Clase 11

jueves, 20 de septiembre de 2018 12:20

Métodos de Evaluación de Test

Cuando se realiza una docimasia de hipótesis con

$$H_0 : \theta \in \Theta_0$$

$$H_1 : \theta \in \Theta_0^c$$

se pueden producir errores de decisión, de los cuales tenemos dos tipos

Si $\theta \in \Theta_0$ pero decidimos que rechazamos H_0 entonces estamos cometiendo un Error tipo I; en cambio si $\theta \in \Theta_0^c$ y aceptamos H_0 estamos cometiendo un error tipo II

Decision

Verdad	Acept. H_0	Rechazo H_0
Acept. H_0	decisión correcta	Error tipo I
Rechazo H_0	Error tipo II	decisión correcta

Supongamos que R es la región crítica para un test. Si $\theta \in \Theta_0$ se cometerá un error si $x \in R$, luego la probabilidad de error tipo I es $P_\theta(X \in R)$

Si $\theta \in \Theta_0^c$ la probabilidad de error tipo II es $P_\theta(X \in R^c)$ y

$$P_\theta(X \in R^c) = 1 - P_\theta(X \in R)$$

Luego $P_\theta(X \in R)$ contiene toda la información del test con región crítica R

Si $\theta \in \Theta_0 \Rightarrow P_\theta(X \in R)$ es la probabilidad de error tipo I

Si $\theta \in \Theta_0^c \Rightarrow P_\theta(X \in R)$ es 1 menos la probabilidad de error tipo II

Definición 64 La función de potencia de un test de hipótesis con región crítica R a la función

$$\beta(\theta) = P_\theta(X \in R)$$

La función de potencia ideal es aquella que vale 0 para todo $\theta \in \Theta_0$ y 1 para todo $\theta \in \Theta_0^c$. El problema es que excepto situaciones triviales esta función ideal no se obtiene. Cualitativamente un buen test es aquel que tiene función de potencia cercana a 0 para $\theta \in \Theta_0$ y cercana a 1 para $\theta \in \Theta_0^c$

Ejemplo $X_1, \ldots, X_n \sim N(\theta, \sigma^2)$ (σ^2 conocido). Es fácil ver que el test LRT de $H_0: \theta \leq \theta_0$ $H_1: \theta > \theta_0$ tiene región crítica $\frac{\bar{X}-\theta_0}{\sigma/\sqrt{n}} > c$

$$\beta(\theta) = P_\theta\left(\frac{\bar{X}-\theta_0}{\sigma/\sqrt{n}} > c\right) = P\left(\frac{\bar{X}-\theta}{\sigma/\sqrt{n}} > c + \frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right)$$

$$= P\left(Z > c + \frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right) \quad y \quad Z \sim N(0,1)$$

Se tiene que $\lim_{\theta \to -\infty} \beta(\theta) = 0$ $\lim_{\theta \to +\infty} \beta(\theta) = 1$

y si α es tal que $P(Z > c) = \alpha$ entonces $\beta(\theta_0) = \alpha$

En general $\beta(\theta)$ depende de n, por lo que a veces la función de potencia puede usarse para elegir el tamaño muestral n apropiado para un experimento en cuestión.

Para una muestra de tamaño fijo en general es imposible hacer que ambos errores tengan probabilidad arbitrariamente chica. Es común restringirse a test que controlen el error tipo 1 a un nivel específico. Dentro de esa familia de test se elige aquel que tenga mayor potencia, (o sea hacer la prob de error tipo II lo menor posible)

Definición 65: Para $0 \leq \alpha \leq 1$, un test con función de potencia $\beta(\theta)$ se llama test tamaño α si

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$$

Definición 66: Para $0 \leq \alpha \leq 1$, un test con función de

potencia $\beta(\theta)$ se llama test de nivel α si
$$\sup_{\theta\in\Theta_0}\beta(\theta)\leq\alpha$$

