

CS 237: PROBABILITY IN COMPUTING

INSTRUCTORS: JOHN BYERS, TIAGO JANUARIO



TOP HAT QUESTION

You roll two dice. Let X_1 be the number on the 1st die, X_2 be the number on the 2nd die, and $X = |X_1 - X_2|$. Find $\mathbb{E}(X)$.

A. 0

B. 1

C. $\frac{70}{36}$

D. 2

E. 3.5

F. None of the above

LINEARITY OF EXPECTATION

- ▶ **Theorem:** For **any** random variables X and Y

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

- ▶ The theorem holds even if X and Y are dependent!
- ▶ More generally:

$$\mathbb{E}(aX + b) = a \cdot \mathbb{E}(X) + b \quad \text{for any constants } a, b \text{ and r.v. } X$$

$$\mathbb{E}(c_1X_1 + c_2X_2 + \cdots + c_nX_n) = c_1\mathbb{E}(X_1) + c_2\mathbb{E}(X_2) + \cdots + c_n\mathbb{E}(X_n)$$

for any **constants** c_1, \dots, c_n and random variables X_1, \dots, X_n



SUMS OF RANDOM VARIABLES

- ▶ **Experiment:** roll two fair dice
- ▶ Let X be the number on the 1st die and Y be the number on the 2nd die
- ▶ Let S be the sum of the two rolls, i.e., $S = X + Y$
- ▶ If X and Y are independent, we can find $\mathbb{E}(S)$ using the definition and a fair bit of work (try it as an exercise)
- ▶ It turns out that there is a much easier way:

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y) = 7/2 + 7/2 = 7$$

INDICATOR RANDOM VARIABLES

- ▶ An **indicator random variable** takes on two values: 0 and 1.
- ▶ **Lemma.** For an indicator random variable Y ,
$$\mathbb{E}(Y) = \Pr(Y = 1).$$

Proof:

INDICATOR RANDOM VARIABLES

- ▶ An **indicator random variable** takes on two values: 0 and 1.

- ▶ **Lemma.** For an event E , an **indicator random variable Y for an event E** is

Proof:

$$Y = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

Method for computing expectation of a random variable X :

Represent X as a sum of indicators and use linearity of expectation.

TOP HAT QUESTION

You have a coin that shows HEADS with probability $3/4$. Let X be the number of HEADS in 1000 tosses of your coin. You represent X as the sum: $X = X_1 + X_2 + \dots + X_{1000}$. What is X_1 ?

- A. $3/4$.
- B. The number of HEADS.
- C. The probability of HEADS in toss 1.
- D. The number of heads in toss 1.
- E. None of the above.

TOP HAT QUESTION

You have a coin that shows HEADS with probability $3/4$. Let X be the number of HEADS in 1000 tosses of your coin.

What is the expectation of X ?

- A. $3/4$.
- B. $4/3$.
- C. 500.
- D. 750.
- E. None of the above.

RANDOM PROOFS

- ▶ Infinite monkey theorem:



Given enough time, a hypothetical chimpanzee typing at random would, as part of its output, almost surely produce one of Shakespeare's plays (or any other text).

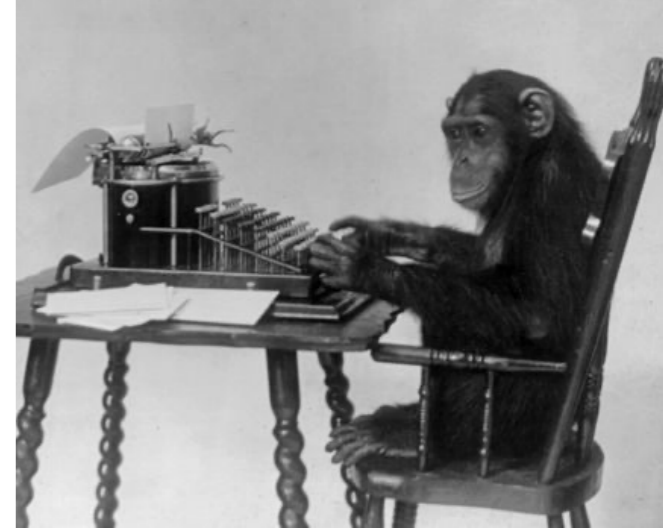
https://en.wikipedia.org/wiki/Infinite_monkey_theorem



RANDOM PROOFS

- ▶ Suppose that the monkey types on a 26-letter keyboard
- ▶ Each letter is chosen independently and uniformly at random from the 26 letters
- ▶ Suppose that the monkey types $n = 1,000,000$ letters
- ▶ Let X = number of times the sequence "proof" appears
- ▶ What is $\mathbb{E}(X)$?

RANDOM PROOFS



TOP HAT QUESTION

3 students exchange their hats, so that everybody gets a random hat. Let X be the number of students that got their own hats. **What's the expectation of X ?**

A. $3/4$

B. 1

C. $3/2$

D. None of the above.

TOP HAT QUESTION

If students 1,2,3 got hats 2,1,3 then $X=1$.

