Lecture 9: Discrete versus Continuous Stochastic Processes

Scribe: Kwai Hung Henry Lam

Department of Statistics, Harvard University

October 14, 2006

In all the previous lectures we focused on discrete random walks and their central limit theorems and corrections. In this lecture we would start discussing the connection between discrete and continuous random walks. More specifically, we want to find results as the time spent in each step goes infinitesimally small. The approach involves PDE (namely diffusion equation). Its main advantages are:

1. In higher dimensions the analysis of discrete random walks becomes overwhelmingly difficult. In contrast its continuous counterpart through PDE remains tractable.

2. Continuous version is easier to solve for boundary value problems.

1 Continuum Approximation to Discrete Random Walks

Here we discuss how to approximate the PDF of discrete random walks by PDE. The key here is to define a function of distance x and time t that is the continuous analog of discrete case where the PDF depends on x and step N. We achieve this by letting the time step go small, forming a diffusion equation.

First of all let us consider random walks with IID steps with all moments of the step distribution being finite. We have

$$X_N = \sum_{n=1}^{N} \triangle x_n$$

and recall from past lectures that

$$\hat{P}_N(k) \sim e^{N\psi(k)} \quad \text{as} \quad N \to \infty$$
$$\sim e^{N \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} c_n} \quad \text{as} \quad k \to 0 \tag{1}$$

Note that if not all moments of p_n exist, then we can only expand the cumulant generating function up to a finite number of terms (and with the possibility of fractional power). Another important point to note is that the expansion is useful only for the "central region", and outside the region more corrections terms in the expansion may even worsen our approximation, as was discussed previously.

1.1 The Fokker-Planck Equation

We introduce the continuous function $\rho(x, t)$, where $t = N\tau$ and τ is the time step, as

$$\rho(x, N\tau) = P_N(x)$$

We also define

$$D_n = \frac{c_n}{n!\tau}$$

for $n \geq 1$. Then from (1) we have

$$\hat{\rho}(k, t) \sim e^{[\sum_{n=1}^{\infty} (ik)^n D_n]t}$$

So

$$\frac{\partial \hat{\rho}}{\partial t}(k, t) \sim \sum_{n=1}^{\infty} (ik)^n D_n \hat{\rho}(k, t) = \sum_{n=1}^{\infty} (-1)^n D_n \frac{\widehat{\partial^n \rho}}{\partial x^n}(k, t)$$

where the last equality follows from the fact that

$$\frac{\widehat{\partial^n \rho}}{\partial x^n}(k, t) = (-ik)^n \hat{\rho}(k, t)$$

Now clearly this is the Fourier transform of the PDE

$$\frac{\partial \rho}{\partial t} = \sum_{n=1}^{\infty} (-1)^n D_n \frac{\partial^n \rho}{\partial x^n}$$

or

$$\frac{\partial \rho}{\partial t} + D_1 \frac{\partial \rho}{\partial x} = D_2 \frac{\partial^2 \rho}{\partial x^2} - D_3 \frac{\partial^3 \rho}{\partial x^3} + D_4 \frac{\partial^4 \rho}{\partial x^4} - \cdots$$

It turns out that only the terms up to second order dominate as $t \to \infty$. So we have the PDE

$$\frac{\partial \rho}{\partial t} + D_1 \frac{\partial \rho}{\partial x} = D_2 \frac{\partial^2 \rho}{\partial x^2}$$

One has to be careful about letting t tend to ∞. The existence of a valid PDE holds usually but not always.

   Two points to note:

   1. If we put $\tilde{x} = x - D_1 t$, and let $\tilde{\rho}(\tilde{x}, t) = \rho(\tilde{x} + D_1 t, t) = \rho(x, t)$, then

   $$\frac{\partial \tilde{\rho}}{\partial t}(\tilde{x}, t) = \frac{\partial \rho}{\partial t}(\tilde{x} + D_1 t, t) = \frac{\partial \rho}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial \rho}{\partial t} = D_1 \frac{\partial \rho}{\partial x} + \frac{\partial \rho}{\partial t}$$

   and

   $$\frac{\partial \rho}{\partial x}(x, t) = \frac{\partial \rho}{\partial x}(\tilde{x} - D_1 t, t) = \frac{\partial \rho}{\partial \tilde{x}}(\tilde{x} - D_1 t, t) = \frac{\partial \tilde{\rho}}{\partial \tilde{x}}(\tilde{x}, t)$$

   So we also have

   $$\frac{\partial^2 \rho}{\partial x^2}(x, t) = \frac{\partial^2 \tilde{\rho}}{\partial \tilde{x}^2}(\tilde{x}, t)$$

   Hence we get

   $$\frac{\partial \tilde{\rho}}{\partial t} = D_2 \frac{\partial^2 \tilde{\rho}}{\partial \tilde{x}^2}$$

   which is the undrifted diffusion equation.

