Best Available Copy
TECHNICAL REPORT NO. 10

ESTIMATING A LINEAR FUNCTIONAL RELATION

Prepared Under Contract No. DA-36-034-ORD-1517 (RD)
(Experimental Designs for Industrial Research)

Dept. of Army Project No. 5B99-01-004
Ordnance Project No. TB2-0001
CRR Project No. 832

Philadelphia Ordnance District
Department of the Army, Department of Defense
with
Institute of Statistics
North Carolina State College of
The University of North Carolina
Raleigh, North Carolina

Technical Supervisor
Ballistics Research Laboratories
Aberdeen Proving Ground
Aberdeen, Maryland

H. F. Smith
Author of Report
January 1956
1. Introduction and review

How to evaluate from observations, all subject to error, an estimate of a functional relation has been a persistent statistical problem for 80 years. When the relation is linear, when errors of observation are normally distributed, and when either nothing is postulated about distribution of the underlying hypothetical variables or they are assumed to be also normally distributed random variables, a consistent solution is possible only if the ratios of variances and covariances of the errors are known (or alternatively all but one of the second moment parameters of the error distributions). We shall consider here a linear relation between two variates with homogeneous error variances. Usually one assumes the covariance to be zero and the ratio of variances known. With that condition the solution was given by Kummell in 1879 and has been several times rediscovered. Nevertheless the theoretical foundation for Kummell's solution has remained ambiguous; it is known to be consistent, but beyond that its statistical properties such as bias and efficiency have not been investigated. Miss Dent (1935) seems to have been the only writer to attempt to evaluate the sampling variance of the estimated slope and her solution is far from satisfactory. It ignores distinction between parameters and statistics, it is based on a Taylor expansion which is not always convergent, and as it seeks the variance of tan (2\(\beta\)) it degenerates toward infinity in the most important region where the slope, tan \(\beta\), is near unity.

The purpose of this paper is to show why Kummell's solution is unique; thence to prove that it is efficient and unbiased, with respect to the angle of the line with either coordinate axis; and to obtain its sampling distribution.
Consider the following model. Two variables, $y_1$ and $y_2$, are linearly related:

$$y_1 = A + B y_2$$  \hspace{1cm} (1)

or

$$y_1 \cos \beta - y_2 \sin \beta - \kappa = 0$$  \hspace{1cm} (2)

Experimentation yields paired observations

$$y_{pi} = y_{pi} + \delta_{pi}$$

$$p = 1, 2; \ i = 1 \ldots n$$

The errors of observation, $\delta_1$, $\delta_2$, are assumed to be random variables normally independently distributed with zero means and common variance $\sigma_0^2$. Except where otherwise stated nothing is postulated about the distribution of $y_1$ (or equivalently, owing to the relation (1), of $\eta_2$). The model is illustrated in fig. 1 where circles represent equal frequency contours of the distribution of $\delta_p$. In particular we will consider the circles with radius equal to $\sigma_0$. Only two sub-populations are shown in the figure although usually $n$ would be substantially greater than two. The relationship (1) is represented by the line $AA'$.  

We shall consider also the more general model where errors of observations are not independently distributed with equal variances. Let the observations then be denoted $x_{pi}$, and assume them to be normally distributed with variances $\sigma_{11}$, $\sigma_{22}$ and covariance $\sigma_{12}$ around centers $\xi_{pi}$ which obey the relation

$$\xi_1 = A' + B' \xi_2$$  \hspace{1cm} (3)

This model is diagrammed in fig. 2. We consider in particular the contour ellipse

$$\sigma_{22} \xi_1^2 - 2 \sigma_{12} \xi_1 \xi_2 + \sigma_{11} \xi_2^2 = \sigma_{11} \sigma_{22} - \sigma_{12}^2$$  \hspace{1cm} (4)
where \( \varepsilon_p = x_p - \frac{c}{p} \). The length of any radius of this ellipse is the standard deviation of any section of the bivariate frequency distribution in the same direction; and the projection of the ellipse on any line is the standard deviation of the marginal distribution of the sub-population projected onto that line.

Provided we know either two of the quantities \( \sigma_{pq} \), or two ratios among them, this model can be transformed to the previous one by

\[
y_1 = a_1x_1 + a_2x_2
\]

\[
y_2 = b_1x_1 + b_2x_2
\]

subject to

\[
a^2_1\sigma_{11} + 2a_1a_2\sigma_{12} + a^2_2\sigma_{22} = b^2_1\sigma_{11} + 2b_1b_2\sigma_{12} + b^2_2\sigma_{22}
\]

\[
a_1b_1\sigma_{11} + (a_1b_2 + a_2b_1)\sigma_{12} + a_2b_2\sigma_{22} = 0
\]

a transformation which can be made in many ways. All deductions about the "y" model can be transferred to the "x" model by the transformation.

