NATIONAL BUREAU OF STANDARDS REPORT

9011

NOTES ON THE USE OF PROPAGATION OF ERROR FORMULAS

U.S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

THE NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards is a principal focal point in the Federal Government for assuring maximum application of the physical and engineering sciences to the advancement of technology in industry and commerce. Its responsibilities include development and maintenance of the national standards of measurement, and the provisions of means for making measurements consistent with those standards; determination of physical constants and properties of materials; development of methods for testing materials, mechanisms, and structures, and making such tests as may be necessary, particularly for government agencies; cooperation in the establishment of standard practices for incorporation in codes and specifications; advisory service to government agencies on scientific and technical problems; invention and development of devices to serve special needs of the Government; assistance to industry, business, and consumers in the development and acceptance of commercial standards and simplified trade practice recommendations; administration of programs in cooperation with United States business groups and standards organizations for the development of international standards of practice; and maintenance of a clearinghouse for the collection and dissemination of scientific, technical, and engineering information. The scope of the Bureau's activities is suggested in the following listing of its four Institutes and their organizational units.

Institute for Basic Standards. Applied Mathematics. Electricity. Metrology. Mechanics. Heat. Atomic Physics. Physical Chemistry. Laboratory Astrophysics.* Radiation Physics. Radio Standards Laboratory:* Radio Standards Physics; Radio Standards Engineering. Office of Standard Reference Data.

Institute for Materials Research. Analytical Chemistry. Polymers. Metallurgy. Inorganic Materials. Reactor Radiations. Cryogenics.* Materials Evaluation Laboratory. Office of Standard Reference Materials.

Institute for Applied Technology. Building Research. Information Technology. Performance Test Development. Electronic Instrumentation. Textile and Apparel Technology Center. Technical Analysis. Office of Weights and Measures. Office of Engineering Standards. Office of Invention and Innovation. Office of Technical Resources. Clearinghouse for Federal Scientific and Technical Information.**

Central Radio Propagation Laboratory.* Ionospheric Telecommunications. Tropospheric Telecommunications. Space Environment Forecasting. Aeronomy.

---

* Located at Boulder, Colorado 80301.

** Located at 5285 Port Royal Road, Springfield, Virginia 22171.

NATIONAL BUREAU OF STANDARDS REPORT

NBS PROJECT		NBS REPORT
20503-12-2050131	December 1, 1965	9011

NOTES ON THE USE OF PROPAGATION OF ERROR FORMULAS

by

Harry H. Ku

U.S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

Notes on the Use of Propagation of Error Formulas

by

H. H. Ku

ABSTRACT

The "law of propagation of error" is a tool which physical scientists have conveniently and frequently used in their work for many years, yet an adequate reference is difficult to find. In this paper an expository review of this topic is presented, particularly in the light of current practices and interpretations. Examples on the accuracy of the approximations are given. The reporting of the uncertainties of final results is discussed.

KEY WORDS: Approximation, Error, Formula, Imprecision, Law of error, Products, Propagation of error, Random, Ratio, Systematic, Sum

CONTENTS

Notes on the Use of Propagation of Error Formulas

H. H. Ku
National Bureau of Standards

Introduction

In the December, 1939, issue of the American Physics Teacher, Raymond T. Birge wrote an expository paper on "The Propagation of Errors." In the introductory paragraph of his paper, Birge remarked:

> "The question of what constitutes the most reliable value to be assigned as the uncertainty of any given measured quantity is one that has been discussed for many decades and, presumably, will continue to be discussed. It is a question that involves many considerations and by its very nature has no unique answer. The subject of the propagation of errors, on the contrary, is a purely mathematical matter, with very definite and easily ascertained conclusions. Although the general subject of the present article is by no means new,[1/] many scientists still fail to avail themselves of the enlightening conclusions that may often thus be reached, while others frequently use the theory incorrectly and thus arrive at quite misleading conclusions."

1/ See, for instance, M. Merriman, Method of Least Squares, (ed. 8, 1910), pp.75-79.

Birge's remark twenty-seven years ago still sounds fitting today. In what follows, a review of this topic will be made, particularly in the light of current practice and interpretations. Some examples on the accuracy of the approximations will also be given.

In Section I, we consider the two distinct situations under which the propagation of error formulas can be used. The mathematical manipulations are the same, yet the interpretations of the results are entirely different. In Section II the notations are defined and the general formulas given. A group of frequently used special formulas are listed at the end of the section for convenient reference. In Section III the accuracies of the approximations are discussed, together with suggestions on the use of the errors propagated. Section IV contains suggestions on the reporting of final results.

I. Statistical Tolerancing vs Imprecision of a Derived Quantity.

1.1 Propagation of error formulas are frequently used by engineers in the type of problem called "Statistical tolerancing." In such problems, we are concerned with the behavior of the characteristic W of a system as related to the behavior of a characteristic X of its component. For instance, an engineer may have designed a circuit. A property W of the circuit may be related to

the value X of the resistance used. As the value of X is changed, W changes, and the relationship can be expressed by a mathematical function

$$W = F(X)$$

within a certain range of the values of X.

Suppose our engineer decides on $W = w_o$ to be the desired property of the circuit, and specifies $X = x_o$ for this purpose. He realizes, however, that there will be variations among the large lot of resistors he ordered, no matter how tight his specifications are. Let x denote the value of any one of the resistors in the lot, then some of the time x will be below x_o, while at other times x will be above x_o. In other words, x has a distribution of values somewhat clustered about x_o. As x varies with each resistor, so does w with each circuit manufactured.

If our engineer knows the mean and standard deviation (or variance) of x, based on data from the history of their manufacture, then he can calculate the approximate mean and variance of w by the propagation of error formulas:

$$\text{mean } (w) \doteq F(\text{mean } x), \text{ and}$$

$$\text{variance } (w) \doteq \left[\frac{dF}{dX} \right]^2 \text{var}(x), \qquad (1.1)$$

where the square brackets signify that the derivatives within the brackets are to be evaluated at the mean of x. The approximations computed refer to the mean and variance of an individual unit in the collection of circuits that will be manufactured from the lot of resistors. The distribution of values of w, however, is still far from being determined since it depends entirely on the functional form of the relation between W and X, as mathematical variables, and the distribution of x itself, as a random variable. This type of approach has been used frequently in preliminary examinations of the reliability of performance of a system, where X may be considered as a multi-dimensional variable.

1.2 Let us consider now the second situation under which propagation of error formulas are used. This situation is the one considered in Birge's paper, and is the one that will be discussed in the main part of this paper.

A physicist may wish to determine the "true" value w_o of interest, for example, the atomic weight of silver. He makes n independent measurements on some related quantity x and calculates

$$\bar{x}_n = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$ as an estimate of the true value x_o and

$$s^2 = \frac{1}{n-1} \sum_{1}^{n} (x_i - \bar{x}_n)^2$$ as an index of dispersion of his measured values.

