Random Walks, Hitting Probabilities, and the SPRT

Recurrence of Centered Random Walks

Let $X_n = x+\sum_{i=1}^n Z_i$ be a $\mathbb{R}^d$-valued random walk with *i.i.d.* steps Z_i, with mean $\mathsf{E}\, Z_i = 0$ and covariance $\mathsf{E}\, Z_i Z_j = Q_{ij}$. Then X_n is itself an $\mathbb{R}^d$-valued martingale, and $b \cdot X_n$ is a real-valued martingale for each d-vector b, but what about $|X_n|$? To study the recurrence of X_n it's useful to locate a number α for which $|X_n|^\alpha$ is a martingale. Fix any α and set $f(x) = |x|^\alpha = \Big(\sum_i (x_i)^2\Big)^{\alpha/2}$.

A second-order Taylor expansion yields

$$\begin{aligned}
f(x+z) &\approx f(x) + \sum_{i=1}^d z_i \frac{\partial f(x)}{\partial x_i} + {}^1\!/\!_2 \sum_{i=1}^d \sum_{j=1}^d z_i z_j \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \\
&= |x|^\alpha + \sum_{i=1}^d z_i\, \alpha |x|^{\alpha-2} x_i + {}^1\!/\!_2 \sum_{i,j=1}^{d,d} z_i z_j \Big[\alpha(\alpha-2)|x|^{\alpha-4} x_i x_j + \alpha |x|^{\alpha-2}\delta_{ij}\Big] \\
&= f(x) + \alpha|x|^{\alpha-2} \sum_{i=1}^d z_i x_i + \frac{\alpha|x|^{\alpha-2}}{2}\Big[(\alpha-2)\sum_{i,j=1}^{d,d} \frac{z_i z_j x_i x_j}{|x|^2} + \sum_{i=1}^d (z_i)^2\Big]
\end{aligned}$$

and, to second-order,

$$\begin{aligned}
&\mathsf{E}[f(X_{n+1}) - f(X_n) | \mathcal{F}_n] \\
&\qquad \approx \alpha|X_n|^{\alpha-2}\, \mathsf{E}\, Z_{n+1}\cdot X_n + \frac{\alpha|x|^{\alpha-2}}{2}\Big[(\alpha-2)\sum_{i,j=1}^{d,d} \mathsf{E}\frac{Z_i Z_j (X_n)_i (X_n)_j}{|X_n|^2} + \sum_{i=1}^d (Z_i)^2\Big] \\
&\qquad = \frac{\alpha|x|^{\alpha-2}}{2}\Big[(\alpha-2)\sum_{i,j=1}^{d,d} \frac{Q_{ij}(X_n)_i (X_n)_j}{|X_n|^2} + \sum_{i=1}^d Q_{ii}\Big]
\end{aligned}$$

For the symmetric random walk, $Q_{ij} = \delta_{ij}/d$ and hence

$$\begin{aligned}
&\mathsf{E}[f(X_{n+1}) - f(X_n) | \mathcal{F}_n] \\
&\qquad = \frac{\alpha|X_n|^{\alpha-2}}{2}\Big[\sum_{i=1}^d \frac{(\alpha-2)(X_n)_i^2}{d\,|X_n|^2} + \sum_{i=1}^d \frac{1}{d}\Big] \\
&\qquad = \frac{\alpha|X_n|^{\alpha-2}}{2}\Big[\frac{\alpha-2}{d} + 1\Big],
\end{aligned}$$

or zero if $\alpha = 2-d$; thus for any centered d-dimensional random walk, $|X_n|^{2-d}$ is (approximately) a martingale. In two dimensions this isn't an interesting process (it's constant), but a similar calculation shows that $\log|X_n|$ is a martingale when $d=2$.

For any $0 < r < R < \infty$, starting X_n at some at $x \in \mathbb{R}^d$ satisfying $r < |x| < R$. By the martingale property, the probability $p_r^R(x)$ of hitting $|X_n| = r$ before $|X_n| = R$ must satisfy

$$\begin{aligned}
|X_0|^{2-d} = |x|^{2-d} &= \mathsf{E}\big[|X_{T_{r,R}}|^{2-d} \mid X_0 = x\big] \\
&= r^{2-d} p_r^R(x) + R^{2-d}\Big[1 - p_r^R(x)\Big], \text{ so} \\
p_r^R(x) &= \begin{cases} \frac{R^{2-d}-|x|^{2-d}}{R^{2-d}-r^{2-d}} & \text{if } d \neq 2 \\ \frac{\log(R)-\log(|x|)}{\log(R)-\log(r)} & \text{if } d = 2 \end{cases}
\end{aligned}$$

In the limit as $R \to \infty$ this converges to 1 for $d \le 2$ and to $|\frac{x}{r}|^{2-d}$ for $d > 2$, showing that the random walk is recurrent only in dimensions $d \le 2$ and moreover yielding the probability of eventual return as approximately $|\frac{x}{r}|^{2-d}$ for the transient walks.