Efforts to find a means of estimating the functional relation (1) or (3), under conditions stated, have followed one of three lines, namely:

I by considering criteria of consistency, II least squares, III maximum likelihood. (We exclude from consideration here estimates from moments of higher than second order which become available when \( \frac{c}{p} \) is postulated to have a non-normal frequency distribution.) The following review quotes only a few of the papers on the topic to indicate salient features of the literature.

I: Criteria of consistency are exemplified by the proposals of Gini (1921), Seares (1944, 1945) and Hald (1952). Usually this method seeks to apply an adjustment to the regression lines. If the hypothetical values were known
with means zero we would have

$$B = \sum_1^n \eta_1 \eta_2 / \sum_1^n \eta_2^2$$

Consistent estimators for the numerator and denominator are $\sum_1^n y_1 y_2$ and $\sum_1^n y_2^2 = n \sigma_0^2$ respectively. There are variants. With certain assumptions the Kummell line may be indicated in this way (Lindley, 1947, sec. 7.3).

II. The method of least squares is the commonest approach. The usual idea has been to minimize a sum of squares of deviations of observations from the fitted line. The problem has been to determine in what direction should the deviations be measured.

Adcock (1878-79) and Pearson (1901) minimized sum of squares perpendicular to the line without attention to the ratio of error variances. Roos (1937) however pointed out that this produces a solution which fails to be invariant under change of scale. He considered that the direction in which deviations are measured should depend only on the precisions of the observations and be independent of the slope of the line. He therefore proposed to use deviations at $45^\circ$ to either coordinate axis when variates are scaled so as to have equal precisions. But Lindley (1947, sec. 8.2) pointed out that even this gives consistent estimates only under rather special conditions.

Kummell (1879) and Deming (1931-43)(assuming $\sigma_{12} = 0$) proposed to minimize

$$S = \sum \left\{ \frac{(x_1 - \bar{x}_1)^2}{\sigma_{11}} + \frac{(x_2 - \bar{x}_2)^2}{\sigma_{22}} \right\}$$

(6)
or, equivalently, a proportionate expression using only the ratio 
\[ \lambda = \frac{\sigma_{11}}{\sigma_{22}} \], subject to the restriction (3). Kummell showed that a solution is obtainable only if \( \lambda \) be known and that it is equivalent to minimizing the sum of squares of deviations perpendicular to the line when the variables are scaled so as to have equal error variances, that is when transformed to the form (1). He reached the well known solution of the quadratic equation

\[ (S_{11} - \lambda S_{22})B + (\lambda - B^2)S_{12} = 0 \]  

(7)

where \( S_{11} = \sum (x_1 - \bar{x}_1)^2 \), etc.

To fit curves and planes he proposed an approximate method which shows interesting variation on usual procedure. When residuals are not linear in the parameters the classical least squares method begins by expressing the residuals (before squaring) as the linear terms of a Taylor expansion in adjustments to trial parameter values. Kummell first expresses \( S \) as a Taylor expansion linear only in the deviations \( (x_p - \xi_p) \) and uses this to eliminate the \( \xi \) and obtain for the residuals (e.g., \( x_1 - A - B'x_2 \)) so-called weights which are usually functions of the parameters. The whole expression \( S \), with \( \xi \) thus eliminated, is then expressed as a Taylor expansion in adjustments to trial parameter values, and with the 'weights' remaining as functions of the parameters in subsequent differentiations. (Roos criticizes laxity in his mathematical arguments.)

Deming's approximating procedure is more similar to the usual least squares approximation but differs in that the expansion of the residuals contains simultaneously terms both in \( (x_p - \xi_p) \) and in parameter adjustments. It leads to 'weighting' of the observable residuals similar to Kummell's
formulation, but with the weights expressed as functions of the trial parameter values and hence entering differentiation as constants. In classical procedure the parameter adjustments tend to zero as iteration proceeds with convergence on the required estimates. Doming's expansion differs in the material respect that the discrepancies \( x_p - \hat{x}_p \) do not tend to zero and cannot be made to do so without losing contact with the observations which must remain as the anchor for computations. His book repeatedly reiterates that the solutions obtained will differ from the true ones only in squares of the residuals, but it evades enunciating the corollary that since these same squares of residuals constitute the function on whose minimization the solution depends they are not negligible. When the fitted relation is linear the procedure leads to one of the regression lines (albeit with a modified estimate of the error variances). In this case it fails to distinguish (as it appears to purport to do) between regression and the functional relation, and the proposed weighting procedure is wasted effort. When a curved relation is to be estimated it does allow that observations in regions of very steep slope do not get the very high weights which regression would in effect assign to them. But just what may be accomplished seems not to have been threshed out. Recognition of the order of magnitude of neglected terms suggests that bias may still be about as great as by any simpler method; it suggests indeed that the method may be strictly appropriate only when residuals are so small that almost any method of fitting will yield a satisfactory result. (A similar comment was made by J. H. Smith, 1945.)
Lindley (1947) seeks a least squares solution by considering the squares of residuals defined by the form in which the relation happens to be written; in particular using (3) he writes the residuals as

\[ (x_1 - A^t - B^t x_2) \]

