

# Contrast coding in linear (mixed-effects) models

Nora Wickelmaier

Last modified: April 23, 2025

## Categorical predictors

- For all regression models (including mixed-effects model), we need to make a choice in how to treat categorical predictors

## Categorical predictors

- For all regression models (including mixed-effects model), we need to make a choice in how to treat categorical predictors
- We need to assign numeric values often called contrasts to our categorical variables to enter them into a regression model

## Categorical predictors

- For all regression models (including mixed-effects model), we need to make a choice in how to treat categorical predictors
- We need to assign numeric values often called contrasts to our categorical variables to enter them into a regression model
- Contrasts allow for comparisons between one or more levels of a categorical variable

## Categorical predictors

- For all regression models (including mixed-effects model), we need to make a choice in how to treat categorical predictors
- We need to assign numeric values often called contrasts to our categorical variables to enter them into a regression model
- Contrasts allow for comparisons between one or more levels of a categorical variable
- For any categorical variable with  $n$  levels,  $n - 1$  contrasts (or dummy variables) can be defined

## Categorical predictors

- For all regression models (including mixed-effects model), we need to make a choice in how to treat categorical predictors
- We need to assign numeric values often called contrasts to our categorical variables to enter them into a regression model
- Contrasts allow for comparisons between one or more levels of a categorical variable
- For any categorical variable with  $n$  levels,  $n - 1$  contrasts (or dummy variables) can be defined
- The default contrast in R is the treatment contrast (or dummy coding)

## Categorical predictors

- For all regression models (including mixed-effects model), we need to make a choice in how to treat categorical predictors
- We need to assign numeric values often called contrasts to our categorical variables to enter them into a regression model
- Contrasts allow for comparisons between one or more levels of a categorical variable
- For any categorical variable with  $n$  levels,  $n - 1$  contrasts (or dummy variables) can be defined
- The default contrast in R is the treatment contrast (or dummy coding)
- For ANOVAs, so-called sum or effects coding is usually used

## ① 2 × 2 example for linear mixed-effects model



## Example by Brehm and Alday (2022)

- We will look at the impact of different contrasts with a small simulated example

## Example by Brehm and Alday (2022)

- We will look at the impact of different contrasts with a small simulated example
- We have two manipulated factors:

## Example by Brehm and Alday (2022)

- We will look at the impact of different contrasts with a small simulated example
- We have two manipulated factors:
  1. *Utensils* with levels *Fork* and *Spoon*

## Example by Brehm and Alday (2022)

- We will look at the impact of different contrasts with a small simulated example
- We have two manipulated factors:
  1. *Utensils* with levels *Fork* and *Spoon*
  2. *Foods* with levels *Salad* and *Soup*

## Example by Brehm and Alday (2022)

- We will look at the impact of different contrasts with a small simulated example
- We have two manipulated factors:
  1. *Utensils* with levels *Fork* and *Spoon*
  2. *Foods* with levels *Salad* and *Soup*
- The dependent measure is *RT* (speed of eating in minutes)

## Example by Brehm and Alday (2022)

- We will look at the impact of different contrasts with a small simulated example
- We have two manipulated factors:
  1. *Utensils* with levels *Fork* and *Spoon*
  2. *Foods* with levels *Salad* and *Soup*
- The dependent measure is *RT* (speed of eating in minutes)
- We will simulate 20 subjects and 10 food ingredients (items)

## Example by Brehm and Alday (2022)

- We will look at the impact of different contrasts with a small simulated example
- We have two manipulated factors:
  1. *Utensils* with levels *Fork* and *Spoon*
  2. *Foods* with levels *Salad* and *Soup*
- The dependent measure is *RT* (speed of eating in minutes)
- We will simulate 20 subjects and 10 food ingredients (items)
- The data is generated by the following model

$$RT = \beta_0 + \beta_1 \textit{Utensils} + \beta_2 \textit{Foods} + \beta_3 \textit{Utensils} \times \textit{Foods} + v_0 + \eta_0 + \varepsilon$$

with  $v_0 \sim N(0, \sigma_v^2)$ ,  $\eta_0 \sim N(0, \sigma_\eta^2)$ ,  $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ , all i.i.d.

