What Is Probability?

EES 4891/5891

Probability & Statistics for Geosciences

Jonathan Gilligan

Class #2: Thursday, January 09 2025

Review of Mathematical Notation

Review of Mathematical Notation

- Logic:

  E, F: Propositions that are either true or false

  $\bar{E}, \bar{F}$: Negations: the opposites of E and F.

  $E \cup F$: E or F is true (includes both true)

  $E \cap F$: E and F are both true

  $E|F$: The truth of E if F is known to be true

- Probability:

  $P(E)$: The probability that E is true

  $P(E|F)$: The probability that E is true, given that we know F is true

What Is Probability?

What Is Probability?

- What does it mean to say that the probability of something has some value?
  - The probability of a coin coming up “heads” is 0.5 (50%)
  - The probability of a die coming up “5” is 0.17 (17%, or 1 in 6)
  - The probability that it will snow in Nashville tomorrow is 0.97 (97%)
  - Nate Silver’s prediction in 2016 that the probability of Hillary Clinton winning the election was 0.714 (71.4%)
  - The probability that there will be a caldera-forming supereruption at Yellowstone in 2025 is one in 730,000 (according to USGS)
  - The probability of a nuclear war in the next ten years is around 0.1 (10%) (according to international security expert Martin Hellman)

Two schools of thought:

- **Frequentist:** Probability is about averages for repeated identical events:
  - If I toss a coin many times, it will come up “heads” 0.5 of the time
  - If I look at all the days the weather forecast said 97% probability of snow, it would snow on 97% of those days
- **Bayesian:** Probability reflects my degree of confidence that something is true:
  - It makes no sense to think about repeating the 2016 election thousands of times and counting the number of times Clinton won.
  - Assigning a 71.4% probability that Clinton would win reflected what Silver knew at the time.

Frequentist Interpretation

- For **frequentists**, probabilities are about frequencies, not about truth.
  - Frequentist statistical tests report confidence with a p value:

    > *A clinical trial found that patients taking the experimental drug lived longer than patients taking a placebo, with* $p = 0.001$

    - This means that if the drug doesn't actually cause patients to live longer, if you repeated the clinical trial many times, about one time in 1000, purely by chance, the patients taking the drug would live longer than those who didn't.
    - This doesn't mean there's 99.9% confidence that the drug works.

Bayesian Interpretation

- For **Bayesians**, probabilities are about knowledge. Every time you make a new measurement, you learn something, and update the probability to reflect the new state of your knowledge.
  - Suppose you begin life not knowing whether the sun will ever rise again, after it sets.
    - After the first day, you might assume there's a 50% probability that the sun will rise again.
    - After the sun rises on the second day, you update the probability to reflect your new knowledge that you've seen it rise twice.
    - After thousands of sunrises, your updated knowledge gives a very high probability that it will continue to rise every day.

Mathematical treatment

- Let H be a hypothesis that is either true or false, and H_0 is an alternate *null hypothesis*.
  - H says patients who take the drug live longer, and H_0 says there is no difference between those who take the drug and those who don't.
- Let x be the evidence collected in an experiment
- Frequentists consider likelihood: $P(x|H_0)$: the probability that you'd see the evidence from the experiment if the null hypothesis were true
- Bayesians consider $P(H|x)$: the probability that the hypothesis is true, based on the evidence from the experiment.

Bayes's Theorem

- Suppose there are N possible hypotheses, H_1, H_2, ... H_N.
- Initially, you assign probabilities $P(H_1)$, $P(H_2)$, ..., $P(H_N)$.
  - These are called ***prior probabilities***. If you don't know anything, you might choose a **uniform prior**, giving each hypothesis the probability $1/N$.
- You observe new data D, and want to update your estimates of the probability for each hypothesis:

$$P(H_i|x) = \frac{P(x|H_i)P(H_j)}{\sum_{j=1}^{N} P(x|H_j)P(H_j)}$$

Bayes's Theorem (contd.)

$$P(H_i|x) = \frac{P(x|H_i)P(H_i)}{\sum_{j=1}^{N} P(x|H_j)P(H_j)}$$

- $P(H_i|x)$ is called the ***posterior probability*** of H_i, after updating it to account for the new data
- $P(H_i)$ is called the **prior probability_** of H_i, before observing the data
- D is the new data
- $P(D|H_i)$ is the likelihood of observing D, if H_i is true
- $\sum_{j=1}^{N} P(D|H_j)P(H_j)$, is equal to $P(D)$, the probability of observing D regardless which hypothesis is true, and it's called the ***evidence***.

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Worked Examples

Prosecutor's fallacy

- A crime has been committed in Nashville, and the person who committed it left a smudged, partial fingerprint at the scene.
- A suspect is apprehended.
  - H is the hypothesis that the suspect committed the crime
  - H_0 is the null hypothesis, that the suspect is innocent
- Fingerprint experts: the suspect's fingerprint matches the partial closely enough that there is only one chance in 100,000 an innocent person would match it that well:

$$P(x|H_0) = 0.001$$

- Does this mean there is a 99.999% chance that the suspect is the criminal?
- There are over 700,000 people in Nashville.
  - How many innocent people in Nashville would match the partial fingerprint? 7
  - How many criminals are there? 1
  - 1 criminal and 7 innocents.
    - $P(H) = 1/8$, and there's a $7/8$ probability the suspect is innocent.

