Title
LEAST-SQUARES FITTING OF A GAUSSIAN FUNCTION EVALUATION OF THE ERRORS OF THE COEFFICIENTS

Authors
Mitchell, Richard
Madey, Richard.

Publication Date
1953-06-01

UNIVERSITY OF CALIFORNIA

Radiation Laboratory

BERKELEY, CALIFORNIA

UCRL-2783
Unclassified Physics

UNIVERSITY OF CALIFORNIA
Radiation Laboratory
Berkeley, California

Contract No. W-7405-eng-48

LEAST-SQUARES FITTING OF A GAUSSIAN FUNCTION and EVALUATION OF THE ERRORS OF THE COEFFICIENTS

Richard Mitchell and Richard Madey

June 1, 1953

Printed for the U.S. Atomic Energy Commission

LEAST-SQUARES FITTING OF A GAUSSIAN FUNCTION and EVALUATION OF THE ERRORS OF THE COEFFICIENTS

Contents

LEAST-SQUARES FITTING OF A GAUSSIAN FUNCTION and EVALUATION OF THE ERRORS OF THE COEFFICIENTS

Richard Mitchell and Richard Madey

Radiation Laboratory, Department of Physics
University of California, Berkeley, California

June 1, 1954

I. Curve Fitting a Gaussian by the Least-Squares Method

A. Statement of Problem

We want to represent a set of observational points, (R_i, ℓ_i), by the gaussian function

$$\bar{R} = R_o \exp\left[-\frac{(\ell - \ell_o)^2}{2\beta^2}\right] \tag{1}$$

The observed function value, R_i, is measured at an accurately known argument value, $\ell_i - \ell_o$, with an estimated standard deviation, δR_i. In order to represent these points by the smooth function $\bar{R}$, we shall use the least-squares method to evaluate the best arbitrary constants R_o and β for $\bar{R}$.

The best arbitrary constants in the least-squares sense (Appendix A) are such that the sum of the squares of the weighted residuals is a minimum; that is,

$$\sum_i \left(\frac{R_i - \bar{R}_i}{\delta R_i}\right)^2 = \text{a minimum}, \tag{2}$$

where R_i is the observed function value, δR_i the observed experimental standard deviation of R_i at ℓ_i, and $\bar{R}_i$ is calculated at ℓ_i using the best arbitrary constants in the least-squares sense. Thus, we write

$$\sum_i \left(\frac{R_i - R_o \exp\left[-\frac{1}{2}\left(\frac{\ell_i - \ell_o}{\beta}\right)^2\right]}{\delta R_i}\right) = \text{a minimum} \tag{3}$$

This is the least-squares criterion for the best constants R_o and β.

We will transform this criterion to an equivalent criterion in which the arbitrary constants appear linearly. The linear criterion is then differentiated and equated to zero. This process gives us two linear equations in the arbitrary constants, which are easily solved by determinants.

B. Linear Form and Definition of Weight

The gaussian representation in Eq. (1) transforms to the linear form

$$\bar{y} = a_o + a_1 x, \tag{4}$$

where

$$y = \ell n\, R, \tag{5a}$$

$$x = (\ell - \ell_o)^2, \tag{5b}$$

$$a_o = \ln R_o, \tag{5c}$$

$$a_1 = -\frac{1}{2\beta^2}. \tag{5d}$$

Here a_o and a_1 are the new arbitrary constants.

The least-squares criterion for the new function is

$$\sum_i \left(\frac{y_i - \bar{y}_i}{\delta y_i}\right)^2 = \text{a minimum}, \tag{6}$$

where $\bar{y}_i$ is calculated from a_o and a_1 at x_i, and δy_i is the standard deviation of the function y_i.

We now wish to show that

$$\sum_i \left(\frac{R_i - \bar{R}_i}{\delta R_i}\right)^2 = \sum_i \left(\frac{y_i - \bar{y}_i}{\delta y_i}\right)^2, \tag{7}$$

since we are interested in finding the best gaussian representation $\bar{R}_i$. (A more detailed proof appears in Appendix B.) If we relate the residual $R_i - \bar{R}_i$ to the residual $y_i - \bar{y}_i$ by the law of the propagation of small errors (Appendix C), we get

$$(y_i - \bar{y}_i)^2 = \left(\frac{dy}{dR_i}\right)^2 (R_i - \bar{R}_i)^2 \tag{8}$$

$$= \left(\frac{R_i - \bar{R}_i}{R_i}\right)^2 \tag{9}$$

Equation (9) follows by substituting for $\frac{dy}{dR}$ from Equation (5a). With the aid of Eq. (9), the left-hand member of Eq. (7) can be written

$$\sum_i \left(\frac{R_i - \bar{R}_i}{\delta R_i}\right)^2 = \sum_i \left(\frac{y_i - \bar{y}_i}{\delta R_i/R_i}\right)^2 \tag{10}$$

Now, if we identify δy_i by the law of propagation of small errors,

$$(\delta y_i)^2 = (\delta \ln R_i)^2 = \left(\frac{\delta R_i}{R_i}\right)^2, \tag{11}$$

the right-hand member of Eq. (10) becomes identical to that in Eq. (7).

