

## STATISTICS - WORKSHEET 3

Khushboo Pandey  
(FlipRobo Internship 26)

### Marked Answers in Bold

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

1. 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

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

3. How many outcomes are possible with Bernoulli trial?

**a) 2**

b) 3

c) 4

d) None of the mentioned

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

5. Level of significance is also called:

a) Power of the test

**b) Size of the test**

c) Level of confidence

d) Confidence coefficient

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

a) Decrease

**b) Increase**

c) Both of them

d) None

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

a) Probability

**b) Hypothesis**

c) Causal

d) None of the mentioned

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**

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

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

10. What Is Bayes' Theorem?

**Ans.** It is a mathematical formula for determining conditional probability. Conditional probability is the likelihood of an outcome occurring, based on a previous outcome occurring.

$$P(A|B) = [P(B|A) * P(A)] / P(B)$$

A, B = events

P(A|B) = probability of A given B is true

P(B|A) = probability of B given A is true

P(A), P(B) = the independent probabilities of A and B

11. What is z-score?

**Ans.**  $z = (x - \mu) / \sigma$

Z = standard score

x = observed value

$\mu$  = mean of the sample

$\sigma$  = standard deviation of the sample

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 deviations from the mean. If a Z-score is 0, it indicates that the data point's score is identical to the mean score. A Z-score of 1.0 would indicate a value that is one standard deviation from the mean. Z-scores may be positive or negative, with a positive value indicating the score is above the mean and a negative score indicating it is below the mean.

12. What is t-test?

**Ans.** A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. The t-test is one of many tests used for the purpose of hypothesis testing in statistics

$$t = (m - \mu) / (s / \sqrt{n})$$

t = Student's t-test

m = mean

$\mu$  = theoretical value

s = standard deviation

{n} = variable set size

13. What is percentile?

**Ans.** In statistics, a percentile is a score below which a given percentage of scores in its frequency distribution fall or a score at or below which a given percentage fall. For example, the 50th percentile

is the score below which 50% or at or below which 50% of the scores in the distribution may be found.

14. What is ANOVA?

**Ans.** Analysis of variance (ANOVA) is an analysis tool used in statistics that splits an observed aggregate variability found inside a data set into two parts: systematic factors and random factors. The systematic factors have a statistical influence on the given data set, while the random factors do not.

**$F = \text{MST} / \text{MSE}$**

F → Anova Coefficient

MST → Mean Sum of Squares due to treatment.

MSE → Mean Sem of Squares due to error.

15. How can ANOVA help?

**Ans.** The ANOVA test is used to determine the influence that independent variables have on the dependent variable in a regression study.