Estadística. Grupo m3

Hoja 6. Contrastes de razón de verosimilitudes y Bayesianos

Mayte Rodríguez

Ejercicio 1

Para una observación de la función de densidad $f_\theta(x) = \frac{2}{\theta^2}(\theta - x)$ si $0 < x \leq \theta(\theta > 0)$, encontrar el test de razón de verosimilitudes para contrastar $H_0 : \theta = \theta_0$ frente a $H_1 : \theta \neq \theta_0$, para un nivel de significación α.

La función de verosimilitud es

$$L(\theta) = \frac{2}{\theta^2}(\theta - x)I_{[x,\infty)}(\theta)$$

que alcanza un valor máximo en $\hat{\theta}_{MV} = 2x$.

Ejercicio 1

Entonces, el estadístico de contraste del método de la razón de verosimilitudes es

$$\lambda(x) = \frac{\max_{\theta=\theta_0} L(\theta)}{\max_{\theta\in\Theta} L(\theta)} = \frac{L(\theta_0)}{L(\hat{\theta}_{MV})} = \frac{\frac{2}{\theta_0^2}(\theta_0 - x)I_{[x,\infty)}(\theta_0)}{\frac{2}{4x^2}(2x - x)} = \frac{4x(\theta_0 - x)I_{(0,\theta_0]}(x)}{\theta_0^2}$$

La región de rechazo es

$$\begin{aligned} R &= \{x : \lambda(x) \leq c\} = \{x : x \leq k\} \cup \{x : x \geq \theta_0 - k\} \\ &= \{x : x \in (0, k] \cup [\theta_0 - k, \theta_0]\} \end{aligned}$$

Hay que hallar k para que el contraste tenga tamaño α.

Ejercicio 1

$$\begin{aligned}\alpha &= \sup_{\theta=\theta_0} P_\theta(R) = P_{\theta_0}(R) = P_{\theta_0}((X \leq k) \cup (X \geq \theta_0 - k)) \\ &= \int\limits_0^k \frac{2}{\theta_0^2}(\theta_0 - x)dx + \int\limits_{\theta_0-k}^{\theta_0} \frac{2}{\theta_0^2}(\theta_0 - x)dx = \frac{2k}{\theta_0}\end{aligned}$$

Por tanto $k = \frac{\alpha\theta_0}{2}$ y

$$R = \left(0, \frac{\alpha\theta_0}{2}\right] \cup \left[\theta_0 - \frac{\alpha\theta_0}{2}, \theta_0\right]$$

Ejercicio 2

Sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de una población $Exp(\theta)$. Se desea contrastar $H_0 : \theta \geq 1$ frente a $H_1 : \theta < 1$.

a) Si adoptamos como región de rechazo $R = \{(x_1, \dots, x_n) : x_{(1)} \geq c\}$, determinar el valor de c para que el test tenga tamaño α.

$$\alpha = \sup_{\theta \geq 1} P_\theta(R) = \sup_{\theta \geq 1} P_\theta(X_{(1)} \geq c) = \sup_{\theta \geq 1} \exp\{-n\theta c\} = \exp\{-nc\}$$

Entonces $c = -\frac{\ln \alpha}{n}$ y la región de rechazo es $R = \{(x_1, \dots, x_n) : x_{(1)} \geq -\frac{\ln \alpha}{n}\}$.

Ejercicio 2

b) Si tomamos como densidad a priori para θ, $\pi(\theta) \sim Exp(1)$, calcular la probabilidad a posteriori de la región que constituye la hipótesis nula, cuando la muestra consta de una única observación.

La distribución a posteriori es

$$\pi(\theta|x) \propto f_\theta(x)\pi(\theta) = \theta \exp\{-\theta x\} \exp\{-\theta\} = \theta \exp\{-\theta(x+1)\}$$

y entonces

$$\pi(\theta|x) \sim Gamma(a = x+1, p = 2).$$

Ejercicio 2

La probabilidad a posteriori de la hipótesis nula es

$$P(\theta \geq 1|x) = \int_1^\infty \pi(\theta|x)d\theta = \int_1^\infty \frac{(x+1)^2}{\Gamma(2)}\theta \exp\{-\theta(x+1)\}d\theta = (x+2)\exp\{-(x+1)\}$$

Se rechaza H_0 cuando

$$P(\theta \geq 1|x) < P(\theta < 1|x)$$

La región de rechazo es

$$R = \{x : (x+2)\exp\{-(x+1)\} < 1/2\}$$

Ejercicio 3

Sea $(X_1, X_2, ..., X_n)$ una muestra aleatoria simple de un modelo $Exp(\theta)$, donde $\theta \in \{1, 2\}$. Contrastar las hipótesis $H_0 : \theta = 1$ frente a $H_1 : \theta = 2$ siendo la distribución a priori sobre el parámetro la dada por $\pi(\theta = 1) = 3/4$. Se ha observado una muestra aleatoria simple de tamaño 10 con los siguientes datos:

0.8	1.4	0.4	0.3	1.8	0.7	0.9	1.2	0.6	1.1

¿Debe rechazarse H_0?

