STA 610 Homework 1

Yuren Zhou

September 14, 2024

Question 1.

a. (1 point) Since $y_{1,A}, \dots, y_{n_A,A} \overset{iid}{\sim} N(\theta_A, \sigma^2)$, we know $\bar{y}_A$ also follows a normal distribution, with

$$\mathbb{E}[\bar{y}_A] = \frac{1}{n_A}\sum_{i=1}^{n_A}\mathbb{E}[y_{i,A}] = \theta_A, \quad \mathbb{V}[\bar{y}_A] = \frac{1}{n_A^2}\sum_{i=1}^{n_A}\mathbb{V}[y_{i,A}] = \frac{\sigma^2}{n_A}.$$

Similarly, $\bar{y}_B \sim N(\theta_B, \frac{\sigma^2}{n_B})$.

b. (1 point) Since $\bar{y}_A, \bar{y}_B$ follow independent normal distributions, we know

$$Z := \frac{\bar{y}_A - \bar{y}_B}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}$$

also follows a normal distribution, with

$$\mathbb{E}[Z] = \frac{\mathbb{E}[\bar{y}_A] - \mathbb{E}[\bar{y}_B]}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} = \frac{\theta_A - \theta_B}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}, \quad \mathbb{V}[Z] = \frac{\mathbb{V}[\bar{y}_A] + \mathbb{V}[\bar{y}_B]}{\sigma^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)} = 1.$$

c. (1 point) For testing $H : \theta_A = \theta_B$ at significance level α, we compute test statistic Z and use confidence interval $(z_{\alpha/2}, z_{1-\alpha/2})$ for z_α denoting the α-quantile of standard normal distribution.

d. (2 points) Since sample mean is an unbiased estimator of population mean, we have

$$\begin{aligned}\mathbb{E}\left[\frac{SSW}{n_A + n_B - 2}\right] &= \frac{n_A - 1}{n_A + n_B - 2} \cdot \frac{\sum_{i=1}^{n_A}(y_{i,A} - \bar{y}_A)^2}{n_A - 1} + \frac{n_B - 1}{n_A + n_B - 2} \cdot \frac{\sum_{i=1}^{n_B}(y_{i,B} - \bar{y}_B)^2}{n_B - 1} \\ &= \frac{n_A - 1}{n_A + n_B - 2} \cdot \sigma^2 + \frac{n_B - 1}{n_A + n_B - 2} \cdot \sigma^2 = \sigma^2.\end{aligned}$$

We further notice that $\frac{SSW}{n_A+n_B-2}$ follows $\chi^2(n_A + n_B - 2)$ distribution. Hence to test $H : \theta_A = \theta_B$ at significance level α with σ^2 unknown, we compute test statistic

$$T := \frac{\bar{y}_A - \bar{y}_B}{\sqrt{\frac{SSW}{n_A+n_B-2}} \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}$$

and use interval $(t_{\alpha/2}, t_{1-\alpha/2})$ for t_α denoting the α-quantile of student t distribution with degree of freedom $n_A + n_B - 2$.

Question 2.

a. (1 point)

$$\text{Var}[y_{i,j}|\mu,\tau^2,\sigma^2] = \text{Var}[\mu+\alpha_j+\epsilon_{i,j}|\mu,\tau^2,\sigma^2] = \text{Var}[\alpha_j|\tau^2] + \text{Var}[\epsilon_{i,j}^2|\sigma^2] = \tau^2+\sigma^2.$$

b. (2 points)

$$\text{Cov}[y_{i,j}, y_{i',j}|\mu,\tau^2,\sigma^2] = \text{Cov}[\mu+\alpha_j+\epsilon_{i,j}, \mu+\alpha_j+\epsilon_{i',j}|\mu,\tau^2,\sigma^2] = \text{Cov}[\alpha_j,\alpha_j|\tau^2] = \tau^2.$$

c. (2 points)

$$\text{Cov}[y_{i,j}, y_{i',j'}|\mu,\tau^2,\sigma^2] = \text{Cov}[\mu+\alpha_j+\epsilon_{i,j}, \mu+\alpha_{j'}+\epsilon_{i',j'}|\mu,\tau^2,\sigma^2] = 0.$$