

Contents

1. Linear Regression	1
1.1 Introduction	1
1.2 The Method of Least Squares	3
1.2.1 Correlation version	7
1.2.2 Large-sample limit	8
1.3 The origins of regression	9
1.4 Applications of regression	11
1.5 The Bivariate Normal Distribution	14
1.6 Maximum Likelihood and Least Squares	21
1.7 Sums of Squares	23
1.8 Two regressors	26
Exercises	28
2. The Analysis of Variance (ANOVA)	33
2.1 The Chi-Square Distribution	33
2.2 Change of variable formula and Jacobians	36
2.3 The Fisher F-distribution	37
2.4 Orthogonality	38
2.5 Normal sample mean and sample variance	39
2.6 One-Way Analysis of Variance	42
2.7 Two-Way ANOVA; No Replications	49
2.8 Two-Way ANOVA: Replications and Interaction	52
Exercises	56
3. Multiple Regression	61
3.1 The Normal Equations	61

3.2	Solution of the Normal Equations	64
3.3	Properties of Least-Squares Estimators	70
3.4	Sum-of-Squares Decompositions	73
3.4.1	Coefficient of determination	79
3.5	Chi-Square Decomposition	80
3.5.1	Idempotence, Trace and Rank	81
3.5.2	Quadratic forms in normal variates	82
3.5.3	Sums of Projections	82
3.6	Orthogonal Projections and Pythagoras's Theorem	85
3.7	Worked examples	89
	Exercises	94
4.	Further Multilinear Regression	99
4.1	Polynomial Regression	99
4.1.1	The Principle of Parsimony	102
4.1.2	Orthogonal polynomials	103
4.1.3	Packages	103
4.2	Analysis of Variance	104
4.3	The Multivariate Normal Distribution	105
4.4	The Multinormal Density	111
4.4.1	Estimation for the multivariate normal	113
4.5	Conditioning and Regression	115
4.6	Mean-square prediction	121
4.7	Generalised least squares and weighted regression	123
	Exercises	125
5.	Adding additional covariates and the Analysis of Covariance	129
5.1	Introducing further explanatory variables	129
5.1.1	Orthogonal parameters	133
5.2	ANCOVA	135
5.2.1	Nested Models	139
5.3	Examples	140
	Exercises	145
6.	Linear Hypotheses	149
6.1	Minimisation Under Constraints	149
6.2	Sum-of-Squares Decomposition and F-Test	152
6.3	Applications: Sequential Methods	157
6.3.1	Forward selection	157
6.3.2	Backward selection	158
6.3.3	Stepwise regression	159
	Exercises	160

7.	Model Checking and Transformation of Data	163
7.1	Deviations from Standard Assumptions	163
7.2	Transformation of Data	168
7.3	Variance-Stabilising Transformations	171
7.4	Multicollinearity	174
	Exercises	177
8.	Generalised Linear Models	181
8.1	Introduction	181
8.2	Definitions and examples	183
8.2.1	Statistical testing and model comparisons	185
8.2.2	Analysis of residuals	187
8.2.3	Athletics times	188
8.3	Binary models	190
8.4	Count data, contingency tables and log-linear models	193
8.5	Over-dispersion and the Negative Binomial Distribution	197
8.5.1	Practical applications: Analysis of over-dispersed models in R®	199
	Exercises	200
9.	Other topics	203
9.1	Mixed models	203
9.1.1	Mixed models and Generalised Least Squares	206
9.2	Non-parametric regression	211
9.2.1	Kriging	213
9.3	Experimental Design	215
9.3.1	Optimality criteria	215
9.3.2	Incomplete designs	216
9.4	Time series	219
9.4.1	Cointegration and spurious regression	220
9.5	Survival analysis	222
9.5.1	Proportional hazards	224
9.6	$p \gg n$	225
	Solutions	227
	Dramatis Personae: Who did what when	269
	Bibliography	271
	Index	279