4 Generating functions

Even quite straightforward counting problems can lead to laborious and lengthy calculations. These are often greatly simplified by using generating functions. [28]

Definition 4.1. Given a collection of real numbers $(a_k)_{k\ge0}$, the function

$$G(s) = G_a(s) \stackrel{\mathsf{def}}{=} \sum_{k=0}^{\infty} a_k\, s^k \tag{4.1}$$

is called the *generating function* of $(a_k)_{k\ge0}$.

☞ Why do we care? If the generating function $G_a(s)$ of $(a_n)_{n\ge0}$ is analytic near the origin, then there is a one-to-one correspondence between $G_a(s)$ and $(a_n)_{n\ge0}$; namely, a_k can be recovered via [29]

$$a_k = \frac{1}{k!}\,\frac{d^k}{ds^k}\,G_a(s)\,\Big|_{s=0}\ . \tag{4.2}$$

This result is often referred to as the uniqueness property of generating functions.

Definition 4.2. If X is a discrete random variable with values in $\mathbb{Z}^+ \stackrel{\mathsf{def}}{=} \{0,1,\dots\}$, its (*probability*) *generating function*,

$$G(s) \equiv G_X(s) \stackrel{\mathsf{def}}{=} \mathsf{E}\big(s^X\big) = \sum_{k=0}^{\infty} s^k \mathsf{P}(X=k)\,, \tag{4.3}$$

is just the generating function of the pmf $\big\{p_k\big\} \equiv \big\{\mathsf{P}(X=k)\big\}$ of X.

Recall that the moment generating function[30] $M_X(t) \stackrel{\mathsf{def}}{=} \mathsf{E}(e^{tX})$ of a random variable X is just [31] $\sum_{k\ge0} \frac{\mathsf{E}(X^k)}{k!} t^k$. ☞ Why do we introduce both $G_X(s)$ and $M_X(t)$?

The following result illustrates one of the most useful applications of generating functions in probability theory:

✍ **Theorem 4.3.** *If X and Y are independent random variables with values in $\{0,1,2,\dots\}$ and $Z \stackrel{\mathsf{def}}{=} X+Y$, then their generating functions satisfy* [32]

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

Example 4.4. *If X_1, X_2, …, X_n are independent identically distributed random variables* [33] *with values in $\{0,1,2,\dots\}$ and if $S_n = X_1+\dots+X_n$, then*

$$G_{S_n}(s) = G_{X_1}(s)\dots G_{X_n}(s) \equiv \big[G_X(s)\big]^n\,.$$

[28] introduced by de Moivre and Euler in the early eighteenth century.

[29] this and a several other useful properties of power series can be found in Sect. A.4 below.

[30] we might have $M_X(t) = \infty$ for $t \ne 0$!

[31] ie., it is the generating function of the sequence $\mathsf{E}(X^k)/k!$.

☞ [32] recall: if X and Y are discrete random variables, and f, $g : \mathbb{Z}^+ \to \mathbb{R}$ are arbitrary functions, then $f(X)$ and $g(Y)$ are independent random variables and $\mathsf{E}\big[f(X)g(Y)\big] = \mathsf{E}f(X)\cdot\mathsf{E}g(Y)$;

[33] from now on we shall often abbreviate this to just i.i.d.r.v.

☞ **Example 4.5.** *Let X_1, X_2, …, X_n be* i.i.d.r.v. *with values in* $\{0,1,2,\dots\}$ *and let* $N \ge 0$ *be an integer-valued random variable independent of* $\{X_k\}_{k\ge1}$. *Then*[34] $S_N \stackrel{\mathsf{def}}{=} X_1 + \dots + X_N$ *has generating function*

$$G_{S_N}(s) = G_N\big(G_X(s)\big)\,. \tag{4.4}$$

Solution. This is a straightforward application of the partition theorem for expectations. Alternatively, the result follows from the standard properties of conditional expectations: $\mathsf{E}\big(z^{S_N}\big) = \mathsf{E}\big[\mathsf{E}\big(z^{S_N} \mid N\big)\big] = \mathsf{E}\big(\big[G_X(z)\big]^N\big) = G_N\big(G_X(z)\big)$. □

