

# Probability Theory and Mathematical Statistics

A WILEY PUBLICATION IN MATHEMATICAL STATISTICS

# Probability Theory and Mathematical Statistics

MAREK FISZ

*Late Professor of Mathematics  
New York University*

THIRD EDITION

John Wiley & Sons, Inc.  
New York • London • Sydney

*Published in collaboration with PWN, Polish Scientific Publishers*

AUTHORIZED TRANSLATION FROM THE POLISH.

TRANSLATED BY R. BARTOSZYNSKI

THE FIRST AND SECOND EDITIONS OF THIS BOOK  
WERE PUBLISHED AND COPYRIGHTED IN POLAND UNDER THE TITLE

RACHUNEK PRAWDOPODOBIEŃSTWA I STATYSTYKA  
MATEMATYCZNA

COPYRIGHT © 1963

BY

JOHN WILEY & SONS, INC.

---

*All Rights Reserved*

*This book or any part thereof must not  
be reproduced in any form without the  
written permission of the publisher.*

TRANSLATIONS INTO OTHER LANGUAGES ARE PROHIBITED EXCEPT BY  
PERMISSION OF PWN, POLISH SCIENTIFIC PUBLISHERS

THIRD PRINTING, JANUARY, 1967

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 63-7554

PRINTED IN THE UNITED STATES OF AMERICA

To Olga and Aleksander



## Preface to the English Edition

The opening sentence of this book is “Probability theory is a part of mathematics which is useful in discovering and investigating the regular features of random events.” However, even in the not very remote past this sentence would not have been found acceptable (or by any means evident) either by mathematicians or by researchers applying probability theory. It was not until the twenties and thirties of our century that the very nature of probability theory as a branch of mathematics and the relation between the concept of “probability” and that of “frequency” of random events was thoroughly clarified. The reader will find in Section 1.3 an extensive (although, certainly, not exhaustive) list of names of researchers whose contributions to the field of basic concepts of probability theory are important. However, the foremost importance of Laplace’s *Théorie Analytique des Probabilités*, von Mises’ *Wahrscheinlichkeit, Statistik und Wahrheit*, and Kolmogorov’s *Grundbegriffe der Wahrscheinlichkeitsrechnung* should be stressed. With each of these three works, a new period in the history of probability theory was begun. In addition, the work of Steinhaus’ school of *independent functions* contributed greatly to the clarification of fundamental concepts of probability theory.

The progress in foundations of probability theory, along with the introduction of the theory of characteristic functions, stimulated the exceedingly fast development of modern probability theory. In the field of limit theorems for sums of random variables, a fairly general theory was developed (Khintchin, Lévy, Kolmogorov, Feller, Gnedenko, and others) for independent random variables, whereas for dependent random variables some important particular results were obtained (Bernstein, Markov, Doeblin, Kolmogorov, and others). Furthermore, the theory of stochastic processes became a mathematically rigorous branch of

probability theory (Markov, Khintchin, Lévy, Wiener, Kolmogorov, Feller, Cramér, Doob, and others).

Some ideas on application of probability theory that are now used in mathematical statistics are due to Bayes (estimation theory), Laplace (quality control of drugs), and to Gauss (theory of errors). However, it was not until this century that mathematical statistics grew into a self-contained scientific subject. In order to restrict myself to the names of just a few of the principal persons responsible for this growth, I mention only K. Pearson, R. A. Fisher, J. Neyman, and A. Wald, whose ideas and systematic research contributed so much to the high status of modern mathematical statistics.

At present, the development of the theory of probability and mathematical statistics is going on with extreme intensity. On the one hand, problems in classical probability theory unsolved, as of now, are attracting much attention, whereas, on the other hand, much work is being done in an attempt to obtain highly advanced generalizations of old concepts, particularly by considering probability theory in spaces more general than the finite dimensional Euclidean spaces usually treated. That probability theory is now closely connected with other parts of mathematics is evidenced by the fact that almost immediately after the formulation of the distribution theories by L. Schwartz and Mikusinski their probabilistic counterparts were thoroughly discussed (Gelfand, Itô, Urbanik). Moreover, probability theory and mathematical statistics are no longer simply "customers" of other parts of mathematics. On the contrary, mutual influence and interchange of ideas between probability theory and mathematical statistics and the other areas of mathematics are constantly going on. One example is the relation between analytic number theory and probability theory (Borel, Khintchin, Linnik, Erdős, Kac, Rényi, and others), and another example is the use of the theory of games in mathematical statistics and its stimulating effect on the development of the theory of games itself (v. Neumann, Morgenstern, Wald, Blackwell, Karlin, and others).

