Test basados en razones y/o p-valores

- El test rechaza H_0, si los datos son "extraños" a la muestra.

- El p-valor es interpretado como una medida de "sorpresa" al proponer H_0 para los datos observados.

- Según el abordaje Fisheriano, rechazar H_0, no implica "aceptar" H_1.

Ejercicio $X \sim$ Modelo, $n=1$

$\mathcal{P} = \{P_0, P_1\}$ — f_0 y f_1

$$H_0: f = f_0(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

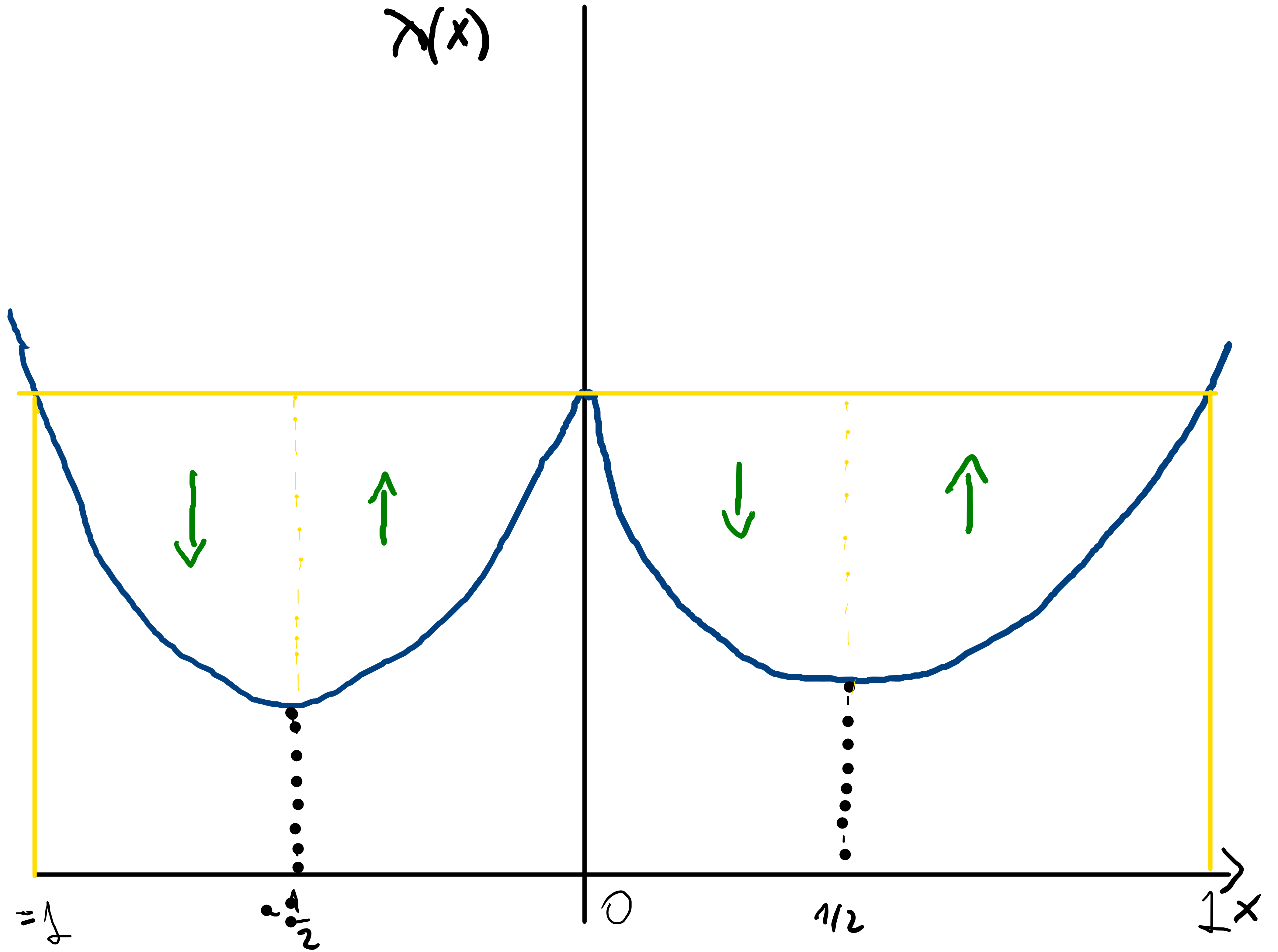
$$H_1: f = f_1(x) = 4^{-1} e^{-|x|/2}$$

$$T(x) = \begin{cases} 1, & \lambda(x) = f_1(x)/f_0(x) > c \\ 0, & c.c \end{cases}$$

