[Watch the latest videos on YouTube.com](#)

Subsections

- [A One-Dimensional Random Walk](#)
- [Multi-Dimensional Random Walks](#)
- [The Wiener Process (Brownian Motion)](#)
- [Properties of the Wiener Process](#)
    - [Scale Invariance](#)
    - [Expectation Values](#)
    - [White Noise](#)
- [Martingales](#)
- [The Langevin Equation](#)

Stochastic Processes : An Elementary Introduction

In this chapter, I will present the theory of stochastic processes in an elementary manner sufficient for understanding the theory presented in the following chapters. A reader interested in a more rigorous approach could consult Ross[4]. This chapter mostly follows Roepstorff[5] which has a more physical description of stochastic processes.

A One-Dimensional Random Walk

This is probably the simplest example of a discrete time stochastic process. One way of defining it would be :

$$P(X_{i+1} - X_i = 1) = P(X_{i+1} - X_i = -1) = \tfrac{1}{2}, \ i \in \mathbb{Z}_+$$ which simply means that the variable X (the subscript stands for the time) has an equal probability of increasing or decreasing by 1 at each time step. For historical reasons, the term Brownian motion is also used for random walks. This is usually described as a ``drunken ant'' with a clock and a coin. At each tick of the clock, the ant

tosses the coin. If it turns out heads, the ant moves one step towards the right and if it turns out tails it moves one step to the left.

This process is both homogeneous (since the transition probability is only dependent on the distance between the initial and final points) and isotropic (the transition probability is independent of the direction of movement).

The random walk can be taken to be a Markov chain with a transition matrix P (for one time step) where the element P_{ij} stands for the probability that an ant with initial position i will end up at j. For n time steps, the accumulated transition probability is given by P^n.

It can be easily seen that the following difference equation holds

$$P(X_{i+1} = x) = \tfrac{1}{2}\left(P(X_i = x+1) + P(X_i = x-1)\right) \tag{5}$$

This is equivalent to

$$P(X_{i+1} = x) - P(X_i = x) = \tfrac{1}{2}\left(P(X_i = x+1) - 2P(X_i = x) + P(X_i = x-1)\right) \tag{6}$$

Now, if we re-scale so that the step size is h and the time step is τ, the above becomes

$$\frac{P(x,t+\tau) - P(x,t)}{\tau} = \frac{h^2}{2\tau}\frac{P(x+h,t) - 2P(x,t) + P(x-h,t)}{h^2} \tag{7}$$

where I have used the more intuitive but less rigorous notation where $P(x,t)$ is the probability of the ant being at x at time t (it is important to note that these are still discrete variables).

In the limit $h \to 0, \tau \to 0$, $\frac{P(x,t+\tau)-P(x,t)}{\tau}$ becomes $\frac{\partial P(x,t)}{\partial t}$ (now, of course, x and t are continuous) and $\frac{P(x+h,t)-2P(x,t)+P(x-h,t)}{h^2}$ becomes $\frac{\partial^2 P(x,t)}{\partial x^2}$. Hence, in this limit, we get the following equation for the one-dimensional random walk

$$\frac{\partial P(x,t)}{\partial t} = D\frac{\partial^2 P(x,t)}{\partial x^2} \tag{8}$$

where D is $\lim_{h\to 0, \tau \to 0} \frac{h^2}{2\tau}$. Since $\frac{h^2}{\tau}$ tends to a finite limit when $h \to 0$ and

$\tau \to 0$, we see that the ``velocity" of the ant $\frac{h}{\tau} \to \infty$. This shows that the velocity of a particle undergoing a random walk is infinity [?].

Equation 2.4 is, of course, the diffusion equation which should not come as too much of a surprise as diffusion is the result of the random movement of molecules. The reader should also be familiar with the fact that the solution of the diffusion equation with the initial condition $P_0(x,0) = \delta(x)$ is given by the Gaussian

$$P_0(x,t) = \frac{1}{2\sqrt{\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right), t > 0 \tag{9}$$

Since the discrete random walk actually follows a binomial distribution, we can also think of this as following from the classical theorem of Laplace and De Moivre about the convergence of the binomial distribution towards the normal distribution.

