Plan Diffusions and Brownian Motion

1. Finish Point Processes
2. The Botanists Tale
3. Brownian Motion
4. Properties
5. Diffusions

Reading: G&S: 13.1–13.3
Next Time: We'll see how far we get.

Homework 9 available next time

Question 1. Suppose that Pittsburgh city buses arrive a bus stop every $\Delta > 0$ minutes and that a passanger arrives at the bus stop at some random point uniformly distributed over the day. How long on average does the passenger wait?

Suppose instead that the arrival times of the busses at the stop are described by a Poisson random measure on the line with mean measure $\nu = \text{Leb}/\Delta$. How long on average does the passenger wait?

What's going on here?

Question 2. Suppose the positions of trees in a forest are well characterized by a Poisson random measure N with mean measure $\nu = \lambda \text{Leb}$ for $\lambda > 0$. What is the distribution of the distance from an arbitrary point in the forest to the nearest tree? What is the distance from that point that you can see to the east?

Example 3. Let M be a non-negative integer-valued random variable with distribution G. Let μ be a probability measure on a space $\mathcal{S}$.

Given any $k-1$ disjoint sets $A_1, \ldots, A_{k-1}$, let $A_k = (\overset{k-1}{\underset{1}{\cup}} A_j)^c$. Let

$$\mathsf{P}\{N(A_j) = n_j, j = 1, \ldots, k\} = \binom{M}{n_1, \ldots, n_k} \mu^{n_1}(A_1) \cdots \mu^{n_k}(A_k). \tag{1}$$

What is the marginal distribution of $N(A_1)$ here? Note that $M = N(\mathcal{S})$.

We care not just about the distributions of $N(A)$ but also moments. For example, if A_1 and A_2 partition $\mathcal{S}$:

$$\mathsf{E}(N(A_1)N(A_2) \mid M) = M(M-1)\mu(A_1)\mu(A_2). \tag{2}$$

So $\mathsf{Cov}(N(A_1), N(A_2)) = c_{[2]}\mu(A_1)\mu(A_2)$, where $c_{[2]}$ is second factorial cumulant of M.

These distributions and moments are easiest to study with generating functions.

Practical Motivation 4. There are two threads to keep hold of when delving into the theory of point processes. The first is that we want to be able to characterize the implications of our assumptions for the structure of the point scatter. The abstract tools are designed to capture different features of this structure to help us understand it better. The second thread is that in practice we are faced with point scatters with different types of structure, and we want to be able to *generate* models that match what we observe.

The following properties, which will be more fully developed in homework, give insight into different features of the process.

Stationarity 5. A point process N is *stationary* if for all bounded, measurable sets $A_1, \ldots, A_k$ for $k \geq 1$, the joint distribution of $N(A_1 + t), \ldots, N(A_k + t)$ does not depend on t over all t such that the shifted sets $A_i + t$ remain in the domain of N.

Question 6. Given an example of a stationary point process.

Definition 7. Covariance Measure.

Let N be a point process with mean measure ν, and define

$$C_2(A, B) = \mathsf{Cov}(N(A), N(B)). \tag{3}$$

What does this mean? Can we simplify this in the stationary case?

Definition 8. Avoidance Measure.

Let N be a point process. Write

$$V(A) = \mathsf{P}\{N(A) = 0\}. \tag{4}$$

What is the avoidance measure for the Poisson random measure?

Definition 9. The K function.

The K-function $K(t)$ for a stationary point process is defined as the expected number of points within a distance t of a typical point, divided by the overall intensity of the process.

For a homogeneous Poisson scatter in d dimensions, $K(t) = v_d t^d$, where v_d is the volume of the unit ball in d dimensions.

Idea 10. Integrating Random Measures

We have seen that with a measure, one can compute the measure of sets *or* integrate functions and that the latter is a more general representation.

If M is a random measure, then $\int f\, dM$ is a random variable which gives for each $\omega \in \Omega$, $\int f\, dM(\omega, \cdot)$.

It is often useful to represent random measures via their integrals.

Definition 11. Laplace Functionals If M is a random measure, then

$$L[f] = \mathsf{E} e^{-\int f\, dM}, \tag{5}$$

for non-negative, measurable functions f, is called the Laplace functional of the process.

The Laplace functional determines the distribution of the process.

Question 12. What is the Laplace functional of the Poisson random measure?

Definition 13. Probability Generating Functional

Let N be a point process and define

$$G_N[h] = \mathsf{E} e^{\int \log h\, dN} = \mathsf{E} \prod_i h(X_i), \tag{6}$$

where $0 \le h \le 1$ is a function with $1 - h$ vanishing outside a compact set and where the latter product is over the "points" of the process.

