Mathematical Preliminaries

Introduction to Probability - Moment Generating Function

The *moment generating function* of X, denoted $M_X(\theta)$ is given by

$$M_X(\theta) = \mathbb{E}\left[e^{\theta x}\right] = \int_{\mathbb{R}} e^{\theta x} p(x)\, dx$$

provided the expectation exists. We can expand as a power series to obtain

$$M_X(\theta) = \sum_{n=0}^{\infty} \frac{\theta^n \mathbb{E}(X^n)}{n!}$$

so the n^{th} moment is the coefficient of $\theta^n/n!$, or the n^{th} derivative evaluated at zero.

How do we arrive at this result?

We use the Taylor series expansion for the exponential function: $\int_{\mathbb{R}} e^{\theta x} p(x)\, dx =$

$$\int_{\mathbb{R}} \left(1 + \theta x + \frac{(\theta x)^2}{2!} + \frac{(\theta x)^3}{3!} + \ldots\ldots\right) p(x)\, dx$$

$$\begin{aligned}
&= \underbrace{\int_{\mathbb{R}} p(x)\, dx}_{1} + \theta \underbrace{\int_{\mathbb{R}} x p(x)\, dx}_{\mathbb{E}(X)} + \tfrac{\theta^2}{2!} \underbrace{\int_{\mathbb{R}} x^2 p(x)\, dx}_{\mathbb{E}(X^2)} + \\
&\quad \tfrac{\theta^3}{3!} \underbrace{\int_{\mathbb{R}} x^3 p(x)\, dx}_{\mathbb{E}(X^3)} + \ldots. \\
&= 1 + \theta \mathbb{E}(X) + \tfrac{\theta^2}{2!} \mathbb{E}\left(X^2\right) + \tfrac{\theta^3}{3!} \mathbb{E}\left(X^3\right) + \ldots. \\
&= \sum_{n=0}^{\infty} \frac{\theta^n \mathbb{E}(X^n)}{n!}.
\end{aligned}$$

Calculating Moments

The k^{th} moment m_k of the random variable X can now be obtained by differentiating, i.e.

$$\begin{aligned} m_k &= M_X^{(k)}(\theta); \quad k = 0, 1, 2, ... \\ M_X^{(k)}(\theta) &= \left.\frac{d^k}{d\theta^k} M_X(\theta)\right|_{\theta=0} \end{aligned}$$

So what is this result saying? Consider $M_X(\theta) = \sum_{n=0}^{\infty} \frac{\theta^n \mathbb{E}(X^n)}{n!}$

$$M_X(\theta) = 1 + \theta \mathbb{E}[X] + \frac{\theta^2}{2!} \mathbb{E}\left[X^2\right] + \frac{\theta^3}{3!} \mathbb{E}\left[X^3\right] + ... + \frac{\theta^n}{n!} \mathbb{E}[X^n]$$

As an example suppose we wish to obtain the second moment; differentiate twice with respect to θ

$$\frac{d}{d\theta} M_X(\theta) = \mathbb{E}[X] + \theta \mathbb{E}\left[X^2\right] + \frac{\theta^2}{2} \mathbb{E}\left[X^3\right] + ... + \frac{\theta^{n-1}}{(n-1)!} \mathbb{E}[X^n]$$

and for the second time

$$\frac{d^2}{d\theta^2} M_X(\theta) = \mathbb{E}\left[X^2\right] + \theta \mathbb{E}\left[X^3\right] + ... + \frac{\theta^{n-2}}{(n-2)!} \mathbb{E}[X^n].$$

Setting $\theta = 0$, gives

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

which captures the second moment $\mathbb{E}[X^2]$. Remember we will already have an expression for $M_X(\theta)$.

A useful result in finance is the MGF for the normal distribution. If $X \sim N(\mu, \sigma^2)$, then we can construct a standard normal $\phi \sim N(0,1)$ by setting $\phi = \dfrac{X - \mu}{\sigma} \Longrightarrow X = \mu + \sigma\phi$.

The MGF is

$$\begin{aligned} M_X(\theta) &= \mathbb{E}\left[e^{\theta x}\right] = \mathbb{E}\left[e^{\theta(\mu + \phi\sigma)}\right] \\ &= e^{\theta\mu} \mathbb{E}\left[e^{\theta\sigma\phi}\right] \end{aligned}$$

So the MGF of X is therefore equal to the MGF of ϕ but with θ replaced by $\theta\sigma$.This is much nicer than trying to calculate the MGF of $X \sim N(\mu, \sigma^2)$.

$$\begin{aligned} \mathbb{E}\left[e^{\theta\phi}\right] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\theta x} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\theta x - x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(x^2 - 2\theta x + \theta^2 - \theta^2\right)} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-\theta)^2 + \frac{1}{2}\theta^2} dx \\ &= e^{\frac{1}{2}\theta^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-\theta)^2} dx \end{aligned}$$

Now do a change of variable - put $u = x - \theta$

$$\begin{aligned} \mathbb{E}\left[e^{\theta\phi}\right] &= e^{\frac{1}{2}\theta^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du \\ &= e^{\frac{1}{2}\theta^2} \end{aligned}$$

Thus

$$\begin{aligned} M_X(\theta) &= e^{\theta\mu}\mathbb{E}\left[e^{\theta\sigma\phi}\right] \\ &= e^{\theta\mu+\frac{1}{2}\theta^2\sigma^2} \end{aligned}$$

To get the simpler formula for a standard normal distribution put $\mu = 0,\ \sigma = 1$ to get $M_X(\theta) = e^{\frac{1}{2}\theta^2}$.

