Discussion 5C

Tarang Srivastava - CS70 Summer 2020

Mini Review

Lecture Highlights

$P[X=i]$

Random Variable

The function $X(\cdot)$ is not random, not a ~~variable~~.

rant

$X \sim \text{Bernoulli}(p)$

$X \sim \text{Bernoulli}(p)$. This is also called an indicator variable

Future

Gaussian (μ, σ)

X: a random variable that takes $\{0,1\}$

$$X = \begin{cases} 1 \\ 0 \end{cases}, \quad Pr[X=i] = \begin{cases} p & \text{if } i=1 \\ 1-p & \text{if } i=0 \end{cases}$$

$E[X]$

binomial

$X \sim \text{Binomial}(p)$

Biased coin

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

X heads - p

Tails - 1-p

HHH TTT...TT

Flip it 10 times

$$P[X=3] = \binom{n}{k} p^3 (1-p)^{10-3}$$

"weighted average"

$$E[X] = \sum_{a \in A} a \times \Pr[X = a]$$

1\$ Heads

0\$ Tails

$$E[X] = \sum_{a \in A} a \times P[X=a]$$

\$10

Linearity of Expectations

$X =$

$$E[c_1 X_1 + \cdots + c_n X_n] = c_1 E[X_1] + \cdots + c_n E[X_n]$$

POWERFUL

anything

goal

understand this by today

Expectation of Binomial Distribution:

$$\boxed{E[X] = np}$$

Lecture Highlights

Lecture Highlights

Plan:
Do question 1, 3 and 4. Visit 2 if time permitting

Schedule a 1-on-1 session if you are finding these discussion questions too difficult.

Question 1

Question 1: Head Count

Consider a coin with $\mathbb{P}(\text{Heads}) = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(a) Name the distribution of X and what its parameters are.

(b) What is $\mathbb{P}(X = 7)$?

(c) What is $\mathbb{P}(X \geq 1)$? Hint: You should be able to do this without a summation.

(d) What is $\mathbb{P}(12 \leq X \leq 14)$? ← you'll need some sort of summation

a) Binomial(20, 2/5)

$$P(X=k) = \binom{n}{k}\left(\frac{2}{5}\right)^k\left(1-\frac{2}{5}\right)^{n-k}$$

Question 1: Head Count

Consider a coin with $\mathbb{P}(\text{Heads}) = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(a) Name the distribution of X and what its parameters are.

(a) Since we have 20 independent trials, with each trial having a probability $2/5$ of success, $X \sim \text{Binomial}(20, 2/5)$.

Question 1: Head Count

Consider a coin with $\mathbb{P}(\text{Heads}) = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(b) What is $\mathbb{P}(X = 7)$?

$$P(X=7) = \binom{20}{7} \frac{2}{5}^{7} \left(1-\frac{2}{5}\right)^{20-7}$$

$$\mathbb{P}(X = 7) = \binom{20}{7} \left(\frac{2}{5}\right)^{7} \left(\frac{3}{5}\right)^{13}.$$

Question 1: Head Count

Consider a coin with $\mathbb{P}(\text{Heads}) = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(c) What is $\mathbb{P}(X \geq 1)$? Hint: You should be able to do this without a summation.

$$\sum_{i=0}^{n} P(x=i) = 1$$

$$1 - P(X=0) = \sum_{i=1}^{n} P(X=i)$$

$$\left(\frac{3}{5}\right)^{20}$$

$$\cancel{\binom{20}{0}\left(\frac{2}{5}\right)^{0}}\ \boxed{\left(\frac{3}{5}\right)^{20}}$$

1

0.6

$(0.6)^2 = 0.36$

95% success

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \left(\frac{3}{5}\right)^{20}.$$

Question 1: Head Count

Consider a coin with $\mathbb{P}(\text{Heads}) = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(d) What is $\mathbb{P}(12 \leq X \leq 14)$?

$$= P(X=12) + P(X=13) + P(X=14)$$

exactly

$$\sum_{i=12}^{14} P(X=i)$$

$$= \binom{20}{i}\left(\frac{2}{5}\right)^{i}\left(\frac{3}{5}\right)^{n-i}$$

$$\begin{aligned}\mathbb{P}(12 \leq X \leq 14) &= \mathbb{P}(X=12)+\mathbb{P}(X=13)+\mathbb{P}(X=14) \\ &= \binom{20}{12}\left(\frac{2}{5}\right)^{12}\left(\frac{3}{5}\right)^{8}+\binom{20}{13}\left(\frac{2}{5}\right)^{13}\left(\frac{3}{5}\right)^{7}+\binom{20}{14}\left(\frac{2}{5}\right)^{14}\left(\frac{3}{5}\right)^{6}.\end{aligned}$$

Question 2

Question 2: Family Planning

Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

(a) Determine the sample space, along with the probability of each sample point.

