SIMULTANEOUS TOLERANCE INTERVALS
IN REGRESSION

By
GERALD J. LIEBERMAN and RUPERT G. MILLER, JR.

TECHNICAL REPORT NO. 59
APRIL 30, 1962

SUPPORTED BY THE ARMED SERVICES UNDER
CONTRACT Nonr-225 (53) (NR-042-002)
WITH THE OFFICE OF NAVAL RESEARCH
SIMULTANEOUS TOLERANCE INTERVALS
IN REGRESSION
by
Gerald J. Lieberman and Rupert G. Miller, Jr.

TECHNICAL REPORT NO. 59
April 30, 1962

Supported by the Army, Navy and Air Force under
Contract Nonr-225(53) (NR-042-002)
with the Office of Naval Research

Gerald J. Lieberman, Project Director

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

APPLIED MATHEMATICS AND STATISTICS LABORATORIES
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
0. Summary

Joint prediction intervals (based upon the original fitted model) for \( K \) future responses at each of \( K \) separate settings of the independent variable have been treated by Lieberman [1]. When \( K \) is unknown and possibly arbitrarily large, these results do not apply. A solution to the problem of arbitrary \( K \) is given in terms of tolerance intervals on the distributions of future observations, the intervals being (probabilistically) simultaneous in each possible value of the independent variable. Four alternative techniques are proposed and compared for their applicability in different situations. The first is the simultaneous extension of the Wallis [2] technique. The other three are based on Scheffé simultaneous confidence principles. One gives intervals for a fixed central proportion \( P \) of the distribution which are simultaneous in all values of the independent variable; the other two give intervals simultaneous in the independent variable and different central proportions \( P \). A numerical example is analyzed, and some remarks are made on the applicability of the Scheffé techniques to the detection of outliers.
1. Introduction

Determining the relationship between several variables is a problem which often arises in engineering and other fields. In particular, a linear relationship of the form

$$E(Y) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p$$

is usually assumed, and estimates of $\beta_0, \beta_1, \ldots, \beta_p$ are obtained from the sample data. More important, however, these results are sometimes used to make statements about future values of $Y$ for given values of the $x$'s. For example, consider the following hypothetical situation. The speed of a missile is a critical factor which can be determined only by measurement after firing. On the other hand, the orifice opening of the valve which admits the fuel is easily obtained by bench tests. Suppose that there exists an underlying relationship between these two variables of the form

$$\text{Expected speed} = \beta_0 + \beta_1 (\text{orifice opening}),$$

where speed corresponds to $Y$ and the orifice opening corresponds to $x_1 (p = 1)$. $N$ missiles are produced and fired so that $N$ pairs of observations, $(x_{11}, y_1), (x_{12}, y_2), \ldots, (x_{1N}, y_N)$ are taken. Estimates of $\beta_0$ and $\beta_1$ can then be obtained by the method of least squares.

If a new missile is produced, a prediction of its speed based on the knowledge of its orifice opening may be desirable. The usual techniques can be used to obtain a prediction interval such that the speed of the missile, having a given orifice opening $x^*$, will lie in the interval with preassigned probability, say 0.95. Expressions for such an interval can be obtained from any standard text (see [3], Chapter 9).
Unfortunately, it is often the case that this statement is not the appropriate one for the problem at hand; e.g., many new missiles are to be produced, and on the basis of the original \( N \) pairs of observations, it is desired to predict the speed of each new missile from the individual orifice openings. Using the results for the one missile case over and over again is incorrect since the prediction intervals are not independent. Furthermore, even if the intervals were independent, the number of future statements may be unknown so that an exact probability statement is not possible.

The problem of determining the joint prediction interval for the responses at each of \( K \) (known) separate settings of the independent variable when all \( K \) predictions are based upon the original fitted model has been treated by Lieberman [1], but these results do not extend to the case where \( K \) is unknown and possibly arbitrarily large. However, this problem can be solved by means of statistical tolerance intervals on the normal distributions of future observations, the intervals being (probabilistically) simultaneous in each possible value of the independent variable.

Tolerance intervals \([L_x(P), U_x(P)]\) are sought which are based upon the same estimated linear regression and which have the property that the interval \([L_x(P), U_x(P)]\) contains 100P\% of the normal distribution centered at \( \beta_0 + \beta_1x \) for any \( x \) and any or fixed \( P \) with confidence coefficient \( 1 - \alpha \). Interpreted in terms of predictions, a future value of \( Y \), given the independent variable \( x \), will lie in \([L_x(P), U_x(P)]\) with probability at least \( P \) with confidence \( 1 - \alpha \) for any \( x \) and any or fixed \( P \). This will enable the experimenter to use the same
estimated regression line over and over, and still make exact probability
statements associated with his prediction intervals.

Four alternative techniques are proposed and compared for their
applicability in different situations. The first is the simultaneous ex-
tension of the Wallis [2] technique. Wallis has considered the problem
where the future value of $x$ is fixed (as is the value of $P$). His
method is extended to the case where $x$ is allowed to vary.

The second technique is based on the Scheffé simultaneous confidence
principle [4]. Tolerance intervals are obtained for central (i.e., sym-
metric about the mean) proportions $P$ of the distribution which are
simultaneous in all values of the independent variable $x$ and all values
of the proportion $P$.

The third technique is also based on the Scheffé simultaneous con-
fidence principle. Tolerance intervals are obtained for a fixed central
proportion which are simultaneous in the values of the independent vari-
able $x$.

The last technique is the simplest of the four. It combines the
Working-Hotelling-Scheffé simultaneous confidence region for the regres-
sion line with a one-sided confidence interval on the standard deviation
through a Bonferroni inequality. The resulting intervals are simultane-
ous in both $x$ and $P$.

No special tables other than the cumulative normal and chi-square
distributions are required for the simultaneous Wallis procedure although
in any application it is necessary to solve numerically a transcendental
equation. The last technique requires only the cumulative $F$ and
chi-square tables. The other two techniques do require previously
untabulated percentile constants. For the first of these the constant \( c^* \) is tabulated for degrees of freedom \( DF = 1(1)30, 30(5), 50, 50(10)100 \) and \( \alpha = .001, .01, .05, .10, .30 \) and \( .50 \) in the case \( p = 1 \) (i.e., \( E(Y) = \beta_0 + \beta_1 x \)). For the other the constant \( c^{**} \) is tabulated for \( N = 4(2)20, P = .50, .75, .90, .95, .99, \) and \( .999, \) and \( \alpha = .01, .05, \) and \( .10 \) in the case \( p = 1 \).

In the next to last section these four procedures are compared in a numerical example with hypothetical data in the speed-orifice problem. The lengths of these intervals are contrasted with the shorter intervals for the non-simultaneous Wallis procedure. The last section contains a short discussion on the application of the Scheffé techniques to the detection of outliers in regression.
2. Underlying model

Throughout this paper it will be assumed that

\[ Y = X\beta + e = (1, X_1)\beta + e \]

where \( X \) is a known \( N \times (p + 1) \) matrix of rank \( p + 1 \) \((N > p + 1)\) and \( \beta = \text{col} (\beta_0, \beta_1, \ldots, \beta_p) \) is the \((p + 1) \times 1\) vector of unknown regression parameters. The first column of \( X \) consists entirely of ones (i.e., \( \beta_0 \) is a mean constant) and without loss of generality it will be assumed that the means of the other columns \( \bar{x}_1, \ldots, \bar{x}_p \) are all zero. The error vector \( e \) is assumed to be normally distributed with zero mean and covariance matrix \( \sigma^2 I \) (i.e., independence).

It should be noted in passing that the results of the subsequent sections can easily be modified to the case where \( \beta_0 \) is absent, i.e., where the regression plane is forced through the origin.

The tolerance intervals will be based on the customary (least squares) estimators:

\[
\begin{align*}
\mathbf{b} &= (X' X)^{-1} X' \bar{y} + (X_1' X_1)^{-1} X_1' y \\
\sigma^2 &= \frac{1}{N-p-1} \sum_{i=1}^{N} (y_i - \bar{y} - x_{i1} b_1 - \cdots - x_{ip} b_p)^2
\end{align*}
\]

The vector \( \mathbf{b} \) is normally distributed with mean vector \( \beta \) and covariance matrix

\[
\sigma^2 (X'X)^{-1} = \sigma^2 \begin{pmatrix}
\frac{1}{N} & 0 \\
0 & (X_1' X_1)^{-1}
\end{pmatrix}.
\]
The variable \((N-p-1)s^2/\sigma^2\) has a \(\chi^2\) distribution on \(N-p-1\) degrees of freedom and is independent of \(b\).

