STATISTICAL METHODS

STATISTICS:

The term statistics refers to a set of mathematical procedures for organizing, summarizing, and interpreting information.

Researcher uses different statistical methods or procedures. Although the researchers have developed a variety of different statistical procedures to organize and interpret data, these different procedures can be classified into two general categories. The first category, *descriptive statistics,* the second general category of statistical technique is called *inferential statistics.*

TYPES OF STATISTICS

- ➢ Descriptive Statistics
- ➢ Inferential Statistics
- ➢

DESCRIPTIVE STATISTICS :

Descriptive statistics is a discipline of quantitatively describing the main features of a collection of data, or the quantitative description of itself. Descriptive statistics are used to summarize, organize and simplify data. Descriptive statistics are techniques that take raw scores and organize or summarize them in a form that is more manageable. Often the scores are organized ina table or a graph so that it is possible to see the entire set of scores. Descriptive statistics are not developed on the basis of probability theory. Different measures are used to describe descriptive statistics.

SCALES OF MEASUREMENTS:

Several types of scales are associated for measurement. The distinction among the scales is important because they underscore the limitations of certain data collected on the same scales but not on others. Four types of scales are used for measurement which are :

- o The Nominal Scale
- o The Ordinal Scale
- o The Interval Scale
- o The Ratio Scale

Nominal Scale

Nominal scale consists of a set of categories that have different names. Measurement on a nominal scale label and categorize observations, but do not make any quantitative distinctions between observations.

❖ If two individuals are measured on a nominal scale, it is possible to determine measurements are different one cannot determine how big is the difference. So one cannot say that one measurement is more or less than the other.

The Ordinal Scale

The ordinal scale consists of a set of categories that are organized in an ordered sequence. Measurement on an ordinal scale rank observation in terms of size or magnitude.

❖ Often ordinal scale consists of a series of ranks (first, second and third .. so on). Often categories are identified by verbal labels like small, medium and large.

The Interval Scale

An interval scale consists of a set of categories that are all intervals of exactly the same size. With an interval scale, equal differences between numbers on the scale reflect equal differences in magnitude. However, ratios of magnitude are not meaningful. Another characteristics of an interval scale is that it has an arbitrary zero point. Specifically, zero does not indicate the total absence of the variable being measured.

❖ Like in temperature, 0 degrees temperature does not mean that there is no temperature.

The Ratio Scale

The ratio scale is an interval scale with additional features of an absolute zero point. With a ratio scale, ratios of numbers do reflect ratios of magnitude. In ratio scale, value of zero, indicates none (a complete absence) of the variable being measured.

❖ For example one can say that one measurement is three times larger than other, or one can say that one score is only half as large as another.

Some major measures that are commonly used to describe descriptive statistics are :-

- Frequency distribution
- Measures of Central Tendency
- Measures of Variability or Dispersion

Frequency distribution includes frequency distribution tables, graph, shapes and stem and leaf display. Measures of Central Tendencyinclude Mean, Median and Mode. While Measures of Variability include Standard deviation (Variance), the minimum and maximum values of variables, kurtosis and skewedness.

FREQUENCY DISTRIBUTION :

Frequency distribution is an organized tabulation of the number of individuals located in each category on the scale of measurement. Frequency distribution takes a disorganized set of scores and places them in order from highest to lowest. Grouping together all individuals who have the same scores. A frequency distribution can be structured either as a table or a graph but in either case the distribution presents the same two elements:

1. The set of categories that make up the original measurement scale.
2. A record of the frequency, or number of individuals in each category.

FREQUENCY DISTRIBUTION TABLES:

The simplest frequency distribution table presents the measurement scale by listing the different measurement categories(*X* values) in a column from highest to lowest.

