_1^{n_1''-1} \ldots. p_k^{n_k''-1}, \qquad \ldots (2.2.3)$$

with posterior mean $\hat{p}_i = \frac{n_i''}{m''}$ (i= ١, ٢, …, 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

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

٢.٣ 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 ٢.٢. The stopping risk (the posterior expected loss) of the terminal decision d_i when the posterior distribution for $\underline{p}$ has parameters $(n_1'', n_2'', \ldots, n_k''; m'')$, 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:

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

$$= k^* \underset{\pi(\underline{p}|\underline{n})}{E}[\text{Þ}_{[k]} - \text{Þ}_i]$$

$$= k^*[\underset{\pi(\underline{p}|\underline{n})}{E}(\text{Þ}_{[k]}) - \hat{p}_i]$$

$$= k^*\left[\underset{\pi(\underline{p}|\underline{n})}{E}(\text{Þ}_{[k]}) - \frac{n_i''}{m''}\right]. \quad \text{... (٢.٣.١)}$$

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 $n_1'', n_2'', ..., n_k''$ is $n_{[1]}'' \le n_{[2]}'' \le, ..., \le n_{[k]}''$. The marginal posterior probability density function of p_i if $p_i = p_{[k]}$ is

$$f(p_{[k]}) = \frac{(m''-1)!}{(n_{[k]}''-1)!(m''-n_{[k]}''-1)!}\, p_{[k]}^{n_{[k]}''-1}(1-p_{[k]})^{m''-n_{[k]}''-1} \quad \text{... (٢.٣.٢)}$$

and the cumulative density function is

$$\text{Putting } y = p_{[k]}, \quad F(y) = \int_0^y \frac{z^{n_{[k]}''-1}(1-z)^{m''-n_{[k]}''-1}}{B(n_{[k]}'', m''-n_{[k]}'')}dz$$

Using integration by parts,

Let $u = (1-z)^{m''-n''_{[k]}-1}$ then $du = -(m''-n''_{[k]}-1)(1-z)^{m''-n''_{[k]}-2}dz$ and

$$dv = z^{n''-1}dz \quad \text{which leads} \quad v = \frac{1}{n''_{[k]}} z^{n''_{[k]}}.$$

$$F(y) = \frac{1}{B(n''_{[k]}, m''-n''_{[k]})}\left\{\frac{y^{n''_{[k]}}}{n''_{[k]}}(1-y)^{m''-n''_{[k]}-1} + \int_0^y z^{n''_{[k]}}(1-z)^{m''-n''_{[k]}-2}dz\right\}$$

$$= \frac{1}{B(n''_{[k]}, m''-n''_{[k]})}\left\{\frac{y^{n''_{[k]}}}{n''_{[k]}}(1-y)^{m''-n''_{[k]}-1} + \frac{(m''-n''_{[k]}-1)}{n''_{[k]}(n''_{[k]}+1)} y^{n''_{[k]}+1}.\right.$$

$$(1-y)^{m''-n''_{[k]}-2} + \frac{(m''-n''_{[k]}-1)(m''-n''_{[k]}-2)}{n''_{[k]}(n''_{[k]}+1)(n''_{[k]}+2)} y^{n''_{[k]}+2}(1-y)^{m''-n''_{[k]}-3}$$

$$\left. + \ldots\ldots\ldots\ldots\ldots + \frac{(m''-n''_{[k]}-1)!}{n''_{[k]}(n''_{[k]}+1)\ldots\ldots(m''-1)} y^{m''-1}\right\}$$

$$= \frac{(m''-1)!}{n''_{[k]}(n''_{[k]}-1)!(m''-n''_{[k]}-1)!} y^{n''_{[k]}}(1-y)^{m''-n''_{[k]}-1} +$$

$$\frac{(m''-1)!(m''-n''_{[k]}-1)}{n''_{[k]}(n''_{[k]}+1)(n''_{[k]}-1)!(m''-n''_{[k]}-1)!}.y^{n''_{[k]}+1}(1-y)^{m''-n''_{[k]}-2} +$$

$$\frac{(m''-1)!(m''-n''_{[k]}-1)(m''-n''_{[k]}-2)}{n''_{[k]}(n''_{[k]}+1)(n''_{[k]}+2)(n''_{[k]}-1)!(m''-n''_{[k]}-1)!} y^{n''_{[k]}+2}(1-y)^{m''-n''_{[k]}-3} +$$

$$\ldots\ldots\ldots + \frac{(m''-1)!(m''-n''_{[k]}-1)!y^{m''-1}}{(n''_{[k]}-1)!(m''-n''_{[k]}-1)!n''_{[k]}(n''_{[k]}+1)\ldots.m(''-1)}$$

$$= \frac{(m''-1)!}{n''_{[k]}!(m''-n''_{[k]}-1)!} y^{n''_{[k]}}(1-y)^{m''-n''_{[k]}-1} + \frac{(m''-1)!}{(n''_{[k]}+1)!(m''-n''_{[k]}-2)!} y^{n''_{[k]}+1}.$$

$$(1-y)^{m''-n''_{[k]}-2} + \frac{(m''-1)!}{(m''+2)!(m''-n''_{[k]}-3)!} y^{n''_{[k]}+2}(1-y)^{m''-n''_{[k]}-3}$$

$$+ \ldots\ldots\ldots\ldots\ldots + \frac{(m''-1)!}{(m''-1)!} y^{m''-1}$$

$$= \sum_{j=n''_{[k]}}^{m''-1} \frac{(m''-1)!}{j!(m''-1-j)!} . y^{j}(1-y)^{m''-1-j} \quad , \quad \text{where } j = n''_{[k]}, \ldots, m''-1 .$$

Then

$$g(p_{[k]}) = k\ f(p_{[k]})\left(F(p_{[k]})\right)^{k-1}$$

$$= k \frac{(m''-1)!}{(n''_{[k]}-1)!(m''-n''_{[k]}-1)!} p_{[k]}^{n''_{[k]}-1}(1-p_{[k]})^{m''-n''_{[k]}-1} .$$

$$\left\{ \sum_{j=n''_{[k]}}^{m''-1} \frac{(m''-1)!}{j!(m''-1-j)!} p_{[k]}^{j}(1-p_{[k]})^{m''-1-j} \right\}^{k-1} .$$

Also we have

$$\underset{\pi(\underline{p}|\underline{n})}{E}(p_{[k]}) = \int_0^1 p_{[k]} \frac{k(m''-1)!}{(n''_{[k]}-1)!(m''-n''_{[k]}-1)!} p_{[k]}^{n''_{[k]}-1}(1-p_{[k]})^{m''-n''_{[k]}-1}$$

$$= \frac{k\left[(m''-1)!\right]^k}{(n''_{[k]}-1)!(m''-n''_{[k]}-1)!}\int_0^1 p_{[k]}^{n''_{[k]}+1-1}(1-p_{[k]})^{m''-n''_{[k]}-1}\,dp_{[k]} \qquad .$$

$$\left\{\sum_{j=n''_{[k]}}^{m''-1}\frac{\left(\frac{p_{[k]}}{1-p_{[k]}}\right)^j}{j!(m''-1-j)!}\right\}^{k-1}\left\{(1-p_{[k]})^{m''-1}\right\}^{k-1}dp_{[k]}$$

$$= \frac{k\left[(m''-1)!\right]^k}{(n''_{[k]}-1)!(m''-n''_{[k]}-1)!}\left\{\int_0^1\left[\sum_{j=n''_{[k]}}^{m''-1}\frac{\left[\frac{p_{[k]}}{(1-p_{[k]})}\right]^j}{j!(m''-1-j)!}\right]^{k-1} \right.$$

$$\left. p_{[k]}^{n''_{[k]}}(1-p_{[k]})^{km''-n''_{[k]}-k}\right\}dp_{[k]}$$

Since

$$,\left[\sum_{j=n''_{[k]}}^{m''-1}\frac{\left(\frac{p_{[k]}}{(1-p_{[k]})}\right)^j}{j!(m''-1-j)!}\right]^2 = \sum_{j_1=n''_{[k]}}^{m''-1}\sum_{j_2=n''_{[k]}}^{m''-1}\left[\frac{\left(\frac{p_{[k]}}{(1-p_{[k]})}\right)^{j_1+j_2}}{j_1!\,j_2!(m''-j_1-1)!(m''-j_2-1)!}\right]$$

then using mathematical induction, we can prove that

$$\left[\sum_{j=n''_{[k]}}^{m''-1}\frac{\left(\frac{p_{[k]}}{1-p_{[k]}}\right)^j}{j!(m''-1-j)!}\right]^{k-1} = \sum_{j_1=n''_{[k]}}^{m''-1}\sum_{j_2=n''_{[k]}}^{m''-1}\cdots\cdots\sum_{j_{k-1}=n''_{[k]}}^{m''-1} .$$

$$\left\{\left(\frac{p_{[k]}}{1-p_{[k]}}\right)^{j_1+j_2+\ldots+j_{k-1}} \Big/ j_1!(m''-j_1-1)!\, j_2!(m''-j_2-1)!\ldots j_{k-1}!(m''-j_{k-1}-1)!\right\}.$$

So that

$$\underset{\pi(\underline{p}|\underline{n})}{E}(\text{þ}_{[k]}) = \frac{k[(m''-1)!]^k}{(n''_{[k]}-1)!(m''-n''_{[k]}-1)!} \sum_{j_1=n''_{[k]}}^{m''-1} \sum_{j_2=n''_{[k]}}^{m''-1} \ldots\ldots \sum_{j_{k-1}=n''_{[k]}}^{m''-1} \int_0^1 \left(\frac{p_{[k]}}{1-p_{[k]}}\right)^{j_1+j_2+\ldots+j_{k-1}}$$

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

We have

$$\int_0^1 p_{[k]}^{j_1+j_2+\ldots+j_{k-1}+n''_{[k]}}(1-p_{[k]})^{km''-n''_{[k]}-k-j_1-j_2-\ldots-j_{k-1}}dp_{[k]}$$

$$= \frac{(j_1+j_2+\ldots+j_{k-1}+n''_{[k]})!(km''-n''_{[k]}-k-j_1-\ldots-j_{k-1})!}{(km''-k+1)!}$$

which is Beta function.

