

Tools from Stochastic Analysis for Mathematical Finance: A Gentle Introduction

Laura Ballotta¹ Gianluca Fusai²

May 23, 2018

¹Faculty of Finance, Cass Business School, City University London, UK; Email: L.Ballotta@city.ac.uk; Cass Experts: Laura Ballotta

²Faculty of Finance, Cass Business School, City University London, UK; Email: Gianluca.Fusai.1@city.ac.uk; Cass Experts: Gianluca Fusai, and Dipartimento SEI, Università del Piemonte Orientale, Novara, Italy; Email: gianluca.fusai@uniupo.it; Upobook: Gianluca Fusai

FOREWORD

The idea of this document is to provide the reader with an intuitive, yet rigorous and comprehensive introduction to the main tools in stochastic analysis required in Finance to understand the modern modelling, pricing and hedging techniques.

We would like to emphasize that this document is very much work in progress and we would like to encourage readers to get in touch with us with feedback, comments, suggestions for additions and, of course, corrections of typos. All of these will be gratefully acknowledged in the future releases of this document.

Few implemented examples for each model discussed here can be found on the website
<https://sites.google.com/uniupo.it/stochasticcalculus/home?authuser=1>.

If you have suggestions please contact us at the following email: stochasticcalculus@gmail.com

Laura Ballotta and Gianluca Fusai

This Version: 1.1

PART I

A Gentle Introduction to Stochastic Calculus in Continuous Time

CONTENTS

I A Gentle Introduction to Stochastic Calculus in Continuous Time	3
1 Aims	7
2 The Brownian motion	9
2.1 Defining the Brownian Motion	10
2.1.1 The odd properties of the Brownian Motion	10
2.1.2 Density of the Brownian Motion at different Time Horizons	11
2.1.3 The (auto)-covariance function	12
2.1.4 Correlated Brownian motions	13
2.1.5 Martingale property	14
2.1.6 Scaling property	15
2.1.7 Markov property	16
2.1.8 Simulation of Brownian sample paths	17
2.1.9 Total Variation	17
2.1.10 Quadratic Variation	19
2.1.11 Properties of the Increments of the BM	23

3 The Stochastic Integral and Stochastic Differential Equations	25
3.1 Introduction	26
3.2 Defining the Stochastic Integral	26
3.3 The Ito stochastic integral as a mean square limit of suitable Riemann-Stieltjes sums	27
3.4 A motivating example: Computing $\int_0^t W(s)dW(s)$	28
3.4.1 Example: Computing $\int_0^t W(s)dW(s)$ by simulation	28
3.5 Properties of the Stochastic Integral	28
3.6 Stochastic integrals as Time Changed Brownian Motions	32
3.7 Ito process and Stochastic Differential Equations	36
3.8 ABM with deterministic volatility	36
3.8.1 Matlab: Simulating the Arithmetic Brownian Motion	39
3.9 Solving Stochastic Integrals and/or Stochastic Differential Equations	42
3.9.1 Deterministic vs. Stochastic Differential Equations	43
3.9.2 Examples of stochastic differential equations	45
3.9.3 Existence and uniqueness of the solution	45
4 Introducing Ito's formula	47
4.1 A fact from ordinary calculus.	48
4.2 Itô's formula when $Y = g(X), g(X) \in C^2$	49
4.3 Guiding principle	50
4.4 Itô's formula when $Y(t) = g(t, X), g(t, X) \in C^{1,2}$	50
4.5 The Lamperti transformation	52
4.5.1 Example	53
4.6 The Multivariate Ito's lemma when $Z = g(t, X, Y)$	54
4.6.1 Examples	56
5 Important SDEs	58
5.1 The Geometric Brownian Motion $GBM(\mu, \sigma)$	60
5.1.1 Solving the ODE $dX(t) = \mu X(t)dt$	61

5.1.2	Solving the SDE $dX = \mu X dt + \sigma X dW$	61
5.1.3	Matlab Implementation: Simulating GBM	64
5.1.4	Remark. GBM with deterministic drift and volatility	68
5.2	The Vasicek Mean-Reverting process	75
5.2.1	A Note: The Ordinary Differential Equation $dx(t) = \alpha(\mu - x(t)) dt$	76
5.2.2	Solving the SDE $dX(t) = \alpha(\mu - X(t)) dt + \sigma dW(t)$	79
5.2.3	The (auto)-covariance function	80
5.2.4	Matlab: Simulation of the Vasicek model	81
5.2.5	Extension: MR with deterministic volatility	87
5.3	The Cox-Ingersoll-Ross (CIR) model	89
5.3.1	Solving the SDE $dX(t) = \alpha(\mu - X(t)) dt + \sigma\sqrt{X(t)}dW(t)$	89
5.3.2	Computing the expectation of the CIR model	90
5.3.3	Computing the variance of the CIR model	91
5.3.4	Again on computing the variance and the auto-covariance of the CIR model	92
5.3.5	The distribution of the short rate in the CIR model	94
5.3.6	The Extended-CIR model	97
5.3.7	Simulating the CIR model	100
5.4	The Constant Elasticity of Variance (CEV) model	106
5.5	CIR, CEV and the Bessel Process	116
5.6	The Brownian Bridge	118
5.6.1	A Note: The Ordinary Differential Equation $dx(t) = (b - x(t))/(T - t)dt$	118
5.6.2	Solving the SDE $dX_t = (b - X(t))/(T - t)dt + dW(t)$	119
5.6.3	Matlab Implementation: Simulating Brownian motions (part 2)	121
5.7	The Heston Stochastic Volatility model	126
5.7.1	The characteristic function of the log-price	128
5.7.2	Heston model: Monte Carlo Simulation	134
5.7.3	Recovering the density function in the Heston model via inversion of the characteristic function	138
5.7.4	Heston model and option pricing	144
5.7.5	Heston model and option pricing: Main findings	147

5.8 The Heston app	152
6 Stochastic Processes with Jumps	155
6.1 Preliminaries	156
6.1.1 The Poisson process	156
6.1.2 The Compound Poisson process	158
6.1.3 The Gamma process	159
6.1.4 Simulation of Jump Processes	160
6.2 Jump Diffusion (JD) processes	163
6.2.1 The Merton Jump Diffusion Process	163
6.2.2 The Kou process	173
6.3 Subordinated Brownian motions	176
6.3.1 The Variance (VG) Gamma process	177
6.4 Final Remark: Lévy processes	184
6.5 Jumps versus Stochastic Volatility	185
7 Changes of Measure	191
7.1 Preliminaries	192
7.2 Girsanov Theorem for Brownian motions	196
7.2.1 Application 1: Risk neutral measures	198
7.2.2 Application 2: The Numéraire Pair	200
8 References	213
9 Matlab LiveScripts	216
II Appendix	218
10 A Quick Review of Distributions Relevant in Finance	219

10.1 The Normal Distribution	221
10.2 The Lognormal distribution	221
10.3 The Chi-Square distribution	227
10.4 The noncentral chi-squared distribution	229
10.5 The Poisson distribution	233
10.6 The Exponential distribution	234
10.7 The Gamma distribution	238
10.8 The multivariate Gaussian distribution	241
10.8.1 The bivariate Normal distribution	242
10.9 Simulating Random Variables	250
10.9.1 Example: Sampling from the Normal Distribution	252
10.9.2 Example: Sampling from the Lognormal Distribution	255
10.9.3 Example: Sampling from the Poisson Distribution	256
10.9.4 Example: Sampling from the Gamma Distribution	256
10.9.5 Example: Sampling from the Chi-Square Distribution	256
10.9.6 Example: Sampling from the Non-Central Chi-Square Distribution	256
10.9.7 Example: Sampling from the Multivariate Normal Distribution	258
11 Quadrature Methods	262
11.1 Quadrature Rules	262
11.2 Newton-Cotes integration	263
11.2.1 Rectangle and Mid-Point Rule	264
11.2.2 Trapezoid Rule	266
11.2.3 Error analysis for polynomial rate of convergence	267
11.2.4 An example	268
11.2.5 Newton-Cotes integration	269
11.2.6 Limits of Newton-Cotes rules	270
11.3 Gaussian Quadrature Formulae	272
11.4 Adaptive Quadrature	273

11.5 An example	275
11.5.1 Newton-Cotes integration	275
11.6 Gaussian quadrature	277
11.6.1 Adaptive Quadrature	279
12 Inversion of the characteristic function and the COS method	281
12.1 Characteristic function and density function	281
12.2 Characteristic function and option pricing	282
12.2.1 FFT method in MATLAB® Algorithms	287
12.2.2 Inversion of the characteristic function and the COS method	291

LIST OF FIGURES

2.1	Density of the Brownian Motion at different time horizons.	12
2.2	Simulating the Brownian Motion: Excel example.	18
2.3	Simulated Paths of the Brownian Motion.	19
2.4	Movie of simulated variance in the Heston model.	20
2.5	Total and Quadratic Variation of Brownian motions	22
3.1	Computing by simulation the stochastic integral $\int_0^t W(s)dW(s)$	29
3.2	Simulated sample paths of the ABM(0.2,0.3).	42
3.3	Simulated and theoretical expected value of $X(t)$ of the ABM(0.2,0.3) at different horizons.	42
3.4	Simulated and theoretical variance of the ABM(0.2,0.3) at different horizons.	43
3.5	Simulated and theoretical standard deviation of the ABM(0.2,0.3) at different horizons.	43
3.6	Simulated distribution of the ABM(0.2,0.3) at different horizons.	44
3.7	Time evolution of the exact distribution of the ABM(0.2,0.3).	44
5.1	Sample Paths of the GBM(0.2,0.3).	68
5.2	Density of the GBM(0.2,0.3) at different horizons.	68
5.3	Theoretical and Monte Carlo estimated expected value of the GBM(0.2,0.3) at different horizons.	69
5.4	Theoretical and Monte Carlo estimated variance of the GBM(0.2,0.3) at different horizons.	69
5.5	Theoretical and estimate mean of the GBM(0.2,0.3) at different horizons.	70
5.6	Theoretical and estimate variance of the GBM(0.2,0.3) at different horizons.	70
5.7	Simulated GBM paths fitting in average the term structure of futures prices	75

5.8 Mean reversion and expected change in the state variable (here an interest rate).	77
5.9 Convergence to the long-run value of a mean-reverting process	78
5.10 The Vasicek model: sample trajectories	82
5.11 Mean reverting processes: densities	83
5.12 Movie with simulated paths of the Vasicek model starting above the long term level.	85
5.13 Movie with simulated paths of the Vasicek model starting at the long term level.	86
5.14 Movie with simulated paths of the Vasicek model starting below the long term level.	87
5.15 CIR model: densities	97
5.16 Asymptotic density in the CIR model	98
5.17 Violation of the Feller condition: density of the CIR process	98
5.18 CIR process: sample trajectories via Euler scheme	103
5.19 CIR process: sample trajectories via exact moments	104
5.20 CIR process: sample trajectories via sampling from non-central chi-square distribution	105
5.21 Density function of the CEV model	112
5.22 Density function of log-returns in the CEV model	112
5.23 Implied Volatility in the CEV model for different values of β	114
5.24 Simulated Paths of the Brownian motion via Brownian Bridge.	123
5.25 Scenario Generation via Brownian bridge	124
5.26 Mean, standard deviation, skewness and kurtosis of the Heston model varying the correlation coefficient ρ	133
5.27 Term structure of skewness in the Heston model for different values of the correlation coefficient ρ	134
5.28 Term structure of kurtosis in the Heston model for different values of the correlation coefficient ρ	134
5.29 Skewness in the Heston model for different values of the vol-vol parameter ε	135
5.30 Kurtosis in the Heston model for different values of the vol-vol parameter ε	135
5.31 Mean, standard deviation, skewness and kurtosis of the Heston model varying the volatility coefficient ε	136
5.32 Movie comparing the Euler and the exact scheme.	139
5.33 Movie of simulated log-returns in the Heston model.	140
5.34 Movie of simulated variance in the Heston model.	141
5.35 Heston vs Black-Scholes: comparing densities	145
5.36 Heston PDF changing ρ	149

5.37 Heston PDF changing ε	149
5.38 Heston PDF changing κ	149
5.39 Heston PDF changing θ	149
5.40 Implied Volatility as a function of ρ in the Heston model	150
5.41 Implied Volatility as a function of ϵ in the Heston model	150
5.42 Implied Volatility as a function of θ in the Heston model	151
5.43 Implied Volatility as a function of κ in the Heston model	151
5.44 The app can be downloaded from here https://sites.google.com/uniupo.it/stochasticcalculus/download-apps . Its functioning has been checked on Windows 64 bit.	154
6.1 ABM vs JD: sample trajectories	157
6.2 Poisson, Compound Poisson and Gamma processes: sample trajectories	162
6.3 Merton JD: sample trajectories	170
6.4 Merton JD vs the Brownian motion: density and returns	171
6.5 VG: term structure of skewness and excess kurtosis	180
6.6 VG vs Brownian motion: densities	181
6.7 VG process: sample trajectories	182
6.8 Implied Volatility as a function of θ in the VG model	188
6.9 VG PDF changing θ	188
6.10 Implied Volatility as a function of k in the VG model	188
6.11 VG PDF changing σ	188
6.12 Implied Volatility as a function of σ in the VG model	189
6.13 VG PDF changing k	189
6.14 Convergence of Lévy processes to the Brownian motion	190
7.1 Girsanov theorem: Densities and distribution under change of measure	211
7.2 Girsanov theorem: simulated paths under changes of measure	212
10.1 Density and Cumulative Distribution of the Gaussian random variable.	223
10.2 Densities of the Gaussian and of the Lognormal random variables.	223

10.3 Density and Cumulative Distribution of the Lognormal random variable.	226
10.5 Density of the Chi-Square random variable varying n	228
10.4 Density and Cumulative Distribution of the Chi-Square random variable with $n = 20$ degrees of freedom.	229
10.6 Density and CDF of the noncentral chi-squared distribution with $n = 5$ and $d = 2$	231
10.7 Density of the non-central chi-square distribution varying the parameter of non centrality d	232
10.8 Density and CDF of the Poisson distribution with $\lambda = 5$	234
10.9 PDF (top) and CDF (bottom) of the Poisson distribution varying the rate of arrival λ	235
10.10 Density and CDF of the Exponential distribution with $\lambda = 1.5$	237
10.11 Density and CDF of the Gamma distribution with $\alpha = 5, \lambda = 0.5$	239
10.12 PDF and CDF of the Gamma distribution varying the shape parameter α and the rate parameter λ	240
10.13 Bivariate standard Gaussian distribution and contour levels. $\rho = 0.9$	245
10.14 Bivariate standard Gaussian distribution and contour levels. $\rho = 0.0$	246
10.15 Bivariate standard Gaussian distribution and contour levels. $\rho = -0.9$	247
10.16 Jointly simulated standard Gaussian random variables with different correlation	248
10.17 Gaussian margins, non Gaussian joint distribution	250
10.18 Non Gaussian convolution	250
10.19 Inverse transform method for a hypothetical CDF F_X	251
10.20 Sampling from the standard Gaussian distribution.	253
10.21 Sampling from the Gaussian distribution with $\mu = 0.2$ and $\sigma = 0.1$	254
10.22 Sampling from the Lognormal distribution with parameters $\mu = 0.05$ and $\sigma = 0.2$	255
10.23 Sampling from the Poisson distribution with $\lambda = 10$	257
10.24 Sampling from the Gamma distribution with $\alpha = 2, \lambda = 0.5$	258
10.25 Sampling from the bivariate Gaussian distribution (positive correlation).	261
11.1 Left: Rectangle rule, Right: Mid-point.	265
11.2 Approximation of the integral using trapezium.	266
11.3 Matlab adaptive routine quad: integration of $\sin(1/x)$	274
12.1 Black–Scholes price–FFT price	290
12.2 COS Method: error in function of N in approximating the standard Normal distribution	297

12.3 COS Method: error in function of N in approximating the standard Normal distribution	298
12.4 COS Method: error in function of N in approximating the Merton Jump Diffusion distribution	299
12.5 COS Method: error in function of N in approximating the Merton Jump Diffusion distribution	300

LIST OF TABLES

2.1	Calculus with dW	24
3.1	Properties of the stochastic integral.	33
3.2	Properties of the arithmetic Brownian motion process.	38
3.3	Examples of deterministic and stochastic differential equations.	45
5.1	Relevant SDEs for Financial Modelling	60
5.2	Properties of the geometric Brownian motion process.	64
5.3	Light Sweet Crude Oil (WTI) Futures prices quoted at CME as of march 12, 2013	72
5.4	Monte Carlo simulation: Futures prices	74
5.5	Properties of the Vasicek model	84
5.6	Properties of the Square-Root model	102
5.7	Moments of log-returns in the CEV model varying the leverage parameter β	113
5.8	Properties of the Constant Elasticity of Variance CEV(μ, σ, β)	115
5.9	Moments of the Heston model	132
5.10	Moments computation by inversion of the characteristic function in the Heston model	144
5.11	Prices of European at-the-money call and put options in the Heston model	147
5.12	Properties of the Heston model.	153
6.1	Properties of the Gamma process.	161
6.2	Sample moments of daily log-price changes in oil price in year 2012.	167

6.3 Calibrated parameters of the MJD model to sample moments of daily log-price changes in oil price for year 2012.	168
6.4 Properties of the Merton Jump-Diffusion model.	172
10.1 Probability intervals of the Gaussian distribution.	222
10.2 Moments of the Gaussian distribution.	222
10.3 Properties of the Lognormal distribution.	225
10.4 Moments of the Chi-square distribution.	228
10.5 Moments of the noncentral chi-squared distribution.	231
10.6 Moments of the Poisson distribution.	233
10.7 Moments of the Exponential distribution.	236
10.8 Moments of the Gamma distribution.	239
11.1 Numerical integration of the integral in 11.7.	270
11.2 Numerical integration error in the computation of the integral in 11.7 using different rules.	270
11.3 Comparing quadrature methods	271
11.4 Trapezoid Rule: Estimate of the order of convergence	271
11.5 Numerical integration of the integral in 11.7.	276
11.6 Numerical integration error in the computation of the integral in 11.7 using different rules.	276
11.7 Comparing quadrature methods	277
11.8 Gauss-Legendre integration with 5 points	278
11.9 Numerical Integration of the function 11.7 usign a Gauss-Legendre rule	279
11.10 Syntax of the MATLAB adaptive quadrature routines.	279
12.1 Elementary properties of the characteristic function.	282
12.2 Pseudo-code for option pricing via the FFT algorithm.	287
12.3 Merton Jump Diffusion distribution: COS approximation	301
12.4 Pricing options with COS: accuracy	302

CHAPTER 1. AIMS

1. To introduce the concept of Brownian motion.

- What $W(t)$ and $dW(t)$ are;
- What the properties of $W(t)$ are.

2. Explaining the meaning of the stochastic differential equation

$$dX = \mu(X, t) dt + \sigma(X, t) dW(t).$$

3. Explaining the meaning of the stochastic integral, specifically

- To give a meaning to

$$X(T) = X(0) + \int_0^T \mu(X_s, s) ds + \int_0^T \sigma(X_s, s) dW_s;$$

- Properties of $X(T)$: distribution, expected value, variance.
- How to construct a (continuous) martingale.

4. Explaining the Ito's formula.

- For example, how to relate

$$dS = \mu S dt + \sigma S dW$$

to

$$d \ln S = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW$$

5. Examples of common stochastic differential equations.

- Arithmetic Brownian Motion;
- Geometric Brownian Motion;
- Mean-Reverting Gaussian Model (Vasicek);
- Mean-Reverting Square-Root Model (CIR);
- Constant Elasticity of Variance Model (CEV);
- Stochastic Volatility Model (Heston);

6. Examples of jump processes.

- Poisson Process;
- Poisson Compound Process;
- Gamma Process;
- Jump Diffusion Process (Merton);
- Jump Diffusion Process (Kou double exponential);
- Subordinated Brownian motions (Variance Gamma Model);

CHAPTER 2. THE BROWNIAN MOTION

Contents

2.1 Defining the Brownian Motion	10
2.1.1 The odd properties of the Brownian Motion	10
2.1.2 Density of the Brownian Motion at different Time Horizons	11
2.1.3 The (auto)-covariance function	12
2.1.4 Correlated Brownian motions	13
2.1.5 Martingale property	14
2.1.6 Scaling property	15
2.1.7 Markov property	16
2.1.8 Simulation of Brownian sample paths	17
2.1.9 Total Variation	17
2.1.10 Quadratic Variation	19
2.1.11 Properties of the Increments of the BM	23

In this chapter we define the Brownian Motion (BM) and we will present its main properties. Let us recall that a stochastic process (s.p.) is a family of indexed random variables (r.v.'s). Usually, but it is not necessary, the index refers to time. If we use the time index, as we will do, this means that at each instant of time (if the stochastic process is continuous in time) or at discrete times (if the stochastic process is discrete in time) we have to draw a r.v.. Once we have extracted them, we observe a path or a trajectory of the s.p.. We have to stress that in order to fully characterize the s.p. we have to specify the distribution of the r.v. at different times as well as their time dependence properties, i.e. how the r.v. that will be extracted in 1 month will affect the value of the r.v. that will be extracted in 2 months. The simplest process in continuous time is the Brownian motion.

A comprehensive introduction to the BM, stochastic differential equations and and their properties with a range of financial applications is provided by the book of S. Shreve Stochastic Calculus for Finance II, Springer Finance, 2004.

2.1. DEFINING THE BROWNIAN MOTION

Fact 1 (Brownian Motion) *The stochastic process $W := \{W(t) : t \geq 0\}$ is called Brownian motion if*

- i) $W(0) = 0$
- ii) *for $s \leq t$, $W(t) - W(s)$ is independent of the past history of W until (and including) time s , i.e. the Brownian motion has increments which are independent of the σ -algebra $\mathcal{F}(s)$ generated by the process up to time s .*
- iii) *for $0 \leq s \leq t$, $W(t) - W(s)$ and $W(t-s)$ have the same distribution, which is Gaussian with mean zero and variance $(t-s)$, i.e.*

$$W(t) - W(s) \sim W(t-s) \sim \mathcal{N}(0, t-s).$$

In other words. the Brownian motion has stationary increments.

- iv) *W has continuous sample paths: $W(t), t \geq 0$ is a continuous functions of t (a.s.).*

2.1.1. THE ODD PROPERTIES OF THE BROWNIAN MOTION

- $W(t)$ is continuous in t (by definition); it is also a continuous state space process, i.e. its trajectories do not show sudden jumps.

- $W(t)$ is not monotone in any interval, no matter how small the interval is (like in a fractal structure). This property means that, even over very small time intervals, we cannot say if the BM is increasing or decreasing.
- $W(t)$ is nowhere differentiable. This property means that we cannot predict the movement of the BM over the next time interval.
- $W(t)$ is a process of unbounded variation, i.e. the space it travels over any finite time period is infinite. This property is a consequence of the two previous properties. BM paths are so irregular that if you try to measure their length you do not obtain a finite measure.
- $W(t)$ is a process of bounded quadratic variation. Quadratic variation, as we will see shortly, is connected to the variance of the process. Therefore, even being a very irregular process, BM at each finite time has a finite variance.
- Brownian motion has the Markov property. Broadly speaking, this means that the information we have on the past paths of a BM is not useful in determining the current value of the process.
- $\mathbb{E}(W(t)W(s)) = \min(s, t)$. This property, discussed later, is telling us that the time dependence of the Brownian motion depends on what happened up to $\min(s, t)$.
- The BM will eventually hit any and every real value no matter how large, or small. The BM paths are so dispersed that the BM can reach any desired level in finite time with positive probability.
- Once a BM hits a value, it immediately hits it again infinitely often.

2.1.2. DENSITY OF THE BROWNIAN MOTION AT DIFFERENT TIME HORIZONS

Fact 2 (Density function of the Brownian Motion) *From the definition it follows that*

$$W(t) - W(s) \sim \mathcal{N}(0, t-s).$$

and therefore the density function of the increment of the BM between time $s > 0$ and time $t > s$ is

$$\phi_{0,\sqrt{t-s}}(x) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{1}{2} \left(\frac{x}{\sqrt{t-s}}\right)^2\right), x \in \mathfrak{R}.$$

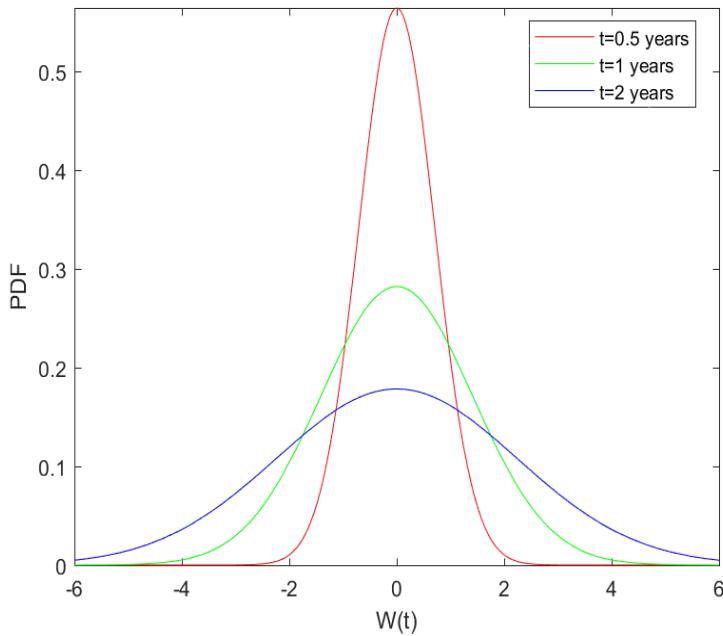


Figure 2.1: Density of the Brownian Motion at different time horizons.

Remark 3 We observe that $W(t) = W(t) - W(0)$, from which it follows that $W(t) \sim \mathcal{N}(0, t)$. The dispersion of the BM, as measured by the variance, increases with the time horizon t . This is shown in Figure (2.1).

2.1.3. THE (AUTO)-COVARIANCE FUNCTION

- The (auto)-covariance function of a stochastic process X having expectation μ_X is defined as

$$c_X(t, s) = cov(X_t, X_s) = E((X_t - \mu_X(t))(X_s - \mu_X(s))), t, s \in T.$$

- The variance function of X is just $c_X(t, t)$.

Fact 4 (Covariance function of the Brownian Motion) *A Brownian Motion has covariance function given by*

$$c_W(t, s) = \min(t, s).$$

Proof: The (auto)-covariance function of the BM

- Let us consider two time instants, t and s , $t < s$. We have

$$\begin{aligned} c_W(t, s) &= \text{cov}(W(t), W(s)) \\ &= \text{cov}(W(t), (W(s) - W(t)) + W(t)) \\ &= \text{cov}(W(t), (W(s) - W(t))) + \text{cov}(W(t), W(t)) \\ &= 0 + c_W(t, t) \\ &= t, \end{aligned}$$

where the penultimate equality follows from the definition of BM.

- With a similar reasoning, if we take t and s with $s < t$, we have

$$c_W(t, s) = s.$$

- Therefore

$$c_W(t, s) = \min(t, s).$$

2.1.4. CORRELATED BROWNIAN MOTIONS

Proposition 5 *In order to construct two Brownian motions, $W(t)$ and $X(t)$, such that*

$$\text{Corr}(W(t), X(t)) = \rho, \text{ i.e. } \text{Cov}(W(t), X(t)) = \rho t,$$

we may take

$$W(t) = \rho X(t) + \sqrt{1 - \rho^2} Z(t),$$

where $Z(t)$ is another Brownian motion independent of $X(t)$.

To see this we note the following.

- Any process of the form $W(t) = aX(t) + bZ(t)$ possesses the stationary, independent increments property, has continuous sample paths and starts from 0 at time t .
- The distribution of $W(t) = aX(t) + bZ(t)$ is $N(0, (a^2 + b^2)t)$.
- Choosing $a = \rho$, $b = \sqrt{1 - \rho^2}$ implies that $\text{Var}(W(t)) = t$, so that $W(t)$ is a standard Brownian motion.
- Further

$$\text{Cov}(W(t), X(t)) = \text{Cov}(\rho X(t) + \sqrt{1 - \rho^2} Z(t), X(t)) = \rho \text{Var}(X(t)) = \rho t.$$

2.1.5. MARTINGALE PROPERTY

Remark 6 A stochastic process $X = (X(t) : t \geq 0)$ is a martingale with respect to the probability measure \mathbb{P} if the following conditions are satisfied.

- i) $\mathbb{E}|X(t)| < \infty$ for all $t \geq 0$;
- ii) $\mathbb{E}[X(t) | \mathcal{F}(s)] = X(s)$ for all $s \leq t$.

- In other words, a martingale is a random process whose future variations are completely unpredictable given the current information set.
- For this reason, the best forecast of the change in X over an arbitrary interval is zero, i.e. $\mathbb{E}[X(t) - X(s) | \mathcal{F}(s)] = 0$.
- A martingale represents a fair game: given the knowledge we currently have, on average the return produced by the investment is what we invested in it.

Fact 7 (Martingale Property) The Brownian motion is a martingale.

Proof: The martingale property of the BM

We check that the Brownian motion satisfies the two properties of a martingale process. Hence.

- For the first property, direct calculations show us

$$\begin{aligned}\mathbb{E}|W(t)| &= \int_{-\infty}^0 \frac{-y}{\sqrt{2\pi t}} \exp\left(-\frac{1}{2}\left(\frac{y}{\sqrt{t}}\right)^2\right) dy + \int_0^\infty \frac{z}{\sqrt{2\pi t}} \exp\left(-\frac{1}{2}\left(\frac{z}{\sqrt{t}}\right)^2\right) dz \\ &= 2 \int_0^\infty \frac{z}{\sqrt{2\pi t}} \exp\left(-\frac{1}{2}\left(\frac{z}{\sqrt{t}}\right)^2\right) dz \\ &= \sqrt{\frac{2t}{\pi}} < \infty.\end{aligned}$$

- Let us consider two time instants, s and t , $s < t$. We have

$$\mathbb{E}[W(t)|\mathcal{F}(s)] = \mathbb{E}[W(t) - W(s) + W(s)|\mathcal{F}(s)] = \mathbb{E}[W(t) - W(s)] + W(s) = W(s),$$

where the last two equalities follow from the definition of BM.

2.1.6. SCALING PROPERTY

Fact 8 (Scaling Property) *Let $c > 0$ and $W(t)$ a Brownian motion. Then the process $B(t) = W(ct)/\sqrt{c}$ is a Brownian motion.*

Proof: The scaling property of the BM

We check that the process $B(t)$ satisfies the definition of Brownian motion. Indeed it starts from zero and it is Gaussian; further.

- The moments are.

$$\begin{aligned}\mathbb{E}(B(t)) &= \mathbb{E}(W(ct))/\sqrt{c} = 0 \\ &\quad \text{as } W \text{ is a Brownian motion} \\ \mathbb{V}ar(B(t)) &= \mathbb{V}ar(W(ct))/c = t.\end{aligned}$$

- Finally, the process $B(t)$ has independent and stationary increments due to again the fact that $W(t)$ is a Brownian motion.

Remark 9 We can rewrite the scaling property given above as $\sqrt{c}B(t) = W(ct)$. Intuitively, this relation suggests that rescaling a Brownian motion by a positive constant is equivalent to observing the Brownian motion evolving according a rescaled time. The expression $W(ct)$, in fact, shows that the process W does not evolve according to the traditional calendar clock t , but a rescaled clock ct . The bigger c , the faster the rescaled clock, the bigger the variance of $\sqrt{c}B(t)$.

We will discuss in further details the change of clock from t to a rescaled version later in relation to solving stochastic differential equations. Indeed, the above result is a particular case of the Dambis, Dubins-Schwarz Theorem given in section 3.6.

2.1.7. MARKOV PROPERTY

Remark 10 Let $\mathcal{F}(s)$ denote the σ -field generated by the process up to time s . The process $\{X\}_{t \geq 0}$ has the Markov property if, for $t \geq s$, the conditional distribution of $X(t)$ given $\mathcal{F}(s)$ is the same as the conditional distribution of $X(t)$ given $X(s)$

$$P(X(t) \leq y | \mathcal{F}(s)) = P(X(t) \leq y | X(s)), \quad a.s.$$

- In practice, this means the process does not remember how it got to the current state x .

Fact 11 (Markov Property) Brownian motion has the Markov property.

- This property is also important for simulating the Brownian motion, because we can iteratively add to the last simulated value of the Brownian motion a new independently simulated increment. The increment of the BM in the time interval $(t, t + dt)$ does not depend on the past history up to t . Therefore, the value of the process at time $t + dt$ will be given by the value up to time t plus the independent increment in $(t, t + dt)$.

2.1.8. SIMULATION OF BROWNIAN SAMPLE PATHS

In order to simulate a Brownian motion sample path $W(t)$ on the interval $[0, T]$ we can adopt the following steps.

1. Choose an integer n and let $\Delta t = \frac{T}{n}$, so that $t_i = i\Delta t$, for $i = 0, 1, \dots, n$.
2. Generate a sequence $\varepsilon_1, \dots, \varepsilon_n$ of iid standard normal rvs:
 - (a) Generate a sequence U_1, \dots, U_n of Uniform r.v.'s in the interval $(0, 1)$;
 - (b) Set $\varepsilon_i = \Phi^{-1}(U_i)$, where $\Phi^{-1}(x)$ is the inverse cumulative distribution of the standard Gaussian distribution;
 - (c) Set $dW_i = \varepsilon_i \sqrt{\Delta t}$.
3. Finally recursively construct the sample path of the BM letting
 - $W(0) = 0$
 - $W(t_i) = W(t_{i-1}) + dW_i, \quad i = 1, \dots, n$

Remark 12 Notice that the simulation of the BM as described above is convenient if you are using a spreadsheet or programming language like C, VBA. This is illustrated in Figure (2.2). If you are using Matlab a different approach, avoiding for cycles, is preferred, see Figure (2.3) and the accompanying code.

A movie showing simulated paths of the Brownian motion is shown in figure 2.4.

2.1.9. TOTAL VARIATION

- Given a partition of the interval $[x_0, x_n]$, the variation of a function $f(x)$ measures the total amount of up and down movements, i.e.

$$TV(f) = \sum_{i=0}^N |f(x_i) - f(x_{i-1})|.$$

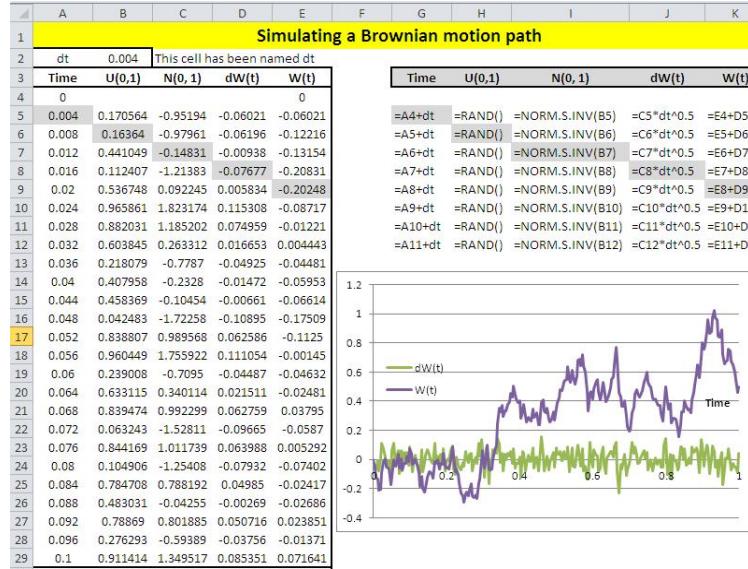


Figure 2.2: Simulating the Brownian Motion: Excel example.

- The notion of unbounded variation plays an important role in stochastic calculus since the majority of continuous time stochastic processes used to represent asset prices have trajectories with unbounded variation i.e. they are very irregular.
- Heuristically, functions of bounded variation are not excessively irregular. In fact, any smooth function will be of bounded variation.

Fact 13 (Total Variation of the Brownian Motion) *The Brownian Motion has unbounded total variation, i.e.*

$$TV(W) = \sum_{i=1}^N |\Delta W(t_i)| \rightarrow \infty$$

as $N \rightarrow \infty$ where t_0, t_1, \dots, t_N represents a discrete partition of the interval $[0, t]$.

Matlab Code

```
%%%%%%%%%%%%%%
%%%SIMULATING THE BROWNIAN MOTION%%%%%
%%%%%
nsimul=50; %Assigning the number of simulated paths
expiry=1; %time to maturity (expiry)
nsteps=250; %number of steps
dt=expiry/nsteps;%time step (dt)
timestep=[0:dt:expiry]';%observation times (timestep)
%Simulate increments of the BM setting:
dw=randn(nsteps,nsimul) * dt^0.5;
%Simulate Brownian motion process:
%use cumulative sum of the increments
cdW=zeros(1,nsimul); cumsum(dw);
%Plot simulated paths:
h=figure('Color',[ 1 1 1]); plot(timestep, cdW)
title('Simulated Paths of the Wiener Process');
xlabel('Time (years)')
%Save the figure in a png format
print(h,'-dpng','FigBMPaths')
```

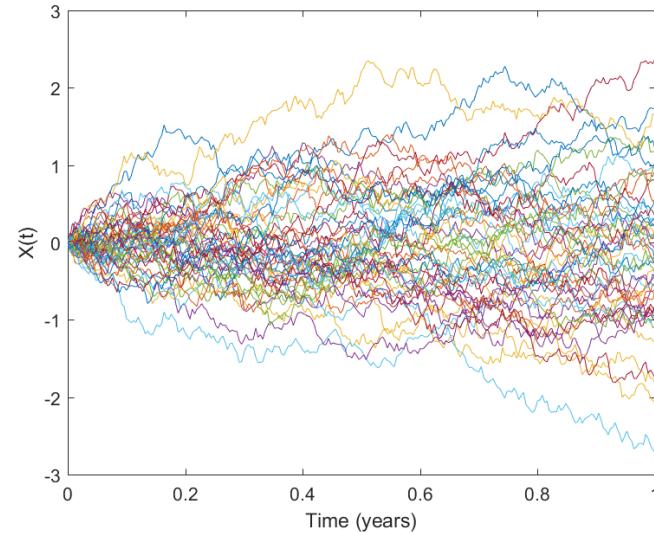


Figure 2.3: Simulated Paths of the Brownian Motion.

2.1.10. QUADRATIC VARIATION

- Quadratic variation is linked to the variance of the process.
- Given a partition of the interval $[t_0, t_n]$, the quadratic variation of a function $f(t)$ is given by

$$QV(f) = \sum_{i=0}^N (f(t_i) - f(t_{i-1}))^2.$$

- QV plays a major role in stochastic calculus, but it is hardly ever meet in standard calculus due to the fact that smooth functions have zero quadratic variation. This makes sense if QV relates to the variance.



Figure 2.4: Movie of simulated variance in the Heston model.

- For example, consider the function $X(t) = at$. This function has QV equal to $\sum_{i=1}^N (a\Delta t_i)^2 = a^2 N \Delta t^2 = a^2 N (\frac{t}{N})^2 \rightarrow 0$ as $N \rightarrow \infty$ (here we have set $\Delta t = t/N$).

Fact 14 (Quadratic Variation of the Brownian Motion) *The quadratic variation of the BM over the time interval $[0, t]$ tends to t*

$$QV(W) = \sum_{i=1}^N (\Delta W(t_i))^2 \rightarrow t$$

as $N \rightarrow \infty$ where t_0, t_1, \dots, t_N represents a discrete partition of the interval $[0, t]$.

Total Variation and Quadratic Variation of the BM: Matlab example

- Figures (2.5) illustrate, by simulation, that:
 - the Total Variation of the BM grows unbounded as we refine the partition of the interval $(0,1)$. Notice that given

$$\mathbb{E} |dW(t_i)| = \sqrt{dt \frac{2}{\pi}},$$

so that (given that $dt = t/N$)

$$\mathbb{E}(TV) = N \sqrt{dt \frac{2}{\pi}} = \sqrt{TN \frac{2}{\pi}},$$

that diverges as we increase the number of points of the partition, for a given horizon T . The blue line in the left panel represents the expected value of the total variation and its simulated values (in red) as the number of points in the partition increases.

- the Quadratic Variation converges to the length of the interval as we refine the partition of the interval $(0,1)$. In this case, it is easy to show that the expected value of the Quadratic Variation converges to T as we increase N .

- The Matlab code to generate these figures is as follows.

Matlab Code

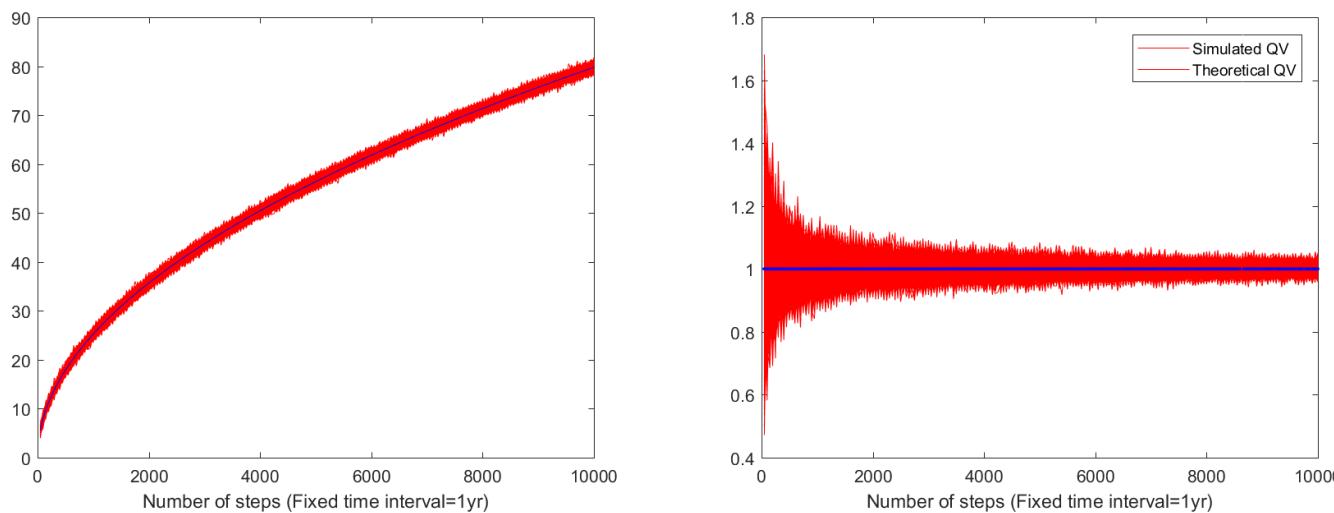


Figure 2.5: Left: the Brownian motion has unbounded total variation. Right: the Brownian motion has bounded quadratic variation

```
%%%%%%%
%%%COMPUTING BY SIMULATION TOTAL VARIATION%%%%%
%%%AND QUADRATIC VARIATION OF THE BM%%%%%
%%%%%%%
clear all;
i=1; %Counter to be used in the for cycle
expiry=1; %Time horizon
nsimul=500;%Number of simulations
N=(50:50:10000);%number of points in the partition
for jstep=N %increase the number of steps (refine the partition)
    dt=expiry/jstep; %reduce the time step
    %simulate increments of the BM
    dW=[zeros(1,nsimul); randn(jstep,nsimul)*dt^0.5];
```

```
%compute total variation
TV(i,:)=sum(abs(dW));
%compute quadratic variation
QV(i,:)=sum(dW.^2);
i=i+1; %update the counter
end
h=figure('Color',[1 1 1]);
plot(N, TV, 'r')
hold on
plot(N, (N*expiry*2/pi).^0.5, 'b') %this is the plot of the E[TV]
title('BM has unbounded total variation');
xlabel('Number of steps (Fixed time interval=1yr)')
print(h, '-dpng', 'LecBM_fig_TV')

h = figure('Color',[1 1 1]);
plot(N,QV,'r',N,expiry,'b.')
legend('Simulated QV', 'Theoretical QV')
title('BM has bounded quadratic variation');
xlabel('Number of steps (Fixed time interval=1yr)')
```

2.1.11. PROPERTIES OF THE INCREMENTS OF THE BM

In the following we set equal to 0 quantities that are $o(dt)$, i.e. that tend to zero faster than dt . Further, we read the quantity ' $dW(t)$ ' as the infinitesimal increment of the Brownian motion (i.e. the increment of the process over an infinitesimal period of time dt).

Fact 15 (Calculus with dW) *We have*

1. $dW(t) \sim \mathcal{N}(0, dt)$ (by definition);
2. $\mathbb{E}(dW(t)) = 0$; (by definition);
3. $\mathbb{E}(dW(t) dt) = 0$; (by the linearity of the expectation);

4. $\mathbb{E}(dW^2(t)) = dt$; (this is the second moment of the increment of the BM: i.e. its variance as the mean is zero by definition);
5. $\text{Var}(dW^2(t)) = \mathbb{E}_t(dW^4(t)) - \mathbb{E}_t(dW^2(t))^2 = 3dt^2 - dt^2 = o(dt)$; (as dW is Gaussian);
6. $dW^2(t) = dt$ (it follows by properties (4) and (5)); i.e. the square of the increment can be approximated by its mean value as the error of doing so, which is measured by $\text{Var}(dW^2(t))$ is negligible.
7. $\mathbb{E}((dW(t)dt)^2) = \mathbb{E}(dW^2(t))dt^2 = dt^3 = o(dt)$ it follows by (4);
8. $\text{Var}(dW(t)dt) = \mathbb{E}((dW(t)dt)^2) - \mathbb{E}((dW(t)dt))^2 = \mathbb{E}(dW^2(t))dt^2 - o(dt) = o(dt)$;
9. $dW(t)dt = o(dt)$ (it follows by (3) and (8): the argument is the same as for point (6) above).
10. We can summarize the above results by using Table (2.1).

\times	dt	$dW(t)$
dt	0	0
$dW(t)$	0	dt

Table 2.1: The cells content represents the product of the quantities appearing in the first row and in the first column. We have set equal to 0 the quantities that are of order dt^n , $n > 1$.

CHAPTER 3. THE STOCHASTIC INTEGRAL AND STOCHASTIC DIFFERENTIAL EQUATIONS

Contents

3.1	Introduction	26
3.2	Defining the Stochastic Integral	26
3.3	The Ito stochastic integral as a mean square limit of suitable Riemann-Stieltjes sums	27
3.4	A motivating example: Computing $\int_0^t W(s)dW(s)$	28
3.4.1	Example: Computing $\int_0^t W(s)dW(s)$ by simulation	28
3.5	Properties of the Stochastic Integral	28
3.6	Stochastic integrals as Time Changed Brownian Motions	32
3.7	Ito process and Stochastic Differential Equations	36
3.8	ABM with deterministic volatility	36
3.8.1	Matlab: Simulating the Arithmetic Brownian Motion	39
3.9	Solving Stochastic Integrals and/or Stochastic Differential Equations	42
3.9.1	Deterministic vs. Stochastic Differential Equations	43
3.9.2	Examples of stochastic differential equations	45
3.9.3	Existence and uniqueness of the solution	45

3.1. INTRODUCTION

- The Ito or Stochastic integral is one way of defining sums of uncountable and unpredictable random increments over time:

$$\underbrace{\int_0^T \sigma(X(s), s) dW(s)}_{\text{sum of iid noises}} \approx \sum_{j=0}^{n-1} \underbrace{\sigma(X(t_j), t_j)}_{\text{vol scaling factor in } t_j} \underbrace{[W(t_{j+1}) - W(t_j)]}_{\text{noise in } [t_j, t_{j+1}]}$$

- Recall that the sample paths of BM are nowhere differentiable and have unbounded variation.
- This has major consequences for the definition of a stochastic integral with respect to Brownian sample paths.

Fact 16 (Problem in defining the stochastic integral) *If $\int_0^T f(s) dg(s)$ exists as a Riemann-Stieltjes integral for all continuous functions f on $[0, T]$, then g necessarily has bounded variation. Unfortunately, the BM has unbounded variation.*

3.2. DEFINING THE STOCHASTIC INTEGRAL

- Last observation is telling us that if our aim is to define the stochastic integral $\int_0^T f(s) dW(s)$ for all continuous deterministic functions f on $[0, T]$, the pathwise integration suggested by the Riemann-Stieltjes integral approach fails since it does not allow the integration of a large class of integrable functions f .
- We will define the integral as a probabilistic average. This will lead us to the so called Ito stochastic integral.
- We define the Ito stochastic integral as a mean square limit of suitable Riemann-Stieltjes sums. This is equivalent to say that the variance of the random error $\varepsilon_n = X_n - X$, goes to zero as we refine the partition.
- This has the disadvantage that we loose the intuitive interpretation of an integral which is naturally provided by a pathwise integral.

Definition 17 (Mean square convergence) *Let us consider a sequence of random numbers $X_1, X_2, \dots, X_n, \dots$ with $X_n \in L^2$ i.e. $\mathbb{E}(X_n^2) < \infty$. We say that the sequence of r.v. X_n converges in mean square to the random variable X , and we write $X_n \xrightarrow{L^2} X$, iff $\lim_{n \rightarrow \infty} \mathbb{E}(|X_n - X|^2) = 0$*

3.3. THE ITO STOCHASTIC INTEGRAL AS A MEAN SQUARE LIMIT OF SUITABLE RIEMANN-STIELTJES SUMS

- Let:

1. $\sigma(u)$ is an adapted process¹;
2. $\sigma(u)$ is square integrable, i.e. $E\left(\int_0^T \sigma^2(u) du\right) < \infty$.

- Then the following limit

$$\lim_{\|\Pi\| \rightarrow 0} \sum_{j=1}^n \sigma(t_{j-1}) [W(t_j) - W(t_{j-1})] = I_\sigma(0, T)$$

exists (in the mean square sense) and is independent of the partitions used to take the limit.

- By definition, this limit is the Ito integral of $\sigma(t)$ with respect to BM over $[0, T]$.

Example 18 Let us consider the Ito integral of the identity function with respect to BM over $[0, T]$. Given a partition $\Pi = \{0 = t_0 < t_1 < \dots < t_{n-1} < t_n = t\}$, we define the Riemann-Stieltjes sums

$$S_n = \sum_{j=1}^n (W(t_j) - W(t_{j-1})).$$

As S_n is a telescopic sum, then $S_n = W(T)$, therefore

$$\int_0^T dW(u) = W(T) - W(0) = W(T).$$

¹i.e. $\sigma(u)$ is known once we know the whole history of the BM up to time t

3.4. A MOTIVATING EXAMPLE: COMPUTING $\int_0^t W(s)dW(s)$

- Given a partition $\Pi = \{0 = t_0 < t_1 < \dots < t_{n-1} < t_n = t\}$, let us define the Riemann-Stieltjes sums:

$$S_n = \sum_{j=1}^n W(t_{j-1})(W(t_j) - W(t_{j-1})).$$

- Using the binomial formula for $(W(t_j) - W(t_{j-1}))^2$, S_n can be written as:

$$S_n = \frac{W^2(t)}{2} - \frac{1}{2} \sum_{i=1}^n (W(t_j) - W(t_{j-1}))^2 = \frac{W^2(t)}{2} - \frac{1}{2} QV_n(W)(t).$$

- But if we refine the partition τ_n , then $QV_n(W)(t) \rightarrow t$, so that

$$\mathbb{V}ar \left(S_n - \frac{W^2(t) - t}{2} \right) \rightarrow 0,$$

i.e. S_n converges in mean square to $\frac{W^2(t) - t}{2}$, which we can take as the value of the integral $\int_0^t W(s)dW(s)$.

3.4.1. EXAMPLE: COMPUTING $\int_0^t W(s)dW(s)$ BY SIMULATION

The following code illustrates the one-to-one relationship between $\int_0^t W(s)dW(s)$ and its value $\frac{W^2(t) - t}{2}$. See also Figure (3.1): the blue line in the plot has a unit slope and goes through the origin. This confirms, by simulation, that the integral, albeit being a random variable, can be written as $\frac{W^2(t) - t}{2}$.

3.5. PROPERTIES OF THE STOCHASTIC INTEGRAL

Fact 19 *The stochastic integral $I_\sigma(0, T)$ has the following properties:*

```
%%%%%%%%%%%%%
%%%Computing Integral[W(s) dW(s),{s,0,T}]%%%%%
%%%%%%%%%%%%%
clear all;
expiry=1;
nstep=10000;
nsimul=10;
%fix the time step
dt=expiry/nstep;
%simulate increments of the BM
dW=randn(nstep,nsimul)*dt^0.5;
%simulate the BM by cumulating the increments
W=[zeros(1,nsimul); cumsum(dW)];
%Approximated value of the stoch. integral by a sum
StocInt=sum(W(1:end-1,:).*dW)';
%Exact Value
ExactInt=(W(end-1,:).^2-expiry)/2;
%Plot the result:
%The simulated values stay on the 45 degree red line.
h = figure('Color',[1 1 1]);
axis equal
plot(StocInt, ExactInt, '.')
hold on
fplot(@(x) x, [-1,6], 'red')
ylabel('\int_0^t W(s)dW(s)')
print(h, '-dpng', 'FigStocIntegral')
```

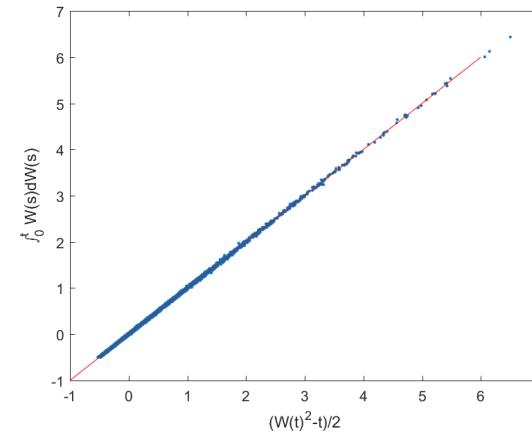


Figure 3.1: Computing by simulation the stochastic integral $\int_0^t W(s)dW(s)$ and comparing to its value $\frac{W(t)^2-t}{2}$.

1. $\mathbb{E}[I_\sigma(0, T)] = 0$.
2. $\mathbb{E}[I_\sigma^2(0, T)] = \mathbb{E}\left[\int_0^T \sigma^2(u)du\right]$ (*Itô isometry*).

3. The Itô integral is Gaussian, for any deterministic integrand function $\sigma(t)$

$$I_\sigma(0, T) \sim \mathcal{N} \left(0, \int_0^T \sigma^2(u) du \right).$$

4. $I_\sigma(0, T)$ is a \mathbb{P} -martingale.

5. The QV of $I_\sigma(0, T)$ over the time interval $[0, T]$ is

$$QV(I_\sigma) = \sum_{i=1}^N \sigma^2(t_{i-1}) (W(t_i) - W(t_{i-1}))^2 \rightarrow \int_0^T \sigma^2(u) du$$

as $N \rightarrow \infty$.

Proof: Properties of the Stochastic Integral

The above mentioned properties can be shown using the definition of Brownian motion, the fact that the integrand function is an adapted process and the properties listed in section 2.1.11.

1. We use the Tower property of the conditional expectation, so that

$$\begin{aligned} \mathbb{E}[I_\sigma(0, T)] &= \int_0^T \mathbb{E}[\sigma(u)dW(u)] \\ &= \int_0^T \mathbb{E}[\mathbb{E}_u(\sigma(u)dW(u))] \\ &= \int_0^T \mathbb{E}[\sigma(u)\mathbb{E}_u(dW(u))] = 0, \end{aligned}$$

where the last equality follows from property 2.1.11.2

2. Given a partition $\Pi = 0 = t_0 < t_1 < \dots < t_{n-1} < t_n = t$, consider the Riemann-Stieltjes sums

$$S_n = \sum_{j=1}^n \sigma(t_{j-1})(W(t_j) - W(t_{j-1})).$$

It follows

$$S_n^2 = \sum_{j=1}^n \sigma^2(t_{j-1})(W(t_j) - W(t_{j-1}))^2 + \sum_j \sum_{i \neq j} \sigma(t_{j-1})\sigma(t_{i-1})(W(t_j) - W(t_{j-1}))(W(t_i) - W(t_{i-1})).$$

Without loss of generality, assume $t_i < t_j$; we note

$$\begin{aligned} \mathbb{E}(S_n^2) &= \mathbb{E}\left[\sum_{j=1}^n \mathbb{E}_{t_{j-1}}\left(\sigma^2(t_{j-1})(W(t_j) - W(t_{j-1}))^2\right)\right] \\ &\quad + \mathbb{E}\left[\sum_j \sum_{i \neq j} \mathbb{E}_{t_{j-1}}\left(\sigma(t_{j-1})\sigma(t_{i-1})(W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1}))\right)\right] \\ &= \mathbb{E}\left[\sum_{j=1}^n \sigma^2(t_{j-1})(t_j - t_{j-1})\right] \end{aligned}$$

where the last equality follows from the definition of Brownian motion. The result follows.

3. If the integrand function is deterministic, the stochastic integral turns out to be a weighted sum of (independent) Gaussian increments, and therefore it is Gaussian as well. The mean and the variance of the stochastic integral follows from the above properties.

4. We note that $I_\sigma(0, T) = I_\sigma(0, t) + I_\sigma(t, T)$. Then

$$\begin{aligned} \mathbb{E}_t[I_\sigma(0, T)] &= I_\sigma(0, t) + \mathbb{E}_t[I_\sigma(t, T)] \\ &= I_\sigma(0, t) + \mathbb{E}_t\left[\int_t^T \mathbb{E}_u(\sigma(u)dW(u))\right] \\ &= I_\sigma(0, t). \end{aligned}$$

5. Given a partition of the interval $[0, T]$, by definition the QV of the process I_σ is given by

$$\begin{aligned} QV(I_\sigma) &= \sum_{i=1}^N (I_\sigma(0, t_i) - I_\sigma(0, t_{i-1}))^2 \\ &= \sum_{i=1}^N \sigma^2(t_{i-1}) (W(t_i) - W(t_{i-1}))^2 \\ &\rightarrow \int_0^T \sigma^2(u) du. \end{aligned}$$

Remark 20 As the Ito integral is a zero mean process, the Ito isometry implies that

$$\mathbb{V}ar(I_\sigma(0, T)) = \mathbb{E} \left[\int_0^T \sigma(u)^2 du \right].$$

Why is this result relevant in finance? It is related to the interpretation of the implied volatility of an option as the average variance of the underlying stock return over the remaining life of the option.

3.6. STOCHASTIC INTEGRALS AS TIME CHANGED BROWNIAN MOTIONS

The following result provides the equivalent for stochastic integrals of the scaling property that we have seen to apply to the Brownian motion (Section 2.1.6).

In order to simplify notation let us denote the QV over the period $[0, t]$ of a stochastic process X as

$$[X]_t = QV(X).$$

Fact 21 (Dambis, Dubins-Schwarz Theorem) Let $M(t)$ be a continuous martingale such that $M(0) = 0$ and $[M]_\infty = \infty$. There exists a Brownian motion W such that, for every $t \geq 0$, $M(t) = W([M]_t)$.

IMPORTANT PROPERTIES OF THE STOCHASTIC INTEGRAL

$I_\sigma(0, t)$ is a stochastic process with continuous sample paths.

$I_\sigma(0, t)$ is a martingale wrt to the Brownian filtration.

$I_\sigma(0, t)$ has zero expectation

$$\mathbb{E}(I_\sigma(0, t)) = 0.$$

$I_\sigma(0, t)$ satisfies the isometry property:

$$\mathbb{V}ar(I_\sigma(0, t)) = \mathbb{E} \left[\int_0^t \sigma^2(u, w) du \right].$$

For constants α and β , $I_\sigma(0, t)$ is linear

$$I_{\alpha\sigma_1+\beta\sigma_2}(0, t) = \alpha I_{\sigma_1}(0, t) + \beta I_{\sigma_2}(0, t).$$

and for adjacent intervals $0 \leq t \leq T$

$$I_\sigma(0, T) = I_\sigma(0, t) + I_\sigma(t, T).$$

Table 3.1: Properties of the stochastic integral.

An intuitive argument for the Dambis, Dubins-Schwarz Theorem

- We have seen in Fact 19 that a continuous martingale can be written in terms of the stochastic integral

$$\int_0^t \sigma(u) dW(u).$$

- Also, it follows from Fact 19 that

$$\left[\int \sigma(u) dW(u) \right]_t = \int_0^t \sigma^2(u) du$$

(Itô isometry)

- According to the Dambis, Dubins-Schwarz Theorem given above, there exists a Brownian motion W such that

$$\int_0^t \sigma(u) dW(u) = W \left(\int_0^t \sigma^2(u) du \right).$$

Note that the two Brownian motions entering both side of the previous equations are not the same; this is a slight abuse of notation.

- Indeed, let $M(t) = W \left(\int_0^t \sigma^2(u) du \right)$. Then, the properties of the Brownian motion imply

$$\mathbb{E}_t(M(t)) = \mathbb{E}_t \left(W \left(\int_0^t \sigma^2(u) du \right) \right) = 0,$$

from which it follows that $\mathbb{E}(M(t)) = 0$.

- Further, the properties of the Brownian motion also imply

$$\text{Var}_t(M(t)) = \text{Var}_t \left(W \left(\int_0^t \sigma^2(u) du \right) \right) = \int_0^t \sigma^2(u) du;$$

consequently $\text{Var}(M(t)) = \mathbb{E} \left(\int_0^t \sigma^2(u) du \right)$ as postulated by the Itô isometry.

- Finally, consider two points in time $s < t$. Then

$$\begin{aligned}\mathbb{E}_s(M(t)) &= \mathbb{E}_s(M(t) - M(s)) + M(s) \\ &= \mathbb{E}_s\left(\mathbb{E}_t\left(W\left(\int_0^t \sigma^2(u)du\right) - W\left(\int_0^s \sigma^2(u)du\right)\right)\right) + M(s) \\ &= M(s)\end{aligned}$$

where the second equality follows from the Tower property (or law of iterated conditional expectations), and the last equality follows from the fact that the increments of the Brownian motion have zero mean. Hence the process $M(t)$ is a martingale.

The meaning of the Dambis, Dubins-Schwarz Theorem is the following.

- Let $T(t) = [M]_t$.
- In analogy with what observed for the scaling property of the Brownian motion in Section 2.1.6, we can interpret the stochastic integral $\int_0^t \sigma(u)dW(u)$ as a Brownian motion evolving on a time scale which is not governed by the standard calendar clock, but by a stochastic clock $T(t)$, i.e. $W\left(\int_0^t \sigma^2(u)du\right)$.
- Thus, if we model prices using a stochastic integral, implicitly we recognize that price changes are random due to two reasons.
- Firstly, there is uncertainty related to the time at which the next investor enters the market placing a transaction which is going to alter the current level of the prices.
- Secondly, there is uncertainty related to the magnitude of the change.
- $T(t)$ models the first effect. The Brownian motion W models the second one.
- In Finance the process $T(t)$ is referred to as business time.

The process $W(T(t))$ is referred to as Time Changed Brownian Motion, $T(t)$ is the time change, and the function $\sigma^2(t)$ in this context is referred to as the activity rate of the stochastic clock. The more the process $\sigma^2(t)$ varies, the faster the clock $T(t)$, the more volatile the process $W(T(t))$.

3.7. ITO PROCESS AND STOCHASTIC DIFFERENTIAL EQUATIONS

Fact 22 An Ito process is a stochastic process of the form

$$X(t) = X(0) + \int_0^t f(X(u), u) du + \int_0^t \sigma(X(u), u) dW(u),$$

where $X(0)$ is a random variable, and $f(X(t), t), \sigma(X(t), t)$ are adapted processes satisfying some regularity conditions.

The Ito process $X(t)$ can also be written in differential form (as a shorthand notation)

$$\begin{cases} dX(t) = f(X(t), t) dt + \sigma(X(t), t) dW(t), \\ X(0) = X. \end{cases}$$

Meaning of $f(X(t), t)$ and $\sigma(X(t), t)$

- We observe that

$$\begin{aligned} \mathbb{E}_t(dX(t)) &= f(X(t), t) dt, \\ \mathbb{V}ar_t(dX(t)) &= \sigma^2(X(t), t) dt, \end{aligned}$$

hence, we can interpret

- $f(X(t), t) dt$ as the expected instantaneous change in X over the time period $(t, t + dt)$; $f(X(t), t)$ is called DRIFT COEFFICIENT.
- $\sigma^2(X(t), t) dt$ as the variance of the instantaneous changes; $\sigma(X(t), t)$ is called DIFFUSION COEFFICIENT.

3.8. ABM WITH DETERMINISTIC VOLATILITY

- An Arithmetic Brownian Motion is an Ito process defined as

$$X(t) = X(0) + \int_0^t \mu(s) ds + \int_0^t \sigma(s) dW(s),$$

for any deterministic functions $\mu(t), \sigma(t)$ such that $\int_0^T \sigma^2(u)du < \infty$ for any fixed $T < \infty$.

- In virtue of the properties of the stochastic integral, it follows that

$$X(t) \sim \mathcal{N} \left(X(0) + \int_0^t \mu(s) ds, \int_0^t \sigma^2(s) ds \right).$$

- The conditional distribution of $X(t)$ given $\mathcal{F}(s)$ is

$$X(t) | \mathcal{F}(s) \sim \mathcal{N} \left(X(s) + \int_s^t \mu(s) ds, \int_s^t \sigma^2(u) du \right).$$

- The expressions of the variance obtained above are obtained using the Ito isometry.
- The corresponding stochastic differential equation is

$$dX(t) = \mu(t) dt + \sigma(t) dW(t).$$

Example 23 Let us assume that $\mu(t) = \mu \in \mathbb{R}$ and $\sigma(t) = \sigma > 0$, then the ABM is given by

$$X(t) = X(0) + \mu t + \sigma W(t),$$

or, equivalently

$$dX(t) = \mu dt + \sigma dW(t).$$

From the above, the facts listed in Table (3.2) hold.

ARITHMETIC BROWNIAN MOTION: FACTS

The SDE

$$dX(t) = \mu dt + \sigma dW(t), X(0) = x_0.$$

The solution

$$X(t) = X(0) + \mu t + \sigma W(t).$$

The distribution of $X(t)$

$$X(t) \sim \mathcal{N}(X(0) + \mu t, \sigma^2 t).$$

The moments of $X(t)$

$$\mathbb{E}_0(X(t) - X(0) - \mu t)^p = (\sigma\sqrt{t})^p (p-1)!! \text{ if } p \text{ is even.}$$

In particular:

if $n = 1$

$$\mathbb{E}_0(X(t)) = X(0) + \mu t.$$

if $n = 2$

$$\mathbb{E}_0(X^2(t)) = (X(0) + \mu t)^2 + \sigma^2 t.$$

The auto-covariance of $X(t)$

$$c_X(t, s) = \sigma^2 \min(t, s).$$

Table 3.2: Properties of the arithmetic Brownian motion process.

3.8.1. MATLAB: SIMULATING THE ARITHMETIC BROWNIAN MOTION

Sample paths of the ABM are presented in Figure (3.2) (see the accompanying code below). Figure (3.6) shows the distribution of the ABM at different time horizons.

