Measurement Uncertainty Evaluation for Least Squares Best-Fit Lines

Ricardo A. Nicholas
The Boeing Company
Primary Standards Engineering
Seattle, WA 98124
P.O. Box 3707, M/C 2T-40
Phone: 206 544-0569, Fax: 206 544-5907
E-mail: ricardo.a.nicholas@boeing.com

Abstract: Least squares best-fit Line calculations have become a routine measurement tool. Unfortunately, although of great importance, the evaluation of uncertainty consistent with the Guide to the Expression of Uncertainty in Measurement (*Guide*) for parameters of the line, has not been routine. All too often, the estimation of uncertainty has simply involved the product of a coverage factor and the calculated standard deviation, with little consideration for the level of confidence. This paper presents simple formulas in worksheet format that enables the calculation of uncertainty for best-fit line parameters. Covered are uncertainties for the slope of the line, the data points of the line, the interpolated points, and the line as a whole. The uncertainty calculations are consistent with the (*Guide*), in that they have both an interval of confidence and level of confidence as in traditional Type A measurement uncertainty evaluations. In addition, examples are provided which are suitable for use as models.

1 Preface

This paper is based on the guidance provided by the uncopyrighted, public domain, National Institute of Standards and Technology Handbook 91, "Experimental Statistics", by Mary Gibbons Natrella, originally published in 1963 and reprinted in 1966, with corrections. What follows is a compilation of verbatim, adapted quotations or paraphrases of Handbook 91, and original material provided by this paper's author. The arrangement and type of material in Handbook 91 is of practical value to evaluator's of measurement uncertainty. Sections from the Handbook have therefore been left essentially intact except for certain editorial modifications required for NCSLI publication.

2 Introduction

In many situations it is desirable to know something about the relationships between two characteristics of a material, product, or process. In some cases, it may be known from theoretical considerations that two properties are functionally related, and the problem is to find out more about the structure of this relationship. In other cases, there is interest in investigating whether there exists a degree of association between two properties, which could be used to advantage. For example, in specifying methods of test for a material, there may be two tests available, both of which reflect performance, but one of which is cheaper, simpler, or quicker to run. If a high degree of association exists between the two tests, we might wish to run regularly only the simpler test.

In this paper, we deal only with linear relationships. It is worth noting that many nonlinear relationships may be expressed in linear form by a suitable transformation (change of variable). For example, if the relationship is of the form $Y = aX^b$, then $\log Y = \log a + b \log X$. Putting $Y_T = \log Y$, $b_0 = \log a$, $b_1 = b$, $X_T = \log X$, we have the linear expression $X_T = b_0 + b_1 X_T$ in terms of the new (transformed) variables X_T and Y_T. A number of common linearizing transformations are summarized in Table 1-4 and are discussed in Paragraph 1-4.4.

3 Plotting the Data

Where only two characteristics are involved, the natural first step in handling the experimental results is to plot the points on graph paper. Conventionally, the *independent variable X* is plotted on the horizontal scale and the *dependent variable Y* is plotted on the vertical scale. There is no substitute for a plot of the data to give some idea of the general spread and shape of the results. A pictorial indication of the probable form and sharpness of the relationship, if any, is indispensable and sometimes may save needless computing. When investigating a structural relationship, the plotted data will show whether a hypothetical linear relationship is borne out; if not, we must consider whether there is any theoretical basis for fitting a curve of higher degree. When looking for an empirical association of two characteristics, a glance at the plot will reveal whether such association is likely or whether there is only a patternless scatter of points.

In some cases, a plot will reveal unsuspected difficulties in the experimental setup which must be ironed out before fitting any kind of relationship. An example of this occurred in measuring the time required for a drop of dye to travel between marked distances along a water channel. The channel was marked with distance markers spaced at equal distances, and an observer recorded the time at which the dye passed each marker. The device used for re cording time consisted of two clocks hooked up so that when one was stopped, the other started: Clock 1 recorded the times for Distance Markers 1, 3, 5, etc.; and Clock 2 recorded times for the even-numbered distance markers. When the elapsed times were plotted, they looked somewhat as shown in Figure 1. It is obvious that there was a systematic time difference between odd and even markers (presumably a lag in the circuit connecting the two clocks). One could easily have fitted a straight line to the odd-numbered distances and a different line to the even-numbered distances, with approximately constant difference between the two lines. The effect was so consistent, however, that the experimenter quite properly decided to find a better means of recording travel times before fitting any line at all.

If no obvious difficulties are revealed by the plot, and the relationship appears to be linear, then a line $Y = b_o + b_1 X$ ordinarily should be fitted to the data, according to the procedures given in this course. Fitting by eye usually is inadequate for the following reasons:

(a) No two people would fit exactly the same line, and, therefore, the procedure is not objective;

(b) We always need some measure of how well the line does fit the data, and of the uncertainties inherent in the fitted line as a representation of the true underlying relationship — and these can be obtained only when a formal, well-defined mathematical procedure of fitting is employed.

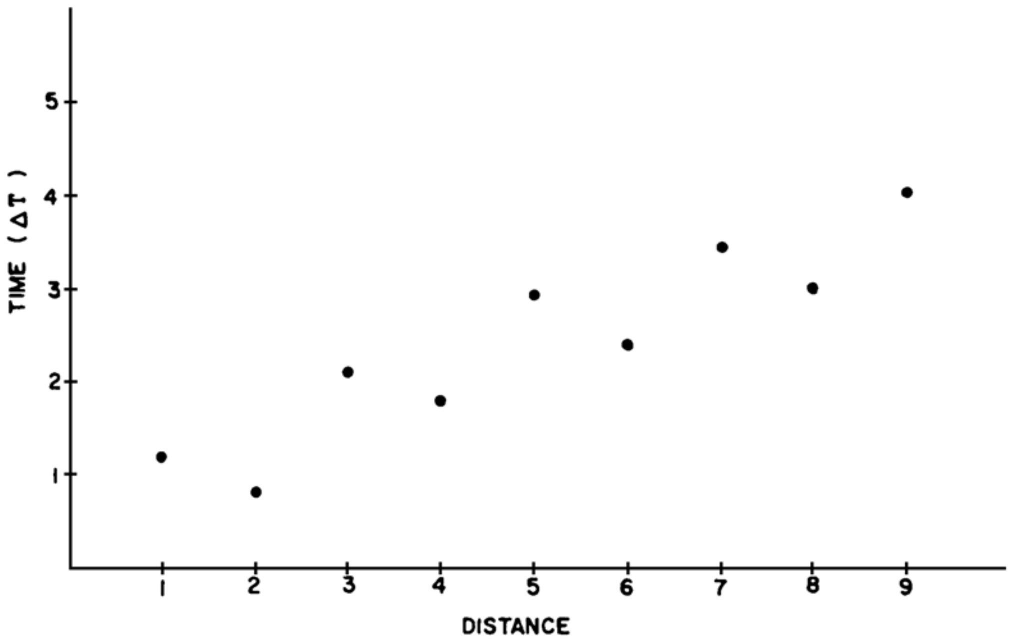

Figure 1. Time required for a drop of dye to travel between distance markers.

4 Two Important Systems of Linear Relationships

Before giving the detailed procedure for fitting a straight line, we discuss different physical situations, which can be described by a linear relationship between two variables. The methods of description and prediction may be different, depending upon the underlying system. In general, we recognize two different and important systems, which we call *Statistical* and *Functional*. It is not possible to decide which is the appropriate system from looking at the data. The distinction must be made before fitting the line— indeed, before taking the measurements.

4.1 Functional Relationships

In the case of a Functional Relationship, there exists an exact mathematical formula (Y as a function of x) relating the two variables, and the only reason that the observations do not fit this equation exactly is because of disturbances or errors of measurement in the observed values of one or both variables. We discuss two cases of this type:

FI—Errors of measurement affect only one variable (Y). (See Fig. 2).
FII—Both variables (X and Y) are subject to errors of measurement. (See Fig. 3).

