# STATISTICS IN PSYCHOLOGY

mendura
academic innovators

# CHAPTER 1

## INTRODUCTION TO STATISTICS

### 1.1: DEFINITION

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

Statistics are used to organize and summarize the information so that the researcher can see what happened in the research study and can communicate the results to others. Statistics help the researcher to answer the questions that initiated the research by determining exactly what general conclusions are justified based on the specific results that were obtained. Statistical procedures help ensure that the information or observations are presented and interpreted in an accurate and informative way. In addition, statistics provide researchers with a set of standardized techniques that are recognized and understood throughout the scientific community. Thus, the statistical methods used by one researcher will be familiar to other researchers, who can accurately interpret the statistical analyses with a full understanding of how the analysis was done and what the results signify.

### 1.2: POPULATION AND SAMPLE

**A population is the set of all the individuals of interest in a particular study, whereas a sample is a set of individuals selected from a population, usually intended to represent the population in a research study.**

A population can be quite large—for example, the entire set of women on the planet Earth. A researcher might be more specific, limiting the population for study to women who are registered voters in India. Perhaps the investigator would like to study the population consisting of women who are in a particular state of India too. Populations can obviously vary in size from extremely large to very small, depending on how the investigator defines the population. The population being studied should always be identified by the researcher. In addition, the population need not consist of people—it could be a population of rats, corporations, parts produced in a factory, or anything else an investigator wants to study. In practice, populations are typically very large.

Because populations tend to be very large, it usually is impossible for a researcher to examine every individual in the population of interest. Therefore, researchers typically select a smaller, more manageable group from the population and limit their studies to the individuals in the selected group. In statistical terms, a set of individuals selected from a population is called a sample. A sample is intended to be representative of its population, and a sample should always be identified in terms of the population from which it was selected.

## 1.3: VARIABLES

**A variable is a characteristic or condition that changes or has different values for different individuals.**

Typically, researchers are interested in specific characteristics of the individuals in the population (or in the sample), or they are interested in outside factors that may influence the individuals. For example, a researcher may be interested in the influence of the weather on people's moods. As the weather changes, do people's moods also change? Something that can change or have different values is called a variable.

Variables can be characteristics that differ from one individual to another, such as height, weight, gender, or personality. Also, variables can be environmental conditions that change such as temperature, time of day, or the size of the room in which the research is being conducted.

## 1.4: DATA

**Data (plural) are measurements or observations. A data set is a collection of measurements or observations. A datum (singular) is a single measurement or observation and is commonly called a score or raw score.**

To demonstrate changes in variables, it is necessary to make measurements of the variables being examined. The measurement obtained for each individual is called a datum, or more commonly, a score or raw score. The complete set of scores is called the data set or simply the data.

## 1.5: PARAMETERS AND STATISTICS

**A parameter is a value, usually a numerical value, that describes a population. A parameter is usually derived from measurements of the individuals in the population. A statistic is a value, usually a numerical value, that describes a sample. A statistic is usually derived from measurements of the individuals in the sample.**

When describing data, it is necessary to distinguish whether the data come from a population or a sample. A characteristic that describes a population—for example, the average score for the population—is called a parameter. A characteristic that describes a sample is called a statistic. Thus, the average score for a sample is an example of a statistic.

## 1.6: DESCRIPTIVE AND INFERENTIAL STATISTICS

**Descriptive statistics are statistical procedures used to summarize, organize, and simplify data, where as inferential statistics consist of techniques that allow us to study samples and then make generalizations about the populations from which they were selected.**

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, consists of statistical procedures that are used to simplify and summarize 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 in a table or a graph so that it is possible to see the entire set of scores. Another common technique is to summarize a set of scores by computing an average. Note that even if the data set has hundreds of scores, the average provides a single descriptive value for the entire set.

The second general category of statistical techniques is called inferential statistics. Inferential statistics are methods that use sample data to make general statements about a population.

## 1.7: SAMPLING ERROR

**Sampling error is the naturally occurring discrepancy, or error, that exists between a sample statistic and the corresponding population parameter.**

Because populations are typically very large, it usually is not possible to measure everyone in the population. Therefore, a sample is selected to represent the population. By analyzing the results from the sample, we hope to make general statements about the population. Typically, researchers use sample statistics as the basis for drawing conclusions about population parameters. One problem with using samples, however, is that a sample provides only limited information about the population. Although samples are generally representative of their populations, a sample is not expected to give a perfectly accurate picture of the whole population. There usually is some discrepancy between a sample statistic and the corresponding population parameter. This discrepancy is called sampling error, and it creates the fundamental problem inferential statistics must always address.

## 1.8: DESCRETE AND CONTINUOUS VARIABLES

**A discrete variable consists of separate, indivisible categories. No values can exist between two neighboring categories, whereas For a continuous variable, there are an infinite number of possible values that fall between any two observed values. A continuous variable is divisible into an infinite number of fractional parts.**

A discrete variable consists of separate, indivisible categories. For this type of variable, there are no intermediate values between two adjacent categories. Consider the values displayed when dice are rolled. Between neighboring values—for example, seven dots and eight dots—no other values can ever be observed. Discrete variables are commonly restricted to whole, countable numbers—for example, the number of children in a family or the number of students attending class.

On the other hand, variables such as time, height, and weight are not limited to a fixed set of separate, indivisible categories. You can measure time, for example, in hours, minutes, seconds, or fractions of seconds. These variables are called continuous because they can be divided into an infinite number of fractional parts.

## 1.9: SCALES OF MEASUREMENTS

In data collection we make measurements of our observations. Measurement involves assigning individuals or events to categories. The categories can simply be names such as male/female or employed/unemployed, or they can be numerical values such as 68 inches or 175 pounds. The categories used to measure a variable make up a scale of measurement, and the relationships between the categories determine different types of scales. The distinctions among the scales are important because they identify the limitations of certain types of measurements and because certain statistical procedures are appropriate for scores that have been measured on some scales but not on others. There are four types of scales of measurements.

### 1.9.1: Nominal Scale

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

The word nominal means “having to do with names.” Measurement on a nominal scale involves classifying individuals into categories that have different names but are not related to each other in any systematic way. The measurements from a nominal scale allow us to determine whether two individuals are different, but they do not identify either the direction or the size of the difference. Examples of nominal scales include classifying people by race, gender, or occupation.

### 1.9.2: Ordinal Scale

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

The categories that make up an ordinal scale not only have different names (as in a nominal scale) but also are organized in a fixed order corresponding to differences of magnitude. Often, an ordinal scale consists of a series of ranks (first, second, third, and so on) like the order of finish in a horse race. Occasionally, the categories are identified by verbal labels like small, medium, and large drink sizes at a fast-food restaurant. In either case, the fact that the categories form an ordered sequence means that there is a directional relationship between categories. With measurements from an ordinal scale, you can determine whether two individuals are different and you can determine the direction of difference. However, ordinal measurements do not allow you to determine the size of the difference between two individuals.

### 1.9.3: Interval Scale

**An interval scale consists of ordered categories that are all intervals of exactly the same size. Equal differences between numbers on scale reflect equal differences in magnitude. However, the zero point on an interval scale is arbitrary and does not indicate a zero amount of the variable being measured.**

Interval scale consists of a series of ordered categories (like an ordinal scale) with the additional requirement that the categories form a series of intervals that are all exactly the same size. Thus, the scale of measurement consists of a series of equal intervals, such as inches on a ruler. Other examples of interval and ratio scales are the measurement of time in seconds, weight in pounds, and temperature in degrees Fahrenheit. Note that, in each case, one interval (1 inch, 1 second, 1 pound, 1 degree) is the same size, no matter where it is located on the scale. The fact that the intervals are all the same size makes it possible to determine both the size and the direction of the difference between two measurements. For example, you know that a measurement of 80° Fahrenheit is higher than a measure of 60°, and you know that it is exactly 20° higher.

An interval scale has an arbitrary zero point. That is, the value 0 is assigned to a particular location on the scale simply as a matter of convenience or reference. In particular, a value of zero does not indicate a total absence of the variable being measured. For example a temperature of 0º Fahrenheit does not mean that there is no temperature, and it does not prohibit the temperature from going even lower. Interval scales with an arbitrary zero point are relatively rare.

### 1.9.4: Ratio Scale

**A ratio scale is an interval scale with the additional feature of an absolute zero point. With a ratio scale, ratios of numbers do reflect ratios of magnitude.**

A ratio scale is anchored by a zero point that is not arbitrary but rather is a meaningful value representing none (a complete absence) of the variable being measured. The existence of an absolute, non-arbitrary zero point means that we can measure the absolute amount of the variable; that is, we can measure the distance from 0. This makes it possible to compare measurements in terms of ratios. Ratio scales are quite common and include physical measures such as height and weight, as well as variables such as reaction time or the number of errors on a test.

## CHAPTER 2

## FERQUENCY DISTRIBUTION GRAPH

**A frequency distribution is an organized tabulation of the number of individuals located in each category on the scale of measurement.**

### 2.1: FREQUENCY DISTIBUTION GRAPH

**A frequency distribution graph is basically a picture of the information available in a frequency distribution table.** Even though there are several different types of graphs, but all start with two perpendicular lines called axes. The horizontal line is the X-axis, or the abscissa (ab-SIS-uh). The vertical line is the Y-axis, or the ordinate. The measurement scale (set of X values) is listed along the X-axis with values increasing from left to right. The frequencies are listed on the Y-axis with values increasing from bottom to top. As a general rule, the point where the two axes intersect should have a value of zero for both the scores and the frequencies. A final general rule is that the graph should be constructed so that its height (Y-axis) is approximately two-thirds to three-quarters of its length (X-axis).

#### 2.1.1: Graphs for Interval or Ratio Scale

When the data consist of numerical scores that have been measured on an interval or ratio scale, there are two options for constructing a frequency distribution graph. The two types of graphs are called histograms and polygons.

#### 2.1.2: Graphs for Nominal or Ordinal Scale

When the scores are measured on a nominal or ordinal scale (usually non-numerical values), the frequency distribution can be displayed in a bar graph.

### 2.2: PERCENTILE RANK

**The percentile rank of a particular score is defined as the percentage of individuals in the distribution with scores at or below the particular value. When a score is identified by its percentile rank, the score is called a percentile.**

The percentile rank refers to a percentage and that percentile refers to a score. Also the rank or percentile describes the exact position within the distribution.

## CHAPTER 3

## MEASURES OF CENTRAL TENDENCY

**Central tendency is a statistical measure to determine a single score that defines the center of a distribution. The goal of central tendency is to find the single score that is most typical or most representative of the entire group.**

Central tendency attempts to identify the “average” or “typical” individual. This average value can then be used to provide a simple description of an entire population or a sample. In addition to describing an entire distribution, measures of central tendency are also useful for making comparisons between groups of individuals or between sets of data.

Statisticians have developed three different methods for measuring central tendency: the mean, the median, and the mode. They are computed differently and have different characteristics. To decide which of the three measures is best for any particular distribution, you should keep in mind that the general purpose of central tendency is to find the single most representative score.

### 3.1: THE MEAN

**The mean for a distribution is the sum of the scores divided by the number of scores.**

The mean, also known as the arithmetic average, is computed by adding all the scores in the distribution and dividing by the number of scores. The mean for a population is identified by the Greek letter mu, **μ** (pronounced “mew”), and the mean for a sample is identified by M.

### 3.2: THE MEDIAN

**If the scores in a distribution are listed in order from smallest to largest, the median is the midpoint of the list. More specifically, the median is the point on the measurement scale below which 50% of the scores in the distribution are located.**

The second measure of central tendency we will consider is called the median. The goal of the median is to locate the midpoint of the distribution. Unlike the mean, there are no specific symbols or notation to identify the median. Instead, the median is simply identified by the word median. In addition, the definition and computations for the median are identical for a sample and for a population.

Defining the median as the midpoint of a distribution means that that the scores are being divided into two equal-sized groups. We are not locating the midpoint between the highest and lowest X values. To find the median, list the scores in order from smallest to largest. Begin with the smallest score and count the scores as you move up the list. The median is the first point you reach that is greater than 50% of the scores in the distribution. The median can be equal to a score in the list or it can be a point between two scores.

## 3.3: THE MODE

**In a frequency distribution, the mode is the score or category that has the greatest frequency.**

The final measure of central tendency that we will consider is called the mode. In its common usage, the word mode means “the customary fashion” or “a popular style.” The statistical definition is similar in that the mode is the most frequent observation among a group of scores. As with the median, there are no symbols or special notation used to identify the mode or to differentiate between a sample mode and a population mode. In addition, the definition of the mode is the same for a population and for a sample distribution. The mode is a useful measure of central tendency because it can be used to determine the typical or most frequent value for any scale of measurement, including a nominal scale.

