A Statistical Biography of George Udny Yule

A Statistical Biography of George Udny Yule:

A Loafer of the World

By

Terence C. Mills

Cambridge
Scholars
Publishing

A Statistical Biography of George Udny Yule: A Loafer of the World

By Terence C. Mills

This book first published 2017

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2017 by Terence C. Mills

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-4438-5067-5
ISBN (13): 978-1-4438-5067-4

CONTENTS

List of Tables ix

List of Figures xiii

Preface xviii

Chapter One 1
Udny Yule: A Sketch of the Life and Career of a Statistician

Chapter Two 11
Skew Curves, Epidemiology and Accident Distributions

Chapter Three 42
Correlation and Regression

Chapter Four 96
The Association of Attributes

Chapter Five 128
Vital Statistics

Chapter Six 187
Nonsense Correlations between Time Series

Chapter Seven 244
Periodicity in Time Series and the Lambdagram

Chapter Eight 277
Heredity and Evolution

Chapter Nine 338
The Statistical Analysis of Literary Style and Other Research

Chapter Ten 387
An Introduction to the Theory of Statistics

Chapter Eleven 468
Modern Perspectives on Yule's Legacy to Statistics

References 483

Index 496

LIST OF TABLES

Table 3.1	Tables illustrating the correlation of the rate of total pauperism with the proportion of out-relief given.
Table 3.2	Means of successive rows of the surfaces in Table 3.1.
Table 3.3	Correlation table of the rate of total pauperism with the proportion of out-relief given for “mainly rural” unions.
Table 3.4	Summary statistics and correlation coefficients for pauperism data.
Table 3.5	Table of the regression of pauperism on other variables for all groups of unions.
Table 3.6	Table of balance sheets.
Table 4.1	Yule’s small-pox tables.
Table 4.2	Proportions calculated from the three tables of Table 4.1.
Table 5.1	Reduction of the Legitimate Birth-Rate to (1) the Total Proportion of Married Women in 1901; (2) the Proportion of Married Women under 45 years of Age.
Table 5.2	Correction of the Legitimate Birth-Rate using Tait’s Law (Yule, 1906b, Table V).
Table 5.3	Showing the number of registrations districts in England and Wales exhibiting (1) a given proportion of male births, (2) a given total number of births during the decade 1881-90.
Table 5.4	Sample statistics from the correlation table of Table 5.3.

Table 5.5 ENGLAND AND WALES: Census populations, 1801 to 1911, and the values given by logistics fitted by the three different methods: (1) Three censuses; (2) Sums of reciprocals; (3) Percentage increases; together with the errors of the logistic values.

Table 5.6 ENGLAND AND WALES: Percentage increases of the population over successive periods of 10 years, 20 years, etc., up to 50 years.

Table 6.1 Deviations from the mean of the sample in samples of 10 terms from a random series, averaging separately samples in which the first deviation is positive and samples in which the first deviation is negative: average of first deviations taken as +1000.

Table 6.2 Coefficients of the terms in the deviations from the mean of the sample, in a sample of 10 terms from a series with random differences *a*, *b*, *c*, …, *l*.

Table 6.3 Coefficients between deviations from the mean of the sample, in a sample of 10 terms from a series of random differences.

Table 6.4 Deviations from the mean of the sample in samples of 10 terms from a series with random differences, averaging separately samples in which (*a*) first deviation is +, (*b*) first deviation is –, (*c*) last deviation is +, (*d*) last deviation is –. The average of first or last deviations, respectively, called +1000.

Table 6.5 Coefficients of the terms in the deviations from the mean of the sample, in a sample of 10 terms from a series of which the second differences are random.

Table 6.6 Coefficients between deviations from the mean of the sample, in a sample of 10 terms from a series of which the second differences are random.

Table 6.7 Deviations from the mean of the sample, in samples of 10 terms from a series of which the second differences are random, averaging separately samples in which (*a*) first deviation is +, (*b*) first deviation is –, (*c*) last deviation is +, (*d*) last deviation is –. The average of first or last deviations, respectively, called +1000.

Table 6.8 Comparison of serial correlations for three series with random differences, with fitted arithmetical progressions.

Table 6.9 Comparison of serial correlations for three series with correlated differences, with fitted cubic series.