We now define the weight w_i of a residual, $y_i - \bar{y}_i$, at x_i as

$$w_i = \frac{r^2}{(\delta y_i)^2} = \frac{r^2}{(\delta R_i/R_i)^2}, \tag{12}$$

where r^2 is an arbitrary constant so chosen that the w_i's have a range of values that are convenient in computing the normal equations. Once a value of r^2 is assigned, all w_i's are determined. Note that r is the standard deviation of an observation of unit weight.

The least-squares criterion finally is given by

$$\sum_i w_i (y_i - \bar{y}_i)^2 = \text{a minimum}. \tag{13}$$

C. Normal Equations and their Solution

In order to satisfy Eq. (13), the partial derivatives with respect to the arbitrary constants should be zero:

$$\frac{\partial}{\partial a_o}\left[\sum_i w_i (y_i - a_o - a_1 x_i)^2\right] = 0, \tag{14}$$

$$\frac{\partial}{\partial a_1}\left[\sum_i w_i (y_i - a_o - a_1 x_i)^2\right] = 0. \tag{15}$$

Expanding these sums, we get the normal equations

$$a_o \sum_i w_i + a_1 \sum_i w_i x_i = \sum_i w_i y_i , \tag{16}$$

$$a_o \sum_i w_i x_i + a_1 \sum_i w_i x_i^2 = \sum_i w_i y_i x_i . \tag{17}$$

These two equations for the two unknowns a_o and a_1 can be solved by the method of determinants. The solution is

$$a_o = \frac{\begin{vmatrix} \sum_i w_i y_i & \sum_i w_i x_i \\ \sum_i w_i y_i x_i & \sum_i w_i x_i^2 \end{vmatrix}}{D} = \frac{\left(\sum_i w_i y_i\right)\left(\sum_i w_i x_i^2\right) - \left(\sum_i w_i y_i x_i\right)\left(\sum_i w_i x_i\right)}{D}, \tag{18}$$

$$a_1 = \frac{\begin{vmatrix} \sum_i w_i & \sum_i w_i y_i \\ \sum_i w_i x_i & \sum_i w_i y_i x_i \end{vmatrix}}{D} = \frac{\left(\sum_i w_i\right)\left(\sum_i w_i y_i x_i\right) - \left(\sum_i w_i x_i\right)\left(\sum_i w_i y_i\right)}{D}, \tag{19}$$

where

$$D = \begin{vmatrix} \sum_i w_i & \sum_i w_i x_i \\ \sum w_i x_i & \sum w_i x_i^2 \end{vmatrix} = \left(\sum_i w_i\right)\left(\sum_i w_i x_i^2\right) - \left(\sum_i w_i x_i\right)^2 . \tag{20}$$

The best arbitrary constants for a gaussian function are, from Eqs. (5c) and (5d),

$$R_o = e^{a_o} , \quad (21)$$

$$\beta = \sqrt{\frac{-1}{2a_1}} . \quad (22)$$

In conclusion, Eqs. (21) and (22) give the best constants as defined by Eq. (3).

Note that if ℓ_o is not known exactly but is a parameter to be adjusted, we cannot use the techniques of this paper. An alternative technique is the method of differential correction (Appendix D). Also, if ℓ_i has a significantly large error, we are confronted by a multi-error problem whose solution requires knowledge not only of δR_i but of $\delta \ell_i$ as well.[2]

II. The Estimation of Standard Deviations

A. Statement of the Problem

The problem of estimating the standard deviation of the values obtained for the arbitrary constants can be discussed in two parts. The first part concerns the error expected on the basis of the statistical error of the observed points and the propagation of these errors through calculations. This kind of error is called the error based on internal consistency[1]. The second part concerns the error computed on the basis of the deviation of the approximating function from the observational points. This kind of error is called the error based on the external consistency[1].

B. Weighted Observations and Internal Consistency

It is of value at this point to show the relation between weights and errors. When we weight a given observational value, we say that one observation with a small estimated standard deviation is equivalent to the average value of a number of observations with large estimated deviations. The standard deviation r_{y_i} of the ith observations is related to its weight w_i through the equation (Appendix E)

$$r_{y_i} = \frac{r}{\sqrt{w_i}} , \tag{23}$$

where r is the standard deviation of an observation of unit weight. Equation (23) is the same as Eq. (12), where $r_{y_i} \equiv \delta y_i$. Since the weight of the weighted average of n observations is $\sum_i w_i$, the standard deviation $\bar{r}_y$ of the average value of the weighted observations is

$$\bar{r}_y = \frac{r}{\sqrt{\sum_i w_i}} . \tag{24}$$

Now $\bar{r}_y$ is the error based on internal consistency, since it is computed on the basis of the estimated observational errors.

