Stochastic differential equations in financial modelling

Mathematics, IC, London

Prof. Damiano Brigo
Mathematical Finance Section
Dept. of Mathematics

Imperial College London

Course Layout I

1. **PART 1: PROBABILITY and SDEs**
2. Probability space $(\Omega, \mathcal{F}, \mathbb{P})$
   - Sample space
   - Sigma-fields
   - The probability measure
3. Random variable on $(\Omega, \mathcal{F}, \mathbb{P})$
   - Common random variables
   - Moment generating and characteristic functions
   - Other facts on random variables
   - Central limit theorem and normal r.v. moments
   - Multivariate random variables
   - Covariance and correlation
4. Convergence of random variables
   - Almost sure and L^p convergence
   - Convergence in mean square and in probability

Course Layout II

- Weak convergence or convergence in law/distribution
- Relationship between types of convergence

5. Stochastic processes
   - Brownian Motion or Wiener Process W_t
   - Some intuition on Brownian motion
   - Simulation of Brownian motion
6. Ordinary Differential Equations (ODEs)
7. Stochastic Differential Equations (SDEs)
   - Stochastic Integrals: Ito and Stratonovich
   - Existence and uniqueness of solutions
   - Diffusion processes
   - Itô's formula
   - Quadratic variation
   - Example: Arithmetic to Geometric Brownian Motion
   - Example: Geometric to Arithmetic Brownian Motion

Course Layout III

- Example: Ornstein-Uhlenbeck process
- Square root process (CIR model)
- Examples of mean–reverting SDEs

8. Stochastic differential of a product & quadratic covariation
9. Martingales
10. PDE solutions as expectations of SDEs: Feynman-Kac theorem
11. Change of measure: Girsanov's theorem
12. **PART 2: SDEs FOR OPTION PRICING**
13. No arbitrage and option pricing / hedging
    - The Black Scholes and Merton Analysis
    - Contingent Claims
    - Strategies, Value process, Gains
    - Self financing Strategies, attainable claims & arbitrage
    - Call and Put options
    - The Black-Scholes PDE

Course Layout IV

- Black and Scholes fomula for Call Options
- In-, at-, and out-of-the money options
- The Feynman Kac theorem, Girsanov and the Martingale Measure
- Fundamental Theorems: No arbitrage and completeness
- Martingales and numeraires
- Zero-coupon bond
- Put Call parity and forward contracts
- Black and Scholes formula for a Put option
- Hedging
- Incomplete Markets
- The Greeks / Sensitivities

14. Volatility smile modeling
    - Introducing the volatility smile
    - Bachelier Model

Course Layout V

- Displaced Diffusion Model
- CEV model
- Anticipating the smile type based on the return volatility
- The mixture diffusion dynamics model
- The shifted mixture dynamics model
- Local volatility models & future smile flattening
- Hint at Stochastic volatility models (Heston model)

15. Summary so far and... Crises
    - Origin, size and relevance of derivatives markets
    - 10 $\times$ planet GDP: Thales, Bachelier, Black & Scholes
    - What does it all mean?
    - From 1997 Nobel to crises: ... 1998, 2007, 2008...
    - Crises: Barings collapse
    - Crises: Risk Measures and Default risk

16. **PART 3: RISK MEASURES**

Course Layout VI

17. Properties of financial time series under the measure P
    - Mean, STD, skewness & kurtosis of historical returns
    - Historical vs Implied Volatilities
    - Histograms and QQ Plots
18. Value at Risk
    - VaR definition and intuition
    - VaR drawbacks and Expected Shortfall
19. Expected Shortfall
    - Definition
    - Drawbacks of ES
20. Risk measures: numerical examples and codes
    - VaR and ES: Short Straddle
    - VaR and ES: Risk Reversal
    - VaR and ES: Bull call spread
    - VaR and ES: Options on different correlated stocks

Course Layout VII

- Volatilties and correlations

21. **PART 4: NUMERICAL SOLUTION OF SDEs**
22. Euler-Maruyama numerical schemes for SDEs
    - General Euler scheme
    - Example: Geometric Brownian motion
    - Example: $dX_t = m\,dt + \sigma X_t dW_t$
    - Example: study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$
23. **MOCK EXAMS WITH SOLUTIONS**
24. Exam 1
    - Problem 1. SDEs: $dX = m\,dt + \sigma X\,dW$
    - Problem 2. Straddle Option Pricing - Black Scholes.
    - Problem 3. Straddle Option Pricing - Displaced Diffusion.
    - Problem 4. Risk measures: VaR for a stock position
25. Exam 2
    - Problem 1. SDEs: $dX = m(X-a)dt + \sigma(X-a)dW$

Course Layout VIII

- Problem 2. Option pricing Black Scholes - 2 stocks
- Problem 3. Option pricing with smile - 2 stocks
- Problem 4. Risk measures - 2 stocks & Bachelier model

26. Exam 3
    - Problem 1: SDEs with cubic root
    - Problem 2: Asset/Cash or nothing options in Black Scholes
    - Problem 3: Asset/Cash or nothing options with smile
    - Problem 4: Risk Measures - ES for a stock in Bachelier
27. Exam 4
    - Problem 1: SDEs $dX = v^2 X(1 + X^2)dt + v(1 + X^2)dW$
    - Problem 2: Option pricing - no arbitrage & put-call parity
    - Problem 3: Smile modeling - butterfly spread in BS & Bachelier
    - Problem 4. Risk measures: long Bond & short forward in B&S
28. Exam 5
    - Problem 1: SDEs - OU Process

Course Layout IX

- Problem 2: Option pricing - Short Risk Reversal in Black Scholes
- Problem 3: Option pricing - Bear Call Spread with DD
- Problem 4: Risk measures. Bond & Stock with different maturities

29. Exam 6
    - Problem 1: SDEs - $dX = X/2dt + \sqrt{1+X^2}dW$
    - Problem 2: Option pricing - Bull call spread in Black Scholes
    - Problem 3: Option pricing - Bull Call Spread with Bachelier
    - Problem 4. Risk measures: Short Bond & long forward in B&S
30. Exam 7
    - Problem 1: CIR SDE $dX = k(\theta - X)dt + \sigma\sqrt{X}dW$
    - Problem 2: Long call calendar spread in Black Scholes
    - Problem 3: Long call calendar spread in Bachelier
    - Problem 4: Risk Measures. Forwards with different T in BS
31. Mastery Questions

Course Layout X

32. Mastery Question 2025-2026
33. Coursework 2025-2026
34. **APPENDIX: PYTHON CODES**
    - Simulation of Brownian motion
    - Short straddle
    - Portfolio of call and put on two correlated stocks
    - Bull Call Spread
    - Risk Reversal
    - Euler scheme for Geometric Brownian Motion
    - Euler scheme for $dX = mdt + \sigma XdW$
    - Euler scheme $dX = (-3kX + 3X^{1/3}\sigma^2)dt + 3\sigma X^{2/3}dW_t$
    - More efficient Python codes with vectorization
    - Vectorized simulation of Brownian motion
    - Vectorized Short straddle
    - Vectorized portf. of call & put on correlated stocks

Course Layout XI

- Vectorized Bull Call Spread
- Vectorized Risk Reversal
- Vectorized Euler scheme for Geometric Brownian
- Vectorized Euler scheme for $dX = mdt + \sigma XdW$
- Vectorized Euler $dX = (-3kX + 3X^{1/3}\sigma^2)dt + 3\sigma X^{2/3}dW_t$

A personal note on health

I suffer from a progressive health condition affecting my ears.

Six years ago a virus affected my inner ear, with problems of balance, hearing loss and tinnitus.

To this day I have extremely loud tinnitus (electric whistling in the ears), ear pain, and I struggle to hear. I also have ear sensitivity, so I need to wear ear protection.

Please when you talk to me from close, for questions etc, keep your voice low.
I'm always available for questions on the course, or material, or office hours but I need to limit conversations and I can't do much small talk or other interaction, career advice, projects, projects supervision, unfortunatelty, on top of the work I already do for the MSc. I can however chat in teams or email or ED.

Format of lecture notes

The slides you are seeing are the lecture notes. I have written them in this form to maximise visual memory from attending the lectures.

Please note that these are slides written in LaTeX with the package “beamer” and are not a “power-point presentation”.

There are occasional requests of students to have slides more in a book format, but this would require to rewrite everything in book format and would lose the visual memory benefit when revising.

Going to the initial table of contents, the different parts of the course are indexed with a click, so it is easy to access topics and mock exams.

If some students don’t attend lectures, the visual memory benefit will be lost and a few non-attending might prefer a book format, but I designed these notes assuming that students attend lectures.

PART 1: PROBABILITY and SDEs

In this part we recall the main notions from probability theory we will need and we introduce and explain Stochastic Differential Equations (SDEs)

This part is NOT meant to be an extensive exposition, but more an informal road map for students to make them aware of the relevant notions with basic intuition.

The sample space

Definition (Ω – sample space)

It is the set of all elementary outcomes of a random experiment.

Example

Suppose the experiment is rolling a dice, then $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Suppose a second experiment is measuring the length of a mechanical piece within a few decimal places. Then we could take $\Omega = \mathbb{Q}$.

Other times we can take, more generally, $\Omega = \mathbb{R}$.

Sigma-fields and events

An event A is a subset of the sample space Ω, $A \subseteq \Omega$. Usually we will not be interested in all subsets, but only in some subsets that belong in some set of subsets. This set of subsets A of Ω we are interested in, the set of events $\mathcal{F}$ containing all A's we are interested in, has the structure of a sigma-field, which we define below.

Definition ($\mathcal{F}$ – sigma-field)

Also denoted by σ-field or called sigma-algebra, a sigma-field $\mathcal{F}$ is the set of all possible subsets A of Ω ($A \subseteq \Omega$) we are interested in, called events, which satisfies the following properties:

- is non-empty $\mathcal{F} \neq \Phi$ and contains Ω: $\Omega \in \mathcal{F}$
- is closed under COMPLEMENTATION; if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$
- is closed under COUNTABLE UNIONS; if $A_i \in \mathcal{F}$ for $i \in \mathbb{N}$ then $\cup_i A_i \in \mathcal{F}$

Sigma-fields and events

Given that $\Omega \in \mathcal{F}$, by the second property $\Omega^c = \Phi \in \mathcal{F}$. This leads to the trivial σ-field $\mathcal{F} = \{\Omega, \Phi\}$, the smallest possible $\mathcal{F}$. Another one is the power set $\wp(\Omega)$ also denoted by $\mathcal{P}(\Omega)$, namely the set of all subsets of Ω, which is the largest possible $\mathcal{F}$. Prove that both are sigma-fields.

An important σ-field when $\Omega = \mathbb{R}$ is the Borel σ-field, namely the σ-field of $\mathbb{R}$ generated by open intervals in $\mathbb{R}$. This is called BOREL SET of $\mathbb{R}$ and sometimes denoted by $\mathcal{B}(\mathbb{R})$. It contains all possible countable unions of intervals among other subsets of $\mathbb{R}$.

The reasons for taking intervals in the above definition is that we know how to measure intervals in $\mathbb{R}$. The measure of $[a\ b]$ is simply $b - a$. This is called Lebesgue measure and those of you who studied measure theory know about this and Lebesgue integration, but we won't insist on that here.

Sigma-fields and events

An event A is a set of elementary outcomes, $A \subseteq \Omega$. If the random experiment produces ω_1 and $\omega_1 \in A_1$, then we say that the event A_1 happened.

Example

Assume, as in the dice experiment, $\Omega = \{1, 2, 3, 4, 5, 6\}$ and suppose you are interested in the events:

- $A = \{$even numbers: $2, 4, 6\}$;
- $B = \{$odd numbers: $1, 3, 5\}$;
- $C = \{$numbers smaller than 3: $1, 2\}$.

If $\omega = 2$, then A and C happened, but not B.

Sigma-fields and events

Example

Assume we are measuring a mechanical piece in meters up to the 6th decimal digit. We take $\Omega = \mathbb{R}$ (we could also take $\Omega = \mathbb{Q}$) and suppose you are interested in the events:

- $A = [0,\ 1]$;
- $B = [0,\ 0.5] \cup [1,\ 1.5]$;
- $C = [1,\ 1.5]$.

If our measurement comes out as $\omega = 0.143234$, then A and B happened, but not C, because $\omega \in A$, $\omega \in B$ but ω is not in C. Also, $A \cup B$ happened, $B \cap C$ did not happen.

Sigma-fields and events

Given a set of subsets of Ω, say $\mathcal{A}$, we denote by $\sigma(\mathcal{A})$ the sigma-field generated by $\mathcal{A}$. This is the smallest sigma-field on Ω containing $\mathcal{A}$. It will contain all the countable unions, intersections and complementations of elements in $\mathcal{A}$ and combinations of these operations. Consider for example $\mathcal{A} = \{A\}$, one subset of Ω in $\mathcal{F}$.

$$\sigma(\{A\}) = \{A, A^c, \Omega, \Phi\}.$$

Consider for example $\mathcal{A} = \{A, B\}$, two subsets of Ω in $\mathcal{F}$. We have

$$\sigma(\{A, B\}) = \{A \cap B^c, A \cap B, A^c \cap B, A, B, A \cup B, (A \cap B^c) \cup (B \cap A^c),$$

$$A^c \cup B, A^c \cup B^c, A \cup B^c, A^c, B^c, A^c \cap B^c, (A^c \cup B) \cap (B^c \cup A), \Omega, \Phi\}$$

The probability measure

Definition ($\mathbb{P}$ - probability measure)

Is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$, i.e. it associates to every event $A \in \mathcal{F}$ a number $\mathbb{P}(A)$ between 0 and 1, which is interpreted as the probability of that event A. To be a probability measure, $\mathbb{P}$ should satisfy the following properties:

- $\mathbb{P}(\Phi) = 0$ and $\mathbb{P}(\Omega) = 1$;
- $\mathbb{P}$ is COUNTABLY ADDITIVE,
  i.e. if $A_i \in \mathcal{F}$, where $i \in I$, with I countable (e.g. $I = \mathbb{N}$), and $A_i \cap A_j = \Phi$ for $i \neq j$, then $\mathbb{P}(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mathbb{P}(A_i)$.

The random variable

Definition (X – random variable)

A random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$ is a function $X : \Omega \to \mathbb{R}$ that is MEASURABLE wrt the Borel σ-field on $\mathbb{R}$. This means that

$$B \text{ Borel set of } \mathbb{R} \implies X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}.$$

The world "measurable" is meant to convey the fact that X brings back things that I can measure in $\mathbb{R}$ (with Lebesgue measure) to things that I can measure in $\mathcal{F}$ (with $\mathbb{P}$). If I can measure B in $\mathbb{R}$, and I can because it is in the Borel set, then bringing B back with X^{-1} gives me something I can measure in $\mathcal{F}$.
Indeed, since $X^{-1}(B) \in \mathcal{F}$ we can "measure" it by computing its probability:

$$\mathbb{P}\big(X^{-1}(B)\big) = \mathbb{P}\big(\{\omega \in \Omega : X(\omega) \in B\}\big)$$

CDF and PDF of X

For the particularly important case of $B = (-\infty, b]$:

Definition (Cumulative Distribution Function of X)

$$\mathbb{P}\big(X^{-1}((-\infty, b])\big) = \mathbb{P}\big(\{\omega \in \Omega : X(\omega) \leq b\}\big) =: F_X(b)$$

or shortly, $\mathbb{P}(X \leq b) =: F_X(b)$.

Definition (Probability Density Function of X)

If $\exists$ a function $p_X : \mathbb{R} \to \mathbb{R}^{\geq 0}$ such that

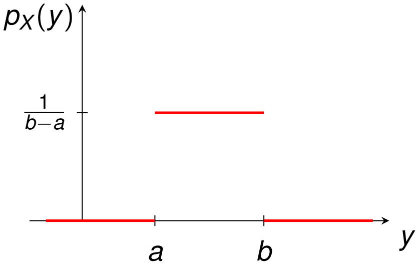
$$F_X(b) = \int_{-\infty}^{b} p_X(y)\,\mathrm{d}y,$$

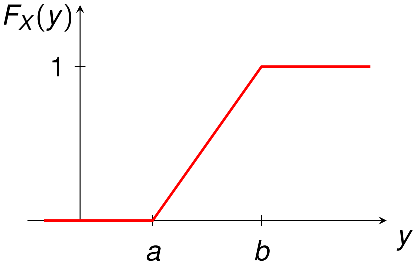
then p_X is the Probability Density Function of X.
If F_X is differentiable, then $F_X' = p_X$. For more on densities and on discrete vs continuous r.v.'s see appendix.

Uniform random variable on $[a, b]$

$X \sim \mathcal{U}(a, b)$, "X is distributed as a uniform random variable in $[a, b]$" if

$$p_X(y) = \begin{cases} \frac{1}{b-a} & \text{if} \quad y \in [a, b] \\ 0 & \text{otherwise} \end{cases} \qquad F_X(y) = \begin{cases} 0 & \text{if} \quad y < a \\ \frac{y-a}{b-a} & \text{if} \quad y \in [a, b] \\ 1 & \text{if} \quad y \geq b \end{cases}$$





Expected value and variance

Expected value of a random variable X is

$$\mu_X = \mathbb{E}[X] = \int_\Omega X(\omega)\, \mathrm{d}\mathbb{P}(\omega) = \int_\mathbb{R} y\, dF_X(y) \;= \int_\mathbb{R} y\, p_X(y)\, \mathrm{d}y$$

The variance is

$$\mathrm{Var}[X] = \sigma_X^2 = \mathbb{E}[(X - \mu_X)^2] = \mathbb{E}[X^2] - \mu_X^2$$

The standard deviation is σ_X, sometimes denoted $\mathrm{Std}(X)$, the positive square root of variance, and is an index of dispersion of the random variable values around the mean. Expected value is linear:

$$\mathbb{E}[\alpha X + Y] = \alpha \mathbb{E}[X] + \mathbb{E}[Y], \;\; \alpha \in \mathbb{R}.$$

$VAR[\alpha X] = \alpha^2 VAR[X]$ but VAR and standard deviation are not linear, $VAR[X + Y]$ may be different from $VAR[X] + VAR[Y]$ in general.

Exercise: compute mean and standard deviation of $X \sim \mathcal{U}(a, b)$.

Expected value and variance of a uniform I

Do this on your own and check. For $X \sim \mathcal{U}(a, b)$,

$$\mathbb{E}[X] = \int_{\mathbb{R}} x p_X(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^2}{2}|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

which is, not surprisingly, the middle point between a and b, being the probability density uniform between a and b and zero elsewhere.

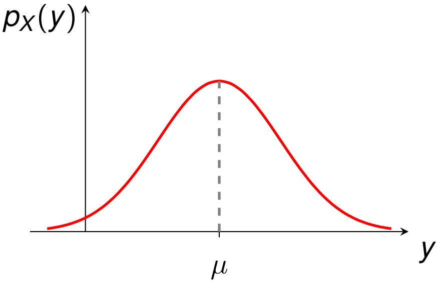
$$Var[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \int_a^b x^2 \frac{1}{b-a} dx - \left(\frac{a+b}{2}\right)^2 =$$

$$= \frac{1}{b-a} \frac{x^3}{3}|_a^b - \frac{a^2 + b^2 + 2ab}{4} = \frac{b^3 - a^3}{b-a} - \frac{a^2 + b^2 + 2ab}{4} =$$

$$= \frac{a^2 + ab + b^2}{3} - \frac{a^2 + b^2 + 2ab}{4} = \frac{(b-a)^2}{12}.$$

Normal or Gaussian random variable $X \sim \mathcal{N}(\mu, \sigma^2)$

$$p_X(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(y-\mu)^2}{\sigma^2}\right)$$



Note that μ is the MEAN, MEDIAN and MODE, while σ^2 is the VARIANCE and $\sigma = \sqrt{\sigma^2}$ the STANDARD DEVIATION.

$$\mu = \mathbb{E}[X] = \int_\Omega X(\omega)\, d\mathbb{P}(\omega) = \int_\mathbb{R} y\, p_X(y)\, dy,$$

$$\mathrm{Var}[X] = \sigma^2 = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mu^2,$$

skewness (asymmetry in tails):

$$\frac{\mathbb{E}[(X-\mu)^3]}{\sigma^3} = 0,$$

excess kurtosis (fatness of tails):

$$\frac{\mathbb{E}[(X-\mu)^4]}{\sigma^4} - 3 = 0.$$

Moment generating and characteristic functions

In general, for a random variable X its moment generating function is defined as

$$M_X(u) = \mathbb{E}[\exp(uX)].$$

The name "moment generating function" is due to the fact that derivatives of M_X computed at $u = 0$ provide the moments of X. The n-th moment of X is defined as $\mathbb{E}[X^n]$. We have

$$\frac{d}{du} M_X(u)|_{u=0} = \mathbb{E}[X], \; \frac{d^2}{du^2} M_X(u)|_{u=0} = \mathbb{E}[X^2] \; \text{ etc}$$

Not all random variables admit a moment generating function. A generalization is the characteristic function, which is defined as $\phi_X(u) = \mathbb{E}[\exp(i\, uX)]$ where i is the imaginary unit number $0 + 1i \in \mathbb{C}$.

We will use occasionally the moment generating function of a normal random variable $X \sim \mathcal{N}(\mu, \sigma^2)$. This is

$$M_{\mathcal{N}(\mu,\sigma^2)}(u) = e^{(u\mu + \frac{1}{2} u^2 \sigma^2)}$$

Exponential Random Variable with intensity γ

This is a random variable X with cumulative distribution function

$$F_X(y) = 1 - \exp(-\gamma y), \quad y \geq 0; \quad F_X(y) = 0 \ \ \textit{if} \ \ y < 0$$

and with density

$$p_X(y) = \gamma \ \exp(-\gamma y) \quad \textit{for} \ \ y \geq 0; \quad p_X(y) = 0 \ \ \textit{if} \ \ y < 0$$

We have

$$\mu = \mathbb{E}[X] = 1/\gamma; \quad \sigma^2 = \mathbb{E}[(X-\mu)^2] = 1/\gamma^2.$$

Important property: Lack of memory. Conditional probability:

$$\mathbb{P}(X > x + y | X > y) := \frac{\mathbb{P}(\{X > x + y\} \cap \{X > y\})}{\mathbb{P}(X > y)} = \mathbb{P}(X > x).$$

This property is important when X models arrival times, for example. In finance X is used to model default times.

Lognormal r.v.

Lognormal random variable

Y is lognormal if $Y = e^X$ with $X \sim \mathcal{N}(\mu, \sigma^2)$. We can compute the expectation of a lognormal r.v. as

$$\mathbb{E}[e^{\mathcal{N}(\mu,\sigma^2)}] = M_{\mathcal{N}(\mu,\sigma^2)}(1) = e^{\mu + \frac{1}{2}\sigma^2}$$

via the moment generating function of a normal. For the variance we have

$$\mathrm{Var}(e^X) = \mathbb{E}[(e^{\mathcal{N}(\mu,\sigma^2)})^2] - \mathbb{E}[e^{\mathcal{N}(\mu,\sigma^2)}]^2 = \mathbb{E}[e^{2\mathcal{N}(\mu,\sigma^2)}] - (e^{\mu+\frac{1}{2}\sigma^2})^2 =$$

$$= M_{\mathcal{N}(\mu,\sigma^2)}(2) - e^{2\mu+\sigma^2} = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2}.$$

Independent random variables X and Y

Two r.v. X and Y are INDEPENDENT (see also Appendix) if

$$\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B) \quad \forall\, A, B \in \mathcal{B}(\mathbb{R}).$$

For independent X and Y: $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$,
$Var[X + Y] = Var[X] + Var[Y]$.

Sum of independent normals is normal

$X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$; $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$; X_1 and X_2 independent

$$\Longrightarrow\; X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Sum and products of independent lognormals

X_1, X_2 independent lognormals $\Longrightarrow$ $X_1 + X_2$ is not lognormal but

$$X_1 X_2 \text{ is lognormal : think of } X_1 = e^{\mathcal{N}_1}, X_2 = e^{\mathcal{N}_2}$$

Central limit theorem and normal r.v. moments I

Going back to normal random variables, the normal is one of the most important r.v. in statistics, due to the central limit theorem. This states that if $X_1, X_2, \ldots, X_i, \ldots$ is a sequence of independent and identically distributed (i.i.d.) random variables, X_i with finite mean μ and finite variance σ^2 , each with the same distribution that *need not be normal*, and the sample mean of the first n r.v.'s is $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, then we have the following convergence in distribution or in law (see below "convergence of random variables" to see what this means exactly)

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \xrightarrow[n\uparrow\infty]{law} \mathcal{N}(0,1)$$

Central limit theorem and normal r.v. moments II

In other words, the rescaled sample mean $\bar{X}_n$ converges to a normal even if the single r.v's X_i are not normal. This is one of the reasons why we use the name "normal" for the Gaussian, it is the normal type of limit you find for any type of random variables sample mean.

Another result we will use from normal random variables is their moments. Assume $X \sim \mathcal{N}(\mu, \sigma^2)$. Then the central moments of the normal X are given by

$$\mathbb{E}[(X-\mu)^n] = \begin{cases} 0 & \text{for } n \text{ odd} \\ \sigma^n\,(n-1)!! & \text{for } n \text{ even} \end{cases}$$

where the semi-factorial of an odd integer m is defined as

$$m!! = m(m-2)(m-4)\cdots 3\;1 \quad \text{for } m \text{ odd.}$$

For example, $\mathbb{E}[(X-\mu)^3] = 0$, $\mathbb{E}[(X-\mu)^4] = 3\sigma^4$, $\mathbb{E}[(X-\mu)^5] = 0$, $\mathbb{E}[(X-\mu)^6] = 15\sigma^6$ etc.

Multivariate random variables

These are formally defined as measurable $X : \mathcal{F} \to \mathbb{R}^n$, meaning that B is Borel in $\mathbb{R}^n \Rightarrow X^{-1}(B) \in \mathcal{F}$.
The components can be written as X_i, so that $X = [X_1, X_2, \ldots, X_n]$ can be put in vector form.
The cumulative distribution function for a multivariate random variable is

$$F_X(x_1, x_2, \ldots, x_n) = \mathbb{P}(X_1 \le x_1 \cap X_2 \le x_2 \cap \ldots \cap X_n \le x_n).$$

If there exists a function $p_X : \mathbb{R}^n \to \mathbb{R}^{\ge 0}$ such that

$$F_X(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \ldots \int_{-\infty}^{x_n} p_X(y_1, y_2, \ldots, y_n)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \ldots \mathrm{d}y_n$$

then p_X is the Probability Density Function of X.
If F_X is differentiable, then $\frac{\partial^n F_X}{\partial x_1 \ldots \partial x_n} = p_X$.

Covariance and correlation

Take a bivariate random variable $X = [X_1, X_2]$. The covariance between X_1 & X_2 is defined as

$$\mathrm{cov}(X_1, X_2) = \mathbb{E}[(X_1 - \mathbb{E}[X_1])(X_2 - \mathbb{E}[X_2])] = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]$$

and is a number expressing how much X_1 and X_2 vary together, similarly to how the variance of X is a number expressing how much X varies (due to randomness).

Note that if X_1 and X_2 are independent then we have immediately $\mathrm{cov}(X_1, X_2) = 0$ from $E[X_1 X_2] = E[X_1]E[X_2]$. Similarly, if at least one of X_1 or X_2 is a deterministic constant, the covariance is zero.

The opposite is not always true: we can have two r.v. with zero covariance but that are not independent, see example with X and X^2 below.

Covariance and correlation

Correlation. The correlation between X_1 and X_2 is defined as

$$\rho_{1,2} = \mathrm{cor}(X_1, X_2) = \frac{\mathrm{cov}(X_1, X_2)}{\sqrt{Var(X_1)}\sqrt{Var(X_2)}} = \frac{\mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]}{Std(X_1)\ Std(X_2)}.$$

Correlation, due to Schwartz inequality in L^2, is always in the interval $[-1, 1]$. Two independent random variables have zero correlation but the opposite is not always true, in that two random variables with zero correlation are not necessarily independent.

Take for example $X \sim \mathcal{N}(0, 1)$ and set $X_1 = X, X_2 = X^2$. The correlation is zero because
$E[XX^2] - E[X]E[X^2] = E[X^3] - 0\ 1 = 0 - 0 = 0$ where we used the moments of a normal random variable. Now, even with zero correlation, the two random variables are clearly dependent, as one is the square of the other one. Square is not one-to-one, but still expresses a clear dependence, and yet the correlation is zero.

Covariance and correlation

Similarly, we can find two random variables that are totally dependent in a one-to-one positive relationship but whose correlation is less than one. Take again $X \sim \mathcal{N}(0,1)$ and set $X_1 = X, X_2 = X^3$. We have

$$\rho_{1,2} = \frac{\mathbb{E}[X\,X^3] - \mathbb{E}[X]\mathbb{E}[X^3]}{Std(X)\;Std(X^3)} = \frac{E[X^4]}{\sqrt{Var[X^3]}}.$$

Now, recalling the formula for the central moments of the Gaussian and taking into account that X has zero mean μ and $\sigma = 1$, we get $Var[X^3] = E[(X^3)^2] - E[X^3]^2 = E[X^6] - 0^2 = 15 - 0 = 15$ and $E[X^4] = 3$. Thus

$$\rho_{1,2} = \frac{E[X^4]}{\sqrt{Var[X^3]}} = \frac{3}{\sqrt{15}} = 0.77.... < 1$$

Hence we see that even if X and X^3 are related by a one to one increasing and invertible transformation, their correlation is less than one.

Covariance and correlation

This is because correlation expresses the *linear* dependence between two random variables. If correlation is equal to 1, the two variables are totally related by a positive slope linear transformation, whereas if $\rho = -1$ the slope is negative. All other cases measure partial linear dependence:

$$\rho_{1,2} = 1 \text{ for } X_2 = mX_1 + q, \;\; \rho_{1,2} = -1 \text{ for } X_2 = -mX_1 + q$$

where m is a positive real constant.

Covariance and correlation

For an example of two correlated random variables, given three independent random variables X, Y, Z, each with mean zero and variance 1, set

$$X_1 = \sqrt{\alpha}X + \sqrt{1-\alpha}\,Y, \quad X_2 = \sqrt{\alpha}X + \sqrt{1-\alpha}Z$$

for a positive constant $\alpha \in [0,1]$. It is easy to show $\rho_{1,2} = \alpha$ based on linearity. For example,
$E[X_1] = E[\sqrt{\alpha}X + \sqrt{1-\alpha}\,Y] = \sqrt{\alpha}E[X] + \sqrt{1-\alpha}E[Y] = 0$ because $E[X] = E[Y] = 0$. Then

$$E[X_1\,X_2] - E[X_1]E[X_2] = E[(\sqrt{\alpha}X + \sqrt{1-\alpha}\,Y)(\sqrt{\alpha}X + \sqrt{1-\alpha}Z)] - 0 \cdot 0 =$$

$$= \alpha E[X^2] + \alpha\sqrt{1-\alpha}E[XZ] + \alpha\sqrt{1-\alpha}E[YX] + (1-\alpha)E[YZ]$$

$$= \alpha 1 + \underbrace{\alpha\sqrt{1-\alpha}E[X]E[Z]}_{X\ \&\ Z\ independ.} + \underbrace{\alpha\sqrt{1-\alpha}E[Y]E[X]}_{Y\ \&\ X\ independ.} + \underbrace{(1-\alpha)E[Y]E[Z]}_{Y\ \&\ Z\ independ.} = \alpha$$

as $E[X] = E[Y] = E[Z] = 0$ and $E[X^2] = E[X^2] - E[X]^2 = Var[X] = 1$.

Covariance and correlation

Along similar lines, one shows that $Var[X_1] = 1,\ Var[X_2] = 1$ and then $\rho_{1,2} = \alpha$. For example,

$$Var[X_1] = Var[\sqrt{\alpha}X + \sqrt{1-\alpha}\,Y] = Var[\sqrt{\alpha}X] + Var[\sqrt{1-\alpha}\,Y] = \dots$$

where we used that the variance of the sum of independent r.v. is the sum of the variances. Then, as $\mathrm{Var}[\beta X] = \beta^2 Var[X]$,

$$\dots = \alpha Var[X] + (1-\alpha) Var[Y] = \alpha \cdot 1 + (1-\alpha) \cdot 1 = 1$$

as both X and Y have variance 1.

However, for nonlinear transformation, this type of correlation may not work well. Indeed, we have X and X^2 that are clearly related but have correlation zero, and X and X^3 that are totally related by a one-to-one transformation whose correlation is less than one.

Covariance and correlation

The above $\rho_{1,2}$ is called Pearson (linear) correlation. There are other correlations that solve the above problems, called rank correlations (e.g. Kendall tau or Spearman rho) but we won't study them here.

Covariance and correlation matrices

The matrix $\text{cov}(X) = (\text{cov}(X_i, X_j))_{i,j=1,\dots,n}$ is called covariance matrix of the random vector X.

The matrix $\rho = (\rho_{i,j})_{i,j}$ is called the correlation matrix of the random vector X.

These matrices are positive semidefinite.

The covariance can be expressed in terms of standard deviations of the two variables X_i and X_J and of their correlation $\rho_{i,j}$ as $\text{cov}(X_i, X_j) = \sigma_i \sigma_j \rho_{i,j}$.

Multivariate normal

We can now define a multivariate normal random variable in dimension n. Let $\mu = [\mu_1, \mu_2, \ldots, \mu_n]$ be the vector for the means and let V be a covariance matrix, $V_{i,j} = \sigma_i \sigma_j \rho_{i,j}$.

We say that X follows a multivariate normal distribution in dimension n and we write
$X = [X_1, \ldots, X_n] \sim \mathcal{N}(\mu, V) \sim \mathcal{N}((\mu_i)_{i=1,\ldots,n}, (\sigma_i \sigma_j \rho_{i,j})_{i,j=1\ldots n})$ if

$$p_X(y) = p_{X_1,\ldots,X_n}(y_1, \ldots, y_n) = \frac{(2\pi)^{-n/2}}{\sqrt{det(V)}} \exp\left(-\frac{1}{2}(y-\mu)V^{-1}(y-\mu)^T\right)$$

where $(y-\mu)^T$ denotes the column vector obtained by transposition of the row vector $(y-\mu) = [y_1 - \mu_1, \ldots, y_n - \mu_n]$.

Multivariate normal

We note that
$X = [X_1, \ldots, X_n] \sim \mathcal{N}(\mu, V) \sim \mathcal{N}((\mu_i)_{i=1,\ldots,n}, (\sigma_i \sigma_j \rho_{i,j})_{i,j=1\ldots n})$ implies that the components are normal too, namely $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$.
We further point out that sum of components of a multivariate normal is still normal, namely

$$\sum_i X_i \sim \mathcal{N}(\sum_i \mu_i, \sum_{i,j} \sigma_i \sigma_j \rho_{ij}).$$

Be careful here: if we only know that X_1 and X_2 are normal but we don't know that $[X_1, X_2]$ is a bivariate normal, it does NOT follow that $X_1 + X_2$ is normal.

Convergence of Random Variables

Suppose we have, on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a sequence of random variables $X_1, X_2, \ldots, X_n, \ldots$. The sequence can converge to a limit random variable $\bar{X}$ on the same space.

Almost sure (a.s.) convergence. We say that $X_n \xrightarrow{a.s.} \bar{X}$ if

$$\mathbb{P}\{\omega : \lim_n X_n(\omega) = \bar{X}(\omega)\} = 1.$$

In other terms, the set of $\omega \in \Omega$ where the sequence converges is measurable and has probability one.

L^p **convergence.** We say that $X_n \xrightarrow{L^p} X$ $(p \geq 1)$ if $\mathbb{E}\{|X_n|^p\} < +\infty$ for all n and

$$\lim_n \mathbb{E}[|X_n - \bar{X}|^p] = 0.$$

Convergence of Random Variables

A special case of L^p convergence is $p = 2$, which is called mean squared convergence:

Mean square (m.s.) convergence. We say that $X_n \xrightarrow{m.s.} \bar{X}$ if

$$\mathbb{E}[|X_n|^2] < +\infty \text{ and } \lim_n \mathbb{E}[|X_n - \bar{X}|^2] = 0.$$

Convergence in probability. We say that $X_n \xrightarrow{\mathbb{P}} \bar{X}$ if

$$\text{for all } \epsilon > 0, \lim_n \mathbb{P}\{\omega : |X_n(\omega) - \bar{X}(\omega)| > \epsilon\} = 0.$$

Convergence of Random Variables

Weak convergence / convergence in law / in distribution. Here the random variables X_n can be even defined on different probability spaces. We say that we have convergence in law $X_n \xrightarrow{law} \bar{X}$ if

$$\lim_n \mathbb{E}[f(X_n)] = \mathbb{E}[f(\bar{X})] \;\text{ for all } f \text{ continuous and bounded.}$$

Equivalently,

$$\lim_n F_{X_n}(x) = F_{\bar{X}}(x) \;\text{ for all } x \text{ where } F_{\bar{X}} \text{ is continuous.}$$

Convergence of Random Variables

The strongest convergence is the a.s. convergence and the L^p convergence. They don't imply each other.

Both of them imply convergence in probability but convergence in probability does not imply convergence in L^p ($p \geq 1$) or *a.s.*

Convergence in probability implies in law but not vice versa.

$$X_n \xrightarrow{a.s} \bar{X} \ \Rightarrow \ X_n \xrightarrow{\mathbb{P}} \bar{X} \ \Rightarrow \ X_n \xrightarrow{law} \bar{X}$$

$$(p > q) \ \ X_n \xrightarrow{L^p} \bar{X} \Rightarrow X_n \xrightarrow{L^q} \bar{X} \ \Rightarrow X_n \xrightarrow{\mathbb{P}} \bar{X} \Rightarrow X_n \xrightarrow{law} \bar{X}$$

$$X_n \xrightarrow{a.s} \bar{X} \ \not\Rightarrow X_n \xrightarrow{L^p} \bar{X}, \ \ X_n \xrightarrow{L^p} \bar{X} \not\Rightarrow X_n \xrightarrow{a.s} \bar{X}$$

Convergence of Random Variables

Example. Let $X_n = \frac{1}{n} U$ where U is a standard uniform random variable. Does X_n converge to a random variable, and in case to which one?

We can see that X_n converges almost surely to 0. Indeed, take any $\omega \in \Omega$ and set $u = U(\omega)$. As U is standard uniform, its values are in $[0, 1]$, so that u is a number between 0 and 1.

If $u = 0$, we have $X_n(\omega) = \frac{1}{n} U(\omega) = 0$ for all n, so in this case $X_n(\omega) \to 0$.

If $0 < u \leq 1$, we have $X_n(\omega) = \frac{1}{n} U(\omega) = \frac{1}{n} u$ for all n, so in this case $X_n(\omega) \to 0$ again since $\lim_n u/n = 0$ for all finite u.

So we conclude by saying that for all ω we have $X_n(\omega) \to 0$ and we conclude that we have almost sure convergence. This implies convergence in probability and convergence in law.

Convergence of Random Variables

We can check that we have convergence in L^p too for any integer $p \geq 1$, and convergence in mean square in particular. To check this, we need to compute

$$\mathbb{E}[|X_n - 0|^p] = \mathbb{E}[\frac{1}{n^p} U^p] = \frac{1}{n^p}\mathbb{E}[U^p]$$

For a standard uniform U we know that $E[U^p] = \frac{1}{p+1}$ so that

$$\mathbb{E}[|X_n - 0|^p] = \frac{1}{(p+1)n^p}$$

and it is easy to see that this tends to 0 as $n \uparrow \infty$, so that we have convergence in L^p.

Convergence of Random Variables

We now give examples showing that reverse implications do not hold.

Convergence in law does not imply convergence in probability.
Take $X \sim \mathcal{N}(0,1)$, and set $X_n = -X$ for all n. Given symmetry, $X_n \sim \mathcal{N}(0,1)$ for all n. We thus have $F_{X_n} = F_X = F_{N(0,1)}$ for all n and trivially $X_n \xrightarrow{law} X$. However, we can see that there is no convergence in probability.
Indeed, $\mathbb{P}\{|X_n - X| > \epsilon\} = \mathbb{P}\{|-2X| > \epsilon\} = \mathbb{P}\{|\mathcal{N}(0,1)| > \epsilon/2\} > 0$ also in the limit where $n \uparrow +\infty$, so we cannot have convergence in probability.

Convergence of Random Variables

Convergence in probability does not imply convergence in mean square. Take $X_n = \sqrt{n} 1_{\{U<1/n\}}$ where U is a standard uniform in $[0,1]$. Let's show that $X_n \xrightarrow{\mathbb{P}} 0$. Calculate

$$\mathbb{P}\{|X_n - 0| > \epsilon\} = \mathbb{P}\{\sqrt{n} 1_{\{U<1/n\}} > \epsilon\} = \mathbb{P}\{0 \leq U < 1/n\} = \frac{1}{n}$$

and this tends to 0 when n goes to infinity, so we have convergence in probability to 0. However, for convergence in mean square, we have

$$\mathbb{E}[|X_n - 0|^2] = \mathbb{E}[n 1_{\{U<1/n\}}] = n\mathbb{E}[1_{\{U<1/n\}}] = n\,\mathbb{P}\{U < 1/n\} = n\,\frac{1}{n} = 1$$

and this does not converge to 0 when $n \uparrow +\infty$, so we don't have convergence in mean square.

Putting time in the picture

We are going to define now a stochastic process, which is, roughly speaking, a family of random variables indexed by time.

In a sense we have already seen a stochastic process. In a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ the index n can be seen as a discrete time index.

In this sense X_1 would be the stochastic process at time 1, X_2 would be the process at time 2, X_n the process at time n, etc.

However, while in Econometrics and often in Statistics one uses discrete time stochastic processes, in Mathematical Finance we use continuous time processes, where time is not in a discrete set like $\mathbb{N}$ but in a continuous set like $\mathbb{R}$.

Continous time stochastic processes

Definition (Stochastic process)

Is a COLLECTION OF RANDOM VARIABLES X_t INDEXED by $t \in \mathbb{R}^{\geq 0}$

$$\{X_t \text{ random variable in } (\Omega, \mathcal{F}), t \in \mathbb{R}, t \geq 0\}$$

satisfying minimal consistency conditions (see e.g. Kolmogorov construction and the notion of separability).

Definition (Filtration $\{\mathcal{F}_t\}_{t\geq 0}$)

A filtration in $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing family of sub-σ-fields of $\mathcal{F}$:

$$\mathcal{F}_t \subseteq \mathcal{F},\ \mathcal{F}_t \subseteq \mathcal{F}_{t+h} \ \text{ for } t \geq 0,\ \ h \geq 0$$

Ideally, $\mathcal{F}_t$ models the events that are known at time t.

Filtrations and information flow

A filtration in $(\Omega, \mathcal{F}, \mathbb{P})$ is meant to model the flow of information, which increases over time. $\mathcal{F}_t$ models the information we know at time t.

Filtrations can be assigned a priori or can be generated by a stochastic process. For example, take the stochastic process X_t above taking values in $\mathbb{R}$ (if we had $\mathbb{R}^n$ then B would be the generic Borel set of $\mathbb{R}^n$). We can define the sigma-field generated by the stochastic process up to time t as

$$\mathcal{F}_t^X := \sigma\left(\{X_s^{-1}(B) : \ B \text{ Borel set in } \mathbb{R}, \ s \le t\}\right).$$

We see that when we know $\mathcal{F}_t^X$ we know everything about the process X up to time t. This is called the natural filtration of the process X and is often denoted, for brevity, by

$$\mathcal{F}_t^X = \sigma(\{X_s, \ s \le t\}).$$

Brownian Motion or Wiener Process W_t in $(\Omega, \mathcal{F}, \mathbb{P})$

Definition (W_t – Brownian motion)

The Brownian Motion (BM) in $(\Omega, \mathcal{F}, \mathbb{P})$ is a stochastic process which satisfies the following conditions:

- $W_0 = 0$;
- has CONTINUOUS PATHS $t \mapsto W_t(\omega)$;
- has INDEPENDENT INCREMENTS under $\mathbb{P}$, i.e. for all $s < t < u$, $W_u - W_t$ independent of $W_t - W_s$;
- has STATIONARY INCREMENTS under $\mathbb{P}$, i.e. distribution of $W_{t+h} - W_t$ does NOT depend on t, but only on h, for $h > 0$;
- W is a GAUSSIAN PROCESS with distribution $W_t \sim \mathcal{N}(0, t)$ and $W_t - W_s \sim \mathcal{N}(0, t - s)$ for all $t > 0, s > 0, t > s$ under $\mathbb{P}$.

Brownian Motion or Wiener Process W_t I

Brownian motion is very important, because it is the source of randomness in Stochastic Differential Equations. It is the random engine in the equation. Please familiarize yourself very well with the definition above. Now we will look at the meaning of these properties. We will not be fully rigorous but will reason in a roughly intuitive way.

First of all, note that any definition of Brownian motion is related to a probability measure. Because when we say independent increments, stationary increments, normal distributed increments, all these conditions in the definition are related to the probability space we are in, $(\Omega, \mathcal{F}, \mathbb{P})$ in our original definition. Indeed, take for example independence of increments. Two random variables (increments) that are independent under a probability measure P might not be so under a different probability measure Q. Independence of increment $W_u - W_t$ and increment $W_t - W_s$ $(s < t < u)$ means

Brownian Motion or Wiener Process W_t II

$P(W_u - W_t \in A \cap W_t - W_s \in B) = P(W_u - W_t \in A)P(W_t - W_s \in B)$ for any two Borel sets A and B, and as you see this depends on P. If we change measure to a different probability $\mathbb{Q}$, properties of W like independent increments or Gaussian law or stationary increments might not hold and W would not be a Brownian motion under $\mathbb{Q}$.

Next, I call your attention to two properties that are somewhat at odds.

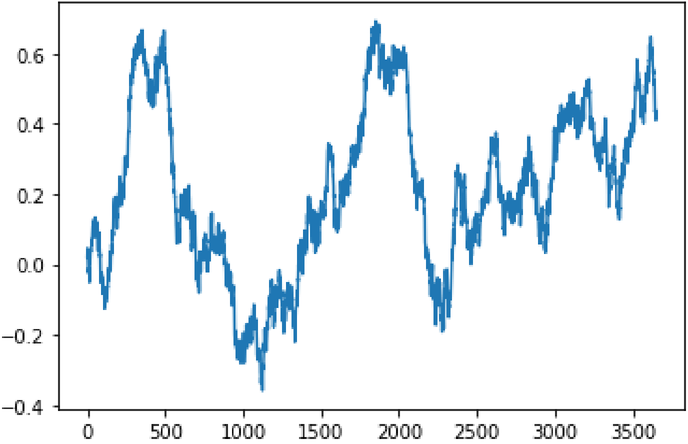
1. Paths are **continuous functions of time**, $t \mapsto W_t(\omega)$ is continous in t for almost every $\omega \in \Omega$. Continuity, intuitively, is associated with being somewhat foreseeable, because if the path is continuous it does not jump and cannot surprise you entirely, so it is in some way "**foreseeable**".

Brownian Motion or Wiener Process W_t III

2. The increments are **independent.** So if we have the history of Brownian motion up to time t, namely the path $s \mapsto W_s(\omega)$ for all times $s \leq t$, the next step $dW_t = W_{t+dt} - W_t$ is independent of the path up to t. This means that W_{t+dt} can take any value compared to the previous path up to t, and as such one would think it to be somewhat **"unforeseeable"**.

It turns out that it is possible for the two above properties to co-exist but the consequence is that Brownian motion paths have **unbounded variation.** In intuitive imprecise language, this means the paths swing a lot and change direction all the time, zig-zagging extremely. A consequence is that the paths are nowhere differentiable.

Brownian Motion or Wiener Process W_t IV



One Brownian path, $t \mapsto W_t(\omega)$, for a given $\omega \in \Omega$. t is on the x axis, while W_t is on the y axis. Note the unbounded variation features.

Brownian Motion or Wiener Process W_t V

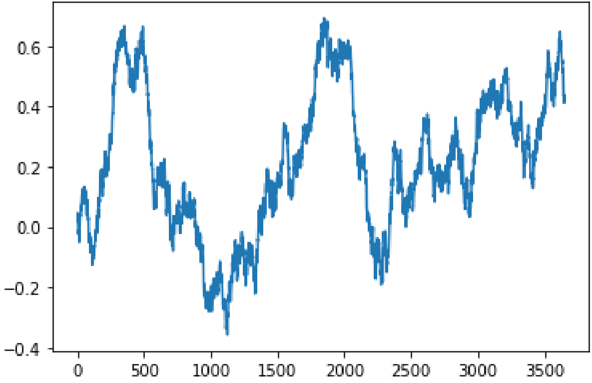
In a precise mathematical sense, unbounded variation means that if we take a finite interval $[0, T]$ and partitions $\{t_0 = 0, t_1, t_2, \ldots, t_i, t_{i+1}, \ldots, t_n = T\}$ with mesh size tending to zero as n grows to infinity, we get that Brownian motion has infinite variation with probability one, namely

$$\mathbb{P}\left\{\omega \in \Omega : \sup_{mesh\downarrow 0 \ as \ n\uparrow+\infty} \sum_{j=0}^{n-1} \left| W_{t_{j+1}}(\omega) - W_{t_j}(\omega) \right| = +\infty \right\} = 1$$

Think of any regular function you are used to, they all have bounded variation in close bounded intervals $[0, T]$ where they are continuous, and the above sup is usually finite. An exponential e^t, a logarithm $\ln(t+1)$, a power function t^n, etc. They all have finite variation.

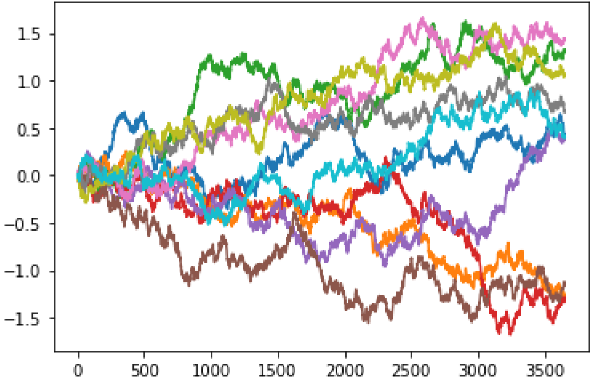
Simulation of Brownian Motion

To get some further intuition on Brownian motion, let us code a simulation of Brownian motion paths up to 1 year, with a time step of 1/10 of a day, or equivalently 1/3650 years.
The simulation of a single path is the following plot we saw earlier.



Simulation of Brownian Motion

Note the extremely swinging and zig-zagging nature of the signal. This is unbounded variation. Notice also the lack of differentiability everywhere. This path is not smooth, it is rough. We can also show ten different paths of Brownian motion, each with a different colour:



Simulation of Brownian Motion

In the appendix we give the python code that we used for simulating the Brownian motion. In the code we also check that the final standard deviation is 1 and the final mean is 0, plus skewness and excess kurtosis being 0, as should be, given that the final W_{1y} is Gaussian. To get these statistics more or less right you need to increase the number of simulations to $n = 10000$ or more. Recall

$$W_t \sim \mathcal{N}(0, t) \;\Rightarrow\; E[W_t] = 0; \;\; \mathrm{Std}(W_t) = \sqrt{t}.$$

So in particular $E[W_{1y}] = 0; \;\; \mathrm{Std}(W_1) = \sqrt{1} = 1.$

Finally (not required for exam), note that the distribution of Brownian motion W_T being normal for all T is to be expected. We can write Brownian motion W_T as a telescopic sum of its increments on finer and finer partitions. As increments are i.i.d., the central limit theorem would apply as we increase the number of increments by making the partition finer, leading to a Gaussian W_T in the limit.

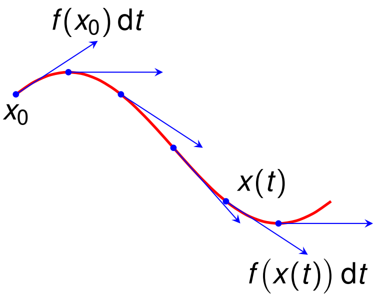
Ordinary Differential Equation (ODE)

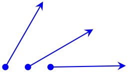
An Ordinary Differential Equation (ODE) for a deterministic signal $X(t)$ reads:

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = f\big(X(t)\big), \quad X(0) = x_0.$$

We can write it in differential form:

$$\mathrm{d}X(t) = f\big(X(t)\big)\,\mathrm{d}t \quad \text{INTERPRETATION: } X(t+\mathrm{d}t) = X(t) + f\big(X(t)\big)\,\mathrm{d}t$$



Solving an ODE: given  and x_0, find .

The solution to an ODE exists and it is unique under some conditions on $f(\cdot)$ (Lipschitz continuity, linear growth).

1st order ODEs: example 1

Consider the ODE

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = \mu X(t), \quad X(0) = x_0.$$

Here μ is a real constant. To solve this equation we separate variables:

$$\frac{\mathrm{d}X}{X} = \mu \mathrm{d}t, \quad X(0) = x_0.$$

Integrate both sides

$$\int_{x_0}^{X(T)} \frac{\mathrm{d}X}{X} = \int_0^T \mu \mathrm{d}t, \quad X(0) = x_0.$$

We get $(\ln X)|_{x_0}^{X(T)} = \mu T$ from which $\ln(X(T)/x_0) = \mu T$ or

$$X(T) = x_0 \exp(\mu T)$$

for all $T \geq 0$. We will show a numerical example of this equation below.

1st order ODEs: example 2

It will be useful to recall the solution for the linear (affine) differential equation

$$\frac{dX(t)}{dt} = B(t) - A(t)X(t),$$

where A and B are functions of time. The solution is given in any textbook as

$$X(t) = \exp\left(-\int_0^t A(s)ds\right)\left[\int_0^t \exp\left(\int_0^u A(s)ds\right)B(u)du + X(0)\right]$$

This will be useful later to solve the Ornstein-Uhlenbek SDE.

1st order ODEs: Autonomous and non-autonomous

A 1st order ODE is called autonomous (or non-autonomous, respectively) if it is of the form

$$\frac{dX(t)}{dt} = f(X(t)) \quad \left(\frac{dX(t)}{dt} = f(t, X(t)) \right).$$

So for example the ODE

$$\frac{dX(t)}{dt} = \mu \, X(t)$$

is autonomous, whereas the ODE

$$\frac{dX(t)}{dt} = B(t) - A(t)X(t)$$

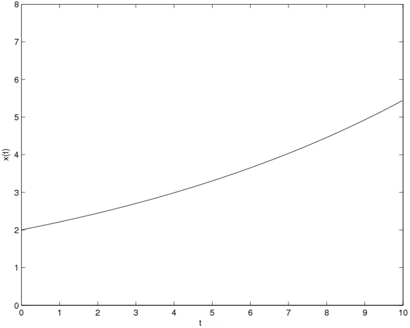
is non-autonomous.
The same nomenclature will be applied to SDEs

Example

Suppose we have the following toy model for population growth

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = 0.1X(t), \quad X(0) = 2.$$

This ODE tells us that the instantaneous change in population size at a given time is 1/10 of the population size at that same time.
Solving this ODE amounts to finding the whole time evolution of $X(t)$, given only its instantaneous change in all possible points and its initial position $X(0)$.

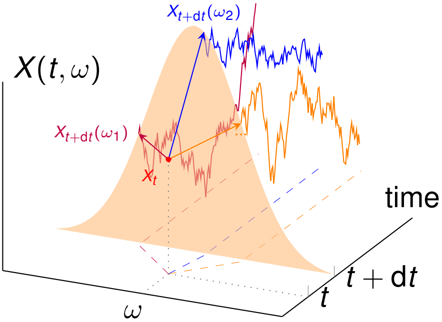


Solution, $X(t) = 2e^{0.1t}$, up to 10 years.

Stochastic Differential Equation (SDE)

A SDE is a generalization of an ODE where RANDOMNESS is added to the system:

$$dX_t = \underbrace{f(t, X_t)\,dt}_{\text{local mean}} + \underbrace{\sigma(t, X_t)}_{\substack{\text{local standard}\\ \text{deviation}}} \quad \underbrace{dW_t}_{\substack{\text{BM increments:}\\ \text{stationary, independent}\\ \text{from the past and } \sim\mathcal{N}(0,dt)}}$$



Can we make it rigorous? Big problem with the paths:

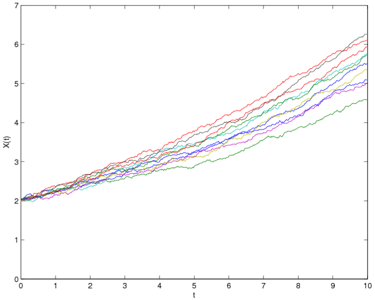
- UNBOUNDED VARIATION;
- NOWHERE DIFFERENTIABLE with probability 1.

We'll come back to this important problem later. Now we present some examples to develop intuition

Example

For a 2D visualization, let's add a random component to our toy model for population growth

$$dX(t) = 0.1X(t)\,dt + \sigma\,dW_t, \quad X(0) = 2.$$



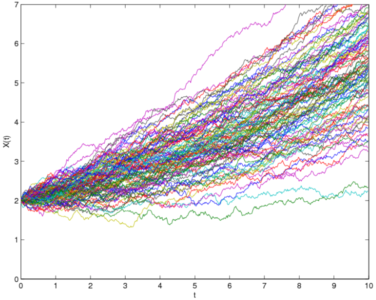


Figure: on the left 10 paths $\sigma = 0.1$ and on the right 100 paths for $\sigma = 0.2$.

Further intuition behind SDEs

Let us look at another example, for example a stock price.

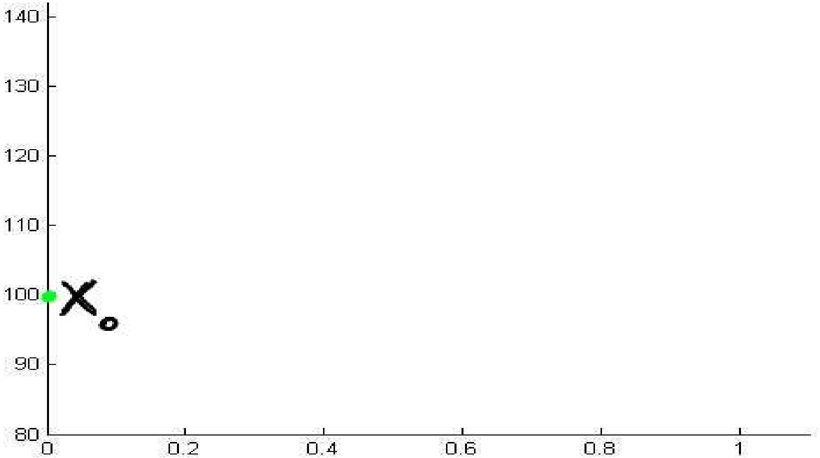
This will be a stochastic process X_t described by a stochastic differential equation.

Let us suppose this is the future price of an asset with return 5% and see how this varies with σ.
We will take $\sigma = 0.1 = 10\%$, but we will also plot the case $\sigma = 0.04 = 4\%$.

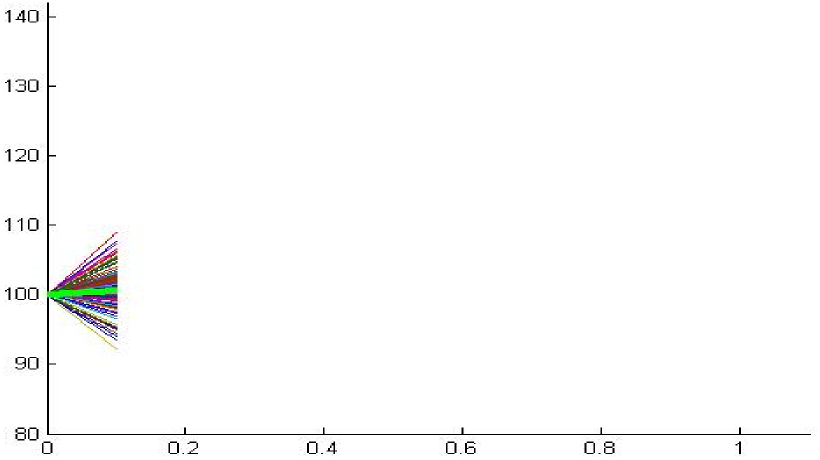
The initial value of the SDE is $X_0 = 100$.

$$dX_t = 0.05 X_t dt + 0.1 X_t dW_t, \; X_0 = 100.$$

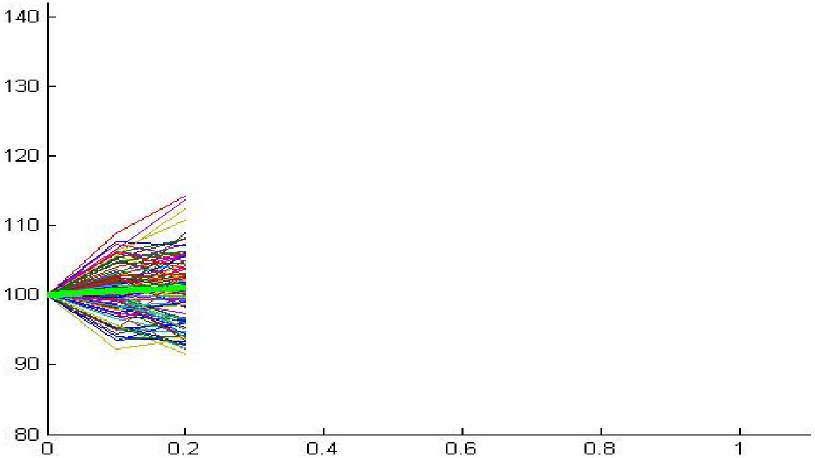
SDE: $dX_t = 0.05 X_t dt + 0.1 X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05 X_t dt$



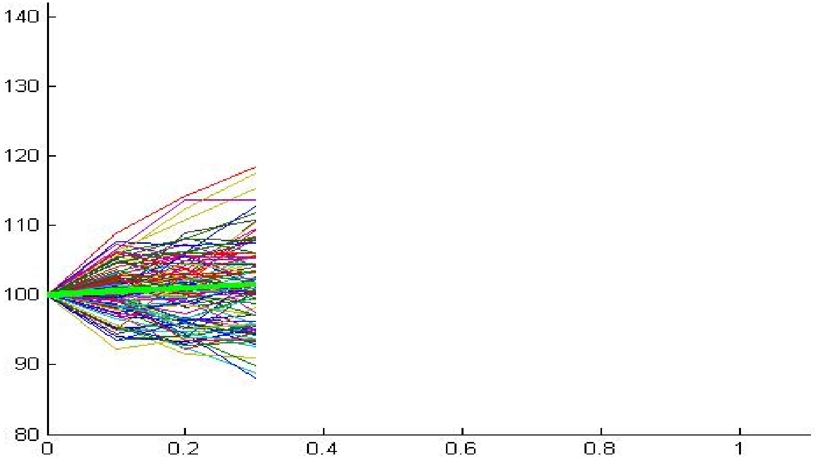
SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



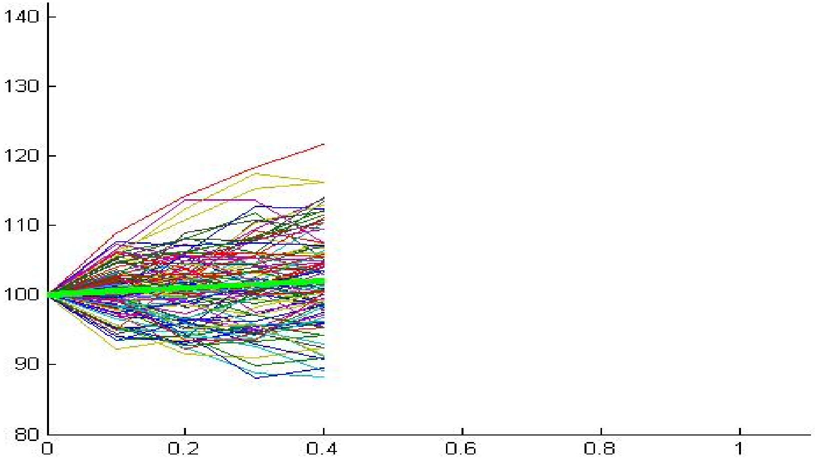
SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



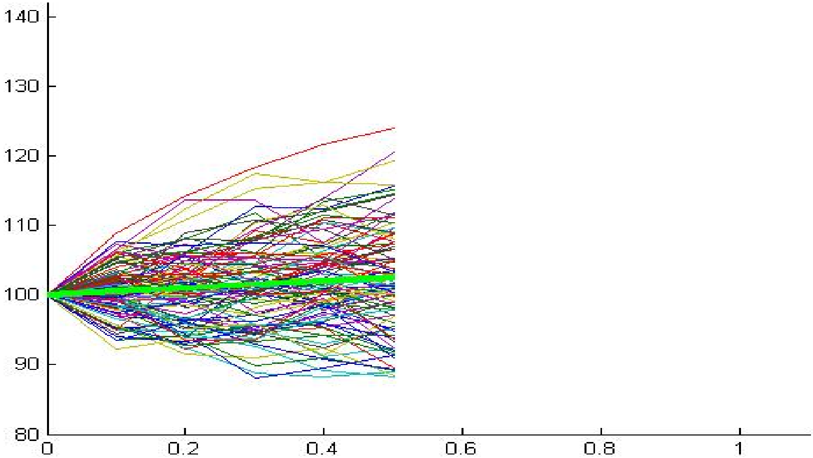
SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



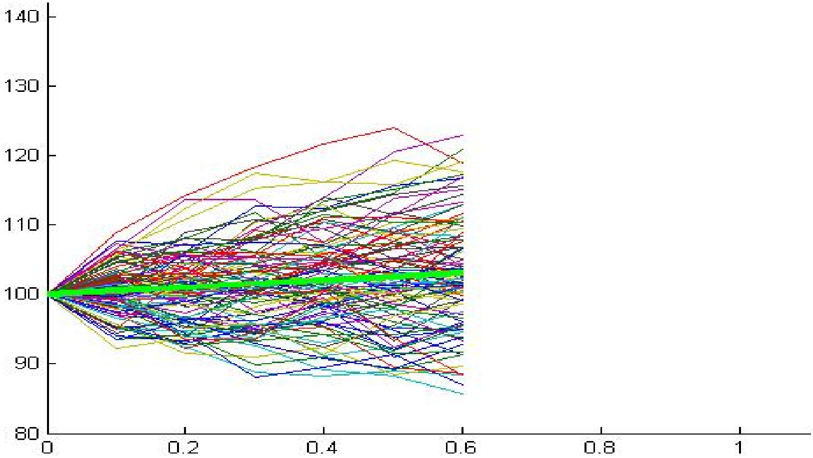
SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



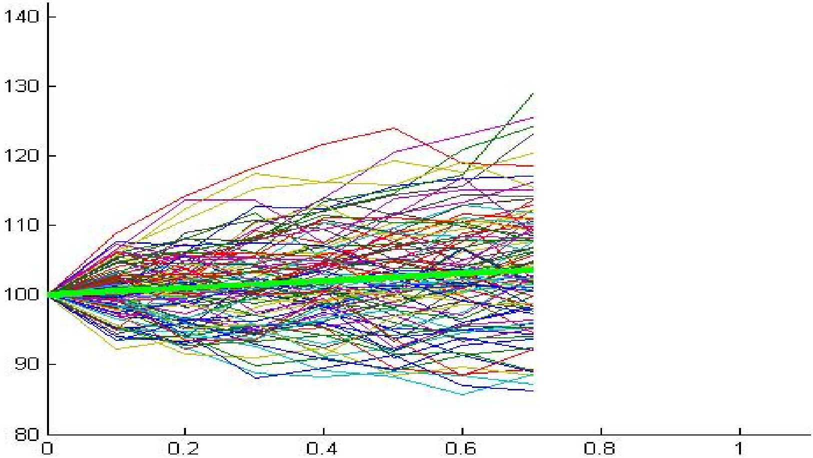
SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



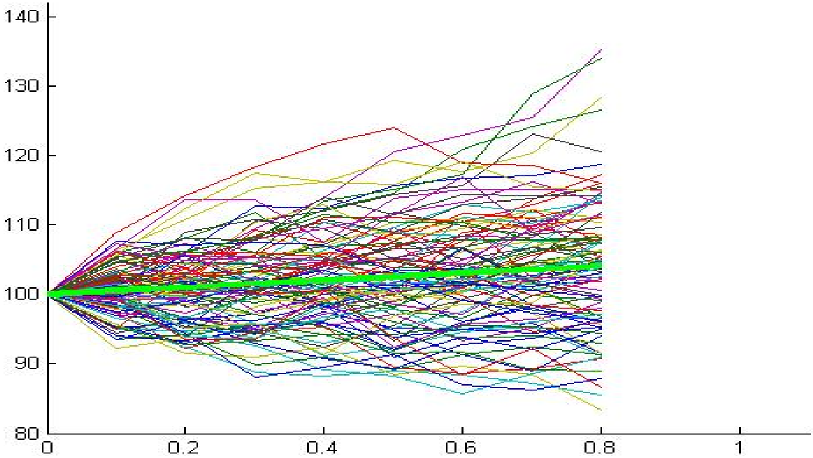
SDE: $dX_t = 0.05 X_t dt + 0.1 X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05 X_t dt$



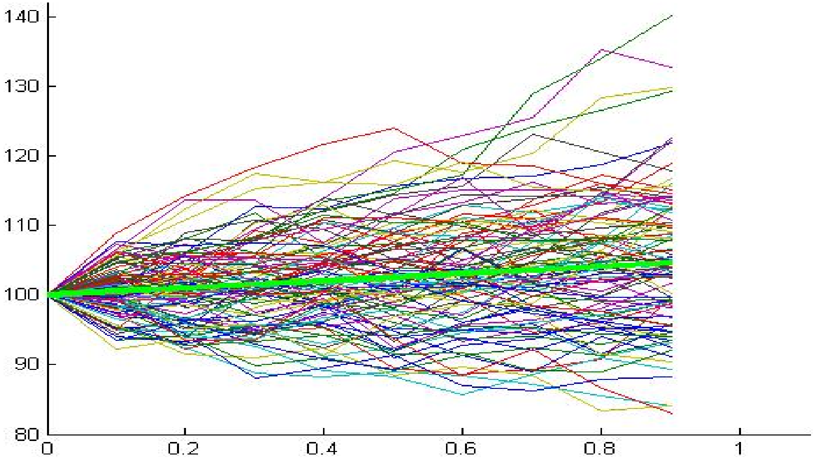
SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



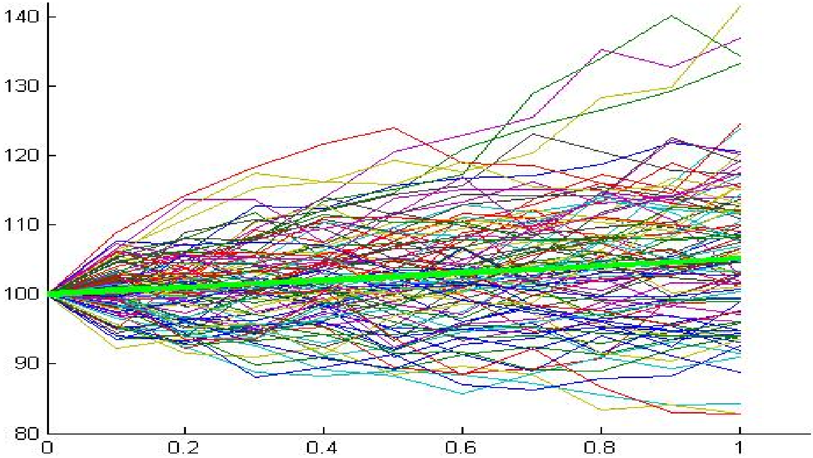
SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



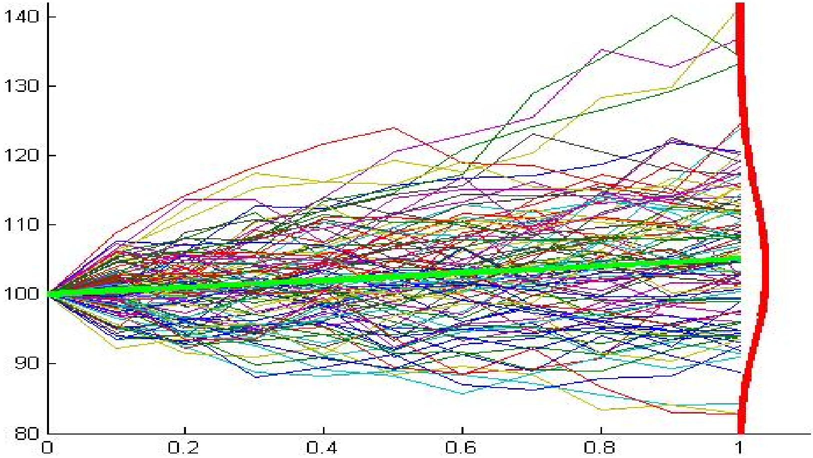
SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$



SDE: $dX_t = 0.05X_t dt + 0.1X_t dW_t$, $X_0 = 100$, ODE $dX_t = 0.05X_t dt$. Randomness, Dynamics & Prob



Probability density function

Here we plotted the probability density function (in red) at time $t = 1 = 1y$, $x \mapsto p_{X_{1y}}(x)$, where the horizontal axis is actually vertical.

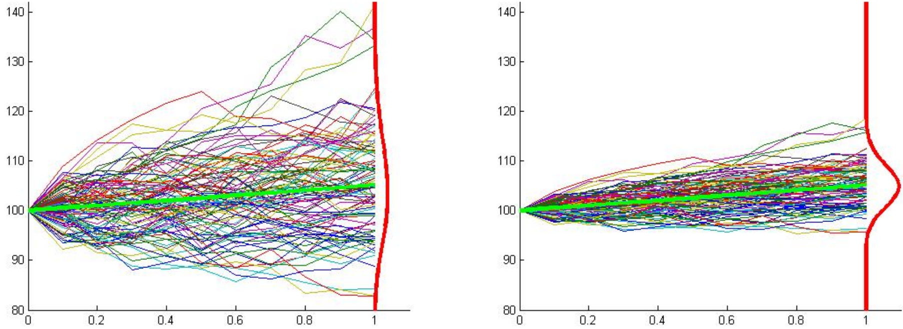
You can see that where the final coloured scenarios are more concentrated, the red density curve is higher (goes more to the right hand side), see for example values around 100-110.

Where the colored scenarios are sparse, the density curve is smaller, like for examples in values near 130-140 or 80-85. .

In the next picture we compare the case $\sigma = 0.1$ with $\sigma = 0.04$. We will see that the larger sigma, the more randomness the system has, and the more spread the paths and the final density will be.

$dX_t = 0.05X_t dt + \sigma X_t dW_t$, $X_0 = 100$:

$\sigma = 0.1$ vs $\sigma = 0.04$



SDEs: Mathematical problems

We now go back to the mathematical definition of an SDE.

Our problem is that in defining a SDE we write dW_t but it turns out that the paths $t \mapsto W_t(\omega)$, while being continous, have *unbounded variation* for almost every $\omega \in \Omega$ (with probability one). A consequence is that Brownian motion is nowhere differentiable:

$$\mathbb{P}\left\{\omega \in \Omega : \ \frac{dW_t(\omega)}{dt} \text{ does not exist for any } t\right\} = 1$$

If $\frac{\mathrm{d}W_t}{\mathrm{d}t}$ is not well defined, we cannot interpret it as a differential or a time derivative. Then what does $\mathrm{d}W_t$ in our SDE really mean?
We need to define $\mathrm{d}X_t = f(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t$ as an INTEGRAL EQUATION

$$X_t = x_0 + \int_0^t f(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}W_s$$

SDEs: Definition as integral equations

Thus, definining an SDE amounts to define a stochastic integral. However, since W has unbounded variation, we cannot define $\int_0^t \sigma(s, X_s)\,\mathrm{d}W_s$ as a Stieltjes integral on the paths. Fixing a path $\omega \in \Omega$, the integral

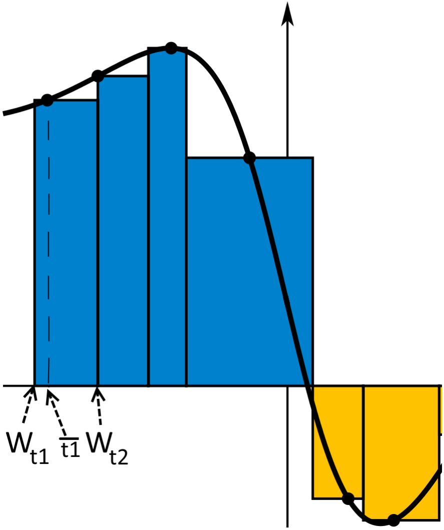
$$\int_0^t \sigma(s, X_s(\omega))\,\mathrm{d}W_s(\omega)$$

does not exist as a Riemann Stiltjes integral in t.

To simplify exposition, we focus on the autonomous case $\sigma(t, x) = \sigma(x)$ (and $f(t, x) = f(x)$) but this generalizes straightforwardly.

If W were differentiable, we could write a standard limit of Stiltjes sums to define a Stiltjes type integral. Let $t_0^n = 0, t_1^n, \ldots, t_n^n = T$ be an increasing partition of $[0, T]$ for all n, that grows finer as n increses. The partitions are assumed to be nested for different n. Suppose that the mesh size of the partition tends to 0 as $n \uparrow +\infty$.

The stochastic integral as a Stiltjes integral?



In a Stiltjes integral one has $\int_0^T \sigma(X_s)dW_s =$

$$= \lim_n \sum_{i=1}^{n} \sigma(X(\bar{t}_i))(W_{t_{i+1}} - W_{t_i})$$

for ANY choice $\bar{t}_i \in [t_i, t_{i+1})$.

However, for Brownian motion this does not work since W has unbounded variation.

Add an extra specification: *we need to explicitly decide which point $\bar{t}_i$ is considered.*

SDEs: Definition as integral equations

In a standard Stiltjes integral one has that the following limit converges

$$\int_0^T \sigma(X_s) dW_s = \lim_n \sum_{i=1}^{n} \sigma(X(\bar{t}_i))(W_{t_{i+1}} - W_{t_i})$$

for ANY possible choice of $\bar{t}_i \in [t_i, t_{i+1})$.

However, for Brownian motion this does not work since W has unbounded variation and is not differentiable.

It turns out that one can still define the stochastic integral in a Riemann Stiltjes way adding an extra specification that is not needed for ordinary Stiltjes integrals. *We need to explicitly decide at which point $\bar{t}_i$ in each limit interval $[t_i, t_{i+1})$ the integrand $\sigma(X_t)$ is evaluated.*

Two types of stochastic integrals: Ito and Stratonovich

- 2 main definitions of stochastic integrals: Initial point vs mid point

$$\int_0^T \sigma(X_s)dW_s = \lim_n \sum_{i=1}^n \sigma(X(t_i))(W_{t_{i+1}} - W_{t_i}) \quad \text{(Itô)}$$

$$\int_0^T \sigma(X_s)\circ dW_s = \lim_n \sum_{i=1}^n \sigma\left(X\left(\frac{t_i + t_{i+1}}{2}\right)\right)(W_{t_{i+1}} - W_{t_i})\text{(Stratonovich)}$$

- (a more general definition for Stratonovich would be

$$\int_0^T \sigma(X_s)\circ dW_s = \lim_n \sum_{i=1}^n \frac{\sigma(X(t_i)) + \sigma(X(t_{i+1}))}{2}(W_{t_{i+1}} - W_{t_i}) \quad)$$

where it is understood that as n tends to infinity the mesh size of the partition $\{[0,t_1),[t_1,t_2),\ldots,[t_{n-1},t_n = T]\}$ of $[0,T]$ tends to 0.

- Stratonovich integral looks into the future, Ito does not.

Two types of stochastic integrals: Ito and Stratonovich

- We note immediately that if $\sigma(X_t)$ does not depend on X and is a constant σ, then the Ito and Stratonovich integrals coincide. This also holds if $\sigma = \sigma(t)$ is a deterministic function of time. In this case the integral is called a *Wiener integral*.
- What do we mean by "look into the future"? In terms like $\sigma(X(t_i))(W_{t_{i+1}} - W_{t_i})$ the *future* increment is $W_{t_{i+1}} - W_{t_i}$ and we interpret t_i as the present time. In the Ito integral, $\sigma(X)$ is evaluated at the present time t_i, whereas the Stratonovich integral $\sigma\left(X\left(\frac{t_i+t_{i+1}}{2}\right)\right)(W_{t_{i+1}} - W_{t_i})$ is evaluated at a future time $\frac{t_i+t_{i+1}}{2}$ which is after the present time t_i.
- In finance we cannot know the future, so we use Ito.
- Stochastic integrals are limits in probability, and sometimes, under more strict assumptions on the integrand, they can also be defined as limits in mean square or almost sure. This means that in the above definitions what we meant really was that

- limit in $\mathbb{P}$: $\bar{t}_i = t_i$ for Ito and $\bar{t}_i = \frac{t_i + t_{i+1}}{2}$ for Stratonovich: for all $\epsilon > 0$ we have

$$\lim_{mesh \downarrow 0 \ as \ n \uparrow +\infty} \mathbb{P}\left\{ \left| Integral - \sum_{i=1}^{n} \sigma\left(X\left(\bar{t}_i\right)\right)\left(W_{t_{i+1}} - W_{t_i}\right) \right| > \epsilon \right\} = 0$$

- limit in mean square: $\bar{t}_i = t_i$ for Ito and $\bar{t}_i = \frac{t_i + t_{i+1}}{2}$ for Stratonovich:

$$\lim_{mesh \downarrow 0 \ as \ n \uparrow +\infty} \mathbb{E}\left[\left| Integral - \sum_{i=1}^{n} \sigma\left(X\left(\bar{t}_i\right)\right)\left(W_{t_{i+1}} - W_{t_i}\right) \right|^2 \right] = 0$$

- limit a.s. : $\bar{t}_i = t_i$ for Ito and $\bar{t}_i = \frac{t_i + t_{i+1}}{2}$ for Stratonovich:

$$\mathbb{P}\left\{ \omega : \lim_{mesh \downarrow 0 \ as \ n \uparrow +\infty} \sum_{i=1}^{n} \sigma\left(X(\bar{t}_i)(\omega)\right)\left(W_{t_{i+1}}(\omega) - W_{t_i})(\omega\right) = Integral(\omega) \right\} = 1$$

Once the integral is defined, the SDE notation is shorthand for an integral equation, and it can be either in Itô or Stratonovich form:

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t \quad \stackrel{Ito}{\Longleftrightarrow} \quad X_t = X_0 + \int_0^t f(X_s)ds + \int_0^t \sigma(X_s)dW_s$$

$$dX_t = f(X_t)dt + \sigma(X_t) \circ dW_t \quad \stackrel{Str}{\Longleftrightarrow} \quad X_t = X_0 + \int_0^t f(X_s)ds + \int_0^t \sigma(X_s) \circ dW_s$$

In this course we will use Itô's (without going into more detail) because it does not look into the future. Two properties to keep in mind, coming from $\bar{t}_i = t_i$ and by independence of W increments, so that $W_{t_{i+1}} - W_{t_i}$ is independent of $\sigma(X_{t_i})$, are:

$$\mathbb{E}\left[\int_0^t \sigma(X_s)\,\mathrm{d}W_s\right] = 0 \qquad \text{Itô's zero mean property}$$

$$\mathbb{E}\left[\left(\int_0^t \sigma(X_s)\,\mathrm{d}W_s\right)^2\right] = \int_0^t \mathbb{E}[\sigma(X_s)^2]\,\mathrm{d}s \qquad \text{Itô's isometry}$$

Ito integral has zero mean

We give an heuristic argument to show some intuition on why the Ito (but not the Stratonovich) integral has zero mean. Consider first the Riemann-Stieltjes sums of which the Ito integral is the limit (say in Probability).

$$\int_0^T \sigma(X_s)dW_s = \lim_{mesh\downarrow 0 \ as \ n\uparrow +\infty} \sum_{i=1}^{n} \sigma\left(X\left(t_i\right)\right)(W_{t_{i+1}} - W_{t_i})$$

where the limit is the limit of convergence in Probability. On the right hand side, take an expected value:

$$\mathbb{E}[\sum_{i=1}^{n} \sigma\left(X\left(t_i\right)\right)(W_{t_{i+1}} - W_{t_i})] = \sum_{i=1}^{n} \mathbb{E}[\sigma\left(X\left(t_i\right)\right)(W_{t_{i+1}} - W_{t_i})] =$$

now keep in mind that X_{t_i} is independent of the future Brownian increment $W_{t_{i+1}} - W_{t_i}$.

Ito integral has zero mean

This is because all the randomness that affects X_{t_i} is coming from all the process X history up to time t_i. This randomness was generated by the only sources of randomness entering the sde for X, or its integral version, namely increments ΔW_t for $t < t_i$ and possibly X_0 if random, but X_0 is assumed independent of W. This means that all the randomness in X_{t_i} is coming from past Brownian increments ΔW and, since Brownian increments are independent, the next increment $W_{t_{i+1}} - W_{t_i}$ is independent of (i) all past increments ΔW up to t_i and (ii) of X_0, and thus of X_{t_i} whose randomness is entirely driven by (i) and (ii). Given independence, we can factor the expectation

$$= \sum_{i=1}^{n} \mathbb{E}[\sigma(X(t_i))]\mathbb{E}[(W_{t_{i+1}} - W_{t_i})] = 0$$

because the expectation of the Brownian increment is zero by definition of Brownian motion. Given that the mean is zero for every partition, it remains zero when we move to the limit for mesh size tending to zero.

Ito integral has zero mean

Note that the same does not hold for the Stratonovich integral, because $X_{\frac{t_i+t_{i+1}}{2}}$ and $W_{t_{i+1}} - W_{t_i}$ are not independent, given that X is evaluated at a time after the increment starts.

Finally, we will not give an argument for the Ito isometry, which we give without proof.

We just mention that all the above definitions and discussions, and in particular the zero mean property of the Ito integral and the Ito isometry, also hold in the non-autonomous case when $f(x)$ is replaced with $f(t,x)$ and $\sigma(x)$ with $\sigma(t,x)$.

Existence and uniqueness of solutions

Consider the Ito SDE $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = Z$

where the random initial condition Z is independent of $\sigma(\{W_t, \ t \le T\})$ and $\mathbb{E}[Z^2] < +\infty$. The functions $\mu : [0, T] \times \mathbb{R} \to \mathbb{R}$ (the drift) and $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ (the diffusion coefficient) are assumed measurable. Assume μ and σ satisfy global Lipschitz continuity

$$|\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| \le K|x - y| \text{ for all } t \in [0, T]$$

and all $x, y \in \mathbb{R}$, and linear growth

$$|\mu(t, x)| + |\sigma(t, x)| \le K'(1 + |x|)| \text{ for all } t \in [0, T] \text{ and all } x \in \mathbb{R}$$

for two constants K, K'. Then for the SDE above there exists a unique global solution X_t for all $t \in [0, T]$ that is $(\mathcal{F}_t^{W,Z})_t$ adapted and continous in t, satsifying $\mathbb{E}[\int_0^T X_t^2 dt] < +\infty$.

Existence and uniqueness of solutions

A stochastic process $X_t(\omega)$ is adapted to a filtration $\mathcal{F}_t$ if X_t is $\mathcal{F}_t$ mesurable for all t. This means that the information in $\mathcal{F}_t$ includes the value of X_t. By construction X_t is $\mathcal{F}_t^X$ adapted and the theorem says it is also $\mathcal{F}_t^{W,Z}$ adapted, meaning that the value of X_t is known if we know this information: we know Z and the history of W up to time t.

The above conditions for existence and uniqueness are very strong sufficient (but not necessary) conditions. Existence of local solutions can be obtained by requiring local versions of these conditions. Also, for some SDEs that do not satisfy Lipschitz, e.g. the square root process $dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$, the Yamada-Watanabe theorem (that we won't cover) works as an alternative.

Finally, note that for autonomous SDEs, where the SDE drift and diffusion coefficient are $\mu(t,x) = \mu(x)$ and $\sigma(t,x) = \sigma(x)$, not depending on t, then global Lipschitz continuity implies linear growth, so one only has to check for Lipschitz.

Existence and uniqueness of solutions

If you have an autonomous SDE just check Lipschitz continuity but *do mention* that this implies linear growth too. To show that Lipschitz $x \mapsto f(x)$ implies linear growth f, note that, from the triangular inequality
$|f(x)| = |f(x) - f(0) + f(0)| \leq |f(x) - f(0)| + |f(0)|$ so that
$|f(x) - f(0)| \geq |f(x)| - |f(0)|$ and then

$$|f(x)| - |f(0)| \leq |f(x) - f(0)| \leq K|x - 0| \Rightarrow |f(x)| - |f(0)| \leq K|x| \Rightarrow$$

$$\Rightarrow |f(x)| \leq K|x| + |f(0)| \Rightarrow |f(x)| \leq \max(K, f(0))(1 + |x|)$$

which is a linear growth condition.
This holds for $f(x)$ but does not hold for $f(t, x)$ that is Lipschitz in x. If in doubt, verify both conditions anyway to be safe.
We will not delve further into existence theorems here, but it is good to have at least the above example of theorem, with global Lipschitz continuity and linear growth.

Existence and uniqueness of solutions

From the theory of deterministic differential equations we can see why these two conditions are needed. Let's look at two examples.

Growth more than linear (no existence). Take the ODE

$$\frac{dX_t}{dt} = KX_t^2, \quad X_0 = 1$$

for positive constant K and integrate it by separation of variables,

$$\frac{dX}{X^2} = Kdt$$

and integrate both sides. One gets

$$X_t = \frac{1}{1 - K\ t}.$$

In this case the solutions is not global because it is only defined for $t < 1/K$. Hence with quadratic growth we have no global existence of a solution for $t \in [0, T]$.

Existence and uniqueness of solutions

Lack of Lipschitz continuity (no uniqueness). Consider the ODE

$$\frac{dX_t}{dt} = 3X_t^{2/3}, \ X_0 = 0.$$

Integrating again by separation of variables we get

$$\int_0^{X_t} X^{-2/3} dX = 3t, \ \ 3(X_t)^{1/3} = 3t \ \ X_t = t^3.$$

However, in dividing we have assumed $X \neq 0$. What if $X = 0$ in some interval? Certainly $X_t = 0$ is a solution too. So there are at least two solutions, $X_t = t^3$ and $X_t = 0$. However, we also have all the other solutions

$$X_t = 0 \ 1_{\{t \leq a\}} + (t-a)^3 \ 1_{\{t > a\}}$$

for any positive a less then T. Thus without Lipschitz continuity we have lack of uniqueness of solutions.

Diffusion processes

The solution of a SDE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = Z$$

is sometimes called a *diffusion process*, or shortly diffusion. This is also why the term $\sigma(t, X_t)$ is called the diffusion coefficient of the SDE. In practice, as we have seen in the example of SDE simulation for the stock price with the coloured scenarios, the Brownian motion presence allows the paths to *diffuse* in space, according to some probability law which is locally Gaussian as dW_t is normally distributed.

Itô's formula

So, now $\mathrm{d}X_t = f(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t$ has a meaning as an integral equation. Given this equation, can we find $\mathrm{d}\varphi(t, X_t)$ where $\varphi(\cdot, x)$ is a smooth function?

Chain Rule

If X(t) is differentiable, then

$$\mathrm{d}\varphi(t, X_t) = \frac{\partial \varphi}{\partial t}\,\mathrm{d}t + \frac{\partial \varphi}{\partial x}\,\mathrm{d}X_t$$

However, for SDE's with Itô's integrals, this is modified as follows:

Itô's formula

$$\mathrm{d}\varphi(t, X_t) = \frac{\partial \varphi}{\partial t}\,\mathrm{d}t + \frac{\partial \varphi}{\partial X}\,\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2 \varphi}{\partial X^2}\mathrm{d}X_t\,\mathrm{d}X_t$$

A more rigorous expression for $\mathrm{d}X_t\,\mathrm{d}X_t$ is the QUADRATIC VARIATION $\mathrm{d}\langle X \rangle_t$. Here we present an informal account.

Quadratic variation of Brownian motion

When we write $dW_t dW_t = dt$ we actually mean to express a limit. A more rigorous expression for this is $d\langle W\rangle_t = dt$, or $\langle W\rangle_t = t$. The quadratic variation may be defined for example as a limit in mean square. Given nested partitions $t_0^n, t_1^n, \ldots, t_n^n$ with $t_0^n = 0$ and $t_n^n = t$ as we took in the definition of stochastic integrals, we get quadratic variation $\langle W\rangle_t$ defined as the mean square limit

$$\lim_{mesh\downarrow 0 \text{ as } n\uparrow\infty} \mathbb{E}\left[\left(\sum_{i=0}^{n-1}(W_{t_{i+1}^n} - W_{t_i^n})^2 - \langle W\rangle_t\right)^2\right] = 0,$$

or $X_n = \sum_{i=0}^{n-1}(W_{t_{i+1}^n} - W_{t_i^n})^2 \xrightarrow{mean\ square} \langle W\rangle_t$.
One can also show that with nested partitions $t_0^n, t_1^n, \ldots, t_n^n$,
$X_n = \sum_{i=0}^{n-1}(W_{t_{i+1}^n} - W_{t_i^n})^2 \xrightarrow{a.\ s.} \langle W\rangle_t$.
For Brownian motion $\boxed{\langle W\rangle_t = t, \ \text{ or } \ d\langle W\rangle_t = dt}$

Quadratic variation of Brownian motion

Let's take this formulation, based on nested partitions $X_n = \sum_{i=0}^{n-1}(W_{t_{i+1}^n} - W_{t_i^n})^2 \xrightarrow{a.\ s.} \langle W \rangle_t$. This means that

$$\mathbb{P}\left\{ \lim_{mesh \downarrow 0 \ as \ n \uparrow +\infty} \sum_{j=0}^{n-1} \left| W_{t_{j+1}^n} - W_{t_j^n} \right|^2 = t \right\} = 1$$

So, there is probability one that the quadratic variation of Brownian motion is t. Recall that the 1-variation of Brownian motion is infinite with probability one:

$$\mathbb{P}\left\{ \sup_{mesh \downarrow 0 \ as \ n \uparrow +\infty} \sum_{j=0}^{n-1} \left| W_{t_{j+1}^n} - W_{t_j^n} \right| = +\infty \right\} = 1$$

Roughly, if we sum all the $|dW_{t_i}|$ over a finite time interval $[0, t]$, in the sup limit we get infinity. But if we sum the squares, $|dW_{t_i}|^2$, in the limit we get a finite number t called quadratic variation.

Stratonovich and the Chain rule

For the Stratonovich version $dX_t = f(t, X_t)\,dt + \sigma(t, X_t) \circ dW_t$ has a meaning as a Stratonovich integral equation. Given this equation, can we find $d\varphi(t, X_t)$ where $\varphi(\cdot, x)$ is a smooth function?

