SELECTION PROCEDURES FOR THE t BEST POPULATIONS*

by

Raymond J. Carroll
Department of Statistics
University of North Carolina at Chapel Hill

Shanti S. Gupta
Purdue University

Deng-Yuan Huang
Purdue University and
Academia Sinica, Taiwan

Institute of Statistics Mimeo Series No. 993

April, 1975

* This research was supported by the Office of Naval Research, contract N00014-67-A-0226-00014 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

SELECTION PROCEDURES FOR THE t BEST POPULATIONS*

Raymond J. Carroll
University of North Carolina at Chapel Hill

Shanti S. Gupta
Purdue University

and

Deng-Yuan Huang
Purdue University and
Academia Sinica, Taiwan

SUMMARY

Suppose one has k populations and wishes eventually to choose the "best" one. It is often desirable to use a multi-stage experiment, where at each stage, at least t (≥1) populations are selected for further study. This paper is concerned with selecting at least t(≥2) populations, presenting generalizations of two methods in existence for t = 1. One method has as its goal the selection of good populations, while the other eliminates bad populations. Two special cases considered are the problems of selecting large normal means or small normal variances, and tables are constructed for implementing the procedures. In the case of normal variances, the tables are useful in the classical indifference zone problem. Applications are presented in a variety of fields.

*This research was supported by the Office of Naval Research, contract N00014-67-A-0226-00014 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

1. Introduction

In many problems where the eventual goal is to select the "best" population, the evaluation criteria may not be strictly numerical. Examples (see also Section 4) include problems of choosing industrial components (where time and availability of materials are also considered), deciding on an effective drug to reduce tumor size (where side effects are an important factor), and choosing students for admission to a professional school (where the subjective personal interview plays an important part). A possible method for solving such problems is to select a subset of the populations on the basis of a numerical criterion for further study of their other effects. For want of a better term, these other effects will be called subjective effects. Note that one will usually need to study more than one population after the first stage of the experiment since it is entirely possible that the population which ranks highest on the numerical score may be unacceptable. Beside the examples mentioned above, another illustration would be an industrial process which is unacceptable in that it causes excessive amounts of pollution. By choosing more than one population for further study, one will thus often save substantial amounts of time.

In this paper we will have k populations and the goal will be to select a subset of these populations of size at least $t(\geq 1)$. Two approaches will be studied. In Section 2 a procedure of Desu (1970) for $t = 1$ which eliminates "bad" populations will be generalized. In Section 3, Gupta's rule (1965), which selects "good" populations, will also be generalized. Tables will be given for implementing these procedures in the special cases of normal means and normal variances; the tables for the latter will be useful in computing the probability of a correct selection for the t smallest normal variances using Bechhofer's (1954) indifference zone approach. Other work related to this paper include results of Gupta and Sobel (1962), Deverman (1969), Deverman and Gupta (1969), Panchapekesan (1969), and Santner (1974).

2. Eliminating the "Non t Best" Populations.

Consider a problem where one has k drugs and wishes (i) to select at least t (≥ 1) of these for further study and (ii) to make sure "ineffective" drugs are eliminated. One possible method, introduced by Desu (1970) for the case $t = 1$, is to eliminate all populations which are not "sufficiently close" to the t best populations. To be precise, suppose that for $i = 1,\ldots,k$, we have a random variable X_i (usually a sufficient statistic) from the population π_i with distribution $F(x-\theta_i)$, where F is known. Denote by $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ the true (unknown) ordering of the parameters. We say that π_i is *strictly non-t-best* if $\theta_{[k-t+1]} - \theta_i \geq \Delta > 0$, where Δ is chosen by the experimenter. A correct decision (CD) is defined to be the selection of a subset of $\pi_1,\ldots,\pi_k$ which excludes all strictly non-t-best populations. We use the rule R_1, which reduces to Desu's rule (1970) when $t = 1$:

$$R_1\text{: Reject } \pi_i \text{ if } X_i < X_{[k-t+1]} - (\Delta - d_1) \quad (0 < d_1 < \Delta),$$

where $X_{[1]} \leq \ldots \leq X_{[k]}$ is the ordered sample. Letting Ω denote the parameter space, the problem reduces, for given values of P^* and Δ, to finding an admissible value d_1 such that $P_\theta(CD|R_1) \geq P^*$ if $\underline{\theta} = (\theta_1,\ldots,\theta_k) \in \Omega$.

Lemma 2.1. For the rule R_1, if $t \leq k-t$,

$$\inf_{\Omega} P_{\underline{\theta}}(CD|R_1) = t \int_{-\infty}^{\infty} F^{k-t}(x+d_1)[1-F(x)]^{t-1} dF(x), \tag{2.1}$$

so that the required value of d_1 is the smallest number d_1 (less than Δ) for which the expression in (2.1) exceeds P^*.

