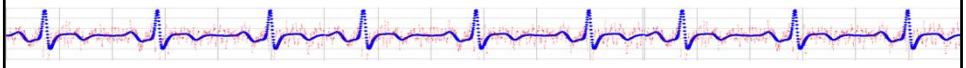
Empirical Research Methods in Information Science

IS 4800 / CS 6350

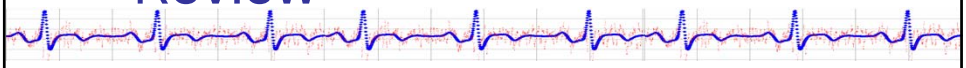


Lecture 23

Multi-factor Between-Subjects Designs

1

Review



- Within-subjects design
  - What is it?
  - How do you do one?
  - Criteria for choosing over between-subjects design?
  - Power analysis?
  - Analysis?

2

Types of Study Designs

- Qualitative
  - Ethnography
- Quantitative
  - Descriptive
  - Correlational
  - Demonstrative
  - Experimental
    - Between-subjects
      - Single factor, two-level
    - Within-subjects
      - Single factor, two-level

3

Accompanying Statistics

- Experimental
  - Between-subjects
    - Single factor, N-level (for N>2)
      - One-way Analysis of Variance (ANOVA)
    - Two factor, two-level (or more!)
      - Factorial Analysis of Variance
      - AKA N-way Analysis of Variance (for N IVs)
      - AKA N-factor ANOVA
  - Within-subjects
    - Repeated-measures ANOVA (not discussed today)
      - AKA within-subjects ANOVA

4

Basic Logic of ANOVA

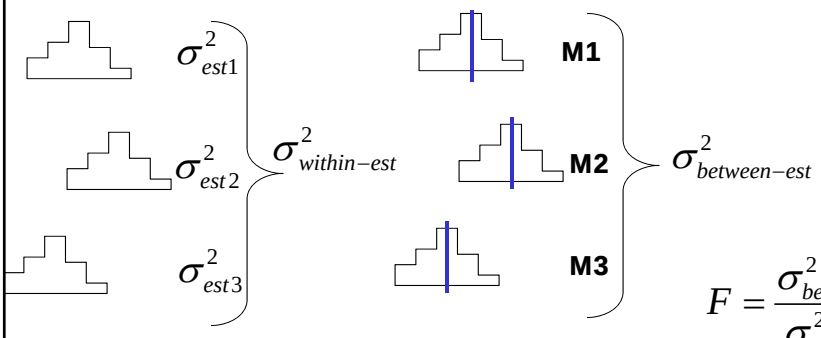
- Null hypothesis
  - Means of all groups are equal.
- Test: do the means differ more than expected given the null hypothesis?
- Terminology
  - Group = Condition = Cell

5

One-Way ANOVA – Assuming Null Hypothesis is True...

Within-Group Estimate Of Population Variance

Between-Group Estimate Of Population Variance



$$F = \frac{\sigma^2_{between-est}}{\sigma^2_{within-est}}$$

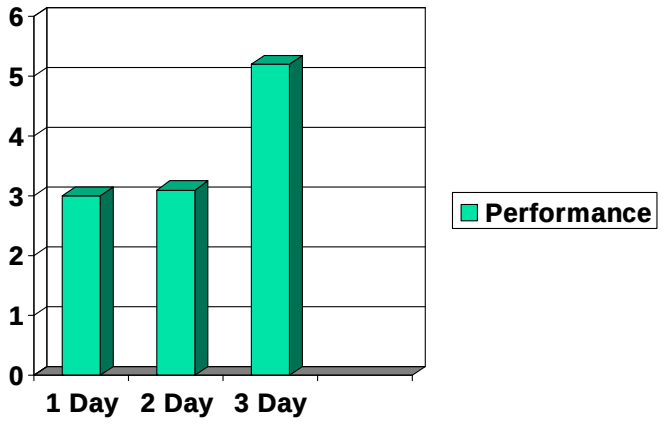
7

One-way ANOVA in R

	SID	TrainingDays	Performance
1	1	1.00	5.00
2	1	2.00	4.00
3	1	3.00	4.00
4	2	1.00	2.50
5	2	2.00	3.00
6	2	3.00	6.50
7	3	1.00	2.00
8	3	2.00	1.50
9	3	3.00	6.00
10	4	1.00	3.00
11	4	2.00	4.50
12	4	3.00	4.50
13	5	1.00	2.00
14	5	2.00	2.00
15	5	3.00	4.00
16	6	1.00	2.00
17	6	2.00	3.00
18	6	3.00	4.00

8

Data



9

One-way ANOVA in R

```
> one$TrainingDays <- factor(one$TrainingDays)
> res <- aov(one$Performance ~ one$TrainingDays)
> summary(res)
                 Df Sum Sq Mean Sq F value   Pr(>F)
one$TrainingDays  2 24.812  12.406  9.4417 0.001188 **
Residuals        21 27.594   1.314
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’
0.1 ‘ ’ 1
```

