Randomness in Computing

LECTURE 5

CS 537

Last time

- Random variables
- Linearity of expectation
- Jensen's inequality

Today

- Bernoulli and binomial RVs
- Conditional expectation
- Branching process
- Geometric RVs

Bernoulli random variables

- A Bernoulli random variable with parameter p:

$$\begin{cases} 1 & \text{with probability } p; \\ 0 & \text{otherwise.} \end{cases}$$

- The expectation of a Bernoulli R.V. X is

$$\mathbb{E}[X] = p \cdot 1 + (1-p) \cdot 0 = p.$$

The distribution of a Bernoulli random variable is called the Bernoulli distribution.

Binomial random variables

The binomial distribution with parameters n and p, denoted $\text{Bin}(n, p)$, is the distribution of the number of HEADS in n tosses of a coin with bias p.

A random variable $\mathrm{X} \sim \text{Bin}(n, p)$ is a binomial R.V.

Notation $\sim$: ``*has probability distribution*'' or ``*is distributed according to*''

- Lemma. The probability distribution of $\mathrm{X} \sim \text{Bin}(n, p)$ is

$$\Pr[X = j] = \binom{n}{j} p^j (1-p)^{n-j} \qquad \text{for all } j = 0,1, \dots, n.$$

- Lemma. The expectation of $X \sim \text{Bin}(n, p)$ is

$$\mathbb{E}[X] = np.$$

Review Question

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 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. Bin (n, m)

B. Bin $(m, 1/n)$

C. Bin $\left(m, \frac{n-1}{n}\right)$

D. a binomial distribution, but none of the above.

E. not a binomial distribution.

Review Question

Throw m balls into n bins. Let Y be the number of empty bins.

Compute $\mathbb{E}[\mathrm{Y}]$.

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.

Review question

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

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

B. $\mathbb{E}[a\mathrm{X}+b\mathrm{Y}] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for all $a, b \in \mathbb{R}$.

C. $\mathbb{E}[\mathrm{XY}] = \mathbb{E}[X]\ \mathbb{E}[Y]$

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 $\mathbb{E}[XY]$.

Conditional expectation: definition

- 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 $\mathbb{E}[X_1 \mid X = 5]$

Linearity of conditional expectation

- 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 $c \in \mathbb{R}$,

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

Exercise

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 $\mathbb{E}[X \mid A]$

CS 537 Recall: Law of total probability

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

- 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 Ω and let $E_1, \ldots, E_n$ be mutually disjoint events whose union is Ω.

- Law of total expectation:

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

- Notable special case of law of total expectation.
  For any two random variables X and Y,

$$\mathbb{E}[X] = \sum_{y \in Range(Y)} \mathbb{E}[X \mid Y = y] \Pr[Y = y] .$$

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

- Recall: For any two random variables X and Y,

$$\mathbb{E}[X] = \sum_{y \in Range(Y)} \mathbb{E}[X \mid Y = y] \Pr[Y = y].$$

- In other words,

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

- 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 $\mathbb{E}[H]$.

$$\mathbb{E}[H] = \mathbb{E}\big[\mathbb{E}[H|N]\big] = \mathbb{E}\left[\frac{N}{2}\right] = \frac{3.5}{2} = 1.75.$$

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, $\mathbb{E}[Y] = \sum_{i=0}^{\infty} \mathbb{E}[Y_i] = \sum_{i=0}^{\infty}(np)^i = \begin{cases} \frac{1}{1-np} & \text{if } np < 1 \\ \text{unbounded } o.w. \end{cases}$

Geometric random variables

The geometric distribution with parameter p, denoted Geom(p), is the distribution of the number of tosses of a coin with bias p until it shows HEADS.

- Lemma. The probability distribution of $X \sim$Geom(p) is
$$\Pr[X = n] = (1 - p)^{n-1} p$$
for all $n = 1,2, \ldots$.
- Lemma. The expectation of $X \sim$Geom(p) is
$$\mathbb{E}[X] = 1/p.$$