Estadística Teorica

$$\mathrm{Var}\, X = E(X-EX)^2 = \begin{cases} \int (x-EX)^2 p(x)\,dx \\ \sum_{x \in \mathcal{X}_0} (x-EX)^2 p(x) \end{cases}$$

$$\mathrm{Var}\, X = EX^2 - (EX)^2$$

Vectores Aleatorios

Si $X_1, \dots, X_p$ son variables aleatorios

$$\mathbb{X} : \mathcal{E} \to \mathbb{R}^p \quad \text{def} \quad \mathbb{X}(e) = \begin{pmatrix} X_1(e) \\ X_2(e) \\ \vdots \\ X_p(e) \end{pmatrix} \quad \text{Vector aleatorio}$$

su medida de distribución

$P_{\mathbb{X}}$ definida $P_{\mathbb{X}}(B) = P(\mathbb{X} \in B) = P(\{e \in \mathcal{E} : \mathbb{X}(e) \in B\})$

$\mathbb{X}$ se dice absolutamente continuo con densidad $p_{\mathbb{X}}$ si

$P_{\mathbb{X}} \ll$ Lebesgue en cuyo caso

$$P(\mathbb{X} \in B) = \int \dots \int_B p_{\mathbb{X}}(x_1, \dots, x_p)\, dx_1 \dots dx_p$$

$\mathbb{X}$ se dice discreta si $P(\mathbb{X} \in \mathcal{X}_0) = 1$ para algun $\mathcal{X}_0$ numerable

Medidas Producto

Sea μ una medida en $(\mathcal{X}, \mathcal{A})$
ν — — en $(\mathcal{Y}, \mathcal{B})$

Entonces $\exists!$ medida $\mu \times \nu$ llamada 'medida producto' en $(\mathcal{X} \times \mathcal{Y}, \mathcal{C})$, donde $\mathcal{C}$ es la σ-algebra mas chica que contiene todos los subconjuntos de $\mathcal{A} \times \mathcal{B}$.

$$\mu \times \nu (A \times B) = \mu(A) \cdot \nu(B)$$

$$\forall A \in \mathcal{A},\ B \in \mathcal{B}$$

Ej 1: μ y ν son Lebesgue en $\mathbb{R}^p$ y $\mathbb{R}^q$
$\mu \times \nu$ es — — $\mathbb{R}^{p+q}$

2: μ y ν son conteo en $\mathcal{X}$ e $\mathcal{Y}$.
$\mu \times \nu$ es conteo $\mathcal{X} \times \mathcal{Y}$

Teorema Fubini (Parte A)

Si f nonnegativa y (m) entonces

$$\int f \, d(\mu \times \nu) = \int \left\{ \int f(x,y) \, d\nu(y) \right\} d\mu(x)$$

$$= \int \left\{ \int f(x,y) \, d\mu(x) \right\} d\nu(y)$$

Independencia

Dos vectores aleatorios $X \in \mathbb{R}^p$ y $Y \in \mathbb{R}^q$ son independientes si

$$P(X \in A \text{ y } Y \in B) = P(X \in A)\, P(Y \in B)$$

Si $Z = \binom{X}{Y}$ $\quad P(Z \in A \times B) = P(X \in A)\, P(Y \in B)$

$X \perp\!\!\!\perp Y$ $\quad P_Z(A \times B) = P_X(A)\, P_Y(B)$

Supongamos que P_X tiene densidad p_X c.r. μ
P_Y ✓ ✓ p_Y c.r. ν

entonces $Z = \binom{X}{Y}$ tiene densidad

$$p_Z(z) = p_X(x)\, p_Y(y)$$

DEM $P(Z \in A) = \int 1_A(z)\, dP_Z(z) = \int 1_A(z)\, d(P_X \times P_Y)$

$$= \int \left\{ \int 1_A(x,y)\, dP_X(x) \right\} dP_Y(y)$$

$$= \int \left\{ \int 1_A(x,y)\, p_X(x)\, d\mu(x) \right\} p_Y(y)\, d\nu(y)$$

$$= \int 1_A(x,y) \underbrace{\left[p_X(x)\, p_Y(y) \right]} d(\mu\times\nu)(x,y)$$

$$= \int 1_A(z)\, p_Z(z)\, d(\mu\times\nu)(z) \qquad \square$$

Distribuciones Conjuntas, Marginales y Condicionales

Tenemos $Z = \binom{X}{Y}$ con $X \in \mathbb{R}^p$ e $Y \in \mathbb{R}^q$
y medidas μ y ν resp.

