ESTIMATION OF THE MEANS OF DEPENDENT VARIABLES

By Olive Jean Dunn[1]

Statistical Laboratory, Iowa State College

1. Summary. Methods are given for constructing sets of simultaneous confidence intervals for the means of variables which follow a multivariate normal distribution.

In section (3), a set of confidence intervals is obtained for each of two special cases; first when the variances are assumed to be known, and second when the variances are assumed to be equal. These two sets have the property that the confidence is known exactly, rather than merely being bounded below. In the case of known variances, the intervals are of fixed lengths (i.e., the lengths are the same from sample to sample); when the variances are unknown, the intervals are of variable lengths. It may be surprising to note that nothing need be known about the covariances in order to obtain confidence intervals of fixed lengths whose confidence coefficient is exact. These intervals are long, and do not make use of all the information provided by the sample.

Each of sections (4) to (7) considers a different method for obtaining confidence intervals of bounded confidence level. In each section a set of fixed lengths is obtained when the variances are assumed to be known, while a set of variable lengths is obtained when the variances are unknown but equal. In section (5) the set of variable lengths applies to the general multivariate normal distribution, all the other confidence intervals in this paper require some assumption concerning the variances.

In section (8) the sets of intervals are compared on the basis of length. One of the bounded confidence level methods, which has been established only for two or three variables or for an arbitrary number of variables with a special type of correlation matrix, is shown to yield the best possible set. Another of the bounded confidence level methods, whose use is established in general, is shown to be almost as good as the best set for confidence coefficients of practical interest.

It is interesting to notice that intervals with bounded confidence level, are found which are much shorter than the ones whose confidence level is exact. This need not surprise us, however. In the case of just one variable, we might easily find that the 95% confidence intervals for the mean using the t-statistic were shorter on the average than 94% confidence intervals using order statistics. Moreover, since in admitting sets of confidence intervals with bounded confidence level we consider a much broader class of methods, we might almost expect that some of them would give better intervals.

2. Introduction. The problem of estimating the unknown means of dependent variables arises frequently in situations where repeated measurements are made

Received February 15, 1957; revised March 25, 1958.

[1] Now at the University of California at Los Angeles.

on the same individuals, and the assumption of independence is unjustifiable. In biological research, for example, growth data are often obtained with measurements taken on n individuals at k different times; the measurements would be highly correlated. The psychologist might measure n individuals' responses to k different levels of a stimulus; again, a high degree of dependence would be expected. The point estimates chosen for the means would be the same as for independent variables; in this paper we wish to develop simultaneous confidence intervals for the means.

Let $y_1, \cdots, y_k$ be k jointly distributed variables whose means are $\mu_1, \cdots, \mu_k$ respectively. A set of simultaneous confidence intervals for $\mu_1, \cdots, \mu_k$ with confidence coefficient $1 - \alpha$ consists of $2k$ functions of the sample values, say g_i and h_i, $i = 1, 2, \cdots, k$, with the following property: if E_i is the event that the interval g_i to h_i covers μ_i, $i = 1, 2, \cdots, k$, then the probability that E_1, $E_2, \cdots, E_k$ occur simultaneously is greater than or equal to $1 - \alpha$, where $0 < \alpha < 1$. Symbolically,

$$P(E_1E_2, \cdots, E_k) = P(g_1 < \mu_1 < h_1, \cdots, g_k < \mu_k < h_k) \geqq 1 - \alpha.$$

If the inequality sign holds, the set is of bounded confidence level.

Paul G. Hoel has in a recent paper [1] given a method for estimating a mean regression curve and a confidence band for it which is applicable to the situations we have in mind provided one assumes the existence of a polynomial regression curve of a given degree. In this paper we shall assume that the experimenter is actually interested in the regression curve, but is either unwilling to make the necessary laborious calculations or else is unable to make the necessary assumptions concerning its form. He knows that there exist methods for studying linear contrasts among the means, but this is not what he wishes to do. He might indeed decide to make k different 95% confidence intervals, one for each of the k means; this is satisfactory only when he focuses on one individual mean.

We shall assume, then, that he will welcome a set of k confidence intervals, one for each mean, being assured, with a high probability, that such a set covers all k means simultaneously.

Another situation in which such a set of intervals would be useful arises when a regression line, curve, or surface has been fitted, and several predictions are made on the basis of it.

Suppose, for example, that the assumption has been made that the variables x_i are normally distributed with means $\alpha + \beta t_i$ and variances σ^2, and that the maximum likelihood estimate $\hat{\alpha} + \hat{\beta} t_i$ has been calculated from a sample of size m.

At any particular value of t, say t_0, one can obtain a prediction interval for x_0, an observation drawn at random from the x's belonging to t_0, by using the fact that $u_0 = x_0 - \hat{\alpha} - \hat{\beta} t_0$ is normally distributed. But the research worker is cautioned not to do this for more than one value of t, and of course this is exactly what he wishes to do.

If he goes ahead and gets such intervals at k different points, say $t_1^*, \cdots, t_k^*$, he has the same unsatisfactory situation as with repeated tests of significance. The variables $u_i^* = x_i^* - \hat{\alpha} - \hat{\beta} t_i$, where x_i^* is an observation chosen at random from the x's at $t = t_i^*$, $i = 1, 2, \cdots, k$, are normally distributed and are correlated; thus the methods of this paper may be used to give simultaneous prediction intervals for the points $x_i^*, \cdots, x_k^*$.

3. Confidence regions using independent linear combinations.

3.1. Assuming first known variances, we seek independent linear combinations of the sample values which can be used to give a set of confidence intervals of fixed lengths whose confidence level is exact.

The observations $y_{1j}, y_{2j}, \cdots, y_{kj}$, $j = 1, \cdots, n$, are a random sample of n observations from $n_k(y_1, \cdots, y_k)$, the multivariate normal distribution with unknown means, $\mu_1, \cdots, \mu_k$, known variances, $\sigma_1^2, \cdots, \sigma_k^2$, and unknown covariances λ_{is}, $i \neq s$.

Let $z_i = \sum_{j=1}^n a_{ji} y_{ij}$, $i = 1, \cdots, k$, with the following restrictions on the a_{ji}:

$$(1) \quad \sum_{j=1}^n a_{ji} = 1, \qquad i = 1, \cdots, k$$

$$(2) \quad \sum_{j=1}^n a_{ji}\, a_{js} = 0, \qquad i \neq s$$

$$(3) \quad \sum_{j=1}^n a_{ji}^2 = c^2, \qquad i = 1, 2, \cdots, k.$$

The means, variances, and covariances of the z_i may then be calculated, remembering that $E(y_{ij} - \mu_i)(y_{sj} - \mu_s) = \lambda_{is}$, but that (since two observations in a random sample are independent) $E(y_{ij} - \mu_i)(y_{sj'} - \mu_s) = 0$ for $j \neq j'$. The means of the z_i are calculated to be μ_i, $i = 1, \cdots, k$, their variances are proportional to $\sigma_1^2, \cdots, \sigma_k^2$, and their covariances are zero.

