THE ELEMENTS OF PROBABILITY THEORY and some of its applications

HARALD CRAMÉR

HARALD CRAMÉR

THE ELEMENTS OF PROBABILITY THEORY

AND SOME OF ITS APPLICATIONS

WILEY PUBLICATIONS IN STATISTICS

Walter A. Shewhart, Editor

Mathematical Statistics

CRAMÉR · The Elements of Probability Theory and Some of Its Applications
SAVAGE · Foundations of Statistics
BLACKWELL and GIRSHICK · Theory of Games and Statistical Decisions
HANSEN, HURWITZ, and MADOW · Sample Survey Methods and Theory, Volume II
DOOB · Stochastic Processes
RAO · Advanced Statistical Methods in Biometric Research
KEMPTHORNE · The Design and Analysis of Experiments
DWYER · Linear Computations
FISHER · Contributions to Mathematical Statistics
WALD · Statistical Decision Functions
FELLER · An Introduction to Probability Theory and Its Applications, Volume I
WALD · Sequential Analysis
HOEL · Introduction to Mathematical Statistics, *Second Edition*

Applied Statistics

BUSH and MOSTELLER · Stochastic Models for Learning
FRYER · Elements of Statistics
BENNETT and FRANKLIN · Statistical Analysis in Chemistry and the Chemical Industry
COCHRAN · Sampling Techniques
WOLD and JURÉEN · Demand Analysis
HANSEN, HURWITZ, and MADOW · Sample Survey Methods and Theory, Volume I
CLARK · An Introduction to Statistics
TIPPETT · The Methods of Statistics, *Fourth Edition*
ROMIG · 50-100 Binomial Tables
GOULDEN · Methods of Statistical Analysis, *Second Edition*
HALD · Statistical Theory with Engineering Applications
HALD · Statistical Tables and Formulas
YOUDEN · Statistical Methods for Chemists
MUDGETT · Index Numbers
TIPPETT · Technological Applications of Statistics
DEMING · Some Theory of Sampling
COCHRAN and COX · Experimental Designs
RICE · Control Charts
DODGE and ROMIG · Sampling Inspection Tables

Related Books of Interest to Statisticians

ALLEN and ELY · International Trade Statistics
HAUSER and LEONARD · Government Statistics for Business Use

THE ELEMENTS OF PROBABILITY THEORY

AND SOME OF ITS APPLICATIONS

By

HARALD CRAMÉR

PROFESSOR OF ACTUARIAL MATHEMATICS
AND MATHEMATICAL STATISTICS,
UNIVERSITY OF STOCKHOLM

JOHN WILEY & SONS, NEW YORK
ALMQVIST & WIKSELL, STOCKHOLM

SECOND PRINTING, JUNE, 1955

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

This book is a revised and extended translation of a Swedish textbook which was published first in 1926 and then in entirely rewritten form in 1949. Starting with a historical introduction to the subject, the book covers the elements of the mathematical theory of probability, with the main emphasis on the theory of random variables and probability distributions. Applications to various fields, particularly to modern statistical methods, are discussed and illustrated by a number of examples. The problems offered for the reader's solution include simple exercises as well as important complements to the theories and methods given in the text.

The book is essentially an *elementary* treatise, and does not aim at a complete and rigorous mathematical development of the subject from an axiomatic point of view. In this respect it can only serve as an introduction to more advanced treatises, such as FELLER's *Probability Theory*, or the present author's *Random Variables and Probability Distributions* and *Mathematical Methods of Statistics*. Occasionally, for the complete proof of some theorem, reference will be made to one of the two last-mentioned works, which will be briefly quoted as *Random Variables* and *Mathematical Methods*.

As for the applications of mathematical probability, any book of moderate size must be content with offering a small sample of problems and methods from the immense field of existing applications, which is continually expanding. Any selection of such a sample by an individual author will necessarily be influenced by a strong personal bias, and most certainly this rule will be found to apply here. Although Part III has been named "Applications", a number of important applications will also be found among the examples and problems that illustrate the theoretical developments in Parts I and II.

It will be assumed that the reader has some working knowledge of analytic geometry, calculus, and algebra, including determinants. Probably the book could be read without difficulty by a student at the junior level in an average American college. Some parts are, of course, more elementary than the others: in particular the first part, which could no doubt be read at an even earlier stage.

