

Compulsory Task 1

1. State whether the mean, median, or mode would be useful in the following scenarios:

a. You are doing population statistics. You are asked to give an estimate of the typical income of a single person in the country. There is one snag: wealth distribution is out of whack, and 10% of the population holds 70% of the nation's wealth.

In this scenario, the median would be a more useful measurement since wealth is highly skewed. The mean is distorted to higher values by the presence of a small group of individuals holding a significant portion of the wealth. On the other hand, the median provides a better representation of the typical income for a person in a country. The mode would not be useful here as there might be a clear mode income data.

b. You are running a restaurant, and you are reviewing your menu. You have a list of all orders over the last six months. You are trying to find out which item you should keep based on what customers seem to like the most.

In this scenario, the mode would be the most useful measurement as it identifies the most frequent ordered dishes. Additionally, it provides insights into customer preferences by indicating the items that the mean and median do not directly reveal. While the mean and median can help identify the most popular dish, the mode offers a more comprehensive understanding of customer preferences.

c. You have been buying electricity once a month for the first six months of the year. You are trying to budget your electricity for the rest of the year and therefore need to estimate how much you will spend for the remainder of the year.

In this scenario, the mean would be the most useful measurement. You are looking to average your monthly expenditure over time, which makes the mean the most appropriate measure for estimating future spending.

d. You work in healthcare insurance. You are asked to provide an estimate of the typical amount of money spent on healthcare. This is taking into account the fact that there are a few people who spend a large amount of money on medical healthcare due to major issues.

In this scenario, the median would be the most useful measurement. Few people have extremely high healthcare costs, and their mean would be skewed. The median provides a better representation of the typical healthcare spending of the majority of people. The mode would not be useful in this context.

2. Variance vs Standard Deviation

a. You are choosing a new Internet provider. You find two providers with the same mean speed, but you want to have a more stable connection. You get a list of all reported speeds over the last month and are trying to find the provider that doesn't move too much from the mean value.

Standard deviation is the appropriate measure to use.

Compulsory Task 2

1. Conditional Probabilities

a. You work for a risk analysis insurer. You have read that this year, out of all drivers on the road, 5% have had accidents under the age of 25. You have also read that 10% of all drivers are under the age of 25. A new client approaches you and states that their age is 22. You want to calculate the chance that this driver has had an accident this year based on their age.

$P(B)$: Probability of being under 25 = 10% (0.10).

$P(A \text{ and } B)$: Probability of being under 25 and having an accident = 5% (0.05).

$P(A|B)$: $0.05/0.10 = 0.5$, 50% chance a 22-year-old driver had an accident this year

b. Your friend told you that they would buy you lunch if you can flip a coin and have it land on heads twice. You flip it the first time, and it lands on heads. What are your chances now of it landing on heads again?

$P(B)$: Probability of first flip being heads = 50% (0.5).

$P(A \text{ and } B)$: Probability of both flips being heads = 25% (0.25).

$P(A|B)$: $0.25/0.50$, 50%.

c. You were always told that knowing Maths helps you to achieve 80% in Computer Science. You read some statistics showing that 30% of all Computer Science graduates took Maths and achieved 80%. Overall, 60% of all Computer Science graduates took Maths. Considering you took Maths, what are your chances of achieving 80%?

$P(B)$: Probability of taking Maths = 60% (0.60).

$P(A \text{ and } B)$: Probability of taking Maths and achieving 80% = 30% (0.30).

$P(A|B)$: $0.30/0.60 = 0.50 \rightarrow$ 50% chance of scoring 80% in CS if you took Maths

d. Bayes' Theorem

$P(H)$: Probability of actually having COVID

Based on the diagram: 70 out of 100

$P(H) = 0.7$

$P(E|H)$: Probability of testing positive if you have COVID

Based on diagram: 63 out of 100 infected people tested positive

$$P(E|H) = 0.63$$

P(E): Overall probability of testing positive

From diagram: 63 true positives + 1 false positives = 64 positive results

$$P(E) = 0.64$$

P(H|E): Probability of actually having COVID given a positive test

$$P(H|E): (0.63 \times 0.70) / 0.64 = 0.689, 68.9\%$$