Proof. We may assume $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k$. Define (for $m = 0,1,\ldots,k-t$)

$$\Omega_m = \{\underline{\theta}: \theta_1 \leq \ldots \leq \theta_m \leq \theta_{k-t+1} - \Delta < \theta_{m+1} \leq \ldots \leq \theta_k\}. \tag{2.2}$$

Then an argument similar to that in Desu (1970) shows that $P_{\underline{\theta}}(CD|R_1)$ attains its minimum on Ω_m when $\theta_1 = \ldots = \theta_{k-t} = \theta_{k-t+1} - \Delta = \ldots = \theta_k - \Delta$. This means that on Ω_m,

$$P(CD|R_1) \geq \Pr\{\max_{i \leq j \leq m} X_j < X_{[k-t+1]} - (\Delta - d_1)\}, \tag{2.3}$$

the last probability being under the above parameter configuration. Since this configuration does not change as m increases, the infimum is attained when $m = k-t$, in which case (2.3) reduces to (2.1).

Corollary 2.1. Suppose F is the normal distribution with mean 0 and known variance σ^2, and that samples of size n have been taken from each population. If $d_1 = d\sigma/\sqrt{n}$ and rule R_1 is used,

$$\inf_{\Omega} P(CD|R_1) = t \int_{-\infty}^{\infty} \Phi^{k-t}(x+d)[1-\Phi(x)]^{t-1} d\Phi(x) .$$

The values of d are tabulated in Bechhofer (1954) for various k,t, and P^*. Note that n must be large enough so that $d\sigma/\sqrt{n} < \Delta$.

In the scale parameter case, where the populations have distributions $F(x/\theta_1), \ldots, F(x/\theta_k)$, one might be interested in small scale parameters. For example, one might wish to select that one of several industrial processes with the smallest variation. A population π_i will be said to be strictly non-t-best in this situation if $\theta_{[t]} < \Delta\theta_i$ $(0 < \Delta < 1)$. In analogy with rule R_1, we use:
R_2: Eliminate π_i if $X_i > d_2 X_{[t]}/\Delta$ $(\Delta < d_2 < 1)$.

Lemma 2.2. Using the rule R_2,

$$\inf_{\Omega} P(CD|R_2) = t \int_0^{\infty} [1-F(xd_2)]^{k-t} F^{t-1}(x) dF(x). \tag{2.4}$$

As in the case of Lemma 2.1, the above integral is the infimum of the P(CD) for selecting the smallest scale parameter using the indifference zone approach of Bechhofer (1954). When one is interested in the smallest normal variance, F in (2.4) becomes the distribution of a chi-square random variable with $r = n-1$ degrees of freedom. Table 1 gives values of (2.4) in this case for $r = 2(2)8$ and various values of d_2, k, t. These tables will be useful not only for the problems of this paper but also for the classical indifference zone problem.

3. Selecting a Subset Containing the t Best Populations.

The procedures of this section have as their goal the selection of at least t populations, and are thus an analogue of the subset selection procedures of Gupta (1965). An example of the type of problem of interest here is the problem of student selection for a professional school, where at least t students are to be selected.

We again first consider the case where the populations $\pi_1,\ldots,\pi_k$ have distributions $F(x-\theta_1),\ldots,F(x-\theta_k)$ where F is known. A correct decision (CD) is defined to be the selection of a subset of size at least t which includes the t populations giving rise to $\theta_{[k-t+1]},\ldots,\theta_{[k]}$. The rule R_3 used here reduces to that of Gupta (1965) when $t = 1$.

R_3: Select π_i if $X_i \geq X_{[k-t+1]} - d_3 \quad (d_3 \geq 0)$.

Letting Ω denote the parameter space, the problem is to find, for given P^*,k,t the value d_3 for which $P_{\underline{\theta}}(CD|R_3) \geq P^*$ if $\underline{\theta}$ belongs to Ω .

In general in ranking theory, one attempts to find explicitly a "least favorable configuration" $\underline{\theta}_0$ at which $P_{\underline{\theta}}(CD|R_3)$ attains its infimum on Ω , as for example in Section 2. For R_3, we have not been able to find a least favorable configuration, but, as a first step, we present a lower bound.

Lemma 3.1. For the rule R_3,

$$\text{(3.1)} \qquad \inf_{\Omega} P_{\underline{\theta}}(CD|R_3) \geq t\int_{-\infty}^{\infty} F^{k-t}(x+d_3)[1-F(x)]^{t-1}dF(x) \ .$$

Proof. Again assume $\theta_1 \le \ldots \le \theta_k$. Let

$$A = \{\min(X_{k-t+1}, \ldots, X_k) \ge \max(X_1, \ldots, X_k) - d_3\} .$$

Then A is a subset of the set of correct decisions, so that $P_{\underline{\theta}}(CD|R_3) \ge P_{\underline{\theta}}(A|R_3)$. From Bechhofer (1954), this last probability has as its infimum the right hand side of (3.1). Note that (3.1) is precisely the same as (2.1).

