The Development of Statistical Design and Analysis Concepts at Rothamsted

a British perspective on the history of statistics in agricultural research

Roger Payne,

VSN International,Hemel Hempstead, Herts, UK

(and Rothamsted Research, Harpenden, Herts UK).

email: Roger.Payne@vsni.co.uk

ROTHAMSTED RESEARCH

ASA – CSSA – SSSA 5th November 2007

100 years ago

- significance test (Arbuthnot 1710)
- Bayes theorem (Bayes 1763)
- least squares (Gauss 1809, Legendre 1805)
- central limit theorem (Laplace 1812)
  - distributions of large-samples tend to become Normal
- "Biometric School" Karl Pearson, University College
  - correlation, chi-square, method of moments
- t-test (Gosset 1908)
- Fisher in *Stats Methods for Research Workers*
  - "..traditional machinery of statistical processes is wholly unsuited to the needs of practical research. Not only does it take a cannon to shoot a sparrow, but it misses the sparrow!"

..

ROTHAMSTED RESEARCH

Rothamsted

- Broadbalk – set up by Sir John Lawes in 1843 to study the effects of inorganic fertilisers on crop yields
- had some traces of factorial structure, but no replication, no randomization and no blocking
- and not analysed statistically until 1919..

ROTHAMSTED RESEARCH

Ronald Aylmer Fisher

- in 1919 Rothamsted Director John Russell was faced with "great files of records"
  - looked for a "young mathematician ... prepared to examine our data and elicit further information"
- obtained a reference from Fisher's tutor at Caius College Cambridge
  - The answer was that he could have been a first class mathematician had he "stuck to the ropes" but he would not.
- conclusion
  - That looked like the type of man we wanted, so I invited him to join us. I had only £200, and suggested he stay as long as he thought that should suffice. ... It took me a very short time to realise that he was more than a man of great ability, he was in fact a genius who must be retained. So I set about obtaining the necessary grant.
- Fisher was at Rothamsted from 1919-1935.

ROTHAMSTED RESEARCH

William Sealy Gosset

Photo from the MacTutor History of Mathematics archive (John J O'Connor and Edmund F Robertson) http://www-groups.dcs.st-and.ac.uk/~history/

- worked at Guinness in Dublin
  - published under the pseudonym "Student"
  - devised the t-test (Biometrika 1908)
- approved of Fisher's appointment at Rothamsted
  - "there should be lots of interesting work to be done there and they might easily have got someone there who would have been worse than useless"
- had a long collaboration and correspondence with Fisher
  - .. but when I come to "Evidently" I know that means at least two hours hard work before I can see why (see J.F. Box, p.115)

..