For SDE's with Stratonovich integrals, the chain rule still holds:

$$d\varphi(t, X_t) = \frac{\partial \varphi}{\partial t}\,dt + \frac{\partial \varphi}{\partial X} \circ dX_t$$

More generally, one can still use standard calculus, contrary to the Itô case. However, it is no longer true that the Stochastic integral has zero mean, and its probabilistic properties are not so good.

Stratonovich is good for geometry and for convergence of processes with regular noise when the noise converges to Brownian motion (Wong Zakai).

Ito-Stratonovich transformation

Given a SDE in Itô form, it is possible to write a Stratonovich SDE with the same solution, and vice versa. This is known as Itô - Stratonovich transformation. The following two SDEs

$$dX_t = f(t, X_t)dt + \sigma(t, X_t)dW_t \quad \rightarrow \quad dX_t = \tilde{f}(t, X_t)dt + \sigma(t, X_t) \circ dW_t$$

$$\tilde{f} = f - \frac{1}{2}\sigma\frac{\partial \sigma}{\partial x}$$

have the same solution X. Note that when the diffusion coefficient is deterministic or constant, then the Ito and Stratonovich SDEs for X coincide (and have the same solution). Let σ be a deterministic constant $\sigma(t, x) = \sigma$ or a deterministic function of time only $\sigma(t, x) = \sigma(t)$: then $\tilde{f} = f$.

In this course and in Mathematical Finance in general one uses Itô calculus because of the non-anticipative (not looking into the future) property and of the good probabilistic properties (zero mean).

Itô's formula

$$d\varphi(t, X_t) = \frac{\partial \varphi}{\partial t}\, dt + \frac{\partial \varphi}{\partial X}\, dX_t + \frac{1}{2}\frac{\partial^2 \varphi}{\partial X^2} dX_t\, dX_t$$

The three key rules to remember in Itô's formula are

$$\boxed{dt\, dt = 0, \quad dt\, dW_t = 0, \quad dW_t\, dW_t = dt}$$

Notice that, if V_t is differentiable, then

$$dV_t\, dV_t = V_t'\, dt\; V_t'\, dt = (V_t')^2 \underbrace{dt\, dt}_{0} = 0$$

This does not happen with W_t because it is NOT differentiable. To remember Itô's formula one should keep in mind the three key rules and a 2nd order Taylor expansion.

Example: Arithmetic to Geometric Brownian Motion

Consider $dX_t = \mu_t\, dt + \sigma_t\, dW_t$ (ARITHMETIC BM), with μ_t and σ_t DETERMINISTIC functions of time.

Set $Y_t = e^{X_t} =: \varphi(X_t)$ and find the SDE for Y_t.

$$\boxed{\frac{\partial\varphi}{\partial t} = 0 \qquad \frac{\partial\varphi}{\partial X} = e^{X_t} \qquad \frac{\partial^2\varphi}{\partial X^2} = e^{X_t}}$$

$$\begin{aligned}
d\varphi(X_t) &= \frac{\partial\varphi}{\partial t}\, dt + \frac{\partial\varphi}{\partial X}\, dX_t + \frac{1}{2}\frac{\partial^2\varphi}{\partial X^2}\, dX_t\, dX_t \\
&= 0\, dt + e^{X_t}\, dX_t + \frac{1}{2} e^{X_t}\, dX_t\, dX_t \\
&= e^{X_t}(\mu_t\, dt + \sigma_t\, dW_t) + \frac{1}{2} e^{X_t}(\mu_t\, dt + \sigma_t\, dW_t)^2 \\
&= e^{X_t}(\mu_t\, dt + \sigma_t\, dW_t) + \frac{1}{2} e^{X_t}(\mu_t^2 \underbrace{dt\, dt}_{0} + \sigma_t^2 \underbrace{dW_t\, dW_t}_{dt} + 2\mu_t\sigma_t \underbrace{dt\, dW_t}_{0})
\end{aligned}$$

$$dY_t = \left(\mu_t + \frac{1}{2}\sigma_t^2\right) Y_t\, dt + \sigma_t Y_t\, dW_t \quad \text{GEOMETRIC BM}$$

Example: Arithmetic to Geometric Brownian Motion I

It can be important to derive the mean and variance of the ABM X_t, and to memorize this formula. We assume X_0 to be deterministic. From $dX_t = \mu_t dt + \sigma_t dW_t$, integrating both sides from 0 to T we get

$$X_T = X_0 + \int_0^T \mu_t dt + \int_0^T \sigma_t dW_t. \tag{1}$$

We give an intuitive argument why the Wiener integral $\int_0^T \sigma_t dW_t$ is Gaussian. This integral is the limit of sums of the type

$$\lim_n \sum_{i=1}^n \sigma_{t_i}(W_{t_{i+1}} - W_{t_i})$$

but now recall that each Brownian increment $W_{t_{i+1}} - W_{t_i}$ is Gaussian, and even rescaling for a *deterministic* quantity σ_{t_i} it remains Gaussian.

Example: Arithmetic to Geometric Brownian Motion II

Furthermore, as Brownian increments are independent, all those terms in the sum are independent Gaussian variables. The sum of independent Gaussians is Gaussians and this carries on to the limit. So the resulting Wiener integral is Gaussian.

We already saw intuition on why the stochastic integral has zero mean (each Brownian increment has zero mean), and the Ito isometry in this case simplifies into $E[(\int_0^T \sigma_t dW_t)^2] = \int_0^T \sigma_t^2 dt$ as σ_t is deterministic and no expectation is needed.

Going back to Eq (1), as everything else apart from the Wiener integral we just examined is deterministic, it follows that it is added to the mean but does not change the variance:

$$X_T \sim \mathcal{N}\left(X_0 + \int_0^T \mu_t dt, \int_0^T \sigma_t^2 dt\right)$$

Example: Arithmetic to Geometric Brownian Motion III

from which we see that

$$E[X_T] = X_0 + \int_0^T \mu_t dt, \quad Var[X_T] = \int_0^T \sigma_t^2 dt.$$

In the case where μ and σ are constant, we get

$$X_T = X_0 + \mu \; T + \sigma \; W_T,$$

$$X_T \sim \mathcal{N}\left(X_0 + \mu T, \sigma^2 T\right)$$

$$E[X_T] = X_0 + \mu T, \quad Var[X_T] = \sigma^2 \; T.$$

Example: Geometric to Arithmetic Brownian Motion

Now, consider $dY_t = m_t Y_t\,dt + \nu_t Y_t\,dW_t$ (GEOMETRIC BM), with m_t and ν_t DETERMINISTIC functions of time.

Set $Z_t = \ln(Y_t) =: \varphi(Y_t)$ and find the SDE for Z_t.

$$\boxed{\frac{\partial\varphi}{\partial t} = 0 \qquad \frac{\partial\varphi}{\partial Y} = \frac{1}{Y_t} \qquad \frac{\partial^2\varphi}{\partial Y^2} = -\frac{1}{Y_t^2}}$$

$$\begin{aligned}
d\varphi(Y_t) &= \frac{\partial\varphi}{\partial t}\,dt + \frac{\partial\varphi}{\partial Y}\,dY_t + \frac{1}{2}\frac{\partial^2\varphi}{\partial Y^2}\,dY_t\,dY_t \\
&= 0\,dt + \frac{1}{Y_t}\,dY_t - \frac{1}{2Y_t^2}\,dY_t\,dY_t \\
&= \frac{1}{Y_t}(m_t Y_t\,dt + \nu_t Y_t\,dW_t) - \frac{1}{2Y_t^2}(m_t^2 Y_t^2 \underbrace{dt\,dt}_{0} + \nu_t^2 Y_t^2 \underbrace{dW_t\,dW_t}_{dt} + \\
&\qquad + 2m_t\nu_t Y_t \underbrace{dt\,dW_t}_{0})
\end{aligned}$$

$$dZ_t = \left(m_t - \frac{1}{2}\nu_t^2\right)dt + \nu_t\,dW_t \quad \text{ARITHMETIC BM}$$

Example: Geometric Brownian Motion I

We can use this to actually solve the SDE for geometric Brownian motion. Indeed, we saw that $Y_t = \exp(Z_t)$ and Z can be integrated directly by integrating both sides between 0 and T:

$$Z_T - Z_0 = \int_0^T \left(m_t - \frac{1}{2}\nu_t^2\right)\mathrm{d}t + \int_0^T \nu_t \,\mathrm{d}W_t.$$

Remembering the previous discussion, we know the stochastic (Wiener) integral on the right is Gaussian with mean 0 and Variance $\int_0^T \nu_t^2 \, dt$. Assuming from now on Z_0 is deterministic, given that everything else is deterministic, we have

$$Z_T \sim \mathcal{N}\left(Z_0 + \int_0^T \left(m_t - \frac{1}{2}\nu_t^2\right)\mathrm{d}t, \int_0^T \nu_t^2 \, dt\right) =: \mathcal{N}(Z_0 + M_T, V_T^2)$$

Example: Geometric Brownian Motion II

Now, given that $Y_T = e^{Z_T}$, bringing Z_0 to the right and given $Y_0 = e^{Z_0}$

$$Y_T = Y_0 \exp\left(\int_0^T \left(m_t - \frac{1}{2}\nu_t^2\right)\mathrm{d}t + \int_0^T \nu_t \,\mathrm{d}W_t\right) = Y_0 e^{\mathcal{N}(M_T, V_T^2)}. \quad (2)$$

It can be helpful to memorize the formulae for the mean and the variance of a GBM like $dY_t = m_t Y_t dt + \nu_t Y_t dW_t$. Let's compute the mean of Y_T using Eq. (2).

$$E[Y_T] = E[Y_0 e^{\mathcal{N}(M_T, V_T^2)}] = Y_0 E[e^{\mathcal{N}(M_T, V_T^2)}]$$

and now we can use the moment generating function of a Gaussian computed at point 1, giving

$$E[Y_T] = Y_0 e^{M_T + \frac{1}{2}V_T^2} = Y_0 \exp\left(\int_0^T m_t \, dt\right).$$

Example: Geometric Brownian Motion III

Thus *the expectation at time* 0 *of a GBM with drift rate* m_t *and deterministic initial condition* Y_0 *is the initial condition times the exponential of the integral of the drift rate* m*. This is an important formula to remember.*

For the variance, we can compute

$$Var(Y_T) = E[Y_T^2] - E[Y_T]^2 = E[\left(Y_0 e^{\mathcal{N}(M_T, V_T^2)}\right)^2] - \left(Y_0 \exp\left(\int_0^T m_t \, dt\right)\right)^2$$

$$= E[Y_0^2 e^{2\mathcal{N}(M_T, V_T^2)}] - Y_0^2 \exp\left(2 \int_0^T m_t \, dt\right)$$

Example: Geometric Brownian Motion IV

Now the first expected value is the moment generating function of a normal computed on the point 2 so we get

$$Var(Y_T) = Y_0^2 e^{2M_T + (4/2)V_T^2} - Y_0^2 \exp\left(2\int_0^T m_t \, dt\right)$$

or

$$Var(Y_T) = Y_0^2 \exp\left(\int_0^T 2m_t dt\right)\left(\exp\left(\int_0^T \nu_t^2 dt\right) - 1\right).$$

In the special case where m and ν are constants we get

$$Y_T = Y_0 \exp\left(\left(m - \frac{1}{2}\nu^2\right) T + \nu W_T\right)$$

and

$$E[Y_T] = Y_0 e^{mT}, \quad Var(Y_T) = Y_0^2 e^{2mT}\left(e^{\nu^2 T} - 1\right).$$

Example: Ornstein-Uhlenbeck process

Consider $\mathrm{d}X_t = (b_t - a_t X_t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t$ (ORNSTEIN-UHLENBECK process), with b, a and σ DETERMINISTIC functions of time (called Vasicek model in interest rates).

As σ is not a function of X, we know that this same equation holds in Stratonovich form and we can write

$$\mathrm{d}X_t = (b_t - a_t X_t)\,\mathrm{d}t + \sigma_t \circ \mathrm{d}W_t$$

Now, with Stratonovich we know that the formal rules of standard calculus still hold. This means we can treat dW as if it were differentiable in solving the Stratonovich SDE. Write it like this

$$\mathrm{d}X_t = (b_t + \sigma_t \circ \frac{\text{“}\mathrm{d}W_t\text{''}}{dt} - a_t X_t)\,\mathrm{d}t$$

Example: Ornstein-Uhlenbeck process

$$\frac{\text{“}dX_t\text{''}}{dt} = b_t + \sigma_t \circ \frac{\text{“}dW_t\text{''}}{dt} - a_t X_t$$

Call $A(t) = a(t)$ and $B(t) = b_t + \sigma_t \circ \frac{\text{“}dW_t\text{''}}{dt}$, and recall the solution of the linear-affine ODE we have seen earlier. Substituting A and B in

$$X(t) = e^{-\int_0^t A(s)ds}\left[\int_0^t \exp\left(\int_0^u A(s)ds\right) B(u)du + X(0)\right] \text{ we get}$$

$$X(t) = e^{-\int_0^t a(s)ds}\left[\int_0^t \exp\left(\int_0^u a(s)ds\right)\left(b_u + \sigma_u \circ \frac{\text{“}dW_u\text{''}}{du}\right) du + X(0)\right]$$

Re-writing the “derivative” of W as a differential we get

$$X(t) = e^{-\int_0^t a(s)ds}\left[\int_0^t \exp\left(\int_0^u a(s)ds\right)(b_u\, du + \sigma_u\, dW_u) + X(0)\right]$$

where we also replaced the Stratonovich integral with an Ito one, since they are the same here (again σ does not depend on X).

Example: Ornstein-Uhlenbeck process

$$X(t) = e^{-\int_0^t a(s)ds}\left[\int_0^t \exp\left(\int_0^u a(s)ds\right)(b_u\, du + \sigma_u\, dW_u) + X(0)\right]$$

We note that the only random term in the solution is

$$e^{-\int_0^t a(s)ds}\left[\int_0^t \exp\left(\int_0^u a(s)ds\right)\sigma_u\, dW_u\right]$$

This is a Wiener integral, as the integrand $\exp\left(\int_0^u a(s)ds\right)\sigma_u$ is a deterministic function of time. As seen in the previous examples, these integrals have a Gaussian distribution. Intuitively, this happens because all dW_u at different times u are independent and each is Gaussian. It follows that also all terms of the type $\alpha(u)dW_u$ (where α is a determinitic function of time) are Gaussian and independent of each other, so that adding them up gives a Gaussian distribution. It follows that $X(t)$ is Gaussian if the initial condition $X(0)$ is Gaussian and independent of W or deterministic.

Special case of Ornstein-Uhlenbeck: Vasicek model

A special case of Ornstein-Uhlenbeck process in mathematical finance is the Vasicek model for interest rates, $X_t = r_t$ where $r_t(\omega)$ is the stochastic process for the short term interest rate. In that case

$$dX_t = k(\theta - X_t)dt + \sigma dW_t, \quad x_0$$

where k, θ, σ are constant in time. We get then the special solution. We have $b(t) = k\theta$, $a(t) = k$ and $\sigma_t = \sigma$.

$$X(t) = e^{-kt}\left[\int_0^t \exp(ku)\,(k\theta\, du + \sigma\, dW_u) + X(0)\right]$$

Simplifying

$$X(t) = x_0 e^{-kt} + \theta(1 - e^{-kt}) + \sigma \int_0^t e^{-k(t-u)} dW_u$$

Special case of Ornstein-Uhlenbeck: Vasicek model

Given our previous discussion, we know that X_t will be Gaussian with

$$X(t) \sim \mathcal{N}\left(x_0 e^{-kt} + \theta(1-e^{-kt}), \frac{\sigma^2}{2k}\left[1-e^{-2kt}\right]\right)$$

where the variance is computed as $\mathrm{Var}[X_t] =$

$$= \mathrm{Var}\left[x_0 e^{-kt} + \theta(1-e^{-kt}) + \sigma\int_0^t e^{-k(t-u)} dW_u\right] = \mathrm{Var}\left[\sigma\int_0^t e^{-k(t-u)} dW_u\right] =$$

$$= E\left[\left(\sigma\int_0^t e^{-k(t-u)} dW_u\right)^2\right] - E\left[\sigma\int_0^t e^{-k(t-u)} dW_u\right]^2 =$$

$$= \sigma^2\int_0^t (e^{-k(t-u)})^2 du - 0^2 = \sigma^2\int_0^t e^{-2k(t-u)} du = \frac{\sigma^2}{2k}\left[1-e^{-2kt}\right]$$

where we used Ito's isometry and the fact that the Ito integral has zero mean (martingale).

Example: Square of Ornstein-Uhlenbeck process

Back to $dX_t = (b_t - a_t X_t)\,dt + \sigma_t\,dW_t$ (ORNSTEIN-UHLENBECK process), with b, a and σ DETERMINISTIC functions of time.

Set $Y_t = X_t^2 =: \varphi(X_t)$ and find the SDE for Y_t.

$$\boxed{\frac{\partial\varphi}{\partial t} = 0 \qquad \frac{\partial\varphi}{\partial X} = 2X_t \qquad \frac{\partial^2\varphi}{\partial X^2} = 2}$$

$$\begin{aligned}
d\varphi(X_t) &= \frac{\partial\varphi}{\partial t}\,dt + \frac{\partial\varphi}{\partial X}\,dX_t + \frac{1}{2}\frac{\partial^2\varphi}{\partial X^2}\,dX_t\,dX_t \\
&= 2X_t[(b_t - a_t X_t)\,dt + \sigma_t\,dW_t] + [(b_t - a_t X_t)^2 \underbrace{dt\,dt}_{0} + \sigma_t^2 \underbrace{dW_t\,dW_t}_{dt} + \\
&\qquad + 2(b_t - a_t X_t)\sigma_t \underbrace{dt\,dW_t}_{0}] \\
&= (2b_t X_t - 2a_t X_t^2 + \sigma_t^2)\,dt + 2\sigma_t X_t\,dW_t \qquad X_t = \pm\sqrt{Y_t} \quad \text{take pos. sol.}
\end{aligned}$$

$$dY_t = (\sigma_t^2 + 2b_t\sqrt{Y_t} - 2a_t Y_t)\,dt + 2\sigma_t\sqrt{Y_t}\,dW_t$$

If $b_t = 0$, this is a “square root process” (called CIR model in interest rates).

Square root process (CIR model)

With constant coefficients, a more general square root process is the following,

$$dy_t = \kappa(\mu - y_t)dt + \nu\sqrt{y_t}dW_t, \quad y_0$$

used in finance to model either stochastic interest rates, $y_t = r_t$ (Cox Ingersoll Ross model, CIR), or stochastic volatilities, $\sqrt{y_t} = v_t$ (Heston model). This is more general than the squared Ornstein Uhlenbeck model with $b_t = 0$ and constant parameters because we do not require that $\mu = \nu^2$. Here μ can be general.

The model can never go negative but in some cases it can hit zero. The Feller condition

$$2\kappa\mu > \nu^2$$

ensures that $y_t > 0$ and 0 is never hit.

Mean reverting properties of Vasicek & CIR I

The parameters of CIR have the same interpretation as the parameters of Vasicek:

$$\text{Vasicek model} \;\; dX_t = k(\theta - X_t)dt + \sigma dW_t, \;\; x_0.$$

$$\text{CIR model} \;\; dy_t = \kappa(\mu - y_t)dt + \nu\sqrt{y_t}dW_t, \;\; y_0.$$

k, κ: speed of mean reversion
θ, μ: long term mean reversion level
σ, ν: volatility parameter.
The formula for the mean of Vasicek and CIR is the same:

$$E[X_t] = x_0 e^{-kt} + \theta(1 - e^{-kt})$$

$$E[y_t] = y_0 e^{-\kappa t} + \mu(1 - e^{-\kappa t})$$

Mean reverting properties of Vasicek & CIR II

$$\lim_{t\uparrow+\infty} E[X_t] = \theta, \quad \lim_{t\uparrow+\infty} E[y_t] = \mu.$$

This is why θ and μ are called the long term means of the two models.

From the mean formulas above $E_0[X_t], E_0[y_t]$ it is clear that the parameters k and κ determine how quickly the mean converges to the long term means θ and μ.
One can also compute the variance of the processes. We have

$$\mathrm{Var}(X_t) = \frac{\sigma^2}{2k}\left[1 - e^{-2kt}\right],$$

$$\mathrm{Var}(y_t) = y_0\frac{\nu^2}{\kappa}(e^{-\kappa t} - e^{-2\kappa t}) + \mu\frac{\nu^2}{2\kappa}(1 - e^{-\kappa t})^2$$

Mean reverting properties of Vasicek & CIR III

We see that for $t \uparrow +\infty$ the variance has a finite limit.

$$\lim_{t\uparrow+\infty} \mathrm{Var}(X_t) = \frac{\sigma^2}{2k}, \quad \lim_{t\uparrow+\infty}, \quad \mathrm{Var}(y_t) = \mu\frac{\nu^2}{2k}.$$

After a long time the processes X and y reach (asymptotically) a stationary distribution (Gaussian for X, non-central chi squared for y) around their mean θ, μ and with a corridor of variance $\sigma^2/(2k)$, $\mu\nu^2/(2\kappa)$. Models like this are called “mean reverting”, because the mean reverts to a constant value over time and the variance remains finite, with a finite limit. Intuitively, when the shocks bring the processes away from their long term means θ and μ, the SDE drift $k(\theta - X_t)$ and $\kappa(\mu - y_t)$ will bring the process back towards the means θ and μ, with a variance that does not grow to infinity in time.

Mean reverting properties of Vasicek & CIR IV

The parameters k, κ are called “speed of mean reversion”, as they determine how quickly the means $E[X_t]$ and $E[y_t]$ tend to θ and μ. The larger the speed, the faster the convergence.

The parameters σ and ν are the local volatilties parameters of the two models. They determine how much randomness enters the system instant by instant. Note however that also the speeds k, κ impact the total variance of X and y, as you see from the formulas for $\mathrm{Var}(X_t)$ and $\mathrm{Var}(y_t)$.

The larger k, κ, the faster the processes converge to the stationary distributions. So, ceteris paribus, increasing k, κ reduces the total volatility and variance of X and y.

Mean reverting properties of Vasicek & CIR V

The larger θ, μ, the higher the long term means, so the model will tend to higher X and y in the future on average.

The larger σ, ν, the larger the instantaneous volatilities. Notice however that speeds k, κ fight instantaneous volatilities σ, ν as far as the influence on total volatility and variance is concerned.

Mean reverting properties of Vasicek & CIR

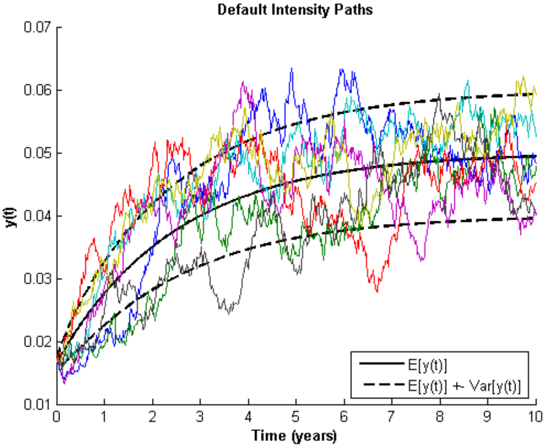


Figure: $y0 = 0.0165, \kappa = 0.4, \mu = 0.05, \nu = 0.04$

Example: Exponential of Ornstein-Uhlenbeck process

Exercise

If X_t is Ornstein-Uhlenbeck, find the SDE for $Z_t = e^{X_t} =: \varphi(X_t)$ (exponential Vasicek model).

We have $\frac{\partial \varphi}{\partial t} = 0$, $\frac{\partial \varphi}{\partial X} = e^X$, $\frac{\partial^2 \varphi}{\partial X^2} = e^X$.

$$\begin{aligned}
\mathrm{d}\varphi(X_t) &= \frac{\partial \varphi}{\partial t}\,\mathrm{d}t + \frac{\partial \varphi}{\partial X}\,\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2 \varphi}{\partial X^2}\,\mathrm{d}X_t\,\mathrm{d}X_t \\
&= 0dt + \; e^{X_t} dX_t + \frac{1}{2} e^{X_t} dX_t dX_t \\
&= Z_t[(b_t - a_t X_t)dt + \sigma_t dW_t] + \frac{1}{2} Z_t \sigma_t^2 dt
\end{aligned}$$

as we have seen in the previous slide that $dXdX = \sigma^2 dt$. We conclude

$$dZ_t = Z_t\left[b_t - a_t \ln Z_t + \frac{\sigma_t^2}{2}\right] dt + \sigma_t Z_t dW_t.$$

Product rule for SDEs, quadratic covariation & instantaneous correlation I

Given two processes X_t and Y_t, the differential of the quadratic covariation between the two processes, written informally $dX_t\, dY_t$ and more rigorously $d\langle X, Y\rangle_t$ is the differential of the quantity $\langle X, Y\rangle_t$ defined as the limit (for example in mean square) over nested partitions $t_0^n, t_1^n, \ldots, t_n^n$ of $[0, t]$ with $t_0^n = 0$ and $t_n^n = t$, as we took in the definition of stochastic integrals:

$$\lim_{mesh\downarrow 0 \text{ as } n\uparrow\infty} \mathbb{E}\left[\left(\sum_{i=0}^{n-1}(X_{t_{i+1}^n} - X_{t_i^n})(Y_{t_{i+1}^n} - Y_{t_i^n}) - \langle X, Y\rangle_t\right)^2\right] = 0,$$

or $Z_n = \sum_{i=0}^{n-1}(X_{t_{i+1}^n} - X_{t_i^n})(Y_{t_{i+1}^n} - Y_{t_i^n}) \xrightarrow{mean\ square} \langle X, Y\rangle_t$ in case this limit exists.

Product rule for SDEs, quadratic covariation & instantaneous correlation II

The quadratic variation of a process X is a special case of the quadratic covariation: $\langle X\rangle_t = \langle X, X\rangle_t$.

Now that we have defined the quadratic variation for a general process X, we can reformulate Ito's formula for the transformation of a SDE X more rigorously as

Itô's formula

$$\mathrm{d}\varphi(t, X_t) = \frac{\partial\varphi}{\partial t}\,\mathrm{d}t + \frac{\partial\varphi}{\partial X}\,\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2\varphi}{\partial X^2}d\langle X\rangle_t$$

Product rule for SDEs, quadratic covariation & instantaneous correlation III

Whenever we write expressions like

$$dt\, dt = 0, \;\; dt\, dW_t = 0, \;\; dW_t dW_t = dt, \;\; dW_t^{(1)} dW_t^{(2)} = \rho\, dt$$

this is an informal notation for the rigorous expressions

$$d\langle t \rangle_t = 0, \;\; d\langle t, W \rangle_t = 0, \;\; d\langle W \rangle_t = dt, \;\; d\langle W^{(1)}, W^{(2)} \rangle_t = \rho\, dt$$

or, even more precisely,

$$\langle t \rangle_t = 0, \;\; \langle t, W \rangle_t = 0, \;\; \langle W \rangle_t = t, \;\; \langle W^{(1)}, W^{(2)} \rangle_t = \rho\, t.$$

We are particularly interested in the last case, the quadratic covariation of two Brownians $W^{(1)}$ and $W^{(2)}$.

Product rule for SDEs, quadratic covariation & instantaneous correlation IV

While $d\langle W\rangle_t = dW_t dW_t = 1\ dt$, it may happen that two different Brownian motions have quadratic co-variation less than 1, or even 0. The differential of the quadratic covariation of two different Brownian motions is written informally $dW_t^{(1)} dW_t^{(2)}$ or more precisely $d\langle W^{(1)}, W^{(2)}\rangle_t$ and is defined as the mean square limit, given nested partitions $t_0^n, t_1^n, \ldots, t_n^n$ of $[0, t]$ with $t_0^n = 0$ and $t_n^n = t$ as we took in the definition of stochastic integrals:

$$\lim_{mesh\downarrow 0\ as\ n\uparrow\infty} \mathbb{E}\left[\left(\sum_{i=0}^{n-1}(W^{(1)}_{t^n_{i+1}} - W^{(1)}_{t^n_i})(W^{(2)}_{t^n_{i+1}} - W^{(2)}_{t^n_i}) - \langle W^{(1)}, W^{(2)}\rangle_t\right)^2\right] = 0,$$

or $Z_n = \sum_{i=0}^{n-1}(W^{(1)}_{t^n_{i+1}} - W^{(1)}_{t^n_i})(W^{(2)}_{t^n_{i+1}} - W^{(2)}_{t^n_i}) \xrightarrow{mean\ square} \langle W^{(1)}, W^{(2)}\rangle_t$.

Product rule for SDEs, quadratic covariation & instantaneous correlation V

With this definition in mind, it's quick to check that if W and $\widetilde{W}$ are independent Brownian motions, then

$$\langle W, \widetilde{W} \rangle_t = 0$$

or informally $dW_t d\widetilde{W}_t = 0$ or again $d\langle W, \widetilde{W} \rangle_t = 0$. Now, set

$$dW_t^{(1)} = dW_t, \;\; dW_t^{(2)} = \rho dW_t + \sqrt{1-\rho^2} d\widetilde{W}_t,$$

with $\rho \in [-1, 1]$ and where W and $\widetilde{W}$ are again two independent Brownian motions. It is easy to see that $W^{(1)}$ is a Brownian motion and so is $W^{(2)}$, but they are not independent unless $\rho = 0$.

Product rule for SDEs, quadratic covariation & instantaneous correlation VI

The quadratic co-variation is

$$dW_t^{(1)} dW_t^{(2)} = dW_t(\rho dW_t + \sqrt{1-\rho^2} d\widetilde{W}_t)$$

$$= \rho dW_t dW_t + \sqrt{1-\rho^2} dW_t d\widetilde{W}_t = \rho dt + \sqrt{1-\rho^2} 0 = \rho dt.$$

Hence we conclude $d\langle W^{(1)}, W^{(2)}\rangle_t = \rho \, dt$. This is also the general case of two correlated Brownian motions. We say that two Brownian motions have quadratic covariation $\rho \in [-1, 1]$ if $dW_t^{(1)} dW_t^{(2)} =$

$$= d\langle W^{(1)}, W^{(2)}\rangle_t = \rho \, dt \ \text{ i.e. } \sum_{i=0}^{n-1} (W^{(1)}_{t^n_{i+1}} - W^{(1)}_{t^n_i})(W^{(2)}_{t^n_{i+1}} - W^{(2)}_{t^n_i}) \xrightarrow{m.s.} \rho \, t.$$

Product rule for SDEs, quadratic covariation & instantaneous correlation VII

Informally, $\text{“}corr\text{''}(dW_t^{(1)}, dW_t^{(2)}) = \dfrac{dW_t^{(1)} dW_t^{(2)}}{\sqrt{(dW_t^{(1)} dW_t^{(1)})(dW_t^{(2)} dW_t^{(2)})}} = \rho.$

This is also true for SDEs driven by $W^{(1)}$ and $W^{(2)}$. If

$$dX_t^{(1)} = \mu_1(t, X_t^{(1)})dt + \sigma_1(t, X_t^{(1)})dW_t^{(1)}, \ x_0^{(1)}$$

$$dX_t^{(2)} = \mu_2(t, X_t^{(2)})dt + \sigma_2(t, X_t^{(2)})dW_t^{(2)}, \ x_0^{(2)}$$

$$\Longrightarrow \quad dX_t^{(1)} dX_t^{(2)} = d\langle X^{(1)}, X^{(2)}\rangle_t = \sigma_1(t, X_t^{(1)})\sigma_2(t, X_t^{(2)})\rho \, dt,$$

$$\text{“corr”}(dX_t^{(1)}, dX_t^{(2)}) = \frac{dX_t^{(1)} dX_t^{(2)}}{\sqrt{(dX_t^{(1)} dX_t^{(1)})(dX_t^{(2)} dX_t^{(2)})}} = \rho.$$

Product rule for SDEs

Suppose you have two SDEs describing two stochastic processes X and Y. The product rule reads

$$d(X_t Y_t) = Y_t\, dX_t + X_t\, dY_t + d\langle X,\ Y\rangle_t.$$

The last term, often written informally as $dX_t\ dY_t$, is again specific to Ito calculus and is missing in standard calculus.

The last term is computed with the usual rules
$d\langle t\rangle_t = dt\ dt = 0,\ \ d\langle t, W\rangle_t = dt\ dW_t = 0,$
$d\langle W\rangle_t = dW_t\ dW_t = dt,\ \ d\langle W^{(1)}, W^{(2)}\rangle_t = dW_t^{(1)}\ dW_t^{(2)} = \rho\ dt'.$

Note: This rule can be proven by applying Ito's formula to the function of two variables $\varphi(X, Y) = XY$ starting from the two-dimensional SDE for (X_t, Y_t). It is a standard application of Ito's formula to bivariate diffusions.

Martingale SDEs

A martingale is a stochastic process X_t such that $\mathbb{E}\{|X_t|\} < +\infty$ and

$$\mathbb{E}[X_t|\sigma(\{X_u,\ 0 \leq u \leq s\})] = X_s, \quad s \leq t.$$

This reads: the expected value of X_t given all the information on X from time 0 to time s is equal to X_s. On average, X stays constant. It is a process with zero local trend, so it does neither increase nor decrease locally, on average.
For a regular SDEs (unique strong solution exists, no explosion, finite second moment...) the solution of the SDE is a martingale if the SDE has zero drift.

$$dX_t = 0\ dt + \sigma(t, X_t)dW_t.$$

Martingale SDEs

Martingales can be defined on more general filtrations than $\sigma(\{X_u,\ 0 \le u \le s\})$.

Remember that the Ito stochastic integral is a martingale.

Martingales are used in finance to model “fair games” and no-arbitrage.

Martingale SDEs

As a first exercise, let's prove that the solution to the SDE with zero drift $dX_t = \sigma_t dW_t,\ x_0$ is a martingale. This is a special case of an arithmetic Brownian motion seen earlier (here $\mu_t = 0$). We can simply integrate both sides directly,

$$X_t = X_0 + \int_0^t \sigma_u \, dW_u.$$

Let's show that X satisfies the classic martingale definition, namely

$$\mathbb{E}[X_t|\mathcal{F}_s^X] = X_s, \quad s \leq t. \text{ We can write } X_t = X_s + \int_s^t \sigma_u \, dW_u$$

Take then

$$\mathbb{E}[X_t|\mathcal{F}_s^X] = \mathbb{E}[X_s + \int_s^t \sigma_u \, dW_u|\mathcal{F}_s^X] = X_s + \mathbb{E}[\int_s^t \sigma_u \, dW_u|\mathcal{F}_s^X] =$$

Martingale SDEs

Now note that all future increments dW_u of Brownian motion for $u > s$ are independent of $\mathcal{F}_s^X$, so that the conditioning can be removed,

$$= X_s + \mathbb{E}[\int_s^t \sigma_u \, dW_u] = X_s$$

because we know that the expectation of the Ito stochastic integral is zero (coming from the fact that each independent dW has expectation zero).

More generally, one can show that the Ito integral seen as a stochastic process $\{M_t = \int_0^t b(X_s(\omega))dW_s(\omega), \ t \geq 0\}$ is a martingale.

Martingale SDEs

As a second exercise, let's prove that the solution to the SDE $dY_t = \nu Y_t dW_t,\ y_0$ is a martingale. We know the solution of this equation from our previous example of GBM (here $m = 0$).

$$Y_t = y_0 \exp\left(-\frac{1}{2}\nu^2\ t + \nu W_t\right).$$

Let's show that Y satisfies the classic martingale definition, namely

$$\mathbb{E}[Y_t|\mathcal{F}_s^Y] = Y_s,\ \ s \le t.$$

We can write

$$Y_t = Y_s \frac{Y_t}{Y_s} = Y_s \exp\left(-\frac{1}{2}\nu^2\ (t-s) + \nu(W_t - W_s)\right)$$

Martingale SDEs

Then

$$\mathbb{E}[Y_t|\mathcal{F}_s^Y] = \mathbb{E}[Y_s \exp\left(-\frac{1}{2}\nu^2\,(t-s) + \nu(W_t - W_s)\right)|\mathcal{F}_s^Y] =$$

$$= Y_s \exp\left(-\frac{1}{2}\nu^2\,(t-s)\right)\mathbb{E}[\exp\left(\nu(W_t - W_s)\right)|\mathcal{F}_s^X] =$$

Now $W_t - W_s$ is indepedent of $\mathcal{F}_s^X$ due to the properties of Brownian motion, so that

$$= Y_s \exp\left(-\frac{1}{2}\nu^2\,(t-s)\right)\mathbb{E}[\exp(\nu(W_t - W_s))] =$$

and the expectation is $\mathbb{E}[\exp(\nu\mathcal{N}(0, t-s))]$ which is the moment generating function of a normal with 0 mean and variance $\nu^2(t-s)$, namely $\exp\left(\frac{1}{2}\nu^2\,(t-s)\right)$. Plugging this in we get

$$= Y_s \exp\left(-\frac{1}{2}\nu^2\,(t-s)\right)\exp\left(\frac{1}{2}\nu^2\,(t-s)\right) = Y_s$$

PDEs and Expectations: The Feynman-Kac theorem

We will use partial differential equations (PDEs) later. An important result allows to express the solution of a parabolic PDE as the expected value of a related SDE. Given suitable regularity and integrability conditions, the solution of the PDE with terminal condition f

$$\frac{\partial V}{\partial t}(t,x)+\frac{\partial V}{\partial x}(t,x)\boxed{\boxed{b(x)}}+\frac{1}{2}\frac{\partial^2 V}{\partial x^2}(t,x)\sigma^2(x)=\boxed{r}V(t,x),\quad V(T,x)=f(x) \tag{3}$$

can be expressed as the expected value

$$V(t,x)=e^{-\boxed{r}(T-t)}\mathbb{E}^Q[f(X_T)|\mathcal{F}_t^X,X_t=x] \tag{4}$$

where the diffusion process X has dynamics starting from x at time t

$$dX_s=\boxed{\boxed{b(X_s)}}ds+\sigma(X_s)dW_s^Q\;,s\geq t,\;X_t=x \tag{5}$$

under the probability measure $\mathbb{Q}$ under which the expectation $\mathbb{E}^Q[\cdot]$ is taken. The process W^Q is a standard Brownian motion under $\mathbb{Q}$.

Change of measure: Girsanov's theorem

In general, the statistics of a random variable X on $(\Omega, \mathcal{F})$ depend on the probability measure. Recall that $F_X(y) = \mathbb{P}(X \leq y)$. If we change probability measure $\mathbb{P}$ with a new measure $\mathbb{Q}$, while X is still the same random variable as a measurable map $X : \Omega \to \mathbb{R}$, its distribution F, and also the mean, variance, etc… will change.

In the theory of SDEs, we have seen that

$$dX_t = f^{\mathbb{P}}(X_t)\,dt + \sigma(X_t)\,dW_t^{\mathbb{P}},$$

where the local mean and the BM are all under the probability measure $\mathbb{P}$. Similarly to the case of random variables X, we can change the probability measure. The process X will remain the same but its statistics, and in particular the local mean, will change. The local standard deviation needs to be the same to apply the Girsanov theorem we are going to introduce, and for the two measures to be equivalent in particular.

Equivalent measures

Definition (Equivalent measures)

Two measures $\mathbb{P}$ and $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ are equivalent if they agree on which events have zero probability.

This implies the measures also agree on which events have probability 1, which is an equivalent definition.
$\mathbb{P}$ and $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ are equivalent when, given $A \in \mathcal{F}$

$$\mathbb{P}(A) = 0 \iff \mathbb{Q}(A) = 0 \text{ or equivalently } \mathbb{P}(A) = 1 \iff \mathbb{Q}(A) = 1.$$

In other words, the two measures agree on what is almost impossible (probability zero) or almost sure (probability one). The two definition are equivalent because

$$A \in \mathcal{F} \iff A^c \in \mathcal{F} \text{ and } \mathbb{P}(A) = 0 \iff \mathbb{P}(A^c) = 1.$$

Consider now the following problem: what happens to an SDE of a process X_t if we change $\mathbb{P}$ with an equivalent measure $\mathbb{Q}$?

Definition (Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$)

It is a random variable such that, for any other random variable X

$$\mathbb{E}^{\mathbb{Q}}[X] = \int X \, d\mathbb{Q} = \int X \frac{d\mathbb{Q}}{d\mathbb{P}} \, d\mathbb{P} = \mathbb{E}^{\mathbb{P}}\left[X \frac{d\mathbb{Q}}{d\mathbb{P}}\right]$$

The Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is the mathematical tool used to express the measure Q given the measure P. As in the formula above, it allows us for example to characterize all Q expectations starting from P expectations.

We will now show the change of measure needed to change the Brownian motion and the local mean (drift) of the given process X that solves the original SDE. X will be the same but its statistics will change. For more intuition see Appendix on Girsanov.

Theorem (Girsanov's theorem)

If we define for all $t \in [0, T]$

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} := \exp\left[-\frac{1}{2}\int_0^t \left(\frac{f^{\mathbb{Q}}(X_s) - f^{\mathbb{P}}(X_s)}{\sigma(X_s)}\right)^2 ds + \int_0^t \frac{f^{\mathbb{Q}}(X_s) - f^{\mathbb{P}}(X_s)}{\sigma(X_s)}\, dW_s^{\mathbb{P}}\right]$$

then

$$dX_t = f^{\mathbb{Q}}(X_t)\, dt + \sigma(X_t)\, dW_t^{\mathbb{Q}}$$

(recall $dX_t = f^{\mathbb{P}}(X_t)\, dt + \sigma(X_t)\, dW_t^{\mathbb{P}}$) *where* $W^{\mathbb{Q}}$ *is a* BM UNDER $\mathbb{Q}$, *and the measures* $\mathbb{P}$ *and* $\mathbb{Q}$ *are* EQUIVALENT.
A sufficient condition for this to hold is the Novikov condition
$\mathbb{E}\left\{\exp\left(\frac{1}{2}\int_0^T \left(\frac{f^{\mathbb{Q}}(X_s)-f^{\mathbb{P}}(X_s)}{\sigma(X_s)}\right)^2 ds\right)\right\} < +\infty$

So, WE CAN CHANGE THE DRIFT (LOCAL MEAN) OF A SDE BY CHANGING THE EQUIVALENT PROBABILITY MEASURE. THE LOCAL VARIANCE STAYS THE SAME.

In finance, an important version of the Girsanov theorem is the CHANGE OF NUMERAIRE.

PART 2: SDEs FOR OPTION PRICING

In this part we introduce no-arbitrage theory in continuous time, the Black-Scholes SDE option pricing model, and a few volatility smile SDE models.

NO ARBITRAGE, OPTION PRICING AND DERIVATIVES MARKETS

We start now the mathematical finance part of the course.

We introduce no-arbitrage, the Black Scholes and Merton result, their precursors (Bachelier, DeFinetti...) and the refinements of their initial theory (Harrison, Kreps, Pliska....) into no-arbitrage valuation, pointing out its significance, successes and failures.

We also look at the derivatives markets and their significance

Later on we will address some effects of the financial crises of the past, introducing risk measures, although for now we focus on options pricing and hedging.

The Black Scholes and Merton Analysis

We will follows these steps:

- Arbitrage as self-financing trading strategy with zero initial cost attaining a positive payout at maturity.
- Portfolio replication theory plus Ito's formula to derive the Black and Scholes PDE for the option price under certain assumptions on the dynamics of the underlying stock price.
- The Feynman-Kac theorem to interpret the solution of the Black and Scholes PDE as an expected value of a function of the stock price with a modified dynamics.
- The Girsanov theorem to interpret the modified dynamics of the stock price as a dynamics under a different (martingale) probability measure.
- No-arbitrage theorem (Harrison, Kreps and Pliska): There is no arbitrage opportunity if and only if there exists a martingale measure.

Description of the economy

We consider:

- A probability space with a (right continuous) filtration $(\Omega, \mathcal{F}, (\mathcal{F}_t : 0 \le t \le T), P)$
- (& assume $\mathcal{F}_T = \mathcal{F}$). In the given economy, two securities are traded continuously from time 0 until time T. The first one (a bank account, or cash, or riskless bond) is riskless and its (deterministic) price B_t evolves according to

$$dB_t = B_t r \, dt, \quad B_0 = 1, \tag{6}$$

which is equivalent to

$$B_t = e^{rt}, \tag{7}$$

where r is the short term or instantaneous interest of the bank account and it is assumed to be a nonnegative number.

Description of the economy

- As for the second one, given the $(\mathcal{F}_t, \mathbb{P})$-Brownian motion (or Wiener process) W_t, consider the following stochastic differential equation $dS_t = S_t[\mu dt + \sigma dW_t]$, or

$$dS_t = \mu\, S_t dt + \sigma\, S_t dW_t, \quad 0 \le t \le T, \tag{8}$$

with deterministic initial condition $S_0 > 0$, and where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}$, $\sigma > 0$. This is a Geometric Brownian Motion. As seen earlier, Equation (8) has a unique solution which is given by

$$S_t = S_0 \exp\left\{ \left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma W_t \right\}, \quad 0 \le t \le T. \tag{9}$$

The Black & Scholes Assumptions

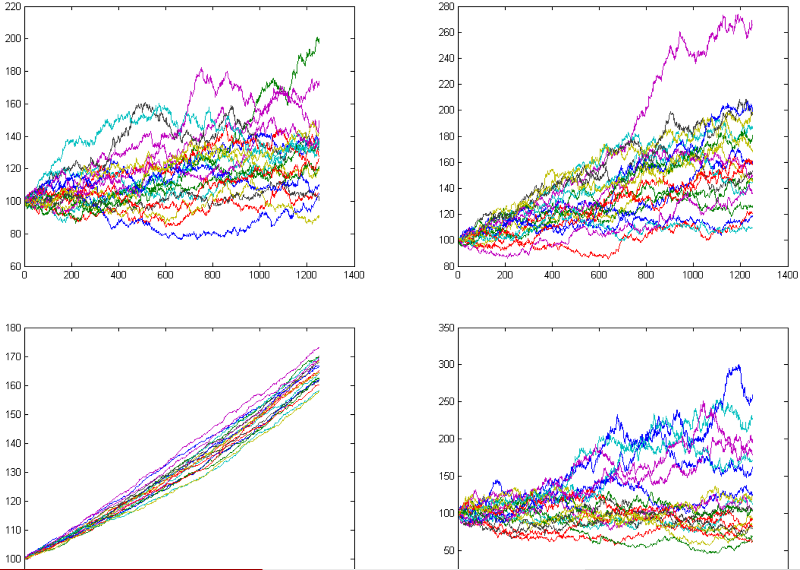
$$
\begin{aligned}
dB_t &= B_t r dt, \quad B_0 = 1, \\
dS_t &= \mu\, S_t\, dt + \sigma\, S_t\, dW_t, \quad 0 \le t \le T,
\end{aligned}
$$

The second asset (a stock) is risky and its price is described by the process S_t. Furthermore, it is assumed that

- (i) there are no transaction costs in trading the stock;
- (ii) the stock pays no dividends or other distributions;
- (iii) shares are infinitely divisible;
- (iv) short selling is allowed without any restriction or penalty. Short selling: investor borrows a security and sells it on the market, planning to buy it back later for less money to give it back to the lender and make a profit. This assumes total absence of credit /default risk.

These assumptions are Black & Scholes' *ideal conditions*.
Example of risky asset dynamics over 5 years:

$S_0 = 100, \quad (\mu, \sigma) = (5\%, 10\%), (10, 10), (10, 1), (1, 20)$



Contingent claims, pricing problem

A **contingent claim** Y for the maturity T is any square-integrable ($\mathbb{E}[Y^2] < +\infty$) and positive random variable in $(\Omega, \mathcal{F}_T, P)$, which is in particular $\mathcal{F}_T$–measurable.

In this derivation we limit ourselves to *simple* contingent claims, i.e. claims of the form $Y = f(S_T)$, measurable functions of the risky asset at the final maturity T.

The idea behind a claim is that it represents an amount that will be paid at maturity to the holder of the contract. I can go to a bank and "buy" the contract for the claim Y. I will pay the initial price at time 0, and at time T the bank will give me the amount $Y(\omega)$ from the claim payoff.

The **Pricing Problem** is giving a fair price to such a contract: how much should the bank charge me as a fair price at time 0?

Trading strategies, Value process, gain process I

A **trading strategy** is a pair of stochastic processes $\phi = (\phi^B, \phi^S)$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t : 0 \leq t \leq T), P)$ that are locally bounded and predictable (and, therefore, $\mathcal{F}_t$–adapted). The pair (ϕ_t^B, ϕ_t^S) represents respectively amounts of bond and stock to be held at time t.
Predictability is assumed to reduce the investor freedom at jump times and assumes that the value ϕ_t will be known immediately before t. However, in our Black Scholes setting, where the paths of the assets are continuous, this issue is not relevant and you need not worry about this assumption. We can just assumed adapted.

The **value process** is the process V describing the value of the portfolio constructed by following the strategy ϕ,

$$V_t(\phi) = \phi_t^B \ B_t + \phi_t^S \ S_t \ .$$

Trading strategies, Value process, gain process II

The **gain process** is defined as

$$G_t(\phi) = \int_0^t \phi_u^B \, dB_u + \int_0^t \phi_u^S \, dS_u \, .$$

and represents the income one obtains thanks to price movements in bond and stock when following the trading strategy ϕ.

The strategy is **self–financing** if $V_t(\phi) \geq 0$ for all t and $V_t(\phi) = V_0(\phi) + G_t(\phi)$, or

$$\phi_t^B \, B_t + \phi_t^S \, S_t \; - (\phi_0^B \, B_0 + \phi_0^S \, S_0) \; = G_t(\phi) \, ,$$

or, in differential terms, $d \; V_t(\phi) = d \; G_t(\phi)$, i.e.

$$d(\phi_t^B \, B_t + \phi_t^S \, S_t) \; = \; \phi_t^B \, dB_t + \phi_t^S \, dS_t \, . \tag{10}$$

Self–financing strategies, arbitrage I

$$d(\phi_t^B\, B_t + \phi_t^S\, S_t) \;=\; \phi_t^B\, dB_t + \phi_t^S\, dS_t\;.$$

Intuitively, this means that the changes in value of the portfolio described by the strategy ϕ are only due to gains/losses coming from price movements, i.e. to changes in the prices B and S, without any cash inflow and outflow.

An important use of self-financing strategies is in defining arbitrage. An **arbitrage opportunity** is a self–financing strategy ϕ such that (recall $V_t(\phi) \geq 0$)

$$\phi_0^B\, B_0 + \phi_0^S\, S_0 \;= 0\;,\;\; \mathbb{P}(\phi_T^B\, B_T + \phi_T^S\, S_T > 0) > 0\;.$$

Self–financing strategies, arbitrage II

Basically, an arbitrage opportunity is a strategy which creates a strictly positive cash inflow from nothing with strictly positive proability and never creates a loss or negative value.

In other terms you have an arbitrage if, with zero initial money, you can only draw or win, and there is a strictly positive probability that you win. “Money from nothing”.

We say that the market is **arbitrage–free**, or simply that we have **no arbitrage**, if there are no arbitrage opportunities.

We will work under no-arbitrage conditions, because no one wants to be in an arbitrageable market, as that means some arbitrageurs can make money from nothing and one would be at a disadvantage against these arbitrageurs and lose money to them.

Self–financing strategies, arbitrage III

In reality there are different definitions of no-arbitrage, involving very complex mathematical issues, including the specific space of trading strategies, its structure, probabilistic poperties and topology, etc.

Based on this, there are more refined definitions related to no arbitrage such as “no free lunch with vanishing risk” (NFLVR), which is slightly more restrictive than the definition of no-arbitrage given above (see the works of Delbaen and Schachermayer).

We have that NFLVR implies no-arbitrage as given above, but not viceversa, although the two definitions are quite close.

For practical purposes, in this introduction by saying “no arbitrage” we will refer to either no-arbitrage as above or to NFLVR, depending on the context.

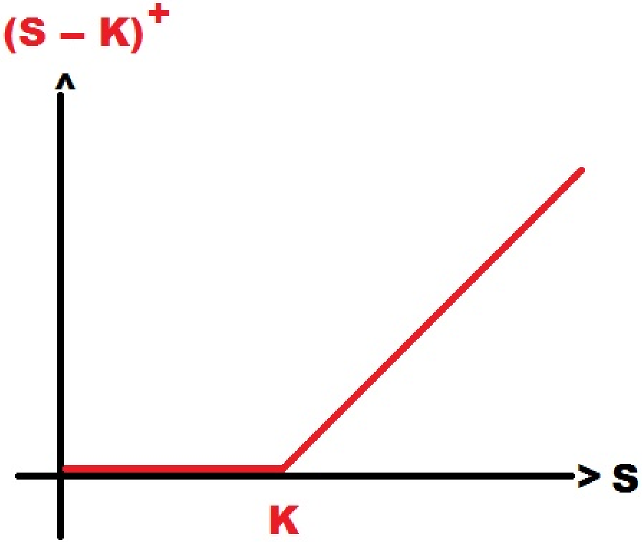
Self–financing strategies, attainable claims, price

Self-financing strategies are important also because they allow us to define **attainable contingent claims**. A contingent claim Y is attainable if $\exists$ self-financing ϕ such that $V_T(\phi) = Y$.

We say that ϕ **generates** Y, & $V_t(\phi)$ is the **price** at time t for Y.

We thus have a **first notion of price** of a claim Y as the value of a self-financing trading strategy attaining (or sometimes we'll say "replicating") the claim Y.

Example of Claim: European Call Option



Example of Claim: European Call Option I

Suppose we have to price a simple claim $Y = f(S_T)$ at time t.

We focus on the case of a European call option: Let K be its strike price and T its maturity. The option payoff (to a long position) is represented by $Y^{Call} = f^{Call}(S_T) = (S_T - K)^+ = \max(S_T - K, 0)$.

This is a contract which at maturity-time T pays nothing if the risky–asset price S_T is smaller than the strike price K, whereas it pays the difference between the two prices in the other case. In other terms, if S_T grows above K by time T, we receive the growth/ difference $S_T - K$, otherwise we receive 0.

Example of Claim: European Call Option II

Example of a defensive use of a call option: Suppose now we are at time 0 and we plan to buy one share (unit) of a certain stock at time T. We wish to pay this stock the same price $K = S_0$ it has now, rather than the price it will have at time T, which could be much higher. We want to be protected by a price increase in the future when we will buy. What one can do in this situation is to buy a call option on the stock with maturity time T and strike price S_0.

He then buys the stock at time T paying S_T and receives $(S_T - S_0)^+$ from the option payoff. Clearly, the total amount he pays in T is then $S_T - (S_T - S_0)^+$ which equals S_T if $S_T \leq S_0$ and equals S_0 if $S_T \geq S_0$. Therefore, an European call option can be seen as a contract which locks the stock price at a desired value to be paid at maturity time T. This *locking* has of course a price, which we wish to determine.

Example of Claim: European Call Option III

An alternative use of the call options is an offensive use and is speculation. If we have a view that the stock price S_T will grow a lot in the future, we can capitalize by buying a call option with strike $K = S_0$, the current stock price.

We pay the price of the option at time 0, but if we are right, the payoff $\max(S_T - S_0, 0)$ will be very large as S_T has grown a lot compared to S_0. This allows us to make a lot of money by time T, as the payoff we get will be much larger than the price we paid for the option at time 0. Of course, if S goes down instead, we will get nothing from the option payoff and we will just lose the option price we paid at time 0.

Example of Claim: European Put Option I

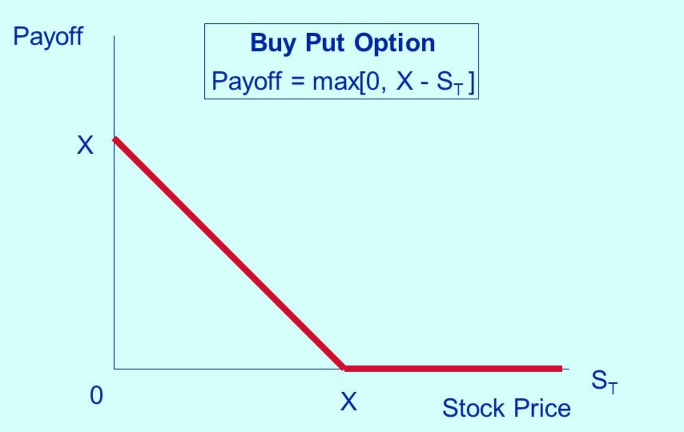


Figure: A one-year maturity Gamble on an equity stock going down. Put Option.

Example of Claim: European Put Option II

The put option is a contingent claim that pays the payoff $Y^{Put} = f^{Put}(S_T) = \max(X - S_T, 0) = (X - S_T)^+$ where X is the strike price (for example, we could take $X = S_0$, the current stock price). The option will pay us at the future time maturity T the difference between the strike X and the stock price at maturity if this difference is positive, and zero otherwise. In other terms, if the stock decreases under X by maturity T, we are paid the gap $X - S_T > 0$, otherwise we are paid 0.

The put option can also be used to protect or speculate.

Defensive use. Suppose we bought a lot of stock. We are concerned that if the stock price goes down too much we could lose a lot. So we

Example of Claim: European Put Option III

buy a put option with strike S_0. The total of holding one stock and one put option at maturity T is

$$S_T + \max(S_0 - S_T, 0) = S_0 \text{ if } S_0 > S_T \text{ and } S_T \text{ if } S_T \leq S_0 = \max(S_0, S_T).$$

So with a put option added to our stock, at maturity T we get always the best between the final stock price S_T and the initial one S_0. So if the stock goes down we are protected and we get the initial S_0, while if it goes up we keep the final stock S_T. Of course this protection will cost us the option price.

Offensive use. Put options can be used for speculation too, in an offensive fashion. If we have a view that the stock price will go down, we can profit it by buying a put option with strike $X = S_0$, the current stock price.

Example of Claim: European Put Option IV

We pay the price of the option at time 0, but if we are right, the payoff $\max(S_0 - S_T, 0)$ will be very large as S_T has gone down a lot compared to S_0. This allows us to make a lot of money by time T, as the payoff we get will be much larger than the price we paid for the option at time 0. Of course, if S goes up instead, we will get nothing from the option payoff and we will just lose the option price we paid at time 0.

We now sketch a derivation of the Black Scholes PDE for the "attainable-claim" price of an option $Y = f(S_T)$, and we will consider in particular the case of a call option $f(S_T) = (S_T - K)^+$, although the derivation is general, if the simple contingent claim is not a call option then this still works by putting $V(T, S) = f(S)$ as terminal condition of the PDE below. This is a relatively informal derivation.

The Black & Scholes PDE for a simple claim

We now assume that the value of the simple claim at time t is a function of the underlying stock S at the same time, namely $V_t = V(t, S_t)$. This is the candidate claim (option) value at time t. Assume the function $V(t, S)$ of time t and of the stock price S to have regularity $V \in C^{1,2}([0, T] \times \mathbb{R}^+)$. In other terms, we assume V is twice continously differentiable with respect to S and once continuously differentiable with respect to t. Apply Ito's formula to V:

$$dV(t, S_t) = \frac{\partial V}{\partial t}(t, S_t)\, dt + \frac{\partial V}{\partial S}(t, S_t)\, dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(t, S_t)\, dS_t\, dS_t. \quad (11)$$

Substituting the equation for $dS_t = \mu S_t dt + \sigma S_t dW_t$ and recalling that $dt\ dt = 0,\ dt\ dW_t = 0,\ dW_t dW_t = dt$, we get

$$\begin{aligned} dV(t, S_t) &= \left(\frac{\partial V}{\partial t}(t, S_t) + \frac{\partial V}{\partial S}(t, S_t)\mu S_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(t, S_t)\sigma^2 S_t^2\right) dt \\ &+ \frac{\partial V}{\partial S}(t, S_t)\sigma S_t dW_t. \end{aligned} \quad (12)$$

The Black and Scholes PDE

Now, if we are looking to attain our claim with a self financing strategy ϕ^S, ϕ^B, so that $V(t, S_t) = \phi_t^S S_t + \phi_t^B B_t$, this will have to satisfy the self financing condition, namely

$$dV(t, S_t) = \phi_t^B dB_t + \phi_t^S \, dS_t. \tag{13}$$

Compare this last Eq to Eq (11) in the previous slide and match the dS_t terms. We get, for each $0 \leq t \leq T$,

$$\phi_t^S = \frac{\partial V}{\partial S}(t, S_t), \quad \phi_t^B = (V_t - \phi_t^S S_t)/B_t. \tag{14}$$

where the first equation comes from the matching, and the second equation comes by construction, as the value of the strategy at time t must be V itself, and clearly $V(t, S_t) = \phi_t^B \, B_t + \ \phi_t^S \, S_t$. In other terms, to get the second equation solve $V(t, S_t) = \phi_t^B \, B_t + \ \phi_t^S \, S_t$ in ϕ_t^B.

The Black and Scholes PDE

Now we can explicit the self financing condition for ϕ:

$$\begin{aligned} dV_t &= \phi_t^B dB_t + \phi_t^S dS_t \qquad (15) \\ &= \left[V(t,S_t) - \frac{\partial V}{\partial S}(t,S_t)S_t \right] rdt + \frac{\partial V}{\partial S}(t,S_t)S_t(\mu dt + \sigma dW_t). \end{aligned}$$

Then by equating (12) and (15) (ITO + SELF FINANCING), we obtain that V_t satisfies

$$\frac{\partial V}{\partial t}(t,S) + \frac{\partial V}{\partial S}(t,S)rS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(t,S)\sigma^2 S^2 = rV(t,S), \qquad (16)$$

which is the celebrated Black and Scholes partial differential equation with terminal condition $V(T,S) = (S-K)^+$ (or more generally $V(T,S) = f(S)$).

Black and Scholes' famous formula

The strategy (ϕ^B, ϕ^S) has final value equal to the claim Y we wish to price (terminal condition of the PDE), and during its life the strategy does not involve cash inflows or outflows (self–financing condition). As a consequence, its initial value V_t at time t must be equal to the unique claim price to avoid arbitrage opportunities.
The solution of the above equation is given by

$$V_{BS}(t) = V_{BS}(t, S_t, K, T, \sigma, r) := S_t \Phi(d_1(t)) - K e^{-r(T-t)} \Phi(d_2(t)), \quad (17)$$

where

$$d_1(t) := \frac{\ln(S_t/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}},$$
$$d_2(t) := d_1(t) - \sigma\sqrt{T-t},$$

and $\Phi(\cdot)$ denotes the cumulative standard normal distribution function.

Black and Scholes' famous formula

Expression (17) is the celebrated Black and Scholes option pricing formula which provides the unique no-arbitrage price for the given European call option.

Notice that the coefficient μ does not appear in (17), indicating that investors, though having different risk preferences or predictions about the future stock price behaviour, must yet agree on this unique option price.

MORE ON THE SIGNIFICANCE OF THIS LATER.

Numerical example

Suppose the current stock value is $S_0 = 100$.
Suppose the risk free interest rate is $r = 2\% = 0.02$.
Suppose that the strike $K = 100$ (at the money option).
Assume the volatility $\sigma = 0.2 = 20\%$.
Take a maturity of $T = 5y$. CALL PRICE IS $V_{BS}(0) = 22.02$.

For example, in Matlab this is obtained through commands

```
S0=100; sig=0.2; r=0.02; K=100; T=5;
d1 = (r + 0.5*sig*sig)*T/(sig*sqrt(T));
d2 = (r - 0.5*sig*sig)*T/(sig*sqrt(T));
V0 = S0*normcdf(d1)-K*exp(-r*T)*normcdf(d2);
```

The same calculation with lower volatility $\sigma = 0.05 = 5\%$ would give

$$V_{BS}(0)|_{\sigma=0.05} = 10.5943, \quad V_{BS}(0)|_{\sigma=0.0001} = 9.52.$$

The last value is very close to the value $S_0 - Ke^{-rT}$.

Numerical example

- Acme today is worth $S_0 = 100$.
- The more the value of acme goes up in 5 years, the more we gain as $S_{5y} - S_0$ grows. In a scenario where $S_{5y} = 200$, we gain 100.
- If however Acme goes down instead, $S_{5y} - S_0$ goes negative but the option $(S_{5y} - S_0)^+$ caps it at zero and we lose nothing. For example, in a scenario where Acme goes down to 60, we get $(60 - 100)^+ = (-40)^+ = 0$ ie we lose nothing
- With the original data, entering the gamble costs initially 22 USD out of 100 of stock notional. It is expensive. On the other hand, it is a gamble where we can only win and in principle have scenarios with unlimited profit.
- You will notice that:

$$\uparrow \sigma \Rightarrow V_{CallBS} \uparrow, \quad \uparrow S_0 \Rightarrow V_{CallBS} \uparrow, \quad \downarrow K \Rightarrow V_{CallBS} \uparrow \ldots.$$

Numerical example

- Indeed, if we lower the volatility to from $\sigma = 0.2$ to $\sigma = 0.05$, the initial option price $V_{BS}(0)$ drops from 22.02 to 10.59.
- With lower volatility the call option price we pay at time 0 is smaller. However, being the volatility smaller, there is less randomness in the system and it's less likely that the stock moves a lot above K at maturity T, so our profit scenarios at T are more limited, and correspondingly the option price at time 0 is lower.
- When the volatility goes very close to 0, for an option that is at the money $K = S_0$ we get a value close to the value (take the limit for $\sigma \downarrow 0$ in the BS formula) $V(0) = 9.52 \approx S_0 - S_0 e^{-rT}$. This is the lowest possible call price when $K = S_0$.

In–, at–, and out–of–the money options I

A call option is in-the-money if the initial stock price is larger than the strike, $S_0 > K$. This means that if the call were exercised immediately, at time 0, we would have a positive payoff $(S_0 - K)^+ > 0$. The immediate-exercise value $(S_0 - K)^+$ is called the intrinsic value of the call option. So in-the-money means positive intrinsic value.

A call option is at-the-money if the initial stock equals the strike, $S_0 = K$. This means that if the call were exercised immediately, we would have a zero payoff $(S_0 - S_0)^+ = 0$, but as soon as the stock moves we go either in the money or out of the money. The intrinsic value is 0.

A call option is out-of-the-money if the initial stock is smaller than the strike, $S_0 < K$. If the call were exercised immediately, we would have a zero payoff $(S_0 - K)^+ = 0$. The intrinsic value is zero.

In–, at–, and out–of–the money options II

For the put options, a put option intrinsic value is again the immediate exercise value at time 0, $(K - S_0)^+$. The put will be in-the-money if $S_0 < K$ ($(K - S_0)^+ > 0$ intrinsic value), at the money if $K = S_0$ ($(K - S_0)^+ = 0$ intrinsic value), and out-of-the-money if $K < S_0$ ($(K - S_0)^+ = 0$ intrinsic value).

The above definitions are modified in “In-, at- or out-of-the-money-forward options” when S_0 is replaced by $S_0 e^{rT}$. More specifically:

A call option will be in-the-money-forward if $K < S_0 e^{rT}$ (implying the payoff at time 0 with a T discounted strike K to be $(S_0 - Ke^{-rT})^+ > 0$), it will be at the money forward if $K = S_0 e^{rT}$, i.e. the strike being equal to the forward stock price at time 0 for maturity T (this is equivalent to $(S_0 - Ke^{-rT}) = (S_0 - Ke^{-rT})^+ = 0$), with the option going

In–, at–, and out–of–the money options III

in-the-money-forward or out-of-the-money forward as soon as the stock moves; the name “forward stock price at time 0 for maturity T” will be explained later when introducing the forward contract), and the call option will be out-of-the-money-forward if $K > S_0 e^{rT}$ (implying $(S_0 - Ke^{-rT})^+ = 0$).

The definitions for the put option are analogous: a put option will be in-the-money-forward if $K > S_0 e^{rT}$ (implying the payoff at time 0 with a T discounted strike to be $(Ke^{-rT} - S_0)^+ > 0$), it will be at the money forward if $K = S_0 e^{rT}$ (equivalent to $(Ke^{-rT} - S_0) = (Ke^{-rT} - S_0)^+ = 0$, with the option going in-the-money-forward or out-of-the-money forward as soon as the stock moves), and it will be out-of-the-money-forward if $K < S_0 e^{rT}$ (implying $(Ke^{-rT} - S_0)^+ = 0$).

In–, at–, and out–of–the money options IV

The difference between at-the-money and at-the money-forward is that we look for the strike that makes the payoff inside the option equal zero when priced at time 0, but in the first case we don't discount the strike from the option maturity T, while in the second case we do. This coincides with defining the strike for at the money forward as the forward price at time 0 for maturity T:

$$S_0 - K = 0 \Rightarrow \text{ at-the-money } K,\ K = S_0.$$

$$S_0 - Ke^{-rT} = 0 \Rightarrow \text{ at-the-money-forward } K,\ K = S_0 e^{rT} = \text{forward price.}$$

Verifying the Self financing condition

Going back to the general Black Scholes result, we then prove that the strategy (for a call option)

$$\phi_t^S = \frac{\partial V_{BS}}{\partial S}(t, S_t), \quad \phi_t^B = (V_{BS}(t) - \phi_t^S S_t)/B_t \tag{18}$$

$$\left(V_{BS}(t) = V_{BS}(t, S_t, K, T, \sigma, r) := S_t \Phi(d_1(t)) - K e^{-r(T-t)} \Phi(d_2(t)) \right),$$

is indeed self-financing. By Ito's Lemma, in fact, we have

$$dV_{BS}(t) = \frac{\partial}{\partial t} V_{BS}(t) dt + \frac{\partial}{\partial S} V_{BS}(t) dS_t + \frac{1}{2} \frac{\partial^2}{\partial S^2} V_{BS}(t) \sigma^2 S_t^2 dt. \tag{19}$$

Verifying the Self financing condition

We need straightforward differentiation of the V_{BS} expression plus several calculations and the following Lemma

$$\textbf{Greeks Lemma}. \quad S_t\phi(d_1(t)) = Ke^{-r(T-t)}\phi(d_2(t)) \tag{20}$$

whose proof is not difficult but long, and will be presented later. The lemma is easy to remember if written $S_t\phi(d_1(t)) - Ke^{-r(T-t)}\phi(d_2(t)) = 0$, it's the Black Scholes formula for a call with the normal PDF ϕ replacing the CDF Φ (recall $\Phi'(x) = \phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$) set to zero. You do not need to fill these calculations, although you may be asked similar calculations on sensitivity of the option price to time t or time to maturity $T-t$. We get

$$\frac{\partial}{\partial t}V_{BS}(t) = -\frac{S_t\phi(d_1(t))\sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}\Phi(d_2(t)), \tag{21}$$

$$\frac{\partial}{\partial S}V_{BS}(t) = \Phi(d_1(t)), \quad \frac{\partial^2}{\partial S^2}V_{BS}(t) = \frac{\phi(d_1(t))}{S_t\sigma\sqrt{T-t}},$$

Verifying the Self financing condition

To conclude it is enough to substitute the ϕ^S and ϕ^B expressions given in (18) into $dV_{BS}(t) = \phi_t^S dS_t + \phi_t^B dB_t$ to obtain from (19), with Eqs. (21) plugged in, that this holds, and this is the self–financing condition in differential form.

The calculation of the Delta, $\frac{\partial}{\partial S} V_{BS}(t) = \Phi(d_1(t))$ will be proven in detail later.

The Feynman Kac theorem for Risk Neutral Valuation

Different interpretation: the Feynman-Kac Theorem allows to interpret the solution of a parabolic PDE such as the Black and Scholes PDE in terms of expected values of a diffusion process.
The solution of the PDE

$$\frac{\partial V}{\partial t}(t,S)+\frac{\partial V}{\partial S}(t,S)\boxed{\boxed{rS}}+\frac{1}{2}\frac{\partial^2 V}{\partial S^2}(t,S)\sigma^2 S^2=\boxed{r}V(t,S),\quad V(T,S)=f(S)$$

can be expressed as

$$V(t,S_t)=e^{-\boxed{r}(T-t)}\mathbb{E}^Q[f(S_T)|\mathcal{F}_t^S] \tag{22}$$

where the diffusion process S has dynamics starting from S_t at time t

$$dS_u=\boxed{\boxed{rS_u}}du+\sigma S_u dW_u^Q\ ,u\geq t,\ S_u|_{u=t}=S_t \tag{23}$$

under the probability measure $\mathbb{Q}$ under which the expectation $\mathbb{E}^Q[\cdot|\mathcal{F}_t^S]$ is taken. The process W^Q is a standard Brownian motion under $\mathbb{Q}$.

Risk Neutral interpretation of the B e S's formula

We obtained that the unique no-arbitrage price of the integrable contingent claim $Y = f(S_T) = (S_T - K)^+$ (European call option) at time t, $0 \leq t \leq T$, is given by

$$V_{BS}(t) = \mathbb{E}^Q \left(e^{-r(T-t)} f(S_T) | \mathcal{F}_t^S \right). \tag{24}$$

The expectation is taken with respect to the so-called **martingale measure** $\mathbb{Q}$, i.e. a probability measure $\mathbb{Q} \sim \mathbb{P}$ under which the risky–asset price $S_t/B_t = e^{-rt} S_t$ measured with respect to the risk-free asset price B_t is a martingale, which is equivalent to S having drift rate $\boxed{r}$ under $\mathbb{Q}$:

$$dS_t = S_t[\boxed{r}dt + \sigma dW_t^Q], \quad 0 \leq t \leq T, \tag{25}$$

Deriving Black-Scholes via risk neutral valuation I

We now derive the Black Scholes formula

$$V_{BS}(0) = e^{-rT} E_0^Q[(S_T - K)^+]$$

under the assumption

$$dS_t = rS_t dt + \sigma S_t dW_t^Q, \quad s_0.$$

We have solved Geometric Brownian Motion equations like this several times and we know that the solution is

$$S(t) = s_0 \exp\left(\sigma W_t^Q + (r - \frac{\sigma^2}{2})t\right)$$

We will omit the Q in W^Q for brevity.

Deriving Black-Scholes via risk neutral valuation II

Compute the distribution of the random variable in the exponent. It is Gaussian, since it is a Brownian motion W_t (which is Gaussian with zero mean and variance t) plus a deterministic quantity. Thus

$$E[\sigma W_t + (r - \sigma^2/2)t] = 0 + (r - \sigma^2/2)t$$

and the variance (recall Var$(X + \mathit{constant})$ =Var(X))

$$\mathrm{Var}\left[\sigma W_t + (r - \sigma^2/2)t\right] = \mathrm{Var}[\sigma W_t] = \sigma^2 t.$$

We thus have

$$I(T) := \sigma W_T + (r - \frac{\sigma^2}{2})T \sim m + V\mathcal{N}(0,1), \;\; m = (r - \frac{1}{2}\sigma^2)T, \;\; V^2 = \sigma^2 T$$

Deriving Black-Scholes via risk neutral valuation III

Recall that we have

$$S_T = s_0 \exp(I(T)) = s_0 e^{m+V\mathcal{N}(0,1)}$$

Compute the option price (omitting for now the discounting e^{-rT} to be added later)

$$E^Q[(S_T - K)^+] = E^Q[(s_0 e^{m+V\mathcal{N}(0,1)} - K)^+]$$
$$= \int_{-\infty}^{+\infty} (s_0 e^{m+Vy} - K)^+ p_{\mathcal{N}(0,1)}(y) dy = \ldots$$

Note that $s_0 \exp(m + Vy) - K > 0$ if and only if

$$y > \frac{-\ln\left(\frac{s_0}{K}\right) - m}{V} =: \bar{y}$$

Deriving Black-Scholes via risk neutral valuation IV

so that

$$\ldots = \int_{\bar{y}}^{+\infty} (s_0 \exp(m + Vy) - K) p_{\mathcal{N}(0,1)}(y) dy =$$

$$= s_0 \int_{\bar{y}}^{+\infty} e^{m+Vy} p_{\mathcal{N}(0,1)}(y) dy - K \int_{\bar{y}}^{+\infty} p_{\mathcal{N}(0,1)}(y) dy =$$

Deriving Black-Scholes via risk neutral valuation V

$$= s_0 \frac{1}{\sqrt{2\pi}} \int_{\bar{y}}^{+\infty} e^{-\frac{1}{2}y^2 + Vy + m} dy - K(1 - \Phi(\bar{y}))$$

$$= s_0 \frac{1}{\sqrt{2\pi}} \int_{\bar{y}}^{+\infty} e^{-\frac{1}{2}(y-V)^2 + m + \frac{1}{2}V^2} dy - K(1 - \Phi(\bar{y})) =$$

$$= s_0 e^{m + \frac{1}{2}V^2} \frac{1}{\sqrt{2\pi}} \int_{\bar{y}}^{+\infty} e^{-\frac{1}{2}(y-V)^2} dy - K(1 - \Phi(\bar{y})) =$$

$$= s_0 e^{m + \frac{1}{2}V^2} \frac{1}{\sqrt{2\pi}} \int_{\bar{y}-V}^{+\infty} e^{-\frac{1}{2}z^2} dz - K(1 - \Phi(\bar{y})) =$$

$$= s_0 e^{m + \frac{1}{2}V^2} (1 - \Phi(\bar{y} - V)) - K(1 - \Phi(\bar{y})) =$$

$$= s_0 e^{m + \frac{1}{2}V^2} \Phi(-\bar{y} + V) - K\Phi(-\bar{y}) =$$

$$= s_0 e^{rT} \Phi(d_1) - K\Phi(d_2), \quad d_{1,2} = \frac{\ln \frac{s_0}{K} + (r \pm \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

Deriving Black-Scholes via risk neutral valuation VI

This is an expression for $E_0^Q[(S_T - K)^+]$. Now we need to add the discount term:

$$V_{BS}(0, S_0, K, T, \sigma, r) = e^{-rT} E_0^Q[(S_T - K)^+] = s_0 \Phi(d_1) - K e^{-rT} \Phi(d_2).$$

$$d_{1,2} = \frac{\ln \frac{s_0}{K} + (r \pm \frac{1}{2}\sigma^2) T}{\sigma \sqrt{T}}$$

This confirms the price formula we had given earlier for the call option, without proof, via the PDE (Ito formula plus self financing condition) method.

The above formula for a call option has to be learned by heart, as it is the most used formula by traders in derivatives markets. While one has to know how to derive the formula, it is important to have it ready in memory for many applications.

Deriving Black-Scholes via risk neutral valuation VII

We now compute the call option Delta, namely the sensitivity of the option price to the initial condition. This is our term $\phi^S(0)$ in the trading strategy we used to calculate the option price with the PDE.
This is done with a partial derivative

$$\phi^S(0) = \Delta_0 = \frac{\partial V_{BS}(0, S_0, K, T, \sigma, r)}{\partial S_0} =$$

$$= \frac{\partial\ s_0 \Phi(d_1)}{\partial S_0} - \frac{\partial K e^{-rT} \Phi(d_2)}{\partial S_0} = \ldots$$

Please note that d_1 and d_2 also depend on S_0. We have

$$\ldots = \Phi(d_1) + s_0 \frac{\partial \Phi(d_1)}{\partial S_0} - K e^{-rT} \frac{\partial \Phi(d_2)}{\partial S_0} =$$

$$= \Phi(d_1) + s_0 \Phi'(d_1) \frac{\partial d_1}{\partial S_0} - K e^{-rT} \Phi'(d_2) \frac{\partial d_2}{\partial S_0} =$$

Deriving Black-Scholes via risk neutral valuation VIII

$$= \Phi(d_1) + s_0\phi(d_1)\frac{1}{S_0\sigma\sqrt{T}} - Ke^{-rT}\phi(d_2)\frac{1}{S_0\sigma\sqrt{T}} \tag{26}$$

$$= \Phi(d_1) + \frac{1}{\sigma\sqrt{2\pi T}}e^{-\frac{d_1^2}{2}} - \frac{Ke^{-rT}}{s_0\sigma\sqrt{2\pi T}}e^{-\frac{d_2^2}{2}}$$

$$= \Phi(d_1) + \frac{1}{\sigma\sqrt{2\pi T}}\exp\left(-\frac{1}{2}\left[\frac{\ln\frac{S_0}{K} + (r+\sigma^2/2)T}{\sigma\sqrt{T}}\right]^2\right)$$

$$-\frac{Ke^{-rT}}{s_0\sigma\sqrt{2\pi T}}\exp\left(-\frac{1}{2}\left[\frac{\ln\frac{S_0}{K} + (r-\sigma^2/2)T}{\sigma\sqrt{T}}\right]^2\right)$$

Deriving Black-Scholes via risk neutral valuation IX

$$= \Phi(d_1) + \frac{1}{\sigma\sqrt{2\pi T}} \exp\left(-\frac{1}{2}\left[\frac{\ln\frac{S_0}{K} + (r+\sigma^2/2)T}{\sigma\sqrt{T}}\right]^2\right)$$

$$-\frac{1}{\sigma\sqrt{2\pi T}} \exp\left(-rT - \ln\frac{S_0}{K} - \frac{1}{2}\left[\frac{\ln\frac{S_0}{K} + (r-\sigma^2/2)T}{\sigma\sqrt{T}}\right]^2\right)$$

$$= \Phi(d_1) + 0 = \Phi(d_1)$$

as the two terms after $\Phi(d_1)$ cancel each other. We conclude

$$\phi^S(0) = \Delta_0 = \frac{\partial V_{BS}(0, S_0, K, T, \sigma, r)}{\partial S_0} = \Phi(d_1).$$

This formula should be memorized as part of the Black Scholes setting. Please note that in differentiating with respect to s_0 it's as if we

Deriving Black-Scholes via risk neutral valuation X

had only differentiated with respect to s_0 in the boxed term but skipped the other s_0 terms in d_1 and d_2:

$$\Delta_0 = \frac{\partial \left(\boxed{s_0} \Phi(d_1) - K e^{-rT} \Phi(d_2) \right)}{\partial S_0}.$$

Indeed, in that case we would get immediately $\Phi(d_1)$ without all the other calculations we did above. However, the real reason why we can ignore d_1 and d_2 for the delta, when differentiating, is rather subtle and we won't explain it here. Moreover, as we will see in a minute, the fact that all reduces to $\Phi(d_1)$ only, allows us to prove the Greeks Lemma.

Note that the delta of the call option is always positive. This means that the option price increases when the underlying stock s_0 increases.

Deriving Black-Scholes via risk neutral valuation XI

In calculating Delta we have proven the Greeks Lemma Eq. (20)
Recall the lemma written for $t = 0$ for simplicity:

$$\textbf{Greeks Lemma}. \quad S_0 \phi(d_1) - K e^{-rT} \phi(d_2) = 0.$$

We have proven this as part of our Delta calculation. Indeed, in the step given by Eq. (26) we have that

$$\Delta_0 = \Phi(d_1) + \frac{1}{S_0 \sigma \sqrt{T}} \left(s_0 \phi(d_1) - K e^{-rT} \phi(d_2) \right)$$

but we also know, having completed the calculation later, that $\Delta_0 = \Phi(d_1)$ which then implies

$$\frac{1}{S_0 \sigma \sqrt{T}} \left(s_0 \phi(d_1) - K e^{-rT} \phi(d_2) \right) = 0$$

from which the Lemma follows immediately.

Deriving Black-Scholes via risk neutral valuation XII

Finally, one important point that we clarify for the avoidance of doubt. For some options, the strike K is a function of S_0, for example if the option is at-the-money, $K = S_0$. A misconception is that, when we compute the delta, we need to differentiate also with respect to the S_0 in the strike. This is not true. We differentiate the option price keeping K fixed and substitute a posteriori $K = S_0$.
Indeed, when we compute the delta of a call option, for example, we want to measure how the option value changes for small changes of the stock price at time 0, but not for changes in the strike. The strike of the option does not change, even if it had been fixed to S_0 itself. So if we were to write the delta of an at-the-money option as a limit, $\Delta_0 =$

$$\lim_{\Delta S \downarrow 0} \frac{V_{BS}(0, S_0 + \Delta S, K, T, \sigma, r) - V_{BS}(0, S_0, K, T, \sigma, r)}{\Delta S}\Bigg|_{K=S_0} = \Phi(d_1)|_{K=S_0}.$$

The Risk Neutral measure via Girsanov's theorem I

We apply Girsanov's theorem to move from

$$d\ S_t = \mu S_t\ dt + \sigma S_t\ dW_t^P \ \text{ to } \ d\ S_t = r S_t\ dt + \sigma S_t\ dW_t^Q$$

and obtain the Radon Nykodym derivative connecting $\mathbb{Q}$ with $\mathbb{P}$.

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t^S} = \exp\left\{-\frac{1}{2}\int_0^t \left(\frac{\mu S_u - r S_u}{\sigma S_u}\right)^2 du - \int_0^t \frac{\mu S_u - r S_u}{\sigma S_u} dW_u^P\right\}. \tag{27}$$

simplifying to

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t^S} = \exp\left\{-\frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 t - \frac{\mu - r}{\sigma} W_t^P\right\}. \tag{28}$$

Note that dQ/dP is a random variable when seen from time 0 (notice W_t^P).

The Risk Neutral measure via Girsanov's theorem II

Note that the Novikov condition needed for Girsanov to work is satisfied trivially:

$$E\left[\exp\left(\frac{1}{2}\int_0^T\left(\frac{\mu S_t - rS_t}{\sigma S_t}\right)^2 dt\right)\right] = \exp\left(\frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 T\right) < +\infty$$

The quantity

$$\frac{\mu - r}{\sigma}$$

is called "market price of risk", or sometimes, in finance circles, the Sharpe ratio. It tells us how much better the stock S is doing with respect to the risk free rate, divided by the volatility. In the real world the stock local growth rate or "return" is μ. So $\mu - r$ is the difference between S's return and the risk free rate, telling us how much better S is doing than a cash account B.

No arbitrage: Main steps followed so far I

1. Self Financing Condition (Portfolio replication theory) plus Ito's formula to derive the Black and Scholes PDE for any simple attainable payout claim $f(S_T)$ in S_T:

$$d\ S_t = \mu S_t\ dt + \sigma S_t\ dW_t$$

$$\frac{\partial V}{\partial t}(t, S_t) + \frac{\partial V}{\partial S}(t, S_t) r S_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(t, S_t)\sigma^2 S_t^2 = rV(t, S_t),$$

$$V_T = f(S_T)$$

2. If each such claim can be replicated/attained with a unique self financing strategy then there is a unique claim price equal to the initial cost of the strategy (and given by the PDE).

No arbitrage: Main steps followed so far II

3. The Feynman-Kac theorem to interpret the price solution of the Black and Scholes PDE as an expected value of a function of the stock price with modified dynamics

$$V(t, S_t) = \mathbb{E}^Q\{e^{-r(T-t)} f(S_T) | \mathcal{F}_t\}$$

$$d\ S_t = rS_t\ dt + \sigma S_t\ dW_t^Q$$

4. The Girsanov theorem to interpret the modified dynamics of the stock price as a dynamics under a new (Risk neutral or martingale) probability measure $\mathbb{Q}$:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left\{-\frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 T - \frac{\mu - r}{\sigma} W_T\right\}.$$

No arbitrage: Main steps followed so far III

5. Hence the notion of attainable/replication claim price obtained from the PDE (self-financing condition & Ito's formula) coincides with a second notion of price: expectation of the claim payout under a risk neutral measure where the risky asset local mean grows at the risk free rate. This is a second way to express no-arbitrage via the condition that S/B is a martingale (more on martingales in a minute), ie a fair game. Hence no arbitrage will be related to the market for the underlying risky asset S to be a fair game.

No arbitrage: Main steps followed so far IV

6. We will now generalize this result presenting the fundamental theorems of no arbitrage. This part of the course is meant to give a brief flavour of what no-arbitrage theory deals with, without being fully rigorous or exhaustive, as doing that would require much stronger pre-requisites and would take away time from applications. A few students are occasionally interested in knowing more about attainable claims, self financing strategies, arbitrage opportunities, fundamental theorems, etc. For these students I recommend the book “Arbitrage theory in continuous time” by the late Tomas Bjoerk.

Fundamental Theorems I

Pricing, no arbitrage, complete markets

The two approaches to no-arbitrage pricing, (i) attainable claim PDE and (ii) Risk neutral expectation, were connected by the Feynman-Kac theorem, that gave us (ii) from (i). But (i) and (ii) are more generally related by the full theory of no-abritrage of Harrison, Kreps and Pliska, and following extensions such as Dalang, Morton & Willinger and Delbaen & Schachermayer, and they are equivalent to the absence of arbitrage opportunities as defined earlier. Without specifying fully all the technical details, we report a high level summary. Here $\exists$ means "there exists" and $\exists!$ means "there exists a unique...".

First Fundamental Theorem of Asset Pricing. We call a martingale meaure a probability measure under which any risky assets divided by the risk free asset is a martingale. Then
$\exists$ a martingale measure $\mathbb{Q} \iff$ we have no arbitrage opportunities.

Fundamental Theorems II

Pricing, no arbitrage, complete markets

Thus existence of a martingale measure is equivalent to having an arbitrage-free market. Note that in Black-Scholes, under $\mathbb{Q}$, the risky asset divided by the risk free asset, S_t/B_t, is a martingale as we will see in a minute, so we have that Black-Scholes is arbitrage free.

The precise statements on the right hand side would be "we have no free lunch with vanishing risk" (NFLVR).

Fundamental: if $\mathbb{Q}\ \exists$, there is no arbitrage opportunity, i.e. there is no self–financing ϕ with $V_t(\phi) \geq 0$ for all t, producing positive final wealth $V_T(\phi) > 0$ with positive probability > 0 with zero initial cost $V_0(\phi) = 0$. Vice versa: if there are no arbitrage opportunities ϕ then a martingale measure $\mathbb{Q}$ exists.

Fundamental Theorems III

Pricing, no arbitrage, complete markets

Theorem ("Pricing with discounted $\mathbb{Q}$ expectations"). $\exists$ a martingale measure $\mathbb{Q} \Rightarrow \exists!$ no-arbitrage attainable-claim price that can be computed as a $\mathbb{Q}$ expectation of the discounted claim.

The above theorem is fundamental, and is very general. It doesn't work just for Black Scholes, but for any SDE model for dS_t such that there exists a measure Q where S_t/B_t is a martingale (see smile modeling later).

There is a second result related to the uniqueness (rather than existence) of the martingale measure. This is related to complete markets.

A market is **complete** if every contingent claim is attainable.

Fundamental Theorems IV

Pricing, no arbitrage, complete markets

Second fundamental theorem of asset pricing
Market is arbitrage free & complete $\iff$ $\exists!$ martingale measure $\mathbb{Q}$

If the market is arbitrage free but not complete, the price of any attainable claim is still uniquely given, either as the value of the replicating strategy or as the risk neutral expectation under *any* equivalent martingale measure.

The Black Scholes market (B_t, S_t) we have seen above is arbitrage free and complete.

In reality markets are never complete, as many risks are not directly associated with tradable assets, so one has to find ways to deal with market incompleteness.

Fundamental Theorems V

Pricing, no arbitrage, complete markets

The above framework can be applied easily to markets with n diffusive underlying assets $S_1, \ldots, S_n$, each similar to the Black Scholes equity process, and with a bank or cash account B_t. The definitions and results on arbitrage opportunities, attainable claims, price, martingale measure, market completeness extend to the n-dimensional case easily and also to non-simple claims that are path dependent or early exercise.

The idea behind the martingale approach

Why martingales?

A martingale is a stochastic process representing a fair game, as on average it doesn't go either up or down. Loosely speaking, the above proposition states that in order to price under uncertainty one must price in a world where the probability measure is such that the risky asset evolves as a fair game when expressed in units of the risk–free asset.

Hence in our case S_t/B_t must be a fair game, ie a martingale.

martingales: local mean =0

As seen earlier, for regular diffusion processes X_t martingale means "zero-drift", no up or down local direction: $dX_t = 0dt + \sigma(t, X_t)dW_t$.

Indeed, show that the drift of the SDE for $d(S_t/B_t)$ is zero under $\mathbb{Q}$.

S_t/B_t is a martingale under $\mathbb{Q}$

We show that S_t/B_t is a martingale under $\mathbb{Q}$ (that's why sometimes $\mathbb{Q}$ is called the martingale measure) by showing that the SDE for S_t/B_t has zero drift under $\mathbb{Q}$, recalling $dS_t = rS_t dt + \sigma S_t dW_t^Q$ and $B_t = e^{r\ t}$.

Let $Y_t = S_t/B_t = S_t/e^{r\ t} = e^{-r\ t}S_t$.

$$dY_t = d(e^{-rt})S_t + e^{-rt}dS_t = \cdots$$

Note that there is no quadratic covariation term, $d(e^{-r\ t})dS_t = 0$ because $dt\ dt = 0$ and $dt\ dW_t = 0$.

$$\cdots = -re^{-rt}S_t\ dt + e^{-rt}dS_t = -re^{-rt}S_t\ dt + e^{-rt}(rS_t dt + \sigma S_t dW_t^Q) =$$

$$= e^{-rt}\sigma S_t dW_t^Q = \sigma(S_t/e^{r\ t})dW_t^Q = \sigma Y_t dW_t^Q.$$

Hence

$$dY_t = 0dt + \sigma Y_t dW_t^Q$$

so the drift is indeed 0 and $Y = S_t/B_t$ is a martingale under Q.

S_t/B_t is a martingale under $\mathbb{Q}$ I

The above proof is not limited to Black Scholes, but to any stock model (see later smile models) for dS where we can find a measure Q such that S/B is a martingale. This is equivalent to put in our chosen stock price model the drift rS_t under Q. We prove this now.
Let's choose the SDE $dS_t = \mu^P(t, S_t)dt + v(t, S_t)dW_t^P$ under the measure P for our stock price and assume the Q measure SDE for the same dS_t is $dS_t = \mu^Q(t, S_t)dt + v(t, S_t)dW_t^Q$ (this could be one of the smile models we'll look at later, Bachelier, CEV, DD, MDD, Heston).
Let's impose that $S_t/B_t = e^{-rt}S_t$ is a martingale under Q and see what we get.

$$d\frac{S_t}{B_t} = d(e^{-rt}S_t) = -re^{-rt}S_tdt + e^{-rt}dS_t =$$

$$= -re^{-rt}S_tdt + e^{-rt}[\mu^Q(t, S_t)dt + v(t, S_t)dW_t^Q]$$

$$= e^{-rt}(\mu^Q(t, S_t) - rS_t)dt + e^{-rt}v(t, S_t)dW_t^Q.$$

S_t/B_t is a martingale under $\mathbb{Q}$ II

Now, we want this to be a martingale, which means that the drift of the SDE must be zero: $\mu^Q(t, S_t) - rS_t = 0 \;\Rightarrow\; \mu^Q(t, S_t) = rS_t$. Thus, for any model for dS_t following a regular SDE, whatever the volatility function $v(t, S_t)$, the condition for S/B to be a martingale under Q is that the drift under Q of the SDE is rS_t. In other words, the SDE for pricing under this dS model must be

$$dS_t = rS_t dt + v(t, S_t) dW_t^Q.$$

Then using the theorem "Pricing with discounted $\mathbb{Q}$ expectations" we introduced earlier, we have that the price at time t of any contingent claim with payoff $f(S_T)$ under the model dS_t above, is the $\mathbb{Q}$ discounted expectation of the payoff $f(S_T)$:

$$\boxed{V_t = E_t^Q[e^{-r(T-t)} f(S_T)], \;\; dS_t = rS_t dt + v(t, S_t) dW_t^Q.} \qquad (29)$$

The idea behind the martingale approach

Numeraire

When we consider S_t/B_t we may say that we are looking at S measured with respect to the numeraire B_t.
In general it is possible to adopt any non-dividend paying asset price as numeraire, and price under the particular probability measure associated with that numeraire. However, the canonical numeraire is the bank account B we have used now and the probability measure associated with the numeraire B is the risk neutral measure $\mathbb{Q}$.

