

A Random Walk with Drift: Interview with Peter J. Bickel

Ya'acov Ritov

I met Peter J. Bickel for the first time in 1981. He came to Jerusalem for a year; I had just started working on my Ph.D. studies. Yossi Yahav, who was my advisor at this time, busy as the Dean of Social Sciences, brought us together. Peter became my chief thesis advisor. A year and a half later I came to Berkeley as a post-doc. Since then we have continued to work together. Peter was first my advisor, then a teacher, and now he is also a friend. It is appropriate that this interview took place in two cities. We spoke together first in Jerusalem, at Mishkenot Shaananim and the Center for Research of Rationality, and then at the University of California at Berkeley. These conversations were not formal interviews, but just questions that prompted Peter to tell his story.

The interview is the intellectual story of a post-war Berkeley statistician who certainly is one of the leaders of the third generation of mathematical statisticians, a generation which is still fruitful today.

The conversation was soft spoken, a stream of memories, ordered more by association than by chronology. In fact, I led Peter to tell his story in a reverse direction, starting from the pure science and ending with the personal background. So, please sit back, and imagine you are part of the chat.

Peter, if you try to summarize the many stages of your career, how do you characterize the different periods?

A random walk with drift. Shall I start with the very beginning?

No, for now can you tell us about your academic career?

I would say that the first period was almost exclusively theoretical work, but actually, I think, almost from the beginning, driven as much by people as by a focus on the subject.

I did my thesis with Erich Lehmann. From the thesis, I published two papers on multivariate analogues of Hotelling's T^2 (Bickel, 1964; Bickel, 1965a), really not knowing much about multivariate analysis at all, learning asymptotics as I went along. After the thesis,

Ya'acov Ritov is ???, Mount Scopus, Jerusalem, 91905, Israel (e-mail: yaacov@mssc.huji.ac.il).

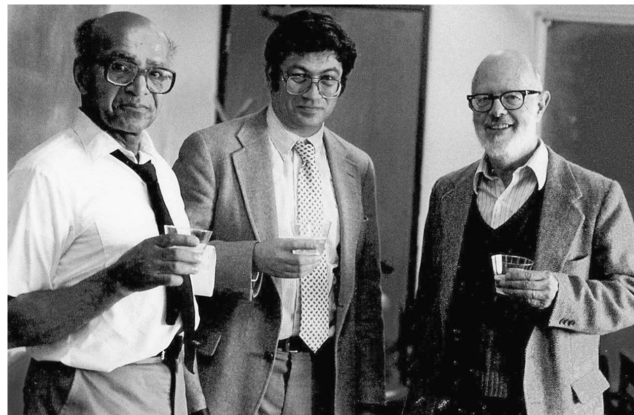


FIG. 1. David Blackwell, Peter Bickel and Erich Lehmann at a party celebrating Peter's election to the National Academy of Science.

partly talking with Peter Huber, and partly talking with Erich Lehmann, I did some interesting things on questions of robust estimation. I had a paper on trimmed means and how they compare to the mean and median (Bickel, 1965b); again, the results were in the spirit of Hodges and Lehmann.

The next stage happened by a curious accident due to Govindarajulu. He asked me if I had ever thought about investigating linear combinations of order statistics. I said, no, but I had some ideas, having learned about weak convergence of stochastic processes. Initially, he said he wanted to work with me, but, at the same time, he was talking with people in Stanford; and then he carried the problem to Le Cam. The result was, finally, that the problem was attacked with three different approaches. One was the approach of the Stanford group, growing from the work Herman Chernoff did on rank statistics, mine, using weak convergence of the quantile process (Bickel, 1967), and Le Cam's, which used the Hájek projection technique. We all got results.

Within this work, I was very pleased about the result that the covariance of two order statistics is non-negative, which Richard Savage claimed was a long unsolved problem. Then it turned out that it was an inequality in Hardy, Littlewood and Polya. This work also led to work in multivariate goodness of fit tests and

1 other problems to which I applied the new notions of
 2 weak convergence of processes.

3 I went through my Ph.D. studies very quickly, which
 4 left me unfamiliar with many parts of statistics. I tried
 5 to fill the gaps later on. After my thesis Yossi Yahav
 6 and I started talking about his notion of asymptotic
 7 pointwise sequential analysis purely from a theoretical
 8 point of view. He had solved some special cases, and I
 9 realized that there was a general pattern we could use.
 10 We have two or three papers in that direction (Bickel
 11 and Yahav, 1967; Bickel and Yahav, 1968; Bickel and
 12 Yahav, 1969a; Bickel and Yahav, 1969b).

13 Erich Lehmann characterizes people as problem
 14 solvers, like Joe Hodges, and system builders, like
 15 Erich himself. I fall somewhere in between, but I am
 16 primarily a problem solver.

17 An area that I started to work on in the seventies
 18 with van Zwet and Götze was second order asymptot-
 19 ics (Bickel, Götze and van Zwet, 1985, 1986). It was
 20 prompted by Hodges and Lehmann’s paper on defi-
 21 ciency. I got an idea on how to prove things for one
 22 sample rank tests. Van Zwet, independently, got fur-
 23 ther, but was stumped by two samples tests. We talked
 24 and, using a method of Renyi, we got a complete an-
 25 swer for rank test statistics and permutation test statis-
 26 tics (Albers, Bickel and van Zwet, 1976). Later, I was
 27 asked to give what is now known as an IMS Medal-
 28 lion Lecture, for which I had to give a topic. I pro-
 29 posed Edgeworth expansions and nonparametric statis-
 30 tics, some of which I knew how to do, at least formally.
 31 But then I had to have a real example, so I chose U -
 32 statistics. Something like a month before the lecture I
 33 realized that there was a substantial difficulty with my
 34 arguments. Luckily, just in time, I found an idea which
 35 worked. Not quite the right idea, but it did give the first
 36 Berry–Esseen bound for U -statistics (Bickel, 1974).
 37 Subsequently, Jon Weirman in Seattle got it right.



