THE IMPACT OF PROBABILITY AND STATISTICAL THEORY ON THE DEVELOPMENT OF FORECASTING

N. H. BINGHAM, Imperial College London

Conference on the History of Business and Economic Forecasting,
Reading, 22 March 2013

Anything important enough becomes political (Couve de Murville).

Politics is not an exact science (Bismarck).

Mathematics is an exact science.

1. Pre-1900.

General reference:

[S] S. M. STIGLER, *The history of statistics. The measurement of uncertainty before 1900*. Harvard UP, 1986.

The English school in Statistics:

Francis GALTON (1822-1911): Regression towards mediocrity in hereditary stature. *J. Anthropological Institute 15* (1986), 246-263; *Natural inheritance*, Macmillan, London, 1889. Galton introduced and named the concepts of regression and correlation (originally 'reversion' and 'co-relation') in his classic study of the heredity of human height. He found the bivariate normal model empirically.

Karl Pearson, 1857-1936. *The Grammar of Science*, 1892 (2nd ed. 1900, 3rd ed. 1911). Pearson became Professor of Applied Mathematics at University College London in 1884, and Galton Professor of Eugenics there in 1911. He re-discovered and named the χ^2-distributions,

and used them as the basis for his *chi-squared goodness-of-fit test* of 1900 – arguably, the beginning of modern Statistics; see [S] Ch. 10.
Francis Ysidro Edgeworth, 1845-1926) – taught at King's College London from 1880. He worked on statistics from 1885 on, influenced in particular by Galton and Pearson. In 1892 and 1893 he gave the distribution theory of the multivariate normal distribution, basic to modern statistics; see [S] Ch. 9.
Sir Arthur SCHUSTER (1851-1934): periodogram 1897, 1899 (sunspot data): spectral methods in Time Series.

2. 1900s.
London School of Economics (& Political Science) founded 1895; in University of London 1900.
The Welfare State: David Lloyd George (1863-1945): Chancellor, 1908-1915; People's Budget 1909 (PM 1916-22);
Sir William BEVERIDGE (1879-1963); Director, LSE, 1919-37; Beveridge Report 1942 (Beveridge's statistical assistant was Harold Wilson, PM 1964-70 and 70-74; President RSS, 1972-73).
Sir Arthur BOWLEY (1869-1957), LSE 1895-1936: economic statistics.
G. Udny YULE (1871-1951): spurious correlation, 1926.
Sir Ronald (R. A.) FISHER (1890-1962), Rothamsted, 1920s, Studies in crop variation: agricultural time series.
Early books on Time Series (TS): Yule & Kendall (1st ed. 1911– 14th ed. 1950: TS, Ch. 26, 27); Hannan 1960; Whittle 1963; Kendall & Stuart Vol. III, 1966/68/76.

The new era, 1970:
E. J. HANNAN, *Multiple time series*;
G. E. P. BOX & G. M. JENKINS, *Time series: forecasting and control*.
'Box-Jenkins' moved TS from something regarded as for experts only, and that economists were scared of, to something economists could *use*. ARMA models/model-fitting/model-checking procedures led to an explosive growth in the application of TS, by economists and econometricians.
Kalman filter (state-space models)
Rudolf KALMAN (1930-); Kalman filter, 1960, in engineering (used in real-time control of manned spacecraft); in TS, A. C. Harvey, 1984 on.
Wavelets
Wavelets appear in 'time-frequency analysis', combining 'time domain' and 'frequency domain'. Continuous wavelet transform, 1983 on; discrete WT, 1988 on (Morlet; Mallat; Meyer). Digitization of FBI fingerprint bank, Coifman.

3. Demography; social statistics; the actuarial profession

If it be now, 'tis not to come; if it be not to come, it will be now; if it be not now, yet it will come; the readiness is all: ... [*Hamlet*, V.II].

The founding father of demography is arguably John Graunt (1620-1674), with his *Bills of Mortality* in 1662 (a little after the Pascal-Fermat correspondence of 1654, which marks the 'official beginning' of probability theory).

Let me remind you that the word statistics is cognate with state. As states and government grew more complicated, the data needed to plan and to govern did likewise, leading to the emergence of mathematical statistics in the late 19th C.

