Conditional expectation

Expectation of a r.v. $X$ conditioned on an event $B$

$$
\mathbb{E}[X|B] = \frac{\sum_{w \in B} X(w) p(w)}{P(B)}
$$

Our setup: consider a random walk $X_n$ on $\mathbb{Z}$

Equivalently

Suppose $S$ is a r.v. that depends on $X_1, \ldots, X_N$ (i.e., $S$ depends on the first $N$ coin tosses)

For instance, $S = 2^{X_N}$ (stock price)

Want: obtain the best possible estimate of $S$ after only $m \leq N$ tosses

Def: Let $1 \leq m \leq N$, at time $m$ we know the realized coin losses

$w_1, \ldots, w_m \in \{H, T\}$

(no longer r.v.)

There are $2^{N-m}$ possible realizations of $w_{m+1}, \ldots, w_N$

(we have to toss coin $N-m-1$ times to find out)

Let $#H(w_{m+1}, \ldots, w_N)$ be the number of heads in $w_{m+1}, \ldots, w_N$

$#T(w_{m+1}, \ldots, w_N)$ tails

Define

$$
\mathbb{E}_m[S](w_1, \ldots, w_m) = \sum P_{(w_{m+1}, \ldots, w_N)} (1-p)
$$
We call $E_m(S)$ the conditional expectation of $S$ based on the information at time $m$.

Based on what we know at time 0, $E_0(S)$ is a random variable, i.e. its value depends on the first $m$ tosses.

**Extreme cases:**

$E_0(S) = E(S)$

$E_N(S) = S$

**Properties:** Let $Y_1, \ldots, Y_N$ be random variables that depend on the first $N$ tosses.

We have:

1. **Linearity:** $E_n(cY_1 + c_2 Y_2) = c_1 E_n(Y_1) + c_2 E_n(Y_2)$
   
   $c_1, c_2 \in \mathbb{R}$

2. **Tacking out what is known:** if $Y$ actually depends on the first $m < N$ tosses, then

   $E_m(Y) = Y E_m[2]$  

3. **Iterated conditioning:** if $0 \leq n \leq m \leq N$

   $E_n(E_m[2]) = E_n(2)$

   $\Rightarrow$ allows to simplify various expressions

   In particular, $E[E_m(2)] = E[2]$

4. **Independence:** if $Y$ depends only on the tosses $m+1, \ldots, N$ then

   $E_n(Y) = E(Y)$

**Ex:**

```
X: 0 2/3 2 1/2 1/2 2/3 2 3 3/2 1/2 ...

"we" n:
0 3/5 2/3 2/5 1/3 1/5 2/3 2/5 3/2 3/5 ...

"we" T:
0 1/3 1/3 1/3 1/3 1/3 1/3 1/3 1/3 1/3 ...

n=0 n=1 n=2 n=3...
```
Random variable: \( S = 2^X \) (i.e. \( N = 2 \))

Let's find \( \mathbb{E}_t[S] \)

At time \( n = 0 \) \( \mathbb{E}_t[S] = 2^0 = 1 \) r.v.

\[ \mathbb{E}_t[S](T) = \frac{2}{3} \cdot 2^0 + \frac{1}{3} \cdot 2^{-2} = \frac{2}{3} + \frac{1}{12} = \frac{3}{4} \]

\[ \mathbb{E}_t[S](M) = \frac{2}{3} \cdot 2^0 = \frac{2}{3} \]

\( \Rightarrow \mathbb{E}_t[S] \) is a r.v. taking value 3 with prob. \( \frac{2}{3} \)

Best estimate of \( S \) given the information at \( n = 1 \)

\( \frac{3}{4} \) with prob. \( \frac{1}{3} \)

**Martingales**

**Def.** Let \( M_0, M_1, M_2, \ldots \) be r.v.'s, with each \( M_n \) depending only on the first \( n \) tosses.

Then \( M_0, M_1, M_2, \ldots \) is called adapted stochastic process.

**Remark.** This definition means that \( M_n \) do not depend on future events.

**Def.** An adapted process is called a martingale if

\( M_n = \mathbb{E}_n[M_{n+1}] \) for all \( n \).

Martingale has no tendency to rise or fall.

(Stock price is a martingale \( \iff \) no tendency to rise or fall)

**Homework:**

Let \( X_0, X_1, \ldots \) be an adapted process.

Show: if for any function \( f \) there is a function \( g \) such that

\[ \mathbb{E}_n[f(X_{n+1})] = g(X_n) \]

then \( X_n \) is a Markov chain ("no memory").