2. Our $P_N(x) = \rho(x, N\tau)$ solves the PDE with the initial value constraint $\rho(x, 0) = \delta(x)$. Hence the "Central Limit Theorem" gives the Green function $G(x, t)$ for this linear PDE. For the general solution where initial condition is $\rho(x, 0) = f(x)$, we have $\rho(x, t) = G * f = \int_{-\infty}^{\infty} G(x - x', t) f(x') dx'$.

Note that, as in the discrete case, $G(x, t)$ accurately describes $P_N(x)$ only in the "central region". i.e. for x where $|x - D_1 t| = O(\sigma\sqrt{t})$.

An advantage of the PDE approach over discrete analysis is that even if there is boundary condition, since the distance from any point to the boundary becomes infinitely large (relative to scaling) as τ goes to 0, whatever boundary condition will not affect our result too much. In other words, the operators of PDE are local. However, the real issue here is whether this PDE approach gives us accurate approximation to the discrete problem that we are ultimately facing.
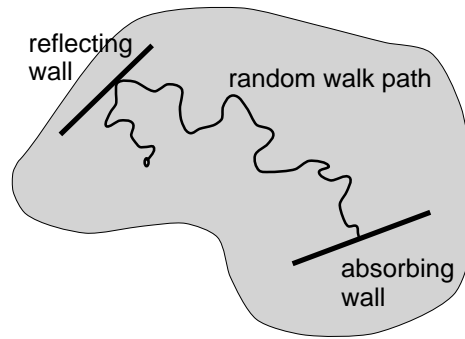
1.2 Occurrences of Diffusion Equation

Diffusion equation arises in many areas of physics. One example is Fick's first and second laws that describe the diffusion of particles. The first law states that
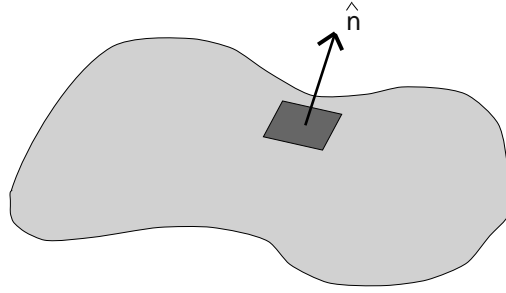
$$\vec{F} = -D_2 \nabla \rho$$

where $\vec{F}$ is the diffusion flux, ρ is the concentration of particles and D_2 is the diffusion coefficient (which turns out to be the same D_2 as we have defined in the previous context). Then, by conservation law, we have

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \vec{F} = D_2 \nabla^2 \rho \tag{2}$$

i.e.

$$\frac{\partial \rho}{\partial t} = D_2 \nabla^2 \rho$$

which is Fick's second law. This equation also arises as Fourier's law, where F becomes heat flux and ρ is temperature.

The first equality in (2) comes from the following. Let

$$Q = \int_V \rho \, dv$$

be the total amount of particles (or heat flow). Then

$$\int_V \frac{\partial \rho}{\partial t} dv = \frac{dQ}{dt} = -\int_S \hat{n} \cdot F \, da = -\int_V \nabla \cdot \vec{F} \, dv$$

Letting $V \to 0$, we get the stated diffusion equation.

More generally, we have

$$\vec{F} = \vec{D}_1 \rho - D_2 \nabla \rho + D_3 \nabla^2 \rho$$

and we get

$$\frac{\partial \rho}{\partial t} + \vec{D}_1 \cdot \nabla \rho = D_2 \nabla^2 \rho$$

Higher corrections to PDE come from higher derivatives in the corresponding flux law.

2 Fat Tail Case

So far we have focused only on light tail distribution. The case for fat tail distribution is, in fact, qualitatively different. Recall in the past lecture that in the fat tail case, i.e. random walk in which higher moments of $p_n(x)$ are infinite,

$$\hat{P}_N(k) \sim e^{N[\sum_{n=1}^\infty \frac{(ik)^n c_n}{n!} + c_\alpha |k|^\alpha + \cdots]} \quad \text{as} \quad k \to 0 \quad \text{and} \quad N \to \infty$$

where α is not an even integer.