The physicist is mainly concerned in obtaining an estimate $\hat{w}$ of w_o, and of the standard deviation of $\hat{w}$ as a measure of precision of his result. He therefore computes by the propagation of error formulas:

$$\hat{w} = F(\bar{x}_n)$$
$$\hat{\text{var}}(\hat{w}) = \left[\frac{dF}{dX}\right]^2 \frac{s^2}{n} \qquad (1.2)$$
$$\hat{\sigma}_{\hat{w}} = \sqrt{\hat{\text{var}}(\hat{w})}$$

Often he assumes that $\hat{w}$ is distributed at least approximately in accordance with normal law of error and gives probability limits to the statistical uncertainty of his estimate $\hat{w}$ based on the standard deviation calculated ($\hat{\sigma}_{\hat{w}}$) and this assumption.

Cramér (1946) has shown that under very general conditions, functions of sample moments are asymptotically normal, with mean and variance given by the respective propagation of error formulas.[1/] Since $\bar{x}_n$ is the first sample moment, the estimate $\hat{w}$ will be approximately normaly distributed for large n. Hence our physicist is interested in the variance (or the standard deviation) of the normal distribution which the distribution of $F(\bar{x}_n)$ approximates as n increases. (Note that both estimators $\hat{w}$ and $\hat{\text{var}}(\hat{w})$ are functions of n.) For n large, the distribution of $\hat{w}$ can be assumed to be approximately normal and probability statements can be made about $\hat{w}$.

1/ A brief summary is given in paragraph 2.2 below.

1.3 Hence, we have the two cases:

(1) The problem of determining the mean and variance (or standard deviation) of the actual distribution of a given function F(x) of a particular random variable x, and

(2) The problem of estimating the mean and variance (or standard deviation) of the normal distribution to which the distribution of $F(\bar{x}_n)$ tends asymptotically.

As examples of problems studied under the first case, we can cite Fieller (1932) on the ratio of two normally distributed random variables, and Craig (1937) and Goodman (1962) on the product of two or more random variables. Tukey, in three Princeton University reports, extended the classical formulas through the fourth order terms for the mean and variance, and propagated the skewness and elongation of the distribution of F(x) as well. These reports present perhaps the most exhaustive treatment of statistical tolerancing to date.

From now on we shall be concerned in this paper with the second case only, i.e., the problem of estimating the mean and variance, or standard deviation, of the normal distribution to which the distribution of $F(\bar{x}_n)$ tends as n increases indefinitely, and hence also the problem of using approximations to the mean and variance computed from a finite number of measurements. Since the mean and standard

deviation are the parameters that specify a particular normal distribution, our problem is by its very nature less complicated than that of statistical tolerancing where the actual distribution of the function may have to be specified. We shall, however, utilize formulas given in Tukey's reports to check on the adequacy of some of the approximations.

II. Propagation of Error Formulas

2.1 Definitions and Notations

1. X, Y, Z in capitals stand for the mathematical variables to be measured; x, y, z in lower cases stand for the measured values of these variables; x_i, y_j, z_k with subscripts stand for the particular values of the i^{th} measurement on x, the j^{th} on y, and the k^{th} on z, respectively.

2. $W = f(X,Y,Z)$ is a continuous function of the variables X, Y, Z , with derivatives

$$\frac{\partial W}{\partial X} , \frac{\partial^2 W}{\partial X \partial Y} , \text{ etc.}$$

3. All derivatives appearing in square brackets, for example $\left[\frac{\partial W}{\partial X}\right]$, $\left[\frac{\partial W}{\partial Y}\right]$, stand for the values of these derivatives evaluated at the means of $\overset{\text{and}}{x \wedge y}$, if known, or at the sample averages of $\overset{\text{and}}{x \wedge y}$, if the means are not known.

4. In order to emphasize the fact that the mean M, variance σ^2 and other population parameters are usually not known, we list here symbols for both the estimators of population values and the population values. For a particular set of valuesof x, the values computed from these estimators are estimates, or computed values of these estimators.

Estimators of parameters	Corresponding Population Parameters
$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$	M_x (mean = first moment)
$s_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$	σ_x^2 (variance = second central moment)
$= \frac{1}{n-1} \left\{ \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n} \right\}$	
$s_{xy} = s_{yx} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$	$\sigma_{xy} = \sigma_{yx}$ (covariance)
$= \frac{1}{n-1} \left\{ \Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n} \right\}$	
$r_{xy} = \frac{s_{xy}}{s_x s_y} = r_{yx}$	ρ_{xy} (correlation coefficient)
s_x	σ_x (standard deviation of x about M_x)
$s_{\bar{x}} = \frac{1}{\sqrt{n}} s_x$	$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$ (standard deviation of the average $\bar{x}$, or standard error)
$v_x = \frac{s_x}{\bar{x}}$	$\sigma_{\bar{M}}$ (coefficient of variation or relative standard deviation)

In addition, we use $|\Delta x|$ to denote the bound for possible systematic errors on the measurements of x. The bound of these errors, unknown in sign, is usually established or conjectured by the experimenter and its value is not based on the measurements in hand.

2.2 General Theorem and Remarks

As mentioned briefly in paragraph 1.2, the propagation of error formulas are special applications of results obtained in the study of properties of distributions of functions of sample moments. Doob (1935), Hsu (1949), and others have investigated the limiting distribution of functions of sample means relating to hypothesis testing. Curtiss (1943) derived the limiting means and variances of the several functions of variables in connection with transformations used in the analysis of variance. Cramér, in Chapters 27 and 28 of his classical treatise, proved two theorems and also discussed the asymptotic properties of distributions of functions of sample moments in detail. For convenient reference we shall phrase his theorems and remarks in terms of functions of sample averages, to serve as a basis of justification for the use of propagation of error formulas.

Theorem (Cramér, pp.366, 352-356)

If, in some neighborhood of the point $X = M_x$, $Y = M_y$, the function $F(X,Y)$ is continuous and has continuous derivatives of the first and second order with respect to the arguments X and Y, the random variable $\hat{w} = F(\bar{x},\bar{y})$ is asymptotically normal, the mean and variance of the limiting normal distribution being given by:

$$\text{mean } \hat{w} = F(M_x, M_y) \tag{2.1}$$

$$\text{var } \hat{w} = \left[\frac{\partial F}{\partial X}\right]^2 \frac{\sigma_x^2}{n} + \left[\frac{\partial F}{\partial Y}\right]^2 \frac{\sigma_y^2}{n} + 2\left[\frac{\partial F}{\partial X}\right]\left[\frac{\partial F}{\partial Y}\right] \frac{\sigma_{xy}}{n} \tag{2.2}$$

Remark 1. (Cramér, p.367)

It follows from this theorem that any function of sample averages is, for large values of n, approximately normally distributed about the value of the function determined by the mean values of the basic variables, with a variance of the form C/n, provided only that expressions (2.1) and (2.2) yield finite values for the mean and the variance of the limiting distribution.

Remark 2. (Cramér, pp.367, 415, also Doob, Hsu)

In general, the constant C in the expression of the variance will have a positive value. However, in exceptional cases C may be zero, which implies that the variance is of a smaller order than n^{-1}. Then some expression of the form

$$n^p\left\{\hat{w} - F(M_x, M_y)\right\}, \quad p > \tfrac{1}{2},$$

may have a definite limiting distribution, but this is not necessarily normal.