49 FIG. 2. From left to right, Peter Bickel, Juliet Shaffer, Erich
 50 Lehmann, Kjell Doksum and David Freedman, celebrating the 65th
 51 birthday of Erich Lehmann.



FIG. 3. Willem van Zwet, Jon Wellner and Jianqing Fan. Bick-
 elfest, Princeton 2005.

I took part in the Princeton Robustness year in 1971
 (Andrews et al., 1972). Unfortunately, I didn’t under-
 stand Tukey most of the time. He had his own language
 that one had to follow. But there were other interesting
 people I could talk with easily. Peter Huber was there,
 David Andrews, and Frank Hampel. One of the things
 that came up was the issue of adaptation. Peter Huber
 and I agreed that it was symmetry that was doing the
 trick. Then I decided to think more about this question.
 But we’ll return to that in a moment.

Then, surprisingly, I moved into something genu-
 inely applied.

Somehow, through Joe Hodges, I got interested in
 finding out more about the university, so I joined some-
 thing called the Graduate Council. Eugene Hammel,
 the Associate Dean of Graduate Studies, presided over
 the Council. One day, Hammel told me he had a very
 strange problem: he had analyzed data on graduate ad-
 missions, because he was worried that the government
 would cut funding, on the grounds that Berkeley was
 biased.

Indeed, he found strong evidence of gender bias. So
 he looked to find the departments or units where the
 bias was, since decisions were made on the department
 or unit level; he couldn’t find them. I told him that there
 is no contradiction and I gave him an example of what
 I later learned was the Simpson paradox. Eventually,
 we made several tests for conditional independence by
 units, one of which we later found out was equivalent to
 the Mantel-Haenszel test. It was an enjoyable paper; it
 appeared in *Science* (Bickel, Hammel and O’Connell,
 1975).

I had some other excursions into applications. I was
 recommended by Betty Scott to the National Research

52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102



FIG. 4. *Katerina Kechris, Haiyan Huang, Friedrich Goetze and Peter Bickel. Bickelfest, Princeton 2000.*

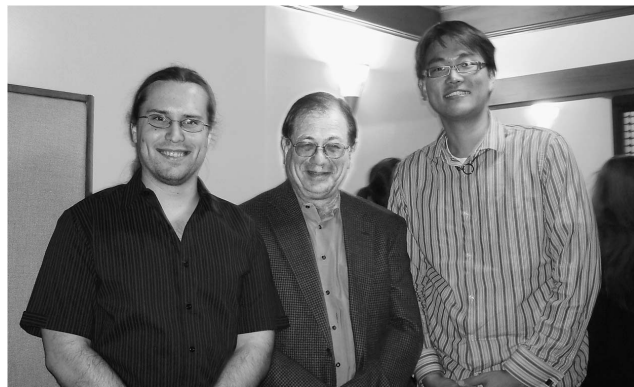


FIG. 5. *Peter Bickel and two recent students, Ben Brown (left) and Choongsoon Bae.*

Council and served on two committees which studied two insurance problems. The first problem was how to implement a mud slide insurance program. The government already ran a flood insurance program in which it subsidized insurance companies to give insurance to flood areas, provided the communities would agree that people would not build on the flood plains. When Southern California had a period of extreme rain, there were mud slides all over. Consequently, the representatives of the districts with mud slides got a law passed in Congress requiring the government to construct a mud slide insurance program. But nobody knew how to do that.

So they convened the mud slide panel. It turned out that the main problem was the nature of the data. They had very extensive aerial photographs of the extent of mud slides, but nobody knew whether they had happened last year or a thousand years ago, so there was no way to set premiums. Basically, the panel knew what had to be done. They proposed that teams of engineers be collected to look at the candidate areas. The engineers would scratch their heads, and they would come up with an insurance rate, given the information available. I couldn't see what else could be done. However, I suggested, "Why don't you have different groups of engineers rate the same area and see if you can test for consistency?" Nobody else on the panel agreed.

The funny thing is that a year or two later I was put on another panel dealing with the same issue. This time it was about flood insurance. The problem was that the Federal Insurance Administration contracted out to different government agencies to assess flood risk. The US Geological Survey and the Army Corps of Engineers were assessing adjacent areas. The border was in the middle of a flood plain, but they came up with different rates for the two halves of the flood plain! Any-

way that was fun. I enjoyed it. But I never did any serious data analysis, in part because I didn't trust myself to be sufficiently observant.

Another line of research that had an impact on my later interests was a paper on the maximum deviation of kernel density estimates that I worked on with Murray Rosenblatt (Bickel and Rosenblatt, 1973). Murray came to my office back in the 70s and asked whether empirical process theory with weak convergence would work on extrema. I started to think about this and realized that it couldn't work that way, because the limit is white noise. Eventually I saw a way that you could attack the problem with Skorohod embedding. While working with Rosenblatt and van Zwet, I read an old paper of Hodges and Lehmann where they looked at minimaxity subject to restrictions, where they did some explicit calculations for the binomial. At some point I learned an identity that Larry Brown pointed to: you can relate the Bayes risk in the Gaussian shift model to the Fisher information of the marginal distribution. This led to papers on semiparametric robustness (Bickel, 1984) and on estimation of a normal mean (Bickel, 1983), which had a surprising follow-up in the work of Donoho and Johnstone.

At about the same time, I looked at the question of adaptation again. A paper on that subject (Bickel, 1982), as well as the preceding work on asymptotic restricted minimax, was developed during a Miller professorship, and that work was given in my Wald lectures. Then I came to Israel on sabbatical and I had the good fortune of having you as a student working on Bayesian robustness.

