Lecture Nov 15

$Y_i^1 = U_{1i}^1 - U_{0i}^1$ diff in scores for subject i from group 1

$Y_i^2 = U_{1i}^2 - U_{0i}^1$ ———————— group 2

$Y_1^1, \ldots, Y_n^1$ difference scores of the n-subjects from group 1

$\sim$ iid $N(\mu_1, \sigma^2)$

$Y_1^2, \ldots, Y_n^2$ ———————— group 2

$\sim$ iid $N(\mu_2, \sigma^2)$

$\delta = \mu_2 - \mu_1$

parameter

$\delta \neq \overbrace{U_{1i} - U_{0i}}^{\text{data}}$

t-test ...

Full Model $\quad Y_i = \mu_i + \varepsilon_i \;,\; \varepsilon_i \text{ iid } N(0,\sigma^2)$

$$\mu_i = \begin{cases} \mu_{(1)} , & \text{if } i = 1 \dots, n \\ \mu_{(2)} , & \text{if } i = n+1, \dots, 2n \end{cases}$$

In matrix notation:

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \\ \hline Y_{n+1} \\ \vdots \\ Y_{n+n} \end{pmatrix} = \begin{pmatrix} \mathbb{1} & \underline{0} \\ \hline \underline{0} & \mathbb{1} \end{pmatrix} \begin{pmatrix} \mu_{(1)} \\ \mu_{(2)} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \\ \hline \varepsilon_{n+1} \\ \vdots \\ \varepsilon_{2n} \end{pmatrix}$$

$$\hat{\mu}_{(1)} = \bar{Y}_1 = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

$$\hat{\mu}_{(2)} = \bar{Y}_2 = \frac{1}{n} \sum_{i=n+1}^{2n} Y_i$$

$$SSE(1) = \| \underline{R}^{(1)} \|^2 = \Big\| \underline{Y} - \underbrace{\begin{pmatrix} \mathbb{1}\, 0 \\ \hline 0\, \mathbb{1} \end{pmatrix} \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{pmatrix}}_{\begin{pmatrix} \mathbb{1} \cdot \bar{Y}_1 \\ \hline \mathbb{1}\, \bar{Y}_2 \end{pmatrix}} \Big\|^2$$

$$SSE(1) = \sum_{i=1}^{n} (Y_i - \bar{Y}_1)^2 + \sum_{i=n+1}^{2n} (Y_i - \bar{Y}_2)^2$$

$$df_1 = 2n - 2$$

<u>Reduced Model</u> $H_0: \delta = 0 \iff \mu_{(1)} = \mu_{(2)} := \mu$

$$\underset{\sim}{Y} = \mathbb{1}\mu + \underset{\sim}{\varepsilon}$$

$$\hat{\mu} = \bar{Y} = \frac{1}{2n} \sum_{i=1}^{2n} Y_i$$

$$= \frac{n}{2n}\bar{Y}_1 + \frac{n}{2n}\bar{Y}_2$$

$$= \frac{1}{2}\bar{Y}_1 + \frac{1}{2}\bar{Y}_2$$

$$SSE(0) = \| \underset{\sim}{R}^{(0)} \|^2$$

$$= \sum_{i=1}^{2n} (Y_i - \bar{Y})^2$$

$$df0 = 2n - 1$$

$$F = \frac{\frac{SSE(0) - SSE(1)}{1}}{\frac{SSE(1)}{df1}}$$

Reject H_0 if $F_{stat} > F(.95; 1, 2n-2)$

Demonstrate that this is equivalent to the t-test!

$$SSE(0) = \sum_{i=1}^{2n} (Y_i - \bar{Y})^2$$

$$= \underbrace{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}_{(*)} + \underbrace{\sum_{i=n+1}^{2n} (Y_i - \bar{Y})^2}_{(**)}$$

$$(*) = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

$$= \sum_{i=1}^{n} \left(Y_i - \frac{1}{2}(\bar{Y}_1 + \bar{Y}_2) \right)^2$$

$$= \sum_{i=1}^{n} \left(Y_i - \underbrace{\frac{1}{2}\bar{Y}_1} - \frac{1}{2}\bar{Y}_2 - \underbrace{\frac{1}{2}\bar{Y}_1} + \frac{1}{2}\bar{Y}_1 \right)^2$$

$$= \sum_{i=1}^{n} \left[(Y_i - \bar{Y}_1) + \frac{1}{2}(\bar{Y}_1 - \bar{Y}_2) \right]^2$$

$$= \sum_{i=1}^{n} (Y_i - \bar{Y}_1)^2 + \sum_{i=1}^{n} \left(\frac{1}{2}(\bar{Y}_1 - \bar{Y}_2) \right)^2$$

$$+ \; 2\left(\frac{1}{2}\right) \underline{\sum_{i=1}^{n} (Y_i - \bar{Y}_1)} (\bar{Y}_1 - \bar{Y}_2)$$

$$= \sum_{i=1}^{n} (Y_i - \bar{Y}_1)^2 + \frac{n}{4}(\bar{Y}_1 - \bar{Y}_2)^2$$

Similarly for $(**)$

$$\sum_{i=n+1}^{2n} (Y_i - \bar{Y})^2 = \sum_{i=n+1}^{2n} (Y_i - \bar{Y}_2)^2 + \frac{n}{4}(\bar{Y}_1 - \bar{Y}_2)^2$$

$\therefore$ SSE(0) − SSE(1)

$$= \sum_{i=1}^{n} (Y_i - \bar{Y}_1)^2 + \sum_{i=n+1}^{2n} (Y_i - \bar{Y}_2)^2 + \frac{n}{2} (\bar{Y}_1 - \bar{Y}_2)^2$$

$$- \left(\sum_{i=1}^{n} (Y_i - \bar{Y}_1)^2 + \sum_{i=n+1}^{2n} (Y_i - \bar{Y}_2)^2 \right)$$

$$= \frac{n}{2} (\bar{Y}_1 - \bar{Y}_2)^2$$

Next:

$$\frac{SSE(1)}{2n-2} = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y}_1)^2 + \sum_{i=n+1}^{2n} (Y_i - \bar{Y}_2)^2}{2n-2}$$

$$= \frac{n-1}{2n-2} \cdot S_1^2 + \frac{n-1}{2n-2} \cdot S_2^2$$

$$= S_p^2$$

$$\Rightarrow F = \frac{\frac{n}{2}(\bar{Y}_1 - \bar{Y}_2)^2}{S_p^2}$$

$$= \left(\frac{(\bar{Y}_1 - \bar{Y}_2)}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{n}\right)}} \right)^2 = \left(t_{STAT} \right)^2$$