50.034 - Introduction to Probability and Statistics

Week 2 – Cohort Class

January–May Term, 2019

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

CC Week 2 Assigned Readings: (Revise: 2.1, 2.2, 2.3, 3.1, 3.2)

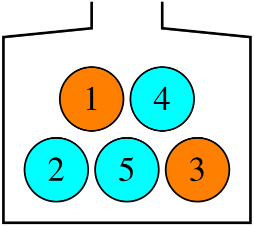
Outline of Cohort Class

Exercises on the following topics:

- Law of total probability
- Bayes' theorem
- Probability mass function (pmf)

What is the probability?

You are blindfolded, and you randomly select a ball from a jar. You have looked at all the balls in the jar before you were blindfolded.



What is the probability that you get **ball 3**?

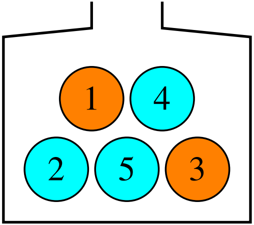
Choices: A) $\frac{2}{5}$, B) $\frac{1}{2}$, C) $\frac{1}{5}$.

Answer: C) $\frac{1}{5}$, because each ball is equally likely to be selected.

What is the probability?

You are blindfolded, and you randomly select a ball from a jar. You have looked at all the balls in the jar before you were blindfolded.

What is the probability that you get an **orange ball**?



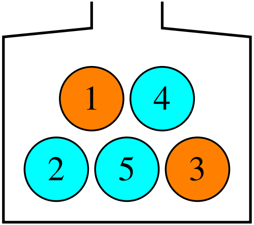
Choices: A) $\frac{2}{5}$, B) $\frac{1}{2}$, C) $\frac{1}{5}$.

Answer: A) $\frac{2}{5}$, because each ball is equally likely to be selected and there are two orange balls.

What is the probability?

You are blindfolded, and you randomly select a ball from a jar. You have looked at all the balls in the jar before you were blindfolded.

Your friend tells you that you have selected an orange ball. With this new information, what is the probability that you get **ball 3**?



Choices: A) $\frac{2}{5}$, B) $\frac{1}{2}$, C) $\frac{1}{5}$.

Answer: B) $\frac{1}{2}$, because ball 3 can only be one of the two orange balls selected. This is the **conditional probability** of having ball 3, given that the ball is orange.

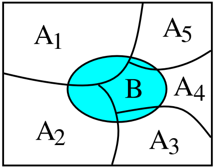
Recall: Law of total probability and Bayes theorem

Let $A_1, \ldots, A_k$ be **mutually exclusive** and **exhaustive** events in some sample space Ω.

- $A_1, \ldots, A_k$ are exhaustive if $A_1 \cup A_2 \cup \cdots \cup A_k = \Omega$.
- $A_1, \ldots, A_k$ are mutually exclusive if $A_i \cap A_j = \varnothing$ for all $i \neq j$.

The **law of total probability** states that for any event B,

$$\Pr(B) = \sum_{i=1}^{k} \Pr(B|A_i)P(A_i)$$



Bayes' theorem states that if B is an event such that $\Pr(B) > 0$, then for every $j = 1, \ldots, k$,

$$\Pr(A_j|B) = \frac{\Pr(B|A_j)\Pr(A_j)}{\Pr(B)} = \frac{\Pr(B|A_j)\Pr(A_j)}{\sum_{i=1}^{k}\Pr(B|A_i)\Pr(A_i)}$$

Exercise 1 (20 mins)

Suppose the probability that a missile is launched and the radar detects it is 0.4, while the probability that a missile is not launched but the radar incorrectly detects it is 0.1.

Given that the radar detects a missile, what is the probability that a missile is launched?

Useful notation:

- Let M be the event that a missile is launched.
- Let M^c be the event that a missile is not launched.
- Let R be the event that the radar detects a missile.

Hint: $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$.

Exercise 1 - Solution

Our goal is to find $\Pr(M|R)$, and we are given the following:

$$\Pr(M \cap R) = 0.4 \text{ and } \Pr(M^c \cap R) = 0.1.$$

By the definition of conditional probability,

$$\Pr(M \cap R) = \Pr(R|M)\Pr(M) \text{ and } \Pr(M^c \cap R) = \Pr(R|M^c)\Pr(M^c).$$

Since M and M^c are mutually exclusive and exhaustive, we can apply the law of total probability:

$$\begin{aligned} \Pr(R) &= \Pr(R|M)\Pr(M) + \Pr(R|M^c)\Pr(M^c) \\ &= \Pr(M \cap R) + \Pr(M^c \cap R) \\ &= 0.5. \end{aligned}$$

Then, by Bayes' theorem,

$$\Pr(M|R) = \frac{\Pr(R|M)\Pr(M)}{\Pr(R)} = \frac{\Pr(M \cap R)}{\Pr(R)} = \frac{0.4}{0.5} = 0.8$$

Exercise 2 (20 mins)

You have recently launched a mobile app with in-app purchases.

