Test de hipó tesis

Ximena Fernández

Probabilidades y Estadística (Computació n)

7 de julio de 2020

Test de hipótesis

¿θ? $\rightarrow$ $\hat{\theta}$

$\rightarrow$ I_θ

$\rightarrow$ Hipótesis $\boxed{\theta = \theta_0}$

- Una hipótesis estadística es una afirmación que se hace sobre una o más características de una población.

- Un test o contraste de hipótesis es algún procedimiento para aceptar o rebatir dicha hipótesis o afirmación.

- La propuesta metodológica de Fisher era una especie de falsacionismo aplicado a la estadística: se trata de rechazar aquellas hipótesis para las cuales las observaciones sean relativamente inverosímiles.

Contraste de hipótesis

El contraste de hipótesis es un método para discernir entre dos hipótesis: la hipótesis nula H_0 y la hipótesis alternativa H_1. Conduce a aceptar o rechazar la hipótesis nula planteada (aceptando, en este último caso, la hipótesis alternativa). Pueden darse las siguientes situaciones:

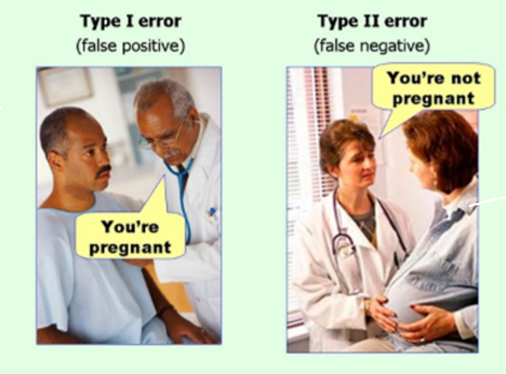
- Se acepta la hipótesis nula siendo verdadera.
- Se rechaza la hipótesis nula siendo falsa.
- Se rechaza la hipótesis nula siendo verdadera. Cometemos error de tipo I. La probabilidad de cometer este error viene dada por el nivel de significación α, fijado de antemano.
- Se acepta la hipótesis nula siendo falsa. Cometemos error de tipo II. La probabilidad de cometer este error se representa por β, y la probabilidad $1 - \beta$ se llama potencia del contraste, que cuantifica la probabilidad de rechazar la hipótesis nula cuando es falsa.

Errores

	Rechazo H_0	No rechazo H_0
H_0 falsa	:)	Error de tipo II
H_0 verdadera	Error de tipo I	:)

H_0: no emba-razadx

H_1: embarazadx



H_0: V
H_1: F

H_0: F
H_1: V

Pasos a seguir en un contraste de hipótesis

1. Especificar las hipótesis nula y alternativa. H_0 vs H_1
2. Tomar una muestra $X_1, X_2 \dots, X_n$.
3. Elegir un estadístico pivote $T(X_1, X_2 \dots, X_n)$ (con distribución conocida) para el realizar contraste (para medir la discrepancia entre lo observado y lo teórico)
4. Realiza una observación $(x_1, x_2, \dots, x_n)$ y evaluar el estadístico en la observación $T(x_1, x_2, \dots, x_n)$
5. Calcular el p-valor
   $p = P(T$ tenga un valor 'más extremo' que $T(x_1, x_2, \dots, x_n) | H_0)$.
   - $H_0 : \theta = \theta_0$ vs. $H_1 : \theta > \theta$,
     $$p = P(T \geq T(x_1, x_2, \dots, x_n))$$
   - $H_0 : \theta = \theta_0$ vs. $H_1 : \theta < \theta$,
     $$p = P(T \leq T(x_1, x_2, \dots, x_n))$$
   - $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta$,
     $$p = P(|T| \geq T(x_1, x_2, \dots, x_n)).$$

Pasos a seguir en un contraste de hipó tesis

6. Fijar un nivel de significació n α (probabilidad de error de tipo I) y realizar la siguiente comparació n:

   - si $p < \alpha$ se rechaza H_0
   - si $p \geq \alpha$ se acepta H_0

   Los valores de α con los que se suele trabajar son $0.1, 0.05$ o 0.01. El más habitual es $\alpha = 0.05$

- Para rechazar la hipó tesis nula se requiere que la probabilidad del fenó meno que se produce por casualidad (error de tipo I) sea inferior a α.
- No rechazar una hipó tesis no prueba que sea totalmente cierta. Podemos cometer un error de tipo II.

