Algorithm AS 18

Evaluation of Marginal Means

By J. C. Gower

Rothamsted Experimental Station

Language

Algol 60.

Description and Purpose

This procedure reads, or transfers in some other way, the values of an n-way table, allotting space for margins, and then fills the marginal cells with marginal means. It uses the subroutine package described in AS 1 (Gower, 1968) for handling multiway tables as though they were one-way. The revised form of *checkset* (see remark AS R1) is assumed.

Method

The method is illustrated for a three-way table with values x_{pqr} and levels $0-P$, $0-Q$, $0-R$ respectively, where P, Q and R are the levels corresponding to the margins. Suffices for marginal cells are written in the usual statistical notation, e.g. $x_{p..}$. First evaluate $x_{.qr} = 1/P\sum_{p=0}^{P-1} x_{pqr}$ for all $0 \leqslant q < Q, 0 \leqslant r < R$. Secondly evaluate $x_{p.r} = 1/Q\sum_{q=0}^{Q-1} x_{pqr}$ for all $0 \leqslant p \leqslant P, 0 \leqslant r < R$. Finally evaluate $x_{pq.} = 1/R\sum_{r=0}^{R-1} x_{pqr}$ for all $0 \leqslant p \leqslant P, 0 \leqslant q \leqslant Q$. Note that the marginal levels are excluded in all summations and that, when evaluating $x_{p.r}$, all terms of the form $x_{..r}$ are also evaluated because p is allowed to take its marginal value P. Similarly when evaluating $x_{pq.}$ we also get $x_{p..}$, $x_{.q.}$ and $x_{...}$. The exclusion of marginal levels is controlled by the "exclude" parameter j in the final *scan* call of the procedure. This avoids additions whose results will be overwritten at a later stage (e.g. the first set of summations would include $\sum_{p=0}^{P-1} x_{p.r}$ if q were allowed to take the value Q) and is essential when there is any possibility of the marginal cells initially containing values in non-standard form for real (floating point) numbers; no value is taken from any marginal cell until a valid mean has been placed in it.

Structure

The procedure call statement is:

means (n, F, X, *transfer*, *ifault*)

Formal parameters

n	Integer	input:	the number of factors.
F	Integer array	input:	$F[i]$ is the number of levels of factor i (excluding marginal levels). When every factor has the same number of levels L (say) the device (described in

			AS 1) may be used, of restricting F to a single element $F[1]$ set to $L+RC$.
X	Real array	output:	a one-way array of length $\prod_{i=1}^{n}(F[i]+1)$ containing the values $x(l_1, l_2, \ldots, l_n)$ of an n-way table stored in the order implied by equation 4 of algorithm AS 1.
ifault	Integer	output:	the fault settings are the same as those for procedure *setup* described in AS 1. The procedures *address* and *scan* are also used, but with a dummy fault indication because, provided *setup* was called correctly, no further faults can occur. Thus *ifault* $= 1$ if $n \leqslant 0$ otherwise *ifault* $= 0$.
transfer	Procedure		see auxiliary algorithms, below.

Auxiliary algorithms

The procedures *setup*, *scan* and *address* described in AS 1 are used together with the global identifiers which they operate on, i.e. **integer array** L, M; **integer** s, and the local constants RC (regular case) and FS (factor set).

The user must supply a procedure *transfer*(y) which returns the next data-value in the standard order implied by equation 1 of AS 1.

Formal parameters

y	Real	output:	the name of the identifier to which the next value is to be assigned.

Transfer may, for example, represent an external transfer using read statements, or an internal transfer from a table without marginal spaces to X which has margins. If the latter, it is often convenient to store the data-values initially at the bottom end of X and use *transfer* to spread them out over X, allowing the correct gaps for the margins.

After *transfer* has been used, the procedure *address* (see AS 1) can be used to give any cell, including margins, of X; the marginal level corresponding to factor i is $F[i]$.

References

Gower, J. C. (1968). Simulating multidimensional arrays in one dimension (Algorithm AS 1). *Appl. Statist.*, **17**, 180–185.

—— (1969). Remark AS R1, on AS 1 subroutine package. *Appl. Statist.*, **18**, 116.

```
procedure means(n, F, X, transfer, ifault);

comment Algorithm AS18 J.R.statist.Soc. C, (1969), Vol.18, No.2;

value n; integer n, ifault; integer array F; real array X;
procedure transfer;

comment The procedures  setup,  address and scan  described in AS1 are
used. FS, RC and s are global identifiers described in AS1. The values
are first transferred to X by the procedure  transfer, supplied by the
user;

  begin
  setup(n, F, 1, ifault);
```

```
if ifault = 0 then
  begin integer i, j, t; real a;
  j := 2 ∧ n - 1;
L:transfer(X[s]); scan(0, j + FS, L, 1, t);

  comment The marginal means are next filled in by taking the
  factors in turn from 1 to n. Margins not yet filled in are
  excluded from the summations by the parameter j. This avoids
  unnecessary summations and is essential if there is any
  possibility of the marginal cells containing values not in the
  standard form for real (floating point) numbers;

  for i := 1 step 1 until n do
    begin
  P:a := 0;
SUM:a := a + X[s]; scan(i, i, SUM, 0, t);
    X[address(i, F, 0, t)] :=
    a / (if F[1] ≥ RC then F[1] - RC else F[i]);
    scan(i, j + FS, P, 1, t); j := j ÷ 2
    end
  end
end means
```

Algorithm AS 19

Analysis of Variance for a Factorial Table

By J. C. Gower

Rothamsted Experimental Station

Language

Algol 60.

Description and Purpose

Given an n-way table with margins filled with the marginal means, this procedure produces tables of corrected sums of squares, and associated degrees of freedom, for all main effects and interactions. The subroutine package AS 1 (Gower, 1968) for handling multi-way tables as if they were one-way is used. The modified version of *checkset* described in AS R1 is assumed.

Method

The method is illustrated for a three-way table with values x_{pqr} and levels $0-P$, $0-Q$, $0-R$ respectively, where P, Q and R are the levels corresponding to the margins. First crude sums of squares are evaluated in the order

$$S_0, S_R, S_Q, S_{RQ}, S_P, S_{PR}, S_{PQ}, S_{PQR},$$

where S_0 is the correction factor $PQR\,x^2...$, $S_{RQ} = P\sum_{q=0}^{Q-1}\sum_{r=0}^{R-1} x^2_{.qr}$, and similarly for the other terms. These values could be obtained in the order $S_0, S_P, S_Q, S_{PQ}, \ldots$ by renaming the factors.

The crude sums of squares and degrees of freedom both form 2^3 tables and can be converted into corrected tables by an operation similar to the Yates algorithm for computing the effects of a 2^n experiment. The successive steps, the results of which overwrite the previous values, are: