STOCHASTIC PROCESSES, SDES, AND OPTION PRICING

1. Measure Theory in 5 Minutes

Definition 1. Ω is a set called the sample/event space. A σ-algebra $\mathcal{F}$ is a family $\mathcal{F}$ of subsets of Ω with the following properties

(1) $\emptyset \in \mathcal{F}$
(2) $F \in \mathcal{F} \Rightarrow F^C \in \mathcal{F}$ i.e. $\mathcal{F}$ is closed under complementation.
(3) $A_i \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ i.e. $\mathcal{F}$ is closed under countable unions.

Definition 2. The pair $(\Omega, \mathcal{F})$ is called a measurable space and the subsets F of Ω which belond to $\mathcal{F}$ are called $\mathcal{F}$-measurable sets.

Definition 3. A *measure* is a function function $\mu : \mathcal{F} \to \mathbb{R}_+$ such that $\mu(\emptyset) = 0$ and μ is countably additive

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} E_i$$

. It "weighs" sets. The triple $(\Omega, \mathcal{F}, \mu)$ is appropriately called a *measure space*. If $\mu(\Omega) < \infty$ then $(\Omega, \mathcal{F}, \mu)$ is called a *finite-measure space*.

Definition 4. Let $(\Omega, \mathcal{F}, \mu)$ and $(\Omega', \mathcal{F}', \mu')$ be measure spaces. Then a function $X : \Omega \to \Omega'$ is $\mathcal{F} - \mathcal{F}'$*-measurable* if for every $U \in \mathcal{F}'$

$$X^{-1}(U) = \{\omega \in \Omega; X(\omega) \in U\} \in \mathcal{F}$$

I.e. pre-images of measurable sets in the σ*-algebra* associated with the codomain are measurable sets in the σ*-algebra* associated with the domain.

Talk a little about defining the Lebesgue integral in terms of simple functions and approximating continuous functions by simple functions.

Definition 5. The *abstract Lebesgue integral* of an $\mathcal{F} - \mathcal{F}'$-measurable function $X : \Omega \to \Omega'$ on the set $A \in \mathcal{F}'$ is denoted

$$\int_A X d\mu(\omega)$$

and roughly corresponds to inverse of the typical understanding of the Riemann integral: fix a value $X = x$ for $x \in A$, find the inverse image $X^{-1}(x) = \{\omega : X(\omega) = x\}$, and find the measure/mass of that set $\mu\left(X^{-1}(x)\right) = \mu(\{\omega : X(\omega) = x\})$ under the measure on the domain σ-algebra $\mathcal{F}$, i.e. μ. The *Lebesgue integral* is an abstract Lebesgue integral with $\mu = \lambda$ the Lebesgue measure[1] on $\mathbb{R}$.

Definition 6. If $\mu(\Omega) = 1$ then $(\Omega, \mathcal{F}, \mu)$ is called a *probability space* and convention is to write μ as P. Furthermore P should be such that for A_i such that $A_i \cap A_j = \emptyset$ for $i \neq j$ it's the case that $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{n} P(A_i)$, called countable additivity.

Definition 7. For $(\Omega, \mathcal{F}, P)$ and $(\mathbb{R}^n, \mathcal{B}^n)$ a *random variable* is an $\mathcal{F}$*-measurable* function $X : \Omega \to \mathbb{R}^n$.

Every random variable induces a probability measure μ_X on $\mathbb{R}^n$, defined by

$$\mu_X(A) = P\left(X^{-1}(A)\right)$$

where $B \in \mathcal{B}^n$ and $\mathcal{B}^n$ is the σ-algebra over $\mathbb{R}^n$. μ_X is called the *distribution* or *law of* X.

2. Stochastic Processes

Definition 8. A stochastic process is a parameterized collection of random variables $\{X_t\}_{t \in T}$ with each X_t defined on the same measure space (probability space) $(\Omega, \mathcal{F}, P)$ and $X_t : \Omega \to \mathbb{R}^n$.

The parameter space for T and the range of X_t determine taxonomical classifiction of the process with each possibly being either continuous or discrete.

For fixed t it's the case that $X_t(\omega)$ is just a plain random variable

$$X_t(\omega) : \Omega \to \mathbb{R}^n$$

and for fixed ω it's the case that $X_\omega(t)$ is just a function on $\mathbb{R}^n$

$$X_\omega(t) : T \to \mathbb{R}^n$$

called a *path*. Note that this gives a way to identify a function with each $\omega \in \Omega$ and therefore we may treat Ω as a subset of $\tilde{\Omega} \subset (\mathbb{R}^n)^T$ the space of *all* functions from T to $\mathbb{R}^n$.

[1] The unique translation invariant measure that assigns $\lambda([a, b)) = b - a$.