Remark 3. (Cramér, pp.366, 213-214)

The function $F(\bar{x}, \bar{y})$ may be asymptotically normal even though the mean and variance of $F(\bar{x}, \bar{y})$ do not exist, or do not tend to the mean and variance of the limiting normal form. Generally, if the distribution of a random variable w depends on a

parameter n, and if two quantities M and σ can be found such that the distribution function of the variable $\frac{w-M}{\sigma}$ tends to $\Phi(t)$ (normal distribution function with mean zero and standard deviation one) as $n \to \infty$, we shall say that w is asymptotically normal (M, σ). This does not imply that the mean and the standard deviation of w tends to M and σ, nor even that these moments exist, but is simply equivalent to saying that for any interval (a,b) not depending on n,

$$\lim_{n \to \infty} \text{Prob.} \; (M + a\sigma < w < M + b\sigma) = \Phi(b) - \Phi(a) \; .$$

Example: If x is from a normal population the function $\ln \bar{x}$ is not even defined for all values of $\bar{x}$, and therefore the mean of the function $\ln \bar{x}$ does not exist; yet when the mean of $\bar{x}$ has a positive value, (2.1) and (2.2) give the mean and variance of the limiting normal distribution of $\ln \bar{x}$.

2.3 Propagation of Error Formulas

Fortified with the general theorem stated in the preceding paragraph, we shall proceed to derive the traditional propagation of error formulas in an elementary manner, making some comments and assumptions that may be of interest. It will be helpful, however, to explain first what is meant here by the term "random error" in a measurement process.

2.3.1 Random Errors

In a measurement situation, we consider random errors typically to be the sum total of all the small negligible independent errors over which we have no control - interpolation in reading scales, slight fluctuation in environment conditions, imperfection and non-constancy of our senses, etc. Thus for a stable measurement process, we find that:

1. The measured values do follow a distribution, with small errors occurring more frequently than larger ones, and with positive and negative errors about balancing one another, and
2. There is no obvious trend or pattern in the sequence of measurements.

Let us denote the i-th measurement of x to be

$$x_i = M_x + \epsilon_i$$

where M_x is the mean of all measurements for the measurement process, and ϵ_i the random error of measurement x_i. Then for condition 1, we assume A_1: The distribution of errors is symmetrical and bell-shaped, with mean zero and standard deviation σ_x, or

$$\begin{aligned} \text{mean } \epsilon_i &= 0 \\ \text{mean } x_i &= M_x \\ \text{var } x &= \text{var } \epsilon_i \\ &= \text{mean } \epsilon_i^2 = \sigma_x^2 \end{aligned}$$

And for condition 2, we assume A_2: The errors in the measurements of $x_i (i = 1,2,\ldots,n)$ are statistically independent; in particular these errors are not correlated or associated in any way, i.e.,

$$\text{mean}(\epsilon_i \cdot \epsilon_j) = 0 , \quad i \neq j .$$

Thus for $\bar{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$, the mean of $\bar{x}$ is M_x. Furthermore,

$$\bar{x} - M_x = \frac{\epsilon_1 + \epsilon_2 + \ldots + \epsilon_n}{n} .$$

By definition, the variance of $\bar{x}$ is:

$$\text{mean}(\bar{x} - M_x)^2 = \text{mean}\left(\frac{\epsilon_1 + \epsilon_2 + \ldots + \epsilon_n}{n}\right)^2$$

$$= \frac{1}{n^2}\left\{\text{mean}\left(\sum_{i=1}^{n} \epsilon_i^2\right) + \text{mean}\left(\sum_{i\neq j} \epsilon_i \epsilon_j\right)\right\}$$

$$= \frac{1}{n^2}\left\{n \ \text{mean}(\epsilon_i)^2 + \sum_{i\neq j} \text{mean}(\epsilon_i \epsilon_j)\right\} .$$

Using assumption A_1 and A_2, we obtain

$$\text{var}(\bar{x}) = \frac{1}{n} \sigma_x^2$$

or the variance of the average of n <u>independent</u> measurements is $\frac{1}{n}$ of the variance of an individual measurement. If, however, the measurements are not independent, then this formula is incorrect since the means of products $(\epsilon_i \epsilon_j)$ are not equal to zero. In that case let

$$\text{mean}(\epsilon_i \epsilon_j) = \rho_{ij} \sigma_x^2 , \text{ and } \bar{\rho} = \sum_{i\neq j} \rho_{ij}/n(n-1) , \text{ then}$$

$$\text{var}(\bar{x}) = \frac{\sigma_x^2}{n}\left\{1 + (n-1)\bar{\rho}\right\} .$$

Here the average $\bar{x}$ is a linear function of the individual x's, and the exact expression of mean

and variance of an average in term of that of the individual values are well-known. For functions that are not linear in the x's, we expand the function about the mean of x by the Taylor series, and assume that the function in the neighborhood of the mean can be approximated by the lower order terms. For example, let

$$W = F(X,Y),$$

$$x = M_x + \epsilon_x ,$$

$$y = M_y + \epsilon_y ,$$

where each of ϵ_x and ϵ_y satisfies assumptions A_1 and A_2, then we can write

$$F(x,y) = F(M_x,M_y) + \left[\frac{\partial F}{\partial X}\right]\epsilon_x + \left[\frac{\partial F}{\partial Y}\right]\epsilon_y + \frac{1}{2!}\left\{\left[\frac{\partial^2 F}{\partial X^2}\right]\epsilon_x^2 + 2\left[\frac{\partial^2 F}{\partial X \partial Y}\right]\epsilon_x\epsilon_y + \left[\frac{\partial^2 F}{\partial Y^2}\right]\epsilon_y^2\right\} + \text{terms of higher orders in } \epsilon_x \text{ and } \epsilon_y. \quad (2.3)$$

or, neglecting terms of higher order than ϵ_x^2 and ϵ_y^2 ,

$$F(x,y) - F(M_x,M_y) \doteq \left[\frac{\partial F}{\partial X}\right]\epsilon_x + \left[\frac{\partial F}{\partial Y}\right]\epsilon_y + \frac{1}{2!}\left\{\left[\frac{\partial^2 F}{\partial X^2}\right]\epsilon_x^2 + 2\left[\frac{\partial^2 F}{\partial X \partial Y}\right]\epsilon_x\epsilon_y + \left[\frac{\partial^2 F}{\partial Y^2}\right]\epsilon_y^2\right\} .$$

Since the means of ϵ_x and ϵ_y are 0, if we take mean on both sides,

$$\text{mean}\left\{F(x,y)-F(M_x,M_y)\right\} \doteq \frac{1}{2}\left\{\left[\frac{\partial^2 F}{\partial X^2}\right]\sigma_x^2 + 2\left[\frac{\partial^2 F}{\partial X \partial Y}\right]\sigma_{xy} + \left[\frac{\partial^2 F}{\partial Y^2}\right]\sigma_y^2\right\} . \quad (2.4)$$

Thus the mean of function of values always differs from the value of function of means by a quantity represented by (2.4), approximately. If the function of means $F(M_x, M_y)$ is the value of interest, then to approximate $F(M_x, M_y)$ by the mean of $F(x,y)$ would introduce an error, or bias, the magnitude of which depends on the functional form, the variances of and the covariance between x and y. If, however, we use the function of averages, $F(\bar{x}, \bar{y})$, then

$$\text{mean } \hat{w} = \text{mean } F(\bar{x}, \bar{y})$$

$$\doteq F(M_x, M_y) + \frac{1}{2}\left\{\left[\frac{\partial^2 F}{\partial X^2}\right]\frac{\sigma_x^2}{n} + 2\left[\frac{\partial^2 F}{\partial X \partial Y}\right]\frac{\sigma_{xy}}{n} + \left[\frac{\partial^2 F}{\partial Y^2}\right]\frac{\sigma_y^2}{n}\right\}, \quad (2.5)$$

and the bias is only 1/n times that of the mean of function of individual values. When n becomes large, this bias tends to zero, and (2.1) results.

