SCS 1307
Probability & Statistics

Comprehensive Exam Notes with Solved Examples

Created by: M.Y. Hanan Mohamed

Based on Dr. Dilshani Tissera's Course Material

University of Colombo, School of Computing

October 4, 2025

Contents

1 Introduction to Statistics **3**
1.1 Basic Terminology . . . 3
1.2 Types of Variables . . . 3

2 Measures of Central Tendency **4**
2.1 Mean . . . 4
2.2 Median . . . 4
2.3 Mode . . . 4

3 Measures of Dispersion **5**
3.1 Range and IQR . . . 5
3.2 Quartiles and Percentiles . . . 5
3.3 Variance and Standard Deviation . . . 6
3.4 Skewness . . . 7

4 Probability Theory **8**
4.1 Basic Concepts . . . 8
4.2 Probability Axioms . . . 8

5 Conditional Probability **10**
5.1 Law of Total Probability . . . 11
5.2 Bayes' Theorem . . . 11

6 Independence **13**

7 Random Variables **15**
7.1 Probability Distribution . . . 15
7.2 Expectation (Mean) . . . 16
7.3 Variance and Standard Deviation . . . 17

8 Practice Problems **20**

9 Key Formulas Summary **23**

10 Important Notes and Tips **24**

1 Introduction to Statistics

What is Statistics?

Statistics is the science of:

- Collecting data
- Organizing data
- Analyzing data
- Presenting data
- Drawing conclusions in the best possible way

1.1 Basic Terminology

Fundamental Terms

Population The complete collection of individuals or objects that are of interest to study.

Sample A subset of the population.

Variable An observable, measurable, or recordable characteristic; the basic unit of analysis.

Observation A value that a variable assumes for a single element.

Data The set of observations collected for the variable.

Parameter A numerical summary measure used to describe a characteristic of a **population**.

Statistic A numerical summary measure used to describe a characteristic of a **sample**.

1.2 Types of Variables

Classification of Variables

1. **Categorical (Qualitative) Variables**
   - Variables having categories or classifications that are not numerical
   - Examples: Gender (Male, Female), Quality (Good, Bad), Social class (High, Medium, Low)
2. **Numerical (Quantitative) Variables**

- Variables whose values are numerical in nature
- Subdivided into:
  - **Discrete**: Countable values with gaps (e.g., number of customers)
  - **Continuous**: Uncountable values without gaps (e.g., time, weight)

2 Measures of Central Tendency

2.1 Mean

Arithmetic Mean

The average value of observations:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Weighted Mean:

$$X = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

where w_i are weights assigned to each observation.

2.2 Median

Median (M)

The value that divides the ordered data set into two equal parts.
Steps to find median:

1. Arrange all observations in order (smallest to largest)
2. If n is odd: $M = \left(\frac{n+1}{2}\right)^{\text{th}}$ value
3. If n is even: $M = \frac{\left(\frac{n}{2}\right)^{\text{th}} \text{value} + \left(\frac{n}{2}+1\right)^{\text{th}} \text{value}}{2}$

2.3 Mode

Mode

The most frequent observation in the data set.

- There may be more than one mode (bimodal, multimodal)
- There may be no mode
- The only measure used for both categorical and numerical data

Mean vs. Median

- **Symmetric distribution:** Mean = Median
- **Right-skewed:** Mean > Median
- **Left-skewed:** Mean < Median
- **Median** is resistant to outliers; **Mean** is not

3 Measures of Dispersion

3.1 Range and IQR

Range and Interquartile Range

Range: $R = \text{Maximum} - \text{Minimum}$
Interquartile Range (IQR): $\text{IQR} = Q_3 - Q_1$

- Range of the middle 50% of values
- Not affected by outliers

3.2 Quartiles and Percentiles

Quartiles

Divide the ordered data into four equal parts:

$$\begin{aligned} Q_1 &= \frac{1(n+1)}{4}\text{th value} \quad \text{(25th percentile)} \\ Q_2 &= \frac{2(n+1)}{4}\text{th value} \quad \text{(50th percentile = Median)} \\ Q_3 &= \frac{3(n+1)}{4}\text{th value} \quad \text{(75th percentile)} \end{aligned}$$

