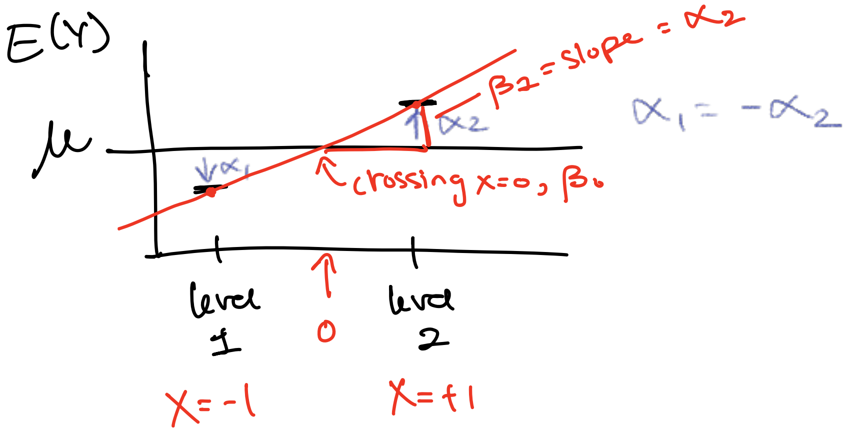
Rep : Anova, 1 factor, 2 levels

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0, \sigma^2) \qquad \sum \alpha_j = 0$$



with only 2 levels effect coding is very simple, and easy to interpret

$$y = \beta_0 + \beta_1 X + \varepsilon$$

$X = -1$ for level 1 (low level)

$X = +1$ for level 2 (high level)

$\beta_0 =$ global mean

$\beta_1 = \alpha_2$

$$\underbrace{2\beta_1}_{\uparrow \text{"main effect"}} = \text{effect of going from level 1 (low level) to level 2 (high level)}$$

Aim : Estimate main effects of several 2-level factors, (and also interaction effects).

Summing up:

- k factors, 2 levels each. Aim: estimate and test factor effects on some measurable response y.
- For each factor, levels are coded as
  -1 (low) and +1 (high)
- 2^k factor combinations
- model (no interactions)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i$$

$\varepsilon_i \sim N(0, \sigma^2)$

$2\beta_j$ = effect on $E(Y)$ when x_j changes from low to high level (other factors held constant)

- Design matrix, one experimental run

$$X \text{ has dim } \underset{\#\text{factor combinations}}{2^k} \times \underset{\text{intercept} + k \text{ factors}}{(k+1)}$$

Example $k=3$, $2^3 = 8$

$$X = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & +1 & -1 \\ 1 & -1 & +1 & +1 \\ 1 & +1 & -1 & -1 \\ 1 & +1 & -1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & +1 & +1 \end{bmatrix} \begin{array}{l} \leftarrow \text{all factors low level} \\ \\ \\ \\ \\ \\ \\ \leftarrow \text{all factors high level.} \end{array}$$

- One experimental run: observe y for each row of X in random order.

- For one (or with repetitions) experimental run:

  columns of $\underline{X}$ are orthogonal

  $$(\underline{X}^T\underline{X})^{-1} = \frac{1}{n} I$$

  $$\hat{\beta} = (\underline{X}^T X)^{-1} X^T Y \sim MVN\left(\beta, \frac{\sigma^2}{n} I\right)$$ equal var, independent estimators

  $\uparrow$ Variance reduced with n, e.g. add replicates.

- $\hat{\beta}_0 = \bar{y}$

  $$\hat{\beta}_1 = \frac{1}{n}\left(\sum_{i:x_{i1}=+1} y_i - \sum_{i:x_{i1}=-1} y_i\right)$$

  $$= \frac{1}{2}\left(\underset{x_1 \text{ high level}}{\text{mean } y} - \underset{x_1 \text{ low level}}{\text{mean } y}\right)$$

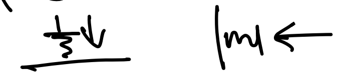
- Main effect factor A (x_1): $\hat{A} = 2\hat{\beta}_1$ etc.

  Testing main effects: $H_0: \beta_j = 0$ vs $H_1: \beta_j \neq 0$

  $j = 1, \dots, k$.

Project examples:

Ex: 1) circumference cookies $\left(\begin{array}{l}\text{sugar (amount)} \\ \text{butter/margarine} \\ \text{baking soda/powder}\end{array}\right)$

2) Drill time (type of material
type of screw
Angle )

SSE, SSR in 2^k DOE

Recall lecture 17, orthogonal columns of X:

$$SS_R = \hat{\beta}_1^2 \underbrace{\sum_i x_{1i}^2}_{\text{here} = n} + \hat{\beta}_2^2 \sum_i x_{2i}^2 + \dots + \hat{\beta}_k^2 \sum_i x_{ki}^2$$

$$= \left(\hat{\beta}_1^2 + \hat{\beta}_2^2 + \dots \hat{\beta}_k^2\right) n$$

$n\hat{\beta}_j^2$ is the variation explained by factor j.

$$\Rightarrow SS_E = SS_T - SS_R = \Sigma_i (y_i - \bar{y})^2 - n \sum_{i=1}^{k} \hat{\beta}_j^2 .$$

Note: conducted 2^2 factorial experiment in class, see R code

---