Common situations that may be described by Functional Relationships include calibration lines, comparisons of analytical procedures, and relationships in which time is the *X* variable.

For instance, we may regard Figure 2 as portraying the calibration of a straight-faced spring balance in terms of a series of weights whose masses are accurately known.
By Hooke's Law, the extension of the spring, and hence the position y of the scale pointer, should be determined exactly by the mass x upon the pan through a linear functional relationship[1] $y = b_0 + b_1 x$. In practice, however, if a weight of mass x_1 is placed upon the pan repeatedly and the position of the pointer is read in each instance, it usually is found that the readings Y_1 are not identical, due to variations in the performance of the spring and to reading errors. Thus, corresponding to the mass x_1 there is a distribution of pointer readings Y_1; corresponding to mass x_2, a distribution of pointer readings Y_2; and so forth—as indicated in Figure 2. It is customary to assume that these distributions are normal (or, at least symmetrical and all of the same form) and that the mean of the distribution of Y_i's coincides with the true value $y = b_0 + b_1 x_i$.

If, instead of calibrating the spring balance in terms of a series of accurately known weights, we were to calibrate it in terms of another spring balance by recording the corresponding pointer positions when a series of weights are placed first on the pan of one balance and then on the pan of the other, the resulting readings (*X* and *Y*) would be related by a linear structural relationship FII, as shown in Figure 3, inasmuch as both *X* and *Y* are affected by errors of measurement. In this case, corresponding to the repeated weighings of a single weight w_1 (whose true mass need not be known), there is a joint distribution of the pointer readings (X_1 and Y_1) on the two balances, represented by the little transparent *mountain* centered over the true point (x_1, y_1) in Figure 3; similarly at points (x_2, y_2) and (x_3, y_3), corresponding to repeated weighings of other weights w_2 and w_3, respectively.

Finally, it should be noticed that this FII model is more general than the FI model in that it does not require linearity of response of each instrument to the independent variable w, but merely that the response curves of the two instruments be linearly related, that is, that $X = a + b * f(w)$ and $Y = c + d * f(w)$, where f(w) may be linear, quadratic, exponential, logarithmic, or whatever.

Table 1 provides a concise characterization of FI and FII relationships.

Detailed problems and procedures with numerical examples for FI relationships are given in Paragraphs 4.1 and 4.2, and for FII relationships in Paragraph 4.3.

---

[1] *Note on Notation for Functional Relationships:*
We have used x and y to denote the true or accurately known values of the variables and *X* and *Y* to denote their values measured with error. In the FI Relationship, the independent variable is always without error and therefore in our *discussions* of the FI case and in the paragraph headings we always use x. In the Worksheet, and Procedures and Examples for the FI case, however we use *X* and *Y* because of the computational similarity to other cases discussed in this Chapter (i.e., the computations for the Statistical Relationships). In the FII case, both variables are subject to error, and clearly we use X and Y everywhere for the observed values.

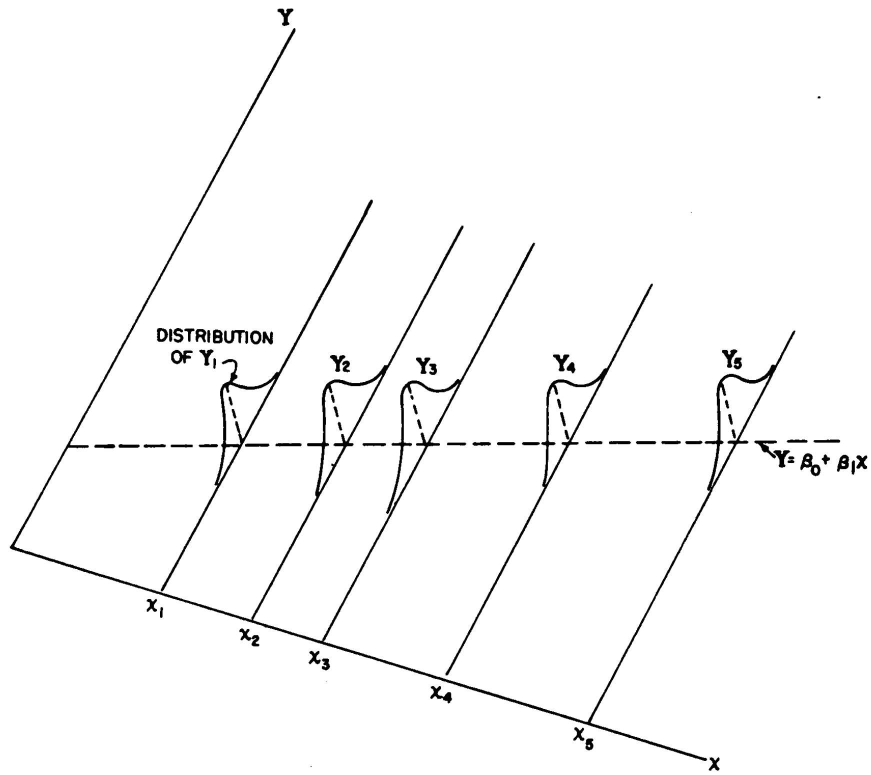

Figure 2. Linear functional relationship of Type FI (only Y affected by measurement errors).

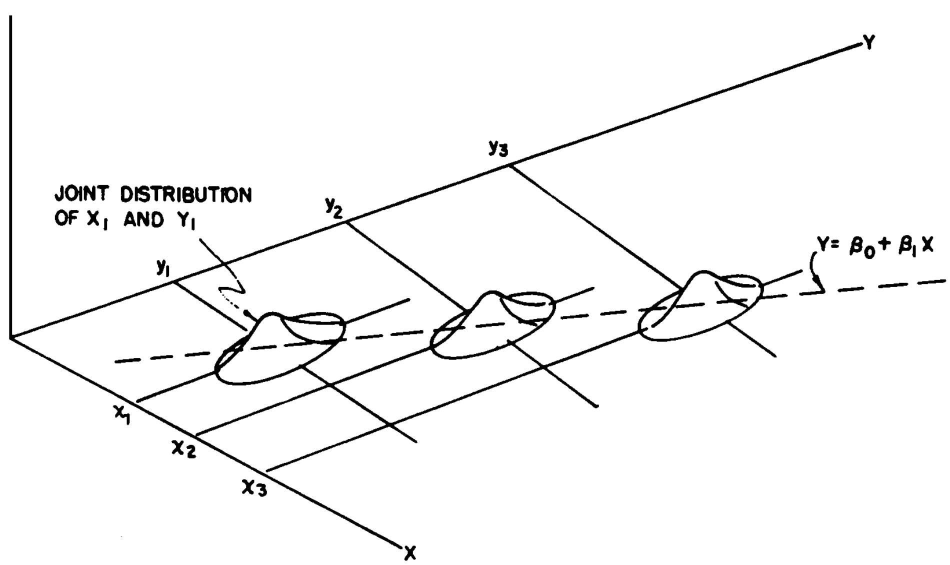

Figure 3. Linear functional relationship of Type FII (X and Y affected by measurement errors).

4.2 Statistical Relationships

In the case of a Statistical Relationship, there is no exact mathematical relationship between X and Y; there is only a statistical association between the two variables as characteristics of individual items from some particular population.

