ENGG 107: Bayesian Statistical Modeling and Computation

The Bayesian Approach (I)

Meeting #2

Objectives for today:

1. Review of key concepts of probability
2. Introduction to the Bayesian Approach

Draft Schedule

Week	Week starting	Class 1	Class 2	Class 3	Notes	Assignment out	Assignment due
1	January 2, 2023	No Class	Introduction to the class	Bayesian Approach I			none
2	January 9, 2023	Bayesian Approach II	Picking a project	Setting up Computation		1	none
3	January 16, 2023	No class (MLK day)	Analytical Solutions / Kalman	Precalibration /(Bayes) Monte Carlo		2	1
4	January 23, 2023	MCMC Part 1	MCMC Part 2	Bayesian Workflow	X hour class	3	2
5	January 30, 2023	Students pitch project ideas	Writing a method section	Catching up / review		4	3
6	February 6, 2023	Convergence diagnostics	Checking Assumptions	Deep Uncertainty		5	4
7	February 13, 2023	Geeting Solid Priors	Links to Decision-Making	Links to Decision-Making		6	5
8	February 20, 2023	Model Choice	Emulation	Sensitivity analysis		7	6
9	February 27, 2023	Communication of results	Student Presentations	Student Presentations		8	7
10	March 6, 2023	Class Debriefing / Resarch links	No Class	No Class		none	8

Recap from last time

What is yodeling?

https://www.youtube.com/watch?v=ByHPs5BQAqQ

The German Yodeling Test

- There is genetic trait (i.e., exposure to German yodeling without headache) with a 0.5 % incidence rate in certain populations (let's assume Germans). You, as a member of the population, test positive, but the test is only 98 % accurate (random error)

- What is the probability that you have the trait after one, two, and three test applications with a positive test result?

Please try to solve this warm-up problem for the first yodeling test application.

Any volunteers to do this on the whiteboard?

Logistics

- Weekly meetings:
  - Lectures: Classes: MWF 11:30 – 12:35 (009 ESCS)
  - X times on Tu 12:15 – 1:05 PM (**on demand, announced**)
  - Problem sets discussions: Fridays 14:00-15:00
    (meet in my office and outside my office in a work area)
  - Office hours: Fridays 15:00-16:30 (in my office IR 385)
- A mixture of
  - lectures to introduce the example problem, the theoretical background, and the possible approaches to solving the problem and
  - problem centered sessions where we tackle example problems in class exercises and code discussions.
- Does everybody have access to a computer?
- What are your experiences with data analysis programs?
- Materials and assignments via CANVAS

Expectations from Students

This class will require your sustained attention.

Students are expected to:

- prepare for the meetings (e.g., by carefully reading and synthesizing the reading assignments and being prepared to present their synthesis in class),
- actively contribute to the group discussions, and
- prepare and submit the assignments in time.

The class is designed with the expectation that students spend roughly three times the lecture contact hours outside the class for readings and assignments.

Academic Integrity

- We are all expected to do our own work on exams.
- Plagiarism is the handing in of materials from other sources as your own work without quoting or clear citation.
- Please review the syllabus on CANVAS and return a signed copy on the assigned due date.

Grading

The overall course grade will be based on a weighted average of the grades for these components:

• Problem sets	40%
• Term Project Presentation	20%
• Written Term Project Report	40%

We adopt the Vikrant Vaze Grading Algorithm (1):

“Your final grade will be decided based on your absolute as well as relative performance in the class, whichever is higher. Out of 100 total points, if you score at least min(90, Median + Standard Deviation) points, you are guaranteed an HP grade. If you score at least min(75, Median - Standard Deviation) points, you are guaranteed at least a P grade. Note that this is even more favorable to you than the standard grading on the curve.”

(1) Personal communication from Vikrant Vaze

Review

Do you know

1. Who is in this course?
2. How this course is designed?
3. What you can learn in this course?
4. What is expected from you in this course?
5. How to solve the German Yodeling Test problem?

End Recap from last time

Outline for today

1. Reading reviews
2. More on motivation
3. Key concepts of probability
4. Conditional probability
5. Bayesian *vs* Frequentist interpretation of probability
6. The prior, the likelihood, and the posterior

Outline for today

1. Reading reviews
2. More on motivation
3. Key concepts of probability
4. Conditional probability
5. Bayesian vs Frequentists
6. The prior, the likelihood, the posterior

Review of Reading Assignments

- Core:
  - Applegate, P. J., & Keller, K. (Eds.). (2016). Risk analysis in the Earth Sciences: A Lab manual. 2nd edition. Leanpub. Open Source and free at: https://leanpub.com/raes (please read the Chapter: "Introduction")
  - Clyde, M., M. Çetinkaya-Rundel, C. Rundel, D. Banks, C. Chai, L. Huang. (2022). An Introduction to Bayesian Thinking. Retrieved from https://statswithr.github.io/book/ accessed 02/07/2022 (Please read Chapter 1).
- Supplementary (if you are interested in background)
  - Freedman, D. (1997). Some Issues in the Foundation of Statistics. In B. C. van Fraassen (Ed.), Topics in the Foundation of Statistics (pp. 19–39). Springer Netherlands. https://doi.org/10.1007/978-94-015-8816-4_4
  - Jefferys, W. H., & Berger, J. O. (1992). Ockham's razor and Bayesian analysis. American Scientist, 80(1), 64–72. https://www.jstor.org/stable/pdf/29774559.pdf
  - Clark, J. S. (2004). Why environmental scientists are becoming Bayesians: Modelling with Bayes. Ecology Letters, 8(1), 2–14. https://doi.org/10.1111/j.1461-0248.2004.00702.x

Reading Review Questions

1. What are the main points?
2. What remains unclear?
3. What questions does this raise?
4. What does this mean for your research / project design?