$$\lambda(x) = \frac{\sqrt{2\pi}}{4} e^{\frac{x^2}{2} - \frac{|x|}{2}} > c$$

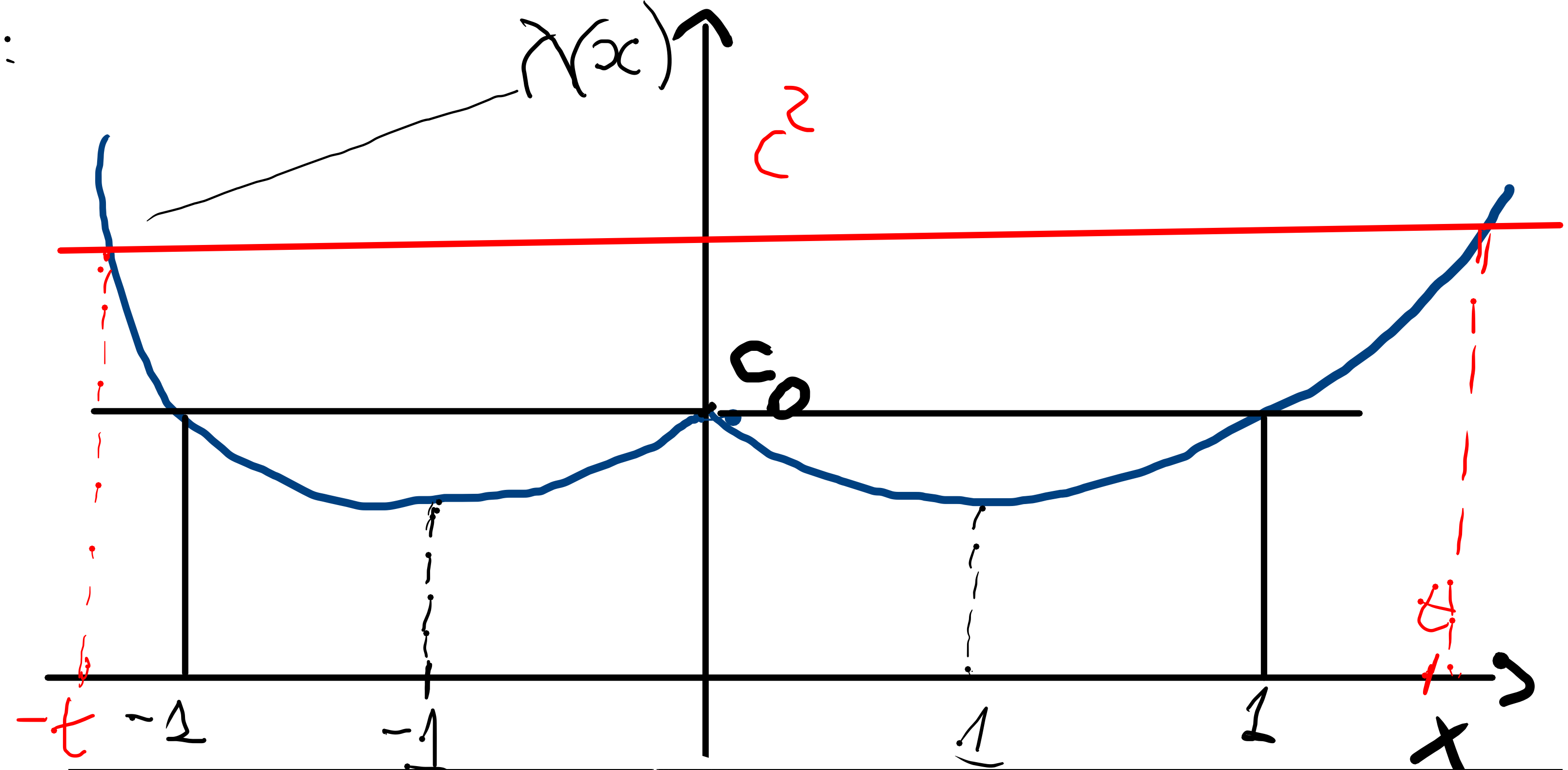
$$\Rightarrow \boxed{\frac{\pi}{8} e^{x^2 - |x|} > c^2}$$