This bias can be calculated by (2.5) and compared to the standard deviation of $\hat{w}$. In practice, if σ_x and σ_y are small, the bias is often of a magnitude that is negligible.

To propagate the variance, we note that if ϵ_x and ϵ_y are small in the sense that the second and higher order terms in (2.3) can be collectively neglected in comparison to terms involving ϵ_x and ϵ_y only, then

$$F(x,y) - F(M_x, M_y) \doteq \left[\frac{\partial F}{\partial X}\right]\epsilon_x + \left[\frac{\partial F}{\partial Y}\right]\epsilon_y \quad ,$$

and the variance of $F(x,y)$ is, approximately,

$$\text{mean}\left\{F(x,y) - F(M_x, M_y)\right\}^2 \doteq \text{mean}\left\{\left[\frac{\partial F}{\partial X}\right]\epsilon_x + \left[\frac{\partial F}{\partial Y}\right]\epsilon_y\right\}^2$$

$$= \left[\frac{\partial F}{\partial X}\right]^2 \sigma_x^2 + \left[\frac{\partial F}{\partial Y}\right]^2 \sigma_y^2 + 2\left[\frac{\partial F}{\partial X}\right]\left[\frac{\partial F}{\partial Y}\right] \sigma_{xy} \quad . \tag{2.6}$$

And for $\hat{w} = F(\bar{x}, \bar{y})$, the variance of $\hat{w}$ is

$$\text{var}(\hat{w}) \doteq \left[\frac{\partial F}{\partial X}\right]^2 \frac{\sigma_x^2}{n} + \left[\frac{\partial F}{\partial Y}\right]^2 \frac{\sigma_y^2}{n} + 2\left[\frac{\partial F}{\partial X}\right]\left[\frac{\partial F}{\partial Y}\right] \frac{\sigma_{xy}}{n} \tag{2.7}$$

the limiting form of which is (2.2).

Finally, if σ_x^2, σ_y^2, and σ_{xy} are not known, we substitute their estimators in formulas (2.5) and (2.7), resulting in:

$$\hat{\text{mean}}(\hat{w}) \doteq F(M_x, M_y) + \frac{1}{2}\left\{\left[\frac{\partial^2 F}{\partial X^2}\right] \frac{s_x^2}{n} + \left[\frac{\partial^2 F}{\partial Y^2}\right] \frac{s_y^2}{n} + 2\left[\frac{\partial^2 F}{\partial X \partial Y}\right] \frac{s_{xy}}{n}\right\} \quad , \tag{2.8}$$

and

$$\hat{\text{var}}(\hat{w}) \doteq \left[\frac{\partial F}{\partial X}\right]^2 \frac{s_x^2}{n} + \left[\frac{\partial F}{\partial Y}\right]^2 \frac{s_y^2}{n} + 2\left[\frac{\partial F}{\partial X}\right]\left[\frac{\partial F}{\partial Y}\right] \frac{s_{xy}}{n} \quad . \tag{2.9}$$

If we assume further that the random errors in measurements of x and y are independent, then $\sigma_{xy} = 0$, and the terms involving σ_{xy} in (2.5), (2.6) and (2.7) vanishes. If this is the case, the terms involving s_{xy} in (2.8) and (2.9) should also be dropped. This reduced version of the formula for independent x and y,

$$\mathrm{var}(\hat{W}) \doteq \left[\frac{\partial F}{\partial \bar{X}}\right]^2 \frac{\sigma_x^2}{n} + \left[\frac{\partial F}{\partial \bar{Y}}\right]^2 \frac{\sigma_y^2}{n} , \qquad (2.10)$$

is of the form given in Birge's paper and in other textbooks on statistical analysis of data (Mandel, 1964, pp.72-76).

For $W = F(X,Y,Z)$, there will be three variance and three covariance terms in (2.5) and (2.7). Extension to more than three variables presents no new problems.

2.3.2 Extension to more than one function of the variables.

Let

$$U = g(X, Y, Z), \quad \text{and}$$

$$V = h(X, Y, Z) .$$

Then in addition to the above formulas, we have

$$\begin{aligned}\sigma_{uv} \doteq & \left[\frac{\partial U}{\partial X}\cdot\frac{\partial V}{\partial X}\right]\sigma_x^2 + \left[\frac{\partial U}{\partial Y}\cdot\frac{\partial V}{\partial Y}\right]\sigma_y^2 + \left[\frac{\partial U}{\partial Z}\cdot\frac{\partial V}{\partial Z}\right]\sigma_z^2 \\ & + \left\{\left[\frac{\partial U}{\partial X}\cdot\frac{\partial V}{\partial Y}\right] + \left[\frac{\partial U}{\partial Y}\cdot\frac{\partial V}{\partial X}\right]\right\}\rho_{xy}\sigma_x\sigma_y \\ & + \left\{\left[\frac{\partial U}{\partial Y}\cdot\frac{\partial V}{\partial Z}\right] + \left[\frac{\partial U}{\partial Z}\cdot\frac{\partial V}{\partial Y}\right]\right\}\rho_{yz}\sigma_y\sigma_z \\ & + \left\{\left[\frac{\partial U}{\partial Z}\cdot\frac{\partial V}{\partial X}\right] + \left[\frac{\partial U}{\partial X}\cdot\frac{\partial V}{\partial Z}\right]\right\}\rho_{zx}\sigma_z\sigma_x \quad .\end{aligned} \tag{2.11}$$

]

Expression (2.11) may be convenient to use to get $\sigma(\hat{w})$ where $W = F(U,V)$, and U and V are known functions of X,Y, and Z .

2.3.3 Some Frequently Used Formulas

For convenience, a few special formulas for commonly encountered functions are listed in table 1 with x,y assumed to be independent. These may be derived from the above formulas.

