Stochastic Differential Equations in Financial Modelling

Arnav Singh

April 21, 2024

Part I
Probability and SDEs

Definition 1.1 (Σ - sample space)**.** It is the set of all elementary outcomes of a random experiment

Definition 1.2 ($\mathcal{F}$ - sigma-field)**.** Also denoted by σ-field, it is a collection of subsets of Σ that satisfies the following properties:

1. $\emptyset \in \mathcal{F}$ and $\mathcal{F}$
2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
3. If $A_1, A_2, \ldots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$, i.e. $\mathcal{F}$ is closed under countable unions

Trivial σ-filed given by $\mathcal{F} = \{\Omega, \varnothing\}$
Power set of Σ is the largest σ-field, denoted by $\mathcal{P}(\Sigma)$ containing all possible subsets of Σ

Definition 1.3 (Borel set)**.** The smallest σ-field that contains all open sets in $\mathbb{R}$ is called the Borel σ-field, denoted by $\mathcal{B}(\mathbb{R})$

Definition 1.4 (Borel σ-field)**.** σ-Field generated by open sets in $\mathbb{R}$ is called the Borel σ-field, denoted by $\mathcal{B}(\mathbb{R})$, containing all open, closed, half-open, half-closed, and countable unions of these sets. It is the smallest σ-field that contains all open sets in $\mathbb{R}$

$$\mathcal{B}(\mathbb{R}) = \sigma\left(\{(a, b) \mid a, b \in \mathbb{R}\}\right) \tag{1}$$

Definition 1.5 (Probability measure)**.** A function $P : \mathcal{F} \to [0, 1]$ is called a probability measure if it satisfies the following properties:

1. $P(\Omega) = 1,\ P(\varnothing) = 0$
2. $P(A) \geq 0$ for all $A \in \mathcal{F}$
3. If $A_1, A_2, \ldots \in \mathcal{F}$ are pairwise disjoint, then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

Definition 1.6 (Random variable)**.** A random variable is a function $X : \Omega \to \mathbb{R}$ that maps the sample space to the real numbers.
It is measurable if for all $B \in \mathcal{B}(\mathbb{R})$, the set $X^{-1}(\mathcal{B}) = \{\omega \in \Omega \mid X(\omega) \in B\} \in \mathcal{F}$

Definition 1.7 (Cumuative distribution function)**.** The cumulative distribution function (CDF) of a random variable X is defined as $F_X(x) = P(X \leq x)$ for all $x \in \mathbb{R}$

$$P(X^{-1}((-\infty, x])) = P(\{\omega \in \Omega \mid X(\omega) \leq x\}) = F_X(x)$$

Definition 1.8 (Probability density function)**.** If there exists a function $f_X : \mathbb{R} \to \mathbb{R}$ such that for all $x \in \mathbb{R}$, $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$, then f_X is called the probability density function (pdf) of X

Or equivalently if the CDF differentiable the pdf is: $f_X(x) = \frac{d}{dx}F_X(x)$

Definition 1.9 (Expectation)**.** The expectation of a random variable X is defined as:

$$E[X] = \int_{\Omega} X(\omega)dP(\omega) \tag{2}$$
$$= \int_{\mathbb{R}} ydF_x(y) \tag{3}$$
$$= \int_{\mathbb{R}} yf_X(y)dy \tag{4}$$

Definition 1.10 (Variance)**.** The variance of a random variable X is defined as:

$$\begin{aligned} Var(X) &= E[(X - E[X])^2] \\ &= E[X^2] - E[X]^2 \end{aligned}$$

Proposition 1.11 (Properties of expectation)**.** *Let X, Y be random variables and $a, b \in \mathbb{R}$, then:*

1. $E[aX + bY] = aE[X] + bE[Y]$
2. *If $X \geq 0$, then $E[X] \geq 0$*
3. *If $X \geq Y$, then $E[X] \geq E[Y]$*

Proposition 1.12 (Properties of variance)**.** *Let X, Y be random variables and $a, b \in \mathbb{R}$, then:*

1. $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$
2. $Var(X) \geq 0$
3. *If $X \geq 0$, then $Var(X) \geq 0$*
4. *If $X \geq Y$, then $Var(X) \geq Var(Y)$*

Definition 1.13 (Moment generating function)**.** The moment generating function (mgf) of a random variable X is defined as:

$$M_X(t) = E[e^{tX}] \tag{5}$$

We have that

$$\frac{d^n}{dt^n}M_X(t) = E[X^n e^{tX}]$$

Definition 1.14 (Characteristic function)**.** Not all random variables have a MGF, we instead have the characteristic function (cf) of a random variable X is defined as:

$$\phi_X(t) = E[e^{itX}] \tag{6}$$

Definition 1.15 (Independence)**.** Two random variables X, Y are independent if for all $A, B \in \mathcal{B}(\mathbb{R})$, we have that:

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

and denoted by $X \perp Y$ We also have that

$$E[XY] = E[X]E[Y]$$
$$Var(X + Y) = Var(X) + Var(Y)$$

Definition 1.16 (Covariance)**.** The covariance of two random variables X, Y is defined as:

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y] \tag{7}$$

Definition 1.17 (Correlation)**.** The correlation of two random variables X, Y is defined as:

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}$$

We have that $-1 \leq Corr(X, Y) \leq 1$. For random variables X_1, X_2 with $X_2 = aX_1 + b$, we have that $Corr(X_1, X_2) = 1$ or -1 depending on the sign of a. Can define a correlation or covariance matrix for a vector of random variables $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ as the matrix of all pairwise correlations or covariances

Theorem 1.18 (Central Limit Theorem)**.** *Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with mean μ and variance σ^2. We have sample mean defined as $\overline{X}_n$*

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$
$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{n\uparrow\infty} N(0, 1)$$

Example 1.19. Take $X \sim N(\mu, \sigma^2)$, then we have that:

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

where $n!!$ is the double factorial defined as $n!! = n(n-2)(n-4)\ldots$

Definition 1.20 (Multivariate random variable)**.** A multivariate random variable is a vector of random variables $\mathbf{X} = (X_1, X_2, \ldots, X_n)$. The joint distribution of $\mathbf{X}$ is defined as:

$$F_{\mathbf{X}}(x_1, x_2, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)$$

With the PDF defined as:

$$f_{\mathbf{X}}(x_1, x_2, \ldots, x_n) = \frac{\partial^n}{\partial x_1 \partial x_2 \ldots \partial x_n} F_{\mathbf{X}}(x_1, x_2, \ldots, x_n)$$

Definition 1.21 (Multivariate Normal)**.** We define for dimension n. Let $\mu = [\mu_1, \mu_2, \cdots, \mu_n]$ 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 $X = [X_1, \cdots, X_n] \sim N(\mu, V) \sim N((\mu_i)i = 1, \ldots, n, (\sigma_i \sigma_j \rho i, j)_{i,j=1\ldots n})$ if

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

Definition 1.22 (Convergence of Random Variables)**.** Let $X, X_1, X_2, \ldots$ be random variables, then we have the following types of convergence:

1. Almost sure convergence: $X_n \xrightarrow{a.s.} X$ if $P(\lim_{n\to\infty} X_n = X) = 1$
2. L^p convergence: $X_n \xrightarrow{L^p} X$ if $E[|X_n - X|^p] \to 0$ as $n \to \infty$

   Making sure that $E[|X_n|^p] < \infty$ for all $n \in \mathbb{N}$
3. Convergence in mean square: $X_n \xrightarrow{m.s.} X$ if $E[(X_n - X)^2] \to 0$ as $n \to \infty$

   Special case of L^p convergence for $p = 2$
4. Convergence in probability: $X_n \xrightarrow{p} X$ if for all $\epsilon > 0$, $P(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$
5. Weak convergence or convergence in distribution: $X_n \xrightarrow{d} X$ if $F_{X_n}(x) \to F_X(x)$ for all $x \in \mathbb{R}$ where F_{X_n} and F_X are the CDFs of X_n and X respectively

Definition 1.23 (Stochastic Process)**.** A stochastic process is a collection of random variables indexed by time, i.e. $\{X_t\}_{t\in T}$ where T is the index set. We have the following types of stochastic processes:

1. Discrete-time: If $T = \{0, 1, 2, \ldots\}$
2. Continuous-time: If $T = [0, \infty)$
3. Markov process: If the future of the process depends only on the present state
4. Martingale: If the expected value of the process at time t given all information up to time s is equal to the value at time s

Definition 1.24 (Filtration)**.** A filtration is a collection of σ-fields $\{\mathcal{F}_t\}_{t\in T}$ such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s \leq t$. It represents the information available at time t

Definition 1.25 (Brownian Motion)**.** A Brownian motion is a stochastic process $\{W_t\}_{t\geq 0}$ with the following properties:

1. $W_0 = 0$
2. **Continuous sample paths:** $t \mapsto W_t(\omega)$
3. **Independent Increments:** For all $0 \leq s < t < u$, $W_t - W_s$ and $W_u - W_t$
4. **Stationary Increments:** Distribution of $W_{t+h} - W_t$ does not depend on t, but only on h for $h > 0$
5. W **is Gaussian:** $W_t \sim N(0,t)$ and $W_t - W_s \sim N(0, t-s)$ for all $0 \leq s < t$ under the probability measure P

Definition 1.26 (Ordinary Differential Equation)**.** An ordinary differential equation (ODE) is an equation involving a function of one variable and its derivatives. It is of the form:

$$\frac{dX(t)}{dt} = f(X(t)), \quad X(0) = x_0$$

where y is the unknown function of t, and f is a given function of t and y
Can rewrite it as follows

$$dX(t) = f(X(t))dt$$

Proposition 1.27 (Solution to Affine ODE)**.** *The solution to an ODE of the following form for A, B functions of time:*

$$\frac{dX(t)}{dt} = B(t) - A(t)X(t)$$
$$X(t) = \exp\left(-\int_0^t A(s)ds\right)\left(X(0) + \int_0^t B(u)\exp\left(\int_0^u A(s)ds\right)du\right)$$

Definition 1.28 (Stochastic Differential Equation)**.** A stochastic differential equation (SDE) is a differential equation in which one or more of the terms is a stochastic process. It is of the form:

$$dX(t) = \underbrace{f(X_t)dt}_{\text{Local drift}} + \underbrace{\sigma(X_t)}_{\text{Local }\sigma} \cdot \underbrace{dW_t}_{\text{Brownian motion}}$$

where $X(t)$ is the unknown function of t, f is the local drift, σ is the local volatility, and W_t is the Brownian motion process with $dW_t \sim N(0, dt)$
Problems:

1. Unbounded variation: Brownian motion has unbounded variation
2. Nowhere differentiable: Brownian motion is nowhere differentiable

So the integral of dW_t is not well defined

Proposition 1.29 (Fixing the SDE)**.** *Can't use the standard Riemann-Stieltjes integral to solve the SDE as the Brownian motion is not of bounded variation.*

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

In a Stiltjes integral one has

$$\int_0^T \sigma(X_s)dW_s = \lim_{n\to\infty}\sum_{i=1}^n \sigma(X(t_i))(W_{t_{i+1}} - W_{t_i})$$

for ANY choice $t_i \in [t_i, t_{i+1})$. We must carefully choose the partition t_i to make the integral well defined.
We have 2 choices:

1. ***Ito's Integral**:*

$$\int_0^T \sigma(X_s)dW_s = \lim_{n\to\infty}\sum_{i=1}^n \sigma(X(t_{i+1}))(W_{t_i} - W_{t_i})$$

where $t_i = i\frac{T}{n}$ and $W_{t_{i+1}} - W_{t_i} \sim N(0, \frac{T}{n})$

2. ***Stratonovich Integral**:*

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

Stratonovich Integral looks into the future to calculate the integral, while Ito's Integral only uses current and past information.
If $\sigma(X_t)$ does not depend on X_t, then the two integrals are the same, we call this a Wiener Integral.

