# Clase 13

jueves, 4 de octubre de 2018 16:50

Definición 78 Un estimador por intervalo de un parámetro $\theta$ real es un par de funciones $L(x_1, \ldots, x_n)$ y $U(x_1, \ldots, x_n)$ de la muestra que satisface $L(x) \leq U(x)$ $\forall x \in \mathcal{X}$. Si $X = x$ es observado la inferencia $L(x) \leq \theta \leq U(x)$ se puede hacer. El intervalo $[L(X), U(X)]$ se llama intervalo de estimación

En general $L(x)$ y $U(x)$ son finitos pero pueden tomar valores infinitos $L(x) = -\infty$ o $U(x) = +\infty$

Definición 79: Para un intervalo de estimación $[L(X), U(X)]$ de un parámetro $\theta$ se llama probabilidad de covertura a la probabilidad de que el verdadero parámetro $\theta$ esté en el intervalo o sea $P_\theta(\theta \in [L(X), U(X)]) = P(\theta \in [L(X), U(X)] / \theta)$

Definición 80: Llamamos coeficiente de confianza de $[L(X), U(X)]$ a

$$\inf_\theta P_\theta(\theta \in [L(X), U(X)])$$

Los intervalos de estimación junto con su coeficiente de confianza suelen llamarse intervalos de confianza.

# Métodos para encontrar intervalos de estimación

a) Invertir un test

Hay una correspondencia entre los test de hipótesis y los intervalos de estimación. En general podemos decir que todo conjunto de confianza corresponde a un test y vice versa

Teorema 81: Para cada $\theta_0 \in \Theta$, sea $A(\theta_0)$ la región de aceptación de nivel $\alpha$ del test $H_0: \theta = \theta_0$

Para cada $x \in \mathcal{X}$ definamos el conjunto $C(x)$ en el espacio de parámetros dado por

$$C(x) = \{\theta_0 : x \in A(\theta_0)\}$$

Entonces el conjunto aleatorio $C(X)$ es un conjunto de confianza $1-\alpha$.

Inversamente sea $C(x)$ un conjunto de confianza $1-\alpha$

Para todo $\theta_0 \in \Theta$ definamos

$$A(\theta_0) = \{x : \theta_0 \in C(x)\}$$

Entonces $A(\theta_0)$ es la región de aceptación de nivel $\alpha$ del test $H_0 : \theta = \theta_0$

Dem: Para la primera parte como $A(\theta_0)$ es la región de aceptación de nivel $\alpha$ se cumple que

$$P_{\theta_0}(X \notin A(\theta_0)) \leq \alpha \Rightarrow P_{\theta_0}(X \in A(\theta_0)) \geq 1-\alpha$$

Dado que $\theta_0$ es arbitrario escribamos $\theta$ en lugar de $\theta_0$ entonces

$$P_\theta(\theta \in C(X)) = P_\theta(X \in A(\theta)) \geq 1-\alpha$$

o sea $C(X)$ es un conjunto de confianza de nivel $1-\alpha$

Para la segunda parte el error tipo I del test $H_0: \theta=\theta_0$ con región de aceptación $A(\theta_0)$ es

$$P_{\theta_0}(X \notin A(\theta_0)) = P_{\theta_0}(\theta_0 \notin C(X)) \leq \alpha$$

luego $\alpha$ es un nivel del test

Ejemplo: $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ con $\sigma^2$ conocido

La región crítica para $H_0: \mu=\mu_0$ es $\left\{x : \frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} > z_{1-\alpha/2}\right\}$

luego la región de aceptación es $\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2}$ o sea

$$\bar{x} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$ luego

el intervalo de confianza $1-\alpha$ es $\left[\bar{x} - z_{1-\alpha/2} \sigma/\sqrt{n}, \bar{x} + z_{1-\alpha/2} \sigma/\sqrt{n}\right]$

b) Cantidades Pivotales

Definición 82: Una variable aleatoria $Q(X,\theta) = Q(X_1, \dots, X_n, \theta)$ es una cantidad pivotal (pivot) si la distribución de $Q(X,\theta)$ es independiente de todos los valores de los parámetros o sea si $X \sim F(x|\theta)$ entonces $Q(X,\theta)$ tiene la misma distribución para todos los valores de $\theta$

La función $Q(x,\theta)$ cumple que para todo conjunto $\mathcal{A}$ $P_\theta(Q(X,\theta)\in\mathcal{A})$ no depende de $\theta$. La técnica para construir conjuntos de confianza a partir de pivots implica encontrar un pivot y un conjunto $\mathcal{A}$ así $\{\theta : Q(x,\theta)\in\mathcal{A}\}$ es un conjunto estimador de $\theta$.

Si $Q(X,\theta)$ es un pivot entonces para un valor $\alpha$ específico podemos encontrar números $a$ y $b$ que no dependen de $\theta$ que satisfagan

$$P_\theta(a\leq Q(X,\theta)\leq b)\geq 1-\alpha$$

$= P_\theta(Q(x,\theta)\in\mathcal{A})$ con $\mathcal{A}=[a,b]$

Entonces para cada $\theta_0\in\Theta$

$$A(\theta_0)=\{x : a\leq Q(x,\theta_0)\leq b\}$$

es la región de aceptación de nivel $\alpha$ del test $H_0: \theta=\theta_0$

Usando el teorema 81 podemos invertir estos test para obtener

$$C(x)=\{\theta_0 : a\leq Q(x,\theta_0)\leq b\}$$

y $C(X)$ es un conjunto de confianza $1-\alpha$ para $\theta$

Si $\theta\in\mathbb{R}$, y si para $x\in\mathcal{X}$, $Q(x,\theta)$ es una función monótona de $\theta$, entonces $C(x)$ será un intervalo. De hecho si $Q(x,\theta)$ es una función creciente de $\theta$, entonces $C(x)$ tiene la forma $L(x,a)\leq\theta\leq U(x,b)$

