Estadística Bayesiana

Clase 4: Modelos multiparamétricos: modelo Normal

Isabel Cristina Ramírez Guevara

Escuela de Estadística
Universidad Nacional de Colombia, Sede Medellín

Medellín, 13 de agosto de 2020

Modelos multiparamétricos

Tenemos $\theta = (\theta_1, \theta_2)$. Si el interés de la inferencia se centra sobre θ_1 solamente, entonces θ_2 es considerado un parámetro de "molestia"(nuisance). Vamos a obtener la posterior conjunta

$$p(\theta_1, \theta_2|y) \propto p(y|\theta_1, \theta_2)p(\theta_1, \theta_2)$$

podemos promediar sobre θ_2

$$\begin{aligned} p(\theta_1|y) &= \int p(\theta_1, \theta_2|y)\, d\theta_2 \\ &= \int p(\theta_1|\theta_2, y)\, p(\theta_2|y)\, d\theta_2 \end{aligned}$$

$$p(A_1 \cap \cdots \cap A_n) = p(A_1)p(A_2|A_1)p(A_3|A_1 \cap A_2) \cdots p(A_n|A_1 \cdots A_{n-1})$$

Modelo normal univariado

Sea **y** un vector de n iid observaciones $N(\theta, \sigma^2)$. La verosimilitud para este modelo es:

$$p(\mathbf{y}|\theta, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right]$$

Modelo normal univariado

Sea **y** un vector de n iid observaciones $N(\theta, \sigma^2)$. La verosimilitud para este modelo es:

$$p(\mathbf{y}|\theta, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right]$$

$$p(\mathbf{y}|\theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}((y_i - \bar{y}) + (\bar{y} - \theta))^2\right]$$

Modelo normal univariado

Sea **y** un vector de n iid observaciones $N(\theta, \sigma^2)$. La verosimilitud para este modelo es:

$$p(\mathbf{y}|\theta, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right]$$

$$p(\mathbf{y}|\theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}((y_i - \bar{y}) + (\bar{y} - \theta))^2\right]$$

$p(\mathbf{y}|\theta, \sigma^2) =$

$$\frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[(y_i - \bar{y})^2 + 2(y_i - \bar{y})(\bar{y} - \theta) + (\bar{y} - \theta)^2\right]\right]$$

Modelo normal univariado

$$p(\mathbf{y}|\theta,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(y_i-\bar{y})^2 + n(\bar{y}-\theta)^2\right]\right]$$

Modelo normal univariado

$$p(\mathbf{y}|\theta,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(y_i-\bar{y})^2 + n(\bar{y}-\theta)^2\right]\right]$$

$$p(\mathbf{y}|\theta,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\theta)^2\right]\right]$$

donde $s^2 = \frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}$.

Escenario 1: *Suponer que la distribución a priori $p(\theta)$ es independiente de la distribución a priori $p(\sigma^2)$ y que ambas distribuciones son* **no** *informativas.*

Modelo normal univariado

Por lo tanto,

$$p(\theta, \sigma^2) = p(\theta)p(\sigma^2)$$

Modelo normal univariado

Por lo tanto,

$$p(\theta, \sigma^2) = p(\theta)p(\sigma^2)$$

Si tenemos como distribución a priori para θ una $\mathrm{N}(\mu_0, \tau_0^2)$, entonces $p(\theta) \propto \exp\left[-\frac{1}{2\tau_0^2}(\theta - \mu_0)^2\right]$. Cuando τ_0^2 se vuelve más grande (no informativa) entonces la forma de la distribución a priori de θ se torna vaga y se puede aproximar a una distribución plana tal que: $p(\theta) \propto$ constante.

Modelo normal univariado

Por lo tanto,

$$p(\theta, \sigma^2) = p(\theta)p(\sigma^2)$$

Si tenemos como distribución a priori para θ una $\mathrm{N}(\mu_0, \tau_0^2)$, entonces $p(\theta) \propto \exp\left[-\frac{1}{2\tau_0^2}(\theta - \mu_0)^2\right]$. Cuando τ_0^2 se vuelve más grande (no informativa) entonces la forma de la distribución a priori de θ se torna vaga y se puede aproximar a una distribución plana tal que: $p(\theta) \propto$ constante.

Si $\sigma^2 \sim$ Gamma-inversa(α, β) entonces $p(\sigma^2) \propto (\sigma^2)^{-(\alpha+1)}e^{-\beta/\sigma^2}$, para que esta distribución sea no informativa $\alpha \to 0$ y $\beta \to 0$, por lo tanto $p(\sigma^2) \propto \frac{1}{\sigma^2}$. Se tiene entonces $p(\theta, \sigma^2) \propto \frac{1}{\sigma^2}$.

