An Overview of Generalizability Theory

1

Section Learning Objectives

1 An Overview of Generalizability Theory

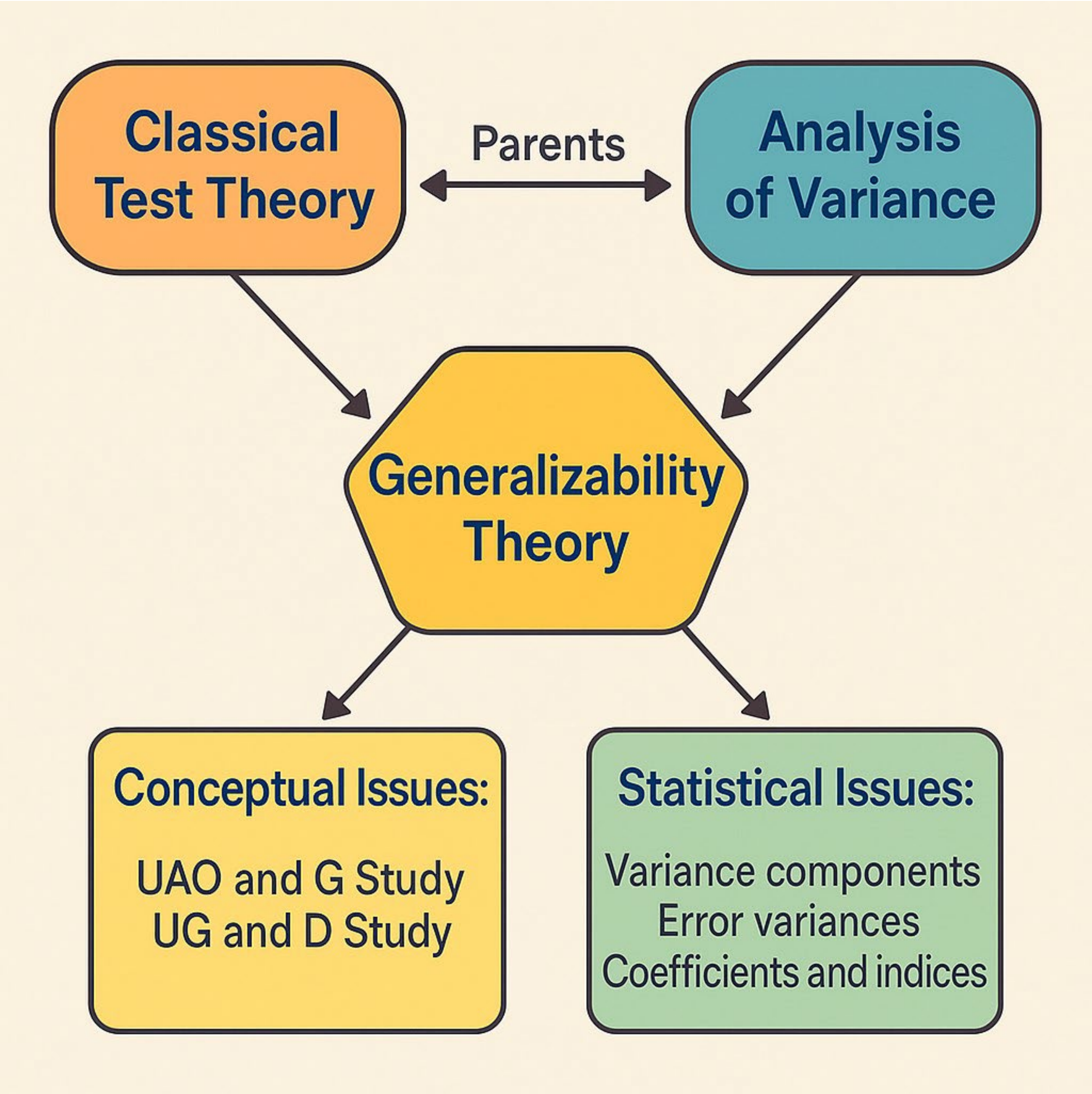
- Describe the fundamental differences between classical test theory and generalizability theory
- Explain key concepts of generalizability theory, including universes of admissible observations and universes of generalization
- Demonstrate how to design and conduct a generalizability (G) study and interpret its variance components
- Apply G study variance components in Decision (D) studies to obtain D-study outcomes

Classical Test Theory and Generalizability Theory

- Classical Test Theory (CTT): $X = T + E$
  - ✓Single undifferentiated error term
  - ✓Classically parallel forms

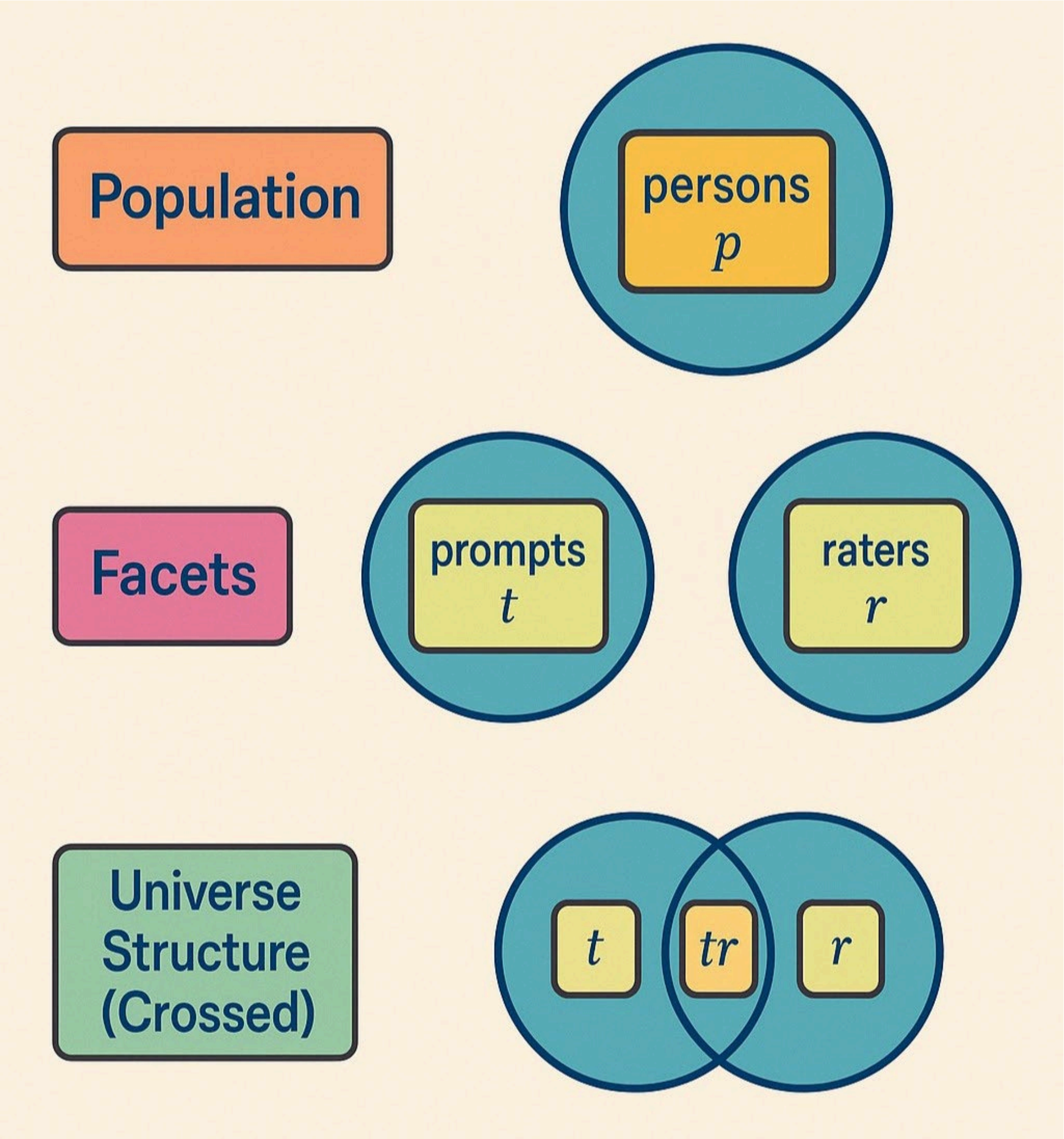
- Generalizability Theory (GT): $X = T + (E_1 + E_2 + \cdots)$
  - ✓Differentiated error terms
  - ✓Randomly parallel forms

Framework of Generalizability Theory

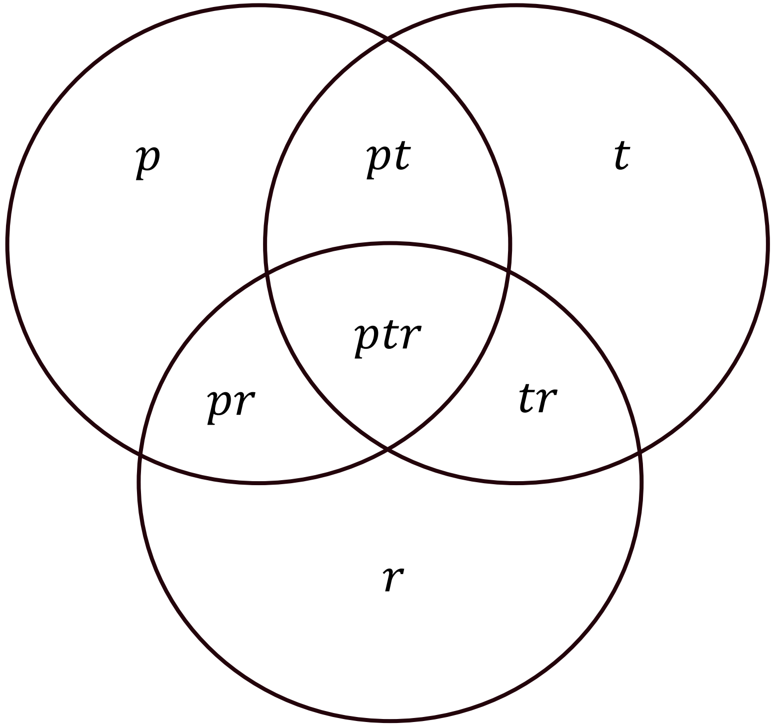


Universe of Admissible Observations (UAO)

- An investigator aims to develop a measurement procedure for a writing assessment targeting a population of persons (p), the objects of measurement.
- The measurement facets may include writing prompts (t) and raters (r).
- In the universe of admissible observations (UAO), both facets are considered infinite.
- The two facets are crossed, denoted $t \times r$.



Generalizability (G) Study $p \times t \times r$ Design



Effect	Variance Component
p	.25
t	.06
r	.02
pt	.15
pr	.04
tr	.00
ptr	.12

- G study sample sizes: n_p=100; $n_t = 4$; $n_r = 2$
- Linear model: $X_{ptr} = \mu + \nu_p + \nu_t + \nu_r + \nu_{pt} + \nu_{pr} + \nu_{tr} + \nu_{ptr}$
- Variance components: $\sigma^2(X_{ptr}) = \sigma^2(p) + \sigma^2(t) + \sigma^2(r) + \sigma^2(pt) + \sigma^2(pr) + \sigma^2(tr) + \sigma^2(ptr)$

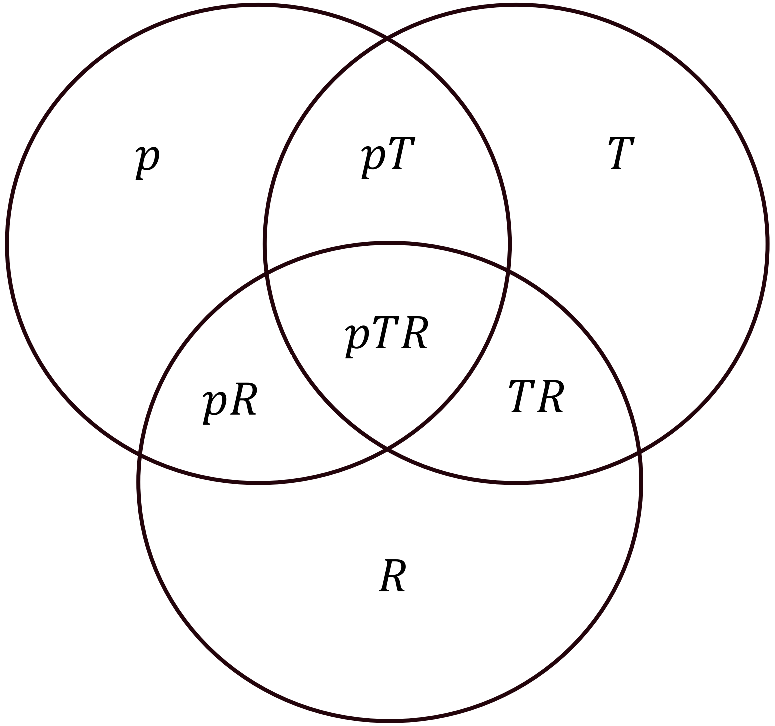
Universe of Generalization and D Studies

- The purpose of a Decision (D) study is to inform substantive decisions about measurement procedures for the objects of measurement.
- Decisions about the objects of measurement (e.g., persons) are based on their ***mean*** scores across the measurement conditions with n'_t and n'_r.
- Variance components for a D study with sample sizes n'_t and n'_r are obtained from the G study variance components.

Universe of Generalization and D Studies

- A universe of generalization (UG) is the universe of ***randomly parallel*** measurement procedures to which an investigator wants to generalize based on the results of a particular D study.
- The design structure of a D study may differ from that of the corresponding G study.

