Diffusion processes

Olivier Scaillet

University of Geneva and Swiss Finance Institute

Outline

Introduction

In the previous chapters, we considered discrete stochastic processes.

Here, we consider continuous-time processes with continuous sample paths.

The main example of such processes is the Brownian motion.

Brownian motion

A Brownian motion $B = \{B_t; t \geq 0\}$ starting at $B_0 = b$ is a process such that

1. B has independent increments
2. $B_{t+h} - B_t$ follows a Normal distribution $N(0, \sigma^2 h)$
3. the sample paths of B are continuous.

The process B is called standard Brownian motion if $\sigma^2 = 1$ and $B_0 = 0$. It is often denoted by W and is also called Wiener process.

Distribution of increments

The increments of a Brownian motion are stationary as the distribution of $B_{t+h} - B_t$ only depends on h.

Conditionally on $B_s = x$, we have that B_t follows a Normal distribution $N(x, t-s)$, that is, $F(t,y,s,x) = P[B_t \leq y \mid B_s = x]$ has a density function

$$f(t,y,s,x) = \frac{\partial F(t,y,s,x)}{\partial y} = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{1}{2}\frac{(y-x)^2}{t-s}\right).$$

Forward and backward equations

The transition density f satisfies the following partial differential equations:

- forward equation: $\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial^2 f}{\partial y^2}$
- backward equation: $\frac{\partial f}{\partial s} = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}$

The forward equation is obtained by fixing $B_s = x$, whereas the backward equation is obtained by fixing $B_t = y$.

First passage time

The first passage time $T(x)$ of the standard Brownian motion at point x is defined by $T(x) = \inf\{t : B_t = x\}$ (the requirement of continuous sample paths is obvious for the definition to make sense).

The random variable $T(x)$ has a density function $f(t) = \frac{|x|}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right)$, $t \geq 0$.

Distribution of the maximum

We can also characterize the density of the maximum $m_t = \max\{B_s; 0 \leq s \leq t\}$.

We note that $T(m) \leq t \iff m_t \geq m$. We could prove that the random variable m_t has a density function $f(m) = \sqrt{\frac{2}{\pi t}} \exp\left(-\frac{m^2}{2t}\right)$, $m \geq 0$.

The sample paths of the Brownian motion satisfy the principles of translation and reflection.

Example: Bachelier's application

One of the first applications of the Brownian motion was proposed by Bachelier (1900). Bachelier used the Brownian motion to describe the evolution of prices on Paris stock exchange.

Bachelier assumed that the infinitesimal price increments dX_t of a financial asset are proportional to the increments dB_t of a standard Brownian motion, that is, $dX_t = \sigma dB_t$.

Starting at an initial value $X_0 = x$, the value of the process at time t is $X_t = x + \sigma B_t$.

Example: Bachelier's application (cont'd)

A major drawback of this specification is that the price has a non zero probability of getting negative.

In order to solve this issue, we rather model the relative increments with respect to the prices (returns) as a standard Brownian motion, that is, $\frac{dX_t}{X_t} = \sigma dB_t$ or equivalently $dX_t = \sigma X_t dB_t$.

The second expression looks like a differential equation. However, there are two difficulties:

1. the variables in the equation are stochastic
2. the sample paths of B_t are not differentiable even though they are continuous.

Example: Bachelier's application (cont'd)

The mathematical solution to this problem was found by Itô in the 40's using a new kind of integral: the stochastic integral.

In particular, it allows to write $X_t = x + \sigma \int_0^t X_s dB_s$ where we integrate with respect to the random element B.

Stochastic Itô integral

The stochastic integral is built in a similar way as the Riemann integral. The integral is first defined on a class of piecewise constant processes and is then extended to a larger class by approximation.

There are nonetheless two major differences between Riemann and Itô integrals. The first one is the convergence type: the Riemann integral converges in $\mathbb{R}$ whereas the Itô integral is approximated by sequences of random variables that converge in L^2, the space of square-integrable random variables (finite variance).

