

Minimax Multiple t-tests for Comparing k Normal  
Populations with a Control\*

by

Shanti S. Gupta  
Purdue University

Klaus J. Miescke  
University of Illinois  
at Chicago

Technical Report #84-44

Department of Statistics  
Purdue University

November 1984

\*This research was supported by the Office of Naval Research Contract  
N00014-84-0167 at Purdue University. Reproduction in whole or in part  
is permitted for any purpose of the United States Government.

Minimax Multiple t-tests for Comparing k Normal  
Populations with a Control

by

Shanti S. Gupta  
Purdue University

Klaus J. Miescke  
University of Illinois  
at Chicago

ABSTRACT

Let  $\pi_1, \dots, \pi_k$  be k normal populations with unknown means  $\theta_1, \dots, \theta_k$ , and a common unknown variance  $\sigma^2 > 0$ . Based on independent samples of sizes  $n_1, \dots, n_k$ , the populations are to be partitioned into two sets, where the first one contains all  $\pi_i$  with  $\theta_i \geq \theta_0$ , and where the other one contains the rest. At first it is assumed that  $\theta_0$  is known. Under an additive "a<sub>i</sub>-b<sub>i</sub>" loss function a minimax procedure is derived which is of a simple natural form. The proof of minimaxity makes use of the Bayes approach and involves a sequence of nonsymmetric priors, which play a similar role as a least favorable prior in simpler problems. Analogous results are presented for the case that  $\theta_0$  is not known. In this case, a control normal population is assumed to exist from which an additional sample of size  $n_0$  can be drawn.

Minimax Multiple t-tests for Comparing k Normal  
Populations with a Control\*

by

Shanti S. Gupta  
Purdue University

Klaus J. Miescke  
University of Illinois  
at Chicago

1. Introduction.

Let  $\pi_1 = N(\theta_1, \sigma^2), \dots, \pi_k = N(\theta_k, \sigma^2)$  be  $k$  normal populations with unknown means  $\theta_1, \dots, \theta_k$  and a common unknown variance  $\sigma^2$ . A population  $\pi_i$  is considered to be "good" if  $\theta_i \geq \theta_0$ , and to be "bad" if  $\theta_i < \theta_0$ ,  $i = 1, \dots, k$ . The control value  $\theta_0$  may either be known or unknown, where in the latter case a control population  $\pi_0 = N(\theta_0, \sigma^2)$  is assumed to be also available. The purpose of this paper is to derive statistical procedures which partition the  $k$  populations into "good" and "bad" ones, respectively, under the minimax criterion.

Let  $\underline{X}_i = (X_{i1}, \dots, X_{in_i})$  be a random sample from  $\pi_i$ ,  $i = 1, \dots, k$ . If  $\theta_0$  is unknown, let  $\underline{X}_0 = (X_{01}, \dots, X_{0n_0})$  be an additional sample from the control population  $\pi_0$ . All samples are assumed to be mutually independent. For notational convenience, let  $\underline{X} = (\underline{X}_1, \dots, \underline{X}_k)$  if  $\theta_0$  is known, and let  $\underline{X} = (\underline{X}_0, \underline{X}_1, \dots, \underline{X}_k)$  if  $\theta_0$  is unknown. In either case, a multiple decision rule can be represented in the form  $\underline{\delta} = (\delta_1, \dots, \delta_k)$ , where, after having observed  $\underline{X} = \underline{x}$ ,  $\delta_i(\underline{x})$  denotes the probability of deciding that " $\pi_i$  is good",  $i = 1, \dots, k$ . Let  $\mathcal{D}$  denote the class of all such rules which are Borel-measurable.

\*Research supported by the Office of Naval Research Contract N00014-84-0167.  
AMS 1980 subject classifications. Primary 62C20; secondary 62F07.  
Key words and phrases. Multiple comparisons with a control, simultaneous t-tests, minimax k-sample tests, Bayes tests, normal location problems.