Now again we let $\rho(x, t) = \rho(x, N\tau) = P_N(x)$, and also

$$D_n = \frac{c_n}{n! \tau}$$

and

$$D_\alpha = \frac{c_\alpha}{\tau}$$

Then we get

$$\hat{P}_N(k) \sim e^{[\sum_{n=1}^\infty (ik)^n D_n + D_\alpha |k|^\alpha + \cdots]t} \quad \text{as} \quad k \to 0 \quad \text{and} \quad N \to \infty$$

Hence

$$\frac{\partial \hat{\rho}}{\partial t} \sim \left[\sum_{n=1}^\infty D_n (ik)^n + D_\alpha |k|^\alpha + \cdots \right] \hat{\rho}$$

From this we get

$$\frac{\partial \rho}{\partial t} + D_1 \frac{\partial \rho}{\partial x} = D_2 \frac{\partial^2 \rho}{\partial x^2} - D_3 \frac{\partial^3 \rho}{\partial x^3} + D_4 \frac{\partial^4 \rho}{\partial x^4} + \cdots - D_\alpha \frac{\partial^\alpha \rho}{\partial |x|^\alpha} + \cdots$$

Here $\frac{\partial^\alpha \rho}{\partial |x|^\alpha}$ is the Riesz fractional derivative defined by

$$\widehat{\frac{\partial^\alpha \rho}{\partial |x|^\alpha}}(k) = -|k|^\alpha \hat{\rho}$$

Explicitly we can write

$$\frac{\partial^\alpha \rho}{\partial |x|^\alpha} = \int_{-\infty}^\infty e^{ikx}(-|k|^\alpha) \left[\int_{-\infty}^\infty e^{-ikx'} \rho(x') dx' \right] \frac{dk}{2\pi}$$

Note that this is a nonlocal operator. Therefore, in contrast to the case where every moments are finite, we have a nonlocal integral PDE here. The local properties of PDE breaks down, and now the approximation to the discrete case is valid only if we have no boundary.

3 A Note on Scaling

Take a look again at the D_n we have described before. It is equal to $\frac{c_n}{n!\tau}$. Some books, especially those in finance, define D_n in terms of moments instead of cumulants. That is, they define it as $\frac{m_n}{n!\tau}$. These two definitions are in fact the same if we have the right "Central Limit Theorem" scaling. We will show this in the one dimension case. The scaling for this case means $m_1, m_2 = O(\tau)$ and $m_n = O(\tau^{n/2})$ for $n > 2$. In the latter definition we would have

$$D_1 = \lim_{\tau \to 0} \frac{m_1}{\tau}$$

$$D_2 = \lim_{\tau \to 0} \frac{m_2}{2\tau}$$

and

$$D_n = \lim_{\tau \to 0} \frac{m_n}{n!\tau} = 0 \quad \text{for} \quad n > 2$$

Now consider the former definition. Since $m_1 = c_1$, we have

$$D_1 = \lim_{\tau \to 0} \frac{c_1}{\tau} = \lim_{\tau \to 0} \frac{m_1}{\tau}$$

Also, since $m_1 = O(\tau)$,

$$D_2 = \lim_{\tau \to 0} \frac{c_2}{2\tau} = \lim_{\tau \to 0} \frac{m_2 - m_1^2}{2\tau} = \lim_{\tau \to 0} \frac{m_2}{2\tau} - \lim_{\tau \to 0} \frac{m_1^2}{2\tau} = \lim_{\tau \to 0} \frac{m_2}{2\tau}$$

In fact, for all $n \geq 0$ we can write

$$c_n = \sum_k A_k \prod_{i=1}^n m_i^{\alpha_{ki}}$$

where the sum is over all products such that $\sum i\alpha_{ki} = n$, and A_k are constants. It is now clear that

$$\lim_{\tau \to 0} \frac{c_n}{n!\tau} = 0$$

for $n > 2$.

Hence the two definitions are the same for all n.

4 Continuous Stochastic Processes

The difference between continuous stochastic process and continuum approximation to discrete stochastic process must be emphasized. All our previous analysis in this lecture has been on continuum approximation. The solution of PDE obtained is thus an approximation to the true PDF of the discrete random walk. In contrast, some stochastic processes are itself continuous. One example is the Wiener process. One can write it as a stochastic integral

$$Z(t) = \int_0^t dZ(t)$$

where $dZ(t)$ is a stochastic differential with $\langle dZ \rangle = 0$ and $\langle dZ^2 \rangle = dt$. The process so defined is continuous everywhere but nowhere differentiable almost surely. Intuitively,

$$\frac{dZ}{dt} = \Gamma(t) = \text{``white noise''}$$

Z(t) = Wiener process