Setting up group models
Part 1

NITP, 2011

What is coming up

- Crash course in setting up models
  - 1-sample and 2-sample t-tests
  - Paired t-tests
  - ANOVA!
- Mean centering covariates
- Identifying rank deficient matrices

Recall

- GLM is flexible
  - One Sample T Test
  - ANOVA
  - Two sample T Test
  - Paired T test
- What do the models look like?

1-Sample T Test

$$X\beta = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \beta$$

Overall mean

$$H_0 : c\beta = 0 \quad \text{where} \quad c = [1]$$

2-Sample T Test

$$\begin{pmatrix} A_1 \\ A_1 \\ A_1 \\ A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_2 \\ A_2 \\ A_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

Mean group 1

Mean group 2

$$H_0 : c\beta = 0 \quad \text{where} \quad c = [1 \quad -1]$$

2-Sample T Test

OR

$$\begin{pmatrix} A_1 \\ A_1 \\ A_1 \\ A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_2 \\ A_2 \\ A_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

Understanding a model

- If you're unsure about a model or the contrasts
  - Plug in numbers
  - Look at graphs (fMRI data)
- Always ask yourself if your model is doing what you want it to

For example…

- For the 2 sample T test
  - Set $\beta_1 = 3 \quad \beta_2 = 5$
  - Then G1=8 and G2=3
  - So β_1 is the mean of group 2 and β_2 is the difference between the groups
  - What are the contrasts to test
    - Mean of G2 $c = [1 \quad 0]$
    - Mean of G1
    - G1-G2

$$X\beta = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

For example…

- For the 2 sample T test
  - Set $\beta_1 = 3 \quad \beta_2 = 5$
  - Then G1=8 and G2=3
  - So β_1 is the mean of group 2 and β_2 is the difference between the groups
  - What are the contrasts to test
    - Mean of G2 $c = [1 \quad 0]$
    - Mean of G1 $c = [1 \quad 1]$
    - G1-G2

$$X\beta = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

For example…

- For the 2 sample T test
  - Set $\beta_1 = 3 \quad \beta_2 = 5$
  - Then G1=8 and G2=3
  - So β_1 is the mean of group 2 and β_2 is the difference between the groups
  - What are the contrasts to test
    - Mean of G2 $c = [1 \quad 0]$
    - Mean of G1 $c = [1 \quad 1]$
    - G1-G2 $c = [0 \quad 1]$

$$X\beta = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

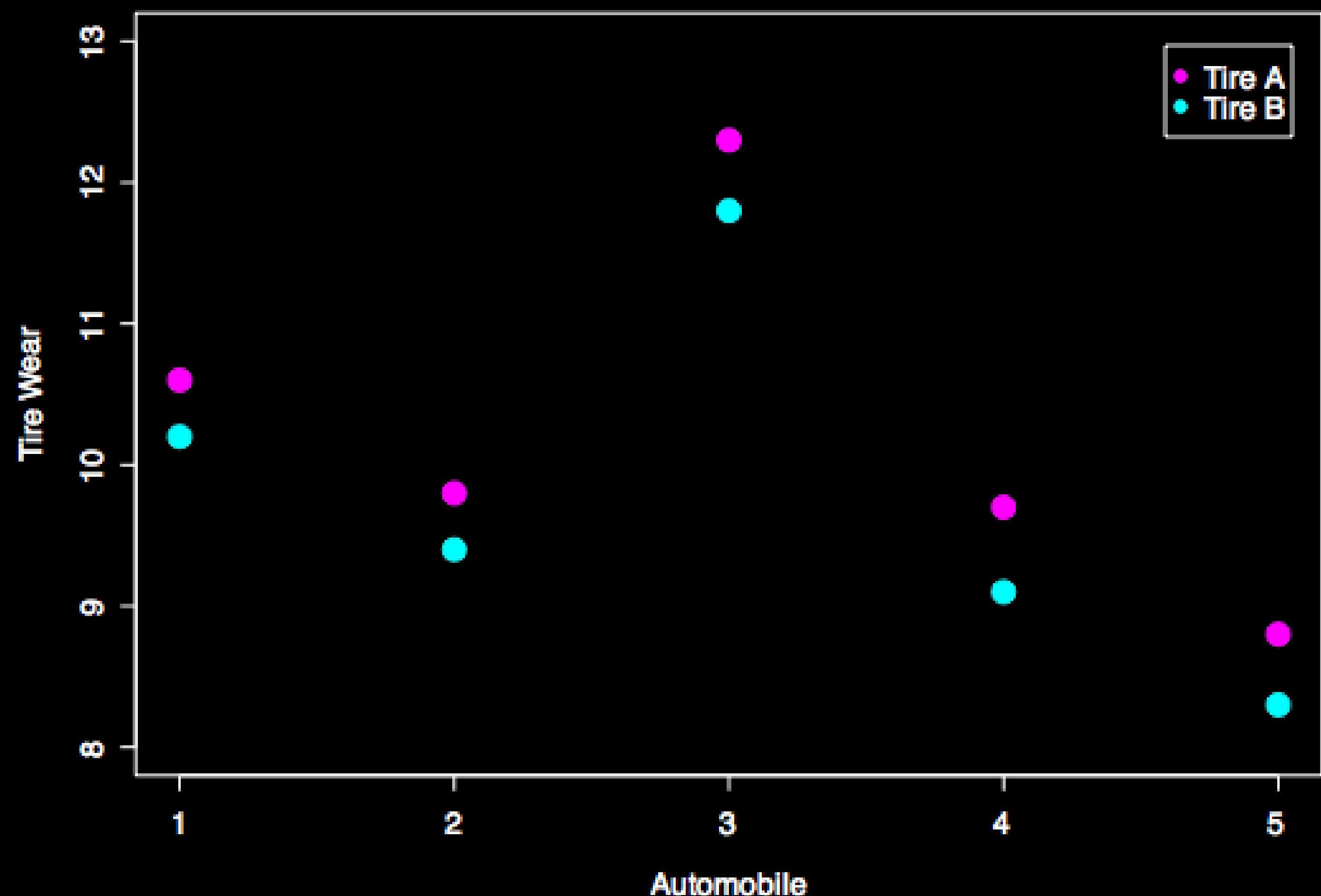
Paired T Test

- A common mistake is to use a 2-sample t test instead of a paired test
- Tire example

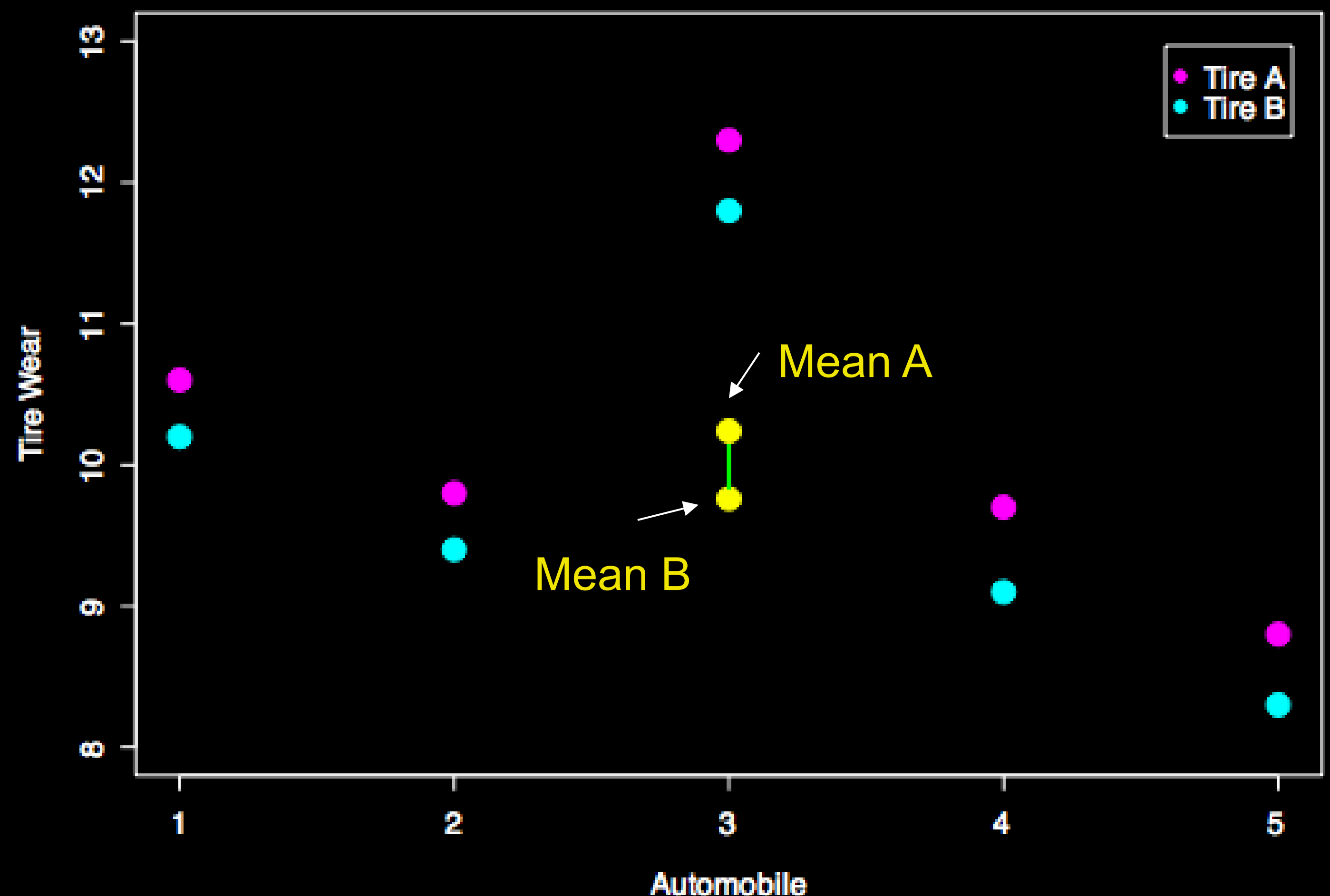
Automobile	Tire A	Tire B
1	10.6	10.2
2	9.8	9.4
3	12.3	11.8
4	9.7	9.1
5	8.8	8.3

- 2-sample T test $p=0.58$
- Paired T test $p<0.001$

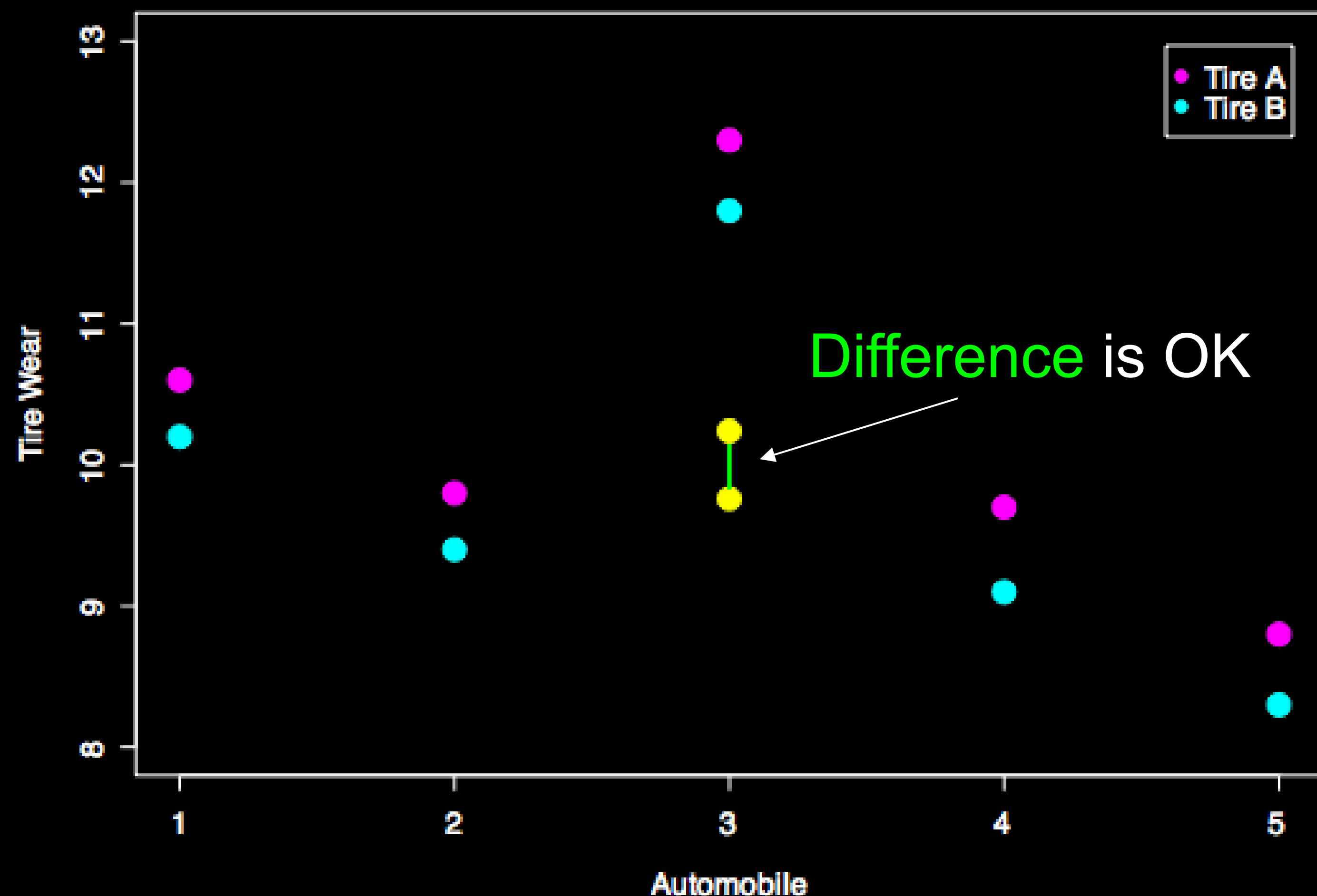
Why so different?