The errors based on internal consistency for the arbitrary constants can be calculated by applying the law of propagation of small errors to the solutions of the normal equations. From Eq. (24) and (18) the internal error in a_o is

$$\bar{r}_{a_o} = r\sqrt{\frac{\Sigma\, w_i x_i^2}{D}} . \tag{25}$$

The derivation of Eq. (25) is given in Appendix F. Similarly, from Eqs. (24) and (19), the internal error in a_1 is

$$\bar{r}_{a_1} = r\sqrt{\frac{\Sigma\, w_i}{D}} . \tag{26}$$

The standard deviation of β is found by again applying the theory of propagation of small errors to Eq. (22). Thus,

$$\bar{r}_\beta = \bar{r}_{a_1}\sqrt{\left(\frac{\beta}{2a_1}\right)^2} . \tag{27}$$

This is the standard deviations of β based on internal consistency. This is not our final answer, for it shows only how observational errors affect the solution and does not show the error introduced by the approximating function.

C. The Relation of the Internal and External Errors

In our problem where the observations are weighted and n different points are observed, the standard deviation based on external consistency is given by the following equations (Appendix G):

$$\epsilon = \sqrt{\frac{\sum_i w_i (y_i - \bar{y}_i)^2}{(n - 2)}} \tag{28}$$

and

$$\bar{\epsilon}_y = \sqrt{\frac{\sum_i w_i (y_i - \bar{y}_i)^2}{(n - 2) \sum_i w_i}} = \frac{\epsilon}{\sqrt{\sum_i w_i}}, \tag{29}$$

where ϵ is the standard deviation of an observation of unit weight, which is estimated from the residual $y_i - \bar{y}_i$, and where $\bar{\epsilon}_y$ is the standard deviation of the linear form of the gaussian function.

We now want to evaluate the external standard deviation of the arbitrary constant a_1. We shall again apply the theory of the law of the propagation of small error to estimate the standard deviation of the arbitrary constants based on external consistency. We finally derive a simple form for these errors by relating external errors to internal errors.

We can calculate the propagation of the external error through the normal equations in the same manner as Eqs. (25) and (26). This calculation is given in Appendix F with the change that ϵ's replace r's. From Eqs. (29) and (18),

$$\bar{\epsilon}_{a_o} = \epsilon \sqrt{\frac{\sum_i w_i x_i^2}{D}}; \tag{30}$$

and from Eqs. (29) and (19),

$$\bar{\epsilon}_{a_1} = \epsilon \sqrt{\frac{\sum_i w_i}{D}}. \tag{31}$$

If we consider the ratio of the internal and external errors, we can re-write these equations. From Eqs. (24) and (29),

$$\frac{\bar{\epsilon}_y}{\bar{r}_y} = \frac{\epsilon (1/\sqrt{\sum_i w_i})}{r(1/\sqrt{\sum_i w_i})} = \frac{\epsilon}{r} \tag{32}$$

or

$$\bar{\epsilon}_y = \frac{\epsilon}{r} \bar{r}_y. \tag{32a}$$

Similarly, from Eqs. (25) and (30) we have the relationship

$$\bar{\epsilon}_{a_o} = \frac{\epsilon}{r} \bar{r}_{a_o}, \tag{33}$$

and from Eqs. (26) and (31)

$$\bar{\epsilon}_{a_1} = \frac{\epsilon}{r} \bar{r}_{a_1}, \tag{34}$$

and also

$$\bar{\epsilon}_{\beta} = \frac{\epsilon}{r} \bar{r}_{\beta}, \tag{35}$$

where

$$\frac{\epsilon}{r} = \frac{1}{r}\sqrt{\frac{\sum_1 w_i (y_i - \bar{y}_i)^2}{(n - 2)}} . \tag{36}$$

These are the estimated standard deviations of the arbitrary constants based on external consistency, because ϵ/r depends on the difference between the observational points and the computed curve.

D. Interpretation of the Ratio of the Internal to the External Error[1]

If the true distribution is a gaussian, then ϵ/r should be unity within some statistical deviation; that is, the standard deviation estimated from the residuals should be equal to the observed standard deviation. If the observed distribution is gaussian-like, the deviation from unity should not be too large. We have to examine two special cases $\epsilon/r < 1$ and $\epsilon/r \gg 1$. If $\epsilon/r < 1$ then the fitting function fits the observational data more closely than is expected on the basis of the statistical errors of the observational points; therefore one should use the internal consistency to estimate errors. Thus in our problem if ϵ/r is less than one,

$$\bar{r}_{\beta} = \bar{r}_{a_1}\sqrt{\left(\frac{\beta}{2 a_1}\right)^2} \tag{27}$$

is taken to be the best estimate of the standard deviation of β.

If ϵ/r is much larger than one, either there are systematic errors in the observations if a gaussian distribution is to be expected, or the gaussian function is not a good representation of the true distribution. In this problem, the argument $\ell - \ell_o$ was assumed to be exact; but a small error in ℓ_o, for example, will make a systematic error in the gaussian representation.

When ϵ/r is greater than unity but not too large, as determined by the problem, we use the external consistency to estimate errors. In our problem

$$\bar{\epsilon}_{\beta} = \frac{\epsilon}{r} \bar{r}_{\beta} \tag{35}$$

is to be taken as the best estimate of the standard deviation of β, if we have $\epsilon/r > 1$ but not too large.

Because of these considerations, the examination of this ratio ϵ/r constitutes an important test in problems of curve fitting.