Stochastic Itô integral (cont'd)

The second difference is the following. Riemann sums approaching the integral of a function $f : [0, T] \to \mathbb{R}$ have the form $\sum_{j=0}^{n-1} f(s_j)(t_{j+1} - t_j)$ with $0 = t_0 < t_1 < \cdots < t_n = T$ and s_j an arbitrary point in $[t_j, t_{j+1}]$ for all j.

The value of the Riemann integral does not depend on the points $s_j \in [t_j, t_{j+1}]$.

In a stochastic context, the sums have the form $I(f_n) = \sum_{j=0}^{n-1} f(s_j)(W_{t_{j+1}} - W_{t_j})$.

The limit of such approximations does depend on the choice of intermediary points $s_j \in [t_j, t_{j+1}]$. In order to solve the ambiguity, we take $s_j = t_j$ for all j. As we choose the left point of the interval, the approximations at a particular date only depend on the information known at this date, and not on future events.

Stochastic Itô integral (cont'd)

Itô's integral is denoted by $\int_0^\infty f(s)dW_s$ and is defined such that $\lim_{n\to\infty} E[\mid \int_0^\infty f(s)dW_s - I(f_n) \mid] = 0$.

The stochastic integral has the following properties:

1. linearity:
   $\int_0^t(\alpha f(u) + \beta g(u))dW_u = \alpha \int_0^t f(u)dW_u + \beta \int_0^t g(u)dW_u$
2. isometry: $E\left[\mid \int_0^t f(u)dW_u \mid^2\right] = E\left[\int_0^t \mid f(u) \mid^2 du\right]$
3. martingale property: $E\left[\int_0^t f(u)dW_u|F_s\right] = \int_0^s f(u)dW_u$

The stochastic integral is the core building block of the definition of diffusion processes.

Diffusion processes

The stochastic integral allows us to define integral equations of the form $X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s$ whose solution is called diffusion process.

We often write this equation in a compact differential form $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$ called stochastic differential equation.

- The term $\mu(t, X_t)$ is called drift
- The term $\sigma(t, X_t)$ is called volatility (or diffusion factor) and corresponds to the instantaneous standard deviation.

Applications in finance

- The geometric Brownian motion is used in Black-Scholes (1975) to model the evolution of stock prices:
  $dS_t = mS_t dt + \sigma S_t dW_t$
- The Ornstein-Uhlenbeck process is used in Vasicek (1979) to model the evolution of the short rate r_t:
  $dr_t = b(a - r_t)dt + s dW_t$
- The square root process is used in Cox-Ingersoll-Ross (1985) to model the evolution of the short rate r_t:
  $dr_t = b(a - r_t)dt + s\sqrt{r_t} dW_t$

Euler discretization

The Euler discretization

$X_{t+h} - X_t = \mu(t, X_t)h + \sigma(t, X_t)\sqrt{h}\epsilon_{t+h}$, $\epsilon_{t+h} \sim N(0, 1)$

corresponds to the discretization of

$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$ with a time step of size h.

It can be used to simulate approximate sample paths of the process by taking h small enough.

Itô's lemma

Let X_t be a diffusion process such that
$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s$, or equivalently
$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$.

We aim at characterizing the evolution of $Y_t = f(t, X_t)$ through its stochastic differential equation. In standard calculus, we would use the total differential or chain rule $dY_t = \frac{\partial f(t,X_t)}{\partial t}dt + \frac{\partial f(t,X_t)}{\partial X}dX_t$.

In stochastic calculus, we get an additional term
$\frac{1}{2}\frac{\partial^2 f(t,X_t)}{\partial X^2}\sigma^2(t, X_t)dt$ in comparison to the usual rule.

Itô's lemma (cont'd)

Itô's lemma gives the following relationship for $Y_t = f(t, X_t)$:

$$dY_t = \frac{\partial f(t,X_t)}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f(t,X_t)}{\partial X^2}\sigma^2(t, X_t)dt + \frac{\partial f(t,X_t)}{\partial X}dX_t.$$

Replacing dX_t by its definition:

$$dY_t = \left[\frac{\partial f(t,X_t)}{\partial t} + \frac{\partial f(t,X_t)}{\partial X}\mu(t, X_t) + \frac{1}{2}\frac{\partial^2 f(t,X_t)}{\partial X^2}\sigma^2(t, X_t)\right] dt + \frac{\partial f(t,X_t)}{\partial X}\sigma(t, X_t)dW_t.$$

The additional factor is a second-order factor that is not negligible in comparison to first-order factors and is due to fast oscillations of the sample path of the process.