Si $Q(x,\theta)$ es decreciente entonces $C(x)$ tiene la forma $L(x,b)\leq\theta\leq U(x,a)$

Ejemplo $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ $\sigma^2$ conocido, iid

entonces $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ es un pivot

Para todo $a$, $P_\mu\left(-a \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq a\right) = P(-a \leq \overset{\sim N(0,1)}{Z} \leq a)$

Luego un intervalo de confianza es

$$\left\{\mu : \bar{x} - a\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + a\frac{\sigma}{\sqrt{n}}\right\}$$

c) Pivoteando una función de distribución continua

Teorema 83: Sea $T$ un estadístico con función de distribución continua $F_T(t|\theta)$. Sean $\alpha_1 + \alpha_2 = \alpha$ con $0 < \alpha < 1$ valores fijos.

Supongamos que para cada $t \in \mathcal{T}$ pueden definirse las funciones $\theta_L(t)$ y $\theta_U(t)$ como:

i) Si $F_T(t|\theta)$ es una función decreciente de $\theta$ para cada $t$ definimos $\theta_L(t)$ y $\theta_U(t)$ por

$$F_T(t|\theta_U(t)) = \alpha_1 \quad F_T(t|\theta_L(t)) = 1-\alpha_2$$

ii) Si $F_T(t)$ es una función creciente de $\theta$ para cada $t$ definimos $\theta_L(t)$ y $\theta_U(t)$ por

$$F_T(t|\theta_U(t)) = 1-\alpha_2 \quad F_T(t|\theta_L(t)) = \alpha_1$$

Entonces el intervalo aleatorio $[\theta_L(T), \theta_U(T)]$ es un intervalo de confianza para $\theta$ de probabilidad $1-\alpha$.

Dem: Se omite

En ausencia de mayor información es común elegir $\alpha_1 = \alpha_2 = \alpha/2$. En el caso de intervalos de un solo lado se suele elegir $\alpha_1$ o $\alpha_2$ igual a 0

En el caso i)

$$F_T(t/\theta_U(t)) = \alpha_1 \qquad F_T(t/\theta_L(t)) = 1-\alpha_2$$

pueden expresarse en términos de la función de densidad de T mediante

$$\int_{-\infty}^{t} f_T(u/\theta_U(t))\,du = \alpha_1$$

$$\int_{t}^{+\infty} f_T(u/\theta_L(t))\,du = \alpha_2$$

El caso ii) es similar (encontrar las ecuaciones)

Ej: (Exponencial) $X_1, \dots, X_n$ iid con $f(x/\mu) = e^{-(x-\mu)} \mathbb{1}_{[\mu,\infty)}(x)$

Se puede demostrar que $Y = \min\{X_1, \dots, X_n\}$ es un estadístico suficiente para $\mu$ con densidad

$$f_Y(y/\mu) = n\, e^{-n(y-\mu)} \mathbb{1}_{[\mu,\infty)}(y)$$

Resolvemos para $\alpha$ fijo $\alpha_1 = \alpha_2 = \alpha/2$

$$\int_{\mu_U(y)}^{y} n e^{-n(z-\mu_U(y))}\,dz = \alpha/2$$

$$\int_{y}^{+\infty} n e^{-n(z-\mu_L(y))}\,dz = \alpha/2$$

Calculando las integrales

$$1-e^{-n(y-\mu_U(y))}=\alpha/2$$

$$e^{-n(y-\mu_L(y))}=\alpha/2$$

que da soluciones

$$\mu_U(y)=y+\frac{1}{n}\log(1-\alpha/2)$$

$$\mu_L(y)=y+\frac{1}{n}\log(\alpha/2)$$

luego el intervalo de confianza $1-\alpha$ es:

$$C(Y)=\{\mu: Y+\frac{1}{n}\log(\alpha/2)\leq\mu\leq Y+\frac{1}{n}\log(1-\alpha/2)\}$$

d) Pivotando una función de distribución discreta

Teorema 84: Sea $T$ un estadístico discreto con función de distribución $F_T(t/\theta)=P(T\leq t|\theta)$

Sean $\alpha_1+\alpha_2=\alpha$ con $0<\alpha<1$ fijos.

Supongamos que para cada $t\in\mathcal{T}$, $\theta_L(t)$ y $\theta_U(t)$ definidas como sigue

i) Si $F_T(t/\theta)$ es decreciente como función de $\theta$ para cada $t$ definimos $\theta_L(t)$ y $\theta_U(t)$ por

$$P(T\leq t|\theta_U(t))=\alpha_1 \qquad P(T\geq t|\theta_L(t))=\alpha_2$$

ii) Si $F_T(t|\theta)$ es creciente como función de $\theta$ para cada $t$ definimos $\theta_L(t)$ y $\theta_U(t)$ por

$$P(T\geq t|\theta_U(t))=\alpha_1 \qquad P(T\leq t|\theta_L(t))=\alpha_2$$

Entonces el intervalo aleatorio $[\theta_L(T),\theta_U(T)]$ es un intervalo de confianza $1-\alpha$ de $\theta$.

Dem: Se omite