Psychology 405: Psychometric Theory: Review of part I + Reliability problem set (with answers)

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

May, 2024

Outline

Types of data: Subjects, Objects, Comparisons

1. The traditional goal in psychological measurement is to assign values to Subjects and sometimes Objects (items)
2. Scaling objects or scaling people by direct or indirect comparisons
3. Focus on just objects: pool values across subjects in order to compare objects
   - Similarities of objects (distance less than some threshold) $|o_i - o_j| < \delta$ (Problem of non-symmetry of similiarity)
   - Ordering of objects $o_i < o_j$
4. Focus on Individuals
   - Similarities of people $|s_i - s_j| < \delta$
   - Ordering of people through tournaments $s_i < s_j$
5. Use objects to order people (conventional measurement)
   - Preferential choice and unfolding($|s_i - o_j| < |s_k - o_m|$) (single peaked function)
   - Monotonic ordering $s_i > o_j$ and we order the objects to order the people

Relating latent scores to observed scores

1. We observe choice (preferences) and performance (ordering)
2. We infer latent values for items and then for people
3. Our inferences reflect the type of assumptions we make about our data
4. Most data are at least (at most?) ordinal
5. Inferences on differences of differences (e.g., interactions) require interval data
6. If our inferences depend upon a particular metric, and change if we rescale/transform the data, perhaps our inferences are incorrect.

Correlation: a most useful statistic

1. Correlation as the (geometric) average regression
2. Correlation as the covariance of standardized scores
3. Correlation as a universal effect size
4. Correlation as the cosine of the angle of two vectors
5. Correlation as predictability of standard scores

Eigen values, Eigen Vectors, Components, and Factors

1. Eigen value decomposition: $\boldsymbol{R} = \boldsymbol{X}\lambda\boldsymbol{X}'$
2. Principal components (aka the square root of a matrix) : $\boldsymbol{R} = \boldsymbol{C}\boldsymbol{C}'$ where $\boldsymbol{C} = \boldsymbol{X}\sqrt{\lambda}$
3. Factor model explains the common part of the correlation matrix but not the unique part: $\boldsymbol{R} = \boldsymbol{F}\boldsymbol{F}' + \boldsymbol{U}^2$
4. Both models account for item covariances (structure)
5. Components are linear sums of variables (and thus are observed)
6. Items are (unobserveable sums of factors) and are thus estimated.

More on factors

1. Rotations and transformations to "simple structure" doesn't change the adequacy of the model
2. Rotations maximize some criterion (but which one?)
3. Oblique solutions lead to a difference of pattern versus structure. $\boldsymbol{R} = \boldsymbol{F}\phi\boldsymbol{F'} + \boldsymbol{U^2}$
   - Pattern $= \boldsymbol{F}$
   - Structure $= \phi\boldsymbol{F'}$
   - If reporting oblique solution, you must report ϕ
4. Higher order solutions (2nd and sometime 3rd stratum factors)
5. Simple structure versus bifactor or more complex solutions

Even more on factors

1. If we know the number of factors, we can find the communalities
2. If we know the communalities, we can find the factors
3. How many factors? No best rule
4. Factor extraction to best fit a certain criterion
5. Maximum likelihood finds the model that maximizes the probability of the data if the model is correct.
6. But what if is incorrect?

Consider a 10 item test with two subtests, A and B. The average correlation in both subtests is .3, and the average item correlation between the two subtests is .2.

Table: Item correlations $\implies$ Scale covariances and correlations

	V1a	V2a	V3a	V4a	V5a	V1b	V2b	V3b	V4b	V5b
V1a	1.0	.3	.3	.3	.3	.2	.2	.2	.2	.2
V2a	.3	1.0	.3	.3	.3	.2	.2	.2	.2	.2
V3a	.3	.3	1.0	.3	.3	.2	.2	.2	.2	.2
V4a	.3	.3	.3	1.0	.3	.2	.2	.2	.2	.2
V5a	.3	.3	.3	.3	1.0	.2	.2	.2	.2	.2
V1b	.2	.2	.2	.2	.2	1.0	.3	.3	.3	.3
V2b	.2	.2	.2	.2	.2	.3	1.0	.3	.3	.3
V3b	.2	.2	.2	.2	.2	.3	.3	1.0	.3	.3
V4b	.2	.2	.2	.2	.2	.3	.3	.3	1.0	.3
V5b	.2	.2	.2	.2	.2	.3	.3	.3	.3	1.0

