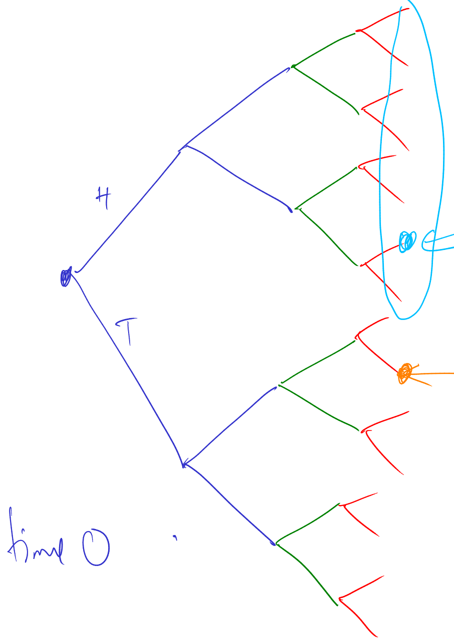
Last time:

$\Omega \to$ Prob space N die rolls.

$\omega = (\omega_1, \dots \omega_N)$

Filtration $\mathcal{F}_n$ = all events that can be determined using only the first $\boxed{n}$ coin tosses

X is $\mathcal{F}_n$ - meas if $X(\omega)$ only depends on $\omega_1, \omega_2 \dots \omega_n$ (and not $\omega_{n+1} \dots \omega_N$).



$\in A$

say $\omega \in A$
$\omega = (H, T, T, H)$

(T, H, H, T)

$A \in \mathcal{F}_1 \to A \subseteq \Omega$

A only depends on the 1^{st} coin toss.

$A = \cancel{\{(H, T, T, H)\}}$

$\Omega = \{$end pts of all these trees$\}$

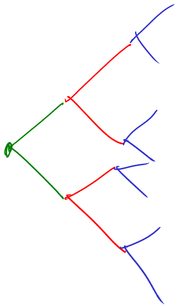
5.3. Conditional expectation.

Definition 5.20. Let X be a random variable, and $n \leqslant N$. We define $\boldsymbol{E}(X \mid \mathcal{F}_n) = \boldsymbol{E}_n X$ to be the *random variable* given by

$$\boxed{\boldsymbol{E}_n X(\omega)} = \frac{\displaystyle\sum_{\omega' \in \Pi_n(\omega)} p(\omega') X(\omega')}{\displaystyle\sum_{\omega' \Pi_n(\omega)} p(\omega')}, \qquad \text{where} \qquad \boxed{\Pi_n(\omega) = \{\omega' \in \Omega \mid \omega'_1 = \omega_1, \; \dots, \omega'_n = \omega_n\}}$$

Remark 5.21. $\boldsymbol{E}_n X$ is the "best approximation" of X given only the first n coin tosses.

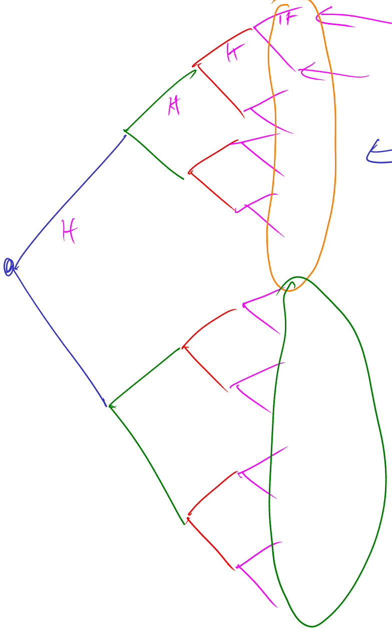
Remark 5.22. The above formula does not generalize well to infinite probability spaces. We will develop a definition that does generalize; after we have that definition we will never ever ever use this formula.



$\omega \in \Omega$. $\Pi_1(\omega) = \{ \text{all } \omega' \in \Omega \mid \omega' = (\omega'_1, \omega'_2, \omega'_3) \;\&\; \omega'_1 = \omega_1 \}$

$\downarrow$

$(\omega_1, \omega_2, \omega_3)$



$X=1$

$X=2$

$\Pi_1((H,\text{—}))$ $\quad X(\omega)$ (fair coin)

$E_1 X$ = cond exp of X given $\mathcal{F}_1$

= best approx of X given only the first coin toss

① If 1st coin is Head

$$E_1 X(\omega) = \left\{ \frac{\displaystyle\sum_{\omega \in \text{top orange block}} X(\omega)\cdot p(\omega)}{\displaystyle\sum_{\omega \in \text{top orange block}} p(\omega)} \right\}$$

② If 1st coin is tails

$X=16$

$E_1 X(\omega) =$ same form

but sum over the bottom green block instead.

Remark 5.23. The conditional expectation $\boldsymbol{E}_n X$ defined by the above formula satisfies the following two properties:

(1) $\boldsymbol{E}_n X$ is an $\mathcal{F}_n$-measurable *random variable.* ⇐ $E_n X(\omega)$ only depends on $\omega_1, \omega_2 \cdots \omega_n$ (& not $\omega_{n+1}, \cdots \omega_N$).

