

Ejercicio 1

Sea $X_i \sim Bi(k, p)$ cond. independientes $i = 1, 2, \dots, n$, pero ambos k (número de ensayos) y p son desconocidos. ¿Cómo haría inferencia bayesiana en este caso? Ejemplo: k individuos de cierta especie de mamífero, con la misma probabilidad p de ser vistos, son muestreados de manera independiente. Los datos son los siguientes: 4 3 1 6 6 6 5 5 5 1 ($n = 10$). ¿Qué puede decir sobre el total k de individuos?

Solución:

Tenemos que la verosimilitud de los datos es

$$\begin{aligned} f(x^n | k, p) &= \prod_{i=1}^n \binom{k}{x_i} p^{x_i} (1-p)^{k-x_i} \\ &= \left(\prod_{i=1}^n \binom{k}{x_i} \right) p^{\sum x_i} (1-p)^{kn - \sum x_i} \end{aligned}$$

Proponiendo las distribuciones a priori de los parámetros como

$$\begin{aligned} \pi(p) &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}, \\ \pi(k) &= \mathbb{1}_{\{X_{(n)}, X_{(n+1)}, \dots\}}(k) \end{aligned}$$

Notemos que la distribución a priori para p es una beta y la distribución para k es una priori impropia cuyo dominio son los naturales mayores o iguales que la estadística de orden superior $X_{(n)}$.

La posterior conjunta dada por el Teorema de Bayes es

$$\begin{aligned} \pi(k, p | x^n) &\propto \left(\prod_{i=1}^n \binom{k}{x_i} \right) p^{\sum x_i} (1-p)^{kn - \sum x_i} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1} \\ &\propto \left(\prod_{i=1}^n \binom{k}{x_i} \right) p^{\sum x_i + \alpha - 1} (1-p)^{kn - \sum x_i + \beta - 1} \end{aligned}$$

Podemos obtener la distribución marginal (salvo una constante) integrando respecto a p . Así

$$\pi(k | x^n) \propto \left(\prod_{i=1}^n \binom{k}{x_i} \right) \int_0^1 p^{\sum x_i + \alpha - 1} (1-p)^{kn - \sum x_i + \beta - 1} dp$$

De la normalización de la distribución beta sabemos que

$$\frac{\Gamma(kn + \alpha + \beta)}{\Gamma(\sum x_i) \Gamma(kn - \sum x_i + \beta)} \int_0^1 p^{\sum x_i + \alpha - 1} (1-p)^{kn - \sum x_i + \beta - 1} dp = 1$$

Por tanto, la marginal tiene la siguiente forma funcional

$$\pi(k | x^n) \propto$$

Ejercicio 2

Partiendo del problema 1 suponga que ahora se plantea decidir el tamaño de muestra n , siendo que cada muestra nos cuesta w pesos.

Ejercicio 3

Supón que con una probabilidad de $1/10$ una señal está presente en cierto sistema en un momento dado y que con una probabilidad $9/10$ la señal no está presente. Una medición hecha en el sistema cuando la señal está presente tiene una distribución Normal con media 50 y varianza 1 y una medición cuando la señal no está presente tiene una distribución Normal con media 52 y varianza 1.

1. Supón que una medición X hecha en el sistema en un momento dado es igual a x . Muestre que la probabilidad posterior de que la señal esté en el sistema es mayor de que no esté, dado $X = x$, si $x < 51 - 0.5 \log 9$.
2. Dado que $X = 50$, demuestre que la esperanza predictiva de una nueva medición Y , independiente de la anterior, es aproximadamente 51.0932.

Solución:

1. Consideremos el evento hay señal como S , al evento no hay señal como NS . Luego $\mathbb{P}(S) = 1/10$ y $\mathbb{P}(NS) = 9/10$. Además

$$\begin{aligned} S &\Rightarrow X \sim N(50, 1) \\ NS &\Rightarrow X \sim N(52, 1) \end{aligned}$$

