DESIGN OF EXPERIMENTS WITH SELECTION AND RANKING GOALS

by

Shanti S. Gupta and S. Panchapakesan
Purdue University Southern Illinois University

Technical Report #95-29C

Department of Statistics
Purdue University

June 1995
(Revised September 1995)

*

DESIGN OF EXPERIMENTS WITH SELECTION AND RANKING GOALS

by

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, IN 47907-1399

and

S. Panchapakesan
Department of Mathematics
Southern Illinois University
Carbondale, IL 62901-4408

Abstract

This paper selectively reviews procedures for selecting the best among several treatments, procedures for comparing experimental treatments with a standard or a control, and simultaneous confidence intervals for comparisons with the best and comparisons with a standard or a control. These procedures are discussed under the assumption of normality using both indifference-zone and subset approaches. Most of the discussions relate to single-factor experiments (with or without blocking) and 2-factorial experiments with or without interaction. A brief discussion deals with models involving Bernoulli and multinomial distributions and restricted families such as IFR and IFRA distributions.

Keywords and phrases: Selection, indifference-zone, subset, normal, Bernoulli, multinomial, single-factor experiments, blocking, factorial experiments, comparisons with control, simultaneous confidence statements.

Design of Experiments With Selection And Ranking Goals

Shanti S. Gupta and S. Panchapakesan

1. Introduction

In many practical situations in everyday life, the experimenter is faced with the problem of comparing $k(\geq 2)$ alternatives with a view to select the "best" among them. These may, for example, be different varieties of wheat in an agricultural experiment, or different coherent systems in engineering models, or different drugs prescribed for treatment of a certain disease. In all these problems, each alternative is characterized by the value of a parameter θ. In the above-mentioned situations, this parameter may be the average yield of a variety of wheat, or the reliability function of a system, or a measure of the effectiveness of a drug.

Consider the well-known balanced one-way layout given by the model:

$$Y_{ij} = \mu_i + \epsilon_{ij}, i = 1, \ldots, k; j = 1, \ldots, n \tag{1.1}$$

where Y_{ij} is the jth response on the ith treatment, the μ_i are *unknown* treatment means, and the ϵ_{ij} are the measurement errors assumed to be independent and normally distributed with mean zero and variance σ^2. The classical approach in this case is generally to test the so-called homogeneity hypothesis $H_0 : \mu_1 = \mu_2 = \ldots = \mu_k$ using the analysis of variance (ANOVA) approach. However, this does not serve the experimenter's real purpose which is not just to accept or reject the homogeneity hypothesis of no treatment differences. In practice, the experimenter will be considering treatments that are indeed different, and with sufficiently large sample, will be rejecting H_0 at any specified level. The real goal of the experimenter to be addressed then is to identify the best alternative (the variety with the largest average yield, the most reliable system, the most effective drug, and so on). Thus the inadequacy of the ANOVA lies in the types of decisions that are made on the basis of the data and not in the design aspects of the procedure. The method of estimating the sizes of the differences between the treatments was often used as an indirect way of reaching a decision regarding the best treatment(s). The attempts to formulate the decision problem to achieve this realistic goal set the stage for the development of the selection and ranking theory.

Selection and ranking problems have generally been formulated adopting one of two main approaches now familiarly known as the *indifference-zone formulation* and the *subset selection formulation.* Consider an experiment with k treatments in which we have n responses for each treatment. We assume the model (1.1).

Let the ordered μ_i be denoted by $\mu_{[1]} \leq \ldots \leq \mu_{[k]}$. It is assumed that there is no information regarding the correct pairing between the ordered and the unordered μ_i.

In the indifference-zone approach due to Bechhofer (1954), the goal is to select the treatment associated with the largest mean $\mu_{[k]}$ (called the best treatment). A selection satisfying the goal is defined to be a *correct selection* (CS). It is required that a correct selection be guaranteed with a probability $P^*(\frac{1}{k} < P^* < 1)$ whenever $\mu_{[k]} - \mu_{[k-1]} \geq \delta^*$, where δ^* is a positive constant. Here δ^* and P^* are specified *in advance* by the experimenter, and P^* is chosen greater than $1/k$ because otherwise we can make a no-data decision by selecting one of the treatments randomly as the best. When $\mu_{[k]} - \mu_{[k-1]} < \delta^*$, two or more treatments including the best are *sufficiently close* and the experimenter is assumed to be indifferent as to setting probability requirement in this case. The region $\Omega_{\delta^*} = \{(\underset{\sim}{\mu}, \sigma^2) | \underset{\sim}{\mu} = (\mu_1, \ldots, \mu_k), \sigma^2 > 0, \mu_{[k]} - \mu_{[k-1]} \geq \delta^*\}$ is called the *preference-zone* and its complement w.r.t. the entire parameter space $\Omega = \{(\underset{\sim}{\mu}, \sigma^2) | -\infty < \mu_i < \infty, i = 1, \ldots, k, \sigma^2 > 0\}$ is the *indifference-zone.* Denoting the probability of a correct selection (PCS) using the rule R by $P(CS|R)$, it is required that any valid rule R satisfy the condition:

$$P(CS|R) \geq P^* \text{ whenever } (\underset{\sim}{\mu}, \sigma^2) \epsilon\ \Omega_{\delta^*}. \tag{1.2}$$

The design aspect of this basic setup is the determination of the minimum (common) sample size n so that the probability requirement (1.2) is satisfied.

In the subset selection approach developed by Gupta (1956), the goal is to select a nonempty subset of the k treatments so that the best treatment will be included in the selected subset with a guaranteed minimum probability P^*. The subset size is not specified in advance; it is random and determined by the data. Formally, any valid rule should satisfy the condition:

$$P(CS|R) \geq P^* \text{ for all } (\underset{\sim}{\mu}, \sigma^2) \epsilon\ \Omega. \tag{1.3}$$

It is obvious that the requirement (1.3) can always be met by including all the treatments

in the selected subset. So the performance of a rule is studied usually in terms of the expected size B of the selected subset. It is expected that a reasonable procedure will tend to select only one treatment when $\mu_{[k]} - \mu_{[k-1]}$ gets large.

Besides being a goal in itself, selecting a subset containing the best can also be considered as the first-stage screening in a two-stage procedure designed to select one treatment as the best; see, for example, Tamhane and Bechhofer (1977, 1979).

The probability requirements (1.2) and (1.3) are also known as the P^*-conditions. An important step in obtaining the constant(s) associated with a proposed rule R so that the P^*-condition is satisfied is to evaluate the infimum of the PCS over Ω or Ω_{δ^*} depending on the approach. Any configuration of $(\underset{\sim}{\mu}, \sigma^2)$ for which the infimum is attained is called a *least favorable configuration* (LFC).

Although we have discussed the selection problem in terms the normal means, the problem in general is to select from k populations $\Pi_1, \ldots, \Pi_k$ characterized by the distribution functions $F_{\theta_i}, i = 1, \ldots, k$, respectively, where the θ_i are unknown parameters taking values in the set Θ. The populations are ranked in terms of the θ_i (There may be other nuisance parameters). The ordered θ_i are denoted by $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ and the population associated with $\theta_{[k]}$ is defined to be the best. To define the preference-zone, one has to define a suitable nonnegative measure $\delta(\theta_i, \theta_j)$ of the separation between the populations Π_i and Π_j. Then $\Omega_{\delta^*} = \{\underset{\sim}{\theta} | \underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k), \delta(\theta_{[k]}, \theta_{[k-1]}) \geq \delta^* > 0\}$.

There are several variations and generalizations of the basic goal in both indifference-zone and subset selection formulations. One can generalize the goal to select at least s of the t best populations with $1 \leq s \leq t \leq k-1$. In the subset selection approach, the size of the selected subset can be random subject to a specified maximum $m (1 \leq m \leq k)$. This approach of *restricted subset selection* studied by Gupta and Santner (1973) and Santner (1975) combines the features of the indifference-zone and subset selection formulations. There have been other attempts in this direction of integrated formulations; for example, see Sobel (1969), Chen and Sobel (1987a, 1987b). An important modification of the goal of selecting the best population is selecting a good population or a subset containing only good populations or containing all the good populations. A good population is defined as one which is close enough to the best within a specified threshold value.

There is now a vast literature on selection and ranking procedures. Several aspects of the theory and associated methodology of these and related procedures have been dealt with in the books by Bechhofer, Kiefer and Sobel (1968), Büringer, Martin and Shriever (1980), Gibbons, Olkin and Sobel (1977), Gupta and Huang (1981), Gupta and Panchapakesan (1979) and Mukhopadhyay and Solanky (1994). A very recent book is by Bechhofer, Santner and Goldsman (1995). A categorical bibliography is provided by Dudewicz and Koo (1982). Besides these books, there have been published review articles dealing with several specific aspects of selection and ranking. The reader is specially referred to Gupta and Panchapakesan (1985, 1988, 1991, 1993), Panchapakesan (1992, 1995a, 1995b), and Van der Laan and Verdooran (1989).

In spite of the vast published literature, there have been only a few papers until the recent years devoted to design models beyond single-factor experiments. In this paper, besides the most common single-factor experiments involving mainly normal distributions, we review significant results involving blocking and factorial designs. The emphasis is not on a total coverage but enough to provide a focus on these problems to help assess the current status and potential for applications and further investigations.

Sections 2 through 5 discuss selection procedures and simultaneous confidence intervals under the assumption that the observed responses for treatments are normally distributed. Selection procedures under both the indifference-zone and the subset formulations are discussed in each section. Section 2 deals with selecting the best treatment and simultaneous confidence statements for comparisons with best in single-factor experiments. Section 3 considers these procedures in experiments with blocking while Section 4 deals with these procedures in factorial experiments. Selection with respect to a standard or control is discussed in Section 5. Selection in experiments involving other models is briefly discussed in Section 6. The models discussed are: Bernoulli, multinomial and restricted families such as IFR and IFRA.

2. Selecting the Best Treatment in Single-Factor Experiments: Normal Theory

Consider $k \geq 2$ treatments $\Pi_1, \ldots, \Pi_k$, where Π_i represents a normal population with mean μ_i and variance σ_i^2. The means μ_i are *unknown*. Different assumptions can be made about the σ_i^2 depending on the context of the experiment. As before, the ordered μ_i are

denoted by $\mu_{[1]} \leq \ldots \leq \mu_{[k]}$ and no prior information is available regarding the true pairing of the ordered and unordered μ_i.

2.1 Indifference-Zone Approach

As described in Section 1, the goal is to identify one of the k treatments as the best (the one associated with $\mu_{[k]}$) with a guaranteed minimum probability P^* of a correct selection whenever $\mu_{[k]} - \mu_{[k-1]} \geq \delta^*$, where $\delta^* > 0$ and $1/k < P^* < 1$ are specified in advance.

Under the assumption that $\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2$ (*known*), Bechhofer (1954) proposed the following single-stage procedure based on samples of common size n. Let $\overline{Y}_i$ denote the mean of the sample responses Y_{ij} $j = 1, \ldots, n$, from $\Pi_i, i = 1, \ldots, k$. His rule is

$$R_1: \quad \text{Select the treatment } \Pi_i \text{ that yields the largest } \overline{Y}_i. \tag{2.1}$$

The LFC for this rule is given by $\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$. For given $(k, \delta^*/\sigma, P^*)$, the minimum sample size n required to meet the P^*-condition is given by

$$n = < 2\left(\frac{\sigma H}{\delta^*}\right)^2 >, \tag{2.2}$$

where $< x >$ denotes the smallest integer $\geq x$, H satisfies

$$Pr\{Z_1 \leq H, \ldots, Z_{k-1} \leq H\} = P^*, \tag{2.3}$$

and the Z_i are standard normal variates with equal correlation $\rho = \frac{1}{2}$. Values of H can be obtained for several selected values of k and P^* from Bechhofer (1954), Gibbons, Olkin and Sobel (1977), Gupta (1963), Gupta, Nagel and Panchapakesan (1973), and Milton (1963).

Hall (1959) and Eaton (1967) have shown that the rule R_1 in (2.1) is the most economical in the sense of requiring fewest observations per treatment among all single-stage location invariant procedures satisfying the P^*-condition.

