TEACHING NOTE 00-05:

BROWNIAN MOTION: FROM DISCRETE TO CONTINUOUS TIME

Version date: July 18, 2008 D:\TN00_05.wpd

This teaching note overlaps and complements TN96-04, Modeling Asset Prices as Stochastic Processes I, TN00-03, Modeling Asset Prices as Stochastic Processes II, 96-05, Itô's Lemma and TN00-04, Girsanov's Theorem in Derivative Pricing.

In order to model an asset price, we must start with a model that reflects the noise produced by random information generated in such a manner that its expectation and volatility is the same through time. That model is typically called a Brownian motion, named for the Scottish scientist, Robert Brown, who supposedly observed it around 1827 in pollen suspended in water. Much of the scientific work for this model was done by Einstein and Norbert Weiner, the latter for whom a form of Brownian motion, the Weiner process, was named. This process is also called a random walk, though technically a random walk is slightly different.[1]

Brownian Motion in a Discrete World

An extremely simple form of the process is a binomial outcome in which a random variable W, starting off at a value of zero, can take on a value of +1 or -1 in the next time period. Thus,

$$W_0 = 0$$

$$W_1 = +1 \text{ or } -1$$

Although some versions of the model permit unequal probabilities, many desirable properties are associated with the case where the probabilities of the up and down moves are ½. If, however, one is attempting to model a process with a given expectation and volatility, then there are formulas that specify the probabilities, which will not generally be ½. In the limiting case, i.e., when the number of time steps is large, however, the formulas converge to ½.

Now let us examine some of the properties of the model. First consider the increment, ΔW_0:

$$E(\Delta W_0) = (1/2)(+1) + (1/2)(-1) = 0.$$

[1]A random walk is a stochastic process consisting of a series of independent and identically distributed random variables. A Brownian motion is a random walk, but all random walks are not Brownian motions.

$$Var(\Delta W_0) = (1/2)(+1 - 0)^2 + (1/2)(-1 - 0)^2 = 1.$$

This means that since $W_1 = W_0 + \Delta W_0$,

$$E(W_1) = E(W_0 + \Delta W_0) = E(W_0) + E(\Delta W_0) = 0 + 0 = 0$$

The variance of W_1 is, by definition, $Var(W_1) = E(W_1^2) - (E(W_1))^2$. Let us examine the expected value of the square of W_1:

$$E(W_1^2) = (1/2)(+1)^2 + (1/2)(-1)^2 = 1.$$

Since $E(W_1) = 0$, we have

$$Var(W_1) = E(W_1^2) - (E(W_1))^2 = 1 - 0 = 1.$$

Another way to get this result is to realize that we want $Var(W_1) = Var(W_0 + \Delta W_0)$. The variance of a sum is the sum of the variances and twice all pairwise covariances, but W_0 is a constant so its variance is zero and, therefore, there is no covariance. Thus, $Var(W_1) = Var(\Delta W_0)$, which we already found as 1.

In this form, however, the model is too simple. We can extend it somewhat by adding time periods. Note, however, that if we just let it move from +1 to +2 or 0 or from -1 to 0 or -2, its variance would obviously increase.[2] With a large number of time periods, we might find ourselves with an unreasonable variance. One way to scale the variance is to establish that we are trying to model a random process over a fixed period of time. We might say, for example, that we wish to model movements over a period t = 1, which might be one year. We might capture these movements with a model of n = 2 periods. We are, therefore, establishing that the time step is $\Delta t = 1/2$. We might be inclined, therefore, to adjust our model so that $\Delta W = \pm 1\Delta t$, but this will cause a problem. Intuitively, we might expect that we can better capture reality if we shrink the time interval such that t is fixed but n is large. In that case, Δt 6 0. What does this leave us with? No motion at all, as it is easy to see that the variance will approach zero.[3]

Alternatively, let us try the model:

$$\Delta W = \pm\sqrt{\Delta t}.$$

Now, let us observe what we have:

---

[2]Now there would be three outcomes at time step 2: +2 with probability 1/4, 0 with probability ½, and -2 with probability 1/4. The expected value is still zero but the variance is now $(1/4)(2 - 0)^2 + (1/2)(0 - 0)^2 + (1/4)(-2 - 0)^2 = 2$.