Supongamos, aparte, que Z tiene densidad $p_Z(z)$ con respecto a $(\mu\times\nu)$.

$P(X \in A) = P(Z \in A\times\mathbb{R}^q) = \int 1_{A\times\mathbb{R}^q}\, p_Z\, d(\mu\times\nu)$

$$= \int \left\{ \int 1_{A\times\mathbb{R}^q}\, p_Z(x,y) \right\} d\nu(y)\, d\mu(x)$$

$$= \int 1_A(x) \underbrace{\int \cancel{1_{\mathbb{R}^q}(y)}\, p_Z(x,y)\, d\nu(y)}_{h(x)}\, d\mu(x)$$

$$p_X(x) = \int p_Z(x,y)\, d\nu(y)$$ densidad marginal de X

$$p_Y(y) = \int p_Z(x,y)\, d\mu(x)$$

Ej Antes recordar

$$\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\, dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = B^{-1}(\alpha,\beta)$$

Supongamos μ conteo en $\{0,1,\dots,k\}$ $\exists \in \mathbb{N}$

ν Lebesgue en $\mathbb{R}$

$$p_Z(x,y) = \begin{cases} \binom{k}{x} y^x (1-y)^{k-x} & x=0,\dots,k \\ & y \in (0,1) \\ 0 & \text{c.c.} \end{cases}$$

$$p_X(x) = \int p_Z(x,y)\, d\nu(y) = \int_0^1 \binom{k}{x} y^x (1-y)^{k-x}\, dy$$

$$= \binom{k}{x} \frac{\Gamma(x+1)\Gamma(k-x+1)}{\Gamma(k+2)} \underbrace{\frac{\Gamma(x+1+k-x+1)}{\Gamma(x+1)\Gamma(k-x+1)} \int_0^1 y^x (1-y)^{k-x}\, dy}_{=1}$$

$$= \underbrace{\binom{k}{x}}_{\frac{k!}{\cancel{x!}\,\cancel{(k-x)!}}} \frac{\cancel{x!}\,\cancel{(k-x)!}}{(k+1)!} = \frac{1}{k+1}$$

$X \sim U(0,\dots,k)$

$$p_X(x) = \begin{cases} \dfrac{1}{k+1} & x = 0,1,\dots,k \\ 0 & \text{c.c.} \end{cases}$$

$$p_Y(y) = \int p_Z(x,y)\, d\mu(x) = \sum_{x=0}^{k} \binom{k}{x} y^x (1-y)^{k-x}$$

$$= \{y + (1-y)\}^k = 1^k = \underline{1}$$

$$Y \sim U(0,1)$$

$$p_Y(y) = \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{c.c.} \end{cases}$$

Distribuciones Condicionales

Definimos la densidad condicional

$$p_{Y|X}(y|x) := \frac{p_Z(x,y)}{p_X(x)}, \quad p_X(x) \neq 0$$

Observar

Caso discreto $P(Y=y \mid X=x) = \dfrac{P(X=x, Y=y)}{P(X=x)}$

Caso continuo $p_{Y|X}(y|x)$ es la densidad condicional de Y dado $X=x$.

