Statistical Science
1993, Vol. 8, No. 3, 331–341

Mentors and Early Collaborators: Reminiscences from the Years 1940–1956 with an Epilogue

E. L. Lehmann

Abstract. These reminiscences extend from the year 1940, in which I arrived in Berkeley, to 1956, the year in which Neyman resigned from the chairmanship of the Berkeley Statistics Department and handed its leadership over to the next generation. They sketch my experiences with six scientists who have influenced me as mentors or collaborators: Evans, Neyman, Wald, Scheffé, Stein and Hodges. The origin of these recollections was a conversation with Agnes Herzberg and Persi Diaconis, which was videotaped under the sponsorship of Pfizer Central Research and the American Statistical Association on April 28, 1992 by the Department of Statistics at the University of Connecticut under the direction of Harry Posten. Although the conversation went in a somewhat different direction and the overlap is moderate, it got me thinking about the people who influenced me in my 20's and 30's and thus led to the present paper.

1. INTRODUCTION

It was exactly fifty years ago that I first heard of statistics as a possible subject of study, and at the urging of Professor Griffith Evans agreed to give it a try. However, the applied nature of the material did not appeal to me. After a semester, I decided to return to pure mathematics but just then was offered a teaching job in the statistics program that I could not refuse. Gradually, I got to like the work and even found the relevance of probabilistic and statistical ideas to so many different aspects of the real world of great interest. So when recently at a meeting with a group of students at the University of Connecticut I was asked to assess my career, my heartfelt response was: "It turned out better than expected."

That it turned out well was partly due to the situation of the field at the time I entered it. The seminal work of Fisher and Neyman-Pearson had laid the foundations; building a superstructure on this basis was a much easier task for which students of Neyman were well prepared. But there was another factor. This was the cohesiveness of the developing profession, the friendship, collegiality and mutual support that the small but growing group of theoretical statisticians extended to one another. The Institute of Mathematical Statistics, which was my scientific home, had fewer than 500 members when I joined it in 1943; *The Annals of Mathematical Statistics*, which published many of my papers and on the editorial board of which I served for many years, in 1943 ran to about 450 pages. As a result, the atmosphere was encouraging rather than intimidating, and many opportunities were available even to a beginner.

During my career, I had the good fortune of enjoying the acquaintance and friendship of many remarkable people who shared their ideas with me and gave me their support. In the early stages, it was my teachers who influenced me; later, I learned from my coworkers; and finally, more recently, it was my students who struck out in new directions and required me to follow them as well as I could. The scientific contributions of many of the figures whom I encountered in the early part of my career have become part of the history of our subject. It therefore seemed that it might be of interest to record the following personal recollections.

E. L. Lehmann is Professor Emeritus, Department of Statistics, University of California at Berkeley, 367 Evans Hall, Berkeley, California 94720.

2. GRIFFITH C. EVANS (1887–1973)

Evans, one of the most distinguished American mathematicians of his generation, made major contributions to functional analysis (the subject of his Colloquium Lectures to the American Mathematical Society in 1916) and potential theory. In recognition of his work, he received many honors, including membership in the National Academy, the American Academy of Arts and Sciences and the American Philosophical Society. He was on the Faculty of Rice University when in 1934 he was asked by the University of California to take over and revitalize the moribund Berkeley Mathematics Department. He was given some new positions and was able to persuade some of the existing faculty to retire. As a result, it was possible for him to hire a group of active young mathematicians: Alfred Foster, Derrick Lehmer, Hans Lewy, Charles Morrey, Anthony Morse and Raphael Robinson. (For a more detailed account of this early history of the Berkeley Mathematics Department, see Rider, 1989. An account of Evans' work can be found in Morrey, 1983.)

Evans, one of whose fields of specialization was mathematical economics, took a broad view of mathematics. He had become interested in statistics through discussions in the summer of 1931 with R. A. Fisher, and "as early as 1935 [had] envisaged California as the place for a really outstanding statistician, if possible of the level of R. A. Fisher himself" (Reid, 1982, p. 142). However, Fisher's visit to Berkeley in 1936 to give the prestigious Hitchcock lectures was not a success, and Evans' choice eventually fell on the Polish statistician Jerzy Neyman.

Neyman who by that time had done (partly in collaboration with E. S. Pearson) his fundamental work on hypothesis testing, confidence intervals and survey sampling, came to Evans' attention through the series of lectures and conferences he gave in Washington, D.C. in 1937. After considerable hesitation and some negotiations, Neyman accepted Evans' offer of a Professorship in the Berkeley Mathematics Department, and at age 44 started work in Berkeley in 1938.

I first met Evans in late December 1940 when on my arrival in Berkeley I presented myself at the office of the Mathematics Department (the staff at that time consisted of a half-time secretary) to see whether I could enroll as a student. Although I had no appointment, I was ushered in to see Evans, a gentle, soft-spoken man who not only on the spot accepted me as a student but suggested that I should start, on a probationary basis, as a graduate student since he thought my studies in Europe had given me the equivalent of an American B.A.

I thus greatly benefited from Evans' informal and personal approach to administrative problems, of which the following story is another example. In his first year at Berkeley, Hans Lewy, who tended to work late into the night, had a morning class for which he sometimes came late and for which occasionally he overslept altogether. One morning when this had happened again, Lewy was awakened by the ringing of his doorbell. In his pajamas and still half asleep, he opened the door to find his chairman, Evans, standing there to remind him that his class was waiting for him. I am told that there was no need for Evans to intervene again.

After my first term had gone satisfactorily, Evans offered me a teaching assistantship. The seven or eight TA's of the Mathematics Department at that time taught their own courses independently, and this included the making up and correcting of the examinations and the assigning of grades. My first assignment was a course in analytic geometry, an ideal choice. It is more advanced and had better students than remedial algebra or trigonometry, which are beset by pedagogical problems; on the other hand, it is a much easier subject to teach than calculus. I believe that Evans chose the course deliberately to make it easy for a student who was new to the program and to the country. It is only one example of his thoughtfulness and his concern for me and the other graduate students. He cared for us and was always available for advice and help when we needed it.

