

David Ruppert

# Statistics and Finance

An Introduction



Springer

---

## Contents

<b>Notation</b> .....	xxi
<b>1 Introduction</b> .....	1
1.1 References .....	5
<b>2 Probability and Statistical Models</b> .....	7
2.1 Introduction .....	7
2.2 Axioms of Probability .....	7
2.2.1 Independence .....	8
2.2.2 Bayes' law .....	8
2.3 Probability Distributions .....	9
2.3.1 Random variables .....	9
2.3.2 Independence .....	10
2.3.3 Cumulative distribution functions .....	10
2.3.4 Quantiles and percentiles .....	11
2.3.5 Expectations and variances .....	12
2.3.6 Does the expected value exist? .....	13
2.4 Functions of Random Variables .....	14
2.5 Random Samples .....	15
2.6 The Binomial Distribution .....	16
2.7 Location, Scale, and Shape Parameters .....	17
2.8 Some Common Continuous Distributions .....	17
2.8.1 Uniform distributions .....	17
2.8.2 Normal distributions .....	18
2.8.3 The lognormal distribution .....	20
2.8.4 Exponential and double exponential distributions .....	21
2.9 Sampling a Normal Distribution .....	21
2.9.1 Chi-squared distributions .....	21
2.9.2 $t$ -distributions .....	22
2.9.3 $F$ -distributions .....	22
2.10 Order Statistics and the Sample CDF .....	23

2.10.1	Normal probability plots	23
2.11	Skewness and Kurtosis	24
2.12	Heavy-Tailed Distributions	28
2.12.1	Double exponential distributions	30
2.12.2	$t$ -distributions have heavy tails	30
2.12.3	Mixture models	31
2.12.4	Pareto distributions	32
2.12.5	Distributions with Pareto tails	34
2.13	Law of Large Numbers and Central Limit Theorem	36
2.14	Multivariate Distributions	37
2.14.1	Correlation and covariance	38
2.14.2	Independence and covariance	40
2.14.3	The multivariate normal distribution	41
2.15	Prediction	42
2.15.1	Best linear prediction	42
2.15.2	Prediction error in linear prediction	43
2.15.3	Multivariate linear prediction	44
2.16	Conditional Distributions	44
2.16.1	Best prediction	45
2.16.2	Normal distributions: Conditional expectations and variance	45
2.17	Linear Functions of Random Variables	46
2.17.1	Two linear combinations of random variables	48
2.17.2	Independence and variances of sums	49
2.17.3	Application to normal distributions	49
2.18	Estimation	49
2.18.1	Maximum likelihood estimation	50
2.18.2	Standard errors	52
2.18.3	Fisher information	53
2.18.4	Bayes estimation*	54
2.18.5	Robust estimation*	56
2.19	Confidence Intervals	60
2.19.1	Confidence interval for the mean	60
2.19.2	Confidence intervals for the variance and standard deviation	61
2.19.3	Confidence intervals based on standard errors	62
2.20	Hypothesis Testing	62
2.20.1	Hypotheses, types of errors, and rejection regions	62
2.20.2	$P$ -values	63
2.20.3	Two-sample $t$ -tests	64
2.20.4	Statistical versus practical significance	65
2.20.5	Tests of normality	66
2.20.6	Likelihood ratio tests	66
2.21	Summary	68
2.22	Bibliographic Notes	70

2.23	References .....	71
2.24	Problems .....	72
<b>3</b>	<b>Returns .....</b>	<b>75</b>
3.1	Introduction .....	75
3.1.1	Net returns .....	75
3.1.2	Gross returns .....	75
3.1.3	Log returns .....	76
3.1.4	Adjustment for dividends .....	77
3.2	Behavior Of Returns .....	78
3.3	The Random Walk Model .....	80
3.3.1	I.i.d. normal returns .....	80
3.3.2	The lognormal model .....	80
3.3.3	Random walks .....	82
3.3.4	Geometric random walks .....	82
3.3.5	The effect of the drift $\mu$ .....	83
3.3.6	Are log returns normally distributed? .....	84
3.3.7	Do the GE daily returns look like a geometric random walk? .....	86
3.4	Origins of the Random Walk Hypothesis .....	89
3.4.1	Fundamental analysis .....	89
3.4.2	Technical analysis .....	91
3.5	Efficient Markets Hypothesis (EMH) .....	93
3.5.1	Three types of efficiency .....	93
3.5.2	Testing market efficiency .....	94
3.6	Discrete and Continuous Compounding .....	95
3.7	Summary .....	96
3.8	Bibliographic Notes .....	97
3.9	References .....	97
3.10	Problems .....	98
<b>4</b>	<b>Time Series Models .....</b>	<b>101</b>
4.1	Time Series Data .....	101
4.2	Stationary Processes .....	102
4.2.1	Weak white noise .....	103
4.2.2	Predicting white noise .....	104
4.2.3	Estimating parameters of a stationary process .....	104
4.3	AR(1) Processes .....	105
4.3.1	Properties of a stationary AR(1) process .....	106
4.3.2	Convergence to the stationary distribution .....	108
4.3.3	Nonstationary AR(1) processes .....	108
4.4	Estimation of AR(1) Processes .....	110
4.4.1	Residuals and model checking .....	111
4.4.2	AR(1) model for GE daily log returns .....	113
4.5	AR( $p$ ) Models .....	115

