

# A Conversation with James O. Berger

Robert L. Wolpert

*Abstract.* James O. Berger was born on April 6, 1950, in Minneapolis, Minnesota, to Orvis and Thelma Berger. He earned his AB, MA and Ph.D. degrees in mathematics from Cornell University in 1971, 1973 and 1974, respectively. He served at Purdue University 1974–1997, as the Richard M. Brumfield Distinguished Professor of Statistics 1986–1997, and at Duke University as the Arts and Sciences Professor of Statistics in the Institute of Statistics and Decision Sciences (ISDS) from 1997 to the present. He is also the Director of the Statistical and Applied Mathematical Sciences Institute (SAMSI), one of the U.S. National Science Foundation’s national institutes in the mathematical and statistical sciences.

Berger served as President of the Institute of Mathematical Statistics (IMS) 1995–1996 and Chair of the Section on Bayesian Statistical Science (SBSS) of the American Statistical Association (ASA) in 1995; he is President of the International Society for Bayesian Analysis (ISBA) in 2004. He has served as Co-Editor of *The Annals of Statistics* and on the editorial boards of the Springer Series in Statistics, of the *Journal of Statistical Planning and Inference*, of *Statistics and Decisions*, of the *International Statistical Review* and of *Test*. He has organized or been on the organizing committees of twenty-eight conferences, including five of the Purdue Symposia on Statistical Decision Theory and Related Topics and four of the Valencia International Meetings on Bayesian Statistics.

Berger is the recipient of a host of honors. He is an elected fellow of the AAAS, the ASA and the IMS and an Elected Member of the International Statistical Institute (ISI). He has earned Guggenheim and Sloan Fellowships and was the 1985 winner of the COPSS Presidents’ Award (joint from IMS, ASA, ENAR, WNAR and CSS) given to an outstanding statistician below forty years of age (it is particularly noteworthy that he won that award at the tender age of 35), and was selected as the COPSS Fisher Lecturer in 2001. He was elected as a foreign member of the Spanish Real Academia de Ciencias in 2002, and to membership in the U.S. National Academy of Sciences in 2003.

Berger has been the advisor of 30 Ph.D. students, has written or edited thirteen books and has numerous other statistical publications. A listing of these publications can be found at his Web site, <http://www.stat.duke.edu/~berger/>.

Berger married Ann Louise Duer (whom he first met when they were in the seventh grade together) in 1970, and they have two children, Jill Berger,

---

*Robert L. Wolpert is Professor, Institute of Statistics and Decision Sciences, and Professor, Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina 27708-0251, USA (e-mail: wolpert@stat.duke.edu). Robert and Jim have known each other since 1969 when they were 19-year-old college roommates.*

who is married to Sascha Hallstein and works as an optical scientist in Silicon Valley, and Julie Gish, who is married to Ryan Gish and works as a consultant in Chicago.

The following conversation began at the home of James and Ann Berger on May 18, 2003, and was completed over the next six months in a wide variety of places on three continents.

### EDUCATION

**Wolpert:** Talk a little about your high school encounters with science and mathematics. How did you come to be recruited by Cornell for the Ford Foundation-sponsored experimental “Six Year Ph.D. program?” Did you know early on what you wanted to do professionally?

**Berger:** I always liked science and mathematics, but my high school did not have much in the way of mathematics, not even any kind of introduction to calculus. So, overall, I had a very enjoyable—but not very academic—high school experience.

As usual, I applied to a number of universities, Cornell among them. When I was informed about the possibility of entering their experimental Six Year Ph.D. program, I thought that sounded exciting; I knew I wanted a Ph.D., and completing all my education within six years was appealing.

My parents had instilled a love of knowledge in me, and so being a professor had always appealed to me as a career. When I left high school, however, I had no real idea as to what I wanted to be a professor of.



FIG. 1. Orvis, James and Thelma Berger (top) with sisters Kathryn, Jo Marie and Nancy (bottom).

**Wolpert:** What sparked your interest in mathematics and statistics at college, and what motivated your decision to study them seriously?

**Berger:** At first, I was trying things other than mathematics and statistics in college. My first year at Cornell I decided to be an economist, and took mostly economics courses. I recall particularly liking a graduate course on microeconomics that was very heavy on math. During the subsequent summer, a family friend convinced me to become an MD-Ph.D. and do medical research, so the first semester of my second year I took all chemistry and biology courses. Again, I found that I liked the mathematical aspects of these courses best, and so started considering choosing mathematics as my field.

At that point, you became my roommate at college, and I remember we started looking at the catalogue deciding what math courses sounded cool. “Stochastic Processes” sounded particularly cool, as you will recall, but the professor, Kioshi Itô, strongly recommended that we have a course on probability theory before we took it. Alas, since it was the second semester, we were out of sequence for taking probability, but Frank Spitzer was very kind and gave us a reading course on introductory probability [out of Chung’s 1968 book, I believe (Chung, 1968)]. It was a fun way to learn probability, having to work lots of things out on our own, yet obtaining weekly inspiration from Frank. Curiously, as I recall, we never did get to take Itô’s course—he went on leave or something.

**Wolpert:** Although it did not have a statistics department at the time, Cornell did have a number of celebrated statisticians; which ones influenced you, and how? Can you recount some memorable moments in class or in your Ph.D. study?

**Berger:** The other main event of that second semester—actually I should say the other main academic event, since that is also when I got married—was taking Jack Wolfowitz’s graduate class on statistical inference. That course got me to start thinking about statistics as a career. Wolfowitz was an entertaining lecturer, and we heard all sorts of funny things about Bayesians (and other undesirables). I eventually decided to choose mathematics for my Ph.D. at Cornell, and eventually decided to write a thesis on statistics.

