THE CALCULATION OF ERRORS BY THE METHOD OF LEAST SQUARES

By Raymond T. Birge
University of California, Berkeley

(Received February 18, 1932)

Abstract

Present status of least squares' calculations.—There are three possible stages in any least squares' calculation, involving respectively the evaluation of (1) the most probable values of certain quantities from a set of experimental data, (2) the reliability or probable error of each quantity so calculated, (3) the reliability or probable error of the probable errors so calculated. Stages (2) and (3) are not adequately treated in most texts, and are frequently omitted or misused, in actual work. The present article is concerned mainly with these two stages.

Validity of the Gaussian error curve.—All least squares' calculations of probable error assume that the residuals follow a Gaussian error curve. This curve is derived from a consideration only of *accidental* errors. Probable errors are, however, evaluated frequently in cases where constant or systematic errors are known to be present. Such a procedure, when used judiciously, is believed by the writer to be better than any alternative procedure, but the results are naturally less reliable than a strict reliance on theory would indicate. The statement is sometimes made that in practise one often gets more *large* residuals than are predicted by theory. This point is tested by 500 measurements of a spectral line, and the resulting Gaussian curve, plotted in Fig. 1, shows no indication of such a deviation. It is possible to account for an excess of large residuals (with the necessary accompaniment of a deficiency of small residuals), by assuming that the various observations used were not in fact of equal reliability. The formulas by which the probable error r may be calculated from the observed residuals, and the reliability of each such value, are briefly considered.

Internal versus external consistency.—Probable errors are calculated usually on the basis of internal consistency, and most texts discuss only this method. Scarborough, in a recent article, claims that there is no logical basis for the calculation from external consistency. This matter is considered in detail, and it is shown that the two methods must necessarily lead to the same result, except for statistical fluctuations, provided that only accidental errors are present. Formulas for the magnitude of the expected fluctuations are given. It is shown that a probable error based on internal consistency (r_i) is virtually a *prediction* and that the probable error based on external consistency (r_e) is the answer to this prediction. When the ratio r_e/r_i exceeds unity by an amount much greater than is to be expected on the basis of statistical fluctuation, one has almost certain evidence of the presence of systematic errors. In such a case new arbitrary weights should be assigned. Then one has available *only* external consistency as a basis for the calculation of errors. The false deductions that result from a failure to note the above facts are illustrated by several examples from the literature, and a numerical problem illustrating all of these relations is presented in detail.

Probable error of a function evaluated by least squares.—The probable error of a function of directly observed quantities is given by the well-known law of "propagation of Error." When, however, one has a function whose coefficients have been evaluated by least squares, the probable error of the function is scarcely mentioned in the literature, and the writer has never seen it calculated in practise. Explicit formulas for the error of a rational integral function of any degree are derived by the writer, from elementary considerations, and are found to agree with the very general formula already known. Practically all of the results presented here find frequent application in the article on probable values of e, h, etc., which immediately follows.

Introduction

THE method of least squares furnishes a precise and reliable process of calculation in much scientific work. That its value is not more generally appreciated seems to be due to the fact that it is often misused or only incompletely used, with resulting discredit to the system as a whole. There are three possible stages in any least squares' calculation, (1) the calculation of the "most probable" values of certain quantities, from a given set of experimental data, (2) the calculation of the "probable error" of each of the quantities just evaluated, (3) the calculation of the reliability, or probable error, of the probable errors so evaluated.

The proper procedure for the first stage is covered quite adequately in almost any good text on the subject. The second stage is covered much less adequately, and certain formulas that it appears to the writer should find frequent use in scientific work have apparently never been stated explicitly and seem never to have been used. The general theory covering these two stages has, however, been known for a long time, so that the working out of explicit formulas is a matter of mere detail. In the case of the third stage, the theory is very incomplete and rather uncertain, even in its simpler aspects.

The present article is concerned chiefly with the second and third stages of the calculation, since most of the misapplication of least squares in the literature has been in connection with these matters. Least squares is a branch of statistical theory and it would seem that there is in statistics a certain elusive character that makes all deductions difficult to state clearly. Certainly the writer must confess that he has consulted practically every available text on least squares, and yet has had considerable difficulty in understanding many points. The numerous instances in the literature of the injudicious or incorrect use of least squares seem to indicate that others have had similar difficulties. This must be the excuse for the discussions to follow, which, at first glance, may appear to be presented in quite unnecessary detail. Most of the facts may be found in any good text, but some of the formulas and points of view are, I believe, new. My object has been to bring out what may crudely be termed the "physics" of the matter, rather than the mathematics. In fact practically no mathematical proofs are given, since these may be found in any number of texts.

The direct incentive to the writing of this article has been two-fold, (1) the publication of a paper by Scarborough[1] entitled "The Invalidity of a Commonly Used Method for Computing a Certain Probable Error," and, (2) the need of certain formulas in my paper[2] on "Probable Values of e, h, e/m and α," which follows the present article. Scarborough's article considers a matter that is ignored in practically all texts, but his main conclusion is unfortunately quite incorrect. In my own article I have had occasion to use formulas that, so far as I know, are not stated in convenient working form

[1] J. B. Scarborough, Proc. Nat. Acad. Sci. **15**, 665 (1929). This article is given, practically verbatim, in his book "Numerical Mathematical Analysis" (Johns Hopkins Press, 1930) pp. 328–332.

[2] This paper on e, h, etc. will hereafter be denoted as G.C. 1932.

anywhere in the literature. The present paper contains the derivation of these formulas, and G.C. 1932, which follows, may be considered as an illustration of nearly every point discussed here.

The Gaussian Error Curve

The first important fact concerning the method of least squares is that it applies strictly only to sets of measurements whose errors follow the Gaussian error curve. It may be proved that this law should be followed if the errors are purely accidental (not systematic or constant), are equally liable to be positive or negative, and are more likely to be small than large, with very large errors entirely lacking.[3] Taken too literally, this requirement would practically eliminate least squares as a method of computation, for in nearly every set of measurements there is presumably a certain amount of systematic or constant error. The probable error calculated by least squares takes account only of accidental errors, but it is always possible to enlarge this probable error to make what seems proper provision for other types of errors. Let us suppose that a given constant has been measured in several different ways, and that the stated probable error of each result makes allowance for the various possible sources of error. If now there is no reason to believe that the errors in one investigation are related, in sign or magnitude, with those of another investigation, it seems quite legitimate to combine by means of least squares the various results, i.e., to weight them according to their stated probable errors, and to derive the probable error in the final weighted average by the usual formulas. This is a very important point which will appear later in the paper and to which some readers may take exception. We assume, in essence, that the constant or systematic errors occurring in a series of unrelated investigations also follow a Gaussian error curve. It must be admitted that in practise one is very unlikely to meet such an ideal distribution of constant errors, and for that reason the final result and its probable error must be considered *less* reliable than the theory of least squares would indicate. What I do maintain is that such a procedure has more to recommend it than any alternative procedure and that, if used with proper discretion, it will result in no serious error. One may summarize the situation by saying that, except in especially favorable cases, least squares' results, and their computed probable errors, are *not* as reliable as indicated by theory. This fact seems unfortunately so well known that many persons have chosen to use other methods (or no method at all!) for calculating their results. I use the word "unfortunately" because these alternative methods are, without exception, inferior to least squares. Certainly one *must* make computations, in scientific work, and the fact that a certain system of computation is not as reliable as over-zealous advocates may claim is no excuse for using in its place a still less reliable system.

