THE STANDARD ERROR OF ANY ANALYTIC FUNCTION OF A SET OF PARAMETERS EVALUATED BY THE METHOD OF LEAST SQUARES

By

WALTER A. HENDRICKS,
Bureau of Animal Industry,
U. S. Department of Agriculture, Washington, D. C.

After fitting a curve to a set of data by the method of least squares, it is occasionally necessary to use the resulting values of some or all of the parameters of the curve in further calculations. Since the estimates of the values of the parameters obtained from a particular set of data are subject to errors of sampling, it follows that the result of any calculation involving those values of the parameters will have a certain standard error. Since the estimated values of the parameters are not independent of each other, the familiar formulas based on the assumption of independence should not be used for the purpose of calculating this standard error from the standard errors of the parameters themselves. The correct approach to the problem involves little more than an application of the methods presented by Schultz (1930) in his excellent paper describing the method of calculating the standard error of a particular function of the parameters, viz., the same function which was used in evaluating the parameters.

Let $y = \varphi(\lambda_1, \lambda_2, \cdots, \lambda_k)$ be an analytic function involving the k parameters, λ_i. This function may not be linear with respect to the parameters, so that if the parameters are to be evaluated by the method of least squares, we have in the general case a function of the form:

$$(1) \quad y = \varphi(\lambda_{1_0}, \lambda_{2_0}, \cdots, \lambda_{k_0}) + \frac{\partial \varphi}{\partial \lambda_{1_0}} \Delta\lambda_1 + \cdots + \frac{\partial \varphi}{\partial \lambda_{k_0}} \Delta\lambda_k + \cdots$$

from which the values of the parameters may be obtained by assuming approximate values and calculating the corrections which must be added to obtain the most probable values.

After the values of the parameters have been obtained, let it be required to find the standard error of a new function, $z = \theta(\lambda_1, \lambda_2, \cdots, \lambda_k)$, involving those values. If z is an analytic function of the parameters, we have to a close approximation:

(2) $$z = \theta(\lambda_{1_0}, \lambda_{2_0}, \cdots, \lambda_{k_0}) + \frac{\partial\theta}{\partial\lambda_{1_0}}\Delta\lambda_1 + \frac{\partial\theta}{\partial\lambda_{2_0}}\Delta\lambda_2 + \cdots + \frac{\partial\theta}{\partial\lambda_{k_0}}\Delta\lambda_k.$$

Any error in z, beyond the insignificant error introduced by the above expansion, will then be due only to errors in $\Delta\lambda_1, \Delta\lambda_2, \cdots\cdots, \Delta\lambda_k$.

Therefore, if

(3) $$f = \frac{\partial\theta}{\partial\lambda_{1_0}}\Delta\lambda_1 + \frac{\partial\theta}{\partial\lambda_{2_0}}\Delta\lambda_2 + \cdots + \frac{\partial\theta}{\partial\lambda_{k_0}}\Delta\lambda_k$$

and S_z and S_f denote the standard errors of z and f, respectively, it is at once apparent that $S_z = S_f$.

The values of $\Delta\lambda_1, \Delta\lambda_2, \ldots, \Delta\lambda_k$ may be expressed in terms of the data from which they were evaluated,

(4) $$\begin{cases} \Delta\lambda_1 = \sigma_1 M_1 + \sigma_2 M_2 + \cdots + \sigma_n M_n \\ \Delta\lambda_2 = \tau_1 M_1 + \tau_2 M_2 + \cdots + \tau_n M_n \\ - - - - - - - - - - - - - - - - - \\ \Delta\lambda_k = \xi_1 M_1 + \xi_2 M_2 + \cdots + \xi_n M_n \end{cases}$$

in which the values, M_i, represent the n observed values of the variable, y; f may then be expressed in the form:

(5) $$f = \sum_{i=1}^{n}\left[\frac{\partial\theta}{\partial\lambda_{1_0}}\sigma_i M_i + \frac{\partial\theta}{\partial\lambda_{2_0}}\tau_i M_i + \cdots + \frac{\partial\theta}{\partial\lambda_{k_0}}\xi_i M_i\right].$$

From the well-known laws of propagation of error and the fact that $S_z = S_f$, it follows that

