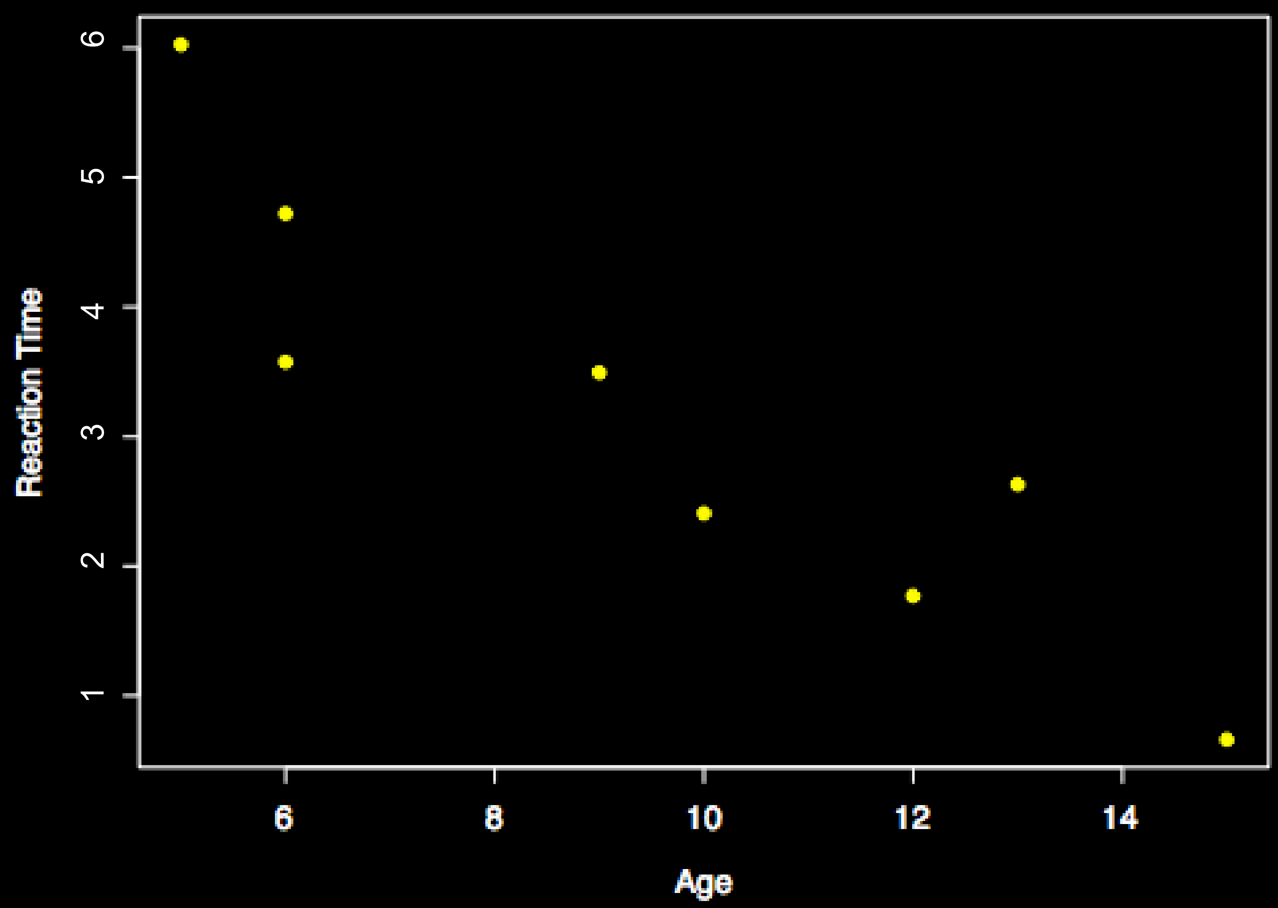
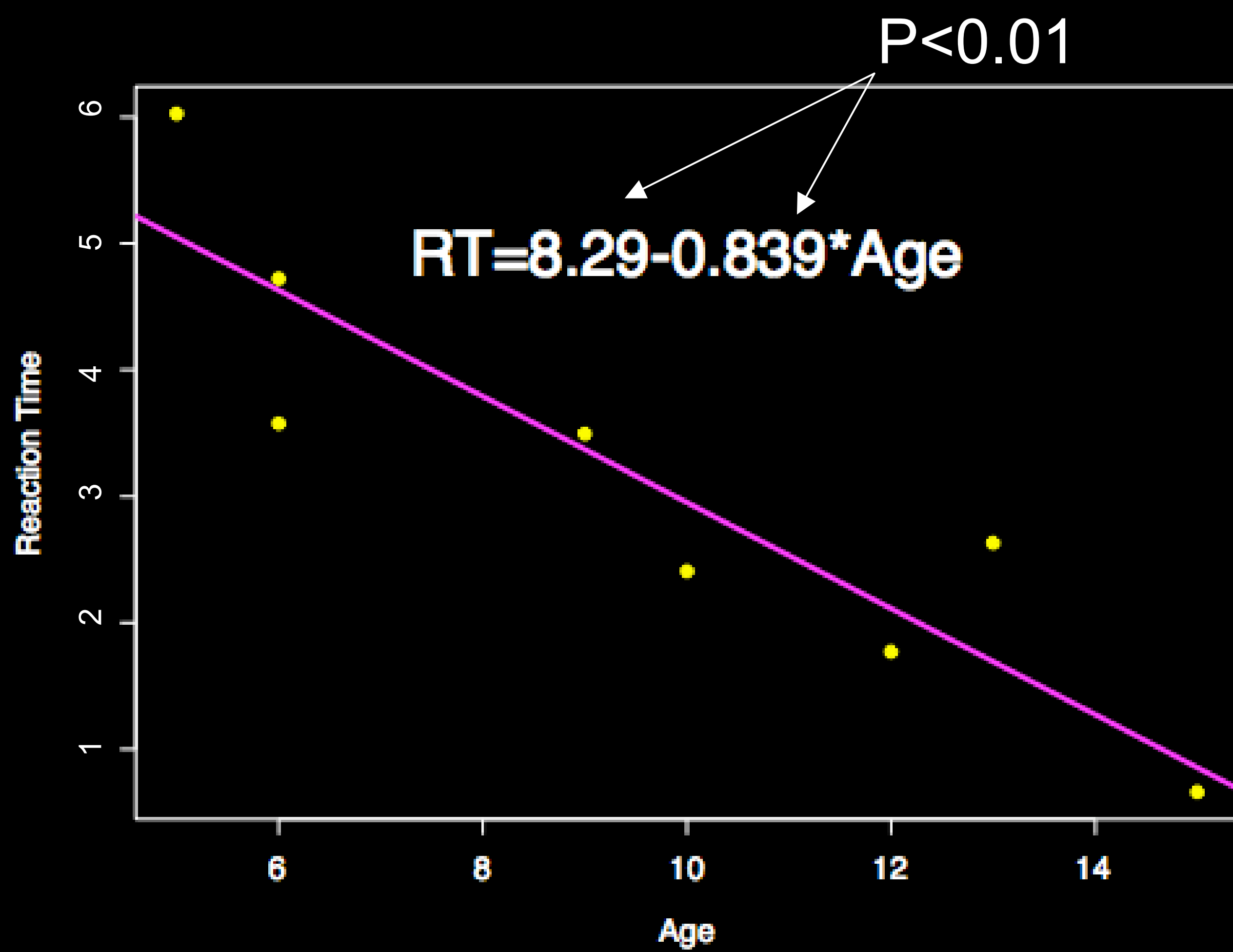
Statistical Modeling and Inference

UCLA Advanced NeuroImaging
Summer School, 2007

Models help tell stories



Models help tell stories



Goal of next 2 hours

- Hour 1
  - Brush up on some stats lingo
  - Review the general linear model (GLM)
  - How to estimate the GLM
- Hour 2
  - Hypothesis testing
  - Building Models

Statistical Terms

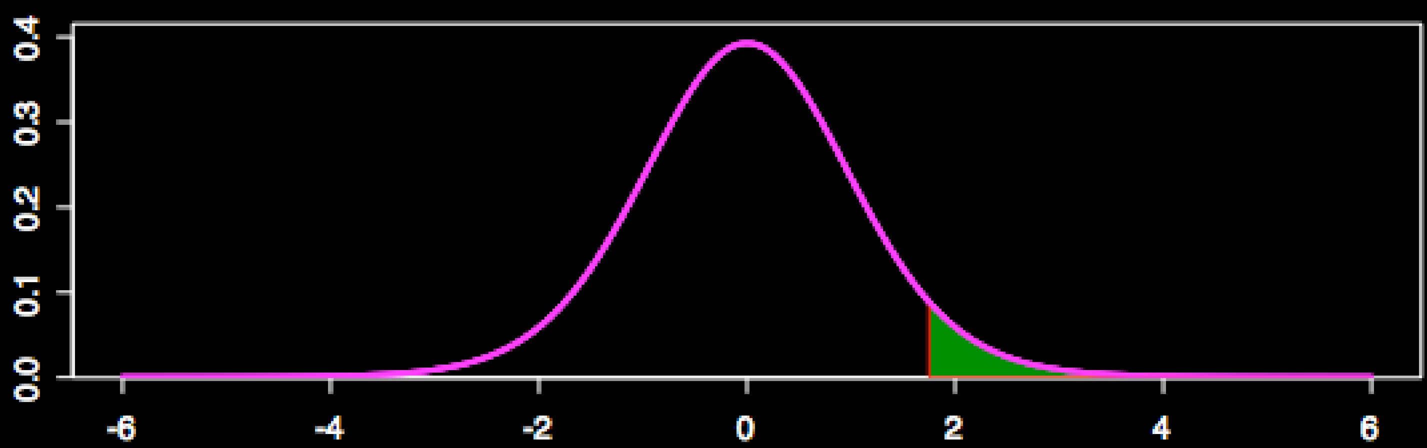
- Probability: The expected relative frequency of a particular outcome.
  – If you flip a "fair" coin a lot of times 50% of the time you'll get "heads"
    - P(heads)=0.5
  – You measure the heights of people and 30% of the time they are taller than 69 inches
    - P(height>69)=0.3

Statistical Terms

- Random variable
  - Values of the variable are different every time it is observed (eg, height)
  - Use capital letters for random variables
  - Use lower case for observed values
  - $P(H>h)$=Probability that height is larger than some observed height h

Statistical Terms

- Probability distribution
  - Describes the distribution of a random variable
  - Defined by density function, f(h)
  - Area under density gives probability

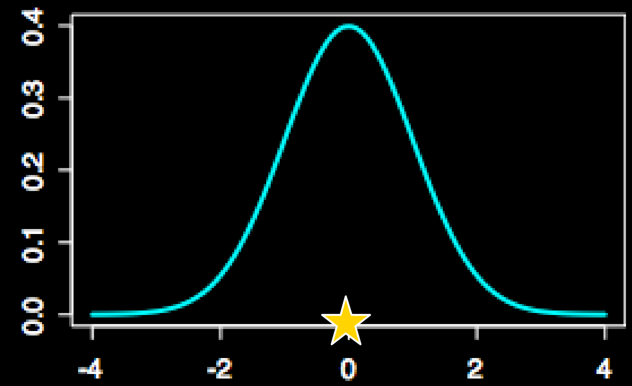


Statistical Terms

- How do we know what the correct distribution of a random variable is?
  - By observation: Most data we deal with is distributed normally
  - Many distributions are related to the normal
    - Chisqare is the square of a normal
    - T uses a normal and a chisquare
    - F is the ratio of two chisquares

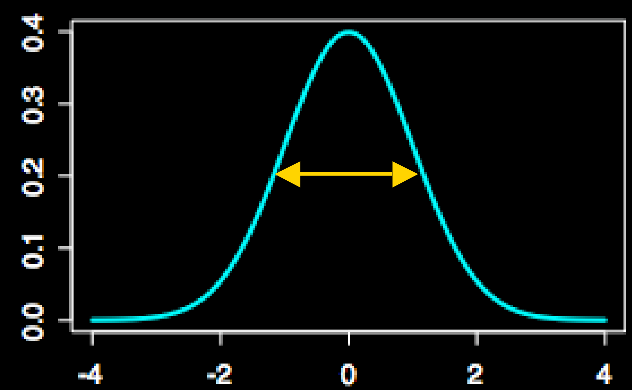
Statistical Terms

- Statistical Independence
  - X and Y are independent if the occurrence of X tells us nothing about Y
- Expected Value
  - The mean of a random variable ($E[Y]$)
  - $E[X+a] = E[X] + a$
  - $E[aX] = aE[X]$
  - $E[X+Y] = E[X] + E[Y]$
  - $E[XY] = E[X]E[Y]$
    - Only if X and Y are independent!!



Statistical Terms

- Variance
  - How the values of the RV are dispersed about the mean
  - $\mathrm{Var}[Y] = E\left[(Y - E[Y])^2\right]$
  - $\mathrm{Var}[aY] = a^2\mathrm{Var}[Y]$
  - $\mathrm{Var}[Y + a] = \mathrm{Var}[Y]$



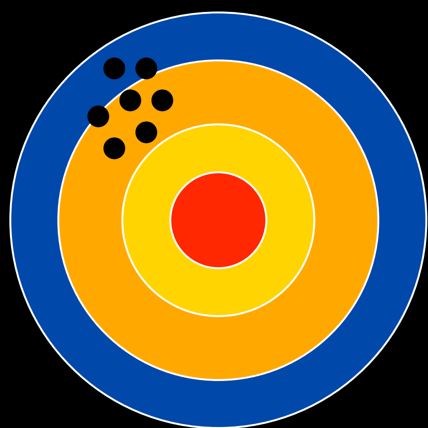
- Covariance
  - How much 2 RV's vary together
  - $\mathrm{Cov}[X, Y] = E\left[(X - E[X])(Y - E[Y])\right]$
  - If 2 RV's are independent, $\mathrm{Cov}[X, Y] = 0$ BUT the opposite is not true

Statistical Terms

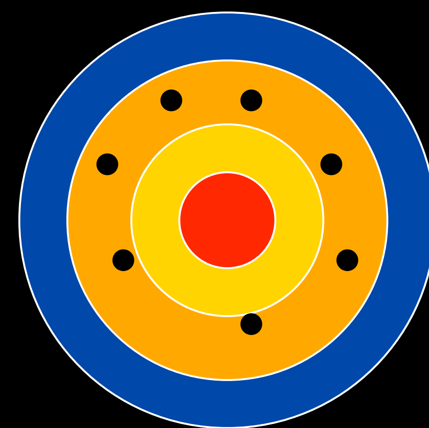
- Bias and Variance can be used to assess an estimator
  - Bias: On average, the estimate is correct
  - Variance: The reliability of the estimate
  - Efficient: The most efficient estimate has the lowest variance among all unbiased estimators

Bias and Variance

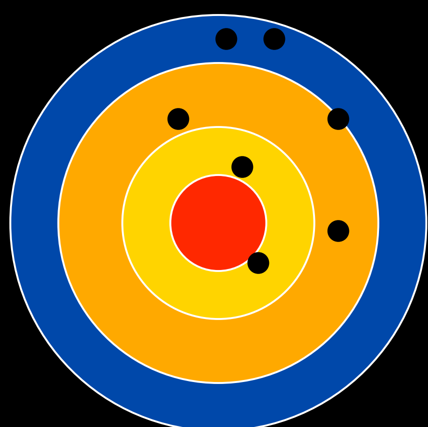
high bias / low variance



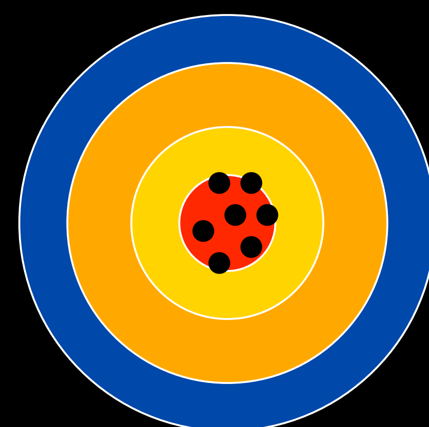
low bias / high variance



high bias / high variance



low bias / low variance



The Model

- For the ith observational unit

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

- Y_i : The dependent (random) variable
- X_i : Predictor variable (not random)
- β_0, β_1 : Model parameters
- ϵ_i : Random error, how the observation deviates from the population mean

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

- Simple linear regression
- Simple: Because there is only 1 regressor and an intercept
- Linear: Because it is linear in its parameters

