50 Years of Genstat ANOVA

Roger Payne

VSN International, 2 Amberside,
Wood Lane, Hemel Hempstead, UK

email: Roger.Payne@vsni.co.uk

VSNi

Rotorua 7th December 2018

QMS ISO 9001 REGISTERED

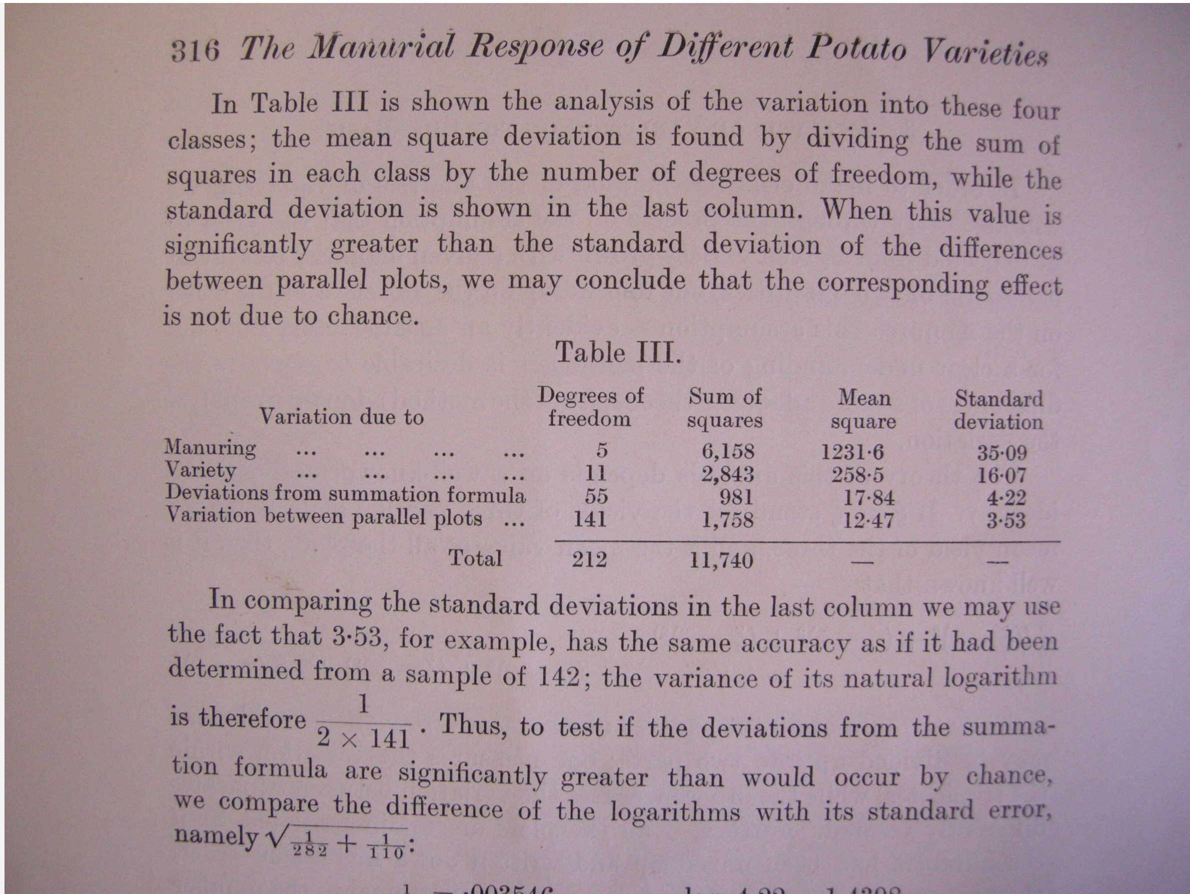
QMS ISO 14001 REGISTERED

It all started with Fisher

- ANOVA
  - Fisher & MacKenzie (1923) Studies in crop variation II. The manurial response of different potato varieties J. Agric. Sci.
  - Fisher (1934) Discussion to 'Statistics in agricultural research' J. Roy. Statist. Soc., Suppl. *"The analysis of variance is not a mathematical theorem, but rather a convenient method of arranging the arithmetic."*
- also started the strong links between Adelaide & Rothamsted
  - Fisher "retired" to Adelaide in 1957
  - John Nelder visited the Waite in 1965-6
  - Graham Wilkinson worked at Rothamsted from 1971-5
  - John Gower worked for CSIRO DMS in Adelaide in 1975
  - I worked for DMS in Adelaide 1978-9 (and visited New Zealand in May 79)
  - ..

VSNi

ANOVA



316 *The Manurial Response of Different Potato Varieties*

In Table III is shown the analysis of the variation into these four classes; the mean square deviation is found by dividing the sum of squares in each class by the number of degrees of freedom, while the standard deviation is shown in the last column. When this value is significantly greater than the standard deviation of the differences between parallel plots, we may conclude that the corresponding effect is not due to chance.

Table III.

Variation due to	Degrees of freedom	Sum of squares	Mean square	Standard deviation
Manuring	5	6,158	1231·6	35·09
Variety	11	2,843	258·5	16·07
Deviations from summation formula	55	981	17·84	4·22
Variation between parallel plots ...	141	1,758	12·47	3·53
Total	212	11,740	—	—

In comparing the standard deviations in the last column we may use the fact that 3·53, for example, has the same accuracy as if it had been determined from a sample of 142; the variance of its natural logarithm is therefore $\frac{1}{2 \times 141}$. Thus, to test if the deviations from the summation formula are significantly greater than would occur by chance, we compare the difference of the logarithms with its standard error, namely $\sqrt{\frac{1}{282} + \frac{1}{110}}$:

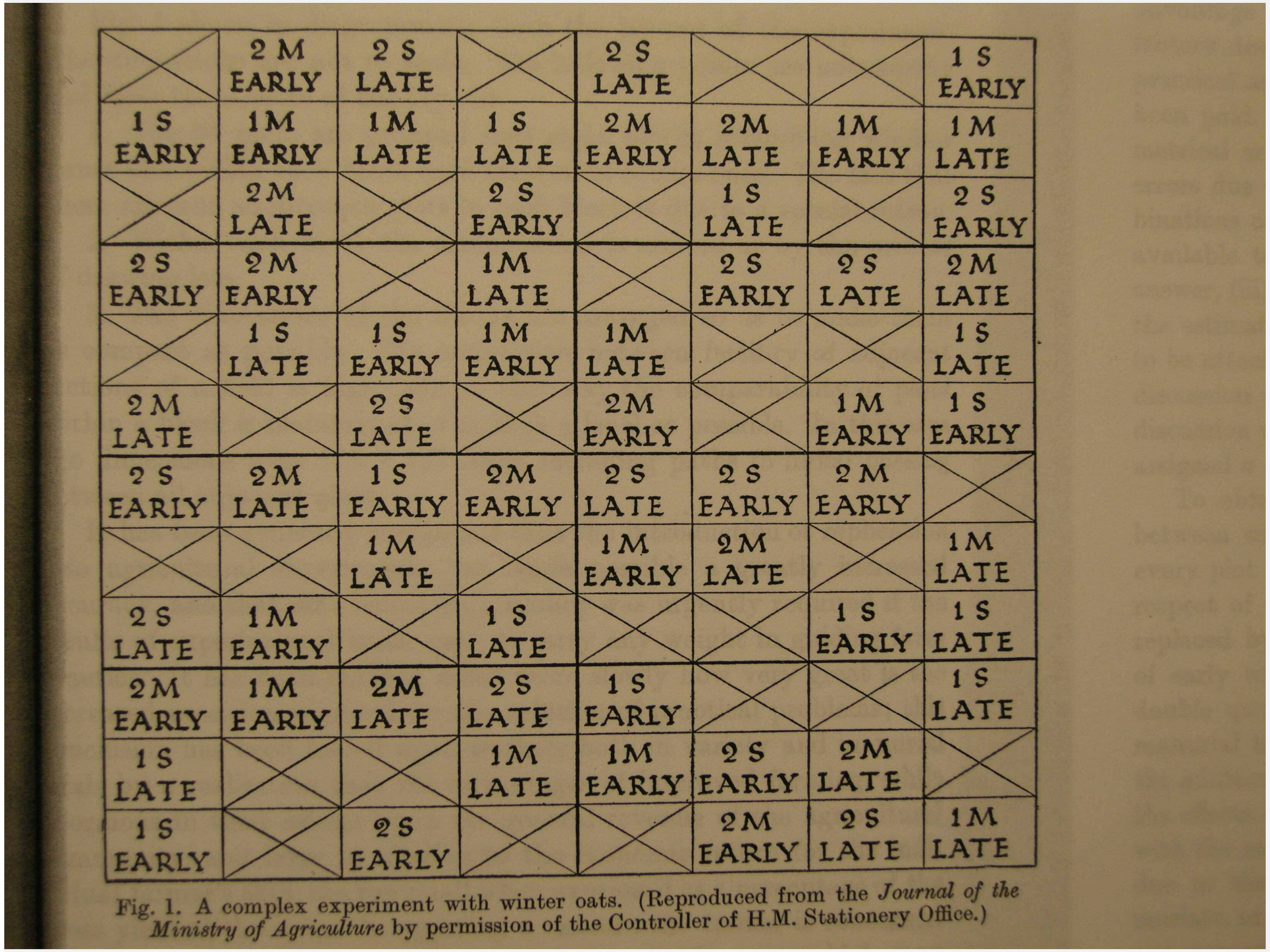
- first reference to analysis of variance (variation *sic.*)
- 12 varieties ×2 dung (+,–) ×3 fertilizers (basal, sulphate, chloride)
- ignored block structure (half-field / (plot * row))
- fitted main effects of variety and manures, and their interaction tested by using approximate Normality of log(variance)
- then fitted a multiplicative model (by eigenvalue decompositions)

Design of Experiments started with Fisher too..