Example:

8, 9, 8, 7, 10, 9, 6, 4, 9, 8, 7, 8, 10, 9, 8, 6, 9, 7, 8, 8

X	F
10	2
9	5
8	7
7	3
6	2
5	0
4	1

We can also find out the proportion *P (P = f/N),* percentage % (% = *p*100).*

FREQUENCY DISTRIBUTION GRAPH:

A frequency distribution graph is basically a picture of the information available in a frequency distribution table. Frequency distribution graph includes:

➢ Histogram (graph for interval or ration data)
   Polygon
➢ Bar graph (graph for nominal or ordinal data)

HISTOGRAM:

In order, to construct histogram, first list the numerical scores (the categories of measurement) along the *X*-axis. Then draw a bar above each *X* value so that

a. The height of the bar corresponds to the frequency for that category.
b. The width of the bar extends to the real limits of the category.
❖ An example of a frequency distribution histogram.

| X	f
6	1
5	2
4	2
3	4
2	2
1	1
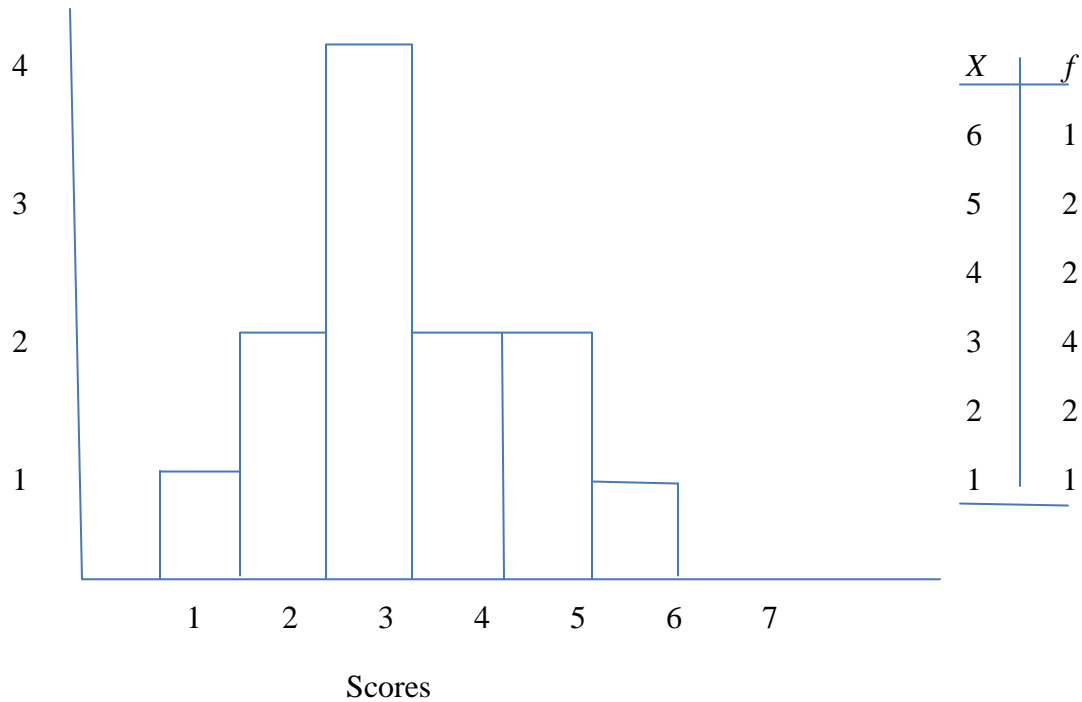
Scores

MODIFIED HISTOGRAM:

In drawing histogram, we draw a bar above each score, the modifications consists of drawing a stack of blocks. Each block represents one individual, so the number of blocks above each score corresponds to the frequency for that score.

❖ An example of modified histogram in which each individual is represented by a block placed directly above the individual's score.

X	f
1	1
2	3
3	3
4	4
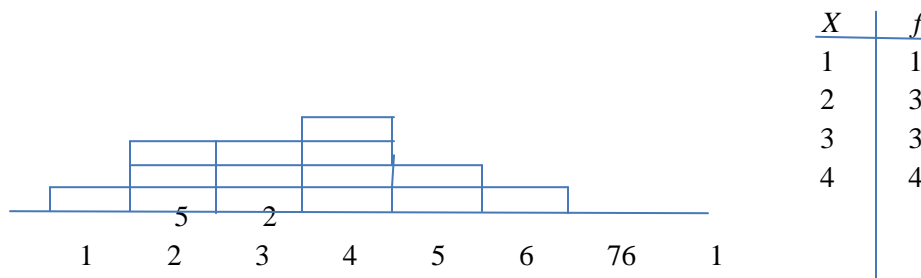
POLYGON:

❖ An example of frequency distribution polygon.