The essence of the actuarial profession is the quantitative aspects of insurance, and the essence of insurance is the sharing of risk. The mathematics involved rests on stochastic processes in general and compound Poisson processes in particular; see below.

4. Stochastic Processes (SPs)

Louis BACHELIER (1870-1946); thesis, *Théorie de la spéculation*, 1900. First use of *Brownian motion* (BM) to model the evolution of stock prices – remarkable, as there was then no mathematical theory for stochastic processes in general or BM in particular.

Filip LUNDBERG (1876-1965), Swedish actuary: *Poisson process* (dynamic counterpart of the Poisson distribution) in 1903, collective risk theory.

A. K. ERLANG (1878-1929), Danish telephone engineer, in 1909: Poisson process, telephone traffic.

Norbert WIENER (1894-1964), American mathematician, in 1923: mathematical theory of BM.

A. Ya. KHINCHIN (1894-1959), Russian mathematician: stochastic processes, 1920s/30s.

A. N. KOLMOGOROV (1903-1987); *Grundbegriffe der Wahrscheinlihkeitsrtechnung*, 1933:

rigorous axiomatic foundation for probability theory, via measure theory.
Discrete v. continuous time SPs; Whittaker-Kotelnikov-Shannon sampling theorem.
Stationary processes.
Gabor SZEGŐ (1895-1985) in 1920-21 (mathematical theory);
Ulf GRENANDER and Gabor SZEGŐ, *Toeplitz forms and their applications*, 1958 (Ch. 10, Applications to probability; Ch. 11, Applications to statistics).
Szegő's theorem: criterion for the influence of the remote past to damp out with time.
Best linear prediction of the future given the entire past, or (harder) a finite section of the past.
NHB, Szegő's theorem and its probabilistic descendants. *Probability Surveys* **9** (2012), 287-324.

5. Stationarity; Granger

A basic principle in accountancy is that one should *discount* everything involving money over time, so as to work with real prices rather than nominal prices. Otherwise, as compound interest, inflation etc. involve exponential growth, time series tend to inflate exponentially.

Stationarity is is a strong assumption! There is a Reading connection here. Sir Harry (H. R.) PITT (1914-2005) was VC here from 1964-78. He was a mathematical analyst but also interested in probability and statistics. He was previously at U. Nottingham, where C. W. J. Granger was his student (PhD 1959). One of the great themes of the work of Sir Clive GRANGER (1934-2006; Nobel Prize 2003) was to warn one not to use methods based on stationarity in non-stationary situations. This can lead, via *spurious regression*, to misleading expert advice to politicians, hence to mistaken macroeconomic policies, and hence to massive and irreversible losses in GDP! Think of

Japan's lost decade in the 90s (lost decades now), the West post-2007/8, etc.
Together with Robert ENGLE (1942-), Granger in 1987 introduced *cointegration* (time series that tend to move together), and extended the ARMA models to allow for varying variance (or volatility). Their models included ARCH (autoregressive conditionally heteroscedastic) and GARCH (generalized ARCH), widely used in econometrics, and in volatility modelling in finance (below). Engle and Granger received the Nobel Prize in 2003.

6. Discrete v. continuous time

Is time discrete or continuous? It is both; which we perceive it as depends on how we measure time (sweep second hand or digital watch).

Continuous time is mathematically harder, but smoother, so we can use calculus. Discrete time is mathematically simpler, but we have to fall back on the Calculus of Finite Differences. Data is discrete (we cannot monitor, e.g., prices continuously in practice). And when implementing anything on a computer, one has to discretize.

Rex (A. R.) BERGSTROM (1925-2005):

Continuous-time econometrics, OUP, 1990.

New Zealander economist; PhD 1955; LSE 1962-64; U. Essex 1970-92.

7. Higher dimensions

Harry MARKOWITZ (1927-); Thesis: Portfolio selection. *J. Finance* **7** (1952), 77-91.

Markowitz gave us two insights, which have become ubiquitous:

(a) Think of risk and return together, not separately (risk corresponds to variance, return to mean, hence mean-variance analysis).

(b) Diversify: hold a range – portfolio – of different assets, *balanced* in that it contains lots of negative correlation (so that when things change, losses on some assets will be compensated by gains on others negatively correlated with them).