Hence,

$$\cdot\ \underset{\pi(\underline{p}|\underline{n})}{E}(\text{þ}_{[k]}) = \frac{k[(m''-1)!]^k}{(n''_{[k]}-1)!(m''-n''_{[k]}-1)!} \sum_{j_1=n''_{[k]}}^{m''-1} \sum_{j_2=n''_{[k]}}^{m''-1} \ldots\ldots \sum_{j_{k-1}=n''_{[k]}}^{m''-1}$$

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

Therefore

$$S_i(n_1'', n_2'', ..., n_k''; m'') = k^* \left\{ \frac{k[(m''-1)!]^k}{(n_{[k]}''-1)!(m''-n_{[k]}''-1)!} \sum_{j_1=n_{[k]}''}^{m''-1} \sum_{j_2=n_{[k]}''}^{m''-1} ..... \sum_{j_{k-1}=n_{[k]}''}^{m''-1} \right.$$

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

$$\left. \frac{n_i''}{m''} \right\} \quad \text{... (٢.٣.٣)}$$

٢.٤ Fully Bayesian (Optimal) Sequential Procedure(BOS)

In this section Bayesian optimal sequential schemes to select the best cell is presented.

٢.٤.١ Construction of the Procedure BOS

The Bayesian decision formulation, given in section ٢.٢, 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 $\underline{\varepsilon}_i = (y_1, y_2, ..., y_k)$ 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 $(n_1'', n_2'', ..., n_k''; m'')$ 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 $(n_1'', n_2'', ..., n_k''; m'')$, let

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

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

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

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

Then the dynamic programming equations for the procedure are

$$B(n_1'', n_2'', ..., n_k''; m'') = c + \sum_{i=1}^{k} \hat{p}_i D(\underline{n''} + \underline{\varepsilon}_i; m'' + 1)$$

, where *c* is constant cost of sampling one observation.

Knowing $S_i(n_1'', n_2'', ..., n_k''; m'')$, (*i*= ١, ..., *k*) and $B(n_1'', n_2'', ..., n_k''; m'')$, the equation for $D(n_1'', n_2'', ..., n_k''; m'')$ is given by

$$D(n_1'', n_2'', ..., n_k''; m'') = \min\{ S(n_1'', n_2'', ..., n_k''; m''), B(n_1'', n_2'', ..., n_k''; m'')\}$$

where

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

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 $(n_1', n_2', ..., n_k'; m')$ until this origin is reached.

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

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

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

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

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

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

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

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

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

٢.٤.٢ Numerical Results

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*=٣ and ٤ are given in tables (٢-١,٢-٢). 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*=٢, ٣, under different values of *c* and k^*.Tables (٢-٣, ٢-٤) show that, the optimal overall risk increases when *c* and k^* increases for different values of prior, for *k*=٣, ٤.

Table (٢- ١)

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= ٣.

c, k^*	sample size / prior	٢	٣
c=١٠٠	(١, ١, ١; ٣)	١٩٦.٧٣٦١	١٦٢.٨٧١٠
=٢٠٠٠ k^*	(١, ١, ٢; ٤)	١٧٥.٨٩٦١	١٤٧.٨٦١٣
	(٠, ٠, ٢; ٢)	٢٣٧.٩١٢١	٢٣١.٧٤٧

	(٠, ٠, ٣; ٣)	٢٢٧.٠٤٧٥	٢٢٢.٦١٨٣
	(٠, ٠, ٢; ٢)	٢٣٧.٩١٢١	٢٣١.٧٤٧
	(١, ٠, ٢; ٤)	١٩١.٤٣٦١	١٢٧.٧٣٤٢
$c=^{٥}٠$	(١, ٠, ١; ٢)	٧٨.٢٥٤٤٤	٦٩.٤٠٤٥٥
$=١٠٠k^{*}$	(١, ٠, ٢; ٣)	٧٧.٥٢٤١٩	٦٣.٣٩٣٢٦
	(٠, ١, ١; ٢)	٧٩.٤١٩٥٤٠	٧١.٥٤٩٠٤
	(٠, ١, ٣; ٤)	٧٨.٥٤٨٨٧	٦٢.٧٧١٧٦
$c=١٠٠$	(١, ٠, ١; ٢)	١٥٢.٢٥٤٤	١٣٦.٩٠٤٦
$=١٠٠k^{*}$	(٢, ١, ١; ٤)	١٥١.٩٠٦٥	١٠٩.٥٦١٠
	(١,٠, ١; ٢)	١٥٣.٢٥٤٤	١٣٦.٩٠٤٦
	(١,٠, ٢, ٣)	١٥٢.٥٢٤٢	١٢٤.١٠٧٥
	(١, ٠, ١; ٢)	١٥٣.٢٥٤٤	١٣٦.٩٠٤٦
	(١, ٢, ١; ٤)	١٥١.٤٢٢٢	١٠٥.٠١٥٥
$c=١٠;$	(٠, ١, ١; ٢)	٢٣.٨٣٩٠٨	١٦.٣٤٨٠٨
$=٢٠٠k^{*}$	(٠, ١, ٢; ٣)	٢١.٧٧٢٥	١٥.٧٧٠٧٦
	(١, ٠, ١; ٢)	٢١.١٠٨٨٨	١٥.٩٤٢٤٣
	(١, ٠, ٢; ٣)	١٩.٥٧٢١٨	١٥.٠١٢٧١

Table (٢- ٢)

The influence of prior information and sample size on the overall risk of the optimal procedure (BOS), under linear loss function,

various priors and various N for k= ٤.

c, k^*	sample size / prior	٢	٣
c=١٠٠ =٢٠٠٠ k^*	(١, ١, ٠, ٠; ٢)	٢٨٣.٣٣٢٦	٢٠٨.٥٧١٤
	(١, ٢, ٠, ٠; ٣)	١٨٥.٧١٤٤	١٧٢.٢٢٢٣
	(٠, ٠, ١, ١; ٢)	٢١٦.١٥٦٤	١٨٠.٠٧٧٩
	(٠, ١, ١, ١; ٣)	١٦١.٦٨٠٨	١٣٢.٥٣٧٢
	(٠, ٠, ١, ٠; ١)	٣١٦.٦٦٦٧	١٨٥.٣٣٣٥
	(٠, ١, ١, ١; ٣)	١٦١.٦٨٠٨	١٣٢.٥٣٧٢
$c=50$ $k^*=1000$	(١, ٠, ١, ٠; ٢)	١٤١.٦٦٦٣	١٠٤.٢٨٥٧
	(١, ٠, ١, ١; ٣)	٨٠.٨٤٠٤١	٦٦.٠١٢٦٧
	(٠, ١, ٠, ٠; ١)	١٥٨.٣٣٣٣	٩١.٦٦٦٧٤
	(٠, ١, ٢, ٠; ٣)	٩٢.٨٥٧١٨	٨٦.١١١٤٠
$c=10$ $k^*=100$	(١, ٠, ١, ٠; ٢)	٢١.٦٦٦٦٣	١٩.٤٢٨٥٧
	(١, ١, ١, ٠; ٣)	٢٠.٣٥٧١٥	١٩.٤٠٤٧٦
	(٠, ١, ٠, ٠; ١)	٢٣.٣٣٣٣٣	١٨.٣٣٣٣٤٠

	(٠, ١, ١, ١; ٣)	١٨.٠٨٤٠٤	١٢.٩٥٨٥٢
	(٠, ١, ٠, ٠; ١)	٢٣.٣٣٣٣٣	١٨.٣٣٣٣٤٠
	(١, ٢, ٠, ٠; ٣)	١٩.٥٢٣٧٨	١٧.٣٨٠٩٤٠
	(٠, ٠, ١, ٠; ١)	٢٣٣.٣٣٣٣	١٨٣.٣٣٣٤
$c=100$	(٠, ١, ١, ١; ٣)	١٣٠.٨٤٠٤	١٢٩.٥٨٥٢
$k^*=1000$	(١, ١, ٠, ٠; ٢)	٢١٦.٦٦٦٣	١٩٤.٢٨٥٧
	(١, ١, ١, ٠; ٣)	٢٠٣.٥٧١٥	١٩٤.٠٤٧٧

Table (٢-٣)

The influence of the loss constant and sampling cost on the optimal overall risk in (BOS), for different prior information, when N= ٢, ٣ and k= ٣.

N=٢	c, k^*	
Prior	١٠, ٢٠٠	١٠٠, ٢٠٠
(٠, ١, ٢; ٣)	٢١.٧٧٢٥	١٥٦.٧٧٢٥
(١, ٠, ١; ٢)	٢١.١٠٨٨٨	١٥٦.١٠٨٩
(٠, ٠, ٢; ٢)	٢٣.٧٩١٢١	١٥٨.٧٩١٢
N=٣	١٠, ٢٠٠	١٠٠, ٢٠٠

(٠, ١, ٢; ٣)	١٥.٧٧٠٧٦	١٢٩.٣٤٢٢
(١, ٠, ١; ٢)	١٥.٩٤٢٤٣	١٣٧.٤٤٢٤
(٠, ٠, ٢; ٢)	١١.٥٠٠٠	١٠١.٥٠٠٠
N=٢	١٠٠, ١٠٠	١٠٠, ٢٠٠٠
(٠, ١, ٢; ٣)	١٥٣.٣٨٦٢	١٧٨.٩٨٨٦
(٠, ١, ١; ٢)	١٥٤.٤١٩٥	٢٣٨.٣٩٠٨
(١, ٠, ١; ٢)	١٥٣.٢٥٤٤	٢٠٧.٤٨٨٨
N=٣	١٠٠, ١٠٠	١٠٠, ٢٠٠٠
(٠, ١, ٢; ٣)	١٢٨.٥٩٩٧	١٤٢.٧٠٧٦
(٠, ١, ١; ٢)	١٤١.٥٤٩	١٥٦.٧٣٠٨
(١, ٠, ١; ٢)	١٣٦.٩٠٤٦	١٤٧.١٢٤٣
N=٢	١٠, ٢٠٠	١٠٠, ٢٠٠٠
(٠, ١, ٢; ٣)	٢١.٧٧٢٥	١٧٨.٩٨٨٦
(٠, ١, ١; ٢)	٢٣.٨٣٩٠٨	٢٣٨.٣٩٠٨
(١, ٠, ١; ٢)	٢١.١٠٨٨٨	٢٠٧.٤٨٨٨
N=٣	١٠, ٢٠٠	٥٠, ١٠٠٠
(٠, ١, ٢; ٣)	١٥.٧٧٠٧٦	١٤٢.٧٠٧٦
(٠, ١, ١; ٢)	١٦.٣٤٨٠٨	١٥٦.٧٣٠٨
(١, ٠, ١; ٢)	١٥.٩٤٢٤٣	١٤٧.١٢٤٣

Table (٢- ٤)

The influence of the loss constant and sampling cost on the optimal overall risk in (BOS), for different prior information, when N= ٢, ٣ and k= ٤.