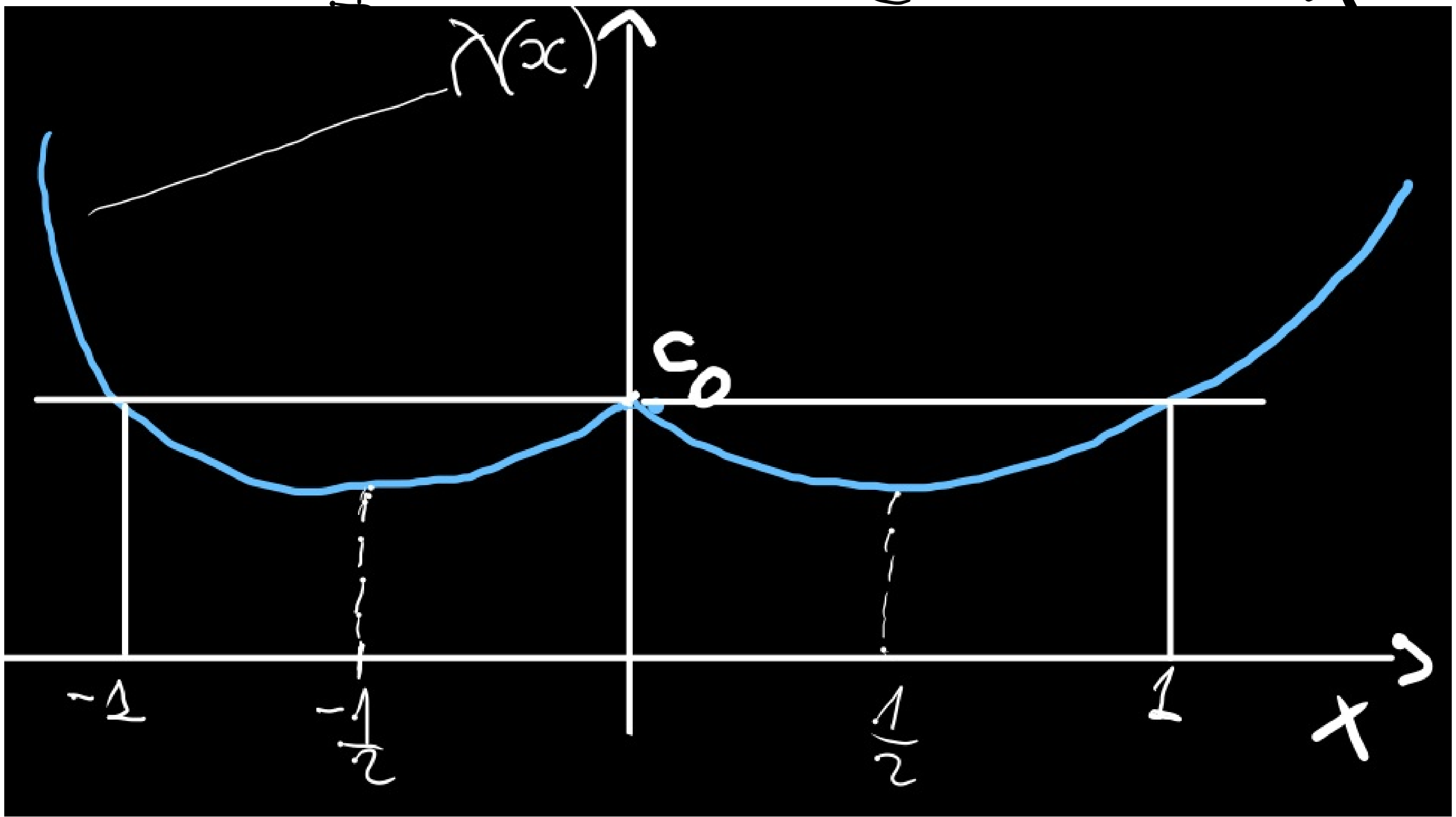
Caso 1: $|x| > t.$



λ(x)
-1
-1/2
0
1/2
1
x

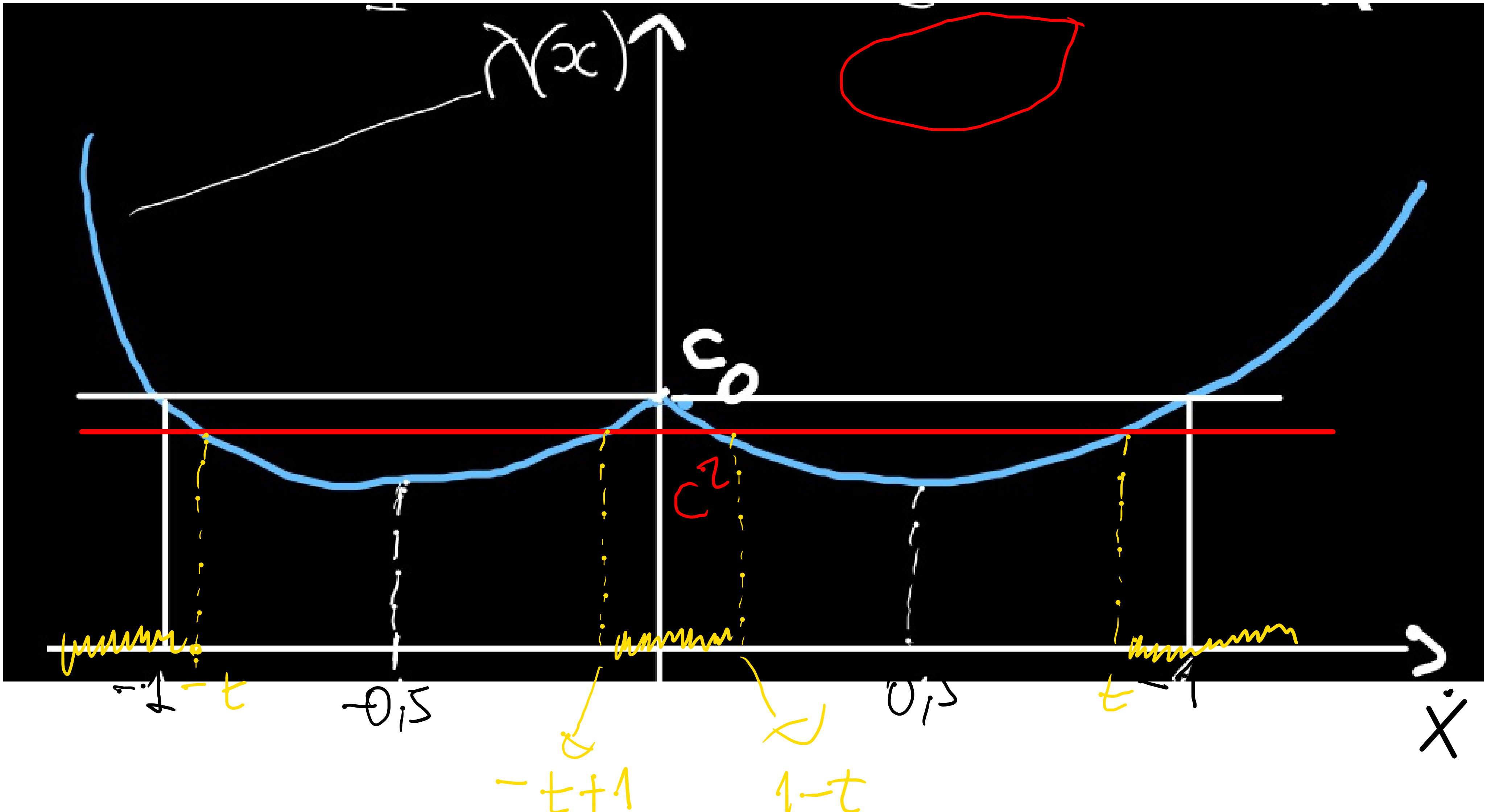
Caso 1:





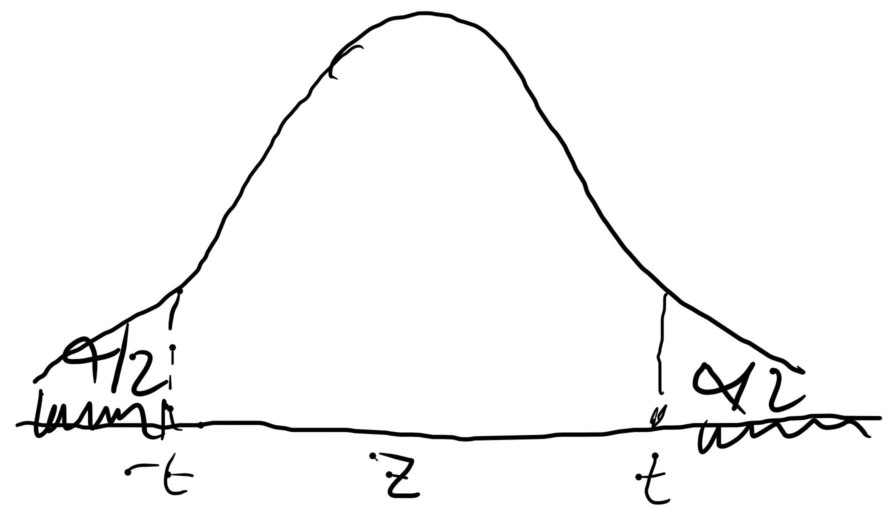
Caso 2:

$\lambda(x) > c^2$



$$T(x) = \begin{cases} 1, & |x| > t_\alpha \text{ or } |x| < 1 - t_\alpha \\ 0, & \text{c.c} \end{cases}$$

$$P_0(T(x) = 1) \leq \alpha$$



caso $t \leq 1$

$$\alpha = 1 - \Phi(t) + \Phi(-t) = 1 - (\Phi(t) - \Phi(-t))$$
$$= 1 - P(-t < Z < t)$$
$$= P(-t < Z < t)^C$$

$$t_\alpha = \Phi^{-1}(1 - \alpha/2)$$

Resumen del ejemplo:

$$X_i \overset{iid}{\sim} N(\mu, 1) \quad i = 1, \dots, n$$

$$H_0: \mu = 0 \ (\mu_0) \ ; \ H_1: \mu = 1 \ (\mu_1)$$

$$\lambda(x) = e^{\sum x_i - \frac{n}{2}} > c$$

$\downarrow$

Monótona (Estrictamente creciente)

- λ es monótona en x.
- $\mu_0 < \mu_1$ ''

Clase 9

Test con razones de verosimilitud monótona.