Regió n de rechazo y funció n de potencia

$$\mathcal{R} = \text{ regió n de rechazo de } H_0$$

La regió n de rechazo de nivel de de significatividad α es el conjunto de resultados de $T(x_1, x_2, \ldots x_n)$ para los cuales H_0 es rechazada en favor de H_1.

$$\pi(\theta) = \text{ funció n de potencia del test}$$

La funció n de potencia del test $\pi(\theta)$ es la probabilidad de rechazar H_0 cuando el valor verdadero del parámetro es θ. Es decir

$$\pi(\theta) = P_\theta(\mathcal{R}).$$

- si θ satisface H_0, $\pi(\theta) = P_\theta(\mathcal{R})$ es la probabilidad de un error tipo I.
- si θ satisface H_1, $1 - \pi(\theta) = P_\theta(\mathcal{R}^c)$ es la probabilidad de un error tipo II.

x = diámetro de la naranja

Ejercicio 1: El problema de las naranjas

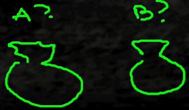
En un supermercado venden dos clases de naranjas, A y B. Las naranjas difieren en el diámetro, siendo en ambos casos de medias 25 y 30 cm, respectivamente. El diámetro sigue una distribución normal que, en ambos casos, tiene un desvío standard de 2 cm. Al llegar a casa, un señor ve que no le han puesto etiqueta a las bolsas, por lo que, en principio, no sabe cuál es cuál. Teniendo en cuenta que las naranjas de la clase A son para él y las de la clase B para el vecino, necesita poder diferenciarlas.

Después de pensar un rato, el señor decide realizar el siguiente proceso para clasificar las naranjas: agarra las de una bolsa (que son 20), y las mide. Aceptará que son del tipo B si la media muestral de las longitudes no supera los 28 cm.

Calcular las probabilidades de los errores que puede cometer el señor.

$$X_A \sim N(25, 2^2)$$

$$X_B \sim N(30, 2^2)$$



$H_0: \boxed{\mu = 30}$

$H_1: \boxed{\mu = 25}$

$X_A \sim N(25, 2^2)$

$X_B \sim N(30, 2^2)$

A ? B ?

$X_1, X_2 \ldots X_{20}$

$x_1, x_2 \ldots x_{20}$

observaciones (20 naranjas)

$T(X_1 \ldots X_{20}) = \bar{X}_{20}$

estadístico

Regla de aceptación:

$\bar{x}_{20} > 28 \longrightarrow$ acepto H_0 (tipo B)

promedio obs

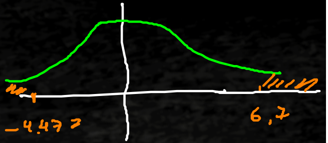
$\bar{x}_{20} \leq 28 \longrightarrow$ rechazo H_0 (Tipo A)

$R = (-\infty, 28]$

$R^c = (28, +\infty)$

$P(\text{error tipo I}) = P(\text{rechazar } H_0 \text{ siendo cierto})$

$= P(\underbrace{\overline{X}_{20} \le 28}\ |\ \underbrace{H_0 \text{ cierto}})$



$\boxed{H_0 \text{ cierto} \Rightarrow \mu = 30}$

$\overline{X}_{20} \sim N\left(30, \left(2/\sqrt{20}\right)^2\right)$

$P(\overline{X}_{20} \le 28 \mid H_0) = P\left(\frac{\overline{X}_{20} - 30}{2/\sqrt{20}} \le \overbrace{\frac{28-30}{2/\sqrt{20}}}^{-4.472}\right)$ $\simeq 0$

— × —

$\boxed{H_0 \text{ falso} \Rightarrow \mu = 25}$ y $\overline{X}_{20} \sim N(25, 4/20)$

