PAK Publishing Group
Growing Knowledge for Future

Proceedings Book of ICETSR, 2014, Malaysia
Handbook on the Emerging Trends in Scientific Research
ISBN: 978-969-9347-16-0

Bayes Risk for Selection the Median Category from Even Sample Size in K-Nomial Distribution

Kawther Fawzi Hamza
University of Babylon, Iraq, Department of Mathematic

Abstract

Bayes risk procedure is proposed for selecting the median (middle most value) in multinomial cell when the number of observation is even .Bayesian decision –theoretic approach with linear loss function and conjunction prior Dirichlet distribution is used to construct this procedure for it we need to deriving $\underset{\pi(\underline{p}|\underline{n})}{E}[\text{þ}_{[e]}]$. Some concluding remarks and suggestions for future work are also included.

1. Introduction

During the early fifties, it was pointed out by several researchers 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.[4]

Consider a k-nomial distribution which is characterized by k events (cells) with probability vector $\underline{p}=(p_1,,.p_h,p_{h+1},..,p_k)$, where p_i is the probability of the event E_i $(1\le i\le k)$ with $\sum_{i=1}^{k}p_i=1$. When m is even observation, the median dependent on n_h,n_{h+1} and p_h,p_{h+1} , where $h=k/2$. Let $n_1,..,n_h,n_{h+1}..,n_k$ be respective frequencies in k cells of the distribution with $\sum_{i=1}^{k}n_i=m$. Further, let $p_{[1]}\le..\le p_{[h]},p_{[h+1]}\le...\le p_{[k]}$ 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 median probable event, that is the event associated with $\underset{\pi(\underline{p}|\underline{n})}{E}[\text{þ}_{[e]}]$ is also called the median cell . According to this formulation we have a multinomial-decision selection problem.

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)[5]. According to this approach, the selection procedure should guarantee the following probability requirements:

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

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 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 . 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[3].

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[6]. 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[7][8]. In this research we constriction approach to selection the median (middle-most value) cell from multinomial distribution where an even observation and used Bayesian procedure with prior Dirichlet distribution.

2. Review of the Median

2.1. History[9]

The idea of the median originated in Edward Wright's book on navigation in 1599 in a section concerning the determination of location with a compass. In 1757, Roger Boscovich developed a regression method based on the L1 norm and therefore implicitly on the median. The distribution of both the sample mean and the sample median were determined by Laplace in the early 1800s. Antoine Cournot in 1843 was the first to use the term median for the value that divides a probability distribution into two equal halves. Gustav Theodor Fechner used the median in sociological and psychological phenomena. Gustav Fechner popularized the median into the formal analysis of data, although it had been used previously by Laplace. Francis Galton used the English term median in 1881, having earlier used the terms middle-most value in 1869 and the medium in 1880.

2.2. Measure of Location[1]

The median is one of a number of ways of summarizing the typical values associated with members of a statistical population; thus, it is a possible location parameter.

In statistics and probability theory, the median is the numerical value separating the higher half of a data sample, a population, or probability distribution, from the lower half. The median of a finite list of numbers can be found by arranging all the observations from lowest value to highest value and picking the middle one. If there is an even number of observations, then there is no single middle value ; the median is then usually defined to be the mean of the two middle values, which corresponds to interpreting the median as the fully trimmed mid-range. The median is of central importance in robust statistics, as it is the most resistant statistic, having a breakdown point 50% : so long as no more than half the data is contaminated , the median will not give an arbitrarily large result. A median is only defined on ordered one-dimensional data, and is independent of any number of dimensions.

2.3. Some Related Concepts with Median and Applications[8]

The median is used primarily for skewed distributions, which it summarizes differently than the arithmetic mean , and it might be seen as a better indication of central tendency (less susceptible to the exceptionally large value in data) than the arithmetic mean .Calculation of medians is a popular technique in summary statistics and summarizing statistical data, since it is simple to understand and easy to calculate, while also giving a measure that is more robust in the presence of outlier values than is the mean.