Table 7.1 Decomposition of the first 30 terms of the simulated series used in Figure 7.2 into complementary function (simple harmonic function) and particular integral (function of the disturbances alone).

Table 7.2 Means and standard deviations of disturbances in successive periods of 42 years. (Y) corresponds to periods investigated by Yule (1927, Table II).

Table 7.3 Serial correlations of the sunspot numbers and the deduced partial correlations for the extended sample period 1700–2011. In the serial correlations, 1 denotes the correlation between x_t and x_{t-1}, i.e., $r(1)$, and so on. In the partial correlations, 2.1 denotes the correlation between x_t and x_{t-2} with x_{t-1} constant, i.e., $r(2 \cdot 1)$, and so on.

Table 8.1 Hybrids of Japanese waltzing mice × albinos paired together: expectation of albino 25%. Observed (roman) and calculated (italic) numbers of litters with 0, 1, 2, … etc. albinos. Data from Darbishire (1904).

Table 8.2 Proportional frequencies of numbers of genera with 1, 2, 3 species after 0, 1, 2, 3 intervals of time, the genera being initially monotypic.

Table 8.3 Giving particulars in respect of four frequency distributions for size of genus.

Table 8.4 Limit values (time infinite) of mean and modal ages and standard deviations for genera of 1 to 100 species $\rho = 1.5$. Unit of time, the doubling period for species within the genus.

Table 8.5 *Chrysomelidæ*. Table showing mean ages, etc., for genera of each size: unit of time, the doubling period for species within the genus in Chrysomelidæ. $\tau = 6.28$, $\rho = 1.925$.

Table 9.1 Frequency distribution of sentence length from Bacon's *Essays.*

Table 9.2 Constants for the distributions of sentence-length in samples from Bacon, Coleridge, Lamb and Macaulay.

Table 9.3 Constants for the distributions of sentence-length in samples of *De Imitatio Christi*, from miscellaneous works of Thomas à Kempis, and from Gerson.

Table 9.4 Showing the number f_X of nouns occurring *X* times in a sample of 8203 occurrences of a noun spread over the miscellaneous works of Thomas à Kempis other than *De Imitatione Christi.*

Table 9.5 Summary of Yule's sampling experiment, based on his Table 4.5 (*ibid.*, page 64).

Table 9.6 Cumulative sampling experiments: summary statistics.

LIST OF FIGURES

Figure 2.1 Frequency distribution of pauperism in 1891 with Type I and normal fits.

Figure 2.2 Number of deaths from enteric fever at different ages.

Figure 2.3 Frequency distribution of bacterial densities when 100 c.c. gives a negative result.

Figure 2.4 Frequency distributions of bacterial densities for 100 c.c. +, 50 c.c. –, the curve with the lower maximum, (2) 100 c.c. –, 50 c.c. +, the curve with the higher maximum.

Figure 2.5 Frequency distributions of bacterial densities for 1 positive out of 2 and 2 positive out of 4 (1 c.c. in each test).

Figure 2.6 Graphs of $_{\kappa}r_{A\lambda}$ as a function of κ. The higher graph has $\overline{A}/\sigma_{\lambda}^2 = 2$; the lower $\overline{A}/\sigma_{\lambda}^2 = 10$.

Figure 5.1 Persons married per thousand living, 1850 – 1901.

Figure 5.2 Fluctuations in the marriage-rate and trade: deviations from 9-year-moving average.

Figure 5.3 Fluctuations in the marriage rate and bank clearings: deviations from 9-year moving average.

Figure 5.4 Fluctuations in the marriage-rate and prices (Sauerbeck's Index Numbers): deviations from 11-year moving average.

Figure 5.5 Fluctuations in the marriage-rate and unemployment (Hartley's, 1904, index numbers): deviations from 11-year moving average.

Figure 5.6 Sauerbeck's Index Numbers.

Figure 5.7 Marriage-rate and price movements: difference of Sauerbeck's Index Numbers for year *n* from that of year $n - m$.

Figure 5.8 Birth-rates per thousand living.

Figure 5.9 Fluctuations (1) in the Marriage-Rate and (2) in the Birth-Rate; Deviations from 11-Year Means.

