RELATIVE RESIDUALS CONSIDERED AS WEIGHTED SIMPLE RESIDUALS IN THE APPLICATION OF THE METHOD OF LEAST SQUARES

WALTER A. HENDRICKS
Junior Biologist, Bureau of Animal Industry
United States Department of Agriculture

In a recent paper the writer[1] discussed some considerations involved in fitting a curve, by the method of least squares, to data in which the magnitude of the errors of measurement was affected by the size of the dependent variable. For the special case in which the percentage errors of measurement were distributed normally, it was shown that the most probable values of the dependent variable could be calculated by minimizing the sum of the squares of residuals of the type, $V - \frac{Y}{f(X)}$, with respect to V, V being the arithmetic mean of the ratios of the observed values of the dependent variable to the corresponding calculated values and equal to unity at that minimum.

The concept of a relative residual has a certain value to the investigator as an aid in visualizing the nature of such a set of data. However, it is possible to use a different method of analysis, based on the theory of weighting, which will yield exactly the same results when applied to such a set of data and in addition possesses the advantage of being applicable to more general problems in which the relation of the errors of measurement to the values of the dependent variable is more complex.

Standard texts on the method of least squares such as that by Merriman,[2] show that if the probability of the occurrence of an error of a given magnitude varies for measurements of successive values of the dependent variable, it is necessary to weight the observation equations when fitting the curve. If the errors of

---

[1]Hendricks, Walter A. 1931. The use of the relative residual in the application of the method of least squares. *Annals of Mathematical Statistics,* 2 (4): 458-478.

[2]Merriman, Mansfield. 1907. The method of least squares. 230 p., illus. John Wiley & Sons, New York.

measurement, made in obtaining one observed value of each of several successive values of a dependent variable, $f(X)$, are influenced by the magnitude of, $f(X)$, the probability of the occurrence of errors of the magnitudes, $x_1, x_2, x_3, \ldots \ldots x_n$, respectively, is given by the following equations:

$$P_1 = k_1 e^{-h_1^2 x_1^2}$$

$$P_2 = k_2 e^{-h_2^2 x_2^2}$$

$$P_3 = k_3 e^{-h_3^2 x_3^2}$$

$$P_n = k_n e^{-h_n^2 x_n^2} \quad \ldots \ldots \quad (1)$$

The probability of the occurrence of the given system of errors is given by the product:

$$P' = k' e^{-(h_1^2 x_1^2 + h_2^2 x_2^2 + h_3^2 x_3^2 + \cdots h_n^2 x_n^2)} \quad (2)$$

in which $P' = P_1 . P_2 . P_3 \cdots P_n$ and $k' = k_1 . k_2 . k_3 \cdots k_n$.

If the exponent of e in equation (2) is divided by a constant measure of precision, h^2, the equation may be written in the form:

$$P' = k' e^{-h^2 (p_1 x_1^2 + p_2 x_2^2 + p_3 x_3^2 + \cdots p_n x_n^2)} \quad (3)$$

in which $h_1^2 = p_1 h^2$, etc., and $p_1, p_2, p_3, \cdots p_n$ are the weights of the corresponding errors. P' will have its maximum value when the value of the expression, $p_1 x_1^2 + p_2 x_2^2 + p_3 x_3^2 + \cdots p_n x_n^2$, is a minimum.

Applying the above principles to curve fitting and substituting a residual, v, for every error, x, to distinguish the residuals from the true errors, it is evident that the constants of a fitted equation must be determined in such a manner that the value of the expression, $p_1 v_1^2 + p_2 v_2^2 + p_3 v_3^2 + \cdots p_n v_n^2$, is a minimum.*

If the equation to be fitted is of the type used in the writer's previous study (loc. cit.), viz.:

$$Y = AX^2 \quad \dots\dots\dots\dots \quad (4)$$

this condition is obviously satisfied by the solution of the following equation:

$$p_1 v_1 \frac{\partial v_1}{\partial A} + p_2 v_2 \frac{\partial v_2}{\partial A} + p_3 v_3 \frac{\partial v_3}{\partial A} + \cdots p_n v_n \frac{\partial v_n}{\partial A} = 0 \quad -- (5)$$

Let Y_i represent any observed value of the dependent variable and let AX_i^2 represent the corresponding most probable value. Then it is evident that:

$$v_i = AX_i^2 - Y_i \quad \dots\dots\dots\dots \quad (6)$$

and

$$\frac{\partial v_i}{\partial A} = x_i^2 \quad \dots\dots\dots\dots \quad (7)$$

All that remains is to find the weight, p_i.

