Probability And Statistics

Assignment set-I

(1) Define Vector space and Vector subspace.

Vector Space :- A vector space is a non-empty set V of objects called vectors on which are defined two operations & called additions and multiplecation by scalers (real Numbers) subject to the axioms (or rules) listed below.-

- The axioms held for all vectors $\bar{u}, \bar{v}, \bar{w}$ in V and for all scalar c and d.

1. The sum of u and V denoted by $\bar{u}+\bar{v}$ is in ie $\bar{u}, \bar{v} \in \bar{V}$ (closure property).
2. $\bar{u}+\bar{v} = \bar{v}+\bar{u}$ (Commutative property)
3. $(\bar{u}+\bar{v})+(\bar{w}) = \bar{u}+(\bar{v}+\bar{w})$ (Associative property)
4. There is a zero vector 0 in V such that $\bar{u}+0 = 0+\bar{u} = \bar{u}$
5. For each $\bar{u}$ in V there is a vector $\bar{u} \in V \ni \bar{u}+(-\bar{u}) = 0$
6. The vector multiplication $c \in F, \bar{u} \in \bar{V} \Rightarrow c\bar{u} \in V$
7. $c(\bar{u}+V) = c\bar{u}+c\bar{v}$. (u, v are vectors). $c \in F, \bar{u} \in \bar{V}$ (c-scalar)
8. $(c+d)u = c\bar{u}+d\bar{u}$ (u-vector). $c, d \in F, \bar{u} \in \bar{V}$
9. $c(d\bar{u}) = (cd)\bar{u}$ $\quad c, d \in F, \bar{u} \in \bar{V}$
10. $1\bar{u} = \bar{u}$

Vector Subspace - A subset of a vector space V is called a subspace of V. if W is itself a vector space under the addition and scalar multiplication defined on V. some properties

(i) The zero vector of V is in H ie $\bar{0} \in H$

(ii) If $\bar{u}.\bar{v} \in H$ then $\bar{u}+\bar{v} \in H$

(iii) $\bar{u} \in H$ and c is any scalar then $c\bar{u} \in H$.

Note:- Let $\bar{u}, \bar{v}, \bar{w} \in H$

$\Rightarrow \bar{u}, \bar{v}, \bar{w} \in V \ (H \subseteq V)$

$(\bar{u}+\bar{v})+\bar{w} = \bar{u}+(\bar{v}+\bar{w})$

And $\bar{u}+\bar{v} = \bar{v}+\bar{u}$

we have $\bar{0} \in H$

$c\bar{u} \in H$

Let $c = -1$

$(-1)\bar{u} = -\bar{u} \in H$

H is subset of V

$\therefore \bar{u} \in H \Rightarrow \bar{u} \in V$

$\therefore c(\bar{u}+\bar{v}) = c\bar{u}+c\bar{v}$

$(c+d)\bar{u} = c\bar{u}+d\bar{u}$

$(cd)\bar{u} = c(d\bar{u})$

we have $c\bar{u}$ H

Put $c = 1$

$\Rightarrow 1\bar{u} = \bar{u} \in H$

(2.) State and prove the addition theorem of probability.

If A and B are two events then the probability of occurance of A or B is given by

$$P(A) \text{ or } P(B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

when A and B are manually exclusive events then

$$P(A \text{ or } B) = P(A) + P(B).$$

Proof:- Since events are sets from set theory, we have

$$n(A \cup B) = n(A) + n(B) - n(A \cap B).$$

Dividing e.g. with $n(S)$; S - Simple space

$$n(A \cup B)/n(S) = n(A)/n(S) + n(B)/n(S) - n(A \cap B)/n(S)$$

Then by the defination of probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

(3.) Explain Sampling Techniques.

The Sampling techniques are used to examine the Selected Sample from the population is known as Sampling Techniques. Sampling Technique is practical and its scope is vast. The whole data is analyzed with better supervision. It require less time and less cost. It gives reliable data the sampling Techniques are as follows:

(i) Random Sampling
(ii) stratified Random Sampling
(iii) Systematic Random Sampling.

(i) Random Sampling – In this sampling, each item in the Population has an equal and likely possibility of getting selected in the sample. "Method of Chance Selection".

(ii) Stratified Random Sampling – In this method, the Population is divided into subgroups to obtain a simple Random Sample from each group and complete the sampling process. the small groups are also ~~allocat~~ called strata.

(iii) Systematic Random Sampling – In this sampling method, the items are chosen from the destination Population by chosing the random selecting point and picking other methods after fixed sample Period. This method is used for when a complete list of population is available.

Ex. - Consider there 1000 persons - from whom we have to chose 10 persons. for the study of any given sample then we number them from 1 to 1000 and make them as 10 intervals and 1st we choose a person from the 1st interval 2nd person is choosen from the 2nd interval and so on,

As we have numbered the persons all the persons are systematically choosen, This kind of method is wed when the population is large.

(iv) Cluster Sample - A cluster sample randomly selected group this- method is wedful when the population quietly dispersed and consists of many natural group such as factories, villages, etc.

(4) write about the testing hypothesis.