4.5.1	AR(6) model for GE daily log returns .....	117
4.6	Moving Average (MA) Processes .....	118
4.6.1	MA(1) processes .....	118
4.6.2	General MA processes .....	118
4.6.3	MA(2) model for GE daily log returns .....	119
4.7	ARIMA Processes .....	120
4.7.1	The backwards operator .....	120
4.7.2	ARMA processes .....	120
4.7.3	Fitting ARMA processes: GE daily log returns .....	121
4.7.4	The differencing operator .....	122
4.7.5	From ARMA processes to ARIMA processes .....	122
4.7.6	ARIMA(2,1,0) model for GE daily log prices .....	123
4.8	Model Selection .....	124
4.8.1	AIC and SBC .....	124
4.8.2	GE daily log returns: Choosing the AR order .....	125
4.9	Three-Month Treasury Bill Rates .....	126
4.10	Forecasting .....	128
4.10.1	Forecasting GE daily log returns and log prices .....	130
4.11	Summary .....	131
4.12	Bibliographic Notes .....	134
4.13	References .....	134
4.14	Problems .....	134
<b>5</b>	<b>Portfolio Theory</b> .....	<b>137</b>
5.1	Trading Off Expected Return and Risk .....	137
5.2	One Risky Asset and One Risk-Free Asset .....	137
5.2.1	Estimating $E(R)$ and $\sigma_R$ .....	140
5.3	Two Risky Assets .....	140
5.3.1	Risk versus expected return .....	140
5.3.2	Estimating means, standard deviations, and covariances .....	141
5.4	Combining Two Risky Assets with a Risk-Free Asset .....	142
5.4.1	Tangency portfolio with two risky assets .....	142
5.4.2	Combining the tangency portfolio with the risk-free asset .....	144
5.4.3	Effect of $\rho_{12}$ .....	146
5.5	Risk-Efficient Portfolios with $N$ Risky Assets* .....	146
5.5.1	Efficient-portfolio mathematics .....	146
5.5.2	The minimum variance portfolio .....	152
5.5.3	Selling short .....	154
5.5.4	Back to the math — Finding the tangency portfolio .....	156
5.5.5	Examples .....	157
5.6	Quadratic Programming* .....	160
5.7	Is the Theory Useful? .....	163
5.8	Utility Theory* .....	164
5.9	Summary .....	165

5.10	Bibliographic Notes .....	166
5.11	References .....	166
5.12	Problems .....	166
<b>6</b>	<b>Regression .....</b>	<b>169</b>
6.1	Introduction .....	169
6.1.1	Straight line regression .....	170
6.2	Least Squares Estimation .....	170
6.2.1	Estimation in straight line regression .....	171
6.2.2	Variance of $\hat{\beta}_1$ .....	172
6.2.3	Estimation in multiple linear regression .....	174
6.3	Standard Errors, $T$ -Values, and $P$ -Values .....	174
6.4	Analysis Of Variance, $R^2$ , and $F$ -Tests .....	177
6.4.1	AOV table .....	177
6.4.2	Sums of squares (SS) and $R^2$ .....	177
6.4.3	Degrees of freedom (DF) .....	178
6.4.4	Mean sums of squares (MS) and testing .....	178
6.4.5	Adjusted $R^2$ .....	179
6.4.6	Sequential and partial sums of squares .....	180
6.5	Regression Hedging* .....	181
6.6	Regression and Best Linear Prediction .....	183
6.7	Model Selection .....	183
6.8	Collinearity and Variance Inflation .....	187
6.9	Centering the Predictors .....	189
6.10	Nonlinear Regression .....	189
6.11	The General Regression Model .....	192
6.12	Troubleshooting .....	193
6.12.1	Influence diagnostics and residuals .....	194
6.12.2	Residual analysis .....	199
6.13	Transform-Both-Sides Regression* .....	206
6.13.1	How TBS works .....	210
6.13.2	Power transformations .....	214
6.14	The Geometry of Transformations* .....	214
6.15	Robust Regression* .....	216
6.16	Summary .....	217
6.17	Bibliographic Notes .....	219
6.18	References .....	220
6.19	Problems .....	220
<b>7</b>	<b>The Capital Asset Pricing Model .....</b>	<b>225</b>
7.1	Introduction to CAPM .....	225
7.2	The Capital Market Line (CML) .....	227
7.3	Betas and the Security Market Line .....	230
7.3.1	Examples of betas .....	231
7.3.2	Comparison of the CML with the SML .....	231