Matlab Code

```
%%%%%%SIMULATING THE ARITHMETIC BROWNIAN MOTIONS%%%%%
%%%%%
clear all
%Assign the number of simulated paths (nsimul)
nsimul=10000;
%Time to maturity (expiry)
expiry=1;
%number of steps (nsteps)
nsteps=250;
%time step (dt) and observation times (timestep):
dt=expiry/nsteps;
timestep=[0:dt:expiry]';
%model parameters
mu=0.2; sigma=0.3;
%Simulate increments ABM dX:
dX=mu*dt+sigma*randn(nsteps,nsimul)*dt^0.5;
%Simulate ABM process: cumulate increments
X=zeros(1,nsimul); cumsum(dX);

%compute theoretical expected value
EX=timestep*mu;

%Plot simulated paths:
h=figure('Color',[1 1 1]);
plot(timestep, X)
hold on
plot(timestep, EX, 'r.')
```

```

title('Simulated Paths of the Arithmetic Brownian Process ABM(0.2, 0.3)')
xlabel('Time')
print(h,'-dpng','LecBMSimulatedABM')

%Sstimate mean of X(t) from the simulations
%and compare to the true one
SimulatedMean=mean(X)');
h=figure('Color',[1 1 1]);
plot(timestep, [SimulatedMean, EX])
xlabel('Time')
legend('Simulated expected value','Theoretical expected value')
title('Comparing the theoretical mean of the ABM with the one estimated via MC')
print(h,'-dpng','LecBM_mean_ABM')

%Now we repeat the calculation for the variance.
%compute theoretical variance
VX=timestep*sigma^2;
%estimate variance from the simulations
SimulatedVar=var(X)';
h=figure('Color',[1 1 1]);
plot(timestep, [SimulatedVar, VX])
xlabel('Time')
legend('Simulated variance','Theoretical variance','location','best')
title('Comparing the theoretical variance of the ABM with the one estimated via MC')
print(h,'-dpng','LecBM_var_ABM')

```

```

%We compare the simulated distribution of the process
% with the Gaussian density with the same mean and variance.
%To plot the simulated distribution we use the
%new Matlab command histogram

%Plot distribution of the ABM at different dates:
%Range of the plot and number of points
bmmin=min(min(X)); bmmax=max(max(X)); nbins=100;

```

```
%Fix the times at which to plot the densities
horizon=[50, 125, 250]+1;

h=figure('Color',[1 1 1]);
for j=1:length(horizon)
ndays=horizon(j);
subplot(3,1,j);
histogram(X(ndays,:),nbins,'Normalization','pdf'); xlabel('ABM')
hold on
plot(sort(X(ndays,:)),normpdf(sort(X(ndays,:)),EX(ndays),VX(ndays)^0.5))
legend('Simulated','Gaussian','location','northwest')
xlim([bmmin, bmmax]);
title(strcat('Density in', [ ' ' num2str(horizon(j)/horizon(end),2)], ' years'))
end
print(h, '-dpng', 'LecBM_dens_ABM')
```

We also plot in a 3D diagram the time evolution of the density of the ABM, see figure 3.7.

```
%Assign number of bins and the edges
abmmmin=mu*timestep(end)-6*sigma*timestep(end)^0.5;
abmmax=mu*timestep(end)+6*sigma*timestep(end)^0.5;
nbins=100;
edges=linspace(abmmmin,abmmax,nbins);

%Build theoretical pdf at each time step
for j=2:nsteps+1
pdf_values(j,:)=pdf('normal',edges, ...
(mu)*timestep(j),sigma*timestep(j)^0.5);
end
steps=20;
%Finally, the plot
h=figure('Color',[1 1 1]);
surf(edges, timestep(steps:steps:end),pdf_values(steps:steps:end,:))
ylabel('Time'); xlabel('ABM Values');
```

```

zlabel('Density');
title('Time evolution of the ABM density')
ylim([timestep(steps) timestep(end)])

```

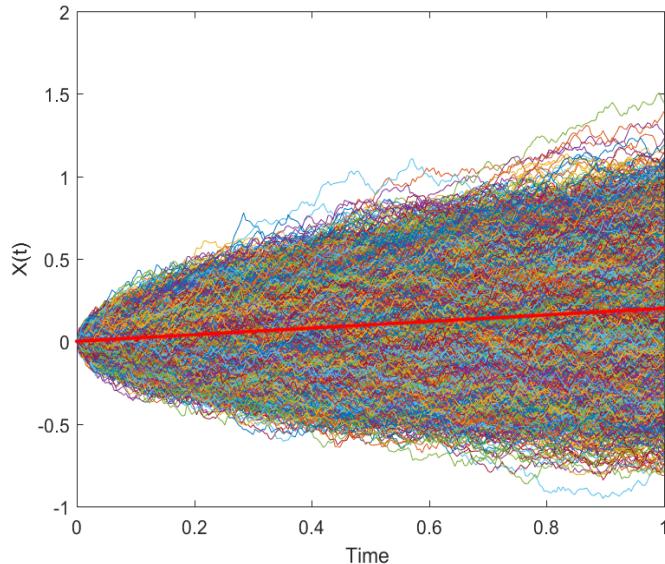


Figure 3.2: Simulated sample paths of the ABM(0.2,0.3).

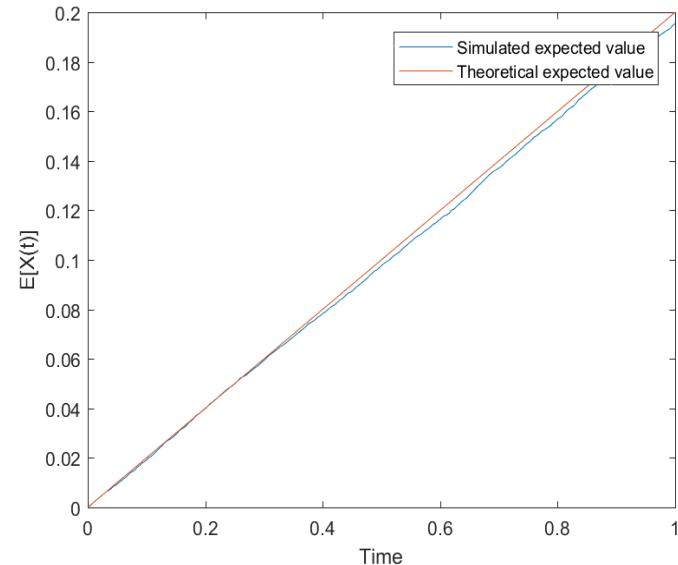


Figure 3.3: Simulated and theoretical expected value of $X(t)$ of the ABM(0.2,0.3) at different horizons.

3.9. SOLVING STOCHASTIC INTEGRALS AND/OR STOCHASTIC DIFFERENTIAL EQUATIONS

- In the previous section we have shown the equivalence between stochastic integrals and stochastic differential equations.
- Consequently, finding the solution to a stochastic integral is equivalent to finding a solution to the corresponding SDE.

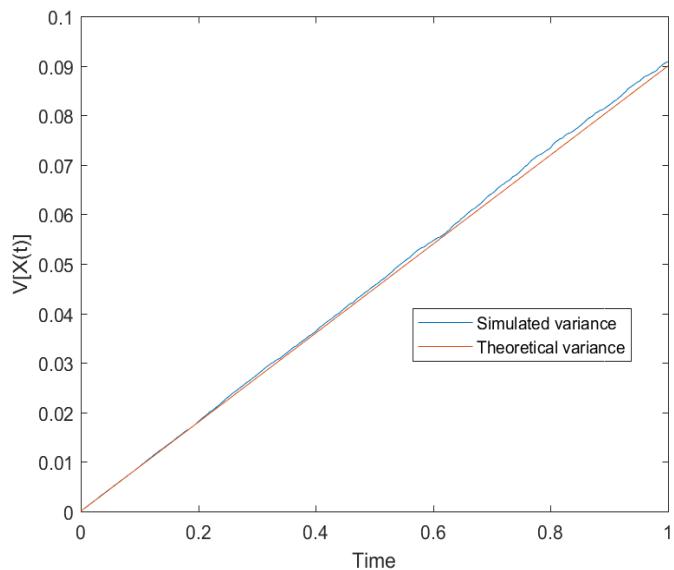


Figure 3.4: Simulated and theoretical variance of the ABM(0.2,0.3) at different horizons.

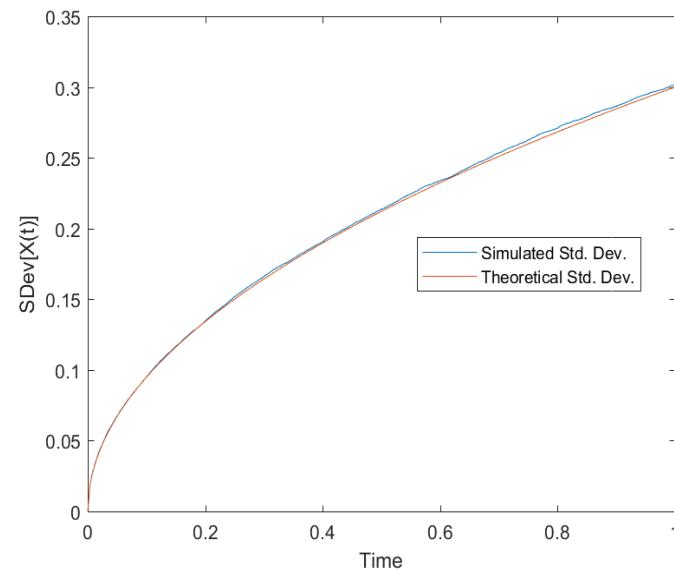


Figure 3.5: Simulated and theoretical standard deviation of the ABM(0.2,0.3) at different horizons.

- As illustrated in section 3.4, finding this solution can be quite complicated.
- However, there is a very useful tool for this task: Ito's lemma. This will be introduced in the next chapter.
- In the remaining of this chapter, we want to summarize some useful facts about SDEs and their solution.

3.9.1. DETERMINISTIC VS. STOCHASTIC DIFFERENTIAL EQUATIONS

Deterministic
Differential form

$$dx(t) = f(x(t), t) dt.$$

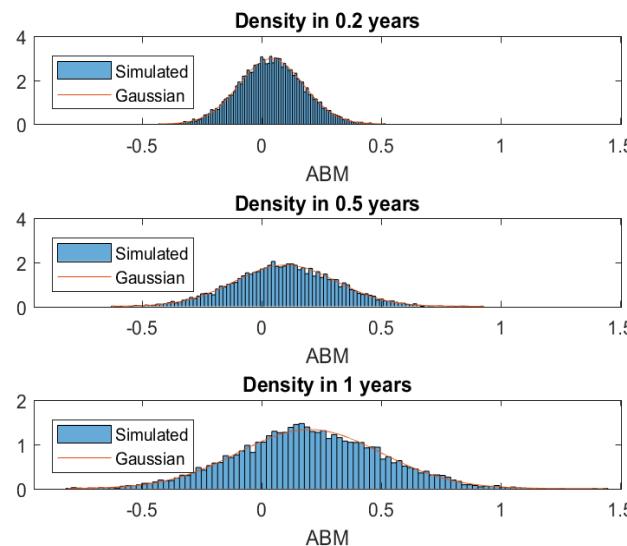


Figure 3.6: Simulated distribution of the ABM(0.2,0.3) at different horizons.

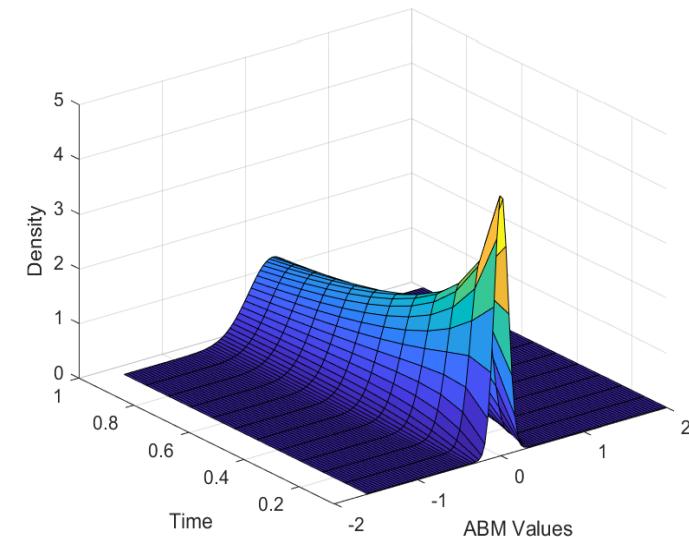


Figure 3.7: Time evolution of the exact distribution of the ABM(0.2,0.3).

Integral form

$$x(t) = x(0) + \underbrace{\int_0^t f(x(u), u) du}_{\text{Riemann Integral}}$$

**Stochastic
Differential form**

$$dX(t) = f(X(t), t) dt + \sigma(X(t), t) dW(t) \quad (3.1)$$

Integral form

$$X(t) = X(0) + \underbrace{\int_0^t f(x(u), u) du}_{\text{Riemann Integral}} + \underbrace{\int_0^t \sigma(X(u), u) dW(u)}_{\text{Stochastic Integral}}. \quad (3.2)$$

3.9.2. EXAMPLES OF STOCHASTIC DIFFERENTIAL EQUATIONS

Few examples of SDEs are given in Table 3.3, for which we also provide the deterministic version.

Deterministic	Stochastic	Name
$dx = \mu dt$	$dX(t) = \mu dt + \sigma dW(t)$	Arithmetic Brownian Motion (ABM)
$dx = \mu x dt$	$dX(t) = \mu X(t) dt + \sigma X(t) dW(t)$	Geometric Brownian Motion (GBM)
	$dX(t) = \mu X(t) dt + \sigma X^{\beta+1}(t) dW(t)$	Constant Elasticity of Variance (CEV)
	$dX(t) = \alpha(\mu - X(t)) dt + \sigma dW(t)$	Ornstein-Uhlenbeck (Vasicek)
$dx = \alpha(\mu - x) dt$	$dX(t) = \alpha(\mu - X(t)) dt + \sigma \sqrt{X(t)} dW(t)$	Square root (Cox-Ingersoll-Ross, CIR)
	$dX(t) = \alpha(\mu - X(t)) dt + \sigma X(t) dW(t)$	Lognormal with mean-reversion
	$dX(t) = \mu X(t) dt + \sqrt{v(t)} X(t) dW(t)$	
	$dv(t) = v(t) dt + \kappa v(t) dW(t)$	Stochastic Volatility (Hull-White model)
	$dv(t) = \alpha(\mu - v(t)) dt + \kappa \sqrt{v(t)} dW(t)$	Stochastic Volatility (Heston model)

Table 3.3: Examples of deterministic and stochastic differential equations.

3.9.3. EXISTENCE AND UNIQUENESS OF THE SOLUTION

Remark 24 (Definition of solution) $X(t)$ is called a strong solution of the SDE (3.1) if for all $t > 0$, $X(t)$ is a function $F(t, (W(s), s \leq t))$ of the given Brownian motion $W(t)$, integrals $\int_0^t f(x(u), u) du$ and $\int_0^t \sigma(X(u), u) dW(u)$ exist and the integral equation (3.2) is satisfied.

Fact 25 (Existence and uniqueness) Assume the initial condition $X(0)$ has finite second moment, $\mathbb{E}(X^2(0)) < \infty$, and is independent of $W(t)$, $t \geq 0$ and that, for all $t \in [0, T]$ and $x, y \in \mathbb{R}$, the coefficient functions $f(x, t)$ and $\sigma(x, t)$ satisfy the following conditions.

1. They are continuous.
2. They satisfy a Lipschitz condition with respect to the first variable.

$$|f(x, t) - f(y, t)| + |\sigma(x, t) - \sigma(y, t)| \leq K|x - y|.$$

Then the Ito stochastic differential equation (3.1) has a unique solution X on $[0, T]$.

In the following we will assume that the considered SDE will also admit a unique solution.

CHAPTER 4. INTRODUCING ITO'S FORMULA

Contents

4.1	A fact from ordinary calculus.	48
4.2	Itô's formula when $Y = g(X)$, $g(X) \in C^2$	49
4.3	Guiding principle	50
4.4	Itô's formula when $Y(t) = g(t, X)$, $g(t, X) \in C^{1,2}$	50
4.5	The Lamperti transformation	52
4.5.1	Example	53
4.6	The Multivariate Itô's lemma when $Z = g(t, X, Y)$	54
4.6.1	Examples	56

- The main tool in stochastic calculus is the Ito's formula, a stochastic version of Taylor formula.
- Given the Ito process $X(t)$

$$dX = \mu(X, t) dt + \sigma(X, t) dW,$$

let $g \in C^2$ (i.e. g is a function whose second order partial derivatives are continuous), and define

$$Y = g(t, X).$$

- We can think of g as the price of a derivative written on $X(t)$, and we aim to find the dynamics of the derivative price.
- What is the stochastic differential of the process $Y(t)$?
- In order to answer to this question, let us recall the basic calculus rules with the BM.

\times	dt	$dW(t)$
dt	0	0
$dW(t)$	0	<u>dt</u>

- Therefore, we have

$$\begin{aligned} (dX)^2 &= (\mu dt + \sigma dW)^2 \\ &= \mu^2 (dt)^2 + 2\mu\sigma dt dW + \sigma^2 (dW)^2 \\ &= o(dt) + o(dt) + \sigma^2 dt \\ &= \sigma^2 dt. \end{aligned}$$

4.1. A FACT FROM ORDINARY CALCULUS.

- Let us consider the function $x = f(t)$ with f a continuously differentiable function. We can write

$$dx(t) = f'(t) dt.$$

- Let us now introduce $y = g(x)$ with f also being continuously differentiable functions. Then

$$y'(t) = \frac{dy}{dt} = \frac{dg}{dx} \times \frac{dx}{dt} = g'(x(t)) f'(t).$$

- We are looking for the dynamics of dy . We have

$$dy(t) = y'(t) dt = g'(x(t)) f'(t) dt = g'(x(t)) dx(t),$$

and over a time period $[0, T]$, we have

$$y(T) = y(0) + \int_0^T dy(t) = y(0) + \int_0^T g'(x(t)) dx(t).$$

- How does this result translate when we deal with stochastic differential equations? We derive the Ito's formula.

4.2. ITÔ'S FORMULA WHEN $Y = g(X)$, $g(X) \in C^2$

- We have $dX(t) = \mu(X, t) dt + \sigma(X, t) dW(t)$, and $y = g(X)$.
 - What is the SDE of Y?**
- Considering the second-order Taylor series expansions (and using the fact seen in expression (4.1), i.e. $(dX)^2 = \sigma^2 dt$)

$$\begin{aligned} dY &= g'(X) dX + \frac{1}{2} g''(X) (dX)^2 \\ &= g'(X) dX + \frac{1}{2} g''(X) (\sigma^2(X, t) dt) \\ &= g'(X) (\mu(X, t) dt + \sigma(X, t) dW(t)) + \frac{1}{2} g''(X) (\sigma^2(X, t) dt) \\ &= \left(g'(X) \mu(X, t) + \frac{1}{2} \sigma^2(X, t) g''(X) \right) dt + \sigma(X, t) g'(X) dW(t). \end{aligned}$$

- The additional term in the drift is due to the rule $dW^2 = dt$.

Therefore, we can say that

Fact 26 *The SDE for $Y = g(X)$ when $dX(t) = \mu(X, t) dt + \sigma(X, t) dW(t)$ is given by:*

$$dY(t) = \left(g'(X) \mu(X, t) + \frac{1}{2} \sigma^2(X, t) g''(X) \right) dt + \sigma(X, t) g'(X) dW(t).$$

Remark 27 *In contrast to the deterministic case, when $X(t)$ is a Brownian motion or an Ito process, the contribution of the second order term in the Taylor expansion is not negligible, since Brownian motion has finite quadratic variation. This fact is the reason for the deviation from the classical chain rule.*

4.3. GUIDING PRINCIPLE

- Write out the Taylor series expansion of g with respect to all its argument.
 1. Take this Taylor series expansion out to first order for every argument that has zero quadratic variation.
 2. Take the expansion out to second order for every argument that has non-zero quadratic variation.
 3. As the variation of order 3 of each argument is zero and the covariation of $W(t)$ and t is zero, the other terms can be neglected.

4.4. ITÔ'S FORMULA WHEN $Y(t) = g(t, X)$, $g(t, X) \in C^{1,2}$

- We have $dX(t) = \mu(X, t) dt + \sigma(X, t) dW(t)$, and

$$Y(t) = g(t, X).$$

- Considering the second-order Taylor series expansions, we have

$$\begin{aligned} dY(t) &= \frac{\partial g(t, X)}{\partial t} dt + \frac{\partial g(t, X)}{\partial x} dX + \frac{1}{2} \frac{\partial^2 g(t, X)}{\partial x^2} (dX)^2 \\ &= \frac{\partial g(t, X)}{\partial t} dt + \frac{\partial g(t, X)}{\partial x} (\mu(X, t) dt + \sigma(X, t) dW) + \frac{\sigma^2(X, t)}{2} \frac{\partial^2 g(t, X)}{\partial x^2} dt. \end{aligned}$$

Therefore.

Fact 28 (Generalized Itô's lemma for $Y(t) = g(t, X)$) The SDE for $Y(t) = g(t, X) \in C^{1,2}$ when $dX(t) = \mu(X, t) dt + \sigma(X, t) dW(t)$ is

$$\begin{aligned} dY(t) &= \frac{\partial g(t, X)}{\partial t} dt + \frac{\partial g(t, X)}{\partial X} dX(t) + \frac{1}{2} \frac{\partial^2 g(t, X)}{\partial X^2} (dX(t))^2 \\ &= \left(\frac{\partial g(t, X)}{\partial t} + \mu(X, t) \frac{\partial g(t, X)}{\partial X} + \frac{1}{2} \sigma^2(X, t) \frac{\partial^2 g(t, X)}{\partial X^2} \right) dt + \sigma(X, t) \frac{\partial g(t, X)}{\partial X} dW(t). \end{aligned}$$

Further, we deduce that.

$$\mathbb{E}_t(dY(t)) = \text{Drift}(dY(t))dt = \left(\frac{\partial g(t, X)}{\partial t} + \mu(X, t) \frac{\partial g(t, X)}{\partial X} + \frac{1}{2} \sigma^2(X, t) \frac{\partial^2 g(t, X)}{\partial X^2} \right) dt.$$

$$\text{Volatility}_t(dY(t)) = \text{Diffusion}(dY(t))\sqrt{dt} = \sigma(X, t) \frac{\partial g(t, X)}{\partial X} \sqrt{dt}.$$

- Ito's Lemma gives us a tool to derive stochastic differential equations of stochastic processes obtained as function of another Ito process.

Example 29 Let us consider the process $Y(t) = W^2(t)$, where $W(t)$ is a BM. By Ito's lemma, the SDE of $Y(t) = g(W)$ is

$$dY(t) = 2W(t)dW(t) + dt.$$

- At the same time, Ito's Lemma can provide a way to solve a given SDE by suitably choosing an auxiliary process with a simpler SDE.

Example 30 Let us consider the process $X(t) = \int_0^t W(s)dW(s)$, or equivalently $dX(t) = W(t)dW(t)$. From the SDE of the process $Y(t) = W^2(t)$ obtained above, it follows that

$$dX(t) = \frac{1}{2} (dY(t) - dt).$$

By integrating both sides, it follows that $X(t) = W^2(t)/2 - t/2$, which verifies the result given in Example 3.4. Indeed

$$\begin{aligned} X(t) &= X(0) + \int_0^t dX(s) \\ &= X(0) + \frac{1}{2} \left(\int_0^t dY(s) - \int_0^t ds \right) \\ &= X(0) + \frac{1}{2} (Y(t) - Y(0) - t) \\ &= 0 + \frac{1}{2} (W^2(t) - t). \end{aligned}$$

4.5. THE LAMPERTI TRANSFORMATION

An application of the Ito's formula is the following.

- Consider a one-dimensional Ito's diffusion $(X(t))_{t \geq 0}$ with time-homogeneous diffusion coefficient, i.e.

$$dX(t) = \mu(X(t), t)dt + \sigma(X(t))dW(t), X(0) = x_0$$

- The Lamperti transform is defined as follows

$$Y(t) = g(X(t)) = \int_a^{X(t)} \frac{du}{\sigma(u)}$$

where a is a suitably chosen coefficient.

- Note that

$$\frac{\partial Y(t)}{\partial x} = \frac{1}{\sigma(x)}, \frac{\partial^2 Y(t)}{\partial x^2(t)} = -\frac{1}{\sigma^2(x)} \frac{\partial \sigma(x)}{\partial x}.$$

- An application of Ito's lemma gives us

$$dY(t) = \left(\frac{\mu(X(t), t)}{\sigma(X(t))} - \frac{1}{2} \frac{\partial \sigma(X(t))}{\partial X(t)} \right)_{X(t)=g^{-1}(Y(t))} dt + dW(t), Y(0) = g(x_0).$$

- This transformation allows us to obtain a process with unit diffusion.
- This is a convenient property in the context of parameter estimation and in the construction of binomial trees based approximations of continuous time diffusion, see for example Nelson and Ramaswamy (1990).

4.5.1. EXAMPLE

- Let us consider the process

$$dX(t) = k(\mu - X(t))dt + \sigma \sqrt{X(t)} dW(t), X(0) = x_0.$$

- In this case, the Lamperti transformation is

$$Y(t) = \int_0^{X(t)} \frac{du}{\sigma\sqrt{u}} = \frac{2}{\sigma} \sqrt{X(t)}.$$

- We have

$$\frac{\partial Y(t)}{\partial x} = \frac{1}{\sigma\sqrt{x}}, \frac{\partial^2 Y(t)}{\partial x^2} = -\frac{1}{\sigma^2\sqrt{x^3}x}.$$

- Therefore, we have a process with constant diffusion coefficient

$$dY(t) = \left(\frac{\mu(X(t), t)}{\sigma(X(t))} - \frac{1}{2} \frac{\partial \sigma(X(t))}{\partial X(t)} \right) dt + dW(t), Y(0) = g(x_0).$$

- The construction of an approximating binomial tree to this process is discussed in the reference given above.

4.6. THE MULTIVARIATE ITO'S LEMMA WHEN $Z = g(t, X, Y)$.

- Let us consider two Brownian motions $W_1(t)$ and $W_2(t)$.
- Their increments $dW_1(t)$ and $dW_2(t)$ satisfy

$$\mathbb{E}(dW_1(t) dW_2(t)) = \rho dt, \quad \rho \in [-1, 1].$$

- Here the coefficient ρ is interpreted as instantaneous correlation coefficient between the increments of the two Brownian motions.
- The multiplicative rules in the multivariate case become

\times	dt	$dW_1(t)$	$dW_2(t)$
dt	0	0	0
$dW_1(t)$	0	dt	ρdt
$dW_2(t)$	0	ρdt	dt

- We have

$$\begin{aligned} dX(t) &= \mu(X, t) dt + \sigma(X, t) dW_1(t), \\ dY(t) &= \mu(Y, t) dt + \sigma(Y, t) dW_2(t). \end{aligned}$$

We consider a function of time t , and of the two variables X and Y , $Z = g(t, X, Y)$ say.

What is the SDE for $Z = g(t, X, Y)$?

- Considering the multivariate second-order Taylor formula, we have

$$\begin{aligned} dZ &= \frac{\partial g(t, x, y)}{\partial t} dt + \frac{\partial g(t, X, Y)}{\partial X} dX + \frac{\partial g(t, X, Y)}{\partial Y} dY \\ &\quad + \frac{1}{2} \frac{\partial^2 g(t, X, Y)}{\partial X^2} (dX)^2 + \frac{1}{2} \frac{\partial^2 g(t, X, Y)}{\partial Y^2} (dY)^2 \\ &\quad + \frac{\partial^2 g(t, X, Y)}{\partial Y \partial X} dXdY, \end{aligned}$$

- Let us use the multiplicative rules of the above Table

$$(dX)^2 = \sigma^2(X, t)dt,$$

$$(dY)^2 = \sigma^2(Y, t)dt,$$

and

$$dXdY = \rho\sigma(X, t)\sigma(Y, t)dt.$$

Fact 31 (Multivariate Ito's Lemma)

$$dZ = \mu(Z, t)dt + \frac{\partial g(t, X, Y)}{\partial X}\sigma(X, t)dW_1(t) + \frac{\partial g(t, X, Y)}{\partial Y}\sigma(Y, t)dW_2(t)$$

where

$$\begin{aligned} \mu(Z, t) &= \frac{\partial g(t, x, y)}{\partial t} + \frac{\partial g(t, X, Y)}{\partial X}\mu(X, t) + \frac{\partial g(t, X, Y)}{\partial Y}\mu(Y, t) \\ &\quad + \frac{1}{2} \frac{\partial^2 g(t, X, Y)}{\partial X^2} \sigma^2(X, t) + \frac{1}{2} \frac{\partial^2 g(t, X, Y)}{\partial Y^2} \sigma^2(Y, t) \\ &\quad + \rho \frac{\partial^2 g(t, X, Y)}{\partial X \partial Y} \sigma(X, t)\sigma(Y, t). \end{aligned}$$

4.6.1. EXAMPLES

Let us consider two examples of the multivariate Ito's formula:

- the product $X Y$;
- the ratio Y/X .

Fact 32 (Product $Z = XY$)

$$\frac{dZ}{Z} = \frac{dX}{X} + \frac{dY}{Y} + \left(\frac{dX}{X} \right) \left(\frac{dY}{Y} \right).$$

Fact 33 (Ratio $Z = Y/X$)

$$\frac{dZ}{Z} = \frac{dY}{Y} - \frac{dX}{X} - \left(\frac{dX}{X} \right) \left(\frac{dY}{Y} \right) + \left(\frac{dX}{X} \right)^2.$$

These properties are important when X and Y are GBM processes. Indeed, they imply that the product and the ratio of (correlated) GBM processes is again a GBM process. This is shown in the following two examples.

Example 34 (Volatilities of Products of GBM) Assume $\mu_i \in \mathbb{R}, \sigma_i > 0$, for $i = X, Y$. Define the processes X, Y as

$$\begin{aligned} \frac{dX}{X} &= \mu_X dt + \sigma_X dW_X, \\ \frac{dY}{Y} &= \mu_Y dt + \sigma_Y dW_Y. \end{aligned}$$

Then, the SDE of the product process $Z = XY$ is given by

$$\begin{aligned}\frac{dZ}{Z} &= (\mu_X + \mu_Y + \rho\sigma_X\sigma_Y) dt + \sigma_X dW_X + \sigma_Y dW_Y. \\ \mathbb{E} \left(\frac{dZ}{Z} \right) &= (\mu_X + \mu_Y + \rho\sigma_X\sigma_Y) dt, \\ \mathbb{V}ar \left(\frac{dZ}{Z} \right) &= (\sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y) dt, \\ \mathbb{V}ol \left(\frac{dZ}{Z} \right) &= \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y} \sqrt{dt}.\end{aligned}$$

Example 35 (Volatilities of Ratios of GBM) Consider the processes X, Y given in the previous example. Then, the SDE for the ratio process $Z = Y/X$ is given by

$$\begin{aligned}\frac{dZ}{Z} &= (\mu_Y - \mu_X - \rho\sigma_X\sigma_Y + \sigma_X^2) dt + \sigma_Y dW_Y - \sigma_X dW_X. \\ \mathbb{E} \left(\frac{dZ}{Z} \right) &= (\mu_Y - \mu_X - \rho\sigma_X\sigma_Y + \sigma_X^2) dt, \\ \mathbb{V}ar \left(\frac{dZ}{Z} \right) &= (\sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y) dt, \\ \mathbb{V}ol \left(\frac{dZ}{Z} \right) &= \sqrt{\sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y} \sqrt{dt}.\end{aligned}$$

CHAPTER 5. IMPORTANT SDEs

Contents

5.1	The Geometric Brownian Motion $GBM(\mu, \sigma)$	60
5.1.1	Solving the ODE $dX(t) = \mu X(t)dt$	61
5.1.2	Solving the SDE $dX = \mu X dt + \sigma X dW$	61
5.1.3	Matlab Implementation: Simulating GBM	64
5.1.4	Remark. GBM with deterministic drift and volatility	68
5.2	The Vasicek Mean-Reverting process	75
5.2.1	A Note: The Ordinary Differential Equation $dx(t) = \alpha (\mu - x(t)) dt$.	76
5.2.2	Solving the SDE $dX(t) = \alpha (\mu - X(t)) dt + \sigma dW(t)$	79
5.2.3	The (auto)-covariance function	80
5.2.4	Matlab: Simulation of the Vasicek model	81
5.2.5	Extension: MR with deterministic volatility	87
5.3	The Cox-Ingersoll-Ross (CIR) model	89
5.3.1	Solving the SDE $dX(t) = \alpha (\mu - X(t)) dt + \sigma \sqrt{X(t)} dW(t)$	89
5.3.2	Computing the expectation of the CIR model	90
5.3.3	Computing the variance of the CIR model	91
5.3.4	Again on computing the variance and the auto-covariance of the CIR model	92
5.3.5	The distribution of the short rate in the CIR model	94
5.3.6	The Extended-CIR model	97
5.3.7	Simulating the CIR model	100
5.4	The Constant Elasticity of Variance (CEV) model	106

5.5 CIR, CEV and the Bessel Process	116
5.6 The Brownian Bridge	118
5.6.1 A Note: The Ordinary Differential Equation $dx(t) = (b - x(t))/(T - t)dt$	118
5.6.2 Solving the SDE $dX_t = (b - X(t))/(T - t)dt + dW(t)$	119
5.6.3 Matlab Implementation: Simulating Brownian motions (part 2)	121
5.7 The Heston Stochastic Volatility model	126
5.7.1 The characteristic function of the log-price	128
5.7.2 Heston model: Monte Carlo Simulation	134
5.7.3 Recovering the density function in the Heston model via inversion of the characteristic function	138
5.7.4 Heston model and option pricing	144
5.7.5 Heston model and option pricing: Main findings	147
5.8 The Heston app	152

In this chapter we will review the most important SDEs in Finance such as the ones illustrated in Table 5.1

Model	SDE
Arithmetic Brownian motion (ABM)	$dX(t) = \mu dt + \sigma dW(t)$
Geometric Brownian motion (GBM)	$dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$
Vasicek (VAS)	$dX(t) = \alpha(\mu - X(t))dt + \sigma dW(t)$
Cox-Ingersoll-Ross (CIR)	$dX(t) = \alpha(\mu - X(t))dt + \sigma\sqrt{X(t)}dW(t)$
Brownian Bridge (BB)	$dX(t) = \frac{b-X(t)}{T-t}dt + dW(t)$
Stochastic Volatility Heston Model (HES)	$dX(t)(t) = \left(\mu - \frac{1}{2}v(t)\right)dt + \sqrt{v(t)}dW_s(t),$ $dv(t) = k(\theta - v(t))dt + \varepsilon\sqrt{v(t)}dW_v(t)$

Table 5.1: Relevant SDEs for Financial Modelling

5.1. THE GEOMETRIC BROWNIAN MOTION $GBM(\mu, \sigma)$

Fact 36 (Geometric Brownian Motion) *The SDE with drift*

$$\mu X, \mu \in \mathbb{R},$$

and diffusion coefficient

$$\sigma X, \sigma \in \mathbb{R}^+,$$

(i.e. $\mu(X, t) = \mu X$, and $\sigma(X, t) = \sigma X$), given by

$$dX(t) = \mu X(t)dt + \sigma X dW(t),$$

is said to be the Geometric Brownian Motion with coefficients μ and σ and is denoted by $GBM(\mu, \sigma)$.

- The GBM process has drift and diffusion which are a linear function of the state variable, as opposed to the case of the ABM for which they are constant.
- In particular, this guarantees that the process remains always positive if it starts from a positive value (a good news if you need to model market prices!).
- Our aim is to use Ito's lemma to solve the above SDE. This is done in section 5.1.1.

5.1.1. SOLVING THE ODE $dX(t) = \mu X(t)dt$

- We would like to solve the SDE $dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$.
- Let us see what happens with the **ordinary differential equation**, i.e. the deterministic version

$$dx(t) = \mu x(t)dt.$$

- We let $y(t) = \ln x(t)$.
- Then, $dy(t) = \frac{1}{x(t)}dx(t)$.
- Therefore $dy(t) = \mu dt$, i.e. $y(t) = y(0) + \mu t$.
- It follows that $\ln x(t) = \ln x(0) + \mu t$,

Fact 37 *The solution of the ODE $dx(t) = \mu x(t)dt$ is*

$$dx(t) = \mu x(t)dt \text{ i.e. } x(t) = x(0) e^{\mu t}$$

5.1.2. SOLVING THE SDE $dX = \mu Xdt + \sigma XdW$

- We want to solve the SDE $dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$.
- By analogy with the ODE $dX(t) = \mu X(t)dt$, let us look for the SDE of $Y(t) = g(X) = \ln X(t)$.

- Using Ito's lemma, we have

$$\begin{aligned}
& dY(t) \\
&= \left(\frac{\partial g(t, X)}{\partial t} + \mu X \frac{\partial g(t, X)}{\partial X} + \frac{\sigma^2 X^2}{2} \frac{\partial^2 g(t, X)}{\partial X^2} \right) dt + \sigma X \frac{\partial g(t, X)}{\partial x} dW(t). \\
&= \left(0 + \mu X \frac{1}{X} - \frac{1}{2} \sigma^2 X^2 \frac{1}{X^2} \right) dt + \sigma X \frac{1}{X} dW(t) \\
&= \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dW(t),
\end{aligned}$$

- This means that $Y(t)$ follows an ABM($\mu - \frac{1}{2}\sigma^2, \sigma$) process, with solution

$$Y(t) = Y(0) + \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W(t)$$

i.e.

$$\ln X(t) = \ln X(0) + \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W(t),$$

which implies

$$\ln X(t) \sim \mathcal{N} \left(\ln X(0) + \left(\mu - \frac{1}{2} \sigma^2 \right) t, \sigma^2 t \right),$$

Fact 38 *The SDE $dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$ has solution*

$$X(t) = X(0) e^{(\mu - \frac{1}{2} \sigma^2)t + \sigma W(t)}.$$

It follows that

$$X(t) \sim \mathcal{LN} \left(\ln X(0) + \left(\mu - \frac{1}{2} \sigma^2 \right) t, \sigma^2 t \right).$$

In particular, we have

$$\mathbb{E}_0(X(t)) = X(0) e^{\mu t}.$$

Remark 39 Notice that the SDE

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$$

is equivalent (e.g. they have the same solution) to the SDE

$$d \ln X(t) = \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dW(t).$$

In general, the first SDE is useful to model prices, whilst the second one is used to model log-returns.

Fact 40 (The exponential martingale) Consider the case of a GBM $X(t)$ with $\mu = 0$. Then $dX = \sigma X dW(t)$ and

$$X(t) = X(0)e^{\sigma W(t) - \frac{\sigma^2}{2}t}.$$

The process is then a martingale.

Indeed: the SDE shows that, due to the absence of the drift, the process is made of the stochastic integral only, which we know to be a martingale. Alternatively, let $s < t$ and consider

$$\begin{aligned} \mathbb{E}_s(X(t)) &= X(0)e^{\sigma W(s) - \frac{\sigma^2}{2}t} \mathbb{E}_s\left(e^{\sigma W(t) - \sigma W(s)}\right) \\ &= X(0)e^{\sigma W(s) - \frac{\sigma^2}{2}t} \mathbb{E}\left(e^{\sigma W(t) - \sigma W(s)}\right) \\ &= X(0)e^{\sigma W(s) - \frac{\sigma^2}{2}t} e^{\frac{\sigma^2}{2}(t-s)} \\ &= X(s) \end{aligned}$$

where the second equality follows from the fact that the Brownian motion has independent increments, and the third equality follows from the fact that the increments of the Brownian motion are stationary with Gaussian distribution.

Simulated trajectories of the GBM are shown in Figure (5.1) (see accompanying code); the resulting distribution originated at different time horizons are presented in Figure (5.2). Properties of the geometric Brownian motion process are given in Table (5.2).

GEOMETRIC BROWNIAN MOTION: FACTS

The SDE

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t), X(0) = x_0.$$

The solution

$$X(t) = X(0) e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}.$$

The distribution of $X(t)$

$$X(t) \sim \mathcal{LN} \left(\ln X(0) + \left(\mu - \frac{1}{2}\sigma^2 \right) t, \sigma^2 t \right).$$

The moments of $X(t)$

$$\mathbb{E}_0(X(t)^n) = X(0)^n e^{n(\mu - \frac{1}{2}\sigma^2)t + \frac{n^2\sigma^2}{2}t}.$$

In particular, if $n = 1$

$$\mathbb{E}_0(X(t)) = X(0) e^{\mu t}.$$

Table 5.2: Properties of the geometric Brownian motion process.**5.1.3. MATLAB IMPLEMENTATION: SIMULATING GBM**

We provide here a code to simulate the GBM process. 100 simulated paths are given in figure 5.1. [Matlab Code](#)

```
%%%%%%%%%%%%%
%%%SIMULATING THE GEOMETRIC BROWNIAN MOTION %%%
%%%%
%Assign the number of simulated paths (nsimul)
nsimul=100;
%Time to maturity (expiry)
expiry=1;
%number of steps (nsteps)
nsteps=250;
%time step (dt) and observation times (timestep):
dt=expiry/nsteps;
timestep=[0:dt:expiry]';
```

```
%model parameters
mu=0.2; sigma=0.3;
%Simulate increments ABM dX:
dY=(mu-sigma^2/2)*dt+sigma*randn(nsteps,nsimul)*dt^0.5;
%Simulate ABM process: cumulate increments
Y=[zeros(1,nsimul); cumsum(dY)];
%Simulate the GBM: exponentiate the levels of the ABM
X0=100; %assign initial level
X=X0*exp(Y);
%Plot the simulated paths and we
%overimpose to it the expected value of the process

%Compute theoretical expected value
EX=X0*exp(timestep*mu);
%Plot simulated paths:
h=figure('Color',[1 1 1]);
plot(timestep, X)
hold on
plot(timestep, EX,'r.')
title('Simulated paths of the geometric Brownian Process GBM(0.2, 0.3)')
xlabel('Time'); ylabel('GBM(t)');
print(h,'-dpng','SimulatedGBM')
```

In order to assess the quality of the simulation we compare the theoretical mean and variance, i.e. $m(t) = X(0)\exp(\mu t)$ and $s^2(t) = X^2(0)\exp(2(\mu - \sigma^2/2)t + 2\sigma^2t) - m^2(t)$, with the ones estimated via simulation. We do this at first for the mean. We leave to the reader to produce the plot.

Matlab Code

```
%%%%%%%%%%%%%
%%%Plot Theoretical and Simulated values of E(X) and V(X)
%%%%%%%%%%%%%
%estimate mean from the simulations
SimulatedMean=mean(X)';
h=figure('Color',[1 1 1]);
```

```

plot(timestep, [SimulatedMean, EX])
xlabel('Time'); ylabel('E(X)')
legend('Simulated expected value','Theoretical expected value','location','best')
title('Comparing the theoretical mean of the GBM with the one estimated via MC')
print(h,'-dpng','LecBM_mean_GBM')

%%Now we repeat the calculation for the variance.
%compute theoretical variance
VX=X0^2*exp(timestep*2*(mu-sigma^2/2)+4*sigma^2/2*timestep)-EX.^2;
%estimate variance from the simulations
SimulatedVar=var(X)';
h=figure('Color',[1 1 1]);
plot(timestep, [SimulatedVar VX])
xlabel('Time');ylabel('V(X)');
legend('Simulated variance','Theoretical variance','location','best')
title('Comparing the theoretical variance of the GBM with the one estimated via MC')
print(h,'-dpng','LecBM_var_GBM')

```

We then can compare the theoretical and simulated distributions of the GBM at different horizons, see figures 5.5 and 5.6.

Matlab Code

```

%%%%%%%%%%%%%
%%%Plot GBM densities at different time horizons%%%%%
%%%%%
%Fix the range and the number of points
gbmmin=0; gbmmax=max(max(X));nbins=100;

%Fix the times at which to plot the densities
horizon=[50, 125, 250]+1;

h=figure('Color',[1 1 1]);
for j=1:length(horizon)
    ndays=horizon(j);
    subplot(3,1,j);
    %theoretical distribution

```

```

histogram(X(ndays,:),nbins,'Normalization','pdf');
xlabel('GBM Values')
hold on
%theoretical distribution
plot(edges,pdf('lognormal',edges, ...
    log(X0)+(mu-sigma^2/2)*timestep(ndays),sigma*timestep(ndays).^0.5))
xlim([gbmmin gbmmax]);
legend('Simulated','Lognormal','location','best')
title(strcat('Density in ', [ ' ' num2str(horizon(j)/horizon(end),2)], ' years'))
end

print(h,'-dpng','LecBMFigGBMDens')

```

We plot the time evolution of the density of the GBM. **Matlab Code**

```

%%%%%%%%%%%%%%%
%%%Time evolution of the GBM density %%%
%%%%%%%%%%%%%%%
%Assign number of bins and the edges
gbmnin=30; gbmmax=270; nbins=100;
edges=linspace(gbmmin,gbmmax,nbins);

%Build theoretical pdf at each time step
for j=2:nsteps+1
pdf_values(j,:)=pdf('lognormal',edges, ...
    log(X0)+(mu-sigma*sigma/2)*timestep(j),sigma*timestep(j)^0.5);
end
steps=20;

%Finally, the plot
h=figure('Color',[1 1 1]);
surf(edges, timestep(steps:steps:end),pdf_values(steps:steps:end,:))
ylabel('Time')
ylim([timestep(steps) timestep(end)])
xlim([gbmmin, gbmmax])

```

```

xlabel('GBM Values');
zlabel('Density');
title('Time evolution of the GBM density')
print(h, '-dpng', 'LecBMFigTimeGBMDens')

```

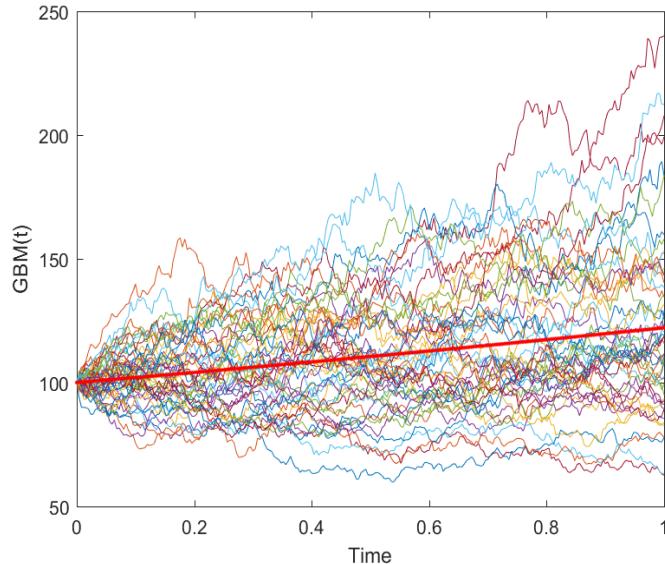


Figure 5.1: Sample Paths of the $\text{GBM}(0.2,0.3)$.

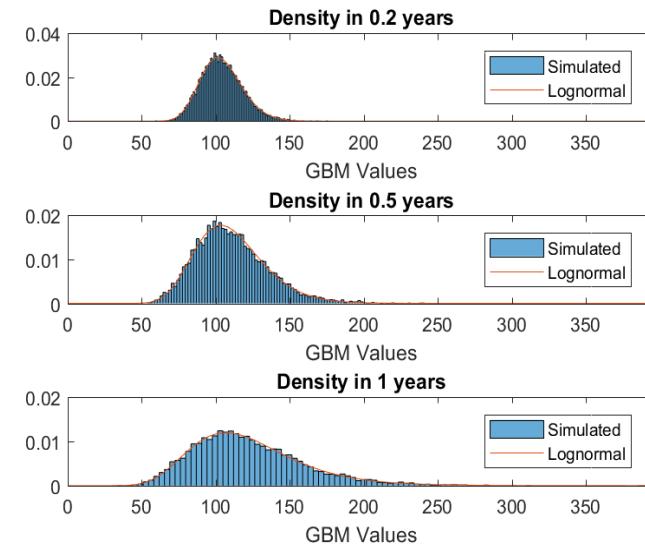


Figure 5.2: Density of the $\text{GBM}(0.2,0.3)$ at different horizons.

5.1.4. REMARK. GBM WITH DETERMINISTIC DRIFT AND VOLATILITY

- The SDE

$$dX(t) = \mu(t) X(t) dt + \sigma(t) X(t) dW(t)$$

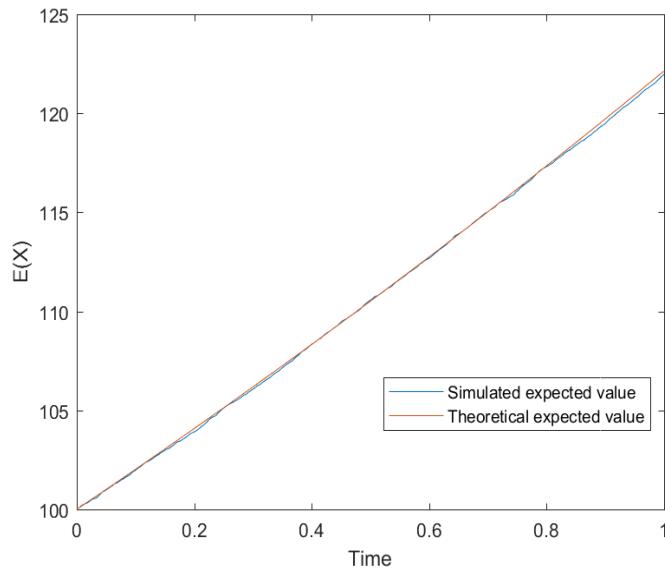


Figure 5.3: Theoretical and Monte Carlo estimated expected value of the $\text{GBM}(0.2,0.3)$ at different horizons.

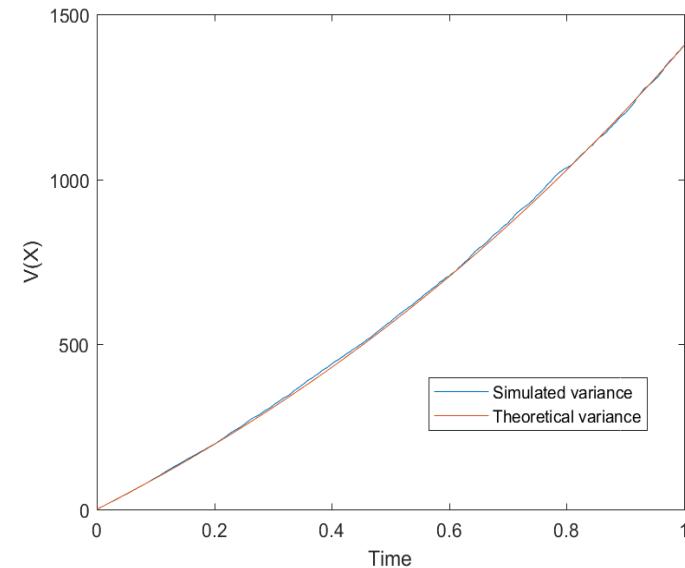


Figure 5.4: Theoretical and Monte Carlo estimated variance of the $\text{GBM}(0.2,0.3)$ at different horizons.

is equivalent to the SDE

$$d \ln X(t) = \left(\mu(t) - \frac{1}{2} \sigma^2(t) \right) dt + \sigma(t) dW(t).$$

- The solution of both equations is

$$X(t) = X(0) e^{\int_0^t (\mu(s) - \frac{1}{2} \sigma^2(s) ds) + \int_0^t \sigma(s) dW(s)},$$

or equivalently

$$\ln X(t) = \ln X(0) + \int_0^t \left(\mu(s) - \frac{1}{2} \sigma^2(s) \right) ds + \int_0^t \sigma(s) dW(s).$$

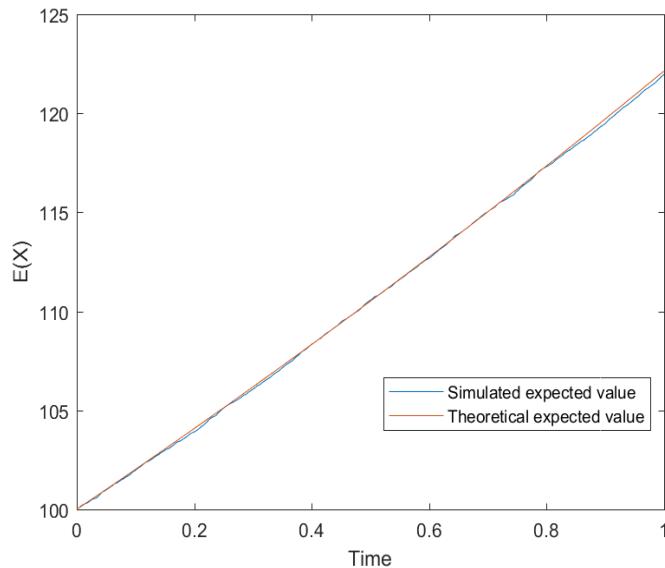


Figure 5.5: Theoretical and estimate mean of the $\text{GBM}(0.2, 0.3)$ at different horizons.

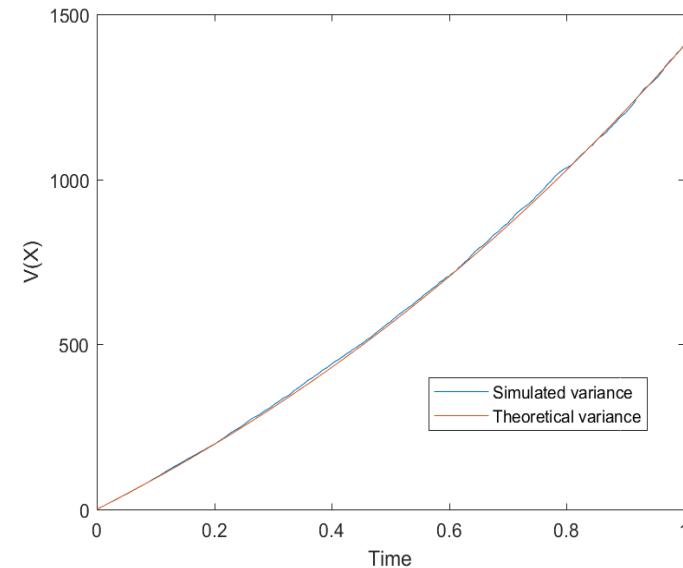


Figure 5.6: Theoretical and estimate variance of the $\text{GBM}(0.2, 0.3)$ at different horizons.

- Moreover

$$\begin{aligned} \ln X(t) &\sim \mathcal{N} \left(\ln X(0) + \int_0^t \left(\mu(s) - \frac{1}{2}\sigma^2(s) \right) ds, \int_0^t \sigma^2(s) ds \right), \\ X(t) &\sim \mathcal{LN} \left(\ln X(0) + \int_0^t \left(\mu(s) - \frac{1}{2}\sigma^2(s) \right) ds, \int_0^t \sigma^2(s) ds \right). \end{aligned}$$

- In particular, we observe that

$$\mathbb{E}(X(t)) = X(0)e^{\int_0^t \mu(s)ds}.$$

- This result can be useful if we interpret $X(t)$ to be the price of some commodity and we observe in the market a term structure of futures prices written on X , say $F_0(t)$.
- Assume we are interested into a GBM process for X taking the observed futures curve as expected value, i.e.

$$\mathbb{E}(X(t)) = F_0(t).$$

- This is possible if we impose

$$X(0)e^{\int_0^t \mu(s)ds} = F_0(t),$$

- Therefore, we can write

$$X(t) = F_0(t)e^{-\frac{1}{2} \int_0^t \sigma^2(s)ds + \int_0^t \sigma(s)dW(s)}$$

- If we are interested into the differential form, we observe that

$$\int_0^t \mu(s)ds = \ln \left(\frac{F_0(t)}{X(0)} \right),$$

must also hold; then by differentiating with respect to t :

$$\mu(t) = \frac{F'_0(t)}{F_0(t)}.$$

Example: Simulation of the oil price according to a GBM fitting the futures term structure

- To make concrete the discussion, let us consider Table (5.3) containing the term structure of futures prices on Light Sweet Crude Oil traded at CME on March, 12, 2013.
- Given the very short maturity of the April contract (only 8 days), we can set

$$X(0) = 91.94.$$

Contract Month	Settlement	Product Code	Open	Days
APR 2013	20-mar	CLJ13	91.94	8
MAY 13	22-apr	CLK13	92.46	41
JUN 13	21-may	CLM13	92.75	70
JLY 13	20-june	CLN13	92.79	100
AUG 13	22-july	CLQ13	93.2	132
SEP 13	20-aug	CLU13	93.13	161
OCT 13	20-sept	CLV13	92.76	192
NOV 2013	22-oct	CLX13	92.13	224
DEC 13	20-nov	CLZ13	91.99	253
JAN 14	19-dec	CLF14	91.91	282
FEB 2014	21-jan	CLG14	92	315
MAR 2014	20-feb	CLH14	91.25	345
APR 2014	20-mar	CLJ14	91.11	373

Table 5.3: Tabe contains Light Sweet Crude Oil (WTI) Futures prices quoted at CME as of March 12, 2013. First column refers to the contract month, the second to the product code; the third provides the exact expiry date of the contract; the fourth column gives the quoted futures price; last column contains the actual number of days to expiration.

- Then we can iteratively simulate X according to

$$X(t_i) = \frac{F_0(t_i)}{F_0(t_{i-1})} X(t_{i-1}) e^{-\frac{1}{2}\sigma^2(t_i-t_{i-1}) + \sigma(W(t_i) - W(t_{i-1}))}, i = 1, \dots,$$

where t_i refers to the expiry dates of different futures contracts, so that $t_0 = 8/365, t_1 = 41/365, t_2 = 70/365, \dots, t_{13} = 373/365$.

- As volatility parameter let us set

$$\sigma = 0.2284,$$

which corresponds to the annualized volatility of log-oil price increments in 2012.

- If we run 100,000 MC simulations, we can verify that the restriction imposed by the futures term structure is satisfied. In table (5.4), we compare for each future date t , the futures price $F_0(t)$ with the expectation $\mathbb{E}_0(X(t))$. The agreement is very good.

t (days)	$\mathbb{E}_0(X(t))$	$F_0(t)$
8	91.94	91.94
41	92.445	92.46
70	92.7386	92.75
100	92.7702	92.79
132	93.1911	93.2
161	93.1345	93.13
192	92.7703	92.76
224	92.1132	92.13
253	91.9431	91.99
282	91.8659	91.91
315	91.951	92
345	91.2335	91.25
373	91.0994	91.11

Table 5.4: The first column refers to the time to maturity (in days) of the different futures contracts. Second column refers to the expected value of the GBM process, i.e. $\mathbb{E}(X(t))$. The third column is the futures price quoted today for maturity t .

- A sample of simulated paths is illustrated in Figure (5.7).

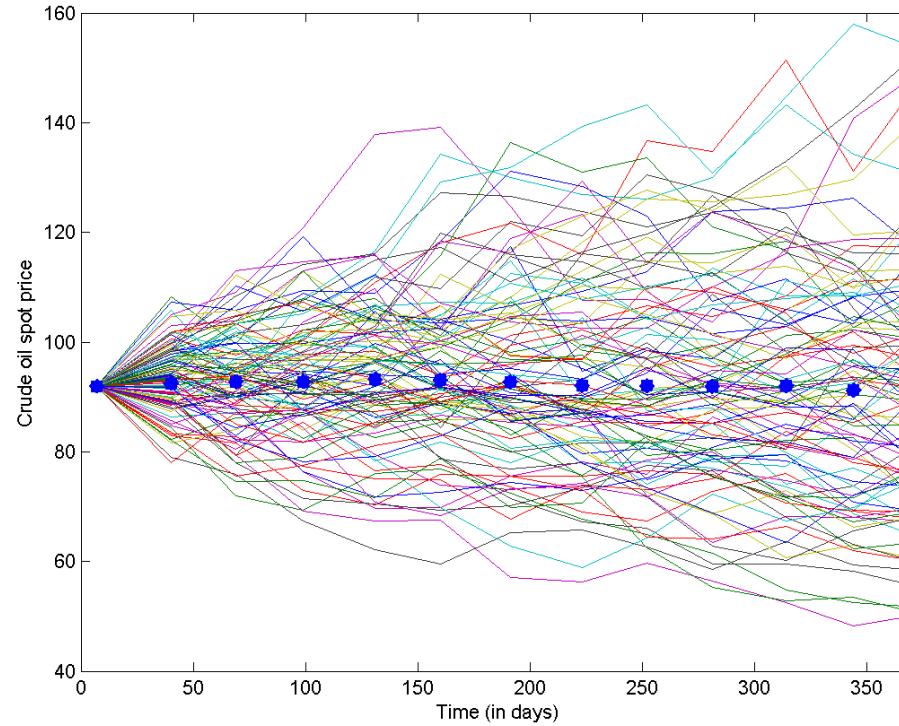


Figure 5.7: Simulated GBM paths fitting in average the term structure of futures prices (blue dotted points).

5.2. THE VASICEK MEAN-REVERTING PROCESS

- An empirical property of several economic variables such as interest rates, inflation rates and even commodity prices, is the tendency towards lower levels (higher levels), when they are too high (low).

- This property is called **mean-reversion** and can be modelled using a so-called **Mean-Reverting (MR) process**. The effect of mean-reversion is described in figure (5.8).
- This process is described by the following SDE

$$dX(t) = \alpha(\mu - X(t)) dt + \sigma dW(t), \mu \in \mathbb{R}; \alpha, \sigma > 0$$

- This process has been introduced in finance by Vasicek to model interest rates. However, its origin traces back to physics in which it is known as the (mean-reverting) Ornstein-Uhlenbeck (OU) process.
- We observe that

$$\mathbb{E}_t[dX(t)] = \alpha(\mu - X(t)) dt,$$

so that, assuming $\alpha > 0$, $\mathbb{E}_t[dX(t)] > 0$ when $X(t) < \mu$, i.e. we expect an increase (decrease) in the interest rate level when we are below (above) the level μ .

- Higher the value of α , faster the return towards the level μ . α is called speed (velocity) of mean reversion , whilst μ determines the long-run mean-level.
- The distribution of X at any future time is Gaussian, so it allows for negative values.
- An extension, that guarantees positive interest rates, has been proposed by Cox, Ingersoll and Ross (CIR model).

5.2.1. A NOTE: THE ORDINARY DIFFERENTIAL EQUATION $dx(t) = \alpha(\mu - x(t)) dt$.

- We would like to solve the SDE in the Vasicek model. Let us start considering the deterministic version.

$$dx(t) = \alpha(\mu - x(t)) dt.$$

- This is a first order ordinary differential equation. The procedure to solve is standard. We recall it here.

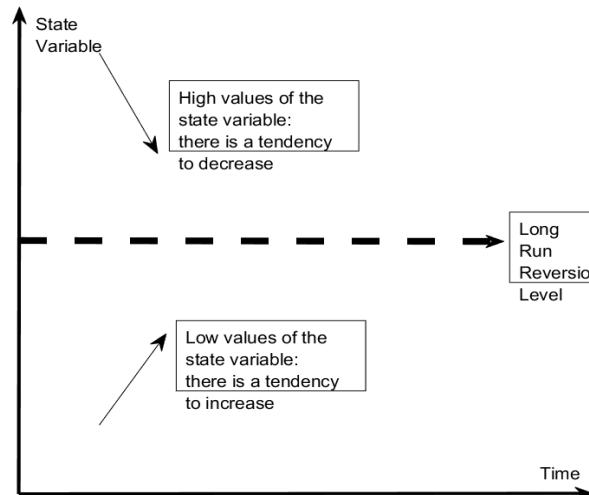


Figure 5.8: Mean reversion and expected change in the state variable (here an interest rate).

- We proceed through the following steps:
 - We let $y(t) = g(t, x) = e^{\alpha t} x(t)$.
 - Then, $dy(t) = \alpha e^{\alpha t} x dt + e^{\alpha t} dx(t)$,
 - Therefore $dy(t) = \alpha e^{\alpha t} x(t) dt + e^{\alpha t} \alpha (\mu - x(t)) dt = e^{\alpha t} \alpha \mu dt$.
 - Finally $y(t) = y(0) + \alpha \mu \int_0^t e^{\alpha s} ds = y(0) + \mu (e^{\alpha t} - 1)$.

We can conclude

Fact 41 *The ODE $dx(t) = \alpha (\mu - x(t)) dt$ admits solution*

$$x(t) = e^{-\alpha t} y(t) \text{ i.e. } x(t) = e^{-\alpha t} x(0) + \mu (1 - e^{-\alpha t}).$$

This is illustrated in Figure (5.9) and in the accompanying Matlab code.

Matlab Code

```
%%%%%%
%%%MEAN REVERSION%%%%%
%%%%%%%
clear all; expiry=10;
timestep=linspace(0,expiry,100)';
mu=100; alpha=1;
sol=@(x0,alpha,timestep) mu+(x0-mu)*exp(-alpha*timestep);

%Plot solutions:
h=figure('Color',[1 1 1]);
fplot(@(timestep) sol(100, alpha,timestep), [0 expiry])
hold on;
fplot(@(timestep) sol(120, alpha,timestep), [0 expiry], 'g')
hold on;
fplot(@(timestep) sol(80, alpha,timestep), [0 expiry], 'b')
hold on;
fplot(@(timestep) sol(120, 2*alpha,timestep), [0 expiry], 'cyan')
hold on;
fplot(@(timestep) sol(80, 2*alpha,timestep), [0 expiry], 'red')
xlabel('Time (years)');
ylabel('X(t)')
legend('Model 1: x_0=100, \mu = 100, \alpha = 1', ...
'Model 2: x_0=120, \mu= 100, \alpha = 1',...
'Model 3: x_0=80, \mu = 100, \alpha = 1',...
'Model 4: x_0=120, \mu = 100, \alpha = 2',...
'Model 5: x_0=80, \mu = 100, \alpha = 2')
title('Convergence of X(t) towards its long-run value')
print(h,'-dpng','LecBMFigMeanReversion.png')
```

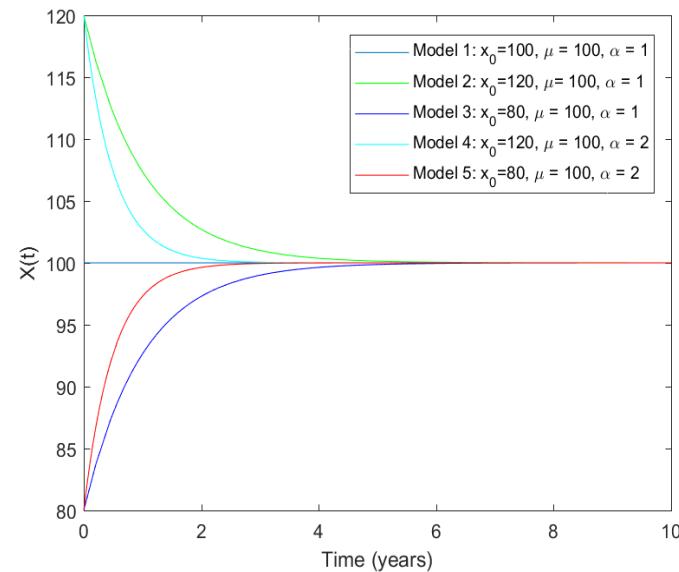


Figure 5.9: Convergence of $X(t)$ towards its long-run value in a mean-reverting model: solution of the equation $dx = \alpha(\mu - x)dt$ changing the initial condition x_0 .

5.2.2. SOLVING THE SDE $dX(t) = \alpha(\mu - X(t))dt + \sigma dW(t)$

- We need to solve

$$dX(t) = \alpha(\mu - X(t))dt + \sigma dW(t).$$

- By analogy with the previous ODE, let us define

$$Y(t) = g(t, Xt) = e^{\alpha t}X(t),$$

and apply Itô's Lemma.

- Then

$$\begin{aligned} & dY(t) \\ &= \left(\underbrace{\frac{\partial g(t, X)}{\partial t}}_{\alpha e^{\alpha t}X} + \alpha(\mu - X) \underbrace{\frac{\partial g(t, X)}{\partial X}}_{e^{\alpha t}} + \frac{1}{2}\sigma^2 \underbrace{\frac{\partial^2 g(t, X)}{\partial X^2}}_0 \right) dt + \sigma \underbrace{\frac{\partial g(t, X)}{\partial X}}_{e^{\alpha t}} dW(t). \\ &= (\alpha e^{\alpha t}X + \alpha(\mu - X)e^{\alpha t}) dt + \sigma e^{\alpha t} dW(t) \\ &= \alpha\mu e^{\alpha t} dt + \sigma e^{\alpha t} dW(t). \end{aligned}$$

- Therefore

$$\begin{aligned} Y(t) &= Y(0) + \int_0^t \alpha\mu e^{\alpha s} ds + \int_0^t \sigma e^{\alpha s} dW(s) \\ &= Y(0) + \mu(e^{\alpha t} - 1) + \int_0^t \sigma e^{\alpha s} dW(s). \end{aligned}$$

Fact 42 The solution of the SDE 5.2.2 is

$$X(t) = e^{-\alpha t}Y(t) = e^{-\alpha t}X(0) + \mu(1 - e^{-\alpha t}) + \sigma \int_0^t e^{-\alpha(t-s)} dW(s).$$

In addition

$$\begin{aligned} X(t) &\sim \mathcal{N}(\mathbb{E}_0(X(t)), \text{Var}_0(X(t))), \\ \mathbb{E}_0(X(t)) &= e^{-\alpha t}X(0) + \mu(1 - e^{-\alpha t}), \\ \text{Var}_0(X(t)) &= \sigma^2 \int_0^t e^{-2\alpha(t-s)} ds = \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha t}); \end{aligned}$$

note that in order to compute the variance, we have exploited the Ito isometry.

