BAYES FOR BEGINNERS? SOME REASONS TO HESITATE

by

David S. Moore
Purdue University

Technical Report #96-33

Department of Statistics
Purdue University
West Lafayette, IN

July 1996

o

Bayes for Beginners? Some Reasons to Hesitate

David S. Moore*

Introduction

The issue I wish to consider is whether ideas and methods of Bayesian inference should be presented in a first statistics course for general students who must later read and sometimes employ statistics in their own disciplines.

This is a quite specific question. I agree that Bayesian methods are increasingly important and should form part of the training of professional statisticians. I also agree that a course for liberal arts students on e.g. "risks and decisions" structured around subjective Bayesian ideas can be stimulating. These are quite different settings from that I have in mind, one more specialized and the other less constrained by student needs. I do not wish to join the foundational debate over whether Bayesian inference is in some sense uniformly preferable to standard inference. That prejudges our question, forcing us to teach what is "right" regardless of customer needs or pedagogical barriers. Most statisticians remain eclectic, willing to employ Bayesian methods where appropriate but unconvinced by universalist claims. That being the case, I am unwilling to settle the content-of-instruction issue by appeals to one or another side in a debate among professionals.

I will offer, as my title says, some down-to-earth reasons to hesitate to present Bayesian ideas in a first course in working statistics.

- Bayesian methods are relatively rarely used in practice. Teaching them has an opportunity cost, depriving students of instruction about methods that are in common use.

- It is unclear what Bayesian methods we should teach. Those who advocate them have not yet agreed on standard approaches to standard problems. And of course, lacking standard methods, we also lack standard software for implementing them.

- A conceptual grasp of Bayesian methods rests on some understanding of conditional probability, a notoriously difficult idea. Although Bayesian conclusions are simple in form, the simplicity disappears when we ask "What do you mean by probability?"

*It is a pleasure to dedicate this paper to Shanti Gupta, who offered me my first position and contributed in many ways to my statistical education.

There is of course no easy path to understanding inference. Nonetheless, the fact that standard inference consistently asks "What would happen if we did this many times?" and answers by displaying a sampling distribution makes standard reasoning more accessible.

- Inference is only part of statistics in practice. Data analysis and the design of data production were always important and have become more so in the past generation. For various reasons (emphasis on inference, ambiguity about the role of designed data production, dependence on conditional probability), Bayes-for-beginners tends to impede the trend toward greater emphasis on developing students' data sense. I like the trend and don't want to impede it.

It seems to me that the weakest conclusion possible is that it is *premature* to make Bayesian methods the focus of basic methodological courses. They haven't sufficiently evolved, been sufficiently widely accepted by users, or even been sufficiently routinized by experts. That situation may change in the future. In the present, academic researchers ought not to impose a still-primitive version of how we think things should in principle be done. Advocates of Bayes for beginners should ponder the similar theoretical cases for basing our first courses on decision theory, or nonparametric methods, or robust inference. To adopt any of these is simply to invite our customers to go elsewhere for beginning instruction.

Are Bayesian Methods Used in Practice?

Our beginning students come to us from other fields of study. They come because their own fields employ statistical ideas and methods. Their first need is to be able to read literature in their own field. We ought to be attentive to our customers' expressed needs, rather than offer them what we imagine they ought to need. What statistics do knowledgeable practitioners in various fields apply? Let us search for data. There follows, in roughly descending order of statistical sophistication, a survey of recent empirical studies on the use of statistics in practice.

Professional statisticians working on applications in the physical sciences may be thought to employ up-to-date and effective methodology. Rustagi and Wright (1995) carried out a census of statisticians employed in Department of Energy National Laboratories. Remarkably, they obtained responses from all 103 members of this population, 100 of whom hold a master's or doctorate degree. Table 1 records the responses to a request to choose from a long list "the three statistical techniques that have been most important to your work/research." Only four of this sophisticated group mentioned Bayesian methods, although 37 reported training in these methods during their university careers. When asked to choose the three most important techniques that were *not* part of their academic training, 19 named Bayesian methods—sixth place behind quality control, reliability, simulation, exploratory data analysis, and graphical display.