III. Sample Calculation

A two-counter telescope was used to measure an accidental coincidence counting rate as a function of the delay line in one input to the coincidence circuit. R_i measures the relative counting with cable length ℓ_i, and δR_i is the statistical error based on total counts.

The problem is to determine the best-fitting gaussian for the observed data, the standard deviation for the gaussian, and the estimated error of the standard deviation. We wish to find the best value of ℓ where

$$\overline{R}_i = R_o \exp\left[-\frac{1}{2}\left(\frac{\ell_i - \ell_o}{\gamma}\right)^2\right],$$

and where $\ell_o = 39.40'$ was determined by the oscillator frequency of the synchrotron. We transform to a linear form,

$$y_i = a_o + a_1 x_i ,$$

where $x_i = (\ell_i - \ell_o)^2$ and $y_i = \ell n\, R_i$. The weight of the residual $y_i - \overline{y}_i$ is

$$w_i = r^2 \left(\frac{R_i}{\delta R_i}\right)^2,$$

where $r = 0.1$ is assumed for convenience in computing.

A complete discussion of the collection of these data is given in Reference (4).

ℓ_i	$\ell_i - \ell_o$	x_i	x_i^2	R_i	y_i	$x_i\, y_i$
34	-5.40	29.16	850.31	32.8	3.4904	101.780
37	-2.40	5.76	33.18	115.2	4.7467	27.341
40	0.60	0.36	0.13	365.0	5.8999	2.124
43	3.60	12.96	167.96	129.0	4.8598	62.983
46	6.60	43.56	1897.47	8.6	2.1518	93.732

δR_i	$R_i/\delta R_i$	w_i	$w_i x_i$	$w_i x_i^2$	$w_i\, y_i$	$w_i\, y_i\, x_i$
2.0	16.400	2.6896	78.429	2287.0	9.3878	273.75
4.4	26.128	6.8550	39.485	227.4	32.5386	187.42
9.8	37.245	13.8719	4.994	1.8	81.8428	29.46
5.1	25.294	6.3979	82.917	1074.6	31.0925	402.96
1.3	6.615	0.4376	19.062	830.3	0.9416	41.02
	Sums	30.2520	224.887	4421.1	155.8033	934.61

$$D = (\sum_i w_i)\ (\sum_i w_i x_i^2)\ -\ (\sum_i w_i x_i)^2 = 83173.0$$

$$D_{a_o} = (\sum_i w_i\, y_i)\ (\sum_i w_i\, x_i^2)\ -\ (\sum_i w_i\, y_i\, x_i)\ (\sum_i w_i\, x_i) = 478640.$$

$$a_o = 5.7547$$

$$D_{a_1} = (\sum_i w_i)\ (\sum_i w_i\, y_i\, x_i)\ -\ (\sum_i w_i\, x_i)\ (\sum_i w_i\, y_i) = -6764.32$$

$$a_1 = -0.081325$$

$$\bar{r}_{a_1} = \sqrt{r\ \frac{\sum_i w_i}{D}} = \pm\ 0.001907 \qquad \gamma = \sqrt{\frac{-1}{2a_1}} = 2.4796$$

$$\bar{r}_\gamma = \bar{r}_{a_1}\sqrt{\left(\frac{\gamma}{2\,a}\right)^2} = \pm\ 0.02907$$

$$\bar{y}_i = a_o + a_1\, x_i \qquad v_i = y_i - \bar{y}_i$$

$\bar{y}_i$	v_i	v_i^2	$w_i v_i^2$
3.3833	0.1071	0.01147	0.0308
5.2863	-0.5396	0.29117	1.9960
5.7254	0.1745	0.03045	0.4224
4.7007	0.1591	0.02531	0.1619
2.2122	-0.0604	0.00365	0.0016

$\sum_i w_i v_i = +0.0012$ $\qquad$ $\sum_i w_i v_i^2 = 2.6127$

$$\frac{\epsilon}{r} = \frac{\epsilon_y}{\bar{r}_y} = \frac{1}{r}\sqrt{\frac{\sum_i w_i v_i^2}{n-2}} = \frac{1}{0.1}\sqrt{\frac{2.6127}{3}} = 9.332$$

$$\bar{\epsilon}_\gamma = \frac{\epsilon}{r}\,\bar{r}_\gamma = \pm\ 0.2714.$$

The final answer is

$$\gamma = 2.48 \pm 0.27.$$

In testing these results, ϵ/r was found to be quite large. This large value of ϵ/r indicates that a gaussian function may not be the best representation of the true distribution. Also, since a chance coincidence counting rate was measured, fluctuations in beam intensity could easily lead to systematic errors in the observations. In the above example, however, the solution was not rejected, as γ was significantly larger than ϵ_γ.