The idea behind the martingale approach

No need to know the real expected return

We noticed earlier that the coefficient μ does not appear in Black Scholes formula (17), indicating that investors, though having different risk preferences or predictions about the future stock price, must yet agree on this unique option price.

This property can also be inferred from (25), the $\mathbb{Q}$ dynamics of S since, under $\mathbb{Q}$, the drift rate of the stock price process equals the risk-free interest rate r rather than μ, while the variance rate is unchanged. For this reason the pricing rule given by discounted E^Q (24) is often referred to as **risk-neutral valuation**, and the measure $\mathbb{Q}$ defines what is called **the risk-neutral world**.

Intuitively, in a risk-neutral world the expected rate of return on all securities is the risk-free interest rate, implying that investors do not require any risk premium for trading stocks.

Weak point of the derivation: Uniqueness of ϕ

The above derivation, however, is still not fully satisfactory, since we have implicitly assumed that (ϕ^B, ϕ^S) is the *unique* self-financing strategy replicating the claim with payoff $f(S_T)$. This uniqueness, anyway, can be obtained by applying the more general theory on complete markets, which is beyond the scope of this introduction.

Zero coupon bonds I

A (default-free) zero coupon bond is the simplest possible contingent claim Y, where at final maturity T the claim pays a fixed amount of currency, the so called bond notional. Assume the bond notional is 1. The bond is called default-free because it is issued by a default-free entity, so that if you buy the bond, you will receive the contingent claim payoff $Y = 1$ at the future time T for sure.

Risk neutral pricing tells us that the price at time t for a zero coupon bond with maturity T is

$$P(t,T) = E_t^Q[e^{-r(T-t)}\ Y] = E_t^Q[e^{-r(T-t)}\ 1] = e^{-r(T-t)}.$$

This is what you pay the bank now at time t to receive $Y = 1$ at future time T.

Zero coupon bonds II

In this course r made its first appearance in the Bank account numerarie $dB_t = r\,B_t\,dt$. We assumed it constant and deterministic. In reality, r can be a stochastic process following a SDE, for example Ornstein Uhlenbeck/Vasicek or a Feller square root process/CIR,

$$dr_t = k(\theta - r_t)dt + \sigma dW_t, \text{ or } \text{ (CIR) } dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t$$

and the bond price formula for a bond at time t with maturity T would then be

$$P(t,T) = E_t^Q[e^{-\int_t^T r_u\,du}\,1] =: E_t^Q[D(t,T)], \quad D(t,T) = e^{-\int_t^T r_u\,du}.$$

If interest rates are stochastic, then $D(t,T)$ is random at time t and is called a stochastic discount factor. $P(t,T)$, being a t expectation of D, is known and not random at time t.

Zero coupon bonds III

Term structure modeling, or modeling of interest rate dynamics dr_t, is beyond the scope of this course. We deal with it in the MSc in Mathematics and Finance. Here we keep assuming that r is constant and deterministic, leading to

$$D(t, T) = P(t, T) = e^{-r(T-t)}.$$

Note that if the bond notional is N instead of one then the bond price is

$$E_t^Q[e^{-r(T-t)} N] = e^{-r(T-t)} N = NP(t, T)$$

and in particular at time 0 we get

$$P(0, T) = e^{-rT}, \quad E_0^Q[e^{-rT} N] = e^{-rT} N = NP(0, T).$$

Forward contracts I

The contingent claim with payoff $Y = S_T - K$ is called a forward contract on the stock S with maturity T and strike K.
The price of this forward contract at a time $t < T$ is

$$V_{FWD}(t, T, S_t, K, r) = e^{-r(T-t)} E_t^Q[S_T - K] = e^{-r(T-t)}(E_t^Q[S_T] - K) = \cdots$$

Let's calculate $E_t^Q[S_T]$. We don't need a model to do that but only the property that S/B must be a martingale under the probability measure Q. This implies, from the definition of martingale,

$$E_t^Q\left[\frac{S_T}{B_T}\right] = \frac{S_t}{B_t}.$$

Recalling that $B_T = e^{rT}$ and $B_t = e^{rt}$ this gives $E_t^Q[S_T] = S_t e^{r(T-t)}$. So

$$\cdots = V_{FWD}(t, T, S_t, K, r) = S_t - e^{-r(T-t)}K.$$

Forward contracts II

Note that we didn't postulate any model for dS to get this but only the martingale property of S/B, which is a general no-arbitrage property of any model.

Hence the forward contract is a contract whose valuation is model-independent. It does not matter which model we will use for dS (Black-Scholes, or models we will see later like Displaced Diffusion, CEV, mixture dynamics...), the price of the forward contract will not depend on the model but will always be $S_t - e^{-r(T-t)}K$. In particular, it does not depend on the volatility.

The forward stock price at time t for maturity T is the value of the strike K for which the forward contract price at time t for maturity T is zero. Namely we solve in K

$$V_{FWD}(t, T, S_t, K, r) = S_t - e^{-r(T-t)}K = 0 \;\Rightarrow K = S_t e^{r(T-t)} =: F_{t,T}.$$

Forward contracts III

We note that *the forward price is a martingale under the measure Q*. Indeed, differentiating wrt t we get

$$dF_{t,T} = d(S_t e^{r(T-t)}) = e^{r(T-t)} dS_t + S_t(-r)e^{r(T-t)} dt =$$

$$= e^{r(T-t)}[rS_t dt + \sigma S_t dW_t^Q] - re^{r(T-t)} S_t \ dt = 0 \ dt + e^{r(T-t)} S_t \sigma dW_t^Q$$

or, in short,

$$dF_{t,T} = \sigma F_{t,T} dW_t^Q.$$

Here we used the Black Scholes dynamics for S, but this is not necessary. Any stock price model dS with drift rS would have worked. We have shown that the forward stock price for a given maturity T is a martingale under Q as its SDE has zero drift.

Forward contracts IV

We can also check the dynamics of $F_{t,T}$ under the measure P, for example in Black Scholes. Let's calculate

$$dF_{t,T} = d(S_t e^{r(T-t)}) = e^{r(T-t)} dS_t + S_t(-r)e^{r(T-t)} dt =$$

$$= e^{r(T-t)}[\mu S_t dt + \sigma S_t dW_t^P] - re^{r(T-t)} S_t\, dt = (\mu - r)e^{r(T-t)} S_t\, dt + e^{r(T-t)} S_t \sigma$$

or, in short,

$$dF_{t,T} = (\mu - r) F_{t,T}\, dt + \sigma F_{t,T} dW_t^P.$$

We have a special case for $t = 0$, and we obtain the price of the forward contract at time 0 for maturity T:

$$V_{FWD}(0, T, S_0, K, r) = S_0 - e^{-rT} K.$$

Forward contracts V

When we introduced the at-the-money-forward call or put options, we said that these are options where the strike K is equal to the forward stock price at time 0 for maturity T, namely $K = F_{0,T} = S_0 e^{rT}$.

This forward stock price is defined as the value of the strike K that makes the price of a forward contract valued at time 0 with maturity T equal zero. In other term, we solve in K

$$V_{FWD}(0, T, S_0, K, r) = S_0 - e^{-rT} K = 0 \;\Rightarrow\; K = S_0 e^{rT} =: F_{0,T}.$$

Put Call parity and Put price I

We have derived the price of a call option, but what about a put option? We could do this through direct integration, as we did with the call, but we will use put call parity instead.

Put call parity is born from the observation that payoff of call minus payoff of put is a straight line, or a “forward contract”. Indeed,

$$(S_T - K)^+ - (K - S_T)^+ = \max(S_T - K, 0) - \max(K - S_T, 0) = S_T - K$$

for all S_T and K . It follows that

$$e^{-rT} E^Q[(S_T - K)^+] - e^{-rT} E^Q[(K - S_T)^+] = e^{-rT} E^Q[S_T - K].$$

or in other terms

CallPrice - PutPrice = ForwardPrice

Put Call parity and Put price II

From which

PutPrice = CallPrice - ForwardPrice

This is true for any model we may use for dS_t. In the specific case of Black Scholes we have

$$V_{BS}^{PUT}(0, S_0, K, T, \nu, r) = \underbrace{S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2)}_{Call\ price} - \underbrace{(S_0 - Ke^{-rT})}_{Forward\ price}$$

or

$$= S_0(\Phi(d_1) - 1) - Ke^{-rT}(\Phi(d_2) - 1) = \cdots$$

Now note that $\Phi(-x) = 1 - \Phi(x)$ so that $\Phi(x) - 1 = -\Phi(-x)$ leading to

$$\cdots = Ke^{-rT}\Phi(-d_2) - S_0\Phi(-d_1)$$

Put Call parity and Put price III

and we conclude

$$V_{BS}^{PUT}(0, S_0, K, T, \nu, r) = Ke^{-rT}\Phi(-d_2) - S_0\Phi(-d_1)$$

with the usual expressions for d_1 and d_2 we have seen in the case of the call.

Finally, we compute the delta of a put option in Black-Scholes.

$$\Delta_0^{PUT} = \frac{\partial V_{BS}^{PUT}(0, S_0, K, T, \nu, r)}{\partial S_0} = \ldots$$

Here we use put-call parity. From PutPrice = CallPrice - ForwardPrice we get

$$\ldots = \frac{\partial[V_{BS}^{CALL}(0, S_0, K, T, \nu, r) - (S_0 - Ke^{-rT})]}{\partial S_0} =$$

Put Call parity and Put price IV

$$= \frac{\partial V_{BS}^{CALL}(0, S_0, K, T, \nu, r)}{\partial S_0} - \frac{\partial (S_0 - Ke^{-rT})}{\partial S_0} = \Delta_{call} - 1 = \Phi(d_1) - 1$$

$$= -\Phi(-d_1).$$

Hence the Delta of a put option in Black Scholes is $-\Phi(-d_1)$ and is always negative. The put price will decrease when the underlying stock S_0 increases.

Dynamic Hedging I

In the process of deriving the BS formula, we have also found a way to perfectly hedge the risk embedded in this contract.

Indeed look at the option pricing problem from the following point of view:

- You are the bank and you just sold a call option to the client.
- At the future time T you will have to pay $(S_T - K)^+$ to your client
- You client pays you V_0 for the option now, at time 0
- Clearly, if the equity goes up a lot in the future, $(S_T - K)^+$ could be very large
- You wish to avoid any risks and decide to hedge away the risk in this contract you sold.
- How should you do that?

Dynamic Hedging II

The answer to this question is in our derivation above.

- You cash in V_0 from the client and use it to buy, at time 0,

$$\frac{\partial V_0}{\partial S_0} = \Phi(d_1(0)) =: \phi_0^S =: \Delta_0 \quad \text{stock and}$$

$$\phi_0^B = (V_0 - \Delta_0 S_0)/B_0 \quad \text{bank account / bond (cash).}$$

- You then implement the self-financing trading strategy, **rebalancing continuously** (hence *dynamic hedging*) your ϕ_t^S, ϕ_t^B amounts of S and B according to

$$\phi_t^S = \frac{\partial V_t}{\partial S_t} = \Phi(d_1(t)) =: \Delta_t \quad \text{stock and}$$

$$\phi_t^B = (V_t - \Delta_t S_t)/B_t \quad \text{bank account / bond (cash).}$$

Dynamic Hedging III

- Because the strategy is self-financing, this rebalancing can be financed thanks to price movements of B and S and you need not add any cash or assets from outside.
- At final maturity we know that the final value will be $V_T = (S_T - K)^+$ as we posed this as boundary condition in our pricing problem.
- Hence by following the above strategy, set up with the initial V_0 and with no subsequent cost, we end up with the payout $(S_T - K)^+$ at maturity.
- We can then deliver this payout to our client and face no risk.
- Basically, our self financing trading strategy in the underlying S, set up with the initial payment V_0, completely replicated the claim we sold to our client.

Dynamic Hedging IV

- An obvious but often overlooked point it this: If we are perfectly hedged, all the money we received from the client (V_0) is spent to set up the hedge, and we as a bank make no gain.
- That's why in reality only partial hedges are often implemented, in an attempt not to erode all potential profit.

The above framework is called **"delta-hedging"**.

Basically one holds an amount of risky asset equal to the sensitivity of the contract price to the risky asset itself (delta).

This strategy is possible only in markets where all risks are directly linked to tradable assets and viceversa (roughly: "complete markets").

Incomplete Markets I

Metatheorem/folklore: A market is complete if there are as many assets as independent sources of randomness.

In reality markets are incomplete, as there are some risks that are covered by no direct assets, and there are more risks than assets.

This can be partly addressed by including a few derivatives themselves among the basic assets, but it is hard to keep the market complete

For example, as we will see in the volatility smile part, in a stochastic volatility model like Heston for the stock price S_t under the measure Q,

$$dS_t = rS_t dt + \sqrt{V_t} S_t dW_t, \;\; s_0, \quad dW dW^V = \rho \, dt$$

$$dV_t = k(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_t^V, \;\; v_0,$$

Incomplete Markets I

$$dS_t = rS_t dt + \boxed{\sqrt{V_t}} S_t dW_t, \quad s_0, \quad dW dW^V = \rho \, dt$$

$$dV_t = k(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_t^V, \quad v_0,$$

we have that now the volatility (see box) in the stock equation, namely $\sqrt{V_t}$, is a second stochastic differential equation driven by a second Brownian motion W^V. In Black Scholes the box would have a deterministic constant σ.

If we hedge only with the stock price S_t, delta hedging does not work because the risk associated with the randomness of the volatility is not covered by the stock, the stock is one asset and can only cover one risk, the risk of W, but not the risk of W^V.

Thus, if our only hedging risky asset is the stock S, in a Heston model the market is incomplete. To make the market complete we need to add another asset to the fundamental assets we start from.

Incomplete Markets II

For example, a specific call option $\bar{C}$ with a given strike $\bar{K}$ and maturity $\bar{T}$ could be added to B_t and the stock, and the market would be complete again, because we would have two risky assets now, S_t and $\bar{C}_t$, to hedge two sources of risk, W and W^V.
A trading strategy would have to be a triple now, $(\phi^B, \phi^S, \phi^{\bar{C}})$.

In reality it's not always possible to find a risky asset matching a given risk, this is partifularly difficult or impossible for some credit risk, liquidity risk, operational risks, etc. Real market remains incomplete.

A further problem is that continuous rebalancing does not happen. Real hedging happens in discrete time and this will imply an hedging error with respect to the idealized case

Incomplete Markets III

In the end hedging is more an art than a science, and it involves many pragmatic choices and rules of thumbs. However, a sound understanding of the idealized case is crucial to appreciate the subtleties in real market applications.

“Greeks” (sensitivities) are often used to deal with hedging, and we briefly look at them now.

The sensitivities (or greeks) I

When hedging derivatives, often sensitivities (or greeks) are used in practice.

A sensitivity is the partial derivative of the price or of another sensitivity with respect to one of the parameters. It tells us how much a small change of the parameter impacts a change in the price or sensitivity we are examining.

We have already met one of the most important sensitivities, delta.

$$\Delta(t) = \frac{\partial V(t)}{\partial S_t},$$

which, for a call option price under Black Scholes, is equal to $\Phi(d_1(t))$, as we have seen above. Delta measures how much the option price V changes when there is a small change in the underlying asset price S.

The sensitivities (or greeks) II

An important point to keep in mind about the greeks is that they are sensitivities only to the chosen parameter, and that they are partial derivatives, not total derivatives. In other terms, assume that the strike of a call option is a function of S_0, e.g. $K = S_0$. In this case the Black Scholes formula for the price at time 0 simplifies, as for example the term $\ln(S_0/K)$ becomes 0. However, if we are calculating the delta of a call at time 0 for $K = S_0$, we need to differentiate the call price keeping K general, obtaining $\Phi(d_1(K))$, and then substitute $K = S_0$ after the differentiation has been made. Instead, if we differentiate the price with $K = S_0$ already plugged in, we are mixing sensitivity to the initial stock with sensitivity to the strike, and this is a mistake.
Let's see the two different approaches for a call at time 0 with $K = S_0$.
Correct approach:

$$\Delta_0 = \frac{\partial V_{BS}(0, S_0, K, T, \sigma, r)}{\partial S_0}|_{K=S_0} = \Phi(d_1(0, S_0, K, T, \sigma, r))|_{K=S_0} =$$

The sensitivities (or greeks) III

$$= \Phi\left(\frac{\ln(S_0/K) + (r+\sigma^2/2)T}{\sigma\sqrt{T}}\right)\Big|_{K=S_0} =$$

$$= \Phi\left(\frac{\ln(S_0/S_0) + (r+\sigma^2/2)T}{\sigma\sqrt{T}}\right) = \Phi\left(\frac{(r+\sigma^2/2)\sqrt{T}}{\sigma}\right).$$

Wrong approach:

$$\Delta_0 = \frac{\partial V_{BS}(0, S_0, S_0, T, \sigma, r)}{\partial S_0} =$$

$$= \frac{\partial}{\partial S_0}\left(S_0\Phi\left(\frac{\ln\frac{S_0}{S_0} + (r+\sigma^2/2)T}{\sigma\sqrt{T}}\right) - S_0 e^{-rT}\Phi\left(\frac{\ln\frac{S_0}{S_0} + (r-\sigma^2/2)T}{\sigma\sqrt{T}}\right)\right)$$

$$= \frac{\partial}{\partial S_0}\left(S_0\Phi\left(\frac{(r+\sigma^2/2)\sqrt{T}}{\sigma}\right) - S_0 e^{-rT}\Phi\left(\frac{(r-\sigma^2/2)\sqrt{T}}{\sigma}\right)\right)$$

The sensitivities (or greeks) IV

$$= \Phi\left(\frac{(r+\sigma^2/2)\sqrt{T}}{\sigma}\right) - e^{-rT}\Phi\left(\frac{(r-\sigma^2/2)\sqrt{T}}{\sigma}\right)$$

In the second case we differentiated wrt S_0 the Black Scholes formula with the strike $K = S_0$ already plugged in before differentiation, and this gives us the joint sensitivity to stock and strike. This is not what we want. The strike is fixed at time 0 and is S_0, but it won't change as times moves on. We wish to see how the option changes when the stock changes, but not the strike. So the first method is correct, the second is wrong.

Going back to greeks in general, a large sensitivity with respect to a parameter means that the trade is quite sensitive to that parameter, and the trader may consider trades that reduce the sensitivity if she wishes to be more prudent with respect to that parameter. If the trader

The sensitivities (or greeks) V

is more aggressive she may decide to trade to increase the sensitivity further.

Other sensitivities or greeks are: Time decay or Θ, negative sensitivity to time to expiry,

$$\Theta_t = -\frac{\partial V(t)}{\partial (T-t)} = \frac{\partial V(t)}{\partial t}, \quad \Theta_t^{Call-BS} = -\frac{S_t \phi(d_1(t))\sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}\Phi(d_2(t)),$$

where ϕ is the standard normal density function, and we had already seen the formula for Θ of a call option in Black Scholes (without proof) when we verified the self-financing condition in the Black-Scholes derivation, see Equation (21). You will notice from the equation that Θ is always negative for a call option. The direct calculation of the partial derivative would give a much long expression, but using Lemmas like Eq. (20), the expression simplifies in the above one. The fact that the

The sensitivities (or greeks) VI

theta is negative for a call means that if you simply hold the option without doing anything and nothing moves but time, you lose money as the option value $V(t)$ decreases in time. This means that for holding the call to be profitable, includinig the cost $V(0)$ you paid at time 0, you are relying on other variables to move, primarily S of course.

Gamma, the sensitivity of delta to the underlying:

$$\Gamma_t = \frac{\partial \Delta(t)}{\partial S} = \frac{\partial^2 V(t)}{\partial S^2}, \quad \Gamma_t^{BS-Call} = \frac{\phi(d_1(t))}{S_t \sigma \sqrt{T-t}},$$

We have seen the expression for gamma also in the Black Scholes derivation when we verified the self-financing condition, see again Equation (21), and note from there that for a call option Gamma is

The sensitivities (or greeks) VII

always positive. At this point we may write an equation linking the three sensitivities just introduced. Recall Ito's formula we have seen earlier

$$dV(t,S_t) = \frac{\partial V}{\partial t}(t,S_t)dt + \frac{\partial V}{\partial S}(t,S_t)dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(t,S_t)\sigma^2 S_t^2 dt.$$

We can rewrite this as

$$dV(t,S_t) = \Theta_t dt + \frac{1}{2}\Gamma_t \sigma^2 S_t^2 dt + \Delta_t dS_t$$

For a call option, Θ is negative, so the option position loses value in time. Γ is positive, so Γ may counterbalance Θ if the market moves considerably in S. In theory, this is still the Δ hedging equation, so in continuous time all should work as a perfect hedge, but in practice hedging happens in discrete time and Gamma / Theta effects show up and need to be taken into account.

The sensitivities (or greeks) VIII

Back to definitions, Vega is the sensitivity to volatility, namely

$$\nu_t = Vega_t = \frac{\partial V(t)}{\partial \sigma}.$$

We calculate the vega for a call or put option in Black Scholes at time 0, as this is often used by traders.

$$Vega_{BS}^{Call} = \frac{\partial V_{BS}^{Call}(0)}{\partial \sigma} = \frac{\partial}{\partial \sigma}\Big(S_0\Phi(d_1(\sigma)) - Ke^{-rT}\Phi(d_2(\sigma))\Big) =$$

$$= S_0\left(\frac{\partial}{\partial \sigma}\Phi(d_1(\sigma))\right) - Ke^{-rT}\frac{\partial}{\partial \sigma}\left(\Phi(d_2(\sigma))\right) = \dots$$

as the only terms depending on σ are $\Phi(d_1)$ and $\Phi(d_2)$.

$$\frac{\partial}{\partial \sigma}\Phi(d_1) = \phi(d_1)\frac{\partial}{\partial \sigma}d_1 = \phi(d_1)\frac{\partial}{\partial \sigma}\left(\frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}\right)$$

The sensitivities (or greeks) IX

$$= \phi(d_1)\frac{(\sigma^2/2 - r)T - \ln(S_0/K)}{\sigma^2\sqrt{T}}$$

where ϕ is the PDF of the standad normal, $\Phi' = \phi$. Similarly,

$$\frac{\partial}{\partial\sigma}\Phi(d_2) = \phi(d_2)\frac{\partial}{\partial\sigma}d_2 = \phi(d_2)\frac{\partial}{\partial\sigma}\left(\frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}\right) =$$

$$= \phi(d_2)\frac{-\ln(S_0/K) - (r + \sigma^2/2)T}{\sigma^2\sqrt{T}}$$

The call vega is thus

$$Vega_{BS}^{Call} = S_0\phi(d_1)\frac{(\frac{\sigma^2}{2} - r)T - \ln\frac{S_0}{K}}{\sigma^2\sqrt{T}} - Ke^{-rT}\phi(d_2)\frac{-\ln\frac{S_0}{K} - (r + \frac{\sigma^2}{2})T}{\sigma^2\sqrt{T}}$$

The sensitivities (or greeks) X

$$= S_0\phi(d_1)\frac{(\frac{\sigma^2}{2} - r)T - \ln\frac{S_0}{K}}{\sigma^2\sqrt{T}} + Ke^{-rT}\phi(d_2)\frac{\ln\frac{S_0}{K} + (r + \frac{\sigma^2}{2})T}{\sigma^2\sqrt{T}}$$

We can simplify this expression further by using the **Greeks Lemma** seen earlier, Eq. (20) at $t = 0$. A consequence is that Vega at time 0 above, can be simplified into

$$\boxed{Vega_{BS}^{Call} = S_0\phi(d_1)\sqrt{T}.}$$

This formula can be memorized for calculating Vega of combinations of options in Black Scholes. We also notice that from the last formula, $Vega_{BS}^{Call} > 0$, so the vega of a call is always positive. The value of a call option always increases with the volatility in the Black Scholes model.

The sensitivities (or greeks) XI

By put call parity, if the two options have the same strike K, same stock, same r, same maturity T, we know that

$$V_{Call}(0,\sigma) - V_{Put}(0,\sigma) = S_0 - Ke^{-rT}$$

Differentiating both sides with respect to σ

$$\frac{\partial V_{Call}(0,\sigma)}{\partial\sigma} - \frac{\partial V_{Put}(0,\sigma)}{\partial\sigma} = 0$$

as the right hand side does not depend on σ. We conclude that

$$Vega_{Call} = Vega_{Put}.$$

This is true for Black Scholes but for other models as well, since put call parity is model independent. In Black Scholes in particular, given

The sensitivities (or greeks) XII

that the Vega of a call is always positive, the same result on positivity will hold for the Vega of a put. F

Another greek is ρ, which is the sensitivity to interest rates r, namely $\rho_t = \frac{\partial V(t)}{\partial r}$. We give a formula for ρ at time 0 without proof for a call option in Black Scholes:

$$\rho_0^{BS-Call} = TKe^{-rT}\Phi(d_2) > 0.$$

The value of a call in Black Scholes always increases with the risk free rate r as ρ is positive.

By put-call parity it is easy to get Θ, Γ and ρ in Black-Scholes for put options, starting from their value for call options.

The sensitivities (or greeks) XIII

Concluding, these greeks can be computed in closed form in Black Scholes for call and put options, for example, and we have seen some examples above. There are further higher order greeks Vanna, Charm, Vomma/volga, Veta, Yoghurt, Speed, Zomma, Color Ultima, Totto... (sounds crazy I know... and one on this list is fake)

The higher the order of the greeks we use, the smoother we are assuming prices to be. For example, Speed = $\partial^3 V/\partial S^3$ requires the price V to be three times differentiable with respect to the underlying S. While this may hold in simple models like Black Scholes for specific payoffs, in general assuming excessive smoothness is not realistic, and therefore using high order greeks has to be done very carefully, especially when the greeks are computed with numerical methods.

Introduction to the volatility smile

The Black-Scholes model (given here under $\mathbb{Q}$, we will write W for W^Q in this part of the volatility smile, so all Brownian motions are under the risk neutral measure)

$$dS_t = rS_t dt + \nu S_t \, dW_t, \quad s_0, \quad t \in [0, T],$$

is a Geometric Brownian motion with density p_{S_t} given by the lognormal density corresponding to

$$\ln S_t \sim \mathcal{N}\left(\ln S_0 + rt - \frac{1}{2}\nu^2 t, \nu^2 t\right), \tag{30}$$

$$\text{or (log return)} \quad \ln \frac{S_t}{S_0} \sim \mathcal{N}\left(rt - \frac{1}{2}\nu^2 t, \nu^2 t\right),$$

which is

$$p_{S_t}(y) = p_{t,\nu}^{lognormal}(y) = \frac{1}{y\nu\sqrt{t\,2\pi}} \exp\left\{-\frac{1}{2\nu^2 t}\left[\ln\frac{y}{S_0} - rt + \tfrac{1}{2}\nu^2 t\right]^2\right\}. \tag{31}$$

Introduction to the volatility smile I

Recall that the price of a European call option with maturity T and strike K paying $Y = (S_T - K)^+$ at time T is

$$V_{BS}(0, S_0, K, T, \nu, r) = E_0^Q[e^{-rT}(S_T - K)^+] =$$

$$= S_0\Phi(d_1(0)) - Ke^{-rT}\Phi(d_2(0)),$$

where

$$d_1(0) := \frac{\ln(S_0/K) + (r + \nu^2/2)T}{\nu\sqrt{T}}, \quad d_2(0) := d_1(0) - \nu\sqrt{T},$$

In particular ν is the volatility of the option and does not depend on K. Important: Volatility is a characteristic of stock S underlying the contract, and has nothing to do with the contract on S and, in particular, it has nothing to do with the strike K of the option we choose to trade.

Introduction to the volatility smile

Now take two different strikes K_1 and K_2. Suppose that the market provides us with the prices of the two call options $\text{MKTCall}(S_0, K_1, T)$ and $\text{MKTCall}(S_0, K_2, T)$.

Does the market follow Black & Scholes formula in a consistent way?

Does there exist a *single* volatility ν such that

$$\text{MKTCall}(S_0, K_1, T) = V_{BS}(0, S_0, K_1, T, \nu, r),$$

$$\text{MKTCall}(S_0, K_2, T) = V_{BS}(0, S_0, K_2, T, \nu, r)?$$

If Black and Scholes is correct, this should happen.
The answer is a resounding **“NO!!!”**
Market option prices do not behave like this.

Introduction to the volatility smile I

Instead two *different* **implied volatilities** $\nu(K_1)$ and $\nu(K_2)$ are required to match the observed market prices if one is to use the Black Scholes formula:

$$\mathrm{MKTCall}(S_0, K_1, T) = V_{BS}(0, S_0, K_1, T, \nu(K_1), r),$$

$$\mathrm{MKTCall}(S_0, K_2, T) = V_{BS}(0, S_0, K_2, T, \nu(K_2), r).$$

In other terms, each market option price requires its own Black and Scholes **implied volatility** $\nu(K)$ depending on the option strike K.

The market therefore uses BS formula simply as a *metric* to express option prices as volatilities. The curve $K \mapsto \nu(K)$ is the so called *volatility smile* of the T-maturity option.

Introduction to the volatility smile

If Black and Scholes model were consistent along different strikes, this curve would be flat, since volatility should not depend on K. Instead, this curve is commonly seen to exhibit "smiley" or "skewed" shapes.

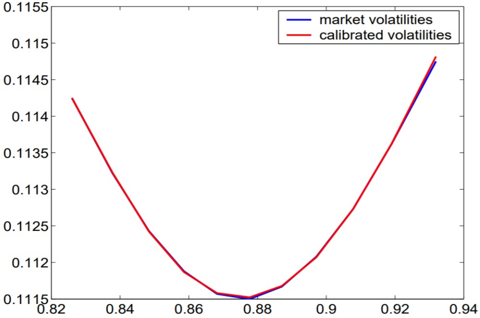


Figure: An example of smile curve $K \mapsto \nu(K)$ from the FX market. K is on the horizontal axis, $\nu(K)$ on the vertical axis

Introduction to the volatility smile

This means that the Black Scholes formula is wrong because, again, ν is a property of S and not of the option I decide to write on S. It should not depend on K if the model were right.

So the solution, when the smile showed up, should have been to ditch the Black Scholes formula looking for a new formula coming from a better model, more in line with market option prices patterns

However, traders were so used to calculate prices and sensitivities (greeks) with Black Scholes that they insisted in retaining the option formula even if the model was wrong. This led to the above definition of implied volatility where traders are willing to change ν when K changes, even for the same S.

Still, even if we keep the Black-Scholes formula on the surface, the real model behind the option prices will be different than Black Scholes. We will explore some alternative models now.

Smile modelling through different SDEs I

Once we understand that the Black-Scholes model cannot be correct if the volatility of an option depends on its strike, we may try to use a different model to see if this different model can account for the drawbacks of the Black - Scholes model.

The alternative models will be able to generate a volatility smile that can be close enough to the market smile curve for practical purposes.

Some terminology first. When we write a model like

$$dS_t = rS_t dt + b(t, S_t) dW_t^Q,$$

then $b(t, S)$ is called the *absolute volatility*. This is in contrast to the *relative or return volatility* $\sigma(t, S)$ defined in

$$dS_t = rS_t dt + \sigma(t, S_t) S_t dW_t^Q.$$

Smile modelling through different SDEs II

The name return volatility comes from the fact that if we write the SDE for the log return $d\ln(S_t/S_0)$, this will have volatility $\sigma(t, S_t)$ or $b(t, S_t)/S_t$.

In the Black Scholes model $dS_t = rS_t dt + \nu S_t dW_t^Q$ with constant ν, the absolute volatility is νS, whereas the relative or return volatility is ν.

Let's look at how an alternative model for the stock dynamics dS_t with relative volatility $\sigma(t, S_t)$ can generate a smile.

Smile modelling through different SDEs III

Alternative SDE model for dS_t can generate a non-flat smile:

1. Set K to a starting value;
2. Compute the model call option price

$$V_{Model}(K) = E_0^Q[e^{-rT}(S_T - K)^+]$$

   with S modeled through an alternative dynamics

$$\text{Model:}\;\; dS_t = rS_t dt + \boxed{\sigma(t, S_t)} S_t \, dW_t^Q, \quad S_0 = s_0$$

3. Invert Black Scholes formula for this strike, i.e. solve

$$V_{Model}(K) = V_{BS}(0, S_0, K, T, \nu(K), r).$$

   in $\nu(K)$, thus obtaining the model implied volatility $\nu(K)$.
4. Change K and restart from point 2.

At the end of this algorithm we have built the smile curve $K \mapsto \nu(K)$ for this model.

Pricing with risk neutral expectations of discounted payoffs

We used the formula $V_{Model}(0, K) = E_0^Q[e^{-rT}(S_T - K)^+]$ for a call option price at time 0 for a Q dynamics
$dS_t = rS_t dt + \boxed{\sigma(t, S_t)} S_t \, dW_t^Q, \quad S_0 = s_0$ that was not Black Scholes.
For this Q-dynamics, S_t/B_t is a martingale, as we saw earlier in the program, as $d(S_t/B_t)$ has zero drift. So, by “the fundamental theorem of asset pricing”, the model is arbitrage free.
More generally, the price at time t of any simple claim $f(S_T)$ under the new dynamics will be

$$V_{Model}(t, K) = E_t^Q[e^{-r(T-t)} f(S_T)].$$

This holds based on Eq. (29) for $v(t, S_t) = \sigma(t, S_t)S_t$ and is based on the theorem “Pricing with discounted $\mathbb{Q}$ expectations”. We will use this in all the following smile models.

Bachelier Model I

We now start exploring alternative SDEs for dS_t that can lead to a volatility smile.

We begin with the model that was proposed in 1900 by Louis Bachelier, PhD student of Henry Poincare. Bachelier is definitely one of the key precursors of the mathematics of option pricing. He is also among the first to study Brownian motion mathematically, anticipating Einstein's 1905 and 1908 famous papers.

At the time Bachelier proposed his model, the no-arbitrage theory was not there. So he did not know that his model should have had as a drift the quantity rS_t. Instead, he worked under the measure P and proposed the model

$$dS_t = \mu\, dt + \sigma dW_t^P.$$

Bachelier Model II

This is an arithmetic Brownian motion. If we were to impose the risk neutral drift rS_t to this process, following the modern theory of no-arbitrage, we would get

$$dS_t = \mu \, dt + \sigma dW_t^P \ \rightarrow \ dS_t = rS_t \, dt + \sigma dW_t^Q.$$

However, this would be a special case of the Ornstein Uhlenbeck process rather than an arithmetic Brownian motion (ABM). Furthermore, with such a change of drift the Radon Nykodym derivative is not guaranteed to exist.

To bypass all these problems, we make the assumption that interest rates are zero. $r = 0$. This way we avoid drift problems related to the

Bachelier Model III

specific shape of the Bachelier model and the Bachelier model can remain an ABM also under $\mathbb{Q}$.

$$\text{Bachelier Model (BaM) } (r = 0) \;\; dS_t = \sigma dW_t, \;\; s_0.$$

The price of a call option is

$$V_{BaM} = E^Q\left[(S_T - K)^+\right].$$

In the BaM we have, integrating $dS_t = \sigma dW_t$,

$$S_T = s_0 + \sigma W_T = s_0 + \sigma\sqrt{T}\mathcal{N}(0,1)$$

where $\mathcal{N}(0,1) =: N$ is a standard normal random variable. It follows that

$$E^Q[(S_T - K)^+] = E^{Q^T}[\left(s_0 + \sigma\sqrt{T}N - K\right)^+] =$$

Bachelier Model IV

$$= E^{Q^T}\left(\sigma\sqrt{T}N - (K - s_0)\right)^+ = \sigma\sqrt{T}E^Q\left(N - \frac{K - s_0}{\sigma\sqrt{T}}\right)^+$$

$$= \sigma\sqrt{T}\int_{-\infty}^{+\infty}\left(x - \frac{K - s_0}{\sigma\sqrt{T}}\right)^+ p_N(x)dx =$$

where as usual Φ is the standard normal CDF and p_N is the standard normal PDF. Now let $y = \frac{K-s_0}{\sigma\sqrt{T}}$ so that

$$= \sigma\sqrt{T}\int_{-\infty}^{+\infty}(x - y)^+ p_N(x)dx = \sigma\sqrt{T}\int_{y}^{+\infty}(x - y)\,p_N(x)dx =$$

since the positive part is non-zero only for $x > y$. Now

$$= \sigma\sqrt{T}\left[\int_{y}^{+\infty} x p_N(x)dx - y\int_{y}^{+\infty} p_N(x)dx\right] =$$

Bachelier Model V

The first integral is trivial, remembering that

$$\frac{d}{dx}[-p_N(x)] = \frac{d}{dx}\left[-\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right] =$$

$$= -\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\frac{d}{dx}(-x^2/2) = x\frac{1}{\sqrt{2\pi}}e^{-x^2/2} = xp_N(x)$$

so that we find that a primitive of the integrand $xp_N(x)$ is $-p_N(x)$:

$$\int_y^{+\infty} xp_N(x)dx = -p_N(x)|_y^{+\infty} = -0 + p_N(y) = p_N(y).$$

The second integral is simply

$$\int_y^{+\infty} p_N(x)dx = 1 - \Phi(y) = \Phi(-y).$$

Bachelier Model VI

Then remembering the definition of y and substituting we have

$$V_{BaM}(0, s_0, K, T, \sigma) = (s_0 - K)\Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}p_N\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right).$$

The smile $K \mapsto \nu(K)$ in the BaM is given, for a fixed σ, as the solution $\nu(K)$ of the following equation for seral values of K:

$$V_{BS}(0, s_0, K, T, \nu(K), r)|_{r=0} = V_{BaM}(0, S_0, K, T, \sigma).$$

The smile is monotonically decreasing, similarly to the smile of the displaced diffusion model we see next when $\alpha < 0$ in that model.

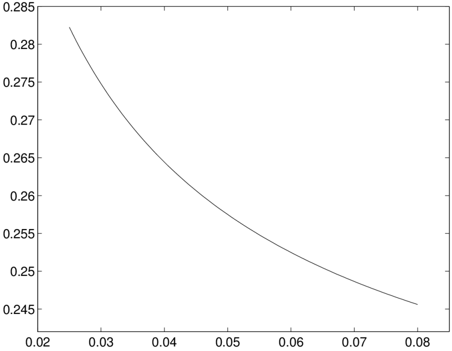
We mentioned that the Bachelier model for the stock S_t can only be used when $r = 0$, as this is the only case where, under both measures, the process remains an arithmetic Brownian motion. However, there is

Bachelier Model VII

another possibility. One could model forward stock prices $F_{t,T}$, which are known to have zero drift under Q as they are martingales, and the zero drift property would be true by definition.
One then expresses the volatility smile for the forward stock price $F_{t,T}$ instead of the stock price itself S_t.
This is occasionally done, but in practice when using Bachelier we will assume $r = 0$ and use the stock price S.

Example of qualitative pattern of the smile for the Bachelier model (call options on interest rates, caplet)

Bachelier Model VIII



In the horizontal axis we have K, in the vertical axis $\nu(K)$.

Bachelier Model IX

It is also interesting to calculate the Delta of a call option in the Bachelier model and see how it compares to Black Scholes. We compute the delta now, namely the partial derivative of the Bachelier call price with respect to the deterministic initial condition $S_0 = s_0$.

$$\frac{\partial V_{BaM}(0)}{\partial s_0} = \frac{\partial\left((s_0 - K)\Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)\right)}{\partial s_0} + \sigma\sqrt{T}\frac{\partial\, p_N\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)}{\partial s_0}.$$

We compute the first partial derivative first.

$$\frac{\partial\left((s_0 - K)\Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)\right)}{\partial s_0} = \Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right) + \frac{s_0 - K}{\sigma\sqrt{T}} p_N\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right)$$

Bachelier Model X

As for the second term,

$$\frac{\partial\, p_N\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)}{\partial s_0} = -\frac{s_0-K}{\sigma\sqrt{T}}\frac{1}{\sigma\sqrt{T}}p_N\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)$$

where we used $p_N'(y) = \frac{1}{\sqrt{2\pi}}\exp(-y^2/2)(-y) = -yp_N(y)$ as the first derivative of the normal PDF. Putting together the pieces we get

$$\frac{\partial V_{BaM}(0)}{\partial s_0} = \Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right) + \frac{s_0-K}{\sigma\sqrt{T}}p_N\left(\frac{s_0-K}{\sigma\sqrt{T}}\right) - \frac{s_0-K}{\sigma\sqrt{T}}p_N\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)$$

$$\frac{\partial V_{BaM}(0)}{\partial s_0} = \Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right). \tag{32}$$

Bachelier Model XI

As already happened for Black Scholes, also for Bachelier the delta can be computed by pretending the only s_0 is the one in the box and differentiating with respect to that, ignoring the other s_0's.

$$\frac{\partial V_{BaM}(0)}{\partial s_0} = \frac{\partial\left(\left(\boxed{s_0} - K\right)\Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)\right)}{\partial s_0} + \sigma\sqrt{T}\frac{\partial\, p_N\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)}{\partial s_0}.$$

The reason is the same as in Black Scholes but it's too subtle to be explained here.

Displaced diffusion model I

We continue exploring alternatives to the Black Scholes SDE. The second model we consider is the displaced diffusion model (DDM). In this model the stock price S under the risk neutral measure is modelled as

$$\text{DDM Def} \;\; S_t = \alpha e^{r\,t} + X_t, \;\; dX_t = rX_t dt + \sigma X_t\, dW_t.$$

To check that this model is indeed arbitrage free we need to see that dS_t has drift $rS_t\, dt$. Calculate

$$dS_t = d(\alpha e^{r\,t}) + dX_t = r\alpha e^{r\,t} dt + rX_t dt + \sigma X_t\, dW_t.$$

Now, substitute $X_t = S_t - \alpha e^{r\,t}$ in this last equation to get

$$dS_t = r\alpha e^{r\,t} dt + r(S_t - \alpha e^{r\,t}) dt + \sigma(S_t - \alpha e^{r\,t})\, dW_t.$$

Displaced diffusion model II

Simplifying terms, we end up with the SDE

$$\text{DDM} \quad \boxed{dS_t = rS_t\,dt + \sigma(S_t - \alpha e^{r\,t})\,dW_t.}$$

We see that the drift is the correct risk neutral drift, so this is arbitrage free.

We also see that if $\alpha = 0$ we get back Black Scholes.

To price a call or put option it's best to resort to Eq. “DDM Def” than “DDM”. Let's see how a call option is priced.

Displaced diffusion model III

$$V_{DDM}(K) = E_0^Q[e^{-rT}(S_T - K)^+] = E_0^Q[e^{-rT}(\alpha e^{r\ T} + X_T - K)^+] =$$

$$= E_0^Q[(X_T - (K - \alpha e^{r\ T}))^+] = E_0^Q[e^{-rT}(X_T - K')^+]$$

where we set $K' = K - \alpha e^{r\ T}$. Now note that X is just a Black Scholes model with volatility σ, so from the last price expectation we have

$$V_{DDM}(K, \sigma) = V_{BS}(0, X_0, K', T, \sigma, r) = X_0 \Phi(d_1(0)) - K' e^{-rT} \Phi(d_2(0)),$$

where

$$d_1(0) := \frac{\ln(X_0/K') + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad d_2(0) := d_1(0) - \sigma\sqrt{T}.$$

Displaced diffusion model IV

Recalling that $X_0 = S_0 - \alpha$, $K' = K - \alpha e^{r\ T}$ and substituting in the last equation we get

$$V_{DDM}(0, S_0, K, T, \sigma, \alpha, r) = (S_0 - \alpha)\Phi(d_1(0)) - (Ke^{-rT} - \alpha)\Phi(d_2(0)),$$

$$d_1(0) := \frac{\ln\left(\frac{S_0-\alpha}{K-\alpha e^{r\ T}}\right) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad d_2(0) := d_1(0) - \sigma\sqrt{T}.$$

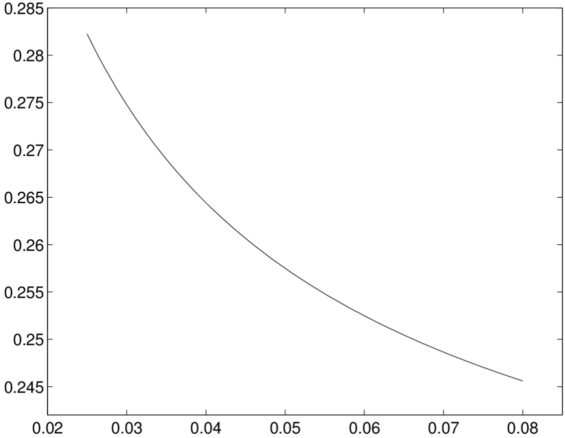
This model generages a smile $K \mapsto \nu(K)$ by assigning the parameters σ and α and solving for all K the following equation in $\nu(K)$:

$$V_{BS}(0, S_0, K, T, \nu(K), r) = V_{DDM}(0, S_0, K, T, \sigma, \alpha, r).$$

Displaced diffusion can generate only monotonically decreasing or increasing smiles (depending on the sign of α). It cannot generate a V

Displaced diffusion model V

shaped smile. Here is an example from the interest rates call options market (caplets)



Displaced diffusion model VI

Introducing $\alpha \neq 0$ has two effects on the smile.

First, it leads to a decreasing ($\alpha < 0$) or increasing ($\alpha > 0$) curve.

Second, it moves the curve upwards ($\alpha < 0$) or downwards ($\alpha > 0$).

More generally, ceteris paribus, increasing α shifts the volatility curve $K \mapsto \nu(K)$ down, whereas decreasing α shifts the curve up.

Shifting a lognormal diffusion can then help in recovering skewed volatility structures. However, such structures are often too rigid, and highly negative slopes are impossible to recover.

Moreover, the best fitting of market data is often achieved for decreasing implied volatility curves, which correspond to negative values of the α parameter, and hence to a support of the stock price density containing unrealistic negative values.

Constant Elasticity of Variance (CEV) model I

The CEV model is given as

$$\text{CEV:} \quad dS_t = r\, S_t dt + \nu S_t^\gamma\, dW_t, \quad S_0 = s_0$$

where γ is a positive exponent, $\gamma > 0$. To avoid issues with explosion (linear growth condition), the exponent should be between 0 and 1, although some empirical studies pointed to an exponent $\gamma = 3/2$, leading to explosion issues.

To use the model safely one needs to assume $\gamma \in (0, 1]$.
There are two special cases.

First, $\gamma = 1/2$ leads to an SDE known as “Feller square-root process”.

$$\text{Feller process:} \quad dS_t = r\, S_t dt + \nu \sqrt{S_t}\, dW_t, \quad S_0 = s_0.$$

Constant Elasticity of Variance (CEV) model II

With $\gamma < 1$ and in the Feller case in paticular, one needs to say what happens at $S = 0$, which is usually taken as an absorbing boundary, meaning that trajectories $t \mapsto S_t(\omega)$ that reach $S = 0$ stay there. The process avoids negative values but can end up in zero. When the stock of a company hits zero, it means the company has defaulted. The model can then be used to model default risk.

For the smile, the model with $\gamma \in (0, 1)$ has a monotonically decreasing smile similar to the smile in the DDM. Here the steepness is mainly decided by γ, whereas in DDM it was mainly decided by α.

Second, if $\gamma = 1$ we get back the Black Scholes model and the smile goes flat at ν.

Constant Elasticity of Variance (CEV) model III

As the CEV model with $\gamma \in (0,1)$ only gives decreasing smiles as the DDM, and calculations require special functions (like Bessel functions) and are much more complicated than in DDM, DDM is usually preferred to CEV for monotonic smiles, unless one insists in keeping S non-negative in all scenarios, in which case CEV can be a better choice than DDM.

Note that if $r = 0$ and if we were allowed to take $\gamma = 0$ we would obtain the Bachelier model as a special case of the CEV model.

Anticipating the smile type based on the return volatility I

The three models we saw so far, Bachelier, DD and CEV, only allow for monotonic smiles (decreasing for Bachelier or CEV with $\gamma < 1$ or DD with $\alpha < 0$ or increasing for CEV with $\gamma > 1$ [problematic/explosion] or DD with $\alpha > 0$). Some traders justify these increasing/decreasing patterns by comparing the return volatilties of the proposed models to the constant return volatilties of Black Scholes. The argument is as follows, but it's a rule of thumb and is not a rigorous argument. The point of our course is also to give a flavour for the type of discussion one may end up having based on rules of thumbs and primitive intuition, while being careful of the claims made. Let's start with Bachelier.

Anticipating the smile type based on the return volatility II

Let's start with Bachelier. Claim: Bachelier return volatility is decreasing while Black scholes is constant, so we expect a decreasing smile curve $K \mapsto \nu(K)$ for Bachelier.
Reasoning. Basically let's say $r = 0$ and let's price an option with the two different models

$$dS = \nu S dW, \quad dS = \sigma \, dW$$

with σ and ν constants. Let us compare the relative/return volatility in the two models,

$$dS = \nu S dW, \quad dS = \frac{\sigma}{S} \, S \, dW.$$

So, to define the smile, we are matching the option price of a model with relative or returns volatility ν to the option price of a model with

Anticipating the smile type based on the return volatility III

relative or returns volatility σ/S. The first relative volatility is constant, the second is a decreasing function of S, that gets smaller and smaller as K increases when we ask $S > K$ in the call option. Therefore, for larger and larger K, the return volatility in Bachelier being smaller and smaller, we will have a smaller ν in Black Scholes to match it, or in other terms as the strike increases the matching ν vol decreases because the return vol σ/S in the Bachelier model is smaller. The smile is therefore expected to be decreasing.

Now we move to CEV. Claim: CEV return volatility is decreasing for $\gamma < 1$ while Black scholes is constant, so we expect a decreasing smile curve $K \mapsto \nu(K)$ for CEV with $\gamma < 1$.

Anticipating the smile type based on the return volatility IV

Reasoning: comparing Black-Scholes and CEV return volatilities we have the two models

$$dS = rSdt + \nu SdW, \quad dS = rSdt + \sigma S^{\gamma-1} S \, dW.$$

Now recall $\gamma < 1$, so $\gamma - 1$ is negative, and so $\sigma S^{\gamma-1}$ is a decreasing function of S. So we are matching an option price of a model with constant relative or returns volatility ν (Black Scholes) to the option price of a model with decreasing relative or returns volatility $\sigma S^{\gamma-1}$. The first relative volatility is constant, the second is a decreasing function of S, that gets smaller and smaller as K increases when we ask $S > K$ in the call option. Therefore, for larger and larger K, the return volatility in CEV with $\gamma < 1$ being smaller and smaller, we will have a smaller ν in Black Scholes to match it, or in other terms as the

Anticipating the smile type based on the return volatility V

strike increases the matching Black Scholes vol decreases, because the return vol in the CEV model is smaller. The smile is expected to be decreasing.

Now we move to DD. Claim: DD return volatility is decreasing for $\alpha < 0$ and increasing for $\alpha > 0$, while Black Scholes is constant, so we expect a decreasing smile curve for $\alpha < 0$ and an increasing smile curve for $\alpha > 0$.
Reasoning: the BS and DD relative volatilities are, respectively, read from

$$dS = rSdt + \nu SdW, \quad dS = rSdt + \sigma S(1 - \frac{\alpha e^{rt}}{S_t})dW.$$

Anticipating the smile type based on the return volatility VI

So, for $\alpha < 0$, the relative volatility of DD is $\sigma(1 - \frac{\alpha e^{rt}}{S_t})$ and it is a decreasing function of S, contrary to the BS relative vol ν that is constant, so we confirm the previous intuition and reason exactly as in CEV and Bachelier, deducing that the smile should be decreasing. For $\alpha > 0$ we get an increasing function of S, $\sigma(1 - \frac{\alpha e^{rt}}{S_t})$ and we expect an increasing smile.

All the earlier models above allow only for monotonic smile pattern $K \mapsto \nu(K)$ that are either increasing or decreasing. That's why we could make the traders analysis above. The following models will not satisfy this assumption, as they will lead to U shaped smiles, more in line with the market.

The Mixture Diffusion Dynamics (MDD) model I

We now present a model originally due to B. and Mercurio, developed from 1998 to 2021 in several versions (univariate, multivariate, shifted, shifted means, local volatility, random volatility...), who developed it in a number of papers and books listed below.

We know that the Black Scholes formula, that traders like so much, comes from a GBM, $dS_t = rS_t dt + \nu S_t dW_t^Q,\ s_0$ with lognormal density $p_{t,\nu}^{lognormal}$ given in Eq. (31) above. The price of say a call option in Black Scholes is given by integration of the option payoff against this lognormal density.

The starting idea of the mixture model is to consider lognormal densities as in the Black Scholes model but for a number N of possible constant deterministic volatilities $\sigma_1, \ldots, \sigma_N$, where we call $p_{i,t} = p_{t,\sigma_i}^{lognormal}$.

The Mixture Diffusion Dynamics (MDD) model II

We wish to build a model

$$dS_t = rS_t dt + \boxed{\sigma_{\text{mix}}(t, S_t)} S_t \, dW_t, \quad S_0 = s_0 \tag{33}$$

where the return volatility $\sigma_{\text{mix}}(t, S_t)$ is built in such a way that the distribution of S_t is a mixture of distributions of the lognormals $p_{i,t}$, or in formula

$$p_{S_t}(y) =: p_t(y) = \sum_{i=1}^{N} \lambda_i p_{i,t}(y) = \sum_{i=1}^{N} \lambda_i p_{t,\sigma_i}^{lognormal}(y)$$

where $\lambda_i \in (0, 1)$ and $\sum_{i=1}^{N} \lambda_i = 1$. The λ_i are the weights of the different densities $p_{i,t}$ on the mixture.

The Mixture Diffusion Dynamics (MDD) model III

Set $\sigma_{\text{mix}}(t,y)^2 = \frac{1}{\sum_j \lambda_j p_{j,t}(y)} \sum_i \lambda_i \sigma_i^2 p_{i,t}(y)$, or more in detail

$$\sigma_{\text{mix}}(t,y)^2 = \frac{\sum_{i=1}^N \lambda_i \sigma_i^2 \, \frac{1}{\sigma_i\sqrt{t}} \exp\left\{-\frac{1}{2\sigma_i^2 t}\left[\ln\frac{y}{S_0} - rt + \frac{1}{2}\sigma_i^2 t\right]^2\right\}}{\sum_{j=1}^N \lambda_j \frac{1}{\sigma_j\sqrt{t}} \exp\left\{-\frac{1}{2\sigma_j^2 t}\left[\ln\frac{y}{S_0} - rt + \frac{1}{2}\sigma_j^2 t\right]^2\right\}},$$

for $(t,y) > (0,0)$; $\sigma_{\text{mix}}(t,y) = \sigma_0$ for $(t,y) = (0,s_0)$.
Here we assumed constant σ_i but a fully rigorous version, as in B. & Mercurio's papers, is to take $t \mapsto \sigma_i(t)$ time dependent and to have them share a common value σ_0 in a very short initial time interval $t \in [0,\epsilon)$, to take then a constant value σ_i shortly after $t = \epsilon$. With this adjustment, the SDE with σ_{mix} has a unique strong solution whose marginal density is the desired mixture $p_{S_t} = \sum_i \lambda_i p_{i,t}$.

The Mixture Diffusion Dynamics (MDD) model IV

If ϵ is small, we can ignore it for the purpose of pricing options and assume the σ_i are constant everywhere. Then we have that, for $(t, y) > (0, 0)$, we can write $\sigma^2_{\text{mix}}(t, y)$ as follows:

$$\boxed{\sigma^2_{\text{mix}}(t, y) = \sum_{i=1}^{N} \Lambda_i(t, y)\sigma_i^2},$$

where $\Lambda_i(t, y) \in (0, 1)$ and $\sum_{i=1}^{N} \Lambda_i(t, y) = 1$.
This tells us that $\sigma^2_{\text{mix}}(t, y)$ is a “weighted average” of the σ_i^2’s with weights Λ_i’s.

The Mixture Diffusion Dynamics (MDD) model V

The weights are indeed

$$\Lambda_i(t, y) = \frac{\lambda_i \; p_{i,t}(y)}{\sum_j \lambda_j \; p_{j,t}(y)}.$$

Notice: $p_{S_t}(\cdot)$ has the correct no-arbitrage Q-expectation:

$$E_0^Q[S_t] = \int y p_{S_t}(y) dy = \sum_{i=1}^{N} \lambda_i \int y p_{t,i}(y) dy = \sum_{i=1}^{N} \lambda_i S_0 e^{rt} = S_0 e^{rt}$$

as in any arbitrage free model under Q. This was already clear from the fact that the SDE for S with σ_{mix} had drift $rS_t \; dt$.

The Mixture Diffusion Dynamics (MDD) model VI

Why is the mixture a good idea? We see the answer when we try and calculate an option price with this model. Take for example a call option on S_T.

$$V_{\text{mix}}^{Call}(0,K,T) = e^{-rT}E^Q\left\{(S_T-K)^+\right\}$$
$$= e^{-rT}\int_0^{+\infty}(y-K)^+p_{S_T}(y)dy = e^{-rT}\int_0^{+\infty}(y-K)^+\sum_{i=1}^N \lambda_i p_{i,T}(y)dy$$
$$= \sum_{i=1}^N \lambda_i e^{-rT}\int (y-K)^+ p_{i,T}(y)dy = \sum_{i=1}^N \lambda_i V_{BS}^{Call}(0,S_0,K,T,\sigma_i,r).$$

We see that the price of the call is a linear (actually convex) combination of Black Scholes prices of calls with volatilities $\sigma_1,\ldots,\sigma_N$ *with weights* $\lambda_1,\ldots,\lambda_N$*.*

The Mixture Diffusion Dynamics (MDD) model VII

So the option price becomes a mix of prices with the given weights and volatilities. **The same holds for put options and all other simple contingent claims with payoff of the type $f(S_T)$.**

Remark [Greeks]. Due to the linearity of the derivative operator, the same convex combination applies also to all option Greeks (sensitivities) like Delta, Gamma, Theta, Rho.

This is an extremely flexible model, as we can fine tune the mumber of components N according to the complexity of the smile. Playing with the parameters σ_i and λ_i we can reproduce most market smiles. The model has been used successfuly in the equity, FX and interest-rate markets.

The Mixture Diffusion Dynamics (MDD) model VIII

Remark [Why a mixture?]. When starting from general dynamics dS_t, it is hard to come up with analytical formulas for European options. The use of analytically-tractable densities $p_{i,t}$, instead, immediately leads to closed-form prices. Moreover, the virtually unlimited number of model parameters can be helpful in the market calibration. Furthermore, traders are used to quote and manage derivatives with the lognormal distribution as a benchmark. Departures from the lognormal distribution are to be kept at a minimum, and also motivated. A mixture of lognormals makes the price of any derivative a linear combination of prices, each under a different lognormal. We obtain a linear combination of Black Scholes prices. This leads to a contained conceptual departure from the lognormal distribution and from the lognormal world.

The Mixture Diffusion Dynamics (MDD) model IX

In the mixture dynamics model, one can show rigorously that the resulting volatility smile curve will have a minimum in the at-the-money-forward price $S_0 e^{rT}$.

See B. & Mercurio's papers for a proof, which is not required in this course.

We will show now an example of smile from the model (in red) calibrated to the market (in blue). In this example both the market and the model smile have the minimum near (but not exactly at) the at-the-money-forward FX rate of 0.88. Note that in the FX market $r = r_{domestic} - r_{foreign}$.

The Mixture Diffusion Dynamics (MDD) model X

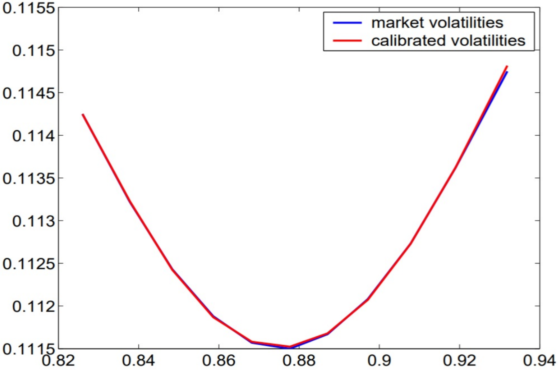


Figure: USD/Euro two-month implied volatilities as of May 21, 2001. The minimum is at $S_0 e^{rT} \approx 0.88$. To see how smiles are actually fitted, see appendix "fitting the smile in practice"

The shifted mixture dynamics model I

Not all markets have the smile with the minimum near $S_0 e^{rT}$. This tends to happen (not always) in the FX market but can be different in other markets like Equity. Can we extend our model to such cases, by adapting the mixture dynamics to volatility smiles that have the minimum away from $S_0 e^{rT}$? To extend our model to arbitrary mimimum points of the smile, we introduce a shift. It's essentially the same idea as in the displaced diffusion model, except that here it is done on the mixture dynamics rather than on Black Scholes.
Let us write a mixture diffusion dynamics model X_t as

$$dX_t = rX_t dt + \sigma_{\text{mix}}(t, X_t) X_t \, dW_t, \quad X_0 = x_0.$$

Let us assume that the asset-price process S_t follows

$$S_t = s_0 \alpha e^{rt} + X_t, \tag{34}$$

where α is a real constant.

The shifted mixture dynamics model II

Differentiating both sides,

$$dS_t = rs_0\alpha e^{rt}dt + dX_t = rs_0\alpha e^{rt}dt + rX_t dt + \sigma_{\text{mix}}(t, X_t)X_t \, dW_t =$$

$$= rs_0\alpha e^{rt}dt + r(S_t - s_0\alpha e^{rt})dt + \sigma_{\text{mix}}(t, S_t - s_0\alpha e^{rt})(S_t - s_0\alpha e^{rt}) \, dW_t$$

(where we used Eq. (34)) and simplifying

$$\boxed{dS_t = rS_t dt + \sigma_{\text{mix}}(t, S_t - s_0\alpha e^{rt})(S_t - s_0\alpha e^{rt}) \, dW_t, \ \ s_0}$$

The shifted mixture dynamics model III

The price of a call option in the shifted mixture dynamics is

$$V^{\text{Call}}_{\text{shift-mix}} = e^{-rT}\sum_{i=1}^{N}\lambda_i\left[\mathcal{S}_0 e^{rT}\Phi\left(\frac{\ln\frac{\mathcal{S}_0}{\mathcal{K}} + \left(r+\frac{1}{2}\sigma_i^2\right)T}{\sigma_i\sqrt{T}}\right)\right.$$
$$\left. - \mathcal{K}\Phi\left(\frac{\ln\frac{\mathcal{S}_0}{\mathcal{K}} + \left(r-\frac{1}{2}\sigma_i^2\right)T}{\sigma_i\sqrt{T}}\right)\right],$$

where $\mathcal{K} = K - s_0\alpha e^{r\,T}$, $\mathcal{S}_0 = s_0(1-\alpha)$
To derive this formula, note that

$$V^{\text{Call}}_{\text{shift-mix}} = e^{-rT}E^Q[(S_T - K)^+] = e^{-rT}E^Q[(s_0\alpha e^{rT} + X_T - K)^+] =$$

$$= e^{-rT}E^Q[(X_T - (K - s_0\alpha e^{rT}))^+]$$

The shifted mixture dynamics model IV

where we used again (34), and now remembering that X follows a mixture dynamics model SDE with initial condition

$$x_0 = s_0 - s_0\alpha = s_0(1-\alpha)$$

(again by (34) at time 0) we have

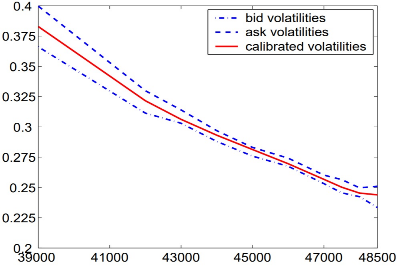
$$V^{\text{Call}}_{\text{shift-mix}} = \sum_{i=1}^{N} \lambda_i V^{Call}_{BS}(0, s_0(1-\alpha), K - s_0\alpha e^{rT}, T, \sigma_i, r)$$

which is the formula given above.

Introducing a non-zero alpha moves the minimum of the volatility smile away from the at-the-money forward $S_0 e^{rT}$ and allows for more general smiles.

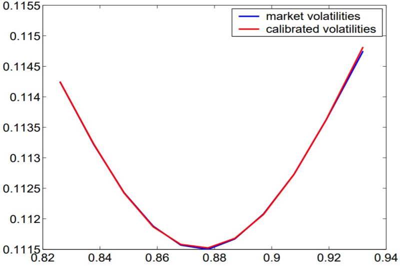
Now we present two examples of calibration to market data.

The shifted mixture dynamics model V



Data: Italian MIB30 equity index on March 29, 2000, at 3,21pm (most liquid puts with the shortest maturity). We set $N=3$, $\lambda_3 = 1 - \lambda_1 - \lambda_2$. We minimize the squared percentage difference between model and market mid prices. We get: $\lambda_1 = 0.201$, $\lambda_2 = 0.757$, $\sigma_1 = 0.019$, $\sigma_2 = 0.095$, $\sigma_3 = 0.229$, $\alpha = -1.852$.

The shifted mixture dynamics model VI



Data: USD/Euro two-month implied volatilities as of May 21, 2001. We set $N=2$, $\lambda_2 = 1 - \lambda_1$. We minimize the squared percentage difference between model and market mid prices. We get: $\lambda_1 = 0.451$, $\sigma_1 = 0.129$, $\sigma_2 = 0.114$, $\alpha = 0.076$.

The shifted mixture dynamics model VII

This has been a quick introduction. Missing:

- Calibrating a whole vol surface with different T's;
- Putting new drifts in the basic processes:

$$dS_t^i = \boxed{\mu_i(t)} S_t^i \, dt + \sigma_i(t) S_t^i dW_t$$

  Increases fitting capability of asymmetric structures
- Analysis of the transition densities and implication on *future* volatility structures.
- Interest rate models... (long story)

The shifted mixture dynamics model VIII

Published Literature:

- Brigo, D., Mercurio, F. (2000) **A Mixed-up Smile**. *Risk*, September, 123-126.
- Brigo, D., Mercurio, F. (2001) **Displaced and Mixture Diffusions for Analytically-Tractable Smile Models**. In *Mathematical Finance - Bachelier Congress 2000*, Geman, H., Madan, D.B., Pliska, S.R., Vorst, A.C.F., eds. *Springer Finance*, Springer, Berlin Heidelberg New York, to appear.
- Brigo, D., Mercurio, F. (2001) **Fitting Volatility Smiles with Analytically Tractable Asset Price Models**, and **Lognormal-Mixture Dynamics and Calibration to Market Volatility Smiles**, in *International Journal of Theoretical and Applied Finance*

The shifted mixture dynamics model IX

- Brigo, D., Mercurio, F. (2001) **Interest Rate Models: Theory and Practice**. Springer Finance. Springer.
- Brigo, D, Mercurio, F & Sartorelli, G (2003), Alternative Asset Price Dynamics and Volatility Smile, Quantitative Finance, vol. 3, no. 3, pp. 173 - 183.
- Brigo, D., Rapisarda, F. and Sridi, A. (2018). The multivariate mixture dynamics: Consistent no-arbitrage single-asset and index volatility smiles, IISE Transactions, 50:1, 27–44
- Brigo, D., Rapisarda, F. and Pisani, C. (2021). The multivariate mixture dynamics model: shifted dynamics and correlation skew. Annals of Operations Research, vol. 299, issue 1, No 56, pp 1435

Local volatility models & future smile flattening I

The volatility models we introduced above are all called *local volatility models*.
In these models the return volatility $\sigma(t, S_t)$ in the SDE

$$dS_t = rS_t dt + \sigma(t, S_t) S_t dW_S^Q(t), \quad s_0$$

is a function of t and S only, and there is no randomness entering the stock price instant by instant except for $dW_S^Q(t)$. We write $dW_S^Q(t)$ instead of dW_t^Q to emphasize this is the Brownian of dS.
We have to be careful what we mean by local volatility. Local volatility means that the return volatility $\sigma(t, S_t)$ seen at time t, *conditional on* $\mathcal{F}_t^{W_S^Q}$, is not random but deterministic. However, this does not mean that the return volatility is deterministic seen from time 0. Conditional on the information at time 0 only, $\sigma(t, S_t)$ is still random because its argument S_t is random seen from time 0.

Local volatility models & future smile flattening II

As we will see, in stochastic volatility models, $\sigma(t, S_t)$ seen at time t, *conditional on* $\mathcal{F}_t^{W_S^Q}$, will be random and not deterministic as in local vol models. In the Heston stochastic vol model we will see later, we have $\sigma(t, S_t) = \sqrt{V_t}$ where V_t is a new stochastic process that is not $\mathcal{F}_t^{W_S^Q}$ measurable, so it is not deterministic given the history of W_S (and thus S) up to time t.

Local volatility models suffer from the fact that their future volatility smile tends to flatten. If you were to calculate the future smile at time T_1 for a maturity $T_1 + T$, conditional on a value of S_{T_1}, and you compared it to the smile at time 0 for the maturity T, you would get a smile that is flatter at T_1.

Let's explain this flat smile problem more in detail.

Local volatility models & future smile flattening III

Condition to a given value for the random stock price (as seen from time 0) S_{T_1}, call it s_1, then run the SDE model for the stock you are using under Q, up to time $T_1 + T$, starting from $S_{T_1} = s_1$ at time T_1.

Get a number of scenarios for S_{T_1+T} from the SDE model as above. These scenarios are conditional on $S_{T_1} = s_1$ because we started the SDE at time T_1 from s_1.

Plug each scenario of S_{T1+T} from the SDE model into the payoff $(S_{T_1+T} - K)^+$, and average to get the Q expectation.

This average is $E^Q[(S_{T_1+T} - K)^+ | S_{T_1} = \bar{S}_1]$, now discount with e^{-rT} and you get the option price for the SDE model at time T_1 conditional on $S_{T_1} = s_1$.

Local volatility models & future smile flattening IV

Equate this price with a Black-Scholes formula at time T_1 for maturity $T_1 + T$ (or time-to-maturity T) & initial stock price s_1 and solve this equation with the Black-Scholes vol as the unknown. Call this vol $\nu(T_1, K, T; s_1)$ (the third parameter is the time to maturity).
If you use a stoch vol model like Heston below, the two smiles

$$K \mapsto \nu(0, K, T; s_0), \quad K \mapsto \nu(T_1, K, T; s_1)$$

will be similar. If you use a local vol model (e.g. CEV or mixture dynamics) the second smile will be flatter.

The reason for this can be investigated theoretically or empirically. For further discussion see the APPENDIX “More on future smile flattening”.

Without proof, we can say that local vol models have this problem because the vol $\sigma(t, S_t)$ has all randomness coming from S_t (or from

Local volatility models & future smile flattening V

W_S), and does not have a random life of its own. A stoch vol model, introducing new randomness in the volatility, can preserve the smile and avoid the flattening. Again see the APPENDIX for more discussion.

Traders don't like the flattening problem and this is the reason why in some markets (especially FX and Equity) local volatility models are rarely used to price and hedge in the front office. They may be used more as risk management or model validation tools, or be used in other markets where the stochasticity of the volatility is not as relevant (e.g. interest rates).

Brief hint at Stochastic Volatility Models (Heston) I

The flattening issue is avoided in stochastic volatility models (SVM). In SVMs the return volatility is a second stochastic process with new randomness. For example, the Heston SVM reads under the measure Q

$$dS_t = rS_t dt + \sqrt{V_t} S_t dW_S(t), \ \ s_0$$

$$dV_t = k(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_V(t), \ \ v_0,$$

$$dW dW_V = \rho \, dt.$$

Note that here the return volatility $\sqrt{V_t}$ of S is based on a new SDE dV with a new Brownian motion W_V, possibly correlated with W_S, the stock price Brownian motion, via ρ, but not identical to it (unless $\rho = 1$). This means that new extra randomness enters dS_t on top of dW_S at every instant.

Brief hint at Stochastic Volatility Models (Heston) II

The process V

$$dV_t = k(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_V(t), \quad v_0,$$

is like the CIR model for interest rates and is mean reverting to θ with speed k and a local variance parametrized by σ_V. In a sense σ_V is the volatility of the (squared) volatility V (traders talk about “vol of vol”).

This model avoids the “flat future smile” problem.
There are other stochastic volatility models like Hull-White, SABR ...
but Heston is one of the best models given the properties of V_t.

What does it all mean

So far we have tried to follow a technical path, but it is time to appreciate the significance of what we have done so far in a broad context, and to revisit some of the assumptions we made.

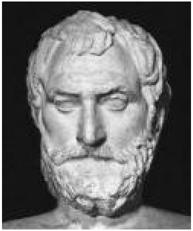
We now ask ourselves: What are the implications of what we have calculated on the big picture?

Quantitative Finance deals in large part with financial derivatives. Options are examples of such derivatives. So, following our derivation above, why are derivatives so important, so popular and, often, unpopular? How did they start?

Options and Derivatives

Derivatives outstanding notional as of June 2011 (BIS) is estimated at **708 trillions USD** (US GDP 2011: 15 Trillions; World GDP: 79 Trillions)

$$708000 \text{ billions}, \quad 708,000,000,000,000, \quad 7.08 \times 10^{14} \ \text{ USD}$$



How did it start? It has always been there. Around 580 B.C., Thales purchased options on the future use of olive presses and made a fortune when the olives crop was as abundant as he had predicted, and presses were in high demand. (Thales is also considered to be the father of the sciences and of western philosophy, as you know).

Options and Derivatives valuation: precursors

- **Louis Bachelier** (1870 – 1946) (First to introduce Brownian motion W_t, first in the modern study of Options);
- **Bruno de Finetti (1906 – 1985) (and Frank Ramsey (1903-1930))** (Fathers of the subjective interpret of probability). BdF shows betting quotients (claim prices?) avoid sure exploitation from gambling broker (market?) if and only if they satisfy axioms of a probability measure.

Modern theory follows Nobel awarded **Black, Scholes and Merton** (and then Harrison and Kreps etc) on the correct pricing of options.

What does it all mean? Call option and Gambling

We saw earlier the call option on a stock (say ACME). This can be a gamble against a bank, where:

- If the future price of the ACME stock in 1y is larger than the value of ACME today, we receive from the bank the difference between the two prices (on a given notional).
- If the future price of the ACME stock in 1y is smaller or equal than the value of ACME today, nothing happens.

The bank will charge us for entering this wage, since we can only win or get into a draw, whereas the bank can only lose or get to a draw.

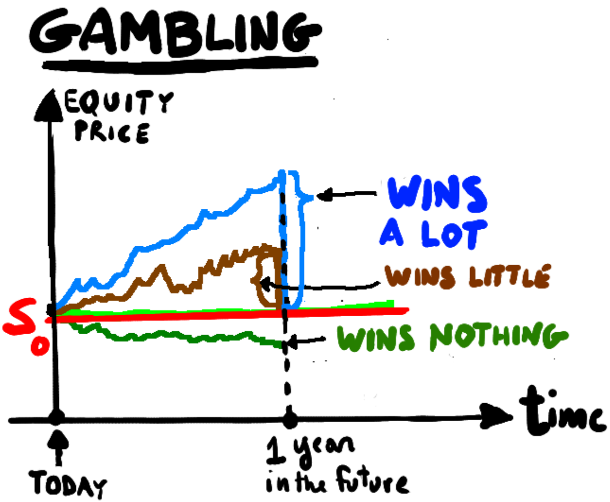


Figure: A one-year maturity Gamble on an equity stock. Call Option.

Call option and Gambling

We have an investor buying a call option on ACME with a 1y maturity.

The Bank's problem is finding the correct price of this option today. This price will be charged to the investor, who may also go to other banks.

This is the option pricing problem, main job of quants, together with hedging, in the past.

Derivatives can be bought to protect or hedge some risk, but also for speculation or "gambling".

Black and Scholes: What does it mean?

We have derived the Black Scholes formula for a call option earlier, and later we also saw the same calculation under different (local volatility smile) models. Let us recall the key points.

Let S_t be the equity price for ACME at time t.
For the value of the ACME stock S_t let us assume, as before, a SDE
$dS_t = \mu S_t dt + \sigma(t, S_t) S_t dW_t$ or also

$$\underbrace{\frac{dS_t}{S_t}}_{\substack{\text{relative change}\\ \text{in stock ACME}\\ \text{between}\\ t \text{ and } t+dt}} = \underbrace{\mu}_{\substack{\text{instantaneous}\\ \text{"mean" return}\\ \text{of ACME}\\ \text{between } t \text{ and } t+dt}} dt + \underbrace{\sigma(t, S_t)}_{\substack{\text{volatility}\\ \text{for ACME}}} \underbrace{dW_t}_{\substack{\text{New}\\ \text{random}\\ \text{shock}}}$$

Black and Scholes: What does it mean?

Then we have seen there exists a formula (Black and Scholes' or one of the smile models formula) providing a unique fair price for the above gamble (option) on ACME in one year.

This Black Scholes or smile model formula **depends on the volatility** σ of ACME, and from the initial value S_0 of ACME today, but **does NOT depend on the expected return** μ of ACME.

This means that two investors with very different expectations on the future performance of ACME (for example one investor believes ACME will grow, the other one that ACME will go down) will be charged the same price from the bank to enter into the option.

The Gamble price does not depend on the investor perception of future markets. One would think that Red Investor should be willing to pay a higher price for the option with respect to Blue Investor. Instead, both will have to pay the gamble according to the green scenarios, where ACME grows with the same returns as a riskless asset

Derivatives prices independent of expected returns???

This seemingly counterintuitive result renders derivatives pricing independent of the expected returns of their underlying assets.

This makes derivatives valuations quite objective, and has contributed to derivatives growth worldwide.

Today, derivatives are used for several purposes by banks and corporates all over the world

A mathematical result has contributed to create new markets that reached 708 trillions (US GDP: 15 Trillions)

But keep in mind that the derivation of the Black Scholes result and of smile models holds only under the 4 ideal conditions and actually many more assumptions:

The Black Scholes Merton analysis assumptions

- **Short selling is allowed without restrictions**
- Infinitely divisible shares
- **No transaction costs**
- No dividends in the stock
- **No default risk of the parties in the deal**
- **No funding costs: Cash can be borrowed or lent at the risk free rate r. Remove this and Valuation becomes Nonlinear** (Semi-Linear PDEs, FBSDEs, see several papers B. & Pallavicini 2011-2015)
- **Continuous time and continuous trading/hedging**
- **Perfect market information, Complete markets**
- **....**

Many of the above assumptions are no longer tenable, especially after 2007-2008, but were already unrealistic well before 2008.

Crisis

After Black Scholes 1973...

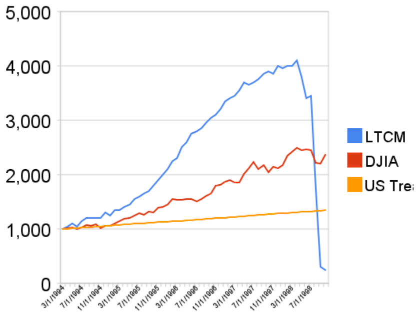
Market players introduced derivatives that may be much more complex functionals of underlying assets and events than the above call option

Gamble/speculate/hedge/protect on anything?

The initial Black Scholes theory of 1973 (Nobel award 1997) has often been extrapolated beyond its limits to address much more complex derivatives. Such derivatives often work on different sectors: Foreign Exchange Rates, Interest Rates, Default Events, Meteorology, Energy, population Longevity...

Aggressive market participants extrapolating the basic theory

One of the most controversial extrapolations is Credit Derivatives and CDOs in particular, linked to the 2008 crisis.



Sometimes the timing of the Nobel committee is funny, and we are not talking about the peace Nobel prize. Warning: anedoctal

1997: Nobel award to Scholes and Merton (Black had passed away).

1998: the US Long-Term Capital Management hedge fund has to be bailed out after a huge loss. The fund had Merton and Scholes in their board and made high use of leverage (derivatives). This leads us to...

Crises and dangers with derivatives

- Metallgesellschaft lost $1.3 billion by entering into long term oil contracts in 1993.
- The Barings Bank collapse of 1995, which was solely down to the fraudulent dealings of one of its traders.
- Long-Term Capital Management's near collapse in 1998 and subsequent bailout overseen by the Federal Reserve. *Somewhat ironically, members of LTCM's board of directors included Scholes and Merton.*
- In 2003, Parmalat collapsed with a EUR 14 billion accounting hole, in what remains Europe's biggest bankruptcy. Parmalat was selling itself credit derivatives (credit-linked notes), placing a bet on its own creditworthiness to conjure up an asset out of thin air.
- The 2007 subprime crisis, triggering the
- 2008 Financial crisis worldwide (8 defaults of financials in 1 month)

Derivatives: the Barings collapse I

The collapse of Barings Bank in Feb 1995 was caused by huge losses of a rogue trader, Nick Leeson. Leeson was head of derivatives in Singapore. He gambled more than $1 billion in non–hedged, unauthorized speculation trading, destroying the venerable bank's reserves.

Derivatives: the Barings collapse II

After fleeing to Malaysia, Thailand and finally Germany, Leeson was arrested in Frankfurt and extradited back to Singapore on 20 November 1995.

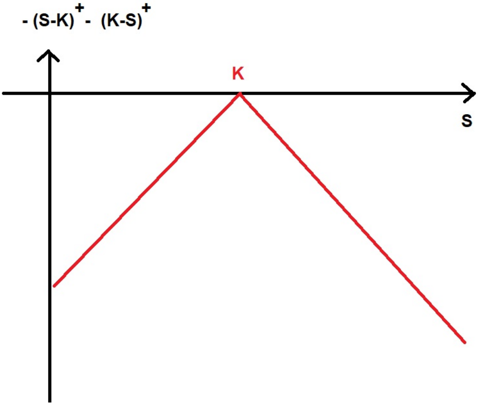
Derivatives: the Barings collapse III

Among Leeson's positions, a famous one is a short Straddle position. This is when you sell a call and a put with the same strike, cashing in the premium for both options.

A short straddle on the stock S with maturity T and strike K has final payoff

$$Y = -(S_T - K)^+ - (K - S_T)^+ = -|S_T - K|.$$

Derivatives: the Barings collapse IV



Basically you will have to pay an amount to the option buyer whatever happens, as any move of the stock from K will trigger the positive payoff $|S_T - K|$ you have to pay.

Derivatives: the Barings collapse V

You cashed in the price of the call and the put at time 0, so your hope is that the stock will move very little, so that the option premium you cashed at time 0 will be much larger than the movement of the stock from K at maturity T.

After a lot of rogue trading in futures, Leeson struggled to further cover his losses and tried to make money by selling a straddle, gambling on the fact that the stock would move little in the future. This way he could cash-in the straddle initial price while hoping to make very small payments in the future.

However, after he sold the straddle, the Kobe earthquake hit Asian markets, creating large price movements in the straddle underlying asset and generating big losses in Leeson's position. To try to get funds and keep his rogue trading going, he entered further futures

Derivatives: the Barings collapse VI

positions but in vain. Losses kept increasing and eventually the bank collapsed.

The Barings collapse was one of the main incidents that convinced regulators that banks needed a risk measure for all their portfolios, and that capital reserves proportional to this measure should be in place when trading the relevant portfolios.

Risk measures as a first response to crises

The introduction of Risk Measures in the late 1990's was a response to the Barings collapse and other incidents. The value at risk (VaR) and later the espected shortfall (ES) were the main risk measures used to respond to the crises. We will look at these measures in part 3.

The 2008 crisis. Credit risk

Later on, following the 7[8] credit events happening to Financials in one month of 2008,

Fannie Mae, Freddie Mac, Lehman Brothers, Washington Mutual, Landsbanki, Glitnir and Kaupthing [and Merrill Lynch]

credit risk in trading could not be ignored anymore.

This led to the introduction of the credit valuation adjustment (CVA), the first of a series of valuation adjustments that had to account for effects often neglected before the 2008 crisis.

The crisis involved also important issues on liquidity, collateral and interest rates, but we won't discuss those here.

We will discuss now risk measures, in Part 3. We will not cover credit risk, collateral, funding costs and valuation adjustments. That is covered in the MSc Mathematics and Finance.

PART 3: RISK MEASURES

In this part we introduce the two key risk measures used in the industry, Value at Risk (VaR) and Expected Shortfall (ES), providing some numerical examples of how these measures are calculated on option portfolios

RISK MEASURES

In this part we look at the problem of risk measurement and management.

So far we discussed mostly valuation and hedging. This is important and is done under the risk neutral measure $\mathbb{Q}$, as we have seen earlier.

Risk Management however is partly based on historical estimation, and is interested in potential losses in the physical world, hence we need to go back to the historical/physical measure $\mathbb{P}$.

We first discuss briefly statistics under the measure $\mathbb{P}$ and then introduce the two fundamental risk measures of Value at Risk (VaR) and Expected Shortfall (ES).

Time series under the measure P I

We have seen that in the Black Scholes model, under the measure P, the stock S_t is

$$S_t = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}, \quad 0 \leq t \leq T.$$

Given a sequence of times $t_0, t_1, t_2, \ldots, t_i, t_{i+1}, \ldots t_N$, where we assume $t_{i+1} - t_i = \delta$ for all i, we can write, from the above equation

$$\log \frac{S_{t_{i+1}}}{S_{t_i}} = \left(\mu - \frac{1}{2}\sigma^2\right)\delta + \sigma(W_{t_{i+1}} - W_{t_i}) \sim \mathcal{N}\left(\left(\mu - \frac{1}{2}\sigma^2\right)\delta, \sigma^2\delta\right)$$

where we used the usual $W_{t_{i+1}} - W_{t_i} \sim \mathcal{N}(0, t_{i+1} - t_i) = \mathcal{N}(0, \delta)$.

The above formula tells us that the log-returns $\log(S_{t_{i+1}}/S_{t_i})$ are Gaussian in Black Scholes. Is this true for real financial data?

Time series under the measure P II

This can be assessed with sample estimators and with a QQ plot.

We introduced at the beginning of the course the skewness and excess kurtosis of a random variable X as

$$\frac{\mathbb{E}[(X-\mu)^3]}{\sigma^3}, \quad \frac{\mathbb{E}[(X-\mu)^4]}{\sigma^4} - 3,$$

both quantities being zero if $X \sim \mathcal{N}(\mu, \sigma^2)$ is Gaussian. So, the first thing we can do if presented with some stock data history $s_{t_0}, s_{t_1}, s_{t_2}, \ldots, s_{t_N}$ is to take log returns

$$x_{t_1} = \log \frac{s_{t_1}}{s_{t_0}}, x_{t_2} = \log \frac{s_{t_2}}{s_{t_1}}, \ldots, x_{t_N} = \log \frac{s_{t_N}}{s_{t_{N-1}}}$$

and to check the sample skewness and kurtosis of the x data.

Time series under the measure P III

There are both unbiased and biased estimators of skewness and excess kurtosis in the literature. Once we choose an estimator, we apply it to the data. If we have that Skewness is significantly different from zero, the probability density function of the data (histogram) will be highly asymmetric, so that our log returns of the stock cannot be normal.

If we have that excess kurtosis is significantly smaller than zero, then the distribution is less dispersed than a normal, and the tails will be thinner than a normal.

If we have that excess kurtosis is significantly larger than zero, then the distribution is more dispersed than a normal, and the tails will be fatter than a normal.

Time series under the measure P IV

Financial time series, for example for stock prices, often exhibit fat tails, i.e. tails that are fatter than a normal. This means that extreme events, which correspond usually to tails of the loss distribution, could have their probability underestimated if modeled as normals.

We can check this with financial data in a couple of examples.
5 years of S&P500 returns 13/3/2017–13/3/2022

If we look at the related Excel spreadsheet (available), Excel has a function "Skew" to compute the skewness, and "Kurt" to compute the excess kurtosis. For log-returns of S&P500 over 5 years we get

$$\mathit{mean} = 0.00044; \;\; \mathit{STD} = 0.01212; \;\; \mathit{skew} = -1.1279; \;\; \mathit{exc\,kurt} = 21.46.$$

We can transform the daily volatility in an annualized one by multiplying by the square root of the number of (252 working) days in a

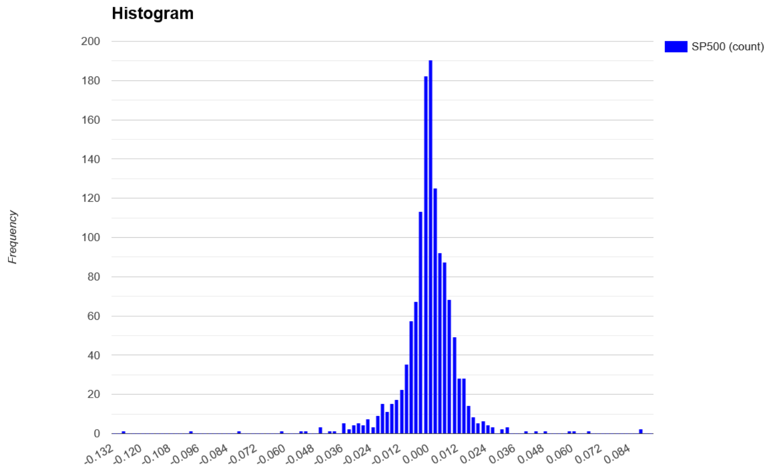
Time series under the measure P V

year. Remember, log-returns are independent, so the variance of their sum over one year is the sum of their variances. This means that to get the annual variance from the daily one we need to multiply by 252, whereas for the standard deviation we have the square root. So

$$\textit{Historical Volat} = \textit{Annualized STD} = \sqrt{252}\,0.01212 = 0.1925 = 19.25\%.$$

This is σ in the Black Scholes model estimated under the measure P. We notice that both skewness and kurtosis are quite different from 0. The kurtosis, in particular, is quite large. We can expect then tails that are much thicker than the normal distribution and we can expect asymmetry. The histogram of the data confirms this:

Time series under the measure P VI



Time series under the measure P I

We re-do the exercise to illustrate how the properties may depend on the particular dataset and especially on the time window of the historical data. We now download from Yahoo finance the S&P500 close prices from March 9, 2021 to March 9, 2022 (1 year of data). We get, for the log returns,

$$\textit{mean} = -0.0004;\ \ \textit{STD} = 0.00912;\ \ \textit{skew} = 0.29;\ \ \textit{exc kurtosis} = 0.65.$$

The daily mean return is quite close to zero. The daily standard deviation is almost 1%.

$$\textit{Historical Volatility} = \textit{Annualized STD} = \sqrt{252}\, 0.00912 = 0.145 = 14.5\%$$

This is σ in the Black Scholes model estimated under the measure P.

Time series under the measure P I

When we looked at smile modeling, we saw the implied volatility, which is the volatility $\sigma_{implied}$ that reproduces the market option price (a Q-expectation) when put in the Black Scholes formula.
On the date 9 March 2022, the implied volatility from S&P500 options was $\sigma_{imp} = 31.8\%$.

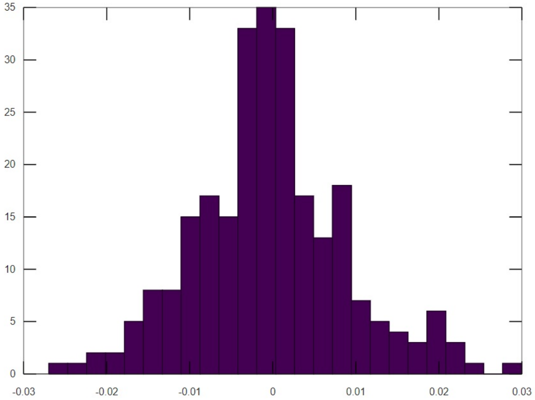
In general, implied volatilities tend to be higher than historical volatilties.

However, if we measure the historical volatility only on the last ten days of the sample, the historical volatility becomes $\sigma_{hist} = 30\%$. If we measure it only on the last 30 days, $\sigma_{hist} = 24\%$ (see spreadsheet).
The historical volatility depends a lot on the history time window on which we calculate it (e.g. 10 days vs 1 month vs 1 year). In all cases, however, in our example, $\sigma_{hist} < \sigma_{implied}$, and in most cases this holds.

Time series under the measure P I

Going now to skewness and excess kurtosis, both numbers are positive. Skew is positive, contrary to the 5y case, whereas kurtosis is positive but much smaller than in the 5y case, so our data are closer to normal now. An histogram may confirm this.

Time series under the measure P II



Histogram is bulky as we only used 1 year of data. We can see that the distribution is more spread than a Gaussian, that it is skewed to the right (right tail longer) and tails are fatter than Gaussian.

Time series under the measure P III

However, if we compare with the 5y histogram, tails are less spread.

A powerful tool that can visualize how far we are from a (standard) normal distribution is the **QQ plot**. This plots the quantiles of a distribution against the quantiles of another distribution, typically a standard normal.

If the two distributions are equal, their quantiles are equal and we get the straight line $y = x$.

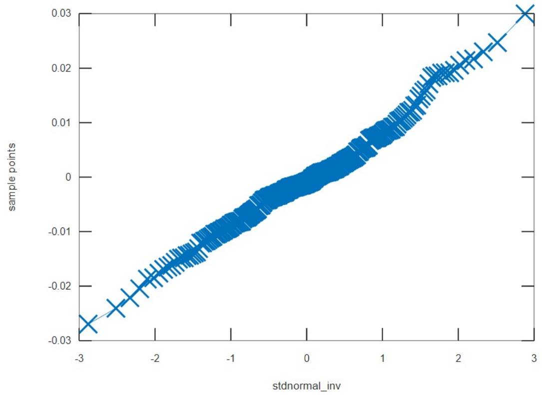
If the two distributions are related by a linear transformation, like a non-standard normal vs a normal, $\mathcal{N}(\mu, \sigma^2) = \mu + \sigma\mathcal{N}(0, 1)$ vs $\mathcal{N}(0, 1)$, then the QQ plot is a straight line but not $y = x$.

Time series under the measure P IV

If the two distributions are not equal nor related by a linear transformation, the QQ plot will depart from a straight line the more the two distributions percentiles, and tails in particular, are different. In general, departure from a linear pattern points to fat tails.

Let's look at the 1y dataset first. From skewness and kurtosis we know that we are not in presence of normal returns. We can look at a QQ plot of the log-return data x against a standard normal.