Table 1

Propagation of Error Formulas for Some Simple Functions

Function Form of $\hat{w}$ [1/]	Approx. Formula for $\hat{Var}(\hat{w})$ (x and y are assumed to be statistically independent)	Term to be added if x and y are correlated, and a reliable estimate of $\sigma_{\bar{x}\bar{y}}$, $s_{\bar{x}\bar{y}}$, can be assumed
$A\bar{x} + B\bar{y}$	$A^2 s_{\bar{x}}^2 + B^2 s_{\bar{y}}^2$	$2AB\, s_{\bar{x}\bar{y}}$
$\left(\frac{\bar{x}}{\sigma_{\bar{x}}^2} + \frac{\bar{y}}{\sigma_{\bar{y}}^2}\right) / \left(\frac{1}{\sigma_{\bar{x}}^2} + \frac{1}{\sigma_{\bar{y}}^2}\right)$ [2/]	$1 / \left(\frac{1}{\sigma_{\bar{x}}^2} + \frac{1}{\sigma_{\bar{y}}^2}\right)$	
$\frac{\bar{x}}{\bar{y}}$	$\left(\frac{\bar{x}}{\bar{y}}\right)^2\left(\frac{s_{\bar{x}}^2}{\bar{x}^2} + \frac{s_{\bar{y}}^2}{\bar{y}^2}\right)$	$\left(\frac{\bar{x}}{\bar{y}}\right)^2\left(-2\,\frac{s_{\bar{x}\bar{y}}}{\bar{x}\bar{y}}\right)$
$\frac{1}{\bar{y}}$	$\frac{s_{\bar{y}}^2}{\bar{y}^4}$	
$\frac{\bar{x}}{\bar{x}+\bar{y}}$	$\left(\frac{\hat{w}}{\bar{x}}\right)^4\left(\bar{y}^2 s_{\bar{x}}^2 + \bar{x}^2 s_{\bar{y}}^2\right)$	$\left(\frac{\hat{w}}{\bar{x}}\right)^4\left(-2\bar{x}\bar{y}\, s_{\bar{x}\bar{y}}\right)$
$\frac{\bar{x}}{1+\bar{x}}$	$\frac{s_{\bar{x}}^2}{(1+\bar{x})^4}$	
$\bar{x}\bar{y}$	$(\bar{x}\bar{y})^2\left(\frac{s_{\bar{x}}^2}{\bar{x}^2} + \frac{s_{\bar{y}}^2}{\bar{y}^2}\right)$	$(\bar{x}\bar{y})^2\left(2\,\frac{s_{\bar{x}\bar{y}}}{\bar{x}\bar{y}}\right)$

Table 1 (Cont.)

Function Form of $\hat{w}$	Approx. Formula for $\hat{Var}(\hat{w})$...	Term to be added...
$\bar{x}^2$	$4\bar{x}^2 s_{\bar{x}}^2$	
$\sqrt{\bar{x}}$	$\frac{1}{4}\frac{s_{\bar{x}}^2}{\bar{x}}$	
$\ln \bar{x}$	$\frac{s_{\bar{x}}^2}{\bar{x}^2}$	
$k\bar{x}^a\bar{y}^b$	$\hat{w}^2\left(a^2 \frac{s_{\bar{x}}^2}{\bar{x}^2} + b^2 \frac{s_{\bar{y}}^2}{\bar{y}^2}\right)$	$(\hat{w})^2\left(2ab \frac{s_{\bar{x}\bar{y}}}{\bar{x}\bar{y}}\right)$
$*e^{\bar{x}}$ 3/	$e^{2\bar{x}} s_{\bar{x}}^2$	
$\frac{\sin\left(\frac{\bar{x}+\bar{y}}{2}\right)}{\sin\frac{\bar{x}}{2}}$	$\frac{1}{4}\left\{\frac{\cos^2\left(\frac{\bar{x}+\bar{y}}{2}\right)}{\sin^2\frac{\bar{x}}{2}} s_{\bar{y}}^2 + \frac{\sin^2\frac{\bar{y}}{2}}{\sin^4\frac{\bar{x}}{2}} s_{\bar{x}}^2\right\}$ ($s_{\bar{x}}$ and $s_{\bar{y}}$ in radians)	$-\frac{\sin\frac{\bar{y}}{2}\cos\left(\frac{\bar{x}+\bar{y}}{2}\right)}{2\sin^3\frac{\bar{x}}{2}} s_{\bar{x}\bar{y}}$
$100\frac{s_{\bar{x}}}{\bar{x}}$ (= coefficient of variation in percent)	$\frac{\hat{w}^2}{2(n-1)}$ (not directly derived from the formulas) 4/	

1/ It is assumed that the value of w is finite and real, e.g., $\bar{y} \neq 0$ for ratios with $\bar{y}$ as denominator, $\bar{x} > 0$ for $\sqrt{\bar{x}}$ and $\ln \bar{x}$.

2/ Weighted mean as a special case of $A\bar{x} + B\bar{y}$, with σ_x and σ_y considered known.

3/ Distribution of $\hat{w}$ is highly skewed and normal approximation could be seriously in error for small n.

4/ See, for example, Statistical Theory with Engineering Applications, p.301, by A. Hald, John Wiley and Sons, 1952.

2.4 Systematic Errors

By a systematic error we mean a fixed deviation that is inherent in each and every measurement of x in a particular sequence of measurements. If the magnitude and direction of the systematic error are known, a correction can be made such that $M_x = x_o$, or the mean of the sequence of measurements is equal to the value sought after. If the sign of the systematic error is not known and the magnitude of the error can be only estimated to be within some reasonable bound $|\Delta x|$, perhaps by experience or judgement, then M_x is within the limits $x_o - \Delta x$ and $x_o + \Delta x$.

For a function of two variables $W = F(X,Y)$ then, a bound $|\Delta w|$ for the systematic error in W is given by:

$$|\Delta w| \doteq \left|\left[\frac{\partial F}{\partial X}\right]\Delta x\right| + \left|\left[\frac{\partial F}{\partial Y}\right]\Delta y\right| , \qquad (2.12)$$

assuming, as before, that Δx and Δy are small such that second and higher order terms in Δx and Δy are collectively negligible in the Taylor series expansion. Since ordinarily we do not know the signs of Δx and Δy, we have no choice but to add the absolute values of the two terms together, even though the signs of the values of the partial derivatives evaluated are known. (If the signs of either Δx or Δy is known, this information, of course, should not be ignored.) If these derivatives are evaluated at the point $\bar{x}$ and $\bar{y}$, then the random components of error of $\bar{x}$ and $\bar{y}$ are required to be small so that these derivatives take approximately the same

values as when evaluated at x_o and y_o .

When there are a number of systematic errors to be propagated, one approach is to take $|\Delta w|$ as the square root of the sum of squares of terms on the right-hand side of (2.12), instead of adding together the absolute values of all the terms. This procedure presupposes that some of the systematic errors may be positive and the others negative, and the two classes cancel each other to a certain extent.

The treatment of inaccuracy due to systematic errors of assignable origin but of unknown magnitudes is discussed in detail in section 4.2 of Eisenhart (1963). Since there is no generally accepted standard method for combining several systematic errors, Eisenhart advised and we quote

> "Therefore, anyone who uses one of these methods for the 'combination of errors' should indicate explicitly which of these (or an alternative method) he has used."

Information on the source and magnitude of each contributing elemental systematic error is, of course, also essential.

III. Practical Accuracies at the Various Stages of Approximations

3.1 From the preceding sections we observe that there are three stages of approximations:

1. In the Taylor series expansion (2.3), terms higher than the first partial derivatives are considered to be negligible.
2. $\hat{w}$ is approximately normally distributed for large n. Is the normal distribution still a good approximation for small n?
3. If σ_x^2 and σ_y^2 are known, we obtain $\sigma_{\hat{w}}^2$ from (2.7), and we can use this value to construct a confidence interval[1/] about $\hat{w}$ with the desired level of confidence (approximate) based on normal theory. If σ_x^2 and σ_y^2 are not known, and s_x^2 and s_y^2 are calculated from a small number of measurements, what can we say about $\hat{w}$ using $\hat{\text{var}}(\hat{w})$ calculated from (2.9)?