My sincere thanks are due to Professor EDWIN HEWITT of the University of Washington, who kindly translated a part of Chapter 16, and to Fil.

Lic. Gunnar Blom, who supplied statistical material for some of the examples. I am further indebted to Professor R. A. Fisher and to Messrs. Oliver and Boyd for permission to reprint the tables of the t and χ^2 distributions from *Statistical Methods for Research Workers*, to Professor George W. Snedecor and to the Iowa State College Press for permission to reprint the tables of the F distribution from *Statistical Methods*, and finally to the Princeton University Press for permission to use some of the diagrams from my own *Mathematical Methods*.

University of Stockholm, September 1954.

H. C.

CONTENTS

Part I

FOUNDATIONS

Part II

RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

Part III

APPLICATIONS

PART I

FOUNDATIONS

CHAPTER 1

HISTORICAL INTRODUCTION

1.1. Origin of mathematical probability theory. — In order to acquire a real understanding of any branch of science it is necessary to study its historical development. All those problems for which ready-made solutions are offered in modern textbooks were once as many unsolved difficulties facing the scientists of bygone times. The long struggle of past generations against all these difficulties, which has finally led up to our modern science, is certainly worth studying, and the labor spent on such a study will be amply rewarded by the insight into the organic growth and structure of science thus obtained.

The particular branch of science with which we are going to deal in this book is no exception to the general rule. In fact, current literature on mathematical probability theory shows many features that can be fully understood only by a reader who knows something about the historic development of this theory. In this introductory chapter we shall give a very brief sketch of the most essential stages of this development. For a more complete account of the origin and early history of probability theory, the reader may be referred to the classical work by TODHUNTER (see list of references p. 265).

The theory of probability, which at the present day is an important branch of pure mathematics, with a field of applications extending over practically all branches of natural, technical, and social science, has developed from a very humble beginning. Its roots lie in a simple mathematical *theory of games of chance* which was founded about three centuries ago.

In the French society of the 1650's, gambling was a popular and fashionable habit, apparently not too much restricted by law. As ever more complicated games with cards, dice, etc., were introduced, and considerable sums of money were at stake in gambling establishments, the need was felt for a rational method for calculating the chances of gamblers in various games. A passionate gambler, the chevalier DE MÉRÉ, had the idea of consulting the famous mathematician and philosopher BLAISE PASCAL in Paris on some questions connected with certain games of chance, and this gave rise to a correspondence between PASCAL and some of his mathematical friends, above all PIERRE FERMAT in Toulouse.

This correspondence forms the origin of modern probability theory. During the rest of the seventeenth century, the questions raised by DE MÉRÉ and further questions of the same kind were discussed among mathematicians. At this early phase of the development, no connected theory of probability had yet been worked out, and the whole subject consisted of a collection of isolated problems concerning various games. In later chapters of this book we shall make detailed acquaintance with these problems. In the present section we shall only make some remarks concerning their general nature.

In all current games of chance with dice, cards, roulettes, and other such apparatus, every single performance of the game must lead to one of a definite number of possible outcomes, represented by the six sides of the die, the 37 cases of the roulette, the 52 cards in an ordinary set of cards, etc. If the gambling apparatus is properly made, and the game is correctly performed, it is not possible to predict in advance which of these possible outcomes will be realized at a particular performance of the game. We cannot predict whether, at the next toss, the coin will fall heads or tails, and similarly in other cases. In fact, this very impossibility of prediction constitutes the *randomness*, the element of uncertainty and chance in the game. On the other hand, there exists between the various possible outcomes of the game a mutual symmetry, which makes us regard all these outcomes as equivalent from the point of view of gambling. In other words, we consider it *equally favourable* for a gambler to risk his stakes on any one of the possible outcomes.