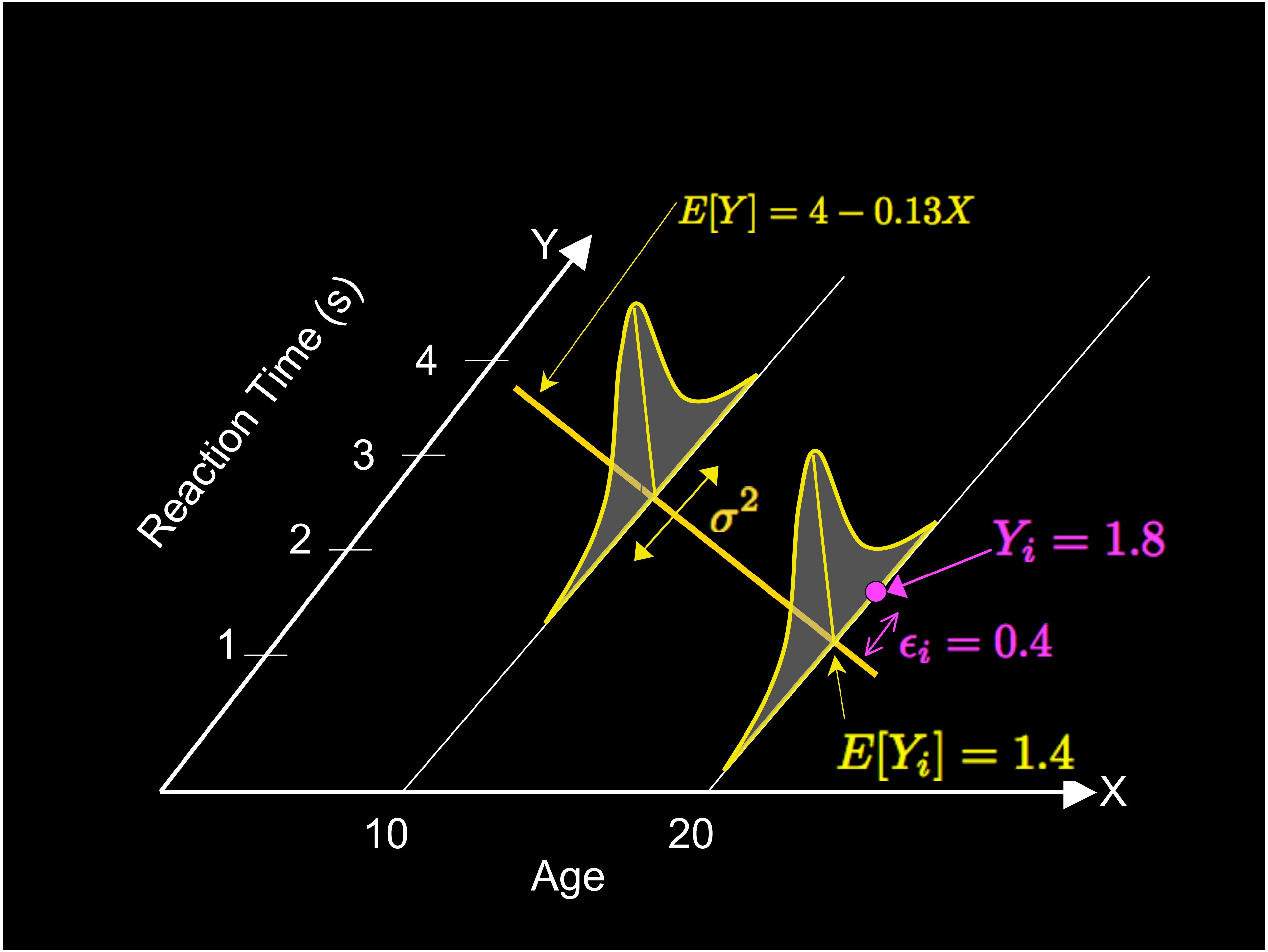
$$Y_i = \beta_0 + \beta_1^2 X_i + \epsilon_i$$

$$Y_i = \beta_0 + \frac{\beta_1}{\beta_0} X_i + \epsilon_i$$

Not linear!

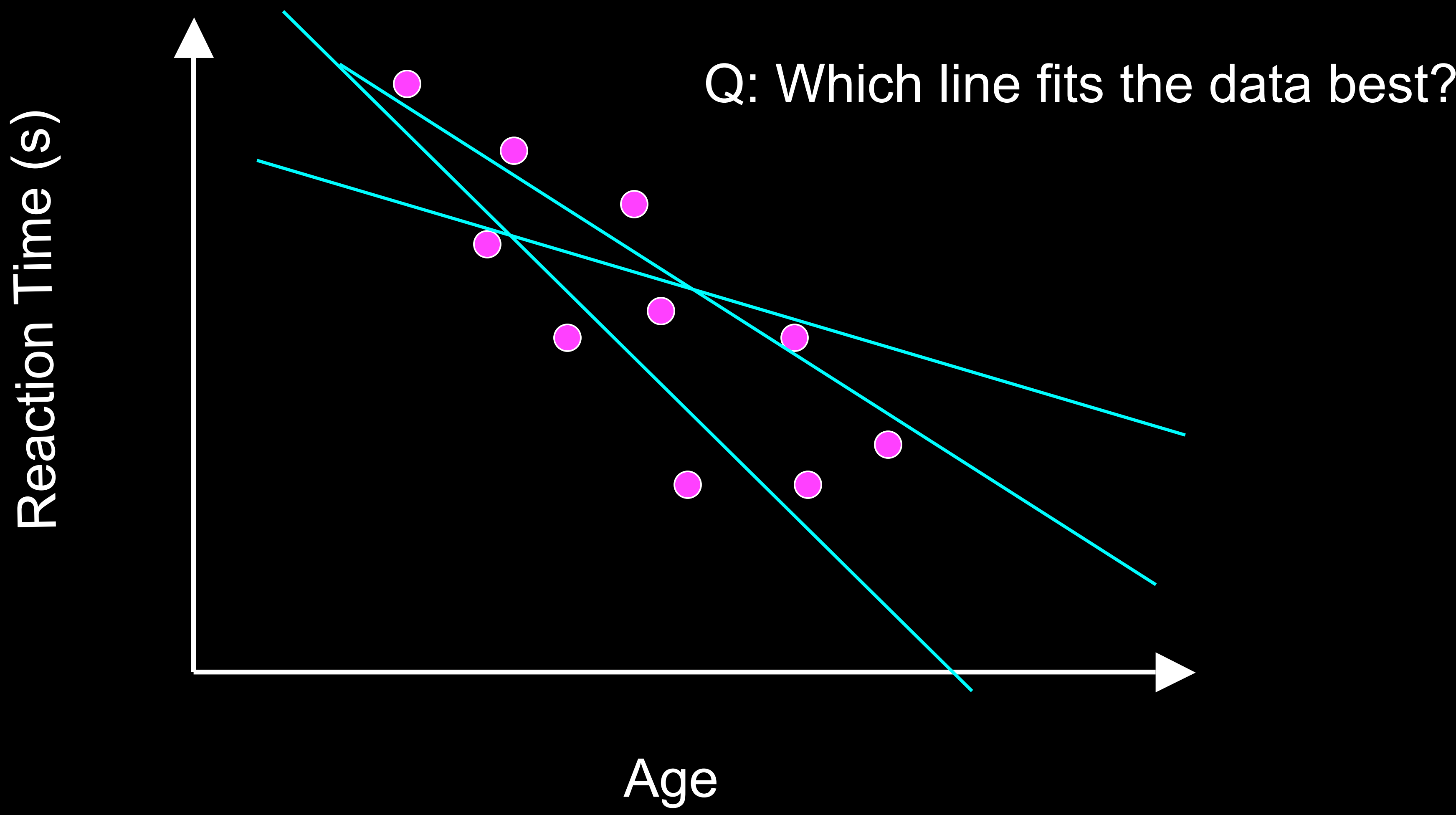
Fixed and Random Parts

- Fixed parts
  - $\beta_0 + \beta_1 X_i$ describes the mean of Y_i, ($E[Y_i]$)
- Random part
  - ϵ_i describes the variability of Y_i
    - ϵ_i has a mean of 0
    - ϵ_i has a constant variance σ^2
    - ϵ_i are uncorrelated
    - It follows that the variance of Y_i is σ^2



$E[Y] = 4 - 0.13X$
Y
Reaction Time (s)
4
3
2
1
σ^2
$Y_i = 1.8$
$\epsilon_i = 0.4$
$E[Y_i] = 1.4$
X
10
20
Age

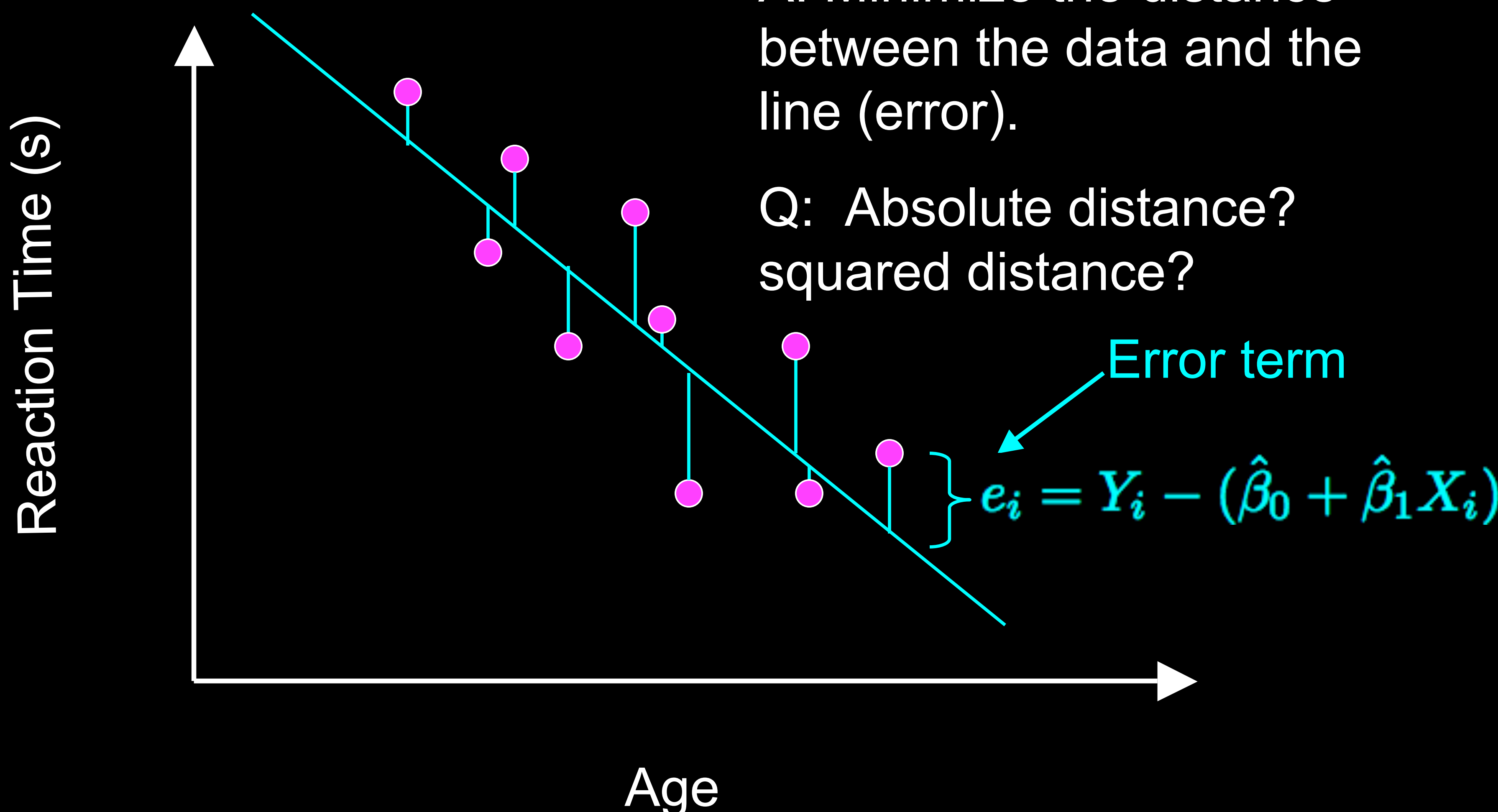
Fitting the Model



Fitting the Model

A: Minimize the distance between the data and the line (error).

Q: Absolute distance? squared distance?



Least Squares

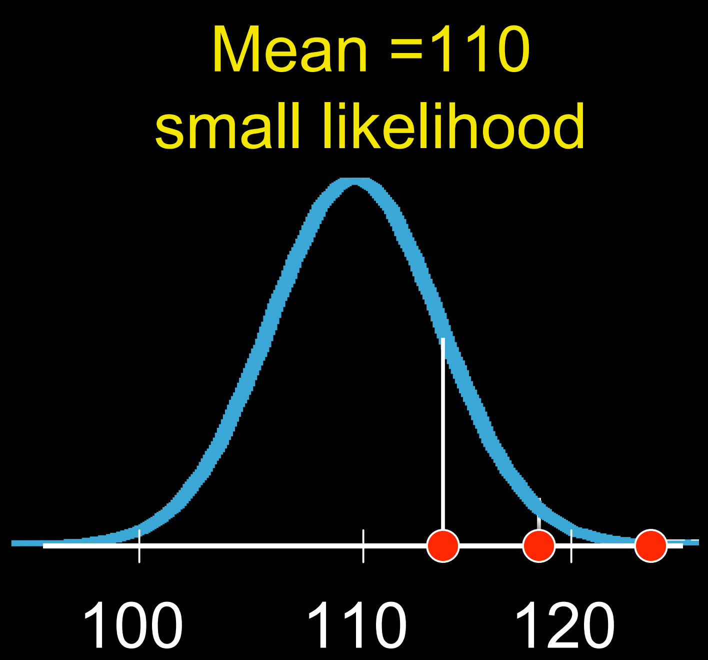
- Minimize squared differences
- Minimize $\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2$
  - Works out nicely distribution-wise
  - You can use calculus to get the estimates
    - $\hat{\beta}_1 = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2}$
    - $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$

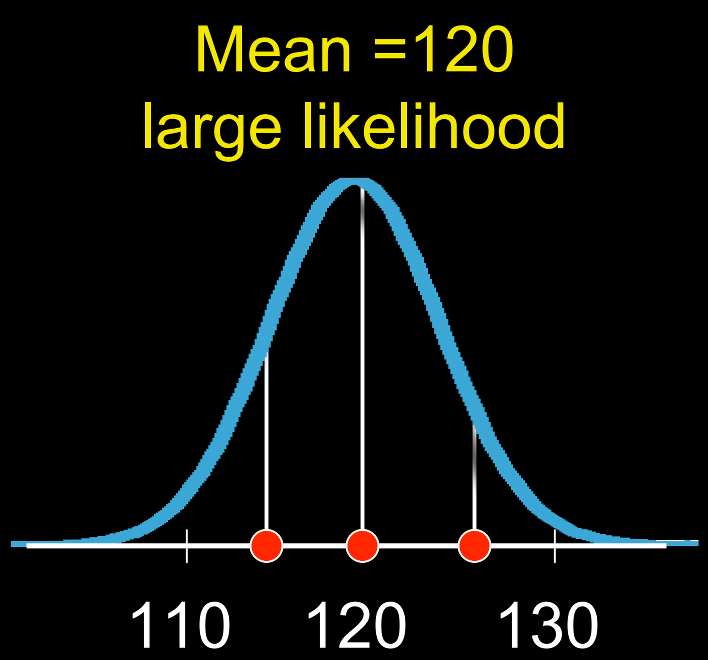
Property of Least Squares

- Gauss-Markov theorem
  - Under the assumptions we’ve made so far (error has mean 0, with constant variance and uncorrelated) the least squares estimators are unbiased and have minimum variance among all unbiased linear estimates
- i.e. The logical way to estimate the model gives really great estimates!

What's maximum likelihood?

- Instead of minimizing error, maximize the likelihood that we would have gotten our data for some given parameters $P(Y|\beta)$





Maximum Likelihood

- When assuming normality, leads to the same results as least squares
- If it is the same, why bring it up?
  - Studying $P(Y|\beta)$ is a Frequentist approach
  - Next week you'll hear about Bayesian methods, which focus on $P(\beta|Y)$

What about the variance?

- We also need an estimate for σ^2
  - Start with the mean square error

$$MSE = \sum(Y_i - \hat{Y}_i)^2 = \sum e_i^2$$

  - Divide by the appropriate degrees of freedom
    - # of independent pieces of information - # parameters in model

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{N - 2}$$

Multiple Linear Regression

- Add more parameters to the model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \epsilon_i$$

- Least square works, but is messy
- Time for linear algebra!

