Remark (Why generating functions?). Probability generating functions are worth the effort: they turn fiddly distribution proofs (sums of Poissons, mean and variance of the geometric, ...) into a few lines of algebra, and they lead naturally into *moment generating functions*, which we will meet later for continuous distributions. Reference: [S3/4] S4 Ch 3.

Packing a Distribution into a Function

A discrete distribution on $\{0, 1, 2, \ldots\}$ is a list of probabilities $p_0, p_1, p_2, \ldots$ The idea of a generating function is to store the whole list as the coefficients of a single power series — a trick you have already seen in combinatorics with ordinary generating functions.

Definition. Let X be a random variable taking values in $\{0, 1, 2, \ldots\}$ with $p_k = \mathbb{P}(X = k)$. The **probability generating function** (PGF) of X is

$$G_X(t) = \mathbb{E}\left[t^X\right] = \sum_{k=0}^{\infty} p_k t^k.$$

The variable t has no probabilistic meaning — it is a formal bookkeeping device. The probability $\mathbb{P}(X = k)$ is simply the coefficient of t^k.

Example

Let X be the score on a fair die. Write down $G_X(t)$.

Fact — For any random variable X,

$$G_X(1) = \sum_k p_k \cdot 1^k = \sum_k p_k = 1.$$

This is a useful sanity check on any PGF you compute — and it guarantees that the series converges at least for $|t| \le 1$.

Two further facts we shall use freely:

- **Uniqueness:** the PGF determines the distribution completely (two variables with the same PGF have the same distribution), because a power series determines its coefficients. This is what makes PGFs so powerful for *identifying* distributions.
- $G_X(0) = p_0 = \mathbb{P}(X = 0)$.

PGFs of the Standard Distributions

Example (Discrete uniform)

Find, in closed form, the PGF of $X \sim U(n)$, uniform on $\{1, 2, \ldots, n\}$.

Example (Bernoulli and binomial)

Find the PGFs of $X \sim B(1, p)$ and of $Y \sim B(n, p)$. Write $q = 1 - p$.

Example (Geometric)

Find the PGF of $X \sim \text{Geo}(p)$, where $\mathbb{P}(X = k) = q^{k-1}p$ for $k = 1, 2, 3, \ldots$

Example (Poisson)

Find the PGF of $X \sim \text{Po}(\lambda)$.

Fact (Standard PGFs) — With $q = 1 - p$ throughout:

Distribution	$\mathbb{P}(X = k)$	$G_X(t)$
Uniform $U(n)$	$\frac{1}{n},\ 1 \le k \le n$	$\frac{t(1-t^n)}{n(1-t)}$
Bernoulli $B(1, p)$	$p^k q^{1-k},\ k \in \{0, 1\}$	$q + pt$
Binomial $B(n, p)$	$\binom{n}{k} p^k q^{n-k}$	$(q + pt)^n$
Geometric Geo(p)	$q^{k-1} p,\ k \ge 1$	$\frac{pt}{1 - qt}$
Poisson Po(λ)	$\frac{e^{-\lambda}\lambda^k}{k!}$	$e^{\lambda(t-1)}$

Extracting the Mean and Variance

Differentiating a PGF term-by-term brings down factors of k — exactly the ingredients of $\mathbb{E}[X]$.

Theorem

If X has PGF G_X, then

$$\mathbb{E}[X] = G_X'(1) \qquad \text{and} \qquad \mathrm{Var}[X] = G_X''(1) + G_X'(1) - [G_X'(1)]^2.$$

Notice that PGFs hand us the *factorial moment* $\mathbb{E}[X(X-1)]$ — the same quantity that made the Poisson variance calculation clean in the last chapter.

Example

Use the PGF to find the mean and variance of $X \sim \mathrm{Po}(\lambda)$.

Example

Use the PGF to prove that $X \sim \text{Geo}(p)$ has $\mathbb{E}[X] = \dfrac{1}{p}$ and $\text{Var}[X] = \dfrac{q}{p^2}$.

Recovering probabilities from a PGF

Since $G_X(t) = p_0 + p_1 t + p_2 t^2 + \cdots$, we can travel in the other direction: given a PGF, expand it as a power series (binomial theorem, known Maclaurin series, or long division) and read off coefficients. Formally, by repeated differentiation at 0,

$$\mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!},$$

which is precisely the Maclaurin coefficient formula.

Example

A random variable X has PGF $G_X(t) = \dfrac{(2+t)^3}{27}$. Find $\mathbb{P}(X = 2)$ and $\mathbb{E}[X]$, and identify the distribution of X.

