Finance 400

A. Penati - G. Pennacchi

The Essentials of Diffusion Processes and Itô's Lemma

I. Introduction

These notes cover the basic properties of continuous time stochastic processes. They describe the characteristics of these processes that are helpful for modeling many financial and economic time series. Modeling a variable as a continuous time process can lead to different implications versus modeling it as a discrete time process. A continuous time stochastic process allows a variable to change randomly and yet be observed at each moment. In contrast, a discrete-time stochastic process implies no change in the value of the variable over a fixed interval or that the change cannot be observed in-between discrete points in time. A natural implication of a discrete time asset price process is that no trading occurs in the asset over the discrete interval. Often this makes problems involving the hedging of an asset's risk difficult, since portfolio allocations cannot be re-balanced over this non-trading interval. Thus hedging the risk of assets may be less than perfect when a discrete time process is assumed.

In contrast, if one assumes that asset prices follow continuous time processes, prices can be observed and trade can take place continuously. This can permit dynamic trading strategies that can fully hedge an asset's risk. Making this continuous hedging assumption often simplifies optimal portfolio choice problems and problems of valuing contingent claims (derivative securities). It allows for a continuous distribution of asset returns (an infinite number of states), yet obtain market completeness because payoffs can be replicated through continuous trading.

The following analysis is done at an intuitive level rather than a mathematically rigorous one. We start with Brownian motion, which is the fundamental building block of diffusion processes. Later, we consider Itô's Lemma, which tells us how to derive the continuous time process of a variable that is a function of another variable that follows a continuous time process.

II. Pure Brownian Motion

Here we show how a Brownian motion process can be defined as the limit of a discrete time process. Consider the following stochastic process observed at date t, $z(t)$. Let Δt be a discrete change in time, that is, some time interval. The change in $z(t)$ over the time interval Δt is given by

$$z(t+\Delta t) - z(t) \;\equiv\; \Delta z \;=\; \sqrt{\Delta t}\,\tilde{\epsilon} \tag{1}$$

where $\tilde{\epsilon}$ is a random variable with $E[\,\tilde{\epsilon}\,] = 0$, $Var[\,\tilde{\epsilon}\,] = 1$, and $Cov[\,z(t+\Delta t) - z(t),\; z(s+\Delta t) - z(s)\,] = 0$ if $(t,\, t+\Delta t)$ and $(s,\, s+\Delta t)$ are non-overlapping intervals. $z(t)$ is an example of a process referred to as a random walk.

Thus, given the assumed moments of $\tilde{\epsilon}$, $E[\Delta z] = 0$ and $Var[\Delta z] = \Delta t$. Note also that $z(t)$ has serially uncorrelated (independent) increments. Now consider the change in $z(t)$ over a fixed interval, from 0 to T. Assume T is made up of n intervals of length Δt. Then

$$z(T) \;-\; z(0) \;=\; \sum_{i=1}^{n} \Delta z_i \tag{2}$$

where $\Delta z_i \;\equiv\; z(i \cdot \Delta t) - z(\,[i-1] \cdot \Delta t) \;\equiv\; \sqrt{\Delta t}\,\tilde{\epsilon}_i$. Hence (2) can also be written as

$$z(T) \;-\; z(0) \;=\; \sum_{i=1}^{n} \sqrt{\Delta t}\,\tilde{\epsilon}_i \;=\; \sqrt{\Delta t}\,\sum_{i=1}^{n} \tilde{\epsilon}_i. \tag{3}$$

Now note that the first two moments of $z(T) - z(0)$ are

$$\begin{aligned} E_0[\,z(T) \;-\; z(0)\,] \;\;&=\;\; \sqrt{\Delta t}\,\sum_{i=1}^{n} E_0[\,\tilde{\epsilon}_i\,] = 0 && (4) \\ Var_0[\,z(T) \;-\; z(0)\,] \;\;&=\;\; \left(\sqrt{\Delta t}\right)^2 \sum_{i=1}^{n} Var_0[\tilde{\epsilon}_i] \;=\; \Delta\, t \,\cdot\, n \,\cdot\, 1 \;=\; T. && (5) \end{aligned}$$

where $E_t\,[\cdot]$ and $Var_t\,[\cdot]$ and the mean and variance operators, respectively, conditional on information at date t. We see that holding T, the length of the time interval, fixed, the mean and variance of $z(T) - z(0)$ are independent of n. Now let us perform the following experiment. Suppose we keep T fixed but let n, the number of intervening increments of length Δt, go to infinity. Can we say something else about the distribution of $z(T) - z(0)$ besides what its first two moments are? Yes we can. Assuming the distributions of $\tilde{\epsilon}_i$ are sufficiently well-behaved, we can state

$$\underset{n\to\infty}{plim}\,(z(T)-z(0))\;=\underset{\Delta\,t\to 0}{plim}\,(z(T)-z(0))\;\sim\;N(0,\,T). \tag{6}$$

In other words, $z(T) - z(0)$ has a *normal distribution* with mean zero and variance T. This follows from the *Central Limit Theorem* which states that the sum of random variables having an arbitrary (but sufficiently well-behaved) probability distribution has a limiting normal distribution. Thus, the distribution of $z(t)$ over any finite interval, $[\,0,\,T\,]$, can be thought of as the sum of infinitely many tiny independent increments, $\Delta\,z_i = \sqrt{\Delta t}\,\tilde{\epsilon}_i$, that have some arbitrary distribution. However, when added together, these increments result in a normal distribution. Therefore, without loss of generality, we can assume that each of the $\tilde{\epsilon}_i$ have a standard (mean 0 variance 1) normal distribution.[1]

The limit of one of these tiny independent increments can be defined as

$$dz(t)\;\equiv\underset{\Delta\,t\to 0}{lim}\,\Delta z\;=\underset{\Delta t\to 0}{lim}\,\sqrt{\Delta\,t}\tilde{\epsilon} \tag{7}$$

where $\tilde{\epsilon} \sim N(0\,1)$. Hence, $E[\,dz(t)\,] = 0$ and $Var[\,dz(t)\,] = \lim_{\Delta t\to 0}\,\Delta\,t = dt$. dz is referred to as a pure Brownian motion process or a Wiener process. We can now write the change in $z(t)$ over any finite interval $[\,0,\,T\,]$ as

$$z(T)-z(0)\;=\;\int_0^T dz(t)\;\sim\;N(0,\,T). \tag{8}$$

The integral in (8) is a *stochastic* (Itô) integral, not the standard Riemann integral. Note that $z(t)$ is a continuous process but constantly changing (by $\tilde{\epsilon}$ over each infinitely small interval Δt), such that over any finite interval it has unbounded variation. Hence, it is nowhere differentiable (very jagged), that is, its derivative $\partial z(t)/\partial t$ does not exist.

Brownian motion provides the basis for more general continuous time stochastic processes, also known as diffusion processes. Let us see how these more general processes can be developed.

[1]Note that sums of normally distributed random variables are also normally distributed. Thus, the Central Limit Theorem also applies to sums of normals.

III. Diffusion Processes

To illustrate how we can build on the basic Wiener process, consider the process for dz multiplied by a constant, σ. Define a new process $x(t)$ by

$$dx(t) \;=\; \sigma\, dz(t) \tag{9}$$

Then over a discrete interval, $[0,\, T], x(t)$ is distributed

$$\int_0^T dx \;=\; x(T) - x(0) \;=\; \int_0^T \sigma\, dz(t) \;=\; \sigma \int_0^T dz(t) \;\sim\; N(0,\, \sigma^2 T). \tag{10}$$

Next, consider adding a deterministic (non-stochastic) change of $\mu(t)$ per unit of time to the $x(t)$ process.

$$dx \;=\; \mu(t)dt + \sigma dz \tag{11}$$

Now over any discrete interval, $[0,\, T]$, we have

$$\int_0^T dx \;=\; x(T) {-} x(0) \;=\; \int_0^T \mu\,(t)dt + \int_0^T \sigma\, dz(t) \;=\; \int_0^T \mu\,(t)dt {+} \sigma \int_0^T dz(t) \;\sim\; N(\int_0^T \mu\,(t)dt,\, \sigma^2 T). \tag{12}$$

For example, if $\mu(t) = \mu$, a constant, then $x(T) - x(0) = \mu T + \sigma \int_0^T dz(t) \sim\; N(\mu T,\, \sigma^2 T)$. Thus, we have been able to generalize the standard trend-less Wiener process to have a non-zero mean or "drift" as well as any desired variance or "volatility." The process $dx \;=\; \mu dt + \sigma dz$ is referred to as arithmetic Brownian motion (with drift).