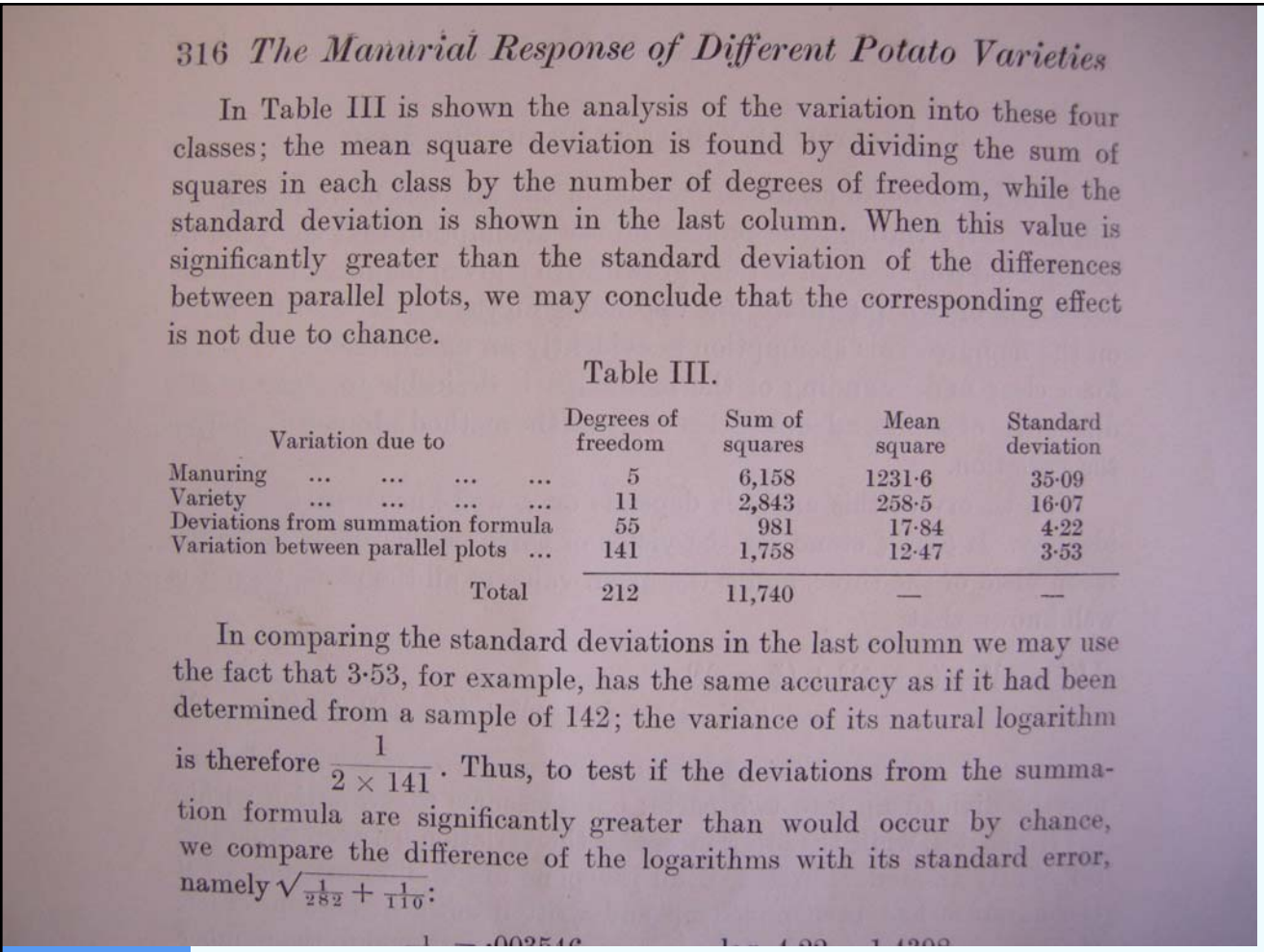
ROTHAMSTED RESEARCH

t-test

- originally derived empirically by Gosset
  - test for small samples that sample mean m = μ
  - if we use sample estimate of s.e. instead of of σ
    cannot assume m – μ will still have a Normal distribution
- caused Fisher to recognise concept of degrees of freedom
  - estimate s.e. by $\sqrt{\{(x - m)^2 / (n - 1)\}}$ with divisor n – 1 not n
- Fisher (1922) Goodness of fit of regression formulae and distribution of regression coefficients. *J. Roy. Statist. Soc.* (1925) Applications of Student's distribution. *Metron* gave full mathematical justification & generalized to
  - differences between means
  - coefficients in regressions and multiple regressions
  - or any statistic of form Normal/√Chisquare
  - and provided an expansion of the cumulative distribution formula of the distribution in inverse powers of n, together with some tables

..

ROTHAMSTED RESEARCH



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 $\sqrt{\frac{1}{282} + \frac{1}{110}}$:

ANOVA

Fisher & MacKenzie (1923) Studies in crop variation II. The manurial response of different potato varieties *J.Agric.Sci.*

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

..

ROTHAMSTED RESEARCH

Analysis of variance

- 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."
- Yates (1937) *Design and Analysis of Factorial Experiments*
  - Yates' algorithm for analysis of variance – orthogonal designs, one error term
- Wilkinson (1970) A general recursive algorithm for analysis of variance. *Biometrika*
  - non-orthogonal (balanced) designs, several error terms
- Payne & Wilkinson (1977). A general algorithm for analysis of variance. *Applied Statistics*
- Payne (1998). Detection of partial aliasing and partial confounding in generally balanced designs. *Comp. Statistics*

..

ROTHAMSTED RESEARCH

Another tea test

Plate 5. *At tea outside the Sample House at Rothamsted.*

- could Dr B. Muriel Bristol tell whether milk had been poured first?
- Chapter II of *The Design of Experiments* (Fisher 1935)
  - used to illustrate concepts of randomization, significance, exact tests

..

ROTHAMSTED RESEARCH

Significance tests

- Arbuthnot (1710) "Divine Providence had intervened in favour of the male sex"
  - probability $1/2^{17}$ of more male than female births in London in 17 consecutive years
- Fisher (1935) *The Design of Experiments* Ch. II
  - Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis
- Yates "those damned stars"

..

ROTHAMSTED RESEARCH

Rothamsted Annual Report 1933

Mean responses and interactions

* 5 per cent Significance. ** 1 per cent. Significance.

