TWO-SAMPLE PROCEDURES IN SIMULTANEOUS ESTIMATION[1]

By W. C. Healy, Jr.[2]

University of Illinois

1. Summary. In this paper, two-sample procedures of the type originated by Stein [4] are developed for a number of problems in simultaneous estimation. The results include the construction of simultaneous confidence intervals of prescribed length or lengths and confidence coefficient $1 - \alpha$ for (1) all normalized linear functions of means, (2) all differences between means, and (3) the means of k independent normal populations with common unknown variance. Simultaneous confidence intervals of length l and confidence coefficients known to be not less than $1 - \alpha$ are constructed for all normalized linear functions of the means of a general multivariate normal population. The single sample analogues of these problems have been discussed by Tukey [5], Scheffé [6] and Bose and Roy [7]. Also, a confidence region having prescribed diameter (or volume) and confidence coefficient $1 - \alpha$ is constructed for the mean vector in the general multivariate normal case.

The procedures depend only on known and tabulated distributions. Illustrative applications from the analysis of variance are described.

2. Introduction. In 1940, Dantzig [2] showed that for the Student problem

$$\text{(2.1)} \qquad \begin{aligned} &\text{Hypothesis:} \quad \mu = \mu_0 \\ &\text{Alternative:} \quad \mu \neq \mu_0 , \end{aligned}$$

where μ is the unknown mean of a normal distribution with unknown variance σ^2, there exists no test having power independent of σ^2 based on a sample of fixed size. More generally, it is shown in [3], Sec. 5.2, that if θ is a location parameter and an unknown scale parameter is present, there exist neither confidence intervals of prescribed length and confidence coefficient nor point estimates with bounded expected squared error for θ. The important general problem has thus been posed: how to conduct experiments in order to obtain a predetermined degree of accuracy in the presence of unknown scale parameters.

In 1945, Stein [4] provided an ingenious solution by sampling in two stages for the case of the Student hypothesis (2.1), and, in fact, for a general linear hypothesis. In his procedure the size of the sample at the second stage depends on the results of the first. He also provided, in the same vein, a two-sample technique obtaining a confidence interval for μ having predetermined length and

Received December 13, 1954.

[1] This work was supported in part by the Office of Ordnance Research, U. S. Army, under Contract No. DA-11-022-ORD-881.

[2] Now with the Ethyl Corporation Research Laboratories, Ferndale, Michigan.

confidence coefficient. By this is meant a rule for constructing an interval (which is a function of the observations) with the two properties that

(a) the length of the interval is equal to l;

(b) the probability that the interval contains the true value of μ is exactly $1 - \alpha$,

where l and $1 - \alpha$ have been specified in advance. Recently, Seelbinder [10] has published tables of the expected total sample size for Stein's procedure.

Problems of simultaneous estimation and simultaneous tests of hypotheses constitute a dilemma to practicing statisticians. A common example was long provided in the application of the analysis of variance when the F-test had rejected the hypothesis of homogeneity of means. The natural desire of the experimenter to make further inferences about the means, such as deciding between which groups of means differences existed, was thwarted by existing statistical theory before 1950. Analysis of variance theory made no provision for such successive inferences, and the experimenter who proceeded anyway accepted the hazard of an unknown significance level for his final conclusions.

Work by Tukey [5], Scheffé [6], Bose and Roy [7], Dunnett [13], and others since 1950 has produced valid techniques for making such simultaneous or successive comparisons, and, importantly, including comparisons suggested by the data themselves. In a number of problems involving normal distributions, and including the F-test dilemma above, the techniques are easy to apply. This contributes to their practical importance.

In this paper, the Stein two-sample idea for obtaining predetermined accuracy is applied to some of the simultaneous confidence interval problems considered in [5], [6], and [7]. Three basic problems which are appropriate to a variety of applications are treated first, with examples. They involve k independent normal populations with unknown means and unknown *common* variance.

Suppose that a (joint) confidence coefficient $1 - \alpha$ is prescribed. Then, Problem I is to construct a system of confidence intervals of prescribed lengths $l_1, l_2, \cdots, l_k$ for the k means. Problem II is to construct a system of confidence intervals each of prescribed length l for the $k(k - 1)/2$ differences between the means. Problem III is to construct a system of confidence intervals, one for each possible normalized linear function of the k means; each interval is to have length l, where l is specified in advance. These systems of intervals are each to have joint confidence coefficient $1 - \alpha$. For a number of applied situations the comparisons of interest are reducible to those of Problems I, II, or III.

In addition, Problems IV and V involving the means of a k-variate normal population with unknown covariance matrix are treated. Problem IV is to construct a system of confidence intervals, one for each possible linear function of the k means. The length and confidence coefficient specifications are as in Problem III. Problem V is to construct a confidence region for the k means, having a prescribed maximum diameter and confidence coefficient. Problems IV and V are thought to be the first multivariate-normal two-sample procedures.

Mention might here be made of the application of the Stein idea to the problem of ranking means of normal populations, made by Bechhofer, Dunnett and Sobel [1].

In Section 3 are stated the distribution results employed for solution of these problems. In Section 4 the univariate problems and solutions are given; in Section 5, three analyses of variance situations are shown, for illustrative purposes, to correspond to Problems I, II, and III, and hence are solved. In Section 6 the multivariate problems and solutions are described, and the solutions are justified.

3. Distribution results.

THEOREM 3.1. (Stein). *Let* $X_{i1}, X_{i2}, \cdots$ $(i = 1, 2, \cdots, k)$ *be mutually independent random variables,* X_{ih} *being distributed* $N(\theta_i, \sigma^2)$. *Let* $n_1, n_2, \cdots, n_k$ *be fixed nonnegative integers. Let* s^2 *be an unbiased estimate of* σ^2 *based on* m *degrees of freedom, distributed independently of* $\sum_{h=1}^{n_i} X_{ih}$ *and* $X_{i,n_i+1}, X_{i,n_i+2}, \cdots$ $(i = 1, 2, \cdots, k)$. *Let* $a_{i1}, a_{i2}, \cdots, a_{iN_i}, a_i$, *and* N_i *be functions of* s^2 *such that* (1) $a_{ih} = a_i$ *for* $h \leqq n_i$ *and* (2) $N_i \geqq \max(n_i, 1)$. *Define*

$$Y_i = \frac{\sum_{h=1}^{N_i} a_{ih}(X_{ih} - \theta_i)}{\left(\sum_{h=1}^{N_i} a_{ih}^2\right)^{1/2}} .$$

Then, $Y_1, Y_2, \cdots, Y_k, s^2$ *are mutually independent random variables,* Y_i *being distributed* $N(0, \sigma^2)$.