Two stage procedures for the problem of selecting the normal treatment with the largest mean assuming a common known variance σ^2 have been studied by Cohen (1959), Alam (1970) and Tamhane and Bechhofer (1977, 1979). These procedures use the subset

selection procedure of Gupta (1956, 1965) to eliminate inferior treatments at the first stage and select the best from among the remaining ones at the second stage. We describe the Tamhane-Bechhofer procedure R_2 below.

R_2: Take a random sample of n_1 observations from each $\Pi_i, i = 1, \ldots, k$. Eliminate from further consideration all treatments Π_i for which $\overline{Y}_i < \overline{Y}_{[k]} - h\sigma/\sqrt{n_1}$, where $\overline{Y}_{[1]} \leq \ldots \leq \overline{Y}_{[k]}$ are the ordered sample means $\overline{Y}_i$, and h is a constant to be determined. If only one treatment remains, then it is selected as the best. If more than one treatment remain, then proceed to the second stage by taking an additional random sample of size n_2 from each of these remaining treatments. Select the treatment that yields the largest sample mean based on the combined sample of $n_1 + n_2$ observations.

The above procedure R_2 of Tamhane and Bechhofer (1977, 1979) involves constants (n_1, n_2, h) to be determined in order to satisfy the P^*-condition. These constants are determined by using a minimax criterion (in addition to the P^*-condition) which minimizes the maximum over the entire parameter space Ω of the expected total sample size required by the procedure. The LFC for this procedure was first established only for $k = 2$. The constants (n_1, n_2, h) tabulated by Tamhane and Bechhofer (1979) for selected values of k, P^*, and δ^*/σ are conservative since they are based on the LFC for a lower bound of the PCS. The fact that the LFC for the PCS is $\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ was proved by Sehr (1988) and Bhandari and Chaudhuri (1990).

A truncated sequential procedure for this problem has been investigated by Bechhofer and Goldsman (1987, 1989). It is designed to have improved performance over an earlier procedure of Bechhofer, Kiefer and Sobel (1968) which is an *open non-eliminating* sequential procedure as opposed to the Bechhofer-Goldsman procedure which is a *closed* but also a non-eliminating procedure. A multi-stage or sequential procedure is called *open* if, in advance of the experiment, no fixed upper bound is set on the number of observations to be taken from each treatment; otherwise, it is called *closed.* An *eliminating* procedure is one which excludes treatments from further sampling if they are removed from further consideration at any stage prior to taking the terminal decision. A non-eliminating procedure, on the other hand, samples from each treatment at each stage whether or not any treatment is removed from the final consideration of selection.

Another well-known procedure for the problem under discussion is that of Paulson (1964) which is a closed procedure with elimination. This procedure was successively improved (by changing the choices for certain constants) by Fabian (1974) and Hartman (1988). For further details regarding various procedures and their performance, see Gupta and Panchapakesan (1991).

We now consider the case of *unknown* common variance σ^2. This is the classical problem of the one-way ANOVA model. If one chooses to define the preference-zone as $\Omega_{\delta^*} = \{(\underline{\mu}, \sigma) | \mu_{[k]} - \mu_{[k-1]} \geq \delta^* \sigma\}$, then the single-stage procedure R_1 in (2.1) can still be used with the minimum required sample size n given by (2.2) with $\sigma = 1$. If we continue with the preference-zone $\Omega_{\delta^*} = \{(\underline{\mu}, \sigma) | \mu_{[k]} - \mu_{[k-1]} \geq \delta^*\}$ as before, it is not possible to devise a single-stage procedure that satisfies the P^*-condition. This is intuitively clear from the fact that the determination of the minimum sample size required depends on the knowledge of σ. In this case of unknown σ^2, Bechhofer, Dunnett and Sobel (1954) proposed an *open two-stage non-eliminating* procedure for selecting the best treatment. The first stage is used to estimate σ^2 and determine the total sample size needed to guarantee the probability requirement. The second stage, if necessary, is used to make the terminal decision. Using in addition the idea of screening, Tamhane (1976) and Hochberg and Marcus (1981) have studied three-stage procedures where the first stage is utilized to determine the additional sample sizes necessary in the subsequent stages, the second stage is used to eliminate inferior populations by a subset rule, and the third stage (if necessary) to make the final decision. Tamhane (1976) also considered a two-stage eliminating procedure which was found to be inferior to the non-eliminating procedure of Bechhofer, Dunnett and Sobel (1954). Later, Gupta and Kim (1984) proposed a *two-stage eliminating* procedure with a new design criterion and obtained a sharp lower bound on the PCS. Gupta and Miescke (1984) studied two-stage eliminating procedures using a Bayes approach. Here we will describe the procedure of Gupta and Kim (1984).

R_3: Take a random sample of size $n_1 (\geq 2)$ from each treatment. Let $\overline{X}_i$ be the sample mean associated with treatment Π_i, $i = 1, \ldots, k$, and S_ν^2 denote the usual pooled sample variance based on $\nu = k(n_1 - 1)$ degrees of freedom. Determine the subset I of $\{\Pi_1, \ldots, \Pi_k\}$ given by

$$I = \{\Pi_i | \overline{X}_i \geq \overline{X}_{[k]} - (dS/\sqrt{n_1} - \delta^*)^+\},$$

where $a^+ = \max(a, 0)$ and d is a constant to be chosen to satisfy the P^*-condition. If I consists of only one treatment, then select it as the best; otherwise, take an additional sample of size $N - n_1$ from each treatment in I, where

$$N = \max\{n_1, < (hS_\nu/\delta^*)^2 >\},$$

$< y >$ denotes the smallest integer $\geq y$, and h is a positive constant to be suitably chosen to satisfy the P^*-condition. Now, select as the best the population in I which yields the largest sample mean based on the combined sample of size N.

There are several possible choices for (n_1, d, h) to satisfy the P^*-condition. Gupta and Kim (1984) used the requirement that

$$Pr\{\text{ the best population is included the subset } I\} \geq P_1^*, \tag{2.4}$$

where $P_1^*(P^* < P_1^* < 1)$ is pre-assigned. Evaluation of these constants is based on a lower bound for the PCS. The Monte Carlo study of Gupta and Kim (1984) shows that their procedure R_3 performs much better than that of Bechhofer, Dunnett and Sobel (1954) in terms of the expected total sample size.

Recently, there have been a series of papers regarding the conjecture of the LFC for the Tamhane-Bechhofer procedure R_2 and some other related procedures for selecting the best normal treatment when the common variance σ^2 is known. As mentioned earlier, the conjecture that the LFC is $\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ has been proved by Sehr (1988) and Bhandari and Chaudhuri (1990). The LFC's for two-stage procedures for more generalized goals have been established by Santner and Hayter (1993) and Hayter (1994). It will be interesting to reexamine the performance of the concerned procedures by using the exact infimum of the PCS.

2.2 Subset Selection Approach

As in Section 2.1, we are still interested in selecting the best treatment. However, we do not set in advance the number of treatments to be included in the selected subset. It is expected that a good rule will tend to select only one population as $\mu_{[k]} - \mu_{[k-1]}$ gets sufficiently large. Gupta (1956) considered the case of known as well as unknown σ^2.

Based on samples of size n from each population, his rule, in the case of known σ^2, is R_4: Select the treatment Π_i if and only if

$$\overline{X}_i \geq \max_{1 \leq j \leq k} \overline{X}_j - \frac{d\sigma}{\sqrt{n}}, \tag{2.5}$$

where $\overline{X}_i$ is the sample mean from Π_i and d is the smallest positive constant for which the PCS $\geq P^*$ for all $(\mu, \sigma) \in \Omega$. (Any larger d would obviously satisfy the P^*-condition but would, if anything, only increase the size of the selected subset). This smallest d is given by $\sqrt{2H}$ where H is the solution of (2.3). Thus d can be obtained from the tables mentioned previously.

When σ^2 is unknown, Gupta (1956) proposed the rule R_5 which is R_4 with σ replaced by S_ν, where S_ν^2 is the usual pooled estimator of σ^2 with $\nu = k(n-1)$ degrees of freedom. To keep the distinction between the two cases, we use d' in the place of d. The smallest d' needed to satisfy the P^*-condition is the one-sided upper $(1 - P^*)$ equicoordinate point of the equicorrelated $(k-1)$-variate cenetral t-distribution with the equal correlation $\rho = 0.5$ and the associated degrees of freedom $\nu = k(n-1)$ The values of d' have been tabulated by Gupta and Sobel (1957) for selected values of k, n, and P^*. They are also available from the tables of Gupta, Panchapakesan and Sohn (1985) corresponding to correlation $\rho = 0.5$.

It should be noted that, unlike in the case of the indifference-zone approach, we do have a single-stage procedure for any specified n when σ^2 is unknown.

We may not have a common variance σ^2 (the heteroscedasticity case) or a common sample size (unbalanced design). These cases have been studied by Gupta and Huang (1976), Chen, Dudewicz and Lee (1976), and Gupta and Wong (1982). In all these cases, the authors have used lower bounds for the infimum of the PCS to meet the P^*-condition.

When the variances are unknown and unequal, and the sample sizes are unequal, Dudewicz and Dalal (1975) proposed a two-stage procedure using both the indifference-zone and subset selection approaches. Sequential subset selection procedures have also been studied which are applicable to the normal model. For a review of these, the reader is referred to Gupta and Panchapakesan (1991).

As we have pointed out previously, several modifications of the basic goal have been

investigated. In particular, we mention here the restricted subset selection approach which includes a specified upper bound for the expected subset size which is otherwise random. Procedures of this type have been proposed by Gupta and Santner (1973) and Santner (1975).

Several authors have also studied the modified goal of selecting good populations. Reference can be made to Gupta and Panchapakesan (1985) and Gupta and Panchapakesan (1991).

2.4 Simultaneous Confidence Intervals for Comparisons with the Best

Related to the selection and ranking objectives is the multiple comparison approach in which one seeks simultaneous confidence sets for meaningful contrasts among a set of given treatments. A comprehensive treatment of this topic can be found in the text by Hochberg and Tamhane (1987). Our main interest here is simultaneous comparisons of all treatments with the best among them. In other words, we are interested in simultaneous confidence intervals for $\mu_i - \max_{j \neq i} \mu_j$, taking a larger treatment effect to imply a better treatment. If $\mu_i - \max_{j \neq i} \mu_j < 0$, then treatment i is not the best and the difference represents the amount by which treatment i is inferior to the best. On the other hand, if the difference is positive, then treatment i is the best and the difference is the amount by which it is better than the second best.

Assume that all (normal) treatments Π_i have a common unknown variance σ^2. Let $\overline{Y}_i$ denote the mean of n independent responses on treatment $\Pi_i, i = 1, \ldots, k$. Let S_ν^2 denote the usual pooled (unbiased) estimator of σ^2 based on $\nu = k(n-1)$ degrees of freedom. Hsu (1984) showed that the intervals

$$[-(\overline{Y}_i - \max_{j \neq i} \overline{Y}_j - C)^-,\ (\overline{Y}_i - \max_{j \neq i} \overline{Y}_j + C)^+],\ i = 1, \ldots, k, \tag{2.6}$$

form $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_i - \max_{j \neq i} \mu_j, i = 1, \ldots, k$. Here $-x^- = \min(x, 0), x^+ = \max(x, 0)$, and $C = cS_\nu/\sqrt{n}$, where c satisfies

$$P[\overline{Y}_k - \max_{j \neq k} \overline{Y}_j > -cS_\nu/\sqrt{n}] = 1 - \alpha \tag{2.7}$$

under the assumption that $\mu_1 = \ldots = \mu_k = 0$.

The intervals in (2.6) are closely related to the selection and ranking methods discussed previously. It was shown by HSU (1984) that the upper bounds of these intervals imply the subset selection inference of Gupta (1956) and the lower bounds imply the indifference-zone selection inference of Bechhofer (1954). Any treatment i for which the upper bound of $\mu_i - \max_{j \neq i} \mu_j$ is zero can be inferred to be not the best. Similarly, any treatment i for which the lower bound of $\mu_i - \max_{j \neq i} \mu_j$ is zero can be inferred to be the best.

The intervals in (2.6) are "constrained" in the sense that the lower bounds are non-positive and the upper bounds are nonnegative. Removing these constraints would require an increase in the critical value. The nonnegativity constraint on the upper bounds does not present a great disadvantage as one will not normally be interested in knowing how bad is a treatment that is rejected as not the best. On the other hand, it will be of interest to assess how much better than others is a treatment inferred to be the best. Motivated by these considerations Hsu (1985) provided a method of unconstrained multiple comparisons with the best which removes the nonpositivity constraint on the lower bounds in (2.6) by increasing the critical value slightly. We describe these simultaneous intervals below.