X	f
6	1
5	2
4	2
3	4
2	2
1	1

Scores
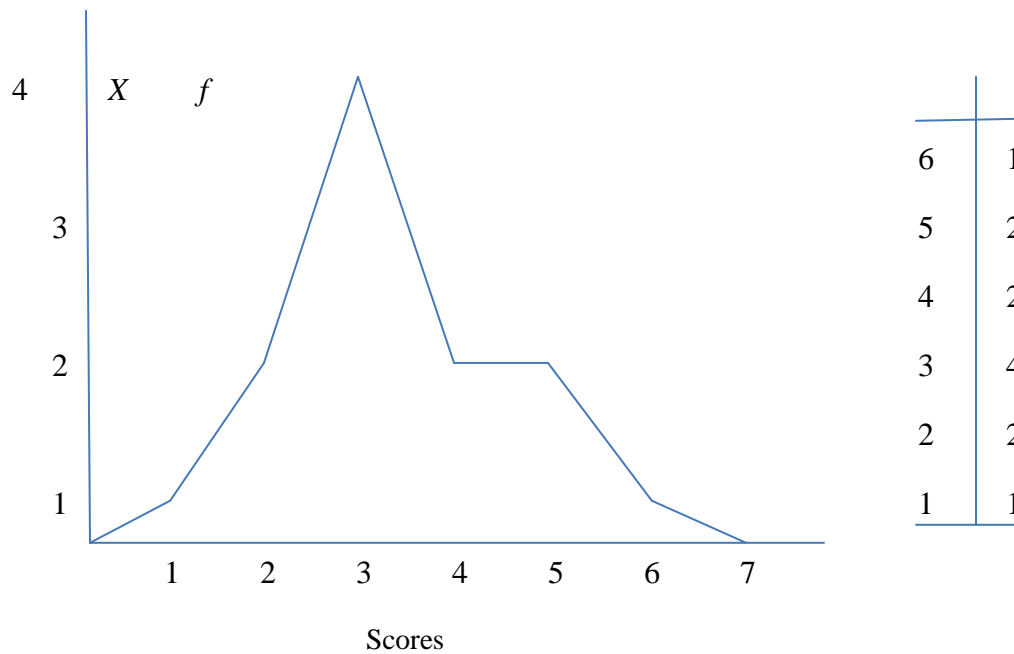
BAR GRAPHS :

A bar graph is essentially the same as a histogram, except that spaces are left between adjacent bars. For a nominal scale, the space between bars emphasizes that the scale consists of separate , distinct categories. For ordinal scales, separate bars are used because one cannot assume that the categories are all the same size.

❖ A bar graph showing the distribution of personality types in a sample of college students. Because personality type is a discrete variable measured on a nominal scale, the graph is drawn with space between the bars.
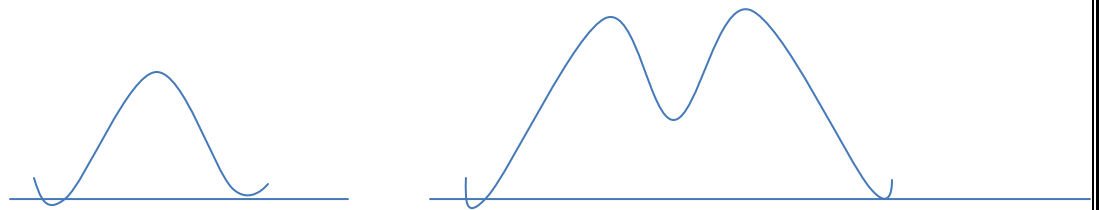

Personality type
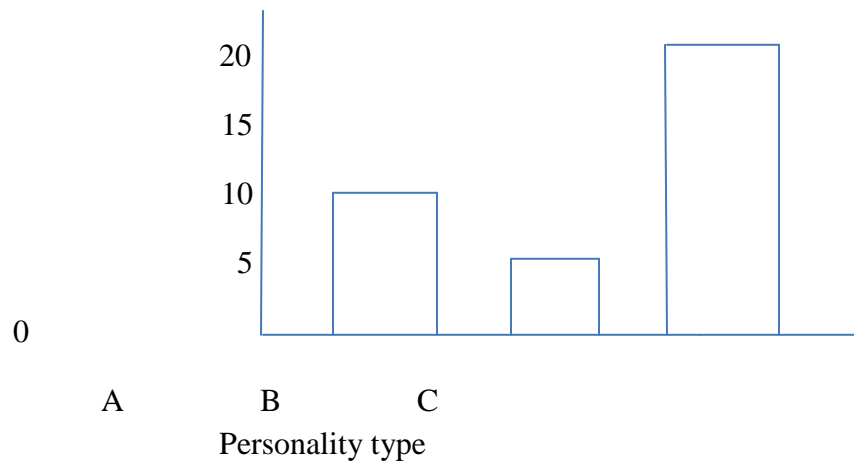
SHAPES OF FREQUENCY DISTRIBUTION:

There are two types of shapes of frequency distribution.