We have not handled the transition from discrete variables to continuous variables rigorously but the steps above should be intuitively reasonable. To do the above rigorously, we only have to identify the discrete probability $P(X_i = x)$ with $\int_{x-\frac{h}{2}}^{x+\frac{h}{2}} P(x', i\tau)dx'$.

Throughout the above discussion, we have assumed that the probability of the ant moving left or right is the same (i.e. we have assumed that the walk is isotropic). This type of random walk is called the simple random walk. To generalize the above discussion, we have to change eqn. (2.1) to

$$P(x, t+\tau) = pP(x+h, t) + qP(x-h, t) \tag{10}$$

where p is the probability of moving to the left and q is the probability of moving to the right (obviously, $p+q=1$). The above analysis can be applied to this system provided the limit $\lim_{h\to 0, \tau\to 0} \frac{h}{\tau}(p-q) = v$ exists. The diffusion equation is then replaced by

$$\frac{\partial P(x,t)}{\partial t} = \left(D\frac{\partial^2}{\partial x^2} + v\frac{\partial}{\partial x}\right) P(x,t) \tag{11}$$

where v can be interpreted as a mean drift velocity. The reader should be easily able to verify that the solution for non-zero v is related to the solution of the simple random walk $P_0(x,t)$ by

$$P_v(x,t) = P_0(x-vt,t) \qquad (12)$$

Hence, any one-dimensional random walk is equivalent to a simple one-dimensional random walk under a Galilean transformation with velocity given by v.

Multi-Dimensional Random Walks

Let us suppose that the ant does not live in a one-dimensional world but in a m-dimensional world. At each step, the ant can move in any of $2m$ directions. If the ant now possesses a $2m$ sided fair dice and a clock, it can walk one step in any direction with equal probability. We again look at the limit where the step size h of the ant tends to zero.

The resulting paths have certain interesting geometrical properties. While these are not important for the purposes of this thesis, they are interesting in their own right. A few of them are described below :

- They are continuous everywhere but differentiable nowhere.
- They are recurrent (i.e. the paths will return to the origin with probability 1) for $m = 1, 2$ and transient (i.e there is a finite probability different

  from 1 that the paths will return to the origin) for $m \geq 3$.

- The path is fractal in nature.

For more interesting properties of random walks, please see section 1.2 of Roepstorff[5].

Equation ([2.3](#)) is now replaced by

$$\frac{P(\mathbf{x},t+\tau) - P(\mathbf{x},t)}{\tau} = \sum_{i=1}^{m} \frac{h^2}{2\tau m} \frac{P(x_i-h,t) - 2P(x_i,t) + P(x_i+h,t)}{h^2} \qquad (13)$$

(the subscripts refer to the components of $\mathbf{x}$) whose continuum limit is

$$\frac{\partial P(x,t)}{\partial t} = D\Delta P(x,t) \qquad (14)$$

where D is now $\lim_{h \to 0, \tau \to 0} \frac{h^2}{2\tau m}$ and Δ is the m-dimensional Laplacian. The

solution to this equation with the initial condition $P(x,0) = \delta(x)$ is given by

$$P_0(\mathbf{x},t) = (4\pi Dt)^{-m/2} \exp\left(-\frac{\mathbf{x}^2}{4Dt}\right) \qquad (15)$$

It is instructive to note that the mean square displacement

$$\langle \mathbf{x}^2 \rangle = \int d^m \mathbf{x} \, \mathbf{x}^2 P_0(\mathbf{x}, t) = 2mDt \tag{16}$$

is proportional to t. In other words, the distribution varies with time according to a $\sqrt{t}$ law.