$N=^{٢}$	c,k^{*}	
Prior	١٠٠, ١٠٠٠	١٠٠, ٢٠٠٠
(١, ١, ٠, ٠; ٢)	٢١٦.٦٦٦٣	٢٨٣.٣٣٢٦
(٠, ١, ١, ١; ٣)	١٣٠.٨٤٠٤	١٦١.٦٨٠٨
(٠, ٠, ١, ٠; ١)	٢٣٣.٣٣٣٣	٣١٦.٦٦٦٧
$N=^{٣}$	١٠٠, ١٠٠٠	١٠٠, ٢٠٠٠
(١, ١, ٠, ٠; ٢)	١٩٤.٢٨٥٧	٢٠٨.٥٧١٤
(٠, ١, ١, ١; ٣)	١٢٩.٥٨٥٢	١٣٢.٥٣٧٢
(٠, ٠, ١, ٠; ١)	١٨٣.٣٣٣٤	١٨٥.٣٣٣٥
$N=^{٢}$	٥٠, ١٠٠٠	١٠٠, ٢٠٠٠
(١, ٠, ١, ٠; ٢)	١٤١.٦٦٦٣	٢٨٣.٣٣٢٦
(٠, ١, ١, ١; ٣)	٨٠.٨٤٠٤١	١٦١.٦٨٠٨
(١, ٠, ٠, ٠; ١)	١٥٨.٣٣٣٣	٣١٦.٦٦٦٧
$N=^{٣}$	٥٠, ١٠٠٠	١٠٠, ٢٠٠٠

(١, ٠, ١, ٠; ٢)	١٠٤.٢٨٥٧	٢٠٨.٥٧١٤
(٠, ١, ١, ١; ٣)	٧٢.٧٦٠٤٧	١٣٢.٥٣٧٢
(١, ٠, ٠, ٠; ١)	٩١.٦٦٦٧٤	١٨٣.٣٣٣٥
N=٢	١٠, ١٠٠٠	٥٠, ١٠٠٠
(١, ٠, ١, ٠; ٢)	٨١.٦٦٦٦٣	١٤١.٦٦٦٣
(٠, ١, ٠, ١; ٣)	٤٦.٢٦٢٦٩٣	٩٦.٢٦٩٦٣
(١, ٠, ٠, ٠; ١)	٩٨.٣٣٣٣٤	١٥٨.٣٣٣٣
N=٣	١٠, ١٠٠٠	٥٠, ١٠٠٠
(١, ٠, ١, ٠; ٢)	٣٢.٢٨٥٧٢	١٠٤.٢٨٥٧
(٠, ١, ٠, ١; ٣)	٣٤.٠٣٨٩٦	٩٠.٠٣٨٩٦
(١, ٠, ٠, ٠; ١)	١٨.٣٣٣٣٤١	٩١.٦٦٦٧٤

٢.٥ The Bayesian (Optimal) Fixed Sample Size Procedure (BOF)

In this section 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.

٢.٥.١ 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 $S_i = S$,

where the stopping risk S_i for decision d_i is given by

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

$$= mc + \underset{\pi(\underline{p}|\underline{n})}{E} \{L(p_{[k]} - p_i)\}. \qquad \text{... (٢.٥.١)}$$

If *m* number of observations have been taken.

٢.٥.٢ Numerical Results

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

The numerical results given in tables (٢-٥, ٢-٦, ٢-٧), when *k*=٣, ٤ and ٥, 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 is decreases.

In tables (٢-٨, ..., ٢-١٣) 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*=٣, ٤ and ٥.

Table (٢-٥)

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=٣, when $c=100$, $k^*=2000$.

$k^*=2000\ c=100$,			
N	Prior		
	(٠, ١, ١; ٢)	(١, ١, ١; ٣)	(٢, ١, ١; ٤)
٢	٥٤٢.٨٥٧١	٤٨٧.١٧٩٥	٢٨٧.١٢١٢
٣	٥٨٧.١٧٩٥	٥٤٦.٢١٢٢	٣٧٠.٥٦٠٩
٤	٦٤٦.٢١٢٢	٦١٥.١٥٣٤	٤٥٩.٠٩١
٥	٧١٥.١٥٣٤	٦٩٠.٩٠٩١	٥٥٠.٧١٨٩
٦	٧٩٠.٩٠٩١	٧٧١.٥٠٣٣	٦٤٤.٣٦١
N	Prior		
	(١, ١, ١; ٣)	(٣, ١, ١; ٥)	(٤, ١, ١; ٦)
٢	٤٨٧.١٧٩٥	٤١٥.١٥٣٤	٣٩٠.٩٠٩٠
٣	٥٤٦.٢١٢٢	٤٩٠.٩٠٩٠	٤٧١.٥٠٣٣
٤	٦١٥.١٥٣٤	٥٧١.٥٠٣٣	٥٥٥.٦٣٩
٥	٦٩٠.٩٠٩١	٦٥٥.٦٣٩	٦٤٢.٤٣٧٩

٦	٧٧١.٥٠٣٣	٧٤٢.٤٣٧٩	٧٣١.٢٨٧٢
N	Prior		
	(١, ٣, ٠; ٤)	(١, ٣, ١; ٥)	(٢, ٣, ١; ٦)
٢	٨٦٦.٦٦٦٦	٤١٥.١٥٣٤	٢٥٩.٠٩١
٣	٨٧١.٤٢٨٨	٤٩٠.٩٠٩٠	٣٥٠.٧١٨٩
٤	٩٠٠.٠٠٠١	٥٧١.٥٠٣٣	٤٤٤.٣٦١
٥	٩٤٤.٤٤٤٤	٦٥٥.٦٣٩	٥٣٩.٣٨٠٢
٦	٩٩٩.٩٩٩٩	٧٤٢.٤٣٧٩	٦٣٥.٣٧٩٥

Table (٢- ٦)

The influence of prior information and N on the Bayes risk of the optimal procedure (BOF), under linear loss function, various priors and various N for k= ٤, when $c=100$, $k^*=2000$.

$k^*=2000$ $c=100$,			
N	Prior		
	(٠, ٠, ٠, ١; ١)	(١, ٠, ٠, ١; ٢)	(٢, ٠, ٠, ١; ٣)
٢	٣٤٦.٠٣١٧	٢٦٨.١٣١٩	٢٣٥.٥٩٥٧

٣	٧٣١.٨٦٨٢	٦٦٤.٤٠٤٤	٣١٩.٤٨١
٤	٧٦٤.٤٠٤٤	٧١٣.٨٥٢٤	٤١٠.٥٩٥٨
٥	٨١٣.٨٥٢٤	٧٧٥.١١٨٥	٥٠٥.٣٢٩٥
N	Prior		
	(١, ٠, ٠, ١; ٢)	(١, ٠, ١, ١; ٣)	(١, ٠, ٢, ١; ٤)
٢	٢٦٨.١٣١٩	٢٣٥.٥٩٥٧	٢١٩.٤٨١٠
٣	٦٦٤.٤٠٤٤	٦١٣.٨٥٢٤	٣١٠.٥٩٥٨
٤	٧١٣.٨٥٢٤	٦٧٥.١١٨٥	٤٠٥.٣٢٩٥
N	Prior		
	(١, ٠, ١, ٠; ٢)	(١, ١, ١, ٠; ٣)	(١, ٢, ١, ٠; ٤)
٢	٦٦٥.٧٣٤٣	٦٢٩.٠٢٩٨	٥٨٦.٩٠٥٤
٣	٧٠٠.٢٥٨٥	٦٩٢.٩٥٠٧	٦٧٠.٤١٩٦
٤	٧٩٢.٩٥٠٧	٧٥٥.٥٧٣٨	٧٤٦.٠١٤٦

Table (٢-٧)

The influence of prior information and N on the Bayes risk of the optimal procedure (BOF), under linear loss function, various priors and various N for k=٥, when $c=100$, $k^*=2000$.

$c=100$, $k^*=2000$			
N	Prior		
	(٠,١,٠,٠,٠;١)	(٠,١,٠,١,٠;٢)	(٠,١,١,١,٠;٣)
٢	١٣٥١.٥١٥	١٢٢١.٤٧٩	١٠٨٥.٩٥٩
٣	١٦٧٥.٠٠٠	١٥٤٤.٤٠٥	١٤٠٧.٩١٨
٤	١٩٠٤.٧٦٢	١٧٨٣.٣٨٥	١٦٥.٤.٣٩٩
N	Prior		
	(٠,١,٠,٠,٠;١)	(٠,٢,٠,٠,٠;٢)	(٠,٣,٠,٠,٠;٣)
٢	١٣٥١.٥١٥	٧٢١.٤٧٨٦	٢٨٥.٩٥٨٦
٣	١٦٧٥.٠٠٠	١١٤٤.٤٠٥	٧٤١.٢٥١
٤	١٩٠٤.٧٦٢	١٤٥٠.٠٥١	١٠٨٢.٩٧
N	Prior		
	(٠,١,٠,٠,٠;١)	(٠,١,٠,١,٠;٢)	(٠,٢,٠,١,٠;٣)
٢	١٣٥١.٥١٥	١٢٢١.٤٧٩	٦٨٥.٩٥٨٦

٣	١٦٧٥.٠٠٠	١٥٤٤.٤٠٥	١٠٧٤.٥٨٤
٤	١٩٠٤.٧٦٢	١٧٨٣.٣٨٥	١٣٦٨.٦٨٥
N	Prior		
	(٠,١,٠,٠,٠;١)	(٠,٢,٠,٠,٠;٢)	(٠,٢,٠,١,٠;٣)
٢	١٣٥١.٥١٥	٧٢١.٤٧٨٦	٦٨٥.٩٥٨٦
٣	١٦٧٥.٠٠٠	١١٤٤.٤٠٥	١٠٧٤.٥٨٤
٤	١٩٠٤.٧٦٢	١٤٥٠.٠٥١	١٣٦٨.٦٨٥

Table (٢-٨)

The influence of loss constant on the optimal procedure

(BOF), for different priors information ,

when N= ٢(١) ١, and k= ٣, $c = 50$.

c, k	N	Prior	
		(٠,١,١;٢)	(١,١,١;٣)
٥٠,٥٠	٢	١٠٨.٥٧١٤	١٠٧.١٧٩٥
	٣	١٥٧.١٧٩٥	١٥٦.١٥٥٣
	٤	٢٠٦.١٥٥٣	٢٠٥.٣٧٨٨
	٥	٢٥٥.٣٧٨٨	٢٥٤.٧٧٢٧

