MEASUREMENT SCALES

INTRODUCTION

Variables differ in *how well* they can be measured, that is, in how much measurable information their measurement scale can provide. There is obviously some measurement error involved in every measurement, which determines the *amount of information* that we can obtain.

Another factor that determines the amount of information that can be provided by a variable is its type of *measurement scale*. Specifically variables are classified under two categories – *qualitative* and *quantitative*.

QUALITATIVE (CATEGORICAL) DATA

Qualitative, also known as *categorical*, data cannot be measured on a numerical scale (quantified). Examples of categorical variables are *gender* (male or female) and *size of T-shirt* (XXS, XS, S, M, L, XL and XXL); yet, these two variables differ in a sense: the first is said to be *nominal* or *purely categorical* whilst the second is known as *ordinal*.

Nominal (purely categorical) data

Nominal variables allow for only qualitative classification. That is, they can be measured only in terms of whether the individual items belong to some distinctively different categories, but we cannot quantify or even rank order those categories. For example, all we can say is that 2 individuals are different in terms of a certain variable (for example, they are of different race), but we cannot say which one *has more* of the quality represented by the variable. Typical examples of nominal variables are gender, race, color, city, marital status, etc.

Example 1

Marital status

1.	Never married	☐	4.	Married/Cohabiting	☐
2.	Divorced	☐	5.	Separated	☐
3.	Widowed	☐			

Clearly, the numbers associated with the options above have no numerical significance so that, not only comparison between values is impossible, but also descriptive statistics like the mean and standard deviation would make no sense if calculated.

Ordinal data

Ordinal variables allow us to rank order the items we measure in terms of which has less and which has more of the quality represented by the variable, but still they do not allow us to say *how much more*. A typical example of an ordinal variable is the socioeconomic status of families. For example, we know that upper-middle is higher than middle but we cannot say that it is, for example, 18% higher. Also, this very distinction between nominal, ordinal, and interval scales itself represents a good example of an ordinal variable. For example, we can say that nominal measurement provides less information than ordinal measurement, but we cannot say *how much less* or how this difference compares to the difference between ordinal and interval scales.

Example 2

Employee's performance

1.	Excellent	☐	4.	Poor	☐
2.	Good	☐	5.	Very poor	☐
3.	Average	☐			

It can be easily deduced that 'Excellent' is *better than* 'Poor', that is, there is a latent scale on which comparison can be made among the various values.

Ordinal data can sometimes be treated as *interval* (see Section B3.2.1 below) for the sake of statistical analysis, provided the assumption is founded. In such a case, the values of the variable are mathematically considered to be 'equidistant' on its scale. In such a case, the numbers associated with each value starts to get some numerical significance so that the mean, though not very convincingly, may be statistically interpreted.

The variable 'Employee's performance' in *Example 2* above can be regarded as *interval* if we assume that the 'distance' between any pair of successive values is *equal* (for example, the distance between 'Excellent' and 'Good' is the same as that between 'Average' and 'Poor'). In such as case, if the average performance score of 100 employees is calculated and found to be, say, 3.2, we may, within some margin of security, conclude that the overall performance of employees is just above 'Average', the latter having been assigned a value of 3.

However, it would be statistically dangerous to assume an interval scale for the following example.

Example 3

Educational level

1.	None	☐	5.	Diploma	☐
2.	Primary	☐	6.	Degree	☐
3.	Vocational	☐	7.	Postgraduate	☐
4.	Secondary	☐	8.	Professional	☐

It is clear that the 'distance' between 'None' and 'Primary' is not equal to that between 'Diploma' and 'Degree'.

QUANTITATIVE (NUMERICAL) DATA

Quantitative data can be easily measured on a numerical scale; variables which can be quantified in terms of units are all quantitative. Examples of quantitative variables are *number of students per class* and *height* (measured in centimetres). Yet again, these two variables differ in their nature: the first is said to be *discrete* whereas the second is *continuous*.

Discrete data

Discrete data occur as definite and separate values; a discrete variable assumes values which are *countable* so that there are *gaps* between its successive values. For example, when counting the number of children in a class, we use *natural numbers* (0, 1, 2, …, *n*).

