Ejemplo

Se desea estimar la probabilidad, θ, de un evento, a partir del resultado de una sucesión de n ensayos Bernoulli. Por lo tanto $y_i = 1$ si se tiene un éxito o $y_i = 0$ si se obtiene un fracaso, con $i = 1, \cdots, n$. Encuentre la distribución posterior de θ si la distribución a priori es una $U(0,1)$

$$X = \sum_{i=1}^{n} y_i \qquad X \sim \text{Binomial}(n, \theta)$$

Verosimilitud:

$$p(X|\theta) = \binom{n}{x} \theta^{x} (1-\theta)^{n-x}$$

A priori: $U(0,1) \to \theta \sim U(\overset{a}{0}, \overset{b}{1})$

$$p(\theta) = 1$$

Posterior

$$p(\theta|X) \propto p(X|\theta)\, p(\theta)$$

$$\propto \theta^{x} (1-\theta)^{n-x} (1)$$

$$\propto \theta^{x} (1-\theta)^{n-x}$$

Distribución	$\text{Beta}(\alpha, \beta)$
Función de densidad	$p(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$

$$p(\theta \mid x) \propto \theta^{x+1-1} (1-\theta)^{n-x+1-1}$$

$$\therefore \quad \theta \mid x \sim \text{Beta}(x+1, n-x+1)$$

Ejemplo

Pepe y Pacho sospechan que una moneda en particular está cargada y que θ la probabilidad de cara es mayor de 0.5. Ellos proponen dos experimentos diferentes para probar estadísticamente esta sospecha. Pepe va a lanzar la moneda 10 veces y su estadístico va a ser el número de caras en esos 10 intentos. Pacho lanzará la moneda hasta obtener 2 sellos y su estadístico va a ser el número de caras que obtenga en los lanzamientos.

θ : probabilidad de cara

$H_0: \theta = 0.5$

$H_A: \theta > 0.5$

Pepe : experimento Binomial
$n = 10$

Pacho : experimento Binomial negativo

Ejemplo

a) *Pepe obtuvo 8 caras en 10 lanzamientos. Obtenga el p-value de esta prueba.*

b) *En el momento en que Pacho obtuvo los dos sellos, había acumulado 8 caras. Encuentre el p-value de esta prueba.*

c) *¿Pepe y Pacho obtienen las mismas conclusiones? Suponga un nivel de significancia de 0.05.*

d) *Caro es bayesiana y decide asumir una distribución a priori no informativa para θ de:*

$$p(\theta) = 1 \quad 0 \leq \theta \leq 1$$

Muestre que la inferencia Bayesiana basada en la distribución posterior de θ va a ser exactamente la misma si Caro utiliza los resultados de Pacho o los de Pepe. Discuta esto como se relaciona con el principio de verosimilitud.

a) p-value $\approx$ 0.0546875

b) p-value = 0.01953

c) Pepe no rechaza Ho, mientras que Pacho si lo rechaza.

d) Pepe → Binomial

Verosimilitud

$$p(x|\theta) = \binom{n}{x} \theta^{x} (1-\theta)^{n-x}$$

A priori: $p(\theta) = 1$

La posterior:

$$p(\theta|x) \propto p(x|\theta)\,p(\theta)$$

$$\propto \theta^{x} (1-\theta)^{n-x} \qquad n = 10 \quad x = 8$$

$$= \theta^{8} (1-\theta)^{2}$$

Pacho → Binomial Negativa

Verosimilitud: $p(x|\theta) = \binom{x-1}{k-1} \theta^{k} (1-\theta)^{x-k}$

k: resultados favorables

Posterior:

