The Statistician (1986) 35, pp. 441–447

Teaching engineering statistics with simulation: a classroom experience

JORGE LUIS ROMEU

Department of Industrial Engineering and Operations Research, Syracuse University and Department of Mathematics State University of New York College at Cortland, NY 13045, U.S.A.

Abstract. Motivating engineering students in statistics courses is often difficult. The present paper discusses our classroom experience teaching Regression and Data Analysis for Engineers—a senior level course. The organisation and support materials of this course are reviewed. The simulation experiments used and their engineering interpretations are described. Practical problems encountered and the adaptive solutions given to them are discussed. Finally the method of evaluating the students and some objective performance measures of the course results are presented.

Introduction

When teaching statistics to engineering students we often encounter two problems: lack of student motivation and lack of time to cover all the coursework material.

The problem of class motivation sometimes arises because many engineering students are essentially practical and result oriented. On the other hand, many statistics service courses are too theoretically oriented, with little or no practical exercises and, in addition, they often try to cover too much in a very short period of time. Finally, many engineering students do not see statistics as a primary area of application in their career and some even take as few courses as are necessary to meet minimal curriculum requirements.

The present paper addresses our classroom experience in teaching IOR529 entitled 'Regression and Data Analysis for Engineers,' which is a senior level non required (but recommended) statistics course for students in the Department of Industrial Engineering and Operations Research at Syracuse University. Pre-requisites include two basic statistics courses, dealing respectively with probability (IOR325) and statistics (IOR326). These two courses were given at the level of the Walpole-Myers (1978) statistics book. In addition, an introductory simulation course was strongly recommended due to the 'slant' of the lectures.

The syllabus was very ambitious (see Table 1). In a period of 13 weeks, meeting at a frequency of two 1½ hours sessions per week, we covered simple and multiple regression, variable selection procedures, analysis of variance and covariance, design of experiments and response surface methodology. The two textbooks were *Regression Analysis by Example*, by Chatterjee & Price (1978), and *Statistics for Experimenters*, by Box *et al.* (1978). In addition, the Regression Analysis book by Draper & Smith (1966), was strongly recommended as a reference text. We were given, however, complete freedom to prepare our lectures from either one of these textbooks and other sources (e.g. recent papers, reports and our own research experiences, e.g. Romeu, 1976, 1978).

The present paper will briefly overview the organisation and support material used in this course. We will describe the simulation experiments written to motivate our engineering students and to enable them to 'learn by doing'. We will discuss some of

Table 1. Course outline

1.	Statistical Model Building
2.	Linear Regression
3.	Analysis of Residuals
4.	Transformations
5.	One and Two Sample Problems
6.	Multiple Comparisons
7.	Multiple Regression
8.	Selection of Variables in Regression
9.	Analysis of Variance
10.	Factorial Designs

the problems encountered while teaching the course and the adaptive solutions we employed. Finally, we will describe the evaluation method and discuss some advantages and disadvantages observed during this 2-year teaching experience.

Course description

To raise the degree of student motivation, which became our first concern, five simulation experiments that we had developed for a systems analysis course were moved to the heart of this course. In this way, the industrial engineering student could work, simultaneously, in simulation and modelling of real systems and in statistical data analysis. This was the main reason for selecting Batch/FORTRAN/GPSS simulation procedures over more recent interactive ones (e.g. MINITAB).

These five 'experiments' presented feasible real-life situations with which an industrial engineer could be faced. They were based upon three different simulation programs—this added a touch of reality to the problems, since the numerical outputs of the simulations were totally unknown to all.

The first simulation program was written in FORTRAN. It modelled a simple queueing system, with one stream of customers, one server, exponential service times and First-in-First-Out (FIFO) queueing disciplines. This model could feasibly represent a machine shop problem (e.g. the operation of a lathe), the stream of customers representing the incoming jobs, the server, the lathe operator and the simulation output, the stream of finished parts. With this model we introduced both simple linear regression and paired comparisons.

The first experiment illustrated the use of simple linear regression models. The students were given the problem of measuring throughput as a function of service time. Since we were using exponential distributions for the generation of the arrival and service times, some regression assumptions were violated. These violations were used to illustrate the importance of checking model assumptions as well as for introducing the concepts of residual analysis and variable transformations.