D Study $p \times T \times R$ Design



$n'_t = 3, n'_r = 2$	
Effect	**Variance Component**
p	.25
T	.02
R	.01
pT	.05
pR	.02
TR	.00
pTR	.02

- Linear model: $\bar{X}_p = X_{pTR} = \mu + \nu_p + \nu_T + \nu_R + \nu_{pT} + \nu_{pR} + \nu_{TR} + \nu_{pTR}$
- $\sigma^2(T) = \frac{\sigma^2(t)}{n'_t}$; $\sigma^2(R) = \frac{\sigma^2(r)}{n'_r}$; $\sigma^2(pT) = \frac{\sigma^2(pt)}{n'_t}$; $\sigma^2(pR) = \frac{\sigma^2(pr)}{n'_r}$; $\sigma^2(TR) = \frac{\sigma^2(tr)}{n'_t n'_r}$; $\sigma^2(pTR) = \frac{\sigma^2(ptr)}{n'_t n'_r}$

D Study $p \times T \times R$ Design: Error Variances and Coefficients

- Universe score variance: $\sigma^2(\tau) = \sigma^2(p)$
- Absolute error variance: $\sigma^2(\Delta) = \sigma^2(T) + \sigma^2(R) + \sigma^2(pT) + \sigma^2(pR) + \sigma^2(TR) + \sigma^2(pTR)$
- Relative error variance: $\sigma^2(\delta) = \sigma^2(pT) + \sigma^2(pR) + \sigma^2(pTR)$
- Generalizability coefficient:

$$E\rho^2 = \frac{\sigma^2(\tau)}{\sigma^2(\tau) + \sigma^2(\delta)}$$

- Index of dependability:

$$\Phi = \frac{\sigma^2(\tau)}{\sigma^2(\tau) + \sigma^2(\Delta)}$$

D Study $p \times T \times R$ Design: Computational Example

$n'_t = 3$, $n'_r = 2$	
Effect	**Variance Component**
p	**.25**
T	**.02**
R	**.01**
pT	**.05**
pR	**.02**
TR	**.00**
pTR	**.02**

- Universe score variance: $\sigma^2(\tau) = \sigma^2(p) = .25$
- Absolute error variance: $\sigma^2(\Delta) = .02 + .01 + .05 + .02 + .00 + .02 = .12$
- Relative error variance: $\sigma^2(\delta) = .05 + .02 + .02 = .09$
- Generalizability coefficient:

$$E\rho^2 = \frac{.25}{.25 + .09} = .74$$

- Index of dependability:

$$\Phi = \frac{.25}{.25 + .12} = .68$$

Summary

- GT explicitly models multiple sources of random error.
- GT involves two main steps:
  - ✓UAO and G studies
  - ✓UG and D studies
- Primary outcomes from G studies:
  - ✓Variance components
- Primary outcomes from D studies:
  - ✓D study variance components
  - ✓Universe score variance
  - ✓Error variances (absolute vs. relative)
  - ✓Reliability-like coefficients (generalizability coefficient vs. index of dependability)

Single-Facet Designs

2

Section Learning Objectives

2 Single-Facet Designs

- Differentiate between crossed and nested designs
- Describe the major purposes of G and D studies
- Explain the difference between relative and absolute errors
- Identify two reliability-like coefficients within Generalizability theory

$p \times i$ Design

- $p \times i$ design (p crossed i)
  - ✓A random sample of n_p persons is selected from the population
  - ✓An independent random sample of n_i items is selected from the universe
  - ✓Each of the n_p persons took each of the n_i items, and the responses (X_{pi}) are obtained

- Some terminology
  - ✓Population: the object of measurement
  - ✓Universe: measurement conditions

$p \times i$ Design

- Universe of admissible observations
  - ✓The complete set of all possible observations or measurement conditions that a researcher considers conceptually acceptable or relevant for measuring a particular construct
  - ✓Under the $p \times i$ design, the universe of admissible observations includes an item facet which consists of all possible items that could have been administered

$p \times i$ Design

	Item1	item2	…	Item k
Person1	5	3		7
Person2	1	2		1
Person3	9	9		8
Person4	4	6		7
Person5	2	3		2
…				
Person g	2	3		5
Mean μ_i	**4.4**	**4.6**		**5.0**

Mean (μ_p)
4.5
2.1
8.7
5.8
2.1
3.7
Mean $\mu = 4.6$

$$X_{p=4,i=1} = \mu + (\mu_p - \mu) + (\mu_i - \mu) + (X_{pi} - \mu_p - \mu_i + \mu)$$

Person effect, ν_p Item effect, ν_i Residual effect, ν_{pi}

Note. k often approaches ∞ unless specified otherwise (e.g., mixed models)

G Study Variance Components for $p \times i$ Design

- The variance component for persons is

$$\sigma^2(p) = E_p\left(\mu_p - E_p\mu_p\right)^2$$
$$= E\nu_p^2$$

- The total score variance is decomposed as

$$\sigma^2\left(X_{pi}\right) = \sigma^2(p) + \sigma^2(i) + \sigma^2(pi)$$

Estimating Variance Components

- Several estimation methods have been developed including the Expected Mean Squares (EMS) approach, Bayesian estimation, and Maximum Likelihood (ML) estimation
- Those approaches usually provide very comparable results unless negative variance components are present
- By definition, variance components are nonnegative
- However, estimated variance components are subject to sampling variability.
- Increasing sample size may resolve the issue
- Often set to zero in subsequent D studies

Estimating Variance Components (EMS)

- The ANOVA procedures for estimating variance components

$$\sum_p \sum_i (X_{pi} - \bar{X})^2$$
$$= n_i \sum_p (X_p - \bar{X})^2 + n_p \sum_i (X_i - \bar{X})^2 + \sum_p \sum_i (X_{pi} - X_p - X_i + \bar{X})^2$$
$$= SS(p) + SS(i) + SS(pi)$$

- Mean squares are:

$$MS(p) = \frac{SS(p)}{n_p - 1}; MS(i) = \frac{SS(i)}{n_i - 1}$$
$$MS(pi) = \frac{SS(pi)}{(n_p - 1)(n_i - 1)}$$

Estimating Variance Components (EMS)

- Expected values of mean squares are

$$EMS(p) = \sigma^2(pi) + n_i\sigma^2(p)$$
$$EMS(i) = \sigma^2(pi) + n_p\sigma^2(i)$$
$$EMS(pi) = \sigma^2(pi)$$

- Solving these questions with mean squares in place of EMS,

$$\hat{\sigma}^2(p) = \frac{MS(p) - MS(pi)}{n_i}$$
$$\hat{\sigma}^2(i) = \frac{MS(i) - MS(pi)}{n_p}$$
$$\hat{\sigma}^2(pi) = MS(pi)$$

D Studies for $p \times I$ Design

- Decisions based on n_i' items
- Universe of generalization
  - Over all items in the infinite universe
  - Replications of a measurement procedure
- Uppercase letters (I) to highlight that decisions are made for average scores (not a single score)

D Studies for $p \times I$ Design

- Linear model for X_{pI}, an examinee's average score over n_i' items

$$X_{pI} = \bar{X} = \mu + \nu_p + \nu_I + \nu_{pI}$$

- Universe score (true score analogue under CTT)

$$\mu_p \equiv E_I X_{pI}$$

- Universe score variance

$$\sigma^2(\tau) = \sigma^2(p) = E_p(\mu_p - \mu)^2$$

D Studies for $p \times I$ Design

- The sample size for the object of measurement, n_p, does not apply
- Other D study variance components are calculated as:

$$\sigma^2(I) = \frac{\sigma^2(i)}{n_i'}$$

$$\sigma^2(pI) = \frac{\sigma^2(pi)}{n_i'}$$

Error Variances

- **Absolute error** reflects all sources of measurement errors (defined by the D study design) that affect the absolute level of a score, often used when making criterion-referenced interpretations of scores

$$\Delta_{pI} = X_{pI} - \mu_p = \nu_I + \nu_{pI}$$

$$\sigma^2(\Delta) = \sigma^2(I) + \sigma^2(pI) = \frac{\sigma^2(i)}{n_i'} + \frac{\sigma^2(pi)}{n_i'}$$

- **Relative error** is concerned with sources of measurement errors that impact the ranking or relative standing of individuals, used when making norm-referenced interpretations

$$\delta_{pI} = \left(X_{pI} - \mu_I\right) - \left(\mu_p - \mu\right) = \nu_{pI}$$

$$\sigma^2(\delta) = \sigma^2(pI) = \frac{\sigma^2(pi)}{n_i'}$$

Error Variances

- $\sigma^2(\Delta) \geq \sigma^2(\delta)$
- Under the classically parallel form assumption, no formal distinction between $\sigma^2(\Delta)$ and $\sigma^2(\delta)$

Reliability-like Coefficients

- Generalizability coefficient

$$E\rho^2 = \frac{\sigma^2(p)}{\sigma^2(p) + \sigma^2(\delta)}$$

  - Expected squared correlation between observed and universe scores
  - Expected squared correlation between pairs of randomly parallel forms
  - $E\rho^2$ for a $p \times I$ design is identical to Cronbach's (1951) alpha, KR-20

- Phi coefficient or index of dependability

$$\Phi = \frac{\sigma^2(p)}{\sigma^2(p) + \sigma^2(\Delta)}$$

- $E\rho^2 \geq \Phi$

$i:p$ Design

- $i:p$ design
  - ✓A random sample of n_p persons is selected from the population
  - ✓Each of the n_p persons is administered a different sample of the n_i items, with all items sampled from the same universe

- Linear model is

$$X_{pi} = \mu + \nu_p + \nu_{i:p}$$

$I\!:\!p$ Design

- Linear model is

$$X_{pI} = \mu + \nu_p + \nu_{I:p}$$

- The variance component for the $\nu_{I:p}$ effect is

$$\sigma^2(I\!:\!p) = \frac{\sigma^2(i\!:\!p)}{n_i'}$$

- Absolute and relative errors:

$$\Delta = \delta = \nu_{I:p}$$

$$\sigma^2(\Delta) = \sigma^2(\delta) = \frac{\sigma^2(i\!:\!p)}{n_i'}$$

$$= \frac{\sigma^2(i) + \sigma^2(pi)}{n_i'}$$

$p \times I$ and $I\!:\!p$ Designs

- The crossed design, $p \times I$, produces a relative error variance less than or equal to the one under the nested design, $I\!:\!p$
- The two designs yield the same absolute error variance
- If the universe of generalization has a crossed design, but data are collected under the nested design, this may lead to biased estimates of relative error variance and the generalizability coefficient
- Sections 3 and 4 will provide more detailed information for multi-facet designs (with more than one facet) in terms of G studies and D studies, respectively

References

- Brennan, R. L. (2001). *Generalizability theory*. New York: Springer-Verlag.
- Lord, F. M., and Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Generalizability (G) Studies for Multi-Facet Designs

3

Section Learning Objectives

3 G Studies

- Define a G study design for a given data collection design
- Draw a Venn diagram for a given G study design
- Explain the difference between main and interaction effects
- Explain the issue of confounding effects

Two-Facet Universes and Designs

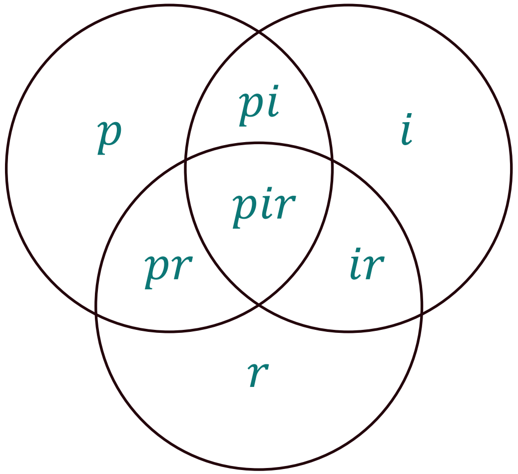
- G study: to obtain estimated variance components for a G study design

- Balanced vs. unbalanced
  - Balanced design: no missing data, or equal sample size for nested facet(s)

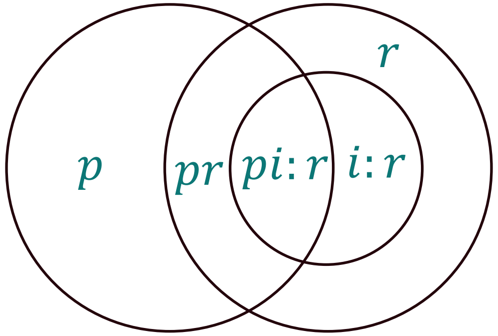
- Random effects models vs. mixed models
  - n = sample size, N= the universe (population) size
  - If $n = N < \infty$ for some facets and $N \to \infty$ for all other facets, the model is the mixed model
  - If $N \to \infty$ for all facets, the model is the random model
  - Multivariate generalizability theory may be preferred (see Brennan, 2001)

Venn Diagrams

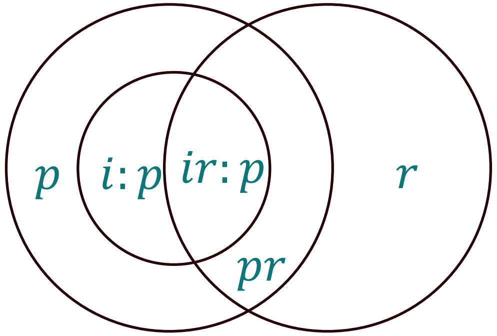
$p \times i \times r$



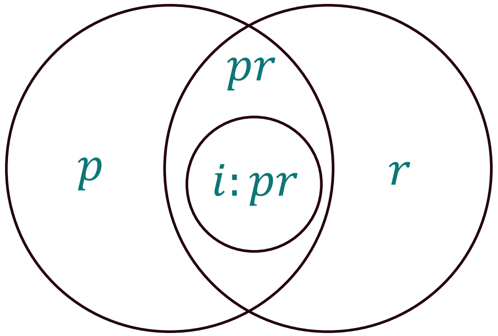
$p \times (i : r)$



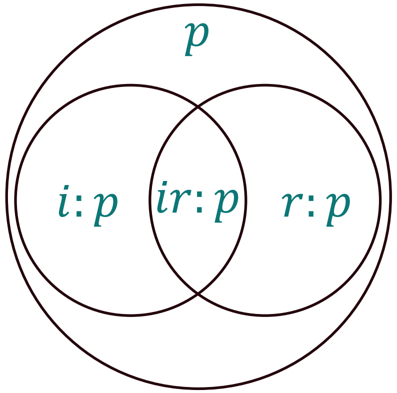
$(i : p) \times r$



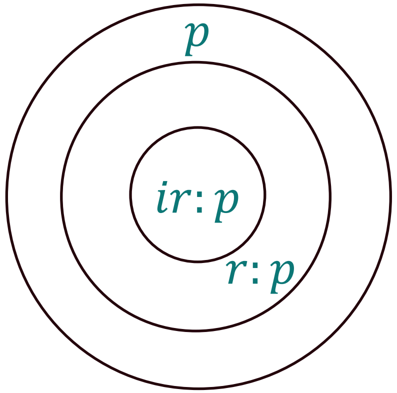
$i : (p \times r)$



$(i \times r) : p$



$i : r : p$

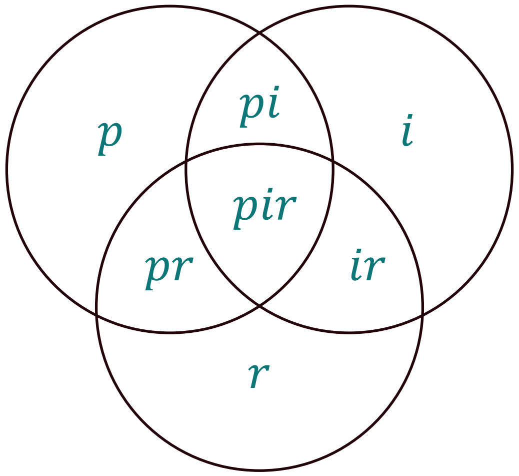


Two-Facet Universes and Designs

- Main effects: represented by a circle
- Interaction effects: combinations for main effects
- Residual error

Design	Main Effects	Interaction Effects
$p \times i \times r$	p, i, r	pi, pr, ir, pir
$p \times (i{:}r)$	$p, r, i{:}r$	$pr, pi{:}r$
$(i{:}p) \times r$	$p, r, i{:}p$	$pr, ir{:}p$
$i{:}(p \times r)$	$p, r, i{:}pr$	pr
$(i \times r){:}p$	$p, i{:}p, r{:}p$	$ir{:}p$
$i{:}r{:}p$	$p, r{:}p, i{:}r{:}p$	

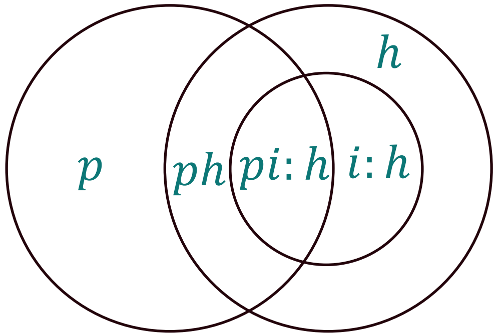
$p \times i \times r$ Design



An Objective Structured Clinical Examination (OSCE) is designed to assess students' clinical skills. In Kreiter et al. (2019), the same two expert raters (r) rated every person (p)'s performance on each of the five stations (i). The station is a specific, timed, and focused assessment point within a structured clinical exam.