This paper will be concerned with tolerance and prediction statements for future observations with mean values given by different combinations of the independent variables, i.e., of the form

\[
E(Y) = \beta_0 + x_1 \beta_1 + \ldots + x_p \beta_p ,
\]

or to avoid confusion with the already observed \(x\) values in the \(X\) matrix,

\[
E(Y) = \beta_0 + \ell_1 \beta_1 + \ldots + \ell_p \beta_p = \ell'\beta .
\]

Since all possible combinations of the independent variables will appear in the probability statements, let

\[\mathcal{L} = \{\ell = \text{col}(\ell_0, \ell_1, \ldots, \ell_p) | \ell_0 = 1, \ell_1, \ldots, \ell_p \text{ arbitrary}\}.\]

The probability statements will then concern \(\ell'\beta\) for all \(\ell \in \mathcal{L}\). For the special case \(p = 1\)

\[\mathcal{L}_1 = \{\ell = \text{col}(1, x) | x \text{ arbitrary}\},\]

and in this case the linear combinations will still be denoted by \(a + \beta x\) instead of \(\beta_0 + \ell_1 \beta_1\).

When \(\beta_0\) is absent the appropriate space of linear combinations for the probability statements is

\[\mathcal{L}_1 = \{\ell = \text{col}(\ell_1, \ldots, \ell_p) | \ell_1, \ldots, \ell_p \text{ arbitrary}\}.\]

For notational convenience let

\[
\Phi(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z} e^{-\frac{z^2}{2}} dz .
\]
3. **Extension of the Wallis procedure for finding tolerance intervals which are simultaneous in all values of** $x$.

Wallis [2] first considered the problem of finding tolerance intervals in linear regression. He treats the case $p = 1$ (i.e., $E(Y) = \beta_0 + \beta_1 x$), and his results are applicable for the situation where the future value of $x$ is fixed (as in the value of $p$). This section will extend his technique to the case where $x$ is allowed to vary by finding a constant $c$ such that for a fixed proportion $P$ the tolerance interval

$$L_x = b_0 + b_1 x - cs \sqrt{\frac{1}{N} + \frac{x^2}{\Sigma x_i^2}},$$

(3.1)

$$U_x = b_0 + b_1 x + cs \sqrt{\frac{1}{N} + \frac{x^2}{\Sigma x_i^2}},$$

has the property that

(3.2) \[ P \left[ A_x \geq P, \text{ for all } x \right] = 1 - \alpha, \]

where

(3.3) \[ A_x = \phi \left( \frac{U_x - \beta_0 - \beta_1 x}{\sigma} \right) - \phi \left( \frac{L_x - \beta_0 - \beta_1 x}{\sigma} \right). \]

Thus, the interval $(L_x, U_x)$ has the property that (with confidence $1 - \alpha$) for all $x$ it captures a proportion $P$ of the normal distribution within it.

Without loss of generality assume $\beta_0 = \beta_1 = 0$, $\sigma = 1$. Let

(3.4) \[ A^*(b_0, b_1, s) = \min x \frac{1}{\sqrt{2\pi}} \int_{b_0 + b_1 x - cs}^{b_0 + b_1 x + cs} \sqrt{\frac{1}{N} + \frac{x^2}{\Sigma x_i^2}} e^{-\frac{x^2}{2}} dz. \]

A constant $c$ is sought such that
\begin{align}
\tag{3.5}
P \{ A^* (b_0', b_1, s) \geq P \} = 1 - \alpha .
\end{align}

Equation (3.5) can be written as

\begin{align}
\tag{3.6}
E_{b_0', b_1} \left[ P \{ A^* (b_0', b_1, s) \geq P | b_0', b_1 \} \right] = 1 - \alpha ,
\end{align}

and the integrand can be expanded in a power series:

\begin{align}
P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \}
= &\ P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \} _{0,0} + b_0 \frac{\partial}{\partial b_0} P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \} _{0,0} \\
+ &\ b_1 \frac{\partial}{\partial b_1} P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \} _{0,0} \\
+ &\ \frac{1}{2} \left[ b_0^2 \frac{\partial^2}{\partial b_0^2} P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \} _{0,0} \\
+ &\ 2 b_0 b_1 \frac{\partial^2}{\partial b_0 \partial b_1} P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \} _{0,0} \\
+ &\ b_1^2 \frac{\partial^2}{\partial b_1^2} P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \} _{0,0} \right] + o (b_0^2, b_1^2).
\end{align}

(The existence of all derivatives can be justified by the implicit function theorem.)

Since simultaneously changing $b_0$ to $-b_0$ and $x$ to $-x$ leaves the value of $A_x$ unchanged, $A^* (-b_0, b_1, s) = A^* (b_0, b_1, s)$; i.e., $A^*(b_0', b_1, s)$ is symmetric in $b_0$ about 0 for each $b_1$ and $s$. Therefore,

\begin{align}
\tag{3.8}
P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \} = P \{ A^*(-b_0, b_1, s) \geq P | -b_0', b_1 \},
\end{align}

and

\begin{align}
\tag{3.9}
\frac{\partial}{\partial b_0} P \{ A^* (b_0, b_1, s) \geq P | b_0', b_1 \} _{0,0} = 0 .
\end{align}
Also, since simultaneously changing $b_1$ to $-b_1$ and $x$ to $-x$ leaves $A_x$ unchanged, $A^*(b_0, -b_1, s) = A^*(b_0, b_1, s)$. Hence, $A^*(b_0, b_1, s)$ is symmetric in $b_1$ about 0 for each $b_0$ and $s$, and

$$\frac{\partial}{\partial b_1} P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}_{0,0} = 0.$$  \hspace{1cm} (3.10)

Similarly by virtue of the symmetry of $A^*(b_0, b_1, s)$ in $b_0$ and $b_1$ about 0 for each $s$,

$$\frac{\partial^2}{\partial b_0 \partial b_1} P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}_{0,0} = 0.$$  \hspace{1cm} (3.11)

From (3.7), (3.9), (3.10) and (3.11), $P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}$ can be written as

$$P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \} = P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}_{0,0}$$

$$+ \frac{1}{2} \frac{\partial}{\partial b_0} \frac{\partial^2}{\partial b_0^2} P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}_{0,0}$$

$$+ \frac{1}{2} \frac{\partial}{\partial b_1} \frac{\partial^2}{\partial b_1^2} P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}_{0,0} + o(b_0^2, b_1^2),$$  \hspace{1cm} (3.12)

and

$$P_{b_0, b_1} \{P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \} = P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}_{0,0}$$

$$+ \frac{1}{2} \frac{1}{N} \frac{\partial^2}{\partial b_0^2} P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}_{0,0}$$

$$+ \frac{1}{2} \frac{1}{\Sigma x_i^2} \frac{\partial^2}{\partial b_1^2} P \{A^*(b_0, b_1, s) \geq P|b_0, b_1| \}_{0,0} + o(\frac{1}{N}),$$  \hspace{1cm} (3.13)

(assuming $\Sigma x_i^2 = O(N)$).
When the two power series expansions are equated,

\[(3.14) \quad P \{ A^*(b_0, b_1, s) \geq P \} = P \{ A^*(b_0, b_1, s) \geq P | b_0 = \frac{1}{\sqrt{N}}, b_1 = \frac{1}{\sqrt{\Sigma x_1^2}} \}, \]

so the problem reduces to choosing \( c \) so that

\[(3.15) \quad P \{ A^* \left( \frac{1}{\sqrt{N}}, \frac{1}{\sqrt{\Sigma x_1^2}}, s \right) \geq P \} = 1 - \alpha. \]

By definition equality holds for the event in the expression (3.15) when

\[(3.16) \quad \min_x \frac{1}{\sqrt{2\pi}} \int \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{\Sigma x_1^2}} x + cs \sqrt{\frac{1}{N} + \frac{x^2}{\Sigma x_1^2}} e^{-\frac{x^2}{2}} dz = P. \]

The minimum in (3.16) is achieved at \( x^* \) where \( x^* \) satisfies the transcendental equation

\[(3.17) \quad \exp \left\{ -\frac{1}{2} \left( \frac{1}{\sqrt{N}} + \frac{x}{\sqrt{\Sigma x_1^2}} + cs \sqrt{\frac{1}{N} + \frac{x^2}{\Sigma x_1^2}} \right)^2 \right\} \left\{ \frac{1}{\sqrt{\Sigma x_1^2}} + cs \sqrt{\frac{1}{N} + \frac{x^2}{\Sigma x_1^2}} \left( \frac{x}{\Sigma x_1^2} \right) \right\} = 0. \]

The solution \( x^* \) will be a function of \( cs \) (i.e., \( x^* = f(cs) \)) and \( A^* \left( \frac{1}{\sqrt{N}}, \frac{1}{\sqrt{\Sigma x_1^2}}, s \right) \) will be a monotonically non-decreasing function of \( cs \) (i.e., \( A^* = g(cs) \)). Since \( g(0) = 0 \) and \( g(+\infty) = 1 \), there exists a constant \( h_0 \) such that \( g(h_0) = P \).
If \( c \) is chosen so that

\[
(3.18) \quad P\left(s > \frac{h_0}{c}\right) = 1 - \alpha,
\]

then

\[
(3.19) \quad P\left(A^* \frac{1}{\sqrt{N}}, \frac{1}{\sqrt{\sum x_i^2}}, s \right) \geq P = 1 - \alpha.
\]

Once \( h_0 \) is found by numerical methods \( c \) can be obtained directly since \((N-2)s^2\) has a chi-square distribution with \( N-2 \) degrees of freedom.