Ejemplo: $X_1, \ldots, X_n$, $X_i \sim \text{Normal}(\theta, \sigma^2)$ ✓

$x = (x_1, \ldots, x_n)$; $H_0: \theta = \theta_0$ y $H_1: \theta = \theta_1$ $(\theta_0 < \theta_1)$

$$\lambda(x) = \frac{f_1(x)}{f_0(x)} \checkmark \; ; \; f_\theta(x) = \prod f_\theta(x_i) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{\sum (x_i - \theta)^2}{2\sigma^2}}$$

$$= \frac{\cancel{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n} e^{-\frac{\sum (x_i - \theta_1)^2}{2\sigma^2}}}{\cancel{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n} e^{-\frac{\sum (x_i - \theta_0)^2}{2\cdot\sigma^2}}}$$

$$= e^{-\frac{\sum (x_i - \theta_1)^2}{2\sigma^2} + \frac{\sum (x_i - \theta_0)^2}{2\sigma^2}} \qquad \text{sumas} \qquad = e$$

$$\text{sumds} = \frac{-\sum (x_i^2 - 2x_i\theta_1 + \theta_1^2)}{2\sigma^2} + \frac{\sum (x_i^2 - 2x_i\theta_0 + \theta_0^2)}{2\sigma^2}$$

$$= \frac{-\cancel{\sum x_i^2} + 2\theta_1 \sum x_i - n\theta_1^2 + \cancel{\sum x_i^2} - 2\theta_0 \sum x_i + n\theta_0^2}{2\sigma^2}$$

$$= \frac{2\sum x_i(\theta_1 - \theta_0) + n(\theta_0^2 - \theta_1^2)}{2\sigma^2}$$

$c_1 = 2(\theta_1 - \theta_0) / 2\sigma^2, \ \theta_0 < \theta_1$

> 0

$c_2 = n(\theta_0^2 - \theta_1^2) / 2\sigma^2 \ < 0$

$$= c_1 \sum_{i=1}^{n} X_i + c_2$$

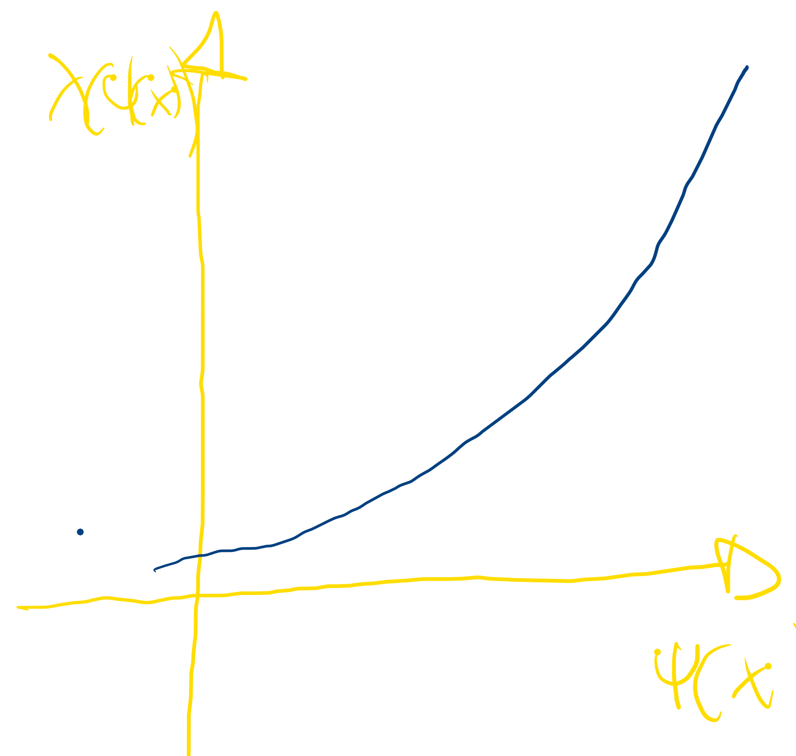
$$= c_2 + c_1 \Psi(X),$$

$$X = (X_1, \ldots, X_n)$$

pendiente es positiva

$$\Rightarrow \lambda(X) = e^{c_2 + c_1 \Psi(X)}$$

es monótona en $\Psi(X) = \sum_{i=1}^{n} X_i$.



Ejemplo: $X_1, \dots, X_n$, $X_i \overset{i.i.d}{\sim} \text{Poisson}(\theta)$

$p_\theta(x_i) = \dfrac{e^{-\theta}\,\theta^{x_i}}{x_i!}$. $H_0: \theta = \theta_0$ v/s $H_1: \theta = \theta_1$

$(\theta_0 < \theta_1)$

$x = (x_1, \dots, x_n)$

$$p_\theta(x) = \prod p_\theta(x_i) = \frac{e^{-n\theta}\,\theta^{\sum x_i}}{\prod x_i!}.$$

$$\lambda(x) = \frac{p_{\theta_1}(x)}{p_{\theta_0}(x)} = \frac{\dfrac{e^{-n\theta_1}\,\theta_1^{\sum x_i}}{\cancel{\prod x_i!}}}{\dfrac{e^{-n\theta_0}\,\theta_0^{\sum x_i}}{\cancel{\prod x_i!}}} = \frac{e^{-n\theta_1}\,\theta_1^{\sum x_i}}{e^{-n\theta_0}\,\theta_0^{\sum x_i}}.$$

$$= e^{n(\theta_0 - \theta_1)} \times (\theta_1/\theta_0)^{\sum x_i}$$

$$= e^{-n(\theta_0-\theta_1)} \times (\theta_1/\theta_0)^{\sum x_i}$$

$$\Theta = \{\theta : \theta > 0\}$$

$$c_1 = e^{-n(\theta_0-\theta_1)} > 0$$

$$c_2 = \theta_1/\theta_0 > 0$$

$$\theta_1 > \theta_0$$

$$\theta_0 < \theta_1 \quad | \quad \theta_0 \neq 0$$

$$1 < \frac{\theta_1}{\theta_0}$$

$$\overline{\hspace{10em}}^{\psi(x)}$$

$$\lambda(x) = c_1 \times c_2$$

es monótona en

$$\psi(x) = \sum_{i=1}^{n} X_i.$$

Función de verosimilitud

independencia

$X_i \overset{iid}{\sim} \text{Poisson}(\theta)$ $\quad P_\theta(x_1, x_2) = P_\theta(x_1)\, P_\theta(x_2)$ $\quad n = 2$