To get some numerical feeling for the closeness of these approximations, we shall simplify matters by making the following assumptions which do not seem to be too restrictive in measurement situations:

B_1: x and y are normally and independently distributed, with the ratio M/σ not less than 10.[2/]

1/ See Natrella (1963), Section 1-7, also Chapters 2 and 3.

2/ For notational convenience, the symbols $w, x, y, \sigma_x, \sigma_y$, etc. are used in this and the subsequent sections. The corresponding symbols for the average could be used by straight substitution.

B_2: The functional forms used are the well-behaved ones that do not possess derivatives assuming unreasonably large values when evaluated at the averages of the individual variables.

Thus for linear functions, such as

$$W = AX + BY ,$$

the second and higher derivatives vanish, and (2.6) is exact.

The adequacy of these approximations are studied in paragraphs 3.2 and 3.3 below. In paragraph 3.4 suggestions are made on the use of standard deviation calculated for $\hat{w}$ when the standard deviations of $\bar{x}$ and $\bar{y}$ are not known. Readers may wish to go directly to paragraph 3.5 for a summary of the conclusions.

3.2 For x,y independently distributed and arbitrary F(x,y), the first correction terms to (2.6) are:

$$\left[\frac{\partial F}{\partial X}\right]\left[\frac{\partial^2 F}{\partial X^2}\right]\gamma_x\sigma_x^3 + \left[\frac{\partial F}{\partial Y}\right]\left[\frac{\partial^2 F}{\partial Y^2}\right]\gamma_y\sigma_y^3, \quad (3.1)$$

where γ is a measure of skewness of the distribution.[1/] Therefore these terms equal zero for x,y symmetrically distributed, a condition satisfied by assumption B_1.

The next order of correction terms involve σ_x^4, σ_y^4 and $\sigma_x^2\sigma_y^2$ and are usually negligible compared to terms in (2.6). These terms are:

$$\begin{aligned}
&\frac{1}{3}\left\{\left[\frac{\partial F}{\partial X}\right]\left[\frac{\partial^3 F}{\partial X^3}\right]\Gamma_x\sigma_x^4 + \left[\frac{\partial F}{\partial Y}\right]\left[\frac{\partial^3 F}{\partial Y^3}\right]\Gamma_y\sigma_y^4\right\} \\
&+ \frac{1}{4}\left\{\left[\frac{\partial^2 F}{\partial X^2}\right]^2(\Gamma_x-1)\sigma_x^4 + \left[\frac{\partial^2 F}{\partial Y^2}\right]^2(\Gamma_y-1)\sigma_y^4\right\} \qquad (3.2) \\
&+ \left\{\left[\frac{\partial F}{\partial X}\right]\left[\frac{\partial^3 F}{\partial X\partial Y^2}\right] + \left[\frac{\partial^2 F}{\partial X\partial Y}\right]^2 + \left[\frac{\partial^3 F}{\partial X^2\partial Y}\right]\left[\frac{\partial F}{\partial Y}\right]\right\}\sigma_x^2\sigma_y^2
\end{aligned}$$

For functions involving powers of x and y

1/ For definition of γ and Γ see (3.3).

less than three, some of the partial derivatives also vanish, . For example, if $W = XY$, the only non-zero term of this order is $\sigma_x^2\sigma_y^2$, or

$$\mathrm{Var}(w) = M_y^2\sigma_x^2 + M_x^2\sigma_y^2 + \sigma_x^2\sigma_y^2 \quad .$$

The contribution of $\sigma_x^2\sigma_y^2$ is less than 1 in 200 if M/σ is larger than ten.

For functional forms such as quotients, roots and logarithms, the accuracy is usually adequate since powers of the means of the variables appear in the denominators of the partial derivatives.

For the exponential function $W = e^X$, the variance of w as given by (2.6) is

$$\mathrm{Var}(w) \doteq e^{2M_x}\sigma^2,$$

whereas the exact formula[1/]

for the variance of w, when x is normaly distributed, is

$$\begin{aligned}\mathrm{Var}(w) &= e^{\sigma^2}e^{2M}\left(e^{\sigma^2} - 1\right)\\ &= e^{\sigma^2}e^{2M}\left(\sigma^2 + \frac{\sigma^4}{2!} + \frac{\sigma^6}{3!} + \cdots\right)\\ &= e^{2M}\sigma^2\left\{e^{\sigma^2}\left(1 + \frac{\sigma^2}{2!} + \frac{\sigma^4}{3!} + \cdots\right)\right\} \quad .\end{aligned}$$

Here the variance of w as given by (2.6) underestimates the true variance by the factor given in the brackets, and the approximation could be seriously in error. (Note,

1/ See, for example, The Lognormal Distribution, p.8, by J. Aitchison and J.A.C. Brown, Cambridge University Press, 1957.

however, the "exact" formula is correct only if x is exactly normally distributed. If x is only approximately normally distributed, then both formulas are approximations.)

For specific functions, formulas (3.1) and (3.2) given in Tukey's report can be used to check on the adequacy of the approximation. We quote Tukey's conclusion in this respect:

> "The most important conclusion is that the classical propagation formula is much better than seems to be usually realized. Examples indicate that it is quite likely to suffice for most work."

3.3 Next we look into the adequacy of the normal approximation. For this purpose we will define the first four central moments of the distribution of w as follows:

$$\begin{aligned} &\text{mean } (w - M_w) = 0 \\ &\text{mean } (w - M_w)^2 = \sigma^2 \\ &\text{mean } (w - M_w)^3 = \gamma\sigma^3 \\ &\text{mean } (w - M_w)^4 = \Gamma\sigma^4 \end{aligned} \qquad (3.3)$$

If w is normally distributed, $\gamma = 0$, and $\Gamma = 3$. Following Tukey, we shall define

$$\text{skewness} = \gamma\sigma^3 \text{ , and}$$

$$\text{elongation} = \Gamma\sigma^4 - 3\sigma^4 \text{ ,}$$

then both skewness and elongation are equal to zero when w is normally distributed.

If x and y are normally distributed as assumed under B_1, then in general $w = F(x,y)$ is not normally distributed unless the function form is linear. By a procedure similar to that used in the last section, the coefficients of skewness β_1 and excess β_2 of w can be calculated where:

$$\beta_1 = \frac{[\text{skewness } w]^2}{[\quad \text{var } w \quad]^3}$$

$$\beta_2 = \frac{\text{elongation } w}{[\quad \text{Var } w \quad]^2} + 3 .$$

If β_1 is close to zero and β_2 is close to 3, the normal approximation may be considered as adequate.