If this statistical association is of bivariate normal type as shown in Figure 4, then the *average* value of the Y's associated with a particular value of X, say $\bar{Y}_X$, is found to depend linearly on X, i.e., $\bar{Y}_x = b_0 + b_1 X$; similarly, the *average* value of the X's associated with a particular value of Y, say, $\bar{X}_Y$ depends linearly on Y (Fig. 4) i.e., $\bar{X}_Y = \beta_0' + \beta_1' Y$; but — and this is important! — the two lines are *not* the same, i.e., $\beta_1' \neq 1/\beta_1$ and $\beta_0' \neq -\beta_0/\beta_1$.[2]

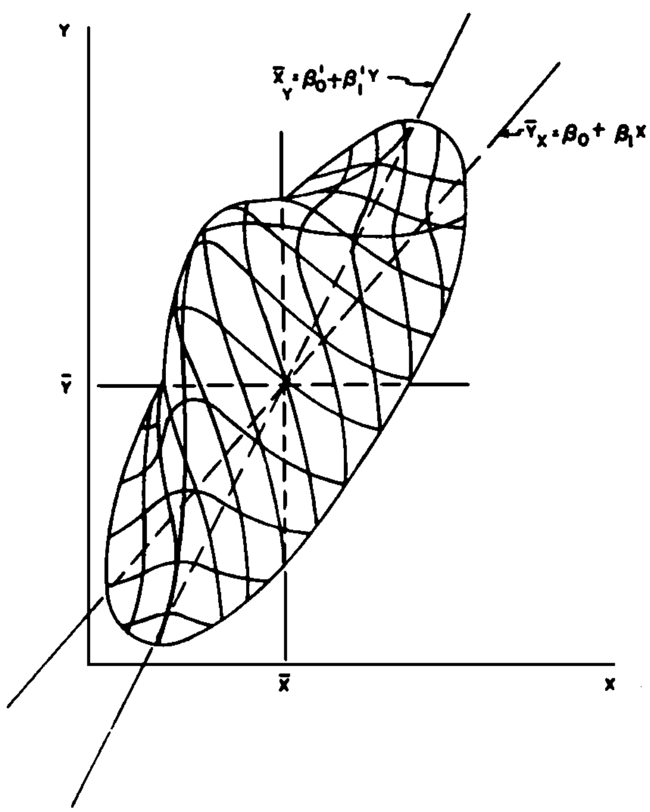

Figure 4. A normal bivariate frequency surface.

---

[2] Strictly, we should write $m_{Y*X} = \beta_0 + \beta_1 X$, and $m_{X*Y} = \beta_0' + \beta_1' Y$, to conform to our notation of using m to signify a population mean, but this more exact notation tends to conceal the parallelism of the curve-fitting processes in the FI and SI situations. Consequently, to preserve appearances here and in the sequel, we use $\bar{Y}_X$ in place of m_{Y*X} and $\bar{X}_Y$ in place of m_{X*Y} — and it should be remembered that these signify *population means*.

If a random sample of items is drawn from the population, and the two characteristics X and Y are measured on each item, then typically it is found that errors of measurement are negligible in comparison with the variation of each characteristic over the individual items. This general case is designated SI. A special case (involving pre-selection or restriction of the range of one of the variables) is denoted by SII.

SI Relationships.

In this case, a random sample of items is drawn from some definite population (material, product, process, or people), and two characteristics are measured on each item.
A classic example of this type is the relationship between height and weight of men. Any observant person knows that weight tends to vary with height, but also that individuals of the same height might vary widely in weight. It is obvious that the errors made in measuring height or weight are very small compared to this inherent variation between individuals. We surely would not expect to predict the exact weight of one individual from his height, but we might expect to be able to estimate the average weight of all individuals of a given height.

The height-weight example is given as one that is universally familiar. Such examples also exist in the physical and engineering sciences, particularly in cases involving the interrelation of two test methods. In many cases there may be two tests that, strictly speaking, measure two basically different properties of a material, product, or process, but these properties are statistically related to each other in some complicated way and both are related to some performance characteristic of particular interest, one usually more directly than the other. Their interrelationship may be obscured by inherent variations among sample units (due to varying density, for example). We would be very interested in knowing whether the relationship between the two is sufficient to enable us to predict with reasonable accuracy, from a value given by one test, the average value to be expected for the other —particularly if one test is considerably simpler or cheaper than the other.

The choice of which variable to call X and which variable to call Y is arbitrary—actually there are two regression lines. If a statistical association is found, ordinarily the variable that is easier to measure is called X. Note well that this is the only case of linear relationship in which it may be appropriate to fit two different lines, one for predicting Y from X and a different one for predicting X from Y, *and* the only case in which the sample correlation coefficient r is meaningful as an estimate of the degree of association of X and Y *in the population* as measured by the population coefficient of correlation $\rho = \sqrt{\beta_1 \beta_1^{'}}$. The six sets of contour ellipses shown in Figure 5 indicate the manner in which the location, shape, and orientation of the normal bivariate distribution varies with changes of the population means (m_X and m_Y) and standard deviations (σ_X and σ_Y) of X and Y and their coefficient of correlation in the population (ρ_{XY}).

If $\rho = \pm 1$, all the points lie on a line and $Y = \beta_0^{'} + \beta_1^{'} X$, and $X = \beta_0^{'} + \beta_1^{'} Y$ coincide.

If $\rho = +1$, the slope is positive, and if $\rho = -1$, the slope is negative. If $\rho = 0$, then X and Y are said to be uncorrelated.

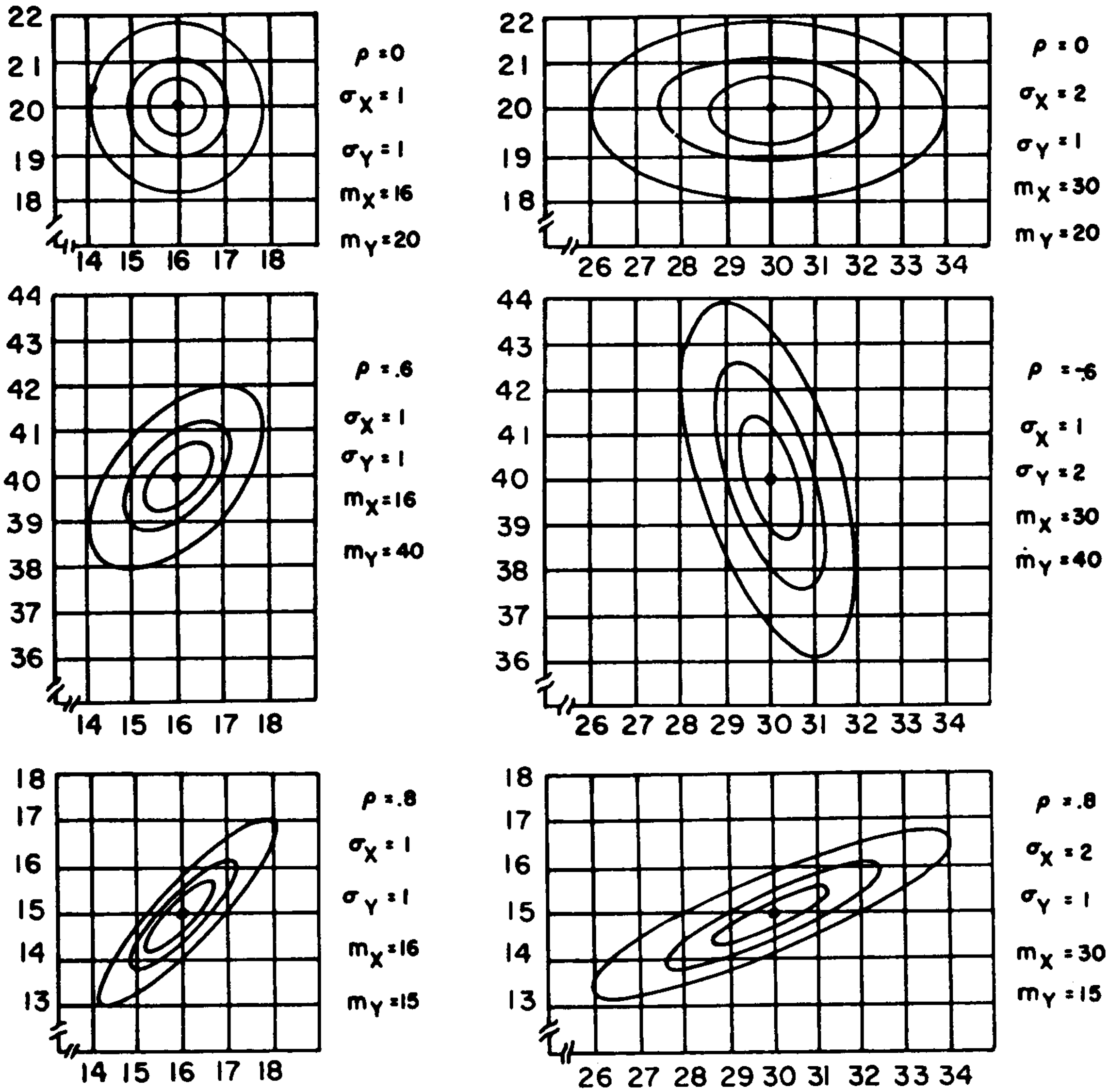

Figure 5. Contour ellipses for normal bivariate distributions having different values of the five parameters m_x, m_y, o_x, a_y, P_{xy}.