$x = (x_1, \ldots, x_n)$ $\quad n \in \mathbb{N}$

$$P_\theta(x) = P_\theta(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P_\theta(x_i)$$

$$= \prod_{i=1}^{n} \left\{ \frac{e^{-\theta}\, \theta^{x_i}}{x_i!} \right\} = \frac{\prod_{i=1}^{n} e^{-\theta} \times \prod_{i=1}^{n} \theta^{x_i}}{\prod x_i!}$$

$$= \frac{e^{-n\theta}\, \theta^{\sum x_i}}{\prod x_i!}$$

$P_{\theta_0}(X)$

Suposición sobre un fenómeno desconocido

↓ Transposición científica

H_0: θ alguna cosa

H_1: θ hay efecto

$\theta = (\beta_0, \beta_1, \beta_2, \ldots)$

$H_0 : P_\theta \in \mathcal{P}_0$

$H_1 : P_\theta \in \mathcal{P}_0$

$H_0 : \theta \in \Theta_0 \rightarrow \theta \leq 0$

$H_1 : \theta \in \Theta_1 \rightarrow \theta > 0$

Θ : mayúsculo

θ : minúsculo

$$T(x) = \begin{cases} 1, & \text{———} \\ \gamma, & \text{———} \\ 0, & \text{———} \end{cases}$$

Ejemplo: $X_i \overset{iid}{\sim} \text{Bernoulli}(p)$

$H_0: p = p_0$ v/s $H_1: p = p_1$ $(p_0 < p_1)$

$$P_\theta(x_i) = p^{x_i}(1-p)^{1-x_i}$$

$$P_\theta(x) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum x_i} \times (1-p)^{n-\sum x_i}$$

$$= \text{función}\left(T(x) = \sum x_i\right)$$

$$Y = \sum_{i=1}^{n} X_i$$

$$\lambda(x) = \frac{P_1(x)}{P_0(x)} = \frac{p_1^{\,y}(1-p_1)^{n-y}}{p_0^{\,y}(1-p_0)^{n-y}}$$

$p_0 < p_1$

$$= \left(\frac{p_1}{p_0}\right)^{y}\left(\frac{1-p_1}{1-p_0}\right)^{n-y}, \quad p_0, p_1, n$$

$$= \left(\frac{p_1}{p_0}\right)^{y}\left(\frac{1-p_1}{1-p_0}\right)^{-y}\left(\frac{1-p_1}{1-p_0}\right)^{n}$$

c_2 $\qquad c_1$

$Y(x) = \sum x_i$

$$c_1 = \left(\frac{1-p_1}{1-p_0}\right)^{n}$$

$$c_2 = \frac{p_1(1-p_0)}{p_0(1-p_1)}$$

$$= \left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)^{y} \cdot c_1 = c_2^{\,y} \cdot c_1$$

$>1 \qquad >0$

$$C_1 = \left(\frac{1-p_1}{1-p_0}\right)^n > 0. \qquad p \in (0,1)$$

$$C_2 = \overset{>1}{\frac{p_1}{p_0}} \times \overset{>1}{\frac{1-p_0}{1-p_1}} > 1$$