Roger Farrell taught us classes on design and multivariate statistics, demonstrating that statistics could



FIG. 2. *Ann and Jim Berger in 1970, the year they were married.*

be done as elegant mathematics. Larry Brown was our teacher in decision theory out of his famous (but still unpublished) notes. In addition to the course itself, I liked his touches such as bringing wine to class the day his first child was born, and just taking the hour to chat about the statistics profession. I did not have Jack Kiefer—the other famous statistician in the math department at the time—in any classes, since he was on medical leave while I was taking courses.

I eventually ended up working with Larry on a thesis in decision theory. I started working on a problem that seemed to me, at the time, to be a “big unsolved problem.” Larry let me work on it for a while, but gradually made me realize that what seems important as an unsolved problem from a mathematical perspective is not the same as what is important as a statistics prob-



FIG. 3. *Robert Wolpert and Jim in 1970 at Cornell.*

lem. This was the beginning of what was to be a long process of my trying to find a personally optimal mix of mathematics and statistics in my work. I finally followed Larry’s advice and worked on a very beautiful statistical problem in admissibility that he suggested.

**Wolpert:** In the early years of your research, name those individuals who had the most influence and why?

**Berger:** I won’t talk about those who were primarily coauthors, such as yourself; coauthors naturally have a profound influence, and that influence is reflected in the work itself. Besides being my advisor, Larry Brown has had an enormous influence on my thoughts; he has always been quite interested in the frequentist–Bayesian interface, and has provided a continual set of crucial insights into this interface. Even after graduation at Cornell, Roger Farrell was extremely helpful; he taught me a lot about statistical professionalism. Curiously, I became good friends with Jack Kiefer after I left Cornell, and especially when I was on sabbatical at Stanford in 1979–80 and he had moved to Berkeley. Jack had been exploring conditioning in the mid to late 1970s, and he was of great value in my own beginning understandings of the subject. We also had fun, ranging from periodically hitting the racetrack to a photo-op that presented itself at a meeting in Japan, of Jack (at maybe 140 pounds) pretending to wrestle with Alexeyev, the world super heavyweight weightlifting champion at the time.

I was, of course, greatly influenced by the faculty at Purdue, from the moment I arrived there in January 1974. Leon Gleser had an office next to mine, and Herman Rubin’s was not far away, and I would continually run to them with questions and would always get answers (and the answers usually turned out to be right!). Herman was the avowed Bayesian in the department, and so had a dramatic influence on my developing into a Bayesian. Finally, I should mention Shanti Gupta at Purdue, who, over a period of 20+ years, was central to much of my career progress and was a role model and mentor on professional involvement in the statistical community. I was just back at Purdue for the Seventh Purdue Symposium, held in memory of Shanti, and it brought back many memories of all he had done for me and the profession.

As I began to enter the Bayesian community, Morris DeGroot, Arnold Zellner, Jack Good, Bruce Hill and Dennis Lindley had a major effect on my development. Morrie, with gentle prodding, continually kept me moving toward a fuller appreciation of Bayesian analysis, when my frequentist upbringing would cause



FIG. 4. *Shanti Gupta sailing with Jim in 1978 in San Diego.*

me to stop only part-way along the path. Arnold introduced me to objective Bayesian analysis, and sold it (and applied Bayesian analysis) to me with remarkable enthusiasm. I had numerous interactions with Jack, Bruce and Dennis that greatly helped me to understand different aspects of Bayesianism.

#### PERSONAL LIFE

**Wolpert:** In high school and college you didn't participate in team sports but you were active in skiing and sailboat racing—you partly worked your way through college by teaching competitive sailing, I believe, and won a host of trophies racing on Minnesota lakes. Do you still sail or ski, competitively or just for fun?

**Berger:** Just for fun, but I don't get out skiing much anymore, and my sailing is pretty much limited to day-sailing at family gatherings [although I did get out in the Bay after the San Francisco Joint Statistical Meetings (JSM) last August, in quite a gale—losing glasses, hats, etc., but luckily no statisticians].

**Wolpert:** You competed in games too, back then—bridge and chess, for example. Your friends were all a little in awe of your memory, especially your playing blindfolded chess, or being able to remember each bid and play at the bridge table days after the game had ended. I know you liked to count cards at blackjack (and even taught your preteen daughters to do it!); do you play cards much these days?

**Berger:** I don't think I ever had a great memory. I somehow was able to recall what had happened in games by thinking about the processes of the games themselves. (Likewise I have always been rotten at remembering math and statistics facts, theorems,

etc.—I continually have to go through a process of rederiving even simple things.)

These days, I try to play bridge a couple times a month (lately with Alan Gelfand), and do have a weakness for blackjack whenever I get near a casino. I just count cards for fun, however—I don't operate with the extreme variation in betting levels and long hours of concentrated work necessary to make any serious money at blackjack.

**Wolpert:** What else do you do other than statistics?

**Berger:** Statistics gives me an active social life (I like socializing with statisticians), and numerous opportunities to travel, which I also like. Luckily, Ann also enjoys these activities (well, at least she appears to enjoy socializing with statisticians). Our children also have always liked travel, so we have many family get-togethers in interesting places.

I also hike a little, find and drink nice wines (a lot), go to movies and watch some TV—so nothing like a serious hobby.

**Wolpert:** So Ann really likes going to statistical meetings?

**Berger:** Yes, she likes seeing former students and acquaintances. And she certainly gets into the spirit of Bayesian meetings. For a while, she obtained her exercise as an aerobic dancing instructor, at a time when I was focusing on Bayesian decision theory, and during that time she loved to say at statistics meetings that she and I worked to opposite ends: her job was to maximize posterior expected loss, while mine was to minimize posterior expected loss. Then, later, at the lively Valencia meeting poster sessions (held in the bar from 10 pm to 1 am), she would walk from poster to poster, stare at the poster for a while and then say, "Hmmm, are you sure that posterior is proper?" Not only would this cause considerable consternation among the poster presenters, but once an author looked at his presentation and said "Oh my gosh (or words to that effect), I forgot a page!"