Although a distribution will have only one mean and only one median, it is possible to have more than one mode. Specifically, it is possible to have two or more scores that have the same highest frequency. In a frequency distribution graph, the different modes will correspond to distinct, equally high peaks. A distribution with two modes is said to be ***bimodal***, and a distribution with more than two modes is called ***multimodal***. Occasionally, a distribution with several equally high points is said to have no mode.

# CHAPTER 4

## VARIABILITY

**Variability provides a quantitative measure of the differences between scores in a distribution and describes the degree to which the scores are spread out or clustered together.**

The term variability has much the same meaning in statistics as it has in everyday language; to say that things are variable means that they are not all the same. In statistics, our goal is to measure the amount of variability for a particular set of scores, a distribution. In simple terms, if the scores in a distribution are all the same, then there is no variability. If there are small differences between scores, then the variability is small, and if there are large differences between scores, then the variability is large.

In general, a good measure of variability serves two purposes:

- Variability describes the distribution. Specifically, it tells whether the scores are clustered close together or are spread out over a large distance. Usually, variability is defined in terms of distance. It tells how much distance to expect between one score and another, or how much distance to expect between an individual score and the mean.
- Variability measures how well an individual score (or group of scores) represents the entire distribution.

Mainly, there are three different measures of variability: ***the range, standard deviation, and variance.***

### 4.1: THE RANGE

**Range is the distance covered by the scores in a distribution, from the smallest score to the largest score.**

The obvious first step toward defining and measuring variability is the range, which is the distance covered by the scores in a distribution, from the smallest score to the largest score. Although the concept of the range is fairly straightforward, there are several distinct methods for computing the numerical value. One commonly used definition of the range simply measures the difference between the largest score ($X_{max}$) and the smallest score ($X_{min}$).

## 4.2: THE STANDARD DEVIATION

**Standard deviation is a measure of the standard, or average, distance from the mean, and describes whether the scores are clustered closely around the mean or are widely scattered.**

The standard deviation is the most commonly used and the most important measure of variability. Standard deviation uses the mean of the distribution as a reference point and measures variability by considering the distance between each score and the mean. In simple terms, the standard deviation provides a measure of the standard, or average, distance from the mean, and describes whether the scores are clustered closely around the mean or are widely scattered.

## 4.3: THE VARIANCE

**Variance equals the mean of the squared deviations. Variance is the average squared distance from the mean.**

The process of squaring deviation scores does more than simply get rid of plus and minus signs. It results in a measure of variability based on squared distances. Although variance is valuable for some of the inferential statistical methods covered later, the concept of squared distance is not an intuitive or easy to understand descriptive measure.

**Standard deviation is the square root of the variance and provides a measure of the standard, or average distance from the mean.**

# CHAPTER 5

## CORRELATION

**Correlation describes the linear relationship between two or more variables.**

Correlation is a statistical technique that is used to measure and describe the relationship between two variables. Usually the two variables are simply observed as they exist naturally in the environment—there is no attempt to control or manipulate the variables. For example, a researcher could check high school records (with permission) to obtain a measure of each student's academic performance, and then survey each family to obtain a measure of income. The resulting data could be used to determine whether there is relationship between high school grades and family income. Notice that the researcher is not manipulating any student's grade or any family's income, but is simply observing what occurs naturally.

### 5.1: TYPES OF RELATIONSHIPS

#### 5.1.1: Positive Relationship

**In a positive correlation, the two variables tend to change in the same direction: as the value of the X variable increases from one individual to another, the Y variable also tends to increase; when the X variable decreases, the Y variable also decreases.**

In positive correlation, an increase in one variable is related to an increase in the other, and a decrease in one is related to a decrease in the other. In this type of correlation, a person who scored low on one variable also scored low on the other, an individual with a mediocre score on one variable had a mediocre score on the other, and anyone who scored high on one variable also scored high on the other. In other words, an increase (decrease) in one variable is accompanied by an increase (decrease) in the other

#### 5.1.2: Negative Relationship

**In a negative correlation, the two variables tend to go in opposite directions. As the X variable increases, the Y variable decreases. That is, it is an inverse relationship.**

The negative correlation indicates that an increase in one variable is accompanied by a decrease in the other variable. This correlation represents an inverse relationship: The more of variable X that we have, the less we have of variable Y.

