1
Mathematics in a Pill

The purpose of this chapter is to give a brief outline of the probability theory underlying the mathematics inside the book, and to introduce necessary notation and conventions which are used throughout.

1.1 PROBABILITY SPACE AND RANDOM VARIABLES

A probability triple (Ω, Σ, P) consists of the following components:

1. A set Ω of elementary outcomes called the *sample space*.
2. A σ-algebra Σ of possible *events* (subsets of Ω).
3. A *probability function* $P : \Sigma \to [0, 1]$ that assigns real numbers between 0 and 1 called probabilities to the events in Σ.

The *conditional probability* of A given B is defined as follows:

$$P(A|B) = P(\mathrm{A} \cap \mathrm{B})/P(B).$$

Two events are said to be *independent* if the following three (equivalent) conditions hold:

1. $P(A \cap B) = P(A)P(B)$
2. $P(A) = P(A|B)$
3. $P(B) = P(B|A)$

A random variable $X : \Omega \to G$ is a measurable function from a probability space Ω into a Banach space G known as the state space.

We say that random two variables X and Y are independent if for all events A and B

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

We define expected (mean) value EX of the random variable X as the integral

$$EX = \int_{\Omega} X(\omega)P(d\omega),$$

and define the variance DX as

$$DX = \int_{\Omega} (X(\omega) - EX) \otimes (X(\omega) - EX)P(d\omega).$$

where $\otimes$ stands for tensor product. We may define the conditional expectation of a random variable X with respect to a σ-algebra $\Xi \subset \Sigma$. It is the only random variable $E(X|\Xi)$ such that for all $A \in \Xi$

$$\int_A X(\omega)P(d\omega) = \int_A E(X|\Xi)(\omega)P(d\omega).$$

If the state space is the real line R, we define the distribution function $F(x)$ (also called the cumulative density function or probability distribution function) as the probability that a real random variable X takes on a value less than or equal to a number x.

$$F(x) = P(X < x).$$

If the function F is differentiable, its derivative $f(x)$ is called the density function:

$$f(x) = F'(x).$$

1.2 NORMAL DISTRIBUTIONS

A normal (Gaussian) distribution on R with mean $EX = \mu$ and variance $DX = \sigma^2$ is a probability distribution with probability function

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\left(\frac{(t-\mu)^2}{2\sigma^2}\right)\right\}. \tag{1.1}$$

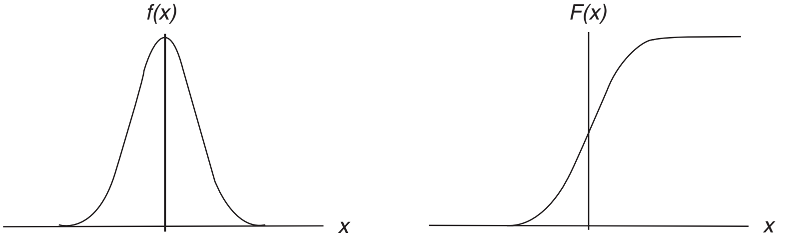

Figure 1.1 Gaussian distribution.

We also have the result that the sum of two normal variables is also a normal variable. A normal variable with mean $\mu = 0$ and variance $\sigma = 1$ is called a standard normal. We denote the cumulative distribution by N. A vector of M normal variables is called a multidimensional normal variable.

1.3 STOCHASTIC PROCESSES

Let $F_t \subset \Sigma$ be a family of increasing σ-algebras. We define the probability quadruple (Ω, F_t, Σ, P) as a standard probability setting for all dynamic models used in this book. A stochastic process is an indexed collection of F_t-measurable random variables $X(t)$, each of which is defined on the same probability triple (Ω, Σ, P) and takes values on the same codomain – in our case the interval $[0, T]$. In a continuous stochastic process the index set is continuous, resulting in an infinite number of random variables. A particular stochastic process is determined by specifying the joint probability distributions of the various random variables $X(t)$.

1.4 WIENER PROCESSES

A continuous-time stochastic process $W(t)$ with the following properties

- $W(0) = 0$,
- W has continuous paths,
- $W(s)$ and $(W(t) - W(s))$ are independent random variables for any $0 < s < t$,
- $W(t)$ has Gaussian distribution with mean 0 and variance t

is called Wiener process or Brownian motion. It was introduced by Louis Bachelier in 1900 as a model of stock prices. A vector of N independent Wiener processes is called a multidimensional Wiener process. The general shape of such a process is seen in the example below.

