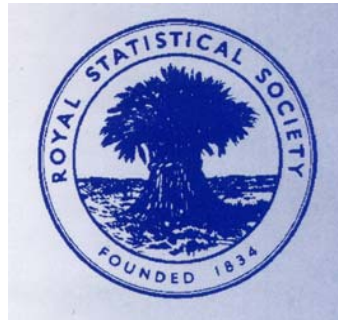


# ***The Royal Statistical Society, Statistics, and the Future***

**David J. Hand  
President: The Royal Statistical Society**



Stockholm 2008

# The Royal Statistical Society: Past

**1833** - The British Association for the Advancement of Science (BA) creates a statistics section, following a presentation by Adolphe Quetelet to fellows, including Thomas Malthus and Charles Babbage.

**1834** - The statistical section of BA establishes a Statistical Society. First officer and Council elected, with Marquis of Lansdowne as President. Original objective '*the collection and classification of all facts illustrative of the present condition and prospects of society, especially as it exists in the British Dominions*'

By end of 1834: 378 Fellows

**1838** - Publication of a Journal.

**1840** - Prince Albert becomes the Society's Royal Patron.

**1858** - Florence Nightingale is the first elected female.

**1887** - The Society is granted a Royal Charter and becomes the Royal Statistical Society.

**1928** - Formation of the first study group (an early version of the Society's sections).

**1934** - The publication of a supplemental journal on statistical methodology, which becomes a series in its own right in 1948.

**1947** - The Society's attempt to gain a supplemental charter to allow professional examinations is blocked by the Royal Economic Society.

**1948** - Society's first conference is held, in Oxford. The Institute of Statisticians (IOS), an organisation dedicated to the interests of professional statisticians is formed.

**1950s** - Leading fellows include economist, J.M. Keynes and Sir William Beveridge, the intellectual godfather of the British welfare state.

**1993** - RSS and IOS merge, retaining the title of the Royal Statistical Society.

## 104 Presidents

The Most Noble The Marquis of Lansdowne, KG, FRS	1834 - 1836
Sir Charles Lemon, Bart, MP, FRS	1836 - 1838
The Rt Hon The Earl Fitzwilliam, FRS	1838 - 1840
The Rt Hon The Viscount Sandon, MP (later Earl of Harrowby)	1840 - 1842
The Most Noble The Marquis of Lansdowne, KG, FRS	1842 - 1843
The Rt Hon The Viscount Ashley, MP (later Earl of Shaftesbury)	1843 - 1845
The Rt Hon The Lord Monteagle	1845 - 1847
The Rt Hon The Earl Fitzwilliam, FRS	1847 - 1849
The Rt Hon The Earl of Harrowby, KG	1849 - 1851
The Rt Hon The Lord Overstone	1851 - 1853
The Rt Hon The Earl Fitzwilliam, KG, FRS	1853 - 1855
The Rt Hon The Earl of Harrowby, KG	1855 - 1857
The Rt Hon The Lord Stanley, MP (later Earl of Derby)	1857 - 1859
The Rt Hon The Lord John Russell, MP, FRS (later Earl Russell)	1859 - 1861
The Rt Hon Sir John S Pakington, Bart, MP, GCB (later Lord Hampton)	1861 - 1863

Nowadays normally two year term of office

Many famous names from the past

- William Farr 1871-1873
- William Guy 1873-1875
- Francis Edgeworth 1912-1914
- George Yule 1924-1926
- Major Greenwood 1934-1936
- William Beveridge 1941-1943
- Austin Bradford Hill 1950-1952
- Ronald Fisher 1952-1954
- Egon Pearson 1955-1957
- Maurice Kendall 1960-1962

# The Royal Statistical Society: Present

- Both a *learned society* and a *professional body*
- The leading UK source of independent advice on statistical issues,
- Through its links with government, academia, corporate and voluntary sectors, one of the most influential.
- Regularly contributes to Royal Commissions, Parliamentary Select Committee and public consultations, offering a unique view on a range of issues, from government targets to sustainable development.
- 18 staff located at the HQ on the edge of the City of London
- Many volunteers

