Aluno: Cleiton Moya de Almeida
EST5104 - Inferência Bayesiana
Prof. Josemar Rodrigues
29/06/2025

Lista 6 - Modelo Hierárquico

Enunciado

1. Justificar o Caso I (μ fixo) do modelo hierárquico;
2. No Caso II (μ desconhecido), obter a $\pi(\mu \mid \boldsymbol{y})$;
3. No Caso II, justificar a $\text{Var}[\theta_j \mid \boldsymbol{y}]$.

Modelo Hierárquico

A ideia geral do modelo hierárquico é relacionar modelos individuais de grupos de dados a um modelo geral, latente, comum a todos estes grupos. Vamos detalhar melhor esta ideia.

Suponha que temos J grupos de dados na seguinte forma:

$$\begin{aligned}
\boldsymbol{y}_1 &= (y_{1,1}, y_{1,2}, \ldots, y_{1,n_1}) \\
\boldsymbol{y}_2 &= (y_{2,1}, y_{2,2}, \ldots, y_{2,n_2}) \\
&\vdots \\
\boldsymbol{y}_j &= (y_{j,1}, y_{j,2}, \ldots, y_{j,n_j}) \\
&\vdots \\
\boldsymbol{y}_J &= (y_{J,1}, y_{J,2}, \ldots, y_{J,n_J})
\end{aligned}$$

Vamos assumir que cada grupo segue uma distribuição normal com variância conhecida e comum a todos os grupos, porém com médias θ_j individuais para cada grupo:

$$y_{j,i} \mid \theta_j, \sigma^2 \sim \mathcal{N}\left(\theta_j, \sigma^2\right) \tag{1}$$

A ideia do modelo hierárquico é então supor que cada θ_j é gerado por um outro modelo, oculto e comum a todos os grupos. Vamos supor que este modelo oculto também é normal:

$$\theta_j \mid \mu, \tau^2 \sim \mathcal{N}\left(\mu, \tau^2\right) \tag{2}$$

Podemos interpretar este modelo como uma priori para cada θ_j. Agora, temos duas possibilidades:

- **Caso I**: Supomos que μ é fixo (conhecido);
- **Caso II**: Supomos que μ é gerado por um outro modelo normal com variância w^2 conhecida:
$$\mu \sim \mathcal{N}\left(0, w^2\right) \tag{3}$$
Ou seja, neste segundo caso estamos adotando uma priori hierárquica para os θ_j.

Em resumo, no modelo hierárquico trabalhamos com 3 estágios:

- **1º estágio**: verosimilhança dos dados: $y_{j,i} \mid \theta_j, \sigma^2 \sim \mathcal{N}\left(\theta_j, \sigma^2\right)$
- **2º estágio**: modelo latente para θ_j: $\theta_j \mid \mu, \tau^2 \sim \mathcal{N}\left(\mu, \tau^2\right)$
- **3º estágio** (Caso II): priori para μ: $\mu \sim \mathcal{N}\left(0, w^2\right)$

Questão 1

No caso em que μ é fixo, a distribuição a posteriori para θ_j é dada por

$$\begin{aligned}
\pi(\theta_j \mid \boldsymbol{y}_j, \mu, \tau^2, \sigma^2) &\propto f(\boldsymbol{y}_j \mid \theta_j, \sigma^2)\pi(\theta_j \mid \mu, \tau^2) \\
&\propto \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_j}(y_{j,i} - \theta_j)^2\right\} \exp\left\{-\frac{1}{2\tau^2}(\theta_j - \mu)^2\right\} \\
&\propto \exp\left\{-\frac{n_j}{2\sigma^2}(\theta_j - \overline{y_j}\,)^2\right\} \exp\left\{-\frac{1}{2\tau^2}(\theta_j - \mu)^2\right\} \\
&\propto \exp\left\{-\frac{n_j}{2\sigma^2}(\theta_j - \overline{y_j}\,)^2 - \frac{1}{2\tau^2}(\theta_j - \mu)^2\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[\frac{n_j}{\sigma^2}(\theta_j - \overline{y_j}\,)^2 + \frac{1}{\tau^2}(\theta_j - \mu)^2\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[\theta_j^2\left(\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}\right) - 2\theta_j\left(\frac{n_j\,\overline{y_j}}{\sigma^2} + \frac{\mu}{\tau^2}\right) + \text{ const.}\right]\right\} \\
&\propto \exp\left\{-\frac{\left(\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}\right)}{2}\left(\theta_j - \frac{\frac{n_j\,\overline{y_j}}{\sigma^2} + \frac{\mu}{\tau^2}}{\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}}\right)^2\right\} \\
&\sim \mathcal{N}\left(\tilde{\mu}_j, \tilde{\sigma}_j^2\right)
\end{aligned} \tag{4}$$

em que $\overline{y_j} = \sum_{i=1}^{n_j} y_{j,i}$ e

$$\tilde{\mu}_j = \frac{\frac{n_j\,\overline{y_j}}{\sigma^2} + \frac{\mu}{\tau^2}}{\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}} = \frac{n_j\tau_j\,\overline{y_j}\, + \mu\sigma^2}{n_j\tau^2 + \sigma^2} \tag{5}$$

$$\tilde{\sigma}_j^2 = \frac{1}{\frac{n_j}{\sigma^2} + \frac{1}{\tau^2}} = \frac{\tau^2\sigma^2}{n_j\tau^2 + \sigma^2} \tag{6}$$

