Quantifying Uncertainty

STA112: Probability I - *From Axioms to Bayes' Theorem*

Department of Statistics

Outline

From Counting to Probability

Recall from Module 1

We learned how to systematically count the number of outcomes in a set.

Now: We assign a numerical measure — a **probability** — to events.

Today's Big Question

How do we build a logical, consistent, and rigorous mathematical system to quantify uncertainty?

Learning Objectives:

- Define sample spaces, events, and the axioms of probability.
- Apply addition rules, conditional probability, and independence.
- Master Bayes' Theorem for "reversing" probabilities.

The Building Blocks: Experiments and Outcomes

Definition (Experiment)

Any process or action that produces an observation or outcome.

Definition (Sample Space S)

The set of all possible, mutually exclusive, and exhaustive outcomes of an experiment.

Discrete Example:

- *Experiment:* Tossing a coin.
- *Sample Space:* $S = \{H, T\}$

Continuous Example:

- *Experiment:* Measuring the lifetime of a car battery.
- *Sample Space:* $S = [0, \infty)$

Events: The Things We Care About

Definition (Event)

Any subset E of the sample space S. An event E **occurs** if the actual outcome of the experiment lies in E.

Example: Rolling a Die

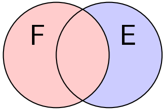
- $S = \{1, 2, 3, 4, 5, 6\}$.
- $E = \{\text{Even numbers}\} = \{2, 4, 6\}$.
- $F = \{\text{Numbers less than 3}\} = \{1, 2\}$.

We can combine events using **set operations** to model more complex ideas.

Set Operations for Events

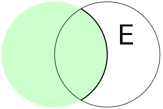
Definition (Union: $E \cup F$)

The event that E **or** F (or both) occurs.



Definition (Intersection: $E \cap F$ or EF)

The event that **both** E **and** F occur.



Definition (Complement: E^c)

The event that E does **not** occur. Everything in S that is not in E.

Definition (Mutually Exclusive (Disjoint))

E and F **cannot** occur together. Their intersection is empty: $EF = \emptyset$.

Axioms: The Bedrock Rules

How do we assign a probability $P(E)$ to every event E in a sample space S?

The Three Axioms of Probability

1. **Non-negativity:** $0 \leq P(E) \leq 1$.
2. **Probability of the Sample Space:** $P(S) = 1$.
3. **Additivity for Disjoint Events:** If $E_1, E_2, \ldots$ are mutually exclusive, then:

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n)$$

Every other rule you learn in probability is a **consequence** of these three simple axioms.

Immediate Consequences of the Axioms

Theorem (Complement Rule)

For any event E:

$$P(E^c) = 1 - P(E)$$

Derivation.

E and E^c are mutually exclusive, and $S = E \cup E^c$.

- Using Axiom 3: $P(S) = P(E \cup E^c) = P(E) + P(E^c)$.
- Using Axiom 2: $1 = P(E) + P(E^c)$.
- Therefore, $P(E^c) = 1 - P(E)$.

$\square$

Example

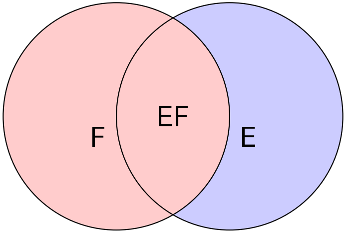
If the probability of rain tomorrow is 0.3, the probability of NO rain is $1 - 0.3 = 0.7$.

The General Addition Rule

Theorem (Addition Rule for Any Two Events)

For any events E and F (not necessarily disjoint):

$$P(E \cup F) = P(E) + P(F) - P(EF)$$



Why subtract $P(EF)$? Because the "intersection" area is counted twice when we add $P(E)$ and $P(F)$.

Solved Example: The Addition Rule

Problem

A card is drawn from a standard 52-card deck. What is the probability it is a **Heart** or a **Face Card**?

Solved Example: The Addition Rule

Problem

A card is drawn from a standard 52-card deck. What is the probability it is a **Heart** or a **Face Card**?

Solution:

- Let E = Heart. $P(E) = 13/52$.
- Let F = Face Card (J, Q, K). $P(F) = 12/52$.
- The intersection is EF = Heart AND Face Card (J, Q, K of Hearts). $P(EF) = 3/52$.