Students re-analysed the same set of experimental data for different batch sizes. Various regression point and interval estimators, for the different batch sizes, were compared and contrasted and the problem of independence of the observations as a model violation, and their consequences in the statistical analysis, were discussed.

In the second experiment, the problem of paired comparisons and the concept of a priority queue were introduced. The students were asked to compare the performance of the two queueing disciplines, FIFO and priority rule, using the same (simulated) stream of customers and the paired t-test. They were to measure, as response variables, system throughput and mean time in the system by customer type. Again, the simulation output was re-analysed for the difference batch sizes (see Table 2 and Fig. 1) and their effect on the independence of the observations and on the statistical tests, contrasted and discussed.

The second program was written in a simulation language: GPSS (IBM). Again, the

Table 2. Batch size trade-off: descriptive statistics of pairwise T tests.

Responses	Batch sizes 200	100	50
No. of batches	5	10	20
Differences detected between the strategies	7·64	8·49	8·71
Standard deviation	2·62	1·37	0·45
T test	2·62	6·18	19·28
$T_{n-1,0.95}$	2·77	2·26	2·09

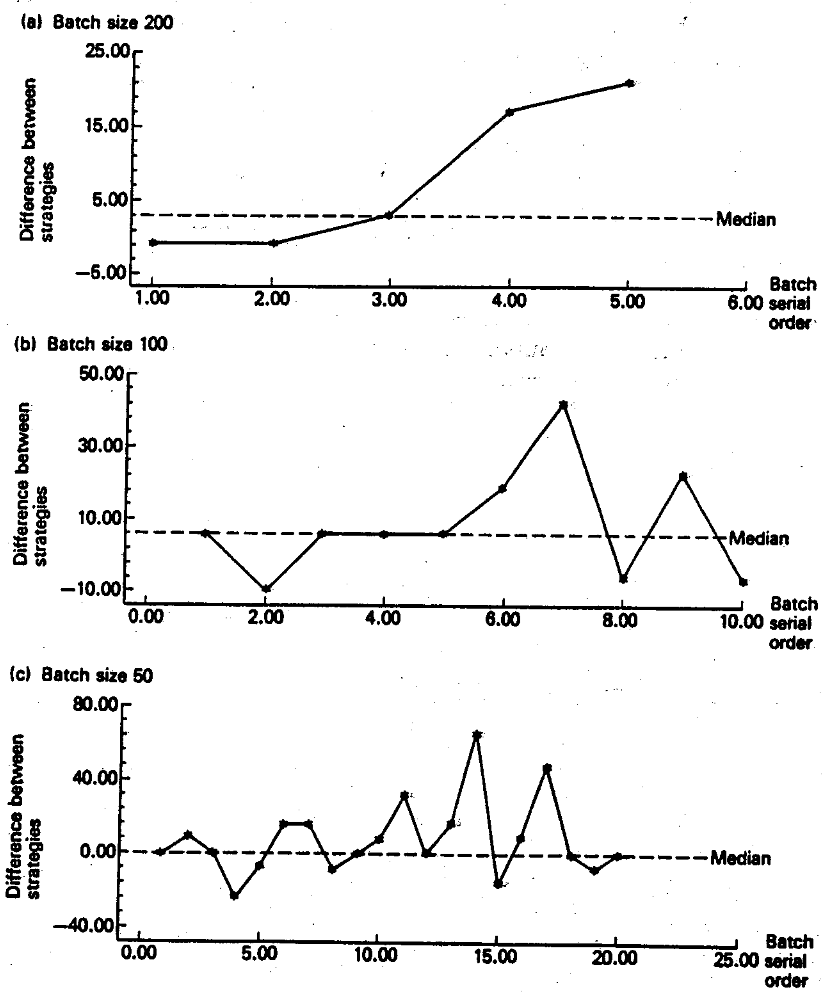

Fig. 1. Pairwise Comparison of the Simulated Service Time for Two Queueing Strategies: (a) using 5 batches of 200 customers; (b) using 10 batches of 100 customers; (c) using 20 batches of 50 customers.

reason for using GPSS and batch procedures was to relate the statistical analysis in question to the systems modelling background of the students. This simulation model consisted of two servers in series, with a buffer in between with limited capacity. The statistical distributions for input rates and service times remained Poisson and exponential. The feature of having two classes of customers in our system was retained.

