Random Variables

$$\Omega = \{(H,H),\ (H,T),\ (T,H),\ (T,T)\}$$

$X =$	2	1	1	0
$Y =$	-2	0	0	2

X # of Heads

Y # of T − # H

$Z = |Y|$

9

Independence and Random Variables

Computer Science Mentors 70

Apr 5 - Apr 9
Spring 2021

Amogh Gupta, Sylvia Jin, Aekus Bhathal, Abinav Routhu, Debayan Bandyopadhyay

Roast us here: `https://tinyurl.com/csm70-feedback20`

1 Independence

1. For each of the following examples, decide whether the listed events are mutually independent, pairwise independent, or neither (if the events are mutually independent, there is no need to also select pairwise independent). Recall the definition of independence–is the probability of both events equal to the product of the probabilities of each event? Should one event influence the probability of the other?

   (a) The event of drawing a jack of hearts from the deck and the event of drawing a jack of clubs from the same deck.

   - Mutually Independent
   - Only Pairwise Independent
   - Neither

   $\frac{1}{52}$ $\frac{1}{51}$ $\frac{1}{52}$

   X_1 X_2 X_3 X_4

   $X_1 \perp\!\!\!\perp X_2$ $X_1 \perp\!\!\!\perp X_3$ $X_2 \perp\!\!\!\perp X_3$

   (b) The event of drawing a jack of hearts from the deck and the event of drawing an ace of diamonds from the same deck.

   - Mutually Independent
   - Only Pairwise Independent
   - Neither

   (c) The outcomes of three consecutive coinflips.

   - Mutually Independent
   - Only Pairwise Independent
   - Neither

   (d) Given 2 random integers x, y, the event that $\underset{A}{x \equiv 5 \pmod n}$, the event that $\underset{B}{y \equiv 7 \pmod n}$, and the event that $\underset{C}{x + y \equiv 20 \pmod n}$.

   - Mutually Independent
   - Only Pairwise Independent
   - Neither

   $Pr[A] = \frac{1}{n}$

   $Pr[B] = \frac{1}{n}$

   $Pr[C] = \frac{1}{n}$

   $Pr[C \mid A] = \frac{1}{n}$ $\quad Pr[C \mid B] = \frac{1}{n}$

   $Pr[C \mid A, B] = 0$

   $5 + y \equiv 20 \pmod n$

   $\Rightarrow y \equiv 15 \pmod n$

2 Random Variables

1. **Intro to Random Variables**

   Suppose that we are flipping 3 coins in a row. Let's try to define some different random variables relating to this process!

   (a) Define the random variable R_1 to represent whether the last coinflip was a $\underset{0}{\text{head}}$ or a $\underset{1}{\text{tail}}$. Draw a mapping from the sample space to the value of R_1 that each event corresponds with.

   $R_1 \in \{0, 1\}$

   $\Omega =$

   HHH → 0
   HTH → 0
   THH → 0
   TTH → 0

   HHT → 1
   HTT → 1
   THT → 1
   TTT → 1

(b) Now, define the random variable R_2 to represent the number of heads seen in our series of flips. What are possible values that R_2 can take on? Draw a mapping from the sample space to the values of R_2 that each event corresponds with.

R_2

HHH → 3
HTH → 2
THH → 2
TTH → 1
HHT → 2
HTT → 1
THT → 1
TTT → 0

$\underline{R_2 = 2}$

{HTH, THH, HHT}

2. **Dice Division** (practice.eecs.org Set 11 Problem 1)
Consider the following game: you roll two standard 6-sided dice, one after the other. If the number on the first dice divides the number on the second dice, you get 1 point. You get 1 additional point for each prime number you roll.

Define the random variable R_1 to be the result of the first roll, and define R_2 to be the result of the second roll. Define the random variable $X = R_1 + R_2$ to be the sum of the numbers that come up on both dice, define the random variable $Y = R_1 \cdot R_2$ to be the product of the numbers that come up on both dice, and define the random variable Z to be the number of points you win in the game.

(a) What values can the random variable Z take on (with nonzero probability)? Give examples of an event that would cause Z to take on that value.

0 (4, 6)
1 (3, 4)
2 (3, 5) (2, 4)
3 (3, 3) (5, 5)

(b) Say that your first roll is a 3 and your second roll is a 6. What is the value of Z?

