

Levels of Measurement and Type I & Type II Errors

[See online here](#)

This article explains the concepts of variables and data levels of measurement. There are two types of variables, continuous and discrete. The four scales of data measurement used in statistics are nominal, ordinal, interval, and ratio measurement. Causality refers to a relation between two variables. The article also discusses the frequency and normal distribution and shows examples. Central limit theorem and type I and II errors are also described in this article.



Measurement

Measurement of data refers to the **quantification of a quality attributed to a data set**. It refers to the scales that are used for measurement in statistics. These scales help in the categorization of data. Each measurement scale has different properties that should be used in case the properties of the scale match with the properties of the data set. Different levels of measurement are used to analyze a data set appropriately.

Variables

A variable in statistics is a value that is changing continuously due to change in different factors. It can be a characteristic, a quantity, or a number that is measurable and continually changing. Some common examples of variables include class, grades,

expenses, vehicles, etc.

There are two types of variables:

Continuous variable

Continuous variables are not countable due to a bulk of values present in data. It refers to a never-ending list of variables with infinite values. For example, the age of a person's pet might be 2 years, 2 months, 5 hours, 5 seconds, 2 milliseconds, and the list goes on. Likewise, the weight of students in a class can be an example of a continuous variable.

Discrete variable

This type of variable is countable in a finite amount of time. For example, the number of plates placed on a table or the grades of students in a class is countable at a specific time point. No matter how long the series of discrete variables is, it is still logically countable with an ending number.

Causality

Causality refers to the **interlink between two variables**. It describes a situation in which one event is caused by another event (i.e. the cause-effect relationship between the two variables). It shows the co-relation in which a change in one variable causes changes in other variables. Or it shows that the occurrence of one event results in the occurrence of a second event. The relationship between the two variables depicts the pattern of causality in statistics.

Example:

The 'hours worked' and 'income earned' are two interlinked factors for which a change in one factor results in a change in the second factor. If a higher number of hours have been worked, it leads to the higher income of a worker and vice versa. The same occurs as an inverse relationship in the case of price and purchasing power. If the price of commodities in the market increases, the purchasing power of consumers goes down.

This concept should be **differentiated from correlation**, which is the statistical measure expressed as a number that describes the size and direction of a relationship between two or more variables.

Levels of Measurement

There are **four scales of data measurement used in statistics**: nominal, ordinal, interval, and ratio. The last two levels of measurement are sometimes known as continuous or scale measurements. Different levels of measurement are used to analyze a dataset appropriately.

Different levels of measurement are shown as follows:

	Nominal	Ordinal	Interval	Ratio
Classification	✓	✓	✓	✓
Magnitude		✓	✓	✓
Equal interval			✓	✓
True zero				✓

Math permissible	count	count, rank	count, rank, add, subtract	count, rank, add, subtract, multiply, divide
-------------------------	-------	-------------	----------------------------	--

Nominal Measurement

Nominal measurement is the first level of data measurement. It is considered the **lowest form of measurement** that refers to the name only of the data set. It deals with the labels, names, and categories of a data set. The data at this level is normally qualitative and simple in nature.

Example:

Suppose a study has been conducted to record the number of individuals with blue eyes in a population. The survey will generate simple responses through nominal measurement in the form of yes or no. Another example of nominal measurement is the number attributed to a football player according to the name and number printed on the back of each player's shirt. This is a form of nominal measurement of a player or data set.

This type of measurement is not helpful as a meaningful contribution to the calculation of the mean or the standard deviation of a data set.

Ordinal Measurement

The second and comparatively advanced level of nominal measurement is named ordinal measurement. This data is **arranged in a sequence and ordered properly**, but it is unable to form a meaningful contribution to the calculation of the mean and the standard deviation.

Example:

The list of top ten football players of the world gives a sequence of the performance level of players from one to ten. Still, the list does not give appropriate reasons or differences for which players are ranked from level 1 to 10. It is not possible to measure the efficiency level of each player. Like nominal measurement data, the ordinal level of data also does not contribute to meaningful statistical calculations.

Interval Measurement

Interval measurement is a more advanced level of measurement for which the data not only can be measured, but the **differences between datasets can also be identified**. This scale of measurement helps to rank data and also to convey information. The distance between values in a data set at the interval measurement level is considered to be uniform, and the averages of data can be computed easily.

Example:

The most common example of interval measurement data is Fahrenheit and Celsius temperature scales. It can be demonstrated that a 20-degree temperature is 60 degrees less than 80 degrees. On the other hand, 0 degrees in both scales represents cold. This level of data is feasible for calculations of the mean, the standard deviation, and other statistical measurements.

Ratio Measurement

Ratio measurement is an advanced form of data measurement that includes **all attributes of interval measurement along with the zero value**. Due to the presence or rationality of the zero factor, it is possible to make a comparison of different data sets, like twice or four times, at the ratio level.

Example:

Distance is the most common example of the ratio measurement level. Zero feet represents no distance, and 2 feet represents twice the distance than 1 foot. This form of data is measurable and comparable. This level of data not only provides the opportunity for the calculation of sums and differences but also for ratio analysis. It results in the meaningful interpretation of data.