Itô's lemma (cont'd)

Itô's formula ressembles a second-order Taylor expansion of $f(t, X_t)$ where

- we replace dX_t by its expression
- we apply the following rules: $(dt)^2 = 0$, $dtdW_t = 0$, $(dW_t)^2 = dt$.

In compact form, we get:

$$
\begin{aligned}
dY &= df \\
&= f_t' dt + f_X' dX + \frac{1}{2}\left(f_{tt}''(dt)^2 + 2f_{Xt}'' dXdt + f_{XX}''(dX)^2\right) \\
&= f_t' dt + f_X'(\mu dt + \sigma dW) + \frac{1}{2}\left(f_{XX}''\sigma^2 dt\right) \\
&= \left(f_t' + f_X'\mu + \frac{1}{2}f_{XX}''\sigma^2\right) dt + \sigma f_X' dW
\end{aligned}
$$

Multidimensional Itô's lemma

Let $X = (X^1, X^2, ...X^d)'$ be a d-dimensional diffusion process:

$$dX_t = \mu_t dt + \sigma_t dB_t$$

with $\mu = (d \times 1)$ vector, $\sigma = (d \times k)$ matrix, and $B = (B^1, B^2, ...B^k)' = (k \times 1)$ vector is a k-dimensional Brownian motion. Take $dY_t = df(t, X_t)$, then

$$dY_t = \left(f_t(t, X_t) + f_x(t, X_t)\mu_t + \frac{1}{2} tr(\sigma_t \sigma_t' f_{xx}(t, X_t)) \right) dt$$

$$+ f_x(t, X_t)\sigma_t dB_t$$

with $f_t = \frac{\partial f}{\partial t}$, f_x = row vector with partial derivatives $\frac{\partial f}{\partial x_i}$, and f_{xx} = matrix with second partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$.

Multidimensional Itô's lemma (cont'd)

More concise form:

$$dY_t = Lf(t, X_t)dt + f_x(t, X_t)\sigma_t dB_t$$

L is called the *infinitesimal* or the *Dynkin's generator*.

$$Lf(t, X_t) = f_t(t, X_t) + f_x(t, X_t)\mu_t + \frac{1}{2}tr(\sigma_t \sigma_t' f_{xx}(t, X_t))$$

Other rule:

$$dY_t = f_t(t, X_t)dt + \sum_i f_{x_i}(t, X_t)dX_t^i + \frac{1}{2}\sum_{i,j} f_{x_i,x_j} dX_t^i dX_t^j$$

and product $dX_t^i dX_t^j$ computed with the *conventions* $dtdt = 0$, $dtdB_t^i = 0$, $dB_t^i dB_t^i = dt$ and $dB_t^i dB_t^j = 0$ if $i \neq j$.

Integration by parts (example)

Let

$$\begin{cases} dX_t^1 = \mu_t^1 dt + \sigma_t^1 dB_t \\ dX_t^2 = \mu_t^2 dt + \sigma_t^2 dB_t \end{cases}$$

Show the integration by parts rule

$$\int_0^t X_s^1 dX_s^2 = X_t^1 X_t^2 - X_0^1 X_0^2 - \int_0^t X_s^2 dX_s^1 - \int_0^t \sigma_s^1 \sigma_s^2 ds$$

and note that it differs from standard differentiation rule $d(uv) = du\, v + u\, dv$ and standard integration by parts rule $uv = \int du\, v + \int u\, dv$.