Percentiles

The k-th percentile P_k: $k\%$ of data are smaller, $(100-k)\%$ are larger.
Procedure:

1. Rank n data values from smallest to largest
2. Calculate $\frac{nk}{100}$
3. If integer a results: $P_k = \frac{a^{\text{th}} + (a+1)^{\text{th}} \text{ values}}{2}$
4. If non-integer: P_k is the next larger integer position value

Example: Finding Quartiles

Exam marks of 50 students. Find Q_1, P_{58}, and Q_3.
Ordered data: 39, 44, 47, 50, 55, 58, 58, 60, 63, 64, 64, 66, 67, 68, 68, 70, 70, 70, 72, 72, 72, 72, 74, 74, 75, 76, 77, 77, 77, 78, 78, 80, 82, 82, 83, 85, 86, 86, 88, 88, 89, 90, 90, 91, 92, 94, 95, 95, 97, 98
Solution for Q_1:

$$\begin{aligned} \frac{nk}{100} &= \frac{50 \times 25}{100} = 12.5 \\ d(Q_1) &= 13 \\ Q_1 &= 13^{\text{th}} \text{ value} = \boxed{67} \end{aligned}$$

Solution for P_{58}:

$$\begin{aligned} \frac{nk}{100} &= \frac{50 \times 58}{100} = 29 \\ P_{58} &= \frac{29^{\text{th}} + 30^{\text{th}} \text{ values}}{2} = \frac{77 + 78}{2} = \boxed{77.5} \end{aligned}$$

Solution for Q_3:

$$\begin{aligned} \frac{nk}{100} &= \frac{50 \times 75}{100} = 37.5 \\ d(Q_3) &= 38 \\ Q_3 &= 38^{\text{th}} \text{ value} = \boxed{86} \end{aligned}$$

3.3 Variance and Standard Deviation

Variance

The average of squared deviations from the mean:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$$

Standard Deviation

The positive square root of variance:

$$S = \sqrt{S^2}$$

Properties:

- If each number increases by constant c: mean increases by c, SD unchanged
- If each number multiplies by constant k: mean multiplies by k, SD multiplies by k

3.4 Skewness

Pearson's Coefficient of Skewness

$$\text{Skewness} = \frac{\text{Mean} - \text{Mode}}{\text{Standard Deviation}} = \frac{3(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

Interpretation:

- Positive skewness: Right-skewed (Mean > Median)
- Negative skewness: Left-skewed (Mean < Median)
- Zero skewness: Symmetric distribution

4 Probability Theory

4.1 Basic Concepts

Fundamental Definitions

Random Experiment
An experiment with uncertain outcome

Outcome The result of an observation

Outcome Space (Ω)
Set of all possible outcomes (sample space)

Event A subset of the outcome space

Elementary Event
An event with just one outcome

Types of Events

1. **Equally Likely Events:** Events with same probability of occurrence
2. **Mutually Exclusive Events:** $A \cap B = \emptyset$ (cannot occur simultaneously)
3. **Exhaustive Events:** $A \cup B = \Omega$ and $P(A \cup B) = 1$

4.2 Probability Axioms

Axioms of Probability

A probability function P satisfies:

1. **Axiom 1:** $P(A) \geq 0$ for any event A
2. **Axiom 2:** $P(\Omega) = 1$
3. **Axiom 3:** $P(A \cup B) = P(A) + P(B)$ when $A \cap B = \emptyset$

Properties of Probability

Property 1: $P(A') = 1 - P(A)$
Property 2 (Addition Rule):

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

Classical Probability

For equally likely outcomes:

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}$$

Example: Coin Toss

Toss a fair coin twice and observe the sequence of Heads and Tails.
Solution:

- Outcome space: $\Omega = \{HH, HT, TH, TT\}$
- Each elementary event has probability $\frac{1}{4}$
- $P(\text{at least one Head}) = P(\{HH, HT, TH\}) = \frac{3}{4}$

Example: Rolling a Die

A fair six-sided die is rolled. Find:

(a) $P(\text{even number})$

(b) $P(\text{number} > 4)$

Solution:

(a) $P(\text{even}) = P(\{2, 4, 6\}) = \frac{3}{6} = \boxed{\dfrac{1}{2}}$

(b) $P(> 4) = P(\{5, 6\}) = \frac{2}{6} = \boxed{\dfrac{1}{3}}$

5 Conditional Probability

Conditional Probability

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

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

Similarly:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) \neq 0$$

Multiplication Rule:

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

Example: Card Drawing

A heart is picked from a deck of 52 cards. Find the probability it's a picture card.
Solution: Let A = picture card, B = heart card

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{3/52}{13/52} \\ &= \frac{3}{13} = \boxed{0.231} \end{aligned}$$

There are 3 picture cards (J, Q, K) among 13 hearts.

Example: Counters in a Bag

A bag contains 10 counters: 7 green and 3 white. A counter is drawn and not replaced, then a second is drawn. Find:

(a) $P(\text{first is green})$

(b) $P(\text{first green and second white})$

(c) $P(\text{different colors})$

Solution:

(a) $P(G_1) = \frac{7}{10} = \boxed{0.7}$

(b) $P(G_1 \cap W_2) = P(G_1) \cdot P(W_2|G_1) = \frac{7}{10} \times \frac{3}{9} = \boxed{\frac{7}{30}}$

(c)

$$\begin{aligned} P(\text{different}) &= P(G_1 \cap W_2) + P(W_1 \cap G_2) \\ &= \frac{7}{10} \times \frac{3}{9} + \frac{3}{10} \times \frac{7}{9} \\ &= \frac{7}{30} + \frac{7}{30} = \boxed{\frac{7}{15}} \end{aligned}$$

5.1 Law of Total Probability

Law of Total Probability

If $B_1, B_2, \ldots, B_n$ are mutually exclusive and exhaustive events, then for any event A:

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

Special case (two events):

$$P(A) = P(A|B) \cdot P(B) + P(A|B') \cdot P(B')$$

Example: Car Starting

Ms. Ferrari's car starting probability depends on rain:

- If it rained: $P(S|R) = 0.6$
- If no rain: $P(S|R') = 0.9$
- Probability of rain: $P(R) = 0.4$

Find probability the car starts on first attempt.
Solution:

$$\begin{aligned} P(S) &= P(S|R) \cdot P(R) + P(S|R') \cdot P(R') \\ &= 0.6 \times 0.4 + 0.9 \times 0.6 \\ &= 0.24 + 0.54 \\ &= \boxed{0.78} \end{aligned}$$

5.2 Bayes' Theorem

Bayes' Theorem

If $A_1, A_2, \ldots, A_n$ are mutually exclusive and exhaustive events, and B is any event with $P(B) \neq 0$, then:

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{P(B)} = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{j=1}^{n} P(B|A_j) \cdot P(A_j)}$$

This allows us to "reverse" conditional probabilities.

Example: Medical Test

A laboratory test for a disease has:

- $P(B|A) = 0.99$ (positive test given disease)
- $P(B|A^c) = 0.005$ (positive test given no disease)
- $P(A) = 0.001$ (0.1% have the disease)

Find $P(A|B)$ = probability of having disease given positive test.
Solution:

$$\begin{aligned} P(A|B) &= \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)} \\ &= \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.005 \times 0.999} \\ &= \frac{0.00099}{0.00099 + 0.004995} \\ &= \frac{0.00099}{0.005985} \\ &= \boxed{0.165 \text{ or } 16.5\%} \end{aligned}$$

Interpretation: Even with a positive test, there's only 16.5% chance of having the disease due to its low prevalence.

6 Independence

Independent Events

Events A and B are independent if the occurrence of one does not affect the probability of the other.
Mathematical Definition:

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

Equivalent Conditions (when $P(A) > 0$ and $P(B) > 0$):

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

Properties of Independence

Property 1: If A and B are independent, then A and B' are also independent.
Property 2: For three events A, B, C to be mutually independent:

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B) \\ P(A \cap C) &= P(A) \cdot P(C) \\ P(B \cap C) &= P(B) \cdot P(C) \\ P(A \cap B \cap C) &= P(A) \cdot P(B) \cdot P(C) \end{aligned}$$

All four conditions must hold!