Definition 9. The *finite-dimensional distributions* $\{\mu_{t_1,\dots,t_k}; k \in \mathbb{N}, t_i \in T\}$ of the process $X = \{X_t\}_{t\in T}$ are the multidimensional distributions of the random vectors $(X_{t_1}, \dots, X_{t_k})$ for any set $\{t_1, \dots, t_k\}$.

These distributions determine many, but not all (and in some important cases crucially), properties the process X. Conversely given a family of distributions $\{\mu_{t_1,\dots,t_k}; k \in \mathbb{N}, t_i \in T\}$(all finite product distributions indexed by all finite subsets of T) on $\underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k}$ one can construct a stochastic process $X = \{X_t\}_{t\in T}$ which will have $\{\mu_{t_1,\dots,t_k}; k \in \mathbb{N}, t_i \in T\}$ as its finite-dimensional distributions given some "natural" consistency conditions.

Theorem 10. *Kolmogorov's extension theorem. For all finite tuples* $(t_1, \dots, t_k)$ *with* $t_i \in T$ *let* $\mu_{t_1,\dots,t_k}$ *be the probability measures defined above. If* $\mu_{t_1,\dots,t_k}$ *satisfy the consistency conditions*

(1) *Invariance under permutation: for all permutations* p *on* $\{1, 2, \dots, k\}$

$$\mu_{p(t_1),\dots,p(t_k)}\left(F_1 \times \cdots \times F_k\right) = \mu_{t_1,\dots,t_k}\left(F_{p^{-1}(1)} \times \cdots \times F_{p^{-1}(k)}\right)$$

(2) *Invariance under marginalization over probability 1 events:*

$$\mu_{t_1,\dots,t_k}\left(F_1 \times \cdots \times F_k\right) = \mu_{t_1,\dots,t_k,t_{k+1},\dots,t_{k+m}}\left(F_1 \times \cdots \times F_k \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{m}\right)$$

then there exists a probability space $(\Omega, \mathcal{F}, P)$ *and a stochastic process* $X = \{X_t\}_{t\in T}$ *on* Ω *with* $\mu_{t_1,\dots,t_k}$ *as its finite-dimensional distributions.*

Kolmogorov's extension is actually an if and only if: any stochastic process trivially has these properties. The power of Kolmogorov extension is that the typical workflow is to define a stochastic process by its finite-dimensional distributions and that is completely sufficient[2].

3. Brownian Motion

Definition 11. Fix $x \in \mathbb{R}$ and define

$$p(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}$$

for $y \in \mathbb{R}$. Then for $0 < t_1 < t_2 < \cdots < t_n$ define a measure $\mu_{t_1,\dots,t_k}$ on $\mathbb{R}^k$ by

$$\begin{aligned}\mu_{t_1,\dots,t_k}\left(F_1 \times \cdots \times F_k\right) &= \int_{F_1} \cdots \int_{F_k} p(t_1, x, x_1)\, p(t_2 - t_1, x_1, x_2) \cdots \\ &\quad p(t_k - t_{k-1}, x_{k-1}, x_k)\, dx_1 dx_2 \cdots dx_k\end{aligned}$$

for $F_i \subset \mathbb{R}$ with the convention that $p(0, x, y)\, dx = \delta_x(y)$ a unit point mass at x. Extend to all finite sequences $t_1, \dots, t_n$ and since $\int_{F_i} p(t, x, y)\, dx = 1$ Kolmogorov's Extension theorem says there exists $(\Omega, \mathcal{F}, P^x)$ and a stochastic process $B = \{B_t\}_{t\geq 0}$ on Ω such that the finite-dimensional distributions of B_t are

$$P^x\left(B_{t_1} \in F_1 \times \cdots \times B_{t_k} \in F_k\right) = \int_{F_1} \cdots \int_{F_k} p(t_1, x, x_1)\, p(t_2 - t_1, x_1, x_2) \cdots p(t_k - t_{k-1}, x_{k-1}, x_k)\, dx_1 dx_2 \cdots dx_k \tag{3.1}$$

Such a process[3] is called a Brownian motion. The paths of a Brownian motion happen to be almost surely continuous[4] (but nowhere differentiable!) so with each $\omega \in \Omega$ you can identify a continuous function $t \to B_t(\omega)$ from $[0, \infty)$ to $\mathbb{R}$. Thus a perspective on Brownian motion is that it's the space of continuous functions $C([0,\infty), \mathbb{R})$ equipped with the measures P^x above[5].

