Statistical Science
1991, Vol. 6, No. 2, 149–162

Research — How to Do It: A Panel Discussion

Peter Kempthorne, Nitis Mukhopadhyay, Pranab K. Sen and Shelemyahu Zacks

Abstract. On May 4, 1990, the Students' Seminar Series of the Department of Statistics at the University of Connecticut organized a "Research Panel Discussion." The students' committee consisting of Tumulesh Solanky (Chair), Tai-Ming Lee and Saibal Chattopadhyay had invited Professors Peter Kempthorne, Pranab K. Sen and Shelemyahu Zacks to serve as guest panelists. Professor Nitis Mukhopadhyay served as the moderator. The informal discussion touched upon many interesting aspects of statistical research and education. The varied opinions and comments of these expert panelists were undoubtedly most informative to the audience present that day, and it is hoped that such comments will also prove to be useful in the future for the general audience. What follows is a slightly edited version of the proceedings of that lively panel discussion.

CURRENT AFFILIATIONS

Tumulesh Solanky: The Students' Seminar Series is proud to present a research panel discussion this afternoon. We look forward to listening to three invited experts and their comments on many important issues. Now, I request Nitis to take over and start the proceedings.

Mukhopadhyay: Thank you Tumulesh. This is a very special afternoon session of our Students' Seminar Series. We have three distinguished researchers among us. On my left is Professor Peter Kempthorne from MIT, on my immediate right is Professor Pranab K. Sen from University of North Carolina, Chapel Hill, and Professor Shelly Zacks from SUNY, Binghamton. We are going to talk about various aspects of doing research and how research is done. Hopefully, we will get some pointers toward the end as to how to do it right. I will first turn to each individual and let them introduce themselves and then we will address various issues. Their answers, I am sure, will lead to more questions and answers, and we will take it from there. Peter, do you want to take over and introduce yourself?

Kempthorne: Certainly. As Nitis said I am at the Massachusetts Institute of Technology and I am based at the Sloan School of Management and I have research activities with the International Financial Services Research Center. I am doing research right now on statistical modeling in finance, in particular, looking at stock market data and building stochastic models of transaction prices in stock markets.

Sen: I am from Chapel Hill, North Carolina. I am jointly in the Department of Statistics and Biostatistics. Although I have been involved in biostatistics for about 25 or 26 years, I have kept some interaction with statistics and mathematical statistics. So, I do have interest in both mathematical statistics and biostatistics, and my research areas include multivariate and sequential nonparametric methods, among other things.

Zacks: I have been working in statistics now for over 30 years. I started in mathematics and sociology and then after getting my BA in sociology, I decided that that's not for me. So I continued in mathematics and also studied operations research and statistics. This was in Israel and there was not much formal statistics at that time. As a student, I gave a seminar talk on Analysis of Variance. I was very much influenced at that time by the book of A. Hald, "Statistical Theory with Engineering Applications." After giving that lecture, I got an opportunity to do some analysis of variance for a plant

Peter Kempthorne is Professor, Sloan School of Management, Massachusetts Institute of Technology, E53 311, Cambridge, Massachusetts 02139. Nitis Mukhopadhyay is Professor, Department of Statistics, University of Connecticut, 196 Auditorium Road, U-120, Storrs, Connecticut 06269. Pranab K. Sen is Cary C. Boshamer Professor, Department of Biostatistics and Statistics, University of North Carolina, Chapel Hill, North Carolina 27599-7400. Shelemyahu Zacks is Professor, Department of Mathematical Sciences, State University of New York at Binghamton, Binghamton, New York, 13901.

physiologist, who worked in the Research Council of Israel. He said, "You know, there is a book by Snedecor." He gave me that book and I read the book and started to do analysis of variance for him. That's how I was initiated into statistics. I also worked with an entomologist and organic chemist. This entomologist, Dr. Baer, worked on the problems of mosquitos and malaria. In particular, he studied how mosquitos developed resistance against DDT and different kinds of other pesticides. The chemist was searching for good synergists. As a statistician, I had to do bioassay and it was very interesting work. We developed Markov chain models of the increase in resistance to pesticides from one generation of mosquitos to another. From these experiences, my interest in statistics grew. So I went to Columbia University and studied statistics and then to Stanford, and so on and so forth.

MOST INFLUENTIAL BOOKS AND BOOK WRITING

Mukhopadhyay: Shelly mentioned the book by Hald. Peter, can you think of one or two books which really interested you most to appreciate statistics?

Kempthorne: My favorite book in statistics is the book on linear regression by Seber. Its text really does not have many words. You will find more equations than words in that book, but I thought that it simplified the theory: both statistical theory and probability theory for regression analysis.

Mukhopadhyay: So, Seber's book really worked for you.

Kempthorne: That's really my favorite. Turning towards broader statistical issues, *Theoretical Statistics* by Cox and Hinkley is an excellent book. I don't think it is a good book to learn from, but it is a good book to gain some perspectives on statistical theories.

Mukhopadhyay: Pranabda, would you talk about one or two books that influenced you most?

Sen: When I was an undergraduate or a graduate student in Calcutta, back in the mid fifties, almost 35 years ago, there were not too many textbooks in statistics to follow. The most notable book was Harald Cramér's *Mathematical Methods of Statistics*, and along that line we had, when we were in the graduate school, C. R. Rao's first book, *Advanced Statistical Methods in Biometric Research*. I guess that these are the two books that influenced me the most. Although when I started working towards my Ph.D. degree in Calcutta, I found Fraser's book *Nonparametric Methods in Statistics* to be very useful, because it had a lot of interesting problems which I could examine very thoroughly that early in the career. So these are the three books which I have enjoyed the most in the fifties, when I was a graduate student starting to do research in Calcutta.

Mukhopadhyay: Now, from the topic on books that influenced you most, I come to the following topic. You are also authors of so many different things. I will turn to Shelly first. Your wrote that comprehensive book on statistical inference. I understand that you have authored two books and you are writing several others right now. How did you get involved in such a mammoth project?

Zacks: That was probably a very ambitious thing to do.

Mukhopadhyay: Was it natural to get into that at that stage of your career, Shelly?

Zacks: Well, the book is now almost 20 years old, so I was quite young when I did that. It took me about 10 years to collect the material. I was a young professor, and a salesman from Holden-Day came to my office one day and said, "Why don't you try to publish your lecture notes on decision theory?" I thought for some time and then I sent them an outline of Chapter 1 and they said, "No, that's too theoretical, we don't want that." Then, I was at Stanford and one thing led to another and eventually I signed a contract with John Wiley. That's how I came to write the book.

