NOTES

This section is devoted to brief research and expository articles, notes on methodology and other short items.

NOTE ON THE ADJUSTMENT OF OBSERVATIONS

By Arthur J. Kavanagh

The Forman Schools, Litchfield, Conn.

The method of least squares has been extended to the adjustment of observations with errors in more than one variable. The history of the development and its principal results have been given by Deming [2], [3], [4], [5]. The basis is the assumption that for the "best" adjustment the sum of the weighted squares of all the residuals (observed values minus adjusted values) must be made a minimum with respect to the adjustments to the observations and with respect to the parameters involved in the conditions the adjusted values must satisfy. In certain problems, such as some arising in the study of relative growth in biology, this assumption is not adequate; it is necessary that the sum to be minimized be generalized to include cross products as well as squares of the residuals.

Suppose we have a set of n universes of q-dimensional points whose centers of gravity are known to satisfy certain conditions; for instance, they might all lie on a certain type of curve. A sample having been taken from each universe, the center of gravity of each sample is taken as the observed center of gravity of the corresponding universe, and it is desired to determine the most probable set of adjustments to the coordinates and the most probable set of parameters involved in the conditions, subject to the requirement that the adjusted values satisfy the conditions exactly. It is assumed that the sampling distribution of the center of gravity in each universe satisfies the multivariate normal law, and that the standard deviations and coefficients of correlation of each sample may with sufficient accuracy be taken as the constants of the corresponding universe. Then by reasoning analogous to that of the derivation of the least squares principle for one variable from the univariate normal law, the probability of getting the observed set of values is proportional to e^{-Q}, where

$$Q = \sum_{i=1}^{n} Q_i \tag{1}$$

Q_i being a homogeneous quadratic function of the errors at the ith centroid and in general involving the cross products as well as the squares of the errors.

The probability will be a maximum when Q is a minimum. Consequently the best estimates for the coordinates of the centroids will be those making Q a minimum, subject to the conditions which the coordinates must satisfy.

For example it may be desired to study the relation between height and weight among growing boys by fitting a curve to the points whose abscissa and ordinate are respectively average height and average weight of a particular age group, one point corresponding to each age group in the study. The data for such a study are obtained from samples of the several age groups. Then the number n of universes is the number of age groups being studied, each universe consisting of the totality of two-dimensional points obtained by pairing the height with the weight of each boy in the age group. The centroid or "average point" of each universe would ideally be obtained from measurements of all the individuals of that age, but since sampling must be resorted to it is necessary to make allowances for the sampling distributions of the centroids. It is known that within each age group there is correlation between height and weight [1]. Consequently the sampling distribution of each centroid will exhibit a correlation which can be expressed in terms of the coefficient of correlation between height and weight of the individuals of the universe from which the sampling distribution arises. The existence of this correlation results in the presence of the cross-product term in the exponent of the bivariate normal formula describing the sampling distribution of the average values, that is in the Q_i of each centroid. If there were no such correlations the cross-product term in each Q_i would vanish and the situation would reduce to that of least squares.

In the general case, let $X_{1i}, X_{2i}, \cdots, X_{qi}$ be the observed coordinates of the ith centroid, $x_{1i}, x_{2i}, \cdots, x_{qi}$ the adjusted values (to be determined), and $V_{ji} = X_{ji} - x_{ji}$. Then Q_i may be written

$$
\begin{aligned}
Q_i = {} & w_{11i}V_{1i}^2 + w_{12i}V_{1i}V_{2i} + \cdots + w_{1qi}V_{1i}V_{qi} \\
& + w_{21i}V_{2i}V_{1i} + w_{22i}V_{2i}^2 + \cdots + w_{2qi}V_{2i}V_{qi} \\
& + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
& + w_{q1i}V_{qi}V_{1i} + w_{q2i}V_{qi}V_{2i} + \cdots + w_{qqi}V_{qi}^2
\end{aligned} \tag{2}
$$

the w's being the weights, with $w_{jki} = w_{kji}$. Thus in the case of two variables, if N_i be the number of items in the ith sample, r_i its coefficient of correlation, and σ_{1i}, σ_{2i} its standard deviations, then

$$
w_{11i} = \frac{N_i}{2(1 - r_i^2)\sigma_{1i}^2}, \qquad w_{12i} = \frac{-N_i r_i}{2(1 - r_i^2)\sigma_{1i}\sigma_{2i}} = w_{21i}, \qquad w_{22i} = \frac{N_i}{2(1 - r_i^2)\sigma_{2i}^2}.
$$

The coefficients of the cross products in Q involve the coefficients of correlation of distributions. If the latter are all zero the cross products vanish and Q reduces to the sum of weighted squares, which is the basic expression of the least squares procedure. Consequently, from this point of view, the least squares assumption is equivalent to the assumption of zero correlation between the errors. The procedure in the more general situation might be called "least quadratics".

The Lagrange method of undetermined multipliers can be used to calculate the values of the adjustments to the coordinates and the values of the parameters. The procedure is the same as for least squares [2], [3], [5], the only difference being the somewhat greater complication of the algebra. We shall summarize the development here.