Here are some of the basic properties of Brownian motion as defined by defn. 3.1:

(1) Fix the range of B_t to be $\mathbb{R}$. For $x = x_0$ B_t is a *Gaussian process,* i.e. for all $0 < t_1 < t_2 < \cdots < t_k$ the random vector $Z = (B_{t_1}, \dots, B_{t_k}) \in (\mathbb{R})^k$ is distributed multinormal. That means there exists a mean vector $\mathbf{m} \in (\mathbb{R})^k$ and covariance matrix $\Sigma = [c_{jm}] \in \mathbb{R}^{k\times k}$ such that

$$\begin{aligned}E^x\left[\exp\left[i\sum_{j=1}^{k} u_j B_{t_j}\right]\right] &:= \int \exp\left[i\sum_{j=1}^{k} u_j B_{t_j}\right] dP^x \\ &= \frac{1}{\sqrt[k]{2\pi t_1 (t_2 - t_1) \cdots (t_k - t_{k-1})}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \exp\left[i\sum_{j=1}^{k} u_j B_{t_j}\right] e^{-\frac{(x_1 - x_0)^2}{2t_1}} e^{-\frac{(x_1 - x_2)^2}{2(t_2 - t_1)}} \cdots \\ &\quad e^{-\frac{(x_{k-1} - x_k)^2}{2(t_k - t_{k-1})}} dx_1 dx_2 \cdots dx_k\end{aligned}$$

---

[2] Kolmogorov extension guarnatees there is a correspondent probability space and process.

[3] One that has these finite-dimensional distributions.

[4] Wiener's theorem.

[5] Which as discussed are product measures on a function space.

Therefore in general[6] the finite dimensional distributions of a Brownian motion are multivariate Gaussians with mean vector $\mathbf{m} = \mathbf{x_0}$ and covariance matrix

$$\mathbf{\Sigma} = \begin{pmatrix} t_1 I_n & t_1 I_n & \cdots & t_1 I_n \\ t_1 I_n & t_2 I_n & \cdots & t_2 I_n \\ \vdots & \vdots & & \vdots \\ t_1 I_n & t_2 I_n & \cdots & t_k I_n \end{pmatrix}$$

where n is the dimension of each B_{t_j}. Therefore, since each component of a multivariate Gaussian is itself Gaussian with means and covariances being a function of $\mathbf{m}$ and $\mathbf{\Sigma}$ it's the case that

$$E^x \left[B_{t_j} \right] = \mathbf{m}_{t_j} = x_0$$

and

$$\mathrm{Var}\left(B_{t_j}\right) = E^x \left[\left(B_{t_j} - x_0\right)^2 \right] = t_j \mathrm{Tr}\left(I_n\right) = n t_j$$

and because of the cascading[7] shape of the covariance matrix

$$\mathrm{Cov}\left(B_{t_j}, B_{t_i}\right) = E^x \left[\left(B_{t_j} - x_0\right)\left(B_{t_i} - x_0\right) \right] = \mathrm{Tr}\left(I_n\right) \min\left(t_j, t_i\right) = n \min\left(t_j, t_i\right)$$

and

$$\begin{aligned} E^x \left[\left(B_{t_j} - B_{t_i}\right)^2 \right] &= E^x \left[\left(B_{t_j} - x_0\right)^2 - 2\left(B_{t_j} - x_0\right)\left(B_{t_i} - x_0\right) + \left(B_{t_i} - x_0\right)^2 \right] \\ &= \mathrm{Tr}\left(I_n\right)\left(t_j - 2\min\left(t_j, t_i\right) + t_i\right) \\ &= \begin{cases} n\left(t_j - t_i\right) & \text{if } t_j \geq t_i \\ n\left(t_i - t_j\right) & \text{otherwise} \end{cases} \quad (3.2) \\ &= n\left|t_i - t_j\right| \quad (3.3) \end{aligned}$$

(2) B_t has independent increments, i.e. for $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k$

$$B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_k} - B_{t_{k-1}}$$

are independent random variables. This follows from the fact that normal RVs are independent if their covariance is zero and for $t_i < t_j$

$$\begin{aligned} E^x \left[\left(B_{t_j} - B_{t_{j-1}}\right)\left(B_{t_i} - B_{t_{i-1}}\right) \right] &= E^x \left[\left(B_{t_j} - B_{t_{j-1}}\right)\left(B_{t_i} - B_{t_{i-1}}\right) \right] \\ &= E^x \left[B_{t_j} B_{t_i} - B_{t_j} B_{t_{i-1}} - B_{t_{j-1}} B_{t_i} + B_{t_{j-1}} B_{t_{i-1}} \right] \\ (3.4) \quad &= E^x \left[\left(B_{t_j} - x_0\right)\left(B_{t_i} - x_0\right) - \left(B_{t_j} - x_0\right)\left(B_{t_{i-1}} - x_0\right) - \left(B_{t_{j-1}} - x_0\right)\left(B_{t_i} - x_0\right) + \left(B_{t_{j-1}} - x_0\right) (\right. \\ &= n\left(t_i - t_{i-1} - t_i + t_{i-1}\right) = 0 \end{aligned}$$

(3) B_t *is* almost surely continuous but we need a little theory to prove it:

Definition. Suppose X_t and Y_t are stochastic processes on $(\Omega, \mathcal{F}, P)$ then X_t is a *modification* of Y_t if for all t

$$P\left(\{\omega; X_t(\omega) = Y_t(\omega)\}\right) = 1$$

Note that X_t and Y_t have the same law[8] and as such are essentially the same, but might have different path properties.