Let $D = dS_\nu/\sqrt{n}$ where d is the solution of

$$Pr\{Z_i \leq Z_k + dS_\nu/\sqrt{n}, i = 1, \ldots, k-2, |Z_{k-1} - Z_k| \leq dS_\nu/\sqrt{n}\} = 1 - \alpha, \qquad (2.8)$$

where the Z_i are independent $N(0,1)$ variables. The $100(1-\alpha)\%$ simultaneous confidence intervals $[D_i^*, D_i^{**}]$ of Hsu (1985) for the differences $\mu_i - \max_{j \neq i} \mu_j$ are defined as follows:

$$D_i^* = \overline{Y}_i - \max_{j \neq i} \overline{Y}_j - D \qquad \text{for } i = 1, \ldots, k,$$

$$D_i^{**} = \begin{cases} (\overline{Y}_i - \max_{j \neq i} \overline{Y}_j + D)^+ \wedge (-D_{[k]}^*) & \text{if } \overline{Y}_i \neq \overline{Y}_{[k]}, \\ \overline{Y}_i - \max_{j \neq i} \overline{Y}_j + D & \text{if } \overline{Y}_i = \overline{Y}_{[k]}. \end{cases} \qquad (2.9)$$

Here $\overline{Y}_{[k]}$ denotes the largest $\overline{Y}_i$ and $a \wedge b = \min(a, b)$. The constrained simultaneous confidence intervals can be implemented by using a computer SAS package; see Gupta and Hsu (1984, 1985), and Aubuchon, Gupta and Hsu (1986). Hsu (1985) has tabulated the values of $d/\sqrt{2}$ for $k = 2(1)5, \nu = 5(1)20, 24, 30, 40, 60, 120, \infty$, and $\alpha = .01, .05$.

2.5 Estimation After Selection

Consider the selection rule R_1 of Bechhofer (1954) defined in (2.1) for selecting the best treatment, namely, the one associated with the largest μ_i. This rule selects the population that yields the largest sample mean $\overline{Y}_i$. Let μ_S denote the treatment mean of the selected population. Then μ_S is a random variable and

$$Pr\{\mu_S = \mu_i\} = Pr\{\overline{Y}_i \geq \overline{Y}_j, j \neq i\}, i = 1, \ldots, k.$$

The experimenter not only wishes to select the treatment with the highest mean but also wants an estimate of the mean for the treatment selected. Of course, the "natural" estimator of μ_S is $\overline{Y}_{[k]}$. For $k = 2, \overline{Y}_{[k]}$ is admissible and minimax under the squared error loss. For $k = 2$ and especially for $k > 2$, $\overline{Y}_{[k]}$ is highly unsatisfactory. It is highly positively biased when the μ_i are equal or close. The bias becomes more severe as k increases and, in fact, it tends to infinity as $k \to \infty$. Sarkadi (1967) and Dahia (1974) have studied this problem for $k = 2$ and known common variance σ^2. Hsieh (1981) also discussed the $k = 2$ case but with unknown σ^2. Cohen and Sackrowitz (1982) considered the case $k \geq 3$ with known σ^2. They have given an estimator which is a convex weighted combination of the ordered sample means $\overline{Y}_{[i]}$ where the weights depend on the adjacent differences in the ordered means.

Jeyaratnam and Panchapakesan (1984) discussed estimation after selection associated with the subset selection rule R_4 defined in (2.5) for selecting a subset containing the best treatment. They considered estimating the average worth of the selected subset defined by $M = \sum_{i \in S} \mu_i / |S|$, where S denotes the selected subset and $|S|$ denotes the size of S. For the case of $k = 2$ and known σ^2, Jeyaratnam and Panchapakesan (1984) considered the natural estimator which is positively biased and some modified estimators with reduced bias.

Cohen and Sackrowitz (1988) have presented a decision-theoretic framework for the combined decision problem of selecting the best treatment and estimating the mean of the selected treatment and derived results for the case of $k = 2$ and known σ^2 with common sample size. Gupta and Miescke (1990) extended this study in several directions. They have considered $k > 2$ treatments, different loss components, and both equal and unequal sample sizes. As pointed out by Cohen and Sackrowitz (1988), the decision-theoretic treatment

of the combined selection-estimation problem leads to "selecting after estimation" rather than "estimating after selection".

Estimation after selection is a meaningful problem which needs further study. The earlier papers of Dahia (1974), Hsieh (1981), and Jeyaratnam and Panchapakesan (1984) dealt with several modified estimators in the case of $k = 2$ treatments. These have not been studied in detail as regards their desirable properties. The decision-theoretic results require too many details to provide a comprehensive view. For a list of references, see Gupta and Miescke (1990).

2.5 Estimation of PCS

Consider the selection rule R_1 of Bechhofer (1954) defined in (2.1) for selecting the treatment with the largest mean μ_i. This rule is designed to guarantee that PCS $\geq P^*$ whenever $\mu_{[k]} - \mu_{[k-1]} \geq \delta^*$. However, the true parametric configuration is unknown. If $\mu_{[k]} - \mu_{[k-1]} < \delta^*$, the minimum PCS cannot be guaranteed. Thus a retrospective analysis regarding the PCS is of importance.

For any configuration of $\underline{\mu}$,

$$PCS = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi(t + \frac{\sqrt{n}(\mu_{[k]} - \mu_{[i]})}{\sigma})\phi(t)dt, \tag{2.10}$$

where Φ and ϕ are the standard normal cdf and density function, respectively. Olkin, Sobel and Tong (1976, 1982) considered the estimator $\hat{P}$ obtained by replacing the $\mu_{[i]}$ by $\overline{Y}_{[i]}$ in (2.10). This estimator $\hat{P}$ is consistent, but its evaluation is not easy. Olkin, Sobel and Tong (1976) have given upper and lower bounds for the PCS that hold for any true configuration, with no regard to any least favorable configuration. They have also obtained the asymptotic distribution of $\hat{P}$ which is a function of $\hat{\delta}_i = \overline{Y}_{[k]} - \overline{Y}_{[i]}, i = 1, \ldots, k$; however, the expression for the variance of the asymptotic distribution is complicated.

Faltin and McCulloch (1983) have studied the small-sample performance of the Olkin-Sobel-Tong estimator $\hat{P}$ of the PCS in (2.10), analytically for $k = 2$ populations and via Monte Carlo simulation for $k \geq 2$. They have found that the estimator tends to overestimate PCS (getting worse when $k > 2$) when the means are close together and tends to underestimate when $\sqrt{n}\ \delta/\sigma$ is large.

Anderson, Bishop and Dudewicz (1977) first gave a lower confidence bound for PCS in the case of the selection rule R_1 of Bechhofer (1954) defined in (2.1). Faltin (1980) provided, in the case of $k = 2$ treatments, a quantile unbiased estimator of PCS which can be regarded as a lower confidence bound for PCS. Later Kim (1986) obtained a lower confidence bound on PCS which is sharper than that of Anderson, Bishop and Dudewicz (1977) and reduces to that of Faltin (1980) in the special case of $k = 2$ treatments. Recently, Gupta, Liao, Qiu and Wang (1994), using a new approach, derived a confidence region for the differences $\mu_{[k-i+1]} - \mu_{[k-i]}, i = 1, \ldots, k-1$, and then obtained a lower bound for PCS which is sharper than that of Kim (1986). They also derived some practical lower bounds by reducing the dimensionality of $\underset{\sim}{\delta} = (\delta_1, \ldots, \delta_{k-1})$, where $\delta_i = \mu_{[k-i+1]} - \mu_{[k-i]}, i = 1, \ldots, k-1$. The lower bound improves as this free-to-choose dimensionally q $(1 \leq q \leq k-1)$ increases and the result for $q = 1$ coincides with that of Kim (1986).

Gupta and Liang (1991) obtained a lower bound for PCS by deriving simultaneous lower confidence bounds on $\mu_{[k]} - \mu_{[i]}, i = 1, \ldots, k-1$, where a range statistic was used. Of the two methods of Gupta and Liang (1991) and Gupta, Liao, Qiu, and Wang (1994), one does not dominate the other in the sense of providing larger PCS values. Generally speaking, for moderate k (say $k \geq 5$), the Gupta-Liang method tends to underestimate PCS.

Finally, Gupta and Liang (1991) obtained a lower bound for PCS also in the case of the two-stage procedure of Bechhofer, Dunnett and Sobel (1954) for selecting the treatment with the largest mean when the common variance σ^2 is unknown.

2.6 Notes and Comments

For the rule R_1 of Bechhofer (1954) defined in (2.1), Fabian (1962) has shown that a stronger assertion can be made without decreasing the infimum of the PCS. We define a treatment Π_i to be *good* if $\mu_i \geq \mu_{[k]} - \delta^*$ and modify the goal to be selection of a good treatment. Now a CS occurs if the selected treatment is a good treatment. Then the procedure R_1 guarantees with a minimum probability P^* that the selected treatment is good no matter what the configuration of the μ_i is.

We have not discussed sequential procedures for selecting the best treatment. There is a vast literature available in this regard. Reference can be made to Gupta and Pancha-

pakesan (1991) besides the books mentioned in Section 1.

For the problem of estimating the PCS, Olkin, Sobel and Tong (1976), Kim (1986), Gupta and Liang (1991), and Gupta, Liao, Qiu and Wang (1994) have obtained their general results for location parameters with special discussion of the normal means case. Gupta, Leu and Liang (1990) have discussed the case of truncated location parameter models.

Finally, robustness of selection procedures is an important aspect. This has been examined in the past by a few authors and there is a renewed interest in recent years. A survey of these studies is provided by Panchapakesan (1995a).

3. Selection in Experiments with Blocking: Normal Theory

There may not always be sufficient quantities of homogeneous experimental material available for an experiment using a completely randomized design. However, it may be possible to group experimental units into blocks of homogeneous material. Then one can employ a traditional blocking design which minimizes possible bias and reduces the error variance.

3.1 Indifference-Zone Approach

Assume that there are sufficient experimental units so that each treatment can be used at least once in each block. Consider the *randomized complete block design* with fixed treatment effects, namely,

$$Y_{ij\ell} = \mu + \tau_i + \beta_j + \epsilon_{ij\ell}, \tag{3.1}$$

where $Y_{ij\ell}$ is the ℓth observation $(1 \leq \ell \leq n)$ on treatment i $(1 \leq i \leq k)$ in block $j(1 \leq j \leq b)$. Here μ is the over-all mean, the τ_i are the treatment effects, the β_j are the block effects, and the errors ϵ_{ijk} are assumed to be iid $N(0, \sigma^2)$. It is assumed without loss of generality that $\sum_{i=1}^{k} \tau_i = \sum_{j=1}^{b} \beta_j = 0$. There is no interaction between blocks and treatments.

Let $\tau_{[1]} \leq \ldots \leq \tau_{[k]}$ denote the ordered τ_i. The goal is to select the best treatment, namely, the one associated B with $\tau_{[k]}$. Let us assume that σ^2 is *known*. Then the procedure R_1 of Bechhofer defined in (2.1) for the completely randomized design can easily be adapted

here. We take n independent observations $Y_{ij\ell}(1 \leq \ell \leq n)$ on treatment Π_i $(1 \leq i \leq k)$ in block $j(1 \leq j \leq b)$. The procedure is based on the estimates $\hat{\tau}_i = \overline{Y}_{i..} - \overline{Y}_{...}$, where $\overline{Y}_{i..}$ and $\overline{Y}_{...}$ are the averages of $Y_{ij\ell}$ over the suffixes replaced by dots. The $\hat{\tau}_i$ are the best linear unbiased estimates (BLUE's) of the treatment effects τ_i. Then the adapted procedure is

$$R_6: \quad \text{Select the treatment } \Pi_i \text{ that yields the largest } \hat{\tau}_i. \tag{3.2}$$

In order to guarantee that the PCS is at least P^* whenever $\tau_{[k]} - \tau_{[k-1]} \geq \delta^*$, the minimum sample size n required is given by

$$n = < \frac{2}{b}(\frac{\sigma H}{\delta^*})^2 >, \tag{3.3}$$

where $< x >$ denotes the smallest integer $\geq x$, and H is given by (2.3).