Centre.	*Mean Yield.*	Mean response to			Standard Error.	Interactions			Standard Error.
		N	P	K		N × P	N × K	P × K	
Roots (washed) : tons per acre									
2. Ipswich ..	*5.17*	+0.02	−0.17	−0.15	±0.17	+0.44	−0.12	+0.06	±0.35
3. Colwick ..	*7.17*	*+1.53***	+0.36	+0.21	±0.46	+0.28	−0.92	−0.30	±0.93
4. Newark ..	*8.29*	+0.10	+0.39	+0.20	±0.36	−1.28	−0.26	+0.24	±0.72
5. Felstead ..	*9.09*	+0.17	+0.26	+0.56	±0.29	−0.44	−0.16	+0.18	±0.58
6. Brigg	*10.74*	*+1.76***	−0.74	+0.10	±0.38	−1.65	−0.06	+0.19	±0.76
7. Poppleton ..	*11.71*	+0.02	−0.02	+0.04	±0.38	−0.04	−0.27	+1.66*	±0.75
8. Bardney ..	*12.32*	+1.26	+0.01	+0.30	±0.71	−0.14	+1.10	−1.48	±1.42
9. Allscott ..	*12.38*	+0.38	−0.16	*+1.16***	±0.28	+0.84	*−2.08***	0.00	±0.55
10. Wissington ..	*13.82*	+0.20	+0.08	−0.39	±0.28	+0.82	+0.09	+0.61	±0.56
11. Peterborough ..	*14.06*	−0.29	−0.03	−0.03	±0.26	−1.26*	−1.44*	−0.10	±0.52
12. King's Lynn ..	*14.36*	−0.39	−0.33	−0.65	±0.33	+0.06	−0.38	+0.16	±0.66
13. Ely	*14.74*	+0.24	+0.63	+0.12	±0.36	−0.21	−1.00	−0.53	±0.73
14. Cantley ..	*16.36*	−0.83	+0.62	+0.36	±0.69	+1.70	−0.08	−0.26	±1.38
Mean	*11.53*	*+0.32*	*+0.07*	*+0.14*		*−0.07*	*−0.43*	*+0.03*	
Sugar Percentage									
2. Ipswich ..	*15.92*	−0.84*	−0.38	+0.34	±0.320	−0.20	+0.10	+0.14	±0.640
3. Colwick ..	*15.10*	−0.13	−0.12	+0.22	±0.322	+0.10	−0.30	−0.26	±0.643
4. Newark ..	*16.23*	*−0.38***	−0.22	+0.12	±0.119	−0.32	−0.15	+0.14	±0.238
5. Felstead ..	*16.72*	*−0.66***	+0.22	+0.14	±0.129	−0.35	+0.02	−0.01	±0.258
6. Brigg	*17.74*	*−0.36***	−0.14	+0.09	±0.096	+0.18	−0.01	−0.17	±0.191
7. Poppleton ..	*17.94*	−0.15	−0.05	+0.25	±0.263	−0.24	+0.24	−0.30	±0.526
8. Bardney ..	*16.15*	−0.62*	+0.01	+0.21	±0.286	+0.32	+0.18	−0.68	±0.573
9. Allscott ..	*15.93*	−0.19	+0.02	*+0.39***	±0.113	−0.08	−0.08	−0.28	±0.226
10. Wissington ..	*16.40*	*−0.62***	+0.12	+0.24	±0.166	−0.23	+0.22	−0.33	±0.331
11. Peterborough ..	*14.84*	−0.15	+0.22	+0.20	±0.222	−0.53	−0.16	+0.63	±0.444
12. King's Lynn	*16.69*	+0.14	+0.63*	−0.24	±0.294	−0.48	+0.32	−0.32	±0.587
13. Ely	*15.89*	−0.07	−0.04	−0.13	±0.189	−0.14	+0.14	−0.26	±0.378
14. Cantley ..	*15.70*	−0.33	+0.07	+0.14	±0.212	−0.06	+0.68	+0.14	±0.424
Mean ..	*16.25*	*−0.34*	*+0.03*	*+0.15*		*−0.16*	*+0.09*	*−0.12*	

ROTHAMSTED RESEARCH

- the first "damned stars"..?