The next stage started when I gave lectures at Johns Hopkins on semiparametric models, and I began to put things in context and realized the connections with Jon Wellner's and Pfanzagl's work and robustness and so

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102

on. Then you, Jon, Chris Klaassen and I started working on our joint book (Bickel et al., 1993 & 1998) and, in the process, solved, separately or jointly, the problems of censored regression and errors-in-variables (Bickel and Ritov, 1987). It was great fun working on the book.

A significant excursion was some work with Leo Breiman. Leo was visiting Berkeley, considering whether to become a faculty member. He spoke about a multivariate goodness of fit test he devised using nearest neighbors in high dimension. Its asymptotic limiting behavior was harder to understand than either of us thought initially, but eventually we worked it out (Bickel and Breiman, 1983). The work appeared in the *Annals of Probability* because the *Annals of Statistics* was then run by David Hinkley, whose views of what constituted statistics were quite different from mine.

Another excursion which was important was to the theory of the bootstrap, on which I early on worked with Freedman, and later in the 90s with Götze and van Zwet (Bickel, Götze and van Zwet, 1995). Brad Efron introduced the bootstrap, following his profound insight into the impact of computing on statistics. It took me a while to realize that the bootstrap could be viewed as Monte Carlo implementation of nonparametric maximum likelihood. So these various things intertwined. I eventually became interested in the interplay between high dimensional data and computing and the tradeoff between computing and efficiency. From the 90s to the present, you and I moved from semiparametrics to nonparametrics, for example, nonparametric testing, the LASSO and that kind of thing. I think it is fair to say that this work was promoted in part by our participation in an unclassified National Security Administration program and in part by conversations with Leo Breiman. As you know, Leo and I got

quite close. I learned a lot from him and I became more sharply aware that high dimensional data and computing had led to a paradigm change in statistics.

During our collaboration, I have tried to keep up with you. Working with graduate students—especially you—is a key part of my life. Ideas come to me when I talk. I have had lots of students, many very good, a few outstanding; all of them were important to me. Just as important have been senior collaborators, including a number of former students and colleagues at Berkeley, Chicago, Stanford, Seattle, Michigan, Harvard, Zurich, Leiden, Bielefeld and Israel.

As I became older, I finally became bolder in starting to think seriously about the interaction between theory and applications, a direction initiated in part by working with my student and, later, collaborator, Liza Levina. Nancy, my wife, thinks, and I think she is right, that getting the MacArthur Fellowship made a change. I never was sure of myself and the MacArthur helped. I became more self-confident. When my student, Niklaus Hengartner, was working on a specific applied problem, we realized that the theoretical semiparametric ideas really helped (Hengartner et al., 1995). Then, the work with John Rice, you, and the Engineering and Computer Science people on transportation problems played an important role. I think my collaboration with John Rice has been very fruitful. He is a wonderful data analyst, has very interesting ideas on the questions that can be asked and is also very knowledgeable about different techniques. I really like the recent work with John and Nicolai Meinshausen, which is coming out in the *Annals of Applied Statistics* (Meinshausen, Bickel and Rice, 2009). As far as



FIG. 6. Peter Buehlmann, John Rice and Niklaus Hengartner. Bickelfest, Princeton, 2005.



FIG. 7. Peter Bickel, Quang Pham, Kang James, Yossi Yahav, Berry James and Nancy Bickel. Bickelfest, Princeton 2005.



FIG. 8. *David Donoho, Iain Johnstone and Peter Bickel, celebrating Iain's election to the National Academy.*

I know, it is the first paper which makes it absolutely clear that a main issue is the tradeoff between the efficiency of a procedure and the amount of computer time required to implement it successfully.

Then there is biology. I was always interested in biology. I met this wonderful guy, Alexander Glazer, who was then the chair of Molecular Biology in Berkeley. At some point, because I was thinking about exploring biology, we talked after a lecture. He said he was unhappy about critiques by phylogeneticists of recent work on some proteins he had long studied. When he gave a talk at Stanford, these critics claimed that, using statistical methods, they had obtained a phylogenetic tree that contradicted his views. I had some doubts about statistical methods in this context, said so, and we started to talk. He was just retiring and closing his lab, becoming a high level administrator, but he wanted to keep his hand in.

I had a very good student, Katerina Kechris, who was just starting. I got her involved in his program. I told her it was risky, but we had a chance to work with a real biologist. It worked out so well that our first paper actually became Alex's inaugural paper in the Proceedings of the National Academy of Sciences when he was elected to the Academy (Kechris et al., 2006). This work taught me something about the limitations of fundamental experimental information. We did some statistical analyses and introduced some funny methods for finding critically important sites in proteins. Since the crystallographic structure of the proteins we were studying was known, we looked to see if the sites we identified statistically were visibly critical to the structure. Not a chance! So in the end we made our case with statistics, by saying that the percentage of time that mutations in our sites have serious consequences is larger than expected by chance, they are closer to

some critical structures then if they were randomly selected and so. But we never were able to say that if you change the amino acid in one of the positions we identified, everything collapses.

Then I put together a proposal with Katerina Kechris and a young colleague, Haiyan Huang, for a special program of the National Science Foundation, funded largely by the National Institute of General Medical Sciences, for problems on the borders of biology and mathematics. The proposal had two parts. One was on a key question of Alex Glazer about lateral gene transfer between different species of bacteria. A second part was on the functional importance of genomic sites conserved between very distant species. We proposed to use a data set from Eddy Rubin's lab. The reviewers liked the conserved sites part of the proposal, but hated the lateral gene transfer part, so our overall score was borderline. Shula Gross was then an NSF program director. She persuaded the NIGMS people to fund the proposal. Katerina, Haiyan, a student, Na Xu, and I started working on the conserved sites problem—with very modest success. Because we got the funding, I was able to attract a student, Ben Brown, from an engineering program and put him to work on the genomics questions. He is both passionate and scholarly about the biology and has mastered a great range of computing techniques on his own. Our results on the Rubin data were still not very satisfying. Fortunately, we were led to change our focus by connecting with a group at the National Human Genome Research Institute.

