The Annals of Statistics
1979, Vol. 7, No. 3, 479–483

HSU'S WORK IN PROBABILITY

By K. L. Chung
Stanford University

Hsu returned from England to Kunming in the middle of the Sino-Japanese war, apparently in 1940. He gave a long course beginning with the theory of measure and integration, through probability theory, and leading to mathematical statistics. For the first part he based his lectures on Carathéodory's *Vorlesungen über Reelle Funktionen*, for the second on Cramér's *Random Variables and Probability Distributions*. He was a polished and vivacious lecturer with complete notes written out beforehand in a notebook he carried. He enjoyed making a fine point in class. For instance, when he was doing the inversion formula for characteristic functions, he took delight in the fact that one could integrate over a single point by the Lebesgue-Stieltjes integral. He had a tattered copy of Cramér's book with many marginal notes and used to say that it was written in a more difficult way than necessary but contained all that was essential for probability. He was a true virtuoso in the method of characteristic functions. His papers [17], [23], [25], [26], [31] and [35] all showed his fascination for, as well as mastery of, this precious tool. Mathematical literature was hard to come by in Kunming during those years, and some of the old volumes shipped from Peking (then called Peiping) were kept in caves to preserve them from air raids. (We did not actually live in caves, but frequently had to run to them for hours at a stretch under raids or alerts.) Among the books so kept was, for example, Kolmogorov's *Grundbegriffe der Wahrscheinlichkeitsrechnung*. At my request Hsu had this book extracted from the caves, but I remember his saying of it: "This is another kind of mathematics". He was not, of course, a probabilist *per se* at a time when such an appellation hardly existed anywhere in the world; by education and inclination he was more fond of problem-solving than generalities. However he never tried to dissuade others from following their own bent and getting interested in those other kinds of things. Indeed, if he could be drawn into a new topic, he would work at it in earnest and quickly produce something of value. For instance, when Borel's work on what he called "denumerable probabilities" attracted my attention, Hsu took an interest also, which led to the paper [21]. This was an early approach to what became known as "recurrent events" under Feller's development. Papers [23] and [30] were probably written under similar circumstances.

But, of course, Hsu's major contribution to probability and its applications flowed from his consummate skill in manipulating characteristic functions. Let me describe and comment on these papers in their chronological order.

Received June 1978.
AMS 1970 *subject classifications.* Primary 01A70; secondary 60E05.
Key words and phrases. Obituary, probability theory.

Paper [17] is, without doubt, one of Hsu's most important works. He wrote it shortly after A. C. Berry obtained the correct order of magnitude of the error estimate in the classical central limit theorem. Namely, if $\{\xi_n, n \geqslant 1\}$ are independent and identically distributed random variables having distribution $P(x)$ with mean 0 and variance 1, and finite third absolute moment β_3, then for all x the distribution

$$F(x) = \Pr\left\{ \frac{1}{n^{\frac{1}{2}}} \Sigma_{r=1}^{n} \xi_r \leqslant x \right\}$$

differs from the standard normal distribution by less than $A\beta_3/n^{\frac{1}{2}}$, where A is an absolute constant. (This result was also proved by Esseen, but Esseen's paper was published later and not known to Hsu at the time.) Hsu extended Berry's method to give a simpler proof of Cramér's theorem on the asymptotic expansion of $F(x)$ when $P(x)$ is nonsingular. But much deeper and farther reaching are the corresponding results when the sample mean

$$\bar{\xi} = \frac{1}{n} \Sigma_{r=1}^{n} \xi_r$$

is replaced by the sample variance

$$\eta = \frac{1}{n} \Sigma_{r=1}^{n} (\xi_r - \bar{\xi})^2.$$

In Hsu's own words: "so much for the known results for the approximate distribution of $\bar{\xi}$. By a purely formal operational method Cornish and Fisher obtain terms of successive approximation to the distribution of any random variable X with the help of semiinvariants. It is hardly necessary to emphasize the importance of turning Cornish and Fisher's formal result (asymptotic expansion without appraisal of the remainder) into a mathematical theorem of asymptotic expansion which gives the order of magnitude of the remainder. In this paper we achieve this for the simplest function next to $\bar{\xi}$, viz. η." The basic idea is as follows. Let