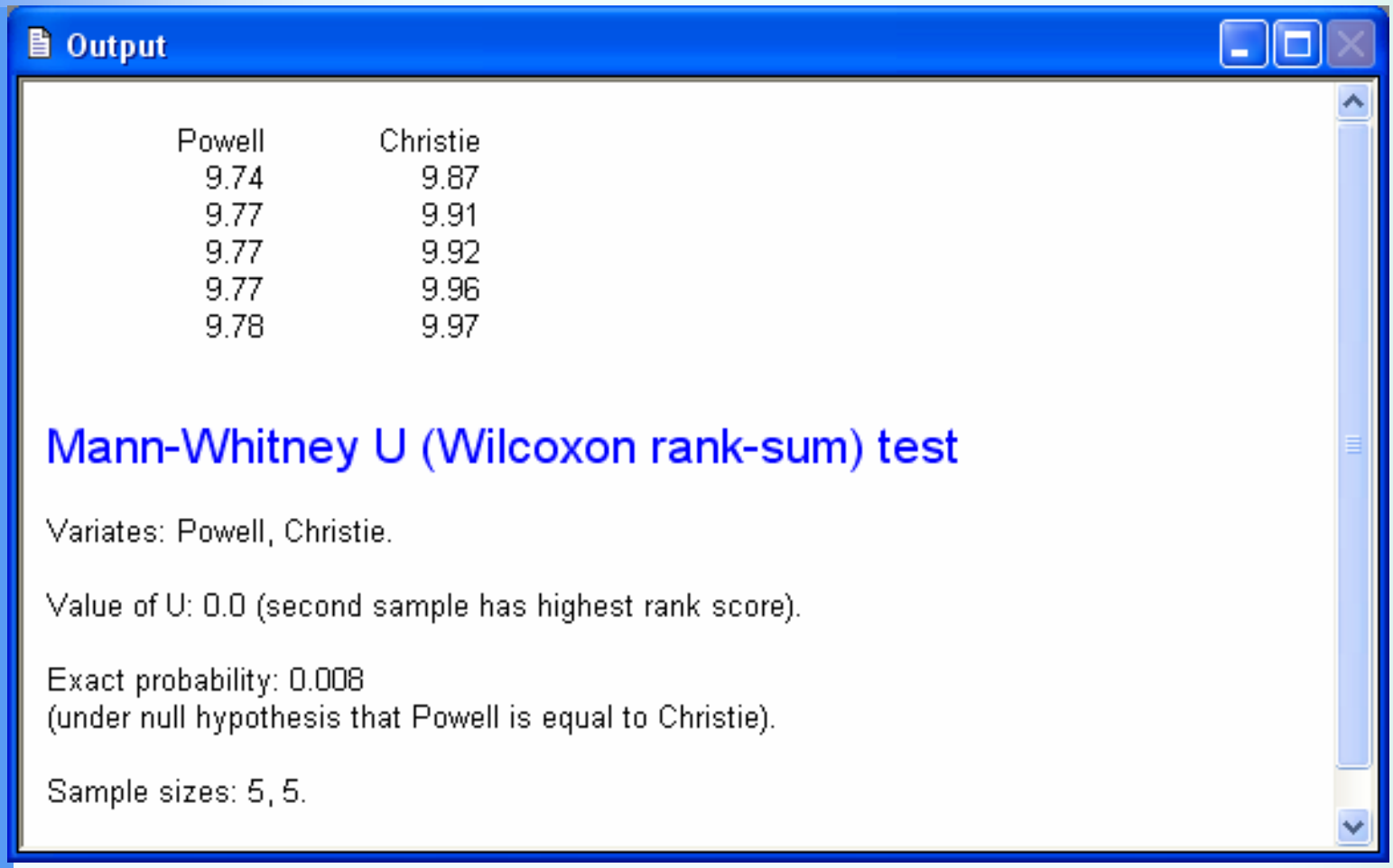
Significance tests

- Arbuthnot (1710) "Divine Providence had intervened in favour of the male sex"
  - probability $1/2^{17}$ of more male than female births in London in 17 consecutive years
- Fisher (1935) *The Design of Experiments* Ch. II
  - Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis
- Yates "those damned stars"
- Payne quoted in Perry (1986) Multiple-comparison procedures a dissenting view. *J. Econ Entomol*
  - .. likened a significance test to the safety net of a tightrope walker: it helps give confidence (that unjustified conclusions are not being drawn from random errors in the data) but should not be part of the act

..

ROTHAMSTED RESEARCH

Significance?



- 5 best 100m. times for Asafa Powell and Linford Christie
  - Powell is significantly faster
  - but would it matter if they wanted to run to catch a bus?
  - must consider *biological* as well as *statistical* significance!

ROTHAMSTED RESEARCH

Randomization

- Fisher (1926) The arrangement of field experiments. *J. Min. Ag. G. Br.*
  - 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.
  - i.e. randomization distribution should approximate Normal distribution
  - demonstrated by Eden & Yates (1933, *J.Agric.Sci*) with uniformity data
- very controversial at the time
  - systematic designs preferred e.g. Russell (1926) *J. Min. Ag. G. Br.* and Student (1936) *J. Roy. Statist. Soc. Suppl.*
- Bailey, Praeger, Rowley & Speed (1981) Generalized wreath products of permutation groups. *Proc. London Math. Soc.*
- Bailey (1983) Restricted randomization. *Biometrika*
  - how to rule out "unfortunate" randomizations

..

ROTHAMSTED RESEARCH

Design of experiments

- Boston Globe (1936)
  - "Fisher taught experimenters how to experiment."
- Fisher & Rothamsted taught Americans applied statistics
  - notice W.G. Cochran (1934-9) & O. Kempthorne (1941-9)
  - also many others on informal research visits

..