With assumptions \( \sigma_{11}/\sigma_{22} = \lambda \) and \( \sigma_{12} = 0 \), he then notes that the variance of these residuals is proportional to \( (\lambda + B^t B) \) and states that this must be introduced as a weight so that the function to be minimized is

\[ \sum \frac{(x_1 - A^t - B^t x_2)^2}{\lambda + B^t B} \]

This produces Kummell's equation, but to imply that it is a weighted least squares solution is, I think, misleading. Weights are introduced into the least squares procedure to take account of variation in precision of the observations. (I omit these here for simplicity. They, Lindley's \( P_k \), \( \sigma_k \), are easily added if required.) Here all the observations have equal weight, and \( (\lambda + B^t B) \) is not a weight in the ordinary sense. Lindley pertinently remarks that his procedure has the advantage "that the redundant \( \xi \) are never mentioned", but he does not explain why they can be thus banished.

III. The maximum likelihood solution has been considered by Dent (1935), Lindley (1947) and Kendall (1950, 1954). They begin by writing the likelihood, assuming \( \sigma_{12} = 0 \), as

\[ (2\pi)^{-n}(\sigma_{11}\sigma_{22})^{-n/2} \exp \left\{ -\frac{1}{2} \sum \left[ \frac{(x_2 - \xi_2)^2}{\sigma_{22}} + \frac{(x_1 - A^t - B^t \xi_1)^2}{\sigma_{11}} \right] \right\} \]

and treat \( \xi_{21} \) as parameters. With no further postulates Lindley and Kendall conclude that all is not well with the resultant equations because they lead to the ratio of estimates of \( \sigma_{22}/\sigma_{11} \) being \( \lambda^2 \) -- an unacceptable result.
While admitting the result to be unacceptable it is scarcely by itself adequate reason for rejecting the method and they do not elucidate why this result appears. Since the likelihood is formulated for 2n observations indeterminance is not due, as has been suggested, to trying to estimate more parameters than there are observations. We shall see later that the trouble is that, relative to distribution of the only deviations which are observable, these two parameters enter as a single unit.

This formulation, with the ratio $\sigma_{11}/\sigma_{22}$ known, is easily seen to be equivalent to Kummell's least squares formulation. The unpleasant feature lies in regarding $\xi_{pi}$ as parameters to be estimated, Neyman and Scott's (1947) "incidental parameters". The word "parameter" as used in statistics has not yet been very specifically defined. Relative to the theory of maximum likelihood it may be defined as a characteristic constant of a probability distribution. To specify a particular parameter we must be able to specify the population of random variables of which it is a characteristic. That being done it is at least theoretically possible to return again and again to resample the specified population, thereby increasing the sample size from which its characters may be estimated. But that is just what, in the problem before us, cannot be done. A basic feature of the problem is that we never know and can never state that any two or more pairs of observations are drawn from the same sub-population with a particular $\xi_{11}$, $\xi_{21}$; they are never definable as the characters of a specifiable population and we can never increase sample sizes for their estimation as is required to demonstrate the optimum properties of maximum likelihood estimators. Furthermore we are not interested in the $\xi_{pi}$ individually, we are concerned only to estimate the line as a whole. The $\xi_{pi}$ are essentially variables, not parameters at all. They may or may not
be random variables, that is variables with an associated probability
distribution, depending on the procedure by which observations are selected.
If they are random variables the parameters of their distribution may become
parameters also of the overall distribution of the observable $x$ and appro-
priate formulation is self evident (appendix 2). If the location of observa-
tions is chosen in a way which precludes assigning a probability distribution
to $\xi$, to treat them as parameters may yet be inappropriate, and we should
seek some other method to eliminate them from the problem.

2. Least squares formulation

The method of least squares seeks estimates of parameters which minimize
a sum of squares of residuals which are usually expressible as

$$x_i - \theta_i$$

(10)

where $x_i$ are observations, and $\theta_i$ are functions of the estimands and known
constants. For the Gauss-Markoff theorem to be applicable, with consequent
nice statistical properties of the estimators, it is necessary (David and
Neyman, 1937) that

(i) $E(x_i) = \theta_i$

(ii) $\theta_i$ be a linear function of the estimands

(iii) the relative weights of $x_i$ be known.

Little seems to be known about the precise statistical properties of estimators
when these conditions are not met. Conditions (i) and (iii) are invariably
assumed. Failure of (ii) creates no difficulty in principle for estimating
the parameters, beyond that the solution may have to be approached by itera-
tion; but estimators may no longer be unbiased, and reliability of approxima-
tions to their variances and covariances, based on linear approximations,
seems uncertain.
When we have to deal with a relation between observations all of which are subject to error we do not obtain a clean separation between observations and estimands as at (10). We must deal with residuals which are mixed functions of the two, for example from (1) or (2);

\[ r = y_1 - A - B y_2 \text{ or } y_1 \cos \beta - y_2 \sin \beta - \alpha. \]

This additional complication has been slurred over by writers who have endeavored to apply the principle of least squares to such situations.