5.2.3. THE (AUTO)-COVARIANCE FUNCTION

- Let us consider two time instants, t and s , $t < s$. We have for $t < s$ (but similarly for $s < t$) that $c_X(t, s)$ is given by

$$\begin{aligned} c_X(t, s) &= \text{cov}\left(\sigma \int_0^t e^{-\alpha(t-u)} dW(u), \sigma \int_0^s e^{-\alpha(s-u)} dW(u)\right) \\ &= \sigma^2 e^{-\alpha(t+s)} \text{cov}\left(\int_0^t e^{\alpha u} dW(u), \int_0^t e^{\alpha u} dW(u) + \int_t^s e^{\alpha u} dW(u)\right) \\ &= \sigma^2 e^{-\alpha(t+s)} \text{Var}\left(\int_0^t e^{\alpha u} dW(u)\right) \\ &\quad \text{by the isometry property} \\ &= \sigma^2 e^{-\alpha(t+s)} \int_0^t e^{2\alpha u} du \\ &= \frac{\sigma^2}{2\alpha} e^{-\alpha(t+s)} (e^{2\alpha t} - 1) \end{aligned}$$

- With a similar reasoning, if we take t and s with $s < t$, we have

$$c_X(t, s) = \frac{\sigma^2}{2\alpha} e^{-\alpha(t+s)} (e^{2\alpha s} - 1).$$

Fact 43 (The autocovariance function of the Vasicek model is given by)

$$c_X(t, s) = \frac{\sigma^2}{2\alpha} e^{-\alpha(t+s)} \left(e^{2\alpha \min(s, t)} - 1 \right)$$

This result can be exploited to generate simultaneously the entire trajectory of the Vasicek model: we can simulate the full path by drawing samples from a multivariate normal distribution with the above covariance matrix.

5.2.4. MATLAB: SIMULATION OF THE VASICEK MODEL

Here we simulate the Vasicek model exploiting the solution in (42). A sample of simulated paths is illustrated in Figure (5.10), whilst the distributions originated at different time horizons are presented in Figure (5.11).

Matlab Code

```
%%%%%%SIMULATING THE VASICEK MODEL%%%%%
%%%%%%%%%%%%%%%
clear all;close all
%Model: dX = a * (mu - X ) * dt + sg * dW
%Assign Inputs
X0=0.05; a=10; mu=0.07; sg=0.1; nstep=250; horizon=1;
nsimul=1000; dt=horizon/nstep; timestep=[0:nstep]*dt;
%Compute the variance of the increments
vol2=(1-exp(-2*a*dt))/(2*a);
Xall=[];%variable where to store the simulations
for j=1:nsimul
    %Initialize the interest rate vector
    X=zeros(nstep+1,1); X(1)=X0;
    %Simulate the increments of the BM
    dW=randn(nstep,1)*vol2^0.5;
    %Start iteration
    for i=1:nstep
        X(i+1)=mu+exp(-a*dt)*(X(i)-mu)+sg*dW(i);
    end
    %store the simulated path
    Xall=[Xall, X];
end
%Plot the sample paths
h=figure('Color',[1 1 1]);
%plot a limited number of simulations for better simulation
plot([0:nstep]*dt,[Xall(:,1:50), mu+(X0-mu)*exp(-a*timestep)]);
xlabel('Time')
legend('Simulated paths','Expected path')
print(h,'-dpng','LecBM_SimVasicek_new')
```

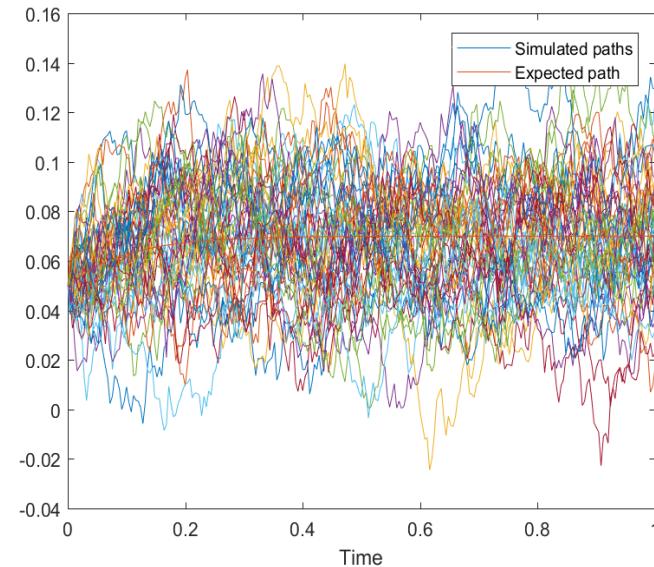


Figure 5.10: Simulated paths of the Vasicek model $dX = \alpha(\mu - X)dt + \sigma dW(t)$. Parameters: $r_0 = 0.05$; $\alpha=10$; $\mu = 0.07$; $\sigma = 0.1$; $nstep=200$; $horizon=1$.

Matlab Code

```
%%%%%%VASICEK densities at different time horizons%%%%%
%%%%%clear all;
%Fix the time horizons
horizon=[0.25 0.5 0.75 1 5];
%Assign new parameters
mu=0.09; sg=0.05; a=0.8; X0=0.04;
%Compute Exp. Value and variance
meanVas=mu+exp(-a*horizon).*(X0-mu);
varVas=sg*sg*(1-exp(-2*a*horizon))/(2*a);
range=linspace(mu-3*sg/(2*a)^0.5,...  

mu+3*sg/(2*a)^0.5,200);
pdfV=[];
for i=1:length(horizon)
meanV=meanVas(i);
stdV=varVas(i).^0.5;
pdfVas=pdf('norm', range, meanV, stdV);
pdfV=[pdfV;pdfVas];
end
h=figure('Color',[1 1 1]);
plot(range,pdfV)
title('PDF of X(t) at different horizons')
xlabel('X(t)');ylabel('pdf')
legend('0.25 yrs','0.5 yrs','0.75 yrs','1 years','5 yrs')
print(h, '-dpng', 'LecBMFpdfVasicek')
```

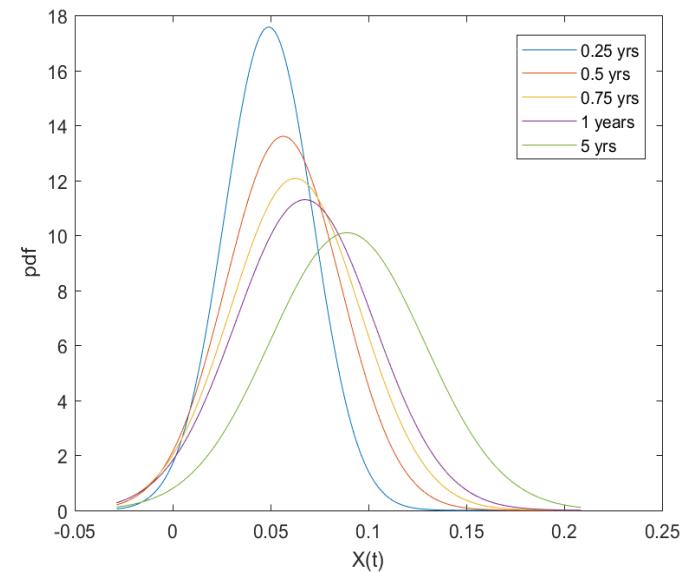


Figure 5.11: Density of the $\text{MR}(0.09, 0.8, 0.05)$ model at different horizons.

Synthetic properties of the model are given in Table 5.5. Simulated paths of the Vasicek model, given different starting values, are shown in the videos in 5.12-5.14.

MEAN-REVERTING PROCESS MR(α, μ, σ): FACTS

The SDE

$$dX(t) = \alpha(\mu - X(t))dt + \sigma dW(t), X(0) = x_0.$$

The solution

$$X(t) = e^{-\alpha t}X(0) + \mu(1 - e^{-\alpha t}) + \sigma \int_0^t e^{-\alpha(t-s)}dW(s).$$

The distribution of $X(t)$

$$X(t) \sim \mathcal{N}(\mathbb{E}_0(X(t)), \text{Var}_0(X(t))).$$

The mean of $X(t)$

$$\mathbb{E}_0(X(t)) = e^{-\alpha t}X(0) + \mu(1 - e^{-\alpha t}).$$

The variance of $X(t)$

$$\text{Var}_0(X(t)) = \sigma^2 \int_0^t e^{-2\alpha(t-s)}ds = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}).$$

The stationary distribution of $X(t)$, ($t \rightarrow \infty$)

$$X(t) \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{2\alpha}\right) \text{ if } \alpha > 0.$$

The auto-covariance of $X(t)$

$$c_X(t, s) = \frac{\sigma^2}{2\alpha}e^{-\alpha(t+s)} \left(e^{2\alpha \min(s,t)} - 1\right).$$

Table 5.5: Properties of the Vasicek model



Figure 5.12: Movie with simulated paths of the Vasicek model starting above the long term level.



Figure 5.13: Movie with simulated paths of the Vasicek model starting at the long term level.



Figure 5.14: Movie with simulated paths of the Vasicek model starting below the long term level.

5.2.5. EXTENSION: MR WITH DETERMINISTIC VOLATILITY

- We can generalize the Vasicek model to a deterministic time-varying volatility.
- The SDE becomes

$$dX(t) = \alpha(\mu - X(t))dt + \sigma(t)dW(t).$$

- It has solution

$$X(t) = X(0) e^{-\alpha t} + \mu (1 - e^{-\alpha t}) + \int_0^t \sigma(s) e^{-\alpha(t-s)} dW(s).$$

- The solution has the following properties (variance and covariance are computed using the isometry property):

$$\begin{aligned} X(t) &\sim \mathcal{N}(\mathbb{E}_0(X(t)), \mathbb{V}ar_0(X(t))), \\ \mathbb{E}_0(X(t)) &= X(0) e^{-\alpha t} + \mu (1 - e^{-\alpha t}), \\ \mathbb{V}ar_0(X(t)) &= \int_0^t \sigma^2(s) e^{-2\alpha(t-s)} ds; \\ \mathbb{C}ov_0(X(t), X(s)) &= \int_0^{\min(t,s)} \sigma^2(u) e^{-\alpha(t+s-2u)} du. \end{aligned}$$

- The case of a time-dependent drift can be treated by the same argument, so that

$$dX(t) = \alpha (\mu(t) - X(t)) dt + \sigma(t) dW(t).$$

whose solution becomes

$$X(t) = X(0) e^{-\alpha t} + \int_0^t \mu(s) e^{-\alpha(t-s)} ds + \int_0^t \sigma(s) e^{-\alpha(t-s)} dW(s).$$

5.3. THE COX-INGERSOLL-ROSS (CIR) MODEL

The SDE is given by

$$dX(t) = \alpha(\mu - X(t))dt + \sigma\sqrt{X(t)}dW(t).$$

This model has been introduced by Cox, Ingersoll and Ross to model the dynamics of the instantaneous interest rate. The peculiar form of the diffusion coefficient has been chosen to ensure that the process does not achieve negative values, still preserving analytical tractability.

5.3.1. SOLVING THE SDE $dX(t) = \alpha(\mu - X(t))dt + \sigma\sqrt{X(t)}dW(t)$

- The SDE of the short rate is given by

$$dX(t) = \alpha(\mu - X(t))dt + \sigma\sqrt{X(t)}dW(t).$$

where $\mu \geq 0$, $X(0) \geq 0$, $\sigma \geq 0$ and $\alpha \in R$.

- If $\mu = 0$ and $X(0) = 0$, the solution of the SDE is $X(t) = 0$, for all t , and $X(t) \geq 0$ for $\mu \geq 0$ and $X(0) \geq 0$.
- This model shares with the Vasicek one the form of the drift term, so that it allows for mean reversion for $\alpha > 0$ and interest rates cannot explode.
 - The mean-reversion property implies also that interest rate displays a steady state distribution.
- The difference with respect to the Vasicek model is the appearance of the square root term \sqrt{X} in the diffusion term.
 - This ensures that the process remains non-negative in every instant of time and a zero rate of interest can become positive again.
 - The level of absolute variance increases with increasing interest rates.

- Unfortunately, this SDE does not admit an explicit solution, as in the case of the Vasicek model. However, a semi-explicit solution is possible by relating $X(t)$ to its full history.
- Indeed, following the same steps as in the Vasicek model, we obtain

$$X(t) = e^{-\alpha t} X(0) + \mu (1 - e^{-\alpha t}) + \sigma \int_0^t e^{-\alpha(t-u)} \sqrt{X(u)} dW(u),$$

where we observe that the current value of X depend on the its path through the stochastic integral term.

- Therefore, we can say that $X(t)$ is a process with memory. This is completely different from the ABM process, which is a memoryless process.
- The CIR model is less tractable than the Vasicek one: the distribution of the short rate is related to the non-central chi-square distribution.
- We can obtain few properties of the solution such as expected value, variance and distribution.
- To do this, we exploit Ito's lemma.

5.3.2. COMPUTING THE EXPECTATION OF THE CIR MODEL

- In particular, the expectation of $X(t)$, $\mu_X(t) = E(X(t))$, is the same as in the Vasicek model and is obtained by solving the ODE

$$d\mu_X(t) = \alpha (\mu - \mu_X(t)) dt.$$

- Therefore, we have

$$\mu_X(t) = \mathbb{E}_0(X(t)) = e^{-\alpha t} X(0) + \mu (1 - e^{-\alpha t}).$$

- We observe that for long horizons ($t \rightarrow \infty$), we have

$$\mu_X(t) \rightarrow \mu.$$

5.3.3. COMPUTING THE VARIANCE OF THE CIR MODEL

- For computing $\mathbb{V}ar_0(X(t))$, we proceed as follows

1. Let us define

$$Y(t) = X^2(t).$$

2. Let us find the SDE for $Y(t)$ by applying Ito's lemma:

$$\begin{aligned} dY(t) &= (2X(t)\alpha(\mu - X(t)) + \sigma^2 X(t))dt + 2X(t)\sigma\sqrt{X(t)}dW(t) \\ &= \left(X(t)(2\alpha\mu + \sigma^2) - 2\alpha Y(t) \right) dt + 2X(t)\sigma\sqrt{X}dW(t). \end{aligned}$$

3. Let us compute $\mu_Y(t) = \mathbb{E}_0(Y(t))$ by solving the ODE

$$d\mu_Y(t) = \left(\mu_X(t)(2\alpha\mu + \sigma^2) - 2\alpha\mu_Y(t) \right) dt.$$

By solving and rearranging terms we obtain

$$\mu_Y(t) = X(0) \left(2\mu + \frac{\sigma^2}{\alpha} \right) e^{-\alpha t} (1 - e^{-\alpha t}) + \left(\mu^2 + \mu \frac{\sigma^2}{2\alpha} \right) (1 - e^{-\alpha t})^2 + X(0)^2 e^{-2\alpha t}.$$

4. This can be done writing

$$d\mu_Y(t) + 2\alpha\mu_Y(t)dt = \mu_X(t)(2\alpha\mu + \sigma^2)dt,$$

and recognizing that the term on the left hand side of the above, once multiplied by $e^{2\alpha t}$, is the derivative of $e^{2\alpha t}\mu_Y(t)$.

5. So, multiply both sides by $e^{2\alpha t}$, integrate and get $\mu_Y(t)$.

6. Finally, the variance of $X(t)$ is obtained as $\mu_Y(t) - \mu_X^2(t)$:

$$\mathbb{V}ar_0(X(t)) = X(0) \frac{\sigma^2}{\alpha} e^{-\alpha t} (1 - e^{-\alpha t}) + \mu \frac{\sigma^2}{2\alpha} (1 - e^{-\alpha t})^2.$$

7. Observe that for long horizons, we have:

$$\mathbb{V}ar_0(X(t)) = \mu \frac{\sigma^2}{2\alpha}.$$

5.3.4. AGAIN ON COMPUTING THE VARIANCE AND THE AUTO-COVARIANCE OF THE CIR MODEL

A different derivation of the variance of the CIR process is illustrated here.

$$\begin{aligned}
 \mathbb{V}ar(X(t)) &= \sigma^2 \mathbb{V}ar_0 \left(\int_0^t e^{-\alpha(t-u)} \sqrt{X(u)} dW(u) \right) \\
 &= \text{by the isometry property} \\
 &= \sigma^2 \int_0^t e^{-2\alpha(t-u)} \mathbb{E}_0(X(u)) du \\
 &= \text{by the expected value of } X(t) \\
 &= \sigma^2 \int_0^t e^{-2\alpha(t-u)} (e^{-\alpha u} X(0) + \mu (1 - e^{-\alpha u})) du \\
 &= \sigma^2 e^{-2\alpha t} \int_0^t (e^{\alpha u} X(0) + \mu (e^{2\alpha u} - e^{\alpha u})) du \\
 &= \sigma^2 e^{-2\alpha t} \left(\frac{e^{\alpha t} - 1}{\alpha} X(0) + \mu \left(\frac{e^{2\alpha t} - 1}{2\alpha} - \frac{e^{\alpha t} - 1}{\alpha} \right) \right) \\
 &= \sigma^2 \frac{e^{-2\alpha t}}{\alpha} \left((e^{\alpha t} - 1) X(0) + \frac{\mu}{2} (e^{2\alpha t} - 2e^{\alpha t} + 1) \right) \\
 &= \sigma^2 \left(\frac{e^{-\alpha t} - e^{-2\alpha t}}{\alpha} \right) X(0) + \sigma^2 \frac{\mu}{2\alpha} (1 - 2e^{-\alpha t} + e^{-2\alpha t}) \\
 &= \sigma^2 \left(\frac{e^{-\alpha t} - e^{-2\alpha t}}{\alpha} \right) X(0) + \sigma^2 \frac{\mu}{2\alpha} (1 - e^{-\alpha t})^2
 \end{aligned}$$

For computing the (auto)-covariance function in the CIR model, we consider two time instants, t and s , $t < s$. We have for $t < s$ (but similarly for $s < t$) that $c_X(t, s)$ is given by

$$c_X(t, s) = cov(X(t), X(s)).$$

Then we proceed in a similar way to the derivation of the variance.

$$\begin{aligned}
 \mathbb{C}ov_0(X(t), X(s)) &= \sigma^2 cov \left(\int_0^t e^{-\alpha(t-u)} \sqrt{X(u)} dW(u), \int_0^s e^{-\alpha(s-u)} \sqrt{X(u)} dW(u) \right) \\
 &= \sigma^2 cov \left(\int_0^t e^{-\alpha(t-u)} \sqrt{X(u)} dW(u), \int_0^t e^{-\alpha(s-u)} \sqrt{X(u)} dW(u) + \int_t^s e^{-\alpha(s-u)} \sqrt{X(u)} dW(u) \right) \\
 &= \sigma^2 cov \left(\int_0^t e^{-\alpha(t-u)} \sqrt{X(u)} dW(u), \int_0^t e^{-\alpha(s-u)} \sqrt{X(u)} dW(u) \right) \\
 &= \sigma^2 e^{-\alpha(t+s)} \mathbb{C}ov \left(\int_0^t e^{\alpha u} \sqrt{X(u)} dW(u), \int_0^t e^{\alpha u} \sqrt{X(u)} dW(u) \right) \\
 &= \sigma^2 e^{-\alpha(t+s)} \mathbb{V}ar \left(\int_0^t e^{\alpha u} \sqrt{X(u)} dW(u) \right) \\
 &= \sigma^2 e^{-\alpha(t+s)} \int_0^t e^{2\alpha u} \mathbb{E}(X(u)) du \\
 &= \sigma^2 e^{-\alpha(t+s)} \int_0^t e^{2\alpha u} (e^{-\alpha u} X(0) + \mu (1 - e^{-\alpha u})) du \\
 &= \sigma^2 e^{-\alpha(t+s)} \int_0^t (e^{\alpha u} X(0) + \mu (e^{2\alpha u} - e^{\alpha u})) du \\
 &= \sigma^2 e^{-\alpha(t+s)} \left(\frac{e^{\alpha t} - 1}{\alpha} X(0) + \mu \left(\frac{e^{2\alpha t} - 1}{2\alpha} - \frac{e^{\alpha t} - 1}{\alpha} \right) \right) \\
 &= \sigma^2 e^{-\alpha(t+s)} \left(\frac{e^{\alpha t} - 1}{\alpha} X(0) + \mu \left(\frac{e^{2\alpha t} - 2e^{\alpha t} + 1}{2\alpha} \right) \right) \\
 &= \sigma^2 \frac{e^{-\alpha s} - e^{-\alpha(t+s)}}{\alpha} X(0) + \sigma^2 \frac{\mu}{2\alpha} e^{-\alpha(t+s)} (e^{\alpha t} - 1)^2
 \end{aligned}$$

In particular, setting $s = t$ we obtain again the expression for the variance.

5.3.5. THE DISTRIBUTION OF THE SHORT RATE IN THE CIR MODEL

A detailed discussion on the properties of the CIR model can be found in Cairns (2004).

- In order to get an intuition on the distribution of the short rate, let us consider the SDE

$$dX(t) = -\frac{\alpha}{2}X(t)dt + \frac{\sigma}{2}dW(t).$$

- We recognize that $X(t)$ is a OU process (see Section 5.2); hence its properties are known. In particular, the process is Gaussian with mean

$$m(t) = e^{-\frac{\alpha t}{2}}X(0)$$

and variance

$$s^2(t) = \sigma^2 \frac{(1 - e^{-\alpha t})}{8\alpha}.$$

- Now let us define $R(t) = X^2(t)$ and let us apply Ito's lemma to it; then

$$dR(t) = \alpha(\mu - R(t))dt + \sigma\sqrt{R(t)}dW(t),$$

with $\mu = \sigma^2/(4\alpha)$. Therefore $R(t)$ follows a square root process.

- To understand what is the distribution of R , we proceed as follows.

1. With a little abuse of notation, let us write

$$X = sZ + m,$$

where Z is a standard Gaussian random variable.

2. It follows that R is the square of a non-standard Gaussian distribution, i.e.

$$R = X^2 = s \left(Z + \frac{m}{s} \right)^2.$$

This implies that

$$\frac{R}{s}$$

has a non central chi-square distribution with 1 degree of freedom and parameter of non-centrality m/s (see Appendix).

- This shows that the distribution of R , i.e. the solution of the square-root SDE, is related to a non-central chi-square distribution with 1 degree of freedom and parameter of non centrality m/s .
- However, if we generalize to

$$R(t) = \sum_i^d X_i^2(t)$$

where the X_i are d iid processes like in the case considered above, with coefficients α_i and σ_i , R will still have a non-central chi-square distribution but now with d degrees of freedom

- The result can be generalized to the more general case of a non integer number d .

Fact 44 (Distribution of X in the CIR model) *Let us assume that $X(t)$ follows the square-root SDE*

$$dX(t) = \alpha(\mu - X(t))dt + \sigma\sqrt{X(t)}dW(t).$$

Let us set

$$k = \frac{\sigma^2 (1 - e^{-\alpha(T-t)})}{4\alpha}.$$

Then, the distribution of

$$\frac{X(T)}{k}$$

conditioned on $X(t)$ is a non-central chi-square distribution with d degrees of freedom and non centrality parameter λ , where

$$d = \frac{4\alpha\mu}{\sigma^2}, \lambda = \frac{4\alpha X(t)}{\sigma^2(e^{\alpha(T-t)} - 1)}.$$

In particular, the density of $X(T)$, conditioned on $X(t)$, is given by

$$\frac{1}{2k} \left(\frac{X(T)e^{\alpha(T-t)}}{X(t)} \right)^{\frac{d/2-1}{2}} e^{-\frac{X(t)e^{-\alpha(T-t)}+X(T)}{2k}} I_{d/2-1} \left(\frac{\sqrt{X(t)X(T)e^{\alpha(T-t)}}}{e^{\alpha(T-t)}k} \right),$$

where $I_v(x)$ is the modified Bessel function of the first type of order v . In addition, applying the properties of the non-central chi-square distribution, the expectation of $X(t)$ given $X(0)$ is

$$\mu_X(t) = k(d + \lambda) = \mathbb{E}_0(X(t)) = e^{-\alpha t} X(0) + \mu(1 - e^{-\alpha t}).$$

and its variance is

$$\mathbb{V}ar_0(X(t)) = 2k^2(d + 2\lambda) = X(0) \left(\frac{\sigma^2}{\alpha} \right) (e^{-\alpha t} - e^{-2\alpha t}) + \mu \left(\frac{\sigma^2}{2\alpha} \right) (1 - e^{-\alpha t})^2.$$

These expressions are the same as the ones found earlier on applying Ito's lemma and the isometry property.

Fact 45 (Stationary Distribution of X in the CIR model) The stationary distribution of $X(t)$ for large t is Gamma, with density function

$$\frac{\omega^\nu X(t)^{\nu-1} e^{-\omega X(t)}}{\Gamma(\nu)},$$

where

$$\omega = \frac{2\alpha}{\sigma^2}, \nu = \frac{2\alpha\mu}{\sigma^2},$$

and $\Gamma(x)$ is the Gamma function. In addition the stationary mean and variance are respectively equal to μ and $\sigma^2\mu/(2\alpha)$.

The distribution of the CIR process, as the time horizon lengthens, is shown in Figure 5.15 (see accompanying Matlab code). The asymptotic density is given in figure 5.16.

Fact 46 (The Feller positivity condition) If

$$\frac{2\alpha\mu}{\sigma^2} > 1 \tag{5.1}$$

then X will never reach zero.

See Figure 5.17 for an illustration of violations of the Feller condition.

Matlab Code Here we provide the code for generating the probability density function in the CIR model, for different horizons.

```
%CIR Parameters & time horizon
alpha=0.03; sigma=0.02; Xt=0.04; mu=0.03;
tau=[1 5 10 15 20 50];
%compute d
d=4 *alpha *mu/sigma^2;
%grid values for X
XT=linspace(0,0.12,200);
for i=1:length(tau)
    lambda=4*alpha* Xt/(sigma^2 *(exp(alpha*tau(i))-1));
    k=sigma^2 *(1-exp(-alpha*tau(i)))/(4*alpha);
    pdfcir(:,i)=pdf('ncx2',XT/k,d,lambda)/k;
    %check expected value
    EXTnum(i)=trapz(XT,XT.*pdfcir(:,i)');
end
h=figure('Color',[1 1 1]);
plot(XT, pdfcir, '.'); xlabel('X'); ylabel('pdf')
legend('T-t=1','T-t=5','T-t=10','T-t=15','T-t=20','T-t=20')
title('The CIR pdf at different horizons')
print(h, '-dpng', 'FigshortrateCIR')
```

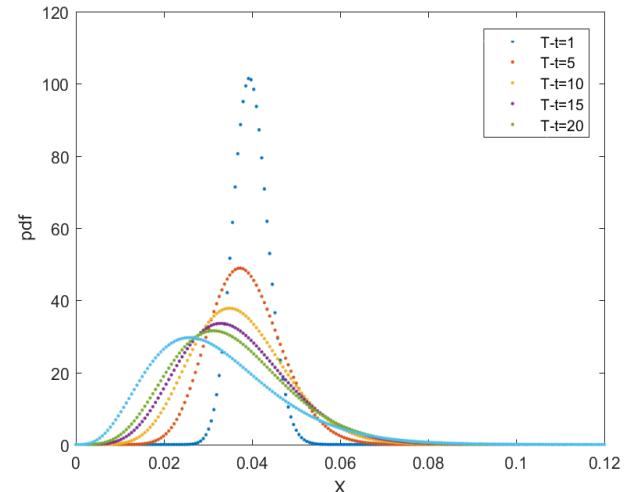


Figure 5.15: Density of the short rate in the CIR model at different horizons T . $\alpha = 0.1$, $\sigma = 0.01$, $r(t) = 0.03$, $\mu = 0.05$.

5.3.6. THE EXTENDED-CIR MODEL

- An interesting generalization of the CIR model is obtained by considering time dependent coefficients.
- This is particularly relevant for the model to be consistent with a set of observed zero-coupon bond prices, i.e. if the instantaneous short rate follows the CIR process under the risk-neutral measure the price of a zero coupon bond can be obtained by computing

$$\tilde{\mathbb{E}}_0 \left(e^{- \int_0^t r(s) ds} \right).$$

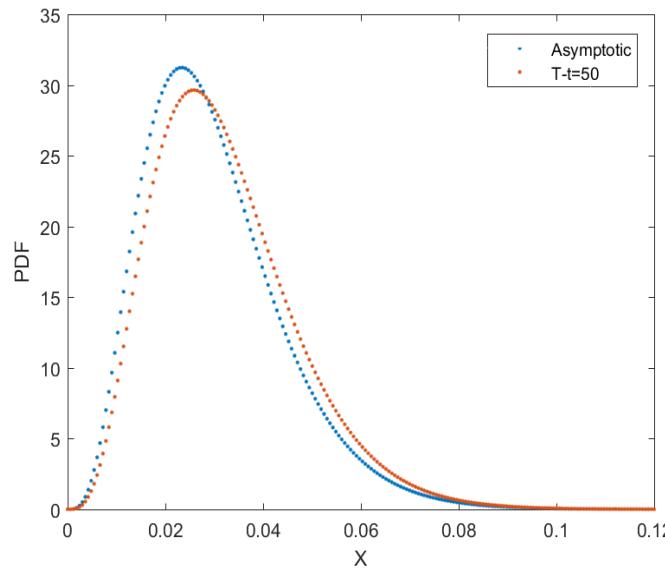


Figure 5.16: Asymptotic density in the CIR model. $\alpha = 0.03$, $\sigma = 0.02$, $X(t) = 0.04$, $\mu = 0.03$.

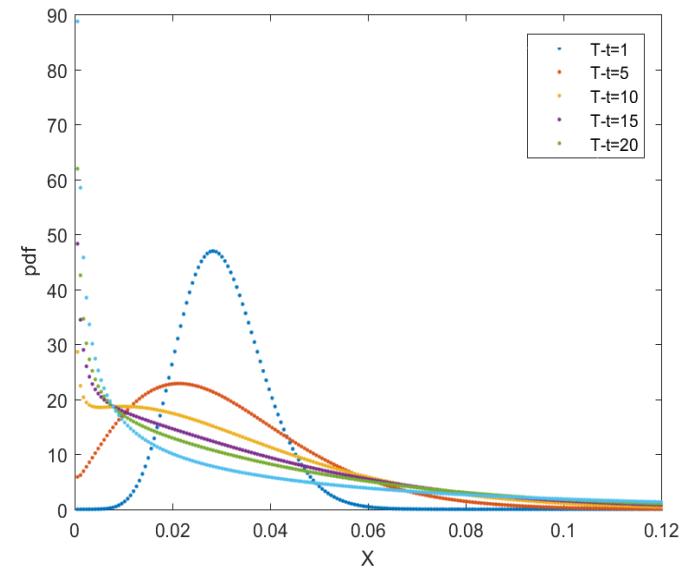


Figure 5.17: Short rate density in the CIR model when the Feller condition (5.1) is violated. $\alpha = 0.01$, $\sigma = 0.05$, $X(t) = 0.03$, $\mu = 0.05$.

- Closed form expression for it are available. Unfortunately, given the limited number of parameters, it will not be possible to price exactly all quoted zero-coupon bonds.
- For this reason, Maghsoudi (1996) examines the so called Extended-CIR (ECIR) model

$$dX(t) = \alpha(t)(\mu(t) - X(t))dt + \sigma(t)\sqrt{X(t)}dW(t) \quad (5.2)$$

- In particular, if

$$\delta(t) = 4\frac{\mu(t)\alpha(t)}{\sigma^2(t)} = d \in N$$

the solution of (5.2) can be represented in terms of a generalized d -dimensional Ornstein-Uhlenbeck process.

- These results also show that the CIR and ECIR models can be regarded as being inherently of multifactor nature.
- In addition, the extended interest rate is always non-negative, and when $d = 1$ it cannot avoid zero in finite time and when $d > 1$ it does avoid zero.
- Finally, future extended spot rates too have a rescaled non-central chi-squared distribution.
- In the original paper by Cox Ingersoll and Ross a square-root process with time dependent drift term is also considered

$$dX(t) = \alpha(\mu(t) - X(t))dt + \sigma\sqrt{X(t)}dW(t).$$

- In this model, the future expected value given becomes

$$\tilde{E}_0(X(t)) = x(0)e^{-\alpha t} + \alpha \int_0^t e^{-\alpha(t-u)}\mu(u)du.$$

- For example, if $X(t)$ represents a commodity spot price and the above expectation is taken under the risk neutral measure, the above expected value must be equal to the commodity forward price. Therefore

$$F(0, t) = F(0, 0)e^{-\alpha t} + \alpha \int_0^t e^{-\alpha(t-u)}\mu(u)du,$$

and then taking the derivative with respect to t we obtain

$$\begin{aligned} \frac{\partial F(0, t)}{\partial t} &= -\alpha \left(F(0, 0)e^{-\alpha t} + \alpha \int_0^t e^{-\alpha(t-u)}\mu(u)du \right) + \alpha\mu(t) \\ &= -\alpha F(0, t) + \alpha\mu(t). \end{aligned}$$

- In conclusion, the time varying drift can be calibrated to the observed forward term structure by imposing

$$\mu(t) = F(0, t) + \frac{1}{\alpha} \frac{\partial F(0, t)}{\partial t}.$$

- The extended CIR process is also considered in Fusai et al. (2008) to make the spot price dynamics consistent with the quoted futures curve and the volatility term structure.
- Still in the context of zero-coupon bond pricing, a tractable approach is suggested in Brigo and Mercurio (2001) where the short rate is modelled according to

$$r(t) = \theta(t) + X(t),$$

i.e. an exogenous time-dependent shift is added to the model to guarantee consistency with an input term structure of rates. Indeed, if $P^{mkt}(0, t)$ are the quoted zero-coupon prices, the model-to-market consistency is guaranteed if

$$e^{-\int_0^t \theta(u)du} = \frac{P^{mkt}(0, t)}{\tilde{E}_0 \left(e^{-\int_0^t r(s)ds} \right)},$$

or

$$\theta(t) = -\frac{\partial \left(\ln P^{mkt}(0, t) - \ln \tilde{E}_0 \left(e^{-\int_0^t r(s)ds} \right) \right)}{\partial t}, \forall t > 0.$$

5.3.7. SIMULATING THE CIR MODEL

- We can simulate the CIR model in at least three different ways:

1. **Euler simulation:** we replace $dW(t)$ by $\sqrt{\Delta t} \times \epsilon(t + \Delta t)$ where $\epsilon(t + \Delta t)$ is a standard Gaussian random variable drawn $+\Delta t$ at time t

$$X(t + \Delta t) = X(t) + \alpha(\mu - X(t))\Delta t + \sigma \sqrt{X(t)\Delta t} \epsilon(t + \Delta t). \quad (5.3)$$

One immediate problem with the Euler scheme above is that the discrete process for X can become negative with non-zero probability, which in turn would make computation of \sqrt{X} impossible and the Euler scheme will fail. The following Modified Euler scheme provides a simple solution to this problem.

2. **Modified Euler simulation:** Among possible remedies, the fix that appears to produce the smallest discretization bias consists in replacing at each step $X(t)$ by $\max(X(t), 0)$, so that

$$X(t + \Delta t) = X(t) + \alpha(\mu - X(t))\Delta t + \sigma \sqrt{\max(X(t), 0) \Delta t} \epsilon(t + \Delta t). \quad (5.4)$$

3. **Gaussian approximation** using exact moments: in the Euler discretization we use the exact mean and the exact standard deviation, rather than the discretised version of the drift and diffusion coefficient

$$X(t + \Delta t) = e^{-\alpha\Delta t} X(t) + \mu \left(1 - e^{-\alpha\Delta t}\right) + \sqrt{\max(\text{Var}_t(X(t + \Delta t), 0))} \epsilon(t + \Delta t), \quad (5.5)$$

where

$$\text{Var}_t(X(t + \Delta t)) = X(t) \left(\frac{\sigma^2}{\alpha}\right) \left(e^{-\alpha\Delta t} - e^{-2\alpha\Delta t}\right) + \mu \left(\frac{\sigma^2}{2\alpha}\right) \left(1 - e^{-\alpha\Delta t}\right)^2.$$

4. **Exact simulation:** we iteratively simulate from a non-central chi-square distribution changing the non-centrality parameter according to the current level of X

$$X(t + dt) = k \chi_{d, \lambda}^2, \quad (5.6)$$

with

$$d = \frac{4\alpha\mu}{\sigma^2}, \lambda = \frac{4\alpha X(t)}{\sigma^2(e^{\alpha\Delta t} - 1)}, k = \frac{\sigma^2(1 - e^{-\alpha\Delta t})}{4\alpha}.$$

so at each time step we have to simulate a non-central chi-square distribution. This can be achieved using the Matlab command `icdf('ncx2', rand, d, lambda)` that exploits the inverse CDF method, that simulates the random variable by applying to a uniform random variable the inverse CDF. An even more direct approach is to use the command `ncx2rnd(d, lambda)` that directly simulates from a noncentral chi-square distribution.

- The third method is exact (i.e. no discretization error and no bias) and ensures positive samples, but is more computationally demanding, as shown in the following table where we provide the CPU time (in seconds) for simulating one path with 200 steps.

Method	CPU (secs)
Euler	0.008382
Gaussian	0.010881
Exact (inverse method)	61.87
Exact (ncx2rnd)	0.438

SQUARE-ROOT MEAN-REVERTING PROCESS $\text{MR}(\alpha, \mu, \sigma)$: FACTS

The SDE

$$dX(t) = \alpha (\mu - X(t)) dt + \sigma \sqrt{X(t)} dW(t), X(0) = x_0.$$

The solution of the SDE

is not explicit.

The distribution of $X(t)$

$$X(t) \sim k \times \chi_{d,\lambda}^2.$$

The mean of $X(t)$

$$\mathbb{E}_0(X(t)) = e^{-\alpha t} X(0) + \mu (1 - e^{-\alpha t}).$$

The variance of $X(t)$

$$\mathbb{V}\text{ar}_0(X(t)) = X(0) \left(\frac{\sigma^2}{\alpha} \right) (e^{-\alpha t} - e^{-2\alpha t}) + \mu \left(\frac{\sigma^2}{2\alpha} \right) (1 - e^{-\alpha t})^2.$$

If $2\alpha\mu > \sigma^2$, the stationary distribution ($t \rightarrow \infty$) of $X(t)$ is Gamma

$$X(t) \sim \frac{\omega^\nu X(t)^{\nu-1} e^{-\omega X(t)}}{\Gamma(\nu)}, \text{ if } \alpha > 0,$$

The stationary mean of $X(t)$, ($t \rightarrow \infty$).

$$\mu.$$

The stationary variance of $X(t)$, ($t \rightarrow \infty$)

$$\mu \frac{\sigma^2}{2\kappa}.$$

Table 5.6: Properties of the Square-Root model

- Therefore, we suggest to use the second method, Gaussian discretization with exact moments. It allows to achieve a good trade-off between accuracy and computational time.
- A sample trajectory is shown in Figure 5.19 (see accompanying Matlab code as well). The Euler discretization simulation is illustrated in Figure 5.18, whilst the exact method is given in Figure 5.20.
- Table 5.6 gives the synthetic properties of the Cox-Ingersoll-Ross model.

Matlab Code

```
%%%%%%SIMULATING THE CIR MODEL: EULER DISCRETIZATION %%%%
%%%%%
clear all; clc;
%Model parameters
alpha=0.01; sigma=0.05; Xt=0.03; mu=0.05;
%Simulation parameters
nstep=200; horizon=1; nsimul=100;
dt=horizon/nstep;
Xall=[];%variable to store simulated values
for j=1:nsimul
    X=zeros(nstep+1,1); X(1)=Xt;
    for i=1:nstep
        X(i+1)=X(i) + alpha*(mu-X(i))*dt+...
            sigma *sqrt(X(i)*dt)*randn;
    end
    Xall=[Xall, X];
end
h=figure('Color',[1 1 1]);
plot(dt*[0:nstep]',Xall)
xlabel('Time')
title('Simulating the CIR model via Euler discretization')
print(h,'-dpng','FigshortrateCIR_Euler')
```

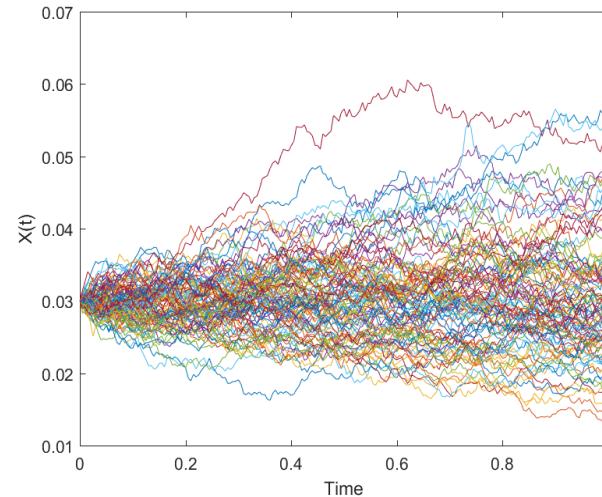


Figure 5.18: Simulated paths of the CIR model. Parameters: $\alpha = 0.01$, $\sigma = 0.05$, $r(t) = 0.03$, $\mu = 0.05$.

Matlab Code

```
%SIMULATING THE CIR MODEL: GAUSSIAN APPROXIMATION %
clear all; clc; close all
alpha=0.01; sigma=0.05; Xt=0.03; mu=0.05;
nstep=200; horizon=1; nsimul=100;
dt=horizon/nstep;
Xall=[];
for j=1:nsimul
    X=zeros(nstep+1,1); X(1)=Xt;
    for i=1:nstep
        m=exp(-alpha*dt)*X(i)+mu*(1-exp(-alpha*dt));
        v=r(i)*(sigma*sigma/alpha)*...
            (exp(-alpha*dt)-exp(-2*alpha*dt))...
            +mu*(sigma*sigma/(2*alpha))*...
            (1-exp(-alpha*dt))^2;
        X(i+1)=m + sqrt(v)*randn;
    end
    Xall=[Xall, X];
end
h=figure('Color',[1 1 1]);
plot(dt*[0:nstep]',Xall)
xlabel('Time')
title('Simulating the CIR model with exact moments')
print(h,'-dpng','FigshortrateCIR_Gaussian')
```

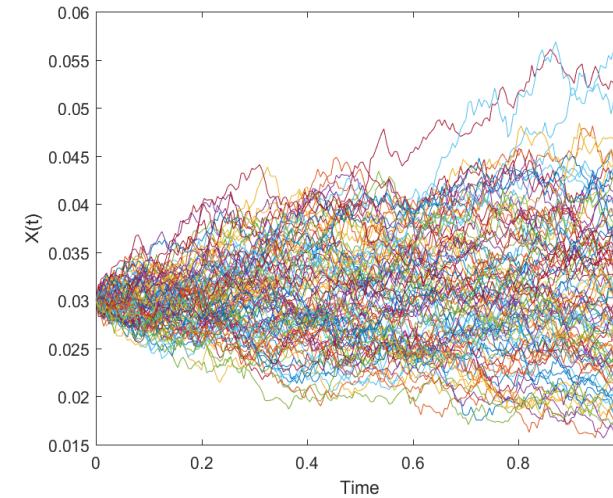


Figure 5.19: Simulated paths of the CIR model with exact moments. Parameters: $\alpha = 0.01$, $\sigma = 0.05$, $r(t) = 0.03$, $\mu = 0.05$.

Matlab Code

```
%SIMULATING THE CIR MODEL: EXACT METHOD
clear all; clc
%CIR Parameters
alpha=0.01; sigma=0.05; Xt=0.03; mu=0.05;
nstep=200; horizon=1; nsimul=20;
dt=horizon/nstep;
X=zeros(nstep+1,1); X(1)=Xt;
Xall=[];
%compute d,k
d=4 *alpha *mu/sigma^2;
k=sigma^2 *(1-exp(-alpha*dt)) / (4*alpha);
tic
for j=1:nsimul
    X=zeros(nstep+1,1); X(1)=Xt;
    for i=1:nstep
        lambda=4*alpha* X(i)/(sigma^2 *(exp(alpha*dt)-1));
        %%use the inverse of the non-central chi2
        X(i+1,:)=icdf('ncx2',rand,d,lambda)*k;
        %use the ncx2rnd command
        %X(i+1,:)=ncx2rnd(d,lambda)*k;
    end
    Xall=[Xall, X];
end
toc
h=figure('Color',[1 1 1]);
plot(dt*[0:nstep]',Xall); xlabel('Time')
title('Simulating the CIR model via Exact Simulation')
print(h,'-dpng','FigshortrateCIR.Exact')
```

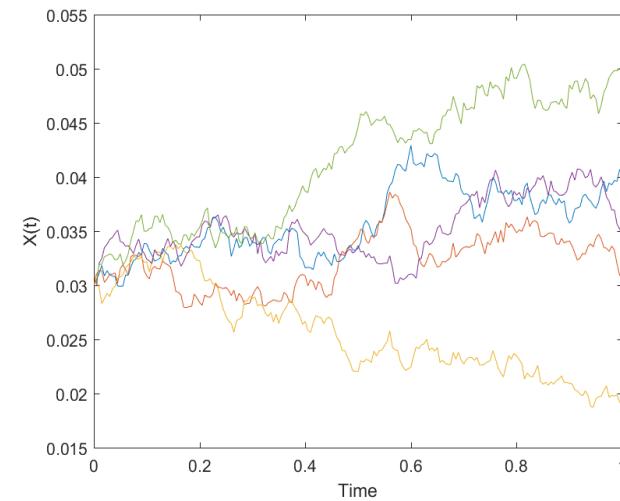


Figure 5.20: Simulated paths of the CIR model sampling from the non-central chi-square distribution. Parameters: $\alpha = 0.01$, $\sigma = 0.05$, $r(t) = 0.03$, $\mu = 0.05$.

5.4. THE CONSTANT ELASTICITY OF VARIANCE (CEV) MODEL

- The CEV model has dynamics

$$dX(t) = \mu X(t)dt + \sigma X^{\beta+1}(t)dW(t). \quad (5.7)$$

- We define the local volatility of the process as

$$\frac{SDev_t(dX(t))}{X(t)} = \sigma X(t)^\beta.$$

- For $\beta > 0$ ($\beta < 0$), the local volatility monotonically increases (decreases) as the asset price increases.
- Therefore, the so called leverage effect, i.e. the inverse relationship between spot price and volatility, can be recovered by taking $\beta < 0$.
- For example, Rubinstein and Jackwerth (1996) find that typical values of the CEV elasticity implicit in the S&P 500 stock index option prices are strongly negative and are as low as $\beta = -4$. They term the corresponding model unrestricted CEV.
- The unrestricted CEV process is used to model the volatility smile effect in the equity index options market.
- The CEV name is due to the fact if we define the elasticity of a function $f(x)$ as

$$E_f(x) = \frac{\partial f(x)}{\partial x} \frac{x}{f(x)},$$

and we set (with a little abuse of notation)

$$f(x) = \mathbb{V}ar_t \left(\frac{dX}{X} \right) = \frac{\sigma^2 X^{2\beta+2}}{X^2} = \sigma^2 X^{2\beta},$$

then

$$E_f(x) = \sigma^2 2\beta X^{2\beta-1} \frac{X}{\sigma^2 X^{2\beta}} = 2\beta,$$

in other words, it does not depend on X , i.e. we have a model for which the elasticity of variance is constant. Henceforth, the name CEV model.

- The CEV model admits as particular cases:
 - The OU Gaussian process when $\beta = -1$;
 - The GBM process when $\beta = 0$, i.e. the elasticity is zero;
 - The square-root (of the CIR model) process when $\beta = -1/2$.
- The CEV SDE does not admit a closed-form solution, however the transition probability density of reaching X_Δ after a time frame of length Δ given that we start at X_0 is

$$p(X_0, X_\Delta; \Delta) := e^{-\mu\Delta} p_0 \left(X_0, e^{-\mu\Delta} X_\Delta; \frac{1}{2\mu\beta} (e^{2\mu\beta\Delta} - 1) \right),$$

with

$$p_0(X_0, X_\Delta; \Delta) = \frac{X_\Delta^{-2\beta-\frac{3}{2}} X_0^{\frac{1}{2}}}{\sigma^2 |\beta| \Delta} e^{-\frac{X_0^{-2\beta} + X_\Delta^{-2\beta}}{2\sigma^2 \beta^2 \Delta}} I_{\frac{1}{2|\beta|}} \left(\frac{X_0^{-\beta} X_\Delta^{-\beta}}{\sigma^2 \beta^2 \Delta} \right),$$

where I_ν is the modified Bessel function of the first kind of order ν .

- The relevant references for the derivation of CEV model are Cox (1975) (when $\beta < 0$), Cox and Ross (1976), Emanuel and MacBeth (1982) (when $\beta > 0$), Schroder (1989) (for the computation of the cumulative distribution), and Delbaen and Shirakawa (2002) (when $-1 < \beta < 0$).
- The derivation of the above density, is based on the following procedure

1. Let us define

$$Z = \frac{X^{-2\beta}}{4\beta^2}$$

and let us apply the Ito's lemma to get the dynamics of Z .

2. We find

$$dZ = \alpha(\theta - Z(t)) + \hat{\sigma}\sqrt{Z}dW(t),$$

where

$$\alpha = 2\beta\mu; \quad \theta = \frac{\sigma^2(2\beta+1)}{4\alpha\beta}; \quad \hat{\sigma} = -\text{sgn}(\beta)\sigma,$$

so $Z(t)$ follows a CIR dynamics.

3. Given that Z is a CIR process, $Z(t)/k$ has a non-central $\chi_{d,\lambda}^2$ distribution, for

$$d = \frac{4(2\beta\mu)}{4\sigma^2\beta^2} \frac{\sigma^2(2\beta+1)}{2\mu} = \frac{2\beta+1}{\beta}, \quad \lambda = \frac{4Z(0)}{\tau(t)}, \quad k = \frac{4\sigma^2\beta^2\tau(t)e^{-2\beta\mu t}}{4} = \sigma^2\beta^2\tau(t)e^{-2\beta\mu t},$$

and

$$\tau(t) = \frac{e^{2\beta\mu t} - 1}{2\beta\mu}.$$

4. The distribution of $X(T)$ therefore can be obtained from the following argument (we assume $\beta > 0$, but a similar argument holds for $\beta < 0$)

$$Pr(X(T) \leq x) = Pr\left(\frac{X(T)^{-2\beta}}{4\beta^2} \leq \frac{x^{-2\beta}}{4\beta^2}\right) = Pr\left(Z(T)/k \leq \frac{x^{-2\beta}}{4k\beta^2}\right),$$

where last term is the CDF of a non-central chi-square distribution.

5. Finally, taking the derivative with respect to x and some algebra we obtain the density of the CEV process.

- When $\beta < 0$, there is a positive probability for the process that starts at X to hit zero at time t

$$1 - \int_0^\infty p(X_0, \xi; \Delta) d\xi.$$

- In this case, Andersen and Andreasen (2000) propose a modified version of the CEV model that never hits 0. Let ε a small number. If

$$S > \varepsilon,$$

the stock-price dynamics remains as in the original CEV model. If

$$S \leq \varepsilon,$$

the stock price dynamics is modified into

$$dX(t) = \mu X(t)dt + \sigma \varepsilon^{\beta+1}(t)dW(t).$$

The modified CEV process converges to the true CEV process as $\varepsilon \rightarrow 0$.

- Similarly, when $\beta > 0$, the price process can go to infinity. In this case, we can modify the process into

$$dX(t) = \mu X(t)dt + \sigma \eta^{\beta+1}(t)dW(t).$$

when $X > \eta$ and η is a large number.

- The regularizations are important to make the process a true martingale (rather than just a local martingale) when $\mu = 0$. However, the density representation is no longer valid.
- A Matlab script p_cev for computing the probability density function of the **spot price** is also available. A plot of the density of the spot price can be obtained by using the commands

```
%CEV PARAMETERS
Param.dt=1; Param.q=0;
Param.rf=0.03; Param.sigma=0.2;
Param.beta=-2;%cev 1.5
fplot(@(ST) p_cev(0.9,ST,Param), [0, 2])
title('CEV density when \beta=-2')
xlabel('Spot price')
```

- In figure 5.21 we plot the density function of the spot price for different values of β .
- The density of log-returns for different values of β (negative, zero and positive) is plotted in figure 5.22. We can see how the parameter β affects the shape of the log-price density.
- These figures are obtained running the following Matlab code in the command window

```
%CEV PARAMETERS
Param.dt=1; Param.q=0;
Param.rf=0.03; Param.sigma=0.2;

%Compute the density of ST for different values of beta
%Build the ST grid
x0=1; xmin=0.01; xmax=4; npoints=400;
%price grid
xTgrid=linspace(xmin,xmax,npoints)';
%return grid
logxTgrid=log(xTgrid);
%beta values
beta=[-2:1:2];

i=1;%counter
for betai=beta
    Param.beta=betai;
    %price density
    cev_dens_pr (:,i)=p_cev(x0,xTgrid,Param)';
    %log-return density
    cev_dens_ret (:,i)=p_cev(x0,exp(logxTgrid),Param)'.*exp(logxTgrid);
    %to build the plot legend
    legendInfo{i} = ['\beta = ' num2str(Param.beta)];
    i=i+1;
end

%Plot the stock price densities
```

```

h=figure('Color',[1 1 1]);
plot(xTgrid,cev_dens_pr); xlim([0 2])
xlabel('Spot Price'); legend(legendInfo)
title('CEV density function of S(T) varying \beta')
print(h, '-dpng','Fig_dens_cev')

h=figure('Color',[1 1 1]);
plot(logxTgrid,cev_dens_ret)
legend(legendInfo); xlim([-1.5 1.5])
title('CEV density function of log(S(T)) varyng \beta')
xlabel('log-returns')
print(h, '-dpng','Fig_denslogret_cev')

```

- Using the same code of the spot price density, we can also compute, via numerical integration using the quadgk function, the expected value, variance, skewness and kurtosis of the **log-price** assigning to the coefficient β the values -4:0.5:4.
- Numerical results are provided in Table 5.7. Notice that for $\beta = 0$, we are back to the Gaussian case (lognormal case for prices), so that the skewness is 0, the kurtosis is 3, the standard deviation 0.2 (the assigned volatility parameter) and the mean 0.01, i.e. $\log(X_0) + r - \sigma^2/2 = 0.03 - 0.2 * 0.2/2$.
- The figures in Table 5.7 are obtained by running the following Matlab code

```

beta=[-2:0.5:2]; STmax=8; S0=1;
Param.dt=1; Param.q=0;
Param.rf=0.03; Param.sigma=0.2;
Tolerance=10^(-9);
for j=1:length(beta)
    Param.beta=beta(j);%cev parameter
    arealogp(j)=quadgk(@(ST) p_cev(S0,ST,Param),0,STmax,'AbsTol',Tolerance);

    exlogp(j)=quadgk(@(ST) log(ST).*p_cev(S0,ST,Param),0,STmax,'AbsTol',Tolerance);

    sdlogp(j)=quadgk(@(ST) ((log(ST)-exlogp(j)).^2).*p_cev(S0,ST,Param),...

```

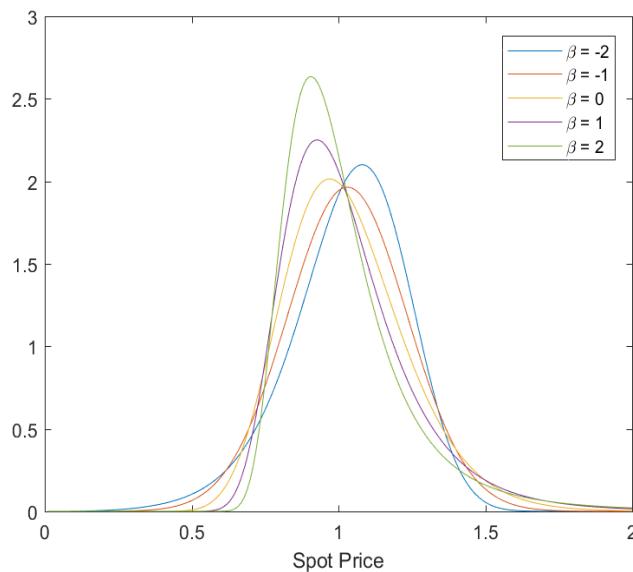


Figure 5.21: Density function of the CEV model for different values of β and $\sigma = 0.2, \mu = 0.03$.

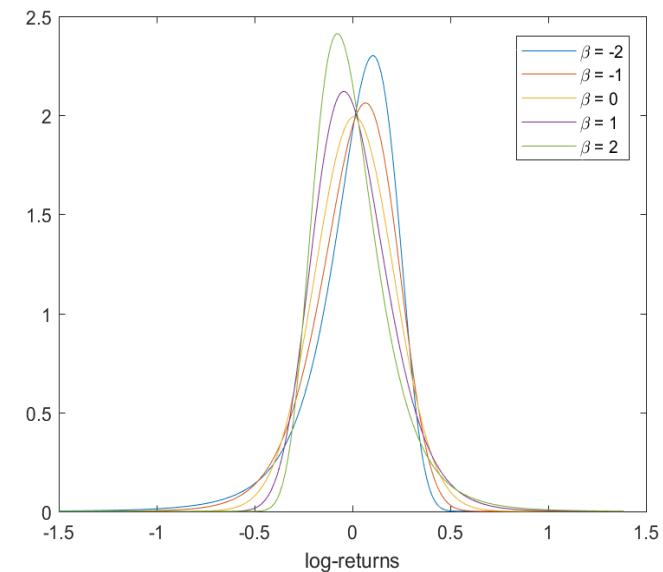


Figure 5.22: Density function of log-returns in the CEV model for different values of β and $\sigma = 0.2, \mu = 0.03$.

```

0, STmax, 'AbsTol', Tolerance)^0.5;

skewlogp(j)=quadgk(@(ST) ((log(ST)-exlogp(j)).^3).*p_cev(S0,ST,Param), ...
0, STmax, 'AbsTol', Tolerance)/sdlogp(j)^3;

kurtlogp(j)=quadgk(@(ST) ((log(ST)-exlogp(j)).^4).*p_cev(S0,ST,Param), ...
0, STmax, 'AbsTol', Tolerance)/sdlogp(j)^4;
end

output=[beta' arealogp' exlogp' sdlogp' skewlogp' kurtlogp']

```

β	Exp. Val.	Stand. Dev.	Skewness	Kurtosis
-2	0.0117	0.2188	-1.5554	8.2524
-1.5	0.0083	0.2188	-1.3355	7.8188
-1	0.0093	0.2084	-0.7168	4.3431
-0.5	0.0099	0.2026	-0.3128	3.2222
0	0.0100	0.2000	0.0000	3.0000
0.5	0.0097	0.1994	0.3031	3.2075
1.0000	0.0089	0.2006	0.6470	4.0076
1.5000	0.0072	0.2025	1.0472	5.6575
2.0000	0.0028	0.1988	1.2719	6.4735

Table 5.7: Moments of log-returns in the CEV model varying the leverage parameter β .

- In figure 5.23 we plot the implied volatility curve. This plot has been obtained pricing European option via numerical quadrature (quadgk in Matlab) and then using the model price to compute the Black-Scholes implied volatility. In particular, large negative values of β generate very steep implied volatility curve, as often observed in the option market.
- Table (5.8) gives the main properties of the CEV model.

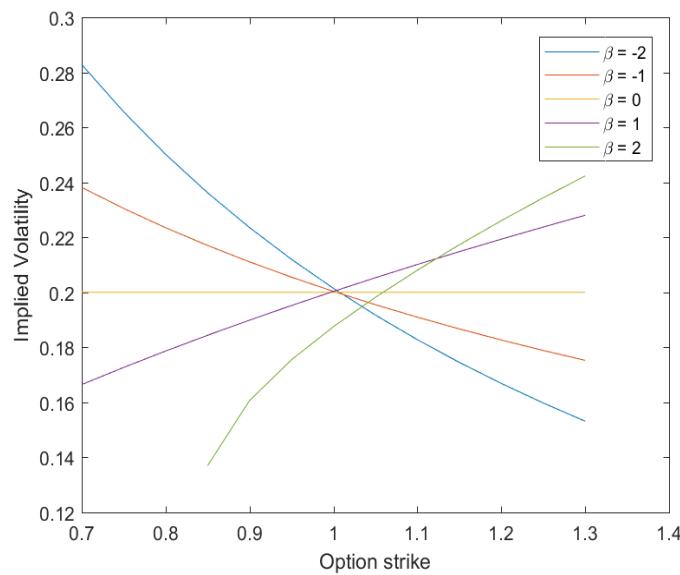


Figure 5.23: Implied Volatility in the CEV model for different values of β .

CONSTANT ELASTICITY OF VARIANCE CEV(μ, σ, β): FACTS

The SDE

$$dX(t) = \mu X(t)dt + \sigma X^{\beta+1}(t)dW(t), X(0) = x_0.$$

The solution of the SDE

is not explicit.

The distribution of

$$\frac{X(t)^{-2\beta}}{4\beta^2} \sim k \times \chi_{d,\lambda}^2.$$

$$d = \frac{2\beta+1}{\beta}, \lambda = \frac{4Z(0)}{\tau(t)}, k = \sigma^2 \beta^2 \tau(t) e^{-2\beta\mu t}, \tau(t) = \frac{e^{2\beta\mu t}-1}{2\beta\mu}.$$

The mean of $X(t)$

$$\mathbb{E}_0(X(t)) = X(0) e^{\mu t}.$$

The moments of $X(t)$

not available as simple expressions

If $\beta < 0$

there is a positive probability of absorption in 0.

If $\beta > 0$

there is a positive probability of explosion.

Skewness of $\log(X(t))$

negative, if $\beta < 0$.

positive, if $\beta > 0$.

Excess of kurtosis of $\log(X(t))$

If $\beta \neq 0$ is positive

zero, if $\beta = 0$

The CEV model admits as particular cases

The O-U Gaussian process if $\beta = -1$

The GBM process if $\beta = 0$

The square-root process if $\beta = -1/2$

The stationary distribution of $X(t), (t \rightarrow \infty)$.

does not exist

Table 5.8: Properties of the Constant Elasticity of Variance CEV(μ, σ, β)

5.5. CIR, CEV AND THE BESSEL PROCESS

The unique strong solution to the following SDE

$$d\rho(t) = \delta dt + 2\sqrt{\rho(t)}dW(t), \rho(0) = \rho_0$$

is called a squared δ -dimensional Bessel process starting at ρ_0 and is denoted by $\text{BESQ}_{\rho_0}^{\delta}$.

- δ is the dimension of the Bessel process and it takes real values.
- If δ is an integer, the process $\rho(t)$ can be represented by the square of the Euclidean norm of δ -dimensional Brownian motion

$$\rho(t) = \sum_{i=1}^{\delta} W_i(t)^2.$$

- $\nu = \delta/2 - 1$ is the index of the process BESQ^{δ} . For $\delta \geq 2$, the BESQ^{δ} process will never reach 0 for $t > 0$, and for $0 \leq \delta < 2$ the process reaches 0 almost surely.
- The square root of $\text{BESQ}_{y^2}^{\delta}$, $\delta \geq 0$, $y \geq 0$ is called the Bessel process of dimension δ starting at y and is denoted by BES_y^{δ} .
- A CIR process X can be represented as a time-changed BESQ in the following way

$$X(t) = e^{-\alpha t} \rho(\tau(t)), \quad \alpha > 0$$

where the time change is given by

$$\tau(t) = \sigma^2 \frac{e^{\alpha t} - 1}{4\alpha}, \quad \tau(0) = 0.$$

- Indeed, we can proceed as follows

1. Solving the BESQ SDE, we have

$$\rho(\tau(t)) = \rho(0) + \delta\tau(t) + 2 \int_0^{\tau(t)} \sqrt{\rho(s)} dW(s).$$

2. We observe that the stochastic integral is a martingale with quadratic variation

$$4 \int_0^{\tau(t)} \rho(s) ds = 4 \int_0^t \rho(\tau(s)) \tau'(s) ds$$

for

$$\tau'(s) = \sigma^2 \frac{e^{\alpha s}}{4}.$$

3. We can rewrite the solution as

$$\rho(\tau(t)) = \rho(0) + \delta\tau(t) + 2 \int_0^t \sqrt{\rho(\tau(s)) \tau'(s)} dW(s),$$

and therefore

$$d\rho(\tau(t)) = \delta\tau'(t) dt + 2\sqrt{\rho(\tau(t)) \tau'(t)} dW(t).$$

4. Set $Y(t) = \rho(\tau(t))$ and apply Ito's lemma to

$$X(t) = e^{-\alpha t} Y(t),$$

so that we have

$$dX(t) = -\alpha X(t) dt + e^{-\alpha t} dY(t),$$

from which it follows

$$dX(t) = -\alpha X(t) dt + e^{-\alpha t} \delta\tau'(t) dt + 2e^{-\alpha t} \sqrt{Y(t) \tau'(t)} dW(t).$$

5. After some algebra, we obtain

$$dX(t) = \theta (\theta - X(t)) dt + \sigma \sqrt{X(t)} dW(t),$$

where

$$\theta = \delta \frac{\sigma^2}{4\alpha}.$$

6. The BESQ process will never reach 0 if $\delta \geq 2$. Similarly, the CIR process will never reach 0 if

$$2\alpha\theta \geq \sigma^2,$$

i.e. the Feller condition previously presented (see Fact 46).

- A CEV process can be transformed in a CIR process. Therefore, it is also related to a time-changed Bessel process.
- Bessel processes are also important to study the integral of a geometric Brownian motion, quantity that arises in the pricing of Asian options.
- A detailed study of the Bessel process can be found in Goings-Jaeschke and Yor (2003) and in Jeanblanc et al. (2009, chapter 4).

5.6. THE BROWNIAN BRIDGE

Fact 47 (Brownian Bridge) Let $W(t)$ be a Brownian motion. Fix $s > 0$ and $T > 0$ with $s < T$, $a \in \mathbb{R}$ and $b \in \mathbb{R}$. We define the Brownian bridge from a to b on $[s, T]$ to be the process $X(t)$ satisfying the SDE

$$dX(t) = \frac{b - X(t)}{T - t} dt + dW(t) \quad (5.8)$$

with initial condition at time s , $X(s) = a$.

5.6.1. A NOTE: THE ORDINARY DIFFERENTIAL EQUATION $dx(t) = (b - x(t))/(T - t)dt$.

- We would like to solve the SDE of the Brownian bridge. Let us start considering the deterministic version.

$$dx(t) = \frac{b - x(t)}{T - t} dt$$

- We proceed through the following steps

- We let $y(t) = g(t, x) = (T - s)/(T - t)x(t)$.
- Then, $dy(t) = b(T - s)/(T - t)$.
- Finally $y(t) = y(s) + b(t - s)/(T - t)$.

We can conclude

Fact 48 *The ODE $dx(t) = (b - x(t))/(T - t)dt$ admits solution*

$$x(t) = a + (b - a)\frac{t-s}{T-s}.$$

5.6.2. SOLVING THE SDE $dX_t = (b - X(t))/(T - t)dt + dW(t)$

- We need to solve

$$dX(t) = \frac{b - X(t)}{T - t}dt + dW(t)$$

- By analogy with the previous ODE, let us define

$$Y(t) = g(t, X(t)) = \frac{T-s}{T-t}X(t),$$

and apply Itô's Lemma.

- Then

$$\begin{aligned} dY(t) &= \left(X(t) \frac{T-s}{(T-t)^2} + \frac{T-s}{T-t} \frac{b - X(t)}{T-t} \right) dt + \frac{T-s}{T-t} dW(t) \\ &= b \frac{T-s}{(T-t)^2} dt + \frac{T-s}{T-t} dW(t). \end{aligned}$$

- Therefore

$$Y(t) = a + b \frac{t-s}{T-t} + (T-s) \int_s^t \frac{1}{T-u} dW(u).$$

Fact 49 *The solution of the SDE 5.8 is*

$$X(t) = a + \frac{t-s}{T-s} (b-a) + (T-t) \int_s^t \frac{1}{T-u} dW(u).$$

Further

$$\begin{aligned} X(t) &\sim \mathcal{N}(\mathbb{E}_0(X(t)), \text{Var}_0(X(t))), \\ \mathbb{E}_0(X(t)) &= a + \frac{t-s}{T-s} (b-a), \\ \text{Var}_0(X(t)) &= \frac{(t-s)(T-t)}{T-s}, \\ c_X(t, z) &= \frac{(t \wedge z - s)(T - t \vee z)}{T-s}. \end{aligned}$$

- The function $a + \frac{t-s}{T-s} (b-a)$, as a function of t , is the line from (s, a) to (T, b) .
- To this line, we add the Brownian bridge from 0 to 0 on $[s, T]$.
- This generates a process that starts in a at time s and terminates in b at time T .

