Lecture 5: Asymptotics with Fat Tails

Scribed by a student (and Martin Z. Bazant)

Department of Mathematics, MIT

February 15, 2005

In this lecture, we continue our discussion of divergent moments associated with a "fat tail", i.e. a slow power-law decay of the PDF,

$$p(x) \sim \frac{A}{x^{1+\alpha}} \quad \text{as} \quad x \to \infty. \tag{1}$$

We focus on cases where the Central Limit Theorem holds ($\alpha > 2$), and yet most (or all) higher moments than the second diverge. The increased frequency of large rare events in the "tail" of the step distribution, many standard deviations away from the mean, has dramatic consequences for the random walk. In previous lectures, we have seen that higher moments (or more precisely, higher cumulants) when they exist, control corrections to the CLT via Gram-Charlier expansions. Here, we will derive a different form for the asymptotic correction to the CLT for symmetric step distributions with fat tails. We will see that the power-law tail amplitude, A, plays the role of a (divergent) cumulant.

1 Singular Characteristic Functions

The issues which arise with fat tails are well illustrated by the following PDF for IID steps in a random walk (introduced in lecture 4):

$$p(x) = \frac{2}{\sqrt{\pi}(1 + x^4)} \tag{2}$$

which has tail of the form (1) with $\alpha = 3$ and $A = \sqrt{2}/\pi$. This is a borderline case where the CLT applies, but the Berry-Esséen Theorem does not, since all "absolute moments" greater than the third diverge:

$$\bar{m}_\beta = \langle |x|^\beta \rangle = \infty \quad \text{for} \quad \beta \geq 3. \tag{3}$$

All of the correction terms in the Gram-Charlier expansion also diverge and, therefore, must take a different form, which we will calculate below.

The Fourier transform (or characteristic function) of the PDF in (2) can be obtained by standard methods of residue calculus:

$$\hat{p}(k) = e^{-|k|/\sqrt{2}}(\cos(k/\sqrt{2}) + \sin(|k|/\sqrt{2})) \tag{4}$$

which has the following asymptotic expansion near the origin:

$$\hat{p}(k) \sim 1 - \frac{k^2}{2} + \frac{|k|^3}{3\sqrt{2}} - \frac{1}{24}k^4 + O(|k|^5) \quad \text{as} \quad k \to 0 \tag{5}$$

Note that this is not a Taylor expansion, since the function is not analytic at the origin, due to the terms involving odd powers of $|k|$. We have seen that when all moments exist, the characteristic function is analytic at the origin, with Taylor coefficients, $i^n m_n/n!$. Here, the leading singularity at order $\alpha = 3$ in the asymptotic expansion as $k \to 0$ is associated with the smallest diverging moment, $\bar{m}_3 = \infty$.

We have seen this sort of behavior before. Recall that on Problem Set 1, we considered the Cauchy PDF,

$$p(x) = \frac{1}{\pi(1 + x^2)} \tag{6}$$

which has a (very) fat tail of the form (1) with $A = 1/\pi$ and $\alpha = 1$. In this case, the variance is infinite, and so is the "absolute mean", $\bar{m}_1 = \langle |x| \rangle$. These properties of the fat tail are reflected in the singular characteristic function,

$$\hat{p}(k) = e^{-|k|} \sim 1 - |k| + k^2/2 - \dots, \quad \text{as} \ \ k \to 0 \tag{7}$$

which has a leading singularity at order $\alpha = 1$. The Cauchy random walk, with IID steps chosen from this PDF, does not satisfy the CLT due to the infinite variance, but the PDF illustrates the general relation between a power-law tail and a singular characteristic function as the first example. (We will study the Cauchy random walk later, as an example of anomalous diffusion in a "Lévy flight".)

For random walks with IID steps, it is convenient to work with the logarithm of the characteristic function. Returning to our example for today, Eq. (2), we have

$$\psi(k) = \log(\hat{p}(k)) \sim \frac{-k^2}{2} + \frac{|k|^3}{3\sqrt{2}} \quad \text{as} \ \ k \to 0 \tag{8}$$

With all moments finite, we interpreted $\psi(k)$ as the cumulant generating function, with Taylor coefficients $i^n c_n/n!$, but now it generates only the finite second cumulant, $c_2 = 1$. The singular terms, however, also carry useful information, as described by the following result.

