View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Clute Institute: Journals

A Procedure For Selecting A Best Multinomial Distribution With Application To Population Income Mobility

Saad T. Bakir, Alabama State University, USA

ABSTRACT

A procedure is developed for selecting a subset which is asserted to contain the "best" of several multinomial populations with a pre-assigned probability of correct selection. According to a pre-chosen linear combination of the multinomial cell probabilities, the "best" population is defined to be the one with the highest such linear combination. As an illustration, the proposed procedure is applied to data relating to the economics of happiness and population income mobility.

Keywords: Economics of happiness; ranking and selection procedures; probability of a correct selection

1. INTRODUCTION

Several procedures have been proposed in the literature for the purpose of selecting the best of several populations. The "best" population is usually defined as the one with the highest parameter of interest such as the location or scale parameter. One of the first selection procedures in the literature is that by Mosteller (1948) who gave a nonparametric test for the null hypothesis of homogeneity (identical populations) against the "slippage" alternative hypothesis (one population has higher location parameter). Paulson (1949) devised a rule for classifying several normal populations into a "superior", and an "inferior" group according to the value of their means and Paulson (1994) gave an eliminating procedure for selecting the best one of several Koopman-Darmois distributions. Bechhofer (1954), Bechhofer and Sobel (1954), Bechhofer (1958), Bechhofer and Blumenthal (1962), Swanepoel and Geertsema (1976), and Turnbull et al. (1978) developed sequential and adaptive sequential procedures for selecting the best of several normal populations. Guttman (1963) proposed a sequential procedure where at each stage one retains fewer populations until a single population is left being the best. Along these lines, Bauer (1989) proposed a multiple testing sequential procedure for eliminating the inferior ones of several populations. Based on the Hodges-Lehann estimators, Swanepoel and Geertsema (1973) developed nonparametric sequential procedures for selecting the best of several populations. Non-sequential procedures were introduced by Gupta and Sobel (1958) for selecting a subset which is asserted to contain the best population with a pre-assigned probability of correct selection. A Baysian approach to the best population problem is adopted by Guttman and Tiao (1964), with special attention to exponential and normal populations. Studden (1967) discussed the selection problem in terms of decision functions and characterizes optimal selection subset rules. Recently, Hayter (2007) developed a combined multiple comparisons and subset selection procedure.

Regarding discrete distributions, Paulson (1952) gave a procedure for determining the best among binomial populations using the inverse sine transformation and Paulson (1967) proposed a sequential procedure for selecting the binomial population with the highest probability of successes. Hoel and Milton (1972) made a comparative study of sequential procedures for selecting the best binomial population. Taheri and Young (1974) investigated two sequential sampling plans (play-the-winner sampling and vector-at-a-time sampling) for selecting the better of two binomial populations; they showed that play-the-winner sampling is uniformly better. Levin and Leu (2007) compared two procedures (for selecting the best binomial population) based on sequential elimination of inferior populations. Bechhofer and Kulkarni (1982) proposed a class of sequential procedures for selecting the best of several Bernoulli populations. Regarding multinomial distributions, Bose and Bhandari (2001), Panchapakesan (2006), and Ng and Panchapakesan (2007) discussed selection procedures for the most probable multinomial event of a single multinomial population. Reference textbooks on ranking and selection procedures include: Balakrishnan et al. (2004), Gupta and Panchapakesan (2002), and Gibbons et al. (1987).

In this paper we propose a procedure for selecting a subset which is asserted to contain the "best" of several multinomial populations with a pre-assigned probability of correct selection, *P*. The proposed procedure assumes that there is a linear combination of the multinomial cell probabilities according to which the experimenter desires to order the populations. The "best" multinomial population is defined as the one with the highest such linear combination. Our proposed procedure may be considered as an extension of Gupta and Sobel (1960) procedure for selecting a subset containing the best of several binomial populations. For binomial populations, the "best" population is usually defined in the literature as the one with the highest probability of success. We will use the tables in Gupta and Sobel (1960) to implement our proposed procedure in practice.

The rest of the paper is arranged as follows: Section 2 contains the development of the proposed procedure for selecting a best multinomial population, Section 3 contains the derivation of the probability of correct selection, and Section 4 contains an illustrative numerical example using data pertaining to the economics of happiness data.