COROLLARY 3.2. *Define* $W_1 = \max_{i \leqq k}(|Y_i|/s)$. *Then,* W_1 *has the distribution of the Studentized Maximum Modulus with* (k, m) *degrees of freedom.*

This distribution is tabulated in [11].

COROLLARY 3.3. *Define* $W_2 = \max_{i,j \leqq k} |Y_i - Y_j|/s$. *Then,* W_2 *has the distribution of the Studentized Range with* (k, m) *degrees of freedom.*

This distribution is tabulated in [8].

COROLLARY 3.4. *Define* $W_3 = \sum_{i=1}^{k} Y_i^2 / s^2$. *Then,* W_3/k *has the F-distribution with* (k, m) *degrees of freedom.*

These three corollaries are the distribution results required for the univariate examples that are to follow. A multivariate analogue of Theorem 3.1 is stated next.

THEOREM 3.5. *Let* $\mathbf{X}_h = (X_{1h}, X_{2h}, \cdots, X_{kh})'$, $(h = 1, 2, \cdots)$ *be mutually independent random vectors,*[3] *the* $\mathbf{X}_h$ *having a multivariate normal distribution with mean vector* $\boldsymbol{\theta}$ *and covariance matrix* $\boldsymbol{\sigma} = (\sigma_{ij})$. *Let* n *be a fixed nonnegative integer. Let* $\mathbf{S} = (S_{ij})$ *be a matrix of unbiased estimates* S_{ij} *of the* σ_{ij}, *the* S_{ij} *having jointly a Wishart distribution with* m *degrees of freedom. Suppose the* S_{ij} *are independent of both* $\sum_{j=1}^{n} \mathbf{X}_j$ *and* $\mathbf{X}_{n+1}, \mathbf{X}_{n+2}, \cdots$. *Let* $a_1, a_2, \cdots, a_n$, a, *and* N *be functions of* $\mathbf{S}$ *such that* (1) $a_j = a$ *for* $j \leqq n$, *and* (2) $N \geqq \max(n, 1)$. *Define*

$$\mathbf{Y} = \sum_{j=1}^{N} a_j(\mathbf{X}_j - \boldsymbol{\theta}) \Bigg/ \left(\sum_{j=1}^{N} a_j^2\right)^{1/2} .$$

[3] A prime will be used to denote the transpose of a vector or matrix.

Then,

$$\frac{m - k + 1}{k} \mathbf{Y}'S^{-1}\mathbf{Y}$$

has the F-distribution with $(k, m - k + 1)$ *degrees of freedom.*

The proof is similar to the univariate case and will be omitted.

4. Univariate problems and solutions.

4.1. PROBLEM I—ESTIMATION OF MEANS. The Problems I, II, and III of this section deal with the following situation: $X_{i1}, X_{i2}, \cdots$ $(i = 1, 2, \cdots, k)$ are mutually independent random variables. The distribution of X_{ih} is $N(\theta_i, \sigma^2)$, θ_i and σ^2 being unknown.

Nonnegative integers $n_1, n_2, \cdots, n_k$ have somehow been determined, and $X_{i1}, X_{i2}, \cdots, X_{in_i} (i = 1, 2, \cdots, k)$ have been observed if $n_i > 0$. If $n_i = 0$, no observations have been taken on the ith distribution.

Let s^2 be an unbiased estimate of σ^2 based on m degrees of freedom, which is independent of both $\sum_{h=1}^{N_i} X_{ih}$, and $X_{i,n_i+1}, X_{i,n_i+2}, \cdots$. A value of s^2 has been observed; it may or may not have been computed from the observations X_{ih}, $h \leqq n_i$.

With this set-up we can now state problems and solutions. First a remark is in order about the statements of the problems. A confidence coefficient $1 - \alpha$ will be prescribed in advance and it can be exactly attained. However, we will require only that the actual confidence coefficient attained shall be $\geqq 1 - \alpha$. The reason is that the solutions then obtained are uniform improvements on the solutions obtained by requiring exactly $1 - \alpha$ confidence coefficient. This same situation is encountered and discussed in Stein's original paper [4].

STATEMENT OF PROBLEM I. Given $0 < 1 - \alpha < 1$ and $l_1, l_2, \cdots, l_k$, with $l_i > 0$, determine joint confidence intervals $I_i(X)$ for $\theta_1, \theta_2, \cdots, \theta_k$ such that

(1) length of $I_i(X) = l_i$,

(2) $\Pr\{\theta_i \varepsilon I_i(X) \text{ for all } i \leqq k\} \geqq 1 - \alpha$ for all $\theta_1, \theta_2, \cdots, \theta_k, \sigma^2$. It is of interest to note that in the single sample analogue of this problem, given, for example, in [7], p. 519, the k confidence intervals are obliged to have the same (random) length, while here the lengths are allowed to differ.

SOLUTION OF PROBLEM I. Determine constants c_i such that

$$\Pr\left\{W_1 \leqq \frac{l_i}{2\sqrt{c_i}}\right\} = 1 - \alpha \qquad i = 1, 2, \cdots, k,$$

when W_1 has the distribution of the Studentized Maximum Modulus with (k, m) degrees of freedom. Determine integers N_i by

$$N_i = \max\{n_i, [s^2 / c_i] + 1\}, \tag{4.1.1}$$

where [ ] means "greatest integer less than."