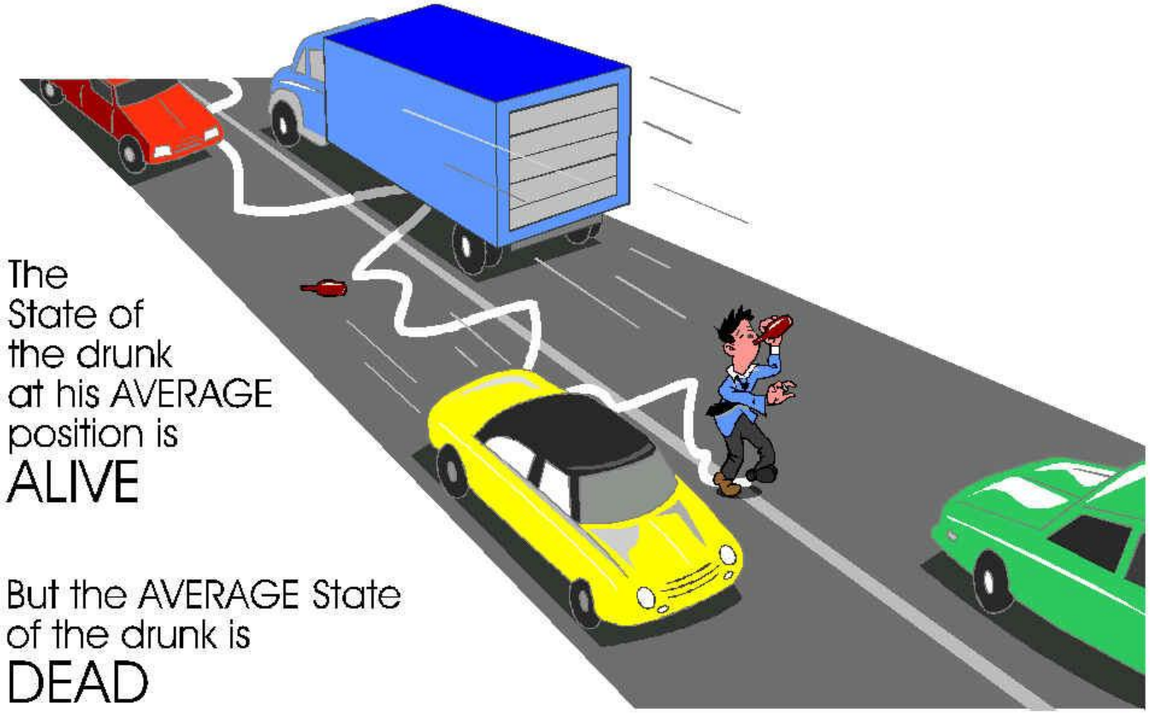
Outline for today

1. Reading reviews
2. More on motivation
3. Key concepts of probability
4. Conditional probability
5. Bayesian vs Frequentists
6. The prior, the likelihood, the posterior

Why may an analysis that neglects uncertainty lead to poor decisions?

Uncertainty can be thy enemy.

Know thy enemy.



http://web.stanford.edu/~savage/flaw/

Why perform quantitative and formal risk analysis?