In general, both μ and σ can be time varying, either as a simple function of time, t, or even a function of the value of the random variable, $x(t)$. In this case, the *stochastic differential equation* describing $x(t)$ is

$$dx(t) \;=\; \mu[x(t),\, t]\, dt \;+\; \sigma[x(t),\, t]\, dz \tag{13}$$

and the corresponding integral equation is

$$\int_0^T dx \;=\; x(T) - x(0) \;=\; \int_0^T \mu[x(t),\, t]\, dt \;+\; \int_0^T \sigma[x(t),\, t]\, dz \tag{14}$$

In this general case, $dx(t)$ could be described as being *instantaneously* normally distributed with mean $\mu[x(t),\, t]\, dt$ and variance $\sigma[x(t),\, t]^2\, dt$, but over any finite interval, $x(t)$ will not, in general, be normally distributed. One needs to know the functional form of $\mu[x(t),\, t]$ and $\sigma[x(t),\, t]$ to determine the discrete time distribution of $x(t)$ implied by its continuous time process. Importantly, however, it can be shown that the discrete distribution of $x(t)$ satisfies particular partial differential equations known as the Kolmogorov backward and forward equations. If we let $p(x, T; x_0, t_0)$ be the probability density function for x at date T given that it equals x_0 at date t_0, where $T \geq t_0$, then it must satisfy the backward Kolmogorov equation[2]

$$\frac{1}{2}\sigma^2(x_0, t_0)\frac{\partial^2 p}{\partial x_0^2} + \mu[x_0,\, t_0]\frac{\partial p}{\partial x_0} + \frac{\partial p}{\partial t_0} = 0 \tag{15}$$

and the forward Kolmogorov equation

$$\frac{1}{2}\frac{\partial^2 \sigma^2(x, T)\, p}{\partial x^2} - \frac{\partial \mu[x,\, T]p}{\partial x} - \frac{\partial p}{\partial T} = 0 \tag{16}$$

where the solutions to (15) and (16) satisfy the boundary equation that $x(T) = x_0$ when $T = t_0$, that is $p(x, t_0; x_0, t_0) = \delta(x - x_0)$, where $\delta(\cdot)$ is the Dirac delta function that equals infinity when its argument is zero, and equals zero everywhere else. Moreover, $\int_{-\infty}^{\infty}\delta(y)\, dy = 1$.

An Itô integral is formally defined as a mean square limit of a sum involving the discrete Δz_i processes. For example, when $\sigma[x(t),\, t]$ is a function of $x(t)$ and t, the Itô integral in (14) is defined from the relationship

$$\lim_{n\to\infty} E_0\left[\sum_{i=1}^{n}\sigma\left[x\left([i-1]\cdot\Delta t\right), [i-1]\cdot\Delta t\right]\Delta z_i - \int_0^T \sigma[x(t),\, t]\, dz\right]^2 = 0 \tag{17}$$

An important Itô integral that will be used below is $\int_0^T\, [dz(t)]^2$. It is defined from

$$\lim_{n\to\infty} E_0\left[\sum_{i=1}^{n}[\Delta z_i]^2 - \int_0^T\, [dz(t)]^2\right]^2 = 0 \tag{18}$$

Recall from (5) that

[2] A explanation for this result will be given in future notes.

$$Var_0\left[z\left(T\right)-z\left(0\right)\right]=Var_0\left[\sum_{i=1}^{n}\Delta z_i\right]=E_0\left[\left(\sum_{i=1}^{n}\Delta z_i\right)^2\right]=E_0\left[\sum_{i=1}^{n}\left[\Delta z_i\right]^2\right]=T \tag{19}$$

