Estadistica Teorica 2012

ENRIQUE ALVAREZ emriqueealvarez@fibertel.com.ar
Teoria Ma 17-21
ealvarez@mate.unlp.edu.ar

Susana Sombielle ssombielle@gmail.com
Practica Jueves

Examenes 2 parciales c/recuperatorio
fechas : ___
___

Bibliografía : Apuntes de Boente - Yohai
Libros 1) Casella Berger "Statistical Inference"

2) Lehmann - Casella
Theory of Point Estimation
2b) Lehmann - Romano
Testing Statistical Hypotheses

3) Theoretical Statistics
Keener

Introducción

En este curso trabajaremos con

- Datos, representados por X, que puede ser una variable aleatoria o más generalmente un vector aleatorio. X se observan

- Parámetro, θ que puede ser escalar o vector
  θ no se conoce ni observa

- X y θ están relacionados por un <u>modelo</u>

  $X \sim P_\theta$. Generalmente a partir de X se nos pide encontrar θ.

- Objetivo: Encontrar alguna función δ, llamada estimador tal que $\delta(X)$ esté cerca de $g(\theta)$ (notar $g(\theta) = \theta$ es habitual)
  Decimos que $\delta(X)$ es un estimador de $g(\theta)$

<u>Ejemplo</u> Giramos una moneda 100 veces

$X = \#$ caras $\qquad \theta =$ probabilidad de cara

$X \sim \text{Binomial}(n=100, \theta) \qquad g(\theta) = \theta$

Un estimador natural es $\delta(X) = \dfrac{X}{100}$ (ah-hoc)

Nos podemos preguntar:

a) ¿Es un buen estimador? ¿Qué significa buen esti?

b) ¿Es óptimo?

c) ¿Existen procedimientos generales para proponer buenos estimadores en situaciones parecidas?

Teoria de la Desición (Cómo decir si δ es bueno de una manera "economicista")

Función de pérdida $L(\theta, d)$ es la pérdida de estimar $g(\theta)$ mediante d.

Debe satisfacer: a) $L(\theta, d) \geq 0$

b) $L(\theta, d) = 0$ ssi $d = g(\theta)$

ej: $L(\theta, d) = (g(\theta) - d)^2$ perdida cuadratica

$L(\theta, d) = |g(\theta) - d|$ perdida absoluta

Dado que $L(\theta, \delta(X))$ es aleatorio, su valor con datos $X = x$, $L(\theta, \delta(x))$ puede ser grande aún para una buena elección de $\delta(X)$ y viceversa.

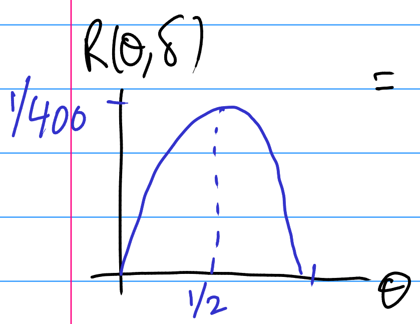
Para evaluar el desempeño de un estimador usaremos la pérdida esperada o riesgo

$$R(\theta, \delta) = E_\theta[L(\theta, \delta(X))]$$

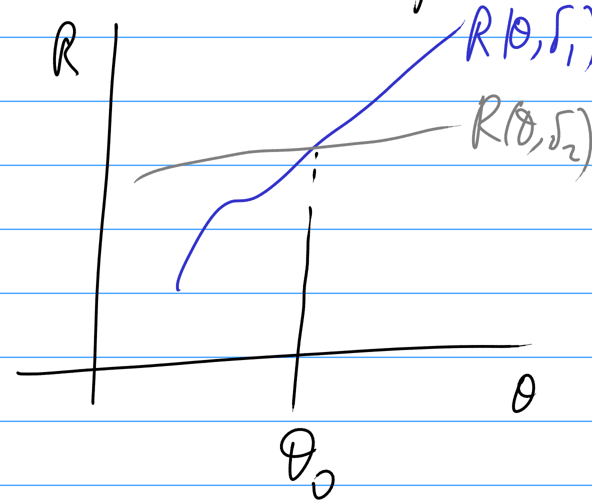
Ej $L(\theta, d) = (\theta - d)^2$

$$R(\theta, \delta) = E_\theta\left(\theta - \frac{X}{100}\right)^2 = \frac{1}{100^2} E_\theta (X - 100\theta)^2$$