Because I have grandchildren in Washington, I decided to find a suitable academic base for visiting in



FIG. 9. *Peter Bickel and Ya'acov Ritov, a day before the interview was taken place. In the background, Mount Olive, Jerusalem.*

1 DC. I looked for a place to work at the National Hu-
2 man Genome Research Institute. I didn't know any-
3 body there, but I was on a committee with biochemist
4 Maynard Olson, who had been a post doc with Alex
5 Glaser. Olson's own post doc, Eric Green, directed
6 a lab group at NHGRI. There, I got involved in the
7 ENCODE (Encyclopedia of DNA) project. Little hap-
8 pened during my first visit, but the next year I brought
9 Ben. That made a big difference. Ben was able to talk
10 with the biologists in their language and carry through
11 computations.

12 I had not expected important statistical methods to
13 come out of the project. My main motivation was learn-
14 ing more about the biology. But, it turned out that
15 we needed to develop a nonparametric model for the
16 genome, which might turn out to be interesting and im-
17 portant. This model is the most nonparametric model
18 you could think of for the genome. In addition, the
19 method of inference that we developed, a modified
20 block bootstrap, gives a check on whether what you
21 are doing is reasonable. On the scales that ENCODE
22 was studying, our theory requires that the statistics on
23 which our methods are based should have approxi-
24 mately Gaussian distribution. By plotting the bootstrap
25 distributions of these statistics, we can see if this as-
26 sumption is roughly valid. Moreover, the model is ro-
27 bust in the sense that, even if the approximation is poor,
28 p -values for tests of association between features are
29 conservative. This work turned out to be very nice the-
30oretically and the biologists seem to like it a lot. I am
31 now somewhat confident that this framework may lead
32 to major contributions.

33 The other direction I've been following is connected
34 with the location of my second set of grandchildren in
35 Boulder, Colorado. I've been visiting at the National
36 Center for Atmospheric Research in Boulder, work-
37 ing with the statistics group headed by Doug Nychka.
38 A basic goal at NCAR is to do relatively short term
39 weather prediction based on computer models. You can
40 say "Why do I need a computer model? I can make
41 predictions using yesterday's weather or other past in-
42 formation." But, because of the high dimensionality
43 of the problem, this approach doesn't work. The com-
44 puter models are valid enough to dramatically improve
45 prediction. However, the computer models themselves
46 produce high dimensional data.

47 Again, through chance, Thomas Bengtsson came to
48 Berkeley. Bengtsson had spent some time in NCAR
49 and, working with the physicists, had been trying to un-
50 derstand how to use these computer models effectively.
51 They had hit a serious problem, the collapse of particle

filters in high dimension. Bengtsson, Snyder, Ander- 52
son, two physicists at NCAR and I were able to analyze 53
this phenomenon. This led to a paper in the *Monthly* 54
Weather Review (Snyder et al., 2008) and some theo- 55
retical papers (Bickel, Li and Bengtsson, 2008). I am 56
now working with Jing Lei, a graduate student, trying 57
to bypass these difficulties of particle filters. 58

As it turns out, both of my current fields of applica- 59
tion, genomics and weather prediction, have fed natu- 60
rally into my theoretical interests in understanding high 61
dimensional data analysis. I still work at a rather ab- 62
stract level, and don't deal well with details, but, for- 63
tunately, my students, post docs and colleagues com- 64
pensate for my shortcomings, so together we're able to 65
make satisfying contributions to both theory and prac- 66
tice. 67

I want to go back to your student years in Berkeley. 68

I started at Caltech, was there for two years, but then 69
transferred to the University of California, Berkeley. 70
I finished my undergraduate work at Berkeley in one 71
year, because I had done five years of high school in 72
Canada, not four as in the US. Caltech paid no atten- 73
tion to the extra year, but Berkeley did give me credit 74
for my fifth high school year, provided I completed my 75
undergraduate degree in mathematics. I had moved to 76
UCB intending to switch to psychology. That's what 77
brought me to a class taught by Joe Hodges. I thought 78
statistics would be necessary for a psychology student. 79
That class drew me into statistics just at the time that 80
a graduate class in mathematical learning theory made 81
me skeptical about psychology. 82

I took a Master's degree in math while I was deciding 83
whether to go into math or statistics. I actually wanted 84
to do my Ph.D. with Hodges, but Hodges insisted that 85
people come to him with their own problems, and I 86
wasn't ready to do that. So he steered me to Lehmann. 87
That was very fortunate for me. Erich Lehmann was 88
really a life guide, not just an academic one. Acad- 89
emically, my progress was a bit funny. I spent only 90
two years in my Ph.D. program. I had already taken 91
the basic graduate probability course. I found statistics 92
interesting and wanted to pursue it in depth. I could 93
really have gained by studying a little bit more, but 94
again chance intervened. The Department had an oral 95
Qualifying Exam for the Ph.D., with three panels of 96
faculty members—in theoretical statistics, applied sta- 97
tistics and probability theory. The students were exam- 98
ined by each of the three panels. My friend Helen Wit- 99
tenberg (now Shanna Swann) needed a study partner 100
and enlisted me. So I took the qualifying exam a year 101
earlier than I otherwise would have done. 102

1 The applied statistics exam was a bit of a farce.
 2 How did one prepare oneself? One read thoroughly
 3 Scheffe's book, *The Analysis of Variance*, a lovely
 4 book, but it's really a theory book, with few examples
 5 of analyses of real data. The panel on applied statistics
 6 consisted of Elizabeth Scott, Jerzy Neyman, Evelyn
 7 Fix and Henry Scheffe. They asked me about the book,
 8 and that was OK. Then Betty Scott actually asked me
 9 something applied, and I didn't know what to say. They
 10 passed me anyway. I was tired of school and wanted
 11 do a thesis right away. Erich gave me a problem, which
 12 I didn't know very much about, but I succeeded. The
 13 Department hired me; so I stayed.