The residuals of classical least squares, and relative to which the principle was developed, are univariate quantities. The residuals with which we have to deal are compounded of two variates, but once compounded in a defined manner the compound becomes again a univariate quantity and should have a univariate probability distribution. The only quantities on which inference must be based are deviations from observed points to the fitted line; there are no such quantities. The hypothetical \( \xi \) or \( \eta \), being irrelevant to the problem, if they can be eliminated, the deviations must be defined by measurement in some specified direction.

It seems natural to assume as a first requirement that the residuals should be formulated so that \( E(r) = 0 \). That condition is satisfied for all formulations which have been proposed when expectation is taken to imply expectation over the bivariate distribution of \( y_1 \) and \( y_2 \), or of \( x_1 \) and \( x_2 \), in each sub-population. But after noting that the residuals are essentially univariate quantities it seems reasonable to consider their distribution about the line as conditioned by the direction in which it has been decided to measure them (cf. app. 1). Suppose one may decide to measure the deviations parallel to some arbitrary direction BB', fig. 2. The mean values of conditional distributions formed by sectioning a bivariate distribution in that
direction then lie on a diameter $CC'$ which bisects all chords of the contour ellipse which are parallel to $BB'$. The conditional expectations of $r$ are then quantities $QR$ whose magnitude depends on the distance of $BB'$ from $\frac{x}{1} \frac{y}{2}$. The net effect of minimizing a sum of squares of such deviations is to bias the estimated line in the direction $CC'$. However such bias can be eliminated if the deviations be measured in one, and only one, particular direction. If the deviations be measured in a direction parallel to tangents to an equi-frequency contour ellipse at its intersection with the functional line $AA'$, that is in the direction $DD'$, fig. 2, the locus of expectations $CC'$ then coincides with $AA'$ and the expectations of deviations ($QR$) are identically zero independently of $\frac{x}{1}, \frac{y}{2}$. When deviations are so measured we may regard them as normally distributed about the line $AA'$, and thus, and only thus, the incidental variables $\frac{z}{5}$ can be eliminated from the problem.

When transformed to our "y" model, figure 1, the condition is seen to be when deviations for independent errors of equal variance are measured perpendicular to the line. Furthermore the variance of such deviations is immediately seen to be $\sigma^2$. In figure 2 the standard deviation is equal to the length of the radius vector, $TD''$, parallel to $DD'$, for the ellipse $(l)$; but the expression for this length is not simple.

3. Maximum likelihood formulation

For simplicity consider the standardized "y" model, figure 1. The deviations to be considered are those perpendicular to the line. From any text book of analytical geometry (or from elementary trigonometry) the distance of a point from the line is

$$ (y_1 - A - By_2)(1 + B^2)^{-1/2} \quad (11) $$
We have seen that these deviations are normally distributed with mean zero and variance $\sigma^2$. The likelihood is therefore

$$L = (2\pi\sigma^2)^{-n/2} \exp \left[ -\frac{1}{2\sigma^2} \sum \frac{(y_1 - A - By_2)^2}{1 + B^2} \right] (12)$$

It is maximized when $\sum (y_1 - A - By_2)^2 / (1 + B^2)$ is minimized. This is Lindley's formulation but we now see that the factor $(1 + B^2)$ (or, more generally, $\lambda + B'^2$) is introduced to evaluate the only deviation from whose distribution the redundant $\xi$ can be eliminated. It is not entered as a weight; that it happens to be proportional to the variance of the marginal distribution when the whole bivariate distribution is projected onto a vertical, in which direction deviations are $(y_1 - A - By_2)$, is an incidental circumstance of the bivariate normal distribution. Since the Kummell solution can now be expressed as a maximum likelihood estimator not involving "incidental parameters", it follows that it is asymptotically efficient.

For further work the equation of the functional line is better expressed in the intercept form (2). In regression analysis the slope of a regression is normally distributed only because values of the independent variable are taken as given constants. Furthermore we cannot evaluate a regression unless these have some spread, that is $\sum (x - \bar{x})^2$ cannot be zero, and the estimated slope cannot approach infinity. In the problem here considered $B$ may approach infinity from either direction while the spread of $x_2$ is still substantial (depending on $\sigma_{22}$). A little consideration shows that the system is circularly symmetric, therefore to obtain a statistic which may be symmetrically distributed and independent of $\beta$ we should consider the angle of the estimated line, rather than its slope whose distribution may be very skew and dependent on $\beta$. 
Taking the parameters to be estimated as $\beta$ and $\alpha$, the log likelihood is now

$$\ln L = \text{const.} - \frac{n}{2} \ln \sigma^2_0 - \frac{1}{2\sigma^2_0} \sum(y_1 \cos \beta - y_2 \sin \beta - \alpha)^2 \quad (13)$$

