Types of Data in Statistics
(Level of measurement)

Level of measurement (Scale) is a classification that describes the nature of information within the values assigned to variables.

Four levels or scales have been supposed as the best-known classification: nominal, ordinal, interval, and ratio unifying both "qualitative" (which are described by "nominal" type) and "quantitative" (to a different degree, all the rest of scales).

Nominal level

The nominal type differentiates between items or subjects based only on their names or meta-categories and other qualitative classifications they belong to. Examples of these classifications include gender, nationality, ethnicity, language, style, biological species, ... Other concrete examples are

- In grammar, the parts of speech: noun, verb, preposition, article, pronoun, etc.
- In biology, the taxonomic ranks below domains: Archaea, Bacteria, and Eukarya
- in software engineering, type of faults: specification faults, design faults, and code faults

Nominal scales were often called qualitative scales, and measurements made on qualitative scales were called qualitative data. However, the rise of qualitative research has made this usage confusing.

If numbers are assigned as labels in nominal measurement, they have no specific numerical value or meaning. No form of arithmetic computation (+, −, ×, etc.) may be performed on nominal measures.

The nominal level is the lowest measurement level used from a statistical point of view.

Central tendency

The mode, i.e. the *most common* item, is allowed as the measure of central tendency for the nominal type.

On the other hand, the median, i.e. the *middle-ranked* item, makes no sense for the nominal type of data since ranking is meaningless for the nominal type.

Ordinal scale

The ordinal type allows for rank order (1st, 2nd, 3rd, etc.) by which data can be sorted, but still does not allow for relative *degree of difference* between them. Examples include, a spectrum of values, such as 'completely agree', 'mostly agree', 'mostly disagree', 'completely disagree' when measuring opinion.

The ordinal scale places events in order, but there is no attempt to make the intervals of the scale equal in terms of some rule. Rank orders represent ordinal scales and are frequently used in research relating to qualitative phenomena. For instance, if x student position in his class is 10 and y student position is 40, it cannot be said that y student position is four times as good as that of x.

Ordinal scales only permit the ranking of items from highest to lowest. Ordinal measures have no absolute values, and the real differences between adjacent ranks may not be equal. All that can be said is that one person is higher or lower on the scale than another, but more precise comparisons cannot be made and without our being able to state how much greater or less. The real difference between ranks 1 and 2, for instance, may be more or less than the difference between ranks 5 and 6. Since the numbers of this scale have only a rank meaning, the appropriate measure of central tendency is the median. A percentile or quartile measure is used for measuring dispersion.

Central tendency

The median, i.e. *middle-ranked*, item is allowed as the measure of central tendency; however, the mean (or average) as the measure of central tendency is not allowed. The mode is allowed. There is no absolute zero.

Interval scale

The interval type allows for the *degree of difference* between items, but not the ratio between them. Examples include *temperature scales* with the Celsius scale, which has two defined points (the freezing and boiling point of water at specific conditions) and then separated into 100 intervals, *date*, *location*, and *direction*. Ratios are not meaningful since 20 °C cannot be said to be "twice as hot" as 10 °C, nor can multiplication/division be carried out between any two dates directly. However, *ratios of differences* can be expressed; for

example, one difference can be twice another. Interval type variables are sometimes also called "scaled variables".

Central tendency and statistical dispersion

The mode, median, and arithmetic mean are allowed to measure central tendency of interval variables, while measures of statistical dispersion include range and standard deviation. Since one can only divide by *differences*, one cannot define measures that require some ratios, such as the coefficient of variation.

Ratio scale

The ratio type takes its name from the fact that measurement is the estimation of the ratio between a magnitude of a continuous quantity and a unit of measurement of the same kind. Most measurement in the physical sciences and engineering is done on ratio scales. Examples include mass, length, duration, plane angle, energy and electric charge. In contrast to interval scales, ratios can be compared using division. Very informally, many ratio scales can be described as specifying "how much" of something (i.e. an amount or magnitude).

Central tendency and statistical dispersion

All statistical measures are allowed because all necessary mathematical operations are defined for the ratio scale.

Qualitative properties are properties that are observed and can generally not be measured with a numerical result. They are contrasted to quantitative properties which have numerical characteristics.