EPSY 5221 Summarizing Educational Data

Measurement presumes that individuals differ in terms of some characteristic – a characteristic we are interested in measuring. We may measure this characteristic because we want to know how much of it a person has or has obtained; we may want to know where a person falls in the population in terms of how much of the characteristic she or he has.

The basis for measurement is that the answers to these questions are unknown, but we assume that individuals differ. We measure primarily to understand these differences (for any purpose).

# Educational Statistics

 We hope to manipulate measurements in a way to summarize some set of measurements taken on our population of interest. We use indicators of central tendency, indicators of variability, and indicators of relationship to accomplish this.

 Moments about the distribution: mean, variance, skewness, kurtosis. Standard normal distribution has a mean = 0, variance = 1, skewness = 0, kurtosis = 0.

# Variability

Variability is an indicator to describe the degree of differences among the measures we have acquired from individuals. Our primary interest is in these differences and variability indicates the extent to which individuals differ.

 Deviations ()

 Problem arises when summing deviations, sum = 0

 Squared Deviations 

 Avoid the sum of deviations problem; other nice properties

 The metric is the original measure squared (not easy to interpret)

 Sum of Squares = *SS* = 

 Indicator of total variation (deviations) in the sample (population)

 Average Sum of Squares = *S*2 = Variance = 

 Standard Deviation (*S*) is the square root of the variance (average sum of squares)

the metric is the same as the metric of the original measure

 Use of *S* implies a normal distribution

See Thorndike Figure 3.3 (p. 83) for number of standard deviations across a normally distributed population
Sometimes we want to compare scores measured on different scales and a convenient indicator is the number of standard deviations each score is away from the mean.

 *z*-score: 

Summary indicator of the deviation – the number of standard deviations a given score is from the mean. It is a standardized score (standardized by the *S*).

A *z* score of 0 is always the mean; a *z* score of +1.0 means that the score is one standard deviation above the mean. This transformation maintains the rank order of scores.

*Relationships*

 Cross Product: 

Covariation: 

 This indicates the degree to which two variables co-vary: that is, do both measures indicate deviation simultaneously – does an individual differ on both measures, in one direction or the other? Covariance is computed by taking the average of the products of deviations on two measures.

 Correlation: = 

 This is a standardized indicator of covariance. Since some variables may be measured on a scale with a wide range, the covariance will also be large. To standardize these effects, the covariance can be standardized by dividing it by the standard deviations of both measures.

 This is similar to standardizing a score (*z*-scores). In fact, if scores are in *z*-score metric, you can take the average products of the *z*-scores to compute the correlation.

 

***A little statistical theory regarding the treatment of items***

***or dichotomously scored variables***

A single item *k* is administered to individual *i* which results in score *Xki*

The mean of an item *k* scored on individuals *i* = 1 to N:

 ; Where *Xk* is a single item = 0 or 1 , 

The variance of an item:

Recall  is the variance of item *k* (any given item), scored on individual *i*

Through some algebraic manipulation, we can derive:



Now we can replace these terms:

 where X = 0 or 1, 

to arrive at:



*p* is the proportion of examinees that answered the item correctly;

(1-*p*) = *q* is the proportion of examinees that answered the item incorrectly.

***A little statistical theory regarding measures of association.***

Pearson’s product-moment correlation coefficient (*r*) is appropriate when variables are continuous and reach an interval or ratio level of measurement. It is a measure of the strength of linear relationship only. When the data are not continuous (dichotomous or ordinal), alternatives are available.

As a simplification, dichotomous variables can be thought of in two ways: (1) truly dichotomous and (2) artificially dichotomous. A truly dichotomous variable has only two values (e.g., yes, no; on, off; male, female). An artificially dichotomous variable is one that has a continuous underlying scale, but has been dichotomized (e.g., tall, short; high scorer, low scorer; stressed, relaxed).

Phi Coefficient ()

 An index of the strength of association between two dichotomous variables;

 This is equivalent to the Pearson correlation coefficient.

Tetrachoric Coefficient (*rtet*)

 An index of strength of association between two artificially dichotomized variables

Point-biserial Coefficient (*rpbis*)

 An index of strength of association between a dichotomous variable and a continuous variable

Biserial Coefficient (*rbis*)

 An index of strength of association between an artificially dichotomized variable and a continuous variable

Spearman Rank-Order Coefficient (*rS*)

 An index of strength of association between two variables measured at the ordinal level of measurement

Kendall’s tau (τ)

 An alternative to Spearman’s rank-order correlation coefficient (more tedious to compute)