Pseudo-median:-For univariate distributions that are symmetric about one median, the Hodges–Lehmann estimator is a robust and highly efficient estimator of the population median; for non-symmetric distributions, the Hodges–Lehmann estimator is a robust and highly efficient estimator of the population pseudo-median, which is the median of a symmetrized distribution and which is close to the population median

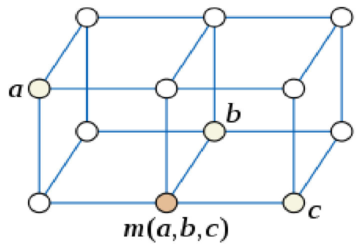
Median filter:- In signal processing, it is often desirable to be able to perform some kind of noise reduction on an image or signal. The median filter is a nonlinear digital filtering technique, often used to remove noise.

Cluster analysis :-In cluster analysis, the k-medians clustering algorithm provides a way of defining clusters, in which the criterion of maximizing the distance between cluster-means that is used in k-means clustering, is replaced by maximizing the distance between cluster-medians.

Median-Median Line:-This is a method of robust regression. The idea dates back to Wald in 1940 who suggested dividing a set of bivariate data into two halves depending on the value of the independent parameter x: a left half with values less than the median and a right half with values greater than the median**.**

Median graph:- In mathematics, and more specifically graph theory, a median graph is an undirected graph in which ever y three vertices a, b, and c have a unique median: a vertex m(a,b,c) that belongs to shortest paths between each pair of a, b, and c.

The median of three vertices in a median graph



Some applications are as follows.

- *the highest flood waters (useful when planning for future emergencies).*
- *In non-parametric statistics, the Theil–Sen estimator, also known as Sen's slope estimator, slope selection, the single median method is a method for robust linear regression that chooses the median slope among all lines through pairs of two-dimensional sample points.*
- *the lowest winter temperature recorded in the last 50 years.*
- *the median price of houses sold in last month.*
- *A medical research team conducts a clinical study comparing the success rates of five different drug regiments for a particular disease.*
- *A median of a triangle is a line segment that joins the vertex of a triangle to the midpoint of the opposite side.*
- *The geometric median of a discrete set of sample points in a Euclidean space is the point minimizing the sum of distances to the sample points.*
- *In object prototype learning and similar tasks, median computation is an important technique for capturing the essential information of a given set of patterns. We extend the median concept to the domain of graphs.*

2.4. Medians of Probability Distributions

For any probability distribution on the real line **R** with cumulative distribution function *F*, regardless of whether it is any kind of continuous probability distribution, in particular an absolutely continuous distribution (which has a probability density function), or a discrete probability distribution, a median is by definition any real number *m* that satisfies the inequalities

$$\mathrm{P}(X \le m) \ge \frac{1}{2} \text{ and } \mathrm{P}(X \ge m) \ge \frac{1}{2}$$

Or, equivalently, the inequalities

$$\int_{(-\infty,m]} dF(x) \ge \frac{1}{2} \text{ and } \int_{[m,\infty)} dF(x) \ge \frac{1}{2}$$

in which a Lebesgue–Stieltjes integral is used. For an absolutely continuous probability distribution with probability density function *f*, the median satisfies

$$\mathrm{P}(X \le m) = \mathrm{P}(X \ge m) = \int_{-\infty}^{m} f(x)\, dx = \frac{1}{2}.$$

Any probability distribution on **R** has at least one median, but there may be more than one median. Where exactly one median exists, statisticians speak of "the median" correctly; even when the median is not unique, some statisticians speak of "the median" informally.

3. Bayesian Procedure for Selecting the Median Multinomial Category (Cell)

3.1. 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 $\Omega_k : \{\underline{p} = (p_1,.., p_h, p_{h+1},..., p_k) : \sum_{i=1}^{k} p_i = 1 \; ; p_i \ge 0\}$ be the parameter space and $D = \{d_1,.., d_h, d_{h+1},..., d_k\}$ be the decision space where in the following terminal *k*-decision rule:

$d_i : p_i$ is the median cell probability $(i = 1,..,h.h+1,...,k)$ .That is, d_i denote the decision to select the event associated with the i^{th} cell as the median probable event, after the sampling is terminated.

