A NEW APPROACH: BAYESIAN SELECTION TO FIND MEDIAN CELL (VALUE) IN MULTINOMIAL POPULATION

نظرة جديدة :الاختيار البيزيني لإيجاد الخلية (القيمة) المتوسطة في مجتمع متعدد الحدود

كوثر فوزي الحسن
كلية التربية (بن حيان)-قسم الرياضيات
جامعة بابل ,2010

ABSTRACT

Median finding is an essential problem in statistics, it is provides a more robust notion of average than the mean. The measure location median was suggested to find the middle among collection multinomial ordered values. Bayesian procedure to select the median population is presented .prior distribution and linear loss function to find the Bayes risk are used. When the number of observation is odd.

الملخص

أيجاد الوسيط مشكلة ضرورية في الإحصائيات, حيث يزود فكرة أكثر متانة للمعدل من الوسط.مقياس الموقع الوسيط أقترح لإيجاد المنتصف من بين العديد من القيم المرتبة. الإجراء البيزيني لاختيار المجتمع المتوسط (الوسيط) قدم . التوزيع السابق ودالة خسارة خطية للوصول إلى الخطورة البيزينية استعملت .عندما عدد المشاهدات هي قيم فردية.

1- INTRODUCTION

Over the last twenty years there has been considerable effort expended to develop statistically valid ranking-and-selection (R&S) procedures to compare a finite number of simulated alternatives. There exist at least four classes of comparison problems that arise in simulation studies: selecting the system with the largest or smallest expected performance measure (selection of the best), comparing all alternatives against a standard (comparison with a standard), and selecting the system with the largest probability of actually being the best
Performer (multinomial selection), and selecting the system with the largest probability of success (Bernoulli selection). The ranking and selection approach is different. It asks given the data on the distributions of these K populations, what is the probability that we can correctly rank them from worst to best? What is the probability that we can choose the best population (perhaps the one with the largest population mean) or at least the best M out of the K populations? [9][5, 1].
Now, in many situations (or problems) we want to select the median value (alternative) from among the alternatives .then, How can we get something dose to the median, reasonably quickly? Just like the "quicker selection ", given an unsorted array, how quickly can one select the median element? Median finding is a special case of the more general selection problem which asks for the *mth* element in sorted order. The median provides a better measure of location than the mean when there are some extremely large or small observations (i.e., when the data are skewed to the right or the left) for this reason, median income is used as the measure of location for the U.S. hoursehold’s income and it is a special case of the more general selection problem which asks for the k th element in sorted order [6].

There are several works in the literature treating the exact median selection problem (cf. [BFPRT73], [DZ99], [FJ80], [FR75], [Hoa61], [HPM97]). Traditionally, the "comparison cost model" is adopted, where the only factor considered in the algorithm cost is the number of key-comparisons. The best upper bound on this cost found so far is nearly 3n comparisons in the worst case (cf. [DZ99]). The algorithm described here approximates the median with high precision and lends itself to an immediate implementation.
The usefulness of such an algorithm is evident for all applications where it is sufficient to find an approximate median for example in some heap sort variants or for median-filtering in image representation. In addition, the analysis of its precision is of independent interest. All the works mentioned above—as well as ours—assume the selection is from values stored in an array in main memory. The algorithm has an additional property which, as we found recently, has led to its being discovered before, albeit for solving a different problem [8].
Because, more the works they found approximate the median selection therefore suggest's the new approach is to slection the median by bayesian selection procedure among multi-populations (categorys) introduce in section 4 . We derive the stopping risks (Bayes risk) of making decision d_i for linear loss function given in section 5.

