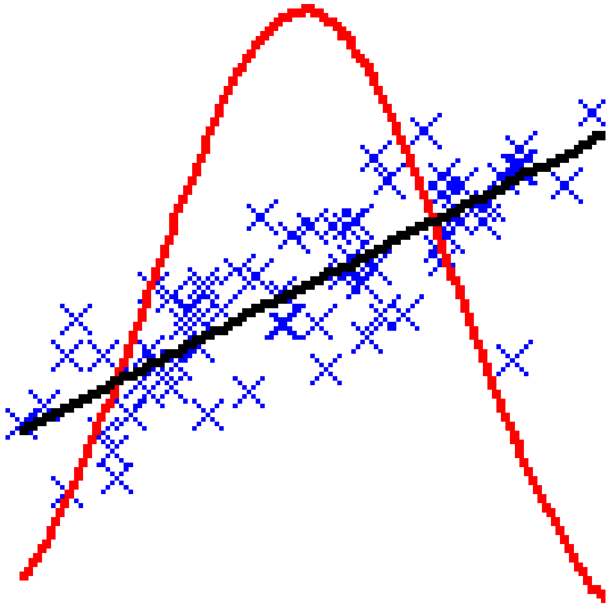
Déjà eu. How we keep on losing and re-inventing statistical theory ’.



Stephen Senn

@stephensenn stephen@senns.uk

Outline of this talk

"No scientific discovery is named after its original discoverer." Stigler's Law of Eponymy[1]

- The roots of the combination of observations in astronomy and other physical sciences.
- Two books on combination of observations.
  - Airy (astronomy).
  - Brunt (astronomy and meteorology).
- What happened in agriculture.
  - Student and his t-test.
  - Fisher.
  - Yates and Cochran.
- What has happened and is happening in medicine.
- Conclusions.

The roots of the combination of observations in the late 18th and early 19th century in the science of astronomy

Stars for significance?

A timeline of Egon Pearson's (1990)

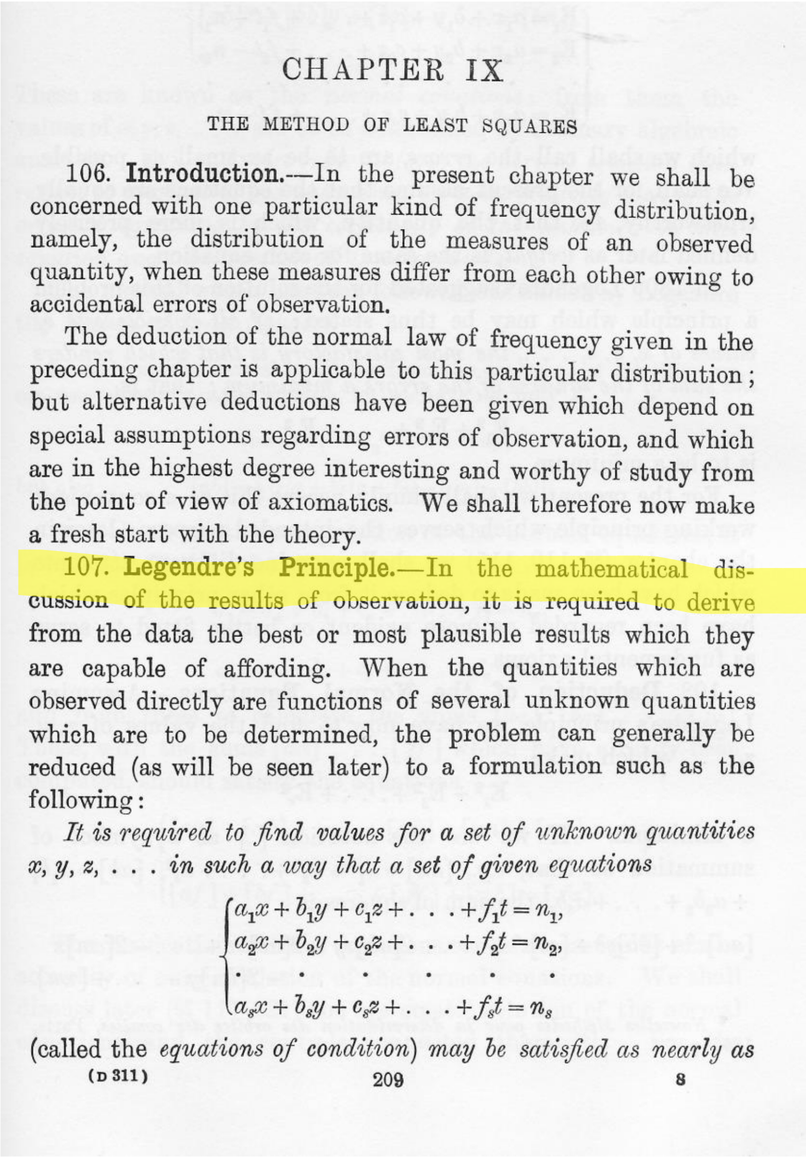
When	Who	What	Where
1764	Thomas Bayes	Inverse probability	*Philosophical Transactions*, Royal Society of London
1774	Pierre Simon Laplace	Inverse probability	*Memoires de l'Academie des Sciences de Paris*
1809	Carl Friedrich Gauss	Least squares	*Theoria motus corpus coelistium*, Hamburg
1810	Laplace	Alternative central limit theorem	*Memoires de l'Academie des Sciences de Paris*
1830	Lubbock & Drinkwater (published anonymously)	*On Probability*	Society for the Diffusion of Useful Knowledge
1837	De Morgan	Theory of probabilities	*Ladner's Cabinet Cyclopedia*
1850	Herschel	Alternative derivation of the law of errors	*Edinburgh Review*
1854	Boole	Challenges inverse probability	*Laws of Thought*, Walton & Maberley, London