Suppose we know the following:

- The probability that a user spends at least 4 hours on the app per week is 0.8.
- The probability that a user spends at least 4 hours per week on the app and at least $10 per week on in-app purchases is 0.5.
- Given that a user spends less than 4 hours per week on the app, the probability that the user spends at least $10 per week on in-app purchases is 0.1.

Kathy is a user of your app. Given that Kathy spends at least $10 per week on in-app purchases, what is the probability that she spends less than 4 hours per week on the app?

Exercise 2 - Solution

Let A and B be the following events:

- $A = \{\text{a user spends} \geq \$10 \text{ per week on in-app purchases}\}$.
- $B = \{\text{a user spends} < 4 \text{ hours per week on the app}\}$.

Our goal is to find $\Pr(B|A)$, and we are given the following:

$$\Pr(B^c) = 0.8, \quad \Pr(A \cap B^c) = 0.5, \quad \Pr(A|B) = 0.1.$$

Note that $\Pr(B) = 1 - \Pr(B^c) = 0.2$. Also, by the definition of conditional probability, we have $\Pr(A \cap B^c) = \Pr(A|B^c)\Pr(B^c)$.

Since B and B^c are mutually exclusive and exhaustive, we can apply the law of total probability:

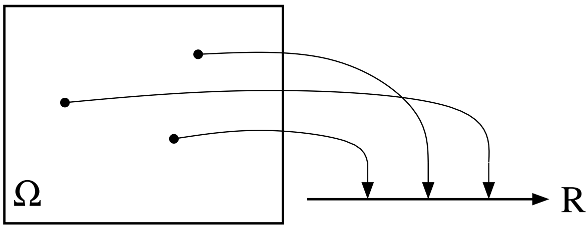
$$\begin{aligned} \Pr(A) &= \Pr(A|B)\Pr(B) + \Pr(A|B^c)\Pr(B^c) \\ &= (0.1)(0.2) + 0.5 = 0.52 \end{aligned}$$

Then, by Bayes' theorem,

$$\Pr(B|A) = \frac{\Pr(A|B)\Pr(B)}{\Pr(A)} = \frac{(0.1)(0.2)}{0.52} = \frac{1}{26}.$$

Random variable (R.V.)

Definition: For a given sample space Ω of some experiment, a random variable is any rule that associates a real number with each outcome in Ω.



- Today, we will look at examples of a special kind of random variable called **discrete random variable**.
- We shall learn more about random variables (inc. non-discrete ones) in **make-up Lecture 4** (this Friday, 2:30–4pm, LT5).

Discrete random variables

A random variable X is called discrete if X can take only a finite number k of different values $x_1, \ldots, x_k$, or, at most, an infinite sequence of different values $x_1, x_2, x_3, \ldots$

Examples:

- A coin is tossed three times. Let X be the number of heads in the experiment. X is a **discrete random variable**; it has four possible values: 0, 1, 2 and 3.
- Consider an experiment that consists of tossing a coin until the first head occurs. Let Y be the number of tosses. Y is also a **discrete random variable**; the possible values form an infinite sequence: $1, 2, 3, \ldots$.

Probability mass function (pmf) for discrete R.V.

Let X be a **discrete** R.V. defined on the sample space Ω.

The probability mass function (pmf) of X is a function $p(x)$, defined on every real number, such that

$$p(x) = \Pr(X = x) = \Pr(\{\omega \in \Omega : X(\omega) = x\}).$$

- In other words, $p(x)$ is the probability of the event $\{X \in \{x\}\}$, i.e. the event consisting of all outcomes whose X-value is x.
- The following five expressions mean the exact same thing:

$$\Pr(X \in \{x\}),\ \ \Pr(X = x),\ \ \Pr(\{X \in \{x\}\}),\ \ \Pr(\{X = x\}),\ \ p(x)$$

Note on terminology: In the textbook, "probability function" is used instead of "probability mass function". Either term is okay.

- Be careful: Whether you use pmf or "probability function", it makes sense only for **discrete** random variables!

Exercise 3 (10 mins)

Suppose X is a discrete R.V. with the following pmf

$$p(x) = \begin{cases} \frac{c}{2^x}, & \text{if } x = 0, 1, 2, \ldots; \\ 0, & \text{otherwise;} \end{cases}$$

where c is an unspecified constant.