2. A PROCEDURE FOR SELECTING A SUBSET CONTAINING THE BEST MULTINOMIAL POPULATION

In this section we develop the proposed procedure for selecting a subset which is asserted to contain the "best" of several multinomial populations with a probability greater than or equal to a pre-assigned value *P*. According to a pre-chosen linear combination of the multinomial cell probabilities, the "best" multinomial population is defined to be the one producing the highest such linear combination. The proposed procedure may be considered as a multinomial extension of Gupta and Sobel's (1960) procedure for selecting the best of several binomial populations, where they define the best binomial population as the one with the highest probability of success.

Suppose we have *m* multinomial populations of *k* classes each, with unknown cell probabilities $\left(\pi_{i1}, \pi_{i2}, \cdots \pi_{ik}\right)$ and probability mass functions (pmf) $f_i(.), 1 \le i \le m,$ given by :

$$f_i\left(n_{i1}, n_{i2}, \cdots, n_{ik}\right) = \frac{n_i!}{\prod_{j=1}^{k} n_{ij}!} \prod_{j=1}^{k} \pi_{ij}^{n_{ij}} \text{; where the random variables } n_{ij} = 0, 1, \cdots, n_i \text{ are}$$

$$\text{subject to } \sum_{j=1}^{k} n_{ij} = n_i \text{, fixed, and the constants } 0 \le \pi_{ij} \le 1 \text{ are such that } \sum_{j=1}^{k} \pi_{ij} = 1. \qquad (1)$$

For arbitrary constants a_1 and a_2 such that $a_1 > a_2$, define the linear combinations

$$h_i = a_1\left(\pi_{i1} + \cdots + \pi_{ir}\right) + a_2\left(\pi_{i,r+1} \cdots + \pi_{ik}\right), \qquad (2)$$

where $i = 1, 2, \cdots, m$ and r is a positive integer such that $1 \le r < k$. The ordered $h_i's$ are denoted by $h_{[1]} \le h_{[2]} \le \cdots \le h_{[m]}$. It is assumed that the correct pairing of the $h_i's$ with the m populations is not known. The "best" multinomial population is defined to be the one associated with $h_{[m]}$, the highest linear combination. Note that there is no loss of generality in assuming that $a_1 > a_2$, since one can renumber the classes of each population so that the first r classes are always associated with the larger a_i $\left(i = 1, 2\right)$ which we call a_1. An instance of the application of this definition of best multinomial is where one is interested in comparing the m populations with respect to a subset, say the first r, of the k classes. Then one chooses a_2 small enough in absolute value, may be zero, so that the contribution due to the remaining *(k-r)* classes to the linear combination *h* is negligible. For example, if the first class is the only class of interest and the population with the highest probability in this class is considered to be the best, then the choice $a_1 = 1$, $a_2 = 0$ and $r = 1$ gives $h_i = \pi_{i1}\left(i = 1, \cdots, m\right)$.

Suppose that independent random samples $\left(n_{i1}, \cdots, n_{ik}\right)$ of sizes $n_i = \sum_{j=1}^{k} n_{ij}\ \left(i = 1, \cdots, m\right)$ are drawn from the multinomial populations under study. Let the unbiased estimators of the h_i be

$$v_i = a_1\left(p_{i1} + \cdots p_{ir}\right) + a_2\left(p_{i,r+1} + \cdots p_{ik}\right), \tag{3}$$

where $p_{ij} = \dfrac{n_{ij}}{n_i}\ \left(i = 1, \cdots, m;\ j = 1, \cdots, k\right)$. Let $v_{\max} = \max\left(v_1, \cdots, v_m\right)$.

Then the proposed procedure is to retain the ith population in best subset if and only if:

$$v_i \geq v_{\max} - c\ , \tag{4}$$

where c is a non-negative constant depending on $m, n_i\ \left(i = 1, \cdots, m\right)$ and P, as will be determined in Section 3.