[3]To obtain the variance we would end up squaring Δt, which would take it to zero when Δt is small.

$$W_1^{++} = +2\sqrt{1/2}$$

$$W_1^{+} = +1\sqrt{1/2}$$

$$W_0 = 0 \qquad\qquad W_1^{+-} = 0\sqrt{1/2} = 0$$

$$W_1^{-} = -1\sqrt{1/2}$$

$$W_1^{--} = -2\sqrt{1/2}$$

Now let us find the expectation and variance of ΔW, W_1 and W_2.[4] First, for the increment:

$$E(\Delta W) = (1/2)(+1\sqrt{1/2}) + (1/2)(-1\sqrt{1/2}) = 0.$$

$$Var(\Delta W) = (1/2)(+1\sqrt{1/2} - 0)^2 + (1/2)(-1\sqrt{1/2})^2 = 1/2.$$

For W_1:

$$E(W_1) = E(W_0) + E(\Delta W) = W_0 + E(\Delta W) = 0 + 0 = 0$$

$$Var(W_1) = Var(W_0) + Var(\Delta W) = 0 + 1/2 = 1/2.$$

For W_2, note that the variance is the sum of the variance of ΔW_0 and ΔW_1, which are the same, and we can ignore any covariance between ΔW_0 and ΔW_1.[5]

$$E(W_2) = E(W_0) + E(\Delta W_0) + E(\Delta W_1) = W_0 + E(\Delta W_0) + E(\Delta W_1) = 0 + 0 + 0 = 0.$$

$$Var(W_2) = Var(W_0) + Var(\Delta W_0) + Var(\Delta W_1) = 0 + 1/2 + 1/2 = 1.$$

Now let us consider the general case of a process W_t:

$$E(\Delta W_t) = (1/2)(+1\sqrt{\Delta t}) + (1/2)(-1\sqrt{\Delta t}) = 0.$$

$$Var(\Delta W_t) = (1/2)(+1\sqrt{\Delta t} - 0)^2 + (1/2)(-1\sqrt{\Delta t} - 0)^2 = \Delta t.$$

Now consider the expectation and variance of W_t. First let us define:

$$W_t = W_0 + \sum_{i=0}^{n-1} \Delta W_i.$$

Then we can find the expected value and variance as

---

[4]We do not need to distinguish ΔW_0 from ΔW_1 as these are the same process.

[5]We could also find the variance of W_2 the traditional way, by probability-weighting the three squares of the outcomes less the expected value.

$$
\begin{aligned}
E(W_t) &= E(W_0) + E\left(\sum_{i=0}^{n-1} \Delta W_i\right) \\
&= 0 + E(\Delta W_0) + E(\Delta W_1) + \ldots + E(\Delta W_{n-1}) \\
&= 0 + 0 + 0 + \ldots + 0 = 0 \\
Var(W_t) &= Var(W_0) + Var\left(\sum_{i=0}^{n-1} \Delta W_i\right) \\
&= 0 + Var(\Delta W_0) + Var(\Delta W_1) + \ldots + Var(\Delta W_{n-1}) \\
&= 0 + \Delta t + \Delta t + \ldots + \Delta t = t.
\end{aligned}
$$

The last result comes from the fact that summing from 0 to n-1 is n items. So n times Δt is t because $\Delta t = t/n$.

The probability distribution for outcomes is, of course, the binomial distribution, which is well-known and which has simple formulas giving the probabilities of various outcomes.