Integration by parts (example)

Take $Y_t = X_t^1 X_t^2$ and apply Itô's lemma with $f = x_1 x_2$.
Hence: $f_t = 0$, $f_x = (x_2, x_1)'$, $f_{xx} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
Since $d = 2$ and $k = 1$, we get $\sigma_t = (\sigma_t^1, \sigma_t^2)'$ which gives the (2×2) matrix $\sigma_t \sigma_t' = \begin{pmatrix} (\sigma_t^1)^2 & \sigma_t^1 \sigma_t^2 \\ \sigma_t^1 \sigma_t^2 & (\sigma_t^2)^2 \end{pmatrix}$

Hence: $\sigma_t \sigma_t' f_{xx} = \begin{pmatrix} \sigma_t^1 \sigma_t^2 & (\sigma_t^1)^2 \\ (\sigma_t^2)^2 & \sigma_t^1 \sigma_t^2 \end{pmatrix}$ and $tr(\sigma_t \sigma_t' f_{xx}) = 2\sigma_t^1 \sigma_t^2$ and

we get $dY_t = (X_t^2 \mu_t^1 + X_t^1 \mu_t^2 + \sigma_t^1 \sigma_t^2)dt + (X_t^2 \sigma_t^1 + X_t^1 \sigma_t^2)dB_t$
or $d(X_t^1 X_t^2) = X_t^2 dX_t^1 + X_t^1 dX_t^2 + \sigma_t^1 \sigma_t^2 dt$ Q.E.D.

Black-Scholes model

The Black-Scholes formula is a European option valuation formula.

The Black-Scholes model considers an economy with two asset classes:

- a riskless asset, T-bill or zero coupon bond
- a risky asset, stock

The value of the riskless asset M_t grows at a continuously compounded constant interest rate r and is normalized such that $M_0 = 1$, that is, $dM_t = rM_t dt$ with solution $M_t = e^{rt}$.

The value of the risky asset S_t follows a geometric Brownian motion, that is, $dS_t = mS_t dt + sS_t dW_t$ with solution $S_t = S_0 \exp\left((m - \frac{1}{2}s^2)t + sW_t\right)$.

Portfolio

Let $F = \{F_t; t \geq 0\}$ be the filtration generated by the observation of the price history, $F_t = \{S_u; 0 \leq u \leq t\}$.

A portfolio is a pair $\alpha = \{\alpha_t; t \geq 0\}$ and $\beta = \{\beta_t; t \geq 0\}$ of F-adapted random processes. The pair (α, β) is a time-dependent portfolio containing α_t units of stock and β_t units of bond.

The value of the (α, β) portfolio at time t is given by the value function $V_t = \alpha_t S_t + \beta_t M_t$.

Valuation by replication

The portfolio is said to be self-financing if $dV_t = \alpha_t dS_t + \beta_t dM_t$.

It means that the changes in the value of the portfolio are only due to variations of the prices: there is no addition or withdrawal of funds.

The self-financing portfolio (α, β) replicates the payoff function of a European call option if its value is equal to the value of the call at maturity, that is, $V_T = (S_T - K)_+$.

By absence of arbitrage (AOA), we then have that the replicating portfolio value must be equal to the call price.

Valuation by replication (cont'd)

Let us consider the value function V_t of the portfolio as a function of time and the stock price
$V_t = V(t, S_t) = \alpha(t, S_t)S_t + \beta(t, S_t)e^{rt}$.

Using Itô's lemma, we get

$$dV_t = \frac{\partial V(t,S_t)}{\partial S} dS_t + \left[\frac{\partial V(t,S_t)}{\partial t} + \frac{1}{2}s^2 S_t^2 \frac{\partial^2 V(t,S_t)}{\partial S^2}\right] dt.$$

Also, the self-financing condition gives
$dV_t = \alpha(t, S_t)dS_t + \beta(t, S_t)re^{rt}dt$.