	٦	٣٠٤.٧٧٢٧	٣٠٤.٢٨٧٦
	٧	٣٥٤.٢٨٧٦	٣٥٣.٨٩١٠
٥٠, ١٠٠	٢	١١٧.١٤٢٩	١١٤.٣٥٩
	٣	١٦٤.٣٥٩	١٦٢.٣١٠٦
	٤	٢١٢.٣١٠٦	٢١٠.٧٥٧٧
	٥	٢٦٠.٧٥٧٧	٢٥٩.٥٤٥٤
	٦	٣٠٩.٥٤٥٤	٣٠٨.٥٧٥٢
	٧	٣٥٨.٥٧٥٢	٣٥٧.٧٨٢
٥٠, ١٥٠	٢	١٢٥.٧١٤٣	١٢١.٥٣٨٥
	٣	١٧١.٥٣٨٥	١٦٨.٤٦٥٩
	٤	٢١٨.٤٦٥٩	٢١٦.١٣٦٥
	٥	٢٦٦.١٣٦٥	٢٦٤.٣١٨٢
	٦	٣١٤.٣١٨٢	٣١٢.٨٦٢٨
	٧	٣٦٢.٨٦٢٨	٣٦١.٦٧٢٩

Table (٢- ٩)

The influence of the sampling cost on the Bayes risk of the optimal procedure (BOF) for different prior information,

when N= ٢(١) ١, and k= ٣, $k^* = 2000$.

c, k^*	N	Prior	
		(٠, ١, ١; ٢)	(١, ١, ١; ٣)
١٠٠, ٢٠٠٠	٢	٥٤٢.٨٥٧١	٤٨٧.١٧٩٥
	٣	٥٨٧.١٧٩٥	٥٤٦.٢١٢٢
	٤	٦٤٦.٢١٢٢	٦١٥.١٥٣٤
	٥	٧١٥.١٥٣٤	٦٩٠.٩٠٩١
	٦	٧٩٠.٩٠٩١	٧٧١.٥٠٣٣
	٧	٨٧١.٥٠٣٣	٨٥٥.٦٣٩
	٨	٩٥٥.٦٣٩	٩٤٢.٤٣٧٩
٥٠٠, ٢٠٠٠	٢	١٣٤٢.٨٥٧	١٢٨٧.١٧٩
	٣	١٧٨٧.١٧٩	١٧٤٦.٢١٢
	٤	٢٢٤٦.٢١٢	٢٢١٥.١٥٣
	٥	٢٧١٥.١٥٣	٢٦٩٠.٩٠٩
	٦	٣١٩٠.٩٠٩	٣١٧١.٥٠٣
	٧	٣٦٧١.٥٠٣	٣٦٥٥.٠٣٩
	٨	٤١٥٥.٦٣٩	٤١٤٢.٤٣٨
٩٠٠, ٢٠٠٠	٢	٢١٤٢.٨٥٧	٢٠٨٧.١٧٩

	٣	٢٩٨٧.١٧٩	٢٩٤٦.٢١٢
	٤	٣٨٤٦.٢١٢	٣٨١٥.١٥٣
	٥	٤٧١٥.١٥٣	٤٦٩٠.٩٠٩
	٦	٥٥٩٠.٩٠٩	٥٥٧١.٥٠٣
	٧	٦٤٧١.٥٠٣	٦٤٥٥.٦٣٩
	٨	٧٣٥٥.٦٣٩	٧٣٤٢.٤٣٨

Table (٢-١٠)

The influence of the loss constant on the Bayes risk of the optimal procedure (BOF) for different prior information,

when N= ٢, ٣, ٤ and k= ٤, $c = 500$.

c, k^*	N	Prior	
		(٠,١,١,١; ٣)	(١,١,١,١; ٤)
٥٠٠, ٥٠٠	٢	١٠٩١.١٠١	١٠٧٨.٤٦٣
	٣	١٥٧٨.٤٦٣	١٥٦٨.٧٨
	٤	٢٠٦٨.٧٨	٢٠٦١.١٦٨
٥٠٠, ٩٠٠	٢	١١٦٣.٩٨٢	١١٤١.٢٣٤
	٣	١٦٤١.٢٣٤	١٦٢٣.٨٠٣
	٤	٢١٢٣.٨٠٣	٢١١٠.١٠٢
٥٠٠, ١٣٠٠	٢	١٢٣٦.٨٦٣	١٢٠٤.٠٠٤

	٣	١٧٠٤.٠٠٤	١٦٧٨.٨٢٧
	٤	٢١٧٨.٨٢٧	٢١٥٩.٠٣٦

Table (٢-١١)

The influence of the sampling cost on the Bayes risk of the optimal procedure (BOF) for different prior information,

when N= ٢, ٣, ٤ and k= ٤, $k^* = 600$.

c, k^*	N	Prior	
		(٠, ١, ١, ١, ٣)	(١, ١, ١, ١, ٤)
٦٠٠, ٦٠٠	٢	١٣٠٩.٣٢١	١٢٩٤.١٥٦
	٣	١٨٩٤.١٥٦	١٨٨٢.٥٣٦
	٤	٢٤٨٢.٥٣٦	٢٤٧٣.٤٠١
١٠٠٠, ٦٠٠	٢	٢١٠٩.٣٢١	٢٠٩٤.١٥٦
	٣	٣٠٩٤.١٥٦	٣٠٨٢.٥٣٦
	٤	٤٠٨٢.٥٣٦	٤٠٧٣.٤٠١
١٤٠٠, ٦٠٠	٢	٢٩٠٩.٣٢١	٢٨٩٤.١٥٦
	٣	٤٢٩٤.١٥٦	٤٢٨٢.٥٣٦
	٤	٥٦٧٣.٤٠١	٥٦٧٣.٤٠١

Table (٢-١٢)

The influence of the loss constant on the Bayes risk of the optimal procedure (BOF) for different prior information,

when N= ٢, ٣, ٤ and k= ٥, $c = 400$.

c, k^*	N	Prior	
		(٠,١,٠,١,٠;٢)	(٠,٠,٢,٠,١;٣)
٤٠٠, ٥٠٠	٢	١٠٥٠.٠٠١	٨٠٥.٨١٠٧
	٣	١٣٠٠.٠٠٠	١٢٠٨.١٣٨
	٤	١٦٠٠.٠٠١	١٦٠٨.٩٥٧
٤٠٠, ٩٠٠	٢	١٢٥٠.٠٠١	٨١٠.٤٥٩٢
	٣	١٣٨٠.٠٠	١٢١٤.٦٤٨
	٤	١٦٠٠.٠٠٢	١٦١٦.١٢٢
٤٠٠, ١٣٠٠	٢	١٤٥٠.٠٠٢	٨١٥.١٠٧٧
	٣	١٤٦٠.٠٠	١٢٢١.١٥٨
	٤	١٦٠٠.٠٠٣	١٦٢٣.٢٨٨

Table (٢-١٣)

The influence of the sampling cost on the Bayes risk of the optimal procedure (BOF) for different prior information , when N= ٢, ٣, ٤ and k= ٥, $k^* = 100$.

c, k^*	N	Prior	
		(٠,١,٠,١,٠;٢)	(٠,٠,٢,٠,١;٣)
١٠٠, ١٠٠٠	٢	٢٥٠.٠٠٠٢	٢٠١.١٦٢١

	٣	٣٢٠.٠٠٠	٣٠١.٦٢٧٦
	٤	٤٠٠.٠٠٠٢	٤٠١.٧٩١٤
٤٠٠, ١٠٠٠	٢	٨٥٠.٠٠٠٢	٨٠١.١٦٢١
	٣	١٢٢٠.٠٠٠	١٢٠١.٦٢٨
	٤	١٦٠٠.٠٠٠	١٦٠١.٧٩١
٧٠٠, ١٠٠٠	٢	١٤٥٠.٠٠	١٤٠١.١٦٢
	٣	٢١٢٠.٠٠	٢١٠١.٦٢٨
	٤	٢٨٠٠.٠٠	٢٨٠١.٧٩١

٢.٦ Comparisons between the Selection Procedures BOS and BOF

In this section we discuss the efficiency of the fully sequential scheme in terms of the percent reduction gained due to the use of BOS scheme risk which is calculated as follows:

$$RD = \frac{Risk\,(BOF) - Risk\,(BOS)}{Risk(BOF)}\% \quad \ldots (٢.٦.١)$$

.

Tables (٢-١٤,…, ٢-١٧) contain these values for k = ٣, ٤ for N= ٢, ٣ priors and various k^*, c. In tables (٢-١٤, ٢-١٥) we note that the percent reduction in risk increases as N increase, when N =٢, ٣ and k = ٣. When k =٤, the percentage reduction risk is also increases when N increases, this found in tables (٢-١٦, ٢-١٧).

The comparisons above assume that observation cost in all cases remains 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 (٢-١٤)

The effect of N on the percentage reduction in risk for different priors when c= ١٠٠, $k^* = 100$ and k= ٣.

$k^* = 100\ c = 100$				
Prior	N	BOF	BOS	RD%
(١, ٠, ١; ٢)	٢	٢٠٧.٨٥٧١	١٥٣.٢٥٤٤	٢٦.٢٦٧٣٤٥
	٣	٣١٤.٣٥٩	١٣٦.٩٠٤٦	٥٦.٤٤٩٦
(٢, ١, ١; ٤)	٢	٢٠٤.٣٥٦١	١٥١.٩٠٦٥	٢٥.٦٦٥٧٨٦
	٣	٣٠٣.٥٢٨	١٠٩.٥٦١٠	٦٣.٩٠٤١٥٣
(١, ٠, ٢; ٣)	٢	٢١٧.٥٠٢٥	١٥٢.٥٢٤٢	٢٩.٨٧٤٧٣٧
	٣	٣٠١.٠١٥٧	١٢٤.١٠٧٥	٤١.٢٢٩٥٧٧
(١, ٢, ١; ٤)	٢	٢١٢.٣١٠٦	١٥١.٤٢٢٢	٢٨.٦٧٨٩٢٦
	٣	٣١٠.٧٥٧٧	١٠٥.٠١٥٥	٦٦.٢٠٦٦٢٩

Table (٢-١٥)

The effect of N on the percentage reduction in risk for

different priors when c=٥٠, $k^* = 100$ and k=٣.

$k^* = 100\ c = 500$				
Prior	N	BOF	BOS	RD%
(١, ٠, ١; ٢)	٢	١٠٧.٨٥٧١	٧٨.٢٥٤٤٤	٢٧.٤٤٦١٨٥
	٣	١٦٤.٣٥٩٠٠	٦٩.٤٠٤٥٥	٤٢.٢٢٧٤١
(١, ٠, ٢; ٣)	٢	١١٧.٥٠٢٥	٧٧.٥٢٤١٩	٣٤.٠٢٣٣٦٩
	٣	١٥١.٠١٥٧	٦٣.٣٩٣٦٢	٥٨.٠٢٢٠٧٣
(٠, ١, ١; ٢)	٢	١١٧.١٤٢٩	٧٩.٤١٩٥٤٠	٣٢.٢٠٢٨٥٦
	٣	١٦٢. ٣٠٧٧	٧١.٥٤٩٠٤	٥٥.٩١٧٦٥٥
(٠, ١, ٣; ٤)	٢	١١٦.٢٣٣٨	٧٨.٥٤٨٨٧	٣٢.٤٢٤١٠٣
	٣	١٥٠.٨٨٤٩	٦٢.٧٧١٧٥	٥٨.٣٩٧٦٢٦

Table (٢-١٦)

The effect N on the percentage reduction in risk for different priors when c= ١٠٠, $k^* = 2000$ and k= ٤.