Ejercicio 3

La distribución a posteriori es

$$\pi(\theta|x_1,\dots,x_n) \propto f_\theta(x_1,\dots,x_n)\pi(\theta) = \theta^n \exp\left\{-\theta\sum_{j=1}^n x_j\right\}\pi(\theta)$$

Entonces

$$\pi(\theta=1|x_1,\dots,x_n) \propto f_{\theta=1}(x_1,\dots,x_n)\pi(\theta=1) = \exp\left\{-\sum_{j=1}^n x_j\right\}\frac{3}{4}$$

y

$$\pi(\theta=2|x_1,\dots,x_n) \propto f_{\theta=2}(x_1,\dots,x_n)\pi(\theta=2) = 2^n \exp\left\{-2\sum_{j=1}^n x_j\right\}\frac{1}{4}$$

Ejercicio 3

Por tanto

$$\pi(\theta = 1|x_1, \ldots, x_n) = \frac{\exp\left\{-\sum_{j=1}^{n} x_j\right\}\frac{3}{4}}{\exp\left\{-\sum_{j=1}^{n} x_j\right\}\frac{3}{4} + 2^n \exp\left\{-2\sum_{j=1}^{n} x_j\right\}\frac{1}{4}} = \frac{3}{3 + 2^n \exp\left\{-\sum_{j=1}^{n} x_j\right\}}$$

Con los datos proporcionados $n = 10$, $\sum_{j=1}^{10} x_j = 9.2$ y la probabilidad de la hipótesis nula es

$$\pi(\theta = 1|x_1, \ldots, x_n) = \frac{3}{3 + 2^{10}\exp\{-9.2\}} = 0.9667 \not\leq 1/2$$

No se rechaza la hipótesis nula.

Ejercicio 4

Sea $(X_1, X_2, ..., X_n)$ una muestra aleatoria simple de $X \sim N\ (\mu, \sigma^2)$. Encontrar el test de razón de verosimilitudes para el constraste $H_0 : \mu \leq \mu_0$ frente a $H_1 : \mu > \mu_0$.

El estadístico de contraste del método de la razón de verosimilitudes es

$$\lambda(x_1, ..., x_n) = \frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)} = \frac{L(\hat{\theta}_0)}{L(\hat{\theta}_{MV})}$$

donde $\theta = (\mu, \sigma^2)$ y $\Theta_0 = \{\theta \in \Theta : \mu \leq \mu_0\}$.

Ejercicio 4

El estimador de máxima verosimilitud es $\hat{\theta}_{MV} = (\bar{x}, v^2) = \left(\frac{1}{n} \sum\limits_{j=1}^{n} x_j, \frac{1}{n} \sum\limits_{j=1}^{n} (x_j - \bar{x})^2 \right)$ y entonces

$$
\begin{aligned}
L(\hat{\theta}_{MV}) &= \left(\frac{1}{\sqrt{2\pi}} \right)^n \left(\frac{1}{n} \sum_{j=1}^{n} (x_j - \bar{x})^2 \right)^{-n/2} \exp\left\{ - \frac{n}{2 \sum\limits_{j=1}^{n} (x_j - \bar{x})^2} \sum_{j=1}^{n} (x_j - \bar{x})^2 \right\} \\
&= \left(\frac{1}{\sqrt{2\pi}} \right)^n \left(\frac{1}{n} \sum_{j=1}^{n} (x_j - \bar{x})^2 \right)^{-n/2} \exp\left\{ - \frac{n}{2} \right\}
\end{aligned}
$$

Ejercicio 4

Bajo $H_0 : \mu \leq \mu_0,$ el punto del espacio paramétrico donde alcanza el valor máximo de la función de verosimilitud es

$$\hat{\theta}_0 = \begin{cases} (\bar{x}, v^2) & \text{si} \quad \bar{x} \leq \mu_0 \\ \left(\mu_0, \frac{1}{n}\sum\limits_{j=1}^{n}(x_j - \mu_0)^2\right) & \text{si} \quad \bar{x} > \mu_0 \end{cases}$$