Valuation by replication (cont'd)

By identification of the factors of dS_t and the factors of dt and using the self-financing constraint, we get that

$$\frac{\partial V(t,S_t)}{\partial S} = \alpha(t,S_t)$$

$$\left[\frac{\partial V(t,S_t)}{\partial t} + \frac{1}{2}s^2 S_t^2 \frac{\partial^2 V(t,S_t)}{\partial S^2}\right] = \beta(t,S_t) r e^{rt}.$$

Multiplying the first equation by rS_t and adding it to the second equation, we obtain

$$\left[\frac{\partial V(t,S_t)}{\partial t} + \frac{1}{2}s^2 S_t^2 \frac{\partial^2 V(t,S_t)}{\partial S^2}\right] + \frac{\partial V(t,S_t)}{\partial S} rS_t \\ = \alpha(t,S_t) rS_t + \beta(t,S_t) r e^{rt}$$

which, by using $\alpha(t,S_t) rS_t + \beta(t,S_t) r e^{rt} = rV(t,S_t)$, leads to the the pricing equation

$$\frac{\partial V(t,S_t)}{\partial t} + \frac{1}{2}s^2 S_t^2 \frac{\partial^2 V(t,S_t)}{\partial S^2} + \frac{\partial V(t,S_t)}{\partial S} rS_t - rV(t,S_t) = 0 .$$

Black-Scholes formula

We get the Black-Scholes formula by solving this parabolic differential equation, using the payoff function of the option as a terminal condition $V(T, S_T) = (S_T - K)_+$:

$$V(t, S_t) = S_t \Phi\left(d_1(T-t, S_t)\right) - Ke^{-r(T-t)}\Phi\left(d_2(T-t, S_t)\right)$$

$$d_1(T-t, S_t) = \frac{\ln(S_t/K) + \left(r + \frac{1}{2}s^2\right)(T-t)}{s\sqrt{T-t}}$$

$$d_2(T-t, S_t) = d_1(T-t, S_t) - s\sqrt{T-t}$$

where $\Phi(x)$ is cdf of the standard normal distribution.

Depending on the derivative product and the assumptions, it is not always as easy to determine an explicit solution, however.

Volatility estimation

The formula depends on the volatility s of the stock returns that we estimate

- either using historical returns (historical volatility)
- or using option prices and the inverse Black-Scholes formula (implied volatility).

Equity-linked life insurance

Concepts used in the Black-Scholes model can be applied to life insurance models.

For usual life insurance contracts, a so-called technical interest is paid.

Insurance companies usually invest part of their reserves in financial markets.

The expected return is higher than the riskless rate but there is an additional risk.

Equity-linked life insurances transfer part of this risk to the owner of the policy.

Equity-linked life insurance (cont'd)

The interest rate is higher than for a standard contract, but it is random.

Suppose that the payoff function of the life insurance depends on a reference index (e.g., the price of a financial asset, portfolio, or market index).

We consider here a contract paying the maximum between some index value and a guaranteed amount b. This approach reduces the risk for the insurer.

Equity-linked life insurance (cont'd)

Let the index value X_t follow a geometric Brownian motion $dX_t = mX_t dt + sX_t dW_t$.

The payment takes place at a random time T that is triggered if the insured dies (or in case of any other event defined in the contract such as a job loss, a divorce, a wedding).

The payoff at time T is $\max(X_T, b) = X_T \vee b$.

Types of life insurances

We consider two types of contracts based on a fixed maturity t_0:

- Term insurance, for which the value $X_T \vee b$ is paid at the time of death $T \in [0, t_0]$.
- Pure endowment insurance, for which the value $X_{t_0} \vee b$ is paid if the insured is still alive at time t_0.

In general, we combine the two types of life insurances, possibly with different amounts b.

Model

Suppose that $T = T_a$ corresponds to the remaining lifetime of an a-years old insured and that it is independent of the index value X.

The probability that the insured does not die in the t years to come is given by $1 - F_{T_a}(t) = P[T_a > t]$.

Also, $F_{T_a}(t) = P[T_a < t]$ gives the probability that the insured dies during the t years to come.

The density of T_a at u is denoted by $f_{T_a}(u)$.

Model (cont'd)

The distribution of T_a depends on the age, the gender, the country, the health, etc.

There are thus two sources of risk in an equity-linked life insurance:

- the financial risk due to the evolution of the index
- the mortality risk of the insured.

These two sources are modelled with X and T_a and have an impact on the insurance payoff.

Model (cont'd)

Note that the payoff function $X_{T_a} \vee b$ can be rewritten as $b + \max(X_{T_a} - b, 0)$, which is equivalent to the payment of a fixed amount b plus a payment $\max(X_{T_a} - b, 0)$.