(6) $$S_z^2 = \sum_{i=1}^{n}\left[\frac{\partial\theta}{\partial\lambda_{1_0}}\sigma_i + \frac{\partial\theta}{\partial\lambda_{2_0}}\tau_i + \cdots + \frac{\partial\theta}{\partial\lambda_{k_0}}\xi_i\right]^2 S_y^2,$$

in which S_y is the standard error of estimate of y based on $n-k$

degrees of freedom. If the right-hand member of equation (6) is expanded, the equation may be written in the form:

$$S_z^2 = \Big\{ \left(\frac{\partial\theta}{\partial\lambda_{1_0}}\right)^2 [\sigma\sigma] + \left(\frac{\partial\theta}{\partial\lambda_{2_0}}\right)^2 [\tau\tau] + \cdots + \left(\frac{\partial\theta}{\partial\lambda_{k_0}}\right)^2 [\xi\xi] +$$
$$(7) \qquad 2\left(\frac{\partial\theta}{\partial\lambda_{1_0}}\right)\left(\frac{\partial\theta}{\partial\lambda_{2_0}}\right)[\sigma\tau] + \cdots + 2\left(\frac{\partial\theta}{\partial\lambda_{1_0}}\right)\left(\frac{\partial\theta}{\partial\lambda_{k_0}}\right)[\sigma\xi] +$$
$$\cdots + 2\left(\frac{\partial\theta}{\partial\lambda_{2_0}}\right)\left(\frac{\partial\theta}{\partial\lambda_{k_0}}\right)[\tau\xi] + \cdots \Big\} S_y^2,$$

in which $[\sigma\sigma] = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$, etc.

The values of the sums of the squares and products multiplying the differential coefficients in equation (7) may be obtained from the normal equations formed for the evaluation of $\Delta\lambda_1, \Delta\lambda_2,$ .... $\Delta\lambda_k$. Let the normal equations be:

$$(8) \qquad \begin{array}{l} [aa]\,\Delta\lambda_1 + [ab]\,\Delta\lambda_2 + \cdots + [al]\,\Delta\lambda_k = [aM] \\ [ba]\,\Delta\lambda_1 + [bb]\,\Delta\lambda_2 + \cdots + [bl]\,\Delta\lambda_k = [bM] \\ - - - - - - - - - - - - - - - - \\ [la]\,\Delta\lambda_1 + [lb]\,\Delta\lambda_2 + \cdots + [ll]\,\Delta\lambda_k = [lM] \end{array}$$

If these equations are solved for $\Delta\lambda_1, \Delta\lambda_2, \cdots, \Delta\lambda_k$ by the method of undetermined multipliers, the first is multiplied by an undetermined constant, α_1, the second by α_2, etc., and the resulting products are added. The conditions for the solution for $\Delta\lambda_1$ are:

$$(9) \qquad \begin{array}{l} [aa]\,\alpha_1 + [ab]\,\alpha_2 + \cdots + [al]\,\alpha_k = 1 \\ [ba]\,\alpha_1 + [bb]\,\alpha_2 + \cdots + [bl]\,\alpha_k = 0 \\ [la]\,\alpha_1 + [lb]\,\alpha_2 + \cdots + [ll]\,\alpha_k = 0 \end{array}$$

To solve for $\Delta\lambda_2$, equations (8) are multiplied by $\beta_1, \beta_2, \cdots, \beta_k$, respectively, added, and the following conditions imposed:

$$(10) \qquad \begin{array}{l} [aa]\,\beta_1 + [ab]\,\beta_2 + \cdots + [al]\,\beta_k = 0 \\ [ba]\,\beta_1 + [bb]\,\beta_2 + \cdots + [bl]\,\beta_k = 1 \\ [la]\,\beta_1 + [lb]\,\beta_2 + \cdots + [ll]\,\beta_k = 0 \end{array}$$

To solve for $\Delta\lambda_k$, equations (8) are multiplied by $\omega_1, \omega_2, \cdots$ ω_k, respectively, added, and the following conditions imposed:

$$
\begin{aligned}
&[aa]\,\omega_1 + [ab]\,\omega_2 + \cdots + [al]\,\omega_k = 0 \\
(11)\qquad &[ba]\,\omega_1 + [bb]\,\omega_2 + \cdots + [bl]\,\omega_k = 0 \\
&[la]\,\omega_1 + [lb]\,\omega_2 + \cdots + [ll]\,\omega_k = 1
\end{aligned}
$$

