BOOK REVIEWS

Correspondence concerning reviews should be addressed to the Book Review Editor, Professor William Kruskal, Department of Statistics, University of Chicago, Chicago 37, Illinois.

A. RÉNYI, *Wahrscheinlichkeitsrechnung, mit einem Anhang über Informationstheorie*, VEB Deutscher Verlag der Wissenschaften, Berlin 1962. Band 52, Hochschulbücher für Mathematik. xi + 547 pp. DM 55.

LEOPOLD K. SCHMETTERER

University of Vienna

This book has its origin in a series of lectures which the author gave at the University of Budapest, beginning in 1948. These lectures were first published in 1954 as a book written in the Hungarian language, a book known to me by reviews only. I think this knowledge is sufficient to assert that the Hungarian and the present German edition differ in several ways. The former contains a chapter on Mathematical Statistics and another one on stochastic processes, both of which are not included in the latter book. On the other hand an appendix on information theory was added to the German edition. It may be that someone knowing the Hungarian edition would especially miss the chapter about stochastic processes. I think the German edition has been improved by the changes.

The present form of Rényi's book reflects his particular mathematical interests. It appears that the author is not only an expert in probability theory but also in many fields of analysis and in the theory of numbers. Applications of probability theory, especially in physics and chemistry, also belong to the sphere of interests of the author. This wide scope of interests is reflected in the numerous problems that are added to each chapter. Some of these problems are very simple exercises, while others are much more difficult and are accompanied by hints. The reader of the book needs prerequisites; some knowledge in the theory of real and complex functions, and the basic facts of Lebesgue measure and integration.

The first short chapter offers an introduction to the algebra of events, with some emphasis on finite Boolean algebras. It is of importance for the foundation of modern probability that each Boolean algebra is isomorphic to an algebra of subsets of a set. This representation theorem of M. Stone is given here, with a proof that goes back to O. Frink. The next chapter deals with the idea of probability. Some examples explain this idea from the intuitive point of view. Kolmogorov's general axiomatic approach is given, after a short treatment of finite probability algebras and of simple combinatorial methods. This chapter also contains some classical examples of the so-called geometrical probabilities. (The connection with Blaschke's integral geometry is indicated by a reference to Blaschke's book only.) The chapter concludes with Rényi's axiomatic approach to conditional probability algebras. Let Ω be a nonempty set and $\mathfrak{A}$ a σ-algebra of subsets of Ω. Let $\mathcal{B}$ be a nonempty otherwise arbitrary subset of $\mathcal{A}$. Let $(A, B) \to P(A \mid B)$ be a map from $\mathcal{A} \times \mathcal{B}$ into the nonnegative real numbers with the following properties: $P(B \mid B) = 1, B \in \mathcal{B}$; $A \to P(A \mid B)$ is a measure on $\mathcal{A}$ for each $B \in \mathcal{B}$; $P(A \mid B) = P(AB \mid C)/P(B \mid C)$ for each $A \in \mathcal{A}$ and B, $C \in \mathcal{B}$ with $B \subseteq C$, where $P(B \mid C) > 0$. [This last condition is erroneously omitted.] The map $(A, B) \to P(A \mid B)$ is called conditional probability in the sense of Rényi (i.s.R.). There are important examples showing the existence of a not necessarily bounded measure μ defined over $(\Omega, \mathcal{A})$ such that the conditional probability i.s.R. is given by $\mu(A \cap B)/\mu(B)$ for all $(A, B) \in \mathcal{A} \times \mathcal{B}$. This is of interest, e.g., in applications in physics.

The third chapter is concerned with discrete random variables, and the fourth with general random variables. These two chapters contain the standard materials found in every book on probability at a high mathematical level. The fifth chapter opens with the definition of random variables on conditional probability algebras i.s.R. Then the ideas of general conditional probability and of conditional expectation (in the sense of Kolmogorov and i.s.R.) are introduced, using the theorem of Radon-Nikodym. (This last theorem is mentioned without proof.) The conditional probability i.s.R. is of course a random variable i.s.R. The usefulness of this last idea is shown by a nice deduction of Maxwell's velocity law. The correlation coefficient and other measures for the degree of dependence of random variables are examined very carefully. Let us mention the following result: Let $\psi(\xi, \eta)$ be the maximal correlation of the random variables ξ, η and let $K_\xi(\eta)$ be the correlation ratio of η and ξ in the classical sense of Karl Pearson. Let $\{\xi_n\}$ be a finite or infinite sequence of random variables. Let $A = \sup_{\Sigma x_n^2 \leqq 1} \sum_{n,m} \psi(\xi_n, \xi_m) x_n x_m$. Then $\sum_n K_{\xi_n}^2(\eta) \leqq A$. The method of proof is an interesting probabilistic generalization of the large sieve of Linnik in the theory of numbers. The chapter concludes with the consistency theorem of Kolmogorov.