Frequency Distribution

Frequency is the **possibility of the occurrence of something over a specific period of time**. For example, John played football on Wednesday morning, Wednesday afternoon, and Friday morning. The frequency of football games played by John is 2 on Wednesday, 1 on Friday, and 3 for the whole week.

Frequency distribution process:

The range of frequencies presented in a table displays how often certain values or outcomes occur. Suppose Sam is a player on a high school football team. Sam has scored various numbers of goals in recent games. The range is given as follows: 2, 3, 1, 2, 1, 3, 3, 2, 4. The frequency of scores is shown as follows:

Score	Frequency
1	2
2	3
3	3
4	1

The above frequency distribution of goals indicates that Sam often scores 2 and 3 goals. The possibility of 1 goal is 2, and the possibility of 4 goals is at least 1.

Normal Distribution

Normal distribution is a well-known type of continuous probability distribution. This type of distribution is helpful in natural and social sciences to **demonstrate random variables and their relationship**. Due to the central limit theorem, the importance of normal distribution has increased.

The curve of normal distribution is called a bell curve as shown below. The data in a normal distribution are distributed around central values with no bias on the left and right sides. The curve of normal distribution is called a bell curve due to its resemblance to the shape of a bell.

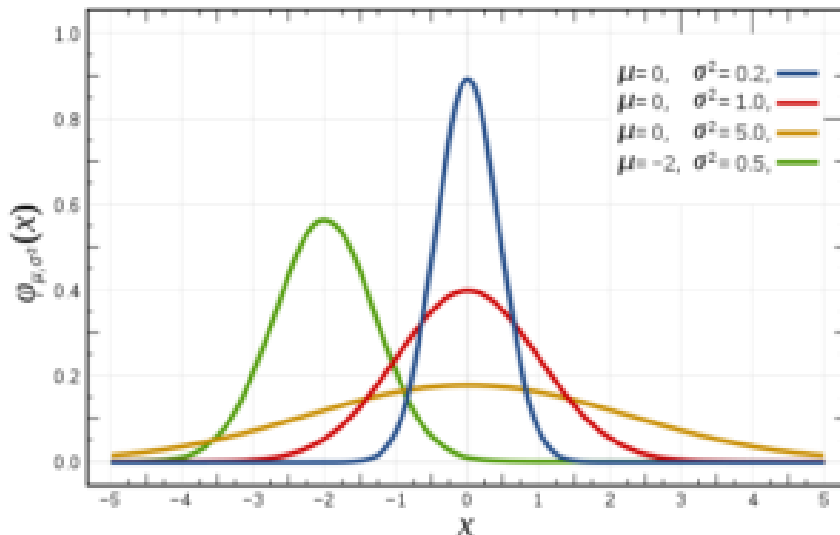


Image: "Probability density function for the normal distribution" by Inductiveload. License: [Public Domain](#)

Central Limit Theorem

The central limit theorem is a statistical theory which states that the mean of **samples taken from a large population with finite variance** is the same as the mean of the whole population. The central limit theorem is based on the normal distribution pattern, which means that the graph of the sample means will look more like a normal distribution curve. Similarly, the **variance of all samples is equal to the variance of the whole population.**

A hypothesis is a statement that predicts the relation between certain population parameters being studied. A null hypothesis (H_0) states that no statistical relation or significance exists between the variables or parameters.

A researcher usually attempts to disprove or reject the null hypothesis (H_0) in a study to support a particular point. The alternative hypothesis (H_a), on the other hand, states that statistically, a significant relation exists between the variables.

Type I and Type II Errors

In statistical hypothetical testing, type I error is the **rejection of a true null hypothesis**. It is **also called false positive**. It states that a supposed relationship between two variables exists when it actually does not exist.

For example, a patient is declared to have a deadly disease who in reality does not have any disease. Another example of a type I error is the buzz of a fire alarm when there is no actual fire.

The following example illustrates type I and type II errors when the null hypothesis is that a man on trial is innocent.

Example:

	H is true and he is truly innocent	H_a is true and he is actually guilty
--	------------------------------------	---

Jury fails to reject H and man is found not guilty	Right decision	Type II error
Jury rejects H and man is found guilty	Type I error	Right decision

A type I error is an error of rejection of the null hypothesis when the hypothesis is actually true. It refers to the observation of a difference when, in reality, there is no difference. It occurs when the null hypothesis (H) is true. It refers to a false hit when a factor or thing actually is not present. The probability of significance level in type I error is denoted by alpha (α). The probability of type I error is denoted by $P(R | H)$. Here R denotes the rejection region of a true null hypothesis.

Type II error or false negative is the opposite of type I error. In this case, the null hypothesis is not rejected when it is actually wrong. The alternative hypothesis, in this case, is a true state of hypothesis instead of a null hypothesis. This error refers to the failure of acceptance of an alternative hypothesis when it is right. It leads to the failure of identification of any difference between the null hypothesis and the alternative hypothesis when there actually is one. The probability of type II error is denoted by $1 - P(R | H)$.

References

[Levels of Measurement](#) via socialresearchmethods.net

[Normal Distribution](#) via mathsisfun.com

[Central Limit Theorem](#) via stat.wmich.edu

Legal Note: Unless otherwise stated, all rights reserved by Lecturio GmbH. For further legal regulations see our [legal information page](#).