Math Revision Session

Statistics (1): Probability

Jukina HATAKEYAMA

The University of Osaka, Department of Economics

February 13, 2025

Sets and Their Definitions

- A **set** is a collection of distinct objects, called elements.
- Notation: A set is usually denoted by a capital letter, e.g., A, B, C.
- Elements of a set are written inside curly brackets:

$$A = \{1, 2, 3, 4, 5\}$$

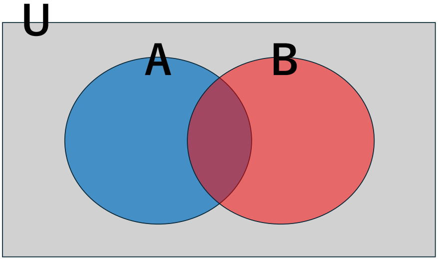
Types of Sets

- **Finite Set**: A set with a limited number of elements.
- **Infinite Set**: A set with infinitely many elements, e.g., the set of natural numbers $\mathbb{N}$.
- **Empty Set** ($\emptyset$): A set with no elements.
- **Universal Set** (U): The set containing all objects under consideration.

Set Operations

- **Union** ($A \cup B$): The set of elements that belong to either A or B (or both).
- **Intersection** ($A \cap B$): The set of elements that belong to both A and B.
- **Difference** ($A - B$): The set of elements in A but not in B.
- **Complement** (A^c): The set of elements not in A (relative to a universal set U).

Example: Venn Diagram



- The overlapping region represents $A \cap B$.
- The entire coloured area represents $A \cup B$.

Definition

Ex.) A Dice Roll

- **Sample space**: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- $\rightarrow$ The set of all possible outcomes.
- **Event**: A subset of the sample space, e.g., $A = \{1, 3, 5\}$ (rolling an odd number).
- $\rightarrow$ A collection of possible outcomes.
- **Probability**: A function that assigns a numerical value between 0 and 1 to each event, representing its likelihood.

Let Ω be the sample space representing the outcome of a dice roll:

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

For each $i \in \{1, 2, 3, 4, 5, 6\}$, define the event $A_i = \{i\}$, meaning that the result of the dice roll is i. Then, assuming a fair dice:

$$Pr(A_1) = Pr(A_2) = \cdots = Pr(A_6) = \frac{1}{6},$$
$$Pr(A_1) + Pr(A_2) + \cdots + Pr(A_6) = 1.$$

Probability Axioms

Kolmogorov's Axioms:

- **Non-negativity**: For any event A, the probability is always non-negative:

$$Pr(A) \geq 0$$

- **Normalization**: The probability of the entire sample space is 1:

$$Pr(\Omega) = 1$$

- **Additivity**: For any two mutually exclusive events A and B,

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

Consequences of Axioms

- **Monotonicity**: If $A \subseteq B$, then $Pr(A) \leq Pr(B)$.
- **Complement Rule**: The probability of the complement of A is:

$$Pr(A^c) = 1 - Pr(A)$$

- **Finite Additivity**: If $A_1, A_2, ..., A_n$ are mutually exclusive,

$$Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} Pr(A_i)$$

Joint Probability

- The **joint probability** of two events A and B is the probability that both events occur simultaneously:

$$Pr(A \cap B)$$

- If A and B are independent, then:

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

Addition Rule of Probability

- The probability of the union of two events is given by:

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

- If A and B are mutually exclusive, then:

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

Conditional Probability

- The probability of event A given that event B has occurred:

$$Pr(A \mid B) = \frac{Pr(A \cap B)}{Pr(B)}, \quad Pr(B) > 0$$

- Describes how probability updates based on new information.

Independence of Events

- Two events A and B are independent if:

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

- This means that knowing one event occurs does not affect the probability of the other.
- If they are independent, we have:

$$Pr(A|B) = Pr(A)$$

Bayes' Theorem

- Bayes' theorem expresses the probability of event A given B in terms of their conditional probabilities:

$$Pr(A \mid B) = \frac{Pr(B \mid A)Pr(A)}{Pr(B)}, \quad Pr(B) > 0$$

- Useful in statistical inference and decision-making.

