STA207 Discussion 3

Jing Lyu

January 2023

One-way ANOVA model:

$$Y_{ij} = \mu_i + \epsilon_{ij}; i = 1, ..., k, j = 1, ..., n_k$$

or

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}; i = 1, ..., k, j = 1, ..., n_k$$

where $\epsilon_{ij} \sim_{iid} N(0, \sigma^2)$. Suppose we have constraints $\sum_{i=1}^k n_i \alpha_i = 0$. $N = \sum_{i=1}^k n_i$. Then LS estimation gives

$$\hat{\mu_i} = \bar{Y}_{i.}, \hat{\mu} = \bar{Y}_{..}, \hat{\alpha_i} = \hat{\mu_i} - \hat{\mu} = \bar{Y}_{i.} - \bar{Y}_{..}, \hat{\sigma}^2 = \frac{SSE}{N-k}$$

1 Properties

(1) **Proof: $\hat{\sigma}^2$ is an unbiased estimator of σ^2**

$$SSE = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i.})^2$$

We just need to calculate $E\left[(Y_{ij} - \bar{Y}_{i.})^2\right]$:

$$\begin{aligned}
E\left[(Y_{ij} - \bar{Y}_{i.})^2\right] = Var(Y_{ij} - \bar{Y}_{i.}) &= Var\left((1 - \frac{1}{n_i})Y_{ij} - \frac{1}{n_i}\sum_{l \neq j} Y_{il}\right) \\
&= (1 - \frac{1}{n_i})^2 Var(Y_{ij}) + \frac{1}{n_i^2}\sum_{l \neq j} Var(Y_{il}) \\
&= (1 - \frac{1}{n_i})^2 \sigma^2 + \frac{1}{n_i^2}(n_i - 1)\sigma^2 \\
&= (1 - \frac{1}{n_i})\sigma^2
\end{aligned}$$

Then we have

$$E[SSE] = \sum_{i=1}^k \sum_{j=1}^{n_i} (1 - \frac{1}{n_i})\sigma^2 = (N-k)\sigma^2$$

(2) **Proof: $E[MSTR] = \sigma^2 \sum_{i=1}^k n_i(\mu_i - \mu)^2/(k-1)$ where $\mu = \frac{\sum_{i=1}^k n_i \mu_i}{N}$**

$$MSTR = \frac{\sum_{i=1}^k n_i(\bar{Y}_{i.} - \bar{Y}_{..})^2}{k-1}$$

We just need to calculate $E\left[(\bar{Y}_{i.} - \bar{Y}_{..})^2\right]$. Notice that $E\left[(\bar{Y}_{i.} - \bar{Y}_{..})^2\right] = Var(\bar{Y}_{i.} - \bar{Y}_{..}) + \left[E(\bar{Y}_{i.} - \bar{Y}_{..})\right]^2$ where

$E(\bar{Y}_{i.} - \bar{Y}_{..}) = \mu_i - \mu$. And

$$\begin{aligned} Var(\bar{Y}_{i.} - \bar{Y}_{..}) &= Var\left(\bar{Y}_{i.} - \frac{\sum_{s=1}^{k} n_s \bar{Y}_{s.}}{N}\right) = Var\left((1-\frac{n_i}{N})\bar{Y}_{i.} - \frac{\sum_{s\neq i} n_s \bar{Y}_{s.}}{N}\right) \\ &= (1-\frac{n_i}{N})^2 Var(\bar{Y}_{i.}) + \sum_{s\neq i} \frac{n_s^2}{N^2} Var(\bar{Y}_{s.}) \\ &= (1-\frac{n_i}{N})^2 \frac{\sigma^2}{n_i} + \sum_{s\neq i} \frac{n_s^2}{N^2}\frac{\sigma^2}{n_s} \\ &= (\frac{1}{n_i} - \frac{1}{N})\sigma^2 \end{aligned}$$

Then we have

$$E\left[SSTR\right] = E\left[\sum_{i=1}^{k} n_i(\bar{Y}_{i.} - \bar{Y}_{..})^2\right] = \sum_{i=1}^{k} n_i\left[(\frac{1}{n_i} - \frac{1}{N})\sigma^2 + (\mu_i - \mu)^2\right] = (k-1)\sigma^2 + \sum_{i=1}^{k} n_i(\mu_i - \mu)^2$$

(3) **Moment properties of the estimation**

$$E(\hat{\mu}) = \mu, Var(\hat{\mu}) = \frac{\sigma^2}{N}$$

$$E(\hat{\mu}_i) = \mu_i, Var(\hat{\mu}_i) = \frac{\sigma^2}{n_i}$$

$$E(\hat{\alpha}_i) = E(\hat{\mu}_i - \hat{\mu}) = \mu_i - \mu = \alpha_i, Var(\hat{\alpha}_i) = Var(\hat{\mu}_i - \hat{\mu}) = Var(\bar{Y}_{i.} - \bar{Y}_{..}) = (\frac{1}{n_i} - \frac{1}{N})\sigma^2$$