To determine the a_{ji}, let $A = (a_{ji})$, an $n \times k$ matrix. The columns of A may be considered to be k vectors in an n-dimensional Euclidean space, each with an end fixed at the origin. The three conditions imply (1) that the k vectors have their endpoints on the plane which passes through the unit points on the coordinate axes, $P: \sum_{i=1}^n a_i = 1$; (2) that they be mutually orthogonal; and (3) that their lengths equal c.

If $n \geqq k$, the columns of

$$D = \begin{bmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & c \\ \hline 0 & 0 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \text{ an } n \times k$$

matrix, are k mutually orthogonal vectors of length c whose endpoints lie on any plane

$$P': \frac{a_1}{c} + \cdots + \frac{a_k}{c} + \frac{a_{k+1}}{m_{k+1}} + \cdots + \frac{a_n}{m_n} = 1.$$

The plane P' can be rotated into the plane P provided the distances of the two planes from the origin are equal; this will be true if

$$c^2 = \frac{k}{n - \frac{1}{m_{k+1}^2} - \cdots - \frac{1}{m_n^2}}.$$

To make the lengths of the confidence intervals formed from the z_i as small as possible, c^2 should be minimized. This is accomplished by choosing for P' the plane $\sum_{i=1}^k (a_i/c) = 1$; then $c^2 = (k/n)$.

The solution is then $A = BCD$, where B is an $n \times n$ orthogonal matrix whose first column consists of the elements $n^{-\frac{1}{2}}$;

$$C = \left[\begin{array}{ccc|ccc} \frac{1}{\sqrt{k}} & \cdots & \frac{1}{\sqrt{k}} & 0 & \cdots & 0 \\ \cdots & & \cdots & & & \\ \cdots & & \cdots & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & & & & & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right], \text{ an } n \times n$$

matrix consisting of zeros except for a $k \times k$ orthogonal matrix in the upper left corner whose first row is $k^{-\frac{1}{2}}, \cdots, k^{-\frac{1}{2}}$; and D is defined as before, with $c = (k/n)$. For C rotates the column vectors of D into vectors whose endpoints lie on the plane $a_1 = n^{-\frac{1}{2}}$. B^{-1} rotates the plane $\sum_{i=1}^n a_i = 1$ into the plane $a_1 = n^{-\frac{1}{2}}$, so that B rotates the k mutually orthogonal vectors of length $(k/n)^{\frac{1}{2}}$ into vectors whose endpoints lie on the plane $\sum_{i=1}^n a_i = 1$. The problem thus reduces to that of writing down a $k \times k$ orthogonal matrix and an $n \times n$ orthogonal matrix.

The $z_1, \cdots, z_k$ are then independently normally distributed with means μ_i and variances $(k/n)\sigma_i^2$. Thus

$$P\left(z_1 - \sqrt{\frac{k}{n}}\,\sigma_1 c_\alpha < \mu_1 < z_1 + \sqrt{\frac{k}{n}}\,\sigma_k c_\alpha, \cdots, z_k - \sqrt{\frac{k}{n}}\,\sigma_k c_\alpha < \mu_k < z_k + \sqrt{\frac{k}{n}}\,\sigma_k c_\alpha\right) =$$

$1 - \alpha$, where c_α is defined by

$$N(c_\alpha) = \frac{1 + (1 - \alpha)^{1/k}}{2},$$

with N the cumulative distribution function of the standard normal variable. The set of confidence intervals is $z_i \pm (k/n)^{\frac{1}{2}}\sigma_i c_\alpha$.

3.2. When the variances are unknown but are assumed to be equal, the same method may be used to construct t-variables whose numerators are independent but which have the same denominator, provided $n > k$. Let $\sigma_i^2 = \sigma^2, i = 1, \cdots, k$.

Let

$$z_i = \sum_{j=1}^n a_{ji} y_{ij}, \qquad i = 1, \cdots, k,$$

and

$$u_m = \sum_{j=1}^{n} b_{jm} y_{rj}, \qquad m = 1, \cdots, n-k,$$

where r is any integer from 1 to k. The choice of r is arbitrary. It may be the same for each u_m, or different r's may be used for the different values of m. The problem is to determine the a_{ji} and b_{jm} so that $z_1, \cdots, z_k, u_1, \cdots, u_{n-k}$ will be independently normally distributed variables with $E(z_i) = \mu_i$, $i = 1, \cdots, k$; $E(u_m) = 0, m = 1, \cdots, n-k, E(z_i - \mu_i)^2 = (k/n)\sigma^2, i = 1, \cdots, k; E(u_m^2) = \sigma^2$, $m = 1, \cdots, n-k$. This will be accomplished provided

(1) $\sum_{j=1}^{n} a_{ji} = 1, i = 1, \cdots, k$, since $E(z_i) = \mu_i \sum_{j=1}^{n} a_{ji} = \mu_i$.

(2) $\sum_{j=1}^{n} a_{ji}^2 = \frac{k}{n}, i = 1, \cdots, k$, since $E(z_i - \mu_i)^2 = \sigma^2 \sum_{j=1}^{n} a_{ji}^2 = \frac{k}{n}\sigma^2$.

(3) $\sum_{j=1}^{n} a_{ji} a_{js} = 0, i \neq s$, since $E(z_i - \mu_i)(z_s - \mu_s) = \lambda_{is} \sum_{j=1}^{n} a_{ji} a_{js} = 0$.

(4) $\sum_{j=1}^{n} b_{jm} = 0, m = 1, \cdots, n-k$, since $E(u_m) = \mu_m \sum_{j=1}^{n} b_{jm} = 0$.

(5) $\sum_{j=1}^{n} b_{ji}^2 = 1, i = 1, \cdots, n-k$, since $E(u_m^2) = \sigma^2 \sum_{j=1}^{n} b_{jm}^2 = \sigma^2$.

(6) $\sum_{j=1}^{n} b_{jm} b_{js} = 0, m \neq s$, since $E(u_m u_s) = E(y_r - \mu_r)(y_{r'} - \mu_{r'}) \sum_{j=1}^{n} b_{jm} b_{js} = 0$.

(7) $\sum_{j=1}^{n} a_{ji} b_{jm} = 0, i = 1, \cdots, k; m = 1, \cdots, n-k$, since $E(z_i - \mu_i)(u_m)$

$$= \lambda_{ir} \sum_{j=1}^{n} a_{ji} b_{jm} = 0.$$

Thus n mutually orthogonal vectors are needed, k of length $(k/n)^{\frac{1}{2}}$ with endpoints on the plane $\sum_{i=1}^{n} a_i = 1$, and $n-k$ of length one with endpoints on the plane $\sum_{i=1}^{n} a_i = 0$.