For a decision theoretic treatment of the problem a loss function has to be specified. Assume that in each  $i$ th component problem a nonnegative loss  $a_i(b_i)$  occurs if  $\pi_i$  is "bad" ("good"), but wrongly classified as "good" ("bad"), and that no loss occurs if the classification is correct,  $i = 1, \dots, k$ . The overall loss is then assumed to be the total sum of these  $k$  losses. Formally, the loss function is thus of the form

$$(1) \quad L(\underline{\theta}, \underline{d}) = \sum_{\substack{i=1 \\ d_i=1}}^k a_i I_{(-\infty, \theta_0)}(\theta_i) + \sum_{\substack{i=1 \\ d_i=0}}^k b_i I_{[\theta_0, \infty)}(\theta_i),$$

where  $\underline{\theta} \in \mathbb{R}^{k+1}$ ,  $\underline{d} \in \{0, 1\}^k$ , and  $d_i = 0(1)$  stands for the decision that  $\pi_i$  is "bad" ("good"),  $i = 1, \dots, k$ .

For the case of  $\theta_0$  known, let  $\underline{\delta}^*$  be the following rule.

$$(2) \quad \delta_i^*(\underline{X}) = 0(1) \text{ iff } n_i^{-1/2}(\bar{X}_i - \theta_0)/S < (>) c_i,$$

where  $S^2$  is the usual unbiased pooled sample estimator of  $\sigma^2$  and  $c_i$  is the lower  $a_i(a_i + b_i)^{-1}$  quantile of a  $t$ -distribution with  $n_1 + \dots + n_k - k$  degrees of freedom,  $i = 1, \dots, k$ .

Analogously, for the case of  $\theta_0$  unknown, let  $\underline{\delta}^{**}$  be given by

$$(3) \quad \delta_i^{**}(\underline{X}) = 0(1) \text{ iff } (n_0^{-1} + n_i^{-1})^{-1/2}(\bar{X}_i - \bar{X}_0)/S < (>) c_i,$$

where  $S^2$  is now derived from  $(\underline{X}_0, \underline{X}_1, \dots, \underline{X}_k)$ , and  $c_i$  is the lower  $a_i(a_i + b_i)^{-1}$  quantile of a  $t$ -distribution with  $n_0 + n_1 + \dots + n_k - k - 1$  degrees of freedom,  $i = 1, \dots, k$ .

The main results to be proved below will confirm that these two procedures are minimax for their associated cases.

The problem of comparing  $k$  normal populations with a control has been considered by many authors. To mention a few of the earlier papers, Paulson (1952), Dunnett (1955), Gupta and Sobel (1958), and Tong (1969) have proposed and studied some natural procedures. Lehmann (1961) and Spjøtvoll (1972) have treated the problem with methods from the theory of testing hypotheses. Randles and Hollander (1971) and Miescke (1981) have derived optimal procedures under the  $\Gamma$ -minimax approach. An overview of this area of research can be found in Gupta and Panchapakesan (1979).

In many of the papers dealing with multiple comparisons with a control, the so-called indifference zones have been adopted, which means that wrong decisions with respect to parameters sufficiently close to  $\theta_0$  do not result in any loss. Thereby, intervals around  $\theta_0$  have to be specified which, together with certain other parameters to be chosen by the experimenter, make the proposed procedures look somewhat complicated.

Our approach to the problem may be more appealing to the experimenter because of its simplicity. There are only  $k$  pairs of losses to be chosen to determine the respective minimax procedure:  $(a_1, b_1), \dots, (a_k, b_k)$ . These losses have a quite natural interpretation which facilitates the experimenter's choice of them. For each  $i = 1, \dots, k$ , the ratio of  $a_i$  and  $b_i$  represents the relative importance of avoiding the two types of possible errors in the  $i$ th component decision problem. After these  $k$  ratios are determined, each pair may still be multiplied by an individual factor. These  $k$  factors may then be chosen in a way to reflect the relative importance of avoiding errors in the  $k$  component decisions.

The method used in this paper to prove minimaxity of  $\underline{\delta}^*$  and  $\underline{\delta}^{**}$  for their respective cases is an asymptotic extension of the standard method, where a procedure is found to be minimax if it is Bayes rule with respect to a least favorable prior. After two technical lemmas are proved in Section 2, minimaxity of  $\underline{\delta}^*$  in the case of a known  $\theta_0$  will be proved in Section 3, and the analogous result for  $\underline{\delta}^{**}$  will be derived in Section 4.

## 2. Two technical lemmas.

These are two main steps in the proof of minimaxity of  $\underline{\delta}^*$  which will be used later. Since they are common for both cases, where  $\theta_0$  is known or unknown, they are presented in this section to avoid repetitions. Also, one may get a fairly clear idea about the proofs to come by just looking at the two lemmas given below.

