RANKING AND SELECTION TUTORIAL: 2-FACTOR NORMAL MEANS PROCEDURES

David Goldman
School of ISyE
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

We discuss sampling procedures for selecting that one of a number of normal populations (with common known variance) which has the largest mean. We present procedures appropriate for single-factor experiments, and then give procedures devised for 2-factor experiments without interaction between the factor-levels. In all cases, the procedures guarantee a prespecified probability of selecting the correct population.

1. INTRODUCTION

This tutorial introduces the reader to ranking and selection, a field that has enjoyed a great deal of popularity in the statistical literature over the last thirty years. We concentrate specifically on the problem of selecting that one of a number of normal populations (with common known variance) which has the largest mean.

The paper is organized as follows. In Section 2, we present some basic notation and formally state the normal means problem. Some procedures for use in single-factor experiments for finding the normal population with the largest mean are outlined in Section 3. Section 4 gives relevant notation and formally states the normal means problem for the 2-factor case. Section 5 generalizes the procedures of Section 3 for use in 2-factor experiments without interaction.

2. THE SINGLE-FACTOR NORMAL MEANS PROBLEM

Suppose $\Pi_1, \ldots, \Pi_k$ are $k$ independent normal populations with unknown means $\mu_i$ and common known variance $\sigma^2$, $i = 1, \ldots, k$. Without loss of generality, we henceforth assume that $\sigma^2 = 1$. Denote the ordered but unknown $\mu_i$’s by $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$. Of course, we do not know the values of the $\mu_i$’s nor how they are paired with the $\Pi_i$’s. Our goal is to determine which of the $\Pi_i$’s has the largest mean $\mu_k$; we call that $\Pi_i$ the "best" population.

A typical procedure for choosing the best population usually requires the experimenter to take a certain number of observations in a prescribed manner from each $\Pi_i$; the selection is then made using statistics calculated from these observations. Since the observations are realizations of random variables, it is possible that the experimenter will not choose the best $\Pi_i$. However, if the best population is indeed chosen, we say that a correct selection (CS) has been made.

We shall limit consideration to procedures that guarantee the following probability requirement on the P[CS] [cf. Bechhofer (1954)].

$$P[CS] \geq P^* \quad \text{whenever} \quad \mu_k - \mu_{k-1} \geq \delta^* \quad (PR)$$

$\delta^*$ and $P^*$ are two constants which must be specified prior to the start of experimentation subject to the restrictions that $\delta^* > 0$ and $1/k < P^* < 1$. $\delta^*$ is chosen as the smallest value of $\mu_k - \mu_{k-1}$ that the experimenter regards as "worth detecting." That is, if $\mu_k - \mu_{k-1}$ were to be $\geq \delta^*$, the experimenter would prefer to
choose the \( \pi_i \) associated with \( \mu_i[k] \); if \( \mu_i[k] < \mu_i[k-1] \) were to be \( < \delta^* \), the experimenter would, for all practical purposes, be indifferent about choosing either the population corresponding to \( \mu_i[k] \) or that corresponding to \( \mu_i[k-1] \). Hence, \[ N_{k+1} = \{ g | \mu_i[k] - \mu_i[k-1] \geq \delta^* \} \] is called the preference-zone, and its complement is the indifference-zone. Procedures which guarantee (PR) are often referred to as indifference-zone procedures.

3. SINGLE-FACTOR INDIFFERENCE-ZONE PROCEDURES

In this section, which closely follows Goldman (1966), we give a number of indifference-zone procedures for the normal means problem discussed in Section 2.

3.1 A Single-Stage Procedure

A single-stage procedure is one for which the size of the sample to be taken from each \( \pi_i \) is a constant determined prior to the start of experimentation. We first examine a single-stage procedure due to Bechhoefer (1954).

