Reminiscences about the Diploma in Mathematical Statistics

P.M.E.Altham, Statistical Laboratory, University of Cambridge.

March 2010

The current academic year, 2009-2010, is the final year in which the MPhil in Statistical Science will be run. From October 2010, its formal structure will essentially be amalgamated with the new Cambridge Master of Advanced Study in Mathematics (MMath/MASt) course: this graduate course being the new name for the age-old 'Part III' of the Mathematical Tripos. The MPhil in Statistical Science course has been run since 1998-9, and this course itself took over from the Diploma in Mathematical Statistics course. The old Diploma course was started in 1947, and as Peter Whittle has written in his History of the Statistical Laboratory, several of the very first Diploma students (and in fact many subsequent such students) became very eminent in the then relatively new field of statistics.

For almost every year since I can remember, that is since 1964-5, a number of our Diploma/MPhil students received public funding to do the course, firstly from the Science Research Council. This became in due course the Science and Engineering Research Council, and lastly from the Engineering and Physical Science Research Council. Applying for and administering these student grants was an important, and sometimes stressful, part of my job. Regrettably such funding for our Masters course in statistics is currently unavailable. For PhD students, the EPSRC now funds intensive short courses in certain modern topics in statistics via the Academy for PhD Training in Statistics, held as week-long courses in various locations such as Bath, Bristol, Warwick and Cambridge.

Following these developments, and the welcome move by the University to give the Part III course the new title of Master of Advanced Study in Mathematics, it clearly makes sense both educationally and administratively to cease to run 'our' MPhil course as a separate entity. Having been involved with both the Diploma and the MPhil courses for very many years (since I myself was a Diploma student in 1964-5!) I cannot help feeling a little pang of regret for this inevitable change, even from my rather semi-detached position of retirement. I hope the reader will forgive me if I indulge in a few personal reflections.

In 1964 I was a naive 3rd year undergraduate reading Mathematics at Newnham College, in particular I was also almost entirely innocent of any knowledge of statistics. I got into the subject because an older friend, Andrew Evans of Trinity College, told me that he was doing this course for which one **could obtain funding** and which would then fit one for a career. Andrew is now Lloyds Register Professor of Transport Risk Management Studies at Imperial College London. My beloved late father had actually been quite revolutionary in his own way in supporting my studies for my first degree. In our extended family almost no-one, let alone any of the women, had been to university before. However, Dad said he was very sorry but he would not pay for me to do a fourth year at Cambridge: he was no doubt thinking of the future bills for the education of my younger brother and sister. I have to say that Dr Sheila Edmonds, my enlightened and caring Director of Studies at Newnham and an analyst herself, encouraged me to do the Diploma, and I am grateful for her guidance. I mention Sheila's positive attitude because at that time, as I learnt later, the study of statistics and probability in the university was regarded with some hostility and suspicion by some members of the mathematical establishment.

The only undergraduate exposure to probability/statistics available when I took my degree in 1964 was the inspirational 24-hour course 'Random Variables' given by David Kendall. Thus it was that I nervously approached David in February 1964 in his basement office at the Lensfield Road Chemistry Department, and in due course I became a Diploma student in October 1964. Figure 1 shows the June 1965 photograph of the 8 Diploma students, together with staff, visitors and research students. (Tragically, my fellow Diploma student Rollo Davidson was killed only a few years later in a mountaineering accident.)

Somehow David found time to do everything: he arranged the SRC awards and also the Applied Projects for the Diploma. My own Applied Project was in Astronomy, very kindly supervised by Peter Freeman at MRC Applied Psychology Unit in Chaucer Road. Peter gave his advice on Principal Components Analysis which we were implementing on astronomical data from the Hertszprung-Russell diagram. (Wikipedia will tell you what the H-R diagram is.) Statistical readers will know that Principal Components Analysis essentially computes the eigen values of the sample covariance matrix. Nowadays one could probably do the whole analysis in a morning, without leaving one's laptop. In 1965 this computation involved several trips by Peter (on his moped as I recall) to EDSAC, or possibly TITAN, at the Computer Laboratory on the Downing Site. I think Peter took the programs as punched paper tape, and since these were even pre-Fortran days, we must have had a program in Cambridge Autocode. This was a homegrown language which of course duly became obsolete, just as one had become expert in it. I certainly remember weeping tears of frustration while trying to punch paper-tapes: this activity required a degree of dexterity and accuracy I simply did not have. I'm sure I benefited a lot from Peter's help, both for the theory and the practice. Peter later went to University College London, and then became Professor at Leicester. My Internal Supervisor was David McLaren, who took endless trouble to see that one had no statistical misunderstandings; David was a pioneering Bayesian and left for Glasgow University in 1970.