$$X = \frac{1}{n^{\frac{1}{2}}} \Sigma_{r=1}^{n} \frac{\xi_r^2 - 1}{(\alpha_4 - 1)^{\frac{1}{2}}}, \qquad Y = \frac{1}{n^{\frac{1}{2}}} \Sigma_{r=1}^{n} \xi_r,$$

where α_4 is the fourth moment of $P(x)$. These are related to η by

$$\left(\frac{n}{\alpha_4 - 1} \right)^{\frac{1}{2}} (\eta - 1) = X - \frac{1}{((\alpha_4 - 1)n)^{\frac{1}{2}}} Y^2$$

and the crux of the method is to approximate the distribution of the random vector (X, Y) through its characteristic function. Even a cursory perusal of the proof would make it obvious that this is a task of a higher degree of complexity and not just a routine analogue of the results of Berry and Cramér, since X and Y are highly correlated. To achieve the goal Hsu literally added a new dimension to the method of approximation used before him. He was still teaching the course mentioned above when he obtained the result, and was obviously elated. I recall his saying that he made several previous attempts at the problem and the breakthrough came when he realized that he must attack the bivariate approximation rather than evading the difficulties by shortcuts. Hsu's method is applicable to many other functions commonly used in statistics such as sample moments of higher order and Student's statistic. Some of these were carried out by his students. Although Hsu's paper was cited by Bikjalis in his work on normal approximations, a recent book on the subject failed to mention it. In the course of writing this article I asked the authors about this glaring omission and received the regretful reply that they were not aware of Hsu's paper! This is a serious omission and suggests that the analytic power generated by Hsu's method has yet to be fully explored.

In paper [23] the main result is that for a sequence of independent and identically distributed random variables with zero mean and finite variance, we have for every $\varepsilon > 0$:

$$\Sigma_n P\left\{ \frac{1}{n}|X_1 + \cdots + X_n| > \varepsilon \right\} < \infty.$$

This is an interesting strengthening of the classical strong law of large numbers in the direction of the Borel-Cantelli lemma. The idea of such a result is probably due to Robbins, but the method of proof is vintage Hsu. The problem is reduced, after a convoluted Fourier inversion, to the estimation of several pieces of an integral involving the characteristic function—a remarkable tour de force. Erdös (1949) later sharpened the result and showed that the conditions on the moments are necessary as well as sufficient. (According to Robbins, their original manuscript also contained a proof of the necessity of the conditions which was not published as being too complicated.) Erdös used a combinatorial type of argument. Baum and Katz extended the result to moments of fractional and higher order.