2- The Median

2.1 A brief of the Median

In probability theory and statistic, a ***median*** is described as the numeric value separating the higher half of a sample, a population, or a 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, so one often takes the mean of the two middle values. In a sample of data, or a finite population, there may be no member of the sample whose value is identical to the median (in the case of an even sample size) and, if there is such a member, there may be more than one so that the median may not uniquely identify a sample member. Nonetheless the value of the median is uniquely determined with the usual definition.
At most half the population has values less than the median and at most half have values greater than the median. If both groups contain less than half the population, then some of the population is exactly equal to the median. For example, if $a < b < c$, then the median of the list $\{a, b, c\}$ is b, and if $a < b < c < d$, then the median of the list $\{a, b, c, d\}$ is the mean of b and c, i.e. it is $(b + c)/2$[12].
The median can be used as a measure of location when a distribution is skewed , when end values are not known, or when one requires reduced importance to be attached to outliers , e.g. because they may be measurement errors. A disadvantage of the median is the difficulty of handling it theoretically
The median—calculated by determining the midpoint of rank-ordered cases can be used with ordinal, interval, or ratio measurements and no assumptions need be made about the shape of the distribution .The median has another attractive feature: it is a resistant measure. That means it is not much affected by changes in a few cases. Intuitively, this suggests that significant errors of observation in several cases will not greatly distort the results. Because it is a resistant measure, outliers have less influence on the median than on the mean. *For example,* notice that the observations 1, 4, 4,5,7,7,8,8,9 have the same
median (7) as the observations 1,4,4,5,7,7,8,8,542. The means (5.89 and 65.44, respectively), however, are quite different because of the outlier, 542 in the
Second set of observations [10].
Hakimi (1964), (1965) was the first to formulate the problem for locating
a single and multi-medians. He also proposed a simple enumeration procedure to solve the problem. The problem is well known to be NP-hard (Garey and Johnson 1979). Several heuristics have been developed for pmedian problems. Some of them are used to obtain good initial solutions or to

calculate intermediate solutions on search tree nodes. Teitz and Bart (1968) proposed simple interchange heuristics (see also (Maranzana 1964)). More complete approaches explore a search tree. They appeared in Efroymson and Ray (1966), Jarvinen and Rajala (1972), Neebe (1978), Christofides and Beasley (1982), Galvão and Raggi (1989) and Beasley (1993). The combined used of Lagrangean relaxation and subgradient optimization in a primal-dual viewpoint was found to be a good solution approach to the problem (Christofides and Beasley 1982), (Galvão and Raggi 1989), (Beasley 1993)[3].

2.2 Theoretical properties [2,12]

- **An optimality property**

A median is also a central point which minimizes the average of the absolute deviations: In the above example, a median would be (1 + 0 + 0 + 0 + 1 + 7) / 6 = 1.5 using the minimum of the absolute deviations; in contrast, the minimize of the sum of squares would be mean, which is 1.944. In the language of statistics, a value of *c* that minimizes

$$E(|X - c|)$$

Is a median of the probability distribution of the random variable X. However, a median *c* need not be uniquely defined. Where exactly one median exists, statisticians speak of "the median" correctly; even when no unique median exists, some statisticians speak of "the median" informally.

- **An inequality relating means and medians**

For continuous probability distributions, the difference between the median and the mean is less than or equal to one standard deviation

Median has different meanings in different contexts:

- *Median* in statistics, a number that separates the lowest-value half and the highest-value half
- *Median(geometry)*, in geometry, a line joining a vertex of a triangle to the midpoint of the opposite side
- *Geometric Median* , a point minimizing the sum of distances to a given set of points
- *Median road* , the portion of a divided highway used to separate opposing traffic
- *Median filter (image processing)* used to reduce noise in images
- an adjective related to the *Medes* , an Iranian people
- *the median language*, the language of the Medes
- the *median nerve*

- **Some Uses[4]**

1- Median _filtering is a commonly used technique in signal processing (is very useful in Vertical Seismic Profile (VSP) data processing and automatic editing of sulfate seismic data. It is also used in seismic processing to enhance linear events.)

2- Find the item which is smaller than half of the items and bigger than half the items.

3- Median crossovers are provided at selected locations on divided highways for crossing by maintenance crossing, traffic service, emergency, and law enforcement vehicles.

4- Median data are used to link adjacent roadways and are similar to template data in that they give a shape to be constructed. Again, a table of typical medians is created.

5- How median spaces are a natural generalization of CAT (0) cube complexes, how one characterizes property (T) and the Haagerup property using median spaces.

6- The median be thought of as geometric middle while the mean is arithmetic middle and the geometric nature of the median results in it not being influenced by a few large numbers at either extreme.

3- The Origin of the Problem

In many situations we need to select the median term such as the median high temperature for the week, the median high among values and variety of averages, When the number of observation is odd. Suppose we have the multinomial distribution with k cells and unknown probabilities of an observation in the ith cell p_i , the probability density function given as below:

$$P(\underline{n} \mid \underline{p}) = \frac{m'!}{n_1!n_2!...n_k!}\prod_{i=1}^{k} p_i^{n_i} \quad ,$$

$$\sum_{i=1}^{k} n_i = m' \text{ And } \sum_{i=1}^{k} p_i = 1 \quad , \qquad \underline{n} = (n_1,...,n_k)$$

(i =1, 2,.., m,.., k), where $\sum_{i=1}^{k} p_i = 1$. It is required to find the cell with the median (bigger than half) probability (best cell in this sense).