For simplicity the simultaneous Wallis technique was presented for the case of simple linear regression, but the technique extends directly to multiple regression. The only substantial modification is that \((3.17)\) is replaced by a system of transcendental equations which must be solved to locate the minimum value.
4. Tolerance intervals for different central proportions \( P \) which are simultaneous in all values of \( x \).

Although the presentation of the previous section pertained to the special case of \( p = 1 \) (i.e., \( E(Y) = \beta_0 + \beta_1 x \)), the results can be generalized for any \( p \). Thus, the technique can be used to construct the tolerance interval

\[
L_\ell = \ell'b - c \sqrt{\ell'(X'X)^{-1}\ell},
\]

\[
U_\ell = \ell'b + c \sqrt{\ell'(X'X)^{-1}\ell},
\]

(4.1)

with the property that

\[
P \{ A_\ell \geq P, \ for \ all \ \ell \in \mathcal{I} \} = 1 - \alpha,
\]

(4.2)

where

\[
A_\ell = \Phi \left( \frac{U_\ell - \ell'\beta}{\sigma} \right) - \Phi \left( \frac{L_\ell - \ell'\beta}{\sigma} \right),
\]

(4.3)

i.e., the interval \((L_\ell, U_\ell)\) has the property that (with confidence \(1 - \alpha\)) it captures a proportion \( P \) of the normal distribution within it. However, this will not be a central proportion (i.e., symmetric about the true mean), and it is valid only for repetitive use with the same \( P \).

The approach to be taken in this section is to construct a confidence interval for the interval

\[
(\ell'\beta - K(P)\sigma, \ell'\beta + K(P)\sigma),
\]

(4.4)

where \( K(P) \) is defined by

\[
P = \Phi(K(P)) - \Phi(-K(P)).
\]

(4.5)
For each $P$ the interval (4.4) is the true interval which contains the proportion $P$ of the normal distribution symmetric about its true mean.

The confidence intervals will be of the form

$$L_{\ell}(P) = \ell'b - c*\sigma \sqrt{\ell' (X'X)^{-1} \ell} + K^2(P),$$

and

$$U_{\ell}(P) = \ell'b + c*\sigma \sqrt{\ell' (X'X)^{-1} \ell} + K^2(P),$$

and will be simultaneous in all values of $\ell$ and $P$, i.e.,

$$P[L_{\ell}(P) \leq \ell' \beta - K(P)\sigma, \ell' \beta + K(P)\sigma \leq U_{\ell}(P), \text{ for all } \ell \in \ell \text{ and all } P] = 1-\alpha.$$ (4.7)

(A related problem for the case $p = 0$ has been treated by Scheffé [5] with a different method.) The probability statement (4.7) on confidence intervals automatically implies that

$$P\{A_{\ell}(P) \geq P, \text{ for all } \ell \in \ell \text{ and all } P\} \geq 1 - \alpha,$$ (4.8)

where

$$A_{\ell}(P) = \Phi \left( \frac{U_{\ell}(P) - \ell' \beta}{\sigma} \right) - \Phi \left( \frac{L_{\ell}(P) - \ell' \beta}{\sigma} \right),$$ (4.9)

which is the corresponding tolerance interval statement.

Intervals of the form (4.6) with property (4.7) permit the investigator to make tolerance or prediction statements at various values of $x$ and not necessarily at the same level $P$ but still have one overall confidence coefficient $1-\alpha$. This may be valuable if for instance the desired percentage of corrected predictions for $x$ on the two extremes need not be as high as for $x$'s in the middle, or vice versa.
The proof of the confidence statement (4.7) depends on the following lemma.

Lemma 4.1: \[ \sum_{i=1}^{N} a_i^2 \leq c^2 \text{ if and only if } \left| \sum_{i=1}^{N} a_i b_i \right| \leq c \left( \sum_{i=1}^{N} b_i^2 \right)^{1/2} \]

for all \( b_2, \ldots, b_N \) (provided \( a_1 b_1 \neq 0 \)).

Proof: (Only if) By the Cauchy-Schwarz inequality

\[ \left| \sum a_i b_i \right| \leq \left( \sum a_i^2 \right)^{1/2} \left( \sum b_i^2 \right)^{1/2} \leq c \left( \sum b_i^2 \right)^{1/2}. \]

(If) \[ \left| \sum a_i b_i \right| \leq c \left( \sum b_i^2 \right)^{1/2}, \]

for all \( b_2, \ldots, b_N \) implies

\[ \left| \sum a_i \left( \frac{a_i b_i}{b_1} \right) \right| \leq c \left( \sum \left( \frac{a_i b_i}{b_1} \right)^2 \right)^{1/2}, \]

for all \( b_2, \ldots, b_N \) so choose \( b_i = a_i b_i / a_1, i = 2, \ldots, N \). This proves the result.

Let \( Z_1, \ldots, Z_{p+1}, \chi^2_{N-p-1} \) be generic random variables with the property that they are all independent, \( Z_1 \) has a unit normal distribution, and \( \chi^2_{N-p-1} \) has a \( \chi^2 \) distribution on \( N-p-1 \) degrees of freedom.

Theorem 4.1:

\[ \mathbb{P} \left( \left| l' \beta + K(P) \sigma - l' b \right| \leq c^* \sqrt{\ell' (X'X)^{-1} \ell} + K^2(P) \right), \text{ for all } \]

\[ l \in I \text{ and all } P = \mathbb{P} \left( \sum_{i=1}^{p+1} Z_i^2 + 1 \leq \frac{1}{N-p-1} \chi^2_{N-p-1} c^2 \right). \]
Proof: Let $Q_1$ be a non-singular matrix such that

$$Q_1(X_1'X_1)^{-1}Q_1' = I.$$ 

Let

$$b_1^{*} = Q_1 b_1, \quad \beta_1^{*} = Q_1 \beta_1, \quad \ell_1^{*} = (Q_1^{-1})' \ell_1,$$

where, for example, $\underline{b_1} = \text{col}(b_1, \ldots, b_p)$. Then $\underline{b_1}^{*}$ is distributed as $N(\beta_1^{*}, \sigma^2 I)$, and

$$\ell_1'(\beta - b) + K(P)\sigma = \ell_o'(\beta_o - b_o) + \ell_1^{*'}(\beta_1^{*} - b_1^{*}) + K(P)\sigma,$$

$$\ell_1'(X'X)^{-1} \ell = \ell_o^2/N + \underline{\ell_1}^{*'}\underline{\ell_1}^{*}.$$ 

Therefore,

$$P \left( \left| \ell_1' \beta - K(P)\sigma - \ell_1'b \right| \leq c^*s \sqrt{\ell_1'(X'X)^{-1} \ell + K^2(P)} \right.$$ 

for all $\ell \in \ell$, P

$$= P \left( \left| \frac{\ell_o}{\sqrt{N}} \cdot \sqrt{N} (\beta_o - b_o) + \ell_1^{*'}(\beta_1^{*} - b_1^{*}) + K(P)\sigma \right| \right.$$ 

$$\leq c^*s \sqrt{\ell_o^2/N + \underline{\ell_1}^{*'}\underline{\ell_1}^{*} + K^2(P)}, \text{ for all } \ell_1^{*}, \text{ P}

$$= P \left( \sum_{1}^{p+1} \sum_{1}^{2} z_{p+1}^2 + 1 \leq \frac{1}{N-p-1} \chi^2_{N-p-1} c^2 \right),$$

by Lemma 4.1.

