Chapter 9

Random Walks - Stopping Times

In the simple random walk, with absorbing barriers at $x = 0$ and $x = N$, the probability of losing (ruin) starting at $x = k$ was shown to be:

$$p_k = \frac{(q/p)^k - (q/p)^N}{1 - (q/p)^N} \qquad 0 \leq k \leq N \qquad . \tag{9.1}$$

where p is probability of winning, $(x \to x+1)$, and q the probability of losing $(x \to x-1)$, at each step.

Suppose D_k defines the 'duration' of the walk. That is, the number of steps before the game ends (win or lose), starting from $x = k$. Clearly D_k is a discrete random variable. Denoting the walker's starting position as S_0, then consider the 'expected' duration of this walk:

$$d_k \equiv \mathbb{E}\left(D_k | S_0 = k\right) \qquad . \tag{9.2}$$

Let B be the event that the *first* game is won: then, $P(B) = p$, $P(B^c) = q$.

Conditioning on the first game, we have:

$$\mathbb{E}\left(D_k\right) = \mathbb{E}\left(D_k|B\right) P(B) + \mathbb{E}\left(D_k|B^c\right) P(B^c) \qquad 1 \leq k \leq N-1. \tag{9.3}$$

We note that, when $k = 0$ or $k = N$, the duration of the walk is zero. That is, if we start with no money, or all the money we want, we don't play any games (take any steps).

$$d_0 = 0 \qquad , \qquad d_N = 0 \quad . \tag{9.4}$$

This provides us with boundary conditions for d_k. It follows that, given we win the first game we move to $k+1$ and continue playing. Thus, given B, we have taken one step right and:

$$\mathbb{E}\left(D_k|B\right) = 1 + \mathbb{E}\left(D_{k+1}\right) = 1 + d_{k+1} \qquad . \tag{9.5}$$

and

$$\mathbb{E}\left(D_k|B^c\right) = 1 + \mathbb{E}\left(D_{k-1}\right) = 1 + d_{k-1} \qquad . \tag{9.6}$$

$$d_k = (1 + d_{k+1})p + (1 + d_{k-1})q \qquad . \tag{9.7}$$

Since $p + q = 1$, this can be written in the form:

$$d_k - pd_{k+1} - qd_{k-1} = 1 \qquad . \tag{9.8}$$

This 'difference' equation is identical to the difference equation for p_k, except that the right-hand-side is non-zero. It is an *inhomogeneous* difference equation.

As previously discussed, the difference equation is a linear equation. In this case, we can write the expression in matrix form:

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ -q & 1 & -p & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \cdots & -q & 1 & -p & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & \cdots & 0 & -q & 1 & -p \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} d_0 \\ d_1 \\ \vdots \\ d_k \\ \vdots \\ d_{N-1} \\ d_N \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} \tag{9.9}$$

A linear equation of the form: $\mathsf{A}\mathsf{d} = \mathsf{b}$, the general solution of the equation is

$$\mathsf{d} = \mathsf{d}(\text{homogeneous.}) + \mathsf{d}(\text{particular solution}) \tag{9.10}$$

The homogeneous solution is the *kernel* (or *nullspace*) of the matrix, A, and this is identical to the solution found before, for p_k.

$$d_k(\text{homo.}) = a_1 1^k + a_2(q/p)^k \tag{9.11}$$

For the particular solution, we 'guess' the following expression:

$$d_k(\text{part}) = bk \qquad , \tag{9.12}$$

where, b, is some constant.It can be found from the original difference equation

$$-pd_{k+1} + d_k - qd_{k-1} = 1 \quad . \tag{9.13}$$

$$-pb(k+1) + bk - qb(k-1) = 1 \quad . \tag{9.14}$$

and after rearrangement this gives,

$$b = \frac{1}{(q-p)} \qquad . \tag{9.15}$$

Therefore the general solution is:

$$d_k = a_1 + a_2(q/p)^k + \frac{k}{(q-p)} \qquad . \tag{9.16}$$

Applying the conditions at the boundary on the left and right, respectively, gives the equations:

$$0 = a_1 + a_2. \quad , \quad 0 = a_1 + a_2(q/p)^N + \frac{N}{(q-p)}. \tag{9.17}$$

and the values of $a_{1,2}$ can be found. This results in the expression:

$$d_k = \frac{k}{(q-p)} - \frac{N}{(q-p)} \frac{(1-(q/p)^k)}{(1-(q/p)^N)} \qquad , \qquad (q/p) \neq 1 \qquad . \tag{9.18}$$

Figure (9.1) displays values of d_k for $N = 10$ for values of $p = 0.4$ and $p = 0.5$.