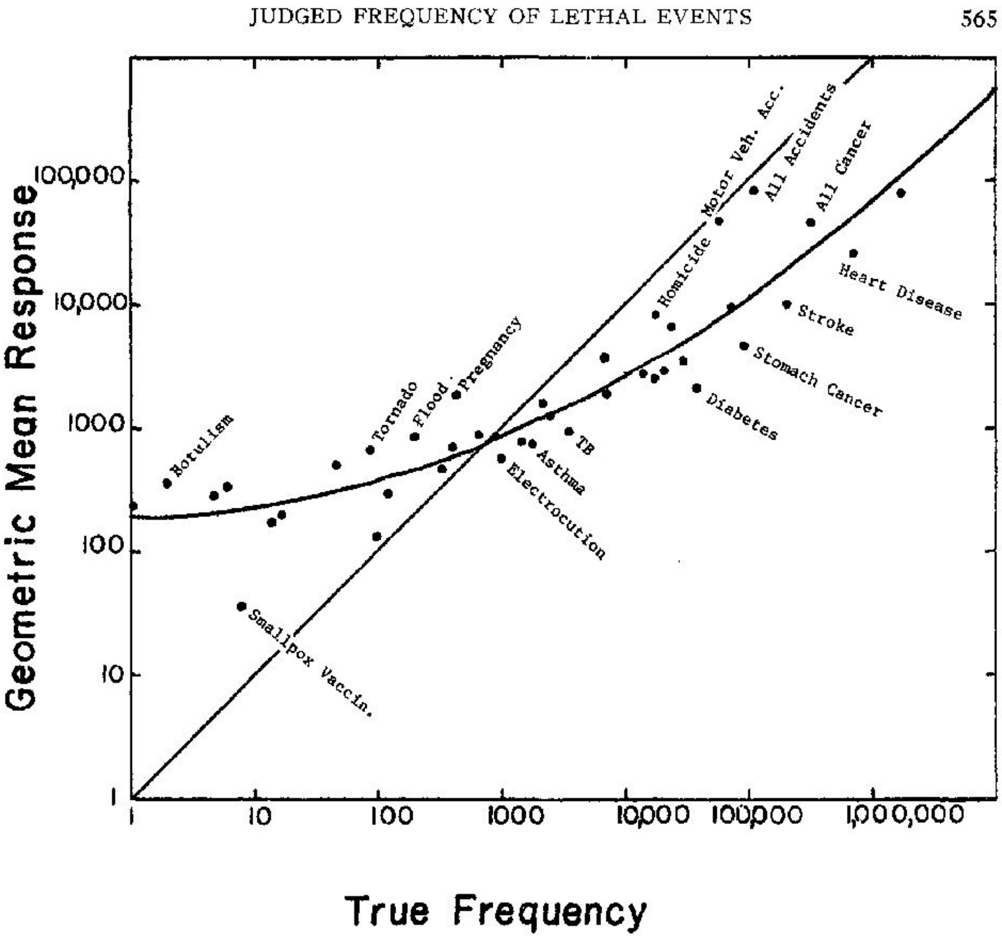


Figure 10. Geometric means (GM) of ratio judgments by motor vehicle accident group subjects as a function of true frequency (TF). (Curved line is best-fitting quadratic: log GM = .07 [log TF]² + .03 log TF + 2.27.)

Lichtenstein 1978

Recap: Do you bet on a new coin?

- Someone approaches you in Times Square in NYC.
- They offer you a bet on a coin resting on their hand.
- You see a head on the upwards side of the coin you can see.
- Bet: Pay 10$ upfront. If there is just one tail in 10 repeated tosses of the coin, you get double your payment.
- Do you accept this bet?
- How would you analyze this decision using the statistics you have learned thus far?

Let's get started very simple: a possibly crooked coin…

Is this coin fair?

- Typical interpretation: "Fair" means that the *probabilities* of "heads" and "tails" are equal.
- Questions:
  - What is probability?
  - How can we estimate a probability?
  - How can we decide this question?
  - How certain are we about our answer?

Outline for today

1. Reading reviews
2. More on motivation
3. Key concepts of probability
4. Conditional probability
5. Bayesian vs Frequentists
6. The prior, the likelihood, the posterior

Summary of Terminology

sample space (Ω):

- The set of possible outcomes of an experiment.
- example once repeated coin toss: {HH, HT, TH, TT}

outcome (ω):

- realization of an experiment
- e.g. ω={HT}

event

- subset of the sample space
- an outcome is an event

probability (P):

A function P that assigns a real number to each event A that satisfy 3 axioms:

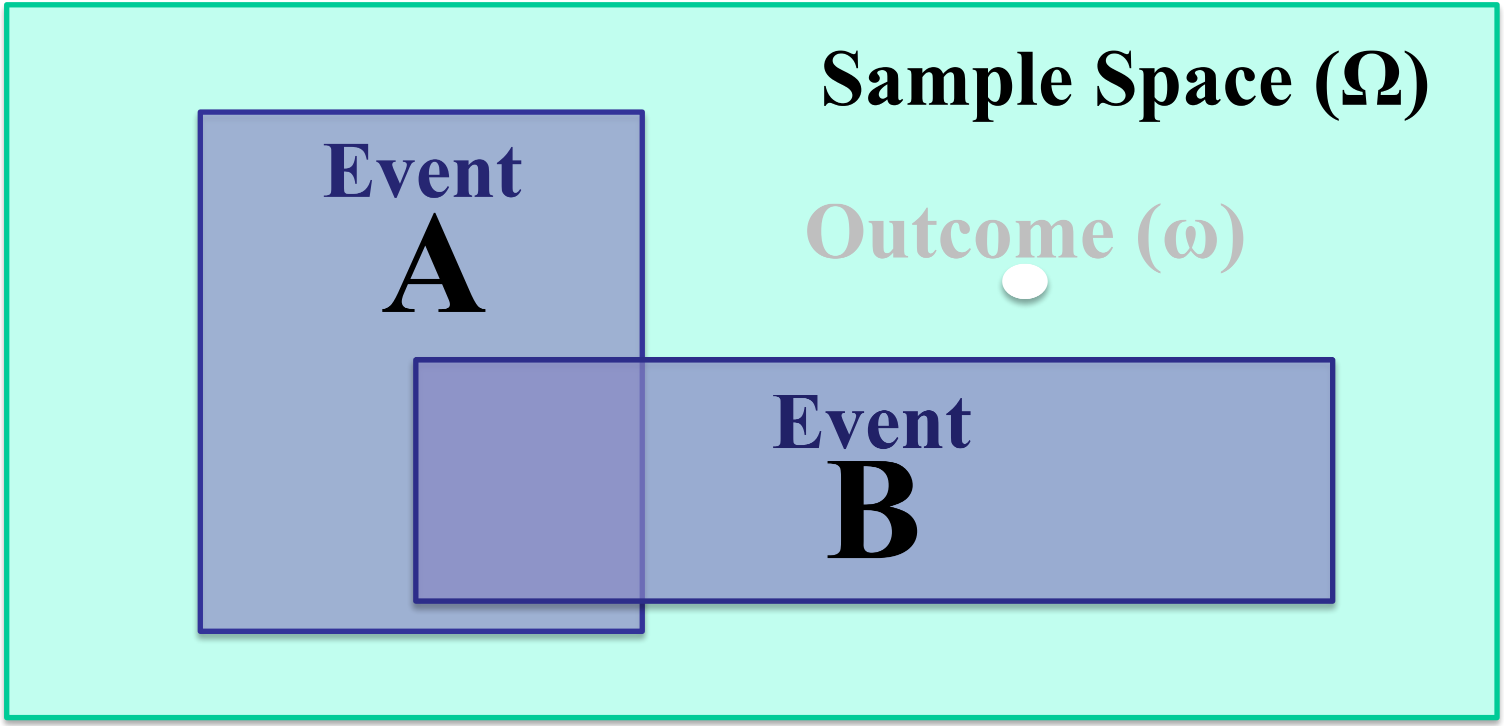
(i) $P(A) \geq 0$ for every A (the probability of an event is nonnegative)