Matrices

- A is a 2x3 matrix

$$A = [a_{ij}] \qquad i = 1, 2; \quad j = 1, 2, 3$$

Row index

Column index

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Matrices

- Square matrix- Same # of rows and columns

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

- Vector- column(row) vector has 1 column(row)

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} \qquad \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix}$$

Matrices

- Special matrices
  - Diagonal Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

  - Identity - I_N

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Matrices

- Transpose: A^T or A'. Swap columns and rows.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \end{pmatrix} \Rightarrow A' = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \end{pmatrix}$$

- Element-wise addition and subtraction
  - Dimensions must match

$$\begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \end{pmatrix} + \begin{pmatrix} 4 & 7 & 3 \\ 1 & 8 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 9 & 6 \\ 6 & 14 & 8 \end{pmatrix}$$

Matrices

- Multiplication: Trickier
  - Number of columns of first matrix must match number of rows of second matrix

$$\xcancel{\begin{pmatrix} 4 & 7 & 3 \\ 1 & 8 & 1 \end{pmatrix} \times \begin{pmatrix} 2 & 3 & 6 \\ 1 & 2 & 1 \end{pmatrix}} \quad \begin{pmatrix} 1 & 2 \\ 4 & 9 \\ 3 & 2 \end{pmatrix} \times \begin{pmatrix} 4 & 2 \\ 1 & 4 \end{pmatrix}$$

Matrices

- Multiplication

$$AB = C \rightarrow c_{ij} = \sum_{n=1}^{cols_A} a_{im} b_{mj}$$

$$\begin{pmatrix} 1 & 2 \\ 4 & 9 \\ 3 & 2 \end{pmatrix} \times \begin{pmatrix} 4 & 2 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} \boxed{} & \\ & \\ & \end{pmatrix}$$

Matrices

- Multiplication

$$AB = C \rightarrow c_{ij} = \sum_{n=1}^{cols_A} a_{im} b_{mj}$$

$$\begin{pmatrix} 1 & 2 \\ 4 & 9 \\ 3 & 2 \end{pmatrix} \times \begin{pmatrix} 4 & 2 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} \boxed{} & \\ & \\ & \end{pmatrix}$$

1x4+

Matrices

- Multiplication

$$AB = C \rightarrow c_{ij} = \sum_{n=1}^{cols_A} a_{im} b_{mj}$$

$$\begin{pmatrix}1 & 2\\ 4 & 9\\ 3 & 2\end{pmatrix} \times \begin{pmatrix}4 & 2\\ 1 & 4\end{pmatrix} = \begin{pmatrix}\boxed{6} & \\ & \\ & \end{pmatrix}$$

1x4+2x1=6

Matrices

- Multiplication

$$AB = C \rightarrow c_{ij} = \sum_{n=1}^{cols_A} a_{im} b_{mj}$$

$$\begin{pmatrix} \boxed{1 \quad 2} \\ 4 \quad 9 \\ 3 \quad 2 \end{pmatrix} \times \begin{pmatrix} 4 & \boxed{\begin{matrix} 2 \\ 4 \end{matrix}} \\ 1 & \end{pmatrix} = \begin{pmatrix} 6 & \boxed{10} \\ & \\ & \end{pmatrix}$$

1x2+2x4=10

Matrices

- Multiplication

$$AB = C \rightarrow c_{ij} = \sum_{n=1}^{cols_A} a_{im} b_{mj}$$

$$\begin{pmatrix} 1 & 2 \\ 4 & 9 \\ 3 & 2 \end{pmatrix} \times \begin{pmatrix} 4 & 2 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 6 & 10 \\ 25 & 44 \\ 14 & 14 \end{pmatrix}$$

Matrix Inverse

- Denoted A^{-1}
- $A^{-1}A = AA^{-1} = I$
- Only for square matrices
- Only exists if matrix is full rank
  - All columns (rows) are linearly independent
- $A^{-1} \neq \left[\frac{1}{a_{ij}}\right]$, but I'll spare the details

Rank Deficient Matrices

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 4 \\ 3 & 3 & 6 \end{pmatrix}$$

2*column1=column3

$$\begin{pmatrix} 1 & 0 & 1 \\ 3 & 1 & 4 \\ 2 & 1 & 3 \end{pmatrix}$$

column1+column2=column3

Pseudoinverse

- If the columns *only* are linearly independent, then $A'A$ is invertible
- Pseudoinverse: $(A'A)^{-1}A'$
- $(A'A)^{-1}A'A = I$
- Matrix doesn't need to be square

Expectation and Variance

- $E[Y] = \begin{pmatrix} E[y_1] \\ E[y_2] \\ E[y_3] \end{pmatrix}$

- $\mathrm{Var}[Y] = \begin{pmatrix} \mathrm{Var}[y_1] & \mathrm{Cov}[y_1, y_2] & \mathrm{Cov}[y_1, y_3] \\ \mathrm{Cov}[y_1, y_2] & \mathrm{Var}[y_2] & \mathrm{Cov}[y_2, y_3] \\ \mathrm{Cov}[y_1, y_3] & \mathrm{Cov}[y_2, y_3] & \mathrm{Var}[y_3] \end{pmatrix}$

Expectation and Variance

- Let RV Y be a nx1 vector and a is a 1xn vector of constants
  - $E[aY] = aE[Y]$
  - $E[a' + Y] = a' + E[Y]$
  - $\mathrm{Var}[aY] = a\mathrm{Var}[Y]a'$
  - $\mathrm{Var}[a' + Y] = \mathrm{Var}[Y]$

Matrix Operations

- A few final properties

$$(AB)' = B'A'$$

$$(A')' = A$$

$$(A^{-1})^{-1} = A \quad \text{(when } A \text{ is invertible)}$$

$$(AB)^{-1} = B^{-1}A^{-1} \quad \text{(when } A \text{ and } B \text{ are invertible)}$$

Back to linear regression

$$\begin{array}{cccccc}
Y_1 = & \beta_0 + & \beta_1 X_{11} + & \beta_2 X_{21} + & \beta_3 X_{31} + & \epsilon_1 \\
Y_2 = & \beta_0 + & \beta_1 X_{12} + & \beta_2 X_{22} + & \beta_3 X_{32} + & \epsilon_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
Y_n = & \beta_0 + & \beta_1 X_{1n} + & \beta_2 X_{2n} + & \beta_3 X_{3n} + & \epsilon_n
\end{array}$$

$$\underset{(nx1)}{\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}} = \underset{(nx4)}{\begin{pmatrix} 1 & X_{11} & X_{21} & X_{31} \\ 1 & X_{12} & X_{22} & X_{32} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} & X_{3n} \end{pmatrix}} \underset{(4x1)}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}} + \underset{(nx1)}{\begin{pmatrix} \epsilon_1 \\ \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}}$$

Back to linear regression

$$
\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \beta_3 X_{31} + \epsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{12} + \beta_2 X_{22} + \beta_3 X_{32} + \epsilon_2 \\
\vdots \quad & \quad \vdots \qquad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad \vdots \\
Y_n &= \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + \beta_3 X_{3n} + \epsilon_n
\end{aligned}
$$

$$\Downarrow$$

$$Y = X\beta + \epsilon$$

(nx1) (nx4) (4x1) (nx1)

Viewing the Design Matrix

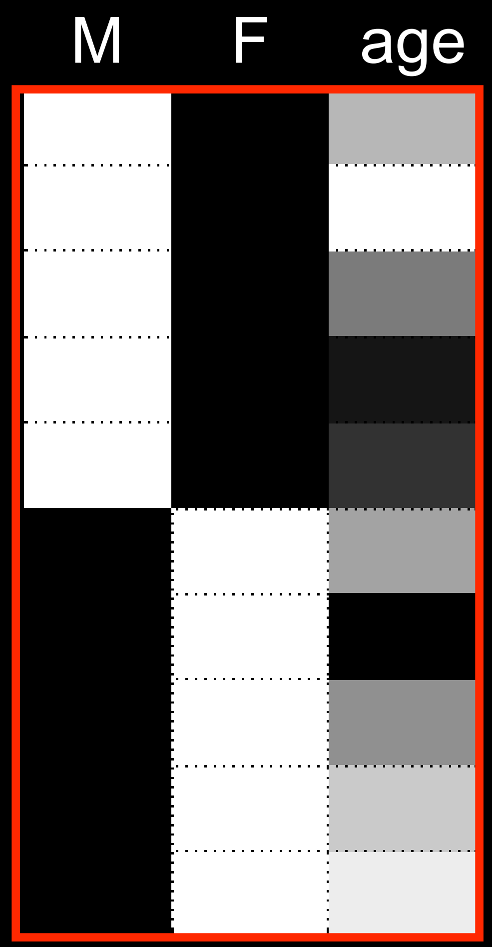
- Look at the actual numbers

$$
\begin{array}{ccc}
M & F & age \\
\end{array}
$$

$$
\begin{pmatrix}
1 & 0 & 29 \\
1 & 0 & 33 \\
1 & 0 & 26 \\
1 & 0 & 22 \\
1 & 0 & 23 \\
0 & 1 & 28 \\
0 & 1 & 21 \\
0 & 1 & 27 \\
0 & 1 & 30 \\
0 & 1 & 32
\end{pmatrix}
$$

Viewing the Design Matrix

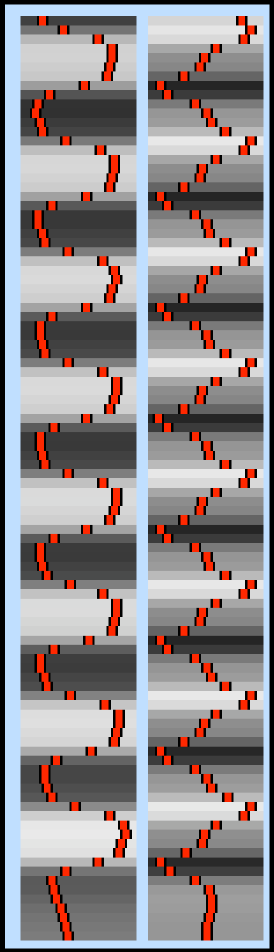
- Look at in image representation
  - Darker=smaller #



Viewing the Design Matrix

- Look at in image representation
  - Darker=smaller #
  - Useful for large fMRI designs

fMRI example (FSL)



Multiple Linear Regression

- The distribution of Y is a multivariate Normal

$$Y \sim N\left(X\beta, \sigma^2 I_n\right)$$

$$\sigma^2 \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} = \begin{pmatrix} \sigma^2 & & & 0 \\ & \sigma^2 & & \\ & & \ddots & \\ 0 & & & \sigma^2 \end{pmatrix}$$

Multiple Linear Regression

- $\hat{\beta}$ is really easy to derive

$$
\begin{aligned}
Y &= X\hat{\beta} \\
X'Y &= (X'X)\hat{\beta} \\
(X'X)^{-1}X'Y &= \hat{\beta}
\end{aligned}
$$

Multiple Linear Regression

- $\hat{\beta}$ is really easy to derive

$$Y = X\hat{\beta}$$

$$X'Y = (X'X)\hat{\beta}$$

$$\underline{\underline{(X'X)^{-1}X'Y = \hat{\beta}}}$$

Same as least squares, but much easier to understand…thanks linear algebra!

Multiple Linear Regression

- $\hat{\sigma}^2 = \dfrac{e'e}{N-p}$

  where $e = Y - X\hat{\beta} = Y - \hat{Y}$

- N=length(Y)
- p=length(β)

Statistical Properties

- $$\begin{aligned} E[\hat{\beta}] &= E[(X'X)^{-1}X'Y] \\ &= (X'X)^{-1}X'E[Y] \\ &= (X'X)^{-1}X'X\beta \\ &= \beta \end{aligned}$$

So the estimate is unbiased

Statistical Properties

- $E[\hat{\beta}] = E[(X'X)^{-1}X'Y]$
  $$= (X'X)^{-1}X'E[Y]$$
  $$= (X'X)^{-1}X'X\beta$$
  $$= \beta$$
  So the estimate is unbiased
- $\mathrm{Var}[\hat{\beta}] = \mathrm{Var}[(X'X)^{-1}X'Y]$
  $$= ((X'X)^{-1}X')\ Var[Y]\ ((X'X)^{-1}X')'$$

Statistical Properties

- $E[\hat{\beta}] = E[(X'X)^{-1}X'Y]$
$$= (X'X)^{-1}X'E[Y]$$
$$= (X'X)^{-1}X'X\beta$$
$$= \beta$$
So the estimate is unbiased

- $\mathrm{Var}[\hat{\beta}] = \mathrm{Var}[(X'X)^{-1}X'Y]$
$$= ((X'X)^{-1}X')\ Var[Y]\ ((X'X)^{-1}X')'$$
$$= (X'X)^{-1}X'\ \sigma^2 I\ X(X'X)^{-1}$$

Statistical Properties

- $E[\hat{\beta}] = E[(X'X)^{-1}X'Y]$
$$\begin{aligned} &= (X'X)^{-1}X'E[Y] \\ &= (X'X)^{-1}X'X\beta \\ &= \beta \end{aligned}$$
So the estimate is unbiased

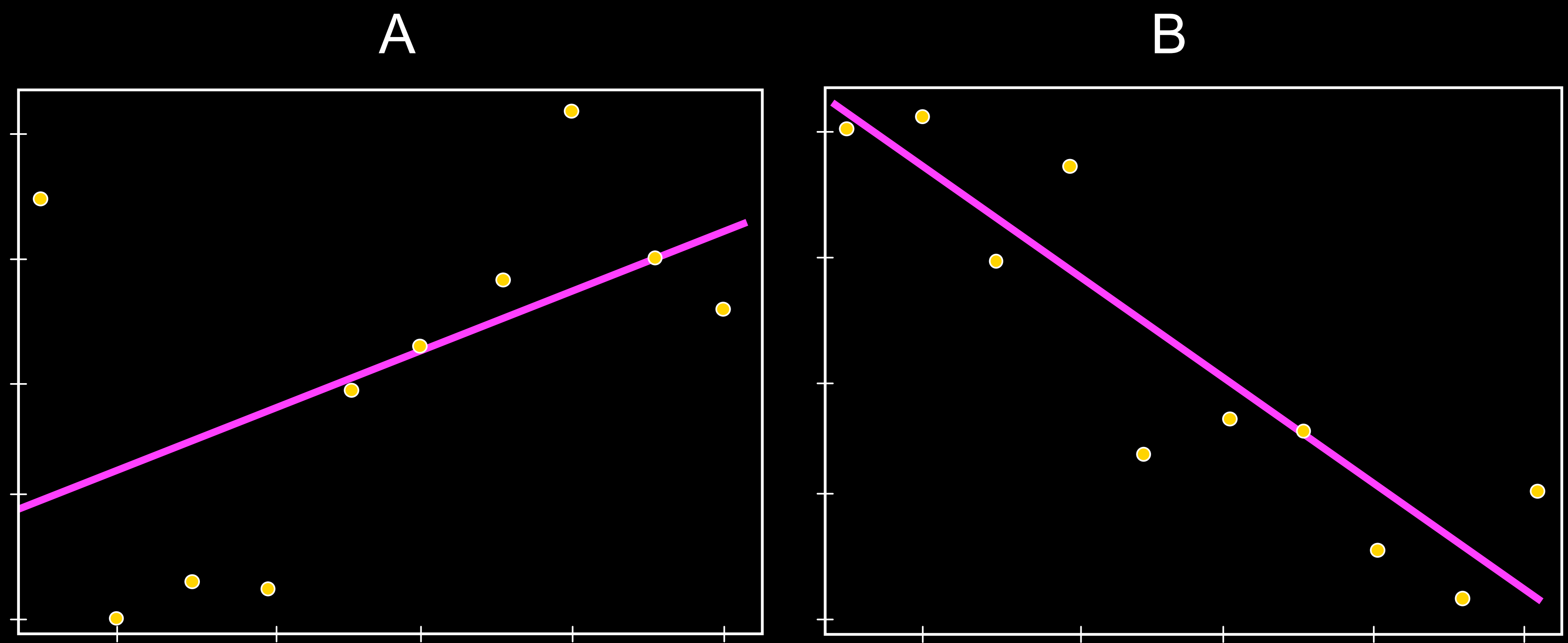
- $\mathrm{Var}[\hat{\beta}] = \mathrm{Var}[(X'X)^{-1}X'Y]$
$$\begin{aligned} &= ((X'X)^{-1}X')\ Var[Y]\ ((X'X)^{-1}X')' \\ &= (X'X)^{-1}X'\ \sigma^2 I\ X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1} \end{aligned}$$

Linear regression is flexible!

- One sample t-test
- Two sample t-test
- Paired t-test
- ANOVA
- ANCOVA
- Correlation analysis
- So, we call it the general linear model (GLM)

Hypothesis Testing

- How we evaluate the estimates
- Fitted model B matches the data better than fitted model A



5 Parts of Hypothesis Tests

- The null hypothesis, H_0
- The alternative hypothesis, H_A
- The test statistic and p-value
- The rejection region
- The conclusion about the hypothesis

H_0 and H_A

- Null Hypothesis, H_0
  - Typically what you want to disprove
  - H_0: My boyfriend is cheating on me
- Alternative Hypothesis, H_A
  - Typically what you want to be true
  - H_A: My boyfriend isn’t cheating on me

How to use H_0 and H_A

- Assuming the null is true (my boyfriend is cheating on me), how likely are my data?
  - Case 1: He buys me gifts, emails me throughout the day, cooks me dinner, tells everybody how awesome it is that he's dating a biostatistician
    - If he were cheating on me, these things wouldn't be very likely…so reject H_0 in favor of H_A

How to use H_0 and H_A

- Assuming the null is true (my boyfriend is cheating on me), how likely are my data?
  - Case 1: He buys me gifts, emails me throughout the day, cooks me dinner, tells everybody how awesome it is that he's dating a biostatistician
    - If he were cheating on me, these things wouldn't be very likely…so reject H_0 in favor of H_A
  - Case 2: He stays out late, never says anything nice to me, keeps talking about his fun female coworker, has lipstick on his collar
    - If he were cheating on me, these things would be very likely…so do not reject H_0.