Various questions about these tests

1. What is the variance of subtest A?
2. What is the covariance of A and B?
3. What is the correlation between A and B?
4. What is the variance of total test?
5. What is coefficient α for subtest A?
6. What is coefficient α for the entire test?
7. What is the worst split half reliability of the test? (Assuming that A and B represent the worst splits).
8. What is the correlation of an item in test A with test A?
9. What is the correlation of an item in test A with the total test?
10. What is ω_h for the entire test?
11. What is the correlation of test A with test B *correcting for reliability*

Te average correlation in test A = .2, in test B = .15, and the average inter-item correlation between the two tests is .1

12. What is the variance of subtest A?
13. What is the covariance of A and B?
14. What is the correlation between A and B?
15. What is the variance of total test?
16. What is coefficient α for subtest A?
17. What is coefficient α for the entire test?
18. What is the worst split half reliability of the test? (Assuming that A and B represent the worst splits).
19. What is the correlation of an item in test A with test A?
20. What is the correlation of an item in test A with the total test?
21. What is ω_h for the entire test?
22. What is the correlation of test A with test B *correcting for reliability*

Think about the items and the tests

Consider a 10 item test with two subtests, A and B. The average correlation in both subtests is .3, and the average item correlation between the two subtests is .2.

Table: Item correlations $\Longrightarrow$ Scale covariances and correlations

	V1a	V2a	V3a	V4a	V5a	V1b	V2b	V3b	V4b	V5b
V1a	1.0	.3	.3	.3	.3	.2	.2	.2	.2	.2
V2a	.3	1.0	.3	.3	.3	.2	.2	.2	.2	.2
V3a	.3	.3	1.0	.3	.3	.2	.2	.2	.2	.2
V4a	.3	.3	.3	1.0	.3	.2	.2	.2	.2	.2
V5a	.3	.3	.3	.3	1.0	.2	.2	.2	.2	.2
V1b	.2	.2	.2	.2	.2	1.0	.3	.3	.3	.3
V2b	.2	.2	.2	.2	.2	.3	1.0	.3	.3	.3
V3b	.2	.2	.2	.2	.2	.3	.3	1.0	.3	.3
V4b	.2	.2	.2	.2	.2	.3	.3	.3	1.0	.3
V5b	.2	.2	.2	.2	.2	.3	.3	.3	.3	1.0

Find the variances and covariances

1. The variance of a composite is the sum of the elements in the composite;
2. $V_A = \Sigma\Sigma(A_{ij}) = \mathbf{1A1'}$
3. $V_A = 5 + 20 * .3 = 11 = V_B$
4. $C_{AB} = 25 * .2 = 5$
5. $r_{AB} = \frac{C_{AB}}{\sqrt{V_A * V_B}} = \frac{5}{\sqrt{11*11}} = .45$
6. $\alpha_A = \frac{V_A - \Sigma v_i}{V_A} * \frac{k_A}{k_A - 1} = \frac{11-5}{11} * \frac{5}{4} = .68$
7. Or, $\alpha = \frac{n\bar{r}}{1+(n-1)\bar{r}} = \frac{5*.3}{1+4*.3} = \frac{1.5}{2.2} = .68$

Covariance

	A	B
A	11	5.0
B	5.0	11

Correlation

A	B.
1.0	.45
.45	1.0

Find the variances and covariances for the total test

Covariance

	A	B
A	11	5.0
B	5.0	11

Correlation

A	B.
1.0	.45
.45	1.0

Total Variance = 32 Sum of Item variances = 10.

$\alpha_{AB} = .76 \quad \alpha_A = .68$

Average r $= \bar{r} = (20 * .3 + 20 * .3 + 25 * .2 + 25 * .2)/90 = .24$

$$\alpha = \frac{V_t - \Sigma(v_i)}{V_t} \frac{k}{k-1} = \frac{32 - 10}{32} \frac{10}{9} = \frac{k\bar{r}}{1 + (k-1)\bar{r}} = \frac{10 * .244}{1 + 9 * .244} = .7$$

Table: Item correlations $\Longrightarrow$ Scale covariances and correlations

	Covariance		Correlation	
	A	B	A	B.
A	11	5.0	1.0	.45
B	5.0	11	.45	1.0

$$\alpha_{1-5} = \alpha = \frac{V_t - \Sigma(v_i)}{V_t} \frac{k}{k-1} = \frac{11-5}{11}\frac{5}{4} = \frac{k\bar{r}}{1+(k-1)\bar{r}} = \frac{5 * .3}{1 + 4 * .3} = .68$$

Worst split $= \frac{2r_{AB}}{1+r_{AB}} = \frac{2*.45}{1+.45} = .62$.
Compare with $\alpha_{1-10} = \frac{32-10}{32} * \frac{10}{9} = .76$ which is greater.