Example

A random variable X has PGF $G_X(t) = \dfrac{t}{2-t}$. Find $\mathbb{P}(X = 3)$, and identify the distribution of X.

Example (OCR S4, June 2014)

The discrete random variable X has probability generating function $\dfrac{t}{a-bt}$, where a and b are constants.

(i) Find a relationship between a and b.

(ii) Use the probability generating function to find $\mathbb{E}[X]$ in terms of a, giving your answer as simply as possible.

(iii) Expand the probability generating function as a power series, as far as the term in t^3, giving the coefficients in terms of a and b.

(iv) Name the distribution for which $\dfrac{t}{a-bt}$ is the probability generating function, and state its parameter(s) in terms of a.

Example (OCR S4, June 2010 (part))

The probability generating function of the discrete random variable X is $\dfrac{e^{4t^2}}{e^4}$. Find

(i) $\mathbb{E}[X]$,

(ii) $\mathbb{P}(X=2)$.

Sums of Independent Random Variables

Here is the headline theorem — the real reason PGFs earn their keep.

Theorem

If X and Y are *independent* random variables, then

$$G_{X+Y}(t) = G_X(t)\,G_Y(t).$$

Adding independent random variables multiplies their PGFs.

The proof is one line, using the fact that independence gives $\mathbb{E}[UV] = \mathbb{E}[U]\,\mathbb{E}[V]$ for functions of X and of Y.

Theorem (Sum of independent Poissons)

If $X \sim \text{Po}(\lambda)$ and $Y \sim \text{Po}(\mu)$ are independent, then $X + Y \sim \text{Po}(\lambda + \mu)$.

Theorem (Sum of independent Bernoullis)

If $X_1, X_2, \ldots, X_n \sim B(1, p)$ are independent, then $X_1 + X_2 + \cdots + X_n \sim B(n, p)$.

Example (Total score of two dice)

Two fair dice are rolled and T is the total score.

(a) Write down the PGF of T.

(b) Use it to find $\mathbb{P}(T = 4)$.

(c) Hence find $\mathbb{E}[T]$.

Example (OCR S4, June 2016)

Andrew has five coins. Three of them are unbiased. The other two are biased such that the probability of obtaining a head when one of them is tossed is $\frac{3}{5}$. Andrew tosses all five coins. It is given that the probability generating function of X, the number of heads obtained on the unbiased coins, is

$$G_X(t) = \tfrac{1}{8} + \tfrac{3}{8}t + \tfrac{3}{8}t^2 + \tfrac{1}{8}t^3.$$

(i) Find $G_Y(t)$, the probability generating function of Y, the number of heads on the biased coins.

(ii) The random variable Z is the total number of heads obtained when Andrew tosses all five coins. Find the probability generating function of Z, giving your answer as a polynomial.

(iii) Find $\mathbb{E}[Z]$ and $\text{Var}[Z]$.

(iv) Write down the value of $\mathbb{P}(Z = 3)$.

Example (Edexcel FS1, June 2023)

The discrete random variable X has probability generating function

$$G_X(t) = \frac{t^2}{(3-2t)^2}.$$

(a) Specify the distribution of X.

(b) A fair die is rolled repeatedly. Describe an outcome that could be modelled by the random variable X.

(c) Use calculus and $G_X(t)$ to find (i) $\mathbb{E}[X]$, (ii) $\text{Var}[X]$.

(d) The discrete random variable Y has probability generating function

$$G_Y(t) = \frac{t^{10}}{(3-2t^3)^2}.$$

Find the exact value of $\mathbb{P}(Y = 19)$.

Exercise. (A classic.) Can two six-sided dice be weighted (not necessarily fairly, not necessarily identically) so that the total T is uniform on $\{2, 3, \ldots, 12\}$? Think about what the factorisation $G_T = G_X G_Y$ would force, and in particular the roots of these polynomials.

> **Remark** (Looking ahead: moment generating functions)**.** Replacing t by e^s gives $M_X(s) = \mathbb{E}[e^{sX}] = G_X(e^s)$, the moment generating function, which makes sense for continuous random variables too and whose derivatives at $s = 0$ generate the moments $\mathbb{E}[X], \mathbb{E}[X^2], \mathbb{E}[X^3], \ldots$ The multiplicative property for independent sums carries over verbatim, and will be our main tool when we prove results like the Central Limit Theorem later in the course.

Textbook Exercises: [S3/4] S4 Ch 3