Estadística Bayesiana

Clase 3: Modelos uniparamétricos: modelo Normal

Isabel Cristina Ramírez Guevara

Escuela de Estadística
Universidad Nacional de Colombia, Sede Medellín

Medellín, 11 de agosto de 2020

Modelo Normal

Estimando la media con varianza conocida. Considere la verosimilitud para una sola observación y:

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

La distribución a priori conjugada para θ es $\mathrm{N}(\mu_0, \tau_0^2)$, por lo tanto $p(\theta) \propto \exp\left(-\frac{1}{2\tau_0^2}(\theta-\mu_0)^2\right)$. Los hiperparámetros μ_0, τ_0^2 son conocidos. Con esto se obtiene la distribución posterior de θ,

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

$$p(\theta|y) \propto \exp\left(-\frac{1}{2\sigma^2}(y-\theta)^2\right)\exp\left(-\frac{1}{2\tau_0^2}(\theta-\mu_0)^2\right)$$

$$p(\theta|y) \propto \exp\left(-\frac{1}{2\sigma^2}(y-\theta)^2\right)\exp\left(-\frac{1}{2\tau_0^2}(\theta-\mu_0)^2\right)$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{(y-\theta)^2}{\sigma^2}+\frac{(\theta-\mu_0)^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left(-\frac{1}{2\sigma^2}(y-\theta)^2\right)\exp\left(-\frac{1}{2\tau_0^2}(\theta-\mu_0)^2\right)$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{(y-\theta)^2}{\sigma^2}+\frac{(\theta-\mu_0)^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{y^2-2y\theta+\theta^2}{\sigma^2}+\frac{\theta^2-2\theta\mu_0+\mu_0^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left(-\frac{1}{2\sigma^2}(y-\theta)^2\right)\exp\left(-\frac{1}{2\tau_0^2}(\theta-\mu_0)^2\right)$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{(y-\theta)^2}{\sigma^2}+\frac{(\theta-\mu_0)^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{y^2-2y\theta+\theta^2}{\sigma^2}+\frac{\theta^2-2\theta\mu_0+\mu_0^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{y^2}{\sigma^2}-\frac{2y\theta}{\sigma^2}+\frac{\theta^2}{\sigma^2}+\frac{\theta^2}{\tau_0^2}-\frac{2\theta\mu_0}{\tau_0^2}+\frac{\mu_0^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left(-\frac{1}{2\sigma^2}(y-\theta)^2\right)\exp\left(-\frac{1}{2\tau_0^2}(\theta-\mu_0)^2\right)$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{(y-\theta)^2}{\sigma^2}+\frac{(\theta-\mu_0)^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{y^2-2y\theta+\theta^2}{\sigma^2}+\frac{\theta^2-2\theta\mu_0+\mu_0^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{y^2}{\sigma^2}-\frac{2y\theta}{\sigma^2}+\frac{\theta^2}{\sigma^2}+\frac{\theta^2}{\tau_0^2}-\frac{2\theta\mu_0}{\tau_0^2}+\frac{\mu_0^2}{\tau_0^2}\right)\right]$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left[-2\theta\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\tau_0^2}\right)+\theta^2\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)\right]\right]$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)\left[\theta^2 - 2\theta\frac{\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\tau_0^2}\right)}{\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)}\right]\right]$$

Para completar el cuadrado sumamos y restamos $\left[\frac{\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\tau_0^2}\right)}{\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)}\right]^2$. Esta cantidad no depende de θ, por lo tanto:

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)\left[\theta^2 - 2\theta\frac{\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\tau_0^2}\right)}{\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)}\right]\right]$$

Para completar el cuadrado sumamos y restamos $\left[\frac{\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\tau_0^2}\right)}{\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)}\right]^2$. Esta cantidad no depende de θ, por lo tanto:

$$p(\theta|y) \propto \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)\left[\theta - \frac{\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\tau_0^2}\right)}{\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)}\right]^2\right]$$

Si $\frac{1}{\tau_1^2} = \left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)$ y $\mu_1 = \frac{\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\tau_0^2}\right)}{\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)}$ se tiene:

$$p(\theta|y) \propto \exp\left[-\frac{1}{2\tau_1^2}\left[\theta-\mu_1\right]^2\right],$$

$$p(\theta|y) \propto \exp\left[-\frac{1}{2\tau_1^2}\left[\theta-\mu_1\right]^2\right],$$

lo que implica que $\theta|y \sim \mathrm{N}(\mu_1, \tau_1^2)$.