For $p = 0.4$, the maximum expected duration is for walks starting at $k = 6, 7.$ for which $d_k \approx 20$ and the shortest walk starts from $k = 1$ with $d_1 = 4.6$ steps.

The duration of the symmetric walk ($q/p = 1$) can again be derived by using L'Hôpital's rule (twice!). A more convenient approach is to let $q/p = 1 + \varepsilon$, and make a series expansion in the small parameter ε, eventually letting it tend to zero. In any case, we arrive at the result:

$$d_k = k(N-k) \quad . \tag{9.19}$$

Not surprisingly, as we would expect, and as is shown in the figure, the maximum duration is a walk that begins at the centre, $k = \frac{1}{2}N$, in which case:

$$d_{N/2} = \tfrac{1}{4}N^2 \quad . \tag{9.20}$$

The important conclusion is that the time to reach the boundary is proportional to the *square* of the distance to the boundary. This is characteristic of *diffusion processes* (also called *Brownian motion*). Compare this with a *deterministic process*, for example, suppose the walker *always* steps right, that is $q = 0$. Then the time taken to reach the boundary starting at the middle would be, with certainty, $d_{N/2} = \frac{1}{2}N$.

9.1 Reflecting boundary and the method of images

It is straightforward to extend this theory to the case where the walker is *not* absorbed at $x = 0$. Instead, suppose that each, and every time, the gambler runs out of money (reaches $x = 0$), a generous benefactor gives him £1. That is, the walker finding himself at $x = 0$, moves back to $x = 1$ on the next step, ready to continue playing. The gambler is never terminates the walk at $x = 0$. Under such conditions, the only place where the walk can terminate is the absorbing boundary at $x = N$. The question arises as to how long, on average, would such a game last.

Suppose R_k is the random variable for the duration of the walk, with a perfect 'reflecting boundary' at $x = 0$, starting from $x = k$. Our interest is in calculating an expression for: $r_k \equiv \mathbb{E}\left(R_k\right)$. This is done by the usual trick of conditioning on the first step away from the starting point, $x = k$.

$$r_k = \mathbb{E}\left(R_k\right) = \mathbb{E}\left(R_k|B\right) P(B) + \mathbb{E}\left(R_k|B^c\right) P(B^c) \qquad 1 \le k \le N-1 \qquad . \tag{9.21}$$

with $r_N = 0$, as a boundary condition at the absorbing boundary. Then we have:

$$\mathbb{E}\left(R_k|B\right) = 1 + \mathbb{E}\left(R_{k+1}\right) = 1 + r_{k+1} \quad , \quad 1 \le k \le N-1 \qquad . \tag{9.22}$$

and following the same argument.

$$\mathbb{E}\left(R_k|B^c\right) = 1 + r_{k-1} \qquad . \tag{9.23}$$

This leads to the same difference equation, when both boundaries are absorbing.

$$r_k = (1 + r_{k+1})p + (1 + r_{k-1})q \quad , \quad 1 \le k \le N-1 \qquad . \tag{9.24}$$

But now the boundary conditions have changed. If the gambler starts at $x = 0$, then - with certainty - they must take one step to $x = 1$ and start the walk from that point. In mathematical terms, a walker starting at $x = 0$ takes one step more than a walker at $x = 1$:

$$r_0 = 1 + r_1 \qquad . \tag{9.25}$$

In a similar manner, the general solution is

$$r_k = \frac{k}{(q-p)} + a_1 + a_2(q/p)^k \quad , \quad q/p \neq 1 \quad . \tag{9.26}$$

with the boundary conditions at the reflecting and absorbing boundaries:

$$a_1 + a_2 = 1 + \frac{1}{(q-p)} + a_1 + a_2(q/p) \qquad , \qquad 0 = \frac{N}{(q-p)} + a_1 + a_2(q/p)^N \qquad . \tag{9.27}$$