$k^* = 2000\ c = 100$				
Prior	N	BOF	BOS	RD%
(١, ١, ٠, ٠; ٢)	٢	٥٣٣.٣٣٠٩	٢٨٣.٣٣٢٦	٤٦.٨٧٤٨٩٥
	٣	٦٣٣.٣٣٠٩٠	٢٠٨.٥٧١٤	٦٧.٠٦٧٥٤٧

(١, ٢, ٠, ٠; ٣)	٢	٤٨٥.٧١٣٣	١٨٥.٧١٤٤	٦١.٧٦٤٦٠٤
	٣	٦٩٢.٩٥٠٧	١٧٢.٢٢٢٣	٧٥.١٤٦٥٢٩
(٠, ١, ٠, ٠; ١)	٢	٨٦٦.٦٦٦٧	٣١٦.٦٦٦٧	٦٣.٤٦١٥٣٦
	٣	٩٦٦.٦٦٤٢	١٨٣.٣٣٣٥	٨١.٠٣٤٤١٧
(٠, ١, ١, ١; ٣)	٢	٥٦٤.٤٠٤٤	١٦١.٦٨٠٨	٧١.٣٥٣٧٣١
	٣	٦١٣.٨٥٢٤	١٣٢.٥٣٧٥	٧٨.٤٠٨٩٤٦

Table (٢-١٧)

The effect of N on the percentage reduction in risk for different priors when c= ١٠, $k^* = 100$ and k= ٤.

$k^* = 100\ c = 10$,				
Prior	N	BOF	BOS	RD%
(١, ٠, ١, ٠; ٢)	٢	٤٣.٢٨٦٧١	٢١.٦٦٦٦٣	٤٩.٩٤٦٢٣
	٣	٥٠.٠١٢٩٣	١٩.٤٢٨٥٧	٦١.١٥٢٩٠٥
(١, ١, ١, ٠; ٣)	٢	٤١.٤٥١٤٩	٢٠.٣٥٧١٥	٥٠.٨٨٩٢٢
	٣	٤٩.٦٤٧٥٤	١٩.٤٠٤٧٦	٦٠.٩١٤٩٦١
(٠, ١, ٠, ٠; ١)	٢	٥٣.٣٣٣٣٤	٢٣.٣٣٣٣٣	٥٦.٢٥٠٠١١
	٣	٦٣.٣٣٣٢١٠	١٨.٣٣٣٣٤	٧١.٠٥٢٠٥٧

(٠,١,٠,١;٢)	٢	٤١.٥٩٣٤١	١٥.٨٠٧٨٢	٦١.٩٩٤٤١٢
	٣	٤٨.٢٢٠٢٢	١٥.٧٧١٠٣	٦٧.٢٩٣٧٤١

٣.١ Summary

In this chapter, 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.

This chapter is organized as follows:

A method to generate multinomial random variates which are used to carry out MC simulations is given in section ٣.٢. Section ٣.٣ presents the description of the MC studies. The construction of the Bayesian suboptimal sequential procedure (BSS) and some numerical results are given in section ٣.٤. Section ٣.٥ contains the construction of the Bayesian suboptimal fixed sample size procedure (BSF) and some comparisons between the schemes (BSS) and (BFS).

٣.٢ Generating Observations from Multinomial Distribution [٢٩]

Given the values of $\underline{p} = (p_1, p_2, ..., p_k)$ we are now in a position to generate observations (by computer) from the multinomial distribution with *k* cells with $p_i, (i = 1,2,..,k)$ as the probability of an observation in the i^{th} cell.

Let $\underline{n} = (n_1, n_2, \ldots, n_k)$ represents the observed frequencies in the *k* cells of the multinomial distribution. Then the problem is to generate a random vector $\underline{n}$ where $\underline{n} \sim M(m, \underline{p})$ where $M(m, \underline{p})$ represents a multinomial conditional on the number of observations $m = \sum_{i=1}^{k} n_i$ and the cell probabilities, $\underline{p}$.

An easy method is generate one observation at a time using the probabilities $p_1, p_2, \ldots, p_k$. If a set of *m* independent observations is generated one at a time from the distribution $M(1, \underline{p})$, the joint distribution of the *m* observations, represented by the vector of cell frequencies is $M(m, \underline{p})$. Hence a straightforward method of obtaining $\underline{n}$ is to use the uniform distribution to generate one observation at a time and accumulate the cell frequencies.

The assignment of a success to one of the *k* cells is as follows:

Generate a uniform variate *u* from *U*(٠, ١) and assign ١ to n_j if

$$\sum_{i=1}^{j} p_i - p_j < u < \sum_{i=1}^{j} p_i \quad \text{for all } j = ١, \ldots, k. \qquad \ldots (٣.٢.١)$$

and $\sum_{i=1}^{k} p_i = 1.$

٣.٣ Description of the MC Studies

In this section, 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 and the expected number of observations. 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=٥٠٠٠), 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:

$$\hat{p}_i = \frac{n_i' + n_i}{m' + m} \quad ,(i = 1,...,k) \qquad \text{... (٣.٣.١)}$$

,

Such that $m = \sum_{i=1}^{k} n_i$ and $m' = \sum_{i=1}^{k} n_i'$ with

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

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

, $\underline{n'} = (n_1', n_2', ..., n_k')$ (prior frequencies) such that $\underline{p} = (p_1, p_2, ..., p_k)$

$\sum_{i=1}^{k} p_i = 1$. 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. As measures of performance of the proposed procedures we shall use the following quantities:

١- P(CS): Probability of Correct Selection

In an MC experimentation the cell that has the greatest probability of event is known to us, so we can check if the procedure gives a correct selection. After *t* repetitions we estimate the probability of correct selection by the fraction of correct selection in the *t* replications. It can be computed as follows:

The proportion of number of times when the procedure stops $P(D_i / D_i)$: and takes decision D_i given decision D_i is true in t repetitions.

An estimate of probability of correct selection is given by

, where $D_i : p_i = p_{[k]}$ (*i*=١, ..., *k*). ... (٣.٣.٢) $P(CS) = \sum_{i=1}^{k} P(D_i / D_i)$

٢- E(M): Expected sample size

where (*M*) denoted the actual number of observations taken from the given population. An estimate of *E*(*M*) is given by

, where M_j denotes the number of observations in cell $E(M) = \sum_{j=1}^{t} M_j / t$ *j*.

٣.٤ Bayesian (suboptimal) sequential procedure (BSS)

٣.٤.١ Construction of the procedure (BSS)

These procedures are constructed using the following sampling rules

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

٢- $R_2(h)$: Group sequential where observations are taken sequentially in group of (*h*) observations at each stage *N* will be a multiple of *h*, in conjunction with the following stopping rule.

Stop sampling at the sample size *m* when

$$\left(\hat{p}_{[k]} / \hat{p}_{[k-1]}\right) \geq \delta_{\circ}, \quad \text{... (٣.٤.١)}$$

where $\delta_{\circ}(1 < \delta_{\circ} < \infty)$ is preassigned and $\hat{p}_{[i]}$ is the i^{th} ordered posterior mean, (*i*=١, …, *k*). In these cases the best cell is chosen to be that with the largest posterior probability with ties being broken by randomization.

٣.٤.٢ Numerical Results

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

$P(CS), P(NCS)$ (١- $P(CS)$)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 (٥٠٠٠) runs.

Tables (٣-١, ٣-٢) 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 ٤-cells and ٥-cells.

From tables (٣-٣, ٣-٤) 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*.

Over these values in all tables $P(NCS)$ decrease as *N*, $\delta_\circ$ increases.

The effect of group size in R_2 is also investigated, some results are presented in tables (٣-٥, ٣-٦) for *N*=١٠٠, ٢٠٠ and group size (١, ٢,٣, ٤, ٥, ١٠, ٢٠, ٢٥). These tables show that $P(CS)$ and *E*(*M*) increases as *h* increases. Also ,we noted that in the same tables the $P(CS)$ and *E*(*M*) increase as $\delta_\circ$ increase.

Table (٣- ١)

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= ٤, h= ٢ and fixed values of $\underline{p} = (p_1,..., p_4)$.

N=١٠٠	: $\underline{p}$=(.03,.31,.19,.47) $\underline{n'}$=(4,7,5,3) Performance Characteristics			
Sampling Rule	δ_c	$P(CS)$	$P(NCS)$	$E(M)$
R_1	١.٧	٥.٦٤٤٠٠٠E-٠١	٤.٣٥٦٠٠٠E-٠١	٥٠.٨٠٤٦
	١.٨	٦.٠١٢٠٠٠E-٠١	٣.٩٨٨٠٠٠E-٠١	٥٧.٢٩١٦
	١.٩	٧.٤٦٦٠٠٠E-٠١	٢.٥٣٤٠٠٠ E-٠١	٧٥.٢٩٣٢
	٢.١	٨.٤٠٤٠٠٠ E-٠١	١.٥٤٦٠٠٠E-٠١	٨٩.٢٦٦
	٢.٢	٨.٤٦٢٠٠٠E-٠١	١.٥٣٨٠٠٠E-٠١	٩١.٥٩٥٤
	٢.٣	٨.٧٤٨٠٠٠E-٠١	١.٢٥٢٠٠٠E-٠١	٩٥.٣١٩٢
	٢.٥	٨.٩٨٦٠٠٠E-٠١	١.٠١٤٠٠٠E-٠١	٩٨.٣٢٨٤
R_2	١.٧	٥.٩٩٦٠٠٠E-٠١	٤.٠٠٤٠٠٠E-٠١	٥٤.٠٨٤٨
	١.٨	٦.١٢٤٠٠٠E-٠١	٣.٧٨٦٠٠٠E-٠١	٥٩.٢٩٣٢
	١.٩	٧.٩٤٦٠٠٠E-٠١	٢.٠٥٤٠٠٠E-٠١	٨٠.٢٧٧٢
	٢.١	٨.٤٢٠٠٠٠E-٠١	١.٥٨٠٠٠٠E-٠١	٨٩.٦٨١٢
	٢.٢	٨.٥٤٦٠٠٠E-٠١	١.٤٥٤٠٠٠E-٠١	٩٢.٢٣٨٨
	٢.٣	٨.٨٥٤٠٠٠E-٠١	١.١٤٦٠٠٠E-٠١	٩٦.١٥٩٦
	٢.٥	٩.٠١٠٠٠٠E-٠١	٩.٩٠٠٠٠١E-٠٢	٩٨.٧٠٢٨

Table (٣- ٢)

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=٥, h=٢ and fixed values of $\underline{p} = (p_1,..., p_5)$.

