Probability

Sean Hellingman ©

Introduction to Statistical Data Analysis (ADSC1000)

shellingman@tru.ca

Fall 2024

THOMPSON RIVERS UNIVERSITY

Topics

Introduction

- **Probability** quantifies the uncertainty that we encounter all around us.
- It is an important element of analytical modelling and decision-making.
- We use probability everywhere:
  - Predicting weather conditions.
  - Modelling the outcomes of medical trials.
  - Strategic business decisions (product launches).
  - Probabilities of winning hands in card games.
  - Determining staffing levels.

Assigning Probability

- In order to assign probability to these events we need a fundamental understanding of probability concepts and probability distributions.
- Basic concepts of probability.
- Random variables and their probability distributions.
- Discrete & Continuous probability distributions.

Basic Concepts & Definitions

- Probabilities are expressed as values between 0 and 1.
  - Often they are converted to percentages.
- An **experiment** is a process that results in some outcome.
  - Examples: rolling dice, observing the weather, or market research study.
- The collection of all possible outcomes of an experiment is called a **sample space**.

Example 1

- Identify the sample spaces for the following experiments:
  - Flipping two fair coins
  - Rolling a six sided die.
  - Flipping one fair coin and rolling one six sided die.

Defining probability

- We can define the probability in three ways:
  1. The outcomes are known and we can use theory to determine probability. (*Classical definition*)
     - Theoretical outcomes of rolling dice or flipping coins.
  2. *Relative frequency* definition based on empirical data.
     - Probabilities may change with more data or changing conditions.
  3. *Subjective* definition is based on personal judgement.
     - *I give Manchester City an 80% chance of winning the Premier League this season.*

Example 2

- Use the *classical definition* to determine the probabilities of the following events:
  - Flipping two fair coins and both resulting in heads.
  - Landing on a 5 when rolling a six sided die.
  - flipping one fair coin and rolling one six sided die and landing on a heads and five.

Example 3

- **Without giving consideration to other potentially important information**, estimate the probability that Real Madrid (RMA) will win (W) their next match using *relative frequency*.

Match	Score	Result
RMA V SEV	3:1	W
RMA V GIR	1:1	D
RAY V RMA	3:2	L
RMA V CAD	2:1	W
VLL V RMA	0:2	W

Rules and Formulas

- The probability associated with any outcome O_i must be between 0 and 1.
- The sum of the probabilities of all possible outcomes must be 1.
- More formally: Assume a sample space with m elements $O_1, O_2, ..., O_m$, then:

$$0 \leq P(O_i) \leq 1 \tag{1}$$

$$P(O_1) + P(O_2) + ... + P(O_m) = 1 \tag{2}$$

- where $P(O_i)$ is the probability of event O_i.

Rule I

- **The probability of any event is the sum of the probabilities of the outcomes that compose that event.**
- Example:
  - What is the probability that throwing two dice will result in a total score of 10 or 5?
  - $3/36 + 4/36 = 7/36$.

Rule II

- The **compliment** of an event A, denoted as A^c is all the outcomes in the sample space that are not in A.
- **The probability of the complement of any event A is $P(A^c) = 1 - P(A)$.**
- Example:
  - What is the probability that throwing two dice will **not** result in a total score of 10 or 5?
  - 1 - $7/36 = 29/36$.

Rule III

- If two events are **mutually exclusive** they have no outcomes in common.
- **If events A and B are mutually exclusive then $P(A \cup B) = P(A) + P(B)$.**
  - May also be written as $P(A \textit{ or } B)$.
- The events in the dice example are mutually exclusive.

Rule IV

- If two events are **not** mutually exclusive then adding their probabilities would lead to double counting.
- **If events** A **and** B **are not mutually exclusive then** $P(A \cup B) = P(A) + P(B)$ - $P(A \cap B)$.