Mukhopadhyay: So it was a natural progression. Pranabda, I know that you have authored many books.

Sen: Not too many [laugh].

Mukhopadhyay: I know about four or five, at different levels. How do you get involved in such projects?

Sen: Actually, it was Professor S. K. Chatterjee with whom I started working on multivariate nonparametric methods in Calcutta. Then I moved to Berkeley in 1964 and he went over to Lucknow University in Uttar Pradesh [India]. After coming over to Berkeley, I met with Professor Puri during one summer, and at that time we thought that the type of work we were doing was of considerable current interest and it would be better to put the materials in the form of a monograph. In the summer of 1966, we planned the whole thing, and it took us five years to get through it. The volume was out in 1971.

Mukhopadhyay: So, that is the evolutionary process behind the famous Puri and Sen book *Nonparametric Methods in Multivariate Analysis*.

Sen: Then, in the seventies, I was more interested in sequential methods. The work I compiled with Malay Ghosh and some other colleagues of mine during 1971–80 was combined in the second book, *Sequential Nonparametrics* in 1981. And then, of course, you are aware of other collaborations that are now currently underway.

Mukhopadhyay: You have anything under the hat, Peter?

Kempthorne: First, going back to the previous topic, one book that has actually had a great impact on me is the book *The Foundations of Statistics* by L. J. Savage. This is a book which, when I first read it, I didn't appreciate much.

Mukhopadhyay: That's a very hard book to go through and understand in the first pass, I think.

Kempthorne: Originally, I didn't think it had much relevance to me. But over the past several years, I have begun collaborations with a mechanical engineer and together we have worked on the problem of defining parameters from a formal point of view. We asked ourselves: What is a statistical parameter? It turns out that one can use ideas of Savage and DeFinetti. Bayesian ideas provide a formal definition for parameters, and if I ever write a book I think it would be on foundations of Bayesian parametric models attempting to leap off of Savage's book.

Mukhopadhyay: We will wait for such a book to come out then.

Kempthorne: Well, don't hold your breath [laugh].

HOW DOES ONE GET RESEARCH IDEAS?

Mukhopadhyay: Now, I move to address some other thoughts. How do you really get ideas and so on? Let me ask this way. You are all researchers, perhaps sitting in an office or facing students. But then, how do you get a research idea? How do you know whether a research idea is going to click and whether or not to pursue a particular idea?

Kempthorne: I think it is very difficult to anticipate what ideas will work and what ideas will not. I think what is very important in undertaking research is having an open mind about what problems you might work on, and to target those areas that you find interesting. Recently, I have chosen to pursue the area of finance and modeling stock prices. It is a very new area for me, and I think I am making a lot of progress there. I think that successful research results largely from perseverance and having a very "can do" attitude about being able to accomplish something new. Very often the first few approaches you have in an area will not be successful and you may hit brick walls. The trick is to be able to redefine problems so that they are solvable, and not to get frustrated. While my degree of perseverance is perhaps longer than others and there is evidence of my willing to be bored with something for much longer than other people, I find it important to have sufficient time to devote to problems.

Mukhopadhyay: Do you want to add something, Pranabda?

Sen: Well, I think that before you move into aspects of research interest, first of all you need to know what made you move into that line instead of choosing some other career. Why would one go for a research-oriented career or teaching plus research-oriented career? In that context, top priority should be given to different kinds of objectives and one has to be particular that such objectives keep up with one's basic goals in life. In India, where we grew up, many people at that time were going for civil services. They were more attracted by various types of opportunities and administrative control. Sacrificing those types of things at that time and going for a research career for a mere 200 rupees [41 U.S. dollars] per month seemed indeed a big issue and that was associated with much uncertainty. In that context, we had to move from one corner to the other corner and prioritize several career paths which were financially more attractive at that time. I think that there are about 8 or 10 principles which I would like to put forward as tentative ones. The number one is that if somebody is interested in a research-oriented career, then he or she should not take anything for granted. They should try to find a "hole" in almost everything. Because if we aren't able to find holes, then it will be very difficult to do something in research. When you look at a published work such as a textbook, a paper or anything like that, it is very likely that it has something there which is not entirely justified in some sense. That is, there is room for doing something extra, and that is the starting point of doing research. The second item is that one should try to understand how the area of statistics got its early start. Shelly had added that he came from sociology, for instance. The first course I offered at Calcutta University about 29 years ago dealt with statistical methods in biological assays, and I literally followed the book *Statistical Methods in Biological Assays* by D. J. Finney. I found that almost everywhere he assumed a normal distribution for the tolerance problems. My experiences at that time, dealing with certain biological assays, showed that relevant data sets were seldom normal, even after making transformations. Now, that was a starting point. The question I asked myself was, what happens if I challenge the underlying assumption of normality.

Mukhopadhyay: Pranadba, so that's what you meant when you were talking about finding "holes" even in many established theories.

Sen: Yes, that's right. Once you take this attitude, then it opens up new avenues, and when such avenues come up one has to examine the extent to which one can proceed, one by one. I am sure that both Peter and Shelly will emphasize Bayesian or empirical Bayesian methodologies later. So I am reserving such related comments for the time being. There is a joke that we should pose as applied statisticians to mathematicians and we should pose as mathematicians to applied statisticians. That could just be a way out for saving our faces in certain situations, but we must realize that our primary motivations come from applications. We need "theory," but we will truly lose out on everything if we get out of "applications." So, as statisticians we need to develop "theory" which is definitely applicable in specific areas, whether you are interested in physical sciences, biological sciences, finance, stock marketing, engineering or something else. Lately, I find lots of interest in the area of neurophysiology. It is a very vast area, where very little has actually been done thus far in developing appropriate statistical methodologies. When you tap a part of the brain which has some millions of cells, you don't really know how the cells are reacting. The question is how you would approach the situation from a statistical point of view. We have some physiologists in our school doing some interesting work in that area. They were trying to find out how the variation changes with the extent of the given stimuli. So they plotted the intensity of the stimuli versus the standard deviation, and they found a very remarkable linear trend. They fitted a straight line and they found a negative intercept and a positive slope. They published their finding in a paper in the *American Journal of Physiology*. After some time, I looked at that paper and I told them that they had done a marvelous job. However, from a statistical point of view it is worthless, because the variance cannot be negative. And that's the "hole." So, you have to find such holes or gaps whenever you look at anything. From one sense something quite nice may have been done, but yet from statistical sense or otherwise there is possibly hope for improvement. There are other points also which I will perhaps address later on.