$$= \frac{1}{100^2} Var_\theta(X) = \frac{1}{100^2} 100\,\theta(1-\theta) = \frac{\theta(1-\theta)}{100}$$



Al comparar estimadores δ_1, δ_2 digamos puede pasar



δ_2 es uniformemente mejor que δ_1 (preferible)

Veremos maneras de responder a la situación $\underline{2}$.

Breve repaso/Introducción a Medida

Una medida μ en un conjunto $\mathcal{X}$ asigna un valor no negativo $\mu(A)$ a muchos, no todos, subconjuntos $A \subset \mathcal{X}$. Debe satisfacer

① $\mu : \mathcal{A} \to [0,\infty]$ $\quad$ $\mathcal{A}$ es un conjunto de subconjuntos de $\mathcal{X}$, dominio de μ

② Si $A_1, A_2, \ldots \in \mathcal{A}$ son disjuntos entonces

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) \quad \text{(aditividad numerable)}$$

Ejemplos: a) Si $\mathcal{X}$ es numerable $\mu(A) = \# A$
μ se llama medida de conteo.

b) si $\mathcal{X} = \mathbb{R}^n$ y definimos (para $A \in \mathcal{a}$)

$$\mu(A) = \int_A \cdots \int dx_1 \ldots dx_n$$

o sea para $n = 1, 2, 3$ largo, área, volumen

μ se la llama medida de Lebesgue

Al margen Ej tomar $\mathbb{Q}$ nros racionales

$$\mathbb{Q} := \{ u = x/y ,\ x \in \mathbb{Z} ,\ y \in \mathbb{Z} \}$$

$$1_{\mathbb{Q}}(s) = \begin{cases} 1 & s \in \mathbb{Q} \\ 0 & \text{c.c.} \end{cases}$$

$\int 1_{\mathbb{Q}}(s)\, ds = 0$ en el sentido de Lebesgue

no es integrable en el sentido de Riemman.

Fin al margen

El dominio $\mathcal{a}$ debe ser una σ-algebra:

1. $\phi, \mathcal{X} \in \mathcal{a}$
2. $A \in \mathcal{a} \Rightarrow A^c \in \mathcal{a}$
3. $A_1, A_2, \ldots \in \mathcal{a}$ entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{a}$, $\bigcap_{i=1}^{\infty} A_i \in \mathcal{a}$.

Decimos que $(\mathcal{X}, \mathcal{a})$ es un espacio medible

$(\mathcal{X}, \mathcal{a}, \mu)$ " " " de medida.

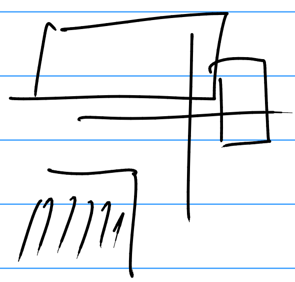
μ se llama 'σ-finita' si existen $A_1, A_2, \ldots$ tal que $\bigcup_i A_i = \mathcal{X}$ y $\forall i$ $\mu(A_i) < \infty$

μ se llama medida de probabilidad si $\mu(\mathcal{X}) = 1$.

En el ejemplo a) $\mathcal{A} = \{\text{todo subconjunto de } \mathcal{X}\}$

$$= 2^{\mathcal{X}}$$

b) $\mathcal{A}$ el σ-algebra de Borel, que es el σ-algebra más chica que incluye todos los rectangulos n-dimensionales.