Moving from a Discrete to a Continuous World

It is possible to demonstrate quite formally that if we hold t fixed and increase n, that in the limit, the random part, ±1, will converge to a standard normal random variable, which of course has expected value of 0 and variance of 1.[6] In that case, our model is as follows:

$$
dW_t = \varepsilon_t \sqrt{dt}
$$

$$
W_t = \int_0^t dW_s
$$

If we substitute the first equation above into the second, the integral looks somewhat strange due to the square root term. This integral is no longer a standard Riemann integral but is instead a stochastic integral. We treat these integrals in TN96-05, but as it turns out, with constant variance, a stochastic integral of this form is, more or less, the same as a Riemann integral. That enables us to say quite comfortably that with $W_0 = 0$,

$$
W_t = \int_0^t dW_s = W_t - W_0.
$$

This result would appear to hold by definition, but a formal proof requires the tools of stochastic integration, which requires defining the integral, not in terms of a limit of an absolute difference,

[6] The formal proof of this result is a variation of the DeMoivre-LaPlace limit theorem, which proves that a binomial distribution converges to a normal distribution in the limit.

but rather the expectation of a mean squared difference. We do not need to concern ourselves with this complication right now.

The variable W_t is a Brownian motion. The requirements for a Brownian motion are

(1) The process starts at zero ($W_0 = 0$) and is continuous

(2) The variable is normally distributed with an expectation of zero and variance of t at time t.

(3) The increments, $W_{t+dt} - W_t = \Delta W_t$, are independent and normally distributed.

As noted, this process is also sometimes called a Wiener process, hence the W notation. Some experts make slight distinctions between Brownian motion and Wiener processes. Some even refer to W_t as the Brownian motion and dW_t as the Wiener process. Technically our discrete time model is not a Brownian motion; it is usually just referred to as a random walk.

Let us find the parameters of the process W_t as well as its increments. Remember that $\varepsilon_t \sim N(0,1)$.

$$E(dW_t) = E(\varepsilon_t \sqrt{dt}) = dtE(\varepsilon_t) = 0$$
$$Var(dW_t) = Var(\varepsilon_t \sqrt{dt}) = dtVar(\varepsilon_t) = dt$$
$$E(W_t) = E\left(\int_0^t dW_j\right) = \int_0^t E(dW_j) = 0$$
$$Var(W_t) = Var\left(\int_0^t dW_j\right) = \int_0^t Var(dW_j) = \int_0^t dj = t.$$

For the last line above, note that since the increments are independent, all covariances between dW_j and $dW_k = 0$, $j \neq k$.

Now suppose we are interested in the covariance between overlapping Brownian motions, W_s and W_t where $t > s$. We have

$$
\begin{aligned}
Cov(W_s, W_t) &= Cov\left(W_0 + \int_0^s dW_j , W_0 + \int_0^t dW_j \right) \\
&= Cov(W_0, W_0) + Cov\left(W_0, \int_0^t dW_j \right) + Cov\left(\int_0^s dW_j , W_0 \right) + Cov\left(\int_0^s dW_j , \int_0^t dW_j \right) \\
&= 0 + 0 + 0 + Cov\left(\int_0^s dW_j , \int_0^t dW_j \right) = Cov\left(\int_0^s dW_j , \int_0^s dW_j + \int_s^t dW_j \right) \\
&= Cov\left(\int_0^s dW_j , \int_0^s dW_j \right) + Cov\left(\int_0^s dW_j \int_s^t dW_j \right) = Cov\left(\int_0^s dW_j , \int_0^s dW_j \right) + 0 \\
&= \int_0^s Cov(dW_j, dW_j) = \int_0^s Var(dW_j) = \int_0^s dt = s.
\end{aligned}
$$

Because the correlation is defined as the covariance divided by the product of the standard deviations, the correlation between W_s and W_t would be

$$
Corr(W_s, W_t) = \frac{s}{\sqrt{s}\sqrt{t}} = \sqrt{s/t}.
$$

This seemingly minor result turns out to play a major role in compound option pricing models where it is necessary to determine a correlation for the underlying asset price on two dates, with the asset price driven by the value of W on the two dates.