The first relation gives the:

$$a_2 = \frac{-2(q/p)}{(1-(q/p))^2} \qquad , \tag{9.28}$$

then

$$a_1 = -\frac{N}{(q-p)} + 2\frac{(q/p)^{N+1}}{(1-(q/p))^2} \qquad . \tag{9.29}$$

$$r_k = \frac{(k-N)}{(q-p)} - 2\frac{[(q/p)^{k+1} - (q/p)^{N+1}]}{(1-(q/p))^2} \tag{9.30}$$

while in the case $p \to q \to \frac{1}{2}$, and applying L'Hôpital's rule we get, for the expected duration of a symmetric walk with perfect reflection at $x = 0$:

$$r_k = (N-k)(N+k) \qquad , \qquad q = p = \tfrac{1}{2} \qquad . \tag{9.31}$$

In particular we have:

$$r_0 = N^2 \qquad . \tag{9.32}$$

The relation of the reflecting barrier to the absorbing barrier (9.20) will be explained in more detail in the next chapter. For the present we note the following important remark on the *method of images*.

Suppose the reflecting barrier (mirror) were replaced by: a transparent barrier (window) at $x = 0$, and an absorbing barrier at $x = -N$. Then a walker arriving at $x = 0$ would either move to $x = +1$ or $x = -1$, with equal probability. That is the walker splits into 'two' compared to the case of the reflected walker: the *real* walker and its *doppelganger* at $x = -1$. This mirror image of the walker. Then the mirror image will imitate (exactly) the random motion of the real walker, so long as we include an extra absorber at $x = -N$. So to calculate the expected time to absorption for a symmetric walk with a reflecting barrier starting at $x = k$. we simply calculate the expected duration of a walk for a absorbing barriers at $x = 0$, and $x = 2N$, starting at $x = N + k$. This is:

$$d_k = (N + k)(2N - N - k) = (N + k)(N - k) \quad , \tag{9.33}$$

as given above (9.31).

9.2 Potential theory

Let us briefly mention the connection with diffusion processes and potential theory; a subject that will be discussed at length later. If we consider the difference equation for the simple random walk (8.12) with symmetry:

$$p_k = \tfrac{1}{2}(p_{k+1} + p_{k-1}) \quad . \tag{9.34}$$

with solution (8.35):

$$p_k = 1 - \left(\frac{k}{N}\right) \tag{9.35}$$

Suppose we now let x be a continuous variable, that is

$$x = k \to x \quad , \quad x = k + 1 \to x + h \quad , \quad x = k - 1 \to x - h \quad .$$

and let $p_k \to p(x)$. Then we have:

$$p(x) = \tfrac{1}{2}(p(x + h) + p(x - h)) \quad . \tag{9.36}$$

Now making a Taylor expansion of the terms on the right-hand side:

$$\begin{aligned} p(x) \quad = \quad & \tfrac{1}{2}\left(p(x) + hp'(x) + \tfrac{h^2}{2!}p''(x) + \tfrac{h^3}{3!}p'''(x) + \tfrac{h^4}{4!}p''''(x) + \cdots\right) \\ & +\tfrac{1}{2}\left(p(x) - hp'(x) + \tfrac{h^2}{2!}p''(x) - \tfrac{h^3}{3!}p'''(x) + \tfrac{h^4}{4!}p''''(x) \cdots\right) \quad . \end{aligned} \tag{9.37}$$

This simplifies to:

$$0 = \frac{h^2}{2!}p''(x) + \frac{h^4}{4!}p''''(x) + \cdots \quad . \tag{9.38}$$

Clearly, in the limit $h \to 0$, we have $0 = 0$. However, if we consider small but finite h we see that, to leading order, we have:

$$\frac{d^2p}{dx^2} = 0 \tag{9.39}$$

which is Laplace's equation (in one dimension). Considering $x = k$ as a continuous variable, we see indeed that the function (9.35):

$$p_x = 1 - \left(\frac{x}{N}\right) \quad , \tag{9.40}$$

is a solution of this equation, as it should be.