Example: Two Die Throws

A die is thrown twice. Find probability of 4 on first throw and odd number on second throw.
Solution:
Let $A = 4$ on first throw, $B =$ odd on second throw
The throws are independent, so:

$$\begin{aligned} P(A) &= \frac{1}{6} \\ P(B) &= \frac{3}{6} = \frac{1}{2} \\ P(A \cap B) = P(A) \cdot P(B) &= \frac{1}{6} \times \frac{1}{2} = \boxed{\frac{1}{12}} \end{aligned}$$

Verification: $P(A \cap B) = \frac{3}{36} = \frac{1}{12}$

Example: Counters with Replacement

A bag has 5 red and 7 black counters. Draw a counter, note color, replace it, then draw again. Are events A (first red) and B (second black) independent?

Solution:

$$P(A) = \frac{5}{12}$$
$$P(B) = \frac{7}{12}$$
$$P(A \cap B) = \frac{5}{12} \times \frac{7}{12} = \frac{35}{144}$$

Check: $P(A) \cdot P(B) = \frac{5}{12} \times \frac{7}{12} = \frac{35}{144} = P(A \cap B)$
Therefore, A and B are $\boxed{\text{independent}}$.

Example: At Least One Event

A fair die is thrown twice. Find probability that:

(a) Neither throw results in 4

(b) At least one throw results in 4

Solution:
Let $A = 4$ on first throw, $B = 4$ on second throw
(a) Neither throw is 4:

$$\begin{aligned} P(A' \cap B') &= P(A') \cdot P(B') \quad \text{(independence)} \\ &= \frac{5}{6} \times \frac{5}{6} \\ &= \boxed{\frac{25}{36}} \end{aligned}$$

(b) At least one 4:

$$\begin{aligned} P(\text{at least one 4}) &= 1 - P(\text{no 4}) \\ &= 1 - \frac{25}{36} \\ &= \boxed{\frac{11}{36}} \end{aligned}$$

7 Random Variables

Random Variable

A random variable X is a function from the outcome space Ω to the real numbers $\mathbb{R}$:

$$X : \Omega \to \mathbb{R}$$

Discrete Random Variable: Takes only countable values (finite or countably infinite).
Sample Space of X: $\Omega_X = \{x : X(\omega) = x \text{ for some } \omega \in \Omega\}$

Example: Coin Tosses

Toss a coin three times. Let X = number of heads, Y = number of tails.
Solution:
$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

ω	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
$X(\omega)$	3	2	2	2	1	1	1	0
$Y(\omega)$	0	1	1	1	2	2	2	3

Therefore: $\Omega_X = \Omega_Y = \{0, 1, 2, 3\}$

7.1 Probability Distribution

Probability Function

The probability function P_X of discrete random variable X is:

$$P_X(x) = \begin{cases} P(X = x) & \text{if } x \in \Omega_X \\ 0 & \text{otherwise} \end{cases}$$

Properties:

1. $P_X(x) \geq 0$ for all x
2. $\sum_{x \in \Omega_X} P_X(x) = 1$

Example: Probability Distribution

Toss a biased coin three times where $P(H) = \frac{1}{3}$ and $P(T) = \frac{2}{3}$. Let X = number of heads. Find the probability distribution of X.
Solution:
$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

Calculate probabilities using independence:

$$P(HHH) = \left(\frac{1}{3}\right)^3 = \frac{1}{27}$$
$$P(HHT) = P(HTH) = P(THH) = \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{2}{27}$$
$$P(HTT) = P(THT) = P(TTH) = \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{4}{27}$$
$$P(TTT) = \left(\frac{2}{3}\right)^3 = \frac{8}{27}$$

Probability Distribution:

x	0	1	2	3
$P_X(x)$	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$

Verification: $\frac{8+12+6+1}{27} = \frac{27}{27} = 1$

7.2 Expectation (Mean)

Expectation

The expectation (expected value or mean) of random variable X:

$$E(X) = \mu = \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\}) = \sum_{x \in \Omega_X} x \cdot P_X(x)$$

Interpretation: The long-run average value of X.