Suppose that we are given a game where this situation is present. Thus every performance of the game will lead to one of a certain number of outcomes, or *possible cases*, and between these possible cases there exists a mutual symmetry of the kind just indicated. Let c denote the total number of these possible cases. Suppose further that, from the point of view of a certain gambler A, the total number c of possible cases can be divided into a group of *favourable cases*, containing a cases, and another group of *unfavourable cases*, containing the remaining $c-a$ cases. By this we mean that, according to the rules of the game, the occurrence of any of the a favourable cases would imply that A wins the game, while the occurrence of any of the $c-a$ unfavourable cases would imply that he loses. If we are interested in estimating the chances of success of A in this game, it then seems fairly natural to consider the ratio a/c between the number a of favourable cases and the total number c of possible cases, and to regard this ratio as a reasonable measure of the chances of gain of A.

Now this is precisely what the classical authors of our subject did. The main difficulty with which they were concerned consisted in the calculation of the numbers a and c of favourable and possible cases in various actual games. As soon as these numbers were known for a given game, their ratio

$$p = \frac{a}{c}$$

was formed. Gradually this ratio came to be known as the *probability* of the *event* which consisted in a gain for the gambler A. This train of thought led to the famous *classical probability definition*, which runs in the following way: *The probability of the occurrence of a given event is equal to the ratio between the number of cases which are favourable to this event, and the total number of possible cases, provided that all these cases are mutually symmetric.*

Though it was not until considerably later that any explicit formulation of a definition of this type appeared, such a definition was more or less tacitly assumed already by PASCAL, FERMAT, and their contemporaries. According to this definition we should say, e. g., that the probability of throwing heads in one toss with a coin is $\frac{1}{2}$, while the probability of obtaining a six in one throw with an ordinary die is $\frac{1}{6}$, the probability of drawing a heart from an ordinary set of 52 cards is $\frac{13}{52} = \frac{1}{4}$, etc.

1.2. Probability and experience. — Already at an early stage, the large mass of empirical observations accumulated in connection with various games of chance had revealed a general mode of regularity which proved to be of the utmost importance for the further development of our subject.

Consider a given game, in each performance of which there are c possible cases which are mutually symmetric in the sense indicated in the preceding section. If this game is repeated under uniform conditions a large number of times, it then appears that *all the c possible cases will, in the long run, tend the occur equally often.* Thus in the long run each possible case will occur approximately in the proportion $1/c$ of the total number of repetitions.

If, e.g., we make a long series of throws with a coin, we shall find that heads and tails will occur approximately equally often. Similarly, in a long series of throws with an ordinary die each of the six sides will occur in approximately $\frac{1}{6}$ of the total number of throws, etc.

If we accept this regularity as an empirical fact, we can draw an important conclusion. Suppose that our gambler A takes part in a game where, in each performance, there are c possible and mutually symmetric cases among which a are favourable to A. Let the game be repeated n times under uniform conditions, and suppose that A wins f times and loses the remaining $n-f$ times. The number f will then be called the *absolute frequency*, or simply the *frequency*, of the event which consists in a gain for A, while the ratio f/n will be denoted as the corresponding *relative frequency* or *frequency ratio*.

Now, if n is a large number, it follows from our fundamental empirical proposition that each of the c possible cases will occur approximately n/c times in the course of the whole series of n repetitions of the game. Since a among these cases are favourable to A, the total number f of his gains should then be approximately equal to an/c. We should thus have, *approximately*, $f=an/c$, or

$$\frac{f}{n}=\frac{a}{c}=p.$$

It thus follows that, according to our empirical proposition, the frequency ratio f/n of A's gains in a long series of games will be approximately equal to the probability p that A wins a game, calculated according to the classical probability definition given in the preceding section.

With a slightly more general formulation, we may express this result by saying that *in the long run any event will tend to occur with a relative frequency approximately equal to the probability of the event.*

Like the classical probability definition, this general principle was not explicitly formulated until a more advanced stage had been reached; already at the time of the chevalier DE MÉRÉ it seems, however, to have been tacitly assumed as an obvious fundamental proposition. As we shall see later (cf. 4.1), one of the questions of DE MÉRÉ was, in fact, directly attached to an application of this general principle to a particular case. In a certain game of chance, DE MÉRÉ had found a lack of agreement between the actually observed frequency ratios of his gains and the value of the corresponding probability of a gain according to his own calculation. It was in order to have this apparent contradiction explained that he consulted PASCAL. However, PASCAL and FERMAT were soon able to show that DE MÉRÉ's calculation of the probability was wrong, and that the correctly calculated probability agreed well with the actually observed frequency ratios, so that no contradiction existed.

