

First and last name

Question 1/20

What do you expect will happen with bias and variance as you increase the size of training data?

- A. Bias increases and Variance increases
- B. Bias decreases and Variance decreases
- C. Bias decreases and Variance increases
- D. Bias increases and Variance decreases

Question 2/20

Efficient portfolios can be defined as those portfolios which for a given level of risk provides

- A. average return
- B. maximum return
- C. minimum return
- D. no gain

Question 3/20

Which of the following is a true statement, for comparing the t distributions with standard normal,

- A. The proportion of area beyond a specific value of "t" is less than the proportion of normal curve
- B. None of the Above
- C. The Normal Curve is symmetrical whereas the t-distributions are slightly skewed
- D. Greater the degree of freedom, the more the t-distribution resembles the standard normal distribution

Question 4/20

Which one of these statistics is unaffected by outliers?

- A. range
- B. interquartile range
- C. standard deviation
- D. Mean

Question 5/20

Skewness of Normal distribution is _____

- A. undefined
- B. positive
- C. negative
- D. 0

Question 6/20

Of what is p the probability if the null hypothesis were true?

- A. p is the probability that the results would be replicated if the experiment was conducted a second time.
- B. p is the probability that the results are due to chance, the probability that the null hypothesis (H_0) is true.
- C. p is the probability that the results are not due to chance, the probability that the null hypothesis (H_0) is false.
- D. p is the probability of observing a test statistic at least as big as the one we have if there were no effect in the population (i.e., the null hypothesis were true).

Question 7/20

Any measure indicating the centre of a set of data, arranged in an increasing or decreasing order of magnitude, is called a measure of:

- A. Central tendency
- B. Symmetry
- C. Skewness
- D. Dispersion

Question 8/20

Suppose you are training a linear regression model. Now consider these points.

1. Overfitting is more likely if we have less data
 2. . Overfitting is more likely when the hypothesis space is small. Which of the above statement(s) are correct?
- A. 1 is false and 2 is true
 - B. 1 is true and 2 is false
 - C. both are true
 - D. both are false

Question 9/20

_____ is a metric to measure how often a randomly chosen element would be incorrectly identified.

- A. Gini Index
- B. Random probability
- C. Entropy
- D. Information Gain

Question 10/20

_____ Statistics uses the data to provide descriptions of the population, either through numerical calculations or graphs or tables.

- A. Qualitative
- B. Inferential
- C. Descriptive
- D. Quantitative

Question 11/20

The pacf is necessary for distinguishing between

- A. An AR and an MA model
- B. An AR and an ARMA model
- C. Different models from within the ARMA family
- D. An MA and an ARMA model

Question 12/20

. What will a factor loading in an orthogonal solution represent?

- A. standard deviation
- B. correlation
- C. covariance
- D. eigenvalues

Question 13/20

In statistical testing of the hypothesis, what happens to the region of rejection when the level of significance α is reduced?

- A. The rejection region is unaltered
- B. The rejection region is increased in size
- C. The rejection region is reduced in size
- D. The answer depends on the value of β

Question 14/20

_____ is an example of a strategy used to reduce the likelihood of committing statistical error.

- A. Excluding outliers in analysis
- B. Altering or otherwise changing the data
- C. Filling in missing data
- D. Including outliers in analysis

Question 15/20

Read the statements given below. Identify the right option from the following for pie chart.

Statement A: To make a pie chart with Matplotlib, we can use the `plt.pie()` function.

Statement B: The `autopct` parameter allows us to display the percentage value using the Python string formatting.

- A. Both the statements are correct
- B. Both the statements are wrong
- C. Statement A is correct
- D. Statement B is correct

Question 16/20

In a linear regression problem, we are using R-squared to measure goodness-of-fit. We add a feature in linear regression model and retrain the same model. Which of the following option is true?

- A. if r squared increases, this variable is significant.
- B. individually r squared cannot tell about variable importance. we cant say anything about it right now.
- C. if r squared decreases, this variable is not significant.
- D. none of these

Question 17/20

What does dimensionality reduction reduce?

- A. entropy
- B. stochastics
- C. collinerity
- D. performance

Question 18/20

Which of the following option is true regarding “Regression” and “Correlation” ?Note: y is dependent variable and x is independent variable.

- A. The relationship is symmetric between x and y in both.
- B. The relationship is not symmetric between x and y in case of correlation but in case of regression it is symmetric.
- C. The relationship is symmetric between x and y in case of correlation but in case of regression it is not symmetric.
- D. The relationship is not symmetric between x and y in both.

Question 19/20

The .ppf() function represents

- A. percentage change in column value
- B. the normal distribution value for which a given probability is the required value.
- C. the probability for a given normal distribution value.
- D. standard normal distribution z - score

Question 20/20

The mean = np and the standard deviation = \sqrt{npq} for

- A. binomial distribution
- B. normal distribution
- C. exponential distribution
- D. poisson distribution