Random walk, gambler's ruin and other good stuff.

Suppose $0 < p < 1$ and let $q = 1 - p$. Let

$$X_1, X_2, \ldots, X_n, \ldots$$

be a sequence of mutually independent random variables with values in $\{-1, 1\}$ such that

$$P(X_n = 1) = p, \qquad i = 1, 2, \ldots, n, \ldots.$$

It is not mathematically obvious that such a sequence of random variables exists; in fact, the sample space must be uncountable. It is, however, intuitively clear that such a sequence, or something like it, exists.

For each $n = 1, 2, \ldots$ let

$$S_n = \sum_{i=1}^{n} X_i.$$

The sequence $S_n,\ n = 1, 2, \ldots$ is an example of what is called a **simple random walk** and is extraordinarily useful.

Consider a game in which you win or lose a dollar at the n-th play if X_n is 1 or -1, respectively. You start playing with a fortune of f dollars with the goal of winning w dollars. You quit if either you are broke or you have won w dollars. What is the probability you will win? Go broke? How long will you play? We now proceed to answer all these questions.

Let

$$T$$

be the duration of play. That is, for each positive integer n, $T = n$ if $-f < S_j < w$ for all $j < n$ and either $S_n = -f$ or $S_n = w$, and $T = \infty$ if you play forever because you never win and are never ruined. Here is an important fact about T: the event $\{T < n\}$ depends only on $X_1, \ldots, X_{n-1}$. This means, by definition, that you can determine the value of T at a sample point if you know the values of $X_1, \ldots, X_{n-1}$ at that sample point, which, I hope, is clearly the case. This implies that the complementary event

$$\{T \geq n\} \text{ is independent of } X_n, X_{n+1}, \ldots. \tag{1}$$

We now prove the basic

Theorem.

$$P(T = \infty) = 0. \tag{2}$$

Proof. Let $N = f + w - 1$ and let $\lambda = 1 - (p^n + q^n)$. Let E_j be the event that not all of $X_{jN+1}, \ldots, X_{(j+1)N}$ are the same; note that

$$P(E_j) = \lambda.$$

Now

$$\{T \geq (j+1)N\} \subset \{T \geq jN\} \cap E_j$$

so

$$P(T \geq (j+1)N) \leq P(\{T \geq jN\} \cap E_j) = P(T \geq jN)\, P(E_j) \leq \lambda\, P(T \geq jN)$$

where we have used (1) to obtain the equality. It follows that

$$P(T \geq jN) \leq \lambda^j \qquad \text{for } j = 0, 1, 2, \ldots.$$

Since $\{T = \infty\} \subset \{T \geq jN\}$ for all nonnegative integers j and since $\lim_{j \to \infty} \lambda^j = 0$ we are done. $\square$

Let us define

$$S_T \;=\; \sum_{i=1}^{T} X_i.$$

By virtue of (2) we find that

$$P(S_T \;=\; f) \;+\; P(S_T \;=\; w) \;=\; 1.$$

Wald's identities.

(3)
$$E(S_T) \;=\; (p-q)\,E(T).$$

(4)
$$\mathrm{Var}(S_T) \;=\; E(T) \qquad\qquad \text{provided } p=q.$$

Proof. Using (1) we obtain

$$\begin{aligned} E(S_T) \;&=\; E(\sum_{i=1}^{\infty} X_i\,1_{\{T\geq i\}}) \\ &=\; \sum_{i=1}^{\infty} E(X_i\,1_{\{T\geq i\}}) \\ &=\; \sum_{i=1}^{\infty} E(X_i)\,E(1_{\{T\geq i\}}) \\ &=\; \sum_{i=1}^{\infty}(p-q)\,P(T\geq i) \\ &=\; (p-q)\,E(T), \end{aligned}$$

