

SURVEY SAMPLING: LEARNING BY DOING. A TWENTY YEARS GRADUATE LEVEL TEACHING EXPERIENCE

Jean-Hugues Chauchat
Université Lumière
France

For the past twenty years, we have been using an original technique to teach statistics and survey sampling methods to postgraduates studying economics and statistics. The students must put their knowledge into practice by carrying out a survey sample for a client who they will have found by themselves. This may include a marketing study for a shop, a brand or a public service, or measuring the audience ratings of a radio station or local television station. More than 100 different surveys have already been carried out by students on this program over the last 20 years. Furthermore, every six years, during the regional parliamentary elections, the entire group (25 students) carries out an estimate of the results for the public local television station, on the basis of the first ballot papers counted in a sample of 300 polling stations; our results are broadcast live on television 30 minutes after the close of polling.

CARRYING OUT A SURVEY SAMPLE FOR A CLIENT

For the past twenty years, we have been using an original technique to teach statistics and survey sampling methods to postgraduates studying economics and statistics at University Lumière, Lyon, France. The students must put their knowledge into practice by carrying out a survey sample for a client who they will have found themselves. This may include a marketing study for a shop, a brand or a public service, or measuring the audience ratings of a radio station or local television station. More than 100 different surveys have already been carried out by students on this program over the last 20 years. These applied studies enable the students to get hands-on experience in the world of work.

BENEFITS FOR STUDENTS

Actually, post-graduate students learn how to

- understand the client's needs,
- propose a tailored study,
- draw up and try out the questionnaire; they can compare the results of open or closed questions, the effect of item list change, etc.;
- organise the sample design, under practical constraints;
- collect information; as they develop the questionnaire themselves, and work themselves as interviewers, they have a better understanding of the relative meaning of the responses they will be statistically working on;
- write up, check, correct and class the results; they have to decide what to do with non-responses, how to code responses to open-questions, etc.;
- calculate the estimates and confidence intervals, according to the (complex) sampling design;
- visualise and translate the results into non-specialist's language for the client.

ESTIMATING REGIONAL ELECTION RESULTS FOR A LIVE TELEVISION BROADCAST

Furthermore, every six years, during the regional parliamentary elections, the entire group (25 students) carries out an estimate of the results for the public local television station, on the basis of the first ballot papers counted in a sample of 300 polling stations; our results are broadcast live on television thirty minutes after the close of polling. The "Rhône-Alpes" region boasts six million inhabitants. The ballot is carried out according to the list system, in eight constituencies; there are roughly 5,000 polling stations in all. This process lasts an entire semester and requires the students to put all their statistical and data-processing knowledge into action, and involves:

- Multidimensional Data Analysis; students have to collect data of the last election results in all polling stations, and analysis them in order to form homogeneous strata;
- Complex survey design; they have to carry out random selection of stations in these strata, with probabilities proportional to the number of registered voters, then to prepare the formulas of calculation for confidence intervals for the results of each list in each constituency (in 1998 there was a total of 72 lists of candidates)
- Visualisation; they have to conceive tables and graphs of the results to be broadcasting live on television; those tables and graphs have to be precise and quickly readable and understandable text by the general public;
- Programming; post-graduate students have to write programs to manage all the data and to calculate estimated confidence intervals for the results of each list in each constituency (in 1998 there was nine lists of candidates on average, for each constituencies, that is to say a total of 72 list results to estimate); forty PC are connected on an intranet are used to record, calculate and draw up tables and graphs of the estimates;
- People management; roughly 300 first and second year students are called upon to observe the counting process in the sample polling stations, and phone in the results; 20 students receive the calls, and capture and check the data;

The estimates are updated several times in the course of the evening, as the counting in our sample of polling stations progresses...

COMPARING A LOT OF CONFIDENCE INTERVALS TO THE TRUE VALUES

In terms of teaching statistics, this technique has a very rare advantage: a real complex survey can be carried out and the 72 estimated results can be quickly compared to the “true population values”, as the official results are published between six to eight hours after the close of polling. The confidence level was more than 99% for each confidence interval, using plus or minus three estimated standard deviations; despite of that, no miscalculation is allowed! The live broadcasting constraint creates a hard stress for the students (and for the professors, too ...), so, they have to implement a “zero-default” policy on calculus, programs, human resources management, etc.

BENEFITS FOR STUDENTS, THE STATISTICS PROGRAM AND THE UNIVERSITY

The benefits are important for every one:

- Post-graduate students are motivated to do a good job; they become self-confident; and, looking for a job, this experience have a great value;
- Under-graduates student are motivated to apply to the “economics and statistics” master program;
- The university participates in the civic instruction of young citizens by associating them to a political election;
- The technical success of the whole statistical operation is a good advertising campaign for the University to politicians and to the whole population.

HOW DID WE BEGIN ?

The self confidence to do this came step to step: twenty years ago, five student where looking for a “client” for the survey study they had to do for my course; they were proposing a marketing study to a small local radio station; at this time, a partial local election campaign was opening and the radio station asked them : “Are you able to make for us an election estimation during the election day night ? So that our station will be the first to give the results and all the candidates will come here to get the results and to debate”. The students asked their professor : “Are we able to do that ? May you help us ?”. We asked them to construct the survey design in order to estimate the final result, on the basis of the first ballot papers counted in a sample of ten out of twenty polling stations. The survey and estimation design was realistic, so that we went on. Only twenty student and two 1980’s PC were involved. The estimation was quite good. We did the same during four successive *local* elections, and we were able to contract with the public

regional television station for the 1982 and 1988 regional *parliamentary* elections. The amount of the last contract, in 1998, was about 165,000 French Francs (about 23,000 US\$).