We can now obtain the first four moments for a standard normal

$$\begin{aligned} m_1 &= \left.\frac{d}{d\theta}e^{\frac{1}{2}\theta^2}\right|_{\theta=0} \\ &= \left.\theta e^{\frac{1}{2}\theta^2}\right|_{\theta=0} = 0 \end{aligned}$$

$$\begin{aligned} m_2 &= \left.\frac{d^2}{d\theta^2}e^{\frac{1}{2}\theta^2}\right|_{\theta=0} \\ &= \left.\left(\theta^2+1\right)e^{\frac{1}{2}\theta^2}\right|_{\theta=0} = 1 \end{aligned}$$

$$\begin{aligned} m_3 &= \left.\frac{d^3}{d\theta^3}e^{\frac{1}{2}\theta^2}\right|_{\theta=0} \\ &= \left.\left(\theta^3+3\theta\right)e^{\frac{1}{2}\theta^2}\right|_{\theta=0} = 0 \end{aligned}$$

$$\begin{aligned} m_4 &= \left.\frac{d^4}{d\theta^4}e^{\frac{1}{2}\theta^2}\right|_{\theta=0} \\ &= \left.\left(\theta^4+6\theta^2+3\right)e^{\frac{1}{2}\theta^2}\right|_{\theta=0} = 3 \end{aligned}$$

The latter two are particularly useful in calculating the skew and kurtosis.

If X and Y are independent random variables then

$$\begin{aligned} M_{X+Y}(\theta) &= \mathbb{E}\left[e^{\theta(x+y)}\right] \\ &= \mathbb{E}\left[e^{\theta x}e^{\theta y}\right] = \mathbb{E}\left[e^{\theta x}\right]\mathbb{E}\left[e^{\theta y}\right] \\ &= M_X(\theta)\,M_Y(\theta)\,. \end{aligned}$$

Review of Differential Equations

Cauchy Euler Equation

An equation of the form

$$Ly = ax^2\frac{d^2y}{dx^2} + \beta x\frac{dy}{dx} + cy = g\left(x\right)$$

is called a Cauchy-Euler equation.

To solve the homogeneous part, we look for a solution of the form

$$y = x^{\lambda}$$

So $y' = \lambda x^{\lambda-1} \rightarrow y'' = \lambda\left(\lambda-1\right)x^{\lambda-2}$, which upon substitution yields the quadratic, A.E.

$$a\lambda^2 + b\lambda + c = 0,$$

where $b = (\beta - a)$ which can be solved in the usual way - there are 3 cases to consider, depending upon the nature of $b^2 - 4ac$.

Case 1: $b^2 - 4ac > 0 \rightarrow \lambda_1,\ \lambda_2 \in \mathbb{R}$ - 2 real distinct roots

$$\text{GS } y = Ax^{\lambda_1} + Bx^{\lambda_2}$$

Case 2: $b^2 - 4ac = 0 \rightarrow \lambda = \lambda_1 = \lambda_2 \in \mathbb{R}$ - 1 real (double fold) root

$$\text{GS } y = x^{\lambda}\left(A + B\ln x\right)$$

Case 3: $b^2 - 4ac < 0 \rightarrow \lambda = \alpha \pm i\beta \in \mathbb{C}$ - pair of complex conjugate roots

$$\text{GS } y = x^{\alpha}\left(A\cos\left(\beta\ln x\right) + B\sin\left(\beta\ln x\right)\right)$$

Diffusion Process

G is called a diffusion process if

$$dG(t) = A(G,t)\,dt + B(G,t)\,dW(t) \tag{1}$$

This is also an example of a Stochastic Differential Equation (SDE) for the process G and consists of two components:

1. $A(G,t)\,dt$ is deterministic – coefficient of dt is known as the *drift* of the process.
2. $B(G,t)\,dW$ is random – coefficient of dW is known as the *diffusion* or *volatility* of the process.

We say G evolves according to (or follows) this process.

For example

$$dG(t) = (G(t) + G(t-1))\,dt + dW(t)$$

is not a diffusion (although it is a SDE)

- $A \equiv 0$ and $B \equiv 1$ reverts the process back to Brownian motion
- Called time-homogeneous if A and B are not dependent on t.
- $dG^{\,2} = B^2 dt$.

We say (1) is a SDE for the process G or a *Random Walk* for dG.

The diffusion (1) can be written in integral form as

$$G(t) = G(0) + \int_0^t A(G,\tau)\,d\tau + \int_0^t B(G,\tau)\,dW(\tau)$$

Remark: A diffusion G is a *Markov* process if - once the present state $G(t) = g$ is given, the past $\{G(\tau),\ \tau < t\}$ is irrelevant to the future dynamics.

We have seen that Brownian motion can take on negative values so its direct use for modelling stock prices is unsuitable. Instead a non-negative variation of Brownian motion called geometric Brownian motion (GBM) is used

If for example we have a diffusion $G(t)$

$$dG = \mu G dt + \sigma G dW \tag{2}$$

then the drift is $A(G,t) = \mu G$ and diffusion is $B(G,t) = \sigma G$.

The process (2) is also called Geometric Brownian Motion (GBM).

Brownian motion $W(t)$ is used as a basis for a wide variety of models. Consider a pricing process $\{S(t) : t \in \mathbb{R}_+\}$: we can model its instantaneous change dS by a SDE

$$dS = a(S,t)\,dt + b(S,t)\,dW \tag{3}$$

By choosing different coefficients a and b we can have various properties for the diffusion process.

A very popular finance model for generating asset prices is the GBM model given by (2). The instantaneous return on a stock $S(t)$ is a constant coefficient SDE

$$\frac{dS}{S} = \mu dt + \sigma dW \tag{4}$$

where μ and σ are the return's drift and volatility, respectively.

