

STATISTICS WORKSHEET-3

Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

Which of the following is the correct formula for total variation?

- a) Total Variation = Residual Variation – Regression Variation
- b) Total Variation = Residual Variation + Regression Variation
- c) Total Variation = Residual Variation * Regression Variation
- d) All of the mentioned

Ans.(b) Total Variation = Residual Variation + Regression Variation

2. Collection of exchangeable binary outcomes for the same covariate data are called _____ outcomes.

- a) random
- b) direct
- c) binomial
- d) none of the mentioned

Ans.(c) binomial

3. How many outcomes are possible with Bernoulli trial?

- a) 2
- b) 3
- c) 4
- d) None of the mentioned

Ans.(a) 2

4. If H_0 is true and we reject it is called

- a) Type-I error
- b) Type-II error
- c) Standard error
- d) Sampling error

Ans(a) Type-I error

5. Level of significance is also called:

- a) Power of the test
- b) Size of the test
- c) Level of confidence
- d) Confidence coefficient

Ans.(a) Power of the test

6. The chance of rejecting a true hypothesis decreases when sample size is:

- a) Decrease
- b) Increase
- c) Both of them
- d) None

Ans(b) Increase

7. Which of the following testing is concerned with making decisions using data?

- a) Probability
- b) Hypothesis
- c) Causal
- d) None of the mentioned

Ans.(b) Hypothesis

8. What is the purpose of multiple testing in statistical inference?

- a) Minimize errors
- b) Minimize false positives

- c) Minimize false negatives
- d) All of the mentioned

Ans.(a) Minimize errors

9. Normalized data are centred at and have units equal to standard deviations of the original data

- a)0
- b)5
- c)1
- d)10

Ans.(a).0

Q10 and Q15 are subjective answer type questions, Answer them in your own words briefly.

10. What Is Bayes' Theorem?

Ans. In statistics and probability theory, the Bayes' theorem (also known as the Bayes' rule) is a mathematical formula used to determine the conditional probability of events. Essentially, the Bayes' theorem describes the [probability](#) of an event based on prior knowledge of the conditions that might be relevant to the event.

The theorem is named after English statistician, Thomas Bayes, who discovered the formula in 1763. It is considered the foundation of the special statistical inference approach called the Bayes' inference.



The diagram illustrates the formula for Bayes' Theorem. On the left, an orange square contains the expression $P(A|B)$. This is followed by an equals sign. To the right of the equals sign, there is a fraction. The numerator of the fraction consists of two teal squares: the first contains $P(B|A)$ and the second contains $P(A)$, with a large dark blue multiplication symbol (\times) between them. The denominator of the fraction is a single teal square containing $P(B)$.

Besides [statistics](#), the Bayes' theorem is also used in various disciplines, with medicine and pharmacology as the most notable examples. In addition, the theorem is commonly employed in different fields of finance. Some of the applications include but are not limited to, modeling the risk of lending money to borrowers or forecasting the probability of the success of an investment.

Formula for Bayes' Theorem

The Bayes' theorem is expressed in the following formula:

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

Where:

$P(A|B)$ – the probability of event A occurring, given event B has occurred

$P(B|A)$ – the probability of event B occurring, given event A has occurred

$P(A)$ – the probability of event A

$P(B)$ – the probability of event B

Note that events A and B are [independent events](#) (i.e., the probability of the outcome of event A does not depend on the probability of the outcome of event B).

A special case of the Bayes' theorem is when event A is a [binary variable](#). In such a case, the theorem is expressed in the following way:

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

Where:

$P(B|A^-)$ – the probability of event B occurring given that event A^- has occurred

$P(B|A^+)$ – the probability of event B occurring given that event A^+ has occurred

In the special case above, events A^- and A^+ are mutually exclusive outcomes of event A.

Example of Bayes' Theorem

Imagine you are a financial analyst at an investment bank. According to your research of [publicly-traded companies](#), 60% of the companies that increased their share price by more than 5% in the last three years replaced their [CEOs](#) during the period.

At the same time, only 35% of the companies that did not increase their share price by more than 5% in the same period replaced their CEOs. Knowing that the probability that the stock prices grow by more than 5% is 4%, find the probability that the shares of a company that fires its CEO will increase by more than 5%.

Before finding the probabilities, you must first define the notation of the probabilities.

$P(A)$ – the probability that the stock price increases by 5%

$P(B)$ – the probability that the CEO is replaced

$P(A|B)$ – the probability of the stock price increases by 5% given that the CEO has been replaced

$P(B|A)$ – the probability of the CEO replacement given the stock price has increased by 5%.

Using the Bayes' theorem, we can find the required probability:

$$P(A|B) = \frac{0.60 \times 0.04}{0.60 \times 0.04 + 0.35 \times (1 - 0.04)} = 0.067 \text{ or } 6.67\%$$

Thus, the probability that the shares of a company that replaces its CEO will grow by more than 5% is 6.67%.

11. What is z-score?

Ans. A z-score is a numerical measurement that describes a value's relationship to the mean of a group of values. Z-score is measured in terms of standard deviation from the mean. If a z-score is 0, it indicates that the data points score is identical to the mean score.

12. .What is t-test?

Ans. A t-test is a statistical test that compares the means of two samples. It is used in hypothesis testing, with a null hypothesis that the difference in group means is zero and an alternate hypothesis that the difference in group means is different from zero.

13. .What is percentile?

Ans. In statistics, a percentile is a term that describes how a score compares to other scores from the same set. While there is no universal definition of percentile, it is commonly expressed as the percentage of values in a set of data scores that fall below a given value.

14. What is ANOVA?

Ans. Analysis of variance (ANOVA) is a statistical technique that is used to check if the means of two

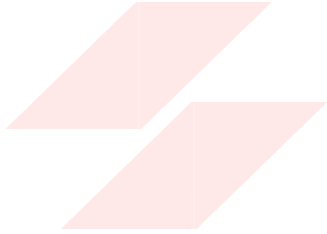
or more groups are significantly different from each other. ANOVA checks the impact of one or more factors by comparing the means of different samples.

15. How can ANOVA help?

Ans. ANOVA is helpful for **testing three or more variables**. It is similar to multiple two-sample t-tests. However, it results in fewer type I errors and is appropriate for a range of issues. ANOVA groups differences by comparing the means of each group and includes spreading out the variance into diverse sources.

For example, let's say you want to know if there's a difference between the average heights of four different types of trees in a forest. Instead of calculating whether each pair is statistically different from one another, you could run one ANOVA test to find out whether any of them are significantly different from one another. We can use ANOVA to test for statistical differences between two or more groups to see if there is a significant difference between the means of those groups. ANOVA determines whether a test is valid by looking at the variation between and within groups.

If a test shows a large standard deviation between groups, then the differences are likely due to random chance; however, if the standard deviation within groups is large, then it may be due to real differences between groups.



FLIP ROBO