EXPECTATION OF A RANDOM VARIABLE

Recall def of Lebesgue integral (measure theory, Durrett 1.4):

$(\Omega, \mathcal{F}, \mu)$ prob. space, $f: \Omega \to \mathbb{R}$ measurable.

- For indicator functions $\int \mathbb{1}_E \, d\mu := \mu(E)$
- For simple functions $\int \left(\sum_1^n a_i \mathbb{1}_{E_i} \right) d\mu := \sum_1^n a_i \mu(E_i)$

  disjoint
- For nonnegative mble functions
  $$\int f \, d\mu := \sup \left\{ \int g \, d\mu : 0 \le g \le f, \ g \text{ mble} \right\}$$
- For $\forall$ mble functions: $f = f^+ + f^-$, $\int f \, d\mu := \int f^+ d\mu - \int f^- d\mu$.

f is <u>integrable</u> if $\int f \, d\mu$ exists $\in \mathbb{R} \cup \{\pm\infty\}$ $\iff$ f^+, f^- integrable $\iff$ $|f|$ integrable.

- $\int_E f \, d\mu := \int f \mathbb{1}_E \, d\mu$.

<u>Def</u> Let X be a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$. Then
$$\mathbb{E}[X] := \int X \, d\mathbb{P}$$

"expected value", "mean", "expectation"

<u>Ex</u> $X = \#$ H's in 2 flips

Prob: $1/4 \quad 1/4 \quad 1/4 \quad 1/4$

$\Omega = \{HH, HT, TH, TT\}$

$X \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$

$\quad 2 \quad 1 \quad 1 \quad 0$

Simple function

$$\mathbb{E}[X] = 2 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} = 1$$

<u>Prop</u> $\mathbb{E}X$ is uniquely determined by the distribution $\mathcal{L}_X$ of X

$\ulcorner$ 1. For simple r.v's $X = a_i$ with prob. p_i, i.e.

If $X = \sum_{i=1}^n a_i \mathbb{1}_{E_i} \Rightarrow \mathbb{E}[X] \underset{\text{def of } \int}{=} \sum_{i=1}^n a_i \mathbb{P}(E_i) = \sum_{i=1}^n a_i \mathbb{P}\{X_i = a_i\} = \sum_{i=1}^n a_i \mathcal{L}_X(\{a_i\})$. ✓

2. General case is determined by simple r.v's $\lrcorner$

Expressing $\mathbb{E}X$ in terms of distribution $\mathcal{L}_X$?

Note: R.v. X on $(\Omega, \mathcal{F}, \mu)$ has the same distr. as r.v. $Y(x) := x$ on $(\mathbb{R}, \mathcal{B}, \mathcal{L}_X)$

$\forall B \in \mathcal{B}$: $\underbrace{\mathbb{P}\{X \in B\}}_{\mathcal{L}_X(B)} \overset{?}{=} \underbrace{\mathbb{P}\{Y \in B\}}_{\mathcal{L}_X(\{x \in \mathbb{R} : \overbrace{Y(x)}^{x} \in B\})}$,

$\mathcal{L}_X(\{x \in \mathbb{R} : x \in B\}) = \mathcal{L}_X(B)$ $\square$

Prop p.18

$$\Rightarrow \quad \mathbb{E}X = \mathbb{E}Y = \int_{-\infty}^{\infty} x \, d\mathcal{L}_X$$

PROP $\forall$ r.v. X,

$$\mathbb{E}X = \int_{-\infty}^{\infty} x \, d\mathcal{L}_X$$

More generally: $\forall$ mble $g : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$:

$$\mathbb{E}g(X) = \int_{-\infty}^{\infty} g(x) \, d\mathcal{L}_X$$

(same proof; note that $X \overset{dist}{=} Y \Rightarrow g(X) \overset{dist}{=} g(Y)$)

Expressing $\mathbb{E}X$ in terms of CDF:

$F(x) = \mathbb{P}\{X \le x\} = \mathcal{L}_X((-\infty, x])$:

$$\mathbb{E}X = \int_{-\infty}^{\infty} x \, dF(x)$$

← (def of Lebesgue-Stieltjes integral)

PROPERTIES OF EXPECTATION

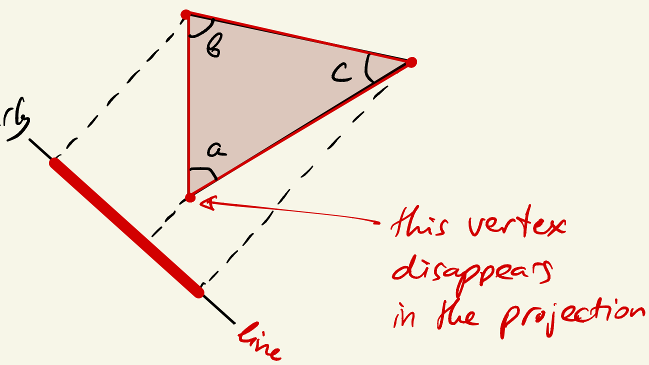
(follow from properties of Lebesgue integral):