Theorem 12. ***Kolmogorov's continuity theorem***. *Suppse for the process X_t there exist α, β, D for all T*

$$E\left[\left|X_t - X_s\right|^{\alpha}\right] \leq D\left|t - s\right|^{1+\beta} \quad \textit{for all } 0 \leq s, t \leq T$$

Corollary 13. *With $\alpha = 4, \beta = 1, D = n(n+2)$ Brownian motion X_t satisfies the criterion for Kolmogorov's continuity theorem and therefore there exists a version Y_t with continuous paths.*

4. Ito Integral

Return to the original problem of finding a reasonable interpretation of

$$\frac{dx}{dt} = b(t, x) + \sigma(t, x) \cdot \text{"noise"}$$

It turns out it's reasonable to model "noise" as some stochastic process W_t and so

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot B_t \quad (4.1)$$

This will be shorthand for

$$X_k = X_0 + \int_0^t b(s, X_s)\, ds + \text{"} \int_0^t \sigma(s, X_s)\, dB_s \text{"} \quad (4.2)$$

---

[6] With $B_{t_j} = \mathbf{B}_{t_j} \in \mathbb{R}^n$

[7] Just think about running your finger down and across the covariance matrix.

[8] The same finite-dimensional distributions

as soon as I define " $\int_0^t \sigma(s, X_s)\, dB_s$".

Definition 14. Let $0 \leq Q < T$ and start by defining

$$\int_Q^T \{\cdot\}\, dB_s(\omega)$$

for simple processes

$$S_n(t, \omega) = \sum_{j=0}^{\infty} a_j(\omega)\, 1_{[j \cdot 2^{-n}, (j+1) \cdot 2^{-n})}(t)$$

You can imagine that S_n is defined for all $t \in [0, \infty)$ by defining the $a_j = 0$ appropriately. So basically chop the real line into intervals of length 2^{-n} and define $S_t(\omega)$ piecewise constant on that mesh. Then define

$$\int_Q^T S_n(s, \omega)\, dB_s := \sum_{j=0}^{\infty} a_j(\omega) \left[B_{s_{j+1}^{(n)}}(\omega) - B_{s_j^{(n)}}(\omega) \right]$$

where

$$s_j^{(n)} := \begin{cases} \frac{j}{2^n} & \text{if } Q \leq j \cdot 2^{-n} \leq T \\ Q & \text{if } j \cdot 2^{-n} < Q \\ T & \text{if } j \cdot 2^{-n} > T \end{cases}$$

which just truncates the sum outside of $[Q, T]$ since for $s_j^{(n)} > T$

$$B_{s_{j+1}^{(n)}}(\omega) - B_{s_j^{(n)}}(\omega) = B_T(\omega) - B_T(\omega) = 0$$

and $T - Q = m \cdot 2^{-n}$ for some m and "around the edges" the error becomes neglibile as $n \to \infty$.

Definition 15. Let $f \in \mathcal{V}(Q, T)$. Choose $S_n \in \mathcal{V}$ according to the approximation such that

$$\lim_{n \to \infty} E\left[\int_Q^T (f - S_n)^2\, ds \right] = \lim_{n \to \infty} \|f - S_n\| = 0$$

Then the *Ito integral* is defined

$$\mathcal{I}[f](\omega) := \int_Q^T f(s, \omega)\, dB_s := \lim_{n \to \infty} \int_Q^T S_n(s, \omega)\, dB_s$$

Alright finally we can actually do an ito integral!

Example 16. Assume $B_0 = 0$. Then

$$\int_0^t B_s dB_s = \frac{1}{2} B_t^2 - \frac{t}{2}$$

Proof. Let $t_j^{(n)} = j \cdot 2^{-n}$ and put $S_n(t, \omega) = \sum_{j=0}^{\infty} B_{t_j^{(n)}}(\omega)\, 1_{\left[t_j^{(n)}, t_{j+1}^{(n)}\right)}(t)$. Then we need to prove convergence in ito norm:

$$E\left[\int_0^t (S_n - B_s)^2\, ds \right] = E\left[\sum_{j=0}^{\infty} \int_{t_j^{(n)}}^{t_{j+1}^{(n)}} \left(B_{t_j^{(n)}} - B_s \right)^2 ds \right] \tag{4.3}$$

$$= \sum_{j=0}^{\infty} \int_{t_j^{(n)}}^{t_{j+1}^{(n)}} E\left[\left(B_{t_j^{(n)}} - B_s \right)^2 \right] ds \tag{4.4}$$

$$= \sum_{j=0}^{\infty} \int_{t_j^{(n)}}^{t_{j+1}^{(n)}} \left(s - t_j^{(n)} \right) ds = \frac{1}{2} \sum_{j=0}^{\infty} \left(t_{j+1}^{(n)} - t_j^{(n)} \right)^2 \tag{4.5}$$