**Wolpert:** Any other statisticians in the family?

**Berger:** Well, my daughters, Jill and Julie, both know Bayes theorem! Jill received her Ph.D. in optical physics and has been doing exciting things with lasers in Silicon Valley, and Julie just received her MBA at Harvard Business School—a former (but, alas, not current) hotbed of Bayesianism. I do have a nephew, Ethan Anderes, in statistics—he is getting his Ph.D. at Chicago.



FIG. 5. Jill, Julie, Ann and Jim at Julie's 2003 graduation from Harvard Business School.

### PROFESSIONAL LIFE

**Wolpert:** You spent 23 years at the Department of Statistics at Purdue University. Tell us about that.

**Berger:** Obviously that period formed most of my academic career, and was a very fruitful time. In addition to the mentors I mentioned earlier at Purdue, who had a lot to do with my development, I had numerous colleagues that were great to collaborate with, including Mary Ellen Bock, Anirban DasGupta, Tom Sellke and Jayanta Ghosh. The university was always very supportive of statistics, which made the environment particularly pleasant. In addition to the well-known Purdue Symposia, a number of important series of workshops got their start at Purdue, including the current popular "OBayes" series of workshops on objective Bayesian analysis.

**Wolpert:** You were influential in the formation of the Institute of Statistics and Decision Sciences at Duke University, and eventually joined us. Tell us about that.

**Berger:** The formation of ISDS at Duke in the mid-1980s was an exciting event for the Bayesian community, since it was to be a department with a heavy Bayesian presence. My main role in this was in solidifying the notion in the Duke administration (instilled by others) that the Institute of Statistics

and Decision Sciences (ISDS) was a very good idea, and helping to suggest personnel. The initial director, John Geweke, and then Mike West (as director for 12 years) and new director Dalene Stangl turned ISDS into an interdisciplinary Bayesian powerhouse, which I decided to join in 1997, in large part because I was very interested in the leading-edge applications, using Bayesian methodology and Markov chain Monte Carlo (MCMC), that were happening here.

**Wolpert:** You headed the effort to bring about the birth of SAMSI, the Statistical and Applied Mathematical Sciences Institute in Research Triangle Park, and are its founding Director. Throughout your career you had somehow avoided the administrative duties that come with chairing a department; how did you decide to accept the challenge this time?

**Berger:** First, I should mention that the formation of SAMSI was a huge joint effort, involving many individuals, especially from Duke University, North Carolina State University (NCSU), the University of North Carolina at Chapel Hill (UNC) and the National Institute of Statistical Sciences (NISS) (which are all formal partners of SAMSI). Tom Banks from NCSU, Alan Karr from NISS and Steve Marron from UNC have been especially crucial as associate directors of SAMSI, and their presence partly answers your

second question—SAMSI was planned so that many of the administrative functions are shared with these individuals, so that I felt comfortable accepting the directorship.

I should also stick in a little advertisement for SAMSI here. Since SAMSI is quite new, many people do not know about the opportunities that are available at the institute. We are very interested in proposals for future research programs of one or two semesters at SAMSI; there are many opportunities to participate in already-planned programs—from long-term visits to attending workshops—and we have a variety of post-doctoral and other positions available. To see what is happening at SAMSI and what opportunities it offers, check out the Web site at <http://www.samsi.info/>.

**Wolpert:** You have organized or coorganized a wide range of important statistics meetings, especially the Purdue Symposia on Statistical Decision Theory and the Valencia conferences on Bayesian Statistics, and have coedited eight volumes of proceedings. What do you find most appealing about conference organizing—the chance to help shape the research agenda for a field, the chance to help develop young people’s careers or the chance to satisfy your sense of duty to the profession?

**Berger:** I never really thought about it before, but I guess all three are the reasons I have done so much conference organizing. Oh, and I would add the fourth reason of staying up-to-date; organizing conferences is a great way to find out what is currently happening in statistics.

**Wolpert:** You have served as IMS President, Section on Bayesian Statistical Science (SBSS) Chair,

Co-Editor of *The Annals of Statistics*, and are the president-elect of the International Society for Bayesian Analysis (ISBA); how have you made use of these opportunities?

**Berger:** These were all quite different experiences. I was one of the first chairs of SBSS, so that was a building effort. Being IMS President reaffirmed to me what a wonderful organization it is—the organization gets so much done, and essentially all on a volunteer basis. The *Annals* coeditorship was, in part, an effort on my part to reconnect with non-Bayesian statistics, after quite a few years of immersion on the Bayesian side. Having Hans Künsch as a coeditor helped immensely in keeping the job interesting and preventing it from being consuming. ISBA is off to a great start, and exciting things are happening this year under Ed George’s presidency, such as the creation of a Bayesian journal with Rob Kass as founding editor, but there are numerous Bayesian communities out there that are not a part of ISBA, and I hope to increase their involvement with ISBA.

**Wolpert:** You have directed the Ph.D. dissertations of thirty students. How are they doing? Any stories?

**Berger:** They are doing great! Twenty are in academics, and publishing lots of wonderful stuff, and quite a few have gone on to leadership roles in statistics. The others have very successful industry careers. I’m sure they have lots of good stories about me, but I better not tell any about them.

**Wolpert:** Congratulations on your recent election to the National Academy of Sciences (NAS). Are there any other Bayesians there? How optimistic are you that the NAS will offer an opportunity to increase the visibility and acceptance of Bayesian methodology in U.S. science?

**Berger:** Thanks and yes, there are certainly other Bayesians in the NAS. David Blackwell, Persi Diaconis, Stephen Fienberg and Brian Skyrms prominently mention Bayes in their academy bios; others, such as Fred Mosteller and Chris Sims, have called themselves Bayesian at one time or another; and many in the NAS have done considerable amounts of Bayesian work.

