Contents

Ρ	Preface				
A	cknowledgements	xi			
Li	List of Figures				
Ir	troduction	1			
1	Using R	3			
	1.1 Installing and using R	3			
	1.2 Documenting your code and results	3			
	1.3 Simple usage of R	6			
	1.4 Graphics	21			
	1.5 Frequency data	28			
	1.6 Tables and arrays for results	3 3			
	1.7 Dates in R	35			
	1.8 Numerical accuracy	37			
	1.9 tidyverse and data.table	39			
2	Measures of disease occurrence	41			
	2.1 Prevalence	41			
	2.2 Mortality rate	42			
	2.3 Incidence rate	44			
	2.4 Standardized mortality ratio	44			
	2.5 Survival	45			
3	Prevalence data—models, likelihood, and binomial regression	47			
	3.1 Likelihood	47			
	3.2 Prevalence by age	55			
	3.3 Comparing different models for the same data	60			
4	Regression models	65			
	4.1 Types of models	65			
	4.2 Normal linear regression model	66			
	4.3 Simple linear regression	67			
	4.4 Multiple regression	68			

vi CONTENT

	4.5	Model formulae in R	74
	4.6	Regression models and generalized linear models	75
	4 7	Collinearity and aliasing	83
	4.8	Logarithmic transformations	87
5	Ana	lysis of follow-up data	93
	5.1	Basic data structure	93
	5.2	Probability model	94
	5.3	Representation of follow-up data	97
	5.4	Splitting the follow-up time along a time-scale	101
	5.5	Smooth age-effects for rates	104
	5.6	SMR	107
	5.7	Time-dependent variables	111
6	Par	ametrization and prediction of rates	119
	6.1	Predictions and contrasts	119
	6.2	Prediction of a single rate	119
	6.3	Categorical variables	121
	6.4	Modelling the effect of quantitative variables	124
	6.5	Two quantitative predictors	134
	6.6	Quantitative interactions	145
7	Ca	e-control and case-cohort studies	159
	7.1	Follow-up and case-control studies	160
	7.2	Statistical model for the odds ratio	163
	7.3	Odds ratio and rate ratio	166
	7.4	Confounding and stratified sampling	166
	7.5	Individually matched studies	170
	7.6	Nested case-control studies	174
	7.7	Case-cohort studies	177
8	Su	rvival analysis	181
	8.1	Introduction	181
	8.2	Life table estimator of survival function	181
	8.3	Kaplan-Meier estimator of survival	184
	8.4	The Cox model	188
	8.5	The time-scale	193
	8.6	Relation between Cox and Poisson models	196

	8.7	Time-dependent covariates	206
	8.8	210	
	8.9	Modelling cause specific rates	214
	8.10	The Fine-Gray approach to competing risks	217
	8.11	Time-dependent variables and competing risks	217
9	Doı	not group quantitative variables	219
	9.1	Problems Caused by Categorizing Continuous Variables	219
Re	ference	223	
In	225		