3. THE PROBABILITY OF CORRECT SELECTION

We say that a correct selection (CS) is made if and only if the retained subset contains the best population. It is required that $\Pr\left(CS \middle| \pi_{ij}\right) \geq P$ for all possible configurations of the true parameters $\pi_{ij}\ \left(i = 1, \cdots, m;\ j = 1, \cdots, k\right)$. The constant c in Eq.4 is chosen to be the smallest nonnegative number such that the infimum of $\Pr\left(CS\right)$ taken over all n_i and $\pi_{ij}\ \left(i = 1, \cdots, m;\ j = 1, \cdots, k\right)$, is greater than or equal to *P*. In order to find $\Pr\left(CS\right)$, we adopt the convention that when there is more than one population associated with $h_{(m)}$ (i.e., more than one best population) we consider one particular "tagged" population as being the best.

To determine the probability of a correct selection, we write the estimator v_i in Eq. 3 as:

$$\begin{aligned} v_i &= a_1\left(p_{i1} + \cdots + p_{ir}\right) + a_2\left(p_{i,r+1} + \cdots + p_{ik}\right) \\ &\frac{1}{n_i}\left[a_1\left(n_{i1} + \cdots + n_{ir}\right) + a_2\left(n_{i,r+1} + \cdots + n_{ik}\right)\right] \\ &= \frac{1}{n_i}\left[(a_1 - a_2)\left(n_{i1} + \cdots + n_{ir}\right) + a_2 n_i\right],\ i = 1, \cdots, m \end{aligned} \tag{5}$$

Define

$$u_i = \frac{n_i\left(v_i - a_2\right)}{a_1 - a_2} = \left(n_{i1} + \cdots + n_{ir}\right). \tag{6}$$

Then it is known that u_i has the binomial distribution $B\left(n_i;Q_{i,r}\right)$, where

$$Q_{i,r}=\pi_{i1}+\cdots\pi_{ir}\equiv\frac{h_i-a_2}{a_1-a_2},\ i=1,\cdots,m. \tag{7}$$

For simplicity we denote $Q_{i,r}$ by Q_i.

Let $Q_{(i)}$, $n_{(i)}, v_{(i)}$, and $u_{(i)}$ be those particular quantities Q_i, n_i, v_i and u_i, respectively, which are associated with the population corresponding to $h_{[i]}\ i=1,\cdots,m$. Note that the subscript notation $(\cdot)$ does not indicate ordered quantities. Then, using the procedure in Eq. 4, a correct selection is made if and only if $v_{(m)}\geq v_{\max}-c$ or $v_{\max}\leq v_{(m)}+c$ which is equivalent to
$v_{(i)}\leq v_{(m)}+c$ for all $i<m$ (for $i=m$ the inequality is satisfied with probability 1). The last inequality can be written as:

$$\frac{n_{(i)}\left(v_{(i)}-a_2\right)}{a_1-a_2}\leq\frac{n_{(i)}\left(v_{(m)}-a_2\right)}{a_1-a_2}+\frac{n_{(i)}c}{a_1-a_2}$$

or

$$u_{(i)}\leq n_{(i)}\left(\frac{u_{(m)}}{n_{(m)}}+\frac{c}{a_1-a_2}\right)\ \text{for all}\ i<m \tag{8}$$

Since the $u_{(i)}$'s are independent binomial $B(n_i;Q_i)$ variables, it can be seen that the probability that (8) holds true, i.e. $\Pr(CS|\pi_{ij})\geq P$ is equal to

$$\sum_{u=0}^{n_{(m)}}\binom{n_{(m)}}{u}Q_{(m)}^{u}\left(1-Q_{(m)}\right)^{n_{(m)}-u}\prod_{i=1}^{m-1}\left\{\sum_{x=0}^{\left[n_{(i)}\left(\frac{u}{n_{(m)}}+\frac{c}{a_1-a_2}\right)\right]}\binom{n_{(i)}}{x}Q_{(i)}^{x}\left(1-Q_{(i)}\right)^{n_{(i)}-x}\right\} \tag{9}$$

Here, $[z]$ denotes the largest integer less than or equal to z.

The problem of interest, now, is to minimize (9).

