Math 170E: Winter 2023

Lecture 8, Fri 27th Jan

Mathematical expectation

Announcements:

1) Midterm 1.

Fri Feb 3rd.

Location: 4000A

Time: 12 – 12:50

↳ 50 mins.

→ Content: lectures ≤ 9

→ Practice uploaded over the weekend.

→ Cheat sheet:

Letter sized 1 sheet.

both sides

→ Basic calculator

2) HW 2 due tonight.

HW3 due Sat. Feb 4th.

Last time:

- A discrete random variable is a function $X : \Omega \to S$, where $S \subseteq \mathbb{R}$ where S is finite or countable
- The PMF of a discrete r.v. X is the map $p_X : S \to [0, 1]$

$$p_X(x) = \mathbb{P}(X = x)$$

- The CDF of a discrete r.v. X is the map $F_X : \mathbb{R} \to [0, 1]$

$$F_X(x) = \mathbb{P}(X \leq x)$$

- $X \sim \text{Uniform}(\{1, 2, \ldots, m\})$ if $p_X(x) = \frac{1}{m}$ for $x \in \{1, 2, \cdots, m\}$

Today:

We'll discuss today:

- How to compute the expected value of a discrete random variable
- What a Bernoulli random variable is
- Properties of the expectation

Definition 2.7: (**Expected value**)

If X is a discrete random variable taking values in a countable set $S \subseteq \mathbb{R}$, its expected value is defined to be

$$\mathbb{E}[X] = \sum_{x \in S} x\, p_X(x),$$

provided the sum converges.

Notation: we also write $\mu_X = E[X]$.

Example 6: Let $p \in (0,1)$. We say that a discrete random variable X is a Bernoulli random variable and write $X \sim \text{Bernoulli}(p)$ if it has PMF

$$p_X(x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = 0 \end{cases}$$

$S = \{0,1\}$.

↳ "Weighted coin flips"

"Success or a failure"

$p = 1/2$: Fair coin flip.

What is $\mathbb{E}[X]$?

$$\mathbb{E}[X] = \sum_{x \in S} x\, p_X(x)$$

$$= \sum_{x \in \{0,1\}} x\, p_X(x) = 0 \cdot p_X(0) + 1 \cdot p_X(1) = p.$$

⚠ $\mathbb{E}(X)$ doesn't have to correspond to any element of S. → If $p = 1/2$, $\mathbb{E}[X] = 1/2$.

Example 7: Uniform random variables

Given $m \geq 1$, let $X \sim \text{Uniform}(\{1, 2, \ldots, m\})$.

What is $\mathbb{E}[X]$?

A) $\dfrac{m(m+1)}{2}$

B) $\dfrac{1}{m}$

C) $\dfrac{m+1}{2}$

D) $\dfrac{m}{2}$

Recall $\sum_{x=1}^{m} x = \frac{m(m+1)}{2}$

$P_X(x) = 1/m$ for all $x \in \{1, \dots, m\}$.

$$\mathbb{E}[X] = \sum_{x=1}^{m} x P_X(x) = \sum_{x=1}^{m} x \cdot \frac{1}{m} = \frac{1}{m} \sum_{x=1}^{m} x$$

$$= \frac{1}{m}(1+2+\dots+m).$$

$$= \frac{1}{\cancel{m}} \cdot \frac{\cancel{m}(m+1)}{2} = \frac{m+1}{2}.$$

average of #s in $\{1, \dots, m\}$.

e.g.: $m=6$, fair d6 roll.

$$\mathbb{E}[X] = \frac{6+1}{2} = 7/2 = 3.5.$$

1 2 3 | 4 5 6

Why $\sum_{x=1}^{m} x = \frac{m(m+1)}{2}$?

Let $S_m = \sum_{x=1}^{m} x = 1 + 2 + \dots + m$

$$= m + (m-1) + \dots + 1.$$

$$\Rightarrow 2S_m = m(m+1)$$

$$\Rightarrow S_m = \frac{m(m+1)}{2}.$$

Proposition 2.8:

If X is a discrete random variable taking values in a countable set $S \subseteq \mathbb{R}$, and $g : S \to \mathbb{R}$ is a function, then the expected value of $g(X)$ is

↪ New r.v.

$$\mathbb{E}[g(X)] = \sum_{x \in S} g(x)\, p_X(x),$$

$g(X)(\omega) = g(X(\omega))$

provided the sum converges.

e.g. $g(x) = x$, $\mathbb{E}[X] = \mathbb{E}[g(X)]$

In particular, $\mathbb{E}[a] = a$ for any $a \in \mathbb{R}$

↪ $g(x) = a$, $a \in \mathbb{R}$ constant.

$g(X) = a$.

$$\mathbb{E}[a] = \mathbb{E}[g(X)] = \sum_{x \in S} a \cdot p_X(x)$$

$$= a \underbrace{\sum_{x \in S} p_X(x)}_{=1}$$

$$= a.$$

Proof: $g: S \to \mathbb{R}$, $S = \{x_1, \ldots\} = \{x_i\}_{i \in \mathbb{N}}$.

Image of g under S $= \{y_1, y_2, \ldots\} = \{y_j\}$.

