## things to know

- PS 3 grading is done
- PS 4's answer key is still in the works
- PS 5 will be assigned this evening and due on Monday
- Drill is on for tomorrow
- There is a script available for today
- April 8 will be skipped
- there is way more in the slides than I can cover today

1

## a note about `emmeans`

- this is a widely-used package in R for the kinds of designs we've been talking about
- it has the following amusing note, early in its FAQ

**I wanted comparisons, but all I get is (nothing)**

This happens when you have only one estimate; and you can't compare it with itself! This is turn can happen when you have a situation like this: you have fitted

```
mod <- lm(RT ~ treat, data = mydata)
```

and `treat` is coded in your dataset with numbers 1, 2, 3, … . Since `treat` is a numeric predictor, `emmeans()` just reduces it to a single number, its mean, rather than separate values for each treatment. Also, please note that this is almost certainly NOT the model you want, because it forces an assumption that the treatment effects all fall on a straight line. You should fit a model like

```
mod <- lm(RT ~ factor(treat), data = mydata)
```

then you will have much better luck with comparisons.

2

## multi-factor designs (part 2)

February 21, 2024

3

## review

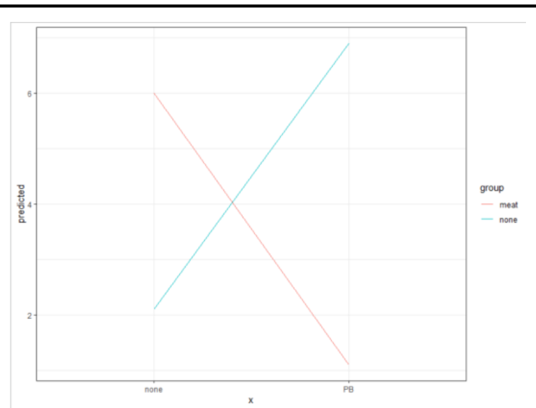
- a factorial design with two factors has
  - *main effects* (the effect of one factor ignoring the other)
  - an *interaction effect* (whether the effect of one factor depends on the value of the other)
- we can analyze a 2 × 2 design with
  - ANOVA
    - (when we move to bigger designs, ANOVA will leave us wanting)
  - contrast codes for the main effects
  - dummy codes for the simple effects/slopes

4

## reminder of the design, results

| | | meat | | |
|---|---|---|---|---|
| | | none | some | |
| PB | none | M = 2.1, $s^2$ = 3.2 | M = 6.0, $s^2$ = 4.2 | M = 4.05 |
| | some | M = 6.9, $s^2$ = 4.8 | M = 1.1, $s^2$ = 2.5 | M = 4.0 |
| | | M = 4.5 | M = 3.55 | |

5




6

## linear model results (dummy codes vs contrast codes)

```
Dummy        Estimate        SE        t    Pr(>F)
(Intercept)    2.1000    0.6071    3.459   0.00141 **
meatD          3.9000    0.8586    4.542  6.03e-05 ***
PBD            4.8000    0.8586    5.590  2.45e-06 ***
meatD:PBD     -9.7000    1.2143   -7.988  1.74e-09 ***

Contrast     Estimate        SE        t    Pr(>F)
(Intercept)    4.0250    0.3036   13.259  2.02e-15 ***
meatC         -0.9500    0.6071   -1.565     0.126
PBC           -0.0500    0.6071   -0.082     0.935
int           -9.7000    1.2143   -7.988  1.74e-09 **
```

7

## reference-reference mean

```
summary(lm(tastiness ~ meatD*PBD, d), t = F)

             Estimate        SE       t    Pr(>F)
(Intercept)    2.1000    0.6071   3.459   0.00141
```

| | | meat: none | meat: some |
|---|---|---|---|
| PB | none | M = 2.1 | M = 6.0 |
| | some | M = 6.9 | M = 1.1 |

8

## simple slope/effect

```
summary(lm(tastiness ~ meatD*PBD, d), t = F)

             Estimate        SE       t    Pr(>F)
meatD          3.9000    0.8586   4.542  6.03e-05
```

| | | meat: none | meat: some |
|---|---|---|---|
| PB | none | M = 2.1 | M = 6.0 |
| | some | M = 6.9 | M = 1.1 |