- Symmetrical Distribution
- Skewed Distribution

In a *symmetrical distribution,* it is possible to draw a vertical line through the middle so that one side of the distribution is a mirror image of the other.

Symmetrical distributions

(+ve)   (-ve)

Skewed distributions

Stem and Leaf Display :

In 1977, tukey presented a technique for organizing data that provides a simple alternative to a frequency distribution table. This technique called *stem and leaf display,* requires that each score be seprated into two parts: the first digit is called stem, and the last digit is called leaf.

❖ Anset of N=5 scores presented as raw data and organized in a stem and leaf display.

DATA	STEM AND LEAF DISPLAY	
21	2	1
15	1	5
32	3	2
27	2	7
24	2	4

MEASURES OF CENTRAL TENDENCY :

Central tendency is a statistical measure to determine a single score that defines the center of distribution. The goal of central tendency is to find the single score that is most typical or most representative of the entire group. Statisticians have developed three different methods for measures of central tendency:mean median and mode.

MEAN:

The mean, is commonly known as athematic average, is computed by adding all the scores in the distribution and dividing by the number of scores. The mean for a population is denoted as μ, pronounced as ' myoo' and the mean for a sample is determined by *M* or *X* (read as " x-bar").

The formula for the population mean is : $\mu = \sum X/N$

The formula for the sample mean is : $M = \sum X/n$

Whereas, n= no. of individuals for sample , and N= no.of individuals for population.

❖ For a population of N = 4 scores,

The mean is

$\mu = \sum X/N = 20/4 = 5.$

But if we have two different samples, then we will find out the mean by weighted mean / overall mean formula.

A weighted mean can be find out by formula:

Overall mean = $M = \sum X$ (overall sum for the combined group)

N (total number in the combined group)

$$= \frac{\sum X + \sum X}{N + N}$$

Whereas $\sum X$ is the sum of scores for the first sample and similarly $\sum X$ is sum of scores for the second sample.

❖ The first sample has $n = 12$ and $M = 6$. Therefore, the total for this sample must be $\sum X = 72$. And the second sample has $n = 8$ and $M = 7$, so the total must be $\sum X = 56$. Using these values, we obtain the overall mean of

Overall mean $= \sum X + \sum X / n + n = 72 + 56 / 12 + 8 = 6.4$

The overall mean is not halfway between the original two samples means. Because two samples are not the same size, one will make a larger contribution to the total group and therefore will carry more weight in determining the overall mean.

Every score in the distribution contributes to the value of the mean. Specifically, every score must be added into the total in order to compute the mean. The total score can be change by introducing a new score or removing a new score in the data. By adding or subtracting a constant from each score, by multiplying or dividing each score by a constant.

MEDIAN:

The second measure of central tendency is median. The median is the score that divides a distribution exactly in half. Exactly 50% of the individuals in a distribution have scores at or below the median. The median is equivalent to the 50th percentile. There are no symbols or notion for the median.