Mukhopadhyay: Shelly, would you add something?

Zacks: The best of my papers were motivated by consulting problems. I was asked by scientists to help with certain statistical problems. In 1963, I was approached by a soil engineer. He wanted to estimate the common mean of two populations and he didn't know anything about the variances. But, a priori from his theory he said that the means should be the same, and here are the two samples from two different soils. So I thought about this problem a little bit and I started to investigate. I realized that there is room for innovation and so I wrote a letter to Herman Chernoff. It was a year after I was at Stanford working with him, and he said, "Yeah, that's really what I would do too but there isn't much known here." So I wrote that paper and sent the paper out to the *Annals*. Then, a whole class of papers have been generated from that paper. Recently I was visiting Tokyo and a very nice Japanese gentleman came and gave me his dissertation. All his dissertation was on this problem area. So, you may start your own line of work from something like this! On change point problems, I wrote a paper with Chernoff in 1963, published in 1964 [*Ann. Math. Statist.*]. That paper has been cited in almost all works in this area. And this actually came about from a filtering problem faced by the researchers at the U.S. Navy who came and asked how to track a missile? The problem was to detect as early as possible the "change point," that is the point of time at which the system got out of control. So we started to work on the tracking problem. As a byproduct we studied the change point problem, which was actually of secondary importance, yet everybody caught on that problem. These ventures were interesting, since these problem areas were posed by persons from outside the field of statistics. Sometimes one also discovers a problem while teaching a class. As a teacher you follow a certain textbook. You talk on the material before the students and then you suddenly realize that there is a "problem" here to work on. I am sure all of you are aware of many such experiences.

WHAT IS RESEARCH?

Mukhopadhyay: You face a student and he or she doesn't have as much experience as you do and you perhaps assign a paper or a book to read. At a certain point, you may say, "Go ahead and read so-and-so's papers in the *Annals* or *JASA*," etc. How would you usually advise a particular student to get even started? Often getting started seems to be the most formidable step. When you give a paper from a journal to so-and-so to read, how would his or her thoughts be guided? How should he or she plan the route because he or she may not even know where to begin. What is research? Do you think you can throw some light, Peter?

Kempthorne: I had the privilege of working with David Cox for my masters degree. That explains why I like the book *Theoretical Statistics* by Cox and Hinkley so much, because whenever I had questions about it, I would consult Cox directly.

Mukhopadhyay: That tells you why I like C. R. Rao's book *Linear Statistical Inference*.

Kempthorne: [Laugh] But I think that an excellent approach when thinking about a research problem, or looking at research papers in the area, is perhaps to read the abstract of a paper first, just to understand the problem that is being addressed, and then to spend an hour or two trying to solve the same problem from first principles, formalizing the problem and observing what approaches you would apply in order to solve it. Whenever I have done that, I always found a different way of thinking about the same problem than the author's. Sometimes the author's approach was much better and then I saw how it could effectively solve the problem. But in other instances, I saw that there were possibilities for research that evolved from the way I had thought about approaching the problem. I think that it is very important to try and think for oneself and not be constrained by the approaches of others. It is very difficult as a student, though, because as a student you are looking at the literature, as is.

Mukhopadhyay: The situation is quite difficult because the student has not been exposed to that much at that stage.

Kempthorne: But, while reading your very first paper from any journal, you can start thinking for yourself and attempt to approach the same problem by yourself. I find such experiences very helpful.

Sen: My feeling is that before you give any problem to a student to look at, you look at the student's own background first. What type of problem is likely to be attractive to him or her? I find that it is essential to do that homework first, before assigning any particular piece of work to look at. For example, if I have a student who has an excellent background in a consultation environment or some applied work, then I would rather motivate him or her to carry out some applied work first and gradually bring in detail methodologies to expand the domain. But, if somebody had a very strong theoretical background to begin with, it may be easier to give a theoretical problem to look at, and then try to induce applications by drawing attention to how the unification can be done with the existing theory and your own generalizations of it.

Mukhopadhyay: Shelly, do you usually follow one of these ways or are your tricks somewhat different when you advise students?

Zacks: I often prepare the students to do research through seminars. A seminar is generally offered around a topic, for example, statistical control theory or whatever. Then, generally I make a list of papers to read from various journals which I think would enlighten the students. I provide them with the list at the beginning of the semester. We assign the students to speak about these papers and ask them to add some of their own ideas and thoughts also. Very often, students come and ask me questions before their presentations and we discuss these. So, you see, one student may have better performance than the other, but you try not to create a very formal setup, so that students can help each other also. Generally, the professor will also go to the blackboard and provide more insight whenever appropriate.

Mukhopadhyay: Shelly, your approach is then to get the students actively involved early on.

Zacks: I had many students who wrote dissertations under my supervision. Sometimes you give certain papers to read to an apparently not so bright student, and he comes back after several months with his dissertation practically ready, and you are surprised. And sometimes you have very good students with all the expectations, but they are very insecure. Every time they come to your office, you have to tell them exactly what to do next. They are afraid to take the next step. So, it is very difficult to have a general answer to the question. It varies from one student to another. I try to do the best with what an individual can really do.

HOW DOES A STUDENT'S RESEARCH GET OFF THE GROUND?

Mukhopadhyay: When students begin a research career, they face enormous difficulties. What are your feelings about having something like a formal course or so where one will teach research methods? Have you thought of that?

Zacks: I told you I studied sociology once, and there I studied a methodology of social research. In chemistry, you study the research methodology of chemistry.

Mukhopadhyay: But, in statistics there is almost nothing like that.

Zacks: There is no such thing, perhaps, because we often are trying to mimic the mathematicians, and that's not good. In operations research you learn the methodology of operations research. There are about 10 steps. You have to discuss the problem first, then you have to list the pertinent variables, and how you measure these pertinent variables. Next, think whether it is possible to make those measurements. After all these steps you start to think about relationships between variables and what's important and what's not. Then you write down certain hypotheses, etc., etc. In operations research it is said that a good formulation of a problem is about 50% of the solution. We should be developing a general methodology of research in statistics.

Mukhopadhyay: The question can be rephrased. Can we even dream of doing that?

