1

BRITISH SOCIETY FOR THE HISTORY OF MATHEMATICS

Volume 19 Issue 1 (2004) (= Bulletin 1)

Articles

Roger Thatcher, 'Mathematics in a non-academic life' [address to the BSHM Christmas meeting 13 December 2003]

https://www.bshm.ac.uk/journal-table-contents

MATHEMATICS IN A NON-ACADEMIC LIFE

ROGER THATCHER

This is a talk which was given to the Christmas meeting of the British Society for the History of Mathematics on 13 December 2003

Introduction

I was invited to give a talk about my career. As you will hear, I was a statistician in industry, then a statistician in the Civil Service, then the Registrar General who ran the census in 1981. After I had retired, I did some international research on human longevity, which again was statistical. However, over the years I also wrote some short papers on three historical problems and I will mention these in the middle of the talk, to make a kind of break from statistics.

Early years

I will begin in October 1944, when I went up to Cambridge to start reading for the maths tripos, but expecting to be called up for the Forces within a year. However, during that year the war ended and the rules were changed. I ended up taking the maths tripos at the end of my second year, then reading economics with statistics in my third year, and then going into the Navy.

In 1947 the Navy was looking for maths and physics graduates to be temporary met officers at naval air stations. I was sent on an intensive course in meteorology. We had to learn how to make our own forecasts and to brief pilots on the weather they should expect, particularly the visibility, wind speeds and heights of the cloud base. In those days there were many aircraft which did not have the equipment to land in really poor

conditions and they were often dependent on the forecasts to plan their flights. If a forecast went badly wrong, the consequences could be very serious. At times it could be a bit hair raising, but it was an exciting life.

Gas Board

When I left the Navy, I wrote to the University Appointments Board to ask what jobs were on offer for people like me. They sent me a list and I chose the job with the highest salary. This was £500 per annum, which was reckoned to be a good starting salary in those days, but which did not actually buy very much. The post was for an assistant statistician with the North Western Gas Board. So I moved to Liverpool.

The work was firmly in the field of applied statistics. I've just got time to mention one of the applications, to give the flavour. In those days gas was stored in gas holders, but these would only hold a limited amount. Production therefore had to be planned very carefully, to match the demand as closely as possible. What the engineers really wanted was a forecast of demand by the hour, not just by the day. This was primarily a problem of time series analysis. There was a long term trend, monthly seasonal cycles, weekly cycles, daily cycles and hourly cycles, which clearly showed special peaks like Bank Holidays and the Monday washday. When all these cycles had been removed, there remained a residual which was strongly correlated with temperature. So the forecast for tomorrow started by obtaining from the local weather station their forecast of the temperature tomorrow and then adding back the cycles. The engineers were delighted. It was very striking how the most elementary statistical methods, just moving averages and correlations, could really pay their way.

Admiralty

In 1952 I decided to transfer from the Gas Board to the Civil Service as a statistician. In those days a degree was not enough and the Civil Service set its own examination in statistics. I got through this and was appointed to the Admiralty, this time as a civilian.

The Admiralty was a large Department with a very long tradition. There was a First Lord of the Admiralty, who was a politician. There were also the Sea Lords, who were senior Admirals. The civil servants were in the Secretariat, which had been founded by Pepys.

The Admiralty had its own attitudes and ethos. I remember an amusing incident when a Welsh nationalist wrote to the Admiralty in Welsh. There was some discussion about how to deal with this and it was decided to send him a reply in Anglo-Saxon.

Many of the statisticians were engaged in manpower planning, which was always important. There was a punched card installation, which was very efficient. This work made a great deal of use of actuarial mathematics. However, there were also opportunities to have spells in the policy branches. Life was not dull at all.

Eventually there came an opportunity to apply some real mathematical statistics. Papers had started to appear on what was sometimes described as the problem of storage and sometimes as the problem of inventory control. The Admiralty had millions of pounds tied up in naval stores, which had to be large enough to guard against the risk of shortages. However, nobody had actually tried to assess the risks scientifically. The aim of the investigation was to apply the theory of probability to do this and to see whether the system could be improved.