One can modify the basic procedure R_1 defined in (2.1) to select the treatment associated with the largest effect $\tau_{[k]}$ in other cases such as the balanced incomplete block design (BIBD) and Latin Square design. However, when the error variance σ^2 is *unknown*, there is no single stage procedure which will accomplish our goal with a guaranteed PCS. Also, Driessen (1992) and Dourleijn (1993) have discussed in detail selection of the best treatment in experiments using more general types of block designs called *connected designs.*

3.2 Subset Selection Approach

Consider model (3.1) with one observation per cell, which we rewrite as

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \tag{3.4}$$

where Y_{ij} is the observation on treatment $\Pi_i(1 \leq i \leq k)$ in block j $(1 \leq j \leq b)$ and the ϵ_{ij} are iid $N(0, \sigma^2)$. Our goal is to select a non-empty subset containing the treatment associated with $\tau_{[k]}$.

When σ^2 is *known*, we can use the procedure R_4 defined in (2.5), which will be in this model, R_7: Select treatment Π_i if and only if

$$\hat{\tau}_i \geq \hat{\tau}_{[k]} - \frac{d\sigma}{\sqrt{b}}, \tag{3.5}$$

where $\hat{\tau}_i = \overline{Y}_{i.} - \overline{Y}_{..} (1 \leq i \leq k)$ and the constant d is the smallest positive constant for which the PCS $\geq P^*$. This constant $d = \sqrt{2}\, H$, where H is the solution of (2.3).

When σ^2 is *unknown*, we use the procedure R_8 which is R_7 with σ replaced by S_ν, where S_ν^2 is given by

$$S_\nu^2 = \sum_{i=1}^{k} \sum_{j=1}^{b} (Y_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y}_{..})^2 / \nu \tag{3.6}$$

based on $\nu = (k-1)(b-1)$ degrees of freedom. To keep the distinction between R_7 and R_8, we denote the constant needed by d' instead of d. The values of d' (as mentioned in the case of R_5 of Section 2.2) have been tabulated for selected values of k, b, and P^* by Gupta and Sobel (1957). They can also be obtained from the tables of Gupta, Panchapakesan and Sohn (1985) corresponding to correlation coefficient $\rho = 0.5$. Gupta and Hsu (1980) have applied the procedure R_8 and its usual analogue for selecting the treatment associated with $\tau_{[1]}$ to a data set relating to motor vehicle traffic fatality rates (MFR) for the forty-eight contiguous states and the District of Columbia for the years 1960 to 1976. Their goal is to select a subset of best (worst) states in terms of MFR.

As in the case of indifference-zone approach, the basic procedures R_4 and R_5 of Gupta (1956) discussed in Section 2.2 can be adapted for other designs such as the BIBD and Latin Square. Driessen (1992) and Dourleijn (1993) have discussed in detail subset selection in experiments involving connected designs.

3.3 Simultaneous Inference with Respect to the Best

In the model (3.4), let us assume that the error variance σ^2 is unknown. Hsu (1982) gave a procedure for selecting a subset C of the k treatments that includes the treatment associated with $\tau_{[k]}$ and at the same time providing simultaneous upper confidence bounds $D_1, \ldots, D_k$ for $\tau_{[k]} - \tau_1, \ldots, \tau_{[k]} - \tau_k$. His procedure is based on the sample treatment means $\overline{Y}_{i.}$ and S_ν^2 given in (3.6). The procedure R_9 of Hsu (1982) defines

$$C = \{\Pi_i : \overline{Y}_{i.} \geq \max_{j \neq i} \overline{Y}_{j.} - (dS_\nu / \sqrt{b})\} \tag{3.7}$$

and

$$D_i = \max\{\max_{j \neq i} \overline{Y}_{j.} - \overline{Y}_{i.} + (dS_\nu / \sqrt{b})\}, i = 1, \ldots, k, \tag{3.8}$$

where the constant $d = (k, b, P^*)$ is to be chosen so that

$$Pr\{\Pi_{(k)} \epsilon\ C \text{ and } \theta_{[k]} - \theta_{[i]} \leq D_i \text{ for } i = 1, \ldots, k\} = P^*$$

and $\Pi_{(k)}$ is the treatment associated with $\tau_{[k]}$. The constant $d = d(k, b, P^*)$ turns out to be the constant d' of the procedure R_5 and it can be obtained for selected values of k, b, and P^* from the tables of Gupta and Sobel (1957) and Gupta, Panchapakesan and Sohn (1985).

For further detailed treatment of multiple comparisons with and selection of the best treatment, the reader is referred to Driessen (1991, 1992).

3.4 Notes and Comments

Rasch (1978) has discussed selection problems in balanced designs. Wu and Cheung (1994) have considered subset selection for normal means in a two-way design. Given b groups, each containing the same k treatments, their goal is to select a non-empty subset from each group so that the probability of simultaneous correct selection is at least P^*. Dourleijn and Driessen (1993) have discussed four different subset selection procedures for randomized designs with illustrated applications to a plant breeding variety trial.

Gupta and Leu (1987) have investigated an asymptotic distribution-free subset selection procedure for the two-way model (3.4) with the assumption that the $\underset{\sim}{\epsilon}_i = (\epsilon_{i1} \ldots, \epsilon_{i\ell})$ are iid with cdf $F(\underset{\sim}{\epsilon})$ symmetric in its arguments. Their procedure is based on the Hodges-Lehmann estimators of location parameters. For a Bayesian treatment of ranking and selection problems in two-way models, reference should be made to Fong (1990), and Fong and Berger (1993).

Hsu (1982) has discussed simultaneous inference with respect to the best treatment in block designs in more generality than what was described previously. He assumes that the ϵ_{ij} are iid with an absolutely continuous cdf F with some regularity conditions. Besides the procedure R_9 based on sample means discussed previously, he considered two procedures based on signed ranks. Finally, Hsu and Nelson (1993) have surveyed, unified and extended multiple comparisons for the General Linear Model.

4. Selection in Factorial Experiments: Normal Theory

Factorial experimentation when employed in ranking and selection problems can produce considerable savings in total sample size relative to independent single-factor experimentation when the probability requirements are comparable in both cases. This was in fact pointed out by Bechhofer (1954) who proposed a single-stage procedure for ranking normal means when no interaction is present between factor-level effects and common known variance is assumed. In this section, we will be mainly concerned with the two-factor model.

Consider a two-factor experiment involving factors A and B at a and b levels, respectively. The treatment means are μ_{ij} $(1 \leq i \leq a, 1 \leq j \leq b)$ are defined by

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}, \tag{4.1}$$

where μ is the over-all mean, the $\alpha_i (1 \leq i \leq a)$ are the so-called row factor main effects, the $\beta_j (1 \leq j \leq b)$ are the column factor main effects, the $(\alpha\beta)_{ij}$ are the two-way interactions subject to the conditions $\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = 0$, $\sum_{j=1}^{b} (\alpha\beta)_{ij} = 0$ for all i, and $\sum_{i=1}^{a} (\alpha\beta)_{ij} = 0$ for all j. The factors A and B are said to be *additive* if $(\alpha\beta)_{ij} = 0$ for *all* i and j, and to *interact* otherwise. In this section, we discuss selection problems under the indifference-zone as well as subset selection approach both when the factors A and B are additive and interacting. Deciding whether or not the additive model holds is an important problem to be handled with caution. We will be content with just referring to Fabian (1991).

4.1 Indifference-Zone Approach

We will first assume an *additive* model. Independent random samples $Y_{ijm}, m = 1, 2, \ldots,$ are taken from normal treatments Π_{ij} $(1 \leq i \leq a, 1 \leq j \leq b)$ with associated unknown means $\mu_{ij} = \mu + \alpha_i + \beta_j$ and common variance σ^2. Here $a \geq 2$ and $b \geq 2$. The goal is to select the treatment combination associated with $\alpha_{[a]}$ and $\beta_{[b]}$; in other words, we seek simultaneously the best levels of both factors. The probability requirement is:

$$P\{CS|\underset{\sim}{\mu}\} \geq P^* \text{ whenever } \underset{\sim}{\mu} \epsilon\ \Omega_{\alpha,\beta}, \tag{4.2}$$

where $\Omega_{\alpha,\beta} = \{\underset{\sim}{\mu} | \alpha_{[a]} - \alpha_{[a-1]} \geq \delta^*_\alpha, \beta_{[b]} - \beta_{[b-1]} \geq \delta^*_\beta\}$.

For the common *known* σ^2 case, Bechhofer (1954) proposed a single-stage procedure based on n independent observations from each Π_{ij}. Let $\overline{Y}_{i..}$ and $\overline{Y}_{.j.}$ denote the means of the observations corresponding to the levels i and j of the factors A and B, respectively. Then the procedure of Bechhofer (1954) is

$$R_{10}: \text{ Select the treatment combination of levels associated with the largest } \overline{Y}_{i..} \text{ and the largest } \overline{Y}_{.j.} \tag{4.3}$$

The LFC for this procedure is given by

$$\alpha_{[1]} = \ldots = \alpha_{[a-1]} = \alpha_{[a]} - \delta^*_\alpha; \ \beta_{[1]} = \ldots = \beta_{[b-1]} = \beta_{[b]} - \delta^*_\beta. \tag{4.4}$$

The PCS for the rule R_{10} at the LFC can be written as a product of the PCS's at the LFC's when the rule R_1 in (2.1) is applied marginally to each factor. This fact enables one to determine the smallest n to guarantee the minimum PCS.

Bechhofer, Goldsman and Hartmann (1993) have studied the performances of the single-stage procedure R_{10} of Bechhoffer (1954) described previously and two other sequential procedures (not discussed here). One of these is a truncated sequential procedure of Bechhofer and Goldsman (1988b) and the other is a closed sequential procedure with elimination by Hartmann (1993).

The procedure R_{10} can be easily generalized to the cae of r factors with levels $k_1, \ldots, k_r$. One is naturally interested in examining the efficiency of an r-factor experiment relative to that of r independent single-factor experiments in the absence of interaction when both guarantee the same minimum PCS P^*. Let n_f and n_r denote the total numbers of observations required for the r-factor experiment and r single-factor experiments, respectively. Then Bawa (1972) showed that the asymptotic relative efficiency (ARE) of the r-factor experiment, defined by $\text{ARE} = \lim_{p^* \to 1} (n_f/n_r)$, is given by

$$ARE = \frac{\max_{1 \le j \le r} \{k_j/\delta_j^{*2}\}}{\sum_{j=1}^{k} (k_j/\delta_j^{*2})}, \tag{4.5}$$

where δ_j^* is the threshold for the preference-zone for the levels of factor j.

Under the additive model, assuming common known σ^2, procedures can be developed for RCBD and split-plot design experiments. The latter has been studied by Pan and Santner (1993).

When the common variance σ^2 is *unknown*, as in the case of a single-factor experiment, we cannot design a single-stage r-factor experiment that will guarantee the probability requirement. Assuming still an *additive* model, Bechhofer and Dunnett (1986) proposed a two-stage procedure which is a straightforward generalization of the open two-stage non-eliminating procedure of Bechhofer, Dunnett and Sobel (1954). It is also assumed by Bechhofer and Dunnett (1986) that the indifference-zone widths are the same for the factors A and B i.e. $\delta_\alpha^* = \delta_\beta^*$.

When σ^2 is unknown, Pan and Santner (1993) have discussed a procedure for the two-factor additive model using a split-plot design.

Factorial Experiments With Interaction

We now consider the model (4.1) in which $(\alpha\beta)_{ij}$ is not zero for some or all pairs (i,j), and assume that σ^2 is *known*. Dudewicz and Taneja (1982) have discussed selecting the combination of the factor levels with the largest mean in an r-factor experiment. Later, Taneja and Dudewicz (1987) discussed in detail the two-factor case. When the interaction is *known* to exist, their procedure is based on the means $\overline{Y}_{ij.}$ of n independent observations taken from each Π_{ij}. Their procedure is

$$R_{11}: \quad \text{Select the } \Pi_{ij} \text{ that yields the largest } \overline{Y}_{ij.}. \tag{4.6}$$

We note that the rule R_{11} can be used even when there is no interaction in which case, however, R_{10} is the best to use (see Dudewicz and Taneja, 1982). This motivates the procedure with a preliminary test, referred to as R_{12} here, proposed by Taneja and Dudewicz (1987) when there is no prior information regarding the presence of interaction. The preliminary test is the likelihood ratio test for $H_0: (\alpha\beta)_{ij} = 0$ for all pairs (i,j) versus H_1 : Not H_0 at level α. If H_0 is rejected, we proceed by applying the selection rule R_{11} in (4.6); otherwise, we apply the rule R_{10} in (4.3).