Now let us look at the probability density for W_t. We know that in general, a normally distributed random variable, X, with mean : and variance Φ^2 has a density of

$$
f(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}.
$$

What we know about W_t is that it is normally distributed with : = 0 and Φ^2 = t. Thus, its density is

$$
f(W_t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}\left(\frac{W_t^2}{t}\right)}.
$$

We also know that dW_t is normally distributed with $\mu = 0$ and Φ^2 = dt. Thus, its density is

$$
f(dW_t) = \frac{1}{\sqrt{2\pi dt}} e^{-\frac{1}{2}\left(\frac{dW_t}{\sqrt{dt}}\right)^2}.
$$

Since we know that dW_t is $\varepsilon_t\sqrt{dt}$, we can substitute into the above and obtain:

$$f(dW_t) = \frac{1}{\sqrt{2\pi dt}} e^{-\frac{1}{2}\varepsilon_t^2}.$$

Now consider two separate Brownian motions, W_x and W_y. Let us examine a process that is a product of these two processes, specifically, $dW_x dW_y$. We know that the two increments are defined as:

$$dW_x = \varepsilon_x \sqrt{dt},$$
$$dW_y = \varepsilon_y \sqrt{dt}.$$

Now let us take the variance of their product. By definition $Var(dW_x dW_y) = E[(dW_x dW_y)^2] - (E[dW_x dWy])^2$. Then $E[dW_x dW_y] = E[\varepsilon_x\sqrt{dt}\varepsilon_y\sqrt{dt}] = dtE[\varepsilon_x\varepsilon_y]$. Using the definition of a covariance, $cov(\varepsilon_x\varepsilon_y) = E[\varepsilon_x\varepsilon_y] - E[\varepsilon_x]E[\varepsilon_y]$, which here reduces to $E[\varepsilon_x\varepsilon_y]$ because the expectation of each ε is zero. Now consider the correlation between the processes ε_x and ε_y, which shall be denoted as Δ_{xy}. Because correlation is the covariance divided by the product of the standard deviations, then $cov(\varepsilon_x\varepsilon_y) = \Delta_{xy}$, since the two standard deviations are each 1.0. So $E[\varepsilon_x\varepsilon_y] = \Delta_{xy}$.[7] So $E[dW_x dW_y] = \Delta_{xy}dt$. But we want the expectation of the squared value of this. Obviously this is zero since we have $\Delta_{xy}^2 dt^2$ and dt^2 goes to zero. It follows that the second term in the variance definition, $(E(dW_x dW_y))^2$, goes to zero since it is $(\rho_{xy}dt)^2$. Thus,

$$Var(dW_x dW_y) = 0.$$

If the variance is zero, then

$$E(dW_x dW_y) = dW_x dW_y = \Delta_{xy}dt.$$

Mathematicians often refer to the derivative with respect to time, in this case dW_t/dt, as the *velocity*.[8] For Brownian motion, however, the velocity does not exist, as shown below:

$$\frac{dW_t}{dt} = \lim_{\Delta t \to 0} \frac{\Delta W_t}{\Delta t} = \lim_{\Delta t \to 0} \frac{\varepsilon_t \sqrt{\Delta t}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\varepsilon_t}{\sqrt{\Delta t}} \to \infty.$$

The intuition is that Brownian motion is characterized by infinitesimally small zig-zags. We cannot take limits while permitting the time increment to shrink. For a derivative to exist, we must be able to take a limit, meaning that the line drawn tangent to the point where we are taking

[7]The covariance is actually the product of the correlation and the two standard deviations of the ε's, which are both 1.0.

[8]The term *velocity* refers to the speed of something. Hence, a derivative such as dW_t/dt would refer to the rate at which W_t

the derivative must converge to a stable value, and this simply does not occur here. This result is, however, not a problem. We shall never need the derivative dW_t/dt.

Perhaps the most important characteristic of the process W_t is the property that its squared increment is no longer stochastic. Let us see how this happens. Consider the squared variable dW_t^2. Let us take its expectation:

$$E(dW_t^{\,2}) = E((\varepsilon_t \sqrt{dt})^2) = dtE(\varepsilon_t^{\,2}) = dt,$$

which results from the fact that $Var(\varepsilon_t) = E(\varepsilon_t^2) - (E(\varepsilon_t))^2 = 1$. But $E(\varepsilon_t)$ is zero. So $E(\varepsilon_t^2) = Var(\varepsilon_t) = 1$.