Medical research is a major and quite sophisticated consumer of statistical analyses. The surveys by Altman (1991) and Emerson and Colditz (1992) document the nature and growth of the use of statistics in medical journals in roughly the decade of the 1980s. In particular, Emerson and Colditz inventory the methods employed in the 115 "Original Articles" appearing in the *New England Journal of Medicine* in 1989, and Altman does the same for the 100 articles appearing in 1990. Table 2 presents some of the findings of Emerson and Colditz. Techniques and ideas from a standard first course predominate—Emerson and Colditz say (referring to a longer time period) that acquaintance with descriptive statistics, t procedures, and contingency tables would give "full access" to 73% of the articles. They identify increasing use of ANOVA, multiple regression, and survival analysis as notable recent trends. Altman points to meta-analysis, new techniques for design of clinical trials, and editorials in several medical journals encouraging more frequent use of confidence intervals. Neither paper contains any mention of Bayesian approaches. Some use could lie hidden in the "other methods" category of their tables, though Emerson and Colditz enumerate several of the "other methods" without using the word Bayes.

In *psychiatry*, Everitt (1987) reprints a table from DeGroot and Mezzich (1985) that surveys 597 papers in the 1980 volumes of three major journals. Of these, 156 were surveys or contained no data. The remaining 441 use statistical methods to some degree. *One* paper among these 441 employed Bayesian methods. Everitt mentions more recent trends toward use of clustering, logistic regression, structural equation models, and Cox regression. Dunn et al. (1993), focusing on depression, point to some of the same innovations along with meta-analysis and an emphasis on controlled clinical trials. Neither they nor any of the discussants to Everitt's survey mention Bayesian approaches as either in use or promising.

Emulating Emerson and Colditz, Hammer and Buffington (1994) survey all articles published in 1992 in six *veterinary medicine* journals. About half contained statistics beyond simple numerical descriptions. The authors summarize: "Knowledge of 5 categories of statistical methods (ANOVA, t-tests, contingency tables, nonparametric tests, and simple linear regression) permitted access to 90% of the veterinary literature surveyed. These data may be useful when modifying the veterinary curriculum to reflect current statistical usage." Multiple regression, epidemiological methods, confidence intervals, and survival analysis fill out the authors' list. Bayesian methods are again unmentioned.

As we move away from the areas that most often involve professional statisticians in their work, the statistical methods employed of course become more traditional. Moreover, the newer techniques that are considered promising vary with the area of application. It would certainly be desirable to have comparable data for other fields, especially the social and behavioral sciences. Nonetheless, the available data suggest that *Bayesian methods are rarely used in any field to which statistics is applied.* Paul Velleman points out that the absence of Bayesian procedures in commercial statistical software is strong evidence of lack of use, as these packages respond quickly to customer demand. Velleman says, "Both the features and the advertising of software packages offer a good measure of what people who actually analyze data really want." The most recent new version announcement I have seen (as of July, 1995) is Minitab Release 11. Minitab claims to have added logistic regression, reliability/survival analysis, polynomial regression, gage R&R, and correspondence analysis. Several of these techniques appear in the lists I have cited. Minitab appears to find more demand for even gage R&R and correspondence analysis than for Bayesian procedures.

The fact that even sophisticated users rarely turn to Bayesian methods deserves attention. Statistical research journals are full of papers advancing and applying Bayesian ideas. Statisticians—especially Bayesians—therefore imagine that use of these ideas in practice is advancing rapidly. I can find no empirical evidence that this is true. The research papers are in the nature of demonstration pieces that suggest the possibilities of Bayesian analyses. These analyses have not yet passed the "Box test" that assesses the usefulness of a method by whether it is actually used.

This survey of empirical evidence for use in practice raises another issue beyond the question of why we should teach beginners an approach that appeals to us in principle but lies unused in practice. *Teaching Bayes to beginners has an opportunity cost.* If Bayesian ideas displace t procedures, contingency tables, regression, or ANOVA, they bar students from access to much literature in any field that applies statistics. If we manage to add Bayes to the list, perhaps by persuading students to elect further courses, we must ask whether the time might better be spent on logistic regression, simulation, survival analysis, or meta-analysis. Bayesian methods are not the only hot field in statistics research, and several others have already passed the Box test.