Let

$$D = \left[\begin{array}{c|c} \begin{matrix} \sqrt{\frac{k}{n}} & 0 & \cdots & 0 \\ 0 & \sqrt{\frac{k}{n}} & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & \sqrt{\frac{k}{n}} \end{matrix} & \begin{matrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 0 \end{matrix} \\ \hline \begin{matrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 0 \end{matrix} & \begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 1 \end{matrix} \end{array}\right] \begin{matrix} \}\, k \\ \\ \}\, n-k \end{matrix},$$

an $n \times n$ matrix whose columns are n mutually orthogonal vectors of the needed lengths.

Let

$$C = \left[\begin{array}{ccc|cccc} \frac{1}{\sqrt{k}} & \cdots & \frac{1}{\sqrt{k}} & 0 & & \cdots & 0 \\ \cdots & \cdots & & \cdots & & & \cdots \\ \cdots & \cdots & & 0 & & \cdots & 0 \\ \hline 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \cdots & & & & & & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{array}\right],$$

an $n \times n$ orthogonal matrix which rotates the first k columns of D into vectors whose endpoints lie on the plane $a_1 = n^{-\frac{1}{2}}$ and which leaves the last $n - k$ columns unchanged.

Let B be an $n \times n$ orthogonal matrix whose first column consists entirely of the elements $n^{-\frac{1}{2}}$. Since B rotates the plane $a_1 = n^{-\frac{1}{2}}$ into the plane $\sum_{i=1}^n a_i = 1$, it must also rotate the parallel plane $a_1 = 0$ into $\sum_{i=1}^n a_i = 0$.

Thus $A = BCD$ is an $n \times n$ matrix whose columns are orthogonal vectors. The first k are of length $(k/n)^{\frac{1}{2}}$ and have endpoints on $\sum_{i=1}^n a_i = 1$; the last $n - k$ are of length one and have endpoints on $\sum_{i=1}^n a_i = 0$.

Then let

$$t_i = \frac{z_i - \mu_i}{\sqrt{\frac{k}{n(n-k)} \sum_{m=1}^{n-k} u_m^2}}, \qquad i = 1, \cdots, k.$$

These are k t-variables whose numerators are independent but whose denominators are the same. Their frequency function is (see [2]):

$$f_{n-k}(t_1, \cdots, t_k) = \frac{\Gamma\left(\frac{n}{2}\right)}{[\pi(n-k)]^{k/2}\Gamma\left(\frac{n-k}{2}\right)} \left(1 + \frac{\sum_{i=1}^k t_i^2}{n-k}\right)^{-(n/2)}$$

If c_α is defined by

$$\int_{-c_\alpha}^{c_\alpha} \cdots \int_{-c_\alpha}^{c_\alpha} f_{n-k}(t_1, \cdots, t_k)\, dt_1, \cdots, dt_k = 1 - \alpha,$$

then $P(-c_\alpha < t_1 < c_\alpha, \cdots, -c_\alpha < t_k < c_\alpha) = 1 - \alpha$. Thus an exact set of confidence intervals of equal but variable lengths is obtained:

$$z_i \pm c_\alpha \sqrt{\frac{k}{n(n-k)} \sum_{m=1}^{n-k} u_m^2}, \qquad i = 1, \cdots, k.$$

4. Intervals of bounded confidence level using the chi-square distribution and Hotelling's T-distribution.

4.1. Known variances. For a sample of size n from the multivariate normal distribution with means $\mu_1, \cdots, \mu_k$ and covariance matrix (λ_{is}), the expression

$n\sum_{i=1}^{k}\sum_{s=1}^{k}\lambda^{is}(\bar{y}_i - \mu_i)(\bar{y}_s - \mu_s)$ follows a Chi-square distribution with k degrees of freedom. Here λ^{is} denotes an element of the inverse matrix $(\lambda^{is}) = (\lambda_{is})^{-1}$, and $\bar{y}_i$ is the sample mean of the observations on y_i. Then

$$\sum_{i=1}^{k}\sum_{s=1}^{k}\lambda^{is}(\bar{y}_i - \mu_i)(\bar{y}_s - \mu_s) = \frac{c_\alpha}{n},$$

where c_α is defined by $U_k(c_\alpha) = 1 - \alpha$, with U_k the cumulative distribution function of a Chi-square variable with k degrees of freedom. In the parameter space of the $\mu_1, \cdots, \mu_k$, this equation defines an ellipsoid, which will be denoted by E. Then

$$P\left(\sum_{i=1}^{k}\sum_{s=1}^{k}\lambda^{is}(\bar{y}_i - \mu_i)(\bar{y}_s - \mu_s) < \frac{c_\alpha}{n}\right) = P[E \text{ covers } (\mu_1, \cdots, \mu_k)] = 1 - \alpha.$$

To obtain a rectangular confidence region of bounded confidence level, a rectangular parallelepiped, say R, with boundary planes parallel to the coordinate planes in the $\mu_1, \cdots, \mu_k$ space is circumscribed around the ellipsoid E. The boundary planes of R are found to be

$$\mu_i = \bar{y}_i \pm \frac{\sigma_i}{\sqrt{n}}\sqrt{c_\alpha},$$

and are not dependent on the correlations.

Then $P[R \text{ covers } (\mu_1, \cdots, \mu_k)] > P[E \text{ covers } (\mu_1, \cdots, \mu_k)] = 1 - \alpha$, thus giving a set of intervals, $\bar{y}_i \pm (\sigma_i/n^{\frac{1}{2}})c_\alpha^{\frac{1}{2}}$, with $U_k(c_\alpha) = 1 - \alpha$.

4.2. Unknown variances. The same method applies when the variances are unknown and $n > k$, using Hotelling's T-statistic. Here E is the ellipsoid $\sum_{i=1}^{k}\sum_{s=1}^{k} l^{is}(\bar{y}_i - \mu_i)(\bar{y}_s - \mu_s) = c_\alpha^2/n$ where (l^{is}) is the inverse of the matrix (l_{is}) and $l_{is} = \sum_{j=1}^{n}(y_{ij} - \bar{y}_i)(y_{sj} - \bar{y}_s)/(n-1)$, $i = 1, \cdots, k$; $s = 1, \cdots, k$. The boundary planes of R, the circumscribed parallelepiped, are $\mu_i = \bar{y}_i \pm (\hat{\sigma}_i/n^{\frac{1}{2}})c_\alpha$, where $\hat{\sigma}_i = l_{ii}^{\frac{1}{2}}$. For c_α defined by $F(c_\alpha) = 1 - \alpha$, with F the c.d.f. of Hotelling's T, the set of confidence intervals is $\bar{y}_i \pm (\hat{\sigma}_i/n^{\frac{1}{2}})c_\alpha$, $i = 1, 2, \cdots, k$.