SII Relationships

The general case described above (SI) is the most familiar example of a statistical relationship, but we also need to consider a common case of Statistical Relationship (SII) that must be treated a bit differently. In SII, one of the two variables, although a random variable in the population, is sampled only within a limited range (or at selected pre-assigned values). In the height-weight example, suppose that the group of men included only those whose heights were between 5'4" and 5'8". We now are able to fit a line predicting weight from height, but are unable to determine the correct line for predicting height from weight. A correlation coefficient computed from such data is not a measure of the true correlation among height and weight in the (unrestricted) population.

The restriction of the range of *X*, when it is considered as the independent variable, does not spoil the estimates of $\bar{Y}_x$ when we fit the line $\bar{Y}_x = b_0 + b_1 X$;. The restriction of the range of the dependent variable (i.e., of *Y* in fitting the foregoing line, or of *X* in fitting the line,
$\bar{X}_Y = b_0' + b_1' Y$), however, gives a seriously distorted estimate of the true relationship.

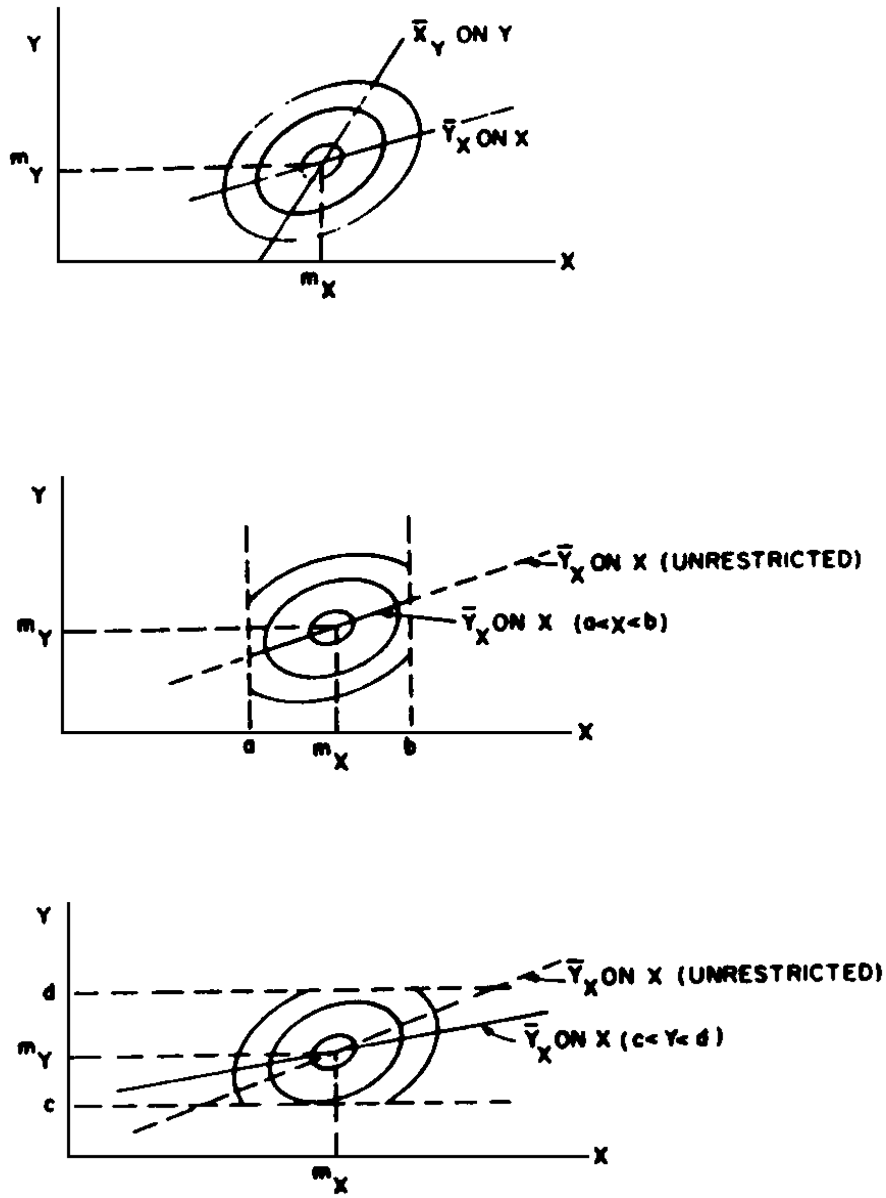

Figure 6. Diagram showing effect of restrictions of X or Y on the regression of Y on X.

This is evident from Figure 6, in which the contour ellipses of the top diagram serve to represent the bivariate distribution of X and Y in the unrestricted population, and the "true" regression lines of $\overline{Y}_X$ on X and $\overline{X}_Y$ on Y are indicated. The central diagram portrays the situation when consideration is restricted to items in the population for which $a < X < b$. It is clear that for any particular X in this interval, the distribution and hence the mean $\overline{Y}'_X$ of the corresponding Y's is the same as in the unrestricted case (top diagram). Consequently, a line of the form $\overline{Y}_X = b_0 + b_1X$ fitted to data involving either a random or selected set of values of X between $X = a$ and $X = b$, but with *no* selection or restrictions on the corresponding Y's, will furnish an unbiased estimate of the *true* regression line $\overline{Y}_X = b_0 + b_1X$ in the population at large.
In contrast, if consideration is restricted to items for which $c < Y < d$, as indicated in the bottom diagram, then it is clear that the mean value, say $\overline{Y}'_X$, of the (restricted) Y's associated with any particular value of $X > m_X$ will be less than the corresponding mean value Y_X in the population as a whole. Likewise, if $X < m_X$, then the mean $\overline{Y}'_X$ of the corresponding (restricted) Y's will be greater than $\overline{Y}'_X$ in the population as a whole. Consequently, a line of the form $\overline{Y}'_X = b_0 + b_1X$ fitted to data involving selection or restriction of Y's will not furnish an unbiased estimate of the true regression line $\overline{Y}'_X = b_0 + b_1X$ in the population as a whole, and the distortion may be serious. In other words, introducing a restriction with regard to X does not bias inferences wit h regard to Y, when Y is considered as the dependent variable, but restricting Y will distort the dependence of $\overline{Y}'_X$ on X so that the relationship observed will not be representative of the true underlying relationship in the population as a whole. Obviously, there is an equivalent statement in which the roles of X and Y are reversed.

As an engineering example of SII, consider a study of watches to investigate whether there was a relationship between the cost of a stopwatch and its temperature coefficient. It was suggested that a correlation coefficient be computed. This was not possible because the watches had not been selected at random from the total watch production, but a deliberate effort had been made to obtain a fixed number of low-priced, medium-priced, and high-priced stopwatches.

In any given case, consider carefully whether one is measuring samples as they come (and thereby accepting the values of both properties that come with the sample) which is an SI Relationship, or whether one selects samples which are known to have a limited range of values of X problems and procedures with numerical exam (which is an SII Relationship).

Table 5 gives a brief summary characterization of SI and SII Relationships. Detailed problems and procedures with numerical examples are given for SI relationships in Paragraph 5-5.1 and for SII relationships in Paragraph 5-5.2 of Handbook 91.