(2) For every $A \in \mathcal{F}_n$, $\displaystyle\sum_{\omega \in A} \boldsymbol{E}_n X(\omega) p(\omega) = \sum_{\omega \in A} X(\omega) p(\omega).$

$A \in \mathcal{F}_n$ → A some event that can be described in terms of the first n coin tosses alone.

Q: Avg of X on the event A = $\displaystyle\sum_{\omega \in \cancel{\Omega}\, A} X(\omega) p(\omega)$ / $\displaystyle\sum_{\omega \in A} p(\omega)$

Q: Avg of $E_n X$ on the event A

$= \displaystyle\sum_{\omega \in A} E_n X(\omega) p(\omega)$ / $\displaystyle\sum_{\omega \in A} p(\omega)$

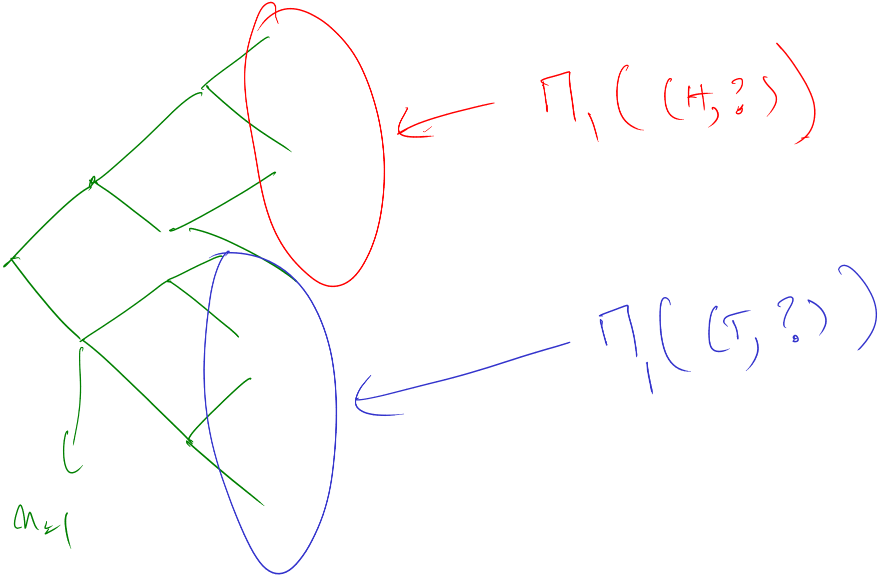
(∵ Venute $A \in \mathcal{F}_n$)

1 Check $E_n X$ (by my formula) satisfies 1 & 2 above.

2 Use this as my def of <u>cond Exp</u>

Proof of (2):

(1) If $A \in \mathcal{F}_n$, then there exist $\omega^1, \ldots, \omega^k \in \Omega$ such that A is the *disjoint union* of $\Pi_n(\omega^1), \ldots, \Pi_n(\omega^k)$.



(2) For any $\omega \in \Omega$, $\displaystyle\sum_{\omega' \in \Pi_n(\omega)} \boldsymbol{E}_n X(\omega') p(\omega') = \sum_{\omega' \in \Pi_n(\omega)} X(\omega') p(\omega')$

$$\sum_{\omega' \in \Pi_n(\omega)} \left(\frac{\displaystyle\sum_{\omega'' \in \Pi_n(\cancel{\omega'})} X(\omega'') p(\omega'')}{\underbrace{\displaystyle\sum_{\omega'' \in \Pi_n(\cancel{\omega'})} p(\omega'')}_{\text{ind of } \omega'}} \right) p(\omega') =$$

QED

$$(3)\quad \text{Hence} \sum_{\underline{\omega \in A}} \boldsymbol{E}_n X(\omega) p(\omega) = \sum_{i=1}^{k} \sum_{\boxed{\omega \in \Pi_n(\omega^i)}} \boldsymbol{E}_n X(\omega) p(\omega) = \sum_{i=1}^{k} \sum_{\omega \in \Pi_n(\omega^i)} X(\omega) p(\omega) = \sum_{\omega \in A} X(\omega) p(\omega)\,.$$

Definition 5.24. Let X be a random variable, and $n \leqslant N$. We define the *conditional expectation of X given $\mathcal{F}_n$*, denoted by $\boldsymbol{E}_n X$, or $\boldsymbol{E}(X \mid \mathcal{F}_n)$, to be the unique *random variable* such that:

(1) $\boldsymbol{E}_n X$ is a $\mathcal{F}_n$-measurable random variable.
(2) For every $A \subseteq \mathcal{F}_n$, we have $\sum_{\omega \in A} \boldsymbol{E}_n X(\omega) p(\omega) = \sum_{\omega \in A} X(\omega) p(\omega)$.

Remark 5.25. This is the definition that generalizes to the continuous case. All properties we develop on conditional expectations will only use the above definition, and not the explicit formula.

Remark 5.26 (Uniqueness). If Y and Z are two $\mathcal{F}_n$-measurable random variables such that $\sum_{\omega \in A} Y(\omega)p(\omega) = \sum_{\omega \in A} Z(\omega)p(\omega)$ for every $A \in \mathcal{F}_n$, then we must have $\boldsymbol{P}(Y = Z) = 1$.