So far the presence of *Eichornia crassipes* in the Nile has not proved disastrous, though it is making the already hard lives of some of the more primitive riverside inhabitants a little harder. So far it has proved advantageous in only the lower reaches of the River Sobat.

H. R. J. DAVIES

Department of Geography,
University of Khartoum.

[1] *Nature*, **182**, 538 (1958).

STATISTICS

Estimation of Linear and Non-linear Structural Relations

THE problem considered by Wayman[1] is one example of a wide class of problems which have given rise to a large body of literature in recent years. Lindley[2] has reviewed the field, and thirty subsequent papers are listed by Barton and David[3]. Of particular importance is the estimation of non-linear structural relationships. Several methods for estimating the unknown parameters are available; we outline below a method which will often give estimates of nearly optimal accuracy.

We suppose the observations to consist of n pairs (x_i, y_i) with:

$$x_i = X_i + u_i, \; y_i = f(X_i) + v_i \; (i = 1, 2, \ldots, n) \qquad (1)$$

where the function $f(X)$ contains unknown parameters to be estimated and where the variances s_i^2, t_i^2 of u_i, v_i may be (*a*) known, (*b*) known up to a constant factor, or (*c*) constant but unknown. Case (*b*) is no more difficult than (*a*); an estimate of the unknown factor can be obtained from the sum of squares of residual deviations. Further, it may be that the values X_i are (i) unknown parameters (to be estimated), or random variables whose common cumulative distribution $P(X)$ is (ii) known, (iii) of known form but with unknown parameters, or (iv) completely unknown.

Even in the linear case, various difficulties arise. Thus in case (*c*) the linear relationship may be *unidentifiable* (see ref. 4). Neyman and Scott[5] have shown that when the number of parameters to be estimated increases indefinitely with n (as in case (i)), the method of maximum likelihood (M–L) (that is, least-squares if the residuals are assumed to be normal (Gaussian)) is not necessarily *consistent*. Cases (ii) and (iii) involve only a fixed number of parameters. Jeffreys'[6] method may be regarded as a special case of (ii) with certain conventional assumptions regarding $P(X)$; it is not consistent unless these assumptions are in fact correct. Kiefer and Wolfowitz[7] have shown that in case (*c*)(iv), assuming identifiability and that the unknown variances are bounded away from zero, M–L yields consistent estimates of the parameters of the line and of $P(X)$.

In general, three methods other than least-squares are already available: (I) Berkson's assumption of the 'controlled variable' (see ref. 8) which reduces case (*c*) to case (*b*); (II) the 'method of moments' in which various relations deducible from $Y = f(X)$ are summed, the sums involving X and Y being then estimated from corresponding sums involving x and y; and (III) obvious extensions of the 'method of dichotomy' due originally to Bose[9]. (I) may be inappropriate; (II) and (III) are consistent (when this is possible) but inefficient.

In cases (ii) and (iii), it is always possible in principle to find the joint distribution of x_i, y_i, and X_i, and to average over X_i; then M–L can be used on the resulting distribution of x_i and y_i. This approach will usually lead to very intractable equations. Barton and David[3] have proposed the following approach, which is certainly workable when $f(X)$ is a polynomial, and is nearly optimal when $f(X)$ is substantially linear.

From (1) and $P(X)$ we can find the mean and variance of y conditional on x, say:

$$E(y \mid x) = a(x), \; \text{Var}\,(y \mid x) = b(x). \qquad (2)$$

Then we minimize:

$$\varphi = \Sigma_i \left\{ \log\, b(x_i) + (y_i - a(x_i))^2 / b(x_i) \right\} \qquad (3)$$

with respect to all the unknown parameters. The argument proving joint asymptotic normality of the resulting estimators follows closely that for M–L estimators, the analytic conditions being a straightforward modification of these. When $f(X)$ is quadratic this procedure leads to three non-linear simultaneous equations which can be solved by iteration.

D. E. BARTON
C. L. MALLOWS

Department of Statistics,
University College,
London.
Sept. 22.

[1] Wayman, P. A., *Nature*, **184**, 77 (1959).
[2] Lindley, D. V., *J.R. Statist. Soc.*, B, **9**, 218 (1947).
[3] Barton, D. E., and David, F. N., Bull. 31st Session I.S.I. (1958) (in the press).
[4] Reiersøl, O., *Econometrica*, **18**, 375 (1950).
[5] Neyman, J., and Scott, E. L., *Ann. Math. Statist.*, **22**, 352 (1951).
[6] Jeffreys, H., 'Theory of Probability', 2nd Edition (Cambridge Univ. Press, 1948).
[7] Kiefer, J., and Wolfowitz, J., *Ann. Math. Statist.*, **27**, 887 (1956).
[8] Geary, R. C., *J. Amer. Stat. Assoc.*, **48**, 94 (1953).
[9] Bose, S. S., *Sankhya*, **3**, 339 (1938).

A Least-Squares Solution for a Linear Relation between Two Observed Quantities

IN a recent communication with the above title, P. A. Wayman[1] presents a solution to the problem of fitting a straight line when both co-ordinates are subject to error. He mentions some previous attempts to solve this problem, but is evidently unaware of the existence of a monograph by W. Edwards Deming[2].

Deming presents a completely general method for fitting experimental results by least squares and, when this general method is applied to the specific problem studied by Wayman, the same result is obtained. The statement that the solution only passes through the centre of gravity if this is found by applying a weight w_r to each point is also found explicitly made by Deming[3] (Wayman's w_r is identically the same as Deming's W).

This republication of a result published first about twenty years ago suggests that Deming's excellent monograph is not as well known to scientists generally as it deserves to be.

B. K. KELLY

Antibiotics Research Station
(Medical Research Council),
4 Elton Road,
Clevedon, Somerset.
Aug. 26.

[1] Wayman, P. A., *Nature*, **184**, 77 (1959).
[2] Deming, W. Edwards, "Statistical Adjustment of Data", (Chapman and Hall, London, 1943).
[3] Deming, W. Edwards, *loc. cit.*, p. 181.