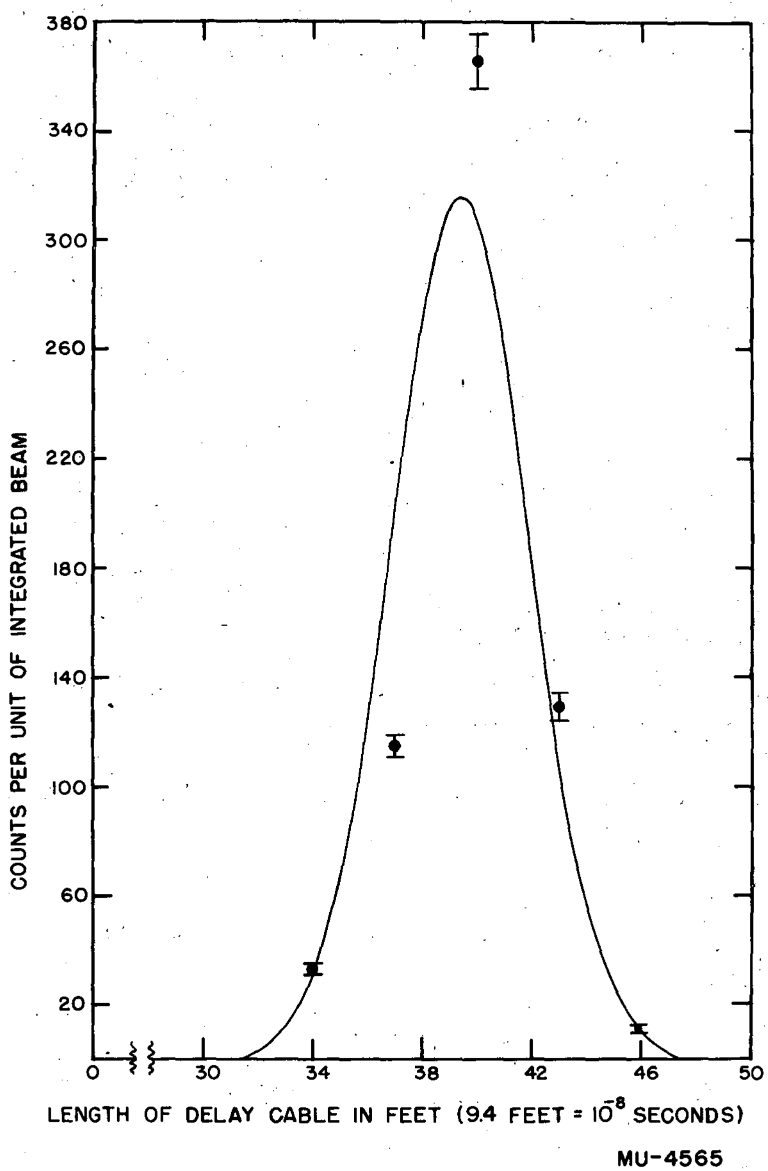

Fig. 1 Plot of calculated function compared to observed points.

IV. Appendix

A. Definition of Least-Squares Criterion[3]

In the simple case where a single quantity x is measured, the probability of observing a given set of random errors is the product of the probabilities for observing each particular error. It is assumed that the errors are normally distributed and that the standard deviation for the error of each measurement is known. The probability that an error $x_i - x$ may be found in the interval $x_i - x$ to $x_i - x + dx_i$ is

$$P_i = \frac{1}{\delta x_i \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_i - x}{\delta x_i}\right)^2\right] dx_i \,. \tag{A1}$$

The probability for the set of errors is

$$P = \prod_i P_i = \prod_i \frac{1}{\delta x_i \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_i - x}{\delta x_i}\right)\right] dx_i \,, \tag{A2}$$

were x is the true value, x_i is the observed value, $x_i - x$ is the observed error, and δx_i is the standard deviation for the ith measurement.

If the true value x is unknown, a reasonable assumption is that the observed set of errors is the most probable set. The probability P will be a maximum, giving the most probable set of errors, if

$$\sum_i \left(\frac{x_i - x}{\delta x_i}\right)^2 = \text{a minimum.} \tag{A3}$$

Now the value for x that would satisfy equation (A3) is defined as $\bar{x}$. The least-squares criterion is

$$\sum_i \left(\frac{x_i - \bar{x}}{\delta x_i}\right)^2 = \text{minimum.} \tag{A4}$$

Let us partially differentiate Eq. (A4) with respect to $\bar{x}$ and equate to zero for a minimum:

$$\sum_i -2\left(\frac{x_i - \bar{x}}{\delta x_i^2}\right) = 0. \tag{A5}$$

Solving Eq. (A5) for $\bar{x}$, we find

$$\bar{x} = \frac{\sum_i \frac{x_i}{\delta x_i^2}}{\sum \frac{1}{\delta x_i^2}} \,, \tag{A6}$$

or that $\bar{x}$ is just the weighted average value of the x_i's.

B. Criterion for this Application

We wish to justify the equation

$$\sum_i \left(\frac{R_i - \bar{R}_i}{\delta R_i}\right)^2 = \sum_i \left(\frac{y_i - \bar{y}_i}{\delta y_i}\right)^2 . \tag{7}$$

Rewrite the left-hand member

$$\Sigma \left(\frac{R_i - \bar{R}_i}{R_i}\right)^2 \frac{1}{(\delta R_i/R_i)^2} = \Sigma \left(1 - \frac{\bar{R}_i}{R_i}\right)^2 \frac{1}{(\delta R_i/R_i)^2} . \tag{B1}$$

Now the identity

$$1 - \frac{\bar{R}_i}{R_i} \equiv 1 - e^{-(\ell n\, R_i - \ell n\, \bar{R}_i)} \tag{B2}$$

becomes

$$1 - \frac{\bar{R}_i}{R_i} = 1 - e^{-(y_i - \bar{y}_i)}, \tag{B3}$$

since $y = \ell n\, R$.