$P(\text{error tipo II}) = P(\text{aceptar } H_0 \text{ siendo falso})$

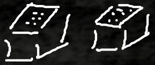
$= P(\overline{x}_{20} > 28 \mid H_0 \text{ falso}) = P(\overline{X}_{20} > 28 \mid \mu = 25)$

$= P\left(\overbrace{\frac{\overline{X}_{20} - 25}{2/\sqrt{20}}}^{N(0,1)} > \overbrace{\frac{28-25}{2/\sqrt{20}}}^{6.7}\right) \simeq 0$

Ejercicio 2: Bond, James Bond

Nos basamos en la película de James Bond "Octopussy" (1982). En el siguiente fragmento (https://vimeo.com/242488892), cuando Kamal Khan, el malo malísimo de la película, está jugando a los dados contra un millonario inglés en un casino de la India. Kamal apuesta a que es capaz de sacar un seis doble al lanzar dos dados. Dice que siempre lo logra, que todo es "cuestió n de muñeca". Desde un punto de vista estadístico, ¿có mo se debería proceder para saber si Kamal está haciendo trampa?

Vamos a suponer que James Bond observa 30 tiradas de los dos dados, y que en esas 30 tiradas el seis doble sale 3 veces.

$p = P(6,6)$ 

$\cancel{H_0}$: $p = \frac{1}{36} = 0.027 \longrightarrow$ inocente

H_1: $p > \frac{1}{36} \longleftarrow$ tramposo

$X_1 \; X_2 \; \dots \; X_{30}$ $\qquad X_i \sim B(p)$

$T = T(X_1, \dots, X_{30}) = \overline{X}_{30}$ (promedio)

$x_1, x_2, \dots, x_{30}$ observaciones $\rightsquigarrow \overline{x}_{30} = \frac{3}{30} = 0.1$

p-valor : $P(\overline{X}_{30} \geq 0.1 \mid H_0) = P\left(\frac{X_1 + X_2 + \cdots + X_{30}}{30} \geq 0.1 \mid H_0\right)$

$p = 1/36$

$P(\underbrace{X_1 + \cdots + X_{30}}_{\sim Bi(30, \frac{1}{36})} \geq 3 \mid H_0) = 1 - P(X_1 + \cdots + X_{30} < 3 \mid H_0)$

$= 1 - \left(\left(\frac{35}{36}\right)^{30} + \binom{30}{1}\left(\frac{35}{36}\right)^{29} \frac{1}{36} + \binom{30}{2}\left(\frac{35}{36}\right)^{28}\left(\frac{1}{36}\right)^{2} \right)$

$\approx 0.0073 < \alpha = 0.05$ (básico)

rechazo H_0

Veamos có mo resuelve James Bond es problema estadístico.
https://vimeo.com/263664023

X = peso de un paquete $\quad X \sim N(\mu, 2^2)$

Ejercicio 3: Lay's miente

Las fabricantes de papas fritas "Lies" aseguran que el peso de sus paquetes sigue una distribució n normal de media 50 gr y desvío 2 gr. El exceso de aire en los paquetes de papas fritas nos hacen sospechar y creemos que en realidad pesan menos que lo que afirman los fabricantes. Decidimos comprar 25 paquetes y pesarlos, obteniendo un promedio de 48.8 gr.

a) Plantear la hipó tesis a testear adecuada, junto con el estadístico del test. muestra $X_1 X_2 \cdots X_{25}$ $\quad T(X_1 \cdots X_{25}) = \overline{X}_{25}$ estadístico

observación $T(x_1, x_2 \cdots x_{25}) = \overline{x}_{25} = 48.8$

$H_0: \mu = 50 \qquad H_1: \mu < 50$

Extra:

¿Có mo quedarían las hipó tesis si sospecháramos que los fabricantes de "Lies" son generosos y ponen papas de más en sus paquetes? $\quad H_0: \mu = 50 \qquad H_1: \mu > 50$

¿Y si la sospecha fuera simplemente que la máquina que empaqueta está descalibrada y no produce paquetes con media 50 gr?

