Session 1: Probability Theory and Lattice Models

Tee Chyng Wen

QF620 Stochastic Modelling in Finance

Properties of Distributions

The **probability density function** (PDF) $f(x)$ contains all information concerning a random variable X. The **expectation** (or mean) of a function $g(X)$ of the random variable X is given by

$$\mu = \quad \mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx.$$

The expectation operation has the following properties:

1. $\mathbb{E}[a] = a$, where a is a constant;
2. $\mathbb{E}[ag(X)] = a\mathbb{E}[g(X)]$, where a is a constant;
3. **Linearity**: if $g(X) = s(X) + t(X)$, then

$$\mathbb{E}[g(X)] = \mathbb{E}[s(X)] + \mathbb{E}[t(X)].$$

Note that in general $\quad \mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$

$$\mathbb{E}[XY] \neq \mathbb{E}[X]\mathbb{E}[Y].$$

Mean and Variance

If X is a random variable with probability density function $f(x)$, then the mean of X is defined as the following expectation

$$\mu = \mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\, dx.$$

The variance of a random variable X is defined as the following expectation

$$V[X] = \mathbb{E}[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx.$$

The variance is also written as σ^2, and has the following properties

1. $V[a] = 0$, where a is a constant;
2. $V[aX+b] = a^2 V[X]$, where a and b are constants.

The variance is always positive, and its positive square root is called the **standard deviation**, denoted as σ. The standard deviation measures the spread of the random variable X around the mean.

Mode and Median

Although the mean discussed previously is the most common measure of the "average" of a random variable, two other useful measures exist which do not require the evaluation of an expectation.

The **mode** of a distribution is the value of the random variable X at which the probability density function $f(x)$ is the greatest.

⇒ If there is more than one of such value then each of them may be called the mode of the distribution.

quintile : 5
decile : 10

2

The **median** M of a distribution is the value of the random variable X at which the cumulative distribution function $F(x)$ takes the value of $\frac{1}{2}$.

4

⇒ Related to this are the lower and upper **quartiles** Q_{lower} and Q_{upper}, where

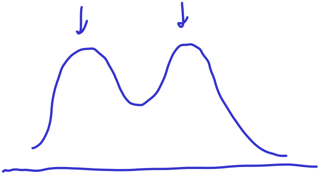
$$F(Q_{lower}) = \frac{1}{4}, \quad F(M) = \frac{1}{2}, \quad F(Q_{upper}) = \frac{3}{4}.$$

⇒ So the median and lower and upper quartiles divide the distribution into 4 regions, each containing a quarter of the probability.

⇒ Smaller subdivisions are also defined, e.g. the n^{th} **percentile**, P_n, of a PDF is defined as $F(P_n) = \frac{n}{100}$.

100

uni-modal



bi-modal

⋮

tri-modal

⋮

multi-modal

Mode and Median

Example I observed the temperature at a place for 110 days:

- For the first 100 days, the temperature is 80 degree.
- From day 101 to 105, the temperature is 85 degree.
- From day 106 to 110, the temperature is 90 degree.

What is the median temperature?

80

$\{1, 1\} = 1$

$\{1, 2\} = 1.5$

$\{1, 1, 2\} = 1$

$\mathbb{E}[X] = \mu$ 1st moment

Moments and Central Moments

The mean of a random variable X is also called the first **moment** of X.

kth moment

Let us introduce the concept of moments: $\mathbb{E}[X^k]$, where k denotes the order (power) of the moment:

$$\mathbb{E}[X^k] = \int_{-\infty}^{\infty} x^k f(x)\, dx.$$

Since

$$V[X] = \mathbb{E}[X^2] - \mu^2,$$

we see that the variance of a random variable can be expressed as its first and second moments.

$\Rightarrow$ In general, it is often a lot easier to calculate the moments of a random variable using the **moment generating function** (MGF) methodology, especially for higher moments where direct evaluation of the expectation becomes laborious.

Moments and Central Moments

The variance $V[X]$ is also called the second **central moment**.

We call this **central** because we subtract μ from x before squaring it, meaning that we are considering the moment about the mean of the distribution, rather than about the origin at $x = 0$.

Now let us introduce the notation $\nu_k = \mathbb{E}\left[(X-\mu)^k\right]$ where k denotes the **order of the central moment**:

$$\begin{aligned}\nu_k &= \mathbb{E}\left[(X-\mu)^k\right] \\ &= \int_{-\infty}^{\infty} (x-\mu)^k f(x)\, dx.\end{aligned}$$

Clearly, the first central moment is $\nu_1 = 0$. The second central moment is the variance by definition.

