Lecture 5

Last time

• Random variables
• Linearity of expectation
• Jensen’s inequality
• Bernoulli and binomial RVs

Today

• Conditional expectation
• Branching process
• Geometric RVs
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
\[
\mathbb{E}[X] = np.
\]
Throw \( m \) balls into \( n \) bins uniformly and independently at random. 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. \( \text{Bin} \left( n, m \right) \)
B. \( \text{Bin} \left( m, \frac{1}{n} \right) \)
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 $\mathbb{E}[Y]$.

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.
For arbitrary random variables $X$ and $Y$, by linearity of expectation:

A. $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$  
B. $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for all $a, b \in \mathbb{R}$.
C. $\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y]$  
D. Both A and B are correct.
E. A, B and C are correct.
Product of independent RVs

• **Theorem.** For any two independent random variables $X$ and $Y$ on the same probability space,
  \[ \mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]. \]

• **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 $\mathbb{E}[X]$, $\mathbb{E}[Y]$ and $\mathbb{E}[XY]$. 
For any random variable $X$ and event $E$, the conditional expectation of $X$ given $E$ is

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

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

You roll two dice. Let $X_1$ be the value on the first die, $X_2$ be the value on the second die, and $X = X_1 + X_2$.

Calculate $\mathbb{E}[X_1 | X = 5]$
• **Theorem.** For all random variables $X$ and $Y$ and all events $A$,

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

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

$$
\mathbb{E}[cX \mid A] = c \cdot \mathbb{E}[X \mid A].
$$
You roll two dice. Let $A$ be the event that you got no sixes. Let $X_1$ be the value on the first die, $X_2$ be the value on the second die, and $X = X_1 + X_2$.

Calculate $\mathbb{E}[X \mid A]$
Law of total expectation

• **Lemma.** For any two random variables $X$ and $Y$,

$$
\mathbb{E}[X] = \sum_y \mathbb{E}[X \mid Y = y] \Pr[Y = y].
$$
Recall

Let $A$ be an event and let $E_1, \ldots, E_n$ be mutually disjoint events whose union is $\Omega$.

- **Law of total probability.**

\[
\Pr[A] = \sum_{i \in [n]} \Pr[A \cap E_i] = \sum_{i \in [n]} \Pr[A \mid E_i] \cdot \Pr[E_i].
\]
Law of total expectation

Let $X$ be a random variable over sample space $\Omega$ and let $E_1, \ldots, E_n$ be mutually disjoint events whose union is $\Omega$. Then

$$
\mathbb{E}[X] = \sum_{i \in [n]} \mathbb{E}[X \mid E_i] \cdot \Pr[E_i].
$$
Conditional expectation: definition

• For random variables X and Y, the **conditional expectation** of X given Y, denoted $\mathbb{E}[X|Y]$, is a random variable that depends on Y. Its value, when $Y = y$, is $\mathbb{E}[X | Y = y]$.

• **Example:** Let $N$ be the number you get when you roll a die. You roll a fair coin $N$ times and get $H$ heads. Find $\mathbb{E}[H|N]$.

\[
\mathbb{E}[H|N = n] = n/2.
\]

\[
\mathbb{E}[H|N] = N/2.
\]
Law of total expectation: compact form

• **Lemma.** For any two random variables $X$ and $Y$,

$$
\mathbb{E}[X] = \sum_y \mathbb{E}[X \mid Y = y] \Pr[Y = y].
$$

• In other words,

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

• **Example:** Let $N$ be the number you get when you roll a die. You roll a fair coin $N$ times and get $H$ heads. Find $\mathbb{E}[H]$.

$$
\mathbb{E}[H] = \mathbb{E}[\mathbb{E}[H \mid N]] = \mathbb{E} \left[ \frac{N}{2} \right] = \frac{3.5}{2} = 1.75.
$$
Law of total expectation: application

Branching Process: A program $P$ tosses $n$ coins with bias $p$ and calls itself recursively for every HEADS.

If we call $P$ once, what is total expected number of calls to $P$ that will be generated?
Branching process

Idea: Consider ``generations” of calls

• Original call is generation 0.
• A recursive call is generation \( i \) if it was called by a call of generation \( i - 1 \).

Random variables: \( Y_i \) = # of recursive calls of generation \( i \), for \( i = 1, 2, \ldots \)

Need to find: \( \mathbb{E}[Y] \), where \( Y = \sum_{i=0}^{\infty} Y_i \).

\( Y_0 = 1 \)

\( Y_1 \sim Bin(n, p) \), so \( \mathbb{E}[Y_1] = \)

By compact form of law of total expectation, \( \mathbb{E}[Y_2] = \mathbb{E}[\mathbb{E}[Y_2|Y_1]] \)

\( \mathbb{E}[Y_2|Y_1 = y_1] = \mathbb{E}[Bin(\quad , \quad)] = \)

\( \mathbb{E}[Y_2|Y_1] = \)

\( \mathbb{E}[Y_2] = \)

Similarly, \( \mathbb{E}[Y_i] = \mathbb{E}[\mathbb{E}[Y_i|Y_{i-1}]] = \)

By linearity of expectation, \( \mathbb{E}[Y] = \sum_{i=0}^{\infty} \mathbb{E}[Y_i] = \sum_{i=0}^{\infty} (np)^i = \left\{ \begin{array}{ll} \frac{1}{1 - np} & \text{if } np < 1 \\ \text{unbounded o.w.} & \end{array} \right. \)
• A geometric random variable with parameter $p$, denoted $\text{Geom}(p)$, is the number of tosses of a coin with bias $p$ until it lands on HEADS.

• **Lemma.** The probability distribution of $X = \text{Geom}(p)$ is

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

for all $n = 1, 2, \ldots$.

• **Lemma.** The expectation of $X = \text{Geom}(p)$ is

\[ \mathbb{E}[X] = 1/p. \]
**Lemma.** The expectation of $X = \text{Geom}(p)$ is $\mathbb{E}[X] = 1/p$.

**Proof:**