In the summer of 1942, half a year after Pearl Harbor and Germany's declaration of war against the U.S., Evans expressed to me his feelings that I could be more useful to the war effort if I left pure mathematics to work in either physics or statistics. Since I greatly respected his judgment and had a very negative attitude toward physics (about which he knew and which he ascribed to a mental block), I agreed to try statistics, a subject of which I had never heard before. He told me to see Professor Neyman, whom I had not met. Thus, as a result of Evans' intervention, I began in the Fall of 1942 what would become a lifelong career in statistics.

Soon after this switch, I became part of Neyman's group and no longer had much contact with Evans. From these later years, only one memory of him stands out. The story is characteristic both of his originality and his self-effacing modesty. When a new building was planned in 1969 which would house the mathematics and statistics departments and their libraries, it was decided to name it in honor of Evans. Even before the cornerstone was laid, a painting of Evans by Erle Loran was unveiled which was intended for the lobby of the new building and which now hangs in the mathematics–statistics library. The unveiling and naming of the building was celebrated by a dinner at which Morrey and others spoke of Evans' accomplishments. When he rose to reply, his opening sentence was rather surprising: "Who was Bacon?" He then proceeded to go through the Bacons listed in the encyclopedia and concluded his short talk by returning to his original question, now slightly amplified: "But who was the Bacon for whom Bacon Hall was named?"[1]

3. JERZY NEYMAN (1894–1981)

Neyman was 44 when in 1938 he accepted the offer to start a statistics program in the Berkeley Mathematics

[1] Bacon Hall was a building on the Berkeley Campus named after Henry Douglas Bacon (1813–1893), a businessman primarily in banking and real estate. He gave his art collection and large library to the University of California, together with some money for a building which later became Bacon Hall (Slutes, 1947).

Department. In Poland, he had headed a small statistical laboratory; in London, he had been a member of Egon Pearson's Department of Statistics at University College. Now he quickly recruited a small staff of assistants from the students in his graduate course to help with the laboratories that were part of his courses and with some of the applied work he was undertaking for faculty members in other departments.

When in 1942 I accepted Evans' suggestion to study statistics, the statistics program consisted essentially of three courses: a one-semester lower division course, and a one-year course each at the upper division and graduate level. In addition, there was a graduate course in probability theory. To take over the teaching of the latter, Neyman had in the Fall of 1942 obtained the services of a young mathematician, Dorothy Bernstein, who was also giving the lower division course.

A more ambitious development of courses and faculty to teach them had to wait until the end of the war but eventually led to the creation during the decade 1946–1956 of the Berkeley Statistics Department, which became one of Neyman's principal American achievements. For many years, it was the leading department of theoretical statistics in the county. Its curriculum set the standards that were followed by many others; it trained hundreds of Ph.D. students from all over the world. In addition, the Berkeley Symposia on Mathematical Statistics and Probability, international meetings that Neyman organized at 5-year intervals from 1945 to 1970, made Berkeley an international center of the first magnitude.

In accordance with Evans' suggestion, I enrolled in Neyman's upper division course but after one semester decided that I did not like statistics. It seemed messy, and the assumptions often appeared to be quite arbitrary. However, before I was able to tell Evans and Neyman of my intention to return to pure mathematics, something happened that changed my mind. Dorothy Bernstein had come to the same conclusion as I and had requested to be released from her contract. Caught in a bind, Neyman asked me whether I would like to take over some of her duties; I was to receive a promotion and a corresponding increase in salary. The offer was too good to refuse.

With this change of status, I became a member of Neyman's group, the small group of budding statisticians who assisted Neyman in the various aspects of his work. This meant in particular the ever present possibility of being drafted for some of the tasks required by Neyman's activities, such as the Symposia and his applied projects. An exciting, but rather scary, example occurred very quickly when Neyman told me that he was leaving for three weeks and wanted me to take over the lectures of the upper division course (my first statistics course!) in which I was a student. He outlined the material to be covered, suggested Uspensky's (1937) book on probability theory as a reference, and then I was on my own.

It turned out not to be too difficult since the course provided only an introduction to probability theory and least squares estimation, roughly corresponding to the first 14 chapters of F. N. David's book *Probability Theory for Statistical Methods* (1949). In hindsight, it is surprising that it contained nothing about testing or confidence intervals. These topics, which Neyman must have considered as still too nonstandard for inclusion in the undergraduate program, were treated along the lines of his own work in the graduate course.

Neyman had an interesting way of teaching. He liked to call students to the board, giving preference to the women in the class ("Ladies first"), and would try to get them to derive the new results under his guidance, which often would essentially turn into dictation. It was somewhat of an ordeal for the hapless victim, and I recall Neyman once coming to class saying that he had just received a letter concerning this practice accusing him of sadism. He clearly was at a complete loss – "Sadism?" – but announced that any student feeling this way could be excused from coming to the board. To the best of my knowledge, no one took advantage of this offer.

In 1944, Neyman recommended me for the Operations Analysis Section that was being formed by the 20th Air Force. Thus, I spent the year 1944–1945 on Guam, where I found myself sharing a tent with my fellow student Joe Hodges and with the statistician George Nicholson.

I returned to Berkeley just in time to participate in the Symposium on Mathematical Statistics and Probability that Neyman organized in 1945 to "mark the end of the war and to stimulate the return to theoretical research." It was a grand event, which gave us graduate students a chance to meet some of the great names in the field. I drew a particularly nice assignment: to be the driver of the probabilists Feller and Doob. I recall a memorable drive to Stanford in which my two passengers entertained each other and me by playing various games such as asking how big a salary it would take for each of them to accept a position in Berkeley. The figure went up and down with the quality of the view and the surroundings. When we reached the Pacific, Feller lowered his figure substantially, but in response Doob raised his since he might be tempted to swim in the ocean and could drown.