Integración

Objetivo: Definir integrales $\int f\,d\mu$ para funciones regulares f que satisfagan:

① $\int 1_A\,d\mu = \mu(A)$

② $\int (cf + dg)\,d\mu = c\int f\,d\mu + d\int g\,d\mu$

③ si $0 \le f_1 \le f_2 \le \cdots$ y $f(x) = \lim_{n\to\infty} f_n(x)\ \forall x$

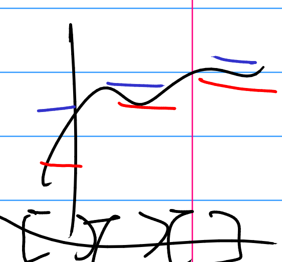
entonces $\int f\,d\mu = \lim_{n\to\infty} \int f_n\,d\mu$

Ej en a) $\int f\,d\mu = \sum_{i=1}^{\infty} f(x_i)$

b) $\int f\,d\mu$ Lebesgue, coincide con Riemann para la mayoría de f.

$\int f(x)\,dx$ $\qquad g_n(x) \nearrow f(x)$ $\qquad h_n(x) \searrow f(x)$

Riemann



$$\underbrace{\sum_i g_n(x_i)\Delta_n}_{} \le \int f(x)\,dx \le \underbrace{\sum_i h_n(x_i)\Delta_n}_{}$$

$\lim_{n\nearrow\infty}$ $(\Delta\to 0)$ $\qquad \int f(x)dx = \lim$

Lebesgue $\sup_{A_i\in\mathcal{G}} \sum \inf_{x\in A_i} f(x)\,\mu(A_i) \le \int f(x)\,d\mu(x) \le \sup_{A_i\in\mathcal{G}} \sum \sup_{x\in A_i} f(x)\,\mu(A_i)$

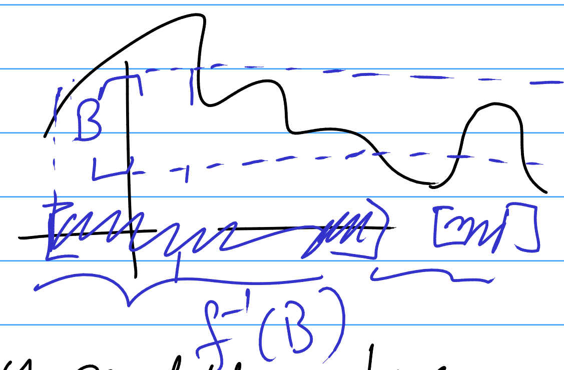
(Ref Billingsley. Probability and Measure)

Una función simple toma solamente un nro finito de valores, y puede escribirse $s = \sum_{i=1}^{m} a_i I_{A_i}$

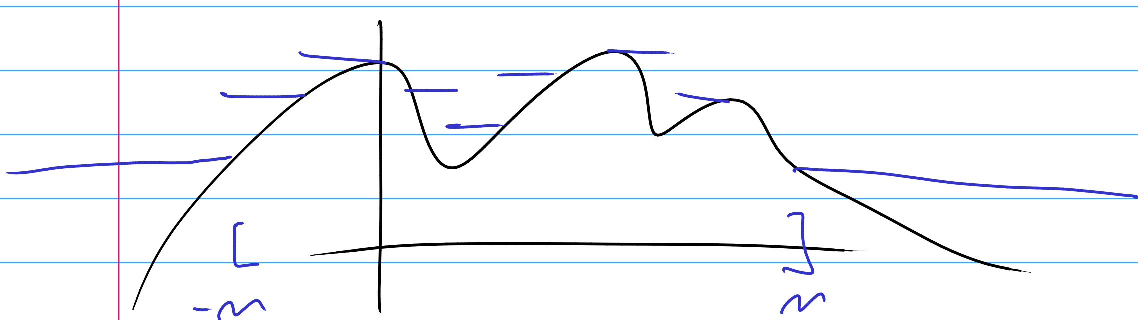
entonces $\int s \, d\mu = \sum_{i=1}^{m} a_i \mu(A_i)$

