Contents

List of Tables xiii

List of Figures xiv

Preface xix

About the Companion Website xxiii

I Linear Time Series Methods 1

R and Time Series Analysis 3
Overview 3
Some Useful R Functions for Time Series Analysis 4

1 Introduction to Linear Time Series Models 7
1.1 Overview 7
1.2 Time Series Data 8
1.3 Patterns in Time Series 9
1.4 Stationary versus Non-Stationary Series 9
1.5 Examples of Univariate Random Processes 12
1.5.1 White Noise Processes 12
1.5.2 Random Walk Processes 13
1.6 Characterizing Time Series 14
1.6.1 The Autocorrelation Function 14
1.6.2 The Sample Autocorrelation Function 16
1.6.3 Non-Stationarity and Differencing 16
1.7 Tests for White Noise Processes 18
1.7.1 Individual Test for H_0: $\rho_k = 0$—Bartlett's Test 18
1.7.2 Joint Test for H_0: $\rho_1 = 0 \cap \rho_2 = 0 \cap \cdots \cap \rho_k = 0$—Ljung & Box's Test 19
1.7.3 A Simulated Illustration—Testing for a White Noise Process 19
1.7.4 A Simulated Illustration—White Noise Tests when the Series is a Random Walk Process 21

Racine, Jeffrey Scott
Reproducible econometrics using R
digitalisiert durch:
IDS Luzern

2 Random Walks, Unit Roots, and Spurious Relationships 23
2.1 Overview 23
2.2 Properties of a Random Walk 24
2.3 The Autocorrelation Function for a Random Walk 25
2.4 Classical Least Squares Estimators and Random Walks 25
2.5 Classical Least Squares Inference and Random Walks 26
2.5.1 Cross-Section (I.I.D. Data) Monte Carlo 26
2.5.2 Time Series (Random Walk) Monte Carlo 27
2.5.3 Time Series (Random Walk with Drift) Monte Carlo 29
2.6 Unit Root Tests 30
2.6.1 Testing for a Unit Root in Spot Exchange Rates 32
2.7 Random Walks and Spurious Regression 33

3 Univariate Linear Time Series Models 37
3.1 Overview 37
3.2 Moving Average Models (MA(q)) 38
3.2.1 Structure of MA(q) Processes 38
3.2.2 Example—Residential Electricity Sales 38
3.2.3 Properties of MA(q) Processes 39
3.2.4 Stationarity of MA(q) Processes 40
3.2.5 The Autocorrelation Function and Identification of MA(q) Processes 40
3.2.6 Forecasting MA(q) Processes 41
3.2.7 Forecasting MA(q) Processes Assuming the ϵ_{T-i} and the θ_i are Known 41
3.2.8 Forecasting MA(q) Processes when the ϵ_{T-i} and the θ_i are Estimated 42
3.2.9 Forecasting MA(q) Processes in the Presence of a Trend 43
3.3 Autoregressive Models (AR(p)) 44
3.3.1 Structure of AR(p) Processes 44
3.3.2 Example—Residential Electricity Sales 45
3.3.3 Properties of AR(p) Processes 46
3.3.4 Stationarity of AR(p) Processes 48
3.3.5 Invertiblity of Stationary AR(p) Processes 50
3.3.6 Identification of AR(p) Processes—The Partial Autocorrelation Function 50
3.3.7 Forecasting AR(p) Processes 51
3.3.8 Forecasting AR(p) Processes when the Parameters ϕ_i are Unknown 53
3.3.9 Forecasting AR(p) Processes in the Presence of a Trend 54
3.4 Non-Seasonal Autoregressive Moving Average Models (ARMA(p, q)) 55