Now let us take the variance of dW_t^2:

$$\begin{aligned} Var(dW_t^{\,2}) &= E((dW_t^{\,2})^2) - (E(dW_t^{\,2}))^2 \\ &= E((\varepsilon_t^{\,2} dt)^2) - (E(\varepsilon_t^{\,2} dt))^2 \\ &= dt^2 E(\varepsilon_t^{\,4}) - dt^2 \\ &= 0. \end{aligned}$$

A key element of this result is remembering that in the limit $dt^k \to 0$ for all $k > 1$.

We shall eventually be interested in generalizing our Brownian motion so that it has a non-zero mean and a variance other than t. Technically, this would no longer be a Brownian motion, but it is often still referred in that manner. Suppose we wish to create a stochastic process, X_t, in which the increment, dX_t, has mean : and variance Φ. Then we simply do the following:

$$dX_t = :dt + \Phi dW_t.$$

The properties of this process are

$$E(dX_t) = :dt$$

$$Var(dX_t) = \Phi^2 dt.$$

The latter result arises because the variance of ΦdW_t is the constant Φ^2 times the variance of dW_t, which is dt. The process, X_t, defined by the stochastic integral

$$X_t = X_0 + \int_0^t dX_j,$$

would have the properties

changes with time, which of course is the speed.

$$E(X_t) = X_0 + E\left(\int_0^t dX_j\right) = X_0 + \int_0^t E(dX_j) = X_0 + \int_0^t \mu dj = X_0 + \mu t$$

$$Var(X_t) = Var\left(\int_0^t dX_j\right) = \int_0^t Var(dX_j) = Var\left(\int_0^t \sigma^2 dj\right) = \sigma^2 t.$$

Note that there are no covariance terms for dX_t in the variance expression in the last line above.

These results give X_t more general characteristics and enable us to use it to model more realistic phenomena.

Changing the Probability Measure with the Radon-Nikodym Derivative in Discrete Time

In this teaching note, we started with a discrete time Brownian motion model. We then extended it to the continuous time case. Before leaving, we need to return to the discrete time case and examine how to change a probability measure. This procedure is extremely important, but is much more difficult in the continuous time case, and we take it up in a separate note, TN 00-04. This material draws heavily on the excellent treatment in Baxter and Rennie (1996).

So let us again go back to the simple world with two outcomes. We can generalize it a little bit without any problems by having our variable W_0 move to W_1^+ or W_1^- with probabilities p and 1 - p, respectively. Now let us suppose that we are interested in changing the probabilities to q and 1 - q. Without further study, one may wonder why we would want to do this or even whether we could do this. Probabilities are, after all, usually given by external phenomena, and we cannot often change them.[9] But assigned probabilities can actually sometimes be changed. As we advance in the study of derivative pricing theory, we shall see that we can create artificial probabilities that lead to correct prices of derivatives but do not require knowledge of the true probabilities, the expected returns of assets, or the utility preferences of investors. This procedure in continuous time, however, is quite complex. We illustrate it initially in the discrete time setting.

So what we have is

W_1^+ with probability p

W_0

W_1^- with probability 1 - p

and what we want is

$$W_1^+ \text{ with probability } q$$

$$W_0$$

$$W_1^- \text{ with probability } 1 - q.$$

Let us define the ratio q/p and (1 - q)/(1 - p) as φ^s where s indicates the state + or -. That is,

$$\varphi^+ = q/p,$$

$$\varphi^- = (1 - q)/(1 - p).$$

The probabilities q and p represent different probability measures. The probabilities p and 1 - p are said to be measure P while q and 1 - q are measure Q. We can take the expected value of W_1 under either measure P or measure Q. Under measure Q, we have, by definition,

$$E^Q(W_1) = qW_1^+ + (1 - q)W_1^-.$$

This statement can be re-written as follows:

$$E^Q(W_1) = p(q/p)W_1^+ + (1 - p)((1 - q)/(1 - p))W_1^+.$$

Therefore, we can write this expectation compactly as,

$$E^Q(W_1) = E^P(\varphi^s W_1).$$

In other words, we can take the expectation under the Q measure by taking the expectation under the P measure, provided we adjust the random process by a new specific stochastic process involving a ratio of probabilities. So φ^s, which is a ratio of probabilities, is a stochastic process itself.