H_0 and H_A in GLM

- Your study
  - How is reaction time associated with age?
  - $RT = \beta_0 + AGE\beta_{age} + \epsilon$
- Two-sided hypothesis
  - As age increases, reaction time changes
  - $H_0 : \beta_{age} = 0 \quad \text{versus} \quad H_A : \beta_{age} \neq 0$
  - Rejection of null means slope is positive or negative

H_0 and H_A in GLM

- One-sided hypothesis test
  - As age increases reaction time decreases
  - $H_0 : \beta_{age} \geq 0 \quad \text{versus} \quad H_A : \beta_{age} < 0$
  - Rejecting null only concludes a negative slope
  - Typically the type of hypothesis test for fMRI

Test Statistic

- Decision about H_0 is based on our data
- We need a statistic with a known distribution!
  - $\hat{\beta}_{age} \sim N(\beta_{age}, \mathrm{Var}(\hat{\beta}_{age}))$
  - Ugh! We don't know $\mathrm{Var}(\hat{\beta}_{age})$

Test Statistic

- We do know

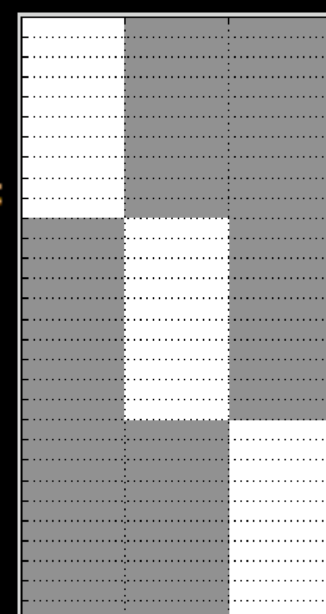
$$t = \frac{\hat{\beta}_{age}}{\sqrt{\widehat{\mathrm{Var}}(\hat{\beta}_{age})}} \sim T_{N-p}$$

Contrasts

- Sometimes we're interested in the sums or differences of 2 parameters
  - Compare G1 to G2

G1 G2 G3

$$X\beta =$$

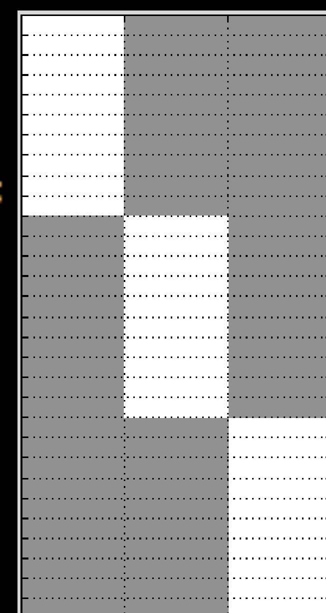


$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta 3 \end{pmatrix}$$

Contrasts

- Sometimes we're interested in the sums or differences of 2 parameters
  - Compare G1 to G2
  - $H_0 : \beta_1 - \beta_2 = 0$

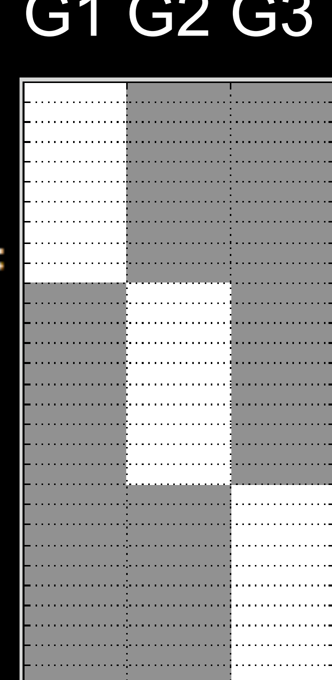
G1 G2 G3

$X\beta =$  $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta 3 \end{pmatrix}$

Contrasts

- Sometimes we're interested in the sums or differences of 2 parameters
  - Compare G1 to G2
  - $H_0 : \beta_1 - \beta_2 = 0$
  - $H_0 : c\beta = 0$
    - $c = [1 \ -1 \ 0]$
    - c is a contrast

$X\beta =$



$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta 3 \end{pmatrix}$$

Test Statistic

- Of course we can test contrasts of parameters as well
  - $H_0 : c\beta = 0$
  - $t = \dfrac{c\hat{\beta}}{\sqrt{c\widehat{\mathrm{Cov}}(\hat{\beta})c'}} \sim T_{N-p}$

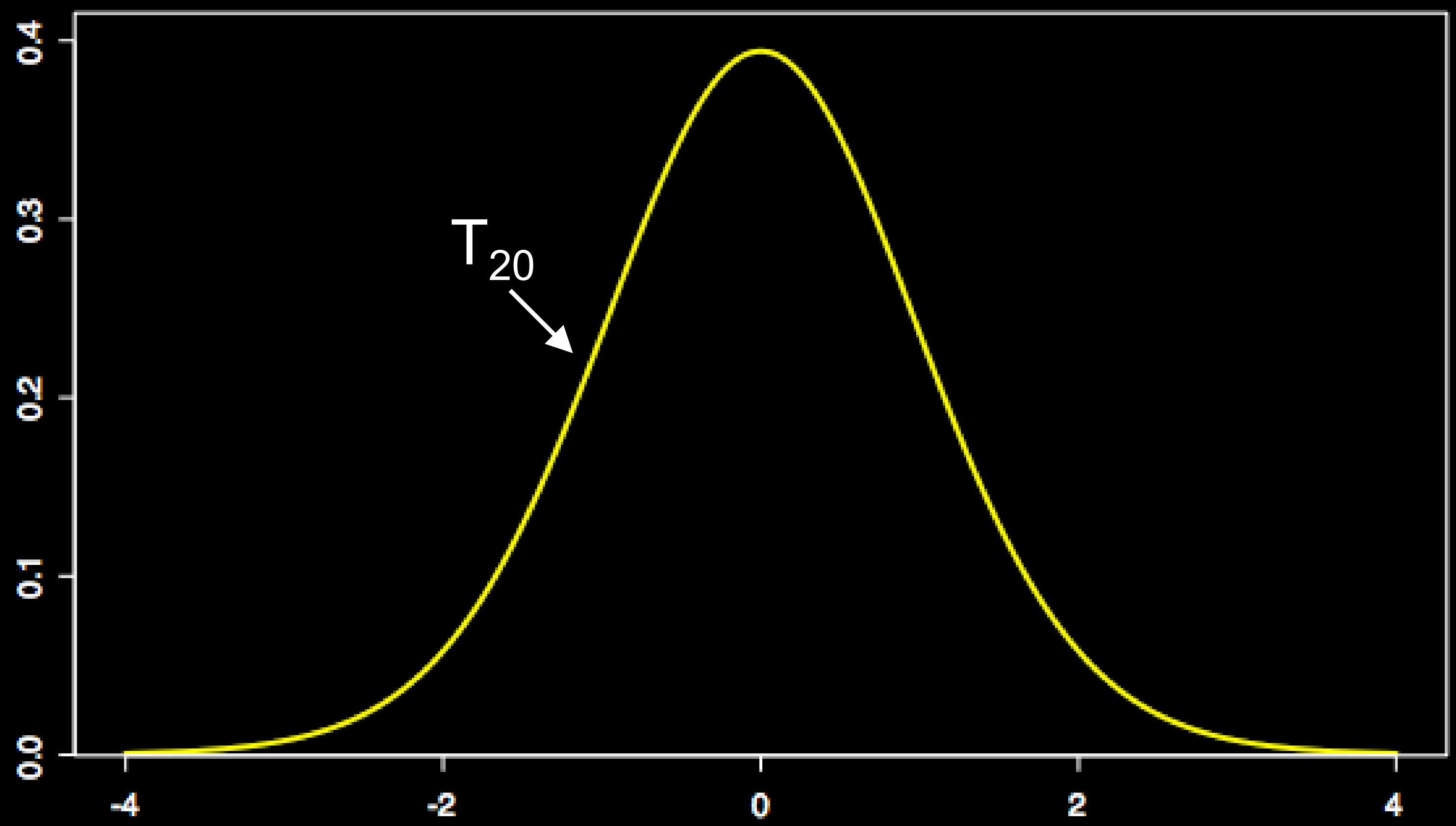
P Values 1-Sided Hypothesis

- Given the null is true, how likely is it to obtain a value more extreme than our statistic?
  - What is meant by ‘more extreme’?

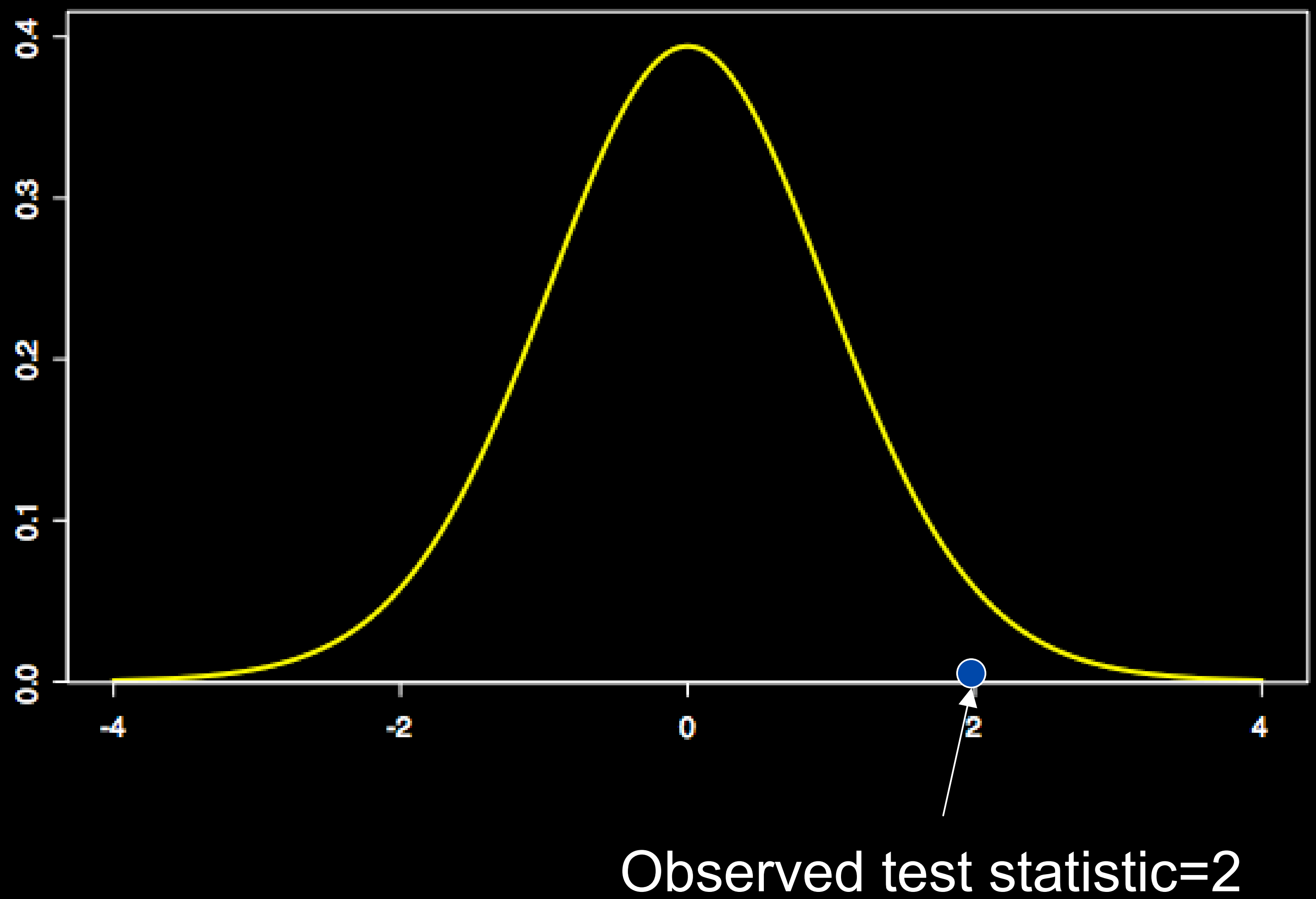
$$H_0 : \beta_{age} \leq 0 \quad \text{versus} \quad H_A : \beta_{age} > 0$$

- Start with the distribution under the null
  - There were 22 observations (N=22)
  - Simple linear regression (p=2)
  - Null is a central T distribution with 20 df

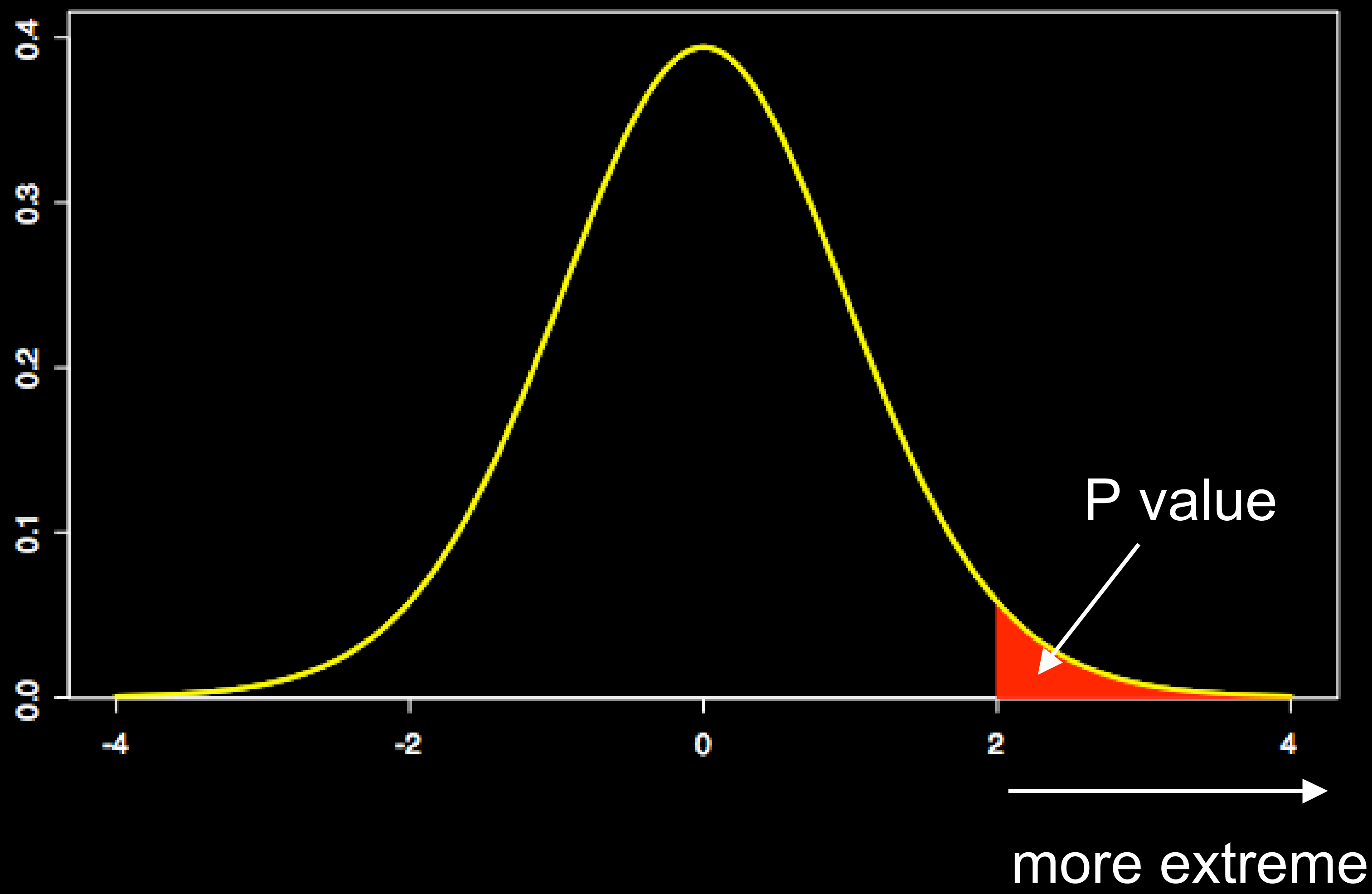
$$H_0 : \beta_{age} \leq 0 \quad \text{versus} \quad H_A : \beta_{age} > 0$$