The first result holds in fact more generally for all  $k$ -decision problems under additive loss. It has been proved in the  $\Gamma$ -minimax approach in Miescke (1981). By allowing  $\Gamma$  to consist of all priors, it can be used also in the minimax approach. For convenience, let us state it below in a form suitable for the present context.

Lemma 1. A decision rule  $\delta^M \in \mathcal{D}$  is minimax if there exists a sequence of priors  $\tilde{p}_m(\underline{\theta}, q)$ ,  $\underline{\theta} \in \mathbb{R}^k$ ,  $q = \sigma^{-2} > 0$ ,  $m = 1, 2, \dots$  such that for every  $i \in \{1, \dots, k\}$  the following holds true: For the  $i$ th component problem there exist Bayes rules  $\delta_{im}^B$  with respect to  $\tilde{p}_m$ ,  $m \in \mathbb{N}$ , such that

$$(4) \quad \sup\{R^{(i)}((\underline{\theta}, q), \delta_i^M) \mid \underline{\theta} \in \mathbb{R}^k, q > 0\} \\ \leq \liminf_{m \rightarrow \infty} r^{(i)}(\tilde{p}_m, \delta_{im}^B),$$

where  $R^{(i)}$  and  $r^{(i)}$  denote the risk function and the Bayes risk, respectively, for the  $i$ th component problem.

Lemma 1 can be used to reduce the  $k$ -decision problem under additive loss to  $k$  individual 1-decision problems, the only common link being the joint priors  $\tilde{p}_m$ ,  $m \in \mathbb{N}$ . As can be anticipated, the second result will now be with respect to a single component problem. Since it may prove to be useful also in other situations, it is given below in a more general form than actually needed in the present context.

Consider the following situation. Let  $Y$  be a sample from a parametric family of probability distributions  $\{P_\theta\}_{\theta \in \mathbb{R}}$ , where we wish to test  $H_0: \theta < \theta_0$  versus  $H_1: \theta \geq \theta_0$ . Let the loss function be  $L(\theta, 1) = L_1(\theta) \geq 0$  if  $\theta < \theta_0$ ,  $L(\theta, 0) = L_2(\theta) \geq 0$  if  $\theta \geq \theta_0$ , and  $L(\cdot, \cdot) = 0$  otherwise. This includes as a special case the 0-1 loss function, where  $L_1 = L_2 = 1$ .

Lemma 2. Let  $\tilde{\pi}$  be a prior density w.r.t. a  $\sigma$ -finite measure  $\mu$  defined on the Borel sets of  $\mathbb{R}$ , such that the following constant  $c$  exists and is not zero:

$$(5) \quad c = \int_{-\infty}^{\theta_0} L_1(\theta) \tilde{\pi}(\theta) d\mu(\theta) + \int_{\theta_0}^{\infty} L_2(\theta) \tilde{\pi}(\theta) d\mu(\theta).$$

Let  $\pi(\theta) = c^{-1} L_1(\theta) \tilde{\pi}(\theta)$  if  $\theta < \theta_0$ , and  $\pi(\theta) = c^{-1} L_2(\theta) \tilde{\pi}(\theta)$  if  $\theta \geq \theta_0$ . Then the Bayes rules under  $L(\cdot, \cdot)$  w.r.t.  $\tilde{\pi}$  coincide with the Bayes rules under the 0-1 loss function w.r.t.  $\pi$ , and the Bayes risks are related to each other through

$$(6) \quad r_L(\tilde{\pi}) = c r_{0,1}(\pi).$$

where the subscript of  $r$  indicates which loss function is assumed.

Proof: Let  $\delta$  be a decision rule and assume, without loss of generality, that it is non-randomized. Under the loss function  $L$ , the Bayes risk of  $\delta$  with respect to a prior  $\tilde{\pi}$ , for which  $c > 0$  exists, is given by

$$\begin{aligned}
 (7) \quad r_L(\tilde{\pi}, \delta) &= \\
 &= \int_{-\infty}^{\theta_0^-} L_1(\theta) P_{\theta}\{\delta(\underline{Y})=1\} \tilde{\pi}(\theta) d\mu(\theta) + \int_{\theta_0}^{\infty} L_2(\theta) P_{\theta}\{\delta(\underline{Y})=0\} \tilde{\pi}(\theta) d\mu(\theta) \\
 &= c \left[ \int_{-\infty}^{\theta_0^-} P_{\theta}\{\delta(\underline{Y})=1\} \pi(\theta) d\mu(\theta) + \int_{\theta_0}^{\infty} P_{\theta}\{\delta(\underline{Y})=0\} \pi(\theta) d\mu(\theta) \right] \\
 &= c r_{0,1}(\pi, \delta),
 \end{aligned}$$

from which the assertions follow immediately.

