

A crash course in generating functions

This sheet is based on notes written by Prof. Wilfrid Kendall and aims to remind you of some material from ST112 Probability B. Any errors, however, are my responsibility. Please send comments and corrections to C.A.Goldschmidt@warwick.ac.uk.

Given a bounded sequence of numbers a_0, a_1, \dots , we can form the *generating function*

$$G(s) = \sum_{n=0}^{\infty} a_n s^n.$$

Here, s is a dummy variable which, for our purposes, will usually take non-negative real values in some interval.¹ You should think of the function $G(s)$ as a convenient way of encoding the sequence. Often, we can figure out a formula for $G(s)$ by using the theory of sums and series. For example:

- If $a_0 = 1/2$, $a_1 = 1/2$ and $a_i = 0$ for all $i \geq 2$, then $G(s) = (1 + s)/2$. Notice that $G(s)$ is well-defined for all $s \in [0, \infty)$.
- If $a_n = 2^{-n-1}$ for $n \geq 0$ then

$$G(s) = \sum_{n=0}^{\infty} 2^{-n-1} s^n = \frac{1}{2-s}.$$

In this case, $G(s)$ is well-defined for $s \in [0, 2)$.

Notice that starting from a generating function $G(s)$ we can get back to the sequence a_0, a_1, \dots by differentiation:

$$\begin{aligned} a_0 &= G(0) \\ a_1 &= G'(0) \\ a_2 &= \frac{1}{2} G''(0) \\ &\vdots \\ a_n &= \frac{1}{n!} G^{(n)}(0), \end{aligned}$$

where $G^{(n)}$ is the n th derivative of G .

We often use generating functions to encode the sequence of probabilities of a non-negative integer-valued random variable:

$$G_X(s) = \sum_{n=0}^{\infty} s^n \mathbb{P}(X = n).$$

¹Since the sequence a_0, a_1, \dots of coefficients is assumed here to be bounded, this range will always include $[0, 1)$.

In such cases, we call $G_X(s)$ the *probability generating function* of X . Notice that we can also think of $G_X(s)$ as $\mathbb{E}[s^X]$.

This is, however, not the only use of generating functions we will see in this course. Consider the definition of $P_{ij}(s)$ in the section of the lecture notes on the first-passage decomposition:

$$P_{ij}(s) = \sum_{n=0}^{\infty} p_{ij}^{(n)} s^n.$$

Although the $p_{ij}^{(n)}$ are probabilities, they are not probabilities of disjoint events concerning a single random variable (just because a Markov chain can travel from i to j in 5 steps, that does not necessarily prevent it from being back at j after 10 steps). In particular, there is no reason that they should sum to 1 when we vary n !

If G_X is the probability generating function of some random variable X taking values in the non-negative integers, then

$$\begin{aligned} G(1) &= 1 \\ G'(1) &= \mathbb{E}[X] \end{aligned}$$

and it's possible to obtain the variance from the values of $G'(1)$ and $G''(1)$. (Of course, if X actually has infinite expectation then we will also find that $G'(1)$ is infinite).

One of the very useful features of probability generating functions is that if X and Y are independent non-negative integer-valued random variables then

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

In many circumstances, this is much easier to handle than the probability of $X+Y$ taking a particular value n . In that case, we would have to deal with a *convolution sum*:

$$\mathbb{P}(X + Y = n) = \sum_{k=0}^n \mathbb{P}(X = k) \mathbb{P}(Y = n - k).$$

Another example of this idea is to be found when we convert the first-passage decomposition into its generating function form:

$$P_{ij}(s) = p_{ij}^{(0)} + F_{ij}(s)P_{jj}(s).$$

One more situation where generating functions are invaluable is when we consider a random sum of independent and identically distributed random variables, X_1, X_2, \dots . Let N be the random number of summands, i.e. consider

$$S = \sum_{k=1}^N X_k.$$

Then an argument using conditional expectation (exercise!) shows that

$$G_S(s) = G_N(G_{X_1}(s)).$$