It is to be noted that this is the only set of intervals given ln this paper for which no assumption has been made concerning the variances. For the other sets, the variances were assumed to be known or else to be unknown but equal.

4.3. More general distribution functions. For n large, T^2 can be assumed to follow a Chi-square distribution with k degrees of freedom, even though the original variables are not normally distributed [3]. A set of confidence intervals for $\mu_1, \cdots, \mu_k$ is then $\bar{y}_i \pm (\hat{\sigma}_i/n^{\frac{1}{2}})c_\alpha^{\frac{1}{2}}$, with c_α the upper α point of the Chi-square distribution with k degrees of freedom.

5. Bounded regions based on linear contrasts. Henry Scheffé [4] obtains simultaneous confidence intervals for the totality of linear contrasts among k means, $\mu_1, \cdots, \mu_k$, using the F distribution. He shows that $P(\hat{\theta} - S\hat{\sigma}_{\hat{\theta}} \leqq \theta \leqq \theta + S\hat{\sigma}_{\hat{\theta}}) = 1 - \alpha$. Here θ is any linear contrast; $S^2 = (k-1)c_\alpha$; c_α is the upper α point of the F distribution with $k - 1$ and ν degrees of freedom; ν is the degrees of freedom of the χ^2 variable used in estimating the variance; and P denotes the probability that all such intervals cover their corresponding contrasts.

It can easily be shown that confidence intervals for the totality of linear combinations of $\mu_1, \cdots, \mu_k$ are similarly obtained from $P(\theta - S\hat{\sigma}_{\hat{\theta}} \leqq \hat{\theta} \leqq \theta + S)\hat{\sigma}_{\hat{\theta}} = 1 - \alpha$, where now $S^2 = kc_\alpha$, with c_α the upper α point of the F distribution with k and ν degrees of freedom. Since the k means $\mu_1, \cdots, \mu_k$ are a subset of the linear combinations, confidence intervals for them follow immediately.

5.1. Variances known. If the variables $y_1, \cdots, y_k$ are normally distributed with unknown means $\mu_1, \cdots, \mu_k$, known variances $\sigma_1^2, \cdots, \sigma_k^2$, and unknown correlations, ρ_{is}, then the χ^2 distribution is used rather than the F distribution, and we have:

$$P\left(\bar{y}_1 - \frac{\sigma_1}{\sqrt{n}}\sqrt{c_\alpha} < \mu_1 < \bar{y}_1 + \frac{\sigma_1}{\sqrt{n}}\sqrt{c_\alpha}, \cdots, \bar{y}_k - \frac{\sigma_k}{\sqrt{n}}\sqrt{c_\alpha} < \mu_k < \bar{y}_k + \frac{\sigma_k}{\sqrt{n}}\sqrt{c_\alpha}\right) \geqq 1 - \alpha.$$

Here c_α is, as in section 4.1, the upper α point of the χ^2 distribution with k degrees of freedom, and the intervals obtained are the same as those of section 4.1.

5.2. Variances unknown but equal. When the variances are unknown but equal, then as an estimate of σ^2 one may use $\hat{\sigma}_1^2 = \sum_{j=1}^n (y_{1j} - \bar{y}_1)^2/(n-1)$. Then

$$P\left(\bar{y}_1 - \sqrt{\frac{k}{n}}\,\hat{\sigma}_1\sqrt{c_\alpha} < \mu_1 < \bar{y}_1 + \sqrt{\frac{k}{n}}\,\hat{\sigma}_1\sqrt{c_\alpha}, \cdots, \bar{y}_k - \sqrt{\frac{k}{n}}\,\hat{\sigma}_1\sqrt{c_\alpha} < \mu_k < \bar{y}_k + \sqrt{\frac{k}{n}}\,\hat{\sigma}_1\sqrt{c_\alpha}\right) \geqq 1 - \alpha,$$

with c_α the upper α point of the F distribution with k and $n - 1$ degrees of freedom. The confidence intervals are $\bar{y}_i \pm (k/n)^{\frac{1}{2}}\hat{\sigma}_1 c_\alpha^{\frac{1}{2}}$.

It may seem unsatisfactory to use only the data from one sample point as an estimate of σ^2; this has been done in order to have a χ^2 variable for the denominator of the F variable.

If one wishes to use a pooled estimate of the variance, $\hat{\sigma}_p^2 = \sum_{i=1}^k \hat{\sigma}_i^2/k$, then $\hat{\sigma}_p^2$ no longer has a χ^2 distribution because of the dependence of the variables. It is possible to show, however, that the F distribution may still be used, provided for degrees of freedom one uses k and $n - 1$ (rather than k and $k(n - 1)$). That the degrees of freedom may not be increased may be seen by examining the extreme case when all the correlations are equal to one.

To establish the necessary inequality for using $\hat{\sigma}_p^2$, one may fix $\hat{\sigma}_1, \cdots, \hat{\sigma}_k$ and consider the conditional probability

$$P(\bar{y}_1 - \sqrt{\frac{k}{n}}\hat{\sigma}_p\sqrt{c_\alpha} < \mu_1 < \bar{y}_1 + \sqrt{\frac{k}{n}}\hat{\sigma}_p\sqrt{c_\alpha}, \cdots, \bar{y}_k - \sqrt{\frac{k}{n}}\hat{\sigma}_p\sqrt{c_\alpha} < \mu_k < \bar{y}_1 + \sqrt{\frac{k}{n}}\hat{\sigma}_p\sqrt{c_\alpha} \mid \hat{\sigma}_1, \cdots, \hat{\sigma}_k\Big)$$

$$\geqq P\left(\bar{y}_1 - \sqrt{\frac{k}{n}\frac{\Sigma\hat{\sigma}_i}{k}}\sqrt{c_\alpha} < \mu_1 < \bar{y}_1 + \sqrt{\frac{k}{n}\frac{\Sigma\hat{\sigma}_i}{k}}\sqrt{c_\alpha}, \cdots, \bar{y}_k - \sqrt{\frac{k}{n}\frac{\Sigma\hat{\sigma}_i}{k}}\sqrt{c_\alpha} < \mu_k < \bar{y}_k + \sqrt{\frac{k}{n}\frac{\Sigma\hat{\sigma}_i}{k}}\sqrt{c_\alpha} \mid \hat{\sigma}_1, \cdots, \hat{\sigma}_k\right)$$

$$\geqq \sum_{i=1}^{k} P\left(\bar{y}_1 - \sqrt{\frac{k}{n}}\hat{\sigma}_i\sqrt{c_\alpha} < \mu_1 < \bar{y}_1 + \sqrt{\frac{k}{n}}\hat{\sigma}_i\sqrt{c_\alpha}, \cdots, \bar{y}_k - \sqrt{\frac{k}{n}}\hat{\sigma}_i\sqrt{c_\alpha} < \mu_k < \bar{y}_k + \sqrt{\frac{k}{n}}\hat{\sigma}_i\sqrt{c_\alpha} \mid \hat{\sigma}_1, \cdots, \hat{\sigma}_k\right)\Big/ k$$