Areas of application of probability theory and mathematical statistics are increasing more and more. Statistical methods are now widely used in physics, biology, medicine, economics, industry, agriculture, fisheries, meteorology, and in communications. Statistical methodology has become an important component of scientific reasoning, as well as an integral part of well-organized business and social work.

The development of probability theory and mathematical statistics and their applications is marked by a constantly increasing flood of scientific papers that are published in many journals in many countries.



Having described briefly the position of modern probability theory and mathematical statistics, I now state the main purpose of this book:

1. To give a systematic introduction to modern probability theory and mathematical statistics.
2. To present an outline of many of the possible applications of these theories, accompanied by descriptive concrete examples.
3. To provide extensive (however, not exhaustive) references of other books and papers, mostly with brief indications as to their contents, thereby giving the reader the opportunity to complete his knowledge of the subjects considered.

Although great care has been taken to make this book mathematically rigorous, the intuitive approach as well as the applicability of the concepts and theorems presented are heavily stressed.

For the most part, the theorems are given with complete proofs. Some proofs, which are either too lengthy or require mathematical knowledge far beyond the scope of this book, were omitted.

The entire text of the book may be read by students with some background in calculus and algebra. However, no advanced knowledge in these fields or a knowledge in measure and integration theory is required. Some necessary advanced concepts (for instance, that of the Stieltjes integral) are presented in the text. Furthermore, this book is provided with a Supplement, in which some basic concepts and theorems of modern measure and integration theory are presented.

Every chapter is followed by "Problems and Complements." A large part of these problems are relatively easy and are to be solved by the reader, with the remaining ones given for information and stimulation.

This book may be used for systematic one-year courses either in probability theory or in mathematical statistics, either for senior undergraduate or graduate students. I have presented parts of the material, covered by this book, in courses at the University of Warsaw (Poland) for nine academic years, from 1951/1952 to 1959/1960, at the Peking University (China) in the Spring term of 1957, and in this country at the University of Washington and at Stanford, Columbia and New York Universities for the last several years.

This book is also suitable for nonmathematicians, as far as concepts, theorems, and methods of application are concerned.

\* \* \*

I started to write this book at the end of 1950. Its first edition (374 pages) was published in Polish in 1954. All copies were sold within a few months. I then prepared the second, revised and extended, Polish

edition, which was published in 1958, simultaneously with its German translation. Indications about changes and extensions introduced into the second Polish edition are given in its preface. In comparison with the second Polish edition, the present English one contains many extensions and changes, the most important of which are the following:

Entirely new are the Problems and Complements, the Supplement, and the Sections: 2.7,C, 2.8,C, 3.2,C, 3.6,G, 4.6,B, 6.4,B, 6.4,C, 6.12,D, 6.12,F, 6.15, 8.11, 9.4,B, 9.6,E, 9.9,B, 10.10,B, 10.11,E, 12.6,B, 13.5,E, 13.7,D, 14.2,E, 14.4,D, 15.1,C, 15.3,C, 16.3,D, 16.6, and 17.10,A. Section 8.10 (Stationary processes) is almost entirely new.

Considerably changed or complemented are Sections: 2.5,C, 3.5, 3.6,C, 4.1, 4.2, 5.6,B, 5.7, 5.13,B, 6.2, 6.4,A, 6.5, 6.12,E, 6.12,G, 7.5,B, 8.4,D, 8.8,B, 8.12, 9.1, 9.7, 10.12, 10.13, 12.4,C, 12.4,D, 13.3, 16.2,C.

These changes and extensions have all been made to fulfill more completely the main purpose of this book, as stated previously.

\* \* \*

J. Lukaszewicz, A. M. Rusiecki, and W. Sadowski read the manuscript of the first edition and suggested many improvements. The criticism of E. Marczewski and the reviews by Z. W. Birnbaum and S. Zubrzycki of the first edition were useful in preparing the second edition; also useful were valuable remarks of K. Urbanik, who read the manuscript of the second edition. Numerous remarks and corrections were suggested by J. Wojtyniak, and R. Zasępa (first edition), and by L. Kubik, R. Sulanke, and J. Włoka (second edition). R. Bartoszynski, with the substantial collaboration of Mrs H. Infeld, translated the book from the Polish. J. Karush made valuable comments about the language. Miss D. Garbose did the editorial work. B. Eisenberg assisted me in the reading of the proofs. My sincere thanks go to all these people.