Table 1. Summary of Four Cases of Linear Relationships

	Functional (F)		Statistical (S)	
	FI	FII	SI	SII
Distinctive Features and Example	x and y are linearly related by a mathematical formula, $y = \beta_0 + \beta_1 x$, or $x = \beta_1' y$, which is not observed exactly because of disturbances or errors in one or both variables. Example: Determination of elastic constant of a spring which obeys Hooke's law: x = accurately known weight applied, Y = measured value of corresponding elongation y.		X = Height Y = Weight Both measured on a random sample of individuals. X is not selected but "comes with" sample unit.	X = Height (pre-selected value) Y = Weight of individuals of pre-selected height X is measured beforehand; only selected values of X are used at which to measure Y.
Errors of Measurement	Measurement error affects Y only.	X and Y both subject to error.	Ordinarily negligible compared to variation among individuals.	Same as in SI.
Form of Line Fitted	$Y = b_0 + b_1 X$	See paragraph 5-4.3 of Handbook 91.	$\overline{Y}_X = b_0 + b_1 X$ $\overline{X}_Y = b_0' + b_1' Y$	$\overline{Y}_X = b_0 + b_1 X$ only.
Procedure for Fitting	See Paragraphs 5.1, 5.2, and basic worksheet.	Procedure depends on what assumptions can be made. See Paragraph 5-4.3 of Handbook 91.	See Paragraph 5-5.1 and basic worksheet of Handbook 91.	See Paragraph 5-5.2 and basic worksheet of Handbook 91.
Correlation Coefficient	Not applicable	Not applicable	Sample estimate is $\gamma = \dfrac{Sxy}{\sqrt{Sxx}\sqrt{Syy}}$ See Paragraph 5-5.1.5 of Handbook 91.	Correlation may exist in the population, but r computed from such an experiment would provide a distorted estimate of the correlation.

Basic Worksheet for all Types of Linear Relationships

X denotes ________________ Y denotes ________________

$\sum X =$ ________________ $\sum Y =$ ________________

$\overline{X} =$ ________________ $\overline{Y} =$ ________________

Number of points: $n =$ ________________

(2) $\left(\sum X\right)\left(\sum Y\right) =$ ________________

Step (1) $\sum XY =$ ________________

(3) $S_{XY} =$ Step (1) – Step (2)

(4) $\sum X^2 =$ ________________

(5) $\dfrac{\left(\sum X\right)^2}{n} =$ ________________

(6) $S_{XX} =$ Step (4) – Step (5)

(7) $\sum Y^2 =$ ________________

(8) $\dfrac{\left(\sum Y\right)^2}{n} =$ ________________

(9) $S_{YY} =$ Step (7) – Step (8)

(10) $b_1 = \dfrac{S_{XY}}{S_{XX}} =$ Step (3) ÷ Step (4)

(11) $\overline{Y} =$ ________________

(12) $b_1\overline{X} =$ ________________

(13) $b_0 = \overline{Y} - b_1\overline{X} =$ Step (11) – Step (12)

(14) $\dfrac{\left(S_{XY}\right)^2}{S_{XX}} =$ ________________

(15) $(n-2)S_Y^2 =$ Step (9) – Step (14)

(16) $S_Y^2 =$ Step (15) ÷ $(n-2)$

$S_Y =$ ________________

Equation of the line: $Y = b_0 + b_1\overline{X}$

____

Estimated variance of the slope:

$S_{b_1}^2 = \dfrac{S_Y^2}{S_{XX}} =$ Step (16) ÷ Step (6)

$S_{b_0} =$ ________________

$S_{b_1} =$ ________________

Estimated variance of intercept:

$S_{b_0}^2 = S_Y^2\left[\dfrac{1}{n} + \dfrac{\overline{Y}^2}{S_{XX}}\right] =$ ________________

Note: The following are algebraically identical:

$S_{XX} = \sum\left(X-\overline{X}\right)^2$; $S_{YY} = \sum\left(Y-\overline{Y}\right)^2$; $S_{XY} = \sum\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)$

Ordinarily, in hand computation, it is preferable to compute as shown in the steps above.
Carry all decimal places obtainable—i.e., if data are recorded to two decimal places, carry four places in Steps (1) through (9) in order to avoid losing significant figures in subtraction.

5. Problems and Procedures for Functional Relationships

5.1 FI Relationships (General Case)

There is an underlying mathematical (functional) relationship between the two variables, of the form $y = \beta_0 + \beta_1 x$. The variable x can be measured relatively accurately. Measurements Y of the value of y corresponding to a given x follow a normal distribution with mean $\beta_0 + \beta_1 x$ and variance $\sigma_{y_x}^2$ which is independent of the value of x. Furthermore, we shall assume that the deviations or errors of a series of observed Y's, corresponding to the same or different x's, all are mutually independent. See Paragraph 4.1 and Table 1. The general case is discussed here, and the special case where it is known that $\beta_0 = 0$ (i.e., a line known to pass through the origin) is discussed in Paragraph 5-4.2 of Handbook 91. The procedure discussed here also will be valid if in fact $\beta_0 = 0$ even though this fact is not known beforehand. However, when it is known that $\beta_0 = 0$, the procedures of Paragraph 5-4.2 of Handbook 91 should be followed because they are simpler and somewhat more efficient. It will be noted that SII, Paragraph 5-5.2 of Handbook 91, is handled computationally in exactly the same manner as FI, but both the underlying assumptions and the interpretation of the end results are different.

Data Sample 5.1—Young's Modulus vs. Temperature for Sapphire Rods

Observed values (Y) of Young's modulus (y) for sapphire rods measured at different temperatures (x) are given in the following table. There is assumed to be a linear functional relationship between the two variables x and y. For computation purposes, the observed Y values were coded by subtracting 4000 from each. To express the line in terms of the original units, add 4000 to the computed intercept; the slope will not be affected. The observed data are plotted in Figure 7.

x = Temperature °C	Y = Young's Modulus	Coded Y = Young's Modulus Minus 4000
30	4642	642
100	4612	612
200	4565	565
300	4513	513
400	4476	476
500	4433	433
600	4389	389
700	4347	347
800	4303	303
900	4251	251
1000	4201	201
1100	4140	140
1200	4100	100
1300	4073	73
1400	4024	24
1500	3999	-1