Missing but major

Whittaker and Robinson
The Calculus of Observations
1924, 1926, 1940, 1944

- Adrien Legendre (1752-1833)
- Identified by Stephen Stigler as being a key figures in the development of least squares.
- Often overlooked but there are exceptions:

"His method involved three observations taken at equal intervals and he assumed that the comet followed a parabolic path so that he ended up with more equations than there were unknowns. He applied his methods to the data known for two comets. In an appendix Legendre gave the least squares method of fitting a curve to the data available.
However, Gauss published his version of the least squares method in 1809 and, while acknowledging that it appeared in Legendre's book, Gauss still claimed priority for himself. This greatly hurt Legendre who fought for many years to have his priority recognised." MacTutor



CHAPTER IX

THE METHOD OF LEAST SQUARES

106. **Introduction.**—In the present chapter we shall be concerned with one particular kind of frequency distribution, namely, the distribution of the measures of an observed quantity, when these measures differ from each other owing to accidental errors of observation.

The deduction of the normal law of frequency given in the preceding chapter is applicable to this particular distribution; but alternative deductions have been given which depend on special assumptions regarding errors of observation, and which are in the highest degree interesting and worthy of study from the point of view of axiomatics. We shall therefore now make a fresh start with the theory.

107. **Legendre's Principle.**—In the mathematical discussion of the results of observation, it is required to derive from the data the best or most plausible results which they are capable of affording. When the quantities which are observed directly are functions of several unknown quantities which are to be determined, the problem can generally be reduced (as will be seen later) to a formulation such as the following:

It is required to find values for a set of unknown quantities $x, y, z, \ldots$ *in such a way that a set of given equations*

$$\begin{cases} a_1x + b_1y + c_1z + \ldots + f_1t = n_1, \\ a_2x + b_2y + c_2z + \ldots + f_2t = n_2, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_sx + b_sy + c_sz + \ldots + f_st = n_s \end{cases}$$

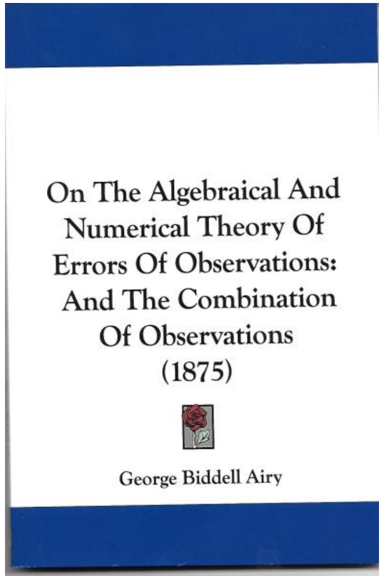
(called the *equations of condition*) *may be satisfied as nearly as*

(D 311) 209 8

Two books on the combination of observations

Light on Airy with the brunt on Brunt

George Biddell Airy



1801-1892

Astronomer Royal from 1835

First edition 1866

This was a book that Student mentioned in correspondence with Fisher

First. The combination-weight for each measure ought to be proportional to its theoretical weight.

Second. When the combination-weight for each measure is proportional to its theoretical weight, the theoretical weight of the final result is equal to the sum of the theoretical weights of the several collateral measures. (pp55-56).

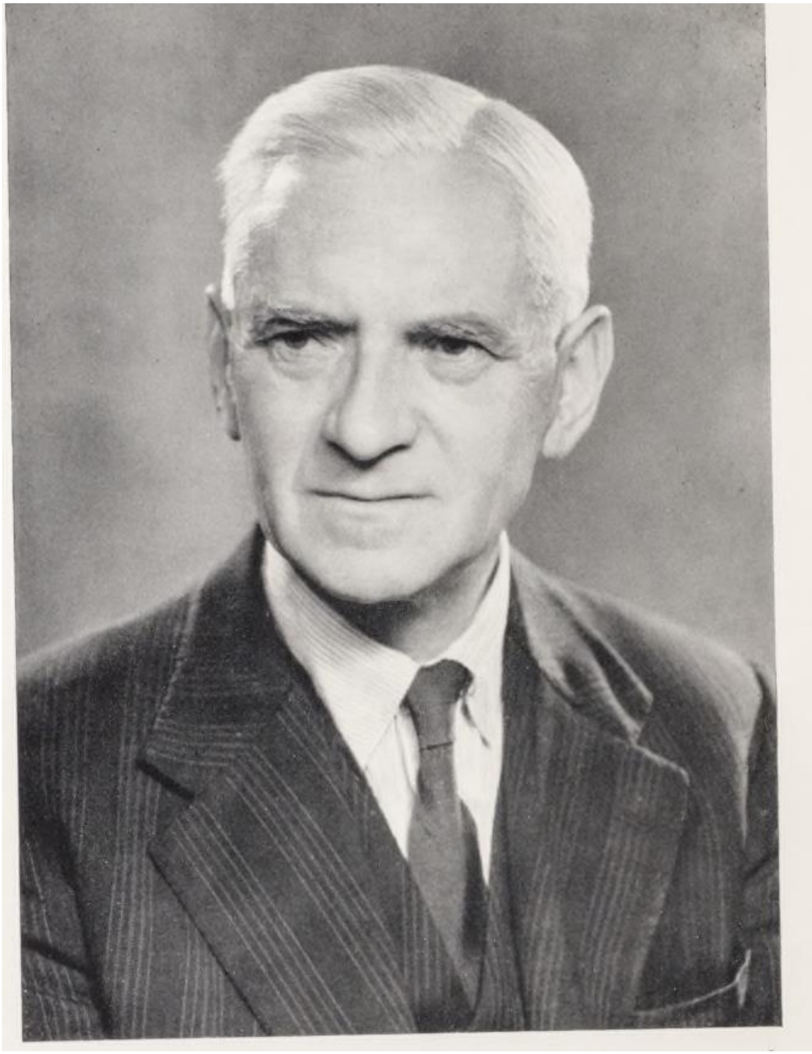
Modern interpretation of Airy's formulae