Remark 50 *Notice that the process*

$$X(t) = a + \frac{t-s}{T-s} (b-a) + (T-t) \int_s^t \frac{1}{T-u} dW(u)$$

is equivalent (e.g. they have the same distribution, mean, variance and (auto)-covariance) to

$$X(t) = a + \frac{t-s}{T-s} (b-a) + (W_t - W_s) - \frac{t-s}{T-s} (W_T - W_s).$$

5.6.3. MATLAB IMPLEMENTATION: SIMULATING BROWNIAN MOTIONS (PART 2)

The primary use for the Brownian bridge in finance is as an aid to Monte Carlo simulation, as the Brownian bridge $X(t)$ represents a Brownian motion on the time interval $[s, T]$, starting at $W(s) = a$ and conditioned to arrive at b at time T .

- To see this, consider a time partition such that $t_i < t_j < t_k$.
- Let

$$\begin{aligned} X &= W(t_j) - W(t_i); \\ Y &= W(t_k) - W(t_j); \\ Z &= W(t_k) - W(t_i) = X + Y. \end{aligned}$$

- Then X and Y are independent; moreover, $X \sim \mathcal{N}(0, \sigma_X^2)$, $Y \sim \mathcal{N}(0, \sigma_Y^2)$ and $Z \sim \mathcal{N}(0, \sigma_Z^2)$, where $\sigma_X^2 = t_j - t_i$, $\sigma_Y^2 = t_k - t_j$, and $\sigma_Z^2 = t_k - t_i = \sigma_X^2 + \sigma_Y^2$.
- Therefore, the conditional density of X given Y is

$$\begin{aligned} f_{X|Z}(x) &= \frac{f_X(x)f_Y(y)}{f_Z(z)} \\ &= \frac{1}{B\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-Az}{B}\right)^2}, \end{aligned}$$

where $A = \sigma_X^2/\sigma_Z^2$ and $B = \sigma_X\sigma_Y/\sigma_Z$.

- Hence, we can claim that, conditioning on the knowledge of the process value at time t_k , $W(t_j) - W(t_i) \sim N(Az, B^2)$.
- From this, it follows that

$$W(t_j) = \frac{t_k - t_j}{t_k - t_i} W(t_i) + \frac{t_j - t_i}{t_k - t_i} W(t_k) + \sqrt{\frac{(t_k - t_j)(t_j - t_i)}{t_k - t_i}} \varepsilon, \quad \varepsilon \sim N(0, 1). \quad (5.9)$$

- But this is the Brownian bridge from $W(t_i)$ to $W(t_k)$ on $[t_i, t_k]$.

Hence, we can simulate the value of the Brownian motion at each time step over $[0, T]$ by using the Brownian bridge according to the following steps.

- Simulate first the value of Brownian motion at time T . Set b equal to this value and set $a = 0$.
- Simulate $W(t_1)$, using 5.9.
- Set $a = W(t_1)$ and repeat for all $t_j \in (0, T)$.

Resulting sample trajectories are shown in Figure (5.24) (see also accompanying Matlab code).

The Brownian bridge can support useful variance reduction techniques for pricing derivative contracts using Monte Carlo simulation, such as stratified sampling. The Brownian bridge, in fact, allows us to pre-specify the value that the Brownian motion needs to achieve at a given point in time. This allows us to ‘bucket’ the value of the Brownian motion within a chosen stratum. For further details, we refer the interested reader to Glasserman (2004).

Because of the above mentioned property, the Brownian bridge finds useful application in the context of scenario generation, when for example a target for stress testing is given by the regulators. An example is shown in Figure 5.25, where we generate a number of trajectories for a stock price, starting at 100 and ending in one year time either at 150 or 50.

Matlab Code

```
%%%%%%SIMULATING THE BROWNIAN BRIDGE %%%%%%
%%%%%Assigning the number of simulated paths
%(nsimul), time to maturity (expiry), number of steps
%(nsteps), time step (dt) and observation times (timestep):
clear all;
nsimul=10000, expiry=1, nsteps=250;
dt=expiry/nsteps;
timestep=[0:dt:expiry]';
Wt=zeros(nsteps+1,nsimul);
%Simulate the Brownian motion at T:
eY = randn(1,nsimul);
Wt(nsteps+1,:)= sqrt(expiry).*eY;
%Simulate the Brownian motion W(t):
for j=2:nsteps
    deltat1=(nsteps+1-j) / (nsteps+1-j+1);
    eYt = randn(1,nsimul);
    Wt (j,:)=deltat1*Wt (j-1,:)+(1-deltat1)*Wt (nsteps+1,:)+sqrt (de
end
Bb=Wt;
%Plot simulated paths:
h=figure('Color',[1 1 1])
plot(timestep, Bb)
title('Simulated Paths of the Brownian motion via Brownian Bridge')
xlabel('Time (years)')
```

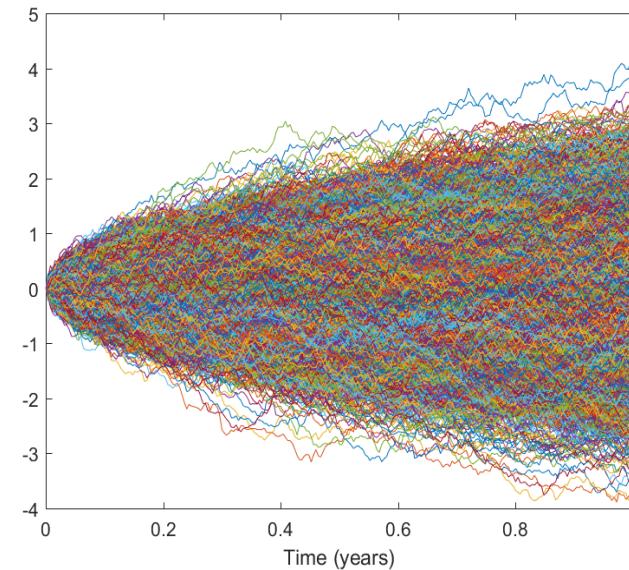


Figure 5.24: Simulated Paths of the Brownian motion via Brownian Bridge.

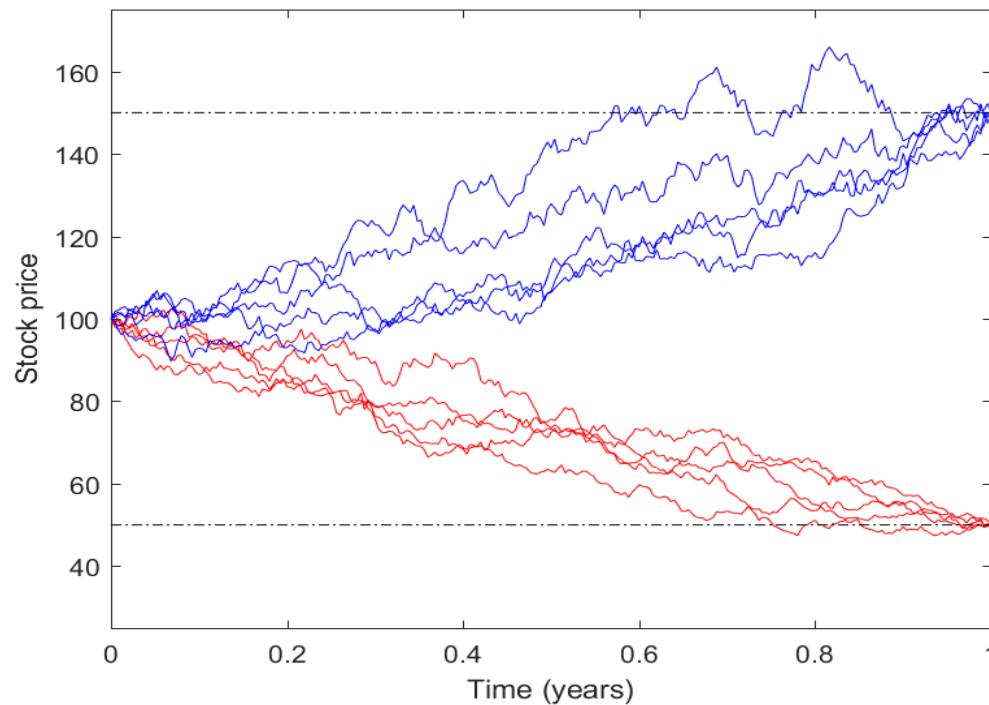


Figure 5.25: Scenario Generation via Brownian bridge. Target stock price: $(150, 50)$; $\mu = 0.1$; $\sigma = 0.2$; time horizon: 1 year; daily monitoring.

BROWNIAN BRIDGE BB($a, b; s, T$): FACTS

The SDE

$$dX(t) = \frac{b-X(t)}{T-t} dt + dW(t); \quad X(s) = a.$$

The solution of the SDE

$$X(t) = a + \frac{t-s}{T-s} (b-a) + (T-t) \int_s^t \frac{1}{T-u} dW(u).$$

The distribution of $X(t)$

$$X(t) \sim \mathcal{N}(\mathbb{E}_0(X(t)), \text{Var}_0(X(t))).$$

The mean of $X(t)$

$$\mathbb{E}_0(X(t)) = a + \frac{t-s}{T-s} (b-a).$$

The variance of $X(t)$

$$\text{Var}_0(X(t)) = \frac{(t-s)(T-t)}{T-s}.$$

The autocovariance of $X(t)$

$$c_X(t, z) = \frac{(t \wedge z - s)(T - t \vee z)}{T-s}.$$

5.7. THE HESTON STOCHASTIC VOLATILITY MODEL

- Stochastic volatility models are widely used in investment banks and financial institutions.
- The Heston model is sufficiently rich for explaining the volatility smile in the option market.
- The dynamics of the log-price $s(t) = \log S(t)$ is

$$ds(t) = \left(\mu - \frac{1}{2}v(t) \right) dt + \sqrt{v(t)} dW_s(t), \quad (5.10)$$

and it is accompanied by a second sde describing the stochastic dynamics of the instantaneous variance

$$dv(t) = k(\theta - v(t))dt + \varepsilon\sqrt{v(t)}dW_v(t), \quad (5.11)$$

with starting values $s(0)$ for the log-price and $v(0)$ for the variance.

- Hence, the Heston model is represented by a bivariate system of SDEs: the first one describing the dynamics of the log-price, the second one the dynamics of the instantaneous variance of the log-price.
- The correlation between the Brownian motions entering the dynamics of the two state variables is

$$\mathbb{E}(dW_s(t)dW_v(t)) = \rho dt.$$

- The model parameters are:
 - $v(0) := v_0$: the initial (time zero) level of the variance;
 - μ : the instantaneous rate of return on the stock. Indeed over a short period we can write $\mathbb{E}_t(dS)/S = \mu dt$;
 - k : the mean-reversion speed for the variance;
 - θ : the long-run level for the variance;

- ε : the volatility of variance
- ρ : the correlation coefficient between the underlying and volatility

- The parameters restrictions are

$$S_0, v_0, k, \theta, \varepsilon > 0, \quad \rho \in [-1, 1]$$

- Notice that the variance process is the same as the mean-reverting CIR dynamics, so it shares the same main properties.
- In particular,

1. the distribution of the variance at a given horizon is related to a non-central chi-square distribution, and it admits a stationary distribution;
2. the variance process is always positive and cannot reach 0, if the Feller condition is satisfied, i.e.

$$\varepsilon^2 \leq 2\kappa\theta.$$

- We also observe that the Heston model can alternatively be written as

$$\begin{aligned} s(t) &= \mu t - \frac{1}{2} \int_0^t v(s) ds + \int_0^t \sqrt{v(s)} dW_s(s), \\ v(t) &= k\theta t - k \int_0^t v(s) ds + \varepsilon \int_0^t \sqrt{v(s)} dW_v(s). \end{aligned}$$

In virtue of the Dambis, Dubins-Schwarz Theorem in Section 3.6, the above can be equivalently written also as

$$\begin{aligned} s(t) &= \mu t - \frac{1}{2} T(t) + W_s(T(t)), \\ v(t) &= k\theta t - kT(t) + \varepsilon W_v(T(t)), \\ T(t) &= \int_0^t v(s) ds. \end{aligned}$$

In other words, the Heston model is given by a system of time-changed ABMs.

5.7.1. THE CHARACTERISTIC FUNCTION OF THE LOG-PRICE

- The Heston model belongs the class of affine processes (AD) as in Duffie, Pan and Singleton (2000): drift and covariances are linear in the state variables, i.e. log-price and variance, (up to a constant).
- In general, in affine models, the density function is not known in closed form, but it can be recovered through a numerical inversion of the characteristic function ψ^H

$$f_H(s(T); s(t), v(t), \tau) = \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[e^{-ius(T)} \psi^H(u; s(t), v(t), \tau) \right] du, \quad (5.12)$$

- In particular, the Heston's characteristic function is available in closed form:

$$\psi^H(u) := \psi^H(u; s(t), v(t), T-t) = \mathbb{E}_t \left[e^{ius(T)} \right] = e^{C(u, T-t) + D(u, T-t)v_t + ius(t)}, \quad (5.13)$$

where the functions C and D are given by in terms of $\tau = T - t$

$$\begin{aligned} C(u, \tau) &= iur\tau + \frac{k\theta}{\varepsilon^2} \left((k - i\rho\varepsilon u - d)\tau - 2 \ln \frac{1 - ge^{-d\tau}}{1 - g} \right), \\ D(u, \tau) &= \frac{1}{\varepsilon^2} (k - i\rho\varepsilon u - d) \frac{1 - e^{-d\tau}}{1 - ge^{-d\tau}}, \end{aligned}$$

with $i = \sqrt{-1}$, $\tau = T - t$ and

$$d = \sqrt{(i\rho\varepsilon u - k)^2 + \varepsilon^2(iu + u^2)}, \quad g = \frac{k - i\rho\varepsilon u - d}{k - i\rho\varepsilon u + d}.$$

- The moments of the log-price can be found by taking the derivative of the characteristic function with respect to u and then setting u equal to 0

$$\mathbb{E}_0(s_T^n) = (-i)^n \left. \frac{\partial^n \psi^H(u)}{\partial u^n} \right|_{u=0}.$$

- The first moment of the log-price in the Heston model is given by

$$\mathbb{E}_0(s_T) = s_0 + \left(\mu - \frac{\theta}{2} \right) \tau + (\theta - v_0) \frac{1 - e^{-\kappa\tau}}{2\kappa},$$

and for large horizons it grows linearly with time.

- The expressions for the higher order moments are very lengthy and not reported here. However, they can be computed using symbolic calculus and the results are available using the implemented Matlab function momHeston.
- This function computes the moments also resorting to the Cauchy integral formula, a result not very well known in the financial community. In practice, this formula allows the computation of a derivative via an...integral!. In fact, we have

$$\mathbb{E}_0(x_T^n) = \frac{n!}{2\pi i^n} \int_0^{2\pi} \frac{\psi^H(e^{iu})}{e^{iun}} du. \quad (5.14)$$

This formula is implemented using the quadgk Matlab function in the function momHeston.

- Concerning the moments of the spot price, their existence is not always guaranteed. They can become infinite at a finite horizon. This is known as moment explosion problem. Conditions under which all the moments of the spot price exist finite are given in Andersen and Piterbarg (2007).

Matlab Code Here we provide the Matlab code for implementing the Heston characteristic function.

```
function cf= cfHeston(u, kappa, theta, sigma, rho, tau, r, s0, v0)
% Heston parameters:
%   kappa = variance mean reversion speed parameter
%   theta = variance long-run level parameter
%   rho   = correlation between two Brownian motions
%   sigma = volatility of variance
%   v0    = initial variance
%   s0    = initial stock price
```

```
% Log of the stock price.
x = log(S0);

% Parameter transformation
a = kappa*theta;
sg2=sigma^2;
d = sqrt((rho*sigma*li*u - kappa).^2 - sg2*(li*u - u.^2));
g = (kappa - rho*sigma*li*u - d)./(kappa - rho*sigma*li*u + d);

% "Little Heston Trap" formulation
D1 = (kappa - rho*sigma*li*u - d)/sg2;
D2 = ((1-exp(-d*tau))./(1-g.*exp(-d*tau)));
D = D1.*D2;
G = (1-g.*exp(-d*tau))./(1-g);
C = li*u*r*tau + a/sg2*((kappa - rho*sigma*li*u - d)*tau - 2*log(G));

% The characteristic function.
cf = exp(C + D*v0 + li*u*x);
```

The numerical computation of the moments is presented in the Matlab code below. Given the code length we only present the code for the computation of the moments via the Cauchy integral and the exact formula for the first moment of the log-price. The code for the remaining moments is not reported here, but can be found inside the script available with the accompanying Matlab files. In the code below, the last argument of the function allows the user to ask for mean, variance, skewness and kurtosis rather than for integer moments.

```
function [moments mcauchy]= momHeston( kappa, theta, sigma, rho, tau, r, S0, v0, central)

% Log of the stock price.
x0 = log(S0);

%Computation of the moments via Cauchy formula
cf=@(u) cfHeston(u, kappa, theta, sigma, rho, tau, r, S0,v0);
for n=1:4
```

```

nf=factorial(n);
mcauchy(n,1)=real(quadgk(@(ss) cf(exp(li*ss))./exp(li*ss*n), 0, 2*pi)*nf/(2*pi*li^n));
end

%Computation of the moments via differentiation of the cf
m1=x0+(r-theta/2)*tau+(theta-v0)*(1-exp(-kappa*tau))/(2*kappa);
m2=...; %Too long expression: see Matlab script
m3=...; %Too long expression: see Matlab script
m4=...; %Too long expression: see Matlab script

if central==0
    moments=[m1,m2,m3,m4];
elseif central==1
    sd=(m2-m1^2)^0.5;
    sk=(m3-3*m1*sd^2-m1^3)/sd^3;
    kur=(m4-4*m1*m3+6*m1^2*sd^2+3*m1^4)/sd^4;
    moments=[m1,sd,sk,kur];
...
end

```

Example For example, if in a Matlab script we write

```

%Heston Parameters
kappa=5; theta=0.05;
sigma=0.5; rho=-0.8
tau=0.5; r=0.03;
S0=1; v0=0.05;
%Compute integer moments

[moments mcauchy]= momHeston( kappa, theta, sigma, rho, tau, r, S0, v0,0)
%mean, std. dev., skewness and kurtosis
[moments mcauchy]= momHeston( kappa, theta, sigma, rho, tau, r, S0, v0,1)

```

we obtain the moments in Table 5.9.

Order	1	2	3	4
Moment	0.0025	0.026301	-0.00395	0.003175
Centered Moment	0.0025	0.158200	-0.047372	3.559633

Table 5.9: The Table returns in the second row the integer moments up to the fourth order. In the third row, we have mean, standard deviation, skewness and kurtosis.

- Figure 5.26 illustrates how the correlation coefficient ρ affects the skewness of the distribution. Approximately, positive (negative) values of ρ determine positive (negative) values of the skewness of the distribution. Mean and variance are not affected by ρ . The excess kurtosis appears, for the given parameter set, to achieve a minimum when the correlation is near 0.
- Figures 5.27 and 5.28 show the term structure of skewness and kurtosis in the Heston model. As previously said, the correlation coefficient affects the sign of skewness. Whatever the correlation coefficient, the Heston model always generates a positive excess kurtosis.
- Figures 5.29 and 5.30 show how the skewness and kurtosis change varying the vol-vol parameter ε . In particular, the larger the vol of vol, the larger the kurtosis: whatever the correlation coefficient, the Heston model always generates a positive excess kurtosis. The skewness also increases by increasing the vol-vol parameter, with a sign that depends on the correlation coefficient ρ .
- As explained by Heston (1993), positive (negative) correlation implies a rise (fall) in variance when the returns are positive. This implies a fattening of the right (left) tail of the distribution, and therefore a positive (negative) skewness.
- In Figure 5.31 we can see that the larger the volatility of the variance, the larger the kurtosis of the distribution. As explained by Heston (1993), when this parameter is large, the variance process is highly dispersed, so the return distribution has a higher kurtosis and fatter tails. The skewness is also affected, but its sign does not depend on ε .

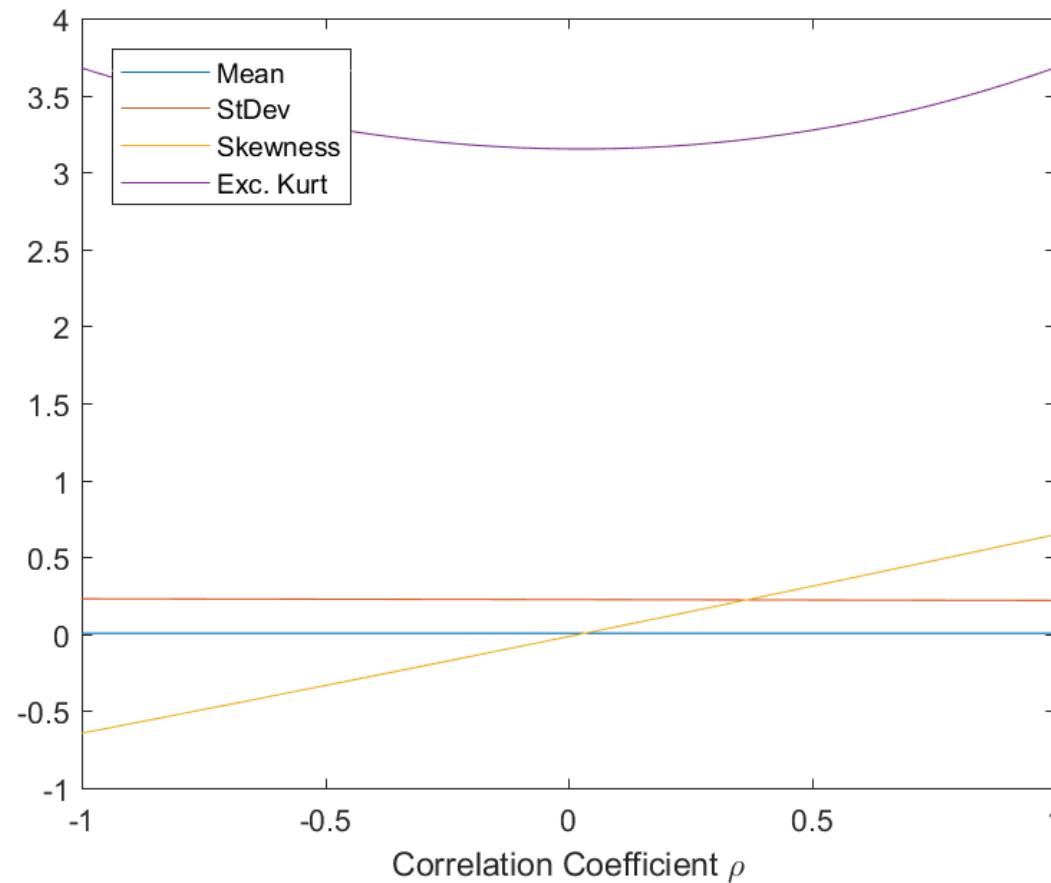


Figure 5.26: Mean, standard deviation, skewness and kurtosis of the Heston model varying the correlation coefficient ρ . Parameter set: see above Matlab script.

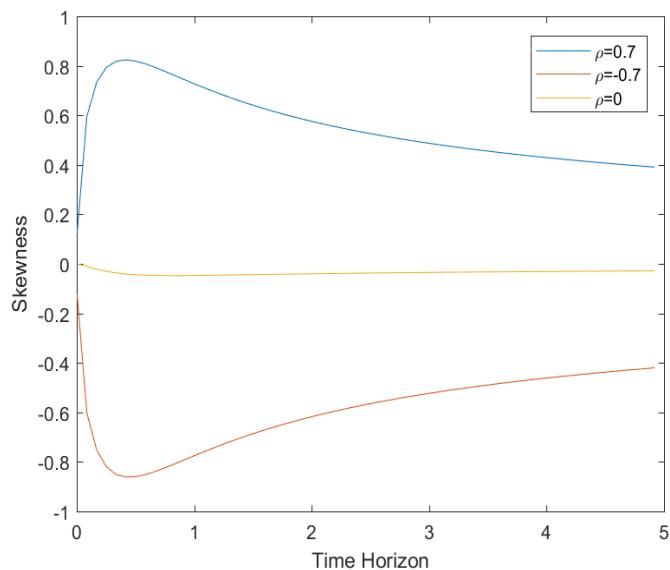


Figure 5.27: Term structure of skewness in the Heston model for different values of the correlation coefficient ρ . Other parameters: $\kappa = 5$; $\theta = 0.05$; $\sigma = 0.5$; $\mu = 0.03$; $S_0 = 1$; $v_0 = 0.05$.

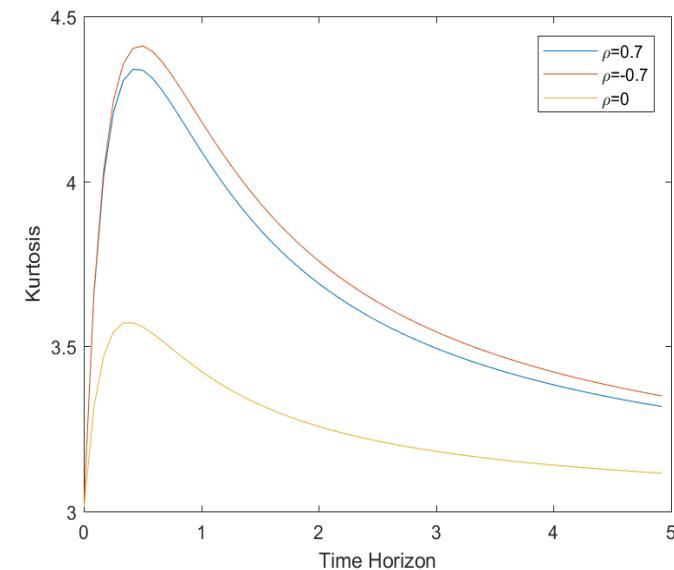


Figure 5.28: Term structure of kurtosis in the Heston model for different values of the correlation coefficient ρ . Other parameters: $\kappa = 5$; $\theta = 0.05$; $\sigma = 0.5$; $\mu = 0.03$; $S_0 = 1$; $v_0 = 0.05$.

5.7.2. HESTON MODEL: MONTE CARLO SIMULATION

The simulation of the Heston model can be performed in different ways¹.

NAIVE SAMPLE PATH GENERATION

- Divide the time interval $[0, T]$ into n equidistant dates t_1, t_2, \dots, t_n subject to spacing $\delta = T/n$.

¹We thanks Ioannis Kyriakou for help in writing this section.

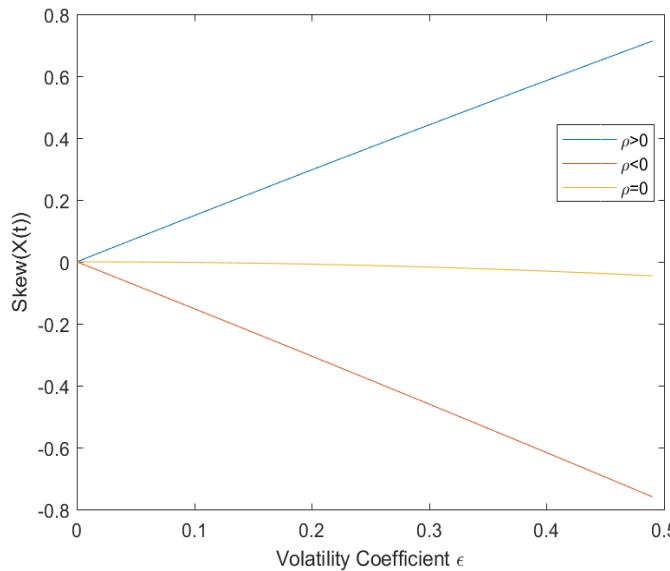


Figure 5.29: Skewness in the Heston model for different values of the vol-vol parameter ε . Other parameters: $\kappa = 5$; $\theta = 0.05$; $\mu = 0.03$; $S_0 = 1$; $v_0 = 0.05$.

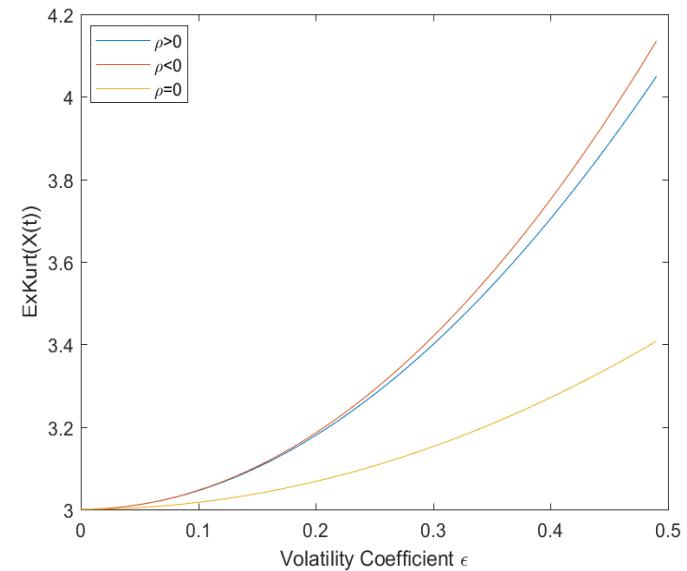


Figure 5.30: Kurtosis in the Heston model for different values of the vol-vol parameter ε . Other parameters: $\kappa = 5$; $\theta = 0.05$; $\mu = 0.03$; $S_0 = 1$; $v_0 = 0.05$.

- **Basic Euler scheme:** discretize (5.10)–(5.11),

$$\begin{aligned} s_{t+\delta} &= s_t + \left(\mu - \frac{v_t}{2} \right) \delta + \sqrt{v_t} \left(\rho \tilde{Z} + \sqrt{1-\rho^2} Z \right) \sqrt{\delta}, \\ v_{t+\delta} &= v_t + \kappa (\theta - v_t) \delta + \eta \sqrt{\varepsilon_t} \tilde{Z} \sqrt{\delta}, \end{aligned}$$

where \tilde{Z}, Z are independent standard normal variables.

- Problem: v samples can become negative with nonzero probability, making the computation of \sqrt{v} impossible and causing the time stepping scheme to fail.

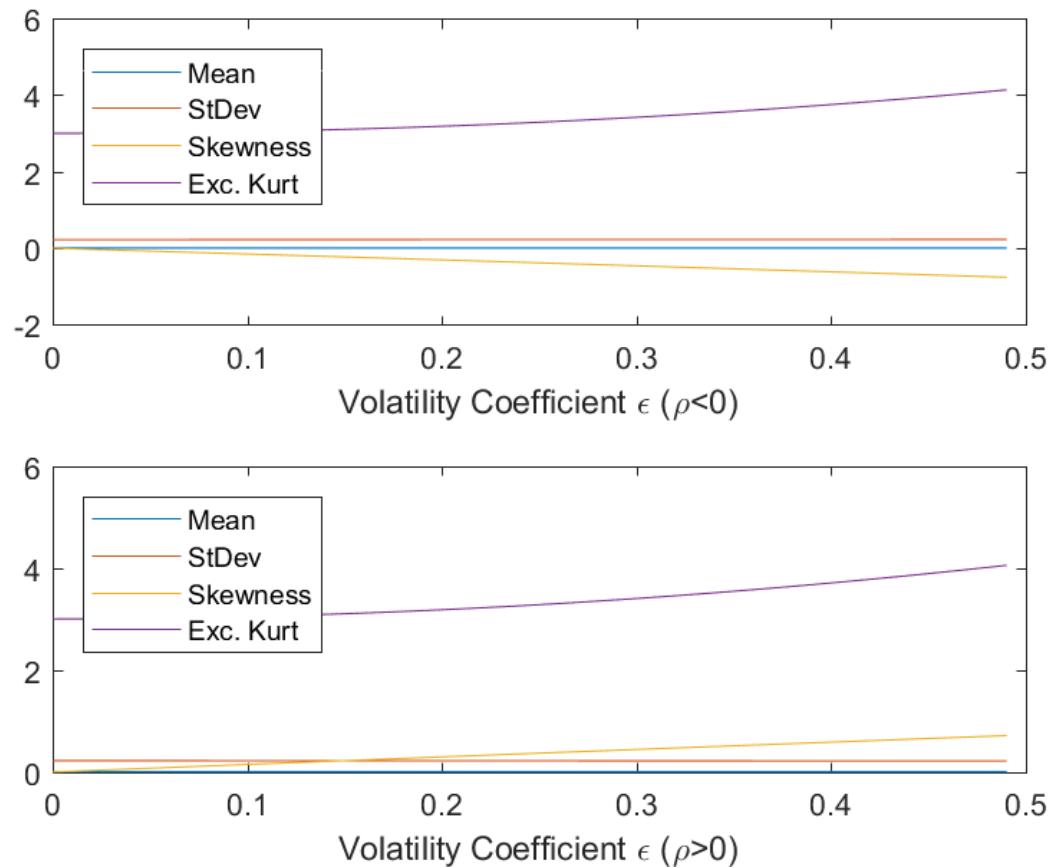


Figure 5.31: Mean, standard deviation, skewness and kurtosis of the Heston model varying the volatility coefficient ϵ . Parameter set: see above Matlab script. Top: $\rho < 0$, Bottom: $\rho > 0$.

- Several remedies have been proposed in the literature: see for example Andersen (2008) and Lord et al. (2010) for a review of various fixes and different methods. For example, $\max(v_t, 0)$.

LOW-BIAS SAMPLE PATH GENERATION

A simulation method having a low bias is as follows. Consider (5.10). Conditional on $v_{t+\delta}$, v_t and $\int_t^{t+\delta} v_s ds$, $X_{t+\delta}$ is normal:

$$\begin{aligned} X_{t+\delta} &= X_t + \left(r - \frac{\rho\alpha\beta}{\eta} \right) \delta + \left(\frac{\rho\alpha}{\eta} - \frac{1}{2} \right) \underbrace{\int_t^{t+\delta} v_s ds}_{\approx \frac{(v_t+v_{t+\delta})\delta}{2}} + \frac{\rho}{\eta} (v_{t+\delta} - v_t) \\ &\quad + \sqrt{(1 - \rho^2) \underbrace{\int_t^{t+\delta} v_s ds}_{\approx \frac{(v_t+v_{t+\delta})\delta}{2}} Z_{\text{randn}}} \end{aligned} \tag{5.15}$$

where $Z \sim \mathcal{N}(0, 1)$ is independent of v . Exact simulation of transition of variance diffusion (*scaled noncentral chi-square*) can be performed as shown in the Cox-Ingersoll-Ross method

$$v_{t+\delta} = \frac{\eta^2 (1 - e^{-\alpha\delta})}{4\alpha} \underbrace{\chi_d'^2(\lambda_{NC})}_{\text{ncx2rnd}(d, \lambda_{NC})}, \quad t > 0,$$

where $\chi_d'^2(\lambda_{NC})$ denotes a noncentral chi-square random variable with

$$d = \frac{4\alpha\beta}{\eta^2}$$

degrees of freedom and noncentrality parameter

$$\lambda_{NC} = \frac{4\alpha e^{-\alpha\delta}}{\eta^2 (1 - e^{-\alpha\delta})}.$$

This path generation method for the variance is exact (no discretization error / no bias) and ensures positive samples generated, but is more computationally demanding. Example: 1,000 samples `ncx2rnd(3, 2.2)` take 0.33 seconds versus `randn` 0.001 seconds.

A comparison between the simulation of the variance process using the Euler scheme and the enhanced exact scheme is illustrated in the movie (5.32). Notice that using the Euler scheme we can achieve negative values of the variance.

A movie illustrating the simulated log-returns and the variance in the Heston model and the corresponding distribution is shown in figures 5.33 and 5.34.

5.7.3. RECOVERING THE DENSITY FUNCTION IN THE HESTON MODEL VIA INVERSION OF THE CHARACTERISTIC FUNCTION

- The numerical computation of the integral in 5.12 allows us to obtain the density function given the characteristic function. Among the variety of methods proposed in the financial literature, we present three methods due to their simplicity.
 1. Numerical integration using the `quadgk` Matlab function²
 2. Numerical integration via the trapezium rule and the Fast Fourier Transform. This method is very convenient if we need to compute the density at a large number of points. This method is discussed in Appendix.
 3. Numerical integration using the COS method. This method is discussed in Appendix.

Method 1: Gaussian quadrature

```
%Heston Parameters
kappa=5; theta=0.05;
sigma=0.5; rho=-0.8
tau=0.5; r=0.03;
S0=1; v0=0.05;
```

²An introduction to quadrature methods with applications to finance can be found in the book by Fusai and Roncoroni (2008). A quick exposition is offered in the Appendix.



Figure 5.32: Movie comparing the Euler and the exact scheme.

```
%Grid for the log-return
npoints=200;
[moments mcauchy]= momHeston( kappa, theta, sigma, rho, tau, r, S0, v0,1)
xmin=moments(1) -9*moments(2);
xmax=moments(1) +9*moments(2);
```



Figure 5.33: Movie of simulated log-returns in the Heston model.

```
ST=S0*linspace(exp(xmin),exp(xmax),npoints);
xT=log(ST);

%define the cf using shorthand notation
cf=@(u) cfHeston(u, kappa, theta, sigma, rho, tau, r, S0, v0);
```



Figure 5.34: Movie of simulated variance in the Heston model.

```
%Method 1: numerical quadrature using quadgk
utr=50; %truncation level of the integral
for j=1:length(xT)
    %fix integration range
    utrunc(j)=fsolve(@(u) real(exp(-li*u*xT(j)).*cf(u))-10^(-15),utr);
    utr=utrunc(j);
```

```
%numerical integration
qgkdf(j)=max(quadgk(@(u) real(exp(-1i*u*xT(j)).*cf(u)),0, utrunc(j))/pi,0);
end
```

Method 2: FFT method

```
%same setting as in other methods
...
%Method 2: FFT method
%(1) define grids
N = npoints;
ugrid = ( (0:N-1) - N/2 ) ; %c.f. grid
xT = ( (0:N-1) - N/2 ) /N *(2*pi) ; %p.d.f. grid
%(2) compute the cf on the cf grid
CharFn=cf(ugrid);
%(3) add +/- to the c.f.
% multiply with a vector h = (-1)^(j-1) forall j in {1,2,...,N}
% this is simply h = {-,+,-,+,...}
f = (-1).^(0:N-1) .* CharFn ;
%(4) fast fourier transform
tmpp = fft(f,N);
%
%(5) get p.d.f.
% switch +/- back
% only use the real part if positive
% divide by 2*pi
fftpdf = max(real( tmpp .* (-1).^(0:N-1) ./ (2*pi) ),0) ;
```

Method 3: COS method

```
%same setting as in other methods
...
```

```
%Method 3: COS method
a=xmin; b=xmax; N=npoints;
xT=linspace(a,b,N);
cf=@(u) cfHeston(u, kappa, theta, sigma, rho, tau, r, S0, v0);
ugrid=(0:N-1)*pi/(b-a);
CharFn=cf(ugrid);
for j=1:length(xT)
    V = (2/(b-a))*cos((xT(j)-a)*(0:N-1)*pi/(b-a));
    COSpdf(j)=max(real(sum(CharFn.*V.*exp(1i*(0:N-1)*pi*(-a)/(b-a)))-0.5*CharFn(1)*1*V(1)),0);
end
```

- A possible assessment of the different methods is to look at the moments originated by the approximation methods.

```
%exact moments
momexact=[1 momHeston( kappa, theta, sigma, rho, tau, r, S0, v0,0)];
%Moments using 1st method
xT=xTgk;
momqgkdf=[trapz(xT,qgkdf) trapz(xT,xT.*qgkdf) ...
trapz(xT,(xT.^2).*qgkdf) trapz(xT,(xT.^3).*qgkdf) ...
trapz(xT,(xT.^4).*qgkdf)];

%Moments using FFT method
xT=xTFFT;
momFFT=[trapz(xT,fftpdf) trapz(xT,xT.*fftpdf) ...
trapz(xT,xT.^2.*fftpdf) trapz(xT,xT.^3.*fftpdf) ...
trapz(xT,(xT.^4).*fftpdf)];

%Moments using COS method
xT=xTCOS;
momCOS=[trapz(xT,COSpdf) trapz(xT,xT.*COSpdf) ...
trapz(xT,(xT.^2).*COSpdf) trapz(xT,(xT.^3).*COSpdf) ...
trapz(xT,(xT.^4).*COSpdf)];

%compare different methods looking at the moments
```

```
cfrmoments=[momexact' momqgkdf' momFFT' momCOS']
err=[ (momqgkdf./momexact-1)' (momFFT./momexact-1)' (momCOS./momexact-1)' ]
```

- With reference to the assigned parameters, Table 5.10 reports the percentage error in computing the first four moments about the origin of the three inversion methods.

Order	GK	FFT	COS
0	0.000188	4.96E-13	0
1	0.009433	-2.5E-11	7.47E-05
2	0.001442	1.82E-10	-2.1E-05
3	-0.00718	1.04E-09	-0.0003
4	0.011467	1.46E-08	-0.00075

Table 5.10: Relative error in computing the first four moments using the density obtained via numerical inversion of the characteristic function.

- Figure 5.35 illustrates the Heston density obtained using the three different methods and compare them with a Gaussian density with same mean (0.25%) and standard deviation (16.21563%).
- Shapes of the density function for different values of the model parameters are illustrated in Figures (5.36)-(5.39).

5.7.4. HESTON MODEL AND OPTION PRICING

- The Heston model is used for better fitting market prices of European call and put option prices.
- The price of an European call option is given by

$$c_H(S, v = \sigma_t^2) = SP_1 - KP(t, T) P_2,$$

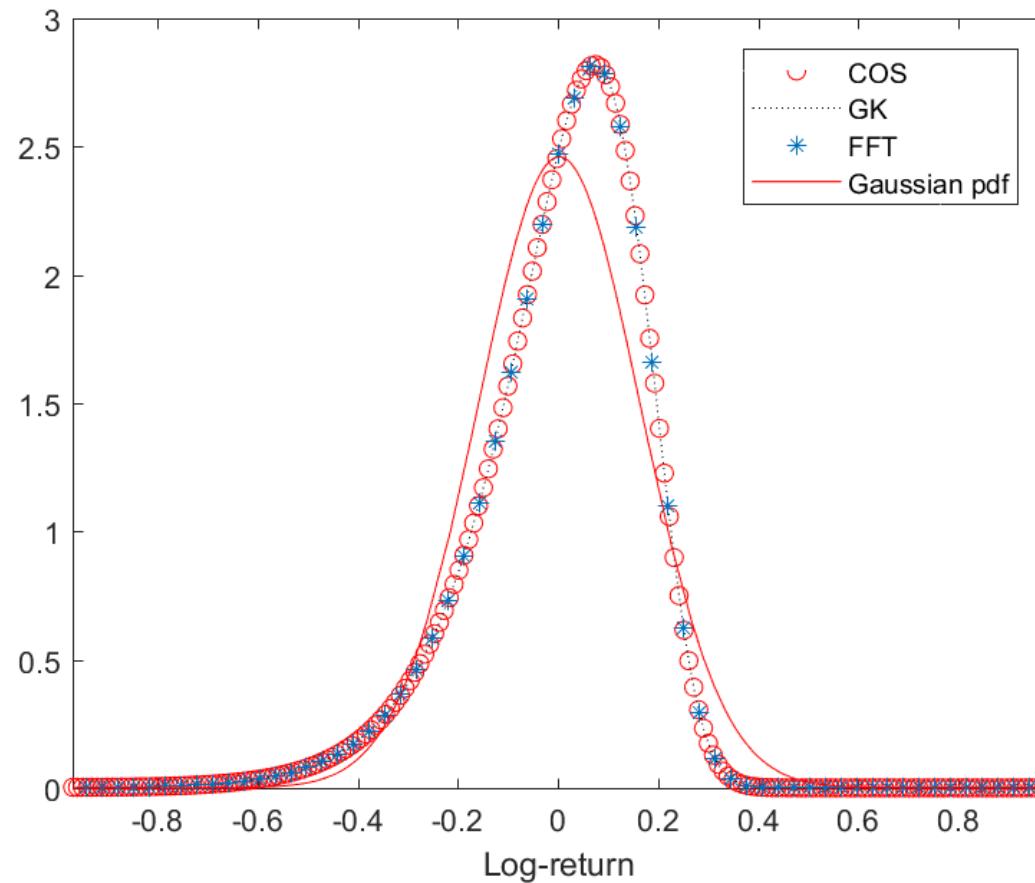


Figure 5.35: Heston PDF computed via different numerical inversion of the cf and comparison with the Gaussian PDF. Parameters: $\kappa = 5; \theta = 0.05; \sigma = 0.5; \rho = -0.8; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

where

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[\frac{e^{-iu \ln K} f_j(x, v, \tau; u)}{iu} \right] du, j = 1, 2.$$

where S is the stock price, $\tau = T - t$ is the time to maturity, K is strike price, $\operatorname{Re}()$ denotes the real part of a complex variable, $x = \ln(S)$ and $P(t, T) = e^{-r(T-t)}$ is the discount factor, and

$$f_j(x, v, \tau, u) = \exp(C_j(u, \tau) + D_j(u, \tau)v + iux),$$

$i = \sqrt{-1}$, $C_j(u, \tau)$ and $D_j(u, \tau)$ are the functions given in Section 5.7.1.

- In the previous section, we have seen how to recover the density function of the model numerically. Given the density functions computed on a pre-assigned grid, we can compute the option price via trapezium rule.

```
df=exp(-r*tau);
K=1
european=[trapz(xTgk,max(exp(xTgk)-K,0).*qgkdf) trapz(xTgk,max(K-exp(xTgk),0).*qgkdf)
trapz(xTFFT,max(exp(xTFFT)-K,0).*fftpdf) trapz(xTFFT,max(K-exp(xTFFT),0).*fftpdf)
trapz(xTCOS,max(exp(xTCOS)-K,0).*COSpdf) trapz(xTCOS,max(K-exp(xTCOS),0).*COSpdf)]*exp(-r*tau)
%Compute BS implied vol
ivol=blsimpv(S0,K,r,tau,european(:,1))
```

- European at-the-money call and put prices are given in the following Table. Last two rows give the Black-Scholes prices and the corresponding implied volatility. The Black-Scholes price is computed using the model standard deviation (16.21563%).
- The put option price can be computed via the put-call parity.
- However, a more convenient representation of the option price via a single univariate Fourier inversion formula is given by the so called Carr and Madan (1999) formula. This is discussed in the Appendix.
- The implied volatility shapes generated by the Heston model are shown in Figures 5.40-5.43 for different values of the model parameters.

	Call	Put
GK	0.068777	0.05384
FFT	0.068476	0.053588
COS	0.068691	0.053803
BS	0.053209	0.038321
IV	21.824%	21.824%

Table 5.11: Prices of European at-the-money call and put options.

- Depending on the sign of the correlation coefficient, the smile slope changes of sign, from negative (if $\rho < 0$) to positive (if $\rho > 0$). If $\rho = 0$, the smile turns out to be approximately symmetric, see Figure 5.40.
- Larger the vol-vol parameter, steeper the smile (the slope still depends on the sign of the correlation coefficient), see Figure 5.41.
- The long-run level θ affects the level of the implied volatility curve. In Figure 5.42 we have that the initial level of the variance is 0.05 and we consider increasing values for θ . This moves up the smile curve.
- Finally, the parameter κ does not seem to affect in a considerable way the smile curve, see Figure 5.43. Indeed, very often in the calibration of the Heston model, this parameter is assigned a fixed value.

5.7.5. HESTON MODEL AND OPTION PRICING: MAIN FINDINGS

- For ATM options, the Heston model gives the same result as the BS model.
- Bear in mind that
 - we require negative correlation between the two processes in order to generate an asymmetric distribution.
 - a higher value of the volatility of volatility parameter ε gives a higher kurtosis in the distribution.

- Introducing a non-zero correlation, the mispricing of OTM options is considerably reduced. This is true also in terms of hedging error.
- The Heston stochastic volatility model does not generate a considerable kurtosis and skewness over the short period.
- Heston stochastic volatility does not generate enough skewness and kurtosis in the short time. To accurately fit short term options we need to add a jump component to the return equation.
- Table 5.12 gives the main properties of the Heston model.

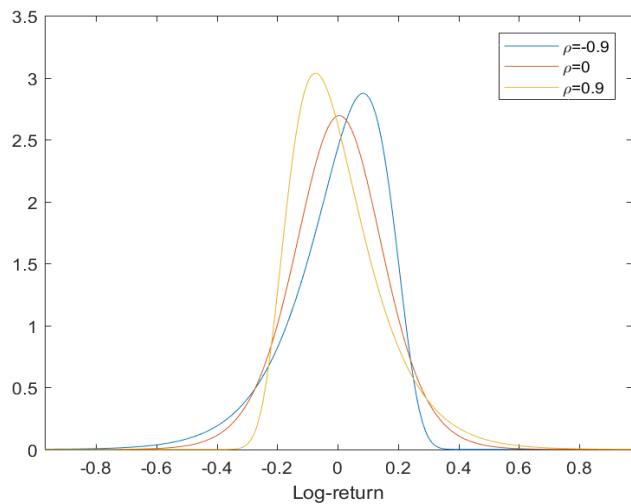


Figure 5.36: Heston PDF changing ρ . Parameters: $\kappa = 5; \theta = 0.05; \sigma = 0.5; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

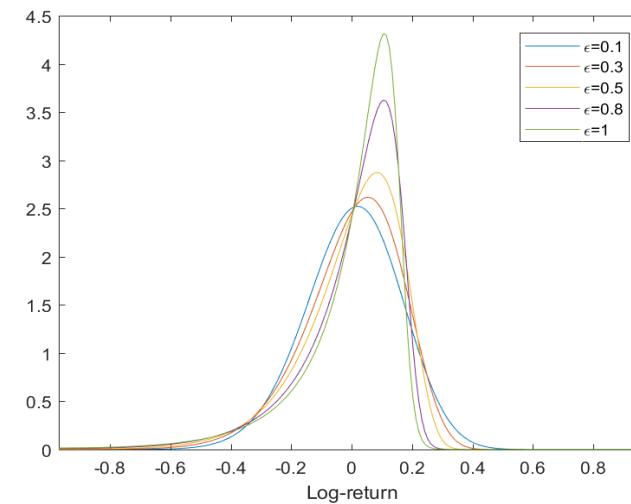


Figure 5.37: Heston PDF changing ϵ . Parameters: $\kappa = 5; \theta = 0.05; \rho = -0.8; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

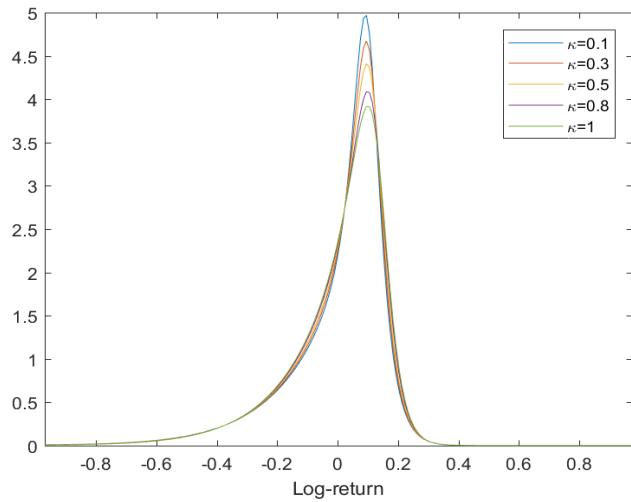


Figure 5.38: Heston PDF changing κ . Parameters: $\theta = 0.05; \sigma = 0.5; \rho = -0.8; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

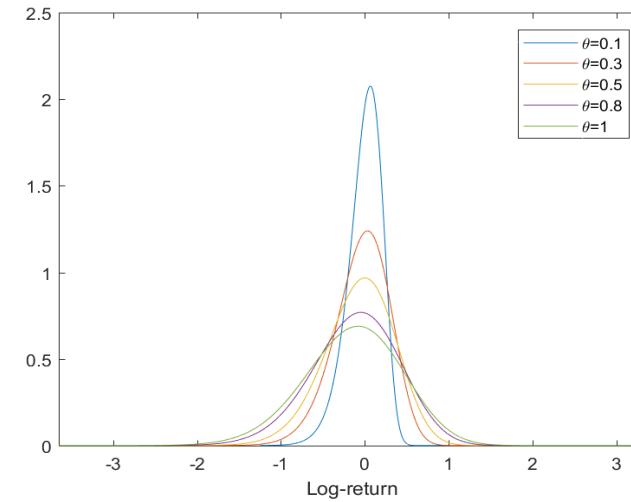


Figure 5.39: Heston PDF changing θ . Parameters: $\kappa = 5; \rho = -0.8; \sigma = 0.5; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

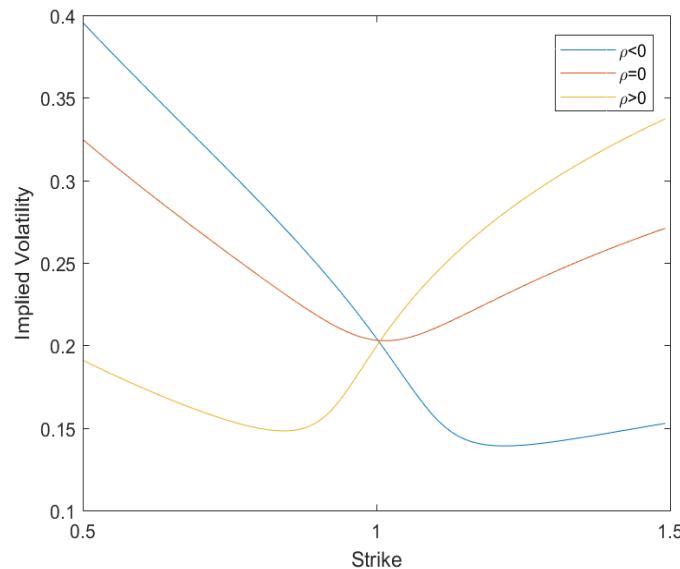


Figure 5.40: Black-Scholes implied volatility changing ρ in the Heston model.
 Parameters: $\kappa = 0.1; \theta = 0.05; \sigma = 0.5; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

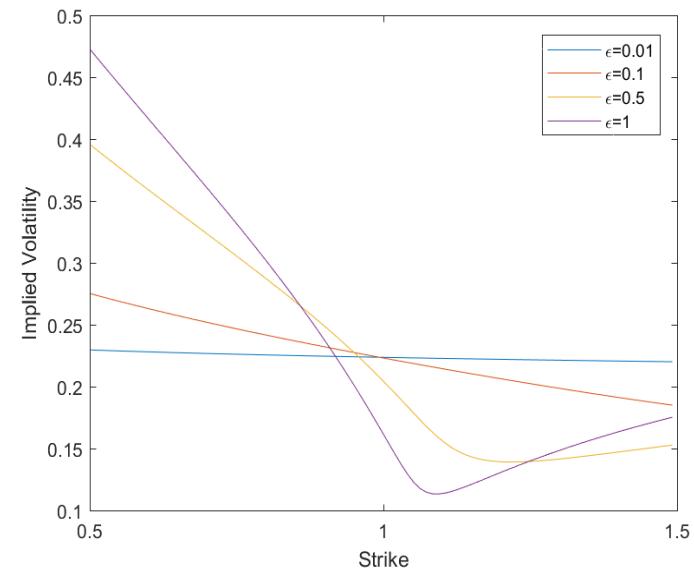


Figure 5.41: Black-Scholes implied volatility changing ϵ in the Heston model.
 Parameters: $\kappa = 0.1; \theta = 0.05; \rho = -0.8; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

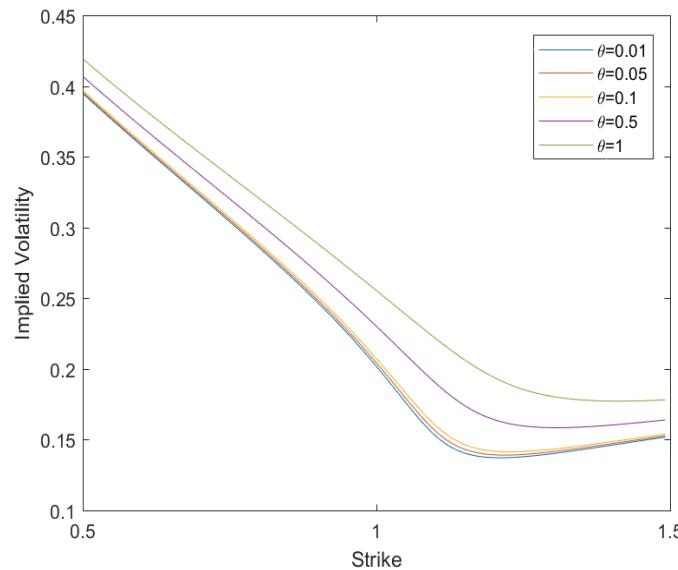


Figure 5.42: Black-Scholes implied volatility changing θ in the Heston model.
 Parameters: $\kappa = 0.1; \rho = -0.8; \sigma = 0.5; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

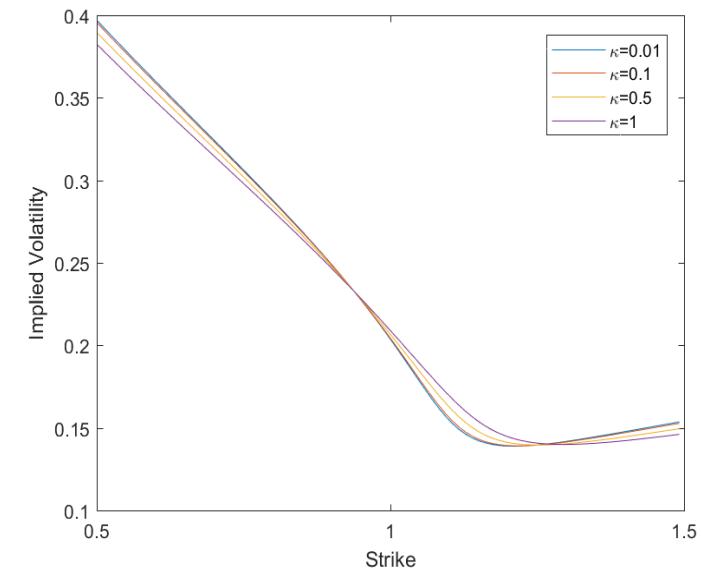


Figure 5.43: Black-Scholes implied volatility changing κ in the Heston model.
 Parameters: $\theta = 0.5; \rho = -0.8; \sigma = 0.5; t = 0.5; r = 0.03; x_0 = 0; v_0 = 0.05$.

5.8. THE HESTON APP

The reader can download from the web site associated to this document the Heston app, that allows him to play with the model parameters and to assess their effect on the shape of the density function and of the implied volatility smile. The user can also download in Excel the density and option prices corresponding at different strikes. A screenshot of the app is given in figure ??

HESTON MODEL $H(\mu, k, \varepsilon, \rho)$: FACTS

The (system of) SDEs

$$\begin{aligned} ds(t) &= \left(\mu - \frac{1}{2}v(t) \right) dt + \sqrt{v(t)} dW_s(t) && \text{(log-price)} \\ dv(t) &= k(\theta - v(t)) dt + \varepsilon \sqrt{v(t)} dW_v(t) && \text{(inst. variance)} \\ \mathbb{E}(dW_s(t) dW_v(t)) &= \rho dt \end{aligned}$$

with starting values $s(0)$ for the log-price and $v(0)$ for the variance.

The solution of the (system of) SDE

is not explicit.

The distribution

is not explicit

The mean of $s(t)$

$$\mathbb{E}_0(s_T) = s_0 + \left(\mu - \frac{\theta}{2} \right) \tau + (\theta - v_0) \frac{1 - e^{-\kappa\tau}}{2\kappa}$$

The moments of $X(t)$

not available as simple expressions but computable via 5.14
or by differentiating the characteristic function

The characteristic function of $X(t)$

see expression 5.13

Skewness of $s(t)$

negative, if $\rho < 0$
zero, if $\rho = 0$.
positive, if $\rho > 0$

Excess of kurtosis of $s(t)$

always positive

The stationary distribution of $s(t)$, ($t \rightarrow \infty$).

does not exist

Table 5.12: Properties of the Heston model.

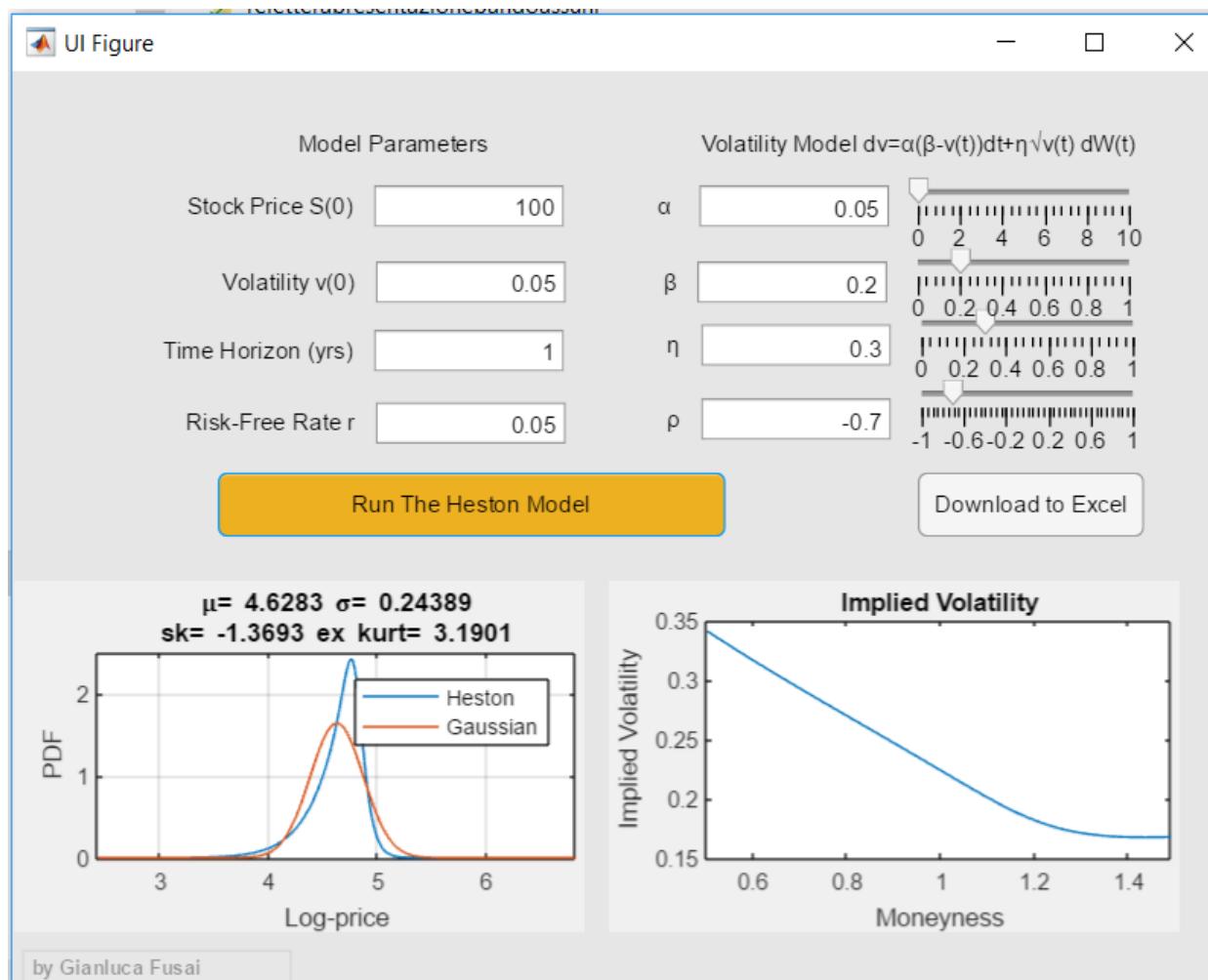


Figure 5.44: The app can be downloaded from here <https://sites.google.com/uniupo.it/stochasticcalculus/download-apps>. Its functioning has been checked on Windows 64 bit.

CHAPTER 6. STOCHASTIC PROCESSES WITH JUMPS

Contents

6.1	Preliminaries	156
6.1.1	The Poisson process	156
6.1.2	The Compound Poisson process	158
6.1.3	The Gamma process	159
6.1.4	Simulation of Jump Processes	160
6.2	Jump Diffusion (JD) processes	163
6.2.1	The Merton Jump Diffusion Process	163
6.2.2	The Kou process	173
6.3	Subordinated Brownian motions	176
6.3.1	The Variance (VG) Gamma process	177
6.4	Final Remark: Lévy processes	184
6.5	Jumps versus Stochastic Volatility	185

- The Brownian motion is a process which is continuous in time and space.
- As a consequence, it cannot capture extreme movements, like market crashes for example.
- The Brownian motion is, in fact, Gaussian, i.e. it has symmetric distribution with zero excess kurtosis.
- Skewness and excess kurtosis can be originated by allowing, for example, discontinuity in space, i.e. introducing jumps.
- Possible examples of such processes are
 - Jump Diffusion (JD) processes, like the Merton JD or the Kou JD;
 - Subordinated Brownian motions, like the VG process.
- The construction of these processes requires some preliminary facts listed in the following section.

Note that in the remaining of this chapter, we assume the following.

- The dynamics of the log-price is

$$s(t) = \log S(t),$$

where

$$s(t) = at + X(t).$$

- $X(t)$ is the stochastic process of interest.

6.1. PRELIMINARIES

6.1.1. THE POISSON PROCESS

Fact 51 A Poisson process is an increasing, positive stochastic process $N(t)$ on \mathbb{N} with independent and stationary increments which are Poisson distributed with instantaneous rate of arrival $\lambda > 0$. In other words, for any $0 < s < t$ the following hold.

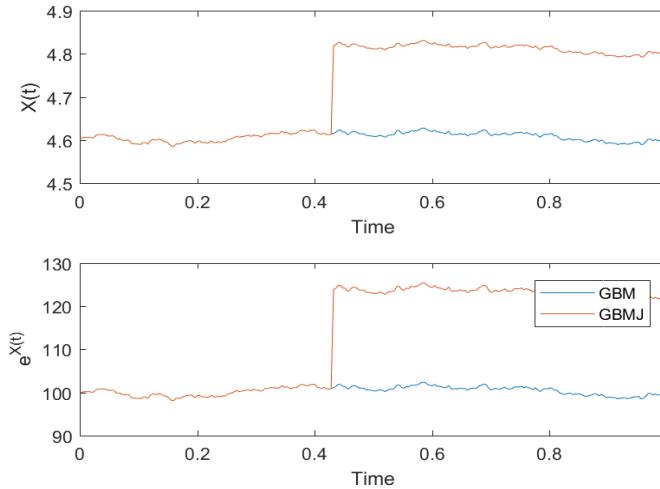


Figure 6.1: Sample trajectories of a log-price (top) and stock price (bottom) in the cases in which $X(t)$ is either an Arithmetic Brownian motion or a jump-diffusion process of the form $X_t = \mu t + \sigma W_t + \sum_{k=1}^{N_t} Z_k$. In this example, the jump size Z is Gaussian (Merton, 1976). The process X is obtained from the Arithmetic Brownian motion (the continuous parts are identical) by superimposing the compound Poisson process.

1. $N(0) = 0$;
2. $N(t) - N(s)$ is independent of the information set $\mathcal{F}(s)$ generated up to time s ;
3. $N(t) - N(s) \sim N(t-s) \sim Poi(\lambda(t-s))$.

Moreover, the characteristic function of $N(t)$ is

$$\phi_N(u; t) = e^{\lambda t(e^{iu}-1)},$$

where $i = \sqrt{-1}$ is the imaginary unit. Further,

$$\mathbb{E}(N(t)) = \lambda t$$

and

$$\text{Var}(N(t)) = \lambda t.$$

Hence, we note the following.

- It follows from properties (1) and (3) above that $N(t) \sim Poi(\lambda t)$.
- It follows from the definition of the Poisson distribution (see Appendix), that the increments can only take values 1 or 0 according to whether an arrival occurs or not.
- Hence, the Poisson process counts the arrivals in a system, like calls at a call center or shocks in the market.
- By definition of Poisson distribution, there cannot be more than one jump per time period.
- Hence, the Poisson process can only generate a finite number of jumps over a finite time horizon.
- For this reason, the Poisson process is said to have finite activity.