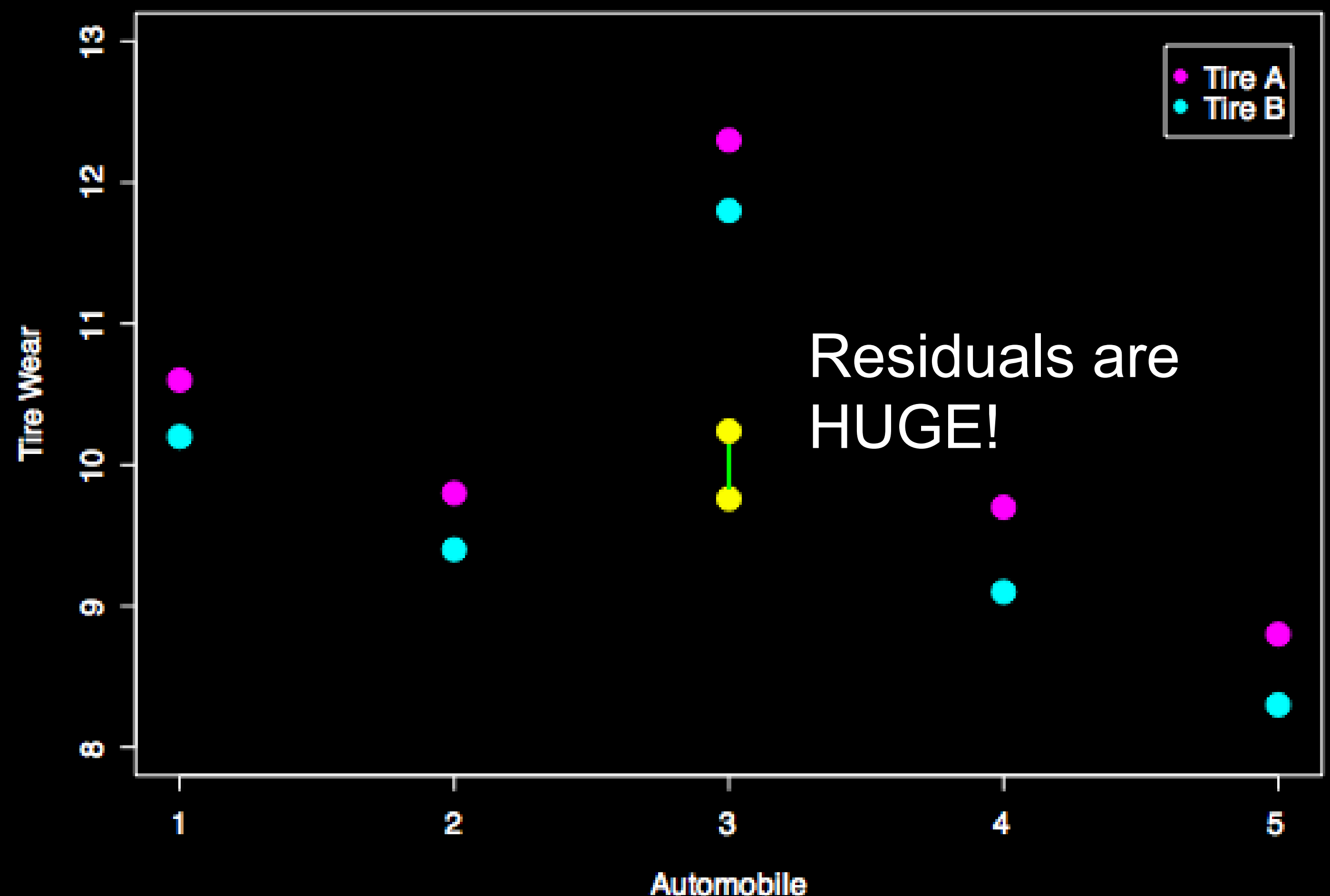
Why so different?



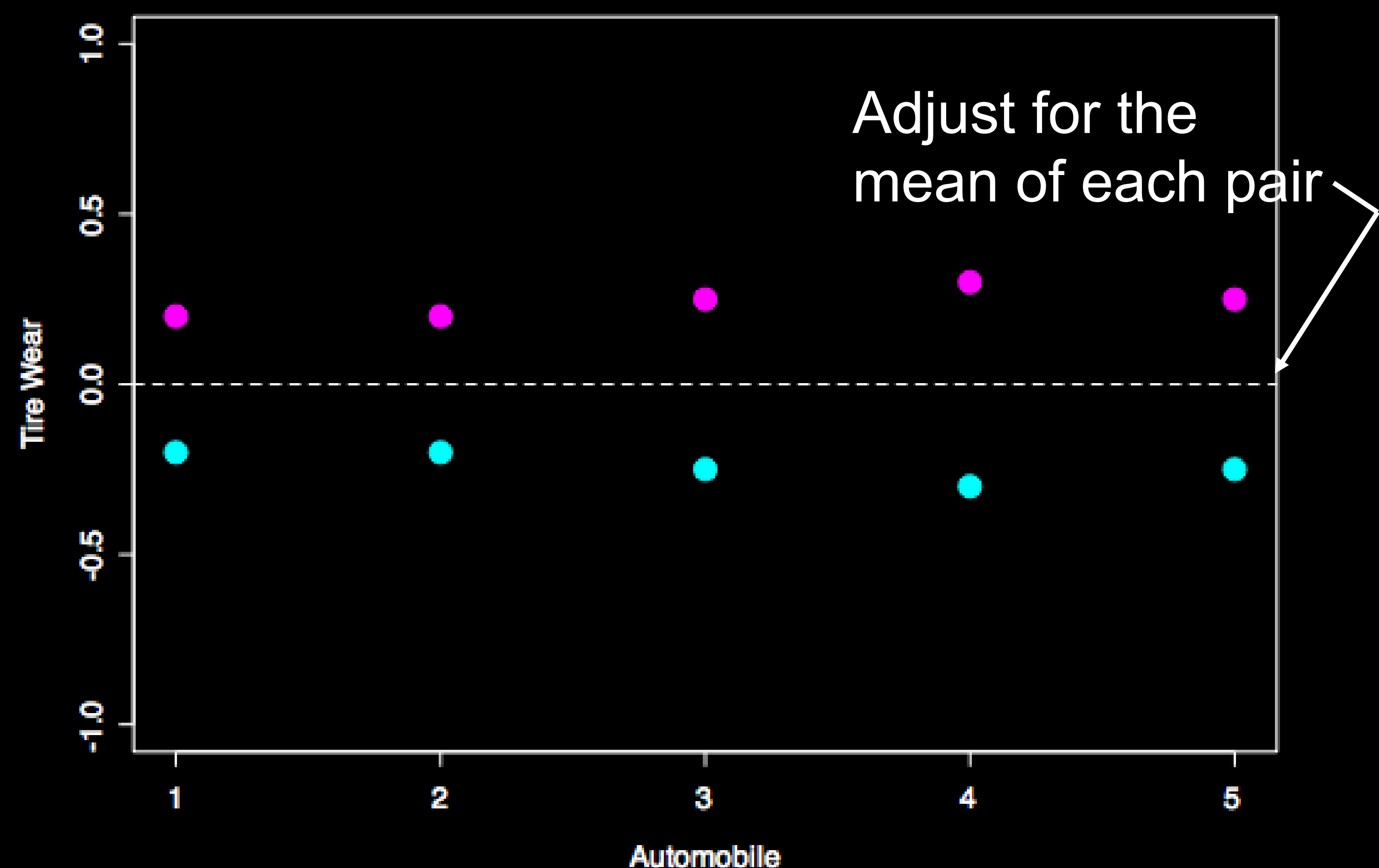
Why so different?



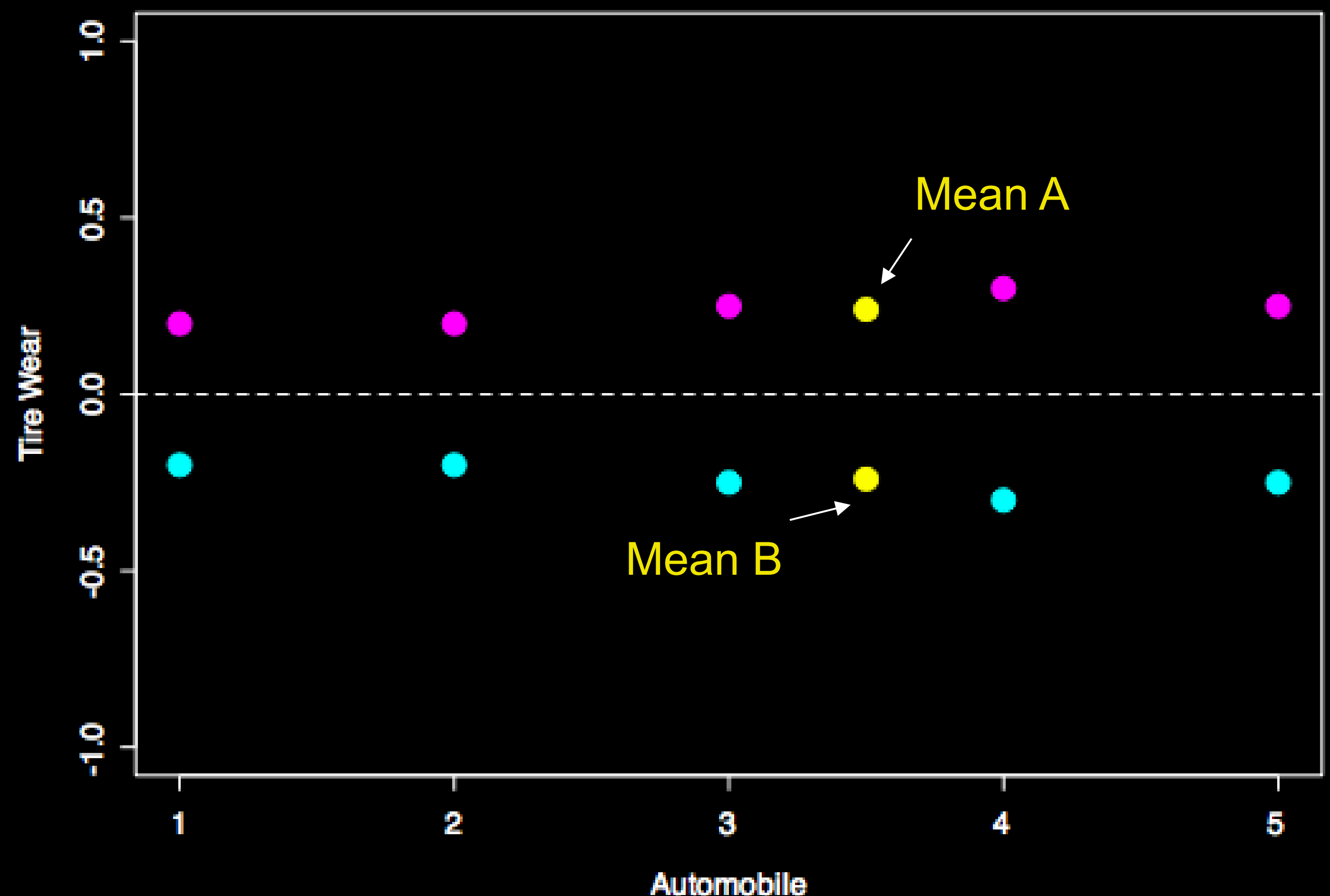
Why so different?



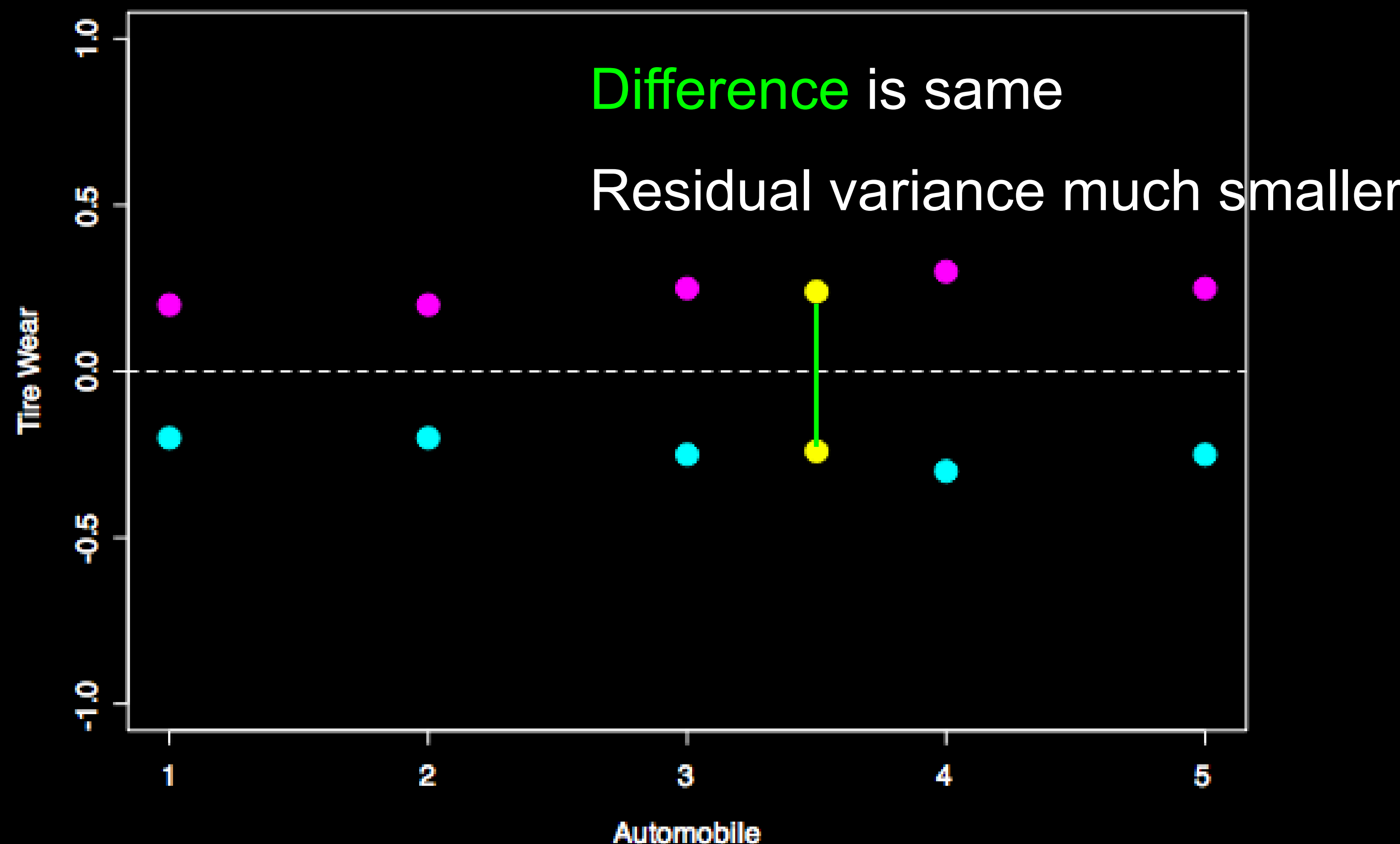
Paired T Test



Paired T Test



Paired T Test



Paired T Test GLM

$$\begin{pmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \\ A_3 \\ B_3 \\ A_4 \\ B_4 \\ A_5 \\ B_5 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