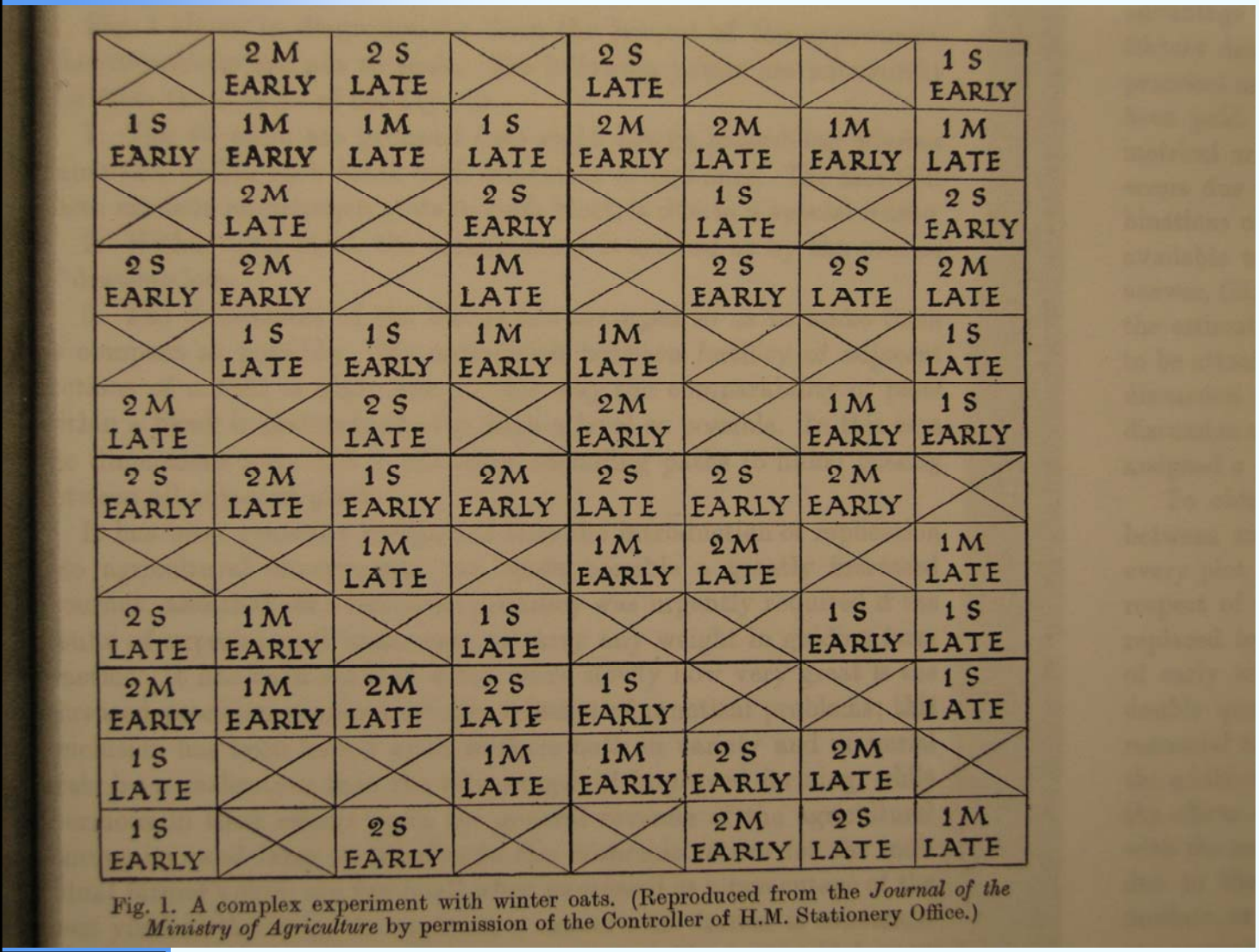
Design of experiments

- Boston Globe (1936)
  - "Fisher taught experimenters how to experiment"
- Fisher (1926) introduced factorial design
  - "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."
  - if there is no interaction, we have combined several one-at-a-time expts
  - while, if there is an interaction, we are much wiser!
- Eden & Fisher (1927) Studies in crop variation IV The experimental determination of the value of top dressings with cereals *J. Agric. Sci.*
  - first of Fisher's design to be used in practice

..

