How we Train Students for Statistical Consulting

Wang Jinglong - Shanghai, China

As we all know, statisics is a science applicable almost everywhere. Wherever there are data, there are statistics. In order to make students understand and know how to apply statistical techniques well, we pay special attention to training students for statistical consulting. In our department we put this training into practice by means of the following three aspects of activities.

(i) *Statistical practice:* The complete undergraduate programme in our department runs for four years. The senior student has a course "Statistical Practice". In this course, we often invite technicians or research workers to talk to our students on applications of statistical techniques in their work. Besides this, a group consisting of five or six students is sent to some factory to try to solve some problem raised by the factory. Though our department was established in the fall of 1984, we had students majoring in statistics in the year before and our first batch of graduates came out in July 1987. That means, by now, we have had this course for four terms. Up to the present, fourteen problems which were raised by eleven factories have been studied by students. In the study of these problems, regression analysis, experimental design, discriminant analysis, optimisation, quality control, forecasting and some other statistical techniques have been used.

All the problems are interesting. The most interesting one was raised by Qi Bao Winery located in the suburb of Shanghai. Through fermentation and some other processes with rice, plain liquor is produced, but in this raw state, the liquor is classified into several kinds, according to the degree of its alcoholic content and the order in which it is distilled. By blending several kinds of these raw products, the drinkable liquor is obtained. In Qi Bao Winery some mellow drinkable liquor is obtained through blending two kinds of the raw product. The traditional blending proportion is 30:70 on the basis of past experience. The quality of liquor depends on quantities of some trace elements in the liquor, and on the blending proportion. Is the blending proportion 30:70 the best? The winery has never analysed this problem scientifically.

In the winery, the output of two kinds of the raw product is in the proportion of 70:30, just in the opposite direction of the blending proportion 30:70, and much of the first distillation is left unused. The winery suffers losses from this left-over, hence they raised the question how to blend two kinds of raw product in order to compound a liquor of good quality yet retain a blending proportion close to the proportion of output.

Students gave a mathematical model for this problem:

$$\min \; |X_1/X_2 - P| ;$$
$$a_i \leq Y_i(X_1, X_2) \leq b_i, \quad i = 1, ..., q,$$

where X_1, X_2 are the quantities of two kinds of raw product when they are blended; P is the output ratio of the two kinds of distillates, here P = 70/30; $Y_i(X_1, X_2)$ is the quantity of the i-th trace element in 100ml drinkable liquor when the quantities of two kinds of raw product X_1 and X_2 are properly blended, i = 1,...,q, here q = 9. In accordance with analysis of famous liquors and material accumulated through practice for a long time, optimal control range $a_i \leq Y_i \leq b_i$ is found with a_i and b_i known, i = 1,...,q.

Using regression analysis, regression equations $Y_i = Y_i(X_1, X_2)$ are obtained, i = 1,...,q; then using the optimal calculating method, a better blending proportion 38:68:61:32 is obtained. To blend the two kinds of raw product according to this proportion, the quality of drinkable liquor is improved, production of drinkable liquor is increased, and a million yuan of income pours in. All these beneficial results make the winery very prosperous.

Through statistical practice, students learn the process of solving problems using statistical techniques. It is very useful in helping them find good positions after they have graduated from the university.

(ii)   *Bachelor paper - a term paper in the senior year:* Each student is required to write a paper for graduation with the degree of Bachelor of Science in Statistics. We pay great attention to the choice of the topics for the bachelor papers. We hope students choose the topic by themselves through statistical consulting. Among bachelor papers of the first three years' graduates, there are eleven papers requested by chemical factories on optimisation techniques; two papers by the Shanghai Spring Factory on reliability of springs; three papers by machine factories on design of production; five papers by factories, companies, banks, etc. on the prediction of production of poultry and clocks, the number of tourists, income from tourist trade, sales volume of souvenirs, output value and cost of production, and amount of deposits; two papers by the Departments of Foreign Language and Physics in our university on educational statistics; ten papers on survey sampling by the Shanghai Statistics Bureau, the educational administration of our university, the Shanghai Population Office, and the Shanghai Student Union.