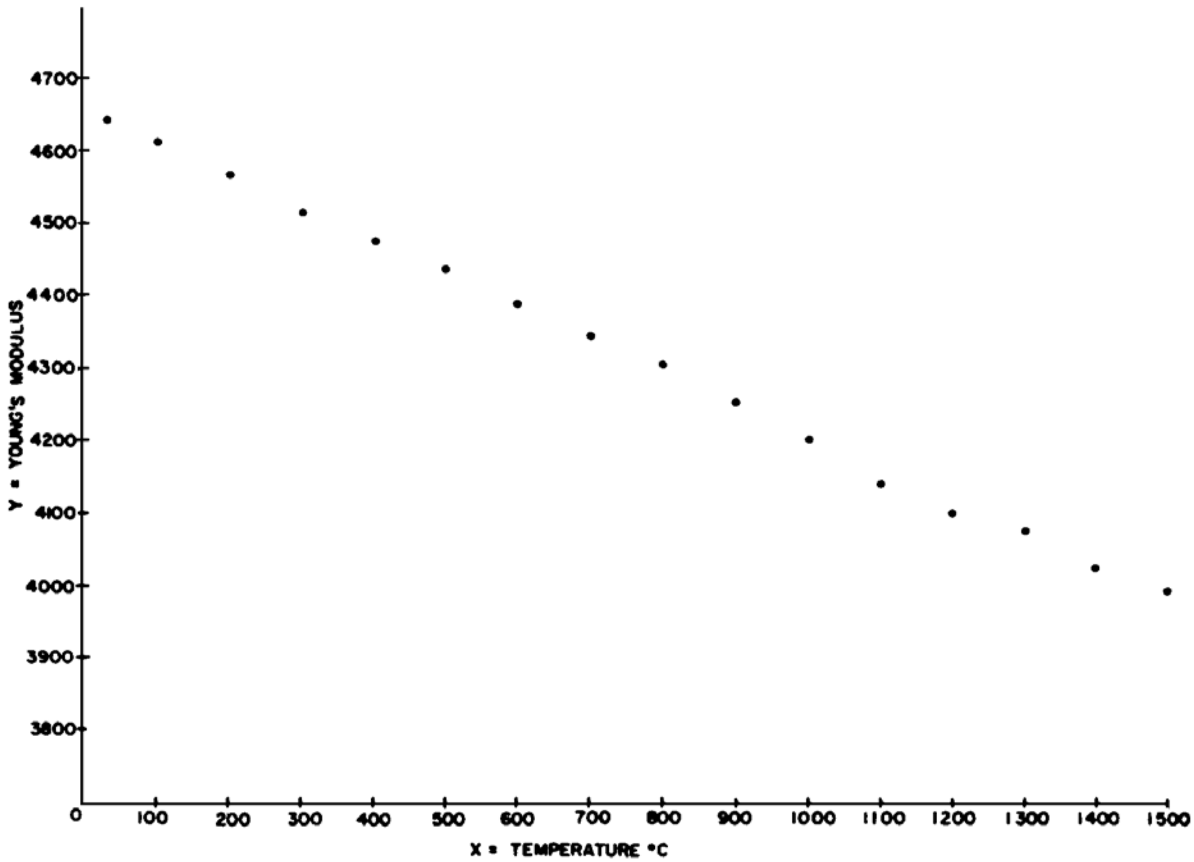

Figure 7. Young's modulus of sapphire rods as a function of temperature—an FI relationship

5.1.1 What is the Best Line to be Used for Estimating *Y* From Given Values of *X*?

Procedure[3]

Using Worksheet (See worksheet 5.1), compute the line $y = b_0 + b_1 x$. This is an estimate of the true equation $y = \beta_0 + \beta_1 x$. The method of fitting a line given here is a particular application of the general method of least squares. From Data Sample 5.1, the equation of the fitted line (in original units) is: $Y = 4654.9846 - 0.44985482x$. The equation in original units is obtained by adding 4000 to the computed intercept b_0. Since the *Y*'s were coded by subtracting a constant, the computed slope b_1 was not affected. In Figure 5-8, the line is drawn and confidence limits for the line (computed as described in Paragraph 5.1.2.1) also is shown.

[3] CAUTION: Extrapolation, i.e., use of the equation of the line for prediction outside the range of data from which the line was computed, may lead to highly erroneous conclusions.

Worksheet 5.1
Example of FI Relationship
Young's Modulus as Function of Temperature

X denotes Temperature, °C

Y denotes Young's Modulus – 4,000

$\sum X =$ 12,030

$\sum Y =$ 5,068

$\overline{X} =$ 751.875

$\overline{Y} =$ 316.75

Number of points: $n =$ 16

(2) $\left(\sum X\right)\left(\sum Y\right) =$ 3,810,502.5

Step (1) $\sum XY =$ 2,300,860

(3) $S_{XY} =$ Step (1) – Step (2) = –1,509,642.5

(4) $\sum X^2 =$ 12,400,900

(7) $\sum Y^2 =$ 2,285,614

(5) $\dfrac{\left(\sum X\right)^2}{n} =$ 9,045,056.25

(8) $\dfrac{\left(\sum Y\right)^2}{n} =$ 1,605,289

(6) $S_{XX} =$ Step (4) – Step (5) = 3,355,843.75

(9) $S_{YY} =$ Step (7) – Step (8) = 680,325

(10) $b_1 = \dfrac{S_{XY}}{S_{XX}} =$ Step (3) ÷ Step (4) = –0.44985482

(14) $\dfrac{\left(S_{XY}\right)^2}{S_{XX}} =$ 679l,119.9614

(11) $\overline{Y} =$ 316.75

(15) $(n-2)S_Y^2 =$ Step (9) – Step (14) = 1,205.0386

(12) $b_1\overline{X} =$ –338.2345959

(16) $S_Y^2 =$ Step (15) ÷ $(n-2)$ = 86.07418908

(13) $b_0 = \overline{Y} - b_1\overline{X} =$ Step (11) – Step (12) = 654.9846

$S_Y =$ 9.277618

b_0 (in original units) = 4654.9846

Equation of the line: $Y = b_0 + b_1 X$

= 4654.9846 – 0.44985482 x

Estimated variance of the slope:

$S_{b_1}^2 = \dfrac{S_Y^2}{S_{XX}} =$ Step (16) ÷ Step (6) = 0.000025649046

$S_{b_1} =$ 0.005064

Estimated variance of intercept:

$S_{b_0}^2 = S_Y^2\left[\dfrac{1}{n} + \dfrac{\overline{X}^2}{S_{XX}}\right] =$ 19.879453

$S_{b_0} =$ 4.458638

Note: The following are algebraically identical:

$S_{XX} = \sum\left(X - \overline{X}\right)^2$; $S_{YY} = \sum\left(Y - \overline{Y}\right)^2$; $S_{XY} = \sum\left(X - \overline{X}\right)\left(Y - \overline{Y}\right)$

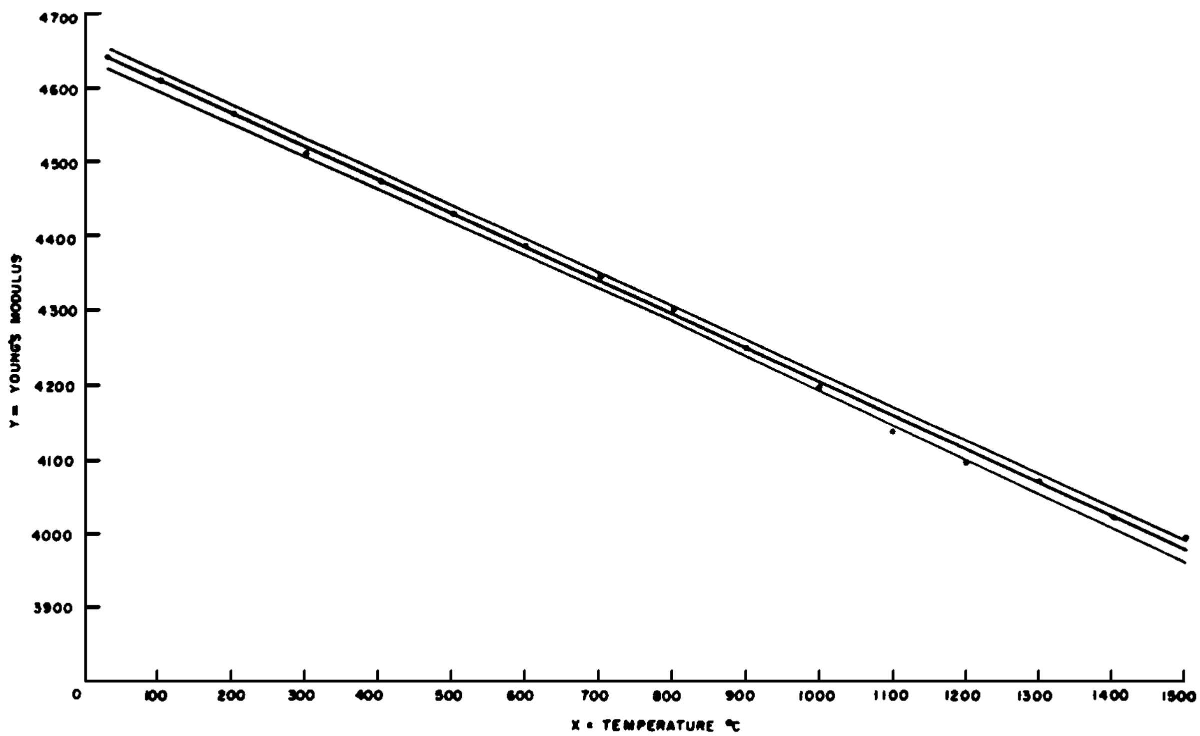

Figure 8. Young's modulus of sapphire rods as a function of temperature—showing computed regression line and confidence interval for the line.

