

Modelos normales bayesianos conjugados

1. Introducción

¡Bienvenido! Esta aplicación permite mostrar el efecto de los datos observados y la distribución a priori sobre la distribución posterior de los parámetros del modelo. Este efecto se ilustra con el modelo normal conjugado. Recordemos que la perspectiva bayesiana cuantifica la creencia en la ocurrencia de un evento a partir de un concepto de probabilidad subjetiva. Para ello establecemos una distribución de probabilidad para las observaciones del suceso $p(\mathbf{y} | \theta)$, y una distribución de probabilidad para los parámetros $p(\theta)$, que se consideran variables aleatorias. Posteriormente, estas se utilizan en el teorema de Bayes con el fin de actualizar la creencia e inferir sobre los parámetros y las observaciones futuras a través de la distribución posterior que pertenece a una familia de distribuciones conocidas:

$$p(\theta | \mathbf{y}) \propto p(\mathbf{y} | \theta) p(\theta)$$

Esta aplicación utiliza los casos del modelo normal conjugado descritos a continuación.

2. Media desconocida: su estimación con varianza conocida.

2.1 Distribución de verosimilitud Supongamos que tenemos $\mathbf{y} = (y_1, \dots, y_n)$ un vector de observaciones independientes tal que $\mathbf{y} \stackrel{\text{i.i.d.}}{\sim} N(\theta, \sigma^2)$ con σ^2 , por lo tanto, la verosimilitud es:

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

2.2 Distribución a priori para θ La distribución a priori para θ está dada por $\theta \sim N(\mu_0, \tau_0^2)$, cuyos hiperparámetros θ and τ_0^2 son conocidos. Luego,

$$p(\theta) = \frac{1}{\sqrt{2\pi\tau_0^2}} \times \exp\left(\frac{-1}{2\tau_0^2}(\theta - \mu_0)^2\right),$$

2.3 Distribución posterior para θ La distribución posterior para θ es $\theta | \mathbf{y} \sim N(\theta | \mu_n, \tau_n^2)$, con esto,

$$p(\theta | \mathbf{y}) \propto \exp\left(\frac{-1}{2\tau_n^2}(\theta - \mu_n)^2\right)$$

donde $\frac{1}{\tau_n^2} = \frac{1}{\sigma^2} + \frac{1}{\tau_0^2}$ y μ_n es la nueva media de θ , compuesta por la media ponderada entre la media a priori y la media muestral con pesos proporcionales a sus respectivas precisiones, $\mu_n = \frac{\sum_{i=1}^n y_i \tau_n^2 + \mu_0 \sigma^2}{\tau_n^2 + \sigma^2}$

2.4 Valores y parámetros conocidos A continuación se describe cada uno de los valores necesarios para el cálculo de las distribuciones:

- **Tamaño muestral (n)**
- **Varianza conocida (σ^2):** Es la variación de la población.

Ambos valores son necesarios para simular el conjunto de datos.

- **Media muestral (\bar{y}_n):** la media a priori se estima a partir de este valor, de la siguiente manera $\mu_0 = \bar{y}_n \pm d\sigma$

- **Número de desviaciones estándar (d):** Indica cuántas desviaciones estándar queremos que tenga la media a priori de la muestra. Así, puede establecerse en un rango de ± 3 desviaciones estándar con pasos de 0.5.
- **Precisión de la distribución a priori (τ_0):** Recordemos que la precisión es la inversa de la varianza, es decir, $\tau_0 = \frac{1}{\sigma^2}$.

3. Varianza desconocida: su estimación con media conocida

3.1 Verosimilitud En este caso, se establece que la distribución muestral para un vector de observaciones $\mathbf{y} = (y_1, \dots, y_n)$ es una distribución Normal con media conocida y varianza desconocida, $\mathbf{y} \stackrel{\text{i.i.d.}}{\sim} N(\theta, \sigma^2)$, es decir, la distribución de probabilidad viene dada por:

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

3.2 Distribución a priori para σ^2 La distribución a priori para σ^2 es $\sigma^2 \sim \text{Inverse-Gamma}\left(\frac{v_0}{2}, \frac{v_0\sigma_0^2}{2}\right)$, luego,

$$p(\sigma^2) = \frac{\left(\frac{v_0\sigma_0^2}{2}\right)^{\frac{v_0}{2}}}{\Gamma\left(\frac{v_0}{2}\right)} (\sigma^2)^{-\frac{v_0}{2}+1} \times \exp\left(-\frac{v_0\sigma_0^2}{\sigma^2}\right)$$