Let $p_{[1]} \le ... \le p_{[h]} \le p_{[h+1]} \le ... \le p_{[k]}$ denote the ordered values of the , $p_i (1 \le i \le k)$ the goal of the experimenter is to select the least cell probability, that is the cell associated with $p_{[e]}$.

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

$$L(d_i, \underline{p}^*) = \begin{Bmatrix} k^*(p_{[e]} - p_i) & \text{if } (p_{[e]} \ne p_i) \\ 0 & \text{if } (p_{[e]} = p_i) \end{Bmatrix} \quad .... (2)$$

That is the loss if decision d_i is made when the true value of $\underline{p} = \underline{p}^*$. Where $p_{[e]}$ is the middle most value or equal to $p_{[h]} + p_{[h+1]} / 2$ and 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 $m', n'_1, n'_h, n'_{h+1}..., n'_k$. The normalized density function is given by

$$\pi(\underline{p}) = \frac{\Gamma\left(\sum_{i=1}^{k} n'_i\right)}{\prod_{i=1}^{k}\Gamma(n'_i)} \prod_{i=1}^{k} p_i^{n'_i-1}, \text{ where } m' = \sum_{i=1}^{k} n'_i \qquad \ldots\ldots (3)$$

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 $\underline{n}' = (n'_1,..,n'_h,n'_{h+1},..,n'_k)$, 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

$$P_r(N_1 = n_1,..,N_h = n_h, N_{h+1} = n_{h+1},.., N_k = n_k \mid p_1,..., p_k) = P(\underline{n} \mid \underline{p})$$

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

Since

$$P(\underline{n} \mid \underline{p}) \propto p_1^{n_1}...p_1^{n_h} p_1^{n_{h+1}}..\ p_k^{n_k} \text{ and } \pi(\underline{p}) \propto p_1^{n'_1-1}..p_h^{n'_h-1} p_{h+1}^{n'_{h+1}-1}..\ p_k^{n'_k-1},$$

then the posterior is $\pi(\underline{p} \mid \underline{n}) \propto p_1^{n_1+n'_1-1}...p_1^{n_h+n'_h-1} p_1^{n_{h+1}+n'_{h+1}-1}...\ p_k^{n_k+n'_k-1}$

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

$$\pi(\underline{p} \mid \underline{n}) = \frac{(m''-1)!}{(n''_1-1)!..(n''_h-1)!(n''_{h+1}-1)!...(n''_k-1)!} p_1^{n''_1-1}...p_h^{n''_h-1} p_{h+1}^{n''_{h+1}-1}...p_k^{n''_k-1}$$

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.

3.2. The Stopping Risks

In this section, we derive the stopping risks (Bayes risk) of making decision d_i for linear loss function. The stopping risk (the posterior expected loss) of the terminal decision d_i when the posterior distribution for $\underline{p}$ has parameters $(n''_1,..,n''_h,n''_{h+1}...,n''_k;m'')$, that is when the sample path has reached $(n''_1,..,n''_h,n''_{h+1}...,n''_k;m'')$ from the origin $(n'_1,..n'_h,n'_{h+1},.,n'_k;m')$, denoted by $S_h(n''_1,..,n''_h,n''_{h+1}...,n''_k;m'')$ can be found as follows.

$$S_i(n''_1,..n''_h,n''_{h+1},..,n''_k;m'') = \underset{\pi(\underline{p}|\underline{n})}{E}[L(d_i,\underline{p}^*)]$$

$$= k^*\left[\underset{\pi(\underline{p}|\underline{n})}{E}(\text{þ}_{[e]}) - \frac{n''_i}{m''}\right] \qquad \ldots\ldots (4)$$

Since the number of observation is even, the value of $\underset{\pi(\underline{p}|\underline{n})}{E}[\text{þ}_{[e]}]$ is equal to $\underset{\pi(\underline{p}|\underline{n})}{E}[f(\text{þ}_{[h]},\text{þ}_{[h+1]})]$ and is derived as follows.