## Example by Brehm and Alday (2022)

- We will look at the impact of different contrasts with a small simulated example
- We have two manipulated factors:
  1. *Utensils* with levels *Fork* and *Spoon*
  2. *Foods* with levels *Salad* and *Soup*
- The dependent measure is *RT* (speed of eating in minutes)
- We will simulate 20 subjects and 10 food ingredients (items)
- The data is generated by the following model

$$RT = \beta_0 + \beta_1 \text{UtensilsSpoon} + \beta_2 \text{FoodsSoup} + \beta_3 \text{UtensilsSpoon} \times \text{FoodsSoup} + v_0 + \eta_0 + \varepsilon$$

with  $v_0 \sim N(0, \sigma_v^2)$ ,  $\eta_0 \sim N(0, \sigma_\eta^2)$ ,  $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ , all i.i.d.



## Example by Brehm and Alday (2022)

```
library(lme4)
set.seed(1012)

# Average time to eat in min
SpoonSoup <- 5
ForkSoup <- 10
SpoonSalad <- 10
ForkSalad <- 5
# Standard deviation for all groups
Groupsd <- 2
# Number of subjects (ps) and ingredients (ii)
ps <- 20
ii <- 10
```

## Example by Brehm and Alday (2022)

```
# Create data frame for given design
ds <- data.frame(
  Utensils      = factor(rep(c("Spoon", "Fork"), each = ps * ii, times = 2)),
  Foods         = factor(rep(c("Soup", "Salad"), each = ps * ii * 2)),
  Participant   = factor(rep(paste0("p", 1:ps), times = ii * 4)),
  Item          = factor(rep(paste0("i", 1:ii), each = ps, times = 4)))

# Simulate data
ds$RT <- c(rnorm(ps * ii, mean = SpoonSoup, sd = Groupsd),
          rnorm(ps * ii, mean = ForkSoup, sd = Groupsd),
          rnorm(ps * ii, mean = SpoonSalad, sd = Groupsd),
          rnorm(ps * ii, mean = ForkSalad, sd = Groupsd))
psre <- rnorm(ps, mean = 0, sd = Groupsd / 5)
iire <- rnorm(ii, mean = 0, sd = Groupsd / 5)
ds$RT <- ds$RT + psre[ds$Participant] + iire[ds$Item]
```

## Treatment coding

- The default coding in R is treatment (or dummy) coding
- The first level of each factor is assigned as the reference category

```
contrasts(ds$Utensils)
#      Spoon
# Fork      0
# Spoon      1
contrasts(ds$Foods)
#      Soup
# Salad     0
# Soup      1
```

## Effects coding

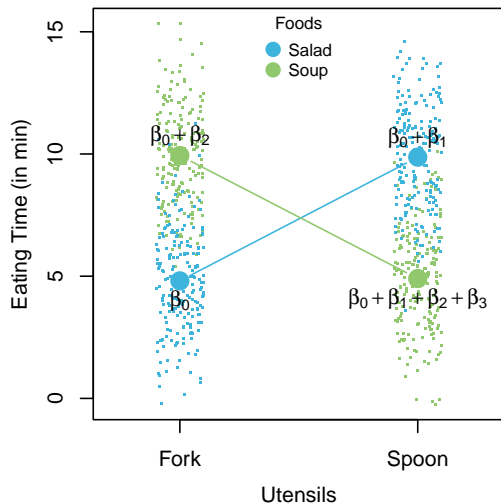
- In effects (or sum) coding, one level is set as negative and one positive, with zero as the mean of the two levels
- For more than two levels, there will also be a reference category, which is the *last* factor level in R

```
contrasts(ds$Utensils)
#           [,1]
# Fork         1
# Spoon        -1
contrasts(ds$Foods)
#           [,1]
# Salad         1
# Soup          -1
```

## Exercise

- Fit the model on slide 4 to your simulated data
  1. with dummy coding
  2. with effects coding
- Interpret the fixed parameters and compare the variance components for the random effects for the two models
- What is the interpretation of the intercept for the two models? Why does it differ?