[3] See, for instance, M. Merriman "Method of Least Squares" 8th edition, articles 17–28, or Whittaker and Robinson "Calculus of Observations" pp. 168–173. The validity of the proposed proofs of the Gaussian error curve has been the subject of endless discussion, but this article is not concerned with the philosophical aspects of least squares.

The equation of the Gaussian error curve is

$$y = \frac{h}{(\pi)^{1/2}} e^{-h^2x^2} \tag{1}$$

where y is the proportional number of observations having an error x, and h is the so-called precision index. The curve applies strictly only to errors, but these are never known, since the true magnitude of the measured quantity is also not known. All that we know is the most probable magnitude of the quantity, or what is often termed its adjusted value, and the deviation of any measurement from this is called the residual v. If a sufficient number of measurements are available to make at all possible the plotting of an error curve, the residuals should follow the same equation and lead to essentially

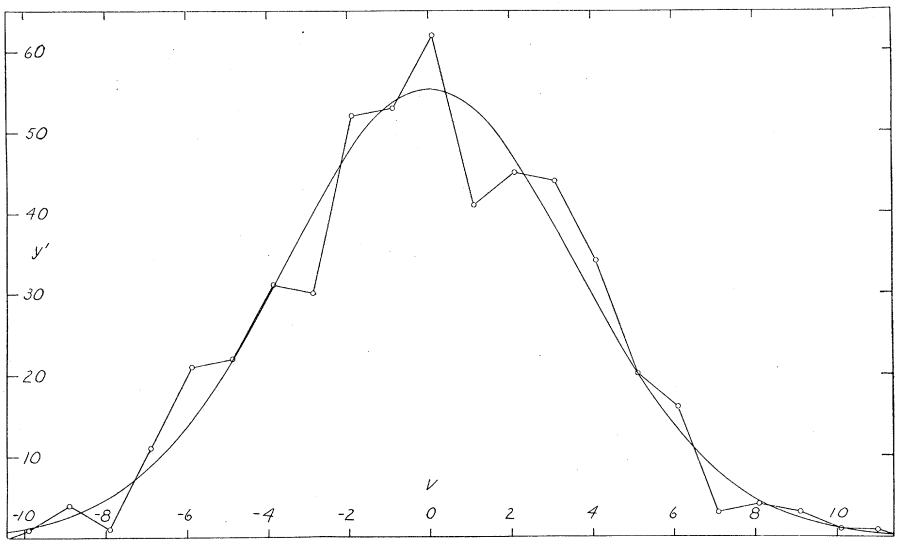

Fig. 1. The distribution of residuals for 500 measurements of a spectral line, compared with the Gaussian error curve evaluated by second moments. Abscissa represents residual v, in 0.001 mm units. Ordinate represents the number of residuals of magnitude v.

the same calculated value of h. In other words, the difference in the results given by errors and by residuals should be negligibly small, and in the following discussion this is assumed to be the case. Now in most scientific work only a comparatively small number of observations are taken, and it is manifestly impossible to make a direct test of the distribution of the residuals. Yet least squares' formulas are applied to these small sets of observations. In such cases the following implicit assumption is made. One assumes, from his own previous experience or from that of others, that if the number of observations were greatly extended, with no change in conditions, it would be found that this large set agreed with the Gaussian error curve.

The statement is sometimes made[4] that in such extensive sets of measure-

[4] See, for instance, Tuttle and Satterly "Theory of Measurements" p. 175. The Frontispiece of G. C. Comstock "Method of Least Squares" shows four error curves, and three of these

ments, one generally finds more large residuals than are predicted by theory. In order to test this statement, I have made a series of 500 cross-hair settings on a very wide[5] but symmetrical solar spectrum line, under conditions as favorable as possible to equal reliability for all observations. Readings were made to the nearest 0.001 mm, and the arithmetic average was adopted as the most probable value. The residuals calculated with respect to this average are scattered over 12 positive and 10 negative values. The resulting distribution is shown in Fig. 1, in which the unit abscissa is 0.001 mm and the plotted ordinate is y' $(=500y)$. y' thus represents the actual number of residuals of a given magnitude.

In fitting these points to Eq. (1), the one adjustable constant h is calculated by moments, normally by second moments.[6] From the definition of y we have as alternative expressions for the second moment of the points, $\Sigma yx^2=\overline{x^2}\cdot\Sigma y=\overline{x^2}=\Sigma x^2/n$, where the last summation is to be extended over the entire n original observations. The second moment of the calculated curve, $\int_{-\infty}^{+\infty}yx^2dx$ equals [7] $1/2h^2$. Hence, by equating the two second moments we obtain

$$\frac{1}{h}=\left(\frac{2\ \Sigma x^2}{n}\right)^{1/2}. \tag{2}$$

If the errors x are replaced by the residuals v, Eq. (2) becomes[8]

$$\frac{1}{h}=\left(\frac{2\ \Sigma v^2}{n-q}\right)^{1/2} \tag{3}$$

where q is the number of constants simultaneously evaluated by the least squares' solution. In the case of a series of observations of a single quantity, $q=1$.

---

do show a slight indication of this excess of large residuals. On the other hand, Merriman, reference 3, page 33, presents 300 astronomical observations that agree very well with the calculated error curve, and he remarks "Whatever may be thought of the theoretical deductions of the law of probability of error, there can be no doubt but that its practical demonstration by experience is entirely satisfactory". Scarborough "Numerical Mathematical Analysis", page 314, gives 470 astronomical observations that show an equally good agreement with theory.

[5] The width was achieved in part by the use of high magnification in the comparator telescope.

[6] The term second moment comes from the close analogy to mechanics. (See T. C. Fry, "Probability and Its Engineering Uses", p. 183). If all the observed points in the negative quadrant of Fig. 1 are folded over onto the positive quadrant, and each point is given a mass equal to its y value, then the second moment of the observations, Σyx^2, is merely their moment of inertia with respect to the y axis. Similarly, if the calculated curve encloses a flat body of constant mass per unit area, such that the total mass of this body equals the total mass of the points, then $\int yx^2dx$ represents its moment of inertia. In the calculation of h by second moments one equates $\int yx^2dx$ to Σyx^2 and thus effectively makes the two moments of inertia equal. Since the masses are necessarily equal, one may say that the "radius of gyration" of the calculated curve is equated to the radius of gyration of the points. Similarly, in the case of *first* moments, the "lever arm" of the force moment of the points is equated to the corresponding lever arm of the calculated curve.