#### 5.1.3: No Relationship

It is also possible to observe no meaningful relationship between two variables. The correlation coefficient for these data is very close to 0.

## 5.1.4: Curvilinear Relationship

A correlation coefficient of 0 indicates no meaningful relationship between two variables. However, it is also possible for a correlation coefficient of 0 to indicate a curvilinear relationship. The strong positive relationship depicted in the left half of the graph essentially cancels out the strong negative relationship in the right half of the graph. Although the correlation coefficient is very low, we would not conclude that there is no relationship between the two variables. As the figure shows, the variables are very strongly related to each other in a curvilinear manner, with the points being tightly clustered in an inverted U shape.

## 5.2: MAGNITUDE

**Magnitude is an indication of the strength of the relationship between two variables.**

The magnitude, or strength, of a relationship is determined by the correlation coefficient describing the relationship. As we saw in Module 6, a correlation coefficient is a measure of the degree of relationship between two variables; it can vary between 1.00 and 1.00. The stronger the relationship between the variables, the closer the coefficient is to either 1.00 or 1.00. The weaker the relationship between the variables, the closer the coefficient is to 0.

A correlation coefficient of either 1.00 or 1.00 indicates a perfect correlation the strongest relationship possible. For example, if height and weight were perfectly correlated (1.00) in a group of 20 people, this coefficient would mean that the person with the highest weight was also the tallest person, the person with the second-highest weight was the second-tallest person, and so on down the line. In addition, in a perfect relationship each individual's score on one variable goes perfectly with his or her score on the other variable.

## 5.3: THE PEARSON CORRELATION

**The Pearson correlation measures the degree and the direction of the linear relationship between two variables.**

The Pearson correlation for a sample is identified by the letter r. The corresponding correlation for the entire population is identified by the Greek letter rho (ρ), which is the Greek equivalent of the letter r. Conceptually, this correlation is computed by

$$r = \frac{\text{degree to which } X \text{ and } Y \text{ vary together}}{\text{degree to which } X \text{ and } Y \text{ vary separately}}$$

$$= \frac{\text{covariability of } X \text{ and } Y}{\text{variability of } X \text{ and } Y \text{ separately}}$$

## 5.4: THE SPEARMAN CORRELATION

When the Pearson correlation formula is used with data from an ordinal scale (ranks), the result is called the Spearman correlation. The Spearman correlation is used in two situations.

- First, the Spearman correlation is used to measure the relationship between X and Y when both variables are measured on ordinal scales. An ordinal scale typically involves ranking individuals rather than obtaining numerical scores.
- In addition to that, Spearman correlation can be used as a valuable alternative to the Pearson correlation, even when the original raw scores are on an interval or a ratio scale.

## 5.5: THE POINT BISERIAL CORRELATION

The point-biserial correlation is used to measure the relationship between two variables in situations in which one variable consists of regular, numerical scores, but the second variable has only two values. A variable with only two values is called a dichotomous variable or a binomial variable.

# CHAPTER 6

## REGRESSION

**The statistical technique for finding the best-fitting straight line for a set of data is called regression, and the resulting straight line is called the regression line.**

The goal for regression is to find the best-fitting straight line for a set of data. To accomplish this goal, however, it is first necessary to define precisely what is meant by “best fit.” For any particular set of data, it is possible to draw lots of different straight lines that all appear to pass through the center of the data points. Each of these lines can be defined by a linear equation of the form $Y = bX + a$ where b and a are constants that determine the slope and Y-intercept of the line, respectively. Each individual line has its own unique values for b and a. The problem is to find the specific line that provides the best fit to the actual data points.

### 6.1: LINEAR EQUATION

In general, a linear relationship between two variables X and Y can be expressed by the equation,

$\mathbf{Y} = \boldsymbol{b}\,\mathbf{X} + \boldsymbol{a}$ ,where ‘***a***’ and ‘***b***’ are fixed constants.

In the general linear equation, the value of b is called the slope. The slope determines how much the Y variable changes when X is increased by one point.

### 6.2: MULTIPLE REGRESSION

The process of using several predictor variables to help obtain more accurate predictions is called multiple regression.