	Station1		Station2		Station3		Station4		Station5	
	Rater1	Rater2	Rater1	Rater2	Rater1	Rater2	Rater1	Rater2	Rater1	Rater2
P1										
P2										
P3										
P4										

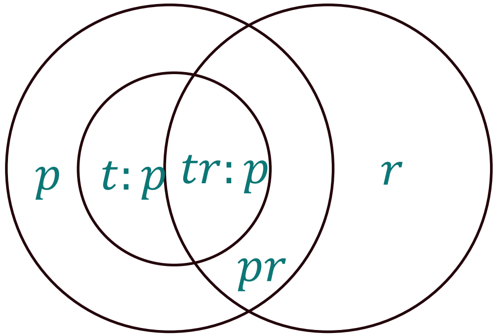
$p \times (i\!:\!h)$ Design



The Iowa Tests of Basic Skills (ITBS) of Maps and Diagrams test for 4th graders (p) had four passages (h), each including 6, 6, 7, and 7 items (i) (Lee & Frisbie, 1999).

	Passage1			Passage2			Passage3			Passage4		
	Item1	…	Item6	Item7	…	Item12	Item13	…	Item19	Item20	…	Item26
P1												
P2												
P3												
P4												

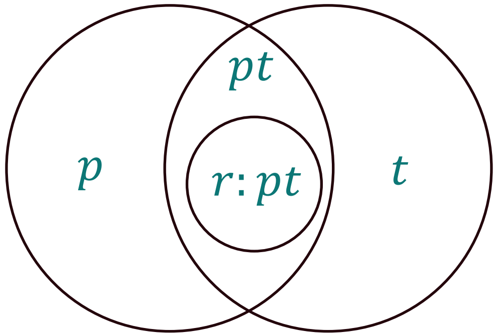
$(t\!:\!p) \times r$ Design



A team of raters (r) evaluated teachers (p) on a series of tasks (t) that were unique to each teacher (McGaw et al., 1971).

		Rater1	Rater2	Rater3
P1	Task1			
	Task2			
P2	Task3			
	Task4			
P3	Task5			
	Task6			
P4	Task7			
	Task8			

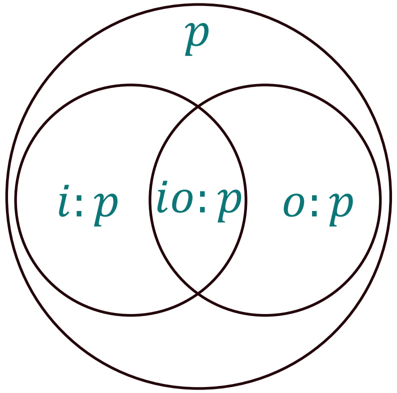
$r:(p \times t)$ Design



For TOEFL, each examinee (p) completed the same set of writing tasks (t), but each "examinee-by-task pair" was rated by different raters (r) (Lee & Kantor, 2005).

	Task1	Task2	Task3	Task4
P1	Rater1	Rater2	Rater3	Rater4
P2	Rater5	Rater6	Rater7	Rater8
P3	Rater9	Rater10	Rater11	Rater12

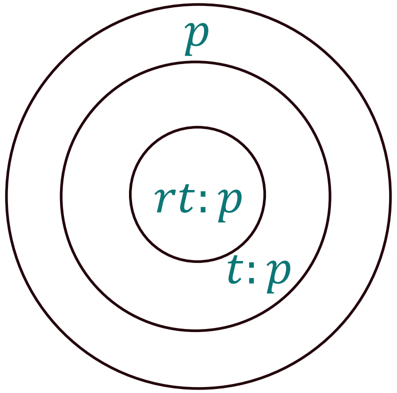
$(i \times o){:}\,p$ Design



Each examinee (p) was administered a set of items (i) on two occasions (o). For an individual person, the same items were administered on each occasion but for different persons, the items were different, and the occasions were different (Brennan, 2002).

	Occasion1		Occasion2		Occasion3		Occasion4		Occasion5		Occasion6	
P1	Item1	Item2	Item1	Item2								
P2					Item3	Item4	Item3	Item4				
P3									Item5	Item6	Item5	Item6

$r\colon t\colon p$ Design



Students (p) completed two different writing tasks (t), and each task was scored by an independent group of two raters (t) (Hathcoat & Penn, 2012).

	Task1		Task2		Task3		Task4		Task5		Task6	
P1	Rater1	Rater2	Rater3	Rater4								
P2					Rater5	Rater6	Rater7	Rater8				
P3									Rater9	Rater10	Rater11	Rater12

The Problem of One

- The problem of one (Brennan, 2017): The data collection design has a single condition for a facet. Effects for at least two facets are confounded.
- It is impossible to estimate variance components for the facet; therefore, subsequent D studies are conducted without including the facet.
- Resulting reliability or error variance statistics will be biased.

References

- Brennan, R. L. (2017). Using G theory to examine confounded effects: “The problem of one”. *Center for Advanced Studies in Measurement and Assessment*.
- Hathcoat, J. D., & Penn, J. D. (2012). Generalizability of Student Writing across Multiple Tasks: A Challenge for Authentic Assessment. *Research & Practice in Assessment*, *7*, 16-28.
- Kreiter, C. D., Zaidi, N. L., & Park, Y. S. (2019). Generalizability theory. In *Assessment in health professions education* (pp. 51-69). Routledge.
- Lee, G., & Frisbie, D. A. (1999). Estimating reliability under a Generalizability theory model for test scores composed of testlets, *Applied Measurement in Education*, *12*(3), 237-255, DOI: 10.1207/S15324818AME1203_2
- Lee, Y. W., & Kantor, R. (2007). Evaluating prototype tasks and alternative rating schemes for a new ESL writing test through G-theory. *International Journal of Testing, 7*(4), 353-385.
- McGaw, B., Wardrop, J. L., & Bunda, M. A. (1972). Classroom observation schemes: Where are the errors?. *American Educational Research Journal, 9*(1), 13-27.
- Peeters, M. J., Cor, M. K., Petite, S. E., & Schroeder, M. N. (2021). Validation evidence using generalizability theory for an objective structured clinical examination. *Innovations in pharmacy, 12*(1), 10-24926.

Decision (D) Studies for Multi-facet Designs

4

Section Learning Objectives

4 D Studies

- Identify three primary factors influencing D study results
- Explain the impact of using a fixed facet in a D study compared to using all random facets
- Compute D study results for a nested design using crossed-design variance components
- Describe traditional reliability coefficients from the perspective of generalizability theory

Introductory Notes

- It is assumed that all facets in the UAO are infinite in size, meaning that they are treated as ***random*** facets and that the G study variance components have been estimated using a random effects model.
- D studies are characterized by:
  - ✓D study sample sizes
  - ✓D study design structure
  - ✓Universe of generalization (i.e., random vs. fixed facets)

Introductory Notes

- The example introduced in Section 1 is extended in this section:
  - ✓A writing assessment involving writing prompts (t) and raters (r)
  - ✓G-study variance components estimated from the $p \times t \times r$ design
- The D study design ($p \times T \times R$) with the same structure as the G study, assuming infinite universe sizes for both T and R (i.e., random), was discussed in Section 1.
- Two additional D study scenarios are considered in this section:
  - ✓A restricted UG, in which facet T is treated as fixed
  - ✓A D study design structure that differs from the G study design

D Study $p \times T \times R$ Design with T Fixed

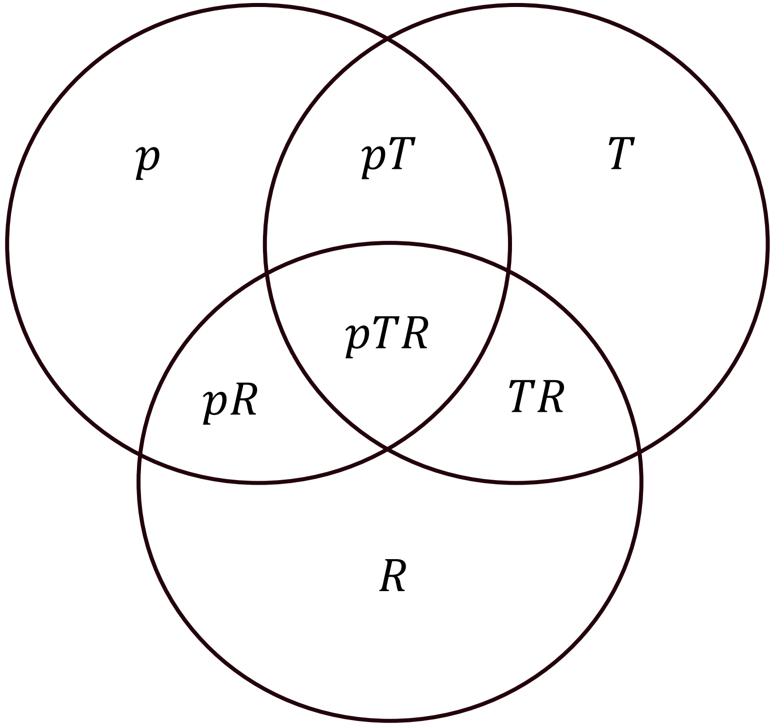
- The investigator is ***not*** interested in generalizing over writing prompts (T), meaning that all parallel measurement procedures will involve the same set of prompts.

- This represents a restricted UG, which differs from the UAO.

- The D study design structure remains the same as that of the G study.

- Under the restricted UG, the universe score variance and error variances are based on different variance components compared to the random $p \times T \times R$ design.

D Study $p \times T \times R$ Design with T Fixed



$n'_t = 3$, $n'_r = 2$			
Effect	**Variance Component**	***T* Random *R* Random**	***T* Fixed *R* Random**
p	.25	τ	τ
T	.02	Δ	
R	.01	Δ	Δ
pT	.05	Δ δ	τ
pR	.02	Δ δ	Δ δ
TR	.00	Δ	Δ
pTR	.02	Δ δ	Δ δ

- $\sigma^2(\tau) = \sigma^2(p) + \sigma^2(pT) = .30$
- $\sigma^2(\Delta) = \sigma^2(R) + \sigma^2(pR) + \sigma^2(TR) + \sigma^2(pTR) = .05$
- $\sigma^2(\delta) = \sigma^2(pR) + \sigma^2(pTR) = .04$

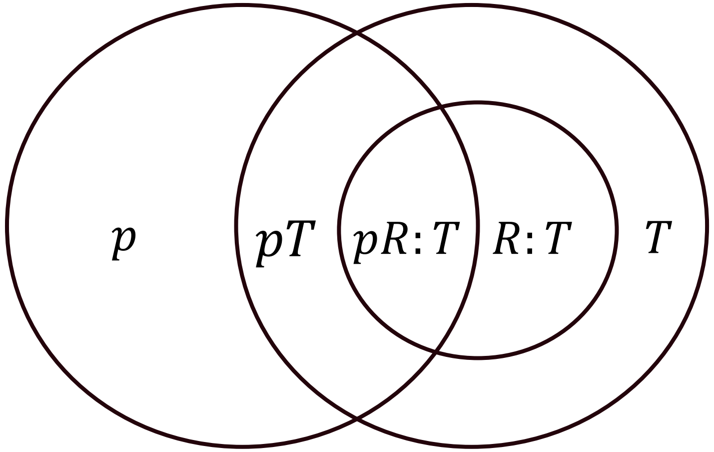
- $E\rho^2 = \frac{\sigma^2(\tau)}{\sigma^2(\tau)+\sigma^2(\delta)} = .88$
- $\Phi = \frac{\sigma^2(\tau)}{\sigma^2(\tau)+\sigma^2(\Delta)} = .86$

D Study $p \times (R\!:\!T)$ Design with T and R Random

- Suppose the investigator wants to use a different set of raters to evaluate each writing prompt. ➔This is a $p \times (R\!:\!T)$ design.

- D study variance components for this design are computed from the G study variance components for the $p \times t \times r$ design using a two-step process:
  1. Apply the confounding rules (Brennan, 2001, pp. 62-63) to convert the G study variance components from the $p \times t \times r$ design into those appropriate for a $p \times (r\!:\!t)$ design.
     - ✓ $\sigma^2(r\!:\!t) = \sigma^2(r) + \sigma^2(tr)$
     - ✓ $\sigma^2(pr\!:\!t) = \sigma^2(pr) + \sigma^2(ptr)$
  2. Use the transformed G study variance components from Step 1 to calculate the D study variance components for the $p \times (R\!:\!T)$ design.

