

---

# Contents

---

## Part I DATA EXPLORATION, ESTIMATION AND SIMULATION

---

<b>1</b>	<b>UNIVARIATE DATA DISTRIBUTIONS</b>	3
1.1	Probability Distributions and their Parameters	3
1.1.1	Standard Probability Distribution Families	3
1.1.2	Estimation from Empirical Data	17
1.1.3	Quantiles and Q-Q Plots	21
1.2	Observations and NonParametric Density Estimation	31
1.2.1	Sample Data	31
1.2.2	Nonparametric Estimation	34
1.2.3	Histograms	38
1.2.4	Kernel Density Estimation	40
1.2.5	Empirical Q-Q Plots	46
1.3	Monte Carlo Computations	47
1.3.1	Generating Random Samples in R	47
1.3.2	Limit Theorems and Monte Carlo Computations	50
1.3.3	Home Grown Random Samples	57
	Problems	59
	Notes & Complements	64
<b>2</b>	<b>HEAVY TAIL DISTRIBUTIONS</b>	67
2.1	A Primer on Extreme Value Theory	67
2.1.1	Empirical Evidence of Extreme Events	67
2.1.2	Pareto Distributions	68
2.1.3	Tidbits of Extreme Value Theory	72
2.2	GEV & GPD Parameter Estimation	79
2.2.1	The Method of L-Moments	79
2.2.2	Maximum Likelihood Estimation	85
2.2.3	An Example chosen for Pedagogical Reasons	88
2.2.4	Implementation of the Block-Maxima Method	91

2.3	Semi Parametric Estimation	92
2.3.1	Threshold Exceedances	93
2.3.2	Semi Parametric Estimation	95
2.3.3	The Example of the PCS Index Revisited	97
2.3.4	The Example of the Weekly S&P Returns	101
2.4	Appendix: Risk Measures: Why and What For?	103
	Problems	107
	Notes & Complements	111
<b>3</b>	<b>DEPENDENCE &amp; MULTIVARIATE DATA EXPLORATION</b>	<b>115</b>
3.1	Multivariate Data and First Measure of Dependence	115
3.1.1	Density Estimation	117
3.1.2	The Correlation Coefficient	119
3.2	The Multivariate Normal Distribution	121
3.2.1	Simulation of Random Samples	123
3.2.2	The Bivariate Case	124
3.2.3	A Simulation Example	124
3.2.4	Let's Have Some Coffee	125
3.2.5	Is the Joint Distribution Normal?	127
3.3	Marginals and More Measures of Dependence	129
3.3.1	Estimation of the Coffee Log-Return Distributions	129
3.3.2	More Measures of Dependence	133
3.4	Copulas	135
3.4.1	Definitions and First Properties	136
3.4.2	Examples of Copula Families	137
3.4.3	Copulas and General Bivariate Distributions	139
3.4.4	Fitting Copulas	141
3.4.5	Monte Carlo Simulations with Copulas	142
3.4.6	A Risk Management Example	144
3.4.7	A First Example from the Credit Markets	147
3.4.8	Higher Dimensional Copulas	148
3.4.9	Multi Name Credit Derivatives and CDOs	153
3.5	Principal Component Analysis	159
3.5.1	Identification of the Principal Components of a Data Set	160
3.5.2	PCA with $R$	162
3.5.3	Effective Dimension of the Space of Yield Curves	162
3.5.4	Swap Rate Curves	165
	Appendix 1: Calculus with Random Vectors and Matrices	168
	Appendix 2: Families of Copulas	170
	Problems	172
	Notes & Complements	181