N=١٠٠ ; $\underline{P}$ = (.11,.15,.18,.20,.36) n' = (6,9,8,3,7)				
Sampling Rule	$\delta_\circ$	Performance Characteristics		
		$P(CS)$	$P(NCS)$	$E(M)$
R_1	١.٧	٩.٨٢٢٠٠٠E-٠١	١.٧٨٠٠٠٠E-٠٢	٧٧.٧٨٨٢
	١.٨	٩.٨٢٨٠٠٠E-٠١	١.٧٢٠٠٠٠E-٠٢	٨٢.٢٨٦
	١.٩	٩.٨٣٢٠٠٠E-٠١	١.٦٨٠٠٠٠E-٠٢	٩٠.١٩١٢
	٢.١	٩.٨٤٢٠٠٠E-٠١	١.٥٨٠٠٠٠E-٠٢	٩٦.١٩١٨
	٢.٢	٩.٨٤٨٠٠٠E-٠١	١.٥٢٠٠٠٠E-٠٢	٩٧.٨٤٣٤
	٢.٣	٩.٨٥٤٠٠٠E-٠١	١.٤٦٠٠٠٠E-٠٢	٩٨.٧١١٢
	٢.٥	٩.٨٥٦٠٠٠E-٠١	١.٤٤٠٠٠٠E-٠٢	٩٣.٣٧٧٢
R_2	١.٧	٩.٨٢٦٠٠٠E-٠١	١.٧٤٠٠٠٠E-٠٢	٧٨.٧٥٨٠
	١.٨	٩.٨٣٤٠٠٠E-٠١	١.٦٦٠٠٠٠E-٠٢	٨٥.٣١٢٨
	١.٩	٩.٨٣٦٠٠٠E-٠١	١.٦٤٠٠٠٠E-٠٢	٩٠.١٩٨٠
	٢.١	٩.٨٥٨٠٠٠E-٠١	١.٤٢٠٠٠٠E-٠٢	٩٦.٣٩٢٤
	٢.٢	٩.٨٦٤٠٠٠E-٠١	١.٣٦٠٠٠٠E-٠٢	٩٧.٨٨٤٨
	٢.٣	٩.٨٦٦٠٠٠E-٠١	١.٣٤٠٠٠٠E-٠٢	٩٨.٨٥٨٤
	٢.٥	٩.٨٨٦٠٠٠E-٠١	١.١٤٠٠٠٠E-٠٢	٩٩.٥٤٤٨

Table (٣-٣)

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= ٤, h= ٢ and N= ٥٠(١٠) ١٠٠.

$\delta_\circ = 1.5$; $P = (.3,.2,.1,.4)$ $n' = (14,11,16,17)$				
Sampling Rule	N	Performance characteristics		
		$P(CS)$	$P(NCS)$	$E(M)$
R_1	٥٠	٩.٠١٠٠٠٠E-٠١	٩.٩٠٠٠٠٠E-٠٢	٤٠.٣٠٨٦٠٠
	٦٠	٩.٠٧٦٠٠٠E-٠١	٩.٢٤٠٠٠٠E-٠٢	٤٦.٢١٥٤٠٠
	٧٠	٩.١١٦٠٠٠E-٠١	٨.٨٤٠٠٠٠E-٠٢	٥١.٧٢٠٦٠٠
	٨٠	٩.٢٤٢٠٠٠E-٠١	٧.٥٨٠٠٠٠E-٠٢	٥٧.٠٥٤٢٠٠
	٩٠	٩.٢٨٠٠٠٠E-٠١	٧.٢٠٠٠٠٠E-٠٢	٦٢.٨٦٣٢٠٠
	١٠٠	٩.٣٤٢٠٠٠E-٠١	٦.٥٨٠٠٠٠E-٠٢	٦٧.٩٢٠٠٠٠
R_2	٥٠	٩.٠١٦٠٠٠E-٠١	٩.٨٤٠٠٠٠E-٠٢	٤٠.٤٩٦٨٠٠
	٦٠	٩.٠٩٨٠٠٠E-٠١	٩.٠٢٠٠٠١E-٠٢	٤٦.٥٦٨٠٠٠
	٧٠	٩.١٥٠٠٠٠E-٠١	٨.٥٠٠٠٠٠E-٠٢	٥٢.٢٢٨٠٠٠
	٨٠	٩.٢٥٤٠٠٠E-٠١	٧.٤٦٠٠٠٠E-٠٢	٥٨.٣٦٦٤٠٠
	٩٠	٩.٢٨٠٠٠٠E-٠١	٧.٢٠٠٠٠٠E-٠٢	٦٤.٠٠٥٦٠٠
	١٠٠	٩.٣٨٦٠٠٠E-٠١	٦.١٤٠٠٠٠E-٠٢	٦٨.٥٦٠٠٠٠

Table (٣-٤)

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=٥ and N=٥٠(١٠)١٠٠.

$\delta_\circ = 2.2$; $\underline{P} = (.11,.18,.15,.20,.36)$ $n' = (4,5,6,2,3)$				
Sampling Rule	N	Performance Characteristics		
		$P(CS)$	$P(NCS)$	$E(M)$

R_1	٥٠	٨.٣٧٠٠٠٠E-٠١	١.٦٣٠٠٠٠E-٠١	٤٩.١٩٠٨٠٠
	٦٠	٨.٨٥٢٠٠٠E-٠١	١.١٤٨٠٠٠E-٠١	٥٨.٥٥٤٨٠٠
	٧٠	٩.١٥٦٠٠٠E-٠١	٨.٤٤٠٠٠٠E-٠٢	٦٧.٨٢٣٢٠٠
	٨٠	٩.٣٩٦٠٠٠E-٠١	٦.٠٤٠٠٠٠E-٠٢	٧٧.٠٠١٤٠٠
	٩٠	٩.٥٤٤٠٠٠E-٠١	٦.٠٤٠٠٠٠E-٠٢	٨٦.٣٣٣٢٠٠
	١٠٠	٩.٧٥٦٠٠٠E-٠١	٣.١٢٠٠٠٠E-٠٢	٩٥.٣٤١٦٠٠
R_2	٥٠	٨.٥١٢٠٠٠E-٠١	١.٤٨٨٠٠٠E-٠١	٤٩.٢٤٠٠٠
	٦٠	٨.٩٢٨٠٠٠ E-٠١	١.٠٧٢٠٠٠E-٠١	٥٨.٦٦٥٦٠٠
	٧٠	٩.١٨٢٠٠٠ E-٠١	٨.١٨٠٠٠٠E-٠٢	٦٨.٠٧١٦٠٠
	٨٠	٩.٤٧٨٠٠٠ E-٠١	٥.٢٢٠٠٠٠E-٠٢	٧٧.٣٦١٦٠٠
	٩٠	٩.٥٩٦٠٠٠E-٠١	٥.٢٢٠٠٠٠E-٠٢	٨٦.٥٤٧٢٠٠
	١٠٠	٩.٧٨٤٠٠٠E-٠١	٢.٨٦٠٠٠٠E-٠٢	٩٥.٧٥٤٨٠٠

Table (٣-٥)

Performance characteristics of the group sequential schemes for k= ٤, using R_2 (h) sampling method for different group and various $\delta_\circ$, and N= ١٠٠.

$n' = (n'_1,...,n'_4)$	N=١٠٠	h	Performance characteristics

$\underline{p}=(p_1,\ldots,p_4)$	$\delta_\circ$		$P(CS)$	$P(NCS)$	$E(M)$
$\underline{n}$=(٤,٧,٥,٣) $\underline{p}$=(.٠٣,.٣١,.١٩,.٤٧)	١.٨	١	٦.٠١٢٠٠٠E-٠١	٣.٩٨٨٠٠٠E-٠١	٥٧.٢٩١٦
		٢	٦.٢١٤٠٠٠E-٠١	٣.٧٨٦٠٠٠E-٠١	٥٩.٢٩٣٢
		٣	٦.٣٩٢٠٠٠E-٠١	٣.٦٠٨٠٠٠E-٠١	٦٣.٤٢٠٢
		٤	٦.٥٧٠٠٠٠E-٠١	٣.٣٤٣٠٠٠E-٠١	٦٣.٤٨٠
		٥	٧.٦٧٢٠٠٠E-٠١	٢.٣٢٨٠٠٠E-٠١	٧٥.٠٧٢
		١٠	٨.٥٨٢٠٠٠E-٠١	١.٤١٨٠٠٠E-٠١	٨٥.١٥٢
		٢٠	٩.١٢٨٠٠٠E-٠١	٨.٧٢٠٠٠٠E-٠٢	٨٩.٩٠٨
		٢٥	٩.٢٧٤٠٠٠E-٠١	٧.٢٦٠٠٠١E-٠٢	٩١.٤٧٥
	٢.٣	١	٨.٧٤٨٠٠٠E-٠١	١.٢٥٢٠٠٠E-٠١	٩٥.٣١٩٢
		٢	٨.٨٥٤٠٠٠E-٠١	١.١٤٦٠٠٠E-٠١	٩٦.١٥٩٦
		٣	٨.٩٢٨٠٠E-٠١	١.٠٧٢٠٠E-٠١	٩٦.٥٧٨٨
		٤	٨.٩٦٦٠٠٠E-٠١	١.٠٣٤٠٠٠E-٠١	٩٧.٢٠٩٦
		٥	٩.٠٤٦٠٠٠E-٠١	٩.٥٤٠٠٠١E-٠١	٩٧.٨٦٤
		١٠	٩.٠٦٤٠٠٠E-	٩.٣٦٠٠٠٠E-	٩٧.٦٥٤
		٢٠			٩٩.٢٢٠
		٢٥			٩٩.٢٥٠

			٠١	٠٢	
			٩.٣٨٨٠٠٠E-٠١	٦.١٢٠٠٠٠E-٠٢	
			٩.٣٩٤٠٠٠E-٠١	٦.٠٦٠٠٠٠E-٠٢	
			٨.٩٨٦٠٠٠E-٠١	١.٠١٤٠٠٠E-٠١	
			٩.٠١٠٠٠٠E-٠١	٩.٩٠٠٠٠٠E-٠٢	
		١	٩.٠٤٦٠٠٠E-٠١	٩.٥٤٠٠٠١E-٠٢	٩٨.٣٢٨٤
		٢			٩٨.٧٠٢٨
		٣	٩.٠٤٨٠٠٠E-٠١	٩.٥٢٠٠٠٠E-٠١	٩٨.٧١٢٤
		٤			٩٨.٧٥٧٦
	٢.٥	٥	٩.١٤٠٠٠٠E-٠١	٨.٦٠٠٠٠٠E-٠٢	٩٩.٤٣٦
		١٠			٩٩.٤٥٦
		٢٠	٩.١٤٢٠٠٠E-٠١	٨.٥٨٠٠٠٠E-٠٢	٩٩.٦٤٠
		٢٥			٩٩.٧١٥
			٩.٣٦٠٠٠٠E-٠١	٦.٤٠٠٠٠١E-٠٢	
			٩.٤٢٢٠٠٠E-٠١	٥.٧٨٠٠٠٠E-٠٢	

Table (٣- ٦)

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=٥,N=٢٠٠..