❖ When N is an odd number

3, 5, 8, 10, 11

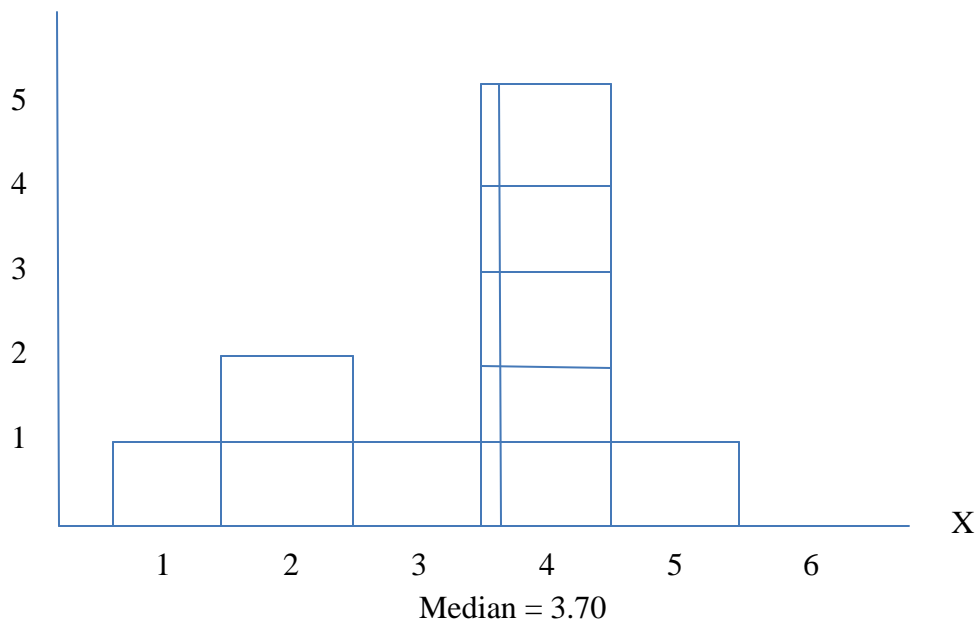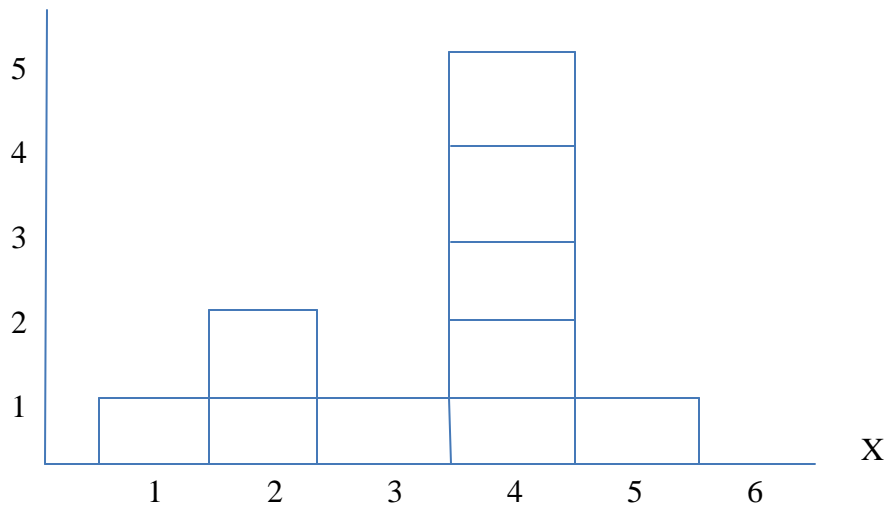Then its median will be $= Median = 8$.

❖ When the N is an even number

3, 3, 4, 5, 7, 8

Median $= 4 + 5 / 2 = 9 / 2 = 4.5$

❖ When there are several values in the middle of the distribution

1, 2, 2, 3, 4, 4, 4, 4, 4, 5

To locate the median with greater precision it will be necessary to split the scores into fractions so that part of each score goes above the median and part goes below the median.

Median = 3.70

❖ A distribution with several scores clustered at the median. The median for this distribution is positioned so that each of the five boxes above $X=4$ is divided into two sections, with 1/5 of each box below the median (to the left) and 4/5 of each box above the median (to the right).

MODE:

The final measure of central tendency is called mode. In a frequency distribution the mode is the score or category that has the greatest frequency.

As with median, there are no symbols or notion used to identify the mode or to differentiate between sample or population mode. In addition, the definition of the mode is the same for a population and for a sample distribution.

The mode is useful measure of the central tendency because it can be used to determine the typical or average value for any scale of measurement, including a nominal scale.

MODE FOR UNGROUPED DATA

We find out the mode for ungrouped data by following way:

❖ A set of scores = X = 5, 5, 4, 4, 4, 4, 4, 3, 3, 3, 2, 2, 1, 1

Now in this data **4** is the mode because it is has the greatest frequency .

MODE FOR GROUPED DATA

We find out the mode for grouped data by following formula :

$$L + \left[\frac{f - f}{2f - f - f} \right] \times h$$

Whereas , L = lower class limit, h = size of class interval, f = freq.of modal class, f = preceding frequency of modal class, f = following frequency of modal class.