14 *How would you describe the people in Berkeley at*
 15 *this time?*

16 It was a very eminent group and there was a lot of
 17 collaboration and a cordial environment. I have en-
 18 joyed these aspects of the department very much from
 19 the beginning. There were many joint papers. Between
 20 Hodges and Lehmann, of course, there was a long col-
 21 laboration. Blackwell and Hodges had papers; Black-
 22 well and Le Cam had papers. I don't know about Ney-
 23 man and Le Cam, but certainly they interacted in-
 24 tensely. Henry Scheffe and Erich also worked together
 25 a lot before my time. Before he moved to Stanford in
 26 the early 50s, Charles Stein worked with Erich. There
 27 were some tensions in the department, but I wasn't
 28 aware of them at that time.

29 Le Cam and Neyman viewed themselves as applied
 30 statisticians, though the rest of the statistical world
 31 might not have agreed. Betty Scott did applied statis-
 32 tics, in astronomy and climatology. Henry Scheffe was
 33 a serious applied statistician. He worked with Cuthbert
 34

Daniel, who was a private consultant and very impres- 52
 sive. Henry brought him to Berkeley for a semester of 53
 lectures, which was very good for all of us. Joe Hodges 54
 was considered the most talented applied statistician in 55
 the department. He had a wonderful sense of data, but 56
 Joe, interestingly enough, didn't want to be an applied 57
 statistician. 58

59 The intellectual center of the Department was cer-
 60 tainly mathematical theory. There was a young group
 61 of probabilists, including David Freedman and Lester
 62 Dubins and the more senior Loeve and Le Cam. David
 63 Freedman eventually switched to statistics. The rela-
 64 tions with Stanford were excellent. We used to have
 65 the Berkeley–Stanford colloquia twice a quarter, one in
 66 Berkeley, and one in Stanford. So, it was a very pleas-
 67 ant place to work.

68 I collaborated with Erich and Joe and with David
 69 Blackwell. Eventually, but much later, I collaborated
 70 with David Freedman, I think that's about it with the
 71 early years group. Subsequently, I collaborated with
 72 later arrivals, Leo Breiman, Rudy Beran and Warry
 73 Millar, as well as, of course, Kjell Doksum, with
 74 whom, in addition to papers, I published a book whose
 75 second edition we are still working on.

76 Not long after I started teaching, in the late 60s and
 77 early 70s, Berkeley was full of turmoil. Things hap-
 78 pened that had nothing to do with statistics. I and most
 79 of my colleagues supported the Free Speech Move-
 80 ment. Later on we supported a student strike by hold-
 81 ing our classes off campus, but we were not personally
 82 engaged. Of course, it was very emotional. People left
 83 the university from both the right and the left. When
 84 Reagan was governor and when Nixon was president,
 85 conflict about the Vietnam War got heated. I remember
 86 teaching class in Dwinelle Hall at noon, and tear gas
 87 coming through the Windows.

88 *How do you define your generation? Erich Lehmann*
 89 *was the leading person in the second generation. The*
 90 *third generation more or less started with you and your*
 91 *colleagues.*

92 To some extent, yes. Moving beyond Berkeley, I've
 93 been struck by a curious observation. A substantial
 94 number of leading figures in my generation came
 95 from Caltech. They include, among others, Brad Efron,
 96 Larry Brown, Chuck Stone and Carl Morris. Nobody
 97 taught statistics at Caltech. But, for some reason, we all
 98 felt we wanted to do things in the real world. Among
 99 us, only Brad claims that he always wanted to do sta-
 100 tistics. Larry went to Cornell and worked with Jack
 101 Kiefer and wrote a statistical thesis. Chuck said he
 102 wanted to do statistics, but he moved to probability



50 FIG. 10. Peter Bickel in his parents, Madeleine and Eliezer
 51 Bickel. Ca. 1943.



FIG. 11. Peter Bickel with his uncle and aunt, Shlomo and Yetta Bickel, and his cousin Alexander Bickel.

as a student of Karlin. Later he got involved with Leo Breiman and went into statistics.

Can you tell about yourself? You once told me, “We are lucky to belong to a generation that didn’t suffer from wars.” I found it an interesting comment from somebody who was born as a Jew in central Europe during WWII. Can you tell us about your history?

I was born in Bucharest in 1940, but I was really not very aware of the war, except that sometimes we had to go to the bomb shelters (when the Americans were bombing the oil field in Ploesti). Once, when we were coming back, I saw broken windows in some office buildings. My father Eliezer (Lothar) Bickel was able to continue to practice medicine during the war. There was a pogrom in Bucharest, which we narrowly avoided by my mother Madeleine’s courageous behavior. Then, after the war and after the communists took over, my parents arranged, with difficulty, to leave Romania legally. But I didn’t realize the difficulties at the time.

We went to France in 1948 and then to Canada in 1949. I studied in France in a public school for ten months. In France my father insisted on giving me English lessons after an eight hour day of school and homework. From Canada we went to California. I could have been drafted in the Vietnam War, but I married young and we had a child. So that probably was the source of my remark.

Can you tell us about the intellectual influence of your family on you?