---

**Part II REGRESSION**

---

<b>4</b>	<b>PARAMETRIC REGRESSION</b> .....	185
4.1	Simple Linear Regression .....	185
4.1.1	Getting the Data .....	186
4.1.2	First Plots .....	187
4.1.3	Regression Set-up .....	188
4.1.4	Simple Linear Regression .....	190
4.1.5	Cost Minimizations .....	193
4.1.6	Regression as a Minimization Problem .....	194
4.2	Regression for Prediction & Sensitivities .....	195
4.2.1	Prediction .....	195
4.2.2	Introductory Discussion of Sensitivity and Robustness .....	197
4.2.3	Comparing L2 and L1 Regressions .....	198
4.2.4	Taking Another Look at the Coffee Data .....	200
4.3	Smoothing versus Distribution Theory .....	201
4.3.1	Regression and Conditional Expectation .....	201
4.3.2	Maximum Likelihood Approach .....	202
4.4	Multiple Regression .....	206
4.4.1	Notation .....	206
4.4.2	The R Function <code>lm</code> .....	207
4.4.3	$R^2$ as a Regression Diagnostic .....	208
4.5	Matrix Formulation and Linear Models .....	210
4.5.1	Linear Models .....	210
4.5.2	Least Squares (Linear) Regression Revisited .....	211
4.5.3	Confidence and Prediction Intervals .....	216
4.5.4	First Extensions .....	217
4.5.5	Testing the CAPM .....	220
4.6	Polynomial Regression .....	222
4.6.1	Polynomial Regression as a Linear Model .....	223
4.6.2	Example of R Commands .....	223
4.6.3	Important Remark .....	224
4.6.4	Prediction with Polynomial Regression .....	225
4.6.5	Choice of the Degree $p$ .....	227
4.7	Nonlinear Regression .....	228
4.8	Term Structure of Interest Rates: A Crash Course .....	231
4.9	Parametric Yield Curve Estimation .....	236
4.9.1	Estimation Procedures .....	237
4.9.2	Practical Implementation .....	238
4.9.3	R Experiments .....	239
4.9.4	Concluding Remarks .....	240
	Appendix: Cautionary Notes on Some R Idiosyncracies .....	241
	Problems .....	244
	Notes & Complements .....	251

<b>5</b>	<b>LOCAL &amp; NONPARAMETRIC REGRESSION</b>	255
5.1	Review of the Regression Setup	255
5.2	Basis Expansion Regression	257
5.2.1	Natural Splines as Local Smoothers	257
5.2.2	Feature Function Expansions	258
5.3	Nonparametric Scatterplot Smoothers	261
5.3.1	Smoothing Splines	261
5.3.2	Locally Weighted Regression	263
5.3.3	A Robust Smoother	264
5.3.4	The Super Smoother	265
5.3.5	The Kernel Smoother	265
5.4	More Yield Curve Estimation	269
5.4.1	A First Estimation Method	269
5.4.2	A Direct Application of Smoothing Splines	270
5.4.3	US and Japanese Instantaneous Forward Rates	270
5.5	Multivariate Kernel Regression	271
5.5.1	Running the Kernel in R	274
5.5.2	An Example Involving the June 1998 S&P Futures Contract	275
5.6	Projection Pursuit Regression	281
5.6.1	The R Function <code>ppr</code>	282
5.6.2	<code>ppr</code> Prediction of the S&P Afternoon Indicators	284
5.7	Nonparametric Option Pricing	289
5.7.1	Nonparametric Pricing Alternatives	289
5.7.2	Description of the Data	290
5.7.3	The Actual Experiment	291
5.7.4	Numerical Results	297
	Appendix: Kernel Density Estimation & Kernel Regression	299
	Problems	301
	Notes & Complements	314

---

**Part III TIME SERIES & STATE SPACE MODELS**

---

<b>6</b>	<b>TIME SERIES MODELS: AR, MA, ARMA, &amp; ALL THAT</b>	321
6.1	Notation and First Definitions	321
6.1.1	Notation	322
6.1.2	Regular Time Series and Signals	322
6.1.3	Calendar and Irregular Time Series	323
6.1.4	Creating and Plotting <code>timeSeries</code> Objects in R	324
6.1.5	High Frequency Data	325
6.1.6	<code>TimeDate</code> Manipulations	329
6.2	Time Dependent Statistics and Stationarity	331
6.2.1	Statistical Moments	331
6.2.2	The Notion of Stationarity	332