An Extension of Itô's Lemma (2D)

Now suppose we have a function $V = V(S,t)$ where S is a process which evolves according to (4). If $S \longrightarrow S + dS$, $t \longrightarrow t + dt$ then a natural question to ask is "what is the jump in V?" To answer this we return to Taylor, which gives

$$\begin{aligned} & V(S+dS, t+dt) \\ = \quad & V(S,t) + \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}dS^2 + O\left(dS^3, dt^2\right) \end{aligned}$$

So S follows

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

Remember that

$$\mathbb{E}(dW) = 0, \quad dW^2 = dt$$

we only work to $O(dt)$ - anything smaller we ignore and we also know that

$$dS^2 = \sigma^2 S^2 dt$$

So the change dV when $V(S,t) \to V(S+dS, t+dt)$ is given by

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}\left[S\mu dt + S\sigma dW\right] + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt$$

Re-arranging to have the standard form of a SDE $dG = a(G,t)\,dt + b(G,t)\,dW$ gives

$$dV = \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S\frac{\partial V}{\partial S}dW. \tag{5}$$

This is Itô's Formula in two dimensions.

Naturally if $V = V(S)$ then (5) simplifies to the shorter version

$$dV = \left(\mu S\frac{dV}{dS} + \frac{1}{2}\sigma^2 S^2\frac{d^2 V}{dS^2}\right)dt + \sigma S\frac{dV}{dS}dW. \tag{6}$$

Examples: In the following cases S evolves according to GBM.

Given $V = t^2 S^3$ obtain the SDE for V, i.e. dV. So we calculate the following terms

$$\frac{\partial V}{\partial t} = 2tS^3, \quad \frac{\partial V}{\partial S} = 3t^2S^2 \to \frac{\partial^2 V}{\partial S^2} = 6t^2 S.$$

We now substitute these into (5) to obtain

$$dV = \left(2tS^3 + 3\mu t^2 S^3 + 3\sigma^2 S^3 t^2\right)dt + 3\sigma t^2 S^3 dW.$$

Now consider the example $V = \exp(tS)$

Again, function of 2 variables. So

$$\begin{aligned} \frac{\partial V}{\partial t} &= S\exp(tS) = SV \\ \frac{\partial V}{\partial S} &= t\exp(tS) = tV \\ \frac{\partial^2 V}{\partial S^2} &= t^2 V \end{aligned}$$

Substitute into (5) to get

$$dV = V\left(S + \mu t S + \frac{1}{2}\sigma^2 S^2 t^2\right) dt + (\sigma S t V)\, dW.$$

Not usually possible to write the SDE in terms of V − but if you can do so - do not struggle to find a relation if it does not exist. Always works for exponentials.

One more example: That is $S(t)$ evolves according to GBM and $V = V(S) = S^n$. So use

$$dV = \left[\mu S \frac{dV}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2V}{dS^2}\right] dt + \left[\sigma S \frac{dV}{dS}\right] dW.$$

$$V'(S) = nS^{n-1} \rightarrow V''(S) = n(n-1) S^{n-2}$$

Therefore Itô gives us $dV =$

$$\left[\mu S n S^{n-1} + \frac{1}{2}\sigma^2 S^2 n(n-1) S^{n-2}\right] dt + \left[\sigma S n S^{n-1}\right] dW$$

$$dV = \left[\mu n S^n + \frac{1}{2}\sigma^2 n(n-1) S^n\right] dt + [\sigma n S^n]\, dW$$

Now we know $V(S) = S^n$, which allows us to write

$$dV = V\left[\mu n + \frac{1}{2}\sigma^2 n(n-1)\right] dt + [\sigma n]\, V dW$$

with drift $= V\left[\mu n + \frac{1}{2}\sigma^2 n(n-1)\right]$ and diffusion $= \sigma n V$.

Important Cases - Equities and Interest Rates

If we now consider S which follows a lognormal random walk, i.e. $V = \log(S)$ then substituting into (6) gives

$$d\left((\log S)\right) = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW$$

Integrating both sides over a given time horizon (between t_0 and T)

$$\int_{t_0}^{T} d\left((\log S)\right) = \int_{t_0}^{T} \left(\mu - \frac{1}{2}\sigma^2\right) dt + \int_{t_0}^{T} \sigma dW \quad (T > t_0)$$

we obtain

$$\log \frac{S(T)}{S(t_0)} = \left(\mu - \frac{1}{2}\sigma^2\right)(T - t_0) + \sigma\left(W(T) - W(t_0)\right)$$

Assuming at $t_0 = 0$, $W(0) = 0$ and $S(0) = S_0$ the exact solution becomes

$$S_T = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma\phi\sqrt{T}\right\}. \tag{7}$$

(7) is of particular interest when considering the pricing of a simple European option due to its non path dependence. Stock prices cannot become negative, so we allow S, a non-dividend paying stock to evolve according to the lognormal process given above - and acts as the starting point for the Black-Scholes framework.

However μ is replaced by the risk-free interest rate r in (7) and the introduction of the risk-neutral measure - in particular the Monte Carlo method for option pricing.

Interest rates exhibit a variety of dynamics that are distinct from stock prices, requiring the development of specific models to include behaviour such as return to equilibrium, boundedness and positivity. Here we consider another important example of a SDE, put forward by Vasicek in 1977. This model has a mean reverting Ornstein-Uhlenbeck process for the short rate and is used for generating interest rates, given by

$$dr_t = (\eta - \gamma r_t)\, dt + \sigma dW_t. \tag{8}$$

So drift $= (\eta - \gamma r_t)$ and volatility $= \sigma$.

γ refers to the reversion rate and $\frac{\eta}{\gamma}(= \overline{r})$ denotes the mean rate, and we can rewrite this random walk (7) for dr as

$$dr_t = -\gamma\, (r_t - \overline{r})\, dt + \sigma dW_t.$$

By setting $\Theta_t = r_t - \overline{r}$, Θ_t is a solution of

$$d\Theta_t = -\gamma \Theta_t dt + \sigma dW_t;\ \Theta_0 = \alpha, \tag{9}$$

hence it follows that Θ_t is an Ornstein-Uhlenbeck process and an analytic solution for this equation exists. (9) can be written as $d\Theta_t + \gamma\Theta_t dt = \sigma dW_t$.

