

STATISTICS WORKSHEET- 6

1. (d) All of the mentioned

2. (a) Discrete

3. (a) pdf

4. (c) mean

5. (a) variance

6. (a) variance

7. (c) 0 and 1

8. (b) bootstrap

9. (b) summarized

10. Histograms are a special kind of bar graph that shows a bar for a range of data values instead of a single value. A box plot is a data display that draws a box over a number line to show the interquartile range of the data. The 'whiskers' of a box plot show the least and greatest values in the data set.

11. Metrics can be selected based on problems. In the problems where you have to focus on the minimal error there you can use MSE and where you have to judge whether the features that you are using are reasonable enough for making prediction of the target variable, in these cases use R^2 score or Adjusted R^2 . this is for regression

For classification, use accuracy if your data is balanced, else you can use precision, recall or f1 score if data is not balanced.

12. To assess statistical significance, you would use hypothesis testing. The null hypothesis and alternate hypothesis would be stated first. Second, you'd calculate the p-value, which is the likelihood of getting the test's observed findings if the null hypothesis is true. Finally, you would select the threshold of significance (α) and reject the null hypothesis if the p-value is smaller than the α — in other words, the result is statistically significant

13. a) Allocation of wealth among individuals.

b) Distributions of income

c) Distributions of house prices

14. When a marks of 11 students given, In this case, it can be observed that some observations are very far from the other observations.

Therefore, the mean will not be an appropriate measure of central tendency but the median will be an appropriate measure of central tendency.

15. The chance of something happening or how likely something is to happen is called as Likelihood.