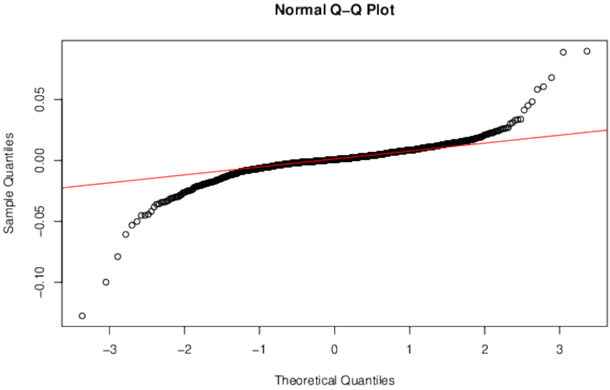
Time series under the measure P V



We see that while the curve is not exactly sitting on a straight line, the departure is not dramatic. Then we can still use the Black Scholes with its normal returns as an approximation for risk management purposes with this dataset.

Time series under the measure P VI

The situation is quite different with the 5y dataset. The QQ plot in that case is



Time series under the measure P VII

Here using Black Scholes, implying normal returns, would be more problematic as the departure from a straight line is more dramatic.

The fat tails in the 5y dataset are confirmed both by the high Kurtosis and by the QQplot, which give consistent findings.

Time series under the measure P VIII

We can summarize both examples in a table

Data set	mean	hist vol last 2 weeks	hist vol whole sample	impl vol	skew	kurt	QQ Plot
1y	- 0.0004	30%	14.5%	31.8%	0.29	0.65	line
5y	0.00044	27.12%	19.25%	31.8%	-1.13	21.46	curved

Table: S&P500 log-returns statistics, 1y dataset from Yahoo finance, period 9 March 2021 - 9 March 2022. 5y dataset from Federal Reserve, period 13 March 2017 - 13 March 2022.

Time series under the measure P IX

For the 5y dataset the model should have fat tails.

Some of the models we have seen, like the mixture dynamics, have fatter tails than the Gaussian and are more consistent both with P (historical skewness and kurtosis) and Q data (market volatility smile). We have seen these models only under Q, but it is easy to formulate them under P by changing their drift to $\mu S_t\, dt$.

We can use a mixture dynamics to see if it can achieve similar patterns to the 5y data. We can choose a mixture dynamics as follows under the measure P

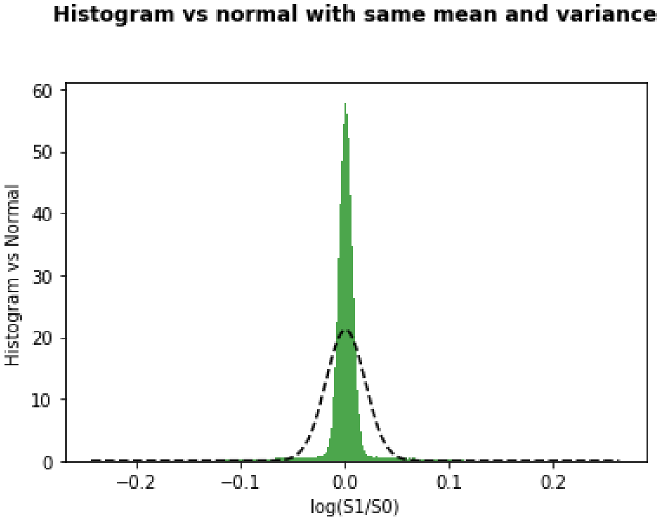
Time series under the measure P X

λ_1	λ_2	σ_1	σ_2	$E[\ln \frac{S_{1d}}{S_0}]$	$\sqrt{252}STD[\ln \frac{S_{1d}}{S_0}]$	skew	kurt
0.9	0.1	0.1	0.9	0.0004	0.30	-0.23	21.08

Table: One example of mixture lognormal dynamics achieving a large Kurtosis. We mix two lognormals. $S_0 = 100, \mu = 0.1511$; Statistics of simulated log-returns over one day

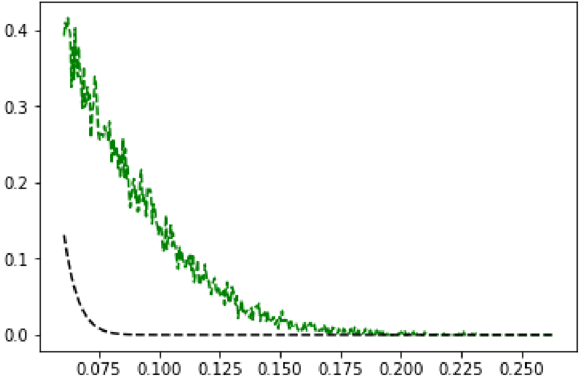
We can show the histogram plot of the mixture return density (green) against a density of a normal with the same mean and variance (dashed black)

Time series under the measure P XI

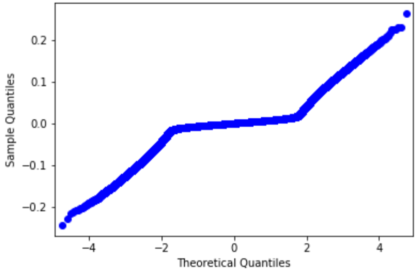


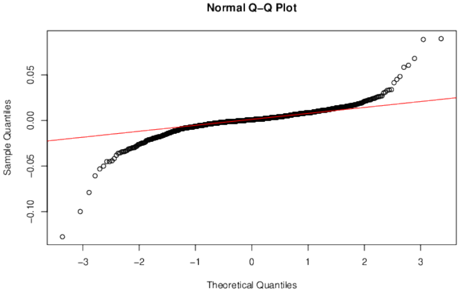
Time series under the measure P XII

The plot does not allow to appreciate the difference in the tails due to scale. We zoom on the fat tails on the right:



Time series under the measure P XIII





On the left hand side we have the mixture dynamics daily return QQplot. On the right hand side we have the 5 year dataset QQplot. They follow a roughly similar pattern.

Time series under the measure P XIV

This shows the mixture dynamics has the potential to model fat tails distributions.

This is just an example. In reality one can do a precise maximum likelihood estimation of the model to historical data. Also, as we have seen previously, one can alternatively fit the model to option smile data under the measure Q. In that case the drift has to be replaced by $rS_t dt$.

This is a key example of how model selection may be guided by empirical analysis of data.

We now introduce the two main risk measures the industry has adopted to respond to incidents like the Barings collapse: Value at Risk and Expected Shortfall.

Value at Risk I

Value at risk (VaR) is a single, summary, statistical measure of possible portfolio losses. It aggregates all of the risks in a portfolio into a single number suitable for use in the boardroom, reporting to regulators, or disclosure in an annual report, and it is the most widely used risk measure in financial institutions according to McNeil et al.

In addition to this, VaR estimates not only serve as a summary statistic, but are also often used as a tool to manage and control risk with institutions changing their market exposure to maintain their VaR at a prespecified level.

The theory behind VaR is quite simplistic, actually too simplistic: VaR is defined as

the loss level that will not be exceeded with a certain confidence level over a certain period of time.

Value at Risk II

Again, this is related to the idea of downside risk, which measures the likelihood that a financial instrument or portfolio will lose value.

Downside risk can be measured by quantiles, which are the basis of the mathematics behind VaR. We now introduce a formal definition of VaR.

Value at Risk III

VaR is related to the potential loss on our portfolio, due to downside risk, over the time horizon H. Define this loss L_H as the difference between the value of the portfolio today (time 0) and in the future H.

$$L_H = \text{Portfolio}_0 - \text{Portfolio}_H.$$

For the scope of this course, we will assume that the loss L_H is a continuous random variable. This will avoid a number of technical subtleties that would lengthen the exposition considerably.
Consistently with earlier notation, we may call $\Pi(t, T)$ the sum of all future cash flows in $[t, T]$, discounted back at t, for our portfolio. These are random cash flows and not yet prices. Price of the portfolio at t is

$$\text{Portfolio}_t = \mathbb{E}_t^{\mathbb{Q}}[\Pi(t, T)].$$

T is usually the final maturity of the portfolio, and typically $H << T$.

Value at Risk IV

For example, if the portfolio is just a stock forward contract where at maturity we pay fixed K and receive the stock S_T, then the payout is written, as we have seen earlier, for $t \leq T$, as

$$\Pi(t, T) = D(t, T)(S_T - K).$$

If our portfolio is for example an amount A of a forward contract on a first stock $S^{(1)}$ with strike K_1 and maturity T_1 and an amount B of put options on a second stock $S^{(2)}$ with strike K_2 and maturity $T_2 > T_1$ we get, $t < T_1$:

$$\Pi(t, T_2) = A\, D(t, T_1)(S^{(1)}_{T_1} - K_1) + B\, D(t, T_2)(K_2 - S^{(2)}_{T_2})^+.$$

Value at Risk V

$\text{VaR}_{H,\alpha}$ with horizon H and confidence level α is defined as that number such that

$$\mathbb{P}[L_H < \text{VaR}_{H,\alpha}] = \alpha$$

or,

$$\boxed{\mathbb{P}[\mathbb{E}_0^{\mathbb{Q}}[\Pi(0,T)] - \mathbb{E}_H^{\mathbb{Q}}[\Pi(H,T)] < \text{VaR}_{H,\alpha}] = \alpha}$$

so that our loss at time H is smaller than $\text{VaR}_{H,\alpha}$ with $\mathbb{P}$-probability α.

In other terms, it is that level of loss over a time H that we will not exceed with $\mathbb{P}$-probability α. It is the α $\mathbb{P}$-percentile of the loss distribution at time H.

From this last equation, notice the interplay of the two probability measures.

Value at Risk VI

From the dialogue by Brigo (2011). "Counterparty Risk FAQ: Credit VaR, PFE, CVA, DVA, Closeout, Netting, Collateral, Re-hypothecation, WWR, Basel, Funding, CCDS and Margin Lending". See also the book by Brigo, Morini and Pallavicini: "Credit, Collateral and Funding", Wiley, March 2013.

A: VaR is calculated through a simulation of the basic financial variables underlying the portfolio under the historical probability measure, commonly referred as $\mathbb{P}$, up to the risk horizon H. At the risk horizon, the portfolio is priced in every simulated scenario of the basic financial variables, including defaults, obtaining a number of scenarios for the portfolio value at the risk horizon.

Value at Risk VII

Q: So if the risk horizon H is one year, we obtain a number of scenarios for what will be the value of the portfolio in one year, based on the evolution of the underlying market variables and on the possible default of the counterparties.

A: Precisely. A distribution of the losses of the portfolio is built based on these scenarios of portfolio values. When we say ”priced” we mean to say that the discounted future cash flows of the portfolio after the risk horizon are averaged conditional on each scenario at the risk horizon but under another probability measure, the Pricing measure, or Risk Neutral measure, or Equivalent Martingale Measure if you want to go technical, commonly referred as $\mathbb{Q}$.

Q: Not so clear... [Looks confused]

Value at Risk VIII

A: [Sighing] All right, suppose your portfolio has a call option on equity, traded with a Corporate client, with a final maturity of two years. Suppose for simplicity there is no interest rate risk, so discounting is deterministic. To get the Var, roughly, you simulate the underlying equity under the P measure up to one year, and obtain a number of scenarios for the underlying equity in one year.

Q: Ok. We simulate under P because we want the risk statistics of the portfolio in the real world, under the physical probability measure, and not under the so called pricing measure Q.

Value at Risk IX

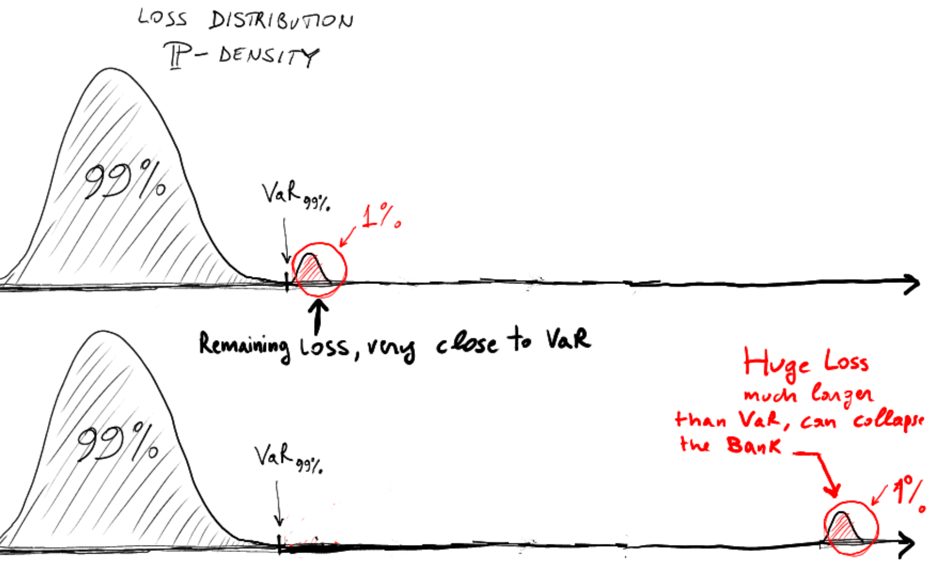
A: That's right. And then in each scenario at one year, we price the call option over the remaining year using for example a Black Scholes formula. But this price is like taking the expected value of the call option payoff in two years, conditional on each scenario for the underlying equity in one year. Because this is pricing, this expected value will be taken under the pricing measure Q, not P. This gives the Black Scholes formula if the underlying equity follows a geometric brownian motion under Q.

VaR drawbacks and Expected Shortfall I

VaR has a number of drawbacks. We list two of them now, starting from the most relevant.

VaR drawback 1: VaR does not take into account the tail structure beyond the percentile.

Consider the following two cases.


LOSS DISTRIBUTION
P- DENSITY
99%
VaR 99%
1%
Remaining Loss, very close to VaR
99%
VaR 99%
Huge Loss much larger than VaR, can collapse the Bank
1%

VaR drawbacks and Expected Shortfall I

From the picture above we see that we may have two situations where the VaR is the same but where the risks in the tail are dramatically different.

In the first case, the VaR singles out a 99% percentile, after which a slightly larger loss follows with 1% probability mass. The bank may be happy to know the 99% percentile in this case and to base its risk decision on that.

In the second case, the VaR singles out the same 99% percentile, after which an enormously much larger loss concentration follows with probability 1%. For example, this is now so large to easily collapse the bank. Would the bank be happy to ignore this potential huge and devastating loss, even if it has a small 1% probability?

VaR drawbacks and Expected Shortfall II

Probably not, and in this second case the bank would not base its risk analysis on VaR at 99%.

The VaR at 99% does not capture this difference in the two distributions, and if the bank does not explore the tail structure, it cannot know the real situation.

The most dangerous situation is the bank computing VaR and thinking it is in the first situation when it is actually in the second one.

VaR drawbacks and Expected Shortfall III

VaR drawback 2: VaR is not sub-additive on portfolios.
Suppose we have two portfolios P_1 and P_2, and a third portfolio $P = P_1 + P_2$ that is given by the two earlier portfolios together.
VaR at a given confidence level and horizon would be sub-additive if

$$\text{VaR}(P_1 + P_2) \leq \text{VaR}(P_1) + \text{VaR}(P_2) \quad \text{(VaR subadditivity. Is it true?)}$$

ie the risk of the total portfolio is smaller than the sum of the risks of its sub-portfolios (benefits of diversification, among other things).

However, *this is not true*. It may happen that

$$\text{VaR}(P_1 + P_2) > \text{VaR}(P_1) + \text{VaR}(P_2) \quad \text{in some cases.}$$

While such cases are usually difficult to see in practice, it is worth keeping this in mind.

VaR drawbacks and Expected Shortfall IV

As a remedy to this sub-additivity problem (and only partly to the first drawback) Expected Shortfall (ES) has been introduced.

ES requires to compute VaR first, and then takes the expected value on the TAIL of the loss distribution for values larger than VaR, conditional on the loss being larger than Value at Risk.

ES is sub-additive (solves drawback 2).

ES looks at the tail after VaR, but only in expectation, without analyzing the tail structure carefully. Hence, it is only a partial solution to drawback 1.

Expected Shortfall Definition I

Recalling that we defined the loss L_H as the difference between the value of the portfolio today (time 0) and in the future H.

$$L_H = \text{Portfolio}_0 - \text{Portfolio}_H,$$

ES for this portfolio at a confidence level α and a risk horizon H is

$$\text{ES}_{H,\alpha} = \mathbb{E}^{\mathbb{P}}[L_H | L_H > \text{VaR}_{H,\alpha}]$$

By definition, ES is always larger than the corresponding VaR.
Note that ES is defined through a conditional expectation. Recall that, by definition of conditional expectation,

$$\text{ES}_{H,\alpha} = \mathbb{E}^{\mathbb{P}}[L_H | L_H > \text{VaR}_{H,\alpha}] = \frac{\mathbb{E}^{\mathbb{P}}[L_H 1_{\{L_H > \text{VaR}_{H,\alpha}\}}]}{\mathbb{P}\{L_H > \text{VaR}_{H,\alpha}\}} =$$

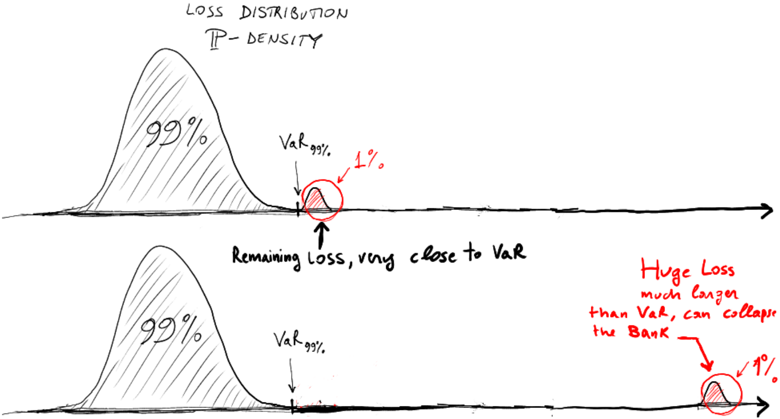
Expected Shortfall Definition II

$$= \frac{\mathbb{E}^{\mathbb{P}}[L_H 1_{\{L_H > \mathrm{VaR}_{H,\alpha}\}}]}{1 - \mathbb{P}\{L_H \le \mathrm{VaR}_{H,\alpha}\}} = \frac{\mathbb{E}^{\mathbb{P}}[L_H 1_{\{L_H > \mathrm{VaR}_{H,\alpha}\}}]}{1 - \alpha}$$

as, by definition, probability that Loss is below VaR at a given confidence level is equal to the confidence level itself.
Be aware of the fact that ES has several other names, and there are other risk measures that are defined very similarly. Names you may hear are:
Conditional value at risk (CVaR), average value at risk (AVaR), and expected tail loss (ETL).

Expected Shortfall drawbacks I

Drawback 1 of ES: tail structure. Go back to drawback 1 of VaR and look at the Figure there.



Expected Shortfall drawbacks II

ES does not fully solve this problem. It takes the average of the tail, so it will signal that the second portfolio is more risky than the first, but it won't show how the risk is structured. It is therefore an improvement over VaR but it is a blunt instrument to assess the risk in the tail.

Drawback 2 of ES (and drawback 3 of VaR): Liquidity risk.
Another problem of ES (and VaR) is that it is homogeneous with respect to the portfolio size. Namely, if k is a positive constant, then

$VaR(k\ Portfolio) = k\ VaR(Portfolio)$
and
$ES(k\ Portfolio) = k\ ES(Portfolio)$.

This is unrealistic and completely neglects liquidity risk and market impact. Selling one million shares is more than one million times risky

Expected Shortfall drawbacks III

than selling one share. Placing the order for selling one million shares will move the whole market and change the share price (theory of maket impact/market microstructure) with potential additional losses due to said market impact, whereas placing the order for one share will not move the market. Liquidity risk strongly disagrees with the homogeneous assumption.

For example, let Ford shares be trading at \$24 each. Placing the order for selling one million shares will alert the market that one important market player thinks Ford will be losing value, as they are trying to sell a huge amount of shares. The market will react instantly and the bid price will go down to \$23. This will cause additional loss to the market player who will face a loss of \$1 for each sold share because this player will receive \$23 instead of \$24 for each sold share. Instead, if a player places the order to sell one share, the price will not move from

Expected Shortfall drawbacks IV

\$24, as one single share does not signal a trend. It follows that one million times the impact of selling one share is zero, whereas the impact of selling a block of one million shares will be \$1 million.

Risk Measures: Numerical examples and software codes

In this part we will look at numerical examples of VaR and ES applied to financial portfolios, with software code provided, so that you can play with the code and come up with your own examples. The code will be available in Octave/Matlab or Python. We will loook at

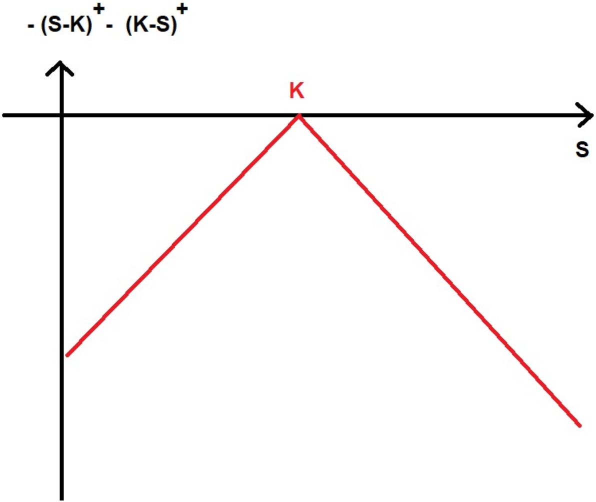
- Short Straddle (remember the Barings collapse)
- Long Risk Reversal
- Long Bull call spread
- Options on different correlated stocks

The mock exams will have further problems on risk measures.

Value at Risk and ES: Short Straddle I

Given a stock S_t, consider a payoff where we sell a call option with strike K and a put option with the same strike, both with maturity T. The payoff is $Y = -(S_T - K)^+ - (K - S_T)^+ = -|S_T - K|$.

Value at Risk and ES: Short Straddle II



Value at Risk and ES: Short Straddle III

If we include the inital price of Y in the payoff itself, the initial price we pay will be negative, as the payoff is always negative or zero, meaning that we will receive a positive cash flow at time 0 from selling the two options. We would then have to shift the payoff plot upwards of an amount equal to the initial price to include the initial price of the trade in the overall payoff.

The above short straddle is famous because it is one of the trades of Leeson that led to the collapse of Barings bank back in 1995. A trader entering a short straddle with payoff Y expects the stock to move very little, so that the initial premium she receives from selling the two options is much larger than the actual payoff that she will have to pay at maturity.

Value at Risk and ES: Short Straddle IV

Note that the more the stock moves away from K (typically $K = S_0$), the larger $|S_T - K|$ becomes, and thus the trader holding the short straddle $Y = -|S_T - K|$ will have to pay more money at maturity.

A trader should enter a short straddle position only if confident that the stock S will not move much.

The maximum gain is the price of the two options the trader sells initially. The maximum loss is potentially unlimited.

Value at Risk and ES: Short Straddle V

The risk factor of this portfolio is the stock price

$$dS_t = \mu S_t dt + \sigma S_t dW_t^P, \quad S_0 = s_0$$

given here under the measure P.
Assume the options have maturity $1y$ and we take a risk horizon $H = 0.25y$ (3 months, 3m).
We know from the Black Scholes formulas that we have seen earlier that the price of the payoff Y at time 0 is minus the price of the call with strike K minus the price of the put, namely

$$-[S_0\Phi(d_1(0,K)) - Ke^{-rT}\Phi(d_2(0,K))] + [S_0\Phi(-d_1(0,K)) - Ke^{-rT}\Phi(-d_2(0,K))].$$

The price of the payoff 3m in the future will be, setting $\bar{T} = T - 3m = T - 0.25$,

Value at Risk and ES: Short Straddle VI

$$-[S_{3m}\Phi(d_1(3m,K)) - Ke^{-r\bar{T}}\Phi(d_2(3m,K))] + [S_{3m}\Phi(-d_1(K)) - Ke^{-r\bar{T}}\Phi(-d_2(K))]$$

where

$$d_{1,2}(t,K) = \frac{\ln(S_t/K) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

For the loss at 3m, this is the portfolio price at time 0 minus the portfolo price at time $3m$. The only random quantity in the loss will be S_{3m}. We write $L_{3m}(S_{3m}) =$

$$-[S_0\Phi(d_1(0,K)) - Ke^{-rT}\Phi(d_2(0,K))] + [S_0\Phi(-d_1(0,K_2)) - K_2e^{-rT}\Phi(-d_2(K_2))]$$

$$+S_{3m}\Phi(d_1(3m,K)) - Ke^{-r\bar{T}}\Phi(d_2(3m,K)) + Ke^{-r\bar{T}}\Phi(-d_2(K,3m)) - S_{3m}\Phi(-d_1(K))$$

Value at Risk and ES: Short Straddle VII

We need to simulate many scenarios of S_{3m} up to 3m **under the measure P** and plug all scenarios in $L_{3m}(S_{3m})$, getting many scenarios for L_{3m}. From these scenarios we can isolate the α percentile, giving VaR, and average the loss conditional on it being larger than VaR, getting expected shortfall.

Simulating S up to 3m is easy as we know its distribution:

$$S_{3m} = S_0 \exp((\mu - \frac{\sigma^2}{2})0.25 + \sigma W_{3m}) = S_0 \exp((\mu - \frac{\sigma^2}{2})0.25y + \sigma\sqrt{0.25y}\mathcal{N}(0,1))$$

where we used that W_{3m} is normally distributed with variance 0.25. It is enough therefore to sample N scenarios from the standard normal distribution $\mathcal{N}(0,1)$, plug each scenario in the exponent of the above formula, get N scenarios for S_{3m} and with those get N scenarios for $L_{3m}(S_{3m})$. Once we have these N scenarios we can select the correct percentiles for VaR and ES.

Value at Risk and ES: Short Straddle VIII

Suppose indeed that we wish to get the $\alpha = 99\% = 0.99$ confidence level VaR for a 3m risk horizon. Assume: $S_0 = 100, K = 100$, $r = 1\% = 0.01$, $T = 1y$, $\mu = 5\%$, $\sigma = 0.5 = 50\%$, $N = 100000$
We use a Matlab/Octave code I made available.

General note. Octave is freeware, can be downloaded and installed for free and I find it more convenient for prototyping than R or Python. It is very good at vectorizing operations and very convenenient for plots and graphs. However, given the emphasis on Python and the fact that you have studied it as part of your education at Imperial, I have written Python codes equivalent to the Octave ones you find here. These Python codes are listed at the end of this slides set. They will also be made available to you as source codes.

Value at Risk and ES: Short Straddle IX

I recommend that you play with the codes, change some parameters and see how the risk measures change, understand why, and even code risk measures for different combinations of put and call options.

Value at Risk and ES: Short Straddle

```
pkg load statistics
S0 =100; k=100; Sigma=0.5; r=0.01; miu=0.05;
T=1;
confidence=0.99;
n=100000;
% call and put at time 0
d1c=(log(S0/k1)+(r+0.5*Sigma^2)*T)/(Sigma*T^0.5);
d1p=(log(S0/k2)+(r+0.5*Sigma^2)*T)/(Sigma*T^0.5);
c0=S0*normcdf(d1c,0,1)
          -k1*exp(-r*T)*normcdf(d1c-Sigma*T^0.5,0,1);
p0=-S0*normcdf(-d1p,0,1)
          +k2*exp(-r*T)*normcdf(-d1p+Sigma*T^0.5,0,1);
v0=-c0-p0;
```

Value at Risk and ES: Short Straddle

```
% computing call and put prices after h years
h= 0.25;
T=T-h;
Zt=normrnd(0,1,1,n);
St=S0*exp((miu-0.5*Sigma^2)*h)*exp(Zt.*Sigma*(h^0.5));
ct=zeros(1,n);
pt=zeros(1,n);
for i=1:n;
d1cnew=(log(St(i)/k1)+(r+0.5*Sigma^2)*T)/(Sigma*T^0.5)
d1pnew=(log(St(i)/k2)+(r+0.5*Sigma^2)*T)/(Sigma*T^0.5)
ct(i)=St(i)*normcdf(d1cnew,0,1)
    -k1*exp(-r*T)*normcdf(d1cnew-Sigma*T^0.5,0,1);
pt(i)=-St(i)*normcdf(-d1pnew,0,1)
 +k2*exp(-r*T)*normcdf(-d1pnew+Sigma*T^0.5,0,1);
end;
vt=-ct-pt;
```

Value at Risk and ES: Short Straddle

```
% vvar is Loss_3m
vvar=v0-vt;
vvar=sort(vvar);
ivar = round(confidence*n);
var = vvar(ivar);
ESv=mean(vvar(floor((confidence)*n):n));
% output histograms
figure(1);
hist(vvar,100);
xlabel('P&L');
ylabel('Frequencies');
title('Histogram of L3m of Portfolio');
var
ESv
```

VaR and ES: Short Straddle I

$\sigma = 0.5$: $\mathrm{Call}_0 = 20.144$; $\mathrm{Put}_0 = 19.149$.
$V0 = -Call_0 - Put_0 = -39.294$.
Ceteris paribus:
$\sigma = 0.5 = 50\% \implies VaR = 41.501;\ \ ES = 55.786;$
$\sigma = 0.7 = 70\% \implies VaR = 69.180;\ \ ES = 95.157;$
$\sigma = 0.2 = 20\% \implies VaR = 13.380;\ \ ES = 17.424;$

Please note how your risk measure depends crucially on the volatility σ. If your assessement of future volatility is wrong being too low, you will suffer much bigger losses (Leeson's case).

Suppose you think the volatility will stay at 20%, so you expect a loss over three months to be below 13.38 millions with 99% confidence . But if the volatility is instead 50% your loss will be below the much larger 41.501 millions that could break the bank.

VaR and ES: Short Straddle II

Leeson however didn't even have VaR or ES measures, so he couldn't run the above scenarios. Not that this would have stopped him, but a risk controller looking at the VaR/ES figures might have.

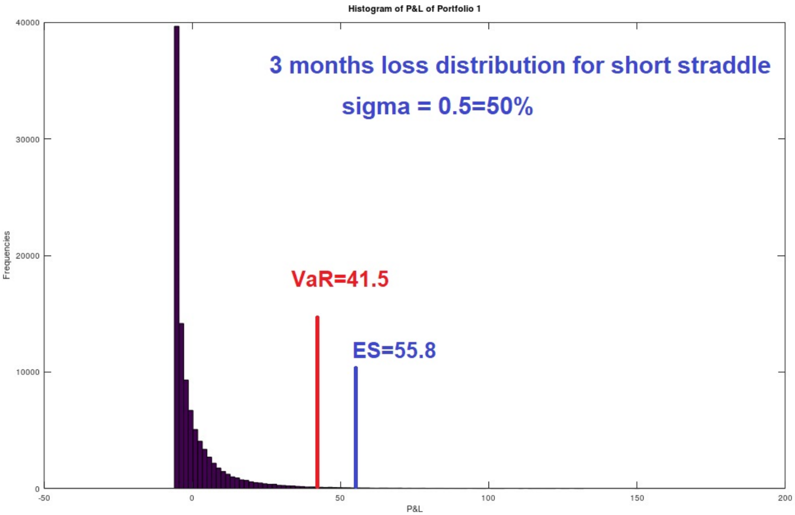
In the next plots we look at the density of the $Loss_{3m}$, namely we plot

$$x \mapsto p_{Loss_{3m}}(x),$$

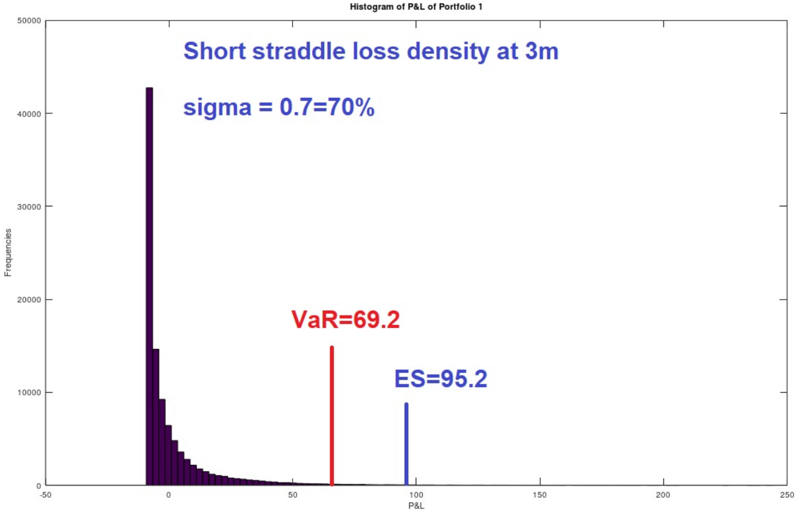
its 99th percentile and the expectation of the loss tail beyond the 99th percentile, conditional on the loss beyond the 99th percentile.

Note that the loss distribution has no left tail. This is because the maximum gain (negative loss) you can make is the initial premium of the options, whereas your maximum loss (right tail) is unlimited.

VaR and ES: Short Straddle III

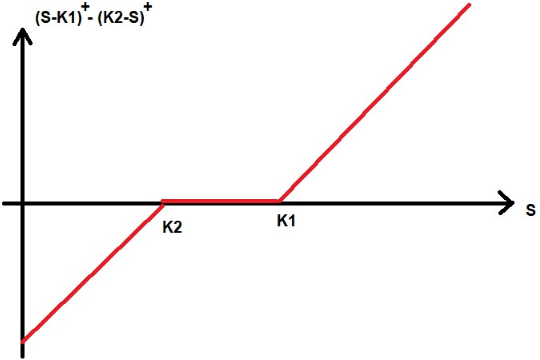


VaR and ES: Short Straddle IV



Value at Risk and ES: Risk Reversal I

Given a stock S_t, consider a payoff given by a call option with strike K_1 minus a put option with strike $K_2 < K_1$, both with maturity T. The payoff is $Y = (S_T - K_1)^+ - (K_2 - S_T)^+$.



Value at Risk and ES: Risk Reversal II

If we include the inital price of Y in the payoff itself, the initial price may be positive or negative depending on the strikes and other parameters. We would then have to shift the plot of the initial price to include the initial price of the trade in the overall payoff.

The above risk reversal is called a “bull risk reversal” or “long risk reversal”. This refers to the fact that a trader buying this is bullish (e.g. is confident) on a price increase for the underlying stock, expecting a large gain.

Value at Risk and ES: Risk Reversal III

Indeed, this payoff allows one to profit with $S_T - K_1$ when the stock goes above K_1 at maturity, to draw and get 0 if the stock stays between K_1 and K_2, and to lose $K_2 - S_T$ if the stock goes below K_2. A bullish trader will expect the first scenario to happen.

Please note that while the potential profit is infinite, as S can grow arbitrarily large in principle, the potential loss is limited, as S can at worst hit zero, so the maximum loss from the payoff is $K_2 - 0$, which may however be quite substantial in practice.

Given the potential for a loss, this portfolio will be obviously cheaper than a pure call options with strike K_1, and some aggressive investors willing to face a limited loss may be willing to pay the put in a bad scenario in order to reduce the option price for potentially benefiting from the call.

Value at Risk and ES: Risk Reversal IV

Essentially the trader buys the call with a view to benefit from a stock price increase and pays for the call initially by selling a put option.

When the risk reversal is used as an aggressive “bull” trade, the trader is essentially putting on a trade for close to no cost or even a credit when the put is more expensive than the call. If the trader is correct, and the stock continues increasing, the short put will become worthless while the long call will increase in value, generating a good profit.

The above use is speculative, but Risk reversals can be used also for protection: it can protect a trader who is short or indebted at the underlying stock (and will thus have to pay the future value of the stock price) from a rising stock price at a limited cost.

Value at Risk and ES: Risk Reversal V

The risk factor of this portfolio is the stock price

$$dS_t = \mu S_t dt + \sigma S_t dW_t^P, \quad S_0 = s_0$$

given here under the measure P.
Assume the options have maturity $5y$ and we take a risk horizon $H = 1y$.
We know from the Black Scholes formulas that we have seen earlier that the price of the payoff Y at time 0 is the price of the call with strike K_1 minus the price of the put with strike K_2, namely

$$S_0\Phi(d_1(0,K_1)) - K_1 e^{-rT}\Phi(d_2(0,K_1)) + [S_0\Phi(-d_1(0,K_2)) - K_2 e^{-rT}\Phi(-d_2(0,K_2))].$$

The price of the payoff at one year in the future will be, $\bar{T} = T - 1y$,

$$S_{1y}\Phi(d_1(1y,K_1)) - K_1 e^{-r\bar{T}}\Phi(d_2(1y,K_1)) + [S_{1y}\Phi(-d_1(K_2)) - K_2 e^{-r\bar{T}}\Phi(-d_2(K_2))]$$

Value at Risk and ES: Risk Reversal VI

where

$$d_{1,2}(t,K) = \frac{\ln(S_t/K) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

For the loss at 1y, this is the portfolio price at time 0 minus the portfolo price at time $1y$. The only random quantity in the loss will be S_{1y}. We write $L_{1y}(S_{1y}) =$

$$S_0\Phi(d_1(0,K_1)) - K_1 e^{-rT}\Phi(d_2(0,K_1)) + [S_0\Phi(-d_1(0,K_2)) - K_2 e^{-rT}\Phi(-d_2(K_2))]$$

$$-\Big(S_{1y}\Phi(d_1(1y,K_1)) - K_1 e^{-r\bar{T}}\Phi(d_2(1y,K_1)) +$$

$$+[S_{1y}\Phi(-d_1(K_2)) - K_2 e^{-r\bar{T}}\Phi(-d_2(K_2))]\Big).$$

Value at Risk and ES: Risk Reversal VII

We need to simulate many scenarios of S_{1y} up to 1y **under the measure P** and plug all scenarios in $L_{1y}(S_{1y})$, getting many scenarios for L_{1y}. From these scenarios we can isolate the α percentile, giving VaR, and average the loss conditional on it being larger than VaR, getting expected shortfall.

Simulating S up to 1y is easy as we know its distribution:

$$S_{1y} = S_0 \exp((\mu - \sigma^2/2)1y + \sigma W_{1y}) = S_0 \exp((\mu - \sigma^2/2)1y + \sigma\sqrt{1y}\mathcal{N}(0,1))$$

where we used the fact that W_{1y} is normally distributed with variance 1.

It is enough therefore to sample N scenarios from the standard normal distribution $\mathcal{N}(0,1)$, plug each scenario in the exponent of the above formula, get N scenarios for S_{1y} and with those get N scenarios for $L_{1y}(S_{1y})$. Once we have these N scenarios (say $N = 10000$) we can select the correct percentiles for VaR and ES.

Value at Risk and ES: Risk Reversal VIII

Suppose indeed that we wish to get the $\alpha = 95\% = 0.95$ confidence level VaR for a 1y risk horizon. Assume: $S_0 = 100, K_1 = 90, K_2 = 110$, $r = 1\% = 0.01$, $T = 5y$, $\mu = 5\%$, $\sigma = 0.2 = 20\%$, $N = 10000$
We use a Matlab/Octave code I made available.

Value at Risk and ES: Risk Reversal

```
pkg load statistics
S0 =100; k1=110; k2=90; Sigma=0.2; r=0.01; miu=0.05;
T=5;
confidence=0.95;
n=10000;
% call and put at time 0
d1c=(log(S0/k1)+(r+0.5*Sigma^2)*T)/(Sigma*T^0.5);
d1p=(log(S0/k2)+(r+0.5*Sigma^2)*T)/(Sigma*T^0.5);
c0=S0*normcdf(d1c,0,1)
            -k1*exp(-r*T)*normcdf(d1c-Sigma*T^0.5,0,1);
p0=-S0*normcdf(-d1p,0,1)
           +k2*exp(-r*T)*normcdf(-d1p+Sigma*T^0.5,0,1);
v0=c0-p0;
c0
p0
```

Value at Risk and ES: Risk Reversal

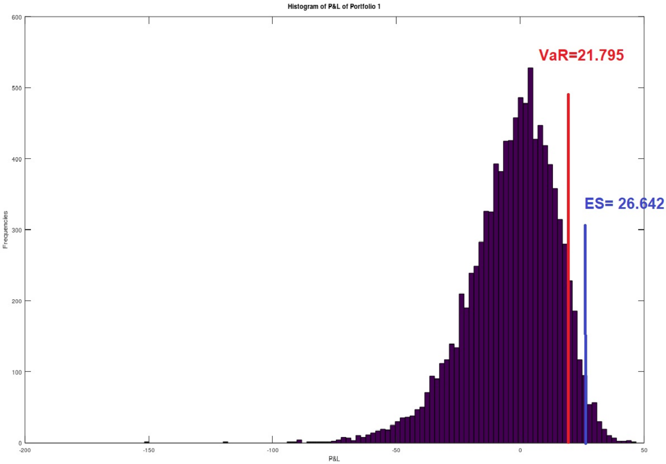
```
% computing call and put prices after one year
T=T-1;
Zt=normrnd(0,1,1,n);
St=S0*exp(miu-0.5*Sigma^2)*exp(Zt.*Sigma);
ct=zeros(1,n);
pt=zeros(1,n);
for i=1:n;
d1cnew=(log(St(i)/k1)+(r+0.5*Sigma^2)*T)/(Sigma*T^0.5)
d1pnew=(log(St(i)/k2)+(r+0.5*Sigma^2)*T)/(Sigma*T^0.5)
ct(i)=St(i)*normcdf(d1cnew,0,1)
    -k1*exp(-r*T)*normcdf(d1cnew-Sigma*T^0.5,0,1);
pt(i)=-St(i)*normcdf(-d1pnew,0,1)
 +k2*exp(-r*T)*normcdf(-d1pnew+Sigma*T^0.5,0,1);
end;
vt=ct-pt;
```

Value at Risk and ES: Risk Reversal

```
% vvar is Loss_1y
vvar=v0-vt;
vvar=sort(vvar);
ivar = round(confidence*n);
var = vvar(ivar);
ESv=mean(vvar(floor((confidence)*n):n));
% output histograms
figure(1);
hist(vvar,100);
xlabel('P&L');
ylabel('Frequencies');
title('Histogram of L1y of Portfolio 1');
var
ESv
```

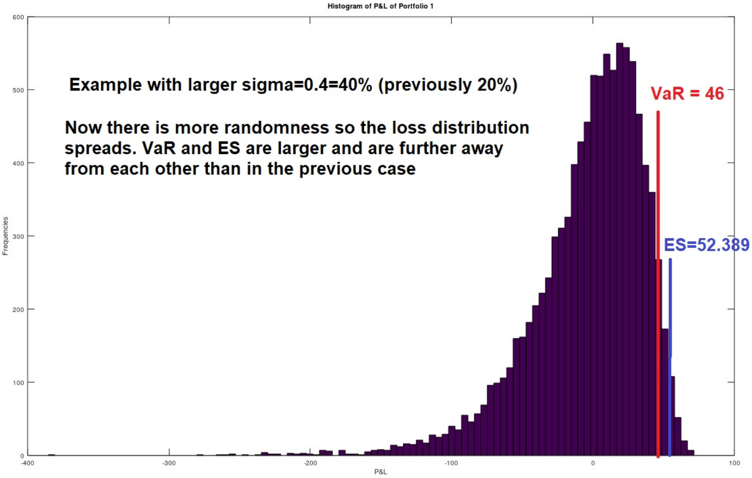
Value at Risk and ES: Risk Reversal

Running the code gives the following L_{1y} density.



Value at Risk and ES: Risk Reversal

Running the code gives the following L_{1y} density (new scale).



Value at Risk and ES: Risk Reversal I

Above we discussed a bull or long risk reversal.

One can also have a bear or short risk reversal. This happens when one sells the call and buys the put, leading to the payoff

$$Y = (K_2 - S_T)^+ - (S_T - K_1)^+.$$

Exercise: draw the payoff of a bear risk reversal; explain why a trader might buy this, what this trader would expect to happen to the underlying stock; how this can be used for speculation or protection in different circustances. You could also adapt the code above to assess value at risk and expected shortfall of a bear risk reversal.

Risk reversals are very popular in the FX market, but are used also in the equity markets.

Value at Risk and ES: Bull call spread I

A bull call spread payoff is the difference between a call with smaller strike and a call with a larger strike. The underlying asset and the option maturities are the same.

In formula: if $K_1 > K_2$, then

$$Y = (S_T - K_2)^+ - (S_T - K_1)^+.$$

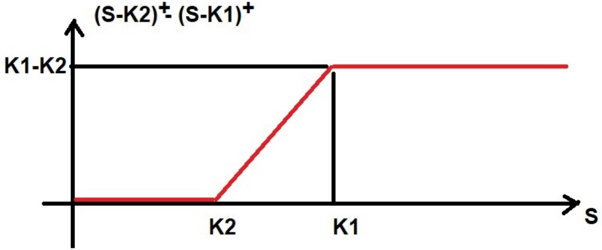
This can also be written as

$$Y = (K_1 - K_2)1_{S_T > K_1} + (S - K_2)1_{K_2 < S_T \leq K_1} + 0\ 1_{S \leq K_2}.$$

This contingent claim consists of one long call with a lower strike price and one short call with a higher strike.
Note that the initial price of Y would be positive to us, since it is an in-the-money call minus an out-of-the-money call. This means that to purchase this payoff we need to pay.

Value at Risk and ES: Bull call spread II

The payoff of a bull call spread, excluding the initial payment needed to buy the product, looks like



If we include the initial price we pay for purchasing the option, this will shift the plot down of that price

.

Value at Risk and ES: Bull call spread III

Who would buy this? A bull call spread profits when the underlying stock rises in price. Profit is limited as the stock price rises above the strike price K_1, and the loss is also limited as the stock price falls below the strike price K_2.

Value at Risk and ES: Bull call spread IV

In this sense it is safer than a risk reversal, as the loss is floored after the stock drops below K_2, but the potential profits are also capped as the stock rises above K_1. Hence this contract will be sought by a trader who does not want excessive risk and who is expecting the stock to increase.

Basically this contract is less expensive than a call option with stike K_2, as it reduces the price of the call by selling another call with a higher strike K_1.

Value at Risk and ES: Bull call spread I

The risk factor of this portfolio is the stock price

$$dS_t = \mu S_t dt + \sigma S_t dW_t^P, \quad S_0 = s_0$$

given here under the measure P.
Assume the options have maturity $5y$ and we take a risk horizon $H = 1y$.
We know from the Black Scholes formulas that we have seen earlier that the price of the payoff Y at time 0 is the price of the call with strike K_2 minus the price of the call with strike K_1, namely

$$S_0\Phi(d_1(0,K_2)) - K_2 e^{-rT}\Phi(d_2(0,K_2)) - [S_0\Phi(d_1(0,K_1)) - K_1 e^{-rT}\Phi(d_2(0,K_1))].$$

The price of the payoff at one year in the future will be, $\bar{T} = T - 1y$

$$S_{1y}\Phi(d_1(1y,K_2)) - K_2 e^{-r\bar{T}}\Phi(d_2(1y,K_2)) - [S_{1y}\Phi(d_1(K_1)) - K_1 e^{-r\bar{T}}\Phi(d_2(K_1))]$$

Value at Risk and ES: Bull call spread II

where

$$d_{1,2}(t,K) = \frac{\ln(S_t/K) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

For the loss at 1y, this is the portfolio price at time 0 minus the portfolo price at time $1y$. The only random quantity in the loss will be S_{1y}. We write $L_{1y}(S_{1y}) =$

$$S_0\Phi(d_1(0,K_2)) - K_2e^{-rT}\Phi(d_2(0,K_2)) - [S_0\Phi(d_1(0,K_1)) - K_1e^{-rT}\Phi(d_2(K_1))]$$

$$-\Big(S_{1y}\Phi(d_1(1y,K_2)) - K_2e^{-r\bar{T}}\Phi(d_2(1y,K_2))$$

$$-[S_{1y}\Phi(-d_1(K_1)) - K_1e^{-r\bar{T}}\Phi(-d_2(K_1))]\Big).$$

Value at Risk and ES: Bull call spread III

We need to simulate many scenarios of S_{1y} up to 1y **under the measure P** and plug all scenarios in $L_{1y}(S_{1y})$, getting many scenarios for L_{1y}. From these scenarios we can isolate the α percentile, giving VaR, and average the loss conditional on it being larger than VaR, getting expected shortfall.

Simulating S up to 1y is easy as we know its distribution:

$$S_{1y} = S_0 \exp((\mu - \sigma^2/2)1y + \sigma W_{1y}) = S_0 \exp((\mu - \sigma^2/2)1y + \sigma\sqrt{1y}\mathcal{N}(0,1))$$

where we used the fact that W_{1y} is normally distributed with variance 1.

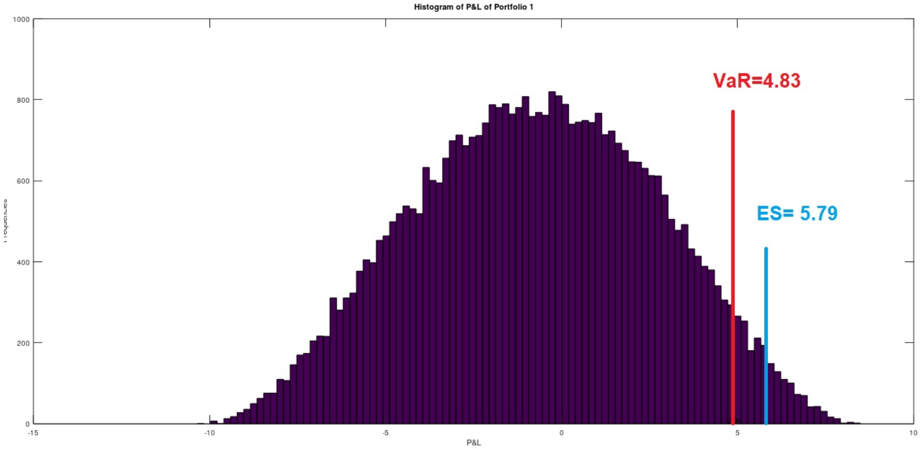
It is enough therefore to sample N scenarios from the standard normal distribution $\mathcal{N}(0,1)$, plug each scenario in the exponent of the above formula, get N scenarios for S_{1y} and with those get N scenarios for $L_{1y}(S_{1y})$. Once we have these N scenarios (say $N = 40000$) we can select the correct percentiles for VaR and ES.

Value at Risk and ES: Bull call spread IV

Suppose indeed that we wish to get the $\alpha = 95\% = 0.95$ confidence level VaR for a 1y risk horizon. Assume: $S_0 = 100, K_2 = 90, K_1 = 110$, $r = 1\% = 0.01$, $T = 5y$, $\mu = 5\%$, $\sigma = 0.2 = 20\%$, $N = 40000$
We use a Matlab/Octave code I made available.

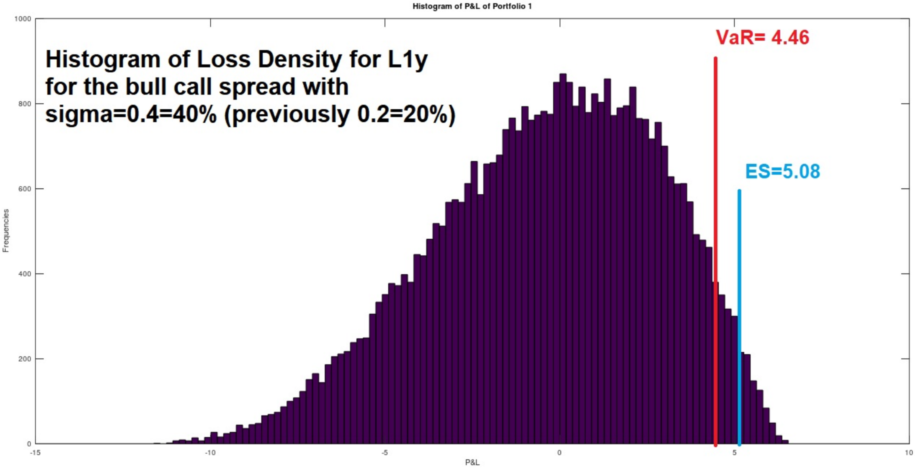
Value at Risk and ES: Bull Call spread

Running the code gives the following L_{1y} density.



Value at Risk and ES: Bull Call spread

Same but with $\sigma = 0.4$ instead of $\sigma = 0.2$.



Value at Risk and ES: Bull Call spread I

In this case increasing the volatility has decreased the risk measure. When we increase randomness of the underlying, it becomes likely that more underlying scenarios approach or even cross the strikes K_1 and K_2. Scenarios approaching or crossing K_1 will lead to higher values for the future payoff, and thus to lower losses. On the contrary, scenarios approaching or crossing K_2 will lead to lower values for the future payoff, and to higher losses.

A careful analysis of the risk of the options separately (try VaR and ES for call with strike K_1 and then VaR and ES for call with strike K_2) shows that, in the range of parameters we are using, the K_1 option risk increases more with the volatility than the K_2 option risk. As we are looking at plus K_2 option minus K_1 option, this means roughly that risk will go down as we increase the vol, because the effect will be stronger in increasing K_1 than K_2 and we are negative K_1.

Value at Risk and ES: Bull Call spread II

The effect of the two options combined will be to decrease the risk, ceteris paribus, when the volatility increases.

One can also have a bear spread call. This happens when one sells the call with the lower strike and buys the call with the higher strike, leading to the payoff

$$Y = (S_T - K_1)^+ - (S_T - K_2)^+$$

$$= 0\ 1_{\{S_T \leq K2\}} + (K_2 - S_T)1_{\{K_2 < S_T < K_1\}} - (K_1 - K_2)1_{\{S_T \geq K_1\}}.$$

Note that the bear spread initial value is negative, as we buy an out of the money call option and sell an in-the-money call option. This means that if we are the bank calculating the initial negative price, selling this payoff to an investor, the investor will be charged a negative price,

Value at Risk and ES: Bull Call spread III

meaning that we will pay him the positive value given by the opposite of the negative price we calculated. This compensates the client for the fact that her payoff at maturity will be always zero or negative.

Or if we are the client, this means that when we purchase Y we will also receive an initial premium at time 0, as the product cost/price is negative to us. This shifts the payoff up of that price amount, if we include it in the payoff.

Exercise: draw the payoff of a bear call spread; explain why a trader might buy this, what this trader would expect to happen to the underlying stock; how this can be used for speculation or protection in different circustances. You could also adapt the code above to assess value at risk and expected shortfall of a bear risk reversal.

VaR and ES: Options on different correlated stocks I

We now consider a portfolio with a call option on a first stock $S^{(1)}$ with strike K_1 and a put option on a second stock $S^{(2)}$ with strike K_2, both options with maturity T.

$$Y = (S_T^{(1)} - K_1)^+ + (K_2 - S_T^{(2)})^+.$$

$$d\ S_t^{(1)} = \mu_1 S_t^{(1)}\ dt + \sigma_1 S^{(1)} dW_t^{(1)},\ s_0^{(1)},$$

$$d\ S_t^{(2)} = \mu_2 S_t^{(2)}\ dt + \sigma_2 S^{(2)} dW_t^{(2)},\ s_0^{(2)},$$

$$dW^1 dW^2 = \rho\ dt.$$

Recall that ρ can be interpreted as an instantaneous correlation between changes in S^1 and S^2,

$$\text{“corr”}(dS_t^1, dS_t^2) = \rho.$$

VaR and ES: Options on different correlated stocks II

We assume

$S_0^{(1)} = 120,\ S_0^{(2)} = 80, \mu_1 = 0.05, \mu_2 = 0.02, \sigma_1 = 0.5, \sigma_2 = 0.2,\ \rho = \text{range}$

$T = 2y,\ \ K_1 = 116,\ \ K_2 = 86,\ \ r = 0.01,\ \ H = 0.25y = 3m,$ conf lev 95%,

We will look at a couple cases and then calculate several cases based on different values of the correlation ρ, ceteris paribus.

Our aim is to see the impact of changing ρ and σ on VaR and ES.

VaR and ES: Options on different correlated stocks III

We know from the Black Scholes formulas that we have seen earlier that the price of the payoff Y at time 0 is the price of the call on the first stock with strike K_1 plus the price of the put on the second stock with strike K_1, namely

$$S_0^{(1)}\Phi(d_1(0,K_1)) - K_1 e^{-rT}\Phi(d_2(0,K_1)) + K_2 e^{-rT}\Phi(-d_2(0,K_2)) - S_0^{(2)}\Phi(-d_1(0,K_2)).$$

The price of the payoff at $H = 3m = 0.25y$ in the future will be

$$S_H^{(1)}\Phi(d_1(H,K_1)) - K_1 e^{-r\bar{T}}\Phi(d_2(H,K_1)) + K_2 e^{-r\bar{T}}\Phi(-d_2(H,K_2)) - S_H^{(2)}\Phi(-d_1(H,K_2$$

where $\bar{T} = T - 0.25y$ and

$$d_{1,2}(H,K_i) = \frac{\ln(S_H^{(i)}/K_i) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-H)}{\sigma\sqrt{T-H}}.$$

VaR and ES: Options on different correlated stocks IV

For the loss at time H, this is the portfolio price at time 0 minus the portfolo price at time H. The only random quantity in the loss will be the stocks $S_H^{1,2}$. We write $L_H(S_H^{1,2}) =$

$$S_0^{(1)}\Phi(d_1(0,K_1)) - K_1 e^{-rT}\Phi(d_2(0,K_1)) + K_2 e^{-rT}\Phi(-d_2(K_2)) - S_0^{(2)}\Phi(-d_1(K_2))$$

$$-S_H^{(1)}\Phi(d_1(H,K_1)) + K_1 e^{-r\bar{T}}\Phi(d_2(H,K_1)) - K_2 e^{-r\bar{T}}\Phi(-d_2(K_2)) + S_H^{(2)}\Phi(-d_1(K_2)).$$

VaR and ES: Options on different correlated stocks V

We need to simulate many scenarios of $S_H^{(1,2)}$ up to H **under the measure P** and plug all scenarios in $L_H(S_H^{(1,2)})$, getting many scenarios for L_H. From these scenarios we can isolate the α percentile, giving VaR, and average the loss conditional on it being larger than VaR, getting expected shortfall.

Simulating $S_H^{(1,2)}$ up to H is easy as we know its distribution:

$$S_H^{(1)} = S_0^{(1)} \exp((\mu_1 - \sigma_1^2/2)H + \sigma_1 W_H^{(1)}) = S_0^{(1)} \exp((\mu_1 - \sigma_1^2/2)H + \sigma_1 \sqrt{H}\mathcal{N}_1)$$

$$S_H^{(2)} = S_0^{(2)} \exp((\mu_2 - \sigma_2^2/2)H + \sigma_2 W_H^{(2)}) = S_0^{(2)} \exp((\mu_2 - \sigma_2^2/2)H + \sigma_2 \sqrt{H}\mathcal{N}_2)$$

VaR and ES: Options on different correlated stocks VI

where $[\mathcal{N}_1, \mathcal{N}_1]$ is a bivariate normal random variable with zero means, variances equal to 1, and correlation or covariance ρ:

$$[\mathcal{N}_1, \mathcal{N}_2] \sim \mathcal{N}\left([0,\ 0], \left[\begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array}\right]\right).$$

This comes from the fact that $[W_H^{(1)}, W_H^{(2)}]$ is jointly normally distributed with mean $[0,\ 0]$ and variance/ covariance matrix

$$H\left[\begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array}\right].$$

It is enough therefore to sample N bivariate scenarios from the bivariate normal distribution $[\mathcal{N}_1, \mathcal{N}_2]$, plug each scenario in the exponents of the above formulas for the two stocks, get N scenarios

VaR and ES: Options on different correlated stocks VII

for $[S_H^{(1)}, S_H^{(2)}]$ and with those get N scenarios for $L_H(S_H^{(1,2)})$. Once we have these N scenarios we can select the correct percentiles for VaR and ES.

If you don't have a generator for correlated normal ranom variables, you can manage by correlating indipendent realizations from a single random number generator.

Assume you wish to have $\mathcal{N}_1$ and $\mathcal{N}_2$ but you only have one generator of standard normals. You generate independent realizations $\mathcal{N}_1^0$ and $\mathcal{N}_2^0$ from the same generator, by repeated simulation, and then mix them as follows to obtain the correlated samples for $\mathcal{N}_1$ and $\mathcal{N}_2$:

$$\mathcal{N}_1 = \mathcal{N}_1^0, \;\; \mathcal{N}_2 = \rho\mathcal{N}_1^0 + \sqrt{1-\rho^2}\mathcal{N}_2^0.$$

It is immediate to check that $\mathcal{N}_1$ and $\mathcal{N}_2$ are jointly normal with zero means, unit variances and correlation ρ.

VaR and ES: Options on different correlated stocks VIII

Calculations are through the Matlab/Octave code I make available below. Let's look at three key correlation cases and explain the pattern.

ρ	-1	0	1
VaR	30.9	26.1	17.4
ES	34.6	29.5	18.0
$ES - VaR$	3.7	3.4	0.6

Both risk measures decrease when ρ increases. This means that total positive correlation is less risky, for our portfolio, than total negative. The total effect of ρ on VaR is 30.9-17.4 = 13.5 and $(30.9 - 17.4)/17.4 = 0.78$. **Correl impacts on VaR is** 78%**.** For ES total effect we have 34.6-18 = 16.6, and $(34.6 - 18)/18 = 0.92$, **Correlation impacts on ES is** 92%**.** We also note that negative correl has larger $ES - VaR$, meaning the tail is deeper, more risk.

VaR and ES: Options on different correlated stocks IX

ρ	-1	0	1
VaR	30.9	26.1	17.4
ES	34.6	29.5	18.0
$ES - VaR$	3.7	3.4	0.6

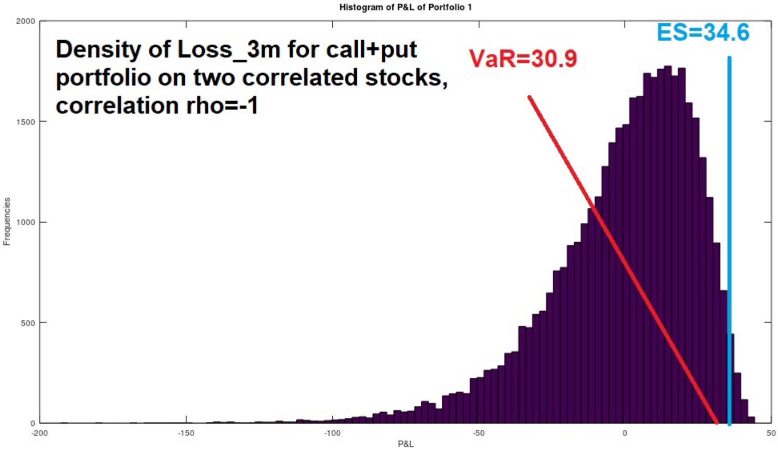
CASE 1. $\rho = -1 \implies$ totally negative correlation, when S^1 goes down, S^2 goes up. When S^1 goes down, this decreases the value of the call as it moves towards out of the money. At the same time S^2 can go only up, due to the extreme negative correlation. When S^2 goes up, the put is worth less as it also moves towards out of the money. So when the call loses money, the put does too due to the negative correlation. This means that we will face larger losses compared to the case where the correlation is less negative or positive, because portfolio$_H$ will be smaller when we subtract it from portfolio$_0$, leading to a bigger loss.

VaR and ES: Options on different correlated stocks X

CASE 2. $\rho = -1 \implies$ totally negative correlation. In this other case, due to total negative correlation, when S^1 goes up, S^2 goes down. Then the call becomes more valuable as it gets more in the money, while the put also becomes more valuable as it goes more in the money. In this case we have a doubly positive option value in H that subtracts a lot to the value of the portfolio at time 0, $\text{Loss}_H = \text{portfolio}_0 - \text{portfolio}_H$, creating a small or even negative loss. However, remember that for risk measures we care about large values of the loss, not negative or small, so the previous CASE 1 is the one that matters for VaR and ES.

Note also that $\rho = -1$ results in the largest $ES - VaR$ difference, which points to a deeper loss-distribution tail.

VaR and ES: Options on different correlated stocks XI



VaR and ES: Options on different correlated stocks XII

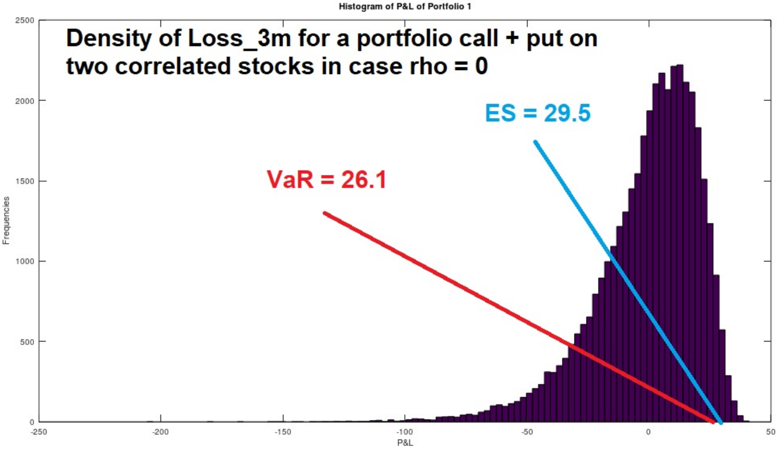
ρ	-1	0	1
VaR	30.9	26.1	17.4
ES	34.6	29.5	18.0
ES $-$ *VaR*	3.7	3.4	0.6

CASE 1: $\rho = 0 \implies$ zero correlation, when S^1 goes down, S^2 can go either up or down, the changes are unrelated. When S^1 goes down, this decreases the value of the call. At the same time S^2 can either up or down. When S^2 goes up, the put is worth less, when it goes down, the put is worth more. So when the call loses money, the put can either lose or gain money. Clearly this is less risky than the previous case. The scenarios that contribute to a larger loss are those where S^1 goes down and S^2 goes up, but due to the zero correlation these scenarios are less than in the case $\rho = -1$, and the loss will be smaller overall.

VaR and ES: Options on different correlated stocks XIII

Note also that $\rho = 0$ results in a smaller $ES - VaR$ difference wrt $\rho = -1$, which points to a less deep loss-distribution tail.

VaR and ES: Options on different correlated stocks XIV



VaR and ES: Options on different correlated stocks XV

ρ	-1	0	1
VaR	30.9	26.1	17.4
ES	34.6	29.5	18.0
ES − *VaR*	3.7	3.4	0.6

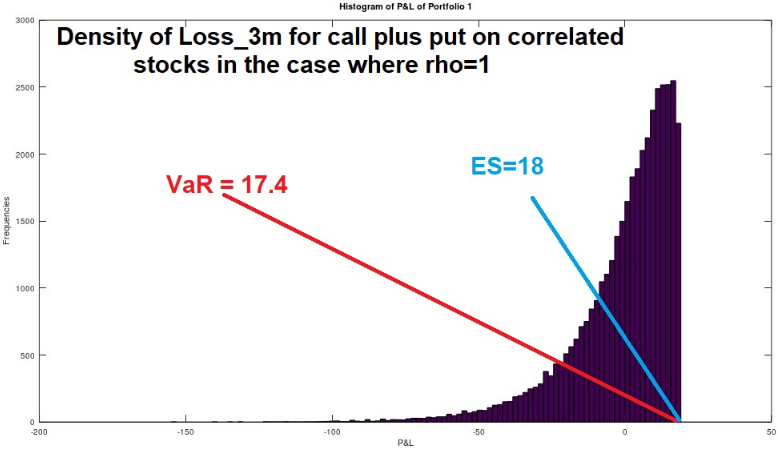
CASE1: $\rho = 1 \implies$ total positive correlation, when S^1 goes down, S^2 will go down too as they are totally correlated. When S^1 goes down, this decreases the value of the call. At the same time S^2 goes down and this increases the value of the put. So when the call loses money, the put makes money. Clearly this is less risky than the previous two cases, because in every scenario where the call has a loss, the put offsets that with a gain.

VaR and ES: Options on different correlated stocks XVI

CASE2: $\rho = 1 \implies$ total positive correlation, when S^1 goes up, S^2 will go up too as they are totally correlated. When S^1 goes up, this increases the value of the call. At the same time S^2 goes up and this decreases the value of the put. So when the call makes money, the put loses money. Clearly this is less risky than the previous two cases, because in every scenario where the call has a loss, the put offsets that with a gain and vice versa.

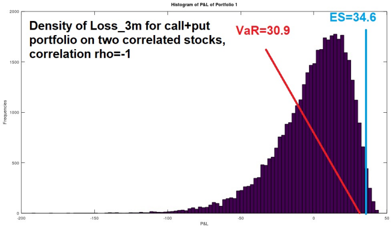
Note also that $\rho = 1$ results in the smallest $ES - VaR$ difference, 0.6 vs previous 3.7 and 3.4, which points to a very thin loss-distribution tail.

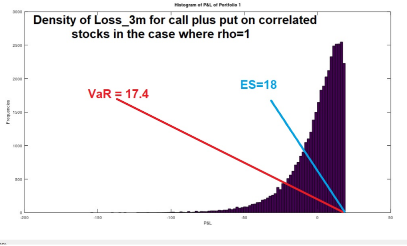
VaR and ES: Options on different correlated stocks XVII



VaR and ES: Options on different correlated stocks XVIII

Note how different the shape of the loss distribution is in the two cases $\rho = -1$ and $+1$. Especially the right tail.





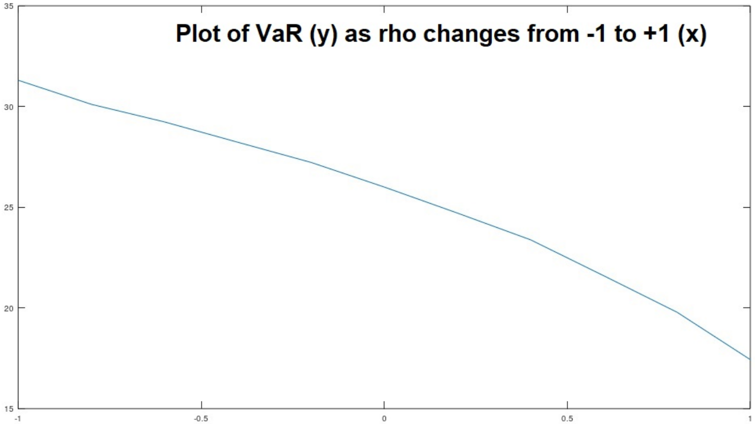
Recall that in the case +1 the put and the call offset each other in all scenarios, bringing the right tail of the loss to an abrupt halt.

VaR and ES: Options on different correlated stocks XIX

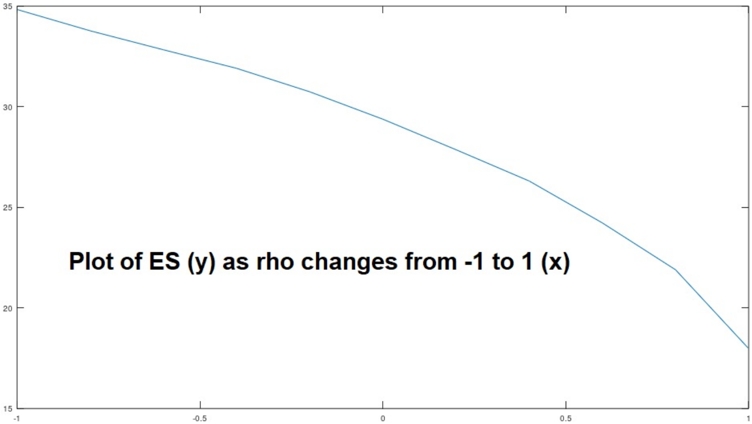
Now we let ρ span the interval $[-1, 1]$, ceteris paribus, and we see how VaR and ES change with ρ. For the impact of ρ on VaR and ES, we can look at the following table (V=VaR, E=ES) and plots

ρ	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1
V	30.9	30.3	29.2	28.2	27.4	26.1	24.9	23.3	21.7	19.8	17.4
E	34.6	34.0	32.8	31.9	30.9	29.5	28.0	26.4	24.4	21.9	18.0

VaR and ES: Options on different correlated stocks XX



VaR and ES: Options on different correlated stocks XXI



VaR & ES: Options on correlated assets - Codes

```
pkg load statistics
S10 =120; S20 =80; k1=116; k2=86;
Sigma1=0.5; Sigma2=0.2;
T=2; r=0.01; miu1=0.05; miu2=0.02;
rho = 0;
n=40000; confidence=0.95; h = 0.25;
% call and put price at time 0
d1c=(log(S10/k1)+(r+0.5*Sigma1^2)*T)/(Sigma1*T^0.5);
d1p=(log(S20/k2)+(r+0.5*Sigma2^2)*T)/(Sigma2*T^0.5);
c0=S10*normcdf(d1c,0,1)
    -k1*exp(-r*T)*normcdf(d1c-Sigma1*T^0.5,0,1);
p0=-S20*normcdf(-d1p,0,1)
   +k2*exp(-r*T)*normcdf(-d1p+Sigma2*T^0.5,0,1);
v0=c0+p0;
```

VaR & ES: Options on correlated assets - Codes

```
% computing the prices at t=h
T=T-h;
Z10t=normrnd(0,1,1,n);
Z20t=normrnd(0,1,1,n);
Z1t = Z10t; Z2t = rho*Z10t + ((1-rho^2)^0.5)*Z20t;
S1t=S10*exp((miu1-0.5*Sigma1^2)*h)
                    *exp(Z1t.*Sigma1*(h^0.5));
S2t=S20*exp((miu2-0.5*Sigma2^2)*h)
                    *exp(Z2t.*Sigma2*(h^0.5));
ct=zeros(1,n); pt=zeros(1,n);
```

VaR & ES: Options on correlated assets - Codes

```
for i=1:n;
    d1cnew=(log(S1t(i)/k1)+(r+0.5*Sigma1^2)*T)
                          /(Sigma1*T^0.5);
    d1pnew=(log(S2t(i)/k2)+(r+0.5*Sigma2^2)*T)
                          /(Sigma2*T^0.5);
    ct(i)=S1t(i)*normcdf(d1cnew,0,1)
        -k1*exp(-r*T)*normcdf(d1cnew-Sigma1*T^0.5,0,1)
    pt(i)=-S2t(i)*normcdf(-d1pnew,0,1)
     +k2*exp(-r*T)*normcdf(-d1pnew+Sigma2*T^0.5,0,1);
end;
vt=ct+pt;
vvar=v0-vt;
vvar=sort(vvar);
```

VaR & ES: Options on correlated assets - Codes

```
ivar = round(confidence*n);
var = vvar(ivar);
ESv=mean(vvar(floor((confidence)*n):n));
% output histogram
figure(1);
hist(vvar,100);
xlabel('P&L');
ylabel('Frequencies');
title('Histogram of Loss H');
rho
var
ESv
```

Volatilities and correlations I

More generally, volatilities, correlation, dynamics and statistical dependencies have a very important impact on risk.

For very large portfolios it is difficult to obtain intuition on why some risk patterns are observed, as there are too many assets and parameters.

A rigorous quantitative analysis of risks is fundamental to have a safe result. However, the assumptions underlying the analysis need to be kept in mind and stress-tested

PART 4: NUMERICAL SOLUTION OF SDEs I

PART 4: NUMERICAL SOLUTION OF SDEs

In the numerical examples of option pricing and risk measures we always used the Black Scholes model.

But what if we were to use a smile model? How would we simulate, for example, the stock price up to the risk horizion $t = H$? With Black Scholes this is easy, we know the solution

$$S_H = S_0 \exp((\mu - \sigma^2/2)H + \sigma\sqrt{H}\boxed{\mathcal{N}(0,1))}$$

and we only need a standard normal generator to simulate this.

PART 4: NUMERICAL SOLUTION OF SDEs II

In this part we look at how we can simulate a SDE that does not have a closed form solution, or whose solution may involve difficult special functions (CEV) or Fourier transforms methods (Heston), or where the solution of the SDE is not known despite its marginal probability law being known (mixture dynamics). If you noticed, in the mixture dynamics case we know the SDE solution is distributed as a mixture of lognormals but we never solved the SDE, because we don't know how to solve it. So we need a numerical method for some payoffs more complex than combinations of calls and put.

Euler scheme for numerical solutions of SDEs I

We illustrate the schemes for one dimensional SDEs. The Heston model, having two SDEs, would require a two-dimensional scheme, but this is easily generalized from the one dimensional scheme.
We start with the Euler Scheme for the general SDE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = Z$$

where Z is a random variable independent of W, or a deterministic constant.
The Euler scheme idea is very simple and consists of replacing differentials with increments. Take a time grid

$$t_0 = 0, t_1, t_2, \ldots, t_n = T$$

Euler scheme for numerical solutions of SDEs II

where T is the final time of the simulation, the final time where we need the SDE solution. Let the time step be $\Delta t = t_{i+1} - t_i = \delta$ for all i and write $\Delta W_{t_i} = W_{t_{i+1}} - W_{t_i}$, $\Delta X_{t_i} = X_{t_{i+1}} - X_{t_i}$
Hence the SDE becomes

$$\Delta X_{t_i} = \mu(t_i, X_{t_i})\Delta t_i + \sigma(t_i, X_{t_i})\Delta W_{t_i}, \quad X_0 = Z$$

and, writing

$$\Delta W_{t_i} = W_{t_{i+1}} - W_{t_i} \sim \sqrt{\delta}\mathcal{N}_i(0,1)$$

where $\mathcal{N}_i(0,1)$ is a standard normal and all normals with different i's are independent (because Brownian increments are independent).

Euler scheme for numerical solutions of SDEs III

Hence

$$X_{t_{i+1}} = X_{t_i} + \mu(t_i, X_{t_i})\Delta t_i + \sigma(t_i, X_{t_i})\sqrt{\delta}\mathcal{N}_i(0,1), \quad X_0 = Z.$$

This is the Euler scheme. Iteratively, given scenarios for X_{t_i} allows you to get a scenario for $X_{t_{i+1}}$ by simulating a standard normal $\mathcal{N}_i(0,1)$, everything else in the equation is known by the previous steps. The inital step in each scenario is sampling the distribution Z as X_0 or taking the constant value of the initial condition if it is deterministic.

Assume we plan to simulate N scenarios of the SDE above. For our notation, we denote the j-th scenario of the SDE solution by X^j, where the upper index does not denote power but scenario. We can write the scheme as

$$X^j_{t_{i+1}} = X^j_{t_i} + \mu(t_i, X^j_{t_i})\Delta t_i + \sigma(t_i, X^j_{t_i})\sqrt{\delta}\mathcal{N}^j_i(0,1), \quad X^j_0 = Z^j.$$

Euler scheme for numerical solutions of SDEs IV

The Euler scheme converges under the sufficient conditions for existence and uniqueness of the global solution of our SDE (Lipschitz continuity and linear growth). Its output has an order of convergence of $1/2$, meaning that if we denote the output of an Euler scheme with step Δt by $X_T^{\Delta t}$, and compare it with the real solution X_T, we have that there exists a positive real number δ_0 such that

$$E\{|X_T^{\Delta t} - X_T|\} \leq C(T)(\Delta t)^{1/2} \text{ for all } \Delta t \leq \delta_0$$

where $C(T) > 0$ is a constant (strong convergence of order $1/2$).

Euler scheme for geometric Brownian motion I

Example: simulate the geometric Brownian motion with Euler scheme; compare with the exact lognormal distribution.

$$dS_t = \mu\, S_t dt + \sigma\, S_t dW_t, \quad S_0 = s_0,$$

with s_0 a deterinistic constant becomes

$$S^j_{t_{i+1}} = S^j_{t_i} + \mu\, S^j_{t_i}\, \Delta t_i + \sigma S^j_{t_i} \sqrt{\delta} \mathcal{N}^j_i(0,1), \;\; S^j_0 = s^j_0.$$

We use the parameters: (naive code & vectorized code respectively)

$$S_0 = 100,\; \mu = 0.05,\; \sigma = 0.2,\; T = 1y,\; \Delta t = 1/200,\; N = 40000$$

$$S_0 = 100,\; \mu = 0.05,\; \sigma = 0.2,\; T = 1y,\; \Delta t = 1/2000,\; N = 100000$$

We show the distribution of the Euler simulated scheme versus the distribution of the one-shot simulated scheme, as we know the SDE

Euler scheme for geometric Brownian motion II

solution in one year and we can simulated it directly one-shot without time steps. This is what we do in the option risk measures examples.

Also, we show mean & standard deviation for the log-return. Recall

$$\ln S_t \sim \mathcal{N}\left(\ln S_0 + \mu t - \frac{1}{2}\sigma^2 t, \sigma^2 t\right), \text{ or (log return) } \ln \frac{S_t}{S_0} \sim \mathcal{N}\left(\mu t - \frac{1}{2}\sigma^2 t, \sigma^2 t\right)$$

log-return over $t = h$ is normal with the given mean and standard deviation. As mean and standard deviation characterize the normal distribution, a good check for the scheme is to visualize the histogram and to check the log returns mean and standard deviation.
We also check that log-returns skewness & excess kurtosis for log both Euler and one-shot schemes are zero.

Euler scheme for geometric Brownian motion III

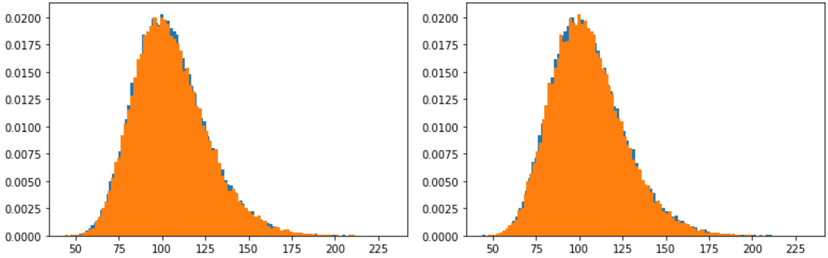


Figure: Naive scheme. Left: Histogram of S_{1y} for Euler (blue) and one-shot (orange); right: Histogram of S_{1y} for Euler (orange) and one-shot (blue)

Euler scheme for geometric Brownian motion IV

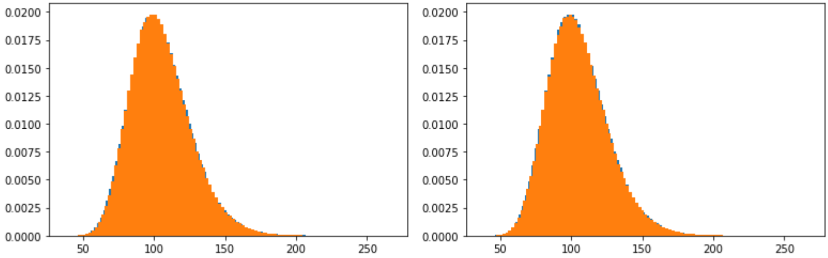


Figure: Vectorized scheme. Left: Histogram of S_{1y} for Euler (blue) and one-shot (orange); right: Histogram of S_{1y} for Euler (orange) and one-shot (blue)

Euler scheme for geometric Brownian motion V

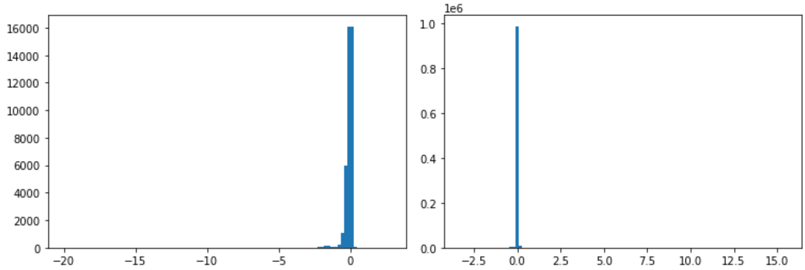


Figure: Left: Naive scheme. Histogram of difference between Euler Scheme and One Shot Scheme solutions at one year. Right: same with vectorized scheme.

Euler scheme for geometric Brownian motion VI

Results Euler Scheme:

Data set	mean log-return	stdev log-return	skewness log-return	kurtosis log-return
Theoretical	4.6352	0.2	0	0
One-shot simulation	4.6358	0.2004	0.0167	0.0182
Euler simulation	4.6351	0.1995	-0.0073	-0.0099
Vectorized:				
One-shot simulation	4.6352	0.2004	0.0008	0.0030
Euler simulation	4.6366	0.2006	0.0090	-0.0146

Table: Statistics of log stock $\log(S_{1y})$ from theory, Euler Scheme and One-shot simulation

Euler scheme for $dX_t = m\,dt + \sigma X_t dW_t$ I

As a second example, we simulate the SDE of Problem 1 of Mock exam 1.

$$dX_t = m\,dt + \sigma X_t dW_t,\ X_0 = x_0$$

where $m \in \mathbb{R}$, $\sigma > 0$ and $x_0 \in \mathbb{R}$ are deterministic.
We don't have a solution for this SDE, so we apply an Euler Scheme.
This reads

$$X^j_{t_{i+1}} = X^j_{t_i} + m\Delta t_i + \sigma X^j_{t_i}\sqrt{\delta}\mathcal{N}^j_i(0,1),\ \ X^j_0 = x_0.$$

We take as values $m = 1, \sigma = 0.4, x_0 = 0$. We take $n = 40000$ scenarios with time step $\Delta t = 1/200$ and final time $T = 2y$.
The histogram of the probability density function of the solution X_{2y} looks like this

Euler scheme for $dX_t = m\,dt + \sigma X_t dW_t$ II

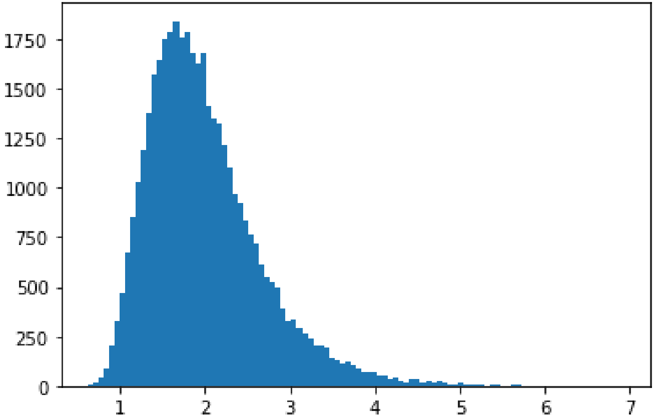


Figure: Histogram of S_{2y} for Euler scheme

Euler scheme for $dX_t = m\,dt + \sigma X_t dW_t$ III

	mean $E[S_{2y}]$	stdev $\text{STDEV}[S_{2y}]$	skewness	kurtosis
Theoretical	2	0.6805	NA	NA
Euler simulation	1.998	0.6712	1.2123	2.4994

Table: Statistics of stock S_{2y} from theory and Euler Scheme

Note that the simulated density is all in the positive axis. This might lead to think that the SDE solution is always positive. This is not true in general. We can see that for $m = 0$ the SDE becomes a Geometric Brownian motion with zero drift, so that the solution would be generally positive, but the initial condition $X_0 = 0$ gives us $X_t = 0$ for all t in this case: $X_t = X_0 \exp(-\sigma^2/2\, t + \sigma W_t)$ vanishes for all t for $X_0 = 0$. But let us try a negative drift. Set $m = -1$ and keep all other parameters equal.

Euler scheme for $dX_t = m\,dt + \sigma X_t dW_t$ IV

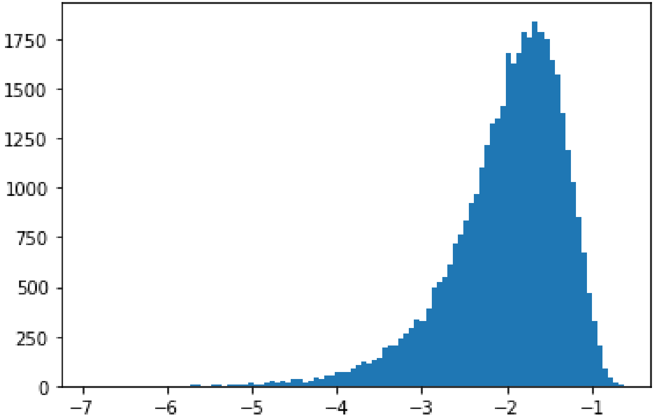


Figure: Histogram of S_{2y} for Euler scheme

Euler scheme for $dX_t = m\,dt + \sigma X_t dW_t$ V

We can see that the density now is entirely in the negative axis, so $X_t < 0$.

	mean $E[S_{2y}]$	stdev $\mathrm{STDEV}[S_{2y}]$	skewness	kurtosis
Theoretical	-2	0.6805	NA	NA
Euler simulation	-1.998	0.6712	-1.2123	2.4994

Table: Statistics of stock S_{2y} from theory and Euler Scheme

Results are the same, only signs change.

Euler scheme for $dX_t = m\,dt + \sigma X_t dW_t$ VI

So it seems that with $X_0 = 0$, the numerical results suggest that:

$$m > 0 \;\Rightarrow\; X_t > 0, \quad m = 0 \;\Rightarrow\; X_t = 0, \quad m < 0 \;\Rightarrow\; X_t < 0.$$

However, we have only numerical results suggesting this, we have not proven it.

It is possible to prove it with a comparison theorem, but this is beyond the scope of this course. The theorem states that, under some conditions that our SDE satisfies, the solution is increasing in m. Given that for $m = 0$ the solution is 0, our result follows.

See for example Theorem 1.1 in

Yamada, T. (1973). On a comparison theorem for solutions of stochastic differential equations and its applications. J. Math. Kyoto Univ. 13-3 (1973) pp 497-512.

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3}dW_t$ I

Consider the SDE

$$dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3}dW_t \tag{35}$$

with deterministic initial condition $x_0 = 0.5$, where $k > 0$ and $\sigma > 0$ are real constants and W is a standard Brownian motion.

a) Say whether the theorem with sufficient conditions for existence and uniqueness of a strong solution of SDEs given in these lecture notes applies to this SDE.

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3}dW_t$ II

b) Study this SDE numerically using an Euler scheme, without checking the conditions for the Euler scheme to converge. Simulate the SDE using an Euler scheme up to $T = 1$, one year, with a time step of 1 working day, $\Delta t = 1/400$ years (remember 1 working day = $1/250$ years). We assume $k = 1$, $\sigma = 0.05$, $x_0 = 0.5$. Plot the density of the solution at $T = 1$, namely the density of X_T using a histogram from the simulation, and comment on its shape. Is it a skewed or symmetric distribution? Are the tails fat? You may answer these questions also by calculating the sample skewness and sample excess kurtosis from the simulated solution.

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$ III

c) Solve the stochastic differential equation. Hint: to solve the SDE, use the transformation $Y_t = X_t^{1/3}$ and solve the SDE for Y, getting back X from Y. If you need to use Ito's formula, it might happen that the transformation and the SDE do not satisfy some assumptions required to apply Ito's formula. Comment on this, and then apply the formula anyway, formally, to find a formal solution.

d) If we change the initial condition to $x_0 = 0$, does the original SDE (35) admit the solution $X_t = 0$ for all t, on top of the solution you found in c)? Are the two solutions different? What does this say with regard to question a)?

e) Going back to the case $x_0 = 0.5$, once you have solved for X in c), see if you can simulate X "one shot" over one year by simulating Y, without any time steps in-between, just a single one-year step from time 0.

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3}dW_t$ IV

f) To make sure you have a consistent picture, compare the two densities of X_T, one-shot and one-day time steps, coming from e) and b), and comment on the differences if any. Plot the two histograms to visualize how close they are.

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3}dW_t$ V

Solutions:

a) The theorem does not apply because the cubic root is not a Lipschitz function. So we cannot ensure through that theorem that the SDE has a unique solution.

b) The Euler scheme for the proposed SDE is, denoting $X_i = X_{t_i}$, $\Delta t = t_i - t_{i-1}$ with $t_0 = 0,\ t_N = 1$,

$$X_{i+1} = X_i - 3kX_i\Delta t + 3X_i^{1/3}\sigma^2\Delta t + 3\sigma X_i^{2/3}(W_{i+1} - W_i), \quad X_0 = x_0.$$

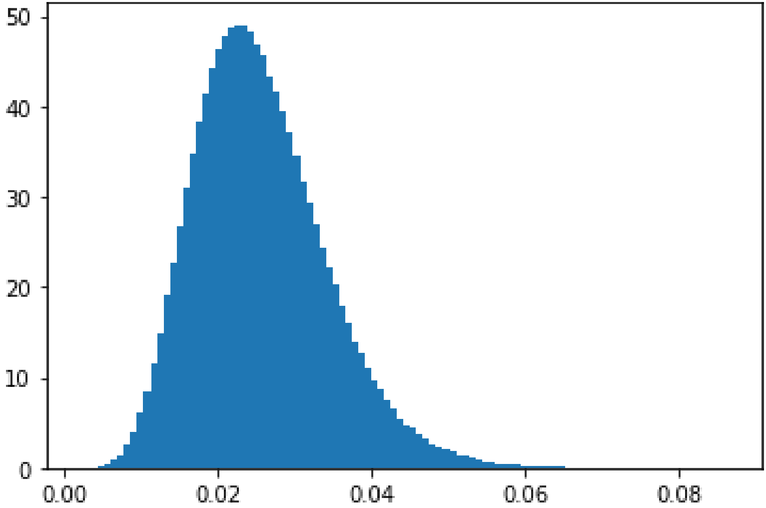
Remembering properties of brownian motion and denoting scenario j with an upper index j we have

$$X_{i+1}^j = X_i^j - 3kX_i^j\Delta t + 3(X_i^j)^{1/3}\sigma^2\Delta t + 3\sigma(X_i^j)^{2/3}\sqrt{\Delta t}N_i^j, \quad X_0 = x_0.$$

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3}dW_t$ VI

where N_i^j are all i.i.d. realizations from a standard normal.
We simulate the scheme using the vectorized version, starting from $x_0 = 0.5$ and using $n = 1000000$ scenarios. The Python code gives the following density.

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$ VII



Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$ VIII

The sample skewness and kurtosis are, using the relevant python routines,

Skew Euler: 0.6578196539807093
Kurtosis Euler: 0.6529821724338789

We see therefore that the Skew is positive, which is confirmed by a visual inspection of the density, that is skewed to the right. We see that there is excess Kurtosis, meaning the tails are fatter than the Gaussian. We will comment more on this after solving the next point.