$1 + 1 = 2$

(c) Say that your first roll is a 4 and your second roll is a 1. What is the value of $X^2 + Y + Z$?

$X = 5$
$Y = 4$
$Z = 0$

$X^2 + Y + Z = 25 + 4 + 0 = 29$

(d) Conditioned on the fact that your second roll is a 1, what is the probability that $Z = 1$?

$\Pr[Z = 1 \mid R_2 = 1] = 4/6 = \frac{2}{3}$

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

(e) Conditioned on the fact that your ~~second~~ first roll is a 1, what is the probability that $Z = 2$?

$\Pr[Z = 2 \mid R_2 = 1] = 0$

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

$\Pr[Z = 2 \mid R_1 = 1] = \frac{3}{6} = \frac{1}{2}$

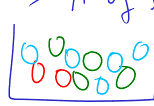
1, 1 1, 4
1, 2 1, 5
1, 3 1, 6

3. **The Binomial to the Multinomial**
A binomial is often a useful distribution when modelling the counts of independent and identically distributed (i.i.d) bernoulli trials. But how can we model trials with more than two possible outcomes?

# of successes ~ Binom(10, 0.5)

The content mentors were split on a very important issue: how to eat corn. Some of them ate the corn horizontally (think

→ # of blue balls



→ # of red
→ # of green

typewriter-style), while others ate the corn in spirals from left to right. To get to the bottom of who was right, they decided to hold a survey of the CS70 student population. But because they were too lazy to create a google form, they replaced each student with a biased coin with probability p of landing heads and flipped it. They then made the arbitrary rule that heads meant the student ate the corn horizontally and tails meant the student ate the corn in spirals.

(a) Suppose there are n students, find the probability that k of them ate the corn horizontally.

X is # of students who ate horizontally.

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

(b) A rumour was recently spread that some people ate corn in a double-helix pattern (like a DNA strand). In the spirit of the scientific method, the content mentors decided they should add this options to their test. But since they couldn't find an object with three faces, they had to code up a simulation in python with probability p of the student eating the corn horizontally, probability q of the student eating the corn in spirals, and probability r of the student eating the corn in a double-helix pattern such that $p+q+r=1$. Find the probability that k student ate the corn in a double-helix pattern.

(c) But the content mentors are hungry for data, so they want to know something more specific: what is the probability that k students ate the corn in a double-helix pattern **and** m student ate the corn in spirals. Simplify your expression.

(b) $Pr[Z=k] = \binom{n}{k} r^k (1-r)^{n-k}$

(c) $n-k-m$ horizontally

$k+m$

$Pr[Z=k \cap Y=m]$

$\binom{n}{n-k-m} p^{n-k-m} (1-p)^{k+m}$

$= Pr[Z=k]\, Pr[Y=m \mid Z=k]$

$= \binom{n}{k} r^k (1-r)^{n-k} \binom{n-k}{m} \left(\frac{q}{p+q}\right)^m \left(1-\frac{q}{p-q}\right)^{n-k-m} = \frac{n!}{k!\,m!\,(n-k-m)!} r^k q^m p^{n-k-m}$

3 Expectation

1. **Introduction to Expectation**

Imagine that Sylvia has two, 3-sided loaded (non-uniform probability) dice, which are represented by the random variables X and Y, respectively. X and Y are distributed as follows:

$$P(X=1) = \frac{1}{2} \qquad P(Y=1) = \frac{1}{6}$$
$$P(X=2) = \frac{1}{4} \qquad P(Y=2) = \frac{1}{6}$$
$$P(X=3) = \frac{1}{4} \qquad P(Y=3) = \frac{2}{3}$$

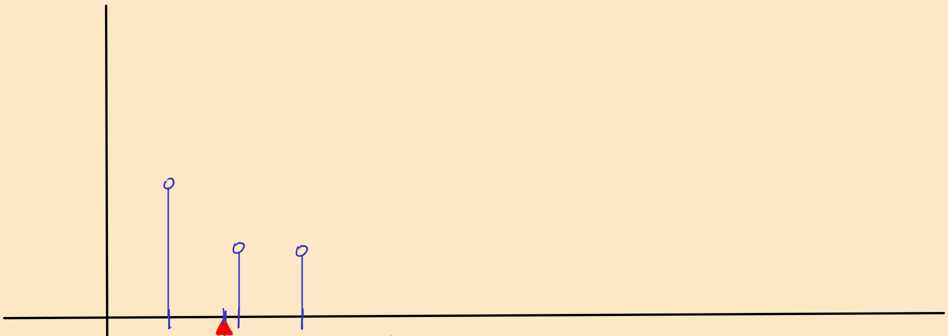
$\sum_x x \cdot Pr[X=x]$

(a) Sylvia rolls the first die, represented by random variable X.What is the expected value of the roll of the first die. What is the probability it will roll the expected value?