ROTHAMSTED RESEARCH

Eden & Fisher (1927)



	2 M EARLY	2 S LATE		2 S LATE			1 S EARLY
1 S EARLY	1M EARLY	1M LATE	1 S LATE	2M EARLY	2M LATE	1M EARLY	1M LATE
	2M LATE		2 S EARLY		1 S LATE		2 S EARLY
2 S EARLY	2M EARLY		1M LATE		2 S EARLY	2 S LATE	2M LATE
	1 S LATE	1 S EARLY	1M EARLY	1M LATE			1 S LATE
2M LATE		2 S LATE		2M EARLY		1M EARLY	1 S EARLY
2 S EARLY	2M LATE	1 S EARLY	2M EARLY	2 S LATE	2 S EARLY	2 M EARLY	
		1M LATE		1M EARLY	2M LATE		1M LATE
2 S LATE	1M EARLY		1 S LATE			1 S EARLY	1 S LATE
2M EARLY	1M EARLY	2M LATE	2 S LATE	1 S EARLY			1 S LATE
1 S LATE			1M LATE	1M EARLY	2 S EARLY	2M LATE	
1 S EARLY		2 S EARLY			2M EARLY	2 S LATE	1M 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.)

ROTHAMSTED RESEARCH

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

..

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")

..

ROTHAMSTED RESEARCH

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

..

ROTHAMSTED RESEARCH

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

..

ROTHAMSTED RESEARCH

Design of expts – some later work

- Yates (1933) The principles of orthogonality and confounding in designed experiments. *J. Agric. Sci*
  Yates (1935) Complex expts. *J. Roy. Statist. Soc., Suppl.*
  - split plots, strip plots, balanced confounding
- Yates (1936) Incomplete randomized blocks. *Ann. Eugenics*
  - balanced incomplete blocks, efficiency factor
- Yates (1936) A new method of arranging variety trials involving a large number of varieties. *J. Agric. Sci*
  - lattice designs, pseudo-factors, efficiency factors
- Paterson (1952) Expts involving sequences of treats *Biometrika*
  & Williams (1976) A new class of resolvable designs *Biometrika*
- Preece (1966) Some bib's for 2 sets of treatments *Biometrika*
- Nelder (1965) The analysis of randomized experiments with orthogonal block structure. *Proc. Roy. Soc. A*
- Payne & Tobias (1992) General balance, combination of information and the analysis of covariance. *Scand. J. Stats.*
- Coombes, Payne, & Lisboa (2002) Comparison of nested simulated annealing and reactive tabu search for efficient experimental designs with correlated data. *COMPSTAT 2002*

..

ROTHAMSTED RESEARCH

Non-Normal data

- Fisher (1924) Case of zero survivors in probit assay. *Ann.Appl.Biol.*
  - maximum likelihood analysis of probit data
- Finney *Probit Analysis* (1947, 1964, 1971)
- Nelder & Wedderburn (1972) Generalized linear models. *J. Roy. Statist. Soc. A*
- Wedderburn (1974) Quasi-likelihood functions, generalized linear models and the Gauss-Newton method. *Biometrika*
- Nelder & Pregibon (1987) An extended quasi-likelihood function. *Biometrika*
- Lee & Nelder (1996) Hierarchical generalised linear models. *J. Roy. Statist. Soc. A*
  - generalized linear models with several error terms, h-likelihood
- Lee & Nelder (2006) Double hierarchical generalized linear models. *Appl. Statist.*
- Payne, Lee, Nelder & Noh (2007) Procedures HGANALYSIS etc
  - most recent GenStat implementation

..

ROTHAMSTED RESEARCH

Likelihood

- Fisher (1912) On an absolute criterion for fitting frequency curves. *Messeng. Math.*
  - used likelihood for estimation
- Fisher (1922) On the mathematical foundations of mathematical statistics. *Phil. Trans. A*
  - maximum likelihood, consistency, sufficiency..
- Fisher (1930) Inverse Probability. *Proc. Camb. Phil. Soc.*
  - fiducial inference
- Wilkinson (1977) On Resolving the Controversy in Statistical Inference. *J. Roy. Statist. Soc. B*
- Ross (1987) *Maximum Likelihood Program*
- Ross (1990) *Nonlinear Estimation*

..

ROTHAMSTED RESEARCH

Variance components

- Fisher (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinb.*
  - introduced concept of analysis of variance components
- Patterson & Thompson (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika*
  - REML (residual/restricted maximum likelihood) for analysis of unbalanced designs and variance-component estimation
- Wilkinson, Eckert, Hancock & Mayo (1983) Nearest Neighbour (NN) Analysis of Field Experiments. *J. Roy. Statist. Soc. B*
- Gilmour, Thompson & Cullis (1995). Average Information REML, an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics*
- Gilmour, Cullis & Verbyla (1997). Accounting for natural extraneous variation in analysis of field experiments. *JABES*
- Verbyla, Cullis, Kenward & Welham (1997). The analysis of designed experiments and longitudinal data using smoothing splines. *Applied Statistics*

..