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$ IX

c) To solve the SDE we do as follows.

$$Y_t = X_t^{1/3}$$

We use Ito's formula but with a warning, the transformation we are taking, the cubic root, is not twice continuously differentiable with respect to x in $x = 0$. So there can be issues, and $X = 0$ in particular might be a problem. We will still apply Ito's formula formally and see what happens.

$$dY = 1/3X^{-2/3}dX + 1/2\ 1/3\ (-2/3)X^{-5/3} 9\sigma^2 X^{4/3} dt$$

$$dY = 1/3X^{-2/3}(-3kXdt + 3X^{1/3}\sigma^2 dt + 3\sigma X^{2/3} dW_t) - \sigma^2 X^{-1/3} dt$$

$$dY = 1/3(-3kX^{1/3}dt + 3\sigma dW_t)$$

$$dY = -kX^{1/3}dt + \sigma dW_t$$

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$ X

so, concluding:

$$dY = -kYdt + \sigma dW_t$$

which is a linear SDE we know how to solve.

$$Y_t = y_0 e^{-kt} + \sigma \int_0^t e^{-k(t-u)} dW_u.$$

As $X_t = Y_t^3$ we get

$$X_t = \left(x_0^{1/3} e^{-kt} + \sigma \int_0^t e^{-k(t-u)} dW_u \right)^3.$$

So to simulate X_T one shot we need to simulate the quantity between round brackets for $t = T = 1$ year and then raise that to the power 3. The quantity between round brackets is Normal, as it's a constant plus

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3}dW_t$ XI

a Wiener integral, and by Ito's isometry we can find its variance at time t as

$$\sigma^2/(2k)(1 - e^{-2kt})$$

and

$$X_t \sim \left(\mathcal{N}(x_0^{1/3}e^{-kt}, \sigma^2/(2k)(1 - e^{-2kt}))\right)^3.$$

d) If we set $x_0 = 0$, we see that the right hand side of the SDE (35) computed at $X = 0$ is 0. This means we have $dX_t = 0$ and so the solution does not change. In the case $X_0 = 0$, the solution $X_t = 0$ for all t is not found by the method in point c). Indeed, in that method we used Ito's formula without the twice differentiability assumptions for drift and diffusion coefficient being satisfied at $x = 0$. Also, we couldn't establish

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$ XII

that the equation has a unique solution. Thus, for $x_0 = 0$, we have at least two solutions, the solution $X_t = 0$ and the solution found in c):

$$X_t = 0, \quad X_t = \left(\sigma \int_0^t e^{-k(t-u)} dW_u\right)^3$$

e) It is easy to simulate one shot over one year from the solution in c):

$$X_t \sim \left(\mathcal{N}(x_0^{1/3} e^{-kt}, \sigma^2/(2k)(1 - e^{-2kt}))\right)^3.$$

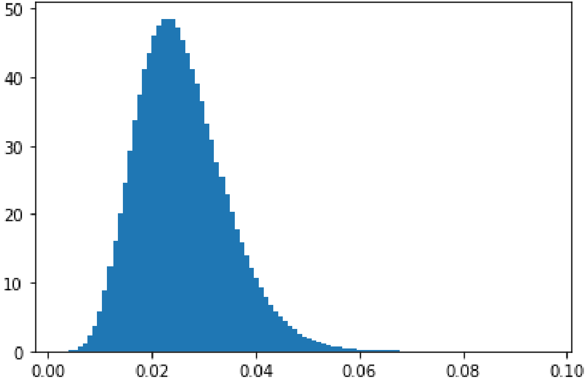
Or

$$X_t^j = \left(x_0^{1/3} e^{-kt} + \sqrt{\sigma^2/(2k)(1 - e^{-2kt})}\,\mathcal{N}^j\right)^3$$

where $\mathcal{N}^j$ are i.i.d. standard normals.

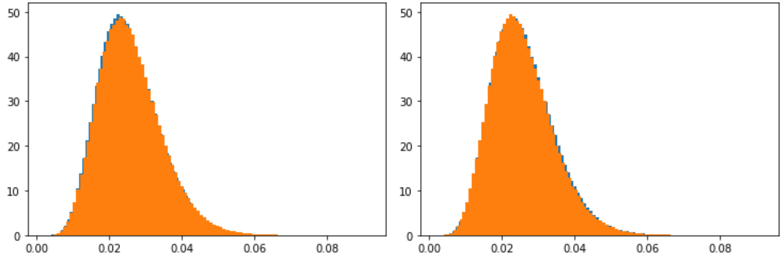
Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$ XIII

f) We run the simulation in e) one shot and we get the density



Visual inspection tells us that the two densities are very close. We can also plot one on top of the other

Study of $dX_t = (-3kX_t + 3X_t^{1/3}\sigma^2)dt + 3\sigma X_t^{2/3} dW_t$ XIV



Problems with solutions I

We now present a few solved problems similar to those to be expected at the exam.

Mock Exam 1 I

Problem 1. Consider the SDE $dX_t = m\,dt + \sigma X_t dW_t,\ X_0 = x_0$ where $m \in \mathbb{R}$, $\sigma > 0$ and $x_0 \in \mathbb{R}$ deterministic.
a) Prove that the SDE admits a unique solution.
b) For the solution X find $E(X_T)$ & $\mathrm{Var}(X_T)$ for any given $T > 0$.
c) If $Y = X_t^3$, find the SDE satisfied by Y. Write down the drift, diffusion coefficient and initial condition for the Y SDE.

Mock Exam 1 II

Solutions. a) A sufficient condition for existence and uniqueness of a strong solution is that we have two conditions regarding Lipschitz continuity and linear growth. We know from the theory that for the SDE $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = Z$ with Z independent of $\sigma(\{W_t,\ t \leq T)\}$ and $\mathbb{E}[Z^2] < +\infty$, and with $\mu : [0, T] \times \mathbb{R} \to \mathbb{R}$ (the drift) and $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ (the diffusion coefficient) being measurable, if we have global Lipschitz continuity

$$|\mu(t,x)-\mu(t,y)|+|\sigma(t,x)-\sigma(t,y)| \leq K|x-y| \text{ for all } t \in [0,T] \text{ and all } x \in \mathbb{R}$$

and linear growth

$$|\mu(t,x)| + |\sigma(t,x)| \leq K'(1 + |x|)| \ \text{ for all } t \in [0,T] \text{ and all } x \in \mathbb{R}$$

for two constants K, K', then our SDE has a unique global solution X_t.

Mock Exam 1 III

Let's check our conditions. In our case $\mu(t,x) = m$, $\sigma(t,x) = \sigma x$ are both measurable functions (being constant and linear respectively), $X_0 = Z = x_0$ is deterministic, and thus trivially independent of W and with finite mean square $E(Z^2) = x_0^2 < \infty$, and we can see that

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| = 0 + \sigma|x - y| \text{for all } t \text{ and } x$$

so that the Lipschitz condition is satisfied by taking $K = \sigma$. Given that the SDE is autonomous (no explicit dependence on t in the coefficients), you may quote the result given in the lectures that "global Lipschitz continuity implies linear growth" and you are done. This is ok for the exam. In case you wish to prove linear growth directly,

$$|\mu(t,x)| + |\sigma(t,x)| = |m| + \sigma|x| \le \max(|m|,\sigma)(1 + |x|) \text{ for all } t \text{ and } x$$

shows that also the linear growth condition is satisfied with $K' = \max(|m|,\sigma)$. Hence our SDE admits a unique global solution.

Mock Exam 1 IV

b) We integrate both sides of the SDE bewtween 0 and T.

$$\int_0^T dX_t = \int_0^T m dt + \int_0^T \sigma X_t dW_t.$$

We get

$$X_T - x_0 = m\,T + \sigma \int_0^T \sigma X_t dW_t.$$

Now we take expected value on both sides, remembering that the Ito integral has zero mean.

$$E[X_T] - x_0 = m\,T + \sigma 0 \implies E[X_T] = x_0 + mT\ .$$

Mock Exam 1 V

For the variance, we know that $\mathrm{Var}(X_t) = E(X_t^2) - E(X_t)^2$. We are missing $E(X_t^2)$. We use Ito's formula to get $d(X_t^2)$. Set $\phi(t,x) = x^2$ and compute

$$d(X_t^2) = \mathrm{d}\phi(t, X_t) = \frac{\partial \phi}{\partial t}\,\mathrm{d}t + \frac{\partial \phi}{\partial X}\,\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2 \phi}{\partial X^2}\mathrm{d}X_t\,\mathrm{d}X_t$$

$$= 0\;dt + 2X_t dX_t + \frac{1}{2}2dX_t dX_t = 2X_t(m\;dt + \sigma X_t dW_t) + \sigma^2 X_t^2 dt$$

$$\text{where}\;\; dXdX = (m\;dt + \sigma X dW)(m\;dt + \sigma X dW) = \sigma^2 X_t^2 dt$$

(recall $dt\;dt = 0$, $dt\;dW = 0$, $dW\;dW = dt$). We thus have

$$d(X_t^2) = (2mX_t + \sigma^2 X_t^2)dt + 2\sigma X_t^2 dW_t.$$

Mock Exam 1 VI

Integrate both sides between 0 and T:

$$\int_0^T d(X_t^2) = \int_0^T (2mX_t + \sigma^2 X_t^2)dt + \int_0^T 2\sigma X_t^2 dW_t$$

to get

$$X_T^2 - x_0^2 = \int_0^T (2mX_t + \sigma^2 X_t^2)dt + \int_0^T 2\sigma X_t^2 dW_t.$$

Take expected value on both sides, recalling that the Ito integral has zero mean:

$$E[X_T^2] - x_0^2 = \int_0^T (2mE[X_t] + \sigma^2 E[X_t^2])dt + 0.$$

Mock Exam 1 VII

Here we used a Fubini type theorem, moving the expected value inside the time integral. We know $E(X_t) = x_0 + m\,t$ from our previous calculation, so that

$$E[X_T^2] - x_0^2 = \int_0^T (2m[x_0 + m\,t] + \sigma^2 E[X_t^2])dt.$$

Set $m_2(t) = E[X_t^2]$ for all t, so that the last integral equation reads

$$m_2(T) - x_0^2 = \int_0^T (2m[x_0 + m\,t] + \sigma^2 m_2(t))dt.$$

Now, to proceed further, we differentiate both sides with respect to T. We get

$$\frac{d}{dT} m_2(T) = 2m[x_0 + m\,T] + \sigma^2 m_2(T).$$

Mock Exam 1 VIII

This is a linear-affine ODE that we have seen in the introduction. From the standard solution, keeping in mind that $m_2(0) = x_0^2$, we get

$$m_2(T) = e^{\sigma^2 T}\left[\int_0^T e^{-\sigma^2 u} 2m(x_0 + mu)du + x_0^2\right].$$

$$= e^{\sigma^2 T}\left[\frac{e^{-\sigma^2 T}(-\sigma^2(2mx_0 + 2m^2 T) - 2m^2)}{\sigma^4} + \frac{\sigma^2 2mx_0 + 2m^2}{\sigma^4} + x_0^2\right]$$

$$= \frac{2m}{\sigma^4}(m + x_0\sigma^2)(e^{\sigma^2 T} - 1) - \frac{2m^2}{\sigma^2}T + e^{\sigma^2 T}x_0^2$$

where we used integration by parts. We can now calculate $\mathrm{Var}(X_T) = E(X_T^2) - E(X_T)^2 = m_2(T) - (x_0 + mT)^2$. Complete the calculations.

Mock Exam 1 IX

c) We now apply Ito's formula with $\varphi(t,x) = x^3$. We have $\frac{\partial\varphi}{\partial t} = 0$, $\frac{\partial\varphi}{\partial x} = 3x^2$, $\frac{\partial^2\varphi}{\partial x^2} = 6x$. We have

$$dY_t = \mathrm{d}\varphi(t, X_t) = \frac{\partial\varphi}{\partial t}\,\mathrm{d}t + \frac{\partial\varphi}{\partial X}\,\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2\varphi}{\partial X^2}\mathrm{d}X_t\,\mathrm{d}X_t$$

$$= 0\ dt + 3X_t^2 dX_t + \frac{1}{2}6X_t dXdX = \ldots$$

Now we know from the previous point that $dXdX = \sigma^2 X_t^2 dt$ so that

$$\ldots = 3mX_t^2\ dt + 3\sigma X_t^3 dW_t + 3\sigma^2 X_t^3\ dt$$

$$dY_t = 3X_t^2(m + \sigma^2 X_t)dt + 3\sigma X_t^3 dW_t.$$

Recalling that $Y = X^3,\ X = Y^{1/3}$, we can write the SDE in Y as

$$dY = 3Y_t^{2/3}(m + \sigma^2 Y_t^{1/3})dt + 3\sigma Y_t dW_t,\ \ Y_0 = x_0^3.$$

Mock Exam 1 X

The drift for Y's SDE is $\mu(t, y) = 3y^{2/3}(m + \sigma^2 y^{1/3})$ while the diffusion coefficient is $\sigma(t, y) = 3\sigma y$. We also have the initial condition $Y_0 = X_0^3 = x_0^3$ which is deterministic.

Mock Exam 1 I

Problem 2. Option pricing in Black Scholes (long Straddle).
Consider a stock market where the stock price S follows the dynamics $dS_t = rS_t dt + \sigma S_t dW_t$ under the risk neutral measure Q, with initial stock price $s_0 > 0$, deterministic. The risk free rate r is a non-negative deterministic constant. Consider a straddle payoff on S, with maturity T and strike $K = S_0 e^{rT}$, namely

$$Y = (S_T - K)^+ + (K - S_T)^+.$$

a) Make a plot of the payoff as a function of S_T. Explain what kind of investor would buy this payoff. What would the investor rely on, to make money, when buying this product?
b) Calculate the price of the straddle at time 0. You can use the formula for a call option in Black-Scholes without deriving it, if you remember it, but derive the put from the call using parity.

Mock Exam 1 II

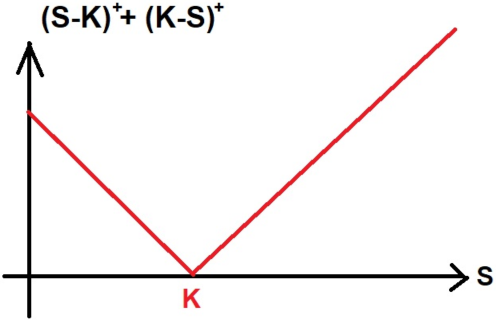
c) How does the straddle price above change with σ? In particular, calculate the Vega of the straddle, namely if V is the straddle price, compute $Vega = \frac{\partial V}{\partial \sigma}$ and discuss your findings.

Mock Exam 1 III

Problem 2 Solutions.

a) Recalling how the call and put payoffs work, the straddle payoff $Y = (S_T - K)^+ + (K - S_T)^+$ is equal to $S_T - K$ if $S_T \geq K$ and to $K - S_T$ if $S_T < K$. It is immediate to see that this is the same as $Y = |S_T - K|$. The plot is therefore an absolute value plot centered in K.

Mock Exam 1 IV



The payoff becomes larger the farther away S_T moves from $S_0 e^{rT}$. This means that a straddle payoff makes money if the stock S moves away from S_0 a lot. An investor will therefore buy a straddle if she expects the stock price to be very volatile. In this case she will gain the

Mock Exam 1 V

large absolute difference between the stock at maturity and K. It doesn't matter if the stock moves up or down, the important point for the straddle investor is that it moves. On the contrary, if the stock moves very little the straddle makes very little money.

An investor who expects the market to move very little should sell rather than buy a straddle. This way the investor cashes in the initial straddle price from the client but will have to pay a very little payoff at maturity if the stock moves very little. In the limit case where the stock at maturity is $S_T = K = S_0 e^{rT}$, the straddle payoff $|S_T - K|$ will be worth 0, so the investor selling the straddle will cash in a large price at time 0 and will pay nothing at maturity T.

Mock Exam 1 VI

b) As the straddle payoff is call plus put payoffs, the price $V_{Str}(0)$ at time 0 will be the price of a call plus the price of a put. Indeed,

$$V_{Str}(0) = E^Q[e^{-rT}((S_T - K)^+ + (K - S_T)^+)] =$$

$$= E^Q[e^{-rT}(S_T - K)^+] + E^Q[e^{-rT}(K - S_T)^+] = V^{call}(0) + V^{put}(0).$$

This is completely general and holds whichever model we use for S. In point b) we are given the Black Scholes model, so we can compute the price as a sum of call and put in Black Scholes. Recall that (and if you don't recall it then derive it as we have done in the lectures)

$$V_{BS}^{CALL}(0, S_0, K, T, \sigma, r) = s_0\Phi(d_1) - Ke^{-rT}\Phi(d_2)$$

where Φ is the CDF of a standard normal and where

$$d_1 = \frac{\ln(s_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \; d_2 = d_1 - \sigma\sqrt{T}$$

Mock Exam 1 VII

Now we can derive the price of a put option by put-call parity. Write the argument and the derivation here, as it has been done in the lecture, using the put call parity and the price of a forward contract. We get

$$V_{BS}^{PUT}(0, S_0, K, T, \sigma, r) = Ke^{-rT}\Phi(-d_2) - s_0\Phi(-d_1).$$

We can now calculate the straddle price as

$$V_{BS}^{STR}(0) = V_{BS}^{CALL}(0) + V_{BS}^{PUT}(0) =$$

$$= s_0\Phi(d_1) - Ke^{-rT}\Phi(d_2) + Ke^{-rT}\Phi(-d_2) - s_0\Phi(-d_1) =$$

$$= s_0(2\Phi(d_1) - 1) - Ke^{-rT}(2\Phi(d_2) - 1)$$

where we used the fact that $\Phi(-x) = 1 - \Phi(x)$.

Mock Exam 1 VIII

Substitute $K = s_0 e^{rT}$, also in $d_{1,2}$ to obtain the final price

$$V_{BS}^{STR}(0) = 2s_0 \left(\Phi\left(\frac{\sigma\sqrt{T}}{2}\right) - \Phi\left(\frac{-\sigma\sqrt{T}}{2}\right) \right)$$

c) To explain how the price of the straddle changes with σ let's first compute the straddle Vega, as requested.
Important: we do the calculation by first substituting $K = s_0 e^{rT}$ into the straddle price, and then differentiating with respect to σ. This is much faster and is not a problem, because the strike does not depend on σ, so we are not inadvertently differentiating twice. In other terms, we do not have the problem we have when computing the delta of an at-the-money option, where we need to hold a general strike K during differentiation and substitute $K = s_0$ only after the differentiation has taken place. Here we can susbtitute before, because while $K = s_0 e^{rT}$

Mock Exam 1 IX

would affect a differentiation with respect to s_0, as it contains s_0, it does not affect a differentiation with respect to σ, as it does not contain σ. Thus

$$Vega = \frac{\partial V_{BS}^{STR}(0)}{\partial \sigma} = \frac{\partial}{\partial \sigma}\left(2s_0\left(\Phi\left(\frac{\sigma\sqrt{T}}{2}\right) - \Phi\left(\frac{-\sigma\sqrt{T}}{2}\right)\right)\right)$$

$$= 2s_0\left(\frac{\partial}{\partial \sigma}\Phi\left(\frac{\sigma\sqrt{T}}{2}\right) - \frac{\partial}{\partial \sigma}\Phi\left(\frac{-\sigma\sqrt{T}}{2}\right)\right)$$

$$= 2s_0\left(\phi\left(\frac{\sigma\sqrt{T}}{2}\right)\frac{\sqrt{T}}{2} + \phi\left(\frac{-\sigma\sqrt{T}}{2}\right)\frac{\sqrt{T}}{2}\right)$$

$$= 2s_0\frac{\sqrt{T}}{2}2\phi\left(\frac{\sigma\sqrt{T}}{2}\right) = 2s_0\sqrt{T}\phi\left(\frac{\sigma\sqrt{T}}{2}\right)$$

Mock Exam 1 X

where $\phi = \Phi'$ is the density of the standard Gaussian distribution and where we used the fact that ϕ is symmetric with respect to the origin, namely $\phi(-x) = \phi(x)$, so $\phi(x) + \phi(-x) = 2\phi(x)$.

Now note that ϕ is always positive, being the normal probability density function. It follows that, since also σ, T and s_0 are positive, vega is always positive.

A positive vega means $\frac{\partial V_{BS}^{Str}}{\partial \sigma} > 0$, and if a function has a positive derivative with respect to a variable, it is increasing with respect to that variable. It follows that with positive vega, V_{BS}^{Str} is increasing in the volatility σ.

Mock Exam 1 XI

We conclude that the straddle price will increase with the stock volatility σ. The larger the volatility, the larger the straddle price. This is consistent with our intuition on the payoff given in point a) above.

An alternative faster solution that would have been accepted is if one remembered the final formula for Vega for a Call option, $Vega_{BS}^{Call} = s_0\phi(d_1)\sqrt{T}$. By put-call parity, the Vega of a put is the same since

$$V^{Put} = V^{call} - V^{forward} = V^{call} - (s_0 - Ke^{-rT})$$

and

$$\frac{\partial V^{Put}}{\partial\sigma} = \frac{\partial V^{Call}}{\partial\sigma} - \frac{\partial(s_0 - Ke^{-rT})}{\partial\sigma} = \frac{\partial V^{Call}}{\partial\sigma} - 0 = \frac{\partial V^{Call}}{\partial\sigma}$$

Mock Exam 1 XII

So Vega call is the same as Vega put and the solution would have been immediately, substituting the strike $K = s_0 e^{rT}$:

$$Vega = Vega^{Call} + Vega^{Put} = 2\, Vega^{call} = 2 s_0 \phi(d_1) \sqrt{T} = 2 s_0 \sqrt{T} \phi \left(\frac{\sigma \sqrt{T}}{2} \right)$$

‘ In all cases we see that Vega is positive and hence the straddle price will be increasing with the volatility, as mentioned earlier.

Mock Exam 1 I

Problem 3. Option pricing Displaced Diffusion (long Straddle).
Consider a stock market where the stock price S follows the displaced diffusion (DD) dynamics

$$dS_t = rS_t dt + \sigma(S_t - \alpha e^{rt})dW_t, \ \ s_0,$$

where α is a deterministic shift, under the risk neutral measure Q, with initial stock price $s_0 > 0$, deterministic. The risk free rate r is a non-negative deterministic constant. Consider a straddle payoff on S, with maturity T and strike $K = S_0 e^{rT}$, namely

$$Y = (S_T - K)^+ + (K - S_T)^+.$$

a) Compute the straddle price at time 0 in the DD model.
b) Check that in the limit case $\alpha = 0$ you get back the Black Scholes price of a straddle.

Mock Exam 1 II

c) Can you comment if, ceteris paribus, adding the shift α increases or decreases the Straddle price compared to the pure Black Scholes case? More generally, what is the impact of α on the straddle price? How is the price sensitive to α?

Mock Exam 1 III

Problem 3 Solutions.
a) For the displaced diffusion (DD) model, it is convenient to write it as in the lectures. We write

$$S_t = \alpha e^{r\ t} + X_t, \quad dX_t = rX_t dt + \sigma X_t dW_t, \quad X_0 = s_0 - \alpha$$

where the dynamics of X is under Q (W is a Brownian motion under Q). It is immediate to check that with this definition we have

$$dS_t = rS_t dt + \sigma(S_t - \alpha e^{rt})dW_t, \quad s_0$$

as given in the problem.
We can now compute the straddle price in the DD model as

$$V_{DD}^{Str} = E^Q[e^{-rT}((S_T - K)^+ + (K - S_T)^+)] =$$

Mock Exam 1 IV

$$= E^Q[e^{-rT}((X_T + \alpha e^{r\ T} - K)^+ + (K - X_T - \alpha e^{r\ T})^+)] =$$
$$= E^Q[e^{-rT}(X_T + \alpha e^{r\ T} - K)^+] + E^Q[e^{-rT}(K - X_T - \alpha e^{r\ T})^+] =$$
$$= E^Q[e^{-rT}(X_T - K')^+] + E^Q[e^{-rT}(K' - X_T)^+], \ \ K' = K - \alpha e^{r\ T}.$$

The first expectation is a call option for the stock X with strike K'. Given that X follows a Black Scholes model with volatility σ and initial stock price $x_0 = s_0 - \alpha$, we can use a Black Scholes call option formula for this. We have $E^Q[e^{-rT}(X_T - K')^+] =$

$$= x_0\Phi(d_1') - K'e^{-rT}\Phi(d_2') = (s_0 - \alpha)\Phi(d_1') - (K - \alpha e^{r\ T})e^{-rT}\Phi(d_2'),$$

where

$$d_{1,2}' = \frac{\ln\frac{x_0}{K'} + (r \pm \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = \frac{\ln\frac{s_0-\alpha}{K-\alpha e^{r\ T}} + (r \pm \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

Mock Exam 1 V

Recall that the put option is obtained by put-call parity (which holds for all models, as it is model independent) as

$$V_{DD}^{PUT} = V_{DD}^{CALL} - (S_0 - Ke^{-rT})$$

so that

$$V_{DD}^{STR} = V_{DD}^{CALL} + V_{DD}^{PUT} = V_{DD}^{CALL} + V_{DD}^{CALL} - (S_0 - Ke^{-rT}) =$$

$$= 2V_{DD}^{CALL} - (S_0 - Ke^{-rT}),$$

so that the straddle price at time 0 is

$$V_{DD}^{STR} = 2(s_0 - \alpha)\Phi(d_1') - 2(K - \alpha e^{r\ T})e^{-rT}\Phi(d_2') - s_0 + Ke^{-rT}$$

Mock Exam 1 VI

Substituting $K = s_0 e^{rT}$, also in $d'_{1,2}$, we get

$$V_{DD}^{STR} = 2(s_0 - \alpha)\Phi(\bar{d}'_1) - 2(s_0 e^{rT} - \alpha e^{r\ T})e^{-rT}\Phi(\bar{d}'_2) - s_0 + s_0 e^{rT} e^{-rT}$$

$$= 2(s_0 - \alpha)(\Phi(\bar{d}'_1) - \Phi(\bar{d}'_2))$$

where

$$\bar{d}'_{1,2} = \pm\frac{1}{2}\sigma\sqrt{T}$$

b) We first compute the straddle price in a Black Scholes model. (This has been done in Problem 2 above, follow the same steps). We obtain

$$V_{BS}^{STR} = 2s_0(\Phi(\bar{d}'_1) - \Phi(\bar{d}'_2)) = 2s_0(\Phi(\bar{d}'_1) - \Phi(\bar{d}'_2))$$

$$\bar{d}'_{1,2} = \pm\frac{1}{2}\sigma\sqrt{T}.$$

Mock Exam 1 VII

Consider now, from point a) above,

$$V_{DD}^{STR} = 2(s_0 - \alpha)(\Phi(\bar{d}_1') - \Phi(\bar{d}_2')) = \dots$$

by substituting $\alpha = 0$ we obtain

$$\dots = 2s_0(\Phi(\bar{d}_1') - \Phi(\bar{d}_2'))$$

which is exactly the Black Scholes straddle price $V_{BS}^{STR}(0)$ above.
c) Compare the two prices

$$V_{BS}^{STR} = 2s_0\left(\Phi\left(\frac{\sigma\sqrt{T}}{2}\right) - \Phi\left(\frac{-\sigma\sqrt{T}}{2}\right)\right) =: 2s_0 A,$$

$$V_{DD}^{STR} = 2(s_0 - \alpha)A, \quad A = \left(\Phi\left(\frac{\sigma\sqrt{T}}{2}\right) - \Phi\left(\frac{-\sigma\sqrt{T}}{2}\right)\right)$$

Mock Exam 1 VIII

We see that the price comparison between the two models depends on the sign of α. Given that

$$A = \Phi\left(\frac{\sigma\sqrt{T}}{2}\right) - \Phi\left(\frac{-\sigma\sqrt{T}}{2}\right) > 0,$$

as Φ is increasing, we obtain that

$$V_{DD}^{STR} > V_{BS}^{STR} \iff 2(s_0 - \alpha)A > 2s_0A \iff \alpha < 0,$$

where we divided both sides for the positive quantity A. So with negative shift $\alpha < 0$ the straddle price will be larger than the basic Black Scholes price. With positive α, the straddle price will be smaller than the Black Scholes case. For $\alpha = 0$ we recover the Black Scholes case.

Mock Exam 1 IX

We now analyze the impact of the shift α on the price. From the formula for V_{DD}^{STR} we see that the price is linear in α. We can easily compute

$$\frac{\partial V_{DD}^{STR}}{\partial \alpha} = \frac{\partial 2(s_0 - \alpha)A}{\partial \alpha} = -2A < 0$$

from which we see that the straddle price is decreasing with respect to the shift α. Thus, we conclude that increasing the shift α will decrease the straddle price, and decreasing the shift will increase the straddle price.

Mock Exam 1 I

Problem 4: Risk Measures.
Consider the dynamics of an equity asset price S in the Black and Scholes model, under both probability measures P (the Physical or Historical measure) and Q (the risk neutral measure).

a) Define Value at Risk (VaR) for a time horizon H with confidence level α for a general portfolio.

b) Compute VaR for horizon H and confidence level α for a portfolio with N units of equity, where the equity price follows the Black Scholes process above.

c) Explain at least one drawback of VaR as a risk measure

d) Is the equity dynamics you used for VaR the same you would have used to price an equity call option in Black Scholes?

Mock Exam 1 II

Problem 4: Solutions.

a)

VaR is related to the potential loss on our portfolio over the time horizon H. Define this loss L_H as the difference between the value of the portfolio today (time 0) and in the future H.

$$L_H = \text{Portfolio}_0 - \text{Portfolio}_H.$$

VaR with horizon H and confidence level α is defined as that number $q = q_{H,\alpha}$ such that

$$P[L_H < q] = \alpha$$

so that our loss at time H is smaller than q with P-probability α. In other terms, it is that level of loss over a time H that we will not exceed with probability α. It is the α P-percentile of the loss distribution over H.

Mock Exam 1 III

b)

In Black Scholes the equity process follows the dynamics

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where μ, σ are positive constants and W is a brownian motion under the physical measure P.

We know that S_H can be written as

$$S_H = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) H + \sigma W_H\right\}, \tag{36}$$

and recalling the distribution of W_H,

$$S_H = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) H + \sqrt{H}\sigma \mathcal{N}(0,1)\right\} \tag{37}$$

Mock Exam 1 IV

so that in our case $L_H = N(S_0 - S_H)$, namely

$$L_H = NS_0 \left(1 - \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) H + \sqrt{H}\sigma\mathcal{N}(0,1)\right\}\right)$$

Hence

$$\alpha = P[L_H < q] = P\left[\left(1 - \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) H + \sqrt{H}\sigma\mathcal{N}(0,1)\right\}\right) < \frac{q}{NS_0}\right]$$

$$= P\left[\left(\mu - \frac{1}{2}\sigma^2\right) T + \sqrt{H}\sigma\mathcal{N}(0,1) > \ln\left(1 - \frac{q}{NS_0}\right)\right]$$

$$= P\left[\mathcal{N}(0,1) > \frac{\ln\left(1 - \frac{q}{NS_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right) H}{\sqrt{H}\sigma}\right] =$$

Mock Exam 1 V

$$= 1 - \Phi\left(\frac{\ln\left(1 - \frac{q}{NS_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)H}{\sqrt{H}\sigma}\right)$$

$$= \Phi\left(-\frac{\ln\left(1 - \frac{q}{NS_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)T}{\sqrt{H}\sigma}\right)$$

So we have obtained

$$\alpha = \Phi\left(-\frac{\ln\left(1 - \frac{q}{NS_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)H}{\sqrt{H}\sigma}\right)$$

or

$$\Phi^{-1}(\alpha) = -\frac{\ln\left(1 - \frac{q}{NS_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)H}{\sqrt{H}\sigma}$$

Mock Exam 1 VI

and therefore

$$\exp\left(-\sqrt{H}\sigma\Phi^{-1}(\alpha)+\left(\mu-\frac{1}{2}\sigma^2\right)H\right)=\left(1-\frac{q}{NS_0}\right)$$

$$q=NS_0\left[1-\exp\left(-\sqrt{H}\sigma\Phi^{-1}(\alpha)+\left(\mu-\frac{1}{2}\sigma^2\right)H\right)\right]$$

Optional (not required as an answer to this question in the exam). Some students may be confused by the lack of a maturity T in this example in the stock position, as well as the lack of any Q expectation to evaluate the loss. After all, in all the numerical VaR examples we always had a risk horizon H for a portfolio of options with maturity $T > H$. But what happens here, with a pure stock position?

Mock Exam 1 VII

The stock is not a derivative, it is a fundamental asset of our economy, and as such it doesn't need a risk neutral expectation to compute its value at any time. The value at any time H is simply S_H. This is why the loss in this problem is immediately given by $L_H = N(S_0 - S_H)$ without any need to take risk neutral expectations.
But suppose now we introduce a maturity for the stock position, for the sake of the argument. Suppose we are in a position where the value of the amount N of stock S will be paid only at maturity $T > H$, and we wish to calculate the VaR at the risk horizon H. In this case then we calculate the loss as usual as $Loss_H = Portfolio_0 - Portfolio_H$ or

$$Loss_H = E_0^Q[e^{-rT} N S_T] - E_H^Q[e^{-r(T-H)} N S_T] =$$

$$= N E_0^Q[e^{-rT} S_T] - N E_H^Q[e^{-r(T-H)} S_T] = N B_0 E_0^Q[S_T/B_T] - N B_H E_H^Q[S_T/B_T] =$$

Mock Exam 1 VIII

where we have used the definition $B_0 = 1$ and $B_t = e^{rt}$. Now, as we know that under Q the ratio S/B is a martingale, we have that

$$E_0^Q[S_T/B_T] = S_0/B_0, \;\; E_H^Q[S_T/B_T] = S_H/B_H.$$

Substituting in the Loss_H calculation we get

$$= NB_0 \; S_0/B_0 - NB_H \; S_H/B_H = N(S_0 - S_H)$$

namely, the loss for a stock with maturity T at a risk horizon H is the same as the loss of the stock computed directly, or the loss at a different maturity T_1.

So as you see the "stock maturity" is irrelevant, because the stock price at H is S_H and even if we were to consider it at a future maturity T, its present value or risk neutral price at the earlier time H would still be S_H. So the loss is still $S_0 - S_H$ for every possible maturity of the

Mock Exam 1 IX

stock, as it is the same as taking the stock as fundamental asset which has its direct price at time H as S_H, without a maturity, as it is a fundamental asset and not a derivative. This is similar to the reason why a forward contract of any maturity T is worth $S_0 - Ke^{-rT}$ at time 0, regardless of the maturity: as you see, no matter the maturity T of the forward, the stock value is always S_0 at time 0, it's only the K part that changes with T through discounting.

Mock Exam 1 X

c) VaR is not subadditive, hence it does not recognize the benefit of diversification. Also, VaR ignores the structure of the loss distribution after the percentile. So if 99% VaR is 10 billions, we can have the remaining 1% loss concentrated

(i) either on 10.1 billions,

(ii) or on 10 trillions,

as two stylized cases, without VaR being able to tell us anything on whether we are in case (i) or (ii).

d) No the dynamics is not the same, to price an option we need to use the risk neutral dynamics, where the drift parameter μ of S is replaced by the risk free rate r of the bank account.

Mock Exam 2 I

Problem 1. Consider the Ito SDE

$$dX_t = m(X_t - a)dt + \sigma(X_t - a)dW_t, \quad x_0$$

where a and m are deterministic constants, $\sigma > 0$ and x_0 is deterministic. We assume $x_0 > a$.
a) Does this equation admit a unique global solution?
b) Write the equation in Stratonovich form
c) If the Equation admits solutions, find a solution. Extra points if you give two possible ways to get a solution.
d) Compute the following probability for the solution you found in point c): $P\{X_t > x_0\}$.

Mock Exam 2 II

Solutions. a) We can check the sufficient conditions for global existence and uniqueness given by measurable drift and diffusion coefficient, finite second moment for the initial condition, and Lipschitz continuity and linear growth.
First of all the equation has drift $\mu(t,x) = m(x-a)$ and diffusion coefficent $\sigma(t,x) = \sigma(x-a)$. Both are trivially measurable functions, as they are linear.
The initial condition is deterministic, so that the condition on the second moment $E[X_0^2] < \infty$ is trivially satisfied, given that $X_0^2 = x_0^2$ is a finite deterministic constant.
Next we check the Lipschitz condition and linear growth conditions:

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| = |mx - my| + |\sigma x - \sigma y| =$$

$$= |m||x-y| + \sigma|x-y| \leq K|x-y|$$

Mock Exam 2 III

for all x, y and all t provided that $K = 2\max(|m|, \sigma)$. Hence Lipschitz continuity holds. Given that the SDE is autonomous (no explicit dependence on t in the coefficients), you may quote the result given in the lectures that "global Lipschitz continuity implies linear growth" and you are done. This is ok for the exam. In case you wish to prove linear growth directly,

$$|\mu(t,x)| + |\sigma(t,x)| = |m(x-a)| + |\sigma(x-a)| =$$

$$\le |m+\sigma||x-a| \le |m+\sigma|(|a|+|x|) \le \max(|a|,1)|m+\sigma|(1+|x|) = K'(1+|x|)$$

for all t and x if we set $K' = \max(|a|, 1)|m + \sigma|$. So the linear growth condition is satisfied. It follows that the equation admits a unique global solution.

Mock Exam 2 IV

b) To write the SDE in stratonovich form we recall the transformation rule. The following two SDEs

$$dX_t = f(X_t)dt + v(X_t)dW_t \quad \rightarrow \quad dX_t = \tilde{f}(X_t)dt + v(X_t) \circ dW_t$$

$$\tilde{f} = f - \frac{1}{2} v \frac{\partial v}{\partial x}$$

have the same solution X. In our case $f(x) = m(x-a)$ and $v(x) = \sigma(x-a)$. We see immediately that $\frac{\partial v}{\partial x} = \sigma$ and

$$\tilde{f}(x) = m(x-a) - \frac{1}{2}\sigma(x-a)\sigma.$$

So the equivalent Stratonovich SDE with the same solution is

$$dX_t = (m - \frac{\sigma^2}{2})(X_t - a)dt + \sigma(X_t - a) \circ dW_t.$$

Mock Exam 2 V

c) We established in a) that the equation has a unique solution.
We can approach the solution in two ways.
First, we can try to solve the SDE using the Stratonovich form. For the Strat form, the formal rules of calculus hold. We present two possible approaches based on Stratonovich calculus, each considered a satisfactory solution to the exam.

(i) Separation of variables.

$$dX_t = (m - \frac{\sigma^2}{2})(X_t - a)dt + \sigma(X_t - a)\circ dW_t \;\Rightarrow\; \frac{dX}{X - a} = (m - \frac{\sigma^2}{2})dt + \sigma \circ dW_t$$

Mock Exam 2 VI

For this to work we need to assume $X_t \neq a$. We only know that $x_0 > a$. For other $t > 0$, we will check this a posteriori. Now, in the second SDE above we can integrate both sides immediately:

$$\int_{x_0}^{X_T} \frac{dX}{X - a} = \int_0^T \left((m - \frac{\sigma^2}{2})dt + \sigma \circ dW_t \right)$$

As $dX = d(X - a)$, we can rewrite the left side integral as

$$\int_{x_0}^{X_T} \frac{d(X - a)}{X - a} = \int_0^T \left((m - \frac{\sigma^2}{2})dt + \sigma \circ dW_t \right)$$

leading to

$$\ln(X - a)|_{x_0}^{X_T} = (m - \frac{\sigma^2}{2})T + \sigma W_T$$

Mock Exam 2 VII

or

$$\ln\left(\frac{X_T - a}{x_0 - a}\right) = (m - \frac{\sigma^2}{2})T + \sigma W_T.$$

Exponentiating both sides we get

$$\frac{X_T - a}{x_0 - a} = \exp\left[(m - \frac{\sigma^2}{2})T + \sigma W_T\right]$$

or

$$X_T = (x_0 - a)\exp\left[(m - \frac{\sigma^2}{2})T + \sigma W_T\right] + a$$

for all $T > 0$. We need to check that $X_T \neq a$ for all $T > 0$. Given the first term in the solution is strictly positive (because $x_0 > a$ by assumption and thus $x_0 - a > 0$, and the exponential is strictly positive too) and we add a to it, as a second term, the total will be larger than a

Mock Exam 2 VIII

as it is a plus a strictly positive number. Thus $X_T > a$ for all T and we are fine.

(ii) We can simplify the SDE by taking logs. Set $Y_t = \ln(X_t - a)$. This is only possible if $X_t > a$. *We will have to check this a-posteriori, once the solution is found.* Differentiating

$$dY_t = \frac{1}{X_t - a} \circ d(X_t - a) = \frac{1}{X_t - a} \circ dX_t =$$

$$= \frac{1}{X_t - a}((m - \frac{\sigma^2}{2})(X_t - a)dt + \sigma(X_t - a) \circ dW_t) = (m - \frac{\sigma^2}{2})dt + \sigma \circ dW_t.$$

So

$$dY_t = (m - \frac{\sigma^2}{2})dt + \sigma \circ dW_t.$$

Mock Exam 2 IX

This last equation is the same in Ito form, since the diffusion coefficient does not depend on Y. So we can write in Ito form

$$dY_t = (m - \frac{\sigma^2}{2})dt + \sigma dW_t.$$

This is an arithmetic Brownian motion and is easily integrated as

$$\int_0^T dY_t = \int_0^T (m - \frac{\sigma^2}{2})dt + \int_0^T \sigma dW_t$$

leading to

$$Y_T = Y_0 + (m - \frac{\sigma^2}{2})T + \sigma W_T.$$

Mock Exam 2 X

To go back to X_t, we recall that $Y_t = \ln(X_t - a)$ from which

$$X_t = e^{Y_t} + a = \exp(Y_0 + (m - \frac{\sigma^2}{2})T + \sigma W_T) + a$$

$$= \exp(Y_0)\exp((m - \frac{\sigma^2}{2})T + \sigma W_T) + a = (x_0 - a)\exp((m - \frac{\sigma^2}{2})T + \sigma W_T) + a$$

Now we can check that $X_t > a$ as requied to do the log transformation. We know that $x_0 > a$ so that $x_0 - a > 0$ and the exponential is also positive. Hence X_t is the sum of a positive term plus a, and as such is larger than a as required.

A further metod to derive the solution *without going the Stratonovich way* would be to note that we can set $Z_t = X_t - a$ and derive the Ito SDE for Z.

Mock Exam 2 XI

$$dZ_t = d(X_t - a) = dX_t - 0 = m(X_t - a)dt + \sigma(X_t - a)dW_t =$$
$$= mZ_t dt + \sigma Z_t dW_t.$$

Hence

$$dZ_t = mZ_t dt + \sigma Z_t dW_t, \quad Z_0 = x_0 - a$$

is a Geometric Brownian Motion and we know how to integrate it (take log, apply Ito's formula, write all the steps). We obtain

$$Z_t = Z_0 \exp((m - \sigma^2/2)t + \sigma W_t).$$

As $X_t = Z_t + a$, we obtain

$$X_t = a + (x_0 - a) \exp((m - \sigma^2/2)t + \sigma W_t).$$

Mock Exam 2 XII

We can see that this coincides with the solution obtained with the Stratonovich transformation.

d) Compute

$$P[X_t > x_0] = P[a + (x_0 - a)\exp((m - \sigma^2/2)t + \sigma W_t) > x_0] =$$

$$P[(x_0 - a)\exp((m - \sigma^2/2)t + \sigma W_t) > x_0 - a] = \dots$$

Now as $x_0 - a > 0$, we can divide both sides of the inequality inside the probability by $x_0 - a$ without changing the verse of the inequality. We get

$$\dots = P[\exp((m - \sigma^2/2)t + \sigma W_t) > 1] = P[(m - \sigma^2/2)t + \sigma W_t > 0] = \dots$$

Mock Exam 2 XIII

where we took log on both sides, which does not change the inequality as log is a strictly increasing function. Then

$$\ldots = P[\sigma W_t > -(m - \sigma^2/2)t/\sigma] = P[\sqrt{t}\mathcal{N}(0,1) > -(m - \sigma^2/2)t/\sigma] =$$

$$= P\left[\mathcal{N}(0,1) > -\frac{(m - \sigma^2/2)t}{\sigma\sqrt{t}}\right] = \ldots$$

where we used that Brownian motion $W_t \sim \mathcal{N}(0,t) \sim \sqrt{t}\mathcal{N}(0,1)$. Then

$$\ldots = 1 - P\left[\mathcal{N}(0,1) \leq -\frac{(m - \sigma^2/2)\sqrt{t}}{\sigma}\right] = \ldots$$

Mock Exam 2 XIV

where we used the fact that for any event A, $P[A^c] = 1 - P[A]$ where A^c is the complement of A.

$$\ldots = 1 - \Phi\left(-\frac{(m - \sigma^2/2)\sqrt{t}}{\sigma}\right) = \Phi\left(\frac{(m - \sigma^2/2)\sqrt{t}}{\sigma}\right)$$

where Φ is the CDF for the standard normal and we usd the property $\Phi(-x) = 1 - \Phi(x)$.

Mock Exam 2 I

Problem 2. We now consider a portfolio with a call option on a first stock $S^{(1)}$ with strike K_1 and a put option on a second stock $S^{(2)}$ with strike K_2, both options with maturity T. The final payoff of the portfolio is

$$Y = (S_T^{(1)} - K_1)^+ + (K_2 - S_T^{(2)})^+.$$

The risk-free rate is assumed to be a positive deterministic constant $r > 0$.

a) Assume both stocks follow a Black Scholes model. Specifically, the stocks dynamics under the measure P are

$$d\ S_t^{(1)} = \mu_1 S_t^{(1)}\ dt + \sigma_1 S^{(1)} dW_t^{(1)},\ s_0^{(1)},$$

$$d\ S_t^{(2)} = \mu_2 S_t^{(2)}\ dt + \sigma_2 S^{(2)} dW_t^{(2)},\ s_0^{(2)},$$

$$dW^1 dW^2 = \rho\ dt.$$

Mock Exam 2 II

Recall that ρ can be interpreted as an instantaneous correlation between changes in S^1 and S^2,

$$\text{“corr”}(dS_t^1, dS_t^2) = \rho.$$

Write the stocks dynamics under the risk neutral measure Q.
b) Calculate the price V_{BS} of the portfolio at time 0. You can use the formula for a call option in Black Scholes without deriving it, but for the put option derive it with put-call parity from the call.
c) Does the price V_{BS} of the portfolio depend on the correlation ρ? Elaborate and provide intuition on your answer. Also, calculate the sensitivity $\frac{\partial V_{BS}}{\partial \rho}$.
d) How is the portfolio price sensitive to the volatility of the second stock? Calculate $\frac{\partial V_{BS}}{\partial \sigma_2}$. Take the special case of at-the-money-forward options, namely $K_1 = s_0^{(1)} e^{rT}$ and $K_2 = s_0^{(2)} e^{rT}$, and find the sensitivity

Mock Exam 2 III

$\frac{\partial V_{BS}}{\partial \sigma_2}$ in this special case. Describe how the portfolio price changes with σ_2 in this special case. [Hint: if you proceed by differentiation wrt σ_2, you can use the lemma $s_0^{(2)}\phi(d_1^{(2)}) = K_2 e^{-rT}\phi(d_2^{(2)})$. Or, if you remember the formula for $Vega_{Call} = Vega_{Put}$ in Black Scholes, you can use it directly for the second option.]
e) How is the portfolio price sensitive to both volatilities σ_1 and σ_2 of the two stocks? Calculate $\frac{\partial^2 V_{BS}}{\partial \sigma_1 \, \partial \sigma_2}$. Explain your answer.

Mock Exam 2 IV

Solutions Problem 2.
a) Under the risk neutral measure Q, the drift rates μ_1 and μ_2 are replaced by the risk-free rate r. We thus get

$$d\ S_t^{(1)} = rS_t^{(1)}\ dt + \sigma_1 S^{(1)} dW_t^{(1),Q},\ s_0^{(1)},$$

$$d\ S_t^{(2)} = rS_t^{(2)}\ dt + \sigma_2 S^{(2)} dW_t^{(2),Q},\ s_0^{(2)},$$

$$dW^{1,Q} dW^{2,Q} = \rho\ dt$$

where W^Q are Brownian motions under Q. The reason why $dW^1 dW^2$ stays the same comes from Girsanov theorem written for Brownians (see last formula in the Girsanov Appendix plus $dt\ dW = 0$ & $dt\ dt = 0$)

$$dW^{1,Q} dW^{2,Q} = \left(\frac{\mu_1 - r}{\sigma_1} dt + dW^{1,P}\right)\left(\frac{\mu_2 - r}{\sigma_2} dt + dW^{2,P}\right) = dW^{1,P} dW^{2,P}.$$

Mock Exam 2 V

b) We can price Y as risk neutral expectation of discounted payoff.

$$V_{BS} = E^Q[e^{-rT}Y] = E^Q[e^{-rT}[(S_T^{(1)} - K_1)^+ + (K_2 - S_T^{(2)})^+]]$$

$$= E^Q[e^{-rT}(S_T^{(1)} - K_1)^+ e^{-rT}(K_2 - S_T^{(2)})^+]$$

$$= E^Q[e^{-rT}(S_T^{(1)} - K_1)^+] + E^Q[e^{-rT}(K_2 - S_T^{(2)})^+].$$

The first expectation is the Black Scholes price of a call option on stock S_1. We know this is

$$E^Q[e^{-rT}(S_T^{(1)} - K_1)^+] = s_0^{(1)}\Phi(d_1^1) - K_1 e^{-rT}\Phi(d_2^1)$$

$$d_{1,2}^1 = \frac{\ln(s_0^{(1)}/K_1) + (r \pm \sigma_1^2/2)T}{\sigma_1\sqrt{T}}$$

Mock Exam 2 VI

The second expectation is the price of a put option on she stock S_2. We derive this from the price of the call through put-call parity. Put call parity tells use that

$$E^Q[e^{-rT}(K_2 - S_T^{(2)})^+] = V_{BS}^{PUT_2} = V_{BS}^{CALL_2} - (s_0^{(2)} - K_2 e^{-rT}),$$

or

$$V_{BS}^{PUT_2} = s_0^{(2)}\Phi(d_1^2) - K_2 e^{-rT}\Phi(d_2^2) - (s_0^{(2)} - K_2 e^{-rT}),$$

$$= K_2 e^{-rT}\Phi(-d_2^2) - s_0^{(2)}\Phi(-d_1^2)$$

$$d_{1,2}^2 = \frac{\ln(s_0^{(2)}/K_2) + (r \pm \sigma_2^2/2)T}{\sigma_2\sqrt{T}}.$$

Mock Exam 2 VII

Adding call and put, $V_{BS} = V_{BS}^{CALL_1} + V_{BS}^{PUT_2}$, or

$$V_{BS} = s_0^{(1)}\Phi(d_1^1) - K_1 e^{-rT}\Phi(d_2^1) + K_2 e^{-rT}\Phi(-d_2^2) - s_0^{(2)}\Phi(-d_1^2).$$

c) Inspecting carefully the formula for V_{BS} we see that ρ appears nowhere. This means that the portfolio price does not depend on ρ and we get $\frac{\partial V}{\partial \rho} = 0$.
Why is that? The reason is in the shape of the payoff, which can be decomposed additively in the sum of two options, each depending only on one stock. In other terms, there is never an expectation involving both stocks at the same time, so that the joint statistics of the two stocks doesn't play a role, and the correlation does not play a role. Indeed, given that

Mock Exam 2 VIII

$$V_{BS} = E^Q[e^{-rT}(S_T^{(1)} - K_1)^+] + E^Q[e^{-rT}(S_T^{(2)} - K_2)^+]$$

we see that the first expectation will depend only on the first stock statistics, whereas the second only on the second stock. There is no term depending on both stocks together.

d) Compute $\frac{\partial V_{BS}}{\partial \sigma_2}$ as

$$\frac{\partial}{\partial \sigma_2}\left(s_0^{(1)}\Phi(d_1^1) - K_1 e^{-rT}\Phi(d_2^1) + K_2 e^{-rT}\Phi(-d_2^2) - s_0^{(2)}\Phi(-d_1^2)\right)$$

$$= 0 + \frac{\partial}{\partial \sigma_2}\left(K_2 e^{-rT}\Phi(-d_2^2) - s_0^{(2)}\Phi(-d_1^2)\right) = \dots$$

Mock Exam 2 IX

as the first option does not depend on σ_2, which is only found in the d^2 terms.

$$\ldots = K_2 e^{-rT} \frac{\partial}{\partial \sigma_2} \Phi(-d_2^2) - s_0^{(2)} \frac{\partial}{\partial \sigma_2} \Phi(-d_1^2) =$$

$$= -K_2 e^{-rT} \phi(-d_2^2) \frac{\partial}{\partial \sigma_2}(d_2^2) + s_0^{(2)} \phi(-d_1^2) \frac{\partial}{\partial \sigma_2}(d_1^2) = \ldots$$

where we used the chain rule $\frac{d}{d\sigma}\Phi(f(\sigma)) = \Phi'(f(\sigma))\frac{df}{d\sigma}$ and the fact that $\Phi' = \phi$, the probability density function of the standard normal.

$$\ldots = -K_2 e^{-rT} \phi(-d_2^2) \frac{\partial}{\partial \sigma_2} \left(\frac{\ln(s_0^{(2)}/K_2) + (r - \sigma_2^2/2)T}{\sigma_2 \sqrt{T}} \right)$$

$$+ s_0^{(2)} \phi(-d_1^2) \frac{\partial}{\partial \sigma_2} \left(\frac{\ln(s_0^{(2)}/K_2) + (r + \sigma_2^2/2)T}{\sigma_2 \sqrt{T}} \right)$$

Mock Exam 2 X

Calculating the derivatives we obtain

$$\frac{\partial V_{BS}}{\partial \sigma_2} = -K_2 e^{-rT} \phi(-d_2^2) \left(\frac{(-\sigma_2^2/2 - r)T - \ln(s_0^{(2)}/K_2)}{\sigma_2^2 \sqrt{T}} \right)$$

$$+ s_0^{(2)} \phi(-d_1^2) \left(\frac{(\sigma_2^2/2 - r)T - \ln(s_0^{(2)}/K_2)}{\sigma_2^2 \sqrt{T}} \right)$$

Using the lemma in the hint and the fact that $\phi(-x) = \phi(x)$, we get, after a few simplifications:

$$\frac{\partial V_{BS}}{\partial \sigma_2} = s_0^{(2)} \phi(d_1^2) \sqrt{T}$$

Mock Exam 2 XI

which is the classic $Vega_{Put} = Vega_{Call}$ in Black Scholes. If you rememberd this Vega formula, you could have cited it directly and it would have been accepted as solution.

In the special at-the-money-forward (ATMF) case $K_2 = s_0^{(2)} e^{rT}$ the above formula specializes to

$$\frac{\partial V_{BS}}{\partial \sigma_2}|_{ATMF} = s_0^{(2)} \phi(d_1^2)\sqrt{T} = s_0^{(2)}\sqrt{T}\phi\left(\sigma_2\sqrt{T}/2\right) > 0$$

meaning that $V_{BS}|_{ATMF}$ increases with σ_2, as its derivative is positive. The portfolio value will increase when σ_2 increases and will decrease when σ_2 decreases.

Mock Exam 2 XII

e) It is immediate to see that the second derivative

$$\frac{\partial^2 V_{BS}}{\partial\sigma_1\partial\sigma_2} = 0.$$

Indeed, we have computed previously $\frac{\partial V_{BS}}{\partial\partial\sigma_2}$ and by inspection we can see that this derivative does not depend on σ_1. Thus

$$\frac{\partial^2 V_{BS}}{\partial\sigma_1\partial\sigma_2} = \frac{\partial}{\partial\sigma_1}\left(\frac{\partial V_{BS}}{\partial\sigma_2}\right) = \frac{\partial}{\partial\sigma_1}(\text{quantity without } \sigma_1) = 0.$$

The reason for this is similar to the reason we have seen in point c). The price is the sum of two prices: the first price depends only on σ_1 (the call) while the second price depends only on σ_2 (the put). There is no term depending on σ_1 and σ_2 jointly, so that when we differentiate we never find both variables and the derivative is zero.

Mock Exam 2 I

Problem 3. Consider a portfolio long a call option on a first stock $S^{(1)}$ with strike K_1 and short a put option on a second stock $S^{(2)}$ with strike K_2, both options with maturity T. The final payoff of the portfolio is

$$Y = (S_T^{(1)} - K_1)^+ - (K_2 - S_T^{(2)})^+.$$

The risk-free rate is assumed to be a non-negative deterministic constant $r \geq 0$. We assume the strikes are the at-the-money-forward strikes, $K_{1,2} = S_0^{(1,2)} e^{rT}$.

a) Assume that the market volatilty smile curve is roughly constant for the first stock, and is decreasing for the second one. Choose suitable models for the first and second stock that are consistent with this pattern.

Mock Exam 2 II

b) With the chosen models, price the portfolio at time 0.

c) If the market smile pattern for the second stock had been V shaped, what smile model would have you chosen for the second stock? You are not requested to calculate the price with the chosen model, only to discuss the model.

Mock Exam 2 III

Problem 3: Solutions. a) A constant market volatility smile curve is in line with the Black Scholes model, so we can choose Black Scholes for the first Stock. Under the risk neutral measure Q, we write:

$$d\ S_t^{(1)} = rS_t^{(1)}\ dt + \sigma_1 S^{(1)} dW_t^{(1),Q},\ s_0^{(1)}.$$

A decreasing smile curve is consistent with three models we have seen: Bachelier, Displaced Diffusion with negative shift and CEV with exponent smaller than one. Any of these models can be chosen. It is preferable to choose a model with more parameters so as to be able to explain the market smile better. Bachelier only has one parameter, the absolute volatility σ, whereas both CEV and DD have two parameters. We know that DD is more tractable, so we chooose the displaced diffusion model for the second stock,

$$d\ S_t^{(2)} = rS_t^{(2)}\ dt + \sigma_2(S_t^{(2)} - \alpha e^{rt}) dW_t^{(2),Q},\ s_0^{(2)},$$

Mock Exam 2 IV

$$dW^{1,Q}dW^{2,Q} = \rho\, dt$$

where W^Q are Brownian motions under Q.

b) To price the portfolio we compute

$$V = E^Q[e^{-rT}Y] = E^Q\{e^{-rT}[(S_T^{(1)} - K_1)^+ - (K_2 - S_T^{(2)})^+]\}$$

$$= E^Q\{e^{-rT}(S_T^{(1)} - K_1)^+ - e^{-rT}(K_2 - S_T^{(2)})^+\} =$$

$$= E^Q\{e^{-rT}(S_T^{(1)} - K_1)^+\} - E^Q\{e^{-rT}(K_2 - S_T^{(2)})^+\}$$

so that the price is the call price on S_1 with Black Scholes minus the Put price on S_2 with DD.
The call price is easy and it is the usual Black Scholes call price (write down the formula, no need to derive it if you remember it).

Mock Exam 2 V

For the put price, we use put-call parity.

$$V_{DD}^{PUT_2} = V_{DD}^{CALL_2} - (s_0^{(2)} - K_2 e^{-rT})$$

We now compute

$$V_{DD}^{CALL_2} = E^Q\{e^{-rT}(S_T^{(2)} - K_2)^+\} = \dots$$

We know from the lectures that it is often convenient to write the displaced diffusion model as

$$S_t^{(2)} = X_t + \alpha e^{rt}, \quad dX_t = rX_t dt + \sigma_2 X_t dW_t^{(2),Q}, \; X_0 = S_0^{(2)} - \alpha.$$

This leads to

$$\dots = E^Q\{e^{-rT}(X_T + \alpha e_T^{rT} - K_2)^+\} = E^Q\{e^{-rT}(X_T - K')^+\}, \; K' = K_2 - \alpha e^{rT}.$$

Mock Exam 2 VI

The last expectation is a call option in the Black Scholes model X and is given by the standard call option price formula

$$V_{DD}^{CALL_2} = E^Q\{e^{-rT}(X_T - K')^+\} = x_0\Phi(d_1') - K'e^{-rT}\Phi(d_2') =$$

$$= (s_0 - \alpha)\Phi(d_1') - (K_2 - \alpha e^{rT})e^{-rT}\Phi(d_2'),\ d_{1,2}' = \frac{\ln \frac{s_0^{(2)} - \alpha}{K_2 - \alpha e^{rT}} + (r \pm \frac{1}{2}\sigma_2^2)T}{\sigma_2\sqrt{T}}$$

We can get the put by parity,

$$V_{DD}^{PUT_2} = V_{DD}^{CALL_2} - (s_0^{(2)} - K_2 e^{-rT})$$

$$= (s_0^{(2)} - \alpha)\Phi(d_1') - (K_2 - \alpha e^{rT})e^{-rT}\Phi(d_2') - (s_0^{(2)} - K_2 e^{-rT})$$

Finally, to obtain the whole portfolio price we need to add the call on S_1 with Black Scholes,

Mock Exam 2 VII

$$V = s_0^{(1)}\Phi(d_1^1) - K_1 e^{-rT}\Phi(d_2^1) - \Big[(s_0^{(2)} - \alpha)\Phi(d_1') - (K_2 - \alpha e^{rT})e^{-rT}\Phi(d_2')$$

$$-(s_0^{(2)} - K_2 e^{-rT})\Big]$$

where

$$d_{1,2}^1 = \frac{\ln\frac{s_0^{(1)}}{K_1} + (r \pm \frac{1}{2}\sigma_1^2)T}{\sigma_1\sqrt{T}}$$

We can now specialize the solution to the at-the-money forward strikes $K_{1,2} = S_0^{(1,2)}e^{rT}$. We obtain

$$V = s_0^{(1)}\Phi(\bar{d}_1^1) - s_0^{(1)}\Phi(\bar{d}_2^1) - (s_0^{(2)} - \alpha)\Phi(\bar{d}_1') + (s_0^{(2)} - \alpha)\Phi(\bar{d}_2')$$

Mock Exam 2 VIII

where

$$\bar{d}^1_{1,2} = \frac{\ln \frac{s_0^{(1)}}{s_0^{(1)} e^{rT}} + (r \pm \frac{1}{2}\sigma_1^2)T}{\sigma_1\sqrt{T}} = \pm\frac{1}{2}\sigma_1\sqrt{T}$$

$$\bar{d}'_{1,2} = \frac{\ln \frac{s_0^{(2)}-\alpha}{e^{rT}(s_0^{(2)}-\alpha)} + (r \pm \frac{1}{2}\sigma_2^2)T}{\sigma_2\sqrt{T}} = \pm\frac{1}{2}\sigma_2\sqrt{T}.$$

Simplifying further

$$V = s_0^{(1)}(\Phi(\bar{d}^1_1) - \Phi(\bar{d}^1_2)) - (s_0^{(2)} - \alpha)(\Phi(\bar{d}'_1) - \Phi(\bar{d}'_2))$$

$$= s_0^{(1)}\left(2\Phi\left(\frac{1}{2}\sigma_1\sqrt{T}\right) - 1\right) - (s_0^{(2)} - \alpha)\left(2\Phi\left(\frac{1}{2}\sigma_2\sqrt{T}\right) - 1\right)$$

where we used $\Phi(x) - \Phi(-x) = \Phi(x) - (1 - \Phi(x)) = 2\Phi(x) - 1$.

Mock Exam 2 IX

c) The only model we discussed in detail that has a V shaped smile is the mixture dynamics model, so in this case we would choose that model. Another option would be using a stochastic volatility model like Heston, but we have not discussed this model in detail.

Mock Exam 2 I

Problem 4. Risk measures on a portfolio with two stocks.
Consider two stocks in a market with zero interest rate $r = 0$. We assume both stocks follow a Bachelier model under the measure P:

$$d\ S_t^{(1)} = \mu_1\ dt + \sigma_1 dW_t^{(1)},\ s_0^{(1)},$$

$$d\ S_t^{(2)} = \mu_2\ dt + \sigma_2 dW_t^{(2)},\ s_0^{(2)},$$

$$dW^1 dW^2 = \rho\ dt$$

where W's are Brownian motions under P.
a) Consider a portfolio long an amount N_1 of stock $S^{(1)}$, short an amount N_2 of stock $S^{(2)}$, both with maturity T, and long a zero coupon bond with maturity T and notional N_B. The portfolio payoff at maturity T is

$$Y = N_1 S_T^{(1)} - N_2 S_T^{(2)} + N_B.$$

Mock Exam 2 II

Compute the value at risk of this portfolio when $\rho = 0$ (i.e. the two Brownian motions W^1 and W^2 are independent and so are the two stocks), for a risk horizon $H < T$ at a confidence level α.

b) Thinking of the Barings collapse, describe a situation that could put the bank at serious risk when trading this portfolio. For example, would an extremely large N_1 with small N_2, N_B be dangerous? Large N_2 with small N_1, N_B? Large N_B with small $N_{1,2}$? Analyze the three cases and discuss.

Mock Exam 2 III

Problem 4: solutions. a) Recall the definition: VaR is related to the potential loss on our portfolio over the time horizon H. Define this loss L_H as the difference between the value of the portfolio today (time 0) and in the future H.

$$L_H = \text{Portfolio}_0 - \text{Portfolio}_H.$$

VaR with horizon H and confidence level α is defined as that number $q = q_{H,\alpha}$ such that

$$P[L_H < q] = \alpha$$

so that our loss at time H is smaller than q with P-probability α.
We know that $S_H^{(1,2)}$ can be written as

$$S_H^{(1,2)} = S_0^{(1,2)} + \mu_{1,2} H + \sigma_{1,2} W_H^{(1,2)}, \tag{38}$$

Mock Exam 2 IV

and recalling the distribution of $W_H^{(1,2)}$,

$$S_H^{(1,2)} = S_0^{(1,2)} + \mu_{1,2} H + \sqrt{H}\sigma_{1,2}\mathcal{N}_{1,2}(0,1) \qquad (39)$$

where $\mathcal{N}_1$ and $\mathcal{N}_2$ are independent standard normals. Let us calculate the price of our portfolio at time 0. The stock and bond positions are trivial, there is no option, so we can write the portfolio price using the stock prices at time 0, $S_0^{(1)}$ and $S_0^{(2)}$, and the bond price at time 0, e^{-rT}. We get

$$\text{Portfolio}_0 = N_1 S_0^{(1)} - N_2 S_0^{(2)} + N_B e^{-rT}.$$

Similarly, at time $t = H$ we can write the portfolio price using the stock prices at time H, $S_H^{(1)}$ and $S_H^{(2)}$, and the bond price at time H, $e^{-r(T-H)}$.

$$\text{Portfolio}_H = N_1 S_H^{(1)} - N_2 S_H^{(2)} + N_B e^{-r(T-H)}.$$

Mock Exam 2 V

Then

$$L_H = \text{Portfolio}_0 - \text{Portfolio}_H =$$

$$= N_1(S_0^{(1)} - S_H^{(1)}) - N_2(S_0^{(2)} - S_H^{(2)}) - N_B(e^{-r(T-H)} - e^{-rT})$$

The only random parts in this portfolio are $S_H^{(1)}$ and $S_H^{(2)}$. Let us consider

$$X = -N_1 S_H^{(1)} + N_2 S_H^{(2)},$$

$$K = N_1 S_0^{(1)} - N_2 S_0^{(2)} - N_B(e^{-r(T-H)} - e^{-rT}).$$

This way

$$L_H = X + K$$

where X is random and K is deterministic.

$$X = -N_1 S_H^{(1)} + N_2 S_H^{(2)}$$

Mock Exam 2 VI

$$= -N_1(S_0^{(1)} + \mu_1 H + \sqrt{H}\sigma_1\mathcal{N}_1) + N_2(S_0^{(2)} + \mu_2 H + \sqrt{H}\sigma_2\mathcal{N}_2)$$

Then

$$L_H = X + K = -N_1(\mu_1 H + \sqrt{H}\sigma_1\mathcal{N}_1) +$$
$$+N_2(\mu_2 H + \sqrt{H}\sigma_2\mathcal{N}_2) - N_B(e^{-r(T-H)} - e^{-rT}) =$$
$$= Z + C$$

where

$$Z = -N_1\sqrt{H}\sigma_1\mathcal{N}_1 + N_2\sqrt{H}\sigma_2\mathcal{N}_2$$
$$C = -N_1\mu_1 H + N_2\mu_2 H - N_B(e^{-r(T-H)} - e^{-rT}).$$

Z is random while C is deterministic. Let's derive the distribution of Z. If $\mathcal{N}_1$ is a standard normal, also $-\mathcal{N}_1$ is a standard normal. We can thus say that

Mock Exam 2 VII

$$Z_1 := -N_1\sqrt{H}\sigma_1\mathcal{N}_1 = N_1\sqrt{H}\sigma_1(-\mathcal{N}_1) \sim$$
$$\sim Normal_1(0, N_1^2 H\sigma_1^2)$$

and

$$Z_2 := N_2\sqrt{H}\sigma_2\mathcal{N}_2 \sim Normal_2(0, N_2^2 H\sigma_2^2).$$

As the two normals are independent,

$$Z = Z_1 + Z_2 = Normal_1(0, N_1^2 H\sigma_1^2) + Normal_2(0, N_2^2 H\sigma_2^2) \sim$$

$$\sim Normal(0, N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2) \sim \sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2}\ Normal(0,1)$$

as the sum of two independent normals is a normal with mean the sum of the means and with vaiance the sum of variances.

Mock Exam 2 VIII

As $L_H = Z + C$, we get

$$L_H \sim C + \sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2}\, Normal(0,1)$$

We can now calculate

$$\alpha = P[L_H < q] = P\left[C + \sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2}\, Normal(0,1) < q\right]$$

$$= P\left[Normal(0,1) < \frac{q - C}{\sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2}}\right]$$

$$= \Phi\left(\frac{q - C}{\sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2}}\right).$$

Mock Exam 2 IX

From

$$\alpha = \Phi\left(\frac{q - C}{\sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2}}\right)$$

apply Φ^{-1} to both sides to get

$$\Phi^{-1}(\alpha) = \frac{q - C}{\sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2}}$$

$$\text{and} \quad q = VaR_{\alpha,H} = \Phi^{-1}(\alpha)\sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2} + C$$

and, recalling the expression of C,

$$VaR_{\alpha,H} = \Phi^{-1}(\alpha)\sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2} +$$

Mock Exam 2 X

$$-N_1\mu_1 H + N_2\mu_2 H - N_B(e^{-r(T-H)} - e^{-rT}).$$

Now recall that we are assuming zero interest rates, $r = 0$. This means that the VaR is

$$VaR_{\alpha,H} = \Phi^{-1}(\alpha)\sqrt{N_1^2 H\sigma_1^2 + N_2^2 H\sigma_2^2} - N_1\mu_1 H + N_2\mu_2 H.$$

In other terms, the bond component of VaR

$$N_B(e^{-r(T-H)} - e^{-rT}) = N_B(1-1) = 0$$

and this is due to the interest rate being 0. Indeed, with zero interest rates the bond has constant value

$$P(t,T) = e^{-r(T-t)} = e^0 = 1$$

Mock Exam 2 XI

so that holding a bond position adds no risk, as the bond never changes value.

b) Thinking of the Barings collapse, we see what kind of positions would lead to a huge VaR.
We first assume the position in the first stock to be very large compared to the other two positions. In other terms

$$N_1 >> N_2, \quad N_1 >> N_B.$$

write

$$VaR_{\alpha,H} = \Phi^{-1}(\alpha) N_1 \sqrt{H\sigma_1^2 + (N_2/N_1)^2 H\sigma_2^2} +$$
$$-N_1 \left(\mu_1 H - \frac{N_2}{N_1}\mu_2 H + \frac{N_B}{N_1}(e^{-r(T-H)} - e^{-rT}) \right).$$

Mock Exam 2 XII

As $N_1 >> N_2, \;\; N_1 >> N_B$, the ratios N_2/N_1 and N_B/N_1 will be close to 0. In the limits where N_1 is extremely large these ratios are negligible, and the VaR becomes close to

$$VaR_{\alpha,H} \approx \Phi^{-1}(\alpha) N_1 \sqrt{H\sigma_1^2} - N_1 \mu_1 H = N_1 \sqrt{H}(\Phi^{-1}(\alpha)\sigma_1 - \mu_1 \sqrt{H}).$$

A very large positive VaR is dangerous, while a very negative VaR would be good. Therefore, all depends on the sign of $\Phi^{-1}(\alpha)\sigma_1 - \mu_1\sqrt{H}$. If this is positive, namely $\sigma_1 > \frac{\mu_1 \sqrt{H}}{\Phi^{-1}(\alpha)}$, then very large N_1 will be very dangerous, leading to large potential losses. As an example, suppose we are taking a 99% confidence VaR, with a 1 year risk horizon. $H = 1y$, $\alpha = 0.99$. We have $\Phi^{-1}(0.99) = 2.33$ and we are in danger with large N_1 if

$$\sigma_1 > \frac{\mu_1}{2.33} \;\; \text{or} \;\; \mu_1 < 2.33\sigma_1.$$

Mock Exam 2 XIII

So if the instantaneous trend (drift) of the first stock is smaller than a given proportion 2.33 of the volatility of the same stock, we can be in trouble. Note that we are long the first stock, so if the first stock goes up we are good, whereas if it goes down we face a loss. The stock is more likely to go up if its instantaneous growth rate, or drift, is positive and large. Indeed, we see that if $\mu_1 \geq 2.33\sigma_1$ our VaR will be zero or negative even for very large N_1, meaning that we are not in trouble. This would be because our huge position N_1 is on a stock that has positive trend that dominates the volatility (times 2.33) of that stock. So we expect the stock to grow rather than decrease, statistically, and we can have a negative or zero VaR from this long position.

Mock Exam 2 XIV

A similar analysis carries out when N_2 is exceedingly large compared to N_1 and N_2. We get

$$N_2 >> N_1, \quad N_2 >> N_B.$$

Write

$$VaR_{\alpha,H} = \Phi^{-1}(\alpha) N_2 \sqrt{(N_1/N_2)^2 H\sigma_1^2 + H\sigma_2^2} +$$

$$+N_2 \left(-\frac{N_1}{N_2}\mu_1 H + \mu_2 H - \frac{N_B}{N_2}(e^{-r(T-H)} - e^{-rT}) \right).$$

As $N_2 >> N_1, \quad N_2 >> N_B$, the ratios N_1/N_2 and N_B/N_2 will be close to 0. In the limits where N_2 is extremely large these ratios are negligible, and the VaR becomes close to

$$VaR_{\alpha,H} \approx \Phi^{-1}(\alpha) N_2 \sqrt{H\sigma_2^2} + N_2 \mu_2 H = N_2 \sqrt{H} (\Phi^{-1}(\alpha)\sigma_2 + \mu_2 \sqrt{H}).$$

Mock Exam 2 XV

The dangerous situation is when *VaR* is positive and very large. This now happens in all cases, if we assume that both the volatility and the expected growth term (drift) μ_2 are positive.
The risk is only absent if we have a negative term

$$\Phi^{-1}(\alpha)\sigma_2 + \mu_2\sqrt{H} < 0 \iff \mu_2 < -\frac{\Phi^{-1}(\alpha)\sigma_2}{\sqrt{H}}.$$

We see that μ would have to be negative and below a negative value proportional to the volatility for the position. Again, if $\alpha = 0.99$ and $H = 1$ we get that VaR is negative or zero (no danger) if

$$\mu_2 < -2.33\sigma_2.$$

So extremely large positions in the second stock with small positions in the first stock and the bond can lead to extremely large VaR in all

Mock Exam 2 XVI

cases, unless the drift of the second stock is negative and below a given negative proportion of the volatility.
Indeed, note that we are short the second stock, so that we will face a huge loss if the second stock increases and we will avoid the loss if the second stock goes down. For the stock to go down, statistically, it needs a negative drift and this has to be below a given proportion of the volatility.

Finally, we consider the case where $N_B >> N_1, \; N_B >> N_2$.

$$VaR_{\alpha,H} = N_B \left(\Phi^{-1}(\alpha) \sqrt{\left(\frac{N_1}{N_B}\right)^2 H\sigma_1^2 + \left(\frac{N_2}{N_B}\right)^2 H\sigma_2^2 +} \right.$$

$$\left. - \frac{N_1}{N_B}\mu_1 H + \frac{N_2}{N_B}\mu_2 H - (e^{-r(T-H)} - e^{-rT}) \right).$$

Mock Exam 2 XVII

We see that as N_B grows exceedinly large compared to N_1 and N_2 we get

$$VaR_{\alpha,H} \approx -N_B\left(e^{-r(T-H)} - e^{-rT}\right) = -N_B e^{-rT}(e^{rH} - 1).$$

Now we assumed $r = 0$, so this last term is zero, and the VaR is zero. There is no risk in holding very large bond positions with zero interest rates. What would happen if $r > 0$? This is not required for the solution, but let's discuss it anyway. If $r > 0$ and we are left only with the bond position, this would give us always a negative VaR, meaning no danger, given that the term in brackets is always positive, given $H > 0$ and $r > 0$. Indeed, as we have a long position in the bond and the bond always increases value in time, as $P(t, T) = e^{-r(T-t)}$ is increasing in t for $r > 0$, we are good.

Mock Exam 2 XVIII

Hence a long bond position with zero or positive rates is never dangerous for VaR in the modeling context of this problem. Negative rates, on the other hand, $r < 0$, would make the position dangerous.

Mock Exam 3 I

Consider the Ito SDE

$$dX_t = \frac{1}{3}(X_t)^{1/3}dt + (X_t)^{2/3}dW_t, \; X_0 = x_0$$

where the initial condition is a deterministic constant.
a) Do not try to prove existence and uniqueness of a solution a priori, invoking a theorem. Try to find one explicit solution using calculus and then check it is fine a posteriori. Hint: you may tranform in Stratonovich form and then use separation of variables and / or change of variables.
b) Check a posteriori that the solution you found satisfies the given Ito SDE.
c) Take the case $X_0 = 0$. Show that the solution is not unque by providing a second solution. [Hint: you can find easily a constant second solution in this case.] Why was it reasonable not to expect uniqueness in the first place?

Mock Exam 3 II

Problem 1: Solutions.
a) Consider

$$dX_t = \frac{1}{3}(X_t)^{1/3}dt + (X_t)^{2/3}dW_t,\ X_0 = x_0.$$

$\sigma(x) = (x)^{2/3}$. The equivalent Stratonovich SDE is obtained by changing the drift by

$$\frac{1}{3}x^{1/3} \to \frac{1}{3}x^{1/3} - \frac{1}{2}\sigma(x)\frac{d}{dx}\sigma(x) =$$

$$= \frac{1}{3}x^{1/3} - \frac{1}{2}(x)^{2/3}\frac{2}{3}x^{-1/3} = \frac{1}{3}(x)^{1/3} - \frac{1}{3}(x)^{1/3} = 0.$$

So the equivlent Stratonovich SDE has zero drift,

$$dX_t = X_t^{2/3} \circ dW_t.$$

Mock Exam 3 III

Now we know that the Stratonovich SDE obeys the formal rules of calculus. Let's try to solve the SDE by separating variables:

$$\frac{dX_t}{X_t^{2/3}} = 1 \circ dW_t$$

Integrate both sides

$$\int_{X_0}^{X_t} \frac{dX}{X^{2/3}} = \int_0^t 1 \circ dW_t$$

leading to

$$3(X_t^{1/3} - X_0^{1/3}) = W_t$$

and therefore

$$X_t^{1/3} = \frac{W_t}{3} + X_0^{1/3}$$

Mock Exam 3 IV

Taking the cube on both sides, our solution is

$$X_t = \left({X_0}^{1/3} + \frac{W_t}{3} \right)^3.$$

b) Let's check this is correct. Write

$$X_t = \left({X_0}^{1/3} + \frac{W_t}{3} \right)^3 = Z_t^3, \;\; Z_t = {X_0}^{1/3} + \frac{W_t}{3}.$$

Let's differentiate X_t as a function of Z_t using Ito's formula.

$$dX_t = 3Z^2 dZ + \frac{1}{2} 6Z \, dZdZ = 3Z^2 dZ + 3Z \, dZdZ$$

Mock Exam 3 V

i.e.

$$dX_t = 3Z^2 dZ + \frac{1}{3} Z_t dt$$

As $X = Z^3$, we have $Z = X^{1/3}$ and

$$dX_t = 3X_t^{2/3} d \underbrace{\left({X_0}^{1/3} + \frac{W_t}{3} \right)}_{Z_t} + \frac{1}{3} X_t^{1/3} dt$$

or

$$dX_t = \frac{1}{3} X_t^{1/3} dt + X_t^{2/3} dW_t$$

which is our initial Ito SDE.

Mock Exam 3 VI

c) Consider

$$dX_t = \frac{1}{3}(X_t)^{1/3}dt + (X_t)^{2/3}dW_t, \ X_0 = 0.$$

Let's try a constant solution $X_t = k$, for a constant k. Given that $X_0 = 0$ and that the solution is constant in time, we need to have $k = 0$. Then $dX_t = dk = d0 = 0$ and the SDE reads

$$0 = \frac{1}{3}(0)^{1/3}dt + (0)^{2/3}dW_t$$

leading to the identity

$$0 = 0$$

Mock Exam 3 VII

so that the equation is satisfied and indeed $X_t = 0$ is a solution. Hence in the case $X_0 = 0$ we have at least two solutions: the previous solution we found

$$X_t = \left({X_0}^{1/3} + \frac{W_t}{3} \right)^3 = \left(\frac{W_t}{3} \right)^3 .$$

The new solution we found is

$$X_t = 0.$$

The two solutions are clearly different.
It was reasonable not to expect uniqueness as the drift and diffusion coefficients $x^{1/3}$ and $x^{2/3}$ do not satisfy the Lipschitz condition. In particular, note that the coefficients do not admit first derivative in zero, as the derivative grows larger and larger as we approach zero from

Mock Exam 3 VIII

either direction, and this makes the Lipschitz condition fail near zero. Linear growth still works, as we have

$$|\mu(t, X)| + |\sigma(t, X)| = 1/3|X^{1/3}| + |X^{2/3}| \leq 2(1 + |X|)$$

having both terms on the left hand side smaller than $1 + |X|$. However, linear growth alone is not enough. Existence and uniquness require both conditions, and as Lipschitz fails, we had no guarantee that the equation admitted a unique solution.

Mock Exam 3 I

Problem 2. Consider an asset or Nothing option (ANO) and the related cash-or-nothing option (CNO) with final maturity T and strike K on an equity stock with price S in a market with constant and deterministic interest rates $r \geq 0$. Define $Y^{AN} = S_T 1_{\{S_T > K\}}$, $Y^{CN} = K 1_{\{S_T > K\}}$ as the final payoff of the ANO and CNO options at maturity T, respectively. Assume that the stock price S follows a Black Scholes model, so that it is a geometric Brownian motion under the risk neutral measure, with volatility σ and with deterministic initial value s_0 at time 0, namely

$$dS_t = rS_t dt + \sigma S_t dW_t, \quad S_0 = s_0$$

where r, σ are positive constants and W is a brownian motion under the risk neutral measure Q.

Mock Exam 3 II

a) Draw the payoffs of the ANO and CNO as functions of the underlying stock S_T. Discuss what kind of investor would find a ANO attractive and what investor would find a CNO attractive and why.
b) Consider the payoff of a portfolio long one ANO and short one CNO on the same stock with the same strike K and maturity T. Is this portfolio payoff equivalent to a familiar option payoff? Which option? Show the detailed reasoning.
c) Derive a formula for the price of the CNO.
d) Derive a formula for the price of the ANO. Hint: you may be helped by combining solutions of b) and c) above.

Mock Exam 3 III

Problem 2 Solutions.
a) The ANO payoff is

$$Y^{AN} = S_T 1_{\{S_T > K\}} = \begin{cases} S_T & \text{if } S_T > K \\ 0 & \text{if } S_T \leq K \end{cases} .$$

The CNO payoff is

$$Y^{CN} = K 1_{\{S_T > K\}} = \begin{cases} K & \text{if } S_T > K \\ 0 & \text{if } S_T \leq K \end{cases} .$$

The ANO is an option that pays the stock at maturity only if the stock is above a threshold K, and pays nothing otherwise. Clearly, this is an option that is attractive to someone who foresees the stock to grow or stay above the level K in the future. Here are two examples of investors who would be interested in a ANO, both valid as an answer.

Mock Exam 3 IV

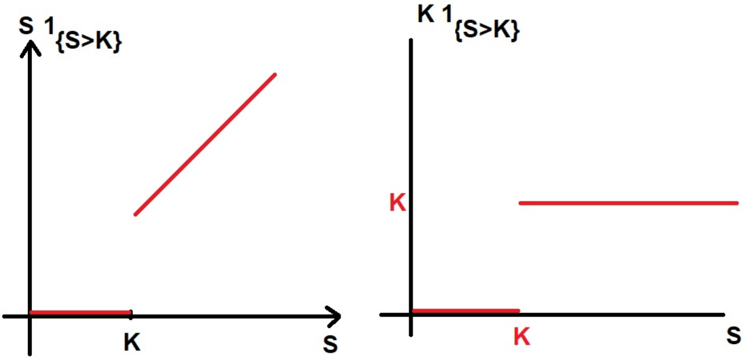
(i) ANO can be used by someone who is short the stock. Suppose you are short selling the stock with maturity T. At time T, you will have to pay S_T times the amount of stock you have been short-selling to the client. If S_T grows too much, you will have to pay too much money. If you are not comfortable paying stock prices above the level K, you can buy an amount of ANO with strike K equal to the amount of stock you have been short selling. This way, at maturity, if S_T is above K, you will receive S_T from the ANO you purchased and you can pay your client to whom you owe S_T using the ANO income. If S_T is below K the payment you have to make to your client is below the threshold level K and you are ok, which is fine since in this case the ANO gives you nothing.

(ii) ANO can also be a speculative investment for a trader who strongly believes the stock price will grow above K. Why would such a trader

Mock Exam 3 V

prefer the ANO to directly buying the stock? This is because the ANO will be cheaper. Indeed, the stock will pay S_T at time T in any market condition. The ANO will pay the stock at time T also as S_T but only if $S_T > K$. In this sense the ANO is less likely to pay than the stock itself, because for $S_T \leq K$ the stock will still pay S_T but the ANO will pay 0, whereas for $S_T > K$ they both pay S_T. So the payoff of the ANO is lower or equal to the payoff of the stock at maturity T in all scenarios, and therefore the ANO price will be smaller than the stock price. Thus a speculative investor who wants to speculate on the stock, and is confident this will rise above K, can buy a ANO, which is cheaper than the stock itself.

Mock Exam 3 VI



For the CNO we provide two examples but one is enough.

Mock Exam 3 VII

(i) The CNO is like a bond, it gives you a constant payoff K at maturity but only if the stock is above a level K. This will be attractive to people who are confident the stock will be above K at maturity T. Note that the CNO is less expensive that a zero coupon bond with the same maturity and notional K. The bond payoff is K at maturity always, whereas the CNO will be K only if $S_T > K$. This will translate in a lower price of the CNO than the bond, because there are scenarios where the CNO will pay nothing, whereas the bond will always pay K and hence is worth more.

(ii) The CNO can be sold to raise cash at time 0 (cashing the initial CNO price) by an investor who is confident that the stock price will never exceed K. This way, the investor cashes in the option price at time 0 and, if they are right and $S_T \leq K$, they will pay nothing at maturity, realizing a profit.

Mock Exam 3 VIII

b) The portfolio payoff would be

$$Y = S_T 1_{\{S_T > K\}} - K 1_{\{S_T > K\}} = (S_T - K) 1_{\{S_T > K\}} = (S_T - K) 1_{\{S_T - K > 0\}} = \ldots$$

This latest payoff is equal to $S - K$ if $S - K$ is positive, and zero otherwise. This is a call option:

$$\ldots = (S_T - K)^+.$$

c) Compute

$$V^{CN}_{BS}(0) = \mathbb{E}^Q \left[e^{-rT} K 1_{\{S_T > K\}} \right] = e^{-rT} K \mathbb{E}^Q \left[1_{\{S_T > K\}} \right] = e^{-rT} K \mathbb{Q}(S_T > K)$$

Mock Exam 3 IX

since from basic probability we know that $E^Q[1_A] = \mathbb{Q}(A)$. We are now left with computing $\mathbb{Q}(S_T > K)$. We recall the SDE for S under the risk neutral measure $\mathbb{Q}$:

$$dS_t = rS_t dt + \sigma S_t dW_t, \quad S_0.$$

Ito's formula for the natural logarithm $\ln S_t$ gives easily (exercise, write this in detail)

$$d\ln(S_t) = (r - \sigma^2/2)dt + \sigma dW_t$$

from which, writing in integral form and recalling that $W_0 = 0$

$$\ln S_T - \ln S_0 = (r - \sigma^2/2)T + \sigma W_T \sim (r - \sigma^2/2)T + \sigma\sqrt{T}\mathcal{N}(0,1)$$

Now we write

$$\mathbb{Q}(S_T > K) = \mathbb{Q}(\ln S_T > \ln K) = ...$$

Mock Exam 3 X

because logarithm is an increasing function; by substituing our expression for $\ln S_T$

$$... = \mathbb{Q}(\ln S_0 + (r - \sigma^2/2)T + \sigma\sqrt{T}\mathcal{N}(0,1) > \ln K) =$$

$$= \mathbb{Q}(\sigma\sqrt{T}\mathcal{N}(0,1) > -\ln(S_0/K) - (r - \sigma^2/2)T) =$$

$$= \mathbb{Q}\left(-\mathcal{N}(0,1) < \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}\right) =$$

$$= \Phi\left(\frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}\right) = \Phi(d_2(0))$$

where Φ is the cdf of the standard normal and where we used the fact that the opposite of a standard normal is still a standard normal.

Mock Exam 3 XI

An alternative approach to calculating $\mathbb{Q}(S_T > K)$, for a student who recalls the solution of a GBM by heart, is to write directly

$$S_T = S_0 \exp\left((r - \sigma^2/2)T + \sigma W_T\right)$$

and calculating $\mathbb{Q}(S_T > K)$ starting from there, taking logs on both sides of the probability argument and proceeding similarly to above.

The final price of the CNO is

$$V_{BS}^{CN}(0) = e^{-rT} K \Phi(d_2).$$

We can now confirm that this is smaller than the price of a bond with maturity T and notional K, an mentioned in point a)(i) for the CNO. Indeed, this bond would be worth

$$V_{Bond} = E^Q[e^{-rT}K] = e^{-rT}K > e^{-rT}K\Phi(d_2) = V_{BS}^{CN}(0)$$

Mock Exam 3 XII

as Φ is smaller than one, being a CFD.

d) We have shown earlier that the payoff of a ANO minus the payoff of a CNO is equal to a call option,

$$(S_T - K)^+ = Y^{AN} - Y^{CN}.$$

It follows that $\mathbb{E}^Q\left(e^{-rT}(S_T - K)^+\right) =$

$$= \mathbb{E}^Q\left(e^{-rT}(Y^{AN} - Y^{CN})\right) = \mathbb{E}^Q\left(e^{-rT}Y^{AN}\right) - \mathbb{E}^Q\left(e^{-rT}Y^{CN}\right)$$

or

$$V_{BS}^{CALL}(0) = V_{BS}^{AN}(0) - V_{BS}^{CN}(0).$$

Going back to our formula for the CNO we know that

$$V_{BS}^{CN}(0) = e^{-rT}K\mathbb{Q}(S_T > K) = e^{-rT}K\Phi(d_2(0)).$$

Mock Exam 3 XIII

Recall the BS formula for a Call option, written at time 0

$$V_{BS}^{CALL}(0) = S_0\Phi(d_1) - e^{-rT}K\Phi(d_2).$$

Hence

$$V_{BS}^{CALL}(0) = S_0\Phi(d_1(0)) - V_{BS}^{CN}(0).$$

Since we have seen a few lines above than

$$V_{BS}^{CALL}(0) = V_{BS}^{AN}(0) - V_{BS}^{CN}(0)$$

it follows immediately by inspection that

$$V_{BS}^{AN}(0) = S_0\Phi(d_1(0)).$$

We can now verify that indeed the price of a ANO is smaller than the price of the stock, as we mentioned in point a)(ii) for the ANO. Indeed, $S_0\Phi(d_1(0)) < S_0$ because the normal CDF Φ is always smaller than 1 for a finite argument $d_1(0)$.

Mock Exam 3 I

Problem 3. Asset/Cash or nothing options with smile.
Consider a cash-or-nothing option (CNO) with final maturity T and strike K on an equity stock with price S in a market with deterministic interest rates $r \geq 0$. Define $Y^{CN} = K 1_{\{S_T > K\}}$ as the final payoff of the CNO option at maturity T.
a) If we observe a V shaped smile for the stock in the market, with the minimum at the at-the-money-forward level, choose a volatility smile model that is suited to this situation and calculate the price of the CNO with this model.
b) Calculate the Delta of the CNO price, namely the partial derivative of the CNO price with respect to the initial stock price S_0.
c) Analyze the sign of Delta and draw some conclusions on the behaviour of the CNO price with respect to the underlying asset.

Mock Exam 3 II

d) Assume now that the CNO is at-the-money-forward, with $K = S_0 e^{rT}$. Specialize the formulas for the price and the Delta to this situation and discuss the changes.