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

Where $g(p_{[h]}, p_{[h+1]})$ be the joint probability density function of the median order statistics $p_{[h]}$, $p_{[h+1]}$. $g(p_{[h]}, p_{[h+1]}) = \frac{k!}{(h-1)!(k-h-1)!}\left[F(p_{[h]})\right]^{h-1}\left[1-F(p_{[h+1]})\right]^{k-h-1} f(p_{[h]}) f(p_{[h+1]})$

and

$f(p_{[h]}, p_{[h+1]}) = \frac{p_{[h]} + p_{[h+1]}}{2}$. Let the ordered values of $n''_1,..,n''_h, n''_{h+1},..,n''_k$ is $n''_{[1]} \le,.. \le n''_{[h]}, \le n''_{[h+1]} \le,..,\le n''_{[k]}$. The marginal posterior probability density function of $p_{[h]}$ & $p_{[h+1]}$ are respectively

$$f(p_{[h]}) = \frac{(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!} p_{[h]}^{n''_{[h]}-1}(1-p_{[h]})^{m''-n''_{[h]}-1} ,$$

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

and the cumulative density function of $p_{[h]}$ & $p_{[h+1]}$ are respectively

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

$$F(p_{[h+1]}) = \sum_{L=n''_{[h+1]}}^{m''-1} \frac{(m''-1)!}{L!(m''-1-L)!}.p_{[h+1]}{}^{L}(1-p_{[h+1]})^{m''-1-L}$$

Then,

$$g(p_{[h]}, p_{[h+1]}) = \frac{k!}{(h-1)!(k-h-1)!}\left[\sum_{j=n''_{[h]}}^{m''-1} \frac{(m''-1)!}{j!(m''-1-j)!}.p_{[h]}{}^{j}(1-p_{[h]})^{m''-1-j}\right]^{h-1}\left[1-\left\langle \sum_{l=n''_{[h+1]}}^{m''-1} \frac{(m''-1)!}{l!(m''-1-l)!} p_{[h+1]}{}^{l}(1-p_{[h+1]})^{m''-1-l}\right\rangle\right]^{k-h-1} \left\langle \frac{(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!} p_{[h]}^{n''_{[h]}-1}(1-p_{[h]})^{m''-n''_{[h]}-1}\right\rangle \left\langle \frac{(m''-1)!}{(n''_{[h+1]}-1)!(m''-n''_{[h+1]}-1)!} p_{[h+1]}^{n''_{[h+1]}-1}(1-p_{[h+1]})^{m''-n''_{[h+1]}-1}\right\rangle . \quad \text{...........(5)}$$

$$\underset{\pi(\underline{p}|\underline{n})}{E}[f(p_{[h]}, p_{[h+1]})] = \int_0^1\int_0^1 \left\{\left(\frac{p_{[h]}+p_{[h+1]}}{2}\right)\left(\frac{k!}{(h-1)!(k-h-1)!}\right)\left[\sum_{j=n''_{[h]}}^{m''-1} \frac{(m''-1)!}{j!(m''-1-j)!}.p_{[h]}{}^{j}(1-p_{[h]})^{m''-1-j}\right]^{h-1} \left[1-\left\langle \sum_{l=n''_{[h+1]}}^{m''-1} \frac{(m''-1)!}{l!(m''-1-l)!}.p_{[h+1]}{}^{l}(1-p_{[h+1]})^{m''-1-l}\right\rangle\right]^{k-h-1} \left\langle \frac{(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!} p_{[h]}^{n''_{[h]}-1}(1-p_{[h]})^{m''-n''_{[h]}-1}\right\rangle\right.$$

Hence,

$$\left.\left\langle \frac{(m''-1)!}{(n''_{[h+1]}-1)!(m''-n''_{[h+1]}-1)!} p_{[h+1]}^{n''_{[h+1]}-1}(1-p_{[h+1]})^{m''-n''_{[h+1]}-1}\right\rangle\right\} dp_{[h]}dp_{[h+1]}$$

