Analyzing Experiments

Data Analysis for Psychology in R 2

dapR2 Team

Department of Psychology
The University of Edinburgh

Week's Learning Objectives

1. Understand the different types of experimental design and the data that results from them.
2. Be able to link experimental designs to linear model specifications
3. Understand the distinction between simple, main and interaction effects.
4. Be able to test main effects (via F-tests)
5. Interpret interactions with effects coding.

Experimental Design: manipulation

- A key feature of experimental designs is that we actively manipulate our predictor (IV).
- The intention is that changing the predictor will result in changes in the outcome (DV).
- That is our manipulation will lead to variation in the outcome.
- Our experiments can fail because we design these manipulations poorly.
- The predictors in an experiment are (primarily) experimental conditions.

Conditions/Factors & levels

- **Conditions**:
  - Are part of our experimental designs.
  - They are what is manipulated.
- **Factors**
  - The resultant variables in our data set that code the experimental conditions are typically called factors.
  - Generally the terms conditions and factors are used interchangeably.
  - But it is useful to differentiate the design (conditions) and the data that represents aspects of the design (factors)
- Factors can have **levels**
  - These are the number of ways we vary or manipulate the condition

Between vs Within Person

- Two broad choices of study structure:
  - **Between person**: Participants only appear on one level/condition
  - **Within person**: Participants appear in multiple level/conditions

Between vs Within Person

- Two broad choices of study structure:
  - **Between person**: Participants only appear on one level/condition
  - **Within person**: Participants appear in multiple level/conditions
- The labels we use to refer to kinds of studies reflect the number of conditions and whethert the conditions are between vs within.
  - One-way between person
  - Two-way within person
  - etc.

A new study

- Suppose we wanted to look at the number of reading errors caused by noise distraction.
- We might devise a task where participants had to read a passage of text and put a cross through all verbs.
- Our outcome, or dependent variable, is the number of verbs correctly crossed out.
- Our predictor, or independent variable, is the noise level.

One-way Between Person

Noise Level	
None	**Noise**
Tom	Adam
Aja	Fiona
Alex	Simon
Brandy	Tasha
Darren	Josh
Lucy	Charlotte

One-way Between Person (more levels)

Noise Level		
None	**Moderate**	**Loud**
Tom	Darren	Adam
Aja	Lucy	Fiona
Alex	Josh	Simon
Brandy	Charlotte	Tasha

Two-way Between Person

	Noise		
Distraction	**None**	**Moderate**	**Loud**
Words	Tom	Darren	Adam
Words	Aja	Lucy	Fiona
No Words	Alex	Josh	Simon
No Words	Brandy	Charlotte	Tasha

One-way Within Person

Noise Level	
None	**Noise**
Tom	Tom
Aja	Aja
Alex	Alex
Brandy	Brandy
Darren	Darren
Lucy	Lucy
Josh	Josh
Charlotte	Charlotte

Two-way Within Person

	Noise Level	
Distraction	**None**	**Noise**
Word	Tom	Tom
Word	Aja	Aja
Word	Alex	Alex
Word	Brandy	Brandy
No Word	Tom	Tom
No Word	Aja	Aja
No Word	Alex	Alex
No Word	Brandy	Brandy

Mixed Designs

	Noise Level	
Distraction	**None**	**Noise**
Word	Tom	Tom
Word	Aja	Aja
Word	Alex	Alex
Word	Brandy	Brandy
No Word	Darren	Darren
No Word	Lucy	Lucy
No Word	Josh	Josh
No Word	Charlotte	Charlotte

Models and Experiments

- Our linear model can be simply stated as:

$$outcome = model + error$$

- When we have an experiment:

$$outcome = design + error$$

- The design is simply sets of categorical variables.

$$y = b_0 + \underbrace{(b_1 E_1 + b_2 E_2)}_{\text{Conditin1}} + \underbrace{b_3 E_3}_{\text{Condition2}} + \underbrace{b_4 E_{13} + b_5 E_{23}}_{\text{Interactions}} + \underbrace{\epsilon_i}_{\text{error}}$$

- So to analyse an experiment, we are simply analysing a linear model with categorical predictors.
- From here on we will focus on between person designs

Time for a break

Welcome Back!

