Writing TPE2

George Casella

As I write this memory of working with Erich Lehmann on the second edition of *Theory of Point Estimation* (TPE2), I am still trying to process the news from Bin Yu about Erich's death. The passing of Erich Lehmann not only signals the passing of a great statistician and a great man, but also the end of an era. Post World War II statistics, from approximately the late 1940s until the middle 1970s, was an era of "mathematical statistics". It was marked by the development of much of the underlying theory, but also marred by a marginalization - some thought that statistics was merely a branch of applied mathematics.

Not Erich. Although his work (some with Scheffé and Hodges) was very mathematical, he strove to solve real statistical problems, and create the underlying theory of statistics: *theoretical statistics*, not *mathematical statistics*. He was never concerned with elegant proofs, or math for math's sake, but to provide sound and understandable solutions to real problems. As an example, take a look at Erich's 1983 JASA paper "Estimation with Inadequate Information", that shows how (and why) unbiasedness can sometimes produce an absurd result.

But to the Story of TPE2. In 1994 I got an email from John Kimmel, Editor of the Springer-Verlag Statistics Textbook Series, asking if I would be interested in working with Erich Lehmann on a revision of *Theory of Point Estimation* (TPE1), which was originally published in 1983. As you might expect, this was an offer that no one could refuse! I talked with Erich on the phone about this (he was at Berkeley and I was at Cornell), and we agreed that I would come to Berkeley to discuss how to proceed. I worked for about two months putting together a plan for what new material to put in, and for what old material to take out. I was working under the assumption that John Kimmel had suggested this collaboration to Erich, and that I would have to convince him to take me on.

G. Casella

Distinguished Professor, Department of Statistics, University of Florida, Gainesville, FL 32611
e-mail: casella@stat.ufl.edu

After meeting with Erich the first time (I took me over a year to get up enough nerve to address him as “Erich”), as I was in the mode of trying to convince him to work with me, he stopped and said, “George, you don’t understand. Kimmel didn’t suggest you to me, I suggested you to him. So just relax.” It turned out that Erich was a fan of *Statistical Inference* (Casella and Berger) and decided that he wanted to have me work on TPE2. Wow.

So now the “experienced” text book writer (me- I had already done two books) really learned how to write a textbook. Erich was adamant that whatever was included in the text had to be done right. That is, the entire story must be told - introduce the problem, show how it can be solved, explain the scope and limitations of the solutions, and provide examples and exercises. A text book must have the whole story. We don’t say, “for details see some paper in the literature”, we give, and explain, the details. It was Erich’s insistence that any good qualified teacher, who may not be totally familiar with the topic, could pick up the textbook and teach a course out of it.

A number of people had suggestions for the second edition. Persi Diaconis told me that one of the charms of TPE1 was that it was not an enormous tome - Erich had done a wonderful job of distilling the important concepts. Persi urged me to not make the second edition too big, or else it would lose its charm. Although that was a little difficult (the second edition grew by 100 pages) another suggestion, by Larry Brown, caused me a bit more difficulty. Larry did not like the way the Information Inequality was stated, and wanted me to completely revise the treatment. The problem, Larry said, is that in the statement of the Information Inequality in TPE1, the conditions were on the estimator (it had to allow differentiation under the integral sign), and this restricted the class. Larry wanted the conditions on the density. I liked this idea, and proposed it to Erich. Erich felt that changing the assumptions made the presentation a bit too difficult, and did not want to do it. So there I was caught in the middle, between Lehmann and Brown. The final result? See Section 2.5, where there are two information inequalities.

In other places the material changed, with perhaps the most significant change being the addition of Chapter 5: Average Risk Optimality, which is all about Bayesian estimation. Erich always wanted the main chapters to reflect optimality, hence the title. Here is one interesting anecdote about this chapter. I wrote a first draft in Ithaca, and sent it to Erich in Berkeley. After a week or so he called me to discuss what I had written, and in his gentle way he informed me that I had gotten it all wrong! What happened was that I had structured the chapter according to distributions, rather than types of estimators. Erich told me that as I rewrote and rearranged everything I would see how things would flow better and, of course, he was right.

Other new material included much on Stein estimation and the accompanying calculations, the addition of complete class theorems, and some material on EM and Gibbs. Erich was not initially in favor of the computational material, but I convinced him that it was necessary so that we could tell the entire story!

In any second edition it is necessary to not only add new material, but also to delete some material from the first edition. Of course, one could just add, but a good textbook is not an encyclopedia, but rather a filter of what the authors believe comprises a thorough treatment of the topic. One major deletion was material on robust estimation, not because it is not important but because with the appearance of some excellent

textbooks on the subject, we could defer this material and make room for the Bayes chapter - which we both wanted to have in the book. Other topics, or sections, were judiciously edited (for example, Section 5.1 in TPE1 is Section 1.8 in TPE 2), since hindsight always allows for a more coherent flow of material.

We worked together for four years, 1994-1998, during which I made about two trips per year to Berkeley, where I always stayed with Erich and his wife Juliet Shaffer, and was treated to good food and good conversation (and hard work). Julie was on sabbatical one year at Princeton, and during that year I could drive from Ithaca to Princeton for more frequent meetings, and one time Erich and Julie came and stayed at my house in Ithaca. (Much later, after I moved to Florida, I convinced Erich and Julie to visit, and Erich gave a talk at the department. As a reward, I took them on a hike to a nature preserve where we saw about fifty alligators.)

TPE2 is still, as I understand, a mainstay in many graduate programs, and has been translated into Chinese. For the most part, the material still fits well into modern graduate programs because it is core material. To do good statistics, whether theory or applications (with the latter always needing some theory) it is necessary to have a solid theoretical core. TPE2 fills that role. Will there be a TPE3? I expect so. As we monitor the change in our field, we see trends that sometimes lead to necessary changes in the core curriculum. When that happens, TPE3 will be there.