"WHAT DOES THIS HAVE TO DO WITH US?" TEACHING STATISTICS TO ENGINEERS

Richard J. Wilson
University of Queensland
Australia

Teaching statistics to students aspiring to other professions can be both frustrating and rewarding. The frustration arises from (a) having limited time to cover everything from introductory to advanced material, (b) receiving little input from staff in the client profession, (c) the concepts of unpredictability and randomness being alien to students' thinking, particularly for engineering students, and (d) students not having the background knowledge or skills necessary to understand the methods fully. The reward comes from seeing students understand both basic and advanced concepts and methods, and from requests for assistance with later work as former students discover the relevance of statistics. This paper will address some methods used to overcome the frustration and to enhance the rewards in teaching a first course in statistics to engineering students. Although situations vary, these ideas will hopefully provide helpful tools.

INTRODUCTION

Engineering degree programmes typically consist of two types of courses: those that provide background information and tools (such as mathematics, statistics, physics and chemistry), and those, which provide engineering concepts and methods. Such programmes generally have too much material to cover and it is usually the service areas, which are allocated too little, time to cover all the topics required. For applied statistics courses, this typically means covering everything from basic probability and data analysis through to experimental design and multiple regression (and sometimes beyond) in maybe twenty lectures.

This setting creates an environment in which staff may be tempted to rush through material, skipping examples and demanding too much of the students. Students react adversely to such courses, querying the relevance of the course (due to the lack of examples in their area of interest) and complaining about the amount of work required in comparison to other courses. The end result is that students find engineering statistics courses like these difficult and do not learn the concepts and methods, and staff find the experience frustrating and try to avoid being allocated such courses whenever possible.

In this paper, some ideas will be presented which may help alleviate the frustration and assist in communicating the material. Firstly, a brief discussion of the characteristics of the client students (and staff) will be given. Naturally, generalisations will be made which will not be true for all students but describe a typical student. The nature of engineering statistics courses in general will then be discussed, with attention restricted to courses covering applied statistics. The main section covers some methods, which have been tried and seem to work well, although there is always room for improvement (particularly if more time was available). It is not intended to imply that there is only one way of teaching applied statistics to engineering students but rather to suggest some approaches which others may not have tried. Finally, some concluding remarks will be made, indicating some of the benefits, which may accrue.

ENGINEERING STUDENTS – A FIELD GUIDE

Engineering students often have their first introduction to probability and statistics in the second or third year of their programme. The first year of their programme usually covers basic science and mathematics, as well as introductions to the various fields of engineering. As such, the students have a grasp of deterministic systems and seem to believe that "error" in measurement is precisely that, an error, and not due to unpredictability in the measured variable. Consequently, dealing with uncertainty seems superfluous and random models incomprehensible. (This, unfortunately, may be true of some engineering staff who have rarely used statistical methodology or who use statistical methods in a deterministic way or as a black box, perhaps in reaction to the way they were taught.) So, the first obstacle to teaching statistics to engineering students can be a negative attitude to the material before the course commences.

By the time engineering students are introduced to statistical modelling, they are often fixed in their choice of field of engineering. As a consequence, they only wish to see material, which relates to that field and tend to be inattentive whenever examples from other fields are used. As courses in engineering statistics are typically heterogeneous (comprising students from different fields of engineering such as chemical, environmental, mechanical, civil, mining and minerals processing engineering), this can lead to students missing key sections of the work. As well, some areas of probability and statistics are more relevant for some fields than for others: for example, random processes are particularly relevant for electrical engineers but less so for chemical engineers whereas the reverse applies with experimental design. Hence, some sub-groups of a lecture group may be disinterested at times and so be disruptive. The issue of catering for students with diverse backgrounds is currently gaining more attention (see Shaw & Shaw (1997) and is difficult to handle.

Students who choose to study engineering often do so because they are "active"; that is, they enjoy "doing things" and are able to "do things". This also affects the way they learn new concepts and methods. Engineering students typically learn through action; that is, they learn *kinematically*. Hence, they learn best by doing something while being taught and often find it difficult to learn just by listening and reading. This has implications for how information and understanding is transferred from staff member to student.

Engineering students also tend to have reasonable visualisation skills. They require these in order to visualise what is happening with a process or design. This is also one of the key skills required of a statistician – much of statistics involves being able to recognise patterns in data. Although this skill may be underdeveloped by the students, staff should still take advantage of it and utilise graphs and so on in communicating key concepts and methods to the students.

As remarked earlier, the initial stages of an engineering programme usually include a reasonable amount of mathematics. As mathematics provides the key tools for statistics, again staff need to take advantage of what the students should know, especially with regard to those areas which are relevant to applied statistics (such as computational methods). Unfortunately, students' mathematical skills are also often underdeveloped due to the pressure of so much to learn in so little time.