My father was born in Bukovina, a German speaking province of the Austro-Hungarian Empire. He had a traditional Jewish and then a secular education. In high school, under the influence of one of the high school



FIG. 12. The young Nancy and Peter Bickel, and their daughter and son Amenda and Stephen. This picture was used as new year greeting card.

teachers, he and other Jewish students became involved in one of the many intellectual groups of the time. He became, basically, a disciple of a German philosopher called Constatin Brunner, a son of the grand rabbi of Hamburg, who rebelled against his father. Brunner was involved in elaborating a philosophical system based on Spinoza. He believed strongly that the Jews should assimilate to German culture. One of his books, called “*Unser Christus oder Das Wesen des Genie,*” or in English, “*Our Christ, or The Essence of Genius,*” among other things, advocated assimilation to German culture, including Christianity. Like Brunner, my father favored assimilation. In Romania, we never celebrated Jewish holidays, and in fact, we celebrated Christmas, but in a nonreligious way.

My father became the leader of the Brunner group; he was treated almost as the equivalent of a Hasidic rabbi. He pursued two fields at the same time—medicine and philosophy. He was able to study medicine at the University of Bucharest, even though very few Jews were admitted because there was a “*numerus clausus.*” Also, as he was the second son, his father wanted him to run the family store and wouldn’t pay for his further education. He rebelled and had to struggle on his own. My father went to Germany to do post graduate study in medicine. At the same time he was able to meet and study with Brunner. He was successful in both fields. I found out later that he was an experimentalist, publishing 23 papers while in Germany. Then, and later, he published books in philosophy.

If Hitler had not come to power, I might not be here. My father would have stayed in Germany and become a professor in Berlin. But when Hitler expelled the foreign Jews, my father returned to Bucharest, married my

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102

1 mother, and I came to be. My mother was seriously ill
 2 during the first few years of my life, but I had a loving
 3 set of grandparents and a nurse who took care of me
 4 and, as far as I can remember, was happy.

5 I was eleven when my father died. By then we were
 6 living in Canada. Although he was very ill with heart
 7 disease, he was studying very hard to qualify as a doc-
 8 tor in Canada. My relations with him were never easy.
 9 He kept a notebook of anecdotes about me from ages
 10 one to five. When I translated it for my wife Nancy,
 11 I saw as I read that the anecdotes are all instances in
 12 which the father humiliates the child. After he died,
 13 I tried to help my mother, in the house, and by taking a
 14 job delivering for the local drugstore. She coped won-
 15 derfully even though her life in Romania had not pre-
 16 pared her for the role of relatively poor single mother.
 17 Our relations were very close. I would act as confidante
 18 and counselor and was very proud of what help I could
 19 give. She remarried, and that's how we got to Califor-
 20 nia.

21 I found school work and languages very easy as a
 22 child, too easy, as I discovered when I got to Caltech
 23 and had to compete with many who were as quick or
 24 quicker than I was. I wanted to be a scientist ever since
 25 I read the books of George Gamow. But I was broadly
 26 interested in physics on the one hand and physiolo-
 27 gy and biochemistry on the other. I avoided medicine
 28 and philosophy, since my father had been a physician
 29 and philosopher. Fortunately, I found my way through
 30 mathematics to statistics, which has allowed me to dip
 31 into almost every science.

32 *Your uncle was a lawyer?*

33 In Romania my uncle Shlomo Bickel was a lawyer,
 34 but he, like my father, and most of his generation of
 35 Jews, was part of a movement. He was a Yiddishist
 36 and a Zionist. He was able to get out to the States in
 37 1938. He couldn't practice law, so became a journalist
 38 and wrote weekly columns for *The Day*, one of the two
 39 large Yiddish papers in New York. He also wrote sev-
 40 eral books; one chapter of his book "Rumania," later
 41 translated into English, was about my father and other
 42 rebellious Bickels. I felt very close to my uncle Shlomo
 43 and aunt Yetta. Their household was full of intellectual
 44 and literary discussion. They showed great affection to
 45 each other and to me, particularly after my father died.
 46 They showed me how loving and intellectually lively
 47 family life could be. Like my uncle, I've been fortu-
 48 nate to have a family life full of love and discussion.

49 **ACKNOWLEDGMENTS**

50 The encouragement of Nancy Bickel was more than
 51 helpful both to Peter and me. The pictures collection

was done by her. The pictures were taken by friends,
 family, and students and staff in Berkeley and Prince-
 ton. I apologize that I cannot give personal acknowl-
 edgments for them.