Observe $X_{i,n_i+1}, X_{i,n_i+2}, \cdots, X_{i,N_i}$ $(i = 1, 2, \cdots, k)$ if $N_i > n_i$, and estimate θ_i by the interval

$$\frac{1}{N_i}\sum_{h=1}^{N_i} X_{ih} \pm \frac{l_i}{2}. \tag{4.1.2}$$

4.2 Problem II—Estimation of Differences.

Statement of Problem II. Here we will take $n_1 = n_2 = \cdots = n_k = n$. Given $0 < 1 - \alpha < 1$ and $l > 0$, determine joint confidence intervals $I_{ij}(X)$ for the $k(k-1)/2$ differences $\theta_i - \theta_j$, $i < j \leqq k$, such that

(1) length of $I_{ij}(X) = l$,

(2) $\Pr\{\theta_i - \theta_j \,\varepsilon\, I_{ij}(X)$ for all $i < j \leqq k\} \geqq 1 - \alpha$ for all $\theta_1, \theta_2, \cdots, \theta_k, \sigma^2$.

Solution of Problem II. Determine a constant c such that

$$\Pr\left\{W_2 \leqq \frac{l}{2\sqrt{c}}\right\} = 1 - \alpha,$$

when W_2 has the distribution of the Studentized Range with (k, m) degrees of freedom. Determine N by

$$N = \max\,(n, [s^2 / c] + 1). \tag{4.2.1}$$

Observe $X_{i,n+1}, X_{i,n+2}, \cdots, X_{iN}$ $(i = 1, 2, \cdots, k)$ if $N > n$, and estimate $\theta_i - \theta_j$ by the interval

$$\frac{1}{N}\sum_{h=1}^{N} (X_{ih} - X_{jh}) \pm \frac{l}{2}. \tag{4.2.2}$$

4.3 Problem III—Estimation of Contrasts.

Statement of Problem III. Again take $n_1 = n_2 = \cdots = n_k = n$. Given $0 < 1 - \alpha < 1$ and $l > 0$, determine a system of simultaneous confidence intervals $I_\nu(X)$ for the elements of the set L of all linear functions $\sum_{i=1}^k C_{i\nu}\theta_i$ with $\sum_{i=1}^k C_{i\nu}^2 = 1$. (The index ν denotes a particular element of L.)

The intervals are to have the properties

(1) length of $I_\nu(X) = l$,

(2) $\Pr\{\sum_{i=1}^k C_{i\nu}\theta_i \,\varepsilon\, I_\nu(X)$ for all $\nu\} \geqq 1 - \alpha$ for all $\theta_1, \theta_2, \cdots, \theta_k, \sigma^2$.

Solution of Problem III. Determine a constant c such that

$$\Pr\{W_3 \leqq l^2/4c\} = 1 - \alpha$$

when W_3/k has the F-distribution with (k, m) degrees of freedom. Determine N by

$$N = \max(n, [s^2 / c] + 1). \tag{4.3.1}$$

Observe $X_{i,n+1}, X_{i,n+2}, \cdots, X_{iN}$, if $N > n$, and estimate $\sum_{i=1}^k C_{i\nu}\theta_i$ by the interval

$$\frac{1}{N}\sum_{i=1}^{k} C_{i\nu}\sum_{h=1}^{N} X_{ih} \pm \frac{l}{2}. \tag{4.3.2}$$

Note that to estimate some linear function $\sum_{i=1}^k d_{i\nu}\theta_i$, where $\sum_{i=1}^k d_{i\nu}^2 \neq 1$, we simply employ the interval

$$\frac{1}{N}\sum_{i=1}^{k} d_{i\nu}\sum_{h=1}^{N} X_{ih} \pm \sqrt{\sum_{i=1}^{k} d_{i\nu}^2}\,\frac{l}{2}.$$

The argument required to justify these solutions is essentially the same as that originally given by Stein, and details will be omitted.

5. Three applications. In this section we apply the previous section to the solution of the two-sample versions of three simultaneous estimation examples treated by Bose and Roy and by Scheffé for the single sample case.

5.1. 2^r *Factorial experiment.* In this example we will utilize Problem I of the preceding section. Suppose that in an experiment involving r factors, each at two levels, it is desired to obtain joint confidence intervals of fixed length for the r main effects and $r(r - 1)/2$ two-factor interactions. Suppose also that the experimental situation is replicable as many times as desired. An example might be an experiment to discover the effect on the yield of a chemical reaction of the addition or nonaddition of different reagents.

Let $\theta_{11}, \theta_{22}, \cdots, \theta_{rr}$ be the true values of the main effects, and let $\theta_{12}, \theta_{13}, \cdots, \theta_{r-1,r}$ be the true two-factor interactions. Denote the factors by $A_1, A_2, \cdots, A_r$ and let the symbolic product $(a_{i_1} a_{i_2} \cdots a_{i_r})$ denote the true yield when factors $A_{i_1}, A_{i_2}, \cdots, A_{i_r}$ are at their upper levels and all other factors are at their lower levels. For the chemical illustration, $(a_1a_2a_3)$ means the true yield when reagents 1, 2, 3 have been added, and no others.

Then θ_{ii} is defined by the result of multiplying out the expression

$$\theta_{ii} = \frac{1}{2^r - 1}(a_1 + 1)(a_2 + 1) \cdots (a_{i-1} + 1)(a_i - 1)(a_{i+1} + 1) \cdots (a_r + 1),$$

and θ_{ij} is defined by the result of multiplying out the expression

$$\theta_{ij} = \frac{1}{2^r - 1}(a_1 + 1)(a_2 + 1) \cdots (a_{i-1} + 1)(a_i - 1)(a_{i+1} + 1) \cdots (a_{j-1} + 1)(a_j - 1)(a_{j+1} + 1) \cdots (a_{r-1} + 1)(a_r + 1),$$

where the dots indicate terms with plus 1's. For further details on factorial experiments, the reader is referred to [9], for example.