It was impossible to predict the demand for stores with accuracy. In effect this was a stochastic process and we had to deal with probabilities. I spent a great deal of time examining the data and eventually found that for many items the probabilities could be approximated by a negative binomial distribution, which was a tabulated function. This was an important discovery in those days, before computers. It was then possible to construct a working model which could be used to study alternative strategies and look for simple practical rules which would improve on the previous methods.

I wrote this up and produced several reports for the Admiralty and a paper which was read to the Royal Statistical Society. At this point I was moved to another Department, but my successors carried on the work.

Robbins Committee

After the Admiralty I moved to the Central Statistical Office, where my main job was to assist the Robbins Committee on Higher Education. This Committee had its own statistician, who was Professor Claus Moser, now Lord Moser. He is a statistician of the first rank. My task was to act as a liaison and to assemble all the official statistics which might be useful to the Committee. There were also some special studies.

When the Committee started, about 7 per cent of the age group were going into higher education. The Committee were extremely thorough and considered all the evidence they could find about the pool of ability and the expected future demand for graduates. They recommended that the 7 per cent should be increased to 15 per cent of the age group. Several new universities were built as a result.

It all seems a very far cry from today, when the current proportion of the age group going into higher education is about 30 per cent and the target is 50 per cent, though this target is not based on any investigation like that made by Robbins. As this is a Christmas meeting, I hope that I may be allowed to tell a politically incorrect joke. This came from an MP, so I suppose that it is not really seen as incorrect by all politicians. The joke is that the target for higher education is 50 per cent of the age group, but the trouble is that 30 per cent of the age group do not know what "50 per cent" means.

Ministry of Labour

After the Central Statistical Office I moved to what was then the Ministry of Labour, where I was the Deputy Director of Statistics. The Ministry of Labour was concerned with such things as employment, unemployment, industrial disputes and wage claims. Official statistics were needed on all these subjects, which were often debated in Parliament. In those days the Ministry of Labour also produced the Retail Prices Index. In order to compile the RPI the prices had to be collected every month throughout the whole country, and for this we used the staff in the Ministry of Labour's own network of local offices, the Labour Exchanges. The Ministry also set up the Family Expenditure Survey, originally in order to provide the weights for the RPI.

This was collecting statistics on a very large scale. These were statistics in the original sense of the word, numerical data required by the state. There are plenty of problems about collecting statistics in bulk which are not mentioned in textbooks on statistics.

When I was there, there were some very experienced staff who knew all about the statistics which had been collected in the past, but whose knowledge would be lost when they retired. Luckily I had the opportunity to produce an official publication British Labour Statistics: Historical Abstract 1886-1968

Department of Employment

In 1968 the Ministry of Labour became the Department of Employment and Productivity (later shortened to Department of Employment), and I was appointed the Director of Statistics. There was economic planning, British style. This was the time of incomes policies and there was tremendous pressure to improve the statistics on earnings. To meet these, I set up a completely new type of survey. This covered a very large random sample of the whole working population, consisting of all people whose national insurance numbers ended in certain combinations of digits. Their employers then supplied the data on their earnings in a particular week. This survey had to be approved by the Prime Minister. At the time this was Harold Wilson, who used to be a statistician himself when he started life working for Beveridge.

I made some analyses of the earning statistics and eventually produced two papers for the Royal Statistical Society. One interesting result concerned the earnings of manual men. These followed a lognormal distribution. That is to say, the logarithm of their earnings has a normal distribution, which is now sometimes called the Bell curve (though earlier it was known as the Gaussian distribution, originally discovered by de Moivre). There had been earlier surveys of manual earnings at various dates right back to 1886 and these had been lognormal too. What was extremely unexpected, though, was that these lognormals all had almost exactly the same variance. This meant that although the level of earnings had changed enormously since 1886 in terms of £ per week, the percentage differences between higher paid and lower paid workers had hardly changed at all.