D Study $p \times (R{:}T)$ Design with T and R Random



- $\sigma^2(\tau) = \sigma^2(p) = .25$
- $\sigma^2(\Delta) = \sigma^2(T) + \sigma^2(pT) + \sigma^2(R{:}T) + \sigma^2(pR{:}T) = .1$
- $\sigma^2(\delta) = \sigma^2(pT) + \sigma^2(pR{:}T) = .077$
- $E\rho^2 = \frac{\sigma^2(\tau)}{\sigma^2(\tau)+\sigma^2(\delta)} = .76$
- $\Phi = \frac{\sigma^2(\tau)}{\sigma^2(\tau)+\sigma^2(\Delta)} = .71$

$p \times t \times r$	
Effect	Var Comp
p	.25
t	.06
r	.02
pt	.15
pr	.04
tr	.00
ptr	.12

$p \times (r{:}t)$	
Effect	Var Comp
p	.25
t	.06
$r{:}t$	.02
pt	.15
$pr{:}t$	.16

$p \times (R{:}T)$	
Effect	Var Comp
p	.250
T	.020
$R{:}T$	.003
pT	.050
$pR{:}T$	.027

$n'_t = 3$
$n'_r = 2$

Comparison of Three D Studies

$n'_t = 3$, $n'_r = 2$			
	$p \times T \times R$ T Random R Random	$p \times T \times R$ T Fixed R Random	$p \times (R{:}T)$ T Random R Random
$\sigma^2(\tau)$	.25	.30	.25
$\sigma^2(\Delta)$	.12	.05	.10
$\sigma^2(\delta)$	.09	.04	.08
$E\rho^2$	.74	.88	.76
Φ	.68	.86	.71

Traditional Reliability Coefficients

- Test-retest (TR)
  - ✓Correlation of scores from the same test form administered on different occasions
  - ✓Items do not contribute to the error, but occasions do
- Parallel-forms (PF)
  - ✓Correlation of scores from different test forms administered on different occasions
  - ✓Both items and occasions contribute to the measurement error
- Internal consistency (IC)
  - ✓Single administration of a test form

Traditional Reliability Coefficients

- Test-retest reliability (TR)
  - ✓G study: $p \times i \times o$ with $n_o = 2$
  - ✓D study: $p \times I \times O$ with $n'_o = 1$; O random and I fixed
  - ✓$E\rho^2 = \frac{\sigma^2(p)+\sigma^2(pI)}{\sigma^2(p)+\sigma^2(pI)+[\sigma^2(pO)+\sigma^2(pIO)]}$
- Parallel-forms reliability (PF)
  - ✓G study: $p \times (i{:}o)$ with $n_o = 2$
  - ✓D study: $p \times (I{:}O)$ with $n'_o = 1$; both I and O random
  - ✓$E\rho^2 = \frac{\sigma^2(p)}{\sigma^2(p)+[\sigma^2(pO)+\sigma^2(pI{:}O)]} = \frac{\sigma^2(p)}{\sigma^2(p)+[\sigma^2(pO)+\sigma^2(pI)+\sigma^2(pIO)]}$

Traditional Reliability Coefficients

- Internal consistency (IC)
  - ✓G study: $p \times i \times o$ or $p \times (i{:}o)$ with $n_o \geq 2$
  - ✓D study: $p \times I \times O$ or $p \times (I{:}O)$ with $n'_o = 1$; O fixed and I random
  - ✓$E\rho^2 = \frac{\sigma^2(p)+\sigma^2(pO)}{\sigma^2(p)+\sigma^2(pO)+[\sigma^2(pI)+\sigma^2(pIO)]}$

- It is almost always the case that

$$PF \leq TR \leq IC$$

Other Issues

- Large sample sizes (e.g., n_t, n_r) are highly desirable for a G study.
- A fully crossed G-study design is generally preferable.
- Not discussed in this module:
    - ✓Unbalanced designs
    - ✓Multivariate GT

References

- Brennan, R. L. (2001). *Generalizability theory*. New York: Springer-Verlag.

Conducting G and D Studies in R

5

Section Learning Objectives

5 R Demonstration

Prepare data for G study analysis in R for both crossed $p \times i$ and $p \times (r{:}\,i)$ nested designs.

Use *lmer()* function to estimate variance components in G study.

Compute D study statistics for various test conditions using R.

Interpret G study and D study results.

Computing Variance Components Using *lmer()* Function

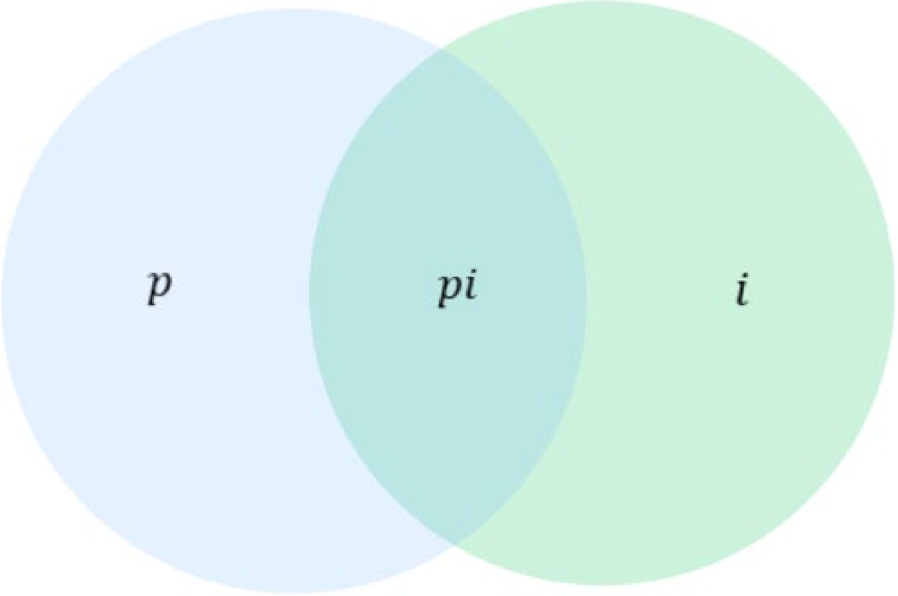
- *lmer()* is a function from the lme4 package in R (Bates et al., 2015).
- *lmer()* function fits linear mixed-effects models using restricted maximum likelihood (REML) estimation.
- It allows for the decomposition of variance into components attributable to different random factors.
- In the $p \times i$ design, person and items are specified as random effects to estimate variances due to individual difference, item variability, and person-by-item interaction.

$$X_{pi} = \mu + v_p + v_i + v_{pi}$$

Sample Dataset for G Study $p \times i$ Design

DATA STRUCTURE:

- 36 persons (rows)
- 30 dichotomous items (columns)
- Crossed design: every person answered every item
- Total observations: 36 x 30 = 1,080



1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	0	1	1
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	1	0	1	1
1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	0	1	1	1	0
1	0	0	0	1	1	1	0	1	0	0	0	0	1	1	0	0	1	0	0	0	1	1	1	0	0	1	0	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	0	0	1	1	1	1	1
1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1
1	1	1	0	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0	0	1	0	0	0	1	0
0	1	1	0	0	1	0	1	0	0	0	1	1	1	1	0	1	1	0	0	0	1	1	0	1	1	0	0	1	1
1	0	0	1	0	1	1	1	1	1	0	1	1	1	0	0	1	1	0	0	0	1	1	0	1	0	1	0	1	1
1	1	1	1	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1
1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	1	0
1	0	0	0	0	0	0	1	1	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	0	0	0	0	1	1
1	1	0	0	0	1	1	1	1	0	0	1	1	1	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	1
1	1	1	0	1	1	1	1	1	0	0	1	1	0	1	0	1	1	0	1	1	0	0	1	1	0	1	0	1	1
1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	0	1	1	0	1	0	0	0	0	1	1	0	0	1	1
1	0	0	0	1	0	1	1	1	0	1	0	1	0	1	0	1	1	0	1	0	0	1	0	1	1	0	1	1	1
1	1	1	0	0	0	0	1	0	0	1	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	1	1	1	0
1	1	1	1	0	0	0	0	1	0	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1
1	1	0	0	1	1	1	0	1	0	1	1	0	1	0	0	1	1	0	1	0	1	1	0	1	0	1	0	1	1
0	1	0	0	1	1	1	1	0	0	0	0	1	0	1	0	1	1	0	0	1	1	0	1	1	0	0	0	1	1
1	1	0	1	0	1	0	1	1	0	1	0	0	1	1	0	1	1	0	0	1	1	0	0	1	1	1	0	1	1
1	1	1	0	0	1	0	1	1	1	1	1	1	0	0	0	1	0	1	1	0	1	1	0	1	0	0	0	1	1
1	0	1	1	1	1	1	0	1	0	0	0	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	0	1	1
1	0	1	0	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	0	0	1	1	0	1	1	1
0	1	1	0	1	1	0	0	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	0	0	0	0	1	0
1	1	0	0	1	1	1	1	1	0	1	0	1	1	1	0	0	0	0	1	1	0	0	0	1	0	0	0	1	1
1	1	0	1	0	1	0	0	0	1	1	0	1	0	1	1	0	1	1	0	1	0	0	0	1	1	0	0	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1
1	1	0	0	0	1	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	0	1	1
1	1	0	0	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1	0
1	1	1	0	0	1	0	1	0	0	1	1	1	1	1	0	1	1	0	0	1	1	0	0	1	0	1	0	1	1
1	1	0	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	0	1	1	1	0	0	1	0	1	0	1	0
0	1	0	0	0	1	1	0	0	0	1	0	1	0	1	1	1	0	0	1	1	0	0	0	1	0	1	0	1	0
1	1	1	0	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	0	1	0	1	0	0	1	1	1
0	0	1	0	0	0	1	1	1	0	1	0	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	0	1	1

Prepare the Data for G Study $p \times i$ Design in R

Data Preparation: Convert the wide format (person x item) into long format with one row per response.

```
install.packages("tidyr")
library(tidyr)
data <- read.table("C:/Users/data/data1", header = FALSE, sep = "", fill = TRUE)
colnames(data) <- paste0("item", 1:30)
data$person <- 1:nrow(data)
data_long <- pivot_longer(data,
                          cols = starts_with("item"),
                          names_to = "item",
                          values_to = "score")
```

	item1	item2	item3	item4	item5	item6	item7	item8	item9	item10
1	1	1	1	0	0	1	1	1	1	0
2	0	0	0	1	0	0	0	0	0	0
3	1	1	1	0	1	1	1	1	1	1
4	1	1	1	0	1	1	1	0	1	0
5	1	0	0	0	1	1	1	0	1	0
6	1	1	1	1	1	1	1	1	1	1
7	1	1	0	1	1	1	1	0	1	0
8	1	1	1	0	0	0	0	1	1	1
9	0	1	1	0	0	1	0	1	0	0
10	1	0	0	1	0	1	1	1	1	1

	person	item	score
1	1	item1	1
2	1	item2	1
3	1	item3	1
4	1	item4	0
5	1	item5	0
6	1	item6	1
7	1	item7	1
8	1	item8	1
9	1	item9	1
10	1	item10	0
11	1	item11	0
12	1	item12	0
13	1	item13	1
14	1	item14	1
15	1	item15	1
16	1	item16	1
17	1	item17	1
18	1	item18	1
19	1	item19	1

1062	36	item12	0
1063	36	item13	1
1064	36	item14	1
1065	36	item15	1
1066	36	item16	1
1067	36	item17	1
1068	36	item18	0
1069	36	item19	0
1070	36	item20	1
1071	36	item21	1
1072	36	item22	1
1073	36	item23	0
1074	36	item24	1
1075	36	item25	1
1076	36	item26	0
1077	36	item27	1
1078	36	item28	0
1079	36	item29	1
1080	36	item30	1

Using *lmer* Function for G Study Analysis

```
##Install Package##
install.packages("lme4")
library(lme4)
##Linear mixed-effects model##
model <- lmer(score ~ 1 + (1|person) + (1|item), data = data_long)
summary(model)
```

- Mixed-model formula used to model the two crossing factors:
  - ✓ $1 + (1 \mid g1) + (1 \mid g2)$ where $g1$ and $g2$ represent person and item factors.
- Output: Variance components for person, item, and interaction.

G Study Output from R

```
> model <- lmer(score ~ 1 + (1|person) + (1|item), data = data_long)
> summary(model)
Linear mixed model fit by REML ['lmerMod']
Formula: score ~ 1 + (1 | person) + (1 | item)
   Data: data_long

REML criterion at convergence: 1388.6

Scaled residuals: 
    Min      1Q  Median      3Q     Max 
-2.1767 -0.9894  0.3699  0.7466  2.0325 

Random effects:
 Groups   Name        Variance Std.Dev.
 person   (Intercept) 0.01462  0.1209  
 item     (Intercept) 0.03185  0.1785  
 Residual             0.19242  0.4387  
Number of obs: 1080, groups:  person, 36; item, 30

Fixed effects:
            Estimate Std. Error t value
(Intercept)  0.61296    0.04057   15.11
```

D Study for the $p \times I$ Design

- Define the test lengths.
- Compute the following values for each test lengths using the estimated variance components from the G study:
  - ✓Relative error variance: $\sigma^2(\delta) = \frac{\sigma^2(pi)}{n_i'}$
  - ✓ Absolute error variance:

$$\sigma^2(\Delta) = \frac{\sigma^2(i)}{n_i'} + \frac{\sigma^2(pi)}{n_i'}$$

  - ✓G Coefficients: $E\rho^2 = \frac{\sigma^2(p)}{\sigma^2(p)+\sigma^2(\delta)}$
  - ✓Phi Coefficients: $\Phi = \frac{\sigma^2(p)}{\sigma^2(p)+\sigma^2(\Delta)}$

```r
## G study results
sigma2_p <- 0.0146
sigma2_i <- 0.0318
sigma2_pi <- 0.1924
n_p <- 36
## Define test lengths
n_items_list <- c(5, 10, 15, 20, 25, 30, 40, 50)
## Vectors for storing results
sigma2_delta <- numeric(length(n_items_list))
sigma2_Delta <- numeric(length(n_items_list))
Ep2 <- numeric(length(n_items_list))
phi <- numeric(length(n_items_list))
## Loop
for (i in seq_along(n_items_list)) {
  n_i <- n_items_list[i]

  sigma2_delta[i] <- sigma2_pi / n_i
  sigma2_Delta[i] <- sigma2_i / n_i + sigma2_pi / n_i
  Ep2[i]          <- sigma2_p / (sigma2_p + sigma2_delta[i])
  phi[i]          <- sigma2_p / (sigma2_p + sigma2_Delta[i])
}
## Final output
d_study_table <- data.frame(
  n_i = n_items_list,
  sigma2_delta = round(sigma2_delta, 4),
  sigma2_Delta = round(sigma2_Delta, 4),
  Ep2 = round(Ep2, 4),
  phi = round(phi, 4)
)
```