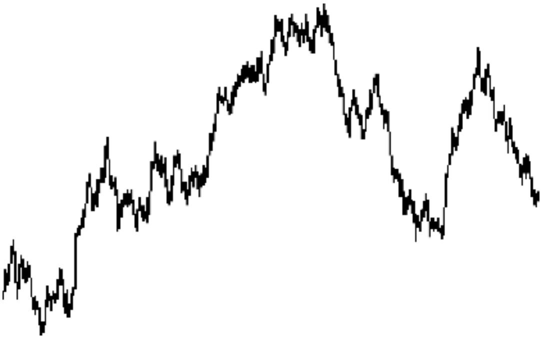

Figure 1.2 Wiener process.

1.5 GEOMETRIC WIENER PROCESSES

The following stochastic process

$$X(t) = X(0) \exp\left\{\mu t + \sigma W(t) - \frac{\sigma^2}{2} t\right\} \tag{1.2}$$

is called geometric Wiener process. The coefficient μ is called the drift and the coefficient σ is called the volatility.

1.6 MARKOV PROCESSES

A stochastic process X whose future probabilities are determined by its most recent values is called or is said to be Markov. This can be described mathematically in the following manner

$$P(X(T) \in A | X(s), s \leq t) = P(X(T) \in A | X(t)).$$

1.7 STOCHASTIC INTEGRALS AND STOCHASTIC DIFFERENTIAL EQUATIONS

If Y is a predictable stochastic process such that

$$P\left(\int_0^t |Y(s)|^2 \, ds < \infty \right) = 1,$$

we may define the stochastic integral with respect to the Wiener process $W(t)$ to be

$$C(t) = \int_0^t Y(s) \cdot dW(s). \tag{1.3}$$

If the process Y is deterministic then C is Gaussian with independent increments. The stochastic integral has the following properties:

$$EC(t) = 0 \text{ and } EC^2(t) = E \int_0^t |Y(s)|^2 \, ds.$$

We say that Y satisfies the Ito stochastic differential equation

$$\begin{aligned} dY(t) &= f(t, Y(t))dt + g(t, Y(t)) \cdot dW(t), \\ Y(0) &= y. \end{aligned} \tag{1.4}$$

If

$$Y(t) = Y(0) + \int_0^t f(s, Y(s))ds + \int_0^t g(s, Y(s)) \cdot dW(s).$$

If f and g are deterministic functions with properties that ensure uniqueness of solution, then the process Y is a Markov process. A Geometric Wiener process satisfies the following stochastic equation:

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t). \tag{1.5}$$

1.8 ITO'S FORMULA

Let the process Y satisfy the Ito equation:

$$dY(t) = f(t)dt + g(t) \cdot dW(t)$$

and let F be a smooth function. By applying the Ito formula we produce the stochastic equation satisfied by the process $F(t, Y(t))$:

$$dF(t, Y(t)) = \left(\frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial Y^2} |g(t)|^2 \right) dt + \frac{\partial F}{\partial Y} dY(t). \tag{1.6}$$

1.9 MARTINGALES

The N-dimensional stochastic process $M(t)$ is a martingale with respect to F_t if $E|C(t)| < \infty$ and the following property also holds:

$$M(t) = E\left(M(T)|F_t\right).$$

Every stochastic integral (and hence any Wiener process) is a martingale. However, a Geometric Wiener process is a martingale only if $\mu = 0$. Any continuous martingale M can be represented as an Ito integral, i.e.

$$M(t) = \int_0^t Y(s) \cdot dW(s)$$

for some predictable process Y. A martingale can be considered as a model of a fair game and therefore can be considered a proper model of financial markets.

1.10 GIRSANOV'S THEOREM

Let M be a positive continuous martingale, such that $M(0) = 1$. Then there exists a predictable stochastic process $\sigma(t)$ such that

$$dM(t) = -\sigma(t)M(t)dW(t)$$

or, equivalently

$$M(t) = \exp\left\{ -\frac{1}{2} \int_0^t \sigma^2(s)ds - \int_0^t \sigma(s)dW(s) \right\}.$$

If we now define new probability measure E_T by

$$P_T(A) = \int_\Omega I_A(\omega)M(T, \omega)P(d\omega).$$

then P_T is a probability measure under which the stochastic process

$$W_T(t) = W(t) + \int_0^t \sigma(s)ds$$

is a Wiener process.

1.11 BLACK'S FORMULA (1976)

Let the stochastic process X satisfy the equation:

$$dX(t) = \sigma X(t) dW(t).$$

Let C represent the (undiscounted) payoff from a European call option, so that $C = E\,(X(T) - K)^+$. Then C is given by the Black'76 formula:

$$C = X(0)N(d_1) - KN(d_2), \tag{1.7}$$

where

$$d_1 = \frac{\ln(X(0)/K) + \sigma^2 T/2}{\sigma\sqrt{T}},$$

$$d_2 = d_1 - \sigma\sqrt{T}.$$

1.12 PRICING DERIVATIVES AND CHANGING OF NUMERAIRE

We can introduce a general abstract approach to derivatives pricing as follows: We are given a set of positive continuous stochastic processes $X_0(t), X_1(t), \ldots, X_N(t)$ representing market quantities; these could be stock prices, interest rates, exchange rates, etc. We assume that the market is arbitrage-free, so that the quantities $M_1(t) = \frac{X_1(t)}{X_0(t)}, \ldots, M_N(t) = \frac{X_N(t)}{X_0(t)}$ are martingales, where $X_0(t)$ is called a basic asset – a numeraire. Pricing European derivatives maturing at time T consists of calculating functionals of the form:

$$Price = E\left\{\frac{\varsigma}{X_0(T)}\right\},$$

where ς is a random variable representing the payoff at time T. The process $X_0(t)$ is understood as the time value of money, i.e. comparable to a savings account, so we have to assume that $X_0(0) = 1$. If we define N new probability measures by

$$P_i(A) = X_i^{-1}(0) \int_\Omega I_A(\omega) M_i(T, \omega) P(d\omega).$$

then this leads to the following theorem:

Theorem. The processes $\frac{X_0(t)}{X_i(t)}, \frac{X_1(t)}{X_i(t)}, \ldots, \frac{X_N(t)}{X_i(t)}$ are martingales under the measure P_i.

Proof. Let ς be an F_t-measurable random variable.

$$X_i(0)E_i\left\{\frac{X_j(t)}{X_i(t)}\varsigma\right\} = E\left\{\frac{X_j(t)}{X_i(t)}\frac{X_i(t)}{X_0(t)}\varsigma\right\} = EM_j(t)\varsigma = E\varsigma E\left(M_j(T)\,|F_t\right)$$
$$= EE\left(M_j(T)\varsigma\middle|F_t\right) = EM_j(T)\varsigma = E\left\{\frac{X_j(T)}{X_i(T)}\frac{X_i(T)}{X_0(T)}\varsigma\right\} = X_i(0)E_i\left\{\frac{X_j(T)}{X_i(T)}\varsigma\right\}.$$

This simple theorem is extremely important. In pricing derivatives the savings account $X_0(t)$ can be replaced by any other tradable asset – we can change the numeraire, which may allow us to simplify certain calculations, for example we have

$$Price = E\left\{\frac{\varsigma}{X_0(T)}\right\} = E\left\{\frac{\varsigma}{X_1(T)}\frac{X_1(T)}{X_0(T)}\right\} = X_1(0)E_1\left\{\frac{\varsigma}{X_1(T)}\right\}.$$

1.13 PRICING OF INTEREST RATE DERIVATIVES AND THE FORWARD MEASURE

The theory of interest rate derivatives is in some sense simple because it relies only on one basic notion – the time value of money. Let us start with some basic notions: denote by $B(t, T)$ be discount factors on the period $[t, T]$ – understood as value at time t of an obligation to pay \$1 at time T. Payment of this dollar is certain; there is no credit risk involved. This obligation is also called a zero-coupon bond. We assume that zero-coupon bonds with all maturities are traded and this market is absolutely liquid – there are no transaction spreads. These assumptions are quite sensible since the money, bond and swap markets are very liquid with spreads not exceeding several basis points. Notice several obvious properties of discount factors:

$$0 < B(t, T) \le B(t, S) \le 1 \text{ if } S \le T \text{ and } B(T, T) = 1.$$

Let $X_0(t)$ be the savings account then all tradable assets $\xi(t)$ satisfy the arbitrage property that

$$\frac{\xi(t)}{X_0(t)} \text{ is a martingale.}$$

In particular we have that

$$M(t, T) = \frac{B(t, T)}{X_0(t)B(0, T)}$$

is a positive continuous martingale. We assume that the savings account is a process with finite variation – existence and uniqueness of a savings account may be a subject to a fascinating mathematical investigation. Since this problem is completely irrelevant to pricing issues – we refer to Musiela and Rutkowski (1997b) stating only that it is satisfied for all practical models. The savings account is of little interest because it is not a tradable asset, hence its importance is rather of mathematical character and practitioners try get rid of all notions not related to trading as soon as possible. We adopt this principle and will shortly remove the notion of savings account from our calculations.

There exists a d-dimensional stochastic process $\Sigma(t,T)$ a d-dimensional Brownian motion and such that

$$dB(t,T) = -B(t,T)\left(-d\ \ln X_0(t) + \Sigma(t,T)\cdot dW(t)\right)$$

and

$$dM(t,T) = -M(t,T)\Sigma(t,T)\cdot dW(t).$$

Remark. The d-dimensional representation is not unique, however uniqueness does hold for the single dimensional representation. Since most financial models are multidimensional we have chosen the less elegant d-dimensional representation. The dot stands for scalar product.