I am just beginning to understand the workings and the role of the NAS in influencing research and providing scientific input into public policy. I expect to have the opportunity periodically to say—“Hey, that is a societal problem that really needs heavy involvement of the statistical profession”—but that is about all I can say at this point.



FIG. 6. Jim and Larry Brown in 1995 at a conference at Purdue University organized by Jim in honor of Shanti Gupta.



FIG. 7. At the 1993 JSM in San Francisco, Jim got together with some of his Ph.D. students, left to right and top to bottom: Jim Albert, Gene Hwang, Dipak Dey, Mark Berliner, S. Sivaganesan, Duncan Fong, Jean Francois Angers, Sudip Bose, Keying Ye, Dongchu Sun, Ming Hui Chen and Chunfu Qiu.

**Wolpert:** Even before election to the U.S. NAS, I noticed that you were elected in 2002 as a foreign member of the Spanish Real Academia de Ciencias. How did that come about?

**Berger:** Part of it is that I have worked extensively with many Spanish statisticians, including Susie Bayarri, José Bernardo, Elias Moreno and David Ríos Insua, and organized many meetings in Spain. But probably most importantly is that statistics in general, and Bayesian statistics in particular, has very significant visibility within the Real Academia, in large part due to the efforts of members such as Javier Girón.

Ann and I were over for the induction ceremony in Madrid last December. The trip included a very enjoyable tour of Spanish universities, and one of the funniest things was reading the resulting newspaper articles about what Bayesian statistics meant and the wonders it would do for the world. (Not that it won't do wonders; what was funny is which wonders the newspaper people said it would do!)

## RESEARCH

### Theory or Methodology or Application

**Wolpert:** During your career the focus of your statistics articles seems to have shifted from theorems

and proofs to models, methodology and explication. Why has this happened and is it important?

**Berger:** When I was primarily a theoretician, the practical impact of a theorem certainly made my top five list of virtues, but was not number 1 or 2. Gradually, "having direct practical impact" rose up my list of virtues to number 1.

Another factor in this change is that, early on, I had an extremely optimistic view of the applicability of my theorems, a view of which I gradually became disabused the more I worked on applications. With my current stricter view of what has practical impact, I find it a lot harder to find theorems that are both fundamentally interesting and that have immediate practical impact, and so I tend to focus more on methodology. Still, I must admit that nothing makes me quite as happy as finding a really unexpected theorem that also has immediate practical value, such as my recent result with Marilena Barbieri that the median probability model, and not the maximum probability model, is typically the best model to use for prediction.

**Wolpert:** The traditional mathematical objection to Bayesian methods is that they seem to offer fewer exciting mathematical challenges than classical statistics: fewer theorems and proofs. Is this a real distinc-

tion between frequentist and Bayesian methods, or is it a sociological or demographical distinction?

**Berger:** I first heard the viewpoint stated by a prominent statistician, something like: “If the Bayesian paradigm is right, I would find something more interesting to do than being a statistician.” Behind this notion is that, after Bayes’s theorem, everything else is just application, so what is the challenge?

A superficial answer to your question is to simply note that, over the past quarter century, Bayesian analysis has been an extremely fruitful source for deep and interesting theorems. There is another answer, however, that I think is more relevant to being a Bayesian. The most common results in classical statistics are of the form: “Here is a demonstration of Property  $A$  of Procedure  $X$ .” In contrast, the Bayesian approach is to put the work into modeling the data and the unknowns in the problem, and then let Bayes’s theorem automatically construct the procedure and guarantee good properties (if the hard modeling work was done well). This constructive Bayesian process involves a sophisticated interplay of data modeling, prior choice and computation, and I generally find it more interesting than dreaming up statistical procedures and proving theorems about them.

**Wolpert:** Most statisticians probably think of you as a theoretician, but you have done lots of interdisciplinary and applied work. What mix do you like, and how has it influenced your research?

**Berger:** It wasn’t until the 1990s that I routinely started doing large applications, the first major one being a hierarchical Bayesian analysis of fuel efficiency standards with Richard Andrews (who sadly died recently) and Murray Smith. I am now working a lot with astronomers (especially Bill Jefferys) and on a major NISS project (led by Jerry Sacks) involving computer model validation with a host of applications. Solving interdisciplinary problems can be quite fun in itself, of course, and is crucial for the health of our profession, but my involvement with such problems has had the ulterior motive of making sure my theoretical and methodological developments are on track. Modern Bayesian analysis requires a melding of theory, modeling and computation, as I just mentioned, and the only way to get a feel for the needed mix is to be involved constantly with major applied problems.

### Shrinkage Estimation and Robustness

**Wolpert:** Early in your career you worked mostly on minimaxity and admissibility, maybe inspired by Stein’s astonishing result that the usual (minimax)

sample-average estimator of a multivariate normal mean is not admissible. What made you gradually shift to research about Bayesian methods?

**Berger:** Working on shrinkage estimation naturally causes one to start thinking a little bit like a Bayesian (where does one shrink to?), and most work in admissibility primarily utilizes Bayesian tools, so I was getting lots of exposure to Bayesian concepts. I really didn’t start thinking of myself as a Bayesian, however, until I realized that minimaxity is not compatible with practical Bayesian shrinkage in nonsymmetric situations. (Minimaxity essentially requires shrinking the noisiest coordinates the least, while Bayesian reasoning suggests that they should be shrunk the most.)

By the way, my interest in admissibility never really ended. For instance, I have done recent work with Christian Robert and Bill Strawderman on developing objective prior distributions for hierarchical modeling. Hierarchical modeling is pervasive in statistical practice today, and getting the hyperpriors right can be very important, and I feel that admissibility is a key tool that allows us to get the hyperpriors right.