We have the following property for Ito's Integral:

$$E\left[\int_0^T \sigma(X_s)dW_s\right] = 0$$

Proposition 1.30 (Ito's Isometry)**.** *We have that:*

$$E\left[\left(\int_0^t \sigma(X_s)dW_s\right)^2\right] = E\left[\int_0^t \sigma(X_s)^2 ds\right]$$

Definition 1.31 (Adapted to Filtration)**.** A stochastic process $\{X_t\}_{t\in T}$ is adapted to a filtration $\{\mathcal{F}_t\}_{t\in T}$ if for all $t \in T$, X_t is $\mathcal{F}_t$-measurable, i.e. X_t is known at time t

Theorem 1.32 (Existence and Uniqueness of Solutions)**.** *Consider Ito SDE of the form:*

$$dX(t) = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \;\; X_0 = Z$$

where μ, σ are Globally Lipschitz continuous and follow linear growth in X and $\sigma \neq 0$. Z a random variable independent $\sigma(\{W_t : t < T\})$ and $E[Z^2] < \infty$. Then there exists a unique global solution to the SDE on the interval $[0, T]$ that is adapted to the filtration generated by the Brownian motion, $\mathcal{F}_t^W$ and is square integrable.

- ***Lipschitz continuity:*** *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz continuous if there exists a constant $L > 0$ such that for all $x, y \in \mathbb{R}^n$, we have that:*

$$||f(x) - f(y)|| \leq L||x - y||$$

- ***Linear growth:*** *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ has linear growth if there exists a constant $K > 0$ such that for all $x \in \mathbb{R}^n$, we have that:*

$$||f(x)|| \leq K(1 + ||x||)$$

Definition 1.33 (Ito's formula)**.** Given

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t$$

we have that for a smooth function $\phi(t, x)$, the Ito's formula is given by:

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

We have that:

- $dtdt = 0$
- $dW_t dW_t = dt$
- $dtdW_t = 0$

Note for V_t differentiable, we have that:

$$dV_t dV_t = V'(t)dtV'(t)dt = V'(t)^2 \underbrace{dtdt}_{=0} = 0$$

Definition 1.34 (Quadratic Variation)**.** The quadratic variation of a stochastic process $\{X_t\}_{t\in T}$ is defined as:

$$[X]_t = \lim_{n\to\infty}\sum_{i=1}^{n}(X_{t_{i+1}} - X_{t_i})^2$$

where the limit is taken in the mean square sense. We have that:

$$[W]_t = t$$

We have that for 2 independent Brownian motions $W, \widetilde{W}$, we have that:

$$[W, \widetilde{W}]_t = 0$$

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

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

Proposition 1.35 (Ito-Stratonovich Transformation)**.** *We have that:*

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t \to dX_t = \widetilde{f}(X_t)dt + \sigma(X_t)\circ dW_t$$

where $\circ$ *denotes the Stratonovich integral and* $\widetilde{f} = f - \frac{1}{2}\sigma\frac{\partial\sigma}{\partial x}$. *Here both integrals are equivalent. Note for* σ *deterministic, the two integrals are the same.*

Definition 1.36 (Geometric Brownian Motion)**.** A geometric Brownian motion is a stochastic process $\{S_t\}_{t\geq 0}$ with

$$dS_t = m_t S_t dt + \nu_t S_t dW_t$$

where m_t is the drift, ν_t is the volatility, and W_t is the Brownian motion process. We have that:

$$S_t = S_0 \exp\left(\underbrace{\int_0^t (m_s - \frac{1}{2}\nu_s^2)ds}_{M_t} + \underbrace{\int_0^t \nu_s dW_s}_{V_t^2}\right) = S_0 e^{\mathcal{N}(M_t, V_t^2)}$$

- Expectation: $E[S_t] = S_0 \exp(\int_0^t m_s ds)$
- Variance: $Var(S_t) = S_0^2 \exp(2\int_0^t m_s ds)\left(\exp(\int_0^t \nu_s^2 ds) - 1\right)$

Definition 1.37 (Arithmetic Brownian Motion)**.** An arithmetic Brownian motion is a stochastic process $\{S_t\}_{t\geq 0}$ with

$$dS_t = \mu_t dt + \sigma_t dW_t$$

where μ_t is the drift, σ_t is the volatility, and W_t is the Brownian motion process. We have that:

$$S_t = S + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$$

We have that:

$$S_t \sim N\left(S + \int_0^t \mu_s ds, \int_0^t \sigma_s^2 ds\right)$$

Definition 1.38 (Ornstein-Uhlenbeck Process)**.** An Ornstein-Uhlenbeck process is a stochastic process $\{X_t\}_{t\geq 0}$ with

$$dX_t = (b_t - a_t X_t)dt + \sigma_t dW_t$$

where b_t is the drift, a_t is the mean reversion, σ_t is the volatility, and W_t is the Brownian motion process. We have that:

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

Definition 1.39 (Vasicek Model)**.** A special case of the Ornstein-Uhlenbeck process is the Vasicek model, where we have that:

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

Where κ is the mean reversion rate, θ is the long-term mean, σ is the volatility, and W_t is the Brownian motion process. We have $b(t) = \kappa\theta, a(t) = \kappa$ and $\sigma_t = \sigma$. Then:

$$\begin{aligned} X(t) &= e^{-\kappa t}\left[\exp(\kappa u)\left(\kappa\theta du + \sigma dW_u\right) + X(0)\right] \\ &= x_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}) + \sigma \int_0^t e^{-\kappa(t-u)} dW_u \end{aligned}$$

We know X_t will be Gaussian with mean $x_0 e^{-kt} + \theta(1 - e^{-\kappa t})$ and variance $\frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})$

