REPORT NO. 1986

CONFIDENCE BOUNDS FOR THE GENERAL LINEAR MODEL

Malcolm S. Taylor
J. Richard Moore

May 1977

Approved for public release; distribution unlimited.

USA ARMAMENT RESEARCH AND DEVELOPMENT COMMAND
USA BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MARYLAND
Destroy this report when it is no longer needed. Do not return it to the originator.

Secondary distribution of this report by originating or sponsoring activity is prohibited.

Additional copies of this report may be obtained from the National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia 22151.

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.
In this paper, for the general linear model \( Y = X\beta + e \), we consider the construction of confidence bounds about the entire regression line. To accomplish this we exploit a powerful theorem of Scheffé. A procedure often encountered is one in which a set of confidence intervals about \( E(y|x) \) or prediction intervals for future observations are determined and then the end points are connected in such a fashion as to describe an envelope. The belief is that what has been accomplished is precisely what Scheffé's theorem allows one to do.
In addition, we present some extensions concerning confidence bounds about combinations of regression lines and suggest a useful application of these results. Specifically, we propose to use the confidence bounds about the difference of regression lines to make a quantitative assessment of when and where independent sets of data characterizing the same phenomena are in agreement or disagreement.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>5</td>
</tr>
<tr>
<td>2. CONFIDENCE REGION FOR A POLYNOMIAL</td>
<td>7</td>
</tr>
<tr>
<td>3. THE TWO SAMPLE CASE</td>
<td>9</td>
</tr>
<tr>
<td>4. THE k-SAMPLE CASE</td>
<td>12</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>16</td>
</tr>
<tr>
<td>DISTRIBUTION LIST</td>
<td>18</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

It is appropriate at the onset that we devote a few paragraphs to the introduction of the general linear hypothesis model of full rank. We want to consider uncorrelated observations \( y_1, y_2, \ldots, y_n \) that satisfy the relation

\[
y_i = \sum_{j=1}^{p} x_{ij} \beta_j + e_i, \quad i = 1, 2, \ldots, n \quad (1.1)
\]

and are linear in the unknown parameters \( \beta_1, \beta_2, \ldots, \beta_p \) with known coefficients \( x_{ij} \) and random term \( e_i \) satisfying

\[
E(y_i) = \sum_{j=1}^{p} x_{ij} \beta_j,
\]

and

\[
\text{Var}(y_i) = \sigma^2.
\]

In other words, the random term \( e_i \) is a random variable with expected value \( E(e_i) \) equal to zero and unknown variance \( \text{Var}(e_i) \) equal to \( \sigma^2 \).

The problem, in its most general sense, involves determining point and interval estimates of several quantities of interest of the model and the testing of various statistical hypotheses.

For compactness of notation and ease of manipulation let

\[
Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}
\]

then we can write the system of relations (1.1) as

\[
Y = X\beta + e
\]

and proceed to define the general linear hypothesis model of full rank as follows:
Definition 1.1. The model $Y = X\beta + e$ where $Y$ is a random observed vector, $e$ is a random vector, $X$ is an $n \times p$ matrix of known fixed quantities, and $\beta$ is a vector of unknown parameters is called the general linear hypothesis model of full rank, provided the rank of $X$ is equal to $p$ where $p \leq n$.

In the present inquiry we restrict our consideration to the normal theory case, which means the random vector $e$, already satisfying $E(e) = 0$ and $\text{cov}(e) = \sigma^2_i$, will, in addition, be assumed to be normally distributed.

The problem most frequently addressed is that of estimating the unknown parameters $\beta_j$ on the basis of the observations $y_i$. These estimates of $\beta_j$, denoted by $\hat{\beta}_j$, are functions of $y_i$; and, as such, are themselves random variables about which confidence intervals can be constructed. These ideas are fully developed in a number of textbooks.\(^1\) A point not so widely expounded is that the usual frequency interpretation of a confidence interval based on a single sample $y_1, y_2, \ldots, y_n$ holds only for a single coefficient $\beta_j$; if the same data are used to determine confidence intervals for both $\beta_i$ and $\beta_j$, $i \neq j$, the probability is not $1-\alpha$ that the confidence intervals thus constructed will simultaneously contain $\beta_i$ and $\beta_j$. The complexity is advanced by the fact that the interval estimates are not independent; so, in general, only a single confidence statement can be made from a single set of observations.