$$= \int_0^1\int_0^1 \left\{\left(\frac{k!}{2(h-1)!(k-h-1)!}\right)(p_{[h]}+p_{[h+1]})\left\langle \frac{(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!} p_{[h]}^{n''_{[h]}-1}(1-p_{[h]})^{m''-n''_{[h]}-1}\right\rangle \left\langle \frac{(m''-1)!}{(n''_{[h+1]}-1)!(m''-n''_{[h+1]}-1)!} p_{[h+1]}^{n''_{[h+1]}-1}(1-p_{[h+1]})^{m''-n''_{[h+1]}-1}\right\rangle \left[\sum_{j=n''_{[h]}}^{m''-1} \frac{(m''-1)!}{j!(m''-1-j)!}.p_{[h]}{}^{j}(1-p_{[h]})^{m''-1-j}\right]^{h-1} \left[1-\left\langle \sum_{l=n''_{[h+1]}}^{m''-1} \frac{(m''-1)!}{l!(m''-1-l)!}.p_{[h+1]}{}^{l}(1-p_{[h+1]})^{m''-1-l}\right\rangle\right]^{k-h-1}\right\} dp_{[h]}dp_{[h+1]}$$

$$= \int_0^1 \int_0^1 \left(\frac{(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!} \right) \left(\frac{(m''-1)!}{(n''_{[h+1]}-1)!(m''-n''_{[h+1]}-1)!} \right) \left(\frac{k!}{2(h-1)!(k-h-1)!} \right)$$

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

$$\left(p_{[h]} + p_{[h+1]} \right) \left\langle p_{[h]}^{n''_{[h]}-1} (1-p_{[h]})^{m''-n''_{[h]}-1} \right\rangle \left\langle p_{[h+1]}^{n''_{[h+1]}-1} (1-p_{[h+1]})^{m''-n''_{[h+1]}-1} \right\rangle dp_{[h]} dp_{[h+1]}$$

$$= \left(\frac{(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!} \right) \left(\frac{(m''-1)!}{(n''_{[h+1]}-1)!(m''-n''_{[h+1]}-1)!} \right) \left(\frac{k!}{2(h-1)!(k-h-1)!} \right)$$

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

$$\left. \left\langle p_{[h]}^{n''_{[h]}} - p_{[h]}{}^{m''-1} \right\rangle \left\langle p_{[h+1]}^{n''_{[h+1]}-1} - p_{[h+1]})^{m''-2} \right\rangle dp_{[h]} dp_{[h+1]} \right\} + \left\{ \int_0^1 \int_0^1 \left[\sum_{j=n''_{[h]}}^{m''-1} \frac{(m''-1)!}{j!(m''-1-j)!} . p_{[h]}{}^{j} (1-p_{[h]})^{m''-1-j} \right]^{h-1} \right.$$

$$\left. \left[1 - \left\langle \sum_{l=n''_{[h+1]}}^{m''-1} \frac{(m''-1)!}{l!(m''-1-l)!} . p_{[h+1]}{}^{l} (1-p_{[h+1]})^{m''-1-l} \right\rangle \right]^{k-h-1} \left\langle p_{[h]}^{n''_{[h]-1}} - p_{[h]}{}^{m''-2} \right\rangle \left\langle p_{[h+1]}^{n''_{[+1h]}} - p_{[h+1]})^{m''-1} \right\rangle \right\} dp_{[h]} dp_{[h+1]}$$

Then,

$$\underset{\pi(\underline{p}|\underline{n})}{E} [f(\text{þ}_{[h]}, \text{þ}_{[h+1]})] = \left(\frac{(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!} \right) \left(\frac{(m''-1)!}{(n''_{[h+1]}-1)!(m''-n''_{[h+1]}-1)!} \right) \left(\frac{k!}{2(h-1)!(k-h-1)!} \right)$$

$$\left\{ \int_0^1 \int_0^1 \left[(1-p_{[h]})^{(m''-1)(h-1)} \sum_{j_1=n''_{[k]}}^{m''-1} \sum_{j_2=n''_{[k]}}^{m''-1} \cdots\cdots \sum_{j_{k-1}=n''_{[h]}}^{m''-1} \frac{\left(\frac{p_{[h]}}{1-p_{[h]}} \right)^{j_1+j_2+\ldots+j_{h-1}}}{j_1!(m''-j_1-1)!\ldots j_{h-1}(m''-j_{h-1}-1)!} \right] \left[\sum_{J=0}^{k-h-1} \binom{k-h-1}{J} \right. \right.$$