Borowiak and De los Reyes (1992) considered the class C of rules R_{12} obtained by

varying α in $[0,1]$. In the special case of $a = b = 2$, they showed that the rule R_{11} (which corresponds to $\alpha = 1$) maximizes the minimum PCS in the class C.

When the common variance σ^2 is unknown, as in the case of single-factor experiments, no single-stage procedure can guarantee the probability requirement. In the case of additive model, Bechhofer (1977) proposed a two-stage procedure analogous to that of Bechhofer, Dunnett and Sobel (1954). The decision rule of this procedure parallels R_{10}. In the case of interaction being present, Taneja (1986) proposed a similar two-stage procedure whose decision rule parallels R_{11}.

For the remainder of this subsection, we consider the model in (4.1) with interaction and let, for convenience, $\gamma_{ij} = (\alpha\beta)_{ij}$. Assume that the common variance σ^2 is known. We consider a new goal, namely, selecting the treatment combination associated with the largest γ_{ij}. Let $\gamma_{[1]} \leq \ldots \leq \gamma_{[ab]}$ denote the ordered interaction effects.

It is required that

$$P(CS|\gamma) \geq P^* \text{ whenever } \gamma_{[ab]} \geq \Delta^* \text{ and } \gamma_{[ab]} - \gamma_{[ab-1]} \geq \delta^*, \tag{4.7}$$

where the event [CS] occurs if and only if the treatment combination corresponding to $\gamma_{[ab]}$ is selected, and the constants δ^*, Δ^* and P^* satisfy

$$\delta^* > 0, \frac{(a-1)(b-1)}{(a-1)(b-1)-1}\delta^* < \Delta^* \text{ and } \frac{1}{ab} < P^* < 1.$$

This problem was first considered by Bechhofer, Santner and Turnbull (1977) in the case of $a = 2$ with an arbitrary b. Their study was expanded by Santner (1981) to arbitrary levels a and b. Take n observations on each treatment combination and let $\hat{\gamma}_{ij} = \overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...}$ $(1 \leq i \leq a, 1 \leq j \leq b)$, where a dot replacing a subscript indicates that an average has been computed over the elements for that subscript. The single-stage procedure of Santner (1981) is

$$R_{13}: \text{ Select the treatment combination that yields the largest } \hat{\gamma}_{ij}. \tag{4.8}$$

The LFC for this procedure has been shown by Santner (1981) to be a solution to a non-linear programming problem. However, the computation of the PCS is difficult. The LFC can be determined more easily in some special cases.

4.2 Subset Selection Approach

Consider the model in (4.1) with no interaction, i.e. $(\alpha\beta)_{ij} = 0$ for all pairs (i,j). We assume that the common variance σ^2 is *unknown*. Our goal is to select a subset of the treatment combinations which contains the combination of levels associated with $\alpha_{[a]}$ and $\beta_{[b]}$. A correct selection occurs if a subset is selected consistent with this goal. Generalizing the single-factor procedure of Gupta (1956), the following procedure was proposed by Bechhofer and Dunnett (1987).

R_{14} : Include in the selected subset all treatment combinations for which

$$\overline{Y}_{i..} \geq \max_{1\leq \ell \leq a} \overline{Y}_{\ell..} - C_A S_\nu/\sqrt{bn}$$

and (4.9)

$$\overline{Y}_{.j.} \geq \max_{1\leq \ell \leq b} \overline{Y}_{.\ell.} - C_B S_\nu/\sqrt{an}.$$

Here n is the number of observations taken on each treatment combination, $\overline{Y}_{i..}$ and $\overline{Y}_{.j.}$ are the appropriate means (as defined for the rule R_{13}), and S_ν^2 is the usual unbiased estimator of σ^2 on $\nu = abn - a - b + 1$ degrees of freedom. The LFC for the rule R_{14} is:

$$\alpha_1 = \ldots = \alpha_a \text{ and } \beta_1 = \ldots = \beta_b. \tag{4.10}$$

For given n, a, b and P^*, equating the PCS at the LFC in (4.10) to P^*, one can solve for C_A and C_B. The solution is not unique. Tables C.1 through C.3 of Bechhofer and Dunnett (1987) give a particular solution $g_1 = C_A\sqrt{2} = C_B\sqrt{2}$ for $P^* = 0.80, 0.90, 0.95$ respectively, for $a = 2(1)5, b = a(1)7, 10$ and $\nu = 5(1)30(5)50, 60(20)120, 200, \infty$.

Pan and Santner (1993) have discussed subset selection procedures under the additive model when a split-plot design is used.

One can consider subset selection procedures when interaction is present. A natural goal to consider is selecting a subset of the ab treatment combinations which contains the one associated with the largest γ_{ij}. If the experimenter is unsure whether or not interaction is present, then the selection can be done by using an appropriate procedure depending on the conclusion of a preliminary test for interaction. These problems have not been studied in detail.

4.3 Notes and Comments

Pan and Santner (1993) have also discussed selection procedures under the indifference-zone approach as well as the subset selection approach for blocked strip-plot experiments.

When interaction is present in (4.1), one may be interested in selecting the treatment combination associated with the largest absolute interaction. Procedures for this goal have been considered by Bechhofer, Santner and Turnbull (1977) and Santner (1981) for completely randomized two-factor experiments.

Bechhofer and Goldsman (1988a) have studied sequential selection procedures for multi-factor experiments involving Koopman-Darmois population under the additivity assumption. Federer and McCulloch (1984, 1993) obtained simultaneous confidence intervals for the sets of differences $\alpha_i - \alpha_{[a]}, i = 1, \ldots, a,$ and $\beta_j - \beta_{[b]}, j = 1, \ldots, b,$ when the experiment is conducted using split-plot, split-split-plot and split-block designs. Taneja (1987) considered nonparametric selection procedures in complete factorial experiments. Finally, reference should be made to Driessen (1992).

5. Selection With Reference to A Standard or A Control: Normal Theory

In the preceding sections we discussed decision procedures for choosing the best from among a given set of $k(\geq 2)$ treatments. Although the experimenter is generally interested in selecting the best one of the competing treatments, in certain situations even the best may not be good enough to warrant its selection. This typically happens in experiments involving comparison of a set of experimental treatments with a standard or a control treatment.

Let $\Pi_1, \ldots, \Pi_k$ be k experimental (normal) treatments with unknown means $\mu_1, \ldots, \mu_k$, respectively. These mean responses are compared with μ_0 which is either a *specified* (known) standard or the *unknown* mean of a control treatment Π_0. When the comparison is with a standard, the decision is based on data from the k experimental treatments. When the comparison involves a control population, we take samples from all the $k+1$ treatments. For convenience, we will refer to the standard also as Π_0.

5.1. Indifference-Zone Approach

Let $\mu_{[1]} \leq \ldots \leq \mu_{[k]}$ denote the ordered means of the k experimental treatments. Our goal here is to select the treatment associated with $\mu_{[k]}$ if $\mu_{[k]} > \mu_0$, or to select Π_0 if $\mu_{[k]} \leq \mu_0$. Any selection consistent with the goal is a correct selection (CS). The probability requirement for any valid procedure is:

$$Pr\{\Pi_0 \text{ is selected }\} \geq P_0^* \text{ whenever } \mu_0 \geq \mu_{[k]} + \delta_0^* B \tag{5.1}$$

and

$$Pr\{\Pi_{(k)} \text{ is selected }\} \geq P_1^* \text{ whenever } \mu_{[k]} \geq \mu_0 + \delta_1^* \text{ and } \mu_{[k]} \geq \mu_{[k-1]} + \delta_2^*, \tag{5.2}$$

where $\delta_0^*, \delta_1^*, \delta_2^*, P_0^*$, and P_1^* are constants with $0 < \{\delta_1^*, \delta_2^*\} < \infty, -\delta_1^* < \delta_0^* < \infty$, $2^{-k} < P_0^* < 1$, $(1 - 2^{-k})/k < P_1^* < 1$, and $\Pi_{(k)}$ denotes the treatment associated with $\mu_{[k]}$.

We assume that the treatments have a common *known* variance σ^2. Let μ_0 be the specified standard. In this case, Bechhofer and Turnbull (1978) proposed a single-stage procedure based on $\overline{Y}_i, i = 1, \ldots, k$, the means of sample of size n from each treatment. Their procedure is

$$\begin{aligned} R_{15}: \ &\text{Choose } \Pi_0 \text{ if } \overline{Y}_{[k]} < \mu_0 + c; \text{ otherwise, choose the population that} \\ &\text{yields } \overline{Y}_{[k]} \text{ as the one associated with } \Pi_{(k)}. \end{aligned} \tag{5.3}$$

Bechhofer and Turnbull (1978) have given simultaneous equations for obtaining (n_e, c) such as the requirements (5.1) and (5.2) are satisfied for the common sample size $n = [n_e] + 1$, where $[x]$ denotes the largest integer $\leq x$.

In many applications, we take $\delta_0^* = 0$ and $\delta_1^* = \delta_2^* = \delta^*$ (say). This is the formulation considered earlier by Paulson (1952). If we let $h = \sqrt{n_e}c/\sigma$ and $g = \sqrt{n_e}\delta^*/\sigma$, then h and g satisfy:

$$\Phi^k(h) = P_0^* \tag{5.4}$$

and

$$\int_{h-g}^{\infty} \Phi^{k-1}(y+g)\phi(y)dy = P_1^* \tag{5.5}$$

where $\Phi(y)$ and $\phi(y)$ denote the standard normal c.d.f. and density function, respectively. The values of h and c satisfying (5.4) and (5.5) are tabulated by Bechhofer and Turnbull (1978) for $k = 2(1)15$ and selected values of P_0^* and P_1^*.

When σ^2 is unknown, Bechhofer and Turnbull (1978) proposed a two-stage sampling procedure which is an analogue of the two-stage sampling procedure of Bechhofer, Dunnett and Sobel (1954). This procedure of Bechhofer and Turnbull (1978) and other early procedures for comparisons with a control are discussed in Gupta and Panchapakesan (1979, Chapter 20).

5.2. Subset Selection Approach

Here our goal is to select a subset of the treatments that includes *all* those treatments which are better than the standard or the control treatment, i.e., all those experimental treatments for which $\mu_i > \mu_0$.

Let us first consider a specified standard μ_0 and assume that the common variance σ^2 is *known*. Any valid rule R is required to satisfy

$$P(CS|R) \geq P^* \text{ for all } \underset{\sim}{\mu} = (\mu_1, \ldots, \mu_k). \tag{5.6}$$

Let $Y_{ij}, j = 1, \ldots, n_i$, be independent sample responses from treatment $\Pi_i (i = 1, \ldots, k)$. Based on the sample means $\overline{Y}_i$, Gupta and Sobel (1958) proposed the rule

$$\begin{aligned} R_{16}: \text{ Include } \Pi_i \text{ in the selected subset if and only if} \\ \overline{Y}_i > \mu_0 - d\sigma/\sqrt{n_i} \end{aligned} \tag{5.7}$$

where $d > 0$ is the smallest number so that the requirement on the PCS can be met, and it is given by $\Phi(d) = (P^*)^{1/k}$.

When σ^2 is *unknown*, we replace σ in (5.7) by S_ν where S_ν^2 is the usual pooled estimator of σ^2 with $\nu = \sum_{i=1}^{k}(n_i - 1)$ degrees of freedom and $n_i \geq 2$ for $1 \leq i \leq k$. In this case, the constant d is given by

$$\int_0^\infty \Phi^k(yd)\, q_\nu(y) dy = P^* \tag{5.8}$$

where $q_\nu(y)$ is the density of $Y = S_\nu/\sigma$. The d-values satisfying (5.8) can be obtained from the tables of Dunnett (1955) for selected values of k, ν, and P^*. It should be noted that d

is the one-sided upper-$(1 - P^*)$ equicoordinate point of the equicorrelated $(k-1)$-variate central t-distribution with the equal correlation $\rho = 0$ and ν degrees of freedom.