- Linearity: $\mathbb{E}\left[\sum_{i=1}^{N} a_i X_i\right] = \sum_{i=1}^{N} a_i \mathbb{E}X_i$
- Constant: $X = a$ a.s $\Rightarrow \mathbb{E}X = a$.
- Monotonicity: $X \le Y$ a.s $\Rightarrow \mathbb{E}X \le \mathbb{E}Y$
- Dominated Convergence: $X_n \to X$ a.s & $|X_n| \le Y$ a.s. $\forall n$ & $\mathbb{E}|Y| < \infty$

  $\Rightarrow \mathbb{E}[X_n] \to \mathbb{E}[X]$

Examples: (a) $\boxed{\sum \text{angles of } \forall \text{ triangle} = \pi}$ A probabilistic proof:

- Project the Δ onto a random line (uniform orientation):
- \# disappearing vertices $= 1$ always
- $X_a := \begin{cases} 1 & \text{if vertex } a \text{ disappears} \\ 0 & \text{otherwise} \end{cases}$; X_b, X_c similarly

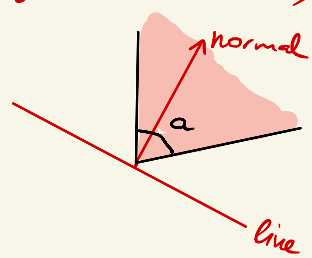


$$\Rightarrow X_a + X_b + X_c = 1 \quad \text{a.s.}$$

$$\Rightarrow \mathbb{E}[X_a] + \mathbb{E}[X_b] + \mathbb{E}[X_c] = 1$$

- But $\mathbb{E}[X_a] = \mathbb{P}\{\text{vertex } a \text{ disappears}\}$ (def of Lebesgue integral: $\int \mathbb{1}_E \, d\mathbb{P} = \mathbb{P}(E)$)

$$= \mathbb{P}\{\text{normal of the line} \in \text{cone}(a)\} = \frac{a}{\pi}$$



$$\Rightarrow \frac{a}{\pi} + \frac{b}{\pi} + \frac{c}{\pi} = 1 \quad \Rightarrow \quad \boxed{a + b + c = \pi}$$

(b) (Derangements p. 11)

What is the expected $\underbrace{\text{\# of students who receive their own exam?}}_{= X}$

$$X = \sum_{i=1}^{n} X_i \quad \text{where} \quad X_i = \begin{cases} 1 & \text{if student } i \text{ receives own exam} \\ 0 & \text{if not} \end{cases}$$

$$\Rightarrow \mathbb{E}X = \sum_{i=1}^{n} \mathbb{E}X_i = \sum_{i=1}^{n} \underbrace{\mathbb{P}\{\text{student } i \text{ receives own exam}\}}_{= 1/n} = 1.$$

DISCRETE DISTRIBUTIONS

Def X has discrete distr. if $\mathcal{L}_X$ is supported on a finite or countable set

$\Updownarrow$

X takes on finite or countable # of values $x_1, x_2, \dots$

$\Rightarrow$ <u>Probability</u> <u>mass function</u> $x_i \longmapsto P\{X=x_i\}$ determines distr of X.

$$EX = \int_{-\infty}^{\infty} x \, d\mathcal{L}_X = \boxed{\sum_i x_i P\{X_i = x_i\}}$$

More generally: $\boxed{Eg(X) = \sum_i g(x_i) P\{X=x_i\}}$

Ex

(a) <u>Bernoulli</u> distribution: $\boxed{X \sim Ber(p)}$ if $P\{X=1\}=p$, $P(X=0)=1-p$.

$EX = 1\cdot p + 0\cdot(1-p) = p$.

(b) e.g. <u>indicator</u> $X = \mathbb{1}_E$ of an event E. $X \sim Ber(p)$ with $p = P(E)$.

(c) $X = $ # Heads in 2 flips

$$EX = 0 \cdot P\{X=0\} + 1 \cdot P\{X=1\} + 2 \cdot P(X=2) = 0\cdot\frac{1}{4} + 1\cdot\frac{1}{2} + 2\cdot\frac{1}{4} = 1.$$

(d) $X \sim Unif\{x_1, \dots, x_n\} \Rightarrow P\{X=x_i\} = \frac{1}{n} \Rightarrow EX = \frac{1}{n}\sum_1^n x_i$ = arithmetic mean

(e) (<u>St. Petersburg</u> <u>Paradox</u>)

The initial stake (money in the pot) is \$2.

\$2

The player tosses a fair coin.

If T $\Rightarrow$ game ends, the player wins whatever is in the pot.