The Wiener Process (Brownian Motion)

We are now in a position to discuss the Wiener process which is of fundamental importance in the theory of option pricing. This process was first discussed by Einstein in his description of Brownian motion and was first put into a rigorous form by Wiener.

For simplicity, we will assume $D = \frac{1}{2}$ in the rest of this discussion. This can always be achieved by rescaling the time and space coordinates.

Consider a particle which was initially at the origin performing a m-dimensional simple random walk. In other words, $\mathbf{x} = \mathbf{0} \in \mathbb{R}^m$ at time $t = 0$.

The position of the particle at time t can be considered as a random variable X_t (the term *random vector* might be more appropriate as it emphasis the fact that X_t has several components). Now, if we can find the probabilities $P(X_t \in A)$ for all $A \subset \mathbb{R}^m$ with non-zero measure, we would have described the path followed by the particle completely.

We have actually found the answer to this in eqn. (2.11). To make this clearer, we can express the answer in the following manner

$$P(X_t \in A) = \int_A K(\mathbf{x}, t) d^m \mathbf{x} \tag{17}$$

with density

$$K(\mathbf{x}, t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{\mathbf{x}^2}{2t}\right) \tag{18}$$

Hence, X_t follows a normal distribution.

A stochastic process is a map $t \mapsto X_t$, where t ranges over some interval ($[0,\infty)$ in this case). To properly define a stochastic process, we need to be able to determine the probabilities of general events. Before we can do this, it is necessary to consider compound events of the form `` $X_{t_1} \in A_1, X_{t_2} \in A_2, \ldots, X_{t_n} \in A_n$'', where $0 < t_1 < t_2 < t_3 < \cdots < t_n$, $A_i \subset \mathbb{R}^m$, $n > 0$ and to devise rules that determine their probability.

We denote the probability of a compound event by $P(X_{t_1} \in A_1, X_{t_2} \in A_2, \ldots, X_{t_n} \in A_n)$ Varying $A_1, A_2, A_3, \ldots, A_n$ we get the joint distribution of the random variables $X_t, t \in T = (t_1, t_2, \ldots, t_n)$. This is also called the distribution of the process with base T and may be abbreviated as $P_T(A)$ where $A = A_1 \times A_2 \times \cdots \times A_n \subset \mathbb{R}^{nm}$. The distribution is said to be finite-dimensional (of order n) since the base is finite (has n elements).

The stochastic process X_t is said to be the Wiener process[?] if the finite-dimensional distributions are of the form

$$\int_{A_n} d^m\mathbf{x}_n \cdots \int_{A_2} d^m\mathbf{x}_2 \int_{A_1} d^m\mathbf{x}_1 K(\mathbf{x}_n - \mathbf{x}_{n-1}, t_n - t_{n-1}) \ldots K(\mathbf{x}_2 - \mathbf{x}_1, t_2 - t_1) K(\mathbf{x}_1, t_1) \quad (19)$$

with $K(\mathbf{x}, t)$ given by (2.14) and if the initial distribution is

$$P(X_0 \in A_0) = \begin{cases} 1 & \text{if } 0 \in A_0 \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

Hence, the particle starts at the origin (in other words, $X_0 = 0$ with certainty).

We can rewrite eqn. (2.15) as

$$P(X_{t_1} \in d\mathbf{x}_1, \ldots, X_{t_n} \in d\mathbf{x}_n) = \prod_{k=1}^{n} d^m\mathbf{x}_k \, K(\mathbf{x}_k - \mathbf{x}_{k-1}, t_k - t_{k-1}) \quad (21)$$

($\mathbf{x}_0 = 0, t_0 = 0$) with density $K(\mathbf{x}, t)$ given by (2.14).

The fact that the right hand side of (2.17) is a product tells us that the Wiener process is a Markov (memory-independent) process. This is reassuring as the present state then contains all the information that is relevant for the future

which we have seen is true for an efficient market.

