Random Walk

Rabi Bhattacharya, The University of Arizona, USA

The simple random walk $\{S_n : n = 0, 1, ...\}$, starting at an integer x, is a stochastic process on the integers, given by $S_0 = x$, $S_n = x + X_1 + ... + X_n (n \geq 1)$, where $X_n, n \geq 1$, is an independent Bernoulli sequence: $P(X_n = 1) = p$, $P(X_n = -1) = 1 - p = q$, $0 < p < 1$. In the case, $p = q = 1/2$, it is called the *simple symmetric random walk*, while if $p \neq 1/2$, it is *asymmetric*. By the binomial theorem, $P(S_n = y \mid S_0 = 0) = C^n_{\frac{(n+y)}{2}} p^{(n+y)/2} q^{(n-y)/2}$, if y and n are of the same parity, i.e., if either both are odd or both are even. Otherwise, $P(S_n = y \mid S_0 = 0) = 0$. Here $= C^n_m = n!/(m!(n-m)!)$.

For $c \leq x \leq d$ integers, the probability $\pi(x)$ that a simple random walk, starting at x, reaches c before d satisfies the equation

$$\pi(x) = p\pi(x+1) + q\pi(x-1) \text{ for } c < x < d, \ \pi(c) = 1, \ \pi(d) = 0, \qquad (1)$$

as shown by conditioning on the first step X_1. For the symmetric walk, the solution of this equation is $\pi(x) = (d-x)/(d-c)$. Since $\pi(x) \to 1$ as $d \to \infty$, the symmetric walk will reach the state c, starting from any state $x > c$, with probability one. By symmetry, it will reach every state with probability one. Iterating this argument one sees that, with probability one, the symmetric random walk visits every state infinitely often. That is, the walk is *recurrent*. For the asymmetric walk, the solution to (1) is $\pi(x) = (1-(p/q)^{d-x})/(1-(p/q)^{d-c})$. If $p < 1/2$, then the limit of this is 1 as $d \to \infty$ and, with probability one, the random walk will visit c, starting from $x > c$. On the other hand, if $p > 1/2$, then $\pi(x) \to (q/p)^{x-c} < 1$, as $d \to \infty$. The probability of ever reaching d, starting from $x < d$ is obtained by symmetry as 1 if $p > 1/2$ and $(p/q)^{d-x}$ if $p < 1/2$. The asymmetric simple random walk is thus *transient*. Indeed, it follows from the strong law of large numbers $(SLLN)$ that if $p > 1/2$, then $S_n \to \infty$ with probability one as $n \to \infty$; and $S_n \to -\infty$, with probability one, if $p < 1/2$. For these and other properties of the random walk, such as those described below, see Feller (1968), Chapter 3, Bhattacharya and Waymire (2009), Chapter 1, or Durrett (1995), Chapter 3. For additional information, refer to Billingsley (1968), and Spitzer (1964).

For computation of various probabilities associated with a simple random walk, the following result proved by D. Andre in 1887 is very useful: Consider the polygonal path of the random walk joining successive points (j, S_j), $(j+1, S_{j+1})$ $(j = 0, 1, ..., n-1)$ by line segments. Let $y > 0$. Then (a) the set of paths from $(0,0)$ to $(n, y-1)$ (n and $y-1$ of the same parity) which touch or cross the level y, is in one-one correspondence with (b) the set of all paths from $(0,0)$ to $(n, y+1)$ (*Reflection principle*). To prove this, let τ be the first time a path of the type (a) touches the level y prior to time n. Then replace the segment of the path from (τ, y) to $(n, y-1)$ by its mirror reflection about

1

the level y. This gives a path of the type (b). Conversely, given any path of the type (b), reflect about y the segment of the path from (τ, y) to $(n, y + 1)$. This gives a path of the type (a). Here is an application of this principle.

Example 1 (First passage time distribution of a simple random walk). Let y be a positive integer, and $F_{n,y}$ the event that the random walk, starting at zero, reaches y for the first time at time n, i.e., $F_{n,y} = \{S_j \neq y, \text{for } 0 \leq j < n, S_n = y\}$, n and y of the same parity. Altogether there are $C^n_{\frac{n+y}{2}}$ paths from $(0, 0)$ to (n, y), each having probability $p^{(n+y)/2}q^{(n-y)/2}$. Of these, the number which cross or touch the level y prior to time n and for which $S_{n-1} = y - 1$ is, by the reflection principle, $C^{n-1}_{\frac{n+y}{2}}$. Also the number for which $S_{n-1} = y + 1$ is $C^{n-1}_{\frac{n+y}{2}}$. Subtracting these two from the number $C^n_{\frac{n+y}{2}}$ of all paths, one obtains, for all $y \neq 0$ (treating the case $y < 0$ by symmetry),

$$P(F_{n,y}) = (C^n_{\frac{n+y}{2}} - 2C^{n-1}_{\frac{n+y}{2}})p^{(n+y)/2}q^{(n-y)/2} = (|y|/n)C^n_{\frac{n+y}{2}}p^{(n+y)/2}q^{(n-y)/2} \quad (2)$$

$(n = |y|, |y| + 2, |y| + 4, ...)$.

