Lecture 15 – Foundations of Probability

DSC 40A, Spring 2023

Announcements

- No homework due this week!
- Janine is not holding office hours today.
- Welcome to Part 2 of the course!

Agenda

- Probability: context and overview.
- Complement, addition, and multiplication rules.
- Conditional probability.

Probability: context and overview

From Lecture 1: course overview

Part 1: Learning from Data

- Summary statistics and loss functions; mean absolute error and mean squared error.
- Linear regression (incl. linear algebra).
- Clustering.

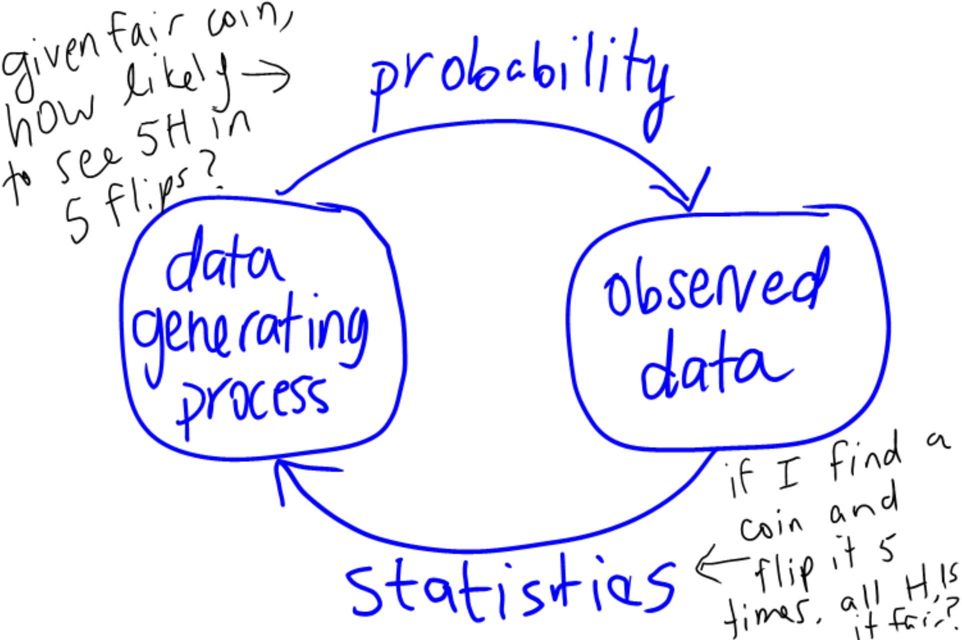
Part 2: Probability

- Probability fundamentals. Set theory and combinatorics.
- Conditional probability and independence.
- Naïve Bayes (uses concepts from both parts of the class).

Why do we need probability?

- So far in this class, we have made predictions based on a dataset.
- This dataset can be thought of as a **sample** of some population.
- For a prediction rule to be useful in the future, the sample that was used to create the prediction rule needs to look similar to samples that we'll see in the future.

Probability and statistics



Statistical inference

Given observed data, we want to know how it was generated or where it came from, for the purposes of

- predicting outcomes for other data generated from the same source.
- knowing how different our sample could have been.
- drawing conclusions about our entire population and not just our observed sample (i.e. generalizing).

Probability

Given a certain model for data generation, what kind of data do you expect the model to produce? How similar is it to the data you have?

- Probability is the tool to answer these questions.
- You need probability to do statistics, and vice versa.
- Example: Is my coin fair?

?

fair coin

5 flips
5 H

Terminology

$S = \{H, T\}$

$S = \{1, 2, 3, 4, 5, 6\}$

- An **experiment** is some process whose outcome is random (e.g. flipping a coin, rolling a die).
- A **set** is an unordered collection of items. $|A|$ denotes the number of elements in set A.

$|A| = 3$ $\quad$ $A = \{1, 3, 4\} = \{3, 4, 1\}$

- A **sample space**, S, is the set of all possible outcomes of an experiment.
  - Could be finite or infinite!
- An **event** is a subset of the sample space, or a set of outcomes.
  - Notation: $E \subseteq S$. subset

want prob. of an event

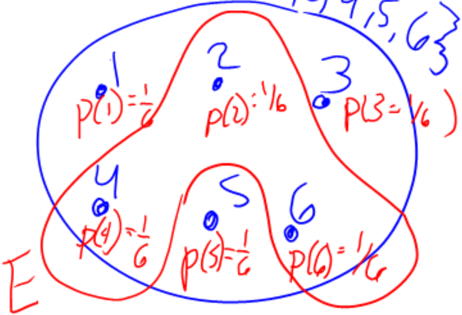
Probability distributions

- A probability distribution, p, describes the **probability** of each outcome s in a sample space S.
  - The probability of each outcome must be between 0 and 1: $0 \leq p(s) \leq 1$. ← number between 0 and 1 associated with each outcome in sample space
  - The sum of the probabilities of each outcome must be exactly 1: $\boxed{\sum_{s \in S} p(s) = 1.}$
- The probability of an **event** is the sum of the probabilities of the outcomes in the event.
  - $\boxed{P(E) = \sum_{s \in E} p(s).}$