We are going to move on to look at the type of statistical tests we typically make in experiments, and how to do them with linear models

Hypotheses we test in experimental studies

- One-way designs:
  - **Main effect**: Tests overall effect of a condition (F-tests)
  - **Contrasts**: Tests differences between specific group means (based on coding schemes and associated β)

Hypotheses we test in experimental studies

- One-way designs:
  - **Main effect**: Tests overall effect of a condition (F-tests)
  - **Contrasts**: Tests differences between specific group means (based on coding schemes and associated β)
- Factorial designs:
  - Main effects & Contrasts
  - **Interactions**: Categorical*categorical and usually based on effects (sum to zero) coding (F-tests & β)
  - **Simple contrasts/effects**: Effects of one level in one condition, across levels of another condition.

Example

- We will keep with out hospital example so we have familiar data and can make easy links and comparisons to previous weeks.
- A researcher was interested in whether the subjective well-being of patients differed dependent on the post-operation treatment schedule they were given, and the hospital in which they were staying.
- **Condition 1**: `Treatment` (Levels: TreatA, TreatB, TreatC).
- **Condition 2**: `Hosp` (Levels: Hosp1, Hosp2).
- Total sample n = 180 (30 patients in each of 6 groups).
  - Between person design.
- **Outcome**: Subjective well-being (SWB)
  - An average of multiple raters (the patient, a member of their family, and a friend).
  - SWB score ranged from 0 to 20.

The data

```
hosp_tbl <- read_csv("hospital.csv", col_types = "dff")
hosp_tbl %>%
  slice(1:10)
```

```
## # A tibble: 10 x 3
##       SWB Treatment Hospital
##     <dbl> <fct>     <fct>
##  1   6.2 TreatA    Hosp1
##  2  15.9 TreatA    Hosp1
##  3   7.2 TreatA    Hosp1
##  4  11.3 TreatA    Hosp1
##  5  11.2 TreatA    Hosp1
##  6   9   TreatA    Hosp1
##  7  14.5 TreatA    Hosp1
##  8   7.3 TreatA    Hosp1
##  9  13.7 TreatA    Hosp1
## 10  12.6 TreatA    Hosp1
```

Table of means

```
mean(hosp_tbl$SWB)
```

```
## [1] 9.880556
```

```
aggregate(SWB ~ Treatment + Hospital,
  hosp_tbl, mean)
```

```
##   Treatment Hospital       SWB
## 1    TreatA    Hosp1 10.800000
## 2    TreatB    Hosp1  9.430000
## 3    TreatC    Hosp1 10.103333
## 4    TreatA    Hosp2  7.853333
## 5    TreatB    Hosp2 13.116667
## 6    TreatC    Hosp2  7.980000
```

```
aggregate(SWB ~ Hospital,
  hosp_tbl, mean)
```

```
##   Hospital      SWB
## 1    Hosp1 10.11111
## 2    Hosp2  9.65000
```

```
aggregate(SWB ~ Treatment,
  hosp_tbl, mean)
```

```
##   Treatment       SWB
## 1    TreatA  9.326667
## 2    TreatB 11.273333
## 3    TreatC  9.041667
```

Table of means

- All of the above gives us a full table of means

	Hosp1	Hosp2	Marginal
TreatA	10.80	7.85	9.33
TreatB	9.43	13.11	11.27
TreatC	10.10	7.98	9.04
Marginal	10.11	9.65	9.88

- We are going to start with loking at the testing of main effects using F-tests.
  - i.e. exactly what we have done before.

One way main effects

- As we have an experiment, we typically use efefcts coding:

```
contrasts(hosp_tbl$Treatment) <- contr.sum
contrasts(hosp_tbl$Hospital) <- contr.sum
```

- Run the model:

```
m1 <- lm(SWB ~ Treatment, data = hosp_tbl)
anova(m1)
```

```
## Analysis of Variance Table
##
## Response: SWB
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## Treatment   2  177.02  88.511  14.037 2.196e-06 ***
## Residuals 177 1116.08   6.306
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

One way main effects

```
summary(m1)
```

```
##
## Call:
## lm(formula = SWB ~ Treatment, data = hosp_tbl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.373 -1.987 -0.300  1.838  7.173
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.8806     0.1872  52.791  < 2e-16 ***
## Treatment1   -0.5539     0.2647  -2.093   0.0378 *
## Treatment2    1.3928     0.2647   5.262 4.09e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.511 on 177 degrees of freedom
## Multiple R-squared:  0.1369,	Adjusted R-squared:  0.1271
## F-statistic: 14.04 on 2 and 177 DF,  p-value: 2.196e-06
```

Table of means