Now, let μ_0 be the *known* mean of the control treatment. We assume that all treatments have a common variance σ^2 . Let $\overline{Y}_i, i = 0, 1, \ldots, n$, be the means of random samples of size n_0 from the control treatment and of size n from each of the experimental treatments. When σ^2 is *unknown*, the Gupta-Sobel procedure is

R_{17} : Include Π_i in the selected subset if and only if

$$\overline{Y}_i > \overline{Y}_0 - d\sigma\sqrt{\frac{1}{n} + \frac{1}{n_0}}, \tag{5.9}$$

where the smallest $d > 0$ for which the minimum PCS is guaranteed to be P^* is the solution of

$$Pr\{Z_1 \leq d, \ldots, Z_k \leq d\} = P^*, \tag{5.10}$$

and the Z_i are equicorrelated standard normal variables with equal correlation $\rho = n/(n + n_0)$. The d-values are tabulated by Gupta, Nagel and Panchapakesan (1973) for selected values of k, P^* and ρ. When $n = n_0$, then $d = H$, given by (2.3) with $k - 1$ replaced by k, and thus can be obtained from the tables mentioned in that case.

When σ^2 is *unknown*, we use rule R_{17} with σ replaced by s_ν, where s_ν^2 is the usual pooled estimator of σ^2 based on $\nu = k(n-1) + (n_0 - 1)$ degrees of freedom. In this case, d is the one-sided upper-$(1 - P^*)$ equicoordinate point of the equicorrelated (k)-variate central t-distribution with the equal correlation $\rho = n/(n + n_0)$ and ν degrees of freedom. Values of d are tabulated by Gupta, Panchapakesan and Sohn (1985) and Bechhofer and Dunnett (1988) for selected values of k, P^*, ν and ρ.

5.3. Simultaneous Confidence Intervals

In some applications, the experimenter may be interested in the differences between the experimental treatments and μ_0, which is either a specified standard or the unknown mean of a control treatment. Let us first consider the case of comparisons with a standard μ_0. Assume that the treatments have a common *known* variance σ^2. Let $\overline{Y}_i (i = 1, \ldots, k)$ denote the mean of a random sample of size n_i from Π_i. Define

$$I_i = (\overline{Y}_i - \mu_0 - d\sigma/\sqrt{n_i}, \overline{Y}_i - \mu_0 + d\sigma/\sqrt{n_i}) \tag{5.11}$$

for $i = 1, \ldots, k$, where d is the α-quantile of the standard normal distribution with $\alpha = [1 + (P^*)^{1/k}]/2$. Then

$$Pr\{\mu_i - \mu_0 \in I_i, i = 1, \ldots, k\} \geq P^*. \tag{5.12}$$

When σ^2 is *unknown*, let I_i' be the interval obtained by replacing σ in (5.11) with S_ν, where S_ν^2 is the pooled estimator of σ^2 with $\nu = \sum_{i=1}^{k}(n_i - 1)$ degrees of freedom. In this case, the joint confidence statement (5.12) holds by taking d as the two-sided upper-$(1 - P^*)$ equicoordinate point of the equicorrelated k-variate central t-distribution with the equal correlation $\rho = 0$ and ν degrees of freedom.

When μ_0 is the *unknown* mean of a control treatment Π_0, let $\overline{Y}_0$ be the mean of a random sample from Π_0. When the common variance σ^2 is *unknown*, Dunnett (1955) obtained one-sided and two-sided confidence intervals for $\mu_i - \mu_0, i = 1, \ldots, k$, with joint confidence coefficient P^*. The lower joint confidence limits are given by

$$\overline{Y}_i - \overline{Y}_0 - d_i S_\nu \sqrt{\frac{1}{n_i} + \frac{1}{n_0}}, i = 1, \ldots, k, \tag{5.13}$$

where s_ν^2 is the pooled estimator σ^2 based on $\nu = \sum_{i=0}^{k}(n_i - 1)$ degrees of freedom and the constants d_i are chosen such that

$$Pr\{t_i < d_1, \ldots, t_k < d_k\} = P^* \tag{5.14}$$

where the joint distribution of the t_i is the multivariate t. If $n_1 = \cdots = n_k = n$, then the t_i are equicorrelated with correlation $\rho = n/(n + n_0)$. In this case, $d_1 = \cdots = d_k = d$. For selected values of k, P^*, ν and ρ, the value of d can be obtained from the tables of Gupta, Panchapakesan and Sohn (1985). Dunnett (1955) has tabulated d-values in the case of $n_0 = n_1 = \cdots = n_k$ (i.e. $\rho = 0.5$).

Similar to (5.14), we can write upper confidence limits and two-sided limits. In the equal sample sizes case, Dunnett (1964) has tabulated the constant needed.

When the n_i are unequal, there arises a problem of optimal allocation of observations between the control and the experimental treatments. The optimality is in the sense of maximizing the confidence coefficient for fixed $N = \sum_{i=0}^{k} n_i$. This problem has been studied

by several authors. For a detailed discussion, see Gupta and Panchapakesan (1979, Chapter 20, Section 10).

5.4. Notes and Comments

Chen and Hsu (1992) proposed a two-stage procedure which involves selecting in the first stage the best treatment provided it is better than a control and testing a hypothesis in the second stage between the best treatment selected (if any) at the first stage and the control.

There are studies in which several treatments and a control are administered to the same individuals (experimental units) at different times. The observations collected from the same unit under these treatments are no longer independent. This type of design is called repeated measurements design. Chen (1984) has considered selecting treatments better than a control under such a design.

Bechhofer, Dunnett and Tamhane (1989) have studied two-stage procedures for comparing treatment with a control. In the first stage, they employ the subset selection procedure of Gupta and Sobel (1958) to eliminate "inferior" treatments. In the second stage, joint confidence statement is made for the treatment versus control differences (for those treatments retained after the first stage) using Dunnett's (1955) procedure.

Bechhofer and Tamhane (1981) developed a theory of optimal incomplete block designs for comparing several treatments with a control. They proposed a general class of designs that are balanced with respect to test treatments (BTIB).

Bofinger and Lewis (1992) considered simultaneous confidence intervals for normal treatments versus control differences allowing unknown and unequal treatment variances. Gupta and Kim (1980) and Gupta and Hsiao (1983) have studied subset selection with respect to a standard or control using decision-theoretic and Bayesian formulations. Hoover (1991) generalized the procedure of Dunnett (1955) to comparisons with respect to two controls.

In our discussion of selecting treatments that are better than a standard or control, we assumed that there was no information about the ordering of the treatment means μ_i. In some situations, we may have partial prior information in the form of a simple or

partial order relationship among the unknown means μ_i of the experimental treatments. For example, in experiments involving different dose levels of a drug, the treatment effects will have a known ordering. In other words, we know that $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$ even though the μ_i are unknown. For the goal of selecting all populations for which $\mu_i \geq \mu_0$, we would expect any reasonable procedure R to have the property: If R selects Π_i, then it selects all treatments Π_j with $j > i$. This is the *isotonic* behavior of R. Naturally, such a procedure will be based on the isotonic estimators of the μ_i. Such procedures have been investigated by Gupta and Yang (1984) in the case of normal treatment means allowing the common variance σ^2 to be known or unknown.

6. Selection in Experiments Involving Other Models

Thus far we discussed selection procedures and simultaneous confidence intervals under the assumption that the treatment responses are normally distributed. In this section, we briefly mention some other models for which these problems have been investigated.

6.1 Single-Factor Bernoulli Models.

The Bernoulli distribution serves as an appropriate model in experiments involving manufacturing processes and clinical trials. In these experiments, response variables are qualitative giving rise to dichotomous data such as defective-nondefective or success-failure. Thus we are interested in comparing Bernoulli populations in terms of their success probabilities. The initial and basic contributions to this problem were made by Sobel and Huyett (1957) under the indifference-zone formulation and Gupta and Sobel (1960). There are many interesting aspects of the Bernoulli selection problems. For specification of the preference-zone one can use different measures for the separation between the best and the next best population, namely, $p_{[k]} - p_{[k-1]}, p_{[k]}/p_{[k-1]}$ and $[p_{[k]}(1-p_{[k-1]})]/[(1-p_{[k]})p_{[k-1]}]$. The last measure is the *odds ratio* used in biomedical studies. Besides the usual fixed sample size procedures and purely sequential procedures, the literature includes inverse sampling procedures and so-called Play-the-Winner sampling rules. For a detailed review of these procedures, reference may be made to Gupta and Panchapakesan (1979, 1985).

6.2 Multinomial Models

The multinomial distribution, as a prototype for many practical problems, is a very useful model. When observations from a population are classified into a certain number of categories, it is natural to look for categories that occur very often or rarely. Consider a multinomial distribution on m cells with probabilities $p_1, \ldots, p_m$. Selecting the most and the least probable cells are two common goals. The early investigations of Bechhofer, Elmaghraby and Morse (1959), Gupta and Nagel (1967), and Cacoullos and Sobel (1966) set the pace for a considerable number of papers that followed. The investigations of multinomial selection problems reveal an interesting picture regarding the structure of the LFC which, it turns out, is not similar for the two common goals mentioned previously and also depends on whether a ratio or a difference is used to define the preference-zone. For further discussion and additional references, see Gupta and Panchapakesan (1993).

Although selecting the best cell from a single multinomial population has been investigated over a period of close to forty years, selecting the best of several multinomial populations has not received enough attention until recently except for the paper by Gupta and Wong (1977). For ranking multinomial populations, we need a measure of diversity within a population. Selection procedures have been studied in terms of diversity measures such as Shannon's entropy and the Gini-Simpson index. An account of these procedures is given in Gupta and Panchapakesan (1993).

6.3. Reliability Models

In experiments involving life-length distributions, many specific distributions such as the exponential, Weibull and gamma have been used to characterize the life-length. Panchapakesan (1995b) provides a review of selection procedures for the one- and two-parameter exponential distributions. How the life-length distribution is described as a member of a family characterized in terms of failure rate properties. The IFR (increasing failure rate) and IFRA (increasing failure rate on the average) families are well-known examples of such families. Selection procedures for distributions belonging to such families have been investigated substantially by several authors. A review of these investigations is provided by Gupta and Panchapakesan (1988).

7. Concluding Remarks

As we have pointed out in Section 1, our review of design of experiments with selection and ranking goals covers mainly basic normal theory for single-factor experiments with and without blocking and 2-factorial experiments with and without interaction. We have referred to a few authors who have studied the problem using a Bayesian approach. There have also been a number of investigations under an empirical Bayes approach. Some useful additional references in this connection are: Berger and Deely (1988), Fong (1992), Gupta and Liang (1987) and Gupta, Liang and Rau (1994).

In Section 2.4, we referred to a computer SAS package of Aubuchon, Gupta and Hsu (1986) for implementing simultaneous confidence intervals for the difference between each treatment mean and the best of the other treatment means. This package can also be used for selecting the best treatment using the indifference-zone and the subset selection approaches. There are a few other statistical packages such as CADEMO and MINITAB which contain modules for selection procedures. A commercially distributed package exclusively devoted to selection procedures is RANKSEL; see Edwards (1985, 1986) for details. There are also programs developed by several researchers in the course of their investigations. Rasch (1995) has given a summary of available software for selection procedures with specific description of each. Several FORTRAN programs needed for investigation and implementation of selection procedures are given in the recent book by Bechhofer, Santner and Goldsman (1995).

In the foregoing sections, we have discussed, as alternatives to tests of hypotheses among treatment means, three types of formulations: indifference-zone approach, subset selection approach, and multiple comparisons approach. We have mainly considered single-stage fixed sample size procedures. In some cases we have described two-stage procedures. Sequential procedures have only been referred to. In all these cases, we have not described every available procedure for a given goal. As such we have not gone into efficiency comparisons of competing procedures. However, brief comments have been made in certain cases where a procedure was improved upon or bested by another at a later date. It should however be emphasized that the procedures we have described are viable as yet.

Acknowledgement

This research was supported in part by US Army Research Office Grant DAAH04-95-1-0165.

References

Alam, K. (1970). A two-sample procedure for selecting the populations with the largest mean from k normal populations. *Ann. Inst. Statist. Math.* **22**, 127-136.