Complication: g is a function so each x_i gets mapped to a y_j, but many x_i could give the same y_j

$\hookrightarrow$ g doesn't have one-to-one.

$$\sum_{x \in S} g(x) P_X(x) = \sum_i g(x_i) P_X(x_i)$$

$$= \sum_j \sum_{i : g(x_i) = y_j} g(x_i) P_X(x_i)$$

$$= \sum_j \sum_{i : g(x_i) = y_j} y_j P_X(x_i)$$

$$= \sum_j y_j \sum_{i : g(x_i) = y_j} P_X(x_i)$$

$S = \{x_1, x_2, x_3\}$

$\{y_1, y_2\}$

$g(x_1) = g(x_2) = y_1$

$g(x_3) = y_2$.

$$\sum_{i \in \{1,2,3\}} g(x_i) P_X(x_i)$$

$$= \sum_{j \in \{1,2\}} \sum_{i : g(x_i) = y_j} \bigcirc$$

$$= \underbrace{\sum_{i : g(x_i) = y_1} \bigcirc}_{j=1} + \underbrace{\sum_{i : g(x_i) = y_2} \bigcirc}_{j=2}$$

$$= \sum_j y_j \, \mathbb{P}(g(X) = y_j).$$

$$= \sum_j y_j \, P_{g(X)}(y_j) = \mathbb{E}[g(X)].$$

Example 8: Let $X : \Omega \to \{-1, 0, 1\}$ be a discrete r.v. such that

$$p_X(-1) = 0.2, \quad p_X(0) = 0.5, \quad p_X(1) = 0.3. \to 0.98.$$

0.01 0.01

What is $\mathbb{E}[X^2]$?

Approach 1: Use a transformation

$g : \{-1,0,1\} \to \mathbb{R}$, $g(x) = x^2$

$\hookrightarrow g(X) = X^2$.

$g(X(\omega)) = X(\omega)^2$.

So by Prop$^{\underline{n}}$ 2.8,

$$\mathbb{E}[X^2] = \mathbb{E}[g(X)] = \sum_{x \in \{-1,0,1\}} g(x) p_X(x).$$

$$= \sum_{x \in \{-1,0,1\}} x^2 p_X(x) = (-1)^2 p_X(-1) + 0^2 \cdot p_X(0) + 1^2 \cdot p_X(1),$$

$$= p_X(-1) + p_X(1) = 0.5.$$

Approach 2: New r.v.

Define $Y = X^2$. disc. r.v.

$\hookrightarrow Y \in \{0,1\}$.

Compute PMF for Y:

$$p_Y(0) = \mathbb{P}(Y=0) = \mathbb{P}(X=0) = 0.5.$$

$$p_Y(1) = \mathbb{P}(Y=1) = \mathbb{P}(X^2=1) = \mathbb{P}(\{X=-1\} \cup \{X=1\}),$$

$$= \mathbb{P}(X=-1) + \mathbb{P}(X=1).$$

$$= 0.2 + 0.3 = 0.5.$$

$$\Rightarrow \mathbb{E}[X^2] = \mathbb{E}[Y] = \sum_{y \in \{0,1\}} y\, p_Y(y) = 1 \cdot (0.5) = 0.5.$$

Proposition 2.9: If X is a discrete r.v.taking values in a countable set $S \subseteq \mathbb{R}$. If $a, b \in \mathbb{R}$ and $g, h : S \to \mathbb{R}$, then

$$\mathbb{E}[\underbrace{ag(X) + bh(X)}] = a\mathbb{E}[g(X)] + b\mathbb{E}[h(X)]$$

"Expectation is linear"

Let $f(x) = ag(x) + bh(x)$.

$\hookrightarrow$ New r.v. $f(X) = ag(X) + bh(X)$.

Proof:

$$\mathbb{E}[ag(X) + bh(X)] = \mathbb{E}[f(X)].$$

$$= \sum_{x \in S} f(x) p_X(x).$$

$$= \sum_{x \in S} (ag(x) + bh(x)) p_X(x)$$

$$= a \underbrace{\sum_{x \in S} g(x) p_X(x)}_{\mathbb{E}[g(X)].} + b \underbrace{\sum_{x \in S} h(x) p_X(x)}_{\mathbb{E}[h(X)]}$$

Example 9: Let X be a discrete random variable. What is $\mathbb{E}[X - \mathbb{E}[X]]$?

with $|\mathbb{E}[X]| < +\infty$

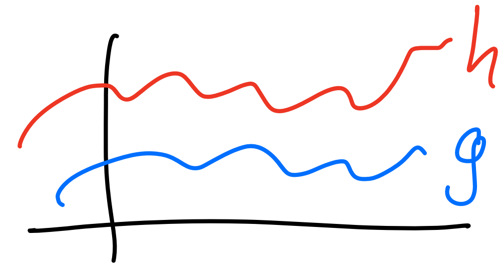
constant $\leftarrow$ not random

$\hookrightarrow$ By linearity,

$$\mathbb{E}[X - \mathbb{E}[X]] = \mathbb{E}[X] + \mathbb{E}[-\mathbb{E}[X]]$$
$$= \mathbb{E}[X] - \mathbb{E}[X] = 0.$$

Proposition 2.10: If X is a discrete r.v.taking values in a countable set $S \subseteq \mathbb{R}$.

If $g, h : S \to \mathbb{R}$ such that $g(x) \leq h(x)$ for all $x \in S$, then



$$\mathbb{E}[g(X)] \leq \mathbb{E}[h(X)]$$

Proof:

$$\mathbb{E}[g(X)] = \sum_{x \in S} g(x) \overbrace{p_X(x)}^{\text{all non-negative for all } x \in S}$$

$$\leq \sum_{x \in S} h(x) p_X(x) = \mathbb{E}[h(X)].$$

Example 10: Let X be a discrete random variable. Show that

$$-1 \leq \mathbb{E}[\sin(X)] \leq 1$$

We have $-1 \leq \sin(x) \leq 1$ for all $x \in \mathbb{R}$.

$$-1 = \mathbb{E}[-1] \leq \mathbb{E}[\sin(X)] \leq \mathbb{E}[1] = 1$$