$$p_0 < p_1 \quad | \cdot (-1)$$

$$-p_0 > -p_1 \quad | +1$$

$$1-p_0 > 1-p_1 \quad | \; 1-p_0 > 0$$

$$\boxed{1 > \frac{1-p_1}{1-p_0} > 0.}$$

$$\frac{1-p_0}{1-p_1} > 1$$

$$p_0 < p_1$$

$$\frac{p_1}{p_0} > 0$$

Ejemplo: determinar c_α en el caso Normal

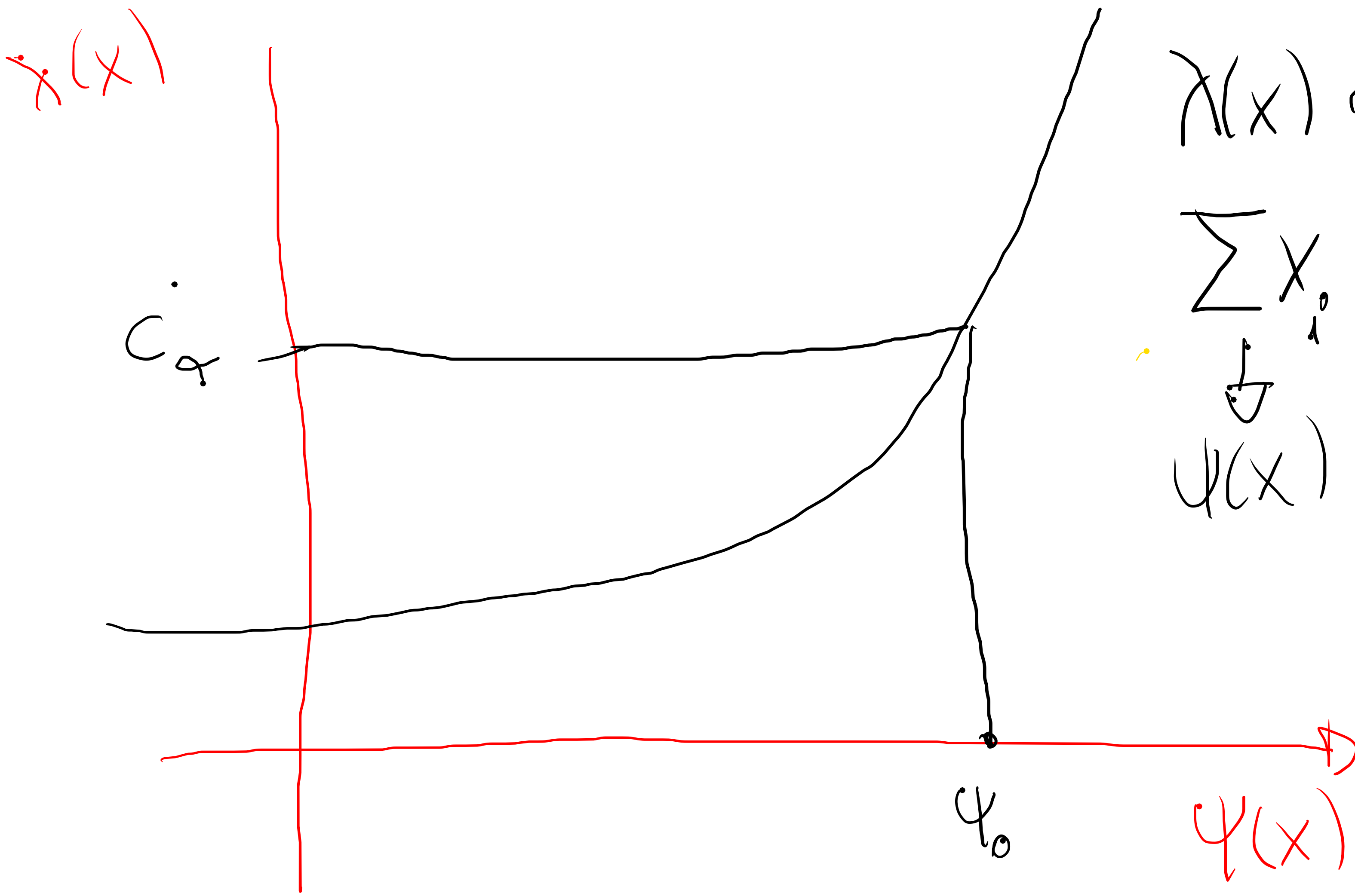
$X_i \overset{iid}{\sim} \text{Normal}(\mu, 1)$ $H_0: \mu = 0$ v/s $H_1: \mu = 1$

$$\lambda(x) = \frac{f_1(x)}{f_0(x)} = e^{\sum x_i - \frac{n}{2}}$$

$$T(x) = \begin{cases} 1, & e^{\sum x_i - \frac{n}{2}} > c_\alpha \\ 0, & \text{c.c} \end{cases}$$

es TUP.

$$\lambda(x) = e^{\sum X_i - \frac{n}{2}} > c \;,\; P_0\left(e^{\sum X_i - \frac{n}{2}} > c_\alpha\right) \leq \alpha .$$



$\lambda(x) \sim$ Modelo

$\sum X_i \sim \text{Normal}(\mu, \sigma^2)$

$\downarrow$

$\psi(x)$

$$\lambda(x) = e^{\sum x_i - \frac{n}{2}} > c \quad \Big/ \log \quad \Longleftrightarrow$$

$$\log \lambda(x) = \sum x_i - \frac{n}{2} > \log(c) \quad \Big/ + \frac{n}{2} \quad \Longleftrightarrow$$

$$\log \lambda(x) + \frac{n}{2} = \sum_{i=1}^{n} x_i > \underbrace{\log(c) + \frac{n}{2}}_{k} \quad \Longleftrightarrow$$

$$T'(x) = \begin{cases} 1, & e^{\sum x_i - \frac{n}{2}} > c \\ 0, & c.c \end{cases} \Leftrightarrow T'(x) = \begin{cases} 1, & \sum x_i > k \\ 0, & c.c. \end{cases}$$

$$P\left(e^{\sum X_i - \frac{n}{2}} > c\right) = P\left(\sum X_i > K\right)$$

identicos en probabilidad

La región crítica o de rechazo $\alpha = 0{,}05$

$$A_1 = \left\{ x \in \Delta : \sum_{i=1}^{n} x_i > K_\alpha \right\}$$

$$P_0\left(T(x) = 1\right) \leq 0{,}05$$

K_α

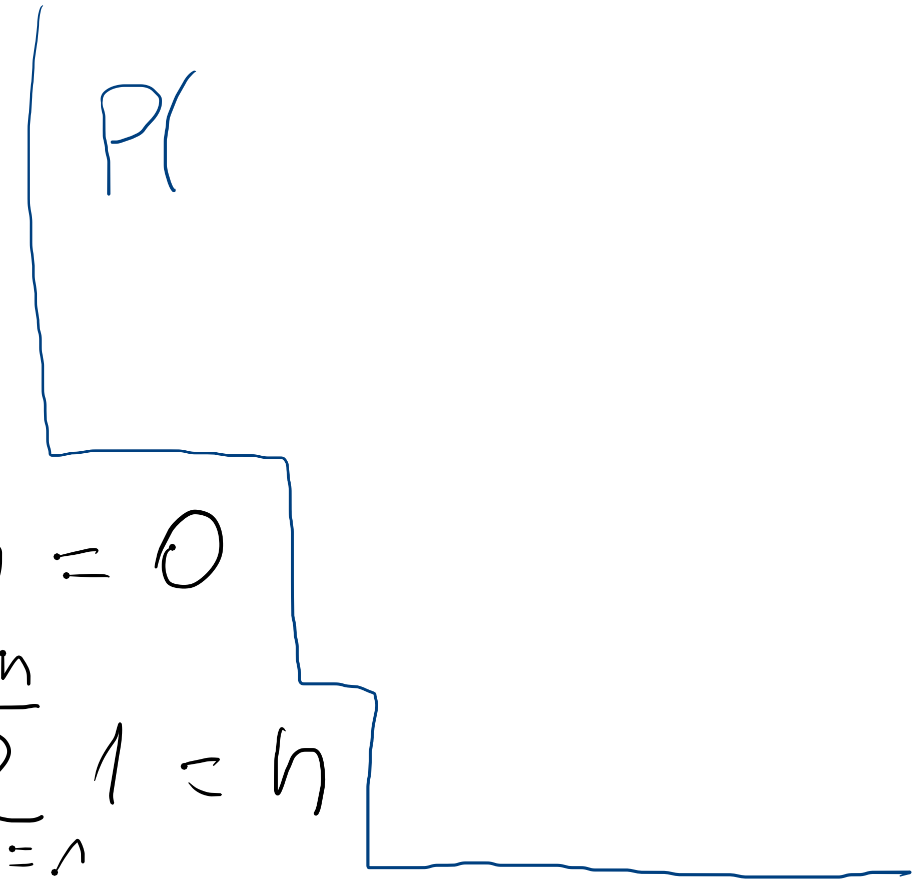
caso continuo $P_0\left(T(x) = 1\right) = 0{,}05$

