

HANDBOOK OF APPLICABLE MATHEMATICS

Chief Editor: Walter Ledermann

Volume VI: Statistics

PART A

Edited by

Emlyn Lloyd

University of Lancaster

A Wiley-Interscience Publication

JOHN WILEY & SONS

Chichester – New York – Brisbane – Toronto – Singapore

Contents

PART A

INTRODUCTION	xvii
1. INTRODUCTION TO STATISTICS	1
1.1. The meaning of the word 'statistics'	1
1.2. Sampling distribution; statistic; estimate	4
1.3. The subject-matter of this book	7
1.3.1. Sampling distributions; distribution-free methods	8
1.3.2. Estimates, tests, decisions	9
1.3.3. Bayesian inference	12
1.3.4. Multivariate analysis	13
1.3.5. Time series	13
1.3.6. Bibliography and references	13
1.3.7. Appendix: statistical tables	14
1.3.8. Some topics that are not covered	14
1.4. Conventions and notations	15
1.4.1. Mathematical conventions	15
1.4.2. Statistical and probabilistic conventions and abbreviations	15
(i) Abbreviations; (ii) Nomenclature of standard distributions; (iii) capital letter convention for random variables; (iv) Notation for moments and related quantities; (v) Non-standard usages: induced r.v., statistical copy; (vi) An ambiguity in the meaning of $P(A/K)$: 'the probability of A on K'; (vii) Nomenclature for tabular values: percentage points	
1.5. Further reading	18
2. SAMPLING DISTRIBUTIONS	19
2.1. Moments and other statistics	19
2.1.1. Statistic	19
2.1.2. Moments	22
(a) Population moments; (b) Sample moments; (c) Sample variance and sample standard deviation; (d) Bivariate samples	

2.2. Sampling distributions; introductory definitions and examples	26
2.3. Sampling moments of a statistic	29
2.3.1. Lower sampling moments of the sample mean	30
2.3.2. Lower sampling moments of the sample variance	30
2.3.3. Sampling covariance between the sample mean $\bar{x}$ and the sample variance v	31
2.3.4. Sampling moments of higher sample moments	32
2.3.5. Sampling moments of the sample standard deviation	32
2.3.6. Sampling moments of the sample skewness	32
2.4. Distributions of sums of independent and identically distributed variables	33
2.5. Sampling distributions of functions of Normal variables	33
2.5.1. The Normal distribution	33
2.5.2. Effect of a linear transformation; standardization	34
2.5.3. Linear functions of Normal variables	36
(a) Linear functions of independent Normals	36
(b) Sampling distribution of the sample mean	37
(c) Linear function of correlated Normals	37
(d) Several linear functions of correlated Normals	37
(e) Independent linear functions of correlated Normals	38
(f) Independent linear functions of i.i.d. Normals	38
2.5.4. Quadratic functions of Normal variables	39
(a) The chi-squared distribution. Sum of squares of independent standard Normals. Quadratic forms in Normal variables	39
(b) Independence of the sample sum of squares and the sample mean in Normal samples	42
(c) Tables of the χ^2 distribution	42
(d) Sampling distribution of the sample variance	44
(e) Sampling distribution of the sample standard deviation	45
2.5.5. Student's distribution (the 't distribution')	48
2.5.6. The variance ratio (F) distribution	50
(a) F as a weighted ratio of χ^2 variables	50
(b) Relation between the F distribution and the Beta distribution	52
(c) Approximation to the F distribution when one degree of freedom is very much larger than the other	52
2.5.7. The sample correlation coefficient	53
2.5.8. The independence of quadratic forms: the Fisher-Cochran theorem; Craig's theorem	55
2.5.9. The range and the Studentized range	58
2.6. Approximate sampling distribution of $\bar{x}$ and of non-linear functions of $\bar{x}$	59
2.7. Approximation to sampling expectation and variance of	