Thus for the unconditional probability one has:

$$P\left(\bar{y}_1 - \sqrt{\frac{k}{n}}\hat{\sigma}_p\sqrt{c_\alpha} < \mu_1 < \bar{y}_1 + \sqrt{\frac{k}{n}}\hat{\sigma}_p\sqrt{c_\alpha}, \cdots, \bar{y}_k - \sqrt{\frac{k}{n}}\hat{\sigma}_p\sqrt{c_\alpha} < \mu_k < \bar{y}_k + \sqrt{\frac{k}{n}}\hat{\sigma}_p\sqrt{c_\alpha}\right) \geqq 1 - \alpha.$$

6. Regions based on a bonferroni inequality. Confidence regions can be obtained very simply using a Bonferroni inequality [5]. The use of this inequality in a related situation was suggested by E. Paulson [6].

6.1. Variances known. Let $n_k(y_1, \cdots, y_k; \mu_i, \sigma_i^2, \rho_{is})$ be the frequency function of k normally distributed variables with means $\mu_1, \cdots, \mu_k$, known variances $\sigma_1^2, \cdots, \sigma_k^2$, and unknown correlations ρ_{is}. Let $\bar{y}_i$ be the mean of a random sample of size n, $y_{i1}, \cdots, y_{in}$.

Let $z_i = ((\bar{y}_i - \mu_i)n^{\frac{1}{2}})/\sigma_i)$, $i = 1, \cdots, k$. Then the joint frequency function of $z_1, \cdots, z_k$ is $n_k(z_1, \cdots, z_k; 0, 1, \rho_{is})$, and

$$P(-c < z_1 < c, \cdots, -c < z_k < c) = \int_{-c}^{c} \cdots \int_{-c}^{c} n_k(z_1, \cdots, z_k; 0, 1, \rho_{is})\, dz_1 \cdots dz_k.$$

Using a Bonferroni inequality, this integral is greater than or equal to $1 - 2k(1 - N(c))$, where N is the c.d.f. of a standard normal variable. Setting this expression equal to $1 - \alpha$, c_α may be defined by $N(c_\alpha) = 1 - (\alpha/2k)$. Then

$$P\left(-c_\alpha < \frac{(\bar{y}_1 - \mu_1)\sqrt{n}}{\sigma_1} < c_\alpha, \cdots, -c_\alpha < \frac{(\bar{y}_k - \mu_k)\sqrt{n}}{\sigma_k} < c_\alpha\right) = P[R \text{ covers } (\mu_1, \cdots, \mu_k)] \geqq 1 - \alpha,$$

where R is bounded by

$$\mu_1 = \bar{y}_i \pm \frac{\sigma_i}{\sqrt{n}} c_\alpha.$$

6.2. Variances unknown but equal. Let y_i, $i = 1, \cdots, k$, have the joint frequency function $n_k(y_1, \cdots, y_k; \mu_i, \sigma^2, \rho_{is})$, where the variances are unknown but equal. Let $z_i = ((\bar{y}_i - \mu_i)n^{\frac{1}{2}})/\sigma$, $i = 1, \cdots, k$.

We wish to define Student t-variables $t_1, \cdots, t_k$ using $z_1, \cdots, z_k$ in the numerators and using the same Chi-square variable in the denominators. If $u_i = \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2/\sigma^2$, then u_i is a Chi-square variable with $n - 1$ degrees of freedom. Since the u_i are not independent of each other, we choose one, say u_1, to use in all the denominators, rather than use their sum which does not have a Chi-square distribution.

Then

$$t_i = \frac{\sqrt{n-1}\, z_i}{u_1^{1/2}} = \frac{\sqrt{n}(\bar{y}_i - \mu_i)}{\hat{\sigma}_1}, \qquad i = 1, \cdots, k,$$

are Student t-variables with the same denominators. Their distribution function [3] is;

$$f_{n-1}(t_1, \cdots, t_k; \rho_{is}) = \frac{\Gamma\left(\frac{k+n-1}{2}\right)}{(n-1)^{k/2}\pi^{k/2}\Gamma\left(\frac{n-1}{2}\right)} \, |(\rho^{is})|^{1/2}\left[1 + (n-1)^{-1}\sum_{i=1}^{k}\sum_{s=1}^{k}\rho^{is}t_it_s\right] - \frac{k+n-1}{2}$$

where ρ^{is} is an element of $(\rho^{is}) = (\rho_{is})^{-1}$, and $|(\rho^{is})|$ is the determinant of (ρ^{is}).

As in 6.1,

$$P(-c < t_1 < c, \cdots, -c < t_k < c)$$
$$= \int_{-c}^{c} \cdots \int_{-c}^{c} f_{n-1}(t_1, \cdots, t_k; \rho_{is})\, dt_1 \cdots dt_k \geqq 1 - 2k(1 - H_{n-1}(c)),$$

where H_{n-1} is the c.d.f. of a t-variable with $n - 1$ degrees of freedom.

The set of confidence intervals is then

$$\bar{y}_i \pm \frac{\hat{\sigma}_1}{\sqrt{n}} c_\alpha,$$

where

$$H_{n-1}(c_\alpha) = 1 - \frac{\alpha}{2k},$$

and

$$\hat{\sigma}_1 = \sum_{j=1}^{n} (y_{1j} - \bar{y}_1)^2/(n-1).$$

As in section 5.2, it is possible in these confidence intervals to replace $\hat{\sigma}_1$ by $\hat{\sigma}_p$, the pooled estimate of the variance; $n - 1$ must be retained as the degrees of freedom.

7. Regions with bounded confidence level using inequalities between dependent and independent cases.

7.1. Variances known. For $y_1, \cdots, y_k$ independently normally distributed with unknown means $\mu_1, \cdots, \mu_k$ and known variances, $\sigma_1^2, \cdots, \sigma_k^2$, let x_i be defined by $x_i = (n^{\frac{1}{2}}(\bar{y}_i - \mu_i))/\sigma_i$, where $\bar{y}_i$ is the mean of the n observations on the ith variable. Then

$$P(-c_\alpha < x_1 < c_\alpha, \cdots, -c_\alpha < x_k < c_\alpha) = \prod_{i=1}^{k} P(-c_\alpha < x_i < c_\alpha) = 1 - \alpha,$$

where c_α is defined by $N(c_\alpha) = \frac{1}{2}[1 + (1 - \alpha)^{1/k}]$, with N the c.d.f. of the univariate normal distribution. The set of simultaneous confidence intervals whose exact confidence level is $1 - \alpha$ is then $\bar{y}_i \pm \sigma_i c_\alpha / n^{\frac{1}{2}}$.