Moments

$k=1$: $\mathbb{E}[X]$

$k=2$: $\mathbb{E}[X^2]$

$k=3$: $\mathbb{E}[X^3]$

$\vdots$ $\vdots$

Central Moments

$\mathbb{E}[(X-\mu)] = 0$

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

$\mathbb{E}[(X-\mu)^3]$

$\vdots$

Moments and Central Moments

Example Let X be a random variable. Express X's 3^{rd} central moment in terms of its moments.

Solution

nu

$$\nu_3 = \mathbb{E}[(X-\mu)^3]$$
$$= \mathbb{E}[X^3] - 3\mathbb{E}[X]\mathbb{E}[X^2] + 2\mathbb{E}[X]^3. \quad \triangleleft$$

$$(X-\mu)^3$$

$$= X^3 - 3X^2\mu + 3X\mu^2 - \mu^3$$

$$\mathbb{E}[X^3] - 3\mathbb{E}[X^2]\mu + 3\mathbb{E}[X]\mu^2 - \mu^3$$

Moments and Central Moments

A normalized and dimensionless central moment is also defined as

gamma →
$$\gamma_k = \frac{\nu_k}{\sigma^k}.$$

Here, γ_3 and γ_4 are known as **skewness** and **kurtosis** of the distribution, respectively.

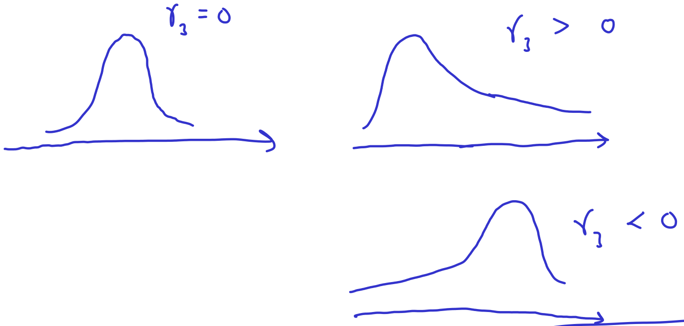
The skewness γ_3 is 0 if the distribution is symmetrical around the mean. If the distribution is skewed to values of x smaller than the mean then we have $\gamma_3 < 0$ (**negative skewed**), and vice versa.

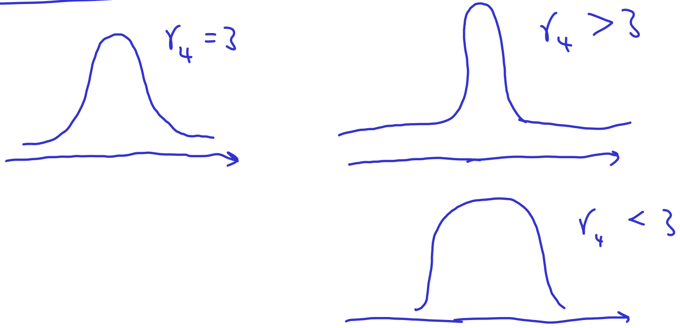
Additional Example Q3, Q4

The kurtosis of a normal distribution $N(\mu, \sigma^2)$ is given by

$$\gamma_4 = \frac{v_4}{\sigma^4} = \frac{3\sigma^4}{\sigma^4} = 3.$$

Using normal distribution as a reference, it is common to define the **excess kurtosis** of a distribution as $\gamma_4 - 3$. A positive excess kurtosis implies a relatively narrower peak and wider tails than the normal distribution with the same mean and variance, and vice versa.


$\gamma_3 = 0$
$\gamma_3 > 0$
$\gamma_3 < 0$


$\gamma_4 = 3$
$\gamma_4 > 3$
$\gamma_4 < 3$

Moment Generating Functions

The **moment generating function** (MGF) of a random variable X following the probability density function $f(x)$ is defined as

$$M_X(\theta) = \mathbb{E}\left[e^{\theta X}\right] = \int_{-\infty}^{\infty} e^{\theta x} f(x)\, dx.$$