D Study Output for $p \times I$ Design

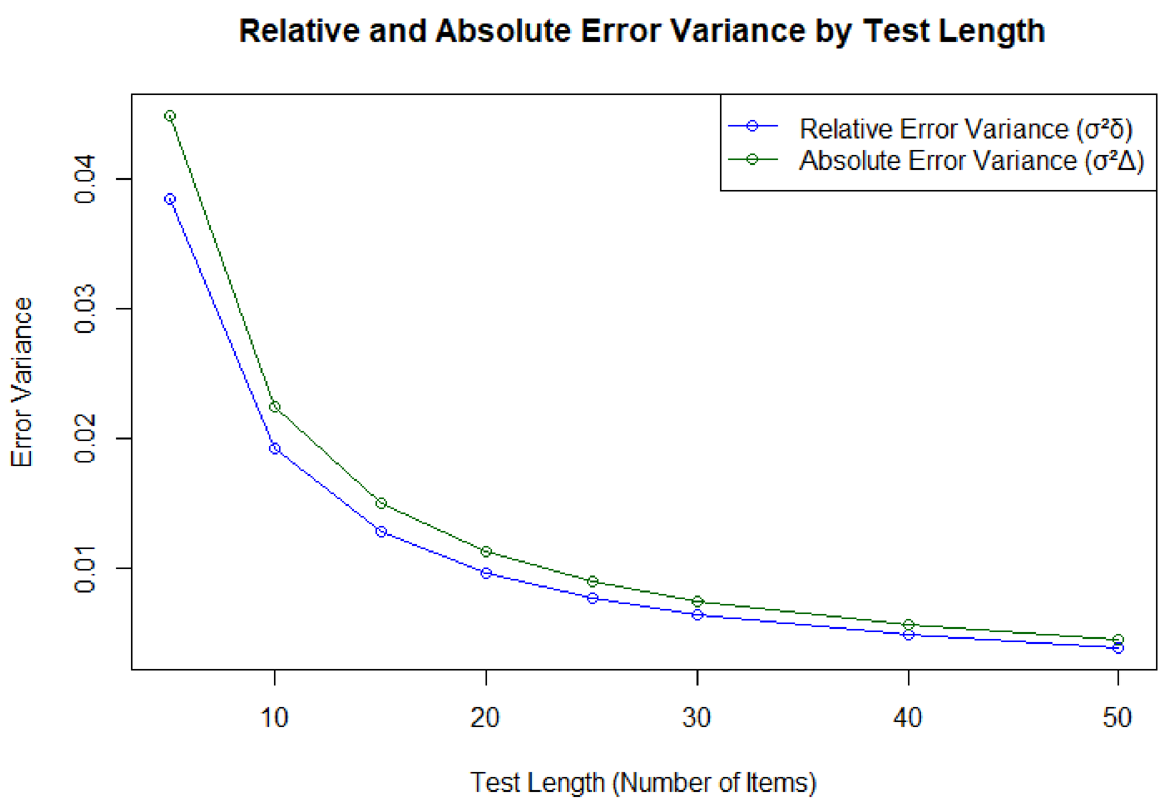
```
> print(d_study_table)
  n_i sigma2_delta sigma2_Delta    Ep2    phi
1   5       0.0385       0.0448 0.2751 0.2456
2  10       0.0192       0.0224 0.4314 0.3944
3  15       0.0128       0.0149 0.5323 0.4941
4  20       0.0096       0.0112 0.6028 0.5657
5  25       0.0077       0.0090 0.6548 0.6195
6  30       0.0064       0.0075 0.6948 0.6614
7  40       0.0048       0.0056 0.7522 0.7226
8  50       0.0038       0.0045 0.7914 0.7650
```

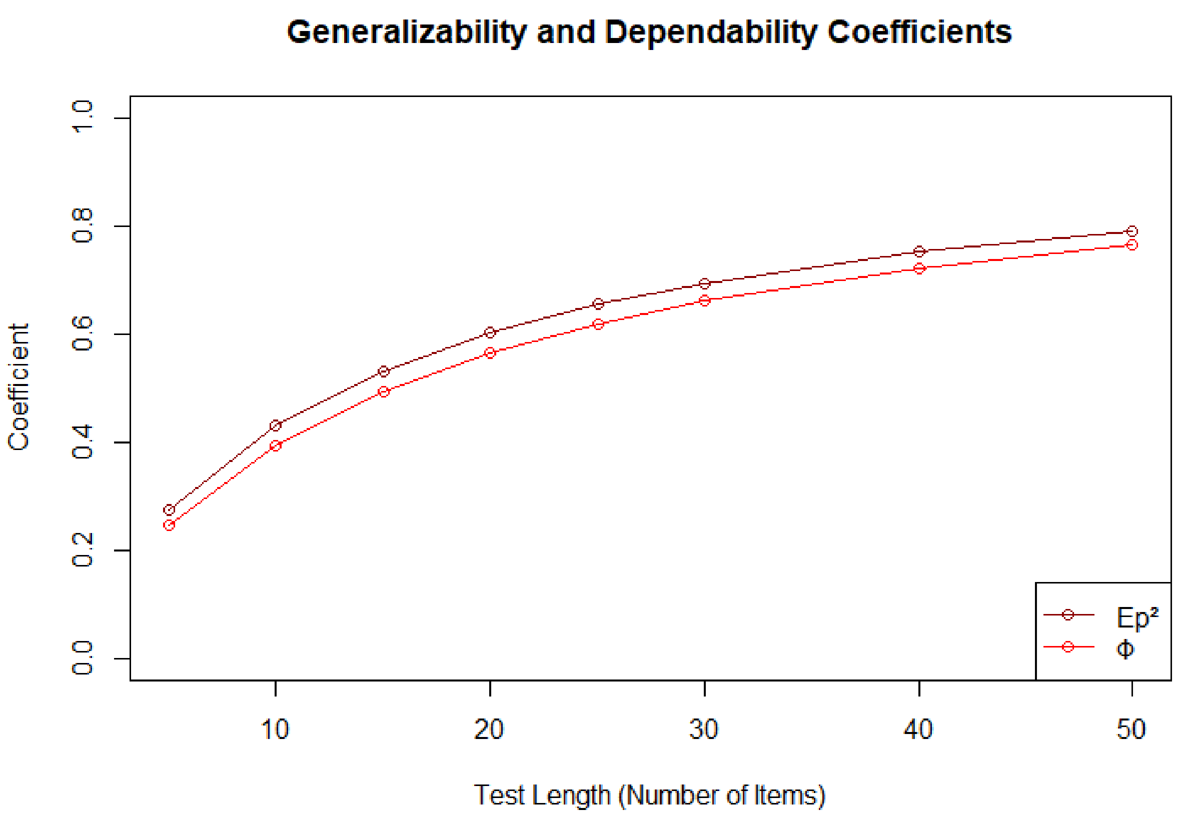
- ✓ The output table on the left provides estimated D study results for 5, 10, 15, 20, 25, 30, 40, and 50 items.
- ✓ Optional Step: Plotting the error variances and the coefficients in R for visualization.

```
# Plot 1: Error variances
plot(n_items_list, sigma2_delta, type = "o", col = "blue", ylim = range(c(sigma2_delta, sigma2_Delta)),
     xlab = "Test Length (Number of Items)", ylab = "Error Variance",
     main = "Relative and Absolute Error Variance by Test Length")
lines(n_items_list, sigma2_Delta, type = "o", col = "darkgreen")
legend("topright", legend = c("Relative Error Variance (σ²δ)", "Absolute Error Variance (σ²Δ)"),
       col = c("blue", "darkgreen"), lty = 1, pch = 1)

# Plot 2: G and Phi coefficients
plot(n_items_list, Ep2, type = "o", col = "darkred", ylim = c(0,1),
     xlab = "Test Length (Number of Items)", ylab = "Coefficient",
     main = "Generalizability and Dependability Coefficients")
lines(n_items_list, phi, type = "o", col = "red")
legend("bottomright", legend = c("Ep²", "Φ"),
       col = c("darkred", "red"), lty = 1, pch = 1)
```

Visualizing the Error Variances and G Coefficients Results





2 $p \times (r\!:\!i)$ Design

In this design, we will utilize one of the Genova sample datasets: *Synthetic Data Set No. 4*. The G-study design is denoted as $p \times (r\!:\!i)$, indicating that raters are nested within items, and examinees are crossed with items in the universe of admissible observations. Both raters and items are considered random facets and were sampled independently. Raters were randomly assigned to specific items, ensuring no overlap in rater assignments across items.

Dataset Overview and $p \times (r\!: i)$ Design

Data Structure:

✓10 persons

✓3 polytomous items

✓4 raters nested in each of the 3 items

✓Persons crossed with items

✓Raters nested within items

Table Data Structure for the $p \times (r\!: i)$ Design

	$i1$				$i2$				$i3$			
Person	$r1$	$r2$	$r3$	$r4$	r5	$r6$	$r7$	$r8$	$r9$	$r10$	$r11$	$r12$
1	5	6	5	5	5	3	4	5	6	7	3	3
2	9	3	7	7	7	5	5	5	7	7	5	2
3	3	4	3	3	5	3	3	5	6	5	1	6
4	7	5	5	3	3	1	4	3	5	3	3	5
5	9	2	9	7	7	7	3	7	2	7	5	3
6	3	4	3	5	3	3	6	3	4	5	1	2
7	7	3	7	7	7	5	5	7	5	5	5	4
8	5	8	5	7	7	5	5	4	3	2	1	1
9	9	9	8	8	6	6	6	5	5	8	1	1
10	4	4	4	3	3	5	6	5	5	7	1	1

Prepare Data for G Study Analysis Using *lmer* for $p \times (r{:}i)$ Design

```
library(tidyr)
library(dplyr)
data <- read.table("C:/Users/data/Genova data.txt", header = FALSE, sep = "", fill = TRUE)
colnames(data) <- c(paste0("I1_R", 1:4), paste0("I2_R", 1:4), paste0("I3_R", 1:4))
data$Person <- factor(1:nrow(data))
long_data <- data %>%
  pivot_longer(
    cols = -Person,
    names_to = c("Item", "Rater"),
    names_pattern = "I(\\d)_R(\\d)",
    values_to = "Score"
  )
long_data <- long_data %>%
  mutate(
    Item = factor(paste0("I", Item)),
    Rater = factor(Rater),
    RaterInItem = factor(paste0("I", Item, "_R", Rater))
  )
```

I1_R1	I1_R2	I1_R3	I1_R4	I2_R1	I2_R2	I2_R3	I2_R4	I3_R1	I3_R2	I3_R3	I3_R4	Person
5	6	5	5	5	3	4	5	6	7	3	3	1
9	3	7	7	7	5	5	5	7	7	5	2	2
3	4	3	3	5	3	3	5	6	5	1	6	3
7	5	5	3	3	1	4	3	5	3	3	5	4
9	2	9	7	7	7	3	7	2	7	5	3	5
3	4	3	5	3	3	6	3	4	5	1	2	6
7	3	7	7	7	5	5	7	5	5	5	4	7
5	8	5	7	7	5	5	4	3	2	1	1	8
9	9	8	8	6	6	6	5	5	8	1	1	9
4	4	4	3	3	5	6	5	5	7	1	1	10

	Person	Item	Rater	Score	RaterInItem
1	1	I1	1	5	II1_R1
2	1	I1	2	6	II1_R2
3	1	I1	3	5	II1_R3
4	1	I1	4	5	II1_R4
5	1	I2	1	5	II2_R1
6	1	I2	2	3	II2_R2
7	1	I2	3	4	II2_R3
8	1	I2	4	5	II2_R4
9	1	I3	1	6	II3_R1
10	1	I3	2	7	II3_R2
11	1	I3	3	3	II3_R3
12	1	I3	4	3	II3_R4
13	2	I1	1	9	II1_R1
14	2	I1	2	3	II1_R2
15	2	I1	3	7	II1_R3

Showing 1 to 16 of 120 entries, 5 total columns

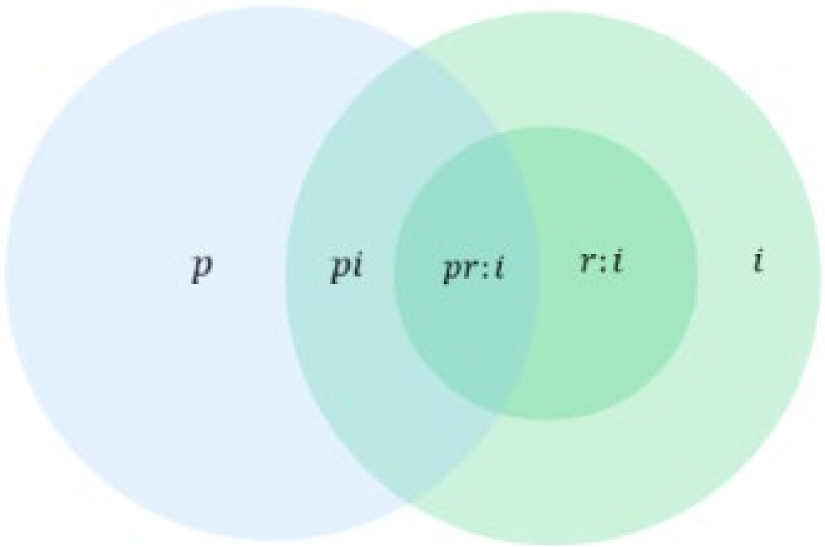
G Study Using $lmer$ for $p \times (r{:}\,i)$ Design

STEPS:

- Model equation:

$$X_{pir} = \mu + v_p + v_i + v_{r:i} + v_{pi} + v_{pr:i}$$

- Decomposition of the Variance components:
  - ✓ σ_p^2
  - ✓ σ_i^2
  - ✓ $\sigma_{r:i}^2$
  - ✓ σ_{pi}^2
  - ✓ $\sigma_{pr:i}^2$




```r
library(lme4)
# Fit G-study model with p × (r : i) design
g_model <- lmer(Score ~ 1 +
                      (1 | Person) +
                      (1 | Item) +
                      (1 | RaterInItem) +
                      (1 | Person:Item),
                    data = long_data)
# Get variance components
vc <- as.data.frame(VarCorr(g_model))
print(vc)
```

G Study Output for $p \times (r\!:\!i)$ Design

Estimated Variance Components:

```
> print(vc)
           grp        var1 var2      vcov     sdcor
1  Person:Item (Intercept) <NA> 0.5595499 0.7480307
2 RaterInItem (Intercept) <NA> 0.6474889 0.8046669
3       Person (Intercept) <NA> 0.4730233 0.6877669
4         Item (Intercept) <NA> 0.3253211 0.5703693
5     Residual        <NA> <NA> 2.3802862 1.5428176
```

D Studies for $p \times (R{:}I)$ Design

STEPS:

- Define D study conditions.
- Use G study estimated variance components to:
  - ✓ Compute the relative and absolute variances.
  - ✓ Compute the G and Phi coefficients for each condition.

```
# Function to compute D-study results for given items and raters
compute_G_Phi <- function(n_i, n_r) {
  rel_error <- (var_pi / n_i) + (var_pir / (n_i * n_r))
  abs_error <- (var_i / n_i) + (var_r_i / (n_i * n_r)) + (var_pi / n_i) + (var_pir / (n_i * n_r))

  G_coef <- var_p / (var_p + rel_error)
  Phi_coef <- var_p / (var_p + abs_error)

  return(data.frame(
    Relative_Error = rel_error,
    Absolute_Error = abs_error,
    G_Coefficient = G_coef,
    Phi_Coefficient = Phi_coef
  ))
}
# Apply D-study conditions
dstudy_conditions <- data.frame(
  Items = c(1, 2, 3, 4, 5, 6),
  Raters = c(12, 6, 4, 3, 2, 2)
)

# Compute for each row
d_results <- do.call(rbind, apply(dstudy_conditions, 1, function(row) {
  compute_G_Phi(as.numeric(row["Items"]), as.numeric(row["Raters"]))
}))
# Combine with the original D-study table
final_results <- cbind(dstudy_conditions, round(d_results, 4))
print(final_results)
```

D Study Summary Output Table for $p \times (R{:}I)$ Design

```
print(final_results)
```

Items	Raters	Relative_Error	Absolute_Error	G_Coefficient	Phi_Coefficient
1	12	0.7579	1.1372	0.3843	0.2938
2	6	0.4781	0.6948	0.4973	0.4051
3	4	0.3849	0.5473	0.5514	0.4636
4	3	0.3382	0.4735	0.5831	0.4997
5	2	0.3499	0.4798	0.5748	0.4965
6	2	0.2916	0.3998	0.6186	0.5420

Note. The selection of D-study conditions was informed by Table 4.6 in Brennan (2001), ensuring that the product of n_i, and n_r (numbers of items and raters) remained as close to 12 as possible without exceeding it.