<i>non-linear statistics; variance-stabilizing transformations;</i>	
<i>Normalizing transformations</i>	61
2.7.1. Approximations	61
(a) Functions of a single r.v.	61
(b) Functions of two random variables	63
2.7.2. Variance-stabilizing transformations	64
(a) A general formula	64
(b) Poisson data: the square root transformation	65
(c) Binomial data: the arc sine (or angular) transformation	66
(d) Stabilizing the variance of a regressed variable by weighting	66
2.7.3. Normalizing transformations	67
(a) The log transform: positively skewed positive variables	67
(b) The log transform for variables bounded above and below; the inverse hyperbolic tangent transformation for the correlation coefficient	67
(c) Normalizing transforms of χ^2	68
(d) The probability integral transform: probits	69
2.7.4. Curve-straightening transformations	70
2.7.5. A Student-distributed transform of the sample correlation coefficient when $\rho = 0$	71
2.7.6. A χ^2 -distributed transform of a uniformly distributed variable	71
2.8. <i>Non-central sampling distributions</i>	72
2.8.1. Non-central chi-squared	72
2.8.2. Non-central F	73
2.8.3. The non-central Student distribution	73
2.9. <i>The multinomial distribution in sampling distribution theory</i>	74
2.9.1. Binomial, Trinomial and multinomial (m) distributions	74
(i) Binomial; (ii) Trinomial; (iii) Multinomial	74
2.9.2. Properties of the Multinomial (m)	76
(a) Low-order moments	76
(b) Marginal distributions	76
2.9.3. The multinomial as a conditional of the joint distribution of independent Poisson variables	76
2.9.4. Frequency tables	76
2.10. <i>Further reading and references</i>	77
 3. <i>ESTIMATION: INTRODUCTORY SURVEY</i>	79
3.1. <i>The estimation problem</i>	79
3.2. <i>Intuitive concepts and graphical methods</i>	84
3.2.1. Introduction	84

3.2.2.	Frequency tables, histograms and the empirical c.d.f.	84
(a)	Discrete data	84
(b)	Bar chart and histogram of discrete data	87
(c)	Continuous data	90
(d)	Histogram of continuous data	92
(e)	Sample analogue of the c.d.f.; Probability graph papers	93
3.3.	<i>Some general concepts and criteria for estimates</i>	94
3.3.1.	Introduction: Dimension, exchangeability, consistency, concentration	94
(a)	Dimensions	95
(b)	Exchangeability	95
(c)	Consistency	96
(d)	Concentration (high local probability)	98
3.3.2.	Unbiased estimates and minimum-variance unbiased estimates	99
3.3.3.	Efficiency: the Cramér–Rao ('C–R') bound	103
(a)	The C–R inequality: random sample from a univariate one-parameter distribution	103
(b)	Attainment of the C–R bound. Efficiency of an estimate	106
(c)	Regularity conditions	108
(d)	The C–R inequality for independent vector observations on a 1-parameter multivariate distribution	109
(e)	The C–R inequality for observations that are not independent and/or not identically distributed	111
(f)	A generalization of the C–R inequality to the case of several parameters. The information matrix	111
3.4.	<i>Sufficiency</i>	114
3.4.1.	Definition of sufficiency	114
3.4.2.	The Factorization criterion and the exponential family	117
3.4.3.	Sufficiency and minimum-variance unbiased estimates	121
(a)	The Rao–Blackwell theorem	121
(b)	Minimum-variance unbiased estimates and sufficiency	123
3.4.4.	Sufficiency in the multi-parameter case	124
3.5.	<i>Practical methods of constructing estimates: an introduction</i>	125
3.5.1.	Graphical methods	125
3.5.2.	Minimum variance unbiased estimates and the method of least squares	128
3.5.3.	The method of moments	130
3.5.4.	The method of maximum likelihood	132
3.5.5.	Normal linear models in which maximum likelihood and least squares estimates coincide	135
3.6.	<i>Further reading</i>	136

4. INTERVAL ESTIMATION	137
4.1. Introduction: the problem	137
4.1.1. An intuitive approach	137
4.1.2. Standard error	139
4.1.3. Probability intervals	140
(a) Probability intervals for continuous random variables	140
(b) Probability intervals for discrete random variables	142
4.2. Confidence intervals and confidence limits	143
4.3. Constructing a confidence interval by using a pivot	147
4.4. Factors affecting the interpretation of a confidence interval as a measure of the precision of an estimate	149
4.5. Confidence intervals when there are several parameters	151
4.5.1. Individual confidence intervals	151
4.5.2. Confidence intervals for functions of two parameters, including the difference between parameters and the ratio of parameters. (Fieller's theorem)	156
4.6. Construction of confidence intervals without using pivots	162
4.7. Approximate confidence intervals when the underlying distribution is discrete	168
(a) A crude approximation; (b) A better approximation	169
4.8. Distribution-free confidence intervals for quantiles	175
4.9. Multiparameter confidence regions	176
4.9.1. Exact confident regions	176
4.9.2. Elliptical confidence regions for the expectation vector of a bivariate Normal distribution; and approximate confidence regions for a pair of maximum likelihood estimates	182
4.10. Confidence interval obtainable from a large sample, using the likelihood function	184
4.10.1. The likelihood function	184
4.10.2. Approximate confidence intervals, using the derivative of the log-likelihood	186
4.10.3. Confidence intervals obtained with the aid of an (approximately) Normalizing transformation	188
4.11. Confidence band for an unknown continuous cumulative distribution function	189
4.11.1. The empirical c.d.f.	189
4.11.2. The Kolmogorov–Smirnov distance between the c.d.f. and the empirical c.d.f.	191
4.11.3. The sampling distribution of the Kolmogorov–Smirnov statistic; confidence limits for the c.d.f.	192
4.12. Tolerance intervals	195
4.13. Likelihood intervals	196
4.13.1. Likelihood	196
(a) Likelihood and log-likelihood: definitions and examples	196
(b) Likelihood and sufficiency	199