$$p(\theta|x) \propto \theta^{k} (1-\theta)^{x-k} \qquad k = 8, \quad x = 10$$

$$9-1 \leftarrow \quad = \theta^{8} (1-\theta)^{2} \rightarrow 3-1$$

$\therefore\ \theta|x \sim Beta(9, 3)$

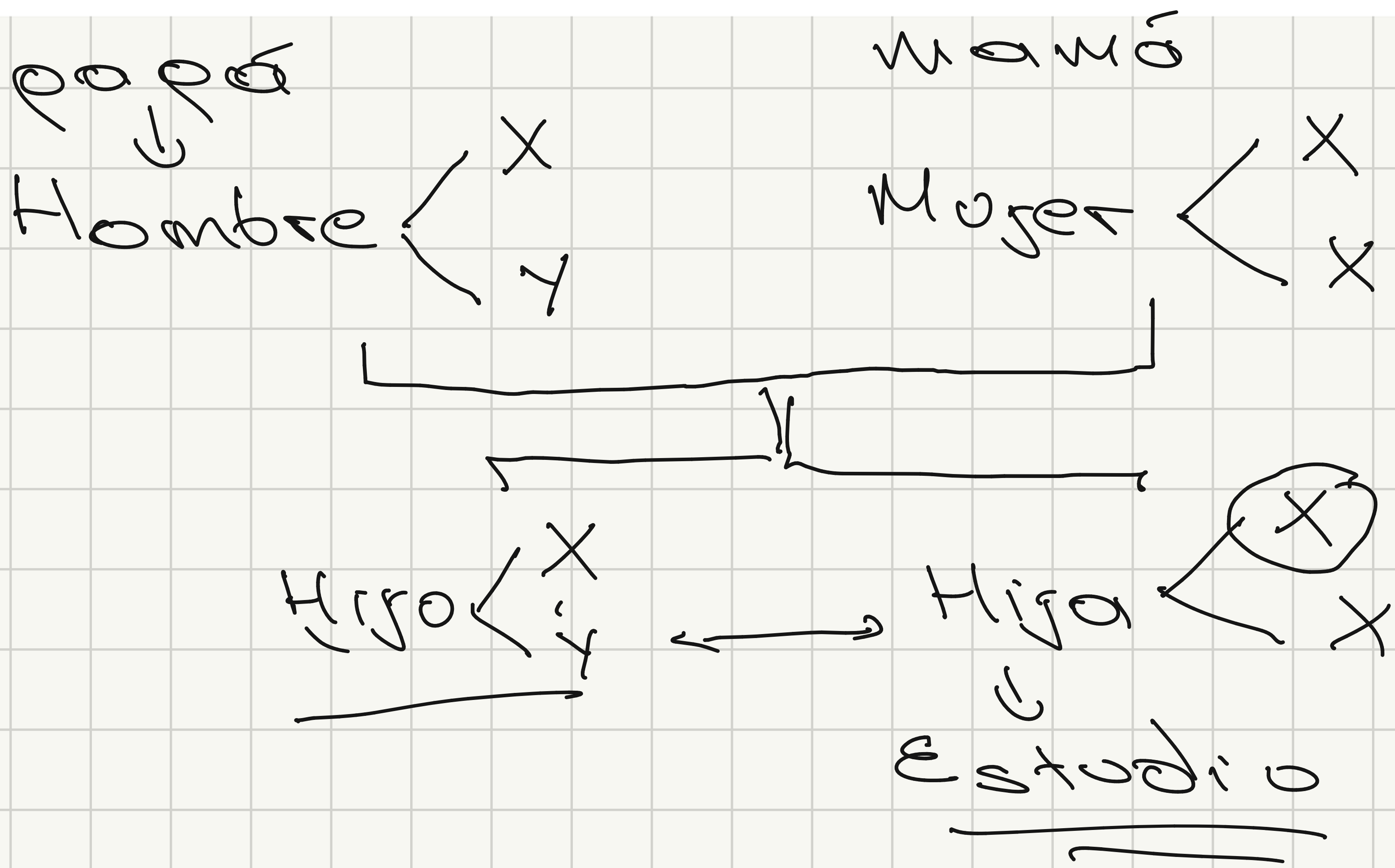
Ejemplo

La hemofilia es una enfermedad que muestra herencia recesiva ligada al cromosoma X, entonces un hombre que tiene el gen que causa la enfermedad en el cromosoma X es afectado, pero las mujeres que tienen el gen en uno de sus cromosomas X no es afectada. La enfermedad es fatal para la mujeres que heredan los dos genes y es un caso raro.

Suponga que una mujer tiene un hermano afectado por la enfermedad, por lo tanto la mujer tiene un 50 % de tener el gen "malo". Nos interesa el estado de la mujer. Los datos utilizados para determinar el estado de la mujer es la información de los dos hijos de la mujer los cuales están sanos.

i. *¿Cuál es la probabilidad posterior de que la mujer esté enferma?.*

ii. *La mujer tiene un tercer hijo sano, actualice la probabilidad encontrada en i.*



θ: La condición de la mujer

$\theta = 1$ si la mujer tiene el gen "malo"
$\theta = 0$ " " " no " " " "

$$\underbrace{p(\theta=1) = p(\theta=0) = 0.5}_{\text{Distribución a priori}}$$

Verosimilitud:

Sea y_i: el estado de los hijos

$y_i = 1$ si tienen el gen "malo"
$y_i = 0$ " NO " " " "

$i = 1, 2.$

$p(y_1 = 0, y_2 = 0 \mid \theta = 1)$

Independencia condicional $= p(y_1 = 0 \mid \theta = 1) \times p(y_2 = 0 \mid \theta = 1)$

$= 0.5 \times 0.5 = 0.25$

$p(y_1=0, y_2=0 \mid \theta=0)$

Independencia condicional $= p(y_1=0 \mid \theta=0)\, p(y_2=0 \mid \theta=0)$

$= (1)(1) = 1$

La distribución posterior de θ

$\mathbf{y} = (y_1, y_2)$

$$p(\theta=1 \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \theta)\, p(\theta)}{p(\mathbf{y})}$$

$$= \frac{0.25\,(0.5)}{p(\mathbf{y} \mid \theta=1)\, p(\theta=1) + p(\mathbf{y} \mid \theta=0)\, p(\theta=0)}$$

$$= \frac{0.25\,(0.5)}{0.25\,(0.5) + (1)(0.5)} = 0.20$$

ii) Nuestra a priori:

$p(\theta=1) = 0.20 \qquad p(\theta=0) = 0.8$