Modelo Estadístico

Definir con rigor el modelo

Teoría de Probabilidad

ML

Modelo sea desconocido

f(datos)

Modelo específico

Modelo Probabilistico

$$(\Omega, \mathcal{F}, P_0)$$

$\Omega \to \omega$

$P_0 \to$ conocida y única.

- $P(B)$: conocida $\forall B \in \mathcal{F}$
- $f(P) = \int \omega P$ Valor medio.

Sea X una función aleatoria

variable

$$X : \Omega \to \mathbb{R}$$
$$\omega \to x$$

X: induce un modelo de probabilidad

$\mathcal{X}$: espacio muestral de X

$\mathcal{G}$: σ-algebra de $\mathcal{X}$

P_X: la med. de prob. inducida x X

Original: $(\Omega, \mathcal{F}, P_0)$

Si $A \in \mathcal{G}$.

$$P_X(X \in B) = P_0(\{\omega \in \Omega : X(\omega) \in A\})$$

Modelo de Prob. inducido por X

$$(\mathcal{X}, \mathcal{G}, P_X)$$

$\hookrightarrow$ conocida.
Totalmente especificada.

$X:$ Discreta ✓

Continua ✓ $P_X \to$

Mixta

- densidad de probabilidad.

- función de probabilidad.

En el caso continuo

$$P_X(A) = \int_A f_X(x)\,dx. \quad \forall A \in \mathcal{G}.$$

$(\mathcal{X}, \mathcal{G}, f_X)$ Model Estad.

Agregado.

Modelo prob. Gaussiano

$$\left(\mathbb{R}, \mathcal{B}, P_X\right), \qquad \mu = 0, \quad \sigma = 1$$

$\forall\ (a,b] = A$

donde $$P_X(A) = \int_A \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}_{f_X} dx$$ ".

Exp. aleatorio

Incertidumbre: en el resultado del experimento.

$(\mathbb{R}, \mathcal{B}, P_x)$

Estadístico: no hay incertidumbre sobre el exp. x datos.

$(\mathbb{R}, \mathcal{B}, P_x)$

Conozco sus propiedades pero no sé cuál en específico es.

¿Como represento mi incertidumbre sobre P_x?

Explicar P_x

- Paramétrico Clásico
- Bayesiano
- No paramétrico

Predecir un funcional de P_x

- Machine Learning

Familia de medidas de Prob.

$$\mathcal{P} = \{ P_1, P_2, \dots, P_{100} \};\ P_i \neq P_j$$

$i \neq j.$
$i, j = 1 - 100.$

Familia parametrizada de medidas de prob.

$$\mathcal{P}_{\Theta} = \{ P_\theta ; \theta \in \Theta \}$$

$\Theta \to$ espacio paramétrico.

Reprsentación de mi incertidumbre acerca de PX

θ: escalar o vector determinístico.

Modelo Bayesiano Paramétrico

X: var. aleatoria $\quad P_X = ¿?$ → P datos

θ: — —

Π: conocida "

Modelo a priori para θ

$(\Theta, \mathcal{G}_{\Theta}, \Pi)$.

$\Pi(C) = P(\theta \in C)$.

$$\mathcal{P}_{\Theta} = \{ P_{\theta} ; \theta \in \Theta \text{ y } \theta \text{ es una variable aleatoria} \}.$$

$$\frac{\text{Modelo Bayesiano posterior}}{\text{Modelo de Prob.}}$$

$$\mathcal{P}_{\Theta} \longrightarrow P_{\theta|x} . \left(\Theta , g_{\Theta} , P_{\theta|x} \right)$$

Modelo paramétrico clásico

$$\left(\mathcal{X}, \mathcal{G}, \mathcal{P}_{\Theta} = \{P_\theta : \theta \in \Theta\}\right)$$

θ: escalar o vector

Modelo Bayesiano paramétrico

$$\left(\mathcal{X}, \mathcal{G}, \mathcal{P}_{\Theta} = \{P_\theta : (\theta, X) \text{ es aleatorio}\}\right)$$

(θ, X) es aleotorio.

- $f(\theta_0, x)$

$f(\theta_0)$

$f_{\theta|X}$

Modelo Bayesiano Paramétrico

$\left(\Theta, g_{\Theta}, P_{\theta|X}\right)$ → Modelo de Probabilidad.

