Randomness in Computing

CS 537

Lecture 4

Last time

- Randomized min-cut algorithm

Today

- Random variables
- Expectation
- Linearity of expectation
- Jensen’s inequality

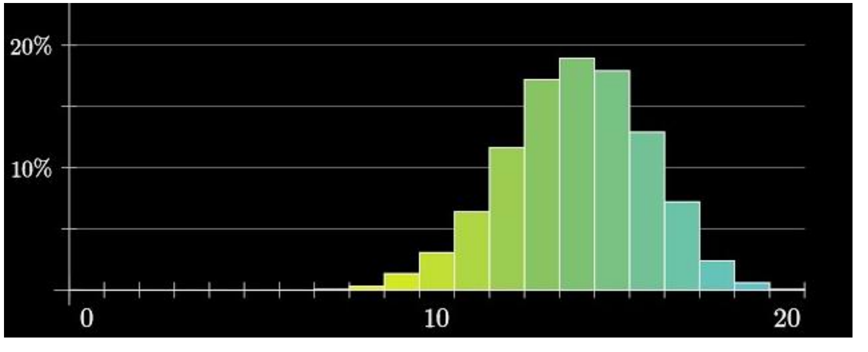
Measurements in random experiments

- Example 1: coin flips
  - Measurement X: number of heads.
  - E.g., if the outcome is HHTH, then X=3.

- Example 2: permutations
  - n students exchange their hats, so that everybody gets a random hat
  - Measurement X: number of students that got their own hats.
  - E.g., if students 1,2,3 got hats 2,1,3 then X=1.

Recall: random variables

- A random variable X on a sample space Ω is a function $X: \Omega \to \mathbb{R}$ that assigns to each sample point $\omega \in \Omega$ a real number $X(\omega)$.
- For each random variable, we should understand:
  - The set of values it can take.
  - The probabilities with which it takes on these values.
- The distribution of a discrete random variable X is the collection of pairs $\{(a, \Pr[X = a])\}$.



Question

You roll two dice. Let X be the random variable that represents the sum of the numbers you roll.

What is the probability of the event X=6?

A. 1/36
B. 1/9
C. 5/36
D. 1/6
E. None of the above.

Question

You roll two dice. Let X be the random variable that represents the sum of the numbers you roll.

How many different values can X take on?

A. 6

B. 11

C. 12

D. 36

E. None of the above.

Question

You roll two dice. Let X be the random variable that represents the sum of the numbers you roll.

What is the distribution of X?

A. Uniform distribution on the set of possible values.

B. It satisfies $\Pr[X = 2] < \Pr[X = 3] < \cdots < \Pr[X = 12]$.

C. It satisfies $\Pr[X = 2] > \Pr[X = 3] > \cdots > \Pr[X = 12]$.

D. It satisfies $\Pr[X = 2] < \Pr[X = 3] < \cdots < \Pr[X = 7]$ and $\Pr[X = 7] > \Pr[X = 8] > \cdots > \Pr[X = 12]$.

E. None of the above is true.

	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

CS 537

Independent RVs: definition

- Random variables X and Y are independent if
$$\Pr[(X = x) \cap (Y = y)]$$
$$= \Pr[X = x] \cdot \Pr[Y = y]$$

for all values x and y.

- Random variables $X_1, X_2, \ldots, X_n$ are mutually independent

if for all subsets of $I \subseteq [n]$ and all values x_i, where $i \in I$,
$$\Pr[\cap_{i \in I} (X_i = x_i)\ ]$$
$$= \prod_{i \in I} \Pr[X_i = x_i]\,.$$

Question

You roll one die. Let X be the random variable that represents the result.

What value does X take, on average?

A. 1/6

B. 3

C. 3.5

D. 6

E. None of the above.

Random variables: expectation

- The expectation of a discrete random variable X over a sample space Ω is

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega) \cdot \Pr[\omega].$$

- We can group together outcomes ω for which $X(\omega) = a$:

$$\mathbb{E}[X] = \sum_{a} a \cdot \Pr[X = a],$$

where the sum is over all possible values a taken by X.

- The second equality is more useful for calculations.

Example: random hats

- Example: permutations
  - n students take off their hats, then everybody gets a random hat
  - R.V. X: the number of students that got their own hats.
  - E.g., if students 1,2,3 got hats 2,1,3 then X=1.
- Distribution of X:

$$\Pr[X = 0] = 1/3, \Pr[X = 1] = 1/2, \Pr[X = 3] = 1/6.$$

- What's the expectation of X?