- They are equivalent to the formula we use for fixed effects meta-analysis.
- They are analogous *mutatis, mutandis* to the formula for Bayesian updating of the prior distribution to the posterior using the likelihood.

$$\text{Estimate}_{posterior} = \frac{\text{Precision}_{prior} \times \text{Estimate}_{prior} + \text{Precision}_{data} \times \text{Estimate}_{data}}{\text{Precision}_{prior} + \text{Precision}_{data}}$$

and

$$\text{Precision}_{posterior} = \text{Precision}_{prior} + \text{Precision}_{data}$$



See: Sutton, O.G., *David Brunt, 1886-1965*. 1965, The Royal Society London.

David Brunt (1886-1965)

- Born into a Welsh speaking family in Montgomeryshire.
- Did not learn English until the age of 10.
- Studied maths in Cambridge at Trinity College.
- Interested in astronomy.
- Became involved in meteorology during WWI.
- Worked in the meteorological office.
- Was later a professor of meteorology at Imperial College London.
- His book, *The Combination of Observations* is a thorough treatment of the theory of least squares as it was in 1917.
- Second edition 1931 was already outdated.

The Combination of Observations 1917 & 1931

I. Errors of observation
II. The law of error
III. The case of one unknown
IV. Observations of different weight
V. The general problem of the adjustment of indirect observations involving several unknown quantities
VI. Evaluation of the most probable values of the unknowns, their weights and probable errors
VII. The adjustment of conditioned observations
VIII. The rejection of observations
IX. Alternatives to the Normal law of error
X. Correlation
XI. Harmonic analysis from the standpoint of least squares
XII. The periodogram

The first edition is 219 pages long with 90 references (some repeated).
The second edition is 239 pages long with 112 references (some repeated).

The Combination of Observations

use equation (3). In all other cases it is safer to use equation (2); i.e. to calculate the P.E. from the residuals.

EXAMPLES.

1. Given the following six determinations of the parallax of the star Lalande 21185, find the weighted mean and its P.E.

Parallax	Weight (p)	x	px	v	v^2	pv^2
0″·507	8	107	856	104	10816	86528
·438	5	38	190	35	1225	6125
·381	2	−19	−38	−22	484	968
·371	8	−29	−232	−32	1024	8192
·350	13	−50	−650	−53	2809	36517
·402	20	2	40	−1	1	20
	56		166			138350

Let the parallax of the star be $0''\cdot 4 + x'' \times 10^{-3}$ Then the six values of x are given in the third column. Each value of x is multiplied by the corresponding weight, and the result written in the fourth column. The weighted mean gives for x,

$$x = \tfrac{166}{56} = 3.$$

The adopted parallax is therefore 0″·403.

The residual v obtained by subtracting the weighted mean from x is written in the fifth column, v^2 in the sixth column, and pv^2 in the seventh column. The sum of the last column yields $[pvv]$.

The P.E. of the weighted mean is

$$0\cdot6745 \sqrt{\frac{[pvv]}{[p](n-1)}} = 0\cdot6745 \sqrt{\frac{138350}{56 \times 5}} = 15\cdot0$$

in units of the third decimal place.

The final result may be written

$$0''\cdot403 \pm \cdot015.$$

EXAMPLES.

1. Given the following six determinations of the parallax of the star Lalande 21185, find the weighted mean and its P.E.

Parallax	Weight (p)	x	px	v	v^2	pv^2
0″·507	8	107	856	104	10816	86528
·438	5	38	190	35	1225	6125
·381	2	−19	−38	−22	484	968
·371	8	−29	−232	−32	1024	8192
·350	13	−50	−650	−53	2809	36517
·402	20	2	40	−1	1	20
	56		166			138350

Brunt in summary

Approach

- Weights the same way that Airy does.
    - As in a fixed effects estimator.
- Calculates probable error between studies.
    - As in a random effects estimator.
- Assumes Normality
    - No concept of degrees of freedom.