$$H_0 : \beta_{age} \leq 0 \quad \text{versus} \quad H_A : \beta_{age} > 0$$

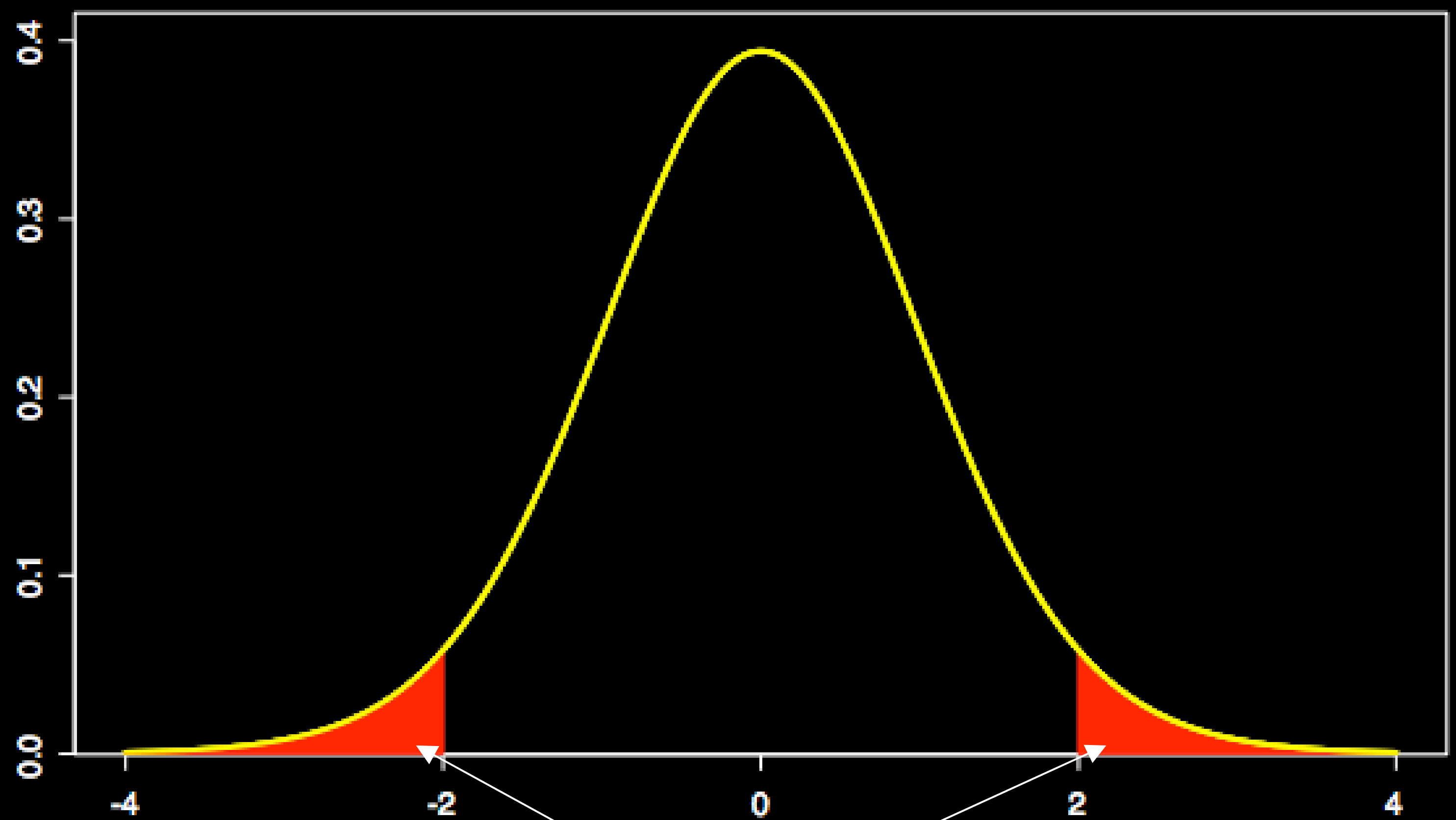


$$H_0 : \beta_{age} \leq 0 \quad \text{versus} \quad H_A : \beta_{age} > 0$$



$$H_0 : \beta_1 = 0 \quad \text{versus} \quad H_A : \beta_1 \neq 0$$

$$H_0 : \beta_1 = 0 \quad \text{versus} \quad H_A : \beta_1 \neq 0$$



Sum=P value

Assessing a P Value

- $0.1<p$
  - Data support the null
- $0.05<p<0.1$
  - Weak evidence against the null
- $0.01<p<0.05$
  - Some evidence against the null
- $0.001<p<0.01$
  - Good evidence against the null
- $p<0.001$
  - Really good evidence against the null

Notes About P Values

- The P value is not the probability that the null is true $p \neq P(H_0)$
  - $P(T_{N-p} > t|H_0)$ (one sided)

- 1-p is not the probability that the alternative is true

Rejection Region

- We need to choose a threshold
- A p value is significant if it falls below the threshold
- Denoted by α , typically set at 0.05 or 0.01
  - The probability that the null is rejected when it is true
  - For $\alpha = 0.05$ if 100 independent tests were conducted and the null was true, 5 times we'd reject the null

Types of Error

	Null Hypothesis TRUE	Null Hypothesis FALSE
Decision		
Reject null		
Accept null		

Types of Error

	Null Hypothesis: TRUE	Null Hypothesis: FALSE
Decision: Reject null		Correct!
Accept null	Correct!	

Types of Error

Null Hypothesis

Decision	TRUE	FALSE
Reject null	Type I Error α	Correct!
Accept null	Correct!	Type II Error β

Power

- Probability of rejecting the null, when the alternative is true
  - $\text{Power} = 1 - \beta$
- Ideal situation has low α and high power
  - Power is a function of α
  - Increasing α increases power

Testing Multiple Contrasts

- You can test multiple contrasts simultaneously
  - Are any of my beta's 0?
  - $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$
  - Use a contrast matrix $\longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
  - Turns into an F test
    - $f = (c\hat{\beta})'[r * c(\widehat{\mathrm{Cov}}(\hat{\beta})c']^{-1}(c\hat{\beta}) \sim F_{r,N-p}$
    - r=rank(c)

F tests are great!

- If the F test isn't significant, then none of the individual t tests will be significant
- I've heard of reviewers getting angry when two insignificant t tests were reported as opposed to 1 F test
- Why does it matter how many tests we run?

Multiple Testing Problems

- What if we perform many hypothesis tests?
- 'Confidence coefficient' $= 1 - \alpha = 0.95$
- Joint confidence coefficient for 5 independent tests
  - $(1 - \alpha)^5 = 0.95^5 = 0.77$
  - Much smaller than we'd like

Multiple Testing Problems

- Bonferroni method
  - Use $\alpha^* = \alpha/n = 0.05/5 = 0.01$
  - $(1-\alpha^*)^n = 0.99^5 = 0.951$
  - If $P(Y_i \text{ passes}|H_0) \leq \alpha/n$ then $P(\text{some } Y_i \text{ passes}|H_0) \leq \alpha$
  - With fMRI data multiple testing a a huge problem and Bonferroni is too conservative…stay tuned

Let's talk about models!

- Focus on residuals and degrees of freedom
- Goal…make our t stat as big as we can without using too many DF

$$t = \frac{c\hat{\beta}}{\hat{\sigma}^2 c(X'X)^{-1}c'} \sim T_{N-p}$$

Let's talk about models!

- Focus on residuals and degrees of freedom
- Goal…make our t stat as big as we can without using too many DF

Can't do much about these pieces

$$t = \frac{c\hat{\beta}}{\hat{\sigma}^2 c(X'X)^{-1}c'} \sim T_{N-p}$$

Let's talk about models!

- Focus on residuals and degrees of freedom
- Goal…make our t stat as big as we can without using too many DF

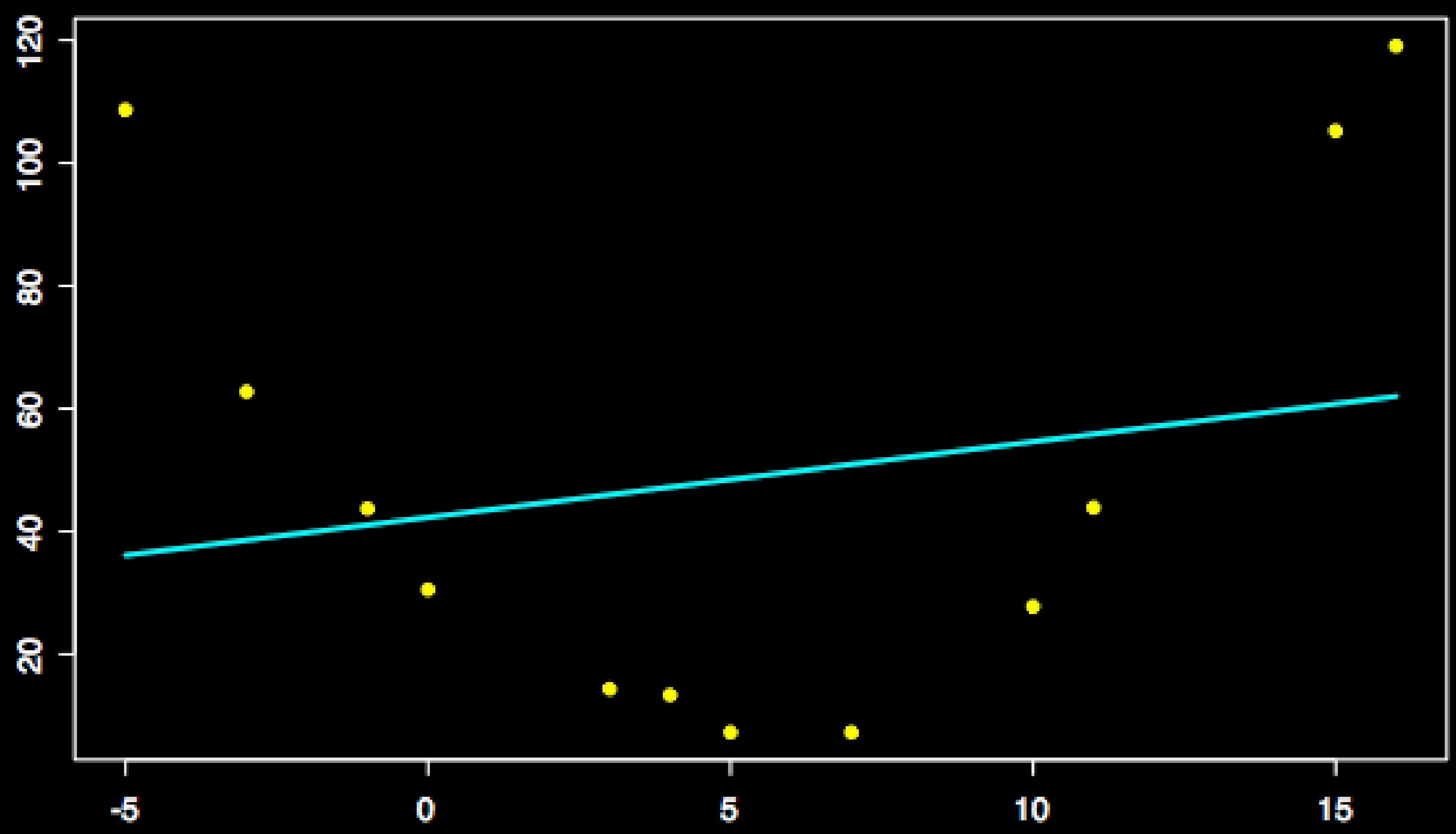
$$t = \frac{c\hat{\beta}}{\hat{\sigma}^2 c(X'X)^{-1}c'} \sim T_{N-p}$$

Try to decrease this estimate

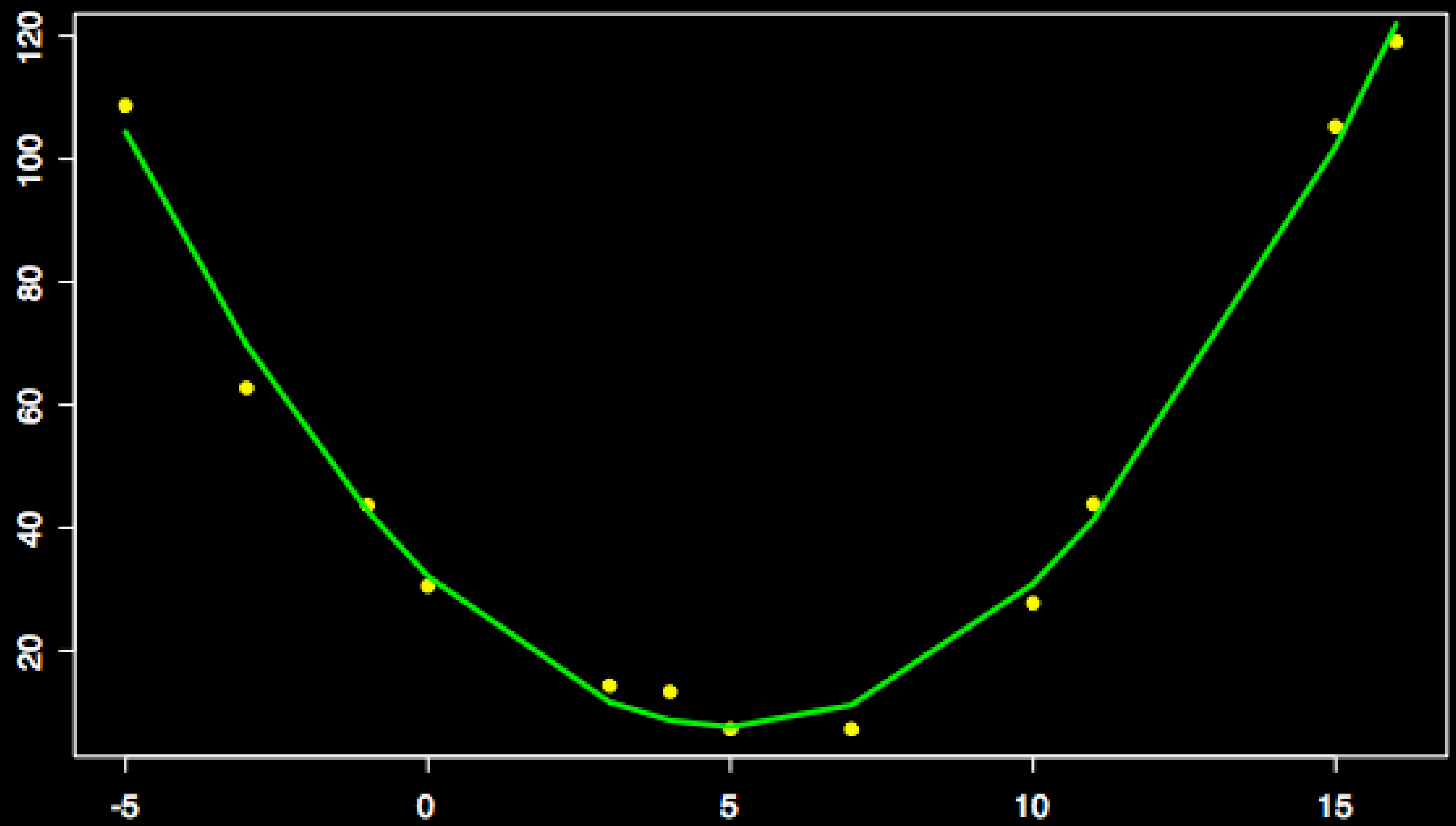
Variance Estimate

- Recall $\hat{\sigma}^2 = \dfrac{\sum (Y_i - \hat{Y}_i)^2}{N - p}$
- If we make our model fit better, the estimate will decrease
  - Add in regressors to model confounding factors (age, gender, etc)
  - Make sure the regressors you do have capture the trends you are modeling

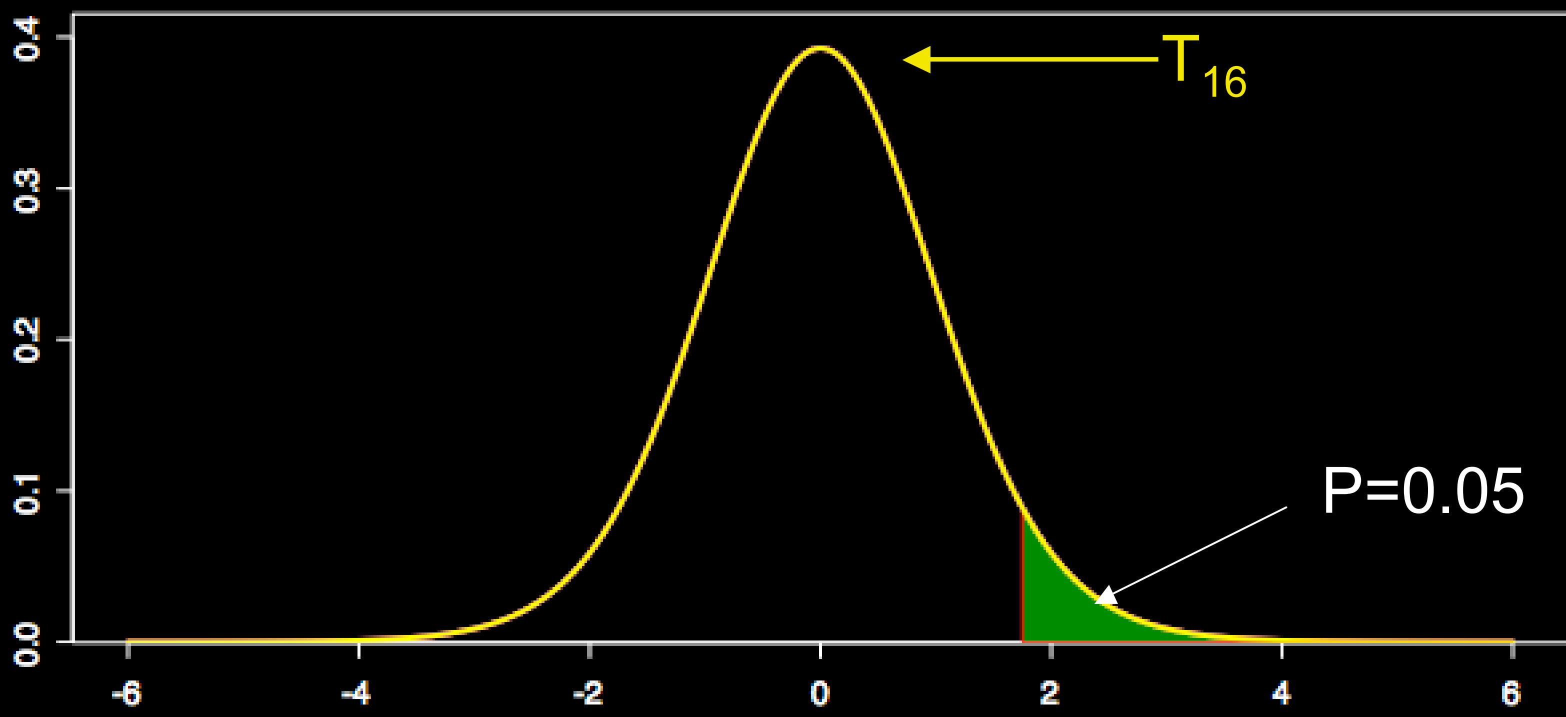
Linear regressor is not significant (p=0.5)