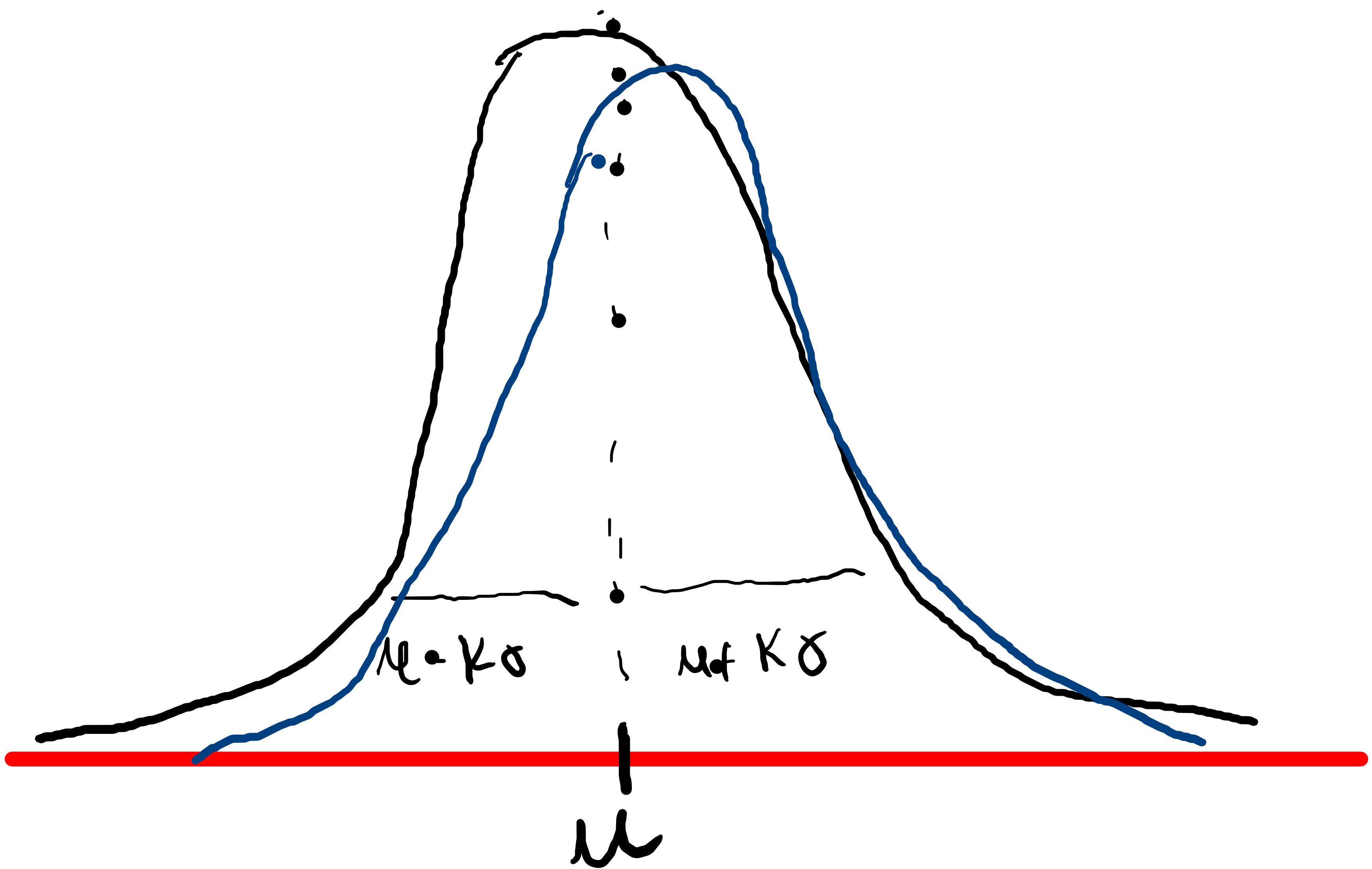
$$\mathcal{P} = \{ P : P \text{ una prop específica} \}$$

<u>Población</u>: P_0 que generó los datos (muestra) $x = (x_1, \ldots, x_n)$ y que está en una clase $\mathcal{P}$.

Simetria

$X \sim \text{Normal}(\mu, \sigma^2)$: suposición fuerte.

Y es simetrica: menos fuerte.