Difference

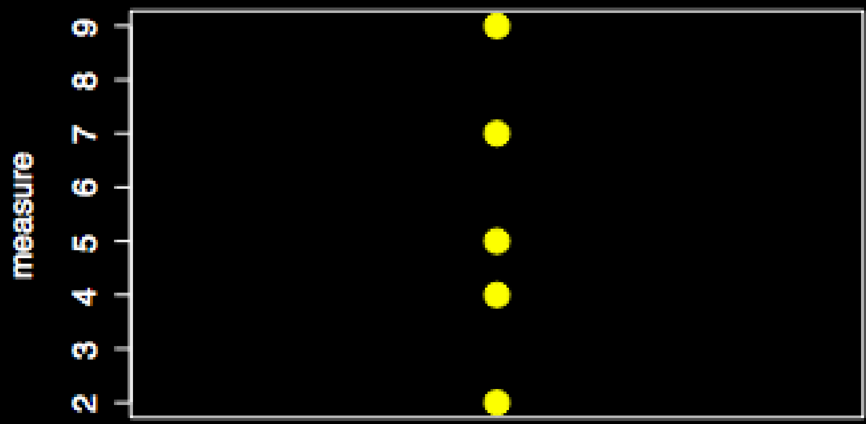
Mean of each pair

$$H_0 : \text{Paired difference} = 0$$
$$H_0 : c\beta = 0, \quad c = [1\ 0\ 0\ 0\ 0\ 0]$$

ANOVA

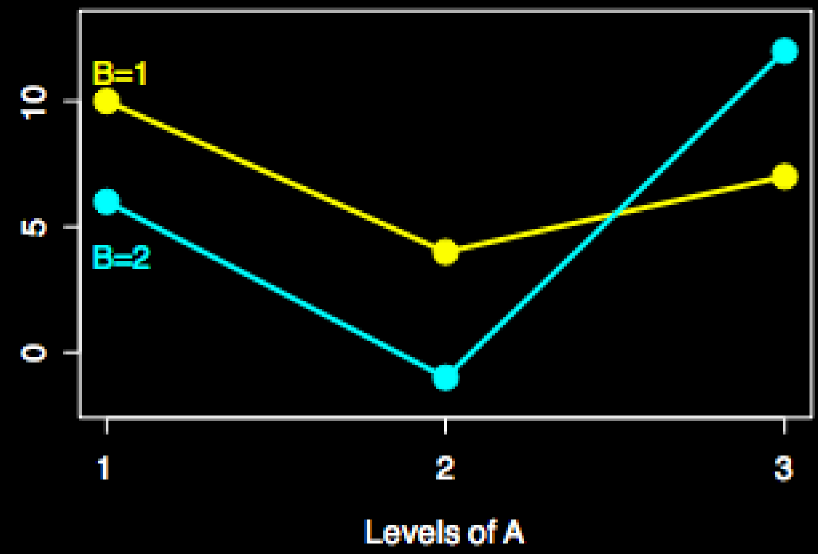
1-way ANOVA

μ_1
μ_2
$\vdots$
μ_N



2-way ANOVA

	B	
A	μ_{11}	μ_{12}
	μ_{21}	μ_{22}
	μ_{31}	μ_{32}



Modeling ANOVA with GLM

- Cell means model
  - Model a mean for each “cell”

μ_1
μ_2
$\vdots$
μ_N

	B	
	μ_{11}	μ_{12}
A	μ_{21}	μ_{22}
	μ_{31}	μ_{32}

Modeling ANOVA with GLM

- Factor effects model
  - Model each factor as a set of regressors
  - One regressor for overall mean, other regressors describe how factor effects relate to the overall mean

Modeling ANOVA with GLM

- Factor effects model
  - Model the overall mean and have regressors for each factor
  - Hypothesis tests from this model correspond to standard ANOVA hypotheses
    - Eg, if group (2 levels) and stimulus type (3 levels) are modeled you can test for a main group effect, main stimulus type effect and interaction effect

1 Way ANOVA - Cell Means

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$H_0 : G2 - G3 = 0$$

1 Way ANOVA - Cell Means

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$H_0 : G2 - G3 = 0$$

$$H_0 : c\beta = 0 \quad \text{where} \quad c = [0 \;\; 1 \quad -1 \;\; 0]$$

1 Way ANOVA - Cell Means

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$H_0 : G1 = G2 = G3 = G4 = 0$$

1 Way ANOVA - Cell Means

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$H_0 : G1 = G2 = G3 = G4 = 0$$

$$H_0 : c\beta = 0 \quad \text{where} \quad c = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1 Way ANOVA factor effects

- Always start with overall mean (column of 1s)
- Number of regressors for a specific factor is number of levels – 1
  - Why?

1 Way ANOVA factor effects

- Always start with overall mean (column of 1s)
- Number of regressors for a specific factor is number of levels – 1
  - Why?
  - If I know the sum of 4 numbers is 10, then I only need to know 3 of the numbers to figure out what the fourth is.

1 Way ANOVA - Factor Effects

- In general
  - # of regressors for a factor = # levels-1
  - Factor with 4 levels
    - $$X_i = \begin{array}{rl} 1 & \text{if case from level i} \\ -1 & \text{if case from level 4} \\ 0 & \text{otherwise} \end{array}$$

1 Way ANOVA - Factor Effects

- In general
  - # of regressors for a factor = # levels-1
  - Factor with 4 levels
    - $$X_i = \begin{array}{rl} 1 & \text{if case from level i} \\ -1 & \text{if case from level 4} \\ 0 & \text{otherwise} \end{array}$$

Note: I'm assuming a balanced design!!

1 Way ANOVA - Factor Effects

mean

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

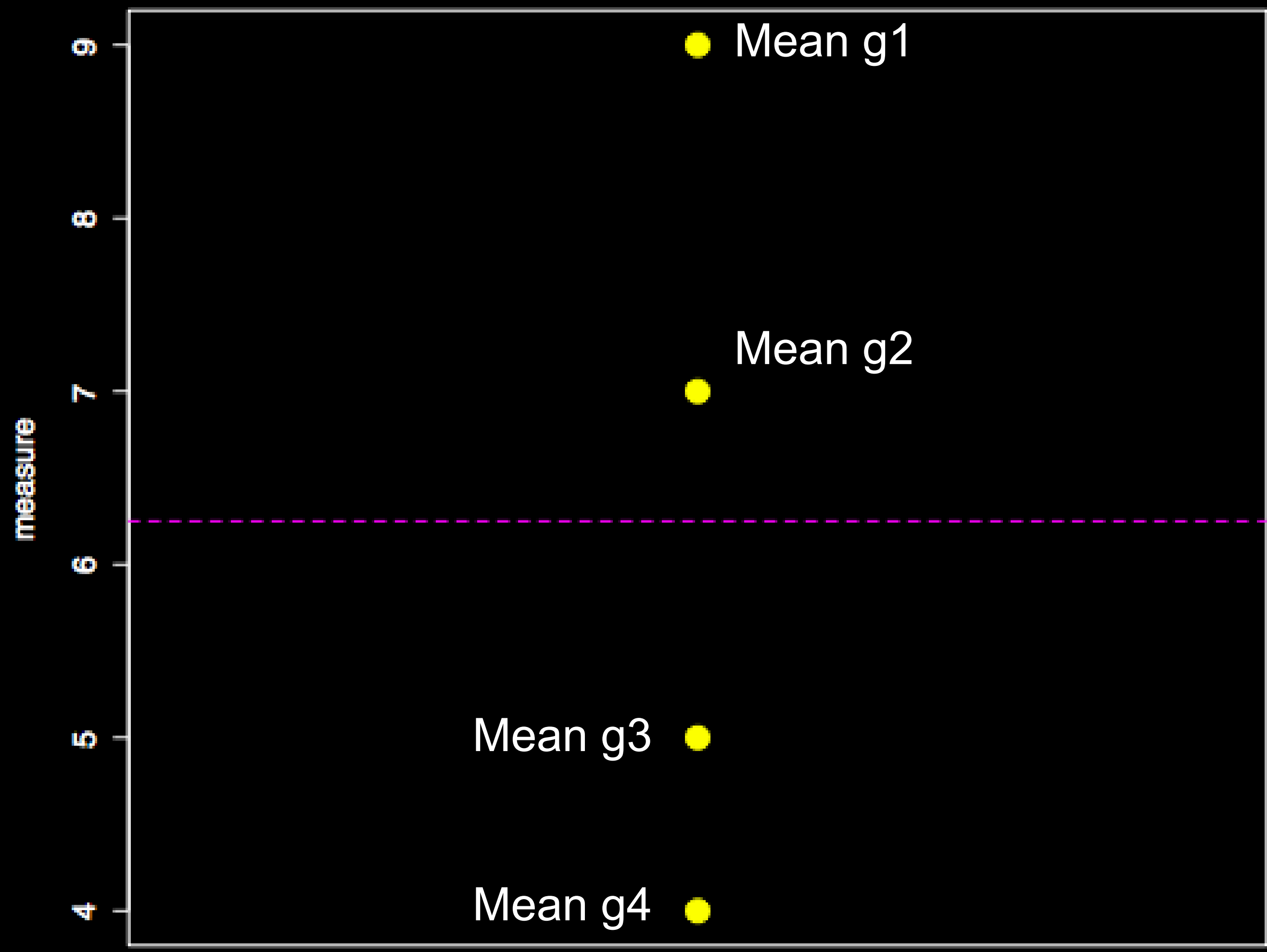
1 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

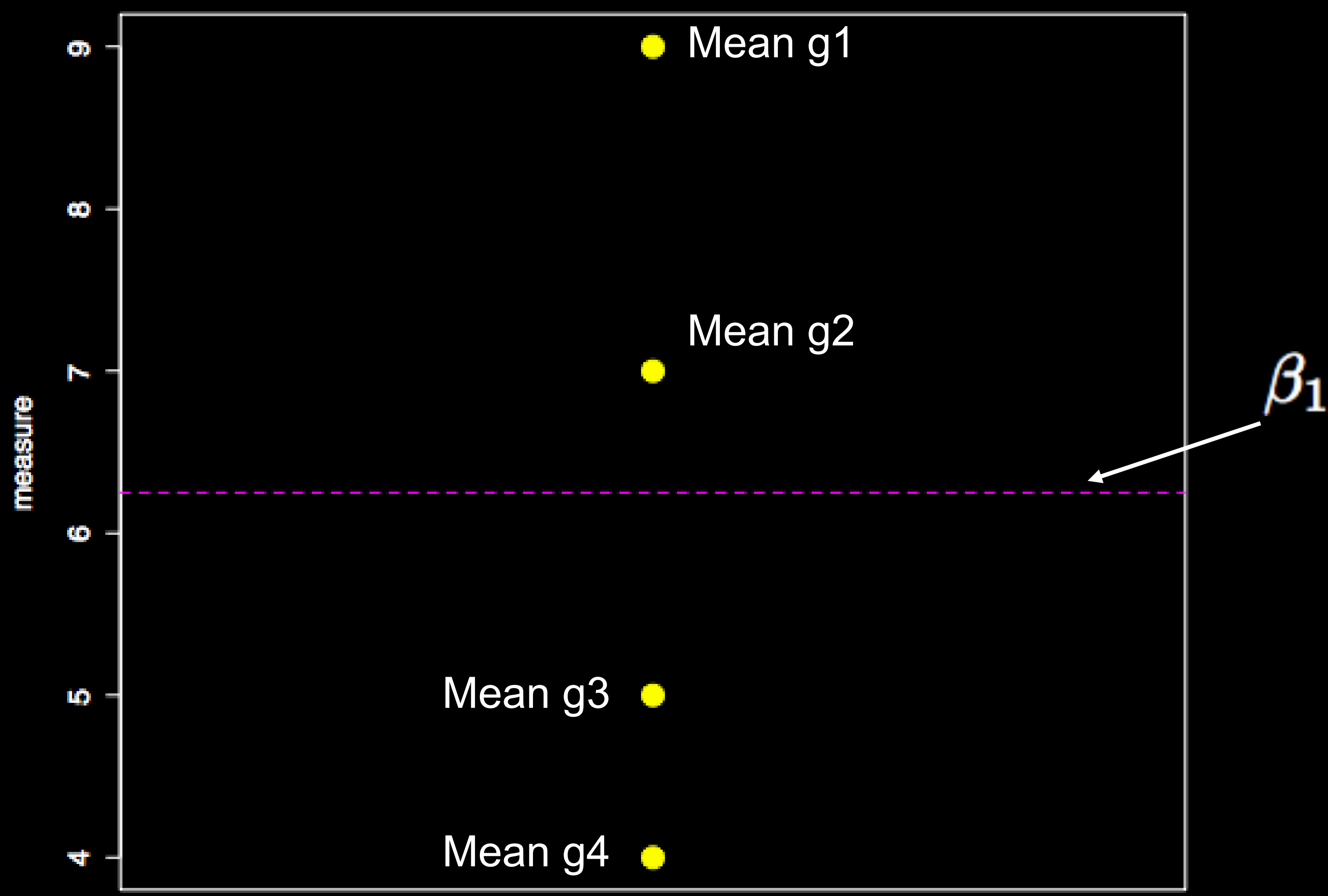
mean

$$G1 = \beta_1 + \beta_2$$
$$G2 = \beta_1 + \beta_3$$
$$G3 = \beta_1 + \beta_4$$
$$G4 = \beta_1 - \beta_2 - \beta_3 - \beta_4$$