Sen: Nowadays, in many places there is something along the lines of what you are possibly doing here today. Ask the students to take the initiative in reporting their way of looking at problems via students' seminar series. Make students accustomed to presenting basic ideas in the form of seminars. During such preparation, one may come up with some plausible ways of resolving certain unresolved problems. At Chapel Hill, particularly in statistics and biostatistics departments, we have some advanced courses which are particularly geared towards this direction. That's been very helpful. But still there is a limit beyond which the faculty member may not be able to contribute towards the students' learning. There has to be a seriously developing interaction between the faculty and the students. I think that for the faculty and students to effectively interact the basic requirement is openness from both sides. If that can be instituted, then we would be able to solve this complex problem in a reasonable way.

Mukhopadhyay: What do you think, Peter?

Kempthorne: I would have a hard time formulating a program for teaching people how to undertake research. I partially agree with what Shelly was saying. A difficult part of the problems lies in just defining a problem worth working on. There is a joke that the error of the third kind is testing the wrong hypothesis—a variation would be working on the wrong problem. The trick to finding problems that you can work on lies in addressing some applied areas of interest and just trying to work on certain problems. Very soon thereafter you will encounter worthy statistical problems, which the existing literature has not yet solved. During my masters work, I was given some data on a clinical trial involving breast cancer. I approached the problem of how does one explain the incidence of breast cancer by means of a logistic regression model. I was looking at the problem of selecting important explanatory variables for the logistic regression model. Well, I looked at the literature on variable selection methods for logistic regression. There really wasn't much available. There was a lot more on variable selection for linear regression models but none of these approaches were very compelling. So to make things simpler I posed the problem: How should one select variables in linear regression models? I pursued that research problem for several years. I think that in almost any applied area, if you try and formulate a statistical solution to the problem, you will encounter certain aspects where there is no existing statistical theory providing guidance and you will actually have to develop it. I think that's a great source for problems. In terms of teaching people how to undertake research, I can tell you that an approach which has been very helpful to me is to work with senior faculty on problems. I had the privilege of working with D. R. Cox at Imperial College and I also worked with Jack Kiefer when I was a Ph.D. student at Berkeley. Just hearing the senior professors discuss how they might formulate a problem, sometimes even just a few accounts, were enough to instigate my own thinking in research. I think that being very curious, inquisitive and taking a responsibility to be inquisitive and to come forth with your own suggestions just to get feedback are very important features. From my own experience as a faculty member, I find the back and forth interactions between students and faculty to be very rewarding. As a student, you may see it all coming one way, from faculty. On the other side, I think faculty members enjoy themselves most when the students participate in stimulating discussions and when they push the faculty in new directions. That's the challenge for all the students.

WHEN TO LEAVE A PROJECT AND MOVE ON

Mukhopadhyay: In a certain problem area, suppose that you are working on a particular well-defined project. You start doing the right things and later interesting results develop and so on. When do you know or get the feeling that the work is done? I am asking this because often some of us have difficulties and we keep on refining our results and never get satisfied with the particular product at any particular stage. Do you have any clue as to how you handle this important aspect of research?

Sen: Well, as a matter of fact, you have to look at the work from several angles. You have to consider the novelty or originality of the work, plus technical difficulties and its publishability. If the chances are that it is publishable and it has certain amount of novelty plus certain amount of technicalities which are not trivial, then I would rather go for it and say that it is really the phase one of this particular research. That means you polish it and let it go. If I have doubt on either count, then I would like to see more work to improve one or both of the two requirements, namely, novelty or technicalities.

Zacks: At a certain point of time, to terminate a particular study, often perhaps unwillingly, because research is truly never ending. If you sit on a problem three or four years, you get tired of it. It also becomes somewhat obsolete, and people who needed some answers are not interested in them anymore. You have to somehow finish your studies. For example, if you try to develop some new types of adaptive estimation, you may start with several procedures. You may prove first order efficiency properties very easily. Second order efficiency properties may be a little bit harder. But if you are satisfied with what you have obtained, you publish it. Another question is whether you would continue to do research on the same topic? Quite often I have seen people who actively pursue the third order or even fourth order efficiency of procedures. This involves lengthy calculations. I know that if you sit on a problem long enough, you'll get another expression, followed by yet another expression, etc. Obviously, it is interesting to some individuals but it is not interesting to me. That's why I have worked on many different problems and did research in different fields of probability and statistics. Altogether I have done research in about 11 different fields.

Mukhopadhyay: Shelly, you have always moved from one topic to another. That is right and it is remarkable.

Zacks: I have often returned, perhaps 10 years later, to a previous area of research with new outlooks, by bringing in new techniques from other fields.

Kempthorne: I think it is very hard to leave a problem and to say that's it, that's finished. What is helpful is to look at papers that are published in the journals and to see how much detail they provide on particular problems, to gauge and scout for how extensive the analyses of particular problems need to be. Also, it is very helpful to start writing up the results of some research early on, because very often, when you start writing a paper on a particular problem, you then discover that there is really a lot more to say than you originally thought. Waiting until you think you have finished the problem before starting to write is a mistake, one I have made several times. The Ph.D. thesis is an excellent time to start writing right away, as soon as one has started working on a problem seriously. Based on my own Ph.D. experience and that of directing Ph.D. students, it seems that people often wait too long to start writing up their thesis—until they think they have enough material to comprise a whole thesis. The process of writing typically takes about a year. Even one year is a short period of time and something like two years would be best. If you want to finish a Ph.D. degree program in four years, it means that at the end of your second year you should start writing things down so that you will finish in good time. Writing does take a very long time.

MOST INFLUENTIAL CONTRIBUTIONS

Mukhopadhyay: I don't want to put any one of you on the spot but I have to ask this question. In the last 50 years, what are the top five research areas in statistics that had the most impact in your opinion?

Sen: Sam Kotz and Norman Johnson are editing a volume entitled *Breakthroughs in Statistics, 1890–1989*. That's 100 years. They have included 36 published manuscripts with additional editorial comments. I have been asked to give some detailed editorial comments on one of these papers which is number one in my opinion.

Mukhopadhyay: So, Pranabda, it appears that you are in a very good position to comment on this topic.