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). *Fitting linear mixed-effects models using lme4*. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01
- Brennan, R. L. (2001). *Generalizability theory*. Springer-Verlag.

Conducting G and D Studies in GENOVA

6

Section Learning Objectives

6 GENOVA Demonstration

Prepare data for GENOVA analysis

Create GENOVA control cards for $p \times i$ design

Create GENOVA control cards for $p \times (r:i)$ design

Interpret the output generated by GENOVA

Data Preparation

Data structure

1. $p \times i$ design: (data 1)

✓ 36 person, 30 items.

2. $p \times (r\!:\!i)$ design: (data 2)

✓ 10 person, 4 raters nested within 3 items.

Data 1 Item →

Person ↓

1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	0	1	1
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	1	0	1	1
1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	0	1	1	1	0
1	0	0	0	1	1	1	0	1	0	0	0	0	1	1	0	0	1	0	0	0	1	1	1	0	0	1	0	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	0	0	1	1	1	1	1
1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1
1	1	1	0	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0	0	1	0	0	0	1	0
0	1	1	0	0	1	0	1	0	0	0	1	1	1	1	0	1	1	0	0	0	1	1	0	1	1	0	0	1	1
1	0	0	1	0	1	1	1	1	1	0	1	1	1	0	0	1	1	0	0	0	1	1	0	1	0	1	0	1	1
1	1	1	1	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1
1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	1	0
1	0	0	0	0	0	0	1	1	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	0	0	0	0	1	1
1	1	0	0	0	1	1	1	1	0	0	1	1	1	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	1
1	1	1	0	1	1	1	1	1	0	0	1	1	0	1	0	1	1	0	1	1	0	0	1	1	0	1	0	1	1
1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	0	1	1	0	1	0	0	0	0	1	1	0	0	1	1
1	0	0	0	1	0	1	1	1	0	1	0	1	0	1	0	1	1	0	1	0	0	1	0	1	1	0	1	1	1
1	1	1	0	0	0	0	1	0	0	1	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	1	1	1	0
1	1	1	1	0	0	0	0	1	0	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1
1	1	0	0	1	1	1	0	1	0	1	1	0	1	0	0	1	1	0	1	0	1	1	0	1	0	1	0	1	1
0	1	0	0	1	1	1	1	0	0	0	0	1	0	1	0	1	1	0	0	1	1	0	1	1	0	0	0	1	1
1	1	0	1	0	1	0	1	1	0	1	0	0	1	1	0	1	1	0	0	1	1	0	0	1	1	1	0	1	1
1	1	1	0	0	1	0	1	1	1	1	1	1	0	0	0	1	0	1	1	0	1	1	0	1	0	0	0	1	1
1	0	1	1	1	1	1	0	1	0	0	0	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	0	1	1
1	0	1	0	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	0	0	1	1	0	1	1	1
0	1	1	0	1	1	0	0	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	0	0	0	0	1	0
1	1	0	0	1	1	1	1	1	0	1	0	1	1	1	0	0	0	0	1	1	0	0	0	1	0	0	0	1	1
1	1	0	1	0	1	0	0	0	1	1	0	1	0	1	1	0	1	1	0	1	0	0	0	1	1	0	0	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1
1	1	0	0	0	1	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	0	1	1
1	1	0	0	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1	0
1	1	1	0	0	1	0	1	0	0	1	1	1	1	1	0	1	1	0	0	1	1	0	0	1	0	1	0	1	1
1	1	0	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	0	1	1	1	0	0	1	0	1	0	1	0
0	1	0	0	0	1	1	0	0	0	1	0	1	0	1	1	1	0	0	1	1	0	0	0	1	0	1	0	1	0
1	1	1	0	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	0	1	0	1	0	0	1	1	1
0	0	1	0	0	0	1	1	1	0	1	0	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	0	1	1

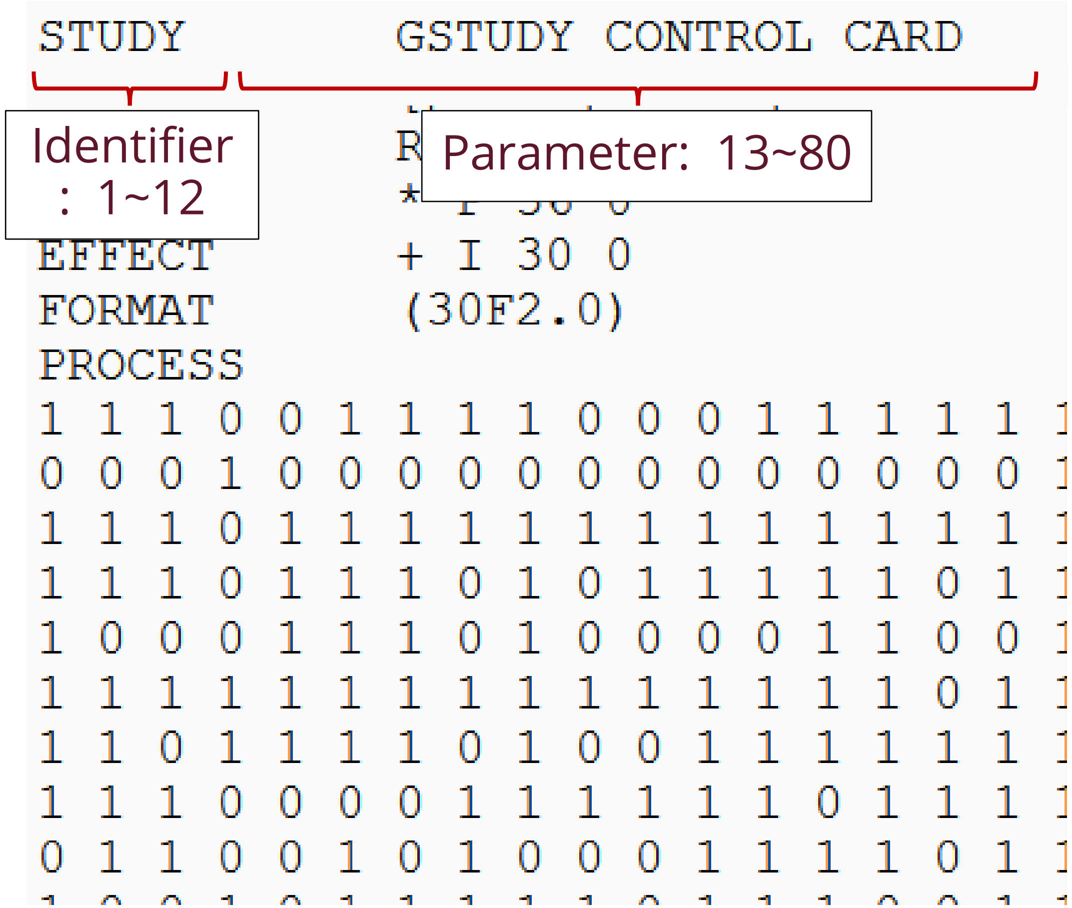
Data 2

5	6	5	5	5	3	4	5	6	7	3	3
9	3	7	7	7	5	5	5	7	7	5	2
3	4	3	3	5	3	3	5	6	5	1	6
7	5	5	3	3	1	4	3	5	3	3	5
9	2	9	7	7	7	3	7	2	7	5	3
3	4	3	5	3	3	6	3	4	5	1	2
7	3	7	7	7	5	5	7	5	5	5	4
5	8	5	7	7	5	5	4	3	2	1	1
9	9	8	8	6	6	6	5	5	8	1	1
4	4	4	3	3	5	6	5	5	7	1	1

Creating Control Cards for GENOVA

Control cards for $p \times i$, G study

- **Format**
  - ✓ The control card should be prepared in a text (.txt) format.
  - ✓ Each card consists of an Identifier and Parameters.
  - ✓ The Identifier must be located in columns 1–12, and the Parameters in columns 13–80.



Creating Control Cards for GENOVA

Control cards for $p \times i$, G study

1. STUDY

✓ Can be used interchangeably with 'GSTUDY'.

✓ User-defined alphanumeric heading that appears on output page.

2. COMMENT

✓ comments or notes for output file.

```
STUDY        GSTUDY CONTROL CARD
COMMENT      [p x i] design
OPTIONS      RECORDS 2
EFFECT       * P 36 0
EFFECT       + I 30 0
FORMAT       (30F2.0)
PROCESS
1 1 1 0 0 1 1 1 1 0 0 0 1 1 1 1 1 1
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 0 1 1 1 0 1 0 1 1 1 1 1 0 1 1
1 0 0 0 1 1 1 0 1 0 0 0 0 1 1 0 0 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1
1 1 0 1 1 1 1 0 1 0 0 1 1 1 1 1 1 1
1 1 1 0 0 0 0 1 1 1 1 1 1 0 1 1 1 1
0 1 1 0 0 1 0 1 0 0 0 1 1 1 1 0 1 1
```

Creating Control Cards for GENOVA

Control cards for $p \times i$, G study

3. OPTIONS

- RECORDS
  - ✓ ALL : print all input records.
  - ✓ NONE : suppress the printing of input records.
  - ✓ NUMEREC : print only the first and last NUMEREC input records.
- EMS / ALGORITHM
  - ✓ Specify which estimated variance components from the G-study are to be used in D-study.
  - ✓ Default = EMS

```
STUDY      GSTUDY CONTROL CARD
COMMENT    [p x i] design
OPTIONS    RECORDS 2
EFFECT
EFFECT
FORMAT
PROCESS
1 1 1 0 0 1 1 1 1 0 0 0 1 1 1 1 1 1
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 0 1 1 1 0 1 0 1 1 1 1 1 0 1 1
1 0 0 0 1 1 1 0 1 0 0 0 0 1 1 0 0 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1
1 1 0 1 1 1 1 0 1 0 0 1 1 1 1 1 1 1
1 1 1 0 0 0 0 1 1 1 1 1 1 0 1 1 1 1
0 1 1 0 0 1 0 1 0 0 0 1 1 1 1 0 1 1
```

RECORD

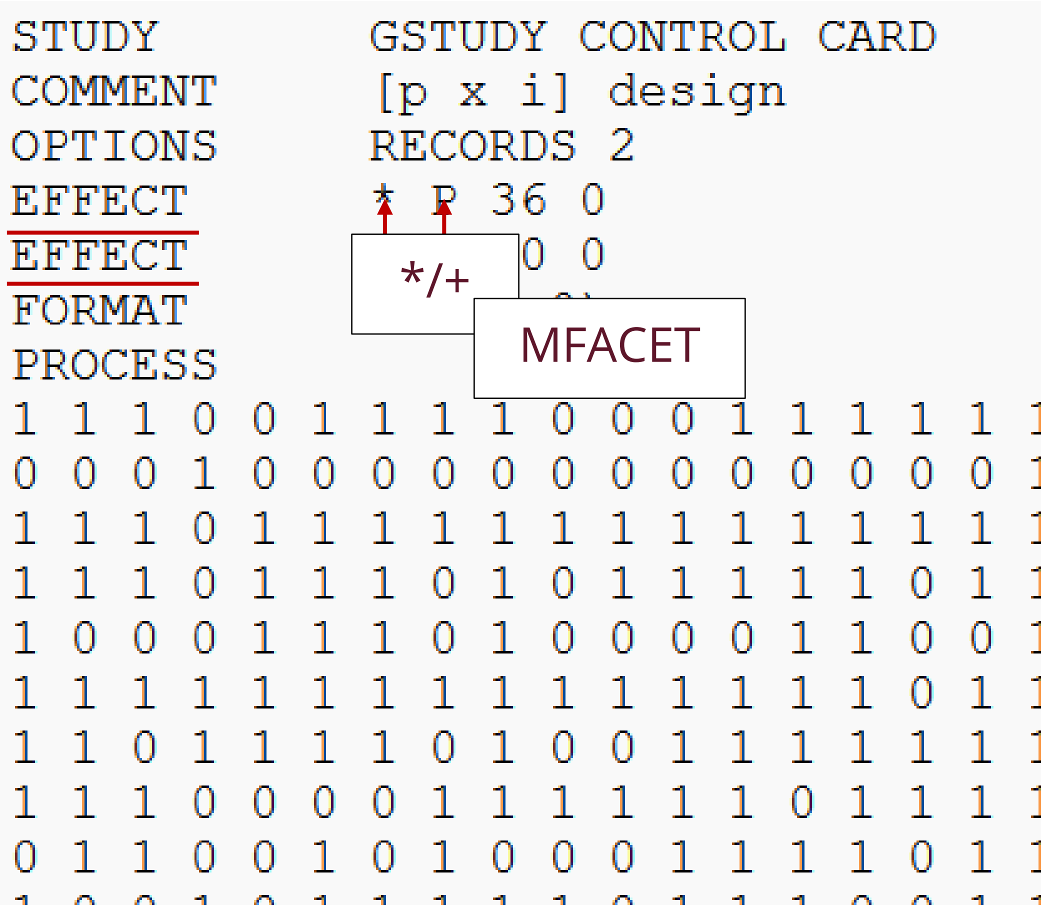
EMS / ALGORITHM

Creating Control Cards for GENOVA

Control cards for $p \times i$, G study

4. EFFECT

- */+ (Optional)
  - ✓ *: Marks the facet that defines each record.
  - ✓ +: Print cell mean scores of facet.
- MFACET
  - ✓ Facet for the study.
  - ✓ Letters (A~Z) and colon (:).