proving (3). To prove (4), we assume that $p=q$ and use (3) and (1) to calculate

$$\begin{aligned} \mathrm{Var}(S_T) \;&=\; E(S_T^2) \\ &=\; E(\sum_{i=1}^{\infty} X_i\,1_{\{T\geq i\}} \sum_{j=1}^{\infty} X_j\,1_{\{T\geq j\}}) \\ &=\; \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} E(X_i\,1_{\{T\geq i\}}\,X_j\,1_{\{T\geq j\}}) \\ &=\; \sum_{i=1}^{\infty} E(X_i^2\,1_{\{T\geq i\}}^2) \\ &=\; \sum_{i=1}^{\infty} E(\,1_{\{T\geq i\}}) \\ &=\; E(T), \end{aligned}$$

where we have used that fact that

$$E(X_i\,1_{\{T\geq i\}}\,X_j\,1_{\{T\geq j\}}) \;=\; 0 \qquad\qquad \text{if } i\neq j;$$

indeed, if $i<j$,

$$E(X_i\,1_{\{T\geq i\}}\,X_j\,1_{\{T\geq j\}}) \;=\; E(X_i\,1_{\{T\geq i\}}\,1_{\{T\geq j\}})\,E(X_j) \;=\; 0.$$

Let

$$p_{\text{ruin}} \;=\; P(S_T \;=\; -f) \qquad \text{and} \qquad p_{\text{win}} \;=\; P(S_T \;=\; w).$$

All questions answered in case $p = q$. It follows directly from (3) that

$$-f\,p_{\text{ruin}} + w\,p_{\text{win}} = E(S_T) = 0$$

and (2) implies

$$p_{\text{ruin}} + p_{\text{win}} = 1.$$

Consequently,

$$p_{\text{ruin}} = \frac{w}{f+w} \qquad \text{and} \qquad p_{\text{win}} = \frac{f}{f+w}.$$

This in turn implies that

$$E(T) = \operatorname{Var}(S_T) = E(S_T^2) = f^2\,p_{\text{ruin}} + w^2\,p_{\text{win}} = f\,w.$$

All questions answered in case $p \neq q$. We now need to vary f and w. For each integer n greater than $-f$ and less than w let

$$g_n$$

be the probability of winning starting with an initial fortune of $f+n$ dollars and trying to win $w-n$ dollars. We also

$$g_{-f} = 0 \quad \text{and} \quad g_w = 1.$$

The key fact is that

$$g_n = p\,g_{n+1} + q\,g_{n-1}, \qquad -f < n < w; \tag{5}$$

we leave it to the reader to prove this by conditioning on the outcome of the first play; make sure you understand this as it's the basis of everything that follows. Since $p+q=1$ this gives

$$g_{n+1} - g_n = \rho\,(g_n - g_{n-1}) \qquad -f < n < w$$

which implies that

$$g_{n+1} - g_n = \rho^{f+n}\,g_{-f+1}, \qquad -f \le n < w,$$

where we have set

$$\rho = \frac{q}{p} \neq 1.$$

Using the formula for summing a geometric progression we find that

$$g_n = g_{-f+1}\,\frac{1-\rho^{f+n}}{1-\rho}, \qquad -f \le n \le w.$$

Since $g_w = 1$ we find that

$$g_{-f+1} = \frac{1-\rho}{1-\rho^{f+w}}$$

so that

$$g_n = \frac{1-\rho}{1-\rho^{f+w}}\,\frac{1-\rho^{f+n}}{1-\rho} = \frac{1-\rho^{f+n}}{1-\rho^{f+w}}, \qquad -f \le n \le w.$$

Thus

$$p_{\text{win}} = g_0 = \frac{1-\rho^f}{1-\rho^{f+w}} \quad \text{and} \quad p_{\text{ruin}} = 1 - p_{\text{win}} = 1 - \frac{1-\rho^f}{1-\rho^{f+w}} = \rho^f\,\frac{1-\rho^w}{1-\rho^{f+w}}.$$

Using L'Hôpital's rule we note that these converge as $\rho \to 1$ to $\frac{f}{f+w}$ and $\frac{w}{f+w}$, respectively, which is nice. This immediately gives

$$E(S_T) = -f p_{\text{ruin}} + w p_{\text{win}}.$$

From (3) we infer that

$$E(T) = \frac{1}{p-q}(-f p_{\text{ruin}} + w p_{\text{win}}).$$