For application to simple linear regression ($p=1$) this theorem becomes
Corollary 4.1:

\[ P \left( a + b x - c^*s \sqrt{1/N + x^2/\Sigma x_i^2 + K(P)} \leq \alpha + \beta x - K(P)\sigma \right), \]

\[ \alpha + \beta x + K(P)\sigma \leq a + bx + c^*s \sqrt{1/N + x^2/\Sigma x_i^2 + K(P)}, \text{ for all } x, P \]

\[ = P \left( z_1^2 + z_2^2 + 1 \leq \frac{1}{N-2} x_{N-2}^2 c^*^2 \right). \]

The distribution of the random variable \( (N-p-1)(\Sigma z_i^2 + 1)/x^2 \) can easily be written down in integral form. For example, in the case \( p = 1, d = c^*/\sqrt{N-2} \) is defined by

\[ 1 - \alpha = P \left\{ \frac{z_1^2 + z_2^2 + 1}{x_{N-2}^2} \leq d^2 \right\} = \int_0^\infty P \left( z_1^2 + z_2^2 \leq u d^2 - 1 \right) dP(x_{N-2}^2 \leq u) \]

\[ = \frac{1}{\Gamma\left(\frac{N-2}{2}\right) \Gamma\left(\frac{N-2}{2}\right)} \int_0^\infty u^{(N/2)-2} e^{-u/2} du \]

\[ - \frac{e^{1/2}}{\Gamma\left(\frac{N-2}{2}\right) \Gamma\left(\frac{N-2}{2}\right) (1+d^2)^{(N-2)/2}} \int_0^\infty u^{(N/2)-2} e^{-u/2} du. \]

However, evaluation of these percentile points requires numerical computation. A table of the \( 1 - \alpha \) percentile points \( c^* \) for various combinations of \( \alpha \) and \( N \) was computed on the Stanford University Burroughs 220 electronic computer and is given in Table 1 below.

When \( \beta_0 \) is absent from the regression model then there is no distinction between the case of a fixed proportion \( P \) and that of a variable proportion since the simultaneous inequalities remain invariant under a scale change on one variable. The analogous theorem to the one
Theorem 4.2:

\[
P \left[ \left| \ell_1 \beta_1 + \beta(\ell) \sigma - \ell_1' b_1 \right| \leq c * s \sqrt{\ell_1' (X'X)^{-1} \ell_1 + \lambda^2(\ell)} \right],
\]

for all \( \ell \in \mathcal{I} \), and all (or any fixed) \( \ell \)

\[
= P \left( \frac{z_1^2 + \ldots + z_p^2 + 1}{n-p} \leq \frac{1}{n-p} \cdot \frac{X_{\ell}^2}{n-p} \cdot \sigma^2 \right).
\]
<table>
<thead>
<tr>
<th>Degrees of Freedom</th>
<th>$\alpha = .5$</th>
<th>$\alpha = .3$</th>
<th>$\alpha = .1$</th>
<th>$\alpha = .05$</th>
<th>$\alpha = .01$</th>
<th>$\alpha = .001$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.399</td>
<td>4.264</td>
<td>13.16</td>
<td>26.40</td>
<td>132.1</td>
<td>1321.</td>
</tr>
<tr>
<td>2</td>
<td>1.958</td>
<td>2.792</td>
<td>5.275</td>
<td>7.604</td>
<td>17.26</td>
<td>54.75</td>
</tr>
<tr>
<td>4</td>
<td>1.748</td>
<td>2.280</td>
<td>3.471</td>
<td>4.329</td>
<td>6.851</td>
<td>12.54</td>
</tr>
<tr>
<td>5</td>
<td>1.711</td>
<td>2.189</td>
<td>3.204</td>
<td>3.894</td>
<td>5.781</td>
<td>9.569</td>
</tr>
<tr>
<td>6</td>
<td>1.686</td>
<td>2.130</td>
<td>3.059</td>
<td>3.635</td>
<td>5.183</td>
<td>8.951</td>
</tr>
<tr>
<td>7</td>
<td>1.668</td>
<td>2.088</td>
<td>2.928</td>
<td>3.463</td>
<td>4.805</td>
<td>7.148</td>
</tr>
<tr>
<td>8</td>
<td>1.654</td>
<td>2.056</td>
<td>2.848</td>
<td>3.341</td>
<td>4.546</td>
<td>6.555</td>
</tr>
<tr>
<td>9</td>
<td>1.643</td>
<td>2.032</td>
<td>2.788</td>
<td>3.251</td>
<td>4.357</td>
<td>6.139</td>
</tr>
<tr>
<td>10</td>
<td>1.634</td>
<td>2.013</td>
<td>2.741</td>
<td>3.181</td>
<td>4.214</td>
<td>5.831</td>
</tr>
<tr>
<td>11</td>
<td>1.627</td>
<td>1.997</td>
<td>2.704</td>
<td>3.126</td>
<td>4.102</td>
<td>5.594</td>
</tr>
<tr>
<td>12</td>
<td>1.620</td>
<td>1.984</td>
<td>2.673</td>
<td>3.080</td>
<td>4.012</td>
<td>5.408</td>
</tr>
<tr>
<td>13</td>
<td>1.615</td>
<td>1.973</td>
<td>2.647</td>
<td>3.043</td>
<td>3.939</td>
<td>5.257</td>
</tr>
<tr>
<td>14</td>
<td>1.610</td>
<td>1.964</td>
<td>2.626</td>
<td>3.011</td>
<td>3.877</td>
<td>5.132</td>
</tr>
<tr>
<td>15</td>
<td>1.606</td>
<td>1.956</td>
<td>2.607</td>
<td>2.984</td>
<td>3.825</td>
<td>5.027</td>
</tr>
<tr>
<td>16</td>
<td>1.602</td>
<td>1.949</td>
<td>2.591</td>
<td>2.961</td>
<td>3.780</td>
<td>4.938</td>
</tr>
<tr>
<td>17</td>
<td>1.599</td>
<td>1.942</td>
<td>2.577</td>
<td>2.941</td>
<td>3.741</td>
<td>4.862</td>
</tr>
<tr>
<td>18</td>
<td>1.596</td>
<td>1.937</td>
<td>2.564</td>
<td>2.923</td>
<td>3.707</td>
<td>4.795</td>
</tr>
<tr>
<td>19</td>
<td>1.594</td>
<td>1.932</td>
<td>2.553</td>
<td>2.907</td>
<td>3.677</td>
<td>4.737</td>
</tr>
<tr>
<td>20</td>
<td>1.591</td>
<td>1.928</td>
<td>2.544</td>
<td>2.893</td>
<td>3.650</td>
<td>4.685</td>
</tr>
<tr>
<td>21</td>
<td>1.589</td>
<td>1.924</td>
<td>2.535</td>
<td>2.880</td>
<td>3.626</td>
<td>4.630</td>
</tr>
<tr>
<td>22</td>
<td>1.587</td>
<td>1.920</td>
<td>2.527</td>
<td>2.869</td>
<td>3.605</td>
<td>4.598</td>
</tr>
<tr>
<td>23</td>
<td>1.585</td>
<td>1.917</td>
<td>2.519</td>
<td>2.858</td>
<td>3.585</td>
<td>4.561</td>
</tr>
<tr>
<td>24</td>
<td>1.584</td>
<td>1.913</td>
<td>2.513</td>
<td>2.849</td>
<td>3.567</td>
<td>4.527</td>
</tr>
<tr>
<td>25</td>
<td>1.582</td>
<td>1.911</td>
<td>2.507</td>
<td>2.840</td>
<td>3.551</td>
<td>4.496</td>
</tr>
<tr>
<td>26</td>
<td>1.581</td>
<td>1.908</td>
<td>2.501</td>
<td>2.832</td>
<td>3.536</td>
<td>4.468</td>
</tr>
<tr>
<td>27</td>
<td>1.579</td>
<td>1.906</td>
<td>2.496</td>
<td>2.825</td>
<td>3.523</td>
<td>4.443</td>
</tr>
<tr>
<td>28</td>
<td>1.578</td>
<td>1.904</td>
<td>2.491</td>
<td>2.818</td>
<td>3.510</td>
<td>4.419</td>
</tr>
<tr>
<td>29</td>
<td>1.577</td>
<td>1.902</td>
<td>2.487</td>
<td>2.812</td>
<td>3.498</td>
<td>4.397</td>
</tr>
<tr>
<td>30</td>
<td>1.576</td>
<td>1.900</td>
<td>2.482</td>
<td>2.806</td>
<td>3.488</td>
<td>4.377</td>
</tr>
<tr>
<td>35</td>
<td>1.571</td>
<td>1.892</td>
<td>2.466</td>
<td>2.782</td>
<td>3.443</td>
<td>4.295</td>
</tr>
<tr>
<td>40</td>
<td>1.568</td>
<td>1.886</td>
<td>2.453</td>
<td>2.764</td>
<td>3.411</td>
<td>4.235</td>
</tr>
<tr>
<td>45</td>
<td>1.566</td>
<td>1.882</td>
<td>2.443</td>
<td>2.750</td>
<td>3.386</td>
<td>4.189</td>
</tr>
<tr>
<td>50</td>
<td>1.563</td>
<td>1.878</td>
<td>2.436</td>
<td>2.739</td>
<td>3.366</td>
<td>4.153</td>
</tr>
<tr>
<td>60</td>
<td>1.560</td>
<td>1.873</td>
<td>2.424</td>
<td>2.723</td>
<td>3.336</td>
<td>4.099</td>
</tr>
<tr>
<td>70</td>
<td>1.558</td>
<td>1.869</td>
<td>2.416</td>
<td>2.712</td>
<td>3.316</td>
<td>4.062</td>
</tr>
<tr>
<td>80</td>
<td>1.556</td>
<td>1.866</td>
<td>2.410</td>
<td>2.703</td>
<td>3.300</td>
<td>4.034</td>
</tr>
<tr>
<td>90</td>
<td>1.555</td>
<td>1.864</td>
<td>2.405</td>
<td>2.696</td>
<td>3.288</td>
<td>4.013</td>
</tr>
<tr>
<td>100</td>
<td>1.554</td>
<td>1.862</td>
<td>2.401</td>
<td>2.691</td>
<td>3.279</td>
<td>3.996</td>
</tr>
</tbody>
</table>

Table 1

Values of $c^*$

for simple linear regression ($p=1$).
5. Tolerance intervals for a fixed central proportion $P$ which are simultaneous in all values of $\ell$.