Thus we had at our disposal the following independent variables:

—input rate (by customer class)
—service times (for each of the 2 servers)
—buffer capacity

In addition we had the following response variables:

—time in the system (by customer class)
—percent utilisation for each server
—throughput

The model allowed for the illustration of concepts and problems in multiple regression, variable selection procedures and hypothesis testing in multiple regression analysis. Several examples of multiple regression were developed. They dealt with forward selection, backwards elimination and stepwise regression, as well as the analysis of residuals and model assumption validation.

The fourth experiment was introduced using the simulation program described above. This time it also included two queueing disciplines (FIFO and priority) as a qualitative variable. This feature allowed for the discussion of the analysis of covariance model and the introduction of the General Linear Model. Students were required to check model assumptions through residual analysis using graphical procedures (e.g. plotting the residuals against the different variables) and performing tests of normality for their distributions.

The third and last simulation program was also written in GPSS. It consisted of several servers in parallel, having equal service rates, one stream of customers and Poisson and exponential distributions. It provided the model for developing the last experiment and for the final project.

The last (fifth) experiment illustrated the concepts of factorial designs and introduced the student to elementary response surface methodology. The final project summarised the most important skills learned throughout the course.

Operation

The course duration was 13–14 weeks, with two meetings per week. Considering the mid-term exam and the final project, this schedule provided the opportunity of assigning the students one 'homework' every weekend. These alternated between a (simple) textbook example and one of the five mentioned experiments. The final project was introduced during the last week of class and was due one week after the course concluded.

As mentioned earlier, our main goal was to motivate our engineering students. This was achieved by giving practical interpretations to the experiments. The first two experiments were illustrated through the example of a lathe operator. The third experiment exemplified the operation of an assembly plant, the first server representing the complete assembly process with the second server representing the quality control stage. In the fourth experiment, the assembly plant was complicated with two different production lines such as would be found in the assembly of cars and trucks. Buffer space would represent warehouse capacity. Finally, the fifth example represented the operation of the checkout counter of a supermarket. It could also model the operation of tellers in a bank, of telephone calls in a switch board, etc.

At the beginning of each 'homework' the experiment was explained, so the student would understand the statistical nature of the problem, and the engineering applications of the methods were emphasised. Each student was then provided with the simulation program and a *different* random seed, which would generate an individualised experimental result. In this way it was possible to allow the students to interact amongst themselves, without fear of plagiarism, since not only the numerical results of each would be different, but occasionally even the experimental outcomes. An example of this occurred during the comparison of the two queueing disciplines. Due to the sampling process several students obtained experimental results leading to the acceptance of opposite hypotheses and this allowed the possibility of cooperating and exchanging information, this providing the student group with an additional incentive in their work.

Students' experimental results were graded on analysis approach, assumption validation and result interpretations, rather than on numerical results. Complete freedom was given to the students within the domain of an experiment. For example, when dealing with the problem of variable selection in multiple regression, each student was free to choose between stepwise regression, backward elimination, forward selection or all possible regressions and in many cases, and completely on their own, students tried more than one method, and compared their results.

In addition, not all the students were asked to analyse the same response variable. It was then possible to assign more than one student per simulation run. For example, in the case of the study of queueing discipline one set of students analyse the servers' utilisation whilst another group does the same for mean time in the system, by customer class. By cooperating amongst themselves and exchanging information the students enriched their learning experience. Also, students were able to look at additional facets of the problem which they would not have had time to look at individually.

The final project was a 2 week assignment. Students were given two management policies to evaluate and for this task they were required to use a simulation model (usually the third one described). They had to perform an experiment, which included applying the General Linear Model, validate the assumptions of the statistical model, perform hypothesis tests and finally, write a management report containing their conclusions. The management report should also discuss the technical aspects involved in the experiment.

The course evaluation method weighted equally its four components: homeworks, experiments, mid-term exam and final project. With this grading system it was extremely difficult to fail or excel by pure chance. The majority of the students worked enthusiastically during the course, achieving very satisfying results. During these 2 years, only one student failed the course, less than 5% of those who undertook it.

Problems of applying such a course layout

The course work described assumes the availability of a mainframe computer and some computer budget to run the simulation programs. However, if a mainframe is not available it is also possible to run simulation programs and use statistical software available for the less expensive personal computer, for example the IBM-PC.