❖ An example for grouped data of mode is given below :

CI	f
5 - 10	3
10 – 15	5
15 – 20	7
20 – 25	2
25 – 30	4

By putting values in formula

$$\text{Mode} = 15 + \left[\frac{7 - 5}{2 \times 7 - 5 - 2} \right] \times 5$$

= **76.42**

WHEN TO USE MEAN
The goal of central tendency is to find the best representative value of distribution, so mean has good advantage of being closely related to variance and standard deviation, the most common measure of variability is mean. This relationship makes mean a valuable measure for purposes of inferential statistics.

WHEN TO USE MEDIAN
When a distribution has a few extreme scores, scores that are very different in value from most of the others, then the mean may not be good representative of distribution. The one or to extreme values can have a large influences and cause the mean to be displaced. So in this situation we use median. And when the values are undetermined, then we also use median for the distribution. If the distribution is said to be open – ended when there is no upper limit or lower limit for one of the categories. Then we also use median.

WHEN TO USE MODE
Mode is use in three situations
Nominal scales
Discrete variables
Describing shapes.

MEASURES OF VARIABILITY OR DISPERSION :

VARIABILITY

Variability provides a quantitative measures of the degree to which scores in a distribution are spread out or clustered together.

Variability describes the distribution. Specifically it tells where the scores are clustered close together or are spread out over a large distance. Usually variability defines in terms of distance. Variability measures how well an individual's score (or group of scores) represents the entire distribution. It also tells that how much error to expect if one is using to represent a population.

THE RANGE

The range is the distance between the largest score (*X max*)and the smallest score (*Xmin)* in the distribution. Therefore, the range is difference between upper real limit and lower real limit :

Range = URL *Xmax* – LRL *Xmin*

- ❖ For example, consider the following data :
  3, 7, 12, 8, 5, 10
  Range = URL *Xmax* – LRL *Xmin*
       = 12.5-2.5 = 10

STANDARD DEVIATION AND VARIANCE FOR A POPULATION:

Standard deviation is the most commonly used and the most important measures of variability. Standard deviation uses mean of the distribution as reference point and measures

Deviation is a distance from the mean:

Deviation score = $X - \mu$

For a distribution of scores with $\mu = 50$, if the score X = 53. Then deviation score will be

$X - \mu = 53 - 50 = 3$

POPULATION VARIANCE

Population variance = mean squared deviation. Variance is the mean of the squared deviation scores.

Process of squaring deviation scores does more than simply get rid of plus and minus sign. It results in a measure of variability based on squared distances. Although variance is valuable for some of the inferential statistical methods. The mean squared distance is not the best descriptive measure for variables.

Standard deviation = $\sqrt{}$ variance

Technically the standard deviation is the square root of the mean squared deviation

Whereas variance = square mean deviation = sum of squared deviations/ number of scores

SS, or **sum of squares**, is the sum of the squared deviation scores.

Definitional formula: $SS = \sum (X - \mu)2$

- ❖ the following set of N = 4 scores. These scores have a sum of $\sum X = 8$, so the mean is μ 2. For each scores , we have computed the deviation and the squared deviation. The squared deviation add up to SS = 22. These sets are illustrated as

xx-μ	(x-μ)2		
			$\sum X$ = 8
1	-1	1	N = 4
0	-2	4	μ = 2
6	+4	16	
1	-1	1	
		22 = $\sum (X -\mu)2$	

Now in order to find the SS :

X	X2	
1		1
0		0
6		36
1		1
$\sum X = 8$		$\sum X = 38$

SS = $\sum X - (\sum X)2 / N$

= 38 – (8)2 / 4

= 38 – 64 /4

= 38- 16

= 22.

INFERENTIAL STATISITCS:

Inferential statistics consists of techniques that allow us to study samples and then make generalizations about the population from which they were selected.

It is usually not possible to measure everyone in the population. Because population is typically very large, a sample is selected form population which represents population. So by analyzing the results from sample, we hope to make general statements about population.

USES OF INFERENTIAL STATISTICS:

✓ Many top-tiered journals will not publish articles that do NOT use inferential statistics.

✓ Allows you to generalize your findings to the larger Population.

✓ Can determine not just what CAN happen, but what tends to happen in programs like yours?

