Math 464/564, O. Knill, **April 1997**

1D walk (2), Ballot and Arc Sin theorem.

REFLECTION PRINCIPLE. For $a, b \in \mathbb{N} \setminus \{0\}$, one has

$$P[a + S_n = b \mid T_{-a} \leq n] = P[S_n = a + b] .$$

PROOF. The number of paths from a to b passing zero is equal to the number of paths from $-a$ to b which is the number of paths from zero to $a + b$.

DISTRIBUTION OF THE RUIN TIME. $P[T_{-a} \leq n] = P[S_n \leq -a] + P[S_n > a]$.
$P[T_{-a} = n] = \frac{a}{n} P[S_n = a]$.

PROOF. a) Use the reflection principle

$$\begin{aligned} P[T_{-a} \leq n] &= \sum_{b \in \mathbb{Z}} P[T_{-a} \leq n | a + S_n = b] \\ &= \sum_{b \leq 0} P[a + S_n = b] + \sum_{b > 0} P[S_n = a + b] \\ &= P[S_n \leq -a] + P[S_n > a] \end{aligned}$$

b) From $P[S_n = a] = \left(\begin{array}{c} n \\ \frac{a+n}{2} \end{array} \right)$ we get

$$\frac{a}{n} P[S_n = a] = \frac{1}{2}(P[S_{n-1} = a - 1] - P[S_{n-1} = a + 1]) .$$

$$\begin{aligned} P[S_n > a] - P[S_{n-1} > a] &= P[S_n > a \mid S_{n-1} \leq a] + P[S_n > a \mid S_{n-1} > a] - P[S_{n-1} > a] \\ &= \frac{1}{2}(P[S_{n-1} = a] - P[S_{n-1} = a + 1]) \end{aligned}$$

and analoguously

$$P[S_n \leq -a] - P[S_{n-1} \leq -a] = \frac{1}{2}(P[S_{n-1} = a - 1] - P[S_{n-1} = a]) .$$

Therefore, using a)

$$\begin{aligned} P[T_{-a} = n] &= P[T_{-a} \leq n] - P[T_{-a} \leq n - 1] \\ &= P[S_n \leq -a] - P[S_{n-1} \leq -a] + P[S_n > a] - P[S_{n-1} > a] \\ &= \frac{1}{2}(P[S_{n-1} = a] - P[S_{n-1} = a + 1]) + \frac{1}{2}(P[S_{n-1} = a - 1] - P[S_{n-1} = a]) \\ &= \frac{1}{2}(P[S_{n-1} = a - 1] - P[S_{n-1} = a + 1]) = \frac{a}{n} P[S_n = a] \end{aligned}$$

BALLOT THEOREM. $P[S_n = a \mid S_1 > 0, \ldots, S_{n-1} > 0] = \frac{a}{n} \cdot P[S_n = a]$.
PROOF. Reverse time: the number of paths from 0 to a of length n which do no more hit 0 is the number of paths of length n which start in a and for which $T_{-a} = n$.

DISTRIBUTION OF FIRST RETURN TIME.

$$P[T_0 > 2n] = P[S_{2n} = 0] .$$

PROOF.

$$\begin{aligned}
P[T_0 > 2n] &= \frac{1}{2}P[T_{-1} > 2n-1] + \frac{1}{2}P[T_1 > 2n-1] \\
&= P[T_{-1} > 2n-1] \qquad\qquad (\text{ by symmetry}) \\
&= P[S_{2n-1} > -1 \text{ and } S_{2n-1} \leq 1] \\
&= P[S_{2n-1} \in \{0,1\}] \\
&= P[S_{2n-1} = 1] = P[S_{2n} = 0]
\end{aligned}$$

REMARK. We see that $\lim_{n\to\infty} P[T_0 > 2n] = 0$ which restates that the random walk is recurrent. However, the expected return time is very long:

$$E[T_0] = \sum_{n=0}^{\infty} nP[T_0 = n] = \sum_{n=0}^{\infty} P[T_0 > n] = \sum_{n=0}^{\infty} P[S_n = 0] = \infty$$

since by Stirling's formula

$$P[S_{2n} = 0] \sim (\pi n)^{-1/2} .$$

DEFINITION. The stopping time

$$L = \max\{0 \leq n \leq 2N \mid S_n = 0\}$$

is the **last visit of the random walk in** 0 **before time** $2N$. Interpretation. If the random walk describes a game between two players, who play over a time $2N$, then L is the time when one of the two players does no more give up his or her leadership.

ARC-SIN LAW. a) L has the discrete arc-sin distribution:

$$P[L = 2n] = 2^{-2N} \binom{2n}{n} \binom{2N-2n}{N-n}$$

b) For $N \to \infty$, we have

$$P[\frac{L}{2N} \leq z] \to \frac{2}{\pi} \arcsin(\sqrt{z}) .$$

PROOF.

$$P[L = 2n] = P[S_{2n} = 0] \cdot P[T_0 > 2N - 2n] = P[S_{2n} = 0] \cdot P[S_{2N-2n} = 0]$$

which gives a).
Stirling gives $P[S_{2k} = 0] \sim \frac{1}{\sqrt{\pi k}}$ so that

$$P[L = 2k] = \frac{1}{\pi} \frac{1}{\sqrt{k(N-k)}} = \frac{1}{N} f(k/N)$$

with

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}} .$$

It follows that

$$P[L/2N \leq z] \to \int_0^z f(x)\, dx = \frac{2}{\pi} \arcsin(\sqrt{z}) .$$

INTERPRETATION. From the shape of the arc-sin distribution, One has to expect that the winner takes the final leading position either very early or very late.