Questão 2

No Caso II, assumimos que μ é desconhecido e tem uma distribuição a priori $\mu \sim \mathcal{N}(0, w^2)$, onde a variância w^2 é conhecida. Nosso objetivo é calcularmos a posteriori $\pi(\mu \mid \boldsymbol{y})$.

Como de praxe, partimos do Teorema de Bayes:

$$\pi(\mu \mid \boldsymbol{y}) \propto f(\boldsymbol{y} \mid \mu)\pi(\mu) \tag{7}$$

Como os J grupos são condicionalmente independentes na presença de μ:

$$\pi(\mu \mid \boldsymbol{y}) \propto \left[\prod_{j=1}^{J} f(\boldsymbol{y}_j|\mu)\right]\pi(\mu) \tag{8}$$

Agora, a alternativa mais direta seria obter a verosimilhança marginal $f(\boldsymbol{y}_j|\mu)$ integrando os θ_j:

$$f(\boldsymbol{y}_j \mid \mu) = \int f(\boldsymbol{y}_j|\theta_j)\pi(\theta_j|\mu)d\theta_j \tag{9}$$

Porém, este caminho envolve muito cálculo. Vamos tomar um atalho utilizando as propriedades da normalidade e as leis da esperança e variância totais.

Primeiramente, note que $f(\boldsymbol{y}_j|\mu) \propto f(\overline{y_j} \; |\mu)$, pois podemos re-escrever a verosimilhança em termos da estatística suficiente, como fizemos na questão anterior.

Agora,a segunda ideia fundamental é notar que $f(\overline{y_j} \; |\mu)$ possui distribuição normal, pois $\overline{y_j}$ é uma combinação linear de variáveis aleatórias normais:

$$\overline{y_j} = \frac{1}{n_j}\sum_{i=1}^{n_j} y_{j,i} \tag{10}$$

Resta então calcularmos $\mathbb{E}[\overline{y_j} \; |\mu]$ e $\mathrm{Var}[\overline{y_j} \; |\mu]$.

Para calcularmos $\mathbb{E}[\overline{y_j} \; |\mu]$, utilizamos a Lei da Esperança Total na forma condicional:

$$\mathbb{E}[\overline{y_j} \; |\mu] = \mathbb{E}\left[\; \mathbb{E}[\overline{y_j} \; |\mu, \theta_j] \mid \mu\right] \tag{11}$$

Como $\overline{y_j}$ não depende de μ na presença de θ_j:

$$\mathbb{E}[\overline{y_j} \; |\mu, \theta_j] = \mathbb{E}[\overline{y_j} \; |\theta_j] = \theta_j \tag{12}$$

e então a eq. (11) fica:

$$\mathbb{E}[\overline{y_j} \; |\mu] = \mathbb{E}[\theta_j \mid \mu] = \mu \tag{13}$$

De forma análoga, calcularmos $\mathrm{Var}[\overline{y_j} \; |\mu]$, utilizamos a Lei da Variância Total na forma condicional:

$$\begin{aligned}\mathrm{Var}[\overline{y_j} \; |\mu] &= \mathbb{E}\left[\; \mathrm{Var}[\overline{y_j} \; |\mu, \theta_j] \mid \mu\right] + \mathrm{Var}\left[\; \mathbb{E}[\overline{y_j} \; |\mu, \theta_j] \mid \mu\right] \\ &= \mathbb{E}\left[\; \mathrm{Var}[\overline{y_j} \; |\theta_j] \mid \mu\right] + \mathrm{Var}\left[\theta_j \mid \mu\right]\end{aligned} \tag{14}$$

Mas $\mathrm{Var}[\overline{y_j} \; |\theta_j] = \sigma^2/n_j$ e $\mathrm{Var}[\theta_j|\mu] = \tau^2$. Então:

$$
\begin{aligned}
\mathrm{Var}[\overline{y_j} \mid \mu] &= \mathbb{E}\left[\frac{\sigma^2}{n_j} \mid \mu\right] + \mathrm{Var}\left[\theta_j \mid \mu\right] \\
&= \frac{\sigma^2}{n_j} + \tau^2
\end{aligned} \tag{15}
$$

Concluímos então que

$$
\overline{y_j} \mid \mu \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n_j} + \tau^2\right). \tag{16}
$$