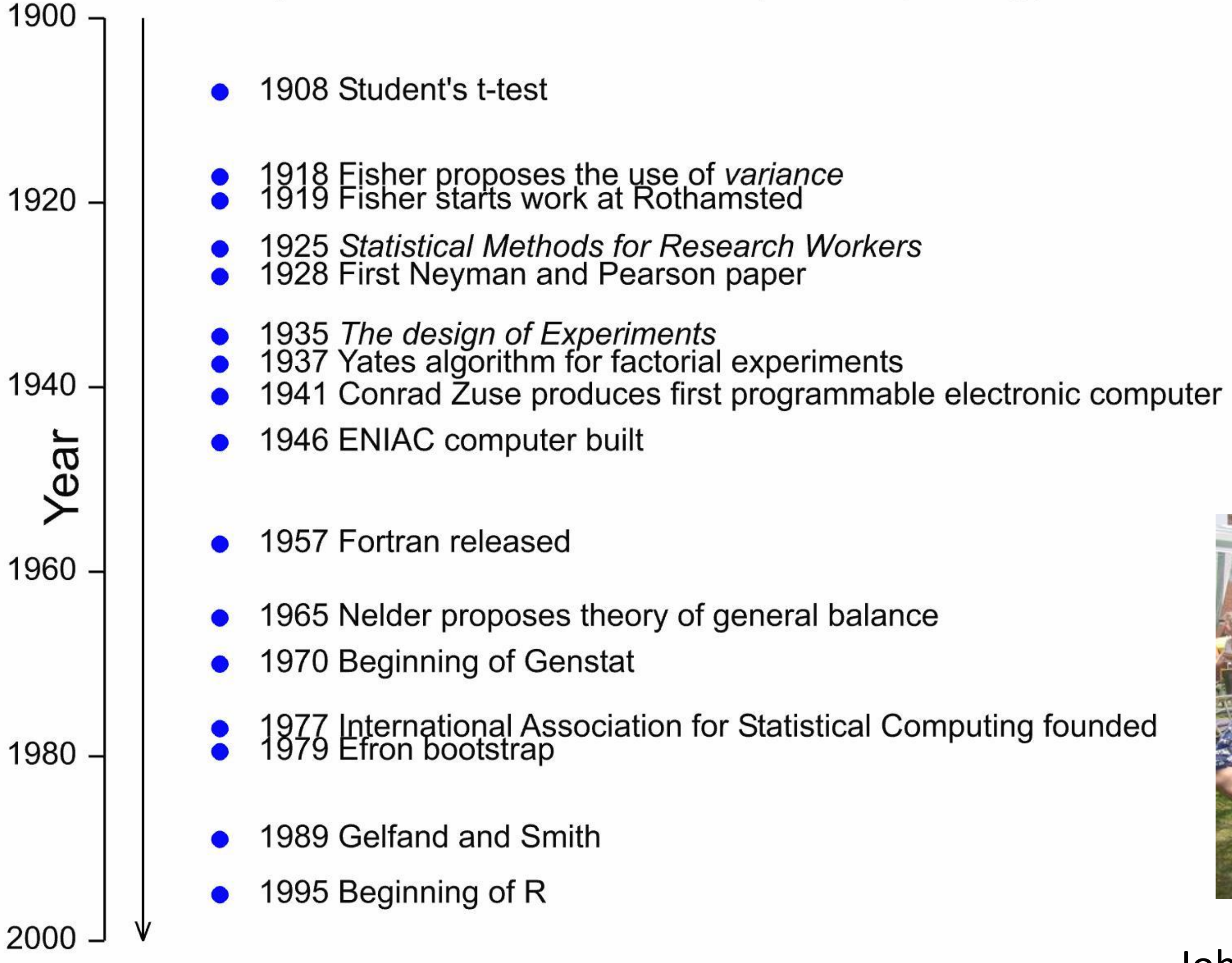
Citations

- Well aware of the long tradition in astronomy stretching back to Gauss and Laplace.
    - But, curiously, does not cite Airy's book despite the phrase *The Combination of Observations* appearing in both.
- Unaware of what Student and then Fisher did for small samples.
    - Forgivable in 1917.
    - Outdated by 1931.

What Happened in Agriculture

Research on fields and a field of research

A 20th century time line of statistics (& computing)



John Nelder & Michael Healy

William Sealy Gosset 1876-1937

- Born Canterbury 1876
- Educated Winchester and Oxford
- First in mathematical moderations 1897 and first in degree in Chemistry 1899
- Starts with Guinness in 1899 in Dublin
- Autumn 1906-spring 1907 with Karl Pearson at UCL
- 1908 publishes ‘The probable error of a mean’
- First method available to judge ‘significance’ in small samples

"Student" (1867-1937)

- We now know that Gosset was anticipated by Jakob Lüroth# in 1876.
- However, he developed the t-test independently and used it regularly in his work, so deserves the credit he is given.
- Was put in touch with Fisher by Stratton, a tutor at Caius where, Fisher was when he was at Cambridge.
- Some years later a letter from Isidor Greenwald informed Fisher that the data he had copied from Student and which were used in *Statistical Methods for Research Workers* had been mislabelled.

THE INTERPRETATION OF EXPERIMENTAL RESULTS.

BY T. B. WOOD, *School of Agriculture,*

AND

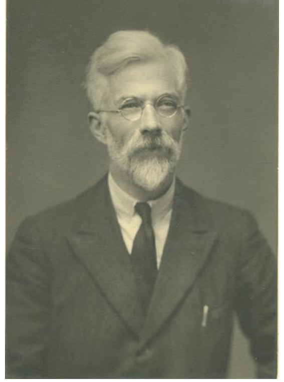
F. J. M. STRATTON, *Caius College, Cambridge.*

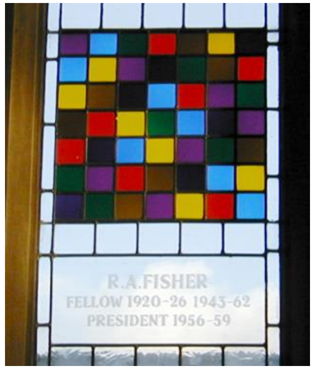
Probable Error.