To derive the estimators assume that $y_1, y_2$ are measured from their means so that $\bar{y}_1, y_p = 0$.

$$\frac{\partial \ln L}{\partial \alpha} = \frac{1}{\sigma^2_0} \sum (y_1 \cos \beta - y_2 \sin \beta - \alpha)$$

whence $\hat{\alpha} = 0$.

$$\frac{\partial \ln L}{\partial \beta} = \frac{1}{\sigma^2_0} \sum (y_1 \cos \beta - y_2 \sin \beta - \alpha)(y_1 \sin \beta + y_2 \cos \beta) \quad (15)$$

which is zero when

$$\sum \left( \frac{1}{2} (y_1^2 - y_2^2) \sin 2\beta + y_1y_2 \cos 2\beta \right) = 0$$

$$\tan 2\hat{\beta} = \frac{2 \sum y_1y_2}{\sum y_2^2 - \sum y_1^2} \quad (16)$$

which is equivalent to Krummel's solution (7).

Further deductions are easier in terms of transformed variables obtained by rotating to coordinate axes parallel and perpendicular to the theoretical line. Define

$$u_1 = y_1 \cos \beta - y_2 \sin \beta \quad (17)$$

$$u_2 = y_1 \sin \beta + y_2 \cos \beta$$

$$E(u_1) = \eta_1 \cos \beta - \eta_2 \sin \beta = \alpha \quad \text{(Constant independently of } i)$$

$$E(u_{2i}) = \gamma_1^i \sin \beta + \gamma_2^i \cos \beta = u_{1i} \quad \text{say} \quad (18)$$
U is an alternative variable to \( \eta_1 \) or \( \eta_2 \) and represents the distance of hypothetical points as measured along the line, ST in figure 1. We then have

\[
\ln L = \text{const.} - \frac{n}{2} \ln \sigma^2 - \frac{1}{2 \sigma^2} \sum (u_1 - \lambda)^2
\]

(19)

\[
\frac{\partial \ln L}{\partial \sigma^2} = - \frac{n}{2 \sigma^2} + \frac{1}{2 \sigma^4} \sum (u_1 - \lambda)^2
\]

= 0 when \( \sigma^2 = \frac{1}{n} \sum (u_1 - \lambda)^2 \) (20)

Since \( du_1/\partial \beta = -u_2 \), and \( u_2 \) is a constant relative to each conditional distribution along a fixed line DD', the elements of the information matrix are

\[
\begin{align*}
E \left( \frac{\partial^2 \ln L}{\partial \lambda^2} \right) &= - \frac{\partial^2 \ln L}{\partial \sigma^2} = \frac{n}{\sigma^4} \\
E_0 \left( \frac{\partial^2 \ln L}{\partial \lambda / \partial \beta} \right) &= - \frac{\partial^2 \ln L}{\partial \lambda / \partial \beta} = \frac{\lambda u_2}{\sigma^2} \\
E_0 \left( \frac{\partial^2 \ln L}{\partial \beta^2} \right) &= - \frac{\partial^2 \ln L}{\partial \beta^2} = \frac{\sum u_2^2}{\sigma^4} \\
E \left( \frac{\partial^2 \ln L}{\partial \sigma^2} \right) &= - E \left( \frac{\partial^2 \ln L}{\partial \sigma^2} \right) = \frac{n}{2 \sigma^4}
\end{align*}
\]

(21)

where \( E_0 \) implies conditional expectation given \( u_2 \). However we have here the peculiarity that as \( \beta \) is varied there is a shift in the directions along which the deviations are measured. On more accurately evaluating \( \text{var} \left( \frac{\partial \ln L}{\partial \beta} \right) \) it will turn out, as we might intuitively expect, that a better approximation to the variance-covariance matrix is given by replacing \( u_2 \) by \( U \) wherever it occurs in (21). A peculiar feature of this formulation is that, allowing \( u_2 \) to be a variable, \( E(u_{21}) = U_1 \) and \( du_2/\partial \beta = u_1 \), we obtain

\[
- E \left( \frac{\partial^2 \ln L}{\partial \beta^2} \right) = \frac{\sum u_2^2}{\sigma^2}
\]

as required; but the alternative form \( E \left( \frac{\partial \ln L}{\partial \beta} \right)^2 \) still remains as \( \frac{\sum u_2^2 + n \sigma^2}{\sigma^2} \). The product
term $\frac{\partial \ln L}{\partial \alpha} \cdot \frac{\partial \ln L}{\partial \beta}$ on the other hand has the same expectation as $\frac{\partial^2 \ln L}{\partial \alpha \partial \beta}$.

Full theoretical implication of that inconsistency eludes me; it is evidently associated with the condition that the direction in which the postulated sampling distribution of observations is defined "wobbles" with errors in estimating its parameters. We shall see later that the condition under which the indicated variances are approached is not $n \to \infty$, but that the ratio $(\sigma^2 / \text{second moment of } U) \to 0$.