- Fisher (1926) The arrangement of field experiments. J. Min. Ag. G. Br.
  - systematic designs introduce *"a flagrant violation of the conditions upon which a valid estimate {of error} is possible" ... "The estimate of error is valid because, if we imagine a large number of different results obtained by different* ***random*** *arrangements, the ratio of the real to the estimated error, calculated afresh for each of these arrangements, will be actually distributed in the theoretical distribution by which the significance of the result is tested."*
  - *"It would be exceedingly inconvenient if every field trial had to be preceded by a succession of even ten uniformity trials; consequently since the only purpose of these trials is to provide an estimate of the standard error, means have been devised for obtaining such an estimate from the actual yields of the trial year. The method adopted is that of* ***replication****."*
  - *"No aphorism is more frequently repeated in connection with field trials, than that we should ask nature few questions or, ideally, one question, at a time. The writer is convinced that this view is mistaken. Nature, he suggests, will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other topic has been discussed."*
  - also introduced **blocking** and use of the Latin square

VSNi

First design used in practice



	2 M EARLY	2 S LATE		2 S LATE			1 S EARLY
1 S EARLY	1 M EARLY	1 M LATE	1 S LATE	2 M EARLY	2 M LATE	1 M EARLY	1 M LATE
	2 M LATE		2 S EARLY		1 S LATE		2 S EARLY
2 S EARLY	2 M EARLY		1 M LATE		2 S EARLY	2 S LATE	2 M LATE
	1 S LATE	1 S EARLY	1 M EARLY	1 M LATE			1 S LATE
2 M LATE		2 S LATE		2 M EARLY		1 M EARLY	1 S EARLY
2 S EARLY	2 M LATE	1 S EARLY	2 M EARLY	2 S LATE	2 S EARLY	2 M EARLY	
		1 M LATE		1 M EARLY	2 M LATE		1 M LATE
2 S LATE	1 M EARLY		1 S LATE			1 S EARLY	1 S LATE
2 M EARLY	1 M EARLY	2 M LATE	2 S LATE	1 S EARLY			1 S LATE
1 S LATE			1 M LATE	1 M EARLY	2 S EARLY	2 M LATE	
1 S EARLY		2 S EARLY			2 M EARLY	2 S LATE	1 M LATE

Fig. 1. A complex experiment with winter oats. (Reproduced from the *Journal of the Ministry of Agriculture* by permission of the Controller of H.M. Stationery Office.)

- factorial in randomized blocks
- 2^3 + control (replicated 4 times)

- Eden & Fisher (1927) Studies in crop variation IV The experimental determination of the value of top dressings with cereals *J. Agric. Sci.*

VSNi

Analysis of variance

Table IV.

	Degrees of freedom	Sum of squares	
		Grain	Straw
Blocks	7	2,286·4	27,556·8
Treatments	8	387·0	18,667·1
Errors	24	773·2	5,491·2
	56	2,508·8	18,556·3
Total	95	5,955·4	70,271·4

- distinguished between
  - plot error (24 d.f.) from within-block replicates of null control
  - block-treatment interaction ("differential responses")

..

Analysis of variance

Output

Analysis of variance

Variate: Grain

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Blocks stratum	7	2286.44	326.63	10.14	
Blocks.Plots stratum					
Treatments	8	387.01	48.38	1.50	0.209
Blocks.Treatments	56	2508.72	44.80	1.39	0.189
Residual	24	773.21	32.22		
Total	95	5955.39			

- block-treatment interaction not significant
  - so can combine errors

..

VSNi

Analysis of variance

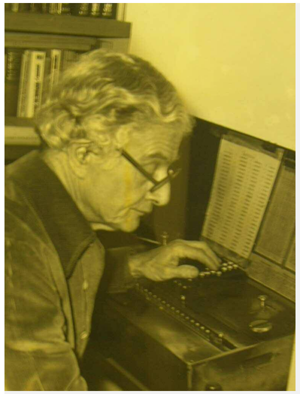
Output

Analysis of variance

Variate: Grain

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Blocks stratum	7	2286.44	326.63	8.04	
Blocks.Plots stratum					
Anytreat	1	191.75	191.75	4.72	0.033
Anytreat.Amount	1	24.84	24.84	0.61	0.437
Anytreat.Timing	1	76.84	76.84	1.89	0.173
Anytreat.Type	1	17.67	17.67	0.43	0.512
Anytreat.Amount.Timing	1	35.07	35.07	0.86	0.356
Anytreat.Amount.Type	1	7.48	7.48	0.18	0.669
Anytreat.Timing.Type	1	23.31	23.31	0.57	0.451
Residual	81	3292.00	40.64		
Total	95	5955.39			

- significant effect of nitrogen
  - but not of differences in Amount, Timing or Type



Later work - Frank Yates

- Yates (1933) The principles of orthogonality and confounding in designed experiments. *J. Agric. Sci*
  - (1933) *"since it is logically impossible that an interaction should exist without a main effect, the significance of main effects should be tested strictly on the assumption that their interactions are negligible"* (c.f. Nelder, 1977, A reformulation of linear models. *JRSS A*)
  - confounding of main effects – split plots, strip plots, Latin square with additional treatment factors applied to rows, and to columns
  - confounding of interactions (to avoid blocks becoming too large)
- Yates (1935) Complex experiments. *J. Roy. Statist. Soc., Suppl.*
  - partial and balanced confounding
- Yates (1936) Incomplete randomized blocks. *Ann. Eugenics*
  - balanced incomplete blocks, **efficiency factor**
- Yates (1937) *Design and Analysis of Factorial Experiments*
  - includes **Yates' ANOVA algorithm** – orthogonal designs, one error term

The Yates algorithm

- Yates (1937) *Design and Analysis of Factorial Experiments,* p.15
- calculates *Yates effects* for 2***n* designs
- form a variate of treatment totals
- form next variate with sums of pairs of values in upper half, differences of pair of values in lower half
- do this *n* times
- final column has response totals for the treatment terms
- divide by replication of main effects to get Yates effects
- also works for factors with more than 2 levels, e.g. for 3 estimates linear and quadratic contrasts (see e.g. *Encyclopedia of Statistical Sciences*, Volume 9, pp. 659-662)
- essentially provides a formal process for the estimation of effects from differences of means

..

VSNi

The Yates algorithm

```
MATRIX    [ROWS=8; COLUMNS=8; VALUES=\
           1, 1, 0, 0, 0, 0, 0, 0,\
           0, 0, 1, 1, 0, 0, 0, 0,\
           0, 0, 0, 0, 1, 1, 0, 0,\
           0, 0, 0, 0, 0, 0, 1, 1,\
          -1, 1, 0, 0, 0, 0, 0, 0,\
           0, 0,-1, 1, 0, 0, 0, 0,\
           0, 0, 0, 0,-1, 1, 0, 0,\
           0, 0, 0, 0, 0, 0,-1, 1] SumAndDiff
" Form treatment totals "
TABULATE [CLASS=D,K,N] Yield; TOTALS=Totals
" Put totals into a variate (in standard order) "
VTABLE    Totals; VARIATE=Col[0]; CLASS=Class
VARIATE    [NVALUES=NVALUES(Col[0])] Col[1...4]
CALCULATE Col[1] = SumAndDiff *+ Col[0]
CALCULATE Col[2] = SumAndDiff *+ Col[1]
CALCULATE Col[3] = SumAndDiff *+ Col[2]
CALCULATE Col[4] = Col[3] / 16
PRINT      Class['N','K','D'],Col[]; FIELD=3(1),5(8);\
           HEAD=3(' '),'Yield','(1)','(2)','Total','Effect'
```