```r
mean(hosp_tbl$SWB)
```

```
## [1] 9.880556
```

```r
aggregate(SWB ~ Treatment,
  hosp_tbl, mean)
```

```
##   Treatment       SWB
## 1    TreatA  9.326667
## 2    TreatB 11.273333
## 3    TreatC  9.041667
```

```r
m1$coefficients
```

```
## (Intercept)  Treatment1  Treatment2
##   9.8805556  -0.5538889   1.3927778
```

Hypotheses we test in Factorial Designs

- Main effects
    - An overall, or average, effect of a condition.
    - Is there an effect of `Treatment` averaged over `Hospital`?
    - Is there an effect of `Hospital` averaged over `Treatment`?
- Interactions (categorical*categorical)
    - A change in the effect of some condition as a function of another.
    - Does the effect of `Treatment` differ by `Hospital`?

Hypotheses we test in Factorial Designs

- Main effects
  - An overall, or average, effect of a condition.
  - Is there an effect of `Treatment` averaged over `Hospital`?
  - Is there an effect of `Hospital` averaged over `Treatment`?
- Interactions (categorical*categorical)
  - A change in the effect of some condition as a function of another.
  - Does the effect of `Treatment` differ by `Hospital`?
- Simple contrasts/effects
  - An effect of one condition at a specific level of another.
  - Is there an effect of `Hospital` for those receiving `Treatment A`? (...and so on for all combinations.)

Our model and coefficients

- Remember whichever coding scheme we use, we have k-1 variables representing the condition.
  - So for `Treatment` we have 2 predictors (E1 & E2)
  - And for `Hospital` we have 1 predictor (E3)
- We can write the linear model more explicitly as:

$$y_{ijk} = b_0 + \underbrace{(b_1E_1 + b_2E_2)}_{\text{Treatment}} + \underbrace{b_3E_3}_{\text{Hospital}} + \underbrace{b_4E_{13} + b_5E_{23}}_{\text{Interactions}} + \epsilon_i$$

For effects coding

$$y_{ijk} = b_0 + \underbrace{(b_1 E_1 + b_2 E_2)}_{\text{Treatment}} + \underbrace{b_3 E_3}_{\text{Hospital}} + \underbrace{b_4 E_{13} + b_5 E_{23}}_{\text{Interactions}} + \epsilon_i$$

```
## # A tibble: 6 x 7
##   Treatment Hospital    E1    E2    E3   E13   E23
##   <chr>     <chr>    <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 A         Hosp1        1     0     1     1     0
## 2 A         Hosp2        1     0    -1    -1     0
## 3 B         Hosp1        0     1     1     0     1
## 4 B         Hosp2        0     1    -1     0    -1
## 5 C         Hosp1       -1    -1     1    -1    -1
## 6 C         Hosp2       -1    -1    -1     1     1
```

Factorial main effects and interaction

- Run the model:

```
m2 <- lm(SWB ~ Treatment*Hospital, data = hosp_tbl)
anova(m2)
```

```
## Analysis of Variance Table
##
## Response: SWB
##                     Df Sum Sq Mean Sq F value    Pr(>F)
## Treatment            2 177.02  88.511 21.5597 4.315e-09 ***
## Hospital             1   9.57   9.568  2.3306    0.1287
## Treatment:Hospital   2 392.18 196.088 47.7635 < 2.2e-16 ***
## Residuals          174 714.34   4.105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using model comparisons

- The F-test table can be thought of as containing the results of a set of model comparisons between the following models:

```
comp1 <- lm(SWB ~ Treatment, data = hosp_tbl)
comp2 <- lm(SWB ~ Hospital, data = hosp_tbl)
comp3 <- lm(SWB ~ Treatment + Hospital, data = hosp_tbl)
comp4 <- lm(SWB ~ Treatment + Hospital + Treatment*Hospital, data = hosp_tbl)
```

