

Contents

1	The Problem of Simultaneous Inference	1
1.1	Sources of Multiplicity	3
1.2	Multiple Hypotheses Testing	4
1.2.1	Measuring and Controlling Errors	4
1.2.2	Structured Systems of Hypotheses	8
1.3	Relationships to Other Simultaneous Statistical Inference Problems	9
1.4	Contributions of this Work	11
	References	12

Part I General Theory

2	Some Theory of p-values	17
2.1	Randomized p -values	20
2.1.1	Randomized p -values in Discrete Models	20
2.1.2	Randomized p -values for Testing Composite Null Hypotheses	21
2.2	p -value Models	22
2.2.1	The iid.-Uniform Model	22
2.2.2	Dirac-Uniform Configurations	24
2.2.3	Two-Class Mixture Models	25
2.2.4	Copula Models Under Fixed Margins	26
2.2.5	Further Joint Models	26
	References	27
3	Classes of Multiple Test Procedures	29
3.1	Margin-Based Multiple Test Procedures	30
3.1.1	Single-Step Procedures	30
3.1.2	Stepwise Rejective Multiple Tests	32
3.1.3	Data-Adaptive Procedures	35
3.2	Multivariate Multiple Test Procedures	37
3.2.1	Resampling-Based Methods	37
3.2.2	Methods Based on Central Limit Theorems	38

3.2.3	Copula-Based Methods	38
3.3	Closed Test Procedures	40
	References	43
4	Simultaneous Test Procedures.	47
4.1	Three Important Families of Multivariate Probability Distributions	50
4.1.1	Multivariate Normal Distributions	50
4.1.2	Multivariate t -distributions.	51
4.1.3	Multivariate Chi-Square Distributions	51
4.2	Projection Methods Under Asymptotic Normality.	52
4.3	Probability Bounds and Effective Numbers of Tests	56
4.3.1	Sum-Type Probability Bounds	57
4.3.2	Product-Type Probability Bounds	58
4.3.3	Effective Numbers of Tests	61
4.4	Simultaneous Test Procedures in Terms of p -value Copulae	62
4.5	Exploiting the Topological Structure of the Sample Space via Random Field Theory.	65
	References	68
5	Stepwise Rejective Multiple Tests	71
5.1	Some Concepts of Dependency	72
5.2	FWER-Controlling Step-Down Tests.	74
5.3	FWER-Controlling Step-Up Tests.	76
5.4	FDR-Controlling Step-Up Tests	80
5.5	FDR-Controlling Step-Up-Down Tests	82
	References	89
6	Multiple Testing and Binary Classification	91
6.1	Binary Classification Under Sparsity.	93
6.2	Binary Classification in Non-Sparse Models	96
6.3	Feature Selection for Binary Classification via Higher Criticism	99
	References	101
7	Multiple Testing and Model Selection	103
7.1	Multiple Testing for Model Selection	104
7.2	Multiple Testing and Information Criteria	106
7.3	Multiple Testing After Model Selection	108
7.3.1	Distributions of Regularized Estimators.	108
7.3.2	Two-Stage Procedures.	111
7.4	Selective Inference	112
	References	114

8	Software Solutions for Multiple Hypotheses Testing.	117
8.1	The R Package <code>multcomp</code>	118
8.2	The R Package <code>multtest</code>	118
8.3	The R-based μ TOSS Software	119
8.3.1	The μ TOSS Simulation Tool	120
8.3.2	The μ TOSS Graphical User Interface	122
	References	124

Part II From Genotype to Phenotype

9	Genetic Association Studies.	129
9.1	Statistical Modeling and Test Statistics	130
9.2	Estimation of the Proportion of Informative Loci	133
9.3	Effective Numbers of Tests via Linkage Disequilibrium	134
9.4	Combining Effective Numbers of Tests and Pre-estimation of π_0	137
9.5	Applicability of Margin-Based Methods	138
	References	139
10	Gene Expression Analyses.	141
10.1	Marginal Models and p -values	141
10.2	Dependency Considerations	143
10.3	Real Data Examples	146
10.3.1	Application of Generic Multiple Tests to Large-Scale Data	146
10.3.2	Copula Calibration for a Block of Correlated Genes	147
10.4	LASSO and Statistical Learning Methods	149
10.5	Gene Set Analyses and Group Structures.	150
	References	151
11	Functional Magnetic Resonance Imaging.	155
11.1	Spatial Modeling	156
11.2	False Discovery Rate Control for Grouped Hypotheses	157
11.2.1	Clusters of Voxels	157
11.2.2	Multiple Endpoints per Location	159
11.3	Exploiting Topological Structure by Random Field Theory	160
11.4	Spatio-Temporal Models via Multivariate Time Series	161
11.4.1	Which of the Specific Factors have a Non-trivial Autocorrelation Structure?	164
11.4.2	Which of the Common Factors have a Lagged Influence on Which X_i ?	165
	References	165

Part III Further Applications in the Life Sciences

12 Further Life Science Applications	169
12.1 Brain-Computer Interfacing	169
12.2 Gel Electrophoresis-Based Proteome Analysis	172
References	174
Index	177

<http://www.springer.com/978-3-642-45181-2>

Simultaneous Statistical Inference
With Applications in the Life Sciences

Dickhaus, T.

2014, XIV, 180 p. 19 illus., 10 illus. in color., Hardcover

ISBN: 978-3-642-45181-2