Laplace's equation is used to describe the *potentials* of (incompressible) fluid flow, static electric fields and gravitational fields. For example the equation for the flow of inviscid water along a pipe is described by the equation:

$$\frac{d^2\phi}{dx^2} = 0 \quad , \tag{9.41}$$

where x is the distance along the pipe and, ϕ is the fluid velocity potential, related to the velocity v, by $v = d\phi/dx$. The study of Laplace's equation (in many variables and various geometries) and its solutions describes a branch of mathematics called *potential theory.*

Let us compare this equation with the equation describing the equilibrium temperature distribution of a bar (thin rod) u as a function of position, x:

$$\frac{d^2u}{dx^2} = 0 \qquad . \tag{9.42}$$

Again, this is Laplace's equation with solution:

$$u(x) = Ax + B \qquad . \tag{9.43}$$

Suppose the rod is of length N and we have the boundary conditions in which one end of the rod, $x = 0$, has a temperature $u = 1$ and the other end, $x = N$, is maintained at zero temperature: $u = 0$. Then we can determine the constants A and B with the solution:

$$u(x) = 1 - \left(\frac{x}{N}\right) \qquad . \tag{9.44}$$

Comparing equations (9.39) and (9.42), we see that mathematically, calculating the temperature distribution of a metal rod is identical to calculating the probability of ruin (tossing a fair coin). This might seem far-fetched, but it is mathematically true. It is simply that the meaning and interpretation is different. This has several important advantages. If we make appropriate transformations, the physics of heat/fluid flow can assist us in visualising the flow of probability. Indeed the powerful mathematical (and computational) techniques used to study fluid dynamics can be transferred to the theory of probability. The analogy will be reinforced later when we discuss the Kolmogorov equation for time-dependent (dynamic) probability.

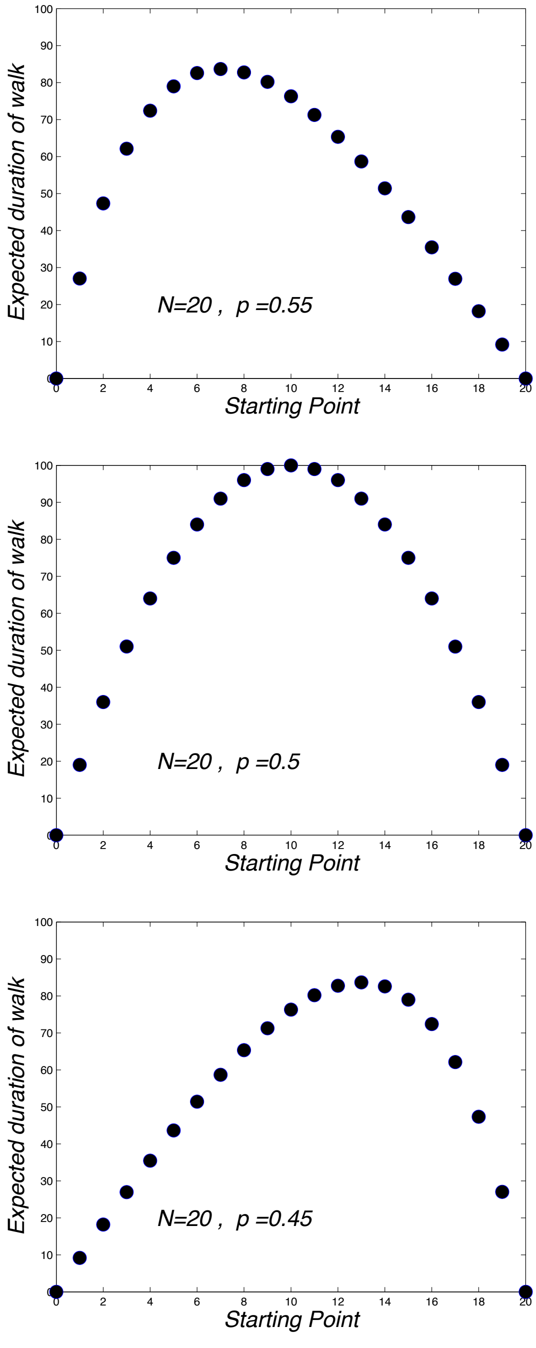

Figure 9.1: . Expected duration of the walk d_k as a function of starting point, k, for $N = 20$. Top; $p = 0.55$, (biased in favour), Middle $p = 0.50$ (fair game), Bottom $p = 0.45$ (biased against).