Es típico elegir $\alpha=0{,}01$, $\alpha=0{,}05$ y $\alpha=0{,}10$

Esta propiedad a veces nos permite determinar el test específico

Ejemplo: Vimos que para $X_1,\dots,X_n\sim N(\mu,1)$
$$H_0\ \mu=\theta_0$$
$$H_1\ \mu\neq\theta_0$$
el test LRT obtiene una región crítica
$$R=\left\{x: |\bar{x}-\theta_0|\geq\sqrt{\frac{-2\log c}{n}}\right\}$$

Sabemos que como $X_1,\dots,X_n\sim N(\mu,1)$ entonces $\bar{X}$ tiene distribución normal con media μ y varianza $1/n$

$\beta(\theta)=P_\theta(X\in R)$. Si consideramos que $\theta\in\Theta_0$ entonces es la probabilidad de error tipo I

En nuestro caso $\Theta_0=\{\theta_0\}$ entonces
$$\sup_{\theta\in\Theta_0}\beta(\theta)=\beta(\theta_0)=P_{\theta_0}(X\in R)=$$
$$=P_{\theta_0}\left(|\bar{X}-\theta_0|>\sqrt{\frac{-2\log c}{n}}\right)$$

Si queremos un test que cumpla $\sup_{\theta\in\Theta_0}\beta(\theta)=\alpha$ entonces debe ser
$$P_{\theta_0}\left(|\bar{x}-\theta_0|>\sqrt{\frac{-2\log c}{n}}\right)=\alpha$$

Como $\bar{X}\sim N(\theta_0,1/n)$ bajo H_0 entonces $\dfrac{\bar{X}-\theta_0}{1/\sqrt{n}}\sim N(0,1)$ luego

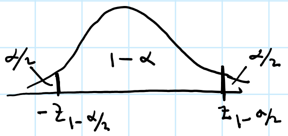
debe ser
$$P_{\theta_0}\left(|\bar{X}-\theta_0|>\sqrt{\frac{-2\log c}{n}}\right)=P_{\theta_0}\left(\sqrt{n}\,|\bar{x}-\theta_0|>\sqrt{-2\log c}\right)=\alpha$$

y debe cumplirse que $\sqrt{n}(\bar{X}-\theta_0)\in(-z_{1-\alpha/2},z_{1-\alpha/2})$

o sea $\sqrt{-2\log c}=z_{1-\alpha/2}$ entonces $c=e^{-\frac{z^2_{1-\alpha/2}}{2}}$

con $z_{1-\alpha/2}$ definido por $P(Z\leq z_{1-\alpha/2})=1-\alpha/2$

o sea $P(Z > z_{1-\alpha/2}) = \alpha/2$, y $Z \sim N(0,1)$



Definición 67: Un test con función de potencia $\beta(\theta)$ es insesgado si $\beta(\theta') \geq \beta(\theta'')$ para todo $\theta' \in \Theta_0^c$ y $\theta'' \in \Theta_0$

Definición 68: Sea $\mathcal{C}$ una familia de test para $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_0^c$

Un test en la familia $\mathcal{C}$, con función de potencia $\beta(\theta)$ es un test de la clase $\mathcal{C}$ de mayor potencia uniformemente (UMP) si $\beta(\theta) \geq \beta'(\theta)$ para todo $\theta \in \Theta_0^c$ y para toda $\beta'(\theta)$ que es una función de potencia de otro test en la familia $\mathcal{C}$

Nosotros trabajaremos con la familia $\mathcal{C}$ de los test de nivel α.

Teorema 69: (Lema de Neyman-Pearson)

Consideremos el test $H_0: \theta = \theta_0$ contra $H_1: \theta = \theta_1$, donde la función de probabilidad o de densidad de θ_i es $f(x|\theta_i)$; $i = 0, 1$. Supongamos que el test tiene región de rechazo R satisface

$$x \in R \text{ si } f(x/\theta_1) > k f(x/\theta_0)$$

$$\circledast$$

y

$$x \in R^c \text{ si } f(x/\theta_1) < k f(x/\theta_0)$$

para algún $k \geq 0$ y

$$\alpha = P_{\theta_0}(X \in R) \quad (**)$$

Entonces:

a) (Suficiencia) Todo test que satisface $\circledast$ y $(**)$ es un test UMP de nivel α.