Conjecture. $p(x)$ is a symmetric PDF satisfying Eq. (1) for $\alpha > 2$ (not an even integer) if and only if

$$\psi(k) \quad = \quad \log \hat{p}(k) \sim f(k) + g(k), \quad \text{as} \ \ k \to 0, \tag{9}$$
$$\text{where} \ \ f(k) \quad \sim \quad -c_2 k^2/2 \ \ \text{is analytic} \tag{10}$$
$$\text{and} \ \ g(k) \quad \sim \quad c|k|^\alpha \ \ \text{is singular} \tag{11}$$
$$\text{with} \ \ c \quad = \quad \frac{\pi A}{\Gamma(\alpha + 1) \sin(\alpha\pi/2)} \tag{12}$$

If α is an even integer, then logarithmic factors may appear in Eq. (11) at leading order.

This result has been suggested by Prof. Bazant (without proof), although it may appear in the literature on "Tauberian theorems". Later in the class, when we discuss Lévy flights and continuous-time random walks, we will encounter similar Tauberian theorems for Fourier and Laplace transforms, respectively, which relate the *dominant* singular behavior of a transform near the origin ($f(k) = 0$) with a power-law tail of the original function via the same relation. Here, the singularity appears at higher order, $g(k) = o(f(k))$ where $f(k) \sim -c_2 k^2/2$, since some moments exist, at least up to second order. In the next lecture, we will provide a derivation to motivate this result, and here we note that it is satisfied by both of our examples.

2   Generalized Gram-Charlier Expansion for Fat Tails

We now consider the PDF of the position of a random walk with IID steps, whose PDF is symmetric has a fat tail of the form (1) with $\alpha > 2$ (so the CLT holds):

$$P_N(x) = \int_{-\infty}^{\infty} e^{ikx} e^{N\psi(k)} \frac{dk}{2\pi} \tag{13}$$

As $N \to \infty$ in the "central region", we employ Laplace's method of asymptotic expansion (which does not require $\psi(k)$ to be analytic) as before:

$$P_N(x) \sim \int_{-\infty}^{\infty} e^{ikx} e^{N(-k^2/2 + c|k|^3)} \frac{dk}{2\pi} \tag{14}$$

where $c = 1/3\sqrt{2}$ for the example in Eq. (2). As in our derivation of the CLT, we introduce the usual scaling:

$$w^2 = k^2 N, \; z = \frac{x}{\sqrt{N}}$$

and consider the PDF of the rescaled variable, $Z_N = X_N/\sqrt{N}$,

$$\phi_N(z) = \sqrt{N} P_N(z\sqrt{N}) \tag{15}$$

which satisfies

$$\phi_N(z) \sim \int_{-\infty}^{\infty} e^{iwz} e^{\frac{-w^2}{2} + \frac{c|w|^3}{\sqrt{N}}} \frac{dw}{2\pi}. \tag{16}$$

as $N \to \infty$ with z fixed. As before, the "central region" is defined by $z_N = x_N/\sqrt{N} = O(1)$. In this limit, we Taylor expand the exponential,

$$\phi_N(z) \sim \int_{-\infty}^{\infty} e^{iwz} e^{-\frac{w^2}{2}} \left(1 + \frac{c|w|^3}{\sqrt{N}} + ... \right) \frac{dw}{2\pi} = \phi(z) + \sum_{m=1}^{\infty} \frac{g_m(z)}{N^{m/2}} = \phi(z) + \frac{g_1(z)}{\sqrt{N}} + \dots \tag{17}$$

where the leading Gaussian term expresses the CLT,

$$\phi(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \tag{18}$$

while the first correction,

$$g_1(z) = \frac{c f_3(z)}{\pi}, \tag{19}$$

is given by the following integral:

$$f_3(z) = \int_{-\infty}^{\infty} e^{iwz} e^{-\frac{w^2}{2}} |w|^3 \frac{dw}{2}. \tag{20}$$

The expansion in Eq. (17) looks very similar to the one leading to the Gram-Charlier expansion, with the crucial difference that the terms w^m for odd m are replaced by $|w|^m$.

