Probabilidade e Inferência Estatística I

Aula 1 - 05 de março de 2013 Luiz Gustavo Esteves

- Entre 13 e 14 listas de exercícios. As listas serão disponibilizadas na pasta 13, no xerox do bloco B.

Probabilidade?

- Uma medida de incerteza
- um limite de frequências relativas
- física (massa, volume, área, ...)

Experimento	Espaço Amostral (Ω)
Lançamento de um dado	$\{1,2,3,4,5,6\}$
Campeonato paulista	$\{PAL, SP, PP, COR, \dots\}$ (20 elementos)
Temperatura	$[10,40]$. um dos conjuntos possíveis para Ω
Mega Sena	$\{A \subseteq \{1,\dots,60\} : \lvert A\rvert = 6\}$

Evento

$\{2,4,6\}$ $\{3,4,5,6\}$

$\{SP, COR, PAL\}$ $\{PP, GUA\}$

$[35,50]$, $[-10,30]$, $\{28,7\}$

$\{\{1,2,3,4,5,6\}\}$ $\{\{1,2,3,4,5,7\},\{1,2,3,4,5,8\},\dots\}$ considerando que o premiado será $\{1,2,3,4,5,6\}$

Conjunto: coleção de objetos distintos. Denotaremo-os por letras maiúsculas A, B, C, ..., com exceção do espaço amostral, que será denotado por Ω (também denotado por S por alguns autores).

Relação de Pertinência: $\omega \in A$, $\omega \notin A$ Conjunto vazio: $\emptyset$ $A^c = \{\omega \in \Omega : \omega \notin A\}$

$A, B \to A \cup B$ união de conjuntos A e B

$$A_1 \cup A_2 \cup \ldots \cup A_n = \bigcup_{i=1}^{n} A_i$$

$$A_1 \cup A_2 \cup \ldots = \bigcup_{i=1}^{\infty} A_i$$

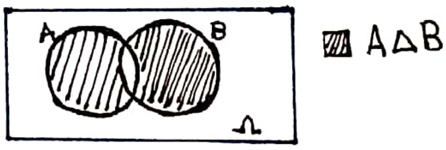
De modo mais geral, Λ conjunto de índices: $\bigcup_{\lambda \in \Lambda} A_\lambda$

$A \cap B$

$$A_1 \cap A_2 \cap \ldots \cap A_n = \bigcap_{i=1}^{n} A_i$$

$$A_1 \cap A_2 \cap \ldots = \bigcap_{i=1}^{\infty} A_i$$

De modo mais geral, Λ conjunto de índices: $\bigcap_{\lambda \in \Lambda} A_\lambda$



$A \triangle B = (A \cap B^c) \cup (A^c \cap B)$

Diferença simétrica

$A \triangle B = \{x \in \Omega : x \in A \cup B \text{ e } x \notin A \cap B\}$

A_1, A_2 são disjuntos se $A_1 \cap A_2 = \emptyset$. (mutuamente exclusivos)

Sequência de Conjuntos: $A_1, A_2, A_3, A_4, \ldots$

$(A_n)_{n \geq 1}$ sequência de conjuntos

Exemplo:

$\Omega = \mathbb{N}$

1. $A_n = \{1, \ldots, n\}$
2. $B_n = \{0, n\}$
3. $C_n = \{2n, 2n+2, 2n+4, \ldots\}$
4. $D_n = \begin{cases} \{0, n\}, & n \text{ par} \\ \{1, n\}, & n \text{ ímpar} \end{cases}$
5. $E_n = \begin{cases} \{1, \ldots, n\}, & n \text{ ímpar} \\ \{n\}, & n \text{ par} \end{cases}$

- Dizemos que uma sequência de conjuntos $(A_n)_{n\geq 1}$, é monótona se $A_n \subseteq A_{n+1}$, ou $A_n \supseteq A_{n+1}$, $\forall n \geq 1$.

- Se $A_n \subseteq A_{n+1}$, $\forall n\geq 1$ dizemos que A_n é uma sequência não-decrescente (Notação: $A_n \uparrow$).

- Se $A_n \supseteq A_{n+1}$, $\forall n \geq 1$ dizemos que A_n é uma sequência não-crescente. (Notação: $A_n \downarrow$)

- Quando $A_n\uparrow$, dizemos que existe $\lim\limits_{n\to\infty} A_n = \bigcup\limits_{n=1}^{\infty} A_n$.

  No exemplo 1., $\lim\limits_{n\to\infty} A_n = \mathbb{N}/\{0\}$

- Quando $A_n\downarrow$, dizemos que existe $\lim\limits_{n\to\infty} A_n = \bigcap\limits_{n=1}^{\infty} A_n$.

  No exemplo 3., $\lim\limits_{n\to\infty} A_n = \emptyset$.

Recordemos que

$$\underbrace{\inf\{a_n, a_{n+1}, \dots\}}_{b_n\uparrow} \leq a_n \leq \underbrace{\sup\{a_n, a_{n+1}, \dots\}}_{c_n\downarrow}$$

$$\lim b_n = \sup_{n\geq 1} \inf_{k\geq n} a_k \qquad \lim c_n = \inf_{n\geq 1} \sup_{k\geq n} a_k$$

Da maneira análoga, podemos raciocinar para sequência de eventos. Notemos que

$$\underbrace{A_n \cap A_{n+1} \cap A_{n+2} \cap \dots}_{B_n} \subseteq A_n \subseteq \underbrace{A_n \cup A_{n+1} \cup A_{n+2} \cup \dots}_{C_n}$$

$B_n \subseteq B_{n+1}$ $\qquad$ $C_n \supseteq C_{n+1}$ $\qquad$. Logo $B_n\uparrow$ e $C_n\downarrow$.

Como $(B_n)_{n\geq 1}$ é monótona, $\exists \lim\limits_{n\to\infty} B_n = \bigcup\limits_{n=1}^{\infty} B_n = \bigcup\limits_{n=1}^{\infty}\bigcap\limits_{k=n}^{\infty} A_k = \liminf\limits_{n\to\infty} A_n$.

Da mesma forma, $\exists \lim\limits_{n\to\infty} C_n = \bigcap\limits_{n=1}^{\infty} C_n = \bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty} A_k = \limsup\limits_{n\to\infty} A_n$.

$B_n \subseteq A_n \subseteq C_n$

Se $\liminf A_n = \limsup A_n$, dizemos que existe

$$\lim_{n\to\infty} A_n = \liminf A_n = \limsup A_n$$

- Interpretação

$\liminf A_n$: conjunto de pontos que estão em todos os A_n, exceto em um número finito deles.