# Mission Statement

*The purpose of the Royal Statistical Society is to develop, foster and disseminate statistical knowledge, methodology and good practice for the benefit of all society*

## **In more detail:**

*To **nurture the discipline of statistics** by publishing a Journal, organising meetings, setting and maintaining professional standards, accrediting university courses and operating examinations*

*To **promote the discipline of statistics** by disseminating and encouraging statistical knowledge and good practice with both producers and consumers of statistics, and in society at large*

*To **provide effective and efficient services** to our members which support their professional and academic interests and their endeavours to advance the other objectives of the Society.*

# Membership

Over 7000 members

About 1/4 outside the UK

Student membership scheme

'Linked associate' membership scheme to allow joint membership by groups of people from particular employers

# Sections

To study a particular area of methodology or applications

- Business and Industrial
- Environmental Statistics
- General Applications
- Medical Statistics
- Official Statistics
- Quality Improvement
- Research
- Social Statistics
- Statistical Computing
- New: 'Young Statisticians' Section

# Local Groups

Avon  
West Midlands  
East Kent  
East Midlands  
Edinburgh  
Glasgow  
Hertfordshire & Bedfordshire  
Highlands  
Lancashire & East Cumbria  
Leeds & Bradford  
Manchester

North Eastern  
Merseyside  
North Ireland  
Oxford  
Reading  
Sheffield  
South Wales  
South West  
Sussex  
Three Country Corner  
York

## Study Groups

To investigate new areas of statistical development or application. May go on to become sections.

- 2011 Census
- Bioinformatics
- Primary Healthcare
- Six Sigma

# Statistics User Forum

A discussion body of organisations which use statistical data

- British Society of Criminology
- British Urban and Regional Information Systems Association
- Business Statistics User Group
- Chartered Institute of Library and Information Professionals
- Demographics User Group
- Financial Statistics User Group
- Fire and Rescue Services User Group
- Health Statistics User Group
- International Trade Statistics User Group
- Labour Market Statistics User Group
- Local Authority Research and Intelligence Association
- Output Area Classification User Group
- National Accounts User Group
- Society of Business Economists
- Transport Statistics User Group
- Market Research Society

# Young Statisticians

(= *Career young*)

Events and activities for students and those in the first 10 years of their career

Two main annual conferences:

- Young Statisticians Conference
- Research Students Conference

# Meetings

Over 150 meetings and events per year throughout the UK

***Ordinary meetings:*** papers read before the Society, and the paper and discussion published in the journal

***Special named lectures:*** Beveridge Lecture, Caradog Jones Lecture, Cathie Marsh Lecture

# Conferences

An annual conference

Normally alternate years general and special topic

- 2008 in Nottingham 1-5 September
- 2009 in Edinburgh 7-11 September  
(the 175th Anniversary Conference)

# Publications

## ***Journal of the Royal Statistical Society: Series A: Statistics in Society***

Material for specialist and non-specialist readers

## ***Journal of the Royal Statistical Society: Series B: Statistical Methodology***

Technical, statistical methodology

## ***Journal of the Royal Statistical Society: Series C: Applied Statistics***

Novel applications of statistical techniques

## ***RSS News***

The Society's Newsletter. 10 editions per year

## ***Significance***

To communicate and demonstrate in an entertaining and thought-provoking way the practical use of statistics in all walks of life and to show how statistics benefit society

## ***And soon ....***

A glossy annual report for public consumption

## External relations

Network of statisticians to act as media contacts

Press releases on statistical topics

Working parties to consider topical issues

Development of considered viewpoints on various issues

e.g.

