Statistical Worksheet

Q.1 Answer ; A) True

Q.2 Answer ; C) Centroid Limit Theorem

Q.3 Answer ; B) Modeling bounded count data

Q.4 Answer ; A) The exponent of a normally distributed random variables follows what is called the log- normal distribution

Q.5 Answer ; C) Poisson

Q.6 Answer ; B) False

Q.7 Answer ; B) Hypothesis

Q.8 Answer ; A) 0

Q.9 Answer ;

10. Answer :

Normal Distribution : A distribution is said to be normal distribution if the curve of the data is symmetrical or the mean , median and mode of the distribution falls at the same point .

11. Answer :

To handel the missing values , first I check wheather the dataset has the missing value or not by the help of isnull().sum().

Then I write a program as –

```
# from scipy.stats import zscore

 Z = numpy.abs(zscore(xyz_dataset)

Z
```

Statistical Worksheet

Threshold = 3

Print(numpy.where(Z>3))

12.Answer ;

13.Answer ; Throught Mean imputation of missing data is a straightforward method , It is not acceptable as the best practice .

14.Answer ; In statistics , Linear Regression is a powerful tool for understanding and predicting the relationships between variables . Essentially it fits a straight line to the data on the x – y plane in a way that it minimize the errors between predicted and actual values .

I f n input and n output are there , then the Linear regression formula is

Y = a + b1X1 + b2X2 + …… + bnXn + (error)

15.Answer :

Mainly there are two branches are there of statistics . Those are ---

1. Descriptive statistics
2. Inferential statistics

1. Descriptive statistics : It is devided in to two types -------

    a) Central Tendency : Mean , Median, mode
    b) Dispersion of data : Range , Quartile deviation , Variance , standard deviation , skewness .

2. Inferential statistics : It use Hypothesis testing , regression analysis etc to draw conclusions about a population on a sample of data

---------- x ----------