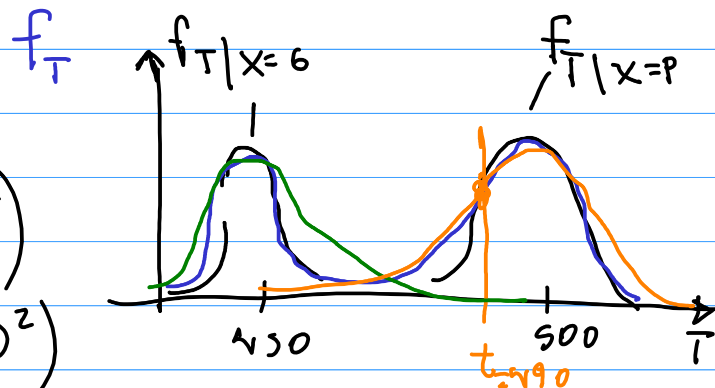
1. Sonia está trabajando en un proyecto de inteligencia artificial que utiliza redes neuronales para clasificar imágenes. Para entrenar el modelo, utiliza un conjunto de datos de imágenes etiquetadas con 2 clases: "Perro" y "Gato".

La distribución de los tamaños de las imágenes en el conjunto de datos sigue una distribución normal con una media de 500 píxeles y una desviación estándar de 50 píxeles para las imágenes de "Perro", y una media de 450 píxeles y una desviación estándar de 40 píxeles para las imágenes de "Gato".

Suponiendo que la cantidad de imágenes de clase ' Perro" y "Gato" está balanceada, si Sonia toma una nueva imagen y la misma tiene un tamaño de 490 píxeles. ¿Cuál es la probabilidad de que pertenezca a la clase "Perro"?

$$\boxed{P(X=P) = P(X=G) = 1/2}$$



$$(*) \begin{cases} T \mid X=P \sim N(500, 50^2) \\ T \mid X=G \sim N(450, 40^2) \end{cases}$$

Mi muestra aleatoria es T. (observé $t=490$)

$$P(X=P \mid T=t) = \frac{P(T=t \mid X=P)\cdot \cancel{P(X=P)}}{P(T=t\mid X=P)\cdot\cancel{P(X=P)} + P(T\mid X=G)\cancel{P(X=G)}}$$

$$= \frac{P(T=t \mid X=P)}{P(T=t\mid X=P) + P(T=t\mid X=G)} = \text{"}\frac{0}{0}\text{"}$$

por L'Hôpital

$$\overset{?}{=} \frac{f_{T|X=P}(t)}{f_{T|X=P}(t) + f_{T|X=G}(t)}$$

reemplazo la función de densidad (*) y evalúo en $t=490$

$$\approx \boxed{0,56}$$

Un desarrollador web está interesado en determinar si el tiempo promedio que los usuarios pasan en un sitio web que construyó ha aumentado en los últimos meses. Para ello, recopiló una muestra de $n=$ 50 usuarios y registró el tiempo que cada uno pasó en el sitio web durante su última visita, obteniendo una media muestral $\bar{x}=$ de 6.5 minutos y un desvío estándar muestral $s=$ de 1.5 minutos.

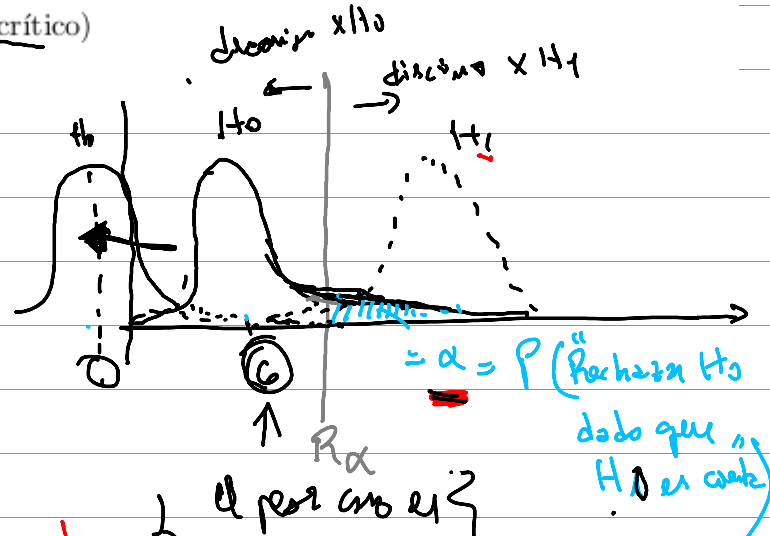
El desarrollador sabe que sabe que la media histórica del tiempo de permanencia en el sitio web ha sido $\mu=$ de 6 minutos. Se desea determinar con un nivel de significancia del 5 % si el tiempo promedio ha aumentado.

$\alpha = 0{,}05$

a) Escribir las hipótesis nula y alternativa

b) Calcular el test (Definir el estadístico a usar y el valor crítico)

c) ¿Cuál es la conclusión del desarrollador web?