In the previous section confidence intervals were constructed for $\ell'\beta + K(P)\sigma$ which were simultaneous in $\ell$ and $P$. The simultaneity in $P$ allows the experimenter to vary the accuracy with which he wishes to predict future responses as $\ell$ varies. In this section the analogous confidence intervals for $\ell'\beta + K(P)\sigma$ which are simultaneous in $\ell$ for a fixed value of $P$ will be obtained. These are applicable when the experimenter is always interested in bracketing the same central proportion of the distribution or in predicting all future responses with the same accuracy.

For a fixed $P$ confidence intervals of the form

$$L_\ell = \ell'b - c**s \sqrt{\ell'(X'X)^{-1}\ell},$$

(5.1)

$$U_\ell = \ell'b + c**s \sqrt{\ell'(X'X)^{-1}\ell},$$

will be derived for $(\ell'\beta - K(P)\sigma, \ell'\beta + K(P)\sigma)$ which are simultaneous for $\ell \in \mathcal{L}$. Whereas previously $c^*$ depended on just $p$ and $N-p-1$, now $c**$ will also depend on $K(P)$. The derivation is contained in the following theorem.

Theorem 5.1:

$$P \left( |\ell'\beta + K(P)\sigma - \ell'b| \leq c**s \sqrt{\ell'(X'X)^{-1}\ell} \text{ for all } \ell \in \mathcal{L} \right)$$

$$= P \left( (Z_1 + \sqrt{N-K(P)})^2 + Z_2^2 + \ldots + Z_{p+1}^2 \leq \frac{1}{N-p-1} \chi^2_{N-p-1} c^{**2} \right).$$
Proof: Choose

\[ Q = \begin{pmatrix} \sqrt{N} & 0 \\ 0 & Q_1 \end{pmatrix}, \]

where \( Q_1 \) is non-singular, so that \( Q(X'X)^{-1} Q' = I \).

Let

\[ \beta^* = Q\beta, \ b^* = Qb, \ \text{and} \ \ell^* = (Q^{-1})' \ell. \]

Then

\[ \ell^* \beta + K(P)\sigma - \ell^* b = \ell^{**} \beta^* + K(P)\sigma - \ell^{**} b^* \]

\[ = \ell_0^{**} \left( \beta_0^{**} - b_0^{**} + (K(P)/\ell_0^{**})\sigma \right) + \ell_{1}^{**} \left( \ell_1^{**} - \ell_{1}^{**} \right), \]

\[ \ell^* (X'X)^{-1} \ell = \ell^{**} \ell^*. \]

where \( \ell_0^{**} = 1/\sqrt{N} \) and \( b^* \) is distributed as \( N(\beta^*, \sigma^2 I) \).

By Lemma 4.1

\[ P \left( \left| \ell^{**} \beta^* + K(P)\sigma - \ell^{**} b^* \right| \leq c^{**} \sqrt{\ell^{**} \ell^*}, \text{ for all} \right. \]

\[ \ell_{1}^{**}, \ldots, \ell_{p}^{**} \]

\[ = P((Z_1 + \sqrt{N} K(P))^2 + Z_2^2 + \cdots + Z_{p+1}^2 \leq \frac{1}{N-p-1} x^2 \chi^2_{N-p-1} c^{**^2}). \]

For application to simple linear regression this theorem becomes

**Corollary 5.1:**

\[ P \left( |\alpha + \beta x + K(P)\sigma - a - b x| \leq c^{**s} \sqrt{\frac{1}{N} + \frac{x^2}{\Sigma x^2}}, \text{ for all x} \right) \]

\[ = P \left((Z_1 + \sqrt{N} K(P))^2 + Z_2^2 \leq \frac{1}{N-2} x^2 \chi^2_{N-2} c^{**^2} \right). \]

Thus, all that is required to determine the intervals (5.1) is to evaluate the percentile points of the distribution of
\[
(5.2) \quad \frac{(z_1 \pm \sqrt{\frac{N}{N} k(P)})^2 + z_2^2 + \ldots + z_{p+1}^2}{\chi_{N-p-1}^2 / N-p-1}.
\]

The distribution of (5.2) is not exactly a non-central F because of \( \pm \sqrt{\frac{N}{N} k(P)} \), i.e., both plus and minus signs. However, the percentile can be evaluated in special cases by numerical computation. When \( p = 1 \) (i.e., \( E(y) = \alpha + \beta x \)) \( c^{**} \) is defined by

\[
1 - \alpha = P \left\{ (z_1 + \sqrt{\frac{N}{N} k(P)})^2 + z_2^2 \leq \frac{1}{N-2} \chi_{N-2}^2 c^{**2} \right\}
\]

\[
\int_{N}^{\infty} \left[ \Phi \left( \sqrt{\frac{N}{N} k(P)} \right) - \Phi \left( \sqrt{\frac{N}{N} k(P)} \right) \right] \frac{dP}{N} \left\{ \frac{\chi_{N-2}^2 c^{**2}}{N-2} - \frac{Z_2^2}{N-2} \leq t \right\}
\]

and in the special case of \( N \) even \( (N = 2k) \) for \( t > 0 \)

\[
dP \left\{ \chi_{N-2}^2 D^2 - Z_2^2 \leq t \right\}
\]

\[
= \frac{e^{-t/2D^2}}{\Gamma(k/2) \Gamma(k-1/2)} \int_0^\infty u^{-1/2} (u+t/D^2)^{k-2} e^{-(1+D^2)u/2} du dt.
\]

\[
(5.4)
\]

Various upper percentile points have been computed on the Stanford University Burroughs 220 electronic computer for a range of values of \( N \) and are presented in Table 2.
<table>
<thead>
<tr>
<th>N</th>
<th>$P = .50$</th>
<th>$P = .75$</th>
<th>$P = .90$</th>
<th>$P = .95$</th>
<th>$P = .99$</th>
<th>$P = .999$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>5.380</td>
<td>7.453</td>
<td>9.698</td>
<td>11.150</td>
<td>14.014</td>
<td>17.363</td>
</tr>
<tr>
<td>12</td>
<td>5.023</td>
<td>7.221</td>
<td>9.586</td>
<td>11.112</td>
<td>14.121</td>
<td>17.634</td>
</tr>
<tr>
<td>16</td>
<td>5.197</td>
<td>7.586</td>
<td>10.150</td>
<td>11.803</td>
<td>15.057</td>
<td>18.856</td>
</tr>
<tr>
<td>18</td>
<td>5.300</td>
<td>7.786</td>
<td>10.449</td>
<td>12.165</td>
<td>15.542</td>
<td>19.484</td>
</tr>
<tr>
<td>20</td>
<td>5.408</td>
<td>7.987</td>
<td>10.747</td>
<td>12.526</td>
<td>16.023</td>
<td>20.104</td>
</tr>
</tbody>
</table>

$\alpha = .10$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>5.933</td>
<td>8.281</td>
<td>10.831</td>
<td>12.484</td>
<td>15.750</td>
<td>19.568</td>
</tr>
<tr>
<td>10</td>
<td>5.728</td>
<td>8.080</td>
<td>10.632</td>
<td>12.286</td>
<td>15.553</td>
<td>19.369</td>
</tr>
<tr>
<td>12</td>
<td>5.684</td>
<td>8.093</td>
<td>10.701</td>
<td>12.391</td>
<td>15.724</td>
<td>19.619</td>
</tr>
<tr>
<td>14</td>
<td>5.711</td>
<td>8.194</td>
<td>10.880</td>
<td>12.617</td>
<td>16.045</td>
<td>20.050</td>
</tr>
</tbody>
</table>

$\alpha = .05$

<table>
<thead>
<tr>
<th>N</th>
<th>24.466</th>
<th>33.019</th>
<th>42.398</th>
<th>48.620</th>
<th>60.500</th>
<th>74.642</th>
</tr>
</thead>
</table>

$\alpha = .01$

<table>
<thead>
<tr>
<th>N</th>
<th>42.398</th>
<th>48.620</th>
<th>60.500</th>
<th>74.642</th>
</tr>
</thead>
</table>

Table 2

Values of $c^*$
in simple linear regression ($p = 1$).

