Random Walk Algorithms: Lecture 4

David A. Meyer

In which the notion of independence is revisited and extended to multiple events, an instance of which—a sequence of coin flips—is subsequently analyzed.

Independence

So far we have discussed "algorithm" and "random" (or "probabilistic"). Our goal in the next two lectures is to define "random walk". First we need to extend our previous definition of independent events:

DEFINITION. A set of events $\mathcal{A} = \{A_i \subseteq \Omega \mid i \in \{1, \ldots, n\}\}$ is pairwise independent if for all $1 \leq i < j \leq n$, $\Pr(A_i \cap A_j) = \Pr(A_i) \Pr(A_j)$. \mathcal{A} is mutually independent if for any subset of events $\{A_{i_1}, \ldots, A_{i_k}\} \subseteq \mathcal{A}$,

$$\Pr(A_{i_1} \cap \cdots \cap A_{i_k}) = \Pr(A_{i_1}) \cdots \Pr(A_{i_k}).$$

Mutually independent implies pairwise independent, but the converse is generally false. These definitions carry over to sets of random variables, just as did independence of pairs of events to pairs of random variables.

Binomial distributions

Suppose we have a sequence of mutually independent coin flips at times $t \in \{1, ..., n\}$, with each having Pr(head) = p. The sample space is $\Omega = \{f_1 \dots f_n \mid f_t \in \{H, T\}\}$. Define n random variables

$$H_t(\omega) = \begin{cases} 1 & \text{if } f_t = \mathbf{H}; \\ 0 & \text{if } f_t = \mathbf{T}. \end{cases}$$

Since the coin flips are mutually independent, so is the set $\{H_t \mid t \in \{1, \ldots, n\}\}$, and each random variable has $\Pr(H_t = 1) = p$. Let

$$K_n = \sum_{t=1}^n H_t.$$

Then

$$\Pr(K_n = k) = \Pr(k \text{ heads out of } n \text{ flips}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

This is a *binomial* probability distribution function.

We can compute the mean and variance of a binomial probability distribution function using the properties that we deduced in the last lecture. For the expectation value we have:

$$\mathsf{E}[K_n] = \mathsf{E}[H_1 + \dots + H_n] = \mathsf{E}[H_1] + \dots + \mathsf{E}[H_n] = np.$$

And for the variance, since the H_t are mutually independent,

$$\mathsf{Var}[K_n] = \mathsf{Var}[H_1 + \dots + H_n] = \mathsf{Var}[H_1] + \dots + \mathsf{Var}[H_n] = np(1-p).$$

To define a random walk we will need a new set of random variables defined on the same sample space Ω . Let the t^{th} "step" be

$$S_t = \begin{cases} +1 & \text{if } f_t = \mathbf{H}; \\ -1 & \text{if } f_t = \mathbf{T}. \end{cases}$$

Notice that $S_t = 2H_t - 1$. Also, let

$$X_n = \sum_{t=1}^n S_t = 2K_n - n.$$

Since X_n is a function of K_n , it is easy to compute its probability distribution function:

$$\Pr(X_n = x) = \Pr(2K_n - n = x) = \Pr(K_n = \frac{n+x}{2}) = \binom{n}{(n+x)/2} p^{(n+x)/2} (1-p)^{(n-x)/2},$$

for n + x even; $Pr(X_n = x) = 0$ for n + x odd. In case this is not clear, we will discuss it further in Lecture 6.

The mean and variance of X_n are also easy to compute by writing it in terms of K_n :

$$E[X_n] = E[2K_n - n] = 2np - n = n(2p - 1)$$

Var[X_n] = Var[2K_n - n] = 4Var[K_n] = 4np(1 - p),

using the properties of mean and variance we derived in Lecture 3.