Lecture 11:

Oct. 18

Analysis of Variance

Recall: Two independent sample t-test

Pop 1: $N(\mu_1, \sigma^2) \Rightarrow$ Random Sample $Y_1^1, \dots, Y_n^1 \overset{iid}{\sim} N(\mu_1, \sigma^2)$

Pop 2: $N(\mu_2, \sigma^2) \Rightarrow$ Random Sample $Y_1^2, \dots, Y_n^2 \overset{iid}{\sim} N(\mu_2, \sigma^2)$

Let $\delta = \mu_2 - \mu_1$ and its estimate $\hat{\delta} = \bar{Y}_2 - \bar{Y}_1$.

Observe that $\underset{\substack{\uparrow \\ \text{linear combination of Normals}}}{\hat{\delta} \sim} N\left(E(\hat{\delta}) = \mu_2 - \mu_1 = \delta,\ \underset{\substack{\uparrow \\ \text{independent samples}}}{Var(\hat{\delta}) =} \frac{\sigma^2}{n_2} + \frac{\sigma^2}{n_1} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$

Thus,

$$\frac{\hat{\delta} - \delta}{\sqrt{\underset{\substack{\uparrow \\ \text{unknown, so estimate}}}{\sigma^2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

$$S_p^2 = \underbrace{\boxed{\frac{(n_1-1)}{n_1+n_2-2}}\, S_1^2 + \boxed{\frac{(n_2-1)}{n_1+n_2-2}}}_{\text{weights sum to 1.}}\, S_2^2$$

$$\therefore \frac{(n_1+n_2-2)\, S_p^2}{\sigma^2} = \frac{\sum_{i=1}^{n_1} (Y_i^1 - \bar{Y})^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (Y_i^2 - \bar{Y})^2}{\sigma^2} \sim \chi^2_{(n_1+n_2-2)}$$

So,

$$\frac{\hat{\delta} - \delta}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{(n_1+n_2-2)}$$

And under Ho,

$$T = \frac{\hat{\delta}}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{Ho}{\sim} t_{(n_1+n_2-2)}$$

Decision Rule:

Reject Ho if

$$|T| > t_{(0.975,\, n_1+n_2-2)}$$

Linear Regression Framework

$$\left.\begin{array}{ll} Y_i^1 = \mu_1 + \epsilon_i & \epsilon_i \overset{iid}{\sim} N(0,\sigma^2) \\ Y_j^2 = \mu_2 + \epsilon_j & \epsilon_j \overset{iid}{\sim} N(0,\sigma^2) \end{array}\right\} \text{ Model Assumption}$$

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon}$$

$$\underline{\beta} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

$$\begin{bmatrix} Y_1^1 \\ \vdots \\ Y_{n_1}^1 \\ \hdashline Y_{n_1+1}^2 \\ \vdots \\ Y_{n_1+n_2}^2 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_1 \\ \hdashline \mu_2 \\ \vdots \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n_1} \\ \hdashline \epsilon_{n_1+1} \\ \vdots \\ \epsilon_{n_1+n_2} \end{bmatrix} \quad \text{where } \underline{\epsilon} \sim N(0, I_n \otimes \sigma^2)$$

Note:

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_1 \\ \hdashline \mu_2 \\ \vdots \\ \mu_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbb{1} & 0 \\ \hdashline 0 & \mathbb{1} \end{bmatrix}}_{X} \underbrace{\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}}_{\underline{\beta}}$$

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon}$$

And by construction of X (X is full column rank – actually orthogonal), $(X^TX)^{-1}$ exists.

$$\therefore \ \hat{\underline{\beta}} = (X^TX)^{-1}X^TY = \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{bmatrix}$$

$$\delta = \mu_2 - \mu_1 = \underbrace{(-1 \ \ 1)}_{\underline{c}} \underline{\beta}$$

$$\hat{\delta} = \bar{Y}_2 - \bar{Y}_1 = \underline{c}\,\hat{\underline{\beta}},$$