Def f es medible si $f^{-1}(B) \in \mathcal{A}$ para todo boreliano B. Abreviamos m

$$f: \mathbb{R} \to \mathbb{R}$$
$$B$$



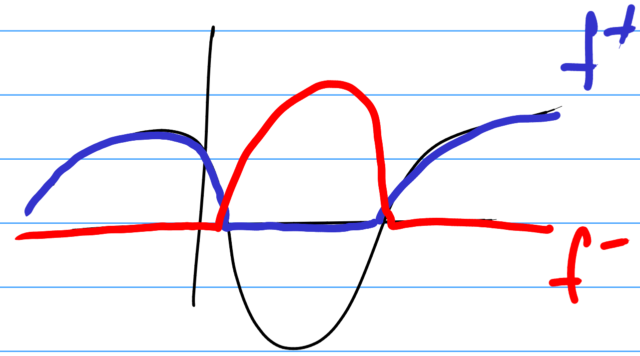
Teorema Si f es no negativa y medible entonces existen funciones $0 \le f_1 \le f_2 \le \dots$, todos simples, tales que
$$\lim_{n\to\infty} f_n(x) = f(x) \text{ para todo } x$$



Entonces definimos $\int f \, d\mu = \lim_{n\to\infty} \int f_n(x) \, d\mu(x)$.

Finalmente para f medible (no nec positiva)
sea

no negativa $\begin{cases} f^+ = \max\{f, 0\} \\ f^- = \max\{-f, 0\} \end{cases}$



notar $f = f^+ - f^-$

Def $\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$

Si da $\infty - \infty$ decimos que no está definida

Variables Aleatorias

Sea P una medida de probabilidad en $(\mathcal{E}, \mathcal{B})$
los conjuntos $B \in \mathcal{B}$ se llaman eventos
puntos $e \in \mathcal{E}$ " " resultados
$P(B)$ es la probabilidad de B.

Def Una función medible $X : \mathcal{E} \to \mathbb{R}$ se llama variable aleatoria.

La medida P_X definida por

$$P_X(A) = P\left(\{e \in \mathcal{E} : X(e) \in A\}\right)$$ se llama medida de distribución de X y se escribe

$$X \sim P_X$$

$$P_X(A) = P(X \in A) = P(\{\varepsilon \in \mathcal{E} : X(\varepsilon) \in A\})$$

en donde A es un boreliano.

La función de distribución acumulada de X es

$$F_X(x) = P(X \leq x) = P_X((-\infty, x]) = P(\{e \in \mathcal{E} : X(e) \leq x\})$$

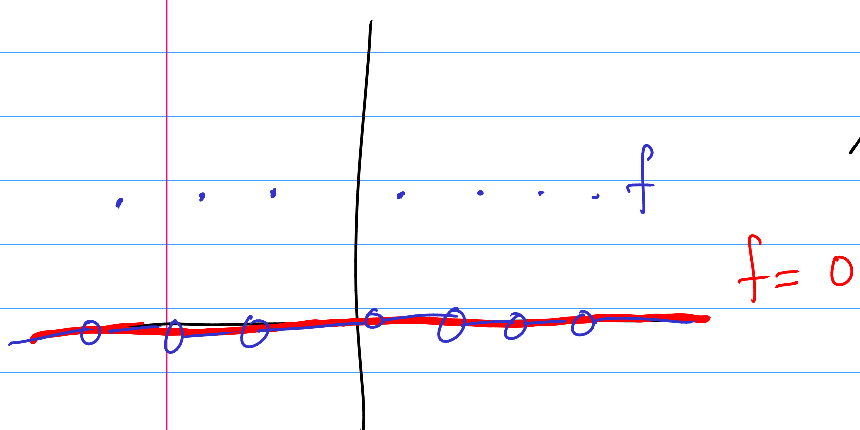
Conjuntos nulos

Sea μ una medida en $(\mathcal{X}, \mathcal{B})$. Un conjunto N se llama nulo si $\mu(N) = 0$.