(ii) $P(\Omega) = 1$ (the probability of the sample space is unity)

(iii) P(A or B)=P(A)+P(B), if A and B are mutually exclusive

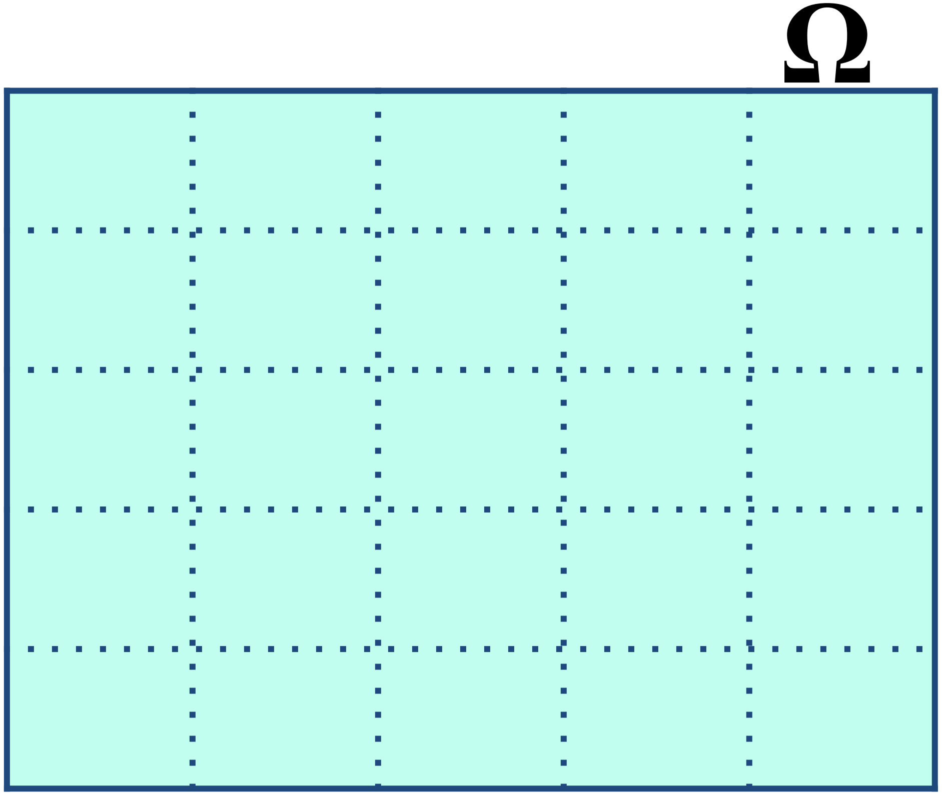
Wasserman (2004) Wilks (1995)