			Yield	(1)	(2)	Total	Effect
_	_	_	425	851	3172	9331	583.2
n	_	_	426	2321	6159	333	20.8
_	k	_	1118	2679	86	2271	141.9
n	k	_	1203	3480	247	105	6.6
_	_	d	1283	1	1470	2987	186.7
n	_	d	1396	85	801	161	10.1
_	k	d	1673	113	84	-669	-41.8
n	k	d	1807	134	21	-63	-3.9

The Yates algorithm

	Yield	(1)	(2)	Total	Effect
_ _ _	425	851	3172	9331	583.2
n _ _	426	2321	6159	333	20.8
_ k _	1118	2679	86	2271	141.9
n k _	1203	3480	247	105	6.6
_ _ d	1283	1	1470	2987	186.7
n _ d	1396	85	801	161	10.1
_ k d	1673	113	84	-669	-41.8
n k d	1807	134	21	-63	-3.9

```
 43
 44  TREATMENTSTRUCTURE N * K * D
 45  ANOVA    [PRINT=effects; TWOLEVEL=Yates] Yield
```

Tables of effects

Variate: Yield

N Y-effect 20.8, s.e. 6.48, rep. 16

K Y-effect 141.9, s.e. 6.48, rep. 16

D Y-effect 186.7, s.e. 6.48, rep. 16

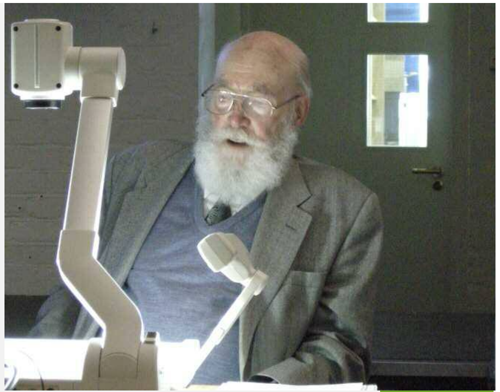
N.K Y-effect 6.6, s.e. 6.48, rep. 8

N.D Y-effect 10.1, s.e. 6.48, rep. 8

K.D Y-effect -41.8, s.e. 6.48, rep. 8

VSNi

Later work – John Nelder



- General balance
    - Nelder, J.A. (1965a,b). The analysis of randomized experiments with orthogonal block structure. *Proceedings of the Royal Society*, *Series A*, 283, 147-178.
    - Payne, R.W. (2014) Developments from analysis of variance through to generalized linear models and beyond. *Ann. Appl. Biol.* 164, 11–17. (*historical context*)
- conditions
    - block terms mutually orthogonal
    - treatment terms mutually orthogonal
    - single efficiency factor, for each treatment term, in each stratum
- close relationship to designs analysable by ANOVA
    - clarified by Payne & Tobias (1992)

..

VSNi

The ANOVA algorithm

- Wilkinson (1970) A general recursive algorithm for analysis of variance. *Biometrika*, 57, 19-46.
  - non-orthogonal designs
  - several error terms
  - K^{th} *order balance* (several efficiency factors for each model term)
- James & Wilkinson (1971). Factorisation of the residual operator and canonical decomposition of non-orthogonal factors in analysis of variance. *Biometrika*, 58, 279-294.
  - underlying theory
  - canonical efficiency factors
- Payne & Wilkinson (1977). A general algorithm for analysis of variance. *Applied Statistics*, 26, 251-260.
  - new dummy analysis
  - first-order balance only
  - pseudo-factors used for terms with more than one efficiency factor
- Payne & Tobias (1992) General balance, combination of information and the analysis of covariance. *Scand. J. Stats.*, 19, 3-23.
  - relationship of first order balance to general balance – does not require orthogonal block or treatment structures (but need these in practice for a full analysis)

VSNi

..

ANOVA in Genstat V Release 4

```
GENSTAT V    RELEASE 4.04B
COPYRIGHT 1984 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)

'REFERENCE'         ANOVA(4)
''
ANALYSIS OF SPLIT PLOT DESIGN
    (YATES,F: THE DESIGN AND ANALYSIS OF FACTORIAL EXPERIMENTS,
    TECHNICAL COMMUNICATION NO. 35 OF THE COMMONWEALTH
    BUREAU OF SOILS, P74)
A VARIETAL AND MANURIAL TRIAL OF OATS
6 BLOCKS OF 3 WHOLE PLOTS, ONE VARIETY GROWN ON EACH WHOLE PLOT
WHOLE PLOTS SPLIT INTO 4 SUB-PLOTS FOR NITROGEN
AREA OF EACH SUB PLOT WAS 1/80 ACRE
''
'UNITS' $72
'NAMES' N=0-CWT,0.2-CWT,0.4-CWT,0.6-CWT
      : V=VICTORY,GOLDRAIN,MARVLOUS
'FACTORS' BLOCKS$6
        : PLOTS$3
        : SUBPLOTS$4
        : NITROGEN$N=4,3,2,1,1,2,4,3,1,2,3,4,3,1,2,4,4,1,2,3,2,1,3,4,
                     2,3,4,1,4,2,3,1,1,4,2,3,3,4,1,2,1,3,4,2,2,3,4,1,
                     4,1,3,2,3,4,1,2,3,4,2,1,3,1,4,2,4,3,1,2,1,2,3,4
        : VARIETY$V=4(3,1,2,3,1,2,2,3,1,3,2,1,2,1,3,1,2,3)
'GENERATE' BLOCKS,PLOTS,SUBPLOTS
'READ/PRIN=DE' YIELD
'CALC' YIELD=(YIELD*80)/(112*4)
'HEADING' H='' OF OATS IN CWT. PER ACRE''
'DESCRIBE' YIELD $ 2;H
'BLOCKS' BLOCKS/PLOTS/SUBPLOTS
'TREATMENTS' VARIETY*NITROGEN
'ANOVA/PR=12012' YIELD
'PAGE'
'RUN'

156 118 140 105 111 130 174 157 117 114 161 141
104 70 89 117 122 74 89 81 103 64 132 133
108 126 149 70 144 124 121 96 61 100 91 97
109 99 63 70 80 94 126 82 90 100 116 62
96 60 89 102 112 86 68 64 132 124 129 89
118 53 113 74 104 86 89 82 97 99 119 121
'EOD'
```

ANOVA in Genstat V Release 4

```
***** ANALYSIS OF VARIANCE *****

VARIATE: YIELD OF OATS IN CWT. PER ACRE

SOURCE OF VARIATION                DF        SS      SS%         MS        VR

BLOCKS STRATUM                      5   506.227    30.54    101.245

BLOCKS.PLOTS STRATUM
  VARIETY                           2    56.963     3.44     28.482     1.485
  RESIDUAL                         10   191.751    11.57     19.175
TOTAL                              12   248.714    15.00     20.726

BLOCKS.PLOTS.SUBPLOTS STRATUM
  NITROGEN                          3   638.409    38.51    212.803    37.686
  VARIETY.NITROGEN                  6    10.260     0.62      1.710     0.303
  RESIDUAL                         45   254.106    15.33      5.647
TOTAL                              54   902.774    54.46     16.718

GRAND TOTAL                        71  1657.715   100.00

GRAND MEAN                           18.57
TOTAL NUMBER OF OBSERVATIONS            72


***** TABLES OF EFFECTS AND RESIDUALS *****

VARIATE: YIELD OF OATS IN CWT. PER ACRE


*** BLOCKS STRATUM ***

  BLOCKS                       RESIDUALS:          REP         12    SE      2.652

        BLOCKS       1        2        3        4        5        6
                  5.60    -1.04     0.59    -2.33    -1.44    -1.38
```

VSNi

ANOVA in Genstat V Release 4

```
*** BLOCKS.PLOTS STRATUM ***

 VARIETY                     EFFECTS:              REP         24   ESE       0.894

      VARIETY  VICTORY GOLDRAIN MARVLOUS
                 -1.13     0.09     1.04


 BLOCKS.PLOTS                RESIDUALS:            REP          4    SE       1.632

        PLOTS        1        2        3
       BLOCKS
            1    -2.04     2.50    -0.47
            2    -1.60    -0.06     1.66
            3     0.98     1.46    -2.44
            4    -2.05     0.72     1.33
            5    -1.73    -1.26     2.99
            6    -0.07    -1.17     1.24



*** BLOCKS.PLOTS.SUBPLOTS STRATUM ***

 NITROGEN                    EFFECTS:              REP         18   ESE       0.560

     NITROGEN    0-CWT  0.2-CWT  0.4-CWT  0.6-CWT
                 -4.39    -0.91     1.83     3.47


 VARIETY.NITROGEN            EFFECTS:              REP          6   ESE       0.970

     NITROGEN    0-CWT  0.2-CWT  0.4-CWT  0.6-CWT
      VARIETY
      VICTORY    -0.28    -0.51     0.53     0.26
     GOLDRAIN     0.01    -0.16    -0.01     0.16
     MARVLOUS     0.26     0.68    -0.51    -0.42
```