It should again be obvious that $M(0) = 1$. MGF is also useful for obtaining the moments of a random variable. Note that

$$\begin{aligned}\mathbb{E}\left[e^{\theta X}\right] &= \mathbb{E}\left[1 + \theta X + \frac{\theta^2 X^2}{2!} + \frac{\theta^3 X^3}{3!} + \cdots\right] \\ &= 1 + \mathbb{E}[X]\theta + \mathbb{E}[X^2]\frac{\theta^2}{2!} + \mathbb{E}[X^3]\frac{\theta^3}{3!} + \cdots .\end{aligned}$$

Using this expression, it becomes clear that we can derive the moments of a random variable in terms of its MGF by

$$\mathbb{E}[X^n] = \left.\frac{d^n M_X(\theta)}{d\theta^n}\right|_{\theta=0}.$$

Taylor Series e^x : $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$

$$M_x(\theta) = 1 + \theta\,\mathbb{E}[X] + \frac{\theta^2\,\mathbb{E}[X^2]}{2!} + \frac{\theta^3\,\mathbb{E}[X^3]}{3!} + \ldots$$

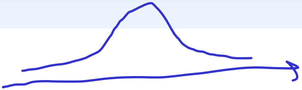
$$\frac{dM_x(\theta)}{d\theta} = 0 + 1\cdot\mathbb{E}[X] + \theta\,\mathbb{E}[X^2] + \frac{\theta^2\,\mathbb{E}[X^3]}{2!} + \ldots$$

$$\frac{dM_x(0)}{d\theta} = \mathbb{E}[X]$$

$$\frac{dM_x^2(\theta)}{d\theta^2} = 0 + 0 + \mathbb{E}[X^2] + \frac{\theta\,\mathbb{E}[X^3]}{1} + \ldots$$

$$\frac{dM_x^2(0)}{d\theta^2} = \mathbb{E}[X^2]$$

Integrating e^{-x^2}



The following integral is one of the most important integrals in quantitative finance—it is related to the normal distribution

$$I = \int_{-\infty}^{\infty} e^{-x^2}\,dx.$$

But this 1-dimensional integral is intractable. We need to consider the 2-dimensional case by switching over to polar coordinates

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-x^2}\,dx \times \int_{-\infty}^{\infty} e^{-y^2}\,dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,dx\,dy \\ &= \int_{0}^{2\pi}\int_{0}^{\infty} e^{-r^2} r\,dr\,d\phi \\ &= \pi. \end{aligned}$$

$\int_0^{2\pi}\int_0^{a} e^{-r^2} r\,dr\,d\phi$

Hence $I = \sqrt{\pi}$ (why do we take the positive root?).

Integrating e^{-x^2}

Since both integration regions are infinite, we can establish a proof of this result by inequality. Let

$$I^2(a) = \int_{-a}^{a}\int_{-a}^{a} e^{-(x^2+y^2)}\,dx\,dy.$$

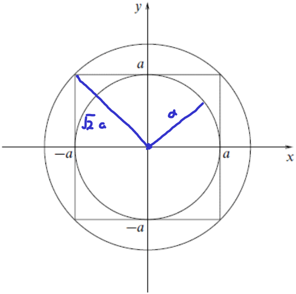
Looking at the integration region, we have the inequality:

$$\pi\left(1-e^{-a^2}\right) < I^2(a) < \pi\left(1-e^{-2a^2}\right)$$

Taking limits on both upper and lower boundaries, we have

$$\lim_{a\to\infty}\pi\left(1-e^{-a^2}\right) = \lim_{a\to\infty}\pi\left(1-e^{-2a^2}\right) = \pi.$$

This completes the proof for $I^2(\infty) = \pi$.



Probability Density Function of Normal Distribution

From this we can obtain the probability density function for **normal distribution**. Starting with

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi},$$

if we let $y = \sqrt{2}x$, we can obtain the result

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\,dy = 1,$$

where $\phi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ is the standard normal probability density function.
Similarly, if we let $y = \sqrt{2}\sigma x + \mu$, we can obtain the result

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}\,dy = 1,$$

where $f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$ is the normal probability density function.

let $y = \sqrt{2}\, x \rightarrow \frac{y}{\sqrt{2}} = x$

$\frac{dy}{dx} = \sqrt{2} \rightarrow \frac{dy}{\sqrt{2}} = dx$

$$\Rightarrow \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \cdot \frac{dy}{\sqrt{2}} = \sqrt{\pi}$$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 1$$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