It may be proved that:

$$
\begin{aligned}
&\alpha_1 = [\sigma\sigma] \qquad \beta_1 = [\tau\sigma] \qquad \omega_1 = [\xi\sigma] \\
(12)\qquad &\alpha_2 = [\sigma\tau] \qquad \beta_2 = [\tau\tau] \qquad \omega_2 = [\xi\tau] \\
&\alpha_k = [\sigma\xi] \qquad \beta_k = [\tau\xi] \qquad \omega_k = [\xi\xi].
\end{aligned}
$$

The method of deriving equations (12) is indicated in the well-known text on the method of least squares by Merriman (1907) in which a detailed proof of the fact that β_2 is equal to $[\tau\tau]$ is presented. The other relations may be derived in analogous fashion. It may be observed that $[\sigma\tau] = [\tau\sigma]$, etc.

The required quantities to be substituted in equation (7) may, therefore, be calculated by solving the sets of simultaneous equations, (9), (10), and (11).

This completes the solution of the general problem presented in the first part of this paper. Some confusion may arise in regard to the proper application of the methods described above if one or both of the functions, y and z, happens to be in a linear form with respect to the parameters. It may be shown that the formulas given will hold in any of these special cases. Although Taylor's theorem may be applied to such functions, such a treatment is superfluous. If either or both of the functions, y and z, is linear with respect to the parameters, the expression for S_z^2 is identical with equation (7) even though the linear function, or functions, was not first expanded by Taylor's theorem. Furthermore, if y is linear with respect to the parameters, the values of the coefficients. $[\sigma\sigma]$, etc., in equation (7) will be the same regardless of whether the parameters were evaluated directly or whether y was first expanded. The latter statement is evident from an inspection of equations (9), (10), and (11) and a consideration of the law of formation of normal equations.

As an example of the application of the methods presented in this paper to a specific problem, consider a set of data given by Spillman (1933) relating to the yields of potatoes obtained from four plots of ground which had been treated with different amounts of potash.

YIELDS OF POTATOES FROM FOUR PLOTS OF GROUND RECEIVING DIFFERENT AMOUNTS OF POTASH

x (Units of K_2O)	y (Bushels of potatoes)
0	91
1	251
2	331
3	381

When a simple exponential equation of the form,

$$y = A - Be^{-\kappa x} \tag{13}$$

was fitted to this set of data, the most probable values of the parameters, A, B, and κ, were found to be as follows:

$$\begin{aligned} A &= 432.801 \pm 11.637 \\ B &= 341.393 \pm 11.406 \\ \kappa &= 0.6195918 \pm 0.0462871 \,. \end{aligned}$$

The value of the product of the parameters, A and κ, happens to be of some interest, at least to the author of this paper, since it gives the value of the first derivative of y with respect to x at the point where the curve crosses the x-axis. In the present example it represents the increase in yield, per unit increase in amount of potash applied, which would be possible if certain inhibiting influences, which seem to be proportional to the yield, had no effect. For the particular data under consideration, the value of this product is 268.160.

In order to calculate the standard error of this value, equation (7) was applied as follows:

$$S_z^2 = \left(\kappa^2[\sigma\sigma] + A^2[\xi\xi] + 2A\kappa[\sigma\xi] \right) S_y^2 , \tag{14}$$

from which the standard error of $A\kappa$ was found to be equal to $\pm$ 13.331.

The familiar formula for calculating the standard error of the product of two independent quantities, when employed for the purpose of calculating the standard error of $A\kappa$ may be written in the form:

$$S_z^2 = \left(\kappa^2[\sigma\sigma] + A^2[\xi\xi] \right) S_y^2 . \tag{15}$$

Equation (15) gives a value of $\pm$ 21.291 for the standard error of $A\kappa$, which deviates considerably from the correct value given by equation (14). The discrepancy is due entirely to the fact that the estimated values of A and κ are not independent.

REFERENCES

MERRIMAN, MANSFIELD. 1907. The Method of Least Squares. John Wiley & Sons, New York.

SCHULTZ, HENRY. 1930. The standard error of a forecast from a curve. Jour. Amer. Stat. Assoc., 25 (N. S. 17): 139-185.

SPILLMAN, W. J. 1933. Use of the exponential yield curve in fertilizer experiments. U. S. D. A. Tech. Bull. 348.