Graphical Review of Terminology



Review of Probability

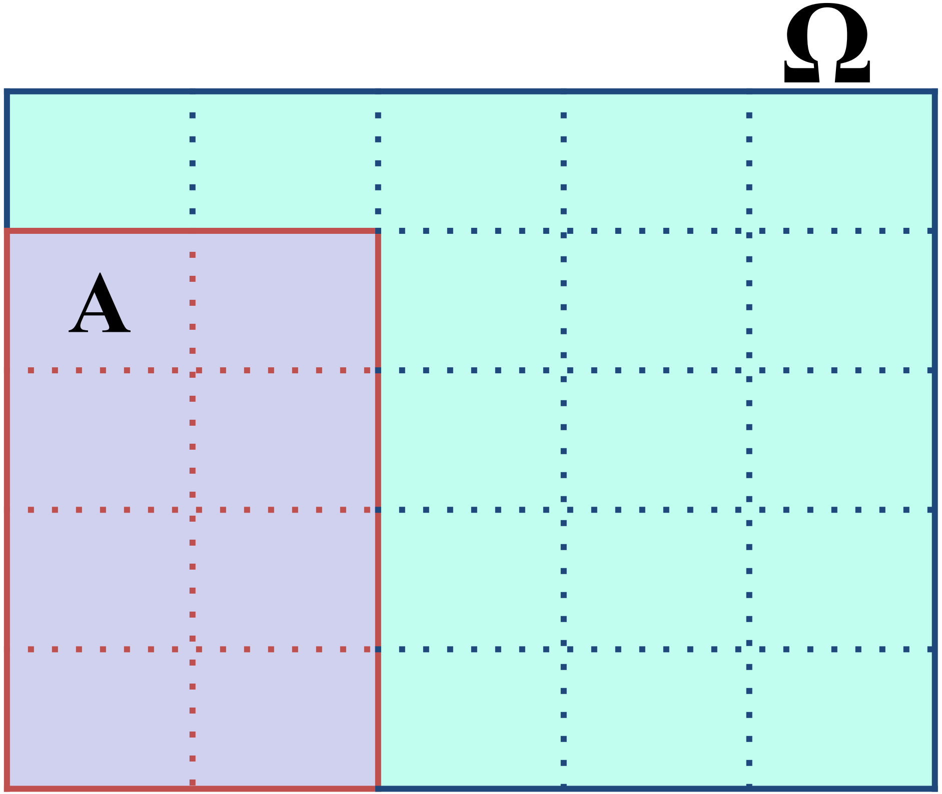
$$P(\Omega) = \quad ?$$



Review of Probability

$$P(\mathbf{\Omega}) = 1$$

$$P(\mathbf{A}) = ?$$

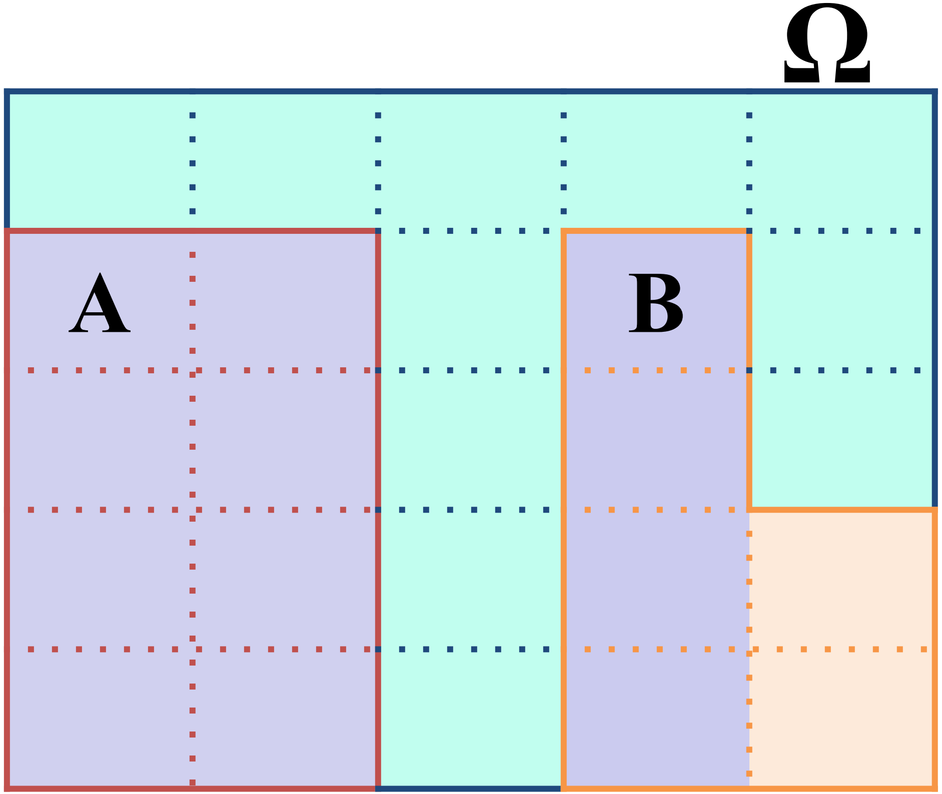


Review of Probability

$$P(\Omega) = 1$$

$$P(A) = \frac{8}{25}$$

$$P(B) = ?$$



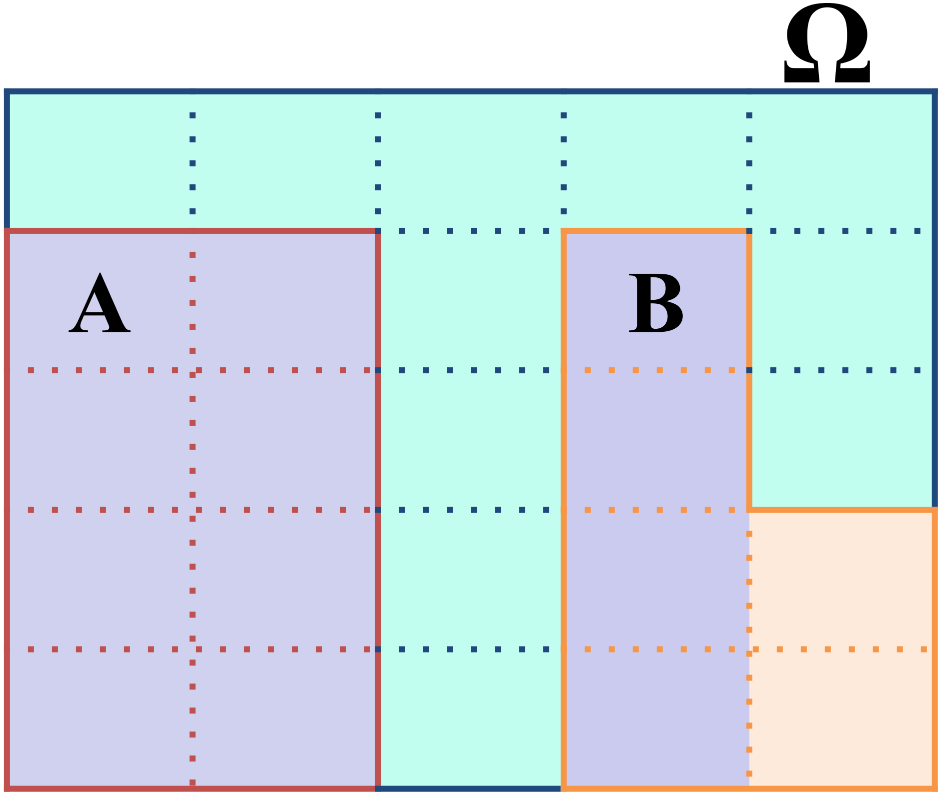
Review of Probability

$$P(\Omega) = 1$$

$$P(\mathbf{A}) = \frac{8}{25}$$

$$P(\mathbf{B}) = \frac{6}{25}$$

$$P(\mathbf{A} \text{ or } \mathbf{B}) = ?$$