b) (Necesidad) Si existe un test que satisface $\circledast$, $(**)$ con $k > 0$, entonces todo test UMP de nivel α es un test de tamaño α y todo test UMP de nivel α satisface $\circledast$ excepto tal vez en un conjunto A que satisface $P_{\theta_0}(X \in A) = P_{\theta_1}(X \in A) = 0$

Dem: Lo demostraremos en el caso de densidades

Observemos primero que como $\Theta_0=\{\theta_0\}$ todo test que satisface (**) es un test de nivel α y un test de tamaño α

Definamos para cada test con región crítica R
$$\phi(x)=\mathbb{1}_{\{x\in R\}} \quad \text{para } x\in\mathcal{X}$$

Sea ahora un test que cumple (*) y (**) y sea ϕ su función asociada. Sea ahora ϕ' la función asociada a otro test de nivel α y notemos por $\beta(\theta)$ y $\beta'(\theta)$ sus funciones de potencia.

Se cumple que:

$$(***)\ (\phi(x)-\phi'(x))(f(x/\theta_1)-k f(x/\theta_0))\geq 0 \ \forall x\in\mathcal{X}$$

pues si $x\in X$ es tal que $f(x/\theta_1)>kf(x/\theta_0)$ entonces $\phi(x)=1$ y como $0\leq\phi'(x)\leq 1$ entonces $(***)\geq 0$

si $x\in X$, tal que $f(x/\theta_1)<kf(x/\theta_0)$ entonces $\phi(x)=0$ entonces $(***)\geq 0$.

Luego
$$0\leq\int(\phi(x)-\phi'(x))(f(x/\theta_1)-kf(x/\theta_0))\,dx=$$

$$=\underbrace{\int\phi(x)f(x/\theta_1)\,dx}_{\int \mathbb{1}_{(x\in R)}f(x/\theta_1)dx=\int_R f(x/\theta_1)dx=P_{\theta_1}(x\in R)=\beta(\theta_1)}-\int\phi'(x)f(x/\theta_1)\,dx$$

$$-k\int\phi(x)f(x/\theta_0)\,dx+k\int\phi'(x)f(x/\theta_0)\,dx=$$

$$=\beta(\theta_1)-\beta'(\theta_1)-k\beta(\theta_0)+k\beta'(\theta_0)$$

o sea $0\leq\beta(\theta_1)-\beta'(\theta_1)-k(\beta(\theta_0)-\beta'(\theta_0))$

como ϕ' corresponde a un test de nivel α y ϕ a un test de tamaño α entonces
$$\beta(\theta_0)-\beta'(\theta_0)=\alpha-\beta'(\theta_0)\geq 0$$
como $k>0$ entonces debe ser
$$\beta(\theta_1)-\beta'(\theta_1)\geq 0 \quad \text{lo que}$$
implica que ϕ corresponde al test de nivel α UMP.
esto prueba a)

Para probar b) sea ϕ' una función que corresponde a un test de nivel α UMP. Sea un test que cumple (*) y (**) y ϕ la función indicatriz correspondiente. Por la parte a) este test también es un test de nivel α UMP o sea debe ser $\beta(\theta_1)=\beta'(\theta_1)$

Ya vimos que
$$\beta(\theta_1)-\beta'(\theta_1)-k(\beta(\theta_0)-\beta'(\theta_0))\geq 0$$
luego como $k>0$ y $\beta(\theta_1)=\beta'(\theta_1)$ debe ser:
$$\alpha-\beta'(\theta_0)=\beta(\theta_0)-\beta'(\theta_0)\leq 0$$
luego $\alpha\leq\beta'(\theta_0)$

Pero ya vimos que β' corresponde a un test de nivel α o sea $\beta'(\theta_0) \leq \alpha$ de donde debe ser $\beta'(\theta_0) = \alpha$ o sea es un test de tamaño α y la integral calculada vale 0.