Multiply both sides by an integrating factor $e^{\gamma t}$

$$\begin{aligned} e^{\gamma t}\,(d\Theta_t + \gamma\Theta_t)\, dt &= \sigma e^{\gamma t} dW_t \\ d\left(e^{\gamma t}\Theta_t\right) &= \sigma e^{\gamma t} dW_t \end{aligned}$$

Integrating over $[0, t]$ gives

$$\begin{aligned} \int_0^t d\,(e^{\gamma s}\Theta_s) &= \int_0^t \sigma e^{\gamma s} dW_s \\ e^{\gamma s}\Theta_s\big|_0^t &= \int_0^t \sigma e^{\gamma s} dW_s \to e^{\gamma t}\Theta_t - \Theta_0 = \int_0^t \sigma e^{\gamma s} dW_s \end{aligned}$$

$$\Theta_t = \alpha e^{-\gamma t} + \sigma \int_0^t e^{\gamma(s-t)} dW_s. \tag{10}$$

By using integration by parts, i.e. $\displaystyle\int v\ du = uv - \int u\ dv$ we can simplify (10).

$$\begin{aligned} &u = W_s \\ &v = e^{\gamma(s-t)} \to dv = \gamma e^{\gamma(s-t)} ds \end{aligned}$$

Therefore

$$\int_0^t e^{\gamma(s-t)} dW_s = W_t - \gamma \int_0^t e^{\gamma(s-t)} W_s\ ds$$

and we can write (10) as

$$\Theta_t = \alpha e^{-\gamma t} + \sigma \left(W_t - \gamma \int_0^t e^{\gamma(s-t)} W_s\ ds \right)$$

allowing numerical treatment for the integral term.

Higher Dimensional Itô

Consider the case where N shares follow the usual Geometric Brownian Motions, i.e.

$$dS_i = \mu_i S_i dt + \sigma_i S_i dW_i,$$

for $1 \leq i \leq N$. The share price changes are correlated with correlation coefficient ρ_{ij}. By starting with a Taylor series expansion

$$V\left(t+\delta t, S_1+\delta S_1, S_2+\delta S_2, \ldots.., S_N+\delta S_N\right) =$$
$$V\left(t, S_1, S_2, \ldots.., S_N\right) + \tfrac{\partial V}{\partial t} + \sum_{i=1}^{N} \tfrac{\partial V}{\partial S_i} dS_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=i}^{N} \tfrac{\partial^2 V}{\partial S_i \partial S_j} + \ldots.$$

which becomes, using $dW_i dW_j = \rho_{ij} dt$

$$dV = \left(\frac{\partial V}{\partial t} + \sum_{i=1}^{N} \mu_i S_i \frac{\partial V}{\partial S_i} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=i}^{N} \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} \right) dt + \sum_{i=1}^{N} \sigma_i S_i \frac{\partial V}{\partial S_i} dW_i.$$

We can integrate both sides over 0 and t to give

$$\begin{aligned} V\left(t, S_1, S_2, \ldots.., S_N\right) &= V\left(0, S_1, S_2, \ldots.., S_N\right) + \\ &\int_0^t \left(\tfrac{\partial V}{\partial \tau} + \sum_{i=1}^{N} \mu_i S_i \tfrac{\partial V}{\partial S_i} + \tfrac{1}{2}\sum_{i=1}^{N}\sum_{j=i}^{N} \sigma_i \sigma_j \rho_{ij} S_i S_j \tfrac{\partial^2 V}{\partial S_i \partial S_j} \right) d\tau \\ &+ \int_0^t \sum_{i=1}^{N} \sigma_i S_i \tfrac{\partial V}{\partial S_i} dW_i. \end{aligned}$$