One may also consider the simple symmetric random walk $S_0 = x$, $S_n = x + X_1 + ... + X_n$ $(n \geq 1)$, in dimension $d \geq 1$, as a stochastic process on the d-dimensional lattice Z^d, with $X_n (n \geq 1)$ i.i.d. random vectors, taking values $\pm e_j (j = 1, ..., d)$, each with probability $1/2d$. Here e_j is the vector whose j-th coordinate is 1 and the remaining $d - 1$ coordinates are zero. It was proved by G. Polya in 1921 that this walk is recurrent in dimensions 1,2, and transient in higher dimensions.

De Moivre (1756) obtained the normal approximation to the binomial probability $P(S_n = y \mid S_0 = 0)$, as a combinatorial result. The full potential of this was realized by Laplace (1812) who formulated and derived the far reaching central limit theorem (CLT). Apparently, Gauss knew about the normal distribution as early as 1794, and assuming this as the distribution of errors of measurement, he obtained his famous method of least squares. Hence the name Gaussian distribution is often used for the normal distribution. The final version of the CLT for a *general random walk* $S_n = X_1 + ... + X_n$ $(n \geq 1)$, where X_n are arbitrary independent identically distributed (*i.i.d.*) random variables with mean zero and finite variance $\sigma^2 > 0$, was obtained by Le'vy (1925): $n^{-1/2}(X_1 + ... + X_n)$ converges in distribution to the normal distribution $N(0, \sigma^2)$ with mean zero and variance σ^2, as $n \to \infty$. In physical terms, this result says the following: if time and length are rescaled so that in one unit of rescaled time there are a large number n of i.i.d. displacements of small rescaled lengths of order $1/\sqrt{n}$, then the random walk displacements over a period of time t will appear as Gaussian with mean zero and variance $t\sigma^2$, the increments over disjoint intervals being independent. That such a Gaussian process exists with continuous sample paths was proved rigorously by N.Wiener in 1923. This process is

called the *Brownian motion*, following its implicit use by A. Einstein in 1905-6 to describe the kinetic motion of colloidal molecules in a liquid, experimentally observed earlier by the botanist R. Brown. Interestingly, even before Einstein, Bachelier (1900) described the random movements of stocks by this Gaussian process. The statement that the rescaled random walk $S_n (n = 0, 1, 2, ...)$ converges in distribution to Brownian motion was proved rigorously by M. Donsker in 1951, and this result is known as the functional central limit theorem ($FCLT$). Both the CLT and the $FCLT$ extend to arbitrary dimensions d.

As consequences of the $FCLT$, one can derive many asymptotic results for the simple symmetric random walk given by the corresponding result for the limiting Brownian motion. Conversely, by evaluating combinatorially some probability associated with the random walk, one may derive the corresponding probability for the Brownian motion. A Brownian motion with variance parameter $\sigma^2 = 1$ is called a standard Brownian motion, and denoted $\{B_t : t \geq 0\}$ below.

Example 2 (Boundary hitting probability of Brownian motion). Let $c \leq x \leq d$ be arbitrary reals. Then, using the corresponding result for the scaled random walk, one obtains

$$P(\{B_t : t \geq 0\} \text{ reaches c before } d \mid B_0 = x) = (d - x)/(d - c). \qquad (3)$$

Example 3 (Arcsine law). Let U denote the amount of time in $[0, 1]$ the Brownian motion spends above zero, i.e., U = Lebesgue measure of the set $\{t : 0 \leq t \leq 1 : B_t > 0\}$, given $B_0 = 0$. Consider the polygonal path of the simple symmetric random walk $S_j (j = 0, 1, ...n)$, starting at zero. By combinatorial arguments, such as the reflection principle, one can calculate exactly the proportion of times the polygonal path lies above zero and, by the $FCLT$, this yields

$$P(U \leq x) = (2/\pi) sin^{-1} \sqrt{x} \ (0 \leq x \leq 1). \qquad (4)$$

REFERENCES

Bachelier, L. (1900). The'orie de la speculation. *Ann. Sci. Ecole Norm. Sup.* **17**, 21-86.

Bhattacharya, R.N. and Waymire, E.C. (2009).*Stochastic Processes with Applications.* SIAM Classics in Applied Mathematics, vol. 61. SIAM, Philadelphia.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

De Moivre, A. (1756). *Doctrine of Chance.* London.

Durrett, R. (1995). *Probability: Theory and Examples.* 2nd edition. Duxbury, Belmont.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications.* Vol.1. 3rd edition. Wiley, New York.

Laplace, P.S. (1812). *The The'orie Analitique des Probabilite's.* Paris.

Le'vy, P. (1925). *Calcul des Probabilite's.* Gauthier-Villars. Paris.

Spitzer, F. (1964). *Principles of Random Walk.* Van Nostrand. Princeton, NJ.