Differentiation of Integrals

The chain rule for partial differentiation states that if we have a bivariate function $g(u(x), v(x))$, where u and v are both functions of x, then

$$\frac{dg}{dx} = \frac{\partial g}{\partial u}u'(x) + \frac{\partial g}{\partial v}v'(x).$$

Now consider the case where the integrand is $f(x,t)$ and t is the dummy variable. Let F denote the anti-derivative of f, we can write it as

$$\int_{u(x)}^{v(x)} f(x,t)dt = F(x, v(x)) - F(x, u(x)).$$

Let's define

$$I(v(x), u(x), x) = F(x, v(x)) - F(x, u(x)),$$

we can see that

$$\frac{\partial I}{\partial v} = f(x, v(x)), \qquad \text{and} \qquad \frac{\partial I}{\partial u} = -f(x, u(x)).$$

Leibnitz's Rule

$I(v(x), u(x), x)$

Now differentiating I with respect to x, we obtain

$$
\begin{aligned}
\frac{dI}{dx} &= \frac{\partial I}{\partial v} v'(x) + \frac{\partial I}{\partial u} u'(x) + \frac{\partial I}{\partial x} \cancelto{1}{\frac{dx}{dx}} \\
&= f(x, v(x)) v'(x) - f(x, u(x)) u'(x) + \frac{\partial}{\partial x} \int_{u(x)}^{v(x)} f(x,t)\, dt \\
&= f(x, v(x)) v'(x) - f(x, u(x)) u'(x) + \int_{u(x)}^{v(x)} \frac{\partial f(x,t)}{\partial x}\, dt.
\end{aligned}
$$

Under the special case where the integrand isn't a function of x, then we recover the corollary integration relationship:

$$
\frac{d}{dx}\left[\int_{u(x)}^{v(x)} f(t)\, dt\right] = f(v(x)) v'(x) - f(u(x)) u'(x).
$$

Leibnitz's Rule

$f(x,t) = \frac{\sin(xt)}{t}$

Example Differentiate the following with respect to x:

$v(x) = x^2$

$$\int_x^{x^2} \frac{\sin(xt)}{t} dt$$

$u(x) = x$

Solution

$$\begin{aligned}\frac{dI}{dx} &= \frac{\sin(x^3)}{x^2}(2x) - \frac{\sin(x^2)}{x}(1) + \int_x^{x^2} \frac{t\cos(xt)}{t} dt \\ &= \frac{1}{x}(3\sin(x^3) - 2\sin(x^2))\end{aligned}$$

Leibnitz's Rule

PDF

A standard normal probability density function is given by

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

CDF

Integrating the probability density function yields the cumulative distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

In probability theory, integrating ϕ yields Φ, while differentiating Φ yields ϕ. Indeed, differentiating $\Phi(x)$ with respect to x, we obtain, using Leibnitz's Rule

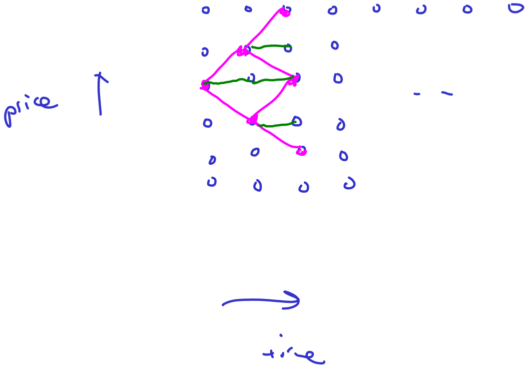
$$\frac{d\Phi(x)}{dx} = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}.$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$v(x) = x$, $u \to -\infty$, $f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

$$\frac{d\Phi(x)}{dx} = f(x) \cdot \frac{dv(x)}{dx} - f(-\infty) \cdot \frac{du}{dx} + \int_{-\infty}^{x} \frac{\partial}{\partial x} f(t)\, dt$$

$$= f(x) \cdot 1 - 0 + 0$$


price
time

Two-State Model

Let us now take a look at the pricing of vanilla options using the concept of a **lattice model** (i.e. a tree model).

$\Rightarrow$ The aim is to illustrate the concepts of **arbitrage**, **risk-neutrality** and **replication**.