[7] See Merriman, reference 3, page 73.

[8] See Merriman, reference 3, pp. 79–82.

The probable error r of one observation (i.e., the value of $|x|$ that divides the Gaussian curve into two equal areas) is related to h by $r=0.4769/h$, so that

$$r = 0.6745\left(\frac{\Sigma v^2}{n-q}\right)^{1/2}. \tag{4}$$

If h is calculated by first moments (by equating $\int yx dx$ to Σyx), one obtains

$$r = 0.8453\frac{\Sigma|x|}{n} \tag{5}$$

which is known as Peter's formula.

Gauss was able to show[9] that the value of h, or of r, when calculated by second moments is more reliable than that calculated by moments of any other degree. His general formula gives for the proportional probable error in h, or in r, when obtained by second moments,

$$u = \frac{0.4769}{(n)^{1/2}}. \tag{6}$$

If r is calculated by first moments, the numerator in Eq. (6) becomes 0.5096, if by third moments 0.497, etc.

This seems to be the extent of the theory, so far as concerns the third stage of the calculations, mentioned in the introduction. The derivation of Eq. (6) was made by Gauss in connection with the curve of *errors*, and it is still uncertain[10] whether n or $n-1$ should be used in actual calculations with residuals. It seems to me that when one is fitting a curve containing q constants, so that the probable error of an observation of unit weight[11] is given by Eq. (4), one should use $n-q$ in place of n, in Eq. (6). I am not, however, certain about this, and I shall therefore continue to use Eq. (6), as I have done in the past. It seems safe to conclude, in any case, that the value of u, aside from other sources of uncertainty, is unreliable to the amount calculated by replacing n by $n-q$ in Eq. (6).

The smooth curve drawn in Fig. 1 corresponds to $h=0.1965$, the value calculated from second moments. By Eq. (6) its uncertainty is 2.13 percent. The value of h from first moments is 0.1959, a change of only 0.3 per cent, and the reliability of this latter value is 2.28 percent. These two methods for computing h thus yield almost identical results, in this case, and always have almost equal reliability. The two calculated curves agree too closely to be conveniently shown together in Fig. 1.

This figure certainly gives no evidence of an excess of large residuals, and in fact the entire course of the observations follows the Gaussian curve with

[9] See Whittaker and Robinson, reference 3, p. 201.

[10] See Merriman, reference 3, p. 208.

[11] The question of weighting, and the meaning of "unit weight" is considered later in this paper.

surprising accuracy.[12] The one noticeably "low" point at $v=+1.1$, bordered by the "high" point at $v=+0.1$ is in all probability due to a slight irregularity of the comparator screw, which causes one setting to be favored at the expense of an adjacent setting. This incidentally furnishes a good illustration of the fact that errors, other than accidental, are *always* present to greater or less degree, in actual scientific work. I believe that the apparent excess of large residuals sometimes noted in practise is due usually to the unsuspected presence of observations of unequal reliability. If a portion of the observations has actually one precision index h, and another portion has another, it is easy to show that a single plot of all the residuals *should* exhibit an apparent excess of large residuals. Now the area under the observed curve[13] and under the calculated Gaussian error curve is necessarily the same (unity for each). Hence it follows that any observed excess of large residuals must be compensated by a numerically equal deficiency of small residuals. In the case of the large residuals, however, the *proportional* excess is large, while the proportional deficiency of small residuals is small. Thus the observed excess stands out, while the correlated deficiency may easily pass unnoticed.[14]

Internal Versus External Consistency

Scarborough's article[1] is concerned with the relative merits of internal and external consistency, in the calculation of probable errors, although he does not employ these particular designations. This is a matter that is very inadequately treated in text-books, and concerning which there appears to be wide-spread misunderstanding. It is accordingly presented in considerable detail.

Let us assume that we have made n direct observations, each of the

[12] N. R. Campbell, on page 182 of his book "Measurement and Calculation", remarks concerning published distributions of residuals, "The examples offered are never typical of those occurring in physics (of other sciences, such as astronomy or geodesy, I say nothing). In my experience, the residuals in physical examples usually depart notably from the law, the curve representing them having a flatter top than the Gaussian. This is what our theory would predict; for there is no reason to believe that all errors considerably less than the maximum have not equal chances."

The last sentence is quite correct. By *sufficiently* restricting the range of x, the Gaussian error curve is itself essentially "flat". The data plotted in Fig. 1 are typical of the sort of work on which the writer is engaged. Whether or not it is physics may be left to the reader. In this connection it should be noted that Campbell advocates, in place of least squares, the "method of averages," discovered by Tobias Mayer in 1748 and in common use by statisticians at the present time (see Whittaker and Robinson, reference 3, pp. 258–259, H. L. Rietz "Handbook of Mathematical Statistics", p. 64, and Scarborough, reference 1, pp. 357–363). This method was "rediscovered" by Campbell in 1920 (N. R. Campbell, Phil. Mag. **39**, 177, 1920), and named by him the "method of zero sum", a name he still employs (Phil. Mag. **10**, 745, 1930). In 1924 he remarked (Phil. Mag. **47**, 816, 1924), "so far as I know, the method is not habitually used by any one but myself".

[13] If, as in Fig. 1, the observed points are connected by straight lines, ending on the v-axis as shown, the area under this series of lines (the "observed" curve) is n, if y' is plotted as in Fig. 1, or unity if the proportional number of residuals y is plotted. The total area under the calculated Gaussian curve, Eq. (1), is of course unity.

[14] This is very well illustrated by the curves in Comstock's text, mentioned in footnote 4.

same reliability, of a certain quantity, and have calculated the arithmetic average and the deviations (residuals) from this average. Then, by Eq. (4), the probable error in any one observation is

$$r = F\left(\frac{\sum v^2}{n-1}\right)^{1/2} \tag{7}$$

where[15] $F = 0.6745$. This equation is essentially a statement of certain facts. As has been brought out in the preceding discussion, it is assumed that the residuals actually follow a Gaussian error curve, or would do so if the observations were sufficiently numerous. From the observed distribution of the residuals we then deduce, by means of Eq. (7), that particular residual r that divides the area under the Gaussian curve into two equal parts, since this equation expresses the most reliable relation between r and the residuals v.

