Generating Functions and Random Walks

We start with one more quick example of probability generating functions. Suppose that we flip a biased coin (with success probability p), and record the number of flips up until and including the first success. The time until first success is a geometric random variable. (We came across such a random variable before in the coupon collector's problem.) What is the expected time until the first success? Let Y be the time of first success. Then, $\mathbb{P}(Y = k) = p(1-p)^{k-1}$ for $k \geq 1$.

Exercise 1. *Let $p_k = \mathbb{P}(Y = k)$, and write the probability generating function $q(x) = \sum_{k=0}^{\infty} p_k x^k$ of the above distribution in closed form.* 2

Exercise 2. *Show that $\mathbb{E}(Y) = q'(1)$, and use your function from the previous exercise to find $\mathbb{E}(Y)$ as a function of p.* 2

Standard 1D random walk

Now, we'd like to examine a standard, balanced 1-dimensional walk in more detail, and in particular show that it is guaranteed to return to the origin in finite time. We start at the origin at time 0, and we let $X_1, X_2, \ldots$ take values -1 and 1 with equal probability, independently of each other. We set $S_0 = 0$, and $S_n = \sum_{i=1}^{n} X_i$. This indicates the position of the random walker at time n.
We let

$$p_{2n} = \mathbb{P}(S_{2n} = 0),$$

the probability that the walker hits 0 after $2n$ steps. Note that $p_{2n+1} = 0$ always.

Exercise 3. *Show that we have* 1

$$p_{2n} = \binom{2n}{n} 2^{-2n}.$$

Now, let q_{2n} be the probability that the walker hits 0 for the first time after $2n$ steps (not counting the start at 0: let $q_0 = 0$).

Exercise 4. *Show that for $n \geq 1$, we have* 1

$$p_{2n} = q_0 p_{2n} + q_2 p_{2n-1} + \ldots + q_{2n} p_0.$$

Let the probability generating functions be defined as usual, as

$$p(x) = \sum_{n=0}^{\infty} p_{2n} x^n, \quad q(x) = \sum_{n=0}^{\infty} q_{2n} x^n.$$

Exercise 5. *Use Stirling's approximation to show that $\binom{2n}{n} \sim \frac{4^n}{\sqrt{\pi n}}$.* 2

Exercise 6. *Show that $p(x)$ converges in the interval $(-1, 1)$.* 1

Exercise 7. *Show that for $n \geq 1$, we have* 1

$$p_{2n} = q_0 p_{2n} + q_2 p_{2n-1} + \ldots + q_{2n} p_0.$$

By this above relation, we see that

$$p(x) = 1 + q(x)p(x).$$

We add the 1, because $p_0 = 1$, and the recursive formula was only valid for $n \geq 1$.

Exercise 8. *Use standard series to show that, for small enough x,* 3

$$p(x) = \frac{1}{\sqrt{1-x}}$$

and

$$q(x) = 1 - \sqrt{1-x}.$$

Exercise 9. *Show that $q'(x) = \frac{p(x)}{2}$, and use this to derive a direct formula for q_{2n}.* 2

We continue the analysis of the 1-dimensional walk. We're interested in the probability of eventual return to the origin. We let η_n be the probability that our random walker has returned to 0 by time n (either at time n or earlier). Then, we let

$$\eta = lim_{n\to\infty}\eta_n.$$

Compare this to the extinction probability η in the branching processes. The sequence $\eta_1, \eta_2 \ldots$ is increasing and bounded (by 1), and therefore has a limit. We have

$$\eta_{2n} = \sum_{i=1}^{n} q_{2n}, \quad \eta = \sum_{i=1}^{\infty} q_{2n}.$$

From the definition of η, we should have that $q(x)$ converges at 1, and we obtain

$$\eta = q(1) = 1.$$

(Slightly more precisely, we have $\eta = \sum_{n=0}^{\infty} q_{2n} = \lim_{x\uparrow 1} q(x)$, and then we use the fact that $q(x)$ converges and is left-continuous at 1.)

Number of returns in a 1D random walk

So far we have only thought about the probability of returning in some finite time: in 1D the walker always returns in finite time, implying that it returns infinitely often. The next natural question is: how often to we expect the walker to return in the time interval $(0, t)$?

Let r_{2n} be the number of returns on all walks of length $2n$ (note this is not a random variable). We let R_{2n} be the random variable that counts the number of returns in a random walk of length $2n$. Then

$$\mathbb{E}(R_{2n}) = \frac{r_{2n}}{2^{2n}}.$$

As usual, we consider the generating function

$$r(x) = \sum_{k=0}^{\infty} r_{2k} x^k.$$

We partition all walks of length $2n$ into sets W_{2k}: the set of walks with a first return at time $2k$, and a set Z (the walks of length $2n$ that do not return). The set of walks of length $2n$ is the set

$$W_2 \cup W_4 \cup \ldots \cup W_{2n} \cup Z.$$

We have

$$|W_{2k}| = q_{2k} 2^{2n}.$$

(Why?) The total number of returns in the set of walks W_{2k} is

$$|W_{2k}| + 2^{2k} q_{2k} r_{2n-2k}.$$

Exercise 10. *Show that* 2

$$r_{2n} = \sum_{k=0}^{\infty} \left(|W_{2k}| + 2^{2k} q_{2k} r_{2n-2k} \right) = \sum_{k=0}^{\infty} \left(q_{2k} 2^{2n} + 2^{2k} q_{2k} r_{2n-2k} \right).$$

Exercise 11. *Show that* $q_{2n} = p_{2n-2} - p_{2n}$, *and then show that this implies that* $\sum_{k=0}^{n} q_{2k} = 1 - p_{2n}$. 2

Exercise 12. *Show that this gives us* 2

$$r(x) = \frac{1}{1-4x} - p(4x) + q(4x) r(x),$$

Exercise 13. *Finally, solve for* $r(x)$ *to obtain* 2

$$r(x) = \frac{1}{(1-4x)^{3/2}} - \frac{1}{1-4x}.$$

Exercise 14. *It is helpful to recognize the first term as* $\frac{1}{2} p'(4x)$. *Use this to find the direct expression for* r_{2n} *as* 2

$$r_{2n} p_{2n+2} 2^{2n+1}(n+1) + 2^{2n} = \frac{1}{2} \binom{2n+2}{n+1} (n+1) - 2^{2n} \sim \sqrt{\frac{2}{\pi}} \sqrt{2n},$$

using Stirling approximation for the last step. Notice that the number of returns only grows as the square root of time.