23
6. **Central tolerance intervals simultaneous in** \( x \) **and** \( P \) **from a Bonferroni inequality.**

The Bonferroni inequality \( P(AB) \geq 1 - P(A^c) - P(B^c) \) where \( A^c(B^c) \) denotes the complement of \( A(B) \) has frequently proved to be useful in applied statistical theory. In this context it can be employed to combine confidence intervals on the regression mean and standard deviation in order to construct a central tolerance interval.

A confidence region for the entire regression line was first derived by Working and Hotelling [6], and the same result can also be obtained from the Scheffé technique. The appropriate probability statement is the following:

\[
(6.1) \quad P(|\varepsilon'(b-\beta)| \leq s \left( \frac{F_{p+1,N-p-1}^{1-\alpha/2}}{p+1,N-p-1} \right)^{1/2} \sqrt{\varepsilon'(X'X)^{-1}\varepsilon'},
\]

for all \( \ell \in \mathcal{R} \) = 1 - \( \alpha/2 \),

where \( F_{p+1,N-p-1}^{1-\alpha/2} \) is the (upper) \( 1-\alpha/2 \) percentile point of the \( F \) distribution on \( p+1 \) and \( N-p-1 \) degrees of freedom. In the special case \( p = 1 \) the expression (6.1) reduces to

\[
(6.2) \quad P(|\alpha + \beta x - a - bx| \leq s \left( \frac{2F_{2,N-2}^{1-\alpha/2}}{2F_{2,N-2}^{1-\alpha/2}} \right)^{1/2} \sqrt{1/N + x^2/\sum x^2},
\]

for all \( x \) = 1 - \( \alpha/2 \).

An upper bound on \( \sigma \) can easily be obtained from a one-sided chi-square confidence interval.

\[
(6.3) \quad P\left( \sigma \leq \left( \frac{N-p-1}{\chi^2_{N-p-1}} \right)^{1/2} \right) = 1 - \alpha/2,
\]
where $\chi^2_{N-p-1}$ is the (lower) $\alpha/2$ percentile point of the $\chi^2$ distribution on $N-p-1$ degrees of freedom.

The confidence statements (6.1) and (6.3) can now be combined into a joint confidence statement with probability greater than or equal to $1-\alpha$ by the aforementioned Bonferroni inequality. This in turn implies the following probability statement:

$$
(6.4) \quad P[|\ell' \beta + K(P)\sigma - \ell'b| \leq s \left[ \left( \frac{(p+1)F_{1-\alpha/2, p+1, N-p-1}}{\sqrt{\ell'(X'X)\ell}}\right)^{1/2} + K(P)\left(\frac{N-p-1}{\chi^2_{N-p-1}}\right)^{1/2}\right], \text{ for all } x, P \geq 1 - \alpha.
$$

In particular, for simple linear regression

$$
(6.5) \quad P[|\alpha + \beta x + K(P)\sigma - a - bx| \leq s \left[ \left( \frac{1-\alpha/2}{2F_{1-\alpha/2, N-2}}\right)^{1/2} \sqrt{1/N + x^2/\Sigma x^2_1} + K(P)\left(\frac{N-2}{\chi^2_{N-2}}\right)^{1/2}\right], \text{ for all } x, P \geq 1 - \alpha.
$$

The probability statements (6.4) and (6.5) are not exact (i.e., equal to $1 - \alpha$) as for the first three methods. Nevertheless, in many instances methods derived from a Bonferroni inequality have proved just as efficient as theoretically more exact methods.
7. Numerical example

Fifteen hypothetical pairs of values \((x_i, y_i)\) were selected for the speed-orifice problem and are given below:

<table>
<thead>
<tr>
<th>Orifice opening (inches)</th>
<th>Speed (miles/hr)</th>
<th>Orifice opening (inches)</th>
<th>Speed (miles/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.310</td>
<td>4360</td>
<td>1.360</td>
<td>5550</td>
</tr>
<tr>
<td>1.313</td>
<td>4590</td>
<td>1.364</td>
<td>5390</td>
</tr>
<tr>
<td>1.320</td>
<td>4520</td>
<td>1.373</td>
<td>5670</td>
</tr>
<tr>
<td>1.322</td>
<td>4770</td>
<td>1.376</td>
<td>5490</td>
</tr>
<tr>
<td>1.338</td>
<td>4760</td>
<td>1.384</td>
<td>5810</td>
</tr>
<tr>
<td>1.340</td>
<td>5070</td>
<td>1.395</td>
<td>6060</td>
</tr>
<tr>
<td>1.347</td>
<td>5230</td>
<td>1.400</td>
<td>5940</td>
</tr>
<tr>
<td>1.355</td>
<td>5080</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\bar{x} = 1.3531, \quad \bar{y} = 5219.3, \quad \sum (x_i - \bar{x})^2 = 0.011966,
\]

\[
a = -19041.9, \quad b = 17930, \quad s = 130.5.
\]

At the three points \(x_1, \bar{x},\) and \(x_{15}\) the total constant which multiplies \(s\) in the tolerance interval was computed for the original Wallis technique \((W)\), the simultaneous Wallis technique \((SW)\), the simultaneous \(x\) and \(P\) central procedure \((CSxP)\), the simultaneous \(x\) (but fixed \(P\) central procedure \((CSx)\), and finally the simultaneous technique from the Bonferroni inequality \((CBSxP)\). Three different combinations of \(\alpha\) and \(P\) were chosen for the comparison.
Table 3

Constants multiplying s in the example.

<table>
<thead>
<tr>
<th>P</th>
<th>α</th>
<th>x</th>
<th>W</th>
<th>SW</th>
<th>CSxP</th>
<th>CSx</th>
<th>CBSxP</th>
</tr>
</thead>
<tbody>
<tr>
<td>.95</td>
<td>.01</td>
<td>1.310</td>
<td>3.84</td>
<td>6.58</td>
<td>7.94</td>
<td>7.23</td>
<td>5.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.353</td>
<td>3.60</td>
<td>3.61</td>
<td>7.79</td>
<td>3.97</td>
<td>4.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.400</td>
<td>3.88</td>
<td>6.99</td>
<td>7.97</td>
<td>7.70</td>
<td>5.77</td>
</tr>
<tr>
<td>.75</td>
<td>.05</td>
<td>1.310</td>
<td>1.90</td>
<td>3.23</td>
<td>3.78</td>
<td>3.89</td>
<td>3.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.353</td>
<td>1.77</td>
<td>1.77</td>
<td>3.59</td>
<td>2.13</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.400</td>
<td>1.92</td>
<td>3.43</td>
<td>3.82</td>
<td>4.14</td>
<td>3.43</td>
</tr>
<tr>
<td>.50</td>
<td>.10</td>
<td>1.310</td>
<td>1.02</td>
<td>1.73</td>
<td>2.18</td>
<td>2.42</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.353</td>
<td>.95</td>
<td>.95</td>
<td>1.91</td>
<td>1.33</td>
<td>1.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.400</td>
<td>1.04</td>
<td>1.84</td>
<td>2.22</td>
<td>2.56</td>
<td>2.39</td>
</tr>
</tbody>
</table>

Legend:  
W - Wallis  
SW - Simultaneous Wallis  
CSxP - Central intervals simultaneous in x and P.  
CSx - Central intervals simultaneous in x.  
CBSxP - Central Bonferroni intervals simultaneous in x and P.
Some tentative conclusions can be drawn from this small numerical example, but their general validity can be substantiated only by further experience with these techniques.

The non-simultaneous Wallis procedure yields the shortest intervals, and except for two instances in the stringent case $P = .95, \alpha = .01$ the simultaneous Wallis procedure gives the second shortest intervals. The non-simultaneous Wallis intervals must, of course, be shorter than the corresponding simultaneous Wallis ones, and the fact that these latter intervals should in most cases be shorter than the other simultaneous intervals agrees with one's intuition. The Wallis technique only aims at bracketing a proportion $P$ of the underlying distribution whereas the other simultaneous procedures are choosier and require the proportion to be symmetrically located about the true mean. That the simultaneous Wallis intervals are not universally shorter than the others is somewhat surprising.

The most illuminating fact to be found in the table is the strength of the Bonferroni intervals. Although they are crudely obtained by adding together the endpoints of separate confidence intervals, in more than one case they are shorter than the other simultaneous central intervals and in the two instances mentioned above are even better than the simultaneous Wallis intervals. The Bonferroni intervals will be particularly strong competitors for small $\alpha$ since then the inequalities in (6.4) and (6.5) will practically be equalities.