Let $p_{[1]} \le p_{[2]} \le .. \le p_{[m]} \le .. \le p_{[k]}$ denote the order values of the $p_i (1 \le i \le k)$, the goal of experimenter is to select the median cell probability that is the cell associated with $p_{[m]}$.

4- The Bayesian Median Procedure

Before we introduce the Bayesian procedures, we introduce some standards definitions and notations which are needed to construct the procedures. Let $\Omega_k : \{\underline{p} = (p_1, p_2,..., p_k) : \sum_{i=1}^{k} p_i = 1 \ ; p_i \ge 0\}$ 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 median cell probability (*i*=1, 2, m,.., *k*).

That is, d_i denote the decision to select the event associated with the i^{th} cell as the bigger than half 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.

$$L(d_i, \underline{p}^*) = \begin{cases} k^*(p_{[m]} - p_i) & \text{if } (p_{[m]} \ne p_i) \\ 0 & \text{if } (p_{[m]} = p_i) \end{cases} \qquad \text{……….. (4.1)}$$

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

The Bayesian approach requires that we specify a prior probability density function $\pi(\underline{p})$, expressing our beliefs about $\underline{p}$ before we obtain the data. From a mathematical point of view, it would be convenient if $\underline{p}$ is assigned a prior distribution which is a member of a family of distributions closed under multinomial sampling or as a member of the conjugate family. The conjugate family in this case is the family of Dirichlet distribution. Accordingly, let $\underline{p}$ is assigned Dirichlet prior distribution with parameters $m', n_1', n_2',..., n_k'$. The normalized density function is given by[11]

$$\pi(\underline{p}) = \frac{\Gamma\left(\sum_{i=1}^{k} n_i'\right)}{\prod_{i=1}^{k} \Gamma(n_i')} \prod_{i=1}^{k} p_i^{n_i'-1}, \text{ where } m'' = \sum_{i=1}^{k} n_i' \quad \ldots\ldots\ldots \quad (4.2)$$

and the marginal distribution for p_i is Beta density

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

Here $\underline{n}' = (n_1', n_2', ..., n_k')$, are regarded as hyper parameters specifying the prior distribution. They can be thought of "imaginary counts" from prior experience. If N_i be the number of times that category i is chosen in m independent trials, then $\underline{N} = (N_1, ..., N_k)$ has a multinomial distribution with probability mass function

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

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

Since

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

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

This is a member of the Dirichlet family with parameters

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

Hence, the posterior distribution has density function

$$\pi(\underline{p} \mid \underline{n}) = \frac{(m''-1)!}{(n_1''-1)!(n_2''-1)!....(n_k''-1)!} p_1^{n_1''-1} ... p_m^{n_m''-1} ... p_k^{n_k''-1} \quad \ldots\ldots\ldots \quad (4.3)$$

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

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

5- The Stopping Risks of the Median Procedure

Consider a multinomial distribution which is characterized by k events (cells) with probability vector $\underline{p} = (p_1, p_2, ..., p_k)$, where p_i is the probability of the event E_i $(1 \le i \le k)$ with $\sum_{i=1}^{k} p_i = 1$. Let $n_1, n_2, .., n_m, ..., n_k$ be respective frequencies in k cells of the distribution with $\sum_{i=1}^{k} n_i = m$. Further, let $p_{[1]} \le p_{[2]} \le ... \le p_{[m]} \le ...... \le p_{[k]}$ 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 bigger than half probable event, that, is the event associated with $p_{[m]}$, also called the median cell.

The stopping risk (the posterior expected loss) of the terminal decision d_i when the posterior distribution for $\underline{p}$ has parameters $(n''_1, n''_2, .., n''_m, ..., n''_k; m''')$, that is when the sample path has reached $(n''_1, n''_2, .., n''_m, ..., n''_k; m''')$ from the origin $(n'_1, n'_2, .., n'_m, .., n'_k; m'')$, denoted by $S_i(n''_1, n''_2, ., n''_m, .., n''_k; m''')$, can be found as follows.

$$S_i(n''_1, n''_2, .., n''_m, .., n''_k; m''') = \underset{\pi(\underline{p}|\underline{n})}{E}[L(d_i, \underline{p}^*)]$$

$$= k^*\left[\underset{\pi(\underline{p}|\underline{n})}{E}(þ_{[m]}) - \frac{n''_i}{m'''}\right] \qquad \text{...... (5.1)}$$

The value of $\underset{\pi(\underline{p}|\underline{n})}{E}[þ_{[m]}]$ is derived as follows.