If, now, the $y_1, \cdots, y_k$ are defined as above except that now there may be correlations among them, the same confidence intervals can be used as a set with bounded confidence level, provided it can be proved that

$$P(-c_\alpha < x_1 < c_\alpha, \cdots, -c_\alpha < x_k < c_\alpha) \geqq 1 - \alpha.$$

The proof of the following theorem establishes this inequality for certain cases.

THEOREM. If $x_1, \cdots, x_k$ are normally distributed with zero means, unit variances, and correlations ρ_{is}, then

$$\int \underset{C}{\cdots} \int n_k(x_1, \cdots, x_k; 0, 1, \rho_{is})\, dx_1 \cdots dx_k \geqq \left[\int_{x=-c}^{x=c} n_1(x; 0, 1)\, dx\right]^k,$$

provided (1) $k = 2$ or 3; or (2) $\rho_{is} = b_i b_s$, for $i, s = 1, 2, \cdots, k, i \neq s$ and with $0 < b_i < 1, i = 1, 2, \cdots, k$. The region of integration C is the region bounded by the planes $x_i = \pm c, i = 1, \cdots, k$; $n_k(x_1, \cdots, x_k; 0, 1, \rho_{is})$ is the frequency function of $x_1, \cdots, x_k$; and $n_1(x; 0, 1)$ is the standard univariate normal frequency function.

PROOF. (1) $k = 2, 3$. For brevity the proof is merely outlined. The expression $\int \underset{C}{\cdots} \int n_k(x_1, \cdots, x_k; 0, 1, \rho_{is})\, dx_1, \cdots, dx_k$ may be regarded as a function of the ρ_{is}, say $F(\rho_{is})$. The proof consists in showing that for all admissible ρ_{is}, $F(\rho_{is})$ has an absolute minimum at the origin of the ρ_{is} space.

First it must be shown that there is a relative minimum at the origin. This can be shown for any k by considering the various first and second partial derivatives with respect to the correlations.

The first partial derivative with respect to ρ_{12}, say F_{12}, can be shown to be:

$$F_{12} = 2 \int_{x_3=-c}^{x_3=c} \cdots \int_{x_k=-c}^{x_k=c}$$
$$\cdot [n_k(c, c, x_3, \cdots, x_k; 0, 1, \rho_{is}) - n_k(c, -c, x_3, \cdots, x_k; 0, 1, \rho_{is})]\, dx_3, \cdots, dx_k \cdot$$

Similarly, the second derivative with respect to ρ_{12} and ρ_{pq}, say $F_{12,pq}$, is

$$F_{12,pq} = 2 \int_{x_3=-c}^{x_3=c} \cdots \int_{x_k=-c}^{x_k=c} n_k(c, c, x_3, \cdots, x_k; 0, 1, \rho_{is})$$
$$\cdot \left[\rho^{pq} + \left(\sum_{\substack{i=1\\x_1=c\\x_2=c}}^{k} \rho^{pi} x_i\right)\left(\sum_{\substack{i=1\\x_1=c\\x_2=c}}^{k} \rho^{qi} x_i\right)\right]$$
$$dx_3, \cdots, dx_k \text{—a similar integral with } x_1 = c, x_2 = -c.$$

When all the ρ_{is}'s are zero, it is easily seen that F_{12} vanishes. Further, $F_{12,pq}$ vanishes also at the origin unless $p = 1$ and $q = 2$, while $F_{12,12}$ is seen to be positive.

Thus in the expansion of $F(\rho_{is})$ about the origin, the first degree terms vanish and the second degree terms form a positive definite quadratic form, so that $F(\rho_{is})$ has a relative minimum at the origin for any k.

The next part of the proof is to show from the form of the first derivative, that at any point beside the origin, at least one of the first derivatives differ from zero. This was done only for $k = 2$ and 3.

The set of all admissible points (points such that (ρ_{is}) is positive definite and $0 < |(\rho_{is})| < 1$), together with the boundary points, form a compact set, so that $F(\rho_{is})$ must assume an absolute minimum either at an admissible point or at a boundary point. Hence if it can be shown that no point on the boundary of the set yields an absolute minimum, then the absolute minimum of F must be at the origin.

For $k = 2$, the boundary points are just $\rho_{12} = \pm 1$, and they actually yield absolute maxima for $F(\rho_{12})$.

For $k = 3$, a boundary point, say $(\rho_{12}, \rho_{13}, \rho_{23})$ was considered. It was shown that for m sufficiently close to 1 but less than 1, $(m\rho_{12}, m\rho_{13}, \rho_{23})$ is an admissible point, and that the derivative of F at $(m\rho_{12}, m\rho_{13}, \rho_{23})$ in the direction of $(\rho_{12}, \rho_{13}, \rho_{23})$ is positive. Hence $(\rho_{12}, \rho_{13}, \rho_{23})$ cannot yield an absolute minimum of F.

This completes the outline of the proof for $k = 2$ and 3, with any correlation matrix.

(2) For any k, if $\rho_{is} = b_i b_s$, with $0 < b_i < 1$ for $i = 1, \cdots, k$, a proof may be given which is adapted from the proof of a similar theorem by C. W. Dunnett and M. Sobel [3].

For $y_0, y_1, \cdots, y_k$ independently normally distributed, with zero means and unit variances, define

$$x_i = \sqrt{1 - b_i^2}\, y_i - b_i y_0, \qquad i = 1, \cdots, k.$$

Then the x_i's are normally distributed with means zero, unit variances, and correlations $\rho_{is} = b_i b_s$.