If the difference $y_i - \bar{y}_i$ is small compared with unity,

$$1 - e^{-(y_i - \bar{y}_i)} \approx y_i - \bar{y}_i , \tag{B4}$$

then

$$\Sigma \left[\frac{R_i - \bar{R}_i}{\delta R_i}\right]^2 \approx \Sigma \left[\frac{y_i - \bar{y}_i}{\delta y_i}\right]^2 \tag{7}$$

where $\delta y_i = \delta R_i/R_i$. Thus, the criterion for this application is that the residuals $y_i - \bar{y}_i$ are small compared to unity.

C. The Law of the Propagation of Small Errors[5]

Given the function

$$y = F(x_i\ x_2\ x_3\ x_4 \ldots x_k) \tag{C1}$$

and given the error δx_i in x_i, the observed function is

$$y + \delta y = F(x_i + \delta x_1,\ x_2 + \delta x_2,\ \ldots,\ x_k + \delta x_k), \tag{C2}$$

where δy is the error in y. Expanding in a Taylor's series about $x_1, x_2, x_3, \ldots$ and assuming second-order terms and higher terms are negligible, we get

$$\delta y_j = \sum_{i=1}^{k} \frac{\delta F}{\delta x_i} \delta x_{ij}, \tag{C3}$$

where subscript j differentiates independent observations of the same function. Let us square δy_j and sum over the j observations. Now dividing equations by the number of independent observations n, we get the mean square error, δy^2.

$$\delta y^2 = \frac{\sum_j^n \delta y_j^2}{n} = \sum_j^n \sum_i^k \left(\frac{\partial F}{\partial x_i}\right)^2 \frac{\delta x_{ij}^2}{n}. \tag{C4}$$

We note that the average value of a cross product terms in δy_j^2 should become vanishingly small when averaged over a large number of observations, because the probability for positive and for negative values of an error δx_{ij} are equal; for example, for large n,

$$\sum_{j=1}^{n} 2 \frac{\partial F}{\partial x_i} \frac{\partial F}{\partial x_{i+1}} \frac{\delta x_{ij} \, \delta x_{i+1\,j}}{n} \approx 0. \tag{C5}$$

Letting δx_i^2 represent the mean square error in x_i.

$$\delta y^2 = \sum_i^k \left(\frac{\partial F}{\partial x_i}\right)^2 \sum_j^n \frac{\delta x_{ij}^2}{n} = \sum_i^k \left(\frac{\partial F}{\partial x_i}\right)^2 \delta x_i^2. \tag{C6}$$

The root mean square error is

$$\delta y = \sqrt{\sum_i^k \left(\frac{\partial F}{\partial x_i}\right)^2 \delta x_i^2}. \tag{C7}$$

Because values of δx_{ij} are required only to be small as implied by equation (C3), and positive as frequently as negative, as required by equation (C5), the δx's may be probable errors, and then δy is a probable error; if δx's are standard deviations, then δy is a standard deviation; if δx's are residuals, then δy represents a residual.

D. Method of Differential Correction[3]

If the arbitrary constant ℓ_o is not precisely known then we cannot reduce Eq. (1),

$$\bar{R} = R_o \exp\left[-\frac{(\ell - \ell_o)^2}{2\beta^2}\right],$$

directly to a linear form in the arbitrary constants. Let us define new arbitrary constants as follows:

$$R_o = R'_o + \Delta R_o, \tag{D1}$$

$$\beta = \beta' + \Delta\beta, \tag{D2}$$

$$\ell_o = \ell'_o + \Delta\ell_o, \tag{D3}$$

where R'_o, β', ℓ'_o are approximations to R_o, β, ℓ_o; and where the new arbitrary constants ΔR_o, $\Delta\beta$, $\Delta\ell_o$ are small correction terms. By a Taylor expansion of the gaussian, we get

$$R = R(R'_o\ \beta'\ \ell'_o) + \frac{\partial R(R'_o\ \beta'\ \ell'_o)}{\partial R'_o}\Delta R_o + \frac{\partial R(R'_o\ \beta'\ \ell'_o)}{\partial \beta'}\Delta\beta + \frac{\partial R(R'_o\ \beta'\ \ell'_o)}{\partial \ell'_o}\Delta\ell_o, \tag{D4}$$

where second- and higher-order terms are neglected. R is now a linear function in the new arbitrary constants ΔR_o, $\Delta\beta$, $\Delta\ell_o$. The residual $R_i - \bar{R}_i$ is

$$R_i - \bar{R}_i = R_i - \left\{R'_o \exp\left[-\frac{1}{2}\left(\frac{\ell_i - \ell'_o}{\beta'}\right)^2\right] + \exp\left[-\frac{1}{2}\left(\frac{\ell_i - \ell'_o}{\beta'}\right)^2\right]\Delta R_o + \frac{(\ell_i - \ell'_o)^2}{\beta'^3} R'_o \exp\left[-\frac{1}{2}\left(\frac{\ell_i - \ell'_o}{\beta'}\right)^2\right]\Delta\beta + \frac{\ell_i - \ell'_o}{\beta'^2} R'_o \exp\left[-\frac{1}{2}\left(\frac{\ell_i - \ell'_o}{\beta'}\right)^2\right]\Delta\ell_o\right\}. \tag{D5}$$