**Wolpert:** Did your original interest in minimaxity lead to your research in Bayesian robustness?

**Berger:** Yes and no. After understanding that minimaxity did not, in general, provide the right type of shrinkage, it was a natural step to say—okay, let’s consider Bayesian shrinkage estimators that are not strictly minimax, but which are robust in various senses related to minimax ideas, such as gamma minimaxity or restricted risk Bayes procedures.

Later, after I had come to grips with the notion of conditioning and became oriented toward true Bayesian inference conditional on the data, I became excited about a quite different type of Bayesian robustness—robustness of the (conditional) Bayesian answer with respect to uncertainty in the prior. In some sense, I believe that this is the fundamentally correct paradigm for statistics—admit that the prior (and model and utility) are inaccurately specified, and find the range of implied conclusions—and so I did a lot of work on this in the mid-1980s to early 1990s with Dipak Dey, Mark Berliner, Siva Sivaganesan, Tony O’Hagan and Susie Bayarri among others.

### Conditioning and Testing

**Wolpert:** You have raised the conditioning issue—when did you become aware of it, and how crucial was it in your evolving perspective?

**Berger:** It happened in the later stages of writing the first edition of my book on decision theory (Berger, 1980), when I ran across the famous papers, such as Birnbaum's (1962), on the subject. It took me many years before I fully understood how central conditioning should be to all that we do, but even early on I recognized that I needed to stop what I was doing and understand conditioning. That is when I embarked with you on writing our book *The Likelihood Principle* (Berger and Wolpert, 1984, 1988); if you realize you really need to *understand* something, write a book on it!

**Wolpert:** (Interrupting.) Would you recommend that readers buy that book?

**Berger:** (Laughing.) Well I know the question isn't motivated from self-interest, since we don't get any royalties—the proceeds go to a worthy cause (the IMS). I will say that the book collects, in one place, most of the thought-provoking and illuminating examples and “paradoxes” about conditioning that were developed by previous authors, and hence is an easy way to learn about the topic.

**Wolpert:** Conditioning plays a central role in the debate between Bayesians and frequentists; the Bayesians condition on what is observed and the frequentists condition on . . .

**Berger:** My turn to interrupt, because although conditioning has historically been central to the debate, it is far from clear that it *should* separate Bayesians and frequentists. Frequentists can also condition, and indeed there is a long history of such conditioning, starting with procedures such as the Fisher exact test and attaining status as a formal paradigm through the work of Jack Kiefer and Larry Brown in the mid-to late 1970s. There was also lots of interest in conditioning in the frequentist shrinkage literature at the time when I was still active there, for example, the paper in which Hwang and Casella (1982) showed the relevance of the issue by producing 95% shrinkage-based confidence intervals that could have zero width!

I won't go into any more detail here, because the article in this volume by Bayarri and Berger (2004; henceforth BB) addresses the issue in considerable depth. I should, however, recount two stories that highlighted interesting aspects of my personal journey in understanding conditioning.

The first story involves Morrie DeGroot. Between editions of our *Likelihood Principle* book, I remember the long discussion we had one evening (I suspect scotch was involved) with Morrie and Susie Bayarri at a meeting at the Ohio State University about whether

it was truly possible to define appropriate conditioning outside of a Bayesian context (I recall that we decided in the second edition of our book that it was possible); this led to Morrie's famous line, delivered in a talk on the subject at the Fourth Purdue Symposium: “The most important thing to determine is—where is the bar? A Bayesian always knows” [referring technically to whether one should condition on unobserved variables  $y$ , leading to  $f(x|y, z)$  as the likelihood function, or not, leading to  $f(x, y|z)$  as the likelihood—lexigraphically, the choice becomes one of where to put the conditioning bar “|”].

The second story revolves around what I consider the most surprising success of conditioning—in our reconciliation of frequentist and Bayesian testing. I recall giving a talk in 1991 at a workshop organized by George Casella at Cornell University; the talk focused on the issue that testing did *not* seem to be reconcilable for Bayesians and frequentists, except in special situations involving symmetry. Then, at lunch after the talk, Larry Brown, you and I started chatting about whether it really was irreconcilable (as it was lunch, I do not believe scotch was involved) and, lo and behold, we shortly found that unification *was* possible in simple versus simple testing (Berger, Brown and Wolpert, 1994), a result which has since been extended to much more general testing in work with Ben Boukai, Yinping Wang and Sarat Dass (Berger, Boukai and Wang, 1997, 1999; Dass and Berger, 2003). Since then I have always been very wary of saying that there is anything fundamentally irreconcilable between Bayesians and frequentists.

**Wolpert:** In your 2001 Fisher Lecture at the Atlanta JSM (Berger, 2003) you presented our unified approach to testing as an answer to the rhetorical question of whether Fisher, Jeffreys and Neyman could have agreed on statistical testing. The hope of course is not just for reconciliation, but to find a way to make available to frequentist practitioners some of the advantages of Bayesian procedures, like the stopping-rule independence for sequential procedures, which could offer enormous benefits in clinical trial design and similar areas. Is it realistic to hope for such a reconciliation?

**Berger:** In the long run, I think that reconciliation is inevitable, as it becomes better understood that *good* Bayesian and *good* frequentist viewpoints are simply two illuminations on what I think are the central core truths of statistics. An encouraging sign is that I observe many avowed non-Bayesians grabbing Bayesian tools and many Bayesians grabbing frequentist tools to

handle the complex problems that we are encountering these days.

### ***p*-Values and Model Selection**

**Wolpert:** Your work with Bayarri on Bayesian *p*-values has perplexed both Bayesians, who wonder why you concern yourself with *frequentist* measures of significance, and frequentists, who wonder what a *Bayesian* perspective can offer them. The work generated enormous controversy at the 1998 Sixth Valencia conference (see Bayarri and Berger, 1999, and its discussion). Can you say a few words to those perplexed Bayesian and frequentist statisticians?