- Working Party on Statistical Issues in *First-in-Man Studies*
- Establishment of the *UK Statistics Authority* to oversee independence of the GSS from the Government
- Performance monitoring
- Teaching statistics to 14-19 year-olds
- The Sally Clark case

# Honours

*Guy medals* - for advances in statistics

*Chambers medal* - for services to the Society

*Bradford Hill medal* - for advances in medical statistics

*Greenfield industrial medal*

*West medal* - for advances in social or official statistics

*Research prize* - outstanding original contribution, early career

*Honorary fellowship* - to non-statisticians

## Statistical education

- Associate Schools and Colleges Scheme: network of teachers
- Education workshops
- Careers fairs and talks

Accreditation of university courses

As a **Professional Body** the RSS also has its own qualifications

- *Ordinary Certificate* (final school level)
- *Higher Certificate* (early university level)
- *Graduate Diploma* (university degree level)

And runs a **Continuing Professional Development Centre** with a wide range of courses

- *GradStat*
- *CStat*

# The RSS Centre for Statistical Education

Currently based at Nottingham Trent University

*Aim: To promote the improvement of statistical education, training and understanding at all ages*

Carries out a number of research and collaborative projects

# The RSS Schools Lecture (the 'Guy Lecture')

Aimed at 16-18 year olds, repeated around the country  
Previous years can also be repeated

2007: Neil Sheldon: *Statistics and stories: insight, not numbers*

2006: Susan Starkings: *Can you see the wood for the trees?  
Survey sampling, estimation, and approximation*

2005: John Haigh: *Taking chances*

2004: Frank Duckworth: *Lies and statistics*

2003: Peter Holmes: *Statistics in society*

2002: Adrian Bowman: *A world of difference: a rough guide on  
why statisticians count*

***I'm very keen to strengthen links with  
other European statistical societies***

# ***Statistics and the future***

## Statistics is:

- *the technology of extracting meaning from data*
- *the technology of handling uncertainty*
- *the discipline used for predicting the future or for making inferences about the unknown*
- *the discipline of producing convenient summaries of data.*
- *greater statistics [is] everything related to learning from data, from the first planning or collection to the last presentation or report*
- *applied philosophy of science*

# **A unique discipline because of its role in all aspects of life**

- economics
- social policy
- medicine
- science
- etc

# The invention of a discipline

Origins - in the mists of time

e.g.1: King David I of Scotland, 1150 AD, defined the inch as the average of the width of the thumbs of a big man, a medium man, and a small man measured at the base of the thumbnail.

e.g.2: The rood is an old British unit of length connected to the foot.

*The surveyor should stand by a church door on Sunday. When the service ends, he should “bid sixteen men to stop, tall ones and short ones, as they happen to come out ...”. ...These chosen men should be made to stand in line with “their left feet one behind the other”. The sum of the sixteen actual left feet constitutes the length of “the right and lawful rood” and the sixteenth of it constitutes “the right and lawful foot”.*

16th century account



***Mathematical development*** of least squares,  
probability, and other foundational elements:

Gauss, Laplace, Bernoulli, De Moivre, Simpson, Bayes, Poisson

***Society's needs*** to monitor and understand

Quetelet, Jevons, Lexis, Galton

***Other streams*** from empirical analysis

William Farr, John Snow, Florence Nightingale,

**But the merger that produced modern statistics was not without its obstacles:**

***‘...This singular idea of a calculus applicable to things where the ignorance and passion of men are intermingled in an imperfect light is dangerously illusory...’***

Louis Poinsot, 1836

## **Statistics: an evolving discipline**

*Journal* launched in 1838

the early papers are non-mathematical descriptions of pressing social questions: strikes, education, crime, population

By 1870s, strong interest in economic and commercial data

The mathematisation of the discipline began around the end of the 19th Century:

- 1885 paper by Edgeworth appeared which included *'probability, the normal curve, use of the modulus, and the fluctuation, discusses  $n$  or  $n-1$  as divisor, ...discusses significance tests, use of the median, parametric versus non-parametric tests, describes normal and Poisson approximations to the binomial, ... and deals with the tendency of a mean towards normality'*
- 1893 papers:
  - bivariate normal (Edgeworth)
  - correlation coefficient (Galton)
  - standard deviation and skewness (Pearson)

- Fisher (1890-1962)
- Jeffreys (1891-1989)
- Egon Pearson (1895-1980)
- Ramsey (1903-1930)
- Savage (1917-1971)
- de Finetti (1906-1985)
- Neyman (1894-1981)
- etc etc etc

But then a ***new change*** began, roughly around 1960, with the advent of the computer

Replace *months of error-prone hand calculation* by *split-second production accurate results*