Procedure \( P_{SS} \):

1. For given \( k \) and known \( \sigma^2 = 1 \), specify \( \{ \delta^*, P^* \} \), with \( \delta^* > 0 \) and \( 1/k < P^* < 1 \).
2. Determine \( n_{SS} \), the smallest integer constant satisfying \[ \int_{-\infty}^{0} \Phi(x + \delta^* n_{SS})^{k-1} dx \geq P^* , \]
   where \( \Phi(\cdot) \) is the standard normal c.d.f. \[ n_{SS} \] can also be calculated easily from the tables in, e.g., Bechhoefer (1954).
3. Take \( n_{SS} \) independent observations from each of the \( k \) populations.
4. Let \( x_{ij} \) denote the \( j \)th observation from \( \pi_i \), \( i = 1, \ldots, k \) and \( j = 1, \ldots, n_{SS} \). Calculate \[ A_1 = \sum_{j=1}^{n_{SS}} x_{ij}, \ i = 1, \ldots, k. \]
5. Select as best that population corresponding to the largest \( A_1 \).

Example 1: Suppose that \( k = 3 \), \( \sigma^2 = 1 \), \( P^* = 0.75 \), and \( \delta^* = 0.2 \). From the tables in Bechhoefer (1954), we can calculate that \( n_{SS} = 52 \). So we must take 52 independent observations from each of the three populations.

Although a single-stage procedure such as \( P_{SS} \) is very easy to implement, it has the drawback of being conservative; that is, it cannot react to "favorable" configurations of \( \mu \) by reducing the number of observations to be taken. Sequential procedures, to be discussed next, can react to such conditions.

3.2 An Open Sequential Procedure

A sequential procedure is one which can involve more than a single stage of sampling. Such a procedure is open if, prior to the start of sampling, we can not place an upper bound on the number of observations to be taken from each \( \pi_i \). Otherwise, the procedure is closed. The next procedure is an open sequential procedure from Bechhoefer, Kiefer, and Sobel (1968).

Procedure \( P_{BKS} \):

1. For given \( k \) and known \( \sigma^2 = 1 \), specify \( \{ \delta^*, P^* \} \), with \( \delta^* > 0 \) and \( 1/k < P^* < 1 \).
2. Set the stage counter \( n = 0 \).
3. Set \( n = n + 1 \). Take one observation \( x_{i1} \) from each \( \pi_i \), \( i = 1, \ldots, k \). [Thus, we actually take observations a vector-at-a-time.]
4. For all \( i \), calculate \[ A_{i1} = \sum_{j=1}^{n_{SS}} x_{ij}. \]
   Further, let \( A_{\max} = \max_{1 \leq i \leq k} A_{i1} \).
5. Calculate \[ Z_n = \sum_{i=1}^{k} \exp\{ -\delta^* (A_{\max} - A_{i1}) \} - 1. \]
6. If \( Z_n \leq (1-P^*)/P^* \), stop sampling, and select as best that \( \pi_i \) corresponding to \( A_{\max} \). Otherwise, go to Step 3.

Let \( n_{BKS} \) denote the random number of stages required for \( P_{BKS} \) to terminate (i.e., the value of the counter \( n \) at termination). The distribution
of \( N_{\text{BKS}} \) depends on the underlying configuration of the population means \( \mu \); two configurations are of special interest:

**Least favorable configuration (LFC):**
\[
\mu[1] = \mu[k-1] = \mu[k] - \delta^*.
\]

**Equal means configuration (EMC):**
\[
\mu[1] = \mu[k].
\]

It can be shown that the LFC maximizes \( E[N_{\text{BKS}}|\mu] \) over all \( \mu \in A_* \). Thus, the LFC is the "worst case" configuration for \( \mu \in A_* \). Similarly, the EMC is the "worst case" configuration over the entire space of \( \mu \)'s since \( E[N_{\text{BKS}}|\mu] \leq E[N_{\text{BKS}}|\mu=\text{EMC}] \) \( \forall \mu \). If (unknown to the experimenter) \( \mu \) is in the EMC, \( P_{\text{BKS}} \) might require an inordinate number of stages.