4.13.2. Plausible values and likelihood intervals	200
(a) Equiplausible values of θ	200
(b) One value of θ more plausible than another	200
(c) Conventionally implausible values of θ : likelihood intervals	201
(d) Likelihood intervals for θ and for $g(\theta)$	205
4.13.3. The two-parameter case	205
4.14. <i>Bayesian intervals</i>	206
4.15. <i>Further reading and references</i>	207
 5. <i>STATISTICAL TESTS</i>	209
5.1. <i>What is meant by a significance test?</i>	209
5.2. <i>Introduction to tests involving a simple null hypothesis in discrete distributions</i>	209
5.2.1. A two-sided (Binomial) test: ingredients, procedure and interpretation	210
(a) Probability model	210
(b) Condensing the data: the test statistic	210
(c) Null hypothesis, null distribution	211
(d) Alternative hypothesis	212
(e) Compatibility of sample with H	212
(f) Significance set, significance level (significance probability). Critical region	213
(f ₁) distance ordering	214
(f ₂) probability ordering	214
(f ₃) likelihood ratio ordering	216
(g) Interpretation of the significance level	218
(h) Strength of evidence	219
5.2.2. Conventional interpretation of significance levels; practices used in specifying significance levels; Critical region	219
5.2.3. A one-sided (Binomial) test	221
5.2.4. Tests with Poisson distributions	222
5.2.5. Tests with continuous distributions	222
5.2.6. Choosing a test statistic	225
5.3. <i>Criteria for a test</i>	227
5.3.1. Sensitivity function	227
5.3.2. Sensitivity function of the one-tailed test when σ is unknown, in a sample from a Normal (θ, σ) distribution	231
5.3.3. Sensitivity function of a two-tailed test	233
5.4. <i>Tests involving a composite null hypothesis</i>	235
5.4.1. Conditional tests: equality of Binomial proportions; ratio of Poisson parameters	235
5.4.2. Tests for independence in 2×2 contingency tables; Fisher's 'Exact Test'	241

5.5. Tests involving more than one parameter; generalized likelihood ratio tests	247
5.6. Large sample approximation to the significance level of a likelihood ratio test	251
5.7. Randomization tests	254
5.8. Standard tests with the Normal distribution model	255
5.8.1. The significance of a sample mean when the variance is known	255
5.8.2. The significance of a mean when the variance is unknown): Student's <i>t</i> -test	256
5.8.3. The Fisher-Behrens test for the significance of the difference between two means	259
5.8.4. The <i>t</i> -test for the significance of the difference between two means (variances being equal)	261
5.8.5. The <i>t</i> -test for the significance of a regression coefficient in simple linear regression	262
5.8.6. Test for the equality of two variances	264
5.8.7. Tests for the equality of several means: introduction to the analysis of variance	266
(a) Introduction	266
(b) Comparison of two means as analysis of variance	266
(c) The case of <i>k</i> samples	269
5.9. Tests for Normality	272
5.10. Test for the equality of <i>k</i> variances (Bartlett's test)	275
5.11. Combining the results of several tests	277
5.12. Neyman-Pearson theory	278
5.12.1. Sampling inspection schemes	278
5.12.2. The Neyman-Pearson theory of tests	278
5.13. Further reading and references	282
 6. MAXIMUM-LIKELIHOOD ESTIMATES	 283
6.1. Introduction	283
6.2. The method of maximum likelihood	284
6.2.1. The likelihood and log-likelihood functions	284
6.2.2. The maximum likelihood estimate (MLE)	293
6.2.3. MLE: intuitive justification	295
6.2.4. MLE: kinds of maxima	296
6.2.5. Theoretical justification: for the MLE properties of the likelihood function	297
(a) Asymptotic distribution of the derivative of the log-likelihood function	297
(b) Confidence interval for an estimated parameter: first method	299