$$\left. (-1)^J ((m''-1)!)^J \sum_{J_1=n''_{[1]}}^{m''-1} \sum_{J_2=n''_{[1]}}^{m''-1} \cdots\cdots\cdots \sum_{J_l=n''_{[1]}}^{m''-1} \left(\frac{p_{[h+1]}}{1-p_{[h+1]}} \right)^{J_1+J_2+\ldots+J_l} \frac{p_{[h+1]}^{n''_{[h+1]}} (1-p_{[h+1]})^{m''-n''_{[h+1]}-1}}{j_1!(m''-j_1-1)!\ldots j_l!(m''-j_l-1)!} \right]$$

$$\left. \left\langle p_{[h]}^{n''_{[h]}} - p_{[h]}{}^{m''-1} \right\rangle \left\langle p_{[h+1]}^{n''_{[h+1]}-1} - p_{[h+1]})^{m''-2} \right\rangle dp_{[h]} dp_{[h+1]} \right\} + \left\{ \int_0^1 \int_0^1 \left[(1-p_{[h]})^{(m''-1)(h-1)} \sum_{j_1=n''_{[k]}}^{m''-1} \sum_{j_2=n''_{[k]}}^{m''-1} \cdots\cdots \sum_{j_{k-1}=n''_{[h]}}^{m''-1} \right. \right.$$

$$\left. \frac{\left(\frac{p_{[h]}}{1-p_{[h]}} \right)^{j_1+j_2+\ldots+j_{h-1}}}{j_1!(m''-j_1-1)!\ldots j_{h-1}(m''-j_{h-1}-1)!} \right] \left[\sum_{J=0}^{k-h-1} \binom{k-h-1}{J} (-1)^J ((m''-1)!)^J \sum_{J_1=n''_{[1]}}^{m''-1} \sum_{J_2=n''_{[1]}}^{m''-1} \cdots \sum_{J_l=n''_{[1]}}^{m''-1} \left(\frac{p_{[h+1]}}{1-p_{[h+1]}} \right)^{J_1+J_2+\ldots+J_l} \right.$$

$$\left. \left. \frac{p_{[h+1]}^{n''_{[h+1]}} (1-p_{[h+1]})^{m''-n''_{[h+1]}-1}}{j_1!(m''-j_1-1)!\ldots j_l!(m''-j_l-1)!} \right] \left\langle p_{[h]}^{n''_{[h]-1}} - p_{[h]}{}^{m''-2} \right\rangle \left\langle p_{[h+1]}^{n''_{[h+1]}} - p_{[h+1]})^{m''-1} \right\rangle \right\} dp_{[h]} dp_{[h+1]}$$

$$\underset{\pi(\underline{p}|\underline{n})}{E} [\text{þ}_{[e]}] = \left(\frac{(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!} \right) \left(\frac{(m''-1)!}{(n''_{[h+1]}-1)!(m''-n''_{[h+1]}-1)!} \right) \left(\frac{k!}{2(h-1)!(k-h-1)!} \right)$$

$$\left\{ \sum_{j_1=n''_{[k]}}^{m''-1} \sum_{j_2=n''_{[k]}}^{m''-1} \cdots \sum_{j_{k-1}=n''_{[h]}}^{m''-1} 1 \big/ j_1!(m''-j_1-1)!\ldots j_{h-1}(m''-j_{h-1}-1)! \; \left\langle \Gamma(n''_{[h]} + j_1 + \ldots + j_{h-1}) \right. \right.$$

$$\left. \frac{\Gamma(m''h-h-m''-j_1-\ldots-j_{h-1})}{\Gamma(n''_{[h]}+m''h-m''-h)} - \frac{\Gamma(m''+j_1+\ldots+j_{h-1}-1)\Gamma(m''h-h-m''-j_1-\ldots-j_{h-1})}{\Gamma(m''h-h)} \right\rangle$$

$$\left[\sum_{J=0}^{k-h-1} \binom{k-h-1}{J} (-1)^J ((m''-1)!)^J \sum_{J_1=n''_{[1]}}^{m''-1} \sum_{J_2=n''_{[1]}}^{m''-1} \cdots \sum_{J_l=n''_{[1]}}^{m''-1} 1 \big/ j_1!(m''-j_1-1)!\ldots j_l!(m''-j_l-1)! \left\langle . \Gamma(n''_{[h+1]} + J_1 + .. \right. \right.$$

