Page No. _____
Date

Ques 3. What are different - - - - - - ..

Ans - The different sampling technique are:

(1) Stratified sampling - The sampling divides the data set into subgroups and after that selection is done.
ex - division of covid people for vaccination by age and recording the side effects.

(2) Cluster sampling - Groups are formed where a group contains similar kind of features
ex - Recording vaccination side effects on pregnant women, girls, boys etc.

(3) Systematic Sampling - It is similar to random sampling but random sampling is done on regular intervals
ex - Selecting from every 100 records

(4) Random Sampling - In this type of sampling every record has equal chance to be selected.
ex - Randomly selecting data for covid-19 vaccine.

Page No.
Date

For example of stratified sampling the people are divided on age groups of 0-18, 18-45, 45+ to record the sideffect of vaccination.

Page No.
Date

Ques- For the simplified ----.

Ans- $y = f(\theta, X, \varepsilon)$

Firstly we will try to create a relation between length weight and age. We will try forming this relation from the use of the already known length weight and age data.

After that we will take input from the user and give our ans based on the relation we devised earlier.

Let the relation be

Len = a (Weight) + b(Age) + Error(e)

a & b are ~~parameters~~ coefficient of weight and age respectively.

Ques LONG ANSWER

Ans - Joint probability distribution allows us to compute probabilities of events involving both discrete and continuous variables and understand the relation between them.

Page No.
Date

① Discrete Case

Suppose X and Y are two discrete random variables and that X takes values $(x_1, x_2 \ldots x_n)$ and Y takes $(y_1, y_2 \ldots y_m)$. The joint probability mass function of X and Y is the function of $p(x_i, y_i)$ giving the probability of joint outcome, $X = x_i$ & $Y = y_i$

X \ Y	y_1	y_2	y_3	...	y_m
x_1	(x_1, y_1)	(x_1, y_2)			
x_2	(x_2, y_1)				
x_3					
⋮					
x_n					(x_n, y_m)

It must satisfy

① $0 \leq p(x_i, y_i) \leq 1$

② $\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_i) = 1$

② Continuous Case

In continuous we replace discrete sets by values of continuous intervals

If X take $[a, b]$ and Y $[c, a]$ then pair (X, Y) takes $[a, b] \times [c \times d]$

It must satisfy

(i) $0 \leq f(x,y)$

(ii) $\int_c^d \int_a^b f(x,y)\,dx\,dy = 1$

Bayes rule

According to bayes, if we have an Event A whom we want to calculate probablity of, and B is the new evidence related to A in some way

so the probablity of event A occuring if B has already happen is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Independence is the condition when occuring of one event does not effect the other.

For such event A & B

$$P(A \cap B) = P(A)\,P(B)$$

For jointly distributed data

$$F(X,Y) = F_X(x)\,F_Y(y)$$

For predicting error in regression we use mean squared Error to aggregate the prediction error

Page No. _____
Date

In Regression, the prediction error can be predicted with.

(1) Mean average error – It is not a good measure in case when sum of deviation from mean is zero.

(2) Mean absolute error – It is absolute value of deviation from mean.

$$MAE = \frac{\sum |x_i - \bar{x}|}{N}$$

(3) Mean squared error = $MSE = \frac{\sum |x_i - \bar{x}|^2}{N}$ but the unit is unit^2

(4) RMSE (root mean square error)

$$RMSE = \sqrt{\frac{\sum |x_i - \bar{x}|^2}{N}}$$. It has same unit as of data.

For classification –

(1) Confusion matrix –

Predicted \ Actual	T	F
T	TP	FP
F	FN	TN

Confusion metrics gives the accuracy, f-factor of classification model.

$$Accuracy = \frac{TP + TN}{total}$$