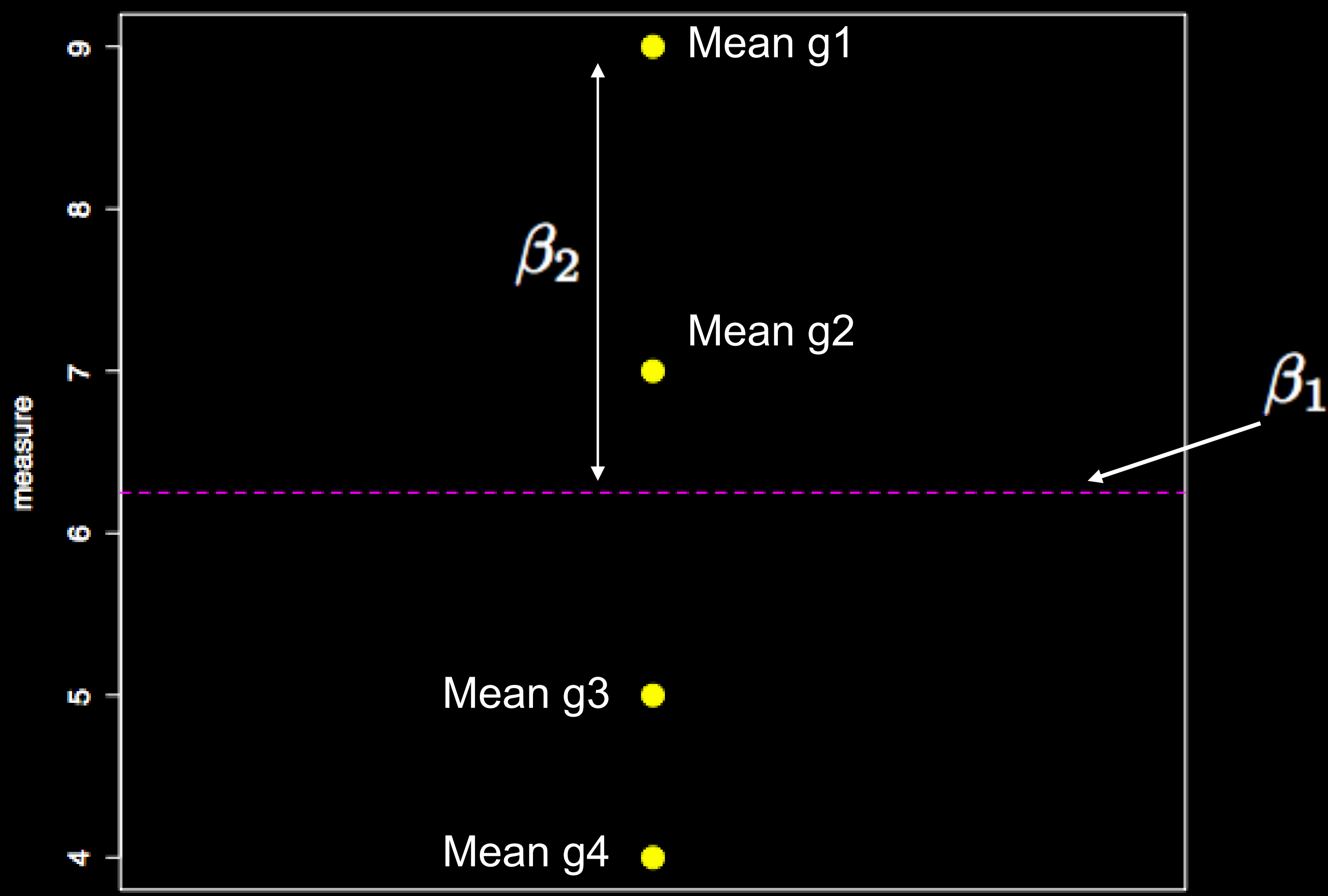
1 Way ANOVA - Factor Effects



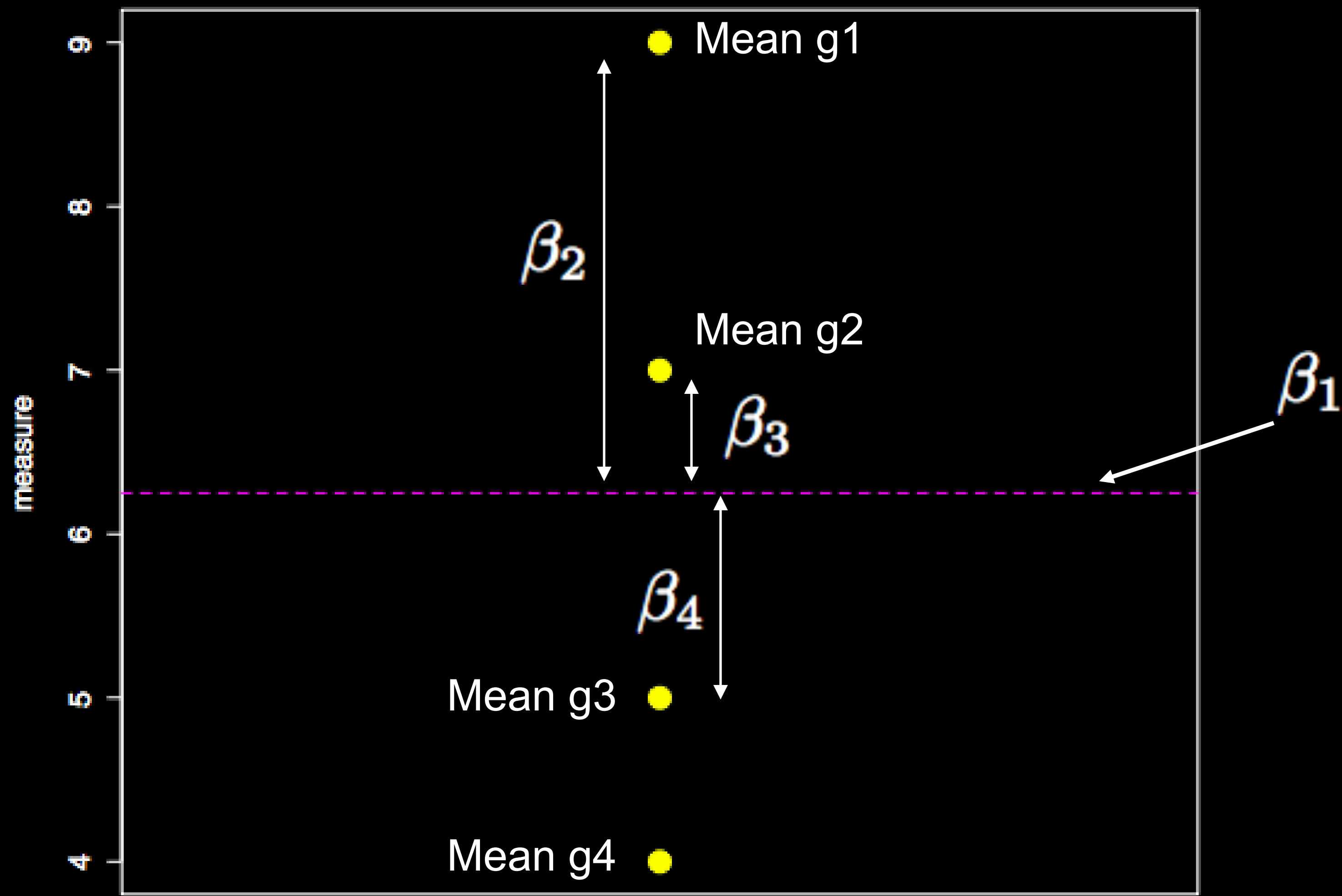
1 Way ANOVA - Factor Effects



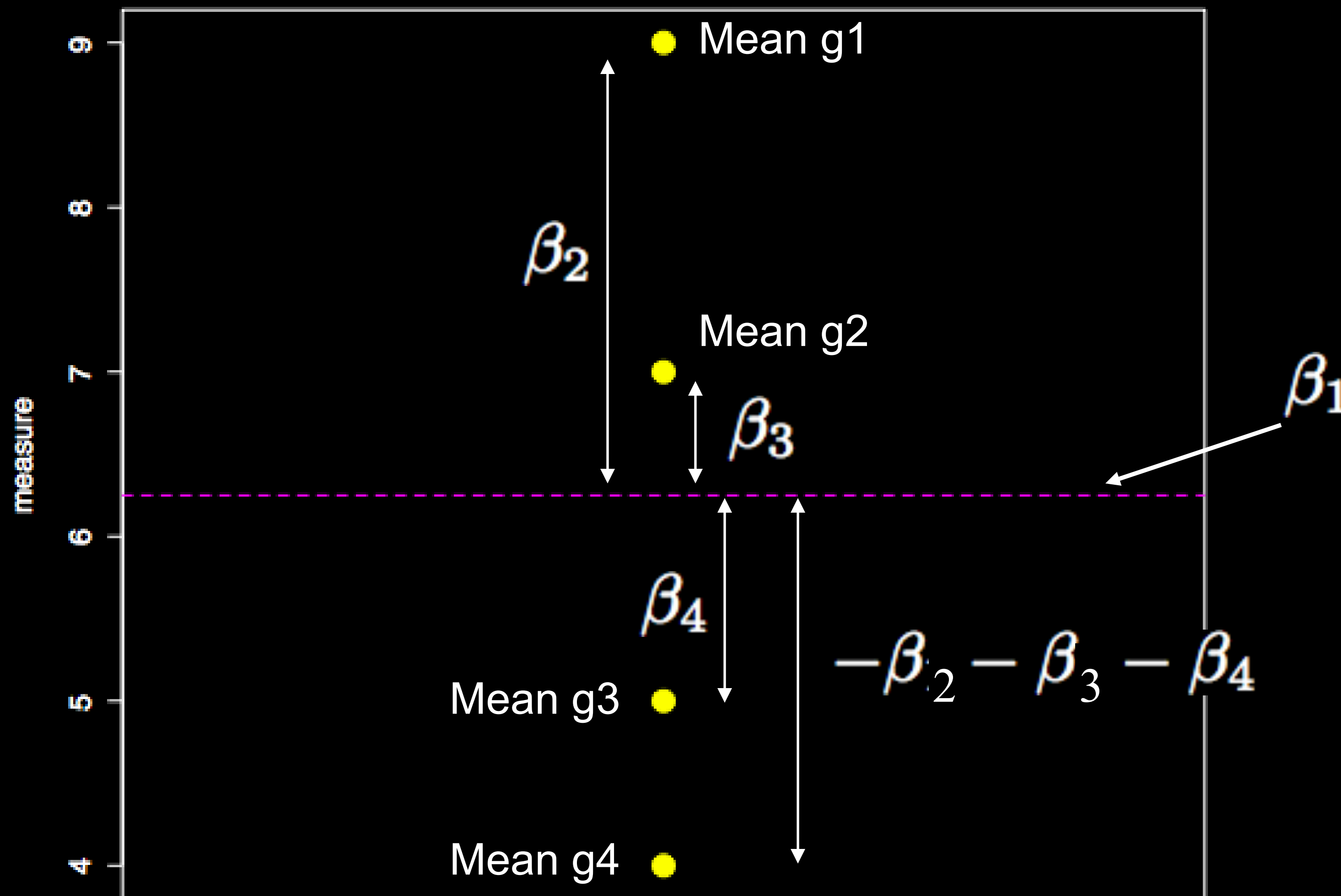
1 Way ANOVA - Factor Effects



1 Way ANOVA - Factor Effects



1 Way ANOVA - Factor Effects



1 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

mean

Test main effect of factor A?

$$c = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$H_0 : \text{mean of G1} = 0$$

1 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$H_0 : \text{mean of G1} = 0$$

$$H_0 : c\beta = 0 \quad \text{where} \quad c = [1 \;\; 1 \;\; 0 \;\; 0]$$

1 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$H_0 : G1 - G4 = 0$$

1 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1 \\ A_1 \\ A_2 \\ A_2 \\ A_3 \\ A_3 \\ A_4 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$H_0 : G1 - G4 = 0$$

$$c = (1 \;\; 1 \;\; 0 \;\; 0) - (1 \;\; -1 \;\; -1 \;\; -1) = (0 \;\; 2 \;\; 1 \;\; 1)$$

2 Way ANOVA (3x2)

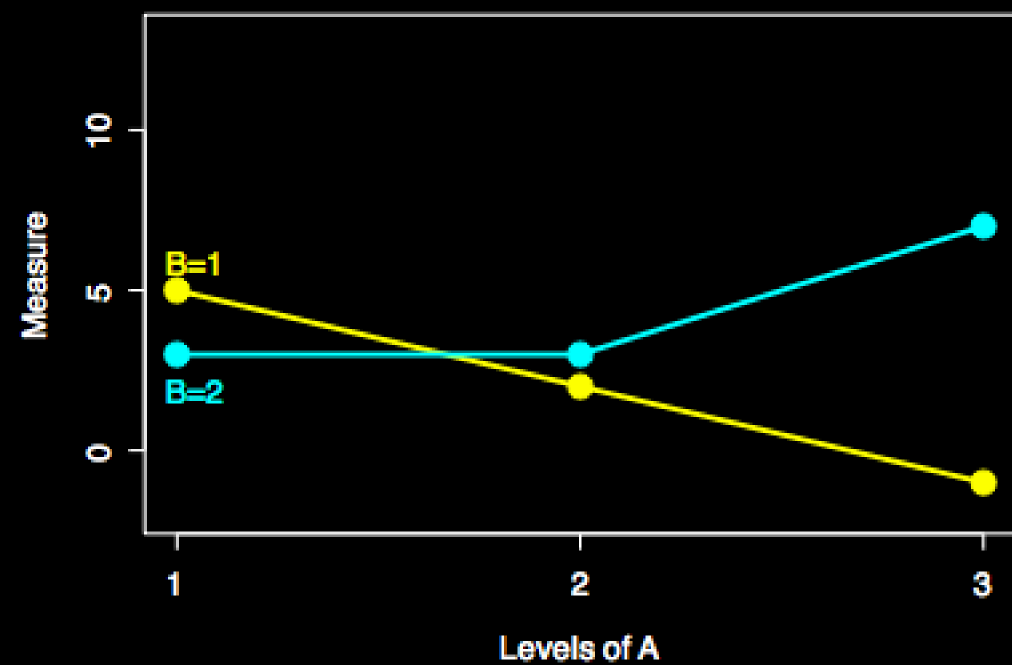
$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{main factor A effect} = 0$$

2 Way ANOVA (3x2)

$$H_0 : \text{main factor A effect} = 0$$

	B1	B2	
A1	5	3	8
A2	2	3	5
A3	-1	7	6
	6	13	19



No effect means the marginals would be the same

Null: A1=A2=A3 equivalently A1-A3=0 and A2-A3=0

2 Way ANOVA (3x2)

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{main factor A effect} = 0$$

2 Way ANOVA (3x2)

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{main factor A effect} = 0$$

$$H_0 : c\beta = 0 \quad \text{where} \quad c = \begin{pmatrix} 1 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{pmatrix}$$

2 Way ANOVA (3x2)

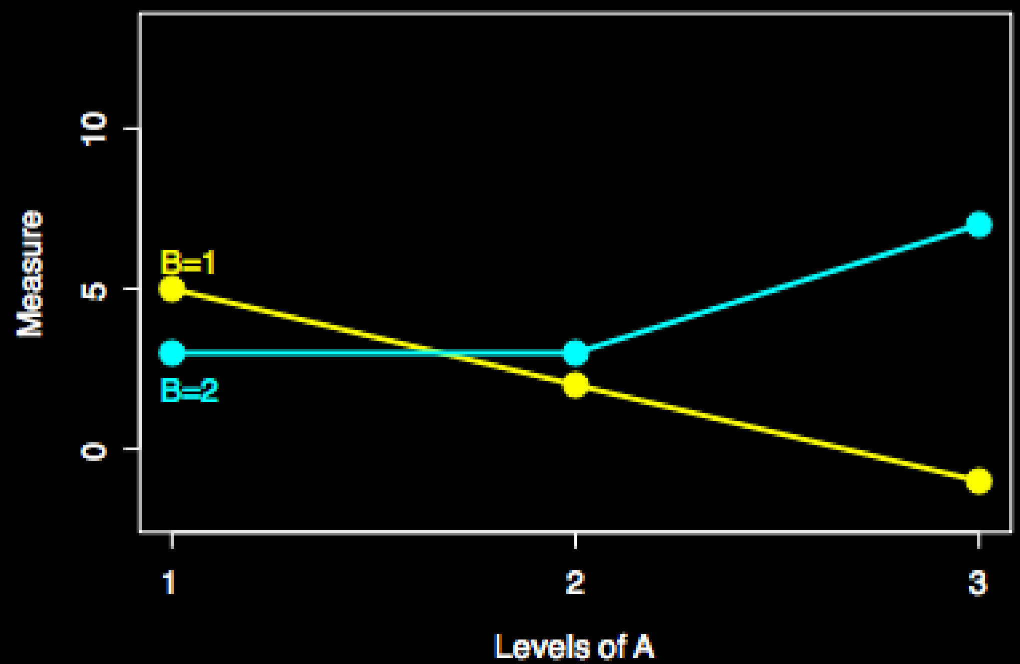
$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{interaction effect} = 0$$