I remember that one day a whizz-kid from Downing Street came round to tell us that there was going to be a brand new incomes policy, which this time would change the distribution of earnings. I said: Well, it is a lognormal distrtibution. What shape do you want it to be?

The Department was also involved in the economic forecasts made by the Treasury, on which they based the budget and many other decisions. There were many econometric models, which were highly mathematical. Eventually I noticed an empirical law, which is that econometric models always fit the data perfectly right up to the date when they are published, but mysteriously cease to work so well when you try to use them later.

Life in the Ministry of Labour and the Department of Employment was always absolutely fascinating. We had a ringside seat to watch British politics in action. The Ministers included colourful characters like Michael Foot, Barbara Castle and Willie Whitelaw. They all took a personal interest in the statistics, particularly unemployment and prices.

The Duration of Play

Now, to make a change from statistics, I come to my three historical problems. The first was a problem which goes back to Huygens, known as the gambler's ruin. There is a more difficult extension, known as the duration of play. There are two gamblers, A and B, who start with different numbers of counters and play a succession of games in which their chances are not equal. The probability that A will run out of counters before B is known as the probability of the gambler's ruin. The extension is to find the probability that this will happen within n games.

In the historical library of the Royal Statistical Society I found a copy of de Moivre's book The Doctrine of Chances. I was so fascinated by this that I eventually bought my own copy. De Moivre found the first general solution of the problem in 1711, but there was a delay in publication and he was pipped at the post by de Montmort, who found a completely different solution which he published in 1713, together with a still more general solution found by Nicholas Bernoulli. De Moivre then solved the problem in two further ways, and in the process found results on recurring series, generating functions and the original derivation of the normal distribution.

I wrote a note on all this and also showed how the results foreshadowed several more recent problems, such as the diffusion process, absorption at a barrier in a random walk, the probability that a dam will overflow within a given time and that a linear sequential test will end within n trials. I submitted the paper to Biometrika and it was published as one of their series on "Studies in the history of probability and statistics". Later, it was republished in the book of the same name which was edited by Pearson and Kendall.

The Rule of Succession

My second historical problem was Laplace's "rule of succession". This started with an apparently simple problem about drawing black and white balls from an urn, as in a lottery. Laplace maintained that if you draw n balls from the urn, and find that m of them are white, then the probability that the next ball will be white is (m+1)/(n+2).

This was known as the "rule of succession" and the reason why it became important was that it was used in philosophical debates about the validity of scientific induction. It showed the probability that predictions based on past observations will be right.

Laplace's solution depends on what is now called a Bayesian argument. This requires us to choose a prior probability function, to represent our state of knowledge before the first ball was drawn from the urn. Personally I was brought up on the frequency definition of probability, and was taught that prior probabilities are flawed. Much later I read the book on probability by Jeffreys, which put the other point of view. There was a considerable controversy between frequentists and Bayesians and it struck me that it would be interesting to compare the numerical predictions which are given by the two methods, in the particular case of Laplace's problem of the balls in the urn.

The mathematics is quite complicated and I only have time to give the punch line. This is that in this particular problem the difference between the confidence limits for predictions made by the frequency method, and the corresponding limits for predictions made by the Bayesian method, is no larger than the effect of making one extra

observation. I wrote a paper on this and it was discussed at a meeting of the Royal Statistical Society.

Diophantus

I now come to the third of my historical problems. This is in the theory of numbers and it all started years ago when I read Heath's History of Greek Mathematics and got completely hooked by his chapter on Diophantus. I decided to buy Heath's edition of Diophantus and found to my surprise that the 1910 edition was still in print.

Then one day, many years later, I came across a reference to a problem which had been posed by a Mr T Baker in the Gentleman's Diary and Mathematical Repository in 1839. The problem was to find five integers such that the sum of any pair is a square. This was just the sort of problem which appealed to Diophantus and I thought that it would be fun to try to solve it by his methods.