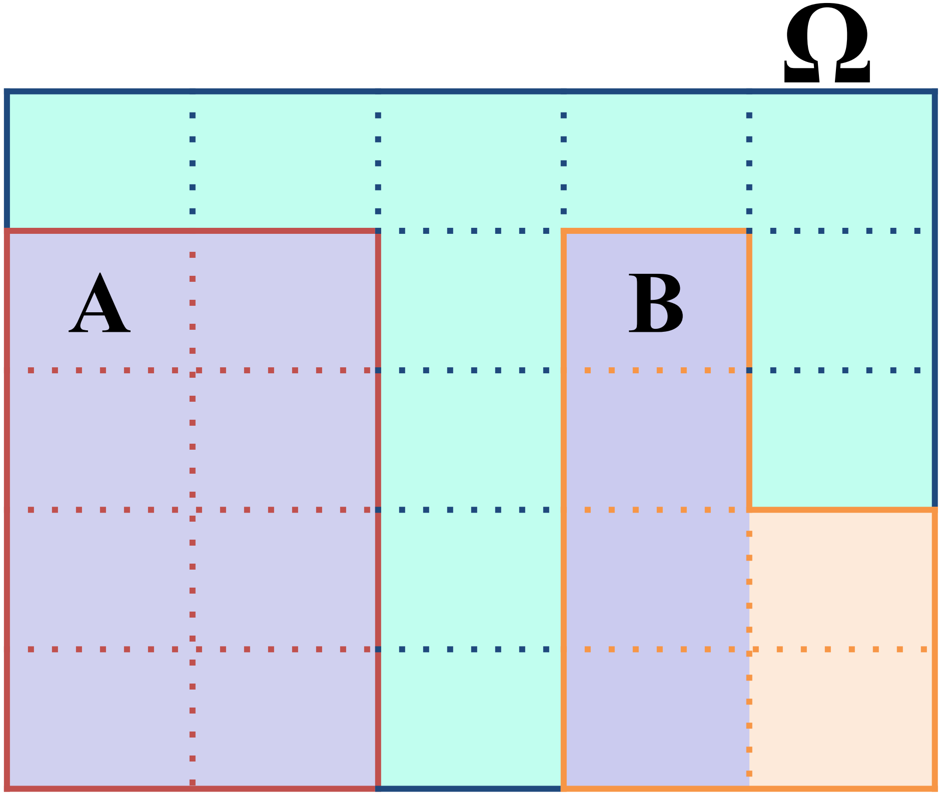
Review of Probability

$$P(\Omega) = 1$$

$$P(A) = \frac{8}{25}$$

$$P(B) = \frac{6}{25}$$

$$P(A \text{ or } B) = \frac{14}{25}$$



Basic Properties of Probabilities

$P(\varnothing) = 0$

$A \subset B \Rightarrow P(A) \leq P(B)$

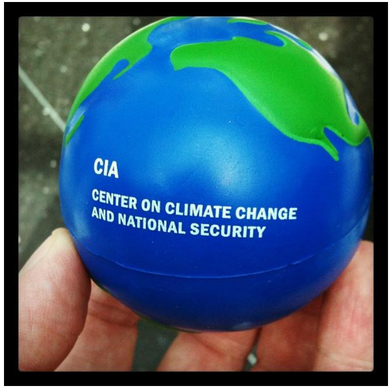
$0 \leq P(A) \leq 1$

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

P(A ∩ B) = P(A) • P(B)
(probability of A and B)
(if A and B are independent)

Wasserman (2004)

What is (deep) uncertainty for a dice outcome?