Properties of the Wiener Process

Scale Invariance

By noting that $K(\mathbf{x}, t)$ is only dependent on the combination $\frac{\mathbf{x}^2}{t}$, we can see that $Y_t = lX_{t/l^2}, l > 0$ defines another Wiener process which is indistinguishable from X_t except for the scale. In other words, the Wiener process is scale invariant (provided we re-scale time accordingly) or self-similar. This follows from the fractal nature of Brownian motion.

Expectation Values

Since any probability distribution is normalized, the zeroth moment of $K(\mathbf{x}, t)$ is 1. Further, the first moment is zero as $K(\mathbf{x}, t) = K(-\mathbf{x}, t)$. We have already found the second moment in (2.12). All the above statements can be rephrased in terms of expectation values. We see that

$$E(X_t) = \int d^m\mathbf{x}\, \mathbf{x} K(\mathbf{x}, t) = \mathbf{0} \tag{22}$$

This might seem strange in terms of the Markovian property of the Wiener process as the origin seems to be special. However, $E(X_t)$ ought to be interpreted as the conditional expectation of X_t given the information that $X_0 = \mathbf{0}$ (the initial condition). Hence, the origin plays a distinguished role only by convention.

For the rest of this discussion, we will restrict ourselves to the case $m = 1$. The generalization to higher dimensions is self-evident.

We will be particularly interested in the following expected value

$$E(X_t X_{t'}) = G(t, t') \tag{23}$$

where $G(t, t')$ is called the covariance matrix of the process. We can calculate $G(t, t)$ using (2.12). Since we are assuming that $D = \frac{1}{2}$, we obtain the simple result $G(t, t) = t$. Further G is symmetrical by definition. To calculate

$G(t,t'), 0 \le t \le t'$, we consider the two increments $X_t - X_0$ and $X_{t'} - X_t$ which are independent with zero mean. Hence, $E((X_t - X_0)(X_{t'} - X_t)) = 0$. We also note that $E(X_0) = 0$. Finally, we write

$$E(X_t X_{t'}) = E((X_t - X_0)(X_{t'} - X_t)) + E(X_t^2) = t \tag{24}$$

Hence, we obtain the final result

$$G(t,t') = \min(t,t') \tag{25}$$

(the min arises due to the assumption $t \le t'$ which can be made without loss of generality as G is symmetric).

White Noise

Now, we are in a position to define white noise as a time derivative of the Wiener process. We note that

$$\frac{\partial^2}{\partial t \partial t'} \min(t,t') = \delta(t - t') \tag{26}$$

Hence, $W = \dot{X}_t$ [?] has covariance

$$E(W_t W_{t'}) = \delta(t - t'). \tag{27}$$

W is termed white noise which emphasizes that the Fourier transformed covariance is constant.

A precise meaning can be given to W_t by considering the concept of generalized stochastic processes. A good discussion of this concept can be found in Roepstorff[5].

Martingales

A martingale is an extension of the concept of a fair game. Let us assume that we have a gambler who tosses a coin at each time step. If he calls correctly, he gets $1 and loses $1 otherwise. If we represent the winnings at time step i by Y_i, then $P(Y_i = 1) = P(Y_i = -1) = \frac{1}{2}$. If X_i represents a random variable representing the amount of money that a gambler has at time step i, then the the expected value $\langle X_{i+1} \mid X_1 = x_1, X_2 = x_2, \ldots, X_i = x_i \rangle$ is x_i. So, the

gambler has zero expected gain in each time step. This is exactly what is meant by a martingale.

A discrete stochastic process $\{X_n, n \geq 0\}$ is said to be a martingale with respect to a process $\{Y_n, n \geq 0\}$ if, for all $n \geq 0$,

$$E[|X_n|] < \infty \qquad \text{and} \qquad E[X_{n+1} \mid Y_0, \ldots, Y_n] = X_n \qquad (2.1)$$

$\{X_n\}$ is called a sub-martingale with respect to $\{Y_n\}$ if, for all $n \geq 0$, X_n is a function of $(Y_0, \ldots, Y_n)$ and

$$E[|X_n^+|] < \infty \qquad \text{and} \qquad E[X_{n+1} \mid Y_0, \ldots, Y_n] \geq X_n \qquad (2.2)$$

where $X_n^+ = \max(0, X_n)$.

