
Contents

Part I DATA EXPLORATION, ESTIMATION AND SIMULATION

1	UNIVARIATE EXPLORATORY DATA ANALYSIS	3
1.1	Data, Random Variables and Their Distributions	3
1.1.1	The PCS Data	4
1.1.2	The S&P 500 Index and Financial Returns	5
1.1.3	Random Variables and Their Distributions	7
1.1.4	Examples of Probability Distribution Families	8
1.2	First Exploratory Data Analysis Tools	13
1.2.1	Random Samples	13
1.2.2	Histograms	14
1.3	More Nonparametric Density Estimation	16
1.3.1	Kernel Density Estimation	17
1.3.2	Comparison with the Histogram	19
1.3.3	S&P Daily Returns	19
1.3.4	Importance of the Choice of the Bandwidth	22
1.4	Quantiles and Q-Q Plots	23
1.4.1	Understanding the Meaning of Q-Q Plots	24
1.4.2	Value at Risk and Expected Shortfall	25
1.5	Estimation from Empirical Data	28
1.5.1	The Empirical Distribution Function	28
1.5.2	Order Statistics	29
1.5.3	Empirical Q-Q Plots	30
1.6	Random Generators and Monte Carlo Samples	31
1.7	Extremes and Heavy Tail Distributions	35
1.7.1	S&P Daily Returns, Once More	35
1.7.2	The Example of the PCS Index	37
1.7.3	The Example of the Weekly S&P Returns	41
	Problems	43
	Notes & Complements	46

2	MULTIVARIATE DATA EXPLORATION	49
2.1	Multivariate Data and First Measure of Dependence	49
2.1.1	Density Estimation	51
2.1.2	The Correlation Coefficient	53
2.2	The Multivariate Normal Distribution	56
2.2.1	Simulation of Random Samples	57
2.2.2	The Bivariate Case	58
2.2.3	A Simulation Example	59
2.2.4	Let's Have Some Coffee	60
2.2.5	Is the Joint Distribution Normal?	62
2.3	Marginals and More Measures of Dependence	63
2.3.1	Estimation of the Coffee Log-Return Distributions	64
2.3.2	More Measures of Dependence	68
2.4	Copulas and Random Simulations	70
2.4.1	Copulas	71
2.4.2	First Examples of Copula Families	72
2.4.3	Copulas and General Bivariate Distributions	74
2.4.4	Fitting Copulas	76
2.4.5	Monte Carlo Simulations with Copulas	77
2.4.6	A Risk Management Example	80
2.5	Principal Component Analysis	84
2.5.1	Identification of the Principal Components of a Data Set	84
2.5.2	PCA with <i>S-Plus</i>	87
2.5.3	Effective Dimension of the Space of Yield Curves	87
2.5.4	Swap Rate Curves	90
	Appendix 1: Calculus with Random Vectors and Matrices	92
	Appendix 2: Families of Copulas	95
	Problems	98
	Notes & Complements	101

Part II REGRESSION

3	PARAMETRIC REGRESSION	105
3.1	Simple Linear Regression	105
3.1.1	Getting the Data	106
3.1.2	First Plots	107
3.1.3	Regression Set-up	108
3.1.4	Simple Linear Regression	111
3.1.5	Cost Minimizations	114
3.1.6	Regression as a Minimization Problem	114
3.2	Regression for Prediction & Sensitivities	116
3.2.1	Prediction	116
3.2.2	Introductory Discussion of Sensitivity and Robustness	118

3.2.3	Comparing L2 and L1 Regressions	119
3.2.4	Taking Another Look at the Coffee Data	121
3.3	Smoothing versus Distribution Theory	123
3.3.1	Regression and Conditional Expectation	123
3.3.2	Maximum Likelihood Approach	124
3.4	Multiple Regression	129
3.4.1	Notation	129
3.4.2	The S-Plus Function <code>lm</code>	130
3.4.3	R^2 as a Regression Diagnostic	131
3.5	Matrix Formulation and Linear Models	133
3.5.1	Linear Models	134
3.5.2	Least Squares (Linear) Regression Revisited	134
3.5.3	First Extensions	139
3.5.4	Testing the CAPM	142
3.6	Polynomial Regression	145
3.6.1	Polynomial Regression as a Linear Model	146
3.6.2	Example of S-Plus Commands	146
3.6.3	Important Remark	148
3.6.4	Prediction with Polynomial Regression	148
3.6.5	Choice of the Degree p	150
3.7	Nonlinear Regression	150
3.8	Term Structure of Interest Rates: A Crash Course	154
3.9	Parametric Yield Curve Estimation	160
3.9.1	Estimation Procedures	160
3.9.2	Practical Implementation	161
3.9.3	S-Plus Experiments	163
3.9.4	Concluding Remarks	165
	Appendix: Cautionary Notes on Some S-Plus Idiosyncracies	166
	Problems	169
	Notes & Complements	172
4	LOCAL & NONPARAMETRIC REGRESSION	175
4.1	Review of the Regression Setup	175
4.2	Natural Splines as Local Smoothers	177
4.3	Nonparametric Scatterplot Smoothers	178
4.3.1	Smoothing Splines	179
4.3.2	Locally Weighted Regression	181
4.3.3	A Robust Smoother	182
4.3.4	The Super Smoother	183
4.3.5	The Kernel Smoother	183
4.4	More Yield Curve Estimation	186
4.4.1	A First Estimation Method	186
4.4.2	A Direct Application of Smoothing Splines	188
4.4.3	US and Japanese Instantaneous Forward Rates	188