where in line 4.4 we use the indicator in the definiton of S_n and in line 4.6 we use that since the limits of integration for s are $t_j^{(n)}$ and $t_{j+1}^{(n)}$ it's the case that $s > t_j^{(n)}$. Finally as the mesh is refined $\frac{1}{2} \sum_{j=0}^{\infty} \left(t_{j+1}^{(n)} - t_j^{(n)} \right)^2 \to 0$. So

$$\int_0^t B_s dB_s = \lim_{n \to \infty} \int_0^t S_n dB_s = \lim_{n \to \infty} \sum_{j=0}^{\infty} B_{t_j^{(n)}} \left(B_{t_{j+1}^{(n)}} - B_{t_j^{(n)}} \right)$$

□

Proof. Let $\Delta B_{t_j^{(n)}} := B_{t_{j+1}^{(n)}} - B_{t_j^{(n)}}$ and $\Delta\left(B_{t_j^{(n)}}^2 \right) = B_{t_{j+1}^{(n)}}^2 - B_{t_j^{(n)}}^2$. Then

$$\int_0^t B_s dB_s = \lim_{n \to \infty} \sum_{j=0}^{\infty} B_{t_j^{(n)}} \Delta B_{t_j^{(n)}}$$

So we have to evaluate this limit. First

$$\begin{aligned}\Delta\left(B^2_{t^{(n)}_j}\right) &= \left(B_{t^{(n)}_{j+1}} - B_{t^{(n)}_j}\right)^2 + 2B_{t^{(n)}_j}\left(B_{t^{(n)}_{j+1}} - B_{t^{(n)}_j}\right) \\ &= \left(\Delta B_{t^{(n)}_j}\right)^2 + 2B_{t^{(n)}_j}\Delta B_{t^{(n)}_j}\end{aligned}$$

The last term in this derivation is what we're interested in because it appears in out limit. Now since B_0

$$B_t^2 = \sum_{j=0}^{\infty}\Delta\left(B^2_{t^{(n)}_j}\right) = \sum_{j=0}^{\infty}\left(\Delta B_{t^{(n)}_j}\right)^2 + 2\sum_{j=0}^{\infty}B_{t^{(n)}_j}\Delta B_{t^{(n)}_j}$$

or

$$\sum_{j=0}^{\infty}B_{t^{(n)}_j}\Delta B_{t^{(n)}_j} = \frac{1}{2}B_t^2 - \frac{1}{2}\sum_{j=0}^{\infty}\left(\Delta B_{t^{(n)}_j}\right)^2$$

Finally by unbounded variation (thm **??**)

$$\lim_{n\to\infty}\sum_{j=0}^{\infty}\left(\Delta B_{t^{(n)}_j}\right)^2 = t$$

and so

$$\lim_{n\to\infty}\sum_{j=0}^{\infty}B_{t^{(n)}_j}\Delta B_{t^{(n)}_j} = \frac{1}{2}B_t^2 - \frac{1}{2}t$$

□

Note the extra $\frac{1}{2}t$ term, exhibiting the difference from the standard integration rule of $\int x = \frac{1}{2}x^2$.

5. Ito's Formula

To derive a calculus of stochastic integrals we take a counterintuitive approach. Note that

$$B_t = \int_0^t dB_s$$

and recall that example 21 shows

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$$

or

$$\frac{1}{2}B_t^2 = \int_0^t B_s dB_s - \frac{1}{2}t$$

So the function $g(x) = \frac{1}{2}x^2$ does not map the Ito integral $x = B_t = \int_0^t dB_s$ into another Ito integral; in fact it's the combination of two integrals

$$\frac{1}{2}B_t^2 = \int_0^t B_s dB_s - \int_0^t \frac{1}{2}ds \tag{5.1}$$

Hence define the class of *Ito processes*

Definition 17. Let B_t be a 1-D Brownian motion on $(\Omega, \mathcal{F}, P)$. An Ito process (or *stochastic integral*) is a stochastic process X_t on $(\Omega, \mathcal{F}, P)$ of the form

$$X_t = X_0 + \int_0^t u(s,\omega)\,ds + \int_0^t v(s,\omega)\,dB_s \tag{5.2}$$

where $v \in \mathcal{V}(0,T)$ and

$$P\left(\int_0^t v^2 ds < \infty \text{ for } t \geq 0\right) = 1$$

and

$$P\left(\int_0^t |u|\,ds < \infty \text{ for } t \geq 0\right) = 1$$

This class of processes is closed under smooth maps. Eqn. 5.2 is more typically written in "differential form"

$$dX_t = udt + vdB_t \tag{5.3}$$

For example eqn. 5.1 is written

$$d\left(\frac{1}{2}B_t^2\right) = \frac{1}{2}dt + B_t dB_t$$

The main tool in stochastic calculus is the Ito formula

Theorem 18. *(Ito formula) Let X_t be an ito process and $g(t,x) \in C^2([0,\infty) \times \mathbb{R})$ then*

$$Y_t = g(t, X_t)$$

is again an ito process and

$$dY_t = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \cdot (dX_t)^2 \tag{5.4}$$

where $(dX_t)^2 = (dX_t) \cdot (dX_t)$ is computed according to

$$dt \cdot dt = dt \cdot dB_t = dB_t \cdot dt = 0 \quad dB_t \cdot dB_t = dt$$

6. SDEs

Finally let's solve some god damn stochastic differential equations!

