9

Conditional Probability, Principle of Inclusion/Exclusion, and Independence

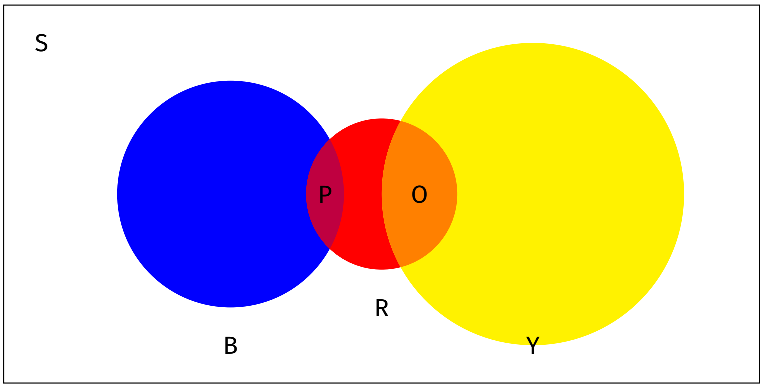
Computer Science Mentors 70

Nov 2 - Nov 6
Fall 2020

Prepared by: Aishani Sil, Austin Lei, Agnibho Roy, Debayan Bandyopadhyay, Abinav Routhu

Probability Pictures

1. (Practice Bank Problem 2) Consider the drawing of the probability space S below. Here, the blue/purple region is the set of events B, the red/purple/orange region is the set of events R, and the yellow/orange region is the set of events Y. The set of events P is the set of events in both B and R, and is represented by the purple region. The set of events O is the set of events in both R and Y, and is represented by the orange region.



$\frac{P[R \cap Y]}{P[Y]}$

Assume that we are sampling from S uniformly at random. Please answer the foll owing enumerateple choice questions about this space, selecting all that apply.

(a) What is $P[R]$, the probability that an element from S is in R?

Hint: Recall the definition of probability–the probability of an event X is the number of outcomes for which X occurs divided by the total number of outcomes.

i. $\frac{|R|-|P|-|O|}{|S|}$

ii. $\frac{|R|}{|S|}$

iii. $|R|$

(b) What is $P[R|Y]$, the probability that an element from S is in R given that it is also in Y?

Hint: Recall the definition of probability–the probability of an event X is the number of outcomes for which X occurs divided by the total number of outcomes. When we condition on an event Z, we are restricting the total outcomes to be outcomes for which Z occurs.

i. $\frac{|O|}{|S|}$

ii. $\frac{|O|}{|Y|}$

iii. $\frac{|R \cap Y|}{|Y|}$

(c) What is $P[R|O]$, the probability that an element from S is in R given that it is in O?

Hint: Recall the definition of probability–the probability of an event X is the number of outcomes for which X occurs divided by the total number of outcomes. When we condition on an event Z, we are restricting the total outcomes to be outcomes for which Z occurs.

i. 1

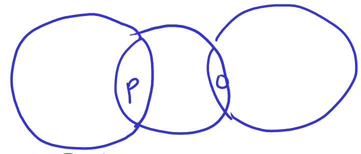
ii. $\frac{|R \cap O|}{|O|} = \frac{|O|}{|O|}$ $\quad R \cap O = O$

iii. $\frac{|R \cap O|}{|S|}$

(d) What is $P[B \cup R \cup Y]$, the probability that an element of S is in B or R or Y?

Hint: Recall the definition of probability–the probability of an event X is the number of outcomes for which X occurs divided by the total number of outcomes. Be careful not to double count!

i. $\frac{|B|+|R|+|Y|-|P|-|O|}{|S|}$



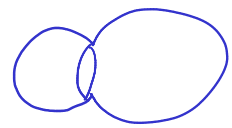
ii. $\frac{|B \setminus P|+|R|+|Y \setminus O|}{|S|}$ $\quad (X \setminus Z = \{x \ s.t. \ x \in X \text{ and } x \notin Z\})$

iii. $\frac{|B|+|R|+|O|}{|S|}$ $\quad B \setminus R = B \setminus P \quad Y \setminus R = Y \setminus O$

(e) What is $P[O|R \cup Y]$, the probability that an element of S is in O given that it is also in R or Y?

Hint: Recall the definition of probability–the probability of an event X is the number of outcomes for which X occurs divided by the total number of outcomes. When we condition on an event Z, we are restricting the total outcomes to be outcomes for which Z occurs. Be careful not to double count!

i. $\frac{|O|}{|S|-|B|}$ $\quad O \cap (R \cup Y) = O$



ii. $\frac{|O|}{|Y|+|R|-|O|}$

iii. $\frac{|O|}{|Y|+|R|}$

Mixed Probability

1. **Generating Graphs**

Imagine that we have a graph on n unique vertices. Consider a random generation of undirected edges on this graph

(a) What is the probability that the graph is complete?

(b) What is the probability that the graph is a complete bipartite graph? (A complete bipartite graph is a graph whose vertices can be partitioned into two sets U and V, such that there is an edge between two vertices if and only if one vertex is from U and the other vertex is from V.)

2. **Playlist**

Leanne has 5 songs on her playlist, and a song is chosen uniformly at random to be played. If her favorite song is played, she plays it again with 100%. Otherwise, she chooses her next song uniformly at random from all songs on her playlist.