Properties of Expectation

Property 1: For constants a and b:

$$E(aX + b) = a \cdot E(X) + b$$

Property 2: For random variables X and Y on same outcome space:

$$E(X + Y) = E(X) + E(Y)$$

Property 3: For independent X and Y:

$$E(XY) = E(X) \cdot E(Y)$$

Example: Calculating Expectation

Random variable X has probability distribution:

x	-2	-1	0	1	2
$P(X = x)$	0.3	0.1	0.2	0.1	0.3

Find: (a) $E(X)$ (b) $E(X^2)$ (c) $E(3X-5)$

Solution:

(a) $E(X)$:

$$\begin{aligned} E(X) &= \sum x \cdot P(X=x) \\ &= (-2)(0.3) + (-1)(0.1) + (0)(0.2) + (1)(0.1) + (2)(0.3) \\ &= -0.6 - 0.1 + 0 + 0.1 + 0.6 \\ &= \boxed{0} \end{aligned}$$

(b) $E(X^2)$:

$$\begin{aligned} E(X^2) &= \sum x^2 \cdot P(X=x) \\ &= (4)(0.3) + (1)(0.1) + (0)(0.2) + (1)(0.1) + (4)(0.3) \\ &= 1.2 + 0.1 + 0 + 0.1 + 1.2 \\ &= \boxed{2.6} \end{aligned}$$

(c) $E(3X-5)$:

$$\begin{aligned} E(3X-5) &= 3E(X) - 5 \\ &= 3(0) - 5 \\ &= \boxed{-5} \end{aligned}$$

7.3 Variance and Standard Deviation

Variance

The variance of random variable X measures spread around the mean:

$$\text{Var}(X) = \sigma_X^2 = E[(X-\mu)^2] = \sum_{x \in \Omega_X} (x-\mu)^2 \cdot P_X(x)$$

Computational Formula:

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2$$

Standard Deviation:

$$\sigma_X = \sqrt{\text{Var}(X)}$$

Properties of Variance

Property 1: For constant c:

$$\text{Var}(cX) = c^2 \cdot \text{Var}(X)$$

Property 2: For constant c:

$$\text{Var}(X+c) = \text{Var}(X)$$

Property 3: For independent X and Y:

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$$
$$\text{Var}(X-Y) = \text{Var}(X) + \text{Var}(Y)$$

Example: Variance Calculation

Using the previous example where $E(X) = 0$ and $E(X^2) = 2.6$, find $\text{Var}(X)$ and σ_X.

Solution:

$$\begin{aligned}\text{Var}(X) &= E(X^2) - [E(X)]^2 \\ &= 2.6 - 0^2 \\ &= \boxed{2.6}\end{aligned}$$

$$\begin{aligned}\sigma_X &= \sqrt{2.6} \\ &= \boxed{1.612}\end{aligned}$$

Example: Game with Die

A game is played with a fair die. A player wins \$20 if a 2 appears, \$40 if a 4 appears, loses \$30 if a 6 appears, and neither wins nor loses otherwise. Find the expected winnings.

Solution:

Let X = winnings. The probability distribution is:

x	-30	0	20	40
$P(X=x)$	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Calculate expectation:

$$\begin{aligned}E(X) &= (-30)\left(\frac{1}{6}\right) + (0)\left(\frac{3}{6}\right) + (20)\left(\frac{1}{6}\right) + (40)\left(\frac{1}{6}\right) \\ &= -5 + 0 + \frac{20}{6} + \frac{40}{6} \\ &= -5 + \frac{60}{6} \\ &= -5 + 10 \\ &= \boxed{\$5}\end{aligned}$$

Interpretation: On average, the player wins \$5 per game.

Example: Coin Toss Game

Toss an unbiased coin once. Win $\$kifHeads, lose\$k$ if Tails. Find the standard deviation of profit X.

