THE TRAINING OF BRUNEL UNIVERSITY UNDERGRADUATES WHO INTEND TO BECOME STATISTICS TEACHERS

G.W. Goodall and F.R. Jolliffe
Brunel University. U.K.

This paper deals with some experiences with an undergraduate course in Mathematical Studies with Education which has in recent years been offered at Brunel University, U.K. We confine ourselves in particular to a statistics input to a final year specialist module titled "Mathematical Education". As a brief preliminary, we should explain that our Department of Mathematics and Statistics and Department of Computer Science co-operate closely in offering many undergraduate courses organised flexibly in a "unit" or "modular" pattern. The undergraduates reading Mathematical Studies with Education may be specialising in mathematics, in statistics, or in computer science; they take all of the usual topics in these areas, and in addition study specialist material in Education. They also undertake teaching practice in schools, and at the end of the course they are eligible both for their degree and for an award in Education, being recommended for recognition as qualified teachers by the U.K.'s Department of Education and Science.

Most of the material in Education is the responsibility of the University's Education Department, but colleagues (such as ourselves) from the Mathematics and Statistics Department and from Computer Science contribute to the Mathematical Education subject module; broadly, the intention is to re-examine elementary ideas from the more advanced viewpoint appropriate to final-year undergraduates, and to consider implications for the teaching of this material at school level. The students are. of course, not necessarily final-year specialist statisticians; it may (and usually does) happen that some will have done no more than the basic first-year statistics course. But this is no drawback, because most of the statistical work likely to be covered in a school is contained in the first-year course; and indeed confers a positive advantage, for our audience of students usually includes some whose knowledge of statistics does not extend much beyond that of a potential sixth former taking it, which has helped us in seeing problems with the perspective of the target group of school pupils.

Models for "elementary" probability

The study of statistics is founded on that of probability, and it is common to commence statistics courses by covering what we are pleased to call "elementary" probability. However, it is also common for it to be precisely this part of a course which causes most difficulty. Evidently, many people do not have any good intuitive grasp of probability (at least one of the authors readily confesses to inclusion in this category!). The concept of probability also often bemuses those who prefer to develop their ideas from a more formal mathematical standpoint. The difficulties seem to apply no matter what the age of the students. And while we might bemoan mathematical education in general, with too much emphasis on deterministic modelling, we still have to do as best we can.

Common practice is to try to get some intuitive understanding by using ideas such as tossing coins and rolling dice – situations which have the virtues of familiarity and simplicity. But, their simplicity clearly does not make such situations intuitively obvious – how many people, on hearing that a fair coin happens to have come down "heads' five times in a row, will aver that the chance of a "head' next time is something other than 1/2? And, further, the very simplicity of these examples can be a disadvantage, for they can appear trivial and inconsequential. Now, we know that probability theory is of extremely wide applicability; so one of us has used the course to explore how real, practical situations might be used as vehicles for the illustration of probability.

We have been able to make a lot of progress using ideas from quality control. The concept of a repetitively-made industrial product is very familiar, and can be used to demonstrate the existence of variability – even though the process is repetitive, and no matter how well controlled it is. the items produced are not all the same. So there will inevitably be some "defectives", and an inspection scheme of some kind is needed. We can here discuss methods of sampling; we can talk about consequences of "wrong decisions", firmly looking forward to what we would later call type I and II errors; we can use simple counting techniques to build up what we might later formally call the hypergeometric or binomial distributions; and we can actually construct Operating Characteristics, and use them to compare inspection schemes.

We have also been able to work similarly in genetics, with some of the simpler facets of classical Mendelian theory; this demonstrates probability ideas, and also leads naturally towards inference, comparing observed and predicted frequencies. We have even been able to discover that a test statistic of the usual χ^2 form could be appropriate, though of course we cannot "discover" what to actually do with this test statistic. And we have investigated other situations also, often suggested by the students themselves.

