***Final Term***

***Name: M Fahad Kamal***

***Id: 13068***

***Subject: Statistical Inference***

***Submitted to: Ms Wajiha Amin***

***Q:1***

**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:**

**Solution:**

**Part a:**

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)

1. 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

**(2) Part b:**

P (7<x<13.8)

= P (-0.784) < Z < 1.618)

= 0.9463 – 0.2177

**=** 0.7286

**(3)Part c:**

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

**Q:2**

1. **Discuss any 3 characteristics of normal distribution and discuss its uses in the business world**

 **Ans:**

The characteristics of normal distribution are as follows:

The distribution is symmetric and its skewness measure is zero and apart from that

the entire family of normal distribution is defined by its mean and its standard

deviation.

The highest point on the normal curve is at the mean, which is also the

median and the mode and the mean either be positive negative or even zero.

The standard deviation decides the width of curve and larger values result in

wider curves.

A business needs proper information in time so that it can make timely

and accurate decisions manage all the risks and confirms the effective application

of decisions. Now for this purpose our business heavily depends on statistics as it

can provide qualitative and quantitative information for a specific purpose and then

after that it analyzes and interprets the collected information and all of it aids in

decision making.

Normal distribution is of utmost importance it helps to decide certain

characteristics of data and also provides as a base for using other statistical tools

for decision making. It helps in quality control, cost managemen and business

operations just to determine the most sensitive part of the variable.

**Part B**

**(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**

**Ans:**

**: (i) simple random sampling**

In this technique, each member of the population has an equal chance of being selected as subject. The entire process of sampling is done in a single step with each subject selected independently of the other members of the population. ... Another way would be to let a computer do a random selection from your population.

1. **Stratified sampling**

A sample may be selected from a population through a number of ways, one of which is the stratified random sampling method. A stratified random sampling involves dividing the entire population into homogeneous groups called strata (plural for stratum). Random samples are then selected from each stratum

**(iiI) cluster sampling Think about what attributes of the student population make sense to stratify vs. cluster.**

1. In stratified sampling, a sample is drawn from each strata (using a random sampling method like simple random sampling or systematic sampling). ...

2. In cluster sampling, the sampling unit is the whole cluster; Instead of sampling individuals from within each group, a researcher will study whole clusters

**Q:3**

**: (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.**

**Ans:**

1) cluster

2) Systematic

3) Simple random

4) Convenience

**Part: B**

**Differentiate between**

* **Descriptive statistics and inferential statistics**
* **Variance and standard deviation**
* **Cluster and strata**

**Ans:**

* **Descriptive statistics and inferential statistics**

 Descriptive statistics describes sets of data and inferential statistics draws

conclusions about the sets of data based on sampling. Descriptive and inferential

they both give different insights into the nature of the data gathered one alone just

cannot provide you with the whole picture

**Descriptive Statistics**

Descriptive statistics describe a sample. That’s pretty straightforward. You simply take a group that you’re interested in, record data about the group members, and then use summary statistics and graphs to present the group properties. With descriptive statistics, there is no uncertainty because you are describing only the people or items that you actually measure. You’re not trying to infer properties about a larger population

**Inferential Statistics**

Inferential statistics takes data from a sample and makes inferences about the larger population from which the sample was drawn. Because the goal of inferential statistics is to draw conclusions from a sample and generalize them to a population, we need to have confidence that our sample accurately reflects the population. This requirement affects our process.

* **Variance and standard deviation**

 The variance measures the degree to which

each point differs from the mean while on the other hand the standard deviation

looks at how spread out a group of numbers is from the mean just by looking at the

square root of the variance.

**Variance**

The variance is the average of the squared differences from the mean. To figure out the variance, first calculate the difference between each point and the mean; then, square and average the results

The variance measures the average degree to which each point differs from the mean—the average of all data points.

**Standard Deviation**

Standard deviation is a statistic that looks at how far from the mean a group of numbers is, by using the square root of the variance. The calculation of variance uses squares because it weighs outliers more heavily than data closer to the mean. This calculation also prevents differences above the mean from canceling out those below, which would result in a variance of zero.

Standard deviation is calculated as the square root of variance by figuring out the variation between each data point relative to the mean.

Standard deviation looks at how spread out a group of numbers is from the mean, by looking at the square root of the variance.

* **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

**Cluster:**

Cluster sampling is a sampling plan used when mutually homogeneous yet internally heterogeneous groupings are evident in a statistical population. ... In this sampling plan, the total population is divided into these groups (known as clusters) and a simple random sample of the groups is selected.

**Strata:**

In statistics, a stratum (plural strata) refers to a subset (part) of the population (entire collection of items under consideration) which is being sampled. Stratification thus consists of dividing the population into strata within each of which an independent sample can be chosen