The above lemma will be applied in Section 3 in the following way. Let  $L_1(\theta) = a$  and  $L_2(\theta) = b$ , respectively. Consider a (normal) prior density  $\pi$  w.r.t. the Lebesgue measure, which is symmetric w.r.t.  $\theta_0$ . Under 0-1 loss, its Bayes rule turns out to be very simple. It will be used later for the  $\delta_{im}^B$ 's in (4). Under the loss function  $L$ , it is also Bayes rule w.r.t. the prior density  $\tilde{\pi}$  given by  $\tilde{\pi}(\theta) = 2b(a+b)^{-1} \pi(\theta)$  if  $\theta < \theta_0$ , and  $\tilde{\pi}(\theta) = 2a(a+b)^{-1} \pi(\theta)$  if  $\theta \geq \theta_0$ . In this case we have  $c = 2ab(a+b)^{-1}$ .

### 3. Known Control $\theta_0$ .

As a natural first step, let us derive the Bayes rules for the given  $k$ -decision problem with respect to the standard family of conjugate priors. Although they are interesting in their own, only the Bayes rule for the case



of  $a_i = b_i$ ,  $i = 1, \dots, k$ , will prove to be useful for the problem under concern. Reconsidering this rule through Lemma 2 as a Bayes rule w.r.t. a non-symmetric prior, it will be used in connection with Lemma 1 to prove minimaxity of  $\underline{\delta}^*$ .

Following DeGroot (1970), ch. 9.6, let  $q = \sigma^{-2}$  be the precision, and let  $\theta_1, \dots, \theta_k$  and  $Q$  denote the random parameters in the Bayes approach, which are assumed to have the following prior density w.r.t. the Lebesgue measure:

$$(8) \quad p(\underline{\theta}, q) = \prod_{i=1}^k p^{(i)}(\theta_i | q) g(q), \quad \underline{\theta} \in \mathbb{R}^k, q > 0,$$

where  $p^{(i)}(\cdot | q)$  is a  $N(\mu_i, (\tau_i q)^{-1})$  density with known  $\mu_i \in \mathbb{R}$  and  $\tau_i > 0$ ,  $i = 1, \dots, k$ , and where

$$(9) \quad g(q) = \beta^\alpha \Gamma(\alpha)^{-1} q^{\alpha-1} e^{-\beta q}, \quad q > 0,$$

is the density of a  $\Gamma$ -distribution with known parameters  $\alpha > 0$  and  $\beta > 0$ .

Standard analysis leads to the following posterior distributions at  $\underline{X} = \underline{x}$ . Given  $Q = q$ ,  $\theta_1, \dots, \theta_k$  are independent  $N((\tau_i \mu_i + n_i \bar{x}_i) (\tau_i + n_i)^{-1}, (q(\tau_i + n_i))^{-1})$ ,  $i = 1, \dots, k$ , and marginally,  $Q$  follows a  $\Gamma$ -distribution with parameters  $\alpha + 2^{-1}n$  and  $\beta'$ , where  $n = n_1 + \dots + n_k$  and

$$(10) \quad \beta' = \beta + 2^{-1} \sum_{i=1}^k \left\{ \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \tau_i n_i (\tau_i + n_i)^{-1} (\bar{x}_i - \mu_i)^2 \right\},$$

and where  $\bar{x}_i$  denotes the sample mean of  $\underline{x}_i$ ,  $i = 1, \dots, k$ .

For  $i \in \{1, \dots, k\}$  fixed, by looking at the posterior joint density of  $\theta_i$  and  $Q$ , it can be seen that the posterior marginal density of  $\theta_i$  is a t-distribution with  $\rho = n + 2\alpha$  degrees of freedom, with location parameter  $(\tau_i \mu_i + n_i \bar{x}_i) (\tau_i + n_i)^{-1}$ , and scale parameter  $\gamma$ , where  $\gamma^2 = 2\beta' (\tau_i + n_i)^{-1} \rho^{-1}$ .