$H_0: \mu = 0 \quad \forall X_i$

$$\sum_{i=1}^{n} X_i \sim \text{Normal}(0, n)$$

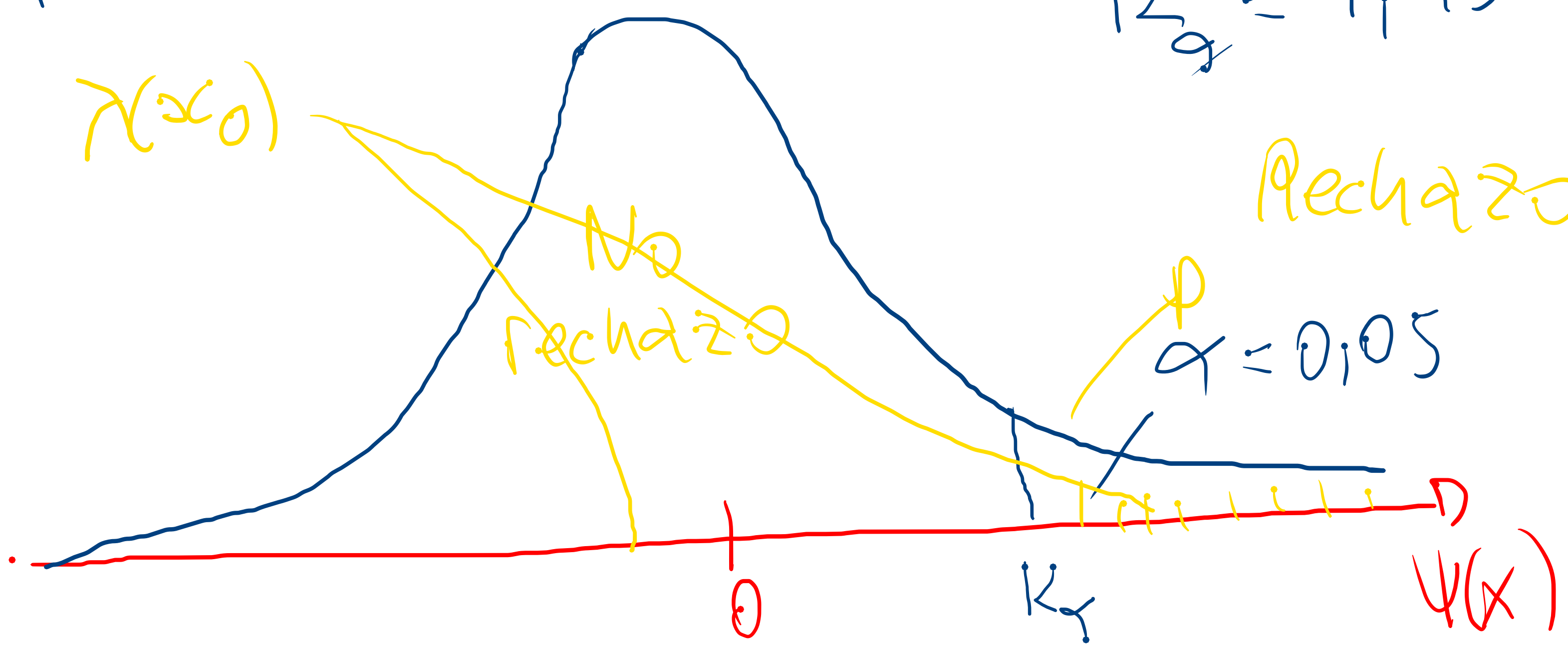
$$E_0 \sum X_i = \sum E_0(X_i) = \sum_{i=1}^{n} 0 = 0$$

$$\text{Var}_0 \sum X_i = \sum \text{Var}(X_i) = \sum_{i=1}^{n} 1 = n$$



$n = 9$

$K_\alpha = 4{,}93$



$X_i \overset{iid}{\sim} \text{Normal}(\mu, 1)$

$H_0: \mu = 0$ v/s $H_1: \mu = 1$

$$T'(X) = \begin{cases} 1, & \sum_{i=1}^{n} X_i > 4{,}93 \\ 0, & c.c \end{cases}$$

es de nivel $\alpha = 0{,}05$ y TUP por el LNP.

