

Statistical and Computational Methods in Data Analysis

SIEGMUND BRANDT

*Institute of High Energy Physics
Heidelberg University
Heidelberg, Germany*



1970

NORTH-HOLLAND PUBLISHING COMPANY — AMSTERDAM · LONDON
AMERICAN ELSEVIER PUBLISHING COMPANY, INC. — NEW YORK

CONTENTS

Preface	IX
List of frequently used symbols	XI
 1. Introduction	 1
 2. Probabilities	 4
2-1. Experiments, events, sample space	4
2-2. The concept of probability	5
2-3. Rules of probability calculus; conditional probability	7
 3. Random variables; distributions of a random variable	 9
3-1. Random variables	9
3-2. Distributions of one random variable	9
3-3. Functions of one random variable, expectation value, variance, moments	11
3-4. Chebychev's inequality	17
 4. Distributions of several random variables	 19
4-1. Distribution function and probability density of two variables; conditional probability	19
4-2. Expectation values, variances, covariances and correlation coefficient	21
4-3. More than two variables; vector and matrix notation	24
4-4. Transformation of variables	27
4-5. Linear and orthogonal transformations; propagation of errors	31
 5. Some important distributions and theorems	 37
5-1. Binomial and multinomial distributions	37
5-2. Frequency; the law of large numbers	40
5-3. Hypergeometric distribution	41
5-4. Poisson distribution	45
5-5. Uniform distribution	50
5-6. The characteristic function of a distribution	51
5-7. The Laplace model of errors	53

5-8. Normal distribution	56
5-9. Quantitative properties of the normal distribution	58
5-10. Multivariate normal distribution	61
5-11. The central limit theorem	67
5-12. Experimental errors and normal distribution; Herschel's model	68
5-13. Convolution of distributions	71
6. Sampling	74
6-1. Random sampling; distribution of a sample; estimates	74
6-2. Sampling from continuous populations	77
6-3. Sampling from partitioned distributions	78
6-4. Sampling without replacement from finite populations; mean square deviation; degrees of freedom	82
6-5. Sampling from normal distributions; χ^2 -distribution	87
6-6. χ^2 and empirical variance	91
7. The method of "maximum likelihood"	94
7-1. Likelihood quotient; likelihood function	94
7-2. The concept of maximum likelihood	96
7-3. Information inequality, minimum variance and sufficient estimates	99
7-4. Asymptotic properties of the likelihood function and of maxi- mum likelihood estimators	106
7-5. Solution of the likelihood equation by iteration	108
7-6. Simultaneous estimation of several parameters	109
7-7. Uniqueness of the method; confidence interval	113
7-8. Bartlett's S -function	115
8. Testing of statistical hypotheses	119
8-1. F -test on equality of variances	121
8-2. Student's test; comparison of means	125
8-3. Some aspects of a general theory of tests	130
8-4. Neyman-Pearson theorem and applications	135
8-5. The likelihood ratio method	138
8-6. The χ^2 -test on goodness of fit	143
9. The method of least squares	151
9-1. Direct measurements with equal or different accuracy	151
9-2. Indirect measurements	156
9-2.1. The linear case	156
9-2.2. The non-linear case	164
9-2.3. Properties of the least squares solution; χ^2 -test	170

9-3. Constrained measurements	173
9-3.1. The method of elements	174
9-3.2. The method of Lagrangian multipliers	176
9-4. The general case of least squares fitting	181
9-5. A FORTRAN program for general least squares fitting; examples	184
10. Some remarks on minimization	202
10-1. Parameter estimation and minimization	202
10-2. Different minimization procedures	203
11. Analysis of variance	210
11-1. One-way classification	210
11-2. Some aspects of two-way classification	215
11-3. A FORTRAN program for two-way classification	224
12. Linear regression	229
12-1. Linear regression as a simple case of least squares	229
12-2. Confidence intervals	233
12-3. Testing of hypotheses	234
12-4. Linear regression and analysis of variance	235

APPENDICES

A. Some elements of the FORTRAN programming language	237
B. Short review of matrix calculus	248
B-1. Definitions of matrices and vectors	248
B-2. Equality, addition, subtraction and multiplication of matrices	251
B-3. Determinant and inverse of a square matrix; solution of matrix equations	255
B-4. FORTRAN programs for matrix handling	262
C. Elements of combinatorial analysis	272
D. Euler's gamma-function	275
E. Collection of important formulae	277

F. Statistical tables	297
F-1. Poisson distribution	297
F-2. Normal distribution function	300
F-3. Fractiles of the normal distribution	303
F-4. χ^2 -distribution function	306
F-5. Fractiles of χ^2 -distribution	308
F-6. F -test	309
F-7. Fractiles for Student's test	314
F-8. Random numbers	315
 Bibliography and references	 316
Subject index	319
Index to FORTAN statements and FORTAN subprograms used in this book	 323