Markowitz was awarded the Nobel prize for economics in 1990.

So by Markowitzian diversification, an investor should hold a (balanced) *portfolio* of assets. Their evolution over time will be a *multiple time series* (Hannan's book, 1970). The Szegő theory extends to higher dimensions:

NHB, Multivariate prediction and matrix Szegő theory, *Probability Surveys* **9** (2012), 325-339.
1960s, capital asset pricing model (CAPM, "cap-emm"): compares the performance of individual assets to that of the market as a whole, using linear *regression*. For statistical background here, see e.g.
[BF] NHB and John M. FRY, *Regression: Linear models in statistics*, Springer, 2010.

8. Finance and derivatives

Options give one the right but not the obligation to buy/sell a risky asset (stock, say) at a specified price at some specified time in the future. They are widely used for speculation and hedging. Before 1973, it was thought that the question of how much an option was worth was ill-formulated: it would depend on the attitude to risk (utility function) of the prospective purchaser. It turns out that, under assumptions, this is not so. The principal assumption is of *no arbitrage (NA)* (arbitrage is the ability to extract riskless profit from the market, which is parasitical, and renders equilibrium impossible). The *Black-Scholes formula* of 1973 (Fischer BLACK (1938-95); Myron SCHOLES (1941-) gave a formula for option pricing – which changed the world, and has promoted a great growth in the trading of options and other financial derivatives. See e.g.

[BK] NHB and Rüdiger KIESEL, *Risk-neutral*

Valuation: Pricing and Hedging of Financial Derivatives. Springer, 1998/2004.
Black and Scholes used PDE methods. It was shown by Robert MERTON (1944-; Nobel Prize, with Merton, 1997) that *martingales* are the key. These are stochastic processes modelling 'fair games'. They have a special calculus, stochastic calculus or *Itô calculus* (Kiyosi ITÔ (1915-2008) in 1944).
The resulting field of *mathematical finance* (in great demand at MSc level!) rests on three pillars: *arbitrage, martingales, numerics*.
In many markets – stock, bond or forex – the volume of trade in derivatives such as options is far bigger than that in the underlying.
All this is highly mathematical/probabilistic/statistical, but can be summed up in one sentence: (i) discount everything; (ii) take the real-world probability measure P; throw it away, and replace it by a mathematical fiction, Q, equivalent to it (same things possible and impossible), under which discounted asset prices become martingales [BK].

9. Prediction(s)

I gave a talk in Sheffield on 26.3.2013, 'Mathematical finance after the Crash' (on my Imperial homepage, under Talks). I close with some thoughts drawn from it.

The 'big two' of Economics last century were J. M. KEYNES (1883-1946), and one of Friedrich HAYEK (1899-1992, Nobel Prize 1974) and Milton FRIEDMAN (1912-2006, Nobel Prize 1976, U. Chicago 1946-74). Friedman was a fine statistician when young, and is the intellectual driving force behind the 'Chicago boys' (widely regarded as having captured academic Economics), and the neo-cons (Thatcher, Reagan etc.) in politics. He influenced Alan GREENSPAN (1926-; Chairman of the Fed., 1987-06). Read Greenspan's autobiography, *The age of turbulence* (2007) – written just before the roof fell in in the US. It is largely written in a spirit of Panglossian optimism: All is for the best in this best of all possible worlds (Voltaire, *Candide*); markets rule OK and are self-correcting; so nothing much can go wrong. Everything went wrong – and still is wrong.

The real contrast is between the sophistication of the detail here (many of the financial derivatives later called toxic were highly mathematical) and the crudity of the policy mistakes. Here are a few obvious predictions:

bubbles burst, eventually (Mr Greenspan became 'the world's favourite central banker' by presiding over a long asset-price bubble);

gross geo-economic/geo-financial imbalances are unsustainable (from Reagan on, the US ran a large trade deficit; from the Chinese 'New Economic Policy' on, China has a large trade surplus, looking for a good home; Chinese money, aggressively lent by US banks, financed the housing bubble that ended with the sub-prime mortgage crisis in 2007);

big multinational companies can and do outwit national governments over tax (Starbucks, Amazon etc.); this is unsustainable, and needs international action; etc.