Example 19. The population growth model from Ch. 1 is

$$\frac{dN_t}{dt} = a_t N_t = (r_t + \alpha W_t) N_t$$

where W_t is white-noise. The Ito interpretation of this model is

$$dN_t = (r_t N_t dt + \alpha N_t dB_t)$$

or

$$\frac{dN_t}{N_t} = r_t dt + \alpha dB_t \tag{6.1}$$

or

$$\int_0^t \frac{dN_s}{N_s} = rt + \alpha B_t$$

To evaluate the left-hand side use the Ito formula with $g(t, X_t) = \ln(X_t)$ to get

$$\begin{aligned} d(\log N_t) &= \frac{1}{N_t} dN_t + \frac{1}{2}\left(-\frac{1}{N_t^2}\right)(dN_t)^2 \\ &= \frac{dN_t}{N_t} - \frac{1}{2N_t^2}\alpha^2 N_t dt = \frac{dN_t}{N_t} - \frac{1}{2}\alpha^2 dt \end{aligned}$$

and hence

$$\frac{dN_t}{N_t} = d(\ln N_t) + \frac{1}{2}\alpha^2 dt$$

and using eqn. 6.1

$$\begin{aligned} \int_0^t \left(d(\ln N_s) + \frac{1}{2}\alpha^2 ds \right) &= rt + \alpha B_t \\ \ln N_t - \ln N_0 + \frac{1}{2}\alpha^2 t &= rt + \alpha B_t \end{aligned}$$

or

$$N_t = N_0 e^{\left(r - \frac{1}{2}\alpha^2\right)t + \alpha B_t}$$

which is called *geometric Brownian motion*. Sanity check: on average the noise should filter out and this result should agree with the deterministic population growth ODE[9], i.e.

$$E[N_t] = E[N_0] e^{rt}$$

as if there were no noise term in the driving force $r_t dt + \alpha dB_t$. Indeed this is true. Let $Y_t = e^{\alpha B_t}$ and apply Ito's formula

$$dY_t = \frac{1}{2}\alpha^2 e^{\alpha B_t} dt + \alpha e^{\alpha B_t} dB_t$$

or

$$Y_t - Y_0 = \frac{1}{2}\alpha^2 \int_0^t e^{\alpha B_s} ds + \alpha \int_0^t e^{\alpha B_s} dB_s$$

and since $E\left[\int_0^t e^{\alpha B_s} dB_s\right] = 0$ by property **??** or Ito integral we have that