ANOVA in Genstat V Release 4

```
***** TABLES OF MEANS *****

VARIATE: YIELD OF OATS IN CWT. PER ACRE


    GRAND MEAN    18.57

       VARIETY  VICTORY GOLDRAIN MARVLOUS
                  17.43    18.66    19.61


      NITROGEN    0-CWT  0.2-CWT  0.4-CWT  0.6-CWT
                  14.18    17.66    20.40    22.03


      NITROGEN    0-CWT  0.2-CWT  0.4-CWT  0.6-CWT
       VARIETY
       VICTORY    12.77    16.01    19.79    21.16
      GOLDRAIN    14.29    17.59    20.48    22.29
      MARVLOUS    15.48    19.38    20.92    22.65



***** STANDARD ERRORS OF DIFFERENCES OF MEANS *****

TABLE                  VARIETY     NITROGEN      VARIETY
                                                NITROGEN
---------------------------------------------------
REP                         24           18            6
SED                      1.264        0.792        1.735
EXCEPT WHEN COMPARING MEANS WITH SAME LEVEL(S) OF:
  VARIETY                                          1.372


***** STRATUM STANDARD ERRORS AND COEFFICIENTS OF VARIATION *****

STRATUM                    DF            SE          CV%

BLOCKS                      5         2.905         15.6
BLOCKS.PLOTS               10         2.189         11.8
BLOCKS.PLOTS.SUBPLOTS      45         2.376         12.8
```

VSNi

Advantages of ANOVA

- works by an efficient sequence of *sweeps*, which
  - calculate and subtract effects from a working variate
  - project (i.e. *pivot*) effects into the space of a block-term
- no matrix inversion, other than for covariate estimation
- effects with natural (weighted) sum-to-zero constraints
- provides effects from each stratum where a term is estimated
- intuitive algorithm (simple combinability) to calculate treatment effects combining information from several strata, with d.f.
- provides an efficient engine for permutation tests etc
- clear and comprehensive output

..

VSNi

Sweeps

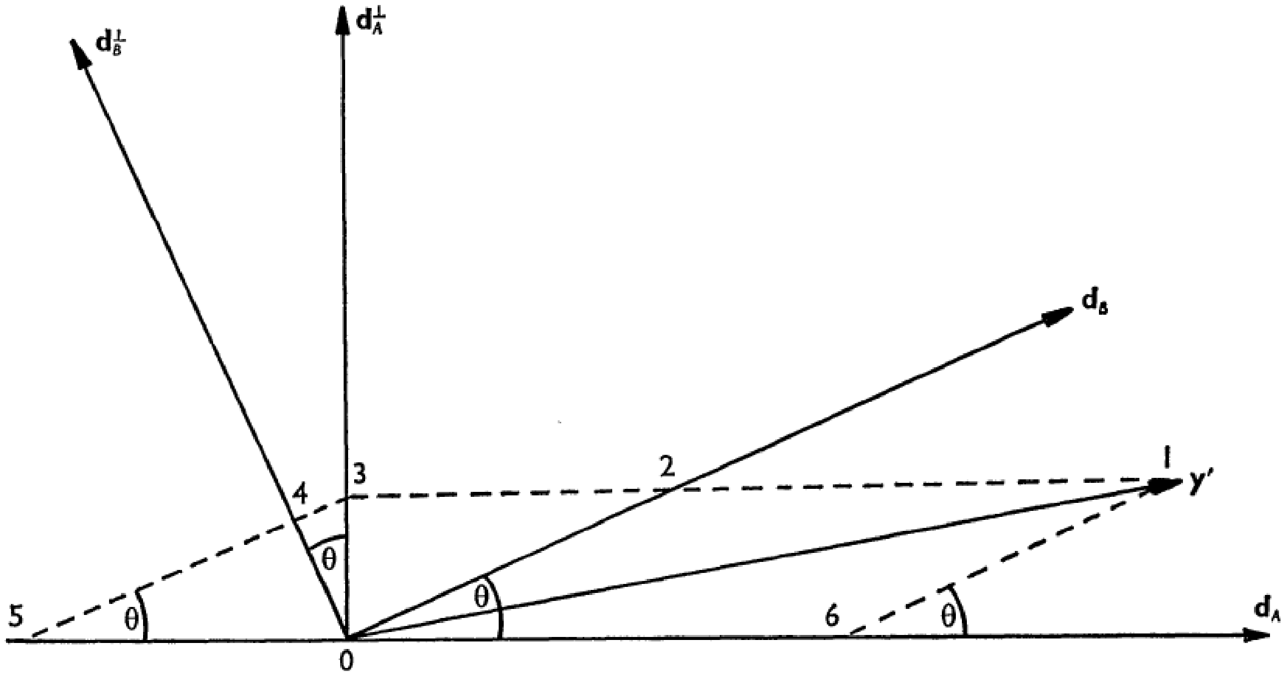


FIG. 1. A geometric representation of the sweeps required to analyse non-orthogonal terms *A* and *B*.

- sweep 1
  - estimate A as 31, projects y → 03 (= 01 – 31)
- sweep 2
  - estimate B should be 35 = 04 / sin(θ) = 43 / $\sin^2(\theta)$
    (so $\sin^2(\theta)$ is the efficiency factor), projects y → 05 (= 03 – 35)
- sweep 3
  - reanalysis sweep for A, y → 0 (= 05 – 05)

Example

- Example C-C-397: Balanced lattice in 4 replicates
  - 9 treatments
  - 3 blocks of size 3 in each replicate

Treatments on each unit of the design

Replicates	Blocks / Plots	1	2	3
1	1	1	4	7
	2	2	5	8
	3	3	6	9
2	1	1	2	3
	2	4	5	6
	3	7	8	9
3	1	1	2	3
	2	5	6	4
	3	9	7	8
4	1	1	2	3
	2	6	4	5
	3	8	9	7

Spreadsheet [Book;1]*

Row	_0	_1	_2	_3
1	1	1; 1; 1	1; 2; 4	1; 3; 7
2	2	1; 1; 2	1; 2; 5	1; 3; 8
3	3	1; 1; 3	1; 2; 6	1; 3; 9
4	4			
5	5	2; 1; 1	2; 2; 2	2; 3; 3
6	6	2; 1; 4	2; 2; 5	2; 3; 6
7	7	2; 1; 7	2; 2; 8	2; 3; 9
8	8			
9	9	3; 1; 1	3; 2; 2	3; 3; 3
10	10	3; 1; 5	3; 2; 6	3; 3; 4
11	11	3; 1; 9	3; 2; 7	3; 3; 8
12	12			
13	13	4; 1; 1	4; 2; 2	4; 3; 3
14	14	4; 1; 6	4; 2; 4	4; 3; 5
15	15	4; 1; 8	4; 2; 9	4; 3; 7
16	Row coordinates	Factors in table	Replicates; Blocks; Treatments	
17	Column coordinates	1	2	3

- represent treatments by factorial combinations of 2 pseudo-factors, A & B
- confounds A in rep 1, B in rep 2, AB in rep 3, AB^2 in rep 4

..

VSNi

Example

- Example C-C-397: Balanced lattice in 4 replicates

```
" sweep for grand mean "
ASWEEP  workvar; RESIDUAL=workvar; RSS=ss['Total']
" sweep for Replications "
ASWEEP  [TERM=Replications] workvar; RESIDUAL=workvar; EFFECTS=EffR; \
        SS=ss['Replications']
" sweep for Replications.Blocks "
ASWEEP  [TERM=Replications.Blocks] workvar; RESIDUAL=workvar; EFFECTS=EffRB
" sweep for Treatments with efficiency 0.75 "
ASWEEP  [TERM=Treatments; EFFICIENCY=0.75] workvar; RESIDUAL=workvar; \
        SS=ss['Treatments (0.75)']
" reanalysis sweep for Replications.Blocks "
ASWEEP  [TERM=Replications.Blocks] workvar; RSS=ss['Residual']

" Replications.Blocks stratum: pivot "
ASWEEP  [TERM=Replications.Blocks; EMETHOD=given; RMETHOD=replace] \
        workvar; RESID=workvar; EFF=EffRB
" sweep for Treatments "
ASWEEP  [TERM=Treatments; EFFICIENCY=0.25] workvar; RESIDUAL=workvar; \
        SS=ss['Treatments (0.25)']
" reanalysis sweep (pivot) for Replications.Blocks "
ASWEEP  [TERM=Replications.Blocks] workvar; RESIDUAL=workvar; \
        SS=ss['Replications.Blocks Residual']
```

VSNi

Example output

s.s. from asweep

Source	s.s.
Replications	0.07739
Treatments (0.25)	2.14478
Replications.Blocks Residual	0.00000
Treatments (0.75)	2.50193
Residual	0.83398
Total	5.96090