ENGINEERING STATISTICS COURSES

Engineering statistics courses are typically "short" courses covering everything from introductory through to advanced topics. They are usually numerical in nature and require a reasonable amount of mathematical background to cover theoretical modelling aspects and visualisation skills to cover applied aspects.

The course uses a considerable amount of subject specific language ("jargon") and introduces alien concepts (randomness and ways of describing it). These are obstacles to the students. These obstacles have often been enlarged by the use of supposedly simple examples, including the traditional and unfortunate coins, cards, dice and balls in urns, which are not seen as relevant or interesting to the students. Alternative examples with engineering contexts (such as classifying items as conforming or non-conforming) are much better received by students. These examples also have the advantage of introducing more complex issues, such as independence, in a natural way.

Applied statistics courses involve a reasonable amount of computational work, usually achieved through a statistics package. As there are usually time limitations, there is not time to teach the use of the package in any detail. In addition, statistical work involves the interpretation of results and the presentation of both results and interpretation in a written report. Interpretation is often difficult (even for the expert statistician) and report writing, especially of a statistical analysis, is a task with which most engineering students struggle. Familiarity with word processing and an appropriate package is also essential.

The mechanisms through which information and skills are taught can include lectures, tutorials, laboratory sessions and texts, with tutorial exercises, quizzes, assignments, laboratory work (with reports) and, becoming more commonly used, group projects (see, for example, MacGillivray (1998)) as a means of providing problems for students to learn material and methods, and for providing guidance to aid individual students in their learning. How these are used will obviously affect how successful a course is in training students.

Finally, maybe one of the most disappointing aspects of teaching statistics to engineering students is the lack of opportunity and time to make links with other engineering courses. This is often compounded by the lack of input provided by those in engineering fields who may give ready assent to requests for the provision of contexts and data for examples but then fail to deliver. It is essential that this input is obtained as the communication of the statistical models and methods must be in the context of engineering and this requires a reasonable understanding of key engineering concepts and methods by the statistics staff member. Perhaps the problem is that it also requires a reasonable understanding of key statistical concepts by the engineering staff member.

TRIED AND PROVEN TECHNIQUES

One of the first steps to take in teaching a successful engineering statistics course may be to develop strong links with members of the client departments and professional engineers, especially through collaborative research projects and consulting. This achieves multiple goals. Firstly, such activity establishes an environment of collaboration in which it is easier to determine the expectations of staff in engineering departments with regard to which material is important and therefore needs to be covered. It also assists in defusing any problems that might arise. Such collaborative activity can help provide relevant modern contexts and case studies for use in the course. The use of such real problems seems to be an increasing trend as indicated by Garfield (1995). This is essential as students will learn the statistical concepts more readily in recognizable engineering problems – it assists them to "own" the data and problems. In addition, it helps overcome the initial negativity and bias many feel towards the intrusion of the statistics course in their engineering programme. If possible, engineering staff should be encouraged to develop links between their courses and the statistics course.

After agreeing on the syllabus, there is still a reasonable amount of flexibility with regard to the specifics and order of the topics. One useful approach is to commence with case studies for each field of engineering of students in the course. This helps students see the relevance of the course at the start and can generate interest in the following material. Often, it is difficult to cover enough case studies, so it is useful to include additional ones that may not be covered by the lectures but can be read by students.

From this point, key introductory concepts on data properties and collection can be covered. This would include data types, key aspects of a statistical study, the basics of experimental design and so on. It is critical to introduce difficult new concepts as early as possible to give students time to absorb them. The concepts of unpredictability and randomness are of particular importance. They should be introduced with different graphical means and with numerous data sets for each of the case studies already described, with emphasis on describing location, spread and shape. Advantage should be taken of their visualisation skills and, by using a large number of data sets, students start to grasp the idea that they do need models which take into account uncertainty and so model the shape of the data.

As it is essential that students have a firm grasp of the implications of uncertainty in the context of inference, introducing basic inference early is also necessary (before students are overloaded with material – about half way through semester). One method of doing so is to introduce the normal distribution as a reasonable model for symmetric bell-shaped data sets (emphasising that other models are required for data sets with different shapes), covering basic probability for the normal model and then commencing on basic inference (one and two sample confidence intervals, prediction intervals and tests of hypotheses). As well as covering new and difficult concepts early, this also helps to cover much of the specialised language at the start so that the students can become used to its usage.

It is then a short step to develop these concepts for multiple samples, quality control, experimental design and regression settings, including the introduction of residual analyses. Each of these areas can be covered in a similar way so that students start to see the general pattern of dealing with data in both simple and complex situations. They also start to understand the general principles of interpreting statistical results and how to transfer this interpretation back to the engineering context. Returning then to more general probability material (discrete and continuous distributions) and other possible distributions (binomial, geometric, Poisson, uniform, exponential, Weibull and others if time is available) is then straight forward following the format established when covering the normal distribution. This should include basic inference for these as well.