In the following we describe two problems of survey sampling which were requested by the Shanghai Statistics Bureau and educational administration in our university. In 1987, we carried out a survey on behalf of the Shanghai Statistics Bureau. In this survey, the population consisted of all the staff members doing statistics in the factories, companies, farms, government branches etc. in Shanghai. The Bureau wanted to know what calculating machines (abacus, calculator and computer) the staff members used and wanted to use, and which of the 46 subjects of statistics the staff members had studied, used and wanted to study. All this information was necessary for the Bureau to work out a professional training plan for the staff. Four students did this survey and

wrote a joint bachelor paper on it. They designed questionnaires and decided to use stratified random sampling. They regarded each department where the staff worked as a stratum, and a list of all statistical staff members in the stratum as a frame. They determined the sample size in each stratum according to the required accuracy of estimation. From each frame they drew a sample, and mailed questionnaires to members of the sample. Response rate was 82%. Finally, a report of the survey was written. The report provided useful information for the Bureau.

In 1984 the credit-hour system was adopted by our university. In 1988, we carried out a survey on behalf of the educational administration of our university. The population consisted of all the students who were enrolled in our university in September 1984 and would graduate in July 1988. The survey was about the implementation of the credit-hour system. The administration wanted to know what the students thought of the credit-hour system and what improvements could be made. Four students carried out this survey and wrote a joint bachelor paper on it. They decided to use stratified, two-stage sampling without replacement with unequal probabilities. They divided the 18 departments of our university into five strata, then selected departments from each stratum, using unequal probabilities sampling without replacement. Finally, a simple random sample of students was selected from each sampled department. Data were collected in this survey through personal interviews. Classmates of these four statistical surveyors were invited to act as interviewers. In their report they put forward five proposals about improvements to the credit-hour system.

(iii)   *Extracurriculum groups:* There are two voluntary extracurriculum groups, the Accurate Data Analysis Group (ADAG) and the Survey Sampling Group (SSG), in our department. They launch statistical consultation in our university. ADAG performs analysis of data which are collected by clients. In September 1988 the Department of Psychology performed psychological evaluations of about 1100 freshmen students in our university. There were 399 items in this evaluation. At the request of the Department of Psychology, ADAG performed an analysis of these nearly 45,000 items. ADAG gave the Department of Psychology not only a report of the analysis, but also efficient computer software to deal with this kind of data. The Measuring Centre of our university hoped ADAG could scan the yield rate of radioactive elements of different wavelengths at different times, in order to easily discover their characteristics. ADAG studied methods for drawing the picture in Euclidean three-dimensional space on the PC, and worked out a program for drawing the picture using oblique axonometry.

The SSG was established for junior students doing the course Survey Sampling. The students who join the SSG practise survey sampling as they study the course. In 1988 and 1989, each SSG student completed one task of survey sampling. One sampling survey was on income and expenditure of students in our university; another was on the number of freshmen and junior students undertaking night study. The two tasks were requested by the general affairs administration and educational administration of our university. They all used stratified random sampling. The President of our university gave considerable weight to the results of these two surveys. Now a third survey is underway, requested by SSG itself. The students who joined the SSG want to know what the opinions are of students of our university about study, love, recreation and sports, and jobs. Our university is very interested in this survey. It uses stratified, two-stage simple random sampling.

The students who joined the SSG practise techniques ranging from questionnaire design to data collection, from data analysis to completion of a report in survey sampling. They have learned facts that they are not taught by the teacher in class, but are useful in practice. Students do survey sampling with much enthusiasm.

The *Shanghai Multivariate System Optimisation Corporation* aims to popularise and apply pattern recognition and computer optimisation using multivariate statistical methods and other efficient methods. Recently, the Corporation requested that our department join them in launching free consultation to all interested organisations. Now we are doing the preparation for this free consultation. This consultation will begin this year. We are looking forward to this, and hoping that this consultation will train our students for a wider range of statistical consulting.