This also determines the distribution of the process.

Reminder 14. What does "distribution of the process" of the process mean?

Brief History 15. The Botanist's Tale

- Robert Brown (1773-1858), botanist
  Examined pollen grains and moss spores in water under microscope,
  Observed a jittery motion. This had been observed before and attributed to biological causes. But when he looked at dust suspended in water, he saw the same motion, which ruled out the biological. Brown didn't explain the phenomena but noted the irregularity of paths and the seeming independence of distinct particle paths. Others had made similar (though less systematic) observations before.
- Many attempts to explain the phenomenon in physical terms, including by Bachelier, Boltzmann, and Gibbs.
  Bachelier, in particular, attempted to describe fluctuations in stock prices mathematically in 1900 and might have had priority on some of Einstein's 1905 results.
- Einstein, apparently unaware of the earlier attempts, predicted the phenomenon on theoretical grounds: according to the kinetic theory of gases (fluids), water molecules undergo repeated random collisions, so a small object in water would experience impacts of random size and direction.
  Einstein's "major aim was to find facts which would guarantee as much as possible the existence of atoms of definite finite size." At the time, the atomic theory remained controversial.
  Einstein formulated a quantitative model as follows. Let $p(x,t)$ be the probability density near position x and time t for a "Brownian" particle. Einstein made some intuitive probabilistic assumptions:

  A. The probability that the particle moves from x to $x+h$ in a small time τ is $g(h,\tau)\,dh$.
  B. $g(h,\tau) = g(-h,\tau)$, so $\int hg(h,\tau)\,dh = 0$
  C. $\int h^2 g(h,\tau)\,dh = D\tau$ for some $D > 0$.

  It follows that in one dimension

$$p(x,t+\tau) = \int p(x-y,t)g(y,\tau)\,dy \tag{7}$$

$$= \int \left(p(x,t) - \frac{\partial p(x,t)}{\partial x} y + \frac{1}{2}\frac{\partial^2 p(x,t)}{\partial x^2} - \cdots \right) g(y,\tau)\,dy \tag{8}$$

$$\approx p(x,t) + \frac{D}{2}\frac{\partial^2 p(x,t)}{\partial x^2}. \tag{9}$$

  The same argument in three dimensions gives the diffusion equation

$$\frac{\partial p}{\partial t} = \frac{D}{2}\Delta p(x,t), \tag{10}$$

  where Δ is the Laplacian and D is a positive constant – the coefficient of diffusivity. If we take the initial position of the particle to be 0, then

$$p(x,t) = (2\pi Dt)^{-3/2} e^{\frac{-|x|^2}{2Dt}}, \tag{11}$$

  a Gaussian$(0, Dt)$ distribution on $\mathbb{R}^3$.
  The second part of Einstein's argument is physical and intuitive. He relates D to other physical quanties such as the Boltzmann constant k, the temperature T, the particle's mass m, and another constant β with dimensions of frequencies.

$$D = \frac{kT}{m\beta} = \frac{RT}{N\phi}, \tag{12}$$

where R is the gas constant, N is Avogadro's number, and ϕ is another constant. If the Brownian particles are spheres of radius r, then Stoke's theory of friction gives $m\beta = 3\pi\eta r$, where η is the coefficient of viscosity of the fluid and

$$D = \frac{kT}{3\pi\eta r}. \tag{13}$$

Both the temperature and viscosity can be measured and a population of uniform particles created, which is one way to determine Boltzmann's constant.

- Langevin, Smoluchowski, Ornstein, and Uhlenbeck built on this to develop a dynamical theory of diffusion.
- Weiner systematized the basic mathematical theory of Brownian motion – which is often called a Weiner process and denoted by W – and prepared the way for the theory of stochastic differential equations.

Process 16. Gaussian Processes

We know what it means for a scalar random variable to have a Normal (i.e., Gaussian) distribution.

A finite collection $X_1, \ldots, X_n$ has a (joint) Gaussian distribution if $\sum_{i=1}^n a_i X_i$ is a scalar Gaussian random variable for all choices of $a_1, \ldots, a_n \in \mathbb{R}$.

We know that a scalar Gaussian is determined by its mean and variance. Note that

$$\mathsf{E} \sum_i a_i X_i = \sum_i a_i \mathsf{E} X_i \tag{14}$$

$$\mathsf{Var} \sum_i a_i X_i = \sum_i \sum_j a_i a_j \,\mathsf{Cov}(X_i, X_j). \tag{15}$$

Hence, the mean vector $(\mathsf{E}X_1, \ldots, \mathsf{E}X_n)$ and covariance matrix $(\mathsf{Cov}(X_i, X_j))_{1 \le i,j \le n}$ determine the distribution of the finite collection.