From this observed value of the probable error of a single observation, we now proceed to predict the probable error in the arithmetic average. This is done by applying the general equation of the Propagation of Errors (see Eqs. (24) and (25) ahead). The result is

$$(R)_{int} = F\left(\frac{\sum v^2}{(n-1)n}\right)^{1/2}. \tag{8}$$

I think that the most important requirement in the present discussion is a recognition of the essential difference between Eqs. (7) and (8). In the case of Eq. (7) we have a number of residuals, and as just indicated, this equation is merely a compact and useful expression for the distribution of these residuals. We have, however, only *one* arithmetic average and Eq. (8) is therefore a direct prediction. It states in effect the following. Let us imagine that instead of *one* set of n observations of a certain constant reliability, we have N such sets, each composed of n observations of this same reliability. We may then calculate the N arithmetic averages, which will hereafter be called *points*, and these points, just as in the case of the n observations of any set, should disagree more or less with one another. If the arithmetic average of the set of points is calculated, and also the deviations V of each point from the average, this new set of residuals should in turn follow a Gaussian error curve, for which the probable error is[16]

[15] This nomenclature for the factor 0.6745 is introduced merely to save type-setting.

[16] A great deal of my own earlier confusion regarding internal and external consistency has been due to ambiguous nomenclature. To get the facts clearly in mind, it seems necessary to make a sharp distinction between the *points* on a curve and the *observations* of which each point is the average. To aid in drawing this distinction, I have used lower case type for all letters connected with the observations (such as v, n, r) and capitals for letters concerned with points. In the next section of this paper, and in general work, no such distinction seems necessary, and accordingly in G.C. 1932 and elsewhere I use the customary lower case letters for both points and observations.

$$(R)_{ext} = F\left(\frac{\sum V^2}{N-1}\right)^{1/2}. \tag{9}$$

Eq. (9) is thus the *answer* (actual or hypothetical) to the *prediction* made by Eq. (8). $(R)_{int}$ is the probable error of a point, as calculated by *internal* consistency, and $(R)_{ext}$ is the probable error of this same point, or any other of the N points, as calculated by *external* consistency.

If now the experimental conditions are such as to justify the application of the theory of least squares, on which the predicted value $(R)_{int}$ is based, the two values $(R)_{ext}$ and $(R)_{int}$ should agree, *except for statistical fluctuations*.[17] In the simplified situation assumed here, it is easy to show just what the equality of $(R)_{ext}$ and $(R)_{int}$ indicates.[18] In Eq. (8) $\Sigma v^2/(n-1)$ actually represents a certain *average* value of e^2, where e is the true error of an observation,[19] i.e.,

$$\frac{\sum v^2}{n-1} = \overline{e^2}. \tag{10}$$

Similarly, in Eq. (9),

$$\frac{\sum V^2}{N-1} = \overline{E^2}. \tag{11}$$

Let us now form the ratio of $(R)_{ext}$ to $(R)_{int}$. From Eqs. (8), (9), (10) and (11) we get

$$\frac{(R)_{ext}}{(R)_{int}} = \left(\frac{\overline{E^2}}{\overline{e^2}/n}\right)^{1/2}. \tag{12}$$

Hence the statement that $(R)_{ext}$ should equal $(R)_{int}$, except for statistical fluctuations, is merely the statement that the root-mean-square average of the errors of the N points should be only $1/(n)^{1/2}$ as large as the root-mean-square average of the errors of the n observations of any set. But this is exactly the relation predicted by Propagation of Errors, and already used in deriving Eq. (8) from Eq. (7).

Before discussing the expected variation from unity of the $(R)_{ext}/(R)_{int}$ ratio, let us consider a more general situation. We assume now that the N sets of observations from which the N points are calculated, are formed of varying numbers n_i of individual observations, and that the reliability of the observations also varies from one set to another. In that case we calcu-

[17] The expected magnitude of such fluctuations will be considered presently.

[18] A. Palmer "Theory of Measurements" pp. 66–71, gives a more abstract discussion of the equality of these two probable errors. This is the only reference in English to this important question that seems worthy of quotation.

[19] This has been pointed out in connection with Eqs. (2) and (3). Theory gives directly the value of r (or of h) in terms of the errors e. The theory further indicates that if the observable Σv^2 is substituted for the hypothetical Σe^2, the *most probable* relation is $\Sigma e^2/n = \Sigma v^2/(n-1)$ in the case of one unknown quantity, or $\Sigma v^2/(n-q)$ in the case of q unknowns.

late the probable error of *each* point by Eq. (8), and denote any representative point by $A_k \pm R_k$, where

$$R_k = F\left(\frac{\sum v^2}{(n_k - 1)n_k}\right)^{1/2}. \tag{13}$$

This, it should be noted, is the probable error of the point, by *internal* consistency, and will henceforth be called $(R_k)_{int}$.

We next desire to combine together the N points, and it is now not necessary to assume that these N points are measurements of one single quantity. We shall, instead, assume that the N points are connected by some functional relation, i.e., that they are the values of y, for a certain set of N values of x, in $y=f(x)$. In the previous example, the N points lay on the simplest of all curves, the horizontal line $y = constant$.

Before proceeding to combine our N points, we must first *weight* each point according to its probable error, $(R_k)_{int}$, assuming that this probable error is confined *solely* to the measurement of the y coordinate.[20] Now the probable error of a measured quantity may be termed an absolute measure of its reliability. The weight of such a quantity, on the other hand, is only a relative measure of its reliability. Moreover, the weight P is defined in such a manner that it is not directly proportional to the probable error R but is connected with it by the following relation,[21]

$$R = \left(\frac{C}{P}\right)^{1/2} \quad \text{or} \quad P = \frac{C}{R^2} \tag{14}$$

where C is an arbitrary constant. Furthermore, in Eq. (13), the reciprocal of the term $\Sigma v^2/(n_k-1)$ may be termed a measure of the reliability of a single observation, as shown in connection with Eq. (10), while n_k measures the number of observations. Hence for observations of equal reliability, the probable error of the arithmetic average varies as $1/(n)^{1/2}$, and by Eq. (14), the probably error varies also as $1/(P)^{1/2}$. Thus the weight of the arithmetic

[20] This assumption of zero error (or infinite weight) in the x coordinate is made implicitly in the case of every equation presented in this article. Fortunately the condition is actually fulfilled in many of the measurements of physical science, such as the plotting of spectral frequencies against the values of a quantum number. Moreover, when all observations have the same probable error in the y coordinate, and a similarly constant error in the x coordinate, it is practically immaterial how these errors compare with one another. In particular, one may assume (1) zero error in the x coordinate, and any constant probable error in the y coordinate, or (2) zero error in the y coordinate and any constant error in the x coordinate. I have found, with typical sets of data conforming to a linear equation, that the values of the constants given by solutions (1) and (2) differ by only about one-fifth of the probable errors of the constants. Such differences are of course devoid of any real significance. On the other hand, if the probable errors for either the x or the y coordinates are not constant, solutions (1) and (2) may be appreciably different. In the case of a non-linear equation, such a difference is likely to exist whether the probable errors are constant or not. I am indebted to Dr. W. E. Deming for the foregoing general conclusions, which he has obtained in the course of a private correspondence on this subject. The basic equations on which they rest appear in his paper, Phil. Mag. **11**, 146 (1931).

