Prof. Lorenzo Naranjo

Class 3

Probability Generating Functions

Most random variables in this course are sums: a random walk is a sum of flips, defaults in a loan book are a sum of indicators, and an annual log return is a sum of daily ones. A generating function encodes a whole distribution in one function and turns sums of independent variables into products. For X taking values in $\mathbb{N} = \{0, 1, 2, \ldots\}$, the **probability generating function** (PGF) is

$$G_X(z) = \mathrm{E}(z^X) = \sum_{k=0}^{\infty} p_X(k) z^k, \qquad G_X(1) = 1. \tag{1}$$

The coefficient of z^k is $\mathrm{P}(X = k) = G_X^{(k)}(0)/k!$, so the PGF determines the distribution: if we recognize a PGF, we know the distribution. Setting $z = 1$ in the derivatives gives the moments,

$$\mathrm{E}(X) = G_X'(1), \qquad \mathrm{V}(X) = G_X''(1) + G_X'(1) - G_X'(1)^2. \tag{2}$$

Three PGFs cover most of what we need: Bernoulli $1 - p + pz$, geometric (flips until the first head) $pz/(1 - qz)$ with $q = 1 - p$, and Poisson $e^{\lambda(z-1)}$.

Sums, Random Sums and the Poisson Limit

If X and Y are independent, then $G_{X+Y}(z) = G_X(z)\, G_Y(z)$. Multiplying the power series performs every convolution sum $\sum_j p_X(j) p_Y(k - j)$ at once. A sum of n Bernoulli defaults has PGF $(1 - p + pz)^n$, the binomial distribution, and independent Poisson counts add up to a Poisson count with the summed parameter.

When the number of terms N is itself random and independent of the iid terms $X_1, X_2, \ldots$, the **random sum** $S = X_1 + \cdots + X_N$ has

$$G_S(z) = G_N(G_X(z)), \qquad \mathrm{E}(S) = \mathrm{E}(N)\mathrm{E}(X). \tag{3}$$

Thinning a Poisson(λ) count, keeping each event with probability p, gives Poisson(λp).

With n loans that each default with probability $p = \lambda/n$, the binomial PGF is $(1 + \lambda(z - 1)/n)^n \to e^{\lambda(z-1)}$. By the continuity theorem for PGFs the probabilities converge too: many independent rare events produce an approximately Poisson count with parameter equal to the expected number of events, the **law of rare events**.

Moment Generating Functions

Substituting $z = e^\theta$ handles variables with negative or continuous values. The **moment generating function** (MGF) is $M_X(\theta) = \mathrm{E}(e^{\theta X})$, wherever it is finite. When it is finite near zero:

1. $\mathrm{E}(X^k) = M_X^{(k)}(0)$.
2. $M_{a+bX}(\theta) = e^{a\theta} M_X(b\theta)$.
3. If X and Y are independent, $M_{X+Y}(\theta) = M_X(\theta) M_Y(\theta)$.
4. The MGF determines the distribution.

The key cases are

$$T \sim \text{Exp}(\lambda) : \frac{\lambda}{\lambda - \theta}, \qquad X \sim \text{Gamma}(\alpha, \lambda) : \left(\frac{\lambda}{\lambda - \theta}\right)^\alpha, \qquad X \sim \mathcal{N}(\mu, \sigma^2) : e^{\mu\theta + \frac{1}{2}\sigma^2\theta^2}, \quad (4)$$

the first two for $\theta < \lambda$. Products of these give three facts at once: a sum of n independent Exp(λ) waiting times is Gamma(n, λ), $Z_1^2 + \cdots + Z_k^2$ is chi-squared Gamma$(k/2, 1/2)$, and independent normals sum to a normal, so independent normal daily log returns give a normal annual log return. Lognormal moments are values of the normal MGF, $\mathrm{E}(e^{\alpha X}) = M_X(\alpha)$. The lognormal variable e^X itself has no MGF, and neither does a Student t.

Cumulant Generating Functions

The **cumulant generating function** $K_X(\theta) = \ln M_X(\theta)$ turns independent sums into sums, $K_{X+Y} = K_X + K_Y$, and

$$K_X'(0) = \mathrm{E}(X), \qquad K_X''(0) = \mathrm{V}(X). \quad (5)$$

The normal has $K(\theta) = \mu\theta + \frac{1}{2}\sigma^2\theta^2$, so its cumulants beyond the variance are zero, and the Poisson has $K(\theta) = \lambda(e^\theta - 1)$. For iid steps with cumulant generating function K, the

process $e^{\theta S_n - nK(\theta)}$ is an **exponential martingale**, the tool we will use for stopping times and changes of measure.

Characteristic Functions and the CLT

The **characteristic function** $\phi_X(u) = \mathrm{E}(e^{iuX}) = \mathrm{E}(\cos(uX)) + i\,\mathrm{E}(\sin(uX))$ exists for every random variable, since $|\phi_X(u)| \le 1$. It shares the MGF properties: $\phi_{a+bX}(u) = e^{iau}\phi_X(bu)$, independent sums multiply, it determines the distribution, and $\phi_X^{(k)}(0) = i^k\,\mathrm{E}(X^k)$. For integer-valued X, $\mathrm{P}(X = k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-iuk}\phi_X(u)\,du$. The normal has $\phi(u) = e^{i\mu u - \frac{1}{2}\sigma^2 u^2}$ and a fair ± 1 coin has $\phi(u) = \cos(u)$.

