Statistical Analysis of Factorial Designs

⌘ Research Hypotheses for Factorial Designs
⌘ The F-tests of a Factorial ANOVA
⌘ Using LSD to describe the pattern of an interaction

RH: for Factorial Designs

Research hypotheses for factorial designs may include

- RH: for main effects
  - involve the effects of one IV, while ignoring the other IV
  - tested by comparing the appropriate marginal means
- RH: for interactions
  - usually expressed as "different differences" -- differences between a set of simple effects
  - tested by comparing the results of the appropriate set of simple effects
  - That's the hard part -- determining which set of simple effects gives the most direct test of the interaction RH:

Sometimes the Interaction RH: is explicitly stated

- when that happens, one set of SEs will provide a direct test of the RH: (the other won't)

Here's an example:

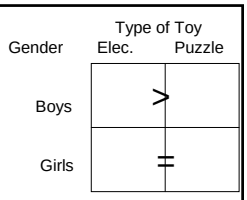
Easy tasks will be performed equally well using paper or computer presentation, however, hard tasks will be performed better using computer presentation than paper.

Task Diff.	Presentation: Comp	Presentation: Paper
Easy	=	
Hard	>	

This is most directly tested by inspecting the simple effect of paper vs. computer presentation for easy tasks, and comparing it to the simple effect of paper vs. computer for hard tasks.

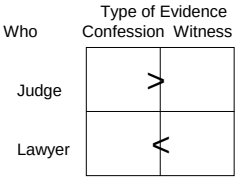
Your Turn...

Young boys will rate playing with an electronic toy higher than playing with a puzzle, whereas young girls will have no difference in ratings given to the two types of toys.



ANCOVA, cont.

Judges will rate confessions as more useful than eyewitness testimony, whereas Lawyers will rate eyewitness testimony as more useful than confessions.



Sometimes the set of SEs to use is "inferred" ...

Often one of the IVs in the study was used in previous research, and the other is "new".

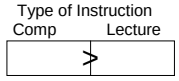
- In this case, we will usually examine the simple effect of the "old" variable, at each level of the "new" variable
- this approach gives us a clear picture of the replication and generalization of the "old" IV's effect.

e.g., Previously I demonstrated that computer presentations lead to better learning of statistical designs than does using a conventional lecture. I would like to know if the same is true for teaching writing.

Let's take this "apart" to determine which set of SEs to use to examine the pattern of the interaction...

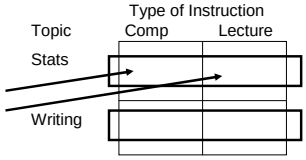
Previously I demonstrated that computer presentations lead to better learning of statistical designs than does using a conventiona lecture. I would like to know if the same is true for teaching writing.

Here's the design and result of the earlier study about learning stats.



Here's the design of the study being planned.

What cells are a replication of the earlier study ?



So, which set of SEs will allow us to check if we got the replication, and then go on to see of we get the same results with the new topic ?

Yep, SE of Type of Instruction, for each Topic ...

Your turn ..

I have previously demonstrated that rats learn Y-mazes faster than do hamsters. I wonder if the same is true for radial mazes ?

Type of Rodent	
Rat	Hamster
<	

Maze	Type of Rodent: Rat	Type of Rodent: Hamster
Y	<	
Radial	?	

I've discovered that Psyc and Soc majors learn statistics about equally well. My next research project will also compare these types of students on how well they learn research ethics.

Major	
Psyc	Soc
=	

Topic	Major: Psyc	Major: Soc
Stats	=	
Ethics	?	

Sometimes the RH: about the interaction and one about the main effects are "combined"

- this is particularly likely when the expected interaction pattern is of the > vs. > type (the most common pattern in Psyc)

Here's an example…

Group therapy tends to work better than individual therapy, although this effect is larger for patients with social anxiety than with agoraphobia.

Anxiety	Type of Therapy: Group	Type of Therapy: Indiv.
Social	>	
Agora.	>	

Int. RH: >

Main effect RH:

So, we would examine the interaction by looking at the SEs of Type of Therapy for each type of Anxiety.