Si una proposición es cierta para todo x en $\mathcal{X} - N$ donde N es nulo se dice que la proposición es cierta "casi en todas partes" a.e.

Ej $\quad f = 0 \quad$ a.e. $\mu \iff f(x) = 0$ salvo en $x \in N$ nulo

$$\mu\{x : f(x) \neq 0\} = 0$$



Si P es una medida de probabilidad y una proposición es cierta casi en todas partes (P), se dice que la proposición es cierta con probabilidad 1 wp1

Hecho útiles sobre funciones medibles

① $f = 0$ a.e. $\Rightarrow \int f \, d\mu = 0$

② $f \geq 0$ y $\int f \, d\mu = 0 \Rightarrow f = 0$ a.e.

③ si $f = g$ a.e. entonces $\int f \, d\mu = \int g \, d\mu$
si cualquiera de ellas existe.

Densidades

Def Sean P y μ medidas en $(\mathcal{X}, \mathcal{A})$
P se llama absolutamente continua con respecto a μ si

$$\mu(N) = 0 \Rightarrow P(N) = 0$$

escribimos $P \ll \mu$

Teorema (Radon-Nikodym):
Si $P \ll \mu$ entonces existe una función no negativa f tal que

$$P(A) = \int_A f \, d\mu = \int f \, 1_A \, d\mu$$

f se llama densidad de P con respecto a μ.

Ejemplo. Variables absolutamente continuas
Si X tiene densidad p con respecto a Lebesgue decimos que X, o P_X o F_X tiene densidad p

en cuyo caso $F_X(x) = \int_{-\infty}^{x} p(s)\,ds$ $\quad p(x) = F_X'(x)$

Ejemplo. Variables discretas

Sea $\mathcal{X}_0$ un subconjunto numerable de $\mathbb{R}$ y sea μ conteo definido por

$$\mu(B) = \#(\mathcal{X}_0 \cap B) \qquad B \text{ boreliano}$$

en este caso $\int f\,d\mu = \sum_{s \in \mathcal{X}_0} f(s)$

Supongamos que X es una variable aleatoria y $P(X \in \mathcal{X}_0) = 1$ entonces X se llama variable aleatoria discreta

La densidad p de P_X con respecto a μ satisface

$$P_X(A) = P(X \in A) = \int_A p \, d\mu = \sum_{s \in A \cap \mathcal{X}_0} p(s)$$

tomando $A = \{x\}$ con $x \in \mathcal{X}_0$ se tiene

$P(X = x) = p(x)$ generalmente llamada función de probabilidad puntual de la variable discreta.

Resumen

X se dice absolutamente continua si $P_X << \text{Lebesgue}$ y en este caso

$$P(X \in A) = \int_A p(x) \, dx$$

X se dice discreta si $P_X <<$ medida de conteo en algún $\mathcal{X}_0$ numerable. En este caso P_X tiene densidad

$$p(x) = P(X = x) \quad \text{y} \quad P(X \in A) = \sum_{x \in A \cap \mathcal{X}_0} p(x)$$

Esperanza

Def Sea X una v.a. en $(\mathcal{E}, \mathcal{B}, P)$.

$$EX = \int X \, dP = \int X \, dP_X$$

(en part $= \int x f(x) \, dx$ cont.

$= \sum x \, p(x)$ discreto)

Resultados útiles

1) Si $Y = f(X)$, $EY = \int f(x)\, dP_X(x)$

2) Si P_X tiene densidad p con respecto a μ

$$\int f\, dP_X = \int f(x)\, p(x)\, d\mu(x)$$

$$\left(\begin{array}{lll} \text{caso continuo} & Ef(X) & = \int f(x)\, p(x)\, dx \\ \text{discreto} & & = \sum_x f(x)\, p(x) \end{array} \right)$$

3) $E(aX+bY) = a\,EX + b\,EY$ linealidad de integrales

4) Si $X \leq Y$ entonces $EX < EY$ salvo que $X = Y$ a.e. en cuyo caso $EX = EY$.