$$E[Y_t - Y_0] = \frac{1}{2}\alpha^2 \int_0^t E\left[e^{\alpha B_s}\right] ds$$

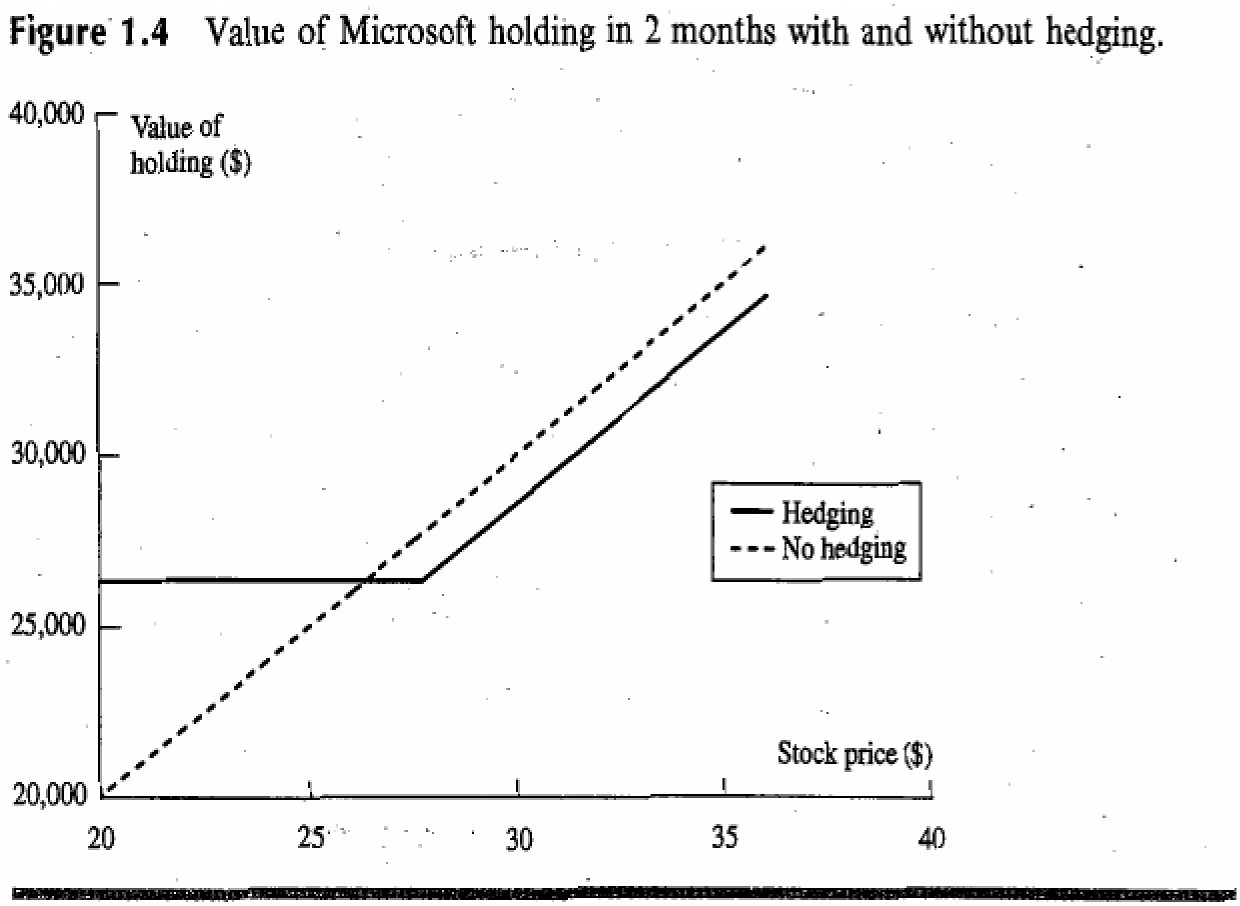
i.e.

$$\frac{d}{dt} E[Y_t] = \frac{1}{2}\alpha^2 E[Y_t], E[Y_0] = 1$$

which necessarily implies

$$E[Y_t] = e^{\frac{1}{2}\alpha^2 t}$$

[9] $\frac{dn}{dt} = an$



investor the right to sell a total of 1,000 shares for a price of \$27.50. If the quoted option price is \$1, then each option contract would cost $100 \times \$1 = \$100$ and the total cost of the hedging strategy would be $10 \times \$100 = \$1{,}000$.

FIGURE 7.1. Hedging

which itself implies that

$$E[N_t] = E[N_0]e^{rt}$$

since

$$N_t = N_0 e^{\left(r-\frac{1}{2}\alpha^2\right)t+\alpha B_t} = N_0 e^{\left(r-\frac{1}{2}\alpha^2\right)t} Y_t$$

and by independence

$$\begin{aligned} E[N_t] &= E[N_0]e^{\left(r-\frac{1}{2}\alpha^2\right)t}E[Y_t] \\ &= E[N_0]e^{\left(r-\frac{1}{2}\alpha^2\right)t}e^{\frac{1}{2}\alpha^2 t} \\ &= E[N_0]e^{rt} \end{aligned}$$

7. Option pricing

Definition 20. An *option* is a type of *derivative* that is a contract the affords the right but not the obligation to take a position against an *underlying* (stock, commodity, equity). The derivative derives its value from the value of the underlying. A *call option* provides the right to buy and a *put option* provides the right to sell the underlying at the *strike price.* Exercising the option is called *writing the contract.* A European call option is which can only be written at the time of expiration. An American option can be exercised at any time up to, and including, the time of expiry.

Example 21. Here is an example of using call options to hedge

So how much should you pay for such a right (either call or put)? Alternatively how much can you charge if you're selling such a contract? The basic idea in figuring out how to price such a contract is to assume that no arbitrages exist in the market and use delta hedging (fractional shares) to construct a perfect hedge (one that eliminates all risk from a portfolio). Since the portfolio is riskless it must earn interest at the risk-free rate. From this you can calculate the value of the portfolio at expiration and work backwards to see how much the option should have the buyer.

7.1. **Black-Scholes.** The Black-Scholes model was constructed by Fischer Black and Myron Scholes in 1973, based on work by Robert Merton. Merton and Scholes won the Nobel in Economics in 97 (Black was ineligible having died in 95). It forms the basis for much more complicated investment strategies used by hedge funds.