We will conduct the experiment by taking some number of replications of the 2^r factorial design. Let Y_{ijh} represent the usual estimate of θ_{ij} from the hth replition. That is, Y_{ijh} is obtained by substituting the observed yields for the true yields in the expression defining θ_{ij}. Then the usual assumptions are that Y_{ij1}, $Y_{ij2}, \cdots, i \leqq j \leqq r$, are mutually independent random variables with Y_{ijh} distributed $N[\theta_{ij}, (\sigma^2/2^{r-2})]$, where σ^2 is the variance of a single observed yield. Suppose that an unbiased estimate of σ^2 can be obtained from the sum of squares due to replications.

Having decided upon a confidence coefficient $1 - \alpha$ and a common length l, the problem before us is to produce $r + (r - 1)r/2$ confidence intervals of length l, one each for the $\theta_{ij}, i \leqq j \leqq r$. Except for notational changes, this is exactly Problem I with a common length l for all the intervals. It remains to adapt the solution of Problem I to the present case.

The first step is to obtain a preliminary estimate of $\sigma^2/2^{r-2}$. There are a variety of ways to accomplish this. A simple one to describe is:

Choose an integer n and perform n replications of the 2^r factorial design; compute the replications sum of squares, say T, which will be based on $(n - 1)2^r$ degrees of freedom; and estimate $\sigma^2/2^{r-2}$ by $T/(n - 1)2^r 2^{r-2} = T/(n - 1)2^{2r-2}$.

Determine a constant c such that

$$\Pr\left\{W_1 \leqq \frac{l}{2\sqrt{c}}\right\} = 1 - \alpha$$

when W_1 has the distribution of the Studentized Maximum Modulus with $r + r\,(r - 1)/2, 2^r(n - 1)$ degrees of freedom; determine N by

$$N = \max\left\{n, \left[\frac{T}{(n - 1)2^{2r-2}c}\right] + 1\right\};$$

perform $N - n$ further replications of the 2^r factorial design; and estimate θ_{ij} by the interval

$$\frac{1}{N}\sum_{h=1}^{N} Y_{ijh} \pm \frac{l}{2}, \qquad i \leqq j \leqq r.$$

Of course, it is not necessary to replicate the entire design in order to estimate $\sigma^2/2^{r-2}$. In the event one does replicate only a portion of the design for this purpose, a question not encountered before can arise; namely, what to do if the total number of replications required, based on the estimate of σ^2, is smaller than the number of replications already obtained of a portion of the design. This question seems too special to discuss further than to point it out, at this time.

In practice it would be unlikely that all main effects and 2-factor interactions would be of equal importance. It would be tempting to specify different lengths as in Problem I. However, the factorial design requires that each combination be replicated equally often in order to get orthogonal estimates; and the estimates all have a common variance. In this case, one might end up choosing N based on the smallest length and would then get the same results as though all lengths had been specified equal to the smallest.

5.2. *Randomized blocks experiment.* In a randomized block setup for comparing k treatments, suppose it is desired to obtain joint confidence intervals of fixed length and confidence coefficient for the $k(k - 1)/2$ differences between the true treatment means. This example utilizes the solution of Problem II.

We will assume the following conventional model for any given block, say block number h:

$$Y_{ih} = \mu + \theta_i + b_h + e_{ih},$$

where Y_{ih} is the observation on treatment i in block h,

μ is a constant,

θ_i is the contribution from the ith treatment,

b_h is the contribution from the hth block,

e_{ih} are mutually independent, each distributed $N(0, \sigma^2)$ for $i = 1, 2, \cdots, k$; $h = 1, 2, \cdots$.

Assume that the experiment can be replicated in as many blocks as desired.

The problem is this: given a confidence coefficient $1 - \alpha$ and a length l, produce $k(k-1)/2$ simultaneous confidence intervals each of length l, one each for the differences $\theta_i - \theta_j$, $(i < j \leqq k)$. To recognize this as Problem II, let

$$X_{ih} = Y_{ih} - \mu - b_h. \tag{5.2.1}$$

Then $X_{i1}, X_{i2}, \cdots, (i = 1, 2, \cdots, k)$ are mutually independent random variables, X_{ih} being distributed $N(\theta_i, \sigma^2)$. In terms of the X_{ih} the problem is exactly Problem II. Although we cannot observe the X_{ih}, we can nevertheless write down the solution from (4.2) in terms of the X_{ih}; we will then discover that the solution to the original problem depends only on the original observations Y_{ih}.

Following (4.2), the experiment can be conducted as follows:

Choose an integer n and perform the randomized block experiment with n blocks and estimate σ^2. The conventional estimate, which we adopt and call T, is based on $(k-1)(n-1)$ degrees of freedom. This estimate is computed from the Y_{ih} that are observed.

Determine a constant c such that

$$\Pr\left\{W_2 \leqq \frac{l}{2\sqrt{c}}\right\} = 1 - \alpha$$

when W_2 has the distribution of the Studentized Range with k, $(k-1)(n-1)$ degrees of freedom; determine N by

$$N = \max\{n, [T/c] + 1\};$$

perform a second randomized block experiment with $N - n$ blocks, if $N > n$; and estimate $\theta_i - \theta_j$ by the interval

$$\frac{1}{N}\sum_{h=1}^{N}(X_{ih} - X_{jh}) \pm \frac{l}{2}, \qquad i < j \leqq k.$$

This is the solution in terms of the X_{ih}. To get the solution in terms of the Y_{ih}, note from (5.2.1) that each interval is

$$\frac{1}{N}\sum_{h=1}^{N}(Y_{ih} - \mu - b_h - \{Y_{jh} - \mu - b_h\}) \pm \frac{l}{2},$$

or

$$\frac{1}{N}\sum_{h=1}^{N}(Y_{ih} - Y_{jh}) \pm \frac{l}{2}.$$

5.3. *Two-way analysis of variance with replications.* Consider a situation appropriately represented by a two-way classification, say by rows and by columns, and suppose there are k rows and p columns. Suppose further that the situation is replicable as many times as desired.