What about an infinite collection?

A real-valued stochastic process $X = (X_t)$ is said to be a *Gaussian process* if all its finite-dimensional distributions are Gaussian.

That is, X is Gaussian if for all choice of $n \ge 1$ and $t_1, \ldots, t_n$, the finite collection $X_{t_1}, \ldots, X_{t_n}$ is Gaussian. The mean vector and covariance matrix of this collection can depend on the ts.

The mean vector of the finite dimensional distributions of X is determined by the *mean function* $\mu(t) = \mathsf{E}X_t$. The covariance matrix of the finite dimensional distributions of X is determined by the *autocovariance function* $\rho(s,t) = \mathsf{Cov}(X_s, X_t)$.

Thus, the "distribution" of a Gaussian process is determined by μ and ρ.

ρ must be a positive definite function: $\sum_{j,k} \rho(t_j, t_k) z_j \overline{z}_k \ge 0$ for all $n \ge 1$, $t_1 < \cdots < t_n$, and complex numbers $z_1, \ldots, z_n$.

Theorem 17. A Gaussian process X is a Markov process if and only if

$$\mathsf{E}(X_{t_n} \mid X_{t_{n-1}} \text{ near } u_{n-1}, \ldots, X_{t_1} \text{ near } u_1) = \mathsf{E}(X_{t_n} \mid X_{t_{n-1}} \text{ near } u_{n-1}), \tag{16}$$

for all $u_1, \ldots, u_{n-1}$ and all times $t_1 < \ldots < t_n$.

Question 18. What might the Markov property look like for a continuous-time process?

Process 19. Brownian Motion, aka The Weiner Process

A *Weiner process* (or *Brownian motion*) $W = (W_t)_{t \geq 0}$ is a real-valued Gaussian process satisfying:

1. W has independent and stationary increments
2. $W_{t+h} - W_t$ has a $\text{Normal}\langle 0, \sigma^2 h \rangle$ distribution, for all $t, h \geq 0$ and a constant $\sigma^2 > 0$.
3. The sample paths of W are almost surely continuous.

A *standard Weiner process* (or *standard Brownian motion*) has $W_0 = 0$ and $\sigma^2 = 1$.

Among the interesting properties of W:

1. W is a Gaussian process.
2. W is a Markov process.
3. W is a Martingale.
4. The sample paths of W are nowhere differentiable.
5. In some sense to be discussed, W is integrated white noise.
6. Extends easily to multiple dimensions.

Exercise 20. Derive the autocovariance function of W.

Detail 21. Does the Weiner process exist?

Yes, though the proof is nontrivial. It is "easy" to show that a process with the distributional properties exists; the hard part is showing that such a process has continuous sample paths.

Two interesting constructions follow.

Construction 22. The Random Walk Construction

Let S_n denote a simple symmetric random walk, where the particle moves left or right a distance $\Delta x > 0$ at each step. Moreover, assume that each time point takes time $\Delta t > 0$.

Let U_t be the position of the particle at time $t = n\Delta t$ for $n \geq 0$. We have that

$$\begin{aligned} U_t &= S_n \Delta x + (n - S_n)(-\Delta x) && (17) \\ &= (2S_n - n)\Delta x. && (18) \end{aligned}$$

Moreover,

$$\begin{aligned} \mathsf{Var}(U_t) &= (\Delta x)^2 \, \mathsf{Var}(2S_n - n) && (19) \\ &= 4(\Delta x)^2 \frac{n}{4} && (20) \\ &= n(\Delta x)^2 && (21) \\ &= t\frac{(\Delta x)^2}{\Delta t}. && (22) \end{aligned}$$

Now, we are going to let $\Delta x \to 0$ and $\Delta t \to 0$ while keeping

$$\frac{(\Delta x)^2}{\Delta t} = D, , \tag{23}$$

for a constant $D > 0$. The number of steps of the random walk within any fixed time will increase to infinity.

Then, we have that

$$\begin{aligned} U_t &= (2S_n - n)\Delta x && (24) \\ &= \left(\frac{S_n - n/2}{\sqrt{n/4}}\right) \sqrt{n}\Delta x && (25) \\ &= \left(\frac{S_n - n/2}{\sqrt{n/4}}\right) \sqrt{t}\frac{\Delta x}{\sqrt{\Delta t}} && (26) \\ &= \left(\frac{S_n - n/2}{\sqrt{n/4}}\right) \sqrt{Dt} && (27) \\ &\xrightarrow{d} N(0, Dt), && (28) \end{aligned}$$

as $n \to \infty$.