Solved Example: The Addition Rule

Problem

A card is drawn from a standard 52-card deck. What is the probability it is a **Heart** or a **Face Card**?

Solution:

- Let E = Heart. $P(E) = 13/52$.
- Let F = Face Card (J, Q, K). $P(F) = 12/52$.
- The intersection is EF = Heart AND Face Card (J, Q, K of Hearts). $P(EF) = 3/52$.

Applying the Rule

$$P(\text{Heart} \cup \text{Face}) = \frac{13}{52} + \frac{12}{52} - \frac{3}{52} = \frac{22}{52} = \frac{11}{26}$$

The Concept of Conditioning

Motivating Question

If you roll two dice and I tell you the first die is a 6, what's the probability the sum is 10?

The Concept of Conditioning

Motivating Question

If you roll two dice and I tell you the first die is a 6, what's the probability the sum is 10?

- Without information: $P(\text{sum} = 10) = 3/36 = 1/12$ (outcomes: (4,6), (5,5), (6,4)).
- With the information: The new sample space is only outcomes where the first die is 6: $\{(6,1), \dots, (6,6)\}$.
- In this reduced world, only (6,4) gives a sum of 10. So probability is $1/6$.

The Concept of Conditioning

Motivating Question

If you roll two dice and I tell you the first die is a 6, what's the probability the sum is 10?

- Without information: $P(\text{sum} = 10) = 3/36 = 1/12$ (outcomes: (4,6), (5,5), (6,4)).
- With the information: The new sample space is only outcomes where the first die is 6: $\{(6,1), \ldots, (6,6)\}$.
- In this reduced world, only (6,4) gives a sum of 10. So probability is $1/6$.

Definition (Conditional Probability)

The probability of E given that F has occurred is:

$$P(E|F) = \frac{P(EF)}{P(F)}, \quad \text{for } P(F) > 0$$

The Multiplication Rule

Theorem (Multiplication Rule)

Rearranging the definition of conditional probability gives us a way to find joint probabilities:

$$P(EF) = P(F) \cdot P(E|F)$$

This extends naturally to more events:

$$P(E_1 E_2 E_3) = P(E_1)P(E_2|E_1)P(E_3|E_1 E_2)$$

Solved Example: Drawing Balls Without Replacement

An urn has 3 red (R) and 2 blue (B) balls. We draw 2 balls. What is $P(\text{both red}) = P(R_1 R_2)$?

The Multiplication Rule

Theorem (Multiplication Rule)

Rearranging the definition of conditional probability gives us a way to find joint probabilities:

$$P(EF) = P(F) \cdot P(E|F)$$

This extends naturally to more events:

$$P(E_1 E_2 E_3) = P(E_1)P(E_2|E_1)P(E_3|E_1 E_2)$$

Solved Example: Drawing Balls Without Replacement

An urn has 3 red (R) and 2 blue (B) balls. We draw 2 balls. What is $P(\text{both red}) = P(R_1 R_2)$?

- $P(R_1) = 3/5$.
- Given the first was red, $P(R_2|R_1) = 2/4$.
- $P(R_1 R_2) = (3/5) \times (2/4) = 6/20 = 0.3$.

Independence: An Extremely Important Concept

Definition (Statistical Independence)

Events E and F are **independent** if the occurrence of one does not affect the probability of the other. Equivalently:

$$P(EF) = P(E)P(F)$$

- If $P(F) > 0$, this is also equivalent to $P(E|F) = P(E)$.

Critical Warning

Disjoint ($EF = \emptyset$) and **Independent** are totally different!

- If E and F are disjoint, they are highly **dependent**. If one happens, the probability of the other becomes ZERO.
- Independence means knowing one happened tells you *nothing* about the other.

Independence in Practice

Example

Roll two fair dice. Let $E = \{\text{sum is } 7\}$, $F = \{\text{first die is } 4\}$. Are E and F independent?

Independence in Practice

Example

Roll two fair dice. Let $E = \{\text{sum is } 7\}$, $F = \{\text{first die is } 4\}$. Are E and F independent?

Solution:

- $P(EF) = P(\{(4,3)\}) = 1/36$.
- $P(E) = 6/36 = 1/6$. $P(F) = 6/36 = 1/6$.
- Since $(1/6) \times (1/6) = 1/36 = P(EF)$, the events **ARE** independent.