7.4	The Security Characteristic Line .....	232
7.4.1	Reducing unique risk by diversification .....	234
7.4.2	Can beta be negative? .....	235
7.4.3	Are the assumptions sensible? .....	235
7.5	Some More Portfolio Theory .....	236
7.5.1	Contributions to the market portfolio's risk .....	236
7.5.2	Derivation of the SML .....	237
7.6	Estimation of Beta and Testing the CAPM .....	238
7.6.1	Regression using returns instead of excess returns .....	241
7.6.2	Interpretation of alpha .....	241
7.7	Using CAPM in Portfolio Analysis .....	242
7.8	Factor Models .....	242
7.8.1	Estimating expectations and covariances of asset returns .....	244
7.8.2	Fama and French three-factor model .....	245
7.8.3	Cross-sectional factor models .....	246
7.9	An Interesting Question* .....	246
7.10	Is Beta Constant?* .....	250
7.11	Summary .....	252
7.12	Bibliographic Notes .....	254
7.13	References .....	254
7.14	Problems .....	255
<b>8</b>	<b>Options Pricing .....</b>	<b>257</b>
8.1	Introduction .....	257
8.2	Call Options .....	258
8.3	The Law of One Price .....	259
8.3.1	Arbitrage .....	259
8.4	Time Value of Money and Present Value .....	260
8.5	Pricing Calls — A Simple Binomial Example .....	260
8.6	Two-Step Binomial Option Pricing .....	263
8.7	Arbitrage Pricing by Expectation .....	264
8.8	A General Binomial Tree Model .....	266
8.9	Martingales .....	268
8.9.1	Martingale or risk-neutral measure .....	269
8.9.2	The risk-neutral world .....	269
8.10	From Trees to Random Walks and Brownian Motion .....	270
8.10.1	Getting more realistic .....	270
8.10.2	A three-step binomial tree .....	270
8.10.3	More time steps .....	271
8.10.4	Properties of Brownian motion .....	273
8.11	Geometric Brownian Motion .....	273
8.12	Using the Black-Scholes Formula .....	276
8.12.1	How does the option price depend on the inputs? .....	276
8.12.2	Early exercise of calls is never optimal .....	277

8.12.3	Are there returns on nontrading days? .....	278
8.13	Implied Volatility .....	279
8.13.1	Volatility smiles and polynomial regression* .....	280
8.14	Puts .....	284
8.14.1	Pricing puts by binomial trees .....	284
8.14.2	Why are puts different from calls? .....	287
8.14.3	Put-call parity .....	287
8.15	The Evolution of Option Prices .....	288
8.16	Leverage of Options and Hedging .....	289
8.17	The Greeks .....	290
8.17.1	Delta and Gamma hedging .....	294
8.18	Intrinsic Value and Time Value* .....	295
8.19	Summary .....	295
8.20	Bibliographic Notes .....	297
8.21	References .....	298
8.22	Problems .....	298
<b>9</b>	<b>Fixed Income Securities .....</b>	<b>301</b>
9.1	Introduction .....	301
9.2	Zero-Coupon Bonds .....	302
9.2.1	Price and returns fluctuate with the interest rate .....	302
9.3	Coupon Bonds .....	304
9.3.1	A general formula .....	305
9.4	Yield to Maturity .....	305
9.4.1	General method for yield to maturity .....	306
9.4.2	MATLAB functions .....	307
9.4.3	Spot rates .....	308
9.5	Term Structure .....	309
9.5.1	Introduction: Interest rates depend upon maturity .....	309
9.5.2	Describing the term structure .....	310
9.6	Continuous Compounding .....	314
9.7	Continuous Forward Rates .....	315
9.8	Sensitivity of Price to Yield .....	316
9.8.1	Duration of a coupon bond .....	317
9.9	Estimation of a Continuous Forward Rate* .....	317
9.10	Summary .....	322
9.11	Bibliographic Notes .....	323
9.12	References .....	324
9.13	Problems .....	324
<b>10</b>	<b>Resampling .....</b>	<b>327</b>
10.1	Introduction .....	327
10.2	Confidence Intervals for the Mean .....	328
10.3	Resampling and Efficient Portfolios .....	332
10.3.1	The global asset allocation problem .....	332