Sequential Probability Ratio Test (SPRT)

Let Y_i be independent, identically-distributed random vectors whose probability distribution $f(y|\theta)$ is partially unknown: θ is known to be *either* $\theta = \theta_0$ *or* $\theta = \theta_1$. To simplify notation we'll denote $f(y|\theta_j)$ by $f_j(y)$, and by H_j the *hypothesis* that $\theta = \theta_j$. Both Bayesian and Frequentist statisticians recommend choosing between H_0 and H_1 by using variations on the Sequential Probability Ratio Test, in which we calculate the Likelihood Ratio

$$L_n = \frac{\prod_{i=1}^n f_1(Y_i)}{\prod_{i=1}^n f_0(Y_i)}$$

and regard high values of L_n as evidence against H_0, low values as evidence in favor of H_0. A common sequential test for the hypothesis H_0 proceeds by choosing some numbers $A < 1 < B$ and taking data until either $L_n \geq B$, whereupon we quit and reject H_0, or $L_n \leq A$, whereupon we quit and fail to reject H_0. I want to think of L_n and its logarithm, $\lambda_n = \log(L_n) = \sum_{i=1}^n \log\big(f_1(Y_i)/f_0(Y_i)\big)$, as stochastic processes and study L_n's hitting probabilities of $A < 1$ and $B > 1$ or, equivalently, $\lambda_n = \log(L_n)$'s hitting probabilities of $-a = \log A$ and $b = \log B$, respectively.

Likelihood ratios as Positive Martingales

Under H_0, L_n is a positive martingale starting at $L_0 = 1$:

$$\begin{aligned}
\mathsf{E}[L_{n+1}|\mathcal{F}_n, \theta = \theta_0] &= L_n \int \frac{f_1(y)}{f_0(y)} f_0(y)\,dy \quad = L_n \int f_1(y)\,dy = L_n, \text{ so by Doob's theorem,} \\
1 = \mathsf{E}[L_0] &= A \times \mathsf{P}[L_n \leq A \text{ before } L_n \geq B | \theta = \theta_0] \\
&\quad + B \times \mathsf{P}[L_n \geq B \text{ before } L_n \leq A | \theta = \theta_0]
\end{aligned}$$

and $\mathsf{P}[L_n \geq B \text{ before } L_n \leq A | \theta = \theta_0] = \frac{1-A}{B-A}$. As $B \to \infty$, $\mathsf{P}[L_n \leq A \text{ eventually} \mid H_0] \to 1$ while even if $A \to 0$, $\mathsf{P}[L_n \geq B \text{ eventually} \mid H_0] \leq 1/B$. In the symmetric case $A = 1/B$, the (pre-experimental) probabilities of "Type-I"and "Type-II" errors are both about $\alpha = \beta = \frac{1-1/B}{B-1/B} = \frac{1}{1+B}$; by choosing large B (and $A = 1/B$ small) a test of any desired *size* and *power* can be had. For the Bayesian L_n is the *Bayes factor* after any (fixed or random) number n of trials, so the posterior probability of hypothesis H_0 is

$$\mathsf{P}[\theta = \theta_0 | Y_1, \ldots, Y_n] = \frac{\pi_0}{\pi_0 + L_n \pi_1}$$

with prior probabilities $\pi_i = \mathsf{P}[\theta = \theta_i]$, or $1/(1 + L_n)$ in case $\pi_0 = \pi_1 = 1/2$. Upon stopping the SPRT this would be $\frac{1}{1+A}$ or $\frac{1}{1+B}$, depending upon which stopping point we reached. The Bayesian who *does* hit $L_n \geq B$ would report a posterior probability for H_0 of $\frac{1}{1+B}$, exactly the same as the p-value for a Frequentist who used $A = 1/B$, but the Frequentist would report other (pre-experimental!) p-values for other values of A even if B is hit first.

Log likelihood ratios as Random Walks

Under both H_0 and H_1, the process λ_n is a random walk starting at zero with independent steps $Z_i = \log\big(f_1(Y_i)/f_0(Y_i)\big)$. The expectation and variance of Z_i will depend on θ, of course: under H_1, the expectation of Z_i is

$$\begin{aligned}
\mu &= \mathsf{E}[\log\left(\frac{f_1(Y)}{f_0(Y)}\right) \mid \theta = \theta_1] \\
&= \int \log\left(\frac{f_1(y)}{f_0(y)}\right) f_1(y)\,dy \\
&= K_L(f_0, f_1) > 0,
\end{aligned}$$

the (positive) Kullback-Liebler "distance" from f_0 to f_1; similarly $\mathsf{E}[Z_i \mid \theta = \theta_0] = -\mu$, so in each case we have a random walk *with drift* whose hitting probabilities force the SPRT to give the correct answer with high probability if a and b are large.