Are There Standard Bayesian Methods?

The content of a first course that aims to provide understanding and useful tools to students from other disciplines should, I think, consist mainly of standard material well-accepted by the profession and widely used in practice. The first section argued that no Bayesian methods are widely used. This section will suggest that even if we listen only to Bayesians, standard methods for standard problems are not yet agreed upon.

Much of the disagreement concerns the essential element distinguishing Bayesian from standard models, namely prior distributions for unknown parameters. There is a continuing tension in Bayesian circles between use of priors that try to reflect the actual partial knowledge of a decision-maker and default priors that are automatically generated from the sampling distribution, taking no account of what partial knowledge may exist. The former class of priors are "informative;" the latter are generally chosen to be "noninformative." The use of conjugate priors, which specify a parametric family of prior distributions on the basis of analytic convenience, but allow prior knowledge to choose the parameters of this family, lies between these extremes. Because knowledge of the mechanism that generates the unknown parameter is rarely complete enough to determine a prior distribution, informative priors are usually subjective. Simplifying a bit, we can imagine several Bayes-for-beginners approaches.

A. Purist Bayes Relevant prior information is always available and should be expressed in an informative (usually subjective) prior distribution tailored to the problem at hand. This view is consistent, easy to explain, and in many settings intellectually attractive. Expositions of the advantages of Bayesian analysis often emphasize the use of genuine prior information and the subjective interpretation of probability. Alas, there can then be no standard Bayesian analysis for standard problems, because every problem is potentially unique. We are left with a tool of great power in non-standard problems, but which is unlikely to ever be widely used in standard settings. Beginners come away with ideas but few usable tools.

B. Accessible Bayes Emphasize the Bayesian Big Idea: express prior information in an informative prior distribution, use data to update this information to form a posterior distribution, base all inference on the posterior distribution. Restrict the settings considered to those in which beginners can implement the Big Idea, mainly discrete or conjugate priors. In a binomial problem, for example, the prior information just happens to be expressed by a beta distribution. Emphasize estimation rather than testing, and rejoice that the one-way ANOVA setting is beyond the scope of the course. We can teach a beginner-friendly course that does provide usable tools for simple settings. However, the tools taught may not reflect actual Bayesian practice (let alone prevailing statistical practice). As Robert (1994, p. 98) notes, "the use of conjugate priors is strongly suspicious for most Bayesians since it is mainly justified on technical grounds rather than for fitting properly the available prior information."

C. Auto-Bayes In practice we do need standard methods for standard problems. We can get them by employing noninformative default priors that are determined by the sampling distribution. In effect, we first present beginners with an explanation of the role of prior information and perhaps even of the machinery for making use of it in the simplest cases. When we come to practical settings, however, we tell our students to ignore prior information. If our students are a bit sophisticated, we may explicitly argue (Box and Tiao, 1973, p. 2) that "In problems of scientific inference we would usually, were it possible, like the data 'to speak for themselves.' Consequently, it is usually appropriate to conduct the analysis as if a state of relative ignorance existed *a priori*."

There is no Bayesian consensus on the relative place of purist, accessible, and automated methods for dealing with specific standard settings. It was once common (e.g., Lindley 1971) to begin with axioms for coherent inference, show that these imply the existence of a subjective prior distribution, and insist that use of these subjective priors is essential to the Bayesian approach. Many Bayesians now criticize this purist stance (e.g., Berger 1985, pp. 198–199). Current opinion among Bayesians seems rather to favor some version of default priors for common statistical settings. But Lindley (1971, pp. 71) is not alone in his criticism of "ready-made Bayesian analyses in which θ is just a parameter."