It might seem at first that no two branches of study could be more widely separated than Agriculture and Astronomy. A moment's consideration, however, will show that they have one point in common: both are at the mercy of the weather. The astronomer's measurements come short of absolute accuracy because of a great number of varying atmospheric conditions, each of which is equally likely to make any one result high or low. He has to obviate this unavoidable lack of accuracy by making many independent observations, and taking their average. This is, or should be, the method followed by the agriculturist.

Wood and Stratton 1910

RA Fisher (1890-1962)



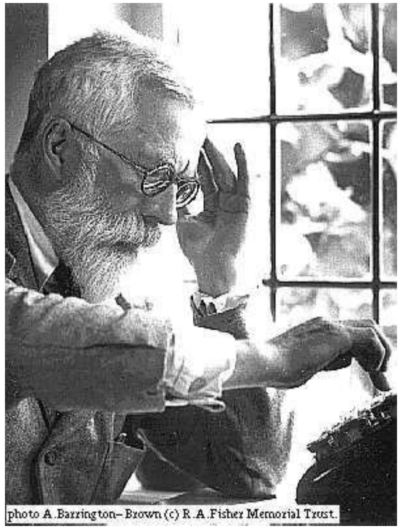


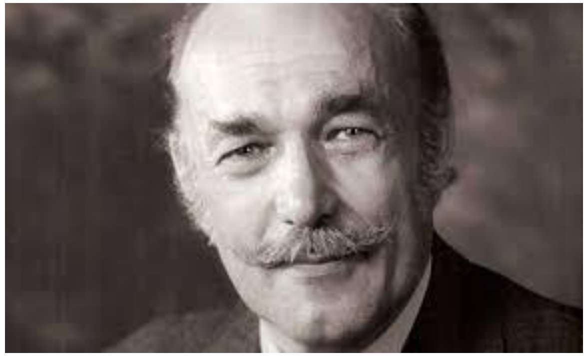
- The single most important figure in modern statistics.
- Studied at Cambridge in Caius, where Stratton was a tutor.
- Worked for many years at Rothamsted.
- Famed as geneticist* as well as statistician.
- Was professor of genetics at University College London and Cambridge but never held a chair in statistics.

* Referred to by Richard Dawkins as the greatest of Darwin's successors. (*The Blind Watchmaker*, 1986, 1988, p 199.)

- Developed Student's small t test to apply to:
  - Two samples (not just one)
  - Regression problems.
- Generalised the approach to deal with:
  - More than two treatments
  - More than one level of variation.
- Linked the parametric approaches based on the Normal distribution to symmetries induced by (or calling for) randomisation.
- Stressed not just good point estimation but valid estimates of error.

The Rothamsted School





RA Fisher
1890-1962
Variance, ANOVA
Randomisation, design,
significance tests

Frank Yates
1902-1994
Factorials, recovering
Inter-block information

John Nelder
1924-2010
General balance, computing
Genstat®

and Frank Anscombe, David Finney, Rosemary Bailey, Roger Payne etc

Frank Yates (1902-1994)

- Fisher's deputy at Rothamsted who succeeded him when Fisher left.
- Made many important contributions to the design and analysis of experiments. In particular:
    - Factorial experiments
    - Incomplete block designs
    - Recovering inter-block information