# Treatment coding



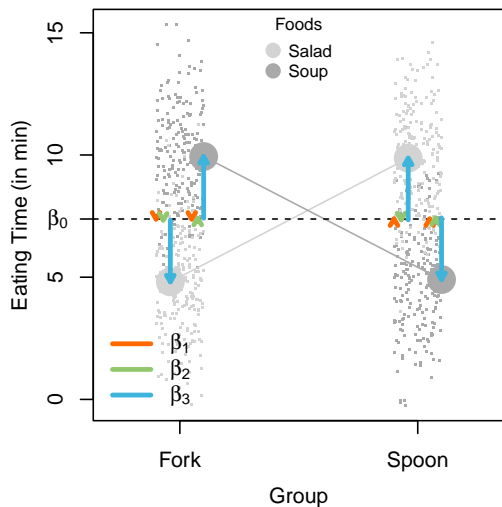
- Means from simulated data

	Salad	Soup
Fork	4.81	9.94
Spoon	9.87	4.90

- Reference categories: Salad and Fork
- Estimated coefficients

	$\beta$	SE	$t$
(Intercept)	4.81	0.21	23.45
UtensilsSpoon	5.06	0.20	25.19
FoodsSoup	5.13	0.20	25.52
UtensilsSpoon:FoodsSoup	-10.10	0.28	-35.52

## Effects coding



- Means from simulated data

	Salad	Soup
Fork	4.81	9.94
Spoon	9.87	4.90

- Estimated coefficients

	$\beta$	SE	$t$
(Intercept)	7.38	0.16	44.97
Utensils1	-0.01	0.07	-0.09
Foods1	-0.04	0.07	-0.57
Utensils1:Foods1	-2.52	0.07	-35.52

## ② Defining contrasts



## General linear model

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \varepsilon_i$$

which corresponds to

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ 1 & x_{31} & x_{32} & \dots & x_{3p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & \dots & x_{Np} \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_N \end{pmatrix}$$

## Example by Schad et al. (2020)

- In order to keep it reasonably complex, let us consider a design with one between-subjects factor with four levels

## Example by Schad et al. (2020)

- In order to keep it reasonably complex, let us consider a design with one between-subjects factor with four levels
- The four levels  $F_1$  to  $F_4$  reflect levels of word frequency with levels low, medium-low, medium-high, and high frequency words, and the dependent variable reflects some response time

## Example by Schad et al. (2020)

- In order to keep it reasonably complex, let us consider a design with one between-subjects factor with four levels
- The four levels  $F_1$  to  $F_4$  reflect levels of word frequency with levels low, medium-low, medium-high, and high frequency words, and the dependent variable reflects some response time
- From four factor levels, we can build three contrasts

## Example by Schad et al. (2020)

- Let us consider 8 subjects for our 4-level between-subjects design

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \end{pmatrix}$$

- For the four means of the factor levels, we get

$$\boldsymbol{\mu} = \mathbf{X} \boldsymbol{\beta}$$

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$

## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$
- $\mu$  is our data and will always be the same

## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$
- $\mu$  is our data and will always be the same
- $\mathbf{X}$  depends on the contrast coding we choose and, therefore, so does  $\beta$



## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$
- $\mu$  is our data and will always be the same
- $\mathbf{X}$  depends on the contrast coding we choose and, therefore, so does  $\beta$
- We will go through different possibilities, all testing different hypotheses

## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$
- $\mu$  is our data and will always be the same
- $\mathbf{X}$  depends on the contrast coding we choose and, therefore, so does  $\beta$
- We will go through different possibilities, all testing different hypotheses
  1. Treatment contrasts

## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$
- $\mu$  is our data and will always be the same
- $\mathbf{X}$  depends on the contrast coding we choose and, therefore, so does  $\beta$
- We will go through different possibilities, all testing different hypotheses
  1. Treatment contrasts
  2. Sum contrasts

## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$
- $\mu$  is our data and will always be the same
- $\mathbf{X}$  depends on the contrast coding we choose and, therefore, so does  $\beta$
- We will go through different possibilities, all testing different hypotheses
  1. Treatment contrasts
  2. Sum contrasts
  3. Helmert contrasts