2 Way ANOVA (3x2)

$$H_0 : \text{interaction effect} = 0$$

	B1	B2	
A1	5	3	8
A2	2	3	5
A3	-1	7	6
	6	13	19

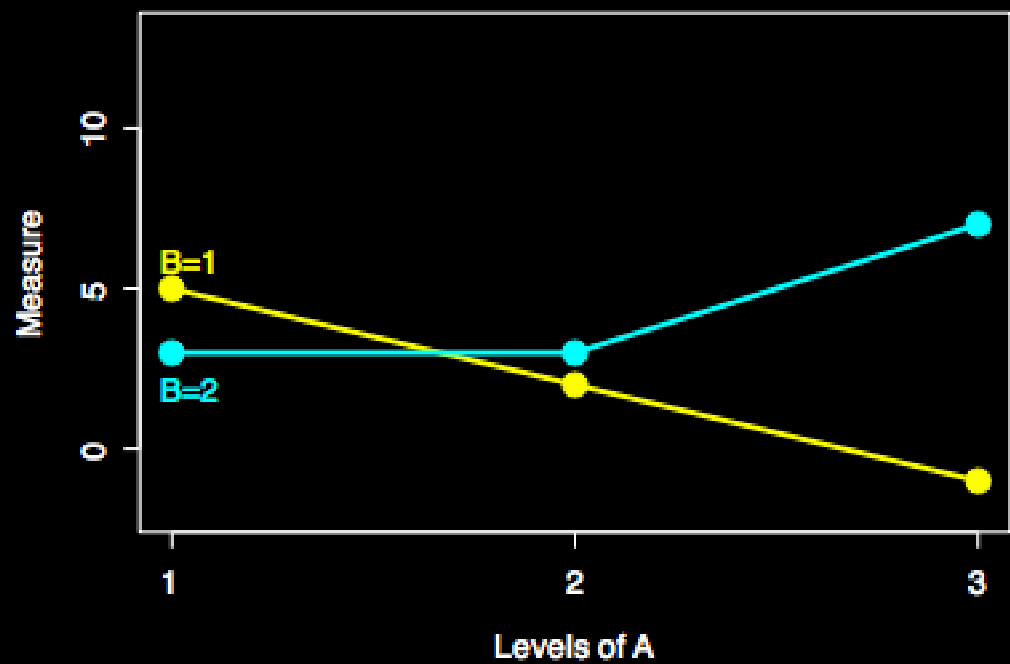


No effect means the lines would be parallel

2 Way ANOVA (3x2)

$$H_0 : \text{interaction effect} = 0$$

	B1	B2	
A1	5	3	8
A2	2	3	5
A3	-1	7	6
	6	13	19



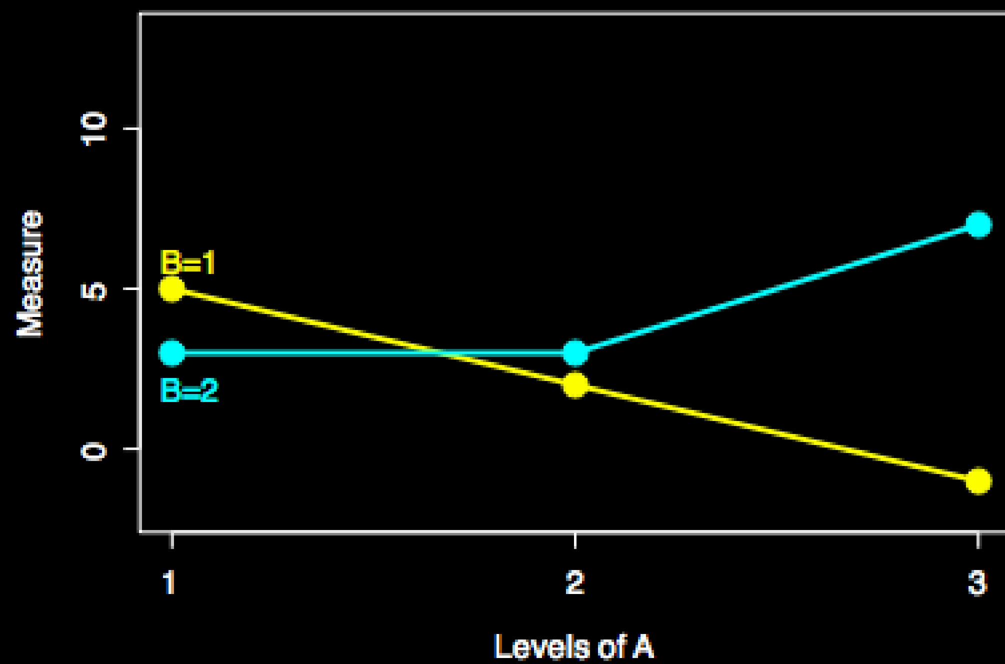
No effect means the lines would be parallel

A1B1-A1B2=A2B1-A2B2=A3B1-A3B2

2 Way ANOVA (3x2)

$$H_0 : \text{interaction effect} = 0$$

	B1	B2	
A1	5	3	8
A2	2	3	5
A3	-1	7	6
	6	13	19



No effect means the lines would be parallel

A1B1-A1B2-A3B1+A3B2 = A2B1-A2B2-A3B1+A3B2 = 0

2 Way ANOVA (3x2)

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{interaction effect} = 0$$

$$H_0 : c\beta = 0 \quad \text{where} \quad c = \begin{pmatrix} 1 & -1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{pmatrix}$$

By the way…

- If you feel confused by that last example, that was my goal!

2 Way ANOVA - Factor Effects

- Recall for factor effects, a factor with n levels has regressors set up like

$$X_i = \begin{array}{rl} 1 & \text{if case from level i} \\ -1 & \text{if case from level n} \\ 0 & \text{otherwise} \end{array}$$

- A has 3 levels, so 2 regressors
- B has 2 levels, so 1 regressor

2 Way ANOVA - Factor Effects

$$
\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix}
=
\begin{pmatrix}
1 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 0 & -1 & -1 & 0 \\
1 & 1 & 0 & -1 & -1 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & -1 & 0 & -1 \\
1 & 0 & 1 & -1 & 0 & -1 \\
1 & -1 & -1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & -1 & -1 \\
1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & -1 & 1 & 1
\end{pmatrix}
\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}
$$

A B AB

2 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{main factor A effect} = 0$$

2 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{main factor A effect} = 0$$

$$H_0 : c\beta = 0 \quad \text{where} \quad c = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

2 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{interaction effect} = 0$$

2 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{interaction effect} = 0$$

$$H_0 : c\beta = 0 \quad \text{where} \quad c = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

2 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{mean cell } A_1B_1 = 0$$

2 Way ANOVA - Factor Effects

$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}$$

$$H_0 : \text{mean cell } A_1B_1 = 0$$

$$H_0 : c\beta = 0 \quad \text{where} \quad c = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

For more examples

- The FSL folks have a bunch of great examples
  - http://www.fmrib.ox.ac.uk/fsl/feat5/detail.html
- Check the FSL help list regularly
  - Subscribe at jiscmail
  - Often others have already asked your questions!

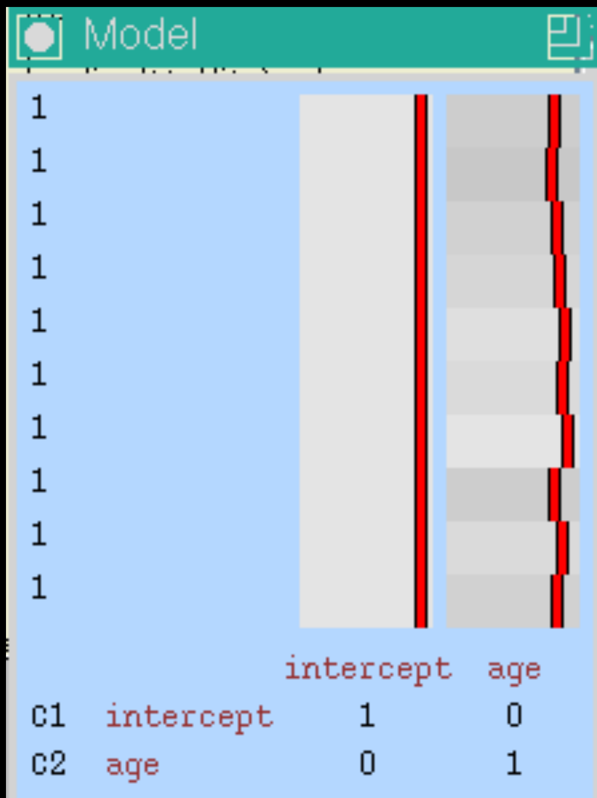
Mean centering covariates

- Why do we mean center?
- When should we mean center?
- What does it do to the parameter estimate interpretation?

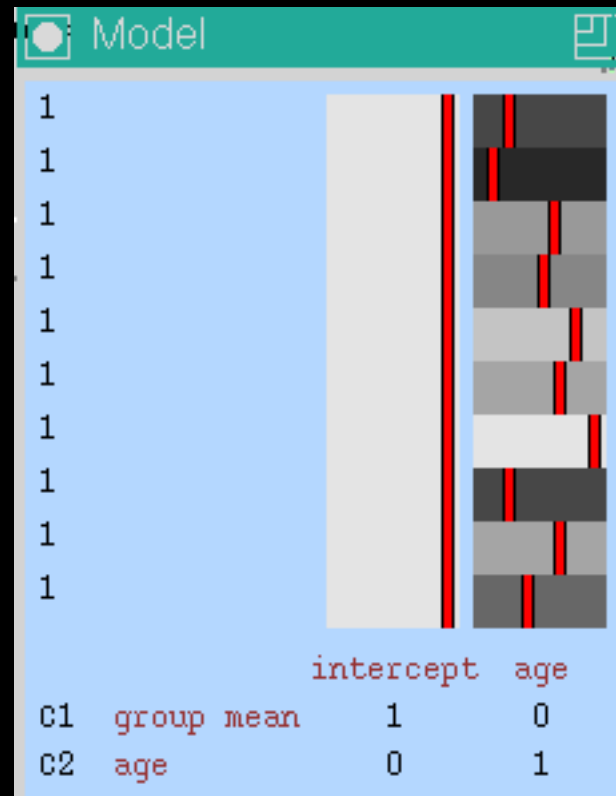
Single group with covariate

- You have a single group of subjects and you also have measured age. You would like to see if there is an age effect.
  - What would the model look like?
  - What contrasts would you specify for the age effect?
  - Can you still obtain the overall mean from this model?

Use age

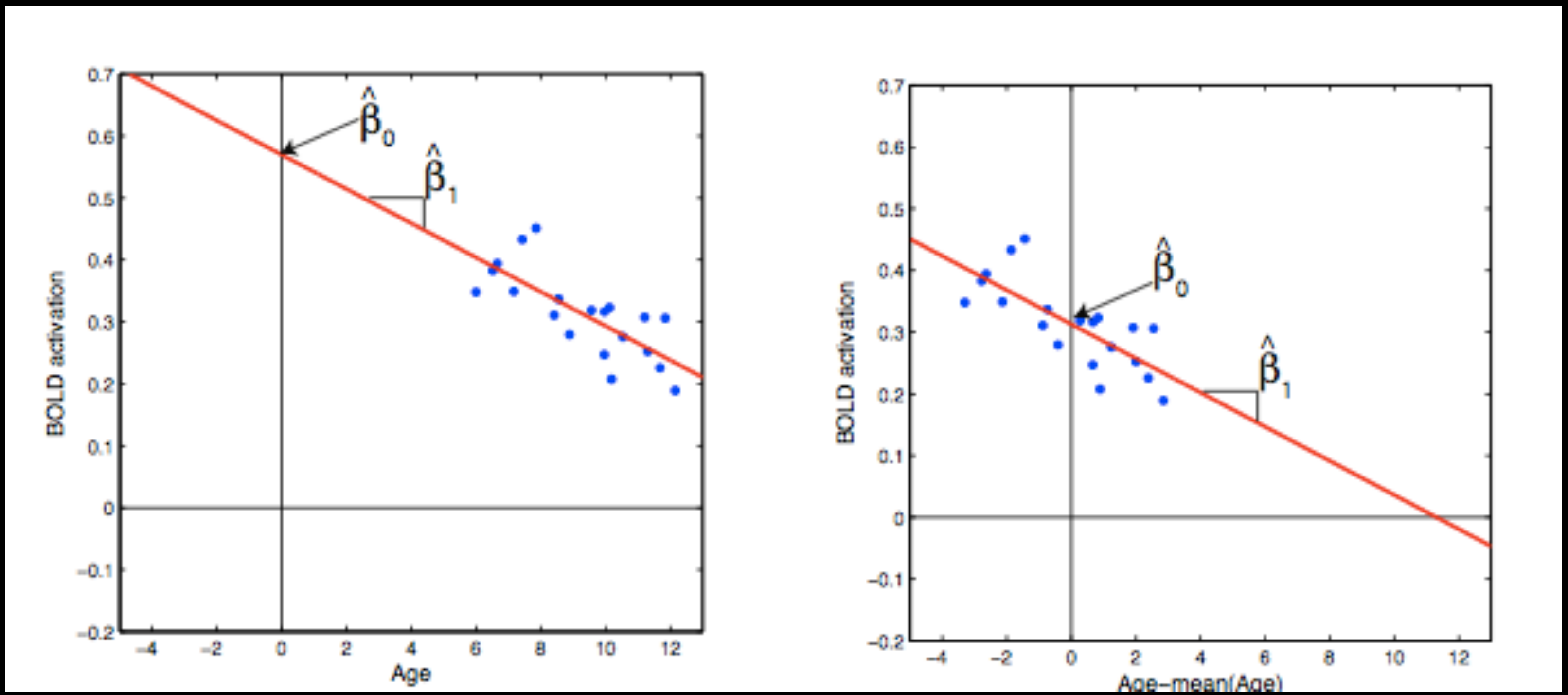


Use demeaned age



- Both models will give exactly the same result for C2, but C1 will be different.