ROTHAMSTED RESEARCH

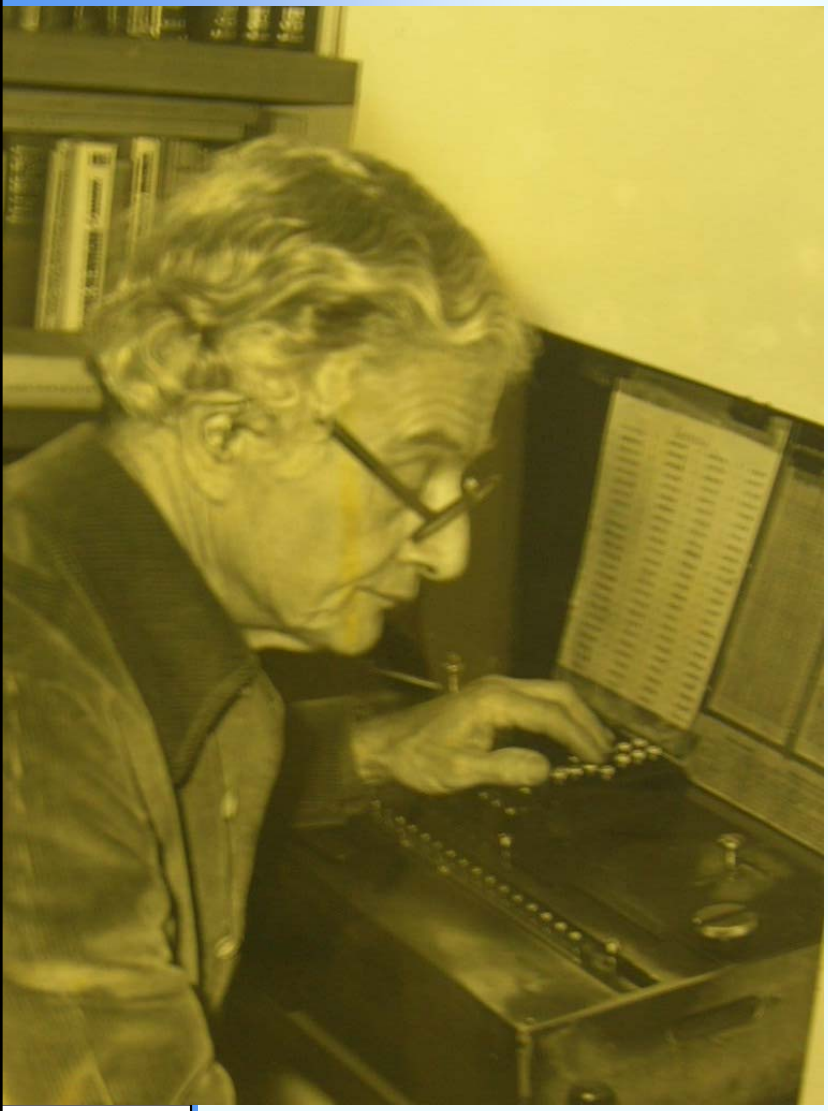
REML – subnote

- also used in plant science
  - e.g. Baird, Johnstone, & Wilson (2004) Normalization of microarray data using a spatial mixed model analysis which includes splines. *BioInformatics*
- providing a link to Fisher's interests in genetics (not discussed here)
  - R. Dawkins (1986) *The Blind Watchmaker* "the greatest of his [Darwin's] successors"
  - L.J. Savage (1976) *Annals of Statistics* "I occasionally meet geneticists who ask me whether it is true that the great geneticist R.A. Fisher was also an important statistician"

..

ROTHAMSTED RESEARCH

Statistical computing



- Millionaire calculating machine
  - bought by Fisher at a cost >£200
  - used here by his successor Frank Yates (Head of Department 1933-1968)
  - Fisher: "most of my statistics has been learned on the machine"

..