3.4.1 Structure 55
3.5 Non-Seasonal Autoregressive Integrated Moving Average Models (ARIMA(p, d, q)) 55
3.5.1 Structure 55
3.5.2 Stationarity of ARIMA(p, d, q) Models 56
3.5.3 Identification of ARIMA(p, d, q) Processes 57
3.5.4 Estimation of ARIMA(p, d, q) Processes 57
3.5.5 Forecasting ARIMA(p, d, q) Processes 58
3.5.6 Trends, Constants, and ARIMA(p, d, q) Models 59
3.5.7 Model Selection Criteria, Trends, and Stationarity 62
3.5.8 Model Selection via `auto.arima()` 65
3.5.9 *Diagnostics for ARIMA(p, d, q) Models* 67
3.6 Seasonal Autoregressive Integrated Moving Average Models (ARIMA$(p, d, q)(P, D, Q)_m$) 67
3.6.1 Example—Modelling and Forecasting European Quarterly Retail Trade 71
3.6.2 Example—Modelling Monthly Cortecosteroid Drug Sales 72
3.7 ARIMA$(p, d, q)(P, D, Q)_m$ Models with External Predictors 74
3.8 Assessing Model Accuracy on Hold-Out Data 77

Problem Set **81**

II Robust Parametric Inference **85**

R, The Bootstrap and the Jackknife **87**
Overview 87
Some Useful R Functions for Data-Driven Inference 87

4 Robust Parametric Inference **89**
4.1 Overview 89
4.2 Analytical Versus Numerical, i.e., Data-Driven, Procedures 90
4.2.1 Drawbacks of the Analytical Approach 91
4.2.2 An Illustrative Example—Testing for a Unit Root 91
4.3 Alternatives to Analytical Approaches 92
4.3.1 Motivating Example—Compute the Standard Error of $\bar{X}$ 92
4.4 An Introduction to Efron's Bootstrap 93
4.4.1 Bootstrapping a Standard Error for the Sample Mean 93
4.4.2 Bootstrap Implementations in R 94
4.5 Jackknifing—Background and Motivating Example 95
4.6 Jackknife and Bootstrap Estimates of Bias 97
4.7 To Bootstrap or Jackknife? 99

4.8 Data-Driven Covariance Matrices 99
4.8.1 Bootstrap Heteroskedasticity Consistent Covariance Matrix Estimation 100
4.9 The Wild Bootstrap 102
4.10 Bootstrapping Dependent Processes 104
4.11 Bootstrap Confidence Intervals 105
4.11.1 Example—Nonparametric Confidence Intervals for the Population Mean 106
4.12 Bootstrap Inference 108
4.12.1 How Many Bootstrap Replications? 109
4.12.2 Generating $\hat{\theta}^*$ Under the Null 110
4.12.3 Example—The Two-Sample Problem 111
4.12.4 Example—Regression-Based Bootstrap Inference 112
4.12.5 Example—Unit Root Testing 114

Problem Set **117**

III Robust Parametric Estimation **119**

R and Robust Parametric Estimation **121**
Overview 121
Some Useful R Functions for Robust Parametric Estimation 121

5 Robust Parametric Estimation **123**
5.1 Overview 123
5.2 Robust Estimation Basics 124
5.2.1 Outlier 124
5.2.2 Breakdown Point 125
5.2.3 Sensitivity Curve 125
5.2.4 Contamination Neighborhoods 126
5.2.5 Influence Function 128
5.3 Unmasking Univariate Outliers 129
5.3.1 L_1 and L_2-norm Estimators of Central Tendency 130
5.3.2 Robustness versus Efficiency 132
5.3.3 M-Estimator Methods 133
5.3.4 Optimal Robustness 135
5.3.5 Huber's M-Estimator of Location—A More Efficient Robust Location Estimator than the Median 135
5.3.6 Rousseeuw and Croux's Q_n Estimator of Scale—A More Efficient Robust Scale Estimator than MAD_n 137
5.3.7 M-Estimators of Scale 139
5.3.8 Unmasking Univariate Outliers—The *three-sigma edit* rule 142

5.4 Unmasking Multivariate Outliers 142
5.5 Unmasking Regression Outliers 148
5.5.1 Outliers in the Y Direction 148
5.5.2 Outliers in the X Direction 150
5.5.3 Leverage Points 150
5.5.4 Dealing with Outlying Observations and Leverage Points 152
5.6 Robust Regression 160
5.6.1 Robust Residuals and High Breakdown Diagnostics 164
5.7 Some Useful Points to Remember 165

Problem Set **167**

IV Model Uncertainty **171**

R and Model Uncertainty **173**
Overview 173
Some Useful R Functions for Model Uncertainty 173

