**Course Title:**

**Statistical Inference**

**Submitted**

**to**

**Ms. Wajiha Amin**

**Submitted by**

**Farhad khan**

**I’d**

**13673**

**IQRA NATIONAL University Peshawar**

**Q1:The average rainfall in an area recorded is 9.22cm for a month. Given the distribution to be normally distributed with a standard deviation of 2.83cm,**

* Find the probability that the rainfall in the next month will be less than 1.84cm.
* Rainfall will be between 7 cm and 13.8cm.
* Rainfall is more than 11.05 cm

**Ans;**

 **DATA**

H= 9.22cm σ = 2.83cm

P ( x< 1.84)= ?

Z = $\frac{1.84-9.22}{2.83}$

Z = -2.607

P (x< 1.84) = P ( Z < -2.607

 = 0.0038)

 2) P (7<x<13.8) = ?

Z1 = $\frac{x-u }{o}$

= $\frac{7-9022}{2.83}$

= - 0.784

P (x>7) = ?

P (x>7) = P (Z> -0.784)

 = 0.2177

P (13.8>x) = ?

Z2 = $\frac{x-u }{o}$ = $\frac{13.8-9.22}{2.83}=1.618$

P ( x<13.8) = P (Z<1.618)

= 0.9463

P (7<x<13.8)

= P (-0.784) < Z < 1.618)

= 0.9463 – 0.2177

= 0.7286

 (3)

P (x > 11.05)

Z=$\frac{x-u }{o}$

Z= $\frac{11.05-9.22}{2.83}$

Z= 0.646

P (x > 11.05)

P ( 2 > 0.646)

= 0.7389

**Q2: (a) Discuss any 3 characteristics of normal distribution and discuss its uses in the business world.**

**Answer ;**

1. **Mean , median , mode**
2. **Symmetrical**
3. **Asymptotic**

****

**Mean** **(Average)**

Mean gives very true picture of central value in case of symmetrically or normally distributed data. As majority of distributions are symmetrical, mean becomes the most widely used measure of central tendency in any industrial or non-industrial set up. Example: Mean height, weight, pressure, score etc.

Median

Consider the income distribution in a country in which income of 20% rich people is roughly equal to income of 80% remaining people.

In this case, mean income will not give a true picture as it will get jacked up due to high income of 20% people.

Median will give typical income that will be more realistic.

**Mode**

Mode is most easily understandable and quickly visible. Many decisions in day-to-day life is taken on modal concept. Like, going to watch a movie where most people are going. Or choosing colour of car as the one most people are buying.

Symmetrical distribution occurs when the values of variables occur at regular frequencies and the [mean](https://www.investopedia.com/terms/m/mean.asp), median and mode occur at the same point. In graph form, symmetrical distribution often appears as a [bell curve](https://www.investopedia.com/terms/b/bell-curve.asp). If a line were drawn dissecting the middle of the graph, it would show two sides that mirror each other. Symmetrical distribution is a core concept in technical trading as the price action of an asset is assumed to fit a symmetrical distribution curve over time.

**Example**

Symmetrical distribution is most often used to put price action into context. The further the price action wanders from the value area one standard deviation on each side of the mean, the greater the probability that the underlying asset is being under or overvalued by the market. This observation will suggest potential trades to place based on how far the price action has wandered from the mean for the time period being used. On larger time scales, however, there is a much greater risk of missing the actual [entry](https://www.investopedia.com/terms/e/entry-point.asp) and [exit points](https://www.investopedia.com/terms/e/exit-point.asp).

**Asymmetrical distribution**

Asymmetrical distributionis a situation in which the values of variables occur at irregular frequencies and the mean, median, and mode occur at different points. An asymmetric distribution exhibits [skewness](https://www.investopedia.com/terms/s/skewness.asp). In contrast, a Gaussian or [normal distribution](https://www.investopedia.com/terms/n/normaldistribution.asp), when depicted on a graph, is shaped like a bell curve and the two sides of the graph are symmetrical.

 **(b) Suppose you are going to be conducting a study on students, asking for their opinion on an issue of interest to you (could be related to the university, or a wider societal issue).Describe how you would carry out the sampling of students using the following methods:**

(i) simple random sampling

(ii) stratified sampling

1. cluster sampling Think about what attributes of the student population make sense to stratify vs. cluster.

**Answer :** Simple random sampling: Simple random sampling is a subset of a statistical population in which each number of the subset has an equal probability of being chosen. A simple random sample is meant to be an unbiased representation of a group.

 It would be the names of 25 students being chosen out of a hat from a batch of 250 students. In this case the population is all 250 employees and the sample is random because each employee has an equal chance of being chosen. Random sampling is used in science to conduct randomized control test or for blinded experiments.

Stratified sampling: Stratified sampling is a method of sampling that involves the division of a population in to smaller sub-groups known as strata.

I will now divide all of the students into two groups then conduct a questioner test then will select the ones with the best results .

This is conducted to save time and money.

The result of a subset of population which represents the entire population.

 Cluster sampling: It is known as the probability sampling technique . Researchers divide the population in to multiple groups called clusters . Then random groups are selected with a systematic random sanpling technique for data analysis.

 I will divide the students in different clusters then select the best group for the test.

I think the cluster formation will be best for the students division.

**Q3: (a) Determine the type of sampling used (simple random, stratified, systematic, cluster, or convenience).**

**1.** A group of test subjects is divided into twelve groups; then four of the groups are chosen at random.

**2.** A market researcher polls every tenth person who walks into a store.

**3.** The first 50 people who walk into a sporting event are polled on their television preferences.

**4.** A computer generates 100 random numbers, and 100 people whose names correspond with the numbers on the list are chosen.

Answer :

1. cluster
2. Systematic
3. Convenience
4. Simple random

**b)** Differentiate between

Descriptive statistics and inferential statistics

Answer :

**Descriptive statistics:**

* Summarize data – visually, numerically
* Obtain insight about population
* Select suitable modeling approach

**Inferential Statistics:**

* Understand patterns in the data
* Drawing inferences about the larger population

**Variance and standard deviation**

Answer;

Standard deviation is in the original measurement units, so it has practical relevance. It describes the typical amount that values stray from the mean.

The variance is in squared units, so it has little use in actually describing data. It's usefulness is really only as critical component of procedures that analyze how multiple variables are related, such as ANOVA, linear regression, etc. Essentially, working with sums of squared deviations, which is the numerator of the variance formula, has better mathematical properties.

 **Differentiate between Cluster and strata**

**Answer**:

**Cluster and strata:** The main difference between cluster sampling and stratified sampling is that in cluster sampling the cluster is treated as the sampling unit so sampling is done on a population of clusters (at least in the first stage). In stratified sampling, the sampling is done on elements within each stratum