The terms up to order of σ^4 in the propagation of skewness for $w = F(x,y)$, with x,y independent, are

$$\text{skewness } w \doteq \left[\frac{\partial F}{\partial X}\right]^3 \gamma_x \sigma_x^3 + \left[\frac{\partial F}{\partial Y}\right]^3 \gamma_y \sigma_y^3$$
$$+ \frac{3}{2}\left[\frac{\partial F}{\partial X}\right]^2 \left[\frac{\partial^2 F}{\partial X^2}\right](\Gamma_x - 1)\sigma_x^4 + \frac{3}{2}\left[\frac{\partial F}{\partial Y}\right]^2 \left[\frac{\partial^2 F}{\partial Y^2}\right](\Gamma_y - 1)\sigma_y^4$$
$$+ 6\left[\frac{\partial F}{\partial X}\right]\left[\frac{\partial F}{\partial Y}\right]\left[\frac{\partial^2 F}{\partial X \partial Y}\right]\sigma_x^2 \sigma_y^2 \quad . \qquad (3.4)$$

For x,y normally distributed, only terms of order σ^4 remain. If we take $w = xy$ again as an example, then

$$\text{skewness } w \doteq 6\, M_x M_y\, \sigma_x^2 \sigma_y^2$$

$$\left[\quad - \frac{\qquad\qquad}{\qquad\qquad}\right.$$

$$\beta_1 \doteq \frac{36\, M_x^2 M_y^2 \sigma_x^4 \sigma_y^4}{[M_y^2\sigma_x^2 + M_x^2\sigma_y^2 + \sigma_x^2\sigma_y^2]^3} \quad .$$

Neglecting $\sigma_x^2\sigma_y^2$ in the brackets in the denominator, and taking $M/\sigma = 10$, β_1 is computed to be .045. Hence, for $\hat{w} = \bar{x}\bar{y}$, where $\bar{x}$ and $\bar{y}$ are averages of four, the coefficient of skewness is reduced by a factor of four or equals .011 approximately.

Similarly, terms up to order σ^4 for the elongation of $w = f(x,y)$, with x,y independent, are:

$$\text{elongation } w \doteq \left[\frac{\partial F}{\partial X}\right]^4 (\Gamma_x - 3)\sigma_x^4 + \left[\frac{\partial F}{\partial Y}\right]^4 (\Gamma_y - 3)\sigma_y^4 \qquad (3.5)$$

which is zero for x,y normal.

Hence $\beta_2 = \frac{\text{elongation } w}{(\text{Variance } w)^2} + 3 = 3$, and no correction for elongation is necessary here.

If we look up a table [1/] of percentage points of the distribution of standardized variate $\frac{\hat{w} - M_{\hat{w}}}{\sigma_{\hat{w}}}$ with given β_1 and β_2, we note that the changes of values are rather sensitive to β_1 and much less so to β_2. Thus the coefficient of elongation is usually not as much a source of worry in the normal approximation as is the coefficient of skewness.

1/ See Table 42, Biometrika Tables for Statisticians, Vol.1, edited by E.S. Pearson and H.O. Hartley, The University Press, 1958. Also pp.79-84.

Table 2. Departures From Normal Approximations

x,y independently , distributed, with

$\gamma \doteq 0$, $\Gamma \doteq 3$, and $(M/\sigma) = 10$.

$\hat{w}$	Skewness from (3.4)	β_1 computed	Percentage point of $\frac{\hat{w} - M_{\hat{w}}}{\sigma_{\hat{w}}}$ lower 2.5%	upper 2.5%
$A\bar{x}+B\bar{y}$	0	0	-1.96	+1.96 1/
$\bar{x}\bar{y}$	$6M_xM_y \frac{\sigma_x^2\sigma_y^2}{n^2}$	$\frac{4.5}{100n}$		
n=1		.045	-1.84	+2.06
n=4		.011	-1.91	+2.01
n=10		.0045	-1.93	+1.99
$\bar{x}^2$	$24M^2 \frac{\sigma^4}{n^2}$	$\frac{9}{100n}$		
n=10		.009	-1.90	+2.00
$\bar{x}^3$	$162M^5 \frac{\sigma^4}{n^2}$	$\frac{36}{100n}$		
n=4		.09	-1.80	+2.09
$\frac{\bar{x}}{\bar{y}}$ *	$6\left(\frac{M_x}{M_y}\right)^3\left(\frac{\sigma_y^4}{M_y^4}+\frac{\sigma_x^2\sigma_y^2}{M_x^2M_y^2}\right)$	$\frac{18}{100n}$		
n=10		.018	-1.89	+2.03
$\ln \bar{x}$ **	$\frac{3}{M^4} \cdot \frac{\sigma^4}{n^2}$	$\frac{9}{100n}$		
n=10		.009	-1.90	+2.00
$e^{\bar{x}}$	$3e^{3M} \frac{\sigma^4}{n^2}$	$\frac{9\sigma^2}{n}$	depends on σ and n (both skewness and β_1 underestimated for $\sigma/\sqrt{n} > .2$)	

* $\bar{y} > 0$

** $\bar{x} > 0$

1/ Exact when x and y are normally distributed.

Formulas (3.4) and (3.5) and the table of percentage points allow us to check how good the normal approximation is for a given number of measurements in the variables x and y. Table 2 gives some examples of results of such calculations.

3.4 The third approximation concerns the use of the sample variance s^2 as an estimate of the population variance σ^2. If we know the precision of the processes for the measurements of x and y, i.e., we know σ_x and σ_y, σ_w can be computed from (2.7) and a confidence interval about w can be constructed with the desired confidence coefficient $1-\alpha$ by using the table of the normal probability integral. If σ_x and σ_y are not known, then even if $\hat{\sigma}_{\hat{w}}$ can be computed from (2.9), the constants to be used for constructing a confidence interval with confidence coefficient $1-\alpha$ will be different from those for known σ.

To offer some guideline to the solution of this problem, we again assume measurements on x and y to be independently and normally distributed. If the number of measurements is large (a rule of thumb could be $n > 30$), then (2.7) can be

used assuming σ_x^2 , σ_y^2 and σ_{xy} are known.

Of course one can always compute the half-widths of the respective 100 $(1-\alpha)$ percent confidence intervals for M_x and for M_y by the use of the Student's t statistic, and use (2.12) to get the half-width of the interval for M_w, i.e., set

$$\Delta x = t_{(1-\frac{\alpha}{2}),\ n-1} \frac{s_x}{\sqrt{n}} \quad \text{and} \quad \Delta y = t_{(1-\frac{\alpha}{2}),\ k-1} \frac{s_y}{\sqrt{k}} \ ,$$

and use (2.12) to get Δw. Then the interval $w \pm \Delta w$ is a confidence interval for M_w for a confidence coefficient of at least $(1-\alpha)$. This procedure, however, may be criticized on the ground of gross inefficiency in using the data.

We may write (2.9) as

$$\hat{var}(\hat{w}) = \lambda_1 s_x^2 + \lambda_2 s_y^2$$

where $\lambda_1 = \frac{1}{n}\left[\frac{\partial F}{\partial X}\right]^2$ and $\lambda_2 = \frac{1}{k}\left[\frac{\partial F}{\partial Y}\right]^2$ are two constants. For given degrees of freedom for s_x, n-1, and s_y, k-1, and given ratios of $\frac{\lambda_1 s_x^2}{\lambda_1 s_x^2 + \lambda_2 s_y^2}$, values of a "v" statistic have been tabulated[2/] for confidence coefficients of .99,.98,.95, and .90 . The interval

$$\hat{w} \pm v\sqrt{\lambda_1 s_1^2 + \lambda_2 s_2^2} \qquad (3.6)$$

2/ See Table 11, Biometrika Tables for Statisticians, Vol.1; also Further critical values for the two-means problem, W.H. Trickett, B.L. Welch, and G.S. James, Biometrika 43, 1956, pp.204-5.

is a confidence interval with confidence coefficient $1 - \alpha$.