Definition 1.40 (CIR Model)**.** A special case of the Ornstein-Uhlenbeck process

$$\begin{aligned} \text{Vasicek:} &\quad dX_t = \kappa(\theta - X_t)dt + \sigma dW_t,\ x_0 \\ \text{CIR:} &\quad dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t,\ x_0 \end{aligned}$$

Where CIR is used to model interest rates $Y_t = r_t$ or volatility $Y_t = \nu_t$. Model can never be negative, but can be zero. We have the Feller condition for the CIR model to be positive:

$$2\kappa\mu \geq \nu^2$$

We have for both models that:

- κ : Mean reversion rate (speed of mean reversion rate)
- θ : Long-term mean reversion level
- σ : Volatility

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

$$Var(X_t) = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}), \qquad Var(Y_t) = y_0 \frac{\sigma^2}{\kappa}(e^{-\kappa t} - e^{-2\kappa t}) + \frac{\theta\sigma^2}{2\kappa}(1 - e^{-\kappa t})^2$$

$$\lim_{t\to\infty} Var(X_t) = \frac{\sigma^2}{2k} \qquad \lim_{t\to\infty} Var(Y_t) = \frac{\theta\sigma^2}{2\kappa}$$

Definition 1.41 (Product Rule for SDEs)**.** Given two stochastic processes X_t, Y_t with SDEs:

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t$$

Last term computed with usual rules.

Definition 1.42 (Equivalent Measures)**.** Say 2 measures $\mathbb{P}, \mathbb{Q}$ on $(\Sigma, \mathcal{F})$ are equivalent if they agree on which events have probability 0 or 1. We write $\mathbb{P} \sim \mathbb{Q}$

Definition 1.43 (Radon-Nikodym derivative)**.**

$$\mathbb{E}^{\mathbb{Q}}[X] = \int_\Omega X d\mathbb{Q} = \int_\Omega X \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} = \mathbb{E}^{\mathbb{P}}[X \frac{d\mathbb{Q}}{d\mathbb{P}}]$$

Definition 1.44 (Girsanov's Theorem)**.** Define for all $t \in [0, T]$

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

Then

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

Definition 1.45 (Poisson Process)**.** A Poisson process is a counting process $\{N_t\}_{t\geq 0}$ with the following properties:

1. $N_0 = 0$
2. All jumps are of size 1
3. **Right continuity:** $t \mapsto N_t(\omega)$ is right continuous
4. **Independent increments:** For all $0 \leq s < t$, $N_t - N_s$ is independent of $\{N_u\}_{u \leq s}$

5. **Stationary increments:** Distribution of $N_{t+h} - N_t$ does not depend on t, but only on h for $h > 0$

6. **Poisson distribution:** For all $0 \leq s < t$, $N_t - N_s \sim Pois(\lambda(t-s))$

Poisson Distribution: $X \sim P(\lambda)$

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

Part II
SDEs for Option Pricing

Definition 2.46 (Economy)**.** A probability space $\Omega, \mathcal{F}, (\mathcal{F}_t : 0 \leq t \leq T), P$
Assume $\mathcal{F} = \mathcal{F}_t$. We have 2 assets traded: a stock price S_t and a bond price B_t with the following:

- **Stock price:** $dS_t = \mu S_t dt + \sigma S_t dW_t$
  $\implies S_t = S_0 \exp\left((\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\right), \quad 0 \leq t \leq T$
- **Bond price:** $dB_t = rB_t dt$
  $\implies B_t = e^{rt}$

Assumptions:

- **No transaction costs:** No costs to buy or sell assets
- **No dividends:** Stock does not pay dividends
- **Shares are infinitely divisible:** Can buy any fraction of a share
- **Short selling allowed:** Can sell assets you do not own
- **No default risk**
- **No funding costs:** Cash can be borrowed or lent at the risk free rate r
- **Continuous time and continuous trading/hedging**
- **Perfect market info, complete markets**

Definition 2.47 (Contingent Claim)**.** A contingent claim Y for 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. We limit ourselves to simple contingent claims, i.e. claims of the form $Y = f(S_T)$, where f is a measurable function of the risky asset at maturity.

Definition 2.48 (Trading Strategy)**.** A trading strategy is a pair of processes (φ^B, φ^S) on $\Omega, \mathcal{F}, (\mathcal{F}_t : 0 \leq t \leq T), P$ that are locally bounded and predictable. Representing the number of shares of the bond and stock respectively. We have the value of the portfolio at time t given by:

$$V_t(\varphi) = \varphi_t^B B_t + \varphi_t^S S_t$$

Definition 2.49 (Gain process)**.**

$$G_t(\phi) = \int_0^t \phi_s^S dS_s + \int_0^t \phi_s^B dB_s$$

representing the income from the trading strategy ϕ up to time t

Definition 2.50 (Self-financing)**.** A strategy is self-financing if $V_t(\phi) \geq 0$ for all $t \in [0, T]$ and

$$V_t(\phi) = V_0(\phi) + G_t(\phi)$$

Or equivalently:

$$dV_t(\phi) = \phi_t^S dS_t + \phi_t^B dB_t = dG_t(\phi)$$

i.e only changes in value of portfolio come from changes in the value of the assets.

Definition 2.51 (Arbitrage)**.** An arbitrage is a self-financing trading strategy ϕ such that:

$$\phi_0^B B_0 + \phi_0^S S_0 = 0, \quad \mathbb{P}(V_t(\phi) > 0) > 0$$

i.e. a strategy that has no initial cost and has a positive probability of making a profit.

Definition 2.52 (Attainable contingent claims)**.** A contingent claim Y is attainable if there exists a self-financing trading strategy ϕ such that:

$$V_T(\phi) = Y$$

Say that ϕ generates $Y \wedge V_t(\phi)$ is the price at time t for Y.

Definition 2.53 (European Call)**.** A European call option is a contingent claim with payoff:

$$Y = (S_T - K)^+$$

where K is the strike price.