Discrete Time Random Walks

When simulating a random walk we write the SDE given by (6) in discrete form

$$\delta S = S_{i+1} - S_i = r S_i \delta t + \sigma S_i \phi \sqrt{\delta t}$$

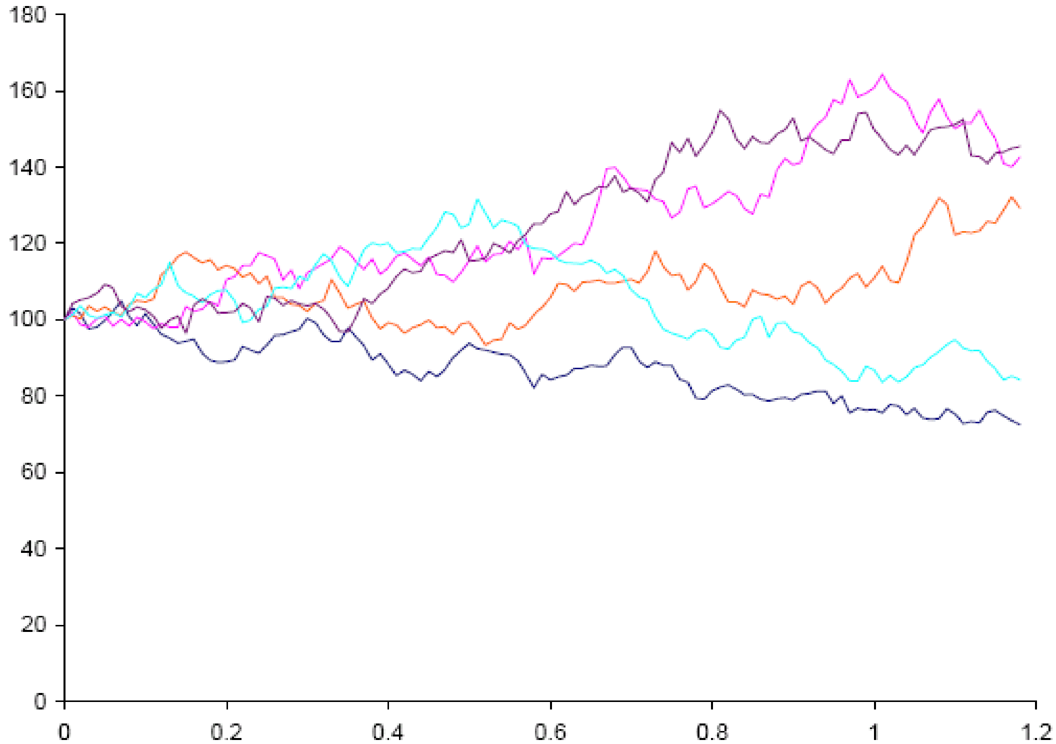
which becomes

$$S_{i+1} = S_i \left(1 + r\delta t + \sigma \phi \sqrt{\delta t}\right). \tag{11}$$

This gives us a time-stepping scheme for generating an asset price realization if we know S_0, i.e. $S(t)$ at $t = 0$. $\phi \sim N(0, 1)$ is a random variable with a standard Normal distribution.

Alternatively we can use discrete form of the analytical expression (7)

$$S_{i+1} = S_i \exp\left\{ \left(r - \frac{1}{2}\sigma^2\right) \delta t + \sigma \phi \sqrt{\delta t} \right\}.$$



So we now start generating random numbers. In C++ we produce uniformly distributed random variables and then use the Box Muller transformation (Polar Marsaglia method) to convert them to Gaussians.

This can also be generated on an Excel spreadsheet using the in-built random generator function RAND(). A crude (but useful) approximation for ϕ can be obtained from

$$\sum_{i=1}^{12} \mathrm{RAND}\,() - 6$$

where $\mathrm{RAND}() \sim U\,[0,1]$.

A more accurate (but slower) ϕ can be computed using $\mathrm{NORMSINV}(\mathrm{RAND}\,())$.

Dynamics of Vasicek Model

The Vasicek model

$$dr_t = \gamma\left(\overline{r} - r_t\right)dt + \sigma dW_t$$

is an example of a *Mean Reverting Process* - an important property of interest rates. γ refers to the *reversion rate* (also called the speed of reversion) and $\overline{r}$ denotes the *mean rate.*

γ acts like a "spring". Mean reversion means that a process which increases has a negative trend (γ pulls it down to a mean level $\overline{r}$), and when r_t decreases on average γ pulls it back up to $\overline{r}$.

In discrete time we can approximate this by writing (as earlier)

$$r_{i+1} = r_i + \gamma\left(\overline{r} - r_i\right)\delta t + \sigma\phi\sqrt{\delta t}$$

To gain an understanding of the properties of this model, look at dr in the absence of randomness

$$\begin{aligned} dr &= -\gamma\left(r - \overline{r}\right)dt \\ \int \frac{dr}{\left(r - \overline{r}\right)} &= -\gamma \int dt \\ r\left(t\right) &= \overline{r} + k\exp\left(-kt\right) \end{aligned}$$

So γ controls the rate of exponential decay.

One of the disadvantages of the Vasicek model is that interest rates can become negative. The Cox Ingersoll Ross (CIR) model is similar to the above SDE but is scaled with the interest rate:

$$dr_t = \gamma\left(\overline{r} - r_t\right)dt + \sigma\sqrt{r_t}dW_t.$$

If r_t ever gets close to zero, the amount of randomness decreases, i.e. diffusion $\longrightarrow 0$, therefore the drift dominates, in particular the mean rate.

Producing Standardized Normal Random Variables

Consider the RAND() function in Excel that produces a uniformly distributed random number over 0 and 1, written $\mathbf{Unif}_{[0,1]}$. We can show that for a large number N,

$$\lim_{N\to\infty}\sqrt{\frac{12}{N}}\left(\sum_{1}^{N}\mathbf{Unif}_{[0,1]}-\tfrac{N}{2}\right)\sim N\left(0,1\right).$$

Introduce $\mathbf{U}_i$ to denote a uniformly distributed random variable over $[0,1]$ and sum up. Recall that

$$\begin{aligned}\mathbb{E}\left[\mathbf{U}_i\right] &= \tfrac{1}{2}\\ \mathbb{V}\left[\mathbf{U}_i\right] &= \tfrac{1}{12}\end{aligned}$$

The mean is then

$$\mathbb{E}\left[\sum_{i=1}^{N}\mathbf{U}_i\right]=N/2$$

so subtract off $N/2$, so we examine the variance of $\left(\sum_{1}^{N}\mathbf{U}_i-\frac{N}{2}\right)$

$$\begin{aligned}\mathbb{V}\left[\sum_{1}^{N}\mathbf{U}_i-\tfrac{N}{2}\right] &= \sum_{1}^{N}\mathbb{V}\left[\mathbf{U}_i\right]\\ &= N/12\end{aligned}$$

As the variance is not 1, write

$$\mathbb{V}\left[\alpha\left(\sum_{1}^{N}\mathbf{U}_i-\tfrac{N}{2}\right)\right]$$

for some $\alpha\in\mathbb{R}$. Hence $\alpha^2\frac{N}{12}=1$ which gives $\alpha=\sqrt{12/N}$ which normalises the variance. Then we achieve the result

$$\sqrt{\frac{12}{N}}\left(\sum_{1}^{N}\mathbf{U}_i-\tfrac{N}{2}\right).$$

Rewrite as

$$\frac{\left(\sum_{1}^{N}\mathbf{U}_i-N\times\tfrac{1}{2}\right)}{\sqrt{\tfrac{1}{12}}\sqrt{N}}.$$

and for $N\to\infty$ by the Central Limit Theorem we get $N\left(0,1\right)$

Generating Correlated Normal Variables

Consider two uncorrelated standard Normal variables ε_1 and ε_2 from which we wish to form a correlated pair ϕ_1, & ϕ_2 $(\sim N(0,\ 1))$, such that $\mathbb{E}[\phi_1\phi_2] = \rho$. The following scheme can be used