Paper [35] forms Appendix III in the translation of Gnedenko-Kolmogorov (1968) and its story may be worth telling here. Hsu had always been impressed with the necessary and sufficient conditions for the general form of the central limit theorem due to Feller, if only for the adroit use of characteristic functions there. In a letter dated May 12, 1947, from Chapel Hill, he told me that he had just solved the same problem for all symmetric stable laws. He wrote: "I believe that in a domain of the weak laws this is the best theorem yet. The method in this work also points to solution of the most general problem of weak law, namely [when the limit is] an infinitely divisible law. But further difficulty is great. . . . The problem in my hand is so natural that I am in constant worry of clashing with people. So, whenever you communicate with probability people, please spread the news." A latter dated May 26, 1947, announced the final result to me. It gives the necessary and sufficient conditions under which the row sums of a triangular array of infinitesimal random variables, independent in each row, converge in distribution to a given infinitely divisible distribution. His conditions differ somewhat from those in Gnedenko (1944) in that the "two tails" are combined. The proof is direct and makes use of the kernel $(\sin t/t)^3$. The complete manuscript containing this result and its self-contained proof is paper [35]. It was given to me shortly before Hsu's departure for China in the summer of 1947. (I have erroneously placed the year to be 1946 in the preface to the translation cited above.) Despite his premonition, Hsu did not learn of Gnedenko's 1944 paper until later. In the only letter he wrote me after his return to China, dated March 19, 1950, he acknowledged Gnedenko's priority and urged me in most emphatic terms to keep the only copy of his paper for him. He had clearly staked his own strength on solving this challenging problem, the culmination of a series of papers by such names as Lévy, Khintchine, Kolmogorov, Marcinkiewicz (whom Hsu met), Feller and Gnedenko. He toiled hard and triumphed quickly. Does it matter that another striver has climbed the same peak before? Hsu's method is direct and "starts from nothing". The last three words give a free rendering of a more elusive Chinese saying which Hsu used on another occasion, meaning minimal reliance on precedents. He had always considered this nonreliance as a virtue of mathematical work, and it is apparent in all his papers where use of previous results for intermediate steps are shunned so far as possible. The new edition of the translation of Gnedenko-Kolmogorov (1968) appeared in the middle of the cultural revolution in China; I do not know if Hsu ever saw this paper in print.

Papers [25] and [26] were reviewed in some detail by me in the *Mathematical Reviews*, Vol. 17 (1956), page 274. Since this is easily available, I need not repeat the contents here. Suffice it to say that they are again products of his expertise with characteristic functions. Paper [31] falls in the same category but apparently was never reviewed. In it it is proved that every distribution in the class L is absolutely continuous. So far as I know this was a new result. It was rediscovered by Fisz and Varadarajan in 1963 (*Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, Vol. 1).

Paper [30] stands out in Hsu's work in probability as it has nothing to do with a characteristic function. Like [31], it is published in Chinese with an English abstract, appeared in the same issue of the journal, and apparently was never reviewed. It deals with the differentiability properties of the transition probability function of a homogeneous Markov process whose state space is Euclidean and whose sample paths are purely jumping. Such processes were studied by D. G. Kendall in 1955, who analysed the differentiation of the transition probability function $P(t, x, E)$ at $t = 0$. In the case of Markov chains, where the state space is discrete, the term-by-term differentiability of the Chapman-Kolmogorov equations:

$$P'(s + t) = P'(s)P(t), \qquad s > 0,$$
$$P'(s + t) = P(s)P'(t), \qquad t > 0,$$

was proved by Austin analytically, and by myself probabilistically (for references to D. G. Kendall and Austin see Chung (1967).) Hsu extended these equations to the Euclidean case, "using more elementary methods and obtaining more precise results". For the first equation above he actually derived an integral representation of $P(t, x, E)$ in terms of the post-exit process, thus generalizing my approach but by purely analytic means. The intimate relation between the analytic and stochastic structures is implicit in his work, and can easily be made explicit.

I did not learn of Hsu's death until my arrival in Peking in June 1972. Hsiao-fu Tuan told me that Hsu spent his last years re-doing a lot of classical combinatorial analysis in his own way and left copious notes. So far the contents of these notes remain unknown.

REFERENCES

CHUNG, K. L. (1967). *Markov Chains With Stationary Transition Probabilities, 2nd ed.* Springer-Verlag.

ERDÖS, P. (1949). On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20** 286–291.

GNEDENKO, B. V. (1944). Limit theorems for sums of independent random variables. *Uspehi Mat. Nauk.* **10** 115–165. English translation: Translation No. 45, *Amer. Math. Soc.*, New York, 1951.

GNEDENKO, B. V. and KOLMOGOROV, A. N. (1968). *Limit Distribution of Sums of Independent Random Variables*, rev. ed. (translated by K. L. Chung, with Appendices by J. L. Doob and P. L. Hsu). Addison-Wesley.

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305