The symposium was such a success that Neyman soon started organizing a second one, which took place in 1950. From then on, the Berkeley Symposia became a regular event every fifth year, the sixth and last being held in 1970. In addition, there were many visitors who came to give summer courses or to substitute for a faculty member on leave. This gave us a chance to get to know such luminaries as Cramér, H. B. Mann, Cochran, Wald, the mathematicians Besicovich and Bochner, and others.

Pleasing his guests was an avocation for Neyman; his hospitality had an international reputation. One of the toasts he would propose at all social events was "To the international intellectual community!" I now realize—although I didn't at the time—how close to Neyman's heart this was. Like many emigrants, he was not completely at home anywhere; the international intellectual community had become his true home, and the toast celebrated this fact.

During the year 1945–1946, I wrote my thesis (on a problem in the Neyman–Pearson theory suggested to me by P. L. Hsu), and the next academic year was appointed Instructor, the first regular faculty member in statistics after Neyman. In the succeeding years, faculty appointments were extended to other students as they completed their degrees: Hodges, Barankin, Scott, Fix, Le Cam. Retention of Berkeley Ph.D.'s on the faculty generally was frowned upon, but an exception was made for Neyman since no one else was providing the training he wanted for his faculty. He did, however, gradually also make some important outside appointments: Stein, Loève, Scheffé and Blackwell, who strengthened and broadened the group enormously.

The year 1947 brought Neyman two great victories. From the beginning, he had envisaged an independent statistics department, separate from the mathematics department. Another aim was a journal under his control that would provide a means of publication for himself and his students and associates. He was in a strong position to confront the administration with these demands when he received an offer from Abraham Wald to join him in the new statistics department that Wald was then forming at Columbia.

Evans adamantly opposed a separate department since he believed in a "greater mathematics department" that would include all the mathematical sciences. A compromise was now reached that left statistics within the mathematics department but with a separate budget that no longer required Evans' approval, and with the right to make its own research appointments although Evans would still have a say on teaching appointments.

A journal of his own had been of great importance to Neyman since in England he had seen the efforts of Karl Pearson to block publication of Fisher's papers, and even more so after the painful experience of having his own fundamental paper on confidence intervals rejected for *Biometrika* by his friend and close collaborator Egon Pearson. It turned out that a journal was not feasible, but Neyman was given something nearly as good: the series *University of California Publications in Statistics*, which for many years stayed essentially under his control.

Neyman's dislike of editorial restrictions was illustrated by an occasion which also brought out his somewhat puckish sense of humor. While working on the proofs of the second edition of his *Lectures and Conferences on Mathematical Statistics* (1952), a very engaging book which is still enjoyable today, he wanted to add an insert of several pages. Since this would have violated the editorial instructions, he cut the proofsheet in question and glued the insert onto the upper and lower part of the sheet, folding it so that when opened it flowed accordion-like to the floor. Since I was going to Washington at the time, he asked me to deliver these proofs to Ed Deming, who was in charge of the project. I left them at Deming's office without waiting for his reaction.

Neyman's wish for a completely independent Department of Statistics had to wait a few more years. When Evans retired in 1954, his successor Charles Morrey had no desire to keep a substantial group of statisticians in the Mathematics Department against their wishes and recommended the creation of a separate statistics department. Thus, in 1955 Neyman finally obtained his own department which consisted of seven tenured faculty members (Barankin, Blackwell, Hodges, Lehmann, Loève, Neyman, Scheffé), three tenure-track assistant professors (Fix, Le Cam, Scott) and several lecturers and visitors. There was a substantial course program including at the graduate level such specialized courses as "Stochastic Processes," "Nonparametric Inference," "Decision Theory," "Experimental Design" and "Large-Sample Theory."

In achieving this expansion in the short span of ten years, Neyman was helped by the fact that in this, as in other campaigns, he was always convinced of the righteousness of his cause. This justified for him the relentless pursuit of his goal. Each compromise reached, each partial success, provided the basis from which to launch his next demand. No wonder that one of his Deans was once heard to complain that he would rather grant whatever Neyman wanted than have to discuss it with him.

When the Berkeley Chancellor finally recommended to President Sproul the creation of a Department of Statistics, Sproul asked an assistant to summarize the history of the issue. As part of his summary, the assistant wrote: "Here a willful, persistent and distinguished director has succeeded, step by step, . . . against the original wish of his department chairman and dean in converting a small 'laboratory' or institute into, in terms of numbers of students taught, an enormously expensive unit; and he then argues that the unit should be renamed a "department" because no additional expense will be incurred. . . . How, in the future, can a bureau, laboratory, or institute director be restrained from enlarging his empire against the judgement of his administrative superiors?" (quoted from Reid, 1982, p. 239).

What Sproul's assistant left out of the account, but which provided the basis for Neyman's vision, was the enormous developing growth of statistics which would soon make it into an extensive scientific field of worldwide importance. By establishing his program immediately after the war, at the beginning of this development, Neyman was able to obtain for Berkeley a leading position that it was to maintain for many years. The particular charge of inordinate expense in terms of the numbers of students taught, although true at the time it was made, lost its validity as gradually the new department took over the teaching of all lower division statistics courses and as a result soon regularly taught statistics to almost 5000 students a year.

The creation of the Department was followed by an unexpected development. Disputes between Neyman and some of his faculty (including myself) that had been simmering for some time now broke into the open – disputes which he himself later characterized as father–son conflicts. Neyman rightly considered the Department his creation. He wanted to continue to run it as a family enterprise of which he was the father and head, a benevolent dictator who would consult with other members but who in the end had sole responsibility and – with the interests of the group in mind – would make the final decisions.