Anderson, P. O., T. A. Bishop and E. J. Dudewicz (1977). Indifference-zone ranking and selection: confidence intervals for true achieved $P(CD)$. *Comm. Statist. A - Theory Methods* **6**, 1121-1132.

Aubuchon, J. C., S. S. Gupta and J. C. Hsu (1986). PROC RSMCB: A procedure for ranking, selection, and multiple comparisons with the best. *Proc. SAS Users Group International Conf.* **11**, 761-765.

Bawa, V. S. (1972). Asymptotic efficiency of one R-factor experiment relative to R one-factor experiments for selecting the best normal population. *J. Amer. Statist. Assoc.* **67**, 660-661.

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.

Bechhofer, R. E. (1977). Selection in factorial experiments. In: H. J. Highland, R. G. Sargent and J. W. Schmidt, eds., *Proceedings of the 1977 Winter Simulation Conference*, National Bureau of Standards, Gaitherburg, Maryland, 65-77.

Bechhofer, R. E. and C. W. Dunnett (1986). Two-stage selection of the best factor-level combination in multi-factor experiments: common known variance. In: C. E. McCulloch, S. J. Schwager, G. Casella and S. R. Searle, eds., *Statistical Design: Theory and Practice.* Biometrics Unit, Cornell University, Ithaca, New York, 3-16.

Bechhofer, R. E. and C. W. Dunnett (1987). Subset selection for normal means in multi-factor experiments. *Comm. Statist. Theory Methods* **16**, 2277-2286.

Bechhofer, R. E. and C. W. Dunnett (1988). Percentage points of multivariate Student

t distributions. In: The Institute of Mathematical Statistics, ed., *Selected Tables in Mathematical Statistics, Vol. 11.* American Mathematical Society, Providence, Rhode Island.

Bechhofer, R. E., C. W. Dunnett and M. Sobel (1954). A two-sample multiple-decision procedure for ranking means of normal populations with a common known variance. *Biometrika* **41**, 170-176.

Bechhofer, R. E., C. W. Dunnett and A. C. Tamhane (1989). Two-stage procedures for comparing treatments with a control: elimination at the first stage and estimation at the second stage. *Biometrical J.* **5**, 545-561.

Bechhofer, R. E., S. Elmaghraby and N. Morse (1959). A single-sample multiple decision procedure for selecting the multinomial event which has the highest probability. *Ann. Math. Statist.* **30**, 102-119.

Bechhofer, R. E. and D. M. Goldsman (1987). Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the normal population which has the largest mean. *Comm. Statist. Simulation Comput.* **16**, 1067-1092.

Bechhofer, R. E. and D. M. Goldsman (1988a). Sequential selection procedures for multi-factor experiments involving Koopman-Dormois populations with additivity. In: S. S. Gupta and J. O. Berger, eds., *Statistical Decision Theory and Related Topics*, Springer-Verlag, New York, 3-21.

Bechhofer, R. E. and D. M. Goldsman (1988b). Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the normal population which has the largest mean (II): 2-factor experiments with no interaction. *Comm. Statist. Simulation Comput.* **17**, 103-128.

Bechhofer, R. E. and D. M. Goldsman (1989). A comparison of the performances of procedures for selecting the normal population having the largest mean when the variances are known and equal. In: L. J. Gleser et al., eds., *Contributions to Probability and Statistics: Essays in Honor of Ingram Olkin.* Springer-Verlag, New York, 303-317.

Bechhofer, R. E., D. M. Goldsman and M. Hartmann (1993). Performances of selection procedures for 2-factor additive normal populations with common known variance. In:

F. M. Hoppe, ed., *Multiple Comparisons, Selection, and Applications in Biometry.* Dekker, New York, 209-224.

Bechhofer, R. E., J. Keifer and M. Sobel (1986). *Sequential Identification and Ranking Procedures (with Special Reference to Koopman-Dormois Populations).* The University of Chicago Press, Chicago and London.

Bechhofer, R. E., T. J. Santner and D. M. Goldsman (1995). *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons.* Wiley, New York.

Bechhofer, R. E., T. J. Santner and B. W. Turnbull (1977). Selecting the largest interaction in a two-factor experiment. In: S. S. Gupta and D. S. Moore, eds., *Statistical Decision Theory and Related Topics.* Academic Press, New York, 1-18.

Bechhofer, R. E. and A. C. Tamhane (1981). Incomplete block designs for comparing treatments with a control: general theory. *Technometrics* **23**, 45-57.

Bechhofer, R. E. and B. W. Turnbull (1978). Two $(k+1)$-decision selection procedure for comparing k normal means with a specified standard. *J. Amer. Statist. Assoc.* **73**, 385-392.

Berger, J. and J. J. Deely (1988). A Bayesian approach to ranking and selection of related means with alternatives to analysis-of-variance methodology. *J. Amer. Statist. Assoc.* **83**, 364-373.

Bhandari, S. K. and A. R. Chaudhuri (1990). On two conjectures about two-stage selection problems. *Sankhyā Ser. A.* **52**, 131-141.

Bofinger, E. and G. J. Lewis (1992). Simultaneous comparisons with a control and with the best: two stage procedures (with Discussion). In: E. Bofinger et al., eds., *The Frontiers of Modern Statistical Inference Procedures.* American Sciences Press, Columbus, Ohio, 25-45.

Borowiak, D. S. and J. P. De Los Reyes (1992). Selection of the best in 2×2 factorial designs. *Comm. Statist. Theory Methods* **21**, 2493-2500.

Büringer, H., H. Martin and K.-H. Schriever (1980). *Nonparametric Sequential Selection*

Procedures Birkhauser, Boston, Massachusetts.

Cacoullos, T. and M. Sobel (1966). An inverse-sampling procedure for selecting the most probable event in multinomial distribution. In: P. R. Krishnaiah, ed., *Multivariate Analysis.* Academic Press, New York, 423-455.

Chen, H. J. (1984). Selecting a group of treatments better than a standard in repeated measurements design. *Statist. Decisions* **2**, 63-74.

Chen H. J., E. J. Dudewicz and Y. J. Lee (1976). Subset selection procedures for normal means under unequal sample sizes. *Sankhyā Ser. B* **38**, 2449-255.

Chen, P. and L. Hsu (1992). A two-stage design for comparing clinical trials. *Biometrical J.* **34**, 29-35.

Chen, P. and M. Sobel (1987a). An integrated formulation for selecting the t best of k normal populations *Comm. Statist. Theory Methods* **16**, 121-146.

Chen, P. and M. Sobel (1987b). A new formulation for the multinomial selection problem. *Comm. Statist. Theory Methods* **16**, 147-180.

Cohen, A. and H. B. Sackrowitz (1982). Estimating the mean of the selected population. In: S. S. Gupta and J. O. Berger, eds., *Statistical Decision Theory and Related Topics III, Vol. I.* Academic Press, New York, 243-270.

Cohen, A. and H. B. Sackrowitz (1988). A decision theory formulation for population selection followed by estimating the mean of the selected population. In: S. S. Gupta and J. O. Berger, eds., *Statistical Decision Theory and Related Topics IV, Vol. 2.* Springer-Verlag, New York, 33-36.

Cohen, D. D. (1959). *A Two-Sample Decision Procedure for Ranking Means of Normal Populations with a Common Known Variance.* M.S. Thesis. Dept. of Operations Research, Cornell University, Ithaca, New York.

Dahiya, R. C. (1974). Estimation of the mean of the selected population. *J. Amer. Statist. Assoc.* **69**, 226-230.

Dourleijn, C. J. (1993). *On Statistical Selection In Plant Breeding.* Ph.D. Dissertation. Agricultural University, Wageningen, The Netherlands.

Dourleijn, C. J. and S. G. A. J. Driessen (1993). Subset selection procedures for randomized designs. *Biometrical J.* **35**, 267-282.

Driessen, S. G. A. J. (1991). Multiple comparisons with and selection of the best treatment in (incomplete) block designs. *Comm. Statist. Theory Methods* **20**, 179-217.

Driessen, S. G. A. J. (1992). *Statistical Selection: Multiple Comparison Approach.* Ph.D. Dissertation. Eindhovan University of Technology, Eindhovan, The Netherlands.

Dudewicz, E. J. and S. R. Dalal (1975). Allocation of observations in ranking and selection with unequal variances. *Sankhyā Ser. B* **37**, 28-78.

Dudewicz, E. J. and J. O. Koo (1982). *The Complete Categorized Guide to Statistical Selection and Ranking Procedures.* Series in Mathematical and Management Sciences, Vol. 6, American Sciences Press, Columbus, Ohio.

Dudewicz, E. J. and B. K. Taneja (1982). Ranking and selection in designed experiments: complete factorial experiments. *J. Japan Statist. Soc.* **12**, 51-62.

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.* **50**, 1096-1121.

Dunnett, C. W. (1964). New tables for multiple comparisons with a control. *Biometrics* **20**, 482-491.

Eaton, M. L. (1967). Some optimum properties of ranking procedures. *Ann. Math. Statist.* **38**, 124-137.

Edwards, H. P. (1985). RANKSEL - An interactive computer package of ranking and selection procedures (with discussion). In: E. J. Dudewicz, ed., *The Frontiers of Modern Statistical Inference Procedures.* American Sciences Press, Columbus, Ohio, 169-184.

Edwards, H. P. (1986). The ranking and selection computer package RANKSEL. *Amer. J. Math. Management Sci. 6*, 143-167.

Fabian, V. (1962). On multiple decision methods for ranking population means. *Ann. Math. Statist.* **33**, 248-254.

Fabian, V. (1974). Note on Anderson's sequential procedures with triangular boundary.

Ann. Statist. **2**, 170-176.

Fabian, V. (1991). On the problem of interactions in the analysis of variance (with Discussion), *J. Amer. Statist. Assoc.* **86**, 362-375.

Faltin, F. W. (1980). A quantile unbiased estimator of the probability of correct selection achieved by Bechhofer's single-stage procedure for the two population normal means problem. Abstract. *IMS Bulletin* **9**, 180-181.

Faltin, F. W. and C. E. McCulloch (1983). On the small-sample properties of the Olkin-Sobel-Tong estimator of the probability of correct selection. *J. Amer. Statist. Assoc.* **78**, 464-467.

Federer, W. T. and C. E. McCulloch (1984). Multiple comparison procedures for some split plot and split block designs. In: T. J. Santner and A. C. Tamhane, eds., *Design of Experiments: Ranking and Selection.* Dekker, New York, 7-22.

Federer, W. T. and C. E. McCulloch (1993). Multiple comparisons in split block and split-split plot designs. In: F. M. Hoppe, ed., *Multiple Comparisons, Selection, and Applications in Biometry.* Dekker, New York, 47-62.

Fong, D. K. H. (1990). Ranking and estimation of related means in two-way models - a Bayesian approach. *J. Statist. Comput. Simul.* **34**, 107-117.

Fong, D. K. H. (1992). Ranking and estimation of related means in the presence of a covariance - a Bayesian approach. *J. Amer. Statist. Assoc.* **87**, 1128-1135.

Fong, D. K. H. and J. O. Berger (1993). Ranking, estimation and hypothesis testing in unbalanced two-way additive models - a Bayesian approach. *Statist. Decisions* **11**, 1-24.

Gibbons, J. D., I. Olkin and M. Sobel (1977). *Selecting and Ordering Populations: A New Statistical Methodology.* Wiley, New York.

Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Mimeo. Ser. 150, Institute of Statistics, University of North Carolina, Chapel Hill, North Carolina.

Gupta, S. S. (1963). Probability integrals of the multivariate normal and multivariate t. *Ann. Math. Statist.* **34**, 792-828.

Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. *Technometrics* **7**, 225-245.

Gupta, S. S. and P. Hsiao (1981). On gamma-minimax, minimax, and Bayes procedures for selecting populations close to a control. *Sankhyā Ser. B* **43**, 291-318.

Gupta, S. S. and J. C. Hsu (1980). Subset selection procedures with application to motor vehicle fatality data in a two-way layout. *Technometrics* **22**, 543-546.

Gupta, S. S. and J. C. Hsu (1984). A computer package for ranking, selection, and multiple comparisons with the best. In: S. Sheppard, U. Pooch and C. D. Pegden, eds., *Proceedings of the 1984 Winter Simulation Conference.* Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, 251-257.