Analysis of variance

Variate: Y

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replications stratum	3	0.07739	0.02580		
Replications.Blocks stratum					
Treatments	8	2.14478	0.26810		
Replications.Blocks.*Units* stratum					
Treatments	8	2.50193	0.31274	4.05	0.008
Residual	16	1.23681	0.07730		
Total	35	5.96090			

VSNi

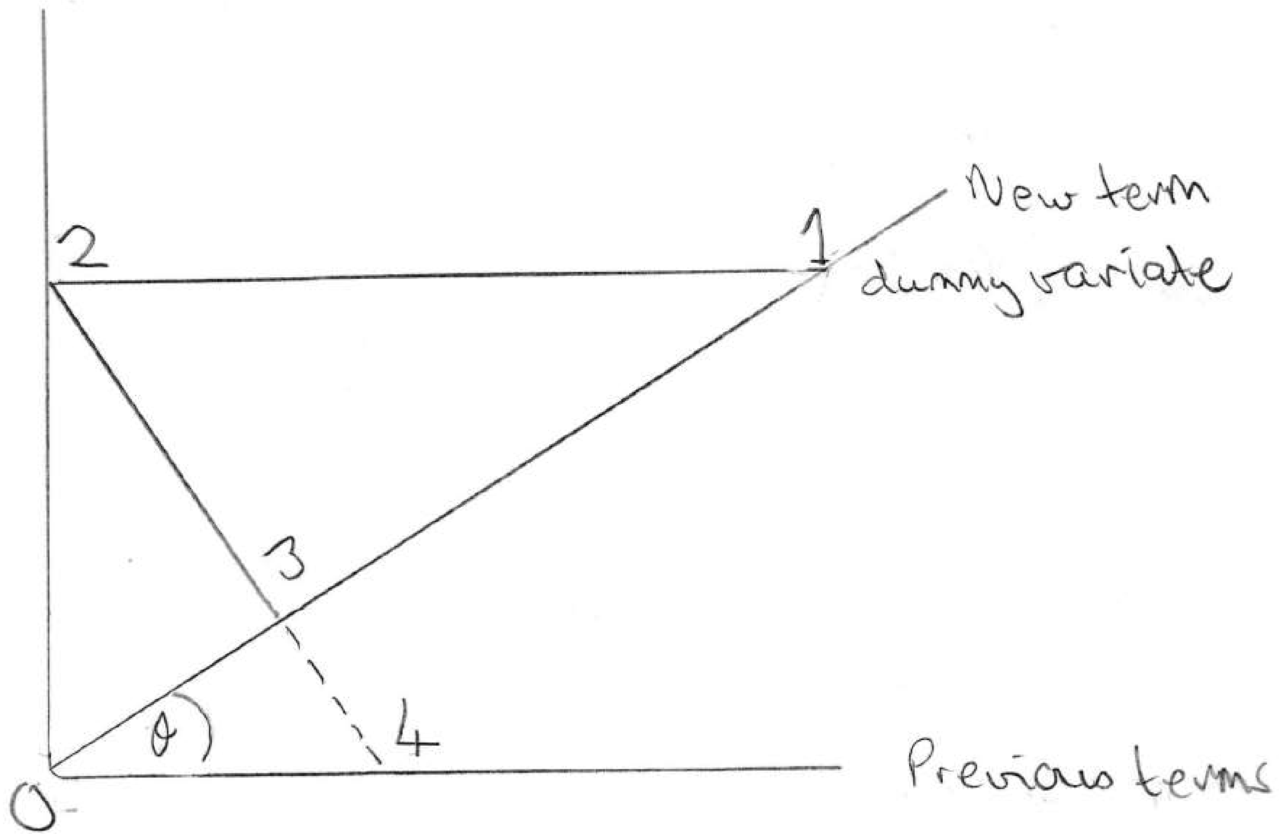
Original dummy analysis

(as I remember it!)

- for each model term, in turn
  - generate a dummy variate in its design space, note initial s.s. = ss_0
  - sweep for the previous terms
  - calculate s.s. for current term = ss_1
  - efficiency factor is ratio of original effects to new effects, calculate as sqrt(ss_1/ss_0)
  - sweep for the current term with efficiency factor
  - reanalysis sweeps for previous terms (non-orthogonal terms have non-zero s.s.)
  - term is balanced if final s.s. = 0
- (time-consuming for large models)

..

Original dummy analysis



$$O3/O1 = O2 \sin\theta / O1 = \sin^2\theta$$

VSNi

Original dummy analysis

```
" original dummy analysis e.g. balanced-incomplete-block design
  (Rogers & Wilkinson, 1974, Institute of Maths & Applications) "
FACTOR [LEVELS=3] B,T; VALUES=!(2(1,2,3)),!(2,3,1,3,1,2)
" form dummy variate (1) "
VARIATE [VALUES=2,4,0,4,0,2] D
" remove marginal term (grand mean) to put into space of T effects "
ASWEEP  D; RESIDUAL=Dummy
PRINT   B,T,D,Dummy
```

B	T	D	Dummy
1	2	2.000	0.000
1	3	4.000	2.000
2	1	0.000	-2.000
2	3	4.000	2.000
3	1	0.000	-2.000
3	2	2.000	0.000

Original dummy analysis

```
" sweep for T to get initial effects and s.s. (0->1) "
ASWEEP  [TERM=T] Dummy; SS=ss0; EFFECTS=eff0
" sweep for B (1->2) "
ASWEEP  [TERM=B] Dummy; RESIDUAL=Dummy
" sweep for T (2->3)  but leave Dummy at 2 "
ASWEEP  [TERM=T] Dummy; SS=ss1; EFFECTS=eff1
" calculate efficiency factor "
CALCULATE ef = SQRT(ss1/ss0)
PRINT   ss0,ss1,ef & eff0,eff1
```

ss0	ss1	ef
16.00	9.000	0.7500

T	eff0	eff1
1	-2.000	-1.500
2	0.000	0.000
3	2.000	1.500

```
" sweep for T with efficiency factor (2->4) "
ASWEEP  [TERM=T; EFFICIENCY=ef] Dummy; RESIDUAL=Dummy
" repeat sweep for B (4->0) "
ASWEEP  [TERM=B] Dummy; RESIDUAL=Dummy; SS=Bss; RSS=Check
" note non-zero B s.s. => non-orthogonal, zero Check => balanced "
PRINT   Bss,Check,Dummy
```

Bss	Check	Dummy
4.000	0	0
		0
		0
		0
		0
		0

Original dummy analysis

```
25  " confirm with ANOVA "
26  BLOCKS B
27  TREATMENTS T
28  ANOVA
```

Analysis of variance

Source of variation	d.f.
B stratum	
T	2
B.*Units* stratum	
T	2
Residual	1
Total	5

Information summary

Model term	e.f.	non-orthogonal terms
B stratum		
T	0.250	
B.*Units* stratum		
T	0.750	B

VSNi

New dummy analysis

- generate dummy variate containing all effects (Cauchy random numbers)
- sweep for grand mean, block terms and then treatment terms
- when a (new) term is found with a non-null sweep
    - remember effects = eff0, and s.s. = ss0 for new term
    - reanalyse for previous terms (note non-zero s.s. as non-orthogonal to new term)
    - re-estimate effects = eff1 and s.s. = ss1 for new term
    - if ss1 ≠ 0 term is non-orthogonal: efficiency = 1 − √(ss1 / ss0); term is balanced if ratio of new to old effects (eff1 / eff0) is constant; AN-1 if unbalanced
    - sweep for new term with efficiency factor
    - reanalyse for previous terms found to be non-orthogonal to new term
    - look for next new term
- for each block term (stratum
    - pivot into stratum and do sweeps as above, except that reanalysis sweep for the block term is a pivot
- (fewer sweeps, and better numerical accuracy than original method)