Example: probability of rolling an even number on a 6-sided die

$1/2 = 3/6$

$S = \{1, 2, 3, 4, 5, 6\}$



$E = \{2, 4, 6\} \subseteq S = \{1, 2, 3, 4, 5, 6\}$

subset

$$p(E) = \sum_{s \in E} p(s) = p(2) + p(4) + p(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = 1/2$$

Equally-likely outcomes

- If S is a sample space with n possible outcomes, and all outcomes are equally-likely, then the probability of any one outcome occurring is $\frac{1}{n}$.

S

- The probability of an event E, then, is

$$P(E) = \underbrace{\frac{1}{n} + \frac{1}{n} + \ldots + \frac{1}{n}}_{\text{one term for every element of } E} = \frac{\text{\# of outcomes in E}}{\text{\# of outcomes in S}} = \frac{|E|}{|S|}$$

- **Example:** Flipping a coin three times.

$S = \{HHH, HHT, HTH, \ldots\} \to$ size 8

$= |E|$ terms

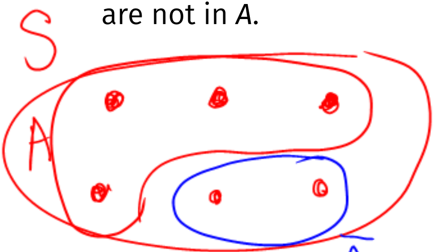
$E =$ see exactly 2 H's $= \{HHT, HTH, THH\}$

$\frac{|E|}{|S|} = \frac{3}{8} \to$ only for fair coin

Complement, addition, and multiplication rules

Complement rule

- Let A be an event with probability $P(A)$.
- Then, the event $\bar{A}$ is the **complement** of the event A. It contains the set of all outcomes in the sample space that are not in A.



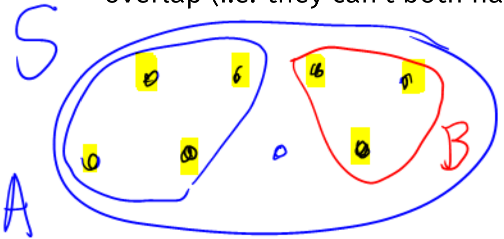
$\rightarrow$ $P(A) + P(\bar{A}) = P(S) = 1$

- $P(\bar{A})$ is given by

$$\boxed{P(\bar{A}) = 1 - P(A)}$$

Addition rule

- We say two events are **mutually exclusive** if they have no overlap (i.e. they can't both happen at the same time).



$$P(\underbrace{A \cup B}) = \sum_{s \text{ in } A \text{ or } B} p(s)$$

$$= \sum_{s \text{ in } A} p(s) + \sum_{s \text{ in } B} p(s)$$

$$= p(A) + p(B)$$

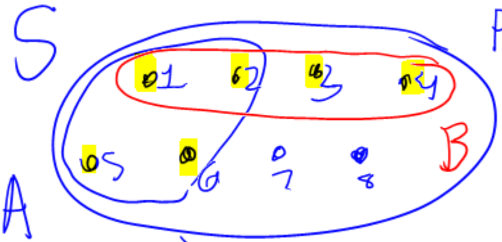
- If A and B are mutually exclusive, then the probability that A or B happens is

$$P(A \cup B) = P(A) + P(B)$$

union ← "or"

Principle of inclusion-exclusion

- If events A and B are not mutually exclusive, then the addition rule becomes more complicated.



$P(A \cup B) = p(1) + p(2) + p(3) + p(4) + p(5) + p(6)$

$P(A) + P(B) - P(A \cap B) = p(1) + p(2) + p(5) + p(6) + p(1) + p(2) + p(3) + p(4) - (p(1) + p(2))$

- In general, if A and B are any two events, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

intersect ("and")

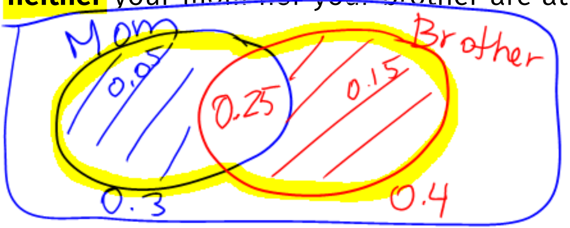
Discussion Question

Each day when you get home from school, there is a

- 0.3 chance your mom is at home.
- 0.4 chance your brother is at home.
- 0.25 chance that both your mom and brother are at home.

