TEACHING STATISTICS TO LAW STUDENTS

Michael O. Finkelstein, Columbia Law School, USA

There is a growing need in the legal community for statistical education. More and better materials would be welcome. Education should begin at least at the law school level, with examples drawn from the law. The methods of the author's law school class, concluding with a mock trial in which statisticians serve as witnesses, are described. The teaching moral is: tie statistical learning to real law issues. And to persuade laymen: keep it simple.

Statistics has become an increasingly important part of the legal universe. In litigation, statistical proof appears in a wide variety of cases. Statistics are also used routinely by legal scholars in appraisals of the legal system. There is even a journal called Jurimetrics published by the American Bar Association that is devoted (but not exclusively) to the subject.

A few years ago the United States Supreme Court in the *Daubert* case held that federal district judges must evaluate proffered scientific testimony and rule out mere personal opinion parading as science. Since most judges don't know any science, this provoked some judicial anxiety. The Federal Judicial Center responded by setting up courses for judges in scientific evidence--most of it involving statistics--and producing a Reference Manual on Scientific Evidence, which prominently features statistics. The Center gave away the Reference Manual to all federal judges and then commercial publishers sold an astonishing 70,000 copies to lawyers. A new edition will appear shortly.

I have read the Reference Manual and moderated several of the teaching sessions for judges. I think both are useful, but certainly there is room for improvement. Statistics teachers should take note. However, one has to remember that the professional lives of district judges are too crowded and demanding to allow much time for an entirely new discipline--particularly one that doesn't make a daily appearance in their courtroom. So it is of law students and their education that I wish to speak.

I have been teaching statistics to law students, primarily at Columbia Law School in New York City, for over 30 years. The course is for only one term because the law school curriculum is crowded and no more can be justified. Within that short period one must teach not only basic ideas, but also their applications in actual cases. This is a quantum leap in difficulty for both teacher and student. Many applications in litigation involve fairly complex facts or advanced techniques that are beyond the scope of beginning courses. There is everything from simple probability calculations to survival analysis to logistic regression. And even law students who have had some statistics are for the most part far from the subject.

With the passage of time the class has evolved with the education of the teacher. Nowadays we actually *do* statistics as well as talk *about* it. This is much harder and the enrollment has suffered. One year when I taught the class at Yale Law School, a delegation of students asked me not to give a final examination. Why? Because, as the students told me with charming frankness, in my class there were right and wrong answers and they might get it wrong!

Most lawyers do not relish making calculations, but in my opinion some grappling with actual data is necessary to keep them from being a statistical *luftmensch*. On the other hand, the class is sold to law students not to make them statisticians, but to help them work more effectively with statistical experts. For that role, the ability to ask probing questions and understand the answers may be quite sufficient without the ability to calculate results. And so, while students are required to make some computations to keep them down to earth, the focus is on the basic ideas that they must learn to articulate.

To approach the subject in this way, the class proceeds simultaneously on two levels. Each week students are selected to present problems and give answers while others are selected as critics. The rest of the students read the materials and comment. In effect, students become active workers in statistics only periodically and are commentators much of the time. In the presentations, I insist on a lawyer-like performance: they must state the facts of the case, describe the statistical model they are using, verbalize the answer they have computed, and explain its relevance to the case. They or their commentators must criticize the model. Articulation, which is the lawyers' stock in trade, is a critical part of the process. Particularly early in the course, it is not unusual for a student triumphantly to produce a number and then be unable to describe exactly what he or she has calculated. (There is in fact a noticeable negative correlation between mathematical and verbal ability: the most verbal students tend to be weak calculators while the better mathematicians frequently are less verbal.)

The course and the problems are from *Statistics for Lawyers*, a textbook written by Professor Bruce Levin of Columbia's School of Public Health and me. The book consists of short sections of statistical exposition followed by a number of problems with questions to focus discussion. The problems include snippets of data for simple calculations. The students are supposed to read the sections of exposition and then apply what they have learned to the problems. The problems generally are from actual cases, including in the descriptions at least some of the complicating factors of real life, so that the readers can appreciate both the difficulties and payoff of statistical method in a realistic setting.

