# Module 15 – Foundations of Probability

DSC 40A, Summer 2023

## Announcements

- No homework due this week. Homework 3 will be released tomorrow, will be due Tuesday of next week.
- Final will be closed books/notes/electronics/web. You will be allowed to keep with you two A4-sized sheets (four sides) with any content you want.
- 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.

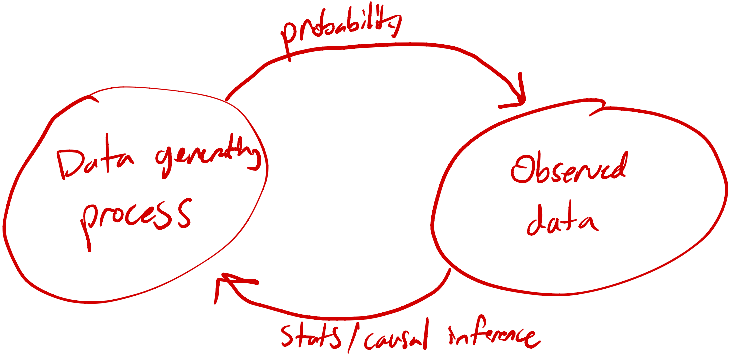
**Part 2: Probability**

- 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. → less of priority for stats
- 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?

## Terminology

- An **experiment** is some process whose outcome is random (e.g. flipping a coin, rolling a die).

  generalize to non-random

- A **set** is an unordered collection of items. $|A|$ denotes the number of elements in set $A$.

  cardinality

  $\{3, 7, 8\} = \{8, 3, 7\}$

- A **sample space**, $S$, is the set of all possible outcomes of an experiment.
  - Could be finite or infinite!

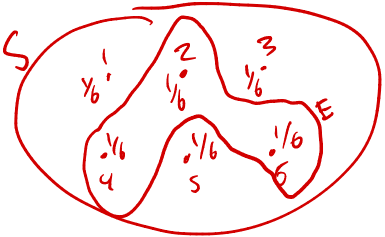
- An **event** is a subset of the sample space, or a set of outcomes.
  - Notation: $E \subseteq S$.

# 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: $0 \leq p(s) \leq 1$.
  - The sum of the probabilities of each outcome must be exactly 1: $\sum_{s \in S} p(s) = 1$.
- The probability of an **event** is the sum of the probabilities of the outcomes in the event.
  - $P(E) = \sum_{s \in E} p(s)$.

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

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




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

# 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 $\frac{1}{n}$.

- The probability of an event $E$, then, is

$$P(E) = \frac{1}{n} + \frac{1}{n} + \ldots + \frac{1}{n} = \frac{\text{\# of outcomes in E}}{\text{\# of outcomes in S}} = \frac{|E|}{|S|}$$

cardinality of event

cardinality of sample space

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

$S = \{HHH, HHT \ldots\} \rightarrow |S| = 8$

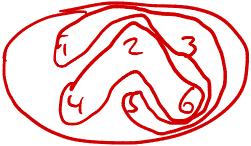
$E = \{HHT, HTH, THH\} \rightarrow |E|\ 3$

$$\frac{3}{8}$$

# 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 $A$.

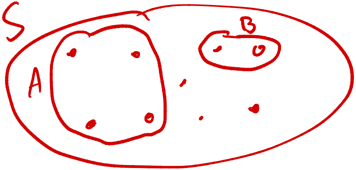



- $P(\bar{A})$ is given by

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

# 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).



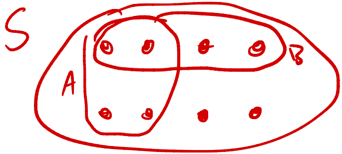

- If $A$ and $B$ are mutually exclusive, then the probability that $A$ or $B$ happens is

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

union...
"or"

# Principle of inclusion-exclusion

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




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

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

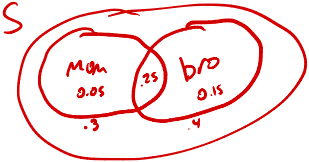
## Discussion Question

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
b) 0.45
c) 0.55
d) 0.7
e) 0.75




$P(\text{mom or bro}) = .3 + .4 - .25 = .45$

$1 - .45 = .55$

# Multiplication rule and independence

"or" generally involves addition

"and" generally involves multiplication

- The probability that events $A$ and $B$ both happen is

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

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

- 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.

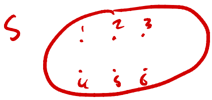
  - For two independent events,

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

# Example: rolling a die

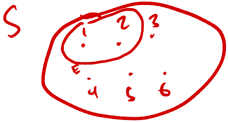
Let's consider rolling a fair 6-sided die. The results of each die roll are independent from one another.

- Suppose we roll the die once. What is the probability that the face is 1 and 2? → 0




→ ~~had it been two rolls ... 11 12 13 ...~~

- 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?

$$P\left(\substack{\text{no 1}\\\text{first roll}} \text{ and } \substack{\text{no 1}\\\text{second roll}} \text{ and } \substack{\text{no 1}\\\text{third roll}}\right) =$$

$$P\left(\substack{\text{no 1}\\\text{first roll}}\right) * P\left(\substack{\text{no 1}\\\text{second roll}}\right) * P\left(\substack{\text{no 1}\\\text{third roll}}\right) =$$

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

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

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

# Example: rolling a die

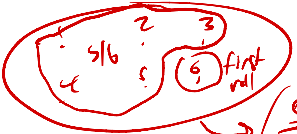

discrete event has prob $\frac{1}{2}$

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

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

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

S = outcomes for second roll




$$\frac{5}{6}$$

second

| 1st | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | x | | | | | |
| 2 | | x | | | | |
| 3 | | | x | | | |
| 4 | | | | x | | |
| 5 | | | | | x | |
| 6 | | | | | | x |

36 total

$$\frac{30}{36}$$

# Summary, next module

# 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,

if not mutually exclusive

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

- The probability that events $A$ and $B$ both happen is

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

can drop if independent

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

# Next module

- 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.