Let Y_{ijh} denote the observation in the ith row, jth column of replication number h. We adopt the following conventional model for the hth replication:

(5.3.1) $Y_{ijh} = \mu + r_i + t_j + b_{ij} + e_{ijh}, \quad i = 1, 2 \cdots, k, j = 1, 2 \cdots, p,$

where μ is a constant,

r_i is the contribution from the ith row,

t_j is the contribution from the jth column,

b_{ij} is the contribution from the (i, j) cell,

e_{ijh} are mutually independent random variables, each distributed $N(0, \sigma^2)$.

Suppose we are interested in comparisons between true row means, i.e., between the quantities

$$\theta_i = \mu + r_i + \frac{1}{p}\sum_{j=1}^{p} b_{ij}$$
$$= \mu + r_i + \bar{b}_{i.}, \qquad i = 1, 2, \cdots, k,$$

where $\bar{b}_{i.} = 1/p \sum_{j=1}^{p} b_{ij}$. (We will use this dot notation in the usual way to indicate summed-out subscripts.)

Suppose further that the situation is such that we cannot tell in advance what row comparisons will be of interest or what rows may turn out to be important. In this situation, since we do not know precisely what we want, it may be desirable to ask for a fixed degree of accuracy for any and all confidence statements that might be made about contrasts between the row means. A contrast is a linear function $\sum_{i=1}^{k} C_i\theta_i$ such that $\sum_{i=1}^{k} C_i = 0$. If we should fix the confidence coefficient at $1 - \alpha$ for the infinite set of all possible contrasts, then for any necessarily finite number of contrasts that we decide to estimate, the joint confidence coefficient must exceed $1 - \alpha$.

It is apparent, however, that requiring the confidence intervals for the various contrasts to have a common fixed length would be asking too much; two contrasts differing only by a constant multiplier and each estimated by an interval of length l are logically incompatible. We will ask instead that the intervals for all contrasts $\sum_{i=1}^{k} C_i\theta_i$, such that $\sum_{i=1}^{k} C_i^2 = 1$, should have fixed length l; i.e., we consider only "normalized" contrasts. This is equivalent to asking that the interval for every contrast $\sum_{i=1}^{k} d_i\theta_i$ should have length $l(\sum_{i=1}^{k} d_i^2)^{1/2}$.

The problem is this: given a joint confidence coefficient $1 - \alpha$ and a length l, produce a system of joint confidence intervals, each of length l, one each for every normalized contrast $\sum_{i=1}^{k} C_i\theta_i$. We have now to reduce this problem to Problem III. To this end, let

(5.3.2) $$\bar{Y}_{i.h} = \frac{1}{p}\sum_{j=1}^{p} Y_{ijh} = \mu + r_i + \bar{b}_{i.} + \bar{t}_{.} + \bar{e}_{i.h},$$

and let

(5.3.3) $$X_{ih} = \bar{Y}_{i.h} - \bar{t}_{.} = \mu + r_i + \bar{b}_{i.} + \bar{e}_{i.h} = \theta_i + \bar{e}_{i.h}.$$

Then, $X_{i1}, X_{i2}, \cdots$ $(i = 1, 2, \cdots, k)$ are mutually independent random variables, X_{ih} being distributed $N(\theta_i, \sigma^2/p)$.

The problem is now in the form of Problem III, except for the present restriction that $\sum_{i=1}^{k} C_i = 0$, since we here are considering only contrasts. This exception can be resolved by the following reduction.

Make an orthogonal transformation from the X_{ih} to Z_{ih}, defined by

$$X_{ih} = \sum_{j=1}^{k} u_{ij} Z_{jh}, \qquad i = 1, 2, \cdots, k, \tag{5.3.4}$$

such that

$$Z_{kh} = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} X_{ih}. \tag{5.3.5}$$

Then, since the inverse of an orthogonal matrix is its transpose, $u_{ik} = 1/\sqrt{k}$ $(i = 1, 2, \cdots, k)$. In terms of the Z_{ih}, substitution from (5.3.4) and (5.3.5) gives

$$\sum_{i=1}^{k} C_i X_{ih} = \sum_{i=1}^{k} \sum_{j=1}^{k} C_i u_{ij} Z_{jh} = \sum_{i=1}^{k} \sum_{j=1}^{k-1} C_i u_{ij} Z_{jh},$$

since $u_{ik} = 1/\sqrt{k}$ and $\sum_{i=1}^{k} C_i = 0$.

Therefore,

$$\sum_{i=1}^{k} C_i X_{ih} = \sum_{j=1}^{k-1} d_j Z_{jh}, \tag{5.3.6}$$

where $d_j = \sum_{i=1}^{k} C_i u_{ij}$.

The Z_{jh}, $j \leqq k$, are independently normally distributed with common variance σ^2/p. Setting $\eta_j = E\{Z_{jh}\}$, we have

$$\sum_{i=1}^{k} C_i \theta_i = \sum_{j=1}^{k-1} d_j \eta_j, \tag{5.3.7}$$

by taking the expectation of (5.3.6).

Also if

$$\sum_{i=1}^{k} C_i^2 = 1, \qquad \text{then} \sum_{j=1}^{k-1} d_j^2 = 1, \tag{5.3.8}$$

by computing the variance of both sides of (5.3.6).

Therefore, in view of (5.3.6), (5.3.7), and (5.3.8) we have reduced the set of all normalized contrasts $\sum_{i=1}^{k} C_i \theta_i$ to the set of all normalized linear functions $\sum_{j=1}^{k-1} d_j \eta_j$, *without* a restriction that $\sum_{j=1}^{k-1} d_j = 0$. This reduction is given in [6].

The problem has become: to construct joint confidence intervals of length l and confidence coefficient $1-\alpha$ for all linear functions $\sum_{j=1}^{k-1} d_j \eta_j$, with $\sum_{j=1}^{k-1} d_j^2 = 1$, based on the random variables $Z_{j1}, Z_{j2}, \cdots$ $(j = 1, 2, \cdots, k-1)$ which are mutually independent, Z_{jh} being distributed $N(\eta_j, \sigma^2/p)$. This is now, in terms of the Z_{jh}, exactly the Problem III, and we proceed to adapt its solution. Again, it will turn out that the solution will depend only on the Y_{ijh}, which can be observed.