Each of the (m – 1) factors in the braces appearing in (9) is a non-increasing function of $Q_{(i)}$ as can be seen by expressing each factor as an incomplete beta function. Recalling that $Q_{(i)}=\dfrac{h_{(i)}-a_2}{a_1-a_2}$ together with the assumption that $a_1 - a_2 > 0$, we see that the ranking $h_{[1]} \le h_{[2]} \le \cdots \le h_{[m]}$ is equivalent to the ranking $Q_{[1]} \le Q_{[2]} \le \cdots \le Q_{[m]}$. Therefore, for a fixed $Q_{[m]}$, each factor in the braces is minimized by taking $Q_{(i)} = Q_{[m]}$, $i = 1, 2, \cdots, m$. Then we consider the infimum of (9) over the range of $Q_{[m]} = Q$, say, which is $0 \le Q \le 1$. To achieve the absolute minimum of $\Pr(CS)$, we must further minimize (9) with respect to $n_{(m)}$ which is an element of the set $\{n_1, n_2, \cdots, n_m\}$. Because all Q_i are taken equal to Q, any manner of pairing the other (m – 1) $n_{(i)}$'s with the remaining n_i's (after selecting a value for $n_{(m)}$) will give exactly the same minimum value for the product of the (m – 1) factors. The condition that the infimum of $\Pr(CS)$ is greater or equal to P, is then

$$\min_{n_{(m)} \in \{n_1, \cdots, n_m\}} \left\{ \inf_{0 \le Q \le 1} \sum_{u=0}^{n_{(m)}} \binom{n_{(m)}}{u} Q^u (1-Q)^{n_{(m)}-u} \prod_{s \ne (m)} \sum_{x=0}^{\left[n_s \left(\frac{u}{n_{(m)}} + \frac{c}{a_1 - a_2} \right) \right]} \binom{n_s}{x} Q^x (1-Q)^{n_s - x} \right\} \ge P \tag{10}$$

For equal sample sizes $n = n_i \ (i = 1, \cdots, m)$, Eq. (10) becomes

$$\inf_{0 \le Q \le 1} \left\{ \sum_{u=0}^{n} \binom{n}{u} Q^u (1-Q)^{n-u} \left[\sum_{x=0}^{u + \frac{nc}{a_1 - a_2}} \binom{n}{x} Q^x (1-Q)^{n-x} \right]^{m-1} \right\} \ge P \tag{11}$$

The constant $\frac{nc}{a_1 - a_2} = d$ is taken to be the smallest non-negative integer such that (11) is satisfied. The values of d satisfying (11) have been tabulated by Gupta and Sobel (1960, pp. 242-45) to carry out their procedure for selecting a subset containing the best binomial population, where they define the best binomial population as the one with the highest probability of success. The tables give the values of d for P = 0.75, .90, .95, 0.99; n = 1 (1) 20, 25 (5) 50, 50 (10) 100, 100 (25) 200, and 200 (50) 500; m =1 (1) 20, 25 (5) 50. After obtaining $d = \frac{nc}{a_1 - a_2}$ from the tables, we solve for c which enables us to carry out the procedure in Eq. (4), that is to retain only those populations for which $v_i \ge v_{\max} - c$. Better still, the procedure in Eq. (4) in the case of equal sample sizes can be put in the form: Retain the ith population if and only if $u_i \ge u_{\max} - \frac{nc}{a_1 - a_2}$, where $u_i = \frac{n(v_i - a_2)}{a_1 - a_2} = (n_{i1} + \cdots + n_{ir})$ and $u_{\max} = \max(u_1, \cdots, u_m)$. This last form of the proposed procedure is more convenient for computations than the first, because it is easier to compute the u_i rather than the $v_i \ (i = 1, \cdots, m)$.

In the case of unequal sample sizes, there is no general rule as to which particular value of $n_{(m)}$ minimizes the left hand side of (10), above. Gupta and Sobel (1960, pp. 230-231) empirically found that for some interval of the constant-value, $[0, d_0]$, the left hand side of (10) is minimized by taking $n_{(m)}$ to be the largest of the

$n_i \left(i = 1, \cdots, m \right)$. But this is not true when d is larger than d_0. In practical applications, they suggested to take the arithmetic mean $\bar{n}$ of the sample sizes as the common sample size and to use the appropriate table with $n = \bar{n}$ to obtain the value of the constant; this value may be improved by further computations depending on the specific situation. Since the present case of multinomial populations is similar to that of binomial populations, we follow this approach of taking $n = \bar{n}$ to deal with situations in which the sample sizes are not equal.