For practical and pedagogical reasons, Bayes for beginners will almost surely employ some mix of conjugate and noninformative priors, focusing on procedures that are computationally feasible. This also reflects a feeling that teaching the Bayesian Big Idea is more important than teaching methods we would actually recommend in practice. Even assuming this, there remain many issues on which a Bayesian consensus has yet to emerge.

- There is no agreement as to which noninformative priors we should use and teach. "Perhaps the most embarrassing feature of noninformative priors, however, is simply that there are often so many of them." (Berger 1985, p. 89) Berger offers *four* choices when θ is the probability of success in the binomial setting, and says, "All four possibilities are reasonable." See Robert (1994, p. 119) for an example due to Berger and Bernardo showing that simply reordering the parameters in the oneway ANOVA setting leads to four different "reference priors" (all of them too messy for beginners to grasp, I might add).

- Noninformative prior distributions are generally improper when the parameter space is not compact. Shall we expose beginners to improper priors? Berry (1996, p. 339) wisely lets pedagogical good sense prevail, saying only that the prior for a normal mean m is "flat over a substantial region of m-values." He also discusses normal (conjugate) priors for the mean m.

- How shall we treat hypothesis testing? The gap between estimation and testing is wider for Bayesian than for standard inference. Bayesians must generally switch priors when moving from estimation to testing, because the continuous priors used for estimation problems put probability zero on a point null hypothesis. Moreover, the results of testing seem to be more sensitive to the choice of prior distribution. The choice of methods is both quite complex and not at all settled. Finally, Bayesian and standard conclusions differ more substantially for testing than for estimation. (Ask any Bayesian about P-values.) Dempster (1971) sees Bayesian "predictive" concepts of probability as suitable for estimation, whereas frequentist "postdictive" probability better fits testing. Like many other variations of Bayesian thinking, Depmster's suggestion seems to have fallen on stony ground. See Kass and Raftery (1995) for a review of Bayesian hypothesis testing. They convince me that this topic is not yet ready for the general public.

- Yet other issues lurk beneath the surface, though we may choose to ignore them in a first course. Some types of noninformative priors depend on the choice of parametrization. We will hide this rather than admit that choosing between probability of success and odds ratio to parametrize a binomial setting changes our automated inference. Once we have made our choice of prior and obtained the posterior distribution, a loss function or utility function usually enables us to complete our inference. We may assume (silently) squared error and 0/1 loss functions for estimation and testing. Or we may simply give posterior distributions and comment informally on what actions they suggest.

It is, of course, possible to give definite answers to these questions. The difficulty is that no one set of answers is fully accepted by Bayesians. Even if we abandon the hope of contact with standard statistical usage, there is as yet no "Bayesian standard" to replace it.

Is Bayesian Reasoning Accessible?

Bayesians generally argue that the *conclusions* of Bayesian inferences are clearer than those of standard inference. I suggest that the *reasoning* of Bayesian inference is considerably less clear. Consider first this partial outline of a "standard" elementary statistics course.

> **A. Data analysis**. We begin with tools and tactics for exploring data. For a single measured variable, the central idea is a *distribution* of values. We meet several tools for graphic display of distributions, and we learn to look for the shape, center, and spread of a distribution. We also learn to describe center and spread numerically.
>
> **B. Data production**. We now distinguish a sample from the underlying population, and statistics from parameters. We meet statistical designs for producing sample data for inference about the underlying population or process. Randomized comparative experiments and random sampling have in common the deliberate use of chance mechanisms in data production. We motivate this as avoiding bias, then study the consequences by asking "What would happen if we did this many times?" The answer is that the statistic would vary. The pattern of variation is given by a distribution, the *sampling distribution.* We can produce sampling distributions by simulation and examine their shape, center, and spread using the tools and ideas of data analysis.
>
> **C. Formal inference**. We want to draw a conclusion about a parameter, a fixed number that describes the population. To do this, we use a statistic, calculated from a sample and subject to variation in repeated sampling from the same population. Standard inference acts as if the sample comes from a randomized data production design. We consistently ask the question "*What would happen if we did this many times?*" and look at sampling distributions for answers. One common type of conclusion is, "If we drew many samples, the interval calculated by this method would catch the true population mean μ in 95% of all samples." Another is, "If we drew many samples from a population for which $\mu = 60$ is true, only 1.2% of all such samples would produce an $\overline{x}$ as far from 60 as this one did. That unlikely occurrence suggests that μ is not 60."