By **Lévy's continuity theorem**, X_n converges in distribution to X exactly when $\phi_{X_n}(u) \to \phi_X(u)$ for every u. For iid X_i with mean zero and variance σ^2, a Taylor expansion gives

$$\phi_{S_n/\sqrt{n}}(u) = \left(1 - \frac{\sigma^2 u^2}{2n} + o\left(\frac{1}{n}\right)\right)^n \to e^{-\frac{1}{2}\sigma^2 u^2}, \tag{6}$$

the **central limit theorem**. Figure 1 shows it for the coin: $\cos(u/\sqrt{n})^n$ approaches $e^{-u^2/2}$. The Poisson limit has the same structure, a transform raised to the n-th power with $(1 + x/n)^n \to e^x$, but rare events give a normal.

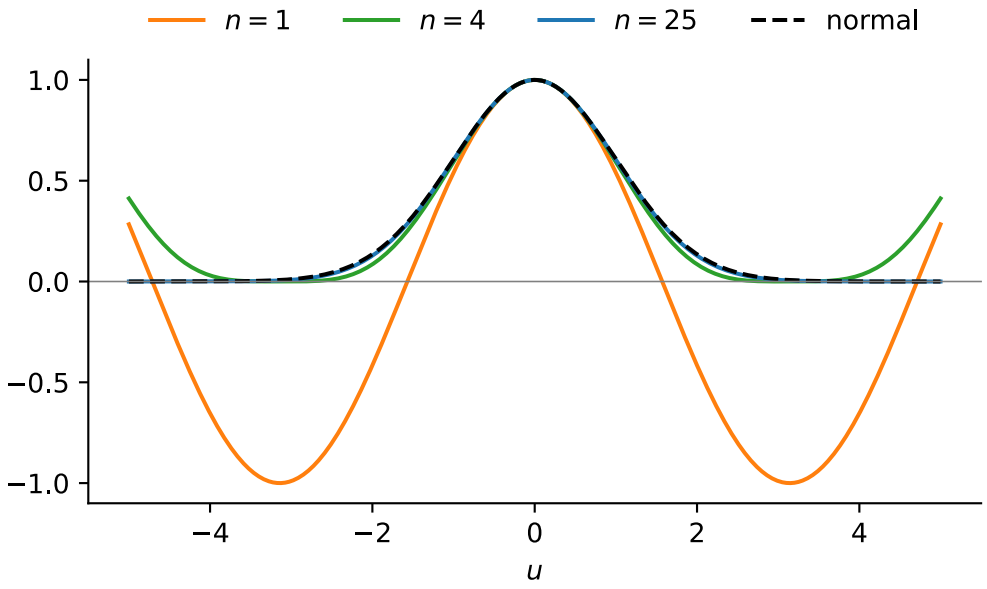


Figure 1: Characteristic function of $S_n/\sqrt{n}$ for a fair ± 1 coin, $\cos(u/\sqrt{n})^n$, against the standard normal characteristic function $e^{-u^2/2}$.

Practice Problems

1. The number of defaults N next year in a loan book has the PGF

$$G_N(z) = e^{3(z-1)}(0.7 + 0.3z)^2.$$

   a. Interpret N as a sum of independent counts: what distributions and parameters are involved?
   b. Compute $\mathrm{E}(N)$ and $\mathrm{V}(N)$.
   c. Compute $\mathrm{P}(N = 0)$ and $\mathrm{P}(N = 1)$.

2. A hedge fund closes at the end of each year with probability 0.2, so the number of years N it operates, including the year it closes, is geometric with $p = 0.2$. In each year of operation the fund beats its benchmark with probability 0.5, independently of other years and of N. Let S be the number of years in which the fund beats its benchmark.

   a. Show that $G_S(z) = \dfrac{1+z}{6-4z}$.
   b. Compute $\mathrm{E}(S)$ and $\mathrm{V}(S)$.
   c. Compute $\mathrm{P}(S = 0)$, and show that $\mathrm{P}(S = k) = \frac{5}{18}\left(\frac{2}{3}\right)^{k-1}$ for $k \geq 1$.

3. A catastrophe bond defaults at the third hurricane landfall in a region. The times between landfalls are independent and exponential with rate $\lambda = 0.5$ per year, and W is the time to the third landfall.

   a. Find the MGF of W and identify its distribution.
   b. Use the cumulant generating function of W to compute $\mathrm{E}(W)$ and $\mathrm{SD}(W)$.
   c. Use the MGF to find the distribution of the same time measured in months, $12W$.

4. A stock's annual log return is $R = 0.08 + 0.20Z - 0.10N$, where $Z \sim \mathcal{N}(0,1)$ and the number of crashes $N \sim \text{Poisson}(0.5)$ are independent. Each crash lowers the log price by 0.10.

   a. Find the cumulant generating function $K_R(\theta)$.
   b. Use it to compute $\mathrm{E}(R)$ and $\mathrm{V}(R)$.
   c. Compute the expected gross return $\mathrm{E}(e^R)$.
   d. Compute the third cumulant $K_R'''(0)$. What does its sign say about the distribution of R, compared with a normal log return?