Suppose a stock is worth 100 today, and will be worth either 110 or 90 tomorrow. Consider also an at-the-money (ATM) call option struck at $K = 100$. Assume interest rate is 0 for now.

$$\begin{tikzcd}
 & \begin{matrix}110\\10\end{matrix} & = \max\{110-100, 0\} \\
\begin{matrix}100\\C_0\end{matrix} \arrow[ur] \arrow[dr] & & \\
 & \begin{matrix}90\\0\end{matrix} & = \max\{90-100, 0\}
\end{tikzcd}$$

A client wishes to buy this ATM call option from us today. If the stock moves up to 110, the option will be exercised, giving the client a profit of 10. If the stock moves down to 90, the option will expire worthless. So the option value has to be between 0 and 10.

Two-State Model

After selling this call option, we should **hedge** by buying some stock. We do not know where the market is heading, but suppose we buy Δ amount of stock today, so that our portfolio is

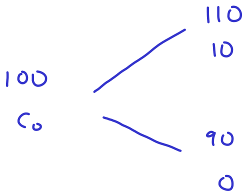
$$\Delta \times \text{stock} - 1 \times \text{call}.$$

We hedge by requiring our portfolio to be worth the same whichever way the market ends up at, i.e.

$$\underbrace{\Delta \times 110 - 1 \times 10}_{\text{up}} = \underbrace{\Delta \times 90 - 1 \times 0}_{\text{down}} \quad \Rightarrow \quad \Delta = \frac{1}{2}.$$

We see that this can be achieved by holding $\frac{1}{2}$ stock today. If we do exactly that, then our portfolio is worth $110 \times \frac{1}{2} - 10 = 90 \times \frac{1}{2} = 45$ in either states of the market. Since this portfolio is **riskfree**, and interest rate is 0, then the portfolio must be worth precisely 45 today also, or there will be arbitrage opportunity. In other words, the **no-arbitrage condition enforces the price**. So the option must be worth

$$100 \times \frac{1}{2} - C_0 = 45 \quad \Rightarrow \quad C_0 = 5.$$



	Cash	Stock	Call
$t=0$	-45	$\Delta \cdot S_0 = \frac{1}{2} \times 100 = 50$	$-C_0 = -5$
$t=1$	-45	up: $\frac{1}{2} \times 110 = 55$ down: $\frac{1}{2} \times 90 = 45$	up: -10 down: 0

Two-State Model

Note that we did not have to use any probability in our analysis above—our result is valid regardless of what probability and up or down jump is.

But what if we know for sure that the stock is going up? This is impossible, because this will contradict the **no-arbitrage condition**:

- The value of a stock tomorrow cannot be guaranteed to grow faster than the risk-free interest rate, otherwise we will borrow at the risk-free rate and invest in stock to arbitrage.
- Hence, the probability of an up jump being 1 is impossible.
- Similarly, the probability of a down jump being 1 is also impossible (why?).
- Hence, the role of the probabilities is just to ensure that both states of the market are possible.

Two-State Model

Can we still use probabilities in our analysis, provided we match the no-arbitrage price obtained in our argument above?

$\Rightarrow$ Definitely. The probabilities that conform to the no-arbitrage price are the **risk-neutral probabilities**.

Let $t = 0$ and $t = 1$ be the timestamps for today and tomorrow, respectively. Let p^* denote the **risk-neutral probability** for an up jump. The expected value of the stock and option will be

$$\begin{cases} \mathbb{E}[S_1] = 110p^* + 90(1 - p^*) & = 100 \\ \mathbb{E}[C_1] = 10p^* & = 5 \end{cases}$$

Observe that

- The only probability that gives the correct no-arbitrage option price is $p^* = \frac{1}{2}$.
- This probability gives an expected stock price of 100, same as today's price, since interest rate is taken to be 0.

Two-State Model

$\mathbb{P}$

$\mathbb{Q}$

Even though the **empirical (real-world) probability** differs from the **risk-neutral probability**, arbitrage is still impossible, since the risk-neutral price by no-arbitrage principle only uses zero and non-zero probabilities.

$\Rightarrow$ This means that if an arbitrage exists for some value of $0 < p < 1$, then it exists for all values of p.

$\Rightarrow$ Conversely, if there are no arbitrage for some value of p, then there are no arbitrages for all values of p.

Hence, if under our **risk-neutral probabilistic approach**, we see that:

$\Rightarrow$ We do not require a premium for the riskiness of the stock over a risk-free bond.

$\Rightarrow$ In the real world, one would expect $p \geq \frac{1}{2}$ due to **risk-aversion**, and the expected value of the option and stock should be greater than the arbitrage-free price.

$\Rightarrow$ The ability to hedge has removed the risk premium.