

Some reminiscences of my friendship with Herman Rubin

Herman Chernoff¹

Harvard University

I first met Herman Rubin in 1947 while I was writing a dissertation in absentia at Columbia University and he was a Fellow at the Institute of Advanced Study at Princeton. I had recently been appointed as a research assistant at the Cowles Commission for Research in Economics which was then located at the University of Chicago. Herman had completed or almost completed his dissertation at the University of Chicago, and we were to be colleagues at the Cowles Commission from June 1948 to September 1949.

While I was at Columbia, I was supposed to investigate the possibility of inverting large matrices by computer, because the method used by the Cowles Commission for estimating the parameters of the Economy, by maximizing a function using a gradient method, involved the inversion of matrices. I worked at the Watson Laboratories which were located then near Columbia and had use of a “Relay Calculator” which could be programmed (with plug boards) to multiply matrices. With the use of the Relay calculator and a card sorter and lots of fancy footwork, it was possible to do the job. At that time the engineers at Watson were beginning to build the electronic computer which was to become one of the bases for the future development of the IBM computers to follow. But I did not have access to that machine. However I did have access to Herman Rubin who came around to kibbitz, and to do some of the fancy footwork. At one point the sorter decided to put the cards with the digit 4 into the box for the digit 7. We counterattacked by instructing the 7 to go into the reject box. That scheme worked for a while, but the sorter replied by putting the 3 into the reject box. I think that we ended up doing some of the card sorting by hand.

At Cowles we had adjacent offices which was not exactly a blessing because Herman had a bad habit. He would come in to the office about 7 AM, pound his calculator (electric and not electronic) for an hour and then prove a few theorems for an hour, and then was ready to discuss important matters with me when I came to work. These important matters were usually how to handle certain bridge hands. Whatever I suggested was usually wrong. That did not bother me as much as the time I had to spend on bridge, a game that I never properly mastered.

I had a few friends in the Mathematics Department at the University. One of them, who had become a long term fixture, related to me how he had thought he was very smart (IQ about 180) when he was an undergraduate, until this little high school kid showed up, and obviously was more capable than most of the graduate students. Needless to say that that enfant terrible was our Herman Rubin.

While we were at Cowles we coauthored a paper, the main object of which was to show that even when not all of the standard conditions were satisfied, large sample theory demonstrated that we could still have confidence in our conclusions. I must admit that my contributions to this effort were only to translate

¹Department of Statistics, Harvard University, Cambridge, MA 02138, USA. e-mail: chernoff@stat.harvard.edu

Herman's work into comprehensible English, and to insist on the admittedly improper use of the word "blithely" to indicate that we could proceed as though all was well.

On another occasion L.J. Savage announced that he had resolved Wald's dilemma in using the Minimax principle, by claiming that what Wald had really meant was "Minimax Regret". In illustrating this principle in a class I was teaching, I discovered that not only could Minimax lead to poor choices, but Minimax Regret violated the principle of "Independence of Irrelevant Alternatives", a principle that had recently been enunciated in Arrow's outstanding result on the Welfare Function. When I confronted Savage with this problem, he first denied that it was serious, but after some discussion, indicated that maybe we should follow up on recent work by De Finetti proposing the Bayesian approach.

In fact I laid out the axioms that I felt should be satisfied by an objective method of rational decision making. The current terminology is "coherent". My results were sort of negative and later published in *Econometrica* after I let them simmer for a few years. The only thing that almost made sense, was that if we neglected one of the axioms, then the rational way to proceed is to treat each unknown state of nature as equally likely. This was an unsatisfactory result for those hoping for an objective way to do inductive inference. In the meantime both Savage and Rubin pursued the Bayesian approach. Savage later became the high priest of the Bayesian revolution. But no one seemed to notice, that two days after the discussion with Savage, Rubin wrote a discussion paper deriving the Bayesian solution. What was special about this paper, was that by omitting unnecessary verbiage, it was about three pages long, and was, unlike most of Herman's writing, eminently readable. Unfortunately, a copy of this paper which I treasured for many years, disappeared in my files, and as far as I know, no copy of it exists today.

I recall going to a seminar in the Mathematics Department, where I confess that I did not understand the lecture. At the end someone asked the speaker whether his results could be generalized to another case. The speaker said that he had thought about it, but was not clear about how to proceed. Herman spoke up, indicating that it was perfectly clear, and explained exactly how the generalization would go. This was one of many examples where it was apparent that Herman could instantly absorb results that were presented to him, and even see further nontrivial consequences of these results. I envied this clarity of thought because my own thinking processes tend to be much more confused and usually some time is needed for me to get things straightened out.