[21] Palmer, reference 18, article 44. Merriman, reference 3, articles 41–43, 61 and 63.

average of n equally reliable observations is proportional to n, and if each observation is arbitrarily assumed to have unit weight, then the weight of the arithmetic average is said to be n. This is the real origin of the idea of weight.

The most explicit expression for the weight P is obtained by substituting for R in Eq. (14) its value from Eq. (13). We then get

$$P = \frac{C}{F^2}\left(\frac{n-1}{\sum v^2}\right)n. \tag{15}$$

This expression states that the weight of an arithmetic average is directly proportional to the number of observations composing it, and also to the reliability of the individual observations.

Thus, from Eqs. (13) and (14), we may calculate a weight P_k for each point from its (internal consistency) probable error $(R_k)_{int}$. Now in many actual situations we do not know the details of the derivation of the values $(R_k)_{int}$. Furthermore it is possible, as discussed earlier, that this probable error is an estimate that includes not only the effect of accidental errors, Eq. (13), but also of other possible types of error. In that case $(R_k)_{int}$ is merely a stated quantity, and is connected with other quantities of interest to us only by Eq. (14).

We shall denote by R_i the probable error of a "point of unit weight."[22] Then by Eq. (14)

$$R_i = (C)^{1/2}. \tag{16}$$

Hence

$$(R_k)_{int} = \left(\frac{C}{P_k}\right)^{1/2} = \frac{R_i}{(P_k)^{1/2}}. \tag{17}$$

Just as in the previous illustration, the value $(R_k)_{int}$ is one based on certain internal evidence and is essentially a prediction. To check this prediction we proceed to fit the N points to the known functional relation[23] $y=f(x)$. This is to be carried out by means of the appropriate least squares' formulas, that is, the q undetermined constants of the function are to be thus evaluated. One then calculates values of y and compares these with the observed values, thus obtaining a set of residuals V. Then the probable error of any point, of weight P_k, as based on *external* consistency, is given by

$$(R_k)_{ext} = F\left(\frac{\sum PV^2}{(N-q)P_k}\right)^{1/2}. \tag{18}$$

[22] This expression may easily be very confusing. It does not refer to any actual point. The constant C is purely arbitrary, and R_i is merely its square root. The magnitude of R_k for any actual point is of course *not* arbitrary. From the real magnitudes R_k and the arbitrary constant C, or R_i, one derives the weights P_k whose absolute values are arbitrary, but whose relative values are not.

[23] In order that our comparison of the values of R_k, based on internal and on external consistency, may be valid, it is necessary that the points actually be connected by the assumed functional relation. Sometimes, as in the case where the points are merely repeated measurements of a given quantity, there is no doubt as to the function. In other cases there may be considerable uncertainty on this matter.

It is convenient, in analogy with Eq. (16), to define the probable error, based on external consistency, of a point of unit weight. This, by Eq. (18) is

$$R_e = F\left(\frac{\sum PV^2}{(N-q)}\right)^{1/2}. \tag{19}$$

Eq. (18) can then be rewritten as

$$(R_k)_{ext} = \frac{R_e}{(P_k)^{1/2}}. \tag{18'}$$

It is possible to calculate also the probable errors of the q constants whose magnitudes have been simultaneously evaluated by the least squares' solution, and of the function itself. To do this we need only the *weight* of each of these quantities, and formulas for these weights will be given later. We are here, however, interested primarily in a comparison of external and internal consistency, and in making such a comparison it is immaterial whether we discuss some particular point on the curve, a hypothetical point of unit weight, one of the constants of the function, or the function itself. This is due to the fact that, as shown by Eqs. (17) and (18'), the appropriate weight of any one of these quantities enters in the same manner into both the external and the internal consistency error, and therefore cancels out when the ratio is taken. We shall, for convenience, consider the hypothetical point of unit weight. Then from Eqs. (16) and (19)

$$\frac{R_e}{R_i} = F\left(\frac{\sum PV^2}{(N-q)C}\right)^{1/2}. \tag{20}$$

This ratio, just as in the case of Eq. (12), should equal unity, except for expected statistical fluctuations. Eq. (20) appears, at first glance, to be quite different in form from Eq. (12), but it is actually only a more general expression.[24]

Since R_i by Eq. (16) is only an arbitrary constant, the uncertainty in the ratio R_e/R_i depends wholly on the uncertainty in R_e. Now the uncertainty in the calculation of the probable error of N points (or observations) is given by Eq. (6).[25] Thus if $N=9$, we may expect that the ratio R_e/R_i of Eq. (20) has itself a probable error of about 16 percent, and therefore we may expect a deviation from unity of this order of magnitude.

[24] This may be seen more clearly by substituting for each P, in the ΣPV^2, its value C/R^2 from Eq. (17). The arbitrary C then cancels out of Eq. (20) and we have under the radical the various $(R_k)_{int}$ which, by Eq. (13), presumably involve the residuals v of the original observations. The right side of Eq. (20) equated to unity, then expresses the general expected relation between the magnitude of the residuals (or errors) of the various sets of observations forming the points, and the magnitude of the residuals (or errors) of the various points, just as Eq. (12) does in one very special case.

[25] The residuals of the N points, from the calculated $f(x)$, should of course follow the Gaussian error curve, regardless of the particular function. The n observations of a single quantity, in connection with which Eq. (6) has been stated, satisfy the particular function $y = constant$.

Let us suppose now that, in a certain situation, the actually calculated ratio differs from unity by several times—say five times—the probable error of Eq. (6). There is approximately only one chance in a thousand that such a deviation is the result of mere statistical fluctuation. We therefore conclude that a definite discrepancy exists between the results based on external and on internal consistency. This discrepancy can be due to only two possible causes, (1) the functional relation to which the points have been fitted is not the true one, (2) the probable errors that have been assigned to at least a portion of the points are false.

The present discussion is not concerned with (1), and we here merely assume that the true functional relation is known. Cause (2) is due, nearly always, to the presence of unsuspected constant or systematic errors in a portion or in all of the points. The evaluation of the ratio R_e/R_i is thus an invaluable aid to the detection of such errors. If now it has been proved, from the value of this ratio, that the stated probable errors of the points are false, then the weights P_k, based on these assigned errors, are entirely misleading. In such an annoying situation, the logical procedure is to discard this set of weights, and to make an arbitrary assignment. Such an assignment is necessarily based on the personal judgment of the computer, and may be very uncertain. It is, however, likely to be better than an assignment that has been proved false. Using the new set of weights, one then makes a new calculation of the constants in $f(x)$, derives a new set of residuals V, and with these calculates a new set of probable errors by means of Eq. (18). In this situation one has available *only* external consistency for the calculation of the errors.