However, even Neyman's own students on the faculty were by now in their mid and later 30's, had national and international reputations, and had their own ideas about appointments and courses which did not always agree with his. He once told me about his reaction to this situation. There is a bird, he said: if someone touches the eggs in its nest, it will no longer have anything to do with them. Feeling the same way, he took a step that came as as complete surprise to all of us. Only a year after the creation of the department for which he had worked so hard, he resigned as its chair and resisted all our efforts and those of the administration to change his mind. For himself he kept the directorship of the statistical laboratory with a small staff, and the university administration agreed to his request that during his lifetime the laboratory would be independent of the Department.

4. ABRAHAM WALD (1902–1950)

The structure that formed the framework for my work in theoretical statistics is due primarily to Jerzy Neyman and Abraham Wald. Although Wald's ideas, unlike Neyman's, came to me not directly from him but via his student Charles Stein, I did have some personal contacts with Wald during the short period between 1947, when I first met him, and his untimely death in an airplane accident three years later.

Wald made revolutionary contributions to sequential analysis, large-sample theory and decision theory. Of these, the last was most relevant to my own interests. In fact, it was in connection with this work that I first met him in 1947. The *Annals* of that year contained a short paper of mine in which I introduced in terms of a very simple example – testing a simple hypothesis in a one-parameter exponential family – a formulation of the testing problem that differed from the traditional Neyman–Pearson approach (Lehmann, 1947). Instead of trying to obtain a single optimum test, I pursued a more modest aim and worked out the smallest family of all possible candidates for this honor, suggesting that beyond that applied considerations would be needed to narrow down the choice. When shortly thereafter I happened to be in New York, I was informed that Wald wanted to meet me. He then told me that he had found my idea applicable to the general decision theory he was developing. He called such families minimal complete and under very general assumptions showed that they consisted exactly of all Bayes procedures together with certain of their limits.

During 1948–1949, Wald spent part of a sabbatical leave in Berkeley, giving a summer course and completing his book on statistical decision theory. Toward the end of that year, Joe Hodges, Charles Stein and I had planned a four-day hiking trip in Yosemite, and we asked Wald whether he would like to join us. Although he was considerably older than we and the hiking promised to be fairly strenuous (the first day we had to cover a distance of 20 miles and climb close to 6000 feet), he was game and turned out to be a good sport and fun to be with. We talked a great deal about statistics, but he was also very much interested in our surroundings. He particularly loved to estimate distances, speeds, heights, etc. I recall how at the bottom of a mountain we were about to climb, he estimated the rise to be 1500 feet. When we arrived at the top he looked down and declared himself satisfied: "Actually it was 1450 feet," he proclaimed, "I wasn't too far off."

Despite his accomplishments and fame, Wald was completely unpretentious, easy going and good natured. He was far less mercurial and competitive than his collaborator Jack Wolfowitz. (Since in statistics it has been traditional to list joint authors alphabetically, it was a standard joke in the profession that Wald had searched a long time before finding a suitable collaborator.) Mathematicians are often classified as being either problem solvers or system builders. Wald excelled at both.

Since I wanted to take a leave from Berkeley during 1950–1951, Wald arranged a visiting appointment at Columbia for me for the Fall semester; I had already made arrangements to spend the spring in Princeton. With Ted Anderson, Howard Levene, Henry Scheffé and Wald and Wolfowitz on the staff, Columbia was an interesting place – together with Berkeley, the strongest group in theoretical statistics in the country. Toward the end of the semester, Wald and his wife left for India, where he was to lecture on his new decision theory. Shortly before Christmas, we began to hear rumors of an airplane accident in which he might have been involved; it took two or three days before the Indian Government confirmed that indeed both Wald and his wife had died in the crash of a local plane that was to have taken them to Nepal. (For further information on Abraham Wald, see Wolfowitz, 1952.)

The death of Wald at age 48 was not only a great personal and scientific loss but also spelled the end of the group he was still in the process of building. Within a few years, Wolfowitz had moved to Cornell, Anderson to Stanford and Scheffé to Berkeley. It was a blow from which the Department still has not recovered, more than 40 years later.

One of the many problems caused by Wald's death was the fate of the students who had been working with him or were planning to do so. Two of the latter group approached me to ask whether I would take them on. So it came about that two Columbia Ph.D.'s, Allan Birnbaum and Jack Laderman, obtained their degrees under the supervision of a Berkeley faculty member.

It has been a great privilege to have known both Neyman and Wald. I only regret that I never met the third of the great founders of modern statistical theory, R. A. Fisher.

5. HENRY SCHEFFÉ (1907–1977)

Besides Evans, Neyman and (indirectly) Wald, the people who had the greatest influence on my development and career were three colleagues with whom I began to collaborate in the years immediately after completing my Ph.D.

The first of these was Henry Scheffé. Having started in analysis with a thesis on asymptotic solutions of certain differential equations, he had switched fields in his mid 30's since he believed statistics to be a more promising area of research. In 1941, to learn the new subject, he joined the famous group of students and associates (containing among others Ted Anderson, Bill Cochran, Fred Mosteller and John Tukey) whom Wilks had gathered at Princeton.

I first met Scheffé in 1946, when he came to spend a year at Berkeley on a Guggenheim Fellowship. Despite the image of a prizefighter conjured up by his picture, Henry turned out to be a very nonbelligerent, sensitive, rather shy person with interests in art, music and literature. He had worked on hypothesis testing problems similar to those considered in my thesis, and his papers had constituted some of my principal references. So we had many common interests and soon became friends.

Although we had dealt with very similar problems, Henry had used the original approach of Neyman and Pearson, while I had employed a method due to P. L. Hsu, who had given me my thesis problem. Discussing these different approaches to the problem of similar regions (i.e., of characterizing the totality of similar tests), we came to understand the common feature that lay behind both methods. It was the existence of sufficient statistics T having a property that we called completeness. For all cases in which such a statistic T exists, we then had a complete solution of the problem of similar regions. The result seemed exciting not only to us but also to Neyman, who arranged for quick publication of our results in the *Proceedings of the National Academy of Sciences* for 1947.