# CHAPTER 7

## PROBABILITY

**The probability for any specific outcome is defined as a fraction or a proportion of all the possible outcomes. If the possible outcomes are identified as A, B, C, D, and so on, then,**

$$\boldsymbol{Probability\ of\ A = \frac{Number\ of\ outcomes\ classified\ as\ A}{Total\ number\ of\ possible\ outcomes}}$$

For example, if you are selecting a card from a complete deck, there are 52 possible outcomes. The probability of selecting the king of hearts is p = 1/52. The probability of selecting an ace is p = 4/52 because there are 4 aces in the deck. The value of probability ranges from 0 to 1.

### 7.1: SAMPLING

**Sampling is the process of selecting representative group from a target population for conducting research.**

If we are to draw reliable and valid conclusions concerning the population, it is imperative that the sample be like the population, that is, a representative sample. With a representative sample we can be fairly confident that the results we find based on the sample also hold for the population. In other words, we can generalize from the sample to the population. There are two ways to sample individuals from a population: probability and nonprobability sampling.

#### 7.1.1: Probability Sampling

**A sampling technique in which each member of the population has an equal likelihood of being selected to be part of the sample.**

When researchers use probability sampling, each member of the population has an equal likelihood of being selected to be part of the sample. Mainly there are three types of probability sampling namely; **simple random sampling, stratified random sampling, and cluster sampling**.

##### 7.1.1.1: Simple Random Sampling

**A sampling technique in which the researchers randomly selecting a certain number of individuals from the population is called simple random sampling.**

Generating a random sample can be accomplished by using a table of random numbers. When using a random numbers table, the researcher chooses a starting place arbitrarily. Once the starting point is determined, the researcher looks at the number say, six counts down six people in the population, and chooses the sixth person to be in the sample. The researcher continues in this manner by looking at the next number in the table, counting down through the population, and including the appropriately numbered person in the sample. For our sample we continue this process until we select a sample of required number. A random sample can be generated in other ways such as by computer or by pulling names randomly out of a hat. The point is that in random sampling each member of the population is equally likely to be chosen as part of the sample.

#### 7.1.1.2: Stratified Random Sampling

**A sampling technique designed to ensure that subgroups or strata are fairly represented.**

Sometimes a population is made up of members of different groups or categories. For instance, both men and women make up the 300 students enrolled in introductory psychology but maybe not in equal proportions. If the researchers want to draw conclusions about the population of introductory psychology students based on our sample, then our sample must be representative of the strata within the population.

One means of attaining such a sample is stratified random sampling. A stratified random sample allows the researchers to take into account the different subgroups of people in the population and to guarantee that the sample accurately represents the population on specific characteristics.

#### 7.1.1.3: Cluster Sampling

**A sampling technique in which clusters of participants that represent the population are used.**

Often the population is too large for random sampling of any sort. In these cases, it is common to use cluster sampling. As the name implies, cluster sampling involves using participants who are already part of a group, or cluster. For example, if a researcher is interested in surveying students at a large university where it might not be possible to use true random sampling, he might sample from classes that are required of all students at the university such as English composition. If the classes are required of all students, they should contain a good mix of students, and if he use several classes, the sample should represent the population.

### 7.1.2: Nonprobability Sampling

**A sampling technique in which the individual members of the population do not have an equal likelihood of being selected to be a member of the sample.**

Nonprobability sampling is used when the individual members of the population do not have an equal likelihood of being selected to be a member of the sample. Nonprobability sampling is typically used because it tends to be less expensive and generating samples is easier.

Mainly there are two types of nonprobability sampling namely; purposive sampling and quota sampling.

#### 7.1.2.1: Purposive Sampling

**A sampling technique in which participants are obtained wherever they can be found and typically wherever it is convenient for the researcher.**

Purposive sampling involves getting participants wherever the researcher can find them and normally wherever is convenient. This method is sometimes referred to as haphazard sampling or convenience sampling. This approach might sound similar to cluster sampling, but there is a difference. With cluster sampling we try to identify clusters that are representative of the population. With purposive sampling, however, researchers simply use who- ever is available as a participant in the study.

#### 7.1.2.2: Quota Sampling

**A sampling technique that involves ensuring that the sample is like the population on certain characteristics but uses purposive sampling to obtain the participants.**

Quota sampling is to nonprobability sampling what stratified random sampling is to probability sampling. In other words, quota sampling involves ensuring that the sample is like the population on certain characteristics. However, even though we try to ensure similarity with the population on certain characteristics, we do not sample from the population randomly. We simply take participants wherever we find them, through whatever means is convenient. Thus, this method is slightly better than convenience sampling, but there is still not much effort devoted to creating a sample that is truly representative of the population nor one in which all members of the population have an equal chance of being selected for the sample.

