

Post-lecture Questions I.2 – Introduction to Psychology Data

Why is it incorrect to say something like: “response time is a valid measure”?

What determines whether a given measure is useful in a given situation?

What are the three main classes of variables? Are all three really different from each other?

Why is “pre-processing” data sometimes required? What are the three basic kinds of pre-processing? In terms of their outputs, what do the three kinds of pre-processing have in common?

What are Descriptive and Inferential Statistics? (Note: make sure that you specify what each makes a statement about, the form of the statement [definitive vs probabilistic], and whether these statements can be wrong.)

Psychologists summarize univariate data in terms of three attributes – what are they? What specific measure do we use for each of the three (in most situations)?

Example multiple-choice questions for this material (taken from Fall, 2010, Exam 1):

1. A single score (for a subject) that is created from a variety of **different** measures is called _____.
 - (A) a combined score
 - (B) a condensed score
 - (C) a summary score
 - (D) a true score
2. Which of the following is **not** part of a complete set of descriptive statistics (as used in psychology)?
 - (A) A numerical measure of the center
 - (B) A numerical measure of the spread
 - (C) A numerical measure of the reliability
 - (D) A name for the shape

A measure does not have validity in general. A measure only has validity with respect to what it is being used to get an estimate of. The measure, itself, is neither valid nor invalid; the use of it for some purpose is what has some amount of validity.

Measures are useful if it provides an estimate of something that is relevant to what you are doing. For example, if the theory being tested makes a prediction for the thing that is estimated by the measure, then it's useful; otherwise, it's irrelevant.

The three main classes of variables are continuous, discrete, and qualitative. The first two are numerical; the third is categorical. The difference between the first two – which is whether all values between two extremes are possible, or only some values are possible – only affects the plotting of the data; it has no real effect on anything else. The difference between the two numerical classes and categorical variables is huge; doing statistics on qualitative data is a total pain (and will not be covered very much in this course).

Pre-processing is sometimes required to get the data into the same format as that for which the theory makes predictions. It is also sometimes used to reduce unreliability. Pre-processing can involve (1) doing nothing, (2) condensing a set of different scores into one score, or (3) summarizing a set of similar scores into one score, usually by finding the mean. In all cases, you almost always end up with a single number or value for each subject; that's what they have in common.

D-stats summarize a set of data. D-stats make definitive statements that are always correct (assuming no calculation errors). I-stats go beyond the sample data and make statements about the population from which the sample was taken and these statements can be wrong (even when all calculations are correct). An example of a D-stat is: "the mean of the sample is 7.89." An example of an I-stat is: "there's a 95% chance that the mean of the entire population is within 1.25 of 7.89."

In most cases, psychologists use measures of (or labels for) the center, spread, and shape to summarize single sets of numbers. We usually use the mean for center, the standard deviation for spread (but variance is also OK), and a name for the shape (e.g., normal, J-shaped, uniform). The second most-popular set of descriptive is median for center, inter-quartile range for spread, and a name for shape.

The correct answer to the first question is B. "Combined score" is not a technical term – it's the general label for both condensed and summary score. A summary score is when the all of the measured being combined are the same, not different. And "true score" refers to something theoretical, unobservable, and unmeasurable; we never can actually calculate a true score.

The correct answer to the second question is C. A complete set of D-stats (in psych) includes a measure of center, a measure of spread, and a name for the shape. Reliability (and unreliability) are attributes of the measure in general; they are not attributes the measure for a particular set of data.