$$...+J_{l-1})\frac{\Gamma(Jm''-J+1-J_1-....-J_{l-1})}{\Gamma(n''_{[h+1]}+Jm''-J)}-\frac{.\Gamma(Jm''-J-J_1...-J_{l-1}+1)\Gamma(m''-1-J_1-....-J_{l-1})}{\Gamma(m''+Jm''-J-1)}\Big\rangle\Big\}$$

$$+\Bigg\{\sum_{j_1=n''_{[k]}}^{m''-1}\sum_{j_2=n''_{[k]}}^{m''-1}.....\sum_{j_{k-1}=n''_{[h]}}^{m''-1}\frac{1}{j_1!(m''-j_1-1)!...j_{h-1}(m''-j_{h-1}-1)!}\Big\langle\Gamma(n''_{[h]}+j_1+....+j_{h-1})$$

$$\frac{\Gamma(m''h-h-m''-2-j_1-....-j_{h-1})}{\Gamma(n''_{[h]}+m''h-m''-h-2)}-\frac{\Gamma(m''+j_1+....+j_{h-1}-1)\Gamma(m''h-h-m''+2-j_1-....-j_{h-1})}{\Gamma(m''h-h-1)}\Big\rangle$$

$$\Bigg[\sum_{J=0}^{k-h-1}\binom{k-h-1}{J}(-1)^J((m''-1)!)^J\sum_{J_1=n''_{[1]}}^{m''-1}\sum_{J_2=n''_{[1]}}^{m''-1}...\sum_{J_l=n''_{[1]}}^{m''-1}\frac{1}{J_1!(m''-J_1-1)!...J_l!(m''-J_l-1)!}\Big\langle\Gamma(2n''_{[h+1]}+J_1$$

$$+...+J_l+1)\frac{\Gamma(m''-n''_{[h+1]}-J_1+...+J_l)}{\Gamma(n''_{[h+1]}+m''+1)}-\frac{\Gamma(m''+n''_{[h+1]}+J_1+...+J_l)\Gamma(m''-n''_{[h+1]}-J_1+...+J_l)}{\Gamma(2m'')}\Big\rangle\Big\}$$

Hence

$$S_i(n''_1,..n''_h,n''_{h+1},.,n''_k;m'')=k^*\Bigg\{\Bigg(\frac{k!(m''-1)!(m''-1)!}{(n''_{[h]}-1)!(m''-n''_{[h]}-1)!2(h-1)!(k-h-1)!(n''_{[h+1]}-1)!(m''-n''_{[h+1]}-1)!}\Bigg)$$

$$.\Bigg\{\sum_{j_1=n''_{[k]}}^{m''-1}\sum_{j_2=n''_{[k]}}^{m''-1}....\sum_{j_{k-1}=n''_{[h]}}^{m''-1}1/j_1!(m''-j_1-1)!...j_{h-1}(m''-j_{h-1}-1)!\ \Big\langle\Gamma(n''_{[h]}+j_1+....+j_{h-1})$$

$$\frac{\Gamma(m''h-h-m''-j_1-....-j_{h-1})}{\Gamma(n''_{[h]}+m''h-m''-h)}-\frac{\Gamma(m''+j_1+....+j_{h-1}-1)\Gamma(m''h-h-m''-j_1-....-j_{h-1})}{\Gamma(m''h-h)}\Big\rangle$$

$$\Bigg[\sum_{J=0}^{k-h-1}\binom{k-h-1}{J}(-1)^J((m''-1)!)^J\sum_{J_1=n''_{[1]}}^{m''-1}\sum_{J_2=n''_{[1]}}^{m''-1}....\sum_{J_l=n''_{[1]}}^{m''-1}1\Big/j_1!(m''-j_1-1)!...j_l!(m''-j_l-1)!\Big\langle.\Gamma(n''_{[h+1]}+J_1+..$$