Gupta, S. S. and J. C. Hsu (1985). RS-MCB: ranking, selection, and multiple comparisons with the best. *Amer. Statist.* **39**, 313-314.

Gupta, S. S. and D.-Y. Huang (1976). On subset selection procedures for the means and variances of normal populations: unequal sample sizes case. *Sankhyā Ser. A* **38**, 153-173.

Gupta, S. S. and D.-Y. Huang (1981). *Multiple Decision Theory: Recent Developments.* Lecture Notes in Statistics, Vol. 6, Springer-Verlag, New York.

Gupta, S. S. and W.-C. Kim (1980). Gamma-minimax and minimax decision rules for comparison of treatments with a control. In: K. Matusita, ed., *Recent Developments in Statistical Inference and Data Analysis.* North-Holland, Amsterdam, 55-71.

Gupta, S. S. and W.-C. Kim (1984). A two-stage elimination type procedure for selecting the largest of several normal means with a common unknown variance. In: T. J. Santner and A. C. Tamhane, eds., *Design of Experiments: Ranking and Selection.* Dekker, New York, 77-93.

Gupta, S. S. and L.-Y. Leu (1987). An asymptotic distribution-free selection procedure for a two-way layout problem. *Comm. Statist. Theory Methods* **16**, 2313-2325.

Gupta, S. S., L.-Y. Leu and T. Liang (1990). On lower confidence bounds for PCS in truncated location parameter models. *Comm. Statist. Theory Methods* **19**, 527-546.

Gupta, S. S. and T. Liang (1987). On some Bayes and empirical Bayes selection procedures. In: R. Viertl, ed., *Probability and Bayesian Statistics.* Plenum Press, New York, 233-246.

Gupta, S. S. and T. Liang (1991). On a lower confidence bound for the probability of a correct selection: analytical and simulation studies. In: Aydin Öztürk and E. C. van der Meulen, eds., *The Frontiers of Statistical Scientific Theory and Industrial Applications.* American Sciences Press, Columbus, Ohio, 77-95.

Gupta, S. S., T. Liang and R.-B. Rau (1994). Bayes and empirical Bayes rules for selecting the best normal population compared with a control. *Statist. Decisions* **12**, 125-147.

Gupta, S. S., Y. Liao, C. Qiu and J. Wang (1994). A new technique for improved confidence bounds for the probability of correct selection. *Statist. Sinica* **4**, 715-727.

Gupta, S. S. and K. J. Miescke (1984). On two-stage Bayes selection procedures. *Sankhyā Ser. B* **46**, 123-134.

Gupta, S. S. and K. J. Miescke (1990). On finding the largest mean and estimating the selected mean. *Sankhyā Ser. B* **52**, 144-157.

Gupta, S. S. and K. Nagel (1967). On selection and ranking procedures and order statistics from the multinomial distribution. *Sankhyā Ser. B* **29**, 1-34.

Gupta, S. S., K. Nagel and S. Panchapakesan (1973). On the order statistics from equally correlated normal random variables. *Biometrika* **60**, 403-413.

Gupta, S. S. and S. Panchapakesan (1979). *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations.* Wiley, New York.

Gupta, S. S. and S. Panchapakesan (1985). Subset selection procedures: review and assessment. *Amer. J. Math. Management Sci.* **5**, 235-311.

Gupta, S. S. and S. Panchapakesan (1988). Selection and ranking procedures in reliability models. In: P. R. Krishnaiah and C. R. Rao, eds., *Handbook of Statistics 7: Quality Control and Reliability.* North-Holland, Amsterdam, 131-156.

Gupta, S. S. and S. Panchapakesan (1991). On sequential ranking and selection procedures. In: B. K. Ghosh and P. K. Sen, eds., *Handbook of Sequential Analysis.* Dekker, New

York, 363-380. A

Gupta, S. S. and S. Panchapakesan (1993). Selection and screening procedures in multivariate analysis. In: C. R. Rao, ed., *Multivariate Analysis: Future Directions.* North-Holland Series in Statistics and Probability, Vol. 5, Elsevier Science Publishers B. V., Amsterdam, 223-262.

Gupta, S. S., S. Panchapakesan and J. K. Sohn (1985). On the distribution of the studentized maximum of equally correlated normal random variables. *Comm. Statist. Simulation Comput.* **14**, 103-135.

Gupta, S. S. and T. J. Santner (1973). On selection and ranking procedures-a restricted subset selection rule. In: *Proceedings of the 39th Session of the International Statistical Institute*, **45**, Book I, 478-486.

Gupta, S. S. and M. Sobel (1957). On a statistic which arises in selection and ranking problems. *Ann. Math. Statist.* **28**, 957-967.

Gupta, S. S. and M. Sobel (1958). On selecting a subset which contains all populations better than a standard *Ann. Math. Statist.* **29**, 235-244.

Gupta, S. S. and M. Sobel (1960). Selecting a subset containing the best of several binomial populations. In: I. Olkin et al, eds., *Contributions to Probability and Statistics.* Stanford University Press, Stanford, California, 224-248.

Gupta, S. S. and W.-Y. Wong (1977). Subset selection for finite schemes in information theory. In: I. Csisár and P. Elias, eds., *Colloquia Mathematica Societatis János Bolyai, 16: Topics in Information Theory*, 279-291.

Gupta, S. S. and W.-Y. Wong (1982). Subset selection procedures for the means of normal populations with unequal variances: unequal sample sizes case. *Selectra Statistica Canadiana* **6**, 109-149.

Gupta, S. S. and H.-M. Yang (1984). Isotonic procedures for selecting populations better than a control under ordering prior. In: J. K. Ghosh and J. Roy, eds., *Statistics: Applications and New Directions.* Indian Statistical Society, Calcutta, 279-312.

Hall, W. J. (1959). The most economical character of Bechhofer and Sobel decision rules.

Ann. Math. Statist. **30**, 964-969.

Hartmann, M. (1988). An improvement on Paulson's sequential ranking procedure. *Sequential Anal.* **7**, 363-372.

Hartmann, M. (1993). Multi-factor extensions of Paulson's procedures for selecting the best normal population. In: F. M. Hoppe, ed., *Multiple Comparisons, Selection, and Applications in Biometry.* Dekker, New York, 225-245.

Hayter, A. J. (1994). On the selection probabilities of two-stage decision procedures. *J. Statist. Plann. Inference* **38**, 223-236.

Hochberg, Y. and R. Marcus (1981). Three stage elimination type procedures for selecting the best normal population when variances are unknown. *Comm. Statist. Theory Methods* **10**, 597-612.

Hochberg, Y. and A. C. Tamhane (1987). *Multiple Comparison Procedures.* Wiley, New York.

Hoover, D. R. (1991). Simultaneous comparisons of multiple treatments to two (or more) controls. *Biometrical J.* **33**, 913-921.

Hsieh, H.-K. (1981). On estimating the mean of the selected population with unknown variance. *Comm. Statist. A. - Theory Methods* **10**, 1869-1878.

Hsu, J. C. (1982). Simultaneous inference with respect to the best treatment in block designs. *J. Amer. Statist. Assoc.* **77**, 461-467.

Hsu, J. C. (1984). Constrained simultaneous confidence intervals for multiple comparisons with the best. *Ann. Statist.* **12**, 1136-1144.

Hsu, J. C. (1985). A method of unconstrained multiple comparisons with the best. *Comm. Statist. Theory Methods* **14**, 2009-2028.

Hsu, J. C. and B. Nelson (1993). Multiple comparisons in the general linear models. Unpublished Report.

Jeyarantnem, S. and S. Panchapakesan (1984). An estimation problem relating to subset selection for normal populations. In: T. J. Santner and A. C. Tamhane, eds., *Design*

of Experiments: Ranking and Selection. Dekker, New York, 287-302.

Kim, W.-C. (1986). A lower confidence bound on the probability of a correct selection. *J. Amer. Statist. Assoc.* **81**, 1012-1017.

Milton, R. C. (1963). Tables of equally correlated multivariate normal probability integral. Tech. Report No. 27, Department of Statistics, University of Minnesota, Minneapolis, Minnesota.

Mukhopadhyay, N. and T. K. S. Solanky (1994). *Multistage Selection and Ranking Procedures: Second-Order Asymptotics.* Dekker, New York.

Olkin, I., M. Sobel and Y. L. Tong (1976). Estimating the true probability of correct selection for location and scale parameter families. Tech. Report No. 174, Department of Operation Research and Department of Statistics, Stanford University, Stanford, California.

Olkin, I., M. Sobel and Y. L. Tong (1982). Bounds for a k-fold integral for location and scale parameter models with applications to statistical ranking and selection problems. In: S. S. Gupta and J. O. Berger, eds., *Statistical Decision Theory and Related Topics III, Vol. 2.* Academic Press, New York, 193-212.

Pan, G. and T. J. Santner (1993). Selection and screening in additive two-factor experiments using randomization restricted designs. Tech. Report 523, Department of Statistics, The Ohio State University, Columbus, Ohio.

Panchapakesan, S. (1992). Ranking and selection procedures. In: N. Balakrishnan, ed., *Handbook of the Logistic Distribution.* Dekker, New York, 145-167.

Panchapakesan, S. (1995a). Robustness of selection procedures. *J. Statist. Planning Inf.*, to appear.

Panchapakesan, S. (1995b). Selection and ranking procedures. In: N. Balakrishnan and A. P. Basu, eds., *The Exponential Distribution: Method, Theory and Applications.* Gordon and Breach, New York, to appear.

Paulson, E. (1952). On the comparison of several experimental categories with a control. *Ann. Math. Statist.* **23**, 239-246.

Rasch, D. (1978). Selection problems in balanced designs. *Biometrical J.* **20**, 275-278.

Rasch, D. (1995). Software for selection procedures. *J. Statist. Planning Inf.,* to appear.

Santner, T. J. (1975). A restricted subset selection approach to ranking and selection problem. *Ann. Statist.* **3**, 334-349.

Santner, T. J. (1981). Designing two factor experiments for selecting interactions. *J. Statist. Plann. Inference* **5**, 45-55.

Santner, T. J. and A. J. Hayter (1993). The least favorable configuration of a two-stage procedure for selecting the largest normal mean. In: F. M. Hoppe, ed., *Multiple Comparisons, Selection and Applications in Biometry.* Dekker, New York, 247-265.

Sarkadi, K. (1967). Estimation after selection. *Studia Sci. Math. Hungr.* **2**, 341-350.

Sehr, J. (1988). On a conjecture concerning the least favorable configuration of a two-stage selection procedure. *Comm. Statist. Theory Methods* **17**, 3221-3233.

Sobel, M. (1969). Selecting a subset containing at least one of the t best populations. In: P. R. Krishnaiah, ed., *Multivariate Analysis - II.* Academic Press, New York, 515-540.

Sobel, M. and M. J. Huyett (1957). Selecting the best one of several binomial populations. *Bell System Tech. J.* **36**, 537-576.

Tamhane, A. C. (1976). A three-stage elimination type procedure for selecting the largest normal mean (common unknown variance case). *Sankhyā Ser. B* **38**, 339-349.

Tamhane, A. C. and R. E. Bechhofer (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. *Comm. Statist. Theory Methods* **6**, 1003-1033.

Tamhane, A. C. and R. E. Bechhofer (1979). A two-stage minimax procedure with screening for selecting the largest normal mean (II): An improved PCS lower bound and associated tables. *Comm. Statist. Theory Methods* **8**, 337-358.

Taneja, B. K. (1986). Selection of the best normal mean in complete factorial experiments with interaction and with common unknown variance. *J. Japanese Statist. Soc.* **16**, 55-65.

Taneja, B. K. (1987). Nonparametric selection procedures in complete factorial experiments. In: W. Sendler, ed., *Contributions to Statistics.* Physcia-Verlag, Heidelberg, Germany, 214-235.

Taneja, B. K. and E. J. Dudewicz (1987). Selection in factorial experiments with interaction, especially the 2×2 case. *Acta Mathematica Sinica, New Series* **3**, 191-203.

Van der Laan, P. and Verdooren, L. B. (1989). Selection of populations: an overview and some recent results. *Biometrical J.* **31**, 383-420.

Wu, K. H. and S. H. Cheung (1994). Subset selection for normal means in a two-way design. *Biometrical J.* **36**, 165-175.