6 Model Uncertainty **175**
6.1 Overview 175
6.1.1 Model Selection References 176
6.1.2 Model Averaging References 176
6.1.3 Resources 177
6.2 A Reflection on Models and Data Generating Processes 177
6.2.1 Model Selection and Averaging--A Simulation 181
6.2.2 Discussion 183
6.3 Kullback-Leibler Distance and Maximum Likelihood Estimation 184
6.4 Model Selection Methods 186
6.4.1 AIC, BIC, C_p and Cross-Validated Model Selection Criteria 186
6.5 Model Averaging Methods 189
6.5.1 Solving for the Optimal Model Average Weights 190
6.5.2 Selecting Candidate Models 191
6.5.3 Pitfalls of Model Selection and Model Averaging 196
6.5.4 An Experimental Robust Regression M-Estimator Model Averaging Procedure 196

Problem Set **201**

V Advanced Topics **207**

R and Advanced Topics **209**
Overview 209
Some Useful R Functions for Advanced Topics 209

7 Advanced Topics **211**
7.1 Overview 211
7.2 Classification Analysis and Support Vector Machines 211
7.2.1 *The Confusion Matrix* 212
7.2.2 Support Vector Machines 213
7.3 Nonparametric Kernel Regression 220

Problem Set **225**

VI Appendix **227**

A R, RStudio, TeX, and Git **229**
A.1 Installation of R and RStudio Desktop 229
A.2 What is R? 229
A.2.1 R in the News 230
A.2.2 *Introduction to R* 230
A.2.3 Econometrics in R 230
A.3 What is RStudio Desktop? 231
A.3.1 Introduction to RStudio 231
A.4 Installation of TeX 231
A.5 *Installation of Git* 231

B R Markdown for Assignments **233**
B.1 Source Code (R Markdown) for this Document 233
B.2 R, RStudio, TeX and git 233
B.3 What is R Markdown? 233
B.4 Creating a New R Markdown Document in RStudio 234
B.5 Including R Results in your R Markdown Document 234
B.6 Reading Data from a URL 234
B.7 Including Plots 235
B.8 Including Bulleted and Numbered lists 236
B.9 Including Tables 237
B.10 Including Verbatim, i.e., Freeform, Text 237
B.11 Typesetting Mathematics 237
B.12 Flexible Document Creation 238
B.13 Knitting your R Markdown Document 238
B.14 Printing Your Assignment for Submitting in Class 238
B.15 Troubleshooting and Tips 239

C Maximum Likelihood Estimation and Inference **243**
C.1 Maximum Likelihood Estimation 243
C.2 Properties of the Maximum Likelihood Estimators 244
C.3 Maximum Likelihood Estimation in Practice 246
C.4 A Simple Example Using Discrete Data 246
C.4.1 Example— 247
C.4.2 Example— 248
C.5 Maximum Likelihood Estimation of the Normal Linear Multivariate Regression Model 249
C.6 Information and the Normal Linear Multivariate Model 252
C.6.1 Example— 253
C.7 Restricted Maximum Likelihood Estimates 254
C.8 Hypothesis Testing in a Maximum Likelihood Framework 254
C.8.1 Example— 255
C.8.2 Example— 256

D Solving a Quadratic Program Using R **259**
D.1 Example 260

E A Primer on Regression Splines **263**
E.1 Overview 263
E.2 Bézier curves 264
E.2.1 Example—A quadratic Bézier curve 264
E.2.2 The Bézier curve defined 265
E.2.3 Example—A quadratic Bézier curve as a linear interpolation between two linear Bézier curves 265
E.2.4 Example—The quadratic Bézier curve basis functions 266
E.3 Derivatives of spline functions 267
E.4 B-splines 267
E.4.1 B-spline knots 267
E.4.2 The B-spline basis function 268
E.4.3 Example—A fourth-order B-spline basis function with three interior knots and its first derivative function 269
E.5 The B-spline function 269
E.6 Multivariate B-spline regression 269
E.6.1 Multivariate knots, intervals, and spline bases 271
E.7 Spline regression 272

Bibliography **273**

Author Index **281**