Continuous data

Continuous data occur as the whole set of *real* numbers or a subset of it. In other words, there are no gaps between successive so that a continuous variable assumes *all* the values (including all the decimals) between given boundaries. *Temperature* is a good example of a continuous variable – though thermometer readings are recorded to the nearest tenth of a degree (Centigrade or Fahrenheit), temperature does not 'jump' from, for example, 17.1^0 C to 17.2^0 C. It passes through all the real numbers between these two values. *Height*, *weight* and *speed* are also continuous variables.

Continuous data can be measured on *interval* and *ratio* scales (explained below).

Interval scale

Interval variables allow us not only to *rank* order the items that are measured, but also to quantify and compare the sizes of differences between them. For example, temperature, as measured in degrees Fahrenheit or Celsius, constitutes an interval scale. We can say that a temperature of 40 degrees is higher than a temperature of 30 degrees, and that an increase from 20 to 40 degrees is twice as much as an increase from 30 to 40 degrees. However, interval scale variables do not have an *absolute zero*. If the temperatures in Singapore and London are 30^0 C and 15^0 C respectively, we cannot say that it is twice as hot in Singapore than it is in London. This is simply because it would not be the case if these temperature were measured in degrees Fahrenheit: 86^0 C and 59^0 F respectively.

Ratio scale

Ratio variables are very similar to interval variables; in addition to all the properties of interval variables, they feature an identifiable absolute zero point, thus they allow for statements such as *x* is two times more than *y*. Typical examples of ratio scales are measures of time or space. For example, as the Kelvin temperature scale is a ratio scale, not only can we say that a temperature of 200 degrees is higher than one of 100 degrees, we can correctly state that it is twice as high. Interval scales do not have the ratio property. Most statistical data analysis procedures do not distinguish between the interval and ratio properties of the measurement scales. Height is also a ratio scale variable since, if a person is twice as tall as another, he/she will remain so, irrespective of the units used (centimetres, inches, etc…).

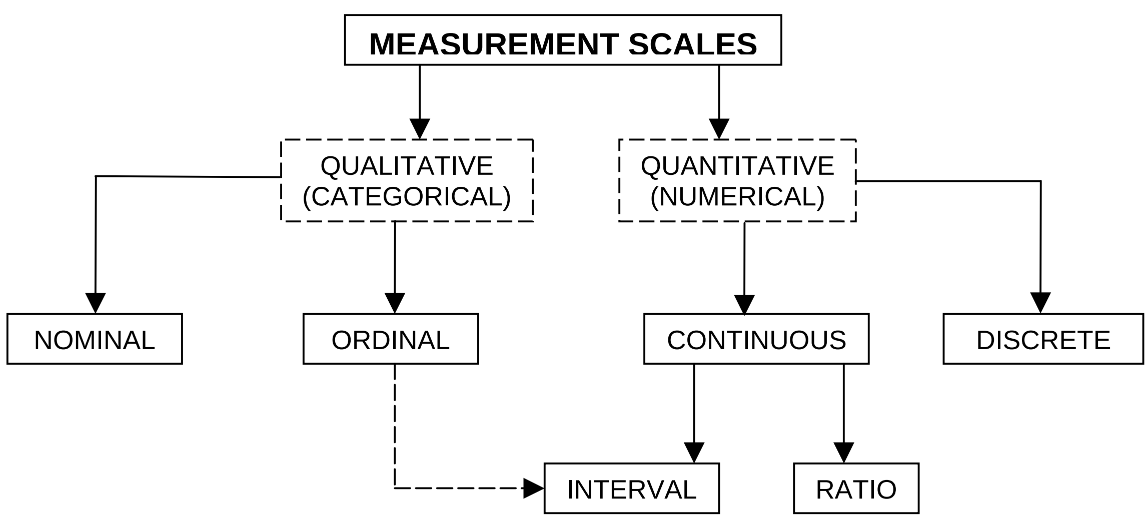

Fig. B3.2.2