Def $P(Y \in A \mid X=x) = \int_A p_{Y|X}(y|x)\, d\nu(y)$

$$E[f(Y) \mid X=x] = \int f(y)\, p_{Y|X}(y|x)\, d\nu(y)$$

Integrabilidad

Una función medible f se dice integrable con respecto a una medida μ si $\int |f|\, d\mu < \infty$

(notar que esto implica ambas $\int f^+ d\mu < \infty$ y $\int f^- d\mu < \infty$)

Fubini (Parte B). Si f es integrable con respecto a $\mu\times\nu$ entonces

$$\int f\, d(\mu\times\nu) = \int\left[\int f(x,y)\, d\mu(x)\right] d\nu(y)$$
$$= \int\left\{\int f(x,y)\, d\nu(y)\right\} d\mu(x)$$

Cor.

$$\underbrace{E\left[f(X,Y)\right]} = E\left[E f(X,Y) \mid X\right] \quad \text{Suavizado}$$
$$= E\left[\underbrace{E f(X,Y) \mid Y}\right]$$

Ej Retomemos el ej. anterior

$$p_2(x,y) = \binom{k}{x} y^x (1-y)^{k-x}$$

$$EY = E\left[EY \mid X\right] = \text{----} \text{ con ej.}$$

Estimación Puntual

Supongamos una muestra $X_1, X_2, \dots, X_n =: \mathbf{X}$ cuya distribución $\mathbf{X} \sim F_{\mathbf{X}} \in \mathcal{F}$

Cuando $F_{\mathbf{X}} = F_{\mathbf{X},\theta}$ y $\theta \in \Theta \subset \mathbb{R}^q \quad \exists q \in \mathbb{N}$

o sea que $\mathcal{F} = \{F_{\mathbf{X},\theta} : \theta \in \Theta\}$

El modelo se llama paramétrica

Ej $\quad X_i \overset{iid}{\sim} N(\mu, \sigma^2) \quad \mathbf{X} = (X_1, \dots X_n) \sim N_n\left(\begin{pmatrix}\mu\\ \vdots\\ \mu\end{pmatrix}, \sigma^2 I\right)$

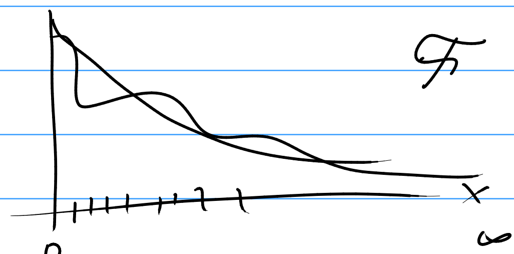
$\Theta = (\mu, \sigma^2) \in \mathbb{R}\times\mathbb{R}^+ \subset \mathbb{R}^2$

$$\mathcal{F} = \left\{ \Phi\left(\frac{x_1-\mu}{\sigma}\right) \cdots \Phi\left(\frac{x_n-\mu}{\sigma}\right) : (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+ \right\}$$

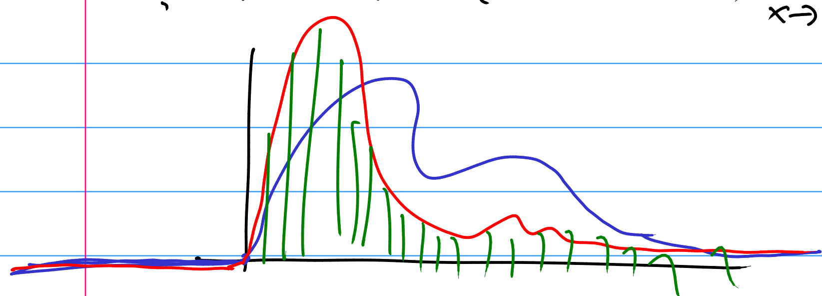
q nro parametros $(=2)$

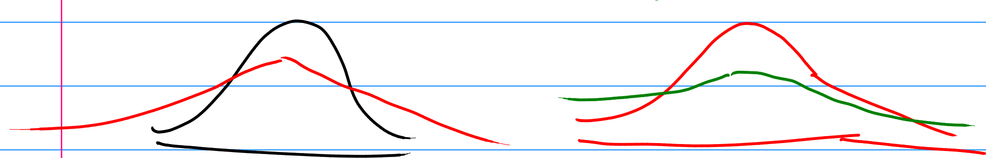
Θ espacio paramétrico.

si $\nexists\, q \in \mathbb{N}$ tal que $\Theta \subset \mathbb{R}^q$.



