STA 360/602L: Module 3.5b

The normal model: joint inference for mean and variance (illustration)

Dr. Olanrewaju Michael Akande

Joint posterior for normal model

- Recall that

$$\pi(\mu,\tau|Y) = \mathcal{N}\left(\mu_n, \frac{1}{\kappa_n \tau}\right) \cdot \text{Gamma}\left(\frac{\nu_n}{2}, \frac{\nu_n \sigma_n^2}{2}\right)$$
$$= \pi(\mu|Y,\tau) \cdot \pi(\tau|Y),$$

where

$$\kappa_n = \kappa_0 + n$$
$$\mu_n = \frac{\kappa_0 \mu_0 + n\bar{y}}{\kappa_n} = \frac{\kappa_0}{\kappa_n}\mu_0 + \frac{n}{\kappa_n}\bar{y}$$

$$\nu_n = \nu_0 + n$$
$$\sigma_n^2 = \frac{1}{\nu_n}\left[\nu_0\sigma_0^2 + s^2(n-1) + \frac{n\kappa_0}{\kappa_n}(\bar{y}-\mu_0)^2\right]$$
$$= \frac{1}{\nu_n}\left[\nu_0\sigma_0^2 + \sum_{i=1}^{n}(y_i-\bar{y})^2 + \frac{n\kappa_0}{\kappa_n}(\bar{y}-\mu_0)^2\right]$$

Back to our examples

- Pygmalion: questions of interest
  - Is the average improvement for the accelerated group larger than that for the no growth group?
    - What is $\Pr[\mu_A > \mu_N | Y_A, Y_N)$?
  - Is the variance of improvement scores for the accelerated group larger than that for the no growth group?
    - What is $\Pr[\sigma^2_A > \sigma^2_N | Y_A, Y_N)$?
- Job training: questions of interest
  - Is the average change in annual earnings for the training group larger than that for the no training group?
    - What is $\Pr[\mu_T > \mu_N | Y_T, Y_N)$?
  - Is the variance of change in annual earnings for the training group larger than that for the no training group?
    - What is $\Pr[\sigma^2_T > \sigma^2_N | Y_T, Y_N)$?

Mildly informative priors

- We will focus on the Pygmalion study. Follow the same approach for the job training data.
- Suppose you have no idea whether students would improve IQ on average. Set $\mu_{0A} = \mu_{0N} = 0$.
- Suppose you don't have any faith in this belief, and think it is the equivalent of having only 1 prior observation in each group. Set $\kappa_{0A} = \kappa_{0N} = 1$.
- Based on the literature, SD of change scores should be around 10 in each group, but still you don't have a lot of faith in this belief. Set $\nu_{0A} = \nu_{0N} = 1$ and $\sigma^2_{0A} = \sigma^2_{0N} = 100$.
- Graph priors to see if they accord with your beliefs. Sampling new values of Y from the priors offers a good check.

Recall the Pygmalion data

- Data:
  - Accelerated group (A): 20, 10, 19, 15, 9, 18.
  - No growth group (N): 3, 2, 6, 10, 11, 5.
- Summary statistics:
  - $\bar{y}_A = 15.2$; $s_A = 4.71$.
  - $\bar{y}_N = 6.2$; $s_N = 3.65$.

Analysis with mildly informative priors

$$\kappa_{nA} = \kappa_{0A} + n_A = 1 + 6 = 7$$
$$\kappa_{nN} = \kappa_{0N} + n_N = 1 + 6 = 7$$
$$\nu_{nA} = \nu_{0A} + n_A = 1 + 6 = 7$$
$$\nu_{nN} = \nu_{0N} + n_N = 1 + 6 = 7$$

$$\mu_{nA} = \frac{\kappa_{0A}\mu_{0A} + n_A\bar{y}_A}{\kappa_{nA}} = \frac{(1)(0) + (6)(15.2)}{7} \approx 13.03$$
$$\mu_{nN} = \frac{\kappa_{0N}\mu_{0N} + n_N\bar{y}_N}{\kappa_{nN}} = \frac{(1)(0) + (6)(6.2)}{7} \approx 5.31$$

$$\sigma^2_{nA} = \frac{1}{\nu_{nA}}\left[\nu_{0A}\sigma^2_{0A} + s^2_A(n_A - 1) + \frac{n_A\kappa_{0A}}{\kappa_{nA}}(\bar{y}_A - \mu_{0A})^2\right]$$
$$= \frac{1}{7}\left[(1)(100) + (22.17)(5) + \frac{(6)(1)}{(7)}(15.2 - 0)^2\right] \approx 58.41$$

$$\sigma^2_{nN} = \frac{1}{\nu_{nN}}\left[\nu_{0N}\sigma^2_{0N} + s^2_N(n_N - 1) + \frac{n_N\kappa_{0N}}{\kappa_{nN}}(\bar{y}_N - \mu_{0N})^2\right]$$
$$= \frac{1}{7}\left[(1)(100) + (13.37)(5) + \frac{(6)(1)}{(7)}(6.2 - 0)^2\right] \approx 28.54$$

