Lecture 5

Last time
- Randomized min-cut algorithm
- Random variables

Today
- Random variables, expectation
- Bernoulli and binomial RVs
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 $X=1$. 
Random variables: definition

• A random variable $X$ on a sample space $\Omega$ 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 $\{(a, \Pr[X = a])\}$. 
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.
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.
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.
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 $i \in I$,
  \[
  \Pr[\cap_{i \in I} (X_i = x_i)] = \prod_{i \in I} \Pr[X_i = x_i].
  \]
You roll two dice. Let $X_1$ be the value on the first die, $X_2$ be the value on the second die, and $X = X_1 + X_2$.

Which statements below are true?

A. $X_1$ and $X_2$ are independent.
B. $X_1$ and $X$ are independent.
C. $X_2$ and $X$ are independent.
D. $X_1, X_2$ and $X$ are mutually independent.
You roll one die. Let $X$ be the random variable that represents the result.

What value does $X$ take, on average?

A. $\frac{1}{6}$
B. 3
C. 3.5
D. 6
E. None of the above.
The expectation of a discrete random variable $X$ over a sample space $\Omega$ is

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

We can group together outcomes $\omega$ for which $X(\omega) = i$:

$$E[X] = \sum_i i \cdot \Pr[X = i],$$

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

The second equality is more useful for calculations.
Example from last lecture: random hats

- **Example:** permutations
  - $n$ students exchange their hats, so that 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 $X$:**
  \[
  \Pr[X = 0] = \frac{1}{3}, \quad \Pr[X = 1] = \frac{1}{2}, \quad \Pr[X = 3] = \frac{1}{6}.
  \]

- **What’s the expectation of $X$?**
Example: roulette

- 38 slots: 18 black, 18 red, 2 green.

- If we bet $1 on red, we get $2 back if red comes up. What’s the expected value of our winnings?
Linearity of expectation

- **Theorem.** For any two random variables $X$ and $Y$ on the same probability space,
  \[ E[X + Y] = E[X] + E[Y]. \]

  Also, for all $c \in \mathbb{R}$,
  \[ E[cX] = c \cdot E[X]. \]
A basket holds 100 chips labeled with integers 1 to 100. Two chips are drawn from the basket at random without replacement.

What is the expected value of their sum?
Indicator random variables

- An indicator random variable takes on two values: 0 and 1.
- Lemma. For an indicator random variable $X$,
  $$E[X] = \Pr[X = 1].$$
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 $X_1$?

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.
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?

B. 4/3.
C. 500.
D. 750.
E. None of the above.
A binomial random variable with parameters \( n \) and \( p \), denoted \( \text{Bin}(n, p) \), is the number of HEADS in \( n \) tosses of a coin with bias \( p \).

**Lemma.** The probability distribution of \( X = \text{Bin}(n, p) \) is

\[
\Pr[X = j] = \binom{n}{j} p^j (1 - p)^{n-j}
\]

for all \( j = 0, 1, \ldots, n \).

**Lemma.** The expectation of \( X = \text{Bin}(n, p) \) is

\[
E[X] = np.
\]
Throw $m$ balls into $n$ bins.
Let $X$ be the number of balls that land into bin 1.
(Recall that $\text{Bin}(n, p)$ is the binomial distribution, i.e., the distribution of the number of HEADS in $n$ tosses of a coin with bias $p$.)

Then the distribution of $X$ is
A. $\text{Bin} (n, m)$
B. $\text{Bin} (m, 1/n)$
C. $\text{Bin} \left( m, \frac{n-1}{n} \right)$
D. a binomial distribution, but none of the above.
E. not a binomial distribution.
Throw $m$ balls into $n$ bins. Let $Y$ be the number of empty bins.

Compute $E[Y]$.

A. $1$
B. $n/2$
C. $\left(1 - \frac{1}{n}\right)^m$
D. $n \left(1 - \frac{1}{n}\right)^m$
E. None of the above.
Theorem. For any two independent random variables $X$ and $Y$ on the same probability space,
\[ E[X \cdot Y] = E[X] \cdot E[Y]. \]

Note. The equality does not hold, in general, for dependent random variables.

Example. We toss two coins.
Let $X =$ number of HEADS, $Y =$ number of TAILS.
Calculate $E[X]$, $E[Y]$ and $E[XY]$. 
Jensen’s inequality

• A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is convex if, for all $x, y$ and all $\lambda \in [0,1]$,
  $$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

• **Jensen’s inequality.** If $f$ is a convex function and $X$ is a random variable, then
  $$E[f(X)] \geq f(E[X]).$$