When you get home from school today, what is the chance that **neither** your mom nor your brother are at home?

a) 0.3
$1-0.45=$ b) 0.45
c) 0.55
d) 0.7
e) 0.75



$P(\text{Mom OR Brother}) = 0.3 + 0.4 - 0.25 = 0.45$

Multiplication rule and independence

- The probability that events A and B both happen is

$$P(A \cap B) = P(A)\boxed{P(B|A)}$$

"and"

- $P(B|A)$ means "the probability that B happens, given that A happened." It is a **conditional probability**.

once you know that

(B is conditioned on A)

- If $P(B|A) = P(B)$, we say A and B are **independent**.
  - Intuitively, A and B are independent if knowing that A happened gives you no additional information about event B, and vice versa.

  - For two independent events,

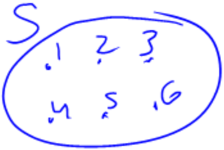
$$P(A \cap B) = P(A)\boxed{P(B)}$$

simplifies when independent

Example: rolling a die

Let's consider rolling a fair 6-sided die. The results of each die roll are independent from one another.

- Suppose we roll the die once. What is the probability that the face is 1 and 2?



prob = 0

- Suppose we roll the die once. What is the probability that the face is 1 or 2?

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

Example: rolling a die

- Suppose we roll the die 3 times. What is the probability that the face 1 never appears in any of the rolls?

$$P\left(\begin{matrix}\text{no 1}\\ \text{on 1st roll}\end{matrix} \text{ AND } \begin{matrix}\text{no 1}\\ \text{on 2nd roll}\end{matrix} \text{ AND } \begin{matrix}\text{no 1}\\ \text{on 3rd roll}\end{matrix}\right)$$

$$= P\left(\begin{matrix}\text{no 1}\\ \text{on 1st roll}\end{matrix}\right) * P\left(\begin{matrix}\text{no 1}\\ \text{on 2nd roll}\end{matrix} \cancel{\Bigg| \begin{matrix}\text{no 1}\\ \text{on 1st}\\ \text{roll}\end{matrix}}\right) * P\left(\begin{matrix}\text{no 1}\\ \text{on 3rd}\\ \text{roll}\end{matrix} \cancel{\Bigg| \begin{matrix}\text{no 1's}\\ \text{on 1st}\\ \text{or 2nd}\\ \text{rolls}\end{matrix}}\right)$$

$$= \left(1 - \frac{1}{6}\right) * \left(1 - \frac{1}{6}\right) * \left(1 - \frac{1}{6}\right) = \left(\frac{5}{6}\right)^3$$

- Suppose we roll the die 3 times. What is the probability that the face 1 appears at least once?

$$1 - \begin{matrix}\text{previous}\\ \text{answer}\end{matrix} = 1 - \left(\frac{5}{6}\right)^3$$

Example: rolling a die

- Suppose we roll the die n times. What is the probability that only the faces 2, 4, and 5 appear?

$$P\binom{2,4,\text{or }5}{\text{on }1^{st}\text{ roll}} * P\binom{2,4,5}{\text{on }2^{nd}\text{ roll}} \times \dots * P\binom{2,4,5}{\text{on }n^{th}\text{ roll}}$$

$$\frac{1}{2} \times \frac{1}{2} \times \dots \times \frac{1}{2} = \left(\frac{1}{2}\right)^n$$

- Suppose we roll the die twice. What is the probability that the two rolls have different faces?

$\frac{5}{6}$

S = outcomes for 2nd roll

P(different) = $\frac{5}{6}$

S = {pairs (1,1)
(1,2), (1,3) ...

|S| = 36

30/36

Conditional probability

Conditional probability

- The probability of an event may **change** if we have additional information about outcomes.

- Starting with the multiplication rule, $P(A \cap B) = P(A)P(B|A)$, we have that

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

  assuming that $P(A) > 0$.

Example: pets

Suppose a family has two pets. Assume that it is equally likely that each pet is a dog or a cat. Consider the following two probabilities:

1. The probability that both pets are dogs given that **the oldest is a dog**.

2. The probability that both pets are dogs given that **at least one of them is a dog**.

Discussion Question

Are these two probabilities equal?

a) Yes, they're equal
b) No, they're not equal

Example: pets

Let’s compute the probability that both pets are dogs given that **the oldest is a dog**.

Example: families

Let’s now compute the probability that both pets are dogs given that **at least one of them is a dog**.

Summary, next time

Summary

- Two events A and B are mutually exclusive if they share no outcomes (no overlap). In this case,

$$P(A \cup B) = P(A) + P(B).$$

- More generally, for any two events,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

- The probability that events A and B both happen is

$$P(A \cap B) = P(A)P(B|A).$$

- $P(B|A)$ is the conditional probability of B occurring, given that A occurs. If $P(B|A) = P(A)$, then events A and B are independent.

Next time

- More probability and introduction to combinatorics, the study of counting.
- **Important:** We've posted **many** probability resources on the resources tab of the course website. These will no doubt come in handy.
    - No more DSC 40A-specific readings, though the Probability Roadmap was written specifically for students of this course.