Changing the role of the data analyst:

- from concern with arithmetic manipulation
- to concern with interpretation and meaning

Two impacts

- can easily do ***what we did before***
- can do ***entirely new things***

## New kinds of tools

- generalised additive models
  - powerful, flexible
- multivariate adaptive regression splines
  - powerful, flexible
- computationally intensive methods
  - jackknife
  - leave-one-out cross-validation
  - bootstrap
- MCMC
- etc

But this had a more fundamental impact on statistics

⇒ advent of ***new data analytic disciplines***, with new perspectives on data analysis

- *database technology*: storage and manipulation of data
- *machine learning*: modelling natural and artificial learning
- *computational learning theory*: theoretical
- *pattern recognition*: for practical classification
- *data mining*: for large data sets

New classes of data analytic tools, initially developed by non-statisticians . . . . .

. . . . . then subsumed within statistics

Examples:

- neural networks
- genetic algorithms
- expert systems
- support vector machines

***Modern statistics is as much a computational discipline as a mathematical one***

But much of this work,

while certainly within the realm of ***greater statistics***,  
is not being conducted by people who would call  
themselves statisticians

***Does this mean future problems for the discipline  
and profession?***

# The role of applications

***Applications*** have been just as fundamental to the development of statistics as have been ***mathematics*** and ***computation***

In steering the direction of statistics

- ***agriculture*** - experimental design
- ***marketing and social policy*** - survey sampling, clustering
- ***medicine*** - survival analysis, graphical models
- ***psychology*** - factor analysis, linear structural equation models, item response theory
- ***ecology*** - ordination, multidimensional scaling

The applications driver is continuing:

- ***bioinformatics***: large  $p$  small  $n$  problems
- ***banking***: dynamic models
- ***robotics***
- trend towards ***evidence-based social policy?***

New kinds of problems, requiring new kinds of solutions

- dynamic, reactive, real-time
- huge data sets
- small  $n$  large  $p$  problems
- new types of data: text, image, metadata
- . . . . .

**One (controversial) example:**

## ***Elections***

**Stage 1: 19th Century**

Distribute manifesto, public speeches

***What matters is what the voters know about the candidates***

## **Stage 2: 20th Century**

Radio, television, ***broadcasting***

At this stage, still

***What matters is what the voters know about the candidates***

## Stage 3: 21st Century ?

Targeted ***narrowcasting***

*microtargeting*

*political sharpshooting*

- (1) Use sophisticated data analytic tools to target the voters who will make a difference
- (2) Tailor the message to match the voter - concentrating on the issues which matter to the particular voter

***What matters is what the candidates know about the voters***

## **Does this matter?**

Elections are often very close, and inherently unstable?

### ***2004 US Presidential Election***

Electoral College: 53% for Bush; 47% for Kerry

⇒ 100% victory for Bush

### ***2005 German Chancellor Election***

35.2% for Merkel; 34.3% for Schröder

⇒ 100% victory for Merkel

### ***2005 UK General Election***

35% for Labour; 32% for Conservatives

⇒ 100% victory for Labour

(Note: percentage of votes may not correspond to percentage of seats)

Distribution of votes across seats varies

- some seats won by an 80:20 majority
- others won by a 51:49 majority

*It is futile spending time and money campaigning in the 80:20 seat*

*It is futile spending time and money trying to persuade people who won't be persuaded*

***⇒ concentrate effort in those seats and on those voters which can be swayed***

# And the future?

## For statisticians

- professionalisation?
- ethics
- challenges from other domains
  
- reduced misunderstanding ?

# For statistics

Computers are still developing

- Moore's Law
- no end in sight

New applications are still appearing

New kinds of problems, new kinds of data

Data quality !

*Statistics is the most exciting of disciplines*

*Statistical science is experiencing a period of unprecedented development*

*We live in the most exciting of times*

# ***END***

**[d.j.hand@imperial.ac.uk](mailto:d.j.hand@imperial.ac.uk)**

**[http://stats.ma.ic.ac.uk/d/djhand/public\\_html/](http://stats.ma.ic.ac.uk/d/djhand/public_html/)**