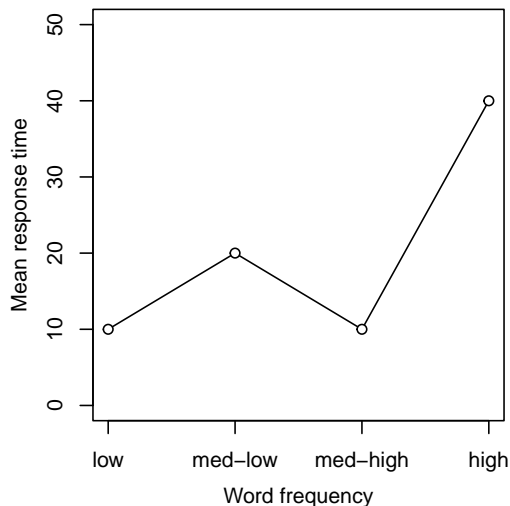
## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$
- $\mu$  is our data and will always be the same
- $\mathbf{X}$  depends on the contrast coding we choose and, therefore, so does  $\beta$
- We will go through different possibilities, all testing different hypotheses
  1. Treatment contrasts
  2. Sum contrasts
  3. Helmert contrasts
  4. Sequential difference contrasts

## Example by Schad et al. (2020)

- The contrast matrix  $\mathbf{X}$  defines how the model parameters are “combined” to estimate  $\mu$
- $\mu$  is our data and will always be the same
- $\mathbf{X}$  depends on the contrast coding we choose and, therefore, so does  $\beta$
- We will go through different possibilities, all testing different hypotheses
  1. Treatment contrasts
  2. Sum contrasts
  3. Helmert contrasts
  4. Sequential difference contrasts
  5. Custom contrasts

## Example by Schad et al. (2020)



```
design <- data.frame(  
  F = factor(c("F1", "F2", "F3", "F4")),  
  mu = c(10, 20, 10, 40))  
  
beta <- lm(mu ~ F, design) |>  
  coef() |>  
  zapsmall()  
  
# Contrast matrix  
mm <- model.matrix(~ F, design)  
  
# Compare to mu  
mm %*% beta
```

## Contrast vs. hypothesis matrix

- In the literature on contrasts, we sometimes find the distinction between contrast and hypothesis matrix (e. g., Schad et al., 2020)



## Contrast vs. hypothesis matrix

- In the literature on contrasts, we sometimes find the distinction between contrast and hypothesis matrix (e. g., Schad et al., 2020)
- When the contrast matrix defines how we add up the parameters to obtain  $\mu$ , then the hypothesis matrix defines how to add up the means to obtain the parameters

$$\mu = \mathbf{X} \beta \quad \Leftrightarrow \quad \beta = \mathbf{X}^{-1} \mu$$

## Contrast vs. hypothesis matrix

- In the literature on contrasts, we sometimes find the distinction between contrast and hypothesis matrix (e. g., Schad et al., 2020)
- When the contrast matrix defines how we add up the parameters to obtain  $\mu$ , then the hypothesis matrix defines how to add up the means to obtain the parameters

$$\mu = \mathbf{X} \beta \quad \Leftrightarrow \quad \beta = \mathbf{X}^{-1} \mu$$

- Hence, the hypothesis matrix is the inverse of the contrast matrix

## Contrast vs. hypothesis matrix

- In the literature on contrasts, we sometimes find the distinction between contrast and hypothesis matrix (e. g., Schad et al., 2020)
- When the contrast matrix defines how we add up the parameters to obtain  $\mu$ , then the hypothesis matrix defines how to add up the means to obtain the parameters

$$\mu = \mathbf{X} \beta \quad \Leftrightarrow \quad \beta = \mathbf{X}^{-1} \mu$$

- Hence, the hypothesis matrix is the inverse of the contrast matrix
- It is sometimes easier to define the hypothesis matrix and then obtain the contrast matrix based on it

# Treatment contrasts

## Contrast Matrix

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

$$\mu_1 = \beta_0$$

$$\mu_2 = \beta_0 + \beta_1$$

$$\mu_3 = \beta_0 + \beta_2$$

$$\mu_4 = \beta_0 + \beta_3$$

## Hypothesis Matrix

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}$$

$$\beta_0 = \mu_1$$

$$\beta_1 = \mu_1 - \mu_2$$

$$\beta_2 = \mu_1 - \mu_3$$

$$\beta_3 = \mu_1 - \mu_4$$

Hypotheses being tested:  $H_{01}: \mu_1 = \mu_2$ ,  $H_{02}: \mu_1 = \mu_3$ ,  $H_{03}: \mu_1 = \mu_4$

