# **WORKSHEET-4**

## **STATISTICS\_ANSWERSHEET-4**

- 1. (d) All of the mentioned
- 2. (a) Discrete
- 3. (a) PDF
- 4. (c) MEAN
- 5. (c) Empirical Mean
- 6. (a) Variance
- 7. (c) 0 and 1
- 8. (b) Bootstrap
- 9. (b) Summarized

## **BoxPlot and Histogram:**

- ♦ Both help to visualize and describe numeric data.
- ◆ Histogram are better in determining the underlying distribution of the data whereas Boxplot allow you to compare multiple data sets better than Histogram as they take up less space.
- Boxplot usually provides the median and clearly separates the points that are considered outliers whereas histogram doesn't.

◆ Histogram indicate the whole frequency distribution of a variable whereas the Boxplot summarizes its most prominent features which includes Median and Outliers.

#### 11. Below action you may take while selecting metrics in stats.

- ◆ Prioritize Objectives
- ◆ Examine which metric consistently predicts their achievement.
- ◆ Identify which activities influence predictors in that order and continuously re-evaluate this process to keep up with the times.

### 12. Steps in Testing for STATISTICAL significance.

- ♦ State the Research Hypothesis.
- ◆ State the Null Hypothesis.
- ◆ Select a Probability of Alpha Level.
- ◆ Select and compute the test for statistical significance.
- ♦ Interpret the Results.
- 13. Exponential distributions do not have a log-normal distribution or a Gaussian distribution. In fact, any type of data that is categorical will not have these distributions as well. Example: Duration of a phone car, time until the next earthquake.

- 14. Income is the classic example of when to use the median instead of the mean because its distribution tends to be skewed. The median indicates that half of all incomes fall below 27581, and half are above it. For these data, the mean overestimates where most household incomes fall.
- 15. The likelihood is the probability that a particular outcome is observed when the true value of the parameter is, equivalent to the probability mass on; it is not a probability density over the parameter.