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

So, if the number of observations is odd then

$$g(p_{[m]}) = k\binom{\text{k -1}}{\text{k} + 1/2}\left[F(p_{[m]})\right]^{k+1/2-1}\left[1 - F(p_{[m]})\right]^{k-k+1/2} f(p_{[m]})$$

$$= k\binom{k-1}{k+1/2}\frac{(m'''-1)!}{(n''_{[m]}-1)!(m'''-n''_{[m]}-1)!}\, p_{[m]}^{n''_{[m]}-1}(1-p_{[m]})^{m'''-n''_{[m]}-1}$$

$$.\left[\sum_{j=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j!(m'''-n''_{[m]}-j)!}.p_{[m]}{}^{j}(1-p_{[m]})^{m'''-1-j}\right]^{k-1/2}.\left[1-\sum_{j=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j!(m'''-1-j)!}.p_{[m]}{}^{j}(1-p_{[m]})^{m'''-1-j}\right]^{k-1/2} \qquad \text{... (5.2)}$$

is the probability density function of the median order statistics $p_{[m]}$ .where the marginal posterior probability density function of p_i if $p_i = p_{[m]}$ is

$$f(p_{[m]}) = \frac{(m'''-1)!}{(n''_{[m]}-1)!(m'''-n''_{[m]}-1)!}\, p_{[m]}^{n''_{[m]}-1}(1-p_{[m]})^{m'''-n''_{[m]}-1} \qquad \text{...... (5.3)}$$

and the cumulative density function is

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

Such that the ordered values of $n''_1, n''_2, ., n''_m, .., n''_k$ is $n''_{[1]} \le n''_{[2]} \le, .., \le n''_{[m]}, ., \le n''_{[k]}$.

Then,

$$\underset{\pi(\underline{p}|\underline{n})}{E}(þ_{[m]}) = \frac{k\left[(m'''-1)!\right]^k}{(n''_{[m]}-1)!(m'''-n''_{[m]}-1)!}\binom{k-1}{k+1/2}\left[\int_0^1 p_{[m]}{}^{n''_{[m]}}(1-p_{[m]})^{(m'''-n''_{[m]}-1)}\,.\right.$$

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

$$\sqrt{\frac{\sum_{J=0}^{k}\left\{\binom{k}{J}(-1)^{J}\left\langle (m'''-1)!(1-y)\right\rangle^{J}\left\langle \sum_{j_1=n''_{[m]}}^{m'''-1}..\sum_{j_J=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j_1!(m'''-n''_{[m]}-j_1)!}..\frac{(m'''-1)!}{j_J!(m'''-n''_{[m]}-j_J)!}.\left(\frac{p_{[m]}}{1-p_{[m]}}\right)^{j_1+..+j_J}\right\rangle\right\}}{1-\sum_{j=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j_k!(m'''-n''_{[m]}-j_k)!}p_{[m]}^{\ j}(1-p_{[m]})^{m'''-1-j}}}\ \Bigg]dp_{[m]}$$

$$=\frac{(m'''-1)!}{j_k!(m'''-n''_{[m]}-j_k)!}\sqrt{\frac{\left[\sum_{j_1=n''_{[m]}}^{m'''-1}...\sum_{j_k=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j_1!(m'''-n''_{[m]}-j_1)!}.....\frac{(m'''-1)!}{j_k!(m'''-n''_{[m]}-j_k)!}.\right]\left[\sum_{J=0}^{m'''-1}\binom{k}{J}(-1)^{J}((m'''-1)!)^{J}\left\langle\sum_{j_1=n''_{[m]}}^{m'''-1}..\sum_{j_J=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j_1!(m'''-n''_{[m]}-j_1)!}..\frac{(m'''-1)!}{j_J!(m'''-n''_{[m]}-j_J)!}\right\rangle\right]}{\left[\sum_{j=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j!(m'''-n''_{[m]}-j)!}\right]\left[1-\sum_{j=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j!(m'''-n''_{[m]}-j)!}\right]}}.$$

$$\int_0^1 p_{[m]}^{\ n''_{[m]}}(1-p_{[m]})^{m'''-1-n''_{[m]}}\sqrt{\left\langle\frac{\left(p_{[m]}^{\ j}(1-p_{[m]})^{m'''-1-j}\right)^{j_1+j_2+...+j_k}}{p_{[m]}^{\ j}(1-p_{[m]})^{m'''-1-j}}\right\rangle\left\langle\frac{\left[(1-p_{[m]})^{J}p_{[m]}^{\ j}(1-p_{[m]})^{m'''-1-j}\right]^{j_1+j_2+...+j_J}}{p_{[m]}^{\ j}(1-p_{[m]})^{m'''-1-j}}\right\rangle}dp_{[m]}$$