1. $\mathbb{E}[\varepsilon_1] = \mathbb{E}[\varepsilon_2] = 0$; $\mathbb{E}[\varepsilon_1^2] = \mathbb{E}[\varepsilon_2^2] = 1$ and $\mathbb{E}[\varepsilon_1\varepsilon_2] = 0$ $(\because\ \varepsilon_1,\ \varepsilon_2$ are uncorrelated$)$.
2. Set $\phi_1 = \varepsilon_1$ and $\phi_2 = \alpha\varepsilon_1 + \beta\varepsilon_2$ (i.e. a linear combination).
3. Now

$$\begin{aligned} \mathbb{E}[\phi_1\phi_2] &= \rho = \mathbb{E}[\varepsilon_1(\alpha\varepsilon_1 + \beta\varepsilon_2)] \\ \mathbb{E}[\varepsilon_1(\alpha\varepsilon_1 + \beta\varepsilon_2)] &= \rho \\ \alpha\mathbb{E}\left[\varepsilon_1^2\right] + \beta\mathbb{E}[\varepsilon_1\varepsilon_2] &= \rho \to \alpha = \rho \end{aligned}$$

$$\begin{aligned} \mathbb{E}\left[\phi_2^2\right] &= 1 = \mathbb{E}\left[(\alpha\varepsilon_1 + \beta\varepsilon_2)^2\right] \\ &= \mathbb{E}\left[\alpha^2\varepsilon_1^2 + \beta^2\varepsilon_2^2 + 2\alpha\beta\varepsilon_1\varepsilon_2\right] \\ &= \alpha^2\mathbb{E}\left[\varepsilon_1^2\right] + \beta^2\mathbb{E}\left[\varepsilon_2^2\right] + 2\alpha\beta\mathbb{E}[\varepsilon_1\varepsilon_2] = 1 \\ \rho^2 + \beta^2 &= 1 \to \beta = \sqrt{1-\rho^2} \end{aligned}$$

4. This gives $\phi_1 = \varepsilon_1$ and $\phi_2 = \rho\varepsilon_1 + \left(\sqrt{1-\rho^2}\right)\varepsilon_2$ which are correlated standardized Normal variables.

Transition Probability Density Functions for Stochastic Differential Equations

To match the mean and standard deviation of the trinomial model with the continuous-time random walk we choose the following definitions for the probabilities

$$\begin{aligned}\phi^{+}(y,t) &= \frac{1}{2}\frac{\delta t}{\delta y^2}\left(B^2(y,t)+A(y,t)\,\delta y\right),\\ \phi^{-}(y,t) &= \frac{1}{2}\frac{\delta t}{\delta y^2}\left(B^2(y,t)-A(y,t)\,\delta y\right)\end{aligned}$$

We first note that the expected value is

$$\begin{aligned}&\phi^{+}(\delta y)+\phi^{-}(-\delta y)+\left(1-\phi^{+}-\phi^{-}\right)(0)\\ =\;&\left(\phi^{+}-\phi^{-}\right)\delta y\end{aligned}$$

We already know that the mean and variance of the continuous time random walk given by

$$dy = A(y,t)\,dt + b(y,t)\,dW$$

is, in turn,

$$\begin{aligned}\mathbb{E}[dy] &= Adt\\ \mathbb{V}[dy] &= B^2dt.\end{aligned}$$

So to match the mean requires

$$\left(\phi^{+}-\phi^{-}\right)\delta y = A\delta t$$

The variance of the trinomial model is $\mathbb{E}[u^2]-\mathbb{E}^2[u]$ and hence becomes

$$\begin{aligned}&(\delta y)^2\left(\phi^{+}+\phi^{-}\right)-\left(\phi^{+}-\phi^{-}\right)^2(\delta y)^2\\ =\;&(\delta y)^2\left(\phi^{+}+\phi^{-}-\left(\phi^{+}-\phi^{-}\right)^2\right).\end{aligned}$$

We now match the variances to get

$$(\delta y)^2\left(\phi^{+}+\phi^{-}-\left(\phi^{+}-\phi^{-}\right)^2\right)=B^2\delta t$$

First equation gives

$$\phi^{+}=\phi^{-}+A\tfrac{\delta t}{\delta y}$$

which upon substituting into the second equation gives

$$(\delta y)^2\left(\phi^{-}+\alpha+\phi^{-}-\left(\phi^{-}+\alpha-\phi^{-}\right)^2\right)=B^2\delta t$$

where $\alpha = A\frac{\delta t}{\delta y}$. This simplifies to

$$2\phi^{-}+\alpha-\alpha^2=B^2\tfrac{\delta t}{(\delta y)^2}$$

which rearranges to give

$$\begin{aligned}\phi^{-} &= \frac{1}{2}\left(B^2\tfrac{\delta t}{(\delta y)^2}+\alpha^2-\alpha\right)\\ &= \frac{1}{2}\left(B^2\tfrac{\delta t}{(\delta y)^2}+\left(A\tfrac{\delta t}{\delta y}\right)^2-A\tfrac{\delta t}{\delta y}\right)\\ &= \frac{1}{2}\tfrac{\delta t}{(\delta y)^2}\left(B^2+A^2\delta t-A\delta y\right)\end{aligned}$$

δt is small compared with δy and so

$$\phi^{-} = \frac{1}{2}\tfrac{\delta t}{(\delta y)^2}\left(B^2 - A\delta y\right).$$

Then

$$\phi^{+} = \phi^{-} + A\tfrac{\delta t}{\delta y} = \frac{1}{2}\tfrac{\delta t}{(\delta y)^2}\left(B^2 + A\delta y\right).$$

Note

$$\left(\phi^{+} + \phi^{-}\right)(\delta y)^2 = B^2\delta t$$

Derivation of the Fokker-Planck/Forward Kolmogorov Equation

Recall that $y',\ t'$ are futures states.