The asymptotic variance-covariance matrix is therefore indicated to be

$$
\sigma^2 \begin{bmatrix}
\frac{1}{n} + \frac{\bar{U}^2}{A} & -\frac{\bar{U}}{A} & 0 \\
-\frac{\bar{U}}{A} & \frac{1}{A} & 0 \\
0 & 0 & \frac{2\sigma^2}{n}
\end{bmatrix}
$$

where $\bar{U} = \frac{1}{n} \sum U/n$, $A = \frac{1}{n} \sum (U - \bar{U})^2$.

Alternative formulation of the likelihood, and of the associated variance-covariance matrix of the estimators, if estimates of individual $U_i$ may be deemed relevant, is indicated in appendix 1. The likelihood formulation if $U$ is also a normally distributed random variable is stated in appendix 2.
4. Summary of additional results

The distribution of \( \hat{\beta} \) is symmetric. By approximating \( (\hat{\beta} - \beta) \) by a power series of variables, whose joint moment generating function can be easily obtained, it has been ascertained that for finite samples the variance of \( \hat{\beta} \) is

\[
\frac{\sigma^2}{\lambda} (1 + w) + \frac{4\sigma^4}{A^2} (2w + w^2) + \frac{6\sigma^6}{3A^3} (-2 + 27w + 48w^2 + 16w^3) + O(n^{-4})
\]

where \( w = \frac{(n - 1)\sigma^2}{\lambda} \).

The parameter of kurtosis is

\[
\gamma_2 = \frac{2\sigma^2}{\lambda} (2 + w - \frac{3w^3}{1 + w^2}) + O(n^{-2})
\]

If

\[
\hat{u}_1 = (y_1 - \bar{y}_1) \cos \hat{\beta} - (y_2 - \bar{y}_2) \sin \hat{\beta}
\]

then

\[
\mathbb{E}(\hat{u}_1^2) = \sigma^2 \left[ \frac{n-2-w}{n-1} (1 - 2w - w^2) - \frac{2w^2}{(n-1)^2} (11 + 16w + 9w^2 + 3w^3) - 0(n^{-3}) \right].
\]

Since \( (n-2) \) will usually be large relative to \( w \) we may for most practical purposes suppose \( (n-2) \) "degrees of freedom". If then we accept as estimator of the error variance

\[
\hat{\sigma}^2 = \frac{\hat{u}_1^2}{(n-2)}
\]

its mean square error is

\[
\mathbb{E}(\hat{\sigma}^2 - \sigma^2)^2 = \frac{2\sigma^4}{n-2} - \frac{\hat{\sigma}^4 (16(w + w^2))}{(n-2)^2} + 0(n^{-3})
\]

If

\[
\hat{u}_2 = (y_1 - \bar{y}_1) \sin \hat{\beta} + (y_2 - \bar{y}_2) \cos \hat{\beta}
\]
a satisfactory estimator for the variance of $\beta$ is

$$\hat{\text{var}}(\hat{\beta}) = Q(1 + (n-1)Q)$$

(a6)

where $Q = \frac{\sum u_1^2}{(n - 2) \sum u_2^2 - n \sum u_1^2}$

(a7)

$$E(\text{var}(\hat{\beta})) = \frac{\sigma^2}{A} (1 + w + \frac{1}{n-1}) + \frac{\sigma^4}{A^2} (9 + 22w - 8w^2) + O(n^{-4})$$

(a8)

which compares reasonably favorably with the value required for an unbiased estimator as indicated by (a1).

An approximate fiducial interval for $\beta$ is given by

$$\sin^2 2\theta = \frac{4t^2}{n-2} \cdot \frac{S_1S_2 - S_c^2}{(S_2-S_1)^2 + 4S_c^2}$$

(a9)

where $\theta = \hat{\beta} - \beta$

$$S_p = \sum (y_p - \bar{y}_p)^2$$

$$S_c = \sum (y_1 - \bar{y}_1)(y_2 - \bar{y}_2)$$

$t = \text{Student's } t$ for the required fiducial probability and $(n-2)$ degrees of freedom.

The error in this interval seems likely to be negligible for the kind of data to which such lines are usually fitted, say for fiducial intervals of less than $n/4$. 
I have been unable to obtain the exact sampling distribution of \( \beta \) under the model so far assumed. But if we add the postulate that \( U \) is normally distributed about zero with variance \( \sigma_U^2 \), then, with notation as follows:

\[
\begin{align*}
\sigma_1^2 &= \sigma^2(u_1) = \sigma^2 \\
\sigma_2^2 &= \sigma^2(u_2) = \sigma_U^2 + \sigma^2 \\
Q &= \sigma_1^2 + \sigma_2^2 \\
C &= \sigma_1^2 \cos^2 \theta + \sigma_2^2 \sin^2 \theta
\end{align*}
\]

where \( n \) is even = \( 2k + 4 \).