✓ Helps assess strength of the relationship between your independent (causal) variables, and you Dependent (effect) variables.

✓ Can assess the relative impact of various program Inputs on your program outcomes/objectives.

INFERENTIAL STATISTICS UTILIZED:

✓ Inferential statistics can only be used under the Following conditions.

  ➤ You have a complete list of the members of the population.

  ➤ You draw a random sample from this population

  ➤ Using a pre-established formula, you determine that your sample size is large enough.

In inferential statistics we use following methods to collect data

- Z – scores
- Hypothesis testing
- T statistics
- Probability
- Correlations and Regression
- Chi – square statistics
- Repeated measures analysis of variance (ANOVA)
- Binomial test

Z–Scores

AZ-Score is a statistical measurement of a score's relationship to the mean in a group of scores. A Z-score of 0 means the score is the same as the mean. A Z-score can also be positive or negative, indicating whether it is above or below the mean and by how many standard deviations.In addition to showing a score's relationship to the mean, the Z-score shows statisticians whether a score is typical or atypical for a particular data set. Z-scores also allow analysts to convert scores from different data sets into scores that can be accurately compared to each other. One real-life application of z-scores occurs in usability testing.

In summary, the process of transforming X values into z – scores serves two useful purposes:

1. Each z – scores will tell the exact location of the original X value within the distribution.
2. The z – score will form a standardized distribution that can be directly compared to other distribution that also has been transformed into z – score.

HYPOTHESIS TESTING

A **statistical hypothesis test** is a method of making decisions using data from a scientific study. In statistics, a result is called statistically significant. if it has been predicted as unlikely to have occurred by chance alone, according to a pre-determined threshold probability, the significant level. The phrase "test of significance" was coined by statistician Ronald Fisher. These tests are used in determining what outcomes of a study would lead to a rejection of the null hypothesis, for a pre-specified level of significance; this can help to decide whether results contain enough information to cast doubt on conventional wisdom, given that conventional wisdom has been used to establish the null hypothesis. The *critical region* of a hypothesis test is the set of all outcomes which cause the null hypothesis to be rejected in favor of the alternative hypothesis. Statistical hypothesis testing is sometimes called confirmatory data analysis, in contrast to exploratory data analysis.which may not have pre-specified hypotheses. Statistical hypothesis testing is a key technique of frequents inference..

Statistical hypothesis tests define a procedure that controls (fixes) the probability of incorrectly *deciding* that a default position null hypothesis is incorrect based on how likely it would be for a set of observations to occur if the null hypothesis were true. Note that this probability of making an incorrect decision is not the probability that the null hypothesis is true, nor whether any specific alternative hypothesis is true. This contrasts with other possible techniques of decision theory, in which the null and alternative hypothesis. are treated on a more equal basis. One naive Bayesian approach to hypothesis testing is to base decisions on the posterior probability but this fails when comparing point and continuous hypotheses. Other approaches to decision making, such as Bayesian Decision Theory attempt to balance the consequences of incorrect decisions across all possibilities, rather than

concentrating on a single null hypothesis. A number of other approaches to reaching a decision based on data are available via decision theory and optimal theory some of which have desirable properties, yet hypothesis testing is a dominant approach to data analysis in many fields of science. Extensions to the theory of hypothesis testing include the study of the power of tests, which refers to the probability of correctly rejecting the null hypothesis when a given state of nature exists. Such considerations can be used for the purpose of sample size determination prior to the collection of data.

In a famous example of hypothesis testing, known as the *Lading Tasting Tea* example, a female colleague of Fisher claimed to be able to tell whether the tea or the milk was added first to a cup. Fisher proposed to give her eight cups, four of each variety, in random order. One could then ask what the probability was for her getting the number she got correct, but just by chance. The null hypothesis was that the Lady had no such ability. The test statistic was a simple count of the number of successes in selecting the 4 cups. The critical region was the single case of 4 successes of 4 possible based on a conventional probability criterion ($< 5\%$; 1 of $70 \approx 1.4\%$). Fisher asserted that no alternative hypothesis was (ever) required. The lady correctly identified every cup, which would be considered a statistically significant result.

T- STATISTICS

In statistics, the t statistics is the ratio of departure of an estimated parameter from its notional value to its standard error. It is used in hypothesis testing, for example in student's t test , in the augmented Dickey-Fuller test, and in bootstrapping.