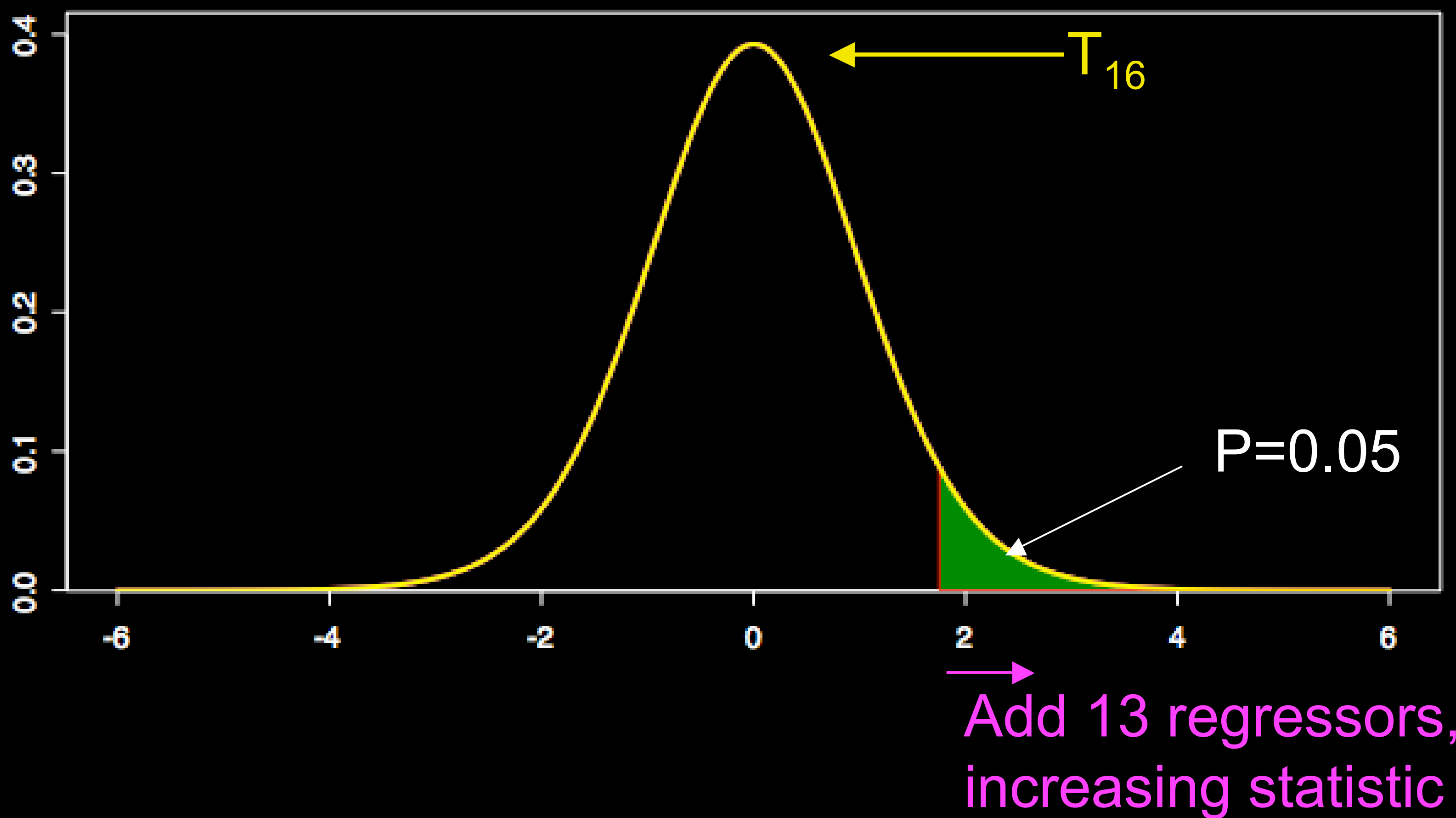
Quadratic regressor is significant ($p<0.0001$)



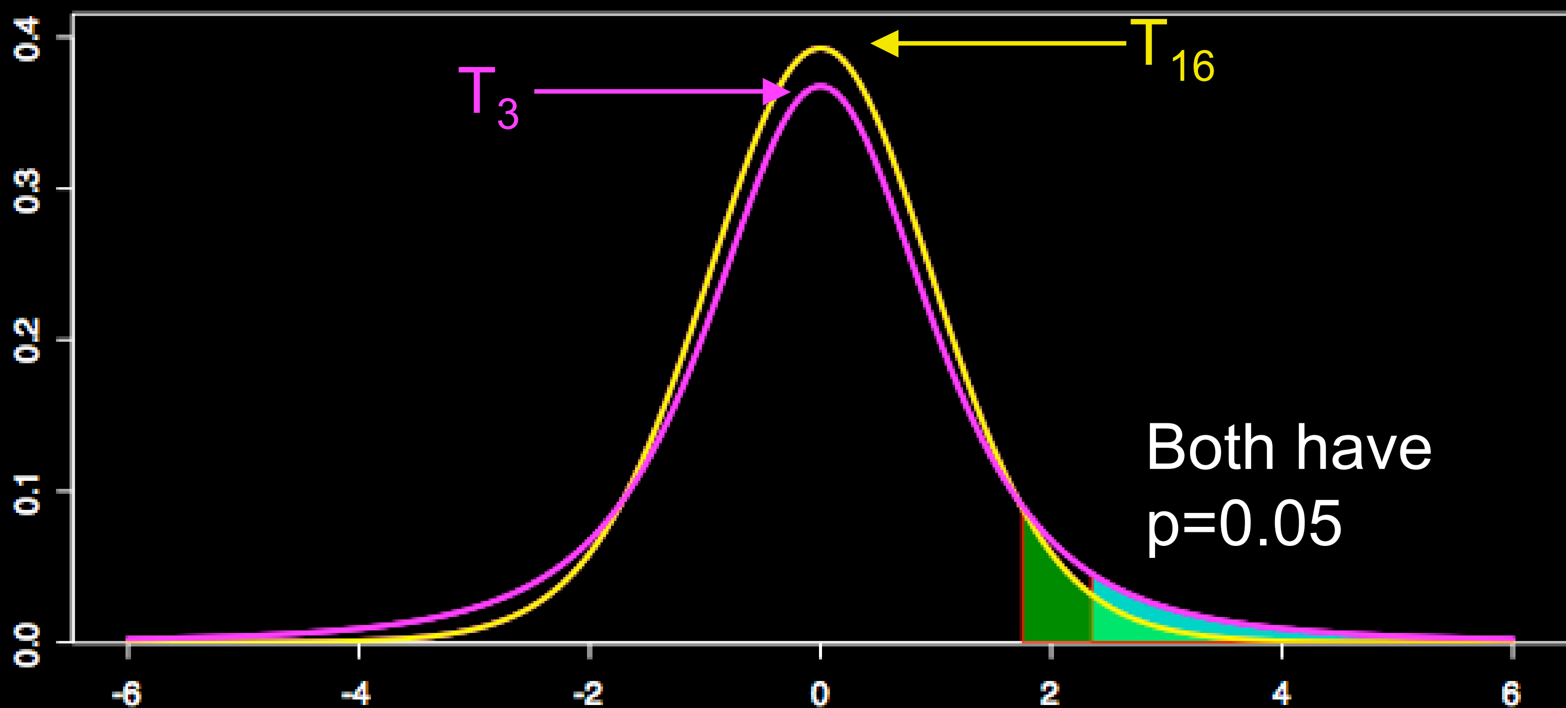
Watch degrees of freedom!



Watch degrees of freedom!



Watch degrees of freedom!



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 $c = [0.5 \quad 0.5]$
    - 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 $c = [0.5 \quad 0.5]$
    - G1-G2 $c = [0 \quad 1]$

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

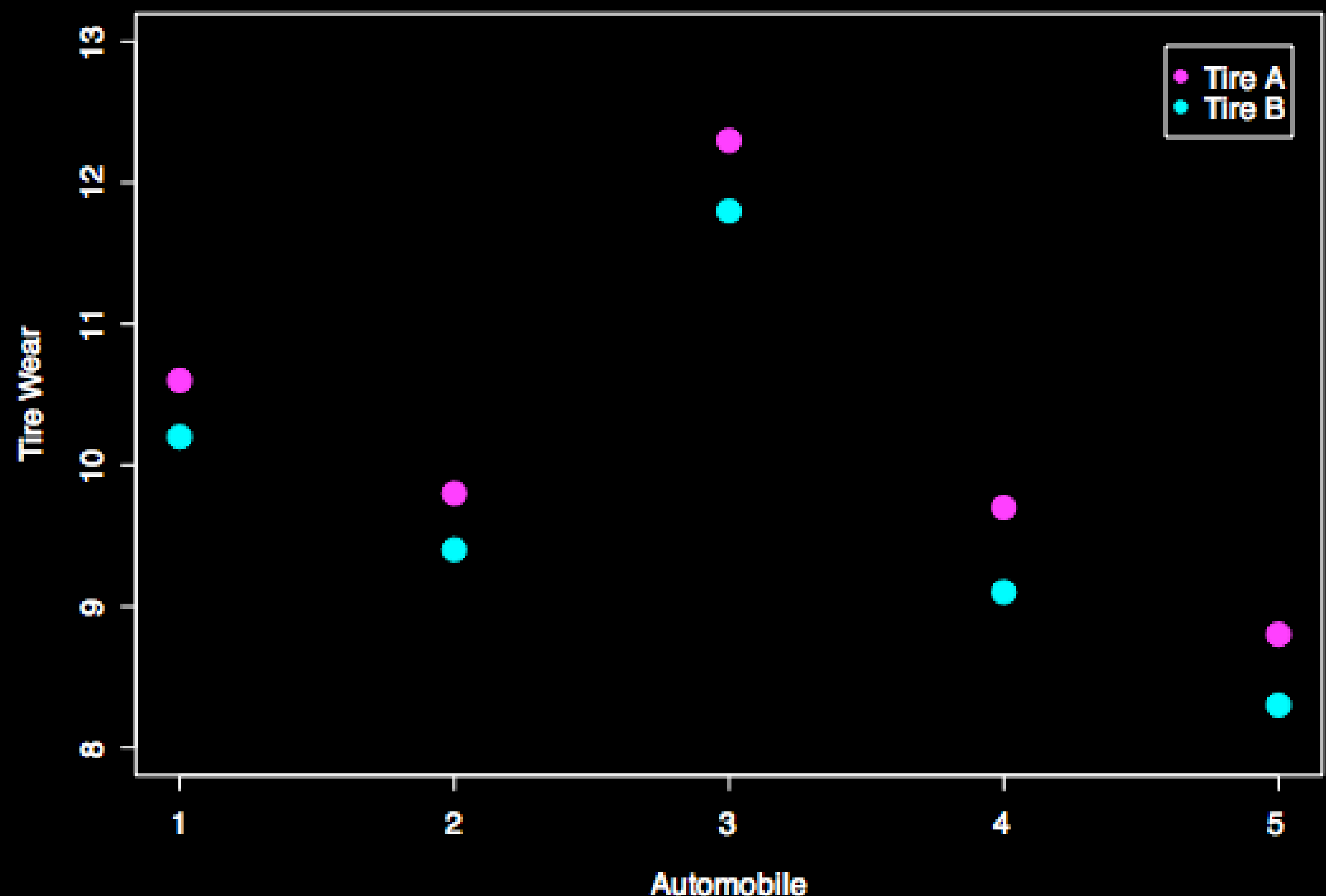
Paired T Test

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

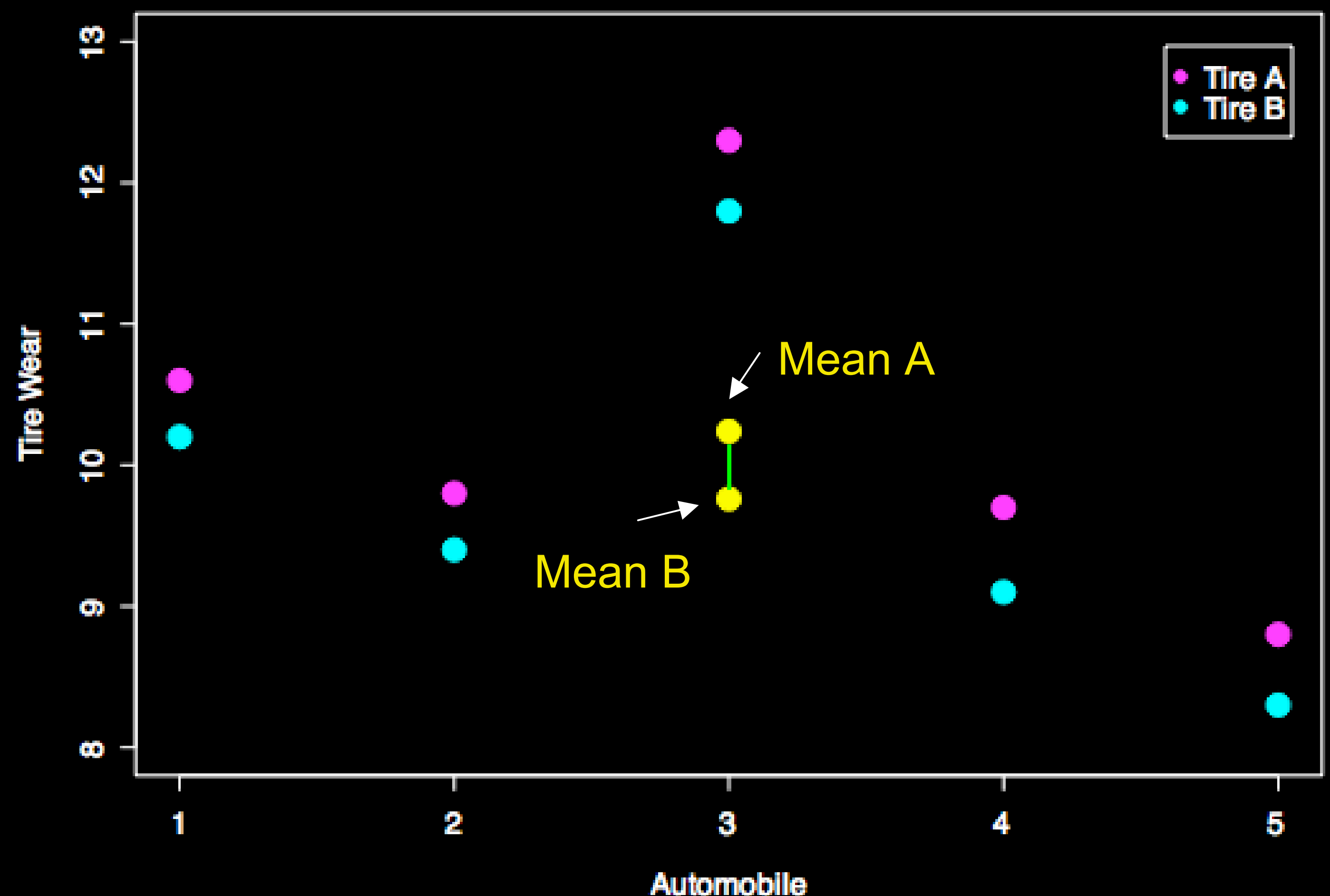
Automobile	Tire A	Tire B
1	10.6	10.2
2	9.8	9.4
3	12.3	11.8
4	9.7	9.1
5	8.8	8.3