If the ratio R_e/R_i does not differ from unity by more than the expected amount, one may use *either* R_i or R_e for the calculation of errors, i.e., either Eq. (17) or Eq. (18′), assuming P_k to represent the appropriate weight of the quantity discussed. In the following article, G.C. 1932, I have adopted the conservative policy of using, in such a situation, the *larger* of the two quantities R_e and R_i. One often encounters the situation in which the number of points N is relatively small. The probable error in R_e is then large, and most writers consequently advocate using R_i. In fact the probable error based on internal consistency is always more reliable than one based on external consistency.[26] An additional advantage in the use of $(R_k)_{int}$ is that it can be calculated far more rapidly than $(R_k)_{ext}$. It must be remembered, however, that this conclusion as to the relative reliability of internal and external consistency depends wholly on the correctness of the assumption that only accidental errors are present, and I again emphasize the point that the

[26] The reliability of any probable error $(R_k)_{int}$ must be judged from its origin. If this probable error is given by Eq. (13), then its reliability is given by Eq. (6) with $n=n_k$. This is also the reliability of $(P_k)^{1/2}$ since in Eq. (17) R_i is merely a constant. On the other hand, the reliability of $(R_k)_{ext}$ is affected not only by $(P_k)^{1/2}$ in Eq. (18), but also by ΣPV^2. The ratio $(R_k)_{ext}/(R_k)_{int}$ is affected merely by the uncertainty in ΣPV^2, and this is given by Eq. (6) with $n=N$. Hence $(R_k)_{ext}$ is always less reliable than $(R_k)_{int}$ and it is apparently for this reason that $(R_k)_{int}$ is commonly employed in the literature.

sole reason for calculating $(R_k)_{ext}$, in addition to $(R_k)_{int}$, is to test this assumption which, unfortunately, will very often be found quite untenable.

The final conclusion of this section is that in the situation where no appreciable systematic errors are present, as shown by a satisfactory value of the ratio R_e/R_i, one may use either external or internal consistency as a basis for the calculation of probable errors. The internal consistency value should then be the more reliable, and is in common use. I advocate, however, as a more conservative policy, the use of the larger of these two errors. On the other hand, when the ratio is not satisfactory, one must make an arbitrary assignment of weights, and with such weighting external consistency furnishes the only basis for the calculation of errors.

The writer hopes that the preceding remarks have made clear to the reader the true situation. The necessity for such a discussion may be emphasized by a few examples of the misstatements and misapplications of this matter appearing in the literature.

(1) Scarborough's article[1] is devoted to a proof of the fact that a probable error based on internal consistency varies directly with the probable errors of the original observations, whereas a probable error based on external consistency involves only ratios of these probable errors, and so is independent of their absolute magnitudes. From these facts he draws the wholly irrelevant deduction that the first method gives a correct result, but the second does not. He states, *in italics* (page 667), that $(R)_{ext}$ is simply a measure of the agreement of the points among themselves, and nothing more. After the preceding discussion it is perhaps unnecessary to remark that such agreement, or better, lack of agreement, constitutes the sole mathematical basis for the calculation of errors. The true relation between the two methods is given most simply in Eq. (12). The first method uses the agreement of the individual observations determining each point to predict what agreement will be found among the various points. The second method gives the actual agreement of the points among themselves, and is therefore an answer to the prediction. Both methods are based on the same theory, and are equally correct in principle.

(2) In Leland's text "Practical Least Squares," page 189, there is given an example of the calculation of a weighted average and of its probable error, from three observations of an angle for each of which the probable error is given. The data are

$$72^\circ\ 47'\ 43.18'' \pm 0.06''$$
$$44.01'' \pm 0.10''$$
$$43.74'' \pm 0.08''.$$

Now a mere glance at these data indicates that they show a much greater variation than is to be expected from the stated probable errors, so that unsuspected errors of some kind are undoubtedly present. Yet Leland proceeds to calculate the weighted average, and its probable error, on the basis of internal consistency. The result is $43.50'' \pm 0.04''$, and he gives this as the

final best answer. I find that $R_e/R_i = 3.81$, and since $n = 3$, the calculated probable error of the ratio is about 27 percent. The observed deviation from unity is thus over ten times the probable error and there is not one chance in a million of this occurring as a result of purely statistical fluctuations.

In his general discussion Leland (loc. cit. pp. 188–189) recognizes the fact that R_e may not equal R_i, but he apparently assumes that in every such case, R_i will give a more reliable result than R_e. We have seen, however, that this is not the case. On the contrary, we can form a mathematical estimate of the expected discrepancy and use this for the detection of previously unsuspected errors. In the above example one should obviously discard the stated probable errors. If, merely for convenience, the same relative weighting is retained, the correct result is $43.50'' \pm 0.16''$, as deduced from external consistency.

(3) In F. W. Clarke's monumental work on the calculation of atomic weights,[27] internal consistency is used exclusively in the calculation of probable errors and in the subsequent weighting. Now until very recently the investigation of atomic weight ratios formed a most fertile field for unsuspected constant and systematic errors. A brief examination of the data presented by Clarke is sufficient to establish this fact. I find, from sample calculations, that the ratio R_e/R_i averages about ten, so that Clarke's stated probable errors average about one-tenth of the most probable values. In certain cases such a system of analysis leads to a clearly false result for the atomic weight itself, as I have pointed out in a previous paper.[28] Thus there is occasionally an atomic weight determination by some particular investigator that is quite at variance with all other known results, but that happens to have high internal consistency. Clarke accordingly gives it a high weight, and this weight carries through to the final result, so that the investigation in question, which should have been discarded entirely, produces an appreciable change in the published final result. All of the recent reports of atomic weight committees seem to recognize the fact that the older determinations are nearly all vitiated by constant errors, and as a result the committee makes an arbitrary assignment of weights. To speak bluntly, it gives zero weight to these older determinations, regardless of their apparent probable errors. This same question arises in the following article, G.C. 1932, and in it I have adopted the same policy.

In concluding this section I give a numerical example of the calculation of errors by internal and by external consistency. For this example a situation is chosen that will incidentally bring out a point regarding weights that appeared implicitly in the preceding discussion.

Merriman[29] lists 24 observations of an angle, measured by the U. S.

[27] F. W. Clarke "A Redetermination of Atomic Weights" 4th edition (1920). Printed as Part 3, Vol. 16 of Memoirs, Nat. Acad. Science.