Sen: That is the paper by Wassily Hoeffding published in 1948 [*Ann. Math. Statist.*]. The title of that paper is "A Class of Statistics with Asymptotically Normal Distribution." If we look at this paper and realize that, before 1948, people assumed either independent and identically distributed variables or at the least, independent but not necessarily identical random variables, then we see the true novelty. Hoeffding encountered this problem with U-statistics where different terms are not independent. However the symmetric structure led to interesting methodologies including projection techniques. In my opinion this is one of the papers which opened up a tremendous and vast area—not only in statistics but also in L_2 projection in probability theory. I consider that paper as one of the most outstanding papers written in the last 50 years. There are other outstanding papers of course in statistics. This list should include much of the celebrated works of Chernoff, Robbins, and lately Efron on bootstrapping. Since 1979, Efron's paper has definitely created a lot of impact. Somebody was mentioning that about 360 papers have appeared on bootstrapping in the last few years. Don't be surprised if you see pages after pages in the *Annals*, *JASA* and *Biometrics* flooded with papers dealing with different aspects of bootstrapping.

Mukhopadhyay: We are already seeing that.

Sen: So, that's a positive indication of how much impact bootstrapping has. But in my mind, I am still not totally convinced about bootstrapping, because it is not a natural way of sampling. You and I can differ because we can have different bootstrap samples. If we manage to take large enough samples we will converge, but in finite cases the conclusions might not be totally in agreement. If we have non identically distributed random variables from time series models or more complex sampling designs, bootstrapping has yet to come up with good solutions. Yet, the start has been very good and I would regard Efron's paper as a very good paper. Chernoff's sequential design problems also made breakthroughs. And of course how about Robbins' empirical Bayes methodologies of 1956 [*Proc. Third Berkeley Symp.*]? That's a very good paper too. These are the four important areas that come to my mind which have clearly been breakthroughs in the last 50 years.

Mukhopadhyay: Shelly, do you want to add something?

Zacks: Well, obviously the paper of Neyman and Pearson was fundamental.

Mukhopadhyay: That's about 50 years ago, right?

Sen: Almost sixty years ago.

Zacks: The papers of Wald in the late forties were fundamental. He developed sequential analysis and decision theory. For a while, the "in thing" was to go after "robustness." This has subsided, and now the profession is leaning heavily towards bootstrapping. This will pass too. Ten years from now, it will be something else. I always thought that we were not doing enough in control theory. The Russians call it cybernetics. There are problems in that field which are more important to the scientists. The scientists are not interested in bootstrapping. They wonder if, after getting only 100 observations from the field, one can create more knowledge by resampling with a computer a million times from these hundred observations. With this computational technique we are only obtaining the sampling distribution of procedures, when we have very limited knowledge of the phenomenon we are studying.

Mukhopadhyay: You mean that one should perhaps use computing skills to investigate preliminaries first in order to grasp certain complex problems.

Zacks: That's true. Suppose that you have a very complicated procedure. You cannot just go to the blackboard and write the sampling distribution from fundamental principles. So you want to obtain the sampling distribution by bootstrapping. If you are speaking about asymptotics, then it is not enough to have a lot of resampling from small samples to begin with. It has really nothing to offer on the phenomenon itself. In order to do statistical research on control theory like the engineers do control theory, we need a lot of knowledge in stochastic differential equations. Many things that we do in asymptotics are irrelevant in control theory, since we are not resampling from the same population. The population changes dynamically all the time. You are seeing at most two or three observations from one population. The same is true in the stock market. You cannot speak there about asymptotics; it is meaningless.

Mukhopadhyay: Shelly, now you are putting me on the spot. I love asymptotics [laugh].

Zacks: I know, I do it also, since asymptotics give us a lot of insight. I had started to do all kinds of work, for example, on catastrophe theory, dynamical systems, etc. These are difficult areas and one cannot get out papers fast. I hope to do even more especially since we have all this equipment now for computations. We can handle very difficult differential equations now.

Mukhopadhyay: Peter, what do you think?

Kempthorne: They said it all.

Mukhopadhyay: Anything to add?

Kempthorne: I am young in the field and I think it is very difficult to look at the field and identify what the major contributions have been. A few years ago, I recall that the National Science Foundation was putting together a blurb on statistics and they wanted people to say what were the best contributions in statistics in the last 10 years. No one I talked with, including professors from Harvard, MIT and elsewhere, really could come up with the really big things statistics had done. I think that in terms of methodological development, it has already been commented upon. I think that the problem of modeling dynamical systems is one that is really going to be very challenging in the future. Perspectives of statistics during the past 50 years have focused on the notion of static models with some physical constants characterizing systems or populations that we need to estimate. Most systems are changing all the time and the notion of changing parameters leads to trying to estimate many quantities at once. Fortunately, the computational resources are available to throw at such problems and I think that they are necessary to solve them. So I think we will be looking at some very exciting research problems in the future based on elaborating the model set that statisticians typically use, making them dynamical. Along that line, I think that papers on the Kalman filter have been very important; the statistics community has shown considerable interest in this area over the past 10 years. Bayesian statistical methods have become very important in modeling dynamical systems. The Bayesian framework provides a compelling formalism for developing such models. Returning to important contributions, I would like to add those of DeFinetti in the thirties being most fundamental. The series of papers on subjective probability edited by Kyborg and Smokler are very important and include a contribution by DeFinetti.

ROLE OF COMPUTERS

Mukhopadhyay: Shelly has already touched upon the computing aspects. My next question is related to that. We all know that in the last 10 or 15 years, the available computing facilities have really opened up new doors not only for statistics and mathematics, but for all areas in hard sciences. I think that computers will play very important roles in the future and we already notice great impacts. I will start with Peter. What do you think would be the effect of the computer-driven research in contrast to research emphasizing theorems, lemmas or proofs? In other words, what is your impression about this computer-driven research? Where are we going and is it a good area to go into?

Kempthorne: My own research has focused on methods that require very intensive computations. I do know that I could not do the work without the powerful computers that are available. From a classical standpoint, I am involved right now in developing very complicated statistical models with time series data. The models I am working with involve specifying a class of stochastic processes describing a time series which have a few hyperparameters. Implementing classical statistical methods like maximum likelihood involve the numerically intensive operations of calculating a marginal density for the observed data given the hyperparameters of this stochastic process. Evaluating high-dimensional integrals and maximizing functions of such integrals are required. Without computers I couldn't think of solving such problems. But now I am able to solve them, and I crave more computational power all the time. Computers will likely play a very big role in the future. They open up the doors to specifying more realistic models in practical settings. In the past, one of the great limitations of statistics has been a restricted set of simple models for fitting data because anything more complicated was impossible to fit. Now, one can formulate stochastic models that are true to a system and use computers to calibrate the parameters of those models.

Mukhopadhyay: In other words, Peter, you will not miss the "theorem and proof" kind of format that much.