*New York*  
*October, 1962*

MAREK FISZ

# Contents

CHAPTER	PAGE
<b>PART 1    PROBABILITY THEORY</b>	
<b>1</b>	<b>RANDOM EVENTS . . . . . 1</b>
1.1	Preliminary remarks . . . . . 3
1.2	Random events and operations performed on them . . . 5
1.3	The system of axioms of the theory of probability . . . 11
1.4	Application of combinatorial formulas for computing probabilities . . . . . 16
1.5	Conditional probability . . . . . 18
1.6	Bayes theorem . . . . . 22
1.7	Independent events . . . . . 24
	Problems and Complements . . . . . 25
<b>2</b>	<b>RANDOM VARIABLES . . . . . 29</b>
2.1	The concept of a random variable . . . . . 29
2.2	The distribution function . . . . . 31
2.3	Random variables of the discrete type and the continuous type 33
2.4	Functions of random variables . . . . . 36
2.5	Multidimensional random variables . . . . . 40
2.6	Marginal distributions . . . . . 46
2.7	Conditional distributions . . . . . 48
2.8	Independent random variables . . . . . 52
2.9	Functions of multidimensional random variables . . . 56

2.10	Additional remarks . . . . .	62
	Problems and Complements . . . . .	62
3	PARAMETERS OF THE DISTRIBUTION OF A RANDOM VARIABLE . . . . .	64
3.1	Expected values . . . . .	64
3.2	Moments . . . . .	67
3.3	The Chebyshev inequality . . . . .	74
3.4	Absolute moments . . . . .	76
3.5	Order parameters . . . . .	77
3.6	Moments of random vectors . . . . .	79
3.7	Regression of the first type . . . . .	91
3.8	Regression of the second type . . . . .	96
	Problems and Complements . . . . .	101
4	CHARACTERISTIC FUNCTIONS	
4.1	Properties of characteristic functions . . . . .	105
4.2	The characteristic function and moments . . . . .	107
4.3	Semi-invariants . . . . .	110
4.4	The characteristic function of the sum of independent random variables . . . . .	112
4.5	Determination of the distribution function by the characteristic function . . . . .	115
4.6	The characteristic function of multidimensional random vectors	121
4.7	Probability-generating functions . . . . .	125
	Problems and Complements . . . . .	126
5	SOME PROBABILITY DISTRIBUTIONS . . . . .	129
5.1	One-point and two-point distributions . . . . .	129
5.2	The Bernoulli scheme. The binomial distribution . . . . .	130
5.3	The Poisson scheme. The generalized binomial distribution	134
5.4	The Pólya and hypergeometric distributions . . . . .	135
5.5	The Poisson distribution . . . . .	140
5.6	The uniform distribution . . . . .	145
5.7	The normal distribution . . . . .	147
5.8	The gamma distribution . . . . .	151
5.9	The beta distribution . . . . .	154
5.10	The Cauchy and Laplace distributions . . . . .	156
5.11	The multidimensional normal distribution . . . . .	158

5.12	The multinomial distribution	163
5.13	Compound distributions	164
	Problems and Complements	170
<b>6</b>	<b>LIMIT THEOREMS</b>	<b>175</b>
6.1	Preliminary remarks	175
6.2	Stochastic convergence	176
6.3	Bernoulli's law of large numbers	179
6.4	The convergence of a sequence of distribution functions	180
6.5	The Riemann-Stieltjes integral	184
6.6	The Lévy-Cramér theorem	188
6.7	The de Moivre-Laplace theorem	192
6.8	The Lindeberg-Lévy theorem	196
6.9	The Lapunov theorem	202
6.10	The Gnedenko theorem	211
6.11	Poisson's, Chebyshev's, and Khintchin's laws of large numbers	216
6.12	The strong law of large numbers	220
6.13	Multidimensional limit distributions	232
6.14	Limit theorems for rational functions of some random variables	236
6.15	Final remarks	239
	Problems and Complements	239
<b>7</b>	<b>MARKOV CHAINS</b>	<b>250</b>
7.1	Preliminary remarks	250
7.2	Homogeneous Markov chains	250
7.3	The transition matrix	252
7.4	The ergodic theorem	255
7.5	Random variables forming a homogeneous Markov chain	263
	Problems and Complements	267
<b>8</b>	<b>STOCHASTIC PROCESSES</b>	<b>271</b>
8.1	The notion of a stochastic process	271
8.2	Markov processes and processes with independent increments	272
8.3	The Poisson process	276
8.4	The Furry-Yule process	281
8.5	Birth and death process	287
8.6	The Pólya process	298
8.7	Kolmogorov equations	301

8.8	Purely discontinuous and purely continuous processes	304
8.9	The Wiener process	309
8.10	Stationary processes	314
8.11	Martingales	323
8.12	Additional remarks	325
	Problems and Complements	327