Mock Exam 3 III

Problem 3 solutions.

a) The only local volatility model we have seen that has been able to obtain a V-shaped smile is the mixture diffusion dynamics model (MDD), and we know this model produces a smile with a minimum occurring at-the-money-forward. Recall the MDD:

$$dS_t = rS_t dt + \sigma_{\text{mix}}(t, S_t) S_t \, dW_t, \quad S_0 = s_0$$

where $\sigma_{\text{mix}}(t, S_t)$ is built in such a way that the distribution of S_t is a mixture of distributions of the lognormals $p_{i,t}$, or in formula

$$p_{S_t}(y) =: p_t(y) = \sum_{i=1}^{N} \lambda_i p_{i,t}(y) = \sum_{i=1}^{N} \lambda_i p_{t,\sigma_i}^{lognormal}(y)$$

Mock Exam 3 IV

where $\lambda_i \in (0,1)$, $\sum_{i=1}^{N} \lambda_i = 1$ and

$$p^{lognormal}_{t,\sigma_i}(x) = \frac{1}{x\sigma_i\sqrt{t\,2\pi}} \exp\left\{-\frac{1}{2\sigma_i^2 t}\left[\ln\frac{x}{S_0} - rt + \frac{1}{2}\sigma_i^2 t\right]^2\right\}.$$

The λ_i are the weights of the different lognormal densities $p_{i,t}$ on the mixture.
We can take just $N = 2$, as we have seen in numerical examples that this is already enough to generate a V-shaped smile. We will thus have λ_1 and $\lambda_2 = 1 - \lambda_1$, as lambdas add up to 1, and we will only have two sigmas, σ_1 and σ_2. The density of this MDD model will be a mixture of two densities

$$p_{S_t}(y) = \lambda_1 p^{lognormal}_{t,\sigma_1}(y) + (1-\lambda_1) p^{lognormal}_{t,\sigma_2}(y)$$

Mock Exam 3 V

Recall also the expression for $\sigma_{\text{mix}}(t,y)^2 = \frac{1}{\sum_{j=1}^{2}\lambda_j p_{j,t}(y)}\sum_{i=1}^{2}\lambda_i\sigma_i^2 p_{i,t}(y)$.
Now we calculate the CNO price with this model.

$$V_{\text{mix}}^{CNO} = e^{-rT}E^Q\left\{K1_{\{S_T>K\}}\right\}$$

$$= e^{-rT}\int_0^{+\infty} K1_{\{y>K\}}p_{S_T}(y)dy = e^{-rT}\int_0^{+\infty} K1_{\{y>K\}}\sum_{i=1}^{2}\lambda_i p_{i,T}(y)dy$$

$$= \sum_{i=1}^{2}\lambda_i\int_0^{+\infty} e^{-rT}K\,1_{\{y>K\}}p_{i,T}(y)dy = \sum_{i=1}^{2}\lambda_i V_{BS}^{CN}(0,S_0,K,T,\sigma_i,r)$$

as the last integral is simply the expectation of the discount CNO payoff under a Black Scholes model with volatility σ_i. This confirms that the price of the CNO in the MDD is a linear (actually convex) combination

Mock Exam 3 VI

of Black Scholes prices of CNOs with volatilities σ_1, σ_2 with weights $\lambda_1, \lambda_2 = 1 - \lambda_1$. We know this holds for every simple claim in MDD.
So the option price becomes a mix of two prices with the given weights and volatilities.
To complete the formula, we now need to calculate the price of the CNO in a Black Scholes model. [This has been done in Problem 2, but rewrite the derivation here.] We have

$$V_{BS}^{CN}(0, S_0, K, T, \sigma_i, r) = Ke^{-rT}\Phi(d_2(\sigma_i)),$$

$$d_2(\sigma_i) = \frac{\ln \frac{S_0}{K} + \left(r - \frac{\sigma_i^2}{2}\right) T}{\sigma_i \sqrt{T}}.$$

Mock Exam 3 VII

So the price is finally

$$V_{\text{mix}}^{CNO} = \sum_{i=1}^{2} \lambda_i K e^{-rT} \Phi(d_2(\sigma_i)).$$

b) For the delta, recall that due to the linearity of differentiation, the same convex combination we found for the CNO price applies also to the CNO Delta. Indeed, differentiate both sides of

$$V_{\text{mix}}^{CNO} = \sum_{i=1}^{2} \lambda_i K e^{-rT} \Phi(d_2(\sigma_i))$$

Mock Exam 3 VIII

by $S_0 = s_0$, obtaining

$$\Delta^{CNO}_{mix} = \frac{\partial V^{CNO}_{\text{mix}}}{\partial S_0} = \sum_{i=1}^{2} \lambda_i K e^{-rT} \frac{\partial}{\partial S_0} \Phi(d_2(\sigma_i)) =$$

$$= \sum_{i=1}^{2} \lambda_i K e^{-rT} \phi(d_2(\sigma_i)) \frac{\partial}{\partial S_0}(d_2(\sigma_i)) =$$

$$= \sum_{i=1}^{2} \lambda_i K e^{-rT} \phi(d_2(\sigma_i)) \frac{1}{S_0 \sigma_i \sqrt{T}},$$

where we have used $\frac{d}{dx}\Phi(x) = \phi(x)$, the probability density function of the standard normal, and the chain rule
$\frac{d}{dS_0}\Phi(f(S_0)) = \phi(f(S_0))\frac{df}{dS_0}(S_0)$.

Mock Exam 3 IX

c) Recall that we found

$$\Delta_{mix}^{CNO} = \sum_{i=1}^{2} \lambda_i K e^{-rT} \phi(d_2(\sigma_i)) \frac{1}{S_0 \sigma_i \sqrt{T}}.$$

As $\lambda \geq 0$, $K > 0$, $e^{-rT} > 0$, $\phi > 0$ and $S_0 > 0$, $\sigma_i > 0$ we get that

$$\Delta_{mix}^{CNO} > 0.$$

This means that V_{mix}^{CNO} is increasing with S_0 as it has a positive derivative wrt S_0. This is intuitive: the option pays K only if $S_T > K$. If we increase S_0, ceteris paribus, scenarios for S_T will become larger, and it will be more likely that $S_T > K$ and that the option pays K, so the option will be worth more.

Mock Exam 3 X

d) In case $K = S_0 e^{rT}$ we get a special value for

$$d_2(\sigma_i, K = S_0 e^{rT}) = \frac{\ln \frac{S_0}{S_0 e^{rT}} + \left(r - \frac{\sigma_i^2}{2}\right) T}{\sigma_i \sqrt{T}} = -\frac{\sigma_i}{2}\sqrt{T}.$$

Then

$$V_{\text{mix}}^{CNO}|_{ATMF} = \sum_{i=1}^{2} \lambda_i K e^{-rT} \Phi(d_2(\sigma_i, K = S_0 e^{-rT})) = \sum_{i=1}^{2} \lambda_i S_0 \Phi\left(-\frac{\sigma_i}{2}\sqrt{T}\right).$$

The ATMF price depends on S_0 only linearly now, while d_2 does not depend on S_0 anymore. To calculate Delta in this special case, we can specialize the previous Delta formula to $K = S_0 e^{rT}$. We obtain

Mock Exam 3 XI

$$\Delta^{CNO}_{mix,ATMF} = \sum_{i=1}^{2} \lambda_i S_0 e^{rT} e^{-rT} \phi(d_2(\sigma_i, K = S_0 e^{rT})) \frac{1}{S_0 \sigma_i \sqrt{T}}$$

$$= \sum_{i=1}^{2} \lambda_i \phi\left(-\frac{\sigma_i}{2}\sqrt{T}\right) \frac{1}{\sigma_i \sqrt{T}}.$$

The ATMF Delta does not depend on S_0, so the ATMF option sensitivity to S_0 will be the same for all S_0's.

Mock Exam 3 I

Problem 4: Risk Measures.
Consider the dynamics of an equity asset price S in the Bachelier model, under both probability measures P (the Physical or Historical measure) and Q (the risk neutral measure), with stock dynamics $dS_t = \mu dt + \sigma dW_t$, with μ and σ deterministic constant, $\sigma > 0$ and where W is a Brownian motion under P. Assume the risk-free interest rate is equal to zero, $r = 0$.

a) Write the risk neutral dynamics of the stock.
b) Define Expected Shortfall (ES) for a time horizon T with confidence level α for a general portfolio.

c) Compute ES for horizon T and confidence level α for a portfolio with N units of equity, where the equity price follows the Bachelier process above.

Mock Exam 3 II

d) Explain one drawback of ES as a risk measure

e) Is the equity dynamics you used for ES the same you would have used to price an equity call option in the Bachelier model?

Mock Exam 3 III

Problem 4: Solutions.
a) We know that, under the risk neutral measure, the drift of a stock is rS_t. Since $r = 0$, our model will have zero drift.

$$dS_t = \sigma dW_t^Q, \;\; s_0$$

where W^Q is a Brownian motion under Q.
b) To define ES we need first to define value at Risk (VaR). VaR is related to the potential loss on our portfolio over the time horizon T. Define this loss L_T as the difference between the value of the portfolio today (time 0) and in the future T.

$$L_T = \text{Portfolio}_0 - \text{Portfolio}_T.$$

Mock Exam 3 IV

VaR with horizon T and confidence level α is defined as that number $q = q_{T,\alpha}$ such that

$$P[L_T < q] = \alpha$$

so that our loss at time T is smaller than q with P-probability α. Recall that ES is then defined as the expectation of the loss conditional on the loss exceeding VaR:

$$\mathrm{ES}_{T,\alpha} = \mathbb{E}^{\mathbb{P}}[L_T | L_T > \mathrm{VaR}_{T,\alpha}] = \frac{\mathbb{E}^{\mathbb{P}}[L_T 1_{\{L_T > \mathrm{VaR}_{T,\alpha}\}}]}{1-\alpha}$$

[see lecture notes for the steps to get to the last expression]
c) In the Bachelier model the equity process follows the dynamics

$$dS_t = \mu \, dt + \sigma \, dW_t, \ \ s_0,$$

Mock Exam 3 V

where μ, σ are positive constants and W is a Brownian motion under the physical measure P.
We know that S_T can be written as

$$S_T = S_0 + \mu T + \sigma W_T, \tag{40}$$

and recalling the distribution of $W_T \sim \sqrt{T}\mathcal{N}(0,1)$,

$$S_T = s_0 + \mu T + \sigma\sqrt{T}\mathcal{N}(0,1) \tag{41}$$

so that in our case $L_T = N(S_0 - S_T)$, namely

$$L_T = N\left(S_0 - (S_0 + \mu T + \sigma\sqrt{T}\mathcal{N}(0,1))\right)$$

$$= N\left(-\mu T - \sigma\sqrt{T}\mathcal{N}(0,1)\right)$$

Mock Exam 3 VI

Hence, if $q = VaR_{T,\alpha}$, we get

$$\alpha = P[L_T < q] = P\left[N\left(-\mu T - \sigma\sqrt{T}\mathcal{N}(0,1)\right) < q\right]$$

$$\alpha = P\left[-\mathcal{N}(0,1) < \frac{\frac{q}{N} + \mu T}{\sigma\sqrt{T}}\right] = \Phi\left(\frac{\frac{q}{N} + \mu T}{\sigma\sqrt{T}}\right)$$

where we used the fact that $-\mathcal{N}(0,1)$ is still distributed as the standard normal. Then, taking Φ^{-1} on both sides,

$$\Phi^{-1}(\alpha) = \frac{\frac{q}{N} + \mu T}{\sigma\sqrt{T}} \tag{42}$$

and therefore

$$q = N(-\mu T + \sigma\sqrt{T}\Phi^{-1}(\alpha)).$$

Mock Exam 3 VII

This is our $VaR_{T,\alpha}$ for the stock position. To compute ES we need to look at

$$\mathrm{ES}_{T,\alpha} = \frac{\mathbb{E}^{\mathbb{P}}[L_T 1_{\{L_T > \mathrm{VaR}_{T,\alpha}\}}]}{1-\alpha}$$

$$= \frac{\mathbb{E}^{\mathbb{P}}[N(S_0 - S_T) 1_{\{N(S_0 - S_T) > \mathrm{VaR}_{T,\alpha}\}}]}{1-\alpha}$$

$$= \frac{\mathbb{E}^{\mathbb{P}}[N\left(-\mu T - \sigma\sqrt{T}\mathcal{N}(0,1)\right) 1_{\{N\left(-\mu T - \sigma\sqrt{T}\mathcal{N}(0,1)\right) > q\}}]}{1-\alpha} = \dots$$

We can compute the expectation through an integral:

$$\mathbb{E}^{\mathbb{P}}[N\left(-\mu T - \sigma\sqrt{T}\mathcal{N}(0,1)\right) 1_{\{N\left(-\mu T - \sigma\sqrt{T}\mathcal{N}(0,1)\right) > q\}}] =$$

Mock Exam 3 VIII

$$= \int_{-\infty}^{+\infty} \left[N\left(-\mu T - \sigma\sqrt{T}x\right) 1_{\{N(-\mu T - \sigma\sqrt{T}x) > q\}} \right] p_{\mathcal{N}(0,1)}(x) dx =$$

$$= \int_{-\infty}^{+\infty} N\left(-\mu T - \sigma\sqrt{T}x\right) 1_{\{x < (-q - N\mu T)/(N\sigma\sqrt{T})\}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \int_{-\infty}^{(-q - N\mu T)/(N\sigma\sqrt{T})} N\left(-\mu T - \sigma\sqrt{T}x\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= -N\mu T \int_{-\infty}^{\frac{-q-N\mu T}{N\sigma\sqrt{T}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - N\sigma\sqrt{T} \int_{-\infty}^{\frac{-q-N\mu T}{N\sigma\sqrt{T}}} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= -N\mu T \Phi\left(\frac{-q - N\mu T}{N\sigma\sqrt{T}}\right) - N\sigma\sqrt{T} \int_{-\infty}^{(-q - N\mu T)/(N\sigma\sqrt{T})} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= -N\mu T \Phi\left(\frac{-q - N\mu T}{N\sigma\sqrt{T}}\right) + N\sigma\sqrt{T} \int_{-\infty}^{(-q - N\mu T)/(N\sigma\sqrt{T})} \frac{1}{\sqrt{2\pi}} d\left(e^{-\frac{x^2}{2}}\right)$$

Mock Exam 3 IX

$$= -N\mu T \Phi\left(\frac{-q - N\mu T}{N\sigma\sqrt{T}}\right) + N\sigma\sqrt{T}\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right)\Big|_{-\infty}^{\frac{-q-N\mu T}{N\sigma\sqrt{T}}} =$$

$$= -N\mu T \Phi\left(\frac{-q - N\mu T}{N\sigma\sqrt{T}}\right) + N\sigma\sqrt{T}\left(\phi(x)\right)\Big|_{-\infty}^{\frac{-q-N\mu T}{N\sigma\sqrt{T}}} =$$

$$-N\mu T \Phi\left(\frac{-q - N\mu T}{N\sigma\sqrt{T}}\right) + N\sigma\sqrt{T}\phi\left(\frac{-q - N\mu T}{N\sigma\sqrt{T}}\right)$$

where ϕ is the density of the standard normal,
$\phi(x) = p_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ and we used that $\lim_{x \to -\infty} \phi(x) = 0$.
Substituting back we get

$$\ldots = ES_{T,\alpha} = \frac{-N\mu T \Phi\left(\frac{-q-N\mu T}{N\sigma\sqrt{T}}\right) + N\sigma\sqrt{T}\phi\left(\frac{-q-N\mu T}{N\sigma\sqrt{T}}\right)}{1-\alpha}$$

Mock Exam 3 X

Taking into account Equation (42) we can further simplify this last expression into

$$\ldots = \frac{-N\mu T\Phi\left(-\Phi^{-1}(\alpha)\right) + N\sigma\sqrt{T}\phi\left(-\Phi^{-1}(\alpha)\right)}{1-\alpha} = \ldots$$

and using

$$\Phi(-\Phi^{-1}(\alpha)) = 1 - \Phi(\Phi^{-1}(\alpha)) = 1 - \alpha$$

we get

$$\ldots = -N\mu T + \frac{N\sigma\sqrt{T}\phi\left(-\Phi^{-1}(\alpha)\right)}{1-\alpha}$$

Mock Exam 3 XI

d) ES does not completely look at the tail structure of the Loss, but does so only in expectation. So if 99% VaR is 10 billions, we can have the remaining 1% loss concentrated

(i) either on 10.1 billions,

(ii) or on 10 trillions,

as two stylized cases, without VaR being able to tell us anything on whether we are in case (i) or (ii).
ES does a little better than VaR, in that it averages the tail. The average in case (ii) will be much larger than the average of case (i), thus alerting one to more risk in case (ii). Still, it won't tell us exactly how the tail risk looks like or where exactly the loss is concentrated on the tail.

Mock Exam 3 XII

Another problem of ES is that it is homogeneous with respect to the portfolio size. Namely, if k is a positive constant, then
$VaR(k\ Portfolio) = k\ VaR(Portfolio)$
and
$ES(k\ Portfolio) = k\ ES(Portfolio)$.
This is unrealistic and completely neglects liquidity risk. Buying one million of shares is more than one million times risky than buying one share. Placing the order for one million shares will move the whole market and change the share price (theory of maket impact/market microstructure) with potential additional losses due to market impact, whereas placing the order for one share willl not move the market. Liquidity risk strongly disagrees with the homogeneous assumption.

e) No the dynamics is not the same, to price an option we need to use the risk neutral dynamics, where the drift parameter μ of S is replaced

Mock Exam 3 XIII

by the risk free rate $r = 0$ of the bank account. So to price an option we need to use the dynamics we found in point a). To compute value at risk or expected shortfall the dynamics that is relevant up to the risk horizon is the dynamics under P, i.e. the dynamics with drift μ.

Mock Exam 4 I

Problem 1: SDEs. Consider the SDE

$$dX_t = v^2 X_t(1+X_t^2)dt + v(1+X_t^2)dW_t,\ X_0 = x_0.$$

where the initial condition is a deterministic constant and v is a positive real constant.
a) Do not try to prove existence and uniqueness of a solution a priori, invoking an existence/uniqueness theorem. Try to find one explicit solution using purely formal calculus, without worring about explosion or singular points of the solutions. Hint: you may tranform in Stratonovich form and then use separation of variables and / or change of variables.
b) Check a posteriori that the solution you found satisfies the original Ito SDE, again in a purely formal sense.

Mock Exam 4 II

c) Now we are not satisfied with a purely formal approach and we wish to discuss the solution we found. Explain whether this is a satisfactory solution in general and highlight problems.

Mock Exam 4 III

Problem 1: Solutions.
a) Consider

$$dX_t = v^2 X_t(1+X_t^2)dt + v(1+X_t^2)dW_t,\ X_0 = x_0.$$

The diffusion coefficient is $\sigma(x) = v(1+x^2)$. The equivalent Stratonovich SDE is obtained by changing the drift by

$$v^2 x(1+x^2) \to v^2 x(1+x^2) - \frac{1}{2}\sigma(x)\frac{d}{dx}\sigma(x) =$$

$$= v^2 x(1+x^2) - \frac{1}{2}v(1+x^2)\frac{d}{dx}(v(1+x^2)) =$$

$$= v^2 x(1+x^2) - \frac{1}{2}v(1+x^2)2vx = 0.$$

Mock Exam 4 IV

So the equivlent Stratonovich SDE has zero drift,

$$dX_t = v(1+X_t^2) \circ dW_t.$$

Now we know that the Stratonovich SDE obeys the formal rules of calculus. Let's try to solve the SDE formally, by separating variables:

$$\frac{dX_t}{1+X_t^2} = v \circ dW_t.$$

Dividing by $1+X^2$ is not a problem, as this is always strictly positive. Integrate both sides

$$\int_{X_0}^{X_t} \frac{dX}{1+X^2} = \int_0^t v \circ dW_t.$$

Mock Exam 4 V

The Stratonovich integral is the same as an Ito one, as the integrand is a constant. Thus

$$\int_0^t v \circ dW_t = \int_0^t v dW_t = v(W_t - W_0) = vW_t.$$

Recalling the basic integral

$$\int_{x_0}^{x_1} \frac{1}{1+x^2} dx = \arctan(x_1) - \arctan(x_0)$$

where arctan is the inverse of the tangent trigonometric function, we obtain

$$\arctan(X_t) - \arctan(x_0) = vW_t,$$

Mock Exam 4 VI

Note that in general there will not be any guarantee that $\arctan(X_t)$ is between $-\pi/2$ and $\pi/2$, given that the Brownian motion W can take arbitrarily large values. Still, let us proceed formally.

$$\arctan(X_t) = \arctan(x_0) + vW_t.$$

Taking the tangent on both sides we have

$$X_t = \tan\left(\arctan(x_0) + vW_t\right).$$

b) Let's check formally that $X_t = \tan\left(\arctan(x_0) + vW_t\right)$ satisfies the original Ito SDE.

Mock Exam 4 VII

Set $Z_t = \arctan(x_0) + vW_t$, so that $dZ_t = vdW_t$, and write the solution as

$$X_t = \tan(Z_t).$$

Let's differentiate both sides of the last equation above using Ito's formula. Recall that $d\tan(z)/dz = 1/\cos^2(z)$, $d^2\tan(z)/dz^2 = 2\sin(z)/\cos^3(z) = 2\tan(z)/\cos^2(z)$. Then

$$dX_t = \frac{1}{\cos^2(Z)} dZ_t + \frac{1}{2} 2 \frac{\tan(Z_t)}{\cos^2(Z_t)} dZ\, dZ$$

i.e.

$$dX_t = \frac{1}{\cos^2(Z)} v dW_t + \frac{X_t}{\cos^2(Z_t)} v^2 dt$$

Mock Exam 4 VIII

as $\tan Z = X$ and $dZ\, dZ = vdWvdW = v^2\, dt$. To complete our calculations we need to evaluate $\cos^2(Z)$. To this end, note that

$$\tan(Z) = \frac{\sin Z}{\cos Z} \implies \tan^2(Z) = \frac{\sin^2 Z}{\cos^2 Z} = \frac{1-\cos^2 Z}{\cos^2 Z}$$

leading to

$$\tan^2(Z) = \frac{1-\cos^2 Z}{\cos^2 Z} \Rightarrow \cos^2 Z = \frac{1}{1+\tan^2 Z}$$

or

$$\cos^2 Z = \frac{1}{1+\tan^2 Z} = \frac{1}{1+X^2}$$

as $\tan Z = X$. Substituting in

$$dX_t = \frac{1}{\cos^2(Z)} vdW_t + \frac{X_t}{\cos^2(Z_t)} v^2 dt$$

Mock Exam 4 IX

we get

$$dX_t = (1 + X_t^2) v dW_t + X_t (1 + X_t^2) v^2 dt$$

which is our original Ito SDE. We have thus shown that our solution $X_t = \tan\left(\arctan(x_0) + vW_t\right)$ satisfies formally the original Ito SDE.

c) The solution

$$X_t = \tan\left(\arctan(x_0) + vW_t\right)$$

has been derived by ignoring a number of potential problems. While from a purely formal point of view its differential satisfies the Ito SDE given initially, there are problems in claiming this is a valid solution. Assume for simplicity that $x_0 = 0$, so that $\arctan(x_0) = 0$ and

$$X_t = \tan\left(vW_t\right).$$

Mock Exam 4 X

The tangent function is not defined when the argument of the function is $\pi/2 + n\pi$ for any integer n. The solution X_t above is *not defined* in the set

$$B_t = \left\{ \omega \in \Omega : \ vW_s(\omega) = \frac{\pi}{2} + n\pi, \text{ for some } n \in \mathbb{Z} \text{ and } s \le t \right\}.$$

Our solution will explode before time t in the set B_t, at the first time s where vW_s hits $\pi/2$ plus or minus integer multiples of π, where the tangent diverges to ∞. We conclude we have not found a proper solution avoiding explosion in finite time. There could still be a possibility that the set B_t has zero probability, so that the SDE has a solution that exists almost surely, in a set of probability 1. We show now that this is not the case.

Mock Exam 4 XI

We can try to compute the probability of explosion. Write B_t as

$$B_t = \left\{ \omega \in \Omega : \;\; W_s(\omega) = \frac{\frac{\pi}{2} + n\pi}{v}, \text{ for some } n \in \mathbb{Z} \text{ \& some } s \le t \right\}.$$

Write B_t as a union of sets for each n:

$$B_t = \bigcup_{n \in \mathbb{Z}} \left\{ \omega \in \Omega : \;\; W_s(\omega) = \frac{\frac{\pi}{2} + n\pi}{v}, \text{ for some } s \le t \right\} =: \cup_{n \in \mathbb{Z}} A_n$$

For a fixed natural number $\bar{n}$ we have that

$$A_{\bar{n}} \subset \cup_{n \in \mathbb{Z}} A_n = B_t \;\Rightarrow\; P[B_t] \ge P[A_{\bar{n}}].$$

If we prove that $A_{\bar{n}}$ has strictly positive probability, B_t will have strictly positive probability too and so explosion will have a positive probability.

Mock Exam 4 XII

$$P[A_{\bar{n}}] = P\left\{\omega \in \Omega : \ W_s(\omega) = \frac{\frac{\pi}{2} + \bar{n}\pi}{v}, \text{ for some } s \le t\right\}$$

$$= P\left\{\omega \in \Omega : \max_{s\in[0,t]} W_s(\omega) \ge \frac{\frac{\pi}{2} + \bar{n}\pi}{v}\right\}$$

where we have used the fact that W_t, starting from $W_0 = 0$ at time 0, will hit $\pi/2 + \bar{n}\pi > 0$ ($\bar{n} \ge 0$) from below at a time $s \le t$ if and only if $\max W_s$ for $s \in [0, t]$ is above $\pi/2 + \bar{n}\pi$. Without further tools we cannot compute this probability, but we can argue that it will be positive, as it is the probability that a continuous random variable taking values in $(-\infty, +\infty)$ is larger than a given real number.

Mock Exam 4 XIII

This calculation below is beyond the scope of our course and would not be given at an exam, but it is possible to compute the law of $\max_{s\in[0,t]} W_s(\omega)$ using a reflection principle. The above probability is

$$P\left[\omega \in \Omega : |W_t| \geq \frac{\frac{\pi}{2} + \bar{n}\pi}{v}\right]$$

and this probability is strictly positive and can be computed using the Gaussian law of Brownian motion. We thus see that there is a strictly positive probability that our formal solution explodes in finite time.

Explosion could not be excluded a priori, because the sufficient conditions guaranteeing existence and uniqueness are violated. Indeed, our drift and diffusion coefficients have more than linear growth. The drift has cubic growth and the diffusion coefficient has

Mock Exam 4 XIV

quadratic growth. This means we could not apply our theorem from global existence and uniqueness, so we could not guarantee existence and uniqueness a priori.

This problem shows that it is not enough to solve a SDE formally. To find a real solution, one needs to check that the solution exists and is unique, without eplosions or other problems.

Mock Exam 4 I

Problem 2: Option pricing and no arbitrage.
Consider the Black and Scholes basic economy given by a bank account and a stock, whose prices are given respectively by

$$dB_t = rB_t dt, \;\; B_0 = 1, \quad dS_t = \mu S_t dt + \sigma S_t dW_t, \; S_0 = 1$$

where r, μ, σ are positive constants and W is a brownian motion under the physical measure P.

Consider two options that are at the money forward, namely having strike

$$K = S_0 e^{rT},$$

respectively a call option and a put option with maturity T and payout

$$(S_T - K)^+ = max(S_T - K, 0) \text{ (Call)}, \;\; (K - S_T)^+ = max(K - S_T, 0) \text{ (Put)}.$$

Mock Exam 4 II

In this case, put call parity tells us that the initial value of the call option, at time 0, must be equal to the initial value of the put option, given that the forward contract value is zero (K is the at-the-money forward strike that sets the forward price to zero).

a) Show that if this put-call parity condition is violated at time 0,

$$\text{CallPrice}_0 = \text{PutPrice}_0 + X$$

for a positive amount $X > 0$, even if the condition holds later for every $t > 0$,

$$\text{CallPrice}_t = \text{PutPrice}_t, \quad \text{for} \quad t > 0$$

then one has arbitrage.

b) Does the existence of arbitrage in point a) require the Black Scholes model or is it more general?

Mock Exam 4 III

c) If put-call parity is violated at time 0 in the opposite direction, namely

$$\text{CallPrice}_0 = \text{PutPrice}_0 - X$$

with positive X, but again holds at times $t > 0$, show that we still have arbitrage.

Mock Exam 4 IV

SOLUTION.
a) Since the price of the call is larger than the one of the put when they should actually be the same, we can try by buying a put and short-selling a call, and buying an at the money forward contract to balance put minus call at maturity. We also buy some bank account with the difference between the call and the put at time 0.

We enter into one position in a put option at time 0, and short sell one call option at time 0, both options with strike K and maturity T. We also enter into a forward contract at the same maturity with strike K. This means that we accept to receive $S_T - K$ at maturity (meaning that if this quantity is positive we receive it, if it is negative we pay its absolute value to our counterparty in the trade). We also buy an amount X of bank account.

The cost of starting this strategy is:

Mock Exam 4 V

- We pay PutPrice_0 to enter into the put option
- we receive $\text{CallPrice}_0 = \text{PutPrice}_0 + X$ by short-selling the Call option
- We pay X to buy a quantity X of bank account B_0 at time 0.
- We pay nothing to enter into the forward contract since its initial cost is $S_0 - Ke^{-rT} = 0$.

These four operations have a total cost of

$$\underbrace{\text{PutPrice}_0}_{buy\ put\ option} \underbrace{-\text{CallPrice}_0}_{sell\ call\ option} + \underbrace{X}_{Buy\ X\ bank\ account} + \underbrace{0}_{enter\ fwd\ contract} = 0$$

where we used our assumption that $\text{CallPrice}_0 = \text{PutPrice}_0 - X$.
So it costs nothing setting up the strategy.

Mock Exam 4 VI

Following the initial setup, we just wait. This clearly preserves the self financing condition a fortiori, since we do not inject external funds or extracts funds from the strategy. We also need to check that the strategy value is non-negative at all times as this is part of the definition of self-financing. As we know that put-call parity holds for $t > 0$, we can write the value of our strategy at t as

$$V_t = V_t^{Put} - V_t^{Call} + V_t^{forward} + Xe^{rt},$$

and because put call parity holds at $t > 0$, we have $V_t^{Call} - V_t^{Put} = V_t^{forward}$ or $V_t^{Put} - V_t^{Call} + V_t^{forward} = 0$ and substituting in the value of the strategy we get

$$V_t = V_t^{Put} - V_t^{Call} + V_t^{forward} + Xe^{rt} = Xe^{rt} > 0,$$

so as requested the value of the strategy is nonnegative for all $t \geq 0$.

Mock Exam 4 VII

At maturity, we have the following cash flows:

- We receive $(K - S_T)^+$ from the Put option.
- We pay $(S_T - K)^+$ for the Call option we have been short-selling
- We receive $S_T - K$ from the forward contract
- We have Xe^{rT} in the bank account.

The total value of this strategy at T is hence

$$\text{PutPayout}_T - \text{CallPayout}_T + \text{FwdContractPayout}_T + XB_T =$$
$$(K - S_T)^+ - (S_T - K)^+ + S_T - K + Xe^{rT} =$$
$$= K - S_T + S_T - K + Xe^{rT} = Xe^{rT} > 0$$

So we have a self-financing trading strategy whose initial cost is zero and that produces a positive final cash flow Xe^{rT} in all scenarios. Hence this is an arbitrage opportunity and the market is arbitrageable.

Mock Exam 4 VIII

b) In the reasoning above, given that we are in the Black Scholes model, the prices of the call and the put would be the Black Scholes prices. However, this is used nowhere in the proof of arbitrage. Indeed, the proof is completely general and would hold under any other model such as Bachelier, displaced diffusion, CEV etc. This reflects the fact that put-call parity is model-independent, as it holds for all models, and thus is not bound by the Black Scholes model.

c) Since the price of the call is smaller than the one of the put when they should actually be the same, we can try by buying a call and short-selling a put, and short–selling an at the money forward contract to balance call minus put at maturity. We also buy some bank account with the difference between the put and the call at time 0.

Mock Exam 4 IX

We enter into one position in a call option at time 0, and short sell one put option at time 0, both options with strike K and maturity T. We also short-sell a forward contract at the same maturity with strike K. This means that we will pay the forward payoff $S_T - K$ at maturity (meaning that if this quantity is positive we pay it, if it is negative we receive its absolute value from our counterparty in the trade). We also buy an amount X of bank account.

The cost of starting this strategy is:

- We pay CallPrice_0 to enter into the call option
- we receive $\text{PutPrice}_0 = \text{CallPrice}_0 + X$ by short-selling the Put option
- We pay X to buy a quantity X of bank account B_0 at time 0.

Mock Exam 4 X

- We pay or receive nothing to short-sell the forward contract since its initial cost is $S_0 - Ke^{-rT} = 0$.

These four operations have a total cost of

$$\underbrace{\text{CallPrice}_0}_{buy\ call\ option} \underbrace{-\text{PutPrice}_0}_{sell\ put\ option} + \underbrace{X}_{Buy\ X\ bank\ account} + \underbrace{0}_{enter\ fwd\ contract} = 0$$

So it costs nothing setting up the strategy.
Following the initial setup, we just wait. This clearly preserves the self financing condition since we do not inject external funds or extracts funds from the strategy. Also, using the fact that put-call parity holds at times $t > 0$, we show easily that the value of the strategy at any $t > 0$ is Xe^{rt}, very similarly to case a), and thus is always non-negative.

At maturity, we have the following cash flows:

Mock Exam 4 XI

- We receive $(S_T - K)^+$ from the call option.
- We pay $(K - S_T)^+$ for the put option we have been short-selling
- We pay $S_T - K$ for the short-selling of the forward contract
- We have Xe^{rT} in the bank account.

The total value of this strategy at T is hence

$$\text{CallPayout}_T - \text{PutPayout}_T - \text{FwdContractPayout}_T + XB_T =$$
$$(S_T - K)^+ - (K - S_T)^+ - (S_T - K) + Xe^{rT} =$$
$$= S_T - K - (S_T - K) + Xe^{rT} = Xe^{rT} > 0$$

So we have a self-financing trading strategy whose initial cost is zero and that produces a positive final cash flow Xe^{rT} in all scenarios. Hence this is an arbitrage opportunity and the market is arbitrageable.

Mock Exam 4 I

Problem 3: Smile modeling - butterfly spread
Consider a butterfly spread option on a stock S with maturity T and initial value S_0. This is a strategy based on buying one in-the-money call option with a low strike price $L = S_0 - X$, selling two at-the-money call options with a middle strike $M = S_0$, and buying one out-of-the-money call option with a higher strike price $H = S_0 + X$, where X is a positive constant. All options have maturity T.
a) Write the payoff of this product and draw it as a function of S_T.
b) Who would buy this product? What are the views on the stock for a client buying this product?
c) Compute the price of this product in a Black Scholes market where the stock follows the dynamics

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 = s_0$$

Mock Exam 4 II

with a deterministic positive s_0 and where the risk free rate r is assumed to be zero.
d) Compute the butterfly delta, namely the sensitivity of the butterfly price with respect to the initial stock price S_0.
e) Compute the butterfly price in a Bachelier model $dS_t = v dW_t^Q$, $S_0 = s_0$.
f) Compute the delta of the butterfly price, namely the sensitivity of the price with respect to S_0, in the Bachelier model.
g) Consider the limit situation when $X \downarrow 0$. Deduce intuitively the price of the butterfly by looking at the plot of the payoff and thinking what happens to the plot when $X \downarrow 0$. Check with a calculation for $X \downarrow 0$ whether the Black and Scholes price and the Bachelier price for the butterfly confirm this intuition.

Solutions.

Mock Exam 4 III

a) The total payoff at maturity T is thus

$$Y = (S_T - L)^+ - 2(S_T - S_0)^+ + (S_T - H)^+.$$

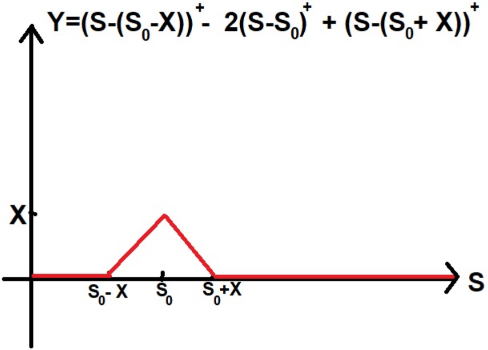
Let us write the payoff looking at different cases.

(i) If $S_T < L$ then all three options expire worthless and $Y = 0$.
(ii) If $L < S_T < S_0$ then the first option pays $S_T - L$ but all other options expire worthless, so $Y = S_T - L = S_T - (S_0 - X)$.
(iii) If $S_0 < S_T < H$ then the first option pays us $S_T - L$, the short selling of the 2 at the money options will require us to pay $2(S_T - S_0)$, while the last option expires worthless, so
$Y = S_T - L - 2(S_T - S_0) = 2S_0 - L - S_T = S_0 + X - S_T$.

Mock Exam 4 IV

(iv) If $S_T > H$ all options have positive payoff, and
$Y = S_T - L - 2(S_T - S_0) + S_T - H = 2S_0 - L - H =$
$S_0 - L + S_0 - H = X - X = 0.$

We can thus draw the plot as



Mock Exam 4 V

b) From the plot we see that the payoff is non-zero only when the final stock S_T is between $S_0 - X$ and $S_0 + X$. The maximum payoff is X, and is achieved when $S_T = S_0$. Therefore, the investor holding the option has a view that the stock will not move much away from S_0, and ideally will stay at S_0 (thus expecting low volatility). This will ensure that the final payoff X is obtained and is larger than the initial premium paid for the butterfly, ensuring a substantial profit. The minimum payoff is zero, and this happens when the final stock value S_T is outside the interval $(S_0 - X,\ S_0 + X)$. Therefore an investor with a view that the stock will move away more than X from its current value S_0, or close to X, in either direction (thus expecting large volatility), will not purchase this product as they would pay a premium at time 0 but gain nothing at maturity T. We further note that this payoff does not allow for a potentially unlimited profit like a call option or a straddle, so the investor cannot aim at a potentially unlimited profit. Also, the payoff exposes

Mock Exam 4 VI

the client buying it to no potential loss, except for the initial price paid to purchase it, and therefore may attract risk averse investors who are not willing to lose anything more than the initially paid price. In addition, this payoff will not be too expensive to purchase, as the price of the two calls that are bought is compensated by the price of the two calls that are sold. Again, this may attract risk averse investors.

c) As

$$Y = (S_T - L)^+ - 2(S_T - S_0)^+ + (S_T - H)^+,$$

it follows that

$$e^{-rT} Y = e^{-rT}(S_T - L)^+ - 2e^{-rT}(S_T - S_0)^+ + e^{-rT}(S_T - H)^+$$

Mock Exam 4 VII

and

$$E^Q[e^{-rT}Y] = E^Q[e^{-rT}(S_T - L)^+] - 2E^Q[e^{-rT}(S_T - S_0)^+] + E^Q[e^{-rT}(S_T - H)$$

or, in other terms,

$$V_{BS}^{Butter}(0) = V_{BS}^{Call}(0, K = L) - 2V_{BS}^{Call}(0, K = S_0) + V_{BS}^{Call}(0, K = H).$$

Recalling that $L = S_0 - X,\ H = S_0 + X$ and $r = 0$, and using the formulas for Call options in Black scholes in this special case, we get

Mock Exam 4 VIII

$$V_{BS}^{Butter}(0) = S_0\Phi\left(\frac{\ln\frac{S_0}{S_0-X}+\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) - (S_0-X)\Phi\left(\frac{\ln\frac{S_0}{S_0-X}-\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) +$$

$$-2S_0\Phi\left(\frac{\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) + 2S_0\Phi\left(\frac{-\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) +$$

$$+S_0\Phi\left(\frac{\ln\frac{S_0}{S_0+X}+\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) - (S_0+X)\Phi\left(\frac{\ln\frac{S_0}{S_0+X}-\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right).$$

The middle d_1 and d_2 terms can be simplified.

Mock Exam 4 IX

d) Recall the delta of a call option in Black Scholes:
$\Delta = \frac{\partial}{\partial S_0} V_{BS}^{Call}(0) = \Phi(d_1)$. Given that, from point c),

$$V_{BS}^{Butter}(0) = V_{BS}^{Call}(0, K = L) - 2V_{BS}^{Call}(0, K = S_0) + V_{BS}^{Call}(0, K = H)$$

we have

$$\frac{\partial}{\partial S_0} V_{BS}^{Butter} = \frac{\partial}{\partial S_0} V_{BS}^{Call}(K = L) - 2\frac{\partial}{\partial S_0} V_{BS}^{Call}(K = S_0) + \frac{\partial}{\partial S_0} V_{BS}^{Call}(K = H)$$

$$= \Phi\left(\frac{\ln \frac{S_0}{S_0 - X} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) - 2\Phi\left(\frac{\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) + \Phi\left(\frac{\ln \frac{S_0}{S_0 + X} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right)$$

where the middle term can be simplified further.

Mock Exam 4 X

e) We recall the price of a call option in the Bachelier model:

$$V^{Call}_{BaM}(0, s_0, K, T, \sigma) = (s_0 - K)\Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}\phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right),$$

where ϕ is the pdf of a standard normal. From our previous points, the payoff Y of a butterfly satisfies

$$E^Q[e^{-rT}Y] = E^Q[e^{-rT}(S_T-L)^+]-2E^Q[e^{-rT}(S_T-M)^+]+E^Q[e^{-rT}(S_T-H)^+$$

or, if S follows the Bachelier model,

$$V^{Butter}_{BaM} = V^{Call}_{BaM}(K = L) - 2V^{Call}_{BaM}(K = M) + V^{Call}_{BaM}(K = H), \qquad (43)$$

or, recalling $r = 0$ and $L = S_0 - X,\ M = S_0,\ H = S_0 + X$

Mock Exam 4 XI

$$
\begin{gathered}
V_{BaM}^{Butter} = X\Phi\left(\frac{X}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}\phi\left(\frac{X}{\sigma\sqrt{T}}\right) \\
-2\sigma\sqrt{T}\phi\,(0) + \\
-X\Phi\left(\frac{-X}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}\phi\left(\frac{-X}{\sigma\sqrt{T}}\right)
\end{gathered}
$$

Note that $\phi(0) = 1/\sqrt{2\pi}$.
Recalling that $\Phi(-x) = 1 - \Phi(x)$ and that $\phi(-x) = \phi(x)$ we get

$$
V_{BaM}^{Butter} = X\left(2\Phi\left(\frac{X}{\sigma\sqrt{T}}\right) - 1\right) - 2\sigma\sqrt{T}\left(\phi(0) - \phi\left(\frac{X}{\sigma\sqrt{T}}\right)\right) \tag{44}
$$

Mock Exam 4 XII

Note that, as X is positive, $\Phi\left(\frac{X}{\sigma\sqrt{T}}\right)$ is calculated in a positive point and is therefore larger than $\frac{1}{2}$, since $\Phi(0) = 1/2$ and Φ is strictly increasing. It follows that $2\Phi\left(\frac{X}{\sigma\sqrt{T}}\right) - 1$ is larger than zero, or that the first term in the V_{BaM} formula above is positive.
As concerns ϕ, the Gaussian standard pdf, this has a maximum in 0, so that the second term in the V_{BaM} formula above being subtracted is also positive. The price is thus the difference of two positive terms related to the Gaussian CDF and PDF respectively:

$$V_{BaM}^{Butter} = \underbrace{X\left(2\Phi\left(\frac{X}{\sigma\sqrt{T}}\right) - 1\right)}_{positive} - \underbrace{2\sigma\sqrt{T}\left(\phi(0) - \phi\left(\frac{X}{\sigma\sqrt{T}}\right)\right)}_{positive}.$$

Mock Exam 4 XIII

f) Careful here. You might be tempted to differentiate directly Formula (44) with respect to s_0 and conclude that the delta is 0, as Eq. (44) does not depend on s_0. However, as explained in the lectures for Black Scholes and as seen in some previous mock exams, if the strike depends on s_0, to calculate the delta you need to differentiate the price with a general K and susbstitute the specific K that is a function of s_0 after you differentiate, not before. This is because we want the sensitivity to s_0 as initial stock and not the joint sensitivity to s_0 and K. Thus, the equation you need to differentiate with respect to s_0 is (43), leaving the strikes L, M, H as generic and not as functions of s_0 yet. We recall the delta for a call with strike K in the Bachelier model from Eq. (32), if you don't remember it derive it as in the derivation of Eq. (32).

Mock Exam 4 XIV

$$\frac{\partial V_{BaM}(0)}{\partial s_0} = \frac{\partial \left((s_0 - K)\Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)\right)}{\partial s_0} + \sigma\sqrt{T}\frac{\partial\, p_N\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)}{\partial s_0} =$$

$$= \Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right).$$

The butterfly spread payoff is call with strike L minus 2 calls with strike M plus a call with strike H. The delta is thus obtained by differentiating the three call prices in Eq (43) and is

$$\frac{\partial V_{BaM}^{Butter}}{\partial s_0} = \Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right)|_{K=L} - 2\Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right)|_{K=M} + \Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right)|_{K=H} =$$

$$\Phi\left(\frac{s_0 - L}{\sigma\sqrt{T}}\right) - 2\Phi\left(\frac{s_0 - s_0}{\sigma\sqrt{T}}\right) + \Phi\left(\frac{s_0 - H}{\sigma\sqrt{T}}\right)$$

Mock Exam 4 XV

$$= \Phi\left(\frac{s_0 - (s_0 - X)}{\sigma\sqrt{T}}\right) - 2\Phi(0) + \Phi\left(\frac{s_0 - (s_0 + X)}{\sigma\sqrt{T}}\right)$$

$$= \Phi\left(\frac{X}{\sigma\sqrt{T}}\right) - 2\frac{1}{2} + \Phi\left(\frac{-X}{\sigma\sqrt{T}}\right)$$

$$= \Phi\left(\frac{X}{\sigma\sqrt{T}}\right) - 1 + 1 - \Phi\left(\frac{X}{\sigma\sqrt{T}}\right) = 0$$

where we used $\Phi(-x) = 1 - \Phi(x)$. So we conclude the delta is zero:

$$\frac{\partial}{\partial S_0} V_{BaM}^{Butter} = 0.$$

We obtained the same result as we would obtained with the wrong method, namely by differentiating Eq. (44), but this is a coincidence. You need to avoid differentiating Eq. (44) directly, even if in this case you would have obtained the same result.

Mock Exam 4 XVI

Zero delta means that the butterfly has always the same price in Bachelier, regardless of the initial stock price S_0. This is due to the special dynamics of the model and to the specific choice of strikes $L = S_0 - X$ and $H = S_0 + X$.

g) Looking at the payoff we see that when $X \downarrow 0$ the payoff tends to be zero everywhere. As such, it will be worth 0 in terms of initial price. The BS and Bachelier prices are continuous in X around $X = 0$, so we can set $X = 0$ direcly in the formulas to compute the limit for $X \downarrow 0$.

$$V_{BS}^{Butter}(X = 0) = S_0 \Phi\left(\frac{\ln \frac{S_0}{S_0} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) - S_0 \Phi\left(\frac{\ln \frac{S_0}{S_0} - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) +$$

Mock Exam 4 XVII

$$-2S_0\Phi\left(\frac{\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) + 2S_0\Phi\left(\frac{-\frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) +$$

$$+S_0\Phi\left(\frac{\ln\frac{S_0}{S_0} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) - S_0\Phi\left(\frac{\ln\frac{S_0}{S_0} - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right).$$

Now all logarithms reduce to $\ln 1 = 0$ and the three terms cancel, so that the price is indeed 0.

In the Bachelier model

$$V_{BaM}^{Butter}(X=0) = 0\left(2\Phi\left(\frac{0}{\sigma\sqrt{T}}\right) - 1\right) - 2\sigma\sqrt{T}\left(\phi(0) - \phi\left(\frac{0}{\sigma\sqrt{T}}\right)\right)$$

$$= (1-1) - 2\sigma\sqrt{T}(\phi(0) - \phi(0)) = 0$$

again, as expected.

Mock Exam 4 I

Consider a porfolio with a zero-coupon bond with notional N and maturity T, and a short position on an amount N of equity forward contract on stock S with strike K and maturity T. In other terms, the payoff at time T is

$$Y = N\,1 - N(S_T - K).$$

The stock is assumed to follow the Black Scholes model

$$dS_t = \mu S_t dt + \sigma S_t dW_t,\ s_0$$

under the measure P. We assume a constant positive risk free rate $r > 0$.

a) Compute the portfolio $VaR_{H,\alpha}$

b) How sensitive is VaR to the stock volatility? Give a quantitative measure of this sensitivity and comment on its sign in the particular case where $\alpha - = 0.95$ and H is either 1 day or 1 year.

Mock Exam 4 II

c) What is the limit VaR when (i) $\sigma \downarrow 0$ and (ii) $\sigma \uparrow +\infty$? Examine how both limits depend on the confidence level α and discuss.

Mock Exam 4 III

Solutions.
a) The loss distribution at H is the portfolio value at time 0 minus the portfolio value at time H, namely

$$L_H = Ne^{-r(T-0)} - N(S_0 - Ke^{-r(T-0)}) - [Ne^{-r(T-H)} - N(S_H - Ke^{-r(T-H)})]$$

$$= N(K+1)(e^{-rT} - e^{-r(T-H)}) - NS_0 + NS_H = A + NS_H$$

where $A = N(K+1)(e^{-rT} - e^{-r(T-H)}) - NS_0$ is deterministic. Note that $A < 0$ as $e^{-rT} < e^{-r(T-H)}$. Here we used the fact that the price of a unit-notional zero coupon bond with maturity T at time $t < T$ is

$$P(t,T) = E_t^Q[e^{-r(T-t)}1] = e^{-r(T-t)}$$

while the price of the forward contract with maturity T at time $t < T$ is

$$E_t^Q[e^{-r(T-t)}(S_T - K)] = S_t - Ke^{-r(T-t)}.$$

Mock Exam 4 IV

$\text{VaR}_{H,\alpha}$ is defined as the quantity q satisfying

$$\alpha = P[L_H < q] = P[A + NS_H < q] = P[S_H < (q - A)/N] =$$

$$= P[S_0 \exp((\mu - \sigma^2/2)H + \sigma W_H) < (q - A)/N] =$$

$$= P[\exp((\mu - \sigma^2/2)H + \sigma W_H) < (q - A)/(NS_0)] = \dots$$

where we have used the usual solution for the Black Scholes geometric-Brownian-motion SDE. As $A < 0$ we have $-A > 0$. We assume $q - A > 0$ so we can take logs on both sides in the last P expression. We have

$$\dots = P\left[W_H < \frac{\ln[(q - A)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma}\right] =$$

Mock Exam 4 V

$$= P\left[\sqrt{H}\mathcal{N} < \frac{\ln[(q-A)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma}\right]$$

$$= P\left[\mathcal{N} < \frac{\ln[(q-A)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma\sqrt{H}}\right]$$

$$\alpha = \Phi\left(\frac{\ln[(q-A)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma\sqrt{H}}\right).$$

Take Φ^{-1} on both sides:

$$\Phi^{-1}(\alpha) = \frac{\ln[(q-A)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma\sqrt{H}},$$

and solve in q:

$$q = A + NS_0 \exp\left(\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2)H\right).$$

Mock Exam 4 VI

This is our VaR. Recall we had assumed $q - A > 0$ to be positive. Let's check it is positive indeed. $q - A$, from the last expression, turns out to be an exponential and as such it is always positive.

b) We see that VaR depends on the volatility through the first and second terms inside exp:

$$VaR_{H,\alpha} = A + NS_0 \exp\left(\Phi^{-1}(\alpha)\boxed{\sigma}\sqrt{H} + (\mu - \boxed{\sigma^2}/2)H\right).$$

To quantify the sensitivity of VaR to σ we compute

$$\frac{\partial VaR}{\partial\sigma} = NS_0 \exp\left(\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \frac{\sigma^2}{2})H\right)\frac{\partial}{\partial\sigma}\left(\Phi^{-1}(\alpha)\sigma\sqrt{H} - \frac{\sigma^2 H}{2}\right)$$

$$= NS_0(\Phi^{-1}(\alpha)\sqrt{H} - \sigma H)\exp\left(\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2)H\right).$$

Mock Exam 4 VII

We investigate the sign of this sensitivity.
The sensitivity will be positive if we have

$$\Phi^{-1}(\alpha)\sqrt{H} - \sigma H > 0 \iff \sigma < \Phi^{-1}(\alpha)/\sqrt{H}.$$

Let's consider $\alpha = 0.95$, corresponding to $\Phi^{-1}(\alpha) \approx 1.65$. The holding period for VaR in years can be one day $H = 1/252$ or one year $H = 1$. In the two cases we have

$$H = 1/252, \ \alpha = 0.95 \Rightarrow \sigma < 26.19.$$

In this case VaR will increase with σ as long as $\sigma < 26.19$ circa, and will decrease otherwise. Volatility of 2619% is really unplausibly huge, so this is a case where only an increasing pattern will show, because the condition will always be true in practice.

Mock Exam 4 VIII

$$H = 1, \ \alpha = 0.95 \Rightarrow \sigma < 1.65.$$

This will almost always be true, as volatilties above 165% are extremely rare. So in this case VaR will always be increasing with σ in practice. It seems that if the holding period H in VaR is long, then we tend to have that VaR will increase with volatility.

c) (i) VaR is continuous in σ. We can therefore see what happens for $\sigma \downarrow 0$ by simply setting $\sigma = 0$ in VaR. We get

$$VaR_{H,\alpha}|_{\sigma=0} = A + NS_0 \exp(\mu H).$$

Note that for $\sigma = 0$ $VaR_{H,\alpha}$ does not depend on the confidence level α anymore. When $\sigma = 0$ VaR is the same for all confidence levels. This

Mock Exam 4 IX

is because there is no risky asset left: with $\sigma = 0$, the stock becomes risk-free.
(ii) When $\sigma \uparrow +\infty$ we get

$$VaR_{H,\alpha}|_{\sigma\uparrow\infty} = \lim_{\sigma\uparrow\infty} \left[A + NS_0 \exp\left(\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2)H\right)\right] = \dots$$

Consider

$$\lim_{\sigma\uparrow\infty} \Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2) =$$

$$= \lim_{\sigma\uparrow\infty} \sigma^2 \left[\frac{\Phi^{-1}(\alpha)\sqrt{H}}{\sigma} + \left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)\right] =$$

$$= \lim_{\sigma\uparrow\infty} \sigma^2 \left[-\frac{1}{2}\right] = -\infty.$$

Mock Exam 4 X

So the total limit is

$$\ldots = A + NS_0 \lim_{\sigma\uparrow\infty} \exp\left(\sigma^2\left[-\frac{1}{2}\right]\right) = A$$

since the exponent in the exponential tends to $-\infty$ and $e^{-\infty}$ tends to zero. Hence the limit is $A = N(K+1)(e^{-rT} - e^{-r(T-H)}) - NS_0$. This limit does not depend on α either. So when the volatility goes to infinity, meaning that the riskyness of the risky asset S goes to infinity, $VaR_{H,\alpha}$ does not depend on α anymore. This is because the infinity risk makes all confidence levels the same, as the volatility is infinite anyway.

Mock Exam 5 I

Problem 1. SDEs - OU Process.
Consider the Ornstein Uhlenbeck (OU) SDE

$$dX_t = (b(t) - a(t)X_t)dt + \sigma(t)dW_t,$$

with $X_0 = x_0$ deterministic and where b, a, σ are smooth deterministic functions of time, with $|a(t)|, |b(t)|$ and $|\sigma(t)|$ all bounded below K for all $t \geq 0$, with K a positive real constant.

a) Prove that this SDE admits a unique global solution.

b) Calculate the solution using Stratonovich calculus.

c) Calculate the solution without using Stratonovich calculus. [Hint: Set $Y_t = \exp(\int_0^t a(s)ds)X_t$ and work with Y]

d) Calculate the expected value of the solution at time $T > 0$.

e) Is it correct to say that the distribution of X_t is Gaussian for every $t > 0$? What is the intuition behind your answer?

Mock Exam 5 II

Problem 1: Solutions.
a) We use the theorem giving sufficient conditions for global existence and uniqueness of solutions for SDEs.
We know from the theory that for the SDE
$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \;\; X_0 = Z$ with Z independent of $\sigma(\{W_t, \; t \leq T)\}$ and $\mathbb{E}[Z^2] < +\infty$, and with $\mu : [0, T] \times \mathbb{R} \to \mathbb{R}$ (the drift) and $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ (the diffusion coefficient) being measurable, if we have global Lipschitz continuity

$$|\mu(t,x)-\mu(t,y)|+|\sigma(t,x)-\sigma(t,y)| \leq K|x-y| \text{ for all } t \in [0,T] \text{ and all } x \in \mathbb{R}$$

and linear growth

$$|\mu(t,x)| + |\sigma(t,x)| \leq K'(1+|x|)| \;\text{ for all } t \in [0,T] \text{ and all } x \in \mathbb{R}$$

for two constants K, K', then our SDE has a unique global solution X_t.

Mock Exam 5 III

Let's check these conditions.
The initial condition has to be squared integrable, $E[X_0^2] < +\infty$, which is true in our case as $X_0 = x_0$ is a finite deterministic constant and $E[X_0^2] = x_0^2 < \infty$. Then we need to prove that the the drift and diffusion coefficient are measurable functions of X, t.
This is trivially true as the drift is a linear affine function of X and smooth in t, $\mu(t, X) = b(t) - a(t)X$, as a and b are smooth, and continuous functions are measurable. Also, the diffusion coefficient $\sigma(t, X) = \sigma(t)$ is trivially measurable as it is a deterministic and continous (as it is smooth) function of t.
Next we need to check the Lipschitz continuity and linear growth condition. The Lipschitz condition reads

$$|\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| =$$

Mock Exam 5 IV

$$= |b(t) - a(t)x - (b(t) - a(t)y)| + |\sigma(t) - \sigma(t)| = |a(t)||x - y|,$$

and recalling that $|a(t)| \le K$, we conclude. The Lipschitz condition is satisfied. As for linear growth,

$$|\mu(t,x)| + |\sigma(t,x)| = |b(t) - a(t)x| + |\sigma(t)|$$

$$\le 2K(1 + |x|) \text{ for all } t \in [0,T] \text{ and all } x \in \mathbb{R}$$

as $|a(t)|, |b(t)|$ and $|\sigma(t)|$ are bounded by K.
So we have a unique global solution.

b) Calculation of the solution using Stratonovich calculus was given in the lecture notes in the “OU Process” example in Part One.

Mock Exam 5 V

c) To calculate the solution without using Stratonovich let's use the hint. Calculate

$$dY_t = d\left(\exp\left(\int_0^t a(s)ds\right)\ X_t\right) =$$

$$= X_t d\exp\left(\int_0^t a(s)ds\right) + \exp\left(\int_0^t a(s)ds\right) dX_t =$$

$$= X_t \exp\left(\int_0^t a(s)ds\right) d\left(\int_0^t a(s)ds\right)$$

$$+ \exp\left(\int_0^t a(s)ds\right) ((b(t) - a(t)X_t)dt + \sigma(t)dW_t) =$$

$$= X_t \exp\left(\int_0^t a(s)ds\right) a(t)dt$$

Mock Exam 5 VI

$$+\exp\left(\int_0^t a(s)ds\right)((b(t)-a(t)X_t)dt+\sigma(t)dW_t)$$

$$=\exp\left(\int_0^t a(s)ds\right)b(t)dt+\exp\left(\int_0^t a(s)ds\right)\sigma(t)dW_t.$$

Thus

$$dY_t=\exp\left(\int_0^t a(s)ds\right)b(t)dt+\exp\left(\int_0^t a(s)ds\right)\sigma(t)dW_t.$$

This is a very easy SDE to integrate, as Y is not on the right hand side. We simply integrate both sides between 0 and T:

$$Y_T-Y_0=\int_0^T\exp\left(\int_0^t a(s)ds\right)b(t)dt+\int_0^T\exp\left(\int_0^t a(s)ds\right)\sigma(t)dW_t.$$

Mock Exam 5 VII

Recalling that $Y_t = \exp\left(\int_0^t a(s)ds\right) X_t$ for all $t > 0$ and substituting this in the last equation above for Y_T we get:

$$\exp\left(\int_0^T a(s)ds\right) X_T - X_0 = \int_0^T \exp\left(\int_0^t a(s)ds\right) b(t)dt$$

$$+ \int_0^T \exp\left(\int_0^t a(s)ds\right) \sigma(t)dW_t.$$

Now multiply both sides for $\exp\left(-\int_0^T a(s)ds\right)$ to get

Mock Exam 5 VIII

$$X_T - e^{-\int_0^T a(s)ds} X_0$$
$$= e^{-\int_0^T a(s)ds} \left[\int_0^T e^{\int_0^t a(s)ds} b(t)dt + \int_0^T e^{\int_0^t a(s)ds} \sigma(t) dW_t \right].$$

leading to

$$X_T = e^{-\int_0^T a(s)ds} \left[X_0 + \int_0^T e^{\int_0^t a(s)ds} b(t)dt + \int_0^T e^{\int_0^t a(s)ds} \sigma(t) dW_t \right]$$

or

$$X_T = e^{-\int_0^T a(s)ds} x_0 + \int_0^T e^{-\int_t^T a(s)ds} b(t)dt + \int_0^T e^{-\int_t^T a(s)ds} \sigma(t) dW_t.$$

Mock Exam 5 IX

d) We compute

$$E[X_T] = E[e^{-\int_0^T a(s)ds} x_0] + E[\int_0^T e^{-\int_t^T a(s)ds} b(t)dt] +$$

$$+E\left[\int_0^T e^{-\int_t^T a(s)ds} \sigma(t)dW_t\right].$$

Now, the first two quantities inside expectations on the right hand side are deterministic, so we can remove expectations. Moreover, we recall that the expected value of an Ito integral is zero. We get

$$E[X_T] = e^{-\int_0^T a(s)ds} x_0 + \int_0^T e^{-\int_t^T a(s)ds} b(t)dt.$$

Mock Exam 5 X

e) The distribution is indeed Gaussian. The intuitive reason is that all terms dW_t in the Ito integral are normal independent random variables, because increments of brownian motion are independent and normally distributed. The expression is then a sum (the integral is essentially a continous sum) of independent random variables multiplied by some deterministic quantities plus other deterministic quantities, leading to a final normal random variable.

Mock Exam 5 I

Problem 2: Option pricing - Short Risk Reversal in Black Scholes
Given a stock with price S_t at time t, $t \geq 0$, consider a payoff Y that is short a call option with strike K_1 and long a put option with strike K_2, with $K_2 < S_0 < K_1$, both options with maturity T. In formula, the payoff is $Y = -(S_T - K_1)^+ + (K_2 - S_T)^+$ and is called a bear (or short) risk reversal payoff.
a) Draw a plot of this payoff as a function of S_T. Explain what type of investor would be interested in buying this payoff and what views on the stock marcket this investor would have.
b) Price the short risk reversal in a Black-Scholes model with stock price dynamics

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \ s_0$$

under the physical measure P, and where interest rates r are constant and deterministic. You can use the formula for a call option without

Mock Exam 5 II

deriving it. Derive the formula for the put option, either through put-call parity or through risk neutral valuation.
c) Calculate the delta of the short risk reversal, namely the sensitivity of its price to the initial stock price s_0. How does the short risk reversal change with s_0?
d) Calculate the vega of the short risk reversal, namely the sensitivity of its price to the volatility σ. [Hint: you can simplify the vega formula obtained through differentiation by using the lemma, for a call option in Black Scholes, $S_0\phi(d_1) = Ke^{-rT}\phi(d_2)$, or if you remember the vegas in Black Scholes you can cite the formula directly without deriving it]

Mock Exam 5 III

a) To draw a plot of Y it is best to re-write it in different areas of the S domain. We note that the two options will have different values depending on $S_T > K_1$ or $S_T > K_2$ so we distinguish three cases:
(i) $S_T < K_2$, (ii) $K_2 < S_T < K_1$, (iii) $K_1 < S_T$.
Let us look at the three cases:
(i) $S_T < K_2 \implies$ the put option is in the money, the call option is out of the money and it is worth zero. The payoff is $Y = K_2 - S_T$.
(ii) $K_2 < S_T < K_1 \implies$ the put option is out of the money, and the call too. So the payoff is zero, $Y = 0$.
(iii) $K_1 < S_T \Rightarrow$ the put option is out of the money and has zero payoff, while the call is in-the-money and has positive payoff $S_T - K_1$, so short the call is $-(S_T - K_1) = K_1 - S_T$.

We get

Mock Exam 5 IV

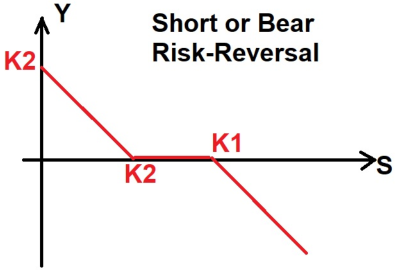
$$Y = -(S_T - K_1)^+ + (K_2 - S_T)^+ = \begin{cases} K_2 - S_T & \text{for } S_T \le K_2 \\ 0 & \text{for } K_2 < S_T < K_1 \\ K_1 - S_T & \text{for } S_T \ge K_1 \end{cases}$$

If we include the initial price of Y in the payoff itself, the initial price may be positive or negative depending on the strikes and other parameters. We would then have to shift the plot of the initial price to include the initial price of the trade in the overall payoff.
We can also write the payoff using indicator functions:

$$Y = (K_2 - S_T)1_{\{S_T \le K_2\}} + (K_1 - S_T)1_{\{S_T > K_1\}}.$$

We can now draw a plot easily.

Mock Exam 5 V



What kind of investor would buy this payoff? The payoff decreases with the stock, except in the interval $[K_2, K_1]$ where it stays constant to zero.

The payoff will make more money if the stock moves below K_2, and the more it moves below K_2 the more money it makes. The extreme case is the stock going to zero, which would give a value of K_2 to the payoff. This is the maximum value the payoff can take.

Mock Exam 5 VI

If the stock is between K_2 and K_1 the payoff is worth nothing, as both options expire out of the money.

Finally, if the stock is larger than K_1 then the put is worth nothing but the short call gives a negative payoff $K_1 - S_T$, and the payoff becomes negative, the more negative the more the stock becomes larger compared to K_1. Note that here the loss is potentially unlimited, as there is no bound for the stock to grow, as opposed for the put options where the stock could not go below zero.

It follows that an investor will buy this payoff only if she expects the stock price to move significantly below K_2 and will not be interested in buying this payoff (or might sell it) if she expects the stock to grow significantly above K_1.

Mock Exam 5 VII

b) To price the risk reversal we need the price of a put with strike K_2 and maturity T minus the price of a call with strike K_1 and maturity T, both prices in the Black Scholes model. For the call we recall that

$$V_{BS}^{CALL}(0, S_0, K_1, T, \sigma, r) = s_0 \Phi(d_1^{(1)}) - K_1 e^{-rT} \Phi(d_2^{(1)})$$

where Φ is the CDF of a standard normal and where

$$d_1^{(1)} = \frac{\ln(s_0/K_1) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \; d_2^{(1)} = d_1^{(1)} - \sigma\sqrt{T}.$$

For the put, we derive its price by put-call parity. Write the argument and the derivation here, as it has been done in the lecture, using the put call parity and the price of a forward contract. We get

$$V_{BS}^{PUT}(0, S_0, K_2, T, \sigma, r) = K_2 e^{-rT} \Phi(-d_2^{(2)}) - s_0 \Phi(-d_1^{(2)}).$$

Mock Exam 5 VIII

$$d_1^{(2)} = \frac{\ln(s_0/K_2) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \; d_2^{(2)} = d_1^{(2)} - \sigma\sqrt{T}.$$

We can now calculate the Short Risk Reversal (SRR) price as

$$V_{BS}^{SRR}(0) = -V_{BS}^{CALL}(0, K_1) + V_{BS}^{PUT}(0, K_2) =$$

$$= -s_0\Phi(d_1^{(1)}) + K_1 e^{-rT}\Phi(d_2^{(1)}) + K_2 e^{-rT}\Phi(-d_2^{(2)}) - s_0\Phi(-d_1^{(2)}) =$$

$$= K_1 e^{-rT}\Phi(d_2^{(1)}) + K_2 e^{-rT}\Phi(-d_2^{(2)}) - s_0\left(\Phi(d_1^{(1)}) + \Phi(-d_1^{(2)})\right).$$

(c) From the fact that the risk reversal is

$$V_{BS}^{SRR}(0) = -V_{BS}^{CALL}(0, s_0, K_1) + V_{BS}^{PUT}(0, s_0, K_2)$$

Mock Exam 5 IX

we can calculate the delta quickly as

$$\frac{\partial V_{BS}^{SRR}(0)}{\partial s_0} = -\frac{\partial V_{BS}^{CALL}(0, s_0, K_1)}{\partial s_0} + \frac{\partial V_{BS}^{PUT}(0, s_0, K_2)}{\partial s_0}.$$

We know from memory (otherwise derive it, see lecture notes) that, in the basic theory of Black Scholes, the delta of a call option is

$$\frac{\partial V_{BS}^{CALL}(0, s_0, K_1)}{\partial S_0} = \Phi(d_1^{(1)}).$$

For the delta of a put, we use again put-call parity to derive the delta of a put from the delta of call and forward contract (see lecture notes). We obtain the formula

$$\frac{\partial V_{BS}^{PUT}(0, s_0, K_2)}{\partial S_0} = -\Phi(-d_1^{(2)}).$$

Mock Exam 5 X

The total delta of the SRR is

$$\frac{\partial V_{BS}^{SRR}(0)}{\partial s_0} = -\frac{\partial V_{BS}^{CALL}(s_0, K_1)}{\partial s_0} + \frac{\partial V_{BS}^{PUT}(s_0, K_2)}{\partial s_0} = -\Phi(d_1^{(1)}) - \Phi(-d_1^{(2)}).$$

Note that the delta of the SRR is negative, meaning that the price of the SRR will go down when the stock S_0 increases. This is in agreement with our intuition of the payoff and the discussion in point a), as the payoff is decreasing in S_T.

(d) The risk reversal vega is

$$\frac{\partial V_{BS}^{SRR}(0)}{\partial \sigma} = \dots$$

Mock Exam 5 XI

Looking at the formula for $V_{BS}^{SRR}(0)$, we see it depends on the volatility only through the terms d_1 and d_2. We will need then to calculate

$$\frac{\partial \Phi(\pm d_1)}{\partial \sigma} = \pm \Phi'(\pm d_1)\frac{\partial d_1}{\partial \sigma} = \pm \phi(\pm d_1)\frac{\partial d_1}{\partial \sigma}$$

$$= \pm \phi(\pm d_1)\frac{\partial}{\partial \sigma}\left(\frac{\ln(s_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}\right)$$

$$= \pm \phi(\pm d_1)\frac{\partial}{\partial \sigma}\left(\frac{\ln(s_0/K) + rT}{\sigma\sqrt{T}} + \frac{\sigma}{2}\sqrt{T}\right)$$

$$= \pm \phi(\pm d_1)\left[-\frac{\ln(s_0/K) + rT}{\sigma^2\sqrt{T}} + \frac{\sqrt{T}}{2}\right],$$

Mock Exam 5 XII

whereas

$$\frac{\partial \Phi(\pm d_2)}{\partial \sigma} = \pm \Phi'(\pm d_2)\frac{\partial d_2}{\partial \sigma} = \pm \phi(\pm d_2)\frac{\partial d_2}{\partial \sigma}$$

$$= \pm \phi(\pm d_2)\frac{\partial}{\partial \sigma}\left(\frac{\ln(s_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}\right)$$

$$= \pm \phi(\pm d_2)\frac{\partial}{\partial \sigma}\left(\frac{\ln(s_0/K) + rT}{\sigma\sqrt{T}} - \frac{\sigma}{2}\sqrt{T}\right)$$

$$= \pm \phi(\pm d_2)\left[-\frac{\ln(s_0/K) + rT}{\sigma^2\sqrt{T}} - \frac{\sqrt{T}}{2}\right],$$

leading to

$$Vega = \frac{\partial V_{BS}^{SRR}(0)}{\partial \sigma} =$$

Mock Exam 5 XIII

$$= K_1 e^{-rT} \partial_\sigma \Phi(d_2^{(1)}) + K_2 e^{-rT} \partial_\sigma \Phi(-d_2^{(2)}) - s_0 \left(\partial_\sigma \Phi(d_1^{(1)}) + \partial_\sigma \Phi(-d_1^{(2)}) \right)$$

$$= -K_1 e^{-rT} \phi(d_2^{(1)}) \left[\frac{\ln(s_0/K_1) + rT}{\sigma^2 \sqrt{T}} + \frac{\sqrt{T}}{2} \right]$$

$$+ K_2 e^{-rT} \phi(-d_2^{(2)}) \left[\frac{\ln(s_0/K_2) + rT}{\sigma^2 \sqrt{T}} + \frac{\sqrt{T}}{2} \right]$$

$$+ s_0 \phi(d_1^{(1)}) \left[\frac{\ln(s_0/K_1) + rT}{\sigma^2 \sqrt{T}} - \frac{\sqrt{T}}{2} \right]$$

$$- s_0 \phi(-d_1^{(2)}) \left[\frac{\ln(s_0/K_2) + rT}{\sigma^2 \sqrt{T}} - \frac{\sqrt{T}}{2} \right] = \dots$$

Mock Exam 5 XIV

Now, using the lemma stating that in the Black Scholes model for options with strike K we have $S_0\phi(d_1) = Ke^{-rT}\phi(d_2)$, we can simplify (recall also that $\phi(-d_{1,2}) = \phi(d_{1,2})$ as ϕ is symmetric)

$$= -s_0\phi(d_1^{(1)})\left[\frac{\ln(s_0/K_1) + rT}{\sigma^2\sqrt{T}} + \frac{\sqrt{T}}{2}\right]$$

$$+ s_0\phi(d_1^{(2)})\left[\frac{\ln(s_0/K_2) + rT}{\sigma^2\sqrt{T}} + \frac{\sqrt{T}}{2}\right]$$

$$+ s_0\phi(d_1^{(1)})\left[\frac{\ln(s_0/K_1) + rT}{\sigma^2\sqrt{T}} - \frac{\sqrt{T}}{2}\right]$$

$$- s_0\phi(d_1^{(2)})\left[\frac{\ln(s_0/K_2) + rT}{\sigma^2\sqrt{T}} - \frac{\sqrt{T}}{2}\right] =$$

Mock Exam 5 XV

$$Vega = s_0\phi(d_1^{(2)})\sqrt{T} - s_0\phi(d_1^{(1)})\sqrt{T}.$$

If one had remembered the vega for a call or put in Black Scholes, given that the SRR is put K_2 minus call K_1, the vega would be vega of put K_2 (that is equal to vega call K_2 by put call parity) minus vega call K_1, and the above formula would have been obtained immediately. This would have been an acceptable solution if one had memorized the vega of a call/put.

Mock Exam 5 I

Problem 3. Option pricing - Bear Call Spread in displaced diffusions
A Bear Call Spread (BeCS) payoff is the difference between a call option payoff with larger strike and a call option payoff with a smaller strike. The underlying asset and the option maturities are the same. In formula: if S_T is the stock price at maturity T, and the strikes are $K_1 > K_2$, then the BeCS payoff is

$$Y = (S_T - K_1)^+ - (S_T - K_2)^+.$$

We assume the initial stock price S_0 to be in-between the two strikes: $K_2 < S_0 < K_1$.
a) Draw a plot of the payoff. Provide your intuition on the payoff and explain what kind of investor might be interested in entering this position.

Mock Exam 5 II

b) Suppose we observe a decreasing volatility smile for the stock S_T. Explain why the displaced diffusion (DD) model is (i) consistent with a decreasing smile and how it can also (ii) produce an increasing volatility smile if needed, changing the parameters. Explain which other models could be used with these two features and why DD is more conveninent.

c) With the DD model chosen in b), price the BeCS. To avoid singularities in the formula, assume

$$|\alpha| << K_1, \ |\alpha| << K_2, \ |\alpha| << S_0.$$

d) Assume now that we take an unrealistically large value for the shift, $\alpha = e^{-rT} K_2^-$, a number smaller than $e^{-rT} K_2$ by an infinitesimal amount. Calculate the new price of the BeCS and its sensitivity to the initial condition S_0. We still assume $S_0 > K_2$ (and hence $S_0 > e^{-rT} K_2$).

Mock Exam 5 III

Solutions of Mock Exam 5. Problem 3: Bear Call Spread with Smile models.

a) The payoff can also be written as follows, by looking at what happens to the two options in the three cases $S_T \leq K_2$, $K_2 < S_T < K_1$, $S_T \geq K_1$.

When $S_T \leq K_2$ we have both call options are out of the money and are worth 0 at maturity.

When $K_2 < S_T < K_1$ we have that the short call with strike K_2 is positive, leading to $-(S_T - K_2)^+ = -(S_T - K_2) = K_2 - S_T$, while the long call with strike K_1 is worth 0.

When $S_T \geq K_1$ both options are in-the-money, so we have $Y = -(S_T - K_2)^+ + (S_T - K_1)^+ = -(S_T - K_2) + (S_T - K_1) = -(K_1 - K_2) < 0$.

Mock Exam 5 IV

Summarizing:

$$Y = -(K_1 - K_2)1_{S_T > K_1} - (S - K_2)1_{K_2 < S_T \leq K_1} + 0\ 1_{S \leq K_2}.$$

or

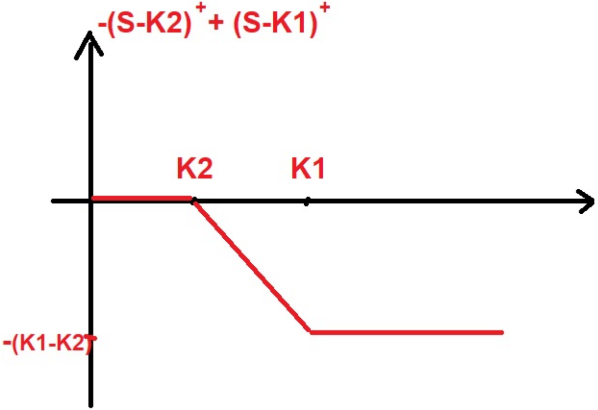
$$Y = -(S_T - K_2)^+ + (S_T - K_1)^+ = \begin{cases} 0 & \text{for } S_T \leq K_2 \\ K_2 - S_T & \text{for } K_2 < S_T < K_1 \\ -(K_1 - K_2) & \text{for } S_T \geq K_1 \end{cases}$$

The payoff is always negative or zero, and therefore its risk neutral discounted expectation will be negative, leading to a negative price. If we enter this position, we pay a negative price, meaning that we will receive money to enter this payoff, as is expected, given that we can only lose money or get nothing at maturity. We would be interested in entering this payoff, cashing in its price at time 0 from the client, if we

Mock Exam 5 V

expected the payoff to expire worthless. In other terms, we expect the stock S to move below K_2 at maturity. This way we pay nothing from the payoff at maturity but we cash in the initial premium from the client at time 0, making an overall profit. The worse that can happen with this payoff is that we need to pay $-(K_1 - K_2)$ to the client at maturity. This is our maximum possible loss. With this payoff the potential loss is bounded.

Mock Exam 5 VI



Mock Exam 5 VII

b) If we observe a decreasing smile, we know from the theory that the models consistent with this are

1. Bachelier
2. Displaced Diffusion (DD) with negative shift;
3. CEV with exponent smaller than 1.

If we wish our model to be able to reproduce also an increasing smile, then we need to rule out Bachelier, which only gives us a decreasing smile. An increasing smile is given by

2. Displaced Diffusion with positive shift;
3. CEV with exponent larger than 1.

Why would we choose the DD? We know that CEV requires special functions and is less tractable than DD.

Mock Exam 5 VIII

To answer (i) and (ii), short answer (acceptable): we have seen in the theory on the DD that it produces a monotonic smile with a decreasing pattern for $\alpha < 0$ and an increasing one with $\alpha > 0$. Being one of the models with monotonic smile patterns and allowing for a decreasing smile by constraining one of the parameters (α) to be negative, the DD is consistent with a negative smile. We also know from the theory that for positive α it produces an increasing smile.

Long answer (if you have time). We note that the DD return/relative volatility is decreasing for $\alpha < 0$ and increasing for $\alpha > 0$, while Black Scholes return/relative volatility is constant, so we expect a decreasing smile curve for $\alpha < 0$ and an increasing smile curve for $\alpha > 0$.

Mock Exam 5 IX

To show this, write the BS and DD models in terms of their relative volatilities.

$$dS = rSdt + \boxed{\nu} SdW, \quad dS = rSdt + \boxed{\sigma(1 - \frac{\alpha e^{rt}}{S})} SdW.$$

So, for $\alpha < 0$, the relative volatility of DD is $\sigma(1 - \frac{\alpha e^{rt}}{S_t})$ and it is a decreasing function of S, contrary to the BS relative vol ν that is constant. As DD relative/return volatility is a decreasing function of S, this vol gets smaller and smaller as K increases when we ask $S > K$ for the call option to have more and more positive value. Therefore, for larger and larger K, the return volatility in DD with $\alpha < 0$ being smaller and smaller, we will have a smaller ν in Black Scholes to match it, or in other terms as the strike increases the matching Black Scholes

Mock Exam 5 X

volatility decreases, because the return vol in the DD model is smaller. The smile is expected to be decreasing.
The opposite reasoning applies if α is positive, since in that case for $\alpha > 0$ we get an increasing function of S for the relative/return volatility, $\sigma(1 - \frac{\alpha e^{rt}}{S_t})$, and we expect an increasing smile.

c) Recall the DD model and write it (under the measure Q) as

$$S_t = X_t + \alpha\ e^{rt}, \ \ dX_t = rX_t dt + \sigma X_t dW_t, \ \ X_0 = S_0 - \alpha.$$

The above is the best form of the model for option pricing, although the model can be written more succinctly as

$$dS_t = rS_t dt + \sigma(S_t - \alpha e^{r\ t})dW_t, \ \ S_0.$$

Mock Exam 5 XI

To avoid singularities or problems with logarithms in the formula, we assume

$$|\alpha| << K_1, \ |\alpha| << K_2, \ |\alpha| << S_0.$$

The price of the Bear Call Spread is the difference of the two call prices in a DD model.

$$V_{DD}^{BeCS} = E_0^Q[e^{-rT}\left(-(S_T - K_2)^+ + (S_T - K_1)^+\right)]$$

$$= -E_0^Q[e^{-rT}(S_T - K_2)^+] + E_0^Q[e^{-rT}(S_T - K_1)^+]$$

$$= -E_0^Q[e^{-rT}(X_T + \alpha e^{rT} - K_2)^+] + E_0^Q[e^{-rT}(X_T + \alpha e^{rT} - K_1)^+]$$

$$= -E_0^Q[e^{-rT}(X_T - K_2')^+] + E_0^Q[e^{-rT}(X_T - K_1')^+] = \dots$$

Mock Exam 5 XII

where $K'_{1,2} = K_{1,2} - \alpha e^{rT}$. By continuing, and remembering that X follows a standard Black Scholes model, we get

$$\ldots = -V_{BS}^{Call}(0, X_0, K'_2, T, \sigma, r) + V_{BS}^{Call}(0, X_0, K'_1, T, \sigma, r)$$

where the call formulas are computed with $X_0 = S_0 + \alpha$ and with modified strikes K':

$$V_{BS}^{Call}(0, X_0, K'_2, T, \sigma, r) = X_0 \Phi(d_1(K'_2)) - K'_2 e^{-rT} \Phi(d_2(K'_2))$$

$$= (S_0 - \alpha)\Phi(d_1(K'_2)) - (K_2 - \alpha e^{rT}) e^{-rT} \Phi(d_2(K'_2))$$

$$d_{1,2}(K'_2) = \frac{\ln \frac{X_0}{K'_2} + \left(r \pm \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}}$$

$$= \frac{\ln \frac{S_0 - \alpha}{K_2 - \alpha e^{rT}} + \left(r \pm \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}}$$

Mock Exam 5 XIII

and similarly for the call with strike K_1':

$$V_{BS}^{Call}(0, X_0, K_1', T, \sigma, r) = (S_0 - \alpha)\Phi(d_1(K_1')) - (K_1 - \alpha e^{rT})e^{-rT}\Phi(d_2(K_1'))$$

$$d_{1,2}(K_1') = \frac{\ln \frac{S_0-\alpha}{K_1-\alpha e^{rT}} + \left(r \pm \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}}.$$

We conclude

Mock Exam 5 XIV

$$V_{DD}^{BeCS} = -(S_0 - \alpha)\Phi\left(\frac{\ln\frac{S_0-\alpha}{K_2-\alpha e^{rT}} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$+(K_2 - \alpha e^{rT})e^{-rT}\Phi\left(\frac{\ln\frac{S_0-\alpha}{K_2-\alpha e^{rT}} + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$+(S_0 - \alpha)\Phi\left(\frac{\ln\frac{S_0-\alpha}{K_1-\alpha e^{rT}} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$-(K_1 - \alpha e^{rT})e^{-rT}\Phi\left(\frac{\ln\frac{S_0-\alpha}{K_1-\alpha e^{rT}} + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

Mock Exam 5 XV

d) If $\alpha = K_2^- e^{-rT}$ we have $K_2 - \alpha e^{rT} = K_2 - K_2^- = 0^+$ and $K_2 = \alpha e^{rT}$. The formula in the previous point becomes

$$V_{DD}^{BeCS}|_{\alpha = K_2^- e^{-rT}} = -(S_0 - K_2 e^{-rT})\Phi\left(\frac{\ln\frac{S_0 - K_2 e^{-rT}}{K_2 - K_2^-} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$+(K_2 - K_2^-)e^{-rT}\Phi\left(\frac{\ln\frac{S_0 - K_2 e^{-rT}}{K_2 - K_2^-} + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$+(S_0 - K_2 e^{-rT})\Phi\left(\frac{\ln\frac{S_0 - K_2^- e^{-rT}}{K_1 - K_2^-} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

Mock Exam 5 XVI

$$-(K_1 - K_2^-)e^{-rT}\Phi\left(\frac{\ln\frac{S_0 - K_2^- e^{-rT}}{K_1 - K_2^-} + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

Given that $K_2 - K_2^- \downarrow 0^+$ and $S_0 - K_2^- e^{-rT} > 0$ we can compute the limits of the logarithms as

$$\lim_{x\downarrow 0^+} \ln\frac{S_0 - K_2^- e^{-rT}}{x} = +\infty$$

leading to $\lim_{z\uparrow+\infty}\Phi(z) = 1$ so that

Mock Exam 5 XVII

$$V_{DD}^{BeCS}|_{\alpha=K_2^- e^{-rT}} = -(S_0 - K_2 e^{-rT}) +$$

$$+(S_0 - K_2 e^{-rT})\Phi\left(\frac{\ln\frac{S_0 - K_2^- e^{-rT}}{K_1 - K_2^-} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$-(K_1 - K_2^-)e^{-rT}\Phi\left(\frac{\ln\frac{S_0 - K_2^- e^{-rT}}{K_1 - K_2^-} + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

or symplifying ($\Phi(x) - 1 = -\Phi(-x)$)

Mock Exam 5 XVIII

$$V_{DD}^{BeCS}|_{\alpha=K_2^- e^{-rT}} = -(S_0 - K_2 e^{-rT})\Phi\left(-\frac{\ln\frac{S_0 - K_2^- e^{-rT}}{K_1 - K_2^-} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$-(K_1 - K_2^-)e^{-rT}\Phi\left(\frac{\ln\frac{S_0 - K_2^- e^{-rT}}{K_1 - K_2^-} + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right).$$

For the sensitivity to initial condition S_0 we need to compute

$$\Delta = \frac{\partial V_{DD}^{BeCS}|_{\alpha=K_2^- e^{-rT}}}{\partial S_0}.$$

Mock Exam 5 XIX

Let's see if we can invoke the known formula for the delta of a call option in Black Scholes to avoid the lengthy calculation we would need to do.
Rewrite the price as

$$V_{DD}^{BeCS}|_{\dots} = -(S_0 - K_2 e^{-rT}) +$$

$$+(S_0 - K_2 e^{-rT})\Phi\left(\frac{\ln\frac{S_0 - K_2^- e^{-rT}}{K_1 - K_2^-} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$-(K_1 - K_2^-)e^{-rT}\Phi\left(\frac{\ln\frac{S_0 - K_2^- e^{-rT}}{K_1 - K_2^-} + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right).$$

Mock Exam 5 XX

Now, apart from the first term $-(S_0 - K_2 e^{-rT})$, this is the price of an option in a Black Scholes model with initial stock price $S_0' = S_0 - K_2 e^{-rT}$, strike $K' = K_1 - K_2$, risk free rate r and volatility σ. Thus

$$V_{DD}^{BeCS}|_{\dots} = -(S_0 - K_2 e^{-rT}) + V_{BS}^{CALL}(0, S_0', K', T, \sigma, r)$$

so

$$\frac{\partial V_{DD}^{BeCS}|_{\dots}}{\partial S_0} = \partial_{S_0}[-(S_0 - K_2 e^{-rT})] + \partial_{S_0} V_{BS}^{CALL}(0, S_0', K', T, \sigma, r)$$

$$= -1 + \frac{\partial V_{BS}^{CALL}(0, S_0', K', T, \sigma, r)}{\partial S_0'} \frac{\partial S_0'}{\partial S_0}$$

$$= -1 + \Phi\left(\frac{\ln \frac{S_0'}{K'} + \left(r + \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}}\right) 1$$

Mock Exam 5 XXI

$$= -1 + \Phi\left(\frac{\ln\frac{S_0 - K_2 e^{-rT}}{K_1 - K_2} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

$$= -\Phi\left(-\frac{\ln\frac{S_0 - K_2 e^{-rT}}{K_1 - K_2} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

where we used the fact that

$$\frac{\partial V_{BS}^{CALL}(0, S_0', K', T, \sigma, r)}{\partial S_0'} = \Phi(d_1'),\ d_1' = \frac{\ln\frac{S_0'}{K'} + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

The delta is negative, meaning that the portfolio value will decrease when S_0 increases. This is in line with our prior intuition on the payoff. The bear call spread payoff decreases or stays constant when the stock increases, so it is expected that increasing S_0 will lead to a lower portfolio price.

Mock Exam 5 I

Problem 4. Risk Measures. Bond & Stock with different maturities. Consider a portfolio with a first bond position in a notional N of zero-coupon bond with maturity U in an economy where we have a deterministic constant risk free interest rate r. Assume that in the same portfolio we are short a bond with maturity $T < U$ on the same notional, and that we hold an amount M of stock S, where the stock price follows the following dynamics under the measure P:
$dS_t = \mu S_t dt + \sigma S_t dW_t,\ s_0$. We assume $M > 0, N > 0$.
a) Compute the Value at risk of this portfolio for a confidence level α at a risk horizon $h < T$.
b) Is VaR increasing or decreasing in the initial stock price S_0? Can you provide financial intuition for your answer?
c) Is VaR increasing or decreasing in the notional N of the bonds? Can you provide financial intuition for your answer?
Solutions.

Mock Exam 5 II

a) Let us analyze the three positions and in particular their value at time h, where we have to assess VaR. A zero coupon bond with maturity U on a notional N promises to pay the notional N at time U. Its value at time $h < U$ is obtained by risk neutral pricing as

$$E_h^Q[e^{-r(U-h)}N] = e^{-r(U-h)}N$$

where we could take away the expectation as there is nothing random in the payoff or in the discount rate r.
A similar approach leads to the price $-e^{-r(T-h)}N$ for the short bond position, where the minus sign is due to the short position.
The value of the stock at time h is simply MS_h, namely the amount of stock we hold times the price of the stock at time h.
Putting all terms together the value of the portfolio is

Mock Exam 5 III

The value of the portfolio at time h is

$$V_h = Ne^{-r(U-h)} - Ne^{-r(T-h)} + MS_h.$$

The value of the portfolio at time 0 is instead, trivially,

$$V_0 = Ne^{-rU} - Ne^{-rT} + MS_0.$$

The loss of the portfolio over the time h is

$$L_h = V_0 - V_h = Ne^{-rU} - Ne^{-rT} + MS_0 - Ne^{-r(U-h)} + Ne^{-r(T-h)} - MS_h.$$

We can set

$$K = Ne^{-rU} - Ne^{-rT} + MS_0 - Ne^{-r(U-h)} + Ne^{-r(T-h)}$$

Mock Exam 5 IV

and rewrite the loss as

$$L_h = K - MS_h.$$

The only random term here is S_h, which in the Black Scholes model is written, under the measure P, as

$$S_h = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) h + \sigma W_h\right)$$

where as usual we recall that $W_h \sim \mathcal{N}(0, h) \sim \sqrt{h}\mathcal{N}(0, 1) \sim \sqrt{h}\mathcal{N}$ where we abbreviate $\mathcal{N} = \mathcal{N}(0, 1)$. To compute $q_{h,\alpha} = VaR_{h,\alpha}$ we need to find the percentile such that

$$\mathbb{P}\{L_h < q_{h,\alpha}\} = \alpha.$$

Mock Exam 5 V

Write $q = q_{h,\alpha}$ for brevity, and calculate

$$\mathbb{P}\{L_h < q\} = P\{K - MS_h < q\} = P\left\{S_h > \frac{K-q}{M}\right\}$$

$$= P\left\{S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) h + \sigma\sqrt{h}\mathcal{N}\right) > \frac{K-q}{M}\right\}$$

$$= P\left\{\mathcal{N} > \frac{\ln\frac{K-q}{MS_0} - \left(\mu - \frac{1}{2}\sigma^2\right) h}{\sigma\sqrt{h}}\right\}$$

$$= 1 - \Phi\left(\frac{\ln\frac{K-q}{MS_0} - \left(\mu - \frac{1}{2}\sigma^2\right) h}{\sigma\sqrt{h}}\right)$$

$$= \Phi\left(-\frac{\ln\frac{K-q}{MS_0} - \left(\mu - \frac{1}{2}\sigma^2\right) h}{\sigma\sqrt{h}}\right).$$

Mock Exam 5 VI

Thus the equation

$$\mathbb{P}\{L_h < q_{h,\alpha}\} = \alpha$$

becomes

$$\Phi\left(-\frac{\ln\frac{K-q}{MS_0} - \left(\mu - \frac{1}{2}\sigma^2\right)h}{\sigma\sqrt{h}}\right) = \alpha$$

or

$$-\frac{\ln\frac{K-q}{MS_0} - \left(\mu - \frac{1}{2}\sigma^2\right)h}{\sigma\sqrt{h}} = \Phi^{-1}(\alpha)$$

from which we can solve in q, obtaining

$$q = VaR_{h,\alpha} = K - MS_0 \exp\left(-\sigma\sqrt{h}\Phi^{-1}(\alpha) + \left(\mu - \frac{1}{2}\sigma^2\right)h\right)$$

Mock Exam 5 VII

b) We need to remember that K depends on S_0 too, so it will contribute to the behaviour of VaR. We have $K = MS_0 + K'$, with $K' = Ne^{-rU} - Ne^{-rT} - Ne^{-r(U-h)} + Ne^{-r(T-h)}$. In particular, K' does not depend on S_0. Hence, from the previous point,

$$VaR_{h,\alpha} = MS_0 - MS_0 \exp\left(-\sigma\sqrt{h}\Phi^{-1}(\alpha) + \left(\mu - \frac{1}{2}\sigma^2\right)h\right) + K'$$

$$= MS_0\left[1 - \exp\left(-\sigma\sqrt{h}\Phi^{-1}(\alpha) + \left(\mu - \frac{1}{2}\sigma^2\right)h\right)\right] + K'.$$

As M is positive, whether VaR is increasing or decreasing will depend on the sign of the quantity between squared brackets. This in turn will depend on the sign of the exponent of the exponential function. If this sign is positive, the exponential will be larger than one and the squared bracket term will be negative, leading to a decreasing VaR. If the sign

Mock Exam 5 VIII

is negative, the opposite will happen, leading to an increasing VaR. So we need to find conditions under which

$$-\sigma\sqrt{h}\Phi^{-1}(\alpha) + \left(\mu - \frac{1}{2}\sigma^2\right) h < 0$$

for VaR to be increasing. This cannot be solved in general but depends on the parameters of the calculation, σ, h, α, μ. So it is not possible to give an answer for the general case, but we can still investigate what happens for typical ranges of values of the parameters, as is usually done in real portfolios.

The confidence level is typically 0.99, so we can assume $\Phi^{-1}(\alpha) = \Phi^{-1}(0.99) = 2.33$. Our inequality then becomes

$$\mu h - 2.33\sigma\sqrt{h} - \frac{1}{2}\sigma^2\ h < 0$$

Mock Exam 5 IX

Now because typically in VaR calculations $h < 1$, we will have that $\sqrt{h} > h$. Also, typical ranges of stock volatilties may go from 0.1 to 0.5. So we see that unless the return μ is expected to be really large, the condition will be satisfied. For example, for a volatility of 0.1 and $h = 0.25$ we get

$$\mu/4 - 2.33\ 0.1/2 - 1/2\ 0.01/4 < 0.$$

For this to be true we require

$$\mu < 0.471$$

or a return smaller than 47%. which is extremely realistic. So the condition will be satisfied and the exponent will be negative, resulting in VaR being increasing in S_0.