4. APPLICATION TO POPULATION INCOME MOBILITY

We illustrate our proposed procedure with the economics of happiness data in Graham and Pettinato (2002) that shows the income mobility up and down the economic ladder of individuals in each of four countries: Peru, USA, Russia, and S. Africa. The economic ladder is divided into five parts (quintiles: Q1, Q2, Q3, Q4, and Q5). Table 1 displays the selected data from Graham and Pettinato (2002, Tables 3-2 through 3-5) that pertains to the income mobility of individuals starting in quintile 1 (at or below poverty level) in each of the four countries. For the data on S. Africa, Graham and Pettinato (2002, Table 3.5), we lumped the data for individuals below 0.5 PL and 1PL into quintile 1, where PL stands for poverty level.

Table 1. Income Mobility Data of Individuals in Quintile 1 in Four Countries

	Q1	Q2	Q3	Q4	Q5	Total
Peru Q1	45	25	19	6	5	100
USA Q1	61	24	9	5	1	100
Russia Q1	39	26	16	10	9	100
S. Africa Q1	66	8	7	17	3	100

Table 1 represents samples of equal size (*n =100*) selected from four (*m = 4*) multinomial populations made up of five (*k = 5*) classes each. In our example, the "best" population is the one that has the highest rate of income mobility of individuals from Q1 (at or below poverty level) to a higher quintile (above poverty level). The best population would then be the one depicting highest cell probabilities in Q2 through Q5. In Table 2, we rearrange Table 1 so that Q2 through Q5 become the first four classes and Q1 becomes the fifth class.

The best multinomial population is the one with the highest linear combination, h_i, given by

$$h_i = a_1\left(\pi_{i1} + \pi_{i2} + \pi_{i3} + \pi_{i4}\right) + a_2\left(\pi_{i4}\right).$$

Choosing $a_1 = 1$ and $a_2 = 0$, we get

$$h_i = \left(\pi_{i1} + \pi_{i2} + \pi_{i3} + \pi_{i4}\right). \tag{12}$$

Let $u_i = n_{i1} + n_{i2} + n_{i3} + n_{i4}$, then the proposed procedure in Section 3, retains the ith population if and only if

$$u_i \geq u_{\max} - d. \tag{13}$$

The values of the u_i and $u_{\max}$ are shown in Table 2.

Table 2. Rearrangement of Table 1

	C1(Q2)	C2(Q3)	C3(Q4)	C4(Q5)	C5(Q1)	Total	u_i
Peru Q1	25	19	6	5	45	100	55
USA Q1	24	9	5	1	61	100	39
Russia Q1	26	16	10	9	39	100	$61 = u_{\max}$
S. Africa Q1	8	7	17	3	66	100	34

Referring to Gupta and Sobel (1960, Table 2) with number of populations = *4*, sample size =*100* and various probabilities (*P*) of correct selection, we find the following values for the constant *d*:

$$d(P=0.75)=9;\ d(P=0.90)=12;\ d(P=0.95)=15;\ d(P=0.99)=19.$$

The decisions are displayed in Table 3.

Table 3. Subset Containing Best Multinomial Population

P	***d***	$u_{\max}-d$	**Subset with Best population**
0.75	9	52	Russia and Peru
0.90	12	49	Russia and Peru
0.95	15	46	Russia and Peru
0.99	19	42	Russia and Peru

In conclusion, Russia and Peru are the countries with the highest probability of income mobility from poverty to a higher economic level.