Derivation

The conditional probability of event A given event B is defined as:

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

Similarly, the conditional probability of event B given event A is:

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

Both equations describe the probability of the intersection of events A and B.

From the definition of conditional probability:

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

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

Since both expressions represent $P(A \cap B)$, we can set them equal to each other:

$$P(A|B)P(B) = P(B|A)P(A)$$

Deriving Bayes' Theorem

Solving for $P(A|B)$, we obtain Bayes' Theorem:

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

This formula allows us to update our belief about A given new evidence B.

Computing $P(B)$ Using the Law of Total Probability

The denominator $P(B)$ can be computed using the law of total probability:

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$$

where A^c represents the complement of A.

Final Form of Bayes' Theorem

Substituting $P(B)$ into Bayes' Theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

This formula helps in various real-world applications such as medical testing, spam filtering, and decision-making under uncertainty.

Problem

In a hospital, the probability of having a certain disease is 1% (0.01). If a person has the disease, the probability of a positive test result is 99% (0.99). If a person does not have the disease, the probability of a positive test result is 5% (0.05). What is the probability that a person who tested positive actually has the disease?

Bayes' Theorem

Bayes' Theorem is given by the following formula:

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

Where:

- $P(A|B)$: The probability that event A occurs given that B has occurred (the probability we want to find)
- $P(B|A)$: The probability that event B occurs given that A has occurred (probability of a positive test given the disease)
- $P(A)$: The probability that event A occurs (probability of having the disease)
- $P(B)$: The probability that event B occurs (probability of a positive test result)

Given Information

The following information is given in this problem:

- $P(A) = 0.01$ (the probability of having the disease)
- $P(B|A) = 0.99$ (the probability of a positive test given the disease)
- $P(A^c) = 0.99$ (the probability of not having the disease)
- $P(B|A^c) = 0.05$ (the probability of a positive test given no disease)

Calculating $P(B)$

First, we calculate the total probability of a positive test result, $P(B)$. This can be broken down as:

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$$

Thus:

$$P(B) = (0.99 \times 0.01) + (0.05 \times 0.99) = 0.0594$$

Applying Bayes' Theorem

Now, we apply Bayes' Theorem to calculate the desired probability $P(A|B)$:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{(0.99 \times 0.01)}{0.0594}$$

After performing the calculation:

$$P(A|B) \approx 0.1667$$

Therefore, the probability that a person who tested positive actually has the disease is approximately 16.67%.

What is a Random Variable?

A random variable is a numerical outcome of a random phenomenon or experiment. It is a function that assigns a real number to each outcome in the sample space of a random experiment.

- A random variable is typically denoted by capital letters such as X, Y, or Z.
- It can take on different values depending on the outcome of the random experiment.
- The probability distribution of a random variable describes how probabilities are assigned to each of its possible values.

Types of Random Variables

Random variables can be categorized into two types:

- **Discrete Random Variable**: A random variable that can take on a finite or countable number of values. For example, the number of heads when flipping a coin multiple times.
- **Continuous Random Variable**: A random variable that can take on an infinite number of values within a given range. For example, the height of a person or the time taken for an event to occur.

Example of a Discrete Random Variable

Consider the example of rolling a fair six-sided die. The random variable X represents the outcome of the roll. The possible values of X are:

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

Each value has a probability of $\frac{1}{6}$, assuming the die is fair. The probability distribution for X is:

$$P(X = x) = \frac{1}{6}, \quad \text{for } x \in \{1, 2, 3, 4, 5, 6\}$$

Example of a Continuous Random Variable

Consider the example of measuring the height of a randomly selected person. The random variable Y represents the height, which can take any value within a certain range, say between 150 cm and 200 cm. Since the height can take infinitely many values within this range, Y is a continuous random variable.

The probability distribution for Y is represented by a probability density function (PDF), which gives the probability of the random variable taking a value within a specific range.

$$f_Y(y) = \text{probability density at height } y$$