Mock Exam 5 X

Let's take another example at the other extreme, take the volatility to be $\sigma = 0.5$ and the risk horizon $h = 1$, and we get

$$\mu - 2.33\ 0.5 - 1/2\ 0.25 < 0,$$

which requires $\mu < 1.04$ so returns smaller than 104%, which is again very realistic, so we have the same conclusions as in the previous case.
It seems that in most realistic situations the exponent will be negative, so the exponential will be smaller than 1, and the total VaR will be a positive constant times S_0 (plus K'), thus resulting in an increasing function of S_0.
So in these realistic examples we see that *VaR* is increasing in S_0.
This means that when the stock price increases, we typically have a larger potential loss at the given confidence level over the given

Mock Exam 5 XI

horizon. This is intuitive because we have a long position in the stock, so whenever the initial stock is larger, we have a larger stock position with the same volatility, so we have that the risk of a potential loss at the same confidence level and risk horizon becomes larger as the portfolio size is larger and with the same volatility.
We emphasize however that this result is not general, and there can be atypical values of the parameters σ, h, α, μ for which this does not hold.

c) The VaR term depending on N is K so we need to establish whether K increases or decreases in N. Recall

$$K = Ne^{-rU} - Ne^{-rT} + MS_0 - Ne^{-r(U-h)} + Ne^{-r(T-h)}$$

Let's take the derivative of K with respect to N:

$$\partial_N K = e^{-rU} - e^{-rT} - e^{-r(U-h)} + e^{-r(T-h)} =$$

Mock Exam 5 XII

$$= e^{-r(T-h)} - e^{-rT} - (e^{-r(U-h)} - e^{-rU}) =$$

Now the first exponential is larger than the second, as it has a less negative exponent. So the difference of the first two exponentials is positive. However, for the same reason the third exponential is larger than the fourth one, so the term in round brackets is positive. We have the difference of two positive terms. As we know $U > T$, let us collect e^{-rT} outside.

$$= e^{-rT}\left[e^{rh} - 1 - (e^{-r(U-T-h)} - e^{-r(U-T)})\right]$$

$$= e^{-rT}\left[1(e^{rh} - 1) - e^{-r(U-T)}(e^{rh} - 1)\right]$$

$$= e^{-rT}(1 - e^{-r(U-T)})(e^{rh} - 1) > 0$$

Mock Exam 5 XIII

because all factors in the last expression are positive. We deduce that $\partial_N K > 0$, so K is increasing in N, and therefore $VaR_{h,\alpha}$ also increases in N.

Financially, this is telling us that increasing the amount of long U-bonds and short T-bonds increases VaR or our potential loss. This makes sense because the T bond is more valuable than the U bond, given $U > T$, as the U bond will pay later and so is worth less. So increasing both bond notionals will impact more the T-bond, which is a short position, than the U bond, which is a long position. With the short position becoming more valuable compared to the long position we are facing a larger potential loss.

Mock Exam 6 I

Problem 1. SDEs - $dX_t = X_t/2dt + \sqrt{1+X_t^2}dW_t$.
Consider the SDE

$$dX_t = \frac{1}{2}X_t dt + \sqrt{1+X_t^2}dW_t,$$

with $X_0 = x_0$ deterministic.
a) Prove that this SDE admits a unique global solution.
b) Calculate the solution (hint: use Stratonovich calculus).
c) Calculate the expected value of the solution, $E_0[X_t]$, for all $t \geq 0$.

Mock Exam 6 II

Problem 1: Solutions.

a) We use the theorem giving sufficient conditions for global existence and uniqueness of solutions for SDEs.

We know from the theory that for the SDE
$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = Z$ with Z independent of $\sigma(\{W_t, \ t \leq T)\}$ and $\mathbb{E}[Z^2] < +\infty$, and with $\mu : [0, T] \times \mathbb{R} \to \mathbb{R}$ (the drift) and $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ (the diffusion coefficient) being measurable, if we have global Lipschitz continuity

$$|\mu(t,x)-\mu(t,y)|+|\sigma(t,x)-\sigma(t,y)| \leq K|x-y| \text{ for all } t \in [0, T] \text{ and all } x \in \mathbb{R}$$

and linear growth

$$|\mu(t,x)| + |\sigma(t,x)| \leq K'(1 + |x|)| \ \text{ for all } t \in [0, T] \text{ and all } x \in \mathbb{R}$$

for two constants K, K', then our SDE has a unique global solution X_t.

Mock Exam 6 III

Let's check these conditions.
The initial condition has to be squared integrable, $E[X_0^2] < +\infty$, which is true in our case as $X_0 = x_0$ is a finite deterministic constant and $E[X_0^2] = x_0^2 < \infty$. Then we need to prove that the the drift and diffusion coefficient are measurable functions of X, t.
This is trivially true as the drift is a linear function of X and does not depend on t, $\mu(t, X) = X/2$. Also, the diffusion coefficient $\sigma(t, X) = \sqrt{1 + X_t^2}$ is trivially measurable as it is a continuous function of X and does not depend on t.
Next we need to check the Lipschitz continuity and linear growth condition. The Lipschitz condition reads

$$|\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| =$$

Mock Exam 6 IV

$$= |x/2 - y/2| + |\sqrt{1+x^2} - \sqrt{1+y^2}| \leq \dots \tag{45}$$

At this point we need to deal with the second term before the inequality, as the first one is trivially $\frac{1}{2}|x-y|$. We claim that

$$|\sqrt{1+x^2} - \sqrt{1+y^2}| \leq |x-y|$$

so that the above condition reduces to

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq \frac{1}{2}|x-y| + |x-y| = \frac{3}{2}|x-y|$$

and we have Lipschitz continuity with constant $3/2$. To prove

$$|\sqrt{1+x^2} - \sqrt{1+y^2}| \leq |x-y|$$

Mock Exam 6 V

note that

$$|\sqrt{1+x^2} - \sqrt{1+y^2}| = \left|\left(\sqrt{1+x^2} - \sqrt{1+y^2}\right)\frac{(\sqrt{1+x^2}+\sqrt{1+y^2})}{\sqrt{1+x^2}+\sqrt{1+y^2}}\right| =$$

$$= \frac{|x^2-y^2|}{\sqrt{1+x^2}+\sqrt{1+y^2}} \leq$$

now we observe that the denominator is larger than $\sqrt{x^2}+\sqrt{y^2} = |x|+|y|$, so that replacing it with this smaller quantity yields a larger fraction,

$$\leq \frac{|x^2-y^2|}{|x|+|y|} = \frac{|x-y||x+y|}{|x|+|y|} \leq |x-y|$$

where the last inequality follows from $|x+y| < |x|+|y|$.

Mock Exam 6 VI

We thus have Lipschitz continuity.

Given that the SDE is autonomous, namely the coefficients do not depend explicitly on t, we invoke the result seen in the lectures that global Lipschitz continuity implies linear growth, so we have both conditions and we are done (write this in the exam, namely that you invoke that result if you don't prove linear growth explicitly, and you are ok).

In case the eager students wishes to prove linear growth directly anyway (which is not necessary if the above point has been made in writing), we need to show that

$$|\mu(t,x)| + |\sigma(t,x)| \leq K'(1+|x|)| \text{ for all } t \in [0,T] \text{ and all } x \in \mathbb{R}$$

Mock Exam 6 VII

We have

$$|\mu(t,x)|+|\sigma(t,x)| = |x|+\sqrt{1+x^2} \le \frac{1}{2}|x|+\sqrt{1}+\sqrt{x^2} = 1+\frac{3}{2}|x| \le \frac{3}{2}(1+|x|$$

$$\text{for all } t \in [0,T] \text{ and all } x \in \mathbb{R}$$

as in general, for two positive real number a and b, $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ (square both sides, that are positive, to convince yourself of this).
So we have a unique global solution.

b) Calculation of the solution using Stratonovich calculus is based on transforming the Ito SDE in a Stratonovich SDE. We know that the transformation changes the drift into

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

Mock Exam 6 VIII

In our specific case,

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

so that the Stratonovich drift becomes

$$\frac{1}{2}x \mapsto \frac{1}{2}x - \frac{1}{2}x = 0.$$

The equivalent Stratonovich SDE with the same solution is therefore

$$dX_t = \sqrt{1+X_t^2} \circ dW_t, \quad x_0.$$

Mock Exam 6 IX

With Stratonovich, we can use formal rules of calculus. We can separate variables as in

$$\frac{dX_t}{\sqrt{1+X_t^2}} = \circ dW_t,$$

and integrate both sides

$$\int_{x_0}^{X_t} \frac{dX}{\sqrt{1+X^2}} = \int_0^t dW_t,$$

leading to

$$\sinh^{-1}(X)|_{x_0}^{X_t} = W_t$$

or

$$\sinh^{-1}(X_t) - \sinh^{-1}(x_0) = W_t$$

Mock Exam 6 X

$$\sinh^{-1}(X_t) = \sinh^{-1}(x_0) + W_t$$

$$X_t = \sinh\left(\sinh^{-1}(x_0) + W_t\right).$$

c) To calculate the expected value we could use the solution directly and proceed to a direct calculation but this is not convenient, as it involves a much more complicated calculation. Let us use the Ito SDE instead. Consider

$$dX_t = \frac{1}{2}X_t dt + \sqrt{1+X_t^2}dW_t,$$

and write it in integral form

$$X_t = x_0 + \int_0^t \frac{1}{2}X_s ds + \int_0^t \sqrt{1+X_s^2}dW_s,$$

Mock Exam 6 XI

and now take the expected value conditional on information at time 0 on both sides. Recall the that expectation of an Ito integral is zero and that x_0 is deterministic, so we get

$$E_0[X_t] = x_0 + E_0[\int_0^t \frac{1}{2} X_s ds] + 0,$$

or

$$E_0[X_t] = x_0 + \int_0^t \frac{1}{2} E_0[X_s] ds + 0,$$

using Fubini's theorem. Let us now call $m_t = E_0[X_t]$ for all $t \geq 0$. The above equation can be written as

$$m_t = x_0 + \int_0^t \frac{1}{2} m_s \, ds.$$

Mock Exam 6 XII

Now differentiate both sides with respect to t, obtaining

$$\frac{dm_t}{dt} = \frac{1}{2}m_t,$$

with initial condition $m_0 = E_0[X_0] = E_0[x_0] = x_0$. The last differential equation is immediate to integrate. We have

$$\frac{dm}{m} = \frac{1}{2}dt$$

so that, integrating both sides

$$\ln(m)|_{m_0}^{m_t} = \frac{1}{2}t$$

or

$$\ln(m_t) - \ln(m_0) = \frac{1}{2}t$$

Mock Exam 6 XIII

or, rearranging,

$$m_t = m_0 \exp\left(\frac{1}{2}t\right)$$

so that

$$E[X_t] = x_0 \exp\left(\frac{1}{2}t\right).$$

Mock Exam 6 I

Problem 2: Option pricing - Bull call spread in Black Scholes

A bull call spread payoff is the difference between a call option with smaller strike and a call option with a larger strike. The underlying asset and the option maturities are the same.
In formula: if the two strikes are $K_1 > K_2$ and the stock at time t is S_t, with the Bull call spread maturity being T, then the bull call spread payoff is

$$Y = (S_T - K_2)^+ - (S_T - K_1)^+.$$

a) Draw a plot of the payoff of a bull call spread as a function of S_T. Describe the type of investor who would buy this payoff and their views on the market.

Mock Exam 6 II

b) Write a formula for the price of the bull call spread at time 0 in the Black Scholes model where the stock evolves, under the measure P, according to

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where W is a Brownian motion under P and the initial condition $S_0 = s_0$ is deterministic. Assume $K_2 < s_0 < K_1$. The risk free rate of the bank account is assumed to be constant and equal to r. Comment on the sign of the price.

c) Compute the delta of the bull call spread, namely its sensitivity to the stock price at time 0. In other words, if V_0 is the price you computed in point b), calculate $\frac{\partial V_0}{\partial s_0}$. What does the sign allow you to conclude on the behaviour of the bull call spread with respect to the

Mock Exam 6 III

underlying stock and does this confirm your intuition on the payoff interpretation in point a)?

d) Compute the Vega of the bull call spread, namely $\frac{\partial V_0}{\partial \sigma}$. If you don't recall the formula for the vega of a call option in the Black Scholes model, apply differentiation and use the following hint:
$s_0 \phi(d_1) = Ke^{-rT} \phi(d_2)$.

Mock Exam 6 IV

Solutions.

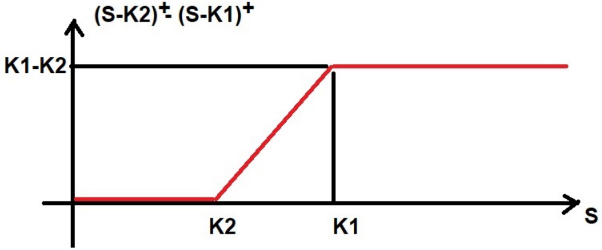
a).
We had already seen this payoff on the part of the course concerning risk measures, but we repeat the analysis here.
The payoff can also be written as

$$Y = (K_1 - K_2)1_{S_T > K_1} + (S - K_2)1_{K_2 < S_T \leq K_1} + 0\ 1_{S \leq K_2}.$$

This contingent claim consists of one long call with a lower strike price and one short call with a higher strike.
Note that the initial price of Y would be positive to us, since it is an in-the-money call minus an out-of-the-money call. This means that to purchase this payoff we need to pay.

Mock Exam 6 V

As for the plot of the payoff of a bull call spread, excluding the initial payment needed to buy the product, the payoff looks like



What type of investor would buy this payoff? A bull call spread profits when the underlying stock rises in price, but profit is limited as the stock price rises above the strike price K_1, and the loss is also limited as the stock price falls below the strike price K_2. So differently from

Mock Exam 6 VI

payoffs like risk reversals, the bull call spread allows for limited gain and limited losses, whichever the model (we know for example that under the Bachelier model a risk reversal can lead to an unlimited loss, potentially). Therefore, as a payoff it will be less risky than a bull or long risk reversal, but at the same time it will allow for less profit in case of strong positive performance of the stock at maturity.

Mock Exam 6 VII

Indeed, the loss is floored after the stock drops below K_2, but the potential profits are also capped as the stock rises above K_1. Hence this contract will be sought by a trader who does not want excessive risk and who is expecting the stock to increase.

The contract may be of interest to a trader with limited funds, because it is less expensive than a call option with stike K_2, as it reduces the price of the call by selling another call with a higher strike K_1. So one finances the purchase of a call with the sale of another call.

b)
To price this payoff, we just need to price the two call options and substract.

Mock Exam 6 VIII

Indeed, we know from the Black Scholes formulas that the price of the payoff Y at time 0 is the Black-Scholes price of the call with strike K_2 minus the Black-Scholes price of the call with strike K_1, namely

$$V_0 = S_0\Phi(d_1(K_2)) - K_2 e^{-rT}\Phi(d_2(K_2)) - [S_0\Phi(d_1(K_1)) - K_1 e^{-rT}\Phi(d_2(K_1))]$$

where

$$d_{1,2}(K) = \frac{\ln(S_t/K) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

The sign of the price is positive, $V_0 > 0$. Indeed, the payoff itself is non-negative in every scenario, so its Q-expectation, leading to the price, will be positive. Another way to look at it is to consider that we are taking the difference of two call options with the same maturity, same stock and different strikes. The first option has a lower strike K_2 with $S_0 > K_2$, so the first call option is in-the-money. The call option we

Mock Exam 6 IX

subtract or sell, has larger strike K_1 and $S_0 < K_1$, so the second call is out of the money. As a call that is in the money is more valuable than a call that is out of the money, everything else being equal, we deduce that the difference will be positive.

c)
To compute the delta, we need to take the partial derivative of V_0 with respect to s_0.
Recall that the delta at time 0 of a call options with stock S, strike K, maturity T, volatility σ and risk free rate r is given by

$$\Delta_{call}(K) = \Phi(d_1(K)), \ \ d_1(K) = \frac{\ln(S_0/K) + \left(r + \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}}.$$

Mock Exam 6 X

Since our bull call spread (BuCS) is the difference of two call options with strikes K_2 and K_1 respectively, we have

$$\Delta_{BuCS} = \frac{\partial V_0}{\partial s_0} = \frac{\partial(\text{CallPrice}(K_2) - \text{CallPrice}(K_1))}{\partial s_0} =$$

$$= \frac{\partial \text{CallPrice}(K_2)}{\partial s_0} - \frac{\partial \text{CallPrice}(K_1)}{\partial s_0} = \Phi(d_1(K_2)) - \Phi(d_1(K_1)).$$

We can discuss the sign of the Delta to see the pattern of the BuCS price with respect to s_0. To do this, we wish to understand whether the delta of a call option, $\Phi(d_1(K))$, is increasing or decreasing in K, or neither. We know Φ is an increasing function as it is the normal CDF,

Mock Exam 6 XI

so the question is whether $d_1(K)$ is increasing, decreasing or neither in K. We can write $d_1(K)$ as

$$d_1(K) = \frac{\ln(S_0) - \ln(K) + \left(r + \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}}.$$

Then

$$\frac{\partial d_1(K)}{\partial K} = \frac{1}{\sigma\sqrt{T}} \frac{\partial(\ln(S_0) - \ln(K) + \left(r + \frac{1}{2}\sigma^2\right) T)}{\partial K} =$$

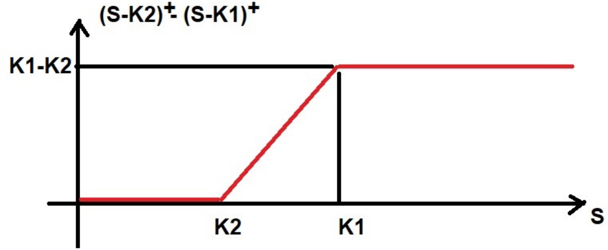
$$= \frac{1}{\sigma\sqrt{T}}(-\frac{1}{K}) < 0.$$

If K is positive, this is a negative numer. This means that $d_1(K)$ is decreasing in K. As BuCS delta is the difference of two call deltas,

$$\Delta_{BuCS} = \Phi(d_1(K_2)) - \Phi(d_1(K_1))$$

Mock Exam 6 XII

with $K_2 < K_1$, and the delta of a call is decreasing in K, it follows that the difference of call deltas is positive. Therefore the delta of a BuCS is positive. This is confirmed by looking at the shape of the payoff.



We see from the picture that if S increases, the value of the payoff increases.

Mock Exam 6 XIII

(d) Here we assume that one does not remember the formula for vega of a call option in Black Scholes. The bull call spread vega is

$$\frac{\partial V_0}{\partial \sigma} = \frac{\partial \mathrm{CallPrice}(K_2)}{\partial \sigma} - \frac{\partial \mathrm{CallPrice}(K_1)}{\partial \sigma} = \ldots$$

Looking at the formula for V_0, we see it depends on the volatility only through the terms d_1 and d_2 of the two call options. We will need then to calculate, as already done for the risk reversal,

$$\frac{\partial \Phi(d_1)}{\partial \sigma} = \Phi'(d_1)\frac{\partial d_1}{\partial \sigma} = \phi(d_1)\frac{\partial d_1}{\partial \sigma}$$

$$= \phi(d_1)\frac{\partial}{\partial \sigma}\left(\frac{\ln(s_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}\right)$$

Mock Exam 6 XIV

$$= \phi(d_1)\frac{\partial}{\partial\sigma}\left(\frac{\ln(s_0/K)+rT}{\sigma\sqrt{T}}+\frac{\sigma}{2}\sqrt{T}\right)$$

$$= \phi(d_1)\left[-\frac{\ln(s_0/K)+rT}{\sigma^2\sqrt{T}}+\frac{\sqrt{T}}{2}\right],$$

where $\Phi'(x)=\phi(x)$ is the standard normal probability density function, whereas

$$\frac{\partial\Phi(d_2)}{\partial\sigma}=\Phi'(d_2)\frac{\partial d_2}{\partial\sigma}=\phi(d_2)\frac{\partial d_2}{\partial\sigma}$$

$$= \phi(d_2)\frac{\partial}{\partial\sigma}\left(\frac{\ln(s_0/K)+(r-\sigma^2/2)T}{\sigma\sqrt{T}}\right)$$

$$= \phi(d_2)\frac{\partial}{\partial\sigma}\left(\frac{\ln(s_0/K)+rT}{\sigma\sqrt{T}}-\frac{\sigma}{2}\sqrt{T}\right)$$

Mock Exam 6 XV

$$= \phi(d_2)\left[-\frac{\ln(s_0/K)+rT}{\sigma^2\sqrt{T}} - \frac{\sqrt{T}}{2}\right].$$

At this point we calculate the vega for a call option with strike K by computing

$$\nu_{Call}(\sigma, K) = \frac{\partial \text{CallPrice}(\sigma, K)}{\partial \sigma} = s_0 \partial_\sigma \Phi(d_1(\sigma, K)) - Ke^{-rT}\partial_\sigma \Phi(d_2(\sigma, K)) =$$

and substituting for the partial derivatives of $\Phi(d_{1,2})$ computed above we get $=$

$$s_0\phi(d_1(K))\left[-\frac{\ln(s_0/K)+rT}{\sigma^2\sqrt{T}} + \frac{\sqrt{T}}{2}\right] - Ke^{-rT}\phi(d_2(K))\left[-\frac{\ln(s_0/K)+rT}{\sigma^2\sqrt{T}} - \frac{\sqrt{T}}{2}\right]$$

Now we use the hint: $s_0\phi(d_1) = Ke^{-rT}\phi(d_2)$ and we simplify the above into $=$

Mock Exam 6 XVI

$$s_0\phi(d_1(K))\left[-\frac{\ln(s_0/K)+rT}{\sigma^2\sqrt{T}}+\frac{\sqrt{T}}{2}\right]-s_0\phi(d_1(K))\left[-\frac{\ln(s_0/K)+rT}{\sigma^2\sqrt{T}}-\frac{\sqrt{T}}{2}\right]$$

$$=s_0\phi(d_1(K))\sqrt{T}.$$

The vega of the bull call spread is

$$\nu_{BuCS}(\sigma,K_1,K_2)=\nu_{Call}(\sigma,K_2)-\nu_{Call}(\sigma,K_1)$$

$$=s_0\phi(d_1(K_2))\sqrt{T}-s_0\phi(d_1(K_1))\sqrt{T}=$$

$$=s_0\sqrt{T}(\phi(d_1(K_2))-\phi(d_1(K_1))).$$

If one had remembered the formula for the vega of a call in Black Scholes, the following much shorter solution would be accepted.

Mock Exam 6 XVII

The SRR is given by a call with strike K_2 minus a call with strike K_1. Thus, given that the partial derivative of a difference is the difference of the partial derivatives, one has

$$\nu_{BuCS}(\sigma, K_1, K_2) = \nu_{Call}(\sigma, K_2) - \nu_{Call}(\sigma, K_1)$$

Recalling that the formula for the vega in a call option in Black Scholes for a strike K is $\nu_{Call}(\sigma, K) = s_0\sqrt{T}\phi(d_1(K))$, one has

$$\nu_{BuCS}(\sigma, K_1, K_2) = s_0\sqrt{T}(\phi(d_1(K_2)) - \phi(d_1(K_1))).$$

Mock Exam 6 I

Problem 3. Option pricing - Bull Call Spread with Bachelier model
A Bull Call Spread (BuCS) payoff is the difference between a call option payoff with smaller strike and a call option payoff with a larger strike, see problem 2 of this exam. The underlying asset and the option maturities are the same.
In formula: if S_T is the stock price at maturity T, and the strikes are $K_1 > K_2$, then the BuCS payoff is

$$Y = (S_T - K_2)^+ - (S_T - K_1)^+.$$

We assume the initial stock price S_0 to be in-between the two strikes: $K_2 < S_0 < K_1$.
a) Draw a plot of the payoff. Provide your intuition on the payoff and explain what kind of investor might be interested in entering this position.

Mock Exam 6 II

b) Suppose we observe a decreasing volatility smile for the stock S_T. Introduce briefly the Bachelier model under the measure P, explain what assumptions it makes on the probability distribution of the stock, explain one important disadvantage of this assumption when modeling stock prices, and explain what particular assumptions you need on the risk free rate r to be able to do the change of measure to Q. Explain why the Bachelier model is consistent with a decreasing smile. Explain which other models could be used to reproduce a purely decreasing smile.

c) With the Bachelier model chosen in b), price the BuCS at time 0. Let V_0 be the price. We assume the risk free rate to be zero, $r = 0$.
(i) First, assume only that $K_2 < s_0 < K_1$ and write the price.

Mock Exam 6 III

(ii) Assume further that s_0 is the mean of the two strikes K_2 and K_1, $s_0 = (K_2 + K_1)/2$. Simplify the price V_0 of the BuCS under Bachelier in this case and prove that $V_0 > 0$.

d) Assuming only (i) $K_2 < s_0 < K_1$, compute the Delta of the BuCS in the Bachelier model, namely $\frac{\partial V_0}{s_0}$. Discuss the sign of the Delta and deduce the behaviour of the BCS price with respect to the underlying stock S_0.

Mock Exam 6 IV

Solutions of Mock Exam 6. Problem 3: Bull Call Spread with the Bachelier models.
a) This question has been answered in Problem 2 of this same mock exam. Refer to the solution given there.

b) The Bachelier model under the measure P postulates the stock dynamics

$$dS_t = \mu dt + \sigma dW_t, \quad S_0 = s_0$$

where μ is the drift, a real constant, $\sigma > 0$ is the volatility, a real constant, and s_0 is a positive deterministic initial condition. This is an arithmetic Brownian motion and as such it has a normal distribution. Indeed, in the Bachelier model the stock is normally distributed, and therefore has the disadvantage of being allowed to take negative values with positive probabilities.

Mock Exam 6 V

To move under the risk neutral measure, we would need to impose the drift rS_t and obtain a model under Q that reads

$$dS_t = rS_t dt + \sigma dW_t^Q, \quad S_0 = s_0.$$

This, however, is no longer an arithmetic Brownian motion, but has become a special case of an Ornstein Uhlenbeck process. It is not desirable to have two different types of processes under the two measures, so the only way to keep an arithmetic Brownian motion under Q is to assume $r = 0$. In that case the Bachelier model under Q reads

$$dS_t = \sigma dW_t^Q, \quad S_0 = s_0$$

and is still an arithmetic Brownian motion.
The *relative* or return volatility function of Bachelier, $\frac{\sigma}{S_t}$, as in $dS_t = \frac{\sigma}{S_t} S_t dW_t^Q$, is a decreasing function of S, whereas the relative

Mock Exam 6 VI

volatility function ν in Black Scholes $dS_t = \nu S_t dW_t^Q$ is a constant ν. When we define the smile, we match the price of our Bachelier model with the price given by a Black Scholes formula, solving in the relative volatility ν of Black Scholes for each strike. So the relative volatility is the correct variable to consider.

The Black Scholes relative volatility is constant, the Bachelier relative volatility is a decreasing function of S, that gets smaller and smaller as K increases when we ask $S > K$ in the call option. Therefore, for larger and larger K, the return volatility in Bachelier being smaller and smaller, we will have a smaller ν in Black Scholes to match it, or in other terms as the strike increases the matching vol ν decreases because the return vol in the Bachelier model is smaller. The smile is expected to be decreasing.

Mock Exam 6 VII

If we observe a decreasing smile, the other models consistent with this are

1. Displaced Diffusion (DD) with negative shift;
2. CEV with exponent smaller than 1.

However we know that CEV requires special functions and is less tractable than DD and Bachelier.

In this case we take the Bachelier model as the problem requires.

c) We price the BuCS by computing the difference between two call prices with different strikes in the Bachelier model. Recall the forula for a call option in the Bachelier model.

Mock Exam 6 VIII

$$V_{BaM}(0, s_0, K, T, \sigma) = (s_0 - K)\Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T} p_N\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right).$$

To obtain the BuCS price, we need to make the difference of a call with strike K_2 minus a call with strike K_1. We get

$$V_0^{BuCS} = (s_0 - K_2)\Phi\left(\frac{s_0 - K_2}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T} p_N\left(\frac{s_0 - K_2}{\sigma\sqrt{T}}\right)$$

$$-(s_0 - K_1)\Phi\left(\frac{s_0 - K_1}{\sigma\sqrt{T}}\right) - \sigma\sqrt{T} p_N\left(\frac{s_0 - K_1}{\sigma\sqrt{T}}\right)$$

Recalling that $s_0 = (K_1 + K_2)/2$, we get that

$$s_0 - K_1 = \frac{K_1 + K_2}{2} - K_1 = -\frac{K_1 - K_2}{2} =: -a,$$

Mock Exam 6 IX

$$s_0 - K_2 = \frac{K_1 + K_2}{2} - K_2 = \frac{K_1 - K_2}{2} =: a, \quad a > 0,$$

where we introduced the constant $a = \frac{K_1 - K_2}{2} = s_0 - K_2 = -(s_0 - K_1)$. Substituting

$$V_0^{BuCS} = a\Phi\left(\frac{a}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}p_N\left(\frac{a}{\sigma\sqrt{T}}\right)$$

$$-(-a)\Phi\left(\frac{-a}{\sigma\sqrt{T}}\right) - \sigma\sqrt{T}p_N\left(\frac{-a}{\sigma\sqrt{T}}\right) = \dots$$

Now we recall that p_N is symmetric, so $p_N(-x) = p_N(x)$ and $p_N(x) - p_N(-x) = 0$, and therefore the p_N terms in the above formula

Mock Exam 6 X

cancel. As for the Φ terms, we recall $\Phi(-x) = 1 - \Phi(x)$. Including all this the last formula simplifies to

$$\ldots = a\Phi\left(\frac{a}{\sigma\sqrt{T}}\right) + a\left(1 - \Phi\left(\frac{a}{\sigma\sqrt{T}}\right)\right) = a = \frac{K_1 - K_2}{2} > 0$$

as $K_1 > K_2$. This proves that the price is positive.
Therefore we can conclude that $V_0^{BuCS} > 0$ in this particular case (ii).

d) We now compute the delta of the BuCS under the Bachelier model. Given that we have the difference of two call options, it is convenient first to compute the delta of a call option in the Bachelier model.
This has been done in the derivation of the Bachelier model in the lecture notes. If you remember the formula by heart, you are allowed to quote it directly without proof, see Eq. (32). If you don't remember you can derive it, see the derivation of (32) in the Bachelier model section.

Mock Exam 6 XI

We have

$$\frac{\partial V_{BaM}(0)}{\partial s_0} = \frac{\partial \left((s_0 - K)\Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)\right)}{\partial s_0} + \sigma\sqrt{T}\frac{\partial\, p_N\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)}{\partial s_0} =$$

$$= \Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right).$$

As the BuCS is the difference between two call options with strikes K_2 and K_1 respectively, with $K_2 < s_0 < K_1$, the delta of the BuCS will be the difference of the two deltas of the call options. We get

$$\frac{\partial V_{BaM}^{BuCS}(0)}{\partial s_0} = \Phi\left(\frac{s_0 - K_2}{\sigma\sqrt{T}}\right) - \Phi\left(\frac{s_0 - K_1}{\sigma\sqrt{T}}\right).$$

Mock Exam 6 XII

As Φ is an increasing function and $K_1 > K_2$, we have $s_0 - K_1 < s_0 - K_2$ and hence the delta will be positive, as the argument of the first Φ is larger than the argument of the subtracted Φ. Hence in the Bachelier model, the value of a BuCS grows with the underlying stock s_0 as the derivative is always positive. This is intuitive looking at the payoff, as we observed in point 2 for Black Scholes, since as the stock grows the payoff becomes more profitable.

Mock Exam 6 I

Consider a porfolio with a short zero-coupon bond with notional N and maturity T, and a long amount N of equity forward contract on stock S with strike K and maturity T. The payoff at time T is

$$Y = -N\,1 + N(S_T - K).$$

The stock is assumed to follow the Black Scholes model

$$dS_t = \mu S_t dt + \sigma S_t dW_t,\ s_0$$

under the measure P. We assume a constant risk free rate $r > 0$.
a) Compute the portfolio $VaR_{H,\alpha}$
b) How sensitive is VaR to the stock volatility? Give a quantitative measure of this sensitivity and comment on its sign and on what this imply on how VaR behaves with respect to σ.
c) What is the limit VaR when (i) $\mu \downarrow -\infty$ and (ii) $\mu \uparrow +\infty$? Examine how both limits depend on the confidence level α and discuss.

Mock Exam 6 II

Solutions.

a) The loss distribution at H is the portfolio value at time 0 minus the portfolio value at time H, namely

$$L_H = -Ne^{-r(T-0)} + N(S_0 - Ke^{-r(T-0)}) - [-Ne^{-r(T-H)} + N(S_H - Ke^{-r(T-H)})]$$

$$= -N(K+1)(e^{-rT} - e^{-r(T-H)}) + NS_0 - NS_H = A - NS_H$$

where $A = N(K+1)(e^{-r(T-H)} - e^{-rT}) + NS_0$ is deterministic. Note that $A > 0$ as $e^{-r(T-H)} > e^{-rT}$. Here we used the fact that the price of a unit-notional zero coupon bond with maturity T at time $t < T$ is

$$P(t,T) = E_t^Q[e^{-r(T-t)}1] = e^{-r(T-t)}$$

while the price of the forward contract with maturity T at time $t < T$ is

$$E_t^Q[e^{-r(T-t)}(S_T - K)] = S_t - Ke^{-r(T-t)}.$$

Mock Exam 6 III

$\mathrm{VaR}_{H,\alpha}$ is defined as the quantity q satisfying

$$\alpha = P[L_H < q] = P[A - NS_H < q] = P[S_H > (A-q)/N] =$$

$$= P[S_0 \exp((\mu - \sigma^2/2)H + \sigma W_H) > (A-q)/N] =$$

$$= P[\exp((\mu - \sigma^2/2)H + \sigma W_H) > (A-q)/(NS_0)] = \dots$$

where we have used the usual solution for the Black Scholes geometric-Brownian-motion SDE. We know that $A > 0$. We temporarily assume $A - q > 0$ so we can take logs on both sides in the last P expression, and we will have to check that this condition is satisfied a posteriori, once we have found q. We have

$$\dots = P\left[W_H > \frac{\ln[(A-q)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma}\right] =$$

Mock Exam 6 IV

$$= P\left[\sqrt{H}\mathcal{N} > \frac{\ln[(A-q)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma}\right]$$

$$= P\left[\mathcal{N} > \frac{\ln[(A-q)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma\sqrt{H}}\right]$$

Recalling that $P[\mathcal{N} > x] = 1 - \Phi(x) = \Phi(-x)$, we have

$$\alpha = \Phi\left(-\frac{\ln[(A-q)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma\sqrt{H}}\right).$$

Take Φ^{-1} on both sides:

$$\Phi^{-1}(\alpha) = -\frac{\ln[(A-q)/(NS_0)] - (\mu - \sigma^2/2)H}{\sigma\sqrt{H}},$$

Mock Exam 6 V

and solve in q:

$$q = A - NS_0 \exp\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2)H\right)$$

This is our VaR. Recall we had assumed $A - q > 0$ to be positive. Let's check it is positive indeed. $A - q$, from the last expression, turns out to be an exponential and as such it is always positive.

b) We see that VaR depends on the volatility through the first and second terms inside exp:

$$VaR_{H,\alpha} = A - NS_0 \exp\left(-\Phi^{-1}(\alpha)\boxed{\sigma}\sqrt{H} + (\mu - \boxed{\sigma^2}/2)H\right)$$

To quantify the sensitivity of VaR to σ we compute

$$\frac{\partial VaR}{\partial \sigma} = -NS_0 \exp\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \frac{\sigma^2}{2})H\right)\frac{\partial}{\partial \sigma}\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H} - \frac{\sigma^2 H}{2}\right) =$$

Mock Exam 6 VI

$$= -NS_0(-\Phi^{-1}(\alpha)\sqrt{H} - \sigma H)\exp\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2)H\right) =$$

$$= NS_0(\Phi^{-1}(\alpha)\sqrt{H} + \sigma H)\exp\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2)H\right) =$$

We investigate the sign of this sensitivity.
The sensitivity will be positive if we have

$$\Phi^{-1}(\alpha)\sqrt{H} + \sigma H > 0.$$

For all usual values of α like $90\% = 0.9,\ 95\% = 0.95,\ 99\% = 0.99$ we have that $\Phi^{-1}(\alpha) > 0$, so that, given that $\sigma > 0$ and $H > 0$, the above condition is always satisfied. The VaR sensitivity to σ will therefore always be positive, meaning that VaR will always increase with σ.
Recall the loss,

$$L_H = A - NS_H.$$

Mock Exam 6 VII

The only random part in the loss is S_H and A does not depend on σ. Therefore by increasing σ we make S_H more volatile, hance more prone to take either smaller or larger values. As we are short S_H, which comes with a minus sign, smaller values of S_H triggered by larger volatilities σ will correspond to potentially larger losses, and hence larger VaR.

c) If we check our expression for VaR,

$$VaR_{H,\alpha} = A - NS_0 \exp\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\boxed{\mu} - \sigma^2/2)H\right)$$

we see it is continuous in μ. μ appears only as an argument of the exponential function in the VaR expression. We can take the two limits for μ and see what happens.

Mock Exam 6 VIII

(i) We can see what happens for $\mu \downarrow -\infty$ by calculating the limit

$$VaR_{H,\alpha}|_{\mu\downarrow-\infty} = \lim_{\mu\downarrow-\infty}\left[A - NS_0 \exp\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2)H\right)\right] = \ldots$$

Consider

$$\lim_{\mu\downarrow-\infty}(\mu - \sigma^2/2) = -\infty$$

So the total limit is a limit of A minus an exponential whose argument tends to minus infinity, so that

$$\ldots = A$$

since the exponent in the exponential tends to $-\infty$ and $e^{-\infty}$ tends to zero. Hence the limit is $A = N(K+1)(e^{-r(T-H)} - e^{-rT}) + NS_0$. This limit does not depend on α. So when μ goes to minus infinity, meaning

Mock Exam 6 IX

that the trend of the risky asset S goes to minus infinity, $VaR_{H,\alpha}$ does not depend on α anymore.
Intuitively and not fully rigorously, this is because the infinite downward trend makes all confidence levels the same, as the stock will always go down to zero by infinitely negative trend in a Black-Scholes - Geometric Brownian motion models. Indeed, recall that the loss is

$$L_H = A - NS_H$$

and that

$$S_H = S_0 \exp\left((\mu - \sigma^2/2)H + \sigma W_H\right)$$

and when $\mu \downarrow -\infty$ this S_H goes to zero for any fixed finite realization of the Brownian motion. So the loss goes

$$L_H|_{\mu\downarrow-\infty} = A - NS_H|_{\mu\downarrow-\infty} = A - 0 = A$$

Mock Exam 6 X

and basically the loss becomes a positive deterministic constant A. As it is deterministic, the loss will be A no matter the confidence level, and VaR will always be A.

(ii) When $\mu \uparrow +\infty$ we get

$$VaR_{H,\alpha}|_{\mu\uparrow\infty} = \lim_{\mu\uparrow\infty}\left[A - NS_0 \exp\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H} + (\mu - \sigma^2/2)H\right)\right] = \dots$$

Consider

$$\lim_{\mu\uparrow\infty}(\mu - \sigma^2/2) = +\infty$$

So the total limit is

$$\dots = -\infty$$

since the exponent in the exponential tends to ∞ and e^{∞} tends to infinity. Hence the VaR limit is $-\infty$. This limit does not depend on α

Mock Exam 6 XI

either. So when μ goes to plus infinity, meaning that the trend of the risky asset S goes to infinity, $VaR_{H,\alpha}$ does not depend on α anymore. Again intuitively and not fully rigorously, this is because the infinite trend makes the stock S_H grow indefinitely,

$$S_H = S_0 \exp\left((\mu - \sigma^2/2)H + \sigma W_H\right)$$

will tend to infinity for $\mu \uparrow \infty$ for any finite fixed realization of W_H, and that makes the loss

$$L_H|_{\mu\uparrow\infty} = A - NS_H|_{\mu\uparrow\infty} = A - \infty = -\infty$$

infinitely negative. This corresponds to the intuition that we are long a forward contract $S - K$, and if S goes to infinity due to the trend, this gives us an infinite gain, and an infinite gain corresponds to a minus infinite loss, that is always minus infinite and thus VaR does not depend on the confidence level.

Mock Exam 7 I

Problem 1. CIR SDE: $dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$.

Consider the Cox Ingersoll Ross or Feller square root process, under the measure P,

$$dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$$

where W is a standard Brownian motion under P, with deterministic initial condition x_0 and where k, θ, σ are positive real constants with $2k\theta > \sigma^2$, ensuring strictly positive solutions. This SDE admits a unique global solution but you are not required to prove it, as this is not possible with the theorem we saw in the course.

a) Derive an expression for the mean at time 0 of the solution at a time $T > 0$, namely calculate $E_0^P[X_T]$. Calculate also $\lim_{T\uparrow\infty} E[X_T]$ and comment on the name "long term mean" that is usually given to the parameter θ.

Mock Exam 7 II

b) Using Ito's formula, derive the Ito differential for X_t^2. [Hint: in this case you are just required to write the Ito differential for X_t^2, not to write an SDE $dY_t = \mu(t, Y_t)dt + \sigma(t, Y_t)dW_t$ for the process $Y = X_t^2$]

c) Derive an expression for the variance at time 0 of the solution X at a time $T > 0$, namely calculate $\text{var}_0^P[X_T]$. Reach an expression involving only elementary functions and operations like exponentials, fractions, sums and products. You don't need to simplify the final expression.

Solutions

a) To derive an expression for the mean, we write the SDE as an integral equation, which is incidentally its true form, by integrating both sides

$$\int_0^T dX_t = \int_0^T k(\theta - X_t)dt + \int_0^T \sigma\sqrt{X_t}dW_t$$

Mock Exam 7 III

leading to

$$X_T - x_0 = \int_0^T k(\theta - X_t)dt + \int_0^T \sigma\sqrt{X_t}dW_t$$

or

$$X_T = x_0 + \int_0^T k(\theta - X_t)dt + \int_0^T \sigma\sqrt{X_t}dW_t.$$

Now take the expected value $E = E_0^P$ at time 0 on both sides

$$E[X_T] = x_0 + E\left[\int_0^T k(\theta - X_t)dt\right] + E\left[\int_0^T \sigma\sqrt{X_t}dW_t\right] =$$

$$= x_0 + \int_0^T E[k(\theta - X_t)]dt + 0 =$$

Mock Exam 7 IV

where we used Fubini's theorem to move the expected value inside the time integral [you are not required to check the assumptions of Fubini's theorem here, you can just use it], we have used the fact that x_0 is deterministic, so its expectation is still x_0, and we have use the zero mean property of the Ito integral, whose expectation is zero.

$$= x_0 + \int_0^T k(\theta - E[X_t])dt$$

where we used the fact that the expectation is a linear operator. We have reached the equation

$$E[X_T] = x_0 + \int_0^T k(\theta - E[X_t])dt.$$

Mock Exam 7 V

To solve this equation, we define the time function $m_t = E[X_t]$ for all $t \geq 0$, so that the above equation is written as

$$m_T = x_0 + \int_0^T k(\theta - m_t)dt.$$

We now differentiate both sides of this last equation with respect to T:

$$\frac{d}{dT} m_T = \frac{d}{dT} \int_0^T k(\theta - m_t)dt.$$

Knowing that the derivative of an integral with respect to the integration upper bound is the integrand valued at the upper bound (fundamental theorem of calculus), we get

$$\frac{d}{dT} m_T = k(\theta - m_T).$$

Mock Exam 7 VI

This is a linear affine ODE,

$$\frac{d}{dT}m_T = k\theta - km_T$$

whose solution we know from the general linear affine ODE,

$$\frac{dm(t)}{dt} = B(t) - A(t)m(t),$$

where A and B are functions of time. The solution is given in any textbook as

$$m(t) = \exp\left(-\int_0^t A(s)ds\right)\left[\int_0^t \exp\left(\int_0^u A(s)ds\right)B(u)du + m(0)\right].$$

Mock Exam 7 VII

So in our case this case reads

$$m_T = m_0 e^{-kT} + \theta(1 - e^{-kT})$$

So we conclude, as $m_0 = E_0[x_0] = x_0$,

$$E[X_T] = x_0 e^{-kT} + \theta(1 - e^{-kT}).$$

We note that if $T \uparrow \infty$ the mean tends to θ, as both exponentials tend to zero, so

$$\lim_{T\uparrow\infty} E[X_T] = \lim_{T\uparrow\infty} [x_0 e^{-kT} + \theta(1 - e^{-kT})] = x_0 * 0 + \theta * (1 - 0) = \theta$$

which indeed justifies the name “long term mean” for θ, as it is the mean of the process after an infinite (long term) time.

Mock Exam 7 VIII

b) We apply Ito's formula to the square transformation $\phi(t, X_t) = X_t^2$. We have

$$d\phi(t, X_t) = \frac{\partial}{\partial t}\phi(t, X_t)dt + \frac{\partial}{\partial X}\phi(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2}{\partial X^2}\phi(t, X_t)dX_t dX_t.$$

As $\phi(t, x) = x^2$ we have

$$\frac{\partial}{\partial t}\phi(t, x) = 0,\ \frac{\partial}{\partial x}\phi(t, x) = 2x,\ \ \frac{\partial^2}{\partial x^2}\phi(t, x) = 2$$

leading to

$$d\phi(t, X_t) = 2X_t dX_t + \frac{1}{2}2dX_t dX_t$$

or, substituting dX,

$$d(X_t^2) = 2X_t[k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t] + dX_t dX_t = \dots$$

Mock Exam 7 IX

Now,

$$dX_t dX_t = \sigma\sqrt{X_t}dW_t\sigma\sqrt{X_t}dW_t = \sigma^2 X_t dt$$

where we used $dt\ dt = 0$, $dt\ dW_t = 0$ and $dW_t\ dW_t = dt$, so that we conclude

$$\ldots = 2X_t[k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t] + \sigma^2 X_t dt$$

Rearranging we have

$$d(X_t^2) = [(2k\theta + \sigma^2)X_t - 2kX_t^2]dt + 2\sigma X_t\sqrt{X_t}dW_t.$$

Mock Exam 7 X

NOTE (not required as part of the solution). If the question had been to write an SDE for $Y_t = X^2$, we would have to substitute in the last equation $X_t = \sqrt{Y_t}$ (we take the positive square root as we know X is always positive) and $\sqrt{X_t} = Y_t^{1/4}$ (where again we used the fact that X is positive, so there is no problem in taking the square root). Then, substituting $X^2 = Y$ and $X = \sqrt{Y}$ and $\sqrt{X} = Y^{1/4}$ would yield the SDE ($\sqrt{Y}Y^{1/4} = Y^{3/4}$)

$$dY_t = [(2k\theta + \sigma^2)\sqrt{Y_t} - 2kY_t^2]dt + 2\sigma Y_t^{3/4} dW_t.$$

Mock Exam 7 XI

c) To compute

$$\mathrm{var}_0^P[X_T] = E_0^P[X_T^2] - E_0^P[X_T]^2$$

we need $E[X_T^2]$, whereas $E[X_T]$ we already have from point a). To compute $E[X_T^2]$ we use the result in point b). Recall that

$$d(X_t^2) = [(2k\theta + \sigma^2)X_t - 2kX_t^2]dt + 2\sigma X_t\sqrt{X_t}dW_t.$$

Now integrate between 0 and T on both sides:

$$\int_0^T d(X_t^2) = \int_0^T [(2k\theta + \sigma^2)X_t - 2kX_t^2]dt + \int_0^T 2\sigma X_t\sqrt{X_t}dW_t,$$

or

$$X_T^2 - x_0^2 = (2k\theta + \sigma^2)\int_0^T X_t dt - 2k\int_0^T X_t^2 dt + \int_0^T 2\sigma X_t\sqrt{X_t}dW_t,$$

Mock Exam 7 XII

and take expectations on both sides

$$E[X_T^2] = x_0^2 + (2k\theta + \sigma^2)E\left[\int_0^T X_t dt\right] - 2kE\left[\int_0^T X_t^2 dt\right] + 0$$

where we used linearity of the expectation and the fact that the expectation of the Ito integral is zero. We now use Fubini's theorem to switch the expectations inside the time integrals:

$$E[X_T^2] = x_0^2 + (2k\theta + \sigma^2)\int_0^T E[X_t]dt - 2k\int_0^T E[X_t^2]dt$$

and finally, we recall that we already have $E[X_t]$ from point a), call it $m(t) = E[X_t] = x_0 e^{-kt} + \theta(1 - e^{-kt})$. So we can write

$$E[X_T^2] = x_0^2 + (2k\theta + \sigma^2)\int_0^T m(t)dt - 2k\int_0^T E[X_t^2]dt$$

Mock Exam 7 XIII

Let us now call $s(T) = E[X_T^2]$ so as to rewrite the last equation as

$$s(T) = x_0^2 + (2k\theta + \sigma^2)\int_0^T m(t)dt - 2k\int_0^T s(t)dt$$

Differentiate wrt T both sides to get

$$\frac{d}{dT}s(T) = 0 + (2k\theta + \sigma^2)\frac{d}{dT}\int_0^T m(t)dt - 2k\frac{d}{dT}\int_0^T s(t)dt$$

where we used the fact that the derivative of a constant is zero, the derivative is linear and now we are going to use the fundamental theorem of calculus again, the derivative of an integral with respect to the upper bound of integration is the integrand evaluated at the upper bound:

$$\frac{d}{dT}s(T) = 0 + (2k\theta + \sigma^2)m(T) - 2ks(T)$$

Mock Exam 7 XIV

Given that $m(T)$ is known, this is a linear affine ODE for $s(T)$ of the type

$$\frac{ds(T)}{dT} = B(T) - A(T)s(T),$$

where A and B are functions of time:

$$B(T) = (2k\theta + \sigma^2)m(T), \quad A(T) = 2k$$

with the textbook solution

$$s(T) = \exp\left(-\int_0^T A(s)ds\right)\left[\int_0^T \exp\left(\int_0^u A(s)ds\right) B(u)du + s(0)\right],$$

which in our case reads

$$s(T) = \exp(-2kT)\left[\int_0^T \exp(2k\ u)\,(2k\theta + \sigma^2)m(u)du + s(0)\right]$$

Mock Exam 7 XV

or

$$s(T) = \exp(-2kT)\left[(2k\theta + \sigma^2)\int_0^T \exp(2k\ u)\, m(u)du + s(0)\right] \quad (46)$$

Let's calculate the integral.

$$\int_0^T \exp(2k\ u)\, m(u)du = \int_0^T \exp(2k\ u)\,[x_0 e^{-ku} + \theta(1 - e^{-ku})]du =$$

$$= \int_0^T \exp(2k\ u)\,[(x_0 - \theta)e^{-ku} + \theta]du = \int_0^T [(x_0 - \theta)e^{ku} + \theta\ e^{2ku}]du =$$

$$= (x_0 - \theta)\frac{1}{k}(e^{kT} - 1) + \theta\frac{1}{2k}(e^{2kT} - 1)$$

Mock Exam 7 XVI

and now substitute in (46):

$$s(T) = e^{-2kT}\left[(2k\theta+\sigma^2)\left(\frac{x_0-\theta}{k}(e^{kT}-1)+\frac{\theta}{2k}(e^{2kT}-1)\right)+x_0^2\right]$$

where we rearranged terms and used $s(0) = E[X_0^2] = E[x_0^2] = x_0^2$ as $X_0 = x_0$ is deterministic.
We can now compute the variance as

$$\mathrm{var}(X_T) = E[X_T^2] - E[X_T]^2 = s(T) - m(T)^2 =$$

$$= e^{-2kT}\left[(2k\theta+\sigma^2)\left(\frac{x_0-\theta}{k}(e^{kT}-1)+\frac{\theta}{2k}(e^{2kT}-1)\right)+x_0^2\right]+$$

$$-(x_0e^{-kT}+\theta(1-e^{-kT}))^2$$

Mock Exam 7 XVII

The above would be accepted as a final answer for this particular question, because the exam is not long enough to do all calculations. However, if one continued, after long and laborious calculations, one would be able to show that the above expression simplifies with the expression given in the lecture notes:

$$\mathrm{var}(X_T) = x_0 \frac{\sigma^2}{k}(e^{-kT} - e^{-2kT}) + \theta \frac{\sigma^2}{2k}(1 - e^{-kT})^2,$$

both expressions being amenable to

$$\mathrm{var}(X_T) = \frac{\theta\sigma^2}{2k} + \frac{\theta\sigma^2 e^{-2kT}}{2k} - \frac{\theta\sigma^2 e^{-kT}}{k} - \frac{\sigma^2 x_0 e^{-2kT}}{k} + \frac{\sigma^2 x_0 e^{-kT}}{k}$$

Mock Exam 7 I

Problem 2. Long Call Calendar Spread (LCCS) in the Black Scholes model.

A long call calendar spread is a strategy where we sell a first call option and buy a second call option on the same stock and with the same strike, typically at-the-money although here we will assume a general strike K. The call we buy has a longer maturity T_1 than the call we sell, with maturity T_2. In practice, maturities are very short, for example the call we sell is close to expiration, with a maturity say of one week. The call we buy has a maturity of a month.
Given that we have two different maturities, there is no way to plot the payoff for a single maturity.
We take the longer maturity for the call we buy as T_1. The call we sell has maturity $T_2 < T_1$. This is not a payoff with a simple maturity but it's the combination of two payoffs with different maturities. Still, we can

Mock Exam 7 II

write the *discounted payoff* at time 0 up to maturity T_1 (sum of all the discounted cash flows in the portfolio up to time T_1) as

$$\Pi(0, T_1) = D(0, T_1)(S_{T_1} - K)^+ - D(0, T_2)(S_{T_2} - K)^+$$

a) Write the price of the LCCS at time 0 in a Black Scholes model with the following dynamics under the risk neutral measure Q,

$$dS_t = rS_t dt + \sigma S_t dW_t^Q,$$

with initial stock price S_0, risk free rate r and strike K.

b) Write the delta of the LCSS, namely, if V_0 is the value of the LCSS from point a), write a formula for $\frac{\partial V_0}{\partial S_0}$.

Mock Exam 7 III

c) Analyze the sign of the LCSS delta, finding whether the delta of the LCSS you found in point b) is always positive, always negative, or can be either positive or negative depending on the parameters S_0, K, σ, r. If both cases are possible, find sufficient conditions based on the parameters S_0 and K for the LCCS Delta to be positive.

SOLUTIONS.

a) The price of the portfolio is the risk neutral expectation of the discounted cash flows, hence

$$V_0 = E_0^Q[D(0,T_1)(S_{T_1} - K)^+ - D(0,T_2)(S_{T_2} - K)^+] =$$

$$= E_0^Q[D(0,T_1)(S_{T_1} - K)^+] - E_0^Q[D(0,T_2)(S_{T_2} - K)^+] =$$

$$= E_0^Q[e^{-rT_1}(S_{T_1} - K)^+] - E_0^Q[e^{-rT_2}(S_{T_2} - K)^+] =$$

Mock Exam 7 IV

$$= V_{BS}^{Call}(0, S_0, K, T_1, \sigma, r) - V_{BS}^{call}(0, S_0, K, T_2, \sigma, r) = \dots$$

as S follows a Black Scholes model. Using the call option formula in Black Scholes we have

$$\dots = S_0\Phi(d_1(T_1)) - Ke^{-rT_1}\Phi(d_2(T_1)) - S_0\Phi(d_1(T_2)) + Ke^{-rT_2}\Phi(d_2(T_2)) =$$

$$= S_0(\Phi(d_1(T_1)) - \Phi(d_1(T_2)) - K(e^{-rT_1}\Phi(d_2(T_1)) - e^{-rT_2}\Phi(d_2(T_2))$$

where we define

$$d_{1,2}(T) := \frac{\ln(S_0/K) + (r \pm \sigma^2/2)T}{\sigma\sqrt{T}}.$$

b) We start from

$$V_0 = V_{BS}^{Call}(0, S_0, K, T_1, \sigma, r) - V_{BS}^{call}(0, S_0, K, T_2, \sigma, r)$$

Mock Exam 7 V

So the delta is

$$\frac{\partial V_0}{\partial S_0} = \frac{\partial}{\partial S_0} V_{BS}^{Call}(0, S_0, K, T_1, \sigma, r) - \frac{\partial}{\partial S_0} V_{BS}^{call}(0, S_0, K, T_2, \sigma, r) = \ldots$$

and we know that the delta for a call in a Black Scholes model is $\Phi(d_1)$, so we obtain

$$\ldots = \Phi(d_1(T_1)) - \Phi(d_1(T_2))$$

c)
To try answering this question we try to determine whether the delta of a call with maturity T in Black Scholes is an increasing or decreasing function of T. To do this, we calculate the derivative wrt T and check

Mock Exam 7 VI

its sign. For a call with maturity T, the delta in Black Scholes is $\Phi(d_1(T))$ so we can calculate

$$\frac{\partial \Phi(d_1(T))}{\partial T} = \Phi'(d_1(T))\frac{\partial d_1(T)}{\partial T} = \phi(d_1(T))\frac{\partial d_1(T)}{\partial T} = \dots$$

where ϕ is the density function of the standard normal. Calculate

$$\frac{\partial d_1(T)}{\partial T} = \frac{\partial}{\partial T}\left[\frac{\ln(S_0/K) + (r+\sigma^2/2)T}{\sigma\sqrt{T}}\right] =$$

$$= \frac{\partial}{\partial T}\left[\frac{\ln(S_0/K)}{\sqrt{T}} + (r+\sigma^2/2)\sqrt{T}\right] = -\frac{\ln(S_0/K)}{2\sqrt{T^3}} + \frac{r+\sigma^2/2}{2\sqrt{T}}$$

Substituting

$$\dots = \phi(d_1(T))\left[-\frac{\ln(S_0/K)}{2\sqrt{T^3}} + \frac{r+\sigma^2/2}{2\sqrt{T}}\right].$$

Mock Exam 7 VII

Now, the term ϕ is a normal density and as such is always positive. The sign then depends on the term between squared brackets. The partial derivative of the delta with respect to T will be positive if $\ln(S_0/K) \leq 0$, namely if $S_0 \leq K$, i.e. if the call option is out of the money or at-the-money. This is because all the other terms in the summation giving the partial derivative of the call delta are positive as well. If the option is in-the-money, we have $S_0 > K$ and thus $\ln(S_0/K) > 0$, so the delta can be either positive or negative.

If $S_0 \leq K$ then the delta is increasing in T, so the delta of the LCCS being $\Phi(d_1(T_1)) - \Phi(d_1(T_2))$ will be positive because $T_1 > T_2$ implies $\Phi(d_1(T_1)) > \Phi(d_1(T_2))$, the delta of the call being increasing. So the sufficient condition for the delta of the LCCS to be positive is $S_0 \leq K$. In the other case $S_0 > K$ both cases are possible.

Mock Exam 7 I

Problem 3. Long Call Calendar Spread (LCCS) in the Bachelier model.

A long call calendar spread is a strategy... [see description on Exam 7 Problem 2]... having the *discounted payoff* at time 0 up to maturity $T_1 > T_2$ (sum of all the discounted cash flows in the portfolio up to time T_1) as

$$\Pi(0, T_1) = D(0, T_1)(S_{T_1} - K)^+ - D(0, T_2)(S_{T_2} - K)^+$$

a) Write the price of the LCCS at time 0 in a Bachelier model with the following dynamics under the risk neutral measure Q,

$$dS_t = \sigma dW_t^Q,$$

with initial stock price $S_0 = s_0$, risk free rate $r = 0$ and strike K.

Mock Exam 7 II

b) Write the delta of the LCSS in the Bachelier model, namely, if V_0 is the value of the LCSS from point a), write a formula for $\frac{\partial V_0}{\partial s_0}$.

c) Analyze the sign of the LCSS delta, finding whether the delta of the LCSS you found in point b) is always positive, always negative, or can be either positive or negative depending on the parameters s_0 and K. If both cases are possible, find sufficient conditions based on the parameters s_0 and K for the LCCS Delta in the Bachelier model to be positive.

SOLUTIONS.

a) The price of the portfolio is the risk neutral expectation of the discounted cash flows, hence

$$V_0 = E_0^Q[D(0,T_1)(S_{T_1} - K)^+ - D(0,T_2)(S_{T_2} - K)^+] =$$

Mock Exam 7 III

Remembering $r = 0$ we have that $D(0, T) = e^{-rT} = 1$ for all T, so

$$= E_0^Q[(S_{T_1} - K)^+] - E_0^Q[(S_{T_2} - K)^+] =$$

$$= V_{BaM}^{Call}(0, s_0, K, T_1, \sigma) - V_{BaM}^{call}(0, s_0, K, T_2, \sigma) = \dots$$

as S follows a Bachelier model (BaM). Using the call option formula in Bachelier (write it down or if you don't remember it derive it) for a call with maturity T

$$V_{BaM}(s_0, K, T, \sigma) = (s_0 - K)\Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T} p_N\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right)$$

we have

Mock Exam 7 IV

$$\ldots = (s_0 - K)\Phi\left(\frac{s_0 - K}{\sigma\sqrt{T_1}}\right) + \sigma\sqrt{T_1}p_N\left(\frac{s_0 - K}{\sigma\sqrt{T_1}}\right) +$$

$$-(s_0 - K)\Phi\left(\frac{s_0 - K}{\sigma\sqrt{T_2}}\right) - \sigma\sqrt{T_2}p_N\left(\frac{s_0 - K}{\sigma\sqrt{T_2}}\right)$$

b) We start from

$$V_0 = V_{BaM}^{Call}(0, s_0, K, T_1, \sigma) - V_{BaM}^{call}(0, s_0, K, T_2, \sigma)$$

So the delta is

$$\frac{\partial V_0}{\partial s_0} = \frac{\partial}{\partial s_0} V_{BaM}^{Call}(0, s_0, K, T_1, \sigma) - \frac{\partial}{\partial s_0} V_{BaM}^{call}(s_0, K, T_2, \sigma) = \ldots$$

Mock Exam 7 V

and we know that the delta for a call with maturity T in a Bachelier Model is

$$\frac{\partial V_{BaM}(0)}{\partial s_0} = \Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right)$$

(if you don't remember it derive it as in Exam 6 Problem 3 point d)) so that

$$\ldots = \Phi\left(\frac{s_0 - K}{\sigma\sqrt{T_1}}\right) - \Phi\left(\frac{s_0 - K}{\sigma\sqrt{T_2}}\right)$$

c)

To try answering this question we try to determine whether the delta of a call with maturity T in Bachelier is an increasing or decreasing function of T. To do this, we calculate the derivative wrt T of the delta

Mock Exam 7 VI

and check its sign. For a call with maturity T, the delta in Bachelier is $\Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)$ so we can calculate

$$\frac{\partial}{\partial T}\Phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)=\Phi'\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)\frac{\partial}{\partial T}\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)=\dots$$

Calculate

$$\frac{\partial}{\partial T}\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)=-\left(\frac{s_0-K}{2\sqrt{T^3}\sigma}\right)$$

and substituting

$$\dots=-\phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)\left[\frac{s_0-K}{2\sqrt{T^3}\sigma}\right]=\phi\left(\frac{s_0-K}{\sigma\sqrt{T}}\right)\left[\frac{K-s_0}{2\sqrt{T^3}\sigma}\right]$$

where ϕ is the standard normal density function.

Mock Exam 7 VII

Now, the term ϕ is a normal density and as such is always positive. The sign then depends on the term between squared brackets. The partial derivative of the delta with respect to T will be positive if $S_0 < K$, zero if $S_0 = K$, and negative if $S_0 > K$.

So if $S_0 < K$ the delta of a call in Bachelier increases with T, and therefore the LCSS delta begin the difference of call deltas with T_1 and T_2 respectively, with $T_1 > T_2$, will be positive. [Note, not needed as part of the answer, in this case the pattern is the same as in the Black Scholes model].

The partial derivative of the delta with respect to T will be zero if $S_0 = K$, meaning that the delta of a call in Bachelier will remain constant in T, and therefore the LCSS delta begin the difference of call

Mock Exam 7 VIII

deltas with T_1 and T_2 respectively, with $T_1 > T_2$, will be zero as the delta is constant in T.

The partial derivative of the delta with respect to T will be negative if $S_0 > K$. So if $S_0 > K$, the delta of a call in Bachelier decreases with T, and therefore the LCSS delta begin the difference of call deltas with T_1 and T_2 respectively, with $T_1 > T_2$, will be negative.

A quicker approach than calculating the derivative would have been to notice that, given the call delta in Bachelier,

$$\Phi\left(\frac{s_0 - K}{\sigma\sqrt{T}}\right)$$

knowing that Φ is an increasing function, we only need to analyze whether its argument increases or decreases. Now, $1/\sqrt{T}$ is a

Mock Exam 7 IX

decreasing function of T, unless there is a negative sign in front of it, then it becomes increasing. So all depends on the sign of $s_0 - K$, and the analysis follows more rapidly than with the formal calculation of the derivative.
Both approaches are valid.

Mock Exam 7 I

Problem 4. Long - short forward contracts with different maturities in Black Scholes.

Consider a portfolio long a forward contract with maturity $T_1 > T_2$ and short a forward contract with maturity $T_2 > 0$, both forward contracts on the same stock, following the Black-Scholes model

$$dS_t = \mu S_t dt + \sigma S_t dW_t^P,$$

$$dS_t = r S_t dt + \sigma S_t dW_t^Q,$$

where we assume $r \geq 0$ and W^P, W^Q to be standard Brownian motions under the physical and risk neutral measures P and Q respectively.
The portfolio discounted cash flows at time 0 up to maturity T_1 are written

Mock Exam 7 II

$$\Pi(0, T_1) = D(0, T_1)(S_{T_1} - K) - D(0, T_2)(S_{T_2} - K),$$

where $D(0, T) = e^{-rT}$ for all T.

a) Calculate the value at risk with confidence level α over a risk horizon $H < T_2$ for the portfolio Π in the particular case where $r = 0$.

b) Calculate the value at risk with confidence level α over a risk horizon $H < T_2$ for the portfolio Π in case $r > 0$.

c) The market has occasionally exhibited negative interest rates $r < 0$ in the past. Calculate the value at risk with confidence level α over a risk horizon $H < T_2$ for the portfolio Π in case $r < 0$

Mock Exam 7 III

d) Now we change the discounted payoff: we are long two forward contracts with maturity T_1 and short a forward contract with maturity T_2, and assume a general $r \neq 0$. Calculate VaR at a confidence level α on a horizon H.

$$\Pi(0, T_1) = 2D(0, T_1)(S_{T_1} - K) - D(0, T_2)(S_{T_2} - K),$$

e) Have you used the Black Scholes model in the different points a)-d)? Was it possible to calculate VaR, and if not why? Discuss your findings.

Mock Exam 7 IV

SOLUTIONS

a) It is convenient to keep r general, and then to specialize the analysis to $r = 0$ only in the end, so we will be able to use the calculations also for the following points. We need to caclulate the loss distribution, which is given by

Loss$_H$=Portfolio$_0$-Porfolio$_H$ $= E_0^Q \Pi(0, T_1) - E_H^Q \Pi(H, T_1) = \dots$

Let us calculate the two parts. First we calculate

$$E_0^Q \Pi(0, T_1) = E_0^Q[D(0, T_1)(S_{T_1} - K)] - E_0^Q[D(0, T_2)(S_{T_2} - K)]$$

namely the difference of the prices at time 0 of the two forward contracts.

Mock Exam 7 V

Remembering that the price of a forward contract at time 0 with maturity T is model independent and is $V^{FWD}(0, T) = S_0 - Ke^{-rT}$, we have

$$E_0^Q \Pi(0, T_1) = E_0^Q[D(0, T_1)(S_{T_1} - K)] - E_0^Q[D(0, T_2)(S_{T_2} - K)] =$$

$$= V^{FWD}(0, T_1) - V^{FWD}(0, T_2) = S_0 - Ke^{-rT_1} - S_0 + Ke^{-rT_2} = K(e^{-rT_2} - e^{-rT_1})$$

As concerns $E_H^Q \Pi(H, T_1)$, we have a similar reasoning. In the course we only derived the value of the forward at time 0. You are allowed to write directly the price of the forward with maturity T at a future time $H < T$ if you remember it, otherwise derive it as follows. The value of a forward contract at time H for maturity T is model-independent and is

Mock Exam 7 VI

given by (remember the bank account $B(t) = e^{rt}$ as the numeraire of the risk neutral measure)

$$V^{FWD}(H,T) = E_H^Q(D(H,T)(S_T - K)) = E_H^Q(e^{-r(T-H)}S_T) - E_H^Q(e^{-r(T-H)}K)$$

$$= E_H^Q\left[\frac{B_H}{B_T}S_T\right] - e^{-r(T-H)}K = B_H E_H^Q\left[\frac{S_T}{B_T}\right] - e^{-r(T-H)}K =$$

now recalling that S/B is a martingale under Q, we have that $E_H^Q[S_T/B_T] = S_H/B_H$ and substituting

$$= B_H\, S_H/B_H - e^{-r(T-H)}K = S_H - e^{-r(T-H)}K$$

Hence

$$V^{FWD}(H,T) = S_H - e^{-r(T-H)}K.$$

Mock Exam 7 VII

So the value of the portfolio in H is

$$E_H^Q \Pi(H, T_1) = E_H^Q[D(H, T_1)(S_{T_1} - K)] - E_H^Q[D(H, T_2)(S_{T_2} - K)] =$$

$$= V^{FWD}(H, T_1) - V^{FWD}(H, T_2) = S_H - e^{-r(T_1 - H)} K - (S_H - e^{-r(T_2 - H)} K)$$

$$= K(e^{-r(T_2 - H)} - e^{-r(T_1 - H)})$$

We can now write the loss ass Loss_H=Portfolio_0-Porfolio_H=

$$= E_0^Q \Pi(0, T_1) - E_H^Q \Pi(H, T_1) = K(e^{-rT_2} - e^{-rT_1}) - K(e^{-r(T_2 - H)} - e^{-r(T_1 - H)}) =$$

$$= K[e^{-r(T_1 - H)} - e^{-rT_1} - (e^{-r(T_2 - H)} - e^{-rT_2})]$$

which is a deterministic constant. In the special case $r = 0$ we have $\text{Loss}_H|_{r=0} = 0$, so the loss is deterministically zero and there is no sensible definition of VaR or ES, as the loss is 0 in every scenario. (For

Mock Exam 7 VIII

those who know distributions [not in the statistical sense but a la Laurent Schwartz], the loss distribution density is a Dirac delta centered at 0)

b) If $r > 0$,

$$Loss_H = K[e^{-r(T_1-H)} - e^{-rT_1} - (e^{-r(T_2-H)} - e^{-rT_2})]$$

$$= K[e^{-rT_1}(e^{rH} - 1) - e^{-rT_2}(e^{rH} - 1)] =$$

$$= K(e^{rH} - 1)(e^{-rT_1} - e^{-rT_2}) < 0$$

as the first term in brackets is positive, while the second is negative. Thus the loss distribution degenerates to a deterministic negative value. Thus VaR and ES are not well defined in this case. We know

Mock Exam 7 IX

with certainty that the loss over the horizon H will be negative, so we will make a positive gain in all scenarios with certainty.

c) In case $r < 0$, we can still use the formula

$$Loss_H = K(e^{rH} - 1)(e^{-rT_1} - e^{-rT_2}) < 0$$

for $r < 0$, as the first factor in brackets is negative while the second is positive. In this case, with negative rates, we make a certain positive gain with the given deterministic value above.

d) Now the discounted payoff is

$$\Pi(0, T_1) = 2D(0, T_1)(S_{T_1} - K) - D(0, T_2)(S_{T_2} - K),$$

Mock Exam 7 X

so the value of the loss changes. We have

$$\text{Portfolio}_0 = E_0^Q[2D(0,T_1)(S_{T_1} - K) - D(0,T_2)(S_{T_2} - K)] =$$

$$= 2E_0^Q[D(0,T_1)(S_{T_1} - K)] - E_0^Q[D(0,T_2)(S_{T_2} - K)]$$

$$= 2V^{FWD}(0,T_1) - V^{FWD}(0,T_2) = 2(S_0 - e^{-rT_1}K) - (S_0 - e^{-rT_2}K)$$

$$= S_0 - 2e^{-rT_1}K + e^{-rT_2}K$$

$$\text{Portfolio}_H = E_H^Q[2D(H,T_1)(S_{T_1} - K) - D(H,T_2)(S_{T_2} - K)] =$$

$$= 2E_H^Q[D(H,T_1)(S_{T_1} - K)] - E_H^Q[D(H,T_2)(S_{T_2} - K)]$$

$$= 2V^{FWD}(H,T_1) - V^{FWD}(H,T_2) = 2(S_H - e^{-r(T_1-H)}K) - (S_H - e^{-r(T_2-H)}K)$$

$$= S_H - 2e^{-r(T_1-H)}K + e^{-r(T_2-H)}K$$

$$Loss_H = \text{Portfolio}_0 - \text{Portfolio}_H =$$

Mock Exam 7 XI

$$= S_0 - 2e^{-rT_1}K + e^{-rT_2}K - S_H + 2e^{-r(T_1-H)}K - e^{-r(T_2-H)}K$$

$$= S_0 - S_H + 2Ke^{-rT_1}(e^{rH} - 1) - Ke^{-rT_2}(e^{rH} - 1)$$

$$= S_0 - S_H + K(e^{rH} - 1)(2e^{-rT_1} - e^{-rT_2}).$$

Now that we have the Loss, $q = \text{VaR}_{H,\alpha}$ is defined as

$$P[Loss_H < q] = \alpha$$

which, in our case, reads

$$P[S_0 - S_H + K(e^{rH} - 1)(2e^{-rT_1} - e^{-rT_2}) < q] = \alpha$$

Write $C = K(e^{rH} - 1)(2e^{-rT_1} - e^{-rT_2})$. Writing under the measure P

$$S_H = S_0 \exp((\mu - \frac{1}{2}\sigma^2)H + \sigma W_H^P) = S_0 \exp((\mu - \frac{1}{2}\sigma^2)H + \sigma\sqrt{H}\mathcal{N}^P)$$

Mock Exam 7 XII

where $\mathcal{N}^P$ is a standard normal under P, we can rewrite the VaR definition as

$$P\left[S_0\left(1-\exp((\mu-\frac{1}{2}\sigma^2)H+\sigma\sqrt{H}\mathcal{N}^P)\right)+C<q\right]=\alpha$$

$$P\left[1-\exp((\mu-\frac{1}{2}\sigma^2)H+\sigma\sqrt{H}\mathcal{N}^P)<\frac{q-C}{S_0}\right]=\alpha$$

$$P\left[\exp((\mu-\frac{1}{2}\sigma^2)H+\sigma\sqrt{H}\mathcal{N}^P)>1-\frac{q-C}{S_0}\right]=\alpha$$

$$P\left[\mathcal{N}^P>\frac{\ln(1-\frac{q-C}{S_0})-(\mu-\frac{1}{2}\sigma^2)H}{\sigma\sqrt{H}}\right]=\alpha$$

$$1-\Phi\left(\frac{\ln(1-\frac{q-C}{S_0})-(\mu-\frac{1}{2}\sigma^2)H}{\sigma\sqrt{H}}\right)=\alpha$$

Mock Exam 7 XIII

$$\Phi\left(-\frac{\ln(1-\frac{q-C}{S_0})-(\mu-\frac{1}{2}\sigma^2)H}{\sigma\sqrt{H}}\right)=\alpha$$

$$-\frac{\ln(1-\frac{q-C}{S_0})-(\mu-\frac{1}{2}\sigma^2)H}{\sigma\sqrt{H}}=\Phi^{-1}(\alpha)$$

and solving in q

$$q=S_0\left[1-\exp\left(-\Phi^{-1}(\alpha)\sigma\sqrt{H}+(\mu-\frac{1}{2}\sigma^2)H\right)\right]+C$$

e) We never used the Black Scholes model under the measure Q to price the portfolios in any of the points a)-d), because the forward contract price at any time is model independent, so there was no need for a model under Q for pricing. As concerns Black-Scholes under the measure P, we discuss it as follows.

Mock Exam 7 XIV

For poins a)-c), having the same stock amount, the stock part of the long-short forward contracts canceled, as the stock part had the same value both at time 0 (S_0) and at time H (S_H), leaving a loss that did not depend on S_H. So there was no need to consider S_H, and the stock appeared nowhere in the loss, so we didn't need Black Scholes under P for the loss statistics calculation either, as the loss did not depend on S. This left only the strike parts of the contract, that differed for different maturities, unless the rate $r = 0$, in which case they were both equal to K and canceled, giving a zero loss. If $r \neq 0$, then the discounting generated a deterministic difference between K cash flows valued at 0 and K cash flows valued at H, generating a unique deterministic scenario for the loss. Given this situation, when the loss is a deterministic number and is not a random variable, it doesn't make sense to talk about confidence levels and percentiles, so it is not possible to sensibly define VaR.

Mock Exam 7 XV

For point d), having now 2 long forwards and 1 short forward, this left a stock of S_0 at time 0 and S_H at time H in the loss, as a difference of the long and short contracts, so the loss depended on S_H. This required us to use the Black Scholes model under P to calculate VaR, as we needed the S_H distribution under P. This allowed us to define sensibly VaR, as now the loss was random and not a deterministic constant.

Mastery Questions

Year 4 and MSc students will get a fifth question in the exam, a mastery question. Here I give some examples. The mastery question topic is specifiied in advance usually and the students know what to expect. For all the mastery questions below, the answers are in the theory above, in the lecture notes.

Other examples could be on the Feynman Kac theorem, on the displaced diffusion model, on the Girsanov theorem, etc.

Mastery Question 1: Black Scholes and No Arbitrage I

a) Write down the equations for the two assets in the Black and Scholes economy, explain the assets nature, and list the assumptions behind the Black and Scholes economy.
[Solution: see theory. Write the equation of B, explain it's the risk free asset, solve the equation, explain what r is, write the equation of S under the measure P, explain it's the risky asset, comment on the type of SDE, etc. Write the assumptions, namely the Black and Scholes ideal conditons]

b) Define a trading strategy in the Black Scholes economy. Define a self-financing trading strategy and explain what it means intuitively. Define an arbitrage and explain the idea behind it.
[Solution: see lecture notes. The intuitive idea of a self financing strategy is that it funds itself with price movements of B and S and you don't need to inject funds to keep it going. An arbitrage is a self

Mastery Question 1: Black Scholes and No Arbitrage II

financing strategy with zero initial value that has positive probability of having strictly positive final value. Since the strategy has always non-negative value, this means having positive money with positive probability at final time, with zero initial investment. So an arbitrage is a market where there can be “money from nothing” with positive probability. Such market is rigged and we don’t want to engage in it, so we will request that there is no arbitrage]

Mastery Question 2: SDEs Ito and Stratonovich I

a) Explain how an SDE is defined in terms of integral equations. Define Ito and Stratonovich SDEs, giving pros and cons of each. [See lecture notes, as W's paths are not differentiable we cannot interpret them as differentials. We then rewrite the SDE as an integral equation. Explain the possible definitions of stochastic integral and how they differ. Explain the difference between Ito and Stratonovich SDEs based on the two different integrals. Ito integral has zero mean and the Ito isometry property. It does not look into the future (give details), and this is good, but it violates the chain rule, replaced by Ito's formula (details). Stratonovich integral looks into the future and does not have zero mean, but Stratonovich SDEs satisfy chain rule]

b) What conditions are sufficient on the coefficients of an SDE for existence and uniqueness of a global solution?
[See lecture notes, global Lipschitz continuity and linear growth]

Mastery Question 3: Smile modeling I

a) Explain the differences between the Black and Scholes model and the Bachelier model, with pros and cons of each
[Black scholes has a stock price that is lognormal (details), whereas Bachelier normal (details). As stock prices are positive, the Black Scholes model is more realistic. However, working with normals is easier than working with lognormals, so in terms of simplicity the Bachelier model is superior. Also, when modeling financial quantities like interest rates, that may go negative, the Bachelier model offers this option whereas the Black Scholes model does not]

b) Explain the basic ideas behind the mixture dynamics.
[See lecture notes. The mixture dynamics idea is to have a SDE for the stock that has a mixture of lognormal densities as density of the stock, at all times. As pricing an option is taking an average of the payoff with respect to this mixture, the price will be a linear combination of Black

Mastery Question 3: Smile modeling II

Scholes prices. This allows to fit market volatility smiles with good precision. Also, under the measure P the mixture dynamics has fat tails and matches the returns of stocks in the market, improving on the normal returns of the basic Black Scholes model. QQ-plots and skewness/kurtosis calculations confirm this in numerical examples.]

Mastery Question 4: Barings Collapse and Risk measures I

a) Write a short summary of the story of the Barings collapse, who caused it and how, why it has happened, what could have prevented it and what the financial industry and regulators did to limit the risk of such a situation repeating itself. [See lecture notes and do some reading on your own online or in the library. Answer is along the following points. Summarize the story of Leeson's trading and his final short straddle position. Talk about the short straddle payoff, what it does, plot it, explain that it was that derivative that acted like the straw that broke the camel's back, explain how the Kobe earthquake caused huge losses on that derivative. The financial industry reacted introducing VaR and later ES. Monitoring a risk measure like Value at Risk or Expected Shortfall on Leeson portfolio could have prevented

Mastery Question 4: Barings Collapse and Risk measures II

him trading like he did and taking too risky positions like the final straddle.]

b) Define value at risk (VaR) and give at least two drawbacks of this risk measure. Define expected shortfall and give at least one drawback. Explain how it would have helped in the Barings case. [See lecture notes, give detailed definitions. Drawback of VaR is that it doesn't see the tail beyond the confidence level. Also, it is not sub-additive. ES overcomes this, but has still the homogeneity assumpion, which is unrealistic for liquidity risk. Give more details. How it would have helped in Barings is explained in the earlier answer.]

Mastery Question 2025-2026 I

This year's mastery question will be on the mixture dynamics model. You won't have to remember all the equations by heart but more answer conceptual questions behind the model, such has for example the one below. Answers of the length of the ones below will be accepted. For preparation, the slides and the questions below are sufficients. If you wish to know more, you can check the papers given in the mixture dynamics section but this is not necessary.

- What is the idea behind the local volatility used in the mixture dynamics SDEs? [It is a local volatility designed in such a way that the density of the stock at very time is a mixture of lognormal densities with different volatilities, equivalent to say that it is a mixture of Black Scholes models densities with different volatilities]

Mastery Question 2025-2026 II

- Does the model fit the market well? [Examples from several asset classes have shown that the model can fit the market smile well in equity, interest-rates and FX. In some cases a shift may need to be introduced, like in the displaced diffusion model, leading to the mixture diffusion dynamics]
- Does the model fit historical data under the measure P if we put a drift $\mu\, S_t dt$ in the SDE to change measure? [QQ plot analysis done in class shows that indeed the mixture dynamics SDE is a big improvement over Black Schooles in fitting sone SP500 returns data, as the QQ plot shape of the model resembles the QQ plot of the data qualitatively]

Mastery Question 2025-2026 III

- Does the model preserve the smile shape in the future, of will future smiles tend to flatten? [As this is a local volatility model, and all these models suffer from the smile flattening problem, this model will have future smiles flattening too]
- Is the mixture dynamics able to reproduce a U-shape smile or only monotonic smiles like displaced diffusion, CEV and Bachelier? [The Mixture dynamics can reproduce U-shaped smiles, indeed it couldn't fit the market as successfully as it does otherwise. We have seen examples in class]

Mastery Question 2025-2026 IV

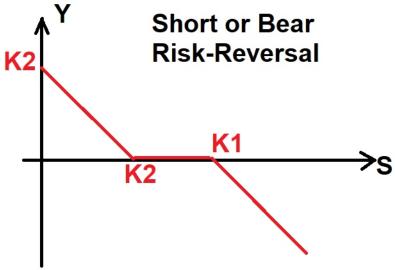
- When was the mixture dynamics introduced in its basic form, before all the subsequent extensions, and by whom? Who helped extend it to multivariate SDEs and what is the benefit? [The mixture dynamics was introduced in the early 2000's by Brigo, Mercurio and other colleagues like Rapisarda and Sartorelli, among others. Sridi and Pisani helped with multivariate extensions of the model in 2018 and 2021. This, given a stock index, allows for a consistent modeling of the single stocks smiles and of the index smile, or to deduce unknown FX rates smiles via FX triangulation starting from two FX rates smiles, to get the third FX rate smile.]

Coursework 2025-2026 I

Consider a bear (or short) risk reversal option payoff on a stock price S, with maturity T, defined as

$$Y = -(S_T - K_1)^+ + (K_2 - S_T)^+$$

with $K_2 < S_0 < K_1$. The payoff is illustrated in the picture below.



Coursework 2025-2026 II

Assume the stock follows a Black Scholes price model under the measure P given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t^P, \quad S_0 = s_0.$$

We have the following values:

$$s_0 = 100; K_2 = 95; K_1 = 105; \mu = 0.1; \sigma = 0.4; r = 0.05; T = 10y.$$

1. Explain what kind of investor would be interested in holding this payoff and what they would expect from the market when purchasing it. **[10pt]**
2. Calculate the price of the bear risk reversal in the Black Scholes model at time 0. **[10pt]**

Coursework 2025-2026 III

3. We wish to calculate the VaR and ES of the bear risk reversal position over a risk horizon $H = 1y$ at 95% confidence level. Write a Python or Matlab/Octave code that does this and present the values of the VaR and expected shortfall you found. Run at least 100,000 scenarios. Use a vectorized code if possible. Append the code as an appendix to the PDF **[20pt]**
4. Produce a histogram of the density of the loss distribution at 1 year and show the VaR and ES points in the loss graph. **[10pt]**
5. Increase the volatility to the following values:

$$a)\ \sigma = 0.6; \quad b)\sigma = 0.8; \quad c)\sigma = 1.$$

   For the three cases, show the VaR and ES figures. **[20pt]**

Coursework 2025-2026 IV

6. More generally, can you deduce a pattern about how the risk of loss with this contract evolves with the volatility over one year? Are VaR and/or ES increasing, decreasing or neither increasing nor decreasing in σ? You may run more cases, possibly with a for loop, to anwer this question. Find also the pattern of the tail depth defined as $TD = ES - VaR$ as the volatility increases. **[30pt]**

Coursework starts on Nov 19 and submission is expected by Dec 3, 2025.

Coursework will contribute 10% of the final mark.

Prepare a PDF file with all the answers, and attach your code as an appendix. Then submit through the Turnitin Dropbox via Blackboard

Coursework 2025-2026 V

Group formation. Students are strongly encouraged to self organize in groups of 4. If 4 is not possible, groups of 3 will be accepted but as a last resort. No groups of 5 or more. **You must have emailed your group information by Nov 13, either as a formed group with a group name, or as an individual who can’t find a group**

Every group elects a group leader. The procedure is to email the Teaching Assistant. Please follow these instructions. *The email should come from the group leader alone* (the other group members should not send emails) and should have a subject reading

Coursework group [chosen name of the group]

The body of the email should read “Hello, this is the email for the group [name of the group]. This group is formed by”

Coursework 2025-2026 VI

	A	B	C	D
1	GROUP:	VIVA ARBITRAGE		
2	Mary	Smith	CID1	Leader
3	John	Woods	CID2	
4	Robert	Jackson	CID3	
5	Louise	James	CID4	
6				

In doing this please use an excel format and copy/paste the excel cells for the table above in cell format on the email body. This will allow the TA to collate the groups more effectively by copying and pasting your group into a general spreadsheet from the email directly.

If you cannot find a group, email the Teaching Assistant with an email whose subject is “Coursework: Can’t find a group” including in the email body you name, family name, CID. The TA will try to insert you in

Coursework 2025-2026 VII

a group of 3 or to form a new group of 4 with the other students who can't find a group.
Also for this, follow the structure of an excel file and copy/paste the single row with your data in the message, in cells format, so that the TA can easily copy/paste this.

Be respectful of other students, polite, and well-behaving in the group. Hear all opinions respectfully. There were previous cases of students who behaved abusively with respect to other students and who decided to discard their opinions on the solutions without reason or discussion. This will not be tolerated and any reports of such situations will be dealt with.

