

## George Udny YULE

b. 18 February 1871 - d. 26 June 1951

**Summary.** Yule's coefficient of association, Yule's paradox (later called Simpson's paradox) and the Yule process all perpetuate the name of a pioneer statistician who also contributed to Mendelian theory and time-series analysis.

George Udny Yule was born in Beech Hill, near Haddington, Scotland. He came from an old-established family of army officers, civil servants and orientalists. Both his father and his uncle had been knighted. He was educated at Winchester College from where, at the age of only 16, he transferred to University College, London, to study engineering. Karl Pearson (q.v.), then Professor of Applied Mathematics, was beginning to develop his own interest in statistics, and in 1892 offered Yule a post as demonstrator. In 1896 Yule was appointed Assistant Professor of Applied Mathematics, a post he held for three years until he resigned it in favour of more remunerative employment.

His continuing interest in statistics, however, led to his appointment as Newmarch Lecturer in Statistics at University College in 1902, a post which he held concurrently with his other work until 1909 and which led to the publication of the book which made his name, *An Introduction to the Theory of Statistics*. The first edition appeared in 1911 [4] and during his lifetime there were thirteen further editions. The eleventh edition was the first to be jointly undertaken with M.G. Kendall [7], and by the time of the fourteenth and last edition of Yule and Kendall, 1950, Kendall's own two-volume *The Advanced Theory of Statistics* was already establishing itself. It, and its present-day descendants, still bear the marks of Yule's pioneering effort.

In 1912 a Lectureship in Statistics was established for Yule by the University of Cambridge, to be held in the Faculty of Agriculture, and this, coupled with a Fellowship at St John's College from 1922, provided him with congenial employment (save for the war years) until 1931 when he retired, by then Reader in Statistics. At that stage he bought an aeroplane and learnt to fly. He kept up his College teaching until the second war, and died in Cambridge in 1951.

A major topic of Yule's early statistical work was pauperism, and he concentrated on the way administrative reforms reduced observed rates. He distanced himself from standard eugenic interpretations.

Yule played an important part in the affairs of the Royal Statistical So-

ciety, of which he was honorary secretary for twelve years and subsequently President (1924-26). He was elected a Fellow of the Royal Society in 1922.

Yule's main contributions in the theoretical field were concerned with regression and correlation, association in contingency tables, Mendelian genetics, epidemiology and time series. In the Pearsonian fields of regression and correlation he gave more prominence than his mentor to the former, perhaps easing the path towards R.A. Fisher's (q.v.) invention of the analysis of variance.

Yule's studies of the correlation of continuous variables led him, in 1900 [1], to study measures of association for discrete variables, in particular the cross-ratio  $c$  ("odds ratio") in a  $2 \times 2$  contingency table and its transform  $Q = (1 - c)/(1 + c)$ , now known as "Yule's coefficient". This led to an altercation with Pearson in which Pearson's capacity for acrimonious and ill-directed criticism was displayed, in marked contrast to Yule's gentler mode of expression. Even Fisher, who as a young man had also felt the sharpness of Pearson's pen, was later moved to remark "Pearson attacked Yule's work at one time much more violently than ever he did mine". In 1903 [3] Yule, making use of his understanding of partial correlation, described what was much later to be termed "Simpson's paradox", in which the pairwise associations at two levels in a  $2 \times 2 \times 2$  table can be seemingly incompatible with the marginal association. In his work on association, all his numerical examples were drawn from biology.

In Mendelian genetics, Yule [2] was the pioneer in suggesting that the observed correlations between parent and offspring could be accounted for by multifactorial Mendelian inheritance, as Fisher fully acknowledged in his classic treatment of the correlation between relatives in 1918. In 1914, in collaboration with F.L. Engledow, Yule invented the method of minimum  $\chi^2$  for estimating a genetic recombination fraction, and the following year, with M. Greenwood, he was the first to recognize that there was something wrong with Pearson's  $\chi^2$  test of association in respect of the degrees of freedom used, as Fisher was later to prove. He introduced the simple birth process of stochastic theory (the "Yule process") in connection with evolution in 1924 [5]. He also introduced the correlogram of a time series and laid the foundations of the theory of autoregressive processes. His name is remembered in the "Yule-Walker equations" which provide least squares estimators of the parameters of an autoregressive process.

F. Yates ended his Royal Society obituary notice of Yule with

“We may ... justly conclude that although Yule did not fully develop any completely new branches of statistical theory, he took the first steps in many directions which were later to prove fruitful lines for further progress. ... In the biological field ... his work provided a corrective to many of the errors committed by the biometric school, and served to spread the use of statistical methods amongst biologists who might otherwise have been wholly repelled by them. He can indeed rightly claim to be one of the pioneers of modern statistics”.

### References

- [1] Yule, G.U. (1900). On the association of attributes in statistics: with illustrations from the material of the Childhood Society, &c. *Philosophical Transactions of the Royal Society of London (A)*, **194**, 257-319.
- [2] Yule, G.U. (1902). Mendel's Laws and their probable relations to intra-racial heredity. *New Phytologist*, **1**, 193.
- [3] Yule, G.U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika*, **2**, 121-134.
- [4] Yule, G.U. (1911). *An Introduction to the Theory of Statistics*. Griffin, London.
- [5] Yule, G.U. (1924). A mathematical theory of evolution, based on the conclusions of Dr J.C.Willis. *Philosophical Transactions of the Royal Society of London (B)*, **213**, 21.
- [6] Yule, G.U. (1927). On a method of investigating the periodicities of disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society (A)*, **226**, 267-298.
- [7] Yule, G.U. and Kendall, M.G. (1937). *An Introduction to the Theory of Statistics*. Griffin, London.

### Bibliography

Yule, G.U. (1971). *Selected Papers of George Udny Yule*. A. Stuart and M.G. Kendall (eds.). Griffin, High Wycombe. (This contains, in particular, a reprinting of Kendall's obituary article from *Journal of the Royal Statistical Society, Series A* **115**(1952), 156-161.)

A.W.F. Edwards