Equations (2) and (3) show that the weights of the errors are proportional to the respective measures of precision, h_1^2, h_2^2, h_3^2, --- h_n^2. It follows from the well-known relation between the measures of precision and variance that any measure of precision, h_i^2, is equal to $\frac{1}{2\sigma_i^2}$ in which σ_i is the standard error of the observed value, Y_i, of the dependent variable. Therefor, any weight, p_i, is given by the equation:

$$p_i = \frac{1}{2h^2 \sigma_i^2} \quad \dots\dots\dots\dots \quad (8)$$

*The above development follows that given by Merriman with a slight change in notation.

in which h^2 is the constant measure of precision in equation (3).

If a set of data were obtained by making one measurement of each of several successive values of the dependent variable, AX^2, and the resulting percentage errors of measurement were distributed normally, it follows that the coefficient of variation of a number of replicate measurements made at any value of AX^2 would be equal to that obtained for every other value of AX^2. In other words, the standard error of every measurement would be directly proportional to the value of AX^2 measured.

Since, by equation (8), the weight of any error of measurement is inversely proportional to the square of its standard error, it follows from the above discussion that this weight must also be inversely proportional to the square of the value of AX^2 measured. Combining all factors of proportionality into a composite constant, C, the relation between any weight, p_i, and the corresponding value of the dependent variable, AX_i^2, may be expressed by the equation:

$$p_i = \frac{C}{A^2X_i^4} \quad \dots\dots\dots \quad (9)$$

If the substitutions suggested by equations (6), (7), and (9) are made in equation (5), this equation may be written:

$$\frac{C(AX_1^2-Y_1)X_1^2}{A^2X^4} + \frac{C(AX_2^2-Y_2)X_2^2}{A^2X^4} + \frac{C(AX_3^2-Y_3)X_3^2}{A^2X^4} + \dots$$

$$\frac{C(AX_n^2-Y_n)X_n^2}{A^2X^4} = 0 \quad \dots\dots\dots\dots\dots \quad (10)$$

Since the constant, C, is common to every term in equation (10), it may be removed by division. The equation may then be reduced to the form:

$$\Sigma \frac{(AX^2-Y)X^2}{A^2X^4} = 0 \quad \dots\dots\dots\dots \quad (11)$$

If there are n observation equations, equation (11) may be written in the form:

$$\frac{n}{A} - \frac{1}{A^2} \Sigma \frac{Y}{X^2} = 0 \quad \ldots\ldots\ldots \quad (12)$$

from which the most probable value of A may be readily calculated.

$$nA - \Sigma \frac{Y}{X^2} = 0$$

or

$$A = \frac{1}{n} \Sigma \frac{Y}{X^2} \quad \ldots\ldots\ldots \quad (13)$$

Equation (13) is identical with the equation obtained by minimizing residuals of the type, $V - \frac{Y}{AX^2}$, reported in the writer's earlier study (loc. cit.). The development given in the present paper is perhaps the better from the purely mathematical point of view since it involves nothing more than a systematic weighting of the observation equations. It can be applied to any problem in curve fitting if the standard error of each observed value of the dependent variable is known or can be deduced from a priori considerations. For example, it often happens that the means of replicated measurements, rather than the individual measurements themselves, are used in fitting the curve. In such cases the standard error of each mean may easily be calculated. The reciprocals of the squares of these standard errors will then be the required weights of the observation equations.

However, if the standard errors are proportional to the values of the dependent variable, it may be desirable to retain the concept of a relative residual. The significance of a percentage error of measurement probably can be appreciated by many investigators in various fields of research, particularly those whose contact with mathematics is more or less incidental, to whom a system of weighting would seem somewhat artificial and arbitrary.

In either event, the necessary computations are identical. The precise procedure described in the present paper, like that developed in the writer's previous study (loc. cit.), cannot be applied when the equation which is to be fitted contains more than one undetermined constant. However, in actual practice it is usually sufficiently accurate to substitute the square of the observed value of the dependent variable for that of the corresponding most probable value in equation (9). If this is done, the method can be applied to any equation which can be fitted by the method of least squares. Using this substitution is equivalent to expressing the errors of measurement as fractions of the observed values of the dependent variable when the standard error of each measurement is proportional to the quantity measured.

Walter A. Hendricks