6.1.2. THE COMPOUND POISSON PROCESS

In order to gain some additional flexibility in modelling the size (severity) of the jumps, the Poisson process can be used to construct a more flexible process by assigning a specific distribution to the severities.

Fact 52 *A compound Poisson process is a stochastic process $Y(t)$ of the form*

$$Y(t) = \sum_{k=1}^{N(t)} Z_k,$$

where $\{Z_k\}_{k \in \mathbb{N}}$ is a sequence of independent and identically distributed random variables which are assumed independent of the Poisson process $N(t)$.

Moreover, the compound Poisson process has characteristic function

$$\phi_Y(u; t) = e^{\lambda t(\phi_Z(u)-1)},$$

where $\phi_Z(u)$ denotes the characteristic function of the random variable Z . It follows that

$$\mathbb{E}(Y(t)) = \lambda \mathbb{E}(Z) t,$$

and

$$\mathbb{V}ar(Y(t)) = \lambda \mathbb{E}(Z^2) t.$$

We can think of the compound Poisson process as follows.

- At time t a jump occurs.
- When this happens, the Poisson process increases of 1 unit.
- At the same time, a random draw Z is taken from a given distribution to quantify the jump size and it is summed up to the value of the process at the previous time point.
- The compound Poisson process has finite activity, like the Poisson process.

6.1.3. THE GAMMA PROCESS

Alternative processes which captures jump arrivals and size simultaneously like the compound Poisson process are available. One example is given by the Gamma process.

Fact 53 A Gamma process is a positive, non decreasing stochastic process $Y(t)$ with independent and stationary increments which follow a gamma distribution (see Appendix), i.e.

- $Y(0) = 0$;
- $Y(t) - Y(s)$ is independent of the information set up to time $s < t$;
- $Y(t) - Y(s) \sim \Gamma(\alpha(t-s), \lambda)$.

The characteristic function of the Gamma process is

$$\phi_Y(u; t) = \left(\frac{\lambda}{\lambda - iu} \right)^{\alpha t};$$

therefore

$$\mathbb{E}(Y(t)) = \frac{\alpha}{\lambda}t,$$

and

$$\mathbb{V}ar(Y(t)) = \frac{\alpha}{\lambda^2}t.$$

Remark 54 The Gamma process differs from the compound Poisson process in two aspects.

1. The Gamma process has infinite activity, as there can be an infinite number of jumps of very small size in any finite time period.
2. In the case of the Gamma process it is not possible to separate the rate of arrival of the jumps from their distribution.

Table (6.1) gives the main properties of the Gamma process.

Sample trajectories of the Poisson process, compound Poisson process and Gamma process are illustrated in Figure 6.2 (see also accompanying Matlab code).

6.1.4. SIMULATION OF JUMP PROCESSES

GAMMA PROCESS $\Gamma(\alpha, \lambda, k)$: FACTS

The process has

independent gamma distributed increments

The distribution of $X(t)$

$$\frac{\lambda^{\alpha t}}{\Gamma(\alpha t)} x^{\alpha t - 1} e^{-\lambda x}.$$

with $\Gamma(z)$ denoting the Gamma function

The characteristic function of $X(t)$

$$\phi_X(u; t) = \left(\frac{\lambda}{\lambda - iu} \right)^{\alpha t}.$$

The mean of $X(t)$

$$\mathbb{E}(X(t)) = \frac{\alpha}{\lambda} t.$$

The variance of $X(t)$

$$\text{Var}(X(t)) = \frac{\alpha}{\lambda^2} t.$$

The autocovariance of $X(t)$

$$c_X(t, s) = \sqrt{\frac{s}{t}}, s < t.$$

Table 6.1: Properties of the Gamma process.

Matlab Code

```
%%%%%%%%%%%%%
%%% SIMULATING JUMP PROCESSES %%%
%%%%%%%%%%%%%
%Assigning the number of simulated paths
%(nsimul), time to maturity (expiry), number of steps
%(nsteps), time step (dt)
%and observation times (timestep):
clear all; nsimul=5, expiry=1, nsteps=250;
dt=expiry/nsteps; timestep=[0:dt:expiry]';
%Assigning parameters
lambdaP=5; muZ=-0.05; sigmaZ=0.1; alpha=5; lambdaG=10;
%Simulate increments of the Poisson process
dN=poissrnd(lambdaP*dt, [nsteps,nsimul]);
%Simulate Poisson process
%(use cumulative sum of the increments):
cdN=zeros(1,nsimul); cumsum(dN);
%1. Simulate increments of the CPP
%for Gaussian jump sizes
dJ=muZ*dN+sigmaZ*sqrt(dN).*randn(nsteps,nsimul);
%2. Simulate CPP process
%(use cumulative sum of the increments):
cdJ=zeros(1,nsimul); cumsum(dJ);
%3. Simulate increments of the Gamma process:
dG=gamrnd(dt*alpha,1/lambdaG, [nsteps,nsimul]);
%4. Simulate Gamma process
%(use cumulative sum of the increments):
cdG=zeros(1,nsimul); cumsum(dG);
%Plot simulated paths:
h=figure('Color',[1 1 1])
subplot(3,1,1);plot(timestep, cdN); xlabel('Time (years)')
title('Simulated Paths of the Poisson Process')
subplot(3,1,2);plot(timestep, cdJ); xlabel('Time (years)')
title('Simulated Paths of the Compound Poisson Process')
subplot(3,1,3);plot(timestep, cdG); xlabel('Time (years)')
title('Simulated Paths of the Gamma Process')
```

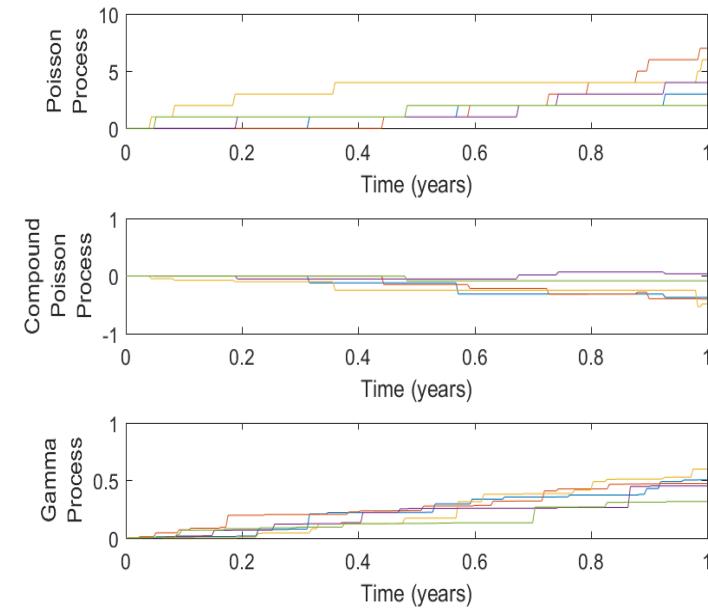


Figure 6.2: Top panel: simulated paths of the Poisson process $N(t) \sim Poi(\lambda t)$ for $\lambda = 5$. Middle panel: simulated paths of the Compound Poisson process with Gaussian jump severities with parameters $\lambda = 5$, $\mu_Z = -0.05$, $\sigma_Z = 0.1$. Bottom panel: simulated paths of the Gamma process $G(t) \sim \Gamma(\alpha t, \lambda)$, with parameters $\alpha = 5$, $\lambda = 10$ (bottom panel).

6.2. JUMP DIFFUSION (JD) PROCESSES

Fact 55 (Jump Diffusion process) A Jump Diffusion process is a stochastic process $X(t)$ with independent and stationary increments which is obtained as the sum of an Arithmetic Brownian Motion and an independent compound Poisson process, i.e.

$$X(t) = \mu t + \sigma W(t) + \sum_{j=1}^{N(t)} Z_k,$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ and

- $W(t)$ is a Brownian motion,
- $N(t)$ is a Poisson process with instantaneous rate of arrival $\lambda > 0$ and independent of $W(t)$,
- $\{Z_k\}_{k \in \mathbb{N}}$ is a sequence of independent and identically distributed random variables, which are independent of both the Brownian motion and the Poisson process.

We can further ‘specialize’ the compound Poisson part of the JD process by specifying the distribution of the jump severities. Common choices for this distribution in financial applications are the Gaussian distribution and the exponential distribution.

6.2.1. THE MERTON JUMP DIFFUSION PROCESS

Fact 56 Let us assume that the jump severities follow a Gaussian distribution, i.e. $Z \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$. Then, the JD process $X(t)$ is a Merton JD process.

- The process takes its name from Robert Merton who first used it for financial applications.
- The choice of modelling dynamics using a JD process is quite common in financial applications due to the following observation.
- Stock prices appear to have small continuous movements most of the time (due, for example, to a temporary imbalance between demand and supply);

- sometimes though they experience large jumps upon the arrival of important information with more than just a marginal impact.
- By its very nature, important information arrives only at a finite number of points in time, hence the jumps have finite activity.

Properties of the Merton JD process

In the Merton model the density function can be obtained in two ways.

1. As weighted average of an infinite number of Gaussian densities with different means and variances; the weights are computed according to the Poisson distribution:

$$f_X(x) = \sum_{j=0}^{\infty} w_j n(x, \mu_j, \sqrt{\sigma_j^2}),$$

where

$$w_j = \frac{e^{-\lambda t} (\lambda t)^j}{j!}, \mu_j = \mu + j\mu_Z, \sigma_j^2 = \sigma^2 + j\sigma_Z^2,$$

and $n(x, m, s)$ is the Gaussian PDF with mean m and standard deviation s . From a computational point of view, the above sum is truncated at N_{max} that can be chosen so that the sum of the weights is near to 1, i.e.¹

$$\sum_{j=0}^{N_{max}} w_j > 1 - 10^{-8}.$$

2. Numerically inverting the characteristic function, that is given by

$$\mathbb{E} \left(e^{iuX(t)} \right) = \exp \left((\mu - \frac{1}{2}\sigma^2 u^2)t + \lambda(\exp(iu\mu_Z - 0.5u^2\sigma_Z^2) - 1)t \right);$$

¹Another possibility is to truncate when the weight becomes very small: $w_{N_{max}} < 10^{-8}$.

The moments of $X(t)$ can be obtained in closed form.

- $\mathbb{E}X(t) = (\mu + \lambda\mu_Z)t$
- $\mathbb{V}ar(X(t)) = (\sigma^2 + \lambda(\mu_Z^2 + \sigma_Z^2))t$
- The indices of skewness and excess kurtosis are respectively

$$\begin{aligned} Skew(t) &= \frac{\lambda\mu_Z(\mu_Z^2 + 3\sigma_Z^2)}{(\sigma^2 + \lambda(\mu_Z^2 + \sigma_Z^2))^{3/2}\sqrt{t}}, \\ \mathbb{E}Kurt(t) &= \frac{\lambda(\mu_Z^4 + 6\mu_Z^2\sigma_X^2 + 3\sigma_Z^4)}{(\sigma^2 + \lambda(\mu_Z^2 + \sigma_Z^2))^2 t}. \end{aligned}$$

- The Matlab script `get_moments_JD` allows the user to compute these quantities

Interpretation of the parameters

- μ = drift of the diffusion component of the process.
- σ = volatility of the Brownian motion.
- λ = rate of arrival of the jumps; it controls the level of excess kurtosis.
- μ_Z = mean of the jump sizes; it controls the sign of the skewness index. Hence, the Merton jump diffusion has a distribution which is skewed to the left if $\mu_Z < 0$ and skewed to the right if $\mu_Z > 0$.
- σ_Z = volatility of the jump sizes. The larger this volatility, the larger the excess kurtosis of the process.

Matlab Code

```
%%%%%%%%%%%%%
%%%Build the density of the Merton model
%%%and compare to the Gaussian one
%%%%%
clear all
%Parameters
lambda=1.5;mu=0.03; sigma=0.2;sigmaZ=0.1; muZ=-0.03;expiry=0.25;
%We assume we cannot have more han 30 jumps in the time frame
N=[0:30];
%Mean and standard deviation to build the x-grid
meanJD=(mu+lambda*muZ)*expiry;
sdJD=((sigma^2+lambda*sigmaZ^2)*expiry)^0.5;
xmin=meanJD-6*sdJD; xmax=meanJD+6*sdJD;
xT=linspace(xmin,xmax,100);
%1. Compute weights
weights=poisspdf(N,lambda*expiry);
%2. Truncate if the sum of weights reaches a threshold
Nmax=max(find(cumsum(weights)<0.9999999));
%3. mean and variance if there are j jumps
meanJ=mu*expiry+[0:Nmax]*muZ;
varJ=sigma^2*expiry+[0:Nmax]*sigmaZ^2;
%4. Build the density
for j=[0:Nmax]
    gaussJD(j+1,:)=weights(j+1)*pdf('norm',xT,meanJ(j+1),varJ(j+1)^0.5);
end
%5. The density
pdfMerton=sum(gaussJD);
%6. A plot
h=figure('Color',[1 1 1]);
plot(xT,pdfMerton)
hold on
plot(xT, pdf('norm',xT,meanJD,sdJD ))
legend('Merton JD','Gaussian')
xlim([xmin xmax])
```

```
xlabel('x')
print(h, '-dpng', 'FigMertonJDDensity')
```

Parameters fitting

- The simplest method (albeit not very accurate) to fit the parameters is to use the method of moments procedure.
- It consists in minimizing the distance between sample moments (such as sample mean, sample variance, sample skewness and sample kurtosis) with theoretical ones.
- For example, over year 2012, the log-return series of crude oil price were characterized by the sample moments in Table (6.2):

Mean	Variance	Skewness	Excess Kurtosis
-0.0003	0.0144	0.1417	4.3605

Table 6.2: Sample moments of daily log-price changes in oil price in year 2012.

- We can solve for the MJD parameters such that theoretical moments fit the ones in Table 6.2.
- The theoretical moments can be computed using the following Matlab function

```
function m=get_moments_JD(mu, sg, lambda, muZ, sgZ, t)

m(1,:)= (mu+lambda*muZ)*t; %mean
m(2,:)= (sg*sg+lambda*(muZ.^2+sgZ.^2))*t; %variance
numsk= lambda.*muZ.* (muZ.^2+3*sgZ.^2);
densk=(sg.^2+lambda.* (muZ.^2+sgZ.^2))^.5*t.^0.5;
m(3,:)=numsk./densk;%skewness
numk= lambda.* (muZ.^4+6*muZ.^2.*sg.^2+3*sgZ.^4);
denk=(sg.^2+lambda.* (muZ.^2+sgZ.^2))^.2*t;
m(4,:)=numk./denk;%excess kurtosis
```

- The fitting can be achieved through the following commands in the Matlab command window.

```
%fitting parameters
ms=[-0.0003 0.0144 0.1417      4.3605];
x0(1)=0; x0(2)=0.05; x0(3)=0.5; x0(4)=0.01; x0(5)=0.17
[xopt fval]= fminsearch(@(x) sum(((get_moments_JD(x(1), x(2), x(3), x(4), x(5),1)-ms').^2)),x0)
mJD=get_moments_JD(xopt(1), xopt(2), xopt(3), xopt(4), xopt(5),1)
```

We obtain the parameter estimates as in Table (6.3).

μ	σ	λ	μ_Z	σ_Z
-0.0037	0.0407	0.5373	0.0064	0.1541

Table 6.3: Calibrated parameters of the MJD model to sample moments of daily log-price changes in oil price for year 2012.

Simulating the Merton JD process

- The simulation procedure for the trajectories of the Merton Jump Diffusion process is based on the following two observations.
 - The increments of the Poisson process are independent and follow a Poisson distribution with rate $\lambda(t_{j+1} - t_j)$;
 - conditioned on the number of jumps occurred between t_j and t_{j+1} , the sum of the jump severities is Gaussian with given mean and variance.
- Hence, the simulation algorithm can be organized as follows.

Step 1. Simulate the continuous part of the JD diffusion process, i.e. the ABM, on the given time partition.

Step 2. Simulate the number of jumps occurring between t_j and t_{j+1} , i.e. $N \sim Poi(\lambda(t_j, t_{j+1}))$.

Step 3. Generate $Z \sim \mathcal{N}(0, 1)$; set $J = \mu_Z N + \sigma_Z \sqrt{N}Z$.

Step 4. Sum the ABM and J .

- Simulated paths are illustrated in Figure 6.3. A comparison between the density of the MJD model and the Gaussian with same mean and variance is given in Figure 6.4.

Matlab Code

```

%%%%% SIMULATING THE MERTON JD Process %%%%
%%%%% Assigning the number of simulated paths
%(nsimul), time to maturity (expiry), number of steps
%(nsteps), time step (dt) and observation times (timestep):
clear all; nsimul=50; expiry=1; nsteps=250;
dt=expiry/nsteps; timestep=[0:dt:expiry]';
%Assigning parameters
mu=-0.0003; sigma=0.0425; lambda=0.5175;
muZ=0.0064; sigmaZ=0.1520;
%1. Simulate increments of the ABM
dW=mu*dt+sigma*sqrt(dt).*randn(nsteps,nsimul);
%2. Simulate increments of the CPP
dN=poissrnd(lambda*dt,[nsteps,nsimul]);
dJ=muZ*dN+sigmaZ*sqrt(dN).*randn(nsteps,nsimul);
dX=dW+dJ;
%3. Simulate MJD process
%(use cumulative sum of the increments):
cdX=zeros(1,nsimul); cumsum(dX);
%Plot simulated paths:
h=figure('Color',[1 1 1])
plot(timestep, cdX); xlabel('Time (years)')
xlabel('Time')
title('Simulated Paths of the Merton JD Process')
print(h,'-dpng','FigMJDPaths')

```

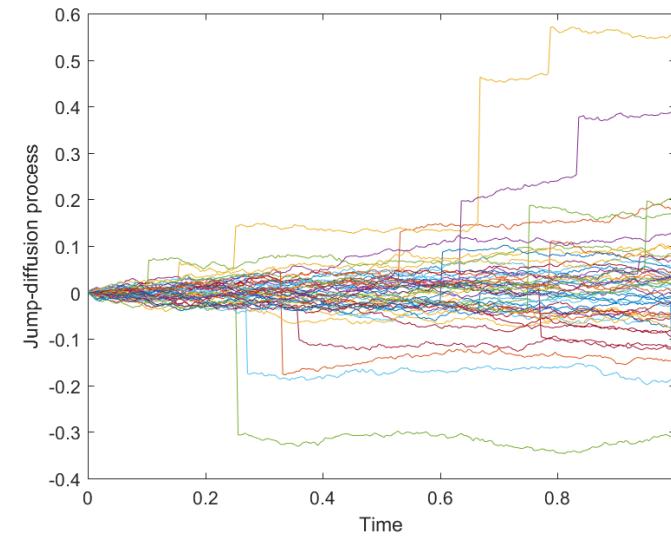


Figure 6.3: Simulated paths of the Merton jump diffusion process. Parameters: $\mu = -0.0003$, $\sigma = 0.0425$, $\lambda = 0.5175$, $\mu_Z = 0.0064$, $\sigma_Z = 0.1520$.

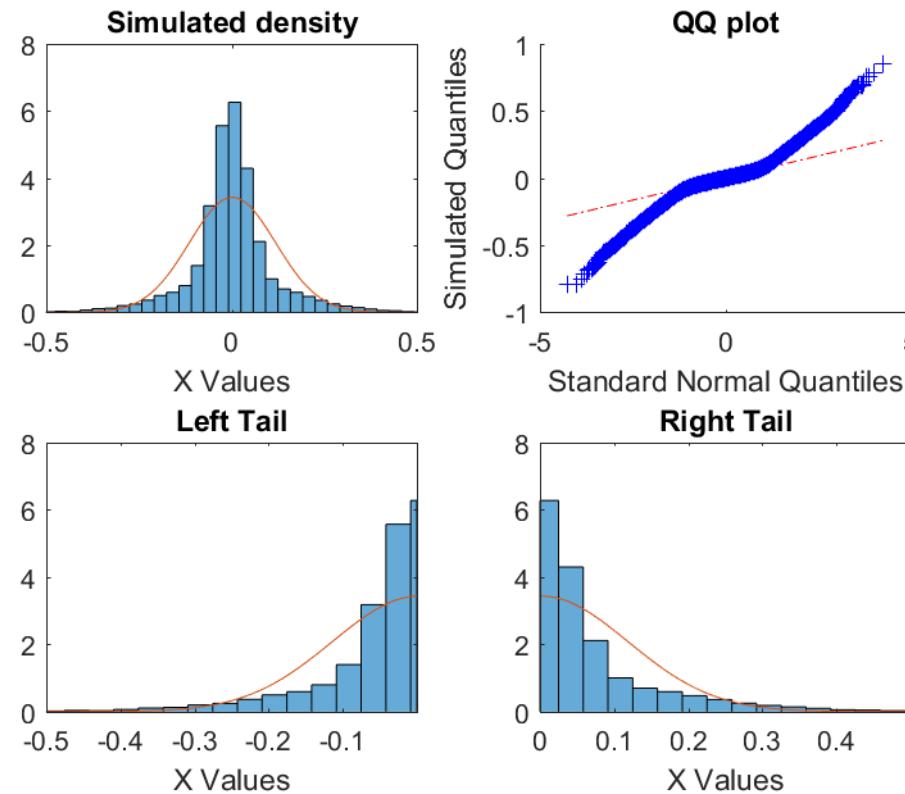


Figure 6.4: Top Left: Simulated density of the Merton jump diffusion process at 1 day horizon and superimposed Gaussian density with same mean and variance. Top right: qqplot of simulated returns. Bottom left: Left tail of the simulated returns versus Gaussian tail. Bottom right: right tail of the simulated returns versus Gaussian tail. Parameters as in Table (6.3).

MERTON JUMP-DIFFUSION($\mu, \sigma, \lambda, \mu_Z, \sigma_Z$): FACTS

The jump-diffusion process

$$X(t) = \mu t + \sigma W(t) + \sum_{j=1}^{N(t)} Z_k, Z_k \sim \mathcal{N}(\mu_Z, \sigma_Z^2),$$

where $N(t)$ is Poisson process with intensity λ

The distribution of $X(t)$

$$f_X(x, t) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{-\frac{1}{2} \left(\frac{(x - (\mu + j\mu_Z)t)^2}{(\sigma^2 + j\sigma_Z^2)t} \right)}$$

i.e. a weighted average of Gaussian pdfs with means $\mu + j\mu_Z$ and variances $\sigma^2 + j\sigma_Z^2$.
 The weights are given by the Poisson probability of having j jumps up to time t

The characteristic function of $X(t)$

$$\psi(u, t) = \mathbb{E}(e^{iuX(t)}) = \exp \left(iu\mu t - \frac{1}{2}u^2\sigma^2 t + \lambda \left(e^{iu\mu_Z} - u^2\sigma_Z^2/2 - 1 \right) t \right)$$

The mean of $X(t)$

$$(\mu + \lambda\mu_Z) t$$

The variance of $X(t)$

$$(\sigma^2 + \lambda(\mu_Z^2 + \sigma_Z^2)) t$$

The skewness of $X(t)$

$$\frac{\lambda\mu_Z(\mu_Z^2 + 3\sigma_Z^2)}{(\sigma^2 + \lambda(\mu_Z^2 + \sigma_Z^2))^{3/2}\sqrt{t}}$$

The excess of kurtosis of $X(t)$

$$\frac{\lambda(\mu_Z^4 + 6\mu_Z^2\sigma_X^2 + 3\sigma_Z^4)}{(\sigma^2 + \lambda(\mu_Z^2 + \sigma_Z^2))^2 t}$$

The stationary distribution of $X(t)$, ($t \rightarrow \infty$)

Does not exist

For large t the skewness and the excess kurtosis go to 0.The auto-covariance of $X(t)$

$$c_X(t, s) = \mathbb{V}ar(X_1) \min(t, s) = (\sigma^2 + \lambda(\mu_Z^2 + \sigma_Z^2)) \min(t, s)$$

Table 6.4: Properties of the Merton Jump-Diffusion model.

6.2.2. THE KOU PROCESS

In the case of the Kou process, the jump sizes follow a double exponential distribution with parameters (p, η_1, η_2) , i.e. their density function is given by

$$p\eta_1 e^{-\eta_1 y} 1_{(y \geq 0)} + (1 - p) \eta_2 e^{\eta_2 y} 1_{(y < 0)}, \quad \eta_1, \eta_2 > 0, p \in [0, 1].$$

Properties of the Kou JD process

The density of this process is not obtainable in closed form in terms of easy to compute functions. However the characteristic function is known in closed form

$$\begin{aligned} \mathbb{E} \left(e^{iuX(t)} \right) &= e^{\lambda t(\phi_Z(u)-1)}, \\ \phi_Z(u) &= p \frac{\eta_1}{\eta_1 - iu} + q \frac{\eta_2}{\eta_2 + iu} \end{aligned}$$

as well as the moments of the log-process. In particular, we have

- $\mathbb{E}(X(t)) = (\mu + \lambda(p/\eta_1 - (1-p)/\eta_2))t$
- $\text{Var}(X(t)) = (\sigma^2 + 2\lambda(p/\eta_1^2 + (1-p)/\eta_2^2))t$
- The indices of skewness and excess kurtosis are respectively

$$\begin{aligned} \text{Skew}(t) &= \frac{6\lambda(p/\eta_1^3 - (1-p)/\eta_2^3)}{(\sigma^2 + 2\lambda(p/\eta_1^2 + (1-p)/\eta_2^2))^{3/2} \sqrt{t}}, \\ \mathbb{E}\text{Kurt}(t) &= \frac{24\lambda(p/\eta_1^4 + (1-p)/\eta_2^4)}{(\sigma^2 + 2\lambda(p/\eta_1^2 + (1-p)/\eta_2^2))^2 t}. \end{aligned}$$

Interpretation of the parameters

- μ = drift of the process.
- σ = volatility of the Brownian motion.
- λ = rate of arrival of the jumps; it controls the level of excess kurtosis.
- p = probability of an upward jump.
- η_1 = parameter of the exponential distribution controlling the upward jumps; therefore, the upward jumps have mean $1/\eta_1$.
- η_2 = parameter of the exponential distribution controlling the downward jumps; therefore, the downward jumps have mean $1/\eta_2$.

KOU JUMP-DIFFUSION($\mu, \sigma, \lambda, \mu_Z, \sigma_Z$): FACTS

The solution

$$X(t) = \mu t + \sigma W(t) + \sum_{j=1}^{N(t)} Z_k, Z_k \sim \mathcal{E}(p, \eta_1, \eta_2),$$

where the density of the jump size Z is

$$f_Z(z) = p\eta_1 e^{-\eta_1 z} 1_{(z \geq 0)} + (1-p)\eta_2 e^{\eta_2 z} 1_{(z < 0)}, \quad \eta_1, \eta_2 > 0, p \in [0, 1].$$

and $N(t)$ is Poisson process with intensity λ

The distribution of $X(t)$

not available in closed form

The characteristic function of $X(t)$

$$\psi(u, t) = \mathbb{E}\left(e^{iuX(t)}\right) = e^{\lambda t(\phi_Z(u)-1)}, \text{ where}$$

$$\phi_Z(u) = p \frac{\eta_1}{\eta_1 - iu} + q \frac{\eta_2}{\eta_2 + iu}$$

The mean of $X(t)$

$$(\mu + \lambda(p/\eta_1 - (1-p)/\eta_2))t$$

The variance of $X(t)$

$$(\sigma^2 + 2\lambda(p/\eta_1^2 + (1-p)/\eta_2^2))t$$

The skewness of $X(t)$

$$\frac{6\lambda(p/\eta_1^3 - (1-p)/\eta_2^3)}{(\sigma^2 + 2\lambda(p/\eta_1^2 + (1-p)/\eta_2^2))^{3/2}\sqrt{t}}$$

The excess kurtosis of $X(t)$

$$\frac{24\lambda(p/\eta_1^4 + (1-p)/\eta_2^4)}{(\sigma^2 + 2\lambda(p/\eta_1^2 + (1-p)/\eta_2^2))^2 t}$$

For large t the skewness and the kurtosis go to 0

The stationary distribution of $X(t)$, ($t \rightarrow \infty$)

Does not exist

The auto-covariance of $X(t)$

$$c_X(t, s) = \mathbb{V}ar(X_1) \min(t, s) = (\sigma^2 + 2\lambda(p/\eta_1^2 + (1-p)/\eta_2^2)) \min(t, s)$$

6.3. SUBORDINATED BROWNIAN MOTIONS

An alternative way of constructing stochastic processes with jumps is to consider an Arithmetic Brownian Motion on a time scale which is not governed by the standard calendar time, but by a random clock. These processes are called Time Changed Brownian Motions and we have already seen an introduction to these processes in Section 3.6.

Fact 57 *A Time Changed Brownian motion is a process of the form*

$$X(t) = \theta G(t) + \sigma W(G(t)), \quad \theta \in \mathbb{R}, \sigma > 0,$$

where $W(t)$ is a Brownian motion and $G(t)$ is a positive, increasing stochastic process. The law of the process G is what allows us to characterize the resulting process X .

Constructing Time Changed Brownian Motions has particular economic appeal as

- this construction finds its rationale in the following: uncertainty in prices changes is originated by the time at which the next investor enters the market with a transaction altering the current price values, and the amount by which this current price is changed by. The random clock models the time at which the next transaction will take place; the ‘size’ of the price change is instead captured by the Brownian motion component.
- Empirical evidence shows that stock log-returns are Gaussian but only under trade time, rather than standard calendar time.
- Further, the time change construction recognizes that stock prices are largely driven by news, and the time between one piece of news and the next is random as is its impact.
- Finally, this construction offers a high degree of mathematical tractability as, once we operate under business time, log-returns are once again Gaussian and therefore the results derived for the Black-Scholes model still hold.
- Both JD and time-changed Brownian motions are characterized by trajectories formed by many little movements with large, rare movements interspersed; however, whilst in the case of JD processes the small increments have Gaussian distribution, in the case of time-changed Brownian motion this distribution has skewness and excess kurtosis.

- We note that if the process chosen as time change has independent and stationary increments, it is called subordinator. The resulting time-changed Brownian motion is then called subordinated Brownian motion.

A subordinated Brownian motion commonly used in finance is the Variance Gamma process.

6.3.1. THE VARIANCE (VG) GAMMA PROCESS

Fact 58 Let us assume that G is a Gamma process with parameters $\alpha = \lambda = k^{-1}$, for any positive constant k , so that $\mathbb{E}G(t) = t$ and $\text{Var}(G(t)) = kt$. Assume G is independent of W . Then, $X(t)$ is a VG process.

We note the following.

- The parameters of the Gamma process are chosen so that $\mathbb{E}(G(t)) = t$, i.e. the process chosen as random clock is an unbiased representation of calendar time.
- The VG process has infinite activity, i.e. it is characterized by an infinite number of jumps in any finite time period.
- The VG process has finite variation, i.e. the process ‘travels’ a finite space in any finite time period.
- The VG process is a subordinated Brownian motion as the Gamma process has independent and stationary increments, i.e. it is a subordinator.

Other examples of subordinated Brownian motions used for financial applications are the Normal Inverse Gaussian and the CGMY process.

Properties of the VG process

- The probability density function is

$$2 \frac{e^{\theta x / \sigma^2}}{k^{t/k} \sigma \sqrt{2\pi} \Gamma(t/k)} \left(\frac{x^2}{\theta^2 + 2\sigma^2/k} \right)^{\frac{t}{2k} - \frac{1}{4}} K_{\frac{t}{k} - \frac{1}{2}} \left(\frac{|x|}{\sigma^2} \sqrt{\theta^2 + 2\sigma^2/k} \right) \quad (6.1)$$

with $K_v(z)$ denoting the modified Bessel function of the third kind with order v and argument z .

- The characteristic function of the VG process is

$$\phi_X(u; t) = \left(1 - iu\theta k + u^2 \frac{\sigma^2}{2} k\right)^{-\frac{t}{k}}, \quad (6.2)$$

- $\mathbb{E}(X(t)) = \theta t$,
- $\mathbb{V}ar(X(t)) = (\sigma^2 + \theta^2 k) t$,
- the indices of skewness and excess kurtosis are respectively

$$\begin{aligned} Skew(t) &= \frac{(3\sigma^2 + 2\theta^2 k) \theta k}{(\sigma^2 + \theta^2 k)^{3/2} \sqrt{t}}, \\ \mathbb{E}Kurt(t) &= \frac{(3\sigma^4 + 12\sigma^2\theta^2 k + 6\theta^4 k^2) k}{(\sigma^2 + \theta^2 k)^2 t}. \end{aligned}$$

- The above quantities are computed via the Matlab function `get_moments_VG`.

Interpretation of the parameters

- $\theta \in \mathbb{R}$: mean of the VG process; it also controls the sign of the skewness index. Hence, the VG process has distribution skewed to the left if $\theta < 0$, and skewed to the right if $\theta > 0$. If $\theta = 0$, the process has symmetric distribution.
- $\sigma > 0$: it controls the variance of the VG process. If $\sigma = 0$, the VG process reduces to the Gamma process.
- $k > 0$: variance rate of the Gamma process. It controls the level of excess kurtosis.
- The term structures of skewness and excess kurtosis are illustrated in Figure 6.5, which shows that as the horizon lengths, the skewness and the excess kurtosis tend to fade away. The moments are computed using the Matlab function `get_moments_VG` given below.

- This is also shown in Figure 6.6 in which we compare the VG density with the Gaussian with same mean and variance. Over short horizons the two distributions are very different, whilst as the horizon lengthens the VG density approaches the Gaussian one. The density has been computed numerically inverting the characteristic function using the COS method.

Matlab Code

```
function m=get_moments_VG(theta, sg, kappa,t)

m(1,:)= theta*t; %mean
m(2,:)= sg*sg*t+theta*theta*kappa*t; %variance
numsk= (3*sg^2+2*theta^2*kappa)*theta*kappa;
densk=(sg^2+theta^2*kappa)^(3/2)*t.^0.5;
m(3,:)=numsk./densk;%skewness
numk=(3*sg^4+12*sg^2*theta^2*kappa+6*theta^4*kappa^2)*kappa;
denk=(sg*sg+theta*theta*kappa)^2*t;
m(4,:)=numk./denk;%kurtosis
```

Simulation of the VG process

The simulation procedure of the Variance Gamma process is based on the following two observations.

1. The increments of the Gamma process are independent and follow a Gamma distribution $\Gamma((t_{j+1} - t_j)/k, 1/k)$;
2. conditioned on the increments of the Gamma clock, the increments of the VG process are Gaussian with given mean and variance.

Hence, the simulation algorithm can be organized as follows.

Step 1. Simulate the increments from t_j to t_{j+1} of the Gamma clock, i.e. $G \sim \Gamma((t_{j+1} - t_j)/k, 1/k)$.

Step 2. Generate $Z \sim \mathcal{N}(0,1)$; set $X = \theta G + \sigma\sqrt{G}Z$.

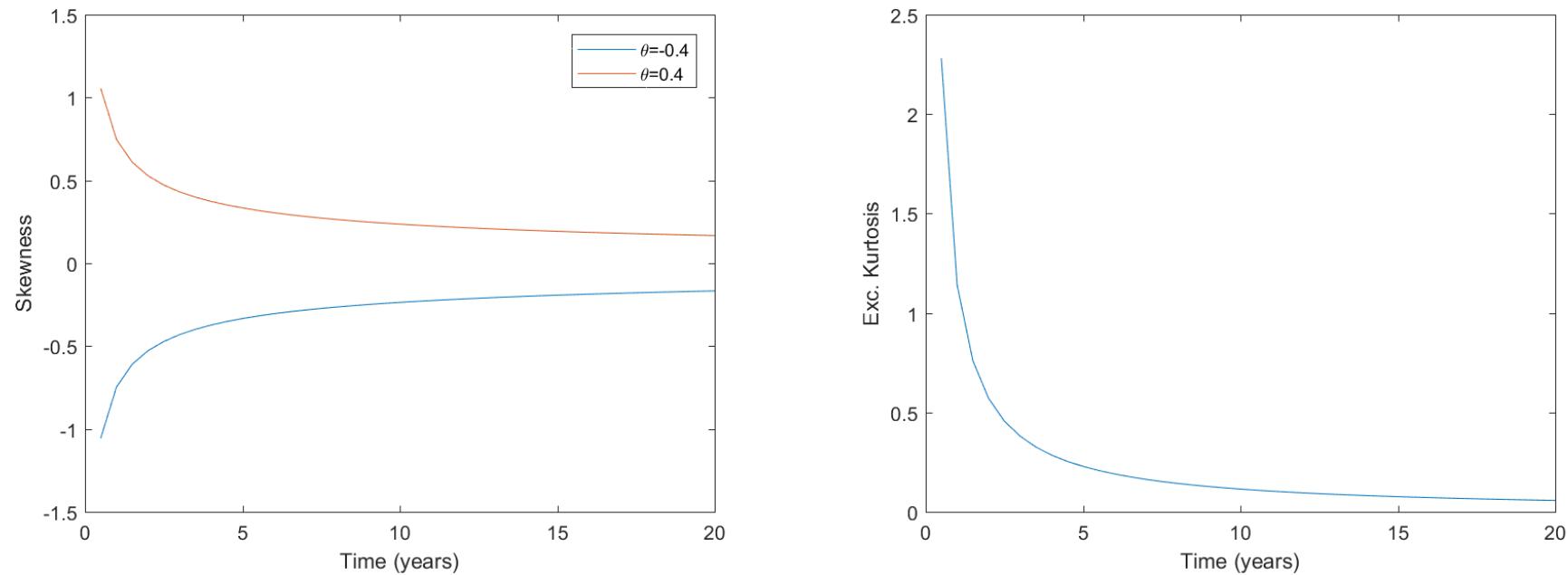


Figure 6.5: Term structures of skewness (**Left**) for two different values of the parameter θ and excess kurtosis (**Right**) in the VG model. Parameters: $\theta = -0.4$, $\sigma = 0.3$, $\kappa = 0.25$.

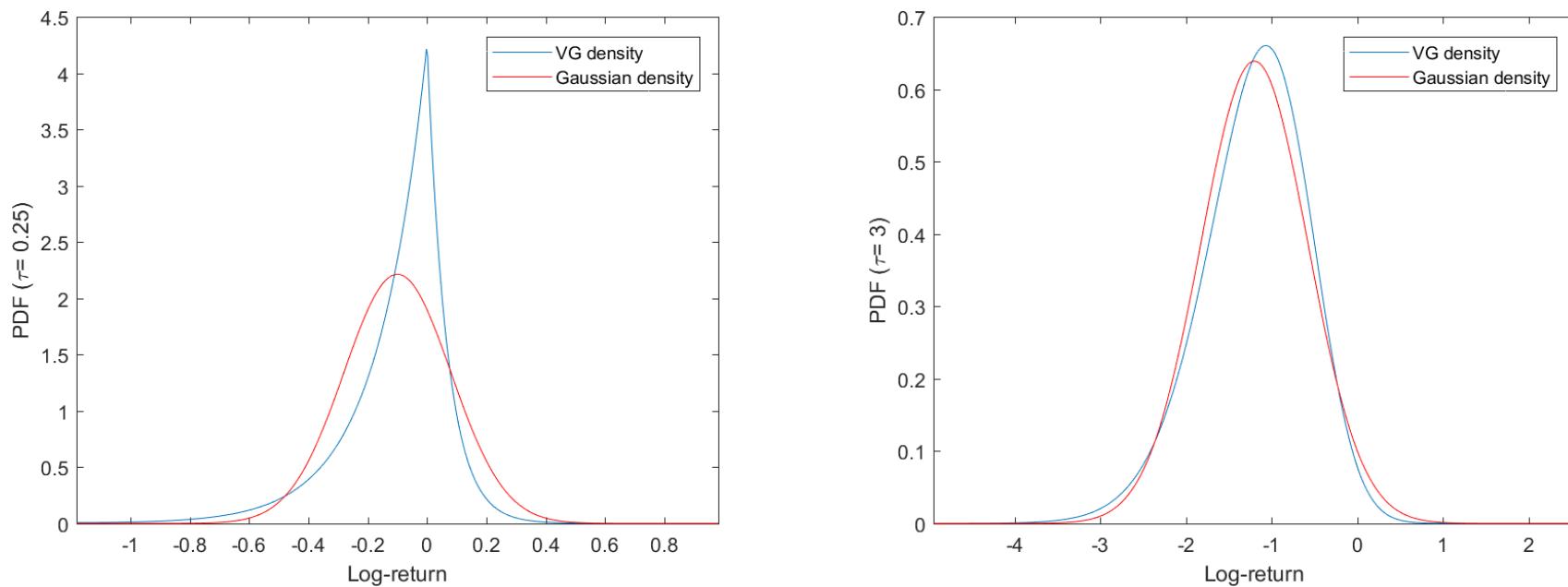


Figure 6.6: VG density and Gaussian with same mean and variance at different horizons: $\tau=3$ m (**Left**) and $\tau=5$ years (**Right**). Parameters: $\theta = -0.4$, $\sigma = 0.3$, $\kappa = 0.25$.

Matlab Code

```
%%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%%
%%% SIMULATING THE VARIANCE GAMMA %%%%
%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%%
%Assigning the number of simulated paths
%(nsimul), time to maturity (expiry), number of steps
%(nsteps), time step (dt) and observation times (timestep):
clear all;
nsimul=50; expiry=1; nsteps=250; dt=expiry/nsteps;
timestep=[0:dt:expiry]';
%Assigning parameters
theta=-0.4; sigma=0.3; kappa=0.25;
%Simulate increments of the Gamma process:
dG=gamrnd(dt/kappa,kappa,[nsteps,nsimul]);
%Simulate increments of the ABM on the Gamma clock scale
dX=theta*dG+sigma*sqrt(dG).*randn(nsteps,nsimul);
%Simulate VG process (use cumulative sum of the increments):
cdX=zeros(1,nsimul); cumsum(dX);
%Plot simulated paths:
h=figure('Color',[1 1 1]);
plot(timestep, cdX)
title('Simulated Paths of the VG Process')
xlabel('Time (years)')
print(h,'-dpng','FigVGPaths.png')
```

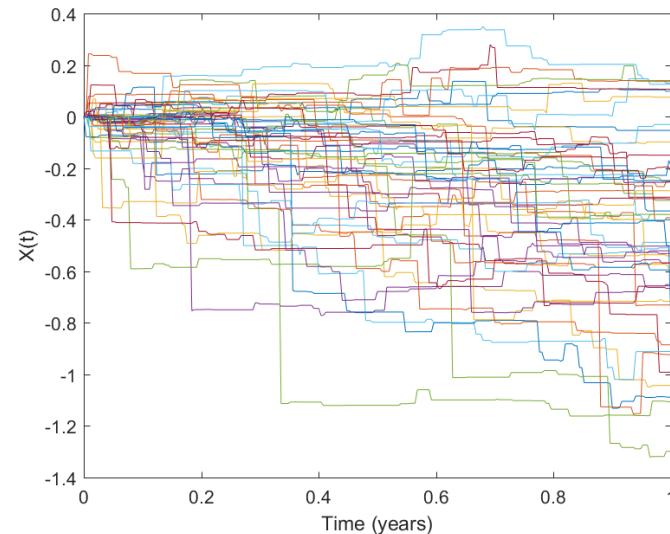


Figure 6.7: Simulated paths of the VG process. Parameters: $\theta = -0.4$, $\sigma = 0.3$, $\kappa = 0.25$.

VARIANCE GAMMA VG(θ, σ, k): FACTS

The solution

$$X(t) = \theta G(t) + \sigma W(G(t)), \quad \theta \in \mathbb{R}, \sigma > 0$$

where $G(t)$ is a Gamma process with parameters $\alpha = \lambda = k^{-1}$, $k > 0$.

The distribution of $X(t)$

$$2 \frac{e^{\theta x/\sigma^2}}{k^{t/k} \sigma \sqrt{2\pi} \Gamma(t/k)} \left(\frac{x^2}{\theta^2 + 2\sigma^2/k} \right)^{\frac{t}{2k}-\frac{1}{4}} K_{\frac{t}{k}-\frac{1}{2}} \left(\frac{|x|}{\sigma^2} \sqrt{\theta^2 + 2\sigma^2/k} \right).$$

with $K_v(z)$ denoting the modified Bessel function of the third kind with order v and argument z .

The characteristic function of $X(t)$

$$\phi_X(u; t) = \left(1 - iu\theta k + u^2 \frac{\sigma^2}{2} k \right)^{-\frac{t}{k}}.$$

The mean of $X(t)$

$$\mathbb{E}(X(t)) = \theta t.$$

The variance of $X(t)$

$$\mathbb{V}ar(X(t)) = (\sigma^2 + \theta^2 k) t.$$

The skewness of $X(t)$

$$\frac{(3\sigma^2 + 2\theta^2 k)\theta k}{(\sigma^2 + \theta^2 k)^{3/2} \sqrt{t}}.$$

The excess kurtosis of $X(t)$

$$\frac{(3\sigma^4 + 12\sigma^2\theta^2 k + 6\theta^4 k^2)k}{(\sigma^2 + \theta^2 k)^2 t}.$$

For large t the skewness and the excess kurtosis go to 0.

Option pricing with VG process

The Variance Gamma has been introduced and applied to option pricing by Madan and Seneta (1990), Madan and Milne (1991) and Madan et al (1998). Madan et al (1998) also report an option pricing formula resembling the Black-Scholes formula, but the probabilities are now special functions expressed in form of indefinite integrals.

Fourier inversion methods, such as the COS method (which will be introduced and discussed in details in the Appendix), allow for fast and accurate pricing of vanilla European calls and puts. In particular, we can explore the impact of the relevant parameters on the shape of the implied volatility.

Results are reported in Figures 6.8 - 6.11, in which we also revisit the behaviour of the density function. We observe the following.

- As previously observed, $\theta \in \mathbb{R}$ controls the sign of the skewness index, so that the VG process has distribution skewed to the left if $\theta < 0$, and skewed to the right if $\theta > 0$. If $\theta = 0$, the process has symmetric distribution. This reflects in the behaviour of the implied volatility in that, by changing θ , the slope of the smile varies from negative skew ($\theta < 0$), to symmetric smile ($\theta = 0$), and positive skew ($\theta > 0$), see figures 6.8 and 6.9.
- The parameter $k > 0$ controls instead the level of excess kurtosis. For a fixed ‘sign’ of the smile slope, changing k translates into steeper volatility slopes, see figures 6.8 and 6.9.
- The parameter $\sigma > 0$ controls the level of the smile. Larger this parameter, larger the implied volatility, see figures 6.12 and 6.11.

6.4. FINAL REMARK: LÉVY PROCESSES

- All the processes presented in this section share the feature of independent and stationary increments.
- The Brownian motion shares the same feature as well.
- They differ in the distribution chosen to model these increments.
- A process with independent and stationary increments is called Lévy process.

- Lévy processes are widely used in financial applications.
- All processes can be enriched in terms of features they can capture by assuming, for example, time dependent parameters, or by using more complex processes as stochastic clocks.
- For example: the instantaneous volatility of any of the processes presented above is constant. This assumption can be relaxed by assuming time dependent parameters. However, the resulting more general process will no longer have independent and stationary increments. This is the case of the Heston model.
- The main problem with Lévy processes is that they cannot capture the volatility clustering effects, which can be captured by other models such as stochastic volatility models.
- Lévy processes and stochastic volatility models complement each other: jump processes have a relative advantage in analytical tractability and they better capture short-term behaviour of financial time series, whilst stochastic volatility models have a richer time dependence structure and are more useful to model long term behaviour.
- Figure 6.14 compares the MJD, the VG and the Gaussian PDF at different horizons (1, 10, 20 and 90 days). We can see that the three densities approach one to the other as time horizon lengthens. This is due to the nature of independence of the increments of the three processes. This allows the central limit theorem to operate: in practice the skewness and the kurtosis fade away very quickly.

6.5. JUMPS VERSUS STOCHASTIC VOLATILITY

- The class of jump-diffusion models augments the ABM returns distribution with a Poisson-driven jump process. The class of subordinated Brownian motion models instead generates jumps by means of the subordinator.
- The class of stochastic volatility models extends the ABM model by allowing the volatility of the return process to itself evolve randomly over time.

- For a given maturity, each class of models could be made consistent with observed degrees of deviation from the ABM model.
- It is less obvious whether the theoretical predictions of either class of models are or can be made consistent with the observed term structures of these deviations.
- The term structures of skewness, kurtosis, and the implied volatility smile that can result in either class of models provides a simple method for distinguishing between the two classes of models
- As observed above, jump-diffusions and subordinated Brownian motions models can generate realistic implied volatility smiles at short maturities, but not at long maturities: the implied volatility smile flattens out too quickly.
- The term structure of implied volatilities of at-the-money forward options in a jump-diffusion and subordinated Brownian motion model is always an increasing function of the time-to-maturity, whilst decreasing or non-monotone term structure patterns frequently arise in practice.
- Stochastic volatility models are not capable of generating high levels of skewness and kurtosis at short maturities under reasonable parametrizations.
- As a consequence, stochastic volatility models cannot generate implied volatility smiles as sharp as those typically observed empirically.
- Moreover, conditional skewness and kurtosis in stochastic volatility models are always hump-shaped in the length of the horizon; indeed, for plausible parameter values, both quantities must be increasing over short to moderate maturities. This implies, contrary to the data, that the smile does not flatten out appreciably as maturity increases.
- On the question of the term structure of implied volatilities of at-the-money forward options, however, stochastic volatility models do well. A variety of patterns are possible in this model: increasing, decreasing, or even non-monotone (for example, U-shaped).
- Taking all this into account, the best solution appears to be the use of a model that combines time-varying volatility with jumps.

- Such a model would, for example, be able to generate adequate kurtosis at short maturities (through the jump component) and at moderate maturities (through the stochastic volatility component). It would also be able to account for a wide range of shapes for the term structure of implied volatilities of at-the-money forward options.
- First, it appears indisputable that the implied volatility smile in most markets is deepest at short maturities and flattens out monotonically as maturity increases.
- Second, a substantial literature has focused on an examination of the term structure of implied volatilities of at-the-money forward options. A variety of shapes (increasing, decreasing, and even sometimes non-monotone) for this curve have been documented in virtually all equity and currency markets.
- A simple and mathematically tractable way of achieving this goal is to consider a Lévy process evolving on a stochastic time scale, governed by some random clock, i.e. a time changed Lévy process.
- Examples of such processes have been studied in Carr et al. (2003), Carr and Wu (2004, 2007, 2016); their calibration performance has been extensively studied by Huang and Wu (2004). Further extensions have been proposed in Ballotta and Rayée (2017).

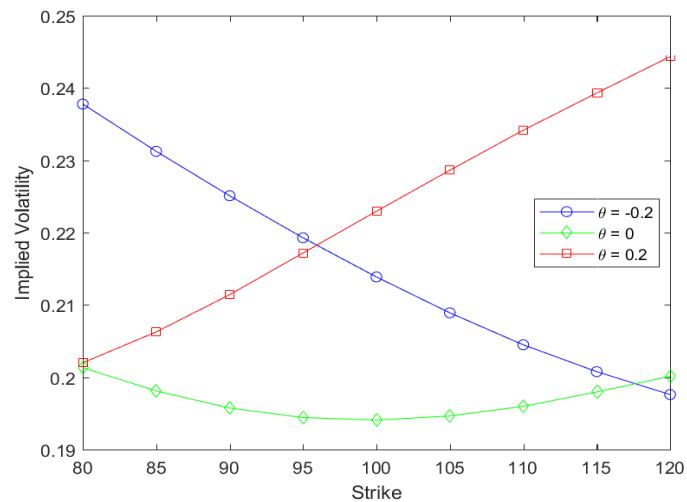


Figure 6.8: Black-Scholes implied volatility changing θ in the VG model. Parameters: $\kappa = 0.25; \sigma = 0.2; t = 1; r = 0.05; q = 0.02$.

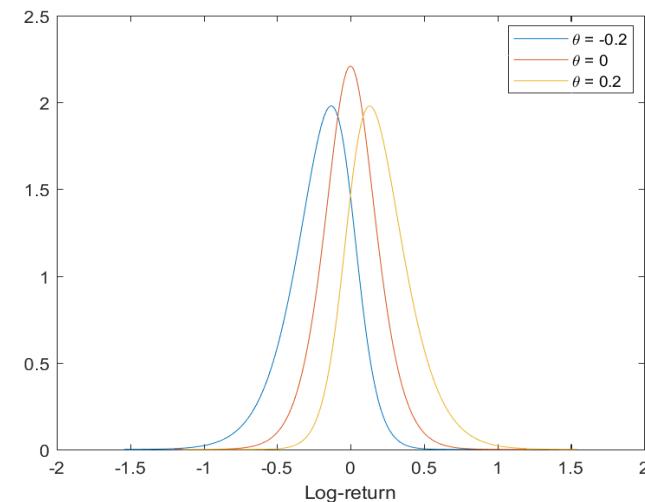


Figure 6.9: VG PDF changing θ . Parameters: $\kappa = 0.25; \sigma = 0.2; t = 1; r = 0.05; q = 0.02$.

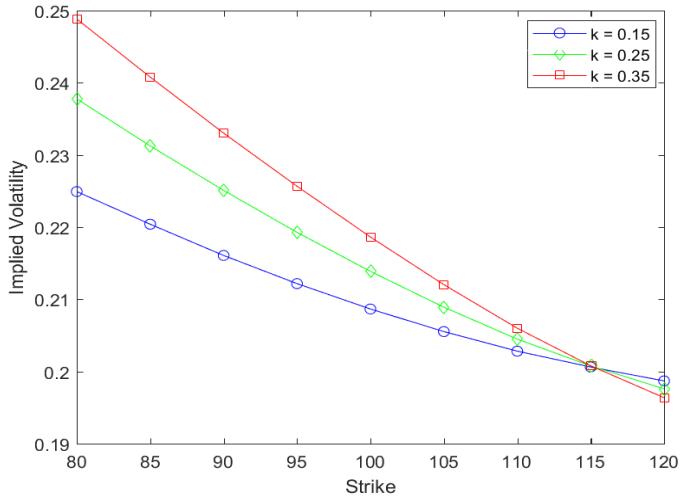


Figure 6.10: Black-Scholes implied volatility changing k in the VG model. Parameters: $\theta = -0.2; \sigma = 0.2; t = 1; r = 0.05; q = 0.02$.

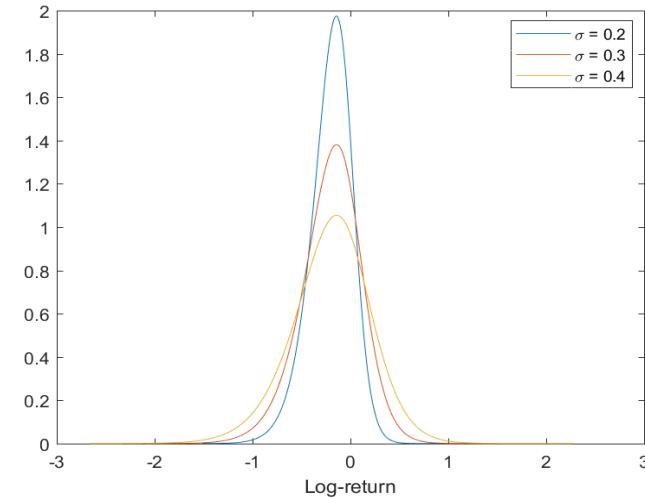


Figure 6.11: VG PDF changing σ . Parameters: $\theta = -0.2; k = 0.2; t = 1; r = 0.05; q = 0.02$.

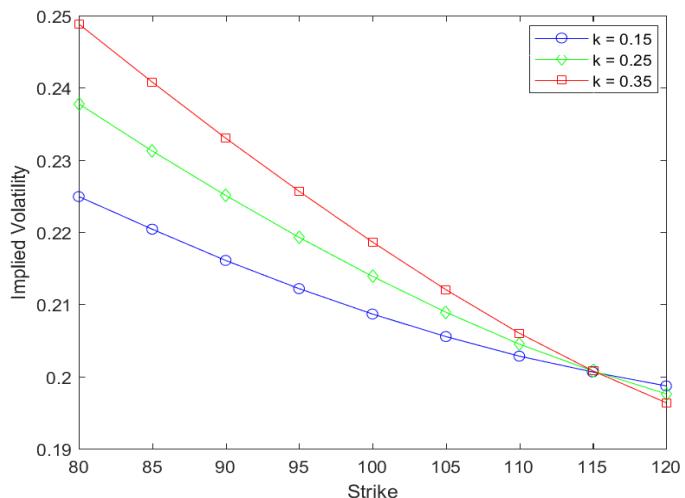


Figure 6.12: Black-Scholes implied volatility changing σ in the VG model. Parameters: $\theta = -0.2; k = 0.2; t = 1; r = 0.05; q = 0.02$.

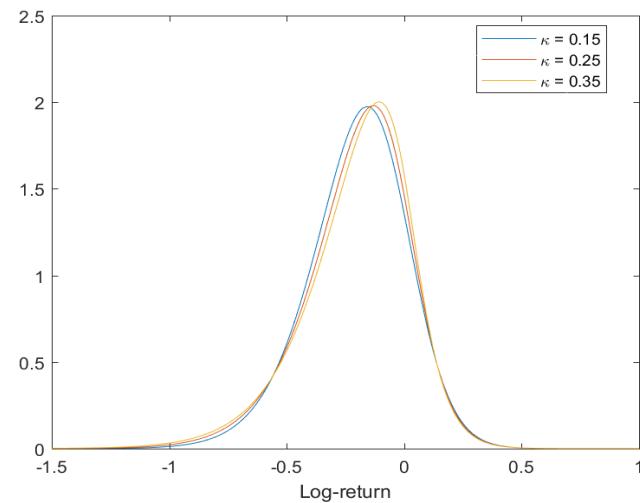


Figure 6.13: VG PDF changing k . Parameters: $\theta = -0.2; \sigma = 0.2; t = 1; r = 0.05; q = 0.02$.

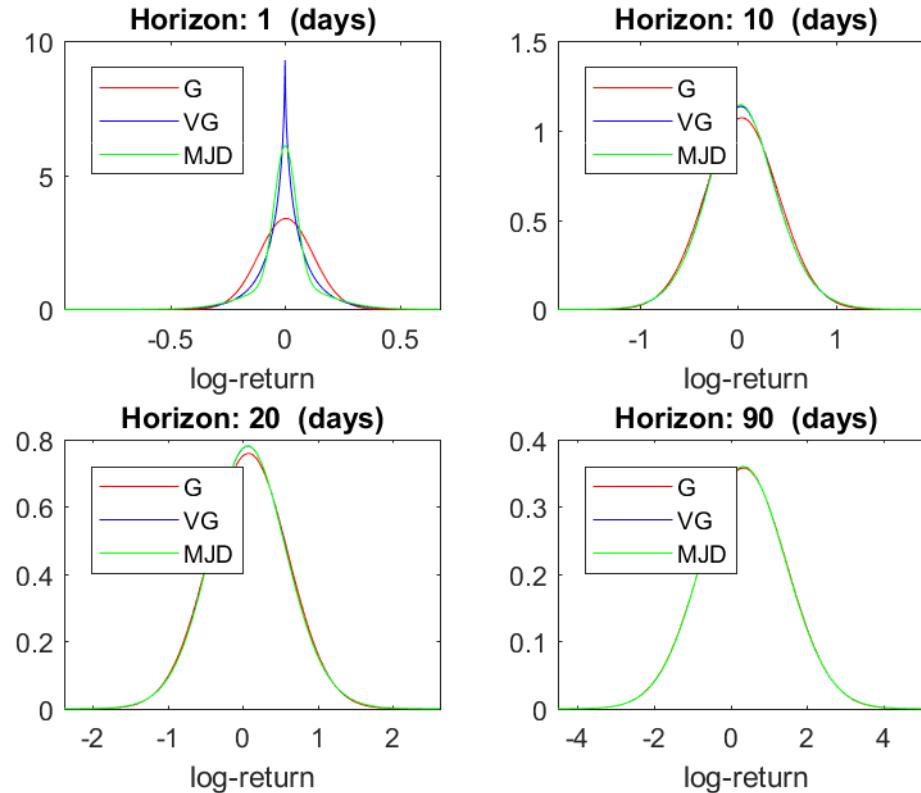


Figure 6.14: PDF of Gaussian (G), Variance-Gamma (VG) and Merton jump-diffusion (MJD) processes at different time horizons. Parameters are chosen to fit sample moments of daily log-returns of oil prices in 2012.

CHAPTER 7. CHANGES OF MEASURE

Contents

7.1 Preliminaries	192
7.2 Girsanov Theorem for Brownian motions	196
7.2.1 Application 1: Risk neutral measures	198
7.2.2 Application 2: The Numéraire Pair	200

Recent methods of derivative asset pricing rest on converting prices of such assets (after appropriate ‘normalization’) into martingales. This is done through transforming the underlying probability distribution using the tools provided by Girsanov Theorem, which we present in the following sections, and which hinge upon considering the basic “density process” of $\hat{\mathbb{P}}$ with respect to \mathbb{P} . The change of measure technique relates to a wide class of processes; thus, to help the intuition of this important result, we focus on Brownian motions. Extensions to more general cases are covered in advanced textbooks such as Cont and Tankov (2004).

7.1. PRELIMINARIES

Example 59 (A motivating intuition) Consider the following simple example.

- Assume that we want to measure the length of a room, and assume we express this measure using the metric system.
- It turns out that the room is 4.30m. long.
- Now assume that we want to change the reference system and express the length of the room in terms of the imperial system (i.e. in feet and inches).
- Then, the room is 14ft. long.
- Note that, in the process of switching from one reference system to the other, the room did not change: it did not shrink; it did not expand.
- The same applies to events and probability measures.

We note that probability measures serve as a tool to indeed ‘measure’ the likelihood of a given event. Hence, the same idea of having more than one reference system as in the previous example follows naturally. The fact that the event of which we want to quantify the likelihood does not change after switching reference system (like in the case of the room in the previous example) is formalized in the following.

Definition 60 (Absolutely continuous/equivalent probability measure) Given two probability measures \mathbb{P} and \mathbb{P}^* defined on the same σ -algebra \mathcal{F} , then.

- i) \mathbb{P} is absolutely continuous with respect to \mathbb{P}^* , i.e. $\mathbb{P} << \mathbb{P}^*$, if $\mathbb{P}(A) = 0$ whenever $\mathbb{P}^*(A) = 0$ for all $A \in \mathcal{F}$.
- ii) If $\mathbb{P} << \mathbb{P}^*$ and also $\mathbb{P}^* << \mathbb{P}$, then $\mathbb{P} \sim \mathbb{P}^*$, i.e. \mathbb{P} and \mathbb{P}^* are equivalent measures. Thus, for $\mathbb{P} \sim \mathbb{P}^*$ the following are equivalent.
 - $\mathbb{P}(A) = 0 \Leftrightarrow \mathbb{P}^*(A) = 0$ (same null sets)
 - $\mathbb{P}(A) = 1 \Leftrightarrow \mathbb{P}^*(A) = 1$ (same a.s. sets)
 - $\mathbb{P}(A) > 0 \Leftrightarrow \mathbb{P}^*(A) > 0$ (same sets of positive measures)

In the context of asset pricing, the above definition is implying that if the market is arbitrage free under the real world probability measure, then it is arbitrage free also under any other equivalent probability measure.

Example 61 Consider a closed interval $[a, b]$, for $0 \leq a \leq b \leq 1$, and consider the experiment of choosing a number from this interval. Define the following

$$\mathbb{P}(\text{the number chosen is in } [a, b]) = \mathbb{P}[a, b] := b - a.$$

But you can also define a different metrics \mathbb{P}^* , according to which

$$\mathbb{P}^*(\text{the number chosen is in } [a, b]) = \mathbb{P}^*[a, b] := b^2 - a^2.$$

As there is a conversion factor that helps you to switch between meters and feet, so that $4.30\text{m} = 14\text{ft}$, there is also a conversion factor between probability measures. If you want to convert meters in feet, you need a “bridge” between the two ($1\text{ ft} = 0.30\text{ meters}$). There is something equivalent to this also for probability measures and it is defined as follows.

Theorem 62 (Radon-Nikodým) If \mathbb{P} and \mathbb{P}^* are two probability measures on (Ω, \mathcal{F}) such that $\mathbb{P} \sim \mathbb{P}^*$, then there exists a random variable $Y \in \mathcal{F}$ such that

$$\mathbb{P}^*(A) = \int_A Y d\mathbb{P} = \mathbb{E}[Y 1_A], \forall A \in \mathcal{F}. \quad (7.1)$$

Y is called the Radon-Nikodým derivative of \mathbb{P}^* with respect to \mathbb{P} and is also written as

$$Y = \frac{d\mathbb{P}^*}{d\mathbb{P}} \quad (7.2)$$

Remark 63 (A matter of notation) Bearing in mind that for any random variable the CDF is such that

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(y) dy, \quad (7.3)$$

where $f(x)$ denotes the density function of X , then the density function can be written as

$$f_X(x) = \frac{dF_X}{dx} = \frac{d\mathbb{P}(\omega)}{dx} \quad \forall x \in \mathbb{R}.$$

Hence, the random variable Y defined in (7.2) is not a proper derivative but actually a likelihood ratio.

Example 64 Consider Example 61. Here we defined two metrics on the interval $[a, b]$, $0 \leq a \leq b \leq 1$

$$\begin{aligned} \mathbb{P}(\text{the number chosen is in } [a, b]) &= \mathbb{P}[a, b] := b - a \\ \mathbb{P}^*(\text{the number chosen is in } [a, b]) &= \mathbb{P}^*[a, b] := b^2 - a^2. \end{aligned}$$

We could be more specific and say that

$$\begin{aligned} \mathbb{P}[a, b] &= \int_a^b d\omega; \\ \mathbb{P}^*[a, b] &= \int_a^b 2\omega d\omega. \end{aligned}$$

Consequently $d\mathbb{P} = d\omega$ and $d\mathbb{P}^* = 2\omega d\omega$, from which it follows that $Y(\omega) = 2\omega$.

Example 65 Consider the usual probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a standard normal random variable X , i.e. $X \sim N(0, 1)$. Define a new random variable Y as $Y = X + \theta$, and let $\hat{\mathbb{P}}(A)$ be another equivalent probability measure on Ω , defined by

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = Z,$$

for

$$Z = e^{-\theta X - \frac{\theta^2}{2}}.$$

Then $Y \sim N(0, 1)$ on $(\Omega, \mathbb{F}, \hat{\mathbb{P}})$. In fact, let $A = \{Y \leq a\}$, $a \in \mathbb{R}$. Let us calculate $\hat{\mathbb{P}}(Y \leq a)$. Note that, by definition of random variable, for $A \in \mathbb{F}$

$$\begin{aligned}\hat{\mathbb{P}}(Y \leq a) &= \int_A Z d\hat{\mathbb{P}} = \int_{\Omega} 1_A Z d\hat{\mathbb{P}} \\ &= \int_{\mathbb{R}} 1_{\{Y \leq a\}} e^{-\theta x - \frac{\theta^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = e^{-\frac{\theta^2}{2}} \int_{-\infty}^{a-\theta} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x^2+2\theta x)}{2}} dx \\ &= \int_{-\infty}^{a-\theta} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+\theta)^2}{2}} dx = \int_{-\infty}^a \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy,\end{aligned}$$

where we made the change of dummy variable $y = x + \theta$ in the last step. Hence

$$\hat{\mathbb{P}}(Y \leq a) = \int_{-\infty}^a \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy,$$

as required.

Fact 66 *Theorem 62 implies that for any random variable X*

$$\mathbb{E}^*[X] = \int X d\mathbb{P}^* = \int XY d\mathbb{P} = \mathbb{E}[XY].$$

A similar relation holds for conditional expectations as well.

Fact 67 (Bayes formula) *Let \mathbb{P} and \mathbb{P}^* be two equivalent probability measures on the same measurable space (Ω, \mathcal{F}) and let*

$$Y = \frac{d\mathbb{P}^*}{d\mathbb{P}}$$

be the Radon-Nikodým derivative of \mathbb{P}^* with respect to \mathbb{P} . Furthermore, let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P}^*)$ such that $\mathbb{E}^* |X| < \infty$ and $\mathcal{G} \in \mathcal{F}$ a sub σ -algebra of \mathcal{F} . Then the following generalised version of the Bayes formula holds

$$\mathbb{E}^* [X | \mathcal{G}] = \frac{\mathbb{E} [XY | \mathcal{G}]}{\mathbb{E} [Y | \mathcal{G}]}.$$

Remark 68 Note that the previous expression can be rewritten as

$$\mathbb{E} [XY | \mathcal{G}] = \mathbb{E}^* [X | \mathcal{G}] \mathbb{E} [Y | \mathcal{G}];$$

in other words the change of measure also allows you to pretend that, conditioned on the available information, the model is devoid of dependence, provided you first operate a suitably chosen change of probability measure.