Modelo normal univariado

La distribución posterior conjunta es:

$$
\begin{aligned}
p(\theta, \sigma^2|\mathbf{y}) &\propto p(\mathbf{y}|\theta, \sigma^2)p(\theta, \sigma^2) \\
&\propto \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\theta)^2\right]\right]\frac{1}{\sigma^2}
\end{aligned}
$$

Modelo normal univariado

La distribución posterior conjunta es:

$$
\begin{aligned}
p(\theta, \sigma^2|\mathbf{y}) &\propto p(\mathbf{y}|\theta, \sigma^2)p(\theta, \sigma^2) \\
&\propto \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\theta)^2\right]\right]\frac{1}{\sigma^2}
\end{aligned}
$$

Por lo tanto la distribución posterior de θ es:

$$
\begin{aligned}
p(\theta|\mathbf{y}) &= \int_0^\infty p(\theta, \sigma^2|\mathbf{y})d\sigma^2 \\
&\propto \int_0^\infty \frac{1}{(\sigma^2)^{n/2+1}} \exp\left[-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\theta)^2\right]\right]d\sigma^2
\end{aligned}
$$

Modelo normal univariado

Haciendo la sustitución $A = (n-1)s^2 + n(\bar{y}-\theta)^2$ y $z = \frac{A}{2\sigma^2}$, por lo tanto $\sigma^2 = \frac{A}{2z}$ $d\sigma^2 = -\frac{A}{2z^2}dz$. Se obtiene la integral gamma no normalizada:

$$p(\theta|\mathbf{y}) \propto \int_0^\infty \left(\frac{2z}{A}\right)^{n/2+1} \exp\left[-\frac{2zA}{2A}\right] \left(\frac{A}{2z^2}\right) dz$$

Modelo normal univariado

Haciendo la sustitución $A = (n-1)s^2 + n(\bar{y}-\theta)^2$ y $z = \frac{A}{2\sigma^2}$, por lo tanto $\sigma^2 = \frac{A}{2z}$ $d\sigma^2 = -\frac{A}{2z^2}dz$. Se obtiene la integral gamma no normalizada:

$$p(\theta|\mathbf{y}) \propto \int_0^\infty \left(\frac{2z}{A}\right)^{n/2+1} \exp\left[-\frac{2zA}{2A}\right] \left(\frac{A}{2z^2}\right) dz$$

$$p(\theta|\mathbf{y}) \propto A^{-n/2} \int_0^\infty z^{n/2-1} e^{-z} dz$$

Modelo normal univariado

Haciendo la sustitución $A = (n-1)s^2 + n(\bar{y} - \theta)^2$ y $z = \frac{A}{2\sigma^2}$, por lo tanto $\sigma^2 = \frac{A}{2z}$ $d\sigma^2 = -\frac{A}{2z^2}dz$. Se obtiene la integral gamma no normalizada:

$$p(\theta|\mathbf{y}) \propto \int_0^\infty \left(\frac{2z}{A}\right)^{n/2+1} \exp\left[-\frac{2zA}{2A}\right] \left(\frac{A}{2z^2}\right) dz$$

$$p(\theta|\mathbf{y}) \propto A^{-n/2} \int_0^\infty z^{n/2-1} e^{-z} dz$$

$$p(\theta|\mathbf{y}) \propto \left[1 + \frac{n(\theta - \bar{y})^2}{(n-1)s^2}\right]^{-n/2}$$

Modelo normal univariado

Haciendo la sustitución $A = (n-1)s^2 + n(\bar{y}-\theta)^2$ y $z = \frac{A}{2\sigma^2}$, por lo tanto $\sigma^2 = \frac{A}{2z}$ $d\sigma^2 = -\frac{A}{2z^2}dz$. Se obtiene la integral gamma no normalizada:

$$p(\theta|\mathbf{y}) \propto \int_0^\infty \left(\frac{2z}{A}\right)^{n/2+1} \exp\left[-\frac{2zA}{2A}\right] \left(\frac{A}{2z^2}\right) dz$$

$$p(\theta|\mathbf{y}) \propto A^{-n/2} \int_0^\infty z^{n/2-1} e^{-z} dz$$

$$p(\theta|\mathbf{y}) \propto \left[1 + \frac{n(\theta-\bar{y})^2}{(n-1)s^2}\right]^{-n/2}$$

$$p(\theta|\mathbf{y}) \propto \left[1 + \frac{1}{n-1}\left(\frac{\theta-\bar{y}}{s/\sqrt{n}}\right)^2\right]^{-\frac{n+1-1}{2}},$$

Modelo normal univariado

por lo tanto $\theta|\mathbf{y} \sim \mathrm{t}_{n-1}(\bar{y}, s^2/n)$. Esta distribución también se puede expresar como: $\left.\frac{\theta-\bar{y}}{s/\sqrt{n}}\right|\mathbf{y} \sim \mathrm{t}_{n-1}$.

Modelo normal univariado

por lo tanto $\theta|\mathbf{y} \sim \mathrm{t}_{n-1}(\bar{y}, s^2/n)$. Esta distribución también se puede expresar como: $\frac{\theta-\bar{y}}{s/\sqrt{n}}\Big|\mathbf{y} \sim \mathrm{t}_{n-1}$.