$\limsup A_n$: está em todos os A_n a partir de um ponto.

Exemplos:

1. $A_n = \{1, \dots, n\}$

$\liminf A_n = \mathbb{N} / \{0\}$

$\limsup A_n = \mathbb{N} / \{0\}$

2. $B_n = \{0, n\}$

$\liminf B_n = \{0\}$

$\limsup B_n = \{0\}$

3. $C_n = \{2n, 2n+2, 2n+4, \dots\}$

$\liminf C_n = \emptyset$

$\limsup C_n = \emptyset$

4. $D_n = \begin{cases} \{0, n\}, & n \text{ par} \\ \{1, n\}, & n \text{ ímpar} \end{cases}$

$\liminf D_n = \emptyset$

$\limsup D_n = \{0, 1\}$

5. $E_n = \begin{cases} \{1, \dots, n\}, & n \text{ ímpar} \\ \{n\}, & \text{par} \end{cases}$ – de um à n e não conjunto n. ímpares

$\liminf E_n = \emptyset$

$\limsup E_n = \mathbb{N} / \{0\}$

Probabilidade e Inferência Estatística I

Aula 2 - 07 de março de 2013 Luiz Gustavo Esteves

Primeira lista - já disponível

Aula 1: Probabilidade, Teoria dos conjuntos, Sequência de eventos

Matematicamente, podemos encarar probabilidade como uma função $P: \mathcal{F} \to [0,1]$
$E \mapsto P(E)$

$\mathcal{P}(\Omega)$: conjunto de todos os subconjuntos de Ω

$P(A_1 \cup A_2 \cup \ldots) = \sum P(A_n)$. Livro do Royden (exemplos em que a eq. acima não é válida)

$P(\Omega)=1$
$P(A^c) = 1 - P(A)$
$P(A_1 \cup A_2) = P(A_1) + P(A_2)$, para $A_1 \cap A_2 = \emptyset$.

Agora, dado $\Omega \neq 0$, temos:

Definição: Dizemos que uma classe $\mathcal{F}$ de subconjuntos de Ω é uma álgebra se

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. $A_1, A_2 \in \mathcal{F} \Rightarrow A_1 \cup A_2 \in \mathcal{F}$.

Exemplos:

1. Ω qualquer

$\mathcal{P}(\Omega) = \{A \subseteq \Omega\}$ Temos que $\Omega \in \mathcal{P}(\Omega)$
. $A \in \mathcal{P}(\Omega) \Rightarrow A^c \in \mathcal{P}(\Omega)$
. $A_1 \cup A_2 \in \mathcal{P}(\Omega)$

Logo, $\mathcal{P}(\Omega)$ é álgebra

2. $\Omega = \mathbb{N}$

$\mathcal{F} = \{A \subseteq \mathbb{N} : A$ finito ou A^c é finito$\}$

1. $\Omega \in \mathcal{F}$

Como $\Omega^c = \emptyset$, finito, segue que $\Omega \in \mathcal{F}$

2. $A \in \mathcal{F} \Rightarrow \begin{cases} A \text{ finito} \Rightarrow (A^c)^c \text{ finito} \Rightarrow A^c \in \mathcal{F} \\ A^c \text{ finito} \Rightarrow A^c \in \mathcal{F} \end{cases}$

3. Se $A_1, A_2 \in \mathcal{F}$

- A_1 e A_2 são finitos $\Rightarrow A_1 \cup A_2$ finito $\Rightarrow A_1 \cup A_2 \in \mathcal{F}$.
- Suponhamos que A_1 não é finito. Então, A_1^c finito.

$A_1^c \cap A_2^c$ é finito e, portanto, $A_1^c \cap A_2^c \in \mathcal{F} \Rightarrow$

$\Rightarrow (A_1^c \cap A_2^c)^c \in \mathcal{F} \Rightarrow A_1 \cup A_2 \in \mathcal{F}$.

Logo $\mathcal{F}$ é uma álgebra.

Consequências:

1. $\emptyset \in \mathcal{F}$
2. $A_1 \cap A_2 \in \mathcal{F}$
3. $A_1, A_2, \dots, A_n \in \mathcal{F} \Rightarrow A_1 \cup A_2 \cup \dots \cup A_n \in \mathcal{F}$.

Definição: Uma classe $\mathcal{F}$ de subconjuntos de Ω ($\Omega \neq \emptyset$) é chamada σ-álgebra se:

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. $(A_n)_{n \geq 1}$, tal que $A_n \in \mathcal{F}, \forall n \geq 1 \Rightarrow \bigcup_{m=1}^{\infty} A_m \in \mathcal{F}$.

Exemplo:

1. Ω

$\mathcal{F}_0 = \{\emptyset, \Omega\}$

$\mathcal{F}_0$ é σ-álgebra de subconjuntos de Ω (σ-álgebra trivial)

2. Ω

$\mathcal{F} = \mathcal{P}(\Omega)$

$\mathcal{F}$ é σ-álgebra de subconjuntos de Ω.

3. $A \subseteq \Omega$, $A \neq \emptyset$

$\mathcal{F} = \{\emptyset, A, A^c, \Omega\} \Rightarrow$ menor de todas as σ-álgebras do qual A é elemento

$\downarrow$

Menor σ-álgebra que tem A como elemento.

que contem $\{A\}$.

4. $\Omega = \mathbb{N}$

$\mathcal{F} = \{A \subseteq \mathbb{N} : A \text{ ou } A^c \text{ é finito}\}$

$A_n = \{2n\}$, $n \geq 1$

$A_n \in \mathcal{F}$, $\forall n \geq 1$

$\underbrace{\bigcup_{n=1}^{\infty} A_n}_{\text{pares}}$ e $\underbrace{\left(\bigcup_{n=1}^{\infty} A_n\right)^c}_{\text{impares}}$ NÃO são finitos, portantos $\bigcup_{n=1}^{\infty} A_n \notin \mathcal{F}$.

Logo, $\mathcal{F}$ não é σ-álgebra.

Comentários: $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{F}$, $\bigcap_{n=1}^{\infty} A_n^c \in \mathcal{F}$

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \in \mathcal{F}$$

$$\liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \in \mathcal{F}$$

Sendo $\mathcal{F}_1$ e $\mathcal{F}_2$ duas σ-álgebras de subconjuntos de Ω, $\mathcal{F}_1 \cap \mathcal{F}_2$ é σ-álgebra.