Figure 5.10 (1) Calculated oscillations in birth-rate due to oscillations in marriage-rate alone. (2) Actual oscillations in birth-rate, smoothed.

Figure 5.11 The logistic curve for growth of population.

Figure 5.12 The logistic curve fitted to the population of England and Wales, 1801-1921.

Figure 5.13 The percentage increases in the population of England and Wales over the half-centuries 1801-51, 1811-61, etc., plotted to the terminal populations of the half centuries, and the line corresponding to the logistic of Figure 5.12.

Figure 5.14 The populations of the United States (1790-1820), England and Wales (1801-1921), and France, including Alsace-Lorraine (1801-1911), fitted to the same logistic drawn to the standard time-scale.

Figure 5.15 The course of the death-rate (lower curve) and the birth-rate (upper curve) in a logistic population in which the death-rates at ages are constant, the population having followed the logistic for an indefinite period in the past.

Figure 5.16 The course of the death-rate (lower curve) and the birth-rate (upper curve) in a logistic population in which the death-rates at ages are constant, when the population, after having attained the stable state, is suddenly jerked on to the logistic at the epoch –4.

Figure 6.1 An idealistic representation of a time series as the sum of trend, cyclical and irregular components.

Figure 6.2 Correlation between standardized mortality per 1,000 persons in England and Wales (circles), and the proportion of Church of England marriages per 1,000 of all marriages (line), 1866–1911. $r = +0.9512$. (Recreated from Yule, 1926, Fig. 1, page 3).

Figure 6.3 Two sine curves differing by a quarter-period in phase, and consequently uncorrelated when the correlation is taken over a whole period.

Figure 6.4 Variation of the correlation between two simultaneous intervals of the sine curves of Figure 6.3, as the centre of the interval is moved across from left to right.

Figure 6.5 Variation of the correlation coefficient between two simultaneous finite intervals of the harmonic curves of Figure 6.3, when the length of the interval is 0.1, 0.3, …, 0.9 of the period, as the centre of the interval is moved across from left to right; only one-eighth of the whole period shown.

Figure 6.6 Frequency distribution of correlations between simultaneous intervals of the sine curves of Figure 6.3 when the interval is, from the top, 0.1, 0.3, 0.5, 0.7 and 0.9, respectively, of the period.

Figure 6.7 Frequency distribution of correlations between two simultaneous intervals of sine curves differing by 60° in phase (correlation over a whole period $+0.5$) when the length of interval is 0.2 of the period.

Figure 6.8 Three random series.

Figure 6.9 Three series with random differences (conjunct series with random differences).

Figure 6.10 Serial correlations up to $r(10)$ for three experimental series (of 100 terms) with random differences.

Figure 6.11 Three series with positively correlated differences (conjunct series with conjunct differences).

Figure 6.12 Serial correlations up to $r(10)$ for three experimental series (of 100 terms) with positively correlated (conjunct) differences.

Figure 6.13 Frequency distribution of 600 correlations between samples of 10 observations from random series.

Figure 6.14 Frequency distribution of 600 correlations between samples of 10 observations from conjunct series with random differences.

Figure 6.15 Frequency distribution of 600 correlations between samples of 10 observations from conjunct series with conjunct differences.

Figure 6.16 Beveridge's index numbers of wheat prices in Western Europe, 1545–1844.

Figure 6.17 Serial correlations up to $r(40)$ for Beveridge's index numbers of wheat prices in Western Europe, 1545–1844.

Figure 6.18 Serial difference correlations $r^h(k)$ for the index numbers of wheat prices in Western Europe; intervals for differencing $h = 1, 5, 6, 11$ and 15 years respectively.

Figure 6.19 Serial difference correlations for $h = 5$ ($r^5(k)$) (dots) and a curve constructed from certain of the periodicities given by Beveridge (dashed line).

Figure 7.1 Graphs of simple harmonic functions of unit amplitude with superposed random fluctuations: (a) smaller fluctuations; (b) larger fluctuations.

Figure 7.2 Graph of a disturbed harmonic function, equation (7.5).

Figure 7.3 Graphs of the sunspots and graduated numbers, and of the disturbances given by equation (7.8): the lines on the disturbance graphs show quinquennial averages.

Figure 7.4 Scatterplot of $x_t + x_{t-2}$ (horizontal) on x_{t-1} (vertical).