This is a problem where any solution in rational numbers can be used to produce a solution in integers. It is therefore sufficient to find five rational numbers which sum in pairs to squares. The pairs can be chosen in ten ways, so we are looking for five rational numbers which satisfy ten conditions simultaneously. I knew that Diophantus himself had managed to find four rational numbers which satisfy eight conditions simultaneously, so I followed his methods.

I worked on this problem at odd moments and eventually found a solution, but it was in very large numbers, as follows:

177 346 641 122
41 897 374 103
6 469 631 522
4 864 525 922
1 675 430 978

If you add together any pair of these numbers, you obtain a square. However, I was sure that there must be better solutions than this, so I wrote to the Mathematical Gazette and offered a small prize. To my amazement, solutions poured in. I heard from schoolmasters, university lecturers, a retired Admiral and a student. One found a solution in which the highest of the five integers had 27 digits. One found a much neater method than mine, which needed only10 digits. However, a student at Cambridge beat us all. He used a university computer outside normal hours and he found the lowest possible solution, which needed only 6 digits:

763 442
177 458
148 583
28 658
7 442

I also learned that the problem had already been solved in France, again on a computer. One of the mathematicians concerned wrote a doctoral thesis on the problem and ended by finding six integers which sum in pairs to squares. This was amazing! So many people were interested that I wrote a paper on the problem and the solutions for the Mathematical Gazette.

A friend of mine, who resigned from the Civil Service to move to the Open University, said that he was glad that I had written a proper mathematical paper at last. He did not count statistics.

OPCS

After this historical interlude, back to the Civil Service! In 1978 I was appointed to a very complicated post with two hats. I became the Registrar General for England and Wales, legally responsible for the registration of births, deaths and marriages, and also legally responsible for conducting the censuses. I was also the Director of a Department with a very long name, the Office of Population Censuses and Surveys, known as OPCS for short. OPCS had the staff which actually carried out the censuses. It also carried out the Government's social surveys, which were used to assess how well policies were working and to provide information for the public.

The registration side was very important and interesting, but there is no time to describe it today.

The first Minister I met when I moved to OPCS was David Ennals, who was a very affable man. His first words were "I admire your social survey interviewers. If I asked women questions like that, I'd get my face slapped"

OPCS was the Government's great source for social statistics. We also produced many medical statistics, based on the causes of death which were recorded in the death registrations. The social surveys gave estimates of migration. We were able to produce demographic analyses of all kinds, and also the basic material for the official population forecasts. OPCS was a very busy Department.

Some countries have population policies, like encouraging births or curbing the growth of population. However, British Ministers here did not want to be involved in anything like that. I know that it sounds like "Yes, Prime Minister" but it is literally true that the policy on population was not to have a policy.

The Census

I now come to the Census of Population. It was my job to run the census in 1981, with the help of some very keen and experienced staff.

Planning the census involves not only planning the questions, but also planning the methodology, the publicity, the public relations, the fieldwork, the processing, the comentaries and the publications. Even more importantly, it involves finding the staff. Because of the census, the staff of OPCS varies over a ten year cycle. When I arrived we were building up to a census peak and there were 1,800 staff in OPCS itself. For the census day itself we also needed 100,000 local enumerators, to deliver and collect the forms. It was like recruiting a small private army.

The cost of the 1981 census was £50 million, which was £1 per head of the population. This had to be agreed with the Treasury and approved by Parliament.