$\underline{n}' = (n_1', \ldots, n_5')$	N=٢٠٠	h	Performance Characteristics		
$\underline{p} = (p_1, \ldots, p_5)$	$\delta_\circ$		$P(CS)$	$P(NCS)$	$E(M)$
$\underline{p} = (0.1, 0.3, 0.5, 0.7, 0.9)$ $\underline{n}' = (200, 200, 200, 200, 200)$	١.٥	١	٩.٤٨٨٠٠٠E-٠١	٥.١٢٠٠٠٠E-٠٢	٦٩.٣٤٤٢
		٢	٩.٥٦٦٠٠٠E-٠١	٤.٣٤٠٠٠٠E-٠٢	٧١.٠٦٨٠
		٣	٩.٦٠٠٠٠E-٠١	٤.٠٠٠٠٠E-٠٢	٧٢.٦٤٢
		٤	٩.٧٥٦٠٠٠E-٠١	٢.٤٤٠٠٠٠E-٠٢	٧٥.٢٩٢٨
		٥	٩.٨٠٤٠٠٠E-٠١	١.٩٦٠٠٠٠E-٠٢	٧٦.٣٤٠٠
		١٠	٩.٨٩٤٠٠٠E-٠١	١.٠٦٠٠٠٠E-٠٢	٧٩.٧٥٢٠
		٢٠	٩.٩٣٦٠٠٠E-٠١	٦.٤٠٠٠٠٠E-٠٣	٨٠.٢٠٠٠
		٢٥	٩.٩٧٨٠٠٠E-٠١	٢.٢٠٠٠٠٠E-٠٣	٨١.٥٥٥٠
	١.٦	١	٩.٨٢٠٠٠٠E-٠١	١.٨٠٠٠٠٠E-٠٢	٩٥.١٨٩٦
		٢	٩.٨٦٦٠٠٠E-٠١	١.٣٤٠٠٠٠E-٠٢	٩٦.٤٧٨٨
		٣	٩.٨٨٦٠٠٠E-٠١	١.١٤٠٠٠٠E-٠٢	٩٨.٢٨٦٦
		٤	٩.٩٠٤٠٠٠E-٠١	٩.٦٠٠٠٠٠E-٠٣	٩٩.٣٩٤٤
		٥	٩.٩١٠٠٠٠E-٠١	٩.٠٠٠٠٠٠E-٠٣	١٠٠.٢٠٣
		١٠	٩.٩٦٦٠٠٠E-	٣.٤٠٠٠٠٠E-٠٣	١٠٣.٥٨٦
		٢٠		٢.٨٠٠٠٠٠E-٠٣	١٠٤.٧٢٠
		٢٥		٦.٠٠٠٠٠٠E-٠٤	١٠٥.٧٧٠

			٠١		
			٩.٧٢٠٠٠٠E-		
			٠١		
			٩.٩٩٢٠٠٠E-		
			٠١		
			٩.٩٦٢٠٠٠E-		
			٠١		
			٩.٩٧٠٠٠٠E-		
		١	٠١	٣.٨٠٠٠٠٠E-٠٣	١١٧.٨٧٦٠
		٢	٩.٩٧٤٠٠٠E-	٣.٠٠٠٠٠٠E-٠٣	١١٩.٣٨٨٦
		٤	٠١	٢.٦٠٠٠٠٠E-٠٣	١٢٢.٣٢١٦
	١.٧	٥	٩.٩٨٠٠٠٠E-	٢.٠٠٠٠٠٠E-٠٣	١٢٣.٧٧٥٠
		١٠	٠١	١.٦٠٠٠٠٠E-٠٣	١٢٦.٤٩٤٠
		٢٠	٩.٩٨٤٠٠٠E-	٦.٠٠٠٠٠٠E-٠٤	١٢٨.٤٩٢٠
		٢٥	٠١	٠.٠٠٠٠٠٠E+٠٠	١٢٨.٩٢٥٠
			٩.٩٩٤٠٠٠E-		
			٠١		
			١.٠٠٠٠٠٠		

٣.٥ Bayesian (Suboptimal) Fixed Procedures (BSF)

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

٣.٥.١ Construction of the procedure (BSF)

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 in the next subsection.

٣.٥.٢ Numerical Results

In tables (٣-٧, ٣-٨), 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 (٣-٧)

The effect of N on the performance characteristics using Sampling methods (R_1 and R_3) for k= ٤, 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	δ^* $= ١.٥$	Performance characteristics	
		N	$P(CS)$	$E(M)$
$n'_1 = 10, n'_2 = 20, n'_3 = 30, n'_4 = 40$ $p_1 = 0.1, p_2 = 0.2, p_3 = 0.3, p_4 = 0.4$	R_1	٥٠	٧.٦٥٨٠٠٠E-٠١	٤٦.٧٨٥٦
	R_3		٨.٠٨٦٠٠٠E-٠١	٥٠.٠٠٠
	R_1	٦٠	٨.٤١٨٠٠٠E-٠١	٥٥.١٩٥٨
	R_3		٨.٩٤٤٠٠٠E-٠١	٦٠.٠٠٠
	R_1	٧٠	٨.٨١٠٠٠٠E-٠١	٦٣.٨٠٩٨
	R_3		٩.٤٨٦٠٠٠E-٠١	٧٠.٠٠٠
	R_1	٨٠	٨.٩٥٨٠٠٠E-٠١	٧١.٢٧٩٤
	R_3		٩.٧٢٢٠٠٠E-٠١	٨٠.٠٠٠

	R_1	٩٠	٩.٠١٢٠٠٠E-٠١	٧٧.٧٦١٨
	R_3		٩.٨٥٨٠٠٠E-٠١	٩٠.٠٠٠
	R_1	١٠٠	٩.٠٦٦٠٠٠E-٠١	٨٤.٠٤٣٤
	R_3		٩.٩١٦٠٠٠E-٠١	١٠٠.٠٠٠
$\underline{n}' = (n_1', ..., n_4')$ $\underline{p} = (p_1, ..., p_4)$	Sampling Rule	δ_0 =١.٥ N	Performance characteristics $P(CS)$	$E(M)$
$\underline{p}^{(1)} = (0.4, 0.3, 0.2, 0.1)$ $\underline{p}^{(2)} = (0.3, 0.4, 0.2, 0.1)$ $\underline{p}^{(3)} = (0.3, 0.2, 0.4, 0.1)$ $\underline{p}^{(4)} = (0.3, 0.2, 0.1, 0.4)$ $\underline{p}^{(5)} = (0.25, 0.25, 0.25, 0.25)$	R_1	٥٠	٨.٠٣٨٠٠٠E-٠١	٤٩.٣٠٥٤
	R_3		٨.٢٢٤٠٠٠E-٠١	٥٠.٠٠٠
	R_1	٦٠	٨.٩٧٦٠٠٠E-٠١	٥٨.٧٧٩٢
	R_3		٩.١٦٨٠٠٠E-٠١	٦٠.٠٠٠
	R_1	٧٠	٩.٣٠٢٠٠٠E-٠١	٦٧.٧٠٥
	R_3		٩.٤٩٦٠٠٠E-٠١	٧٠.٠٠٠
	R_1	٨٠	٩.٤٤٠٠٠٠E-٠١	٧٥.٠٤٧٢
	R_3		٩.٦٣٤٠٠٠E-٠١	٨٠.٠٠٠
	R_1	٩٠	٩.٦٠٤٠٠٠E-٠١	٨٢.٣٣٤٢
	R_3		٩.٧٤٢٠٠٠E-٠١	٩٠.٠٠٠
	R_1	١٠٠	٩.٦٥٤٠٠٠E-٠١	٨٨.٤٠٢٦
	R_3		٩.٧١٤٠٠٠E-٠١	١٠٠

Table (٣-٨)

The effect of N on the performance characteristics using sampling methods (R_1 and R_3) for k=٥, 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	δ_0 =١.١ N	Performance characteristics $P(CS)$	$E(M)$
$\underline{p}^{(1)} = (0.5, 0.2, 0.1, 0.1, 0.1)$ $\underline{p}^{(2)} = (0.2, 0.5, 0.1, 0.1, 0.1)$ $\underline{p}^{(3)} = (0.2, 0.1, 0.5, 0.1, 0.1)$ $\underline{p}^{(4)} = (0.2, 0.1, 0.1, 0.5, 0.1)$ $\underline{p}^{(5)} = (0.2, 0.1, 0.1, 0.1, 0.5)$ $\underline{p}^{(6)} = (0.2, 0.2, 0.2, 0.2, 0.2)$	R_1	١٠٠	٩.٨٧٨٠٠٠E-٠١	١١.٣٨٣
	R_3		٩.٨٩٦٠٠٠E-٠١	١٠٠
	R_1	١١٠	٩.٨٨٢٠٠٠E-٠١	١١.٤٥٠
	R_3		٩.٩١٢٠٠٠E-٠١	١١٠

	R_1 R_3	١٢٠	٩.٨٨٨٠٠٠E-٠١ ٩.٩٣٦٠٠٠E-٠١	١١.٤٨٣٦ ١٢٠
	R_1 R_3	١٥٠	٩.٨٨٨٠٠٠E-٠١ ٩.٩٧٦٠٠٠E-٠١	١١.٥٤٧٤ ١٥٠
	R_1 R_3	١٧٠	٩.٨٩٢٠٠٠E-٠١ ٩.٩٨٢٠٠٠E-٠١	١١.٥٧٤ ١٧٠
$\underline{n}' = (n'_1,...,n'_5)$ $\underline{p} = (p_1,..., p_5)$	Sampling Rule	δ_0 $= ١.٢٢$	Performance characteristics	
		N	$P(CS)$	$E(M)$
$\underline{p} = (0.1, 0.3, 0.5, 0.7, 0.9)$ $\underline{n}' = (5, 5, 5, 5, 5)$	R_1 R_3	١٠٠	٩.٧٧٨٠٠٠E-٠١ ٩.٨٥٦٠٠٠E-٠١	٢٥.٤٩٩٤ ١٠٠
	R_1 R_3	١١٠	٩.٧٩٦٠٠٠E-٠١ ٩.٩٠٢٠٠٠E-٠١	٢٥.٩٤٤٤ ١١٠
	R_1 R_3	١٢٠	٩.٨١٨٠٠٠E-٠١ ٩.٩٣٨٠٠٠E-٠١	٢٦.٣٨٠ ١٢٠
	R_1 R_3	١٤٠	٩.٨٢٨٠٠٠E-٠١ ٩.٨٥٦٠٠٠E-٠١	٢٦.٧٢٩٤ ١٤٠
	R_1 R_3	١٥٠	٩.٨٣٦٠٠٠E-٠١ ٩.٩٦٨٠٠٠E-٠١	٢٦.٨٨٢٨ ١٥٠
	R_1 R_3	١٦٠	٩.٨٤٢٠٠٠E-٠١ ٩.٩٨٢٠٠٠E-٠١	٢٧.٠٩٧٢ ١٦٠
	R_1 R_3	١٧٠	٩.٨٤٤٠٠٠E-٠١ ٩.٩٧٨٠٠٠E-٠١	٢٧.١٤٩٢ ١٧٠

٤.١ Conclusions

١. In this thesis 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.

٢. Since reaching a decision as quickly as possible is desirable it seems sensible to employ sequential technique to achieve the aim.