We can now derive the normal equations and solve for the best least-squares corrections to the original approximations. This correction process is to be repeated until the correction term is less than the estimated error. The process should be repeated at least once just as a check on the numerical results.

E. The Standard Deviation of the Average of a Number of Observations.

In the simple case of a set of n observations of equal accuracy of a single point, the standard deviation of the average value of the point is given by

$$\bar{r} = \frac{r'}{\sqrt{n}}, \tag{E1}$$

where r' is the standard deviation of any one observation. This result follows simply from applying the law of propagation of small errors (Appendix C) to the equation for the simple average of n measurements of equal accuracy.

Let x be the quantity measured and x_1, x_2, x_3, ... , x_n be n measurements of x with equal accuracy so that r' is the standard deviation of each measurement. The average value of x is

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} . \tag{E2}$$

Applying the law of propagation of small errors,

$$\bar{r}^2 = \frac{r'^2}{n^2} + \frac{r'^2}{n^2} + \ldots = \frac{n}{n^2} r'^2, \tag{E3}$$

we confirm Eq. (E1).

Since weights are defined so that an observation with small standard deviation is equivalent to the average value of a number of observations with large standard deviations, it follows that the standard deviation r_i of the ith observation with weight w_i is

$$r_i = \frac{r}{\sqrt{w_i}} \tag{E4}$$

where r is the standard deviation of an observation with unit weight. If we now apply the law of propagation of small errors to the weighted average,

$$\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i}, \tag{E5}$$

the standard deviation of the average of the weighted observations is

$$\bar{r} = \frac{r}{\sqrt{\sum_i w_i}} \tag{E6}$$

F. Propagation of Errors through the Normal Equations

Let us consider the error in a_o. By equation (18), a_o is

$$a_o = \frac{\left(\sum_{i=1}^{n} w_i y_i\right)\left(\sum_{i=1}^{n} w_i x_i^2\right) - \left(\sum_{i=1}^{n} w_i y_i x_i\right)\left(\sum_{i=1}^{n} w_i x_i\right)}{D} . \tag{18}$$

Take the partial derivations of a_o with respect to each y, and square each term, according to the law of propagation of errors, to get

$$(\bar{r}_{a_o})^2 = \sum_{j=1}^{n} (\delta y_j)^2 \left(\frac{\partial a_o}{\partial y_j}\right)^2 \tag{F1}$$

$$(\bar{r}_{a_o})^2 = \sum_j (\delta\ y_j)^2 \left(\frac{w_j \sum_i w_i x_i^2 - w_j x_j \sum_i w_i x_i}{D} \right)^2 \quad \text{(F2)}$$

$$(\bar{r}_{a_o})^2 = \sum (\delta\ y_j)^2 \left[\frac{w_j^2 \left(\sum_i w_i x_i^2 \right)^2 - 2\ w_j^2 x_j \left(\sum_i w_i x_i^2 \right) \left(\sum_i w_i x_i \right)}{D^2} + \frac{w_j^2 x_j^2 \left(\sum w_i x_i \right)^2}{D^2} \right] . \quad \text{(F3)}$$

We have the condition from equation (23) on w_i that

$$w_i = \frac{r^2}{(\delta\ y_i)^2} \quad \text{(23)}$$

or

$$(\delta\ y_j)^2 = \frac{r^2}{w_j} . \quad \text{(23a)}$$

Substituting by equation (23a) for $\delta\ y_j$, equation (F3) now has the form

$$(\bar{r}_{a_o})^2 = r^2 \sum_j \left[\frac{w_j \left(\sum_i w_i x_i^2 \right)^2 - 2\ w_j x_j \left(\sum_i w_i x_i^2 \right) \left(\sum_i w_i x_i \right)}{D^2} + \frac{w_j x_j^2 \left(\sum_i w_i x_i \right)^2}{D^2} \right] . \quad \text{(F4)}$$

After carrying out the summation over j, and since

$$\sum_{j=1}^{n} w_j x_j = \sum_{i=1}^{n} w_i x_i$$

and

$$\sum_{j=1}^{n} w_j = \sum_{i=1}^{n} w_i ,$$

$$(\bar{r}_{a_o})^2 = r^2 \frac{\left(\sum_i w_i \right) \left(\sum_i w_i x_i^2 \right)^2 - \left(\sum_i w_i x_i^2 \right) \left(\sum_i w_i x_i \right)^2}{D^2} . \quad \text{(F5)}$$

Finally, factor out $\sum_i w_i x_i^2$ and D as given in Eq. (20); the error squared is

$$(\bar{r}_{a_o})^2 = r^2 \frac{\sum_i w_i x_i^2}{D} , \tag{F6}$$

and finally

$$\bar{r}_{a_o} = r\sqrt{\frac{\sum_i w_i x_i^2}{D}} . \tag{25}$$

G. Error Estimated from Residuals[3,6]

We can estimate the standard deviation of a single observation on the basis of the consistency of a number of observations.