## PART 2 MATHEMATICAL STATISTICS

9	SAMPLE MOMENTS AND THEIR FUNCTIONS	335
9.1	The notion of a sample	335
9.2	The notion of a statistic	337
9.3	The distribution of the arithmetic mean of independent normally distributed random variables	337
9.4	The $\chi^2$ distribution	339
9.5	The distribution of the statistic $(\bar{X}, S)$	343
9.6	Student's $t$ -distribution	348
9.7	Fisher's $Z$ -distribution	354
9.8	The distribution of $\bar{X}$ for some non-normal populations	357
9.9	The distribution of sample moments and sample correlation coefficients of a two-dimensional normal population	358
9.10	The distribution of regression coefficients	363
9.11	Limit distributions of sample moments	366
	Problems and Complements	368
10	ORDER STATISTICS	372
10.1	Preliminary remarks	372
10.2	The notion of an order statistic	372
10.3	The empirical distribution function	374
10.4	Stochastic convergence of sample quantiles	377
10.5	Limit distributions of sample quantiles	379
10.6	The limit distributions of successive sample elements	384
10.7	The joint distribution of a group of quantiles	387
10.8	The distribution of the sample range	388
10.9	Tolerance limits	388
10.10	Glivenko theorem	390
10.11	The theorems of Kolmogorov and Smirnov	394
10.12	Rényi's theorem	405
10.13	The problem of $k$ samples	407
	Problems and Complements	410

11	AN OUTLINE OF THE THEORY OF RUNS . . . . .	415
11.1	Preliminary remarks . . . . .	415
11.2	The notion of a run . . . . .	415
11.3	The probability distribution of the number of runs . . . . .	416
11.4	The expected value and the variance of the number of runs . . . . .	421
	Problems and Complements . . . . .	423
12	SIGNIFICANCE TESTS . . . . .	425
12.1	The concept of a statistical test . . . . .	425
12.2	Parametric tests for small samples . . . . .	427
12.3	Parametric tests for large samples . . . . .	433
12.4	The $\chi^2$ test . . . . .	436
12.5	Tests of the Kolmogorov and Smirnov type . . . . .	445
12.6	The Wald-Wolfovitz and Wilcoxon-Mann-Whitney tests . . . . .	449
12.7	Independence tests by contingency tables . . . . .	456
	Problems and Complements . . . . .	459
13	THE THEORY OF ESTIMATION . . . . .	461
13.1	Preliminary notions . . . . .	461
13.2	Consistent estimates . . . . .	461
13.3	Unbiased estimates . . . . .	462
13.4	The sufficiency of an estimate . . . . .	465
13.5	The efficiency of an estimate . . . . .	467
13.6	Asymptotically most efficient estimates . . . . .	479
13.7	Methods of finding estimates . . . . .	484
13.8	Confidence intervals . . . . .	490
13.9	Bayes theorem and estimation . . . . .	494
	Problems and Complements . . . . .	499
14	METHODS AND SCHEMES OF SAMPLING . . . . .	503
14.1	Preliminary remarks . . . . .	503
14.2	Methods of random sampling . . . . .	504
14.3	Schemes of independent and dependent random sampling . . . . .	509
14.4	Schemes of unrestricted and stratified random sampling . . . . .	512
14.5	Random errors of measurements . . . . .	520
	Problems and Complements . . . . .	522

15	AN OUTLINE OF ANALYSIS OF VARIANCE . . . . .	524
15.1	One-way classification . . . . .	524
15.2	Multiple classification . . . . .	531
15.3	A modified regression problem . . . . .	535
	Problems and Complements . . . . .	540
16	THEORY OF HYPOTHESES TESTING . . . . .	541
16.1	Preliminary remarks . . . . .	541
16.2	The power function and the <i>OC</i> function . . . . .	541
16.3	Most powerful tests . . . . .	552
16.4	Uniformly most powerful test . . . . .	558
16.5	Unbiased tests . . . . .	560
16.6	The power and consistency of nonparametric tests . . . . .	566
16.7	Additional remarks . . . . .	578
	Problems and Complement . . . . .	578
17	ELEMENTS OF SEQUENTIAL ANALYSIS . . . . .	584
17.1	Preliminary remarks . . . . .	584
17.2	The sequential probability ratio test . . . . .	585
17.3	Auxiliary theorems . . . . .	587
17.4	The fundamental Identity . . . . .	591
17.5	The <i>OC</i> function of the sequential probability ratio test . . . . .	592
17.6	The expected value $E(n)$ . . . . .	595
17.7	The determination of $A$ and $B$ . . . . .	597
17.8	Testing a hypothesis concerning the parameter $p$ of a zero-one distribution . . . . .	597
17.9	Testing a hypothesis concerning the expected value $m$ of a normal population . . . . .	604
17.10	Additional remarks . . . . .	608
	Problems and Complements . . . . .	610
	SUPPLEMENT . . . . .	612
	REFERENCES . . . . .	621
	STATISTICAL TABLES . . . . .	658
	AUTHOR INDEX . . . . .	665
	SUBJECT INDEX . . . . .	671



