

Regression Modeling with Actuarial and Financial Applications

EDWARD W. FREES

University of Wisconsin, Madison

gig CAMBRIDGE
UNIVERSITY PRESS

Contents

<i>Preface</i>	<i>page xiii</i>
1 Regression and the Normal Distribution	1
1.1 What Is Regression Analysis?	1
1.2 Fitting Data to a Normal Distribution	3
1.3 Power Transforms	7
1.4 Sampling and the Role of Normality	8
1.5 Regression and Sampling Designs	10
1.6 Actuarial Applications of Regression	12
1.7 Further Reading and References	13
1.8 Exercises	14
1.9 Technical Supplement – Central Limit Theorem	18
 Part I Linear Regression	
2 Basic Linear Regression	23
2.1 Correlations and Least Squares	23
2.2 Basic Linear Regression Model	29
2.3 Is the Model Useful? Some Basic Summary Measures	32
2.4 Properties of Regression Coefficient Estimators	35
2.5 Statistical Inference	37
2.6 Building a Better Model: Residual Analysis	41
2.7 Application: Capital Asset Pricing Model	46
2.8 Illustrative Regression Computer Output	51
2.9 Further Reading and References	54
2.10 Exercises	54
2.11 Technical Supplement – Elements of Matrix Algebra	62
 3 Multiple Linear Regression – I	70
3.1 Method of Least Squares	70
3.2 Linear Regression Model and Properties of Estimators	76
3.3 Estimation and Goodness of Fit	81
3.4 Statistical Inference for a Single Coefficient	85
3.5 Some Special Explanatory Variables	92
3.6 Further Reading and References	100
3.7 Exercises	101

4	Multiple Linear Regression – II	107
4.1	The Role of Binary Variables	107
4.2	Statistical Inference for Several Coefficients	113
4.3	One Factor ANOVA Model	120
4.4	Combining Categorical and Continuous Explanatory Variables	126
4.5	Further Reading and References	133
4.6	Exercises	133
4.7	Technical Supplement – Matrix Expressions	138
5	Variable Selection	148
5.1	An Iterative Approach to Data Analysis and Modeling	148
5.2	Automatic Variable Selection Procedures	149
5.3	Residual Analysis	153
5.4	Influential Points	160
5.5	Collinearity	165
5.6	Selection Criteria	171
5.7	Heteroscedasticity	175
5.8	Further Reading and References	179
5.9	Exercises	180
5.10	Technical Supplements for Chapter 5	182
6	Interpreting Regression Results	189
6.1	What the Modeling Process Tells Us	190
6.2	The Importance of Variable Selection	196
6.3	The Importance of Data Collection	198
6.4	Missing Data Models	205
6.5	Application: Risk Managers' Cost-Effectiveness	209
6.6	Further Reading and References	218
6.7	Exercises	219
6.8	Technical Supplements for Chapter 6	222
Part II Topics in Time Series		
7	Modeling Trends	227
7.1	Introduction	227
7.2	Fitting Trends in Time	229
7.3	Stationarity and Random Walk Models	236
7.4	Inference Using Random Walk Models	238
7.5	Filtering to Achieve Stationarity	243
7.6	Forecast Evaluation	245
7.7	Further Reading and References	248
7.8	Exercises	249
8	Autocorrelations and Autoregressive Models	251
8.1	Autocorrelations	251
8.2	Autoregressive Models of Order One	254

