Analysis of Variance (ANOVA)

Hirotaka Onishi, MD, MHPE, PhD

Dept. of International Cooperation for Medical Education,
International Research Center for Medical Education,
Graduate School of Medicine, The University of Tokyo

What Is H_0 Meaning?

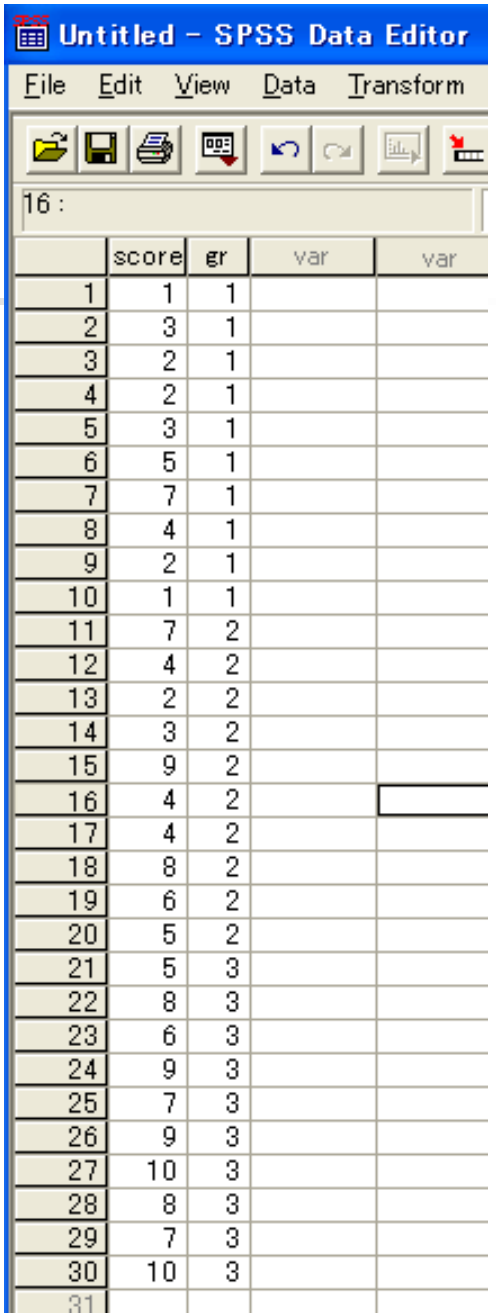
- H_0: $\bar{x}_1 = \bar{x}_2 = \bar{x}_3$
- Even if H_0 is rejected, it does not automatically mean $x_1 < x_2$ or $x_2 < x_3$
- To prove the difference between specific two variables, post-hoc tests will be used.
  - Scheffé test: most stringent
  - Bonferroni: α is divided by the number of tests.
  - Tukey's Honestly Significant Difference (HSD): numbers of cases must be the same

Differentiate Within-subject from Between-subject Data

- Within-subject data
  - Longitudinally (repeatedly) collected data
  - Multiple items in a questionnaire
- Between-subject
  - Different groups of subjects
    - Demographic data: Gender, age group…
    - Study vs control groups

Data Format

control	APN phone	APN ph + visits
1	7	5
3	4	8
2	2	6
2	3	9
3	9	7
5	4	9
7	4	10
4	8	8
2	6	7
1	5	10
$\bar{x}$ = 3.0	$\bar{x}$ = 5.2	$\bar{x}$ = 7.9
sd = 1.89	sd = 2.25	sd = 1.66



Data Format

- In SPSS or Excel
  - Within-subject data are put in parallel → You have to define the group to conduct the analysis
  - Between-subject data are put in tandem

Nation	HiB	HiY	HiR
1	10	6	19
1	12	11	16
1	9	8	19
1	10	9	19
2	8	7	20
2	10	7	20
2	8	9	16
2	9	9	19
3	8	15	9
3	7	19	6
3	11	14	6
3	7	18	8

Select Post-hoc Test

When three or more options are

- Within-subject
  - Sidak is often selected
- Between-subject
  - Tukey is one of the most frequently used post-hoc test

Sphericity

- If within-subject data has three or more options, sphericity might be a problem.
  - When “test of homogeneity of variances” is significant, 3 Steps will be taken:
  - (1) sphericity supposed
  - (2) lowest, and
  - (3) Greenhouse-Geisser’s modification