Si Y∼N(θ,σ²), con σ^2 conocida, y θ ∼N(μ_0, τ_0^2) entonces $\theta|y \sim$ N(μ_1, τ_1^2) con $\mu_1 = \dfrac{\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\tau_0^2}\right)}{\left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)}$ *y* $\frac{1}{\tau_1^2} = \left(\frac{1}{\sigma^2}+\frac{1}{\tau_0^2}\right)$.

Podemos escribir $\mu_1 = \frac{y\tau_0^2+\mu_0\sigma^2}{\sigma^2+\tau_0^2}$. Como ya se ha visto en otros casos, la esperanza posterior μ_1, es un promedio ponderado entre la media a priori y la observación y con pesos proporcionales a las precisiones, es decir, al inverso de la varianza. Si $\tau_0^2 = 0$ la distribución a priori es más precisa que los datos y las distribuciones a priori y posterior están concentradas en μ_0. Si $\sigma^2 = 0$ los datos son precisos y la distribución posterior está concentrada en y.

Distribución predictiva posterior

Para encontrar la distribución predictiva posterior se procede por definición a realizar los siguientes cálculos:

$$
\begin{aligned}
p(\tilde{y}|y) &= \int_{\Theta} p(\tilde{y}|\theta)p(\theta|y)d\theta \\
&\propto \int_{\Theta} \exp\left(-\frac{1}{2\sigma^2}(\tilde{y}-\theta)^2\right)\exp\left(-\frac{1}{2\tau_1^2}(\theta-\mu_1)^2\right)d\theta,
\end{aligned}
$$

y se obtiene que $\tilde{y}|y \sim \text{Normal}(\mu_1, \sigma^2+\tau_1^2)$.

Modelo Normal con un vector de observaciones

Ahora suponga que se tiene $\mathbf{y} = (y_1, \cdots, y_n)$ un vector de observaciones independientes tal que $y_i \sim \mathrm{N}(\theta, \sigma^2)$, por lo tanto la verosimilitud es:

Modelo Normal con un vector de observaciones

Ahora suponga que se tiene $\mathbf{y} = (y_1, \cdots, y_n)$ un vector de observaciones independientes tal que $y_i \sim \mathrm{N}(\theta, \sigma^2)$, por lo tanto la verosimilitud es:

$$
\begin{aligned}
p(\mathbf{y}|\theta) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \theta)^2\right). \\
&\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right)
\end{aligned}
$$

Modelo Normal con un vector de observaciones

Ahora suponga que se tiene $\mathbf{y} = (y_1, \cdots, y_n)$ un vector de observaciones independientes tal que $y_i \sim \mathrm{N}(\theta, \sigma^2)$, por lo tanto la verosimilitud es:

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

$$\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right)$$

Si la distribución a priori es $\mathrm{N}(\mu_0, \tau_0^2)$, se tiene:

Modelo Normal con un vector de observaciones

Ahora suponga que se tiene $\mathbf{y} = (y_1, \cdots, y_n)$ un vector de observaciones independientes tal que $y_i \sim \mathrm{N}(\theta, \sigma^2)$, por lo tanto la verosimilitud es:

$$
\begin{aligned}
p(\mathbf{y}|\theta) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \theta)^2\right). \\
&\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right)
\end{aligned}
$$

Si la distribución a priori es $\mathrm{N}(\mu_0, \tau_0^2)$, se tiene:

$$
\begin{aligned}
p(\theta|\mathbf{y}) &\propto p(\mathbf{y}|\theta)p(\theta) \\
&\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right)\exp\left(-\frac{1}{2\tau_0^2}(\theta - \mu_0)^2\right)
\end{aligned}
$$

Modelo Normal con un vector de observaciones

Se puede probar que $p(\theta|\mathbf{y})=$ N$(\theta|\mu_n, \tau_n^2)$ donde

$$\frac{1}{\tau_n^2} = \frac{n}{\sigma^2} + \frac{1}{\tau_0^2} \qquad \mu_n = \frac{\frac{\bar{y}n}{\sigma^2} + \frac{\mu_0}{\tau_0^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau_0^2}}$$

Modelo Normal con un vector de observaciones

Se puede probar que $p(\theta|\mathbf{y}) = \mathrm{N}(\theta|\mu_n, \tau_n^2)$ donde

$$\frac{1}{\tau_n^2} = \frac{n}{\sigma^2} + \frac{1}{\tau_0^2} \qquad \mu_n = \frac{\frac{\bar{y}n}{\sigma^2} + \frac{\mu_0}{\tau_0^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau_0^2}}$$