Statistical Analysis of 2x2 Factorial Designs

Like a description of the results based upon inspection of the means, formal statistical analyses of factorial designs has five basic steps:

1. Tell IVs and DV
2. Present data in table or figure
3. Determine if the interaction is significant
   - if it is, describe it in terms of one of the sets of simple effects.
4. Determine whether or not the first main effect is significant
   - if it is, describe it
   - determine if that main effect is descriptive or misleading
5. Determine whether or not the second main effect is significant
   - if it is, describe it
   - determine if that main effect is descriptive or misleading

Statistical Analysis of a 2x2 Design

Task Difficulty (b)	Task Presentation (a): Paper	Task Presentation (a): Computer	
Easy	90	70	80
Hard	40	60	50
	65	65	

SE of Presentation for Easy Tasks

SE for Presentation for Hard Tasks

Presentation Main Effect	Difficulty Main Effect	Interaction Effect
$SS_{Presentation}$	$SS_{Dificulty}$	$SS_{Interaction}$
65 vs. 65	80 vs. 50	SE_{Easy} vs. SE_{Hard}

Constructing F-tests for a 2x2 Factorial

$$F_{Presentation} = \frac{(SS_{Presentation} / df_{Presentation})}{(SS_{Error} / df_{Error})}$$

$$F_{Difficulty} = \frac{(SS_{Difficulty} / df_{Difficulty})}{(SS_{Error} / df_{Error})}$$

$$F_{Interaction} = \frac{(SS_{Interaction} / df_{Interaction})}{(SS_{Error} / df_{Error})}$$

Statistical Analyses Necessary to Describe the Interaction of a 2x2 Design

However, the F-test of the interaction only tells us whether or not there is a "statistically significant" interaction…

- it does not tell use the pattern of that interaction
- to determine the pattern of the interaction we have to compare the simple effects
- to describe each simple effect, we must be able to compare the cell means

we need to know how much of a cell mean difference is "statistically significant"

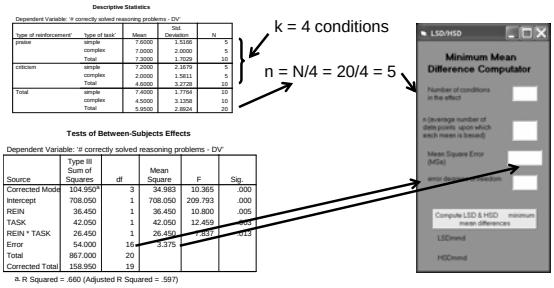
Using LSD to Compare cell means to describe the simple effects of a 2x2 Factorial design

- LSD can be used to determine how large of a cell mean difference is required to treat it as a "statistically significant mean difference"
- Will need to know three values to use the computator
  - df_{error} -- look on the printout or use N – 4
  - MS_{error} – look on the printout
  - n = N / 4 -- use the decimal value – do not round to the nearest whole number!

Remember – only use the lsdmmd to compare cell means. Marginal means are compared using the man effect F-tests.

Using the Pairwise Computator & LSDmmd to Compare cell means to describe the simple effects of a 2x2 Factorial design

For a 2x2 BG Factorial Design



Descriptive Statistics

Dependent Variable: '# correctly solved reasoning problems - DV'

Type of reinforcement	Type of task	Mean	Std. Deviation	N
praise	simple	7.6000	1.5166	5
	complex	7.0000	2.0000	5
	Total	7.3000	1.7029	10
criticism	simple	7.2000	2.1679	5
	complex	2.0000	1.5811	5
	Total	4.6000	3.2728	10
Total	simple	7.4000	1.7766	10
	complex	4.5000	3.1358	10
	Total	5.9500	2.8826	20

k = 4 conditions

n = N/4 = 20/4 = 5

Tests of Between-Subjects Effects

Dependent Variable: '# correctly solved reasoning problems - DV'

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	104.950[a]	3	34.983	10.365	.000
Intercept	708.050	1	708.050	209.793	.000
REIN	36.450	1	36.450	10.800	.005
TASK	42.050	1	42.050	12.459	.003
REIN * TASK	26.450	1	26.450	7.837	.013
Error	54.000	16	3.375		
Total	867.000	20			
Corrected Total	158.950	19			

a. R Squared = .660 (Adjusted R Squared = .597)

Support for Interaction RH:s

To be "fully supported" a RH: about an interaction must correctly specify both of the SEs involved in that RH: test.

Gender	Type of Toy: Elec.	Type of Toy: Puzzle
Boys	>	
Girls	=	

Tell if each RH: is fully, partially or not supported

- Boys will prefer Electric Toys to Puzzles, while girls will prefer Puzzles to Toys. — partial
- Girls will prefer Electric Toys to Puzzles, while boys will show no preference — none
- Boys will prefer Electric Toys to Puzzles, girls will too, but to a lesser extent. — partial
- Boys will prefer Electric Toys to Puzzles, while girls will have no preference — full

Statistical Analyses Necessary to Describe Main Effects of a 2x2 Design

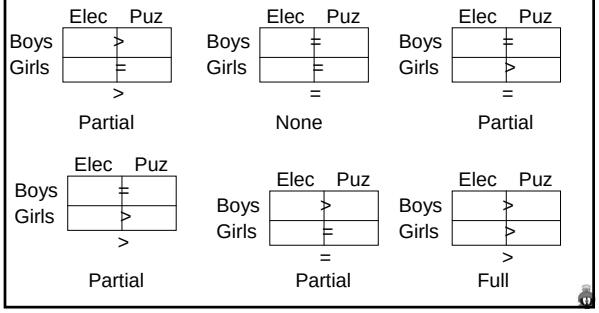
In a 2x2 Design, the Main effects F-tests are sufficient to tell us about the relationship of each IV to the DV…

- since each main effect involves the comparison of two marginal means -- the corresponding significance test tells us what we need to know …
  - whether or not those two marginal means are "significantly different"
  - Don't forget to examine the means to see if a significant difference is in the hypothesized direction !!!