$\{X_n\}$ is called a super-martingale with respect to $\{Y_n\}$ if, for all $n \geq 0$, X_n is a function of $(Y_0, \ldots, Y_n)$ and

$$E[|X_n^-|] < \infty \qquad \text{and} \qquad E[X_{n+1} \mid Y_0, \ldots, Y_n] \leq X_n \qquad (2.3)$$

where $X_n^- = \min(0, X_n)$.

While a martingale describes a fair game, the sub-martingales and super-martingales describe favourable and unfavourable games respectively. If $\{X_i\}$ is a martingale, $E[X_{n+k}] = E[X_n], \forall k \in \mathbb{Z}_+$, while $E[X_{n+k}] \geq E[X_n], \forall k \in \mathbb{Z}_+$ if $\{X_i\}$ is a sub-martingale and $E[X_{n+k}] \leq E[X_n], \forall k \in \mathbb{Z}_+$ if $\{X_i\}$ is a super-martingale.

Some simple examples of martingales are

1. If $Y_0 = 0$ and $\{Y_n, n \geq 1\}$ is a sequence of independent centered random variables (i.e. $E[|Y_n|] < \infty$ and $E[Y_n] = 0$), then $\{X_n\}$ is a martingale with respect to $\{Y_n\}$ where $X_0 = 0$ and $X_n = \sum_{i=1}^{n} Y_i$.

2. If $\{Y_n, n \geq 1\}$ is a sequence of independent random variables with $E[|Y_n|] < \infty$ and $E[Y_n] = m_n \neq 0$ for all $n \geq 1$, then $\{X_n\}$ is a martingale with respect to $\{Y_n\}$ where $X_n = \prod_{i=1}^{n} (Y_i/m_i)$.

The first example is an important, if simple one. It shows that any discrete, and hence the continuous, random walk is a martingale.

Martingales are extremely important in finance due to the concept of risk-neutral valuation. This is due to the fact that the expected growth rate of all securities in a risk-neutral world is the risk-free interest rate. Hence, $e^{-rt}S$ is

a martingale for all securities S in a risk-neutral world. This is why risk-neutral valuation approach is also referred to as using an equivalent martingale measure.

Martingales have several interesting properties and several important limit theorems about them can be proven. These are beyond the scope of this thesis and the interested reader should consult a good textbook on stochastic processes such as Ross[4].

The Langevin Equation

A physical way of discussing stochastic processes is through the Langevin equation. The historical impetus for this equation arose from Einstein's description of Brownian motion.

Langevin considered the equation of motion of a particle in a fluid which is classically given by

$$M\frac{dv}{dt} + \gamma v = 0 \tag{28}$$

where γ is the coefficient of friction. He considered this equation as correct only for the average motion of the particle. In that case, the equation would correctly describe the motion for relatively massive particles (since the random disturbances would be too small to disturb them) but would only describe the long-term motion of lighter particles. Since Brownian motion only occurs for light particles such as pollen grains, this is quite reasonable. Hence, he generalized this equation to

$$M\frac{dv}{dt} + \gamma v = F(t) \tag{29}$$

where $F(t)$ is a stochastic process with zero mean and covariance

$$E[F(t)F(s)] = 2D\delta(t-s) \tag{30}$$

The reader should be able to see that equation (2.28) can also be written as

$$dv = -\frac{\gamma v}{M}dt + \frac{\sqrt{2DW}}{M}dt \tag{31}$$

where W is white noise. We can solve this equation by transforming to the variable $x = ve^{\gamma t/M}$. We obtain