1. $\Omega \in \mathcal{F}_1 \cap \mathcal{F}_2$?

$$\left.\begin{array}{l} \Omega \in \mathcal{F}_1 \\ \Omega \in \mathcal{F}_2 \end{array}\right\} \Rightarrow \Omega \in \mathcal{F}_1 \cap \mathcal{F}_2$$

2. $A \in \mathcal{F}_1 \cap \mathcal{F}_2 \overset{?}{\Rightarrow} A^c \in \mathcal{F}_1 \cap \mathcal{F}_2$

$$A \in \mathcal{F}_1 \cap \mathcal{F}_2 \Rightarrow \left.\begin{array}{l} A \in \mathcal{F}_1 \\ A \in \mathcal{F}_2 \end{array}\right\} \Rightarrow \left.\begin{array}{l} A^c \in \mathcal{F}_1 \\ A^c \in \mathcal{F}_2 \end{array}\right\} \Rightarrow A^c \in \mathcal{F}_1 \cap \mathcal{F}_2$$

3. $A_1, A_2, A_3, \ldots, A_n, \ldots \in \mathcal{F}_1 \cap \mathcal{F}_2 \overset{?}{\Rightarrow} \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_1 \cap \mathcal{F}_2$.

$$A_1, A_2, A_3, \ldots, A_n, \ldots \in \mathcal{F}_1 \cap \mathcal{F}_2 \Rightarrow \left.\begin{array}{l} A_1, A_2, \ldots, A_n, \ldots \in \mathcal{F}_1 \\ A_1, A_2, \ldots, A_n, \ldots \in \mathcal{F}_2 \end{array}\right\} \Rightarrow \left.\begin{array}{l} \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_1 \\ \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_2 \end{array}\right\} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_1 \cap \mathcal{F}_2.$$

Toda coleção enumerável ou não-enumerável de σ-álgebras, a intersecção dessa também será uma σ-álgebra.

No entanto, $\mathcal{F}_1 \cup \mathcal{F}_2$ não é σ-álgebra.

Ex: $\Omega = \{1,2,3,4,5,6\}$

$\mathcal{F}_1 = \{\emptyset, \Omega, \{1,2,3\}, \{4,5,6\}\}$

$\mathcal{F}_2 = \{\emptyset, \Omega, \{1,3,5\}, \{2,4,6\}\}$

$\mathcal{F}_1 \cup \mathcal{F}_2 = \{\emptyset, \Omega, \{1,2,3\}, \{4,5,6\}, \{1,3,5\}, \{2,4,6\}\}$

Notemos que

$\{1,2,3\} \in \mathcal{F}_1 \cup \mathcal{F}_2$. No entanto, $\{1,2,3\} \cup \{1,3,5\} = \{1,2,3,5\} \notin \mathcal{F}_1 \cup \mathcal{F}_2$.

$\{1,3,5\} \in \mathcal{F}_1 \cup \mathcal{F}_2$

Logo $\mathcal{F}_1 \cup \mathcal{F}_2$ não é σ-álgebra de conjuntos de Ω.

- Seja $\Omega \neq \emptyset$ qualquer.
- Seja $\mathcal{E}$ uma classe de subconjuntos qualquer de Ω.
- Note que sempre existe ao menos uma σ-álgebra de subconjuntos de Ω que contém $\mathcal{E}$.
- De fato, $\mathcal{P}(\Omega)$ é σ-álgebra e $\mathcal{E} \subseteq \mathcal{P}(\Omega)$.

Exemplo:

1. $\Omega = \{1,2,3,4,5,6\}$

$\mathcal{E}_1 = \{\{1,2,3,4\}, \{5,6\}\}$ [redacted]

[redacted]

$\mathcal{E}_1 \subseteq \mathcal{P}(\Omega)$

$\mathcal{F}_1 = \{\emptyset, \Omega, \{1,2,3,4\}, \{5,6\}\}$

$\mathcal{F}_2 = \{\emptyset, \{1,2\}, \{3,4\}, \{5,6\}, \{1,2,3,4\}, \{1,2,5,6\}, \{3,4,5,6\}, \Omega\}$

$\mathcal{E}_2 = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$

$\mathcal{E}_2 \subseteq \mathcal{P}(\Omega)$

Para uma dada classe $\mathcal{C}$ de subconjuntos de Ω, seja

$\mathbb{F}(\mathcal{C})$: conjunto de todas as σ-álgebras de Ω que contém $\mathcal{C}$.

$\mathbb{F}(\mathcal{C}) \neq \emptyset$, pois $\mathcal{P}(\Omega) \in \mathbb{F}(\mathcal{C})$

Considere

$$\bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{C})} \mathcal{F} \subseteq \mathcal{F}, \forall \mathcal{F} \in \mathbb{F}(\mathcal{C}) \Rightarrow \bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{C})} \mathcal{F} \text{ é a menor } \sigma\text{-álgebra que contém } \mathcal{C}.$$

$\downarrow$

é uma σ-álgebra que contém $\mathcal{C}$

Nesse caso, $\bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{C})} \mathcal{F}$ é chamada a menor σ-álgebra que contém $\mathcal{C}$.

Alternativamente, é chamada também de "σ-álgebra gerada por $\mathcal{C}$".

Notação: $\sigma(\mathcal{C}) = \bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{C})} \mathcal{F}$.

Exemplos:

1. $\Omega = \{1,2,3,4,5,6\}$

   $\mathcal{C}_1 = \{\{1,2,3,4\},\{5,6\}\}$ $\quad \sigma(\mathcal{C}_1) = \{\emptyset, \Omega, \{1,2,3,4\}, \{5,6\}\}$

   $\mathcal{C}_2 = \{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}\}$ $\quad \sigma(\mathcal{C}_2) = \mathcal{P}(\Omega)$

2. Ω com ao menos 3 elementos

   (A_1, A_2, A_3) partição de Ω

   (isto é, $A_i \cap A_j = \emptyset$, $i \neq j$ e $A_1 \cup A_2 \cup A_3 = \Omega$)

   $\sigma(\{A_1, A_2, A_3\}) = \{\emptyset, \Omega, A_1, A_2, A_3, A_1 \cup A_2, A_1 \cup A_3, A_2 \cup A_3\}$

Em geral, para uma partição $(A_1, A_2, \ldots, A_n)$

$$\sigma(\{A_1, A_2, \ldots, A_n\}) = \{\bigcup_{j \in I} A_j : I \subseteq \{1, \ldots, n\}\}$$

4. $\Omega = \mathbb{R}$

$\mathcal{C} = \{ (-\infty, t] : t \in \mathbb{R} \}$

$\sigma(\mathcal{C})$: a menor σ-álgebra que contém $\mathcal{C}$ $\rightarrow$ $\mathcal{B}(\mathbb{R})$: σ-álgebra de Borel na reta
(borelianos da reta)

$\sigma(\mathcal{C})$

$\mathcal{C}_1 = \{ (-\infty, t) : t \in \mathbb{R} \}$

$\mathcal{C}_2 = \{ (a, b] : a, b \in \mathbb{R}, a < b \}$

Os borelianos podem ser gerados tanto por $\mathcal{C}$, $\mathcal{C}_1$ ou $\mathcal{C}_2$.