We have $p\left(y,t;y',t'\right) =$

$$\begin{aligned}
&\phi^{-}\left(y'+\delta y,t'-\delta t\right)p\left(y,t;y'+\delta y,t'-\delta t\right)\\
&+\left(1-\phi^{-}\left(y',t'-\delta t\right)-\phi^{+}\left(y',t'-\delta t\right)\right)p\left(y,t;y',t'-\delta t\right)\\
&+\phi^{+}\left(y'-\delta y,t'-\delta t\right)p\left(y,t;y'-\delta y,t'-\delta t\right)
\end{aligned}$$

Expand each of the terms in Taylor series about the point y',t' to find

$$p\left(y,t;y'+\delta y,t'-\delta t\right)=p\left(y,t;y',t'\right)+\delta y\frac{\partial p}{\partial y'}+\frac{1}{2}\delta y^2\frac{\partial^2 p}{\partial y'^2}-\delta t\frac{\partial p}{\partial t'}+.....,$$

$$p\left(y,t;y',t'-\delta t\right)=p\left(y,t;y',t'\right)-\delta t\frac{\partial p}{\partial t'}+...,$$

$$p\left(y,t;y'-\delta y,t'-\delta t\right)=p\left(y,t;y',t'\right)-\delta y\frac{\partial p}{\partial y'}+\frac{1}{2}\delta y^2\frac{\partial^2 p}{\partial y'^2}-\delta t\frac{\partial p}{\partial t'}+.....,$$

$$\phi^{+}\left(y'-\delta y,t'-\delta t\right)=\phi^{+}\left(y',t'\right)-\delta y\frac{\partial \phi^{+}}{\partial y'}+\frac{1}{2}\delta y^2\frac{\partial^2 \phi^{+}}{\partial y'^2}-\delta t\frac{\partial \phi^{+}}{\partial t'}+......,$$

$$\phi^{+}\left(y',t'-\delta t\right)=\phi^{+}\left(y',t'\right)-\delta t\frac{\partial \phi^{+}}{\partial t'}+......,$$

$$\phi^{-}\left(y'+\delta y,t'-\delta t\right)=\phi^{-}\left(y',t'\right)+\delta y\frac{\partial \phi^{-}}{\partial y'}+\frac{1}{2}\delta y^2\frac{\partial^2 \phi^{-}}{\partial y'^2}-\delta t\frac{\partial \phi^{-}}{\partial t'}+......,$$

$$\phi^{-}\left(y',t'-\delta t\right)=\phi^{-}\left(y',t'\right)-\delta t\frac{\partial \phi^{-}}{\partial t'}+......,$$

Substituting in our equation for $p\left(y,t;y',t'\right)$, ignoring terms smaller than δt, noting that $\delta y\sim O\left(\sqrt{\delta t}\right)$, gives

$$\frac{\partial p}{\partial t'}=-\frac{\partial}{\partial y'}\left(\frac{1}{\delta y}\left(\phi^{+}-\phi^{-}\right)p\right)+\frac{1}{2}\frac{\partial^2}{\partial y'^2}\left(\left(\phi^{+}-\phi^{-}\right)p\right).$$

Noting the earlier results

$$\begin{aligned}
A &= \frac{\left(\delta y\right)^2}{\delta t}\left(\frac{1}{\delta y}\left(\phi^{+}-\phi^{-}\right)\right),\\
B^2 &= \frac{\left(\delta y\right)^2}{\delta t}\left(\phi^{+}+\phi^{-}\right)
\end{aligned}$$

gives the *forward equation*

$$\frac{\partial p}{\partial t'}=\frac{1}{2}\frac{\partial^2}{\partial y'^2}\left(B^2\left(y',t'\right)p\right)-\frac{\partial}{\partial y'}\left(A\left(y',t'\right)p\right)$$

The initial condition used is

$$p\left(y,t;y',t'\right)=\delta\left(y'-y\right)$$

As an example consider the important case of the distribution of stock prices. Given the random walk for equities, i.e. Geometric Brownian Motion

$$\frac{dS}{S} = \mu dt + \sigma dW.$$

So $A(S', t') = \mu S'$ and $B(S', t') = \sigma S'$. Hence the forward becomes

$$\frac{\partial p}{\partial t'} = \frac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma^2 S'^2 p\right) - \frac{\partial}{\partial S'}\left(\mu S' p\right).$$

This can be solved with a starting condition of $S' = S$ at $t' = t$ to give the transition pdf

$$p(S, t; S', T) = \frac{1}{\sigma S' \sqrt{2\pi(t' - t)}} e^{-\left(\log(S/S') + \left(\mu - \frac{1}{2}\sigma^2\right)(t'-t)\right)^2 / 2\sigma^2(t'-t)}.$$

More on this and solution technique later, but note that a transformation reduces this to the one dimensional heat equation and the *similarity reduction method* which follows is used.

The Steady-State Distribution

As the name suggests 'steady state' refers to time independent. Random walks for interest rates and volatility can be modelled with stochastic differential equations which have steady-state distributions. So in the long run, i.e. as $t' \longrightarrow \infty$ the distribution $p(y, t; y', t')$ settles down and becomes independent of the starting state y and t. The partial derivatives in the forward equation now become ordinary ones and the unsteady term $\frac{\partial p}{\partial t'}$ vanishes.