Definition 2.54 (European Put)**.** A European put option is a contingent claim with payoff:

$$Y = (K - S_T)^+$$

where K is the strike price.

Definition 2.55 (Risk-neutral measure)**.** The risk-neutral measure is a measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ such that the discounted price process $\{e^{-rt}S_t\}_{t\geq 0}$ is a martingale under $\mathbb{Q}$. We have that:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{e^{-rT}S_0}{\mathbb{E}[e^{-rT}S_T]}$$

Definition 2.56 (Black-Scholes-Merton Model)**.** Assume value of simple claim at $t = T$ a function of S_t

$$V_t = V(t, S_t) = \phi_t^S S_t + \phi_t^B B_t$$

and

$$dB_t = rB_t dt, \quad dS_t = \mu S_t dt + \sigma S_t dW_t$$

Assume $V \in C^{1,2}([0,T] \times \mathbb{R}^+)$ i.e 2x differentiable w.r.t S_t and 1x with t. Then by Ito's formula:

$$\begin{aligned} dV_t &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}dS_t^2 \\ dS_t &= \mu S_t dt + \sigma S_t dW_t \\ dtdt &= 0, \quad dW_t dW_t = dt, \quad dtdW_t = 0 \end{aligned}$$

We have that:

$$dV_t = \left(\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S_t \frac{\partial V}{\partial S} dW_t$$

Via the self-financing condition we get:

$$\text{For } 0 \leq t \leq T \quad \phi_t^S = \frac{\partial V}{\partial S}(t, S_t), \quad \phi_t^B = (V_t - \phi_t^S S_t)/B_t$$

Combining the below two equations we get:

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

We now get the Black-Scholes PDE, for terminal condition $V(T, S_T) = f(S_T) = (S_T - K)^+$:

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

We have the following solution:

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

where:

$$d_1(t) = \frac{\log\left(\frac{S_t}{K}\right) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \quad d_2(t) = d_1(t) - \sigma\sqrt{T-t}$$

Definition 2.57 (In-, At-, Out-of-the-money)**.** We have the following definitions for options:

- **In-the-money:** $S_t > K$
- **At-the-money:** $S_t = K$
- **Out-of-the-money:** $S_t < K$

Theorem 2.58 (Feynman-Kac Theorem)**.** *Given a PDE of the form:*

$$\frac{\partial V}{\partial t} + \mu(t,x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 V}{\partial x^2} - rV(t,x) = 0$$

with terminal condition $V(T,x) = f(x)$, then the solution is given by:

$$V(t,x) = e^{-r(T-t)} \cdot \mathbb{E}^Q\left[f(X_T) \mid X_t = x\right]$$

where X_t is the solution to the SDE:

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

Where diffusion process X_t has dynamics starting at $X_t = x$

$$dX_s = b(s, X_s)ds + \sigma(s, X_s)dW_s, s \geq t, X_t = x$$

Take $b(x) = rx, \sigma(x) = \sigma x$ for Black-Scholes model. We get:

$$\begin{aligned} V_{BS}(0, S_0, K, T, \sigma^2, r) &= e^{-rT}\mathbb{E}_0^Q[(S_T - K)^+] \\ &= S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2) \\ d_{1,2} &= \frac{\log\left(\frac{S_0}{K}\right) + (r \pm \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \end{aligned}$$

Proposition 2.59 (Computing Call option delta)**.** *Call option delta - the sensitivity of the option price to changes in the initial stock price:*

$$\phi^S(0) = \Delta_0 = \frac{\partial V_{BS}}{\partial S} = \Phi(d_1)$$

Proposition 2.60 (Risk Netural Measure via Girsanov's Theorem)**.** *Aim to move from*

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

to

$$dS_t = rS_t dt + \sigma S_t dW_t^{\mathbb{Q}}$$

We have that:

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

Call $\frac{\mu - r}{\sigma}$ the market price of risk or Sharpe ratio.

Definition 2.61 (Martingale Measure)**.** A martingale measure is a measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ such that the discounted price process $\{e^{-rt}S_t\}_{t\geq 0}$ is a martingale under $\mathbb{Q}$. We have that:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{e^{-rT}S_0}{\mathbb{E}[e^{-rT}S_T]}$$

which is equivalent to S having drift rate r under $\mathbb{Q}$

$$dS_t = S_t[rdt + \sigma dW_t^{\mathbb{Q}}],\ 0 \leq t \leq T$$

Theorem 2.62 (1st Fundamental Theorem of Asset Pricing)**.** *A market is arbitrage-free if and only if there exists a risk-neutral measure $\mathbb{Q}$ equivalent to the real-world measure $\mathbb{P}$. If there exists a risk-neutral measure $\mathbb{Q}$ equivalent to the real-world measure $\mathbb{P}$, then there exists a unique attainable claim price that can be computed as the expectation of the claim under the risk-neutral measure.*

Definition 2.63 (Complete market)**.** A market is complete if every contingent claim is attainable.

Theorem 2.64 (2nd Fundamental Theorem of Asset Pricing)**.** *A market is complete if and only if there exists a unique risk-neutral measure $\mathbb{Q}$ equivalent to the real-world measure $\mathbb{P}$.*

Definition 2.65 (Numeraire)**.** A numeraire is a risk-free asset whose price is always 1. We can use the numeraire to price other assets.Canonically the numeraire is the bond price B_t with dynamics:

$$dB_t = rB_t dt$$

Definition 2.66 (Zero-Coupon Bonds)**.** A zero-coupon bond is a bond that pays 1 at maturity. The price of a zero-coupon bond at time t is given by:

$$P(t, T) = e^{-r(T-t)}$$

Can take r as a Stochastic process, then we have:

$$P(t, T) = \mathbb{E}^Q[e^{-\int_t^T r_s ds} \mid \mathcal{F}_t]$$

Definition 2.67 (Forward Contracts)**.** A forward contract is an agreement to buy or sell an asset at a future date for a price agreed upon today. The price of a forward contract at time t is given by:

$$F(t, T) = S_t e^{r(T-t)}$$

$$V_{FWD}(S_0, K, r) = S_0 - Ke^{-rT}$$

Definition 2.68 (Put-Call Parity)**.** We have the following relationship between call and put options:

$$\underbrace{(S_T - K)^+}_{\text{CallPrice}} - \underbrace{(K - S_T)^+}_{\text{PutPrice}} = \underbrace{S_T - K}_{\text{ForwardPrice}}$$