Figure 7.5 Graphs of the disturbances given by equation (7.10): the lines on the graphs show quinquennial averages.

Figure 7.6 Graph of the square of a damped harmonic vibration, (7.12).

Figure 7.7 Graph of a series of superposed functions of the form of Figure 7.6, each one starting when the one before reaches its first minimum.

Figure 7.8 Lambdagram for a correlated series formed by summing the terms of a random series in overlapping groups of five.

Figure 7.9 Calculated lambdagrams for a variety of time series.

Figure 8.1 Double logarithmic chart for the frequency distribution of sizes of genera in the *Leguminosæ*.

Figure 8.2 Double logarithmic graph of the frequency distribution of size of genus for $\tau = 1$, $\rho = 1.925$, full line; the broken line showing the limiting form at $\tau = \infty$ for comparison.

Figure 8.3 Double logarithmic graph (full line) of the frequency distribution of size of genus for $\tau = 3$, $\rho = 1.925$.

Figure 8.4 Limiting forms of the frequency distributions of age for genera of $1, 2, 3, \ldots, 10$ species when time is infinite; $\rho = 1.5$.

Figure 8.5 *Chrysomelidæ*. Frequency distributions of age for genera of 1, 2, 3, 10, 60 and 100 species.

Figure 9.1 Vocabulary ratios as a function of sample size.

PREFACE

I had long been aware of George Udny Yule's seminal influence on time series analysis and econometrics but during the research for my books on the history of the subject, *The Foundations of Modern Time Series Analysis* and *A Very British Affair: Six Britons and the Development of Time Series Analysis*, I quickly realized that not only was Yule a wonderful expositor but that he had also published equally important research in an extraordinarily wide range of fields, from developing the theory of correlation and regression to providing mathematical models of evolutionary behavior, and from analyzing data on pauperism to using statistical methods to resolve cases of disputed authorship of medieval manuscripts. This chimed well with my own predilection for applying statistical methods to whatever might pique my interest and for working with colleagues from many disciplines.

Yet little had been written about Yule and his work, apart from a few scattered articles, since his death in 1951 and the two obituaries that appeared in the following year. As I was, at this time, preparing to take "flexible" retirement from my university position, with its consequent reduction in teaching and administrative duties, this was clearly an opportune moment to redress the balance and to embark on the first major study of Yule's statistical research and subsequent legacy. I have enjoyed the project immensely, particularly as modern information technology and software now enables research to be undertaken away from the confines of

the office and library. Imagine my further delight when I read Yule's 1920 article in the *Cambridge Review*, "The wind bloweth where it listeth", in which he memorably places "Tied Research Workers" alongside "Lecturers, Proctors, Vice-Chancellors and other unfortunate drudges", in contrast to "free" researchers, who were able to follow their own interests at their own leisure–a situation that I now found myself in after years of working within the pernicious constraints of the U.K. universities Research Assessment Exercise/Research Excellence Framework! Yule coined the phrase "loafers of the world" to describe such free spirits of academe and I am delighted to have become a member of this increasingly rare breed and to have borrowed this phrase as part of the title of the book.

December 2016
Quorn, Leicestershire, England
Clermont, St. James, Barbados
Svoronata, Kefalonia, Greece

CHAPTER ONE

UDNY YULE: A SKETCH OF THE LIFE AND CAREER OF A STATISTICIAN

Forbears, early life and education

1.1 George Udny Yule was born on February 18, 1871 at Beech Hill, a country house in Morham, near Haddington, East Lothian, some 16 miles (25 kilometres) east of Edinburgh. The tiny hamlet of Morham has now vanished, as has the house, which was destroyed in 1944 by an aircraft crashing into it. Yule was born into an established Scottish family composed of army officers, civil servants, scholars and administrators, although his more distant ancestry had their roots in farming.[1] The name Udny, by which he was always known, derives from an ancestor from the family of Udny of that ilk.[2]

His grandfather, Major William Yule (1764–1839), went to India as an army cadet in 1781, returning home in 1806 after having served as Assistant-Resident of the East India Company in Lucknow and later Delhi.

---

[1] Biographical and educational details are drawn heavily from the two obituaries of Yule written by Maurice Kendall and Frank Yates (see Kendall, 1952, and Yates, 1952, respectively).

