After substitution and equation of (8) and (9) we find that $y_0$ is the solution of

$$(2\alpha-1)\exp(-qy_0)\,\pi(y_0)+(1-\alpha)\exp(-qy_0) = \alpha(2\alpha-1)\,\pi\{(\beta-q)\,y_0/\beta\}+\alpha(1-\alpha), \quad (10)$$

where

$$\alpha = (1-q/\beta)^{n+1}.$$

*Appl. Statist.* (1974),
**23**, No. 1, p. 73.

# An Iterative Computer Procedure for Mixed-up Values in Experiments

By D. A. Preece
*University of Kent at Canterbury*
and
J. C. Gower
*Rothamsted Experimental Station, Harpenden, Hertfordshire*

Summary

Healy and Westmacott (1956) gave an iterative computer procedure for dealing with missing values in experiments. Pearce (1965) and others discussed an improved version of the procedure. This improved version is now extended to deal with inadvertently pooled values.

*Keywords*: COMPUTER ALGORITHM; ITERATIVE PROCEDURE; MISSING VALUES; MIXED-UP VALUES

Healy and Westmacott (1956) gave an iterative computer procedure for dealing with missing values in experiments. This procedure first inserts "guessed" values for the missing units. Residuals are then calculated, and a new value formed for each missing unit by subtracting its current residual from the current estimated value. Residuals are then calculated afresh, and so on, until the residuals for the missing units are small enough. This process always converges as required.

Pearce (1965, Section 7.3), Pearce and Jeffers (1971) and Preece (1971) discussed an improved Healy–Westmacott process. In any iteration of this, the correction subtracted from an estimated missing value is $n/E$ times the corresponding residual, where $n$ is the number of units in the experiment as planned and $E$ is the number of error degrees of freedom as planned.

Several authors have discussed "mixed-up" values, i.e. inadvertently pooled observations whose total value is known but whose individual values are not. Mixed-up values arise, for example, on harvesting a field experiment, if produce from different plots is bulked before the individual plot yields have been noted.

Missing values can be estimated non-iteratively by minimizing the error sum of squares (Yates, 1933); mixed-up values can be estimated similarly, except that the minimization is now subject to the constraint that the estimates must add to the observed total. The residuals corresponding to estimated missing values are all zero; the residuals corresponding to estimated mixed-up values are all the same (but not, in general, zero).

The following algorithm based on the improved Healy–Westmacott process is proposed for the estimation of mixed-up values:

1. Let the initial "estimate" for each of the mixed-up values be $k/n'$, where $n'$ is the number of mixed-up values and $k$ is their total.
2. Calculate residuals.
3. For each mixed-up plot, calculate the "adjusted residual", defined as the residual minus the average of the residuals for all the mixed-up plots.
4. Form a new estimate for each mixed-up plot by subtracting $n/E$ times the current adjusted residual from the current estimate.
5. If the $n'$ adjusted residuals are small enough, stop; otherwise, return to 2.

It can be shown that the estimates from this algorithm always converge to the estimates from the constrained minimization. (Configurations of missing or mixed-up values that produce singularity in the analysis are ignored in this note.) Furthermore, if we use a general multiplier $m$ instead of the multiplier $n/E$, the range of values of $m$ that produces convergence is the same as if the troublesome units were missing, not mixed up. For example, for a randomized block design with $p$ blocks, each a replicate of $q$ treatments, we have convergence as shown in Table 1.

TABLE 1

| *Disposition of missing or mixed-up values* | *Range of convergence* |
|---|---|
| All in the same block | $0<m<2p/(p-1)$ |
| All for the same treatment | $0<m<2q/(q-1)$ |
| All in different blocks and for different treatments | $0<m<2pq/\{(p-1)(q-1)-1\}$ |

Obvious extensions of the algorithm could deal with variates where there is more than one set of mixed-up values or where there are both missing and mixed-up values. It is, however, unlikely that it would be worth while to include such extended procedures in standard computer programs.

## REFERENCES

HEALY, M. J. R. and WESTMACOTT, M. H. (1956). Missing values in experiments analysed on automatic computers. *Appl. Statist.*, **5**, 203–206.

PEARCE, S. C. (1965). *Biological Statistics: An Introduction.* New York: McGraw-Hill.

PEARCE, S. C. and JEFFERS, J. R. N. (1971). Block designs and missing data. *J. R. Statist. Soc.* B, **33**, 131–136.

PREECE, D. A. (1971). Iterative procedures for missing values in experiments. *Technometrics*, **13**, 743–753.

YATES, F. (1933). The analysis of replicated experiments when the field results are incomplete. *Empire J. Exp. Agric.*, **1**, 129–142.