- ▶ PDF of X :

$$\Pr(X = 0) = \frac{1}{3}, \Pr(X = 1) = \frac{1}{2}, \Pr(X = 3) = \frac{1}{6}.$$

EXAMPLE: RANDOM HATS

- ▶ **Permutations**
 - ▶ n students exchange their hats, so that everybody gets a random hat
 - ▶ Let X be the number of students that got their own hats.
- ▶ What's the expectation of X for general n ?

EXPECTATION FOR CONTINUOUS RVS

- ▶ The expectation of a continuous random variable is defined as in the discrete case, but we replace the sum with an integral
- ▶ Let X be a random variable with PDF f_X

discrete: $\mathbb{E}(X) = \sum_{x \in \mathbb{R}} x \cdot f_X(x)$

continuous: $\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$

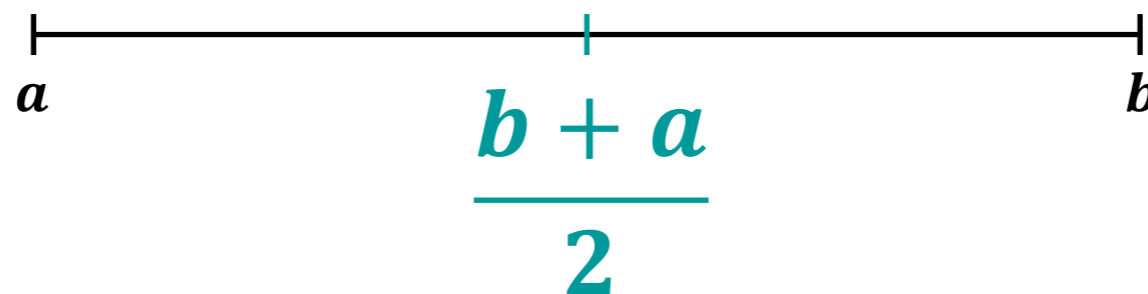
EXAMPLE: UNIFORM RANDOM VARIABLE

- ▶ Suppose X is chosen uniformly at random from the interval $[a, b]$. The PDF is $f_X(x) = \frac{1}{b-a}$ for $x \in [a, b]$

- ▶
$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) dx = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \int_a^b x dx$$

$$= \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b = \frac{1}{b-a} \frac{b^2 - a^2}{2} = \frac{b+a}{2}$$

- ▶ **Note:** $\mathbb{E}(X)$ is the midpoint of the interval $[a, b]$



EXAMPLE: DARTS

- ▶ Suppose we throw a dart at a circular board of radius r . Let D be the distance from the center of the board to the hitting point of the dart. The PDF is $f_X(x) = \frac{2x}{r^2}$ for $x \in [0, r]$

- ▶
$$\begin{aligned}\mathbb{E}(D) &= \int_{-\infty}^{\infty} x \cdot f_X(x) dx = \int_0^r x \cdot \frac{2x}{r^2} dx = \frac{2}{r^2} \int_0^r x^2 dx \\ &= \frac{2}{r^2} \left. \frac{x^3}{3} \right|_0^r = \frac{2}{r^2} \frac{r^3}{3} = \frac{2r}{3}\end{aligned}$$

LAW OF THE UNCONSCIOUS STATISTICIAN (LOTUS)

- ▶ We saw earlier several examples of random variables that are functions of another random variable
- ▶ Examples: $Y = 3X + 1$ $Y = X^2$ $Y = |X|$
- ▶ We also saw how the PDF of Y is related to the PDF of X
- ▶ This leads to the following result for computing expectation:

X is discrete: $\mathbb{E}(g(X)) = \sum_{x \in \mathbb{R}} g(x) \cdot f_X(x)$

X is continuous: $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx$

FUNCTIONS OF RANDOM VARIABLES

- ▶ Let X be a random variable that is uniform over $[0, 1]$

$$f_X(x) = \begin{cases} 1 & \text{if } x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f_X(x) dx = \int_0^1 x^2 dx = \frac{x^3}{3} \Big|_0^1 = \frac{1}{3}$$

LAW OF THE UNCONSCIOUS STATISTICIAN (LOTUS)

- ▶ Let X be uniform over $\{-4, \dots, 4\}$. Then PMF of X is

$$f_X(x) = \begin{cases} \frac{1}{9} & \text{if } x \in \{-4, \dots, 4\} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}(X^2) =$$

LAW OF THE UNCONSCIOUS STATISTICIAN (LOTUS)

- ▶ If $Y = g(X)$ then

discrete: $\mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_{x \in \mathbb{R}} g(x) \cdot f_X(x)$

continuous: $\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx$

- ▶ Let X be uniform over $\{-4, \dots, 4\}$. Then PMF of X is

$$f_X(x) = \begin{cases} \frac{1}{9} & \text{if } x \in \{-4, \dots, 4\} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}(X^2) = \sum_{x \in \mathbb{R}} x^2 \cdot f_X(x) = \sum_{x \in \{-4, \dots, 4\}} x^2 \cdot \frac{1}{9} = \frac{2}{9} (1^2 + 2^2 + 3^2 + 4^2)$$