Lecture 2: Moments, Cumulants, and Scaling

Scribe: Ernst A. van Nierop (and Martin Z. Bazant)

February 4, 2005

Handouts:

- Historical excerpt from Hughes, Chapter 2: Random Walks and Random Flights.
- Problem Set #1

1 The Position of a Random Walk

1.1 General Formulation

Starting at the origin $\vec{X}_0 = 0$, if one takes N steps of size $\Delta\vec{x}_n$, then one ends up at the new position $\vec{X}_N$. This new position is simply the sum of independent random variable steps ($\Delta\vec{x}_n$),

$$\vec{X}_N = \sum_{n=1}^{N} \Delta\vec{x}_n.$$

The set $\{\vec{X}_n\}$ contains all the positions the random walker passes during its walk. If the elements of the set $\{\Delta\vec{x}_n\}$ are indeed independent random variables, then $\{\vec{X}_n\}$ is referred to as a "Markov chain". In a Markov chain, $\vec{X}_{N+1}$ is independent of $\vec{X}_n$ for $n < N$, but does depend on the previous position $\vec{X}_N$.

Let $P_n(\vec{R})$ denote the probability density function (PDF) for the values $\vec{R}$ of the random variable $\vec{X}_n$. Note that this PDF can be calculated for both continuous and discrete systems, where the discrete system is treated as a continuous system with Dirac delta functions at the allowed (discrete) values.

Let $p_n(\vec{r}|\vec{R})$ denote the PDF for $\Delta\vec{x}_n$. Note that in this more general notation we have included the possibility that $\vec{r}$ (which is the value of $\Delta\vec{x}_n$) depends on $\vec{R}$. In Markov-chain jargon, $p_n(\vec{r}|\vec{R})$ is referred to as the "transition probability" from state $\vec{X}_N$ to state $\vec{X}_{N+1}$. Using the PDFs we just defined, let's return to Bachelier's equation from the first lecture.

1.2 Bachelier's Equation

¿From the Markov assumption, the probability of finding a random walker at $\vec{R}$ after N steps depends on the probability of getting to position $\vec{R} - \vec{r}$ after $N-1$ steps, and the probability of stepping $\vec{r}$, so that

$$P_N(\vec{R}) = \int p_N(\vec{r}|\vec{R} - \vec{r}) P_{N-1}(\vec{R} - \vec{r})\, d^d\vec{r}.$$

Now, if we assume that the transition probability is translationally invariant, then $p_n(\vec{r}|\vec{R}) = p_n(\vec{r})$ (i.e. the step does not depend on the current position $\vec{R}$). This reduces the integral to a convolution and gives us Bachelier's equation

$$P_N(\vec{R}) = \int p_N(\vec{r}) P_{N-1}(\vec{R}-\vec{r})\, d^d\vec{r},$$

or

$$P_N = p_N * P_{N-1}$$

where '$*$' denotes the convolution. Note that P_{N-1} is itself the result of a convolution, and we can write P_N in terms of the PDFs of the steps taken $(1 \cdots N)$ and the PDF of the initial position, P_0, so that

$$P_N = p_1 * p_2 * p_3 * \cdots * p_N * P_0.$$

In what follows, we assume that $P_0 = \delta(\vec{x})$, so the random walk always starts at the origin.

1.3 The Convolution Theorem

Convolutions are best treated by Fourier transform, which turns them into simple products. For two functions $f(x)$ and $g(x)$, the Convolution Theorem states that

$$\widehat{f * g}(\vec{k}) = \hat{f}(\vec{k})\hat{g}(\vec{k}).$$

To derive this result, we make use of the definition of the Fourier transform,

$$\begin{aligned}
\widehat{f * g}(\vec{k}) &= \int_{-\infty}^{\infty} e^{-i\vec{k}\cdot\vec{x}} \left[\int_{-\infty}^{\infty} f(\vec{x}')g(\vec{x}-\vec{x}')\, d^d\vec{x}' \right] d^d\vec{x} \\
&= \int_{-\infty}^{\infty} f(\vec{x}') \left[\int_{-\infty}^{\infty} e^{-i\vec{k}\cdot\vec{x}} g(\vec{x}-\vec{x}')\, d^d\vec{x} \right] d^d\vec{x}'
\end{aligned}$$

where we have inverted the order of integration by assuming both integrals are absolutely convergent (which follows from the common assumption, $f, g \in \mathrm{L}_2$). If we now let $\vec{y} = \vec{x} - \vec{x}'$ we obtain

$$\begin{aligned}
\widehat{f * g}(\vec{k}) &= \int_{-\infty}^{\infty} f(\vec{x}') \left[\int_{-\infty}^{\infty} e^{-i\vec{k}\vec{y}} e^{-i\vec{k}\vec{x}'} g(\vec{y})\, d^d\vec{y} \right] d^d\vec{x} \\
&= \hat{f}(\vec{k})\hat{g}(\vec{k}).
\end{aligned}$$

As described in the 2003 notes, the same method can be used to demonstrate the convolution theorem for discrete functions,

$$\begin{aligned}
\hat{f}(k)\hat{g}(k) &= \left(\sum_y e^{-iky} f(y) \right) \left(\sum_z e^{-ikz} g(z) \right) \\
&= \sum_x e^{-ikx} \left(\sum_y g(y) f(x-y) \right) \\
&= \widehat{f * g}(k),
\end{aligned}$$

where all the terms with an exponent equal to z have been collected on the right hand side of the second line. The key property used in both derivations is that $e^{-iky}e^{-ikz} = e^{-ik(y+z)}$.