$X \sim f$ $\qquad \mathcal{F} = \{ f : f(0) < \infty ,\ f(x) \xrightarrow[x\to\infty]{} 0 , \text{cont.} \}$





Cauchy $\qquad f(x) = \dfrac{1}{\pi} \dfrac{1}{1+x^2} \qquad x \in \mathbb{R}$

$$e^x \Phi(x) \xrightarrow[x\to\infty]{} 0$$

$$(x_1 \ldots x_n) = \mathbf{x} \sim F \in \mathcal{F}$$

Estimar los parametros significa intentar identificar el Θ de la muestra $\mathbf{x} \sim F_\Theta \in \mathcal{F}$.

$x_1 \sim \{ N(\mu, 1) \}$ $\qquad x_1 = 5$

Un estadístico $T(\mathbf{x})$ es cualquier función de la muestra

$$T(\mathbf{x}) = (X_1 + \cdots + X_n)/n \quad \checkmark$$

$$T_2(\mathbf{x}) = X_{(n)} - X_{(1)} \quad \checkmark$$

$$T_3(\mathbf{x}) = \text{mediana}\{X_1, \ldots X_n\}$$

$$T_4(X) = \arcsin\left|\frac{X_{(1)}}{X_{(n)}}\right|$$

$$T_5(X) = \frac{X_1 - \mu}{\sigma} \quad \underline{\text{NO}}$$

<u>Def</u> Un estimador es un estadístico que se usa para estimar un parametro.

Ej $\quad X_1, \ldots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$\hat{\mu} = \bar{X}$ estimador de μ

LGN $\quad \bar{X} \xrightarrow{P} \mu \qquad \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

<u>Metodos de Estimación</u>

1) Metodo de los momentos
2) Metodo de Maxima Verosimilitud
3) Mínimos Cuadrados
4) Bayesiano

<u>Def</u> Una muestra simple al azar es un conjunto de observación $X_1 \ldots X_n$ i.i.d.

Método de los Momentos (Karl Pearson)

Sea $X_1 \dots X_n \overset{iid}{\sim} F_\theta \in \mathcal{F}$, $\theta \in \Theta \subset \mathbb{R}^q$

El metodo del los momentos propone un sistema de q ecuaciones

$$\begin{cases} h_1(\theta) = EX = \frac{1}{n}\sum X_i \\ h_2(\theta) = EX^2 = \frac{1}{n}\sum X_i^2 \\ \vdots \\ h_q(\theta) = EX^q = \frac{1}{n}\sum X_i^q \\ \vdots \\ h_{q+m}(\theta) = EX^{m+q} = \frac{1}{n}\sum X_i^{q+m} \end{cases}$$

hasta tener q ecuaciones (que sirvan)

los $\hat{\theta}_1(x) \dots \hat{\theta}_q(x)$ que resuelven el sistema se llaman estimadores de los momentos.

Ej $X_1 \dots X_n \overset{iid}{\sim} \mathcal{E}(\lambda)$ $\quad f_X(x) = \lambda e^{-\lambda x}$ $x>0$

$\lambda \in (0,\infty)$

$$EX = \frac{1}{n}\sum X_i$$

$$\frac{1}{\lambda} = \bar{x} \Rightarrow \boxed{\hat{\lambda} = \frac{1}{\bar{x}}}$$