Calculemos las probabilidades (densidades) posteriores. Del Teorema de Bayes

$$\begin{aligned} \mathbb{P}(S|X=x) &= \frac{\mathbb{P}(X=x|S) \mathbb{P}(S)}{\mathbb{P}(X=x)}, \\ &= \frac{\mathbb{P}(X=x|S) \mathbb{P}(S)}{\mathbb{P}(X=x|S) \mathbb{P}(S) + \mathbb{P}(X=x|NS) \mathbb{P}(NS)} \\ &= \frac{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-50)^2}{2}\right) \cdot \frac{1}{10}}{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-50)^2}{2}\right) \cdot \frac{1}{10} + \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-52)^2}{2}\right) \cdot \frac{9}{10}} \\ &= \frac{\exp\left(-\frac{(x-50)^2}{2}\right)}{\exp\left(-\frac{(x-50)^2}{2}\right) + 9 \exp\left(-\frac{(x-52)^2}{2}\right)} \end{aligned} \quad (1)$$

De forma análoga calculamos la otra probabilidad. Sin embargo es más simple notando que

$$\mathbb{P}(S|X=x) + \mathbb{P}(NS|X=x) = 1.$$

Luego,

$$\mathbb{P}(NS|X=x) = \frac{9 \exp\left(-\frac{(x-52)^2}{2}\right)}{\exp\left(-\frac{(x-50)^2}{2}\right) + 9 \exp\left(-\frac{(x-52)^2}{2}\right)} \quad (2)$$

Busquemos que condiciones es necesaria tal que satisfice

$$\mathbb{P}(S|X=x) > \mathbb{P}(NS|X=x) \quad (3)$$

sustituyendo las ecuaciones (1) y (2) en (3) tenemos

$$\begin{aligned}\exp\left(-\frac{(x-50)^2}{2}\right) &> 9 \exp\left(-\frac{(x-52)^2}{2}\right) \\ -\frac{(x-50)^2}{2} &> \log 9 - \frac{(x-52)^2}{2} \\ \frac{(x-52)^2}{2} - \frac{(x-50)^2}{2} &> \log 9 \\ -4x + 204 &> 2 \log 9\end{aligned}\tag{4}$$

$$x < 51 - \frac{1}{2} \log 9\tag{5}$$

Así, la condición (5) es suficiente para satisfacer (3), lo que concluye la demostración.

2. Sea Y la observación independiente posterior a X . Para calcular la esperanza predictiva, calculamos primero el valor de la distribución predictiva. Requerimos

$$\begin{aligned}f(Y|X) &= \int f(Y|X, \theta) \pi(\theta|X) d\theta \\ &= \end{aligned}\tag{6}$$

En nuestro caso tenemos que la distribución posterior del parámetro es discreta ya que es la distribución de señal presente o señal ausente. De esta forma se reduce a calcular

$$\begin{aligned}f(Y|X) &= f(Y|S) \pi(S|X) + f(Y|NS) \pi(NS|X) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-50)^2}{2}\right) \cdot \pi(S|X) + \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-52)^2}{2}\right) \cdot \pi(NS|X)\end{aligned}\tag{7}$$

Observemos que $\pi(S|x) = \mathbb{P}(S|X=x)$ y $\pi(NS|x) = \mathbb{P}(NS|X=x)$.

La esperanza predictiva es entonces una suma ponderada por las probabilidades posteriores. De esta forma

$$\mathbb{E}[Y|X] = 50\pi(S|X) + 52\pi(NS|X)\tag{8}$$

$$= 50 \cdot 0.4508 + 52 \cdot 0.5491\tag{9}$$

$$= 51.0982.\tag{10}$$

lo que concluye la cuenta.

Ejercicio 4

A scientific journal, in an attempt to maintain experimental standards, insists that all reported statistical results have (classical) error probability of α_0 (or better). To consider a very simple model of this situation, assume that all statistical tests conducted are of the form $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, where θ_0 represents the standard and θ_1 the new proposal. Experimental results are reported in the journal only if the new proposal is verified with an error probability of $\alpha \leq \alpha_0$. (Nota that $\alpha = \mathbb{P}(\text{accepting } H_1)$). Let β denote the power of the test (i.e., $\beta = \mathbb{P}_{\theta_1}(\text{accepting } H_1)$). Assume further that α and β are fixed for all experiments conducted, with α being the specified value α_0 . Let π_0 denote the proportion of all experiments conducted in which θ_0 is correct, and π_1 denote the proportion of experiments in which θ_1 is correct.

1. Show that the proportion of articles published in the journal that have correct results (i.e., $\mathbb{P}(\theta = \theta_1 | \text{the test accepts } H_1)$) is $\pi_1 \beta / [\alpha_0 + \pi_1(\beta - \alpha_0)]$. (Note that many people naively believe that the journal is guaranteeing a proportion of $(1 - \alpha_0)$ of correct articles.)