**Berger:** Some of the perplexity has undoubtedly been due to the fact that, for about a 10-year period—starting with my 1987 *Journal of the American Statistical Association* (JASA) article with Tom Sellke (Berger and Sellke, 1987) and in a number of papers with Mohan Delampady and Julia Mortera—I wrote a lot about how inappropriate it is to use *p*-values to reject null hypotheses. Then, at the Sixth Valencia conference and in subsequent papers, I was writing with Bayarri about how to define and compute *p*-values properly, seemingly a turnaround.

The explanation is that, while I do not believe that *p*-values are useful for quantifying the evidence against a precise hypothesis, I do feel they are useful for answering the question “Is there any indication that something unusual has happened, so that I should spend the effort to try to formally quantify the evidence against the current model or hypothesis?” The problem that Bayarri and I noticed is that many of the *p*-values that were in common use were quite unsuitable for answering this latter question, in that they would not be small even when the model or hypothesis was hopelessly wrong (an example is given in BB). Using Bayesian tools, we then showed how to develop good *p*-values for answering this question. Further confusing matters, we essentially utilized frequentist notions, based on Robins, van der Vaart and Ventura (2000), to define a good *p*-value; but this is just another part of the central theme that optimal statistics requires an interplay between Bayesian and frequentist concepts.

**Wolpert:** If *p*-values shouldn't be used for the purpose, what would you recommend for someone seeking an “objective” way to evaluate or choose among models? Tell us something about your work on Bayesian approaches to model selection in astronomy and other fields. Is it possible or even desirable to choose objectively among prior distributions and models?

**Berger:** In Bayesian model selection, use of the word “objective” becomes strained. That is because the usual type of objective priors (Jeffreys priors, reference priors, . . .) are not typically satisfactory in model selection, and instead one has to choose “conventional proper priors.” The astronomy work to which you referred, with Bill Jefferys and Peter Müller, is of this type.

The approach I and collaborators (especially Luis Pericchi) have taken is to develop priors by, in some sense, either bootstrapping off the data or by bootstrapping off imaginary data arising from a special model. Our approach differs from others primarily in that we insist that it should correspond, in convincing ways, to a Bayesian procedure with a sensible conventional proper prior. Many approaches in use today have no such correspondence, even many that are formally derived using Bayesian tools.

There is really a long way to go here. We are far from sorting out the issues, and Jayanta Ghosh has convinced me that frequentist reasoning is going to play an especially big role in finally obtaining good general objective priors for model selection.

### **Decision Theory Book**

**Wolpert:** Your book *Statistical Decision Theory: Foundations, Concepts and Methods* (Berger, 1980) was immensely popular and influential. Strong books were already available on decision theory; why was it important for you to contribute another? How did you make the decision to invest the enormous amount of time it takes to write such a book?

**Berger:** There were two motivations for me to write the original version of the book. The first was that a lot of work had been done on decision theory (in shrinkage estimation, for example) since the previous decision theory books were written. The second motivation, perhaps surprising for those who read the book today, is because I thought there should be a more “practical” book on frequentist decision theory. Practical books on Bayesian decision theory existed at the time but recall that, when I started writing the book, I was primarily a frequentist, and the existing frequentist books on decision theory were written at a quite advanced level.

There were actually two stages to the process of writing the book. I only came face-to-face with conditioning near the end of writing the first edition of the book, and had not yet fully internalized the issues even as the book went to press. A couple of years later I had

a better understanding of conditioning, and Bayesianism in general, and decided the first edition no longer reflected what I believed—hence the second edition. Amusingly, shortly after the second edition was published, I received a letter from the Springer editor saying that he had received numerous requests to have the first edition reprinted—the second edition had become too Bayesian!

### PROGRESS OF BAYESIAN STATISTICS

**Wolpert:** What was the status of Bayesian statistics when you graduated in 1974?

**Berger:** There was a lot of Bayesian activity in the United Kingdom, with work of Dennis Lindley, Phil Dawid, Adrian Smith and many others, but I was not aware of this activity when I graduated. Indeed, I essentially only knew a bit about the status of Bayesian statistics in the United States—in 1974 it was viewed as rather eccentric and not suitable for most mainstream applications. Probably the most influential Bayesian activity of the time in the U.S. was the series of workshops entitled Seminar on Bayesian Inference in Econometrics and Statistics (SBIES), initiated by Arnold Zellner in 1970. The Workshops were held twice a year, and provided a gathering place for Bayesians in the U.S. My own initial engagement with the Bayesian community was primarily through participation in these workshops. One thing I remember, even from back then, was a vociferous debate between the subjective Bayesians and the objective Bayesians.

Outside of this group, when one heard the word “Bayes” at that time in statistics in the U.S., it was mainly in the phrase “empirical Bayes,” which Brad Efron and Carl Morris were in the process of popularizing.

**Wolpert:** What happened to Bayesian statistics in the 1980s?

**Berger:** There was dramatic growth in research and application. An ever-increasing number of Bayesian meetings gave presence to the field and attracted interested outsiders, especially the Maximum Entropy Workshops, motivated by the Bayesian approach of Ed Jaynes in the physical sciences, and the Valencia meetings started by José Bernardo. (And contrary to widespread rumor, some of those attracted to the Valencia meetings went primarily for the statistics!)

**Wolpert:** How about the 1990s?

**Berger:** While the 1980s was an exciting growth period for the Bayesian community, I would characterize it as a time where Bayesian analysis pretty

much stayed inside that community. During the 1990s something very different happened, with many non-Bayesian statisticians—and even hordes of nonstatisticians—coming to utilize Bayesian methods on a regular basis.

Of course, the Bayesian community itself continued to grow at a rapid rate. ISBA was formed (after considerable debate, centering on the issue of whether it would have an isolationist effect on the community), along with various chapters in a number of countries. In the U.S., SBSS was formed [an effort spearheaded by Jim Press with support of then American Statistical Association (ASA) President Arnold Zellner]. These organizations have had a major effect on Bayesian visibility and opportunities for Bayesians.