Corollary 3.1. Suppose F is the normal distribution with mean 0 and known variance σ^2. If $d_3 = d\sigma/\sqrt{n}$ and rule R_3 is used,

(3.2) $$\inf_{\Omega} P(CD|R_3) \ge t \int_{-\infty}^{\infty} \Phi^{k-t}(x+d)[1-\Phi(x)]^{t-1} d\Phi(x) .$$

Table 2 presents a Monte-Carlo study of the bound (3.2) for $P^* = .90, .95$, $k = 2(2)10$ and the configurations

(3.3) $$\theta_{[1]} = \ldots = \theta_{[k-t]} = \theta_{[k-t+1]} - \delta = \ldots - \theta_{[k]} - \delta ,$$

(3.4) $$\theta_{[i+1]} - \theta_{[i]} = \delta \quad (i=1, \ldots, k-1) ,$$

and for $\delta = 0(1)4$. Even when $\delta = 0$ (that is, when the populations are identical) the bounds are conservative for $t \ge 2$ and $E(S|R_3)$, the expected number of populations selected, tends to be high for (3.3) but behaves well for (3.4).

Because of the conservative nature of the bound (3.2) further study is necessary. If $t = 1$, the least favorable configuration for the location parameter case is the slippage configuration (3.3) with $\delta = 0$. This is *not* true for $t \ge 2$ as is seen in the discussion after Lemma 3.2; however, it is possible to derive the least favorable configuration under the assumption that $\theta_{[k-t+1]} = \ldots = \theta_{[k]}$. Letting $\Omega^* = \{(\theta_1, \ldots, \theta_k) : \theta_{[k-t+1]} = \ldots = \theta_{[k]}\}$, the rule R_3 attains its least favorable configuration over Ω^* in the slippage configuration (3.3) with $\delta = 0$.

Not only is this an important special case in the applications (corresponding to the idea that the good populations are substantially the same), but it also yields a rule for general use which should be very efficient.

Lemma 3.2. If $\Omega^*(\delta)$ is the subset of Ω^* for which $\theta_{[k-t+1]} - \theta_{[k-t]} \geq \delta$, the rule R_3 attains its infimum over $\Omega^*(\delta)$ at the configuration (3.3), and in this configuration (3.3),

$$P(CD|R_3) = t\sum_{j=0}^{t-1}\binom{t-1}{j}\binom{k-t}{t-j-1}A_1(k,t,j) + (k-t)\sum_{j=0}^{t-1}\binom{t}{j}\binom{k-t-1}{t-j-1}A_2(k,t,j)\ , \tag{3.5}$$

where

$$A_1(k,t,j) = \int_{-\infty}^{\infty}[1-F(x)]^j[1-F(x+\delta)]^{t-j-1}[F(x)-F(x-d_3)]^{t-j-1}F^{k-2t+j+1}(x+\delta)dF(x) \tag{3.6}$$

$$A_2(k,t,j) = \int_{-\infty}^{\infty}[1-F(x-\delta)]^j[F(x-\delta)-F(x-\delta-d)]^{t-j}[1-F(x)]^{t-j-1}F^{k-2t+j}dF(x)\ . \tag{3.7}$$

If $\delta = 0$, $A_2(k,t,j) = A_1(k,t,j-1)$, so that in this case

$$P(CD|R_3) = \sum_{j=1}^{t-1}\left\{t\binom{t-1}{j-1}\binom{k-t}{t-j} + (k-t)\binom{t}{j}\binom{k-t-1}{t-j-1}\right\}A_2(k,t,j) + tA_2(k,t,t) + (k-t)\binom{k-t-1}{t-1}A_2(k,t,0)\ . \tag{3.8}$$

Proof. The following decomposition holds:

$$P(CD|R_3) = t\sum_{j=0}^{t-1}\Pr\left\{\begin{matrix}\text{CD and } X_k = X_{[k-t-1]} \text{ and j of the other t best}\\ \text{exceed } X_k\end{matrix}\right\} + (k-t)\sum_{j=0}^{t-1}\Pr\left\{\begin{matrix}\text{CD and } X_1 = X_{[k-t+1]} \text{ and j of the t best}\\ \text{exceed } X_1\end{matrix}\right\}\ . \tag{3.9}$$

Then (3.7) and (3.8) follow.

Table 3 presents, for various values of k,t, and P^*, the value d_3 such that the right hand side of (3.8) equals P^* when $F = \Phi$, the standard normal distribution function. These values were computed by Gaussian quadrature and are correct to two decimals; the values for t = 1 may be found in Bechhofer (1954).

It is interesting and important to note that for $t \geq 2$, $d(k,t,P^*) < d(k-1,t-1,P^*)$. If the configuration (3.3) with $\delta = 0$ actually were the least favorable configuration over the whole parameter space, it would follow that

$$\begin{aligned} P(CD|R_3,\theta_1=\ldots=\theta_k,t) &\leq \lim_{\theta_k\to\infty} P(CD|\theta_1=\ldots=\theta_{k-1},\theta_k,t) \\ &= P(CD|\theta_1=\ldots=\theta_{k-1},t-1) \ . \end{aligned}$$

This inequality is contradicted by the tables, so that the least favorable configuration over Ω is still unknown.