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

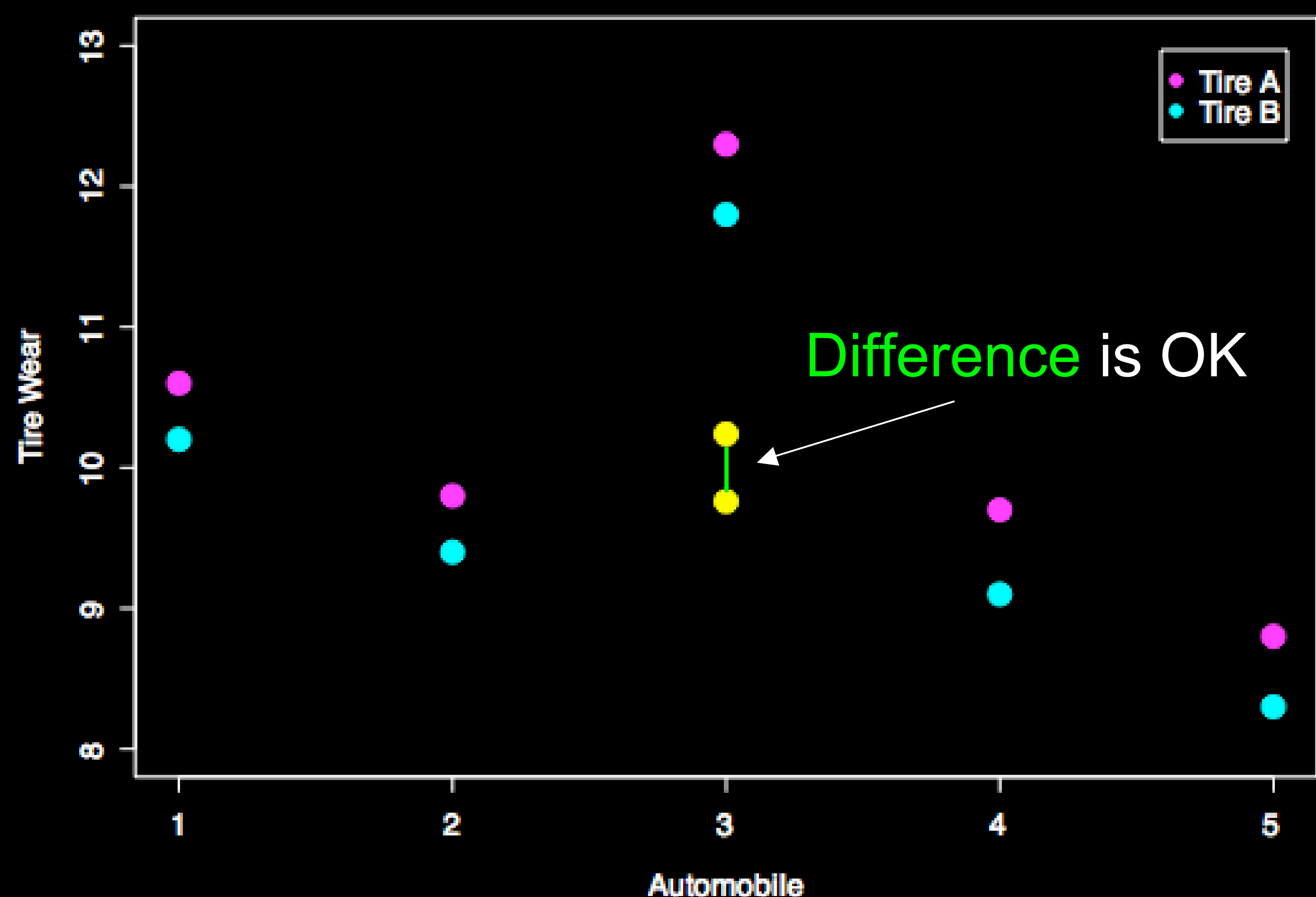
Why so different?



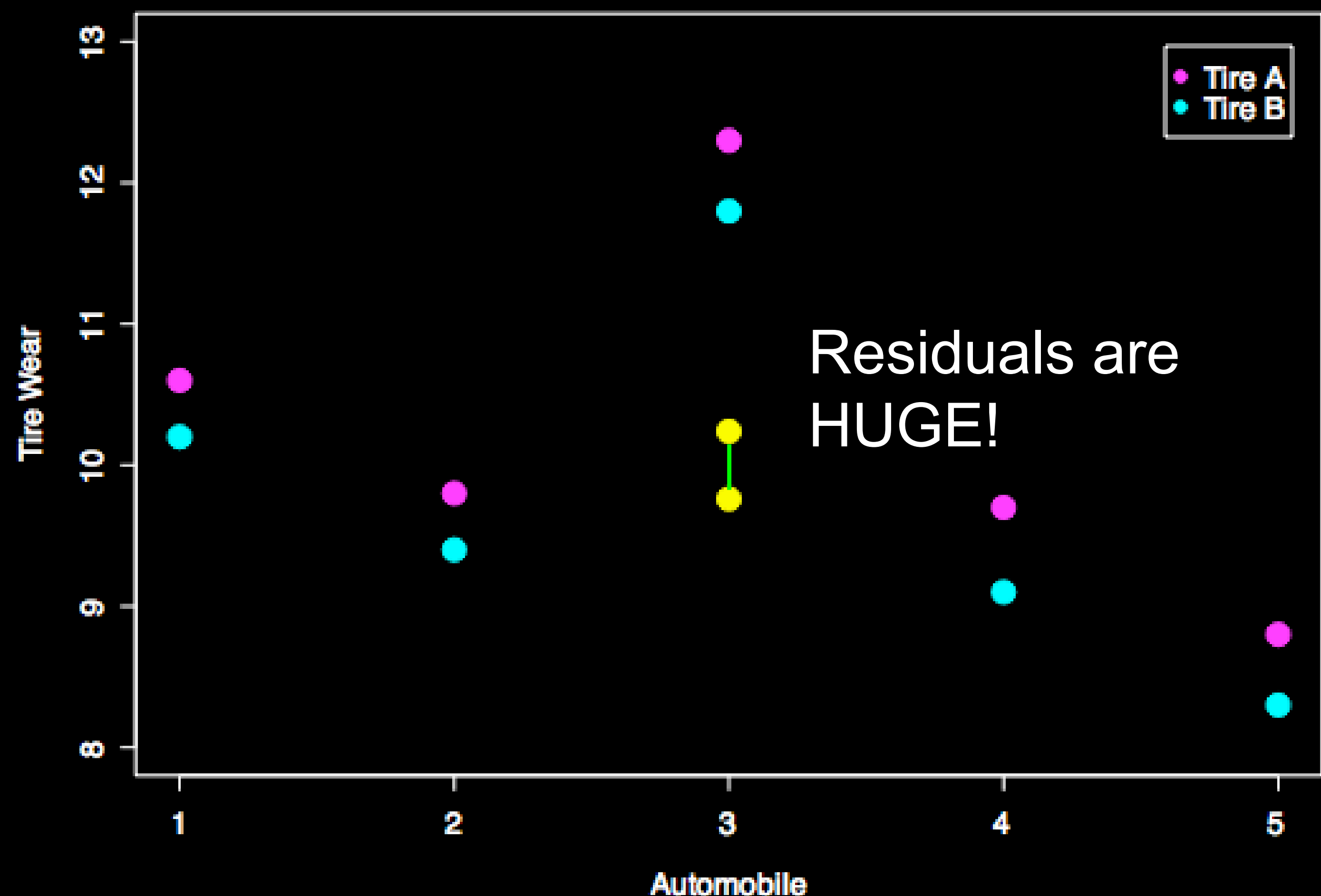
Why so different?



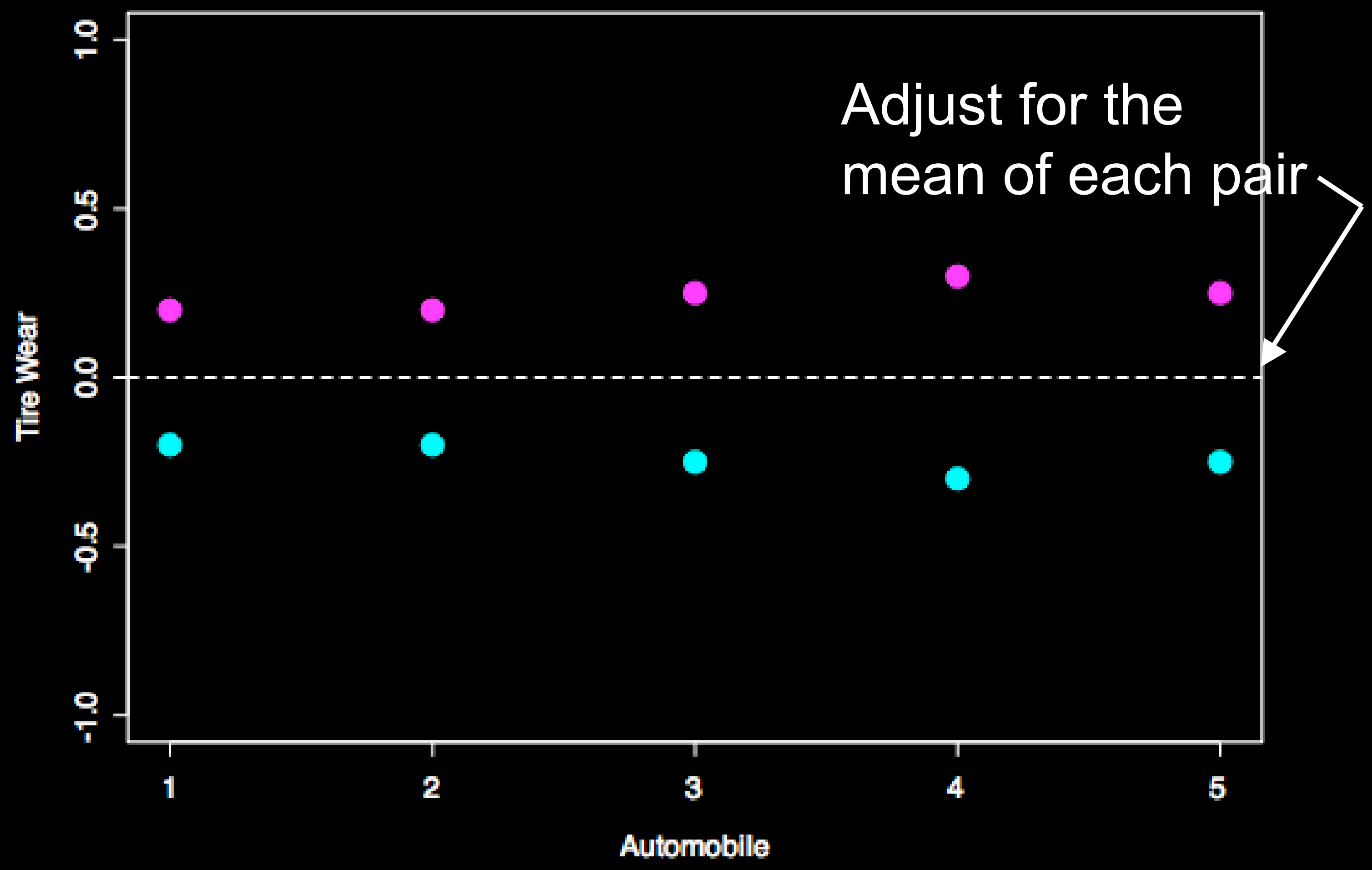
Why so different?



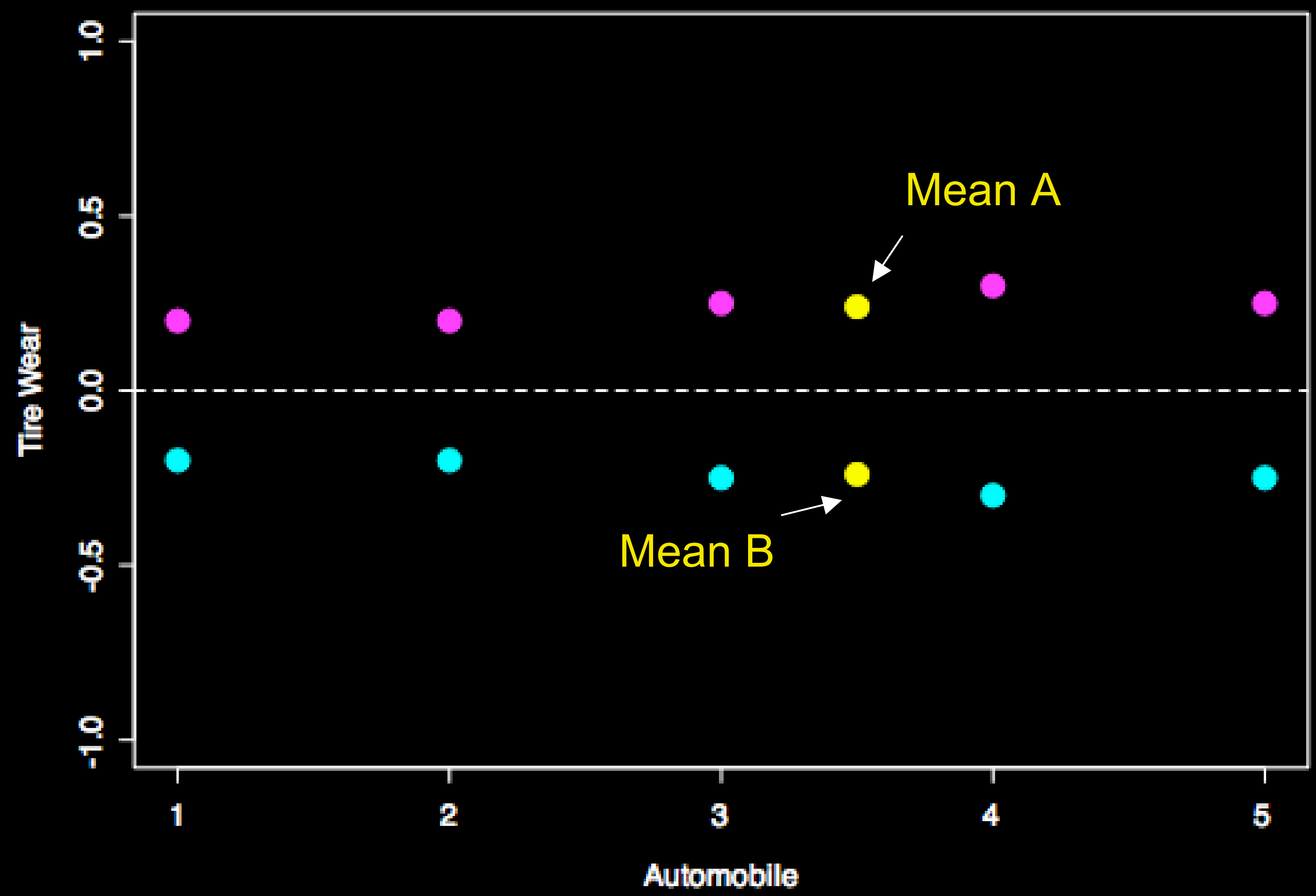
Why so different?