The Bayes rule  $\delta_i^B$  for the  $i$ th component problem can be found by minimizing the associated posterior expected loss. It is given by

$$(11) \quad \delta_i^B(\underline{x}) = 1(0) \text{ iff } P\{\theta_i < \theta_0 | X = \underline{x}\} < (\geq) b_i(a_i + b_i)^{-1}$$

or, by using the results derived above,

$$(12) \quad \delta_i^B(\underline{x}) = 1(0) \text{ iff } (\tau_i \mu_i + n_i \bar{x}_i) (\tau_i + n_i)^{-1} - \theta_0 \geq (<) \gamma e_i,$$

where  $e_i$  is the lower  $a_i(a_i + b_i)^{-1}$  quantile of a t-distribution with  $\rho$  degrees of freedom. Obviously,  $\underline{\delta}^B$  is then the overall Bayes rule for the original  $k$ -decision problem.

For the special case of  $\mu_i = \theta_0$ ,  $a_i = b_i = 1$ ,  $i = 1, \dots, k$ , the Bayes rule turns out to be of a very simple form  $\underline{\delta}^0$ , say, where

$$(13) \quad \delta_i^0(\underline{x}) = 1(0) \text{ iff } \bar{x}_i - \theta_0 \geq (<) 0, \quad i = 1, \dots, k.$$

This is our candidate for the Bayes rules used on the right hand side of (4).

Under the prior distribution given by (8) and (9), assume for a moment that  $Q = q$  is fixed. From the results stated just after (9), it is easy to see that  $\delta_i^0$  is Bayes rule for the  $i$ th component problem, and  $\underline{\delta}^0$  is overall

Bayes rule in this situation, if  $\mu_i = \theta_0$ ,  $a_i = b_i = 1$ ,  $i = 1, \dots, k$ .

Therefore, in view of Lemma 2 and the remarks made at the end of Section 2, it follows that for every  $i \in \{1, \dots, k\}$ ,  $\delta_i^0$  is also Bayes rule for the  $i$ th component problem under the loss function (1) with respect to the prior

$$(14) \quad \tilde{p}(\underline{\theta}|q) = \prod_{i=1}^k 2(a_i+b_i)^{-1} [b_i I_{(-\infty, \theta_0)}(\theta_i) + a_i I_{[\theta_0, \infty)}(\theta_i)] p^{(i)}(\theta_i|q),$$

$\underline{\theta} \in \mathbb{R}^k$ ,  $q > 0$ ,  $\mu_1 = \dots = \mu_k = \theta_0$ . And  $\underline{\delta}^0$  is the overall Bayes rule for the  $k$ -decision problem in this case. Since  $\underline{\delta}^0$  does not depend on  $q$ , the rules are also Bayes rules under the loss function (1) with respect to the unconditional prior

$$(15) \quad \tilde{p}(\underline{\theta}, q) = \tilde{p}(\underline{\theta}|q) g(q), \quad \underline{\theta} \in \mathbb{R}^k, \quad q > 0.$$

Now we are ready to prove the main result of this section.

Theorem 1. Under the loss function (1), the multiple decision rule  $\underline{\delta}^*$ , as given in (2), is minimax. The minimax-value of the problem is equal to

$$\sum_{i=1}^k a_i b_i (a_i + b_i)^{-1}.$$

Proof: Standard arguments show that for every  $i = 1, \dots, k$ , the left hand side of (4) for  $\delta_i^M = \delta_i^*$  is equal to

$$(16) \quad \sup\{R_L^{(i)}((\underline{\theta}, q), \delta_i^*) | \underline{\theta} \in \mathbb{R}^k, q > 0\} = a_i b_i (a_i + b_i)^{-1}.$$

On the right hand side of (4), let  $\underline{\delta}_m^B = \underline{\delta}^0$ , and let  $\tilde{p}_m$  be equal to  $\tilde{p}$  as given in (14) and (15) with  $\tau_1 = \dots = \tau_k = m$ ,  $m \in \mathbb{N}$ . Let  $i \in \{1, \dots, k\}$  be fixed for the rest of the proof. We will show below that

$$(17) \quad \lim_{m \rightarrow \infty} r_L^{(i)}(\tilde{p}_m, \delta_i^0) = a_i b_i (a_i + b_i)^{-1},$$

which clearly completes the proof since under the loss function  $L$  given in (1),  $\delta_i^0$  has been seen to be Bayes rule w.r.t. prior  $\tilde{p}_m$ , for every  $m \in \mathbb{N}$ .