Testing the overall effects

- For the effect of `Treatment`:

```
anova(comp2,comp3)
```

```
## Analysis of Variance Table
##
## Model 1: SWB ~ Hospital
## Model 2: SWB ~ Treatment + Hospital
##   Res.Df    RSS Df Sum of Sq      F   Pr(>F)
## 1    178 1283.5
## 2    176 1106.5  2    177.02 14.078 2.13e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- An effect of Treatment

Testing the overall effects

- For the effect of `Hospital`:

```
anova(comp1, comp3)
```

```
## Analysis of Variance Table
##
## Model 1: SWB ~ Treatment
## Model 2: SWB ~ Treatment + Hospital
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    177 1116.1
## 2    176 1106.5  1    9.5681 1.5219  0.219
```

- No effect of hospital

Testing the overall effects

- For the effect of interaction:

```
anova(comp3, comp4)
```

```
## Analysis of Variance Table
##
## Model 1: SWB ~ Treatment + Hospital
## Model 2: SWB ~ Treatment + Hospital + Treatment * Hospital
##   Res.Df     RSS Df Sum of Sq      F    Pr(>F)
## 1    176 1106.51
## 2    174  714.34  2    392.18 47.764 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- An interaction

Testing the overall effects

- You may have noted using `anova()` for a single model, and for the model comparison approach yeild slightly different results.
    - Sums of squares difference is the same
    - Degrees of freedom are the same
    - F is slightly different for `Treatment` and `Hospital` (and therefore so is p-value)
- Note the main concluions do not change.
- This difference relates to differences in the degrees of freedom associated with the F-test.

Time for a break

In the next section we will use interaction plots for effects coded variables. Please read the Handout on LEARN before watching the next video

Welcome back!