Further, straightforward algebra shows that

$$E_0\left[\sum_{i=1}^{n}\left[\Delta z_i\right]^2-T\right]^2=2T\Delta t \tag{20}$$

Hence, if we take the limit as $\Delta t\to 0$, or $n\to\infty$, of the expression in (20), one obtains

$$\lim_{n\to\infty}E_0\left[\sum_{i=1}^{n}\left[\Delta z_i\right]^2-T\right]^2=\lim_{\Delta t\to 0}2T\Delta t=0 \tag{21}$$

Comparing (18) with (21) implies

$$\begin{aligned}\int_0^T\left[dz\left(t\right)\right]^2 &= T \\ &= \int_0^T dt\end{aligned} \tag{22}$$

Hence, we have $\left[dz\left(t\right)\right]^2=dt$.

To further generalize continuous-time processes, suppose that we have some variable, F, that is a function of the current value of a diffusion process, $x(t)$, and (possibly) also is a direct function of time. Can we then characterize the stochastic process followed by $F(x(t),\,t)$, which now depends on the diffusion process, $x(t)$? The answer is yes, and Itô's lemma shows us to do it.

IV. Functions of Continuous-Time Processes and Itô's Lemma

Itô's Lemma is sometimes referred to as the *fundamental theorem of stochastic calculus*. It gives the rule for finding the differential of a function of one or more variables, each of which follow a stochastic differential equation containing Wiener processes. Here, we state and prove Itô's lemma for the case of a univariate function.

Itô's Lemma (univariate case): Let the variable $x(t)$ follow the stochastic differential equation $dx(t) = \mu(x,t)\,dt + \sigma(x,t)\,dz$. Further, let $F(x(t), t)$ be at least a twice differentiable function. Then the differential of $F(x, t)$ is given by:

$$dF = \frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial t}\,dt + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\,(dx)^2 \tag{23}$$

where the product $(dx)^2 = \sigma(x,t)^2 dt$. Hence, substituting in for dx and $(dx)^2$, the above can be re-written:

$$dF = \left[\frac{\partial F}{\partial x}\,\mu(x,t) + \frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\,\sigma(x,t)^2\right] dt + \frac{\partial F}{\partial x}\,\sigma(x,t)\,dz. \tag{24}$$

<u>Proof:</u> The proof is rather lengthy and will not be detailed here. An intuitive proof is given in Jonathan E. Ingersoll (1987) *Theory of Financial Decision Making*, Rowman and Littlefield, Totowa, NJ, (p.348-349). It involves looking at the change in $F(x, t)$ over a discrete interval of time for a given realization of x, that is, $F(x_j(t+\Delta t), t+\Delta t) - F(x, t)$. Expanding the end of interval value of F in a Taylor series around the (known) value of x and t at the start of the interval, one sees that many of the higher order Taylor series terms go to zero in the limit as the interval shrinks. Taking the expected value and variance of the change in F over this interval provides the relevant drift and stochastic terms for dF. Similar arguments show that

$$\begin{aligned} (dx)^2 &= (\mu(x,t)\,dt + \sigma(x,t)\,dz)^2 \\ &= \sigma(x,t)^2\,(\,dz)^2 = \sigma(x,t)^2 dt \end{aligned} \tag{25}$$

Note from (24) that the process followed by $F(t)$ is similar to that of $x(t)$ in that the stochastic component depends on dz. Thus, while F will have a mean (drift) and variance (volatility) that differs from x, they will both depend on the single source of uncertainty generated by dz.