Let x be the quantity measured and x_1, x_2, . . ., x_n be n measurements of equal accuracy of x. Let $\bar{x}$ denote the average value of the observations,

$$\bar{x} = \frac{\sum_i x_i}{n} . \tag{G1}$$

The residual of the ith observation is

$$v_i \equiv x_i - \bar{x} , \tag{G2}$$

and the true error of the ith observation is

$$\epsilon_i \equiv x_i - x . \tag{G3}$$

The error of the average value is

$$\bar{x} - x = \frac{\sum_i \epsilon_i}{n} . \tag{G4}$$

The true value x is used to relate errors and residuals, since

$$x = \bar{x} - \frac{\sum_i \epsilon_i}{n} = x_1 - \epsilon_1 = x_2 - \epsilon_2 = x_i - \epsilon_i . \tag{G5}$$

By rearrangement of Eq. (G5) the true error of the ith measurement is

$$\epsilon_i = x_i - \left(\bar{x} - \frac{\sum_i^n \epsilon_i}{n} \right) \tag{G6}$$

or

$$\epsilon_i = v_i + \frac{\sum_i \epsilon_i}{n} \tag{G7}$$

or

$$v_i = \epsilon_i - 1/n \sum_{i=1}^{n} \epsilon_i \, . \tag{G8}$$

Thus, the residuals can be written as linear equations in the errors,

$$v_i = (n-1/n)\, \epsilon_i - 1/n \sum_{j \neq i} \epsilon_j \, . \tag{G9}$$

Let us apply to Eq. (G9) the law of propagation of small errors to find the the relation between the standard deviation σ_v of a set of residuals, and the standard deviation ϵ of the corresponding set of errors for a single measurement,

$$\sigma_v = \epsilon \sqrt{\sum_{k=1}^{n} (\partial v_i / \partial \epsilon_k)^2} \, , \tag{G10a}$$

$$\sigma_v = \epsilon \sqrt{(n-1/n^2) + 1/n^2 \, (n-1)} \, , \tag{G10b}$$

$$\sigma_v = \epsilon \sqrt{n-1/n} \, . \tag{G10c}$$

The standard deviation of the residuals is

$$\sigma_v = \sqrt{\sum_i^n v_i^2 / n} \, . \tag{G11}$$

Solve Eqs. (G10c) and (G11) for the standard deviation of the true errors for a single measurement:

$$\epsilon = \sqrt{\sum_{i=1}^{n} v_i^2 / n-1} \tag{G12}$$

In Appendix E we found the standard deviation of the average of n observations of equal accuracy. In this case,

$$\bar{\epsilon} = \frac{\epsilon}{n} \, . \tag{G13}$$

Thus, the estimated standard deviation of the average of the observations is

$$\bar{\epsilon} = \sqrt{\frac{\sum_i v_i^2}{n(n-1)}} \tag{G14}$$

Now take a general case where x is a linear function of q unknown independent parameters and where the observation x_i is given weight w_i. The estimated standard deviation of the weighted mean of the n observations is

$$\bar{\epsilon} = \sqrt{\frac{\sum_i w_i v_i^2}{(n-q) \sum w_i}} \, , \tag{G15}$$

and the standard deviation of an observation of unit weight is

$$\epsilon = \sqrt{\frac{\sum_i w_i v_i^2}{(n-q)}} \quad . \tag{G16}$$

Proof for Eq. (G15) can be found in Reference (6).

For the case of the gaussian function in Eq. (1), the number of independent parameters is two, viz., R_o and β.

Acknowledgments

The authors wish to acknowledge discussions with Edward Vaughan.

This work was performed under the auspices of the U. S. Atomic Energy Commission.

References

1. R. T. Birge, unpublished lecture notes for Physics 204, University of California (Berkeley). (Principal source of the discussion of errors in this paper.)

2. R. J. Trumpler and H. Weaver, Statistical Astronomy, University of California Press, Berkeley and Los Angeles, 1953, Section 1.74.

3. H. Margenau and G. Murphy, The Mathematics of Physics and Chemistry, Van Nostrand, New York, 1943, Chapter 13, Part 3.

4. R. Madey, K. C. Bandtel, and W. J. Frank, "The Radiofrequency Fine Structure of the Photon Beam from the Berkeley Synchrotron," University of California Radiation Laboratory Report No. UCRL-2052; (also Rev. Sci, Instr. 25, 537 (1954)).

5. R. T. Birge, "The Propagation of Errors," American Journal of Physics 7, 351-357 (Dec., 1939).

6. E. Whittaker and G. Robinson, A Calculus of Observations, Blackie, London, Fourth edition, Section 124.