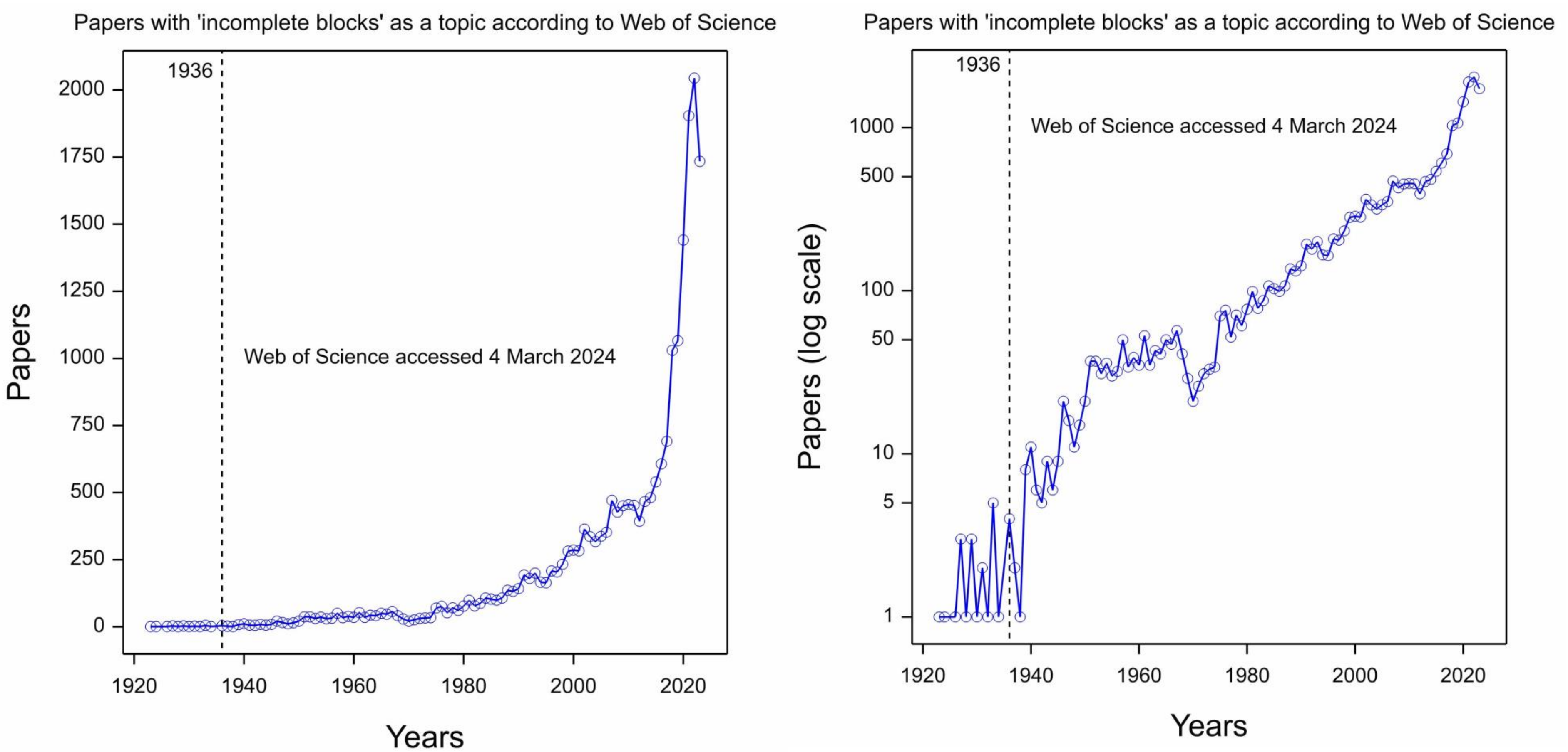
◆ WARTIME WORK ◆

In 1939, war brought grave anxieties about British food supplies. Rothamsted had been to the fore in crop research for a century, but Frank realised how little had been done to collate all evidence on the scope for increasing food production in a country depending so largely on imports. The submarine menace raised questions about the balance between importing food and importing fertilisers. Frank initiated one of the earliest exercises in what today is fashionable as 'metanalysis'. Using the resources of Rothamsted's fine library, and with many voluntary helpers, he abstracted information on yield increases associated with each of the main fertiliser elements, N, P, and K, as evidenced by experiments on all major crops.

A final report (Crowther and Yates, 1941) rested on data from perhaps 5000 experiments. The findings were

David Finney, 1995.