<u>Example:</u>

A process that is used in many applications is the geometric Brownian motion process. It is given by

$$dx = \mu x\,dt + \sigma x\,dz \tag{26}$$

where μ and σ are constants. It is an attractive process because if x starts at a positive value, it always remains positive. This is because its mean and variance are both proportional to its current value, x. Hence a process like dx is often used to model the price of a limited-liability security, such as a common stock. Now consider the following function $y = \ln(x)$. What type of process does y follow? Applying Itô's lemma, we have

$$\begin{aligned} dy &= d\,(\ln x) = \left[\frac{\partial(\ln x)}{\partial x}\,\mu x + \frac{\partial(\ln x)}{\partial t} + \frac{1}{2}\frac{\partial^2(\ln x)}{\partial x^2}\,(\sigma x)^2\right] dt + \frac{\partial(\ln x)}{\partial x}\,\sigma x\, dz \\ &= \left[\mu + 0 - \frac{1}{2}\sigma^2\right] dt + \sigma\, dz. \end{aligned} \tag{27}$$

Thus we see that if x follows *geometric* Brownian motion, then $y = \ln x$ follows *arithmetic* Brownian motion. Since we know that

$$y(T) - y(0) \sim N\left((\mu - \frac{1}{2}\sigma^2)\, T,\, \sigma^2 T\right) \tag{28}$$

then $x(t) = e^{y(t)}$ has a lognormal distribution over any discrete interval (by the definition of a lognormal random variable). Hence, geometric Brownian motion is lognormally distributed over any time interval.

In a number of applications we need to derive the stochastic process for a function of several variables, each of which follows a diffusion process. So suppose we have m different diffusion processes of the form:[3]

$$dx_i = \mu_i\, dt + \sigma_i\, dz_i \qquad i = 1, \ldots, m, \tag{29}$$

and $dz_i dz_j = \rho_{ij} dt$, where ρ_{ij} has the interpretation of a correlation coefficient of the two Wiener processes. What is meant this correlation? Recall that $dz_i dz_i = (dz_i)^2 = dt$. Now the Wiener process dz_j can be written as a linear combination of two other Wiener processes, one being dz_i, and another process that is uncorrelated with dz_i, call it dz_{iu}:

$$dz_j = \rho_{ij} dz_i + \sqrt{1 - \rho_{ij}^2}\, dz_{iu} \tag{30}$$

[3]Note μ_i and σ_i may be functions of calendar time, t, and the current values of x_j, $j = 1, ..., m$.

Then from this interpretation of dz_j, we have

$$
\begin{aligned}
dz_j dz_j &= \rho_{ij}^2 (dz_i)^2 + \left(1-\rho_{ij}^2\right)(dz_{iu})^2 + 2\rho\sqrt{1-\rho_{ij}^2}\, dz_i dz_{iu} \\
&= \rho_{ij}^2 dt + \left(1-\rho_{ij}^2\right) dt + 0 \\
&= dt
\end{aligned}
\tag{31}
$$

and

$$
\begin{aligned}
dz_i dz_j &= dz_i \left(\rho_{ij} dz_i + \sqrt{1-\rho_{ij}^2}\, dz_{iu}\right) \\
&= \rho_{ij} (dz_i)^2 + \sqrt{1-\rho_{ij}^2}\, dz_i dz_{iu} \\
&= \rho_{ij} dt + 0
\end{aligned}
\tag{32}
$$

Thus, ρ_{ij} can be interpreted the proportion of dz_j that is correlated with dz_i.

We can now state (again without proof) a multivariate version of Itô's Lemma.

Itô's Lemma (multivariate version): Let $F(x_1, \ldots, x_m, t)$ be at least a twice differentiable function. Then the differential of $F(x_1, \ldots, x_m, t)$ is given by:

$$dF = \sum_{i=1}^{m} \frac{\partial F}{\partial x_i} dx_i + \frac{\partial F}{\partial t} dt + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 F}{\partial x_i \partial x_j} dx_i \, dx_j \tag{33}$$

where $dx_i \, dx_j = \sigma_i \sigma_j \rho_{ij} \, dt$. Hence, the above can be re-written

$$dF = \left[\sum_{i=1}^{m} \left(\frac{\partial F}{\partial x_i} \mu_i + \frac{1}{2} \frac{\partial^2 F}{\partial x_i^2} \sigma_i^2 \right) + \frac{\partial F}{\partial t} + \sum_{i=1}^{m} \sum_{j>i}^{m} \frac{\partial^2 F}{\partial x_i \partial x_j} \sigma_i \sigma_j \rho_{ij} \right] dt + \sum_{i=1}^{m} \frac{\partial F}{\partial x_i} \sigma_i \, dz_i. \tag{34}$$