The Black–Scholes model assumes that the market consists of at least one risky asset, usually called the stock, and one riskless asset, usually called the money market, cash, or bond. The assumptions on market are

- There is no arbitrage opportunity (i.e., there is no way to make a riskless profit).
- It is possible to borrow and lend any amount, even fractional, of cash at the riskless rate.
- It is possible to buy and sell any amount, even fractional, of the stock.
- The above transactions do not incur any fees or costs (i.e., frictionless market).

The assumptions on the assets are

- The rate of return on the riskless asset is constant and thus called the risk-free interest rate.
- The instantaneous log returns of the stock price is a geometric Brownian motion, and we will assume its drift and volatility is constant.
- The stock does not pay a dividend.

Per the model assumptions above, the price of the underlying asset (typically a stock) follows a geometric Brownian motion. That is

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

We now construct a portfolio Π consisting of some number of European call options on the stock and some number of shares which is a perfect hedge (riskless). First the price of the call option V is a stochastic process and clearly a function of S_t and t. So by Ito's lemma

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

Now delta-hedging portfolio that turns out to be a perfect hedge is to sell 1 option and buy $\frac{\partial V}{\partial S}$ shares of the stock. It will become clear in a moment why this is the portfolio that turns out to be a perfect hedge. The value of this portfolio/holdings is

$$\Pi = -1 \times V + \frac{\partial V}{\partial S} S$$

Why? First note that we're on the sell side of the option. Meaning we have sold someone else the right to sell us something, at the expiration of the option. What happens if the price of the stock exceeds the strike price of the option? Then it wouldn't make sense for the buyer to sell us shares of the stock at the strike price because they would be losing money. Alternatively if the price of the stock is below the strike price then it does make sense for the buyer to sell us shares of the underlying. This is the value V of the option and to us it's a loss (because end up paying above market rate for the shares). Because we're on the sell side of the option (we sold a call option). On the otherhand the value of our holdings in the underlying $\left(\frac{\partial V}{\partial S} S\right)$ is a gain (since the shares simply have their face value).

Discretizing the differentials inΠ: from t to $t + \Delta t$ the net change in the value of the portfolio is

$$\Delta\Pi = -\Delta V + \frac{\partial V}{\partial S}\Delta S$$

Similarly discretising the process that defines S and V

$$\begin{aligned} \Delta S &= \mu S \Delta t + \sigma S \Delta B \\ \Delta V &= \left(\mu S\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) \Delta t + \sigma S \frac{\partial V}{\partial S} \Delta B \end{aligned}$$

Note ratios of differentials haven't been discretised because _ _ _ _ _ _ _. Now substituting into $\Delta\Pi$

$$\Delta\Pi = \left(-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) \Delta t$$

Note that ΔB dependence drops out. The portfolio is now riskless (independent of the stochastic volatility in the market). Therefore the rate of return must be the risk-free rate r (otherwise there would be arbitrage in the market because you could earn money on this riskless portfolio at a greater rate than the risk-free rate). Hence

$$\Delta\Pi = r\Pi\Delta t$$

If this is unrecognizable think

$$\frac{d\Pi}{dt} = r\Pi$$

i.e. continuously compounding interest. Equating the two expressions for $\Delta\Pi$ we get

$$\left(-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) \Delta t = r\left(-V + S\frac{\partial V}{\partial S}\right)\Delta t$$

or

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

the Black-Scholes equation. Solving this PDE with boundary conditions

$$V(0,t) = 0 \text{ for all } t \tag{7.1}$$
$$V(S,t) \to S \text{ as } S \to \infty \tag{7.2}$$
$$V(S,T) = \max\{S-K,0\} \tag{7.3}$$

gives the value of the option for all $t \le T$. To actually solve recognize it as a Cauchy-Euler equation which can be transformed into a diffusion equation with the change of variables

$$\begin{aligned} \tau &= T-t \\ u &= Ve^{r\tau} \\ x &= \ln\left(\frac{S}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)\tau \end{aligned}$$

to produce

$$\frac{\partial u}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2}$$

with boundary condition $V(S,T) = \max\{S-K,0\}$ becoming initial condition

$$u(x,0) = u_0(x) = K\left(e^{\max\{x,0\}} - 1\right)$$

The standard solution is by method of Green's functions

$$\begin{aligned} u(x,\tau) &= \frac{1}{\sigma\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} u_0(y)\, e^{-\frac{(x-y)^2}{2\sigma^2\tau}}\, dy \\ &= Ke^{x+\frac{1}{2}\sigma^2\tau}\Phi(d_1) - K\oplus(d_2) \end{aligned}$$

where Φ is the CDF of a standard normal and

$$d_1 = \frac{1}{\sigma\sqrt{\tau}}\left[\left(x + \frac{1}{2}\sigma^2\tau\right) + \frac{1}{2}\sigma^2\tau\right] \tag{7.4}$$
$$d_2 = \frac{1}{\sigma\sqrt{\tau}}\left[\left(x + \frac{1}{2}\sigma^2\tau\right) - \frac{1}{2}\sigma^2\tau\right] \tag{7.5}$$