Before you Watch

Has anyone ever said to you that something is wicked or sick, meaning it is awesome or great? In chemistry a solution may refer to a liquid mixture, whilst in mathematics a solution may refer to the value which solves an equation or a means of solving a mathematical problem. Ever heard of business terms like hedging or leverage? Engineers may use the word gauge to refer to the width of something (e.g., thickness of sheet metal, or diameter of a screw), whilst scientists may use it to refer to the instrument that measures content (e.g., a rain gauge). Every field has its jargon.

Learning Statistics involves learning the meaning of many new terms AND most importantly learning the correct statistical meanings of a lot of familiar terms.

While one term can be broad and interchangeable, another can be very specific and therefore it may be hard to find correct substitute language or terms.

So look for the keywords and terms. Perhaps start your own glossary so you may more easily check word meanings should you ever get stuck.

The Video Content

What are you interested in measuring? Is it an individual's Blood Glucose Level? Or perhaps their temperature throughout the day?

Perhaps you're interested in the tensile strength of cables made from a certain steel composite? Or even the reaction time of chemicals?
Whatever the measure of interest, it is referred to as being a random variable, as it may assume any number of possible values, or responses, (usually within some range). It is a characteristic of the individual, or item, or subject, we are studying.

The value, or response, for the random variable, or, 'characteristic of interest', will vary depending upon the individual or item we happen to be observing, or the point in time at which we do so. The specific values that we observe for a variable are known as data.

Knowing a variable's TYPE is important as it determines which analyses, or methods of reporting or displaying the data, are appropriate to apply.

The first step is to determine whether the variable is CATEGORICAL or NUMERICAL.

Do the responses to the variable fall into categories? For example, if measuring 'Sex' of each person, the possible responses are either 'Male' or 'Female' … each person falls into one of those two categories. Hence the variable, or 'measure of interest' or 'characteristic' 'Sex' is a categorical variable.

What about when the possible responses are numbers which reflect the quantity, or amount, of something? For example if measuring 'Blood glucose level' of each person, the value we observe is the quantity of glucose in the blood, usually measured in millimoles per Litre. So the variable 'Blood glucose level' is a numerical variable.

The Second step further identifies Categorical variables as either Nominal or Ordinal, and Numerical Variables as either Discrete or Continuous.

If the categorical variable's responses have no natural order to them, then it is called a nominal variable. For example, for the variable 'Sex' the possible responses are Male and Female, and there is no reason why Male should be listed before Female nor vice versa, there is NO natural order. So 'Sex' is a nominal variable.

If the categorical variable's responses have a natural order, then it is called an ordinal variable. For example 'Risk of developing coeliac disease'. If the possible responses are 'Low', 'Medium', or 'High' then we can see the categories have a natural order, Low comes before Medium, which comes before High on the scale of Low through to High.

If the numerical variable's possible responses are numbers arising from a measuring process and can assume only certain or distinct values, then it is called a discrete variable. For example, consider the variable 'Daily number of adverse events' which measures the number of undesirable outcomes occurring in a hospital ward each day. We can observe either 0, or 1, or 2, or 3, and so on events. But we cannot observe 1.35 events- it must be a whole number. This is an example of a discrete variable (it can only assume certain values).

If the numerical variable's possible responses are numbers arising from a measuring process and can assume ANY value within some range or along some continuum, then it is called a continuous variable. For example, Blood Glucose Level. We can record the result as precisely as the measuring device or instrument allows, we could observe values like 5.54 millimoles per Litre or even to more decimal places depending on the measuring instrument. NOTE: it is not the measuring instrument that determines whether a variable is discrete or continuous. The fact that Blood Glucose Level could be measured to be any value within some range is what determines that the variable Blood Glucose Level is continuous.

When we record data for variables, we consider each individual as a record - this is presented as a row in a data table. Then we would record the data for each variable in columns.

So if Person 1 was Male and had a Blood Glucose Level of 5.45 and we recorded similar values for each person, then we see the datatable build.

So, you've now learnt about the four variable types and are getting some data...so it's time to build upon this and see how we might present data arising from each of these variable types...you can learn more about this in the next video in this series.

Now What?

This video is designed to give you an introduction to the concept of a random variable and how to determine a variable's type. You are now ready to look at how we may turn the data for a given variable into readily interpretable information (see the videos Visual Displays - Single Variable, and then Visual Displays - Two Variables).

But, when am I going to use this?

Statistics is essential in all fields of study - it provides a deeper understanding and credibility of investigations within the field. Variation is all around us, Statistics helps us quantity and learn from this variation. Statistics is fundamental to the design and analysis of experiments or investigations, for example, medical trials, …. It's also used in quality control…

Other Links

A YouTube video made by the Statistics Learning Centre in NZ gives an overview of data types using animation that is easy to understand and fun. Note that numerical data types has been given alternative names. Interval and ratio data are used instead of discrete and continuous. It is a slightly different approach but the same principle.

https://www.youtube.com/watch?v=hZxnzfnt5v8

If you really want to understand types of data, along with appropriate statistics and graphs, you can learn on our new Snack-size course. Takes about an hour, and lots of fun!

http://www.statslc.com/snack/