1. Find the value of c.
2. What is $\Pr(X \geq 2)$?

Exercise 3 - Solution

1. The set of possible values for X is $C = \{0, 1, 2, \dots\}$.
Since $p(x)$ is a pmf, it must satisfy

$$\sum_{x \in C} p(x) = \sum_{x \geq 0} p(x) = 1.$$

In other words,

$$\sum_{k=0}^{\infty} \frac{c}{2^k} = c\left(1 + \tfrac{1}{2} + \tfrac{1}{4} + \dots\right) = 2c = 1,$$

therefore $c = \frac{1}{2}$. (Here, we use the formula for geometric series: If $|r| < 1$, then $a + ar + ar^2 + \cdots = \frac{a}{1-r}$.)

2. By definition, $p(x) = \Pr(X = x)$. Thus,

$$\begin{aligned} \Pr(X \geq 2) &= 1 - \Pr(X < 2) = 1 - \Pr(X = 1) - \Pr(X = 0) \\ &= 1 - p(1) - p(0) = 1 - \tfrac{1}{4} - \tfrac{1}{2} = \tfrac{1}{4}. \end{aligned}$$

Exercise 4 (15 mins)

Suppose X is a discrete R.V. with the following pmf:

$$p(x) = \begin{cases} k(x+1), & \text{if } x = 2,4,6; \\ k(x-2), & \text{if } x = 8 \\ 0, & \text{otherwise;} \end{cases}$$

where k is an unspecified constant.

1. Find the value of k.
2. Find the value of $\Pr(0 \le X \le 6)$?

Exercise 4 - Solution

1. Since $p(x)$ is a pmf, it must satisfy $\sum_{x \in D} p(x) = 1$, where D is the set of all possible values. Since $D = \{2, 4, 6, 8\}$, it follows that

$$p(2) + p(4) + p(6) + p(8) = 1.$$

In other words,

$$k(2+1) + k(4+1) + k(6+1) + k(8-2) = 21k = 1,$$

therefore $k = \frac{1}{21}$.

2. By definition, $p(x) = \Pr(X = x)$. Thus,

$$\begin{aligned} \Pr(0 \leq X \leq 6) &= p(2) + p(4) + p(6) \\ &= 1 - p(8) \\ &= 1 - \tfrac{1}{21}(8-2) \\ &= \tfrac{5}{7}. \end{aligned}$$

Exercise 5 (15 mins)

Consider the following functions:

$$p_1(x) = \begin{cases} \frac{1}{x}, & \text{if } x \geq 2, x \text{ is an integer;} \\ 0, & \text{otherwise;} \end{cases}$$

$$p_2(x) = \begin{cases} (-1)^{x+1} \frac{3}{2^x}, & \text{if } x \geq 1, x \text{ is an integer;} \\ 0, & \text{otherwise.} \end{cases}$$

1. Can $p_1(x)$ be the pmf of some discrete random variable?
2. Can $p_2(x)$ be the pmf of some discrete random variable?

Exercise 5 - Solution

1. If $p_1(x)$ is a pmf of some discrete R.V., then the set of possible values must be $\{2, 3, 4, \dots\}$, and the function $p_1(x)$ must satisfy

$$\sum_{x=2}^{\infty} p_1(x) = 1.$$

However,

$$\begin{aligned}
\sum_{x=3}^{\infty} p_1(x) &= \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \dots \\
&= \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right) + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \dots \\
&= \left(\frac{13}{12}\right) + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \dots \\
&> 1,
\end{aligned}$$

therefore $p_1(x)$ **cannot** be the pmf of any discrete random variable. (In fact, the infinite series $\sum_{x=k}^{\infty} \frac{1}{x}$ diverges for any fixed integer k.)

Exercise 5 - Solution

2. If $p_2(x)$ is a pmf of some discrete R.V. X, then by definition, $p_2(x) = \Pr(X = x)$ is the probability of the event $\{X = x\}$, so in particular, it must have a non-negative value.

However, $p(2) = -\frac{3}{4} < 0$, $p(4) = -\frac{3}{16} < 0$, and more generally, $p(n) < 0$ whenever $n = 2k$ is an even integer. Therefore, $p_2(x)$ **cannot** be the pmf of any discrete random variable.

- Some of you may have noticed that

$$\sum_{x=1}^{\infty} p_2(x) = \sum_{x=1}^{\infty} (-1)^{x+1} \tfrac{3}{2^x} = 1.$$

- But remember, for a function $p(x)$ to be an actual pmf of some discrete R.V., it must satisfy **all** of the following:
  - $\sum_{x \in D} p(x) = 1$;
  - $0 \leq p(x) \leq 1$ for all $x \in D$;

  where D is a finite or countably infinite set of possible values.

Summary

Exercises on the following topics:

- Law of total probability
- Bayes' theorem
- Probability mass function (pmf)

Reminders:

- There is **mini-quiz 1** next Cohort Class!
  - Tested on materials from Lecture 1 up to and including Slide 7 ("Mean and variance of binomial R.V.") of Lecture 6.
- Homework Set 2 is also due next Cohort Class.
- **Make-up Lecture 4** is held this Friday, 2:30pm-4pm, LT5.