Vejamos para $\mathcal{C}$ e $\mathcal{C}_1$ a equivalência:

$$\underbrace{(-\infty, t)}_{\text{elemento de } \mathcal{C}_1} = \bigcup_{n=1}^{\infty} (-\infty, t - 1/n] \Rightarrow (-\infty, t) \in \sigma(\mathcal{C}), \forall t \in \mathbb{R} \Rightarrow$$

$\Rightarrow \sigma(\mathcal{C}) \supseteq \sigma(\mathcal{C}_1)$ (i)

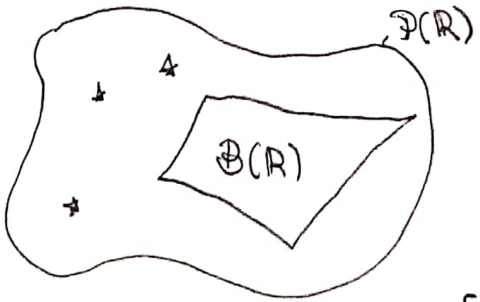
Analogamente

$$(-\infty, t] = \bigcap_{n=1}^{\infty} \left(-\infty, t + \frac{1}{n}\right) \Rightarrow (-\infty, t) \in \sigma(\mathcal{C}_1), \forall t \in \mathbb{R} \Rightarrow$$

$\Rightarrow \sigma(\mathcal{C}_1) \supseteq \sigma(\mathcal{C})$ (ii)

De (i) e (ii): $\sigma(\mathcal{C}) = \sigma(\mathcal{C}_1)$.

- Ver livro How to gamble if you must.
  How and How



$(a,b] = (-\infty, b] \cap (-\infty, a]^c$

Há também $\mathcal{B}(\mathbb{R}^2) = \sigma(\mathcal{E}')$, em que $\mathcal{E}' = \{(-\infty, a] \times (-\infty, b], a, b \in \mathbb{R}\}$

Probabilidade e Inferência Estatística I

Aula 3 - 08 de Março de 2013 — Luiz Gustavo Esteves

Seja Ω: espaço amostral
$\mathcal{F}$: σ-álgebra de subconjuntos de Ω

Ao par $(\Omega, \mathcal{F})$, dá-se o nome de espaço mensurável.

Definição: A função $\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$ é uma medida de probabilidade se:

1. $\mathbb{P}(\Omega) = 1$
2. $(A_n)_{n \geq 1}$, t.q. $A_n \in \mathcal{F}$, $\forall n \geq 1$, $A_i \cap A_j = \emptyset$, $i \neq j$

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) \quad (\Rightarrow \sigma\text{- aditividade da probabilidade})$$

Exemplos:

1. $\Omega = \mathbb{R}$

$\mathcal{F} = \mathcal{B}(\mathbb{R})$

$\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$

$$A \in \mathcal{F} \mapsto \mathbb{P}(A) = \mathbb{I}_A(0) = \begin{cases} 1, & 0 \in A \\ 0, & 0 \notin A \end{cases}$$

(i) $\mathbb{P}(\mathbb{R}) = \mathbb{I}_{\mathbb{R}}(0) = 1$ ✓

(ii) $(A_n)_{n \geq 1}$, $A_n \in \mathcal{B}(\mathbb{R})$, $\forall n \geq 1$; com $A_i \cap A_j = \emptyset$, $i \neq j$

$$0 \in \bigcup_{n=1}^{\infty} A_n \Rightarrow \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{I}_{\bigcup_{n=1}^{\infty} A_n}(0) = 1$$

Além disso, $\exists\, n_0 \in \mathbb{N}$ tal que $0 \in A_{n_0}$ e $0 \notin A_n$, $n \neq n_0$. Logo

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \mathbb{P}(A_{n_0}) = 1$$

$0 \notin \bigcup_{n=1}^{\infty} A_n \Rightarrow \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = 0 \Rightarrow 0 \in \bigcap_{n=1}^{\infty} A_n^c \Rightarrow 0 \in A_n^c, \forall n \geq 1 \Rightarrow$

$\Rightarrow 0 \notin A_n, \forall n \geq 1$ e, portanto, $\mathbb{P}(A_n) = 0, \forall n \geq 1$.

Logo, $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) = 0.$

Em geral, dado $(\Omega, \mathcal{F})$ e $\omega_0 \in \Omega$

$\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$, dado por

$\mathbb{P}(A) = \mathbb{I}_A(\omega_0)$, é probabilidade.

(medida de probabilidade degenerada/concentrada em ω_0).

2. $(\Omega, \mathcal{F})$

Sejam $\mathbb{P}_1, \mathbb{P}_2 : \mathcal{F} \to \mathbb{R}_+$, probabilidades e $\alpha \in (0,1)$

$\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$

$A \in \mathcal{F} \mapsto \mathbb{P}(A) = \alpha \mathbb{P}_1(A) + (1-\alpha)\mathbb{P}_2(A)$

(i) $\mathbb{P}(\Omega) = \alpha \mathbb{P}_1(\Omega) + (1-\alpha)\mathbb{P}_2(\Omega) =$

$= \alpha \cdot 1 + (1-\alpha) \cdot 1 =$

$= 1$ ✓

(ii) $(A_n)_{n\geq 1}$, $A_n \in \mathcal{F}, \forall n \geq 1$, com $A_i \cap A_j = \emptyset$,

$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \alpha \mathbb{P}_1(\bigcup_{n=1}^{\infty} A_n) + (1-\alpha)\mathbb{P}_2(\bigcup_{n=1}^{\infty} A_n) =$

$= \alpha \sum_{n=1}^{\infty} \mathbb{P}_1(A_n) + (1-\alpha) \sum_{n=1}^{\infty} \mathbb{P}_2(A_n) =$

Como $\mathbb{P}_1(\bigcup_{n=1}^{\infty} A_n)$ é um nº entre 0 e 1, $\sum \mathbb{P}(A_n)$ é convergente. Daí, $\alpha \sum \mathbb{P}(A_n) = \sum \alpha \mathbb{P}(A_n)$

$= \sum_{n=1}^{\infty} \alpha \mathbb{P}_1(A_n) + \sum_{n=1}^{\infty} (1-\alpha)\mathbb{P}_2(A_n =$

$= \sum_{n=1}^{\infty} \alpha \mathbb{P}_1(A_n) + (1-\alpha)\mathbb{P}_2(A_n) =$

$= \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

3. Como no exemplo 2, defina:

$$\mathbb{P}'(A) = \mathbb{P}_1(A) + \mathbb{P}_2(A)$$

i. $\mathbb{P}'(\Omega) = \mathbb{P}_1(\Omega) + \mathbb{P}_2(\Omega) = 1 + 1 \neq 1$ ×

Note que $\mathbb{P}'$ atende a condição (ii).