A greater difficulty arises with the availability of simulation packages like GPSS. These are expensive and not so easy to find, nor implement on a PC, as the statistical packages. In addition, there is the problem of the accuracy of the random number generators in 8 and 16-bit machines. It is still possible to write a simple random number generator routine which, inaccurate as it may be, will provide a good substitute in these classroom examples. Also, a discussion of the origin of the statistical reasons (autocorrelation) will provide interesting material for the class to analyse.

Conclusions

The primary objective of the present approach to teaching a statistics service course is to provide a more exciting classroom experience for the industrial engineering student. Some special features (use of GPSS batch simulation routines to perform systems modelling and analysis) are context specific and can easily be modified, if required, to attract the interest of a different engineering student population. Also, the present course layout is labour intensive for the instructor. Our experience indicates it is more appropriate for small classes (between 10 and 15) of intermediate or advanced students than for larger groups of beginners. This was, in fact, the framework in which the course was always taught.

Some objective performance measures can be presented to assess the students' response to this course:

At the beginning of the course, all students were set up with a computer account for running the experiments. The best students invariably ran out of computer money and were provided with extra funds to continue their work. This may give the instructor yet another variable to measure the grade student work: CPU time.

In many cases, and always on their own initiative, the students ran the experiments using more than one statistical procedure and compared the results. In some cases, they worked with their individually assigned response variables as well as with the variables assigned to other students (e.g. they analysed percentage utilisations and also the mean times in the system). We conjecture that this extra work was also motivated by the course layout.

Some students were not satisfied with the statistical software provided with the APL system and requested (and invariably obtained) access to more sophisticated packages (e.g. SAS). We conjecture that this extra interest was also motivated by the course layout.

Finally, our students took this course during the last semester of their senior (last) year, just before leaving with a Bachelor's degree. It was very rewarding to see them depart from the classrooms saying they would like to develop their professional interests in an area where their recently acquired statistical skills could be applied. Some, in fact, obtained such positions. For example, the two top students in the first class were employed by the local power company and a multinational computer manufacturer, respectively.

Acknowledgements

The author gladly acknowledges the Department of Industrial Engineering and Operations Research of Syracuse University for the opportunity and freedom to teach this course and a reviewer for his careful editing and helpful suggestions to improve the presentation of this paper.

References

BOX, G., HUNTER, W. & HUNTER, J. (1978) *Statistics for Experimenters* (New York, John Wiley).
CHATTERJEE, S. & PRICE, B. (1978) *Regression Analysis by Example* (New York, John Wiley).
DRAPER, N. & SMITH, H. (1966) *Applied Regression Analysis* (New York, John Wiley).

IBM CORPORATION. *General Purpose Simulation System/360*. Introductory User's Manual. IBM Applications Program GH20-0304-4.

ROMEU, J.L. (1976) Alqunas Aplicaciones de la Estadistica a la Ingenieria Civil, *Revista de Ingenieria Civil*, No. 2 (La Habana, Cuba).

ROMEU, J.L. (1978) Un Ensayo para el Analisis y la Clasificacion de la Fuentes de Informacion Mediante Metodos de la Estadistica Matematica, *Revista de Ingenieria Civil*, No. 4 (La Habana, Cuba).

WALPOLE, R. & MYERS, R. (1978) *Probability and Statistics for Engineers and Scientists* (London, McMillan).

Appendix: course schedule

Week 1:	Review of statistical Concept Model Building with data
Week 2:	Simple Linear Regression A developed example
Week 3:	Analysis of Residuals Data Transformations
Week 4:	Developed Examples/Experiment #1 Comparison of two populations (T test)
Week 5:	Power of the Test Hypothesis testing
Week 6:	Experiment #2 Mid Term Exam
Week 7:	Analysis of Variance (one way) Residuals/Multiple comparisons
Week 8:	Multiple Regression Regression contd
Week 9:	Experiment #3 Selection of Variables
Week 10:	Dummy Variables (ANCOVA) Experiment #4
Week 11:	Autocorrelation Multicolliniarity
Week 12:	Factorial Designs Experiment #5
Week 13/14:	Confounding/Fractional Factorials Response Surface Methodology Final Project