6.262: Discrete Stochastic Processes
Lecture - Stopping, Wald, and martingales, 5/4/09

- Stopping nodes and stopping rules
- Wald's identity
- Stopped martingales

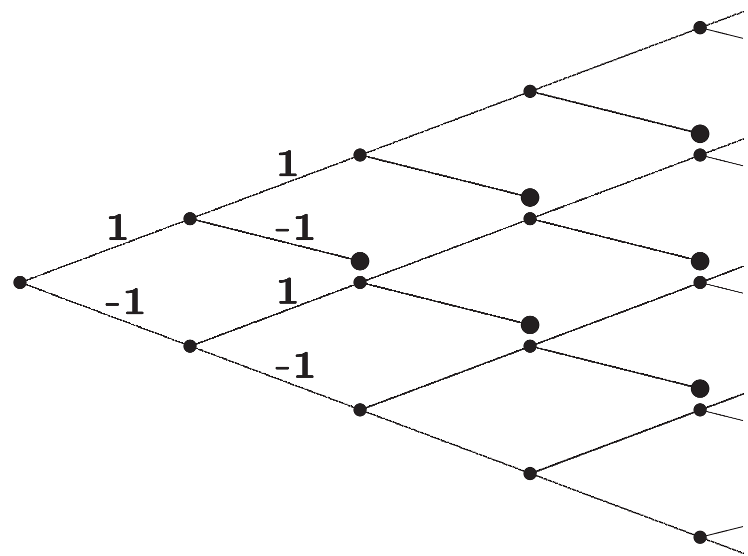
Stopping rules

The concept of stopping is used to capture some event (threshold crossings, a specified string, a combination of events) in a stochastic process.

Stopping is associated with a time at which the event has happened.

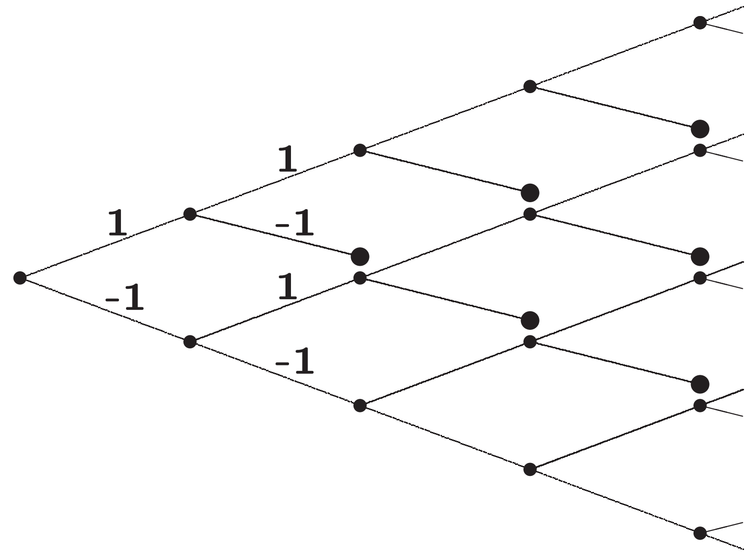
Stopping occurs at time n in a discrete time process $\{X_i; i \geq 1\}$ if the sample sequence $x_1, \dots, x_n$ identifies the given event as having happened and $x_1, \dots, x_{n-1}$ does not identify it.

Example: A tree representing binary sequences: Stop on first occurrence of $(+1, -1)$.



Big dots represent stopping points.

The sample points of any discrete stochastic sequence $\{X_i; i \geq 1\}$ can be represented by a tree. Each node represents an initial segment of a set of sample sequences.



Def: A collection of stopping nodes is a collection of initial segments of sample sequences such that no stopping node is an initial segment of another stopping node.

A stopping rule is a rule specifying a collection of stopping nodes. A stopping time is the length of the initial segment.

How is new rule different from old?

- Contains no probabilities - defined by events.
- Specifies initial segments causing stopping.
- Allows visualization by a tree.
- Permits simple proofs of theorems.
- Allows for defective stopping time rv's in simple way.

Example 1: Random walk with thresholds at $\alpha > 0$, $\beta < 0$.

A stopping node $(n, \mathbf{x})$ is an initial segment such that $s_i = x_1 + \cdots + x_i \in (\alpha, \beta)$ for $1 \leq i < n$ and $s_n \notin (\alpha, \beta)$.