The Game Show Problem

- On the old game show, “Let’s Make A Deal,” a contestant would face three doors.
  - Behind one would be a valuable prize, such as a car
  - Behind the other two would be joke prizes or nothing at all

The Game Show Problem

- After the contestant chose a door, the host, Monty Hall, would often open one of the other doors, to reveal a goat, and ask the contestant, “would you like to switch your choice?”
- What should the contestant do?

Frequentist solution

- The contestant picks door 1, and there’s a 1/3 chance there’s a car behind it.
- There’s a 2/3 chance the car is behind doors 2 or 3.
- Monty Hall opens door 2 to reveal a goat.
- There’s still a 1/3 chance the car is behind door 1, which means there’s a 2/3 chance it’s behind door 3.
  - The contestant should take the offer and pick door 3 instead.
- This logic is very confusing and many experts get it wrong.

Bayesian solution

- 3 possibilities for which door the car is behind: H_1, H_2, H_3.
- At first, you don't know anything, so $P(H_1) = P(H_2) = P(H_3) = 1/3$.
- Let's say you pick door 1, and Monty Hall opens door 2 to reveal a goat.
- D is the observation that Monty Hall opens door 2 to reveal a goat.
  - You know that he will never open the door with the car, so he's not opening doors at random.
  - If the car is behind door 1, he could open doors 2 or 3 to reveal a goat so $P(D|H_1) = 1/2$
  - The car can't be behind door 2, so $P(D|H_2) = 0$
  - If the car is behind door 3, he could only open door 2, so $P(D|H_3) = 1$

$$
\begin{aligned}
P(H_1|D) &= \frac{P(D|H_1)P(H_1)}{\sum_{i=1}^{3} P(D|H_i)P(H_i)} \\
&= \frac{1/2 \times 1/3}{1/2 \times 1/3 + 0 + 1 \times 1/3} \\
&= \frac{1/6}{1/6 + 1/3} = 1/3
\end{aligned}
$$

  - So after door 2 is opened, there is a 1/3 chance the car is beind door 1, and 2/3 chance it's behind door 3.

Bayes’s Theorem in Context

- Traditionally, most statistics has followed frequentist methods and interpretations
  - Many people get confused because it can’t tell you the probability that a hypothesis is true
- Bayes’s theorem offers a straightforward way to update probabilities when new knowledge becomes available.
  - It can be confusing to choose priors if you don’t already know much about the thing you’re studying
  - It can be difficult and computationally expensive to do the necessary calculations.
- Despite the difficulties, Bayesian methods are becoming widely used in geosciences.

Probabilities and Coin Tosses

Rosencrantz & Guildenstern are Dead

Tom Stoppard, *Rosencrantz and Guildenstern Are Dead* (1990)

Mathematical treatment

Is the Coin Fair

- If you toss a coin many times and it comes up heads a certain fraction of the time, what is the probability that it's fair (50% probability)?
- A ***frequentist*** can only tell you the probability that the coin would come up heads that many times if it were fair.
- A Bayesian can tell you the probability that the coin is fair, and can give detailed information about the probability it will come up heads, and the uncertainty about that probability.
- We'll analyze coin tosses using Bayes's theorem in Chapter 5.
- Sophisticated use of Bayesian analysis is beyond the scope of this course, so we'll still mostly be using frequentist methods.

Calculating Probabilities

Calculating Probabilities

- Break down complex problems to simpler ones
- DeMorgan's Laws:

$$E \bar{\cap} F = \bar{E} \cup \bar{F}$$
$$E \bar{\cup} F = \bar{E} \cap \bar{F}$$

- For independent, identically distributed samples:
  - **And rule:** The probability of A and B both happening:

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

  - **Or rule:** The probability of A or B happening, _if A and B are mutually exclusive:

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

Examples

- Easy: What is the probability that if I toss a coin 10 times, it will come up heads all 10 times?
  - Apply the and rule:

$$0.5 \times 0.5 \times 0.5 \times \cdots \times 0.5 = 0.5^{10} = 1/1024$$

Trickier Example:

- Trickier: A plot of land near a river floods on average once every hundred years. What is the probability that it will flood at least once in the next 30 years?
  - Either it will flood, or it won't, so use the or rule:

$$P(\text{flood at least once}) = 1 - P(\text{doesn't flood}).$$

  - And we can solve $P(\text{doesn't flood})$ using the and rule:

$$\begin{aligned}\text{Doesn't flood in 30 years} = \; & \text{Doesn't flood this year} \cap \text{Doesn't flood next year} \cap \cdots \\ & \cdots \text{Doesn't flood 30 years from now}\end{aligned}$$

  - so the probability of not flooding in the next 30 years is

$$P(\text{doesn't flood in one year})^{30} = (1 - 0.01)^{30} = 0.99^{30} = 0.74$$

  - And the probability of flooding at least once in 30 years is

$$P(\text{floods}) = 1 - P(\text{doesn't flood}) = 1 - 0.74 = 0.26$$