The probability density function of \( \Theta = (\hat{\beta} - \beta) \) is

\[
f(\Theta) = \frac{2^{n-2}(n-2)}{n} \frac{(\sigma_1 \sigma_2)^{n-1}}{Q^2} \frac{k}{\gamma} \frac{\gamma(k+1) \left( \frac{C}{\gamma} \right)^{i}}{\left( \frac{k+1}{i} \right)} \quad (a10)
\]

Proof of these statements, and the appendices noted above, will be submitted in a later report.
<table>
<thead>
<tr>
<th>Agency</th>
<th>No. of Technical Copies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Office of Ordnance Research</td>
<td>5</td>
</tr>
<tr>
<td>Box CH, Duke Station</td>
<td></td>
</tr>
<tr>
<td>Durham, North Carolina</td>
<td></td>
</tr>
<tr>
<td>Chief of Ordnance</td>
<td></td>
</tr>
<tr>
<td>Department of the Army</td>
<td>2</td>
</tr>
<tr>
<td>Washington 25, D.C.</td>
<td></td>
</tr>
<tr>
<td>Attn: ORDTB-PS</td>
<td></td>
</tr>
<tr>
<td>Director</td>
<td></td>
</tr>
<tr>
<td>Air University Library</td>
<td>1</td>
</tr>
<tr>
<td>Maxwell Air Force Base, Alabama</td>
<td></td>
</tr>
<tr>
<td>Office of Naval Research</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D.C.</td>
<td>1</td>
</tr>
<tr>
<td>Commanding General</td>
<td></td>
</tr>
<tr>
<td>Aberdeen Proving Ground, Maryland</td>
<td>2</td>
</tr>
<tr>
<td>Attn: ERL</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Picatinny Arsenal</td>
<td>1</td>
</tr>
<tr>
<td>Dover, New Jersey</td>
<td></td>
</tr>
<tr>
<td>Commanding General</td>
<td></td>
</tr>
<tr>
<td>Rock Island Arsenal</td>
<td>1</td>
</tr>
<tr>
<td>Rock Island, Illinois</td>
<td></td>
</tr>
<tr>
<td>Commanding General</td>
<td></td>
</tr>
<tr>
<td>Research and Engineering Command</td>
<td>1</td>
</tr>
<tr>
<td>Army Chemical Center, Maryland</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Diamond Ordnance Fuze Laboratories</td>
<td>2</td>
</tr>
<tr>
<td>Washington 25, D.C.</td>
<td></td>
</tr>
<tr>
<td>Attn: Mr. J. H. Hinman, Jr. Technical Director</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Watertown Arsenal</td>
<td>1</td>
</tr>
<tr>
<td>Watertown 72, Massachusetts</td>
<td></td>
</tr>
<tr>
<td>Technical Reports Library</td>
<td></td>
</tr>
<tr>
<td>SCRL, Evans Signal Corps. Lab.</td>
<td>1</td>
</tr>
<tr>
<td>Belmar, New Jersey</td>
<td></td>
</tr>
<tr>
<td>Agency</td>
<td>No. of Technical Copies</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------</td>
</tr>
</tbody>
</table>
|	Commanding Officer  
Engineering Res. and Dev. Laboratories  
Fort Belvoir, Virginia | 1 |
| Commander  
U.S. Naval Proving Ground  
Dahlgren, Virginia | 1 |
| Chief, Bureau of Ordnance (AD3)  
Department of the Navy  
Washington 25, D. C. | 1 |
| U.S. Naval Ordnance Laboratory  
White Oak, Silver Spring 19, Maryland  
Attn: Library Division | 1 |
| Director  
National Bureau of Standards  
Washington 25, D. C. | 1 |
| National Bureau of Standards  
Washington 25, D. C.  
Attn: Applied Mathematics Division | 1 |
| Corona Laboratories  
National Bureau of Standards  
Corona, California | 1 |
| Commanding General  
Frankford Arsenal  
Brooklyn Station  
Philadelphia 37, Pennsylvania  
Attn: CRDDU-L2 | 1 |
| Technical Information Service  
P.O. Box 62  
Oak Ridge, Tennessee  
Attn: Reference Branch | 1 |
| Commanding General  
Redstone Arsenal  
Huntsville, Alabama  
Attn: CRDDU-MR | 1 |
| Commanding General  
Signal Corps Engineering Lab.  
Fort Monmouth, New Jersey  
Attn: Director of Research | 1 |
<table>
<thead>
<tr>
<th>Agency</th>
<th>No. of Technical Copies</th>
</tr>
</thead>
</table>
| The Director  
Naval Research Laboratory  
Washington 25, D.C.  
Attn: Code 2021 | 1 |
| Jet Propulsion Laboratory  
California Institute of Technology  
1300 Oak Grove Drive  
Pasadena, California  