[28] R. T. Birge "Probable Values of the General Physical Constants," Phys. Rev. Supplement (now called Reviews of Modern Physics) **1**, 1, 1929. See pages 25 and 19–20. This article is referred to as *G.C.* 1929.

[29] M. Merriman, reference 3, page 90.

Coast-Survey, and the data need not be re-copied. The degrees and minutes also are omitted, since these are unnecessary for our purpose. The observations are of presumed equal reliability, and the arithmetic average is therefore the most probable value. This result is $49.64'' \pm 0.28''$, where the probable error has been calculated by Eq. (8).

Let us now, merely for the purpose of this illustration, divide the observations, arranged in their original order, into four sets, each comprising six observations. We may then calculate the arithmetic average of each set, and by means of the residuals,[30] a probable error for each such average. We thus get the following four *points*

$$48.82'' \pm 0.64''$$
$$48.85'' \pm 0.43''$$
$$51.19'' \pm 0.32''$$
$$49.71'' \pm 0.60''$$

Since these four sets each comprise an equal number of observations, of equal reliability, the four probable errors should be equal, except for statistical fluctuations. We may test this as follows. Treating the four errors r_i as if they were four observations of a given quantity, one obtains from Eq. (7) the probable error r' in any one of the four

$$r' = F\left(\frac{\sum(r_i - r_{\text{average}})^2}{3}\right)^{1/2}. \tag{21}$$

The proportional magnitude of r' is thus found to be 19 percent. The predicted value is given by Eq. (6) with $n = 6$, and is 19.5 percent. This very close agreement is of course mainly fortuitous, but it is certainly an indication that the various observations are of equal reliability, and that the fluctuation of the values of r_i is only statistical.

In such a situation one should disregard the variation in r_i and give each of the four points the same weight. To give them unequal weights, on the basis of the unequal values of r_i is in this case just as illogical as to give a series of observations, of presumably constant reliability, a series of weights based on the deviation of each observation from some adopted average.[31] With equal weighting for the four points, one will of course get for the final average merely the first quoted result, $49.64''$.

We now, for the purpose of the illustration, assume that the origin of the four points and of their probable errors is not known, and we proceed to give

[30] It should be noted that a residual now means the deviation of any observation from the average of its own set, *not* from the general average just quoted.

[31] These facts seem to be well recognized. See, for instance, Wright and Hayford "Adjustment of Observations" pp. 74–77, and Tuttle and Satterly, reference 4, pp. 202–203. The former text refers specifically to the Northern Boundary Commission which decided that, when a long series of observations were taken with the same instrument, under comparable conditions, the average results were to be weighted according to the number of observations composing them, and not according to the calculated probable error.

weights (P) to the points according to their probable errors, and to calculate the weighted average, A. The weight of this average is ΣP and the probable error may now be calculated in the usual two ways. By internal consistency it is

$$(R)_{int} = \frac{R_i}{(\sum P)^{1/2}} \cdot \tag{22}$$

By external consistency it is

$$(R)_{ext} = F\left(\frac{\sum PV^2}{(N-1)\sum P}\right)^{1/2} \cdot \tag{23}$$

The numerical results are given in the top row of Table I.

TABLE I.

A	$(R)_{int}$	$(R)_{ext}$	ratio
50.09″	0.22″	0.43″	1.95
49.64″	0.27″	0.15″	0.55
49.38″	0.21″	0.32″	1.52
Average 49.70″	0.23″	0.30″	1.34

Let us now rearrange the observations, forming the first set of six from the 1, 5, 9, etc. observations, the second from 2, 6, 10 etc. We then obtain by a similar process the values in the second row of the table. For a third division I have used observations 1, 3, 5, 7, 9, 11 for the first set, 2, 4, 6, 8, 10, 12 for the second, 13, 15, 17 etc. for the third, and 14, 16 etc. for the fourth. The results are given in the third row, and finally the average of the three methods of division in the last row. The previously quoted result obtained by the direct unweighted average ($49.64'' \pm 0.28''$) is certainly the most reliable, and the table shows that external consistency happens, in this instance, to give for the probable error a slightly more reliable average value than does internal consistency. On the other hand, the fluctuations of $(R)_{ext}$ are greater than those of $(R)_{int}$, in agreement with our previous remarks on this subject. The expected probable error in the ratio, with $N=4$, is 24 percent. From the results given in the last column of the table, the average deviation of the ratio from unity is 64 percent. There is only one chance in fourteen for such a discrepancy, and it is barely possible that all the 24 observations do not have the same reliability, in spite of our previous result to the contrary. This belief is strengthened by the fact that one of the 24 original residuals is exceptionally large, and the entire set forms a rather poor Gaussian error curve. The agreement with theory is, however, on the whole about as good as one may expect from such a small number of observations, and I have deliberately taken for this illustrative problem such a small set, since this is the situation commonly experienced in scientific work. The results show that one should be cautious in drawing any definite conclusion as to the presence of systematic errors, unless the deviation from unity of the ratio R_e/R_i is so much greater than that to be expected that there is a very small probability (say 1 in a 100 or more) that such a deviation is due to pure chance.

The Probable Error of a Function

If one has a function Z of a number of independently observed quantities z_1, z_2, z_3 etc., whose probable errors r_1, r_2, r_3 etc. are known, then the probable error R in Z is given by the well-known formula[32] for the "Propagation of Errors,"

$$R^2 = \left(\frac{\partial Z}{\partial z_1} r_1\right)^2 + \left(\frac{\partial Z}{\partial z_2} r_2\right)^2 + \left(\frac{\partial Z}{\partial z_3} r_3\right)^2 + \text{etc.} \tag{24}$$

If Z is the arithmetic mean of n observed quantities $z_1, \cdots, z_n$, and if all the probable errors r_i are equal, this equation leads to

$$R = \frac{r}{(n)^{1/2}}, \tag{25}$$

a relation used in deriving Eq. (8) from Eq. (7). If $Z = z^n$, Eq. (24) gives the well-known fact, used frequently in G.C. 1932, that the proportional error in Z is n times the proportional error in z.