Taking advantage of symmetry, we express $f_3(z)$ as

$$f_3(z) = 2 \int_0^{\infty} e^{-\frac{w^2}{2}} w^3 \cos(wz) \frac{dw}{2} = \left(-\frac{\partial^3}{\partial z^3} \right) I(z) \tag{21}$$

where

$$I(z) = \int_0^\infty \sin(zw) e^{-\frac{w^2}{2}} dw = \text{Im} \left[\int_0^\infty e^{izw - \frac{w^2}{2}} dw \right] \equiv Im[J(z)]. \tag{22}$$

In carrying out the expansion in Eq. (17), the mth correction to the CLT, $g_m(z)/N^{m/2}$, is a sum of terms of the form, $d^n I/dz^n(z)$ for $n \leq m$, so the only integral we have to do is $I(z)$ to derive the complete asymptotic expansion of $\phi_N(z)$.

By completing the square for $J(z)$,

$$J(z) = e^{-\frac{z^2}{2}} \int_0^\infty e^{-(w-iz)^2/2} dw \tag{23}$$

we arrive at an integral, which we can evaluate by contour integration. Consider the integral

$$\oint e^{-q^2/2} dq = 0 \tag{24}$$

around the closed rectangular contour from 0 to $-iz$ to $R - iz$ to R and back to 0, for some real number $R > 0$. Since e^{-q^2} is entire, Cauchy's theorem implies that the integral is zero. The four contributions to the integral, in the limit $R \to \infty$, are as follows:

$$\int_0^{-iz} e^{-q^2/2} dq = -i \int_0^z e^{t^2/2} dt \tag{25}$$

$$\int_{-iz}^{R-iz} e^{-q^2/2} dq \to e^{z^2/2} J(z) \tag{26}$$

$$\int_{R-iz}^R e^{-q^2/2} dq = O(e^{-R^2/2}) \tag{27}$$

$$\int_R^0 e^{-q^2/2} dq \to -\sqrt{\pi/2} \tag{28}$$

Substituting into Eq. (24) and taking the imaginary part yields the desired result,

$$I(z) = \sqrt{2} D(z/\sqrt{2}) \tag{29}$$

where

$$D(x) = e^{-x^2} \int_0^x e^{t^2} dt \tag{30}$$

is "Dawson's integral" (see Figure 1).

Therefore, we have obtained all of the following functions involved in the generalized Gram-Charlier expansion:

$$f_m(z) = \int_0^\infty w^m \cos(wz) e^{-w^2/2} dw = (-1)^m \frac{d^m I}{dz^m}(z) = \frac{(-1)^m}{2^{(m-1)/2}} \frac{d^m D}{dz^m} \left(\frac{z}{\sqrt{2}} \right) \tag{31}$$

The first three derivatives of Dawson's integral are:

$$\begin{aligned}
D'(x) &= 1 - 2xD(x) \\
D''(x) &= -2x + (4x^2 - 2)D(x) \\
D'''(x) &= 4(x^2 - 1) + (12x - 8x^3)D(x)
\end{aligned} \tag{32}$$

so the first two terms in the expansion for our original example, where $\alpha = 3$, are given by

$$\phi_N(z) \sim \phi(z) - \frac{c D'''(z/\sqrt{2})}{2\pi\sqrt{N}} \tag{33}$$

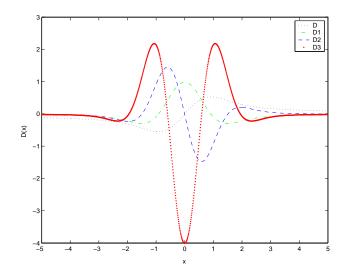which is quite different from the standard Gram-Charlier expansion for finite moments.

Figure 1: Dawson's integral and first three derivatives.