It is an important point that we suggest such examples can (or at least could) be used at a very early stage of a course, and not merely as illustrations later on. Inevitably therefore we are constrained in the techniques we may use. We are limited to cases that are "repetitive" (in the sense that coins and dice are:), so that we can naturally use the usual relative frequency approach to probability; and to cases that are easily seen as "independent trials" (coins and dice again:), for here we seem able to realise that the relative frequencies (and hence probabilities) are simply to be multiplied (whereas conditional probability results do seem beyond the scope of "discovery" by novices without formal instruction).

It does take longer to do the work in this way, if only because it takes longer to describe the real situations; and it requires more preparation. But our Education students have usually responded enthusiastically, often saying that their own understanding has been much enhanced. We would suggest that the advantages, of motivation. interest, general skills in problem formulation, and a look forward to other areas of statistics, are well worth any extra effort involved.

The use of random number tables

Discussion of random number tables and ways of using them has been very successful. This topic provides scope for extending students' understanding of basic concepts, gives them ideas for development of teaching material, and is not too demanding in the amount of background in statistics required. There is no lack of material which could be covered and a selection can easily be made if desired.

We start with a section on random digits which includes looking at printed tables and discussing how to read them, something on how random digits can be generated, consideration of the basic requirements of tables of random digits, and investigation of randomness of a set of digits. Most of this is new for the students, especially using a test in a real-life context, and they find it fairly interesting. Reading of tables of random digits gives students a feel for such as the extent to which odd and even digits alternate, and the lengths and frequencies of runs of digits. Students are encouraged to produce tables of digits and investigate their randomness. These could be tables produced for another purpose, or might be generated as random digits, perhaps by a mechanical method.

The use of random number tables to simulate the outcomes of random experiments is a natural second section. Use of tables to illustrate the stabilisation of relative frequencies is illuminating and can be compared with physical repetitions of experiments. This leads into the solution of probability problems by means of simulation, demonstrating that the method works by looking at very simple problems, and then considering more difficult questions and indicating the method's potential for complex problems where analytical solution may be too difficult. A computer can be used to advantage here. Students are interested in seeing worked examples and often comment that they feel this approach would have helped them at an earlier stage, but they are reluctant to produce their own worked examples. They are more enthusiastic about the use of simulation in a modelling context such as a queueing situation, a birth and death process, or a decision making situation.

It is useful to know how to obtain random observations from probability distributions, and techniques for distributions likely to be met at school level are covered, that is, the binomial, Poisson, exponential and normal distributions. The theory of the probability integral transform is done, and is useful for finding gaps in the students' knowledge. Tables of random numbers from the standard normal distribution, and from the exponential distribution with parameter 1 are examined, and simulation using Minitab is demonstrated. Theoretical distributions and empirical distributions are plotted for comparison purposes, and students are asked to choose a probability distribution and experiment with generating random observations from it. This is usually something they have not done before and they find it interesting.

Obtaining random observations from sampling distributions, with investigation of inference procedures as an important application is something which is worth doing to help future teachers' understanding. However, it takes several hours to do properly since some of the concepts are difficult and the computation involved tends to be heavy. There is usually time only for the revision of the idea of a sampling distribution and a look at confidence intervals for the unknown mean of a normal distribution of known variance.

Organisation in schools and relevance to the school curriculum are considered at every stage, as is also the extent to which practical exercises would be done by pupils, teachers, or a computer, and for which age groups these would be appropriate. Students are encouraged to state their own ideas on these matters. They often report on their teaching practice experiences, but for really worthwhile discussion they must do some of the exercises themselves.

Could we have done anything else ?

As we have indicated, lecturers on the Mathematical Education course have considerable freedom of choice as regards subject matter. We have described the two themes with which we have personal experience. There are many other possibilities which could be tried. One approach might be to base a presentation on the historical development of probability and statistics. This could well include a discussion of some of the famous paradoxes of probability. Many suggestions for teaching statistics through consideration of real data have been made, such as using examples from the media, analysing the results of a fairly complex experiment or social survey, and investigating patterns in data sets. Or the data might be obtained by pupils as results of experiments or projects. Any of these ideas would be suitable for discussion with future statistics teachers. More importantly, students and lecturers are all likely to find consideration of them useful and enjoyable. We hope that others will have the chance to teach on a similar course and develop these and other ideas.