Simulation of Brownian motion: Python code I

```
"""
Brownian motion W: simulation of n paths
@author: Prof. Damiano Brigo
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
from scipy.stats import skew;
from scipy.stats import kurtosis;
from math import exp;
import matplotlib.pyplot as plt;
# Initial value
W0 =0;
# Final time in years
h=1;
# Number of time steps and time step
nt = 3650; dt = h/nt;
# Number of scenarios
n=10000;
""" Generating the normal 0,1 in n scenarios for each time step """
# dW are normals representing W increments, W will be the Brownian
dWt = np.zeros(n);
Wt = W0*np.ones((nt,n));
from random import seed
from random import gauss
#seed random number generator
seed(1)
```

Simulation of Brownian motion: Python code II

```
# loop on scenarios
for j in range(n):
 Wtpredj=0;
# loop on time
 for i in range(nt):
  dWt[j] = gauss(0,1)*(dt**0.5);
  Wt[i,j] =Wtpredj+dWt[j];
  Wtpredj = Wt[i,j];
# endfor
#endor
# mean and standard deviation of the solution
meansim = np.average(Wt[nt-1,0:(n-1)]);
print("mean simulation Euler:",meansim);
#
Stdsim = np.std(Wt[nt-1,0:(n-1)]);
print("Std simulation Euler:",Stdsim);
# Comparing skewness and kurtosis
skewSteuler = skew(Wt[nt-1,0:(n-1)], axis = 0, bias = True);
print("Skew Euler:",skewSteuler);
#
kurtSteuler = kurtosis(Wt[nt-1,0:(n-1)], axis = 0, bias = True);
print("Kurtosis Euler:",kurtSteuler);
tt = list(range(0,nt,1));
plt.plot(tt,Wt)
```

Value at Risk and ES: Short Straddle I

```
"""
SHORT STRADDLE RISK MEASURES: VaR and ES
Created on Mon Feb  7 15:50:34 2022
@author: Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
import matplotlib.pyplot as plt;
# Stock price parameters
S0 =100;
miu=0.05;
Sig=0.5;
r=0.01;
#option strike, number of simulations, maturity
k=100;
n=100000;
T=1;
# Confidence level and risk horizon for VaR and ES
confidence=0.99;
h=0.25;
# Option prices at time 0
d1c=(math.log(S0/k)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
```

Value at Risk and ES: Short Straddle II

```
d1p=(math.log(S0/k)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
c0=S0*norm.cdf(d1c)-k*math.exp(-r*T)*norm.cdf(d1c-Sig*T**0.5);
p0=-S0*norm.cdf(-d1p)+k*math.exp(-r*T)*norm.cdf(-d1p+Sig*T**0.5);
v0=-c0-p0;
""" computing the prices after h years"""
T=T-h;
Zt = np.zeros(n);
St = np.zeros(n);
from random import seed
from random import gauss
#seed random number generator
seed(1)
# Simulating the stock up to time h in n scenarios
for j in range(n):
 Zt[j] = gauss(0,1)
 St[j]=S0*math.exp((miu-0.5*Sig**2)*h)*math.exp(Zt[j]*Sig*(h**0.5))
# Call and put prices at time h in each scenario
ct=np.zeros(n);
pt=np.zeros(n);
for i in range(n):
    d1cnew=(math.log(St[i]/k)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
    d1pnew=(math.log(St[i]/k)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
    ct[i]=St[i]*norm.cdf(d1cnew)-k*math.exp(-r*T)*norm.cdf(d1cnew-Sig*T**(0.5));
    pt[i]=-St[i]*norm.cdf(-d1pnew)+k*math.exp(-r*T)*norm.cdf(-d1pnew+Sig*T**(0.5));
# end for
# value of straddle in each scenario
vt=-ct-pt;
# Loss calculation vvar = loss
```

Value at Risk and ES: Short Straddle III

```
vvar=v0-vt;
vvar=np.sort(vvar);
# Extracting VaR at the right confidence level from the loss
ivar = round((confidence)*n);
var = vvar[ivar];
# Calculating ES
ESv=statistics.mean(vvar[range(math.floor((confidence)*n),n)]);
print("VaR:",var);
print("ES:",ESv);
#histogram of the loss
plt.hist(vvar, bins = 100)
plt.show()
```

VaR and ES: Call & Put on correlated stocks I

```
"""
CALL AND PUT OPTION PORTFOLIO ON CORRELATED STOCKS: VaR and ES
Created on Mon Feb  7 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
import matplotlib.pyplot as plt;
# Stocks data
S01 =120;
S02 =80;
miu1=0.05;
miu2=0.02;
Sig1=0.5;
Sig2=0.2;
rho=-0.8;
r=0.01;
# Options maturity and strikes
T=2;
k1=116;
k2=86;
# number of scenarios, confidence level and risk horizon
```

VaR and ES: Call & Put on correlated stocks II

```
n=40000;
confidence=0.95;
h = 0.25;
# Option prices at time zero
d1c=(math.log(S01/k1)+(r+0.5*Sig1**2)*T)/(Sig1*T**0.5);
d1p=(math.log(S02/k2)+(r+0.5*Sig2**2)*T)/(Sig2*T**0.5);
c0=S01*norm.cdf(d1c)-k1*math.exp(-0.01*T)*norm.cdf(d1c-Sig1*T**0.5);
p0=-S02*norm.cdf(-d1p)+k2*math.exp(-0.01*T)*norm.cdf(-d1p+Sig2*T**0.5);
v0=c0+p0;
print("Call S1 strike K1:",c0);
print("Put S2 strike K2:",p0);
""" computing the option prices after h years"""
T=T-h;
Zt1 = np.zeros(n);
Zt2 = np.zeros(n);
St1 = np.zeros(n);
St2 = np.zeros(n);
from random import seed
from random import gauss
#seed random number generator
seed(1)
# generate Gaussian random values and correlated Brownian motion scenarios
for j in range(n):
 Zt1[j] = gauss(0,1)
 Zt2[j] = rho*Zt1[j] + ((1-rho**2)**(0.5))*gauss(0,1)
# Generating stock scenarios
 St1[j]=S01*math.exp((miu1-0.5*Sig1**2)*h)*math.exp(Zt1[j]*Sig1*(h**0.5))
 St2[j]=S02*math.exp((miu2-0.5*Sig2**2)*h)*math.exp(Zt2[j]*Sig2*(h**0.5))
```

VaR and ES: Call & Put on correlated stocks III

```
# end for
c1t=np.zeros(n);
p2t=np.zeros(n);
# Calculating option prices scenarios in h years
for i in range(n):
    d1c1new=(math.log(St1[i]/k1)+(r+0.5*Sig1**2)*T)/(Sig1*T**0.5);
    d1p2new=(math.log(St2[i]/k2)+(r+0.5*Sig2**2)*T)/(Sig2*T**0.5);
    c1t[i]=St1[i]*norm.cdf(d1c1new)-k1*math.exp(-0.01*T)*norm.cdf(d1c1new-Sig1*T**(0.5));
    p2t[i]=-St2[i]*norm.cdf(-d1p2new)+k2*math.exp(-0.01*T)*norm.cdf(-d1p2new+Sig2*T**(0.5));
# end for
# calculating portfolio value at time h in all scenarios
vt=c1t+p2t;
# calculating loss vvar in all scenarios
vvar=v0-vt;
vvar=np.sort(vvar);
# extracting VaR from loss at the right confidence level
ivar = round((confidence)*n);
var = vvar[ivar];
# Calculating ES
ESv=statistics.mean(vvar[range(math.floor((confidence)*n),n)]);
print("VaR:",var);
print("ES:",ESv);
# Plotting histogram of the loss
plt.hist(vvar, bins = 100)
plt.show()
```

VaR and ES: Bull Call Spread I

```
"""
VaR and ES for BULL CALL SPREAD
Created on Mon Feb  7 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
import matplotlib.pyplot as plt;
# Stock data
S0 =100;
Sig=0.4;
miu=0.05;
r=0.01;
# option strikes and maturity
k1=110;
k2=90;
T=5;
# Number of scenarios, confidence level, risk horizon
n=40000;
confidence=0.95;
h=1;
# Value of the options at time 0
```

VaR and ES: Bull Call Spread II

```
d1c1=(math.log(S0/k1)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
d1c2=(math.log(S0/k2)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
c20=S0*norm.cdf(d1c2)-k2*math.exp(-0.01*T)*norm.cdf(d1c2-Sig*T**0.5);
c10=S0*norm.cdf(d1c1)-k1*math.exp(-0.01*T)*norm.cdf(d1c1-Sig*T**0.5);
v0=c20-c10;
print("Call strike K2:",c20);
print("Call strike K1:",c10);
""" computing the prices after one year"""
T=T-h;
Zt = np.zeros(n);
St = np.zeros(n);
from random import seed
from random import gauss
#seed random number generator
seed(1)
# generate Gaussian random values and Brownian motion scenarios
for j in range(n):
 Zt[j] = gauss(0,1)
 St[j]=S0*math.exp((miu-0.5*Sig**2)*h)*math.exp(Zt[j]*Sig*(h**0.5))
# endfor
c1t=np.zeros(n);
c2t=np.zeros(n);
# Calculating call and put at time h in all scenarios
for i in range(n):
    d1c1new=(math.log(St[i]/k1)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
    d1c2new=(math.log(St[i]/k2)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
    c1t[i]=St[i]*norm.cdf(d1c1new)-k1*math.exp(-r*T)*norm.cdf(d1c1new-Sig*T**(0.5));
    c2t[i]=St[i]*norm.cdf(d1c2new)-k2*math.exp(-r*T)*norm.cdf(d1c2new-Sig*T**(0.5));
```

VaR and ES: Bull Call Spread III

```
# endfor
# Value of portfolio in all scenarios
vt=c2t-c1t;
# Loss scenarios vvar
vvar=v0-vt;
vvar=np.sort(vvar);
# Extracting VaR from loss at the right confidence level
ivar = round((confidence)*n);
var = vvar[ivar];
# Calculating ES
ESv=statistics.mean(vvar[range(math.floor((confidence)*n),n)]);
print("VaR:",var);
print("ES:",ESv);
# Histogram of loss
plt.hist(vvar, bins = 100)
plt.show()
```

VaR and ES: Risk Reversal I

```
"""
VaR and ES for RISK REVERSAL
Created on Mon Feb  7 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
import matplotlib.pyplot as plt;
# Stock data
S0 =100;
Sig=0.2;
r=0.01;
miu=0.05;
# Options strikes and maturities
k1=110;
k2=90;
T=5;
# Number of scenarios, confidence level, risk horizon
n=10000;
confidence=0.95;
h=1;
# Option prices at time 0
```

VaR and ES: Risk Reversal II

```
d1c=(math.log(S0/k1)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
d1p=(math.log(S0/k2)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
c0=S0*norm.cdf(d1c)-k1*math.exp(-r*T)*norm.cdf(d1c-Sig*T**0.5);
p0=-S0*norm.cdf(-d1p)+k2*math.exp(-r*T)*norm.cdf(-d1p+Sig*T**0.5);
v0=c0-p0;
""" computing the prices at time h"""
T=T-h;
Zt = np.zeros(n);
St = np.zeros(n);
from random import seed
from random import gauss
#seed random number generator
seed(1)
# generate some random Gaussian values and the stock scenarios at time h
for j in range(n):
 Zt[j] = gauss(0,1)
 St[j]=S0*math.exp((miu-0.5*Sig**2)*h)*math.exp(Zt[j]*Sig*(h**0.5))
# endfor
ct=np.zeros(n);
pt=np.zeros(n);
# Generating call and put scenarios at time h
for i in range(n):
    d1cnew=(math.log(St[i]/k1)+(r+0.5*Sig**2)*T)/(Sig*(T**0.5));
    d1pnew=(math.log(St[i]/k2)+(r+0.5*Sig**2)*T)/(Sig*(T**0.5));
    ct[i]=St[i]*norm.cdf(d1cnew)-k1*math.exp(-r*T)*norm.cdf(d1cnew-Sig*T**(0.5));
    pt[i]=-St[i]*norm.cdf(-d1pnew)+k2*math.exp(-r*T)*norm.cdf(-d1pnew+Sig*T**(0.5));
# endfor
# final portolio value in all scenarios
```

VaR and ES: Risk Reversal III

```
vt=ct-pt;
# loss scenarios
vvar=v0-vt;
vvar=np.sort(vvar);
# extracting VaR from the loss at the right confidence level
ivar = round((confidence)*n);
var = vvar[ivar];
# calculating ES
ESv=statistics.mean(vvar[range(math.floor((confidence)*n),n)]);
print("VaR:",var);
print("ES:",ESv);
# plotting loss histogram
plt.hist(vvar, bins = 100)
plt.show()
```

Euler scheme for Geometric Brownian Motion I

```
# -*- coding: utf-8 -*-
"""
Euler scheme for Geometric Brownian Motion
Created on Mon Nov  28 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
from scipy.stats import skew;
from scipy.stats import kurtosis;
import matplotlib.pyplot as plt;
# Stock data
S0 =100;
Sig=0.2;
miu=0.05;
# Final time in years
h=1;
# Number of time steps and time step
nt = 200; dt = h/nt;
# Number of scenarios
n=40000;
""" Generating the normal 0,1 in 10000 scenarios for each time step """
```

Euler scheme for Geometric Brownian Motion II

```
# Z are normals, S will be the stock
Zt = np.zeros(n);
St = S0*np.ones(n);
from random import seed
from random import gauss
#seed random number generator
seed(1)
# loop on time
for i in range (nt):
# loop on scenarios
 for j in range(n):
  Zt[j] = gauss(0,1)
  St[j]=St[j]+miu*St[j]*dt+Sig*St[j]*Zt[j]*(dt**0.5)
# endfor
#endor
# One shot final simulation using the knonw GBM solution
Sth = S0*np.ones(n);
# loop on scenarios
for j in range(n):
 Zt[j] = gauss(0,1)
 Sth[j]=S0*math.exp((miu-0.5*Sig**2)*h)*math.exp(Zt[j]*Sig*(h**0.5))
# endfor
# plotting final stock histogram
#Blue is St Euler, Orange is St one shot
plt.hist(St, bins = 100, density=True)
plt.hist(Sth, bins = 100, density=True)
plt.show()
# Orange is St Euler, Blue is St one shot
```

Euler scheme for Geometric Brownian Motion III

```
plt.hist(Sth, bins = 100, density=True)
plt.hist(St, bins = 100, density=True)
plt.show()
#
# plotting histogram of the difference
Stsorted = sorted(St);
Sthsorted = sorted(Sth);
diffe = np.subtract(Stsorted,Sthsorted);
plt.hist(diffe, bins = 100, density=True)
plt.show()
# comparing mean and standard deviation of log returns
meanlogth = np.log(S0)+(miu-0.5*Sig**2)*h;
meanlogsim = np.average(np.log(Stsorted));
meanlogsimh = np.average(np.log(Sthsorted));
print("log mean theoretical:",meanlogth);
print("log mean simulation Euler:",meanlogsim);
print("log mean simulation one shot:",meanlogsimh);
#
Stdth = Sig*h;
Stdsim = np.std(np.log(Stsorted));
Stdsimh = np.std(np.log(Sthsorted));
print("Std theoretical:",Stdth);
print("Std simulation Euler:",Stdsim);
print("Std simulation one shot:",Stdsimh);
# Comparing skewness and kurtosis
skewSteuler = skew(np.log(Stsorted), axis = 0, bias = True);
skewSth = skew(np.log(Sthsorted), axis = 0, bias = True);
print("Skew Euler:",skewSteuler);
```

Euler scheme for Geometric Brownian Motion IV

```
print("Skew one shot:",skewSth);
#
kurtSteuler = kurtosis(np.log(Stsorted), axis = 0, bias = True);
kurtSth = kurtosis(np.log(Sthsorted), axis = 0, bias = True);
print("Kurtosis Euler:",kurtSteuler);
print("Kurtoses one shot:",kurtSth);
```

Euler scheme for for $dX = mdt + \sigma XdW$ I

```
# -*- coding: utf-8 -*-
"""
Euler scheme for dX = m dt + sigma X dW
Created on Mon Nov  28 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
from scipy.stats import skew;
from scipy.stats import kurtosis;
import matplotlib.pyplot as plt;
# Stock data
S0 =0;
Sig=0.4;
miu=1;
# Final time in years
h=2;
# Number of time steps and time step
nt = 200; dt = h/nt;
# Number of scenarios
```

Euler scheme for for $dX = mdt + \sigma XdW$ II

```
n=40000;
""" Generating the normal 0,1 in 10000 scenarios for each time step """
# Z are normals, S will be the stock
Zt = np.zeros(n);
St = S0*np.ones(n);
from random import seed
from random import gauss
#seed random number generator
seed(1)
# loop on time
for i in range (nt):
# loop on scenarios
 for j in range(n):
  Zt[j] = gauss(0,1)
  St[j]=St[j]+miu*dt+Sig*St[j]*Zt[j]*(dt**0.5)
# endfor
#endor
# plotting final stock histogram
#Blue is St Euler, Orange is St one shot
plt.hist(St, bins = 100)
plt.show()
# comparing mean and standard deviation of the solution
meanth = S0+miu*h;
meansim = np.average(St);
print("mean theoretical:",meanth);
print("mean simulation Euler:",meansim);
#
varth = 2*(miu/(Sig**4))*(miu+S0*Sig**2)*(math.exp(h*Sig**2)-1);
```

Euler scheme for for $dX = mdt + \sigma XdW$ III

```
varth = varth - 2*h*(miu/Sig)**2 + math.exp(h*Sig**2)*S0**2;
varth = varth - (S0+miu*h)**2;
Stdth = math.sqrt(varth);
Stdsim = np.std(St);
print("Std theoretical:",Stdth);
print("Std simulation Euler:",Stdsim);
# Comparing skewness and kurtosis
skewSteuler = skew(St, axis = 0, bias = True);
print("Skew Euler:",skewSteuler);
#
kurtSteuler = kurtosis(St, axis = 0, bias = True);
print("Kurtosis Euler:",kurtSteuler);
```

Euler scheme $dX = (-3kX + 3X^{1/3}\sigma^2)dt + 3\sigma X^{2/3}dW_t$ I

```
# -*- coding: utf-8 -*-
"""
Euler scheme for d X = (-3 k X  +3 X^{1/3}  \sigma^{2} )dt+ 3 \sigma X^{2/3} dW_t
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""

import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
from scipy.stats import skew;
from scipy.stats import kurtosis;
from math import exp;
import matplotlib.pyplot as plt;
# Stock data
S0 =0.5;
Sig=0.05;
kk=1;
# Final time in years
h=1;
# Number of time steps and time step
```

Euler scheme $dX = (-3kX + 3X^{1/3}\sigma^2)dt + 3\sigma X^{2/3}dW_t$ II

```
nt = 250; dt = h/nt;
# Number of scenarios
n=40000;
""" Generating the normal 0,1 in 10000 scenarios for each time step """
# Z are normals, S will be the stock
Zt = np.zeros(n);
St = S0*np.ones(n);
from random import seed
from random import gauss
#seed random number generator
seed(1)
# loop on time
for i in range (nt):
# loop on scenarios
 for j in range(n):
  Zt[j] = gauss(0,1)
  St[j]=St[j]-3*kk*St[j]*dt+3*((St[j])**(1/3))*(Sig**2)*dt;
  St[j]=St[j]+3*Sig*((St[j])**(2/3))*Zt[j]*(dt**0.5);
  St[j]
# endfor
#endor
# mean and standard deviation of the solution
meansim = np.average(St);
print("mean simulation Euler:",meansim);
#
```

Euler scheme $dX = (-3kX + 3X^{1/3}\sigma^2)dt + 3\sigma X^{2/3} dW_t$ III

```
Stdsim = np.std(St);
print("Std simulation Euler:",Stdsim);
# Comparing skewness and kurtosis
skewSteuler = skew(St, axis = 0, bias = True);
print("Skew Euler:",skewSteuler);
#
kurtSteuler = kurtosis(St, axis = 0, bias = True);
print("Kurtosis Euler:",kurtSteuler);
# Simulation one shot
Zt1s = np.zeros(n);
St1s = S0*np.ones(n);
for j in range(n):
 Zt1s[j] = gauss(0,1)
 St1s[j]=(S0**(1/3)*math.exp(-kk*1)+Zt1s[j]*((Sig**2/(2*kk))*(1-math.exp(-2*kk)))**(1/2))**3;
# endfor
# plotting final stock histogram
#Blue is St Euler, Orange is St one shot
plt.hist(St, bins = 100, density=True)
plt.hist(St1s, bins = 100, density=True)
plt.show()

plt.hist(St1s, bins = 100, density=True)
plt.hist(St, bins = 100, density=True)
plt.show()
```

More efficient Python codes: Vectorized versions I

The codes above are meant to be educational, and use full loops both on scenarios and time, when time steps are needed. This is because such a code is closer to a C++ code often used in the industry in production libraries.

However, many interpreted languages like Python can become much more efficient when operations are vectorized. In our codes there is no need to loop on scenarios. Python can do operations on vectors directly, so we don't need to loop on components. The following codes use vectorization and are much faster, to the point that one can increase the number of scenarios of the simulation dramatically while staying on a fast run time.

Vectorized Simulation of Brownian: Python code I

```
"""
Brownian motion W: simulation of n paths - vectorized efficient version
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
from scipy.stats import skew;
from scipy.stats import kurtosis;
import matplotlib.pyplot as plt;
# Initial value
W0 =0;
# Final time in years
h=1;
# Number of time steps and time step
nt = 3650; dt = h/nt;
# Number of scenarios
n=40000;
""" Generating the normal 0,1 vector for n scenarios for each time step """
# dW are normals representing W increments, W will be the Brownian
dWt = np.zeros(n);
Wt = W0*np.ones((nt,n));
from random import seed
#seed random number generator
seed(1)
Wtpredj=0;
```

Vectorized Simulation of Brownian: Python code II

```
# loop on time
for i in range(nt):
 dWt = np.random.randn(n)*(dt**0.5);
 Wt[i,:] =Wtpredj+dWt;
 Wtpredj = Wt[i,:];
# endfor
# mean and standard deviation of the solution (should be 0 and 1 from theory)
meansim = np.average(Wt[nt-1,0:(n-1)]);
print("mean simulation Euler:",meansim);
#
Stdsim = np.std(Wt[nt-1,0:(n-1)]);
print("Std simulation Euler:",Stdsim);
# Comparing skewness and kurtosis, should be zero from theory
skewSteuler = skew(Wt[nt-1,0:(n-1)], axis = 0, bias = True);
print("Skew:",skewSteuler);
kurtSteuler = kurtosis(Wt[nt-1,0:(n-1)], axis = 0, bias = True);
print("Kurtosis:",kurtSteuler);
#
nscenariosplot = 10;
tt = list(range(0,nt,1));
plt.plot(tt,Wt[:,0:nscenariosplot])
```

VaR and ES vectorized: Short Straddle I

```
"""
SHORT STRADDLE RISK MEASURES: VaR and ES Vectorized, efficient version
Created on Mon Feb  7 15:50:34 2022
@author: Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
import matplotlib.pyplot as plt;
# Stock price parameters
S0 =100;
miu=0.05;
Sig=0.5;
r=0.01;
#option strike, number of simulations, maturity
k=100;
n=1000000;
T=1;
# Confidence level and risk horizon for VaR and ES
confidence=0.99;
h=0.25;
# Option prices at time 0
d1c=(math.log(S0/k)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
```

VaR and ES vectorized: Short Straddle II

```
d1p=(math.log(S0/k)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
c0=S0*norm.cdf(d1c)-k*math.exp(-r*T)*norm.cdf(d1c-Sig*T**0.5);
p0=-S0*norm.cdf(-d1p)+k*math.exp(-r*T)*norm.cdf(-d1p+Sig*T**0.5);
v0=-c0-p0;
""" computing the prices after h years"""
T=T-h;
Zt = np.zeros(n);
St = np.zeros(n);
from random import seed
#seed random number generator
seed(1)
# Simulating the stock up to time h in n scenarios
Zt = np.random.randn(n);
St=S0*np.exp((miu-0.5*Sig**2)*h)*np.exp(Zt*Sig*(h**0.5))
# Call and put prices at time h
d1cnew=(np.log(St/k)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
d1pnew=(np.log(St/k)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
ct=St*norm.cdf(d1cnew)-k*np.exp(-r*T)*norm.cdf(d1cnew-Sig*T**(0.5));
pt=-St*norm.cdf(-d1pnew)+k*np.exp(-r*T)*norm.cdf(-d1pnew+Sig*T**(0.5));
# end for
# value of straddle in each scenario
vt=-ct-pt;
# Loss calculation vvar = loss
vvar=v0-vt;
vvar=np.sort(vvar);
# Extracting VaR at the right confidence level from the loss
ivar = round((confidence)*n);
var = vvar[ivar];
```

VaR and ES vectorized: Short Straddle III

```
# Calculating ES
ESv=statistics.mean(vvar[range(math.floor((confidence)*n),n)]);
print("VaR:",var);
print("ES:",ESv);
#histogram of the loss
plt.hist(vvar, bins = 100)
plt.show()
```

VaR & ES vectorized: Call & Put on correlated stocks I

```
"""
CALL  PUT OPTION PORTFOLIO ON CORRELATED STOCKS: Vectorized VaR ES
Created on Mon Feb  7 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
import matplotlib.pyplot as plt;
# Stocks data
S01 =120;
S02 =80;
miu1=0.05;
miu2=0.02;
Sig1=0.5;
Sig2=0.2;
rho=1;
r=0.01;
# Options maturity and strikes
T=2;
k1=116;
k2=86;
# number of scenarios, confidence level and risk horizon
```

VaR & ES vectorized: Call & Put on correlated stocks II

```
n=400000;
confidence=0.95;
h = 0.25;
# Option prices at time zero
d1c=(math.log(S01/k1)+(r+0.5*Sig1**2)*T)/(Sig1*T**0.5);
d1p=(math.log(S02/k2)+(r+0.5*Sig2**2)*T)/(Sig2*T**0.5);
c0=S01*norm.cdf(d1c)-k1*math.exp(-0.01*T)*norm.cdf(d1c-Sig1*T**0.5);
p0=-S02*norm.cdf(-d1p)+k2*math.exp(-0.01*T)*norm.cdf(-d1p+Sig2*T**0.5);
v0=c0+p0;
print("Call S1 strike K1:",c0);
print("Put S2 strike K2:",p0);
""" computing the option prices after h years"""
T=T-h;
Zt1 = np.zeros(n);
Zt2 = np.zeros(n);
St1 = np.zeros(n);
St2 = np.zeros(n);
from random import seed
#seed random number generator
seed(1)
Zt1 = np.random.randn(n);
Zt2 = rho*Zt1 + ((1-rho**2)**(0.5))*np.random.randn(n);

St1 = S01 * np.exp((miu1 - 0.5 * Sig1 ** 2) * h) * np.exp(Zt1 * Sig1 * (h ** 0.5));
St2 = S02 * np.exp((miu2 - 0.5 * Sig2 ** 2) * h) * np.exp(Zt2 * Sig2 * (h ** 0.5));
```

VaR & ES vectorized: Call & Put on correlated stocks III

```
# Calculating option prices scenarios in h years
d1c1new=(np.log(St1/k1)+(r+0.5*Sig1**2)*T)/(Sig1*T**0.5);
d1p2new=(np.log(St2/k2)+(r+0.5*Sig2**2)*T)/(Sig2*T**0.5);
c1t=St1*norm.cdf(d1c1new)-k1*np.exp(-0.01*T)*norm.cdf(d1c1new-Sig1*T**(0.5));
p2t=-St2*norm.cdf(-d1p2new)+k2*np.exp(-0.01*T)*norm.cdf(-d1p2new+Sig2*T**(0.5));
# calculating portfolio value at time h in all scenarios
vt=c1t+p2t;
# calculating loss vvar in all scenarios
vvar=v0-vt;
vvar=np.sort(vvar);
# extracting VaR from loss at the right confidence level
ivar = round((confidence)*n);
var = vvar[ivar];
# Calculating ES
ESv=statistics.mean(vvar[range(math.floor((confidence)*n),n)]);
print("VaR:",var);
print("ES:",ESv);
# Plotting histogram of the loss
plt.hist(vvar, bins = 100)
plt.show()
```

VaR & ES Vectorized: Bull Call Spread I

```
"""
VaR and ES for BULL CALL SPREAD, Vectorized efficient version
Created on Mon Feb  7 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
import matplotlib.pyplot as plt;
# Stock data
S0 =100;
Sig=0.2;
miu=0.05;
r=0.01;
# option strikes and maturity
k1=110;
k2=90;
T=5;
# Number of scenarios, confidence level, risk horizon
n=1000000;
confidence=0.95;
h=1;
# Value of the options at time 0
```

VaR & ES Vectorized: Bull Call Spread II

```
d1c1=(np.log(S0/k1)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
d1c2=(np.log(S0/k2)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
c20=S0*norm.cdf(d1c2)-k2*np.exp(-0.01*T)*norm.cdf(d1c2-Sig*T**0.5);
c10=S0*norm.cdf(d1c1)-k1*np.exp(-0.01*T)*norm.cdf(d1c1-Sig*T**0.5);
v0=c20-c10;
print("Call strike K2:",c20);
print("Call strike K1:",c10);
""" computing the prices after one year"""
T=T-h;
Zt = np.zeros(n);
St = np.zeros(n);
from random import seed
#seed random number generator
seed(1)
# block generate Gaussian random vectors and Brownian motion scenarios
Zt = np.random.randn(n);
St=S0*np.exp((miu-0.5*Sig**2)*h)*np.exp(Zt*Sig*(h**0.5))
# Calculating call and put at time h in all scenarios
d1c1new=(np.log(St/k1)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
d1c2new=(np.log(St/k2)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
c1t=St*norm.cdf(d1c1new)-k1*np.exp(-r*T)*norm.cdf(d1c1new-Sig*T**(0.5));
c2t=St*norm.cdf(d1c2new)-k2*np.exp(-r*T)*norm.cdf(d1c2new-Sig*T**(0.5));
# endfor
# Value of portfolio in all scenarios
vt=c2t-c1t;
# Loss scenarios vvar
vvar=v0-vt;
vvar=np.sort(vvar);
```

VaR & ES Vectorized: Bull Call Spread III

```
# Extracting VaR from loss at the right confidence level
ivar = round((confidence)*n);
var = vvar[ivar];
# Calculating ES
ESv=statistics.mean(vvar[range(math.floor((confidence)*n),n)]);
print("VaR:",var);
print("ES:",ESv);
# Histogram of loss
plt.hist(vvar, bins = 100)
plt.show()
```

VaR & ES Vectorized: Long Risk Reversal I

```
"""
VaR and ES for LONG RISK REVERSAL, Vectorized efficient version
Created on Mon Feb  7 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
import statistics;
from scipy.stats import norm;
import matplotlib.pyplot as plt;
# Stock data
S0 =100;
Sig=0.2;
r=0.01;
miu=0.05;
# Options strikes and maturities
k1=110;
k2=90;
T=5;
# Number of scenarios, confidence level, risk horizon
n=1000000;
confidence=0.95;
h=1;
# Option prices at time 0
```

VaR & ES Vectorized: Long Risk Reversal II

```
d1c=(math.log(S0/k1)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
d1p=(math.log(S0/k2)+(r+0.5*Sig**2)*T)/(Sig*T**0.5);
c0=S0*norm.cdf(d1c)-k1*math.exp(-r*T)*norm.cdf(d1c-Sig*T**0.5);
p0=-S0*norm.cdf(-d1p)+k2*math.exp(-r*T)*norm.cdf(-d1p+Sig*T**0.5);
v0=c0-p0;
""" computing the prices at time h"""
T=T-h;
Zt = np.zeros(n);
St = np.zeros(n);
from random import seed
#seed random number generator
seed(1)
# block generate some random vector Gaussian values and the stock scenarios at time h
Zt = np.random.randn(n);
St=S0*np.exp((miu-0.5*Sig**2)*h)*np.exp(Zt*Sig*(h**0.5))
# Block generating call and put scenarios at time h
d1cnew=(np.log(St/k1)+(r+0.5*Sig**2)*T)/(Sig*(T**0.5));
d1pnew=(np.log(St/k2)+(r+0.5*Sig**2)*T)/(Sig*(T**0.5));
ct=St*norm.cdf(d1cnew)-k1*np.exp(-r*T)*norm.cdf(d1cnew-Sig*T**(0.5));
pt=-St*norm.cdf(-d1pnew)+k2*np.exp(-r*T)*norm.cdf(-d1pnew+Sig*T**(0.5));
# final portolio value in all scenarios
vt=ct-pt;
# loss scenarios
vvar=v0-vt;
vvar=np.sort(vvar);
# extracting VaR from the loss at the right confidence level
ivar = round((confidence)*n);
var = vvar[ivar];
```

VaR & ES Vectorized: Long Risk Reversal III

```
# calculating ES
ESv=statistics.mean(vvar[range(math.floor((confidence)*n),n)]);
print("VaR:",var);
print("ES:",ESv);
# plotting loss histogram
plt.hist(vvar, bins = 100)
plt.show()
```

Vectorized Euler scheme for Geometric Brownian I

```
# -*- coding: utf-8 -*-
"""
Euler scheme for Geometric Brownian Motion
Created on Mon Nov  28 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
from scipy.stats import skew;
from scipy.stats import kurtosis;
import matplotlib.pyplot as plt;
# Stock data
S0 =100;
Sig=0.2;
miu=0.05;
# Final time in years
h=1;
# Number of time steps and time step
nt = 2000; dt = h/nt;
# Number of scenarios
n=100000;
""" Generating the normal 0,1 in 10000 scenarios for each time step """
# Z are normals, S will be the stock
Zt = np.zeros(n);
St = S0*np.ones(n);
```

Vectorized Euler scheme for Geometric Brownian II

```
from random import seed
#seed random number generator
seed(1)
# loop on time
for i in range (nt):
  Zt = np.random.randn(n);
  St=St+miu*St*dt+Sig*St*Zt*(dt**0.5);
#endor
# One shot final simulation using the knonw GBM solution
Sth = S0*np.ones(n);
Zt = np.random.randn(n);
Sth=S0*np.exp((miu-0.5*Sig**2)*h)*np.exp(Zt*Sig*(h**0.5))
# endfor
# plotting final stock histogram
#Blue is St Euler, Orange is St one shot
plt.hist(St, bins = 100, density=True)
plt.hist(Sth, bins = 100, density=True)
plt.show()
# Orange is St Euler, Blue is St one shot
plt.hist(Sth, bins = 100, density=True)
plt.hist(St, bins = 100, density=True)
plt.show()
#
# plotting histogram of the difference
Stsorted = sorted(St);
Sthsorted = sorted(Sth);
diffe = np.subtract(Stsorted,Sthsorted);
plt.hist(diffe, bins = 100, density=True)
```

Vectorized Euler scheme for Geometric Brownian III

```
plt.show()
# comparing mean and standard deviation of log returns
meanlogth = np.log(S0)+(miu-0.5*Sig**2)*h;
meanlogsim = np.average(np.log(Stsorted));
meanlogsimh = np.average(np.log(Sthsorted));
print("log mean theoretical:",meanlogth);
print("log mean simulation Euler:",meanlogsim);
print("log mean simulation one shot:",meanlogsimh);
#
Stdth = Sig*h;
Stdsim = np.std(np.log(Stsorted));
Stdsimh = np.std(np.log(Sthsorted));
print("Std theoretical:",Stdth);
print("Std simulation Euler:",Stdsim);
print("Std simulation one shot:",Stdsimh);
# Comparing skewness and kurtosis
skewSteuler = skew(np.log(Stsorted), axis = 0, bias = True);
skewSth = skew(np.log(Sthsorted), axis = 0, bias = True);
print("Skew Euler:",skewSteuler);
print("Skew one shot:",skewSth);
#
kurtSteuler = kurtosis(np.log(Stsorted), axis = 0, bias = True);
kurtSth = kurtosis(np.log(Sthsorted), axis = 0, bias = True);
print("Kurtosis Euler:",kurtSteuler);
print("Kurtoses one shot:",kurtSth);
```

Vectorized Euler scheme for $dX = mdt + \sigma XdW$ I

```
# -*- coding: utf-8 -*-
"""
Euler scheme for dX = m dt + sigma X dW
Created on Mon Nov 28 15:50:34 2022
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
from scipy.stats import skew;
from scipy.stats import kurtosis;
import matplotlib.pyplot as plt;
# Stock data
S0 =0;
Sig=0.4;
miu=1;
# Final time in years
h=2;
# Number of time steps and time step
nt = 200; dt = h/nt;
# Number of scenarios
n=400009;
""" Generating the normal 0,1 in 10000 scenarios for each time step """
# Z are normals, S will be the stock
Zt = np.zeros(n);
```

Vectorized Euler scheme for $dX = mdt + \sigma XdW$ II

```
St = S0*np.ones(n);
from random import seed
#seed random number generator
seed(1)
# loop on time
for i in range (nt):
# loop on scenarios
  Zt=np.random.randn(n);
  St = St+miu*dt+Sig*St*Zt*(dt**0.5);
#endor
# plotting final stock histogram
#Blue is St Euler, Orange is St one shot
plt.hist(St, bins = 100, density=True)
plt.show()
# comparing mean and standard deviation of the solution
meanth = S0+miu*h;
meansim = np.average(St);
print("mean theoretical:",meanth);
print("mean simulation Euler:",meansim);
#
varth = 2*(miu/(Sig**4))*(miu+S0*Sig**2)*(math.exp(h*Sig**2)-1);
varth = varth - 2*h*(miu/Sig)**2 + math.exp(h*Sig**2)*S0**2;
varth = varth - (S0+miu*h)**2;
Stdth = math.sqrt(varth);
Stdsim = np.std(St);
print("Std theoretical:",Stdth);
print("Std simulation Euler:",Stdsim);
# Comparing skewness and kurtosis
```

Vectorized Euler scheme for $dX = mdt + \sigma XdW$ III

```
skewSteuler = skew(St, axis = 0, bias = True);
print("Skew Euler:",skewSteuler);
#
kurtSteuler = kurtosis(St, axis = 0, bias = True);
print("Kurtosis Euler:",kurtSteuler);
```

Vectorized Euler $dX = (-3kX + 3X^{1/3}\sigma^2)dt + 3\sigma X^{2/3}dW_t$ I

```
# -*- coding: utf-8 -*-
"""
Euler scheme for d X = (-3 k X  +3 X^{1/3}  \sigma^{2} )dt+ 3 \sigma X^{2/3} dW_t
Vectorized and efficient
@author: Prof. Damiano Brigo
======
This file contains Python codes.
======
NB you need to run path\to\anaconda\python.exe in a cmd window for this to see the packages numpy
"""
import numpy as np;
import math;
from scipy.stats import skew;
from scipy.stats import kurtosis;
import matplotlib.pyplot as plt;
# Stock data
S0 =0.5;
Sig=0.05;
kk=1;
# Final time in years
h=1;
# Number of time steps and time step
nt = 400; dt = h/nt;
# Number of scenarios
n=1000000;
```

Vectorized Euler $dX = (-3kX + 3X^{1/3}\sigma^2)dt + 3\sigma X^{2/3}dW_t$ II

```
""" Generating the normal 0,1 in 1000000 scenarios for each time step """
# Z are normals, S will be the stock
Zt = np.zeros(n);
St = S0*np.ones(n);
from random import seed
#seed random number generator
seed(1)
# loop on time
for i in range (nt):
  Zt = np.random.randn(n);
  St = St-3*kk*St*dt+3*((St)**(1/3))*(Sig**2)*dt;
  St=St+3*Sig*((St)**(2/3))*Zt*(dt**0.5);
# endfor
# mean and standard deviation of the solution
meansim = np.average(St);
print("mean simulation Euler:",meansim);
#
Stdsim = np.std(St);
print("Std simulation Euler:",Stdsim);
# Comparing skewness and kurtosis
skewSteuler = skew(St, axis = 0, bias = True);
print("Skew Euler:",skewSteuler);
#
kurtSteuler = kurtosis(St, axis = 0, bias = True);
print("Kurtosis Euler:",kurtSteuler);
```

Vectorized Euler $dX = (-3kX + 3X^{1/3}\sigma^2)dt + 3\sigma X^{2/3}dW_t$ III

```
# Simulation one shot
Zt1s = np.zeros(n);
St1s = S0*np.ones(n);
Zt1s = np.random.randn(n);
St1s=(S0**(1/3)*math.exp(-kk*1)+Zt1s*((Sig**2/(2*kk))*(1-math.exp(-2*kk)))**(1/2))**3;
# endfor
# plotting final stock histogram
#Blue is St Euler, Orange is St one shot
plt.hist(St, bins = 1000, density=True)
plt.hist(St1s, bins = 1000, density=True)
plt.show()
#Blue is One shot,, Orange is Euler
plt.hist(St1s, bins = 1000, density=True)
plt.hist(St, bins = 1000, density=True)
plt.show()
plt.show()
```

APPENDIX: Densities, ω's & discrete vs cont. rv's I

We can have discrete random variables (r.v.), continuous rv or hybrid cases. Let us consider the first two.
A discrete random variable is a random variable that takes only a finite or countable number of possible values. Examples of experiments leading to discrete random variables are counting the number of heads in n tosses of a fair coin (Binomial), a random variable describing whether you get head or tail in a single toin coss (Bernoulli), a random variable describing the results of a dice roll, etc. Let us take the example of a single fair coin toss. Here

$$\Omega = \{\omega_1 = \text{head}, \omega_2 = \text{tail}\}$$

and the sigma-field is

$$\mathcal{F} = \{\Omega, \Phi, \{\omega_1\}, \{\omega_2\}\}.$$

APPENDIX: Densities, ω's & discrete vs cont. rv's II

Consider a Bernoulli random variable $X : \Omega \to \{0, 1\} \subset \mathbb{R}$ counting the number of heads in the single coin toss. With a fair coin, the probability of having head or tail is $1/2$. So we have the following probability measure on $(\Omega, \mathcal{F})$.

$$\mathbb{P}(\omega \in \Omega : X(\omega) = 0) = \mathbb{P}(\{\omega_2\}) = \frac{1}{2}, \quad \mathbb{P}(\omega \in \Omega : X(\omega) = 1) = \mathbb{P}(\{\omega_1\}) = \frac{1}{2}.$$

So the random variable only takes two values, 0 and 1, with probability $1/2$ and $1/2$ respectively, and these two values correspond to the events $\{\omega_2\}$ and $\{\omega_1\}$.

Note that, as X is a random variable, it is measurable, so for example

$$\{X = 1\} = \{X \in \{1\}\} = X^{-1}(\{1\}) = \{\omega \in \Omega : X(\omega) = 1\} = \{\omega_2\} \in \mathcal{F}.$$

APPENDIX: Densities, ω's & discrete vs cont. rv's III

Similarly,

$$\{X = 0\} = \{X \in \{0\}\} = X^{-1}(\{0\}) = \{\omega_1\} \in \mathcal{F}.$$

In this example, the correspondence between the random variable values 0 1, and the original elementary events ω and the probabilities is clear. It is also clear to what events in $\mathcal{F}$ the probabilities of the random variable are associated, as shown in the last two equations.

The situation may be a little more confusing with continuous random variables at first.

For continuous random variables, say a standard Gaussian $X \sim \mathcal{N}(0, 1)$, if you calculate $P(X = 0) = P(\omega : X(\omega) = 0) = 0$, you obtain a 0 probability, because the probability that a continuous random variable takes a single value is zero.

APPENDIX: Densities, ω's & discrete vs cont. rv's IV

This is clearly at odds with the example of the discrete random variable we gave above, where a single value of the random variable had non-zero probability equal to $1/2$.

However, the types of minimal set we are interested in for continuous random variables, for computing probabilities, are different, and the smallest ones lead to densities. We need to replace single discrete points, like $\{0\}$ or $\{1\}$ of the Bernoulli case, with infinitesimal intervals around points.

So, for example, now for our Gaussian X the minimalist calculation to do is, for an infinitesimal interval $[x, x + dx)$ which we could take to be $[0, 0 + dx)$,

$$P(X \in [0, 0 + dx)) = P(\omega : X(\omega) \in [0, 0 + dx)) = p_X(0)dx$$

APPENDIX: Densities, ω's & discrete vs cont. rv's V

where $p_X(0)$ is the probability density function of X, or the PDF of a standard Gaussian calculated in 0. More generally,

$$P(X \in [x, x + dx)) = P(\omega : X(\omega) \in [x, x + dx)) = p_X(x)dx$$

where $p_X(x)$ is the Gaussian density in x.

This equality follows remembering that the density is the derivative of the cumulative distribution function. Indeed,

$$F_X(x) = P(X < x) \Rightarrow p_X(x) = \frac{d}{dx} F_X(x) \ \Rightarrow \ p_X(x)dx = dF_X(x) \ \Rightarrow$$

$$p_X(x)dx \approx F_X(x + dx) - F_X(x) = P(X < x + dx) - P(X < x)$$
$$= P(X \in [x, x + dx)).$$

APPENDIX: Densities, ω's & discrete vs cont. rv's VI

Hence we confirm

$$P(\omega \in \Omega : \; X(\omega) \in [x, x + dx)) = p_X(x)dx$$

Here note that even if the interval collapses to a point ($dx = 0$) and you would get $P(X \in [x, x + dx)) = p_X(x)0 = 0$ because $dx = 0$, the density $p_X(x)$ would still be non-zero.

Note also that $[x, x + dx)$ is an interval in $\mathbb{R}$ and thus a Borel set of $\mathbb{R}$. Since X is a random variable, it is measurable, and

$$\{X \in [x, x+dx)\} = X^{-1}([x, x+dx)) = \{\omega \in \Omega : \; X(\omega) \in [x, x+dx)\} \in \mathcal{F}$$

is an event, and we can compute its probability, as above.

APPENDIX: Densities, ω's & discrete vs cont. rv's VII

Now continuous distributions have probability densities like p_X. You can sample from p_X (basic statistics) and get some values for the random variable X that is distributed as p_X. Of course, you will get more often values x where the probability density $p_X(x)$ is more concentrated, or higher, and less values x in areas where the probability $p_X(x)$ is smaller.

When we sample like this it is as if we were sampling the different ω's that lead to the different $p_X(x)$ values in the ways that we showed above. The different colours in the Brownian paths in our SDE simulation early in the course, for example, are different ω's, corresponding to different values of the SDE solution density at different times. For two different x_1 and x_2 you will have the different ω's corresponding to the ω sets associated with $p_X(x_1)$ and $p_X(x_2)$ and in particular with

APPENDIX: Densities, ω's & discrete vs cont. rv's VIII

$X^{-1}([x_1, x_1+dx)) = \{\omega \in \Omega : \ X(\omega) \in [x_1, x_1+dx)\} \in \mathcal{F}$ with probability $p_X(x_1)dx$

and

$X^{-1}([x_2, x_2+dx)) = \{\omega \in \Omega : \ X(\omega) \in [x_2, x_2+dx)\} \in \mathcal{F}$ with probability $p_X(x_2)dx$.

The key thing to keep in mind here is that you need intervals $[x, x + dx)$ instead of points, but being dx infinitesimal, you are close to having a point and, again, to have the analogy with the discrete case of the Bernoulli.

APPENDIX: On the notion of independence I

In the lecture notes we have given the following definition of independent random variables for two random variables X and Y on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
X and Y are independent if

$$P(\{X \in A\} \cap \{Y \in B\}) = P(\{X \in A\})P(\{Y \in B\}) \text{ for all } A, B \in \mathcal{B}(\mathbb{R}).$$

Here A and B are two Borel sets of $\mathbb{R}$, namely two subsets of $\mathbb{R}$ in the sigma-field generated by open intervals of $\mathbb{R}$. As an initial example, you might think of two intervals $A = (a_1, a_2)$ and $B = (b_1, b_2)$ with $a_1 < a_2$ and $b_1 < b_2$ real numbers. However, Borel sets may be much more complex than this.
Given that independence is extremely important in our course (e.g. independent increments of Brownian motion, independent Gaussian random variables, etc) we can make the definition more explicit.

APPENDIX: On the notion of independence II

The above definition is short-hand for

$$P(\{\omega \in \Omega : X(\omega) \in A\} \cap \{\omega \in \Omega : Y(\omega) \in B\}) =$$

$$= P(\{\omega \in \Omega : X(\omega) \in A\}) P(\{\omega \in \Omega : Y(\omega) \in B\}).$$

Note that because X is a random variable (and similarly for Y), X is a measurable function $X : \Omega \to \mathbb{R}$, so that, by definition of measurability,

$$\{X \in A\} = X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\},$$

$$\{Y \in B\} = Y^{-1}(B) = \{\omega \in \Omega : Y(\omega) \in B\}$$

are both in $\mathcal{F}$. Moreover, as $\mathcal{F}$ is a sigma-field, intersection of sets in $\mathcal{F}$ is still in $\mathcal{F}$, so also

$$\{X \in A\} \cap \{Y \in B\} = X^{-1}(A) \cap Y^{-1}(B) =$$

APPENDIX: On the notion of independence III

$$= \{\omega \in \Omega : X(\omega) \in A\} \cap \{\omega \in \Omega : Y(\omega) \in B\}$$

belongs in $\mathcal{F}$. It follows we can compute the probability of all three events of $\mathcal{F}$, and independence is defined as the probability of the intersection event being equal to the product of the probabilities of the intersecting events.

If, after you have read the definition of multivariate random variable and multivariate CDF, you go back to the definition of independence, we can consider, in the above definition, $A = (-\infty, x]$ and $B = (-\infty, y]$. We can see then that independence implies

$$F_{X,Y}(x, y) = F_X(x)F_Y(y),$$

i.e. the joint CDF of the random vector $[X, Y]$ is the product of the two CDFs of the single random variables X and Y. This can also be adopted as an alternative definition of independence.

APPENDIX: On the notion of independence IV

If the two random variables admit a probability density, and the random vector $[X, Y]$ does too, then another way to express independence is

$$p_{[X,Y]}(x, y) = p_X(x)p_Y(y),$$

namely the density of the random vector $[X, Y]$ is the product of the densities of X and Y.

We have noted the important property that the variance of the sum of two independent random variables is the sum of their variances:

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

This is a very important property of independent random variables.

APPENDIX: Intuition behind Girsanov's theorem I

If you are clear about the Girsanov theorem, its purpose, what it does, why it is important, you can skip this part and jump to part 2 on option pricing. However, typically, students when first exposed to this theorem need a little more context and some additional explanation. That is given here in case you need it.

Also, en route we show why the drift and diffusion coefficient of SDEs are the local mean and local standard deviation.

First, recall that any definition of Brownian motion is related to a probability measure: when we introduced Brownian motion, our original W is a Brownian motion under $(\Omega, \mathcal{F}, \mathbb{P})$. This can be emphasized by renaming W as W^P. W^P may not be a Brownian motion under a different probability measure Q, as increments may not be independent under Q, or Q-distribution is not Gaussian, etc.

APPENDIX: Intuition behind Girsanov's theorem II

The fact W^P is only Brownian under P, means that it is possible to re-write the same SDE under a different probability measure Q. It will be the same SDE, for the same process, but it will be expressed under a different measure Q different from P. This means that the statistical properties of the SDE, like the local mean and local standard deviation (drift and diffusion coeff.) will change as we move under Q and use a Brownian motion under Q.

Let us illustrate this first with a static random variable. Take a Bernoulli random variable taking only two values with different probabilities

$$\begin{aligned} \Omega &= \{\omega_1, \omega_2\}, \ \ \mathcal{F} = \{\Phi, \Omega, \{\omega_1\}, \{\omega_2\}\} \\ X &: \ \Omega \to \{0, 1\}, \ \ X(\omega_1) = 0, \ X(\omega_2) = 1. \end{aligned} \tag{47}$$

APPENDIX: Intuition behind Girsanov's theorem III

This random variable X is defined on $(\Omega, \mathcal{F})$. We can easily introduce two different probability measures in this space. Let's start with the masure P:

$$P(X = 0) = P(\{\omega_1\}) = \frac{1}{2}; \quad P(X = 1) = P(\{\omega_2\}) = \frac{1}{2}.$$

Then in $(\Omega, \mathcal{F}, P)$ the mean and variance of X are

$$E^P[X] = 0P(X = 0) + 1P(X = 1) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}.$$

$$Var^P[X] = E^P[X^2] - E^P[X]^2 = 0^2 P(X = 0) + 1^2 P(X = 1) - \left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

APPENDIX: Intuition behind Girsanov's theorem IV

Let's now use a second probability measure Q where

$$Q(X=0) = Q(\{\omega_1\}) = \frac{1}{4}; \quad Q(X=1) = Q(\{\omega_2\}) = \frac{3}{4}.$$

Then in $(\Omega, \mathcal{F}, Q)$ the mean and variance of X are

$$E^Q[X] = 0Q(X=0) + 1Q(X=1) = 0 \times \frac{1}{4} + 1 \times \frac{3}{4} = \frac{3}{4}.$$

$$Var^Q[X] = E^Q[X^2] - E^Q[X]^2 = 0^2 Q(X=0) + 1^2 Q(X=1) - \left(\frac{3}{4}\right)^2 = \frac{3}{16}.$$

Important: X on $(\Omega, \mathcal{F})$ has not changed, it is the same X as in Eq. (47). So it's always the same X. But under the two different measures it has different means $E^P[X]$ vs $E^Q[X]$ and standard deviations (variances) $Var^P[X]$ vs $Var^Q[X]$.

APPENDIX: Intuition behind Girsanov's theorem V

With SDEs it will be the same, the process X that is the solution of the SDE will remain the same, but if expressed under two different probability measures P and Q then the (local) mean $E_t[dX_t]$ and the local variance $Var_t[dX_t]$ of the same process dX_t will be different. We show this now.

So let's say we have a SDE describing a stochastic process X under the measure P. The SDE is

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t^{\mathbb{P}}$$

Now notice that both f and σ depend on the probability measure P. Indeed, for example, informally, mean and variance of the SDE increment conditional on $\mathcal{F}_t$ (denoted E_t^P and Var_t^P) are

APPENDIX: Intuition behind Girsanov's theorem VI

$$E_t^P[dX_t] = E_t^P[f(X_t)dt + \sigma(X_t)dW_t^P] = f(X_t)dt + \sigma(X_t)E_t^P[dW_t^P]$$

$$= f(X_t)dt + \sigma(X_t)\ 0 = f(X_t)dt$$

given that $E^P[dW_t^P] = 0$ and that both $f(X_t)$ and $\sigma(X_t)$ are $\mathcal{F}_t$ measurable and can be brought out of the expectation, and

$$\mathrm{Var}_t^P[dX_t] = \mathrm{Var}_t^P[f(X_t)dt + \sigma(X_t)dW_t^P] = \mathrm{Var}_t^P[\sigma(X_t)dW_t^P]$$

$$= \sigma(X_t)^2\mathrm{Var}_t^P[dW_t^P] = \sigma(X_t)^2 dt$$

given that, conditional on $\mathcal{F}_t$, $f(X_t)$ is a deterministic constant and does not contribute to the variance, $\sigma(X_t)$ is also deterministic and can be taken out of the variance, and the variance of dW is dt.

APPENDIX: Intuition behind Girsanov's theorem VII

So we get that

$$E_t^P[dX_t] = f(X_t)dt, \quad \mathrm{Var}_t^P[dX_t] = \sigma(X_t)^2 dt$$

that justify the names "local mean" for f and "local standard deviation" for σ, and we see clearly that f and σ depend on the measure P we are using, as they are related to the mean and variance <u>under P</u> of dX_t.

To emphasize this we will write $f = f^P$, $\sigma = \sigma^P$.
With this in mind, we can rewrite our SDE for X under P as

$$\mathrm{d}X_t = f^{\mathbb{P}}(X_t)\,\mathrm{d}t + \sigma^P(X_t)\,\mathrm{d}W_t^{\mathbb{P}}$$

In some applications we may want to change the drift (local mean) for various reasons, depending on the field of science (math finance, stochastic filtering, etc).

APPENDIX: Intuition behind Girsanov's theorem VIII

In finance, we will see later that we have the Feynman Kac theorem that will be important to compute option prices as discounted expectations and this will require a drift change in the SDE.

So, for some applications we may need to work with a SDE for the same process X but with a different drift and diffusion coeff. Since drift and diffusion coefficient are induced by the Brownian motion, and the Brownian motion is such only under a particular measure, if we want to change drift and diffusion coefficient we need to change the Brownian motion and hence the measure. Let's write a new SDE for the same dX under a measure Q, $\mathrm{d}X_t = f^Q(X_t)\,\mathrm{d}t + \sigma^Q(X_t)\,dW^Q$.

Now we may wonder what this measure Q is, how to characterize it. In this, the Girsanov theorem comes in. It tells us that if we have two

APPENDIX: Intuition behind Girsanov's theorem IX

different SDEs describing the same process dX_t using Brownian motions under two different probability measures P and Q, namely

$$\mathrm{d}X_t = f^{\mathbb{P}}(X_t)\,\mathrm{d}t + \sigma^{\mathbb{P}}(X_t)\,\mathrm{d}W_t^{\mathbb{P}}, \quad \mathrm{d}X_t = f^{\mathbb{Q}}(X_t)\,\mathrm{d}t + \sigma^{\mathbb{Q}}(X_t)\,\mathrm{d}W_t^{\mathbb{Q}}$$

then we can do this with measures $\mathbb{P}$ and $\mathbb{Q}$ that are equivalent only if $\sigma^P = \sigma^Q = \sigma$. So we need to have the same σ under P and Q for the measures to be equivalent.

What does equivalence of P and Q mean? It means that P and Q agree on which events are certain or impossible.
$Q(A) = 1 \iff P(A) = 1$ or $Q(A) = 0 \iff P(A) = 0$. We don't want things that are certain for the measure P (P probability one) to be uncertain for the measure Q (Q probability less than one). For example, the probability that a stock price X is positive. In Black

APPENDIX: Intuition behind Girsanov's theorem X

Scholes we will see later, with stock X_t, we have $P(X_t > 0) = 1$. So prices X are positive with probability one, almost surely. Equivalence of the measures means that also $Q(X_t > 0) = 1$.

If measures were not equivalent and for example $P(X_t > 0) = 1$, $Q(X_t > 0) = 0.6$, this would spell troubles because a positive stock price X under the original probability measure P could become negative under the pricing measure Q, with Q probability 0.4, and this would be bad, because the stock paths are a given, and we cannot have them be positive and non positive at the same time. So the measures need to agree on this (and other things). Equivalence of the measures, given by the Girsanov theorem, gives us these fundamental agreements.

APPENDIX: Intuition behind Girsanov's theorem XI

Thus assuming equivalence, and hence the same σ, we may write the process dX_t under the two measures as

$$\mathrm{d}X_t = f^{\mathbb{P}}(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t^{\mathbb{P}}, \quad \mathrm{d}X_t = f^{\mathbb{Q}}(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t^{\mathbb{Q}}$$

Then the Girsanov theorem tells us that the measures P and Q will be equivalent and that the relationship between P and Q is described in terms of the difference of the drifts of the two SDEs, divided by the diffusion coefficient $(f^Q - f^P)/\sigma$, called "market price of risk" in finance, and of the original Brownian motion W^P:

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} := \exp\left[-\frac{1}{2}\int_0^t \left(\frac{f^{\mathbb{Q}}(X_s) - f^{\mathbb{P}}(X_s)}{\sigma(X_s)}\right)^2 \mathrm{d}s + \int_0^t \frac{f^{\mathbb{Q}}(X_s) - f^{\mathbb{P}}(X_s)}{\sigma(X_s)}\,\mathrm{d}W_s^{\mathbb{P}}\right]$$

APPENDIX: Intuition behind Girsanov's theorem XII

So the market price of risk, depending on $(f^Q - f^P)/\sigma$, integrated against W^P, defines the Randon Nykodym derivarive dQ/dP defining the measure Q given the measure P.

We can also derive a direct relationship between the two Brownian motions. As $dX_t = dX_t$ we can write

$$f^{\mathbb{P}}(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t^{\mathbb{P}} = f^{\mathbb{Q}}(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t^{\mathbb{Q}}$$

as both sides equal the same dX_t. Solving in dW^Q we get

$$dW_t^Q = \frac{f^{\mathbb{P}}(X_t) - f^{\mathbb{Q}}(X_t)}{\sigma(X_t)}\mathrm{d}t + \mathrm{d}W_t^{\mathbb{P}} \text{ or } W_t^Q = \int_0^t \frac{f^{\mathbb{P}}(X_s) - f^{\mathbb{Q}}(X_s)}{\sigma(X_s)}ds + W_t^{\mathbb{P}}$$

APPENDIX: Process adapted to a filtration I

What do we mean when we say that, $(X_t)_{t\geq 0}$ being adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, means that, for a given t, $\mathcal{F}_t$ "knows" all the information on X_t and of its past X_s, $s < t$?

$(X_t)_{t\geq 0}$ being adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ is defined as X_t being $\mathcal{F}_t$ measurable for all $t \geq 0$. This means that for every Borel set of the real numbers B, we have

$$\{X_t \in B\} = \{\omega \in \Omega : \; X_t(\omega) \in B\} = X_t^{-1}(B) \in \mathcal{F}_t$$

for all t. In particular we can take $B = [x, x + dx)$ and we get

$$\{X_t \in [x, x + dx)\} \in \mathcal{F}_t$$

for all t. So, $\mathcal{F}_t$ "knows", for every real value x, whether X_t is near x or not, as the event $\{\omega \in \Omega : X_t(\omega) \in [x, x + dx)\}$ is included in the

APPENDIX: Process adapted to a filtration II

filtration $\mathcal{F}_t$. As $\mathcal{F}_t$ models the market events/information that are know at time t, this means that X_t is known at time t under $\mathcal{F}_t$ (we know if it is near x for every $x \in \mathbb{R}$).

Now remember that filtrations are increasing sigma fields. This means that if $s < t$ then $\mathcal{F}_s \subseteq \mathcal{F}_t$. Then if we take $(X_t)_{t\geq 0}$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and fix a $s < t$, by definition of "adapted" we have that X_s is measurable or "known" to $\mathcal{F}_s$, namely $\{X_s \in B\} \in \mathcal{F}_s$ for all Borel sets B. But as $\mathcal{F}_s \subseteq \mathcal{F}_t$, it follows also that $\{X_s \in B\} \in \mathcal{F}_t$ for all B, and in particular, using $B = [x, x + dx)$, we can see that $\mathcal{F}_t$ "knows" the x near which X_s will be, as $\{X_s \in [x, x + dx)\} \in \mathcal{F}_t$.

We have illustrated how $\mathcal{F}_t$ knows X_s for all $s \leq t$ if X is adapted.

APPENDIX: Process adapted to a filtration III

Note also that if we consider the filtration generated by the process X itself, namely, for every $t \geq 0$,

$$\mathcal{F}_t^X = \sigma(\{\omega \in \Omega : X_s(\omega) \in B\}, \text{ for all } B \in \mathcal{B}(\mathbb{R}) \text{ and all } s \leq t\}$$

then clearly X is adapted to this filtration, because by construction the filtration is generated by all events $\{X_s \in B\}$ for all $s \leq t$ and all B, which clearly implies that X_t is $\mathcal{F}_t^X$ measurable as obviously $\{X_t \in B\} \in \mathcal{F}_t^X$ for all t. Here remember the notation σ(set of events) means the smallest sigma-field containing all the events in the set. It is given by all possible countable unions, countable intersections, complementations and combinations of these operations on the given set of events.

In the existence and uniqueness theorem for SDEs we have seen that the solution X is $\mathcal{F}^{W,Z}$ adapted. This means that if we take the

APPENDIX: Process adapted to a filtration IV

filtration generated by both the process W and the random variable Z, where Z is the initial condition and W the Brownian motion in the SDE, then X_t is $\mathcal{F}_t^{W,Z}$-measurable for all t, meaning that X_t is known by the filtration $\mathcal{F}_t^{W,Z}$. We explain this. Remember that by definition $\mathcal{F}_t^{W,Z} =$

$$= \sigma(\{\omega \in \Omega : W_s(\omega) \in B_1\}, \{\omega \in \Omega : Z(\omega) \in B_2\} \text{for all } B_1, B_2 \in \mathcal{B}(\mathbb{R}) \ \& \ s \leq t\}.$$

In particular, for every $s \leq t$, the filtration contains the event $\{W_s \in [x, x + dx)\}$ for all x, so that we know, for the whole history of W, to which points x it was close at every time $s \leq t$, so in practice we know the path of W from time 0 to time t. We also know $\{Z \in [x, x + dx)\}$ for all x, so we know the value of the initial condition, as we know the point to which it was infinitesimally close. Knowing the whole path of W up to t and knowing the initial condition Z enables us, intuitively, to know the solution X_t of the SDE at time t since this is built

APPENDIX: Process adapted to a filtration V

starting from $X_0 = Z$ and then propagating the dynamics, adding in every infinitesimal time $[s, s + ds)$, step by step, the increment dW_s which is known to the filtration $\mathcal{F}_t^{W,Z}$.
You can understand this further by writing the SDE solution as a converging Euler scheme that we will see in the final part of the course. The solution to the SDE

$$dX_s = \mu(s, X_s)ds + \sigma(s, X_s)dW_s, \;\; X_0 = Z$$

can be expressed as a converging iterative numerical scheme with time step Δs tending to zero, given by

$$X_{s+\Delta s} = X_s + \mu(s, X_s)\Delta s + \sigma(s, X_s)\Delta W_s$$

APPENDIX: Process adapted to a filtration VI

The scheme starts with $X_0 = Z$ and the first step is

$$X_{0+\Delta s} = Z + \mu(0, Z)\Delta s + \sigma(0, Z)\Delta W_0$$

Now both Z and $\Delta W_0 = W_{\Delta s} - W_0$ are known to $\mathcal{F}_t^{W,Z}$, as $\Delta s < t$, and so $X_{\Delta s}$ is known to $\mathcal{F}_t^{W,Z}$ as $\mathcal{F}_t^{W,Z}$ knows all the parts on the right hand side of the expression for $X_{\Delta s}$. Next step is

$$X_{2\Delta s} = X_{\Delta s} + \mu(\Delta s, X_{\Delta s})\Delta s + \sigma(s, X_{\Delta s})\Delta W_{\Delta s}$$

Again, from the previous step, $X_{\Delta s}$ is known to $\mathcal{F}_t^{W,Z}$, and also $\mu(\Delta s, X_{\Delta s})$ and $\sigma(s, X_{\Delta s})$ (measurable functions of $X_{\Delta s}$) and $\Delta W_{\Delta s} = W_{2\Delta s} - W_{\Delta s}$ are known to $\mathcal{F}_t^{W,Z}$ as $2\Delta s < t$, and so $\mathcal{F}_t^{W,Z}$ knows $X_{2\Delta s}$ as it knows all the parts on the right hand side of the expression for $X_{2\Delta s}$.
Continuing like this until $n\Delta_s = t$ we conclude that $X_{n\Delta s} = X_t$ is known to the filtration $\mathcal{F}_t^{W,Z}$.

APPENDIX: Fitting the smile in practice I

Given a market volatility smile curve, a curve of implied volatilties for some strikes $K_1, \dots, K_m$, the Black-Scholes implied volatilities being $\nu^{mkt}(K_1), \dots, \nu^{mkt}(K_m)$, as explained in the main lectures, how do we find the model parameters that fit the market smile curve $K_i \mapsto \nu^{mkt}(K_i)$?

This is done through model calibration. Recall the steps the model uses to generate the model smile curve (here model can be Bachelier, DD, CEV, Mixture dynamics, Heston),

1. Start with $i = 1$ and thake $K = K_1$ as initial strike;
2. Compute the model ("mod") call price

$$V_{Model}(K_i) = E_0^Q[e^{-rT}(S_T - K_i)^+]$$

with S modeled through a smile dynamics

$$\text{Model:} \;\; dS_t = rS_t dt + \boxed{\sigma(t, S_t)} S_t \, dW_t^Q, \;\; S_0 = s_0$$

APPENDIX: Fitting the smile in practice II

3. Invert Black Scholes formula for this strike, i.e. solve

$$V_{Model}(K_i) = V_{BS}(0, S_0, K_i, T, \nu^{mod}(K_i), r).$$

   in $\nu^{mod}(K_i)$, thus obtaining the model implied volatility $\nu^{mod}(K_i)$.
4. If $i = m$ end, otherwise assign $i = i + 1$ (in the coding sense) and restart from point 2.

At the end of this algorithm we have built the smile curve $K_i \mapsto \nu^{mod}(K_i)$ for this model. Depending on the model (Bachelier, DD, CEV) the curve can be monotonic or also allow (mixure dynamics, Heston) for a U shape.

To fit the model to market, what one does, if β is the vector of model parameters, which in the case of Bachelier is just $\beta = (\sigma)$, for DD $\beta = (\alpha, \sigma)$, for CEV $\beta = (\gamma, \sigma)$, for the shifted mixture dynamics

APPENDIX: Fitting the smile in practice III

$\beta = (\alpha, \lambda_1, \ldots, \lambda_{N-1}, \sigma_1, \ldots, \sigma_N)$, and for the Heston model $\beta = (k, \theta, \sigma_V, v_0, \rho)$, one finds the best fit across all strikes by running a minimization of the squared error, like in nonlinear least squares for example:

$$\beta^* = \text{argmin}_\beta \sum_{i=1}^{m} \left(V_{BS}^{Call}(S_0, K_i, \nu^{mkt}(K_i)) - V_{mod}^{Call}(S_0, K_i; \beta) \right)^2$$

Python or Matlab have packages for the minimization, with local or global methods.
If the minimization error is large, in models like the mixture dynamics you may increase the number of component N of the mixture. In the picture we have seen for the shifted mixture dynamics in the main lecture notes, we could fit almost perfectly the market smile (blue curve) with only two components (red curve), meaning one λ, two σ's,

APPENDIX: Fitting the smile in practice IV

and one α, four parameters in total. Other cases may require 3 components, so 6 parameters in total, etc.
Once the optimization is done, and one has β^*, one knows the best model parameters fitting the market smile in the chosen model, and can generate a plot of the model smile, to be compared with the market one, using the algorithm above with steps 1-4.

APPENDIX: Local vol & future smile flattening I

Here we expand more in detail the explanation on the future smile problem flattening in local volatility models, and explain why it is a problem.

Local volatility models. Models like Bachelier, CEV, Displaced Diffusion and Mixture Dynamics are local volatility models. In these models the return volatility $\sigma(t, S_t)$ in the SDE for dS_t,

$$dS_t = rS_t dt + \boxed{\sigma(t, S_t)} S_t dW_S^Q(t), \quad s_0$$

is a function of t and S only, and there is no other randomness entering the stock price instant by instant, except for dW_S. Here we use notation $dW(t)$ instead of dW_t to avoid confusion for the indices.

APPENDIX: Local vol & future smile flattening II

Compare with a **stochastic volatility model** like Heston. In this model

$$dS_t = rS_t dt + \boxed{\sqrt{V_t}} S_t dW_S^Q(t), \quad s_0$$

$$dV_t = k(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_V(t), \quad v_0,$$

$$dW_S dW_V = \rho \, dt.$$

In this model, the return volatility $\sqrt{V_t}$ in the SDE for dS_t is not just a function of S, and of its randomness driven by W_S. Instead, the return volatility $\sqrt{V_t}$ depends on a new randomness driven by a second Brownian motion, W_V. This is like saying that in stochastic volatility models, the return volatility has a random life of its own, and it is not just driven by S_t alone, or even constant like in Black Scholes. In technical language, we can say that in a local volatility model, the

APPENDIX: Local vol & future smile flattening III

volatility $\sigma(t, S_t)$ is $\mathcal{F}_t^{W_S^Q}$ measurable, whereas in a stochastic volatility model, the volatility $\sqrt{V_t}$ is not $\mathcal{F}_t^{W_S^Q}$ measurable, you need $\mathcal{F}_t^{W_V^Q}$ too.

Local volatility models suffer from the fact that their future volatility smile tends to flatten.
What do we mean by future volatility smile?
This is the volatility smile produced by the local volatility model dS_t but not at time 0 for a maturity T, but at a future time T_1 with maturity $T_1 + T$.
In the main lecture notes we gave a relatively detailed explanation of the future volatility smile, but if some student found that confusing, we now present an alternative definition. Recall as an alternative local volatility model can generate the present smile at $t = 0$.

APPENDIX: Local vol & future smile flattening IV

1. Set K to a starting value;
2. Compute the model call option price at $t = 0$ for time to maturity T

$$V_{Model}(0, K, T) = E_0^Q[e^{-rT}(S_T - K)^+]$$

with S modeled through an alternative dynamics

$$\text{Model:}\ \ dS_t = rS_t dt + \boxed{\sigma(t, S_t)} S_t\, dW_S^Q(t),\ \ t \geq 0,\ \ S_0 = s_0$$

3. Invert Black Scholes formula for this strike, i.e. solve

$$V_{Model}(0, K, T) = V_{BS}(0, S_0, K, T, \nu(0, K, T), r),$$

(where now the T in V_{BS} is the time to maturity and not the maturity) in $\nu(0, K, T)$, the model implied volatility.
4. Change K and restart from point 2.

At the end of this algorithm we have built the smile curve at time 0 for time to maturity T, namely $K \mapsto \nu(0, K, T)$, for this model.

APPENDIX: Local vol & future smile flattening V

This is the smile at time 0 for maturity $0 + T = T$ because the SDE for dS_t is started at time 0 with S_0, and the matching with the Black Scholes option price, to be solved in the implied volatility $\nu(0, K, T)$, is done at time 0 for time to maturity T (and maturity $0 + T = T$).

Let us now see how the future smile, at a future time T_1, for the same time to maturity T (and maturity $T_1 + T$), is generated similarly. In other terms, we want to build the curve $K \mapsto \nu(T_1, K, T)$, for the implied volatility at time T_1 for time to maturity T (and maturity $T_1 + T$).

APPENDIX: Local vol & future smile flattening VI

1. Set K to a starting value;
2. Compute the model call option price at the future time T_1, time to maturity T: $V_{Model}(T_1, K, T; s_1) = E^Q[e^{-rT}(S_{T_1+T} - K)^+ | S_{T_1} = s_1]$ with S modeled through an alternative dynamics

$$\text{Model: } dS_t = rS_t dt + \boxed{\sigma(t, S_t)} S_t \, dW_S^Q(t), \ \boxed{t \geq T_1}, \ S_{T_1} = s_1$$

3. Invert Black Scholes formula for this strike, i.e. solve

$$V_{Model}(T_1, K, T; s_1) = V_{BS}(T_1, s_1, K, T, \nu(T_1, K, T; s_1), r).$$

   in $\nu(T_1, K, T; s_1)$, thus obtaining the model implied volatility.
4. Change K and restart from point 2.

At the end of this algorithm we have built the smile curve at the future time T_1, conditional on the stock value s_1 at T_1, for the time to maturity T, namely the map $K \mapsto \nu(T_1, K, T; s_1)$ for this model.

APPENDIX: Local vol & future smile flattening VII

Note first of all that in the future smile at T_1 the SDE is started at a time T_1, whereas in the current smile at time 0 the SDE is started at time 0.

Flattening of the future smile means the following. Comparing

$$K \mapsto \nu(0, K, T; s_0), \quad \text{with} \quad K \mapsto \nu(T_1, K, T; s_1)$$

means that the second smile will be much flatter than the first, for reasonable choices of the future stock scenario s_1 at T_1. This happens if you use a local volatility model like Bachelier, DD, CEV or the mixture dynamics. It won't happen if you use a stochastic volatility model like Heston.

But why is this happening and why is this important?

APPENDIX: Local vol & future smile flattening VIII

The future smile price is an integral

$$V_{Model}(T_1, K, T; s_1) = E^Q[e^{-rT}(S_{T_1+T} - K)^+ | S_{T_1} = s_1]$$

$$= e^{-rT} \int (x - K)^+ p_{S_{T_1+T}|S_{T_1}=s_1}(x) dx$$

where it is important to notice that the payoff is integrated against the *transition* density $p_{S_{T_1+T}|S_{T_1}=s_1}$ of S between T_1 and $T_1 + T$, as opposed to the marginal density $p_{S_T|s_0}$ that is used at time 0.
A first argument for why the smile flattens is that in local volatility models, we can say that the impact of the diffusion coefficient $\sigma(t, S_t)$ in the SDE will be larger for $p_{S_T|s_0}$ than for $p_{S_{T_1+T}|S_{T_1}=s_1}$, so that, for large T_1, the impact of $\sigma(t, S_t)$ will tend to become close to that of a constant, and thus similar to Black Scholes (flat smile).

APPENDIX: Local vol & future smile flattening IX

To prove this we would need to do some work with the PDE for the model probability density, the Fokker Planck or Kolmogorov equation, but this is beyond the scope of this course.
Other attempts at proving this can be made, one could try to use the theory of large deviations, or generalized versions of the central limit theorem for the log returns. Here I will only say that this is an empirically verifiable feature. I tested this for example in the mixture dynamics by simulation, see table 2 of
`https://arxiv.org/pdf/0812.4052`

Stochastic volatility models like Heston avoid this. For those, one would have to consider the density of the two-dimensional process $p_{S_t,V_t}(x,y)$ and write the Fokker Planck or Kolmogorov PDE for that. The presence of a new stochastic process V_t avoids the pitfalls of local vol models, & the flattening can be avoided or significantly reduced.

APPENDIX: Local vol & future smile flattening X

Again, we can give no proof of this with our current tools, except verify it by simulation, but this is a very important point for traders and a reason why they prefer stochastic volatility models for many asset classes, including equity and FX.

Why is the flattening a bad feature for option traders?
There are options that are much more complicated than the options we saw here, like path dependent options, etc. These options prices depend on future transition densities of the stock, $p_{S_{T_1+T}|S_{T_1}}$ and not only on marginal densities as for the pricing examples we have seen so far.
An example is a ratchet option: this is an option with payoff

$$(S_{T_1+T} - S_{T_1})^+$$

APPENDIX: Local vol & future smile flattening XI

paid at $T_1 + T$. The price is, accordingly,

$$V_0 = e^{-r(T_1+T)} E^Q\{(S_{T_1+T} - S_{T_1})^+\} =$$

$$= e^{-rT_1} E^Q\{e^{-rT} E^Q[(S_{T_1+T} - S_{T_1})^+ | S_{T_1}]\}$$

where we used the tower property. Now, to calculate the internal expectation/future option price, we have to integrate

$$e^{-rT} E^Q[(S_{T_1+T} - S_T)^+ | S_{T_1}] = e^{-rT} \int (x - S_{T_1})^+ p_{S_{T_1+T}|S_{T_1}}(x) dx$$

and this future price at T_1, depending on the transition density $p_{S_{T_1+T}|S_{T_1}}$, will depend on the future smile at T_1 for maturity $T_1 + T$. If this future smile is considerably flatter than the smile at time 0, this will influence the valuation with an assumption that is driven by the model

APPENDIX: Local vol & future smile flattening XII

smile flattening. Unless we have economic or financial reasons to believe that the smile will flatten, we should use a model that keeps the smile more stationary, namely a stochastic volatility model, in order to price the ratchet more realistically.

A further element confirming that the volatility is stochastic comes from the market. The market trades contracts referencing volatility of stock indices, called variance swaps and volatility swaps. If volatility were not stochastic, these products would not exist. See for example the VIX index, where one is allowed to trade a view of volatility for an option in the next 30 days. The fact that the VIX is a tradable asset with a randomness of its own, which is largely independent of the relative stock index randomness, points further to adoption of stochastic volatility models.

Risk Measures: A historical perspective I

This historical perspective is from Brian McHugh's review (2011)

This is an introduction into 'Risk Measures', particularly focusing on Value-at-Risk (VaR) and Expected Shortfall (ES) measures. A brief history of risk measures is given, along with a discussion of key contributions from various authors and practitioners.

What do we mean by "Risk"? I

Risk is defined by the dictionary as 'a situation involving exposure to danger'. It is related to the randomness of uncertainty. Risk is also described as 'the possibilty of financial loss' and this is the definition that will be discussed here.

Risk management, described by Kloman[1] as 'a discipline for living with the possibility that future events may cause adverse effects', is of vital importance to the appropriate day to day running of financial institutions.

Here, downside risk (the probability of loss or less than expected returns) will be the focus of discussion as it is the most crucial area for risk managers. In particular, Value at Risk (VaR) and Expected Shortfall (ES) methodologies of measuring risk will be analysed.

What do we mean by ”Risk”? II

The question that comes to mind is where does this risk come from, and of course there is no single answer.

Risk can be created by a great number of sources, both directly and indirectly, it propogates from government policies, war, inflation, technological innovations, natural phenomena, and many others.

There are a number of risks faced by financial institutions everyday, these include market risk, credit risk, operational risk, liquidity risk, and model risk.

What do we mean by ”Risk”? III

- Market risk includes the unexpected moves in the underlying of the financial assets (stock prices, interest rates, fx rates...)
- Credit risk propogates from the creditworthiness of a counterparty in a contract and the possibility of losses caused by its default.
- Operational risk: possibility of losses occured by internal processes, people, and systems or from other sources externally.
- Liquidity risk stems from the inability, in some cases, to buy or sell financial instruments in sufficient time as to minimise losses.
- Model risk: inaccurate use of valuation and pricing models, for instance inaccurate distributions or unrealistic assumptions. Negative interest rates? (eg Vasicek, Hull White), Models with thin tails instead of fat tails? Bad future volatility structures? Unrealistic correlation patterns? (see discussion on LMM above).

What do we mean by "Risk"? IV

- Finally, all such risks may interact in complex ways and their mutual dependence and contagion is a key aspect of modern research. As these risks are not really completely separable, this classification is purely indicative and not substantial.

[1] H. F. KLOMAN (1990), *Risk Management Agonistes*, Risk Analysis 10:201-205.

A brief history of VaR and Expected Shortfall I

The origins of VaR and risk measures can be traced back as far 1922 to capital requirements the New York Stock Exchange imposed on member firms according to Holton[2].

However, Markowitz's seminal paper 'Porfolio Theory' (1952), which developed a means of selecting portfolios based on an optimization of return given a certain level of risk, was the first convincing if stylized and simplistic method of measuring risk. His idea was to focus portfolio choices around this measurement.

A brief history of VaR and Expected Shortfall II

Risk management methodologies really took off from this point and over the next couple of decades new ideas, such as the Sharpe Ratio, the Capital Asset Pricing Model (CAPM) and Arbitrage Pricing Theory (APT), were being proposed and implemented.

Along with this came the introduction of the Black-Scholes option-pricing model in 1973, which lead to a great expansion of the options market, and by the early 1980s a market for over-the-counter (OTC) contracts had formed.

The related theory had important precursors in Bachelier (1900) and de Finetti (1931)[3]

A brief history of VaR and Expected Shortfall III

Perhaps the greatest consequence of the financial innovations of the 1970s and 1980s was the proliferation of leverage, and with these new financial instruments, opportunities for leverage abounded.

Think of a forward contract with payoff $S_T - K$ that is at the money forward, $K = e^{rT} S_0$. Its price at time 0 is $V_0 = S_0 - e^{-rT} K = S_0 - S_0 = 0$. So it costs nothing to enter this forward contract even on a huge notional, and yet this may lead to very large profit or losses in the future if the stock moves a lot.

Similarly for interest rate swaps, credit default swaps, oil swaps, and a number of other derivatives.

A brief history of VaR and Expected Shortfall IV

Along with academic innovation came technological advances. Information technology companies like Reuters, Telerate, and Bloomberg started compiling databases of historical prices that could be used in valuation techniques.

Financial instruments could be valued quicker with new hi-tech methods such as the Monte Carlo pricing for complex derivatives, and thus trades were being made quicker.

We have now reached super-human speed with high frequency trading, so debated that the EU is considering banning it.

However, in addition to all these innovations and advances came catastrophes in the financial world such as:

A brief history of VaR and Expected Shortfall V

- The Barings Bank collapse of 1995, which was solely down to the fraudulent dealings of one of its traders.
- Metallgesellschaft lost $1.3 billion by entering into long term oil contracts in 1993.
- Long-Term Capital Management's near collapse in 1998 and subsequent bailout overseen by the Federal Reserve. *Somewhat ironically, members of LTCM's board of directors included Scholes and Merton.*

A brief history of VaR and Expected Shortfall VI

For more information on these financial disasters and others see Jorion (2007).[4] Organizations were now more than ever increasingly in need for a single risk measure that could be applied consistently across asset categories in hope that financial disasters such as these could be prevented. However, even this wouldn't be enough, as the Lehman collapse of 2008 has shown. We'll discuss why later.

Q: "What is Basel?"

A: "A city in Europe? Perhaps switzerland?"

The Basel Committee on Banking Supervision was central to the introduction and implementation of VaR on a worldwide scale. The Committee itself does not possess any overall supervising authority, but rather gives standards, guidelines, and recommendations for individual national authorities to undertake.

A brief history of VaR and Expected Shortfall VII

The first Basel Accord of 1988 on Banking Supervision attempted to set an international minimum capital standard, however, according to McNeil at al.[5] this accord took an approach which was fairly coarse and measured risk in an insufficiently differentiated way.

A brief history of VaR and Expected Shortfall VIII

The G-30 (consultative group on international economic and monetary affairs) report in 1993 titled 'Derivatives: Practices and Principles' addressed the growing problem of risk management in great detail.

It was created with help from J.P. Morgans' RiskMetrics system, which measured the firm's risk daily.

The report gave recommendations that portfolios be marked-to-market daily and that risk be assessed with both VaR and stress testing.

While the G-30 Report focused on derivatives, most of its recommendations were applicable to the risks associated with other traded instruments.

A brief history of VaR and Expected Shortfall IX

For this reason, the report largely came to define the new risk management of the 1990's and set an industry-wide standard.

The report is also interesting, as it may be the first published document to use the word "value-at-risk".

Expected shortfall (ES) is a seemingly more recent risk measure, however, Rappoport (1993) [6] mentions a new approach called Average Shortfall in J.P. Morgan's Fixed Income Research Technical Document, which first noted application of the theory of Expected Shortfall in finance.

The later paper of Artzner et al. (1999)[7] introduces four properties for measures of risk and calls the measures satisfying these properties as 'coherent'.

A brief history of VaR and Expected Shortfall X

While such ”coherent” risk measures become ill defined in presence of liquidity risk (especially the proportionality assumption), this was the catalyst for the need of a new ’coherent’ risk measure.

As ES was practically the only operationally manageable coherent risk measure, ES was proposed as a coherent alternative to VaR.

[2] G. A. HOLTON (2002), working paper. *History of Value-at-Risk: 1922-1998*.

[3] Pressacco, F., and Ziani, L. (2010). Bruno de Finetti forerunner of modern finance. In: “Convegno di studi su Economia e Incertezza, Trieste, 23 ottobre 2009”, Trieste, EUT Edizioni Università di Trieste, 2010, pp. 65-84.

[4] P. JORION *Value at Risk: The New Benchmark for Managing Financial Risk 3rd ed.* McGraw-Hill.

[5] A. MCNEIL, R. FREY AND P. EMBRECHTS (2005), *Quantitative Risk Management*, Princeton University Press.

[6] P. RAPPOPORT (1993), *A New Approach: Average Shortfall*, J.P. Morgan Fixed Income Research Technical Document.

[7] P. ARTZNER, F. DELBAEN, J. EBER AND D. HEATH , *Coherent Measures of Risk*, Mathematical Finance Vol.9 No.3.

Solve SDE from Mock Ex 1: $dX = m\,dt + \sigma X dW$ I

We solve $dX_t = m\,dt + \sigma X_t dW_t,\ X_0$. Define the similar SDE given by a Geometric Brownian motion without drift, $d\bar{X}_t = \sigma \bar{X}_t dW_t,\ \bar{X}_0 = 1$. $\bar{X}$ is like X but without drift, so we hope in a simplification when considering $X/\bar{X}$. Define thus $Z_t = 1/\bar{X}_t$ as

$$Z_t = \exp\left(-\sigma W_t + \frac{1}{2}\sigma^2 t\right).$$

Define

$$Y_t := X_t Z_t = X_t/\bar{X}_t.$$

Apply the stochastic Leibnitz rule to Y:

$$dY_t = X_t\,dZ_t + Z_t\,dX_t + dX_t\,dZ_t.$$

Solve SDE from Mock Ex 1: $dX = m\,dt + \sigma XdW$ II

Calculate the terms, using Ito's formula for the exponential and the quadratic covariation

$$dZ_t = Z_t\left(-\sigma\,dW_t + \frac{1}{2}\sigma^2\,dt\right) + \frac{1}{2}Z_t\sigma^2\,dt = Z_t\left(-\sigma\,dW_t + \sigma^2\,dt\right),$$

$$dX_t = m\,dt + \sigma X_t\,dW_t,$$

$$dX_t\,dZ_t = \sigma X_t\,dW_t\,(-\sigma Z_t\,dW_t) = -\sigma^2 X_t Z_t\,dt.$$

Substitute into the equation for dY_t:

$$\begin{aligned} dY_t &= X_t Z_t\left(-\sigma\,dW_t + \sigma^2\,dt\right) + Z_t\,(m\,dt + \sigma X_t\,dW_t) - \sigma^2 X_t Z_t\,dt \\ &= -\sigma X_t Z_t\,dW_t + \sigma^2 X_t Z_t\,dt + mZ_t\,dt + \sigma X_t Z_t\,dW_t - \sigma^2 X_t Z_t\,dt. \end{aligned}$$

Simplifying

$$dY_t = mZ_t dt.$$

Solve SDE from Mock Ex 1: $dX = m\,dt + \sigma X dW$ III

We then have by direct integration

$$Y_t = Y_0 + m\int_0^t Z_s ds$$

Recall now that $X_t = Y_t/Z_t$, so

$$X_t = \frac{Y_0}{Z_t} + \frac{m}{Z_t}\int_0^t Z_s ds$$

and recalling the expression for Z_t we have

$$\boxed{X_t = \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)\left(Y_0 + m\int_0^t \exp\left(-\sigma W_s + \frac{1}{2}\sigma^2 s\right) ds\right)}.$$

Solve SDE from Mock Ex 1: $dX = m\,dt + \sigma X dW$ IV

To check that this is the correct solution, differentiate

$$dX_t = \left(d\ \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)\right)\left(Y_0 + m\int_0^t \exp\left(-\sigma W_s + \frac{1}{2}\sigma^2 s\right) ds\right)$$

$$+ \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right) d\left(Y_0 + m\int_0^t \exp\left(-\sigma W_s + \frac{1}{2}\sigma^2 s\right) ds\right) =$$

$$= \left[\exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)(\sigma dW_t - \frac{1}{2}\sigma^2 dt) + \frac{1}{2}\exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)\sigma^2 dt\right]\cdot$$

$$\cdot\left(Y_0 + m\int_0^t \exp\left(-\sigma W_s + \frac{1}{2}\sigma^2 s\right) ds\right) +$$

$$\exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right) m \exp\left(-\sigma W_t + \frac{1}{2}\sigma^2 t\right) dt$$

Solve SDE from Mock Ex 1: $dX = m\,dt + \sigma X dW$ V

$$= \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)\sigma dW_t \left(Y_0 + m\int_0^t \exp\left(-\sigma W_s + \frac{1}{2}\sigma^2 s\right) ds\right) +$$
$$+mdt$$

and recalling our solution for X above we have indeed

$$dX_t = mdt + \sigma X_t dW_t.$$

Note that in checking the solution we didn't have a covariation term between

$$\exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)$$

and

$$\left(Y_0 + m\int_0^t \exp\left(-\sigma W_s + \frac{1}{2}\sigma^2 s\right) ds\right)$$

Solve SDE from Mock Ex 1: $dX = m\,dt + \sigma XdW$ VI

because, being the second term an integral, it has finite variation and as such it is of "dt" type, and it does not contribute to any covariation term.

Options on stock with continous dividends I

The standard no-arbitrage replication and risk neutral option pricing theory we developed assumes that the stock pays no dividend. However, in real life stocks have dividends. Here we show how the theory can be adjusted for a continuous dividend yield q. The bank account remains $dB_t = rB_t dt$, a risk free asset, because it is not impacted by whatever the risky asset does. In standard Black Scholes, the stock price or risky asset dynamics is $dS_t = \mu S_t dt + \sigma S_t dW_t^P$. Now, however, we assume a dividend yield is taken from the stock: in every small time interval $[t, t + dt)$, the stock pays out a dividend $qS_t dt$ to shareholders. This means that the equation for the stock under the measure P has to be changed to

$$dS_t = (\mu - q)S_t dt + \sigma S_t dW_t^P,$$

Options on stock with continous dividends II

where we are taking into account the dividend payment drain on the stock.

Side note: in the market, real dividends are discrete and paid at given points in time, not a continuous flow. However, we can diffuse a discrete dividend in a kind of equivalent dividend yield, at least in expected value. We use continue dividend yields because it's easier to adjust the theory. For discrete dividend yields the theory can be adjusted too, but the adjustments are much more cumbersome and often numerical methods are needed.

Let's see how the portfolio replicatioin derivation of the Black Scholes PDE changes with continous dividends.

Options on stock with continous dividends III

The most crucial change comes from adjusting the definition of self financing trading strategy. In the standard model, a strategy portfolio with value $V_t(\phi) = \phi_B(t)B_t + \phi_S(t)S_t$ is self-financing if its change in value dV_t comes only from gains or losses coming from movements of the underlying assets B and S: $dV = \phi_S dS + \phi_B dB$.
With a continuous dividend, the ϕ_S shares of stock held in the portfolio generate a continuous cash flow of $\phi_S(t)(qS_t)dt$. For the portfolio to remain self-financing (i.e., no external funds are added or removed), this dividend cash flow must be reinvested into the portfolio, and the self-financing condition becomes:

$$dV = \underbrace{\phi_S dS + \phi_B dB}_{\text{Trading Gains}} + \underbrace{\phi_S (qS) dt}_{\text{Reinvested Dividends}} . \tag{48}$$

Options on stock with continous dividends IV

This new term, $\phi_S qSdt$, accounts for the dividend income being added back into the portfolio.
Replication is still possible, but the underlying dynamics and the resulting PDE are modified.
The definition of the hedge ratio (Delta) for the stock is still the partial derivatives of $V(t, S_t)$ wrt S. This is because Ito's formula has not changed,

$$dV(t,S) = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}dSdS$$

and matching the dS term of this last equation with the dS term of (48), still gives $\phi_S = \frac{\partial V}{\partial S}$.
Thus $\Delta_t = \phi_S(t) = \frac{\partial V(t,S_t)}{\partial S}$
However, the value of V (and thus $\frac{\partial V}{\partial S}$) will be different because V must satisfy a new PDE.

Options on stock with continous dividends V

The replication argument leads to a modified Black-Scholes PDE. We equate the *dt* term in the change in the option's value (from Ito's formula) with the *dt* term in the change in value with the self-financing condition (48), as in the classic derivation, and obtain:
Option Value (Itô's formula):

$$dV = \left(\frac{\partial V}{\partial t} + (\mu - q)S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \left(\sigma S \frac{\partial V}{\partial S}\right) dW^P,$$

Strategy Value (Self-Financing condition):

$$dV = \phi_S dS + \phi_B dB + \phi_S qSdt$$

$$dV = \phi_S[(\mu - q)Sdt + \sigma S dW^P] + \phi_B(rBdt) + \phi_S qSdt$$

$$dV = [\phi_S(\mu - q)S + \phi_B rB + \phi_S qS]dt + [\phi_S \sigma S]dW$$

Options on stock with continous dividends VI

$$dV = [\phi_S \mu S + r(V - \phi_S S)]dt + [\phi_S \sigma S]dW$$

$$dV = [\phi_S(\mu - r)S + rV]dt + [\phi_S \sigma S]dW$$

$$dV = [\frac{\partial V}{\partial S}(\mu - r)S + rV]dt + [\frac{\partial V}{\partial S}\sigma S]dW.$$

We now match the *dt* terms of this last equation and of the Ito's formula equation. This provides the PDE

$$\frac{\partial V}{\partial t} + (\mu - q)S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = \frac{\partial V}{\partial S}(\mu - r)S + rV$$

with the desired terminal condition (for a call $V(T, S) = (S - K)^+$, etc). Substitute $\phi = \frac{\partial V}{\partial S}$:

$$\frac{\partial V}{\partial t} + (\mu - q)S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = \frac{\partial V}{\partial S}(\mu - r)S + rV$$

Options on stock with continous dividends VII

The $\mu S \frac{\partial V}{\partial S}$ terms cancel, leaving the modified Black-Scholes PDE with dividends:

$$\boxed{\frac{\partial V}{\partial t} + (r-q)S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV.}$$

The Risk-Neutral Measure. The numeraire of Q does not change, being still the risk free asset that is not linked to what we do with the risky assets. However, the dividend does change the dynamics of the stock S under this measure Q.

Standard Black Scholes (BS) no dividend: Under Q, the stock must grow at the risk-free rate r, $dS = rSdt + \sigma S dW^Q$.

Options on stock with continous dividends VIII

BS with dividend q: Under Q, the total return of the stock (gains return drift + dividend yield return) must equal the risk-free rate r.

$$\text{ReturnDrift}_Q + \text{ReturnDividendYield} = r$$

$$\text{ReturnDrift}_Q + q = r$$

$$\text{ReturnDrift}_Q = r - q$$

Therefore, the risk-neutral SDE for the stock price becomes:

$$dS = (r - q)Sdt + \sigma SdW^Q. \tag{49}$$

This $(r - q)$ drift is exactly the term that appears in the modified Black-Scholes PDE, confirming the consistency of the framework.
The price of a call option can be derived by solving the PDE or moving to a risk neutral expectation, $E_t^Q[e^{-r(T-t)}(S_T - K)^+]$ where dS follows

Options on stock with continous dividends IX

(49). Alternatively, as we did in the non-dividend case, the expected value version of the call price can be derived from the PDE by applying the Feynman-Kac theorem to the new PDE.
Following a derivation analogous to the one without dividends, based on integration, the formula for a call is

$$V_{Call}^{BS}(t, S_t) = S_t e^{-q(T-t)} \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2),$$

$$d_{1,2} = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r - q \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

We can also investigate the price of the forward contract with payoff $S_T - K$.

Options on stock with continous dividends X

$V^{FWD}(t, T, S_t, K, r)$ is the present value of the expected payoff under the risk-neutral measure. From the general expectation of a Geometric Brownian motion, we know from (49) that

$$E_t^Q[S_T] = e^{(r-q)(T-t)} S_t.$$

In reality this holds for any SDE with drift $(r - q)Sdt$ and is model independent, it doesn't require Black Scholes (it will hold also for smile models). Anyway, using this, we can compute the forward price as

$$V^{FWD}(t, T, S_t, K, r, q) = e^{-r(T-t)} E_t^Q[S_T - K] = e^{-r(T-t)} (E_t^Q[S_T] - K) =$$

$$= e^{-r(T-t)} e^{(r-q)(T-t)} S_t - e^{-r(T-t)} K = e^{-q(T-t)} S_t - e^{-r(T-t)} K.$$

Options on stock with continous dividends XI

Now, with the value of a call and a forward, one can compute the value of a put with dividend by difference, due to put-call parity.

$$V_{Put} = V_{Call} - V_{forward}$$

So

$$V^{BS}_{Put}(t, S_t) = S_t e^{-q(T-t)}\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) - e^{-q(T-t)}S_t + e^{-r(T-t)}K$$

$$= Ke^{-r(T-t)}\Phi(-d_2) - S_t e^{-q(T-t)}\Phi(-d_1)$$

where $d_{1,2}$ are the updated versions with q given above for the call with dividends.

Options on stock with continous dividends XII

The forward price $F_{t,T}$ is, by definition, the specific strike K that makes the contract value zero at time t. Setting $V_t^{FWD} = 0$ and solving for K gives

$$F_{t,T} = S_t e^{(r-q)(T-t)}.$$

Finally, another key insight is that while the bank account B is still the numeraire for the standard risk-neutral measure Q, the discounted stock price, $\frac{S_t}{B_t}$, is no longer the correct quantity that must be a martingale under Q. This is because the stipulation
risky asset/ B
equals martingale only holds for non-dividend paying risky assets.

The quantity that must be a martingale under Q is the dividend-adjusted discounted stock price.

Options on stock with continous dividends XIII

We need to consider the *discounted dividend-including value process*:

$$M_t := \frac{S_t}{B_t} + \int_0^t \frac{qS_u}{B_u}\, du,$$

namely the process where we add back the paid dividend at every point in time over the bank account numeraire at that time, resulting in an integral. Equivalently,

$$M_t = e^{-rt} S_t + \int_0^t e^{-ru} q S_u\, du.$$

Then, differentiating

$$dM_t = -re^{-rt} S_t dt + e^{-rt} dS_t + e^{-rt} q S_t\, dt.$$

Options on stock with continous dividends XIV

Substituting from the dS equation with dividends under Q, (49), we have

$$dM_t = -re^{-rt}S_t dt + e^{-rt}((r-q)S_t dt + \sigma S_t dW_t^Q) + e^{-rt} q S_t\, dt =$$

$$= 0dt + e^{-rt}\sigma S_t dW_t^Q.$$

Given that S_t is well defined and its solution exists unique, and it satisfies suitable integrability conditions, this is a regular stochastic differential with zero drift and as such it is a martingale.
This shows how the numeraire condition is changed with dividends.