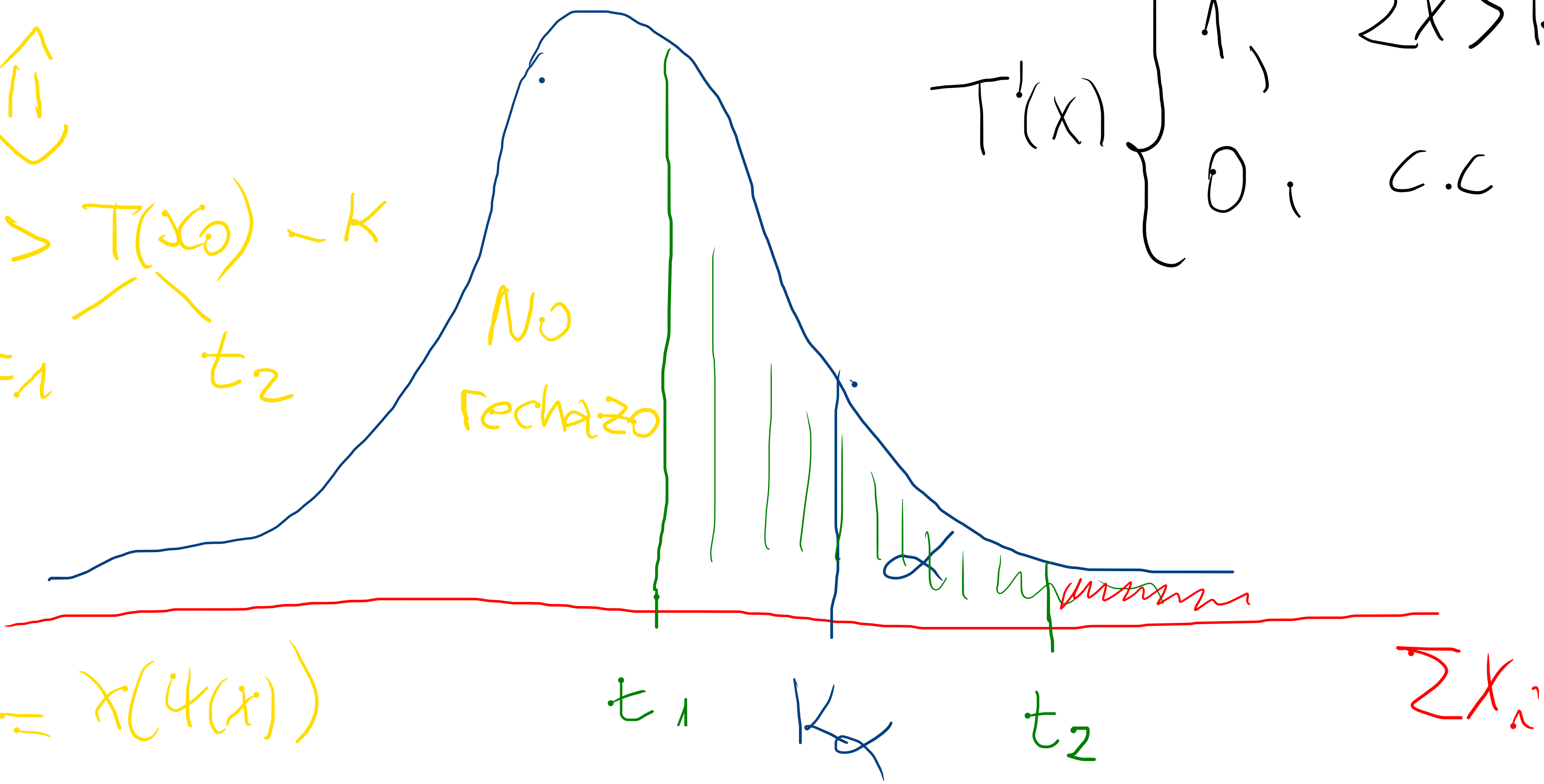
$\lambda(x) > \lambda(x_0) - c$

$\Updownarrow$

$T(x) > T(x_0) - K$

$t_1 \qquad t_2$

$$T'(x) \begin{cases} 1, & \sum X > K_\alpha \\ 0, & c.c \end{cases}$$



$\lambda(x) = \lambda(\psi(x))$

1. $P(\lambda(X) > \lambda(x_0)) = P(T(X) > t_1)$, caso $> \alpha \iff \lambda(x_0) < c_\alpha$

2. $P(\lambda(X) > \lambda(x_0)) = P(T(X) > t_2) < \alpha \iff \lambda(x_0) > c_\alpha$

$$T(x) = \begin{cases} 1, & \sup_{\theta \in \Theta_0} P(\lambda(x) \geq \lambda(x_0)) < \alpha \\ 0, & \text{c.c} \end{cases}$$

Test UP.

Criterio basado en el p_valor

$$T'(x) = \begin{cases} 1, & p\_valor < \alpha \\ 0, & p\_valor \geq \alpha \end{cases}$$

Ejemplo: caso $X_i \sim \text{Bernoulli}(p)$

$H_0: p = p_0$ v/s $H_1: p = p_1$ $(p_0 < p_1)$

$$P_0(p_1(x) = m\, p_0(x)) > 0 \qquad Y = \sum X_i$$

$$T'(x) = \begin{cases} 1, & p_1(x) > m\, p_0(x) \\ \gamma, & p_1(x) = m\, p_0(x) \\ 0, & p_1(x) < m\, p_0(x) \end{cases} \Leftrightarrow \begin{cases} 1, & Y > m \\ \gamma, & Y = m \\ 0, & Y < m \end{cases}$$

$$\lambda(x) \rightarrow \sum_{i=1}^{n} X_i \sim \text{Binomial}(p, n)$$

Error tipo I

$$E_0(T'(Y)) = 1 \times P_0(T'(Y)=1) + \gamma P_0(Y=m)$$

$$= P_0(Y>m) + \gamma P_0(Y=m)$$

$$= \sum_{j=m+1}^{n} \binom{n}{j} p_0^{j}(1-p_0)^{n-j} + \gamma \binom{n}{m} p_0^{m}(1-p_0)^{n-m}.$$

m_α y γ_α : $E_0(T'(Y)) \leq \alpha$ $n=15$ y $p_0=0{,}4$

$m = 9$ y $\gamma = 0{,}05$

$$T'(y) = \begin{cases} 1, & y > 9 \\ 0{,}04, & y = 9 \\ 0, & y < 9. \end{cases}$$

$\alpha = 0{,}05$

$n = 15$

$p_0 = 0{,}4.$

Teorema 17:

$$H_0: \theta \in \Theta_0 = \{\theta : \theta \leq \theta_0\}$$

$$H_1: \theta \in \Theta_1 = \{\theta ; \theta > \theta_0\}$$

$\lambda(x) = \lambda(\psi(x))$ es una RVM

$$\forall\, \theta_0 \in \Theta_0 \text{ y } \theta_1 \in \Theta_1 \; (\theta_0 < \theta_1).$$

$$T(x) = \begin{cases} 1, & \lambda(x) > c \\ 0, & \lambda(x) < c \end{cases}$$

Adelanto $H_0: \theta \leq \theta_0$ vs $H_1: \theta > \theta_0$

Hipótesis generales

$$\Theta_0 \cup \Theta_1 = \Theta \; ; \; \Theta_0 \cap \Theta_1 = \emptyset$$

No existe siempre un TUP.

→ Test de nivel α.

→ Test de la Razón de Verosimilitud Generalizada.