HW

1) Describe what it does, not why, look at output - some patterns → explain those → certain things about hat matrix → write down mathematically

Underlying → understanding what hat matrix is. Any decent linear regression book: chapter on regression diagnostics

→ $HH = H$

$H^n = H$

projection matrix: Take any point in the 3D space, drop perpendicular line to the plane defined by column 1 and 2 of X

$\underset{\sim}{\mu} = H\underset{\sim}{y}$, $\underset{\sim}{y} \in \mathbb{R}^n$

if $n=3$, then we have a point. $X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & n \end{bmatrix} = [\underset{\sim}{c_1} \; \underset{\sim}{c_2}]$

rank: 2

Then, $\underset{3\times1}{\underset{\sim}{\mu}} = X\underset{\sim}{\beta} = \beta_1 \underset{\sim}{c_1} + \beta_2 \underset{\sim}{c_2}$

① In 3 dimensions: $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$

pretty much all models in this class include intercept

off-diagonals are anything

$\begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\mu}_3 \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & \\ & h_{22} & \\ & & h_{33} \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$

diagonals are positive

orthogonal projection

trace = 2 in this case (rank of X)

sum of all elements in $H = n = 3 = 3$ observations

If you project any column in design matrix using the Hat matrix, then you will get that column itself.

$$\begin{array}{cc} H & c_i \\ \left[\qquad\qquad\right] & \left[\;\vdots\;\right] \end{array} = \left(\;\vdots\;\right) \rightarrow H$$

* intercept column = sum of the rows and it will give you the intercept column itself, so the sum of each row is 1

$$\sum_{i=1}^{n} (y_i - \hat{\mu}_i)^2 = (y - \hat{\mu})^T (y - \hat{\mu})$$ ← inner product

$$\sum_{i=1}^{n} a_i^2 = a^T a$$

$$= (y - Hy)^T (y - Hy)$$

$$= \{(I-H)y\}^T \{(I-H)y\}$$

$(ABC)^T = C^T B^T A^T$ ← or vector

$$= y^T (I-H)^T (I-H) y$$

$$= y^T (I-H)(I-H) y$$ ← b/c symmetry

H is symmetric, so I-H is symmetric, so H and H^T are the same.

H is projection. I-H is the residual from that projection, so is projection to a space perdendicular to H

so I-H is idempotent.

$$= y^T (I-H) y$$ ← b/c idempotent ← quadratic form

$$\rightarrow y^T (I-H) = y^T - y^T H \qquad = (y^T - y^T H) y$$

$$= y^T y - y^T H y = y^T y - y^T \hat{\mu}$$

$= \sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} y_i \hat{\mu}_i$

- Good book on Regression — by Seber
"Linear Regression"
old version of book: "Linear Hypothesis"

- another → Christensen

These ideas go back to ANOVA. Write down one-way ANOVA as regression model. Use general ANOVA model

Fixed Layout

$\mu_i = \mu + \alpha_i \longrightarrow \mu_i = \gamma + \alpha_i \qquad i = 1, \dots, K$

singular matrix → $X : \left(\sum_{i=1}^{K} n_i\right) \times (K+1)$ $\qquad n_i$ is # subjects in i^{th} group

The sum of the last K columns is equal to the intercept column

General 2 way ANOVA model:

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$$
$$i = 1, \dots, k_1$$
$$j = 1, \dots, k_2$$

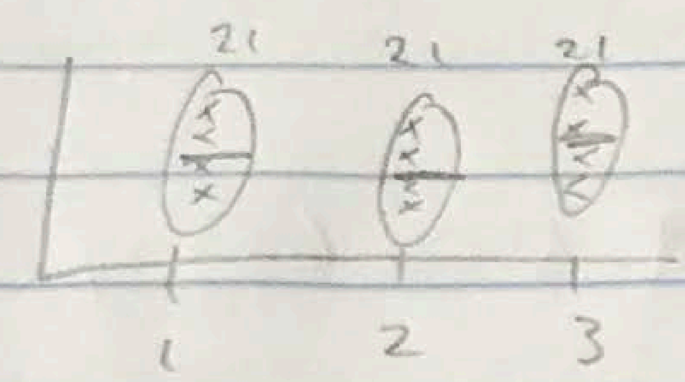
$1 + k_1 + k_2 + k_1 k_2$ ← also singular

More complex 1-way ANOVA: Random layout
↳ Mixed models

HW 1

1) One way ANOVA

increasing trend



$$s.d. = \sqrt{\frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{n_i - 1}}$$ ← sum of squares residuals in ith group

↑ mean of ith group

$$(n_i - 1)\,(s.d._i)^2 = \sum_{j=1}^{n_i} (y_{ij} - y_i)^2$$

Sum of squares within n groups:

1   64.082

\+

199.522 = SSW

y_{ij} group i, observation j in ith group

$$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

$$\bar{y} = \frac{1}{\sum_{i=1}^{K} n_i} \sum_{i=1}^{K} \sum_{j=1}^{n_i} y_{ij}$$

Total corrected for the mean?:

$$\sum_{i=1}^{K} \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$$ ← Total corrected for the mean

← last line

$$\sum_{i=1}^{K} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i + \bar{y}_i - \bar{y} + \cancel{\bar{y}} - \cancel{\bar{y}})^2$$ ← added 0

$$= \sum_{i=1}^{K} \sum_{j=1}^{n_i} \{ \underbrace{(y_{ij} - \bar{y}_i)^2}_{\text{sum of squares within}} + \underbrace{(\bar{y}_i - \bar{y})^2}_{\text{differences between groups}} \}$$ (there is a product term that falls out → zero?)

between: $\sum \sum (\bar{y}_i - \bar{y})^2$

$= \sum_{i=1}^{k} n_i (\bar{y}_i - \bar{y})^2 = 14.90$

$\rightarrow$ 2 df $\rightarrow k-1$

w/in each group: 20 df, $* 3 \rightarrow 60$

Total df: $2 + 60 = 62 = n-1$

F stat 2.24

c) Quadratic regression

$E[Y_{ij}] = \mu_i = \gamma_1 + \gamma_2 i + \gamma_3 i^2$

(a) Any 3 points can be fitted perfectly with a quadratic regression

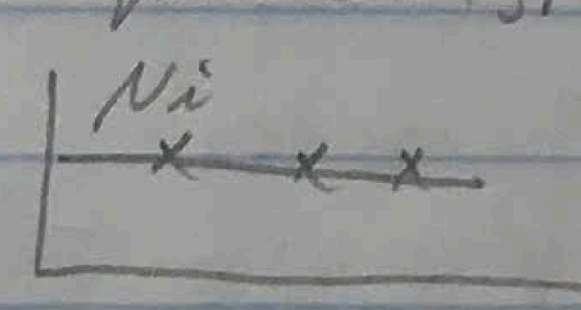
This means that the model here is <u>equivalent</u> to the ANOVA model

↳ 3 group means

This model says the same thing

They are equivalent models.

Original ANOVA:



$E[Y_{ij}] = \mu_i = \gamma_1 + \underbrace{\gamma_2 i + \gamma_3 i^2}_{=0}$

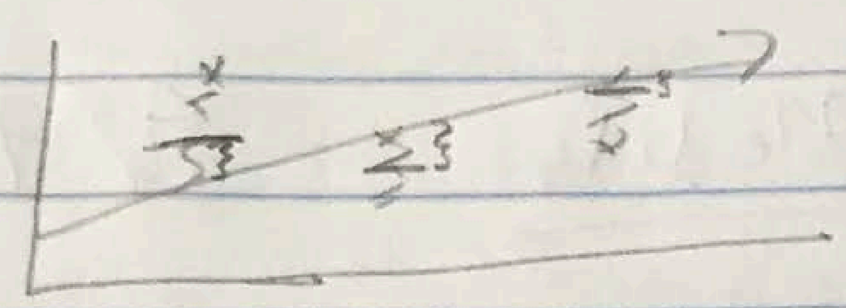
Sample midterm questions

(since the sample sizes are equal. If not equal, use sample size as weight)

Take 3 means + regress them on 1, 2, 3, you will get same slopes

b) The coefficients will be the same

Now, estimate σ^2 ← sum of sq. residuals about the fitted line



line won't go through group means

Linear regression on i

$SSW_{old} + \sum_{i=1}^{3} n_i (\bar{y}_i - \hat{\mu}_i)^2$

↑ from part a) ↑ fitted value from linear regression

Sum of squares about the regression line

- Have to understand them!

Other way: do the linear regression

$X_{63 \times 2} \rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \downarrow & \\ * \, 21 \text{ times} & \end{bmatrix}$

$\underbrace{X^T X}_{2\times 2} = \begin{bmatrix} 63 & \\ & \end{bmatrix}$ ← n

$= 21 \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$

$\underbrace{X^T y}_{2\times 1} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 3 & \cdots & \end{bmatrix} \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_{63} \end{pmatrix} = \begin{pmatrix} \\ n_1\bar{y}_1 + 2n_2\bar{y}_2 + 3n_3\bar{y}_3 \end{pmatrix}$

$= \begin{pmatrix} \\ 21(\bar{y}_1 + 2\bar{y}_2 + 3\bar{y}_3) \end{pmatrix}$

$(X^T X) X^T y \rightarrow \hat{\mu} \rightarrow$ do calculations