4.5	Multivariate Kernel Regression	189
4.5.1	Running the Kernel in S-Plus	192
4.5.2	An Example Involving the June 1998 S&P Futures Contract	193
4.6	Projection Pursuit Regression	197
4.6.1	The S-Plus Function <code>ppreg</code>	198
4.6.2	<code>ppreg</code> Prediction of the S&P Indicators	200
4.7	Nonparametric Option Pricing	205
4.7.1	Generalities on Option Pricing	205
4.7.2	Nonparametric Pricing Alternatives	212
4.7.3	Description of the Data	213
4.7.4	The Actual Experiment	214
4.7.5	Numerical Results	220
	Appendix: Kernel Density Estimation & Kernel Regression	222
	Problems	225
	Notes & Complements	233

Part III TIME SERIES & STATE SPACE MODELS

5	TIME SERIES MODELS: AR, MA, ARMA, & ALL THAT	239
5.1	Notation and First Definitions	239
5.1.1	Notation	239
5.1.2	Regular Time Series and Signals	240
5.1.3	Calendar and Irregular Time Series	241
5.1.4	Example of Daily S&P 500 Futures Contracts	243
5.2	High Frequency Data	245
5.2.1	<code>TimeDate</code> Manipulations	248
5.3	Time Dependent Statistics and Stationarity	253
5.3.1	Statistical Moments	253
5.3.2	The Notion of Stationarity	254
5.3.3	The Search for Stationarity	258
5.3.4	The Example of the CO_2 Concentrations	261
5.4	First Examples of Models	263
5.4.1	White Noise	264
5.4.2	Random Walk	267
5.4.3	Auto Regressive Time Series	268
5.4.4	Moving Average Time Series	272
5.4.5	Using the Backward Shift Operator B	275
5.4.6	Linear Processes	276
5.4.7	Causality, Stationarity and Invertibility	277
5.4.8	ARMA Time Series	281
5.4.9	ARIMA Models	282
5.5	Fitting Models to Data	282
5.5.1	Practical Steps	282

5.5.2	S-Plus Implementation	284
5.6	Putting a Price on Temperature	289
5.6.1	Generalities on Degree Days	290
5.6.2	Temperature Options	291
5.6.3	Statistical Analysis of Temperature Historical Data	294
	Appendix: More S-Plus Idiosyncracies	301
	Problems	304
	Notes & Complements	308
6	MULTIVARIATE TIME SERIES, LINEAR SYSTEMS & KALMAN FILTERING	311
6.1	Multivariate Time Series	311
6.1.1	Stationarity and Auto-Covariance Functions	312
6.1.2	Multivariate White Noise	312
6.1.3	Multivariate AR Models	313
6.1.4	Back to Temperature Options	316
6.1.5	Multivariate MA & ARIMA Models	318
6.1.6	Cointegration	319
6.2	State Space Models	321
6.3	Factor Models as Hidden Markov Processes	323
6.4	Kalman Filtering of Linear Systems	326
6.4.1	One-Step-Ahead Prediction	326
6.4.2	Derivation of the Recursive Filtering Equations	327
6.4.3	Writing an S Function for Kalman Prediction	329
6.4.4	Filtering	331
6.4.5	More Predictions	332
6.4.6	Estimation of the Parameters	333
6.5	Applications to Linear Models	335
6.5.1	State Space Representation of Linear Models	335
6.5.2	Linear Models with Time Varying Coefficients	336
6.5.3	CAPM with Time Varying β 's	337
6.6	State Space Representation of Time Series	338
6.6.1	The Case of AR Series	339
6.6.2	The General Case of ARMA Series	341
6.6.3	Fitting ARMA Models by Maximum Likelihood	342
6.7	Example: Prediction of Quarterly Earnings	343
	Problems	346
	Notes & Complements	351
7	NONLINEAR TIME SERIES: MODELS AND SIMULATION	353
7.1	First Nonlinear Time Series Models	353
7.1.1	Fractional Time Series	354
7.1.2	Nonlinear Auto-Regressive Series	355
7.1.3	Statistical Estimation	356
7.2	More Nonlinear Models: ARCH, GARCH & All That	358

7.2.1	Motivation	358
7.2.2	ARCH Models	359
7.2.3	GARCH Models	361
7.2.4	S-Plus Commands	362
7.2.5	Fitting a GARCH Model to Real Data	363
7.2.6	Generalizations	371
7.3	Stochastic Volatility Models	373
7.4	Discretization of Stochastic Differential Equations	378
7.4.1	Discretization Schemes	379
7.4.2	Monte Carlo Simulations: A First Example	381
7.5	Random Simulation and Scenario Generation	383
7.5.1	A Simple Model for the S&P 500 Index	383
7.5.2	Modeling the Short Interest Rate	386
7.5.3	Modeling the Spread	388
7.5.4	Putting Everything Together	389
7.6	Filtering of Nonlinear Systems	391
7.6.1	Hidden Markov Models	391
7.6.2	General Filtering Approach	392
7.6.3	Particle Filter Approximations	393
7.6.4	Filtering in Finance? Statistical Issues	396
7.6.5	Application: Tracking Volatility	397
	Appendix: Preparing Index Data	403
	Problems	404
	Notes & Complements	408
APPENDIX: AN INTRODUCTION TO S AND S-Plus		411
References		429
Notation Index		433
Data Set Index		435
S-Plus Index		437
Author Index		441
Subject Index		445

Statistical Analysis of Financial Data in S-Plus

Carmona, R.

2004, XVI, 455 p. 138 illus., Hardcover

ISBN: 978-0-387-20286-0