[2] Clan Udny is an armigerous clan (a clan without a chief) originating from the village of Udny (now Udny Green and Udny Station) in Aberdeenshire. The name is often pronounced locally as "Widney", although there is no suggestion that Yule used this pronunciation.

During this time William collected over 250 Persian, Arabic and Hindustani manuscripts and subsequently became widely known for his extensive oriental learning, having translated and printed a lithographic facsimile of the manuscript *Apothegms of Ali the Son of Abu Talib* in 1832.[3] The manuscript collection was later presented by his sons to the British Museum in two tranches, the first in 1847 and the second in 1860 (see Rieu, 1883, pages xviii-xix): a plaster portrait bust of William is still held by the Museum, although it is not currently on display.

Of his three sons, Robert (1817–1857) was killed in action in Delhi commanding the 9th Hussars during the Indian Mutiny. Henry (1820–1889), the youngest brother, became a colonel in the Royal Engineers and was knighted in the year of his death. Retiring from the army in 1862, he later published many travel books, including translations of the work of Marco Polo, for which he was awarded the Founder's Gold Medal of the Royal Geographic Society in 1872, edited several volumes for the Hakluyt Society, a text publishing society of which he was president from 1877 until his death, and was the author of *Hobson-Jobson*, a glossary of Anglo-Indian words and phrases. The eldest son, Udny's father, also christened George Udny (1813–1886), was a member of the Bengal Civil Service, in which, after performing admirably during the Indian Mutiny, he became chief commissioner of Oudh, was subsequently Resident of Hyderabad, and later served on the Governor-General's Council, this distinguished service being rewarded with a knighthood in 1866 before his retirement in

[3] Ali was the cousin and son-in-law of the Prophet Muhammad and the fourth caliphate (successor to Muhammad), reigning from 656 to 661. The question of his right to the caliphate resulted in the split in Islam into Sunni and Shi'ite branches and he is revered by Shi'ites as the first imam, the true successor to the Prophet (see McKenna, 2009, pages 55-57).

1868. Sir George married Henrietta Peach Pemberton (1833–1912), daughter of Captain Robert Boileau Pemberton of the Bengal Army, in 1862; Udny was the youngest of three siblings, his two sisters, Mary Margaret and Henrietta Rose, were born in 1866 and 1870 and pre-deceased him in 1914 and 1927 respectively.

1.2 When Yule was about four years old the family left Scotland and moved to London, living first in Tooting and then at 30 Clanricarde Gardens, Bayswater, where they remained until his father's death following a fall on an icy pavement outside the home.

After attending a day-school in Orme Square, Bayswater, Yule spent three years at a preparatory school at Dunchurch, near Rugby, before, from the age of thirteen, attending the famous public school Winchester College. It was here that Yule developed an interest in physics, although he confessed to not enjoying his schooldays as he was "no use whatever at games or sports" (Yates, 1952, page 309).

Although it had been intended that he should try for an army career in the Royal Engineers, the young Yule objected and pleaded to be allowed to take civilian engineering. As a consequence he left Winchester at the unusually young age of sixteen and entered University College, London (UCL) in 1887, where he was a student for three years. At that time there was no engineering degree and, indeed, few engineering students even tried for a degree, as this meant studying for several more years. Yule himself never contemplated doing this, although he did find the atmosphere at UCL congenial, and it was during this time that he first

made the acquaintance of Karl Pearson, then Professor of Applied Mathematics.

Becoming a statistician

1.3 On leaving UCL, Yule spent a year as a pupil in a small engineering works (George Wailes and Son) and another at the Arc Works, Chelmsford, of Crompton and Co. but, in the summer of 1892, he decided that he was not cut out to be an engineer as he wanted "to find things out, not to make them" (*ibid.*, page 310). As a result of an introduction from Carey Foster, Professor of Physics at UCL, Yule travelled to Bonn to work for a year under Heinrich Hertz, the German physicist who first conclusively proved the existence of electromagnetic waves. It was during this visit that Yule undertook his first scientific research, on the passage of electric waves through electrolytes. On the face of it this was a successful research project, for Yule published several papers based upon it (see, for example, Yule, 1893a, 1893b, 1895a, 1895b), but clearly experimental physics held as little hold on him as engineering did, for he never wrote on either subject again. Indeed, the only influence that Yule's engineering and physics background had on his later work was his careful and expert draughtsmanship, a preference for diagrammatic representation and a facility for constructing various mechanical devices for carrying out what would now be known as "simulation experiments".