So we get that

$$V_{BS}^{PUT}(0, S_0, K, T, \sigma, r) = Ke^{-rT}\Phi(-d_2) - S_0\Phi(-d_1)$$

And for the delta:

$$\Delta_{PUT} = -\Phi(-d_1) = \Phi(d_1) - 1$$

Definition 2.69 (Dynamic Hedging)**.** Dynamic hedging is the process of continuously adjusting the portfolio to maintain a delta-neutral position. We have that:

$$\Delta_t = \frac{\partial V_t}{\partial S_t}$$

And so the number of shares of the stock/cash to hold is given by:

$$\phi_t^S = \Delta_t, \quad \phi_t^B = (V_t - \Delta_t S_t)/B_t$$

Theorem 2.70. *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

Example 2.71. 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 $\mathbb{Q}$,

$$\begin{aligned} dS_t &= rS_t dt + \underline{\sqrt{V_t}} S_t dW_t, \quad \text{so,} \quad dW dW^V = \rho dt \\ dV_t &= k(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^V, \quad V_0, \end{aligned}$$

we have that now the volatility (see underlined) 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.

For example, a specific call option C with a given strike K and maturity T could be added to $\mathcal{B}_t$ and the stock, and the market would be complete again, because we would have two risky assets now, S_t and C_t, to hedge two sources of risk, W and W^V. A trading strategy would then have to be a triple now, (ϕ^B, ϕ^S, ϕ^C). In reality it's not always possible to find a risky asset matching a given risk, this is particularly 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

Definition 2.72 (Sensitivities/Greeks)**.** The Greeks are sensitivities of the option price to changes in the underlying asset price, time, volatility, and interest rate. We have the following Greeks:

- **Delta: (Change to Initial Price)** $\Delta = \frac{\partial V}{\partial S}$
- **Gamma: (Change to Delta)** $\Gamma = \frac{\partial^2 V}{\partial S^2}$
- **Theta: (Time Decay)** $\Theta = \frac{\partial V}{\partial t}$
- **Vega: (Change to Volatility)** $\nu = \frac{\partial V}{\partial \sigma}$
- **Rho: (Change to Interest Rate)** $\rho = \frac{\partial V}{\partial r}$
- **Lambda: (Leverage)** $\lambda = \frac{\Delta S}{V}$
- **Speed: (Change to Gamma)** $S = \frac{\partial^3 V}{\partial S^3}$

Can use the above to rewrite Ito's Formula (for a call option) as follows:

$$dV(t, S_t) = \Theta dt + \Delta_t dS_t + \frac{1}{2}\sigma^2 \Gamma_t dS_t^2$$

If we have $\Theta < 0$ for a call option, the option loses value over time, and if $\Theta > 0$ the option gains value over time. We have Γ to counteract the effect of Θ on the option price.

Intro to Volatility Smile

Definition 2.73 (Volatility Smile)**.** The volatility smile is a pattern that results from the implied volatilities of options with the same underlying asset and expiration date but different strike prices. The smile is so named because it looks like a smile. The volatility smile is a graph of the implied volatility of option contracts at various strike prices.

Definition 2.74 (Implied Volatility)**.** Implied volatility is the estimated volatility of a security's price. In general, implied volatility increases when the market is bearish, when investors believe that the asset's price will decline, and decreases when the market is bullish, when investors believe that the price will rise. It is a function of the options strike price.

Definition 2.75 (Volatility Surface)**.** The volatility surface is a three-dimensional plot of the implied volatility of options at different strike prices and expiration dates. The volatility surface is used to show the relationship between volatility and moneyness.

Proposition 2.76 (Smile Modelling)**.** *Alternative SDE model for dS can generate a non-flat smile:*

1. *Set K to a starting value;*
2. *Compute the model call option price*

$$V_{Model}(K) = E_0^{\mathbb{Q}}\left[e^{-rT}(S_T - K)^+\right]$$

with S modeled through an alternative dynamics (underlined)

$$\text{Model: } dS_t = rS_t dt + \underline{\sigma(t, S_t)} S_t dW_t, \quad S_0 = s_0$$

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.

Definition 2.77 (Bachelier Model)**.** The Bachelier model is a model for the dynamics of a stock price in which the volatility of the stock is constant. The model is used to price European options. The Bachelier model is a special case of the Black-Scholes model in which the volatility is zero.

$$dS_t = \sigma dW_t^Q$$

The price of a call option in the Bachelier model is given by:

$$\mathbb{E}^Q[(S_T - K)^+] = (S_0 - K)N(d)$$

$$V_{BaM}(0, S_0, K, T, \sigma, r) = (s_0 - K)\Phi(d) + \sigma\sqrt{T} p_N(d), \quad d = \frac{s_0 - K}{\sigma\sqrt{T}}$$

We get the smile curve by inverting the Black-Scholes formula for the Bachelier model.

$$V_{BS}(0, S_0, K, T, \nu(K), r)|_{r=0} = V_{BaM}(0, S_0, K, T, \sigma)$$

Definition 2.78 (Displaced Diffusion Model)**.** Here we define:

$$dS_t = rS_t dt + \sigma(S_t - \alpha e^{rt})dW_t$$

See that the drift is now rS_t and so arbitrage free. The price of a call option in the Displaced Diffusion model is given by:

$$V_{DDM}(0, S_0, K, T, \sigma, r) = (S_0 - \alpha)\Phi(d_1(0)) - Ke^{-rT}\Phi(d_2(0))$$

$$d_1(0) = \frac{\log\left(\frac{S_0-\alpha}{K-\alpha e^{rt}}\right) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}, \quad d_2(0) = d_1(0) - \sigma\sqrt{T}$$