For a fixed T_a, we can use the Black-Scholes formula to value a payoff $\max(X_{T_a} - b, 0)$ as it is equivalent to a European call with maturity T_a on the index X and with strike price b.

Model (cont'd)

For a maturity $T_a = u$, the price today is

$$C(u, X_0) = X_0 \Phi\left(d_1(u, X_0)\right) - be^{-ru} \Phi\left(d_2(u, X_0)\right)$$

$$d_1(u, X_0) = \frac{\ln(X/b) + \left(r + \frac{1}{2}s^2\right) u}{s\sqrt{u}}$$

$$d_2(u, X_0) = d_1(u, X_0) - s\sqrt{u}$$

where $\Phi(x)$ is cdf of the standard normal distribution.

In order to compute the insurance premium, we finally weight these prices according to the probability of each possible maturity.

Premium of pure endowment insurance

The payoff of the pure endowment insurance is given by $X_{t_0} \vee b$ at time t_0 if the insured is still alive.

The probability to be alive at time t_0 is given by $P[T_a > t_0] = 1 - F_{T_a}(t_0)$.

The premium today is denoted by Π_0^e and is given by the price today of the unique random payoff at time t_0, that is, the sum of the discounted fixed amount b and the call price $be^{-rt_0} + C(t_0, X_0)$ multiplied by the survival probability $1 - F_{T_a}(t_0)$:
$\Pi_0^e = (be^{-rt_0} + C(t_0, X_0))\,(1 - F_{T_a}(t_0))$.

Premium of term insurance

For the term insurance, the payoff may happen at any time between $[0, t_0]$ if the insured dies.

If we knew that the death would happen at time $T_a = u$, the value today of the security would be $be^{-ru} + C(u, X_0)$.

The premium today is denoted by Π_0^t and is given by the sum of these values weighted by the density $f_{T_a}(u)$ on all possible dates u between 0 and t_0:
$\Pi_0^e = \int_0^{t_0} \left(be^{-ru} + C(u, X_0)\right) f_{T_a}(u) du.$

Merton's intertemporal model

Merton's intertemporal model considers a problem of optimal consumption and portfolio allocation in continuous time.

We aim at maximizing an expected utility by optimizing consumption and investment across time, and having an initial wealth w.

We suppose that there are N financial assets, whose prices are defined by
$dS_t^i = m_i S_t^i dt + s_i S_t^i \sum_{j=1}^{d} dW_t^j$.

The evolution of each financial asset depends on d sources of noise W_t^j, $j = 1, \ldots, d$.

Merton's intertemporal model (cont'd)

We define the price of the riskless asset by $dM_t = rM_t dt$ with $M_0 = 1$.

We may invest in N risky assets and a riskless asset, which gives the multidimensional price process $X = (M, S^1, \dots, S^N)$.

Let c_t be the consumption level at time t and Z the final wealth. We aim at maximizing
$U(c, Z) = E\left[\int_0^t u(c_t, t)dt + F(Z)\right]$,
where u and F stand for the utility of consumption and final wealth, respectively.

Merton's intertemporal model (cont'd)

We use financial assets to transfer wealth from one date to the other and to finance future consumption.

An investment strategy is given by the process $\theta = (\theta^0, \theta^1, \ldots, \theta^N)$. Such a strategy is said to finance the consumption plan $c = \{c_t; t \in [0, T]\}$ and the final wealth Z if it satisfies $\sum_{i=1}^N \theta_t^i S_t^i = w + \int_0^t \sum_{i=1}^N \theta_s^i dS_s^i - \int_0^t c_s ds \geq 0$ for $t \in [0, T]$ and $\sum_{i=1}^N \theta_T^i S_T^i = Z$.

For some initial wealth w, we determine the optimal c, z and θ.

Merton's intertemporal model(cont'd)

We may obtain explicit solutions for some particular utility functions (for inst. $u(w) = w^\alpha/\alpha$) using stochastic optimal control techniques.

Those kind of models are used in pension funds where we invest the received premia in financial assets in order to finance future retirement.