Stopping time is (non-defective) rv.

Example 2: Random walk with threshold at $\alpha > 0$.

A stopping node $(n, \mathbf{x})$ is an initial segment such that $s_i = x_1 + \cdots + x_i < \alpha$ for $1 \leq i < n$ and $s_n \geq \alpha$.

Stopping time is defective unless $\overline{X} > 0$.

Thm: [Wald's identity] Let $\{X_i;\, i \geq 1\}$ **be IID with finite** $\gamma(r) = \ln\{\mathsf{E}\left[e^{rX}\right]\}$ **in** $r_- < 0 < r_+$**. Let** $\alpha > 0$ **and** $\beta < 0$ **be arbitrary real numbers, and let** N **be the smallest** n **for which either** $S_n \geq \alpha$ **or** $S_n \leq \beta$**. Then for all** $r \in (r_-, r_+)$**,**

$$\mathsf{E}\left[\exp(rS_N - N\gamma(r))\right] = 1.$$

Pf: From Lemma 7.1, N **is a rv. The set of stopping nodes** $\mathcal{T}$ **is the set of** $(n, \mathbf{x}) = (n, x_1, \dots, x_n)$ **such that** $s_i \in (\alpha, \beta), i < n$ **and** $s_n \notin (\alpha, \beta)$**.**

Let $p(n, \mathbf{x}) = p(x_1)p(x_2)\cdots p(x_n)$**. Then**

$$\sum_{(n,\mathbf{x})\in\mathcal{T}} p(n, \mathbf{x}) = 1$$

Reason: the set of $(n, \mathbf{x}) \in \mathcal{T}$ **partitions** $\{X_i;\, i \geq 1\}$ **except for set of probability 0.**

Consider another IID probability measure on $\{X_i;\, i \geq 1\}$ **with** $q(x) = p(x) \exp[rx - \gamma(r)]$**, where** $r \in (r_-, r_+)$**.**

The same events cause stopping, but with different probabilities.

Let $q(n, \mathbf{x}) = q(x_1)p(x_2) \cdots q(x_n)$**. Then**

$$\sum_{(n,\mathbf{x}) \in \mathcal{T}} q(n, \mathbf{x}) = 1$$

Reason: the set of $(n, \mathbf{x}) \in \mathcal{T}$ **partitions** $\{X_i;\, i \geq 1\}$ **except for set of probability 0 (Lemma 7.1).**

$$q(n, \mathbf{x}) = p(n, \mathbf{x}) \exp[rs(n, \mathbf{x}) - n\gamma(r)]$$

$$1 = \sum_{(n,\mathbf{x}) \in \mathcal{T}} p(n, \mathbf{x}) \exp[rs(n, \mathbf{x}) - n\gamma(r)]$$

To prove: $\quad \mathsf{E}\,[\exp(rS_N - N\gamma(r))] = 1.$

Note that N and S_N for a sequence $x_1, x_2, \ldots$ are specified by the stopping node $(n, \mathbf{x})$ for that sequence. Thus writing out this expectation as a sum over the stopping nodes,

$$1 = \sum_{(n,\mathbf{x}) \in \mathcal{T}} p(n, \mathbf{x}) \exp[rs(n, \mathbf{x}) - n\gamma(r)],$$

which is what we have already established.

As the proof shows, Wald's identity is much more general than what we have stated.

It applies only to random walks (since that is what allows us to use the 'tilted' probability assignment on segments.

The stopping rule, however, can be anything so long as the stoppiing time is non-defective for both the original probability and for the tilted probability.

The thresholds can be time varying, for example.

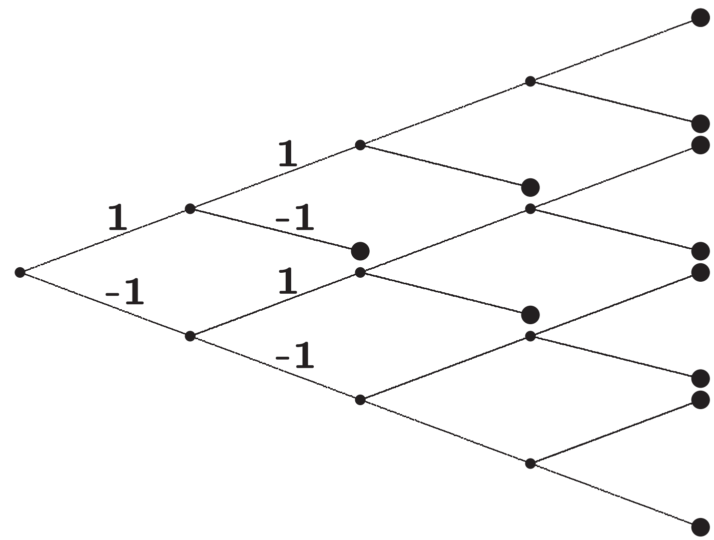
Strangely enough, Wald's equality is slightly harder to prove, using this definition of stopping rule, than Wald's identity.

