

ONLINE FINAL EXAMINATION

Name :- Sadaf Saleem

Class :- BSCS-2nd (M)

Reg # 2019-GUVF-010403

Roll # 2929

Submitted to :- Sir Muhammad Aftab

QUESTION NO: 01

Role of statistics in CS field

⇒ Statistics -

Statistics is used for speech recognition, data mining, vision and image analysis.

A statistical background is essential for understanding algorithms and statistical properties that forms the backbone of computer science.

⇒ Role of Statistics

There are a number of ways the roles of statisticians and computer scientists merge; consider the development of models and data mining. Ultimately, both of them comes together in attempts to solve problems.

- ① statistics involves machine learning in CS field.
- ② Data Mining i.e; Management and Analysis of data can be done using statistics.
- ③ Machine learning is a subset of AI in field of CS that often uses statistical techniques to give computers the ability to learn with data.

→ Data Mining is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics and database systems.

④ Statistics back programs such as Google translate, which uses data to perform online translation.

⑤ Statistics are also used in network traffic modeling.

QUESTION NO :- 02

Difference b/w correlation & Regression.

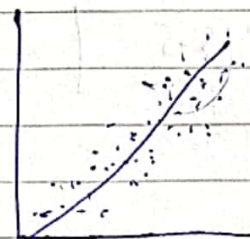
Meaning	
correlation is a statistical measure that determines the association or co-relationship between two variables. uses' pearson	Regression describes how to numerically relate an independent variable to the dependent variable.
Usage	
To represent linear relationship between two variables.	To fit the best line and to estimate one variable based on another.
Dependent / Independent Variable	
No difference	Both variables are different.

Objectives

To find a numerical value expressing the relationship between variables

To estimate values of random variables on the basis of values of a fixed variable.

~~Example~~ Graph.



Example.

x, y, xy, x^2 are same as previous.

X	Y
15	50
20	25
6	30
12	82
4	41
$\Sigma x = 57$	$\Sigma y = 228$

X	Y	XY	X ²	\bar{x}	\bar{y}
-	-	-	-	$57/5 = 11.4$	$228/5 = 45.6$
-	-	-	-		
-	-	-	-		
-	-	-	-		
-	-	-	-		

$$r = \frac{N \Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{(N \Sigma X^2 - (\Sigma X)^2)(N \Sigma Y^2 - (\Sigma Y)^2)}}$$

X	Y	XY	X ²	Y ²
15	50	750	225	2500
20	25	500	400	625
6	30	180	36	900
12	82	984	144	6724
4	41	164	16	1681
57	228	2578	821	12430

Eq. of correlation

$$= \frac{N \sum xy - (\sum x)(\sum y)}{\sqrt{(N \sum x^2 - (\sum x)^2)(N \sum y^2 - (\sum y)^2)}}$$

By putting values;

$$\boxed{= -45.5}$$

Eq. of regression

$$\hat{y} = bx + a$$

$$\therefore b = \left[\frac{\sum xy}{\sum x^2} \right]$$

$$\therefore a = \bar{y} - b\bar{x}$$

$$b = \frac{\sum xy}{\sum x^2}, \quad a = \bar{y} - b\bar{x}$$

$$b = 3.14$$

$$a = 9.8$$

$$\boxed{\hat{y} = 3.14x + 9.8}$$