It is not our intent here to address this problem directly; such an inquiry falls into the general area of simultaneous confidence intervals. It is our intent, however, to consider a ramification of this problem: namely, the construction of a confidence envelope about the entire regression line. We will, in addition, provide some results concerning confidence envelopes about combinations of regression lines and implications of their use.

Toward this end consider the following definition due to Bose:\(^2\)

---


Definition 1.2. A parametric function $\psi$ is called an estimable function if it has an unbiased linear estimate, i.e., if there exists an n-vector $a$ of constant coefficients such that $E(a'y) = \psi$.

If $L$ is a $p$-dimensional space of estimable functions with basis $(\psi_1, \psi_2, \ldots, \psi_p)$ and $\hat{\psi}$ is the least squares estimate of $\psi \in L$, then we have the following theorem due to Scheffé:

Theorem 1.1. Under the general linear hypothesis model (normal case) the probability is $1 - \alpha$ that simultaneously for all $\psi \in L$

$$\hat{\psi} - S \sigma \leq \psi \leq \hat{\psi} + S \sigma$$

where the constant $S = \{pF_{\alpha}(p,n-r)\}^{1/2}$ and rank $X = r$.

The implications of this theorem are far reaching; and in this article we will exploit a single facet, albeit an important and useful one. To facilitate this we need to be aware of the fact that since least squares estimates $\hat{\beta}$ are BLUE, the elements of the vector $\beta$ of the general linear model of full rank form a basis of a space $L$ of estimable functions which includes polynomials as a special case.

2. CONFIDENCE REGION FOR A POLYNOMIAL

To determine a confidence region for a polynomial with observational equations

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}^2 + \ldots + \beta_{p-1} x_{i(p-1)} + e_i, \quad i = 1, 2, \ldots, n$$

in the model $Y = X\beta + e$, the $n \times p$ matrix $X = (x_{ij})$ of known constant coefficients takes the form

The least squares estimate of $\beta$ is given by $\hat{\beta} = (X'X)^{-1}X'Y$. If we choose $\psi_i = \beta_i$, $i = 0, 1, \ldots, p-1$, then $\{\psi_i\} = \{\beta_i\}$ is a set of $p$ linearly independent estimable functions which forms a basis for the space $L$. For any value $x_0$ denote $X'_0 = (1, x_0, \ldots, x_0^{p-1})$. Clearly, $E(y_0) = X'_0\beta \in L$ and has least squares estimate

$$X'_0\hat{\beta} = X'_0(X'X)^{-1}X'Y = \sum_{i=1}^{n} a_i' y_i$$

where the coefficients $a_i$ are the elements of the $1 \times n$ vector $X'_0(X'X)^{-1}X'$. Thus

$$\sum_{i=1}^{n} a_i^2 = X'_0(X'X)^{-1}X'[X'_0(X'X)^{-1}X']'$$

$$= X'_0(X'X)^{-1}X'X(X'X)^{-1}X_0$$

$$= X'_0(X'X)^{-1}X_0$$

so that $\hat{\sigma}_2 = \sigma^2 X'_0(X'X)^{-1}X_0$ with unbiased estimate $\tfrac{1}{\psi} X'_0(X'X)^{-1}X_0$.

From Scheffé's theorem we can assert with probability $1-\alpha$ that simultaneously for all $\psi \in L$ and, in particular, $X'_0\beta \in L$,

$$X'_0\hat{\beta} - \hat{\sigma}_2 \leq X'_0\beta \leq X'_0\hat{\beta} + \hat{\sigma}_2$$

where $S = [pF_\alpha(p,n-p)]^{1/2}$. 

\[
X = \begin{bmatrix}
1 & x_1 & x_1^2 & \ldots & x_1^{p-1} \\
1 & x_2 & x_2^2 & \ldots & x_2^{p-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_n & x_n^2 & \ldots & x_n^{p-1}
\end{bmatrix}
\]
As an illustration, suppose the paired data \((1.20, 0.34), \(1.37, 0.94), \(1.38, 0.99), \(1.65, 1.58), \(1.71, 2.08), \(1.82, 2.25)\) are characterized by the quadratic \(y = -0.31x^2 + 3.97x - 3.95\) over the interval of interest, \(1 \leq x \leq 2\). The 95% confidence region for the entire true line is given by

\[
X_0^{\hat{\beta}} - (5.28)\hat{\sigma}_\psi \leq X_0^{\beta} \leq X_0^{\hat{\beta}} + (5.28)\hat{\sigma}_\psi
\]
as shown in Figure 1.