☞ **Example 4.6.** [Renewals] *Imagine a diligent janitor who replaces a light bulb the same day as it burns out. Suppose the first bulb is put in on day 0 and let X_i be the lifetime of the ith light bulb. Let the individual lifetimes X_i be* i.i.d.r.v.*'s with values in* $\{1,2,\dots\}$ *and have a common distribution with generating function $G_f(s)$. Define* $r_n \stackrel{\mathsf{def}}{=} \mathsf{P}\big(\text{a light bulb was replaced on day } n\big)$ *and* $f_k \stackrel{\mathsf{def}}{=} \mathsf{P}\big(\text{the first light bulb was replaced on day } k\big)$. *Then*

$$r_0 = 1\,,\quad f_0 = 0\,,\qquad \text{and}\qquad r_n = \sum_{k=1}^n f_k r_{n-k}\,,\quad n \ge 1\,.$$

A standard computation implies that $G_r(s) = 1 + G_f(s)\,G_r(s)$ for all $|s| < 1$, so that $G_r(s) = 1/(1 - G_f(s))$.

In general, we say a sequence $(c_n)_{n\ge0}$ is the convolution of $(a_k)_{k\ge0}$ and $(b_m)_{m\ge0}$ (write $c = a \star b$), if

$$c_n = \sum_{k=0}^n a_k\, b_{n-k}\,,\qquad n \ge 0\,, \tag{4.5}$$

The key property of convolutions is given by the following result:

✍ **Theorem 4.7.** [Convolution thm] *If $c = a \star b$, then the generating functions $G_c(s)$, $G_a(s)$, and $G_b(s)$ satisfy $G_c(s) = G_a(s)\,G_b(s)$.*

Example 4.8. *Let $X \sim \mathsf{Poi}(\lambda)$ and $Y \sim \mathsf{Poi}(\mu)$ be independent. Then $Z = X + Y$ is $\mathsf{Poi}(\lambda + \mu)$.*

Solution. A straightforward computation gives $G_X(s) = e^{\lambda(s-1)}$; Theorem 4.3 then implies $G_Z(s) = G_X(s)\,G_Y(s) = e^{\lambda(s-1)}\,e^{\mu(s-1)} \equiv e^{(\lambda+\mu)(s-1)}$, so that the result follows by uniqueness. □

A similar argument implies the following result.

Example 4.9. *If $X \sim \mathsf{Bin}(n,p)$ and $Y \sim \mathsf{Bin}(m,p)$ are independent, then $X + Y \sim \mathsf{Bin}(n + m, p)$.*

Another useful property of probability generating function $G_X(s)$ is that it can be used to compute moments of X:

[34] This is a two-stage probabilistic experiment!

✍ **Theorem 4.10.** *If X has generating function $G(s)$, then*[35]

$$\mathsf{E}\big[X(X-1)\dots(X-k+1)\big] = G^{(k)}(1)\,.$$

Remark 4.10.1. The quantity $\mathsf{E}\big[X(X-1)\dots(X-k+1)\big]$ is called the kth factorial moment of X. Notice also that

$$\mathsf{Var}(X) = G''_X(1) + G'_X(1) - \big(G'_X(1)\big)^2\,. \tag{4.6}$$

Proof. Fix $s \in (0,1)$ and differentiate $G(s)$ k times[36] to get

$$G^{(k)}(s) = \mathsf{E}\big[s^{X-k}X(X-1)\dots(X-k+1)\big]\,.$$

Taking the limit $s \uparrow 1$ and using the Abel theorem,[37] we obtain the result. □

Remark 4.10.2. Notice also that

$$\lim_{s\nearrow 1} G_X(s) \equiv \lim_{s\nearrow 1} \mathsf{E}[s^X] = \mathsf{P}(X<\infty)\,.$$

This allows us to check whether a variable is finite, if we do not know this *apriori.*

Exercise 4.11. Let S_N be defined as in Example 4.5. Use (4.4) to compute $\mathsf{E}\big[S_N\big]$ and $\mathsf{Var}\big[S_N\big]$ in terms of $\mathsf{E}[N]$, $\mathsf{E}[N]$, $\mathsf{Var}[X]$ and $\mathsf{Var}[N]$. Now check your result for $\mathsf{E}\big[S_N\big]$ and $\mathsf{Var}\big[S_N\big]$ by directly applying the partition theorem for expectations.

Generating functions are also very useful in solving recurrences, especially when combined with the following algebraic fact.[38]

Lemma 4.12 (Partial fraction expansion)**.** *Let $f(x) = g(x)/h(x)$ be a ratio of two polynomials without common roots. Let $\deg(g) < \deg(h) = m$ and suppose that the roots $a_1, \dots, a_m$ of $h(x)$ are all distinct. Then $f(x)$ can be decomposed into a sum of partial fractions, ie., for some constants $b_1, b_2, \dots, b_m$,*

$$f(x) = \frac{b_1}{a_1 - x} + \frac{b_2}{a_2 - x} + \dots + \frac{b_m}{a_m - x}\,. \tag{4.7}$$

Remark 4.12.1. Since

$$\frac{b}{a-x} = \frac{b}{a}\sum_{k\ge 0}\Big(\frac{x}{a}\Big)^k = \sum_{k\ge 0}\frac{b}{a^{k+1}}x^k\,,$$

a generating function of the form (4.7) can be easily written as a power series.