F(2,21)=9.44, p<.05

10

Post hoc analysis

- Once the ANOVA indicates there is a significant difference (“omnibus” test), you do either
  - Planned comparisons, or
  - Post hoc tests
- to determine which pairwise comparisons are significantly different
- There are many post hoc tests (B&A 446)
  - Sheffe, Dunnett, Tukey, etc.
    - Very conservative

12

Post-hoc tests in R
Tukey HSD

```
> res <- aov(one$Performance ~ one$TrainingDays)
> TukeyHSD(res)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = one$Performance ~ one$TrainingDays)

$`one$TrainingDays`
      diff        lwr      upr     p adj
2-1 0.0625 -1.3821563 1.507156 0.9934676
3-1 2.1875  0.7428437 3.632156 0.0027729
3-2 2.1250  0.6803437 3.569656 0.0035777
```

13

Publication format

- The overall ANOVA was significant, $F(2,21)=9.44$, $p<.05$, indicating significant differences among the three study treatments.
- Tukey HSD post-hoc tests (at .05 significance) indicated significant differences between 3-day training and the other conditions, but not between 1-day and 2-day training.

14

Example Paper

- Thank you – I did not see that: In-car, speech-based information systems for older adults.
- Critique?

16

Factorial Designs

- Two or more nominal independent variables, each with two or more levels, and a numeric dependent variable.
- Factorial ANOVA teases apart the contribution of each variable separately.
- For N IVs, aka “N-way” ANOVA

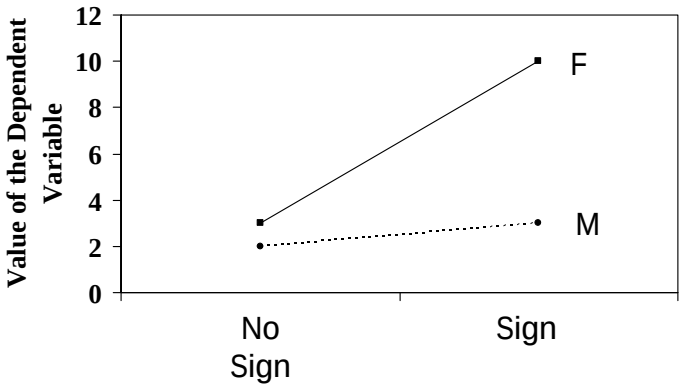
17

Factorial Designs

- Adding a second independent variable to a single-factor design results in a FACTORIAL DESIGN
- Two components can be assessed
  - The *MAIN EFFECT* of each independent variable
    - The separate effect of each independent variable
    - Analogous to separate experiments involving those variables
  - The *INTERACTION* between independent variables
    - When the effect of one independent variable changes over levels of a second
    - Also – when the effect of one variable depends on the level of the other variable.

18

Example of An Interaction - Student Center Sign - 2 Genders x 2 Sign Conditions



19

Two-way ANOVA in R

Book	Instructor	Knowledge
1	1	1.5
2	2	2
1	1	2
1	1	0.5
2	2	2
2	2	2
1	1	2
2	2	2
1	1	1
1	2	6.5

20

Two-way ANOVA in R

```
> out <- aov(Knowledge ~ Book * Instructor,  data=d)
>
> summary(out)
               Df  Sum Sq Mean Sq F value    Pr(>F)
Book            1   1.477   1.477   1.161    0.2941
Instructor      1   0.022   0.022   0.017    0.8975
Book:Instructor  1 123.450 123.450  97.032 4.073e-09 ***
Residuals      20  25.445   1.272
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘
’ 1
```

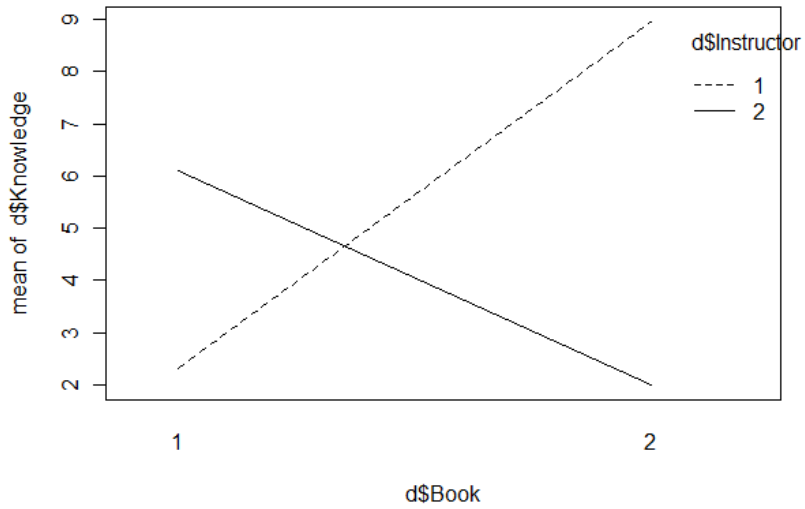
23

Interaction Plot

```
# x.factor = X-axis
# trace.factor = separate plot lines factor
# response = response variable

> interaction.plot(x.factor=d$Book,
          trace.factor=d$Instructor,
          response=d$Knowledge)
```

