

Contents

1	REPETITION OF SOME BASIC NOTIONS	7
1.1	Maximum likelihood estimators	8
1.1.1	Examples of maximum likelihood estimators	9
1.2	Sufficient statistic	14
1.2.1	Examples of sufficient statistics	15
1.3	Fisher information	18
1.3.1	Examples of evaluating Fisher information	19
2	BAYESIAN STATISTICS	23
2.1	Bayes theorem	23
2.1.1	Examples of application of Bayes theorem	29
2.2	Types of priors	34
2.2.1	Conjugate priors	34
2.2.2	Examples of conjugate priors	36
2.2.3	Noninformative priors	41
2.2.4	Examples of noninformative priors	42
2.2.5	Jeffreys' priors	46
2.2.6	Examples of Jeffreys' priors	49
2.3	Predictive density	53
2.3.1	Example of predictive density	55
2.4	Point estimation via loss function	56
2.4.1	Examples of the minimal expected loss estimators	58
3	TESTS OF GOOD FIT	63
3.1	Test of good fit for multinomial distribution	63
3.1.1	Test of good fit with known probabilities	64
3.1.2	Examples of testing good fit with known probabilities	65
3.1.3	Tests of good fit with probabilities depending on unknown parameters	69
3.2	Test of good fit for Poisson distribution	71
3.2.1	Examples of testing good fit for Poisson distribution	73

3.3	Tests of good fit for normality	74
3.3.1	χ^2 tests of normality	74
3.3.2	Test based on skewness and kurtosis	75
3.4	Test of good fit for exponential distribution	76
3.5	Tests of good fit of Kolmogorov-Smirnov type	77
3.5.1	Examples of testing good fit by Kolmogorov-Smirnov test	79
4	CONTINGENCY TABLES	83
4.1	Hypothesis of independence	84
4.1.1	Examples of testing the independence	86
4.2	Test of homogeneity	89
4.2.1	Example of testing the homogeneity	90
4.3	Test of symmetry	91
4.3.1	Examples of testing the symmetry	92
4.4	Four-cell tables	94
4.4.1	Fisher test	96
4.4.2	Example of applying Fisher test	98
4.4.3	McNemara test	99
4.5	Simpson paradox	101
5	SAMPLING FROM FINITE POPULATIONS	105
5.1	Introduction	105
5.2	Basic notions	107
5.3	Sampling techniques	110
5.3.1	Poisson sampling	110
5.3.2	Simple random sampling	116
5.3.3	Rejective sampling	117
5.3.4	Sampford-Durbin sampling	121
5.3.5	Successive sampling	125
5.3.6	Systematic sampling	130
5.3.7	Stratified sampling	131
5.3.8	Multistage sampling	132
5.3.9	Correcting a sample	133
5.4	Estimating methods	137
5.4.1	Representativeness of the sampling strategy	139
5.4.2	Tightness of estimating mean square error	140
5.4.3	Robustness of estimating mean square error	143
5.4.4	Simple linear estimator	144
5.5	Optimal sampling strategy	149

6 APPENDIX A	153
6.1 Combinatorics	153
6.1.1 "Positive" combinatorics	153
6.1.2 "Negative" combinatorics	154
6.1.3 A bit more complicated combinatorics	156
6.2 Testing statistical hypotheses	160
6.3 Tests of good fit	160
7 BIBLIOGRAPHY	163
8 AUTHOR INDEX	167
9 SUBJECT INDEX	169

First of all, let us introduce basic notations. We shall denote by N the set of all positive integers, by R the real line and for any $k \in N$ by R^k the k -dimensional Euclidean space. A (basic) probability space will be denoted by $(\Omega, \mathcal{A}, P)$. For the Borel σ -algebra on the line R we will use $\mathcal{B}$.

All further necessary notations will be given later on.

The first chapter is devoted to the *repetition* of some basic notions we have met with at the introductory course in statistics. We shall need them later, either directly in the explanation of new topics or indirectly as some "hatch marks".

Prior to starting main text, let us make a technical remark. We shall use the word *estimator* for the corresponding statistics (i. e. for a measurable mapping from a probability space to a parameter space, so that it is a random variable or vector) while the word *estimate* for the value of the estimator for given data (i. e. the estimate is a real number, real vector etc.).

In what follows we shall consider mostly random variables which are either continuous or discrete. let us recall that the random variable is a measurable mapping from the space $(\Omega, \mathcal{A})$ to $(R, \mathcal{B})$, i. e.

$$X: (\Omega, \mathcal{A}) \rightarrow (R, \mathcal{B}). \quad (1.1)$$

In statistical textbooks usually capitals (as X, Y, Z etc.) are used for the random variable while $X(\omega), Y(\omega)$ etc. are used for the "observed values" of the corresponding random variables. Finally, small letters are used of real numbers (or vectors). (Putting words *observed values* into converted commas we have indicated that one can feel a difference between observed value and the value of random variable at the point ω .) Since this text is intended for the students of social sciences, we shall use both X as well as $X(\omega)$ etc. for random variables (the former version, at the points where we want to emphasize the fact that the random variable is the mapping (1.1)). In other words, $X(\omega)$ stands for the