Using the Regression Equation for Prediction.

The fitted regression equation may be used for two kinds of predictions:

(a) To estimate the *true* value of the y associated with a particular value of x, e.g., given $x = x'$ to estimate the value of $y' = \beta_0 + \beta_1 x'$; or,

(b) To predict a single new observed value Y corresponding to a particular value of x, e.g., given $x = x'$ to predict the value of a single measurement of y'.

Which prediction should be made? In some cases, it is sufficient to say that the *true* value of y (for given x) lies in a certain interval, and in other cases we may need to know how large (or how small) an individual observed Y value is likely to be associated with a particular value of x. The question of what to predict is similar to the question of what to specify (e.g., whether to specify average tensile strength or to specify minimum tensile strength) and can be answered only with respect to a particular situation. The difference is that here we are concerned with relationships between two variables and therefore must always talk about the value of y, or Y, for fixed x.

The predicted y' or Y' value is obtained by substituting the chosen value (x') of x in the fitted equation. For a particular value of x, either type of prediction ((a) or (b)) gives the same numerical answer for y' or Y'. The uncertainty associated with the prediction, however, does depend on whether we are estimating the *true* value of y', or predicting the value Y' of an individual measurement of y'. If the experiment could be repeated many times, each time obtaining n pairs of (x, Y) values, consider the range of Y values, which would be obtained for a given x. Surely the individual Y values in all the sets will spread over a larger range than will the collection consisting of the average Y's (one from each set).

To estimate the *true* value of y associated with the value x', use the equation

$$y_c' = b_0 + b_1 x'$$

The variance of y_c' as an estimate of the *true* value $y' = \beta_0 + \beta_1 x'$ is

$$\text{Var } y_c' = s_Y^2 + \left[\frac{1}{n} + \frac{\left(X' - \overline{X}\right)^2}{S_{xx}} \right]$$

This variance is the variance of estimate of a point on the fitted line.

For example, using the equation relating Young's modulus to temperature, we predict a value for y at $x = 1200$:

$$y_c' = 4654.9846 - 0.44985482(1200)$$
$$y_c' = 4115.16$$

$$\text{Var } y_c' = 86.074 + \left[\frac{1}{16} + \frac{(1200 - 751.875)^2}{3355843.75} \right]$$
$$= 86.074\ (0.0625 + 0.0598)$$
$$= 86.074\ (0.1223)$$

$$\text{Var } y_c' = 10.53$$

To predict a single observed value of Y corresponding to a given value (x') of x, use the same equation

$$Y_c' = b_0 + b_1 x'$$

The variance of Y_c' as an estimate of a single new (additional, future) measurement of y' is

$$\text{Var } y_c' = s_Y^2 + \left[1 + \frac{1}{n} + \frac{\left(X' - \overline{X}\right)^2}{S_{xx}} \right]$$

The equation for our example is

$Y = 4654.9846 - 0.44985482\,x$

To predict the value of a single determination of Young's modulus at $x = 750$, substitute in this equation and obtain:

$$Y_c' = 4654.9846 - 0.44985482(750)$$
$$= 4317.59$$

$$\text{Var}\, y_c' = s_Y^2\left[1 + \frac{1}{n} + \frac{\left(X' - \overline{X}\right)^2}{S_{xx}}\right]$$

$$= 86.074\left[1 + \frac{1}{16} + \frac{(750 - 751.875)^2}{3355843.75}\right]$$

$$= 86.074\ (1.0625)$$
$$= 91.45$$

5.1.2 What are the Confidence Interval Estimates for: the Line as a Whole; a Point on the Line; a Future Value of *Y* Corresponding to a Given Value of *X*?

Once we have fitted the line, we want to make predictions from it, and we want to know how good our predictions are. Often, these predictions will be given in the form of an interval together with a confidence coefficient associated with the interval—i.e., confidence interval estimates. Several kinds of confidence interval estimates may be made:

(a) A confidence band for the line as a whole.

(b) A confidence interval for a point on the line—i.e., a confidence interval for y' (the *true* value of y and the *mean* value of Y) corresponding to a single value of $x = x'$.

If the fitted line is, say, a calibration line that will be used over and over again, we will want to make the interval estimate described in (a). In other cases, the line as such may not be so important. The line may have been fitted only to investigate or check the structure of the relationship, and the interest of the experimenter may be centered at one or two values of the variables.

Another kind of interval estimate sometimes is required:

(c) A single observed value (Y') of Y corresponding to a new value of $x = x'$.

These three kinds of confidence interval statements have somewhat different interpretations. The confidence interval for (b) is interpreted as follows:

Suppose that we repeated our experiment a large number of times. Each time, we obtain n pairs of values (x_i, Y_I), fit the line, and compute a confidence interval estimate for $y = \beta_0 + \beta_1 x'$, the value of y corresponding to the particular value $x = x'$. Such interval estimates of y' are expected to be correct (i.e., include the *true* value of y') a proportion $(1-\alpha)$ of the time. If we were to make an interval estimate of y'' corresponding to another value of $x = x''$, these interval estimates also would be expected to include y'' the same proportion $(1-\alpha)$ of the time. However, taken together, these intervals do not constitute a joint confidence statement about y' and y'' which would be expected to be correct exactly a proportion $(1-\alpha)$ of the time; nor is the effective level of confidence $(1-\alpha)^2$, because the two statements are not independent but are correlated in a manner intimately dependent on the values x' and x'' for which the predictions are to be made.

The confidence band for the whole line (a) implies the same sort of repetition of the experiment except that our confidence statements are not now limited to one x at a time, but we can talk about any number of x values simultaneously—about the whole line. Our confidence statement applies to the line as a whole, and therefore the confidence intervals for y corresponding to all the chosen x values will simultaneously be correct a proportion $(1-\alpha)$ of the time. It will be noted that the intervals in (a) are larger than the intervals in (b) by the ratio $\sqrt{2F}/t$. This wider interval is the "price" we pay for making joint statements about y for any number of, or for all of, the x values, rather than the y for a single x.

Another *caution* is in order. We cannot use the same computed line in (b) and (c) to make a large number of predictions, and claim that $100(1-\alpha)\%$ of the predictions will be correct. The *estimated* line may be very close to the *true line*, in which case nearly all of the interval predictions may be correct; or the line may be considerably different from the *true line*, in which case very few may be correct. In practice, provided our situation is *in control*, we should always revise our estimate of the line to include additional information in the way of new points.

5.1.2.1 What is the $(1-\alpha)$ Confidence Band for the line as a Whole?

Procedure	**Example**
(1) Choose the desired confidence level, $(1-\alpha)$.	(1) Let: $1-\alpha = 0.95$ $\alpha = 0.05$
(2) Obtain s_y from Worksheet 5.1.	(2) $s_y = 9.277617$
(3) Look up $F_{1-\alpha}$ for (2, n-2) degrees of freedom.	(3) $F_{0.95,2,14} = 3.74$
(4) Choose a number of values of X (within the range of the data) at which to compute points for drawing the confidence band.	(4) Let: $X = 30$ $X = 400$ $X = 800$ $X = 1200$ $X = 1500$, for example

(5) At each selected value of X, compute:

$$Y_c = \overline{Y} + b_1\left(X - \overline{X}\right)$$

and

$$W_1 = s_Y\sqrt{2F}\left[\frac{1}{n} + \frac{\left(X' - \overline{X}\right)^2}{S_{xx}}\right]^{0.5}$$

(5) See Table 2 for a convenient computational arrangement and the example calculations.

(6) A $(1-\alpha)$ confidence band for the whole line is determined by $Y_c = \pm W_1$

(6) See Table 2.

(7) To draw the line and its confidence band, plot Y_c at two of the extreme selected values of X. Connect the two points by a straight line. At each selected value of X, also plot Y_c+W_1 and Y_c-W_1. Connect the upper series of points, and the lower series of points, by smooth curves.