También se puede probar que $\tilde{y}|y \sim \mathrm{Normal}(\mu_n, \tau_n^2 + \sigma^2)$

Ejemplo

Se toma una muestra aleatoria de n estudiantes y se les mide su peso, dando como resultado un peso promedio de 150 libras. Suponga que los pesos en la población están normalmente distribuidos con media θ desconocida y desviación estándar 20 libras. Suponga que la distribución a priori de θ es normal con media 180 y desviación estándar 40.

a) *Escriba la distribución posterior de θ (en función de n).*

Modelo Normal con un vector de observaciones

Se puede probar que $p(\theta|\mathbf{y})= \mathrm{N}(\theta|\mu_n, \tau_n^2)$ donde

$$\frac{1}{\tau_n^2} = \frac{n}{\sigma^2} + \frac{1}{\tau_0^2} \qquad \mu_n = \frac{\frac{\bar{y}n}{\sigma^2} + \frac{\mu_0}{\tau_0^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau_0^2}}$$

También se puede probar que $\tilde{y}|y \sim \mathrm{Normal}(\mu_n, \tau_n^2 + \sigma^2)$

Ejemplo

Se toma una muestra aleatoria de n estudiantes y se les mide su peso, dando como resultado un peso promedio de 150 libras. Suponga que los pesos en la población están normalmente distribuidos con media θ desconocida y desviación estándar 20 libras. Suponga que la distribución a priori de θ es normal con media 180 y desviación estándar 40.

a) *Escriba la distribución posterior de θ (en función de n).*

b) *Se toma el peso de un nuevo estudiante y se encuentra que pesa $\tilde{y}$ libras. Escriba la distribución predictiva posterior de $\tilde{y}$ (en función de n).*

Modelo Normal con un vector de observaciones

Se puede probar que $p(\theta|\mathbf{y})$= $\mathrm{N}(\theta|\mu_n, \tau_n^2)$ donde

$$\frac{1}{\tau_n^2} = \frac{n}{\sigma^2} + \frac{1}{\tau_0^2} \qquad \mu_n = \frac{\frac{\bar{y}n}{\sigma^2} + \frac{\mu_0}{\tau_0^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau_0^2}}$$

También se puede probar que $\tilde{y}|y \sim \mathrm{Normal}(\mu_n, \tau_n^2 + \sigma^2)$

Ejemplo

Se toma una muestra aleatoria de n estudiantes y se les mide su peso, dando como resultado un peso promedio de 150 libras. Suponga que los pesos en la población están normalmente distribuidos con media θ desconocida y desviación estándar 20 libras. Suponga que la distribución a priori de θ es normal con media 180 y desviación estándar 40.

a) *Escriba la distribución posterior de θ (en función de n).*

b) *Se toma el peso de un nuevo estudiante y se encuentra que pesa $\tilde{y}$ libras. Escriba la distribución predictiva posterior de $\tilde{y}$ (en función de n).*

c) *Si n=10 de un intervalo posterior al 95 % para θ y $\tilde{y}$.*

Estimando la varianza con media conocida

En este caso se establece que la distribución de muestreo es una distribución Normal con media conocida y varianza desconocida, esto es:

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

Estimando la varianza con media conocida

En este caso se establece que la distribución de muestreo es una distribución Normal con media conocida y varianza desconocida, esto es:

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

$$\begin{aligned} p(\mathbf{y}|\sigma^2) &\propto (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\theta)^2\right) \\ &= (\sigma^2)^{-n/2} \exp\left(-\frac{n}{2\sigma^2}\nu\right) \end{aligned}$$

donde $\nu = \frac{1}{n}\sum_{i=1}^{n}(y_i-\theta)^2$.

Estimando la varianza con media conocida

En este caso se establece que la distribución de muestreo es una distribución Normal con media conocida y varianza desconocida, esto es:

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

$$\begin{aligned} p(\mathbf{y}|\sigma^2) &\propto (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right) \\ &= (\sigma^2)^{-n/2} \exp\left(-\frac{n}{2\sigma^2}\nu\right) \end{aligned}$$

donde $\nu = \frac{1}{n}\sum_{i=1}^{n}(y_i - \theta)^2$.
La distribución a priori conjugada para σ^2 es la Gamma-inversa$\left(\frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2}\right)$ que es equivalente a decir que $\sigma^2 \sim$ Scale-Inv-$\chi^2(\nu_0, \sigma_0^2)$, o $\frac{\sigma^2}{\nu_0\sigma_0^2} \sim$ Inv-$\chi^2_{\nu_0}$, o $\sigma^2 \sim \frac{\nu_0\sigma_0^2}{\chi^2_{\nu_0}}$.