Inductive inference from uncertain empirical data is a formidable task. There is no simple path. But there may be relatively simpler and more complex paths. Consider these characteristics of the "standard" outline above:

- A parameter is a fixed unknown number that describes the population. It is different in nature from a statistic, which describes a sample and varies when we take repeated samples.

- Inference is integrated with data analysis through the idea of a distribution. The central idea of a sampling distribution can be presented via simulation and studied using the tools of data analysis.

- Probability ideas are motivated by the design of data production, which uses balanced chance mechanisms to avoid bias. The issue of sampling variability arises naturally, and leads naturally to the key question, "What would happen if we took many samples?"

- Probability has a single meaning that is concrete and empirical: "What would happen if we did this many times?"

- Inference consistently asks "What would happen if we did this many times?" Although we use probability language to answer this question, we require almost no formal probability theory. Answers are based on sampling distributions, a concrete representation of the results of repeated sampling.

- For more able students, study of simulation and bootstrapping is a natural extension of the "do it many times" reasoning of standard inference.

This simple outline of standard statistics can legitimately be criticized as lacking generality—standard inference is limited by acting *as if* we did proper randomized data production, for example. For beginners, however, it is clarity rather than generality that we seek. I find that the reasoning of Bayesian inference, though purportedly more general, is considerably more opaque.

A *parameter* does describe the population, but it is a random quantity that has a distribution. In fact, it has two distributions, prior and posterior. We usually avoid calling μ "random" in the same sense that $\overline{x}$ is random, because a distribution for μ reflects our uncertainty, while the sampling distribution of $\overline{x}$ reflects the possibility of actually taking several samples. So in a specific problem setting, the random $\overline{x}$ could vary but the random μ can't? Gotta think about this for a while.

Probability no longer has the single empirical meaning, "What would happen if we did this many times?" Subjective probabilities are conceptually simple, but are not empirical and don't lend themselves to simulation. Because we hesitate to describe the sampling model for the data given the parameter entirely in terms of subjective probability, we must explain several interpretations of "probability." Worse, we often mix them in the same problem.

The core reasoning of Bayesian inference depends not merely on probability but on *conditional probability*, a notoriously difficult idea. Beginners must move from the prior distribution of the parameter and the conditional distribution of the data given the parameter to the conditional distribution of the parameter given the data. (Run that by me once more, will you?)

Garfield and Ahlgren (1988) survey research on the difficulties that probability ideas in general present for students. They note (p. 55) that conditional probability is particularly difficult because "an important factor in misjudgment is misperception of the question being asked." Students find it very difficult to distinguish among $P(A|B)$, $P(A \text{ and } B)$, and $P(B|A)$ in plain-language settings. Bayes-for-beginners must either shortchange the reasoning of inference or (a better tactic) use two-way tables to very carefully introduce conditional probability ideas. Conditional probability and Bayes' theorem, at any level of informality, are certainly less accessible than "What would happen if we did this many times?" at a matching level of informality. Moreover, the need to introduce more probability theory has an opportunity cost in coverage of statistics.

Bayesian conclusions are perhaps not as clear to beginners as Bayesians claim. It is certainly true that users do not speak precisely. They often confuse probability statements about the *method* (standard inference) and probability statements about the *conclusion* (Bayesian inference). "I got this answer by a method that gives a correct answer 95% of the time" easily slides into "The probability that this answer is correct is 95%." If we regard this semantic confusion as important, we ought to ask whether the user of Bayes methods can explain without similar confusion what she means by "probability." Given the multifaceted meanings of probability in Bayesian statistics, a semantically correct interpretation of a credible region may reflect no better understanding than a semantically incorrect interpretation of a confidence region.