The theorem may be restated as follows:

$$P(-c < x_1 < c, \cdots, -c < x_k < c) \geqq \prod_{i=1}^{k} P(-c < x_i < c),$$

or

$$P(-c < \sqrt{1 - b_1^2}\, y_1 - b_1 y_0 < c, \cdots, -c < \sqrt{1 - b_k^2}\, y_k - b_k y_0 < c) \geqq \prod_{i=1}^{k} P(-c < \sqrt{1 - b_i^2}\, y_i - b_i y_0 < c).$$

or

$$P(d_1 < y_1 < e_1, \cdots, d_k < y_k < e_k) \geqq \prod_{i=1}^{k} P(d_i < y_i < e_i),$$

where

$$d_i = \frac{-c + b_i y_0}{\sqrt{1 - b_i^2}}, \qquad e_i = \frac{c + b_i y_0}{\sqrt{1 - b_i^2}}, \qquad i = 1, \cdots, k.$$

This may be written as:

$$\int_{y_0=-\infty}^{y_0=\infty} \left[\int_{y_1=d_1}^{y_1=e_1} \cdots \int_{y_k=d_k}^{y_k=e_k} n_k(y_1, \cdots, y_k; 0, 1, 0)\, dy_1, \cdots, dy_k \right] n_1(y_0; 0, 1)\, dy_0$$
$$\geqq \prod_{i=1}^{k} \int_{y_0=-\infty}^{y_0=\infty} \left[\int_{y_i=d_i}^{y_1=e_i} n_1(y_i; 0, 1)\, dy_i \right] n_1(y_0; 0, 1)\, dy_0,$$

or

$$\int_{y_0=-\infty}^{y_0=\infty} \left[\prod_{i=1}^{k} F_i(y_0) \right] n_1(y_0; 0, 1)\, dy_0 \geqq \prod_{i=1}^{k} \int_{y_0=-\infty}^{y_0=\infty} F_i(y_0) n_1(y_0; 0, 1)\, dy_0,$$

where

$$F_i(y_0) = \int_{d_i}^{e_i} n_1(y_i; 0, 1)\, dy_i.$$

Thus the inequality becomes:

$$E\left(\prod_{i=1}^{k} F_i(y_0)\right) \geqq \prod_{i=1}^{k} E(F_i(y_0)).$$

The expected value of a product of monotone bounded functions is greater than or equal to the product of their expected values [6], so that the last inequality would hold if the F_i were monotone. The functions $F_i(y_0)$, however, are seen to increase from $-\infty$ to 0 and to decrease from 0 to ∞. Since the frequency function of y_0 is symmetric about the origin, the transformation $z = |y_0|$ changes the inequality to

$$E\left(\prod_{i=1}^{k} F_i(z)\right) \geqq \prod_{i=1}^{k} E(F_i(z)),$$

where $F_i(z)$ are monotonically decreasing bounded functions. This completes the proof of the theorem.

7.2. Variances unknown but equal. When the variances are unknown but equal, Student t-variables t_i with the joint frequency function

$$f_{n-1}(t_1, \cdots, t_k; \rho_{is}),$$

as defined in 6.2, are used to form confidence intervals. Using the same methods as in 7.1, the following theorem can be proved:

THEOREM. *For* $k = 2$ or 3,

$$\int \underset{C}{\cdots} \int f_{n-1}(t_1, \cdots, t_k; \rho_{is})\, dt_1 \cdots dt_k \geqq \int \underset{C}{\cdots} \int f_{n-1}(t, \cdots, t_k; 0)\, dt_1 \cdots dt_k$$

For $\rho_{is} = b_i b_s$, with $0 < b_i < 1, i = 1, \cdots, k$,

$$\int \underset{C}{\cdots} \int f_{n-1}(t_1, \cdots, t_k; \rho_{is})\, dt_1 \cdots dt_k \geqq \left[\int_{-c}^{c} f_{n-1}(t)\, dt\right]^k.$$

In this theorem, whose proof follows the same lines as the one in 7.1, C is the region bounded by $t_i = \pm c$, $i = 1, \cdots, k$, $f_{n-1}(t)$ is the density function of a Student t-variable with $n - 1$ degrees of freedom, and $f_{n-1}(t_1, \cdots, t_k; 0)$ is the joint frequency function of the t-variables when all ρ_{is} are zero.

Since

$$t_i = \frac{\sqrt{n}(\bar{y}_i - \mu_i)}{\sigma_1}, \qquad i = 1, \cdots, k,$$

sets of confidence intervals obtained are as follows:

For $k = 2$ or 3, $\bar{y}_1 \pm (\hat{\sigma}_1/n^{\frac{1}{2}})c_\alpha$, where c_α is defined by

$$\int_{-c_\alpha}^{c_\alpha} \cdots \int_{-c_\alpha}^{c_\alpha} f_{n-1}(t_1, \cdots, t_k; 0)\, dt_1 \cdots dt_k = 1 - \alpha.$$

For any k and $\rho_{is} = b_i b_s$, $0 < b_i < 1$, $i = 1, \cdots, k$, the same set is obtained, but with c_α defined by $H_{n-1}(c_\alpha) = (1 + (1 - \alpha)^{1/k})/2$, where H_{n-1} is the c.d.f. of a Student t-variable with $n - 1$ degrees of freedom. As in sections 5.2 and 6.2 one may use $\hat{\sigma}_p^2$ in place of $\hat{\sigma}_1^2$, provided one keeps $n - 1$ as the degrees of freedom.

8. Comparison of confidence intervals. In Table I are listed various sets of confidence intervals, with their properties and restrictions.

One rather obvious way to compare them is by comparing their lengths, or the expected values of the lengths. In Table II are given numerical values of d_α for $1 - \alpha = .95$, where

$$d_\alpha = \frac{\sqrt{n}}{\sigma} \sqrt{E(\tfrac{1}{2}\ell)^2},$$

with ℓ the length of the confidence interval. Throughout Table II, the variances are assumed to be equal.

When the variances are known and equal, and all the correlations are zero, the shortest set of confidence intervals must be those of section 7.1. When nothing is known about the correlations, no shorter set can be obtained. The last column in section 7 of Table II therefore gives the smallest obtainable values for d_α, and may be used as a standard for comparison.

For $1 - \alpha = .95$, the Bonferroni inequality intervals of section 6 are almost as good as the best ones. Indeed for $1 - \alpha$ as low as .80, the values of d_α are still very close, being:

k	Bonferroni	"Best"
1	1.28	1.28
2	1.64	1.61
4	1.96	1.92
6	2.13	2.09
8	2.24	2.20
10	2.33	2.29

