Statistics Worksheet 3

1. B)
2. C)
3. A)
4. A)
5. A)
6. B)
7. B)
8. D)
9. A)
10. Bayes' theorem, named after 18th-century British mathematician Thomas Bayes, is a mathematical formula for determining conditional probability. Conditional probability is the likelihood of an outcome occurring, based on a previous outcome occurring.

Its formula is as follows:

*P*(*A*∣*B*)=*P*(*B*)*P*(*A*⋂*B*)​=*P*(*B*)*P*(*A*)⋅*P*(*B*∣*A*)​

1. Z-scores are a way to compare results to a “normal” population. 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.

Formula: z = (x – μ) / σ

1. 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. It is mostly used when the data sets, like the data set recorded as the outcome from flipping a coin 100 times, would follow a normal distribution and may have unknown variances. A t-test is used as a hypothesis testing tool, which allows testing of an assumption applicable to a population.
2. Percentile is defined as a score *below which* a given percentage of scores in its frequency distribution falls.
3. 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.
4. When ywe want to test a particular hypothesis, we would use ANOVA to help you understand how our different groups respond, with a null hypothesis for the test that the means of the different groups are equal. If there is a statistically significant result, then it means that the two populations are unequal (or different).