$EY = X\beta$

$$\hat{\underline{\beta}} = \underbrace{(X^TX)^{-1}X^T}_{\text{Constant}}Y \sim \mathcal{N}\big(E\hat{\underline{\beta}} = (X^TX)^{-1}X^T EY = \beta,\ Var(\hat{\beta}) = [(X^TX)^{-1}X^T]\,cov(Y)\,[X(X^TX)^{-1}] = (X^TX)^{-1}\otimes\sigma^2\big)$$

utilized $Var(aY) = a^2Y$

$Cov(AY) = A\,Cov(Y)\,A^T$

$$\text{Thus, } \hat{\delta} = \underline{c}\,\hat{\underline{\beta}} \sim \mathcal{N}\Big(E\hat{\delta} = \underline{c}\underline{\beta} = \delta,\ var(\hat{\delta}) = \underline{c}\,cov(\hat{\beta})\,\underline{c}^T = (-1\ \ 1)(X^TX)^{-1}\otimes\sigma^2\begin{pmatrix}-1\\1\end{pmatrix}\Big)$$

$$\vdots$$

$$= \sigma^2\left(\tfrac{1}{n_1} + \tfrac{1}{n_2}\right)$$

$$(X^TX)^{-1} = \begin{bmatrix} \frac{1}{n_1} & 0 \\ 0 & \frac{1}{n_2} \end{bmatrix}$$

$$\therefore \ \hat{\delta} \sim \mathcal{N}\left(\delta, \sigma^2\left(\tfrac{1}{n_1} + \tfrac{1}{n_2}\right)\right)$$

So,

$$\hat{\underline{Y}} = \begin{bmatrix} \hat{Y}_1 \\ \vdots \\ \hat{Y}_{n_1} \\ \hline \hat{Y}_{n_1+1} \\ \vdots \\ \hat{Y}_{n_1+n_2} \end{bmatrix} = X\hat{\underline{\beta}} = \begin{bmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{bmatrix}_{(n_1+n_2)\times 2} \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{pmatrix} = \begin{bmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_1 \\ \hline \bar{Y}_2 \\ \vdots \\ \bar{Y}_2 \end{bmatrix}$$

Our Model: $\underline{Y} = X\underline{\beta} + \epsilon$, where $X\underline{\beta} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_2 \end{bmatrix}$

$$\hat{\underline{Y}} = X\hat{\beta} = \begin{bmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_1 \\ \bar{Y}_2 \\ \vdots \\ \bar{Y}_2 \end{bmatrix}$$

$$\underline{R} = \underline{Y} - \hat{\underline{Y}} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_1 \\ Y_2 \\ \vdots \\ Y_2 \end{bmatrix} - \begin{bmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_1 \\ \bar{Y}_2 \\ \vdots \\ \bar{Y}_2 \end{bmatrix} = \begin{bmatrix} Y_1 - \bar{Y}_1 \\ \vdots \\ Y_1 - \bar{Y}_1 \\ Y_2 - \bar{Y}_2 \\ \vdots \\ Y_2 - \bar{Y}_2 \end{bmatrix}$$

Note:

$$\underline{\epsilon} = \underline{Y} - X\underline{\beta}$$

$$E\|\underline{\epsilon}\|_2 = E\sum_{k=1}^{n_1+n_2} \epsilon_k^2 = (n_1+n_2)\sigma^2$$

$$\text{So, } E\left\{\frac{\|\epsilon\|_2}{n_1+n_2}\right\} = \sigma^2$$

The data analog is:

$$\hat{\sigma}^2 = \frac{\|R\|_2}{n_1+n_2-2} = \frac{\sum_{i=1}^{n_1}(Y_i^1-\bar{Y}_1)^2 + \sum_{j=1}^{n_2}(Y_j^2-\bar{Y}_2)^2}{n_1+n_2-2}$$

$$= \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}, \text{ where } (n_1-1)S_1^2 = \sum_{i=1}^{n_1}(Y_i^1-\bar{Y}_1)^2$$

$$(n_2-1)S_2^2 = \sum_{j=1}^{n_2}(Y_j^2-\bar{Y}_2)^2$$

Stat 200C

ANOVA

An approach for comparing nested models

General Model

FULL $\quad \mathcal{M}_1 : \underline{\beta} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \in \mathbb{R}^2 = p1$

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon}$$

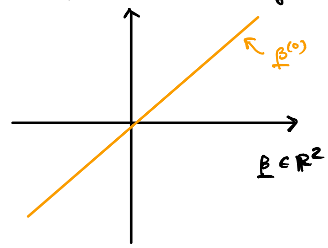
$$X = \begin{bmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{bmatrix}, \quad \underline{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n_1+n_2} \end{bmatrix}$$

Note:

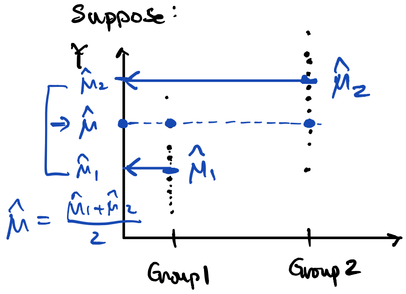
$H_0 : \delta = 0 \Leftrightarrow \mu_1 = \mu_2 := \mu$