7.2. GIRSANOV THEOREM FOR BROWNIAN MOTIONS

In the previous section we have seen that the change of measure for the normal random variable is defined by the “conversion factor”

$$Z = e^{-\lambda X - \frac{\lambda^2}{2}}, \quad \lambda \in \mathbb{R}.$$

In other words the variable Z is the Radon-Nikodým derivative. For the case of changing measure to an entire process, we need to generalize a little this conversion factor; specifically, we define the Radon-Nikodým derivative process (or density process)

$$Z(t) = \mathbb{E} [Z | \mathcal{F}_t], \quad 0 \leq t \leq T.$$

From the basic properties of the conditional expectation, it follows that the process Z is a \mathbb{P} -martingale with $\mathbb{E} [Z(t)] = 1$ for $0 \leq t \leq T$. The density process Z is then used for changing measures to the Brownian motion.

Fact 69 (Girsanov Theorem) Let W be a \mathbb{P} -Brownian motion and $\lambda(t)$ an adapted process with $\mathbb{E} \left(\exp \left(\int_0^t \lambda(s)^2 ds \right) \right) < \infty$ for all $0 \leq t \leq T$. Define the processes $\hat{W}(t), Z(t)$ by

$$\begin{aligned} \hat{W}(t) &= W(t) + \int_0^t \lambda(s) ds \\ Z(t) &= e^{-\frac{1}{2} \int_0^t \lambda(s)^2 ds - \int_0^t \lambda(s) dW(s)} \end{aligned}$$

for all $0 \leq t \leq T$, and define a probability measure $\hat{\mathbb{P}}$ by

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = Z(t).$$

Then \hat{W} is a $\hat{\mathbb{P}}$ -Brownian motion.

Proof. We need to check that \hat{W} is a Brownian motion. To this purpose we use the Lévy characterization of Brownian motions, which means that we need to verify that \hat{W} is a $\hat{\mathbb{P}}$ -martingale with quadratic variation equal to t . To this purpose, set $M(t) = Z(t)\hat{W}(t)$. And then proceed as follows.

- Apply Itô's formula to obtain

$$\begin{aligned} dM(t) &= \hat{W}(t)dZ(t) + Z(t)d\hat{W}(t) + dZ(t)d\hat{W}(t) \\ &= -\hat{W}(t)(Z(t)\lambda(t)dW(t)) + Z(t)dW(t) \\ &= Z(t)\alpha(t)dW(t), \end{aligned}$$

where $\alpha(t) = 1 - \lambda(t)\hat{W}(t)$.

- From this result, it follows that $M(t) = Z(t)\hat{W}(t)$ is a \mathbb{P} -martingale.
- If $\hat{\mathbb{E}}$ denotes the expectation under $\hat{\mathbb{P}}$, and bearing in mind that Z is a \mathbb{P} -martingale, it follows from the Bayes formula that

$$\hat{\mathbb{E}} [\hat{W}(t) | \mathcal{F}(s)] = \frac{\mathbb{E} [Z(t)\hat{W}(t) | \mathcal{F}(s)]}{Z(s)} = \hat{W}(s).$$

- Hence \hat{W} is a $\hat{\mathbb{P}}$ -martingale.
- Further $[\hat{W}]_t = [W]_t = t$.
- Hence, by the Lévy characterization of Brownian motion we conclude that \hat{W} is a standard Brownian motion.

If we reduce the theorem to its bare terms, we understand the following.

- We can construct new probability measures by simply shifting the Brownian motion (i.e. changing the mean of its distribution) of a quantity which depends on the function λ , known as the Girsanov exponent.
- Note that this function is not specified by the theorem. This means that we can consider λ as a ‘degree of freedom’ to characterize the probability measure in order to achieve whatever goal we set.
- For pricing purposes, the Fundamental Theorem of Asset Pricing implies that pricing has to be carried out under a probability measure \hat{P} such that discounted asset prices satisfy the martingale property; this is our rule for choosing the function λ whenever we require to switch from the real probability measure to the risk neutral one.
- Finally, we emphasize the relation between the shift in the Brownian motion and the Girsanov exponent in the density process

$$\begin{aligned}\hat{W}(t) &= W(t) + \int_0^t \lambda(s) ds \\ Z(t) &= e^{-\frac{1}{2} \int_0^t \lambda(s)^2 ds - \int_0^t \lambda(s) dW(s)}\end{aligned}$$

7.2.1. APPLICATION 1: RISK NEUTRAL MEASURES

- Let the risk-free rate of interest at time t to be $r(t)$, so that $B(t) = \exp\left(\int_0^t r(u) du\right)$ is the value at t of £1 invested today at the risk free rate.
- A risk-neutral probability measure is a probability measure \hat{P} such that the discounted price of any tradeable asset behaves as a martingale with respect to \hat{P} .
- If a risk-neutral measure exists, then the value at time s of a (possibly random) amount v paid at time t ($t > s$) is

$$V(s) = B(s) \hat{\mathbb{E}}[B(t)^{-1} v | \mathcal{F}(s)].$$

- In other words $B(t)^{-1}V(t)$ is a martingale under the risk neutral measure $\hat{\mathbb{P}}$.
- To find the risk-neutral measure $\hat{\mathbb{P}}$ we proceed as follows.
- Suppose that the price $S(t)$ of a non-dividend paying share satisfies

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t)$$

- Define $Y(t) = B(t)^{-1}S(t)$.
- Then

$$\begin{aligned} dY(t) &= -r(t)Y(t)dt + (\mu(t)dt + \sigma(t)dW(t))Y(t) \\ &= \sigma(t)Y(t)(\lambda(t)dt + dW(t)), \end{aligned}$$

where $\lambda(t) = (\mu(t) - r(t)) / \sigma(t)$ is the market price of risk.

- Now we see that

$$dY(t) = \sigma(t)Y(t)d\hat{W}(t)$$

for

$$\hat{W}(t) = W(t) + \int_0^t \lambda(s)ds.$$

- Therefore $Y(t)$ is a $\hat{\mathbb{P}}$ -martingale as long as \hat{W} is a $\hat{\mathbb{P}}$ -Brownian motion.
- The Girsanov theorem tells us that this is the case when the density process is

$$\exp \left[- \int_0^T \lambda(t)dW(t) - \frac{1}{2} \int_0^T \lambda(t)^2 dt \right],$$

provided all required integrability conditions are satisfied.

7.2.2. APPLICATION 2: THE NUMÉRAIRE PAIR

The risk neutral valuation principle essentially states that securities in the market can be priced as conditional expectations, under the risk neutral probability measure, of the discounted payoff at maturity. The discount factor is the risk free rate of interest and conditioning is operated using the information set available at the time in which valuation is performed. Hence, referring to the usual notation,

$$\frac{C(t)}{B(t)} = \hat{\mathbb{E}} \left(\frac{C(T)}{B(T)} \middle| \mathcal{F}(t) \right).$$

We observe the following

- The above stated pricing criterion makes use of security prices discounted at the risk free rate or, in more general terms, security prices “normalized” by the money market account.
- In other words, the risk neutral valuation principle applies to security prices which have been expressed in terms of the price of the money market account at the time of valuation.
- Technically speaking, this classifies the money market account as a numéraire.

Fact 70 *A numéraire is an economic term that represents a unit of account. A numéraire is usually applied to a single good, which becomes the base good. All similar goods are then valued and priced against the base good. This comparison makes it possible to identify which goods are worth more than others.*

Hence, if security B is the chosen numéraire, the relative price of security S is

$$\tilde{S} = \frac{S}{B};$$

this is what we also call normalized price.

Remark 71 *As the risk neutral probability measure is defined as the one under which security prices normalized using the money market account are martingales, then the risk neutral probability measure is strictly connected to the choice of the money market account as numéraire. The pair $(B(t), \hat{\mathbb{P}})$ is what goes under the name of (risk neutral) numéraire pair.*

- It is important to notice that the money market account is actually not the only asset in the market which can be chosen as numéraire.
- This implies that the normalization operated above, and therefore the risk neutral probability numéraire pair, is not the only one possible.
- Indeed it is possible to choose almost any asset in the market as a numéraire and consequently, select a suitable probability measure under which normalized security prices (where normalization is operated with respect to this numéraire) are martingales.
- In the context of derivative pricing, the choice of different numéraire pairs can simplify the actual calculations in situations in which there are more than one source of randomness in the market.

Fact 72 (Geman, El Karoui and Rochet, 1995) *Let us assume that market prices are expressed relative to the numéraire $B(t)$ (such that $B(0) = 1$) to which the probability $\hat{\mathbb{P}}$ is associated. Let $X(t)$ be a non-dividend paying numéraire such that $B(t)^{-1}X(t)$ is a $\hat{\mathbb{P}}$ -martingale. Then, there exists a probability measure $\mathbb{P}^{(X)}$ defined by the Radon-Nikodm derivative*

$$\eta(t) = \frac{d\mathbb{P}^{(X)}}{d\hat{\mathbb{P}}} \Big|_{\mathcal{F}_t} := \frac{X(t)B(0)}{B(t)X(0)},$$

such that

- i) the primary security prices discounted with respect to X are $\mathbb{P}^{(X)}$ -martingales;
- ii) if a contingent claim Z has fair price under $(B(t), \hat{\mathbb{P}})$, then it has fair price under $(X(t), \mathbb{P}^{(X)})$ and the prices are the same.

The pair $(X(t), \mathbb{P}^{(X)})$ is called the numéraire pair.

Indeed we observe the following.

- The process $\eta(t)$ is a martingale under the risk neutral measure $\hat{\mathbb{P}}$ by construction.

- Let S denote the price of some traded security. In virtue of the Bayes formula in Fact 67 for $s < t$

$$\begin{aligned}\mathbb{E}_s^{(X)}\left(\frac{S(t)}{X(t)}\right) &= \hat{\mathbb{E}}_s\left(\frac{S(t)}{X(t)}\eta(t)\right)\eta(s)^{-1} \\ &= \hat{\mathbb{E}}_s\left(\frac{S(t)}{X(t)}\frac{X(t)B(0)}{B(t)X(0)}\right)\frac{X(0)B(s)}{B(0)X(s)} \\ &= \frac{S(s)}{X(s)}\end{aligned}$$

as stated in part (i) of Fact 72.

- Further, let consider the contingent claim Z with maturity T , then risk neutral valuation implies

$$\begin{aligned}Z(t) &= B(t)\hat{\mathbb{E}}_t\left(\frac{Z(T)}{B(T)}\right) \\ &= B(t)\hat{\mathbb{E}}_t\left(\frac{Z(T)}{B(T)}\eta(T)\frac{X(0)B(T)}{B(0)X(T)}\right) \\ &= \frac{X(0)B(t)}{B(0)}\hat{\mathbb{E}}_t\left(\frac{Z(T)}{X(T)}\eta(T)\right) \\ &= \frac{X(0)B(t)}{B(0)}\mathbb{E}_t^{(X)}\left(\frac{Z(T)}{X(T)}\right)\hat{\mathbb{E}}_t(\eta(T)) \\ &= X(t)\mathbb{E}_t^{(X)}\left(\frac{Z(T)}{X(T)}\right)\end{aligned}$$

where the last but one equality follows from the Bayes formula in Fact 67 and the final result follows from substituting for $\eta(t)$.

- This is what stated in part (ii) of Fact 72.

More in general, the following holds.

Fact 73 (Change of numéraire) *Let X, Y be numéraires satisfying the assumptions of the previous theorem, and Z a contingent claim. The change of numéraire formula states*

$$X(t)\mathbb{E}^{(X)}\left(\frac{Z(T)}{X(T)} \middle| \mathbb{F}_t\right) = Y(t)\mathbb{E}^{(Y)}\left(\frac{Z(T)}{Y(T)} \middle| \mathbb{F}_t\right).$$

Now, we show how to use the change of numéraire technique. The aim is mainly to simplify calculations.

Example: Revisiting Black-Scholes

- Consider the problem of pricing a European call option in the Black-Scholes economy.
- Risk-neutral valuation implies that

$$C(t) = B(t)\hat{\mathbb{E}}\left[B(T)^{-1}(S(T) - K)^+ \mid \mathbb{F}_t\right].$$

- Let 1_A be the indicator function of the set A ; then

$$\begin{aligned} C(t) &= B(t)\hat{\mathbb{E}}\left[B(T)^{-1}(S(T) - K)1_{(S(T)-K>0)} \mid \mathbb{F}_t\right] \\ &= B(t)\hat{\mathbb{E}}\left[B(T)^{-1}S(T)1_{(S(T)-K>0)} \mid \mathbb{F}_t\right] - e^{-r(T-t)}K\hat{\mathbb{P}}\left[1_{(S(T)-K>0)} \mid \mathbb{F}_t\right]. \end{aligned}$$

- Then the pricing equation can be actually rewritten as

$$C(t) = B(t)\hat{\mathbb{E}}\left[B(T)^{-1}S(T)1_{(S(T)-K>0)} \mid \mathbb{F}_t\right] - e^{-r(T-t)}K\hat{\mathbb{P}}\left(S(T) > K \mid \mathbb{F}_t\right). \quad (7.4)$$

- The first conditional expectation though is not trivial to solve due to the dependence between $S(T)$ and the random variable $1_{(S(T)-K>0)}$.

- However, from Remark 68, we know it is possible to ‘ignore’ dependence, provided we appropriately change the measure, i.e. the numéraire.
- Hence let

$$\gamma(T) = \frac{d\mathbb{P}^{(S)}}{d\hat{\mathbb{P}}}\Big|_{\mathbb{F}_t} = \frac{S(T)}{B(T)S(0)}$$

- The Bayes theorem implies that equation (7.4) can be rewritten as

$$C(t) = S(t)\mathbb{P}^{(S)}(S(T) > K \mid \mathcal{F}(t)) - e^{-r(T-t)}K\hat{\mathbb{P}}(S(T) > K \mid \mathcal{F}(t)).$$

- Further,

$$\begin{aligned}\hat{\mathbb{P}}(S(T) > K \mid \mathcal{F}(t)) &= \hat{\mathbb{P}}\left(S(t)e^{\left(r-\frac{\sigma^2}{2}\right)(T-t)+\sigma\hat{W}(T-t)} > K\right) \\ &= N(d_2); \\ d_2 &= \frac{\ln \frac{S(t)}{K} + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\end{aligned}$$

- As far as

$$\mathbb{P}^{(S)}(S(T) > K \mid \mathcal{F}(t))$$

is concerned, we need to work out first the quantity by which the Brownian motion is shifted with the change of measure/numéraire. Consider the Radon-Nikodým derivative

$$\gamma(T) = \frac{S(T)}{B(T)S(0)} = e^{-\frac{\sigma^2}{2}T+\sigma\hat{W}(T)}.$$

- The Girsanov theorem implies that

$$W^{(S)}(t) = \hat{W}(t) - \sigma t$$

is a $\mathbb{P}^{(S)}$ -Brownian motion.

- The above implies

$$\begin{aligned}
 \mathbb{P}^{(S)}(S(T) > K \mid \mathcal{F}(t)) &= \mathbb{P}^{(S)}\left(S(t)e^{\left(r-\frac{\sigma^2}{2}\right)(T-t)+\sigma\hat{W}(T-t)} > K \mid \mathcal{F}(t)\right) \\
 &= \mathbb{P}^{(S)}\left(S(t)e^{\left(r-\frac{\sigma^2}{2}\right)(T-t)+\sigma(W^S(T-t)+\sigma(T-t))} > K \mid \mathcal{F}(t)\right) \\
 &= \mathbb{P}^{(S)}\left(S(t)e^{\left(r+\frac{\sigma^2}{2}\right)(T-t)+\sigma W^S(T-t)} > K\right).
 \end{aligned}$$

This follows by the property of independent increments. Using the same argument as before

$$\begin{aligned}
 \mathbb{P}^{(S)}\left(S(t)e^{\left(r+\frac{\sigma^2}{2}\right)(T-t)+\sigma W^S(T-t)} > K\right) &= N(d_1) \\
 d_1 &= \frac{\ln \frac{S_t}{K} + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}
 \end{aligned}$$

- Consequently

$$\begin{aligned}
 C(t) &= S(t)N(d_1) - Ke^{-r(T-t)}N(d_2), \\
 d_{1,2} &= \frac{\ln \frac{S(t)}{K} + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}},
 \end{aligned}$$

as we would expect.

- $\mathbb{P}^{(S)}$ is the spot probability measure.

Example: The Delta of the call option price

- The Delta of the call option price can be calculated as

$$\Delta_{call} = N(d_1).$$

- This can also be written as

$$\Delta_{call} = \hat{\mathbb{E}} \left[e^{-rT} \frac{S(T)}{S(0)} \mathbf{1}_{(S(T)>K)} \right] = \mathbb{P}^{(S)} (S(T) > K),$$

where the last equality follows from the change of numéraire.

Example: The generalized framework

- Consider the problem of pricing a European call option via risk-neutral valuation as in Example 7.2.2 but with stochastic rates of interest.
- The money market account is then defined as

$$B(t) = e^{\int_0^t r(s) ds}$$

where r is the short rate process.

- Similarly to the previous example, we can write the option price as

$$C(t) = B(t) \hat{\mathbb{E}} \left[B(T)^{-1} S(T) \mathbf{1}_{(S(T)-K)} \mid \mathbb{F}_t \right] - B(t) K \hat{\mathbb{E}} \left[B(T)^{-1} \mathbf{1}_{(S(T)-K)} \mid \mathbb{F}_t \right]. \quad (7.5)$$

- The above stated conditional expectations would be simpler to solve if we could ‘pretend’ all variables were independent. We know from Remark 68 that this can be achieved by changing measures and numéraires.
- The first conditional expectation can be dealt with by switching to the Spot measure as in Example 7.2.2.
- For the second conditional expectation, let $P(t; \tau)$ be the price at time t of a zero coupon bond expiring at time τ .
- Then set

$$\eta^{(T)}(T) = \frac{d\mathbb{P}^{(T)}}{d\hat{\mathbb{P}}} \Big|_{\mathbb{F}_t} = \frac{P(T; T)}{B(T)P(0; T)}.$$

- The Bayes theorem implies that equation (7.5) can be rewritten as

$$C(t) = S(t)\mathbb{P}^{(S)}(S(T) > K \mid \mathbb{F}_t) - K P(t; T)\mathbb{P}^{(T)}(S(T) > K \mid \mathbb{F}_t). \quad (7.6)$$

- The conditional probabilities can then be calculated once the corresponding changes of measure for the Brownian motion are determined.
- To this purpose, assume that the $\hat{\mathbb{P}}$ -dynamic of the bond price is

$$dP(t; T) = r(t)P(t; T)dt + m(t, T)P(t; T)d\hat{Z}(t),$$

where \hat{Z} is a $\hat{\mathbb{P}}$ -Brownian motion.

- Then

$$\eta^{(T)}(T) = e^{-\int_0^T \frac{m^2(t, T)}{2} dt + \int_0^T m(t, T)d\hat{Z}(t)}.$$

- The Girsanov's theorem implies that

$$Z^{(T)}(t) = \hat{Z}(t) - \int_0^t m(s, T)ds$$

is a $\mathbb{P}^{(T)}$ -Brownian motion.

- $\mathbb{P}^{(T)}$ is the forward-risk-adjusted probability measure.

Example: the Exchange option - Margrabe option

- Let r be the constant risk-free rate of interest and S and A be two risky securities described as

$$\begin{aligned} dS(t) &= rS(t)dt + \sigma_S S(t)d\hat{W}(t), \\ dA(t) &= rA(t)dt + \sigma_A A(t)d\hat{X}(t), \end{aligned}$$

for \hat{W} and \hat{X} correlated standard Brownian motions under the risk-neutral probability measure $\hat{\mathbb{P}}$, i.e.

$$d\hat{W}(t)d\hat{X}(t) = \rho dt,$$

where ρ is the correlation coefficient.

- An exchange option is the right (but not the obligation) to exchange one share in stock A for one share in stock S at a specified time T , i.e.

$$C(T) = \max [S(T) - A(T), 0].$$

- A convenient way to price the Exchange option is to rewrite the payoff as

$$C(T) = A(T) \left(\frac{S(T)}{A(T)} - 1 \right)^+.$$

- This shows that the contract is - up to the factor $A(T)$ - a call option with unit strike price written on the normalized price $S(T)/A(T)$.
- This observation suggests to use the asset $A(T)$ as numéraire, by defining a probability measure $\mathbb{P}^{(A)}$ equivalent to the risk neutral measure using the density process

$$\gamma(T) = \frac{A(T)}{B(T)A(0)}.$$

- The Girsanov theorem implies that

$$\begin{aligned} X^{(A)}(t) &= \hat{X}(t) - \sigma_A t \\ W^{(A)}(t) &= \hat{W}(t) - \rho\sigma_A t \end{aligned}$$

are a $\mathbb{P}^{(A)}$ -Brownian motions.

- Due to the dependence in place, from Section 2.1.4 it follows that

$$\hat{W}(t) = \rho \hat{X}(t) + \sqrt{1 - \rho^2} Z(t)$$

for a Brownian motion $Z(t)$ independent of $\hat{X}(t)$.

- Then proceed as follows

$$\begin{aligned}\hat{W}(t) &= \rho \hat{X}(t) + \sqrt{1 - \rho^2} Z(t) \\ &= \rho \left(X^{(A)}(t) + \sigma_A t \right) + \sqrt{1 - \rho^2} Z(t).\end{aligned}$$

- From which it follows that

$$\hat{W}(t) - \rho \sigma_A t = \rho X^{(A)}(t) + \sqrt{1 - \rho^2} Z(t).$$

- Set $W(t)^{(A)} = \hat{W}(t) - \rho \sigma_A t$, the Girsanov theorem implies this is a Brownian motion under $\mathbb{P}^{(A)}$, correlated with $X^{(A)}(t)$ with correlation coefficient ρ .
- Then, from Example 34, the ratio $R(t) = S(t)/A(t)$ has dynamics

$$dR(t) = \sigma_S R(t) dW_t^{(A)}(t) - \sigma_A R(t) dX^{(A)}(t).$$

- Set

$$\tilde{W}(t) = \frac{\sigma_S dW_t^{(A)}(t) - \sigma_A dX^{(A)}(t)}{\sqrt{\sigma_S^2 + \sigma_A^2 - 2\rho\sigma_S\sigma_A}},$$

and rewrite $R(t)$ in terms of $\tilde{W}(t)$, i.e.

$$dR(t) = \sqrt{\sigma_S^2 + \sigma_A^2 - 2\rho\sigma_S\sigma_A} d\tilde{W}(t).$$

- The price of the exchange option, in virtue of the Bayes formula in Fact 67, is now

$$C(t) = A(t) \mathbb{E}_t^{(A)} \left((R(T) - 1)^+ \right).$$

- By the Black Scholes formula, the price of the exchange option is

$$C(t) = S(t)N(d_1) - A(t)N(d_2)$$

where

$$\begin{aligned} d_1 &= \frac{\ln \frac{S(t)}{A(t)} + \frac{\sigma^2}{2} T}{\sigma \sqrt{T}}; \quad d_2 = d_1 - \sigma \sqrt{T} \\ \sigma^2 &= \sigma_S^2 + \sigma_A^2 - 2\rho\sigma_S\sigma_A. \end{aligned}$$

Finally, in figure 7.1 we represent the PDF and CDF for log-returns and prices using different probability measures: real world, risk-neutral, spot (the asset itself is the numéraire as in Example 7.2.2 - the Black Scholes model revisited) and spot (when another asset, say A, is taken as numéraire as in Example 7.2.2 - the exchange option). It is evident as the change of measure consists of a change of the mean in the distribution of log-returns; however, this only applies when the underlying distribution is Gaussian (i.e. in the case of the Brownian motion). One simulated path for the log-level and the level under the different probability measures is given in figures 7.2.

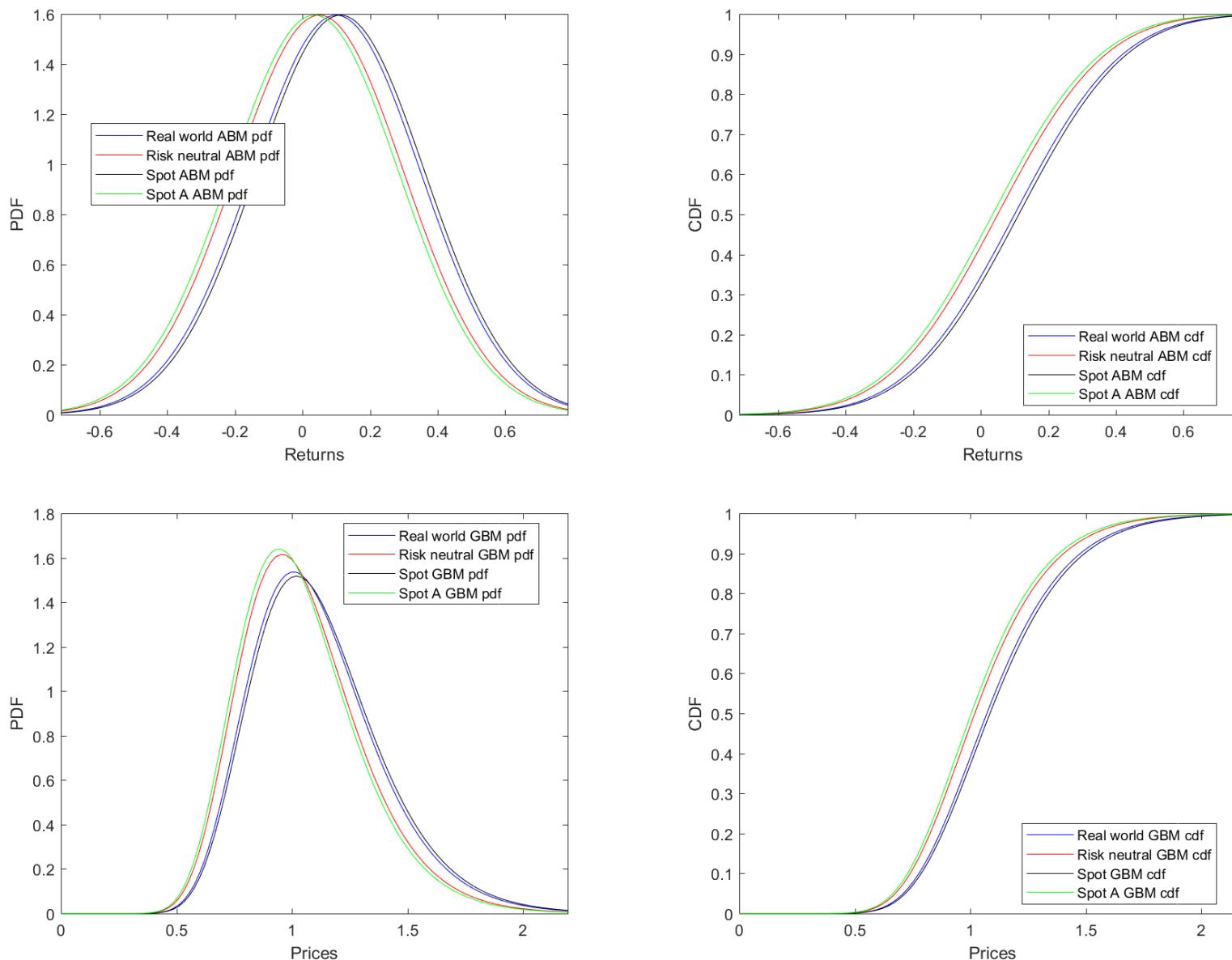


Figure 7.1: PDF (left) and CDF (right) of log-returns (top) and prices (bottom) under different measures. Parameters: $\mu = 0.1, \sigma = 0.25; r = 0.05; t = 1; \sigma_A = 0.15; \rho = -0.4$.

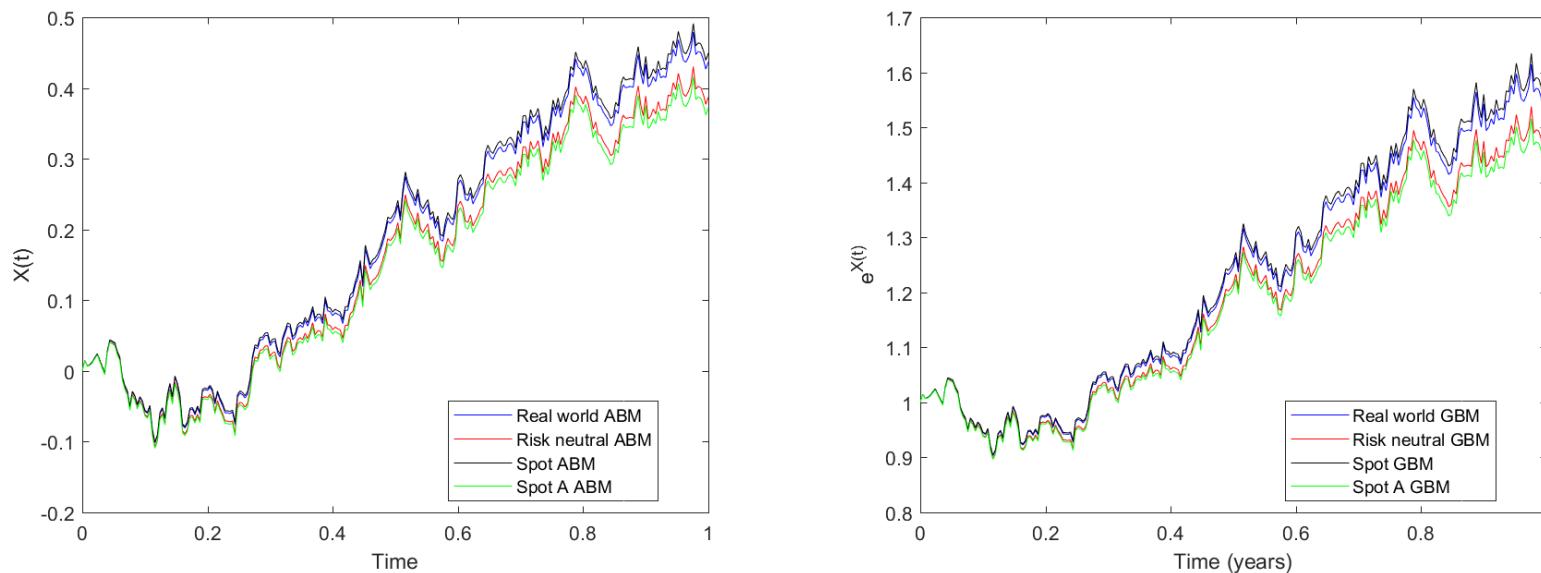


Figure 7.2: Simulated path of the log-level (left) and level (right) under different measures. Parameters: $\mu = 0.1$, $\sigma = 0.25$; $r = 0.05$; $t = 1$; $\sigma_A = 0.15$; $\rho = -0.4$.

CHAPTER 8. REFERENCES

- Andersen, L. (2008). Simple and efficient simulation of the Heston stochastic volatility model. *Journal of Computational Finance* 11(3) 1-42.
- Barndorff-Nielsen, O. E. (1995). Normal inverse Gaussian distributions and the modeling of stock returns. Research report N. 300, Department of Theoretical Statistics, Aarhus University.
- Brigo D. and Mercurio, F. (2006). *Interest Rate Models: Theory and Practice*, 2nd edition, Springer Finance, Heidelberg.
- Cairns, A. (2004). *Interest Rate Models. An Introduction*. Princeton University Press.
- Carr, P., Geman, H., Madan, D. B. and Yor, M. (2002). The fine structure of asset returns: an empirical investigation, *Journal of Business*, 75, 305-32.
- Carr, P., Geman, H., Madan, D.B., Yor, M., (2003). Stochastic volatility for Lévy processes. *Mathematical Finance*, 13, 345–382.
- Carr, P., Madan, D. B. (1999). Option Pricing and the Fast Fourier Transform, July 1999, *Journal of Computational Finance*, 61-73.
- Carr, P. and Wu, L. (2004). Time-changed Lévy processes and option pricing. *Journal of Financial Economics*, 71, 113-141.
- Carr, P. and Wu, L. (2007). Stochastic skew in currency options. *Journal of Financial Economics*, 86, 213–247.

- Carr, P. and Wu, L., (2016). Leverage effect, volatility feedback, and self-exciting market disruptions. *Journal of Financial and Quantitative Analysis*. Forthcoming.
- Cox, J. (1975). Notes on option pricing I: Constant elasticity of variance diffusions. Working paper, Stanford University (reprinted in *Journal of Portfolio Management*, 1996, 23, 15-17).
- Cox, J. and Ross, S. (1976). The valuation of options for alternative stochastic processes. *J. Financial Economics*, 3, 145-166.
- Das, S. J. and Sundaram, R. K. (1999). Of Smiles and Smirks: A Term Structure Perspective. *The Journal of Financial and Quantitative Analysis*, 34(2), 211-239.
- Delbaen, F. and Shirakawa, H. (2002). A Note of Option Pricing for Constant Elasticity of Variance Model. *Asia-Pacific Financial Markets*, 9, 85-99.
- Duffie D., J. Pan, K. Singleton. Transform Analysis and Asset Pricing for Affine Jump-diffusions. *Econometrica*. Volume 68, Issue 6, Pages 1343-1376.
- Emanuel, D. and MacBeth, J. (1982). Further results on the constant elasticity of variance call option pricing model. *J. Financial and Quant. Anal.* 17 533-554.
- Fusai G, Roncoroni, A., Marena, M. (2008). Analytical pricing of discretely monitored Asian-style options: Theory and application to commodity markets. *Journal of Banking and Finance*, 32(10), p. 2033-2045.
- Fusai, G. and Roncoroni, A. (2008). *Implementing Models in Quantitative Finance: Methods and Cases*. Springer Finance.
- Geman, H., El Karoui, N., Rochet, J.C. (1995). Changes of Numeraire, Changes of Probability Measures and Pricing of Options. *Journal of Applied Probability*. 32: 443-458.
- Going-Jaeschke A. and M. Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9:313-349, 2003.
- Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies* 6 (2): 327-343.

- Huang, J.Z., Wu, L., (2004). Specification analysis of option pricing models based on time-changed Lévy processes. *Journal of Finance*, 59, 1405-1439.
- Jeanblanc, M., M. Yor and M. Chesney. (2009). *Mathematical Methods for Financial Markets*, Springer Finance, London.
- Kou, S. G. (2002). A jump-diffusion model for option pricing, *Management Science*, 48, 8, 1086-1101.
- Lord, R., R. Koekkoek, D. Van Dijk. (2010). A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance* 10(2) 177-194.
- Maghsoudi , Y. (1996). Solution of the extended CIR term structure and bond option valuation, *Mathematical Finance*, 6(1), 89-109.
- Madan, D. B., Carr, P., and Chang, E. (1998). The variance gamma process and option pricing, *European Finance Review*, 2, 79-105.
- Madan, D. B. and Milne, F. (1991). Option pricing with VG martingale components, *Mathematical Finance*, 1, 39-45.
- Madan, D. B. and Seneta, E. (1990). The variance gamma (VG) model for share market returns. *Journal of Business*, 63, 511-24.
- Merton, R. (1976). Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics*, 125-144.
- Nelson, D., Ramaswamy, K. (1990). Simple binomial processes as diffusion approximations in financial models. *Rev. Finan. Stud.* 3, 393-430.
- Jackwerth, J., Rubinstein, M. (1996). Recovering Probability Distributions from Option Prices. *J. of Finance* 51(5), 1611-1631.
- Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula. *J. Finance*, 44(1), 211-219.
- Shreve, S. (2004). *Stochastic Calculus for Finance II*, Springer Finance.

CHAPTER 9. MATLAB LIVESCRIPTS

Several examples and figures produced in this document can be reproduced and used for lectures by running the Matlab Live Scripts. In particular, we have produced the following ones

1. SimulatingBrownianMotion.mlx see Section 2.
2. TotalvsQuadraticVariation.mlx see Section 2.1.10.
3. SimulatingStocIntegral.mlx see Section 3.
4. SimulatingArithmeticBrownianMotion.mlx see Section 3.8.
5. SimulatingGeometricBrownianMotion.mlx see Section 5.1.
6. SimulatingGBMFittingFuturesCurve.mlx see Section 5.1.4.
7. UnderstandingMeanReversion.mlx see Section 5.2.
8. SimulatingVasicekModel.mlx see Section 5.2.
9. SimulatingCIR.mlx see Section 5.3.
10. CIRModelProbDensFunction.mlx see Section 5.3.
11. CEVModel.mlx see Section 5.4.

12. SimulatingBrownianBridge.mlx see Section 5.6.
13. MomentsHeston.mlx see see Section 5.7.1.
14. DensityHeston.mlx see Section 5.7.3.
15. OptionPricingHeston.mlx see Section 5.7.4.
16. SimulatingJumpProcess.mlx see Section 6.2.
17. UnderstandingVarianceGamma.mlx see Section 6.3.1.
18. VGOptionpricing.mlx see Section 6.3.1.
19. NumericalQuadrature.mlx see Section 11.
20. AdaptiveQuadrature.mlx see Section 11.
21. InvertingthecharacteristicfunctionviaCOSmethod.mlx see Section 12.
22. TestingCOsoptionpricingonN.mlx see Section 12.2.2.

All of them are freely downloadable from the following website Matlab Live Scripts.

PART II

Appendix

CHAPTER 10. A QUICK REVIEW OF DISTRIBUTIONS RELEVANT IN FINANCE

Contents

10.1 The Normal Distribution	221
10.2 The Lognormal distribution	221
10.3 The Chi-Square distribution	227
10.4 The noncentral chi-squared distribution	229
10.5 The Poisson distribution	233
10.6 The Exponential distribution	234
10.7 The Gamma distribution	238
10.8 The multivariate Gaussian distribution	241
10.8.1 The bivariate Normal distribution	242
10.9 Simulating Random Variables	250
10.9.1 Example: Sampling from the Normal Distribution	252
10.9.2 Example: Sampling from the Lognormal Distribution	255
10.9.3 Example: Sampling from the Poisson Distribution	256
10.9.4 Example: Sampling from the Gamma Distribution	256
10.9.5 Example: Sampling from the Chi-Square Distribution	256
10.9.6 Example: Sampling from the Non-Central Chi-Square Distribution	256
10.9.7 Example: Sampling from the Multivariate Normal Distribution	258

In this Appendix, we quickly review the properties of distributions relevant in finance, like

- Normal distribution,
- Lognormal distribution,
- Chi-Square distribution,
- Non-Central Chi-Square distribution,
- Poisson distribution,
- Exponential distribution,
- Gamma distribution,
- Multivariate Gaussian distribution.

Then we present three important numerical procedures

- Simulation of random variables, via the inverse method.
- numerical integration;
- numerical inversion of the characteristic function.

10.1. THE NORMAL DISTRIBUTION

Fact 74 (Normal Distribution) A normal (Gaussian) random variable on \mathbb{R} with expected value $\mu \in \mathbb{R}$ and standard deviation $\sigma \in \mathbb{R}^+$ has density function (PDF) $\phi_{\mu,\sigma}(x)$ and cumulative distribution (CDF) $\Phi_{\mu,\sigma}(x)$ given by:

$$\phi_{\mu,\sigma}(x) = f_X(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), x \in \mathbb{R},$$

and

$$\Phi_{\mu,\sigma}(x) = F_X(x; \mu, \sigma) = \int_{-\infty}^x \phi_{\mu,\sigma}(s) ds.$$

We write $X \sim \mathcal{N}(\mu, \sigma^2)$. If $\mu = 0$ and $\sigma = 1$ we have the so called unit standard Gaussian random variable. In particular, if $Z \sim \mathcal{N}(0, 1)$, then

$$X = \mu + \sigma Z \sim \mathcal{N}(\mu, \sigma^2).$$

Given the CDF, we can compute the probability that X falls in a given interval

$$\Pr(a < X < b) = \Phi_{0,1}\left(\frac{b-\mu}{\sigma}\right) - \Phi_{0,1}\left(\frac{a-\mu}{\sigma}\right).$$

We can produce Table (10.1). Table (10.2) illustrates the main properties of the Gaussian distribution. The resulting PDF and CDF are shown in Figure (10.1).

10.2. THE LOGNORMAL DISTRIBUTION

Fact 75 (Lognormal Distribution) Let X be a Normal r.v. with mean μ and standard deviation σ . Then the r.v.

$$Y = \exp(X),$$

is said to have a lognormal distribution with parameters μ and σ and density $f_Y(y; \mu, \sigma)$ given by

$$f_Y(y; \mu, \sigma) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right), y \in \mathbb{R}^+.$$

Probability Intervals				
k	$\mu - k\sigma$	$\mu + k\sigma$	$\Pr(\mu - k\sigma < X < \mu + k\sigma)$	
1	-0.3542	0.2742	0.6827	
2	-0.6684	0.5884	0.9545	
3	-0.9826	0.9026	0.9973	
4	-1.2968	1.2168	0.9999	
5	-1.611	1.531	1.0000	

Table 10.1: Probability intervals of the Gaussian distribution.

Properties of the Normal distribution	
Support	$x \in (-\infty; +\infty)$
Mean	$\mathbb{E}(X) = \mu$
Variance	$\text{Var}(X) = \sigma^2$
Skewness	$\frac{\mathbb{E}((X-\mu)^3)}{\sigma^3} = 0$
Kurtosis	$\frac{\mathbb{E}((X-\mu)^4)}{\sigma^4} = 3$
Central odd moments	$\mathbb{E}((X-\mu)^{2n+1}) = 0$
Central even moments	$\mathbb{E}((X-\mu)^{2n}) = \frac{(2n)!}{n!} \frac{\sigma^{2n}}{2^n}$

Table 10.2: Moments of the Gaussian distribution.

Moreover:

$$\mathbb{E}(Y) = e^{\mu + \frac{1}{2}\sigma^2}$$

and

$$\text{Var}(Y) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1).$$

We write $Y \sim \mathcal{LN}(\mu, \sigma^2)$.

The relationship between densities of the Gaussian and Lognormal distribution is depicted in Figure (10.2).

Matlab Code

```
%%%%%%PLOT OF THE GAUSSIAN DISTRIBUTION%%%%%
%%%%%%assign parameters
mu=0.2;
sg=0.1;
%assign x-range
x=linspace(-.4,0.8,100);
%compute pdf and cdf
pdfg=pdf('norm',x,mu,sg);
cdfg=cdf('norm',x,mu,sg);
%make the plot
plot(x,pdfg,'r',x,cdfg,'b');
%define the limits in the x-axis
xlim([-0.4 0.8])
%insert the legend
legend('Gaussian pdf','Gaussian cdf')
%print the figure
print(h,'-dpng','FigGaussianpdf')
```

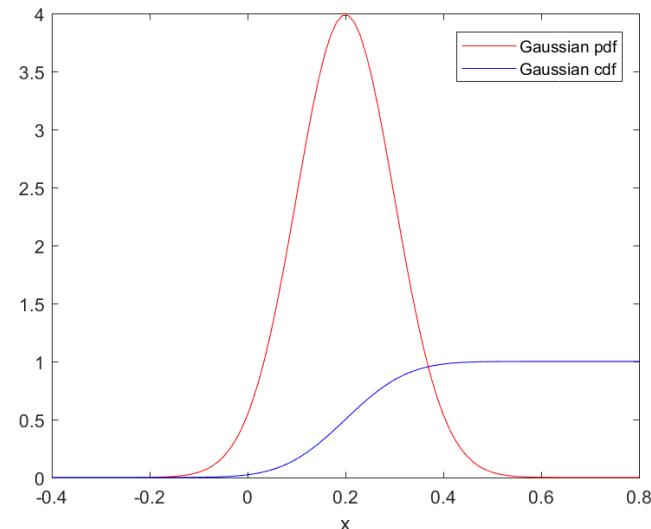


Figure 10.1: Density and Cumulative Distribution of the Gaussian random variable.

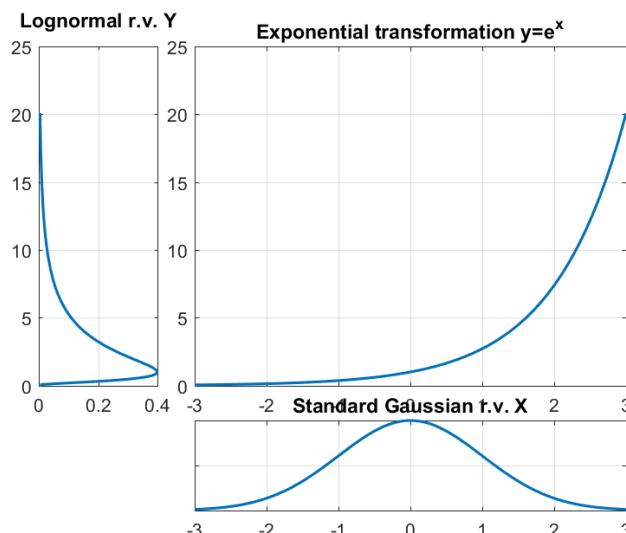


Figure 10.2: Densities of the Gaussian and of the Lognormal random variables.

The cumulative density function of a Lognormal r.v. is given by:

$$\begin{aligned}
 F_Y(y) &= \Pr(Y \leq y) \\
 &= \Pr(e^X \leq y), \text{ where } X \sim \mathcal{N}(\mu, \sigma^2) \\
 &= \Pr(X \leq \ln(y)) \\
 &= \Phi_{\mu, \sigma^2}(\ln(y)) \\
 &= \Phi_{0,1}\left(\frac{\ln(y) - \mu}{\sigma}\right).
 \end{aligned}$$

Given the CDF, we can compute the probability that X falls in a given interval

$$\begin{aligned}
 \Pr(a < Y < b) &= \Phi_{\mu, \sigma^2}(\ln(b)) - \Phi_{\mu, \sigma^2}(\ln(a)) \\
 &= \Phi_{0,1}\left(\frac{\ln(b) - \mu}{\sigma}\right) - \Phi_{0,1}\left(\frac{\ln(a) - \mu}{\sigma}\right).
 \end{aligned}$$

The main properties of the Lognormal distribution are summarized in Table (10.3). The resulting PDF and CDF are shown in Figure (10.3).

LOGNORMAL DISTRIBUTION $Y \sim \mathcal{LN}(\mu, \sigma^2)$ FACTS	
Support	$y \in [0; +\infty)$
PDF	$f_Y(y; \mu, \sigma) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right), y \in \Re^+$
CDF	$\Phi_{0,1}\left(\frac{\ln(y) - \mu}{\sigma}\right)$
Mean	$\mathbb{E}(Y) = e^{\mu + \frac{1}{2}\sigma^2}$
Variance	$\mathbb{V}ar(Y) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$
Skewness	$\frac{\mathbb{E}((Y - \mathbb{E}(Y))^3)}{(\mathbb{V}ar(Y))^{\frac{3}{2}}} = (e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}$
Kurtosis	$\frac{\mathbb{E}((Y - \mathbb{E}(Y))^4)}{(\mathbb{V}ar(Y))^2} = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 3$
Moments	$\mathbb{E}(Y^n) = e^{n\mu + \frac{1}{2}n^2\sigma^2}$

Table 10.3: Properties of the Lognormal distribution.

Matlab Code

```
%%%%%PLOT OF THE LOGNORMAL DISTRIBUTION%%%%%
%assign parameters
mu=0.2; sg=0.1;
%assign x-range
x=linspace(0.0,2,100);
%compute pdf and cdf
pdflg=pdf('lognorm',x,mu,sg);
cdflg=cdf('lognorm',x,mu,sg);
%make the plot
h=figure('Color',[1 1 1])
plot(x,pdflg,'r',x,cdflg,'b');
%define the limits in the x-axis
xlim([0 2]); xlabel('x');
%insert title
title('Lognormal pdf and cdf when \mu=0.2 and \sigma =0.1')
%insert the legend
legend('Log-normal pdf','Log-normal cdf')
%print the figure
print(h,'-dpng','FigLogGaussianpdf.png')
```

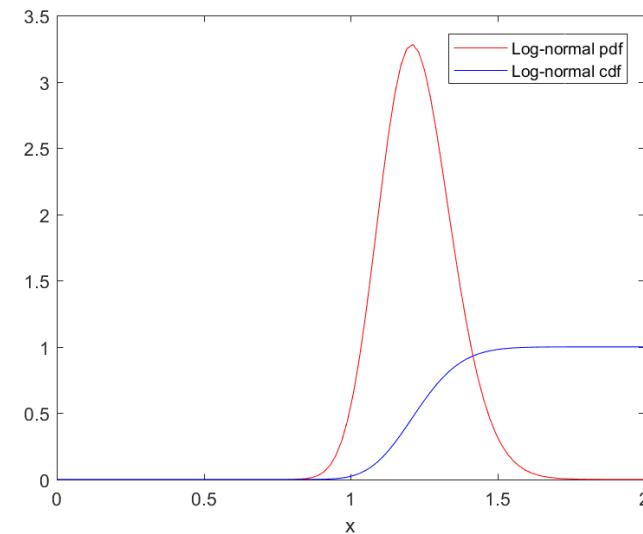


Figure 10.3: Density and Cumulative Distribution of the Lognormal random variable.

10.3. THE CHI-SQUARE DISTRIBUTION

Fact 76 (Chi-Square Distribution) Let $X_i, i=1,\dots,n$, be n independent Normal random variables with zero mean and unit standard deviation. Then the r.v.

$$Y = \sum_{i=1}^n X_i^2,$$

is said to have a *chi-square distribution* with n degrees of freedom and density $f_Y(y; n)$ given by

$$f_Y(y; n) = \frac{e^{-y/2} y^{n/2-1}}{2^{n/2} \Gamma(n/2)}, y \in \Re^+.$$

where $\Gamma(n)$ is the Gamma function

$$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt,$$

and if n is an integer $\Gamma(n) = (n-1)!$. Moreover:

$$\mathbb{E}(X) = n$$

and

$$\mathbb{V}ar(X) = 2n.$$

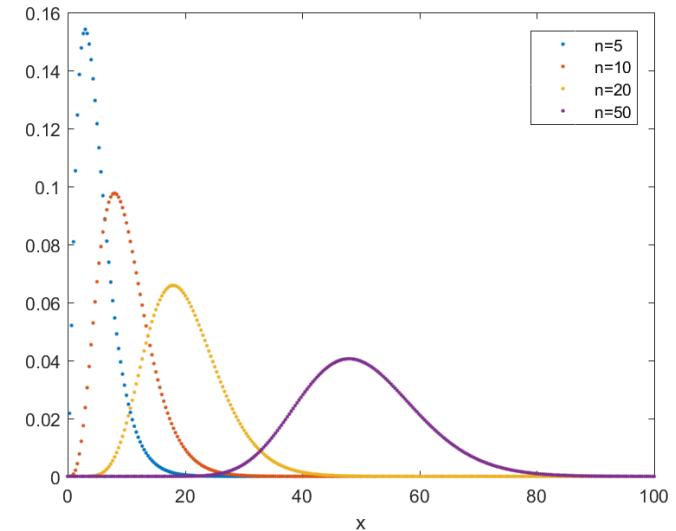
We write $Y \sim \chi_n^2$.

The main properties of the Chi-square distribution are summarized in Table (10.4). The resulting PDF and CDF are shown in Figure (10.4). Figure (10.5) illustrates the different shapes of the chi-square distribution for increasing number of degrees of freedom.

Properties of the chi-square distribution	
Support	$y \in [0; +\infty)$
Mean	$\mathbb{E}(Y) = n$
Variance	$\text{Var}(Y) = 2n$
Skewness	$\frac{\mathbb{E}((Y-n)^3)}{(2n)^{\frac{3}{2}}} = \sqrt{\frac{8}{n}}$
Excess Kurtosis	$\frac{\mathbb{E}((Y-n)^4)}{(2n)^2} - 3 = \frac{12}{n}$
Moments	$\mathbb{E}(Y^m) = 2^m \frac{\Gamma(m+\frac{n}{2})}{\Gamma(\frac{n}{2})}$

Table 10.4: Moments of the Chi-square distribution.**Matlab Code**

```
%%%%%THE Chi-2 DISTRIBUTION AND THE DEGREES OF FREEDOM%%%%%
%%clear all;
%assign parameters
n=[5 10 20 50]; %degrees of freedom
%assign x-range
x=linspace(0,100,300);
%compute pdf and cdf
for i=1:length(n)
    pdfchi2(i,:)=pdf('chi2',x,n,n(i));
end
h=figure('Color',[1 1 1])
plot(x,pdfchi2,'.')
xlabel('x')
legend('n=5','n=10','n=20','n=50')
title('The Chi-Square pdf varying ... the number n of degrees of freedom')
```

**Figure 10.5:** Density of the Chi-Square random variable varying n .

Matlab Code

```
%%%%%%PLOT OF THE Chi-2 DISTRIBUTION%%%%%
%%%%%assign parameters
n=5; %degrees of freedom
%%%%%assign x-range
x=linspace(0,20,100);
%%%%%compute pdf and cdf
pdfchi2=pdf('chi2',x,n);
cdfchi2=cdf('chi2',x,n);
%%%%%make the plot
h=figure('Color',[1 1 1])
plot(x,pdfchi2,'r',x,cdfchi2,'b')
xlabel('x')
title('Chi-Square PDF and CDF')
legend('Chi-Square pdf','Chi-Square cdf')
print(h,'-djpeg','FigChiSquarepdf.jpg')
```

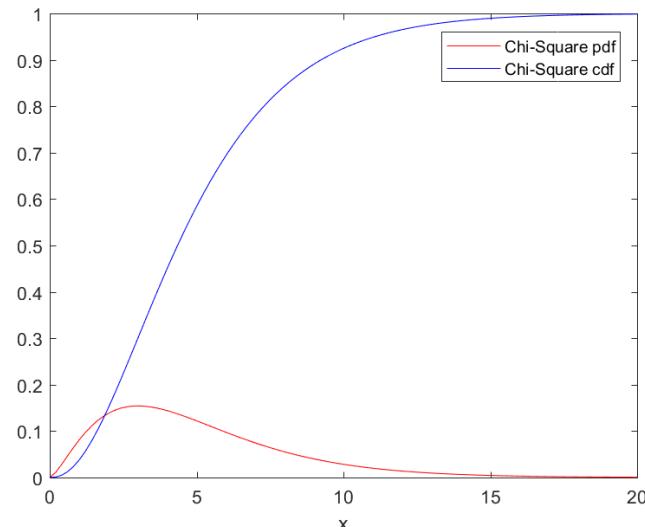


Figure 10.4: Density and Cumulative Distribution of the Chi-Square random variable with $n = 20$ degrees of freedom.

10.4. THE NONCENTRAL CHI-SQUARED DISTRIBUTION

Fact 77 (Noncentral chi-squared distribution) Let $X_i, i = 1, \dots, n$, be n independent Gaussian random variables with zero mean and unit standard deviation. Then the r.v.

$$Y = \sum_{i=1}^n (X_i + \delta_i)^2,$$

is said to have a noncentral chi-squared distribution with n degrees of freedom and non-centrality parameter d

$$d = \sum_i \delta_i^2,$$

and density $f_Y(y; n, d)$ given by

$$f_Y(y; n, d) = e^{-(y+d)/2} \frac{x^{n/2-1}}{2^{n/2}} \sum_{k=0}^{\infty} \frac{d^k}{2^{2k} k! \Gamma(k + 1/2n)}, y \in \mathbb{R}^+.$$

An equivalent representation of the density function is in terms of the Bessel function

$$f_Y(y; n, d) = \frac{e^{-(y+d)/2} y^{(n-1)/2} \sqrt{d}}{2(dy)^{n/4}} I_{n/2-1}(\sqrt{dy}),$$

where $I_n(x)$ is a modified Bessel function of the first kind.

Moreover:

$$\mathbb{E}(Y) = n + d,$$

and

$$\mathbb{V}ar(Y) = 2(n + 2d).$$

We write $Y \sim \chi_{n,d}^2$.

Remark 78 We can make the following observations:

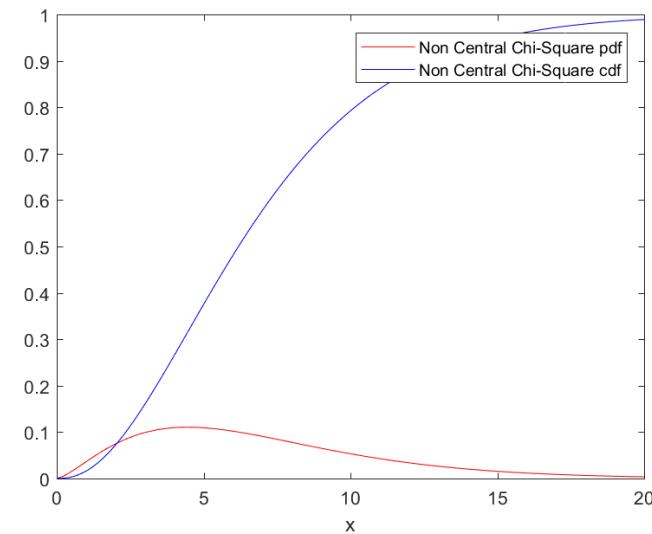
- The noncentrality parameter depends on the sum $d = \sum_i^n \delta_i^2$, not on the individual values of the parameters δ_i ;
- If $d = 0$, the noncentral chi-squared distribution collapses to the chi-squared distribution.
- Here we are assuming that n is an integer parameter: this is not necessary.

Table (10.5) illustrates the main properties of the noncentral chi-squared distribution. The resulting PDF and CDF are shown in Figure (10.6). Figure (10.7) illustrates the different shapes of the density for different values of the noncentrality parameter d .

Properties of the noncentral chi-squared distribution distribution	
Support	$y \in [0; +\infty)$
Mean	$\mathbb{E}(Y) = n + d$
Variance	$\text{Var}(Y) = 2 \times (n + 2d)$
Skewness	$\frac{\mathbb{E}((Y - \mathbb{E}(Y))^3)}{(\text{Var}(Y))^{\frac{3}{2}}} = \frac{2^{3/2}(n+3d)}{(n+2d)^{3/2}}$
Excess Kurtosis	$\frac{\mathbb{E}((Y - \mathbb{E}(Y))^4)}{(\text{Var}(Y))^2} - 3 = \frac{12(n+4d)}{(n+2d)^2}$

Table 10.5: Moments of the noncentral chi-squared distribution.**Matlab Code**

```
%%%%%%PLOT OF THE NONCENTRAL CHI-SQUARED DISTRIBUTION%%%%%
clear all;
%assign parameters
n=5; %degrees of freedom
noncentralpar=2;%non centrality parameter
%assign x-range
x=linspace(0,20,100);
%compute pdf and cdf
pdfncchi2=pdf('ncx2',x,n,noncentralpar );
cdfncchi2=cdf('ncx2',x,n,noncentralpar);
%make the plot
h=figure('Color',[1 1 1])
plot(x,pdfncchi2,'r',x,cdfncchi2,'b')
xlabel('x')
title('Non central Chi-Square pdf and cdf when n=5 and d =2')
legend('Non Central Chi-Square pdf',...
'Non Central Chi-Square cdf')
print(h,'-djpeg','FigNCChiSquarepdf.jpg')
```

**Figure 10.6:** Density and CDF of the noncentral chi-squared distribution with $n = 5$ and $d = 2$.

Matlab Code

```
%%%%%THE NCChi-2 DISTRIBUTION AND THE %%%
%%% PARAMETER OF NON CENTRALITY %%%
%%% clear all;
%assign parameters
n=15; %degrees of freedom
ncp=[0.5 1.5 3 5 10];%non centrality parameter
%assign x-range
x=linspace(0,20,300);
%compute pdf and cdf
for i=1:length(ncp)
    pdfncchi2(i,:)=pdf('ncx2',x,ncp(i));
end
h=figure('Color',[ 1 1 1])
plot(x,pdfncchi2,'.')
xlabel('x')
legend('d=0.5','d=1.5','d=3','d=5','d=10')
title('The Non-Central Chi-Square pdf varying the non ...')
centrality parameter and with n=15 of degrees of freedom')
print(h,'-djpeg','FigNCChiParSquarepdf.jpg')
```

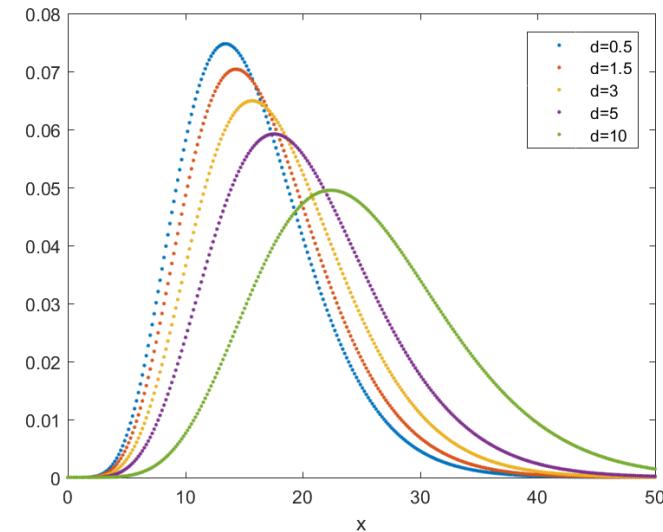


Figure 10.7: Density of the non-central chi-square distribution varying the parameter of non centrality d .

Properties of the Poisson distribution distribution	
Support	$n \in 0, 1, 2, \dots$
Mean	$\mathbb{E}(N) = \lambda$
Variance	$\mathbb{V}ar(N) = \lambda$
Skewness	$\frac{\mathbb{E}((N-\lambda)^3)}{\lambda^{\frac{3}{2}}} = \frac{1}{\sqrt{\lambda}}$
Excess Kurtosis	$\frac{\mathbb{E}((N-\lambda)^4)}{\lambda^2} - 3 = \frac{1}{\lambda}$

Table 10.6: Moments of the Poisson distribution.

10.5. THE POISSON DISTRIBUTION

Fact 79 (Poisson Distribution) *A Poisson random variable N with rate λ has probability mass*

$$p_N(n) = \frac{e^{-\lambda} \lambda^n}{n!}.$$

Moreover, the mean and the variance are respectively:

$$\mathbb{E}(N) = \lambda$$

and

$$\mathbb{V}ar(N) = \lambda.$$

We write $N \sim Poi(\lambda)$.

Table (10.6) illustrates the main properties of the Poisson distribution. The resulting probability mass function and CDF are shown in Figure (10.8). The impact of the rate of arrival λ is represented in Figure (10.9).

Matlab Code

```
%%%%%%PLOT OF THE Poisson DISTRIBUTION%%%%%
%%%%%%%
clear all;
%assign parameters
lambda=5; %rate of arrival
%assign x-range
x=(0:1:20);
%compute pdf and cdf
pdfpoil=poisspdf(x,lambda);
cdfpoil=poisscdf(x,lambda);
%make the plot
h=figure('Color',[1 1 1])
plot(x,pdfpoil,'-o r',x,cdfpoil,'-o b')
xlabel('x')
title('Poisson pdf and cdf when \lambda = 5')
legend('Poisson pdf',...
'Poisson cdf')
print(h,'-djpeg','FigPoisson1pdf.jpg')
```

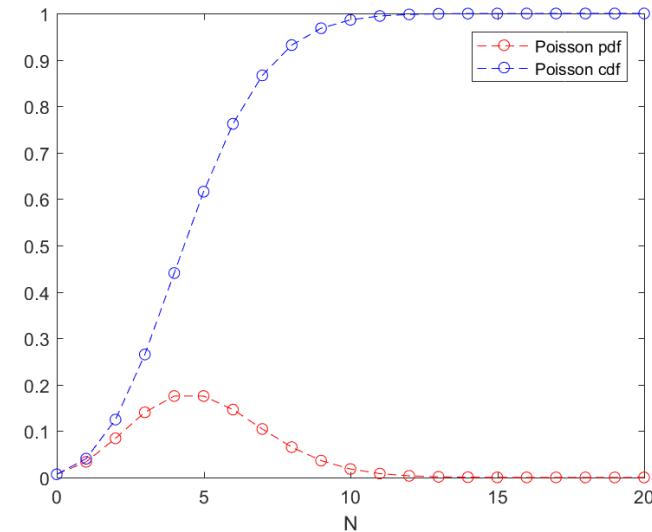


Figure 10.8: Density and CDF of the Poisson distribution with $\lambda = 5$. The functions are only defined at integer values of x . The connecting lines are only guides for the eye.

10.6. THE EXPONENTIAL DISTRIBUTION

Fact 80 (Exponential Distribution) A non negative random variable X is said to have an *exponential distribution* with parameter λ when its density is given by

$$f_X(x; \lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0 & x < 0, \end{cases}$$

and its CDF is given by

$$F_X(x; \lambda) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0 & x < 0, \end{cases}$$

Matlab Code

```
%%%%%%The Poisson DISTRIBUTION and the rate of arrival%%%%%
%%%%%%The Poisson DISTRIBUTION and the rate of arrival%%%%%
%%%%%%The Poisson DISTRIBUTION and the rate of arrival%%%%%
clear all;
%assign parameters
lambda=[0.5 1 5 10]; %rate of arrival
%assign x-range
x=(0:1:20);
%compute pdf and cdf
for i=1:length(lambda)
    pdfpoil(i,:)=poisspdf(x,lambda(i));
    cdfpoil(i,:)=poisscdf(x,lambda(i));
end
%make the plot
h=figure('Color',[1 1 1])
subplot(2,1,1)
plot(x,pdfpoil,'-o')
xlabel('x')
title('Poisson pdf varying the rate of arrival \lambda')
legend('\lambda = 0.5', '\lambda = 1', ...
'\lambda = 5', '\lambda = 10')
subplot(2,1,2)
plot(x,cdfpoil,'o')
xlabel('x')
title('Poisson cdf varying the rate of arrival \lambda')
print(h,'-djpeg','FigPoisson2pdf.jpg')
```

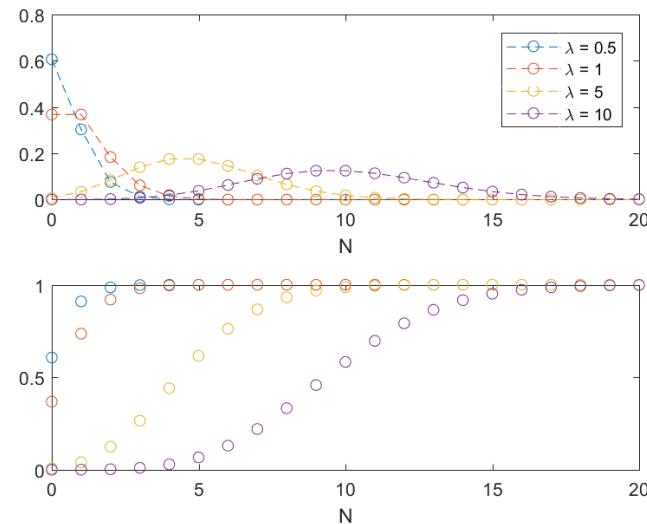


Figure 10.9: PDF (top) and CDF (bottom) of the Poisson distribution varying the rate of arrival λ .

Moreover:

$$\mathbb{E}(X) = \lambda^{-1}$$

Properties of the Exponential distribution distribution	
Support	$x \in [0; +\infty)$
Mean	$\mathbb{E}(X) = \frac{1}{\lambda}$
Variance	$\text{Var}(X) = \frac{1}{\lambda^2}$
Skewness	$\frac{\mathbb{E}((X-\mathbb{E}(X))^3)}{(\text{Var}(X))^{\frac{3}{2}}} = 2$
Kurtosis	$\frac{\mathbb{E}((X-\mathbb{E}(X))^4)}{(\text{Var}(X))^2} = 9$

Table 10.7: Moments of the Exponential distribution.

and

$$\text{Var}(X) = \lambda^{-2}.$$

We write $X \sim \text{Exp}(\lambda)$.

Table (10.7) illustrates the main properties of the Exponential distribution, whilst PDF and CDF are shown in Figure (10.10).