Kempthorne: These new models will lead to theorems. Do maximum likelihood estimates exist for these new models? How can you prove that they exist? You will still need to motivate these models and to demonstrate the "optimality" of computed results.

Mukhopadhyay: So, "proofs" will be computer driven, I suppose. But probably you will not go to that extreme.

Kempthorne: Well, the computer package "Mathematica" is available and it does analytic differentiation, equation solving and other symbolic manipulations. Such tools could be very helpful in "theorem" proving.

Mukhopadhyay: Pranabda, do you want to add anything?

Sen: My point is that no one can deny the benefits of computers, whether in statistics or any other scientific discipline. In medicine and many other areas, I notice that computers have been doing a fantastic job. But the negative aspect is that if we become too much dependent on computers, we may get into the tendency to give up our natural analytical thinking process. Such negativities are being reflected in our high school or even undergraduate curriculum. The students are going away more and more from the mathematical logic and mathematical reasoning. So, the university faculty has the responsibility to make a proper blending of computer reasoning and analytical reasoning. If we give up analytical reasoning, we are giving up a lot actually. I am not opposed to computer-oriented reasonings since that is essential. But, it is also essential to maintain our best resources involving analytical reasoning and then to combine these in a fruitful manner. That's my expectation. We will find out how to do this sort of blending in a few years, I hope.

Mukhopadhyay: Shelly, what are your views?

Zacks: We have to look at the computer as a tool of a scientist or an engineer. A chemist uses a spectrograph to get certain things but he never stops thinking about the "chemistry" involving those molecules and their equations, etc. The chemist or the physicist has a very complicated mathematical theory and from this he predicts if, by constructing a certain experiment, he would observe certain effects. If he can generate that effect, it verifies his theory, and other people should be able to see the effect also. The computer should play a similar role. We have to think first why we are computing something. It is very easy to get reams of printout from the computer and you may not know what to do with the output. The "numbers" may have no meaning.

I work with a very powerful computer all the time. It's my companion on the desk. But, first I have to develop what theory I need, and what I want to compute. I have a recent paper [*J. Statist. Plann. Inference*] where I studied some Bayes sequential procedures for estimating the size of a finite closed population. The procedures are based on sampling-resampling technique. In each stage of the sampling, the sample size can even be random, and then when you write down the Bayes predictive distribution for the next step, you end up with terrible computational problems of determining ratios of expressions whose order of magnitude is about $\exp\{-800\}$. There was a paper of P. R. Freeman in *Biometrika* (1972) on this topic. He used dynamic programming and backward induction, and in fact he actually utilized intensive computations. But, his computations were limited. He had to compute at that time with an IBM 360 system, having certain limited architecture. When you look at these ratios of very small terms, the order of magnitude may be about the same in the numerator and denominator. This is exactly what happened in my case. You may overcome it by inverting a very large polynomial.

Now, you go to basic classical mathematics and open your books on complex analysis and see how to invert those polynomials. Then you do it on the computer to get the results and you obtain actual results. I did it beautifully on the PCs. I didn't need a huge system to compute. But, I had to do the mathematics first. After the mathematical analysis I knew what to compute and how to compute it. That was an important lesson. The conclusion is that, before you go to the computer, start writing your paper and see exactly what you need. Try to solve your problem step by step mathematically before computing numbers.

Kempthorne: I would like to reiterate that there is no substitute for theoretical work to compliment practical work with the computer. In this sense I think that the bootstrap can be quite dangerous because it might be used as a substitute for good thinking about a problem. While it may in fact substitute in a good way, it is very important to have the proper theoretical base before trying to tackle a problem with the computer. An example from my own experience with a simulation study of several model selection procedures in regression was to see which ones were best. The simulation analysis was quite inconclusive. I then went back to the paper and pencil. A part of my Ph.D. thesis was devoted to proving that no model selection procedure was better than another. So there was an analytic way of explaining the computational results. The computational results led to what was to be proven.

A surprising problem in using the computer as a research tool is that when you ask a computer to do very simple things like subtract two numbers, you don't necessarily get the result you want. When you start dealing with computational arithmetic, the difference between two different numbers might be 0 because of the "precision" of the machine. You have to know exactly what you are calculating to be sure that the output is reliable. Some background in numerical methods is very helpful.

START STATISTICS ALL OVER AGAIN?

Mukhopadhyay: Eventually, I want to open the floor for questions from the audience. I will now ask each of you my last question. Think about the process by which you got into statistics: research, teaching and consulting. To begin a career all over again could be quite frightening. But, really if you had to begin your career all over again, with your background, would you come back to statistics? Shelly, do you want to start?

Zacks: That's the most difficult question because so many times I thought if I were in another field, perhaps I could have done better. When I started my study at the University, I didn't even know there was a subject called mathematical statistics. Then I started to mix around with the math students and one student told me that one of the most difficult subjects is statistics and probability.

Mukhopadhyay: They still say that [laugh].

Zacks: Perhaps, then, I should have been scared of it. But then I don't know if it is a combination of random chances, circumstances, constellations or whatever, I was eventually brought into that. But, the influence of professors is important. I was influenced by two or three teachers and possibly because of them I continued the way I did. I am sure that almost everyone here has a similar story about how he or she was influenced by one professor or another. Why didn't they choose chemistry instead? It is very difficult to answer.

Sen: My reaction would be that a lot depends on where you are now and where you were at that time when you first picked this discipline.

Mukhopadhyay: That's why I said the choice all over again could be frightening indeed.

Sen: In America, we face a very depressive picture since the young Americans are not going into graduate school regardless of whether in statistics or any other subject. The final decision to go for a graduate education is questionable now. The graduate schools at American universities are full of people from other countries who are often more serious and may have a completely different outlook. So it depends on whether I would be able to equate my outlook of what I had 25–30 years ago with how I feel now. The decision can be actually quite different. But probably, I would say yes, I will come back to statistics, because after all this is good science, and I have good satisfaction from what I have gotten and this area has good future prospect. So, I am not pessimistic about going back to the start.

Kempthorne: I definitely would. I think that statistics offers the luxury of dabbling in almost any field you like, and where you dabble you are the expert in managing and modeling the data. You can make very significant contributions. I think that the only regret is perhaps not being an expert in another field. If one can do that at the same time, then that would be excellent. I have enjoyed the flexibility the field offers. You aren't constrained to a particular subject matter, apart from having to work with data. It is terrific. If you enjoy solving problems, then it is the best field.

