Statistics Worksheet 1

Q1.

→(A) True

Q2.

→(A) Central Limit Theorem

Q3.

→ (B) Modeling bounded count data

Q4.

→ (D) All of the mentioned

Q5.

→(C) Poisson

Q6.
→(B) False

Q7.
→(B) Hypothesis

Q8.
→(A) 0

Q9.
→(C) Outliers cannot conform to the regression relationship

Q10.
→Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph form, normal distribution will appear as a bell curve.

Q11.
→Missing data can be dealt with in a variety of ways. I believe the most common reaction is to ignore it. Choosing to make no decision, on the other hand, indicates that your statistical programme will make the decision for you.

Imputation is another popular method among people who pay attention. Imputation is the process of filling in missing values with an estimate and then analysing the full data set as if the imputed values were the true observed values.

The following are some of the most prevalent methods:
a) Mean imputation
b) Substitution
c) Regression imputation

Q12.
→A/B testing (sometimes referred to as bucket testing or split-run testing) is a user experience research technique. A/B tests are randomised tests having two variations, A and B. It covers statistical hypothesis testing, sometimes known as "two-sample hypothesis testing" in statistics. A/B testing compares two variations of a single variable, usually by comparing a subject's response to variant A versus variant B and determining which is more effective.

Q13.
→ True, imputing the mean preserves the mean of the observed data. So if the data are missing completely at random, the estimate of the mean remains unbiased. That’s a good thing. Plus, by imputing the mean, you are able to keep your sample size up to the full sample size. That’s good too. This is the original logic involved in mean imputation. If all you are doing is estimating means (which is rarely the point of research studies), and if the data are missing completely at random, mean imputation will not bias your parameter estimate.

Q14.
→ Linear regression is a basic and commonly used type of predictive analysis. The overall idea of regression is to examine two things:
(1) does a set of predictor variables do a good job in predicting an outcome (dependent) variable?
(2) Which variables in particular are significant predictors of the outcome variable, and in what way do they–indicated by the magnitude and sign of the beta estimates–impact the outcome variable?
These regression estimates are used to explain the relationship between one dependent variable and one or more independent variables.

Q15.
→ The two main branches of statistics are descriptive statistics and inferential statistics. Both of these are employed in scientific analysis of data and both are equally important for the statistics.