$$dx = \frac{\partial x}{\partial t}dt + \frac{\partial x}{\partial v}dv + \frac{D}{M^2}\frac{\partial^2 x}{\partial v^2}dt \tag{32}$$
$$= \frac{\sqrt{2D}e^{\gamma t/M}}{M}W\,dt$$

which can be easily solved to obtain $x \sim N(v_0, \frac{D}{M\gamma}(e^{2\gamma t/M} - 1))$ which gives

$$v \sim N\left(v_0 e^{-\gamma T/M}, \frac{D}{M\gamma}\left[1 - e^{-2\gamma t/M}\right]\right) \tag{33}$$

We now solve the Langevin equation formally and check that the solution gives us the same expected value and variance as the solution above. The formal solution for the Langevin equation gives

$$v = v_0 e^{-\gamma t/M} + \frac{1}{M}\int_0^t e^{-\frac{\gamma}{M}(t-\tau)}F(\tau)d\tau \tag{34}$$

so that

$$E[v(t)] = v_0 e^{-\gamma t/M} + \frac{1}{M}\int_0^t e^{\frac{\gamma}{M}(t-\tau)}E[F(\tau)]d\tau = v_0 e^{-\gamma t/M} \tag{35}$$

and

$$E\left[\left(v(t) - v_0 e^{-\gamma t/M}\right)^2\right] = \left(\frac{1}{M}\right)^2\int_0^t\int_0^t e^{-\frac{\gamma}{M}(t-\eta)}e^{-\frac{\gamma}{M}(t-\tau)}E[F(\eta)F(\tau)]d\eta\,d\tau$$
$$= \frac{D}{\gamma M}\left(1 - e^{-2\gamma t/M}\right) \tag{36}$$

Hence, we see that the expectation and the variance calculated using both the methods are the same as should be the case.

As $t \to \infty$, the particle attains equilibrium with its surroundings. Hence, the velocity distribution should be Maxwellian

$$P(v) = \left(\frac{M}{2\pi kT}\right)^{1/2}\exp\left(-\frac{Mv^2}{2kT}\right) \tag{37}$$

which when compared to the solution above yields $D = kT\gamma$ which is the Einstein relation.

The stochastic differential equation for the logarithm of the stock price

$$dx = (r - \frac{\sigma^2}{2})dt + \sigma dz \qquad (38)$$

where z is a Wiener process can also be considered as a Langevin equation

$$\frac{dx}{dt} = r - \frac{\sigma^2}{2} + \frac{\sigma}{\sqrt{2D}}F(t) \qquad (39)$$

(which, after changing variables to $x - \left(r - \frac{\sigma^2}{2}\right)t$ and setting $\gamma = 0$ (implying zero viscosity) and $D = \frac{M^2\sigma^2}{2}$, is the same as the classical Langevin equation)

Equation (2.37) can be readily solved to yield $x \sim N\left(\left(r - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right)$. We can formally solve equation (2.38) as above to obtain

$$x = \left(r - \frac{\sigma^2}{2}\right)t + \frac{\sigma}{\sqrt{2D}}\int_0^t F(\tau)d\tau \qquad (40)$$

from which we get

$$E[x] = \left(r - \frac{\sigma^2}{2}\right)t + \frac{\sigma}{\sqrt{2D}}\int_0^t E[F(\tau)]d\tau = \left(r - \frac{\sigma^2}{2}\right)t \qquad (41)$$

and

$$E\left[\left(x - \left(r - \frac{\sigma^2}{2}\right)t\right)^2\right] = \frac{\sigma^2}{2D}\int_0^t\int_0^t E[F(\tau)F(\eta)]d\tau\,d\eta = \sigma^2 t \qquad (42)$$

The main importance of the Langevin equation is that it gives us a different way of considering stochastic processes. They can also sometimes suggest better methods of solving stochastic differential equations. Mathematically, the two approaches are, of course, equivalent.

Marakani Srikant 2000-08-15