## Sum contrasts

Contrast Matrix

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

$$\mu_1 = \beta_0 + \beta_1$$

$$\mu_2 = \beta_0 + \beta_2$$

$$\mu_3 = \beta_0 + \beta_3$$

$$\mu_4 = \beta_0 - \beta_1 - \beta_2 - \beta_3$$

Hypothesis Matrix

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 3/4 & -1/4 & -1/4 & -1/4 \\ -1/4 & 3/4 & -1/4 & -1/4 \\ -1/4 & -1/4 & 3/4 & -1/4 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}$$

$$\beta_0 = \frac{\mu_1 + \mu_2 + \mu_3 + \mu_4}{4}$$

$$\beta_1 = \frac{3}{4}\mu_1 - \frac{1}{4}(\mu_2 + \mu_3 + \mu_4)$$

$$\beta_2 = \frac{3}{4}\mu_2 - \frac{1}{4}(\mu_1 + \mu_3 + \mu_4)$$

$$\beta_3 = \frac{3}{4}\mu_3 - \frac{1}{4}(\mu_1 + \mu_2 + \mu_4)$$

Hypotheses being tested:  $H_{01}: \frac{\mu_2 + \mu_3 + \mu_4}{3} = \mu_1$ ,  $H_{02}: \frac{\mu_1 + \mu_3 + \mu_4}{3} = \mu_2$ ,  $H_{03}: \frac{\mu_1 + \mu_2 + \mu_4}{3} = \mu_3$

# Helmert contrasts

Contrast Matrix

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 0 & 2 & -1 \\ 1 & 0 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

$$\mu_1 = \beta_0 - \beta_1 - \beta_2 - \beta_3$$

$$\mu_2 = \beta_0 + \beta_1 - \beta_2 - \beta_3$$

$$\mu_3 = \beta_0 + 2 \cdot \beta_2 - \beta_3$$

$$\mu_4 = \beta_0 + 3 \cdot \beta_3$$

Hypothesis Matrix

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ -1/2 & 1/2 & 0 & 0 \\ -1/6 & -1/6 & 1/3 & 0 \\ -1/12 & -1/12 & -1/12 & 1/4 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}$$

$$\beta_0 = \frac{\mu_1 + \mu_2 + \mu_3 + \mu_4}{4}$$

$$\beta_1 = -\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2$$

$$\beta_2 = -\frac{1}{6}(\mu_1 + \mu_2) + \frac{1}{3}\mu_3$$

$$\beta_3 = -\frac{1}{12}(\mu_1 + \mu_2 + \mu_3) + \frac{1}{4}\mu_4$$

Hypotheses being tested:  $H_{01}: \mu_1 = \mu_2$ ,  $H_{02}: \frac{\mu_1 + \mu_2}{2} = \mu_3$ ,  $H_{03}: \frac{\mu_1 + \mu_2 + \mu_3}{3} = \mu_4$

# Sequential difference contrasts

Contrast Matrix

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & -3/4 & -1/2 & -1/4 \\ 1 & 1/4 & -1/2 & -1/4 \\ 1 & 1/4 & 1/2 & -1/4 \\ 1 & 1/4 & 1/2 & 3/4 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

Hypothesis Matrix

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}$$

$$\mu_1 = \beta_0 - \frac{3}{4}\beta_1 - \frac{1}{2}\beta_2 - \frac{1}{4}\beta_3$$

$$\mu_2 = \beta_0 + \frac{1}{4}\beta_1 - \frac{1}{2}\beta_2 - \frac{1}{4}\beta_3$$

$$\mu_3 = \beta_0 + \frac{1}{4}\beta_1 + \frac{1}{2}\beta_2 - \frac{1}{4}\beta_3$$

$$\mu_4 = \beta_0 + \frac{1}{4}\beta_1 + \frac{1}{2}\beta_2 + \frac{3}{4}\beta_3$$

$$\beta_0 = \frac{\mu_1 + \mu_2 + \mu_3 + \mu_4}{4}$$

$$\beta_1 = \mu_2 - \mu_1$$

$$\beta_2 = \mu_3 - \mu_2$$

$$\beta_3 = \mu_4 - \mu_3$$

Hypotheses being tested:  $H_{01}: \mu_1 = \mu_2$ ,  $H_{02}: \mu_2 = \mu_3$ ,  $H_{03}: \mu_3 = \mu_4$

## Custom contrasts

- All contrasts that we looked at so far are standard and directly available in R



## Custom contrasts

- All contrasts that we looked at so far are standard and directly available in R
- It is also possible to construct contrasts testing custom hypotheses

