Comment

H. F. Trotter

There is little that I can add to Professor Le Cam's survey and analysis of the literature on the central limit theorem, but a few remarks based on personal recollection may be of interest. As a student of Feller's, and later a colleague, I of course got some idea of his opinions on various matters, and heard various anecdotes of his experiences. Unfortunately, I never took advantage of the opportunity to ask questions that might have led to a coherent picture of how things were in the thirties, and I never kept notes on what I did hear. Thus the following remarks, based only on scattered recollections, are subject to all of the inaccuracies that affect undocumented memories after a lapse of over 10 years.

Feller certainly had (and expressed) great admiration for Lévy and the brilliant insights contained in his work, and I think that he would agree that much of Lévy's work was slower to receive recognition than it deserved. I do not, however, recall any mention of possible priority on necessary conditions for the central limit theorem.

It is my impression that Feller regarded his papers of 1935 and 1937 as a high point of his early career, and that he took special pleasure in them because they contradicted the opinion of authorities at the time that no sensible necessary conditions could be established. (Unfortunately, I recall no indication of who the "authorities" were.) This suggests that the question was in the air, so that it would not be surprising if Feller and Lévy were working on it independently. (I have a vaguer impression that Feller claimed some originality in introducing, or at least effectively dealing with, fully arbitrary scaling and location factors. For me, this tends to confirm Professor Le Cam's opinion that Feller had not seen Lévy's work of 1931 and 1934.) I personally do not doubt that both were convinced of their respective priority, and legitimately so in the sense that their work was done completely independently.

I have no idea whether Feller was familiar with Kolmogorov's measure-theoretic foundations of probability (Kolmogorov, 1933) in 1935. From what I remember his telling about that time, "random variables" and the like were not clearly defined entities that could be used in any rigorous discussion—a statement of probability theory had to be cast as a proposition in analysis before it could really be proved, and probabilistic notions had only motivational and heuristic value. It would be interesting to trace how, when, and where this attitude changed. Obviously Feller's papers of 1935 and 1937 were written in traditional style. He might have chosen that style for the sake of being more readily understood, even if his own ideas had already been changed by Kolmogorov (1933), but my guess (based only on a general impression, not on anything he ever explicitly said) is that while he might well have read Kolmogorov (1933) by 1935, it had not yet really changed his way of looking at things. Of course the central limit theorem (at least for independent variables) really is just a statement about convolution of distribution functions that can be viewed directly as a proposition in analysis, unlike, say, the strong law of large numbers, which can be much more naturally expressed in genuinely probabilistic language.

Professor Le Cam is entirely correct in describing my own method of proof of Lindeberg's theorem (referred to in the passage he cites from Feller (1971)) as differing from the original mostly by a change of terminology. As it happens, the idea of the proof for the simple case of identically distributed random variables came naturally from work I had done in my thesis on semigroups of operators, and it was only when I looked up Lindeberg (1922) to see whether the idea could be extended to obtain Lindeberg's more general result that I discovered the essential equivalence of the methods. The only proofs I had seen at that time involved characteristic functions, and the point of my paper was only to show that a little very elementary "soft" analysis could substitute for the (slightly) less elementary "hard" analysis involved in proving that convergence of characteristic functions implies convergence of distributions. It is hard now to see why Lindeberg's paper appeared difficult. Because the notion of linear operator was not yet automatically part of the common vocabulary, he needs several pages to establish basic facts that can now be dealt with in a paragraph, but the arguments are clear and straightforward. Perhaps it was the success of characteristic functions, rather than any real difficulty in the paper by Lindeberg (1922) that led to the latter almost dropping out of sight for so long.

There is one matter on which I disagree with Professor Le Cam. His remarks on Gauss are of course peripheral to the main subject of his paper, and evi-

H. F. Trotter is a Professor in the Department of Mathematics, Princeton University, Princeton, NJ 08544.

dently not meant to be taken completely seriously, but I think he goes too far when he accuses Gauss of circular reasoning. I find no foundation for that in what Gauss actually wrote. From the assumption that the arithmetic mean of repeated observations of a single quantity is the maximum likelihood estimate, he derives the normal distribution, and from that the more general principle of least squares for getting maximum likelihood estimates from observations on a number of related quantities. He asserts that the latter principle should be considered to be just as valid as the principle of the arithmetic mean; he does *not* close the circle by claiming that this in turn justifies the principle of the arithmetic mean. (My use of the term "maximum likelihood" is anachronistic; I am not claiming that Gauss had an exact equivalent of the modern notion clearly in mind.)

Gauss argument involves some tacit assumptions, and one can debate the extent to which he was aware of them; the argument is loose by modern standards, but it is not absurd. The extent to which he claimed to be giving a rigorous proof is not clear to me. (Neither Gauss (1809) nor his later extensive exposition of the method of least squares, which appeared in 1821 and 1823 with a supplement in 1826 (Volume 4 of his collected works, pp. 1–93) is at all a treatise in pure mathematics.) He clearly considered the method of least squares to be of great practical value. He was also interested in justifying it philosophically, but he explicitly stated that it cannot be singled out as the only reasonable method on purely theoretical grounds (see Section 186 of Gauss (1809) and Section 6 of the work on least squares mentioned above.)

Professor Le Cam's dissatisfaction with the performance of the hypothetical referee of Gauss (1809) brings to mind one more conversation with Feller, when he expressed some preference for the old days before the present refereeing system. An author who submitted inferior work for publication then ran this risk of damage to his reputation if it appeared. Of course the old system only worked well when the research community was smaller, the volume of publication was much less, and many papers actually got read by at least a few experts, not just counted by deans evaluating candidates for promotion. Gauss' reputation does not seem to have suffered from either the original publication in 1809 or its reprinting over a century and a half later.

I also wonder whether the influence of Gauss' work among nonmathematicians (astronomers, physicists, surveyors) may have played some role (along with the law of eponymy) in attaching his name to the distribution that bears it.

Comment

J. L. Doob

Le Cam's interesting account can be described with only slight exaggeration as a history of (nonrigorous) early research in probability, of probability texts written by mathematicians ignorant of the subject, and finally of frequently clumsy research published before the writers had digested their own work or consulted that of others. Is such the history of all probability research? Of all mathematical research? Should trade secrets be disclosed?

The following quotations are relevant. Lévy, who plays an important role in Le Cam's account, once remarked to me that reading other mathematicians' research gave him actual physical pain. A well known nonprobabilist mathematician remarked to me that the first time a mathematician publishes a research result the treatment is likely to be both wrong and unreadable; the second time the treatment will be correct but obscure; finally, a third treatment may be both correct and clear.

As a partial explanation of the second remark, and of the sometimes unseemly haste to publish, it must be acknowledged that no matter how much a mathematician admires his own work, the writing of it may finally make it so loathesome to his sight that he hastens to send it off for premature publication.

Influential on the nature and speed of probability research is the fact that probabilists, until about the last 30 years, have labored under the psychological disadvantage that their field was not considered a mathematical discipline by their colleagues, who for one thing did not understand why standard mathematical nomenclature was insufficient, why old concepts had to be rechristened "random variables" and "expectations." Moreover, probability books were full of nonmathematical concepts: dice, gambling houses, Peter, and Paul.

Even as late as the 1930s it was not quite obvious to some probabilists, and it was certainly a matter of

J. L. Doob is Emeritus Professor, Department of Mathematics, University of Illinois, Urbana, IL 61801.