At the same time, it is easier than what we did before, and far more insightful.

Wald's equality says $\mathsf{E}[S_N] = \mathsf{E}[X]\,\mathsf{E}[N]$ for a random walk $\{X_n;\, n \geq 1\}$ under the restrictions that $\mathsf{E}[N] < \infty$ and $\mathsf{E}[|X|] < \infty$.

Thus, stopping occurs with probability 1. Since N and S_N are specified by the stopping node, we want to prove that

$$\sum_{(n,\mathbf{x})\in\mathcal{T}} p(n,\mathbf{x})s_n = \sum_{(n,\mathbf{x})\in\mathcal{T}} p(n,\mathbf{x})n\mathsf{E}[X]] = 0$$



A slick way to calculate $\sum_{(n,\mathbf{x})\in\mathcal{T}} p(n,\mathbf{x})s_n$ is to observe that it is the same as summing over all intermediate nodes, say $\mathcal{P}_{int}$ (including the origin) getting

$$\sum_{(n,\mathbf{x})\in\mathcal{T}} p(n,\mathbf{x})s_n = \sum_{(i,\mathbf{x})\in\mathcal{T}_{int}} p(i,\mathbf{x})\mathbf{E}\left[X\right]$$

The other term can be manipulated into the same form.

Stopping rules for martingales

The definition of stopping nodes applies to arbitrary integer time processes $\{Z_n; n \geq 1\}$ as well as to IID sequences.

There is a new twist that is valuable for martingales - instead of being primarily interested in the stopping time N and value Z_n, we create a stopped process that continues in time but stays fixed in Z_n

For a process $\{Z_n; n \geq 1\}$ and a given stopping rule, the stopped process $\{Z_n^*; n \geq 1\}$ is the process satisfying $Z_n^* = Z_n$ for $n \leq N$ and $Z_n^* = Z_N$ for $n > N$.

As one might guess, freezing a martingale (or submartingale or supermartingale) in its tracks does not change the martingale property.

This still has to be proven (because martingales are quite general and often strange animals), but the proofs are straightforward.

Thus the stopped process of a martingale (subm., superm.) is a martingale (subm, superm. respectively).

Recall that a submartingale is defined by

$$\mathbf{E}[|Z_n|] < \infty; \quad \mathbf{E}[Z_n \mid Z_{n-1}, Z_{n-2}, \ldots, Z_1] \geq Z_{n-1}$$

We have seen that it also satisfies

$$\mathbf{E}[Z_n \mid Z_i, Z_{i-1}, \ldots, Z_1] \geq Z_i \quad \text{for } i < n$$

$$\mathbf{E}[Z_n \mid Z_i, Z_{i-1}, \ldots, Z_1] \geq \mathbf{E}[Z_i] \quad \text{for } i < n$$

$$\mathbf{E}[Z_n] \geq \mathbf{E}[Z_i] \quad \text{for } i < n$$

For martingales, the inequalities here become equalities, and for supermartingales, the inequalities are reversed.

When a submartingale is stopped, the expectations of the stopped process lie inside the original process, i.e.,

$$\mathbf{E}[Z_1] \le \mathbf{E}[Z_n^*] \le \mathbf{E}[Z_n]$$

For martingales, the inequalities become equalities, and for supermartingales, the inequalities are reversed.

Pathological cases

Consider the binary product martingale. $Z_n = X_1 X_2 \cdots X_n$ **where** X_n **is IID, equiprobable** $\{0, 2\}$**.**

Stopping at $Z_n = 0$**, the stopped process is the same as the original, so** $\lim_{n\to\infty} Z_n^* = 0$ **W.P.1.**

However $\lim_{n\to\infty} \mathsf{E}[Z_n^*] = 1$**.**

The problem is $Z_n^* = Z_n = 2^n$ **with probability** 2^{-n}**.**