1.3. Defects of the classical definition. — The main difficulties encountered at this early stage of probability theory belong to the domain of combinatorial analysis. Starting from certain elementary cases, which are assumed to be completely symmetric — the six sides of the die, the 52 cards of the set, etc. — it is required to combine these according to the rules of some given game so as to form the cases which are possible in that game, always conserving the symmetry between the cases. As soon as we leave the very simplest types of games, this may be a rather intricate task, and it will accordingly often be found that even persons with a good logical training may easily be led into error when trying to solve problems of this kind. Thus it is by no means surprising to find in the early history of our subject a considerable diversity of opinion with respect to the correct way of forming the possible and favourable cases.

As an example of the controversies which occurred in this connection, let us consider a simple game which is closely related to one of the questions of DE MÉRÉ, and which we shall later (4.3) discuss in a more general form.

A and B play heads and tails with a coin which is assumed to be perfectly symmetric, so that each of the two possible outcomes of any throw has a probability equal to $\frac{1}{2}$. The game consists in two throws with the coin. If heads appears in at least one of these throws, A wins, while in the opposite case B wins. What is the probability that A wins the game?

If we denote heads by 1 and tails by 0, the two throws must give one, and only one, of the following four results:

$$00, \quad 01, \quad 10, \quad 11.$$

FERMAT considered these four possible cases as mutually symmetric or, in other words, *equally possible*. Since all these cases, except the first, are favourable to A, he concluded that the probability of A's winning the game is $\frac{3}{4}$. The same result was obtained in a different way by PASCAL.

However, another contemporary mathematician, ROBERVAL, objected that, in the two cases denoted by 10 and 11, A has already won after the first throw, so that it would not be necessary to throw again. Consequently he would recognize only the three possible cases

$$00, \quad 01, \quad 1,$$

and since the last two of these are favourable to A, the probability of A's winning would according to ROBERVAL amount to $\frac{2}{3}$ instead of $\frac{3}{4}$.

Similar objections against the generally accepted rules of probability theory were advanced at a somewhat later stage by D'ALEMBERT. However, it seems fairly obvious that in this way we should lose the perfect symmetry between the cases which is characteristic of the FERMAT solution.

Controversies of this type show that the classical probability definition cannot be considered satisfactory, as it does not provide any criterion for deciding when, in a given game, the various possible cases may be regarded as symmetric or equally possible. However, for a long time this defect of the definition was not observed, and it was not until much later that the question was brought under serious discussion.

1.4. Generalization of the probability concept. — About the year 1700 a period of rapid development begins for the theory of probability. About this time two fundamental works on the subject appeared, written respectively by JAMES BERNOULLI and ABRAHAM DE MOIVRE.

The former, one of the members of the famous Swiss mathematical family of the BERNOULLIS, wrote a book with the title *Ars conjectandi* (The Art of Conjecture), which was published in 1713, some years after the death of its author. In this work we find among other things the important proposition known as the BERNOULLI theorem, by which for the first time the theory of probability was raised from the elementary level of solutions of particular problems to a result of general importance. This theorem, which will be fully proved and discussed in chapter 6 of the present book, supplies the mathematical background of those regularity properties of certain frequency ratios in a long series of repetitions of a given game which were indicated above in 1.2.

DE MOIVRE was a French Huguenot who on account of his religion had left France and lived as a refugee in England. His work, *The Doctrine of Chances*, with the subtitle *A method of calculating the probabilities of events in play*, appeared in three editions, 1718, 1738, and 1756, which shows the general interest attracted by our subject during this time. Among other things we find here the first statement of the general theorem known as the *multiplication rule* of the theory of probability, which will be proved below in 3.3. In the last two editions of this remarkable book we also find the first indications of the *normal probability distribution* (cf. chapter 7), which at a later stage was to play a very important part in the development of the subject.

In the works of BERNOULLI and DE MOIVRE, the theory of games of chance was further developed on the basis of the (more or less tacitly used) classical probability definition, and various combinatorial and other mathematical methods were applied to the theory. It is a characteristic feature of the early history of probability theory that there was at this