On the questions, we had a real saga. I was determined to simplify the census form by reducing the number of questions. I originally proposed about 24 questions. There was also a strong lobby which wanted a further question on race or ethnic origin, but there were others who were very opposed to it. We held a census test at Haringey and there were actually public demonstrations against an ethnic question. I recommended that this should be dropped, but the political arguments continued. Also, the then Prime Minister was my namesake Mrs Thatcher [no relation]. She wanted to reduce the number of questions even more, to reduce the burden on the public. She went through all the draft questions herself. She asked me for a brief on who wanted each question, how it was used and what would happen if it was not asked. I had a very busy weekend writing this brief. We cut the questions to about 20 and then they were debated in Parliament. One of the questions was put back, and there was an overwhelming vote not to ask an ethnic question in 1981. As you might expect, the debate in Parliament did not take place until the very last minute, when there was just time to send the forms to the printers. We ended up with just 21 questions and I am convinced that the shorter and simpler form was one of the things which helped to make 1981 census go well. I am afraid that my successors did not follow this example and in the 2001 census there were 40 questions.

I think that there is just time for an absolutely true joke about the census. One of the enumerators was leaving a form at a house when the door was opened by a woman with a very long face. She said "It's no use leaving this form here". He asked "Why not?" and she said "Because I'm just going to commit suicide". The enumerator thought very quickly and said "Well, I'll call back on Tuesday in case you've changed your mind".

How many lived before us?

I now leave the census and come to demography. I will begin with a peculiar but quite interesting question, which came up when somebody said that there are more scientists alive today than have ever lived before. Then it wasn't long before somebody else said that there are more people alive today than have ever lived before. So the question was: How many people have ever lived before us? As part of my job I used to skim through the demographic literature and came across a paper which gave some improved estimates of the population of the world at various dates back to 40,000 BC. Earlier, there had been a book which gave estimates of the death rates at various dates back to prehistoric times. It struck me that if you multiplied the populations by the death rates, this would produce estimates of the absolute numbers of deaths at particular dates. With a bit of interpolation, you can then produce an estimate of the total number of deaths from prehistoric times right up to the present. I did the sums and the answer came to 60 thousand million as the number of people who have lived before us. This is ten times as large as the number living today. This calculation was published by the International Statistical Institute and later presented to the British Association.

Human longevity

I now come to a piece of real research. I retired from the Civil Service in 1986, but a few years later I had a completely unexpected invitation to join three people who were planning to do some international research on death rates at very high ages and on the span of human life. The three others were a population historian at Cambridge, a United Nations demographer who had retired to Portugal, and an American professor who was in Denmark but later moved to Germany. It was not at all easy for us to meet.

I must tell you briefly about death rates. After about age 30, death rates start to increase exponentially by about 10 per cent with each extra year of age. This is a law called the Law of Mortality, which was discovered in 1825 by a British actuary called Gompertz . Nowadays the death rate at age 30 is very low, about 0.001. The Gompertz increase brings this up to 0.0011 at age 31, and you will hardly notice an increase of 0.0001. However, exponential increases steadily mount up. They were known to continue until at

least age 90, but there was quite violent disagreement about what happens at still higher ages. The Chart illustrates the three main theories

Print the CHART somewhere near here, reduced in size as desired

The bold line in the centre is an exponential, and shows what happens if the Gompertz law simply continues. Death rates would soon become very high indeed and it would be extremely unlikely that anyone would reach the age of 120, but it would not be absolutely impossible. However, there were some very distinguished biologists who thought that it was impossible. They thought that every species has a natural span of life which it is biologically impossible to exceed. They put the maximum possible age for humans at 115. This meant that death rates must tend to infinity by age 115, so the asymptote would be a vertical. This is the top line on the graph. On the other hand, there were a few actuaries and demographers who thought that at very high ages the death rates follow a logistic curve, which is the bottom line on the graph. This can happen if the population is sufficiently heterogeneous and if it is the fittest who survive to very high ages. The logistic curve tends eventually to a horizontal asymptote. So the question when we started in 1990 was, is the asymptote vertical, or horizontal, or is there no asymptote at all.