VSNi

New dummy analysis

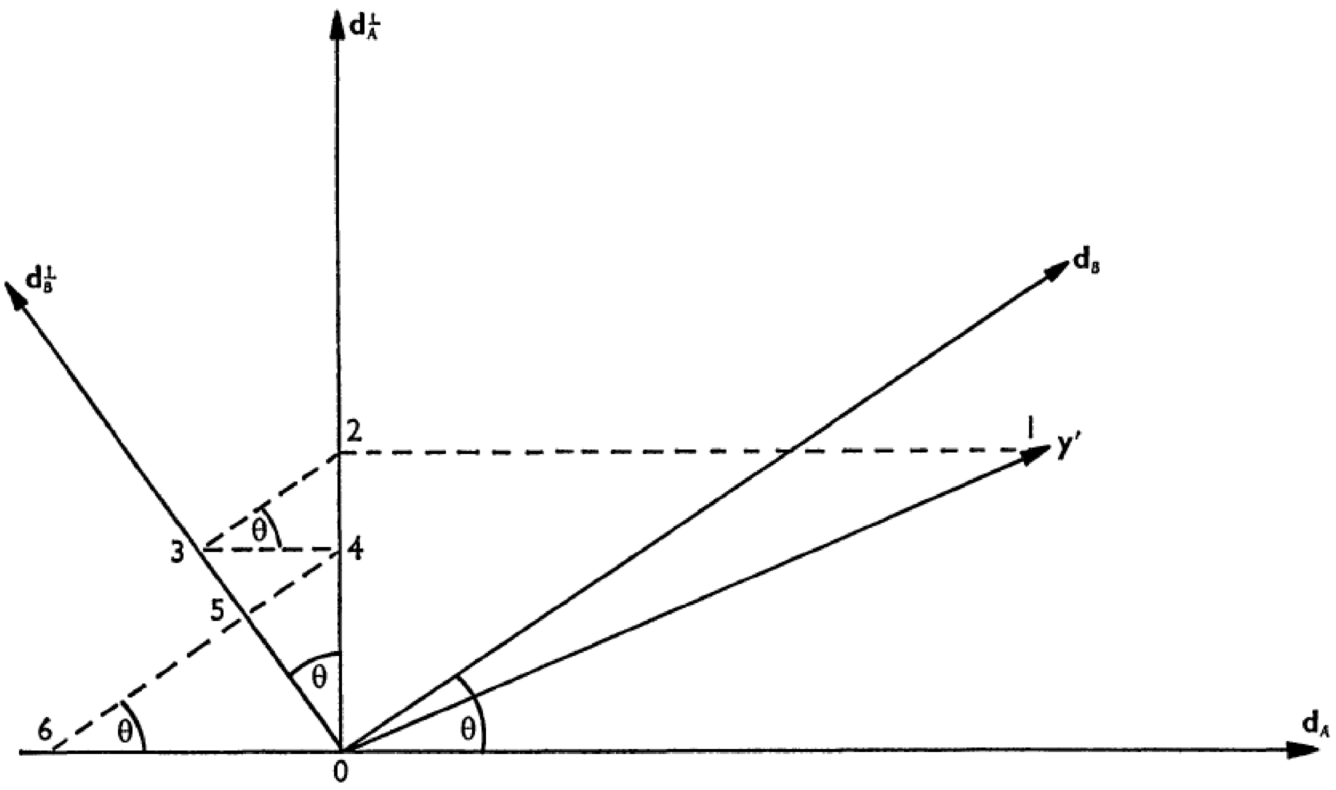


FIG. 2. A geometric representation of the sweeps required to determine the efficiency factor of *B*.

- 54 = 34 cos(θ) = 32 $\cos^2(\theta)$
- efficiency factor = $\sin^2(\theta)$ = 1 – 54 / 32

..

New dummy analysis

```
" new dummy analysis e.g. balanced-incomplete-block design
  (Rogers & Wilkinson, 1974, Institute of Maths & Applications) "
FACTOR [LEVELS=3] B,T; VALUES=!(2(1,2,3)),!(2,3,1,3,1,2)
" form dummy variate (0) "
CALCULATE [SEED=12345] Dummy = GRUNIFORM(NVALUES(T); 0; 1)
&           Dummy = TAN(C('pi')*(Dummy-0.5))
" sweep for grand mean (1) "
ASWEEP  Dummy; RESIDUAL=Dummy
" sweep for B (2) "
ASWEEP  [TERM=B] Dummy; RESIDUAL=Dummy
" sweep for T (3) "
ASWEEP  [TERM=T] Dummy; RESIDUAL=Dummy; SS=ss0; EFFECTS=eff0
" reanalysis sweep for B (4) "
ASWEEP  [TERM=B] Dummy; RESIDUAL=Dummy
" sweep for T (5) to get new s.s. and effects (leave dummy unchanged) "
ASWEEP  [TERM=T] Dummy; SS=ss1; EFFECTS=eff1
" calculate efficiency factor "
CALCULATE ef = 1 - SQRT(ss1/ss0)
" check balance "
CALCULATE Check = SUM((eff1-eff0*SQRT(ss1/ss0))**2)
PRINT   ef,Check,ss0,ss1 & eff0,eff1
```

ef	Check	ss0	ss1
0.7500	0	1.278	0.07987

T	eff0	eff1
1	-0.5393	-0.13482
2	-0.0487	-0.01217
3	0.5880	0.14700

```
" sweep for T with efficiency factor (6) "
ASWEEP  [TERM=T; EFFICIENCY=ef] Dummy; RESIDUAL=Dummy
" reanalysis sweep for B (0) "
ASWEEP  [TERM=B] Dummy; RESIDUAL=Dummy
" and could continue to next term... "
```

VSNi

New dummy analysis

- simplified (faster) dummy analysis for orthogonal designs
- to detect strata and terms estimated in the bottom stratum
  - sweep for grand mean, block terms and then treatment terms (estimable ones will have non-zero s.s.)
  - repeat sweeps and, provided all are null, design is orthogonal "except in particularly complicated designs where terms are aliased"
- to detect terms estimated in the stratum for each block term
  - pivot into stratum, and do sweeps for treatment terms (those estimated there will have non-zero s.s.)
  - repeat sweeps and check that all are null – if so, design is then (still) orthogonal
- if a non-null repeat sweep is found, design is non-orthogonal, and we need to switch to the full dummy analysis

..

VSNi

New dummy analysis

```
" new dummy analysis for orthogonal designs
  e.g. balanced-incomplete-block design
  (Rogers & Wilkinson, 1974, Institute of Maths & Applications
FACTOR [LEVELS=3] B,T; VALUES=!(2(1,2,3)),!(2,3,1,3,1,2)
" form dummy variate "
CALCULATE [SEED=12345] Dummy = GRUNIFORM(NVALUES(T); 0; 1)
&           Dummy = TAN(C('pi')*(Dummy-0.5))
" sweep for grand mean "
ASWEEP  Dummy; RESIDUAL=Dummy
" sweep for B  "
ASWEEP  [TERM=B] Dummy; SS=check; RESIDUAL=Dummy
IF check : PRINT 'B is estimable',check : ENDIF
```

```
                 check
B is estimable   0.3125
```

```
" sweep for T "
ASWEEP  [TERM=T] Dummy; RESIDUAL=Dummy; SS=check; EFFECTS=eff0
IF check : PRINT 'T is estimable',check : ENDIF
```

```
                 check
T is estimable   1.278
```

```
" sweep again for B "
ASWEEP  [TERM=B] Dummy; RESIDUAL=Dummy; SS=check
" sweep again for T  "
ASWEEP  [TERM=T] Dummy; SS=check; EFFECTS=eff1
IF check : PRINT 'design is non-orthogonal',check : ENDIF
```

```
                           check
design is non-orthogonal   0.07987
```

Some problems

- partial aliasing
  - algorithm fits each term as a whole, so it does not know if some of effects are not estimated (except for those with zero replication or already removed by marginal terms)
- partial confounding
  - different contrasts of a term may be fitted in different strata, but the algorithm has no way of knowing how many have been fitted in each one (so d.f. and s.e.'s will then be incorrect)
- single d.f. terms
  - numerical coincidences in the dummy variate may mean that the s.s. for some contrasts may be nearly zero, so 1 d.f. terms may (very occasionally) fail to be detected

..

Partial aliasing and confounding

- Cochran & Cox (1957) Experimental Design, 2nd Edition, page 412
  - 5x5 Simple lattice with different contrasts confounded in replicates 1 & 2
  - this basic design is duplicated in replicates 3 & 4

Spreadsheet [Book;1]*

Row	_0	_1	_2	_3	_4	_5	_6	_7	_8	_9	_10	_11
1	1	1	6	11	16	21		1	2	3	4	5
2	2	2	7	12	17	22		6	7	8	9	10
3	3	3	8	13	18	23		11	12	13	14	15
4	4	4	9	14	19	24		16	17	18	19	20
5	5	5	10	15	20	25		21	22	23	24	25
6	6											
7	7	1	6	11	16	21		1	2	3	4	5
8	8	2	7	12	17	22		6	7	8	9	10
9	9	3	8	13	18	23		11	12	13	14	15
10	10	4	9	14	19	24		16	17	18	19	20
11	11	5	10	15	20	25		21	22	23	24	25
12	Y coordinates	Factors	Variety									
13	X coordinates	1	2	3	4	5	6	7	8	9	10	11

- books generally present unrandomized plans (*not a good idea!*)

..

What could possibly go wrong..?

- (for simplicity analyse just reps 1 & 3)
- number blocks over the whole experiment
- ignore nested block structure
- ignore treatment confounding
- whoops..!

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replicates stratum	1	103.68	103.68	6.62	
BlocksOverExpt stratum					
Variety	24	133.80	5.57		
Residual	-15	108.52			
Replicates.BlocksOverExpt.*Units* stratum					
Variety	24	473.20	19.72	1.26	0.327
Residual	15	234.80	15.65		
Total	49	1054.00			

VSNi

What could possibly go wrong..?

- you've now been warned..!