Ahora la integral calculada anteriormente al ser el integrando no negativo, será cero si y solo si ϕ' satisface ⊛ excepto tal vez en un conjunto A con $\int_A f(x/\theta_i)\,dx = 0$ o sea

$$P_{\theta_1}(X \in A) = \int_A f(x/\theta_1)\,dx = 0 \quad y \quad P_{\theta_0}(X \in A) = \int_A f(x/\theta_0)\,dx = 0$$

□

Corolario 70: Supongamos se cumplen las hipótesis del Lema de N-P. Supongamos que $T(X)$ es un estadístico suficiente para θ y $g(t/\theta_i)$ la función de densidad o de probabilidad de T correspondiente a θ_i $i = 0,1$

Entonces todo test basado en T con región de rechazo S es un test de nivel α UMP si satisface

$$t \in S \quad si \quad g(t/\theta_1) > k\, g(t/\theta_0)$$

y

$$t \in S^c \quad si \quad g(t/\theta_1) < k\, g(t/\theta_0)$$

para algún $k \geq 0$ y donde

$$\alpha = P_{\theta_0}(T \in S)$$

obs S es un subconjunto de los posibles valores de T

Dem: A cargo del lector.

Las hipótesis H_0 y H_1 mostradas en el Lema de N-P especifican un solo valor del parámetro para cada una. Este tipo de hipótesis se denomina simples. Si las hipótesis especifican más de un valor para la distribución de la muestra se denominan hipótesis compuestas

A las hipótesis del estilo $H_0\ \theta \geq \theta_0$ $H_1\ \theta < \theta_0$ se denominan hipótesis de "una cola" o "unilaterales" en tanto las hipótesis $H_0\ \theta = \theta_0$ $H_1\ \theta \neq \theta_0$ se denominan de "dos colas" o "bilaterales"

Definición 71: Una familia de funciones de densidad o probabilidad $\{g(t/\theta) : \theta \in \Theta\}$ para una variable aleatoria univariada T con un parámetro real θ tiene razón de verosimilitud monótona (MLR) si para todo $\theta_2 > \theta_1$, $\dfrac{g(t|\theta_2)}{g(t|\theta_1)}$ es

monótona (creciente o decreciente) como función de t sobre

$\{t : g(t/\theta_1) > 0 \text{ o } g(t/\theta_2) > 0\}$.

Notemos que si $c > 0$, $c/0$ se define como ∞.

Teorema 72: (Karlin-Rubin)

Consideremos un test $H_0 : \theta \leq \theta_0$, $H_1 : \theta > \theta_0$

Supongamos que T es un estadístico suficiente para θ y la familia de funciones de densidad o de probabilidad $\{g(t/\theta) : \theta \in \Theta\}$ de T tiene MLR. Entonces para todo t_0 el test que rechaza H_0 si y sólo si $T > t_0$ es un test de nivel α UMP donde $\alpha = P_{\theta_0}(T > t_0)$.

Demostración: Sea $\beta(\theta) = P_\theta(T > t_0)$ la función de potencia del test.

Fijemos $\theta' > \theta_0$ y consideremos el test

$$H_0' \; \theta = \theta_0$$
$$H_1' \; \theta = \theta'$$

Se puede probar (no lo haremos) que como la familia $\{g(t/\theta)\}$ de T tiene un MLR entonces $\beta(\theta)$ es no decreciente así que

i) $\sup_{\theta \leq \theta_0} \beta(\theta) = \beta(\theta_0) = \alpha$ o sea se tiene un test de nivel α

ii) Si definimos $k' = \inf_{t \in Z} \dfrac{g(t/\theta')}{g(t/\theta_0)}$

donde $Z = \{t : t > t_0 \text{ y } g(t/\theta') > 0 \text{ o } g(t/\theta_0) > 0\}$

se sigue que

$$T > t_0 \iff \frac{g(t|\theta')}{g(t|\theta_0)} > k'$$

Junto con el corolario anterior i) y ii) implican que $\beta(\theta') \geq \beta^*(\theta')$ donde $\beta^*(\theta)$ es la función de potencia de cualquier otro test de nivel α de H_0' esto es, cualquier test que satisfaga $\beta^*(\theta_0) \leq \alpha$

Todo test de nivel α de H_0 satisface

$$\beta^*(\theta_0) \leq \sup_{\theta \in \Theta_0} \beta^*(\theta) \leq \alpha$$

Esto es $\beta(\theta') \geq \beta^*(\theta')$ para todo test de H_0 de nivel α.

Como θ' es arbitrario el test es un test UMP de nivel α