La distribución posterior de σ^2 se obtiene:

$$p(\sigma^2|\mathbf{y}) = \int_{-\infty}^{\infty} p(\theta, \sigma^2|\mathbf{y}) d\theta$$

Modelo normal univariado

por lo tanto $\theta|\mathbf{y} \sim \mathrm{t}_{n-1}(\bar{y}, s^2/n)$. Esta distribución también se puede expresar como: $\frac{\theta-\bar{y}}{s/\sqrt{n}}\Big|\mathbf{y} \sim \mathrm{t}_{n-1}$.

La distribución posterior de σ^2 se obtiene:

$$p(\sigma^2|\mathbf{y}) = \int_{-\infty}^{\infty} p(\theta, \sigma^2|\mathbf{y}) d\theta$$

$$p(\sigma^2|\mathbf{y}) \propto \int_{-\infty}^{\infty} \frac{1}{\sigma^2} \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\theta)^2\right]\right] d\theta$$

Modelo normal univariado

por lo tanto $\theta|\mathbf{y} \sim \mathrm{t}_{n-1}(\bar{y}, s^2/n)$. Esta distribución también se puede expresar como: $\frac{\theta-\bar{y}}{s/\sqrt{n}}\Big|\mathbf{y} \sim \mathrm{t}_{n-1}$.

La distribución posterior de σ^2 se obtiene:

$$p(\sigma^2|\mathbf{y}) = \int_{-\infty}^{\infty} p(\theta, \sigma^2|\mathbf{y})d\theta$$

$$p(\sigma^2|\mathbf{y}) \propto \int_{-\infty}^{\infty} \frac{1}{\sigma^2}\frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\theta)^2\right]\right] d\theta$$

$$p(\sigma^2|\mathbf{y}) \propto$$

$$\frac{\sqrt{2\pi\sigma^2}}{\sigma^2(2\pi\sigma^2)^{n/2}\sqrt{n}} \exp\left[-\frac{1}{2\sigma^2}(n-1)s^2\right] \int_{-\infty}^{\infty} \exp\left[-\frac{n}{2\sigma^2}\left[(\bar{y}-\theta)^2\right]\right] \frac{\sqrt{n}}{\sqrt{2\pi\sigma^2}} d\theta$$

Modelo normal univariado

$$p(\sigma^2|\mathbf{y}) \propto (\sigma^2)^{-\frac{1}{2}(n-1)-1} \exp\left[-\frac{1}{2\sigma^2}(n-1)s^2\right]$$

Por lo tanto $\sigma^2|\mathbf{y} \sim$Gamma-inversa$\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$ o $\sigma^2|\mathbf{y} \sim$Scaled-Inv-$\chi^2(n-1, s^2)$ o $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

Modelo normal univariado

$$p(\sigma^2|\mathbf{y}) \propto (\sigma^2)^{-\frac{1}{2}(n-1)-1} \exp\left[-\frac{1}{2\sigma^2}(n-1)s^2\right]$$

Por lo tanto $\sigma^2|\mathbf{y} \sim$Gamma-inversa$\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$ o $\sigma^2|\mathbf{y} \sim$Scaled-Inv-$\chi^2(n-1, s^2)$ o $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

La distribución predictiva posterior para una nueva observación $\tilde{y}$ se puede escribir como $p(\tilde{y}|\mathrm{y}) = \int\int p(\tilde{y}|\theta, \sigma^2, \mathrm{y})p(\theta, \sigma^2|\mathrm{y})d\theta d\sigma^2$. El primer factor corresponde a la distribución normal para la observación futura dados los valores (θ, σ^2). Se tiene entonces que $\tilde{y}|\mathbf{y} \sim t_{n-1}\left(\bar{y}, \left(1+\frac{1}{n}\right)s^2\right)$.

Modelo normal univariado

Distribución condicional posterior $p(\theta|\sigma^2, \mathbf{y})$. En esta caso vamos a encontrar el mismo resultado de la distribución posterior de θ con varianza conocida y una distribución a priori uniforme, se tiene entonces:

$$\theta|\sigma^2, \mathbf{y} \sim \mathrm{N}\left(\bar{y}, \frac{\sigma^2}{n}\right)$$

Modelo normal univariado

Ejemplo

Simon Newcomb realizó un experimento en 1882 para medir la velocidad de la luz. Midió la cantidad de tiempo requerido para la luz viajar una distancia de 7442 metros. Los datos son desviaciones de 24800 nanosegundos. Se supone un modelo normal (inapropiado). El principal objetivo es hacer inferencia sobre θ.

a) *Suponga que $p(\theta, \sigma^2) \propto \frac{1}{\sigma^2}$. Encuentre un intervalo posterior para θ.*

b) *Suponga que θ es conocida y que es igual a 25. Si se utiliza una a priori conjugada para σ^2 con $\alpha = 0.1$ y $\beta = 0.2$, encuentre un intervalo de credibilidad para σ^2 al 95 %*