4. Como no exemplo 2, defina

$$\mathbb{P}''(A) = \min\{\mathbb{P}_1(A), \mathbb{P}_2(A)\}$$

i. $\mathbb{P}''(\Omega) = \min\{\mathbb{P}_1(\Omega), \mathbb{P}_2(\Omega)\} = \min\{1, 1\} = 1$ ✓

ii. NÃO é verdade!.

Ex: $\Omega = \mathbb{N}$ (Contradição)

$A_n = \{n-1\}, n \geq 1$

$\mathbb{P}_1(A) = \mathbb{I}_A(1)$

$\mathbb{P}_2(A) = \mathbb{I}_A(2)$

$\bigcup_{n=1}^{\infty} A_n = \Omega = \mathbb{N}$ e, portanto,

$$\mathbb{P}''\left(\bigcup_{n=1}^{\infty} A_n\right) = 1$$

No entanto, $\mathbb{P}''(A_n) = 0, \forall n \geq 1$. Assim,

$$\mathbb{P}''\left(\bigcup_{n=1}^{\infty} A_n\right) \neq \sum_{n=1}^{\infty} \mathbb{P}''(A_n)$$

5. $(\Omega, \mathcal{F})$

$\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$ probabilidade

Seja $B \in \mathcal{F}$, tal que $\mathbb{P}(B) > 0$,

Seja $\mathbb{P}_B : \mathcal{F} \to \mathbb{R}_+$ dada por

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$ (medida de probabilidade condicional, dado B)

i. $\mathbb{P}_B(\Omega) = \dfrac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$ ✓

ii. $(A_n)_{n \geq 1}$, $A \in \mathcal{F}\ \forall n \geq 1$, com $A_i \cap A_j = \emptyset, i \neq j$.

$$\mathbb{P}_B\left(\bigcup_{n=1}^{\infty} A_n\right) = \frac{\mathbb{P}\left(\left(\bigcup_{n=1}^{\infty} A_n\right) \cap B\right)}{\mathbb{P}(B)} = \frac{\mathbb{P}\left(\bigcup_{n=1}^{\infty} (A_n \cap B)\right)}{\mathbb{P}(B)} = \frac{\sum_{n=1}^{\infty} \mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_{n=1}^{\infty} \mathbb{P}_B(A_n)$$

Ao trio $(\Omega, \mathcal{F}, \mathbb{P})$ dá-se o nome de espaço de probabilidade, ou modelo probabilístico.

Ω Exp. (Dados)

$\mathbb{P}(A_1) \longrightarrow \mathbb{P}(A_1 | \text{Dados})$

$\mathbb{P}(A_2) \longrightarrow \mathbb{P}(A_2 | \text{Dados})$

$\mathbb{P}(A_3) \longrightarrow \mathbb{P}(A_3 | \text{Dados})$

$\mathbb{P}(D) \longrightarrow \mathbb{P}(D | \text{Dados})$

A medida de probabilidade condicional é útil na abordagem bayesiana.

$(\Omega, \mathcal{F}, P)$ Espaço de probabilidade

Propriedades.

i. $P(\emptyset) = 0$

$$\Omega = \Omega \cup \emptyset \cup \emptyset \cup \emptyset \cup \dots$$

$$P(\Omega \cup \emptyset \cup \emptyset \cup \dots) = \mathbb{P}(\Omega) + \sum_{n=2}^{\infty} \mathbb{P}(\emptyset) \Rightarrow \mathbb{P}(\emptyset) = 0.$$

ii. $A_1, \dots, A_n \in \mathcal{F}$, com $A_i \cap A_j = \emptyset$, $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$

$B_i = A_i$, $i = 1, 2, \dots, n$

$B_i = \emptyset$, $i = n+1, \dots$

$B_i \cap B_j = \emptyset$, $i \neq j$

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{n=1}^{\infty} \mathbb{P}(B_i) \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$$

iii. $\mathbb{P}(A^c) = 1 - P(A)$

Usando (ii), tomamos

$B_1 = A$ e $B_2 = A^c$

$$\mathbb{P}(\underbrace{A \cup A^c}_{\Omega}) = \mathbb{P}(A) + \mathbb{P}(A^c) \Rightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

iv. $A, B \in \mathcal{F}$ com $A \subseteq B$

$$\mathbb{P}(B) = \mathbb{P}((A \cap B) \cup (A^c \cap B)) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B) \overset{A \subseteq B}{\Rightarrow}$$

$$\mathbb{P}(B) = \mathbb{P}(A) + \underbrace{\mathbb{P}(A^c \cap B)}_{\geq 0} \Rightarrow \mathbb{P}(B) \geq \mathbb{P}(A)$$

v. $0 \leq \mathbb{P}(A) \leq 1, \forall A \in \mathcal{F}$

($\emptyset \subseteq A \subseteq \Omega$) usar da propriedade iv.

Defato, se $\emptyset \subseteq A$ então $\mathbb{P}(A) \geq \mathbb{P}(\emptyset) \Rightarrow \mathbb{P}(A) \geq 0$. *

Por outro lado, $A \subseteq \Omega$. Daí, $\mathbb{P}(A) \leq \mathbb{P}(\Omega) \Rightarrow \mathbb{P}(A) \leq 1$, **

De * e **, segue que $0 \leq \mathbb{P}(A) \leq 1, \forall A \in \mathcal{F}$.

vi. $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

Definimos

$$B_1 = A_1$$
$$B_2 = A_1^c \cap A_2$$
$$B_3 = A_1^c \cap A_2^c \cap A_3$$
$$\vdots$$
$$B_n = A_1^c \cap A_2^c \cap \ldots \cap A_{n-1}^c \cap A_n$$

Note que $B_i \cap B_j = \emptyset, i \neq j$.