1.4 Exact Solution for the PDF of the Position

Using the convolution theorem, we now write

$$\begin{aligned}\hat{P}_N(\vec{k}) &= \hat{p}_1(\vec{k})\hat{p}_2(\vec{k})\ldots\hat{p}_N(\vec{k}) \\ &= \prod_{n=1}^{N}\hat{p}_n(\vec{k}).\end{aligned}$$

In the case of identical steps, this simplifies to $\hat{P}_N(\vec{k}) = (\hat{p}_1(\vec{k}))^N$. Using the inverse transform, we find that the PDF of the position of the walker is given by

$$P_N(\vec{x}) = \int_{-\infty}^{\infty} e^{i\vec{k}\cdot\vec{x}} \prod_{n=1}^{N}\hat{p}_n(\vec{k})\frac{d^d\vec{k}}{(2\pi)^d}.$$

This is an exact solution, although the integral cannot be evaluated analytically in terms of elementary functions, except in a few special cases. (See Hughes, Ch. 2.) However, we are generally interested in the "long-time limit" of many steps, $N \to \infty$, where the integral can be approximated by asymptotic analysis. This will be pursued in the next lecture. Here, we focus on scaling properties of the distribution with the number of steps.

2 Moments and Cumulants

2.1 Characteristic Functions

The Fourier transform of a PDF, such as $\hat{P}_N(\vec{k})$ for $\vec{X}_N$, is generally called a "characteristic function" in the probability literature. For random walks, especially on lattices, the characteristic function for the individual steps, $\hat{p}(\vec{k})$, is often referred to as the "structure function". Note that

$$\hat{P}(\vec{0}) = \int_{-\infty}^{\infty} P(\vec{x})\, d^d\vec{x} = 1$$

due to the normalization of $P(\vec{x})$, and that this is a maximum, since

$$|\hat{P}(\vec{k} \neq 0)| \leq \int |e^{-i\vec{k}\cdot\vec{x}}P(\vec{x})|\, d^d\vec{x} = \int P(\vec{x})\, d^d\vec{x} = 1.$$

If $P(\vec{x})$ is nonzero on a set of nonzero measure, then the complex exponential factor causes cancellations in the integral, resulting in $|\hat{P}(\vec{k})| < 1$ for $\vec{k} \neq \vec{0}$.

2.2 Moments

In the 1-dimensional case ($d = 1$), the moments of a random variable, X, are defined as

$$\begin{aligned} m_1 &= \langle x \rangle \\ m_2 &= \langle x^2 \rangle \\ &\vdots \\ m_n &= \langle x^n \rangle. \end{aligned}$$

To get a feeling for the significance of moments, note that m_1 is just the mean or expected value of the random variable. When $m_1 = 0$, the second moment, m_2, allows us to define the "root-mean-square width" of the distribution, $\sqrt{m_2}$. In a multi-dimensional setting ($d > 1$), the moments become tensors,

$$m_n^{(j_1 j_2 \ldots j_n)} = \langle x_{j_1} x_{j_2} \ldots x_{j_n} \rangle,$$

so the n^{th} moment is a tensor of rank n with d^n components. The off-diagonal components measure anisotropy, and generally the moment tensors describe the "shape" of the distribution.

In probability, a characteristic function $\hat{P}(\vec{k})$ is also often referred to as a "moment-generating function", because it conveniently encodes the moments in its Taylor expansion around the origin. For example, for $d = 1$, we have

$$\begin{aligned}\hat{P}(k) &= \sum_{n=0}^{\infty} \frac{(-i)^n k^n m_n}{n!} \\ &= 1 - i m_1 k - \frac{1}{2} m_2 k^2 \ldots\end{aligned}$$

since

$$m_n = i^n \frac{\partial^n \hat{P}}{\partial k^n}(0).$$

For $d > 1$, the moment tensors are generated by

$$m_n^{(j_1 j_2 \ldots j_n)} = i^n \frac{\partial^n \hat{P}}{\partial k_{j_1} \partial k_{j_2} \ldots \partial k_{j_n}}(\vec{0})$$

and from the Taylor expansion,

$$\hat{P}(\vec{k}) = 1 - i \underline{m}_1 \cdot \vec{k} - \frac{1}{2} \vec{k} \cdot \underline{\underline{m}}_2 \cdot \vec{k} + \ldots.$$

Note that $\hat{P}(\vec{k})$ is analytic around the origin if and only if all moments exist and are finite ($m_n < \infty$ for all n). This condition breaks down in cases of distributions with "fat tails", to be discussed in subsequent lectures.