9

## simple slope/effect

```
summary(lm(tastiness ~ meatD*PBD, d), t = F)

            Estimate       SE       t   Pr(>F)
PBD           4.8000   0.8586   5.590 2.45e-06
```

| | | meat: none | meat: some |
|---|---|---|---|
| PB | none | M = 2.1 | M = 6.0 |
| | some | M = 6.9 | M = 1.1 |

10

## interaction effect

```
summary(lm(tastiness ~ meatD*PBD, d), t = F)

            Estimate       SE       t   Pr(>F)
meatD:PBD    -9.7000   1.2143  -7.988 1.74e-09
```

| | | meat: none | meat: some |
|---|---|---|---|
| PB | none | M = 2.1 | M = 6.0 |
| | some | M = 6.9 | M = 1.1 |

11

## practical advice

- which should we use?
- it depends on what you want to know!

12

## what do we want to know?

| | | meat | |
|---|---|---|---|
| | | none | some | |
| PB | none | M = 2.1, $s^2$ = 3.2 | M = 6.0, $s^2$ = 4.2 | M = 4.05 |
| | some | M = 6.9, $s^2$ = 4.8 | M = 1.1, $s^2$ = 2.5 | M = 4.0 |
| | | M = 4.5 | M = 3.55 | |

- if simple slopes/effects, use carefully-chosen dummy codes

13

## what do we want to know?

| | | meat | | |
|---|---|---|---|---|
| | | none | some | |
| PB | none | M = 2.1, $s^2$ = 3.2 | M = 6.0, $s^2$ = 4.2 | M = 4.05 |
| | some | M = 6.9, $s^2$ = 4.8 | M = 1.1, $s^2$ = 2.5 | M = 4.0 |
| | | M = 4.5 | M = 3.55 | |

- if main effects, use carefully chosen contrast codes (or the usual ANOVA)

14

## what do we want to know?

| | | meat | | |
|---|---|---|---|---|
| | | none | some | |
| PB | none | M = 2.1, $s^2$ = 3.2 | M = 6.0, $s^2$ = 4.2 | M = 4.05 |
| | some | M = 6.9, $s^2$ = 4.8 | M = 1.1, $s^2$ = 2.5 | M = 4.0 |
| | | M = 4.5 | M = 3.55 | |

- if the interaction, it doesn't matter much

15

## what is an interaction?

| | | meat | |
|---|---|---|---|
| | | none | some | |
| PB | none | M = 2.1, $s^2$ = 3.2 | M = 6.0, $s^2$ = 4.2 | M = 4.05 |
| | some | M = 6.9, $s^2$ = 4.8 | M = 1.1, $s^2$ = 2.5 | M = 4.0 |
| | | M = 4.5 | M = 3.55 | |

- when the effect of one variable changes across values of another variable
- here, the effect of PB is to *increase* tastiness when there is no meat
- but the effect of PB is to *reduce* tastiness when there is meat

16

## notice that the interaction is really a contrast between simple slopes

| | | meat | | |
|---|---|---|---|---|
| | | none | some | |
| PB | none | M = 2.1, $s^2$ = 3.2 | M = 6.0, $s^2$ = 4.2 | M = 4.05 |
| | some | M = 6.9, $s^2$ = 4.8 | M = 1.1, $s^2$ = 2.5 | M = 4.0 |
| | | M = 4.5 | M = 3.55 | |

- this simple slope is 6.9 – 2.1 = +4.8
- this simple slope is 1.1 – 6.0 = -4.9
- the contrast between the simple slopes is 4.8 – (-4.9) = -9.7

17

## another way to think about this design

- we have a four-group design, which can be depicted as follows

| no PB<br>no meat | no PB<br>some meat | some PB<br>no meat | some PB<br>some meat |
|---|---|---|---|
| 2.1 | 6.9 | 6.0 | 1.1 |

18

## another way to think about this design

- we could analyze this design using the method of subsets

| no PB no meat | no PB some meat | some PB no meat | some PB some meat |
|---|---|---|---|
| 2.1 | 6.9 | 6.0 | 1.1 |
| 3/4 | -1/4 | -1/4 | -1/4 |
| 0 | 1/3 | 1/3 | -2/3 |
| 0 | 1/2 | -1/2 | 0 |