Independence in Practice

Example

Roll two fair dice. Let $E = \{\text{sum is } 7\}$, $F = \{\text{first die is } 4\}$. Are E and F independent?

Solution:

- $P(EF) = P(\{(4,3)\}) = 1/36$.
- $P(E) = 6/36 = 1/6$. $P(F) = 6/36 = 1/6$.
- Since $(1/6) \times (1/6) = 1/36 = P(EF)$, the events **ARE** independent.

Intuition: Knowing the first die is a 4 doesn't make a sum of 7 any more or less likely; you still need exactly a 3 on the second die, which has a 1/6 chance.

The Law of Total Probability

Theorem (Law of Total Probability)

If $F_1, F_2, \ldots, F_n$ are mutually exclusive and exhaustive (their union is S), then for any event E:

$$P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$$

This says: the total probability of E is the **weighted average** of its conditional probabilities, weighted by the probabilities of the "causes" F_i.

Bayes' Theorem: Reversing the Condition

Theorem (Bayes' Theorem)

Using the Law of Total Probability, we can "reverse" a conditional probability:

$$P(F_j|E) = \frac{P(EF_j)}{P(E)} = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}$$

Terminology:

- $P(F_j)$: **Prior** probability (belief before evidence).
- $P(E|F_j)$: **Likelihood** of seeing evidence E.
- $P(F_j|E)$: **Posterior** probability (updated belief).

$$\text{Prior} \xrightarrow{E} \text{Posterior}$$

Solved Example: The Disease Testing Problem

Suppose 1% of a population has a disease. A test is 95% accurate for those who have the disease, and gives a false positive for 2% of healthy people. If a person tests **positive**, what is the probability they actually have the disease?

Solved Example: The Disease Testing Problem

Suppose 1% of a population has a disease. A test is 95% accurate for those who have the disease, and gives a false positive for 2% of healthy people. If a person tests **positive**, what is the probability they actually have the disease?

Step 1: Define Events & Knowns

- D = has disease. $P(D) = 0.01$.
- D^c = healthy. $P(D^c) = 0.99$.
- T^+ = tests positive.
- $P(T^+|D) = 0.95$ (True +).
- $P(T^+|D^c) = 0.02$ (False +).

Step 2: Apply Bayes' Theorem

$$\begin{aligned} P(D|T^+) &= \frac{P(T^+|D)P(D)}{P(T^+|D)P(D) + P(T^+|D^c)P(D^c)} \\ &= \frac{0.95 \times 0.01}{(0.95 \times 0.01) + (0.02 \times 0.99)} \\ &= \frac{0.0095}{0.0095 + 0.0198} = \frac{0.0095}{0.0293} \\ &\approx 0.324 \end{aligned}$$

Solved Example: The Disease Testing Problem

Suppose 1% of a population has a disease. A test is 95% accurate for those who have the disease, and gives a false positive for 2% of healthy people. If a person tests **positive**, what is the probability they actually have the disease?

Step 1: Define Events & Knowns

- D = has disease. $P(D) = 0.01$.
- D^c = healthy. $P(D^c) = 0.99$.
- T^+ = tests positive.
- $P(T^+|D) = 0.95$ (True +).
- $P(T^+|D^c) = 0.02$ (False +).

Step 2: Apply Bayes' Theorem

$$P(D|T^+) = \frac{P(T^+|D)P(D)}{P(T^+|D)P(D) + P(T^+|D^c)P(D^c)}$$

$$= \frac{0.95 \times 0.01}{(0.95 \times 0.01) + (0.02 \times 0.99)}$$

$$= \frac{0.0095}{0.0095 + 0.0198} = \frac{0.0095}{0.0293}$$

$$\approx 0.324$$

The Shocking Result

Despite the positive test, the probability of actually having the disease is only **32.4%**! This is

Module 2: Summary

The Foundation

Probability is a mathematical system built on **three axioms**. Every rule—complements, addition, conditioning—follows logically from them.

Key Concepts to Remember

1. **Conditional Probability** updates our beliefs with evidence.
2. **Independence** means evidence provides no update.
3. **Bayes' Theorem** is the engine for learning from data, taking us from a Prior $P(F)$ to a Posterior $P(F|E)$.

Looking Ahead to Module 3

"Now that we understand events and their probabilities, we will introduce **random variables**—numerical summaries of outcomes—and study their distributions."