In the present edition of the book the answers are given in the back. For the new edition, now in preparation, we have debated whether to continue the present format, or adopt some other strategy, such as giving answers for every other problem, or a separate handbook for instructors. I do not mention to students that answers are in the back of the book, and this unconcealed fact is not generally discovered for some time. So there is a period when students work without answers and seem to do not too badly. However, last year the publisher suddenly ran out of stock and we had to photocopy the book (with permission) for the students. I left out the answers.

As a result, the class fell into semi-chaos. A remarkable correlation developed between being assigned a problem and being absent the next week. At one point I couldn’t resist asking who was going to be absent so I could assign them a problem. Although the experience is not definitive, I think the answers must stay. We are also planning an instructor’s pamphlet, since most law professors don’t have statistical expertise and would be more likely to use the book if it came with a security blanket.

The problems are of various types. One important type calls upon the student to select a statistical model, make calculations, and then critique the model. An example is the problem entitled “Disputed election,” which appears after the exposition of the hypergeometric distribution. Under New York law, a defeated candidate in a primary election is entitled to a rerun if improperly cast votes (it being unknown for whom they were cast) “are sufficiently large in number to establish the probability that the result would be changed by a shift in, or invalidation of, the questioned votes.” In the particular case given, the challenger, who won, had a plurality of 17 votes; among the 2837 votes cast there were 101 invalid votes. What is the probability that the election would be changed if the challenged votes were removed? The subsidiary questions we pose are designed to lead the student through the problem. We first ask what split of the improperly cast votes would change the results if they were removed from the tally. Next we tell the student to use the normal approximation for the cumulative hypergeometric distribution to compute the probability of such a tally. Third, we ask whether the assumptions used in computing the probabilities are reasonable. Finally, we ask whether it matters that the winner is the incumbent or the challenger.

Students have stumbled over all of these questions, but the root problem is conceptualizing invalid votes as random selections from an urn. The difficulty, I think, is that the statistical model, which takes the number of alleged improper votes as given and their distribution among voters as random, reverses the intuitive causality that begins with the voters and thinks of them as generating mistakes. Mistakes wandering around picking voters is not how one intuitively thinks about the process. Once that hurdle is overcome, the students are usually fairly good at using the normal approximation for the cumulative hypergeometric distribution and very good at identifying reasons why the random assumption might not be valid. My hope is that they will do better than the lawyer in one actual disputed election case, who was so flummoxed by the statistical testimony that he could only argue that there were no such thing as a law of probability. Why? Because law was certain (the apple falling from the tree) and probability uncertain. The judge, you will be relieved to know, corrected him.

A similar conceptual difficulty arises in connection with a problem entitled “Epidemic of cardiac arrests,” which appears after the section on the Poisson distribution. An unusual increase in the number of evening cardiopulmonary arrests and deaths in a pediatric intensive care unit led to the prosecution of a nurse. The nurse worked 201 out of 454 evening shifts and was among those on duty at the time of 27 of the 34 deaths. This was the most extensive presence at death times of any of the nurses. In the criminal trial the statistical evidence was introduced over objection. The students are asked, among other things, to compute the probability of a clustering this large if time of death were independent of nurse on evening shift. To compute this probability, the students have to remember the conditioning property of the Poisson distribution and view the deaths as 34 binomial trials with regard to the presence or absence of the suspected nurse. This is the reverse of the usual way of thinking about the situation, in which the presence or absence of the nurse is a given and the occurrence of death is the accidental or random event.