Let us extend this result one more period. Now the process we have is

$$W_2^{++} \text{ with probability } p^2 \text{ or } q^2$$

$$W_1^+ \text{ with probability p or q}$$

$$W_0 \qquad\qquad W_2^{+-} \text{ with probability } 2p(1-p) \text{ or } 2q(1-q)$$

$$W_1^- \text{ with probability } 1 - p \text{ or } 1 - q$$

$$W_2^{--} \text{ with probability } (1 - p)^2 \text{ or } (1 - q)^2$$

Now we must index our ratio of probabilities by time, i.e, φ_t^s. First note that the process starts at $\varphi_0 = 1.0$.[10] The stochastic process of φ_t^s is as follows:

$$\varphi_1^+ = p(q^2/p^2) + (1 - p)q(1 - p)/p(1 - p) = q^2/p + q(1 - q)/p = (q^2 + q - q^2)/p = q/p$$

$$\varphi_1^- = p(1 - q)q/(1 - p)p + (1 - p)(1 - q)^2/(1 - p)^2$$

[9]Consider, however, that Blackjack dealers frequently change decks to alter the probabilities.

[10]At time 0, there is only one state so we do not need to index it by state.

$$= (1 - q)q/(1 - p) + (1 - q)^2/(1 - p) = (1 - q)/(1 - p)$$

The above shows how φ_t is a stochastic process. At time 1 in the + state, there are two upcoming outcomes: φ will equal either q^2/p^2 with probability p or q(1 - q)/p(1 - p) with probability q. A similar statement applies in the - state. So we see that φ_t^s is a stochastic process itself with values

$$\varphi_1^+ = q/p$$

$$\varphi_0 = 1$$

$$\varphi_1^- = (1 - q)/(1 - p)$$

where each value is the expectation of what it will be in the next period. In general

$$\varphi_t = E^P(\varphi_t^s).$$

This process φ_t^s is referred to as the Radon-Nikodym derivative and represents a relationship between two equivalent probability measures. Equivalent probability measures are two probability measures that meet certain requirements, the main one being that any event that is possible under one measure must be possible under the other and vice versa. In other words, the impossible cannot be created from the possible and the impossible cannot be destroyed from the possible. In continuous time, this concept is much more difficult, except that it is more easily seen for what it is, a derivative of one probability distribution with respect to another. In discrete time it is merely a ratio or probabilities, but derivatives are technically not defined in discrete time.[11]

As noted, see TN00-04 for more on the Radon-Nikodym derivative and how it is used in derivative pricing.

The Kolmogorov Equations

There is a useful set of equations that appears in the analysis of movements in a stochastic process, which are called the Kolmogorov equations. These equations are slight variations of a differential equation that is well-known in derivative pricing theory, but instead of relating prices to their derivatives, the Kolmogorov equations relate probabilities to their derivatives. There are two such equations, the forward equation and backward equation.

[11]For a quick peek at the continuous time result, consider a probability measure Q and an equivalent measure P. The density functions are $f^Q(x)$ and $f^P(X)$. These are by definition, $f^Q(X) = dQ/dX$ and $f^P(X) = dP/dX$. Then $(dQ/dX)/(dP/dX) = f^Q(X)/f^P(X)$. This is our φ variable, as φ is defined such that $\varphi df^P(X) = df^Q(X)$, i.e., multiplying the density for P by φ gives us the density for Q. In other words, $\varphi = dQ/dP$.

Consider the following in a discrete time setting. Suppose the random variable of interest starts off in state j at time 0. We are interested in the probability that it is in state k at time n. In the two-state discrete world, the only way it could get to state k at time n is to be at state k+1 at time n-1 and move down with probability ½ or to be in state k-1 at time n-1 and move up with probability ½. If the probability of moving from state j at time 0 to state k+1 at time n-1 is $p_{j,k+1,n-1}$ and the probability of moving from state j at time 0 to state k-1 at time n-1 is $p_{j,k+1,n-1}$, then the probability of being in state k at time n is

$$p_{j,k,n} = p_{j,k+1,n-1}(1/2) + p_{j,k-1,n-1}(1/2).$$

If we express this probability a little more generally as $p(x,t;x_0)$, which is the probability of being in state x at time t, given that we were in state x_0 at time 0, then we can write the above as

