Name: Yuxuan Zhang

Date: September 24, 2024

HW3

1. (a) By the Bayes rule, we know that

$$p(\theta_j|\bar{y}_j) = \frac{p(\bar{y}_j|\theta_j)p(\theta_j)}{p(\bar{y}_j)}.$$

We also know that the prior distribution is

$$p(\theta_j) = \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{(\theta_j - \mu)^2}{2\tau^2}\right),$$

and the likelihood is given by the normal distribution

$$p(\bar{y}_j|\theta_j) = \frac{1}{\sqrt{2\pi\frac{\sigma^2}{n_j}}} \exp\left(-\frac{(\bar{y}_j - \theta_j)^2}{2\frac{\sigma^2}{n_j}}\right).$$

Then, the conditional distribution is

$$\begin{aligned} p(\theta_j|\bar{y}_j) &\propto p(\theta_j)p(\bar{y}_j|\theta_j) \\ &\propto \exp\left(-\frac{(\theta_j - \mu)^2}{2\tau^2} - \frac{(\bar{y}_j - \theta_j)^2}{2\frac{\sigma^2}{n_j}}\right) \\ &= \exp\left(-\frac{1}{2}\left[\frac{n_j}{\sigma^2}(\theta_j - \bar{y}_j)^2 + \frac{1}{\tau^2}(\theta_j - \mu)^2\right]\right). \end{aligned}$$

We then complete the squares:

$$\begin{aligned} \frac{\sigma^2}{n_j}(\theta_j - \mu)^2 + \frac{1}{\tau^2}(\bar{y}_j - \theta_j)^2 &= \frac{n_j}{\sigma^2}(\theta_j^2 - 2\bar{y}_j\theta_j + \bar{y}_j^2) + \frac{1}{\tau^2}(\theta_j^2 - 2\theta_j\mu + \mu^2) \\ &= \left(\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}\right)\theta_j^2 - 2\theta_j(\frac{n_j\bar{y}_j}{\sigma^2} + \frac{\mu}{\tau^2}) + \text{constant} \\ &= \left(\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}\right)\left(\theta_j - \frac{\frac{n_j}{\sigma^2}\bar{y}_j + \frac{1}{\tau^2}\mu}{\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}}\right)^2 + \text{constant}. \end{aligned}$$

Therefore,

$$p(\theta_j|\bar{y}_j) \sim \text{Normal}\left(\frac{\frac{n_j}{\sigma^2}\bar{y}_j + \frac{1}{\tau^2}\mu}{\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}}, \frac{1}{\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}}\right).$$

(b) From part (a), we know that $\mathbb{E}[\theta_j|\bar{y}_j] = \frac{\frac{n_j}{\sigma^2}\bar{y}_j + \frac{1}{\tau^2}\mu}{\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}}$, and $SD[\theta_j|\bar{y}_j] = \sqrt{\frac{1}{\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}}}$. Let $Z = \frac{\theta_j - \mathbb{E}[\theta_j|\bar{y}_j]}{SD[\theta_j|\bar{y}_j]}$. We know that $Z \sim \text{Normal}(0, 1)$. Substitute this into the probability we are solving for, we get

$$\begin{aligned} Pr(\theta_j \in \mathbb{E}[\theta_j|\bar{y}_j] \pm z_{1-\alpha/2}SD[\theta_j|\bar{y}_j]) &= Pr(\mathbb{E}[\theta_j|\bar{y}_j] - z_{1-\alpha/2}SD[\theta_j|\bar{y}_j] \leq \theta_j \leq \mathbb{E}[\theta_j|\bar{y}_j] + z_{1-\alpha/2}SD[\theta_j|\bar{y}_j]) \\ &= Pr(-z_{1-\alpha/2}SD[\theta_j|\bar{y}_j] \leq \theta_j - \mathbb{E}[\theta_j|\bar{y}_j] \leq z_{1-\alpha/2}SD[\theta_j|\bar{y}_j]) \\ &= Pr(-z_{1-\alpha/2} \leq \frac{\theta_j - \mathbb{E}[\theta_j|\bar{y}_j]}{SD[\theta_j|\bar{y}_j]} \leq z_{1-\alpha/2}) \\ &= Pr(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}) \\ &= 1 - \alpha. \end{aligned}$$

(c) The width of the usual z-interval is

$$\bar{y}_j \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n_j}}.$$

The width of the interval from part (b), which is based on the posterior distribution of θ_j, is:

$$\mathbb{E}[\theta_j|\bar{y}_j] \pm z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma^2\tau^2}{n_j\tau^2 + \sigma^2}}.$$

Notice that $\frac{\sigma}{\sqrt{n_j}} = \sqrt{\frac{\sigma^2(\tau^2+1/n_j\sigma^2)}{n_j(\tau^2+1/n_j\sigma^2)}}$, and $\sqrt{\frac{\sigma^2\tau^2}{n_j\tau^2+\sigma^2}} = \sqrt{\frac{\sigma^2\tau^2}{n_j(\tau^2+1/n_j\sigma^2)}}$. Since $\tau^2, \sigma^2, n_j > 0$, $\tau^2 + 1/n\sigma^2 > 0$. Thus, $\frac{\sigma}{\sqrt{n_j}} > \sqrt{\frac{\sigma^2\tau^2}{n_j\tau^2+\sigma^2}}$. This shows that the posterior variance is always smaller than $\frac{\sigma^2}{n_j}$, which means that the interval from part (b) is always narrower than that of the usual z-interval. We know that the shrinkage weight is

$$w = \frac{1/\tau^2}{1/\tau^2 + \frac{n_j}{\sigma^2}}.$$

Hence, the width of the interval from part (b) can also be expressed as

$$2z_{1-\alpha/2} \cdot \sqrt{\frac{1-w}{n_j/\sigma^2}}.$$

We know that the weight controls how much the posterior mean is pulled towards μ. As n_j grows, w decreases and the posterior mean relies more on the data $\bar{y}_j$. The posterior interval is narrower because it reflects this additional information from the prior distribution, leading to a more precise estimation of θ_j.