These tables, however, do not contain values for "v" for n and k less than 10,10,8, and 6 for the respective confidence coefficients, and hence cannot be used for smaller samples. In addition, they are useful only for two independent variables x and y .

Alternatively Welch (1947) has proposed the use of "effective degrees of freedom" for the estimated variance of $\hat{w}$ of the form

$$\hat{var}(\hat{w}) = \Sigma \lambda_i s_i^2 \quad .$$

The effective degree of freedom f is computed from

$$f = \frac{(\Sigma \lambda_i s_i^2)^2}{\Sigma(\lambda_i^2 s_i^4 / f_i)} \qquad (3.7)$$

where f_i is the degrees of freedom for[1/] s_i^2 .

In general f will be fractional. The t value with f degrees of freedom can be found or interpolated from the t table and the confidence interval computed as

$$\hat{w} \pm t_{(1-\frac{\alpha}{2}),\, f}\, \hat{\sigma}_{\hat{w}} \quad .$$

The approximate confidence intervals computed by the use of effective degrees of freedom were found to check the exact confidence intervals given by (3.6) very well over the range of the latter.

1/ If s_i^2 is computed from n_i measurements, the degrees of freedom is $n_i - 1$.

3.5 In summary, the following may be concluded for practical purposes:

1. Terms of order higher than σ^2 in the propagation of error formulas for variance, (2.6) and (2.7), can be neglected if (a) the standard deviations are small in comparison to their respective means, and (b) the second and higher order partial derivatives evaluated at the means do not give rise to abnormally large numbers. This is usually true in the field of physical science, since errors of measurements are usually of the order of 1 part in 1000, or parts per million; furthermore, the functional forms used are usually the well-behaved ones.
2. The normal approximation will be adequate for large n, or if, in addition to (a) and (b) above, (c) the individual variables can be assumed to be normally distributed. For particular functions, the approximate values of the coefficients of skewness and elongation may be calculated and Pearson's table can be used to check the adequacy of the approximation.
3. For the case where the standard deviations of the individual variables are unknown, and are estimated from the data, confidence intervals for the estimate $\hat{w}$ can be constructed either by the use of tabulated values of the "v" statistic or by the use of effective degrees of freedom. These confidence intervals can be considered

as a form of "precision limits" in the sense that if one makes the same sets of measurements a large number of times under the same conditions, and constructs the confidence intervals each time by the same procedure, then a large proportion of the intervals so constructed, $100(1-\alpha)\%$, will bracket the mean of all these sets of measurements. When only one set of measurements will be made, the probability is $1-\alpha$ that this interval will bracket the mean.

IV. Reporting of Results

4.1 Suppose a set of measurement data is available, and, by using the appropriate propagation of error formulas, the following are obtained for the quantity of interest, w_o:

1. The estimate of w_o, $\hat{w}$, based on n values of x,y, etc.;
2. The estimated standard error of $\hat{w}$, $\hat{\sigma}_{\hat{w}}$, and associated degrees of freedom f;
3. Limits to the systematic error in w, Δw .

The estimated standard error of $\hat{w}$ gives a measure of precisior of the experimental results, or a measure of scatter of the values of $\hat{w}$ from the average value of M_w for repeated performance of the particular experiment. But this measure of precision does not indicate at all how close this average value is to the value w_o intended to be measured. The estimation of limits to the systematic error is an essential part of an experiment and need not be discussed

here (Youden, 1961). One may remark generally that systematic errors usually do not pose a serious problem when the "imprecision" is large, since these systematic errors are, so to speak, "swallowed up" by the random errors. The systematic errors, however, play an important role when the precision is excellent and is of about the same order of magnitude as the systematic error. In that case, it is essential that the systematic error, or errors, be reported separately from the imprecision part of the reported value, as measured by the standard error, or the confidence intervals, computed.

In scientific literature, it is not uncommon to come across expressions of results in the form of $M \pm e$, where "M" is an average of some kind and "e" represents the uncertainty of "M" in some vague sense. This type of reporting proves to be most frustrating from the reader's point of view. From the context alone the reader cannot possibly infer whether "e" represents probable error, 3-sigma limits, systematic error, or some combination of random and systematic errors. As a consequence, the quality of the results, and the validity of inference drawn from these results, are to a large extent left to the judgment and guesswork of the reader. Hence, the writer owes to himself, and to his reader, to specify clearly the meaning of "e" as he uses it. In particular, the number

of measurements from which the measure of random error was computed and the manner in which the systematic error was estimated are both essential elements of the reported value and need to be included.

A footnote explaining the role of "e" is often very helpful. Several examples are given below:

> "In the expression of the form $M \pm e$, M is the average and e is the standard error of M based on n measurements (or based on ν degrees of freedom)."

> "The indicated uncertainty limits for M are over-all limits of error based on 95% confidence limits for the mean — and on allowances for effects of known sources of possible systematic error —."

> "The uncertainty given represents 3-sigma limits based on the current accepted value of the standard deviation, known sources of systematic errors being negligible."

Chapter 23 of Natrella (1963) "Expressions of the Uncertainties of Final Results" gives a thorough discussion on this topic, and is an excellent reference for all physical scientists who have occasion to report numerical results of their experiments.

References

[1] Birge, Raymond T., "The Propagation of Errors," The American Physics Teacher, Vol.7, No.6, December 1939.

[2] Craig, C.C., "On Frequency Function of XY," Annals of Mathematical Statistics, Vol.7, (1937),pp. 1-15.

[3] Cramér, Harald, Mathematical Methods of Statistics, Princeton University Press, 1946.

[4] Curtiss, J.H., "On Transformations Used in the Analysis of Variance," Annals of Math. Stat., Vol.14 (1943), pp.107-122.

[5] Doob, J.L., "The Limiting Distribution of Certain Statistics," Annals of Math. Stat., Vol.6, (1935), pp.160-170.

[6] Eisenhart, Churchill, "Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems," Journal of Research, National Bureau of Standards, Vol.67C, 1963, pp.161-187.

[7] Fieller, E.C., "The Distribution of the Index in a Normal Bivariate Population," Biometrika, Vol.24, (1932), pp.428-440.

[8] Goodman, L.A., "The Variance of the Product of K Random Variables," Journal of American Statistical Association, Vol.57, (1962), pp.54-60.

[9] Hsu, P.L., "The Limiting Distribution of Functions of Sample Means and Application to Testing Hypotheses," Berkeley Symposium on Mathematical Statistics and Probability, Univ. of California Press, (1949) pp.359-402.

[10] Mandel, John, The Statistical Analysis of Experimental Data, John Wiley and Sons, 1964.

[11] Natrella, Mary Gibbons, Experimental Statistics, Handbook 91, National Bureau of Standards, 1963.

[12] Tukey, John W., "The Propagation of Errors, Fluctuations and Tolerances," Unpublished Technical Reports No.10, 11, and 12, Princeton University.

[13] Welch, B.L., "The Generalization of 'Student's' problem when several different population variances are involved," Biometrika, 34, (1947), pp.28-35.

[14] Youden, W.J., "Systematic Errors in Physical Constants," Physics Today, Vol.14, September 1961.


DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS