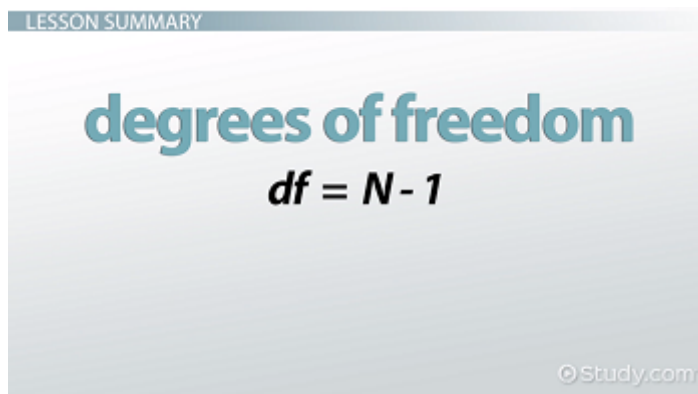## 81. What is the meaning of degrees of freedom (DF) in statistics?

Ans.

Degree of freedom means randomness. Means the maximum number of independent values, which have freedom to vary in the sample data. We calculate degree of freedom by subtracting one from n number of samples.

82. What is the p-value in hypothesis testing?

Ans.

A p-value is a number that describes the probability of finding the observed or more extreme results when the null hypothesis (H0) is True. P-values are used in hypothesis testing to help decide whether to reject the null hypothesis or not. The smaller the p-value, the stronger the evidence that you should reject the null hypothesis.

83. How can you calculate the p-value?

Ans.

We can calculate p-value in two ways

```
   i.      In two-tailed test when population standard deviation is given we calcu
   late z-score, and get the both tail value         from z table. The summation
   of the values will be the value of P.

   ii.     We can calculate using degree pf freedom. By getting the value from t-t
   able when population data is not given.
```

84. If there is a 30 percent probability that you will see a supercar in any 20-minute time interval, what is the probability that you see at least one supercar in the period of an hour (60 minutes)?

Ans.

The probability of seeing at-least a supercar in 20 minutes interval = 0.3

The probability of not seeing at-least a supercar in 20 minutes interval = 1 – 0.3 = 0.7

The probability of not seeing at-least a supercar in 60 minutes interval = $(0.7)^3 = 0.343$

The probability of seeing at-least a supercar in 60 minutes interval = $1 - 0.343 = 0.657$

85. How would you describe a 'p-value'?

Ans.

P-value always represents the significance level. It tells you how many values are not contributing out of whole experiments. (in general words p-value tells you how many experiments are going to fail out of 100)
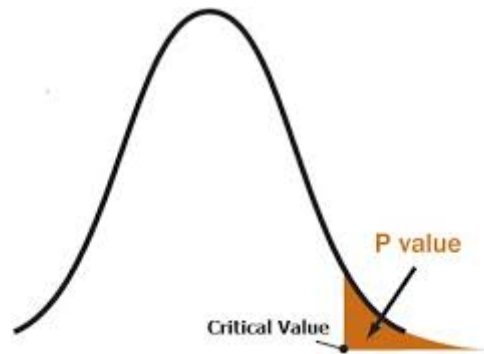
86. What is the difference between type I vs type II errors?

Ans.

Type I error is known as false positive; it occurs when as investigator rejects a null hypothesis that is True. Type II error is known as false negative; it occurs when an investigator fails to reject a null hypothesis that is actually False in the population.

87. When should you use a t-test vs a z-test?

Ans.

We should use t-test instead of z-test when

i.   Population standard deviation is not given

ii.   f l i t20

88. What is the difference between the f test and anova test?

Ans.

The analysis of variance (ANNOVA) is a process of assigning the total variation into its components.

F-test is used for testing any null hypothesis is based on equality of any two variances that may not be the result of an ANNOVA.

89. What is Resampling and what are the common methods of resampling?

Ans.

Resampling is a method of drawing repeated samples from original sample randomly.

Methods of resampling are

● K-fold cross-validation

K-fold cross validation is used to measure the error of the model to evaluate its performance. Where we divide the dataset into k fold. First set will be kept for evaluation and rest all will be used for training of the model. This process will be repeated for k times.

● Bootstrapping

Bootstrap is used for wide range of dataset. We take certain number of observations out of total range for a hypothesis for several times and each time by selecting unique values.

90. What is the proportion of confidence intervals that will not contain the population parameter?

Ans.

The probability proportion of confidence interval will not contain population parameter is = α.

The probability proportion of confidence interval will contain population parameter is = 1- α.

In [ ]:

```
1
```

Thanks

Github: https://github.com/saisubhasish (https://github.com/saisubhasish)

In [ ]:

```
1
```

In [ ]:

```
1
```