Simulated data



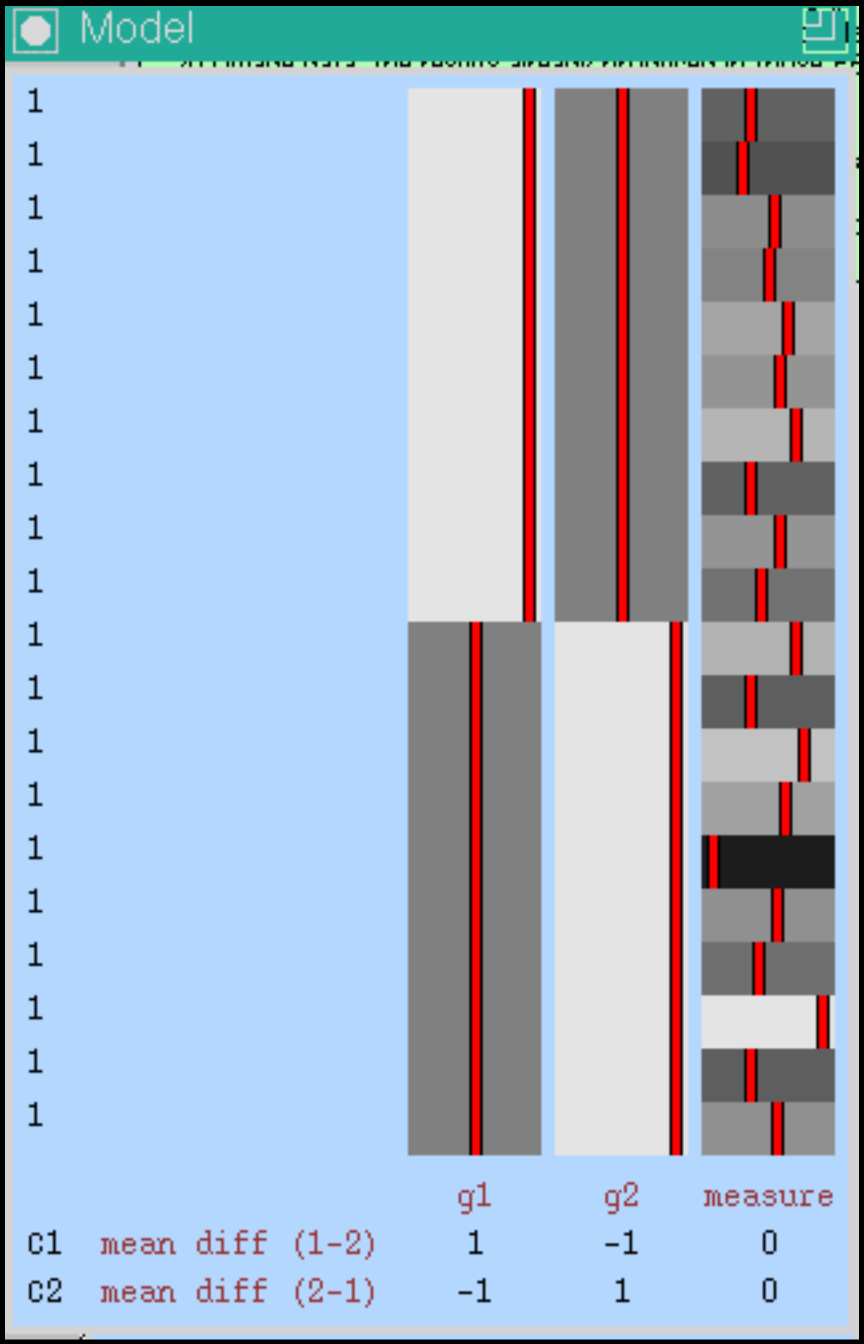
Summary of mean centering

- Only really necessary if you want your PE of column of 1s to be the overall mean
- Often people have rounding errors after demeaning. Double-check this when you do it.

Two groups with continuous covariate

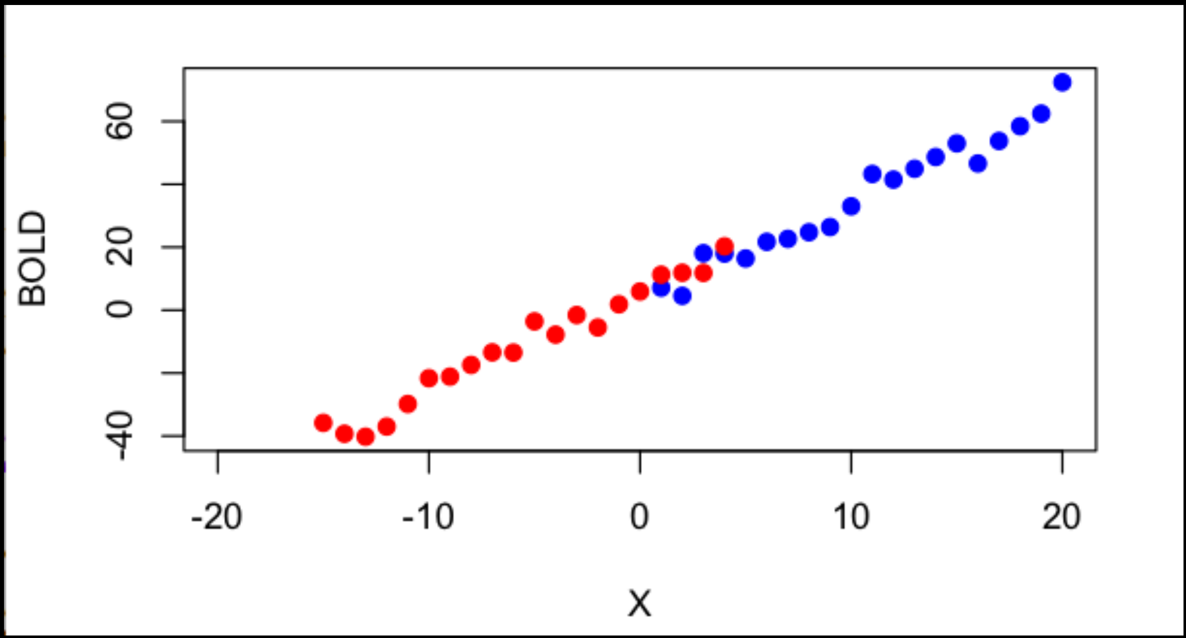
- We have two groups and a confounding covariate (depression). Our primary interest is in the difference of means between the two groups.
  - What is the model to simply test the difference in means
  - If I wanted to make sure this difference wasn't due to <u>between group</u> differences in depression, what would that model look like?
  - What are some restrictions of this model?

- Do not demean the confounding measure within group
  - This removes any confounding effect the measure might have



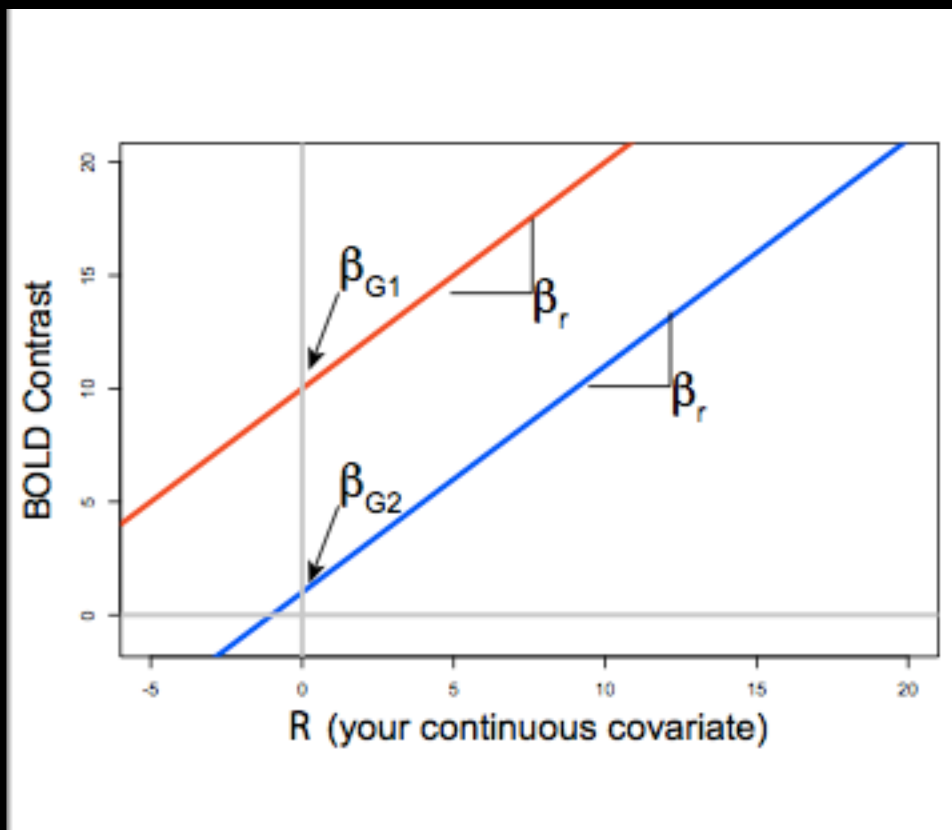
Why you shouldn’t demean within group

- What if this is what your data look like?
  - Difference in means is clearly due to range of X sampled, not the group membership



Will mean centering across all subjects change anything?

Will mean centering across all subjects change anything?

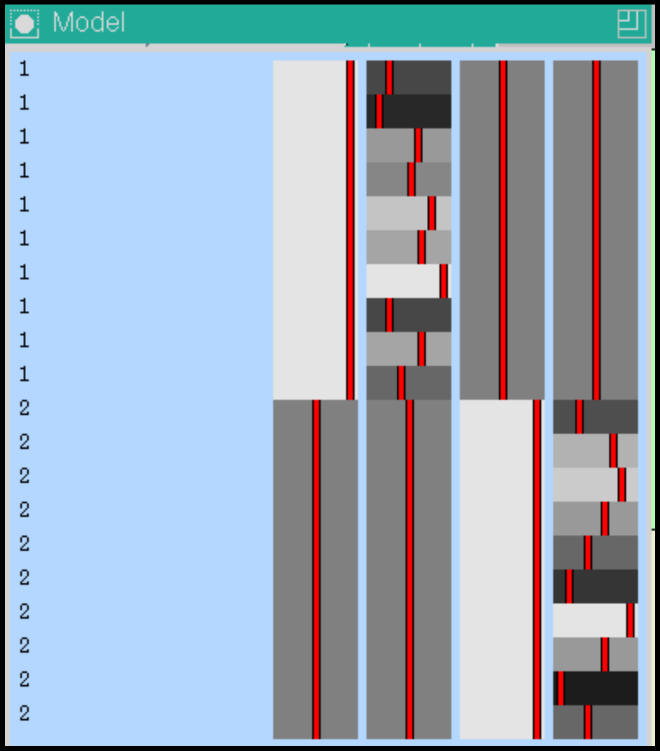


Nope!

Testing the interaction

- Same as previous example, but you'd like to know if the relationship with depression score differs between your groups
- What would the design matrix and contrasts look like?

Design matrix



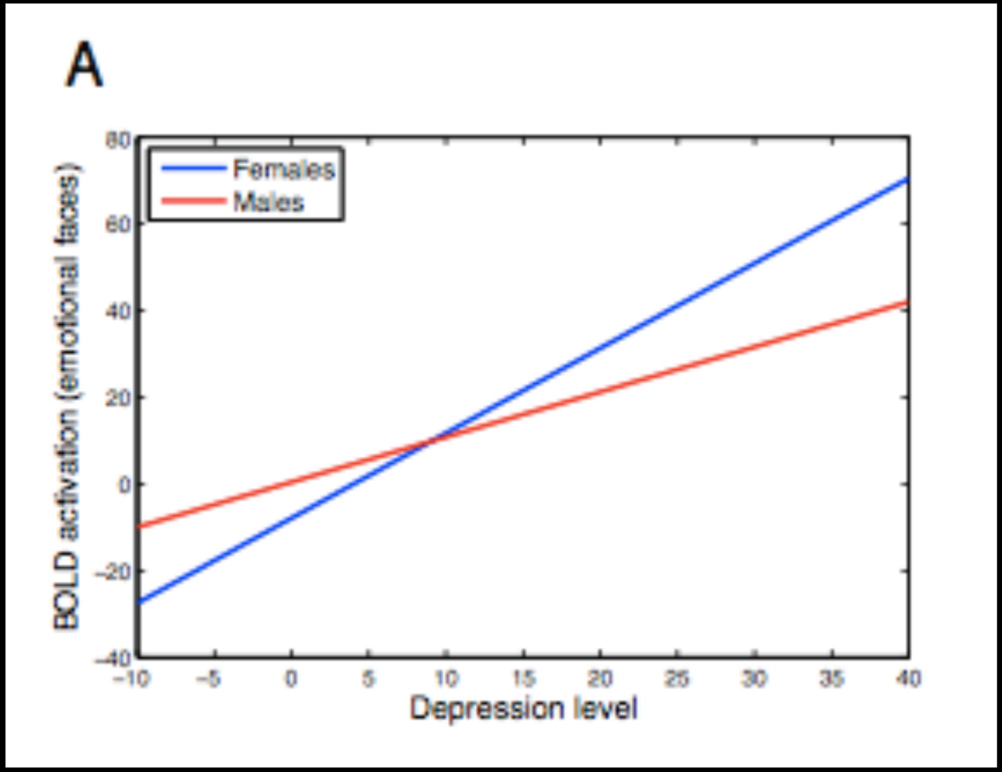
No mean centering!

What contrasts make sense?

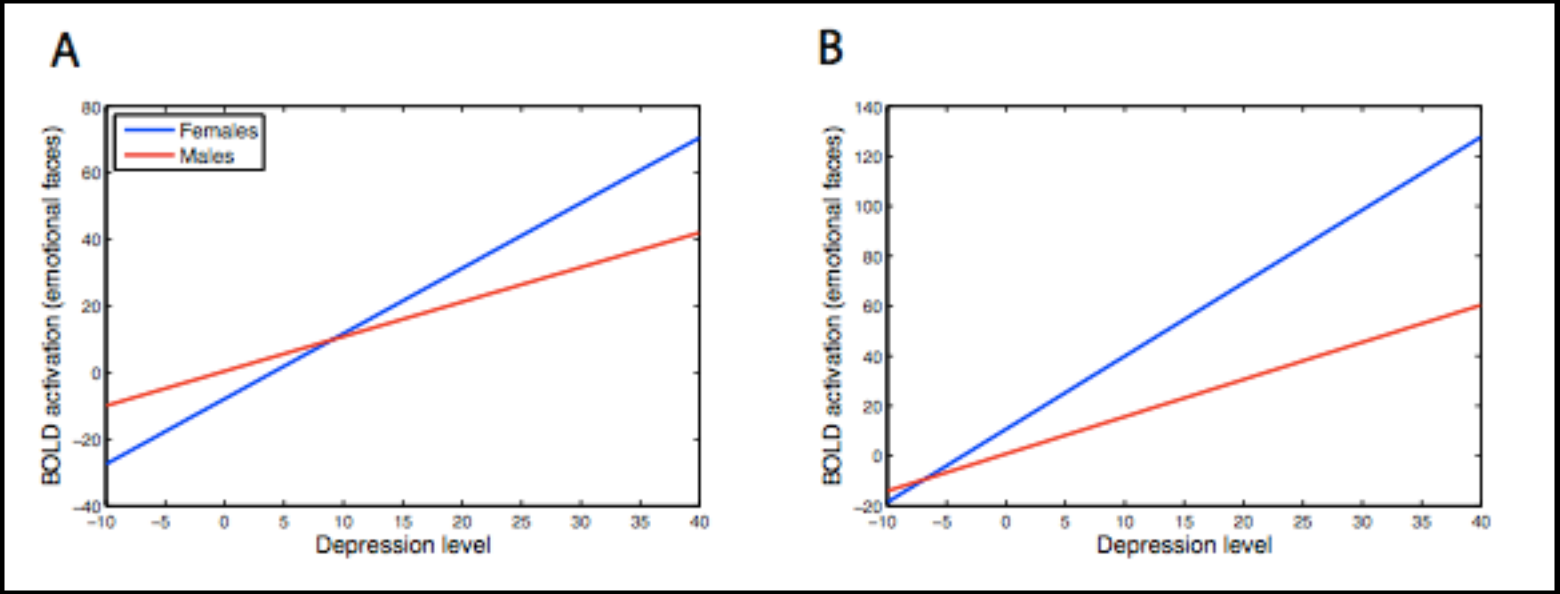
- Do comparisons of means make any sense?
- What is interesting about a significant interaction?

