From: Donald Berry <db@stat.Duke.EDU>
Subject:Teaching elementary Bayesian statistics

I'm reluctant to send anything out over an already crowded Internet. However, I regard the subject of this note to be sufficiently important to risk the wrath of those who may be upset having this message cross their screens.

The subject is teaching statistics from a Bayesian perspective in so-called service courses. A set of papers in The American Statistician a few years addressed this matter. My present note is an extension of my contribution to that exchange. In particular, I am relaying part of my side of a discussion with some faculty in substantive fields at Duke who were reluctant to accept the appropriateness of Bayesian courses for students in their disciplines. They want their students to be familiar with the usual types of statistical presentations in their fields. The story is ongoing, but subsequent discussions lead me to be cautiously optimistic about its end.

I hope my message helps those who find themselves in similar circumstances. Another objective of this message is to seek your views, including your own related experiences. To cut down on traffic, I ask that you respond to me personally <db@stat.duke.edu>. If I receive messages aimed for the broader audience I will post them (including names unless you specify otherwise) in a single mailing.

My regards to all,
Don Berry

E-mail message from Berry to selected faculty:

First, a bit about my statistical focus. I live in a world that requires frequentist statistics. For example, I design and analyze clinical trials in breast cancer for a national cooperative group called the Cancer and Leukemia Group B. As such I deal with regulators and medical journals, both of which are primarily focused on frequentist methods. I use frequentist methods--I must or I wouldn't survive. (I also use some Bayesian methods; these are becoming increasingly used and accepted, but that is another story.) No one knows better than I that students need to be able to use and to understand frequentist methods. In teaching [Stats course], I carry out my responsibility in this regard most diligently. I am about teaching concepts and methods that will serve the students best and in a way they will understand best. I perceive that you and I agree in every way concerning what the students need to know. The only disagreement we may have is how best to get there.

I motivate statistical concepts using a Bayesian perspective, but somewhat paradoxically this enables the students to understand and interpret frequentist measures better than do students who take standard frequentist courses. I exploit the common ground of Bayesian and frequentist methods. Bayesian posterior distributions give rise to frequentist confidence intervals, frequentist P-values, etc. when assuming a flat (or uniform) prior. The students find a 95% probability interval in this case and learn that it is a 95% confidence interval. In most of the examples I use and in most of exercises the students carry out, I ask for confidence intervals and P-values, and I do so using that language. For example, this is a quote from the [Stats course] final exam I gave last year: "Give the 95% confidence interval for d."

When casual application of Bayesian methods and frequentist methods lead to different answers, I come down on the frequentist side. An example is multiple comparisons and other multiplicities such as subgroup analyses. Bayesians come up with answers similar to the frequentist by using hierarchical models, but these models are too advanced for [Stats course]. The routine use (but not hierarchical) of Bayes theorem in analyzing subgroups leads to paying too much heed to the observations in the subgroup at hand. Consistent with the frequentist view, the observations should be shrunk or regressed to those in other subgroups; I take this view and show the students how to do it.

There are three reasons for taking the approach to teaching elementary statistics that I do. One is that the prior distribution in the Bayesian approach allows for bringing in information that happens to not be exchangeable with the evidence in the experiment at hand. This is done by assuming a prior distribution that embodies this other information (and so it is not flat). Bayesian analysis involves synthesis of information from different sources; Bayesian analysis is meta-analysis. So, roughly speaking, flat = frequentist and non-flat allows for addressing larger questions. The Bayesian perspective encompasses the frequentist perspective and goes beyond it.

My second reason is the availability of predictive probabilities in the Bayesian approach. It is especially handy when designing an experiment or for deciding the course of an ongoing experiment to be able to find the probability of future observations, and to do so without having to condition on values of unknown parameters.

My third reason is a frustration with the lack of understanding of statistical concepts on the part of substantive scientists who have taken (and presumably passed) statistics courses. I can attest mostly to this matter as regards MDs and other medical workers, essentially none of whom can give correct definitions and interpretations of confidence intervals and P values. But I have had some exposure to [particular type of scientists] in whom I see the same phenomenon. [Example deleted.] Students in my courses can give proper interpretations. For example, in a typical setting they understand that there is a 95% probability that the treatment benefit is between the 95% confidence limits assuming that other available evidence gives equal support to the various possible treatment benefits.

I teach simple linear regression, and I do this assuming flat priors. This means the answers are the same as in frequentist courses. I also teach a version of analysis of variance, but this is the one setting in which I do not give the full frequentist treatment. The reason is that a frequentist development--even one from a Bayesian perspective--is more advanced than I think the students can handle.

I don't want to get side tracked on the advantages of the Bayesian perspective, but I do want to make a point in this regard. My understanding of the way [particular scientists] do research is that they use statistics in a way that is more flexible than is consistent with the standard frequentist view. For example, since frequentist analyses are tied to the experimental design, any deviation from the design means that--strictly speaking--frequentist methods cannot be used. Of course, no reasonable scientist takes this seriously, and this statement includes scientifically-minded frequentist statisticians. Since the Bayesian approach is flexible in this regard, it is consistent with what scientists do. The Bayesian view allows for drawing inferences from any experiment that has been conducted honestly. This is especially important in many epidemiological studies. Most statisticians have a jaundiced view of retrospective data collections, and I am part of the majority. But the flexibility of the Bayesian approach makes it easier to draw conclusions from these studies. Also, I am aware that statistical hypothesis testing has drawn fire in recent years from some people in the [particular] community. The Bayesian perspective would go a long way to resolving the controversy that has developed.

To reiterate and summarize my message: I do not fail to cover frequentist methodology. Indeed, because these frequentist methods are what university

students will need in their professional lives, frequentist methods
represent a climax in my course.