Solution:
Probability distribution:

x	$-k$	k
$P(X = x)$	0.5	0.5

Calculate mean and variance:

$$E(X) = (-k)(0.5) + (k)(0.5) = 0$$
$$E(X^2) = k^2(0.5) + k^2(0.5) = k^2$$
$$\text{Var}(X) = E(X^2) - [E(X)]^2 = k^2 - 0 = k^2$$
$$\sigma_X = \sqrt{k^2} = \boxed{k}$$

8 Practice Problems

Problem 1: Cards

A card is drawn at random from a standard 52-card deck. Find:

(a) $P(\text{heart})$

(b) $P(\text{not a face card})$

(c) $P(\text{number divisible by 3})$

Solution:
(a) There are 13 hearts:

$$P(\text{heart}) = \frac{13}{52} = \boxed{\frac{1}{4}}$$

(b) Face cards: 12 (J, Q, K in 4 suits), Non-face: 40

$$P(\text{not face}) = \frac{40}{52} = \boxed{\frac{10}{13}}$$

(c) Numbers divisible by 3: 3, 6, 9 in each suit = 12 cards

$$P(\text{divisible by 3}) = \frac{12}{52} = \boxed{\frac{3}{13}}$$

Problem 2: Marbles

Two marbles are drawn simultaneously from a bag with 3 red and 2 blue marbles. Find:

(a) The outcome space

(b) $P(1 \text{ red and } 1 \text{ blue})$

(c) $P(\text{at least } 1 \text{ blue})$

Solution:
(a) Outcome space (by colors):

$$\Omega = \{RR, RB, BB\}$$

Total ways to choose 2 from 5: $\binom{5}{2} = 10$
(b) Ways to get 1R and 1B: $\binom{3}{1} \times \binom{2}{1} = 6$

$$P(1R \text{ and } 1B) = \frac{6}{10} = \boxed{\frac{3}{5}}$$

(c) $P(\text{at least 1 blue}) = 1 - P(\text{no blue})$
Ways to get 2R: $\binom{3}{2} = 3$

$$P(\text{at least 1 blue}) = 1 - \frac{3}{10} = \boxed{\frac{7}{10}}$$

Problem 3: Conditional Probability

A fair coin is tossed three times. Let:

- A = first toss is Head
- B = at least two Heads

Find: (a) $P(A)$ (b) $P(B)$ (c) $P(A \cap B)$ (d) $P(A|B)$

Solution:

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$
$A = \{HHH, HHT, HTH, HTT\}$
$B = \{HHH, HHT, HTH, THH\}$

(a) $P(A) = \frac{4}{8} = \boxed{\frac{1}{2}}$

(b) $P(B) = \frac{4}{8} = \boxed{\frac{1}{2}}$

(c) $A \cap B = \{HHH, HHT, HTH\}$, so $P(A \cap B) = \boxed{\frac{3}{8}}$

(d)

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3/8}{4/8} = \boxed{\frac{3}{4}}$$

Problem 4: Independence Check

Events A and B have $P(A) = \frac{1}{3}$, $P(A|B) = \frac{1}{2}$, $P(B|A) = \frac{2}{3}$.
Find: (a) Are A and B independent? (b) $P(A \cap B)$ (c) Are they mutually exclusive? (d) $P(B)$

Solution:

(a) If independent, then $P(A|B) = P(A)$
Since $P(A|B) = \frac{1}{2} \neq \frac{1}{3} = P(A)$, they are $\boxed{\text{NOT independent}}$.

(b) Using $P(A \cap B) = P(B|A) \cdot P(A)$:

$$P(A \cap B) = \frac{2}{3} \times \frac{1}{3} = \boxed{\frac{2}{9}}$$

(c) If mutually exclusive, $P(A \cap B) = 0$
Since $P(A \cap B) = \frac{2}{9} \neq 0$, they are $\boxed{\text{NOT mutually exclusive}}$.