Matlab Code

```
%%%%%PLOT OF THE Exponential DISTRIBUTION%%%%%
clear all;
%assign parameters
lambda=1.5; %rate of arrival
%assign x-range
x=linspace(0,5,100);
%compute pdf and cdf
pdfexpl=exppdf(x,lambda^(-1));
cdfexpl=expcdf(x,lambda^(-1));
%make the plot
h=figure('Color',[1 1 1])
plot(x,pdfexpl,'r',x,cdfexpl,'b')
xlabel('x')
title('Exponential pdf and cdf when \lambda = 1.5')
legend('Exponential pdf',...
'Exponential cdf')
print(h,'-djpeg','Figexpopdf.jpg')
```

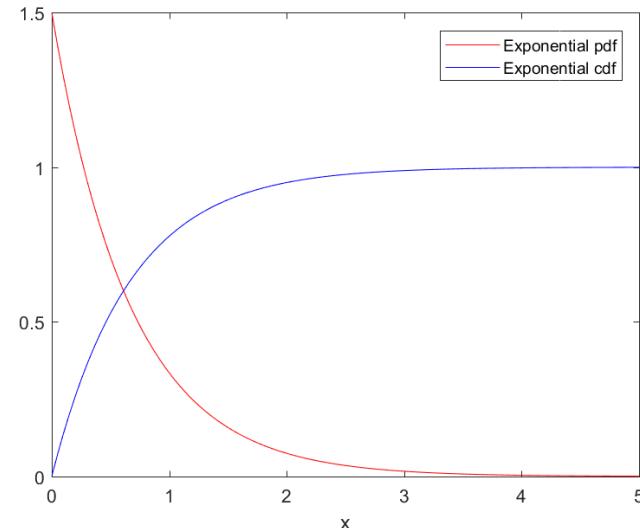


Figure 10.10: Density and CDF of the Exponential distribution with $\lambda = 1.5$.

10.7. THE GAMMA DISTRIBUTION

Fact 81 (Gamma Distribution) A non negative random variable X is said to have a *Gamma distribution* with shape parameter α and rate parameter λ when its density is given by

$$f_X(x; \alpha, \lambda) = \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x},$$

where $\Gamma(\alpha)$ is the Gamma function, which is defined as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx,$$

Moreover:

$$\mathbb{E}(X) = \frac{\alpha}{\lambda}$$

and

$$\mathbb{V}ar(X) = \frac{\alpha}{\lambda^2}$$

We write $X \sim \Gamma(\alpha, \lambda)$.

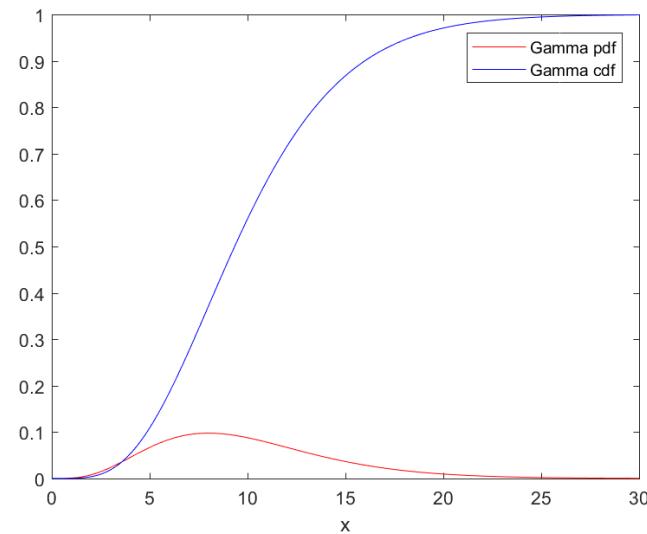
- Remark 82**
- A Gamma random variable with parameter $\lambda = \frac{1}{2}$, i.e. $X \sim \Gamma(\alpha, \frac{1}{2})$, is distributed according to a chi-square distribution with 2α degrees of freedom;
 - A Gamma random variable with $\alpha = 1$, i.e. $X \sim \Gamma(1, \lambda)$, is distributed according to an exponential distribution of parameter λ .

Table (10.8) illustrates the main properties of the Gamma distribution. Figures (10.25) and (10.12) show the Gamma PDF and CDF for different values of the parameters α and λ .

Properties of the Gamma distribution distribution	
Support	$x \in [0; +\infty)$
Mean	$\mathbb{E}(X) = \frac{\alpha}{\lambda}$
Variance	$\text{Var}(X) = \frac{\alpha}{\lambda^2}$
Skewness	$\frac{\mathbb{E}((X-\mathbb{E}(X))^3)}{(\text{Var}(X))^{\frac{3}{2}}} = \frac{2}{\sqrt{\alpha}}$
Kurtosis	$\frac{\mathbb{E}((X-\mathbb{E}(X))^4)}{(\text{Var}(X))^2} = 3 + \frac{6}{\alpha}$

Table 10.8: Moments of the Gamma distribution.**Matlab Code**

```
%%%%%
%%%PLOT OF THE Gamma DISTRIBUTION%%%%%
%%%%%
clear all;
%assign parameters
alpha=5; %shape parameter
lambda=0.5; %rate parameter
%assign x-range
x=linspace(0, 30, 200);
%compute pdf and cdf
pdfgamma=gampdf(x, alpha, lambda^(-1));
cdfgamma=gamcdf(x, alpha, lambda^(-1));
%make the plot
h=figure('Color', [ 1 1 1])
plot(x,pdfgamma, '- r', x,cdfgamma, '- b')
xlabel('x')
title('Gamma pdf and cdf when\alpha =5 \lambda = 0.5')
legend('Gamma pdf',...
'Gamma cdf')
print(h, '-djpeg', 'FigGamma1pdf.jpg')
```



Matlab Code

```

%%%%%
%%The Gamma DISTRIBUTION and the shape/rate parameters%%
%%%%%
alpha=[1 2 5 10];
lambda=[0.5 0.5 0.5 2]; %rate parameter
%assign x-range
x=linspace(0,30,200);
%compute pdf and cdf
for i=1:length(alpha)
    pdfgamma(i,:)=gampdf(x,alpha(i),lambda(i)^(-1));
    cdfgamma(i,:)=gamcdf(x,alpha(i),lambda(i)^(-1));
end
%make the plot
h=figure('Color',[1 1 1])
subplot(2,1,1)
plot(x,pdfgamma,'-')
xlabel('x')
title('Gamma pdf varying \alpha and \lambda')
legend('\alpha = 1, \lambda = 0.5', '\alpha = 2, \lambda = 0.5',
    '\alpha = 5, \lambda = 0.5', '\alpha = 10, \lambda = 2')
subplot(2,1,2)
plot(x,cdfgamma,'-')
xlabel('x')
title('Gamma cdf varying \alpha and \lambda')
legend('\alpha = 1, \lambda = 0.5', '\alpha = 2, \lambda = 0.5',
    '\alpha = 5, \lambda = 0.5', '\alpha = 10, \lambda = 2')
print(h,'-djpeg','FigGamma2pdf.jpg')

```

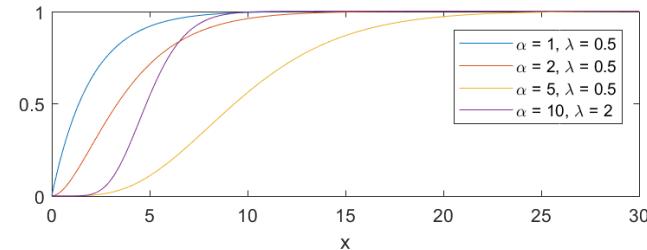
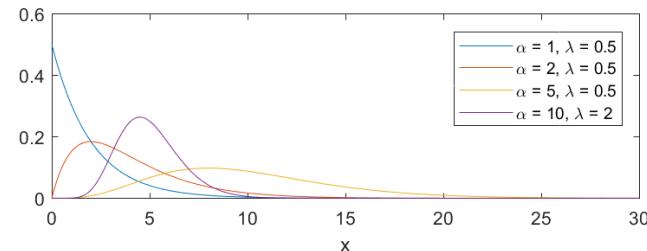


Figure 10.12: PDF (top) and CDF (bottom) of the Gamma distribution varying the shape parameter α and the rate parameter λ .

10.8. THE MULTIVARIATE GAUSSIAN DISTRIBUTION

Fact 83 (Multivariate Gaussian Distribution) Let μ a $N \times 1$ vector and Σ a $N \times N$ matrix, symmetric and positive definite. A random vector $\mathbf{X} = (X_1, \dots, X_N)$ is distributed according to a multivariate normal distribution with parameters μ and Σ , and we write

$$\mathbf{X} \sim \mathcal{N}(\mu, \Sigma),$$

if its density function $f_{\mathbf{X}}(\mathbf{x})$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2} (\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)\right),$$

where $\det(\Sigma)$ is the determinant of the matrix Σ and Σ^{-1} is the inverse of the matrix Σ .

If $\mathbf{X} \sim \mathcal{N}(\mu, \Sigma)$, then

$$\mathbb{E}(\mathbf{X}) = \mu, \mathbb{V}(\mathbf{X}) = \Sigma,$$

and we call μ the expected value (or mean) vector and Σ the covariance matrix. In particular this means that a multivariate Gaussian distribution is determined by its mean vector and covariance matrix.

The mean vector is a vector with N components. The i th component is denoted by μ_i and it represents the expected value of X_i

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_i \\ \vdots \\ \mu_N \end{bmatrix} = \begin{bmatrix} \mathbb{E}(X_1) \\ \mathbb{E}(X_i) \\ \vdots \\ \mathbb{E}(X_N) \end{bmatrix}.$$

The covariance matrix $\Sigma_{N \times N}$:

- is a squared and symmetric matrix;

- it is positive definite, i.e.

$$\mathbf{x}' \boldsymbol{\Sigma} \mathbf{x} > 0 \quad \forall \mathbf{x} \in R^N, \mathbf{x} \neq \mathbf{0}.$$

- it is made of N variances (on the diagonal and denoted by σ_i^2) and $N \times (N - 1) / 2$ covariances (denoted by σ_{ij} with $\sigma_{ij} = \sigma_{ji}$).

$$\boldsymbol{\Sigma}_{N \times N} = \begin{bmatrix} \sigma_1^2 & & \cdots & & \sigma_{1,N} \\ & \sigma_2^2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \sigma_{N-1}^2 & \\ \sigma_{N,1} & & \cdots & & \sigma_N^2 \end{bmatrix}$$

where

$$\sigma_i^2 = \mathbb{V}(X_i); \sigma_{ij} = \text{Cov}(X_i, X_j) = \mathbb{E}((X_i - \mu_i)(X_j - \mu_j)).$$

Definition 84 We say that the random vector $\mathbf{Z} = (Z_1, \dots, Z_N)$ has a multivariate standard normal distribution if

$$\boldsymbol{\mu} = \mathbf{0}_N, \text{ and } \boldsymbol{\Sigma} = \mathbf{I}_N,$$

where $\mathbf{0}_N$ is a vector of zeros and \mathbf{I}_N is the identity matrix of order N . It follows that Z_i and Z_j are uncorrelated and (being normal) also independent. The density function is simply

$$f_Z(\mathbf{z}) = \frac{1}{\sqrt{(2\pi)^N}} \exp\left(-\frac{1}{2}\mathbf{z}'\mathbf{z}\right) = \frac{1}{\sqrt{(2\pi)^N}} \exp\left(-\frac{1}{2}\sum_{i=1}^N z_i^2\right).$$

10.8.1. THE BIVARIATE NORMAL DISTRIBUTION

If $N = 2$, the mean vector and the covariance matrix can be written as

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}; \boldsymbol{\Sigma}_{2 \times 2} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}; \boldsymbol{\Sigma}_{2 \times 2}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix}$$

where $\rho = \text{Corr}(X_1, X_2)$ and the density function becomes

$$f_X(\mathbf{x}) = \frac{\exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)\right)}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}.$$

The corresponding density is bell-shaped with maximum in (μ_1, μ_2) .

The contour lines, i.e. the combinations of x_1 and x_2 such that the density is constant are described by the equation

$$-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right) = k,$$

that it is the equation of an ellipse. Figures (10.13) - (10.15) show the bivariate Normal density (left panel) and the corresponding contour plot (right panel) for different levels of the correlation, ρ .

Matlab Code

```
%%%%%%PLOT OF THE BIVARIATE GAUSSIAN%%%%%
%%%POSITIVE CORRELATION%%%%%
%%%%%
%correlation
rho=0.9;
%variances and covariances
s1=0.3^2; s2=0.5^2; s12=rho*(s1*s2)^0.5;
factor=2*pi*(s1*s2)^0.5;
m=zeros(2,1); %mean vector
npoints1=50; npoints2=50;
x1=linspace(m(1)-3*s1^0.5,m(1)+3*s1^0.5,npoints1);
x2=linspace(m(2)-3*s2^0.5,m(2)+3*s2^0.5,npoints2);
[x1mesh, x2mesh]=...
meshgrid((x1-m(1))/s1^0.5, (x2-m(2))/s2^0.5);
x1x2mesh=-2*rho*x1mesh.*x2mesh;
pdfbiv=exp(-(x1mesh.^2+x1x2mesh+x2mesh.^2)...
/(2*(1-rho^2)))/(factor*(1-rho^2)^0.5);
figure1=figure('PaperSize',[20.98 29.68], 'Color',[1 1 1]);
subplot(1,2,1); surf(x1mesh,x2mesh,pdfbiv)
axis square
title('Bivariate Normal pdf, \rho=0.9')
ylabel('x_1'), xlabel('x_2'), zlabel('PDF')
subplot(1,2,2); contour(x1mesh,x2mesh,pdfbiv)
xlabel('x_1'), ylabel('x_2')
title('Contour Level, \rho=0.9')
axis square
grid on
```

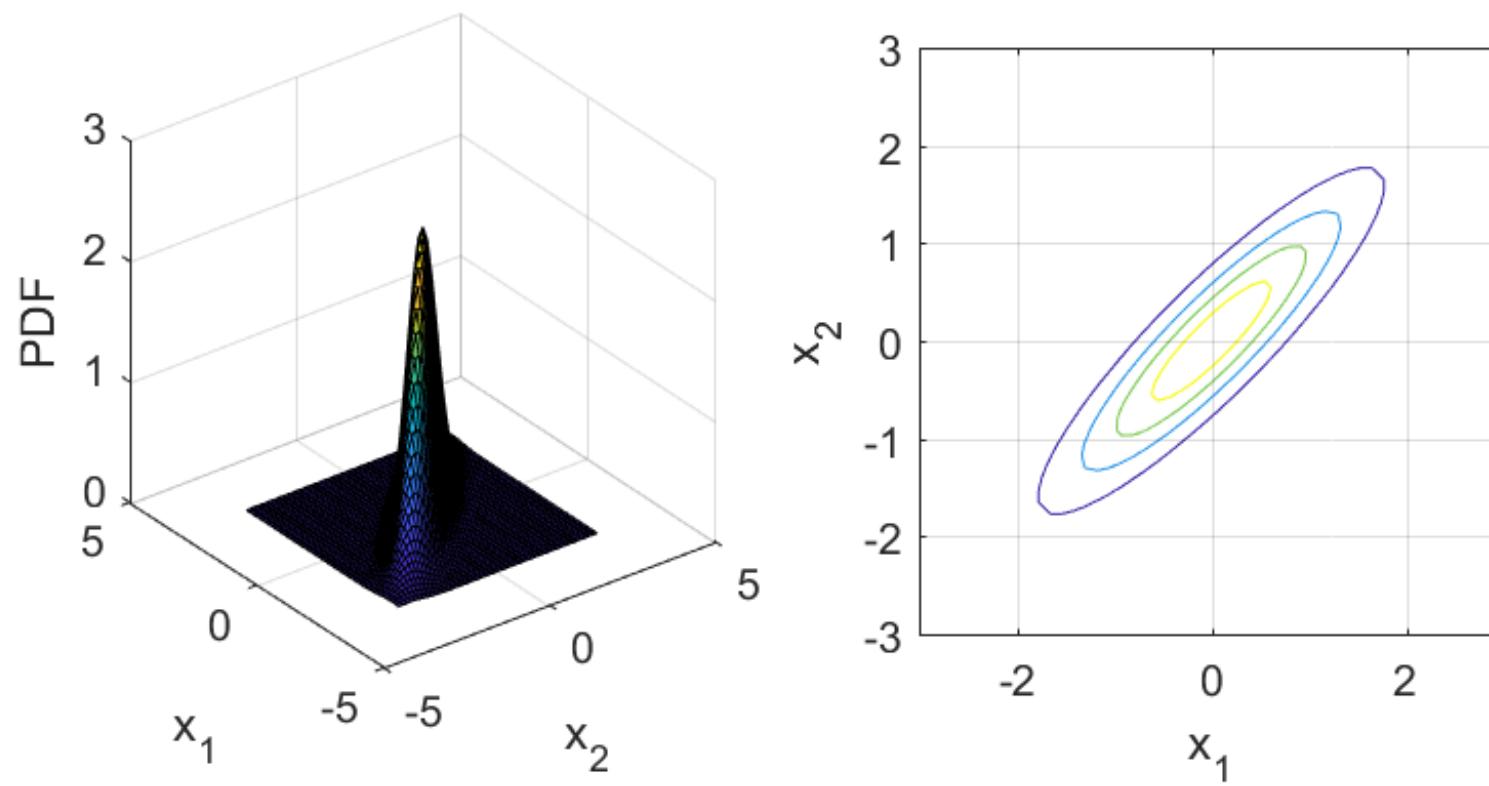


Figure 10.13: Left: Bivariate standard Gaussian distribution. Right: Contour Level. Parameter $\rho = 0.9$.

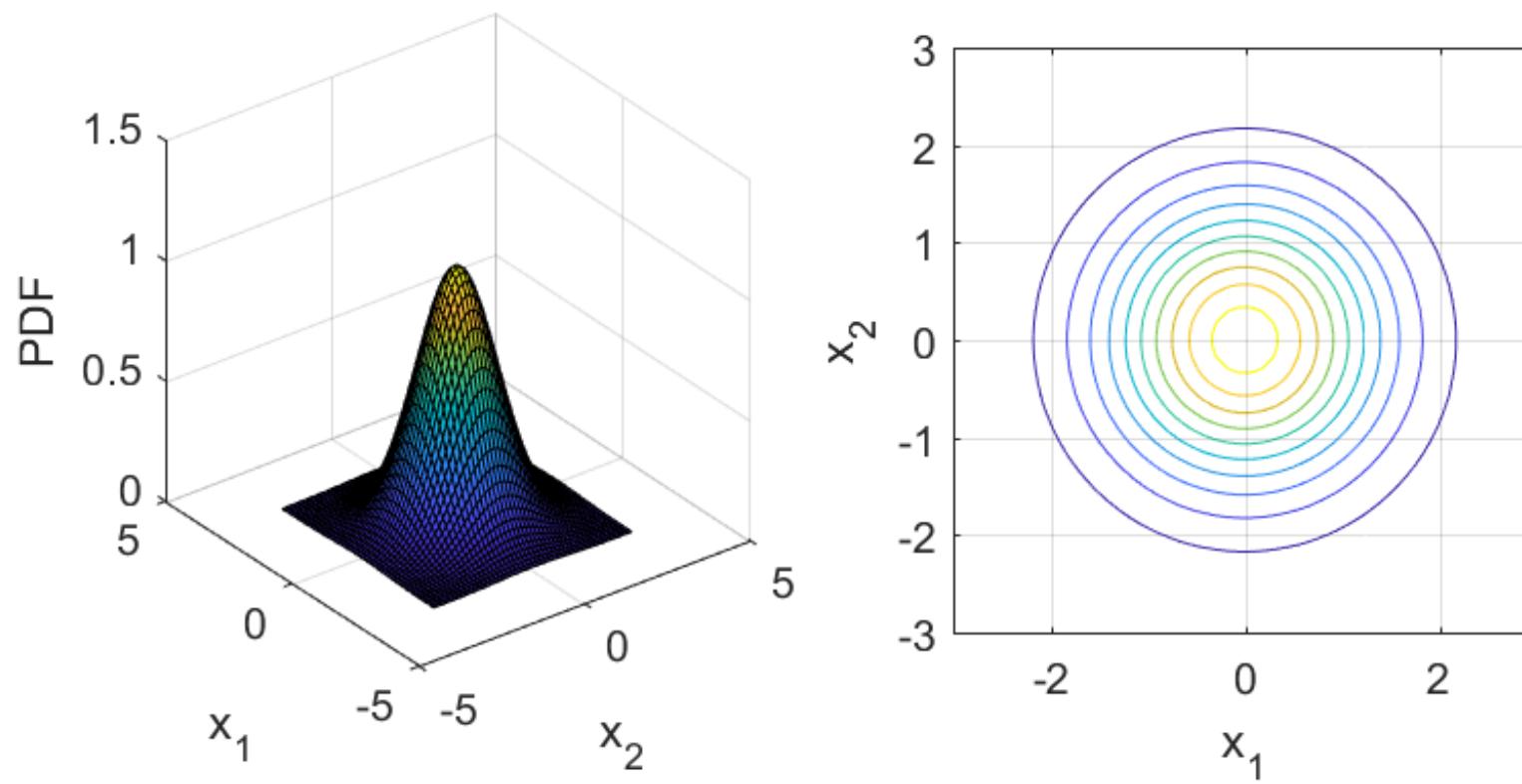


Figure 10.14: Left: Bivariate standard Gaussian distribution. Right: Contour Level. Parameter $\rho = 0.0$.

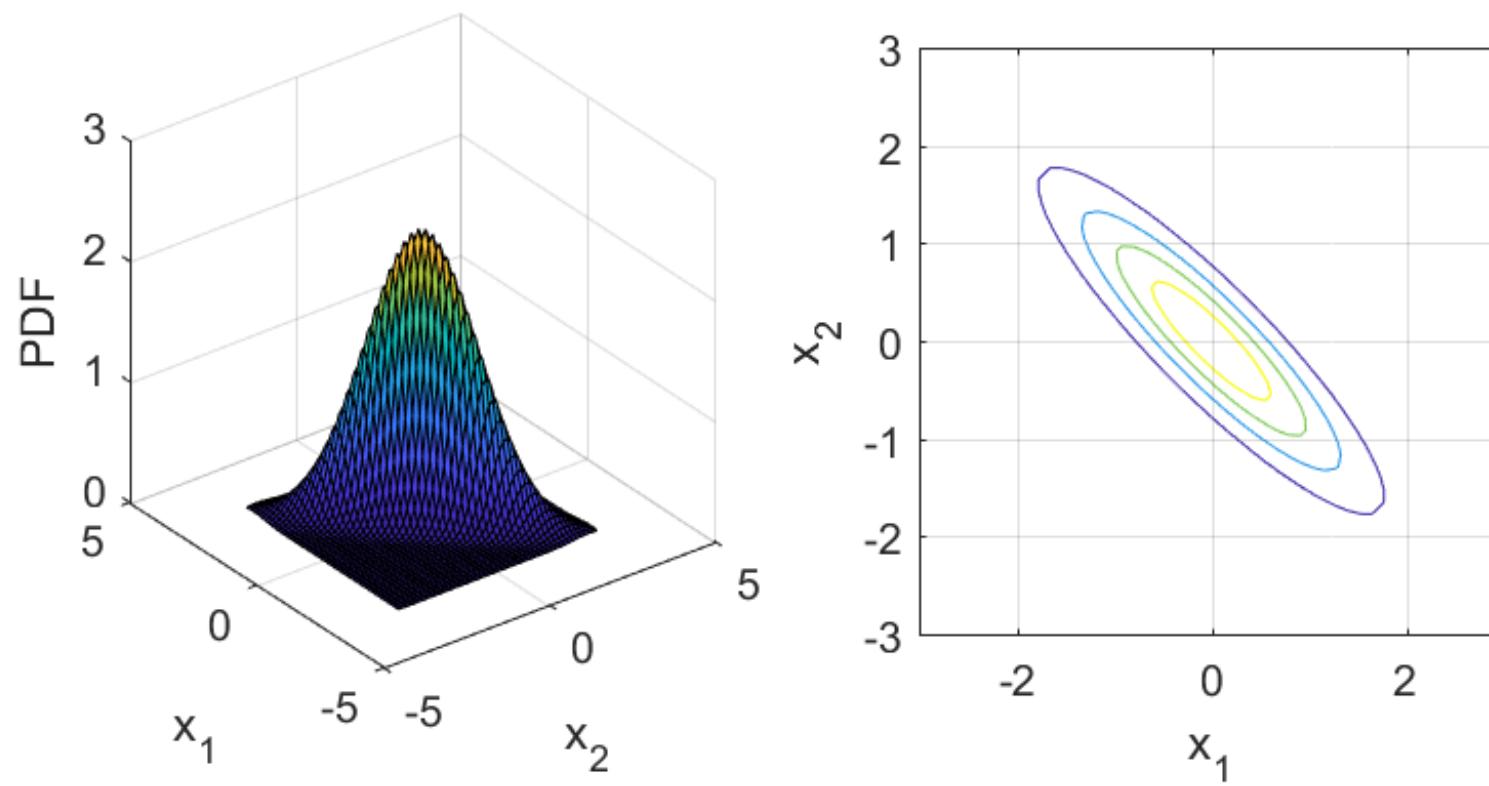


Figure 10.15: Left: Bivariate standard Gaussian distribution. Right: Contour Level. Parameter $\rho = -0.9$.

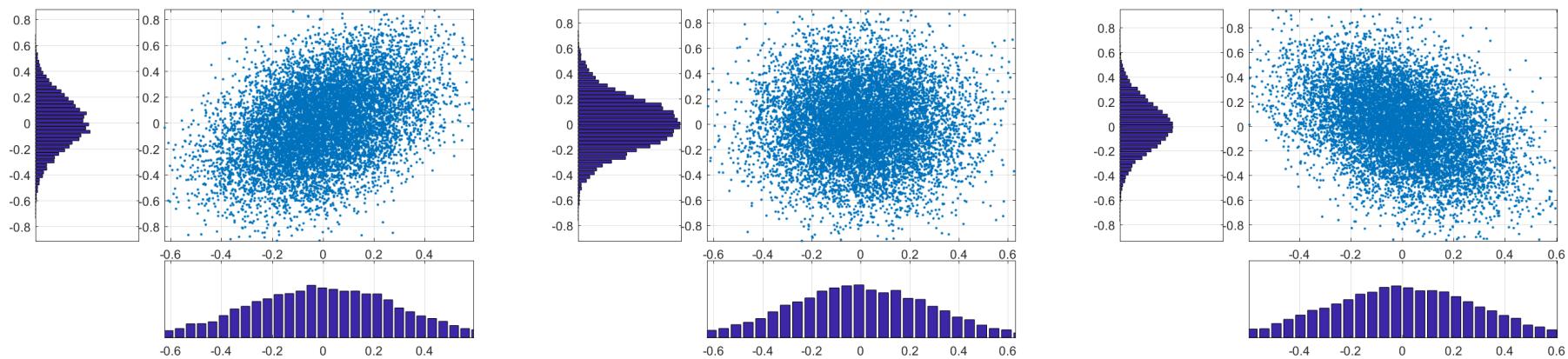


Figure 10.16: Jointly simulated Gaussian random variables (both having zero mean and unit standard deviation) with different correlation. Left: $\rho = 0.9$. Mid: $\rho = 0$. Right: $\rho = -0.9$.

An important fact about sum of Gaussian random variables

Fact 85 *The marginal distributions of a vector \mathbf{X} can all be Gaussian without the joint being multivariate Gaussian.*

Example 86 Let $X_1 \sim \mathcal{N}(0, 1)$ and

$$X_2 = \begin{cases} X_1 & \text{if } -c < X_1 < c \\ -X_1 & \text{elsewhere} \end{cases}$$

We have

$$\Pr(X_2 < x) = \begin{cases} \Pr(X_1 < x) & \text{if } -c < X_1 < c \\ \Pr(X_1 > -x) & \text{elsewhere} \end{cases}$$

and the density is

$$f_{X_2}(x) = \begin{cases} f_{X_1}(x) & \text{if } -c < X_1 < c \\ f_{X_1}(x) & \text{elsewhere} \end{cases} = f_{X_1}(x)$$

and therefore it is Gaussian as well. This is illustrated in Figure (10.17).

Fact 87 However, the sum of two Gaussian random variables is in general not Gaussian.

Example 88 Let us consider the previous example and let us build the new random variable

$$Y = X_1 + X_2 = \begin{cases} 2X_1 & \text{if } -c < X_1 < c \\ 0 & \text{elsewhere} \end{cases}$$

that clearly it is not Gaussian, as shown in Figure (10.18).

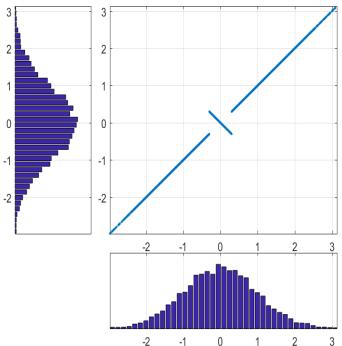


Figure 10.17: Margins are Gaussian, the joint distribution is not Gaussian. We represent the marginal distribution of X_1 and X_2 and their joint distribution as in example 86.

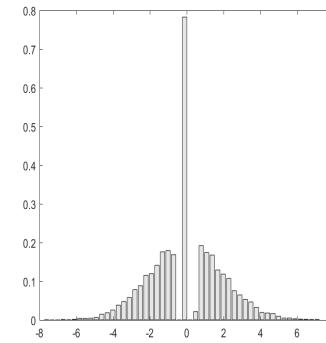


Figure 10.18: The distribution of the sum of $Y = X_1 + X_2$, where both random variables are Gaussian, but they are not jointly Gaussian as discussed in example 88. The random variable sum Y is not Gaussian.

10.9. SIMULATING RANDOM VARIABLES

Many applications entail sampling random variables from pre-specified distributions. A very popular general technique to achieve this is the inverse transform method. This method is implementable if the cumulative density function and its inverse can be computed without difficulties. The idea is formalized in Fact 89 and illustrated in Figure (10.19).

Fact 89 Suppose we need to simulate a random variable X with cumulative distribution function F_X , i.e. such that $\mathbb{P}(X \leq x) = F_X(x)$. Then, the inverse transform method sets

$$X = F_X^{-1}(U), \quad U \sim \text{Unif}[0, 1],$$

where F_X^{-1} is the inverse of F_X and $\text{Unif}[0, 1]$ denotes the uniform distribution on $[0, 1]$.

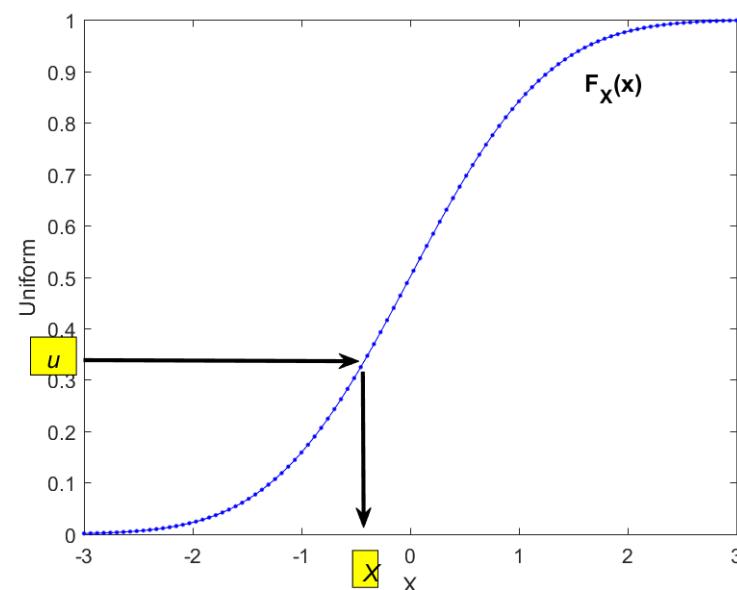


Figure 10.19: Inverse transform method for a hypothetical CDF F_X .

10.9.1. EXAMPLE: SAMPLING FROM THE NORMAL DISTRIBUTION

We refer to the properties presented in section 10.1.

- In Excel simulation of Gaussian r.v.'s can be performed by inputting in a cell the command `=NORMSINV(RAND())`
- In Matlab, we can simulate Gaussian random numbers by simulating Uniform random numbers through the command

```
>U=rand(1,1)
```

and then applying to it the inverse Gaussian CDF

```
>Z=norminv(U, 0, 1);
```

- If we are interested in generating Gaussian random variables with assigned mean μ and standard deviation σ we can use

```
>X=norminv(U, mu, sigma);
```

This is illustrated in detail below at page 253.

- A more direct approach in Matlab is through the command

```
>Z=randn(1,1);
```

- If we are interested in generating Gaussian random variables with assigned mean μ and standard deviation σ we just do

```
>X=mu+Z*sigma;
```

This is illustrated in detail below at page 255.

Matlab Code

```
%%%%%%
% SIMULATION OF STD GAUSSIAN RANDOM VARIABLES BY INVERSION %
%%%%%
clear all; close all
%assign parameters
NSim=10^6;
%simulate a sample of Uniform(0,1) random variables
R=rand(1,NSim);
%simulate the sample of standard Gaussian r.v. by inversion
X = norminv(R,0,1);
%make the plot
h=figure('Color',[1 1 1]);
subplot(4,4,[1 5 9]);
u=linspace(0,1,50);
hist(R,u); view(90,-90)
title('Uniform r.v.')
set(gca,'XTick',[0:0.1:1])
set(gca,'YTickLabel',''); grid on
subplot(4,4,[2:4 6:8 10:12]);
x=linspace(-4,4,100);
f = kSDEnsity(X,x,'function','cdf');
plot(x,f);
title('Inversion of the cdf')
set(gca,'XTick',[-4:1:4])
set(gca,'YTick',[0:0.1:1]); grid on
subplot(4,4,[14:16]);
histfit(X,100)
title('Standard Gaussian r.v.')
set(gca,'XTick',[-4:1:4])
set(gca,'YTickLabel',''); grid on
```

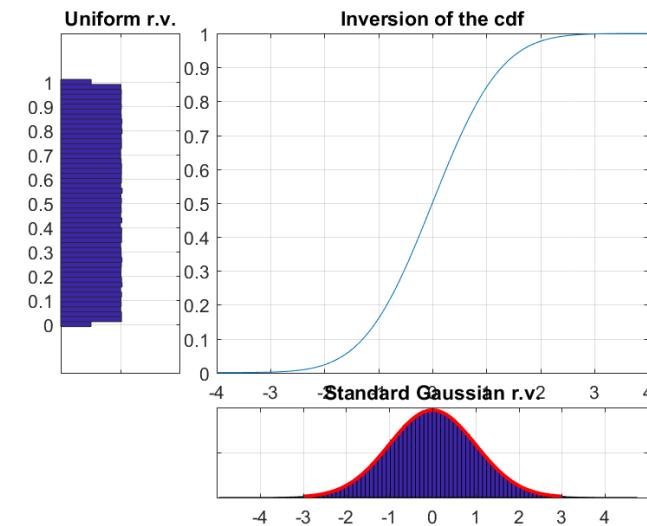


Figure 10.20: Sampling from the standard Gaussian distribution.

Matlab Code

```
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
% SIMULATION OF GAUSSIAN RANDOM VARIABLE BY INVERSION %
% % % % % % % % % % % % % % % % % % % % % % % % % % % % %
clear all; close all
%assign parameters
NSim=10^6; mu=0.2; sg=0.1;
%simulate a sample of Uniform(0,1) random variables
R=rand(1,NSim);
%simulate the sample of non standard Gaussian r.v. by inversion
X = norminv(R,mu,sg);
h=figure('Color',[1 1 1]);
subplot(4,4,[1 5 9]);
u=linspace(0,1,50);
hist(R,u); view(90,-90)
title('Uniform r.v.')
set(gca,'XTick',[0:0.1:1])
set(gca,'YTickLabel',''); grid on
subplot(4,4,[2:4 6:8 10:12]);
x=linspace(-0.3,0.5,100);
f = kSDEnsity(X,x,'function','cdf');
plot(x,f);
title('Inversion of the cdf')
set(gca,'XTick',[-0.3:0.1:0.5])
set(gca,'YTick',[0:0.1:1]); grid on
subplot(4,4,[14:16]);
histfit(X,100)
title('Gaussian r.v.'); xlim([-0.3 0.5]);
set(gca,'XTick',[-0.3:0.1:0.5])
set(gca,'YTickLabel',''); grid on
```

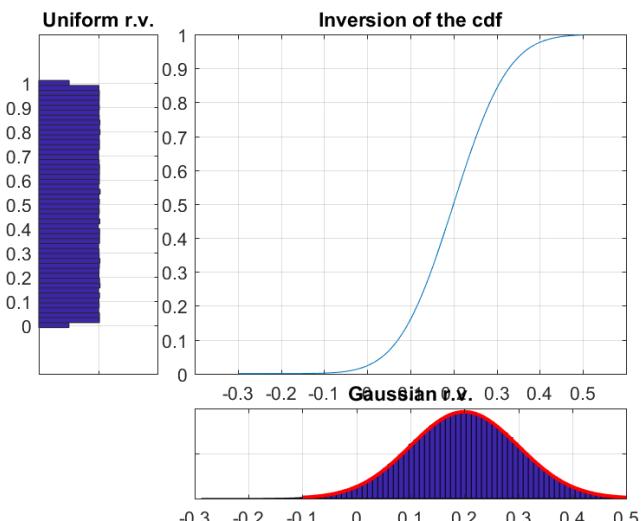


Figure 10.21: Sampling from the Gaussian distribution with $\mu = 0.2$ and $\sigma = 0.1$.

10.9.2. EXAMPLE: SAMPLING FROM THE LOGNORMAL DISTRIBUTION

We can exploit the relationship between Gaussian and Lognormal distribution. In the following lines we simulate 10000 Log-normal variables with parameters $\mu = 0.05$ and $\sigma = 0.2$.

```
>mu=0.05;
>sigma=0.2
>X=mu+Z*randn(10000,1);
>Y=exp(X);
```

Matlab Code

```
%%%%% SIMULATION OF LOGNORMAL RANDOM VARIABLES %%%
clear all;
close all
%assign parameters
NSim=10^6;
mu=0.05;
sigma=0.2;
%simulate the sample of Gaussian r.v.
X = mu+sigma*randn(NSim,1);
%simulate the Lognormal
%by exponential transformation
Y=exp(X);
h=figure('Color',[1 1 1]);
histfit(Y,100,'lognormal');
title('Simulated Lognormal Random Variables')
```

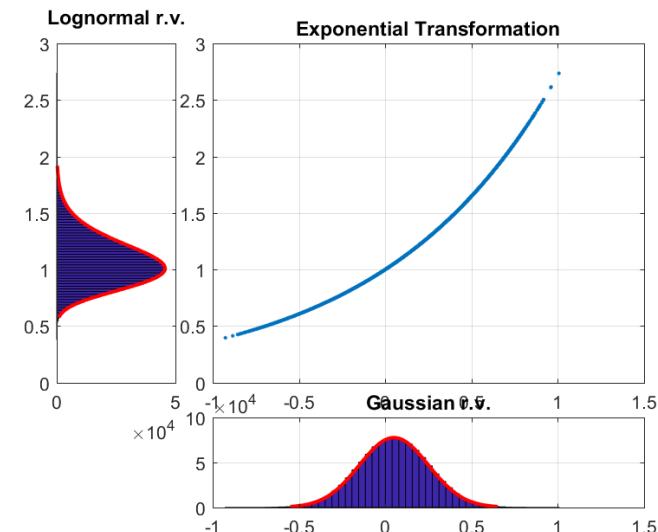


Figure 10.22: Sampling from the Lognormal distribution with parameters $\mu = 0.05$ and $\sigma = 0.2$.

10.9.3. EXAMPLE: SAMPLING FROM THE POISSON DISTRIBUTION

We refer to the properties presented in section 10.5. This is illustrated in detail below at page 257.

10.9.4. EXAMPLE: SAMPLING FROM THE GAMMA DISTRIBUTION

We refer to the properties presented in section 10.7. This is illustrated in detail below at page 258.

10.9.5. EXAMPLE: SAMPLING FROM THE CHI-SQUARE DISTRIBUTION

A Chi-Square distribution is a special case of the Gamma distribution with shape parameter α . Therefore, the simulation turns out to be a special case of the Gamma. However, notice that Matlab has a built-in function to sample directly from the chi-square distribution. In the following lines we simulate 10000 chi-square random variables with degrees of freedom 5.

```
>df=5;
>chi2=chi2rnd(df,10000,1);
```

10.9.6. EXAMPLE: SAMPLING FROM THE NON-CENTRAL CHI-SQUARE DISTRIBUTION

Matlab provides a routine that allows to compute the inverse CDF of this distribution, (ncx2inv(p,df,nc) where p is the probability level, df is the number of degrees of freedom and nc is the parameter of non-centrality). Matlab also provides a random number generator for this distribution. The function ncx2rnd(df,nc, m,n) returns a matrix mxn of random numbers chosen from the non-central chi-square distribution with degrees of freedom df and positive noncentrality parameters nc. In the following lines we simulate 10000 chi-square random variables with degrees of freedom 5.9 and parameter of non-centrality equal to 4.5.

```
>df=5.9;
>nc=4.5
>chi2nc=ncx2rnd(df,nc, 10000,1);
```

Matlab Code

```
%%%%%%
% SIMULATION OF POISSON RANDOM VARIABLE BY INVERSION %
%%%%%
clear all; close all
%assign parameters
NSim=10^6; lambdaP=10;
%simulate a sample of Uniform(0,1) random variables
R=rand(1,NSim);
%simulate the sample of Poisson r.v. by inversion
X = poissinv(R,lambdaP);
h=figure('Color',[1 1 1]);
subplot(4,4,[1 5 9]);
u=linspace(0,1,50);
hist(R,u); view(90,-90)
title('Uniform r.v.')
set(gca,'XTick',[0:0.1:1])
set(gca,'YTickLabel','')
grid on
subplot(4,4,[2:4 6:8 10:12]);
x=(0:1:30);
f = kSDEnsity(X,x,'function','cdf');
plot(x,f,'-o');
title('Inversion of the cdf')
set(gca,'XTick',[0:5:30])
set(gca,'YTick',[0:0.1:1]); grid on
subplot(4,4,[14:16]);
histfit(X,15,'poisson')
title('Poisson r.v.')
set(gca,'XTick',[0:5:30])
set(gca,'YTickLabel',''); grid on
```

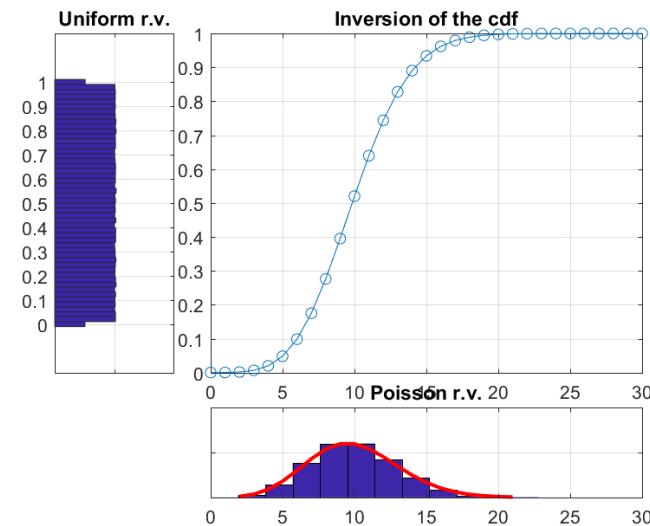


Figure 10.23: Sampling from the Poisson distribution with $\lambda = 10$.

Matlab Code

```
%%%%%%
% SIMULATION OF GAMMA RANDOM VARIABLE BY INVERSION %
%%%%%
clear all; close all
%assign parameters
NSim=10^6; alpha=2; lambdaG=0.5;
%simulate the corresponding sample of Gamma r.v. by inversion
X = gaminv(R,alpha,lambdaG^(-1));
h=figure('Color',[1 1 1]); subplot(4,4,[1 5 9]);
u=linspace(0,1,50); hist(R,u); view(90,-90);
title('Uniform r.v.')
set(gca,'XTick',[0:0.1:1])
set(gca,'YTickLabel',''); grid on
subplot(4,4,[2:4 6:8 10:12]);
x=linspace(0,20,200);
f = kSDEnsity(X,x,'function','cdf');
plot(x,f);
title('Inversion of the cdf')
set(gca,'XTick',[0:2:20])
set(gca,'YTick',[0:0.1:1]); grid on
subplot(4,4,[14:16]); histfit(X,100,'gamma')
title('Gamma r.v.'); xlim([0 20]);
set(gca,'XTick',[0:2:20])
set(gca,'YTickLabel',''); grid on
```

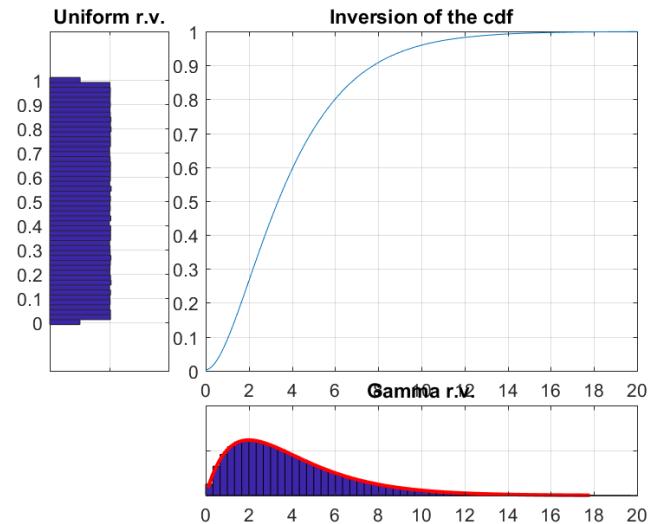


Figure 10.24: Sampling from the Gamma distribution with $\alpha = 2, \lambda = 0.5$.

10.9.7. EXAMPLE: SAMPLING FROM THE MULTIVARIATE NORMAL DISTRIBUTION

We refer to the properties presented in section 10.8. Further, we note the following.

Fact 90 $X \sim \mathcal{N}(\mu, \Sigma)$ if and only if it admits the representation

$$X = \mu + AZ, \quad (10.1)$$

where $\mathbf{Z} = (Z_1, \dots, Z_N)$ is a multivariate standard normal distribution and \mathbf{A} is a $N \times N$ lower triangular matrix with strictly positive diagonal entries.

Fact 91 It follows that:

$$\Sigma = \text{Var}(\mathbf{X}) = \text{Var}(\mathbf{AZ}) = \mathbf{A}\text{Var}(\mathbf{Z})\mathbf{A}' = \mathbf{A}\mathbf{I}\mathbf{A}' = \mathbf{AA}',$$

and we have the so called Cholesky decomposition of Σ

$$\Sigma = \mathbf{AA}'$$

and this guarantees that Σ is semidefinite positive. If $\text{rank}(\mathbf{A}) = N$, Σ is definite positive. This is known as Cholesky decomposition

Example 92 Let

$$\Sigma = \begin{bmatrix} 0.04 & 0.024 \\ 0.024 & 0.09 \end{bmatrix}.$$

In order to find its Cholesky decomposition, we need to look for a 2×2 lower triangular matrix \mathbf{A} such that

$$\begin{bmatrix} 0.04 & 0.024 \\ 0.024 & 0.09 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{11}a_{21} & a_{21}^2 + a_{22}^2 \end{bmatrix},$$

i.e. we have to set

$$\begin{cases} a_{11}^2 = 0.04 \\ a_{11}a_{21} = 0.024 \\ a_{21}^2 + a_{22}^2 = 0.09 \end{cases} \implies \begin{cases} a_{11} = \sqrt{0.04} = 0.2 \\ a_{21} = \frac{0.024}{0.2} = 0.12 \\ a_{22} = \sqrt{0.09 - (0.12)^2} = \sqrt{0.0756} = 0.27495 \end{cases}$$

and therefore

$$\mathbf{A} = \begin{bmatrix} 0.2 & 0 \\ 0.12 & 0.2749545 \end{bmatrix}.$$

- The Cholesky decomposition is at the basis of simulation methods for multivariate Gaussian random variables.

- It can be verified that if

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

then

$$\mathbf{A} = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix}.$$

- This means that in order to perform Monte Carlo simulation we need to simulate Z_1 and Z_2 independently and according to a standard normal random variable and then to set

$$\begin{aligned} X_1 &= \mu_1 + \sigma_1 Z_1, \\ X_2 &= \mu_2 + \rho\sigma_2 Z_1 + \sigma_2\sqrt{1-\rho^2} Z_2. \end{aligned}$$

Example 93 In the bivariate case, we can apply the Cholesky decomposition of previous example and obtain

$$\Sigma = \begin{bmatrix} 0.04 & 0.024 \\ 0.024 & 0.09 \end{bmatrix} \implies \mathbf{A} = \begin{bmatrix} 0.2 & 0 \\ 0.4 \times 0.3 & 0.3 \times \sqrt{1-0.4^2} \end{bmatrix}$$

so that correlated Gaussian r.v's are simulated according to

$$\begin{aligned} X_1 &= \mu_1 + 0.2 \times Z_1, \\ X_2 &= \mu_2 + 0.4 \times 0.3 \times Z_1 + 0.3 \times \sqrt{1-0.4^2} \times Z_2. \end{aligned}$$

In Matlab we can simulate a multivariate Gaussian distribution using the command mvnrnd that takes as parameters the mean vector, the covariance matrix and the number of simulations. This is illustrated in details in the following Matlab script.

Matlab Code

```
%%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%%
%%%%% SIMULATION OF BIVARIATE NORMAL RANDOM VARIABLE %%%%%%
%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%% %%%%%%
clear all
close all
%assign parameters
NSim=10^6;
rho=0.9; mu1=0; mu2=0;
s1=0.3^2; s2=0.5^2; s12=rho*(s1*s2)^0.5;
%simulate the corresponding sample
%of bivariate Gaussian r.v.
Mu=[mu1 mu2]; VC=[s1 s12; s12 s2];
X = mvnrnd(Mu,VC,10000);
X1=X(:,1); X2=X(:,2);
h=figure('Color',[1 1 1]);
subplot(4,4,[1 5 9]);
histfit(X1,100)
set(gca,'YTickLabel','')
view(90,-90)
grid on
subplot(4,4,[2:4 6:8 10:12]);
plot(X2,X1,'+')
grid on
subplot(4,4,[14:16]);
histfit(X2,100)
set(gca,'YTickLabel','')
grid on
```

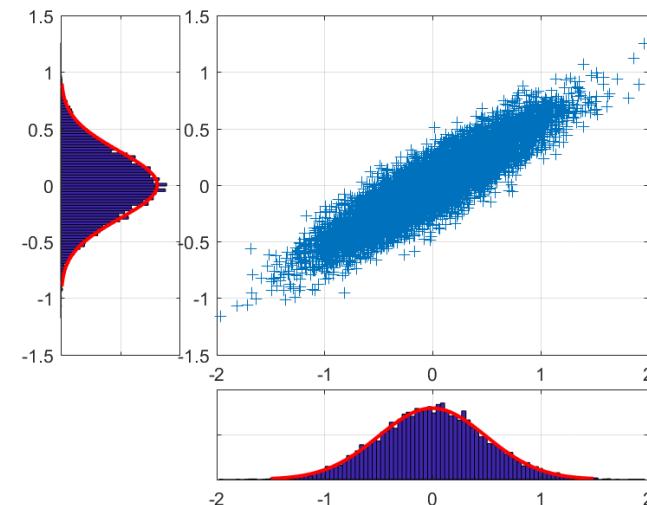


Figure 10.25: Sampling from the bivariate Gaussian distribution (positive correlation).

CHAPTER 11. QUADRATURE METHODS

Quadrature methods allow for numerical computations of integrals of the form

$$I(f) = \int_a^b f(x) dx,$$

where a and b can be finite or infinite, provided that the integral exists and is finite.

In quantitative finance, the numerical computation of the above quantity is important because computing the expected payoff, consists of, whenever the distribution of the underlying variable is available in closed form, computing an integral. The numerical computation of the integral turns out to be very effective for low-dimensional problems.

11.1. QUADRATURE RULES

Consider an integral

$$I(f) = \int_A f(x) dx,$$

where f is a real-valued function defined on a closed and bounded subset A of the real line. This quantity is traditionally computed by using an approximating finite sums corresponding to a suitable partition of the set A .

The basic strategy for computing $I(f)$ is as follows:

- Fix the so-called abscissas or grid points x_i and sample the function f at these points;
- Approximate function f by an interpolating polynomial p_n of order n in a way that $p_n(x_i) = f(x_i)$ for all i 's;
- Integrate p_n and return $I(p_n)$ as approximation to $I(f)$.

Using Newton-Cotes formulae, abscissas x_i are evenly spaced. In Gaussian quadrature formulae, abscissas are selected in a way to maximize the order of the interpolating polynomial. Both methods can be combined with the so called adaptive quadrature, where the abscissas are selected to better capture the behaviour of the function f : a finer grid will be chosen where the function changes very fast, a sparse grid if the function behaviour is very smooth.

11.2. NEWTON-COTES INTEGRATION

In practice, Newton-Cotes formulae are usually not applied to the entire interval of integration $[a, b]$. A popular practice consists of splitting this interval into m evenly spaced subintervals $[x_j, x_{j+1}]$, *i.e.*,

$$x_j = a + (j - 1) h$$

with

$$h = \frac{b - a}{m}.$$

The required integral is given by the sum of the integrals computed on each subinterval, each one being calculated using a Newton-Cotes formula with $n + 1$ equally spaced nodes. The resulting procedure is known as a *composite* Newton-Cotes formula and can be described as follows. The integral

$$I(f) = \int_a^b f(x) dx = \sum_{j=1}^{m-1} \int_{x_j}^{x_{j+1}} f(x) dx$$

is approximated by $I_{n,m}(f)$ defined as

$$I_{n,m}(f) = h \sum_{j=1}^{m-1} \sum_{i=1}^n f_{i,j} w_{i,j},$$

where $f_{i,j} = f(x_i^j)$ and the x_i^j 's are nodes refining subinterval $[x_j, x_{j+1}]$, and $w_{i,g}$ the so called weights that depend on the order of the approximating polynomial. We stress that integers m and n respectively refer to the number of subintervals and to the order of the interpolating polynomial.

Typical examples are the Rectangle ($n = 0$), the Mid-point ($n = 0$) and the Trapezoidal rules ($n = 1$).

11.2.1. RECTANGLE AND MID-POINT RULE

This is the simplest Newton-Cotes approximation and it approximates the function using a piecewise constant polynomial (so that $n = 0$), using the values attained at one of the two vertices of each subinterval (an even better approximation is obtained considering the midpoint value).

The rectangle rule works rather well whenever function f is smooth and the space between points x_i is adequately small (*i.e.*, m is quite large). Function f is approximated by a piecewise constant function that stays steady on each subinterval. The approximated integration formula reads as

$$I_{0,m}(f) = h \sum_{j=1}^{m-1} f_j,$$

where

$$h = \frac{b-a}{m-1}, x_j = a + (j-1) \times h, j = 1, \dots, m-1.$$

For $f \in C^1([a, b])$, the integration error of the rectangle rule can be shown to be

$$I(f) - I_{0,m}(f) = \frac{h}{2} (b-a) f^{(1)}(\xi),$$

with $\xi \in [a, b]$, showing that the error stemming from the rectangle rule approaches zero at the same rate as h .

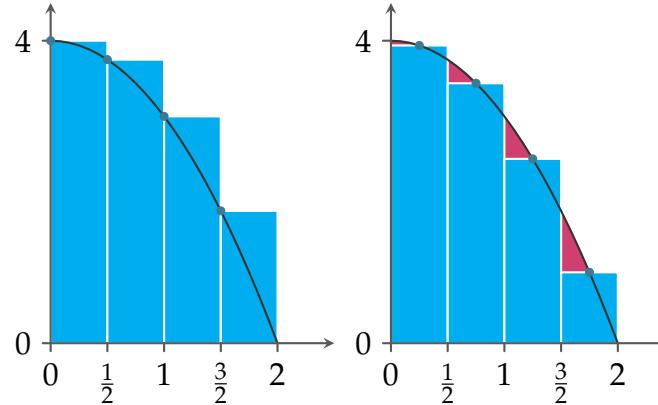


Figure 11.1: Left: Rectangle rule, Right: Mid-point.

In the midpoint approximation, the approximated integration formula reads as

$$I_{1,m}(f) = h \sum_{j=1}^{m-1} f_{j+\frac{1}{2}},$$

where

$$f_{j+\frac{1}{2}} = f\left(\frac{x_{j+1} + x_j}{2}\right).$$

The two methods are illustrated in Figure (11.1).

For $f \in C^2([a, b])$, the integration error of the mid-point rule can be shown to be

$$I(f) - I_{1,m}(f) = \frac{h^2}{24} (b-a) f^{(2)}(\xi),$$

with $\xi \in [a, b]$, showing that the error stemming from the mid-point rule approaches zero at the same rate as h^2 .

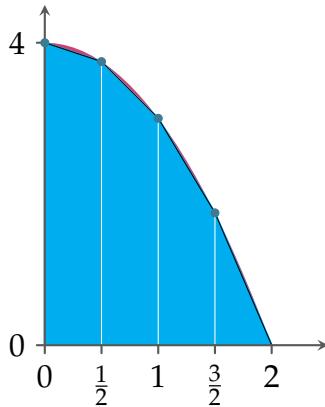


Figure 11.2: Approximation of the integral using trapezium.

11.2.2. TRAPEZOID RULE

This rule approximates the function on each subinterval using linear interpolation (so $n = 1$) as illustrated in Figure 11.2, that is by replacing $\int_{x_j}^{x_{j+1}} f(x) dx$ with

$$(f(x_j) + f(x_{j+1})) \frac{h}{2}.$$

Therefore:

$$\begin{aligned} I_{1,m}(f) &= \frac{h}{2} \sum_{j=1}^{m-1} (f(x_j) + f(x_{j+1})) \\ &= h \left(\frac{1}{2} f(x_1) + f(x_2) + \dots + f(x_{m-1}) + \frac{1}{2} f(x_m) \right). \end{aligned}$$

For $f \in C^2([a, b])$, the integration error of the trapezium rule can be shown to be

$$I(f) - I_{1,m}(f) = \frac{h^2}{12} (b-a) f^{(2)}(\xi),$$

with $\xi \in [a, b]$, showing that the error stemming from trapezoid rule approaches zero at the same rate as h^2 , as for the mid-point rule. For this reason, these rules are said to be of order 2. However, the trapezium rule integrates exactly first-order polynomials, whilst the mid-point rule integrates exactly piecewise constant functions.

If we want to integrate exactly quadratic function we can adopt the Simpson rule. See Fusai and Roncoroni (2008) for more details on this method.

11.2.3. ERROR ANALYSIS FOR POLYNOMIAL RATE OF CONVERGENCE

- Assume that the integration error

$$E_{n,m} = I(f) - I_{n,m}(f) = c_n m^{-p_n}, c \in \mathbb{R}$$

with c and p non depending on m , but on n . Moreover, c depends on the derivative of the function in a internal point of the integration interval.

- For example
 - linear convergence: $p = 1$ (e.g., rectangle rule)
 - quadratic convergence: $p = 2$ (e.g., mid-point rule)
 - quadratic convergence: $p = 2$ (e.g., trapezoidal rule)
 - convergence of order 4: $p = 4$ (e.g., Simpson's rules)

- We write also

$$E_{n,m} = O\left(\frac{1}{m^p}\right), m \rightarrow \infty$$

- In general, estimating an approximation error requires derivatives of f in a internal point of the integration interval.
- If this is not known, we can proceed as follows (see Fusai and Roncoroni, page 163).
 1. Estimate the integral on successively finer grids $m_i < m_{i+1}$, for example by halving the space length h (you can set $m = 2^j + 1, j = 3, \dots, J$)

2. If we use $I_{n,m}$ and $I_{n,2m-1}$ and we have

$$E_{n,m} = c_n \left(\frac{b-a}{m-1} \right)^p \quad \text{and} \quad E_{n,2m-1} = c_n \left(\frac{b-a}{2m-2} \right)^p = c_n \left(\frac{b-a}{m-1} \right)^p \frac{1}{2^p},$$

3. For large m , we have

$$R_{n,m} = \frac{E_{n,m}}{E_{n,2m-1}} = 2^p \rightarrow \log_2(R_{n,m}) \approx p. \quad (11.1)$$

4. Make a plot of $\log_2(m)$ versus $\log_2(R_p)$. You should get, for large m , a constant line at the level p .

- If the value of the integral is not known (as usual!), we proceed as follows:

1. Compute the ratios

$$\tilde{R}_{n,2^j} = \frac{I_{n,2^j+1} - I_{n,2^{j-1}+1}}{I_{n,2^{j+1}+1} - I_{n,2^j+1}},$$

2. Take the logarithm

$$\log_2(\tilde{R}_{n,2^j}) \quad (11.2)$$

3. For large j , this should provide an estimate of p .

4. Make a plot of $\log_2(m)$ versus $\log_2(\tilde{R}(j)_p)$. You should get, for large m , a constant line at the level p .

11.2.4. AN EXAMPLE

Let us consider the computation of the following integral

$$\frac{5}{e^\pi - 2} \int_0^{\pi/2} e^{2x} \cos(x) = 1. \quad (11.3)$$

We begin by defining the function to be integrated using a so called *function handle*.

```
clc; clear all, close all
%define the function to be integrated
fun = @(x) exp(2*x).*cos(x)*5/(exp(pi)-2);
```

11.2.5. NEWTON-COTES INTEGRATION

Then we implement the Rectangle, Midpoint and Trapezium rule. In particular, the Trapezium rule can exploit the Matlab function trapz, that accepts two arguments: the grid and the corresponding function values and applies to each subinterval the trapezium rule¹.

```
i=1; % a counter
for j=2:11
    %number of points in each subinterval
    m(i)=2^j+1;
    %build the grid
    x=linspace(a,b,m(i));
    %grid size
    h=x(2)-x(1);
    %Rectangle Rule
    I_R(i)=sum(fun(x(1:end-1)))*h;
    %Mid-Point Rule
    x_mid=x+h/2;
    I_MP(i)=sum(fun(x_mid(1:end-1)))*h;
    %Trapezium rule
    I_T(i)=trapz(x,fun(x));
    i=i+1;
end
%We compare the results
out=[m', I_R', I_MP', I_T'];
T = table(m',I_R',I_MP',I_T','VariableNames',{'m' 'Rect' 'Mid' 'Trap'})
```

¹Notice, that we can pass to the trapz function a not-evenly spaced grid

Table 11.5 shows the behaviour of approximated values $I_{n,m}$. Table 11.6 shows the behaviour of the errors

$$E_{n,m} = I(f) - I_{n,m}(f). \quad (11.4)$$

m	R	MP	T
5	0.97200380	1.03647822	0.92556504
9	1.00424101	1.00944240	0.98102163
17	1.00684171	1.00238106	0.99523202
33	1.00461138	1.00059655	0.99880654
65	1.00260397	1.00014922	0.99970154
129	1.00137659	1.00003731	0.99992538
257	1.00070695	1.00000933	0.99998134
513	1.00035814	1.00000233	0.99999534
1025	1.00018024	1.00000058	0.99999883
2049	1.00009041	1.00000015	0.99999971

Table 11.1: Numerical integration of the integral in 11.7.

Table 11.7 gives the ratios

$$R_{n,m} = |I(f) - I_{n,m}(f)| / |I(f) - I_{n,2m-1}(f)| \quad (11.5)$$

across alternative values for m . Following estimates of the error for Newton-Cotes formulas, see expressions (11.2.1) and (11.2.2), this ratio is expected to converge, for large values of m , to 4 for the Trapezium rule, and for the Midpoint rule and to 2 for the Rectangle Rule. In addition, in Table 11.4 we show the estimate of the order of convergence of the rule according to the rule described in 11.2.

m	R	MP	T
5	0.0279962	-0.0364782	0.0744350
9	-0.0042410	-0.0094424	0.0189784
17	-0.0068417	-0.0023811	0.0047680
33	-0.0046114	-0.0005965	0.0011935
65	-0.0026040	-0.0001492	0.0002985
129	-0.0013766	-0.0000373	0.0000746
257	-0.0007070	-0.0000093	0.0000187
513	-0.0003581	-0.0000023	0.0000047
1025	-0.0001802	-0.0000006	0.0000012
2049	-0.0000904	-0.0000001	0.0000003

Table 11.2: Numerical integration error in the computation of the integral in 11.7 using different rules.

11.2.6. LIMITS OF NEWTON-COTES RULES

- Newton-Cotes formulae use a fixed number of equally spaced abscissas.

m	R	MP	T
9	6.601299	3.863235	3.922095
17	0.619876	3.965633	3.980377
33	1.483656	3.991398	3.995085
65	1.770908	3.997849	3.998771
129	1.891604	3.999462	3.999693
257	1.947225	3.999866	3.999923
513	1.973955	3.999966	3.999981
1025	1.987062	3.999992	3.999995
2049	1.993552	3.999998	3.999999

Table 11.3: Ratio $R_{n,m} = |I(f) - I_{n,m}(f)| / |I(f) - I_{n,2m-1}(f)|$ for different quadrature methods and different number of grid points.

Table 11.4: Trapezoid Rule: Estimate of the order of convergence

m	$\log_2(m - 1)$	Estimate	Difference	Ratio	$\log_2(Ratio)$
5	2.321928	0.925565			
9	3.169925	0.981022	0.055457		
17	4.087463	0.995232	0.01421	3.902539	1.964413
33	5.044394	0.998807	0.003575	3.975467	1.991124
65	6.022368	0.999702	0.000895	3.993856	1.997782
129	7.011227	0.999925	0.000224	3.998463	1.999446
257	8.005625	0.999981	5.6E-05	3.999616	1.999861
513	9.002815	0.999995	1.4E-05	3.999904	1.999965
1025	10.00141	0.999999	3.5E-06	3.999976	1.999991
2049	11.0007	1	8.74E-07	3.999994	1.999998

- Moreover, they are usually implemented in a composite form: the interval of integration is divided into subintervals of equal size, then the rule is applied to each subinterval, and finally the sum of resulting numbers is taken as an estimate of the overall integral.
- In this framework, convergence to the actual value of the integral can be rather slow.
- Consequently, an accurate estimation may require a excessively large number of subintervals and corresponding evaluations of the integrand function.

11.3. GAUSSIAN QUADRATURE FORMULAE

Newton-Cotes formulae use a fixed number of equally spaced abscissas. Moreover, they are usually implemented in a composite form: the interval of integration is divided into subintervals of equal size, then the rule is applied to each subinterval, and finally the sum of resulting numbers is taken as an estimate of the overall integral. In this framework, convergence to the actual value of the integral can be rather slow. Consequently, an accurate estimation may require a excessively large number of subintervals and corresponding evaluations of the integrand function.

Gaussian quadrature rules select n abscissas and n weights to produce a rule of order $2n - 1$, so that

$$I_n(f) = \sum_{i=1}^n w_i f(x_i).$$

This way we can exactly integrate polynomials of degree $2n - 1$, that is the highest degree for which a polynomial can be integrated using n points. Therefore, compared to the Newton-Cotes formulas, we may freely select the abscissas at which the integrand function is evaluated. Details on the construction of the abscissas and weights of the Gaussian quadrature rule, depending also on the integration domain, can be found in Chapter 6 of Fusai and Roncoroni (2008).

We only observe that the error in the quadrature rule can be evaluated by the following estimate:

$$\int_a^b f(x) dx - \sum_{i=1}^n w_i f(x_i) = \frac{(b-a)^{2n+1} (n!)^4}{(2n+1) (2n!)^3} f^{(2n)}(\xi), \quad (11.6)$$

where $\xi \in (a, b)$, provided that $f^{(2n)}(\xi)$ is continuous on $[a, b]$. The error term can be bounded by an exponential rate of decrease as a function of n . The trapezoidal and the Simpson rules have only polynomial rates of decrease, i.e. $1/n^2$ and $1/n^4$. Gaussian quadrature results to be always better than the trapezoidal rule, except in the case of periodic integrands.

11.4. ADAPTIVE QUADRATURE

This method is based on the following idea. We split the integration range $[a, b]$ and then keep on refining it until the composite quadrature formula produces the required level of accuracy. First, we integrate $f(x)$ by using two numerical methods and come up with approximations I_1 and I_2 . With no loss of generality, we may assume that I_1 is more accurate than I_2 . If the difference of the two approximations is smaller than some prescribed tolerance, one accepts I_1 as the value of the integral. Otherwise, the interval $[a, b]$ is split into $[a, c]$ and $[c, b]$, with $c = (a + b)/2$ and the corresponding two integrals are computed independently of each other. Splitting intervals into subintervals continues until the stopping criterion

$$\frac{|I_1 - I_2|}{|I_1|} < tol$$

is met for each integral. Here, I_1 and I_2 are two estimates for the integral computed on the subinterval under consideration and tol is a predefined tolerance. In order to prevent the procedure from generating an excessive number of subdivisions, the suggested criterion is usually implemented with additional conditions.