3.3 Distribución posterior para σ^2 La distribución posterior de theta viene dada por:

$\sigma^2 | \mathbf{y} \sim \text{Inverse-Gamma}\left(\frac{v_0+n}{2}, \frac{n\nu+v_0\sigma^2}{2}\right)$, así:

$$p(\sigma^2 | \mathbf{y}) \propto (\sigma^2)^{-\frac{n}{2}} \times \exp\left(\frac{-n}{2\sigma^2}\nu\right) (\sigma^2)^{-\frac{v_0}{2+1}} \times \exp\left(\frac{-v_0\sigma_0^2}{2\sigma^2}\right)$$

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

3.4 Valores y parámetros conocidos A continuación se describe cada uno de los valores necesarios para el cálculo de las distribuciones:

- **Tamaño muestral (n):**
- **Media conocida (μ):** Es la varianza poblacional.
- **Varianza muestral (s^2):** Variabilidad que queremos que tenga la distribución de verosimilitud.

Ambos valores son necesarios para simular el conjunto de datos:

- ν : Corresponde a las sumas al cuadrado de las diferencias entre los valores observados y la media conocida. Toma valores de 0,1 a 5, con paso de 0,1 y debe ser ingresada por el usuario.
- **Primer hiperparámetro (α):** $\frac{v_0}{2}$. Es el parámetro de escala de la distribución a priori para σ^2 .
- **Segundo hiperparámetro (β):** $\frac{v_0\sigma_0^2}{2}$. Es el parámetro de forma de la distribución a priori para σ^2 .

4. Media y varianza desconocidas con la distribución a priori de la media en función de la varianza.

4.1 Verosimilitud Este caso se refiere al modelo normal multiparamétrico. Sea $\mathbf{y} = (y_1, \dots, y_n)$ un vector de observaciones independientes que distribuye $\mathbf{y} \stackrel{\text{i.i.d.}}{\sim} N(\theta, \sigma^2)$ con θ y σ^2 desconocidas. Así, la verosimilitud está dada por:

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

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

4.2 Distribuciones a priori La distribución a priori para θ se asume que depende de σ^2 , $p(\theta \mid \sigma^2)$, mientras que la distribución a priori para σ^2 no depende de θ y puede ser escrita como $p(\sigma^2)$. Por lo tanto, este caso tiene dos distribuciones a priori:

- $\theta \mid \sigma^2 \sim N\left(\mu_0, \frac{\sigma^2}{\kappa_0}\right)$
- $\sigma^2 \sim \text{Inverse} - \text{Gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right)$

4.3 Distribuciones posteriores La distribución marginal posterior de θ es:

$$\sigma^2 \mid \mathbf{y} \sim \text{Inverse} - \text{Gamma}\left(\frac{\nu_n}{2}, \frac{\nu_n \sigma_n^2}{2}\right),$$

$$\text{donde } \nu_n = \nu_0 + n, \sigma_n^2 = \frac{\nu_0 \sigma_0^2 + (n-1)S^2 + \frac{n \kappa_0 (\bar{y}_n - \mu_0)^2}{n + \kappa_0}}{n + \nu_0}$$

Y la distribución marginal posterior de θ es:

$$\theta \mid \mathbf{y} \sim t_{n+\nu_0}\left(\mu_n, \frac{\sigma_n^2}{\kappa_0 + n}\right),$$

$$\text{donde } \mu_n = \frac{\mu_0 \kappa_0 + n \bar{y}_n}{n + \kappa_0}$$

4.4 Valores y parámetros conocidos A continuación se describe cada uno de los valores necesarios para el cálculo de las distribuciones:

- **Tamaño muestral** (n)
- **Media muestral** \bar{y}_n
- **Varianza muestral** (s^2): Variabilidad que queremos que tenga la distribución de verosimilitud.

Ambos valores son necesarios para simular el conjunto de datos.

- μ_0 : representa la media de la distribución a priori condicionada a la varianza σ^2 y es un parámetro libre.
- κ_0 : Es la creencia a priori que tenemos sobre el parámetro σ , este parámetro indica el grado de incertidumbre que tenemos sobre el parámetro θ , si hay mucha incertidumbre κ_0 toma valores pequeños, pero si hay suficiente conocimiento sobre θ , κ_0 toma valores grandes. Hay que tener en cuenta que κ_0 solo toma valores positivos.
- α_0 : Es el parámetro de forma de la distribución a priori de sigma. $\alpha_0 = \frac{\nu_0}{2}$
- β_0 : Es el parámetro de escala para la distribución a priori de sigma. $\beta_0 = \frac{\nu_0 \sigma_0^2}{2}$