**Wolpert:** To what extent do you attribute the great increase in visibility of Bayesian statistics today to the following:

1. Dramatic improvements in hardware and software?
2. Increasing interest in nonstandard applications, and suitability of Bayesian analysis for application in complex situations?
3. Answering relevant questions—like evaluating a hypothesis by its posterior probability, which answers “given the observed data, what is the probability this hypothesis is true?” as opposed to the frequentist  $p$ -value, which answers the question “if this hypothesis is true (which it might not be), what is the probability of observing even more extreme data (which we didn’t)?”
4. Foundational issues, like the fact that all admissible procedures are Bayes or the limit of Bayes?
5. Intellectual excitement and lifestyle?

**Berger:** You have made a good list of the reasons. I think the combination of numbers 1 and 2 has been the major cause for the incredible growth of Bayesian statistics in the last 20 years. Twenty years ago I often heard the comment, from would-like-to-be-Bayesians, “Bayesianism is very attractive, but I can’t do the needed computation in complex problems.” In contrast, today I often hear the comment, from avowed non-Bayesians, “I do not believe in Bayesianism philosophically, but the Bayesian approach (with MCMC) was the only way I could analyze this complex problem, so I used the approach.”

That Bayesian statistics allows one to ask any desired question, and obtain an answer to that question, has always appealed to many. I think the appeal is particularly strong among nonstatisticians. Practical classical statisticians become extremely clever at finding

indirect ways to address a question of interest. Your example of testing and  $p$ -values (which is actually an improvement in clarity on the classic similar quotation of Jeffreys) is a great example of this. Good classical statisticians do realize what question the  $p$ -value is answering, and become adept at understanding how to intuitively address the “real question,” utilizing things like power and sample size, in addition to just the  $p$ -value, but this is very difficult for nonstatisticians to do well.

Foundational arguments had a strong impact on many of the old-time Bayesians (within which category you and I, alas, now fall), but they are probably markedly less important than the first three reasons you mention for modern Bayesian growth.

**Wolpert:** By the way, you didn’t say much explicitly about the “Bayesian lifestyle” above. What influence did that have?

**Berger:** Ah, you won’t let that slide by. Well, it is probably true that many neo-Bayesians in the old days found the lifestyle (parties and dancing at meetings, for example) quite attractive, as were the typical locations of the meetings. And, of course, Bayesianism is rather infectious if you hang around Bayesians at all. Now, however, things are very different, with today’s Bayesians being as disparate, in terms of lifestyle preferences, as any statisticians. (But we still have lots of dancing at Bayesian meetings!)

**Wolpert:** Are the walls broken down yet, or must Bayes do more to penetrate good science and policy?

**Berger:** I feel that they are broken down within the statistics profession, in the sense that I think very few people now dismiss Bayesian analysis out of hand. Furthermore, outside the statistics profession there is enough awareness of Bayesian statistics that ignorance is rapidly disappearing as a barrier. Some walls remain, however, from the institutionalization of non-Bayesian procedures, where Bayesian procedures would work much better. My favorite example (or should I say *least favorite*) of this is the use of  $p$ -values to “statistically reject” hypotheses.

**Wolpert:** So, let’s take that example. Why do  $p$ -values still appear for hypothesis testing in scientific publications? Will they eventually die out?

**Berger:** That’s a tough one. I think most statisticians will agree that  $p$ -values are widely misused. At the moment, however, our profession does not agree on an alternative to the use of  $p$ -values for testing and, until we do, it is going to be hard to make progress on eliminating the misuses. I recall being at a national meeting of wildlife scientists several years ago,

in which they argued vociferously (and I think justifiably) that the statistics profession sends out so many mixed signals on this issue that we are to blame for the current state of affairs. In this regard, the “Fisher–Jeffreys–Neyman” synthesis mentioned earlier was introduced primarily as a potential mechanism by which we could speak as a profession with a unified voice.

**Wolpert:** A lot of your career has been devoted to methods of choosing prior distributions that have an appearance of objectivity. This “objective Bayes” approach is at odds with the classical subjective Bayesian view, as espoused by Lindley and others. Where do you see this going? Are objective priors or subjective elicitation going to be more important in the future? Would you say that fields such as public policy, public health and related fields demand the appearance of objectivity?

**Berger:** This question covers a lot of important ground. First, I would quibble with the notion of what is the “classical Bayesian view.” Starting with Bayes and Laplace and until nearly the middle of the twentieth century, the dominant Bayesian view was some version of objective Bayes—the “inverse probability” of Laplace (using constant priors), at the end merging into the Jeffreys approach. During the past 50 years, the subjective Bayes approach has also been prominent, but it should be called “nouvelle Bayes.” (Of course, one could also say that frequentist and Fisherian statistics should be called “nouvelle statistics,” because the truly classical approach is the inverse probability approach of Laplace.)

In a more serious vein, I wish that objective Bayes were not viewed as being at odds with subjective Bayes, but rather as being complementary. As you point out, there are many fields (and many statisticians and other scientists) who (rightly or wrongly) feel that the appearance of objectivity is crucial in statistical analysis and have, therefore, rejected Bayesianism out of hand. How much better it would have been to talk proudly about the objective side of the subject, to allow such people to participate in Bayesian analysis; once involved, they will naturally see the benefits of incorporation of subjective information, when available and when needed. Another point is that, by necessity, most priors used today are objective rather than subjective. Subjective elicitation is just too hard to be done in even a limited way for more than a few unknowns in a problem, so the vast majority of unknowns in complex problems must be handled via objective Bayesian methods (in which I include hierarchical modeling with objective hyperpriors).

Time for another small advertisement. José Bernardo, Dongchu Sun, Ruoyong Yang and I are writing a book on the subject (finished by next year, we hope) to be called *Objective Bayesian Inference*.