Under the 0-1 loss function,  $\delta_i^0$  has also been seen to be Bayes rule w.r.t. prior  $p_m$ , say, which is equal to  $p$  as given in (8) and (9) with  $\mu_1 = \dots = \mu_k = \theta_0$  and  $\tau_1 = \dots = \tau_k = m$ , and this for every  $m \in \mathbb{N}$ . In this setting, the Bayes risk of  $\delta_i^0$  can be written as

$$(18) \quad r_{0,1}^{(i)}(p_m, \delta_i^0) = \int_0^\infty r_{0,1}^{(i)}(p_m, \delta_i^0 | q) g(q) dq, \text{ say, where } q > 0,$$

$$\begin{aligned} r_{0,1}^{(i)}(p_m, \delta_i^0 | q) &= \\ &= \int_{-\infty}^{\theta_0} \Phi((n_i q)^{1/2}(\theta_i - \theta_0)) (mq)^{1/2} \varphi((mq)^{1/2}(\theta_i - \theta_0)) d\theta_i \\ &+ \int_{\theta_0}^{\infty} \Phi((n_i q)^{1/2}(\theta_0 - \theta_i)) (mq)^{1/2} \varphi((mq)^{1/2}(\theta_i - \theta_0)) d\theta_i \\ &= \int_{-\infty}^0 \Phi((n_i/m)^{1/2} w) \varphi(w) dw + \int_0^\infty \Phi(-(n_i/m)^{1/2} w) \varphi(w) dw, \end{aligned}$$

where  $\varphi$  and  $\Phi$  denote the standard normal density and cumulative distribution function, respectively. Clearly for every  $q > 0$ , the sum of the last two

integrals tends to  $1/2$  as  $m$  tends to infinity. Since the value of this sum is always between 0 and 1, uniformly in  $q > 0$  and  $m \in \mathbb{N}$ , it follows by Lebesgues dominated convergence theorem that

$$(19) \quad \lim_{m \rightarrow \infty} r_{0,1}^{(i)}(p_m, \delta_i^0) = 1/2.$$

Applying now Lemma 2, in the way described below of (13), we get

$$(20) \quad r_L^{(i)}(\tilde{p}_m, \delta_i^0) = 2 a_i b_i (a_i + b_i)^{-1} r_{0,1}^{(i)}(p_m, \delta_i^0).$$

From this it follows that (17) holds, and therefore the proof of the theorem is completed.

It should be pointed out that Lehmann (1957) has shown that the minimax-value of the  $i$ th component problem is equal to  $a_i b_i (a_i + b_i)^{-1}$ ,  $i = 1, \dots, k$ . Therefore from (16) it follows that  $\delta_i^*$  is minimax for the  $i$ th component problem,  $i = 1, \dots, k$ . It is a well known fact that student's  $t$ -test is minimax at the suitably chosen level of significance. However, this fact is of no use in the present context, since the overall minimax value may be less than the sum of the  $k$  minimax values of the  $k$  component problems.

As a final remark, let us mention that  $\underline{\delta}^*$  remains minimax if  $S^2$ , the pooled sample estimator of  $\sigma^2$ , is based on a subcollection of observations from  $\underline{X}$ , and if  $c_1, \dots, c_k$  are properly adapted. However, such a modified procedure would have a strictly larger risk, except at  $\theta_1 = \dots = \theta_k = \theta_0$ . This follows from the fact that for every  $i \in \{1, \dots, k\}$ ,  $\delta_i^*$  is the uniformly most powerful unbiased test at its level, whereas the modified procedures'

ith decision rule would only be an unbiased test at the same level of  $b_i(a_i+b_i)^{-1}$ . The modified procedure would thus be inadmissible. Whether or not  $\underline{\delta}^*$  is admissible remains an open question.

#### 4. Unknown Control $\theta_0$ .

In this setting, an additional sample  $X_0$  from population  $\pi_0 = N(\theta_0, \sigma^2)$  is observed. The analogous results to Section 3 can be derived in a similar way. Therefore the treatment of this case will be rather concise.