Now suppose that in the normal means case of Corollary 3.1, the common variance σ^2 is unknown. Suppose that s_r^2 satisfies (i) it is independent of $\bar{X}_1,\ldots,\bar{X}_k$ and (ii) s_r^2/σ^2 has the chi-square distribution G_r with r degrees of freedom. The natural rule to use is

R_5: Select π_i if $\bar{X}_i \geq \bar{X}_{[k-t+1]} - ds_r/\sqrt{n}$.

Then, the conclusion to Corollary 3.1 becomes

$$\inf_{\Omega} P_{\underline{\theta}}(CD|R_3) \geq t \int_0^\infty \int_{-\infty}^\infty \Phi^{k-t}(x+dz)[1-\Phi(x)]^{t-1} d\Phi(x)dG_r(z) \ . \tag{3.10}$$

Lemma 3.2 holds for $\delta = 0$ with

$$A_2(k,t,j) = \int_0^\infty \int_{-\infty}^\infty [1-\Phi(x)]^{t-1}\Phi^{k-2t+j}(x)[\Phi(x)-\Phi(x-dz)]^{t-j}d\Phi(x)dG(z) \ . \tag{3.11}$$

The problem of selecting the smallest scale parameter is quite similar. The natural rule to use is:

R_4: Select π_i if $X_i \le d_4 X_{[t]}$ $(d_4 \ge 1)$.

For this problem, (3.1) becomes

$$\inf_{\Omega} P_{\underline{\theta}}(CD|R_4) \ge t \int_0^\infty [1-F(x/d_4)]^{k-t} F^{t-1}(x)dF(x) \ . \tag{3.12}$$

A result similar to Lemma 3.2 can be found for the rule R_4 in the case where

$$\theta_{[1]} = \dots = \theta_{[t]} = \delta\theta_{[t+1]} = \dots = \delta\theta_{[k]} \quad (0 < \delta \le 1) \ . \tag{3.13}$$

Lemma 3.3. In the slippage case (3.13) using rule R_4 , Lemma 3.2 holds with

$$A_1(k,t,j) = \int_0^\infty F^j(x)[F(d_4x)-F(x)]^{t-j-1}(\delta x)[1-F(\delta x)]^{k-2t+j+1}dF(x) \ . \tag{3.14}$$

$$A_2(k,t,j) = \int_0^\infty F^j(x/\delta)[F(d_4x/\delta)-F(x/\delta)]^{t-j}F^{t-j-1}(x)[1-F(x)]^{k-2t+j}dF(x) \ . \tag{3.15}$$

4. Applications

The following can be seen as reasonable applications of the results of this paper.

4. a. (Student Admittance) A certain professional school admits students on the basis of objective scores and the results of a subject interview. In actual practice, the numerical scores are the first principal components of such factors as SAT scores, high school class rank, and college grade point averages. From a total of 38 in-state applicants in 1973 the school must admit at least 12. The objective scores ranged from -4.43 to 2.74. A reasonable hypothesis (somewhat supported by the data) is the existence of good, average, and poor groups. Thus, if $t = 12$, the following type configuration seems reasonable:

(4.1) $\theta_{[1]}=\ldots=\theta_{[s]}=\theta_{[s+1]}-\delta_1=\ldots=\theta_{[k-t]}-\delta_1=\theta_{[k-t+1]}\delta_1-\delta=\ldots=\theta_{[k]}-\delta_1-\delta$,

where $s \le k-t = 26$ and δ_1 , $\delta \ge 0$. One can show in a straightforward manner that $P\{CD|R_3\}$ is minimized when $\delta_1 = \delta = 0$, so that $P\{CD|R_3\}$ has an infimum equal to the expression in (3.8) with $k = 38$, $t = 12$. An estimate of the variance σ^2 , developed from (4.1), was $\hat{\sigma} = .40$; setting $P^* = .75$, a reasonable value of d obtained by a projection of the tables would be $d(P^*) = 2.5$, so that the rule becomes

R_3: Accept student $\Pi_i \iff X_i \ge X_{[27]} - 1.00$.

Using this rule, 20 of the 38 students would have been retained for further subjective study (namely the personal interview). For the procedure of Section 2 d_1 is approximately 4.5 , so that choosing $\Delta = 5.5$ would have led to similar results. In actuality, the school interviewed all applicants who wished to be seen, causing large expenditures of time; in 1974 their applications number well over 60 which should force them to develop some sort of a screening procedure.

4. b. <u>(Drug Studies)</u> In cancer research, the effect of drugs in often measured by such quantifiable variables as body weight loss, tumor size, and survival time. For example, suppose one is interested in finding the best of a number of dosage levels of a particular drug. The objective criterion could be one of the above variables, a linear combination of them or some other value function (see, for example, Goldin, Venditti, and Mantel (1961). However, even if one dosage seems to maximize the value function, it may be unacceptable because of non-quantitative side-effects not included in the value function. In this case, unlike the previous example, repeated obser-

vations are available. Thus, a sequential type of procedure could be used; given k dosage levels, one applies either of the rules of Sections 2 and 3 to eliminate those levels of poor value and those which have unacceptable side effects, retaining k_2 dosages. This procedure could be repeated or a final decision made after further study.