10.3.2	Uncertainty about mean-variance efficient portfolios...	334
10.3.3	What if we knew the expected returns? .....	337
10.3.4	What if we knew the covariance matrix? .....	338
10.3.5	What if we had more data? .....	339
10.4	Bagging* .....	340
10.5	Summary .....	342
10.6	Bibliographic Notes .....	343
10.7	References .....	343
10.8	Problems .....	343
<b>11</b>	<b>Value-At-Risk .....</b>	<b>345</b>
11.1	The Need for Risk Management .....	345
11.2	VaR with One Asset .....	346
11.2.1	Nonparametric estimation of VaR .....	346
11.2.2	Parametric estimation of VaR .....	348
11.2.3	Estimation of VaR assuming Pareto tails* .....	349
11.2.4	Estimating the tail index* .....	350
11.2.5	Confidence intervals for VaR using resampling .....	353
11.2.6	VaR for a derivative .....	354
11.3	VaR for a Portfolio of Assets .....	355
11.3.1	Portfolios of stocks only .....	355
11.3.2	Portfolios of one stock and an option on that stock ...	356
11.3.3	Portfolios of one stock and an option on another stock	356
11.4	Choosing the Holding Period and Confidence .....	357
11.5	VaR and Risk Management .....	357
11.6	Summary .....	359
11.7	Bibliographic Notes .....	360
11.8	References .....	360
11.9	Problems .....	360
<b>12</b>	<b>GARCH Models .....</b>	<b>363</b>
12.1	Introduction .....	363
12.2	Modeling Conditional Means and Variances .....	364
12.3	ARCH(1) Processes .....	365
12.4	The AR(1)/ARCH(1) Model .....	368
12.5	ARCH( $q$ ) Models .....	370
12.6	GARCH( $p, q$ ) Models .....	370
12.7	GARCH Processes Have Heavy Tails .....	371
12.8	Comparison of ARMA and GARCH Processes .....	372
12.9	Fitting GARCH Models .....	372
12.10	I-GARCH Models .....	377
12.10.1	What does it mean to have an infinite variance? .....	379
12.11	GARCH-M Processes .....	381
12.12	E-GARCH .....	383
12.13	The GARCH Zoo* .....	386

12.14 Applications of GARCH in Finance .....	386
12.15 Pricing Options Under Generalized GARCH Processes* .....	387
12.16 Summary .....	391
12.17 Bibliographic Notes .....	392
12.18 References .....	392
12.19 Problems .....	393
<b>13 Nonparametric Regression and Splines .....</b>	<b>397</b>
13.1 Introduction .....	397
13.2 Choosing a Regression Method .....	400
13.2.1 Nonparametric regression .....	400
13.2.2 Linear .....	400
13.2.3 Nonlinear parametric regression .....	400
13.2.4 Comparison of linear and nonparametric regression ..	401
13.3 Linear Splines .....	405
13.3.1 Linear splines with one knot .....	405
13.3.2 Linear splines with many knots .....	406
13.4 Other Degree Splines .....	407
13.4.1 Quadratic splines .....	407
13.4.2 $p$ th degree splines .....	408
13.5 Least Squares Estimation .....	409
13.6 Selecting the Spline Parameters .....	410
13.6.1 Estimating the volatility function .....	413
13.7 Additive Models* .....	415
13.8 Penalized Splines* .....	418
13.8.1 Penalizing the jumps at the knots .....	420
13.8.2 Cross-validation .....	421
13.8.3 The effective number of parameters .....	423
13.8.4 Generalized cross-validation .....	425
13.8.5 AIC .....	426
13.8.6 Penalized splines in MATLAB .....	427
13.9 Summary .....	431
13.10 Bibliographic Notes .....	431
13.11 References .....	431
13.12 Problems .....	432
<b>14 Behavioral Finance .....</b>	<b>435</b>
14.1 Introduction .....	435
14.2 Defense of the EMH .....	436
14.3 Challenges to the EMH .....	437
14.4 Can Arbitrageurs Save the Day? .....	438
14.5 What Do the Data Say? .....	439
14.5.1 Excess price volatility .....	439
14.5.2 The overreaction hypothesis .....	439
14.5.3 Reactions to earnings announcements .....	440

14.5.4	Counter-arguments to pricing anomalies .....	441
14.5.5	Reaction to non-news .....	442
14.6	Market Volatility and Irrational Exuberance .....	443
14.6.1	Best prediction .....	444
14.7	The Current Status of Classical Finance .....	445
14.8	Bibliographic Notes .....	445
14.9	References .....	446
14.10	Problems .....	447
<b>Glossary</b> .....		449
<b>Index</b> .....		461