The condition equations, supposed ν in number, may be written

$$F^h(x_{11}, \cdots, x_{qn}; p_1, p_2, \cdots, p_r) = 0, \qquad h = 1, 2, \cdots, \nu,$$

where each F^h may in general involve any or all of the numbers x_{ji} as well as any or all of the parameters p_l, whose number we suppose to be r. Let

$$F^h_{ji} = \partial F^h/\partial x_{ji}, \qquad F^h_l = \partial F^h/\partial p_l \tag{3}$$

where the X's have been substituted for the x's after differentiation, and each p_l has been replaced by the best available approximate value p_{l0}. Let F^h_0 be the value of F^h after the same substitution. Also let $v_l = p_{l0} - p_l$. Then if the V's and v's are small the conditions may be written

$$\sum_i \sum_j F^h_{ji} V_{ji} + \sum_l F^h_l v_l = F^h_0, \qquad h = 1, 2, \cdots, \nu. \tag{4}$$

Differentiate Q with respect to the V's and equate the result to zero, eliminating the factor 2. Differentiate (4) with respect to the V's and the v's, multiply each equation by the corresponding undetermined multiplier $-\lambda_h$, and sum the results together with the result from differentiating Q. Collecting coefficients of the differentials δV_{ji} and δv_l, equating to zero and transposing the terms involving λ_h, we get

$$\begin{aligned} w_{11i}V_{1i} + w_{12i}V_{2i} + \cdots + w_{1qi}V_{qi} &= [\lambda_h F^h_{1i}] \\ w_{21i}V_{1i} + w_{22i}V_{2i} + \cdots + w_{2qi}V_{qi} &= [\lambda_h F^h_{2i}] \qquad (i = 1, 2, \cdots, n) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ w_{q1i}V_{1i} + w_{q2i}V_{2i} + \cdots + w_{qqi}V_{qi} &= [\lambda_h F^h_{qi}] \end{aligned} \tag{5}$$

$$[\lambda_h F^h_l] = 0 \qquad l = 1, 2, \cdots, r, \tag{6}$$

where the brackets denote summation with respect to h.

Equations (5) can be written down easily, since the coefficients w_{jki} appear in the same order as in (2). The equations corresponding to each i form a complete set which can be solved independently of those for other values of i. The solution can be expressed

$$\begin{aligned} V_{ji} &= A_{1ji}[\lambda_h F^h_{1i}] + A_{2ji}[\lambda_h F^h_{2i}] + \cdots + A_{qji}[\lambda_h F^h_{qi}] \\ &= \left[\lambda_h \sum_{k=1}^{q} A_{kji} F^h_{ki}\right] \end{aligned} \tag{7}$$

where A_{kji} is $(-1)^{k+j}$ times the minor corresponding to w_{kji}, divided by the principal determinant. By symmetry $A_{kji} = A_{jki}$.

The V's in (4) are to be replaced by their values from (7) and the coefficients of the λ's collected. To facilitate this let

$$L_{jk} = \sum_{i=1}^{n} L_{jki}$$

where

$$L_{jki} = \sum_{s=1}^{q} \sum_{r=1}^{q} A_{rsi} F^{j}_{ri} F^{k}_{si}.$$

Each L_{jki} can be written down easily from the corresponding Q_i as written in (2): in each term $w_{rsi}V_{ri}V_{si}$ replace w_{rsi} by A_{rsi}, V_{ri} by F^{j}_{ri}, and V_{si} by F^{k}_{si}. It is important to preserve the order of the subscripts of the V's in (2), and to treat the diagonal terms $w_{rri}V^2_{ri}$ as though written $w_{rri}V_{ri}V_{ri}$. It is seen that $L_{jki} = L_{kji}$, and $L_{jk} = L_{kj}$. Then the substitution from (7) into (4) gives

$$(8) \qquad \sum_{j=1}^{\nu} L_{jh}\lambda_j + \sum_{l=1}^{r} F^{h}_{l} v_l = F^{h}_{0} \qquad h = 1, 2, \cdots, \nu.$$

Equations (8), with (6), are formally identical with those of the least squares procedure which are called by Deming the "general normal equations", and they can be written schematically in the same manner. The further procedure is identical with that for least squares, involving solution of the general normal equations for the λ's and v's, substitution of the values of the λ's into (7) to obtain the V's, and then adjustment of the observations by use of the V's, and adjustment of the provisional values of the parameters by use of the v's.

A word of appreciation is due Dr. O. W. Richards of The Spencer Lens Company for calling this problem to my attention, and for encouragement in the carrying out of the solution.

REFERENCES

[1] JOSEPH BERKSON, "Growth changes in physical correlation—height, weight, and chest-circumference, males," *Human Biology*, Vol. 1(1929), pp. 462–502.

[2] W. EDWARDS DEMING, "The application of least squares," *Phil. Mag.* 7th Ser., Vol. 11(1931), pp. 146–158.

[3] W. EDWARDS DEMING, "On the application of least squares," *Phil. Mag.* 7th Ser., Vol. 17(1934), pp. 804–829.

[4] W. EDWARDS DEMING, *Proc. Phys. Soc. Lond.*, Vol. 47(1935), pp. 92–106.

[5] W. EDWARDS DEMING, *Some Notes on Least Squares*, Graduate School, Dept. of Agric., Washington.