At the end of the year, Henry left for UCLA, from where the following year he moved to join Wald's Department at Columbia. In the meantime, our joint work continued by correspondence and through occasional visits of mine to Los Angeles and New York and of Henry's to Berkeley. An impetus for a fuller development of our completeness concept came from the realization that it played a crucial role also in the theory of unbiased estimation, where the existence of a complete sufficient statistic T assures that any function of the parameters that has an unbiased estimator has one with uniformly minimum variance, namely the unique unbiased estimator that is a function of T. This is an immediate consequence of the Rao-Blackwell theorem.

Since completeness has come to play such a central role in mathematical statistics, it should be pointed out that we deserve only limited credit for this concept. The basic idea was not original with us but was contained in the applications made in testing by Neyman-Pearson and more explicitly by Hsu, and in unbiased estimation in work by Blackwell, Halmos, Rao and Wolfowitz. Our contribution was to isolate the common feature of these various applications and to provide it with an identity by giving it a name and investigating its properties.

In view of its importance, we decided to explore the properties of the completeness concept (and some of its variants such as bounded and strong completeness) rather fully. The result was a massive two-part paper that we published in the Indian statistical journal, *Sankhyā* (Lehmann and Scheffé, 1950/1955/1956), which we thought would give us the space needed for the detailed discussion and the many examples we wanted to present. The paper also introduced the concept of minimal sufficient statistic, minimality being a necessary condition for completeness. (We found later that a theory of minimal sufficient statistics was being developed at the same time by Dynkin in the Soviet Union.)

In 1953, Neyman brought Henry to Berkeley on a permanent basis as Professor and Assistant Director of the Statistical Laboratory. Both he and I had hoped that this move would lead to further joint work, but to my great regret, although we remained close friends, we never again collaborated on a scientific investigation.

6. CHARLES STEIN (1920-)

When in 1945 I opened the September issue of *The Annals of Mathematical Statistics*, I was excited to see a paper containing a highly original solution of a problem that had long been of interest to Neyman and his students. In it, Charles Stein, in a way that was extremely elegant as well as effective, constructed a two-stage test of the hypothesis specifying the value of a normal mean that achieves a given power, independent of the variance, against an alternative value of the mean. I brought the paper to Neyman's attention, and he was as impressed as I was. The next year when he was visiting Columbia, he sought out Charles (who was then still a graduate student) with the result that in the Fall of 1947 Charles joined the Berkeley Statistics faculty.

Neyman's group was expanding, but the space assigned to it was limited; so Charles, Joe Hodges, Evelyn Fix (who had been Neyman's principal assistant during the war years) and I shared an office barely large enough to accommodate our four desks, which were arranged as a square block in the middle of the room. This enforced proximity of four congenial people with common interests led to many discussions and to joint work of mine with both Charles and Joe.

While Henry and I had been interested in the same kinds of problems, Charles, having studied with Wald, had a more decision theoretic background in which admissibility, minimaxity and least favorable distributions were the central concepts. On the other hand, Charles' two-stage paper of 1945 that had so impressed me was concerned with a problem in hypothesis testing. So it was not difficult to interest him in some of the problems in this theory that concerned me at the time. There was one gap that particularly bothered me. Neyman and Pearson had shown that the t-test was UMP among all similar tests. Was it in fact most powerful among all tests at the given level α?[2] This turned out to be essentially a minimax problem to which Wald's theory of least favorable distributions is applicable. In the t-problem, we were able to determine the least favorable distribution and to show that for $\alpha < 1/2$ (i.e., all levels of interest), the most powerful test does depend on the alternative, so that a UMP test does not exist. (The t-test is UMP for $\alpha \geq 1/2$.) By the same method, we were able to solve the corresponding problem for a number of other hypotheses concerning normal distributions. We had expected to extend the results, published as "Most Powerful Tests of Composite Hypotheses. I. Normal Distributions" (Lehmann and Stein, 1948), to exponential and perhaps other distributions but became more interested in other problems, so that the planned part II was never written.

Our next paper was motivated in part by a survey paper by Scheffé (1943), which called attention to "the need for constructive methods of obtaining 'good' and 'best' tests in the nonparametric case." This was the problem on which Charles and I decided to work next, and it resulted in our paper "On the Theory of Some Nonparametric Hypotheses" (1949). The optimal tests for which we obtained a "constructive method" turned out to be generalizations of Fisher's permutation version of the t-test.

Charles and I wrote two more papers together. One was a short note on completeness in the sequential case (1950); the other (1953) took up a problem raised by our first paper. The fact that at the usual levels the t-test is not UMP poses the question whether it is admissible. Since it is UMP among all similar tests, its admissibility follows immediately in a rather trivial local sense. We now proved that it is also admissible in a much stronger sense, namely against the class of alternatives specifying any fixed value of the standardized mean.

That our collaboration did not continue was mainly due to the fact that after only two years Charles left Berkeley for Chicago and later for Stanford. The reason was an anti-Communist loyalty oath that the Regents of the University of California imposed on the Faculty in 1949. After interminable discussions, the great majority of the Faculty signed, but 31 members refused to do so and were dismissed. A few years later, the California Supreme Court ordered their reinstatement. Some returned, while others remained in the positions they had found in the meantime.

For the statistics group, the controversy resulted in the loss of Charles who left the University rather than waiting to be fired. The rest of us, including Neyman who declared himself to be "color blind" on this issue, had found the oath unpalatable but in itself of no great importance and, though disgusted, had signed it without much difficulty.