Testing of Hypothesis -

A Hypothesis is a statement about the population parameter. Hypothesis-testing is a procedure that helps us to ascertain the likelihood of hypothesized parameter being correct by making use of the sample statistics. The two hypotheses in a statistical test are Normally referred to as,

(i) Null Hypothesis
(ii) Alternative Hypothesis.

(i) Null Hypothesis - Null Hypothesis which is tested to be actually tested for acceptance or rejection is termed as Null Hypothesis. According to R A Fisher, "Null Hypothesis is the hypothesis which is tested for rejection under the assumption that it is true".

The Null Hypothesis is a very useful tool to test the significant of difference. In the process of statistical test, the Hypothesis is rejected or accepted based on the sample drawn from the mean of the population. This hypothesis reveals that the mean of the sample and the mean of the population under study do not show any difference.

A statistical hypothesis is a Null hypothesis if it is accepted. we should take consideration the following while setting up the Null hypothesis:

(a) To test the significance of the difference between the values of the sample and the population, or between two sample values; we set up the Null hypothesis that the difference is not significant. This is beacarse the difference is due to sample fluctuations.

$H_0: \mu = \bar{x}$

where μ = population mean

$\bar{x}$ = Sample mean

(b) To test any statement about the population, we set up the null hypothesis, that is the true.

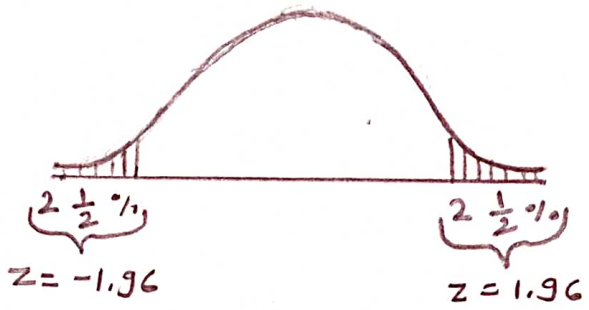
(ii) Alternative Hypothesis:- "Any Hypothesis which is complimentary to the null hypothesis is called an alternative hypothesis.". Rejection of H_0 leads to the acceptance of Alternative hypothesis which is denoted by H_1.

For example, if we want to select the null hypothesis for difference between population mean and Sample mean then we these hypothesis Can be written as follows:

H_0: $\mu = \bar{x}$ (Null Hypothesis)

H_1: $\mu \neq \bar{x}$ (two-tailed alternative Hypothesis)

H_1: $\mu > \bar{x}$ or $H_1: \mu < \bar{x}$ (right tailed and left-tailed tests)



Two-tailed Diagram.

$H_0: \mu = x$

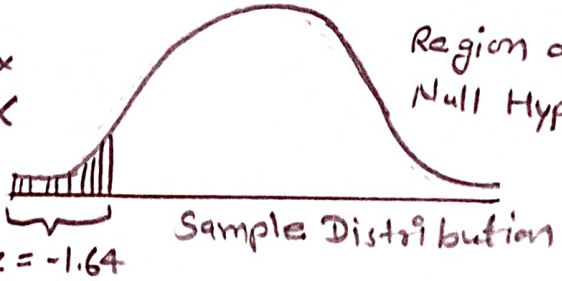
$H_1: \mu = X$

$C = X + z_\sigma \times S.E.$ = Critical value for upper time

$C = X - z_\sigma \times S.E.$ = Critical Value for left tail

$H_0: \mu = x$

$H_1: \mu = X$



The validity of H_0 and H_1 is then ascertained at a certain level of significant. The significance level stands for the confidence with which the experimentor rejects or refains the Null Hypo. Significance of level is customarily expressed as a percentage is in 1 in 20, when it is true.

Having setup the Null and the Alternative Hypothesis and the significance level, the next step is to construct a test criterion. This involves selecting the right probability distribution for the particular test.

5. Difference ~~between~~ Correlation of regression analysis

In many business situations we have to deal with two or more variables. specially, in the analysis and interpretation of data we have to take into account the relationship between demand and price, output and rainfall, volume of sales and expenditure on advertisement, etc. For study of such relationships the two important statistical methods used are Correlation and regression.

These methods are also helpful in forecasting figures for the future. For example, a company planing next year's production may be interested in the forcast of sales for the year. if the marketing manager knows the sales having a relationship with advertising expediture and few other variables such as public expenditure, national income, etc.. he will be able to Predict the value of ~~all~~ sale with these variable Provided the value of all these variable is known. Similarly, a cost accountant can estimate the cost and the ~~price~~ of a product if there are established relationship between the cost and the price of inputs such as labor, meterial, sales ~~pred~~ promotion expediture, etc. In statistics we find these relationships by the methods of correlation and regression.

CORRELATION	REGRESSION
(1) It determines the interconnection or a co-relationship between the variables.	It explains how an independent variable is numerically associated with the dependent variable.
(2) In Correlation, both the independent and dependent values have no difference.	In regression, both the dependent and independent variables are different.
(3) The main objective of Correlation is to find a quantitative or numerical value expressing the asso-ciation between the values.	The main purpose is to calculate the values of a random variable based on the values of a fixed variable.
(4.) It stipulates the degree to which both variables can move together.	It specifies the effect of the change in the unit in the known variable (P) on the evaluated variable (Q).
(5.) It helps to constitute the connection between the two variables.	It helps in estimating a variable's value based on another given value.