Estimando la varianza con media conocida

Por lo tanto:

$$
\begin{aligned}
p(\sigma^2) &= \frac{(\nu_0\sigma_0^2/2)^{\nu_0/2}}{\Gamma(\nu_0/2)}(\sigma^2)^{-(\nu_0/2+1)}\exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right] \\
&\propto (\sigma^2)^{-(\nu_0/2+1)}\exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right]
\end{aligned}
$$

Estimando la varianza con media conocida

Por lo tanto:

$$
\begin{aligned}
p(\sigma^2) &= \frac{(\nu_0\sigma_0^2/2)^{\nu_0/2}}{\Gamma(\nu_0/2)}(\sigma^2)^{-(\nu_0/2+1)}\exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right] \\
&\propto (\sigma^2)^{-(\nu_0/2+1)}\exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right]
\end{aligned}
$$

La distribución posterior es:

$$
\begin{aligned}
p(\sigma^2|\mathbf{y}) &\propto p(\mathbf{y}|\sigma^2)p(\sigma^2) \\
&\propto (\sigma^2)^{-n/2}\exp\left(-\frac{n}{2\sigma^2}\nu\right)(\sigma^2)^{-(\nu_0/2+1)}\exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right] \\
&= (\sigma^2)^{-\left(\frac{\nu_0+n}{2}+1\right)}\exp\left(-\frac{1}{2\sigma^2}(n\nu+\nu_0\sigma_0^2)\right)
\end{aligned}
$$

Estimando la varianza con media conocida

Por lo tanto:

$$
\begin{aligned}
p(\sigma^2) &= \frac{(\nu_0\sigma_0^2/2)^{\nu_0/2}}{\Gamma(\nu_0/2)}(\sigma^2)^{-(\nu_0/2+1)}\exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right] \\
&\propto (\sigma^2)^{-(\nu_0/2+1)}\exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right]
\end{aligned}
$$

La distribución posterior es:

$$
\begin{aligned}
p(\sigma^2|\mathbf{y}) &\propto p(\mathbf{y}|\sigma^2)p(\sigma^2) \\
&\propto (\sigma^2)^{-n/2}\exp\left(-\frac{n}{2\sigma^2}\nu\right)(\sigma^2)^{-(\nu_0/2+1)}\exp\left[-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right] \\
&= (\sigma^2)^{-\left(\frac{\nu_0+n}{2}+1\right)}\exp\left(-\frac{1}{2\sigma^2}(n\nu+\nu_0\sigma_0^2)\right)
\end{aligned}
$$

$$
\sigma^2|\mathbf{y} \sim \text{Gamma-inversa}\left(\frac{\nu_0+n}{2}, \frac{n\nu+\nu_0\sigma_0^2}{2}\right)
$$

$$
\sigma^2|\mathbf{y} \sim \text{Scale-Inv}-\chi^2\left(\nu_0+n, \frac{n\nu+\nu_0\sigma_0^2}{\nu_0+n}\right)
$$

Estimando la varianza con media conocida

Si $Y \sim$*Normal* (θ, σ^2) *siendo* θ *conocida la a priori conjugada para* σ^2 *es la Gamma-inversa*$\left(\frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2}\right)$, *por lo tanto*
$\sigma^2|\mathbf{y} \sim$ *Gamma-inversa* $\left(\frac{\nu_0+n}{2}, \frac{n\nu+\nu_0\sigma_0^2}{2}\right)$ *donde* $\nu = \frac{1}{n}\sum_{i=1}^{n}(y_i - \theta)^2$.

Ejemplo

Se supone que los precios de las acciones de una empresa (Y) se distribuyen normal con media θ *y varianza* σ^2 *desconocida. Se desea hacer inferencia sobre* σ *y con ese fin se toma una muestra aleatoria de tamaño 12. Se registran los siguientes precios: 212, 249, 250, 240, 210, 234, 195, 199, 222, 213, 233 y 251. Si* $\theta = 220$ *y la distribución a priori para* σ^2 *es Gamma-inversa(1100, 250000). Encuentre la distribución posterior de* σ^2.