Module 2: Intro to G Theory

Main topics

- Why G theory?
- Important concepts
  - Universe score and multifaceted measurement error
  - Relative and absolute decisions
  - G coefficients
  - G and D studies

Universe score and measurement error

- A measure (or score) for a particular examinee is a sample from a universe of admissible observations.
  - Universe score is the expected value of the admissible observations
  - Population refers to the collection of examinees
  - Universe refers to the collection of possible measurement conditions, such as rater, form, item…

About facet

- Refers to a set of similar conditions of measurement. That is, a facet can be a certain source of measurement error in generalization of score interpretations, such as rater, occasion, item, task, …
- Can be random or fixed
- Can be nested or crossed with other facets depending on the designs used

An example

- 8 multiple-choice items were given to 20 students.
  - What is the facet involved?
  - What is the design in this G study?

The example: P X I design

- Eight multiple-choice items were given to 20 students.

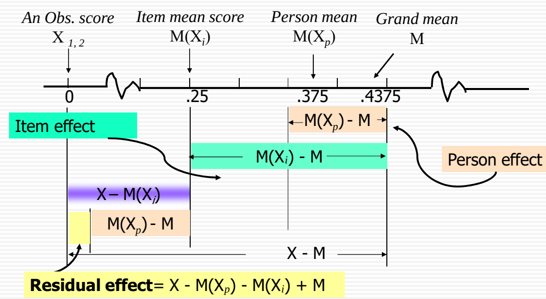
	Item					Person
Person	1	2	3	…	8	means
1	1	0	0	…	1	.375
2	0	0	0	…	1	.375
…	…					…
20	0	0		…	0	.125
Item means	.55	.25			.55	.4375

Variability of observed scores

Conceptual explanation:

What factors affect the variance of students' scores?

Variability of observed scores: When the observed score is 0



Representation of the difference

- So the difference between $X_{1,\,2}$ and the grand mean can be represented as:

X_{pi} – grand mean =

(Mean of X_p – grand mean)	[Person effect]
+ (Mean of X_i - grand mean)	[Item effect]
+ (X_{pi} – Mean of X_p – Mean of X_i + grand mean)	[Residual effect]

Consider the universe and population

	Item 1	2	3	…	k	Person mean
1	$X_{1,1}$	$X_{1,2}$	$X_{1,3}$	…	$X_{1,k}$	$M(x_{1.})$
2	$X_{2,1}$	$X_{2,2}$	$X_{2,3}$	…	$X_{2,k}$	$M(x_{2.})$
…	…					…
N	$X_{n,1}$	$X_{n,2}$	$X_{n,3}$	…	$X_{n,k}$	$M(x_{n.})$
Item mean	$M(x_{.1})$	$M(x_{.2})$	$M(x_{.3})$	…	$M(x_{.k})$	$M(x_{..})$

μ_p

μ_i

Constant μ

Decompose the observed score and its variance in a P x I design

- Observed score

X_{pi} = μ	[grand mean]
+ μ_p μ	[person effect]
+ μ_i μ	[item effect]
+ X_{pi} - μ_i + μ_p μ	[residual effect]

- Variance

$V(X_{pi}) = V(\mu_p) + V(\mu_i) + V(\mu_{pi,\,e})$ or

$\sigma^2(X_{pi}) = \sigma_p^2 + \sigma_i^2 + \sigma_{pi,\,e}^2$

A hypothetical result of the P x I design

Source of Variability	Notation of Variability	Estimate
Person (*p*)	σ_p^2	1.2
Item (*i*)	σ_i^2	.3
Residual (*pi*, e)	$\sigma_{pi,e}^2$	.5

Q: What is the variance due to the individual differences?
Variance due to measurement errors?

Relative and Absolute Decisions

Different decision purposes

- Scores can be used to:
  - Rank/order examinees by comparing examinees' performances
  - Index an examinee's absolute level of knowledge, skill, or attitude without regard to how well or poorly his or her peers performed

An example

- Consider a scenario: A test including 8 items was given to 100 job applicants.
  - The applicants who obtained a total score of 5 and more will be offered positions.
  - The 5 top-scoring applications will be offered positions.

Relative and absolute measurement errors

- Relative and absolute errors are associated with different types of score interpretations.
  - Absolute measurement error includes *all* the error terms
  - Relative measurement error includes the error terms *involving p*.

Relative and absolute G coefficients

- G-coefficients
  - Relative G-coefficients = $\frac{V(P)}{V(P) + V(\text{Rel. Error})}$
  - Absolute G-coefficients = $\frac{V(P)}{V(P) + V(\text{Abs. Error})}$

Calculating G Coefficients

Estimate variance components (I)

- Take the p x i design (random effect) as an example

Effect	SS	df	MS	Expected MS	Estimated G Study Variance Components
Person(*p*)	SS_p	n_p-1	$MS_p = SS_p/df_p$	$E(MS_p) = \sigma^2_{pi,e} + n_i\sigma^2_p$	$\sigma^2_p = [MS_p - MS_{pi,e}]/n_i$
Item(*i*)	SS_i	n_i-1	$MS_i = SS_i/df_i$	$E(MS_i) = \sigma^2_{pi,e} + n_p\sigma^2_i$	$\sigma^2_i = [MS_i - MS_{pi,e}]/n_p$
pi, e	$SS_{pi,e}$	$(n_p-1)(n_i-1)$	$MS_{pi,e} = SS_{pi,e}/df_{pi}$	$E(MS_{pi,e}) = \sigma^2_{pi,e}$	$\sigma^2_{pi,e} = MS_{pi,e}$

Estimate variance components (II)

- Eight test items were given to 20 students. The hypothetical results were presented below.

Effect	SS	df	MS	Estimated G Study Variance Components
Person(p)	SS_p	19	.6839	
Item(I)	SS_i	7	.3752	
pi, e	$SS_{pi,e}$	133	.2103	

- What is the relative error? Absolute error? Relative G coefficient? Absolute G coefficient?

Interpret the results (I)

The results:

- provide estimated variances of the observed scores for a <u>single</u> person on a <u>single</u> item.
  - Square root of an estimated variance (e.g., σ_p is the sqrt of σ_p^2) is the SD for the measurements within a given facet.
- reveal the magnitude of error in generalizing from a person's score on a single item to his/her universe score.

Interpret the results (II)

What to read in order to determine the reliability of a measurement procedure?

- the estimated variances to find out whether errors (abs. or rel.) are small
- the percentages of the components using the estimated variances.
- the G coefficients to determine the reliability in generalizing from a person's score on a single item to his/her universe score.
- Square roots of the estimated variances to determine whether the variation is small for the given scale of the score.

G and D Studies

An unanswered question

- In the p x i design (random effect), we now know the reliability when one item was used. We can also figure out the reliability using 8 items.
- But what about the reliability of other tests (including 10 or 20 items)?

Research purposes

- G study is intended to determine the reliability of a particular design.
- D study is intended to evaluate the effectiveness of alternative designs for minimizing error and maximizing reliability.
  - For example, the reliability when more or less items are included in a test, or the reliability when more or less raters are participated.

Estimate variance components in D studies

- In the p x i design (random effect)

Effect	Estimated Variance Components for 1 item	Proposed D Study Variance Components
Person(p)	$\sigma^2_p = [MS_p - MS_{pi,e}]/n_i$	σ^2_p
Item(I)	$\sigma^2_i = [MS_i - MS_{pi,e}]/n_p$	σ^2_i/n'_i
pi, e	$\sigma^2_{pi,e} = MS_{pi,e}$	$\sigma^2_{pi,e}/n'_i$

They are the estimated variances of the mean score across the items for a single individual.