In Table 4 data (collected for the National Cancer Institute) are given for the weight loss and median survival time in days of mice inoculated with a certain type of tumor. The mice have been given a drug at different dosage levels (A,B,C,D,E,F) in ascending order and on different schedules (on Day 1 only, Days 1-5-9, and Days 1 through 9). Although most combinations were given to 10 mice, in order to make the numbers of observations in each group equal only the first 8 mice in each group were considered. Also, two combinations were left out because they gave rise to non-homogeneous sample variances; however, they both lead to high weight loss and low survival time. We will apply R_3 to find the combination giving rise to the smallest weight loss (largest weight gain); the procedures of Section 2 would also be applicable here.

Let $t = 4$ with $k = 12$ combinations. Then the data indicate something like (3.8) holds. Using $\hat{\sigma} = 19.5$, since $d \doteq 2.50$, $d\hat{\sigma}/\sqrt{n} \doteq 15.2$, so we reject combination Π_i if $\bar{X}_i \geq \bar{X}_{[4]} + 15.2$, since the interest is in the *smallest* weight loss. Thus, the number of combinations left is 6 , and they are marked by (*) in Table 4. Note the interesting but not too surprising fact that the selected combinations have uniformly higher median survival time than the rejected combinations.

4.c. (Choosing Measuring Methods) Suppose an analyst has k different methods of measuring the concentrate of a certain solution, and that for each method the measurements are normally distributed with unknown means μ_i and unknown variances σ_i^2. The best method in this case (assuming $\mu_1=\ldots=\mu_k$) would be the one with the smallest variance; however certain good methods may be inadmissible if they are found to be too time consuming. The methods of Section 2 (and Table 1) are directly applicable here, while approximations can be obtained for the methods of Section 3 by making use of the asymptotic distribution of the natural logarithm of the sample variance.

5. Summary

The methods developed in this paper will be useful in situations where one is attempting to choose one of k populations (a process, a drug, etc.) based on both numerical and what we have called subjective criteria. At the first stage, at least t $(t\geq 1)$ populations will be chosen using the methods of Sections 2 and 3. One method has as its goal the elimination of "poor" populations, the other attempts specifically to include good ones. It is very often useful to have $t \geq 2$ because an obviously superior population (based on numerical criteria) may be unacceptable for other reasons (see Section 4). Tables for the development of the elimination procedure are already available in the special case where one has k normal populations with common known variance; we have also provided tables for use when one would like to select small normal variances, tables which can also be used for the indifference zone problem of Bechhofer (1954). The selection schemes developed in Section 3 are complicated by the fact that we have not been able to find the least favorable configuration; however, a lower bound on P(CD) has been provided as have tables for the slippage case in the normal means problem. For normal variances, a similar bound and discussion of the slippage case have also been provided.

Acknowledgements

We wish to thank Dr. John M. Venditti of the National Cancer Institute and Sandra H. Stone of Arthur D. Little, Inc. for providing the data for the drug example of Section 4. We also thank Carol de Branges for computing the entries of Table 3.

References

[1] 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.

[2] Desu, M.M. (1970). A selection problem. *Ann. Math. Statist.* 41, 1596-1603.

[3] Deverman, J.N. (1969). A general selection procedure relative to the t best populations. Ph.D. Thesis, Purdue University.

[4] Deverman, J.N. and Gupta, S.S. (1969). On a selection procedure concerning the t best populations. *Ann. Math. Statist.* 40, 1870 (abstract).

[5] Goldin, A., Venditti, J.M., and Mantel, N. (1961). Preclinical screening and evaluation of agents for the chemotherapy of cancer: a review. *Cancer Research*, 21, 1334-51.

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

[7] Gupta, S.S. and Sobel, M. (1962). On a statistic which arises in selection and ranking problems. *Ann. Math. Statist.* 28, 956-967.

[8] Panchapakesan, S. (1969). Some contributions to multiple decision (selection and ranking) procedures. Ph.D. Thesis, Purdue University, Mimeo Series #192.

[9] Santner, T.J. (1974). A generalized goal in restricted subset selection theory. Dept. of Operations Research, Cornell University, Mimeo Series #217.