If H $\Rightarrow$ stake doubles, game continues.

What would be a fair price to pay the casino for entering this game?

X = player's winnings.

$P\{X=2\} = 1/2$; $P\{X=4\} = 1/4$, ...

$$E[X] = 2\cdot\frac{1}{2} + 4\cdot\frac{1}{4} + 8\cdot\frac{1}{8} + \dots = \infty$$

Tosses	Winnings \$	Prob.
T	2	1/2
HT	4	1/4
HHT	8	1/8
HHT	16	1/16
...		...

1/2 ← Median

1/2

<u>Ans</u>: <u>$\forall$ price</u>.

<u>Remarks</u>:
- Volatility, financial bubbles.
- <u>Median</u> is more stable. $Med(X) = M$ if $P\{X \le M\} \ge \frac{1}{2}$, $P\{X \ge M\} \ge \frac{1}{2}$.
- In the St. Pete's paradox, $Med(X) \in [2,4]$

(f): Number of cycles in a random permutation

- Recall: permutation = a rearrangement of n distinct objects.
  There are $n!$ permutations

- $\forall$ permutation breaks down into cycles.

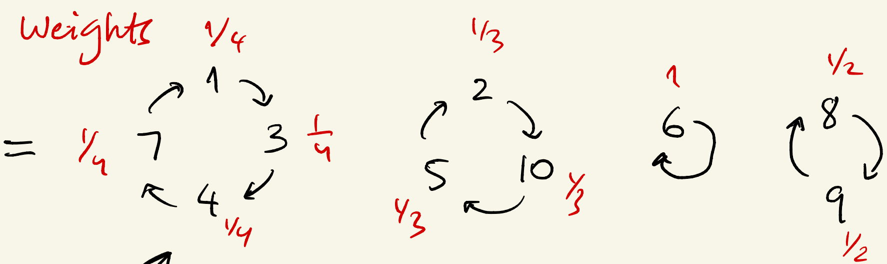
For example:

people

	1	2	3	4	5	6	7	8	9	10
π:	↓	↓	...							
	3	10	4	7	2	6	1	9	8	5

jackets

=



Q! What is $E\,\#(\text{cycles})$ in a random permutation ?

drawn from the set of all $n!$ permutations with uniform probability.

- Intuition: 1^{st} cycle has length $\sim n$ (e.g. $n/2$), next $n/4$, ... $\Rightarrow \log n$.

- Thm The expected # of cycles in a random permutation of n elements is

$$\sum_{k=1}^{n} \frac{1}{k} \approx \ln n + 0.58$$

Proof
- To each element in a cycle of length k, assign weight $:= 1/k$.

Sum of all weights = # cycles

E.g, here sum= $1+1+1+1=4$

• Formally, let

$$Y_j := \text{length of the cycle the element } j \text{ is in};$$

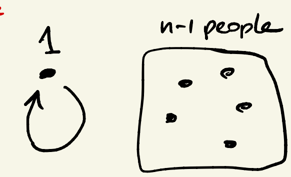
$$X = \sum_{j=1}^{n} \frac{1}{Y_j}$$

Linearity $\Rightarrow$ $$E[X] = \sum_{j=1}^{n} \underbrace{E\left[\tfrac{1}{Y_j}\right]}_{\text{all equal by symmetry}} = n \cdot \underbrace{E\left[\tfrac{1}{Y_1}\right]}_{\overset{\|}{?}}$$

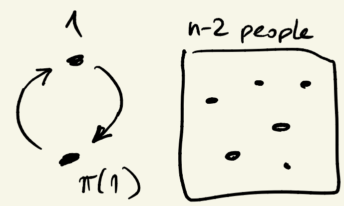
$$E\left[\frac{1}{Y_1}\right] = \sum_{k=1}^{n} \frac{1}{k} \underbrace{P\{Y_1 = k\}}_{\|}$$

$$P\{\text{element } 1 \text{ is in a cycle of length } k\} = ?$$

• $$P\{1 \text{ is in a cycle of length } 1\} = \frac{\overbrace{(n-1)!}^{\#\text{ ways to permute other } n-1 \text{ people}}}{n!} = 1/n$$



$$P\{1 \text{ is in a cycle of length } 2\} = \frac{\overbrace{(n-1)}^{\#\text{ways to choose } \pi(1)}\overbrace{(n-2)!}^{\#\text{ ways to permute other } n-2 \text{ people}}}{n!} = \frac{1}{n}$$



...

$$P\{1 \text{ is in a cycle of length } k\} = \frac{1}{n} \quad \text{(DIY)}$$

• $$\Rightarrow E\left[\frac{1}{Y_1}\right] = \sum_{k=1}^{n} \frac{1}{k} \cdot \frac{1}{n} \Rightarrow E[X] = \sum_{k=1}^{n} \frac{1}{k} \qquad \text{QED}$$