Interaction

	Male	Female	
Young	5	10	15
Old	10	5	15
	15	15	

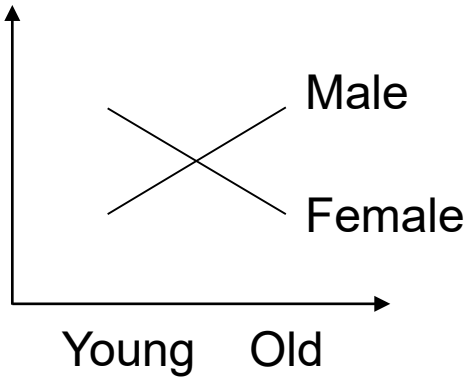
Main effect for Age

Main effect for Sex

H_0: $\mu_{young} = \mu_{old}$
H_1: H_0 is false

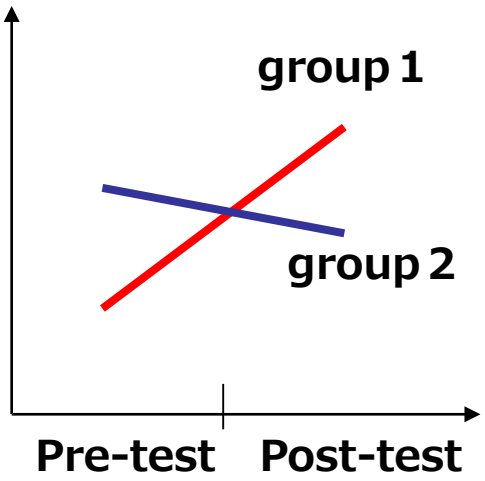
H_0: $\mu_{male} = \mu_{female}$
H_1: H_0 is false

H_0: There is no interaction
H_1: H_0 is false

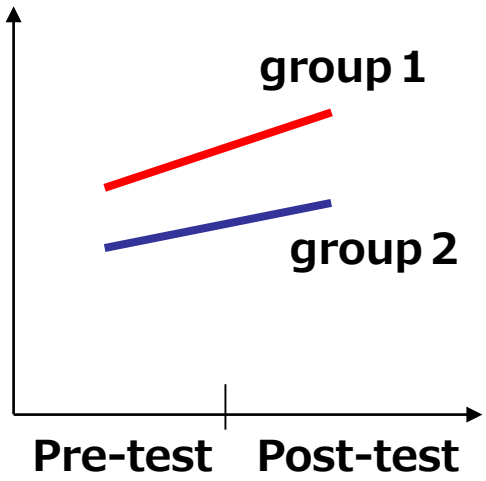


2way, 3way ANOVA

- More than one factors are compared simultaneously.
- Both within-subject and between subject data are combined for one analysis.
- Interaction is another issue for 2way ANOVA or more.



With interaction



Without interaction

検定・分析の種類別の代表的な効果量の指標と大きさの目安

使用される検定（分析）	対象と注意	効果量の指標	効果量の目安 小 (Small)	中 (Medium)	大 (Large)
相関分析		r	.10	.30	.50
重回帰分析		R^2	.02	.13	.26
		f^2	.02	.15	.35
t 検定（t-test）	r と d は対応ありの場合は注意	r	.10	.30	.50
		d	.20	.50	.80
一元配置分散分析 (One-way ANOVA)	全体の検定	η^2	.01	.06	.14
		partial η^2	-	-	-
		ω^2	.01	.09	.25
		f	.10	.25	.40
	多重比較	r	.10	.30	.50
		d	.20	.50	.80
二元配置分散分析 (Two-way ANOVA) 多元配置分散分析* (Multi-way ANOVA) *三元配置以上の分散分析	主効果	η^2	.01	.06	.14
		partial η^2	-	-	-
		ω^2	.01	.09	.25
	交互作用	η^2	.01	.06	.14
		partial η^2	-	-	-
		ω^2	.01	.09	.25
	多重比較	r	.10	.30	.50
		d	.20	.50	.80

水本，竹内．研究論文における効果量の報告のために．英語教育研究 31, 57-66, 2008

Exercise 1

- Open 2wayANOVA.xlsx
- Seek for the relationship (1) age (2) # of delivery and (3) sex of the baby (1 male, 0 female)

対馬栄輝．SPSSで学ぶ医療系データ解析：分析内容の理解と手順解説、バランスのとれた医療統計入門．pp166-174

Exercise 2

- Open repANOVA.xlsx
- 23 young subjects measured height in 3 consecutive years.
- Are they growing?
- Change in the data format might be needed.

対馬栄輝．SPSSで学ぶ医療系データ解析：分析内容の理解と手順解説、バランスのとれた医療統計入門．pp182-188

Exercise 3

- Open water learning.xlsx
- 9 divers participated in the experiments of learning words in 8 conditions
  - 1st digit: reproduction–1, reaffirmation–2
    - Reproduction: recall what was experienced as it was
    - Reaffirmation: confirm that what is being questioned is what was experienced.
  - 2nd digit: learning on the ground–1, in the water–2
  - 3rd digit: testing on the ground–1, in the water–2

山内光哉．心理・教育のための分散分析と多重比較－エクセル・SPSS解説付き．pp199-212

Exercise 4

- Open MixedANOVA.xlsx
- 12 workers from 3 countries (Japan – 1, Korea – 2, US – 3) participated in the physical exercise.
- Conditions
  - Color lighting: blue, red and yellow
  - Temperature: high, normal and low

竹原卓真．SPSSのススメ(2) 3 要因の分散分析をすべてカバー．pp198-302