1.4 Towards the end of his year in Bonn, Karl Pearson offered Yule the post of demonstrator at UCL, which he accepted and in which position

he remained until being promoted to Assistant Professor of Applied Mathematics in 1896. Although Yule looked to continue his research into electric waves on his return to UCL, this was at the time when Pearson, by now also the Gresham Professor of Geometry, was becoming increasingly interested in statistics and giving lecture courses on the subject (details and reflections of the content of these courses were later provided by Yule, 1897a, 1936a, 1938a). Yule was naturally drawn to statistics, partly because of Pearson's enthusiasm for the subject and his inspirational teaching, and partly through having found, at long last, a subject that satisfied both his interests and skills. It was not long before Yule began making fundamental contributions himself, both in the fitting of skew distributions and in the development of a framework for regression and correlation, topics that are treated in detail in Chapters 2 and 3. His long connection with the Royal Statistical Society (RSS) began at this time with his election as a fellow in 1895 and at the time of his death he had been a fellow for 56 years.

1.5 As Assistant Professor, Yule found himself lecturing on subjects such as hydrostatics and spherical astronomy, subjects that he had never taken as a student. The salary scarcely provided a living wage and his marriage in 1899 to May Winifred, daughter of Dr William Hayman Cummings, Principal of the Guildhall School of Music, prompted Yule to find more remunerative employment. He therefore gave up his post at UCL to become Assistant to Sir Philip Magnus, Superintendent of the Department of Technology at the City and Guilds of London Institute. The duties were almost purely secretarial and Yule found the work dull, but his

interest in statistics remained undiminished and he continued to research in the evenings.

Fortunately, his links with UCL were not severed entirely and in 1902 he became the Newmarch Lecturer in Statistics, a position he held concurrently with his post at the City and Guilds Institute until 1909. During these years his research widened to include the theory of association, the subject of Chapter 4, and his lecturing duties included a course on vital statistics, a subject that he became particularly interested in, with his research in this field providing the material for Chapter 5. Much of his teaching involved lecturing in the evenings to a small class, largely made up of civil servants, and these lectures provided the foundation for his textbook, *An Introduction to the Theory of Statistics*, which was published in 1911. As discussed in detail in Chapter 10, the book went through some fourteen editions, the last, by then co-authored with Maurice Kendall, being published in 1950.

In 1907 Yule became an honorary secretary of the RSS, an onerous position that he held for twelve years, and in 1911 the Society awarded him its highest honour, the Guy Medal in Gold.

Cambridge, academic stability and intellectual maturity

1.6 After marrying May Winifred, Yule lived in Dulwich and later at 28 Great Ormond Street, Bloomsbury, but the marriage was not a success and it was annulled in 1912, there being no children. In that year he was offered the new University Lectureship in Statistics at Cambridge, a post that he accepted with some financial sacrifice but no regrets. On his arrival he also became Statistician to the School of Agriculture and, in 1913, was

made a member of St John's College, becoming a Fellow in 1922, the same year in which he was elected a Fellow of the Royal Society. He resided in the college for the rest of his life except during the last years of illness.

On his election to the College Fellowship he was also appointed College Director of Natural Sciences, a post with the very broadly defined duties of looking after, making friends with and advising students taking natural sciences: "many generations of students of the college … have reason to remember with affection the friendly guidance and council of 'dear old Udny', his good fellowship and his infectious laughter" (Yates, 1952, page 311).

1.7 The First World War interrupted Yule's academic career and in 1915 he was seconded from Cambridge to become statistician to the Director of Army Contracts at the War Office, later moving to the Ministry of Food as Director of Requirements. For this work he was awarded the C.B.E. in 1918, although evidently he did not recall this period with any great affection.

Nevertheless, the war years were not unproductive in terms of research, for with his lifelong friend Major Greenwood he was able to engage in some statistical consultancy with the military and this is discussed in Chapter 2.