**Example 1** (continued): Again suppose that \( k = 3 \), \( \sigma^2 = 1 \), \( P^* = 0.75 \), and \( \delta^* = 0.2 \). Behbohfer and Goldman (B-G) (1987b) show that \( P_{\text{BKS}} \) yields \( P[CS|\mu=\text{LFC}] \approx 0.777 \), \( E[N_{\text{BKS}}|\mu=\text{LFC}] \approx 38.84 \), and \( E[N_{\text{BKS}}|\mu=\text{EMC}] \approx 47.98 \). Therefore, for this example, these expected values are less than the 52 observations required by each population for the single-stage \( P_{\text{SS}} \).

Two unappealing features of \( P_{\text{BKS}} \) are the fact that it is open and that it generally gives \( P[CS|\mu=\text{LFC}] \) which is somewhat larger than the desired \( P^* \). (In the previous example, \( P[CS|\mu=\text{LFC}] \approx 0.777 > 0.75 = P^* \).) B-G (1987b) study a closed version \( (P_{\text{BKS-T}}) \) of \( P_{\text{BKS}} \) which employs automatic truncation of sampling after a certain predetermined number of stages have been conducted. The truncation point \( (n_0) \) is the smallest upper bound on the number of stages which still guarantees (PR); tedious Monte Carlo experimentation is required in order to determine \( n_0 \) for given \( k, \sigma^2, \delta^*, P^* \). For the above example, \( P_{\text{BKS-T}} (n_0 = 68 \text{ stages}) \) yields \( P[CS|\mu=\text{LFC}] \approx 0.751 \), \( E[N_{\text{BKS-T}}|\mu=\text{LFC}] \approx 35.94 \), and \( E[N_{\text{BKS-T}}|\mu=\text{EMC}] \approx 41.90 \). \( P_{\text{BKS-T}} \) always results in less probability "overprotection" than \( P_{\text{BKS}} \).

3.3 Other Sequential Procedures

Other single-factor indifference-zone normal means procedures are due to Paulson (1964), Alam (1970), Fabian (1974), and Tantke and Bechhofer (1977, 1979). The reader should also consult the texts of Gibbons, Olkin, and Sobel (1977) [in addition, see Bechhofer (1980)] and Gupta and Panchapakesan (1979), both of which give more comprehensive treatments on ranking and selection techniques (primarily from a single-factor point of view).

4. THE 2-FACTOR NORMAL MEANS PROBLEM

This section parallels the development of Section 2. We are now concerned with 2-factor experiments without interaction. References are B-G (1987a, 1987c).

Suppose that we have a-b independent normal populations \( \pi_{ij} \) \( (i \leq a, \ j \leq b) \) with unknown means \( \mu_{ij} \) and a common known variance \( \sigma^2 \); we henceforth assume without loss of generality that \( \sigma^2 = 1 \). Further, we assume the usual additive fixed-effects ANOVA model,

\[
\mu_{ij} = \mu + \alpha_i + \beta_j,
\]

where \( \Sigma \alpha_i = \Sigma \beta_j = 0 \); that is, there is no interaction between the factor-levels. \( \alpha_i \) (\( \beta_j \)) is the "effect" of the ith (jth) level of factor A (B) on \( \mu \).

Analogous to the single-factor case, denote the ordered but unknown \( \alpha_i \)'s and \( \beta_j \)'s by \( \alpha[1] \leq \cdots \leq \alpha[a] \) and \( \beta[1] \leq \cdots \leq \beta[b] \). We do not know the values of the \( \alpha_i \)'s and \( \beta_j \)'s nor how they are paired with the \( \pi_{ij} \)'s. Our goal is to determine which of the a-b \( \pi_{ij} \)'s has the largest mean (i.e., the \( \pi_{ij} \) associated with \( \alpha[a] \) and \( \beta[b] \)); we call that \( \pi_{ij} \) the 'best' population. If the best population is indeed chosen by a procedure, a CS has been made. We limit consideration to procedures that guarantee the following 2-factor indifference-zone probability requirement on the \( P[CS] \).

\[
P[CS] \geq P^* \text{ whenever } \left\{ \begin{array}{l}
\alpha[a] - \alpha[a-1] \geq \delta^a \\
\beta[b] - \beta[b-1] \geq \delta^b
\end{array} \right.
\]

\( \delta^a, \delta^b, \) and \( P^* \) are three constants which must be specified prior to the start of experimentation.
subject to the restrictions that \( \delta_0^* > 0, \delta_0^* > 0, \) and \( 1/ab < P^* < 1. \) For ease of exposition, we only consider the case \( \delta_0^* = \delta_0^* = \delta^*, \) say.