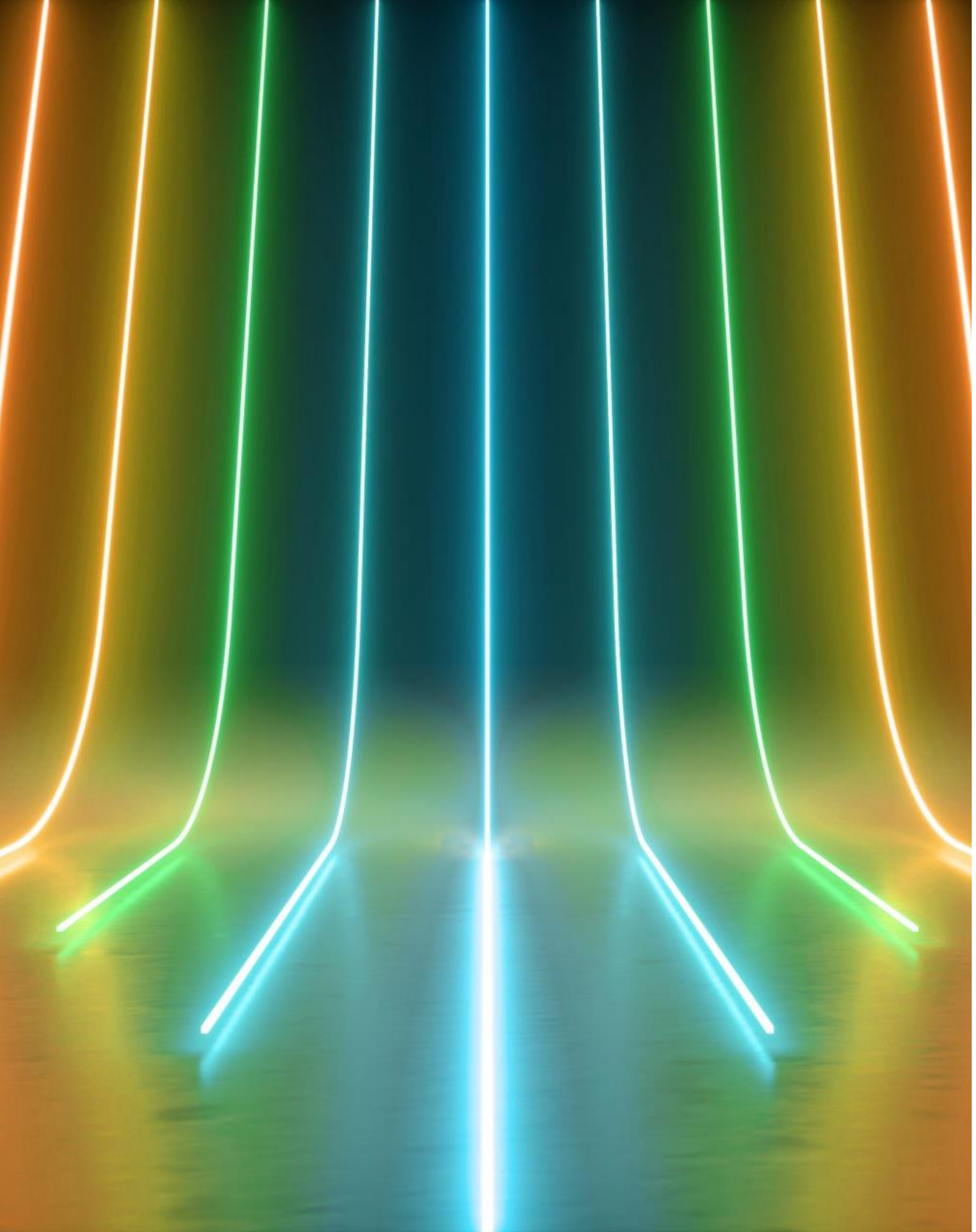
Incomplete Blocks Papers



F. Yates (1936) Incomplete randomized blocks. Annals of Eugenics, 121-140.

Yates and Cochran
On combination of studies

- The analysis of groups of experiments. Journal of Agricultural Science **1938**;*28(4),*556-580.
- Covers combination of results from a series of similar experiments
- Describes the methodological approaches to what we would now call meta-analysis
- Much of the technical details given in Cochran's earlier paper in JRSSB (1937)



Y&C On Representativeness

- “..it is usually impossible to secure a set of sites selected entirely at random…the deliberate inclusion of sites representing extreme conditions may be of value. Lack of randomness is then only harmful in so far as it results in the omission of certain types and in the consequent arbitrary restriction of the range of conditions. In this respect scientific research is easier than technical research.”
- P558

What happened and is happening in medicine

Operating tables and operations with tables

How Agricultural Research Differs from Medical Research

Agriculture

- DF for estimating error usually scarce
- Experiments of similar size
- Many treatments per trial
- Complex treatment structure
- Process of measurement *relatively* easy

Medicine

- DF for estimating error abundant
- Trials of *greatly* differing sizes
- Few treatments per trial
- Simple treatment structure
- Process of measurement difficult
  - Ethics, missing values

Another Important Difference

- The difference between agricultural and medical research is that agricultural research is not done by farmers

