Coefficient Alpha

PSY 600K
Dr. Alyssa Gibbons
February 20, 2018

Agenda

- Start talking to your SMEs!
- Coefficient alpha
  - What it is and where it comes from.
  - Is alpha reliability?
  - Some myths and truths about alpha.
- Practical issues in reliability.
  - Factors affecting reliability.
  - Estimating true scores.
  - Using the standard error of measurement.
  - Correcting for unreliability (?)

Internal Consistency

- Split-half reliability is a form of *internal consistency* reliability.
  - Internal consistency = essentially, to what extent do all of these items "go together."
  - More formal definition in a minute.
  - Addressing error (imprecision) due to the fact that items aren't perfect and aren't all the same.
- But split-half reliability has its own issues.
  - In particular… how do you choose the split?

Coefficient Alpha

- Often called Cronbach's alpha… Guttman-Cronbach alpha if you want to be precise.
- Formalizes internal consistency:
  - The proportion of total variance attributable to a common source; aka
  - The (adjusted) ratio of inter-item covariances to total variance.

Alpha and Item Covariances

- We said before that the total test variance = the sum of all of the item variances and covariances (x 2).
- So consider 2 covariance matrices:

$$\begin{matrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{matrix} = \begin{matrix} 2.1 & \mathbf{4.6} & \mathbf{5.1} & \mathbf{5.4} \\ \mathbf{4.6} & 1.9 & \mathbf{4.7} & \mathbf{4.9} \\ \mathbf{5.1} & \mathbf{4.7} & 2.2 & \mathbf{5.3} \\ \mathbf{5.4} & \mathbf{4.9} & \mathbf{5.3} & 1.8 \end{matrix} \quad \text{vs.} \quad \begin{matrix} \mathbf{7.6} & 2.4 & 3.9 & 3.2 \\ 2.4 & \mathbf{8.4} & 2.8 & 3.6 \\ 3.9 & 2.8 & \mathbf{7.2} & 2.1 \\ 3.2 & 3.6 & 2.1 & \mathbf{8.8} \end{matrix}$$

- Which set of items has more **shared** variance?
  - This is how we calculate alpha!

The Alpha Formula

- $\propto = \frac{p}{p-1}\left[1 - \frac{\sum_{i=1}^{p} Var(X_i)}{Var(X)}\right] = \frac{p}{p-1}\left[\frac{\sum_{i \neq j} Cov(X_i, X_j)}{Var(X)}\right]$
- Literally, the ratio of shared variance (covariance) to total variance.
  - Or the average inter-item covariance per unit of variance.
  - With a scaling factor (p/p-1) to keep the values between 0 and 1.
  - $\begin{bmatrix} 7.6 & \mathbf{2.4} & \mathbf{3.9} & \mathbf{3.2} \\ \mathbf{2.4} & 8.4 & \mathbf{2.8} & \mathbf{3.6} \\ \mathbf{3.9} & \mathbf{2.8} & 7.2 & \mathbf{2.1} \\ \mathbf{3.2} & \mathbf{3.6} & \mathbf{2.1} & 8.8 \end{bmatrix} = \frac{4}{3}\begin{bmatrix} \frac{36}{68} \end{bmatrix} = (1.33) * (0.53) = 0.71$
  - Interpreted just like a correlation-based reliability coefficient.

Is Alpha Reliability?

- Yes and no.
- Practically speaking, we rely on alpha so much as a field that it is often used synonymously with reliability.
- But technically…
  - R & M point out that alpha is equal to reliability (that is, σ_T^2/σ_Y^2) **if and only if:**
    - The items are parallel.
    - The errors are uncorrelated.
  - If errors are uncorrelated but the items aren't parallel, alpha **underestimates** reliability.
    - Alpha approaches reliability [illegible]ctor loadings increase.
  - If the errors are correlated, alpha is just incorrect – we don't know how much or in what direction.