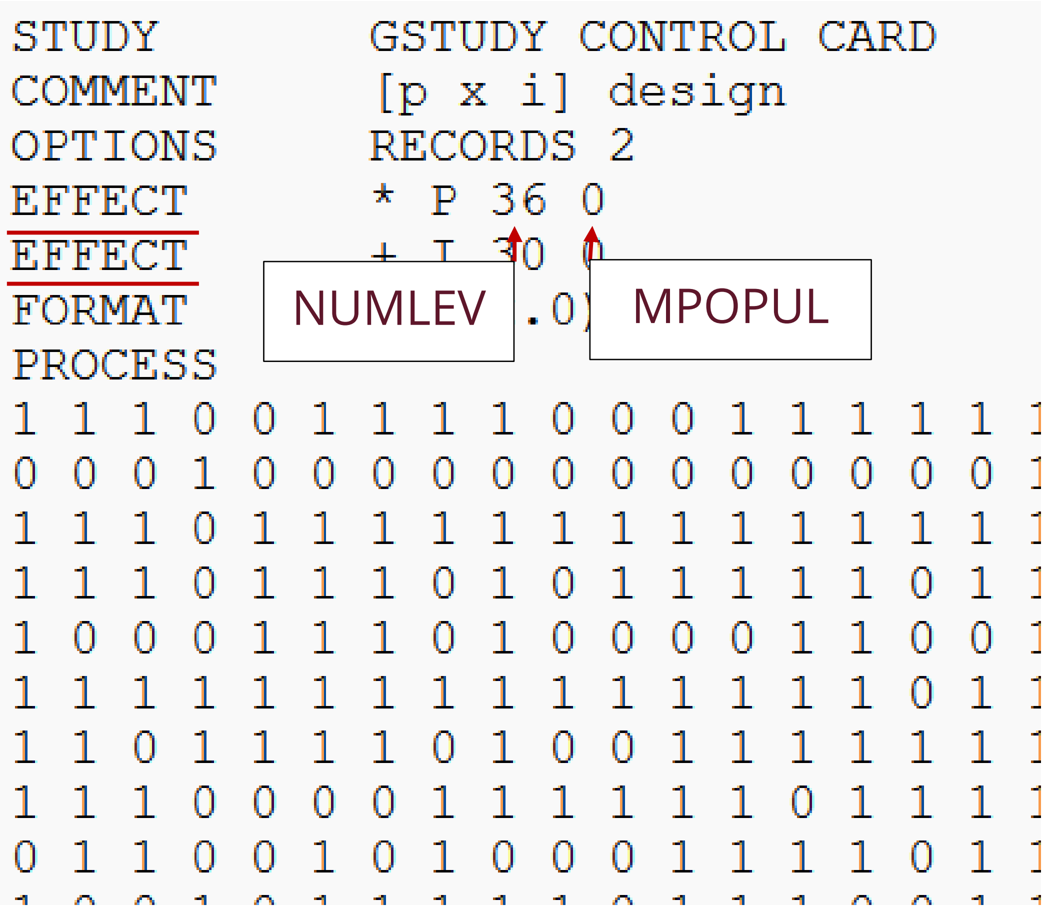
```
STUDY        GSTUDY CONTROL CARD
COMMENT      [p x i] design
OPTIONS      RECORDS 2
EFFECT       * P 36 0
EFFECT           0 0
FORMAT
PROCESS
1 1 1 0 0 1 1 1 1 0 0 0 1 1 1 1 1
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 0 1 1 1 0 1 0 1 1 1 1 1 0 1
1 0 0 0 1 1 1 0 1 0 0 0 0 1 1 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1
1 1 0 1 1 1 1 0 1 0 0 1 1 1 1 1 1
1 1 1 0 0 0 0 1 1 1 1 1 1 0 1 1 1
0 1 1 0 0 1 0 1 0 0 0 1 1 1 1 0 1
```

Creating Control Cards for GENOVA

Control cards for $p \times i$, G study

4. EFFECT

- NUMLEV
  - ✓ Sample size for the facet.
- NPOPUL (Optional)
  - ✓ Population or universe size for the facet.
  - ✓ Default = 0 (infinite universe).



```
STUDY        GSTUDY CONTROL CARD
COMMENT      [p x i] design
OPTIONS      RECORDS 2
EFFECT       * P 36 0
EFFECT       + I 30 0
FORMAT       .0)
PROCESS
1 1 1 0 0 1 1 1 1 0 0 0 1 1 1 1 1
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 0 1 1 1 0 1 0 1 1 1 1 1 0 1
1 0 0 0 1 1 1 0 1 0 0 0 0 1 1 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1
1 1 0 1 1 1 1 0 1 0 0 1 1 1 1 1 1
1 1 1 0 0 0 0 1 1 1 1 1 1 0 1 1 1
0 1 1 0 0 1 0 1 0 0 0 1 1 1 1 0 1
```

Creating Control Cards for GENOVA

Control cards for $p \times i$, G study

5. FORMAT

- Specify FORTRAN format for reading data.
- The parameter must be enclosed within parentheses.
- e.g., (30F2.0)
  - ✓ 30: Total of 30 values.
  - ✓ F2: Each value occupies 2 characters.
  - ✓ .0: No digits after the decimal point.

```
STUDY         GSTUDY CONTROL CARD
COMMENT       [p x i] design
OPTIONS       RECORDS 2
EFFECT        * P 36 0
EFFECT        + I 30 0
FORMAT        (30F2.0)
PROCESS
1 1 1 0 0 1 1 1 1 0 0 0 1 1 1 1 1 1
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 0 1 1 1 0 1 0 1 1 1 1 1 0 1 1
1 0 0 0 1 1 1 0 1 0 0 0 0 1 1 0 0 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1
1 1 0 1 1 1 1 0 1 0 0 1 1 1 1 1 1 1
1 1 1 0 0 0 0 1 1 1 1 1 1 0 1 1 1 1
0 1 1 0 0 1 0 1 0 0 0 1 1 1 1 0 1 1
```

Creating Control Cards for GENOVA

Control cards for $p \times i$, G study

6. PROCESS

- Data used for analysis

1) Enter data directly into the control card

2) Provide data through a separate file
   - ✓ An integer value specifying the logical unit number to be used for reading the input data (excluding 5 and 6).
   - ✓ The data file must end with a blank line (press Enter after the last record).

```
STUDY          GSTUDY CONTROL CARD
COMMENT        [p x i] design
OPTIONS        RECORDS 2
EFFECT         * P 36 0
EFFECT         + I 30 0
FORMAT         (30F2.0)
PROCESS        7
FINISH

```

Unit number

```
1 1 0 0 1 1 1 0 1 0 1 1 0 1 0 0 1 1 0 1 0 1 1 0 1 0 1 0 1 1
0 1 0 0 1 1 1 1 0 0 0 0 1 0 1 0 1 1 0 0 1 1 0 1 1 0 0 0 1 1
1 1 0 1 0 1 0 1 1 0 1 0 0 1 1 0 1 1 0 0 1 1 0 0 1 1 1 0 1 1
1 1 1 0 0 1 0 1 1 1 1 1 1 0 0 0 1 0 1 1 0 1 1 0 1 0 0 0 1 1
1 0 1 1 1 1 1 0 1 0 0 0 1 1 1 0 1 1 1 1 1 1 1 0 1 1 0 0 1 1
1 0 1 0 1 1 0 1 1 0 1 1 1 1 1 0 1 1 1 1 1 0 0 0 1 1 0 1 1 1
0 1 1 0 1 1 0 0 1 1 1 0 1 0 1 1 1 1 0 1 1 0 1 0 0 0 0 0 1 0
1 1 0 0 1 1 1 1 1 0 1 0 1 1 1 0 0 0 0 1 1 0 0 0 1 0 0 0 1 1
1 1 0 1 0 1 0 0 0 1 1 0 1 0 1 1 0 1 1 0 1 0 0 0 1 1 0 0 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1
1 1 0 0 0 1 1 0 0 0 1 1 1 1 1 1 1 1 0 1 1 0 0 1 1 1 1 0 1 1
1 1 0 0 1 1 1 1 0 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 0 1 1 1 0
1 1 1 0 0 1 0 1 0 0 1 1 1 1 1 0 1 1 0 0 1 1 0 0 1 0 1 0 1 1
1 1 0 1 1 1 1 1 1 0 1 0 1 1 1 0 1 0 0 1 1 1 0 0 1 0 1 0 1 0
0 1 0 0 0 1 1 0 0 0 1 0 1 0 1 1 1 0 0 1 1 0 0 0 1 0 1 0 1 0
1 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 1 1 1 1 0 1 0 1 0 0 1 1 1
0 0 1 0 0 0 1 1 1 0 1 0 1 1 1 1 1 0 0 1 1 1 0 1 1 0 1 0 1 1

```

Creating Control Cards for GENOVA

Control cards for $p \times i$, D study

1. DSTUDY

- A user-defined alphanumeric heading that will appear on D study output page.

2. DEFFECT

- $:
  - ✓ Indicates the object of measurement in the D study.
- DFACET:
  - ✓ Facet for D study.
  - ✓ Letters (A~Z) and colon (:).

```
STUDY        GSTUDY CONTROL CARD
COMMENT      [p x i] design
OPTIONS      RECORDS 2
EFFECT       * P 36 0
EFFECT       + I 30 0
FORMAT       (30F2.0)
PROCESS      7
COMMENT      D STUDY CONTROL CARD
DSTUDY       P X I DESIGN
DEFFECT      $ P 36 / 0
DEFFECT        I 5 10 15 20 30 40 50 / 0
ENDDSTUDY
FINISH
```

$

DFACET

Creating Control Cards for GENOVA

Control cards for $p \times i$, D study

2. DEFFECT

- ISAMPD: (Optional)
  - ✓ Values indicating the sample size.
  - ✓ GENOVA automatically repeats the last sample size for a facet if fewer values are given.
  - ✓ Default = sample size from the G study

    0 = Infinite
- IUNIVD: (Optional)
  - ✓ Population or universe size for the facet.
  - ✓ Must have '/' when a IUNIVD value is specified
  - ✓ Default = size from the G study

    0 = Infinite
- e.g., D study results for 36-5, 36-10, 36-15. 36-20, 36-30, 36-40,36-50

```
STUDY        GSTUDY CONTROL CARD
COMMENT      [p x i] design
OPTIONS      RECORDS 2
EFFECT       * P 36 0
EFFECT       + I 30 0
FORMAT       (30F2.0)
PROCESS      7
COMMENT      D STUDY CONTROL CARD
DSTUDY       P X I DESIGN
DEFFECT      $ P 36 / 0
DEFFECT        I 5 10 15 20 30 40 50 / 0
ENDDSTUDY
FINISH
```

ISAMPD

IUNIVD

Creating Control Cards for GENOVA

Control cards for $p \times i$, D study

3. ENDDSTUDY

- Indicate the end of control card.

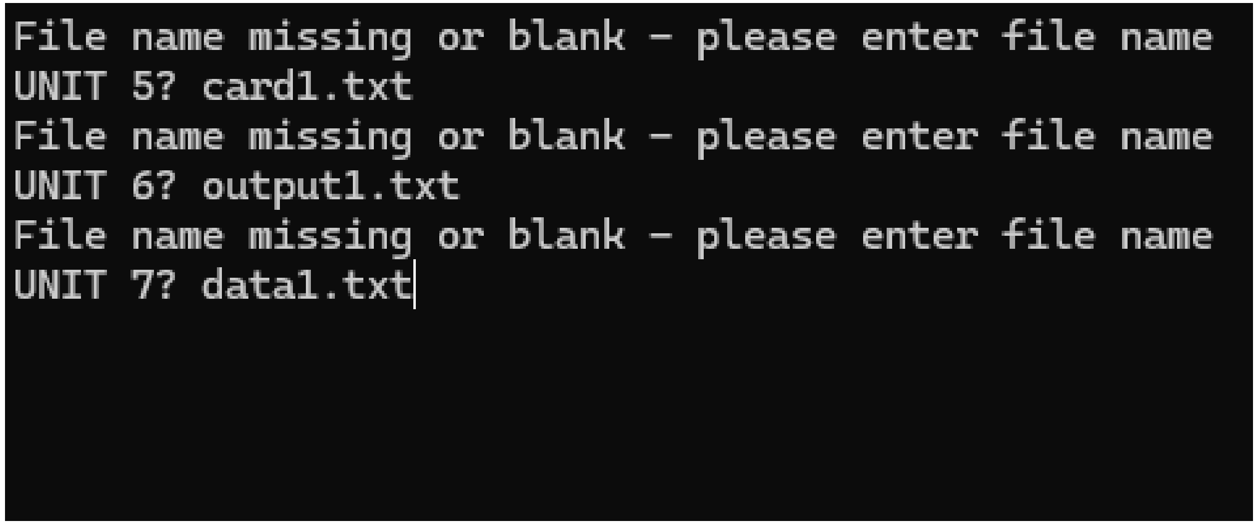
```
STUDY        GSTUDY CONTROL CARD
COMMENT      [p x i] design
OPTIONS      RECORDS 2
EFFECT       * P 36 0
EFFECT       + I 30 0
FORMAT       (30F2.0)
PROCESS      7
COMMENT      D STUDY CONTROL CARD
DSTUDY       P X I DESIGN
DEFFECT      $ P 36 / 0
DEFFECT        I 5 10 15 20 30 40 50 / 0
ENDDSTUDY
FINISH
```

Program Execution: GENOVA

Lauch the GENOVA program

- **Execute the file ‘genova36.exe’**

1. **UNIT 5:**
   - ✓ Enter the name of the control card.
   - ✓ The control card must be located in the same folder as the ‘Genova36.exe’ file.
2. **UNIT 6:**
   - ✓ Enter the name of the output file.
   - ✓ An output file will be generated following successful program execution.
3. **UNIT 7: (Optional)**
   - ✓ Enter the name of the data file.
   - ✓ The unit number is displayed according to the value entered in the PROCESS section of the control card.