Analysis with mildly informative priors

- So our joint posterior is

$$\mu_A | Y_A, \tau_A \sim \mathcal{N}\left(\mu_{nA}, \frac{1}{\kappa_{nA}\tau_A}\right) = \mathcal{N}\left(13.03, \frac{1}{7\tau_A}\right)$$

$$\tau_A | Y_A \sim \text{Gamma}\left(\frac{\nu_{nA}}{2}, \frac{\nu_{nA}\sigma^2_{nA}}{2}\right) = \text{Gamma}\left(\frac{7}{2}, \frac{7(58.41)}{2}\right)$$

$$\mu_N | Y_N, \tau_N \sim \mathcal{N}\left(\mu_{nN}, \frac{1}{\kappa_{nN}\tau_N}\right) = \mathcal{N}\left(5.31, \frac{1}{7\tau_N}\right)$$

$$\tau_N | Y_N \sim \text{Gamma}\left(\frac{\nu_{nN}}{2}, \frac{\nu_{nN}\sigma^2_{nN}}{2}\right) = \text{Gamma}\left(\frac{7}{2}, \frac{7(28.54)}{2}\right)$$

Monte Carlo sampling

- To evaluate whether the accelerated group has larger IQ gains than the normal group, we would like to estimate quantities like $\Pr[\mu_A > \mu_N | Y_A, Y_N)$ which are based on the **marginal posterior** of μ rather than the **conditional distribution**.

- Fortunately, this is easy to do by generating samples of μ and σ^2 from their joint posterior.

Monte Carlo sampling

- Suppose we simulate values using the following Monte Carlo procedure:

$$
\begin{aligned}
\tau^{(1)} &\sim \text{Gamma}\left(\frac{\nu_n}{2}, \frac{\nu_n \sigma_n^2}{2}\right) \\
\mu^{(1)} &\sim \mathcal{N}\left(\mu_n, \frac{1}{\kappa_n \tau^{(1)}}\right) \\
\\
\tau^{(2)} &\sim \text{Gamma}\left(\frac{\nu_n}{2}, \frac{\nu_n \sigma_n^2}{2}\right) \\
\mu^{(2)} &\sim \mathcal{N}\left(\mu_n, \frac{1}{\kappa_n \tau^{(2)}}\right) \\
&\vdots \\
&\vdots \\
&\vdots \\
\tau^{(m)} &\sim \text{Gamma}\left(\frac{\nu_n}{2}, \frac{\nu_n \sigma_n^2}{2}\right) \\
\mu^{(m)} &\sim \mathcal{N}\left(\mu_n, \frac{1}{\kappa_n \tau^{(m)}}\right)
\end{aligned}
$$

Monte Carlo sampling

- Note that we are sampling each $\mu^{(j)}$, $j = 1, \ldots, m$, from its conditional distribution, not from the marginal.
- The sequence of pairs $\{(\tau, \mu)^{(1)}, \ldots, (\tau, \mu)^{(m)}\}$ simulated using this method are independent samples from the joint posterior $\pi(\mu, \tau | Y)$.
- Additionally, the simulated sequence $\{\mu^{(1)}, \ldots, \mu^{(m)}\}$ are independent samples from the **marginal posterior distribution**.
- While this may seem odd, keep in mind that while we drew the μ's as conditional samples, each was conditional on a different value of τ.
- Thus, together they constitute marginal samples of μ.

Monte Carlo sampling

It is easy to sample from these posteriors:

```
aA <- 7/2
aN <- 7/2
bA <- (7/2)*58.41
bN <- (7/2)*28.54
muA <- 13.03
muN <- 5.31
kappaA <- 7
kappaN <- 7
tauA_postsample <- rgamma(10000,aA,bA)
thetaA_postsample <- rnorm(10000,muA,sqrt(1/(kappaA*tauA_postsample)))
tauN_postsample <- rgamma(10000,aN,bN)
thetaN_postsample <- rnorm(10000,muN,sqrt(1/(kappaN*tauN_postsample)))
sigma2A_postsample <- 1/tauA_postsample
sigma2N_postsample <- 1/tauN_postsample
```

Monte Carlo sampling

- Is the average improvement for the accelerated group larger than that for the no growth group?
  - What is $\Pr[\mu_A > \mu_N | Y_A, Y_N)$?

```
mean(thetaA_postsample > thetaN_postsample)
```

```
## [1] 0.9681
```

- Is the variance of improvement scores for the accelerated group larger than that for the no growth group?
  - What is $\Pr[\sigma^2_A > \sigma^2_N | Y_A, Y_N)$?

```
mean(sigma2A_postsample > sigma2N_postsample)
```

```
## [1] 0.8092
```

- What can we conclude from this?

Recall the job training data

- Data:
  - No training group (N): sample size $n_N = 429$.
  - Training group (T): sample size $n_A = 185$.
- Summary statistics for change in annual earnings:
  - $\bar{y}_N = 1364.93$; $s_N = 7460.05$
  - $\bar{y}_T = 4253.57$; $s_T = 8926.99$
- Using the same approach we used for the Pygmalion data, answer the questions of interest.

What's next?

Move on to the readings for the next module!