Comparison of the CSxP and CSx intervals shows that their lengths will criss-cross as a function of $x$. For $x$ values close to $\bar{x}$ the CSx intervals will be shorter than the CSxP intervals, but as
The reverse occurs. This can be deduced from a comparison of the form of the coefficients of $a$ in the two expressions. Thus the selection of either of these techniques depends where most of the $x$ values will arise.

An overall comparison of all the simultaneous procedures with the non-simultaneous one indicates that the lengths of the simultaneous intervals vary from slightly larger to about twice as large as the non-simultaneous intervals. This gives a rough indication of the price the experimenter will have to pay, or should be paying, for simultaneity.
8. Detection of outliers

The detection and correction of outlying or maverick observations in regression analysis or, more generally, in any statistical analysis has received increased attention recently (see *Technometrics*, May 1960). Perhaps the most appealing criterion for the detection of a discrepant observation in a regression analysis would be the magnitude of the maximum absolute Studentized residual (cf., [7], [8]), however, it is not feasible to tabulate the percentile points of this distribution because of its dependence on the particular values of the independent variable $x$ in each problem. This leaves the statistician without any objective numerical criterion for outlier detection. It is still of course possible to correct for outliers without detecting them provided one adopts the principle of always correcting a fixed percentage of the maximum absolute residuals [9].

Anscombe [10] has indicated that a form of approximate analysis is available since bounds can be placed on the variance of each residual. Let $r_i = y_i - a - bx_i, \ i = 1, \ldots, N$. Then

$$0 \leq \text{Var}(r_i) \leq (N - 1)s^2/N,$$

and in the case of equally-spaced $x$'s,

$$\frac{N - 4}s^2/N < \text{Var}(r_1) \leq (N - 1)s^2/N.$$  

However, the complex correlation structure makes it difficult to evaluate even a rough detection procedure which depends upon classifying as mavericks those observations whose absolute residuals are too large.
A proposal for alleviating this situation would be to use the confidence intervals on \((\ell' \beta - K(P)\sigma, \ell' \beta + K(P)\sigma)\) which are simultaneous in \(x\) and \(P\) to evaluate how far out on the tail of its distribution each observation lies. That is, if \(y_1\) is an observation at \(x_1\), then there exists a \(P(y_1)\) which satisfies the equation

\[
|y_1 - a - bx_1| = s \left[ \left( \frac{2F_{\alpha/2, N-2}}{2} \right)^{1/2} \sqrt{\frac{1}{N} + \frac{x_1^2}{\sum x_i^2}} \right. \\
\left. + K(P(y_1)) \left( \frac{N-2}{N-2} \right)^{\alpha/2} \right]^{1/2}.
\]

The area under the tails of the normal distribution for values larger than this residual is less than or equal to \(1 - P(y_1)\) with confidence greater than or equal to \(1 - \alpha\) for all such statements. For the statistician who does not believe in events of small probability it would be reasonable to regard as suspect all values \(y_1\) with very small \(1 - P(y_1)\). The precise formulation of the detection criterion could be left somewhat vague with the statistician being suspicious of a single value with tremendously small \(1 - P(y_1)\) or of a group of values if too many of small probability occur.

Some numerical investigation quickly reveals that as it stands (8.3) is not a very sensitive detection device because a discrepant observation sufficiently inflates the internally estimated standard deviation \(s\) to obviate its detection. This difficulty can be circumvented by judiciously dividing the observations into \(n\) groups. Then the observations in each group can be tested for discrepancy by comparing them with the regression line and standard deviation estimated
from pooling the observations in the other \( n - 1 \) groups. If for each group the comparison is performed with confidence \( 1 - \frac{\alpha}{n} \), then the overall confidence for all the points will be greater than or equal to \( 1 - \alpha \) by the Bonferroni inequality.

For the numerical example suppose the observed value at 1.34 had been 5800 instead of 5070. The data was divided into five groups consisting of consecutive triplets; the level of significance for each triplet was .01 so that the overall confidence was greater than or equal to .95. When the regression line and standard deviation were estimated from the first and last three triplets, the value at 1.34 had a residual of 819.7 and corresponding \( K(P) = 2.27 \). This would mean that for this observation \( 1 - P = .023 \) which would probably make the statistician uneasy and possibly want to correct this value.

The technique of section 4 could similarly be adopted to outlier detection, but in this instance it does not give as sensitive a result.

Before this detection device is widely applied some further work should be done to evaluate its performance. There also remains the unanswered question of what to do with a maverick once it has been spotted. Ignore it, reject it, Winsorize it — these are all possibilities. Discussion of this has appeared elsewhere (see Technometrics, May 1960, and [11]), and the proper choice very likely depends on the problem.
9. **Acknowledgments**

For the computations and programs involved in this paper the authors would like to acknowledge and thank Shirley Eberly, Gay Fischer, Phyllis Groll, and Margaret Korpi.
References


STANFORD UNIVERSITY

TECHNICAL REPORT DISTRIBUTION LIST

CONTRACT No. 225(53)

(IN 042-002)

Chief, Bureau of Ordnance
Quality Control Division QCO
Department of the Navy
Washington 25, D. C.
Attn: Dr. W. R. Fabsi, Jr.

1

Chief, Arizona Air Procurement Dist.
Attn: Quality Control Div.
2675 Sky Harbor Blvd.
Phoenix, Arizona

1

Chief, San Diego Air Procurement Dist.
San Bernardino Air Material Area
Attn: Quality Control Division
P. O. Box 1549
Old San Diego Station
4325 Pacific Highway
San Diego 10, California

1

Chief, Los Angeles Air Procurement Dist.
San Bernardino Air Material Area
Attn: Quality Control Division
Baldwin Bos.
1206 S. Maple Street
Los Angeles, California

1

Chief, San Francisco Air Procurement Dist.
Sacramento Air Material Area
Attn: Quality Control Division
1513 Oak Street
Oakland 12, California

1

Chief, Atlanta Air Procurement Dist.
Warner Robins Air Material Area
Attn: Quality Control Division
441 West Peachtree Street N.E.
Atlanta, Georgia

1

Chief, Boston Air Procurement Dist.
Middletown Air Material Area
Attn: Quality Control Division
Boston Air Terminal
Boston 10, Massachusetts

1

Chief, St. Louis Air Procurement Dist.
Oklahoma City Air Material Area
Attn: Quality Control Division
1114 Market Street
St. Louis 1, Missouri

1

Chief Indianapolis Air Procurement Dist.
Mobile Air Material Area
Attn: Quality Control Division
54 Monument Circle
Indianapolis 6, Indiana

1

Chief, Rochester Air Procurement Dist.
Middletown Air Material Area
Attn: Quality Control Division
20 Symington Place, P.O. Box 1667
Rochester 3, New York

1

Chief, Cleveland Air Procurement Dist.
Mobile Air Material Area
Attn: Quality Control Division
1279 W. Third Street
Cleveland 15, Ohio