Example 4

- Let $A = \{2, 11\}$ and $B = \{\text{Even number}\}$ for the sum when throwing two dice.
  - Are A and B mutually exclusive. Why or why not?
  - Calculate $P(A \cup B)$.

Conditional Probability

- **Conditional Probability** is the probability of event A, given that another event B is known to be true.

- The probability of A given B: $P(A|B)$

- More formally, the conditional probability of an event A given that event B is known to have occurred is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{3}$$

Example 5

Suppose $n = 100$ dogs were given three types of dog food, the first food they ate was recorded. The dogs were categorised in two ways, *large* or *small*. The following cross-tabulation was created:

Dog Size	Food 1	Food 2	Food 3	Total
Small Dog	25	17	21	63
Large Dog	9	6	22	37
Total	34	23	43	100

1. What is the probability that a Small Dog ate Food 2 first?
2. What is the probability that a Large Dog ate Food 3 first?

Conditional Probability

- The conditional probability equation can be rearranged:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A). \tag{4}$$

Example 6

- What is the probability of ending up with *pocket aces* in a game of Texas Hold 'Em poker?
- *Hint:* $P(A \cap B) = P(B|A)P(A)$

Independence

- Two events are said to be **independent** if the occurrence of one does not affect the probability of the occurrence of the other.

- Formally:

$$P(A|B) = P(A) \tag{5}$$

and

$$P(A \cap B) = P(A)P(B) = P(B)P(A). \tag{6}$$

Example 7

- Given you have rolled a 5 on one die, what is the probability that you roll a 4 on the second die?

Exercise 1

- Identify the sample spaces for the following experiments:
  - Flipping a fair coin and rolling an eight sided die.
  - Flipping two fair coins and rolling one six sided die.

Exercise 2

- Calculate the probabilities for the following events:
  - Flipping a heads on a fair coin and rolling a 6 or a 2 on an eight sided die.
  - Flipping two heads on two fair coins and rolling a 3 or a 4 on one six sided die.

Exercise 3

- From example 3, use *Rule II* to determine the probability that Real Madrid (RMA) does **not** win their next match $P(W^c)$.

Exercise 4

- Let $A = \{\text{Red Cards}\}$ and $B = \{\text{Face Cards}\}$ when drawing cards from a standard 52 card deck.
    - Are A and B mutually exclusive? Why?
    - Calculate $P(A \cup B)$.
- Let $A = \{\text{Clubs } \clubsuit\}$ and $B = \{\text{Red Cards}\}$ when drawing cards from a standard 52 card deck.
    - Are A and B mutually exclusive? Why?
    - Calculate $P(A \cup B)$.

Exercise 5

Suppose $n = 50$ cats were given three toys, the first toy they played with was recorded. The cats were categorised in two ways, *long haired* or *short haired*. The following cross-tabulation was created:

Cat	Ball	Bell	Mouse	Total
Long Haired	5	7	10	22
Short Haired	9	6	13	26
Total	14	13	23	100

1. What is the probability that **any** cat chose the mouse?
2. What is the probability that a cat has long hair?
3. What is the probability that a Long Haired cat choose the mouse?
4. What is the probability that a Short Haired cat **did not** choose the mouse?

Exercise 6

- What is the probability of drawing two Clubs ♣ in a row from a well shuffled standard 52 card deck (without replacement)?

Exercise 7

- Are the following events independent? (Why or why not?)
    - Probability of rolling a 6 on a die and drawing a 7 from a deck of cards.
    - The probability that the second card drawn (without replacement) from a deck of cards is red.
    - The probability of a Long haired Cat and the probability of a mouse being selected.

References & Resources

1. Evans, J. R., Olson, D. L., & Olson, D. L. (2007). *Statistics, data analysis, and decision modeling.* Upper Saddle River, NJ: Pearson/Prentice Hall.
2. Devore, J. L., Berk, K. N., & Carlton, M. A. (2012). *Modern mathematical statistics with applications (Second Edition).* New York: Springer.