$$p(x,t;x_0) = p(x-\Delta x,t-\Delta t;x_0)(1/2) + p(x+\Delta x,t-\Delta t;x_0)(1/2),$$

where we are using Δt to represent time steps and Δx to represent state movements. If we expand $p(x,t;x_0)$ in a Taylor series, we obtain

$$p(x+\Delta x,t-\Delta t;x_0) = p(x,t;x_0) - \frac{\partial p}{\partial t}\Delta t + \frac{\partial p}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2 p}{\partial x^2}\Delta x^2 + \ldots,$$

where we have suppressed some of the arguments on the probability for simplicity. If the parameters Δt and Δx are properly specified so that the process accurately reflects the parameters of the desired process, we obtain a continuous time representation of a partial differential equation for the probability density by letting Δt approach zero:

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial x^2}.$$

This is called the *forward equation*. It relates the probability at a forward time to prior paths. The forward equation is a partial differential equation with solution the probability density,

$$p(x,x_0;t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}\frac{(x-x_0)^2}{t}}.$$

This is also recognized as the density for a variable x with mean x_0 and variance t.[12]

Now let us look briefly at the backward equation. Let $p_{j,k,n}$ be the probability of going from state j at time 0 to state k in n steps. Let $p_{j-1,k,n-m}$ be the probability of going from state j-1

[12]That this density solves the differential equation can be verified by taking the partial derivatives with respect to t and x and substituting back into the differential equation.

to k in m steps and $p_{j+1,k,m}$ be the probability of going from state j+1 to state k in m steps. Given that the probability is ½ that we went up to state j+1 at time n-m and ½ that we went down to state j-1 at time n-m, then the probability of going from state j to state k in m+1 steps is

$$p_{j,k,m+1} = p_{j+1,k,m}(1/2) + p_{j-1,k,m}(1/2).$$

Following a similar procedure, we obtain the continuous time backward equation

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial {x_0}^2},$$

where we note that the difference with the forward equation is that here we are viewing the problem from the current position x_0. In the forward equation we are viewing the problem from the future position, x. In problems in finance, the backward equation is more commonly used, because we tend to be positioned at a current point in time, equipped with knowledge of the current value of an asset, and are looking toward a future point in time where the probability of being in a certain state is important for valuing an asset or derivative.

Both the forward and backward equations bear a strong resemblance to the partial differential equation that describes the price of a derivative. This is also commonly known as the heat transfer equation. In fact solving a derivative pricing problem, such as is done with the Black-Scholes model, can be done by transforming the partial differential equation into one almost identical to the one above.[13]

References

Au, K. T., R. Mahrendra and D. C. Thurston. "An Intuitive Explanation of Brownian Motion as a Limit of a Random Walk." *The Journal of Financial Education* 23 (Spring, 1997), 91-94.

Baxter, M. and A. Rennie. *Financial Calculus*. Cambridge: Cambridge University Press (1996), Chs. 2, 3.

Briys, E., M. Bellalah, H. M. Mai, and F. DeVarenne. *Options, Futures, and Exotic Derivatives* Chichester, U. K.: Wiley (1998), Ch. 2.

[13]Those familiar with the Black-Scholes partial differential equation will note that it contains a first-order partial derivative with respect to the underlying variable and a term representing the derivative price itself. These can be eliminated by suitable transformations.

Chance, D. M. "The ABCs of Geometric Brownian Motion." *Derivatives Quarterly* 1 (Winter, 1994), 41-47.

Cox, D. R. and H. D. Miller. *The Theory of Stochastic Processes*. New York: Wiley (1965), Ch. 5.

Karatzas, I. and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, 2nd. ed. Berlin: Springer-Verlag (1991), Chs. 2.

Neftci, S. *An Introduction to the Mathematics of Financial Derivatives*, 2nd. ed. San Diego: Academic Press (2000), Chs. 6, 8.

Nielsen, L. T. *Pricing and Hedging Derivative Securities*. Oxford, U.K.: Oxford University Press (1999), Ch. 1.