QUESTIONS FROM THE FLOOR

Mr. Bani Mallick (Graduate Student, Department of Statistics, University of Connecticut): I will just add something to the last question. Are the panelists philosophically satisfied by the subject?

Kempthorne: Everyone faces the difficulties of foundational questions in statistics and inference. What is evidence? What is conclusive evidence for some position or some fact of nature? I think that statistics and probability theory provide the formal reasoning for processing information. That's what makes it a very exciting field.

Mukhopadhyay: Shelly, in terms of philosophy, are you satisfied? That is the question.

Zacks: I am talking about statistics. If I were a probabilist, there would be no philosophical problem because probability is a branch of mathematics. In statistics, there are major difficulties in the foundations. That's why if you think a lot about the foundations, you have to become more and more Bayesian, because many frequentist procedures are not really free of inconsistencies. Take for example multiple comparisons. You can have procedures there which might lead you to inconsistent decisions with the same data. Most of the Bayesian methods will not lead you to such inconsistencies. This is one thing. There are also difficulties with the Bayesian procedures. But, if we were satisfied, there will be no more progress.

Mukhopadhyay: That's true.

Zacks: Scientists who are satisfied are also complacent. Look at physics or astronomy. The scientists there are turning over their theories all the time with the advancement of new instruments and new revelations, etc. So no one is really satisfied. I don't know if I answered the question about philosophy.

Sen: Well, actually, satisfaction has two phases. First, are you satisfied with the development made so far? What are you contemplating for the future? In one sense, the development in statistics that Shelly has put in very nicely has been mostly around statistical inference. But that's only one minor aspect. The major aspect which we are facing now is statistical modeling. That comes in almost all areas of science. For example, the ideas of biotechnology or genotoxicity are quite central in any environmental problem. Ultimately it boils down to what type of statistical formulation we can have here, and only then we can think of the statistical inference. So, this modeling aspect is a vast open area in almost any discipline such as sociology, science, technology, medicine, engineering, etc. We have of course great need of these developments. There we are not satisfied with the current status and we need a lot more. But it is a positive outlook, not a negative one. So we are looking forward to moving into that vast area. Only then we can say that statistical science has gone beyond the traditional quarters of statistical inference to become more useful in other fields.

Mukhopadhyay: Any other questions?

Professor Bob Bendel (Biostatistician, Department of Animal Science, University of Connecticut): I was going to ask Professor Sen a question about biotechnology and environmetrics. I think you partially answered it. Can you be a little more specific of some of the areas of current research interests?

Sen: In fact, in the School of Medicine at Chapel Hill, we have some colleagues who have been doing some work with certain types of venereal diseases, and they have come to the conclusion that it is the DNA which is really the main contributing agent for everything. Now in DNA, for example, one deals with only linear mapping, yet when a huge set of data was collected, linearity was nowhere in sight. When I was asked to look into the data, I first thought that perhaps they were not taking into account the proper variables or the variance might not be the way it should have been so that some transformation of variables might be needed. And then, instead of using the classical linear models, some other models might be more appropriate. So, the basic issue in such problems seems to be the identification of what should be a reasonable approach to bring in the chance mechanism in addition to the deterministic factors and then implement statistical analysis that way. We cannot assume independent and identically distributed sampling. It is absolutely meaningless to say that these are all independent. They are not. They are not even marginally identically distributed. But it is possible to combine the deterministic factors and the stochastic factors, and then the basic issue is to utilize some appropriate statistical modeling.

Mukhopadhyay: Any other questions or comments?

Professor Harry Posten (Department of Statistics, University of Connecticut): I enjoyed Nitis putting you on the spot to determine the most important developments in the past 50 years. I wonder, Professor Sen, if you would be able to comment on the studies of interdependence and multivariate analysis, at this time.

Sen: Actually, there are several encouraging facts. First, you look at the statistical multivariate literature during the past 30–40 years. Invariably often one assumed multivariate normal distribution. Then, people started questioning to what extent that assumption was justified. The entire area of multivariate nonparametrics came out of such queries. Now, the nonparametrics can be quite inefficient if normality really holds. Recently, there has been some work on robustness of the normal theory model in nonnormal situations. K. T. Fang and T. W. Anderson have edited a very recent book entitled *Statistical Inference in Elliptically Contoured and Related Distributions*. I think that this volume contains a lot of promise of how you can actually go beyond normality without going necessarily to a full nonparametric family. Now, regarding the dependent structures, if you look from the mathematical point of view, a lot of work has been done. Researchers use various names such as mixing processes, star mixing, ϕ mixing, absolute mixing, regular mixing, etc. All sorts of fancy names are there. But, I think that the current works on some of the time series models are very interesting since these bring in certain more practical types of dependence models. In biometrics, this approach should have a very good impact.

Professor Ashish Gangopadhyay (Department of Mathematics, Boston University): I have a question for Peter. Can you tell us the kind of problems which are statistically interesting in the area of finance?

Kempthorne: If you look at any finance journal, you will see many papers that use linear regression models, and hypothesis testing methods for empirical analyses. If you read any of these papers, you will see applications in statistics where you might have approached problems in other ways. There is a very famous model in finance called the Capital Asset Pricing Model, which is simply a linear regression model relating the stock price of any individual stock and the price of a market portfolio in the stock market. This simple linear regression model motivates a great deal of finance theory and financial economic theory. Lately, people have considered elaborations upon this very simple model and the elaborations have been in the direction of multivariate analysis. Factor analytic models have become very important and a new class of models based on them are called Arbitrage Pricing Theory Models. These models try to explain the covariability of individual stock returns in terms of a small number of underlying factors in the market place. A challenging research problem is how should one actually fit such models? If you have 2000 stocks, then you have a 2000 by 2000 covariance matrix that needs estimating. How do you estimate such a covariance matrix? Is it stable over time? If it is not stable over time, then what is to be done? It is sort of a Pandora's Box of problems. But if you are a statistician, then you have a tool box for working on such problems. Right now, I am working on modeling local movements in stock prices, that is, I am looking at transaction-by-transaction activity in a particular stock. We are trying to model the order-flow process of buy-orders and sell-orders and predict the relative likelihoods of different kinds of orders coming to the market place. Many statistical problems arise in this setting.

Gangopadhyay: Do you find any long range stability at all, Peter?

Kempthorne: Absolutely. In working with models that are fitted to a day's worth of transactions data, we are concerned whether the parameters that are estimated for one day are good for another day. In other words, we ask, are the transaction level parameters stable in the long term, across several days? There is some evidence that they are.