(b) Compute the joint distribution of G and C. Fill in the table below.

	$C=1$	$C=2$	$C=3$
$G=0$			
$G=1$			

(c) Use the joint distribution to compute the marginal distributions of G and C and confirm that the values are as you'd expect. Fill in the tables below.

$\mathbb{P}(G=0)$	
$\mathbb{P}(G=1)$	

$\mathbb{P}(C=1)$	$\mathbb{P}(C=2)$	$\mathbb{P}(C=3)$

(d) Are G and C independent?

(e) What is the expected number of girls the Browns will have? What is the expected number of children that the Browns will have?

Question 2: Family Planning

Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

(a) Determine the sample space, along with the probability of each sample point.

Question 2: Family Planning

Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

(b) Compute the joint distribution of G and C. Fill in the table below.

	$C=1$	$C=2$	$C=3$
$G=0$			
$G=1$			

Question 2: Family Planning

Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

	$C=1$	$C=2$	$C=3$
$G=0$	0	0	$\mathbb{P}(bbb)=1/8$
$G=1$	$\mathbb{P}(g)=1/2$	$\mathbb{P}(bg)=1/4$	$\mathbb{P}(bbg)=1/8$

[redacted]

[redacted]

Question 2: Family Planning

Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

(d) Are G and C independent?

[redacted]

Question 2: Family Planning

Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

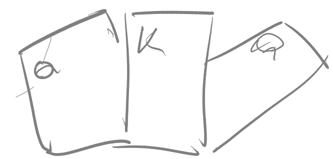
(e) What is the expected number of girls the Browns will have? What is the expected number of children that the Browns will have?

[redacted]

Question 3

Question 3: How Many Queens?

You shuffle a standard 52-card deck, before drawing the first three cards from the top of the pile. Let X denote the number of queens you draw.



(a) What is $\mathbb{P}(X = 0)$, $\mathbb{P}(X = 1)$, $\mathbb{P}(X = 2)$ and $\mathbb{P}(X = 3)$?

(b) What do your answers you computed in part a add up to?

(c) Compute $\mathbb{E}(X)$ from the definition of expectation.

(d) Suppose we define indicators X_i, $1 \leq i \leq 3$, where X_i is the indicator variable that equals 1 if the ith card is a queen and 0 otherwise. Compute $\mathbb{E}(X)$ using linearity of expectation.

(e) Are the X_i indicators independent? Does this affect your solution to part (d)?

Question 3: How Many Queens?

You shuffle a standard 52-card deck, before drawing the first three cards from the top of the pile. Let X denote the number of queens you draw.

(a) What is $\mathbb{P}(X = 0)$, $\mathbb{P}(X = 1)$, $\mathbb{P}(X = 2)$ and $\mathbb{P}(X = 3)$?

- Each hand equally likely

$$\Omega = \binom{52}{3} \leftarrow \text{TAR}$$

- order of the three cards doesn't matter

$$P(X=0) = \frac{\binom{48}{3}}{\binom{52}{3}} \quad \leftarrow \text{not queens}$$

$$P(X=1) = \frac{\binom{4}{1}\binom{48}{2}}{\binom{52}{3}}$$

$$P(X=1) =$$

Alternatively, every 3-card hand is equally likely, so we can use counting. There are $\binom{52}{3}$ total 3-card hands, and $\binom{48}{3}$ hands with only non-queen cards, which gives us the same result.

$$\mathbb{P}(X = 0) = \frac{\binom{48}{3}}{\binom{52}{3}} = \frac{4324}{5525}$$

- We will continue to use counting. The number of hands with exactly one queen amounts to the number of ways to choose 1 queen out of 4, and 2 non-queens out of 48.

$$\mathbb{P}(X = 1) = \frac{\binom{4}{1}\binom{48}{2}}{\binom{52}{3}} = \frac{1128}{5525}$$

- Choose 2 queens out of 4, and 1 non-queen out of 48.

$$\mathbb{P}(X = 2) = \frac{\binom{4}{2}\binom{48}{1}}{\binom{52}{3}} = \frac{72}{5525}$$

- Choose 3 queens out of 4.

$$\mathbb{P}(X = 3) = \frac{\binom{4}{3}}{\binom{52}{3}} = \frac{1}{5525}$$

Question 3: How Many Queens?

You shuffle a standard 52-card deck, before drawing the first three cards from the top of the pile. Let X denote the number of queens you draw.

(b) What do your answers you computed in part a add up to?