TABLE 1

This table lists the values of the right hand side of (2.4) for various r, k, t, and d_2 if F is the chi-square distribution with r^2 degrees of freedom

$\nu = 2$

k	t	d_2 .0200	.0400	.1000	.2000	.3000	.4000	.5000	.6000	.7000	.8000
3	2	.97069	.94268	.86580	.75758	.66890	.59524	.53333	.48077	.43573	.39683
4	2	.94268	.89031	.75758	.59524	.48077	.39683	.33333	.28409	.24510	.21368
	3	.96426	.93028	.83787	.71023	.60809	.52521	.45714	.40064	.35329	.31328
6	2	.89031	.79821	.59524	.39683	.28409	.21368	.16667	.13369	.10966	.09160
	3	.89796	.80992	.60809	.40064	.27921	.20292	.15237	.11747	.09253	.07422
	4	.92107	.85109	.67641	.47747	.34840	.26109	.20003	.15614	.12385	.09962
8	2	.84232	.72005	.48077	.28409	.18797	.13369	.10002	.07769	.06216	.05095
	3	.83787	.71023	.45714	.25000	.15237	.09998	.06921	.04988	.03710	.02828
	4	.85019	.72865	.47747	.26109	.15614	.09962	.06679	.04659	.03358	.02486
10	2	.79821	.65308	.39683	.21368	.13369	.09160	.06676	.05095	.04034	.03293
	3	.78325	.52581	.35329	.16711	.09253	.05662	.03710	.02552	.01820	.01336
	4	.78633	.62898	.34840	.15614	.08112	.04659	.02880	.01883	.01284	.00904
	5	.80137	.65034	.36933	.16655	.08510	.04749	.02833	.01784	.01175	.00805

$\nu = 4$

k	2	d_2 .0200	.0400	.1000	.2000	.3000	.4000	.5000	.6000	.7000	.8000
3	2	.99816	.99304	.96309	.88539	.79553	.70651	.62388	.54964	.48407	.42671
4	2	.99633	.98619	.92882	.79288	.65444	.53374	.43441	.35453	.29080	.23998
	3	.99764	.99111	.95330	.85743	.94967	.64601	.55262	.47119	.40143	.34221
6	2	.99268	.97280	.86707	.65247	.47333	.34299	.25132	.18697	.14132	.10846
	3	.99295	.97376	.87052	.65449	.46865	.33155	.23523	.16843	.12204	.08956
	4	.99445	.97930	.98583	.70946	.53464	.39513	.29039	.21367	.15796	.11754
8	2	.98907	.95979	.81288	.55092	.36312	.24271	.16628	.11691	.08420	.06194
	3	.98830	.95697	.79936	.51811	.32042	.19872	.12560	.08132	.05398	.03668
	4	.98896	.95932	.80797	.52846	.32543	.19853	.12224	.07650	.04881	.03179
10	2	.98549	.94714	.76488	.47411	.28981	.18243	.11912	.08049	.05601	.04000
	3	.98369	.94071	.73752	.42117	.23216	.13065	.07617	.04613	.02892	.01864
	4	.98351	.94003	.73294	.40741	.21446	.11377	.06196	.03488	.02036	.01232
	5	.98455	.94350	.74465	.42049	.22171	.11636	.06205	.03384	.01892	.01089

Table 1 (cont.)

$\nu = 6$

k	2	d_2 .0200	.0400	.1000	.2000	.3000	.4000	.5000	.6000	.7000	.8000
3	2	.99988	.99909	.98915	.94266	.86759	.77874	.68732	.60002	.52023	.44923
4	2	.99975	.99819	.97851	.89215	.76596	.63337	.51271	.41053	.32725	.26071
	3	.99983	.99880	.98583	.92661	.83385	.72754	.62177	.52422	.43816	.36429
6	2	.99950	.99639	.95840	.80684	.61913	.45381	.32666	.23432	.16880	.12257
	3	.99950	.99641	.95880	.80678	.61427	.44219	.30952	.21412	.14779	.10235
	4	.99959	.99711	.96677	.83982	.66824	.50334	.36718	.26314	.18691	.13234
8	2	.99925	.99459	.93924	.73713	.51759	.34769	.23117	.15457	.10460	.07181
	3	.99917	.99494	.93335	.71265	.47621	.29921	.18385	.11275	.06976	.04372
	4	.99920	.99424	.93569	.71891	.48051	.29885	.17985	.10701	.06374	.03832
10	2	.99901	.99281	.92103	.67882	.44300	.27809	.17434	.11078	.07166	.04733
	3	.99884	.99168	.90933	.63652	.38212	.21621	.12078	.06813	.03911	.02281
	4	.99880	.99140	.90652	.62434	.36216	.19480	.10208	.05352	.02853	.01555
	5	.99893	.99185	.91041	.63403	.36985	.19801	.10216	.05208	.02665	.01388