3   Edge of the Central Region

The difference becomes most apparent at the edge of the central region, as $z \to \infty$. Note that Dawson's integral has the following asymptotic expansion as $x \to \infty$:

$$D(x) \sim \frac{1}{2x}\left(1 + \frac{1}{2x^2} + \frac{3}{4x^4} + \sum_{m=3}^{\infty} \frac{(2m-1)!!}{2^m x^{2m}}\right) \tag{34}$$

Therefore, $D'''(x) \sim -3/x^4$, and

$$f_3(z) = -\frac{1}{2}D'''\left(\frac{z}{\sqrt{2}}\right) \sim \frac{6}{z^4} \quad \text{as } z \to \infty \tag{35}$$

The first correction in Eq. (17) thus behaves like

$$\frac{g_1(z)}{\sqrt{N}} = \frac{cf_3(z)}{\pi\sqrt{N}} \sim \frac{6c}{\pi N^{1/2} z^4} = \frac{6cN^{3/2}}{\pi x^4} \tag{36}$$

as one leaves the central region, $z \to \infty$.

The "edge" of the central region is found when this correction term becomes as large as the leading Gaussian term, $\phi(z) = e^{-z^2/2}/\sqrt{2\pi}$. This occurs when

$$z_{max} \approx \sqrt{\log N} \quad \text{or} \quad x_{max} \approx \sqrt{N \log N} \tag{37}$$

which is not much wider than the minimum required by the CLT, $z = O(1)$ and $x = O(\sqrt{N})$. In contrast, we saw in the previous lecture that the central region is much wider when higher cumulants are finite: $x_{max} = O(N^{2/3})$ for $c_3 \neq 0$ or $x_{max} = O(N^{3/4})$ for $c_3 = 0$ and $c_4 > 0$. The collapse of the central region is due to the fat tail.

4 Additivity of Power-Law Tails

Returning to our example, Eq. (2), we have $c = 1/3\sqrt{2}$ and $A = 2/\sqrt{\pi}$, in which case the correction term takes the form

$$\frac{g_1(z)}{\sqrt{N}} \sim \frac{\sqrt{2}}{\pi\sqrt{N}z^4} \tag{38}$$

If this term were to persist *outside* the central region, where it dominates the Gaussian term, then we would have

$$P_N(x) = \frac{1}{\sqrt{N}}\Phi_N\left(\frac{x}{\sqrt{N}}\right) \sim \frac{N\sqrt{2}}{\pi x^4} = \frac{NA}{x^4} \tag{39}$$

which, curiously, is a fat tail of the same power-law exponent ($\alpha = 3$) and N times the tail amplitude (A) as the PDF for the steps, $p(x)$.

Could it be that power-law tail amplitudes are additive, just like cumulants? Our analysis does not apply outside the central region, so this is only a conjecture, but it is consistent with our exact result for the Cauchy random walk from Problem Set 1:

$$p(x) = \frac{1}{\pi(1 + x^2)} \sim \frac{1}{\pi x^2} \quad \text{for } x \to \infty \tag{40}$$

and

$$P_N(x) = \frac{N}{\pi(N^2 + x^2)} \sim \frac{N}{\pi x^2} \quad \text{for } x/N \to \infty. \tag{41}$$

Therefore, it appears that

$$p(x) \sim \frac{A}{|x|^{1+\alpha}} \;\Rightarrow\; P_N(x) \sim \frac{NA}{|x|^{1+\alpha}} \;\text{ as }\; |x| \to \infty. \tag{42}$$

Indeed, the additivity of power-law tails for sums of IID random variables follows from our Conjecture about singular characteristic functions, $\psi(k) = \log \hat{p}(k) \sim -c_2 k^2/2 + \ldots + c|k|^\alpha$, where $c \propto A$:

$$\begin{aligned} P_N(x) &= \int_{-\infty}^{\infty} e^{ikx} e^{N\psi(k)} \frac{dk}{2\pi} \\ &\sim \int_{-\infty}^{\infty} e^{ikx} e^{-c_2 Nk^2/2 + \ldots + cN|k|^\alpha + \ldots} \frac{dk}{2\pi} \end{aligned} \tag{43}$$

Clearly, *all* coefficients in the asymptotic expansion of $\psi(k)$ near the origin are additive, not only the cumulants (coefficients of analytic terms, k^n), but also the power-law tail amplitudes (coefficients of singular terms, $|k|^\beta$).

In the next lecture, we will study the Conjecture in more detail and give intuition regarding the additivity of power-law tails.