Empirical analysis is really important to motivating my research. One basically hypothesizes some theory about how data is generated, then looks at the compatibility of observed data with that model through computational work. When the model doesn't fit very well, one tries to elaborate upon the original model in the "direction" in which the data is varying. The challenge of defining new models is very stimulating.

Mukhopadhyay: I have a follow up on the questions Ashish just asked. As you know, many researchers from other disciplines find statistics very hard. I guess that many statisticians will find finance very hard. Would it be unrealistic for students in statistics to pick up a finance journal and read papers and borrow ideas from them? Roughly, what kind of training is essential, Peter, for casual readers of journals in finance.

Kempthorne: As in any field, many papers don't start from the beginning and they are based upon an extensive literature. The field of finance does have a number of good textbooks which give you the very basics on the various perspectives people are using. Browsing through such texts first would be better than just diving into the finance literature.

Mukhopadhyay: That means, Peter, you are talking about some serious investment. You can't just pick a journal and all of a sudden dive into "finance" and start writing statistical things related to "finance." That would be tough. Any other questions or comments from the floor?

Bani Mallick (Graduate Student, Department of Statistics, University of Connecticut): Which field of statistics will influence most in the next 5 or 10 years?

Mukhopadhyay: This is more on prediction. Would any panelists like to address that?

Kempthorne: I think stochastic models and Bayesian analysis of stochastic models will play key roles. Things are changing all the time. You want to model "changes" in stochastic processes appropriately while collecting data and develop associated statistical theory appropriate to such models.

Zacks: Filtering, not only linear filters but also nonlinear filters, will have great impacts. The engineer is to some extent constrained there in terms of what is available in hardware when he actually puts a network together. He builds a filter. He verifies its performance, etc. But the technologies are also changing; for example, many of the possible nonlinear filters in engineering are not built because they may be infeasible technologically. So if we are doing things in that area, probably we can learn a lot from the electrical engineers. But many things that we are doing today have possibly been done by some engineers 20 years ago. They are far ahead in this area. We really have to catch up with the game.

Sen: I may add that I forgot to mention about one other paper by David Cox published in 1972 [*J. Roy. Statist. Soc. Ser. B*]. Now, this paper was a very readable one but what is more important is that it generated a lot of interaction among researchers during the past 18 years. The latest development has been to bring in the counting process approach in the Cox regression model. The European schools really have contributed a lot in these areas in the last 10 years. Starting with this breakthrough in 1972 and the recent follow up works by the Danish and the Dutch school, I think that it is one area which will remain very useful in biometrics as well as in many engineering disciplines. In this context, we need developments in stochastic processes, linear and nonlinear filtering, potential theory and of course a lot of other types of methodology. But essentially this type of modeling will prove to be the basic criterion which will guide us to subsequent developments. So I am quite positive about the biometrics area for the next five years as a whole.

Kempthorne: I would like to second the comment that the paper by Cox has been very important. It stresses the importance of censored data. As our ability to collect data increases, the nature of the data will very often be censored or have a truncated form. How do you work with that? Models like Cox's are needed.

Mukhopadhyay: The proceedings have been very enjoyable and informative. I think that with all these marvelous thoughts, we will stop here. I wish that we could continue the discussions. Unfortunately, however, we have to stop somewhere. I sincerely thank Shelly, Pranabda and Peter very much for participating in our research panel discussion. I thank you all for listening.

ACKNOWLEDGMENTS

The administrative assistant, Mrs. Cathy Randazzo, spent many hours transcribing the 85 minute audiotape, and only then could the editing start. The authors thank the Students' Seminar Series Committee for its initiatives and Cathy for her invaluable service. They are grateful to the Center for Environmental Health at the University of Connecticut for providing the necessary funds to organize the panel discussion. They also thank the Executive Editor, Professor Carl Morris, and a referee for several valuable comments.

REFERENCES

CHERNOFF, H. (1972). *Sequential Analysis and Optimal Design.* SIAM, Philadelphia.

CHERNOFF, H. and ZACKS, S. (1964). Estimating the current mean of a normal distribution which is subject to changes in time. *Ann. Math. Statist.* **35** 999–1028.

COX, D. R. (1972). Regression models and life tables. *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

COX, D. R. (1972). Regression models and life tables. *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

EFRON, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans.* SIAM, Philadelphia.

FANG, K. T. and ANDERSON, T. W. (1990). *Statistical Inference in Elliptically Contoured and Related Distributions.* Alberton, New York.

FINNEY, D. J. (1964). *Statistical Methods in Biological Assay,* 2nd ed. Griffin, London.

FRASER, D. A. S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

FREEMAN, P. R. (1972). Sequential estimation of the size of the population. *Biometrika* **59** 9–17.

HALD, A. (1952). *Statistical Theory with Engineering Applications.* Wiley, New York.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 546–557.

LAI, T. L. and SIEGMUND, D. (1986). The contributions of Herbert Robbins to mathematical statistics. *Statist. Sci.* **1** 276–284.

NEYMAN, J. and PEARSON, E. S. (1933). The testing of statistical hypotheses in relation to probabilities a priori. *Proc. Cambridge Philos. Soc.* **24** 492–510.

PURI, M. L. and SEN, P. K. (1971). *Nonparametrics Methods in Multivariate Analysis.* Wiley, New York.

RAO, C. R. (1952). *Advanced Statistical Methods in Biometric Research*. Wiley, New York.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.

ROBBINS, H. (1956). An empirical Bayes approach to statistics. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 157–163. Univ. California Press, Berkeley.

SAVAGE, L. J. (1954). *The Foundations of Statistics*. Wiley, New York.

SEBER, G. A. F. (1980). *The Linear Hypothesis: A General Theory*, 2nd ed. Griffin, London.

SEN, P. K. (1981). *Sequential Nonparametrics*. Wiley, New York.

SNEDECOR, G. W. and COCHRAN, W. G. (1937). *Statistical Methods*, 1st ed. Iowa State Univ. Press, Ames, Ia.

WALD, A. (1947). *Sequential Analysis*. Wiley, New York.

WALD, A. (1950). *Statistical Decision Functions*. Wiley, New York.

ZACKS, S. (1971). *The Theory of Statistical Inference*. Wiley, New York.

ZACKS, S., PEREIRA, C. A. and LEITY, J. G. (1990). Bayes sequential estimation of the size of a finite population. *J. Statist. Plann. Inference* **25** 363–380.