1.8 In 1919 Yule returned to Cambridge and the next decade saw him make major theoretical advances in statistics and related topics. Most famous and long-lasting were his papers on time series and these are dealt

with in considerable detail in Chapters 6 and 7. He also had a long-standing interest in heredity and evolution, stemming from the theories of Gregor Mendel that began to gain influence around the turn of the century, and this found an outlet in a series of papers that form the basis of Chapter 8, no doubt nurtured by his association with the agronomists at Cambridge and his friendship with the botanist J.C. Willis, who had retired to the city.

Yule was President of the RSS from 1924 to 1926 and took an active part in the work of other scientific societies. He was a Fellow, and sometime member of the Council, of the Royal Anthropological Institute, from 1928 to 1930 was President of the Cambridge Philosophical Society, and was elected an honorary member of various foreign statistical societies. He was also elected a Fellow of UCL in 1926.

Retirement, illness and new passions

1.9 Yule resigned his University Lectureship at the comparatively early age of sixty in 1931. He claimed that he was weary of lecturing and the demands that it was putting on him, especially as he considered that the subject was becoming increasingly mathematical, a trend that he felt he was unable to follow with the appropriate degree of competency, although this was characteristic of his basic modesty. In the final months of his tenure his post was raised to that of Reader.

After retirement from this post, Yule continued to play an active part in college life for some years. In 1932 he was elected to a College Lectureship and for several years gave courses on statistics for economics students and on vital statistics for medical students. Soon after his retirement, however, his heart, which apparently had never been very

robust, began to give him serious trouble and by 1935 he was a semi-invalid, compelled to be physically inactive. In that year Yule resigned from his College Directorship of Natural Sciences and finally gave up lecturing in 1940.

1.10 Yule was considerably irked by his incapacity and this no doubt affected his research, for he published little of substance during the first half of the 1930s, feeling that the new, primarily mathematical, developments of the subject had overrun him. He was also deeply affected by the sudden death of Karl Pearson in 1936:

> "I feel as though the Karlovingian era has come to an end, and the Piscatorial era which succeeds it is one in which I can play no part." (Kendall, 1952, page 157)[4]

However, as is recounted in more detail in Chapter 10, an opportunity arose to bring in Maurice Kendall as co-author of the *Introduction* and the eleventh edition, published in 1937, and its immediate success, seems to have given Yule a new lease of life. Perhaps the philological interests of his immediate forbears also began to reveal themselves, for Yule now became interested in verifying some doubts cast on the attribution of the classic Catholic devotional book *De Imitatione Christi* to Thomas á Kempis. This prompted a study of the statistical characteristics of an author's style and to other statistical analyses of vocabulary, leading to several papers and his final book, *The Statistics of Literary Vocabulary*,

[4] "Karlovingian" clearly refers to the influence of Karl Pearson on the development of statistics in the late 19th to early 20th century. "Piscatorial" refers to R.A. Fisher and his followers, whose deliberately mathematical approach to the subject Yule found increasingly uncongenial.

published in 1944 in his seventy-third year. Chapter 9 discusses this final area of Yule's research.

1.11 During the 1920s Yule developed a keen interest in motoring, with a taste for fast driving, and this led to a desire to fly, for which he took lessons after his retirement. As it was impossible to arrange insurance on a hired plane for someone over sixty, Yule was forced to buy his own plane before these lessons could begin, but he was nevertheless able to obtain his pilot's flying certificate A in 1931! Unfortunately his flying career was brought to an abrupt and premature end by the occurrence of his heart trouble.

Indeed, as Yule grew older his health steadily deteriorated and even research became too much for him, leading him to spend the last two and a half years of his life in nursing homes, "walking a little, reading a little, corresponding a little, but conscious that his powers were failing, and waiting, not always patiently, for the end" (Kendall, 1952, page 158). This came, in his eighty-first year, on June 26, 1951, in the Evelyn Nursing Home, Cambridge.

1.12 The aim of this sketch of the life of Udny Yule is to provide an introduction to his work and to this book. A full appreciation of his research, his statistical legacy and his personality must wait until Chapter 11.

As can be seen, the book adopts the chapter format of sub-heading and section number for ease of cross-referencing: thus a cross-reference to section **y** of Chapter **x** will be denoted **§x.y** in subsequent chapters.