8.3	Estimation and Diagnostic Checking	256
8.4	Smoothing and Prediction	258
8.5	Box-Jenkins Modeling and Forecasting	260
8.6	Application: Hong Kong Exchange Rates	265
8.7	Further Reading and References	269
8.8	Exercises	270
9	Forecasting and Time Series Models	273
9.1	Smoothing with Moving Averages	273
9.2	Exponential Smoothing	275
9.3	Seasonal Time Series Models	278
9.4	Unit Root Tests	284
9.5	ARCH/GARCH Models	285
9.6	Further Reading and References	288
10	Longitudinal and Panel Data Models	289
10.1	What Are Longitudinal and Panel Data?	289
10.2	Visualizing Longitudinal and Panel Data	291
10.3	Basic Fixed Effects Models	293
10.4	Extended Fixed Effects Models	296
10.5	Random Effects Models	299
10.6	Further Reading and References	301
Part III Topics in Nonlinear Regression		
11	Categorical Dependent Variables	305
11.1	Binary Dependent Variables	305
11.2	Logistic and Probit Regression Models	307
11.3	Inference for Logistic and Probit Regression Models	312
11.4	Application: Medical Expenditures	315
11.5	Nominal Dependent Variables	318
11.6	Ordinal Dependent Variables	325
11.7	Further Reading and References	328
11.8	Exercises	329
11.9	Technical Supplements – Likelihood-Based Inference	337
12	Count Dependent Variables	343
12.1	Poisson Regression	343
12.2	Application: Singapore Automobile Insurance	348
12.3	Overdispersion and Negative Binomial Models	352
12.4	Other Count Models	354
12.5	Further Reading and References	359
12.6	Exercises	360
13	Generalized Linear Models	362
13.1	Introduction	362
13.2	GLM Model	364

13.3	Estimation	367
13.4	Application: Medical Expenditures	371
13.5	Residuals	374
13.6	Tweedie Distribution	375
13.7	Further Reading and References	376
13.8	Exercises	377
13.9	Technical Supplements – Exponential Family	378
14	Survival Models	383
14.1	Introduction	383
14.2	Censoring and Truncation	385
14.3	Accelerated Failure Time Model	390
14.4	Proportional Hazards Model	392
14.5	Recurrent Events	395
14.6	Further Reading and References	397
15	Miscellaneous Regression Topics	399
15.1	Mixed Linear Models	399
15.2	Bayesian Regression	403
15.3	Density Estimation and Scatterplot Smoothing	406
15.4	Generalized Additive Models	409
15.5	Bootstrapping	410
15.6	Further Reading and References	412
Part IV Actuarial Applications		
16	Frequency-Severity Models	417
16.1	Introduction	417
16.2	Tobit Model	418
16.3	Application: Medical Expenditures	421
16.4	Two-Part Model	424
16.5	Aggregate Loss Model	427
16.6	Further Reading and References	429
16.7	Exercises	432
17	Fat-Tailed Regression Models	433
17.1	Introduction	433
17.2	Transformations	434
17.3	Generalized Linear Models	437
17.4	Generalized Distributions	442
17.5	Quantile Regression	446
17.6	Extreme Value Models	448
17.7	Further Reading and References	449
17.8	Exercises	451

18	Credibility and Bonus-Malus	452
18.1	Risk Classification and Experience Rating	452
18.2	Credibility	453
18.3	Credibility and Regression	458
18.4	Bonus-Malus	464
18.5	Further Reading and References	465
19	Claims Triangles	467
19.1	Introduction	467
19.2	Regression Using Functions of Time as Explanatory Variables	471
19.3	Using Past Developments	475
19.4	Further Reading and References	477
19.5	Exercises	478
20	Report Writing: Communicating Data Analysis Results	481
20.1	Overview	481
20.2	Methods for Communicating Data	482
20.3	How to Organize	486
20.4	Further Suggestions for Report Writing	490
20.5	Case Study: Swedish Automobile Claims	491
20.6	Further Reading and References	503
20.7	Exercises	504
21	Designing Effective Graphs	505
21.1	Introduction	506
21.2	Graphic Design Choices Make a Difference	508
21.3	Design Guidelines	513
21.4	Empirical Foundations for Guidelines	520
21.5	Concluding Remarks	526
21.6	Further Reading and References	526
	<i>Brief Answers to Selected Exercises</i>	529
	<i>Appendix 1: Basic Statistical Inference</i>	547
	<i>Appendix 2: Matrix Algebra</i>	551
	<i>Appendix 3: Probability Tables</i>	554
	<i>Index</i>	559