Além disso,

$$\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$$

Assim,

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(B_n) =$$

$$= \underset{\mathbb{P}(A_1)}{\underset{\parallel}{\mathbb{P}(B_1)}} + \sum_{n=2}^{\infty} \underbrace{\mathbb{P}(A_1^c \cap \ldots \cap A_{n-1}^c \cap A_n)}_{\leq \mathbb{P}(A_n)} \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

vii. $A_1, A_2 \in \mathcal{F}$

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$$

$$A_1 \cup A_2 = A_1 \cup (A_1^c \cap A_2)$$

$$A_2 = (A_1 \cap A_2) \cup (A_1^c \cap A_2)$$

Assim,

$$\left.\begin{aligned} \mathbb{P}(A_1 \cup A_2) &= \mathbb{P}(A_1) + \mathbb{P}(A_1^c \cap A_2) \\ \mathbb{P}(A_2) &= \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1^c \cap A_2) \end{aligned}\right\} \Rightarrow \mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$$

Em geral, para $A_1, A_2, \dots, A_n$,

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{i=1}^{n} (-1)^{i+1} \sum_{\substack{I \subseteq \{1,\dots,n\} \\ |I| = i}} \mathbb{P}\left(\bigcap_{j \in I} A_j\right)$$

viii. $(A_n)_{n \geq 1}$ tais que $A_n \subseteq A_{n+1}, \forall n \geq 1$

$$A_n \uparrow \bigcup_{n=1}^{\infty} A_n$$

$$\mathbb{P}(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n) \quad \text{(continuidade da probabilidade)}$$

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n .$$

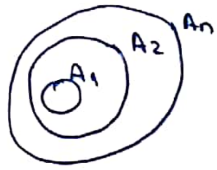
$B_1 = A_1$

$B_2 = A_1^c \cap A_2$

$B_3 = A_1^c \cap A_2^c \cap A_3$

$\vdots$

$B_n = A_1^c \cap A_2^c \cap \dots \cap A_{n-1}^c \cap A_n$



anéis

$$\mathbb{P}(\lim_{n\to\infty} A_n) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(B_n) = \lim_{K\to\infty} \sum_{n=1}^{K} \mathbb{P}(B_n) =$$

$$= \lim_{K\to\infty} [\mathbb{P}(B_1) + \mathbb{P}(B_2) + \dots + \mathbb{P}(B_K)] =$$

$\cdot\, B_1 = A_1, \quad A_2 = (A_1^c \cap A_2) \cup (A_1 \cap A_2) =$

$\quad = B_2 \cup A_1 \Rightarrow \mathbb{P}(B_2) = \mathbb{P}(A_2) - \mathbb{P}(A_1)$

$$= \lim_{K\to\infty} \{\mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1) + \mathbb{P}(A_3) - \mathbb{P}(A_2) + \dots + \mathbb{P}(A_K) - \mathbb{P}(A_{K-1})\} =$$

$$= \lim_{K\to\infty} \mathbb{P}(A_K) = \lim_{n\to\infty} \mathbb{P}(A_n) \quad \blacksquare$$

Agora, se

$$A_n \supseteq A_{n+1}, \forall n \geq 1 \Rightarrow A_n^c \subseteq A_{n+1}^c, \forall n \geq 1$$

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^c\right) = \mathbb{P}(\lim_{n\to\infty} A_n^c) = \lim_{n\to\infty} \mathbb{P}(A_n^c) = \lim_{n\to\infty} (1 - \mathbb{P}(A_n)) =$$

$$= 1 - \lim_{n\to\infty} \mathbb{P}(A_n)$$

Mas,

$$1 - \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\left(\bigcap_{n=1}^{\infty} A_n\right)^c\right) = 1 - \lim_{n\to\infty} \mathbb{P}(A_n) \Rightarrow$$

$$1 - \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = 1 - \lim_{n\to\infty} \mathbb{P}(A_n) \Rightarrow \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

Logo,

$$\mathbb{P}(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

σ- Aditividade

1. $\mathbb{P}(\Omega) = 1$

2. $(A_n)_{n\geq 1}$, disjuntos

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

$\Longleftrightarrow$

Aditividade - Finita

1. $\mathbb{P}(\Omega) = 1$

2'. $A_1, \dots, A_n$ disjuntos

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$$

3'. $(A_n)_{n\geq 1}$, $A_n \downarrow \emptyset$, $\lim_{n\to\infty} \mathbb{P}(A_n) = 0$

(continuidade para o vazio)

com essa, a volta é válida

Vemos que $(2') + (3') \Rightarrow (2)$ $\bigcup_{n=1}^{\infty} A_n = A_1 \cup A_2 \cup \dots \cup A_k \cup \left(\bigcup_{n=k+1}^{\infty} A_n\right)$

Seja $(A_n)_{n\geq 1}$, $A_n \in \mathcal{F}$, $\forall n \geq 1$, com $A_i \cap A_j = \emptyset$, $i \neq j$.

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(A_1 \cup A_2 \cup \dots \cup A_k \cup \left(\bigcup_{n=k+1}^{\infty} A_n\right)\right)$$

$$= \mathbb{P}(A_1) + \mathbb{P}(A_2) + \dots + \mathbb{P}(A_k) + \mathbb{P}\left(\bigcup_{n=k+1}^{\infty} A_n\right) \Rightarrow$$

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \underbrace{\sum_{i=1}^{k} \mathbb{P}(A_i)}_{a_k} + \underbrace{\mathbb{P}\left(\bigcup_{n=k+1}^{\infty} A_n\right)}_{b_k}$$

a_k $\downarrow$ seq. não decrescente

b_k $\downarrow$ seq. não crescente

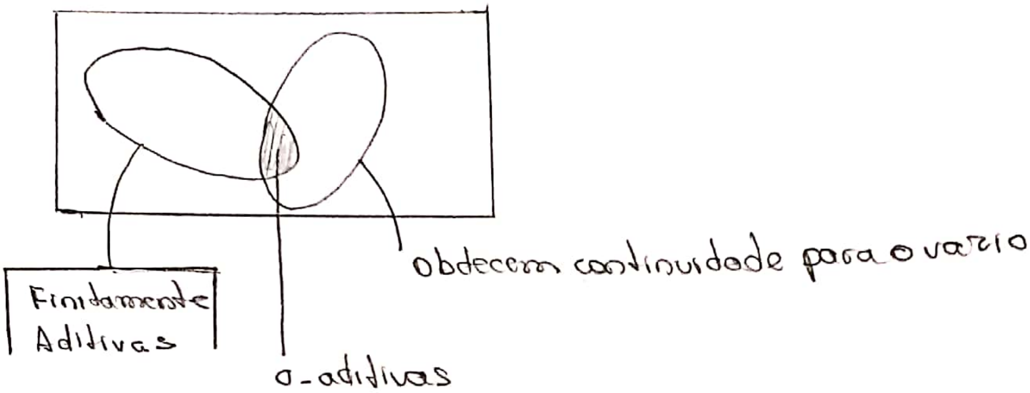
$$\Rightarrow \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \left(\lim_{k\to\infty} \sum_{i=1}^{k} \mathbb{P}(A_i)\right) + \lim_{k\to\infty} \mathbb{P}\left(\bigcup_{n=k+1}^{\infty} A_n\right) =$$

