Homework

I still haven't finished grading the fourth homework assignment, and I haven't received the final versions of your projects yet. (You should get this to me, by hardcopy or email, no later than this Wednesday.)

Please fill out an address on one of the envelopes so I can return your final project and your homeworks to you.

Poisson processes

(the first six subsections of the Poisson processes section are taken from the previous lecture's notes, and are re-included here for your convenience; only the last subsection, on Non-homogeneous Poisson processes, is new)

Bernoulli trials with p small and n large

Taking the limit as $p \to 0$, $n \to \infty$: Poisson processes

CADLAG functions

Memorylessness

The M/M/1 queueing model

Splitting and thinning Poisson processes

Non-homogeneous Poisson processes

Recall that a Poisson process is a counting process that associates some (random) finite non-negative integer with every interval I; the Poisson process of intensity λ is characterized by two properties (plus a few technical hypotheses I won't worry about here):
(1) if $I = [t_1, t_2]$, the expected number of events in the time-interval I is $\lambda(t_2 - t_1)$; and
(2) for any two disjoint time-intervals I_1 and I_2, the number of events occurring in I_1 is independent of the number of events occurring in I_2 (we saw this in the special case of the intervals from 0 to s and from s to t).
This is called the independent increments property.

(If this looks unfamiliar, recall that last time we represented a Poisson process as a function $N(t)$ that signifies, for each $t \geq 0$, the number of events that occurred in $[0, t]$. So the number of events that occurred in $[t_1, t_2]$ is just $N(t_2) - N(t_1)$.)

So for a Poisson process, the expected number of events from time t to time $t+\Delta t$ is exactly $\lambda \Delta t$. That is, the expected number of events from time t to time $t+\Delta t$, divided by Δt, equals λ.

More generally, we can have a counting process with independent increments such that the expected number of events from time t to time $t+\Delta t$, divided by Δt, converges to some function $\lambda(t)$, rather than some constant λ, as Δt goes to 0.

This is called a nonhomogeneousPoissonprocess.

In the case where we have an upper bound Λ on $\lambda(t)$ valid for all t, we can use the thinning trick (also called the sampling trick) discussed earlier: generate an ordinary Poisson process of rate Λ, and if the timer goes off at time t, accept the event with probability $\lambda(t)/\Lambda$ and reject it otherwise. (If $\lambda(t)$ is some constant less than Λ, this is ordinary Poisson thinning.)

Application: Recall that the expected number of Poisson events in a time-interval I is proportional to the length of I. But don't think of I as a time-interval anymore; think of it as an interval in a 1-dimensional space, and think of the Poisson process as defining a way of throwing "darts" at the line and seeing how many of them land in I. Analogously, define a <u>2-dimensional Poisson process</u> with intensity λ as a random variable whose "values" are sets of points in the plane, such that:

(1) for any subset S of the plane with area A, the expected number of darts landing in S is λA; and

(2) for any two disjoint subsets S_1 and S_2 of the plane, the number of darts landing in S_1 is independent of the number of darts landing in S_2 (compare this with the comparable statement in 1 dimension about the intervals I_1 and I_2); and

... (some technical hypotheses I won't include here).

How do we simulate a 2-dimensional Poisson process on the disk of radius R?

Answer: Construct the points radially from the center. Note that the annulus from radius r to radius $r+\Delta r$ has area $\approx 2\pi r\ \Delta r$, so the annuli further out are more likely to contain points than the ones further in, even if they have the same thickness Δr. Sending $\Delta r \to 0$, we find that if we replace the random points in the disk by their distances r from the center, and order these distances by size, the result is a nonhomogeneous Poisson process with rate $\lambda(r) = Cr$ for some suitable constant C. (Note that distance r plays the role of time t here.) To simulate this sequence of distances, simulate a Poisson process of rate CR and apply non-homogeneous thinning, accepting a proposed r with probability r/R. (This ceases to be feasible when $r > R$, since then r/R is not a probability, but this is okay, since we're only interested in points inside the disk of radius R; once our proposed distances from the center exceed R, we can stop generating proposed distances.) Then take the resulting sequence of random distances r_1, r_2, ... and choose a random point uniformly on the circle of radius r_1, a random point uniformly on the circle of radius r_2, etc.; the result will be a finite set of points in the disk governed by the 2-dimensional Poisson distribution of rate λ (so that, in particular, for any subset S of the disk of area A, the expected number of points in the randomly-chosen subset that will lie in S equals λA).