An interesting application of Eq. (24), which is needed in G.C. 1932, is the following. Let us assume that we know two points, $y_1 \pm r_1$, at $x = x_1$, and $y_2 \pm r_2$, at $x = x_2$, and that we pass a straight line,

$$y = a + bx \tag{26}$$

through these two points. What are the resulting errors, r_a, r_b and R, in a, b and the function itself? To get the error in a and in b we must first express these quantities in terms of the original data. The results are

$$a = \frac{y_2 x_1 - x_2 y_1}{x_1 - x_2} \tag{27}$$

$$b = \frac{y_1 - y_2}{x_1 - x_2}. \tag{28}$$

We then apply Eq. (24), where now Z is replaced first by a and then by b, and z_1 and z_2 are replaced by y_1 and y_2. The results are

$$r_a^2 = \left(\frac{x_2 r_1}{x_2 - x_1}\right)^2 + \left(\frac{x_1 r_2}{x_1 - x_2}\right)^2 \tag{29}$$

$$r_b^2 = \left(\frac{r_1}{x_1 - x_2}\right)^2 + \left(\frac{r_2}{x_2 - x_1}\right)^2. \tag{30}$$

To obtain the error R in the function $a + bx$, we must similarly express this function directly in terms of y_1 and y_2. From Eqs. (27) and (28) we get

$$y = a + bx = y_1\left(\frac{x_2 - x}{x_2 - x_1}\right) + y_2\left(\frac{x_1 - x}{x_1 - x_2}\right). \tag{31}$$

[32] Merriman, reference 3, pp. 75–79. Palmer, reference 18, pp. 95–104.

Then, from Eq. (24), with $Z=a+bx$,

$$R^2 = \left[\frac{(x_2 - x)r_1}{x_2 - x_1}\right]^2 + \left[\frac{(x_1 - x)r_2}{x_1 - x_2}\right]^2. \tag{32}$$

When $x=x_1$, $R=r_1$, and when $x=x_2$, $R=r_2$, as should of course be the case. An actual plot of the value of R is shown in connection with curve g, Fig. 7, of G.C. 1932.

Let us now assume a more complex situation. Instead of two points, which uniquely determine a straight line, we have a number of points, through which, by means of least squares, we are to pass the *most probable* straight line. What is now the resulting error in a, in b, and in the function itself? The most general situation occurs when we have any function whatsoever, which is to be evaluated by least squares. Formulas for calculating the errors in the various constants so evaluated are given in most good texts,[33] but I have been able to find in only two places, in English, a formula for calculating the error in the function itself. In both places[34] only an entirely general formula is given, of such a character as to be rather inconvenient in actual work. I have *never* seen this formula used, nor even mentioned, although there is a distinct need for the calculation of such errors, in G.C. 1932, and I am sure in much scientific work. The following relations have therefore been obtained.

Instead of discussing the general equation, it is simpler to deduce, from elementary considerations, a specific formula for the probable error of a linear function. Let us assume that weights p have been assigned to the various points, on the basis of the given probable errors ($p=c/r^2$). Then the least squares' values of a and b, and their probable errors r_a and r_b, are given by [35]

$$a = \frac{(\sum py)(\sum px^2) - (\sum px)(\sum pxy)}{D} \tag{33}$$

$$b = \frac{(\sum p)(\sum pxy) - (\sum px)(\sum py)}{D} \tag{34}$$

where

$$D = (\sum p)(\sum px^2) - (\sum px)^2. \tag{35}$$

[33] See, for instance, Palmer, reference 18, Chap. IX, and Whittaker and Robinson, reference 3, pp. 239–242.

[34] The references are (1) Whittaker and Robinson, reference 3, near top of page 243, and (2) Wright and Hayford, reference 31 near bottom of page 138. Both published formulas give the *weight* of a function y. By Eq. (17) or (18′) one can calculate the resulting probable error, on the basis of internal or of external consistency.

[35] In G.C. 1929, page 5, the sign of both numerator and denominator in $a(=a_{01})$ and $b(=a_{11})$ was reversed, resulting in an apparent negative value of the determinant D. This may have led to confusion when D was used under the radical in Eqs. (36) and (37) ahead. Palmer, reference 18, page 112, gives the general formulas for the probable errors of the constants of any function containing only two constants.

Also

$$r_a = r_e\left(\frac{\sum px^2}{D}\right)^{1/2} \tag{36}$$

$$r_b = r_e\left(\frac{\sum p}{D}\right)^{1/2} \tag{37}$$

where

$$r_e = 0.6745\left(\frac{\sum pv^2}{(n-2)}\right)^{1/2}. \tag{38}$$

If r_a and r_b are calculated by internal consistency, r_e is replaced by $r_i(=c^{1/2})$.

From Eq. (26) it is obvious that r_a, the probable error in a, is also the error in the function y at the particular point $x=0$. We desire the error in y, at *any* point $x=\epsilon$. Let us make, in Eq. (26), the linear transformation $x'=x-\epsilon$, giving

$$y = a' + b'x'. \tag{39}$$

Then the error in y at the desired point $x'=0$ is merely the error in a'. This error, from Eq. (36), is given by

$$r_a' = r_e\left(\frac{\sum px'^2}{D}\right)^{1/2} = r_e\left(\frac{\sum p(x-\epsilon)^2}{D}\right)^{1/2} \tag{40}$$

since the determinant D is invariant to a linear transformation in x. Eq. (40) gives the probable error in any *linear* function of x, at the point $x=\epsilon$. In the case of any other rational integral function (r.i.f.) of x, one can obtain the probable error in the function at $x=\epsilon$ directly from the formula for the error in the absolute term a, by merely substituting $x'=x-\epsilon$ in place of x.[36]

The error r_a' of Eq. (40), as might be expected, is least in the vicinity of the observed points, and increases nearly linearly as one extrapolates the function in either direction. In fact the error of the function is a minimum at the "center of gravity" of the observations, defined by $\epsilon=\Sigma px/\Sigma p$, since this is the condition required to make $\Sigma p(x-\epsilon)^2$ a minimum, considering ϵ as the variable. A plot of the value of r_a' is shown, by broken lines, in each figure of G.C. 1932.

[36] The weight of the first unknown (the absolute term of a r.i.f.) is always D/A_{11}, where A_{11} is the minor determinant formed by omitting the first column and top row of the general determinant D (see Whittaker and Robinson, reference 3, pp. 231 and 241). For a r.i.f. of degree f, D is given by

$$\begin{vmatrix} \sum p & \sum px & \sum px^2 \cdots & \sum px^f \\ \sum px & \sum px^2 & \cdots\cdots\cdots & \\ \vdots & & & \vdots \\ \sum px^f & \cdots\cdots\cdots\cdots\cdots & & \sum px^{2f} \end{vmatrix}$$

I have checked these results by showing that in the case of the r.i.f. of zero, first and second degree, the general formula mentioned in footnote 34 actually reduces to the explicit form D/A_{11}, with x' substituted for x in A_{11}.

In scientific work it is quite customary to evaluate the coefficients of a function by least squares' methods. Rather more rarely, the probable error in each coefficient is similarly evaluated. The accuracy of the resulting function, on the other hand, is usually judged roughly from the deviations of the data (the residuals) from the calculated function. The equations just quoted furnish a quantitative evaluation of this accuracy, not only in the vicinity of the data, but also in the extrapolated region on either side of the data. It seems to the writer that such information must necessarily be of value in much scientific work.