19

## another way to think about this design

- this would answer some interesting questions, but would not test the interaction (nor any main effects)

| no PB no meat | no PB some meat | some PB no meat | some PB some meat |
|---|---|---|---|
| 2.1 | 6.9 | 6.0 | 1.1 |
| 3/4 | -1/4 | -1/4 | -1/4 |
| 0 | 1/3 | 1/3 | -2/3 |
| 0 | 1/2 | -1/2 | 0 |

20

## another way to think about this design

- contrast codes result in a different set of contrasts

| no PB no meat | no PB some meat | some PB no meat | some PB some meat |
|---|---|---|---|
| 2.1 | 6.9 | 6.0 | 1.1 |
| -1/2 | -1/2 | +1/2 | +1/2 |
| -1/2 | +1/2 | -1/2 | +1/2 |
| +1/4 | -1/4 | -1/4 | +1/4 |

21

## interim summary

- despite the factorial nature of this design, it's just a four-group design
- *any* three orthogonal contrasts can be used to analyze it
- but if we are interested in specific questions – including the interaction – we need to carefully choose our contrasts
- contrast or dummy coding main effects (and creating a product term) will allow us to answer the questions of interest

22

## larger two-factor designs

- Factor A: sentence (normal/intact vs scrambled)
- Factor B: presentation rate (300, 450, 600 wpm)
- DV = % correct detection of a word

- this is a 2 (sentence) × 3 (rate) design
- there are six groups
- ultimately, no matter how we create them, we'll need five contrast codes

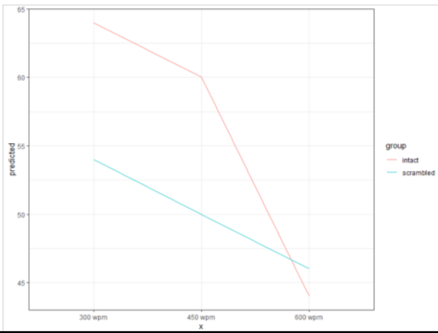
23

## the results (cell, marginal, overall means)

| | 300 | 450 | 600 | |
|---|---|---|---|---|
| intact | 64 | 60 | 44 | 56 |
| scrambled | 54 | 50 | 46 | 50 |
| | 59 | 55 | 45 | 53 |

24

## results, plotted




25

## how to analyze?

- let's generate contrast codes for each factor, ignoring the other factor
- for the sentence factor, there's no decision to be made
- with two levels, we'll use +1/2 and -1/2

26

## filling in some codes

| | intact 300 | intact 450 | intact 600 | scr 300 | scr 450 | scr 600 |
|---|---|---|---|---|---|---|
| T | +1/2 | +1/2 | +1/2 | -1/2 | -1/2 | -1/2 |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |

27

## how to analyze?

- let's generate contrast codes for each factor, ignoring the other factor
- for the rate factor, the researcher thought something interest would happen at the very-high rate relative to the other two
- R1: 300, 450 vs 600
- the other contrast is the only one leftover
- R2: 300 vs 450

28

## filling in some codes: multiply to get interactions

| | intact 300 | intact 450 | intact 600 | scr 300 | scr 450 | scr 600 |
|---|---|---|---|---|---|---|
| T | +1/2 | +1/2 | +1/2 | -1/2 | -1/2 | -1/2 |
| R1 | +1/3 | +1/3 | -2/3 | +1/3 | +1/3 | -2/3 |
| R2 | +1/2 | -1/2 | 0 | +1/2 | -1/2 | 0 |
| T*R1 | +1/6 | +1/6 | -2/6 | -1/6 | -1/6 | +2/6 |
| T*R2 | +1/4 | -1/4 | 0 | -1/4 | +1/4 | 0 |

29

## what do we get?

```
             Estimate        SE      t Pr(>|t|)
(Intercept)        53      0.99  53.62  < 2e-16
T                   6      1.98   3.03  0.00412
R1                 12      2.10   5.72 9.95e-07
R2                  4      2.42   1.65  0.10600
TR1                12      4.19   2.86  0.00655
TR2                 0      4.84   0.00  1.00000
```

| | 300 | 450 | 600 | |
|---|---|---|---|---|
| intact | 64 | 60 | 44 | 56 |
| scrambled | 54 | 50 | 46 | 50 |
| | 59 | 55 | 45 | 53 |

30