Hence $S_i(n_1,n_2,...,n_k;m)$

$$=\left\{\frac{(m'''-1)!}{j_k!(m'''-n''_{[m]}-j_k)!}\sqrt{\frac{\left[\sum_{j_1=n''_{[m]}}^{m'''-1}...\sum_{j_k=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j_1!(m'''-n''_{[m]}-j_1)!}.....\frac{(m'''-1)!}{j_k!(m'''-n''_{[m]}-j_k)!}.\right]\left[\sum_{J=0}^{m'''-1}\binom{k}{J}(-1)^{J}((m'''-1)!)^{J}\left\langle\sum_{j_1=n''_{[m]}}^{m'''-1}..\sum_{j_J=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j_1!(m'''-n''_{[m]}-j_1)!}..\frac{(m'''-1)!}{j_J!(m'''-n''_{[m]}-j_J)!}\right\rangle\right]}{\left[\sum_{j=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j!(m'''-n''_{[m]}-j)!}\right]\left[1-\sum_{j=n''_{[m]}}^{m'''-1}\frac{(m'''-1)!}{j!(m'''-n''_{[m]}-j)!}\right]}}.\right.$$

$$\left.\frac{(n''_{[m]}!)\left[(m'''-j-1)+(j_1-j_2-...-j_k-1)+J\big/(m'''-j-1)+(j_1-j_2-...-j_J-1)\big/2\right]!}{\left\langle n''_{[m]}+1+\left[(m'''-j-1)+(j_1-j_2-...-j_k-1)+J\big/(m'''-j-1)+(j_1-j_2-...-j_J-1)\big/2\right]\right\rangle!}\right\}-\frac{n''_i}{m'''}\ ....(5.4)$$

6 - **Future Work**

- Our plan in future is to produce some numerical results for this procedure and simulation .
- Fully and group Bayesian sequential scheme to selecting the median multinomial selection problem can be developed .
- An upper bound for risks may be found using functional analysis.
- General loss functions may be used, where linear loss is considered as a special case.
- To simplify the formula (5.4) we can use sterling's approximation for large factorials and hence we will get an approximate formula to (5.4) .

7- **References**

[1]. Bechhofer, R. E. and Kulkarni, R.V. (1984). Closed sequential which have the largest probabilities. Communications in statistics-Theory and Methods. A13 (24). pp. 2997- 3031.

[2]. D.Dor, and U. Zwick. Selecting the Median. SIAM Jour. Comp., 28(5):1722–1758, 1999.

[3]. Edson L. F. Sane, lagrangean/surrogate heuristics for p-median problems, Chapter 6, 2000.

[4]. Johnny Wolfe Jay Measures of Central Tendency: Mean, Median, and Mode Examples, August 23, 2001, High School Santa Rosa County Florida .

[5]. Jean Dickinson and M. Sobel, selecting and ordering populations: a new statistical methodology, society for industrial and applied mathematics, 1999.

[6]. Lari Arjomand, Descriptive Statistics 1996

[7]. Roger D. Congleton, The Median Voter Model, Center for Study of Public Choice, George Mason University 2-22-02.

[8]. S. Battiato, D. C, C and G. Cincotti , An Efficient Algorithm for the Approximate Median Selection Problem, Universit`a di Catania, Department di Matematica,Viale A. Doria 6, I-95125 Catania, Italy, Department Of Computer Science , 01609-2280 ,October 28, 2002.

[9]. Seong- Hee, Kim Georgia and Barry L. Nelson, Selecting the Best System , Institute of Technology Northwestern University Proceedings of the 2004, Winter Simulation Conference.

[10]. Werner Grosshans and Eleanor Chelimsky , **Quantitative** Data Analysis: An Introduction, Report To Program Evaluation and Methodology Division, 1992.

[11]. Order statistic
http;//psychology.wikia.com/wiki/order_statistic.

[12]. Median
http;//cn.wikipedia.org/wiki/median