Warning 2, code AN 16, statement 1 on line 15

Command: ANOVA [PRINT=aovtable; FPROBABILITY=yes] Yield
Partial aliasing.
BlocksOverExpt is partially aliasedDegrees of freedom, mean squares and standard errors will be incorrect.

Warning 3, code AN 18, statement 1 on line 15

Command: ANOVA [PRINT=aovtable; FPROBABILITY=yes] Yield
Negative residual degrees of freedom.
-15 residual degrees of freedom in the BlocksOverExpt stratum.

Warning 4, code AN 17, statement 1 on line 15

Command: ANOVA [PRINT=aovtable; FPROBABILITY=yes] Yield
Partial confounding.
Variety is partially confounded

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replicates stratum	1	103.68	103.68	6.62	
BlocksOverExpt stratum					
Variety	24	133.80	5.57		
Residual	-15	108.52			
Replicates.BlocksOverExpt.*Units* stratum					
Variety	24	473.20	19.72	1.26	0.327
Residual	15	234.80	15.65		
Total	49	1054.00			

VSNi

Some solutions

- check for partial aliasing in the orthogonal dummy analysis
  - calculate original s.s. for each term before starting analysis in each stratum
  - save s.s. for that term and its marginal terms during the dummy analysis
  - there is partial aliasing if sum of these s.s. is not equal to the original s.s.
- check that all the numbers of residual d.f. are non-negative
- the sum of the efficiency factors for each term must be ≤1
- ANOVA repeats the dummy analysis with different random numbers, and effects added in for aliased 1 d.f. terms
- define model to identify marginality
  - e.g. blocks/plots not blocks+plots
  - design menus save the correct models
  - analysis menus define the correct models for you
- use pseudo factors to identify partial confounding
  - design menus and commands can form these for you
- Payne, R.W. (1998). Detection of partial aliasing and partial confounding in generally balanced designs. *Computational Statistics*, 13, 213-226.

Getting it right

```
16  BLOCKS Replicates / BlocksOverExpt / Plots
17  TREATMENTS Variety // Blocks
18  ANOVA [PRINT=aovtable; FPROBABILITY=yes] Yield
```

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replicates stratum	1	103.68	103.68	3.82	
Replicates.BlocksOverExpt stratum					
Variety	4	133.80	33.45	1.23	0.422
Residual	4	108.52	27.13	2.31	
Replicates.BlocksOverExpt.Plots stratum					
Variety	20	473.20	23.66	2.02	0.063
Residual	20	234.80	11.74		
Total	49	1054.00			

- recognise the nested block structure
- identify the confounded contrasts

..

Combination of information

- obtain effects combining information from all the strata where a treatment term is estimated
  - i.e. same effects as estimated by REML
  - requires general balance
  - efficient to calculate – weighted average of the effects from each of the strata (simple combinability), weights are efficiency factors divided by stratum variances
- Payne & Tobias (1992) General balance, combination of information and the analysis of covariance. *Scand. J. Stats.*, 19, 3-23.
  - algorithm for estimating stratum variances (and thus variance components)
  - extension to analysis of covariance – combine covariate SSP's to estimate combined covariate regression coefficients

..

VSNi

Combination of information

```
" combination of information "
SPLOAD '%data%/slatehall.gsh'
BLOCK replicates/(rows*columns)
TREATMENTS variety
PDESIGN
ANOVA [PRINT=aovtable,stratumvariance,effects,cbeffect; FPROB=yes] yield
AKEEP replicates.rows + replicates.columns + replicates.rows.columns;\
      STRATUMVARIANCE=rvar,cvar,rcvar; COMPONENT=rvc,cvc,rcvc
AKEEP [STRATUM=replicates.rows] variety; EFFECTS=reffects; EFFICIENCY=refficiency
AKEEP [STRATUM=replicates.columns] variety; EFFECTS=ceffects; EFFICIENCY=cefficiency
AKEEP [STRATUM=replicates.rows.columns]\
      variety; EFFECTS=rceffects; EFFICIENCY=rcefficiency; CBEFFECTS=cbeffects

" show calculation - weighted average of effects from each stratum "
CALCULATE rweight,cweight,rcweight = refficiency,cefficiency,rcefficiency / rvar,cvar,rcvar
&     wdiv = rweight + cweight + rcweight
&     rweight,cweight,rcweight = rweight,cweight,rcweight / wdiv
&     cbcheck = rweight*reffects + cweight*ceffects + rcweight*rceffects
PRINT cbeffects,cbcheck

" compare with REML "
VCOMPONENTS [FIXED=variety] RANDOM=replicates/(rows*columns)
REML  [PRINT=model,components,means] yield
ADISPLAY [PRINT=cbmeans,stratumvariances]
```

VSNi

Combination of information

```
" show calculation - weighted average of effects from each stratum "
CALCULATE rweight,cweight,rcweight = refficiency,cefficiency,rcefficiency / rvar,cvar,rcvar
&      wdiv = rweight + cweight + rcweight
&      rweight,cweight,rcweight = rweight,cweight,rcweight / wdiv
&      cbcheck = rweight*reffects + cweight*ceffects + rcweight*rceffects
PRINT cbeffects,cbcheck
```

variety	cbeffects	cbcheck
1	-1.869	-1.869
2	0.786	0.786
3	-0.495	-0.495
4	-0.186	-0.186
5	0.628	0.628
6	0.570	0.570
7	-0.697	-0.697
8	-0.131	-0.131
9	-1.716	-1.716
10	-2.772	-2.772
11	-1.432	-1.432
12	0.133	0.133
13	1.486	1.486
14	-1.438	-1.438
15	0.276	0.276
16	-1.243	-1.243
17	0.277	0.277
18	1.217	1.217
19	1.991	1.991
20	1.695	1.695
21	0.230	0.230
22	1.739	1.739
23	-1.413	-1.413
24	0.760	0.760
25	1.602	1.602

Combination of information

Table of predicted means for variety

variety	1	2	3	4	5	6	7	8	9
	12.84	15.49	14.21	14.52	15.33	15.27	14.01	14.57	12.99

variety	10	11	12	13	14	15	16	17	18
	11.93	13.27	14.84	16.19	13.27	14.98	13.46	14.98	15.92

variety	19	20	21	22	23	24	25
	16.70	16.40	14.93	16.44	13.29	15.46	16.31

Standard error of differences: 0.6202

```
25  ADISPLAY [PRINT=cbmeans,stratumvariances]
```

Tables of combined means

Variate: yield

variety	1	2	3	4	5	6	7	8	9
	12.836	15.490	14.209	14.519	15.333	15.274	14.007	14.574	12.989

variety	10	11	12	13	14	15	16	17	18
	11.932	13.272	14.838	16.190	13.266	14.980	13.461	14.982	15.922

variety	19	20	21	22	23	24	25
	16.696	16.399	14.934	16.444	13.291	15.465	16.306

Standard errors of differences of combined means

Table	variety
rep.	6
s.e.d.	0.6202
effective d.f.	79.99

VSNi

Related algorithms

- Hemmerle (1974) Northogonal analysis of variance using iterative improvement and balanced residuals. *JASA*, 69, 772-778
  - keep sweeping out fixed effects (all at once) with a common (postulated) efficiency factor
- Worthington (1975) General iterative method for analysis of variance when block structure is orthogonal. *Biometrika*, 63, 113-120
  - similar method extended to orthogonal block structures
- Payne (2003) General balance, large data sets and extensions to unbalanced treatment structures. *CSDA*, 44, 297-304.
  - method defined in context of general balance, first-order balance & sweeps
  - implemented in an unpublished Genstat procedure
  - not really a rival to REML
- Brien (2017) Procedures ACANONICAL etc.
  - implementation of full James & Wilkinson dummy analysis
  - rather slow...

..

VSNi

Why still use ANOVA?

- encourages (requires?) you to design your experiment
  – and use of generally balanced designs
    - condition 1: block terms mutually orthogonal
    - key condition to be able to generate an analysis of variance
- example – split-plot for 4 treatment factors
    - 8 cages of poultry
    - diet factors *Thyroxine* and *Yeast* applied to complete cages
    - factors *Hensfood* and *Sex* on individual chickens
    - response variate is gain in weight in second two weeks after hatching
    - John & Quenouille, 2nd Edition (1977) page 95

..