Grubbs\textsuperscript{5} showed that for the case \(y = \beta_o + \beta_1x\) the confidence bounds resulting from Schéffé's theorem are

\[
\hat{\beta}_o + \hat{\beta}_1x \pm [2F_a(2, n-2)]^{1/2}\left\{\frac{1}{n} + \frac{n(x-x_bar)^2}{A_{xx}}\right\}^{1/2}
\]

where \(S = \left\{\frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{\beta}_o - \hat{\beta}_1x_i)^2\right\}^{1/2}\) and \(A_{xx} = n\Sigma x_i^2 - [\Sigma x_i]^2\).

Note that the value \(x\) appearing in (2.1) is not limited to an \(x_i\) which appears in the observations \((x_i, y_i), i = 1, 2, \ldots, n\).

3. THE TWO-SAMPLE CASE

Suppose two independent sets of data have given rise to two characterizations of the same phenomenon so that we are now confronted with what is, in essence, two models:

\[
Y_1 = X_1^{\beta_1} + e_1, \text{ an } n_1 \times p_1 \text{ problem},
\]

and

\[
Y_2 = X_2^{\beta_2} + e_2, \text{ an } n_2 \times p_2 \text{ problem}.
\]

We can still represent this situation as \(Y = X\beta + e\) where now

\textsuperscript{5} Grubbs, F. E., Linear Statistical Regression and Functional Relations, BRL Report No. 1842, November 1975. (AD #A018651)
FIGURE 1

95% CONFIDENCE LIMITS FOR A LINEAR FUNCTION
\[
Y = \begin{pmatrix}
\begin{array}{c}
Y_1 \\
Y_2
\end{array}
\end{pmatrix}, \quad X = \begin{pmatrix}
\begin{array}{c|c}
X_1 & 0 \\
0 & X_2
\end{array}
\end{pmatrix}, \quad \beta = \begin{pmatrix}
\beta_1 \\
\beta_2
\end{pmatrix}, \quad e = \begin{pmatrix}
e_1 \\
e_2
\end{pmatrix}.
\]

The least squares estimate is given, as before, by \( \hat{\beta} = (X'X)^{-1}X'Y \).

Consider now the difference of two polynomials \( y_1^* - y_2^* \in L \) with LS estimate \( X_1^* \hat{\beta}_1 - X_2^* \hat{\beta}_2 \), where \( X_1^* = (1, x_1, \ldots, x_i^{p_i-1}) \). Rewriting,

\[
X_1^* \hat{\beta}_1 - X_2^* \hat{\beta}_2 = \begin{pmatrix}
\begin{array}{c|c}
X_1^* & -X_2^*
\end{array}
\end{pmatrix} \begin{pmatrix}
\hat{\beta}_1 \\
\hat{\beta}_2
\end{pmatrix} = \begin{pmatrix}
\begin{array}{c|c}
X_1^* & -X_2^*
\end{array}
\end{pmatrix} (X'X)^{-1}X'Y
\]

where the coefficients \( a_i \) are the elements of the \( 1 \times (n_1+n_2) \) vector \( X^*(X'X)^{-1}X' \). Thus,

\[
Ea_i^2 = X^*(X'X)^{-1}X'[X^*(X'X)^{-1}X']' = X^*(X'X)^{-1}X^*X(X'X)^{-1}X^* = X^*(X'X)^{-1}X^*.
\]

Now \( X^* = \begin{pmatrix}
\begin{array}{c|c}
X_1^* & -X_2^*
\end{array}
\end{pmatrix} \) and \( (X'X)^{-1} = \begin{pmatrix}
\begin{array}{c|c}
(X_1'X_1)^{-1} & 0 \\
0 & (X_2'X_2)^{-1}
\end{array}
\end{pmatrix} \)

so \( X^*(X'X)^{-1}X^* = X_1^*(X_1'X_1)^{-1}X_1^* + X_2^*(X_2'X_2)^{-1}X_2^* \).