## Custom contrasts

- All contrasts that we looked at so far are standard and directly available in R
- It is also possible to construct contrasts testing custom hypotheses
- Let us assume for our example, we want to test if the first two means ( $F_1$  and  $F_2$ ) are identical, but that means for levels  $F_3$  and  $F_4$  increase linearly

## Custom contrasts

- All contrasts that we looked at so far are standard and directly available in R
- It is also possible to construct contrasts testing custom hypotheses
- Let us assume for our example, we want to test if the first two means ( $F_1$  and  $F_2$ ) are identical, but that means for levels  $F_3$  and  $F_4$  increase linearly
- Hence, we assume a potential hypothetical outcome, such as  $\mu_1 = 10$ ,  $\mu_2 = 10$ ,  $\mu_3 = 20$ , and  $\mu_4 = 30$

## Custom contrasts

- All contrasts that we looked at so far are standard and directly available in R
- It is also possible to construct contrasts testing custom hypotheses
- Let us assume for our example, we want to test if the first two means ( $F_1$  and  $F_2$ ) are identical, but that means for levels  $F_3$  and  $F_4$  increase linearly
- Hence, we assume a potential hypothetical outcome, such as  $\mu_1 = 10$ ,  $\mu_2 = 10$ ,  $\mu_3 = 20$ , and  $\mu_4 = 30$
- These means can easily be turned into a contrast by centering them and then maybe do some scaling

$\mu_1$	$\mu_2$	$\mu_3$	$\mu_4$
10	10	20	30
-7.5	-7.5	2.5	12.5
-3	-3	1	5

## Custom contrasts

Contrast Matrix

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 1 & -3 \\ 1 & 1 \\ 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

$$\mu_1 = \beta_0 - 3\beta_1$$

$$\mu_2 = \beta_0 - 3\beta_1$$

$$\mu_3 = \beta_0 + \beta_1$$

$$\mu_4 = \beta_0 + 5\beta_1$$

Hypothesis Matrix

$$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ -0.068 & -0.068 & 0.023 & 0.114 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}$$

$$\beta_0 = \frac{\mu_1 + \mu_2 + \mu_3 + \mu_4}{4}$$

$$\beta_1 = -\frac{3}{44}\mu_1 - \frac{3}{44}\mu_2 + \frac{1}{44}\mu_3 + \frac{5}{44}\mu_4$$

Hypothesis being tested:  $H_{01}: -3\mu_1 - 3\mu_2 + \mu_3 + 5\mu_4 = 0$

## Custom contrasts

### Contrast Matrix

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & -3 & -0.511 & -0.533 \\ 1 & -3 & 0.131 & 0.727 \\ 1 & 1 & 0.760 & -0.387 \\ 1 & 5 & -0.380 & 0.194 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

- It is often a good idea to add orthogonal contrasts, so any variance in your data that is independent of your contrast can be accounted for
- R will automatically add orthogonal contrasts if you provide less than  $n - 1$  contrasts for a factor with  $n$  levels

### Hypothesis Matrix

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ -0.068 & -0.068 & 0.023 & 0.114 \\ -0.511 & 0.131 & 0.760 & -0.380 \\ -0.533 & 0.727 & -0.387 & 0.194 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}$$

Hypothesis being tested:  $H_{01}: -3\mu_1 - 3\mu_2 + \mu_3 + 5\mu_4 = 0$

## Simulate data

```
set.seed(1700)

n <- 20
dat <- data.frame(id = 1:n,
                  F = factor(rep(c("F1", "F2", "F3", "F4"),
                                times = n / 4)),
                  DV = rnorm(n, mean = c(10, 20, 10, 40), sd = 5))

datm <- aggregate(DV ~ F, dat, mean)
aggregate(DV ~ F, dat, sd)

lattice::xyplot(DV ~ F, dat, type = c("a", "p"))
```

## Exercise

- Now fit linear models with all contrasts we looked at
  1. Treatment contrasts
  2. Sum contrasts
  3. Helmert contrasts
  4. Sequential difference contrasts
  5. Custom contrasts
- Interpret the fixed parameters
- Predict data with each model and compare the results
- Look again at the intercepts for all models and compare



## References

- Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, 125, 104334.  
<https://doi.org/10.1016/j.jml.2022.104334>
- Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110, 104038. <https://doi.org/10.1016/j.jml.2019.104038>