Interpreting an interaction with effects codes

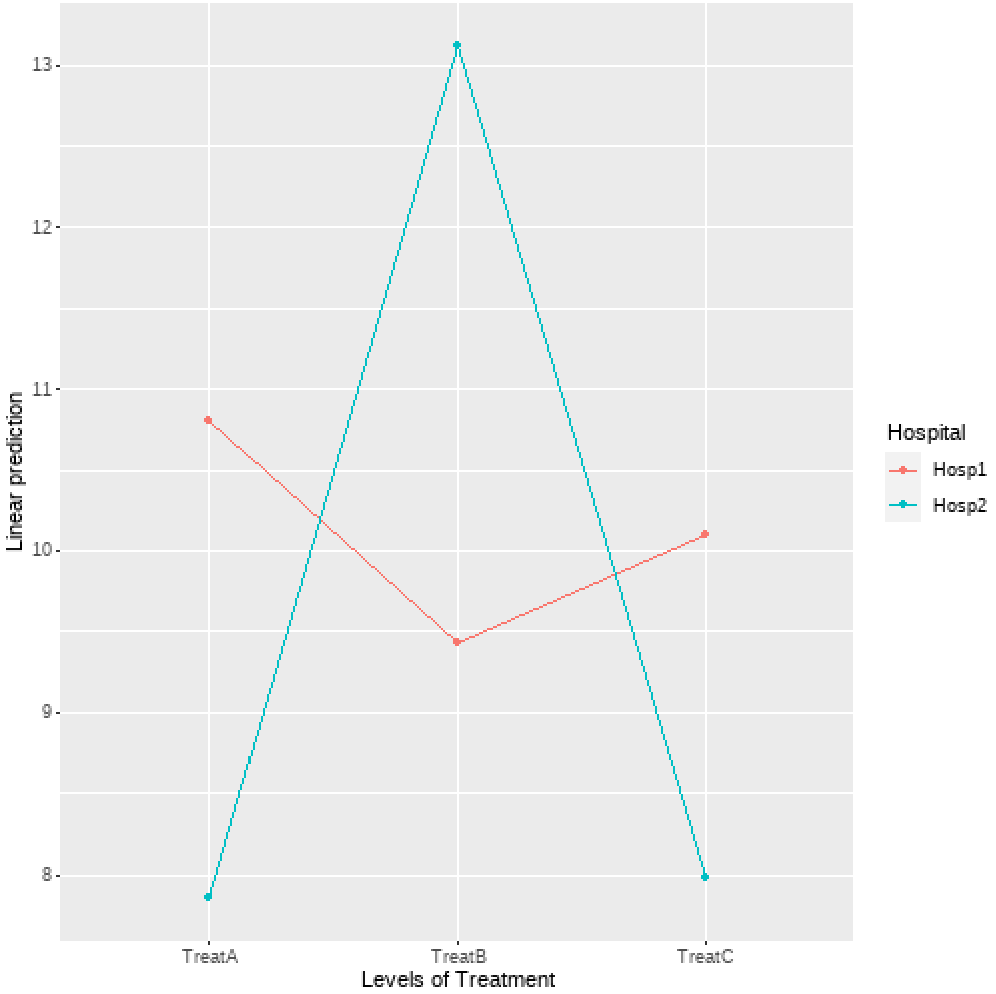
Recap categorical interactions

- When the effects of one predictor on the outcome differ across levels of another predictor.
- Categorical*categorical interaction:
    - There is a difference in the differences between groups across levels of a second factor.

Our results

```
##
## Call:
## lm(formula = SWB ~ Treatment * Hospital, data = hosp_tbl)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.6000 -1.2533  0.1083  1.2650  5.7000
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)            9.8806     0.1510  65.425  < 2e-16 ***
## Treatment1            -0.5539     0.2136  -2.593   0.0103 *
## Treatment2             1.3928     0.2136   6.521 7.30e-10 ***
## Hospital1              0.2306     0.1510   1.527   0.1287
## Treatment1:Hospital1   1.2428     0.2136   5.819 2.79e-08 ***
## Treatment2:Hospital1  -2.0739     0.2136  -9.710  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.026 on 174 degrees of freedom
## Multiple R-squared:  0.4476,    Adjusted R-squared:  0.4317
## F-statistic:  28.2 on 5 and 174 DF,  p-value: < 2.2e-16
```

Visualizing the interaction



Interpretation with effects coding

```
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 9.88       0.15   65.42     0.00
## Treatment1                 -0.55       0.21   -2.59     0.01
## Treatment2                  1.39       0.21    6.52     0.00
## Hospital1                   0.23       0.15    1.53     0.13
## Treatment1:Hospital1        1.24       0.21    5.82     0.00
## Treatment2:Hospital1       -2.07       0.21   -9.71     0.00
```

- b_0 = Grand mean.
- b_1 = Difference between row marginal for treatment A and the grand mean.
- b_2 = Difference between row marginal for treatment B and the grand mean.
- b_3 = Difference between column marginal for Hospital 1 and the grand mean.
- b_4 = Difference between Treatment A and grand mean, in Hospital 1 and Hospital 2
- b_5 = Difference between Treatment B and grand mean, in Hospital 1 and Hospital 2

Interpretation with effects coding

```
##                           Estimate Std. Error
## (Intercept)                   9.88       0.15
## Treatment1                   -0.55       0.21
## Treatment2                    1.39       0.21
## Hospital1                     0.23       0.15
## Treatment1:Hospital1          1.24       0.21
## Treatment2:Hospital1         -2.07       0.21
```

	Hosp1	Hosp2	Marginal
TreatA	10.80	7.85	9.33
TreatB	9.43	13.11	11.27
TreatC	10.10	7.98	9.04
Marginal	10.11	9.65	9.88

- b_0 = Grand mean.
- b_1 = Difference between row marginal for treatment A and the grand mean.
- b_2 = Difference between row marginal for treatment B and the grand mean.
- b_3 = Difference between column marginal for Hospital 1 and the grand mean.
- b_4 = Difference between Treatment A and grand mean, in Hospital 1 and Hospital 2
- b_5 = Difference between Treatment B and grand mean, in Hospital 1 and Hospital 2

Our results

```
m2sum <- summary(m2)
round(m2sum$coefficients,2)
```

```
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  9.88       0.15   65.42     0.00
## Treatment1                  -0.55       0.21   -2.59     0.01
## Treatment2                   1.39       0.21    6.52     0.00
## Hospital1                    0.23       0.15    1.53     0.13
## Treatment1:Hospital1         1.24       0.21    5.82     0.00
## Treatment2:Hospital1        -2.07       0.21   -9.71     0.00
```

Visualizing the interaction

Summary

- This week we had a good amount of material.
- We reviewed experimental designs, and linked this linear models with categorical predictors.
- We defined main effects, simple effects, and interactions in the context of an experiment.
- We used F-tests to explore main effects
- We looked at interactions with effects coding
- Next time we will look more closely at simple effects

Thanks for listening!