(c) The principal practical approximation for the sampling distribution of the MLE	302
(d) Confidence intervals and confidence regions: second method	303
6.2.6. Estimating a function of θ . Invariance	304
6.3. <i>Examples of ML estimation in one-parameter cases</i>	305
6.4. <i>Examples of MLE in multi-parameter cases</i>	314
6.5. <i>Maximum likelihood estimation of linear regression coefficients</i>	321
6.5.1. The meaning of 'regression'.	321
6.5.2. The MLE procedure for linear regression with weighted data	324
6.5.3. Linear regression with equally weighted data	324
(i) The estimate; (ii) Sampling properties of the estimates;	
(iii) Confidence intervals for α and β ; (iv) Significance tests for α and β ; (v) Confidence interval for $y(x)$.	
6.6. <i>Estimation of dosage-mortality relationship</i>	333
6.6.1. Quantal responses	333
6.6.2. Probability model	335
6.6.3. The likelihood surface	338
6.6.4. The maximum likelihood estimator	340
6.6.5. Reliability of the estimates	343
6.6.6. Dose required for a specified response	346
6.7. <i>Maximum likelihood estimates from grouped, censored or truncated data</i>	348
6.7.1. Discrete data	348
(i) Truncation; (ii) Grouping, censoring	
6.7.2. Continuous data: grouping. Sheppard's correction	350
6.8. <i>Further reading</i>	352
 7. <i>CHI-SQUARED TESTS OF GOODNESS OF FIT AND OF INDEPENDENCE AND HOMOGENEITY</i>	355
7.1. <i>Validity of the model</i>	355
7.1.1. Introduction	355
7.1.2. Adequacy of the linear regression model	357
7.2. <i>Karl Pearson's distance statistic: the χ^2 test</i>	358
7.3. <i>Amalgamating low frequency cells: Cochran's criteria</i>	365
7.4. <i>The χ^2 test for continuous distributions</i>	365
7.5. <i>Cross-classified frequency tables (contingency tables). Tests of independence and of homogeneity</i>	370
7.5.1. 2×2 tables; the special case of a single degree of freedom	370
7.5.2. $k \times m$ tables	377
7.6. <i>Index of dispersion</i>	379
7.6.1. Index of dispersion for Binomial samples	379
7.6.2. Index of dispersion for Poisson samples	380
7.7. <i>Further reading and references</i>	381

8. LEAST SQUARES ESTIMATION AND THE ANALYSIS OF VARIANCE	383
8.1. Least squares estimation for general models	383
8.2. Least squares estimation for full-rank linear models, Normal equations. The Gauss–Markov theorem	384
8.2.1. Examples of Least Squares estimation	384
8.2.2. Matrix formulation for linear models	385
8.2.3. Properties of Least Squares estimates	390
8.2.4. Residuals	393
8.2.5. Orthogonal designs and orthogonal polynomials	396
8.2.6. Modifications for observations of unequal accuracy; weighted Least Squares	401
8.2.7. Modification for non-independent observations	403
8.3. Analysis of variance and tests of hypotheses for full-rank designs	404
8.3.1. Normal theory	404
8.3.2. Confidence regions	405
8.3.3. Testing hypotheses	407
8.3.4. The basic identities	408
8.3.5. Tests of hypotheses in orthogonal designs	410
8.3.6. Hypothesis testing in non-orthogonal designs	417
8.3.7. Group orthogonality	427
8.3.8. The general linear hypothesis	429
8.4. Further reading	436
9. THE DESIGN OF COMPARATIVE EXPERIMENTS	437
9.1. Historical introduction	437
9.2. The tea-tasting lady	438
9.3. A more elaborate example: Darwin's experiment	440
(i) Blocking; (ii) Replication; (iii) Balance; (iv) Randomization	440
9.4. Complete randomized block designs	442
9.5. Treatments at one level, and at several levels	443
9.6. The need for devices to reduce block size	443
9.7. Balanced incomplete blocks for comparing single-level treatments	444
9.8. The complete factorial design	445
9.8.1. Main effects and interaction: two factors at two levels	446
9.8.2. Main effects and interactions: three factors at two levels	448
9.8.3. Main effects and interactions in a 2^t design:	450
9.9. Incomplete blocks; confounding	451
9.10. Partial confounding	453
9.11. Factors at three or more levels	454
9.11.1. The Latin square	454
9.11.2. The graeco-Latin square	456
9.12. Further reading and references	457

10. LEAST SQUARES AND THE ANALYSIS OF STATISTICAL EXPERIMENTS: SINGULAR MODELS; MULTIPLE TESTS	459
10.1. Singular models	459
10.1.1. Introduction	459
10.1.2. Estimation. Estimable functions	461
10.1.3. Tests of hypotheses	465
10.1.4. The two-way hierarchical classification	470
10.1.5. The two-way cross-classification	473
10.1.6. Higher order classifications	480
10.1.7. Analysis of covariance	484
10.2. Multiple tests and comparisons	488
10.2.1. Introduction	488
10.2.2. Combination of tests, and overall size	489
10.2.3. Multiple comparisons	491
10.3. The assumptions about the errors	496
10.3.1. The basic assumptions	496
10.3.2. Residual analysis	497
10.4. Further reading and references	498
Appendix	A1
Index	xxiii

Part B contains Chapters 11–20.