Brownian motion

(very loosely adapted from *Introduction to Probability Models* by Sheldon Ross, section 10.1)

Definition of Wiener process

One-dimensional Brownian motion is like one-dimensional random walk, except that the step-sizes and the time-scale on which the steps occur both go to 0 (although, as we'll see, it's important that they go to 0 at different rates).

Suppose at each Δt time-step we go either Δx to the left or Δx to the right, each with probability $\frac{1}{2}$, with successive steps being independent.

Let $X(t)$ be the position of the walker at time t.
The random variable $X(t)$ is a sum of $t/\Delta t$ steps, each with mean 0 and variance $(\Delta x)^2$, and so has mean 0 and variance $(t/\Delta t)(\Delta x)^2$.

If we let $\Delta x = \sqrt{\Delta t}$, then $(t/\Delta t)(\Delta x)^2 = (t/\Delta t)\,\Delta t = t$.

If we send Δt to 0 and apply the Central Limit Theorem, the following properties of the limiting behavior of $X(t)$ seem reasonable:

(1) For all $t \geq 0$, $X(t)$ is normal (aka Gaussian) with mean 0 and variance t.
(2) The process $\{X(t),\ t \geq 0\}$ has independent increments, in the sense that for all $t_1 < t_2 < \ldots < t_n$, the increments $X(t_n) - X(t_{n-1})$, $X(t_{n-1}) - X(t_{n-2})$, ..., $X(t_2) - X(t_1)$, $X(t_1)$ are independent. In fact, each increment $X(t) - X(s)$ (with $s < t$) is normal with mean 0 and variance $t - s$.

It turns out that there is exactly one continuous-time stochastic process satisfying (1) and (2); it is called the (unit-variance) Wiener process, aka Brownian motion.

The technical construction involves a (big!) probability space Ω whose elements are function f from $[0, \infty)$ to $\mathbb{R}$ and a probability measure ("Wiener measure") on Ω such that, for each fixed $0 \leq s < t$, if we pick f from Ω in accordance with Wiener measure, the derived random variable $f(t)-f(s)$ is normal with mean 0 and variance $t-s$.

It can be shown that, with probability 1, such a random f is continuous EVERYWHERE and differentiable NOWHERE.

Gaussians are "universal", in a sense made precise by the Central Limit Theorem: if you add lots of i.i.d. random variables with finite mean and variance, the distribution of the sum looks more and more Gaussian, even if the individual summands didn't have this property. Likewise, Brownian motion is universal: if you look at all the partial sums of an infinite sequence of i.i.d. random variables with finite mean and variance and rescale time and space appropriately, you get Brownian motion in the limit.

Constructing Brownian paths

If we're only interested in graphing $f(\mathtt{t})$ for $\mathtt{t}$ in $\mathbb{Z}$, we can let $f(1) = X_1$, $f(2) = X_1+X_2$, $f(3) = X_1+X_2+X_3$, etc., where $X_1, X_2, X_3, \ldots$ are normal with mean 0 and variance 1 (recall that a Brownian process has independent, normal increments). But what if we want to know $f(\mathtt{t})$ for values of $\mathtt{t}$ not in $\mathbb{Z}$?

E.g., if we have taken $f(1) = B$, how should we pick $f(\frac{1}{2})$?

We can answer this with facts about conditional expectation of Gaussians.

Recall that, before we conditioned on the value of $f(1)$, the definition of Brownian motion told us that $f(\frac{1}{2})-f(0)$ and $f(1)-f(\frac{1}{2})$ are independent Gaussians of mean 0 and variance $\frac{1}{2}$. Let Y_1 and Y_2 denote these independent Gaussians, so that $f(0) = 0$, $f(\frac{1}{2}) = Y_1$, and $f(1) = Y_1+Y_2$. So in conditioning on $f(1)=B$ we are conditioning on $Y_1+Y_2=B$, and in trying to simulate $f(\frac{1}{2})$ we need to know "If Y_1 and Y_2 are independent Gaussians of mean 0 and variance $\frac{1}{2}$, what is the conditional distribution of Y_1, given that $Y_1+Y_2=B$?