Ej 2. $X_1 \dots X_n \sim \text{Gamma}(\alpha, \beta)$

$\alpha, \beta \in \mathbb{R}^{+2}$

$$f_X(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}$$

$$\begin{cases} \alpha\beta = EX = \frac{1}{n}\sum X_i \\ \alpha\beta^2 + (\alpha\beta)^2 = EX^2 = \frac{1}{n}\sum X_i^2 \end{cases}$$

$$m_1 = \frac{1}{n}\sum x_i$$

$$m_2 = \frac{1}{n}\sum x_i^2$$

$$\begin{cases} \alpha\beta = m_1 & (1) \\ \alpha\beta^2 + \alpha^2\beta^2 = m_2 & (2) \end{cases}$$

(2) $\underbrace{(\alpha\beta)}\beta + (\alpha\beta)^2 = m_2$

$$m_1\beta + m_1^2 = m_2$$

$$\boxed{\hat{\beta} = \frac{m_2 - m_1^2}{m_1} = \frac{\frac{1}{n}\sum (x_i - \bar{x})^2}{\frac{1}{n}\sum x_i}} = \frac{\frac{1}{n}\sum x_i^2 - \left(\frac{1}{n}\sum x_i\right)^2}{\frac{1}{n}\sum x_i}$$

(1) $\alpha\beta = m_1$

$$\alpha = \frac{m_1}{\hat{\beta}} = \frac{m_1}{\frac{m_2 - m_1^2}{m_1}} = \frac{m_1^2}{m_2 - m_1^2}$$

$$\boxed{\hat{\alpha} = \frac{\bar{x}^2}{\sum x_i^2/n - \bar{x}^2}}$$

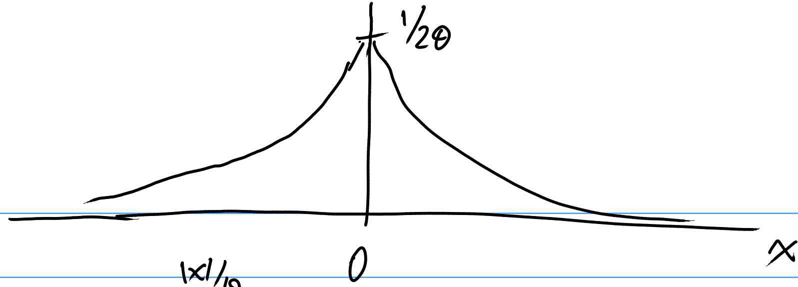
Ej $X_1 \dots X_n \overset{iid}{\sim} P(\lambda)$ $\quad (1-\varepsilon)P(\lambda) + \varepsilon\,\delta_{x_0}$

A) $EX = \bar{X}$ $\quad \hat{\lambda} = \bar{X}$

B) $\tilde{\lambda} = \frac{1}{n}\sum (X_i - \bar{X})^2$

$$E\, g_k(X) = \frac{1}{n}\sum g_k(x_i)$$

Ej $X_1 \dots X_n \overset{iid}{\sim} f_X(x) = \frac{1}{2\theta} e^{-|x|/\theta}$ $\quad x \in \mathbb{R}$, $\theta > 0$



$$f(x) = \frac{1}{2\theta} e^{-|x|/\theta}$$

$$0 = EX = \frac{1}{n} \sum X_i \qquad \cancel{X}$$

$$EX^2 = \frac{1}{n} \sum X_i^2$$

$$EX^2 = \frac{1}{2\theta} \int_{-\infty}^{\infty} x^2 e^{-|x|/\theta} dx = \frac{1}{\theta} \int_0^{\infty} x^2 e^{-x/\theta} dx$$

$$= \theta^2 \int_0^{\infty} \left(\frac{x}{\theta}\right)^2 e^{-x/\theta} \frac{1}{\theta} dx \qquad u = x/\theta$$

$$= \theta^2 \int_0^{\infty} u^2 e^{-u} du$$

$$= \theta^2 \quad \underbrace{\Gamma(1)\ 1'}_{1} \qquad \frac{\int_0^{\infty} u^{1+1} e^{-u} du}{\Gamma(1)\ 1'} \quad 1$$

$$\theta^2 = \frac{1}{n} \sum X_i^2$$

$$\hat{\theta} = \sqrt{\frac{1}{n} \sum X_i^2}$$