In the course of our joint work, I learned from Charles about techniques for proving minimaxity including, in particular, the concept of invariance and the Hunt-Stein theorem. This theorem was included in the lecture notes of my course in hypothesis testing recorded by Colin Blyth (Lehmann, 1948–49), which led an existence as an underground text until the official publication of a much expanded version in 1959 (Lehmann, 1959). The notes refer to the Hunt–Stein paper

[2] I was teaching a graduate course on hypothesis testing, and it seemed to me only natural that some student would ask this question. In a similar way, much of my research arose from questions that came up in preparing my lectures. In contrast, Henry Scheffé told me that most of his research came out of his consulting practice.

"Most Stringent Tests of Statistical Hypotheses" as "to be published." However, by 1959, the paper still had not appeared, so that with Charles' permission I presented the theorem and its proof in my book. The reason for the postponement of publication of the Hunt–Stein theorem (told to me by Charles) is interesting. The theorem requires that the group in question satisfy a certain condition (essentially what is now called amenability). This condition holds for the group of translations and the group of orthogonal transformations which are the groups required for (univariate) analysis of variance.

Charles expected it to hold also for the full linear group, required for Hotelling's T^2-test, and felt uncomfortable publishing the paper without it. By 1959, he had constructed a simple counterexample, but by that time the theorem was pretty widely known, so that separate journal publications no longer held much interest for him.

For Charles, invariance was of interest as a condition that may insure stringency or other minimax properties of tests. For the exposition on which I was working for my course and later the book, it played an additional role. Since UMP tests exist only rarely, I was interested in reasonable conditions of "impartiality" which might lead to tests that are UMP within the class restricted in this way. Unbiasedness, introduced by Neyman and Pearson, was one such condition; invariance now provided another. In this way, the results on UMP invariant tests that I had learned from Charles in connection with most stringent tests became important in their own right and played a central role in my account of hypothesis testing. During the three years in which we worked together, Charles thus became an important influence on my development as a statistician and on my career. (For further information on Charles Stein see DeGroot, 1986.)

7. JOSEPH L. HODGES, JR. (1922-)

The third, last and most extensive of my early collaborations was with my fellow student and Guam tentmate Joe Hodges. Joe had come to Berkeley in 1938 as a beginning undergraduate and had stayed on until he obtained his Ph.D. under Neyman in 1948 with a two-part thesis: "I. Initial Sample Size in the Stein Procedure" and "II. Stringency in Acceptance Sampling." The first part he never published. The second part was the first paper in the series of UC Publications in Statistics (Hodges, 1949). (I think it might have had more of an impact had it appeared in a regular statistical journal.) Joe, like myself two years before, was retained by Neyman for his Faculty.

After we both had learned about Wald's decision theory from our officemate Charles Stein, Joe and I were struck by the dearth of concrete examples for which minimax solutions had been worked out, and we decided to tackle a simple case: estimating binomial p with squared error loss. Theory suggested that as a first step we should look for a Bayes estimator with constant risk, since any such estimator would automatically be minimax. The only priors for which we could see how to obtain the Bayes estimator and its risk explicitly were the beta-distributions, and beginners' luck was with us. There was a beta prior which led to a constant risk estimator, and so we had solved the problem on the first try.

But while our solution was a theoretical success, it was from a practical point of view a disaster. Since p is much easier to estimate accurately when it is close to 0 or 1, a constant risk estimator is typically not what one wants and, except for very small sample sizes, the standard unbiased estimator has much smaller risk than our minimax estimator over most of the range of p. What we had thus obtained, unwittingly, was a counterexample showing how unsatisfactory a minimax estimator can be.

Following this 1950 paper, we published during the next fifteen years roughly a paper a year. In the second paper, we presented a new method of proving admissibility by solving an appropriate differential inequality, an approach that has proved more widely applicable than we would have expected. One of my favorites is the next paper, "The Use of Previous Experience in Reaching Statistical Decisions" (1952), which proposes a compromise between the Bayes and minimax approach, namely to minimize the Bayes risk subject to a bound (somewhat larger than the minimax risk) on the maximum risk. Although the principle seems appealing and has found a number of applications, unfortunately such restricted Bayes solutions tend to be rather messy.

By the mid 1950's, Joe's and my interests had shifted from decision theory to the relatively new methodology of nonparametric inference. We were intrigued by Pitman's surprising result that the asymptotic relative efficiency (ARE) of the Wilcoxon to the t-test is $3/\pi = .955$ in the normal case. Computation of the efficiency in some other cases gave even higher values, which raised the question of how low this efficiency can get. The answer given in our paper "The efficiency of some nonparametric competitors of the t-test" (1956) came as a great surprise. The sharp lower bound for this ARE turned out to be .864. These high values gave a boost to nonparametric tests such as the Wilcoxon which generally had been thought to be inefficient.

What made our collaboration so pleasant and effective was the way we complemented each other. Joe was an outstanding problem solver, who enjoyed tackling a problem from scratch and inventing and developing whatever methods were needed to solve it. For me, it was more useful and important to see how the question fitted into the whole fabric of related results and if possible to be guided to a solution by the context. And I then liked to see what the solution contributed to the whole pattern and to consider how far it could be extended. To this difference between us one could add that Joe tended to think more geometrically and I more in terms of formal manipulation.

Joe and I did much of our joint work on walks with or without trails, in the Berkeley hills. Conversation on these walks was not confined to statistics; we talked about music, literature, politics and worked on composing limericks. Here is one of Joe's finest, in which I had no part, and which I quote with his permission.

The music of Johannes Brahms
Has strange, ineluctable charms.
And sometimes it seems
It might lapse into themes,
But alas, they are all false alarms.