Following 4.3 the experiment can be conducted as follows: Choose an integer n and perform n replications of the two-way layout in order to estimate σ^2/p. Suppose that the replications sum of squares, say T, provides an unbiased estimate of σ^2; T will be based on $kp(n - 1)$ degrees of freedom. Estimate σ^2/p, therefore, by $T/kp^2(n - 1)$.

Determine a constant c such that

$$\Pr\left\{W_3 \leqq \frac{l^2}{4c}\right\} = 1 - \alpha,$$

when $W_3/k - 1$ has the F-distribution with $k - 1$, $kp(n - 1)$ degrees of freedom; determine N by

$$N = \max\left\{n, \left[\frac{T}{kp^2(n - 1)c}\right] + 1\right\};$$

perform $N - n$ further replications of the two-way layout, if $N > n$; and estimate

$$\sum_{j=1}^{k-1} d_j \eta_j = \sum_{i=1}^{k} C_i \theta_i$$

by the interval

$$\frac{1}{N} \sum_{h=1}^{N} \left(\sum_{j=1}^{k-1} d_j Z_{jh}\right) \pm \frac{l}{2}. \tag{5.3.9}$$

To express this interval in terms of the original observations, substitute in (5.3.9) from (5.3.6) and (5.3.3), obtaining

$$\frac{1}{N} \sum_{h=1}^{N} \sum_{i=1}^{k} C_i X_{ih} \pm l/2$$

or

$$\frac{1}{N} \sum_{h=1}^{N} \sum_{i=1}^{k} C_i(\overline{Y}_{i.h} - \bar{t}_{.}) \pm l/2$$

or

$$\frac{1}{N} \sum_{h=1}^{N} \sum_{i=1}^{k} C_i \overline{Y}_{i.h} \pm l/2$$

or

$$\sum_{i=1}^{k} C_i \overline{Y}_{i..} \pm l/2.$$

5.4 *Comments on the examples.* The preceding three examples were chosen for illustrative purposes. Depending on actual circumstances, any analysis of variance design could produce any of the three types of problems we have considered, that is, estimation of independent effects, differences, or general contrasts.

In each example we have specified a way, arbitrarily, to estimate the variance, utilizing the experiment design involved. It is perhaps worth remarking that it is only necessary to have an independent unbiased estimate of the variance. Where is comes from is immaterial and in practice it may come from some other experiment (though in theory it should not have been used for any other purpose). This fact is inherent in the setup of Problems I, II, III, since we allow the initial sample sizes n_i to be 0.

Another feature of the examples is the lack of discussion of how to choose the degrees of freedom on which to base the estimate of variance. Such a discussion would presumably be based on tables of the expected total sample size, but such tables are lacking for these problems.

Also, somewhat more general situations than the preceding examples would indicate are reducible to one of the Problems I, II, or III. In particular, the situations in Scheffé [6], p. 87, and Bose and Roy [7], p. 515–519, when posed as problems of simultaneous confidence intervals of fixed length and confidence coefficient are so reducible. The methods of reduction are essentially the same as indicated in these papers.

6. Multivariate problems and solutions. Throughout the discussion of multivaate problems, we will denote all matrices by boldface letters; primes will denote matrix transposes.

6.1. Problem IV—Estimation of Contrasts. The Problems IV and V of this section deal with the following multivariate situation:

$$\mathbf{X}_h' = (X_{1h}, X_{2h}, \cdots, X_{kh}), \qquad h = 1, 2, \cdots,$$

are mutually independent random vectors, each $\mathbf{X}_h$ having the multivariate normal distribution with mean vector $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\sigma}$.

A nonnegative integer n has somehow been determined, and if $n > 0$, values of $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$ have been observed. If $n = 0$, no observations on $\boldsymbol{X}$ have been taken.

$\boldsymbol{S} = (S_{ij})$ is a matrix of random variables S_{ij}, the S_{ij} having a Wishart distribution with m degrees of freedom; S_{ij} is an unbiased estimate of σ_{ij}, and the S_{ij} are independent of $\sum_{h=1}^{n}\mathbf{X}_h$, and $\mathbf{X}_{n+1}, \mathbf{X}_{n+2}, \cdots$. A value of S_{ij} has been observed, $i \leqq j \leqq k$. These may or may not have been computed from the $\mathbf{X}_h$, $h \leqq n$.

Statement of Problem IV. Given $0 < 1 - \alpha < 1$ and $l > 0$, determine a system of simultaneous confidence intervals $I_\nu(\boldsymbol{X})$ for the elements of the set D of all linear functions $\sum_{i=1}^{k} C_{i\nu}\theta_i$ with $\sum_{i=1}^{k} C_{i\nu}^2 = 1$. The intervals are to have the properties that

$$\text{(1) the length of } I_\nu(\boldsymbol{X}) = l \text{ for all } \nu,$$

$$\text{(2) } \Pr\left\{\sum_{i=1}^{k} C_{i\nu}\theta_i \;\varepsilon\; I_\nu(\boldsymbol{X}) \text{ for all } \nu\right\} \geqq 1 - \alpha \text{ for all } \boldsymbol{\theta} \text{ and } \boldsymbol{\sigma}. \tag{6.1.1}$$

SOLUTION OF PROBLEM IV. Determine a constant c such that

$$\Pr\left\{W_4 \leqq \frac{l^2}{4c}\right\} = 1 - \alpha$$

when $(m - k + 1)\, W_4/k$ has the F-distribution with $(k, m - k + 1)$ degrees of freedom; determine N by

$$N = \max\left\{n, \left[\frac{\lambda}{c}\right] + 1\right\}, \tag{6.1.2}$$

where λ is the largest latent root of S; observe $\mathbf{X}_{n+1}, \mathbf{X}_{n+2}, \cdots \mathbf{X}_N$; and estimate $\sum_{i=1}^{k} C_{i\nu}\theta_i$ by

$$\mathbf{C}_\nu'\overline{\mathbf{X}} \pm l/2, \tag{6.1.3}$$

where $\mathbf{C}_\nu' = (C_{1\nu}, C_{2\nu}, \cdots C_{k\nu})$ and $\overline{\mathbf{X}} = 1/N\sum_{h=1}^{N}\mathbf{X}_h$.