$$= \sum_{i=1}^{\infty} \mathbb{P}(A_i) + \underset{\text{por } (3')}{0}.$$

$\left\{ \bigcup_{n=k+1}^{\infty} A_n \downarrow \emptyset \Rightarrow \lim_{k\to\infty} \mathbb{P}\left(\bigcup_{n=k+1}^{\infty} A_n\right) = 0 \right.$

Logo: $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

Fixados Ω e $\mathcal{F}$:



Probabilidade e Inferência Estatística I

Aula 04 - 12/03/2013

Luiz Gustavo Esteves

- Modelos Discretos (Construção)

Ω: enumerável $\Omega = \{\omega_1, \omega_2, \dots\}$

$\mathcal{F}: \mathcal{P}(\Omega)$: todos os subconjuntos de Ω

$(p_n)_{n \geq 1}$ sequência de números não-negativos, tais que

$$\sum_{n=1}^{\infty} p_n = 1 \quad \left(\sum_{n=1}^{\infty} p_n < \infty\right)$$

Para $A \in \mathcal{P}(\Omega)$

$$\mathbb{P}(A) = \sum_{\{i \geq 1, \omega_i \in A\}} p_i = \sum_{i \geq 1} p_i \mathbb{I}_A(\omega_i), \text{ em que } \mathbb{I}_B(x) = \begin{cases} 1, & x \in B \\ 0, & x \notin B \end{cases}$$

Verificação que $\mathbb{P}$ é, de fato, probabilidade.

1. $\mathbb{P}(\Omega) = \sum_{i \geq 1} p_i \mathbb{I}_{\Omega}(\omega_i) = \sum_{i \geq 1} p_i = 1$

2. $A_1, A_2 \in \mathcal{P}(\Omega)$, com $A_1 \cap A_2 = \emptyset$

$$\mathbb{P}(A_1 \cup A_2) = \sum_{i \geq 1} P_i \, \mathbb{I}_{A_1 \cup A_2}(\omega_i) = \sum_{i \geq 1} P_i \left[\mathbb{I}_{A_1}(\omega_i) + \mathbb{I}_{A_2}(\omega_i) \right]$$

$$= \sum_{i \geq 1} \left[P_i \, \mathbb{I}_{A_1}(\omega_i) + P_i \, \mathbb{I}_{A_2}(\omega_i) \right] =$$

$$= \sum_{i \geq 1} P_i \, \mathbb{I}_{A_1}(\omega_i) + \sum_{i \geq 1} P_i \, \mathbb{I}_{A_2}(\omega_i) =$$

$$= \mathbb{P}(A_1) + \mathbb{P}(A_2).$$

3. $(A_n)_{n \geq 1}$, com $A_n \downarrow \emptyset$

Para todo $k \geq 1$, $\exists\, n_k \in \mathbb{N}$, t.q. $n \geq n_k$,

$$A_{n_k} \subseteq \{\omega_k, \omega_{k+1}, \omega_{k+2}, \dots\}$$

Logo, $\forall k \geq 1 : \mathbb{P}(A_{n_k}) \leq \mathbb{P}(\{\omega_k, \omega_{k+1}, \dots\}) = \sum_{i \geq k} P_i \Rightarrow$

$$\Rightarrow \limsup_k \mathbb{P}(A_k) \leq \lim_{k \to \infty} \mathbb{P}(\{\omega_k, \omega_{k+1}, \dots\})$$

$$= \lim_{k \to \infty} \sum_{i \geq k} P_i = 0$$ (pois a série $\{P_i\}$ é convergente)

Como $\liminf A_{n_k} = \limsup A_{n_k}$, então $\lim_{k \to \infty} \mathbb{P}(A_{n_k}) = 0$.

Logo: $\lim_{n \to \infty} \mathbb{P}(A_n) = 0$.

Então $\mathbb{P}$ é probabilidade.

Exemplos

1. $\Omega = \{1, 2, \dots\}$. $\mathcal{F} = \mathcal{P}(\Omega)$

$P_i = (1-p)^{i-1} p \qquad \mathbb{P}(A) = \sum_{i \geq 1} (1-p)^{i-1} p \, \mathbb{I}_A(i)$

$$\sum_{i=1}^{\infty} P_i = 1$$

$$\frac{p}{1-(1-p)} = 1$$

Logo, $\mathbb{P}$ é probabilidade.

2. $\Omega = \{3, 6, 9, 12, \dots\}$

$\omega_i = 3i$

$$p_i' = \frac{2^{i-1}}{(i-1)!} \qquad \sum p_i' = e^2$$

$$p_i = \frac{p_i'}{e^2} = \frac{e^{-2}\, 2^{i-1}}{(i-1)!}$$

O mesmo argumento vale se Ω é finito:

Se $\Omega = \{\omega_1, \dots, \omega_N\}$, basta especificar $p_1', \dots, p_N' \geq 0$.

$$p_i = \frac{p_i'}{p_1' + p_2' + \dots + p_N'}, \quad i = 1, \dots, N$$

No exemplo dado

$\Omega = \{1, 2, 3, 4, 5, 6\}$

$p_i' = i^2$

$$p_i = \frac{i^2}{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}$$ ("dando credibilidade maior aos números maiores")

- Quando Ω é finito com N elementos,

tomando $p_i = \frac{1}{N}$, $i = 1, \dots, N$

obtemos o Modelo Clássico.

Assim, para $A \in \mathcal{P}(\Omega)$

$$\mathbb{P}(A) = \frac{1}{N}|A| = \frac{|A|}{|\Omega|}$$

Exemplo.

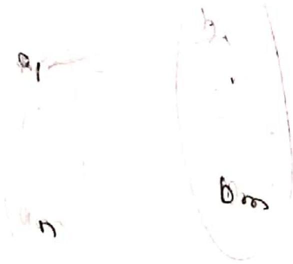
$\mathcal{A} = \{a_1, \dots, a_n\}$ $\quad$ Ω: conjunto de todas as funções de $\mathcal{A}$ em $\mathcal{B}$.