Our most memorable hike took place in Summer 1950. We had planned a four-day hike with Charles Stein in Yosemite. Since Wald was in Berkeley at the time we asked him to join us and, as reported in Section 4, he agreed. On the second day of this walking tour, Charles fell sick, and we decided that we had to return. When our trail crossed the road late that afternoon, Joe and I were delegated to hitch a ride down to the valley where our car was parked and then to return to pick up the other two. Joe and I, neither of us experienced hitchhikers, had no luck and decided to try one at a time. Eventually a car stopped for me, and the two women in it opened the door for me to get in. At that moment Joe, with his big frame and height of 6 feet 4 came lumbering out of the bushes. The women gave a scream, slammed the door shut and sped off. We gave up the effort for that night, which was lucky. The next morning Charles felt all right, and the second part of the hike (over Vogelsang Pass) turned out to be the most scenic one.

8. EPILOGUE

The epilogue to a story deals with the future of its characters. The present section will do this for the three subjects of the present account who continued as members of the Berkeley Statistics Department after 1956: Neyman, Scheffé and Hodges.

Neyman

Neyman's resignation as chair constituted a major crisis for the new Department. The statistical community at large expected open warfare and the end of the Berkeley statistics group. Neyman was, after all, the dominant figure; in addition, as project and laboratory director, he controlled all research funds, the total of which was about equal to the departmental budget. There was thus plenty of opportunity for conflict, but the dire predictions did not come true.

That the transition proceeded peacefully was in the first place due to David Blackwell, who, as first post-Neyman chair of the Department, was able to smooth over potential conflicts and to maintain his close and friendly relationship with Neyman as well as with other members of the Department. However, a great deal of credit is also due to Neyman himself, who neither tried to dominate nor to wash his hands of the new departmental organization. He participated in staff meetings like any other member of the faculty and went out of his way to acknowledge the authority of the chair. This found a curious symbolic expression when nearly 20 years later I became department chair. As a student I, like Neyman's other students, had always called him Mr. Neyman, while he addressed us by our first names. This did not change after I joined the faculty; he never encouraged me to use his first name. Now he started signing his notes to me "Jerry."

Neyman had had such a fundamental influence on the development of statistics through his scientific and organizational work and his personal international efforts that it seemed an account of his life and work—the two are largely synonymous—would be of great value. When in 1978 I read Constance Reid's biography of Courant, it appeared to me that she would be the ideal author for such a project. When I approached her with this suggestion, her first reaction was negative, but a few months later she expressed a cautious interest. We agreed that I would sound Neyman out about his willingness to cooperate in such an undertaking. Neyman had often indicated to me that he had no interest in contemplating the past but wanted to think only about the future. It was therefore with some trepidation that I went to see him, with Reid's (1970) Hilbert and (1976) Courant biographies under my arm. Nevertheless, I was taken aback by his reply to my question of how he felt about such a project. "It's a free country," he said curtly.

I made it clear that this was not a basis on which we could proceed. There was no intention of going against his wishes, and in any case his cooperation would be required. "How much cooperation?", he wanted to know. When I told him that Constance needed to meet with him once or twice for an hour or two, he replied that he always enjoyed talking to young ladies. These initial conversations must have gone well, for a pattern soon developed of a meeting between them every Saturday morning, followed by lunch and a drink, and then more conversation. When after a year, Constance told Neyman that she now had enough information and would instead have to start writing, he was quite disappointed. The book, which came out in 1982, a few months after Neyman's death, not only tells the story of his life and accomplishments, but also gives a very lifelike picture of him as a person.

Scheffé

Blackwell was succeeded as chair of the Statistics Department by Le Cam, and he in turn by Scheffé, who chaired the department from 1965 to 1968. It was a period of great unrest in Berkeley, the time of the Free Speech Movement. Student strikes in particular caused difficult problems for the chair. But in spite of violently conflicting attitudes within both faculty and student body, Scheffé managed to hold the Department together and to keep the atmosphere within the Department pleasant. His fair mindedness was greatly valued by all members of the Department.

Scientifically, soon after our *Sankhyā* paper, Henry's main interests turned from optimality theory to more methodologically oriented work, of which his famous paper on the S-method for judging all contrasts in analysis of variance was the first important accomplishment. He brought to this new orientation the rigor of his mathematical origins. This is particularly noticable in his emphasis on clearly defined and justified models and on the consequences of violations of these model assumptions which is such a striking feature and one of the great strengths of his pioneering book *The Analysis of Variance* (1959). He had planned to revise this work after his retirement, but he died in 1977 from injuries sustained in a bicycle accident without completing the revision. It is a pity that he did not live to see the beautiful optimality property of his S-method, established by Wijsman in 1979. I believe it would have given him great pleasure. (For further information on Henry Scheffé, see the obituary by Daniel and Lehmann, 1979.)

Hodges

After 1956, Joe and I continued our joint work particularly in the area of nonparametric inference. Of this later work, I mention only one paper, "Estimates of Location Based on Rank Tests" (1963), which brought the pairing of our names into the statistical terminology. Here we showed how the rank methods that had proved so unexpectedly successful for hypothesis testing could be transferred to point estimation. The class of estimators defined in this way, the so-called R-estimators, includes in particular the estimator now known as the Hodges–Lehmann estimator, which shares the good efficiency properties of the Wilcoxon test.

Joe and I collaborated not only on research but also joined to write an elementary text, *Basic Concepts of Probability and Statistics* (1964; second edition 1970), which we dedicated to Neyman and which was later translated into Danish, Hebrew, Italian and recently into Farsi. It gave a rigorous development of probability and statistical inference in finite sample spaces without calculus. While the probability part was fairly conventional, I believe it was the first elementary book discussing such topics as the Neyman–Pearson Lemma and optimal design. The lower division course ("Stat. 1") which used it as a text was very different from the more cookbookish methods course "Stat. 2." Henry Scheffé once explained the difference by saying that "Stat. 1" was intended for students who wanted to understand statistics but were not planning to use it, while the reverse was true for "Stat. 2." [Since then "Stat. 2" has changed radically under the influence of our colleagues Freedman, Pisani and Purves, whose text *Statistics* (1978; second edition 1991) has made this into an attractive and intellectually stimulating course.]