Another type of problem involves a critique of some expert’s testimony. Since statistical tools are used by experts in other fields, mistakes abound. And some statistical concepts are just plain hard to keep in mind, at first. Power is an example. We have a section on power and then a problem entitled “Death penalty for rape.” A black man was convicted of raping a white woman in Hot Springs, Arkansas, and sentenced to death. Arguing against that penalty, a noted sociologist collected statistics on race and penalty in Arkansas rape cases. Testifying for the defendant, he presented the results in a series of two-by-two tables showing that the race of defendant was highly associated with the death penalty in rape cases. He concluded that other factors which might account for the correlation--such as prior record--could be ignored because the correlations with race were not statistically significant. Of course, what the expert forgot was the issue of power. We ask the students simply whether the data support the expert’s conclusion. The problem is quite hard because a full answer requires selection of a plausible alternative hypothesis (the logical candidate is the actual data) and computation of power under that hypothesis. But I think most students stumble over this because they don’t have firmly in mind what power is. In the next edition of the book we are going to take the student through the problem with questions leading step by step, as in the Disputed Election problem.

The grand daddy of conceptual difficulties arises from the inconvenience that evidence in the law is defined as a fact which makes the occurrence of a relevant fact more or less likely than it would be without the evidence. This is Bayesian in formulation. But since most statistical evidence is not Bayesian, the statistics don’t tell lawyers what they really want to know. The temptation is strong to misinterpret likelihoods as Bayesian probabilities. Thus in judicial opinions (and even in that most hallowed ground, briefs submitted to the Supreme Court) it is said of a statistically significant result that there is less than a 5% chance that the null hypothesis is true. We pound away at that mistake in various contexts, but I am not sure how deep a dent we make. In most contexts the mistake does no harm, because the same conclusion is reached with either formulation.

The highlight of the course is the mock trial at the end. It is designed to put to work what the students have learned and are still learning. It works this way. The class is divided into groups, each representing a party to a lawsuit, and one group being judges. Each group is assigned a professor of statistics as an expert, with whom the group works. We have a data base and a set of facts defining the case. For the last few years we have used a version of the facts of an actual case involving charges of a pattern and practice of discrimination with respect to pay and promotion against women doctors at a medical school. The student groups meet with their expert, design an approach to the data, write a brief report disclosing their calculations, and disclose it to the other side. The trial before a six-person jury on the question whether plaintiffs have proved by a preponderance of the evidence (the data) that there was discrimination. The method of analysis commonly used is multiple regression in which the dependent variable is salary and the explanatory variables (in addition to sex) are education, experience, publications, and medical specialty. The coefficient for sex is taken as evidence of discrimination. This type of analysis has become standard in actual cases. While they are preparing the case, the class reads the chapter of the book on multiple regression.

The students divide up the trial work: there are opening statements; direct examinations of the expert witnesses; cross-examinations of the witnesses, and closing statements. The judge charges the jury without commenting on the evidence (not permitted in American courts). The jury then retires to deliberate and returns with a verdict. Unlike a real trial, the jurors then tell us why they reached their conclusions. After that there is a party.

Everyone enjoys this exercise. The professor experts are generous in taking time to work with my students and frequently ask for repeat performances; they like to see their learning in action. The jurors are interested, sometimes even passionate, in their views. And most important, the students find the work highly engaging and rewarding. Presenting a technical case is a challenge to their intellect and imagination.

The reaction of the jurors is fascinating. Even though Professor Levin and I cooked up data that had considerable evidence of discrimination, and even though the jurors all have at least college degrees, in not one instance--out of about 10--has a jury come out for the plaintiffs. Testimony about statistical significance is only weakly understood, and the jurors seem mulishly unwilling to find wrongdoing solely on the basis of statistical evidence. They don't believe the multiple regression model. The closest plaintiffs came to a victory was when a jury of three men and three women deadlocked along gender lines. When they asked for food from the party (it becoming late and the party awaiting), we said verdict first. But they stuck to their guns and in the end, to save the pizza, I had to declare a mistrial. Statistical evidence wrapped up in sophisticated manipulation just doesn't seem to persuade.

The teaching moral is this: tie statistical learning to real law issues in as realistic setting as possible. And to persuade laymen: keep it simple.

REFERENCES

Finkelstein, M. and Levin, B. (1990). *Statistics for Lawyers.* Springer-Verlag