Interpreting the GENOVA Output

Output file

- Note: The output file shown in this presentation includes only selected pages from the full output.

```
GENOVA  VERSION 3.1                                                                                                          PAGE       1
                                                  CONTROL CARD INPUT LISTING


COLUMN   1111111111222222222233333333334444444444555555555566666666667777777778
12345678901234567890123456789012345678901234567890123456789012345678901234567890
STUDY       GSTUDY CONTROL CARD
COMMENT     [p x i] design
OPTIONS     RECORDS 2
EFFECT      * P 36 0
EFFECT      + I 30 0
FORMAT      (30F2.0)
PROCESS

GENOVA  VERSION 3.1                                                                                                          PAGE       2
       G STUDY                                       GSTUDY CONTROL CARD

                                          EXPANDED MAIN AND INTERACTION EFFECT TABLE

      (** = INFINITE)    P         I                  TOTAL   DEGREES
      SAMPLE SIZE       36        30        PRIMARY  NUMBER     OF
      UNIVERSE SIZE   ****      ****        INDICES  INDICES  FREEDOM
     ************************************
     *              *         *         *
     *  P           *    1    *    0    *       1        1         35
     *  I           *    0    *    1    *       1        1         29
     *              *         *         *
     ************************************
     *              *         *         *
     *  PI          *    1    *    1    *       2        2       1015
     *              *         *         *
     ************************************
```

Interpreting the GENOVA Output

Output file

- Input data / Mean of input data

```
  GENOVA  VERSION 3.1                                                                                PAGE        3
         G STUDY                                         GSTUDY CONTROL CARD

                                                INPUT RECORD LISTING WITH RECORD MEANS


RECORD #      1      1.00000     1.00000     1.00000     0.00000     0.00000     1.00000     1.00000     1.00000
                     1.00000     0.00000     0.00000     0.00000     1.00000     1.00000     1.00000     1.00000
                     1.00000     1.00000     1.00000     1.00000     1.00000     1.00000     0.00000     0.00000
                     1.00000     0.00000     1.00000     0.00000     1.00000     1.00000     0.70000

RECORD #      2      0.00000     0.00000     0.00000     1.00000     0.00000     0.00000     0.00000     0.00000
                     0.00000     0.00000     0.00000     0.00000     0.00000     0.00000     0.00000     0.00000
                     0.00000     1.00000     1.00000     1.00000     0.00000     0.00000     1.00000     1.00000
                     0.00000     0.00000     1.00000     0.00000     1.00000     1.00000     0.30000

RECORD #     35      1.00000     1.00000     1.00000     0.00000     1.00000     1.00000     1.00000     1.00000
                     1.00000     0.00000     0.00000     1.00000     1.00000     0.00000     1.00000     1.00000
                     1.00000     1.00000     1.00000     1.00000     1.00000     0.00000     1.00000     0.00000
                     1.00000     0.00000     0.00000     1.00000     1.00000     1.00000     0.73333

RECORD #     36      0.00000     0.00000     1.00000     0.00000     0.00000     0.00000     1.00000     1.00000
                     1.00000     0.00000     1.00000     0.00000     1.00000     1.00000     1.00000     1.00000
                     1.00000     0.00000     0.00000     1.00000     1.00000     1.00000     0.00000     1.00000
                     1.00000     0.00000     1.00000     0.00000     1.00000     1.00000     0.60000
```

Interpreting the GENOVA Output

Output file

- Grand mean

```
GENOVA   VERSION 3.1                                                                                      PAGE        4
       G STUDY                                             GSTUDY CONTROL CARD

                                                             CELL MEAN SCORES


*******************************************************************************************************************************

                                           *** GRAND MEAN =             0.6129630 ***

*******************************************************************************************************************************

 MEAN SCORES FOR EFFECT:  I                 SUBSCRIPT NOTATION:   (I)


      (1)  =           0.833333        (2)  =           0.777778        (3)  =           0.555556        (4)  =           0.333333
      (5)  =           0.500000        (6)  =           0.805556        (7)  =           0.638889        (8)  =           0.694444
      (9)  =           0.722222       (10)  =           0.333333       (11)  =           0.611111       (12)  =           0.611111
     (13)  =           0.777778       (14)  =           0.638889       (15)  =           0.750000       (16)  =           0.416667
     (17)  =           0.777778       (18)  =           0.833333       (19)  =           0.472222       (20)  =           0.694444
     (21)  =           0.666667       (22)  =           0.527778       (23)  =           0.361111       (24)  =           0.277778
     (25)  =           0.833333       (26)  =           0.361111       (27)  =           0.583333       (28)  =           0.250000
     (29)  =           0.972222       (30)  =           0.777778
```

Interpreting the GENOVA Output

Output file

- Variance components

```
GENOVA   VERSION 3.1                                                                          PAGE        6
       G STUDY                                          GSTUDY CONTROL CARD

                                                          G STUDY RESULTS

      (** = INFINITE)     P         I
      SAMPLE SIZE        36        30
      UNIVERSE SIZE    ****      ****                                               QFM = QUADRATIC FORM
     -------------------------------------------------------------------------------------
                                  M O D E L   V A R I A N C E   C O M P O N E N T S
                  DEGREES         - - - - -   - - - - - - - -   - - - - - - - - - -
                    OF             USING            USING EMS            STANDARD
      EFFECT     FREEDOM         ALGORITHM          EQUATIONS              ERROR
     -------------------------------------------------------------------------------------
      P                35          0.0146196          0.0146196          0.0048985
      I                29          0.0318482          0.0318482          0.0094500
     -------------------------------------------------------------------------------------
      PI             1015          0.1924174          0.1924174          0.0085329
     -------------------------------------------------------------------------------------
        NOTE: THE "ALGORITHM" AND "EMS" ESTIMATED VARIANCE COMPONENTS WILL BE
              IDENTICAL IF THERE ARE NO NEGATIVE ESTIMATES
```

1. Estimated variance components (EVC)
2. EVC recalculated by replacing any negative estimates with zero

* 1=2 if there are no negative estimates

3. Standard error of EVC

Interpreting the GENOVA Output

Output file

- Variance components
- δ
- Δ
- $E\rho^2$
- Φ

```
GENOVA   VERSION 3.1                                                                                       PAGE      10
       D STUDY                                                P X I DESIGN

                                                     D STUDY DESIGN NUMBER 001-001

    OBJECT OF MEASUREMENT :         P                        FACETS :        I
  G STUDY POPULATION SIZE : INFINITE      G STUDY UNIVERSE SIZES : INFINITE
  D STUDY POPULATION SIZE : INFINITE      D STUDY UNIVERSE SIZES : INFINITE
      D STUDY SAMPLE SIZE :    36           D STUDY SAMPLE SIZES :      5
-----------------------------------------------------------------------------------------------------------------------------------
                          VARIANCE COMPONENTS IN TERMS OF                            VARIANCE COMPONENTS IN TERMS OF
                 G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES             D STUDY UNIVERSE (OF GENERALIZATION) SIZES
            ---------------------------------------------------------- -----------------------------------------------------------
                                                  VARIANCE COMPONENTS                                          VARIANCE COMPONENTS
              VARIANCE     FINITE   D STUDY         FOR MEAN SCORES       VARIANCE     FINITE   D STUDY          FOR MEAN SCORES
             COMPONENTS   UNIVERSE SAMPLING -------------------------- COMPONENTS   UNIVERSE SAMPLING --------------------------
             FOR SINGLE     COR-      FRE-                   STANDARD   FOR SINGLE     COR-      FRE-                     STANDARD
EFFECT      OBSERVATIONS  RECTIONS QUENCIES    ESTIMATES       ERRORS  OBSERVATIONS  RECTIONS QUENCIES    ESTIMATES       ERRORS
-----------------------------------------------------------------------------------------------------------------------------------
P               0.01462     1.0000        1    0.01462       0.00490      0.01462     1.0000        1     0.01462       0.00490
I               0.03185     1.0000        5    0.00637       0.00189      0.03185     1.0000        5     0.00637       0.00189
PI              0.19242     1.0000        5    0.03848       0.00171      0.19242     1.0000        5     0.03848       0.00171
-----------------------------------------------------------------------------------------------------------------------------------
           QFM = QUADRATIC FORM

-----------------------------------------------------------------------------------------------------------------------------------


                                                         STANDARD
                                               STANDARD  ERROR OF
                                     VARIANCE DEVIATION  VARIANCE

                  UNIVERSE SCORE      0.01462   0.12091   0.00490
        EXPECTED OBSERVED SCORE      0.05310   0.23044   0.00509
              LOWER CASE DELTA      0.03848   0.19617   0.00171     GENERALIZABILITY COEFFICIENT =   0.27531   ( 0.37989)
              UPPER CASE DELTA      0.04485   0.21179   0.00251                              PHI =   0.24582   ( 0.32594)
                            MEAN      0.00784   0.08857
-----------------------------------------------------------------------------------------------------------------------------------
      NOTE:  SIGNAL/NOISE RATIOS ARE IN PARENTHESES
```

Interpreting the GENOVA Output

Output file

- Summary of D study results

GENOVA VERSION 3.1 PAGE 24
D STUDY P X I DESIGN

SUMMARY OF D STUDY RESULTS FOR SET OF CONTROL CARDS NO. 001

D STUDY DESIGN NO	INDEX= UNIV.=	SAMPLE SIZES $P INF.	SAMPLE SIZES I INF.	VARIANCES UNIVERSE SCORE	VARIANCES EXPECTED OBSERVED SCORE	VARIANCES LOWER CASE DELTA	VARIANCES UPPER CASE DELTA	VARIANCES MEAN	GEN. COEF.	PHI
001-001		36	5	0.01462	0.05310	0.03848	0.04485	0.00784	0.27531	0.24582
001-002		36	10	0.01462	0.03386	0.01924	0.02243	0.00413	0.43175	0.39463
001-003		36	15	0.01462	0.02745	0.01283	0.01495	0.00289	0.53264	0.49440
001-004		36	20	0.01462	0.02424	0.00962	0.01121	0.00227	0.60311	0.56593
001-005		36	30	0.01462	0.02103	0.00641	0.00748	0.00165	0.69506	0.66167
001-006		36	40	0.01462	0.01943	0.00481	0.00561	0.00134	0.75242	0.72280
001-007		36	50	0.01462	0.01847	0.00385	0.00449	0.00115	0.79162	0.76523

Creating Control Cards for GENOVA

Control cards for $p \times (r{:}\, i)$, G & D study

1. EFFECT

- EFFECT cards must list facets in order from the slowest-moving to fastest-moving in nested design.

2. DEFFECT

- If there are specific conditions to be analyzed in the D study, they can be stated in the DEFFECT card to produce results.
- e.g., Item – Rater: 1-12, 2-6, 3-4, 4-3, 5-2, 6-2

```
STUDY        P X (r:i) DESIGN
COMMENT
OPTIONS      RECORDS 2
EFFECT       * P 10 0
EFFECT       + I 3 0
EFFECT       + R:I 4 0
FORMAT       (12F2.0)
PROCESS      7
COMMENT      D STUDY CONTROL CARDS
DSTUDY       P X (R:I)
DEFFECT      $ P
DEFFECT      I 1 2 3 4 5 6
DEFFECT      R:I 12 6 4 3 2 2
ENDDSTUDY
FINISH
```

Table Data Structure for the $p \times (r{:}\, i)$ Design

	$i1$				$i2$				$i3$			
Person	$r1$	$r2$	$r3$	$r4$	r5	$r6$	$r7$	$r8$	$r9$	$r10$	$r11$	$r12$
1	5	6	5	5	5	3	4	5	6	7	3	3
2	9	3	7	7	7	5	5	5	7	7	5	2
3	3	4	3	3	5	3	3	5	6	5	1	6
4	7	5	5	3	3	1	4	3	5	3	3	5
5	9	2	9	7	7	7	3	7	2	7	5	3

Interpreting the GENOVA Output

Output file

- D study results for 3 items and 4 raters

```
 GENOVA  VERSION 3.1                                                                                         PAGE      14
        D STUDY                                                 P X (R:I)

                                                     D STUDY DESIGN NUMBER 001-003

     OBJECT OF MEASUREMENT :      P                        FACETS :      I          R:I
   G STUDY POPULATION SIZE : INFINITE     G STUDY UNIVERSE SIZES : INFINITE     INFINITE
   D STUDY POPULATION SIZE : INFINITE     D STUDY UNIVERSE SIZES : INFINITE     INFINITE
       D STUDY SAMPLE SIZE :     10          D STUDY SAMPLE SIZES :     3            4
-------------------------------------------------------------------------------------------------------------------------------
                         VARIANCE COMPONENTS IN TERMS OF                              VARIANCE COMPONENTS IN TERMS OF
                G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES              D STUDY UNIVERSE (OF GENERALIZATION) SIZES
            ------------------------------------------------------- -------------------------------------------------------
                                                VARIANCE COMPONENTS                                          VARIANCE COMPONENTS
              VARIANCE     FINITE   D STUDY       FOR MEAN SCORES        VARIANCE     FINITE   D STUDY       FOR MEAN SCORES
             COMPONENTS   UNIVERSE SAMPLING  -------------------------- COMPONENTS   UNIVERSE SAMPLING  --------------------------
             FOR SINGLE     COR-     FRE-                      STANDARD FOR SINGLE     COR-     FRE-                      STANDARD
EFFECT      OBSERVATIONS  RECTIONS QUENCIES    ESTIMATES        ERRORS  OBSERVATIONS  RECTIONS QUENCIES    ESTIMATES        ERRORS
-------------------------------------------------------------------------------------------------------------------------------
P               0.47315     1.0000        1     0.47315        0.38558    0.47315     1.0000        1     0.47315        0.38558
I               0.32515     1.0000        3     0.10838        0.14600    0.32515     1.0000        3     0.10838        0.14600
R:I             0.64753     1.0000       12     0.05396        0.03162    0.64753     1.0000       12     0.05396        0.03162
PI              0.55957     1.0000        3     0.18652        0.12554    0.55957     1.0000        3     0.18652        0.12554
PR:I            2.38025     1.0000       12     0.19835        0.03079    2.38025     1.0000       12     0.19835        0.03079
-------------------------------------------------------------------------------------------------------------------------------
           QFM = QUADRATIC FORM

-------------------------------------------------------------------------------------------------------------------------------
```

Interpreting the GENOVA Output

Output file

- Summary of D study results

GENOVA VERSION 3.1 PAGE 22

D STUDY P X (R:I)

SUMMARY OF D STUDY RESULTS FOR SET OF CONTROL CARDS NO. 001

		SAMPLE SIZES			VARIANCES						
D STUDY DESIGN NO	INDEX= UNIV.=	$P INF.	I INF.	R INF.	UNIVERSE SCORE	EXPECTED OBSERVED SCORE	LOWER CASE DELTA	UPPER CASE DELTA	MEAN	GEN. COEF.	PHI
001-001		10	1	12	0.47315	1.23107	0.75792	1.13704	0.50222	0.38434	0.29385
001-002		10	2	6	0.47315	0.95129	0.47814	0.69468	0.31167	0.49738	0.40515
001-003		10	3	4	0.47315	0.85802	0.38488	0.54722	0.24815	0.55144	0.46370
001-004		10	4	3	0.47315	0.81139	0.33825	0.47350	0.21639	0.58313	0.49982
001-005		10	5	2	0.47315	0.82309	0.34994	0.47972	0.21209	0.57485	0.49655
001-006		10	6	2	0.47315	0.76476	0.29162	0.39977	0.18463	0.61869	0.54203