It can be shown that the conditional distribution of Y_1 is Gaussian with mean $B/2$ (that makes intuitive sense, by symmetry) and with variance $\frac{1}{2}$. So we can pick $f(\frac{1}{2})$ to be $f(0)=0$ plus a random increment governed by a Gaussian distribution with mean $f(1)/2$ and variance $\frac{1}{2}$.

More generally, suppose we have already specified the values of $f(.)$ at points $\ldots < s < u < \ldots$, and we want to specify the value of $f(.)$ at some point t in (s,u) (it could be the midpoint $(s+u)/2$ or it could be something else). Say we have chosen $f(s)=A$ and $f(u)=B$. Then we should pick $f(t)$ to be a Gaussian with mean

$$\frac{u-t}{u-s}A + \frac{t-s}{u-s}B$$

and variance

$$\frac{(u-t)(t-s)}{u-s}.$$

Brownian motion on an interval (s,u), conditioned upon the endpoint values $f(s)=A$ and $f(u)=B$, is called Brownian bridge. Constructing a Brownian motion by repeatedly subdividing intervals via Brownian bridges is called bridge sampling of Brownian motion.

Try it:

```
Write simulation
```

For a picture of Brownian motion, see e.g. http://www.nbi.dk/~tweezer/pics/brownian-motion.jpg .

Something I'm trying to do in my own research is to find a way to construct Brownian motion via birth-death processes (aka random walk on $\mathbb{Z}$) by successively approximating Brownian motion on $\mathbb{R}$ by continuous-time random walks on $\mathbb{Z}$, $\mathbb{Z}/2$, $\mathbb{Z}/4$, ...

Two-dimensional Brownian motion

So far, the Brownian motion we've discussed is taking place in 1 dimension.

If $x(\mathtt{t})$ and $y(\mathtt{t})$ are independently doing 1-dimensional Brownian motion, then $(x(\mathtt{t}), y(\mathtt{t}))$ is said to be doing 2-dimensional Brownian motion.

For a picture, see e.g. http://www.daviddarling.info/images/Brownian_motion.jpg .

Write simulation

2-dimensional Brownian motion can also be derived as a continuum limit of 2-dimensional random walk on a grid, as the grid-spacing gets finer and finer.

It can be shown that 2-dimensional Brownian motion is rotationally invariant. Indeed, each time-increment $(x(\mathtt{t}+\Delta\mathtt{t})-x(\mathtt{t}),\ y(\mathtt{t}+\Delta\mathtt{t})-y(\mathtt{t}))$ is just a 2-dimensional Gaussian, and one of the basic properties of a Gaussian is that a 2-dimensional Gaussian (X,Y) consisting of a pair of independent mean 0, variance 1 Gaussian random variables has rotational symmetry in $\mathbb{R}^2$.

Brownian motion can also be defined in higher dimensions, but one thing that's special about 2-dimensional Brownian motion is that it exhibits conformal invariance. That is, if D and D' are simply-connected subsets of the plane and $\varphi : D \to D'$ is a conformal map between those domains (i.e., an orientation-preserving map that preserves angles), then φ carries Brownian motion on D to Brownian motion on D', modulo time-parametrization. That is, the behavior of Brownian motion under a conformal map is a time-changed Brownian motion that sometimes goes "faster" and sometimes goes "slower" according to where it is, but whose itinerary, viewed as a time-ordered set of points, is statistically indistinguishable from the itinerary of Brownian motion. (I put "faster" and "slower" in quotes because Brownian motion doesn't have speed of the ordinary kind, because with probability 1 it's not differentiable anywhere.)

In two dimensions, ANY two simply-connected compact subsets of the plane are related by a conformal map, so this gives a very large set of symmetries for Brownian motion -- much bigger than the symmetry group of any deterministic path (or any other sort of geometric object) could be! In higher dimensions, there are far fewer conformal maps. So there's something rather special about two-dimensional Brownian motion.

Course feedback

Should the final project count as less than 50% of the final grade?

Should I axe the final project, and just lecture more?