Construction 23. The Haar Construction

Define a process $(\xi_t)_{t \geq 0}$ as follows. Let $(A_n)_{n \geq 0}$ be a standard normal white noise Process, i.e., A_n are IID Normal$\langle 0, 1 \rangle$. Define

$$\xi_t(\omega) = \sum_{n=0}^{\infty} A_n(\omega)\psi_n(t), \tag{29}$$

for a particular complete orthonormal basis $(\psi_n)_{n \geq 0}$.

Now, just as we got a random walk process by taking cumulative sums of a discrete white noise process, we can see what we get when we take cumulative *integrals* of a *continuous* white noise process.

Define

$$W_t = \int_0^t \xi_s \, ds = \sum_{n=0}^{\infty} A_n \int_0^t \psi_n(s) \, ds, \tag{30}$$

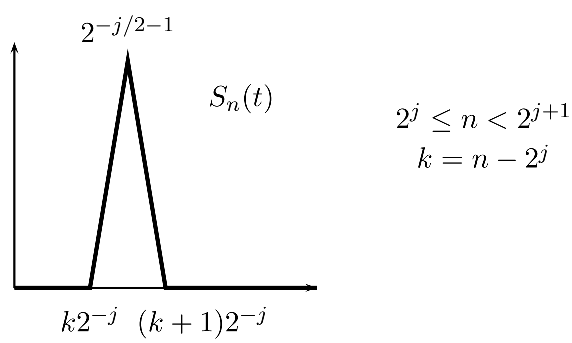
where we choose a specific basis ψ_n.

Now order the Haar functions $(H_0, H_{00}, H_{10}, H_{11}, H_{20}, H_{21}, H_{22}, H_{23}, \ldots)$ and label these as ψ_n for $n \geq 0$. (For $2^j \leq n < 2^{j+1}$ and $j \in \mathbb{Z}_+$, take $k = n - 2^j$ and let $H_n \equiv H_{jk}$.)

For $n \geq 1$,

$$\int_0^t H_n(s) \, ds \equiv S_n(t), \tag{31}$$

called the Schauder function.



It follows that

$$W_t = \sum_{n \geq 0} A_n S_n(t). \tag{32}$$

By the properties of the Schauder functions, we get the following:

$$\mathsf{E}W_t = \sum_{n\geq 0} \mathsf{E}A_n S_n(t) = 0 \tag{33}$$

$$\mathsf{E}W_t^2 = \sum_{n\geq 0} \mathsf{E}A_n^2 S_n(t) = t \tag{34}$$

$$\mathsf{E}W_s W_t = \sum_{n,m\geq 0} \mathsf{E}A_n A_m S_n(t) S_m(s) = \min(s,t) \tag{35}$$

and for $u < s < t$,

$$\mathsf{E}(W_t - W_s)W_u = \sum_{n,m\geq 0} \mathsf{E}A_n A_m (S_n(t) - S_n(s)) S_m(u) = \min(t,u) - \min(s,u) = 0. \tag{36}$$

Moreover, using the characteristic generating functions with $s \leq t$,

$$\mathsf{E}e^{i\lambda(W_t - W_s)} = \mathsf{E}e^{i\lambda \sum_n A_n(S_n(t)-S_n(s))} \tag{37}$$

$$= \prod_{n=0}^{\infty} \mathsf{E}e^{i\lambda A_n(S_n(t)-S_n(s))} \tag{38}$$

$$= \prod_{n=0}^{\infty} e^{-\frac{\lambda^2}{2}(S_n(t)-S_n(s))^2} \tag{39}$$

$$= e^{-\frac{\lambda^2}{2}\sum_n (S_n(t)-S_n(s))^2} \tag{40}$$

$$= e^{-\frac{\lambda^2}{2}(t-2s+s)} \tag{41}$$

$$= e^{-\frac{\lambda^2}{2}(t-s)}, \tag{42}$$

using normality of the A_ns. Hence, $W_t - W_s$ has a Normal$\langle 0, t-s\rangle$ distribution.

This is the Weiner process. But have we shown continuity of the sample paths?

Appendix 24. Properties of Schauder functions

1. Let (a_k) be a real sequence that satisfies $|a_k| = O(k^\gamma)$ for some $0 \leq \gamma < 1/2$. Define $f(t) = \sum_{k\geq 0} a_k S_k(t)$ and $f_n(t)$ be the corresponding partial sum. Then $f_n \to f$ uniformly on $(0,1)$, meaning that $\sup_{0<t<1} |f_n(t) - f(t)| \to 0$.
   (Note: A standard normal white noise sequence A_n satisfies $|A_n| = O(\sqrt{\log n})$ with prob. 1.)
2. If $0 \leq s,t \leq 1$,
$$\sum_{n\geq 0} S_n(s) S_n(t) = \min(s,t). \tag{43}$$