Generating the smile curve by inverting the Black-Scholes formula for the Displaced Diffusion model.

$$V_{BS}(0, S_0, K, T, \nu(K), r) = V_{DDM}(0, S_0, K, T, \sigma, \alpha, r)$$

Definition 2.79 (CEV Model)**.** Constant Elasticity of Variance (CEV) model is a model for the dynamics of a stock price in which the volatility of the stock is a power function of the stock price.

$$dS_t = rS_t dt + \nu S_t^{\gamma} dW_t, \quad S_0 = s_0$$

where γ is the elasticity of variance, we take γ between 0 and 1 . For $\gamma = \frac{1}{2}$ we call it the "Feller Square root process". For $\gamma < 1$, need to say what happens at $S = 0$, usually taken as absorbing boundary.

Definition 2.80 (Mixture Diffusion Dynamics)**.** Wish to build a model

$$dS_t = rS_t dt + \sigma_{mix}(t, S_t)S_t dW_t, \quad S_0 = s_0$$

where $\sigma_{mix}(t, S_t)$ is a mixture of volatilities such that the distribution of S_t is a mixture of log-normal distributions.

$$p_{S_t}(y) =: p_t(y) = \sum_{i=1}^{N} \lambda_i p_{i,t}(y) = \sum_{i=1}^{N} \lambda_i p_{t,\sigma_i}^{lognormal}(y)$$

where $\lambda_i \in (0, 1)$ and $\sum_i \lambda_i = 1$. We have that:

$$\sigma_{mix}(t, y)^2 = \frac{1}{\sum_j \lambda_j p_{j,t}(y)} \sum_i \lambda_i \sigma_i^2 p_{i,t}(y)$$

where $p_{i,t}(y) = \frac{1}{\sigma_i\sqrt{t}} \exp\left\{-\frac{1}{2\sigma_i^2 t}\left(\log\frac{y}{s_0} - (r - \frac{1}{2}\sigma_i^2)t\right)^2\right\}$. Can now write

$$\sigma_{mix}^2(t, y) = \sum_{i=1}^{N} \Lambda_i(t, y)\sigma_i^2, \quad \Lambda_i(t, y) = \frac{\lambda_i p_{i,t}(y)}{\sum_j \lambda_j p_{j,t}(y)}$$

Definition 2.81 (The Shifted Mixture Dynamics model)**.** Write a mixture diffusion dynamics model X_t as

$$dX_t = rX_t dt + \sigma_{mix}(t, X_t)X_t dW_t, \quad X_0 = x_0$$

Assume that the dynamics of S_t are given by

$$S_t = s_0 \alpha e^{rt} + X_t$$

Differentiating to get the dynamics of S_t we get:

$$dS_t = rS_t dt + \sigma_{mix}(t, S_t - s_0\alpha e^{rt})(S_t - s_0\alpha e^{rt})dW_t, \quad S_0 = s_0$$

Price of a call option in the shifted mixture dynamics model is given by:

$$V^{\text{Call}}_{\text{shift-mix}} = e^{-rt}\sum_{i=1}^{N}\lambda_i\left[S_0 e^{rt}\Phi\left(\frac{\ln \mathcal{S}_0/\mathcal{K} + (r + \frac{1}{2}\sigma_i^2)T}{\sigma_i\sqrt{T}}\right) - \mathcal{K}\Phi\left(\frac{\ln \mathcal{S}_0/\mathcal{K} + (r - \frac{1}{2}\sigma_i^2)T}{\sigma_i\sqrt{T}}\right)\right]$$

where $\mathcal{K} = K - s_0\alpha e^{rt}, \mathcal{S}_0 = s_0(1-\alpha)$

Definition 2.82 (Stochastic Volatility Models)**.** The models above are all called local volatility models. In these models the volatility $\sigma(t, S_t)$ in the SDE

$$dS_t = rS_t dt + \sigma(t, S_t)S_t dW_t, s_0$$

is a deterministic function of time and the stock price only. In stochastic volatility models, the volatility is itself a stochastic process.

FILL THE REST OF THIS LATER TOO pg 280 in original notes.

Part III

Risk Measures

Proposition 3.83 (Distribution of log-returns)**.** *From Black-Scholes under measure P we have for stock S_t*

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

Taking logs we get, with $\delta = t_{i+1} - t_i$

$$\log\frac{S_{t_{i+1}}}{S_{t_i}} = \left(\mu - \frac{1}{2}\sigma^2\right)\delta + \sigma(W_{t_{i+1}} - W_{t_i}) \sim \mathcal{N}\left(\left(\mu - \frac{1}{2}\sigma^2\right)\delta, \sigma^2\delta\right)$$

Gaussian distribution! can test this with QQ plots, or sample skewness and kurtosis, both should be 0.

$$\textit{Skewness} = \frac{\mathbb{E}[(X - \mathbb{E}[X])^3]}{\sigma^3}, \quad \textit{Kurtosis} = \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{\sigma^4}$$

Definition 3.84 (VaR - Value at Risk)**.** Defined simply as **the loss level that will not be exceeded with a certain confidence level over a certain period of time**
Value at Risk (VaR) is a measure of the risk of loss for investments. It estimates how much a set of investments might lose (with a given probability), given normal market conditions, in a set time period such as a day. VaR is typically used by firms and regulators in the financial industry to gauge the amount of assets needed to cover possible losses.

$$\mathrm{VaR}_\alpha = -\inf\{x \in \mathbb{R} : P(X \leq x) \geq \alpha\}$$

where α is the confidence level, and X is the loss distribution.