Finalmente, retomando à eq. (8), agora é questão de algebrismo para combinarmos os termos exponenciais:

$$
\begin{aligned}
\pi(\mu \mid \boldsymbol{y}) &\propto \left[\prod_{j=1}^{J} f(\boldsymbol{y}_j|\mu)\right]\pi(\mu) \\
&\propto \left[\prod_{j=1}^{J} f(\overline{y_j} \mid \mu)\right]\pi(\mu) \\
&\propto \exp\left\{-\frac{1}{2}\sum_{j=1}^{J}\frac{(\overline{y_j} - \mu)^2}{\frac{\sigma^2}{n_j} + \tau^2}\right\}\exp\left\{-\frac{\mu^2}{2w^2}\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[\sum_{j=1}^{J}\frac{(\overline{y_j} - \mu)^2}{\frac{\sigma^2}{n_j} + \tau^2} + \frac{\mu^2}{w^2}\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[\mu^2\underbrace{\left(\sum_{j=1}^{J}\frac{1}{\frac{\sigma^2}{n_j} + \tau^2} + \frac{1}{w^2}\right)}_{A} - 2\mu\underbrace{\left(\sum_{j=1}^{J}\frac{\overline{y_j}}{\frac{\sigma^2}{n_j} + \tau^2}\right)}_{B} + \text{const.}\right]\right\} \\
&\sim \mathcal{N}\left(\frac{B}{A}, \frac{1}{A}\right)
\end{aligned} \tag{17}
$$

De onde concluímos que

$$
\pi(\mu \mid \boldsymbol{y}) \sim \mathcal{N}\left(\tilde{\mu}, \tilde{w}^2\right), \tag{18}
$$

$$
\tilde{w}^2 = \left(\sum_{j=1}^{J}\frac{1}{\frac{\sigma^2}{n_j} + \tau^2} + \frac{1}{w^2}\right)^{-1}, \tag{19}
$$

$$
\tilde{\mu} = \tilde{w}^2\left(\sum_{j=1}^{J}\frac{\overline{y_j}}{\frac{\sigma^2}{n_j} + \tau^2}\right). \tag{20}
$$

Questão 3

Nesta questão, o objetivo é calcularmos $\text{Var}[\theta_j \mid \boldsymbol{y}]$. Para isso, usamos a Lei da Variância Total na forma condicional:

$$\text{Var}[\theta_j \mid \boldsymbol{y}] = \mathbb{E}\left[\ \text{Var}[\theta_j|\mu, \boldsymbol{y}] \mid \boldsymbol{y}\right] + \text{Var}\left[\ \mathbb{E}[\theta_j|\mu, \boldsymbol{y}] \mid \boldsymbol{y}\right] \tag{21}$$

Para o primeiro termo da equação, observe que, com μ fixo,

$$\text{Var}[\theta_j|\mu, \boldsymbol{y}] = \text{Var}[\theta_j|\mu] = \tilde{\sigma}_j^2 = \frac{\tau^2\sigma^2}{n_j\tau^2 + \sigma^2} \tag{22}$$

com $\tilde{\sigma}_j^2$ data pela eq. (6) da Questão 1. Observe também que, neste caso, $\mathbb{E}[\tilde{\sigma}_j^2] = \tilde{\sigma}_j^2$.

Agora, para o segundo termo da eq. (21), com μ fixo,

$$\mathbb{E}[\theta_j \mid \mu, \boldsymbol{y}] = \mathbb{E}[\theta_j \mid \mu] = \tilde{\mu}_j \tag{23}$$

com $\tilde{\mu}_j$ calculado na Questão 1, eq. (5). Então:

$$\begin{aligned} \text{Var}\left[\ \mathbb{E}[\theta_j|\mu, \boldsymbol{y}] \mid \boldsymbol{y}\right] &= \text{Var}\left[\frac{n_j\tau_j\,\overline{y_j}\ + \mu\sigma^2}{n_j\tau^2 + \sigma^2} \mid \boldsymbol{y}\right] \\ &= \left(\frac{\sigma^2}{n_j\tau^2 + \sigma^2}\right)^2 \text{Var}[\mu|\boldsymbol{y}] \end{aligned} \tag{24}$$

com $\text{Var}[\mu|\boldsymbol{y}] = \tilde{w}^2$ calculada na Questão 2, eq. (19). Finalmente, juntando as equações (22) e (24):

$$\begin{aligned} \text{Var}[\theta_j \mid \boldsymbol{y}] &= \frac{\tau^2\sigma^2}{n_j\tau^2 + \sigma^2} + \left(\frac{\sigma^2}{n_j\tau^2 + \sigma^2}\right)^2 \tilde{w}^2 \\ &= \frac{\tau^2\sigma^2}{n_j\tau^2 + \sigma^2} + \left(\frac{\sigma^2}{n_j\tau^2 + \sigma^2}\right)^2 \left(\sum_{j=1}^{J} \frac{1}{\frac{\sigma^2}{n_j} + \tau^2} + \frac{1}{w^2}\right)^{-1} \end{aligned} \tag{25}$$