There are numerous other complexities that the teacher of Bayesian methods must face (or choose to ignore). The use of default or reference priors opens a gap between Bayesian principle and Bayesian practice that is not easy to explain to beginners. The need to abandon what seemed satisfactory priors when we move from estimation to testing is annoying. After explaining subjective probability, we may need to deal with the oft-noted conflict between personal probabilities and physical probabilities, and even with the conflict between personal probabilities and the laws of probability. Tversky and Kahneman (1983) show that, "intuitive judgments of all relevant marginal, conjunctive, and conditional probabilities are not likely to be coherent, that is, to satisfy the constraints of probability theory." (p. 313) They also dispute Lindley's claim that coherent personal probabilities can be elicited: "we suspect that incoherence is more than skin deep." And so on.

I understand and have some sympathy for Bayesian claims to employ a single coherent approach that works in very general settings. It is *a priori* unlikely that such a general and powerful method will also be simple. Adding data (years of teaching experience, considerable reading of Bayesian expositions, and the findings of research on such topics as our understanding of probability), I am *a posteriori* convinced that Bayesian reasoning is even harder than the already hard reasoning of standard inference.

What Do We Want Our Students to Learn?

Let me conclude with a less specific—but perhaps more important—objection to basing a first course on Bayesian ideas. The teaching of elementary statistics has only recently moved from an over-emphasis on the parts of our subject that can be reduced to mathematics (probability and formal inference) toward a balanced presentation of data analysis, data production, and inference. See the report of the ASA/MAA joint curriculum committee (Cobb 1992) for a clear statement of trends that I and many others think are healthy. In particular, I believe that introductions to statistics ought to involve constant interaction with real data in real problem settings. Real problem settings often have vaguely defined goals and require the exercise of judgment. This puts me at odds with those who describe the subject matter of statistics as "making decisions under uncertainty."

What do we want our students to learn? In our more realistic moments, we recognize that *many students will not take away from our first courses any clear conceptual grasp of formal probability or of the more subtle varieties of inference.* I would place all flavors of hypothesis testing and all Bayesian reasoning in the "more subtle" category. They will, if we provide the opportunity, take away a number of more valuable messages: Always examine your data carefully, starting with graphs and simple calculations. Always ask what your data say in the context of the setting they describe. Be aware that faulty data production (voluntary response, confounding) often yields worthless data that no fancy analysis can rescue. Understand that an observed association does not imply causation, and that randomized comparative experiments are the gold standard for evidence of causation. Be sure that the data take priority over any model (such as a normal distribution or a linear relationship) used to analyze them. "Data sense" might summarize my primary objectives for a first statistics course.

Intelligent supporters of standard inference recognize that formal inference is not always appropriate. They also recognize that, even when appropriate, inference often plays a "confirmatory" role, confirming by calculation what examination of the data suggests. This understanding contributes to a willingness to reduce the traditional first-course emphasis on inference in favor of hands-on work to develop data sense.

There is, on the other hand, some tendency among Bayesians to neglect data analysis and design of data production in favor of more attention to inference. No doubt this tendency reflects in part the opportunity cost of the need to explain the Bayes machinery. It may also reflect the decision-theoretic bent of many Bayesians, which is clearly reflected in the leading advanced texts, Berger (1985) and Robert (1994). More substantively, Bayesian inference is not as well integrated with the design of data production and with data analysis as is standard inference. Many Bayesians deny the importance of randomization in data production, whereas standard inference sees randomization as validating standard sampling models. The spirit of data analysis (derived from John Tukey) is to minimize prior assumptions and allow the data to suggest models. This spirit fits uneasily with Bayesian emphasis on the importance of prior (prior *to the data*) distributions and clearly structured outcomes. Bayesian thinking seems to start with models rather than with data.

Pursuit of the Bayesian Big Idea is not in principle incompatible with developing data sense. In practice, however, it is likely to turn elementary statistics courses back toward probabilistic formalism and to leave our beginning instruction less accessible, less in contact with practice, and less in contact with data. As George Cobb so nicely puts it, "Bayesian inference offers a way to make a *probability* course deal with statistics."

Acknowledgments

I am grateful to James Berger for discussions about Bayesian viewpoints, and to George Cobb and Paul Velleman for helpful comments on a first version of this paper.