Because $\alpha > \beta$ we conclude the test is not homogeneous.

Item whole correlations

1. We can find item by total correlations by adding up the separate elements.
2. $r_{1,A} = \frac{cov_{1,1..5}}{\sqrt{V_1 V_A}} = \frac{2.2}{\sqrt{1*11}} = .66$
3. $r_{1,(AB)} = \frac{cov_{1,1..10}}{\sqrt{V_1 V_{AB}}} = \frac{3.2}{\sqrt{1*32}} = .56$
4. Correct for overlap: (replace v_i with $\bar{r}$)
   $r_{1,A} = \frac{cov*_{1,1..5}}{\sqrt{V_1 V_A}} = \frac{1.5}{\sqrt{1*11}} = .45$
5. $r_{1,(AB)} = \frac{cov*_{1,1..10}}{\sqrt{V_1 V_{AB}}} = \frac{2.5}{\sqrt{1*32}} = .44$
6. Adjust those for reliability: $\frac{.45}{\sqrt{.68}} = .55$ and $\frac{.44}{\sqrt{.68}} = .54$

Finding ω_h from the structure

1. Normally we want to find the factor loadings of g to ask how much is recreated by the g factor
2. But, we can do this by inspection: Items share a g factor and group factors.
3. Average variance due to general is .2, due to group is .1
4. General factor variance $= 100 * .2 = 20$
5. $\omega_h = \frac{\sigma_g^2}{\sigma_t^2} = \frac{20}{32} = .625$
6. In this particular case, this is the same as the worst split half

Now, we will do this again, but with the average correlation in test A = .2, in test B = .15, and the average inter-item correlation between the two tests is .1

Table: Item correlations $\Longrightarrow$ Scale covariances and correlations

	V1a	V2a	V3a	V4a	V5a	V1b	V2b	V3b	V4b	V5b
V1a	1.0	.2	.2	.2	.2	.1	.1	.1	.1	.1
V2a	.2	1.0	.2	.2	.2	.1	.1	.1	.1	.1
V3a	.2	.2	1.0	.2	.2	.1	.1	.1	.1	.1
V4a	.2	.2	.2	1.0	.2	.1	.1	.1	.1	.1
V5a	.2	.2	.2	.2	1.0	.1	.1	.1	.1	.1
V1b	.1	.1	.1	.1	.1	1.0	.15	.15	.15	.15
V2b	.1	.1	.1	.1	.1	.15	1.0	.15	.15	.15
V3b	.1	.1	.1	.1	.1	.15	.15	1.0	.15	.15
V4b	.1	.1	.1	.1	.1	.15	.15	.15	1.0	.15
V5b	.1	.1	.1	.1	.1	.15	.15	.15	.15	1.0

	Covariance	
	A	B
A	9	2.5
B	2.5	8

Correlation	
A	B.
1.0	.29
.29	1.0

Total Variance = 22 Sum of Item variances = 10.
$\alpha_{AB} = .61 \quad \alpha_A = .55$ Average r =
$\bar{r} = (20 * .2 + 20 * .15 + 25 * .1 + 25 * .1)/90 = .133$

$$\alpha = \frac{V_t - \Sigma(v_i)}{V_t}\frac{k}{k-1} = \frac{22-10}{22}\frac{10}{9} = \frac{k\bar{r}}{1+(k-1)\bar{r}} = \frac{10 * .133}{1 + 9 * .133} = .61$$

Worst split $= \frac{2r_{AB}}{1+r_{AB}} = \frac{2*.29}{1+..29} = .45$. Compare with $alpha_{1-10} = \frac{22-10}{22} * \frac{10}{9} = .61$ which is greater.

Finding ω_h from the structure

1. Normally we want to find the factor loadings of g to ask how much is recreated by the g factor
2. But, we can do this by inspection: Items share a g factor and group factors
3. General = .1 and group = .1 for the first group and .05 for the second group.
4. General factor variance $= 100 * .1 = 10$
5. $\omega_h = \frac{\sigma_g^2}{\sigma_t^2} = \frac{10}{22} = .45$
6. In this particular case, this is the same as the worst split half