JUSTIFICATION OF SOLUTION. We have to establish (6.1.1). Now,

$$\begin{aligned} &\Pr\{|\mathbf{C}_\nu'(\overline{\mathbf{X}} - \boldsymbol{\theta})| \leqq l/2 \text{ for all } \nu\} \\ &= \Pr\left\{N\,|\,\mathbf{C}_\nu'(\overline{\mathbf{X}} - \boldsymbol{\theta})|^2 \leqq \frac{Nl^2}{4} \text{ for all } \nu\right\} \\ &\geqq \Pr\left\{N\,|\,\mathbf{C}_\nu'(\overline{\mathbf{X}} - \boldsymbol{\theta})|^2 \leqq \frac{\lambda l^2}{4c} \text{ for all } \nu\right\}, \end{aligned}$$

since $N \geqq \lambda/c$. Using the fact that

$$\sup_{\mathbf{C}_\nu'\mathbf{C}_\nu = 1} \mathbf{C}_\nu' S \mathbf{C}_\nu = \lambda,$$

where λ is the largest latent root of S, it follows that

$$\Pr\left\{N\,|\,\mathbf{C}_\nu'(\overline{\mathbf{X}} - \boldsymbol{\theta})|^2 \leqq \frac{\lambda l^2}{4c} \text{ for all } \nu\right\}$$

$$\geqq \Pr\left\{\frac{N\,|\,\mathbf{C}_\nu'(\overline{\mathbf{X}} - \boldsymbol{\theta})|^2}{\mathbf{C}_\nu' S \mathbf{C}_\nu} \leqq \frac{l^2}{4c} \text{ for all } \nu\right\} \tag{6.1.4}$$

$$= \Pr\left\{\sup_{\mathbf{C}_\nu'\mathbf{C}_\nu = 1}\left(\frac{N\,|\,\mathbf{C}_\nu'(\overline{\mathbf{X}} - \boldsymbol{\theta})\,|^2}{\mathbf{C}_\nu' S \mathbf{C}_\nu}\right) \leqq \frac{l^2}{4c}\right\}$$

$$= \Pr\left\{N(\overline{\mathbf{X}} - \boldsymbol{\theta})' S^{-1}(\overline{\mathbf{X}} - \boldsymbol{\theta}) \leqq \frac{l^2}{4c}\right\}, \tag{6.1.5}$$

since

$$\sup_{\mathbf{C}_\nu'\mathbf{C}_\nu = 1} \frac{(\mathbf{C}_\nu'\mathbf{u})^2}{\mathbf{C}_\nu' S \mathbf{C}_\nu} = \mathbf{u}' S^{-1}\mathbf{u}.$$

The Theorem 3.5, with $a_h = 1/N$, and the definition of c imply that (6.1.5) is equal to $1 - \alpha$; this establishes (6.1.1).

Problem V is a multivariate analogue of the original Stein procedure.

6.2. Problem V—Estimation of Mean Vector.

Statement of Problem V. Given $0 < 1 - \alpha < 1$ and $l > 0$, to construct a confidence region $R(X)$ for $\boldsymbol{\theta}$ such that

(6.2.1) (1) the maximum diameter of $R(X)$ does not exceed l,

(6.2.2) (2) $\Pr\{\boldsymbol{\theta} \varepsilon R(X)\} = 1 - \alpha.$

Here it is possible to obtain a solution for which the maximum diameter of $R(X)$ is exactly l, but the solution we present is uniformly better.

Solution of Problem V. Determine a constant c such that

$$\Pr\left\{W_5 \leqq \frac{l_2}{4c}\right\} = 1 - \alpha$$

when $(m - k + 1)\, W_5/k$ has the F-distribution with $(k, m - k + 1)$ degrees of freedom. Determine N by

$$N = \max\left\{n, \left[\frac{\lambda}{c}\right] + 1\right\},$$

where λ is the largest latent root of $\boldsymbol{S}$.

Observe $\mathbf{X}_{n+1}, \mathbf{X}_{n+2}, \cdots, \mathbf{X}_N$, and estimate $\boldsymbol{\theta}$ by the set $R(X)$ of points $\mathbf{t}$ satisfying

$$N(\overline{\mathbf{X}} - \mathbf{t})'\boldsymbol{S}^{-1}(\overline{\mathbf{X}} - \mathbf{t}) \leqq \frac{l^2}{4c}, \tag{6.2.3}$$

where $\overline{\mathbf{X}} = 1/N \sum_{h=1}^{N} \mathbf{X}_h$.

A similar problem in which $R(X)$ is required to have predetermined volume is solvable in a similar way but is possibly less useful.

Justification of Solution. We have to establish (6.2.1) and (6.2.2) when $R(X)$ is the set of points $\mathbf{t}$ satisfying (6.2.3). Now (6.2.2) means that

$$\Pr\left\{N(\overline{\mathbf{X}} - \boldsymbol{\theta})'\, \boldsymbol{S}^{-1}(\overline{\mathbf{X}} - \boldsymbol{\theta}) \leqq \frac{l^2}{4c}\right\}$$

when $\boldsymbol{\theta}$ is the true mean vector, and hence follows from the definition of c and Theorem 3.5 with $a_h = 1/N$.

To establish (6.2.1), we first note that (with probability one) (6.2.3) defines the interior and boundary of an ellipsoid in k-dimensional space. Now, if $\mathbf{u}'\boldsymbol{A}\mathbf{u} = 1$ is the equation of an ellipsoid, its maximum diameter is $2\sqrt{\lambda}$, where λ is the largest latent root of $\boldsymbol{A}^{-1}$. Replacing $\boldsymbol{A}$ by $4cN^{-}\ \boldsymbol{S}^{1}/l^2$, it follows that the maximum diameter of the ellipsoid associated with (6.2.3) is $l\sqrt{\lambda}/\sqrt{cN} \leqq l$, since $N \geqq \lambda/c$. This establishes (6.2.1).