ROTHAMSTED RESEARCH

Statistical computing – later work

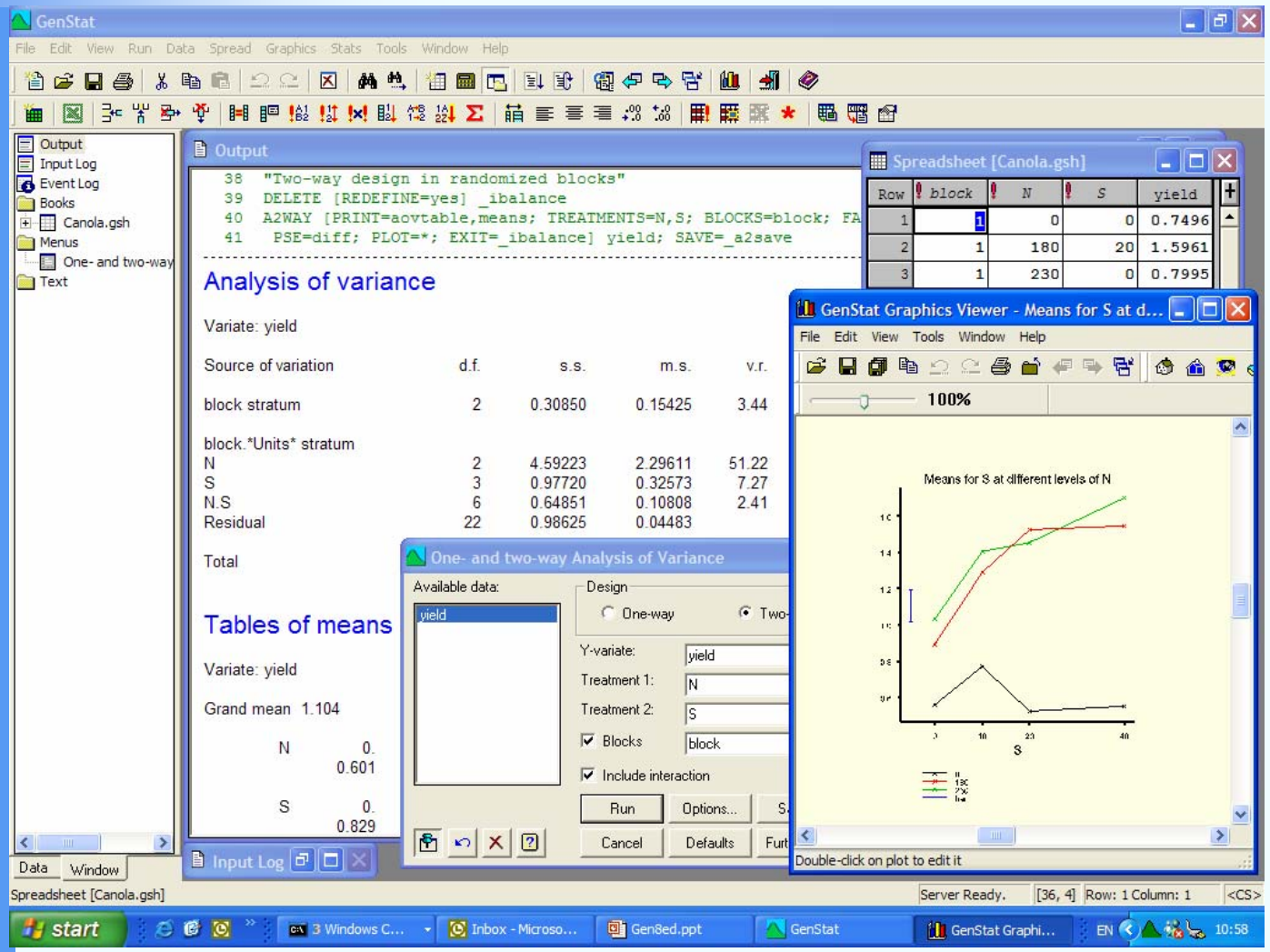
Elliott-NRDC 401 electronic computer

- Elliot 401 computer installed at Rothamsted in 1954
  - Gower (1985) "first computer to be associated primarily with agricultural research and with statistics"
  - Yates (1955) "Having an electronic machine on the spot has made all the difference to developing its applications to research statistical problems"

..

ROTHAMSTED RESEARCH

Statistical computing – later work



- GenStat *for Windows*
  - written by Rothamsted since 1968 (and now VSN) to develop and distribute statistical methodology and research
  - see www.vsni.co.uk (or the Agronomix Stand)..

ROTHAMSTED RESEARCH

Symbosis - stats and agriculture

- the problems posed by the agricultural and biological research at Rothamsted have spurred many advances in statistics
  - tradition started by Fisher, and continued ever since
  - we have helped you, but you have helped us!
- Fisher recommended full statistical involvement
  - "To consult the statistician after an experiment is finished is often merely to ask him to conduct a *post mortem* examination. He can perhaps say what the experiment died of." Presidential Address to the First Indian Statistical Congress, 1938.
- see e.g. Bailey & Payne: Experimental design: statistical research and its application. *Institute of Arable Crops Research Report for 1989*

..

ROTHAMSTED RESEARCH

R.A. Fisher

- "*The Statistics Department had its origin in the appointment of R.A. Fisher in 1919 to study the accumulated results of the classical field experiments. Fisher soon realised the need for improved statistical techniques over the whole range of agricultural and biological research, and the groundwork for modern statistics was laid by him during the 1920s and 1930s.*"
  - Anon (1971) *Rothamsted Experimental Station, Field Experiments and the Work of the Departments*.
- for more information, see Joan Fisher Box (1978) *R.A. Fisher The Life of a Scientist*
  - if you can find a copy!

ROTHAMSTED RESEARCH