certain

uncertain

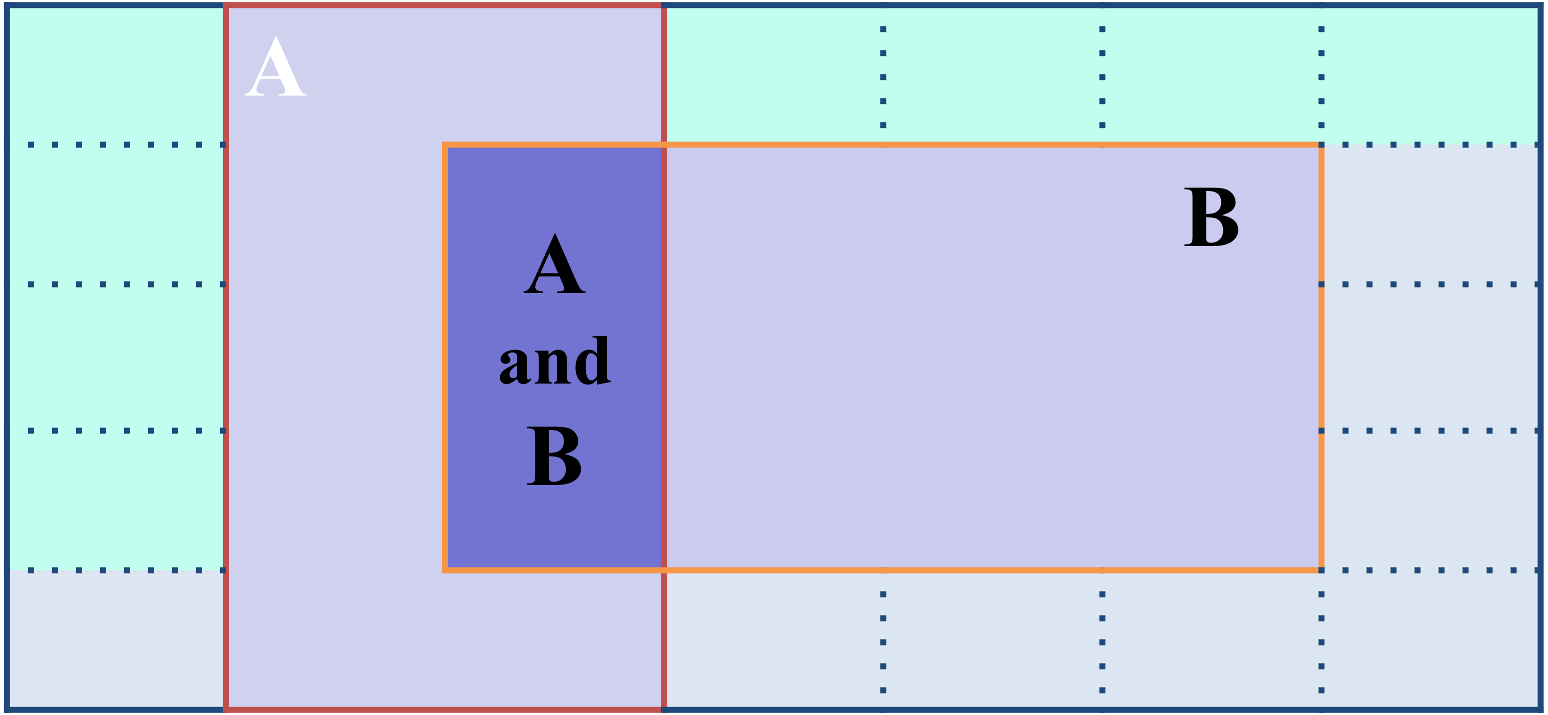
deeply uncertain
(We don’t know how many sides the die we’ll get has)

Credits: Sphere: Klaus Keller, Dice: https://en.wikipedia.org/wiki/Dice#/media/File:Twenty-sided_die_(icosahedron)_with_faces_inscribed_with_Greek_letters_MET_10.130.1158_001.jpg accessed Dec 28 2022 This file was donated to Wikimedia Commons as part of a project by the Metropolitan Museum of Art., Multiple Dice: By Diacritica - This file has been extracted from another file, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=88501597, accesses Dec 28 2022

Outline for today

1. Reading reviews
2. More motivating example problems
3. Key concepts of probability
4. Conditional probability
5. Bayesian vs Frequentists
6. The prior, the likelihood, the posterior

Conditional Probabilities



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

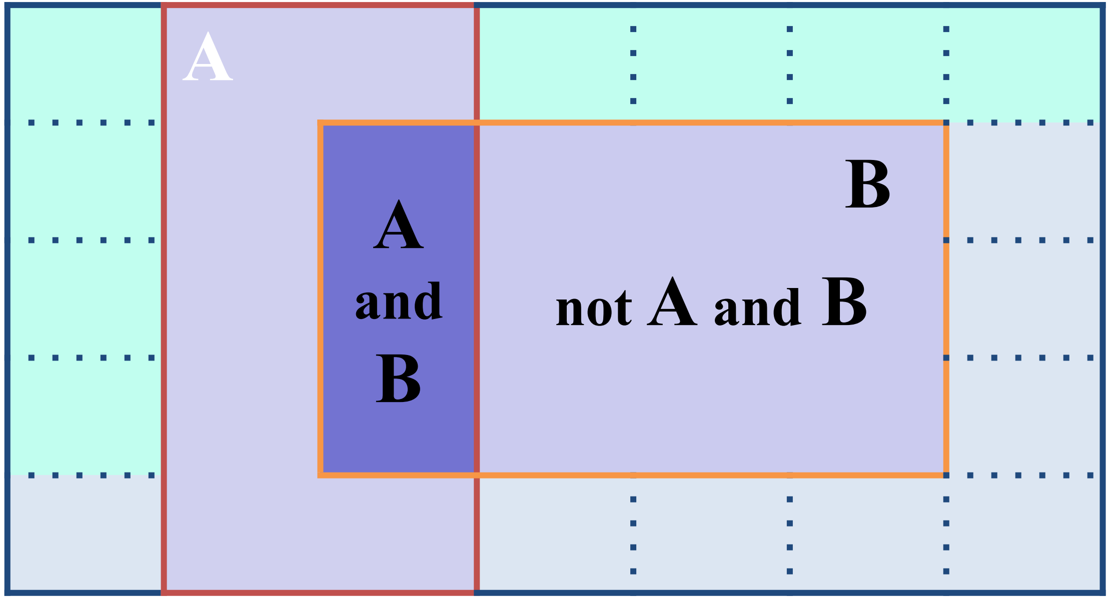
Law of Total Probabilities

Multiplication rule:

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

Law of total probability:

$$P(B) = P(A \cap B) + P(\bar{A} \cap B)$$
$$= P(A|B) \cdot P(B) + P(\bar{A}|B) \cdot P(B)$$



Example of Conditional Probability

- Family, 2 children
- What is the probability of having two girls, if the family has at least one girl?
- D: Both are girls
- C: At least one girl

First Child	Second Child	Number of Girls	Probability
	G2	**2**	**1/4**
G1			
	B2	**1**	**1/4**
	G2	**1**	**1/4**
B1			
	B2	0	1/4

$$P(D \mid C) = \frac{P(C \cap D)}{P(C)}$$

$$= \frac{P(G1 \cap G2)}{P(C)} = \frac{1/4}{3/4} = \frac{1}{3}$$

ftp.stat.duke.edu/Users/db/DOE_in_Charleston.ppt

Outline for today

1. Reading reviews
2. More motivating example problems
3. Key concepts of probability
4. Conditional probability
5. Bayesian vs Frequentists
6. The prior, the likelihood, the posterior

Two interpretations of probability

1) P(H) is the frequency of the outcome H in the limit of an infinitely long experiment.

   – "Frequentist" approach

2) P(H) is the degree of belief of an observer that H will occur in the next experiment.

   – "Bayesian" approach.

- Long and ongoing battle between Frequentists and Bayesians.

Wasserman (2004)

Outline for today

1. Reading reviews
2. More motivating example problems
3. Key concepts of probability
4. Conditional probability
5. Bayesian vs Frequentists
6. The prior, the likelihood, the posterior

Derivation of Bayes' Rule

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

definition of conditional probability

$$= \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

apply, again, the definition of conditional probability

$$= \frac{P(B \mid A) \cdot P(A)}{\mathrm{P(B|A)} \bullet \mathrm{P(A)} + \mathrm{P(B|\bar{A})} \bullet \mathrm{P(\bar{A})}}$$

law of total probability

Voila, Bayes' rule:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{\mathrm{P(B|A)} \bullet \mathrm{P(A)} + \mathrm{P(B|\bar{A})} \bullet \mathrm{P(\bar{A})}}$$

One application of Bayes' rule

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B|A) \bullet P(A) + P(B|\bar{A}) \bullet P(\bar{A})}$$

$A \Rightarrow$ Hypothesis

$\bar{A} \Rightarrow$ Alternative hypothesis

$B \Rightarrow$ Observation(s)

$P(B|A) \Rightarrow$ probability of observation B, given A, a.k.a. the likelihood of B (discussed below)

$P(B|\bar{A}) \Rightarrow$ probability of observation B, given $\bar{A}$

$P(A) \Rightarrow$ Prior probability of A

$P(\bar{A}) \Rightarrow$ Prior probability of $\bar{A}$

Zebras vs horses?

- You stand in Times Square in NYC and hear the sound of hoofs behind you.
- Without turning around, what is your estimate of the probability the this is a zebra or a horse behind you?
- How do you arrive at this estimate?
- Does your estimate change if you were
  - in the zoo in Central Park?
  - in a national park in the Serengeti?
  - at a horse race track?

Solving the zebra problem the Bayesian way

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B|A) \bullet P(A) + P(B|\bar{A}) \bullet P(\bar{A})}$$

$A \Rightarrow$ Hypothesis

$\bar{A} \Rightarrow$ Alternative hypothesis

$B \Rightarrow$ Observation(s)

$P(B|A) \Rightarrow$ probability of observation B, given A, a.k.a. the likelihood of B (discussed below)

$P(B|\bar{A}) \Rightarrow$ probability of observation B, given $\bar{A}$

$P(A) \Rightarrow$ Prior probability of A

$P(\bar{A}) \Rightarrow$ Prior probability of $\bar{A}$

- What are we looking for?
- What is the prior?
- What is the likelihood function?
- What the the posterior?

Reading Assignments for Next Class

- Core
  - van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. Nature Reviews Methods Primers, 1(1), 1–26. https://doi.org/10.1038/s43586-020-00001-2
    This is tough going, we will use this as a map and an example how people write a primer for the entire field.
- Background (if you are interested)
  - Freedman, D. (1997). Some Issues in the Foundation of Statistics. In B. C. van Fraassen (Ed.), Topics in the Foundation of Statistics (pp. 19–39). Springer Netherlands. https://doi.org/10.1007/978-94-015-8816-4_4
  - Jefferys, W. H., & Berger, J. O. (1992). Ockham's razor and Bayesian analysis. American Scientist, 80(1), 64–72. https://www.jstor.org/stable/pdf/29774559.pdf

Review

1. What is probability?
2. What is a Bayesian approach?
3. What are key differences between a Bayesian and a Frequentist approach?
4. How to define and use (in a simple example)
   a. prior
   b. likelihood
   c. probability