As before, \( \text{Var}(\hat{\psi}) \) has the unbiased estimate \( \hat{\sigma}^2 = s^2X^*(X'X)^{-1}X^* \),

where \( s^2 \) is now the pooled estimate of the variance, and with probability \( 1 - \alpha \) simultaneously for all \( \psi \in L \).
\[ X_1^* \hat{\beta}_1 - X_2^* \hat{\beta}_2 - \hat{\sigma}_\psi \leq y_1^* - y_2^* \leq X_1^* \hat{\beta}_1 - X_2^* \hat{\beta}_2 + \hat{\sigma}_\psi \]

with \( S = \left( (p_1 + p_2)^F_a(p_1 + p_2, n_1 + n_2 - (p_1 + p_2)) \right)^{1/2} \).

To illustrate one of the most useful potential applications of this result consider the situation where we are presented with two sets of data, \( \{(x_{11}, y_{11}), (x_{12}, y_{12}), \ldots, (x_{1n}, y_{1n})\} \) and \( \{(x_{21}, y_{21}), (x_{22}, y_{22}), \ldots, (x_{2m}, y_{2m})\} \), collected from the same process; and we want to say something about the similarity or dissimilarity of the two descriptions. Suppose each set is fitted with a quadratic; and we construct the confidence bound about the difference \( y_1^* - y_2^* \), as shown in Figure 2. Over the region \( (1.65 \leq x \leq 2.05) \), where the confidence bounds cover the line \( y = 0 \), we will say the two descriptions are consistent, although the associated probability level cannot be attached without qualification and interpretation.

The extension of Grubbs result (2.1) to this case is direct; the bounds take the form

\[ X_1^* \hat{\beta}_1 - X_2^* \hat{\beta}_2 \pm \left( (p_1 + p_2)^F_a(p_1 + p_2, n_1 + n_2 - (p_1 + p_2)) \right)^{1/2} \cdot S \cdot \left[ \frac{1}{n_1} + \frac{1}{n_2} + \frac{n_1(X^* - \bar{X}_1)}{A_{1x}} + \frac{n_2(X^* - \bar{X}_2)}{A_{2x}} \right]^{1/2} \]

where \( S^2 \) is the pooled estimate of variance and \( A_{xx}^i \) is computed from the \( i \)-th data set.

### 4. THE k-SAMPLE CASE

The straightforward generalization to \( k \) sets of data proceeds as follows:

\[
\begin{pmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_k
\end{pmatrix}
= \begin{pmatrix}
X_1 | 0 | \cdots | 0 \\
0 | X_2 | \cdots | 0 \\
\vdots \\
0 | 0 | \cdots | X_k
\end{pmatrix}
\begin{pmatrix}
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_k
\end{pmatrix}
+ \begin{pmatrix}
e_1 \\
e_2 \\
\vdots \\
e_k
\end{pmatrix}
\]

\[ Y_k \]
FIGURE 2

95% CONFIDENCE LIMITS FOR THE DIFFERENCE BETWEEN TWO QUADRATIC FUNCTIONS
and

\[ \hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} = \begin{pmatrix} (X_1'X_1)^{-1}X_1'Y_1 \\ \vdots \\ (X_k'X_k)^{-1}X_k'Y_k \end{pmatrix} \]

Now for \( X_* = \begin{bmatrix} x_1' & \cdots & x_*' \end{bmatrix} \) where \( X_* = \begin{bmatrix} 1, x, \ldots, x^{p_i-1} \end{bmatrix} \)

we can write \( y_* = X_*' \hat{\beta} = X_*'(X_i'X_i)^{-1}X'_iY_i \), \( i = 1, 2, \ldots, k \).

Suppose we now consider the estimate of \( \Sigma_{i=1}^{n_i} y_* \):

\[ \Sigma_{i=1}^{n_i} y_* = \Sigma_{i=1}^{n_i} X_1' \hat{\beta}_i = \Sigma_{i=1}^{n_i} (X_i'X_i)^{-1}X'_iY_i. \]

Rewriting,

\[ \Sigma_{i=1}^{n_i} X_*' \hat{\beta}_i = \begin{bmatrix} c_1X_1' \\ \vdots \\ c_kX_k' \end{bmatrix} (X'X)^{-1}X'Y = \Sigma_{i=1}^{n_i} a_i y_i \]

where the coefficients \( a_i \) are the elements of the \( 1 \times \Sigma n_i \) vector \( CX_*' (X'X)^{-1}X' \).