PART I

Probability  
Theory



## Random Events

## 1.1 PRELIMINARY REMARKS

**A** Probability theory is a part of mathematics which is useful in discovering and investigating the regular features of random events. The following examples show what is ordinarily understood by the term random event.

**Example 1.1.1.** Let us toss a symmetric coin. The result may be either a head or a tail. For any one throw, we cannot predict the result, although it is obvious that it is determined by definite causes. Among them are the initial velocity of the coin, the initial angle of throw, and the smoothness of the table on which the coin falls. However, since we cannot control all these parameters, we cannot predetermine the result of any particular toss. Thus the result of a coin tossing, head or tail, is a random event.

**Example 1.1.2.** Suppose that we observe the average monthly temperature at a definite place and for a definite month, for instance, for January in Warsaw.<sup>1</sup> This average depends on many causes such as the humidity and the direction and strength of the wind. The effect of these causes changes year by year. Hence Warsaw's average temperature in January is not always the same. Here we can determine the causes for a given average temperature, but often we cannot determine the reasons for the causes themselves. As a result, we are not able to predict with a sufficient degree of accuracy what the average temperature for a certain January will be. Thus we refer to it as a random event.

**B** It might seem that there is no regularity in the examples given. But if the number of observations is large, that is, if we deal with a mass phenomenon, some regularity appears.

Let us return to example 1.1.1. We cannot predict the result of any particular toss, but if we perform a long series of tossings, we notice that the number of times heads occur is approximately equal to the number of times tails appear. Let  $n$  denote the number of all our tosses and  $m$  the number of times heads appear. The fraction  $m/n$  is called the

<sup>1</sup> See example 12.5.1.

frequency of appearance of heads. The frequency of appearance of tails is given by the fraction  $(n - m)/n$ . Experience shows that if  $n$  is sufficiently large, thus if the tossings may be considered as a mass phenomenon, the fractions  $m/n$  and  $(n - m)/n$  differ little; hence each of them is approximately  $\frac{1}{2}$ . This regularity has been noticed by many investigators who have performed a long series of coin tossings. Buffon

TABLE 1.1.1  
FREQUENCY OF BIRTHS OF BOYS AND GIRLS

Year of Birth	Number of Births		Total Number of Births	Frequency of Births	
	Boys	Girls		Boys	Girls
	$m$	$f$	$m + f$	$p_1$	$p_2$
1927	496,544	462,189	958,733	0.518	0.482
1928	513,654	477,339	990,993	0.518	0.482
1929	514,765	479,336	994,101	0.518	0.482
1930	528,072	494,739	1,022,811	0.516	0.484
1931	496,986	467,587	964,573	0.515	0.485
1932	482,431	452,232	934,663	0.516	0.484
Total	3,032,452	2,833,422	5,865,874	0.517	0.483

tossed a coin 4040 times, and obtained heads 2048 times; hence the ratio of heads was  $m/n = 0.50693$ . In 24,000 tosses, K. Pearson obtained a frequency of heads equal to 0.5005. We can see quite clearly that the observed frequencies oscillate about the number 0.5.

As a result of long observation, we can also notice certain regularities in example 1.1.2. We investigate this more closely in example 12.5.1.

**Example 1.1.3.** We cannot predict the sex of a newborn baby in any particular case. We treat this phenomenon as a random event. But if we observe a large number of births, that is, if we deal with a mass phenomenon, we are able to predict with considerable accuracy what will be the percentages of boys and girls among all newborn babies. Let us consider the number of births of boys and girls in Poland in the years 1927 to 1932. The data are presented in Table 1.1.1.

In this table  $m$  and  $f$  denote respectively the number of births of boys and girls in particular years. Denote the frequencies of births by  $p_1$  and  $p_2$ , respectively; then

$$p_1 = \frac{m}{m + f}, \quad p_2 = \frac{f}{m + f}$$

One can see that the values of  $p_1$  oscillate about the number 0.517, and the values of  $p_2$  oscillate about the number 0.483.