٣. 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.

٤. 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 usethe recursive formula when the number of stages are large makes computations particulariy only for small value of k and N. Therefore, Bayesian suboptimal schemes, which are simple and easy to apply, are proposed.

٤.٢ Suggestions for Future Work

Some suggestions for future work are given as follows:

١- Group sequential sampling can be tried where observations are taken in groups to build Bayesian sequential scheme for the selection problem.

٢- The problem of selecting the least probable cell can be attempted.

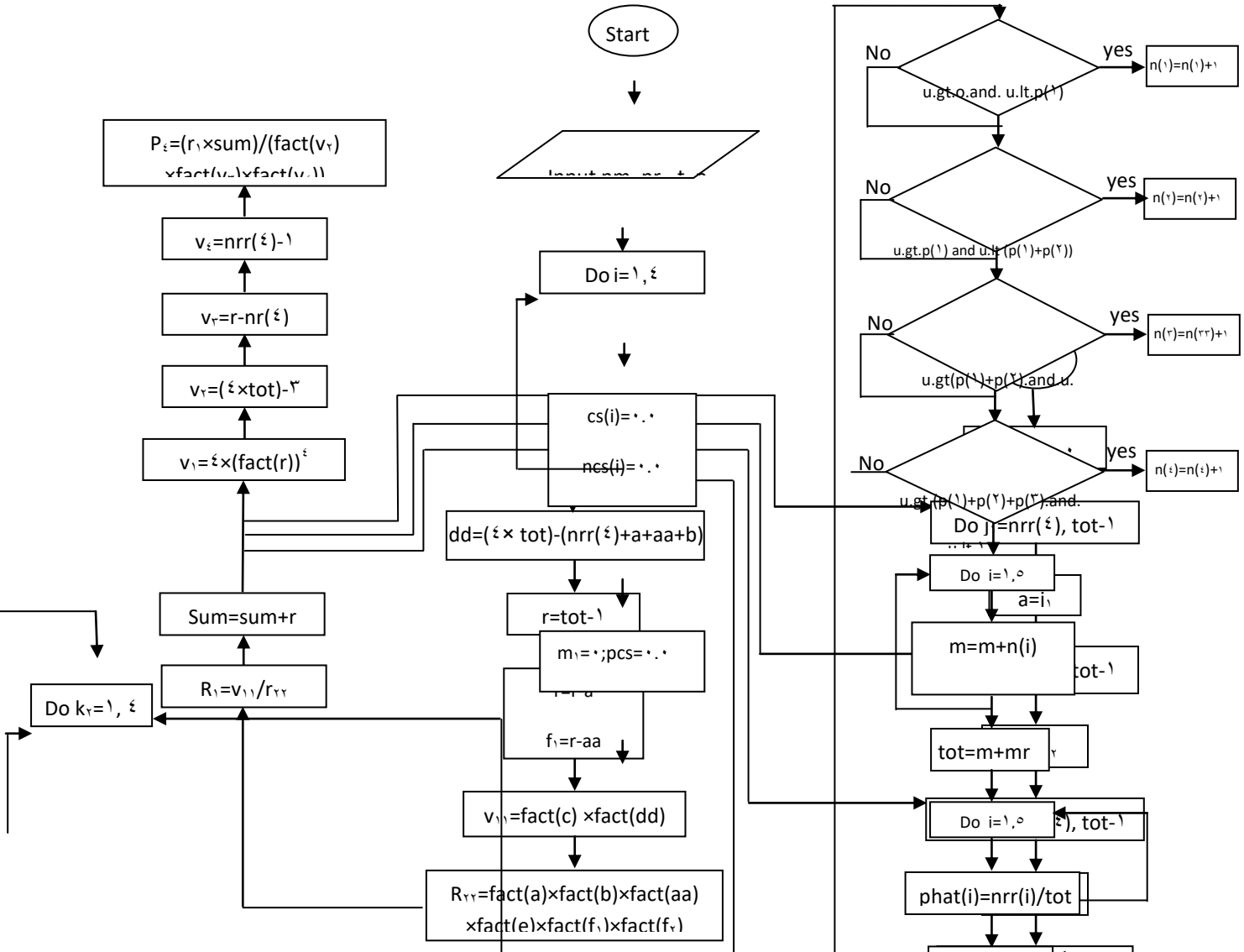
٣- An upper bound for risks may be found using functional analysis.

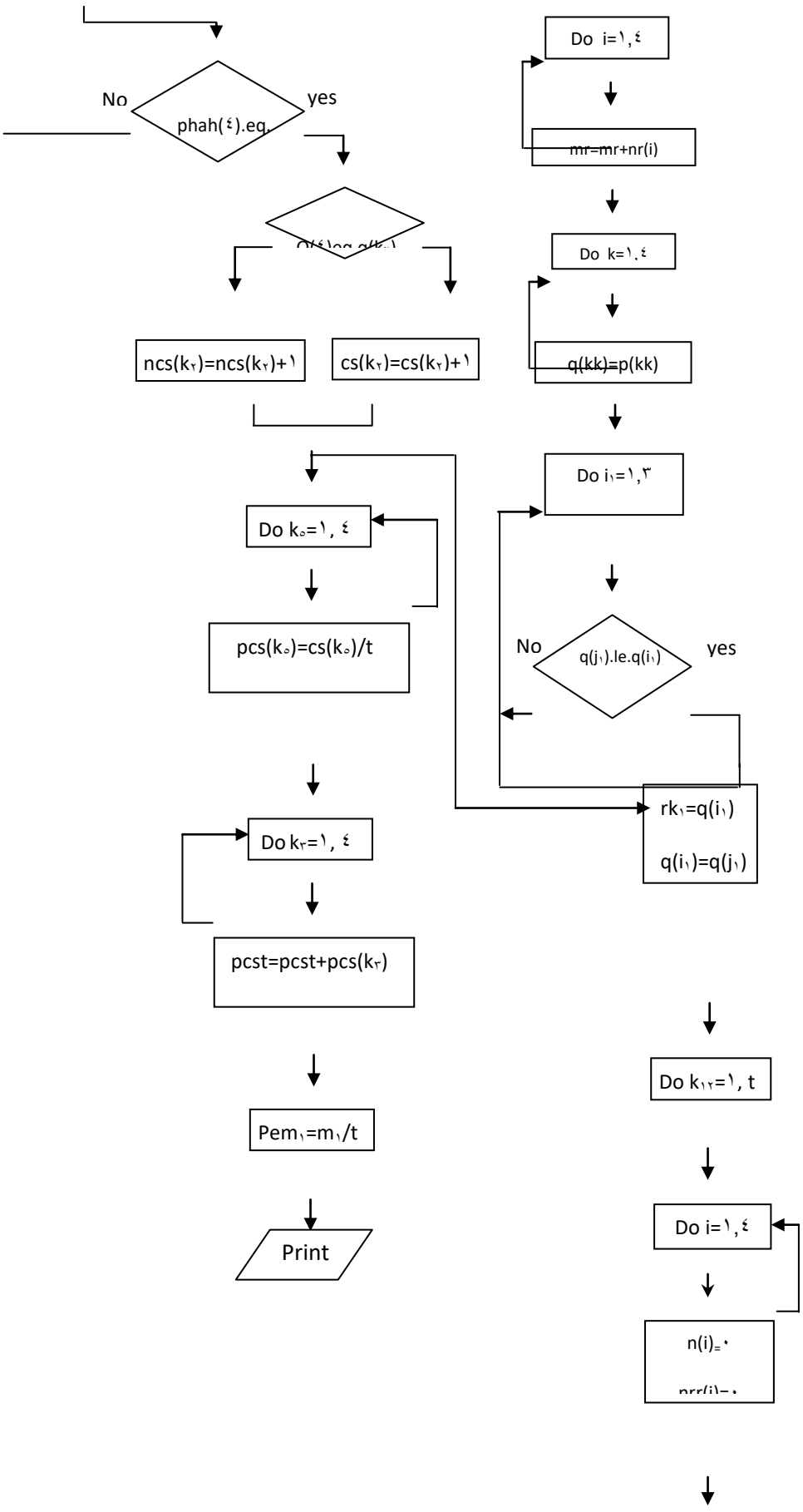
٤- General loss functions may be tried, where linear loss is considered as a special case.

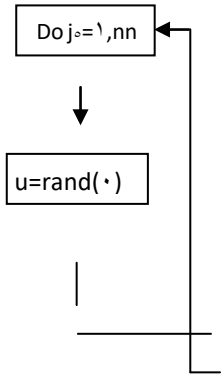
٥- In some problems the experimenter might be interested in selecting a subset of the cells including the best cell. In this problem a correct selection is the selection of any subset including the cell with i^{th} largest probability. Bayesian approach can be used to solve such as a problem.

Flowchart: (A) continued

Flowchart (B): Bayesian (suboptimal) procedures for selecting the best multinomial populations (BSS)




No
phah(٤).eq.
yes
ncs(k٢)=ncs(k٢)+١
cs(k٢)=cs(k٢)+١
Do k٥=١, ٤
pcs(k٥)=cs(k٥)/t
Do k٣=١, ٤
pcst=pcst+pcs(k٣)
Pem١=m١/t
Print
Do i=١,٤
mr=mr+nr(i)
Do k=١,٤
q(kk)=p(kk)
Do i١=١,٣
No
q(j١).le.q(i١)
yes
rk١=q(i١)
q(i١)=q(j١)
Do k١٢=١, t
Do i=١,٤
n(i)=٠



LIST OF TABLES