Estadística Bayesiana

Clase 14: Factor de Bayes

Isabel Cristina Ramírez Guevara

Escuela de Estadística
Universidad Nacional de Colombia, Sede Medellín

Medellín, 5 de octubre de 2020

Factor de Bayes

Sean $M_1, M_2, \cdots, M_q$ un conjunto de modelos bajo consideración, tales que para cada modelo la verosimilitud de los datos viene dada por $p_i(y|\theta_i)$. Los θ_i son desconocidos y tienen dimensión k_i.

Factor de Bayes

Sean $M_1, M_2, \cdots, M_q$ un conjunto de modelos bajo consideración, tales que para cada modelo la verosimilitud de los datos viene dada por $p_i(y|\theta_i)$. Los θ_i son desconocidos y tienen dimensión k_i. Sean $p(M_i)$, $i = 1, \cdots, q$, las probabilidades a priori de que cada modelo sea cierto, y suponga que $\sum_{i=1}^{q} p(M_i) = 1$. Dado un vector de observaciones $\mathbf{y}$, se desea determinar cuál de los modelos tiene una mayor probabilidad posterior. Usando el Teorema de Bayes, se obtiene que

$$\begin{aligned} p(M_i|\mathbf{y}) &= \frac{p(\mathbf{y}|M_i)p(M_i)}{\sum_{j=1}^{q} p(\mathbf{y}|M_j)p(M_j)} \\ &= \frac{m_i(\mathbf{y})p(M_i)}{\sum_{j=1}^{q} m_j(\mathbf{y})p(M_j)} \end{aligned}$$

donde $m_i(\mathbf{y}) = \int p_i(\mathbf{y}|\theta_i)p(\theta_i)d\theta_i$ es la distribución marginal o predictiva del vector de datos $\mathbf{y}$ bajo el modelo i.

Factor de Bayes

Para comparar los modelos se puede calcular el cociente de sus probabilidades posteriores

$$\frac{p(M_i|\mathbf{y})}{p(M_j|\mathbf{y})} = \frac{m_i(\mathbf{y})}{m_j(\mathbf{y})}\frac{p(M_i)}{p(M_j)}$$

Factor de Bayes

Para comparar los modelos se puede calcular el cociente de sus probabilidades posteriores

$$\frac{p(M_i|\mathbf{y})}{p(M_j|\mathbf{y})} = \frac{m_i(\mathbf{y})}{m_j(\mathbf{y})}\frac{p(M_i)}{p(M_j)}$$

Es decir, el cociente de las probabilidades posteriores es el cociente de las probabilidades previas multiplicado por un factor que representa la actualización del conocimiento proporcionado por los datos.

Factor de Bayes

Para comparar los modelos se puede calcular el cociente de sus probabilidades posteriores

$$\frac{p(M_i|\mathbf{y})}{p(M_j|\mathbf{y})} = \frac{m_i(\mathbf{y})}{m_j(\mathbf{y})}\frac{p(M_i)}{p(M_j)}$$

Es decir, el cociente de las probabilidades posteriores es el cociente de las probabilidades previas multiplicado por un factor que representa la actualización del conocimiento proporcionado por los datos. Este factor se denomina factor de Bayes, y será denotado por B_{ij},

$$B_{ij} = \frac{m_i(\mathbf{y})}{m_j(\mathbf{y})} = \frac{\int p_i(\mathbf{y}|\theta_i)p(\theta_i)d\theta_i}{\int p_j(\mathbf{y}|\theta_j)p(\theta_j)d\theta_j}$$

Factor de Bayes

Para comparar los modelos se puede calcular el cociente de sus probabilidades posteriores

$$\frac{p(M_i|\mathbf{y})}{p(M_j|\mathbf{y})} = \frac{m_i(\mathbf{y})}{m_j(\mathbf{y})}\frac{p(M_i)}{p(M_j)}$$

Es decir, el cociente de las probabilidades posteriores es el cociente de las probabilidades previas multiplicado por un factor que representa la actualización del conocimiento proporcionado por los datos. Este factor se denomina factor de Bayes, y será denotado por B_{ij},

$$B_{ij} = \frac{m_i(\mathbf{y})}{m_j(\mathbf{y})} = \frac{\int p_i(\mathbf{y}|\theta_i)p(\theta_i)d\theta_i}{\int p_j(\mathbf{y}|\theta_j)p(\theta_j)d\theta_j}$$

Cuando no existen parámetros desconocidos en los modelos a comparar el factor de Bayes se reduce a la razón de verosimilitudes, coincidiendo así con la inferencia clásica.

Factor de Bayes

Jeffreys proporciona una escala de evidencia para interpretar los valores de un factor de Bayes

B_{10}	$2\log B_{10}$	Interpretación
Por debajo de 1	Negativo	Apoya al modelo 0
1-3	0-2	Evidencia débil a favor del modelo 1 (No suficiente para decidir)
3-20	2-6	Evidencia positiva a favor del modelo 1
20-150	6-10	Evidencia fuerte a favor del modelo 1
Mayor de 150	Mayor de 10	Evidencia decisiva a favor del modelo 1

Factor de Bayes

Ejemplo

Considere nuevamente el problema de determinar la probabilidad de que una mujer sea portadora del gen de la hemofilia ($\theta = 1$) o no ($\theta = 0$). En este caso hay dos modelos que compiten, M_1 : la mujer está afectada y M_2: la mujer no está afectada. Concluya sobre la selección de estos modelos utilizando el factor de Bayes.

Factor de Bayes

Ejemplo

Considere nuevamente el problema de determinar la probabilidad de que una mujer sea portadora del gen de la hemofilia ($\theta = 1$) o no ($\theta = 0$). En este caso hay dos modelos que compiten, M_1 : la mujer está afectada y M_2: la mujer no está afectada. Concluya sobre la selección de estos modelos utilizando el factor de Bayes.

Ejemplo

Se quiere ver si la distribución del genero es equitativa en una población de venados. Con este fin se observaron 28 venados y se encontraron 20 machos y 8 hembras. Sea Y la cantidad de machos y p la proporción de machos, de acuerdo con el objetivo se desea comparar los modelos $M_1 : p = 0.5$ y $M_2 : p \neq 0.5$. Suponga que $p \sim U(0,1)$.