REFERENCES

ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R. (1976). As-
 ymptotic expansions for the power of distribution free tests in
 the one-sample problem. *Ann. Statist.* **4** 108–156. [MR0391373](#)
 ANDREWS, D. F., BICKEL, P. J., HAMPEL, F. R., HUBER, P. J.,
 ROGERS, W. H. and TUKEY, J. W. (1972). *Robust Estimates of*
Location: Survey and Advances. Princeton Univ. Press, Prince-
 ton, NJ. [MR0331595](#)
 BICKEL, P. J. (1964). On some alternative estimates for shift in
 the p -variate one sample problem. *Ann. Math. Statist.* **35** 1079–
 1090. [MR0165624](#)
 BICKEL, P. J. (1965a). On some asymptotically nonparametric
 competitors of Hotelling's T^2 . *Ann. Math. Statist.* **36** 160–173;
 correction, *ibid.* **36** 1583. [MR0181052](#)
 BICKEL, P. J. (1965b). On some robust estimates of location. *Ann.*
Math. Statist. **36** 847–858. [MR0177484](#)
 BICKEL, P. J. (1967). Some contributions to the theory of order
 statistics. In *Proc. Fifth Berkeley Sympos. Math. Statist. and*
Probability (Berkeley, Calif., 1965/66), Vol. I: Statistics 575–
 591. Univ. California Press, Berkeley, CA. [MR0216701](#)
 BICKEL, P. J. (1970). A Hájek–Rényi extension of Lévy's inequal-
 ity and some applications. *Acta Math. Acad. Sci. Hungar.* **21**
 199–206. [MR0261653](#)
 BICKEL, P. J. (1974). Edgeworth expansions in non parametric sta-
 tistics. *Ann. Statist.* **2** 1–20. [MR0350952](#)
 BICKEL, P. J. (1982). On adaptive estimation. *Ann. Statist.* **10** 647–
 671. [MR0663424](#)
 BICKEL, P. J. (1983). Minimax estimation of the mean of a normal
 distribution subject to doing well at a point. In *Recent Advances*
in Statistics 511–528. Academic Press, New York. [MR0736544](#)
 BICKEL, P. J. (1984). Parametric robustness: Small biases can be
 worthwhile. *Ann. Statist.* **12** 864–879. [MR0751278](#)
 BICKEL, P. J. and BREIMAN, L. (1983). Sums of functions of
 nearest neighbor distances, moment bounds, limit theorems and
 a goodness of fit test. *Ann. Probab.* **11** 185–214. [MR0682809](#)
 BICKEL, P. J., GÖTZE, F. and VAN ZWET, W. R. (1985). A simple
 analysis of third-order efficiency of estimates. In *Proceedings*
of the Berkeley Conference in Honor of Jerzy Neyman and Jack
Kiefer, Vol. II (Berkeley, Calif., 1983) 749–768. Wadsworth,
 Belmont, CA. [MR0822063](#)
 BICKEL, P. J., GÖTZE, F. and VAN ZWET, W. R. (1986). The
 Edgeworth expansion for U -statistics of degree two. *Ann. Sta-*
tist. **14** 1463–1484. [MR0868312](#)
 BICKEL, P. J., GÖTZE, F. and VAN ZWET, W. R. (1995). Resam-
 pling fewer than n observations: Gains, losses, and remedies for
 losses. Empirical Bayes, sequential analysis and related topics
 in statistics and probability (New Brunswick, NJ, 1995). *Statist.*
Sinica **7** (1997) 1–31 [MR1441142](#)
 BICKEL, P. J., HAMMEL, E. A. and O'CONNELL, J. W. (1975).
 Sex bias in graduate admissions: Data from Berkeley. *Science*
187 398–404.
 BICKEL, P. J., KLAASSEN, C. A. J., RITOV, Y. and WELLNER,
 J. A. (1993 & 1998). *Efficient and Adaptive Estimation for*

52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102

1	<i>Semiparametric Models. Johns Hopkins Series in the Mathematical Sciences.</i> Johns Hopkins Univ. Press, Baltimore, MD, 1993. Springer, New York, 1998. MR1623559		
2		BICKEL, P. J. and YAHAV, J. A. (1969a). Some contributions to the asymptotic theory of Bayes solutions. <i>Z. Wahrsch. Verw. Gebiete</i> 11 257–276. MR0242298	52
3			53
4	BICKEL, P., LI, B. and BENGTTSSON, T. (2008). Sharp failure rates for the bootstrap particle filter in high dimensions. In <i>Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh</i> 3 318–329. IMS, Beachwood, OH. MR2459233		54
5		BICKEL, P. J. and YAHAV, J. A. (1969b). On an A.P.O. rule in sequential estimation with quadratic loss. <i>Ann. Math. Statist.</i> 40 417–426.	55
6			56
7	BICKEL, P. J. and RITOV, Y. (1987). Efficient estimation in the errors in variables model. <i>Ann. Statist.</i> 15 513–540. MR0888423	HENGARTNER, N., TALBOT, L., SHEPHERD, I. and BICKEL, P. (1995). Estimating the probability density of the scattering cross section from Rayleigh scattering experiments. <i>J. Opt. Soc. Am. A</i> 12 1316–1323.	57
8			58
9	BICKEL, P. J. and ROSENBLATT, M. (1973). On some global measures of the deviations of density function estimates. <i>Ann. Statist.</i> 1 1071–1095. MR0348906		59
10			60
11	BICKEL, P. J. and YAHAV, J. A. (1967). Asymptotically pointwise optimal procedures in sequential analysis. In <i>1967 Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. I: Statistics</i> 401–413. Univ. California Press, Berkeley, CA. MR0212966	KECHRIS, K., LIN, J. C., BICKEL, P. J. and GLAZER, A. N. (2006). Quantitative exploration of the occurrence of lateral gene transfer by using nitrogen fixation genes as a case study. <i>Proc. Natl. Acad. Sci. USA</i> 103 9584–9589.	61
12			62
13		MEINSHAUSEN, N., BICKEL, P. and RICE, J. (2009). Efficient blind search: Optimal power of detection under computational cost constraints. <i>Ann. Appl. Statist.</i> 3 38–60.	63
14			64
15	BICKEL, P. J. and YAHAV, J. A. (1968). Asymptotically optimal Bayes and minimax procedures in sequential estimation. <i>Ann. Math. Statist.</i> 39 442–456. MR0224219	SNYDER, C., BENGTTSSON, T., BICKEL, P. and ANDERSON, J. (2008). Obstacles to high-dimensional particle filtering. <i>Monthly Weather Review</i> 136 4629–4640.	65
16			66
17			67
18			68
19			69
20			70
21			71
22			72
23			73
24			74
25			75
26			76
27			77
28			78
29			79
30			80
31			81
32			82
33			83
34			84
35			85
36			86
37			87
38			88
39			89
40			90
41			91
42			92
43			93
44			94
45			95
46			96
47			97
48			98
49			99
50			100
51			101
			102

THE LIST OF SOURCE ENTRIES RETRIEVED
FROM MATHSCINET

The list of entries below corresponds to the Reference section of your article and was retrieved from MathSciNet applying an automated procedure. Please check the list and cross out those entries which lead to mistaken sources. Please update your references entries with the data from the corresponding sources, when applicable. More information can be found in the support page:
<http://www.e-publications.org/ims/support/mrhelp.html>.