ANOVA Analysis

Analysis of variance

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Cage stratum					
Thyroxine	1	14663.3	14663.3	25.72	0.007
Yeast	1	2592.0	2592.0	4.55	0.100
Thyroxine.Yeast	1	2.5	2.5	0.00	0.950
Residual	4	2280.7	570.2	0.77	
Cage.*Units* stratum					
Sex	1	2329.0	2329.0	3.14	0.079
Hensfood	1	85387.8	85387.8	115.29	<.001
Thyroxine.Sex	1	276.1	276.1	0.37	0.543
Yeast.Sex	1	157.5	157.5	0.21	0.646
Thyroxine.Hensfood	1	19900.1	19900.1	26.87	<.001
Yeast.Hensfood	1	8811.3	8811.3	11.90	<.001
Sex.Hensfood	1	544.5	544.5	0.74	0.393
Thyroxine.Yeast.Sex	1	666.1	666.1	0.90	0.345
Thyroxine.Yeast.Hensfood	1	91.1	91.1	0.12	0.726
Thyroxine.Sex.Hensfood	1	427.8	427.8	0.58	0.449
Yeast.Sex.Hensfood	1	18.0	18.0	0.02	0.876
Thyroxine.Yeast.Sex.Hensfood	1	101.5	101.5	0.14	0.712
Residual	108	79990.6	740.7		
Total	127	218240.0			

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Cage	570.17	4.000	-10.66
Cage.*Units*	740.65	108.000	740.65

VSNi

REML analysis

Estimated variance components

Random term	component	s.e.
Cage	-10.3	26.0

Residual variance model

Term	Model(order)	Parameter	Estimate	s.e.
Residual	Identity	Sigma2	734.8	99.5

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Thyroxine	25.72	1	25.72	4.0	0.007
Yeast	4.55	1	4.55	4.0	0.100
Sex	3.17	1	3.17	109.0	0.078
Hensfood	116.21	1	116.21	109.0	<0.001
Thyroxine.Yeast	0.00	1	0.00	4.0	0.950
Thyroxine.Sex	0.38	1	0.38	109.0	0.541
Yeast.Sex	0.21	1	0.21	109.0	0.644
Thyroxine.Hensfood	27.08	1	27.08	109.0	<0.001
Yeast.Hensfood	11.99	1	11.99	109.0	<0.001
Sex.Hensfood	0.74	1	0.74	109.0	0.391
Thyroxine.Yeast.Sex	0.91	1	0.91	109.0	0.343
Thyroxine.Yeast.Hensfood	0.12	1	0.12	109.0	0.725
Thyroxine.Sex.Hensfood	0.58	1	0.58	109.0	0.447
Yeast.Sex.Hensfood	0.02	1	0.02	109.0	0.876

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Thyroxine.Yeast.Sex	0.91	1	0.91	109.0	0.343
Thyroxine.Yeast.Hensfood	0.12	1	0.12	109.0	0.725
Thyroxine.Sex.Hensfood	0.58	1	0.58	109.0	0.447
Yeast.Sex.Hensfood	0.02	1	0.02	109.0	0.876

General balance - advantages

- condition2: treatment terms mutually orthogonal
  - conclusions do not depend upon order of fitting

Analysis of variance

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Day	1	832.4	832.4	3.14	0.083
A ignoring B	2	2288.8	1144.4	4.31	0.019
A eliminating B	2	1572.7	786.4	2.96	0.062
B ignoring A	1	1900.0	1900.0	7.16	0.010
B eliminating A	1	1183.9	1183.9	4.46	0.067
A.B	2	1688.1	844.0	3.18	0.051
Residual	46	12212.8	265.5		
Total	52	18206.0	350.1		

Analysis of variance

Variate: Y

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Day stratum	1	1441.0	1441.0	5.81	
Day.*Units* stratum					
A	2	1826.4	913.2	3.68	0.034
B	1	1376.0	1376.0	5.55	0.023
A.B	2	3040.0	1520.0	6.13	0.055
Residual	41	10173.4	248.1		
Total	47	17856.8			

VSNi

General balance - advantages

- condition 3: balance – one efficiency factor for each treatment term, in each stratum
  - e.s.e. = √{(stratum variance) / (replication × efficiency factor)}
  - precision depends on replication, not on (random) allocation of treatments to experimental units
  - e.g. give control treatment replication ≈ √{no. test treatments}

..

VSNi

Efficiency of ANOVA

- workspace
  - e.g. balanced-incomplete-block design with *b* blocks, *k* plots per block and *t* treatments
    - general linear-mixed-model algorithm (REML) requires storage of a symmetric matrix with $b+t+1$ rows (for sums of squares and products) and vectors of size *b* and *t* for the block and treatment effects
    - ANOVA requires a working vector of size *bt*, two vectors of size *t* for treatment estimates between and within blocks, and a vector of size *b* if the block residuals are to be stored
  - e.g. lattice design with *r* replicates, *k* blocks per replicate, *k* plots per block and k^2 treatments
    - REML requires storage of a symmetric matrix with $k^2+rk+r+1$ rows and vectors of size k^2, *rk* and *r* for fixed and random effects
    - ANOVA requires storage of vectors of size $r\,k^2$, k^2, k^2, *rk* and *r*
  - conclusion
    - REML workspace depends on the square of number of parameters
    - ANOVA workspace depends linearly on the number of parameters
  - reference
    - Payne & Welham (1990) A Comparison of Algorithms for Combination of Information in Generally Balanced Designs. COMPSTAT 1990, 297-302.

VSNi

Efficiency of ANOVA

- computation speed – can still be an issue e.g. if you have many variates to analyse, or wish to do permutation tests
  - e.g. 10000 y-variates for split-plot for 4 treatment factors

```
spload [print=*] 'JQp95.gsh'
calculate time0 = cputime(0)
blockstructure Cage
treatmentstructure Thyroxine*Yeast*Sex*Hensfood
for [ntimes=10000]
  anova [print=*; factorial=4; fprob=yes; design=savedes] Weight
endfor
calculate anovatime = cputime(0) - time0

calculate time0 = cputime(0)
vcomponents [fixed=Thyroxine*Yeast*Sex*Hensfood] Cage
for [ntimes=10000]
  reml [print=*] Weight
endfor
calculate remltime = cputime(0) - time0

print anovatime,remltime
```

anovatime	remltime
1.172	61.81

Conclusion

- Chris Brien & Roger Payne (2017).
  Graham Neil Wilkinson 1927–2016.
  Journal of the Royal Statistical Society, Series A, 180, 930–931.

  - A key component {of Genstat} was Graham's algorithm for analysis of variance, which exploits the properties of balance in experimental designs to provide a uniquely efficient analysis that allows for different sources of random variation (error strata) and completely general combinations of crossing and nesting between factors.

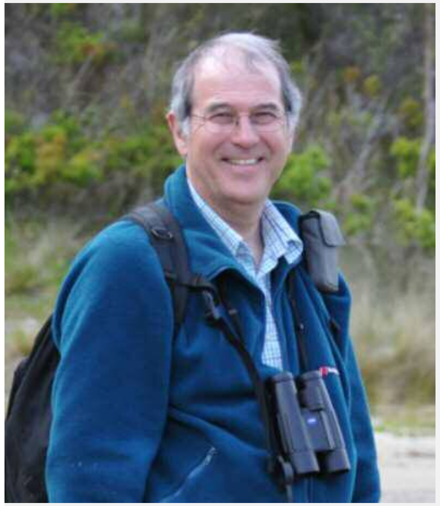
- Heiberger (1981) The Specification of Experimental Designs to ANOVA Programs. American Statistician, 35, 98-104.

  "At this time the GENSTAT ANOVA language provides the most complete capacity for the analysis of generally balanced designed experiments."

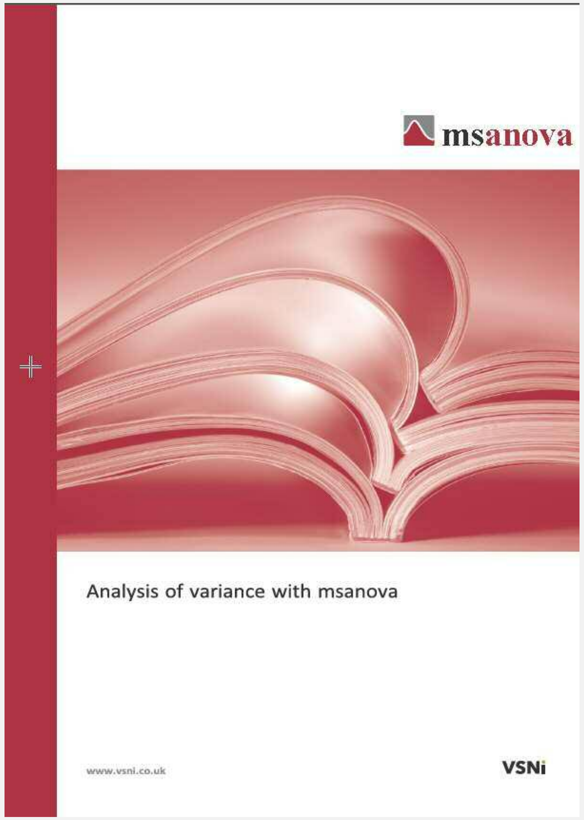
  - and that was even before combination of information, permutation tests, output to speadsheets, automatic reports etc....

VSNi

Conclusion



- so please keep on using ANOVA (and msanova)...!



..

VSNi