---

[35] here, if $G^{(k)}(1)$ does not exists we understand the RHS of the equation as $G^{(k)}(1-) \equiv \lim_{s\uparrow 1} G^{(k)}(s)$, the limiting value of the kth left derivative of $G(s)$ at $s=1$;

[36] As $|G_X(s)| \le \mathsf{E}|s^X| \le 1$ for all $|s| \le 1$, the generating function $G_X(s)$ can be differentiated many times for all s inside the disk $\{s : |s| < 1\}$.

[37] Theorem A.12 below; by footnote 36, it applies to all probability generating functions.

[38] An alternative way would be to use products of matrices; get in touch, if interested!

Example 4.13. *Let a_n be the probability that n independent Bernoulli trials (with success probability p) result in an* even number of successes. *Find the generating function of a_n.*

Solution. The event under consideration occurs if an initial failure at the first trial is followed by an even number of successes or if an initial success is followed by an odd number of successes. Therefore, $a_0 = 1$ and $a_n = q\,a_{n-1} + p\,(1 - a_{n-1})$ for all $n \geq 1$, where $q = 1 - p$. Multiplying these equalities by s^n and adding them we get

$$G_a(s) - 1 = qs\,G_a(s) + p\sum_{n\geq 1} s^n - ps\,G_a(s) = (q-p)s\,G_a(s) + \frac{ps}{1-s}\,,$$

and after rearranging,

$$G_a(s) = \Bigl(1 + \frac{ps}{1-s}\Bigr)/\bigl(1 - (q-p)s\bigr) = \frac{1}{2}\Bigl(\frac{1}{1-s} + \frac{1}{1-(q-p)s}\Bigr)\,.$$

As a result, $a_n = \bigl(1 + (q-p)^n\bigr)/2$. □

Example 4.14. *A biased coin is tossed repeatedly; on each toss, it shows a "head" with probability p. Let r_n be the probability that a sequence of n tosses never has two "heads" in a row. Show that $r_0 = 1$, $r_1 = 1$, and for all $n > 1$, $r_n = qr_{n-1} + pqr_{n-2}$, where $q = 1 - p$. Deduce the generating function of the sequence $(r_n)_{n\geq 0}$.*

Solution. Every sequence of $n \geq 2$ tosses starts either with T or with HT; hence the relation. Multiplying these equalities by s^n and summing, we get

$$G_r(s) = \sum_{n\geq 0} r_n s^n = 1 + s + qs\sum_{n\geq 2} r_{n-1}s^{n-1} + pqs^2\sum_{n\geq 2} r_{n-2}s^{n-2}$$

so that

$$G_r(s) = \frac{1+ps}{1-qs-pqs^2}\,. \qquad \square$$

Theorem 4.15 (Continuity Theorem)**.** *Let for every fixed n the sequence $a_{0,n}$, $a_{1,n}$, ... be a probability distribution, ie., $a_{k,n} \geq 0$ and $\sum_{k\geq 0} a_{k,n} = 1$, and let $G_n(s)$ be the corresponding generating function, $G_n(s) = \sum_{k\geq 0} a_{k,n}s^k$. In order that for every fixed k*

$$\lim_{n\to\infty} a_{k,n} = a_k \tag{4.8}$$

it is necessary and sufficient that for every $s \in [0, 1)$,

$$\lim_{n\to\infty} G_n(s) = G(s)\,,$$

where $G(s) = \sum_{k\geq 0} a_k s^k$, the generating function of the limiting sequence (a_k).

✍ **Remark 4.15.1.** The convergence in (4.8) is known as convergence in distribution!

Example 4.16. *If $X_n \sim \mathsf{Bin}(n,p)$ with $p = p_n$ satisfying $n \cdot p_n \to \lambda$ as $n \to \infty$, then*

$$G_{X_n}(s) \equiv \bigl(1 + p_n(s-1)\bigr)^n \to \exp\{\lambda(s-1)\}\,,$$

so that the distribution of X_n converges to that of $X \sim \mathsf{Poi}(\lambda)$.