One-way ANOVA

$y_{i,j} = \mu + \alpha_i + \varepsilon_{i,j}$ $\quad \varepsilon_{i,j} \overset{iid}{\sim} N(0, \sigma^2)$

$i = 1:a \qquad j = 1:n_i$

"Balanced" model $\quad n_i = n$ for all i

1 factor with a levels

Can we write this model as

$y = X\beta + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I)$?

$$y = \begin{pmatrix} y_{1,1} \\ \vdots \\ y_{1,n_1} \\ y_{2,1} \\ \vdots \\ y_{2,n_2} \\ \vdots \\ y_{a,1} \\ \vdots \\ y_{a,n_a} \end{pmatrix} \qquad \underset{\sim}{\beta} = \begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_a \end{pmatrix}$$

$$\varepsilon = \begin{pmatrix} \varepsilon_{1,1} \\ \vdots \\ \varepsilon_{1,n_1} \\ \vdots \\ \varepsilon_{a,1} \\ \vdots \\ \varepsilon_{a,n_a} \end{pmatrix}$$

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & & & 0 \\ & & \vdots & & & \vdots \\ 1 & 1 & 0 & & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \vdots & \vdots & & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$$

$$\begin{pmatrix} \vdots & & & & & \\ 1 & 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

X is <u>not</u> full rank

what can we do?

we can add restrictions

or look at different parameterizations of the model

a) consider this paramet.

$$y_{i,j} = \mu_i + \varepsilon_{i,j} \qquad \varepsilon_{i,j} \overset{iid}{\sim} N(0,\sigma^2)$$

$i = 1:a$

$j = 1:n_i$

$\mu_1, \ldots, \mu_a$

$$\beta = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_a \end{pmatrix}$$

The resulting X is full

rank

$$\mathbb{X} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & & \vdots \\ 0 & 1 & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 1 & & \vdots \\ \vdots & \vdots & & 0 \\ \vdots & \vdots & & 1 \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$\hat{\beta} = (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'y$$

$$\underset{a \times a}{\mathbb{X}'\mathbb{X}} = \begin{pmatrix} n_1 & 0 & \cdots & 0 \\ 0 & n_2 & & \\ \vdots & & \ddots & \\ 0 & \cdots & & n_a \end{pmatrix}$$

$$(\mathbb{X}'\mathbb{X})^{-1} = \begin{pmatrix} 1/n_1 & 0 & \cdots & 0 \\ 0 & 1/n_2 & & \\ \vdots & & \ddots & \\ 0 & & & 1/n_a \end{pmatrix}$$

$$X'y = \begin{pmatrix} \sum_{j=1}^{n_1} y_{1j} \\ \vdots \\ \sum_{j=1}^{n_a} y_{aj} \end{pmatrix}$$

$$\hat{\beta} = (\cancel{X'X})^{-1} X' y$$

$$= \begin{pmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{pmatrix}$$

$$\hat{\mu}_i = \bar{y}_{i\cdot} =$$

$$= \sum_{j=1}^{n_i} y_{i,j} / n_i$$

Hypothesis testing

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$

vs

H_1 : At least one μ_i is different from the rest

F-test

If we fail to reject then

$y_{i,j} = \mu + \varepsilon_{i,j}$ $\quad \varepsilon_{i,j} \overset{iid}{\sim} N(0, \sigma^2)$

$\hat{\mu} = \bar{y}_{\cdot\cdot}$

b) Add one restriction to the model

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$$

Possible restrictions:

i) $\alpha_{i_0} = 0$ for one i_0

ii) $\sum_{i=1}^{a} \alpha_i = 0$