(d) Using $P(A \cap B) = P(A|B) \cdot P(B)$:

$$\frac{2}{9} = \frac{1}{2} \times P(B) \implies P(B) = \boxed{\frac{4}{9}}$$

Problem 5: Rectangle Dimensions

A machine cuts rectangles. Switch A selects length (2cm or 3cm), Switch B selects width (1cm, 2cm, or 3cm). Given:

(L, W)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
P	0.12	0.13	0.25	0.20	0.15	0.15

Find the probability distribution of perimeter $X = 2L + 2W$.
Solution:
Calculate perimeter for each outcome:

(L, W)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
X	6	8	10	8	10	12
P	0.12	0.13	0.25	0.20	0.15	0.15

Group by perimeter values:

$$P(X = 6) = 0.12$$
$$P(X = 8) = 0.13 + 0.20 = 0.33$$
$$P(X = 10) = 0.25 + 0.15 = 0.40$$
$$P(X = 12) = 0.15$$

Distribution of X:

x	6	8	10	12
$P(X = x)$	$\boxed{0.12}$	$\boxed{0.33}$	$\boxed{0.40}$	$\boxed{0.15}$

9 Key Formulas Summary

Measures of Central Tendency

Mean: $\bar{x} = \frac{\sum x_i}{n}$
Median: Middle value when ordered

- If n odd: $M = \left(\frac{n+1}{2}\right)^{\text{th}}$ value
- If n even: average of two middle values

Mode: Most frequent value

Measures of Dispersion

Range: $R = \max - \min$
IQR: $\text{IQR} = Q_3 - Q_1$
Variance: $S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$
Standard Deviation: $S = \sqrt{S^2}$
Skewness: $\frac{3(\text{Mean}-\text{Median})}{S}$

Probability Rules

Addition Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Conditional Probability: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
Multiplication Rule: $P(A \cap B) = P(A|B) \cdot P(B)$
Independence: $P(A \cap B) = P(A) \cdot P(B)$
Law of Total Probability: $P(A) = \sum_i P(A|B_i)P(B_i)$
Bayes' Theorem: $P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}$

Random Variables

Expectation: $E(X) = \sum x \cdot P(X = x)$
Properties:

- $E(aX + b) = aE(X) + b$
- $E(X + Y) = E(X) + E(Y)$

Variance: $\text{Var}(X) = E(X^2) - [E(X)]^2$
Properties:

- $\text{Var}(cX) = c^2\text{Var}(X)$
- $\text{Var}(X + c) = \text{Var}(X)$
- If independent: $\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$

10 Important Notes and Tips

Exam Tips

1. **Always check units:** Ensure your final answer has appropriate units
2. **Show all work:** Partial credit requires showing calculations
3. **Verify probability axioms:** All probabilities should be between 0 and 1, and sum to 1
4. **Check independence:** Don't assume independence unless stated or proven
5. **Use complement rule:** Often easier to calculate $P(A')$ than $P(A)$
6. **Draw diagrams:** Venn diagrams and tree diagrams help visualize problems
7. **Label clearly:** Define all events and random variables explicitly
8. **Double-check arithmetic:** Simple calculation errors are common under pressure

Common Mistakes to Avoid

- Confusing $P(A|B)$ with $P(B|A)$ - these are generally NOT equal
- Assuming independence without justification
- Forgetting to use $n - 1$ in sample variance formula
- Mixing up mutually exclusive and independent events (they are different!)
- Not ordering data before finding median or percentiles
- Forgetting that $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$ unless independent
- Confusing population parameters with sample statistics
- Incorrectly applying Bayes' theorem without checking all conditions

Key Distinctions

Independent vs. Mutually Exclusive:

- **Independent:** $P(A \cap B) = P(A) \cdot P(B)$ (occurrence of one doesn't affect the other)
- **Mutually Exclusive:** $P(A \cap B) = 0$ (cannot occur simultaneously)
- If A and B are mutually exclusive with $P(A) > 0$ and $P(B) > 0$, they CANNOT be independent!

Parameter vs. Statistic:

- **Parameter:** Describes population (usually unknown) - e.g., μ, σ

- **Statistic:** Describes sample (calculated from data) - e.g., $\bar{x}$, s

Discrete vs. Continuous:

- **Discrete:** Countable values (0, 1, 2, ...)
- **Continuous:** Any value in an interval (time, weight, height)

Good Luck with Your Exam!

"Success is the sum of small efforts repeated day in and day out."
- Robert Collier

Prepared by: M.Y. Hanan Mohamed
University of Colombo School of Computing
SCS 1307 - Probability & Statistics