The resulting forward equation for the steady-state distribution $p_\infty(y')$ is governed by the ordinary differential equation

$$\frac{1}{2}\frac{d^2}{dy'^2}\left(B^2 p_\infty\right) - \frac{d}{dy'}\left(A p_\infty\right) = 0.$$

Example: The Vasicek model for the spot rate r evolves according to the stochastic differential equation

$$dr = \gamma(\overline{r} - r)\, dt + \sigma dW$$

Write down the Fokker-Planck equation for the transition probability density function for the interest rate r in this model.

Now using the steady-state version for the forward equation, solve this to find the **steady state** probability distribution $p_\infty(r')$, given by

$$p_\infty = \tfrac{1}{\sigma}\sqrt{\frac{\gamma}{\pi}} \exp\left(-\tfrac{\gamma}{\sigma^2}\left(r' - \overline{r}\right)^2\right).$$

Solution:

For the SDE $dr = \gamma(\overline{r} - r)\, dt + \sigma dW$ where drift $= \gamma(\overline{r} - r)$ and diffusion is σ the Fokker Planck equation becomes

$$\frac{\partial p}{\partial t'} = \frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial r'^{\,2}} - \gamma \frac{\partial}{\partial r'}\left(\left(\overline{r} - r'\right) p\right)$$

where $p = p(r', t')$ is the transition PDF and the variables refer to future states. In the steady state case, there is no time dependency, hence the Fokker Planck PDE becomes an ODE with

$$\frac{1}{2}\sigma^2 \frac{d^2 p_\infty}{dr^2} - \gamma \frac{d}{dr}\left(\left(\overline{r} - r\right) p_\infty\right) = 0$$

$p_\infty = p_\infty(r)$. The prime notation and subscript have been dropped simply for convenience at this stage. To solve the steady-state equation:

Integrate wrt r

$$\frac{1}{2}\sigma^2 \frac{dp}{dr} - \gamma\left((\overline{r} - r)\, p\right) = k$$

where k is a constant of integration and can be calculated from the conditions, that as $r \to \infty$

$$\begin{cases} \dfrac{dp}{dr} \to 0 \\ p \to 0 \end{cases} \Rightarrow k = 0$$

which gives

$$\frac{1}{2}\sigma^2 \frac{dp}{dr} = \gamma\left((\overline{r} - r)\, p\right),$$

a first order variable separable equation. So

$$\begin{aligned} \frac{1}{2}\sigma^2 \int \frac{dp}{p} &= \gamma \int \left((\overline{r} - r)\right) dr \to \\ \frac{1}{2}\sigma^2 \ln p &= \gamma \left(\overline{r}r - \frac{r^2}{2}\right) + C\,, \qquad C \text{ is arbitrary.} \end{aligned}$$

Rearranging and taking exponentials of both sides to give

$$p = \exp\left(\tfrac{2\gamma}{\sigma^2}\left(\overline{r}r - \frac{r^2}{2}\right) + D\right) = E \exp\left(-\tfrac{2\gamma}{\sigma^2}\left(\frac{r^2}{2} - \overline{r}r\right)\right)$$

Complete the square to get

$$\begin{aligned} p &= E \exp\left(-\tfrac{\gamma}{\sigma^2}\left[(r - \overline{r})^2 - \overline{r}^2\right]\right) \\ p_\infty &= A \exp\left(-\tfrac{\gamma}{\sigma^2}(r' - \overline{r})^2\right). \end{aligned}$$

There is another way of performing the integration on the rhs. If we go back to $-\gamma \int (r - \overline{r})\, dr$ and write as

$$-\gamma \int \tfrac{1}{2}\tfrac{d}{dr}(r - \overline{r})^2\, dr = \frac{-\gamma}{2}(r - \overline{r})^2$$

to give

$$\frac{1}{2}\sigma^2 \ln p = \frac{-\gamma}{2}(r - \overline{r})^2 + C.$$

Now we know as p_∞ is a PDF

$$\begin{aligned} \int_{-\infty}^{\infty} p_\infty\, dr' &= 1 \to \\ A \int_{-\infty}^{\infty} \exp -\left(\tfrac{\gamma}{\sigma^2}(r' - \overline{r})^2\right)\, dr' &= 1 \end{aligned}$$

A few (related) ways to calculate A. Now use the error function, i.e.

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

So put

$$x = \sqrt{\tfrac{\gamma}{\sigma^2}}(r' - \overline{r}) \to dx = \sqrt{\tfrac{\gamma}{\sigma^2}} dr'$$

which transforms the integral above

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

therefore

$$A = \tfrac{1}{\sigma}\sqrt{\frac{\gamma}{\pi}}.$$

This allows us to finally write the steady-state transition PDF as

$$p_{\infty} = \tfrac{1}{\sigma}\sqrt{\frac{\gamma}{\pi}}\exp\left(-\tfrac{\gamma}{\sigma^2}\left(r' - \overline{r}\right)^2\right).$$

The *backward equation* is obtained in a similar way to the forward

$p\left(y, t; y', t'\right) =$

$$\begin{aligned}&\phi^{+}\left(y,t\right)p\left(y+\delta y, t+\delta t; y', t'\right)\\&+\left(1-\phi^{-}\left(y,t\right)-\phi^{+}\left(y,t\right)\right)p\left(y, t+\delta t; y', t'\right)\\&+\phi^{-}\left(y,t\right)p\left(y-\delta y, t+\delta t; y', t'\right)\end{aligned}$$

and expand using Taylor. The resulting PDE is

$$\frac{\partial p}{\partial t} + \frac{1}{2}B^2\left(y,t\right)\frac{\partial^2 p}{\partial y^2} + A\left(y,t\right)\frac{\partial p}{\partial y} = 0.$$