ALBERS, W., BICKEL, P. J., AND VAN ZWET, W. R. (1976). Asymptotic expansions for the power of distribution free tests in the one-sample problem. *Ann. Statist.* **4** 108–156. MR0391373 (52 #12194)

ANDREWS, D. F., BICKEL, P. J., HAMPEL, F. R., HUBER, P. J., ROGERS, W. H., AND TUKEY, J. W. (1972). *Robust estimates of location: Survey and advances*. Princeton University Press, Princeton, N.J. MR0331595 (48 #9927)

BICKEL, P. J. (1964). On some alternative estimates for shift in the p -variate one sample problem. *Ann. Math. Statist.* **35** 1079–1090. MR0165624 (29 #2904)

BICKEL, P. J. (1965). On some asymptotically nonparametric competitors of Hotelling’s T^2 . *Ann. Math. Statist.* **36** (1965), 160–173; correction, *ibid.* **36** 1583. MR0181052 (31 #5281)

BICKEL, P. J. (1965). On some robust estimates of location. *Ann. Math. Statist.* **36** 847–858. MR0177484 (31 #1747)

BICKEL, P. J. (1967). Some contributions to the theory of order statistics. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. I: Statistics*. Univ. California Press, Berkeley, Calif., 575–591. MR0216701 (35 #7530)

BICKEL, P. J. (1970). A Hájek-Rényi extension of Lévy’s inequality and some applications. *Acta Math. Acad. Sci. Hungar.* **21** 199–206. MR0261653 (41 #6266)

BICKEL, P. J. (1974). Edgeworth expansions in non parametric statistics. *Ann. Statist.* **2** 1–20. MR0350952 (50 #3444)

BICKEL, P. J. (1982). On adaptive estimation. *Ann. Statist.* **10** 647–671. MR663424 (84a:62045)

BICKEL, P. J. (1983). Minimax estimation of the mean of a normal distribution subject to doing well at a point. In *Recent advances in statistics*. Academic Press, New York, 511–528. MR736544 (86a:62049)

BICKEL, P. J. (1984). Parametric robustness: small biases can be worthwhile. *Ann. Statist.* **12** 864–879. MR751278 (85m:62051)

BICKEL, P. J. AND BREIMAN, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214. MR682809 (85h:60027)

BICKEL, P. J., GÖTZE, F., AND VAN ZWET, W. R. (1985). A simple analysis of third-order efficiency of estimates. In *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, Calif., 1983)*. Wadsworth Statist./Probab. Ser. Wadsworth, Belmont, CA, 749–768. MR822063 (87k:62043)

BICKEL, P. J., GÖTZE, F., AND VAN ZWET, W. R. (1986). The Edgeworth expansion for U -statistics of degree two. *Ann. Statist.* **14** 1463–1484. MR868312 (88a:62042)

BICKEL, P. J., GÖTZE, F., AND VAN ZWET, W. R. (1997). Resampling fewer than n observations: gains, losses, and remedies for losses. *Statist. Sinica* **7** 1–31. Empirical Bayes, sequential analysis and related topics in statistics and probability (New Brunswick, NJ, 1995). MR1441142 (98g:62079)

Not Found!

BICKEL, P. J., KLAASSEN, C. A. J., RITOV, Y., AND WELLNER, J. A. (1998). *Efficient and adaptive estimation for semiparametric models*. Springer-Verlag, New York. Reprint of the 1993 original. MR1623559 (99c:62076)

BICKEL, P., LI, B., AND BENGTTSSON, T. (2008). Sharp failure rates for the bootstrap particle filter in high dimensions. In *Pushing the limits of contemporary statistics: contributions in honor of Jayanta K. Ghosh*. Inst. Math. Stat. Collect., Vol. **3**. Inst. Math. Statist., Beachwood, OH, 318–329. MR2459233

BICKEL, P. J. AND RITOV, Y. (1987). Efficient estimation in the errors in variables model. *Ann. Statist.* **15** 513–540. MR888423 (88g:62132)

BICKEL, P. J. AND ROSENBLATT, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095. MR0348906 (50 #1400)

BICKEL, P. J. AND YAHAV, J. A. (1967). Asymptotically pointwise optimal procedures in sequential analysis. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. I: Statistics*. Univ. California Press, Berkeley, Calif., 401–413. MR0212966 (35 #3831)

BICKEL, P. J. AND YAHAV, J. A. (1968). Asymptotically optimal Bayes and minimax procedures in sequential estimation. *Ann. Math. Statist.* **39** 442–456. MR0224219 (36 #7263)

BICKEL, P. J. AND YAHAV, J. A. (1969). Some contributions to the asymptotic theory of Bayes solutions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **11** 257–276. MR0242298 (39 #3629)

Not Found!

Not Found!

Not Found!

Not Found!

Not Found!

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102

META DATA IN THE PDF FILE

Following information will be included as pdf file Document Properties:

Title : A Random Walk with Drift: Interview with Peter J. Bickel
Author : Ya'acov Ritov
Subject : Statistical Science, 0, Vol.0, No.00, 1-12
Keywords:
Affiliation:

THE LIST OF URI ADRESSES

Listed below are all uri addresses found in your paper. The non-active uri addresses, if any, are indicated as ERROR. Please check and update the list where necessary. The e-mail addresses are not checked - they are listed just for your information. More information can be found in the support page:

http://www.e-publications.org/ims/support/urihelp.html.

- 200 http://www.imstat.org/sts/ [2:pp.1,1] OK
200 http://dx.doi.org/10.1214/09-ST300 [2:pp.1,1] OK
200 http://www.imstat.org [2:pp.1,1] OK
--- mailto:yaacov@mscc.huji.ac.il [2:pp.1,1] Check skip