# CHAPTER 8

## HYPOTHESIS TESTING

**A hypothesis test is a statistical method that uses sample data to evaluate a hypothesis about a population.**

### 8.1: NULL HYPOTHESIS

**The null hypothesis ($H_0$) states that in the general population there is no change, no difference, or no relationship. In the context of an experiment, H0 predicts that the independent variable (treatment) has no effect on the dependent variable (scores) for the population.**

The first and most important of the two hypotheses is called the null hypothesis. The null hypothesis states that the treatment has no effect. In general, the null hypothesis states that there is no change, no effect, no difference—nothing happened, hence the name null. The null hypothesis is identified by the symbol $H_0$. (The H stands for hypothesis, and the zero subscript indicates that this is the zero-effect hypothesis.)

### 8.2: ALTERNATE HYPOTHESIS

**The alternative hypothesis ($H_1$) states that there is a change, a difference, or a relationship for the general population. In the context of an experiment, $H_1$ predicts that the independent variable (treatment) does have an effect on the dependent variable.**

The second hypothesis is simply the opposite of the null hypothesis, and it is called the scientific, or alternative, hypothesis ($H_1$). This hypothesis states that the treatment has an effect on the dependent variable.

### 8.3: TYPE I ERROR

**A Type I error occurs when a researcher rejects a null hypothesis that is actually true. In a typical research situation, a Type I error means the researcher concludes that a treatment does have an effect when in fact it has no effect.**

A Type I error occurs when a researcher unknowingly obtains an extreme, nonrepresentative sample. In most research situations, the consequences of a Type I error can be very serious. Because the researcher has rejected the null hypothesis and believes that the treatment has a real effect, it is likely that the researcher will report or even publish the research results. A Type I error, however, means that this is a false report. Thus, Type I errors lead to false reports in the scientific literature. Other researchers may try to build theories or develop other experiments based on the false results. A lot of precious time and resources may be wasted.

## 8.4: TYPE II ERROR

**A Type II error occurs when a researcher fails to reject a null hypothesis that is really false. In a typical research situation, a Type II error means that the hypothesis test has failed to detect a real treatment effect.**

Whenever a researcher rejects the null hypothesis, there is a risk of a Type I error. Similarly, whenever a researcher fails to reject the null hypothesis, there is a risk of a Type II error. By definition, a Type II error is the failure to reject a false null hypothesis. In more straightforward English, a Type II error means that a treatment effect really exists, but the hypothesis test fails to detect it.

A Type II error occurs when the sample mean is not in the critical region even though the treatment has an effect on the sample. Often this happens when the effect of the treatment is relatively small. In this case, the treatment does influence the sample, but the magnitude of the effect is not big enough to move the sample mean into the critical region. Because the sample is not substantially different from the original population, the statistical decision is to fail to reject the null hypothesis and to conclude that there is not enough evidence to say there is a treatment effect. The consequences of a Type II error are usually not as serious as those of a Type I error.

## 8.5: DEGREES OF FREEDOM

**Degrees of freedom describe the number of scores in a sample that are independent and free to vary. Because the sample mean places a restriction on the value of one score in the sample, there are (n – 1) degrees of freedom for a sample with n scores**

## 8.6: t STATISTICS

**'t' statistics is a statistical technique used to estimate the mean difference between two variables or between two group on a specific variable.**

## 8.7: ONE-WAY ANALYSIS OF VARIANCE

**One-way analysis of variance (abbreviated one-way ANOVA) is a statistical technique that can be used to compare means of two or more samples (using the F distribution)**

## 8.8: TWO-WAY ANALYSIS OF VARIANCE

**The two-way ANOVA compares the mean differences between groups that have been split on two independent variables (called factors). The primary purpose of a two-way ANOVA is to understand if there is an interaction between the two independent variables on the dependent variable.**

## 8.9: REFERENCES

Gravetter, F. J., & Wallnau, L. B. (2016). *Statistics for The Behavioral Sciences*. Boston, MA: Cengage Learning.

Jackson, S. (2010). *Research Methods: A Modular Approach*. Boston, MA: Cengage Learning.

#Disclaimer: This note is prepared for learning the basics of statistical concepts and based on the PG entrance pattern only. It is not a complete collection of statistical techniques or principles.