John Kingman gave us an excellent course on Applied Probability, leaving us with the (perhaps incorrect) impression that we had understood everything he had said. As far as our introduction to Statistical Theory went, it seems that by the end of the first term we had struggled to understand Type I and Type II errors, and the Neyman-Pearson lemma, perhaps we also progressed to Uniformly Most Powerful Unbiased Estimators. As to textbooks, I remember very little that we found helpful. Lehmann's 'Testing Statistical Hypotheses' was recommended to us, and no doubt it is a classic of its kind: I found it **so** difficult, and almost completely unrelated to anything one might want to do in practical statistics.

In addition, we had our 'Analysis of Data' course, and said Analysis, on a Brunsviga (see Fig 2) or even on a new-fangled electric Facit machine, felt like very hard work, though we were mainly computing only sums and sums of squares. The Brunsviga picture is taken from

`http://www.cl.cam.ac.uk/Relics`

This toil culminated, as I recall, in fitting the linear model to an unbalanced 2-way design. Looking back, I can see why orthogonality in experimental design was so much emphasized. Using the Brunsviga or Facit efficiently was another skill I failed to acquire. Furthermore, confounding in factorial experiments left us all pretty much confounded, even though we analysed the classic Rothamsted NPK experiment (nitrogen, phosphate and potassium).

Back in the 1960's we had an annual intake of only 8-10 Diploma students, almost all of whom came from the UK. We moved from the Lensfield Road basement in 1965, as a 'temporary' measure, to the 16 Mill Lane site, as part of the newly formed Department of Pure Mathematics and Mathematical Statistics: not a name to trip off the tongue! This new home was the converted warehouse for the former Cambridge University Press building in Trumpington Street; it had immensely strong floors, and pillars in funny places, to support the great weight of the loads of paper. Although the
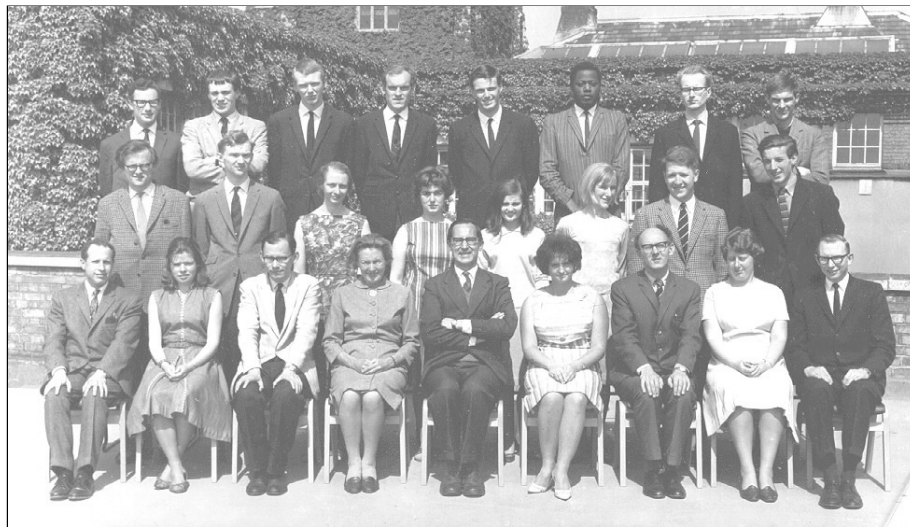
R.Coleman J.Haigh C.M.Goldie M.J.Green J.H.Pollard E.A.C.Thomas G.Gudmundsson D.Mannion

P.M.Lee R.Davidson J.M.O.Speakman M.Hirst C.R.Glencross P.M.E.Gray G.J.M.Shaw M.S.Lawrence

J.B.Kruskal A.C.Calder A.D.McLaren F.J.McWilliams D.G.Kendall M.L.Bagshaw A.M.Walker S.E.A.Prior P.Pyke