1

Contract No. 225(53)
May 1961
<table>
<thead>
<tr>
<th>Commander</th>
<th>Location</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commanding Officer &amp; Director</td>
<td>U.S. Navy Underwater Sound Lab., Fort Trumbull, New London, Connecticut</td>
<td>1</td>
</tr>
<tr>
<td>Commander Mobile Air Material Area</td>
<td>Mobile Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Assistant for Quality Directorate of Maintenance Engineering</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>broccoli Air Force Base, Alabama</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Mobile Air Material Area</td>
<td>Mobile Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Material Quality Division Directorate of Supply &amp; Services</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>broccoli Air Force Base, Alabama</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Mobile Air Material Area</td>
<td>Mobile Air Material Area</td>
<td>2</td>
</tr>
<tr>
<td>Attn: Quality Control Officer, MQ</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>broccoli Air Force Base, Alabama</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Commander Air Material Force, Pacific Area</td>
<td>Air Material Force, Pacific Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Office</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>FEAMCOM Air Base, APO 323</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>San Francisco, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Northern Air Material Area, Pacific Area</td>
<td>Northern Air Material Area, Pacific Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Office</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>FEAMCOM Air Base, APO 323</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>San Francisco, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Sacramento Air Material Area</td>
<td>Sacramento Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Assistant of Quality Directorate of Maintenance Engineering</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Mcclellan AFB, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Sacramento Air Material Area</td>
<td>Sacramento Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Division Directorate of Procurement and Production</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Mcclellan AFB, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Sacramento Air Material Area</td>
<td>Sacramento Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Material Quality Division Directorate of Supply and Services</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Mcclellan AFB, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Sacramento Air Material Area</td>
<td>Sacramento Air Material Area</td>
<td>2</td>
</tr>
<tr>
<td>Attn: Quality Control Office, SQ</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Mcclellan AFB, California</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Commander San Bernardino Air Material Area</td>
<td>San Bernardino Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Assistant for Quality Directorate of Maintenance Engineering</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Norton AFB, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander San Bernardino Air Material Area</td>
<td>San Bernardino Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Division Directorate of Procurement and Production</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Norton AFB, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander San Bernardino Air Material Area</td>
<td>San Bernardino Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Material Quality Division Directorate of Supply and Services</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Norton AFB, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Warner Robins Air Material Area</td>
<td>Warner Robins Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Assistant for Quality Directorate of Maintenance Engineering</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Robins AFB, Georgia</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Warner Robins Air Material Area</td>
<td>Warner Robins Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Office, WRQ</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Robins AFB, Georgia</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Warner Robins Air Material Area</td>
<td>Warner Robins Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Division Directorate of Procurement and Production</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Robins AFB, Georgia</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Oklahoma City Air Material Area</td>
<td>Oklahoma City Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Assistant for Quality Directorate of Maintenance Engineering</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Tinker Air Force Base, Oklahoma City, Oklahoma</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Oklahoma City Air Material Area</td>
<td>Oklahoma City Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Division Directorate of Supply and Services</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Tinker Air Force Base, Oklahoma City, Oklahoma</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Oklahoma City Air Material Area</td>
<td>Oklahoma City Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Division Directorate of Procurement and Production</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Tinker Air Force Base, Oklahoma City, Oklahoma</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Oklahoma City Air Material Area</td>
<td>Oklahoma City Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Quality Control Division, OC</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Tinker AFB, Oklahoma</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Warner Robins Air Material Area</td>
<td>Warner Robins Air Material Area</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Assistant for Quality Directorate of Maintenance Engineering</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Robins AFB, Georgia</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander Pope Air Force Depot</td>
<td>Pope Air Force Depot</td>
<td>1</td>
</tr>
<tr>
<td>Attn: Assistant for Quality Directorate of Maintenance Engineering</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Griffiss AFB, New York</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Contract No. 225(53) May 1961
<table>
<thead>
<tr>
<th>Name</th>
<th>Address</th>
<th>City, State</th>
<th>Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr. A. Chames</td>
<td>The Technological Institute</td>
<td>Northwestern University</td>
<td>Evanston, Illinois</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mr. W. H. Catworthy</td>
<td>Betts Plan, Westinghouse Electric Corporation</td>
<td>Box 1468</td>
<td>Pittsburgh 30, Pennsylvania</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor Paul Clifford</td>
<td>New Jersey State Teachers College</td>
<td>Montclair, New Jersey</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor W. G. Coehran</td>
<td>Department of Statistics</td>
<td>Harvard University</td>
<td>Cambridge, Massachusetts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor C. C. Gockamer</td>
<td>Institute of Statistics</td>
<td>State College Section</td>
<td>North Carolina State College Raleigh, North Carolina</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor Edward P. Coleman</td>
<td>Engineering Department</td>
<td>University of California</td>
<td>Los Angeles 24, California</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dr. J. Louis Court</td>
<td>Division 17</td>
<td>National Bureau of Standards</td>
<td>Washington 25, D. C.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miss Gertrude M. Cox</td>
<td>Institute of Statistics</td>
<td>North Carolina State College</td>
<td>Raleigh, North Carolina</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dr. Joseph Daly</td>
<td>U. S. Census Bureau</td>
<td>Washington 25, D. C.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor Cyrus Derman</td>
<td>Department of Industrial Engineering</td>
<td>Columbia University</td>
<td>New York 27, New York</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mr. H. F. Dodge</td>
<td>Rutgers University</td>
<td>New Brunswick, New Jersey</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dr. Francis Breisch</td>
<td>Stanford Research Institute</td>
<td>333 Ravenswood Avenue Menlo Park, California</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor Achason J. Duncan</td>
<td>Department of Industrial Engineering</td>
<td>Johns Hopkins University</td>
<td>Baltimore 18, Maryland</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor Mayer Dwass</td>
<td>Department of Mathematics</td>
<td>Northwestern University</td>
<td>Evanston, Illinois</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor D. A. S. Fraser</td>
<td>Department of Mathematics</td>
<td>University of Toronto Toronto 5, Canada</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Captain A. E. Chapman</td>
<td>AFPR</td>
<td>Boeing Aircraft Corporation</td>
<td>Seattle, Washington</td>
</tr>
<tr>
<td>Contract No. 225153</td>
<td></td>
<td></td>
<td>May 1961</td>
</tr>
</tbody>
</table>
Professor Norman Rubin
Department of Statistics
Michigan State University
East Lansing, Michigan

Professor Herman Rubin
Department of Statistics
Sacremento State College
Sacremento, California

Miss Marion M. Sandoine
U.S. Department of Agriculture
Western Regional Laboratory
Biometrical Services
Albany, New York

Professor J. Richard Savage
School of Business Administration
University of Minnesota
Minneapolis, Minnesota

Professor L. J. Savage
Mathematics Department
University of Michigan
Ann Arbor, Michigan

Professor Henry Scheffe
Department of Statistics
University of California
Berkeley 4, California

Professor Robert Schlaifer
Graduate School of Business Administration
Harvard University
Boston 53, Massachusetts

Professor Seymour Sherman
Department of Mathematics
Wayne State University
Detroit 2, Michigan

Mr. Walter Shevaniart
Bell Telephone Laboratories, Inc.
Murray Hill, New Jersey

Dr. Rosedith Sitgreaves
Teachers College
Columbia University
New York 27, New York

Dr. Milton Sobel
Statistics Department
University of Minnesota
Minneapolis, Minnesota

Professor Frank Splinter
Department of Mathematics
University of Minnesota
Minneapolis, Minnesota

Mr. Arthur Stein
Cornell Aeromedical Lab., Inc.
P.O. Box 235
Buffalo 21, New York

Professor W. Allen Wallis
School of Business
University of Chicago
Chicago 37, Illinois

Dr. Irving Weiss
The Milne Corporation
Bedford, Massachusetts

Captain Barton L. Weller
AFPR Office
Martin Aircraft Corporation
Denver, Colorado

Professor Mason Wescott
Editor, Industrial Quality Control
Rutgers University
New Brunswick, New Jersey

Professor S. S. Wilks
Department of Mathematics
Princeton University
Princeton, New Jersey

Professor J. Wolfowitz
Department of Mathematics
Cornell University
Ithaca, New York

Dr. Max A. Woodbury
Department of Mathematics
University of New York
New York 53, New York

Distribution via ONR London

Commanding Officer
Branch Office
Navy Rs. 100
Fleet Post Office
New York, New York

Dr. William R. Buckland
22 Ryder Street
London S. W. 1
England

Professor Georges Darpeis
Director, Institute de Statistique
University of Paris
11 Rue Pierre Curie
Paris 5, France

Professor R. Foret
Institut Henri Poincare
Paris, France

Dr. Geoffrey Gregory
4, Osborne Grove
Galley, Chealsea
Chelsea, England

A. Halg
Fennoski, 83
Viron, Denmark

Dr. H. C. Hamaker
Philips Research Laboratories
Eindhoven, Netherlands

Mr. I. D. Hill
Statistical Advisory Unit
Ministry of Supply
London W. 1, England

Professor M. G. Kendall
London School of Economics
London, England

Professor A. Walthar
Technische Hochschule
Darmstadt, Germany

Other Foreign Addresses

Professor Maurice H. Beltz
University of Melbourne
Carlton N. 3
Victoria, Australia

Professor Toyo Kitagawa
Mathematical Institute
Faculty of Science
Kyusyu University
Fukuoka, Japan

Kenichi Koyama
Managing Director
Union of Japanese Scientists
2, 1-chome, Chiba-shi
Chuo-ku, Tokyo, Japan

Dr. Lal Verman
Director
Indian Standards Institute
New Delhi, India

Professor P. C. Mahalanobis
Indian Statistical Institute
203 Balfour Court Road
Calcutta 35, India

N. T. Mathew
SRC Unit
India Statistical Institute
B King George Road
New Delhi, India

Professor Sigeiti Moriguti
Syam Minami-ku
Sakai-ku, Tokyo, Japan

R. O. Nanavat
SRC Unit Calcutta
Indian Statistical Institute
9B Elections East
Calcutta 1, India

H. D. Shurle
Director
National Productivity Council
Golf Links
New Delhi, India

P. V. Sivakumar
SQC Unit, Indian Statistical Inst.
Queens Road, Government Offices
Bombay, India

Srinageshwaran
SQC Unit Bangalore
VS SSKST Technological Inst.
Bangalore 1, India

Professor Eduardo Valenzuela
Republica 517-CHEF
Santiago, Chile

Mr. Cesare Villegas
Inst. de Matemativa y Estadistica
Av. J. Hererra y Reissig
Montevideo, Uruguay

Additional copies for project leaders and assistants, office files, and reserve for future requirements 70

Contract No. 225(53)
May 1961
(425)