Therefore

$$M(t,T) = \exp\left\{-\frac{1}{2}\int_0^t |\Sigma(s,T)|^2 ds - \int_0^t \Sigma(s,T)\cdot dW(s)\right\}$$

and

$$B(t,T) = B(0,T)X_0^{-1}(t)\exp\left\{-\frac{1}{2}\int_0^t |\Sigma(s,T)|^2 ds - \int_0^t \Sigma(s,T)\cdot dW(s)\right\}. \tag{1.8}$$

Since $B(T,T)=1,\ M(T,T)B(0,T) = X_0^{-1}(T)$.

The pricing of European interest rate derivatives consists of finding expectation of discounted values of cash flows

$$E\left(X_0^{-1}(t)\xi\right),$$

where ξ is an F_T-measurable random variable – the intrinsic value of the claim. Define the probability measure E_T by

$$E_T\varsigma = E\varsigma M(T,T)$$

for any random variable ς. By the Girsanov theorem E_T is a probability measure under which the process

$$W_T(t) = W(t) + \int_0^t \Sigma(s,T)ds$$

is a Wiener process. Now

$$EX_0^{-1}(T)\xi = B(0,T)EM(T,T)\xi = B(0,T)E_T\xi.$$

We may take discounting with respect to multiple cash flows as in the case of swaptions. Let δ be accrual period for both interest rates and swaps. For simplicity, we assume it is constant. Define consecutive grid points as $T_{i+1} = T_i + \delta$ for a certain initial $T = T_0 < \delta$. To ease the notation, we set $E_n = E_{T_n}$ and $W_n = W_{T_n}$. The forward compound factors and forward LIBOR rates are defined as

$$\delta L_n(t) + 1 = D_n(t) = \frac{B(t, T_{n-1})}{B(t, T_n)} \tag{1.9}$$

and forward swap rates as

$$S_{nN}(t) = \frac{\sum_{i=n+1}^{N} B(t, T_i) L_i(t)}{A_{nN}(t)} = \frac{B(t, T_n) - B(t, T_N)}{\delta A_{nN}(t)}$$

where

$$A_{nN}(t) = \sum_{i=n+1}^{N} B(t, T_i).$$

Now let

$$C(S_{nN}) = \sum_{i=n+1}^{N} X_0^{-1}(T_i) = \sum_{i=n+1}^{N} B(0, T_i) M(T_i, T_i).$$

Thus the pricing of European swap derivatives consists of finding

$$E(C(S_{nN})\xi),$$

where ξ is an $F_{T_{n+1}}$-measurable random variable – the intrinsic value of the claim. Since $M(t, T)$ is a positive continuous martingale we also have that the following is a positive continuous martingale:

$$M(t, S_{nN}) = \frac{\sum_{i=n+1}^{N} B(0, T_i) M(t, T_i)}{A_{nN}(0)} = \frac{\sum_{i=n+1}^{N} B(t, T_i)}{X_0(t) A_{nN}(0)}.$$

Moreover

$$dM(t, S_{nN}) = -\frac{\sum_{i=n+1}^{N} B(0, T_i) M(t, T_i)}{A_{nN}(0)} \cdot \frac{\sum_{i=n+1}^{N} B(0, T_i) M(t, T_i) \Sigma(t, T_i)}{\sum_{i=n+1}^{N} B(0, T_i) M(t, T_i)} dW(t)$$

$$= -M(t, S_{nN}) \frac{\sum_{i=n+1}^{N} B(t, T_i) \Sigma(t, T_i)}{A_{nN}(t)} dW(t).$$

Therefore E_{nN} defined by

$$E_{nN}\varsigma = E\varsigma M(T_{n+1}, S_{nN})$$

is a probability measure under which the process

$$W_{nN}(t) = W(t) + \int_0^t \frac{\sum_{i=n+1}^{N} B(s, T_i)\Sigma(s, T_i)}{A_{nN}(t)} ds \tag{1.10}$$

is a Wiener process. Hence

$$\begin{aligned} EC(S_{nN})\xi &= \sum_{i=n+1}^{N} B(0, T_i) EM(T_i, T_i)\xi \\ &= \sum_{i=n+1}^{N} B(0, T_i) EM(T_{n+1}, T_i)\xi = A_{nN}(0) E_{nN}\xi. \end{aligned}$$

Moreover

$$M(t, S_{nN}) S_{nN}(t) = \frac{\sum_{i=n+1}^{N} B(t, T_i)}{X_0(t) A_{nN}(0)} \frac{B(t, T_n) - B(t, T_N)}{A_{nN}(t)} = \frac{B(t, T_n) - B(t, T_N)}{X_0(t) A_{nN}(0)}.$$

Therefore $S_{nN}(t)M(t, S_{nN})$ is a martingale under the measure E, and then the forward swap rate $S_{nN}(t)$ is a martingale under E_{nN}.