$H_0: \mu = 50 \qquad H_1: \mu \neq 50$

B) Calcular el p-valor. ¿Cuál es el máximo nivel de significación α para el que se aceptaría la hipótesis nula?

$$\text{p-valor} = P(\bar{X}_{25} \leq \underset{\bar{x}_{25}}{48.8} \mid H_0) =$$

$$\boxed{\begin{array}{l} \bar{X}_{25} \sim N\left(50, \left(\frac{2}{\sqrt{25}}\right)^2\right) \\ H_0: \mu = 50 \\ X \sim N(50, 4) \end{array}}$$

$$= P_{\mu=50}\Bigg(\underbrace{\frac{\bar{X}_{25} - 50}{2/\sqrt{25}}}_{Z \sim N(0,1)} \leq \underbrace{\frac{48.8 - 50}{2/\sqrt{25}}}_{-3} \Bigg)$$

$$= 0.00135$$

α = nivel de significancia

por ej.

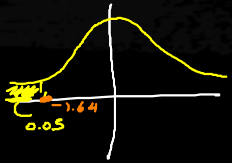
Si $\alpha = 0.05 \Rightarrow$ p-valor $= 0.00135 < \alpha = 0.05$

$\Rightarrow$ Rechazo H_0

Si $\alpha = 0.00135$ o menos $\Rightarrow$ aceptaría H_0

c) Calcular su región de rechazo a nivel 0.05. ¿Cuál es el mínimo valor del promedio muestral para el que se aceptaría la hipótesis nula?

¿Cuándo rechazo H_0?



$$P(\bar{X}_{25} < \bar{x}_{25} \mid H_0) < \overbrace{0.05}^{\alpha}$$

$$P\left(\underbrace{\frac{\bar{X}_{25}-50}{2/5}}_{Z\sim N(0,1)} < \underbrace{\frac{\bar{x}_{25}-50}{2/5}}_{-1.64}\right) < 0.05$$

$$\frac{\bar{x}_{25}-50}{2/5} < -1.64 \iff \bar{x}_{25} < \overbrace{-1.64\cdot\frac{2}{5}+50}^{49.344}$$

$R = (-\infty, 49.344)$

rechazo

$R^c = (49.344, +\infty)$

acepta

d) ¿Cuál es la probabilidad del error de tipo II si el verdadero valor de la media es 49?

Error tipo II :

P(acepto H_0 pero es falso)
$\mu = 49$

$$= P\left(\bar{X}_{25} > -1.64 \cdot \tfrac{2}{5} + 50 \mid \mu = 49\right)$$

$$= P\left(\frac{\bar{X}_{25} - 49}{2/5} > \frac{-1.64 \cdot 2/5 + 50 - 49}{2/5}\right) \quad X \sim N(49, 4)$$

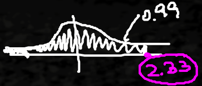
$$= P\Bigg(\underbrace{\frac{\bar{X}_{25} - 49}{2/5}}_{Z \sim N(0,1)} > \overbrace{-1.64 + \frac{50-49}{2/5}}^{0.86}\Bigg) = 0.195$$

$\mu = 49$

e) Calcular la función de potencia del test para una muestra de tamaño n. ¿Qué tamaño debe tener la muestra si se quiere que la potencia del test en este caso sea al menos 0.99?

Función de potencia

$$\pi(\mu) = P_\mu(R)$$



$$P_{\underset{=49}{\mu}}(R) = P\left(\bar{X}_n \le -1.64 \cdot 2/\sqrt{n} + 50\right)$$

$$P\Bigg(\underbrace{\frac{\bar{X}_n - \mu}{2/\sqrt{n}}}_{Z \sim N(0,1)} \le \overbrace{\frac{-1.64 \cdot 2/\sqrt{n} + 50 - \mu}{2/\sqrt{n}}}^{-1.64 + \frac{50-\mu}{2/\sqrt{n}}}\Bigg) \ge 0.99$$

$$-1.64 + \frac{50 - \overset{=49}{\mu}}{2/\sqrt{n}} \ge 2.33 \overset{\text{despejo } n}{\iff} n \ge 63.0436$$

$$\boxed{n \ge 64}$$