### CONCLUDING QUESTIONS

**Wolpert:** You've given us a snapshot of what Bayesian statistics was like thirty years ago and how it has changed up to the present. Bayesians are supposed to be pretty good at prediction; can you peek into the future and tell how you expect the science will change, in light of the obvious changes we can expect in computation and in the demographics of the community of scientists and statisticians, the more intense recent concerns with security, patterns of change in national funding of science, and whatever else you feel may be important?

**Berger:** Wow, that covers a lot of ground. It is actually something I spend a great deal of time thinking about, in my current role as director of SAMSI—we are supposed to be thinking about precisely these kind of issues as we choose the future research programs at SAMSI. To prevent myself from going on and on, let me mention only two rather random things.

One paradoxical problem is that statistics is becoming more and more central to other sciences and applications, but there are so few of us statisticians to go around that we continually see other fields developing their own “metrics” or “informatics” subfields, which then drift further and further away from “statistics central.” As one way of addressing this, we at SAMSI are working hard (along with other organizations such as NISS) to create partnerships between scientists in other disciplines, statisticians (and often applied mathematicians) all working together on the big problems, but with the statisticians maintaining their connection to “statistics central.” There are many barriers to this, both in the other fields, within statistics, and because of funding structures, but we simply have to try to overcome these barriers.

The second thing I wanted to mention is that, with all the concern about statistics adapting to the massive new data streams coming on line, we should not forget that there are still hugely challenging areas, that statistics has barely penetrated, which involve small amounts of data. For instance, one area I am heavily involved with (through a project at NISS), and that is currently getting huge play across science, engineering and industry, is the analysis of complex computer models of processes. In this domain, very limited data

are typically available and the statistical challenges are enormous.

**Wolpert:** What advice do you have for young statisticians? In our generation we learned more mathematics than statistics Ph.D.'s do today. Is this right? Should they learn more computing? More genetics and genomics? What advice would you give?

**Berger:** I sometimes hear it said that today's Ph.D.'s need more mathematics than we learned in our day—because of the much greater current influence of applied mathematical modeling of processes; lots of core probability and statistics—now including several courses in Bayesian analysis; significant computational and computer science grounding and experience; and major interdisciplinary exposure—for example, in genetics and genomics. Having all our students learn all of this in graduate school is simply not feasible.

Another thing I think is wrong is the notion that any of the above desirable knowledge-bases is easy and can be “filled in” if a student has good grounding in some of the others. I don't think any of these subjects is inherently easier than the others, as they each require a significantly different type of sophisticated thinking.

So that leaves a problem without a solution. All we can do is muddle along as we have been, with Ph.D. advisors and mentors finding out what the Ph.D. student or young statistician's interests are, and finding a path through courses that can lead them to useful research in an acceptable length of time. Lots of learning will simply have to occur “on the job,” later in life.

Of course, I won't sensibly leave it at that, but will add the recommendation that Ph.D.'s and young researchers seriously try to look at their problem from a Bayesian perspective, as well as whatever other approach they are being advised to use. Significant insights can emerge from doing so.

### ACKNOWLEDGMENTS

James Berger and Robert Wolpert thank George Casella for a careful reading and helpful suggestions.

### REFERENCES

- BAYARRI, M. J. and BERGER, J. O. (1999). Quantifying surprise in the data and model verification. In *Bayesian Statistics 6* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.) 53–82. Oxford Univ. Press.
- BAYARRI, M. J. and BERGER, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statist. Sci.* **19** 58–80.
- BERGER, J. O. (1980). *Statistical Decision Theory: Foundations, Concepts and Methods*. Springer, New York.

- BERGER, J. O. (2003). Could Fisher, Jeffreys and Neyman have agreed on testing (with discussion)? *Statist. Sci.* **18** 1–32.
- BERGER, J. O., BOUKAI, B. and WANG, Y. (1997). Unified frequentist and Bayesian testing of a precise hypothesis (with discussion). *Statist. Sci.* **12** 133–160.
- BERGER, J. O., BOUKAI, B. and WANG, Y. (1999). Simultaneous Bayesian–frequentist sequential testing of nested hypotheses. *Biometrika* **86** 79–92.
- BERGER, J. O., BROWN, L. D. and WOLPERT, R. L. (1994). A unified conditional frequentist and Bayesian test for fixed and sequential simple hypothesis testing. *Ann. Statist.* **22** 1787–1807.
- BERGER, J. O. and SELKE, T. (1987). Testing a point null hypothesis: The irreconcilability of  $P$  values and evidence (with discussion). *J. Amer. Statist. Assoc.* **82** 112–133, 135–139.
- BERGER, J. O. and WOLPERT, R. L. (1984). *The Likelihood Principle: A Review, Generalizations, and Statistical Implications*. IMS, Hayward, CA. (With discussion.)
- BERGER, J. O. and WOLPERT, R. L. (1988). *The Likelihood Principle: A Review, Generalizations, and Statistical Implications*, 2nd ed. IMS, Hayward, CA. (With discussion.)
- BIRNBAUM, A. (1962). On the foundations of statistical inference (with discussion). *J. Amer. Statist. Assoc.* **57** 269–326.
- CHUNG, K. L. (1968). *A Course in Probability Theory*. Harcourt, Brace and World, New York. (Second and third editions published 1974 and 2001, Academic Press.)
- DASS, S. C. and BERGER, J. O. (2003). Unified conditional frequentist and Bayesian testing of composite hypotheses. *Scand. J. Statist.* **30** 193–210.
- HWANG, J. T. and CASELLA, G. (1982). Minimax confidence sets for the mean of a multivariate normal distribution. *Ann. Statist.* **10** 868–881.
- ROBINS, J. M., VAN DER VAART, A. and VENTURA, V. (2000). Asymptotic distribution of  $p$ -values in composite null models. *J. Amer. Statist. Assoc.* **95** 1143–1156.