

Wolfgang Härdle · Yuichi Mori
Philippe Vieu

Statistical Methods for Biostatistics and Related Fields

With 91 Figures and 57 Tables

 Springer

Contents

I	Biostatistics	1
1	Discriminant Analysis Based on Continuous and Discrete Variables	3
	<i>Avner Bar-Hen and Jean-Jacques Daudin</i>	
1.1	Introduction	3
1.2	Generalisation of the Mahalanobis Distance	4
1.2.1	Introduction	4
1.2.2	Kullback–Leibler Divergence	5
1.2.3	Asymptotic Distribution of Matusita Distance	10
1.2.4	Simulations	12
1.3	Methods and Stopping Rules for Selecting Variables	13
1.4	Reject Option	15
1.4.1	Distributional Result	15
1.4.2	Derivation of the Preliminary Test	18
1.5	Example	22
1.5.1	Location Model	22
1.5.2	Comparison with the Linear Discriminant Analysis	24
1.5.3	Conclusion	24
	Bibliography	25

2	Longitudinal Data Analysis with Linear Regression	29
	<i>Jörg Breitung, Rémy Slama and Axel Werwatz</i>	
2.1	Introduction	29
2.2	Theoretical Aspects	32
2.2.1	The Fixed-effect Model	32
2.2.2	The Random Effects Model	36
2.3	Computing Fixed and Random-effect Models	37
2.3.1	Data Preparation	37
2.3.2	Fixed and Random-effect Linear Regression	38
2.3.3	Options for <code>panfix</code>	38
2.3.4	Options for <code>panrand</code>	39
2.4	Application	40
2.4.1	Results	41
	Bibliography	43
3	A Kernel Method Used for the Analysis of Replicated Micro-array Experiments	45
	<i>Ali Gannoun, Beno Liquefit, Jérôme Saracco and Wolfgang Urfer</i>	
3.1	Introduction	45
3.2	Statistical Model and Some Existing Methods	46
3.2.1	The Basic Model	47
3.2.2	The T-test	47
3.2.3	The Mixture Model Approach	48
3.3	A Fully Nonparametric Approach	49
3.3.1	Kernel Estimation of f_0 and f	50
3.3.2	The Reflection Approach in Kernel Estimation	50
3.3.3	Implementation of the Nonparametric Method	51
3.4	Data Analysis	52
3.4.1	Results Obtained with the Normal Mixture Model	53

3.4.2	Results Obtained with the Nonparametric Approach	53
3.4.3	A Simulation Study	56
3.5	Discussion and Concluding Remarks	58
	Bibliography	59
4	Kernel Estimates of Hazard Functions for Biomedical Data Sets	63
	<i>Ivana Horová and Jiří Zelinka</i>	
4.1	Introduction	63
4.2	Kernel Estimate of the Hazard Function and Its Derivatives	64
4.3	Choosing the Shape of the Kernel	68
4.4	Choosing the Bandwidth	69
4.5	Description of the Procedure	74
4.6	Application	75
	Bibliography	83
5	Partially Linear Models	87
	<i>Wolfgang Härdle and Hua Liang</i>	
5.1	Introduction	87
5.2	Estimation and Nonparametric Fits	89
5.2.1	Kernel Regression	89
5.2.2	Local Polynomial	90
5.2.3	Piecewise Polynomial	93
5.2.4	Least Square Spline	96
5.3	Heteroscedastic Cases	97
5.3.1	Variance Is a Function of Exogenous Variables	98
5.3.2	Variance Is an Unknown Function of T	99
5.3.3	Variance Is a Function of the Mean	99
5.4	Real Data Examples	100
	Bibliography	102

6	Analysis of Contingency Tables	105
	<i>Masahiro Kuroda</i>	
6.1	Introduction	105
6.2	Log-linear Models	105
6.2.1	Log-linear Models for Two-way Contingency Tables	106
6.2.2	Log-linear Models for Three-way Contingency Tables	107
6.2.3	Generalized Linear Models	109
6.2.4	Fitting to Log-linear Models	111
6.3	Inference for Log-linear Models Using XploRe	113
6.3.1	Estimation of the Parameter Vector λ	113
6.3.2	Computing Statistics for the Log-linear Models	113
6.3.3	Model Comparison and Selection	114
6.4	Numerical Analysis of Contingency Tables	115
6.4.1	Testing Independence	115
6.4.2	Model Comparison	119
	Bibliography	124
7	Identifying Coexpressed Genes	125
	<i>Qihua Wang</i>	
7.1	Introduction	125
7.2	Methodology and Implementation	127
7.2.1	Weighting Adjustment	128
7.2.2	Clustering	132
7.3	Concluding Remarks	142
	Bibliography	144

8 Bootstrap Methods for Testing Interactions in GAMs	147
<i>Javier Roca-Pardiñas, Carmen Cadarso-Suárez and Wenceslao González-Manteiga</i>	
8.1 Introduction	147
8.2 Logistic GAM with Interactions	149
8.2.1 Estimation: the Local Scoring Algorithm	150
8.3 Bootstrap-based Testing for Interactions	152
8.3.1 Likelihood Ratio-based Test	153
8.3.2 Direct Test	153
8.3.3 Bootstrap Approximation	153
8.4 Simulation Study	154
8.5 Application to Real Data Sets	156
8.5.1 Neural Basis of Decision Making	156
8.5.2 Risk of Post-operative Infection	159
8.6 Discussion	162
8.7 Appendix	163
Bibliography	165
9 Survival Trees	167
<i>Carmela Cappelli and Heping Zhang</i>	
9.1 Introduction	167
9.2 Methodology	170
9.2.1 Splitting Criteria	170
9.2.2 Pruning	173
9.3 The Quantlet stree	174
9.3.1 Syntax	174
9.3.2 Example	175
Bibliography	179