Support for Main effect RH:s

A RH: about a Main effect is only fully supported if that Main effect is descriptive.

RH: Electric Toys are preferred to Puzzles – tell if each of the following give full, partial or no support …



What statistic is used for which factorial effects????

Age \ Gender	Male	Female	
5	30	30	30
10	20	30	25
	25	30	

There will be 4 statistics

1. F_{Age}
2. F_{Gender}
3. F_{Int}
4. LSD_{mmd}

This design as 7 "effects"

1. Main effect of age
2. Main effect of gender
3. Interaction of age & gender
4. SE of age for males
5. SE of age for females
6. SE of gender for 5 yr olds
7. SE of gender for 10 yr olds

What statistic is used for which factorial effects????

Age	Gender: Male	Gender: Female	
5	50	30	40
10	60	80	70
	25	30	

1. F_{Age} p = .021
2. F_{Gender} p = .082
3. F_{Int} p = .001
4. LSD_{mmd} = 15

Are 40 & 70 different ?	F_{Age}
Are 50 & 30 different ?	LSDmmd
Are 30 & 80 different ?	LSDmmd
Are 50 & 60 differently different than 30 & 80 ?	F_{Int}
Are 50 & 60 different ?	LSDmmd
Are 25 & 30 different ?	F_{Gender}
Are 50 & 30 differently different than 60 & 80 ?	F_{Int}
Are 60 & 80 different ?	LSDmmd

Applying lsd_{mmd} to 2x2 BG ANOVA

Task Difficulty	Task Presentation: Paper	Task Presentation: Computer
Easy	60	90
Hard	60	70

for the interaction
F(1,56) = 6.5, p = .023

lsd_{mmd} = 14

Is there an interaction effect? Based on what?

Yes! F-test of Int

for the following, tell the mean difference and apply the lsd_{mmd}

Simple effect of Task Presentation	30	>
SE of Task Presentation for Easy Tasks SE of Task Presentation for Hard Tasks	10	=
Simple effects of Task Difficulty		
SE of Task Difficulty for Paper Pres.	0	=
SE of Task Difficulty for Comp. Pres.	20	V

Applying lsd_{mmd} to 2x2 BG ANOVA

Task Difficulty	Task Presentation: Paper	Task Presentation: Computer	
Easy	60	90	75
Hard	60	70	65

for Difficulty ME
F(1,56) = 4.5, p = .041

lsd_{mmd} = 14

Is there a Task Difficulty main effect? Based on what?

Yes! F-test of ME

Is main effect descriptive (unconditional) or potentially misleading (conditional)?

Simple effects of Task Difficulty		
SE of Task Difficulty for Paper Pres.	0	=
SE of Task Difficulty for Comp. Pres.	20	V

Descriptive only for Computer presentation; misleading for Paper presentations.

Applying lsd_{mmd} to 2x2 BG ANOVA

Task Difficulty	Task Presentation: Paper	Task Presentation: Computer
Easy	60	90
Hard	60	70
	60	80

for Presentation ME
$F(1,56) = 7.2$, $p = .011$

$lsd_{mmd} = 14$

Is there a Task Presentation main effect? Based on what? Yes! F-test of ME

Is main effect descriptive (unconditional) or potentially misleading (conditional)?

Simple effects of Task Difficulty

SE of Task Presentation for Easy Tasks 30 **<**

SE of Task Presentation for Hard Tasks 10 **=**

Descriptive only for Easy tasks; misleading for Difficult tasks.

Effect Sizes for 2x2 BG Factorial designs

For Main Effects & Interaction (each w/ df=1)

$$r = \sqrt{[\, F / (F + df_{error})]}$$

Rem: This effect size can only be compared with other interaction effects from exactly the same factorial design

For Simple Effects

$$d = (M1 - M2) / \sqrt{Mserror}$$

$$r = \sqrt{\left[\frac{d^2}{d^2 + 4}\right]}$$ (An "approximation formula")

Rem: The effects size for a pairwise comparison can be compared with that pair of conditions from any study.