First, let us find the Bayes rules w.r.t. the standard family of conjugate priors which is essentially the same as (8) and (9), but now with the product in (8) defined over the range  $i = 0, 1, \dots, k$ , since  $\theta_0$  is now an additional random parameter. Of course,  $p^{(0)}(\cdot | q)$  is a  $N(\mu_0, (\tau_0 q)^{-1})$  density with known  $\mu_0 \in \mathbb{R}$  and  $\tau_0 > 0$ . From the results derived below of (9) it can be seen that the posterior distribution at  $\underline{X} = \underline{x}$  has the following properties. For every  $i = 1, \dots, k$ , given  $Q = q$ ,  $\theta_i - \theta_0$  has, marginally, a normal distribution with mean  $\Delta_{i0} = (\tau_i + n_i)^{-1}(\tau_i \mu_i + n_i \bar{x}_i) - (\tau_0 + n_0)^{-1}(\tau_0 \mu_0 + n_0 \bar{x}_0)$  and variance  $q^{-1}((\tau_i + n_i)^{-1} + (\tau_0 + n_0)^{-1})$ . The posterior marginal distribution of  $Q$  is a  $\Gamma$ -distribution with parameters  $\alpha + 2^{-1}(n_0 + n)$  and  $\beta''$ , where  $\beta''$  is the analog to  $\beta'$  given in (10), where the first sum is now defined over the range  $i = 0, 1, \dots, k$ .

For  $i \in \{1, \dots, k\}$  fixed, by looking at the posterior joint density of  $\theta_i - \theta_0$  and  $Q$ , it can be seen that the posterior marginal density of  $\theta_i - \theta_0$  is a t-distribution with  $n_0 + n + 2\alpha$  degrees of freedom with location parameter  $\Delta_{i0}$  and scale parameter  $\gamma_0$ , where  $\gamma_0^2 = 2\beta''((\tau_i + n_i)^{-1} + (\tau_0 + n_0)^{-1})(n_0 + n + 2\alpha)^{-1}$ .

For the  $i$ th component problem the Bayes rule can be found by minimizing the associated posterior expected loss. It is given by

$$(21) \quad \delta_i^B(\underline{x}) = 1(0) \text{ iff } P\{\theta_i \leq \theta_0 | X = \underline{x}\} < (>) b_i(a_i + b_i)^{-1}$$

or, by using the results derived above,

$$(22) \quad \delta_i^B(\underline{x}) = 1(0) \text{ iff } \Delta_{i0} \geq (<) \gamma_0 e_{i0},$$

where  $e_{i0}$  is the lower  $a_i(a_i + b_i)^{-1}$  quantile of a t-distribution with  $n_0 + n + 2\alpha$  degrees of freedom.

For the special case of  $\mu_i = \mu_0$ ,  $n_0\tau_i = n_i\tau_0$ ,  $a_i = b_i$ ,  $i = 1, \dots, k$ , the Bayes rule turns out to be of the simple form  $\underline{\delta}^{00}$ , say, where

$$(23) \quad \delta_i^{00}(\underline{x}) = 1(0) \text{ if } \bar{X}_i - \bar{X}_0 \geq (<) 0, \quad i = 1, \dots, k.$$

Instead of following along the lines below of (13), there is a shorter way to prove minimaxity of  $\underline{\delta}^{**}$  in the present case. The main result of this section is

Theorem 2. Under the loss function (1), the multiple decision rule  $\underline{\delta}^{**}$ , as given in (3), is minimax. The minimax-value of the problem is equal to

$$\sum_{i=1}^k a_i b_i (a_i + b_i)^{-1}.$$

Proof. Again, standard arguments show that for every  $i = 1, \dots, k$ , the left hand side of (4) for  $\delta_i^M = \delta_i^{**}$  is equal to

$$(24) \quad \sup\{R_L^{(i)}((\underline{\theta}, q), \delta_i^{**}) | \underline{\theta} \in \mathbb{R}^{k+1}, q > 0\} = a_i b_i (a_i + b_i)^{-1},$$

where the dimension of the  $\underline{\theta}$ -parameter space is now  $k+1$ .

On the right hand side of (4), instead of choosing  $\delta_m^B$  to be  $\delta^{00}$ , let rather  $\delta_m^B = \delta^0$  as before in the proof of Theorem 1. As to the priors of  $\theta_0, \theta_1, \dots, \theta_k$ , and  $Q$ , assume that  $\theta_0 = \theta_0$  be a fixed known constant  $\theta_0$ , say, and adopt the same priors for  $\theta_1, \dots, \theta_k$ , and  $Q$  as have been used in the proof of Theorem 1. Then for every  $i = 1, \dots, k$ , (17) holds true and  $\delta_i^0$  is Bayes rule with respect to prior  $\tilde{p}_m$  for all  $m \in \mathbb{N}$ . Therefore the proof of Theorem 2 is essentially the same as the proof of Theorem 1.