$\nu = 8$

k	t	d_2 .0200	.0400	.1000	.2000	.3000	.4000	.5000	.6000	.7000	.8000
3	2	.99999	.99988	.99672	.97057	.91236	.83012	.73618	.64069	.55016	.46807
4	2	.99998	.99976	.99347	.94337	.84032	.70959	.57725	.45843	.35882	.27860
	3	.99999	.99984	.99561	.96150	.88786	.78727	.67644	.56796	.46910	.38300
6	2	.99997	.99951	.98710	.89451	.72772	.54824	.39496	.27813	.19404	.13520
	3	.99997	.99951	.98700	.89362	.72302	.53695	.37744	.25685	.17175	.11392
	4	.99994	.99956	.98938	.91216	.76433	.59247	.43493	.30803	.21327	.14567
8	2	.99995	.99927	.98086	.85164	.64287	.44445	.29410	.19127	.12404	.08095
	3	.99994	.99913	.97860	.83536	.60676	.39547	.24271	.14470	.08530	.05016
	4	.99991	.99916	.97908	.83845	.60976	.39472	.23816	.13825	.97882	.04466
10	2	.99993	.99902	.97476	.81354	.47614	.37189	.23067	.14143	.08706	.05445
	3	.99992	.99885	.97040	.78458	.52004	.30567	.16946	.09183	.04941	.02668
	4	.99988	.99875	.96904	.77540	.50060	.28161	.14617	.07447	.03730	.01858
	5	.99989	.99879	.97007	.78115	.50724	.28474	.14681	.07245	.03522	.01711

TABLE 2

If $\pi_1, \ldots, \pi_k$ are normal populations with means (3.4) and common known variance $\sigma^2 = 1$, then the top number represents $P(CS|R_3)$, and the lower number represents $E(S|R_3)$ obtained from a Monte-Carlo study with $P^* = .90$.

k	t	δ = 0	1	2	3	4
2	1	.90	.977	.997	1.000	1.000
		1.80	1.694	1.444	1.200	1.060
4	1	.90	.992	.999	1.000	1.000
		3.60	2.696	1.732	1.356	1.136
	2	.976	.996	.998	1.000	1.000
		3.957	3.635	2.873	2.421	1.134
6	1	.90	.997	1.000	1.000	1.000
		5.40	2.983	1.838	1.439	1.156
	2	.976	.999	1.000	1.000	1.000
		5.919	4.404	2.975	2.494	2.220
	3	.991	.999	1.000	1.000	1.000
		5.977	5.312	4.067	3.536	3.242
8	1	.90	.997	1.000	1.000	1.000
		7.20	3.250	1.964	1.488	1.221
	2	.978	1.000	1.000	1.000	1.000
		7.896	4.644	3.133	2.607	2.302
	3	.991	.999	1.000	1.000	1.000
		7.975	5.765	4.124	3.598	3.320
	4	.995	.999	1.000	1.000	1.000
		7.990	6.749	5.155	4.602	4.296
10	1	.90	.997	1.000	1.000	1.000
		9.00	3.290	1.990	1.510	1.240
	2	.976	.998	1.000	1.000	1.000
		9.871	4.813	3.194	2.587	2.294
	3	.990	1.000	1.000	1.000	1.000
		9.969	5.989	4.285	3.704	3.366
	4	.999	1.000	1.000	1.000	1.000
		9.988	7.021	5.304	4.717	4.398
	5	.997	1.000	1.000	1.000	1.000
		9.994	8.005	6.247	5.685	5.384

If $\pi_1, \ldots, \pi_k$ are normal populations with means (3.4) and common known variance $\sigma^2 = 1$, then the top number represents $P(CS|R_3)$, and the lower number represents $E(S|R_3)$ obtained from a Monte-Carlo study with $P^* = .95$.

K	t	δ = 0	1	2	3	4
2	1	.95 1.900	.991 1.816	.999 1.590	1.000 1.316	1.000 1.118
4	1	.95 3.800	.997 3.048	1.000 1.960	1.000 1.492	1.000 1.224
	2	.990 3.979	.996 3.762	1.000 3.074	1.000 2.572	1.000 2.253
6	1	.95 5.700	..999 3.419	1.000 2.041	1.000 1.552	1.000 1.281
	2	.992 5.967	.999 4.774	1.000 3.182	1.000 2.630	1.000 2.314
	3	.997 5.992	1.000 5.540	1.000 4.261	1.000 3.665	1.000 3.359
8	1	.95 7.600	1.000 3.660	1.000 2.165	1.000 1.641	1.000 1.322
	2	.993 7.962	1.000 5.059	1.000 3.338	1.000 2.738	1.000 2.422
	3	.998 7.991	1.000 6.132	1.000 4.328	1.000 3.730	1.000 3.429
	4	1.000 7.998	1.000 7.032	1.000 5.340	1.000 4.726	1.000 4.396
10	1	.95 9.500	.999 3.710	1.000 2.190	1.000 1.650	1.000 1.340
	2	.990 9.954	1.000 5.252	1.000 3.394	1.000 2.726	1.000 2.399
	3	.997 9.991	1.000 6.330	1.000 4.504	1.000 3.817	1.000 3.474
	4	.998 9.997	1.000 7.385	1.000 5.482	1.000 4.829	1.000 4.508
	5	1.000 9.999	1.000 8.334	1.000 6.449	1.000 5.796	1.000 5.476

If $\pi_1, \ldots, \pi_k$ are normal populations with means (3.3) and common known variance $\sigma^2 = 1$, then the top number represents $P(CS|R_3)$, and the lower number represents $E(S|R_3)$ obtained from a Monte-Carlo study with $P^* = .90$.