Compare means with a significant interaction?

- For what value of depression is the difference in means interesting?



What is interesting when there's an interaction effect?

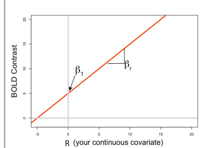
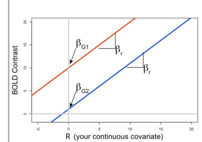
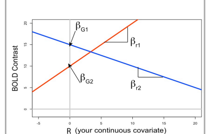


Mean centering when testing an interaction

- It doesn’t make much sense
  - Difference between means varies by continuous covariate when there’s an interaction effect
- Testing the mean difference in general doesn’t make sense
- Only interesting contrast tests difference in slope for continuous covariates between groups

Mean centering guide

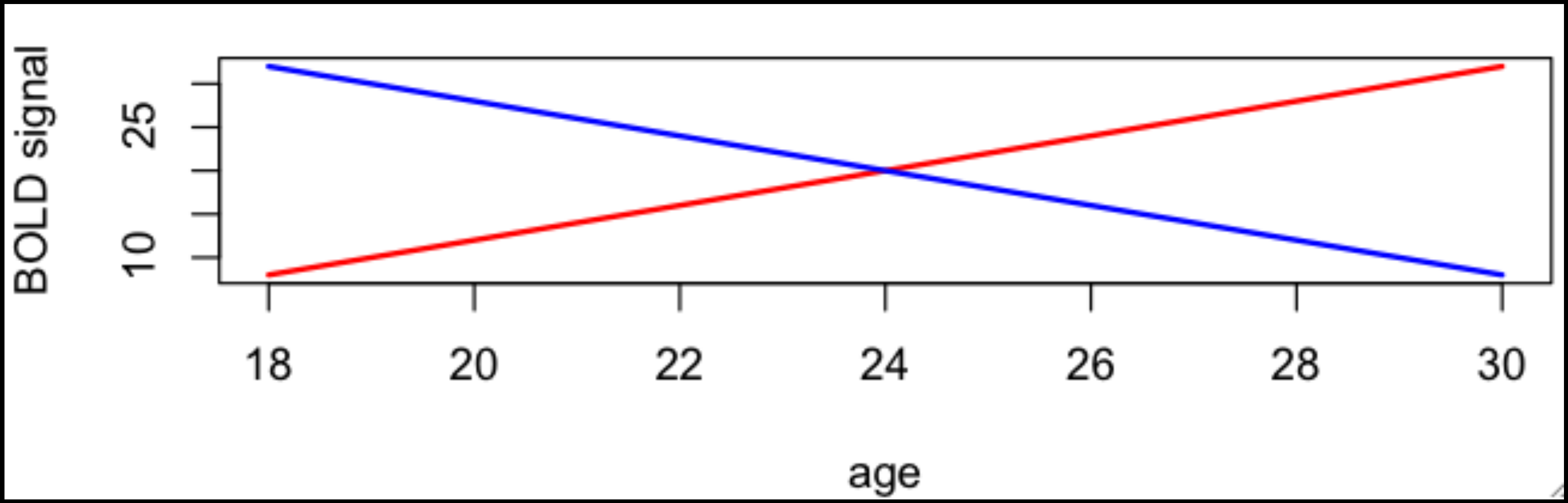
- http://mumford.fmripower.org/mean_centering/

Design matrix $X\beta$	What does the fitted model look like?	Contrast	Is mean centering necessary?
$\begin{pmatrix} 1 & r_1 \\ 1 & r_2 \\ 1 & r_3 \\ 1 & r_4 \\ 1 & r_5 \\ 1 & r_6 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_r \end{pmatrix}$		c=[1 0]	**Yes.** In this case you're interested in the mean of the individual group. Since this changes with the value of R (i.e. the height of the line varies with R) we typically choose to mean center R so that our inference applies to an average subject with the average value of R. Without mean centering you are looking at the mean for the group when R is 0, which is often not interesting.
		c=[0 1]	**No.** Mean centering a covariate will never change the contrast estimate or inference for that covariate (given the intercept is modeled).
$\begin{pmatrix} 1 & 0 & r_1 \\ 1 & 0 & r_2 \\ 1 & 0 & r_3 \\ 0 & 1 & r_4 \\ 0 & 1 & r_5 \\ 0 & 1 & r_6 \end{pmatrix} \begin{pmatrix} \beta_{G1} \\ \beta_{G2} \\ \beta_r \end{pmatrix}$		c=[1 -1 0]	**No.** The lines are parallel and you are testing if the distance between the two fitted lines is significantly different from 0. This distance is the same for all values of R
		c=[0 0 1]	**No.** Mean centering a covariate will never change the inference for that covariate (given the intercept is modeled).
		c=[1 0 0] or c=[0 1 0]	**Yes.** In this case you're interested in the mean of an individual group. Since this changes with the value of R we typically choose to mean center R so that our inferences apply to an average subject with the average value of R. Without mean centering you are looking at the mean for the group when R is 0, which is often not interesting.
$\begin{pmatrix} 1 & 0 & r_1 & 0 \\ 1 & 0 & r_2 & 0 \\ 1 & 0 & r_3 & 0 \\ 0 & 1 & 0 & r_4 \\ 0 & 1 & 0 & r_5 \\ 0 & 1 & 0 & r_6 \end{pmatrix} \begin{pmatrix} \beta_{G1} \\ \beta_{G2} \\ \beta_{r1} \\ \beta_{r2} \end{pmatrix}$		c=[1 -1 0 0]	**Yes.** Sort of. This is a difficult contrast to interpret if the interaction contrast [0 0 1 -1] is significant. This is because a significant interaction means that the distance between the means of your two groups varies with the value of R (as you can see in the plot). So, inference on this contrast when the intereaction is significant almost isn't that helpful. What would be more helpful is knowing where the lines cross and when one is above another. See an introductory linear regression text for information on interpreting interactions. The fact that you are probably running a whole brain analysis compounds the problem. Do not mean center within group (see note in introductory text). NOTE: If the interaction isn't significant, it could be easier to run the model above this unless you're restricted from doing so because you're using the "group" column in FSL.
		c=[1 0 0 0] or c=[0 1 0 0]	**Yes.** In this case you're interested in the mean of an individual group. Since this changes with the value of R we typically choose to mean center R so that our inferences apply to an average subject with the average value of R. Without mean centering you are looking at the mean for the group when R is 0, which is often not interesting. Again, do not mean center within group.
		c=[0 0 1 -1] or c=[0 0 1 0] or c=[0 0 0 1]	**No.** Mean centering a covariate will never change the estimates for that covariate (given the intercept is modeled). The first contrast in this list is typically the most important contrast as it tests the interaction.

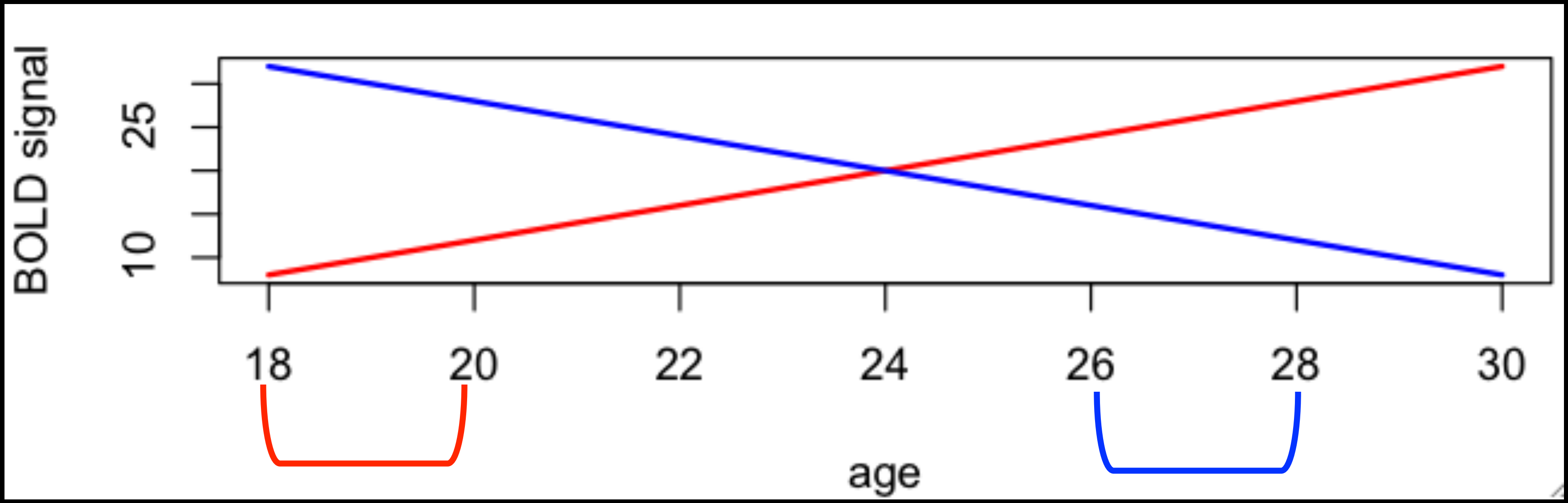
Important assumption

- That the effect tested is linear!
- Relationship between BOLD activation and depression score is linear for both groups
- If your continuous covariate is significantly different between your groups, you must be very careful with interpretation

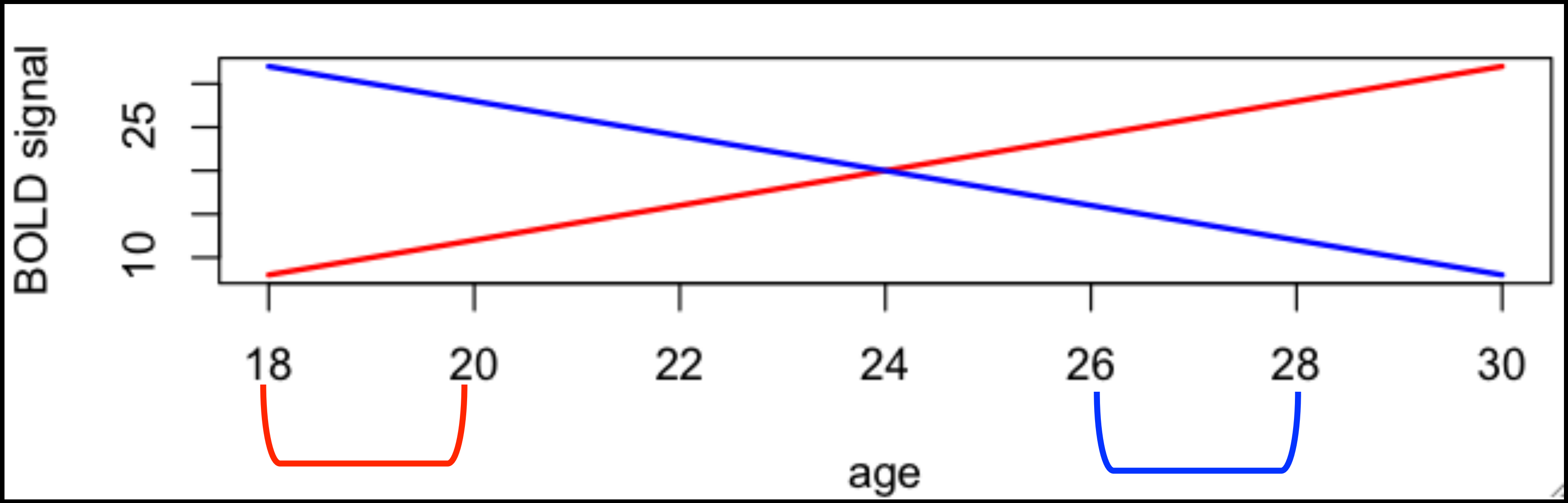
When trends are linear, it doesn't matter where I sample age, I get the same slope

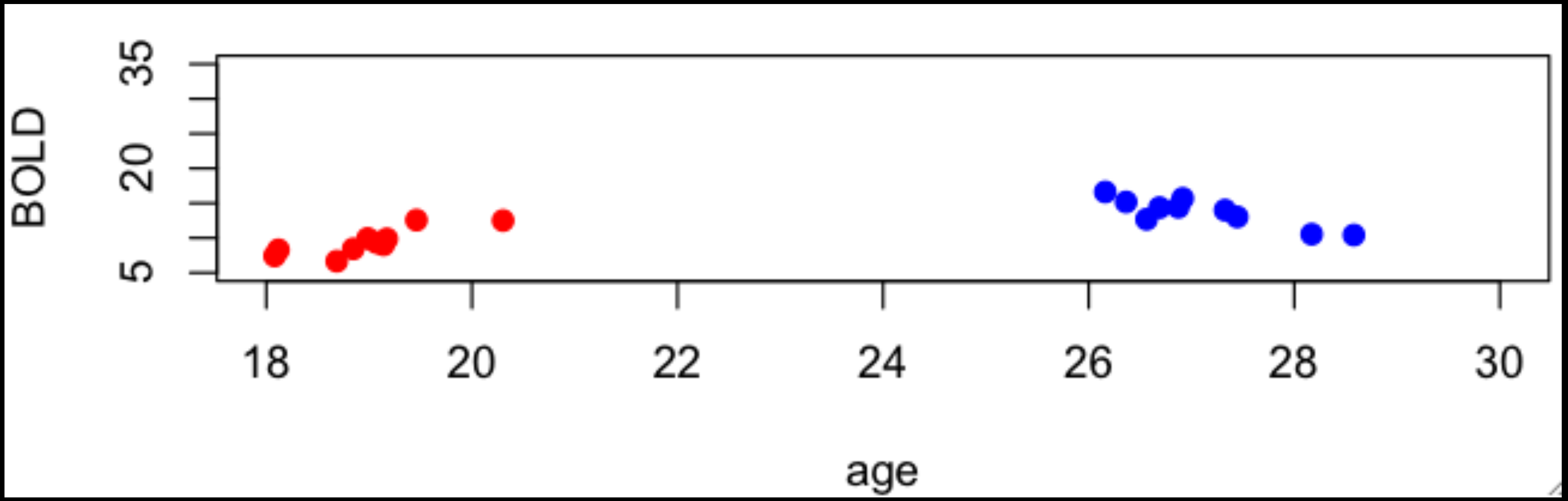


When trends are linear, it doesn’t matter where I sample age, I get the same slope