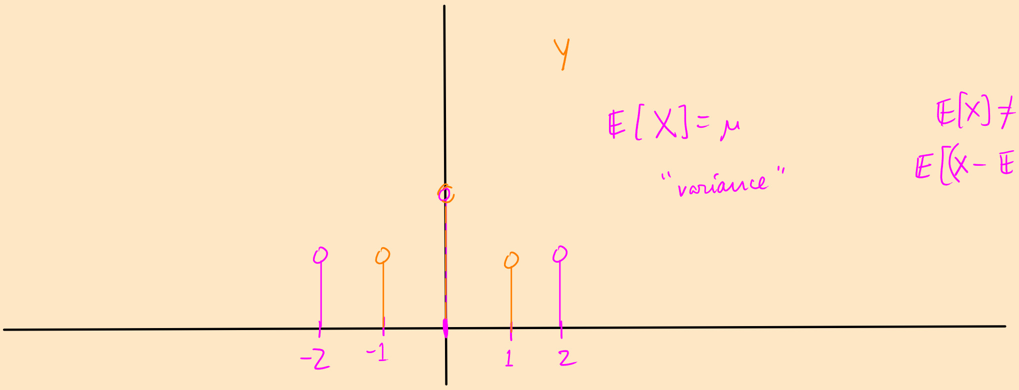
$E[X] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} = \frac{7}{4}$

(b) What is expected value of a roll of the second die, represented by random variable Y?

$\frac{1}{6} + \frac{2}{6} + \frac{6}{3}$



$E[X] = 1.75 = \frac{7}{4}$



$\mathbb{E}[X] = \mu$

"variance"

$\mathbb{E}[X] \neq 0$

$\mathbb{E}[(X - \mathbb{E}[X])^2] = 0$

Y	$(Y - \mathbb{E}[Y])^2$	$Pr[Y=y]$
1	1	1/4
0	0	1/2
-1	1	1/4

$\mathbb{E}[(Y - \mathbb{E}[Y])^2] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

X	$(X - \mathbb{E}[X])^2$	$Pr[X=x]$
2	4	1/4
0	0	1/2
-2	4	1/4

$\mathbb{E}[(X - \mathbb{E}[X])^2] = 1 + 1 = 2$

→ -1 -1 1 1 0 0 0 0 1 0 1 -1

→ 0 -2 2 2 -2 0 2 0 -2 - - -

Q. What's Var of the # of H in 8 coin flips?

$X \sim Binom(8, 1/2)$

$Var(X) = np(1-p) = 8 \cdot \frac{1}{2} \cdot \frac{1}{2} = 2$

$X \sim Binom(n, p)$

$X = \sum_{i=1}^{n} X_i$

$X_i \sim Bernoulli(p)$ $\quad X_i = \begin{cases} 0 & \text{w.p. } 1-p \\ 1 & \text{w.p. } p \end{cases}$

X_i's are iid

$\forall i, \; Var(X_i) = p - p^2 = p(1-p)$

$= \mathbb{E}[X_i^2] - \mathbb{E}^2[X_i]$

$Var(X)$

$= Var\left(\sum_{i=1}^{n} X_i\right)$

bc X_i's are indep

$= \sum_{i=1}^{n} Var(X_i) = np(1-p)$

Z	$Pr[Z=z]$
1	1/12
2	3/24
3	(1/3 + 1/24)
4	
6	
9	

(c) What is the expected value of the product of the two dice?

X, Y $\quad Z = XY$

$$\mathbb{E}[Z] = \mathbb{E}[XY] = \sum z \cdot Pr[Z=z]$$

2. Assume we have 10 coins, each with a different bias towards heads. The first coin has $p = 0.1$ of flipping heads, the second has $p = 0.2$, etc up to the 10th coin which has $p = 1$. What is the expected number of heads from flipping all 10 coins at once?