TABLE I

Confidence Intervals for Means of Dependent, Normally Distributed Variables

Section	Confidence Intervals	Definition of c_α	Conditions
3.1	$\sum_{j=1}^{n} a_{ji} y_{ij} \pm \frac{k}{n} \sigma_i \cdot c_\alpha$	$N(c_\alpha) = \frac{1 + (1-\alpha)^{1/k}}{2}$	$n \geqq k$ (1)
3.2	$\sum_{j=1}^{n} a_{ji} y_{ij} \pm \sqrt{\frac{k}{n(n-k)} \sum_{m=1}^{n-k} u_m^2} \cdot c_\alpha$	$H_{n-k}(c_\alpha) = \frac{1 + (1-\alpha)^{1/k}}{2}$	$n > k$ (2, 3)
4.1	$\bar{y}_i \pm \frac{\sigma_i}{\sqrt{n}} \cdot \sqrt{c_\alpha}$	$U_k(c_\alpha) = 1 - \alpha$	(4)
4.2	$\bar{y}_i \pm \frac{\hat{\sigma}_i}{\sqrt{n}} \cdot c_\alpha$	$F(c_\alpha) = 1 - \alpha$	$n > k$ (5)
5.1	$\bar{y}_i \pm \frac{\sigma_i}{\sqrt{n}} \sqrt{c_\alpha}$	$U_k(c_\alpha) = 1 - \alpha$	(4)
5.2	$\bar{y}_i \pm \frac{\hat{\sigma}_1}{\sqrt{n}} \sqrt{c_\alpha}$	$F_{k,n-1}(c_\alpha) = 1 - \alpha$	(6)
6.1	$\bar{y}_i \pm \frac{\sigma_i}{\sqrt{n}} \cdot c_\alpha$	$N(c_\alpha) = 1 - \frac{\alpha}{2k}$	(1)
6.2	$\bar{y}_i \pm \frac{\hat{\sigma}_1}{\sqrt{n}} \cdot c_\alpha$	$H_{n-1}(c_\alpha) = 1 - \frac{\alpha}{2k}$	(2)
7.1	$\bar{y}_i \pm \frac{\sigma_i}{\sqrt{n}} \cdot c_\alpha$	$N(c_\alpha) = \frac{1 + (1-\alpha)^{1/k}}{2}$ $k = 2, 3,$ or $\rho_{is} = b_i b_s$	(1)
7.2	$\bar{y}_i \pm \frac{\hat{\sigma}_1}{\sqrt{n}} \cdot c_\alpha$	$H_{n-1}(c_\alpha) = \frac{1 + (1-\alpha)^{1/k}}{2}$ $k = 2, 3,$ or $\rho_{is} = b_i b_s$	 (2, 3) (2)

(1) N is the cumulative standard normal distribution function.

(2) H_ν is the cumulative distribution function of a Student t-variable with ν degrees of freedom.

(3) This definition of c_α is approximate. The exact definition is:

$$\int_{-c_\alpha}^{c_\alpha} \cdots \int_{-c_\alpha}^{c_\alpha} f_\nu(t_1, \cdots, t_k)\, dt_1, \cdots, dt_k = 1 - \alpha, \quad \text{where } f_\nu(t_1, \cdots, t_k)$$

$$= \frac{\Gamma\left(\frac{k+\nu}{2}\right)}{\nu^{k/2}\pi^{k/2}\Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{\sum_{i=1}^{k} t_i^2}{\nu}\right]^{-(k+\nu)/2}, \quad \text{where}$$

ν is the degrees of freedom of t_i.

(4) U_k is the cumulative distribution function of a Chi-square variable with k degrees of freedom.

(5) F is the cumulative distribution function of Hotelling's T.

(6) $F_{k,n-1}$ is the cumulative distribution function of an F variable with k and $n - 1$ degrees of freedom.

TABLE II

*Comparison of Lengths of Confidence Intervals for Means of Dependent, Normally Distributed Variables with Equal Variances, $1 - \alpha = .95$**

k	n							
	Variances Unknown						Variances Known	
	Section	4	6	8	10	20	Section	Any
	4.1						4.2	
1		3.18	2.57	2.36	2.26	2.09		1.96
2		10.8	5.52	4.53	4.12	3.55		3.17
4			27.9	11.4	8.63	6.13		4.98
6				49.3	17.8	8.75		6.44
8					77.0	11.8		7.72
10						15.6		8.85
	5.1						5.2	
1		3.18	2.57	2.36	2.26	2.09		1.96
2		7.55	4.16	3.47	3.17	2.74		2.45
4			13.9	6.70	5.21	3.79		3.08
6				20.1	9.12	4.82		3.55
8					26.4	6.01		3.94
10						7.53		4.28
	6.1						6.2	
1		3.18	2.57	2.36	2.26	2.09		1.96
2		4.37	3.40	3.08	2.92	2.66		2.45
4		6.04	4.56	4.06	3.81	3.41		3.08
6		7.32	5.45	4.82	4.50	3.98		3.55
8		8.41	6.21	5.37	5.08	4.46		3.94
10		9.38	6.88	6.03	5.60	4.89		4.28
	7.1						7.2	
1		3.18	2.57	2.36	2.26	2.09		1.96
2		4.17	3.16	2.84	2.68	2.44		2.24
4		5.41	3.80	3.33	3.11	2.76		2.50
6		6.22	4.22	3.64	3.36	2.94		2.64
8		6.92	4.53	3.86	3.55	3.07		2.74
10		7.47	4.77	4.03	3.69	3.17		2.81
	8.1						8.2	
1		3.18	2.57	2.36	2.26	2.09		1.96
2		4.16	3.15	2.83	2.68	2.43		2.24
4		5.35	3.79	3.32	3.10	2.75		2.49
6		6.17	4.20	3.62	3.35	2.94		2.63
8		6.86	4.50	3.84	3.53	3.07		2.73
10		7.40	4.76	4.01	3.67	3.16		2.80

* The figures given in the table are values of $(n^{\frac{1}{2}}/\sigma)\sqrt{E(\frac{1}{2}\ell)^2}$, where ℓ is the length of the confidence interval.

It would be interesting to show that the "best" intervals can be used for arbitrary k and arbitrary correlations, but from a practical viewpoint, for $1 - \alpha$ large enough to be of interest, the Bonferroni regions are good enough.

The regions of section 5, based on the T-distribution and the χ^2 distribution, compare favorably only when k is small and n relatively large. The regions with exact confidence level are everywhere unnecessarily long.

REFERENCES

[1] PAUL G. HOEL, "Confidence bands for polynomial curves," *Ann. Math. Stat.*, Vol. 25, (1954), pp. 534–542.

[2] C. W. DUNNET AND M. SOBEL, "Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution," *Biometrika*, Vol. 42, (1955), pp. 258–260.

[3] H. B. MANN AND A. WALD, "On stochastic limit and order relationships," *Ann. Math. Stat.*, Vol. 14, (1943), p. 224.

[4] HENRY SCHEFFÉ, "A method of judging all contrasts in the analysis of variance," *Biometrika*, Vol. 40, (1953), pp. 87–104.

[5] W. FELLER, *An introduction to probability theory and its applications*. I, p. 75, John Wiley and Sons, 1950.

[6] E. PAULSON, "On the comparison of several experimental categories with a control," *Ann. Math. Stat.*, Vol. 23, (1952), p. 249.

[7] A. W. KIMBALL, "On dependent tests of significance in the analysis of variance," *Ann. Math. Stat.*, Vol. 22, (1951), pp. 600–602.