(a) What is the probability that Leanne's favorite song was played twice?

(b) What is the probability that the second song was Leanne's favorite song?

(c) What is the probability that the first song was Leanne's favorite, given that the second song was Leanne's favorite?

3. **More Dice**

A fair die is rolled once. Given that the number rolled is even, what is the probability that it is prime?

$\frac{1}{3}$ 1 2 3 4 5 6

even

prime

4. **Go Bears!**

Oski the bear has lost his dog in either forest A (with prior probability 0.4) or in forest B (with prior probability 0.6).
On any given day, if the dog is in A and Oski spends a day searching for it in A, the probability that he will find the dog that day is 0.25. Similarly, if the dog is in B and Oski spends a day looking for it there, the probability that he will find the dog that day is 0.15.
The dog cannot go from one forest to the other. Oski can search only in the daytime, and he can travel from one forest to the other only at night.

(a) In which forest should Oski look to maximize the probability he finds his dog on the first day of the search?

(b) Given that Oski looked in A on the first day but didn't find his dog, what is the probability that the dog is in A?

(c) If Oski flips a biased coin (search in A with probability 0.3 and search in B with probability 0.7) to determine where to look on the first day and finds the dog on the first day, what is the probability that he looked in A?

(d) If the dog is alive and not found by the first day of search, it will die that evening with probability $\frac{1}{3}$. Oski has decided to look in A for the first two days. What is the probability that he will find a live dog for the first time on the second day? (Hint: Write out all the events that must simultaneously happen for Oski to find a live dog for the first time on the second day)

5. **Seeding Players**

Imagine that we are trying to pick the first seed for a tournament of n players in a tennis tournament. We want to be efficient with choosing the first seed so we do a rough estimate of who receives it. Think about the following scheme: we line all n players up in a line and ignore the first m players. Then, we go in order starting from the $m + 1$th player and choose the first player who is better than all the previous players since the $m + 1$th player to be the first seed. What is the probability that the first seed is assigned to the best player amongst the n players? You may use the approximation that $\sum_{k=0}^{n} k^{-1} \approx \ln n$ to simplify your answer. You may not have a summation in your answer.(Inspired by an EECS126 HW Problem)

A : first seed is the best player, $B_i = \mathbb{1}\{i^{th}\text{ player is the best}\}$, $1 \le i \le n$

$\mathbb{1}\{X\} = \begin{cases} 1 & \text{if } X \\ 0 & \text{o/w} \end{cases}$

$$Pr[A] = \sum_{i=1}^{n} Pr[A \cap B_i] = \sum_{i=1}^{n} Pr[A|B_i]Pr[B_i]$$

3 4 7 8

$Pr[B_i] = \frac{1}{n}$, $1 \le i \le n$, $Pr[A|B_i] = \frac{m}{i-1}$ $\forall\, m+1 \le i \le n$, 0 o/w

$$Pr[A] = \sum_{i=1}^{n} Pr[A \cap B_i] = \sum_{i=1}^{n} Pr[A|B_i]Pr[B_i]$$

$$= \sum_{i=m+1}^{n} \frac{m}{i-1} \cdot \frac{1}{n} = \boxed{\frac{m}{n} \sum_{i=m+1}^{n} \frac{1}{i-1}} \approx \frac{m}{n}\left(\sum_{i=0}^{n} \frac{1}{i} - \sum_{i=0}^{m} \frac{1}{i}\right) \approx \frac{m}{n} \ln \frac{n}{m}$$

Independence

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.

A B

- Mutually Independent
- Only Pairwise Independent

$P[A] = \frac{1}{52}$, $P[A|B] = \frac{1}{51}$ $P[X_1|X_2] = P[X_1]$

$X_i \perp\!\!\!\perp X_j \;\forall\, i \ne j$ $X_1 \perp\!\!\!\perp X_2$ $X_1 \perp\!\!\!\perp X_3$, $X_2 \perp\!\!\!\perp X_4$

X_1, X_2, X_3, X_4 $P[X_1|X_2, X_3, X_4] = P[X_1]$

- Neither

(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.

A B

- 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 $x = 5 \mod n$, the event that $y = 7 \mod n$, and the event that $x + y = 20 \mod n$.

$\text{Unif}[0, n-1)$ A B C

- Mutually Independent
- Only Pairwise Independent
- Neither

$y = 15 \bmod n$

$P[C] > 0$

$\frac{1}{n}$

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

$x = 13 \bmod n$

$x + y = 12 \bmod n$

$\frac{1}{n}$

$x = 5 \bmod n$

Ω

HHH HHT
HTH HTT
THH THT
TTH TTT

$X(\omega) = \#$ of H in ω

$X(HH) = 2$
$X(HT) = X(TH) = 1$
$X(TT) = 0$

$X : \Omega \to \mathbb{R}$
$\uparrow$
"random variable"

How many H?

What's the prob that I have 2 heads?
" " " " $X = 2$

$\frac{3}{8}$

$E[X]$ $\quad Y$

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

$\mathbb{E}[\ ]$
$\mathbb{E}(\ )$
$E(\ )$
E

$\Pr[X=2, Y=1] = \Pr[X=2 \cap Y=1]$
$= \Pr[X=2 \wedge Y=1]$
$= \Pr[X=2 \text{ and } Y=1]$

$\mathbb{P}$ $\quad P$ $\quad P(\ )$