3. **Binomial Mean**
Show that mean of a Binomial random variable X with parameters n and p has mean np.

4 Indicator Variables

4. **Introduction to Indicators**
Imagine drawing a poker hand of five cards from a deck of cards. What is the expected number of face cards that we get?

$X_i = \begin{cases} 0 & o/w \\ 1 & \text{if smth happens} \end{cases}$ $\quad X = \sum X_i$ $\quad J, K, Q$

$\mathbb{E}[X_i] = Pr[X_i = 1] = \frac{12}{52}$

# of face cards = first card FC? + 2^{nd} card FC? + . . .

count! X $\quad X_i = \begin{cases} 0 & o/w \\ 1 & \text{if } i^{th} \text{ card FC} \end{cases}$ $\quad X = \sum_{i=1}^{5} X_i$ $\quad \mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{5} X_i]$

$= \sum_{i=1}^{5} \mathbb{E}[X_i] = \boxed{5 \cdot \frac{12}{52}}$

5. **Random Homework Party**
N for $N \geq 3$ people join a Homework party. After joining each person is assigned to a breakout rooms randomly. Each assignment was equally likely to be one the $m \geq 2$ breakout rooms and independent of all other assignments. This helps the course staff and students in the way that in each breakout room, every pair of people could help each other. Course staff would like to find the expected number of pairs who are in the the same room.

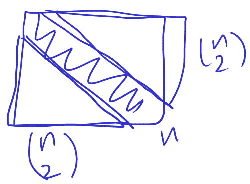
(a) Let the random variable X be the number of pairs of people who are in the same breakout room. Lets model this using indicator variables. Come up with a model for indicator variable for the pair of people without over counting them. Are the random variables mutually independent ?

$\frac{m}{m^2} = \frac{1}{m}$ $\quad i,j$ $\quad i \neq j$ $\quad X_{ij}$ $\quad i < j$

$$X = \sum_{1 \leq i < j \leq N} X_{i,j}$$

(b) Let the random variable X be the number of pairs of people who are in the same breakout room. What is the expectation E[X]

$$\mathbb{E}[X] = \sum_{1 \le i < j \le N} \mathbb{E}[X_{i,j}] = \binom{N}{2} \frac{1}{m}$$



$$n + \frac{n(n-1)}{2} \cdot 2$$

$$= n^2$$

5 Joint Distribution

1. You have a waste-disposal chute that trash is falling through. At the end of the chute, trash either falls into the right bin with probability 3/4, or the left bin with probability 1/4. The right bin can hold 3 pieces of trash, while the left bin can only hold 1 piece. Being the sanitary student that you are, if you notice that a bin is filled, you block off the chute so that no more trash can fall through the chute.

   For this problem, let R be the pieces that fall into the right bin, and L be the number of pieces that fall into the left bin.

1. Determine the sample space (all possible outcomes of the falling trash), along with the probability of each sample point.

   $R \in \{0,1,2,3\}$
   $L \in \{0,1\}$

   $\{\underset{l}{(1,0)}, \underset{rl}{(1,1)}, \underset{rrl}{(1,2)}, \underset{rrr}{(0,3)}\} \subset \{0,1\} \times \{0,1,2,3\}$

2. Compute the Joint Distribution of L and R. Fill in the table:

   $(3/4)^3$

	R=0	R=1	R=2	R=3	
L=0	0	0	0	27/64	27/64
L=1	16/64	12/64	9/64	0	37/64
	16/64	12/64	9/64	27/64	

3. Use the Joint Distribution to compute the Marginal Distribution of L and R. Fill out the tables.

P(L=0)	P(L=1)
27/64	37/64

P(R=0)	P(R=1)	P(R=2)	P(R=3)
16/64	12/64	9/64	27/64

4. Are L and R independent?

   No. Joint is NOT the product of marginals. $\Pr[L=\ell \cap R=r] \neq \Pr[L=\ell]\Pr[R=r]$

5. What is the total expected number of pieces of trash that will fall before you have to empty the bins?

   $$1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{16} + 3 \cdot \left(\frac{9}{64} + \frac{27}{64}\right) ✓$$