Myths & Truths About Alpha

- Alpha does **NOT** tell you whether your items are homogeneous.
  - Homogeneous = measuring the same common factor (unidimensional).
  - Internally consistent = highly interrelated.
  - How could a scale have high interitem covariances but not measure a single common factor?
- Internal consistency is **necessary but not sufficient** for homogeneity.
  - If you want to test homogeneity, you need a factor analysis.

Myths & Truths About Alpha

- Alpha is only **sometimes** a lower bound to reliability.
  - If errors are uncorrelated.
  - Remember that CTT does not **require** uncorrelated errors – we can also model correlated ones.
  - When might you have to worry about correlated errors?
- Alpha **is** equivalent to the average of all possible split-half reliability coefficients.
  - If you take into account item standard deviations.
  - So it is redundant to report both.

Omega: An Alternative to Alpha

- McDonald (1999) proposed a method for estimating reliability more directly based on factor loadings.
  - Coefficient omega: $\omega = \left(\sum \lambda_j\right)^2 / (\sigma_Y^2)$
  - λs = factor loadings.
- This is a much more direct estimate of reliability .
  - Factor loadings allow us to estimate true score variance more precisely than the item covariances because they take into account differing "quality" of items.
  - ω= α when the items are all of the same quality (i.e., true-score equivalent).
  - More on this later. ☺
- Finding ω *does* tell you about homogeneity – because you have to test the factor model in order to calculate ω!

Factors Affecting Reliability

- Of all kinds… not just alpha!
- Reliability is **population dependent**.
  - In a highly homogeneous population, reliability will be lower than in a more diverse population.
  - If there really is little true score variance, your test will have little true score variance!
  - Implication: Reliability is not just a property of a test – it is a property of a test in a particular population.
    - Need to consider whether previous uses of the test have been in the population you are using.

Factors Affecting Reliability

- Reliability estimation does not make sense for speeded (time-limited) tests.
  - Conflates real differences in the constr with difference in processing speed.
- Test length increases reliability.
  - As discussed last time, not in a linear fashion.
  - Even if your items are not unidimensional!

Using Reliability to Estimate True Scores

- Another way to interpret reliability coefficients is as the **correlation between observed score and true score**.
- We can use this in practice to estimate true scores for individuals.
  - $T'_i = \rho_X(X_i - \bar{X}) + \bar{X}$
  - $.80(95 - 80) + 80 = 12 + 80 = 92$
- Is this useful?
  - Rank order will not change.
  - Variance of true score estimates < variance of observed scores.
  - Scores will regress toward the mean.

Standard Error of Measurement

- Remember our description of true scores as the mean of a person's propensity distribution over a very large number of identical tests?
- The SEM estimates the standard deviation of that distribution.
  - How far off, on average, are observed scores from true scores?
- SEM = $\sigma_X\sqrt{1 - \rho_X}$
- We can use the SEM to calculate confidence intervals around an observed score:
  - $= 10\sqrt{(1 - .80)} = 10 * .447 = 4.47$
  - CI for an observed score of 95: +/- 2 SEM:
  - 86.06 – 103.94
- This tells us more about how precise our estimated true scores are.
  - In this case, not very!

Correcting for Unreliability

- Reliability is an index of how much error we have in our measures.
  - Error = noise.
- This noise distorts whatever correlations we might find between our measure and others.
  - In a specific way – it **attenuates** or reduces them.
  - $r_{xy_{obs}} = r_{xy_{true}}(\sqrt{r_{xx}r_{yy}})$
  - $r_{xy_{obs}} = .50(\sqrt{.80 * .85}) = .50(\sqrt{.68}) = .50 * .82 = .41$
- We can reverse this formula to estimate true correlations from the observed correlation and the reliabilities of the measures.
- Should we?
  - What do Schmidt & Hunter (1986) have to say?

Questions?

For next time:

The Common Factor Model

Read: DeVellis pp. 115-125 AND R & M 3.1 – 3.4