References

1. Altman, D. G. (1991), "Statistics in medical journals: developments in the 1980s," *Statistics in Medicine*, 10, 1897–1913.

2. Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis.* New York: Springer.

3. Box, G. E. P. and Tiao, G. C. (1973), *Bayesian Inference in Statistical Analysis.* Reading, Mass: Addison-Wesley.

4. Cobb, G. (1992), "Teaching statistics," in L. A. Steen (ed.), *Heeding the Call for Change: Suggestions for Curricular Action*, MAA Notes 22. Washington, DC: Mathematical Association of America.

5. DeGroot, M. H. and Mezzich, J. E. (1985), "Psychiatric statistics," in A. C. Atkinson and S. E. Fienberg (eds.), *A Celebration of Statistics: The ISI Centenary Volume*, 145–164 . New York: Springer.

6. Dempster, A. P. (1971), "Model searching and estimation in the logic of inference," in V. P. Godambe and V. A. Sprott (eds.), *Foundations of Statistical Inference*, 56–76. Toronto: Holt, Rinehart, and Winston.

7. Dunn, G., Hand, D. J., and Sham, P.C. (1993), "Statistics and the nature of depression," *Journal of the Royal Statistical Society*, A156, 63–87.

8. Emerson, J. D. and Colditz, G. A. (1992), "Use of statistical analysis in the *New England Journal of Medicine*," in J. C. Bailar and F. Mosteller (eds.), *Medical Uses of Statistics*, 2nd ed., 45–57. Boston: NEJM Press.

9. Everitt, B. S. (1987), "Statistics in psychiatry," *Statistical Science*, 2, 107–134.

10. Garfield, J. and Ahlgren, A. (1988), "Difficulties in learning basic concepts in probability and statistics: implications for research," *Journal for Research in Mathematics Education*, 19, 44–63.

11. Hammer, A. S. and Buffington, C. A. (1994), "Survey of statistical methods used in the veterinary medical literature," *Journal of the American Veterinary Medical Association*, 205, 344–345.

12. Kass, R. E. and Raftery, A. E. (1995), "Bayes factors," *Journal of the American Statistical Association*, 90, 773–795.

13. Lindley, D. V. (1971), *Bayesian Statistics: A Review.* Philadelphia: SIAM.

14. Robert, C. (1994), *The Bayesian Choice: A Decision-Theoretic Motivation.* New York: Springer.

15. Rustagi, J. S. and Wright, T. (1995) "Employers' contributions to the training of professional statisticians," *Bulletin of the International Statistical Institute, Proceedings of the 50th Session*, LVI, Book 1, 141–160.

16. Tversky, A. and Kahneman, D. (1983), "Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment," *Psychological Review*, 90, 293–315.

Table 1. Responses of 103 DOE statisticians asked to name the three techniques most important in their work.

Statistical Technique	Top 3 Responses
Regression analysis	63
Basic statistical methods	37
Analysis of variance	26
Design of experiments	26
Probability modeling	22
Sampling, survey sampling	17
Simulation	16
Graphical display and data summary	12
Multivariate analysis	12
Quality control, acceptance sampling	12
Exploratory data analysis	11
Reliability, life data analysis	11
Nonlinear estimation	7
Biostatistics, bioassay	6
Nonparametric methods	5
Numerical analysis	5
Bayesian methods	4
Time series analysis	4
Categorical data analysis	3
Variance components	3
Ranking, paired comparisons	1
Other	5

Source: Rustagi and Wright (1995).

Table 2. Statistical techniques most commonly used in 115 *New England Journal of Medicine* articles, 1989.

Statistical technique	Number of articles
t-tests	45
Contingency tables	41
Survival methods	37
Epidemiologic statistics	25
Nonparametric tests	24
Analysis of variance	23
Pearson correlation	22
Multiple regression	16
Multiway tables	11
Simple linear regression	10
Multiple comparisons	10
Adjustment and standardization	10

Source: Excerpted from Table 4 of Emerson and Colditz (1992).