K	t	δ: 0	1	2	3	4
2	1	.90 1.80	.982 1.707	.998 1.441	1.000 1.183	1.000 1.053
4	1	.90 3.60	.987 3.467	1.000 2.885	1.000 2.048	1.000 1.421
	2	.976 3.957	.992 3.912	.998 3.655	1.000 3.120	1.000 2.529
6	1	.90 5.40	.987 5.146	.999 4.294	1.000 2.984	1.000 1.793
	2	.976 5.919	1.000 5.839	1.000 5.426	1.000 4.566	1.000 3.362
	3	.991 5.977	1.000 5.950	1.000 5.776	1.000 5.255	1.000 4.350
8	1	.90 7.20	.987 6.998	.998 6.019	1.000 4.351	1.000 2.541
	2	.978 7.896	.998 7.792	1.000 7.399	1.000 6.314	1.000 4.577
	3	.991 7.975	1.000 7.943	1.000 7.737	1.000 7.069	1.000 5.726
	4	.995 7.990	1.000 7.980	1.000 7.870	1.000 7.444	1.000 6.431
10	1	.90 9.00	.986 8.700	1.000 7.546	1.000 5.278	1.000 3.055
	2	.976 9.871	.997 9.755	1.000 9.231	1.000 7.882	1.000 5.623
	3	.990 9.969	.999 9.923	1.000 9.691	1.000 8.894	1.000 7.131
	4	.994 9.988	1.000 9.973	1.000 9.863	1.000 9.351	1.000 8.030
	5	.997 9.994	1.000 9.987	1.000 9.924	1.000 9.586	1.000 8.544

If $\pi_1, \ldots, \pi_k$ are normal populations with means (3.3) and common known variance $\sigma^2 = 1$, then the top number represents $P(CS|R_3)$, and the lower number represents $E(S|R_3)$ obtained from a Monte-Carlo study with $P^* = .95$.

K	t	δ = 0	1	2	3	4
2	1	.95 1.900	.994 1.823	1.000 1.597	1.000 1.298	1.000 1.108
	1	.95 3.800	.994 3.708	1.000 3.246	1.000 2.447	1.000 1.657
	2	.990 3.979	.997 3.967	1.000 3.806	1.000 3.394	1.000 2.786
6	1	.95 5.700	.994 5.544	1.000 4.843	1.000 3.622	1.000 2.285
	2	.992 5.967	1.000 5.926	1.000 5.697	1.000 5.012	1.000 3.870
	3	.997 5.992	1.000 5.978	1.000 5.891	1.000 5.540	1.000 4.760
8	1	.95 7.600	.996 7.463	1.000 6.724	1.000 5.213	1.000 3.297
	2	.993 7.962	.999 7.923	1.000 7.704	1.000 6.879	1.000 5.350
	3	.998 7.991	1.000 7.981	1.000 7.876	1.000 7.448	1.000 6.328
	4	1.000 7.998	1.000 7.990	1.000 7.944	1.000 7.681	1.000 6.875
10	1	.95 9.500	.994 9.307	1.000 8.390	1.000 6.397	1.000 3.946
	2	.990 9.954	1.000 9.905	1.000 9.599	1.000 8.658	1.000 6.615
	3	.997 9.991	1.000 9.978	1.000 9.855	1.000 9.354	1.000 7.792
	4	.998 9.997	1.000 9.990	1.000 9.944	1.000 9.655	1.000 8.714
	5	1.000 9.999	1.000 9.994	1.000 9.973	1.000 9.789	1.000 9.094

TABLE 3

Lists the value d such that the expression (3.5) equals P* .

K	t	P* .90	.95	.99
4	2	1.97	2.35	3.10
5	2	2.20	2.58	3.31
6	2	2.36	2.74	3.46
6	3	2.06	2.41	3.08
7	2	2.49	2.86	3.57
7	3	2.23	2.57	3.22
8	2	2.59	2.96	3.67
8	3	2.35	2.69	3.33
8	4	2.14	2.46	3.11
9	2	2.67	3.04	3.76
9	3	2.46	2.79	3.43
9	4	2.26	2.58	3.17
10	2	2.74	3.11	3.82
10	3	2.55	2.89	3.57
10	4	2.36	2.66	3.20

TABLE 4

The combinations marked by (*) are selected.

	Drug Level	Schedule	Average Weight Loss ($\overline{X}$)	Median Survival Time in Days
	A	Day 1	48.57	7.0
	C	Day 1	54.50	6.0
(*)	A	1-5-9	23.25	11.0
	B	1-5-9	46.25	7.0
	C	1-5-9	56.25	6.0
(*)	D	1-5-9	29.25	11.0
(*)	F	1-5-9	11.00	13.0
	A	1-9	45.25	8.0
	B	1-9	47.25	6.0
(*)	D	1-9	33.25	11.0
(*)	E	1-9	17.25	11.5
(*)	F	1-9	18.75	11.0