### Concluding Remarks:

The remarks given at the end of Section 3 hold in an analogous form for the situation considered above. They are omitted for brevity.

For the proofs of the two theorems, the proper choice of priors was crucial. The relevant parameters  $\theta_i$  were assumed to be independent, whenever the nuisance parameter  $Q = q$  was fixed. In the unknown control case, an attempt to use the principle of (location) invariance may not lead to the desired results if one assumes that, a priori,  $\theta_1 - \theta_0, \dots, \theta_k - \theta_0$  are independent. This is due to the fact that at  $\underline{X} = \underline{x}$ , the posterior distribution of each  $\theta_i - \theta_0$  would depend on all given observations. For the case of  $\sigma^2$  known, Randles and Hollander (1971) have given an instructive example.



References

- [1] DeGroot, M. H. (1970). Optimal Statistical Decisions. McGraw-Hill, New York.
- [2] Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. J. Amer. Statist. Assoc. 50, 1096-1121.
- [3] Gupta, S. S. and Panchapakesan, S. (1979). Multiple Decision Procedures. Wiley, New York.
- [4] Gupta, S. S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. Ann. Math. Statist. 29, 235-244.
- [5] Lehmann, E. L. (1957). A theory of some multiple decision problems. II. Ann. Math. Statist. 28, 547-572.
- [6] Lehmann, E. L. (1961). Some Model I problems of selection. Ann. Math. Statist. 32, 990-1012.
- [7] Miescke, K. J. (1981).  $\Gamma$ -minimax selection procedures in simultaneous testing problems. Ann. Statist. 9, 215-220.
- [8] Paulson, E. (1952). On the comparison of several experimental categories with a control. Ann. Math. Statist. 23, 239-246.
- [9] Randles, R. H. and Hollander, M. (1971).  $\Gamma$ -minimax selection procedures in treatments versus control problems. Ann. Math. Statist. 42, 330-341.
- [10] Spjøtvoll, E. (1972). On the optimality of some multiple comparison procedures. Ann. Math. Statist. 43, 398-411.
- [11] Tong, Y. L. (1969). On partitioning a set of normal populations by their locations with respect to a control. Ann. Math. Statist. 40, 1300-1324.

BEFORE COMPLETING FORM		
1. REPORT NUMBER Technical Report #84-44	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Minimax Multiple t-tests for Comparing k Normal Populations with a Control	5. TYPE OF REPORT & PERIOD COVERED Technical	
	6. PERFORMING ORG. REPORT NUMBER Technical Report #84-44	
7. AUTHOR(s) Shanti S. Gupta and Klaus J. Miescke	8. CONTRACT OR GRANT NUMBER(s) N00014-84-0167	
	9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Department of Statistics West Lafayette, IN 47907	
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, DC	10. PROGRAM ELEMENT, PROJECT, TASK, AREA & WORK UNIT NUMBERS	
	12. REPORT DATE October 1984	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	13. NUMBER OF PAGES 15	
	15. SECURITY CLASS. (of this report) Unclassified	
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release, distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Multiple comparisons with a control, simultaneous t-tests, minimax k-sample tests, Bayes tests, normal location problems.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Let $\pi_1, \dots, \pi_k$ be k normal populations with unknown means $\theta_1, \dots, \theta_k$ , and a common unknown variance $\sigma^2 > 0$ . Based on independent samples of sizes $n_1, \dots, n_k$ , the populations are to be partitioned into two sets, where the first one contains all $\pi_i$ with $\theta_i \geq \theta_0$ , and where the other one contains the rest. At first it is assumed that $\theta_0$ is known. Under an additive " $a_i - b_i$ " loss function a minimax		

procedure is derived which is of a simple natural form. The proof of minimaxity makes use of the Bayes approach and involves a sequence of nonsymmetric priors, which play a similar role as a least favorable prior in simpler problems. Analogous results are presented for the case that  $\theta_0$  is not known. In this case, a control normal population is assumed to exist from which an additional sample of size  $n_0$  can be drawn.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)