In 1949, Rubin and Arrow left the Cowles Commission to go to Stanford. Rubin joined the new Statistics Department organized by Albert Bowker with the help of Abraham Girshick. Arrow was joint in Statistics and Economics. I went to the University of Illinois, and was invited to visit Stanford for a semester two years later. I found the department to be an exciting place to be, partly because of the distinguished visitors which included David Blackwell and partly because of the presence of ONR funding for applied and theoretical programs. Herman was teaching courses in measure theory and topology, because the Mathematics Department was busy with other topics and he felt that Statistics students should at least have those basics.

While I was there, Girshick once was teasing Herman about the fact that the news indicated that an African American had just received his Ph.D. at age 18, and Herman had not gotten his degree until he was 19. Herman, taking this teasing seriously, complained that he had spent a year in the Army.

That semester, two topics that arose from the ONR project gave rise to two papers that I wrote and of which I was very proud. They pointed to a direction in optimal experimental design on which I spent much time later. Part of one of these papers involved finding asymptotic upper and lower bounds on the probability that the mean of a sample of independent identically distributed random variables would exceed a certain constant. This paper represented the first application of large deviation theory to a statistical problem. Cramer had derived a much more elegant result in 1938, of which I had been ignorant. My result, involving the infimum of a moment generating function, was less elegant and less general than the Cramer result, but did not require a special condition that Cramer required. Also, my proof could be described as crudely beating the problem to death. Herman claimed that he could get a lower bound much easier. I challenged him, and he produced a short Chebyshev Inequality type proof, which was so trivial that I did not trouble to cite his contribution.

What a mistake! It seems that Shannon had incorrectly applied the Central Limit theorem to the far tails of the distribution in one of his papers on Information theory. When his error was pointed out, he discovered the lower bound of Rubin in my paper and rescued his results. As a result I have gained great fame in electrical engineering circles for the Chernoff bound which was really due to Herman. One consequence of the simplicity of the proof was that no one ever bothered to read the original paper of which I was very proud. For years they referred to Rubin's bound as the Chernov bound, not even spelling my name correctly. I once had the pleasure of writing to a friend who sent me a copy of a paper improving on the Chernov bound, that I was happy that my name was not associated with such a crummy bound. For many years, electrical engineers have come to me and told me that I saved their lives, because they were able to describe the bound on their preliminary doctoral exams. Fortunately for me, my lasting fame, if any, will depend, not on Rubin's bound, but on Chernoff faces.

As I was preparing to return to the University of Illinois to finish off my year in 1952, my wife and I had a long discussion with Herman in which he mentioned that he had certain requirements for marriage. Evidently his proposed wife would have to be a paragon of virtues, beautiful, brilliant, and Jewish. When we returned to Stanford five months later, Herman had discovered this paragon and she was willing and they were already married.

For a few years after I came to Stanford, Rubin and I had neighboring offices at Sequoia Hall. Frequently when I came across a problem that seemed to be one that must have been treated in the literature, I would approach Herman and ask him about it. It was not unusual for him to say that it was not yet in the literature, but that he had already solved that problem. He would then reach into the depths of the mountain of paper on his desk, and pull out his solution. Often I would come to him with a problem on which I was working, and suggest an approach that I might use. His invariable response was "That is the worst way to attack that problem." This response frightened off many students and colleagues, but I found that if I persisted in asking why it was the worst way, he would sometimes explain why and sometimes admit that maybe it was a sensible approach. It required a certain amount of stubbornness, which not everyone had, to confront Herman. But I found that, because Herman was my neighbor, I was often saved from following false trails, often shown what was known, and often encouraged to pursue profitable directions that seemed problematic.

The Japanese have a title of National Treasure which they assign to outstanding artists and scholars. In my opinion, Herman Rubin, the eternal enfant terrible

of Statistics, has served as an American National Treasure by his willingness to counsel those not too frightened to hear "That is the worst way". As I recently became an octogenarian, I realize that Herman is no longer the 20 year old I once knew, but I have no doubt that he is still intellectually a slightly matured 20 year old who has contributed mightily to Statistics and from whom we can expect more.