Also define L_H

$$L_H = \mathrm{Portfolio}_0 - \mathrm{Portfolio}_H$$

and take $\Pi(t,T)$ the sum of all future cash flows from the portfolio in $[t,T]$ discounted back at t, for our portfolio.
This gives us the price of the portfolio at time t, for T final maturity.

$$\mathrm{Portfolio}_t = \mathbb{E}_t^Q[\Pi(t,T)]$$

This gives us $VaR_{H,\alpha}$, for horizon H and confidence level α, satisfying:

$$\mathbb{P}\left(L_H < VaR_{H,\alpha}\right) = \alpha$$

Or equiv.

$$\mathbb{P}\left(\mathbb{E}_0^{\mathbb{Q}}[\Pi(0,T)] - \mathbb{E}_H^{\mathbb{Q}}[\Pi(H,T)] < VaR_{H,\alpha}\right) = \alpha$$

Proposition 3.85 (Drawbacks to VaR)**.** *1.* ***Does not take into account the tail structure beyond the percentile.*** *Refer to figure below:*

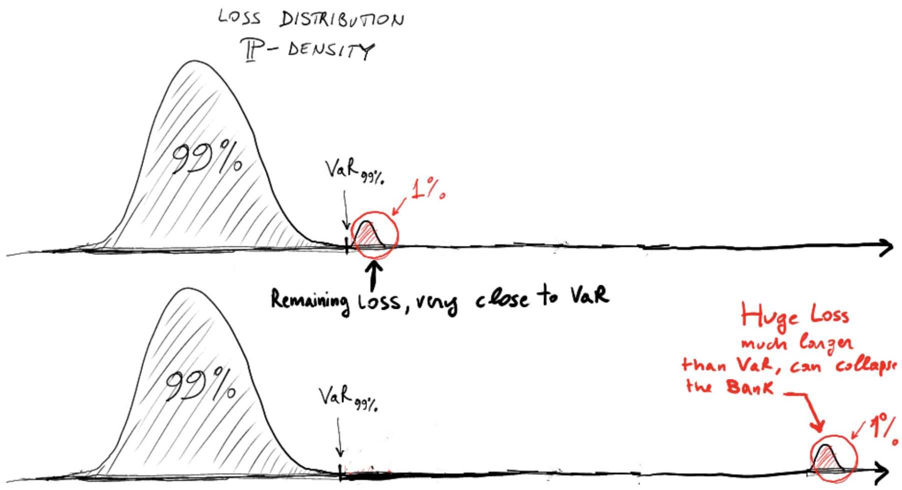

Figure 1: VaR Drawbacks

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?

2. ***VaR is not subadditive****, i.e. the VaR of a portfolio is not the sum of the VaR of the individual assets. This is because VaR is a quantile, and quantiles are not additive.*

Definition 3.86 (ES - Expected Shortfall)**.** Used as a solution to (2) and a partial solution to (1) above.
Expected Shortfall (ES) is a risk measure that quantifies the average loss of the tail of the loss distribution. It is an alternative to Value at Risk that is more sensitive to the shape of the tail of the loss distribution. ES is also known as Conditional Value at Risk (CVaR)

or Expected Tail Loss (ETL).

$$\begin{aligned} ES_\alpha &= \mathbb{E}^{\mathbb{P}}[L_H \mid L_H > VaR_{H,\alpha}] \\ &= \frac{\mathbb{E}^{\mathbb{P}}[L_H \mathbf{1}_{\{L_H > VaR_{H,\alpha}\}}]}{\mathbb{P}\left(L_H > VaR_{H,\alpha}\right)} \\ &= \frac{1}{1-\alpha}\mathbb{E}^{\mathbb{P}}[L_h \mathbf{1}_{\{L_H > VaR_{H,\alpha}\}}] \end{aligned}$$

Proposition 3.87 (Drawbacks to ES)**.** *1. Does not fully solve the problem of (1) above, as it still only considers the average loss in the tail, not the full tail structure.*

*2. **liquidity risk** - common with VaR and ES, as they do not take into account the liquidity of the assets in the portfolio. Namely:*

$$\begin{aligned} VaR(k \cdot \mathit{Portfolio}) &\neq k \cdot VaR(\mathit{Portfolio}) \\ ES(k \cdot \mathit{Portfolio}) &\neq k \cdot ES(\mathit{Portfolio}) \end{aligned}$$

Selling a million shares of a stock will not be as easy as selling one share, and so the risk of the portfolio is not linear with the size of the portfolio.

Part IV
Numerical Solutions of SDEs

Definition 4.88 (Euler Scheme)**.** The Euler scheme is a simple numerical method to solve SDEs. It is a first-order method, and is not very accurate. Let time step be $\Delta t = t_{i+1} - t_i = \delta \forall i$ and write $\Delta W_{t_i} = W_{t_{i+1}} - W_{t_i}, \Delta X_{t_i} = X_{t_{i+1}} - X_{t_i}$. We have:

$$\Delta X_{t_i} = \mu(t_i, X_{t_i})\Delta t_i + \sigma(t_i, X_{t_i})\Delta W_{t_i}, \quad X_0 = Z$$

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 are independent. We get the Euler scheme:

$$X_{t_{i+1}} = X_{t_i} + \mu(t_i, X_{t_i})\delta + \sigma(t_i, X_{t_i})\sqrt{\delta}\mathcal{N}_i(0,1), \quad X_0 = Z$$

Proposition 4.89 (Convergence of Euler Scheme)**.** *The Euler scheme converges to under sufficient conditions for existence and uniqueness of the global solution of our SDE. These are Lipschitz continuity and linear growth conditions on μ and σ. We have an order of convergence of $\frac{1}{2}$. We have that there exists a positive real number δ_0 such that*

$$E\{\left|X_T^{\Delta t} - X_T\right|\} \leq C(T)(\Delta t)^{\frac{1}{2}} \quad \forall \Delta t \leq \delta_0$$

Where $C(T) > 0$ a constant