Attn: Dr. A. J. Stosick | 1 |
| Commanding Officer  
Wentworth Arsenal  
Wentworth 72, Massachusetts  
Attn: OMRO | 1 |
| U.S. Atomic Energy Commission  
Document Library  
19th and Constitution Ave.  
Washington 25, D.C. | 1 |
| Director, Applied Physics Lab.  
Johns Hopkins University  
8621 Georgia Avenue  
Silver Spring 19, Maryland  
Attn: Dr. R. G. Herman | 1 |
| Commanding Officer  
Philadelphia Ordnance District  
128 North Broad Street  
Philadelphia 2, Pennsylvania  
Attn: Operations Division, R and D Branch | 1 |
| Commanding Officer  
Office of Naval Research  
Branch Office  
Navy, 100, PFO  
New York, New York | 2 |
| Commanding General  
Air Res. and Dev. Command  
P. O. Box 1395  
Baltimore 3, Maryland  
Attn: RDR | 1 |
<table>
<thead>
<tr>
<th>Agency</th>
<th>No. of Technical Copies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commanding General Air Res. and Dev. Command</td>
<td>1</td>
</tr>
<tr>
<td>P.O. Box 1395</td>
<td></td>
</tr>
<tr>
<td>Millisboro 3, Maryland</td>
<td></td>
</tr>
<tr>
<td>Attn: ROCLIL(Technical Library)</td>
<td></td>
</tr>
<tr>
<td>Armed Services Tech, Info, Agency</td>
<td>5</td>
</tr>
<tr>
<td>Document Service Center</td>
<td></td>
</tr>
<tr>
<td>Kent Building</td>
<td></td>
</tr>
<tr>
<td>Main and Main Streets</td>
<td></td>
</tr>
<tr>
<td>Dayton 2, Ohio</td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td></td>
</tr>
<tr>
<td>Wright Air Development Center</td>
<td></td>
</tr>
<tr>
<td>Wright-Patterson Air Force Base, Ohio</td>
<td></td>
</tr>
<tr>
<td>Attn: WDCRO</td>
<td></td>
</tr>
<tr>
<td>Chief of Ordnance</td>
<td>1</td>
</tr>
<tr>
<td>Department of the Army</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td></td>
</tr>
<tr>
<td>Attn: ORDU-SE</td>
<td></td>
</tr>
<tr>
<td>For Transmittal to:</td>
<td>1</td>
</tr>
<tr>
<td>Canadian Joint Staff</td>
<td></td>
</tr>
<tr>
<td>2001 Connecticut Avenue N. W.</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td></td>
</tr>
<tr>
<td>Thru ORDCU-SE</td>
<td></td>
</tr>
<tr>
<td>Office of the Chief Signal Officer</td>
<td></td>
</tr>
<tr>
<td>Engineering and Technical Division</td>
<td></td>
</tr>
<tr>
<td>Engineering Control Branch</td>
<td></td>
</tr>
<tr>
<td>Room 2B273, Pentagon Hdg.</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td></td>
</tr>
<tr>
<td>Attn: SICOD</td>
<td>1</td>
</tr>
<tr>
<td>MAC for Aeronautics</td>
<td></td>
</tr>
<tr>
<td>1726 F. Street, N. W.</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td></td>
</tr>
<tr>
<td>Attn: Mr. E. B. Jackson</td>
<td></td>
</tr>
<tr>
<td>Chief, Office of Aeronautical Intelligence</td>
<td>1</td>
</tr>
<tr>
<td>Agency</td>
<td>No. of Technical Copies</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Deputy Chief of Staff for Logistics&lt;br&gt;Department of the Army&lt;br&gt;Washington 25, D. C.&lt;br&gt;Attn: Research Br., R and D Division</td>
<td>1</td>
</tr>
<tr>
<td>Commanding General&lt;br&gt;White Sands Proving Ground&lt;br&gt;Las Cruces, New Mexico&lt;br&gt;Attn: OMDGS-TS-TIB</td>
<td>1</td>
</tr>
<tr>
<td>Chief of Ordnance&lt;br&gt;Department of the Army&lt;br&gt;Washington 28, D. C.&lt;br&gt;Attn: Mr. C. A. Bickel, Design of&lt;br&gt;Experimental Unit&lt;br&gt;R. and D. Division</td>
<td>2</td>
</tr>
<tr>
<td>Commanding Officer&lt;br&gt;Detroit Arsenal&lt;br&gt;Center Line, Michigan&lt;br&gt;Attn: J. C. Brooks</td>
<td>1</td>
</tr>
<tr>
<td>Commanding General&lt;br&gt;Ordnance Weapon Command&lt;br&gt;Rock Island, Illinois&lt;br&gt;Attn: Research Branch</td>
<td>2</td>
</tr>
<tr>
<td>Commanding Officer&lt;br&gt;U. S. Naval Powder Factory&lt;br&gt;Indian Head, Maryland</td>
<td>1</td>
</tr>
</tbody>
</table>