6.2.3	The Search for Stationarity	335
6.2.4	The Example of the $CO_2$ Concentrations	337
6.3	First Examples of Models	340
6.3.1	White Noise	340
6.3.2	Random Walk	344
6.3.3	Auto Regressive Time Series	345
6.3.4	Moving Average Time Series	348
6.3.5	Using the Backward Shift Operator $B$	350
6.3.6	Linear Processes	352
6.3.7	Causality, Stationarity and Invertibility	353
6.3.8	ARMA Time Series	356
6.3.9	ARIMA Models	357
6.4	Fitting Models to Data	357
6.4.1	Practical Steps	357
6.4.2	R Implementation	359
6.5	Putting a Price on Temperature	367
6.5.1	Generalities on Degree Days	368
6.5.2	Temperature Options	369
6.5.3	Statistical Analysis of Temperature Historical Data	371
	Problems	383
	Notes & Complements	391
<b>7</b>	<b>MULTIVARIATE TIME SERIES, LINEAR SYSTEMS &amp; KALMAN FILTERING</b>	<b>395</b>
7.1	Multivariate Time Series	395
7.1.1	Stationarity and Auto-Covariance Functions	396
7.1.2	Multivariate White Noise	396
7.1.3	Multivariate AR Models	397
7.1.4	Back to Temperature Options	401
7.1.5	Multivariate MA & ARIMA Models	403
7.1.6	Cointegration	404
7.2	State Space Models	406
7.3	Factor Models as Hidden Markov Processes	408
7.4	Kalman Filtering of Linear Systems	411
7.4.1	One-Step-Ahead Prediction	411
7.4.2	Derivation of the Recursive Filtering Equations	412
7.4.3	Filtering	415
7.4.4	More Predictions	417
7.4.5	Estimation of the Parameters	418
7.5	Applications to Linear Models	419
7.5.1	State Space Representation of Linear Models	420
7.5.2	Linear Models with Time Varying Coefficients	421
7.5.3	CAPM with Time Varying $\beta$ 's	421
7.6	State Space Representation of Time Series	424

7.6.1	The Case of AR Series . . . . .	425
7.6.2	The General Case of ARMA Series . . . . .	426
7.6.3	Fitting ARMA Models by Maximum Likelihood . . . . .	428
7.7	Example: Prediction of Quarterly Earnings . . . . .	428
	Problems . . . . .	432
	Notes & Complements . . . . .	440
<b>8</b>	<b>NONLINEAR TIME SERIES: MODELS AND SIMULATION . . . . .</b>	<b>443</b>
8.1	First Nonlinear Time Series Models . . . . .	443
8.1.1	Fractional Time Series . . . . .	444
8.1.2	Nonlinear Auto-Regressive Series . . . . .	445
8.1.3	Statistical Estimation . . . . .	446
8.2	More Nonlinear Models: ARCH, GARCH & All That . . . . .	448
8.2.1	Motivation . . . . .	448
8.2.2	ARCH Models . . . . .	449
8.2.3	GARCH Models . . . . .	451
8.2.4	Summary . . . . .	451
8.2.5	R Commands . . . . .	452
8.2.6	Fitting a GARCH Model to Real Data . . . . .	453
8.2.7	Generalizations . . . . .	462
8.3	Stochastic Volatility Models . . . . .	464
8.4	Discretization of Stochastic Differential Equations . . . . .	470
8.4.1	Discretization Schemes . . . . .	471
8.4.2	Monte Carlo Simulations: A First Example . . . . .	472
8.5	Random Simulation and Scenario Generation . . . . .	475
8.5.1	A Simple Model for the S&P 500 Index . . . . .	475
8.5.2	Modeling the Short Interest Rate . . . . .	477
8.5.3	Modeling the Spread . . . . .	479
8.5.4	Putting Everything Together . . . . .	481
8.6	Filtering of Nonlinear Systems . . . . .	483
8.6.1	Hidden Markov Models . . . . .	483
8.6.2	General Filtering Approach . . . . .	484
8.6.3	Particle Filter Approximations . . . . .	485
8.6.4	Filtering in Finance? Statistical Issues . . . . .	489
8.6.5	Application: Tracking Volatility . . . . .	490
	Problems . . . . .	496
	Notes & Complements . . . . .	500
<b>9</b>	<b>APPENDICES . . . . .</b>	<b>503</b>
9.1	Appendix A: A Quick Introduction to R . . . . .	503
9.1.1	Starting R under Windows and under UNIX . . . . .	503
9.1.2	Creating R Objects . . . . .	504
9.1.3	Random Generation and White Noise . . . . .	506
9.2	Appendix B: A Crash Course on Black-Scholes Option Pricing . . . . .	517
9.2.1	Generalities on Option Pricing . . . . .	517

**References** ..... 525

---

**Part IV INDEXES**

---

**Notation Index** ..... 533

**Data Set Index** ..... 537

**R Index** ..... 539

**Author Index** ..... 543

**Subject Index** ..... 547