REFERENCES

1. Balakrishnan, N., Kannan, N., and Nagaraja (2004). Advances in Ranking and Selection, Multiple Comparisons, and Reliability: Methodology and Applications, Editors. Birkhaase. Boston
2. Bauer, P. (1989). A sequential elimination procedure for choosing the best population(s) based on multiple testing. *J. Statist. Plann. Inference* 21(2), 245-252.
3. Bechhofer, R. E. (1954). A single –sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Statist*. 25, 16-39.
4. _______________(1958). A sequential multiple-decision procedure for selecting the best one of several normal populations with a common unknown variance, and its use with various experimental designs. *Biometrics*, 14, 408-429.
5. Bechhofer, R. E. and Blumenthal, S. (1962). A Sequential multiple-decision procedure for selecting the best one of several normal populations with common unknown variance, II: Monte Carlo sampling results and new computing formulae. *Biometrics*, 18, 52-67.
6. Bechhofer, R. E. and Sobel, M. (1954). A Single –sample multiple decision procedure for ranking variances of normal populations. *Ann. Math. Statist*.
7. Bechhofer and Kulkarni (1982). Closed adaptive sequential procedures for selecting the best of k Bernoulli populations. *Statistical decision theory and related topics, III,* Vol. 1, 61-108. Academic Press, New York.
8. Bose, M. and Bhandari, S. K. (2001). Multinomial Subset selection using Inverse Sampling and its Efficiency with Respect to Fixed Sampling. *Sequential Analysis,* 20(3), 165-181.
9. Gibbons, J. D., Olkin, I., Sobel, M. (1987). Selecting and Ordering Populations: A New Statistical Methodology. SIAM.
10. Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Institute of Statistics, Mimeograph Series Report No. 150. Chapel Hill, N. C. University of North Carolina, May 1956.
11. Gupta, S. S., and Panchapakesan, S. (2002). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations. SIAM.
12. Gupta, S. S. and Sobel, M. (1958). On Selecting a subset which contains all populations better than a standard. *Ann. Math. Statist*. 29, 235-244.
13. Gupta, S. S. and Sobel, M. (1960). Selecting a subset containing the best of several binomial populations. Contributions to probability and statistics: Essays in honor of Harold Hotelling, ed. By Ingram Olkin and others (Stanford, California: Stanford University Press. Pp 224.
14. Guttman, I. (1963). A sequential procedure for the best population. Sankhya, Ser. A, 25, 25-28.
15. Guttman, I. and Tiao, G. C. (1964). A Baysian Approach to some best population problems. *Ann. Mat. Statist*., 35, 825-835.

16. Hater, A. J. (2007). A Combination Multiple Comparisons and Subset Selection Procedure to identify Treatments that Are Strictly Inferior to the Best. *J. Statistical Planning and Inference*. 137, 2115-2126.
17. Hoel, D. and Sobel, M. (1972). Comparisons of sequential procedures for selecting the best binomial population. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. IV: *Biology and Health*, pp. 53-69. Univ. of California Press, Berkeley, Calif., 1972.
18. Levin, B and Leu, C-S (2007) A comparison of two procedures to select the best binomial population with sequential elimination of inferior populations. *J. Statist. Plann. Inference*, 137(1), 245-263.
19. Mosteller, F. (1948). A k-sample slippage test for an extreme population. *Ann. Math. Statist*. 19, 58-65.
20. Ng, H. K. T. and Panchapakesan, S. (2007). Is the Selected Multinomial Cell the Best? *Sequential Analysis*, 26, 415-423.
21. Panchapakesan, S. (2006). A Note on a Subset Selection Procedure for the Most Probable Multinomial Event. *Sequential Analysis*. 25, 145-149.
22. Paulson, E. (1949). A multiple decision procedure for certain problems in analysis of variance. *Ann. Math. Statist.* 20, 95-98.
23. Paulson, E. (1952). On the Comparison of Several Experimental Categories with a control. *Ann. Math. Statist*. 23, 239-246.
24. Paulson, E. (1967). Sequential procedures for selecting the best one of several binomial populations. *Ann. Math. Statist.* 38, 117-123.
25. Paulson, E. (1994). Sequential procedures for selecting the best one of K Koopman-Darmois populations. *Sequential Analysis* 13(3) 207-220.
26. Studden, W. J. (1967). On selecting a subset of k populations containing the best. *Ann. Math. Statist*. 38, 1072-1078.
27. Swanepoel, J. W. H. and Geertsema, J. C. (1973). Nonparametric sequential procedures for selecting the best of k populations. *South African Statist. J.* 7, 85-94.
28. Swanepoel, J. W. H. and Geertsema, J. C. (1976). Sequential procedures with elimination for selecting the best of k normal populations. *South African Statist. J*. 10(1), 9-36.
29. Taheri, H. and Young, D. H. (1974). A comparison of sequential sampling procedures for selecting the better of two binomial populations. *Biometrika* 61(3), 585-592.
30. Turnbull, B., Kaspi, H., Smith, R. L. (1978). Adaptive sequential procedures for selecting the best of several normal populations. *J. Statist. Comput. And Simulation* 7(2), 133-150.