Michael Healy

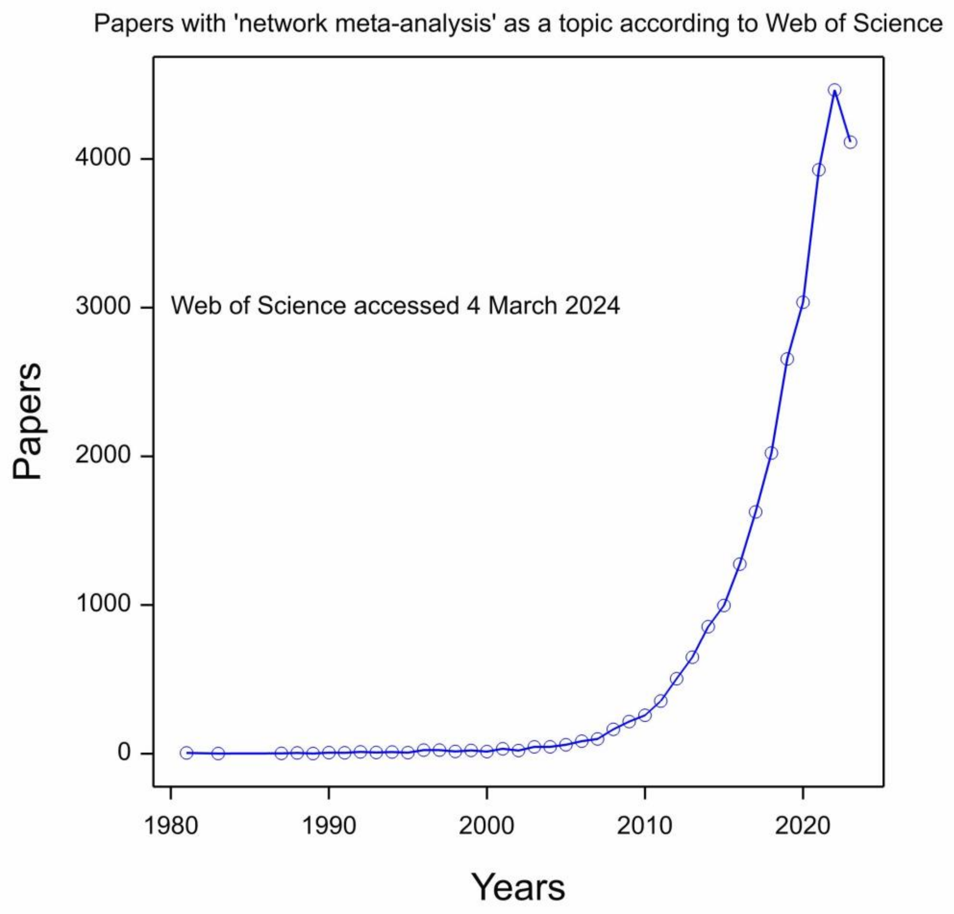
But Some Similarities

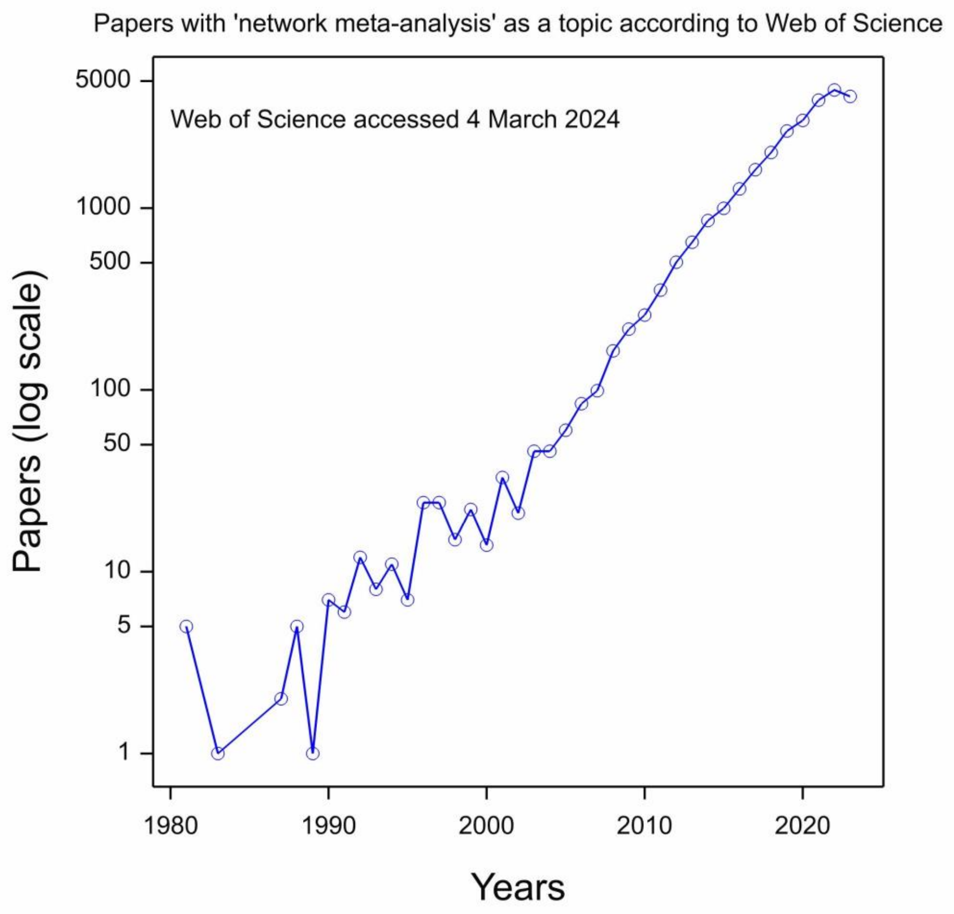
- Variability of material
  - Main effects of fields/centres, experiments/trials
- Limited ability to replicate
  - More serious in medicine
- Experiment/trial by treatment interaction
  - Possibly more serious in agriculture
- Modest treatment differences can bring substantial long term gains

A Timeline of Meta-Analysis in Medicine

When	Who	What
1904	Karl Pearson	Analysis of studies in inoculation for enteric fever
1971	William Cochrane	*Effectiveness and Efficiency: Random reflections on Health Services*
1976	Eugene Glass	Coins *Meta-Analysis*
1985	Hedges & Olkin	*Statistical Methods for Meta-analysis*
1986	DerSimonian & Laird	Random effects meta-analysis
1989	Enkin, Keirse & Chalmers	*A Guide to Effective Care in Pregnancy and Childhood*
1992	Iain Chalmers	Appointed director of the UK Cochrane Centre
1993	Iain Chalmers and others	Set up the Cochrane Collaboration
1999	Irotsu and Yamada	Calculating indirect odds ratios (a form of Network meta-analysis)
2000-		Increasing use of Network MA in health technology assessment

Network Meta-Analysis





Conclusions

Means and ends

Why do we overlook the origins of Meta-Analysis? *Speculation*

- Of course, we are all familiar with regression and its many forms.
- However, the form that students usually approach first is that of determining two unknowns, gradient and intercept, in an overdetermined linear system.
- In fact, the student is more likely to encounter the single parameter problem as estimating the gradient only (regression through the origin) than estimating the intercept.
- But estimating a mean from a number of independent determinations is equivalent to estimating the treatment effects from a number of studies.
- And the former is the simplest application of least squares.

Lessons from the history of combining observations

- The theory was originally developed in connection with astronomy.
  - Gauss, Legendre, Laplace.
- It moved from astronomy into agriculture.
  - Student, Wood and Stratton, Fisher and then Yates and Cochran.
- The astronomers were then soon outdated. They were not aware of what was going on in agriculture.
  - Already when Wood and Stratton applied techniques of astronomy to agriculture, they were behind Student in what they were doing.
- It then moved into medical statistics via education.
  - But the medical statisticians were often unaware of the history and re-invented much that had already been better done in agriculture.
  - Incomplete block theory had much of what now is in network analysis but better and deeper by the 1950s.
  - There was some progress but also some poor work.
- There was a forgetting between disciplines and also within.
- Greater interdisciplinary awareness would be valuable to mankind!