Characteristic functions form the subject of the next chapter. It contains in part standard material, but it is also concerned with theorems about the characterization of the normal distribution. Some of these theorems may hardly be found in textbooks. Let us mention the generalization of Cramér's theorem by Linnik and Singer, and the theorem of Darmois-Skitowitsch. This last theorem can be deduced from the above mentioned theorem of Linnik and Singer. A proof of this is given according to Linnik. It is a basic theorem in mathematical statistics that the mean and standard deviation of random samples from a normal population are independent. The inverse theorem has been proved under conditions of increasing generality by several scholars. Rényi chooses the formulation of E. Lukacs, which assumes the existence of the variance. This chapter also contains the definition of stable and infinitely divisible laws. The formula of Lévy-Khintchine is stated without proof. A paragraph on characteristic functions of distributions on conditional probability algebras i.s.R. is of interest. The characteristic function of such distributions are Fourier transforms of distributions or generalized functions introduced for the first time by L. Schwartz. Rényi deals with the definition of generalized functions as classes of sequences of infinitely often differentiable functions, according to Mikusinki, Lighthill and Temple. All definitions and proofs needed for the above mentioned purpose are given in the book.

The seventh and eighth chapter are concerned with laws of large numbers and with central limit theorems. The weak laws of large numbers are formulated under slightly more general conditions than is usual, the independence of the random variables being replaced by pairwise independence. A short proof is given for a somewhat weaker form of the law of the iterated logarithm: P (lim sup $[\eta_n/(2n \log \log n)^{\frac{1}{2}}] \leqq 1) = 1$ instead of P (lim sup $[\eta_n/(2n \log \log n)^{\frac{1}{2}}] = 1) = 1$. The seventh chapter also contains the definition and basic properties of (strongly) mixing sequences of sets. Here the reader needs a knowledge of the elements of Hilbert space theory. The theory of mixing sequences is used in the following chapter for some generalizations of limit theorems. The chapter on the law of large numbers concludes with Kolmogorov's three series criterion and with some hints concerning the laws of large numbers on conditional probability algebras i.s.R. The central limit theorem is proven in the eighth chapter under the general assumptions of Lindeberg-Feller. It is mentioned too that these conditions are necessary, but without proof. The well known local limit theorem by Gnedenko may also be found in this chapter.

Furthermore, the domain of attraction for the normal distribution is determined according to Khintchine, Feller and Lévy. Only half of the proof of this criterion is carried out. The proof that the criterion is necessary is omitted. This chapter also contains a limit theorem by Erdös and Rényi which generalizes the well known limit theorem by Wald and Wolfowitz which is frequently used in the large sample theory of nonparametric tests. The Wald-Wolfowitz theorem has also been considered by Noether and Hoeffding, but the formulation of Erdös and Rényi is the most general. The convergence to a normal law is assured under Lindeberg-like condition. There are two more sections whose contents are based largely on the author's own work. One is concerned with limit theorems for sums of independent random variables where the number of terms is itself random. The other one is concerned with Rényi's method concerning limit theorems for order statistics. This chapter also contains some material which may be called an introduction to the combinatorial method in probability theory. Results of the Smirnov-Kolmogorov type are presented. The classical results by Pólya on random walks is proven and the Arcsin law is deduced. Sparre Andersen's well known result is given without proof. Spitzer's investigations are mentioned by reference. Only the voluminous eighth chapter contains some facts about Markov chains with finitely many states.

The appendix on information theory is mainly concerned with the axiomatic characterization of several measures of information. Rényi himself has defined information of order α, $0 < \alpha < 1$. The measure of information of Shannon-Wiener is contained in Rényi's definition as a limit when α goes to 1.

The statistical interpretation of the measure of information is illustrated by the simplest coding theorem for noiseless channels. The idea of capacity is not even mentioned. Linnik seems to be the first who has given an information-theoretical proof of the Lindeberg limit theorem; an illustration of this interesting method is given in the appendix. The ergodicity of homogeneous Markov chains is proven by the help of the information of order α.

The book concludes with 20 pages of useful tables and numerous references.

The book is a masterpiece from the didactic point of view. It is very well organized, and the style is lucid. Compared with other standard books in this field, like Feller's and Loève's, it has a position between these two books but nearer to the first than to the second. As far as the selection of topics is concerned it is impossible to satisfy all individual wishes, since the wealth of material in modern probability is overwhelming today. Even a book of 550 pages must reflect the personal tastes of the author. It would perhaps be desirable to treat the theory of Markov chains in somewhat more detail, or that something be said about martingales, whose importance in other branches of mathematics is steadily increasing. But such objections are, as mentioned above, more or less trivial. The book seems to contain only a very small number of misprints and small errors. Let us mention a few of them. On p. 40 Theorem 3, the condition $\mu(B_n) \neq \infty$, $n \geqq n_0$ for some n_0, is missing. Formula (2), p. 165, is correct only almost everywhere. On p. 223, third line from the top, the symbol $P(B \mid CA(\xi))$ is not defined well enough.

This excellent book will certainly attain a widespread diffusion, not only in German-speaking countries, but wherever there is interest in probability theory.