Thus,

\[ \Sigma_{i=1}^{n_i} a_i^2 = CX_*' (X'X)^{-1}X'[CX_*' (X'X)^{-1}X']' \]

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

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

\[ CX^* = \left[ \begin{array}{c|c|c} c_1 X_1^* & \ldots & c_k X_k^* \end{array} \right] \quad \text{and} \quad (X'X)^{-1} = \begin{pmatrix} (X'_1X_1)^{-1} & \vdots \\ \vdots & \vdots \\ (X'_kX_k)^{-1} \end{pmatrix} \]

so \( CX^* (X'X)^{-1} CX^* = \sum_{i=1}^{k} c_i^2 X_i^* (X'_iX_i)^{-1} X_i^* \). The confidence region now assumes the form

\[ \sum_{i=1}^{k} c_i X_i^* \widehat{\beta}_i \leq S \left( \sum_{i=1}^{k} c_i X_i^* \right)^{1/2} \leq \sum_{i=1}^{k} c_i X_i^* \widehat{\beta}_i + S \left( \sum_{i=1}^{k} c_i X_i^* \right)^{1/2}; \]

with \( S = \left( \sum_{i=1}^{k} \Sigma_{n_i} F_{\alpha} (\Sigma_{n_i} n_i - \Sigma_{n_i}) \right)^{1/2} \).

For the linear case we obtain

\[ CX^* (X'X)^{-1} CX^* = \sum_{i=1}^{k} c_i^2 X_i^* (X'_iX_i)^{-1} X_i^* \]

\[ = \sum_{i=1}^{k} c_i \left( \frac{1}{n_i} + n_i (x_i - \bar{x}_i)^2 \right), \]

and the two sample case (Section 3) is obtained by setting \( c_1 = 1 \) and \( c_2 = -1 \).
REFERENCES


5. Grubbs, F. E., "Linear Statistical Regression and Functional Relations", BRL Report No. 1842, November 1975. (AD #A018651)
### DISTRIBUTION LIST

<table>
<thead>
<tr>
<th>No. of Copies</th>
<th>Organization</th>
<th>No. of Copies</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Commander</td>
<td>2</td>
<td>Commander</td>
</tr>
<tr>
<td></td>
<td>Defense Documentation Center</td>
<td>US Army Mobility Equipment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: DDC-TCA</td>
<td>Research &amp; Development Command</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cameron Station</td>
<td>ATTN: Tech Docu Cen, Bldg. 315</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alexandria, VA 22314</td>
<td>DRSME-R2T</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fort Belvoir, VA 22060</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Commander</td>
<td>1</td>
<td>Commander</td>
</tr>
<tr>
<td></td>
<td>US Army Materiel Development and Readiness Command</td>
<td>US Army Armament Research and Development Command</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: DRCDMA-ST</td>
<td>ATTN: DRDAR-TD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5001 Eisenhower Avenue</td>
<td>Dover, NJ 07801</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alexandria, VA 22333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Commander</td>
<td>1</td>
<td>Commander</td>
</tr>
<tr>
<td></td>
<td>US Army Aviation Systems Command</td>
<td>US Army Armament Materiel and Readiness Command</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: DRAV-E</td>
<td>Rock Island, IL 61202</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12th and Spruce Streets</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>St. Louis, MO 63166</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Director</td>
<td>1</td>
<td>Director</td>
</tr>
<tr>
<td></td>
<td>US Army Air Mobility Research and Development Laboratory</td>
<td>US Army TRADOC Systems Analysis Activity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: DRESH-EL</td>
<td>ATTN: ATAA-SA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ames Research Center</td>
<td>White Sands Missile Range</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moffett Field, CA 94035</td>
<td>NM 88002</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Commander</td>
<td>1</td>
<td>Commander</td>
</tr>
<tr>
<td></td>
<td>US Army Electronics Command</td>
<td>US Army Tank Automotive Development Command</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: DROTA-RW</td>
<td>ATTN: DROTA-RWL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fort Monmouth, NJ 07703</td>
<td>Warren, MI 48090</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Commander</td>
<td></td>
<td>Aberdeen Proving Ground</td>
</tr>
<tr>
<td></td>
<td>US Army Missile Research and Development Command</td>
<td>Marine Corps Ln Ofc</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: DRDMD-R</td>
<td>Dir, USAMSAA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Redstone Arsenal, AL 35809</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>