24



25

Degrees of Freedom

- df for between-group variance estimates for main effects
  - Number of levels – 1
- df for between-group variance estimates for interaction effect
  - Total num cells – df for both main effects – 1
  - e.g. 2x2 => 4 – (1+1) – 1 = 1
- df for within-group variance estimate
  - Sum of df for each cell = N – num cells
- Report: "F(*bet-group, within-group*)=*F, Sig.*"

27

Publication format

```
> summary(out)
                Df  Sum Sq Mean Sq F value    Pr(>F)
Book             1   1.477   1.477   1.161    0.2941
Instructor       1   0.022   0.022   0.017    0.8975
Book:Instructor  1 123.450 123.450  97.032 4.073e-09 ***
Residuals       20  25.445   1.272
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
' 1
```

F(1,20)=97.0, p<.05.

There is a significant interaction effect of Book and Instructor on Knowledge gain.

28

Reporting rule

- IF you have a significant interaction
- THEN
  - In general: <u>do not</u> report main effects, even if significant
  - Else: must look at patterns of means in cells to determine whether to report main effects or not.

30

Results?

	Sig.
TrainingDays	0.34
Trainer	0.12
TrainingDays * Trainer	0.41

→ n.s.

31

Results?

	Sig.
TrainingDays	0.34
Trainer	0.12
TrainingDays * Trainer	0.02

Significant interaction between TrainingDays And Trainer, $F(2,22)=.584$, $p<.05$

32

Results?

	Sig.
TrainingDays	0.34
Trainer	0.02
TrainingDays * Trainer	0.41

Main effect of Trainer, $F(1,22)=.001$, $p<.05$

33

Results?

	Sig.
TrainingDays	0.04
Trainer	0.12
TrainingDays * Trainer	0.01

Significant interaction between TrainingDays And Trainer, F(2,22)=.584, p<.05

Do not report TrainingDays as significant

34

Results?

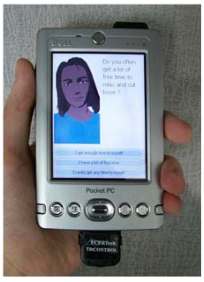
	Sig.
TrainingDays	0.04
Trainer	0.02
TrainingDays * Trainer	0.41

Main effects for both TrainingDays, F(2,22)=7.20, p<.05, and Trainer, F(1,22)=.001, p<.05

35

"Factorial Design"

- Not all cells in your design need to be tested
  - But if they are, it is a "full factorial design", and you do a "full factorial ANOVA"



	Real-Time	Retrospective
Agent	√	√
Text	√	X

36

Higher-Order Factorial Designs

- More than two independent variables are included in a *higher-order factorial design*
  - As factors are added, the complexity of the experimental design increases
    - The number of possible main effects and interactions increases
    - The number of subjects required increases exponentially
    - The volume of materials and amount of time needed to complete the experiment increases

37

ANOVA Assumptions

- Populations are normal
- Populations have equal variances

- More or less..

38

Power Analysis & Multi-factorial designs

- 'N' computed for your criteria for a between-subjects design is for *each cell* of your experimental design
- A two-factor x two-level design has *four cells*
- B&A: Need at least 5 Ss per cell
- See Aron Table 10-16 (pg 410)
  - Example: medium effect size, 2x2, for all effects, requires 33x4 = 132 Ss!

39

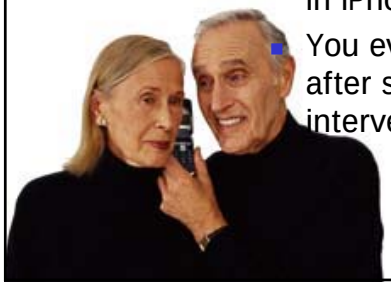
Design Example

- You want to evaluate which of 3 games leads to greatest satisfaction.

40

Example – Best Design?

- You've just developed the "Matchmaker" – a smart phone app that beeps when you are in the vicinity of a compatible person who is also carrying a Matchmaker phone.
- You want to see if you get different results in iPhones vs. Androids.
- You evaluate the life satisfaction of usres after six months of use compared to a non-intervention control group.



41

Group Exercise

- For each problem, write
  1. Kind of study design
  2. Kind of analysis
  3. The 4 populations
  4. Research & Null hypotheses (Means & English)
  5. Test criteria
  6. Plot results
  7. Test results
     - English & Publication format (requires df)
  8. Implications

42

Homework

- Study Presentations II

43