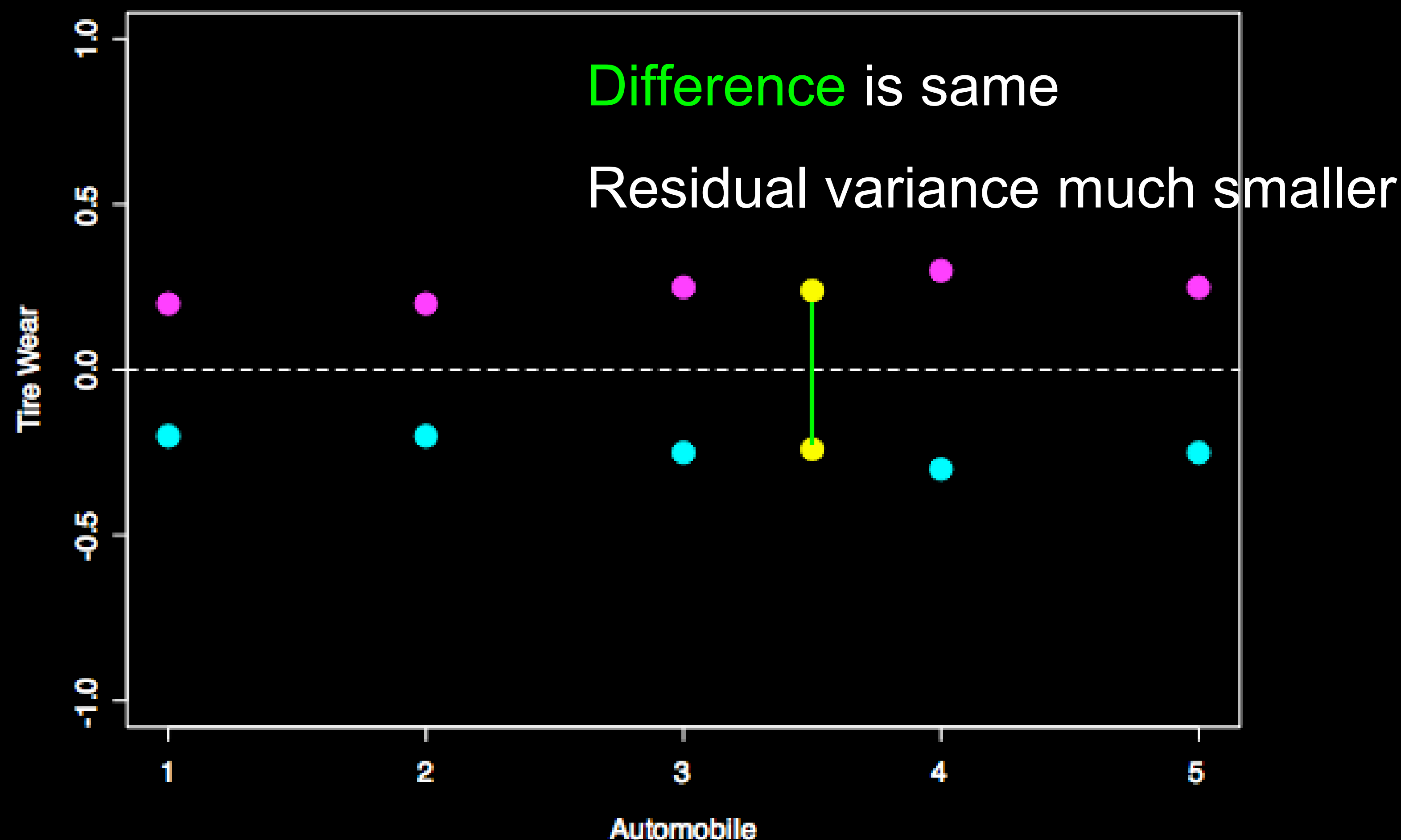
Paired T Test



Paired T Test



Paired T Test



Paired T Test GLM

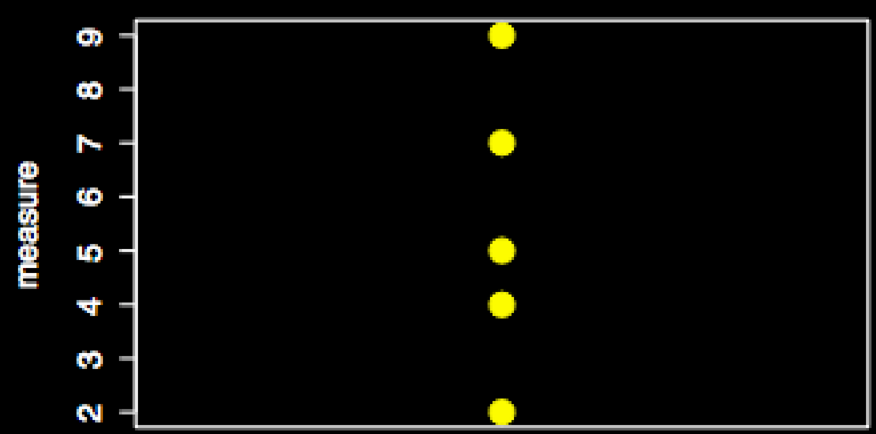
$$
\begin{pmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \\ A_3 \\ B_3 \\ A_4 \\ B_4 \\ A_5 \\ B_5 \end{pmatrix}
=
\begin{pmatrix}
1 & 1 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 \\
-1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 \\
-1 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix}
\begin{array}{l}
\leftarrow \text{Difference} \\
\left.\begin{array}{l} \\ \\ \\ \\ \\ \end{array}\right\} \text{Mean of each pair}
\end{array}
$$

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

ANOVA

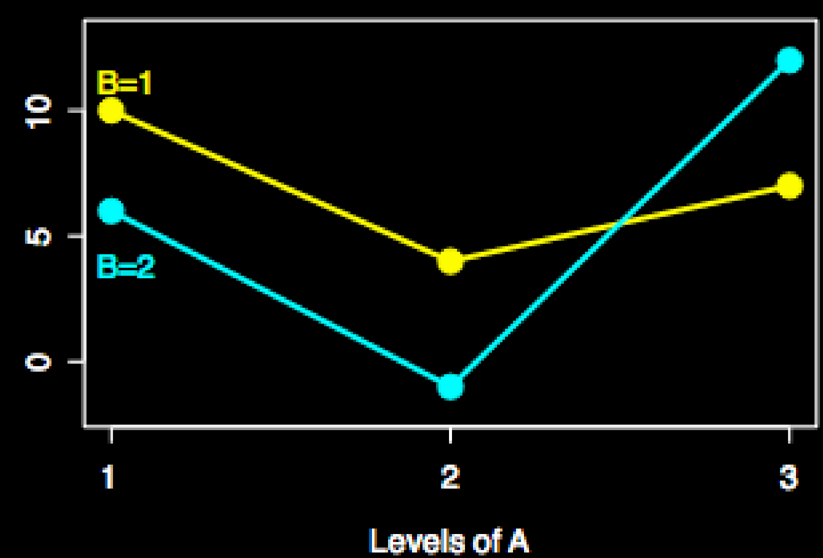
1-way ANOVA

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
  - 1-way ANOVA $Y_{in} = \mu_i + \epsilon_{in}$
  - 2-way ANOVA $Y_{ijn} = \mu_{ij} + \epsilon_{ijn}$
  - EVs are easy, but contrasts are trickier

Modeling ANOVA with GLM

- Cell means model
  - 1-way ANOVA $Y_{in} = \mu_i + \epsilon_{in}$
  - 2-way ANOVA $Y_{ijn} = \mu_{ij} + \epsilon_{ijn}$
  - EVs are easy, but contrasts are trickier
- Factor effects
  - 1-way $Y_{in} = \mu_. + \alpha_i + \epsilon_{in}$
  - 2-way $Y_{ijn} = \mu_. + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijn}$
  - EVs take more thought, but contrasts are easier
- ANOVA = F tests!

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

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

$$H_0 : G2 - G3 = 0$$

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

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

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

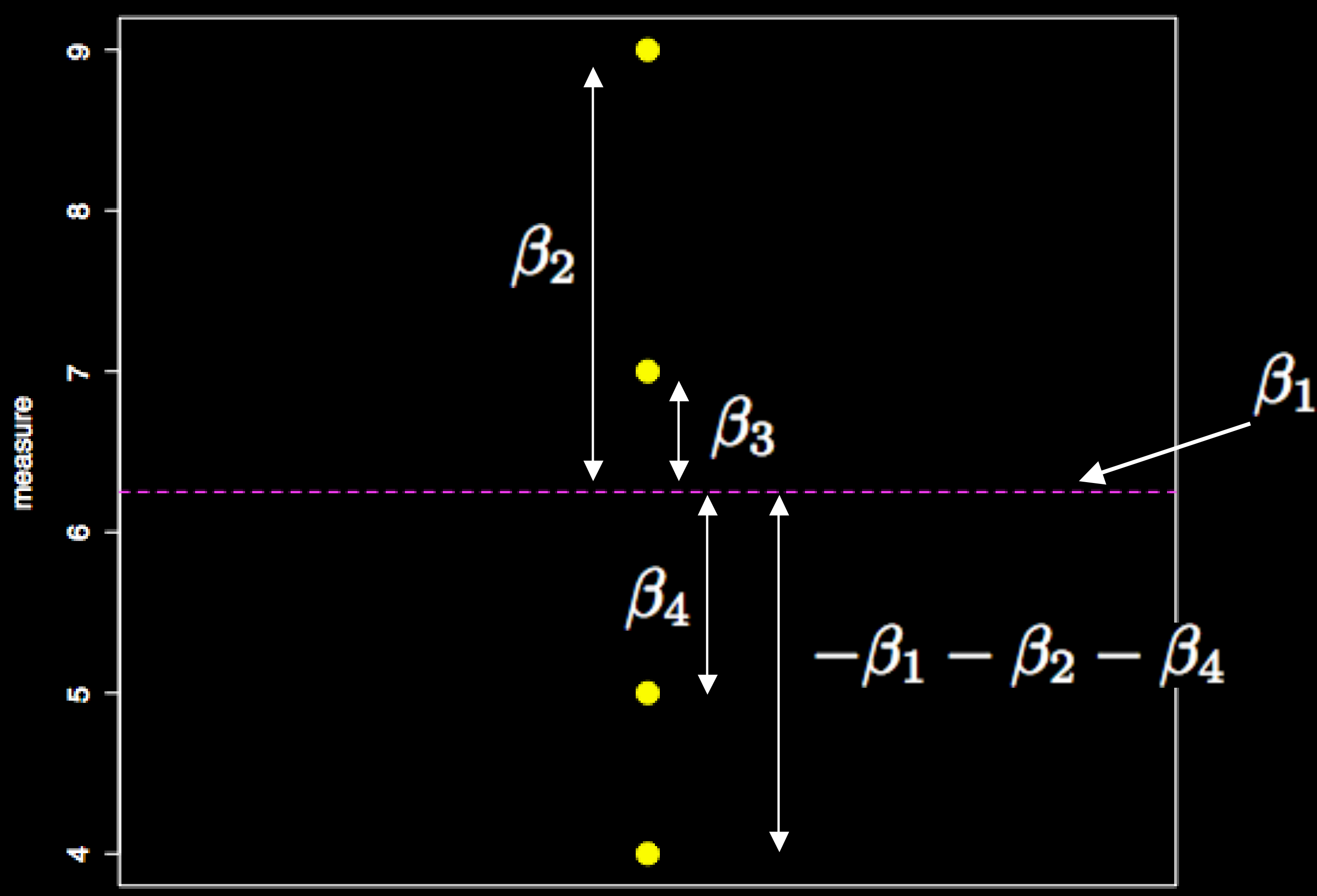
1 Way ANOVA - Factor Effects

mean

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

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

1 Way ANOVA - Factor Effects



1 Way ANOVA - Factor Effects

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

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

1 Way ANOVA - Factor Effects

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

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

$$H_0 : c\beta = 0 \quad \text{where} \quad c = [0.5 \quad 0.5 \quad 0 \quad 0]$$

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

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

$$H_0 : G1 - G4 = 0$$

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

2 Way ANOVA (3x2)

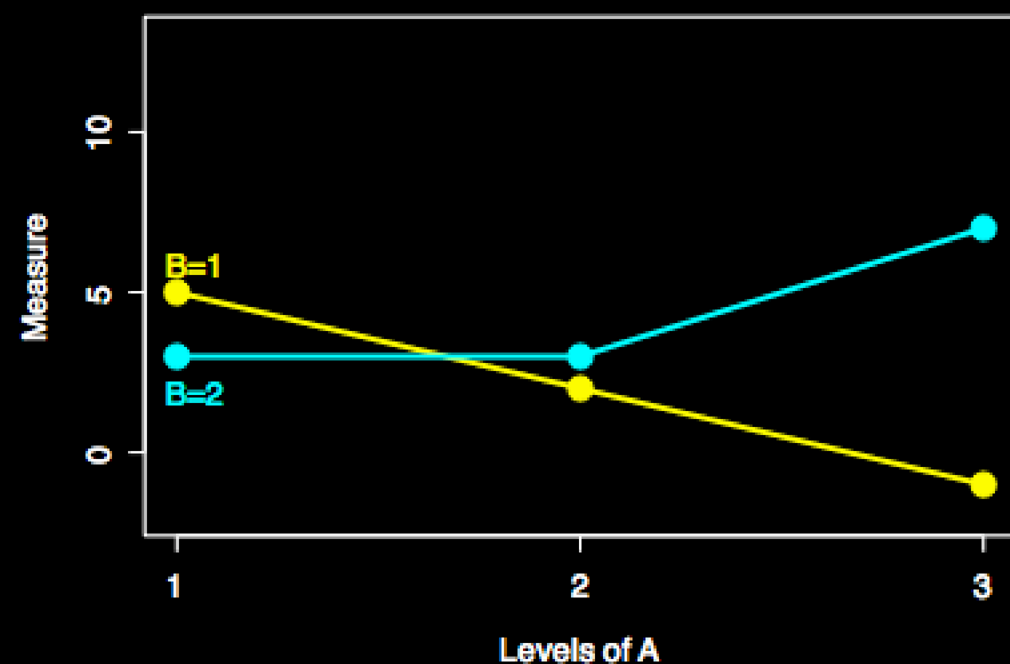
$$
\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 &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

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

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

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)

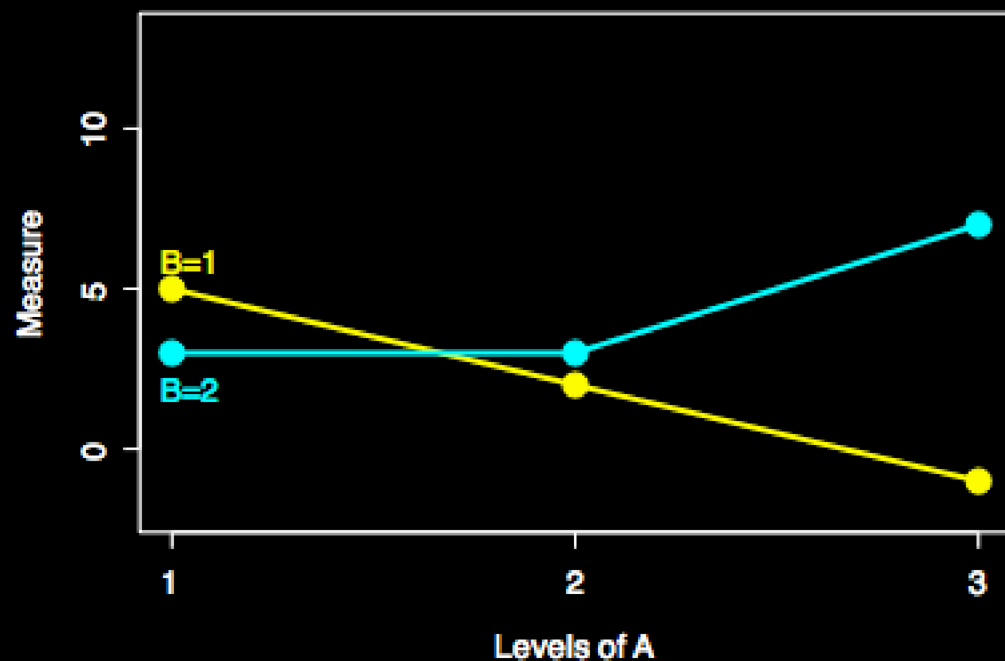
$$\begin{pmatrix} A_1B_1 \\ A_1B_1 \\ A_1B_2 \\ A_1B_2 \\ A_2B_1 \\ A_2B_1 \\ A_2B_2 \\ A_2B_2 \\ A_3B_1 \\ A_3B_1 \\ A_3B_2 \\ A_3B_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1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

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

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

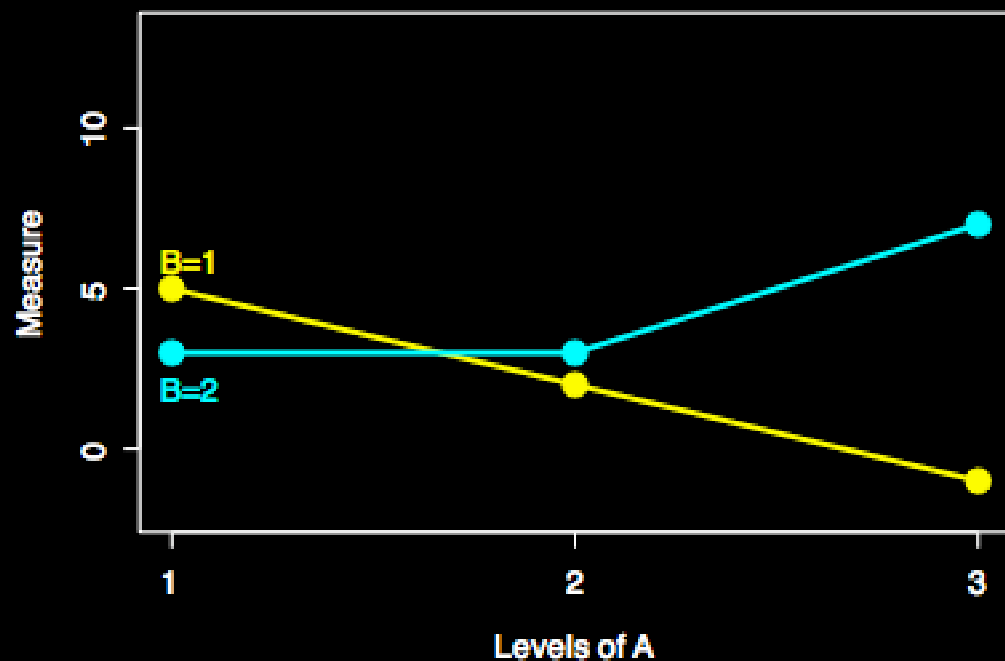


No effect means the lines would be parallel

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)

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

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

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

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

$$A_1 = \beta_1 + \beta_2$$
$$A_2 = \beta_1 + \beta_3$$
$$A_3 = \beta_1 - \beta_2 - \beta_3$$
$$B_1 = \beta_1 + \beta_4$$
$$B_2 = \beta_1 - \beta_4$$
$$A_1 B_1 \text{ interaction effect} = \beta_5$$
$$A_2 B_1 \text{ interaction effect} = \beta_6$$
$$A_1 B_2 \text{ interaction effect} = -\beta_5$$
$$A_2 B_2 \text{ interaction effect} = -\beta_6$$

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!

Why did I just tell you all of this?

- The GLM is a flexible model that allows for a variety of analyses
- Focusing on residuals and degrees of freedom will help you build good models
- Use an F test when appropriate
- A lot of the stats lingo and linear algebra stuff comes up in methods papers

Why did I just tell you all of this?

- The Gauss-Markov theorem tells us our least squares estimates are best if
  - Errors are mean zero, uncorrelated, constant variance
  - fMRI data tend to violate these assumptions
- Multiple comparisons is a huge problem with fMRI data and Bonferroni doesn’t work well