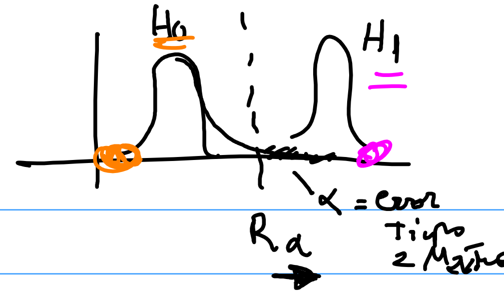


a) $\begin{cases} H_0: \mu \leq 6 \\ H_1: \mu > 6 \end{cases}$

"lo que queremos probar"

Un test con error del T1 acotado.
Un test con Prob. de falsos positivos acotada.

el peor caso es tener $\mu = 6$



U es decreciente en μ

$H_0: \mu \le 6$
$H_1: \mu > 6$ y $U = \frac{\bar{X} - \mu}{S}\sqrt{n} \sim t_{49}$

$$\delta(\underline{x}) = \mathbb{1}\{U > R_\alpha\}$$

= 1 y voy a Rechazar

= 0 y No voy a rechazar

función de la muestra

$$P_{H_0}(\delta(\underline{x}) = 1) \le \underline{P(\delta(\underline{x}) = 1 \mid \mu = 6)} \le \alpha = 0{,}05$$

acotamos por el peor caso

$$P(\delta(\underline{x}) = 1 \mid \mu = 6) = P(U > R_\alpha \mid \mu = 6)$$
$$= 1 - P(U \le R_\alpha \mid \mu = 6)$$
$$= 1 - F_{U|\mu=6}(R_\alpha) \le \alpha$$
$$F_{U|\mu=6}(R_\alpha) \geq 1 - \alpha$$

el mínimo error T2 se obtiene considerando

$$F_{U|\mu=6}(R_\alpha) = 1 - \alpha$$
$$R_\alpha = F^{-1}_{U|\mu=6}(1-\alpha) = F^{-1}_{U|\mu=6}(0{,}995)$$

como $U \sim t_{49}$, $R_\alpha \cong 1{,}67$

$$\delta(\underline{x}) = \mathbb{1}\{U > R_\alpha\} = \left\{ \frac{\overline{X} - 6}{S}\sqrt{n} > 1{,}67 \right\}$$

evaluando en la muestra, $\left(\overline{X} = 6{,}5 \,,\quad S = 1{,}5 \,,\quad n = 50 \right)$

$$\delta(\underline{x}) = \mathbb{1}\{2{,}35 > 1{,}67\} = 1$$

$\therefore$ Rechazo H_0.

3. Supongamos que estás trabajando en una empresa de tecnología que ha lanzado una nueva aplicación de recomendación de contenido. La aplicación utiliza un algoritmo de aprendizaje automático para recomendar posibles series y películas a los usuarios. La empresa cree que el algoritmo es muy efectivo y afirma que la probabilidad de que sus sugerencias sean correctas es del 70 %. "a priori"

Juan ha estado usando la aplicación durante los últimos meses y de las 7 series recomendadas ha visto 5. Él está interesado en saber qué tan efectiva es realmente la aplicación.

Utilizando el método de estimación Bayesiana, estimar la probabilidad de que de las siguientes 10 recomendaciones Juan acepte exactamente 5.

a) Definir una distribución a priori adecuada para la probabilidad de que una recomendación sea aceptada.

b) Calcular la distribución a posteriori para la probabilidad de que una recomendación sea aceptada.

c) A partir de esta distribución a posteriori hallar la probabilidad deseada

Y = "cuantas veo de las que recomienda", $Y \sim Bin$

La m.a. es $\underline{Y} \sim Bin(7, p)$ // $\underline{X_1 \dots X_7 \sim Ber(p)}$

(*) $f_{P|\underline{Y}=\underline{y}}(p) \propto \underbrace{P(\underline{Y}=\underline{y} \mid P=p)}_{\text{"Likelihood"}} \cdot \underbrace{f_P(p)}_{\pi(p)=?}$

Binomial / Bernoulli $\Rightarrow$ conjugada

conocemos

a) $P \sim Beta(a, b)$ x el enunciado, $\mathbb{E}P = 0,7$

$\mathbb{E}P = \dfrac{a}{a+b} = 0,7$

Consideramos $a = 7$ y $b = 3$

b) (*) $P(P=p \mid \underline{Y}=\underline{y}) \propto \underbrace{p^5 (1-p)^2}_{\text{"likelihood"}} \times \underbrace{p^{a-1} (1-p)^{b-1}}_{Beta(a,b)}$

$= p^{11} (1-p)^4$

$\therefore P|\underline{Y} \sim Beta(12, 5)$

$$W \sim Bin(10, \underset{\uparrow \text{v.a.}}{p})$$

Tiene densidad en $[0,1]$

c) $$\hat{P}(\underline{W=5}) = \int_{-\infty}^{+\infty} P(W=5 \mid p) \cdot \overbrace{\underbrace{f_{p|\underline{y}}(p)}_{\text{Beta}(12,5)}} \, dp$$

$$= \int_0^1 \binom{10}{5} p^5 (1-p)^5 \frac{\Gamma(17)}{\Gamma(12)\Gamma(5)} p^{11} (1-p)^4 \cdot dp$$

$$= \frac{\Gamma(17)}{\Gamma(12)\Gamma(5)} \binom{10}{5} \frac{1}{c} \int_0^1 \underbrace{c\, p^{16} (1-p)^9}_{} \, dp$$

es una densidad de una v.a. Beta(17, 10) $= 1$

donde $$c = \frac{\Gamma(27)}{\Gamma(17)\Gamma(10)}$$

$$= \binom{10}{5} \frac{\Gamma(17)\Gamma(10)}{\Gamma(27)} \cdot \frac{\Gamma(17)}{\Gamma(12)\Gamma(5)}$$

$$\simeq 0{,}1$$