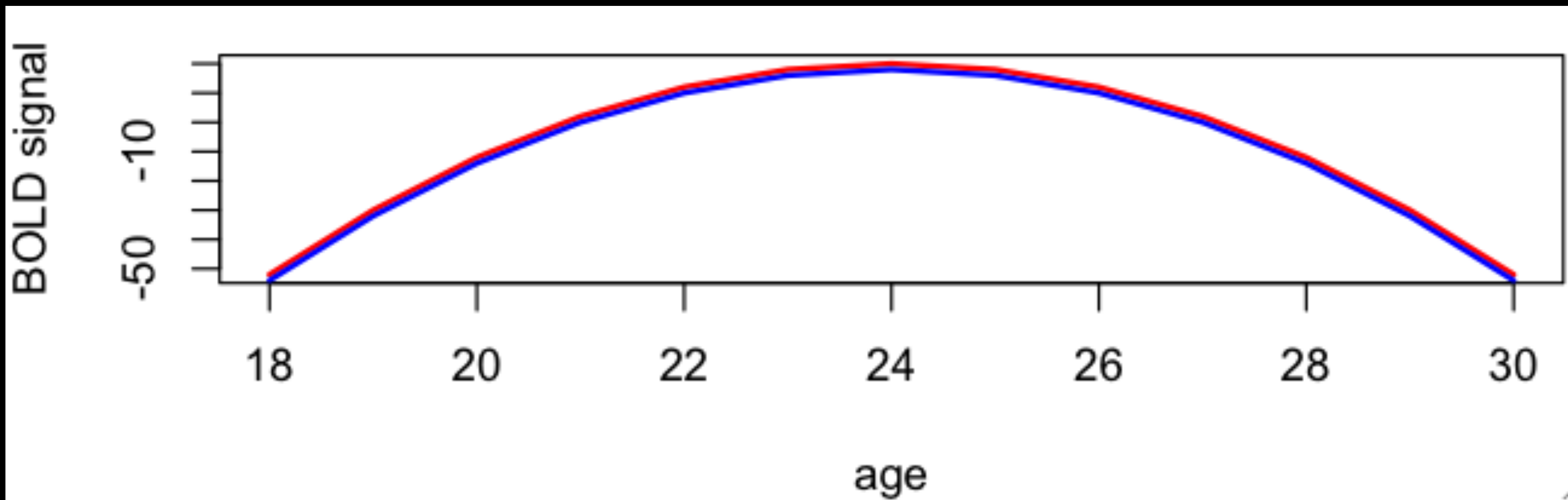


When trends are linear, it doesn't matter where I sample age, I get the same slope

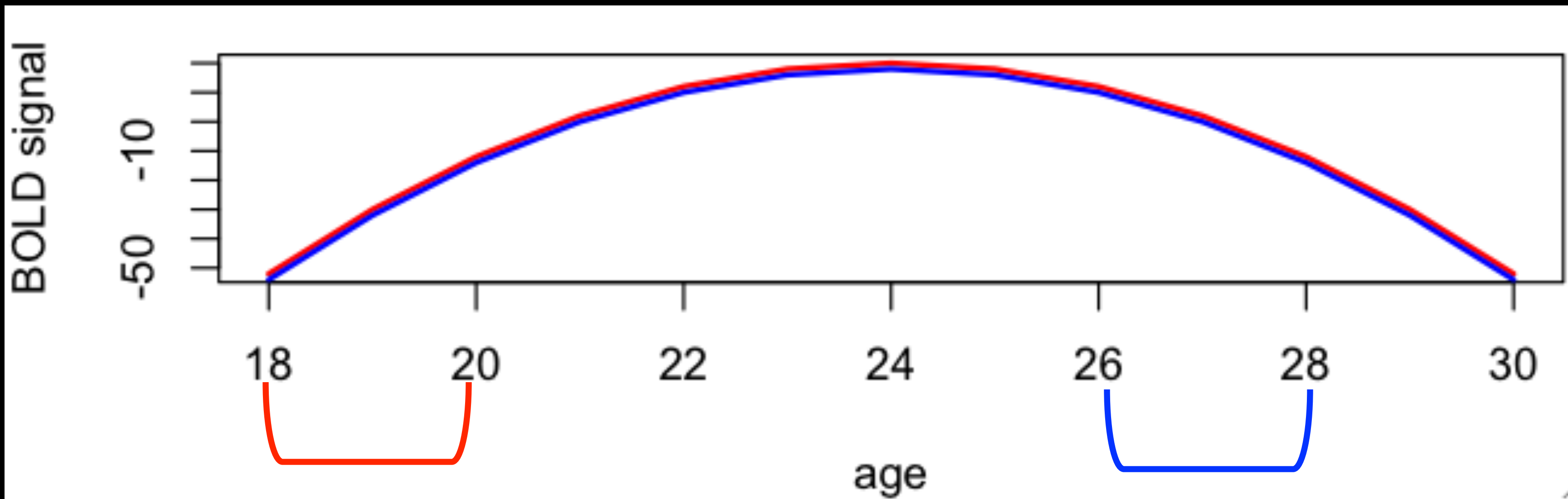




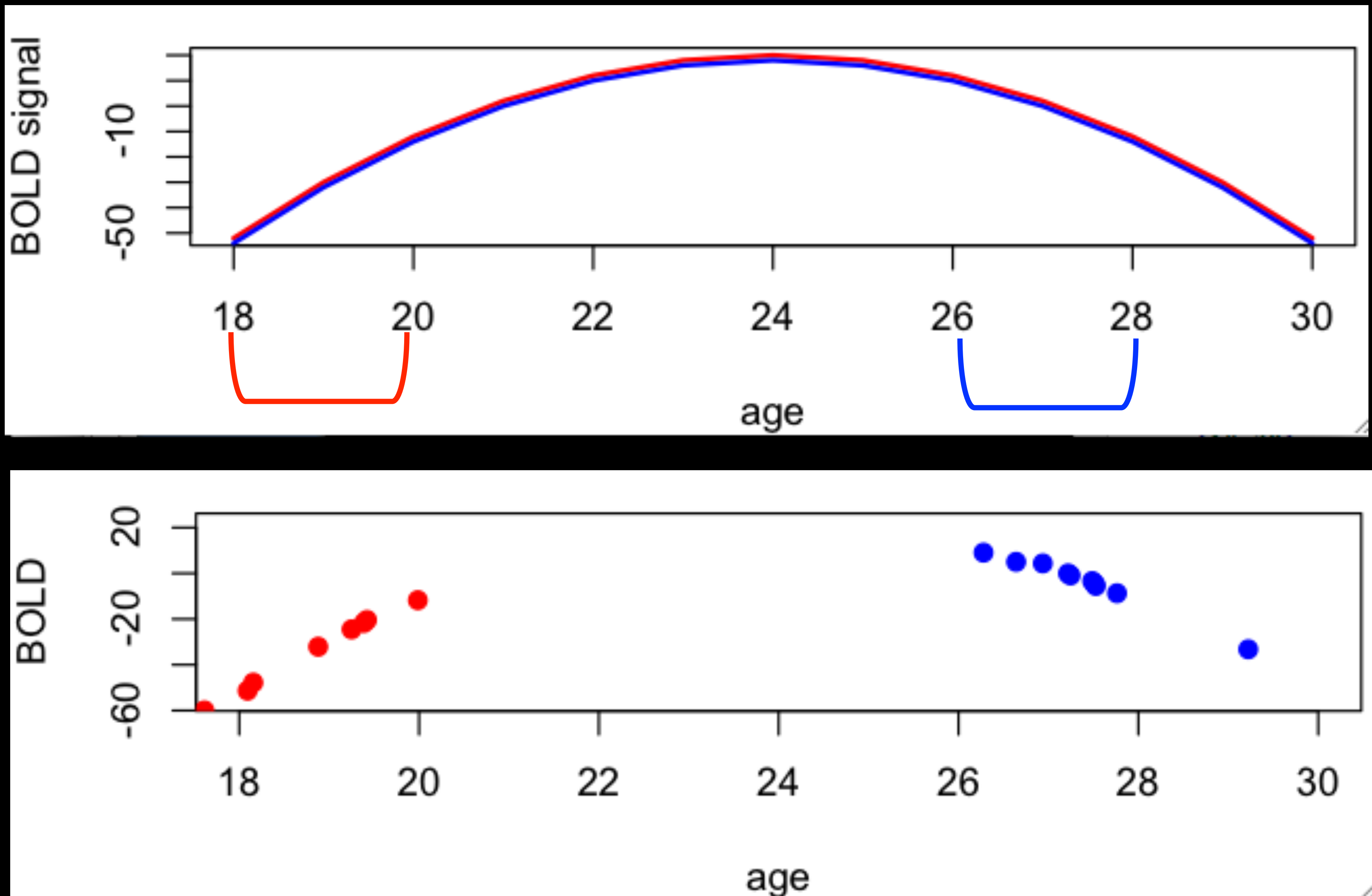
With a nonlinear trend, the trend may be the same in both groups, but sampling different ranges show different trends



With a nonlinear trend, the trend may be the same in both groups, but sampling different ranges show different trends



With a nonlinear trend, the trend may be the same in both groups, but sampling different ranges show different trends



Rank deficient designs

- I highly recommend becoming an expert on spotting rank deficient design matrices
  - Don’t email the FSL list to ask why you got the error message ”at least one EV is (close to) a linear combination of the others”

Rank deficient designs

- I highly recommend becoming an expert on spotting rank deficient design matrices
  - Don’t email the FSL list to ask why you got the error message ”at least one EV is (close to) a linear combination of the others”
    - (the answer to the problem is in the error message!!)

Rank deficient

- One column is a linear combination of some of the other columns
- Linear combination of one set of columns equals a linear combination of another set of columns

Rank deficient

$$\begin{pmatrix} 1 & 0 & 7 \\ 1 & 0 & 7 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 2 \\ 1 & 1 & 5 \\ 1 & 1 & 5 \end{pmatrix}$$

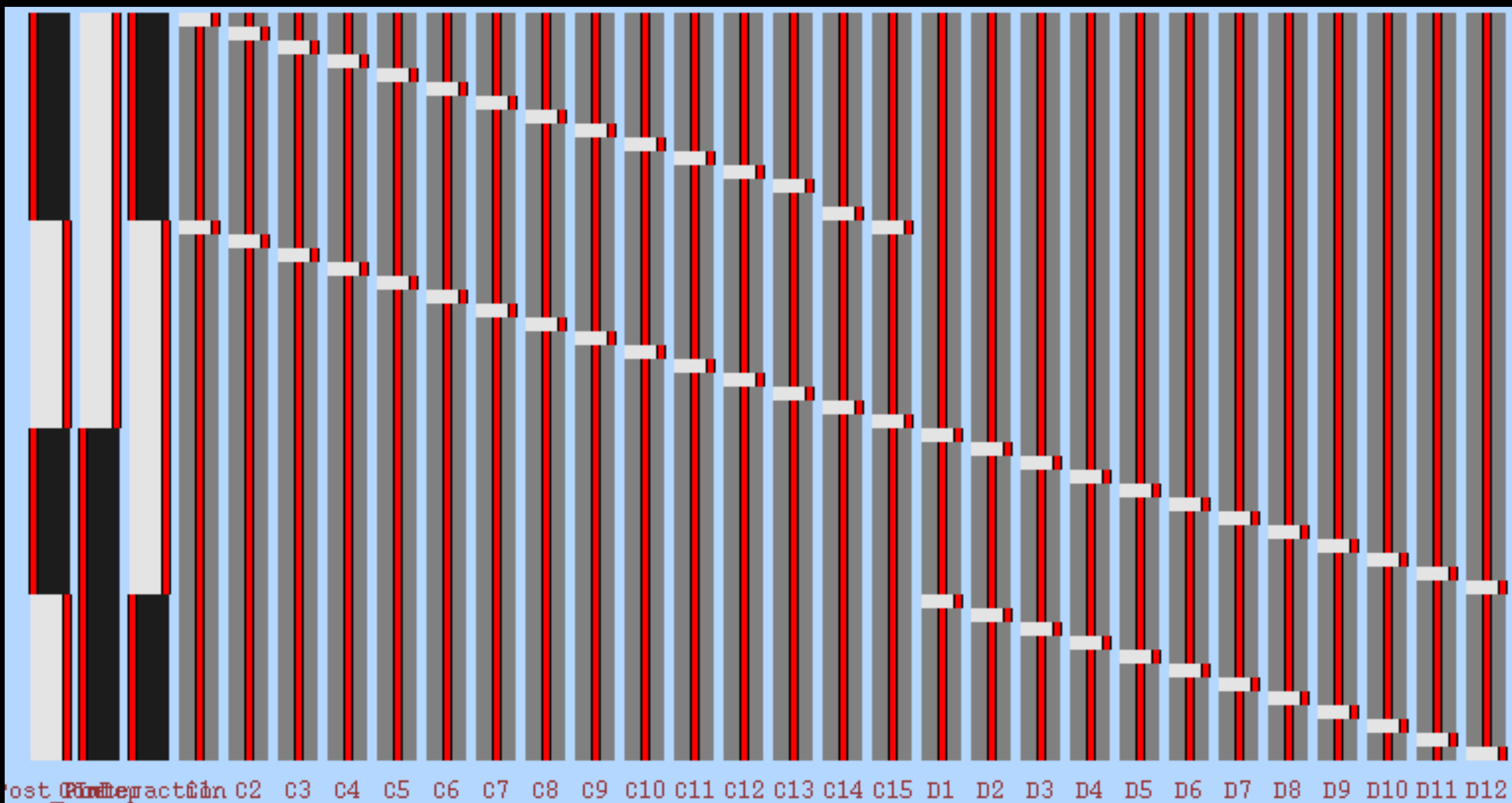
Rank deficient

1	0	1	0
1	0	0	1
1	0	1	0
1	0	0	1
0	1	1	0
0	1	0	1
0	1	1	0
0	1	0	1

Rank deficient?

>> Group	EV1	EV2	EV3	EV4	EV5	EV6	EV7
>>	subj1	subj2	time1	time 2	time 3	time 4	time5
>> 1	1	0	1	0	0	0	0
>> 1	0	1	1	0	0	0	0
>> 1	1	0	0	1	0	0	0
>> 1	0	1	0	1	0	0	0
>> 1	1	0	0	0	1	0	0
>> 1	0	1	0	0	1	0	0
>> 1	1	0	0	0	0	1	0
>> 1	0	1	0	0	0	1	0
>> 1	1	0	0	0	0	0	1
>> 1	0	1	0	0	0	0	1

Rank deficient?



That’s it!

- Next lecture we’ll look at more models. If you have a design that has you stumped, let me know what your data look like and we can all try to figure it out!