X = # of Quee

$$\mathbb{P}(X=0)+\mathbb{P}(X=1)+\mathbb{P}(X=2)+\mathbb{P}(X=3)=\frac{4324+1128+72+1}{5525}=1$$

Question 3: How Many Queens?

You shuffle a standard 52-card deck, before drawing the first three cards from the top of the pile. Let X denote the number of queens you draw.

(c) Compute $\mathbb{E}(X)$ from the definition of expectation.

part (a)

$$E[X] = \sum_{a \in A} a \cdot P(X=a)$$

$$0 \cdot P(X=0) + 1 \cdot P(X=1) \ldots + 3 \cdot P(X=3)$$

part (a)

(c) From the definition, $\mathbb{E}(X) = \sum_{k=0}^{3} k\mathbb{P}(X=k)$, so

$$\mathbb{E}(X) = 0 \cdot \frac{4324}{5525} + 1 \cdot \frac{1128}{5525} + 2 \cdot \frac{72}{5525} + 3 \cdot \frac{1}{5525} = \frac{3}{13}.$$

suspicous

Question 3: How Many Queens?

You shuffle a standard 52-card deck, before drawing the first three cards from the top of the pile. Let X denote the number of queens you draw.

(d) Suppose we define indicators X_i, $1 \le i \le 3$, where X_i is the indicator variable that equals 1 if the ith card is a queen and 0 otherwise. Compute $\mathbb{E}(X)$ using linearity of expectation.

X_1 X_2 X_3

1 2 3

Q NQ Q

$X = X_1 + X_2 + X_3$

$2 = 1 + 0 + 1$

$E[X] = E[X_1] + E[X_2] + E[X_3]$

Indicators $\rightarrow$ $E[X_1] = 0 \cdot P(X = \text{not queen}) + 1 \cdot P(X = \text{queen})$

$\cancel{0 \cdot \frac{48}{52}}$ $\frac{12}{13}$ 0 $\frac{4}{52}$ $\frac{1}{13}$

$E[X_1] = \frac{1}{13}$

$E[X] = \frac{3}{13}$

(d) We know that $\mathbb{E}(X_i) = \mathbb{P}(\text{card } i \text{ is a queen}) + 0 \cdot \mathbb{P}(\text{card } i \text{ is not a queen}) = 1/13$, so

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \mathbb{E}(X_3) = \frac{1}{13} + \frac{1}{13} + \frac{1}{13} = \frac{3}{13}.$$

Notice how much faster it was to compute the expectation using indicators!

Question 3: How Many Queens?

You shuffle a standard 52-card deck, before drawing the first three cards from the top of the pile. Let X denote the number of queens you draw.

not independent

(e) Are the X_i indicators independent? Does this affect your solution to part (d)?

~~Yes~~

No. $P(X_1 = 1, X_2 = 1) = \frac{4}{52} \cdot \frac{3}{51} =$ $\neq$

$P(X=1) \cdot P(X=2) = \frac{4}{52} \cdot \frac{4}{52}$

No, they are not independent. As an example:

$$\mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 1) = \frac{1}{13} \cdot \frac{1}{13} = \frac{1}{169}$$

However,

$$\mathbb{P}(X_1 = 1, X_2 = 1) = \mathbb{P}(\text{the first and second cards are both queens}) = \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}.$$

Even though the indicators are not independent, this does not change our answer for part (g). Linearity of expectation *always* holds, which makes it an extremely powerful tool.

Question 4

Question 4: Linearity

(a) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $1/3$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $1/5$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

(b) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears?

Question 4: Linearity

(a) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $1/3$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $1/5$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

Question 4: Linearity

(b) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears?

WHY use indicator variables!

b o o k

b o o k

b o o k

$$\boxed{\frac{1}{26^4}}$$

$$\frac{1}{26} \cdot \frac{1}{26} \cdot \frac{1}{26} \cdot \frac{1}{26}$$

$$E[A_1] = \cancel{0 \cdot P(A_1 = \text{not book})} + 1 \cdot P(A_1 = \text{Book})$$

$$1 \cdot \frac{1}{26^4}$$

Bernoulli

$$E[x] = p$$

always true

There are $1,000,000 - 4 + 1 = \boxed{999,997}$ places where "book" can appear, each with a (non-independent) probability of $1/26^4$ of happening. If A is the random variable that tells how many times "book" appears, and A_i is the indicator variable that is 1 if "book" appears starting at the ith letter, then

$$\begin{aligned} \mathbb{E}[A] &= \mathbb{E}[A_1 + \cdots + A_{999,997}] \\ &= \mathbb{E}[A_1] + \cdots + \mathbb{E}[A_{999,997}] \\ &= \frac{999,997}{26^4} \approx 2.19. \end{aligned}$$

Powerful