$$...+J_{l-1})\frac{\Gamma(Jm''-J+1-J_1-....-J_{l-1})}{\Gamma(n''_{[h+1]}+Jm''-J)}-\frac{.\Gamma(Jm''-J-J_1...-J_{l-1}+1)\Gamma(m''-1-J_1-....-J_{l-1})}{\Gamma(m''+Jm''-J-1)}\Big\rangle\Big\}$$

$$+\Bigg\{\sum_{j_1=n''_{[k]}}^{m''-1}\sum_{j_2=n''_{[k]}}^{m''-1}.....\sum_{j_{k-1}=n''_{[h]}}^{m''-1}\frac{1}{j_1!(m''-j_1-1)!...j_{h-1}(m''-j_{h-1}-1)!}\Big\langle\Gamma(n''_{[h]}+j_1+....+j_{h-1})$$

$$\frac{\Gamma(m''h-h-m''-2-j_1-....-j_{h-1})}{\Gamma(n''_{[h]}+m''h-m''-h-2)}-\frac{\Gamma(m''+j_1+....+j_{h-1}-1)\Gamma(m''h-h-m''+2-j_1-....-j_{h-1})}{\Gamma(m''h-h-1)}\Big\rangle$$

$$\Bigg[\sum_{J=0}^{k-h-1}\binom{k-h-1}{J}(-1)^J((m''-1)!)^J\sum_{J_1=n''_{[1]}}^{m''-1}\sum_{J_2=n''_{[1]}}^{m''-1}...\sum_{J_l=n''_{[1]}}^{m''-1}\frac{1}{J_1!(m''-J_1-1)!...J_l!(m''-J_l-1)!}\Big\langle\Gamma(2n''_{[h+1]}+J_1$$

$$+.+J_l+1)\frac{\Gamma(m''-n''_{[h+1]}-J_1+.+J_l)}{\Gamma(n''_{[h+1]}+m''+1)}-\frac{\Gamma(m''+n''_{[h+1]}+J_1+.+J_l)\Gamma(m''-n''_{[h+1]}-J_1+.+J_l)}{\Gamma(2m'')}\Big\rangle\Big\}-\frac{n''_i}{m''}\Big\}$$

…….(7)

4. Conclusion and Directions for Future Work

4.1. Conclusions

The median is the one of a number of ways of summarizing the typical values associated with members of a statistical population; thus, it is a possible location parameter. Ranking and selection procedures provides excellent tools for selecting the middle most event when an even of observation 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.

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 any two order statistic $p_{[i]}, p_{[j]}$ probable cells can be attempted.

3- To simplify the formula (7) we can used 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 median cell. In this problem a correct selection is the selection of any subset including the cell with i^{th} median probability. Bayesian approach can be used to solve such as a problem.

7- we can used other measure location eg: range , mid–range,…etc.

References

1. D.Dor, and U. Zwick. Selecting the Median. SIAM Jour. Comp., 28(5):1722–1758, 1999.
2. Edson L. F. Sane, lagrangean/surrogate heuristics for p-median problems, Chapter 6, 2000.
3. Horrace, W. C. Selection procedures for order statistics in Empirical Economic studies, 1-26,2002.
4. Jones, p.w. and Madhi, S.A , Bayesan sequential methods for choosing the best Multinomial cell: Some simulation results. Statistical paper 29, 125-132, 1988.
5. Kim, S.H., and Nelson, B.L, (2003). Selecting the best system. Proceedings Winter simulation conference, Georgia Institute of Technology, North Western Universiy, 2003.
6. Mausumi, B. and Subir, K.B., Selecting the *t* Best cells of a Multinomial distribution. The Indian Journal of statistics, volume 61, series A, pt 1, pp. 139-147, Calcutta, 1999. http://202.54.54.1471/search/61a1/61a16926.pdf
7. Ranking and selection: a comprehensive survey on the theory, Advanced Simulation. http://www.bilkent.edu.tr/~goren/simulation.htmL
8. S. Battiato, D. C, C and G. Cincotti , An Efficient Algorithm for the Approximate Median Selection Problem. Department di Matematica,Viale A. Doria 6, I-95125 Catania, Italy, Department Of Computer Science , 01609-2280 ,October 28, 2002.
9. Median http;//cn.wikipedia.org/wiki/median.