Exchangeability de-Finetti theorem

- Eight-school example revisit:

A B C D E F G H

casual inference based on { observational studies / designed experiments }

e.g. smoking → cancer ?

We throw a specific problem, by restricting a direction for casual inference.

Treatment	Control
$Y_i(1)$	?
?	$Y_{i'}(0)$

problem: each student can only be put into one group.

But ... treatment effect: $\begin{cases} Y_i(1) - Y_i(0) \\ Y_{i'}(1) - Y_{i'}(0) \end{cases}$

Solution:

→ Another exam scheduled. → look at student performance clusters.

SAT score

use this result to decide which student to treatment/control group

After the above preprocessing,

$$\begin{cases} \bar{y}_j \sim N(\theta_j, \sigma_j^2) \\ \quad \text{estimated treatment effect for school } j. \\ \theta_j \sim N(\mu, \tau^2) \quad , \quad p(\mu, \tau^2) \propto p(\tau^2) \end{cases}$$

Choice of τ^2 distribution:

① $\boxed{p(\tau) \propto 1}$ $\Downarrow$ proper posterior. , $\boxed{p(\log \tau) \propto 1 \Leftrightarrow p(\tau) \propto \frac{1}{\tau}}$ $\Downarrow$ improper posterior.

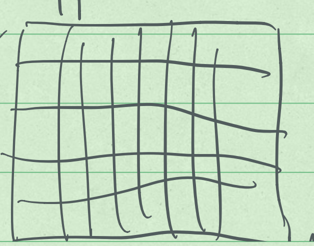
joint posterior dis' $P(\mu, \tau^2, \theta_{1:J} \mid \bar{y}_{1:J})$

~~Method~~ 1, 2, 3

$$p(\mu, \tau^2 \mid \bar{y}_{1:J}) \propto p(\tau) \prod_{j=1}^{J} \left[\frac{1}{\sqrt{\tau^2 + \sigma_j^2}} e^{-\frac{(\bar{y}_j - \mu)^2}{2(\tau^2 + \sigma_j^2)}} \right]$$

$\downarrow$ Grid Approximation

τ^2



$$\mu \mid \tau^2, \bar{y}_{1:J} \sim N\left(\frac{\sum_{j=1}^{J} \bar{y}_j / (\tau^2 + \sigma_j^2)}{\sum_{j=1}^{J} 1/(\tau^2 + \sigma_j^2)} = \hat{\mu}_J, \frac{1}{\sum_{j=1}^{J} \frac{1}{\tau^2 + \sigma_j^2}} \right)$$

Recall in Gaussian–Gaussian single param model:

Another way mention:

$$p(\tau \mid \bar{y}_{1:J}) = \frac{p(\mu, \tau^2 \mid \bar{y}_{1:J})}{p(\mu \mid \bar{y}_{1:J}, \tau^2)}$$

$$\propto \frac{p(\tau) \prod_{j=1}^{J} \frac{1}{\sqrt{\tau^2+\sigma_j^2}} \exp\left(-\frac{\sum_{j=1}^{J}(\bar{y}_j-\mu)^2}{2(\tau^2+\sigma_j^2)}\right)}{\sqrt{\sum_{j=1}^{J}\frac{1}{\tau^2+\sigma_j^2}} \exp\left(-\frac{\sum_{j=1}^{J}\left(\frac{1}{\tau^2+\sigma_j^2}\right)(\mu-\hat{\mu}_J)^2}{2}\right)}$$

Now, examine why $p(\tau) \propto 1 \rightarrow$ proper posterior.

$p(\mu, \tau^2 \mid \bar{y}_{1:J}) \propto p(\tau) \left[\prod_{j=1}^{J} \frac{1}{\sqrt{\tau^2+\sigma_j^2}}\right] e^{-\frac{\Sigma(\ )^2}{2(\ )}}$ (finite integrate ...):

$\int_{\mathbb{R}} d\mu \int_{\mathbb{R}^+} d\tau$

$\rightarrow \tau \to \infty$: integrable

$\searrow \tau \to 0$: still bounded since $\frac{1}{\sqrt{\tau^2+\sigma_j^2}}$

$\downarrow$

only $p(\tau)$ matters,

so if $p(\tau) \propto 1/\tau \Rightarrow$ problem!

& $p(\tau) \propto 1$ ✓

$\mu \to \pm\infty$

$\tau \to 0$

$\tau \to \infty$

Sampling procedure: (× Gibbs) conditional iterations

1. $\tau^{(i)} \sim p(\tau \mid \bar{y}_{1:J})$ grid approximation.
2. $\mu^{(i)} \sim p(\mu \mid \tau^{(i)}, \bar{y}_{1:J})$ Gaussian.
3. $\theta_j^{(i)} \sim p(\theta_j \mid \mu, \tau^{(i)}, \bar{y}_{1:J})$ Gaussian

$\downarrow$

samples $p(\mu, \tau^2, \theta_{1:J} \mid \bar{y}_{1:J})$