2. Show that the proportion of correct published results is never less than π_1 . (Note that $\beta \geq \alpha$ for reasonable tests.)

Solución:

Observemos previamente que las proporciones de artículos asociados a θ_0 mas las proporciones de artículos asociados a θ_1 es uno, es decir, $\pi_0 + \pi_1 = 1$. Luego, del teorema de Bayes y la cantidad de interés se sigue

$$\mathbb{P}(\theta = \theta_1 | \text{test } H_1) = \frac{\mathbb{P}(\text{test } H_1 | \theta = \theta_1) \mathbb{P}(\theta = \theta_1)}{\mathbb{P}(\text{test } H_1)}$$

De la ley de probabilidad total tenemos

$$\mathbb{P}(\theta = \theta_1 | \text{test } H_1) = \frac{\mathbb{P}(\text{test } H_1 | \theta = \theta_1) \mathbb{P}(\theta = \theta_1)}{\mathbb{P}(\text{test } H_1 | \theta = \theta_1) \mathbb{P}(\theta = \theta_1) + \mathbb{P}(\text{test } H_1 | \theta = \theta_0) \mathbb{P}(\theta = \theta_0)}$$

Recordemos que la probabilidad de rechazar H_0 dado que el parámetro es θ_1 es la potencia. Entonces $\mathbb{P}(\text{test } H_1 | \theta = \theta_1) = \beta$. Así, con la notación establecida

$$\begin{aligned} \mathbb{P}(\theta = \theta_1 | \text{test } H_1) &= \frac{\beta \pi_1}{\beta \pi_1 + \alpha \pi_0} \\ &= \frac{\beta \pi_1}{\beta \pi_1 + \alpha(1 - \pi_1)} \\ &= \frac{\beta \pi_1}{(\beta - \alpha) \pi_1 + \alpha} \end{aligned}$$

obteniendo la expresión esperada.

Ejercicio 5

Suppose that $\mathbf{X} = (X_1, \dots, X_n)$ is a sample from a $\mathcal{NB}(m, \theta)$ distribution, and that θ has a $Beta(\alpha, \beta)$ prior distribution. Show that the posterior distribution of θ given \mathbf{x} is $Beta(\alpha + mn, (\sum_{i=1}^n x_i) + \beta)$.

Solución:

Recordemos la función de (masa de) probabilidad de la distribución binomial negativa. Para $X \in \{0, 1, \dots\}$

$$f(x|\theta) = \frac{\Gamma(m+x)}{\Gamma(x+1)\Gamma(m)} \theta^m (1-\theta)^x$$

Por tanto la verosimilitud de las observaciones se representa con la siguiente distribución conjunta

$$\begin{aligned} \mathcal{L}(\theta) &= f(\mathbf{x}^n | \theta) \\ &= \prod_{i=1}^n \frac{\Gamma(m+X_i)}{\Gamma(X_i+1)\Gamma(m)} \theta^m (1-\theta)^{X_i} \\ &= \left(\prod_{i=1}^n \frac{\Gamma(m+X_i)}{\Gamma(X_i+1)\Gamma(m)} \right) \theta^{mn} (1-\theta)^{\sum_{i=1}^n X_i} \end{aligned}$$

Como la distribución a priori es:

$$\pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

Entonces la distribución posterior sigue que

$$\pi(\theta | \mathbf{x}^n) \propto \left(\prod_{i=1}^n \frac{\Gamma(m+X_i)}{\Gamma(X_i+1)\Gamma(m)} \right) \theta^{mn} (1-\theta)^{\sum_{i=1}^n X_i} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

simplificando

$$\pi(\theta|x^n) \propto \theta^{nm+\alpha-1} (1-\theta)^{\sum X_i+\beta-1}$$

que es el kernel de la distribución beta. Así

$$\theta|X^n \sim \text{Beta}(nm + \alpha, \sum X_i + \beta)$$

lo que concluye el ejercicio.

Referencias

- [1] Robert, C. P., Casella, G., and Casella, G. (1999). Monte Carlo statistical methods (Vol. 2). New York: Springer.
- [2] Wasserman, L. (2004). All of statistics: a concise course in statistical inference (p. 413). New York: Springer.