ABSENT: V.R.Cane J.G.Basterfiled D.J.Daley

Figure 1: The Statistical Laboratory photograph for June 1965

old Mill Lane Statslab was far from elegant or even warm, many of us have very happy memories of the Mill Lane days. Our offices were quite cramped, and the computer teaching room was so small as to resemble, at peak times, the Black Hole of Calcutta. But the proximity to Fitzbillies
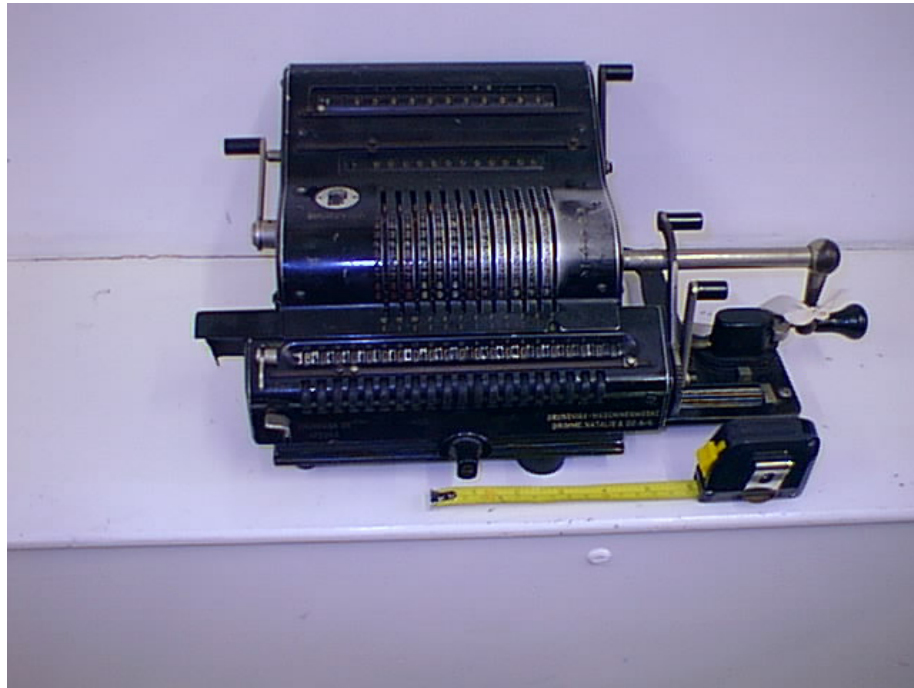
Figure 2: A Brunsviga calculator

meant that we had a very convenient source of excellent Chelsea buns. These were distributed to the long-suffering Diploma students to sustain them in their coffee break during the four-hour Practical examinations in June. Furthermore, if you wanted a refreshing walk by the river, or even

a trip in a punt, it was only 5 minutes' walk away.

But, what you may well ask, what did you actually **do**, for example in your Practical Classes? I suppose that in the 1970's we struggled with Fortran on the cumbersome University mainframe, eventually progressing to Genstat. On our own machines we used BBC Basic, then Glim, via a disk on the BBC computers, then S-Plus via Linux and our own little network, and finally – what a blessed relief – we could move onto the excellent free software R. In my own teaching with such software, I usually had the thoroughly deserved sensation of being 'a couple of beats behind the bar', relying on the speedy learning abilities of the excellent and keen young students whom happily I was paid to teach. Luckily for me, these students were for the most part quite relaxed to see the learning process as a joint effort between their lecturer and themselves. We also learned to be tolerant about logistic problems, such as finding something useful to do when the network crashed; perhaps to prove a theorem, with blackboard and chalk in the lovely reliable old-fashioned way? Of course, over time the standard of presentation of the students' written work, and also of their talks on their Applied Projects, got better and better. We became more-or-less 'web-aware' in 1996, although not actually wifi until the next decade when we were in our new buildings. Meanwhile, the students' own computing resources such as software and laptops improved in quality and bcame much cheaper.

As can be seen from our annual Statslab photos, over the years there have been two very welcome trends. The course numbers increased in size, and the student body itself has become much more international. In 2000, we moved again, now further from the centre of town, to the Centre for Mathematical Sciences at Wilberforce Road. Our 'Pavilion' is palatial, very stylish and warm, with a superb Computer Teaching Room in the basement. Recent MPhil student numbers have hovered around the 15-20 mark. Since almost all of the lecture courses and the corresponding examinations have been available to both Part III and MPhil students, the educational differences between the two 'sorts' of student are very small.

Though I personally do have a certain sadness about the demise of the MPhil course, I do believe that it is entirely rational and sensible that from October 2010, the MPhil in Statistical Science will cease to exist. My colleagues are working out how to retain the sense of 'esprit de corps' among those graduate students we can identify as 'Statslab' people. My colleagues also continue to provide a wide variety of courses in modern statistical science, and it is our privilege to teach very able mathematicians from all over the world.