Figure 11.3 illustrates the way function quad adaptively splits the integration range $[0.1, \pi/2]$ for the purpose of integrating $\sin(1/x)$. As it is clear from the graph, the integrand is more thoroughly sampled in those portions of the domain where integration is more challenging, namely the ones where the integrand function exhibits an oscillatory behaviour

A typical problem arising from the use of adaptive routines is encountered while dealing with discontinuous functions. In this case, a large number of successive evaluations may be produced by the algorithm. In order to circumvent this problem, it may be useful to apply an adaptive quadrature routine on each of the two sides neighbouring the point of discontinuity.

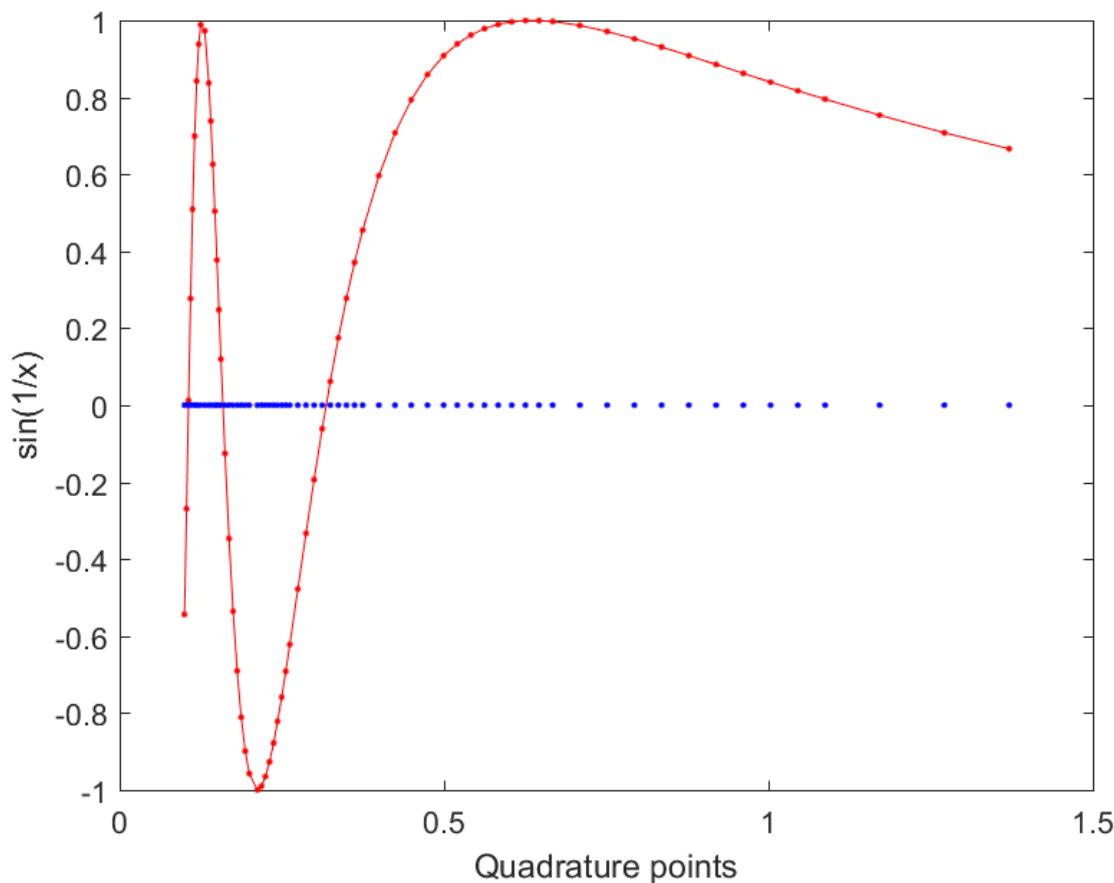


Figure 11.3: Matlab adaptive routine `quad` is used to integrate $\sin(1/x)$ over the interval $[0.1, \pi/2]$. Sampling points are denoted by dots on the x -axis. The corresponding values of the integrand function are indicated by red dots on the red line. Matlab command: `quad(@(x) sin(1./x), 0.1, pi/2, 10^-9, 'trace')`.

11.5. AN EXAMPLE

Let us consider the computation of the following integral

$$\frac{5}{e^\pi - 2} \int_0^{\pi/2} e^{2x} \cos(x) = 1. \quad (11.7)$$

We begin by defining the function to be integrated using a so called *function handle*.

```
clc; clear all, close all
%define the function to be integrated
fun = @(x) exp(2*x).*cos(x)*5/(exp(pi)-2);
```

11.5.1. NEWTON-COTES INTEGRATION

Then we implement the Rectangle, Midpoint and Trapezium rule. In particular, the Trapezium rule can exploit the Matlab function trapz, that accepts two arguments: the grid and the corresponding function values and applies to each subinterval the trapezium rule².

```
i=1; % a counter
for j=2:11
    %number of points in each subinterval
    m(i)=2^j+1;
    %build the grid
    x=linspace(a,b,m(i));
    %grid size
    h=x(2)-x(1);
    %Rectangle Rule
    I_R(i)=sum(fun(x(1:end-1)))*h;
```

²Notice, that we can pass to the trapz function a not-evenly spaced grid

```

%Mid-Point Rule
x_mid=x+h/2;
I_MP(i)=sum(fun(x_mid(1:end-1)))*h;
%Trapezium rule
I_T(i)=trapz(x,fun(x));
i=i+1;
end
%We compare the results
out=[m', I_R', I_MP', I_T'];
T = table(m',I_R',I_MP',I_T','VariableNames',{'m' 'Rect' 'Mid' 'Trap'})

```

Table 11.5 shows the behaviour of approximated values $I_{n,m}$. Table 11.6 shows the behaviour of the errors

$$E_{n,m} = I(f) - I_{n,m}(f). \quad (11.8)$$

m	R	MP	T
5	0.97200380	1.03647822	0.92556504
9	1.00424101	1.00944240	0.98102163
17	1.00684171	1.00238106	0.99523202
33	1.00461138	1.00059655	0.99880654
65	1.00260397	1.00014922	0.99970154
129	1.00137659	1.00003731	0.99992538
257	1.00070695	1.00000933	0.99998134
513	1.00035814	1.00000233	0.99999534
1025	1.00018024	1.00000058	0.99999883
2049	1.00009041	1.00000015	0.99999971

Table 11.5: Numerical integration of the integral in 11.7.

Table 11.7 gives the ratios

$$R_{n,m} = |I(f) - I_{n,m}(f)| / |I(f) - I_{n,2m-1}(f)| \quad (11.9)$$

m	R	MP	T
5	0.0279962	-0.0364782	0.0744350
9	-0.0042410	-0.0094424	0.0189784
17	-0.0068417	-0.0023811	0.0047680
33	-0.0046114	-0.0005965	0.0011935
65	-0.0026040	-0.0001492	0.0002985
129	-0.0013766	-0.0000373	0.0000746
257	-0.0007070	-0.0000093	0.0000187
513	-0.0003581	-0.0000023	0.0000047
1025	-0.0001802	-0.0000006	0.0000012
2049	-0.0000904	-0.0000001	0.0000003

Table 11.6: Numerical integration error in the computation of the integral in 11.7 using different rules.

across alternative values for m . Following estimates of the error for Newton-Cotes formulas, see expressions (11.2.1) and (11.2.2), this ratio is expected to converge, for large values of m , to 4 for the Trapezium rule, and for the Midpoint rule and to 2 for the Rectangle Rule.

m	R	MP	T
9	6.601299	3.863235	3.922095
17	0.619876	3.965633	3.980377
33	1.483656	3.991398	3.995085
65	1.770908	3.997849	3.998771
129	1.891604	3.999462	3.999693
257	1.947225	3.999866	3.999923
513	1.973955	3.999966	3.999981
1025	1.987062	3.999992	3.999995
2049	1.993552	3.999998	3.999999

Table 11.7: Ratio $R_{n,m} = |I(f) - I_{n,m}(f)| / |I(f) - I_{n,2m-1}(f)|$ for different quadrature methods and different number of grid points.

11.6. GAUSSIAN QUADRATURE

- Gaussian quadrature rules select n abscissas and n weights to produce a rule of order $2n - 1$, so that

$$I_n(f) = \sum_{i=1}^n w_i f(x_i).$$

- This way we can exactly integrate polynomials of degree $2n - 1$, that is the highest degree for which a polynomial can be integrated using n points.
- Therefore, compared to the Newton-Cotes formulas, we may freely select the abscissas at which the integrand function is evaluated.

- Details on the construction of the abscissas and weights of the Gaussian quadrature rule, depending also on the integration domain, can be found in Chapter 6 of Fusai and Roncoroni (2008).
- We only observe that the error in the quadrature rule can be evaluated by the following estimate:

$$\int_a^b f(x) dx - \sum_{i=1}^n w_i f(x_i) = \frac{(b-a)^{2n+1} (n!)^4}{(2n+1) (2n!)^3} f^{(2n)}(\xi), \quad (11.10)$$

where $\xi \in (a, b)$, provided that $f^{(2n)}(\xi)$ is continuous on $[a, b]$.

- Experience shows that Gaussian quadrature results to be always better than the trapezoidal rule, except in the case of periodic integrands.
- We here provide an example of the Gauss-Legendre quadrature formula.
- Let us consider the integral in 11.7 and let use a so called Gauss-Legendre quadrature formula,
 - Using a 5 point Gauss-Legendre quadrature, the integral is computed as follows

Table 11.8: Gauss-Legendre integration with 5 points

x	w	$f(x)$	$w * f$
0.073686	0.186082	0.273321	0.05086
0.362485	0.375914	0.456585	0.171637
0.785398	0.446804	0.804496	0.359452
1.208311	0.375914	0.93997	0.353348
1.49711	0.186082	0.34771	0.064703
$I_{g,5}$		0.9999998	

- Using m points we have the results in Table 11.9.

Table 11.9: Numerical Integration of the function 11.7 usign a Gauss-Legendre rule

m	5	9	17	33	65
$I_{g,m}$	0.9999998	0.9999992	0.9999984	0.9999976	0.9999993

- To run the example download gauleg from Gauss Legendre Matlab Code.

```
%Define the function
fun=@(x) exp(2*x).*cos(x)*5/(exp(pi)-2);
%Assign a, b, m
a=0; b=pi/2; m=2;
%Generate abscissas and weights
[x, w] = gauleg(a, b, m); %download gauleg
%sample the function at x
f=fun(x);
%compute the integral
I=sum(f.*w);
```

11.6.1. ADAPTIVE QUADRATURE

Adaptive Quadrature can be implemented in Matlab using the built in functions as reported in Table 11.10

Syntax of adaptive quadrature routines
[I,n] = quad(fun,a,b)
[I,n] = quadl(fun,a,b,tol,trace)
[I,n] = quadgk(fun,a,b)

Table 11.10: Syntax of the MATLAB adaptive quadrature routines.

- Function `quad(fun,a,b)` computes the integral of function `fun` between `a` and `b` by using a recursive adaptive Simpson quadrature.
- Function `quadl(fun,a,b)` computes the integral of function `fun` between `a` and `b` by using a recursive adaptive Lobatto quadrature, i.e. a particular Gaussian quadrature rule that is accurate for all polynomials up to degree $2n^3$ ³,
- Function `quadgk` uses an adaptive Gaussian-Kronrod quadrature formula, in which the evaluation points are chosen so that an accurate approximation can be computed by re-using the information produced by the computation of a less accurate approximation. It is the most efficient for high accuracy and oscillatory integrands.

```
%Adaptive Simpson Rule
[I, evals]=quad(@(x) f(x),a,b,10^-7,'trace on')
%Adaptive Lobatto Rule
[I, evals]=quadl(@(x) f(x),a,b)
%Adaptive Gaussian-Kronrod Rule
[I, errbound]=quadgk(@(x) f(x),a,b)
```

Exercise 94 Let us consider a put option written on a non-dividend paying stock. The stock price is 98, the strike is 96, the risk free rate is 5%, the volatility of the GBM process is 30% and the option maturity is 2 years. Using the above quadrature rules price the option, assuming that the Black Scholes model holds.

³ `quadl(fun,a,b,tol)` allows the user to specify the tolerance level to be adopted in the stopping criterion, whereas `quadl(fun,a,b,tol,trace)` generates a table $[n| a| b - a| I]$, where n is the number of functional evaluations (first column), x_i is the abscissas at which the integrand is computed (second column), $x_{i+1} - x_i$ is the size of each subinterval (third column), and I is the value of the integral in each subinterval (fourth column). `quadl` may issue one of the following warnings: "Minimum step size reached" indicates that the recursive interval subdivision has produced a subinterval whose length is on the order of round-off error in the length of the original interval. A non-integrable singularity is possible. "Maximum function count exceeded" indicates that the integrand has been evaluated more than 10,000 times. A non-integrable singularity is likely. "Infinite or Not-a-Number function value encountered" indicates a floating point overflow or division by zero during the evaluation of the integrand in the interior of the interval.

CHAPTER 12. INVERSION OF THE CHARACTERISTIC FUNCTION AND THE COS METHOD

Several option pricing models generate analytical expressions for the characteristic function of the underlying variables. This allows us to price derivative contracts by using numerical approximations of the required probability distributions as obtained by Fourier inversion. In this section, we define the characteristic function of a r.v. and we show how it can be used for option pricing purposes. Then, we illustrate how quadrature methods can be used for the practical implementation of this important tool. In particular, we present the Fast Fourier Transform (FFT) algorithm and a very handy approach, i.e. the so-called COS method of Fang and Oosterlee (2008).

12.1. CHARACTERISTIC FUNCTION AND DENSITY FUNCTION

The characteristic function $\varphi(\gamma)$ of a random variable Z is defined as

$$\varphi(\gamma) = \mathbb{E}_t(e^{i\gamma Z}),$$

where $i := \sqrt{-1}$ is the imaginary unit and γ is a complex number. Observe that φ is a complex-valued function, except when Z has a symmetric distribution, in which case it is real-valued. If Z has density function $\mathbb{P}_t(Z \in dz)$, we can equivalently write

$$\varphi(\gamma) = \int_{-\infty}^{\infty} e^{i\gamma z} \mathbb{P}_t(Z \in dz).$$

It can be shown that if $\int_{\mathbb{R}} |\varphi(\gamma)| d\gamma < \infty$, then Z has continuous probability density function which can be recovered by the following Fourier inversion formula

$$\mathbb{P}_t(Z \in dz) = \frac{1}{\pi} \int_0^{+\infty} \operatorname{Re}(e^{-i\gamma z} \varphi(\gamma)) d\gamma, \quad (12.1)$$

Table 12.1: Elementary properties of the characteristic function.

(a)	$\varphi(0) = 1$.
(b)	$ \varphi(\gamma) \leq 1$.
(c)	If $Y = -Z$, then $\mathbb{E}_t(e^{i\gamma Y}) = \bar{\varphi}(\gamma)$, where $\bar{\varphi}(\gamma)$ is the complex conjugate of $\varphi(\gamma)$.
(d)	If $Y = a + bZ$, then $\mathbb{E}_t(e^{i\gamma Y}) = e^{i\gamma a} \varphi(b\gamma)$.
(e)	$\mathbb{E}_t(Z^n) = (i)^n \frac{\partial^n \varphi(\gamma)}{\partial \gamma^n} \Big _{\gamma=0}$, $n \in \mathbb{N}$.
(f)	If Z_1 and Z_2 are independent r.v.s with c.f. $\varphi_1(\gamma)$ and $\varphi_2(\gamma)$ then, $Y = Z_1 + Z_2$ has c.f. given by $\varphi_1(\gamma)\varphi_2(\gamma)$.

where $\text{Re}(\cdot)$ stands for the real part of its argument.¹ Expression (12.1) shows that the characteristic function identifies the distribution of Z . More generally, it can be shown that characteristic functions unequivocally identify distribution functions. Elementary properties of the characteristic function are given in Table 12.1.

Example If $Z \sim \mathcal{N}(0, 1)$, then

$$\varphi(\gamma) = e^{-\gamma^2/2}.$$

Moreover, if $Y = \mu + \sigma Z \sim \mathcal{N}(\mu, \sigma^2)$, using property (d) in Table 12.1, we get

$$\mathbb{E}_t(e^{i\gamma Y}) = e^{i\gamma\mu - \gamma^2\sigma^2/2}.$$

12.2. CHARACTERISTIC FUNCTION AND OPTION PRICING

In order to understand the importance of the characteristic function in option pricing, we start with a simple example. Recall that the arbitrage-free price of a contingent claim can be expressed as a conditional expectation of its discounted payoff. If we

¹Sometimes the inversion formula is given by $\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iz\gamma} \varphi(\gamma, T) d\gamma$. The two expressions coincide when the original function is a real function.

consider a standard call option, then the price reads as

$$\begin{aligned} c(K, T) &= \mathbb{E}_t^*(e^{-r(T-t)}(X(T) - K)^+) \\ &= e^{-r(T-t)} \{ \mathbb{E}_t^*((X(T)\mathbf{1}_{\{X(T)>K\}})) - K\mathbb{E}_t^*(\mathbf{1}_{\{X(T)>K\}}) \}. \end{aligned}$$

If we define $z(T) = \ln X(T)$ and $k = \ln K$, we can write

$$\begin{aligned} c(e^k, T) &= e^{-r(T-t)} \{ \mathbb{E}_t^*(e^{z(T)}\mathbf{1}_{\{z(T)>k\}}) - e^k \mathbb{E}_t^*(\mathbf{1}_{\{z(T)>k\}}) \} \\ &= X(t)\Pi_1 - e^k e^{-r(T-t)} \Pi_2. \end{aligned}$$

Quantities Π_1 and Π_2 can be interpreted as “stock-adjusted” and “money-market adjusted” probabilities, i.e., these probabilities have been computed by a change of measure using respectively the stock and the money market account as numéraires. Π_1 and Π_2 both represent the probability of ending up in-the-money at the option expiry. However, they are computed under martingale measures for two different numéraires: Π_1 uses as a numéraire the stock itself, whereas Π_2 uses the money market account, i.e.

$$\begin{aligned} \Pi_1 &= \Pr_t^X(X(T) > K), \\ \Pi_2 &= \Pr_t^B(X(T) > K), \end{aligned}$$

where the apex underlines that in the first case we are using the probability measure under which $e^{rT}/X(T)$ a martingale, whereas in the second case the ratio $X(T)/e^{rT}$ must be a martingale.

The two expectations above can be evaluated using the characteristic function of $Z(T)$. Let us define

$$\begin{aligned} \varphi_1(\gamma) &= \int_{-\infty}^{+\infty} e^{iz}\mathbb{P}_t^X(Z(T) \in dz), \\ \varphi_2(\gamma) &= \int_{-\infty}^{+\infty} e^{iz}\mathbb{P}_t^B(Z(T) \in dz), \end{aligned}$$

where \mathbb{P}_t^X and \mathbb{P}_t^B denotes the density function of $Z(T)$ under the two different numéraires (stock and money market account) and where we omit the dependence of the characteristic function on the option expiry T . The functions $\varphi_1(\gamma)$ and $\varphi_2(\gamma)$ are

the Fourier transforms of the probability density functions $\mathbb{P}_t^X(Z(T) \in dz)$ and $\mathbb{P}_t^B(Z(T) \in dz)$. It can be shown that the probabilities Π_1 and Π_2 can be recovered using the inversion formulas

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi} \int_0^{+\infty} \operatorname{Re} \left(\frac{e^{-i\gamma k} \varphi_1(\gamma)}{i\gamma} \right) d\gamma, \quad (12.2)$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^{+\infty} \operatorname{Re} \left(\frac{e^{-i\gamma k} \varphi_2(\gamma)}{i\gamma} \right) d\gamma, \quad (12.3)$$

and $\operatorname{Re}(\cdot)$ stands for the real part of its argument.

Integrals (12.2) and (12.3) can be computed by quadrature. To this end, some care is necessary if we are to tackle instability issues linked to the oscillatory nature of the integrand function, due to the presence of the complex exponential function. Recalling our previous discussion, the trapezoidal rule is expected to perform much better than other Newton–Cotes rules and should be comparable to Gaussian quadrature. However a more efficient approach has been proposed in Carr and Madan (1998). These authors, instead of computing Π_1 and Π_2 separately, calculate the Fourier transform of an adjusted call option price with respect to the logarithmic strike price k . These authors introduce a damping parameter $\alpha > 0$ and define the quantity

$$c_\alpha(e^k, T) := e^{\alpha k} c(e^k, T), \quad (12.4)$$

where α has to be chosen so that $c_\alpha(e^k, T)$ is square-integrable and therefore admits the Fourier transform

$$\mathcal{F}[c_\alpha](\gamma) = \int_{-\infty}^{+\infty} e^{i\gamma k} c_\alpha(e^k, T) dk.$$

It can be shown that this quantity can be expressed in terms of the characteristic function $\varphi_2(\gamma)$ by

$$\mathcal{F}[c_\alpha](\gamma) = \frac{e^{-rT} \varphi_2(\gamma - \alpha i - i)}{\alpha^2 + \alpha - \gamma^2 + i(2\alpha + 1)\gamma}. \quad (12.5)$$

If the characteristic function of $z(T)$ is known in closed form, we also have an analytical expression for $\mathcal{F}[c_\alpha](\gamma)$ at our disposal. Similarly, for a put option we define

$$p_\alpha(e^k, T) := e^{-\alpha k} p(e^k, T).$$

Its Fourier transform $\mathcal{F}[p_\alpha](\gamma) = \int_{-\infty}^{+\infty} e^{i\gamma k} e^{-\alpha k} p(e^k, T) dk$ can be written as

$$\mathcal{F}[p_\alpha](\gamma) = \frac{e^{-rT} \varphi_2(\gamma + \alpha i - i)}{\alpha^2 - \alpha - \gamma^2 + i(-2\alpha + 1)\gamma}. \quad (12.6)$$

As a last step, Fourier inversion yields the option prices

$$c(e^k, T) = \frac{e^{-\alpha k}}{\pi} \int_0^{+\infty} e^{-i\gamma k} \mathcal{F}[c_\alpha](\gamma) d\gamma, \quad (12.7)$$

$$p(e^k, T) = \frac{e^{\alpha k}}{\pi} \int_0^{+\infty} e^{-i\gamma k} \mathcal{F}[p_\alpha](\gamma) d\gamma. \quad (12.8)$$

It is possible to prove that this method is viable provided that the moment of order $1 + \alpha$ exists and is finite for some $\alpha > 0$, i.e.

$$\mathbb{E}_t^*(X(T)^{1+\alpha}) < \infty.$$

It turns out that any $\alpha \in [1.5, 2]$ works quite well for most cases.

Computing (12.1) and (12.7) (or (12.8)) can be done by using either the trapezoidal rule or the Gauss–Legendre quadrature. The trapezoidal rule can be implemented very efficiently by means of the Fast Fourier Transform (FFT) algorithm. We end this section by illustrating this procedure. Let us consider the problem of computing the $N \times 1$ vector $\mathbf{H} = \{H_0, \dots, H_n, \dots, H_{N-1}\}$ given the $N \times 1$ vector $\mathbf{h} = \{h_0, \dots, h_k, \dots, h_{N-1}\}$, such that

$$H_n = \sum_{j=0}^{N-1} e^{+ijn\frac{2\pi}{N}} h_j, \quad n = 0, \dots, N-1. \quad (12.9)$$

Let $\mathbf{H} = \{H_0, \dots, H_n, \dots, H_{N-1}\}$ be the discrete Fourier transform of the vector $\mathbf{h} = \{h_0, \dots, h_k, \dots, h_{N-1}\}$. Vice versa, using the discrete inverse Fourier transform we can also recover \mathbf{h} from \mathbf{H} as

$$h_k = \frac{1}{N} \sum_{j=0}^{N-1} e^{-ijk\frac{2\pi}{N}} H_j, \quad k = 0, \dots, N-1. \quad (12.10)$$

The only difference between (12.9) and (12.10) is represented by the change of sign in the exponential and the multiplicative constant $1/N$.

In general, if we try to compute \mathbf{H} from \mathbf{h} , or vice versa, we need N^2 multiplications involving complex quantities, plus additional $N(N - 1)$ complex sums. Exploiting the fact that these computations are not independent of each other, in the 1960s Cooley and Tukey discovered an algorithm requiring only $N \ln_2(N)/2$ operations. This algorithm is known as the Fast Fourier Transform and its discovery has greatly stimulated the use of the Fourier transform in several technical disciplines. An important aspect of FFT is that the algorithm is based on a recursive procedure that allows one to express the FFT of length N as the sum of two FFT (each one of length $N/2$). This fact implies that the best choice for N is a power of 2.

In order to exploit the FFT algorithm, we discretize the inversion integral (12.7) using the trapezoidal rule² with step η

$$c(e^k, T) \simeq \frac{e^{-\alpha k}}{\pi} \sum_{j=0}^{N-1} e^{-i\eta j k} \mathcal{F}[c_\alpha](j\eta) w_j \eta. \quad (12.11)$$

Here $w_1 = 1/2, w_2 = 1, \dots, w_{N-2} = 1, w_{N-1} = 1/2$. This quadrature introduces two types of error: first, a truncation error due to the finiteness of the upper limit in the numerical integration; second, a sampling error due to the evaluation of the integrand at grid points only. The FFT returns the option value over a grid of N evenly spaced logarithmic strike prices k_0, \dots, k_{N-1} , with $k_n = k_0 + n\lambda, k_0 = -N\lambda/2$, and $k_{N-1} = N\lambda/2$. By setting $k_n = k_0 + n\lambda$ into expression (12.11), we obtain

$$\begin{aligned} c(e^{k_n}, T) &\simeq \frac{e^{-\alpha k_n}}{\pi} \sum_{j=0}^{N-1} e^{-i\eta j(k_0 + n\lambda)} \mathcal{F}[c_\alpha](j\eta) \eta w_j \\ &= \frac{e^{-\alpha k_n}}{\pi} \sum_{j=0}^{N-1} e^{-i\eta j n \lambda} e^{-i\eta j k_0} \mathcal{F}[c_\alpha](j\eta) \eta w_j, \end{aligned}$$

for $n = 0, \dots, N - 1$. Finally, if we set

$$\lambda\eta \equiv \frac{2\pi}{N},$$

we can apply the FFT algorithm (12.10). The choice $\lambda\eta \equiv 2\pi/N$ highlights a trade-off arising between the accuracy of the integral, which is determined by the sampling rate η of the Fourier transform, and the degree of space refinement as represented

²The trapezoidal rule can be applied to (12.1) or (12.8) as well.

Table 12.2: Pseudo-code for option pricing via the FFT algorithm.

-
- (1) Assign $N, \lambda, \alpha, \varphi(\gamma)$;
 - (2) Construct the Fourier transform of dampened option price $F[c_\alpha](\gamma)$ in (12.7);
 - (3) Construct vector h with components (12.13);
 - (4) Apply the FFT algorithm and obtain a vector H with N components;
 - (5) Multiply the n th component of H by $\exp(-\alpha k_n)N/\pi$, where $k_n = -N\lambda/2 + n\lambda$.
-

by λ . A finer discretization of the strike price space comes with a rougher discretization step η in the transform plane (and vice versa). Unfortunately, numerical experiments are required to determine the best compromise between the two quantities involved in this trade-off. If we choose $\lambda\eta = 2\pi/N$, we then have $\eta k_0 = -N\lambda\eta/2 = -\pi$ and thus:

$$c(e^{k_n}, T) \simeq \frac{e^{-\alpha k_n}}{\pi} \sum_{j=0}^{N-1} e^{-ijn\frac{2\pi}{N}} h_j, \quad (12.12)$$

where

$$h_j = e^{ij\pi} \mathcal{F}[c_\alpha](j\eta) \eta w_j. \quad (12.13)$$

Given the FFT algorithm, option prices can be computed according to the procedure described in Table 12.2.

12.2.1. FFT METHOD IN MATLAB® ALGORITHMS

Matlab® includes built-in routines `fft(x)` and `ifft(x)` which implement discrete Fourier and inverse transforms. The Matlab® FFT code is based on FFTW (The Fastest Fourier Transform in the West) developed at MIT and available from <http://www.fftw.org>. The `fft(x)` Matlab® function operates the following sum

$$X(k) = \sum_{j=1}^N x(j) e^{-i\frac{2\pi}{N}(k-1)(j-1)}$$

and therefore if we need to compute (12.12), we need to construct a vector x having as element at position j exactly the quantity h_j given in (12.13).

Here, we illustrate the Matlab® implementation in the Heston model. The risk neutral characteristic function is given in formula 5.13

Below, we use subfunctions to place all the functions required by the numerical inversion in a single M-file that is named `FFT_Pricing_CallPut.m`. This file can be run from the Matlab® command window. The function returns three vectors containing strikes and corresponding option prices (call and put). The arguments correspond to the spot price (S), the risk-free rate (rf), the volatility (sg), the time to maturity (t), the damping parameter (α), the exponent for determining the power of 2 ($npower$). In particular, if α is positive (negative) we are pricing a call (put) option. Standard values for $npower$ are 12 or 13, whilst the absolute value of α can be 1.5.

```
function [K, callprice, putprice]= FFT_Pricing_CallPut(S,rf,sg,t,alpha,npower)
%find the number of points
N = 2^npower;
%spacing in the strike grid
lambda = 0.01;
% parameters for the cf grid (0,N*eta)
eta = 2*pi/(N*lambda);
% create grid for the cf
g = (0:N-1) * eta;
%create grid for the strike
k0 = -N*lambda/2;
k = k0 + (0:N-1) * lambda;
%compute the ch. function
fc = ft_option(rf,sg,t,alpha,g);
w = [0.5 ones(1,N-2) 0.5]; % trapezoidal rule
h2 = exp(-li*k0*g) .* fc .* w*eta;
%Fourier inversion
g = fft(h2);
g2 = real( exp(-alpha.*k) / pi .* g);
K = S * exp(k); K = K';
Y= exp(-rf*t) * S * g2; callprice= Y';
putprice=callprice-(S-K*exp(-rf*t));
```

Example

The following Matlab[®] commands allow us to price call and put options in the Black-Scholes model via numerical inversion of the Carr and Madan (1999) formula. Figure 12.1 shows that, for the given parameter set, in the range of strikes ($S_0 \times 0.4$, $S_0/0.4$) the pricing error (Black-Scholes price-FFT inversion) is of order 10^{-15} .

```
clear all
close all
clc
%model parameters
S0=1;rf=0.05;sg=0.2;ttm=1;
%Fourier inversion parameters
alpha=1.5; npower=12;
[K, callprice,putprice]= FFT_Pricing_CallPut(S0,rf,sg,ttm,1.5,12);
[cexact pexact]=blsprice(S0,K,rf,ttm,sg,0);
h=figure('Color',[1 1 1])
subplot(2,1,1)
plot(K,[callprice cexact]);
axis([S0*0.4 S0/0.4 0 S0])
xlabel('Strike')
ylabel('Call option price')
legend('FFT inversion','Exact')
subplot(2,1,2)
plot(K,[cexact-callprice]);
xlim([S0*0.4 S0/0.4])
xlabel('Strike')
ylabel('Pricing Error')
title('Pricing Error (call option)')

h=figure('Color',[1 1 1])
subplot(2,1,1)
plot(K,[putprice pexact]);
axis([S0*0.4 S0/0.4 0 S0])
xlabel('Strike')
ylabel('Put option price')
legend('FFT inversion','Exact')
```

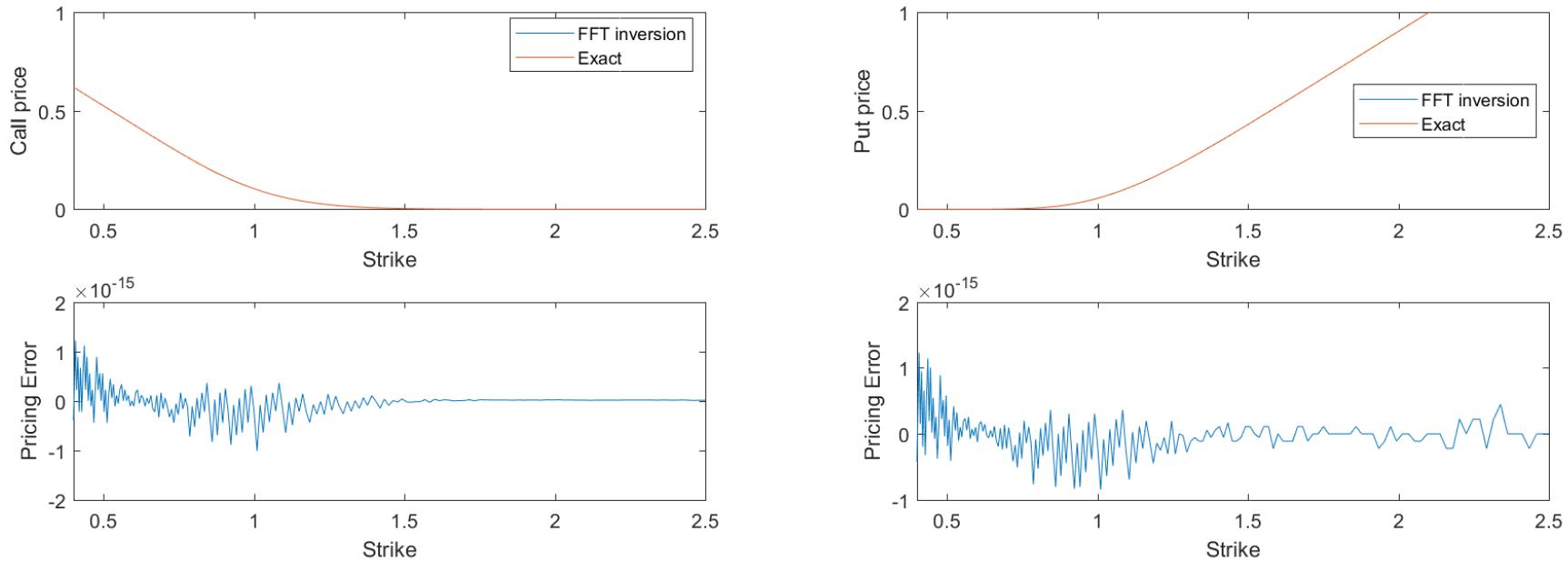


Figure 12.1: Black–Scholes price–FFT price. Parameters: $S(0) = 1, r = 0.05, \sigma = 0.2, T = 1$ year, $\alpha = 1.5, N = 2^{12}$. **Left panel:** call prices. **Right panel:** put prices.

```

subplot(2,1,2)
plot(K,[pexact-putprice]);
xlim([S0*0.4 S0/0.4])
xlabel('Strike')
ylabel('Pricing Error')
title('Pricing Error (put option)')

```

12.2.2. INVERSION OF THE CHARACTERISTIC FUNCTION AND THE COS METHOD

Alternative methods for the inversion of the characteristic function have been developed over the last few years. In particular, we examine the so called COS method proposed by Fang and Oosterlee (2008).

The COS method is based on the Fourier-Cosine expansion of a function $f : [a, b] \rightarrow \mathbb{R}$

$$f(y) = \frac{A_0}{2} + \sum_{k=1}^{\infty} A_k \cos\left(k\pi \frac{y-a}{b-a}\right) \quad (12.14)$$

$$A_k = \frac{2}{b-a} \int_a^b f(y) \cos\left(k\pi \frac{y-a}{b-a}\right) dy. \quad (12.15)$$

As equation (12.15) still contains the unknown function $f(y)$, we observe the following link between f and its Fourier transform (via Euler decomposition)

$$e^{-iv} \varphi(u) = \int_{\mathbb{R}} e^{i(uy-v)} f(y) dy = \int_{\mathbb{R}} (\cos(uy-v) + i \sin(uy-v)) f(y) dy. \quad (12.16)$$

Hence, let

$$\varphi_1(u) = \int_a^b e^{iuy} f(y) dy,$$

be a truncated version of the Fourier transform of f . Then

$$\begin{aligned} A_k &= \frac{2}{b-a} \int_a^b f(y) \cos\left(k\pi \frac{y-a}{b-a}\right) dy \\ &= \frac{2}{b-a} \Re\left(e^{k\pi \frac{-a}{b-a}} \varphi_1\left(\frac{k\pi}{b-a}\right)\right) \\ &\approx \frac{2}{b-a} \Re\left(e^{k\pi \frac{-a}{b-a}} \varphi\left(\frac{k\pi}{b-a}\right)\right) \\ &= F_k; \end{aligned} \quad (12.17)$$

this implies that the Fourier Cosine series of f can be approximated by

$$f(y) \approx f_1(y) = \frac{F_0}{2} + \sum_{k=1}^{\infty} F_k \cos\left(k\pi \frac{y-a}{b-a}\right). \quad (12.18)$$

To the purpose of actual computation, we will need to truncate the infinite sum to the first $N > 0$ terms, so that

$$f(y) \approx f_1(y) \approx f_2(y) = \frac{F_0}{2} + \sum_{k=1}^{N-2} F_k \cos\left(k\pi \frac{y-a}{b-a}\right). \quad (12.19)$$

If f is the density function of a random variable (or a random process), the cumulative distribution function follows by integration

$$F(x) = \int_{-\infty}^x f(y) dy$$

with $f(y)$ given by equation (12.19). By exchanging the order of integration and summation, in fact, it is possible to write

$$F(x) = \sum_{k=0}^{N-1}' F_k \int_{-\infty}^x \cos\left(k\pi \frac{y-a}{b-a}\right) dy$$

where \sum' indicates that the first term is weighted by 0.5. Note that the inner integral can be solved via usual trigonometric representation, and therefore

$$\begin{aligned} F(x) &= \sum_{k=0}^N' F_k \psi_k(a, x) \\ \psi_k(c, d) &= \int_c^d \cos\left(k\pi \frac{y-a}{b-a}\right) dy \\ &= \begin{cases} \frac{b-a}{k\pi} \left(\sin\left(k\pi \frac{d-a}{b-a}\right) - \sin\left(k\pi \frac{c-a}{b-a}\right) \right) & k \neq 0 \\ d - c & k = 0. \end{cases} \end{aligned} \quad (12.20)$$

Remark 95 In order to compute $F(x) = \mathbb{P}(X \leq x)$, then set in (12.20) $c = a, d = x$; if instead $\mathbb{P}(x_1 \leq X \leq x_2)$, set in (12.20) $c = x_1, d = x_2$.

The accuracy of the overall approximation depends on the approximation error originated by using the original characteristic function instead of its truncation to a subinterval $[a, b]$, and the truncation error originated by considering only the first $N > 0$ terms of the summation. In details.

- The first error originates from using equation (12.18) instead of (12.14). As shown in Fang and Oosterlee (2008), this error can be controlled by choosing the truncation range as

$$[a, b] = \left[c_1 - L\sqrt{c_2 + \sqrt{c_4}}, c_1 + L\sqrt{c_2 + \sqrt{c_4}} \right],$$

where c_n denotes the n^{th} -cumulant of the relevant distribution, and L is set to a suitably chosen value which allows us to gauge the error. Fang and Oosterlee (2008) recommend $L = 10$.

- The second error originates from using equation (12.19) instead of (12.18), and it is controlled by gauging the parameter N .

As a test of the goodness of the approximation, we consider the standard Normal distribution, underpinning the Brownian motion. The characteristic function is

$$\varphi(u) = e^{-\frac{u^2}{2}};$$

the error produced by the COS method in reproducing the density function and the cumulative distribution function versus the built-in functions in Matlab are reported in Figures 12.2 and 12.3. We observe the following.

- Figure 12.3 shows that the approximation of the Normal distribution is very accurate for a small L .
- This is mainly due to the fact that the Normal distribution has zero skewness and excess kurtosis.
- A larger value of L is required to cater for these additional features.
- To illustrate the point, we repeat the test for the compound Poisson distribution underpinning the Merton jump diffusion model (MJD) introduced in Section 6.2.1.

- Results are reported in Figures 12.4 and 12.5. The details of the errors are reported in Table 12.3.
- The results support the recommended value $L = 10$.

Here the Matlab code for obtaining the PDF and the CDF given of a random variable given the characteristic function, the mean and the standard deviation of the random variable. N and L are the parameters for the COS method.

```

function [COSpdf, COScdf, x] = get_distrICOS(type,N,L,x0)
%COS Method for inverting the characteristic function
%INPUT
%type: handle to the chosen distribution
%N: number of grid points for sampling the characteristic function
%L: fix the size of the x-grid (-L,L)
%x0: parameters of the r.v. X
%OUTPUT
%The PDF and the CDF and the corresponding x-grid

[c1, c2, ~, c4]=get_cumulants(x0,type); % it obtaines the cumulants of the chosen distribution indicated by type

a=c1-L*sqrt(c2+sqrt(c4)); %Fang and Oosterlee (2008)
b=c1+L*sqrt(c2+sqrt(c4)); %Fang and Oosterlee (2008)
c=a;

xmin=a; xmax=b;

CosCharFn=CharFunc(x0,type,(0:N-1)*pi/(b-a)); % returns the corresponding CF

x=linspace(xmin,xmax,N);
for j=1:length(x)
V = (2/(b-a))*cos((x(j)-a)*(0:N-1)*pi/(b-a));
COSpdf(j)=real(sum(CosCharFn.*V.*exp(1i*(0:N-1)*pi*(-a)/(b-a)))-0.5*CosCharFn(1)*1*V(1));
end

for jj=1:length(x)

```

```

d=x(jj);
psi(1) = d-c;
psi(2:N) = (sin((1:N-1)*pi*(d-a)/(b-a))-sin((1:N-1)*pi*(c-a)/(b-a)))*(b-a)./((1:N-1)*pi);
V = (2/(b-a))*psi;
COScdf(jj)=real(sum(CosCharFn.*V.*exp(1i*(0:N-1)*pi*(-a)/(b-a)))-0.5*CosCharFn(1)*1*V(1));
end
end

```

Example 96 Here we apply the above routine to obtain the probability density function and the cumulative density function of a Normal random variable with mean 0 and variance 1 and we compare our results with the ones obtained from the built-in Matlab function `normpdf`, `normcdf`. We test the accuracy for different value of N by keeping L fixed. This generates Figure 12.2.

```

clear all; close all; clc
nn=4:2:10;%values for N
L=10; % Fang and Oosterlee (2008)

% Define the characteristic function
% We use the Gaussian cf, but it can be replaced by others cf.
% The selection is made by choosing the indicator 'type' according to the
% following
% type = 1 - Gaussian distribution
% type = 2 - Merton Jump Diffusion distribution
% type = 3 - any other distribution of interest (VG for example)

type =1;
meanX=0; sdX=1; x0=[meanX sdX];

%Apply the COS method for different values of N
h=figure('Color',[1 1 1]);

for j=1:length(nn)

```

```
%Number of points
N=2^nn(j);

%get the pdf and the cdf with the COS method
[COSpdf, COScdf, x] = get_distrICOS(type,N,L,x0);

%get the pdf and the cdf from known expressions
Exactpdf= normpdf(x,meanX,sdX);
Exactcdf=normcdf(x,meanX,sdX);

%error in computing the pdf varying N
subplot(2,4,j)
plot(x,COSpdf-Exactpdf)
title(strcat('N = ',num2str(2^nn(j))))
ylabel('COS pdf - Exact pdf')

%error in computing the cdf varying N
subplot(2,4,4+j)
plot(x,COScdf-Exactcdf)
title(strcat('N = ',num2str(2^nn(j))))
ylabel('COS cdf - Exact cdf')
end
print(h, '-dpng', 'figCOS_Gauss')
```

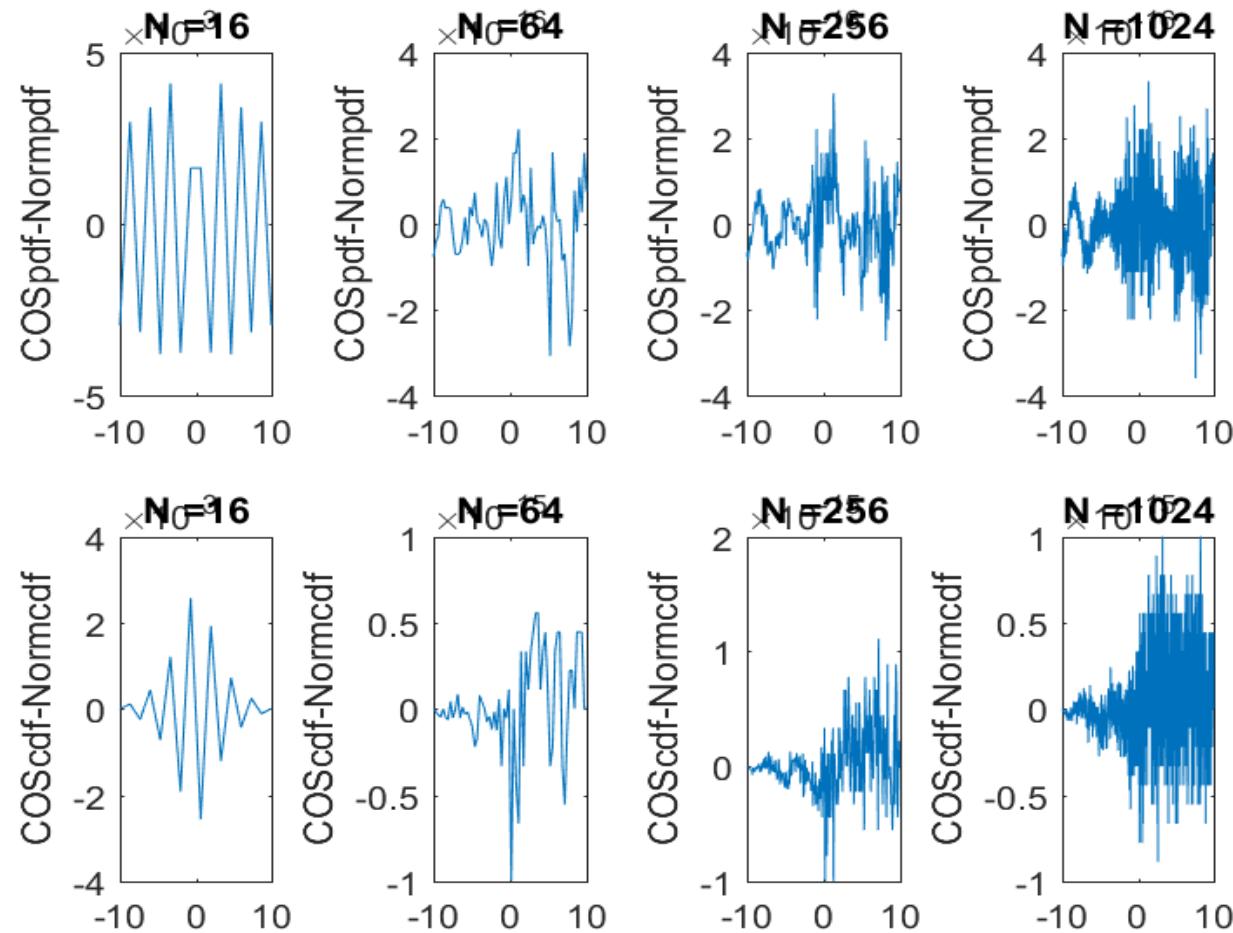


Figure 12.2: COS Method: error in approximating the standard Normal distribution by changing N . $L = 10$.

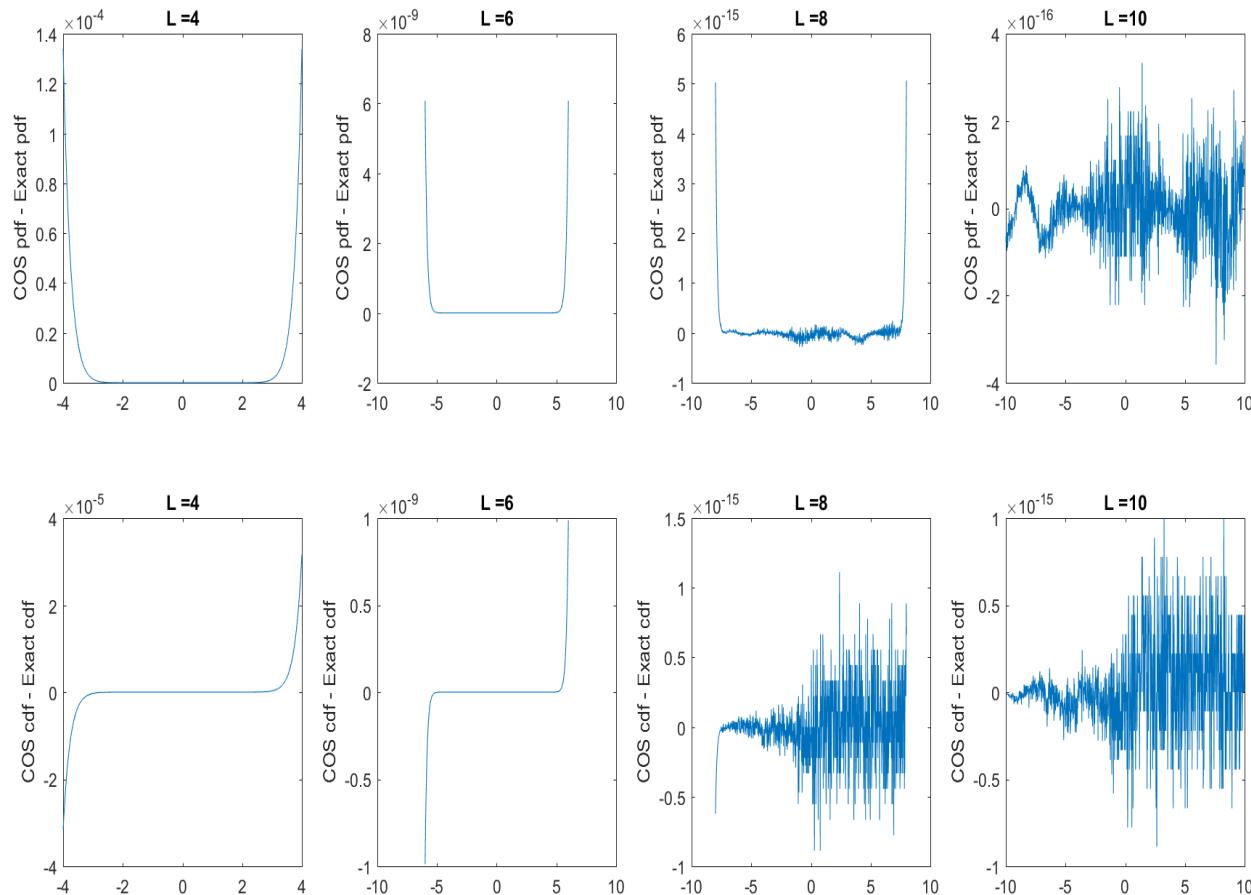


Figure 12.3: COS Method: error in approximating the standard Normal distribution by changing L . $N = 2^{10}$.

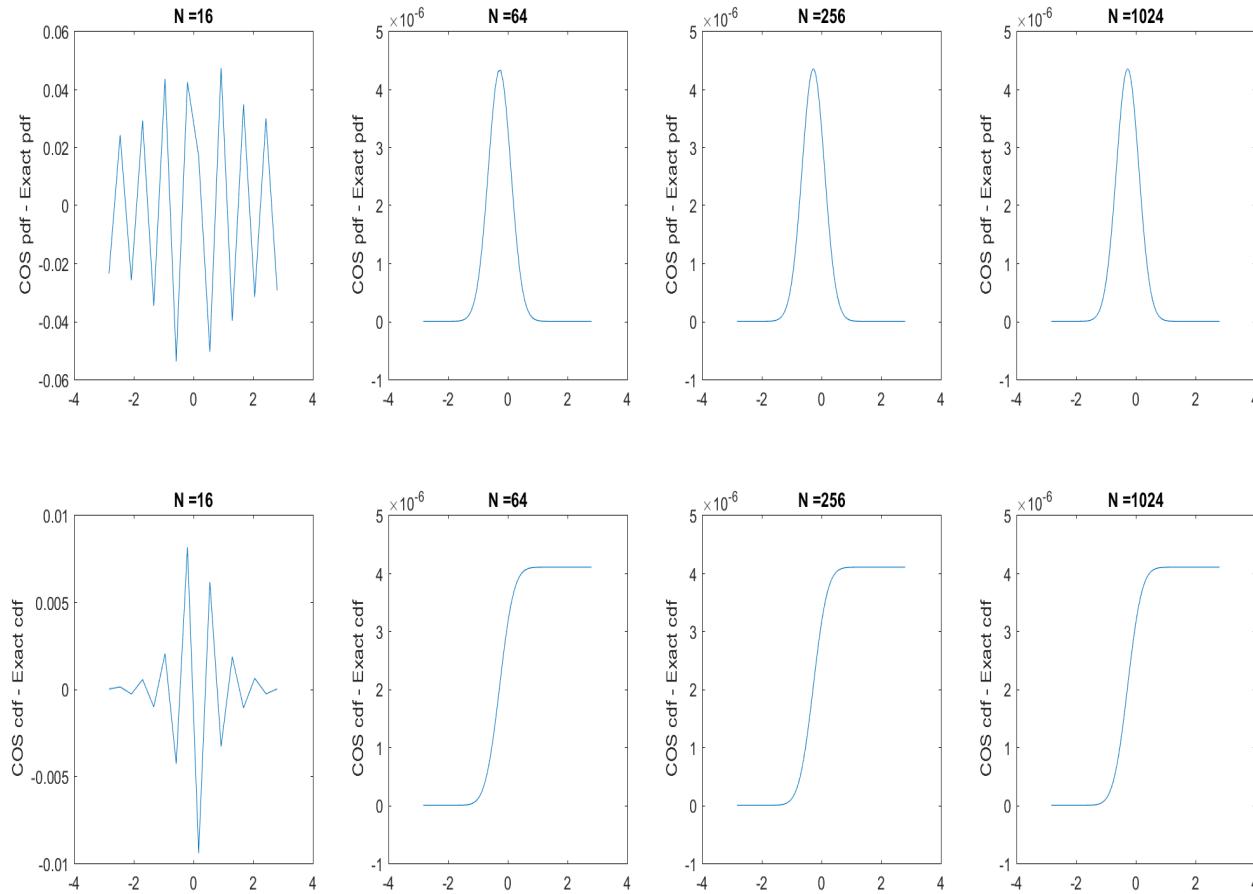


Figure 12.4: COS Method: error in function of N in approximating the Merton Jump Diffusion distribution by changing N . $L = 10$. Parameter set: $\lambda = 1.5$; $\mu = 0.03$; $\sigma = 0.2$; $\sigma_Z = 0.1$; $\mu_Z = -0.03$.

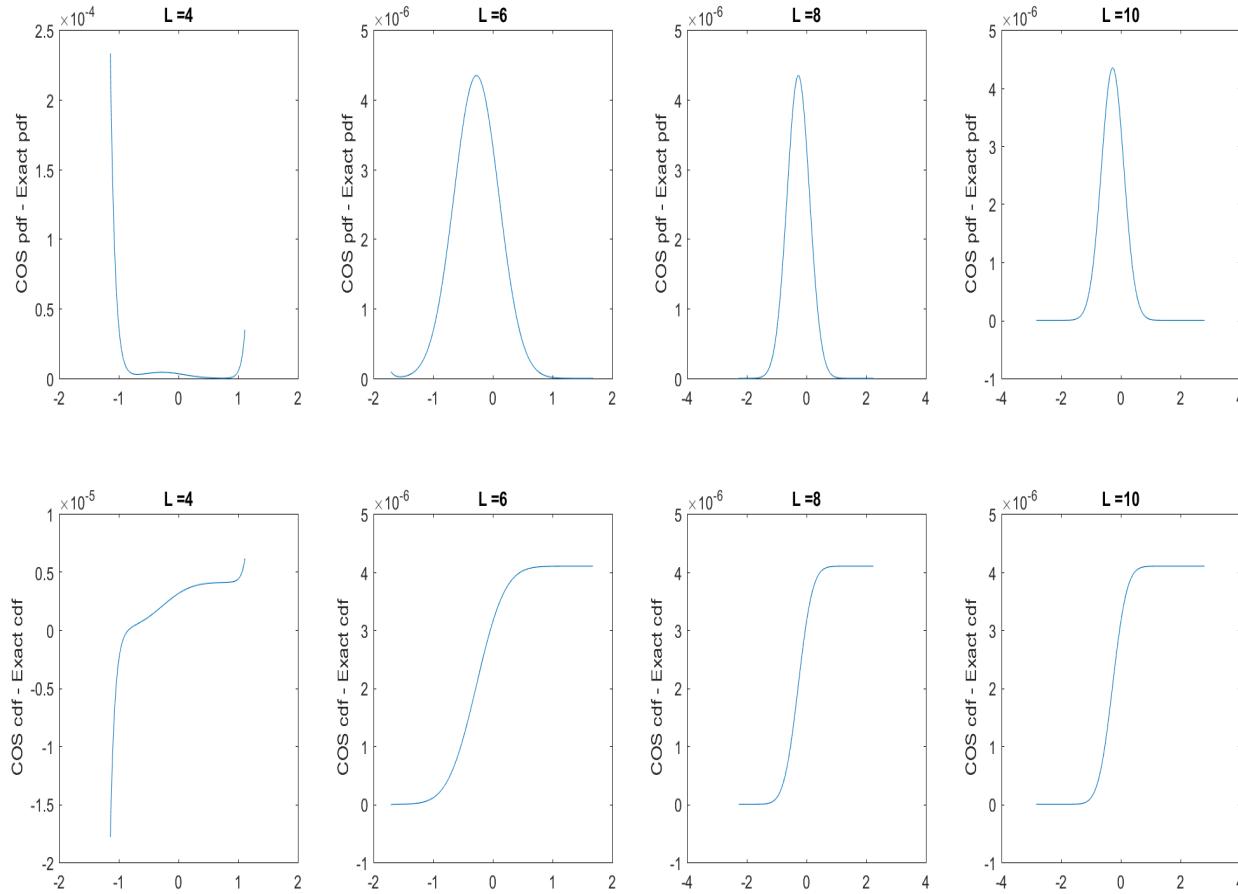


Figure 12.5: COS Method: error in function of N in approximating the Merton Jump Diffusion distribution by changing L . $N = 2^{10}$. Parameter set: $\lambda = 1.5; \mu = 0.03; \sigma = 0.2; \sigma_Z = 0.1; \mu_Z = -0.03$.

L	PDF		CDF	
	$\min(\ COS - EXACT\)$	$\max(\ COS - EXACT\)$	$\min(\ COS - EXACT\)$	$\max(\ COS - EXACT\)$
4	1.59522544524587e-07	0.000233079370812578	5.81205628264782e-09	1.78300160807126e-05
6	9.71507197011898e-11	4.34428993378866e-06	1.66243511491868e-11	4.09755692465375e-06
8	5.47078025186765e-15	4.34426872664151e-06	1.01771141512258e-14	4.09750097718487e-06
10	1.23105315336264e-17	4.34418772576883e-06	2.63121390644560e-18	4.09750097796202e-06

Table 12.3: Column 1: L; Column 2: minimum absolute error in the PDF; Column 3: maximum absolute error in the PDF; Column 4: minimum absolute error in the PDF; Column 5: maximum absolute error in the PDF. Parameter set: $\lambda = 1.5$; $\mu = 0.03$; $\sigma = 0.2$; $\sigma_Z = 0.1$; $\mu_Z = -0.03$; $N = 2^{10}$.

Fang and Osterlee (2008) also show how to use the COS method for option pricing purposes. Let us consider the case of the call option on the stock S , with strike K and maturity T . Set $y = \ln \frac{S_T}{K}$ so that you can re-write the option payoff as $(S_T - K)^+ = K(e^y - 1)^+$. Then, the price of the option is

$$C_0 = \mathbb{E} \left(e^{-rT} (S_T - K)^+ \right) = e^{-rT} K \int_{\mathbb{R}} (e^y - 1)^+ f(y) dy.$$

Using the COS method for the computation of f implies the following.

- By direct integration

$$\begin{aligned} \int_c^d e^y \cos \left(k\pi \frac{y-a}{b-a} \right) dy &= \chi_k(c, d) \\ \chi_k(c, d) &= \left[\frac{e^y}{1 + \left(\frac{k\pi}{b-a} \right)^2} \left(\cos \left(k\pi \frac{y-a}{b-a} \right) + \frac{k\pi}{b-a} \sin \left(k\pi \frac{y-a}{b-a} \right) \right) \right]_{y=c}^{y=d} \end{aligned}$$

- Using the previous results the price of the European call option can be written as

$$C_0 = e^{-rT} K \sum_{k=0}^{N-1} {}'V_k F_k, \quad (12.21)$$

N	Black Scholes - COS	
	Call	Put
16	0.140414254535017	-0.0194277629304489
64	-7.99360577730113e-14	-2.75335310107039e-14
256	-7.99360577730113e-14	-2.75335310107039e-14
1024	-7.99360577730113e-14	-2.75335310107039e-14

Table 12.4: COS option pricing vs Black-Scholes for European vanilla calls and puts. Parameter set: $S = 100$, $K = 95$, $r = 5\%$ p.a., $q = 2\%$ p.a., $T = 1$ year; $\sigma = 0.2$. $L = 10$.

with

$$V_k = \chi_k(0, b) - \psi_k(0, b),$$

and F_k, ψ_k as in equations (12.17) and (12.20).

The accuracy of the COS method applied to vanilla European options in the Black-Scholes model is shown in Table 12.4 where errors are computed in correspondence of different values of the truncation parameter N .

Remark 97 When pricing with Lévy processes, the stock price dynamics has always form

$$S_T = S_0 e^{(r-q-\omega)T + X_T}$$

where X_t is the driving process of choice, $\omega = \ln \phi_X(-i; t) / t$. This adjustment ω guarantees the martingale condition for discounted stock prices. But it also implies that the driving Lévy process has form either

$$X_t = \sigma W_t + \sum_{k=1}^{N_t} Z_k$$

in case of a JD process, or

$$X_t = \theta G_t + \sigma W_{G_t}$$

in case of a subordinated Brownian motion. In other words, the 'drift' parameter is not considered and set to zero. This explains why in the code for the routine executing option pricing using the COS method, the set of parameters which is in-putted for the calculation of the cumulants and the characteristic function is $[0 \ xxx]$, and xxx denotes the vector of parameters of X_t .

For example in the case of the Black-Scholes model, the driver is $X_t = \sigma W_t$, hence there is no drift component. This is helpful if the routine calculating characteristic functions and cumulants has been developed for the more general case of a process of the form $\mu t + X_t$.

The routine is the following.

```

function [Call_COS, Put_COS]=get_VanillaEuro_COS(S,K,r,q,T,xxx,type,N,L)

% type == 1 The Black Scholes model
% type == 2 The VG Model
% type == 3 any other model of choice

%Bounds
[c1, c2, ~, c4]=LProc_cumulants([0 xxx],type,T);
a= c1-L*sqrt(c2+sqrt(c4));
b= c1+L*sqrt(c2+sqrt(c4));

k_in_vect = (0:N-1);
u= k_in_vect *pi/(b-a);

% create char fun and risk neutral adjustment
phi=CharFunc([0 xxx],type,u,T); % calls for the CF of the driftless process
omega=log(CharFunc([0 xxx],type,-li,1)); % computes the RN adjustment
mu=r-q-omega; % computes the risk neutral drift
CF_RN=exp(li * u * log(S/K)).* exp(li*u.*mu*T).*phi;

%—— F_k computation——
Exp_term=exp(-li*pi*a*k_in_vect/(b-a));
Fk_x = real(CF_RN .* Exp_term);

%——V_k computation for Call and Put ——————

```

```

Vk_call = (2 / (b-a)) * (chiFO(0, b, N, a, b) - psiFO(0, b, N, a, b)); % call
Vk_put = -(2 / (b-a)) * (chiFO(a, 0, N, a, b) - psiFO(a, 0, N, a, b)); % put

%——Option price computation———
DF = exp(-r * T);
Call_COS = DF*K*(sum(Fk_x.*Vk_call)-0.5*Exp_term(1)*CF_RN(1)*Vk_call(1));
Put_COS = DF*K*(sum(Fk_x.*Vk_put)-0.5*Exp_term(1)*CF_RN(1)*Vk_put(1));

% Fourier-cosine series coefficients of the terminal payoff function
function ret=chiFO(c,d,N,a,b)
chi = (1./(1+((0:N-1)*pi/(b-a)).^2)).*(cos((0:N-1)*pi*(d-a)/(b-a))*exp(d)-...
cos((0:N-1)*pi*(c-a)/(b-a))*exp(c)+((0:N-1)*pi/(b-a)).*sin((0:N-1)*pi*(d-a)/(b-a))*exp(d)-...
((0:N-1)*pi/(b-a)).*sin((0:N-1)*pi*(c-a)/(b-a))*exp(c));
ret=chi;

function ret=psiFO(c,d,N,a,b)
psi(1) = d-c;
psi(2:N) = (sin((1:N-1)*pi*(d-a)/(b-a))-sin((1:N-1)*pi*(c-a)/(b-a)))*(b-a)./((1:N-1)*pi);
ret=psi;

```

Example 98 The results in Table 12.4 are obtained using the following code.

```

% Testing the accuracy of the COS method versus Black Scholes
% CALLS get_VanillaEuro_COS(S,K,r,q,T,x0,type,N,L)
%
% INPUT
% S: spot
% K: strike
% r: risk free rate (cont. compounded)
% q: dividend yield
% T: maturity

```

```
% x0: parameter set
% type: handle to choose the model process (BM, VG...)
% N, L: COS parameters
%%%%%%%%%%%%%%%
clear all
clc
close all

type=1; % Choose the Black-Scholes model

nn=[4 6 8 10];
L=10;

r = 0.05;
S = 100;
q = 0.02;
K = 95;
T = 1;

sigma=0.25;

% Black Scholes prices
[BS_Call, BS_Put]=blsprice(S,K,r,T,sigma,q);

% Cos option prices
x0=sigma;

for j=1:length(nn)
    N=2^nn(j);
    [COS_Call(j), COS_Put(j)]=get_VanillaEuro_COS(S,K,r,q,T,x0,type,N,L);
    error_call(j)=BS_Call-COS_Call(j);
    error_put(j)=BS_Put-COS_Put(j);
end
```

```
errors=[(2.^nn)' error_call' error_put']
```

Example 99 Figure 6.8 and 6.10 have been developed as follows.

```
% Pricing options using the VG model - Implied Volatility versus \theta and k
%
% INPUT
% S: spot
% K: strike
% r: risk free rate (cont. compounded)
% q: dividend yield
% T: maturity
% x0: parameter set
% type: handle to choose the model process (BM, VG...)
% N, L: COS parameters
%%%%%%%%%%%%%%%
clear all; clc; close all;

type=2;

N=2^10;
L=10;

r = 0.05;
S = 100;
q = 0.02;
K = (80:5:120);
t0=0;
T = 1;
% VG parameters: testing theta
theta = [-0.2 0 0.2];
kappa = 0.25;
sigma = 0.2;
```

```

for jj=1:length(theta)
x0=[theta(jj) sigma kappa];

for j=1:length(K)
[price, ~]=get_VanillaEuro_COS(S,K(j),r,q,T,x0,type,N,L);
CalImpl.Vol(j,jj)=blsimpv(S,K(j),r,T,price,[],q,1e-4);

end
end
figure
plot(K',CalImpl.Vol(:,1), 'b-o',K',CalImpl.Vol(:,2), 'g-d',K',CalImpl.Vol(:,3), 'r-s')
xlabel('Strike')
ylabel('implied Volatility')
legend('\theta = -0.2', '\theta = 0', '\theta = 0.2', 'Location', 'Best');

% VG parameters: testing kappa
theta = -0.2 ;
kappa = [0.15 0.25 0.35];
sigma = 0.2;
for jj=1:length(kappa)
x0=[0 theta sigma kappa(jj)];
omega=log(CharFunc(x0,type,-li,1));
mu=r-q-omega;
param=[mu x0(2:end)];
for j=1:length(K)
[price, ~]=get_VanillaEuro_COS(S,K(j),r,q,T,param,type,N,L);
CalImpl.Vol(j,jj)=blsimpv(S,K(j),r,T,price,[],q,1e-4);
end
end
figure
plot(K',CalImpl.Vol(:,1), 'b-o',K',CalImpl.Vol(:,2), 'g-d',K',CalImpl.Vol(:,3), 'r-s')
xlabel('Strike')
ylabel('implied Volatility')
legend('k = 0.15', 'k = 0.25', 'k = 0.35', 'Location', 'Best');

```