2.3 Cumulants

Certain nonlinear combinations of moments, called "cumulants", arise naturally when analyzing sums of independent random variables. Consider

$$\begin{aligned}P(\vec{x}) &= \int e^{i\vec{k}\cdot\vec{x}} \hat{P}(\vec{k}) \frac{d^d\vec{k}}{(2\pi)^d} \\ &= \int e^{i\vec{k}\cdot\vec{x} + \psi(\vec{k})} \frac{d^d\vec{k}}{(2\pi)^d}\end{aligned}$$

where

$$\psi(\vec{k}) = \log \hat{P}(\vec{k})$$

is called the "cumulant generating function", whose Taylor coefficients at the origin (analogous to the moments above) are the "cumulants", which are scalars for $d = 1$ and, more generally, tensors for $d > 1$

$$\begin{aligned} i^n c_n^{(j_1,j_2,\ldots,j_n)} &= \frac{\partial^n \psi}{\partial k_{j_1} \partial k_{j_2} \ldots \partial k_{j_n}}(\vec{0}) \\ \psi(\vec{k}) &= -i\underline{c}_1 \cdot \vec{k} - \frac{1}{2}\vec{k} \cdot \underline{\underline{c}}_2 \cdot \vec{k} + \ldots. \end{aligned}$$

The cumulants can be expressed in terms of the moments by equating Taylor expansion coefficients in Eq. (2.3). For $d = 1$, we find

$$\begin{aligned} c_1 &= m_1 && \text{or 'mean'} \\ c_2 &= m_2 - m_1^2 = \sigma^2 && \text{or 'variance' } (\sigma = \text{standard deviation}) \\ c_3 &= m_3 - 3m_1m_2 + 2m_1^3 && \text{or 'skewness'} \\ c_4 &= m_4 - 3m_2^2 - 4m_1m_3 + 12m_1^2m_2 - 6m_1^4 && \text{or 'kurtosis'.} \end{aligned}$$

(Note that “skewness” and “kurtosis” are also sometimes used for the normalized cumulants, defined in the next lecture.) Although these expressions increase in complexity rather quickly, the l^{th} cumulant can be expressed in terms of the first l moments in closed form via the determinant of an ‘almost’ lower triangular matrix,

$$c_l = (-1)^{l+1} \begin{vmatrix} m_1 & 1 & 0 & 0 & 0 & 0 & \ldots & 0 \\ m_2 & m_1 & 1 & 0 & 0 & 0 & \ldots & 0 \\ m_3 & m_2 & \binom{2}{1}m_1 & 1 & 0 & 0 & \ldots & 0 \\ m_4 & m_3 & \binom{3}{1}m_2 & \binom{3}{2}m_1 & 1 & 0 & \ldots & 0 \\ m_5 & m_4 & \binom{4}{1}m_3 & \binom{4}{2}m_2 & \binom{4}{3}m_1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ m_{l-1} & m_{l-2} & \ldots & \ldots & \ldots & \ldots & \ddots & 1 \\ m_l & m_{l-1} & \ldots & \ldots & \ldots & \ldots & \ldots & \binom{l-1}{l-2}m_1 \end{vmatrix}.$$

For $d > 1$, the nth cumulant is a tensor of rank n with d^n components, related to the moment tensors, m_l, for $1 \le l \le n$. For example, the second cumulant matrix is given by

$$c_2^{(ij)} = m_2^{(ij)} - m_1^{(i)} m_1^{(j)}.$$

3 Additivity of Cumulants

A crucial feature of random walks with independently identically distributed (IID) steps is that *cumulants are additive*. If we define $\psi(\vec{k})$ and $\psi_N(\vec{k})$ to be the cumulant generating functions of the steps and the final position, respectively, then

$$\begin{aligned} P_N(\vec{x}) &= \int e^{i\vec{k}\cdot\vec{x} + N\hat{\psi}(\vec{k})} \frac{d^d\vec{k}}{(2\pi)^d} \\ &= \int e^{i\vec{k}\cdot\vec{x} + \psi_N(\vec{k})} \frac{d^d\vec{k}}{(2\pi)^d}, \end{aligned}$$

which implies

$$\psi_N(\vec{k}) = N\psi(\vec{k}).$$

Therefore, the n^{th} cumulant tensor $c_{n,N}^{(j_1,j_2,\ldots,j_n)}$ for the position after N steps is related to the n^{th} cumulant tensor for each step by

$$c_{n,N} = N c_n.$$

More generally, if steps are independent but not identical, we still obtain $c_{n,N}$ by simply adding the cumulants of the N steps.

This is a very interesting result in itself, as it brings us back to one of the basic characteristics of normal diffusion. Recall that the standard deviation σ is related to c_2 by $\sigma = \sqrt{c_2}$, and this measures the width of the distribution. After N steps, σ_N will have evolved to $\sqrt{Nc_2} = \sigma\sqrt{N}$. We have recovered the typical "square-root scaling" of normal diffusion (with time or steps taken). As in the first lecture, we see that normal diffusion generally follows from independent (and nearly identical) steps of finite variance (or second moment).

In the next lecture, we will study the long-time asymptotics of $P_N(\vec{x})$ in cases of normal diffusion, with a finite second moment.