Reduced Model $\mathcal{M}_0$ $\quad \underset{\text{Under } H_0}{\underline{\beta}^{(0)}} = \begin{pmatrix} \mu \\ \mu \end{pmatrix} = p^{0} \subset p^{1}$

where we have the following subspace



Suppose:



We consider Two Models:

$\mathcal{M}_1 : \hat{\mu}_1, \hat{\mu}_2$

$\mathcal{M}_0 : \hat{\mu}$

In this case, $\mathcal{M}_1$ seems to have a smaller overall variance that model $\mathcal{M}_0$.

Under $\mathcal{M}_1$:

$$\hat{\underline{\beta}} = \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{pmatrix}$$

$$\hat{\underline{Y}} = X\hat{\beta} = \begin{pmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_1 \\ \bar{Y}_2 \\ \vdots \\ \bar{Y}_2 \end{pmatrix}$$

$$SSE = \| Y - \hat{Y} \|_2$$

$$= \sum_{i=1}^{n_1} (Y_i^1 - \bar{Y}_1)^2 + \sum_{j=1}^{n_2} (Y_j^2 - \bar{Y}_2)^2$$

Under $\mathcal{M}_0$:

$$\underline{Y} = X\underline{\beta}^{(0)} + \underline{\epsilon}$$

$$= \left(\begin{array}{c|c} \mathbb{1} & 0 \\ \hline 0 & \mathbb{1} \end{array} \right) \begin{pmatrix} \mu \\ \mu \end{pmatrix} + \underline{\epsilon}$$

$$= \begin{pmatrix} \mathbb{1}\mu \\ \mathbb{1}\mu \end{pmatrix} + \underline{\epsilon}$$

$$= \underbrace{\mathbb{1}}_{X^0} \underbrace{\mu}_{\beta^0} + \underline{\epsilon}$$

Note: $\hat{\beta}^0 = (X^{0T} X^0)^{-1} X^0 y = \frac{1}{n_1+n_2} \sum_{i=1}^{n_1+n_2} Y_i = \bar{Y}$

$$= \frac{1}{n_1+n_2} \left[\sum_{i=1}^{n_1} Y_i + \sum_{i=n_1+1}^{n_1+n_2} Y_i \right]$$

$$= \frac{n_1}{n_1+n_2} \bar{Y}_1 + \frac{n_2}{n_1+n_2} \bar{Y}_2$$

$$\therefore \; \hat{Y}^0 = X^0 \hat{\beta}^0 = \begin{pmatrix} \bar{Y} \\ \vdots \\ \bar{Y} \end{pmatrix}$$

$$\begin{aligned} SSE^0 &= \| Y - \hat{Y}^0 \|_2 \\ &= \sum_{i=1}^{n_1} (Y_i^1 - \bar{Y})^2 + \sum_{j=1}^{n_2} (Y_j^2 - \bar{Y})^2 \end{aligned}$$

<u>Summary:</u>

M_1 (Full)

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon} \qquad \underline{\beta} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \qquad SSE^1 = \sum_{i=1}^{n_1} (Y_i^1 - \bar{Y}_1)^2 + \sum_{j=1}^{n_2} (Y_j^2 - \bar{Y}_2)^2 \qquad dF = n_1 + n_2 - 2$$

M_0 (Reduced)

$$\underline{Y} = X^0 \underline{\beta}^0 + \underline{\epsilon} \qquad X^0 = \mathbf{1}, \; \underline{\beta}^0 = \mu \qquad SSE^0 = \sum_{i=1}^{n_1} (Y_i^1 - \bar{Y})^2 + \sum_{j=1}^{n_2} (Y_j^2 - \bar{Y})^2 \qquad dF = n_1 + n_2 - 1$$

---

* $SSE^0 \geq SSE^1$ (Reduced Model will be at least as large as the Full Model)

Compare M_1 and M_0 via $SSE(1)$ and $SSE(0)$ by examining:

$$F = \frac{\dfrac{SSE(0) - SSE(1)}{dF_0 - dF_1}}{\dfrac{SSE(1)}{dF_1}} \overset{H_0}{\sim} F(dF = \text{numerator}, \; dF = \text{denominator})$$

F is the test statistic

Decision: Reject H_0 if F is large.

↳ Chose Full Model