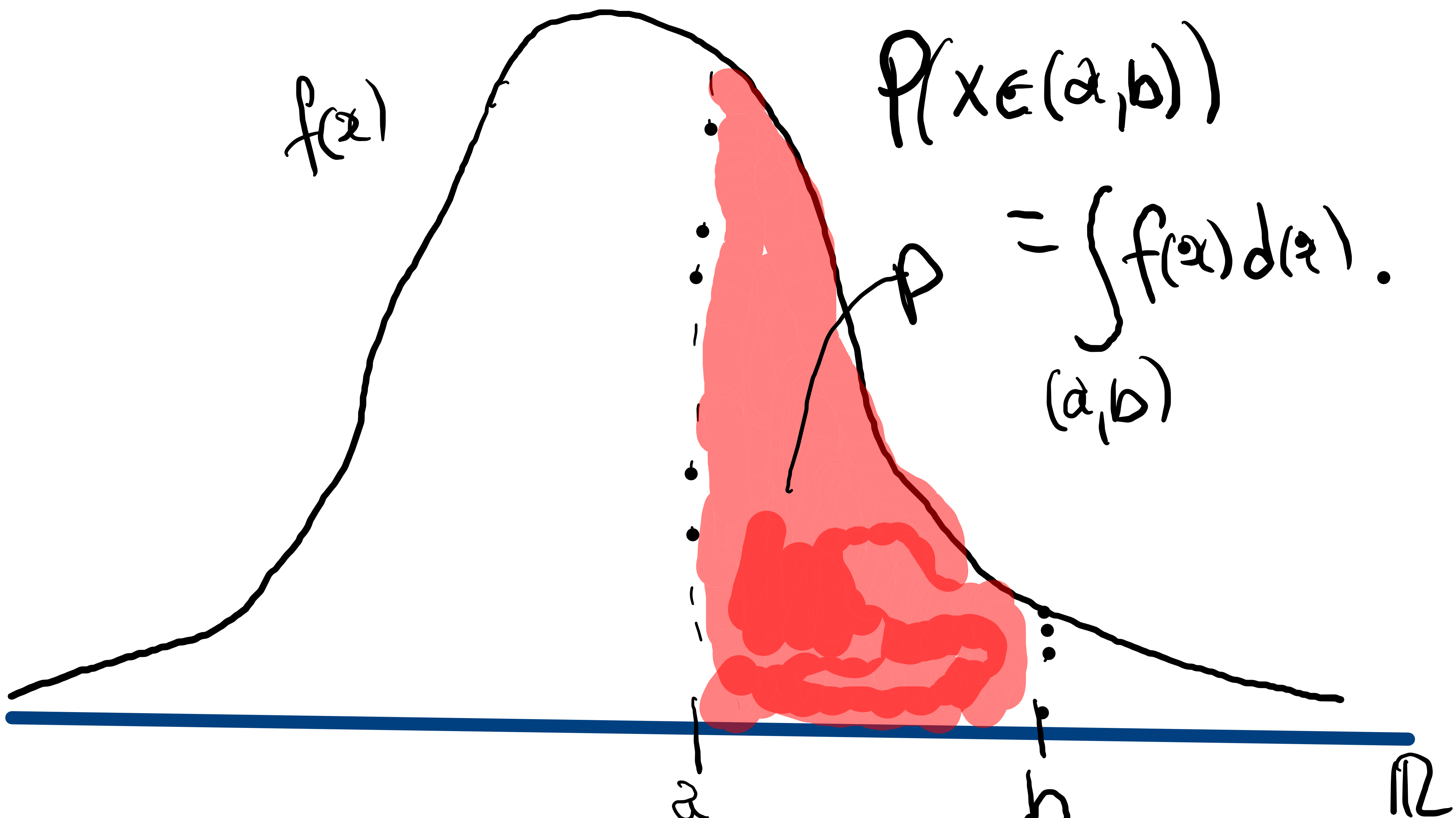


Función de densidad

Una función $f(x)$ real es llamada de f.d.p para X si:

(i) $f(x) \geq 0$

(ii) $P(X \in (a,b)) = \int_{(a,b)} f(x)\,dx < \infty$



$$P(X \in (a,b)) = \int_{(a,b)} f(x)\, d(x).$$

Modelo no paramétrico

$(\mathcal{X}, \mathcal{G}, \mathcal{P})$

↳ no va a tener una descripción en base a θ.

-

① $\mathcal{P} = \{ P : P(A) = \int_A f(x)dx$ y $f(x)$ es simétrica $\}$

(2) $\mathcal{P} = \{ P : E(X) = \int x \, dP < \infty \}$

$\rightarrow$ X tiene valor esperado.

(3) Teoria de estadístico U.

$$F(x) = \int_{-\infty}^{x} f(s)\,ds = P(X < x).$$
$$= P(X \in (-\infty, x))$$

$$\mathcal{P} = \{ P : E(X) = \int x \, dF < \infty \}$$

Valor esperado

$$E(X) = \int x\,dP = \int x\,dF = \begin{cases} +\infty, \\ \\ <\infty, \end{cases}$$

$\downarrow$ L $\qquad\qquad$ $\downarrow$ R-S

Modelo Gaussiano.

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

$$P_\theta(A) = \int_A f_\theta(x)\,dx \quad \text{en que}$$

$$f_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \theta = (\mu, \sigma^2)$$

$$\Theta = \mathbb{R} \times \mathbb{R}_+$$

$$\left(\mathbb{R}, \mathcal{B}, \mathcal{P}_{\Theta} = \left\{ P_\theta(A) = \int_A f_\theta(x)\,dx \right\}\right)$$