ESTIMATED STANDARD ERROR (s M)

The estimated standard error is used as an estimate of the real standard error   when the value of   is unknown. It is computed from the sample variance or sample standard deviation and provides an estimates of the standard distance between a sample mean M and the population mean μ.

$$s = \sqrt{\frac{s}{n}}$$

PROBABILITY

We define probability for any particular outcome as a fraction or proportion. If the probable outcomes are identified as A, B, C, D, and so on, then

Probability of A = number of outcomes classify

 Total number of possible outcomes

CORRELATIONS AND REGRESSION

Correlation is a statistical technique that is used to measure and describe a relationship between two variables. Usually the two variables are simply observed as they exists naturally in the environment. There is no attempt to control or manipulate the variables.

- CHARACTERISTICS OF RELATIONSHIP
  A correlation measures three characteristics of the relationship between X and Y. these three characteristics are as follows:
  1. The Direction of the Relationship: correlation can be classified into two basic categories :

    Positive Correlation : if the two variables X and Y tend to move in the same direction.

    Negative Correlation : if the two variables X and Y tend to go in the opposite directon.

- The Form of the Relationship : correlation have different forms it might be in straight line , scattered form or there may be no correlation.
- The Degree of the Relationship : a correlation measures how well the data fit the specific form being considered. A perfect correlation always is identified by a correlation of 1.00 and indicates a perfect fit.

USE OF CORRELATION :correlation is used in

- Prediction
- Validity
- Reliability
- Theory verification

REGRESSION

Regression (psychology), a defensive reaction to some unaccepted impulses.Regression analysis, a statistical technique for estimating the relationships among variables. There are several types of regression:

- Linear regression model

- Simple linear regression

- Logistic regression

- Nonlinear regression

- Nonparametric regression

- Robust regression

- Stepwise regression

CHI SQUARE STATISITCS

There are basically two types of random variables and they yield two types of data: numerical and categorical. A chi square (X^2) statistic is used to investigate whether distributions of categorical variables differ from one another. Basically categorical variable yield data inthe categories and numerical variables yield data in numerical form. Responses to such questions as "What is your major?" or Do you own a car?" are categorical because they yield data such as "biology" or "no." In contrast, responses to such questions as "How tall are you?" or "What is your G.P.A.?" are numerical. Numerical data can be either discrete or continuous. The table below may help you see the differences between these two variables.

Data Type	Question Type	Possible Responses
Categorical	What is your sex?	male or female
Numerical	Discrete- How many cars do you own?	two or three
Numerical	Continuous - How tall are you?	72 inches

Notice that discrete data arise from a counting process, while continuous data arise from a measuringprocess.

The Chi Square statistic compares the tallies or counts of categorical responses between two (or more) independent groups. (Chi square tests can only be used on actual numbers and not on percentages, proportions, means, etc.)

REPEATED MEASURES ANALYSIS OF VARIANCE (ANOVA)

Analysis of variance (ANOVA) as a hypothesis testing procedure that is used to evaluate mean differences among two or more treatment conditions.The primary advantage of analysis of variance that it allows researcher to test for significant mean differences in situation. Where there are more than two treatment conditions. The F – ratio in the ANOVA compares all the different sample means in a single test statistic using a single alpha level. If the t statistics was used in situation with more than two sample means, it would require multiple t tests to evaluate all the mean differences. Because each t test involves a risk of a Type I error.

F – RATIO FOR THE REPEATED – MEASURES ANOVA

F – ratio compares the actual mean differences between treatments with the amount of difference would be expected just by chance. As always, F – ratio uses variance to measure the size of the difference. Thus, the F – ratio for the repeated – measures ANOVA has the general structure

> *F = Variance (differences) between treatments /variance (differences) expectedbychance/error*

BINOMIAL TEST

Binomial test uses sample data to evaluate hypothesis about the values of p and q for a population consisting of binomial data. In the binomial test, the null hypothesis specifies exact values for the population proportions p and q .The null hypothesis falls into two categories „*Just Chance and No Change or No Difference.*

❖ **What is important for us to know about inferential statistics?**

✓ You should be able to

1. Read and understand computer printout.

2. Construct tables and graphs from the computer printouts.

3. Interpret and explain these tables and graphs to an audience.

4. Make wise decisions based on valid and accurate data