(7) See Figure 8.

If more points are needed for drawing the curves for the band, note that, because of symmetry, the calculation of W_1 at n values of X actually gives W_1 at $2n$ values of X.

For example: W_1 (but not Y_c) has the same value at $X = 400$ (i.e., $\overline{X}$ as at $X = 1103.75$ (i.e., $\overline{X}$ + 351.875).

Table 2. Computational Arrangement for Procedure 5.1.2.1

X	$\left(X - \overline{X}\right)$	Y_c	$\frac{1}{n} + \frac{(X-\overline{X})^2}{S_{xx}}$	$s^2_{Y_c}$	s_{Y_c}	W_1	$Y_c + W_1$	$Y_c - W_1$
30	-721.875	4641.49	.21778	18.7452	4.3296	11.84	4653.33	4629.65
400	-351.875	4475.04	.09940	8.5558	2.9250	8.00	4483.04	4467.04
800	48.125	4295.10	.06319	5.4390	2.3322	6.38	4301.48	4288.72
1200	448.125	4115.16	.12234	10.5303	3.2450	8.88	4124.04	4106.28
1500	748.125	3980.20	.22928	19.7351	4.4424	12.15	3992.35	3968.05

$\overline{X}$ = 751.875 coded $\overline{Y}$ = 316.75 $\overline{Y}$ (original units) = 4316.75 $s^2_Y = 86.07418908$

$\frac{1}{n} = 0.0625$ $b_1 = -0.44985482$ S_{xx} = 3,355,843.75 $\sqrt{2F} = 2.735$

$Y_c = \overline{Y} + b_1\left(X - \overline{X}\right)$ $y_c' = s_Y^2 + \left[\frac{1}{n} + \frac{\left(X' - \overline{X}\right)^2}{S_{xx}}\right]$ $W_1 = 2.735\ s_{Y_c}$

5.1.2.2 Give a $(1-\alpha)$ Confidence Interval Estimate for a Single Point on the Line (i.e., the Mean Value of Y Corresponding to a Chosen Value of x = x')

Procedure

(1) Choose the desired confidence level, $(1-\alpha)$.	(1) Let: $1-\alpha=0.95$ $\alpha=0.05$
(2) Obtain s_y from Worksheet 5.1.	(2) $s_y=9.277617$
(3) Look up $t_{1-\alpha/2}$ for (n-2) degrees of freedom.	(3) $t_{0.975,14}=2.145$
(4) Choose X', the value of X at which we want to make an interval estimate of the mean value of Y.	(4) Let: $X'=1200$
(5) Compute $W_2=t_{1-\alpha/2}s_Y\left[\frac{1}{n}+\frac{\left(X'-\overline{X}\right)^2}{S_{xx}}\right]^{0.5}$ and $Y_c=\overline{Y}+b_1\left(X-\overline{X}\right)$	(5) $W_2=2.145(3.2451)$ $=6.96$ $Y_c=4115.16$
(6) A $(1-\alpha)$ confidence interval estimate for the mean value of Y corresponding to $X=X'$ is given by $Y_c=\pm W_1$.	(6) A 95% confidence interval estimate for the mean value of Y corresponding to $X=1200$ is 4115.16 ± 6.96

Note: An interval estimate of the intercept of the line (β_0) is obtained by setting $X'=0$ in the above procedure.

5.1.2.3 Give a $(1-\alpha)$ Confidence Interval Estimate for a Single (Future) Value (*Y'*) of *Y* Corresponding to a Chosen Value of (*x'*) of *x*.

Procedure

(1) Choose the desired confidence level, $(1-\alpha)$.	(1) Let: $1-\alpha=0.95$ $\alpha=0.05$
(2) Obtain s_y from Worksheet 5.1.	(2) $s_y=9.277617$

(3) Look up $t_{1-\alpha/2}$ for (n-2) degrees of freedom.

(3) $t_{0.975,14} = 2.145$

(4) Choose X', the value of X at which we want to make an interval estimate of the mean value of Y.

(4) Let: $X' = 1200$

(5) Compute

$$W_2 = t_{1-\alpha/2} s_Y \left[1 + \frac{1}{n} + \frac{\left(X' - \overline{X}\right)^2}{S_{xx}}\right]^{0.5}$$

and

$$Y_c = \overline{Y} + b_1\left(X - \overline{X}\right)$$

(5)

$$W_2 = 2.145(9.8288)$$
$$= 21.08$$

$$Y_c = 4115.16$$

(6) A $(1-\alpha)$ confidence interval estimate for Y' (the single value of Y corresponding to X' is $Y_c = \pm W_3$.

(6) A 95% confidence interval estimate for a single value of Y corresponding to $X = 1200$ is 4115.16 ± 21.08

5.1.2.4 What is the Confidence Interval Estimate for β_1, the Slope of the True Line $y = \beta_0 + \beta_1 x$?

Procedure

(1) Choose the desired confidence level, $(1-\alpha)$.

(1) Let: $1-\alpha = 0.95$
$\alpha = 0.05$

(2) Look up $t_{1-\alpha/2}$ for (n-2) degrees of freedom.

(2) $t_{0.975,14} = 2.145$

(3) Obtain s_{b_1} from Worksheet 5.1.

(3) $s_{b_1} = 0.005064$

(4) Compute

$$W_4 = t_{1-\alpha/2}\, s_{b_1}$$

(4)

$$W_4 = 2.145(0.005064)$$
$$= 0.010862$$

(5) A $(1-\alpha)$ confidence interval estimate for b_1 is
$b_1 \pm W_4$

(5) $b_1 = 0.449855$
$W_4 = 0.010862$
A 95% confidence interval for β_1 is the interval -0.449855 ± 0.010862

5.2. Additional Considerations of FI Relationships Treated in Handbook 91

Handbook 91 treats additional FI relationships, as well as limited treatment of FII, and a somewhat fuller treatment of SI and SII relationships, which will not be considered here. Sufficient material has been presented to illustrate the salient features of best-fit line confidence interval estimation to suit the limited scope of this paper. The balance of this paper is devoted to relevance of the proceeding FI treatment to measurement uncertainty evaluation.

6. Measurement Uncertainty

6.1 The Relationship of a Best-Fit Line Confidence Interval to its Best-Fit Line Uncertainty

A measurement uncertainty evaluation process involves the determination of an interval and an associated confidence level[4], which can be considered synonymous to a confidence interval. The significance of that for the purpose of this paper is that determining a confidence interval for a best-fit line is equivalent to determining its measurement uncertainty.

As was shown previously, confidence intervals were determined for the best-fit line as a whole, a measured point on the best-fit line, an additional or future value of a point on the best-fit line, and the slope of the best-fit line, in sections 5.1.2.1, 5.1.2.2, 5.1.2.3 and 5.1.2.4, respectively.

It is worth noting that the confidence intervals and therefore the uncertainty bounds are hyperbolic in nature, which are smallest at the center of the best-fit line, and largest at the ends.

The treatment described in this paper is limited to the FI type, which assumes error in Y only. That assumption is not generally true, although is approximately true if X values are accurately determined relative to the Y values. If that is the case, the FI treatment is sound. If not, other more accurate methods must be employed. One complete discussion of a case in which both X and Y have significant errors can be found in F. S. Acton, *Analysis of Straight-Line Data,* John Wiley & Sons, Inc., New York, N.Y., 1959.

7. Summary

A method for the estimation of confidence intervals about best-fit lines was described. Four parameters were investigated including the best-fit line as a whole, a measured point on the best-fit line, an additional or future value of a point on the best-fit line, and the slope of the best-fit line. Worksheets were provided for calculator use or to serve as a basis for a computer application tool. The link between the best-fit line hyperbolic confidence interval and measurement uncertainty was made and was shown to be equivalent. A reference was provided for the evaluation of uncertainty where both X and Y have significant errors.

---

[4] Confidence level is used here for consistency with Handbook 91, contrary to the convention of the *Guide to the Expression of Uncertainty in Measurement*, which uses level of confidence.