As the course is short (in time), it is essential to provide adequate notes and a relevant text. The notes can be provided in a variety of ways but it is critical that students need to do more than just read them or to listen to lectures. One method that can be used, either through lectures or the web, is to provide a shell, which includes the main descriptions of contexts and examples, graphs and other figures, and any other tedious material to write out. The shell can be provided as a document for downloading, sold/distributed through an appropriate agency, or made available for photocopying (through a library for example). A key feature of this shell is that it has gaps, which the students need to fill in. The gap material should be *key material*. By insisting that students write this material into the shell, it assists them to learn it – they learn kinematically! This approach can overcome the "too much material – too little time" problems. It helps to cover a large number of topics while still emphasising the key concepts and methods.

The choice of text can be difficult. No text will do precisely what you wish (unless you wrote it!) and there seem to be very few which cover adequately and succinctly the material for an applied statistics course for engineers. Traditional texts on probability and statistics for engineers tend to cover too much probability theory, have too many non-engineering contexts and include use of a statistics package as an "add-in" only. One possible text is Vardeman (1994), though many students find that it has more material than they can cope with as the course progresses. It is, however, an excellent reference for students after they finish the course as they can refer to it to fill in some of the details missed by the course due to time limitations. Whichever text is chosen, it is important to tie it into the course and provide page references for students to find the topics as they are covered in the course.

Finally, it is obviously important to use work sheets in such a way as to give students the opportunity to use the material being learnt. These work sheets must link into the lecture notes but can be used to introduce material as well. For instance, laboratory sessions can be used to introduce the statistics package (provided it is a simple to use package, such as MINITAB (registered trademark of Minitab Inc.)) and the structure of a statistical report. They can also be used to introduce alternative inferential methods to those considered in lectures.

It is essential that problems are relevant to students and, if possible, it is desirable to use different problems for the different subgroups within the class (though this is very time-consuming to set up). As engineering programmes are full to overflowing with material and hence susceptible to overwork for the students, it is appropriate to have students complete assessable work during contact time (tutorials and laboratory sessions). This can minimise unproductive collaboration among students, as well as managing their outside contact hours work. It also forces them to start thinking about the material early in the semester and consequently start becoming familiar with concepts and maintain reasonable pace with the course.

Work sheets for such an approach need to be streamlined (for example, they could consist of a handout sheet with questions and space for answers). Laboratory sheets in this format can be arranged to illustrate how to set out a statistical report with engineering problem description, statistical analysis and interpretation, and engineering conclusion. This may take considerable effort to initiate, but reduces later workload. As well, some students may react negatively if they feel they are being disadvantaged by a system in which they have to submit work for assessment at the same time as they are learning the material. However, by the end of the course, most students do acknowledge that this approach seems to work.

CONCLUSION

As a statistician, it is very important to see statistical models and methods being used appropriately by professionals in other areas. It seems that this does not always occur. One of the key ways of changing the *status quo* is to ensure that probability and statistics is taught well to engineering students (and others). Consequently, it is critical that statisticians are involved in teaching these courses, but they must do so using the engineering expertise and examples of their engineering colleagues. As with most statistical work, it is a collaborative effort.

From the staff member's perspective, teaching an engineering statistics course can be frustrating. Even when following the above advice, reactions can be adverse. It is difficult, at an introductory level under pressures of time and student diversity, to translate an engineering problem into a statistical problem, carry out the correct analysis, interpret the results and then translate the interpretation back into the engineering context. This is especially so if some steps have to be repeated. Students will not usually enjoy having to do this, but it is necessary (a) for their learning and (b) for them to realise their limitations and hence realise when to seek assistance or to do further personal research. With regard to the latter point, it is important to encourage students to seek statistical advice after they complete the course (although this may require payment by the client department in order to justify the time and effort required). This continues the teaching commenced in the course and it provides a reward for the staff member as their ex-students discover how to use their statistical skills better. It also has the side benefit of providing additional relevant examples for use in the course and sometimes leads to joint research projects (with attendant papers and grants). It, of course, continues the cycle initiated at the start of Section 4 above.

REFERENCES

Garfield, J. (1995). How students learn statistics. *International Statistical Review, 63,* 25-34.

MacGillivray, H.L. (1998). Developing and synthesizing statistical skills for real situations through student projects. *Proceedings 5th International Conference on Teaching Statistics.* International Statistical Institute, 1149-1155.

Shaw, C.T., & Shaw, V.F. (1997). Attitudes of first-year engineering students to mathematics – a case study. *International Journal of Mathematical Education in Science and Technology*, *28,* 289-301.

Vardeman, S.B. (1994). *Statistics for engineering problem solving.* Boston, USA: PWS Publishing Company.