7. Concluding Remarks. As in the case of Stein's original work, it is possible to modify the solutions of Problems I, II, and III slightly so as to obtain a confidence procedure with exactly $1 - \alpha$ confidence coefficient, and in Problem V so that the maximum diameter is exactly l. Such modifications, however, result in larger expected sample sizes. It should be emphasized that such modifications as these do not make possible the attainment of an exact confidence coefficient in Problem IV, because of the inequality (6.1.4); an intuitive picture of why this is so might perhaps be given by the following remarks. The total sample sizes are the same for Problems IV and V if l and $1 - \alpha$ are the same. Thus, after the completion of sampling, we have the "information" that $\boldsymbol{\theta}$ lies in an ellipsoid like (6.2.3). However, the simultaneous intervals constructed in Problem IV use only the "information" that $\boldsymbol{\theta}$ is in the sphere having the same center as (6.2.3) and with diameter equal to the maximum diameter of (6.2.3). Thus, to the extent that the ellipsoid (6.2.3) is smaller than the sphere, the actual confidence coefficient will exceed $1 - \alpha$.

It is fairly obvious that the use of the confidence region $R(\boldsymbol{X})$ given for Problem V, modified to have maximum diameter exactly l, to test a hypothesis concerning $\boldsymbol{\theta}$ does not yield a test with power independent of the unknown covariance matrix. This is so since the shape of the region $R(\boldsymbol{X})$ is not independent of the unknown covariance matrix. It is possible, but in a wasteful and artificial way, to construct a test of the multivariate hypothesis $\boldsymbol{\theta} = \mathbf{0}$ having power independent of the unknown covariance matrix. This can be done simply by estimating $\sigma_{11}, \sigma_{22}, \cdots, \sigma_{kk}$ from k separate subsamples of the original sample and essentially employing the Stein procedure for the Student hypothesis to the individual hypotheses $\theta_1, = 0, \theta_2 = 0, \cdots, \theta_k = 0$.

There are many other problems which come to mind in connection with the Stein procedure and to which no specific allusion seems to have been made in the literature. One is to obtain a confidence interval of fixed length and confidence coefficient for a given linear fundtion $\sum_{i=1}^{k} C_i\theta_i$ of the means of a general k-variate normal distribution with unknown covariance matrix. This problem is directly solved by Stein's paper [4], since

$$\frac{\sqrt{N} \sum_{i=1}^{k} C_i(\bar{X}_i - \theta_i)}{\left(\sum_{j=1}^{k} \sum_{i=1}^{k} C_i C_j S_{ij}\right)^{1/2}}$$

has the Student t-distribution with m degrees of freedom, when the S_{ij} are unbiased estimates of σ_{ij}, having a Wishart distribution with m degrees of freedom, and $(\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_k)$ is the (independent) sample mean vector based on N observations.

If $\theta_1, \theta_2, \cdots, \theta_k$ are the means of k independent normal populations with unknown (unequal) variances, it is a direct extension of results of Chapman [12] to obtain a confidence interval of fixed length and confidence coefficient for a given linear function $\sum_{i=1}^{k} C_i\theta_i$. The procedure depends on the distribution of the sum of k independent Student-t variables, for which tables do not seem to exist for $k > 2$. However, the normal approximation should be useful except for very small values of k and the degrees of freedom m.

Finally, no study has been made of expected sample sizes for the procedures in this paper. Tables along the lines of Seelbinder [10] but with degrees of freedom $k(n - 1)$ rather than $n - 1$ would be helpful for the univariate problems. Expected sample sizes for the multivariate procedures would be more complicated.

8. Acknowledgment. The author would like to thank Professor W. G. Madow for valuable assistance in the preparation of this paper.

REFERENCES

[1] R. E. Bechhofer, C. W. Dunnett, and M. Sobel, "A two-sample multiple decision procedure for ranking means of normal distributions with a common unknown variance," *Biometrika*, Vol. 41 (1954), p. 170.

[2] G. B. Dantzig, "On the nonexistence of tests of 'Students' hypothesis having power functions independent of σ", *Ann. Math. Stat.*, Vol. 11 (1940), p. 186.

[3] E. L. Lehmann, "Theory of estimation," notes recorded by Colin Blyth, Associated Students' Store, University of California, 1950.

[4] C. Stein, "A two-sample test for a linear hypothesis whose power is independent of the variance," *Ann. Math. Stat.*, Vol. 16 (1945), p. 243.

[5] J. W. Tukey, "The problem of multiple comparisons," unpublished paper.

[6] Henry Scheffé, "A method of judging all contrasts in the analysis of variance," *Biometrika*, Vol. 40 (1953), p. 87.

[7] R. C. Bose and S. N. Roy, "Simultaneous confidence interval estimation," *Ann. Math. Stat.*, Vol. 24 (1953), p. 513.

[8] J. M. May, "Extended and corrected tables of the upper percentage points of the 'Studentized' range," *Biometrika*, Vol. 39 (1952), p. 192.

[9] O. Kempthorne, *The Design and Analysis of Experiments*, John Wiley and Sons, 1952.

[10] B. M. Seelbinder, "On Stein's two-stage sampling scheme," *Ann. Math. Stat.*, Vol. 24 (1953), p. 640.

[11] K. C. S. Pillai and K. V. Ramachandran, "On the distribution of the ratio of the ith observation in an ordered sample from a normal population to an independent estimate of the standard deivation," *Ann. Math. Stat.*, Vol. 25 (1954), p. 565.

[12] D. G. Chapman, "Some two-sample tests," *Ann. Math. Stat.*, Vol. 21 (1950), p. 601.

[13] C. W. Dunnett, "A multiple comparison procedure for comparing several treatments with a control," *J. Amer. Statis. Assoc.*, Vol. 50 (1955), p. 1096.