The great problem was that the data at these very high ages were not reliable. Also, the numbers were so small that no country on its own could make proper statistical tests. What we needed was a large database containing reliable data. Our American member was able to obtain some money to fund a database. We assembled all the published official statistics about deaths at high ages from over 30 countries and devised some checks on their reliability. Only 13 countries passed the tests and these did not include the USA. For our analyses we pooled the 13 countries and this gave us data on 32 million deaths at ages 80 and over, including 120,000 centenarians, the largest database on old age mortality in the world.

The theory that there is an upper limit to life at age 115, the top line on the Chart, was soon exploded when Mme Jean Calment in France reached the age of 122. But this still left a large gap between the Gompertz curve and the logistic curve, and there were several other competing theories in between.

We worked on this together for nearly eight years and finally wrote a book, called “The force of mortality at ages 80 to 120”. This did not run to a paperback edition, but at least it is on the internet. I also wrote two papers, one for the Royal Statistical Society and the other for demographic journals. I will just give a few of the highlights:

a) The data support the logistic curve.

b) Even at enormous ages, there is still about an even chance of living for one more year, though of course the luck always runs out in the end.

c) There is no visible sign of a fixed upper limit to the length of human life. However, although it is not fixed, there is always a highest age which is actually reached in a given country in a given year.

d) This highest age has a probability distribution. The highest age can vary from year to year and in England and Wales it is currently ranging between 109 and 115.

e) The range for the highest observed age depends on the size of the population at risk and this is growing very rapidly.The number of centenarians in England and Wales was about 300 in 1951, is about 8,000 now and is expected to rise to 95,000 by 2066.

f) I have made the prediction that if the numbers indeed increase as much as this, then 80 years from now, in the 2080s, the highest age in England and Wales will be ranging between 116 and 123.

In case anyone wants to look up the details, the references to these and all the other papers I have mentioned, including the historical problems, are on the list below.

Now I wanted to end on a light note, but it's not so easy to find jokes about old age. Looking around and guessing your average age, I would expect that at least one of you, and possibly two or three, will reach the age of 100. To give you some idea what to expect, let me tell you an absolutely true story. A very old lady, a centenarian, was being interviewed on local radio. She was one of those centenarians who look as bright as a button and she was obviously very happy. The interviewer asked "What is it that makes a centenarian feel happy?" She said "The reason why I'm so happy today is that I've just managed to get my daughter into an old folks home".

I can't add to that, so now is the time to stop.

REFERENCES

PAPERS BY A.R.THATCHER

Historical problems

A note on the early history of the problem of the duration of play. Biometrika, 1957, 44, 515-518. Reprinted in Studies in the history of statistics and probability (eds E S Pearson and M G Kendall), 1970 (London: Griffin).

Relationships between Bayesian and confidence limits for predictions. J Roy Statist Soc (B), 1964, 26, 176-192. Discussion on pp 192-210.

Five integers which sum in pairs to squares. Math Gaz, 1978, 62, 25-29.

Statistics

Some results on inventory problems. J Roy Statist Soc (B), 1962, 24, 1-45 (with discussion)

The distribution of earnings of employees in Great Britain. J Roy Statist Soc (A), 1968, 131, 133-180 (with discussion)

Year-to-year variations in the earnings of individuals. J Roy Statist Soc (A), 1971, 134, 374-382.

The 1981 census of population in England and Wales. J Roy Statist Soc (A), 1984, 147, 222-232.

How many people have ever lived on earth? International Statistical Institute, 1983, Contributed papers, 44th Session, 841-843. Presented to British Association, 1984, Annual Meeting, Section Z8.

(With V Kannisto and J W Vaupel) The force of mortality at ages 80 to 120, 1998, (Odense University Press, Denmark). On the internet at www.demogr.mpg.de

The long-term pattern of adult mortality and the highest attained age. J R Statist Soc (A), 1999, 162, 5-43 (with discussion).

The demography of centenarians in England and Wales. Population Trends, 1999, 96, 5-12 (ONS: London). Reprinted (in French) in Population, 2001, 56, 159-180 and (in English) in Population: an English selection, 2001, 13, 139-156 (INED: Paris).