5. **2-FACTOR INDIFFERENCE-ZONE PROCEDURES**

Here, we parallel the development of Section 3. We give a number of indifference-zone procedures for the 2-factor normal means problem discussed in Section 4.

5.1 **A Single-Stage Procedure**

The following single-stage procedure is due to Bechhofer (1954).

**Procedure \( P_{SS2} \):**

1. For given \( a \) and \( b \) and known \( \sigma^2 = 1, \) specify \( \{\delta^*, P^*\}, \) with \( \delta^* > 0 \) and \( 1/ab < P^* < 1. \)

2. Determine \( n_{SS2} \), a constant satisfying

\[
\int_0^\infty \left[ \Phi(x + \delta^* a n_{SS2}) \right]^{a-1} d\Phi(x) \cdot \int_0^\infty \left[ \Phi(x + \delta^* b n_{SS2}) \right]^{b-1} d\Phi(x) \geq P^*
\]

[A table of \( n_{SS2} \)-values is given in B-G (1967c).]

3. Take \( n_{SS2} \) independent observations from each of the \( a \times b \) populations.

4. Let \( x_{ijs} \) denote the \( st \) observation from \( \Pi_{ij}, \) \( i = 1, \ldots, a, j = 1, \ldots, b, \) and \( s = 1, \ldots, n_{SS2}. \) Calculate the quantities

\[
A_i = \sum_{j=1}^{b} \sum_{s=1}^{n_{SS2}} x_{ijs}; \quad i = 1, \ldots, a
\]

and

\[
B_j = \sum_{i=1}^{a} \sum_{s=1}^{n_{SS2}} x_{ijs}; \quad j = 1, \ldots, b.
\]

5. Select as best that \( \Pi_{ij} \) corresponding to the largest \( A_i \) and the largest \( B_j. \)

Example 2: Suppose that \( a = 3, \) \( b = 6, \) \( \sigma^2 = 1, \) \( P^* = 0.0025 = (.975)^2, \) and \( \delta^* = 0.2. \) From a table in B-G (1967c), we find that \( n_{SS2} = 64. \) So we must take 64 independent observations from each of the 18 \( \Pi_{ij} \) 's.

5.2 **An Open Sequential Procedure**

The next procedure is the 2-factor generalization of \( P_{BK3} \) and is itself a special case of a more general treatment of Koopman-Darmois populations without interaction [cf. B-G (1987a)].

**Procedure \( P_{BK3} \):**

1. For given \( a \) and \( b \) and known \( \sigma^2 = 1, \) specify \( \{\delta^*, P^*\}, \) with \( \delta^* > 0 \) and \( 1/ab < P^* < 1. \)

2. Set the stage counter \( n = 0. \)

3. Set \( n = n + 1. \) Take one observation \( x_{ij} \) from each \( \Pi_{ij}, \) \( i = 1, \ldots, a, j = 1, \ldots, b. \) [Thus, we actually take observations a matrix-at-a-time.]

4. Calculate the quantities

\[
A_{in} = \sum_{j=1}^{b} \sum_{s=1}^{n_{SS2}} x_{ijns}; \quad i = 1, \ldots, a,
\]

\[
B_{jn} = \sum_{i=1}^{a} \sum_{s=1}^{n_{SS2}} x_{ijns}; \quad j = 1, \ldots, b,
\]

\[
A_{\text{max}} = \max_{1 \leq i \leq a} A_{in}, \quad \text{and}
\]

\[
P_{\text{max}} = \max_{1 \leq j \leq b} B_{jn}.
\]

5. Calculate

\[
U_n = \sum_{i=1}^{a} \exp \{-\delta^*(A_{\text{max}} - A_{in})\} - 1,
\]

\[
V_n = \sum_{j=1}^{b} \exp \{-\delta^*(B_{\text{max}} - B_{jn})\} - 1,
\]

and

\[
Z_n = U_n + V_n + U_n V_n.
\]

6. If \( Z_n \leq (1-P^*)/P^* \), stop sampling, and select as best that \( \Pi_{ij} \) corresponding to \( A_{\text{max}} \) and \( B_{\text{max}}. \) Otherwise, go to Step 3.