Linearity of expectation

- Theorem. For any two random variables X and Y on the same probability space,

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

Also, for all $c \in \mathbb{R}$,

$$\mathbb{E}[cX] = c \cdot \mathbb{E}[X].$$

Indicator random variables

- An indicator random variable takes on two values: 0 and 1.
- Lemma. For an indicator random variable X,
$$\mathbb{E}[X] = \Pr[X = 1].$$

Question

You have a coin with bias 3/4 (the bias is the probability of HEADS). Let X be the number of HEADS in 1000 tosses of your coin.

You represent X as the sum: $X = X_1 + X_2 + \cdots + X_{1000}$.

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.

Question

You have a coin with bias 3/4 (the bias is the probability of HEADS). 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.

Example: random hats

- Example: permutations
  - n students take off their hats, then everybody gets a random hat
  - R.V. X: the number of students that got their own hats.
- What's the expectation of X for general n?

Jensen's inequality: example

- Exercise: Let X be the length of a side of a square chosen from [99] uniformly at random. What is the expected value of the area?

  Solution: Find $\mathbb{E}[X^2]$.

$$\mathbb{E}[X^2] = \sum_{i=1}^{99} i^2 \cdot \frac{1}{99} = \frac{1}{99} \cdot \frac{99(99+1)(2 \cdot 99 + 1)}{6} = \frac{100 \cdot 199}{6} = \frac{9950}{3}$$

- Comparison. $(\mathbb{E}[X])^2 = \left(\frac{1+2+\cdots+99}{99}\right)^2 = \left(\frac{99 \cdot 50}{99}\right)^2 = 50^2 = 2500$

- In general, $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$

Jensen's inequality

In general, $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$

Proof: Let $\mu = \mathbb{E}[X]$. Consider $Y = (X-\mu)^2$.

$$
\begin{aligned}
0 \leq \mathbb{E}[Y] &= \mathbb{E}[(X-\mu)^2] \\
&= \mathbb{E}[X^2 - 2X\mu + \mu^2] \\
&= \mathbb{E}[X^2] - 2\mu\,\mathbb{E}[X] + \mu^2 \\
&= \mathbb{E}[X^2] - 2\mu^2 + \mu^2 \\
&= \mathbb{E}[X^2] - \mu^2
\end{aligned}
$$

We get: $\mathbb{E}[X^2] \geq \mu^2$

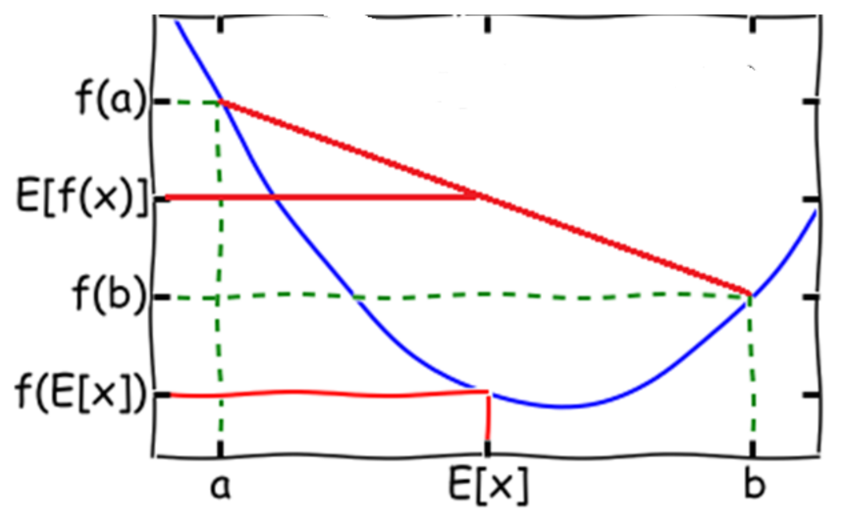
Jensen's inequality

A function $f: \mathbb{R} \to \mathbb{R}$ is convex if, for all $a, b \in \mathbb{R}$ and all $\lambda \in [0,1]$,

$$f(\lambda a + (1-\lambda)b) \le \lambda f(a) + (1-\lambda) f(b).$$

- Jensen's inequality. If f is a convex function and X is a random variable, then

$$f(\mathbb{E}[X]) \le \mathbb{E}[f(X)].$$



Average with secret inputs

Can n students in the class find out their average score on the exam without sharing their scores? (Scores are in $[t]$).

Solution: Let m be an integer larger than nt.

Let s_i be the score of student i, for all $i \in [n]$.

- Each student i picks $X_i[j]$ uniformly at random from 0 to $m-1$ for $j \in [n-1]$ and sets $X_i[n]$ so that
$$s_i = \sum_{j\in[n]} X_i[j] \mod m$$
- Each student $j \in [n]$ gets ``shares'' $X_i[j]$ for all $i \in [n]$, adds them up mod m and shows them to everybody.
- All n sums are added together mod m to obtain the sum of the scores, which is divided by n to obtain the average.