Some years later, Joe wrote another elementary text jointly with the psychologists Krech and Crutchfield. Called *Stat Lab*, it was based on a single large data set which provided the values of a large number of variables on 6^4 families. A third book for a first year upper division course that Joe taught for many years contains many interesting ideas but unfortunately has remained unpublished. The reason is Joe's dislike of the editorial process, which he shared with Neyman. One of the victims of this dislike was a pioneering paper, joint with Fix, on density estimation (Fix and Hodges, 1951). Because of its historical importance, this paper was published with Joe's permission in the *ISI Review* by Silverman and Jones (1989), who also provided an introduction.

After 1970, our joint work decreased since by then much of Joe's efforts had gone into higher administration, as member and then chair of the Budget Committee both at the campus and state level, and also as advisor on academic personnel matters to both the Chancellor at Berkeley and the President of the whole University system.

Joe retired from the University in 1991, taking advantage of a "golden handshake," a special early retirement offer, which—he claims—netted him an extra 32 cents a month.

REFERENCES

Daniel, C. and Lehmann, E. L. (1979). Henry Scheffé 1907–1977. *Ann. Statist.* 7 1149–1161.

David, F. N. (1949). *Probability Theory for Statistical Methods.* Cambridge Univ. Press.

DeGroot, M. (1986). A conversation with Charles Stein. *Statist. Sci.* 1 454–462.

Fix, E. and Hodges, J. L., Jr. (1951). Discriminatory analysis—nonparametric discrimination: consistency properties. Unpublished report.

Freedman, D., Pisani, R. and Purves, R. (1978). *Statistics.* Norton, New York. [2nd ed., 1991, with Adhikari.]

Hodges, J. L., Jr. (1949). The choice of inspection stringency in acceptance sampling by attributes. *Univ. Calif. Publ. Statist.* 1 1–14.

Hodges, J. L., Jr., Krech, D. and Crutchfield, R. S. (1975). *Stat Lab.* McGraw Hill, New York.

Hodges, J. L., Jr. and Lehmann, E. L. (1950). Some problems in minimax point estimation. *Ann. Math. Statist.* **21** 182–197.

Hodges, J. L., Jr. and Lehmann, E. L. (1951). Some applications of the Cramér-Rao inequality. *Proc. Second Berkeley Symp. Math. Statist. Probab.* 13–22. Univ. California Press, Berkeley.

Hodges, J. L., Jr. and Lehmann, E. L. (1952). The use of previous experience in reaching statistical decisions. *Ann. Math. Statist.* **23** 396–407.

Hodges, J. L., Jr. and Lehmann, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. *Ann. Math. Statist.* **27** 324–335.

Hodges, J. L., Jr. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

Hodges, J. L., Jr. and Lehmann, E. L. (1964). *Basic Concepts of Probability and Statistics.* Holden-Day, San Francisco. [2nd ed., 1970.]

Lehmann, E. L. (1947). On families of admissible tests. *Ann. Math. Statist.* **18** 97–104.

Lehmann, E. L. (1948–49). *Theory of Testing Hypotheses.* Notes recorded by Colin Blyth.

Lehmann, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York. [2nd ed., 1986.]

Lehmann, E. L. and Scheffé, H. (1947). On the problem of similar regions. *Proc. Nat. Acad. Sci. U.S.A.* **33** 382–386.

Lehmann, E. L. and Scheffé, H. (1950/1955/1956). Completeness, similar regions, and unbiased estimation. *Sankhyā* **10** 305–340; **15** 219–236; **17** 250 (a correction).

Lehmann, E. L. and Stein, C. (1948). Most powerful tests of composite hypotheses. I. Normal distributions. *Ann. Math. Statist.* **19** 495–516.

Lehmann, E. L. and Stein, C. (1949). On the theory of some nonparametric hypotheses. *Ann. Math. Statist.* **20** 28–45.

Lehmann, E. L. and Stein, C. (1950). Completeness in the sequential case. *Ann. Math. Statist.* **21** 376–385.

Lehmann, E. L. and Stein, C. (1953). The admissibility of certain invariant statistical tests involving a translation parameter. *Ann. Math. Statist.* **24** 473–479.

Morrey, C. B. (1983). *Jerzy Neyman (1894–1981).* Biographical Memoirs. National Academy of Sciences, Washington, DC.

Neyman, J. (1952). *Lectures and Conferences on Mathematical Statistics*, 2nd ed. Graduate School, U.S. Dept. Agriculture, Washington. [1st ed., 1938.]

Reid, C. (1970). *Hilbert.* Springer, New York.

Reid, C. (1976). *Courant in Göttingen and New York.* Springer, New York.

Reid, C. (1982). *Neyman from Life.* Springer, New York.

Rider, R. (1989). An opportune time: Griffith C. Evans and Mathematics at Berkeley. In *A Century of Mathematics in America, Part 2* (P. Duren, ed.) 283–302. Amer. Math. Soc., Providence, RI.

Scheffé, H. (1943). Statistical inference in the nonparametric case. *Ann. Math. Statist.* **14** 305–332.

Scheffé, H. (1959). *The Analysis of Variance.* Wiley, New York.

Silverman, B. W. and Jones, M.C. (1989). E. Fix and J. L. Hodges (1951): An important contribution to nonparametric discriminant analysis and density estimation. *Internat. Statist. Rev.* **57** 233–238.

Shutes, M. H. (1947). Henry Douglas Bacon (1813–1893). *California Historical Quarterly* **26** 193–200.

Uspensky, J. V. (1937). *Introduction to Mathematical Probability.* McGraw Hill, New York.

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

Wijsman, R. A. (1979). Constructing all smallest simultaneous confidence sets in a given class, with applications to MANOVA. *Ann. Statist.* **7** 1003–1018.

Wolfowitz, J. (1952). Abraham Wald, 1902–1950. *Ann. Math. Statist.* **23** 1–13.