Let \( n_{BK3} \) denote the random number of stages required for \( P_{BK3} \) to terminate. The 2-factor LPC is given by

\[
\alpha_1 = \alpha_{[a-1]} - \delta^*,
\]

\[
\beta_1 = \beta_{[b-1]} - \delta^*,
\]

and the EMC is given by

\[
\alpha_{[a]} = \beta_{[b]} = 0.
\]
Example 2 (continued): Again suppose that $a = 3$, $b = 6$, $s^2 = 1$, $P^* = 0.9025$, and $\delta = 0.2$. B-G (1987c) show that $P_{RKS2}$ yields $P_{CS}[P = \text{LFC}] \neq 0.022$, $E[N_{RKS2}^\ast | P = \text{LFC}] \neq 41.80$, and $E[N_{RKS2}^\ast | P = \text{EMC}] \neq 17.76$. Recall that $P_{SS2}$ required 64 observations from each population; so for this example, $P_{RKS2}$ is more parimsomous than $P_{SS2}$ in the LFC.

B-G (1987c) also study a version $(P_{RKS2-T})$ of $P_{RKS2}$ which employs truncation and still guarantees the probability requirement. For the above example, $P_{RKS2-T}$ (with a truncation point of $n_0 = 60$ stages) yielded $P_{CS}[P = \text{LFC}] \neq 0.904$, $E[N_{RKS2-T}^\ast | P = \text{LFC}] \neq 43.53$, and $E[N_{RKS2-T}^\ast | P = \text{EMC}] \neq 70.86$. By design, $P_{RKS2-T}$ always outperforms $P_{RKS2}$ in terms of these and other criteria. Further, the Monte Carlo work in B-G (1987c) suggests that $P_{RKS2-T}$ substantially outperforms $P_{SS2}$ in the LFC, and that $P_{RKS2-T}$ never fares too much worse than $P_{SS2}$ in the EMC.

5.3 Why Conduct 2-Factor Experiments?

One might think that conducting a sequential 2-factor procedure, which samples observations a matrix-at-a-time, is less efficient than carrying out two independent single-factor experiments, which takes observations a vector-at-a-time. A major finding of B-G (1987a) shows that this is not the case. Although the 2-factor procedures take more observations per stage (a+b) than do the corresponding single-factor procedures (a and b, respectively), the 2-factor procedures usually terminate sampling in far fewer stages. Indeed, B-G (1987a) find that, when the situation at hand is appropriate, an experimenter can usually save substantially on the total number of scalar observations by using a 2-factor procedure instead of two single-factor experiments — about 40% when the population means are in the LFC and much more so when the means are in the EMC!

6. CONCLUSIONS

This tutorial served to give the simulation practitioner a flavor of some of the ranking and selection procedures that are available in the literature. In this paper, we concentrated on the single- and 2-factor normal means problem (with common known variance), but ranking and selection procedures exist to solve a wide variety of other interesting problems. In fact, a number of papers have recently appeared which formulate and/or utilize ranking and selection techniques in the simulation environment, e.g., Sullivan and Wilson (1965).

Acknowledgement: The author is pleased to thank Prof. Robert Bechhofer for his comments and encouragement.

REFERENCES


**AUTHOR'S BIOGRAPHY**

DAVE GOLDSMAN is an Assistant Professor of Industrial and Systems Engineering at Georgia Tech. He holds degrees from Syracuse and Cornell, where he spent last summer as a Visiting Scientist. His research interests include simulation output analysis, and ranking and selection.

Dave Goldman  
School of ISyE  
Georgia Tech  
Atlanta, GA 30332  
(404) 894-2365