Entonces

$$L(\hat{\theta}_0) = L(\hat{\theta}_{MV}) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^2\right)^{-n/2} \exp\left\{-\frac{n}{2}\right\}$$

si $\bar{x} \leq \mu_0$

y

$$
\begin{aligned}
L(\hat{\theta}_0) &= \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{n}\sum_{j=1}^{n}(x_j-\mu_0)^2\right)^{-n/2} \exp\left\{-\frac{n}{2\sum_{j=1}^{n}(x_j-\mu_0)^2}\sum_{j=1}^{n}(x_j-\mu_0)^2\right\} \\
&= \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{n}\sum_{j=1}^{n}(x_j-\mu_0)^2\right)^{-n/2} \exp\left\{-\frac{n}{2}\right\}
\end{aligned}
$$

si $\bar{x} > \mu_0$.

Ejercicio 4

Entonces

$$\lambda(x_1,...,x_n)=\begin{cases} 1 & \text{si} \quad \bar{x}\leq\mu_0 \\ \dfrac{\left(\frac{1}{\sqrt{2\pi}}\right)^n\left(\frac{1}{n}\sum\limits_{j=1}^{n}(x_j-\mu_0)^2\right)^{-n/2}\exp\left\{-\frac{n}{2}\right\}}{\left(\frac{1}{\sqrt{2\pi}}\right)^n\left(\frac{1}{n}\sum\limits_{j=1}^{n}(x_j-\bar{x})^2\right)^{-n/2}\exp\left\{-\frac{n}{2}\right\}}=\left(\dfrac{\sum\limits_{j=1}^{n}(x_j-\bar{x})^2}{\sum\limits_{j=1}^{n}(x_j-\mu_0)^2}\right)^{n/2} & \text{si} \quad \bar{x}>\mu_0 \end{cases}$$

Ejercicio 4

Vamos a escribir el contraste en función de $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$

$$\frac{\sum\limits_{j=1}^{n}(x_j-\bar{x})^2}{\sum\limits_{j=1}^{n}(x_j-\mu_0)^2} = \frac{\sum\limits_{j=1}^{n}(x_j-\bar{x})^2}{\sum\limits_{j=1}^{n}(x_j-\bar{x}+\bar{x}-\mu_0)^2} = \frac{\sum\limits_{j=1}^{n}(x_j-\bar{x})^2}{\sum\limits_{j=1}^{n}(x_j-\bar{x})^2+n(\bar{x}-\mu_0)^2}$$

$$= \frac{1}{1+\frac{n(\bar{x}-\mu_0)^2}{\sum\limits_{j=1}^{n}(x_j-\bar{x})^2}} = \frac{1}{1+\frac{n}{n-1}\frac{(\bar{x}-\mu_0)^2}{s^2}} = \frac{1}{1+\frac{1}{n-1}\left(\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right)^2}$$

$$= \frac{1}{1+\frac{1}{n-1}t^2}$$

Ejercicio 4

Entonces

$$\lambda(x_1,...,x_n) = \frac{L(\hat{\theta}_0)}{L(\hat{\theta}_{MV})} = \begin{cases} 1 & \text{si} \quad \bar{x} \leq \mu_0 \\ \left(\frac{1}{1+\frac{1}{n-1}t^2}\right)^{n/2} & \text{si} \quad \bar{x} > \mu_0 \end{cases}$$

La región de rechazo es

$$R = \{(x_1,...,x_n) : \lambda(x_1,...,x_n) \leq c\} = \left\{(x_1,...,x_n) : t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \geq k\right\}$$

Ejercicio 4

Hay que elegir k de forma que el test tenga tamaño α

$$\begin{aligned}
\alpha &= \sup_{\theta\in\Theta_0} P_\theta(R) = \sup_{\mu\le\mu_0} P_\mu\left(\frac{\bar{X}-\mu_0}{S/\sqrt{n}} \ge k\right) \\
&= \sup_{\mu\le\mu_0} P_\mu\left(\frac{\bar{X}-\mu}{S/\sqrt{n}} + \frac{\mu-\mu_0}{S/\sqrt{n}} \ge k\right) = P_{\mu_0}\left(\frac{\bar{X}-\mu_0}{S/\sqrt{n}} \ge k\right)
\end{aligned}$$

y entonces $k = t_{n-1;\alpha}$ y la región de rechazo es

$$R = \left\{(x_1, ..., x_n) : t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \ge t_{n-1;\alpha}\right\}$$

Ejercicio 4

El p-valor del contraste es

$$p = \sup_{\mu \leq \mu_0} P_\mu \left(\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right) = P_{\mu_0} \left(\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right) = P\left(T \geq \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right)$$

donde $T \sim t_{n-1}$