$\mathcal{B} = \{b_1, \dots, b_m\}$ $\quad$ $|\Omega| = m^n < \infty$

$\Omega, \mathcal{P}(\Omega), A \in \mathcal{P}(\Omega) \quad \mathbb{P}(A) = \frac{|A|}{m^n}$

a) Probabilidade de escolher f em Ω, t.q. f é injetora

$$A = \{f \in \Omega : f \text{ é injetora}\}$$



$$\mathbb{P}(A) = \frac{|A|}{m^n} = \frac{m(m-1)\dots(m-(n-1))}{m^n}$$

b) f é sobrejetora

$$A_i = \{f \in \Omega : b_i \in \operatorname{Im} f\},\ i = 1, \dots, m$$

B: f é sobrejetora

$$B = A_1 \cap A_2 \cap \dots \cap A_m$$

$$\mathbb{P}(B) = \mathbb{P}\left(\bigcap_{i=1}^{m} A_i\right) = \mathbb{P}\left(\left(\bigcup_{i=1}^{m} A_i^c\right)^c\right) = 1 - \mathbb{P}\left(\bigcup_{i=1}^{m} A_i^c\right) \quad (*)$$

Mas,

$$\mathbb{P}\left(\bigcup_{i=1}^{m} A_i^c\right) = \sum_{i=1}^{m} (-1)^{i+1} \sum_{\substack{D \subset \{1,\dots,m\} \\ |D| = i}} \mathbb{P}\left(\bigcap_{j \in D} A_j\right)$$

e,

$$\mathbb{P}(A_j^c) = \frac{(m-1)^n}{m^n},\ j = 1, \dots, m$$

$$\mathbb{P}(A_j^c \cap A_k^c) = \frac{(m-2)^n}{m^n},\ \forall\, i, j = 1, \dots, m \text{ com } i \neq j$$

Em geral,

$$\mathbb{P}\left(\bigcap_{j \in D} A_j^c\right) = \frac{(m - |D|)^n}{m^n}$$

Logo

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^c\right) = \sum_{i=1}^{m} (-1)^{i+1} \sum_{\substack{D \subset \{1,\dots,m\} \\ |D| = i}} \frac{(m-i)^n}{m^n} =$$

$$= \sum_{i=1}^{m} (-1)^{i+1} \binom{m}{i} \frac{(m-i)^n}{m^n}$$

Assim, a probabilidade procurada é

$$\mathbb{P}(B) = 1 - \sum_{i=1}^{m} (-1)^{i+1} \binom{m}{i} \frac{(m-i)^n}{m^n}$$

$\Omega = \mathbb{R}$

$\mathcal{F} = \mathcal{B}(\mathbb{R})$

$\mathbb{P}((a,b]) = \int_a^b f(t)\,dt$, onde $f: \mathbb{R} \to \mathbb{R}$ é tal que $\int_{\mathbb{R}} f(t)\,dt = 1$ $(<\infty)$

$$\mathbb{P}((a,b] \cup (c,d]) = \int_a^b f(t)\,dt + \int_c^d f(t)\,dt \;, \quad b<c$$

Seja $\mathcal{C} \subseteq \mathcal{P}(\mathbb{R})$ a classe de subconjuntos que são uniões finitas de intervalos.

Para $C = \bigcup_{i=1}^{k} (a_i, b_i] \in \mathcal{C}$ com $b_i < a_{i+1}$, $i = 1, \dots, k-1$.

$$\mathbb{P}(C) = \sum_{i=1}^{k} \int_{a_i}^{b_i} f(t)\,dt.$$

Para $B \in \mathcal{P}(\mathbb{R})$,

$$\mathbb{P}^*(B) = \inf\left\{ \sum_{i=1}^{\infty} \underbrace{\mathbb{P}(A_i)}_{\text{união finita de intervalos finitos}}, \text{ onde } A_i \in \mathcal{C} \text{ e } B \subset \bigcup_{i=1}^{\infty} A_i \right\}$$

Mas assim,

$$\mathbb{P}^*\left(\bigcup_{n=1}^{\infty} D_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}^*(D_n) \quad \text{(sub-}\sigma\text{-aditividade)}$$

Idéia: Restringir $\mathbb{P}^*$ ao conjunto

$$\mathcal{M} = \{ D \subseteq \mathbb{R} : \mathbb{P}^*(A) = \mathbb{P}^*(A \cap D) + \mathbb{P}(A \cap D^c), \forall A \subseteq \mathbb{R} \}$$

conjunto mensuráveis da reta

Já disponível a 2ª lista.

Probabilidade e Inferência Estatística I

Aula 05 - 19/03/2013 Luiz Gustavo Esteves

Probabilidade Condicional

$(\Omega, \mathcal{F}, \mathbb{P})$ $B \in \mathcal{F}$, $\mathbb{P}(B) > 0$.

· Definição: A probabilidade condicional de A dado B é definida por

$$\mathbb{P}(A/B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Exemplo:

1. Lançamento dado

$$\Omega = \{1,2,3,4,5,6\}, \; |\Omega| = 6$$
$$\mathcal{F} = \mathcal{P}(\Omega)$$
$$\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$$
$$F \in \mathcal{F} \mapsto \mathbb{P}(F) = \frac{|F|}{6}$$

$A = \{2,4,6\}$

$B = \{1,3,4,5,6\}$

$$\mathbb{P}(A/B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{4,6\})}{\mathbb{P}(B)} = \frac{2/6}{5/6} = \frac{2}{5}.$$

$$\mathbb{P}(B/A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{2/6}{3/6} = \frac{2}{3}$$

Da definição, temos

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B/A) \quad \text{(Regra do Produto)}$$

Se $A_1, A_2, \dots, A_n \in \mathcal{F}$, vale que

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1)\mathbb{P}(A_3 | A_1 \cap A_2) \dots \mathbb{P}(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

Exemplo. (Modelo de uma Pólya- Eggenberger)

10 brancas
6 verdes

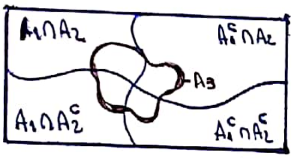
A_i : bola branca retirada na i-ésima extração, $i=1,2,3$.

a.

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2) =$$
$$= \frac{10}{16} \frac{12}{18} \frac{14}{20}$$

Se sai branca, devolve a retirada e adiciona mais duas brancas (o mesmo vale pra verde).

$P(A_3) = ?$



$$P(A_3) = P(A_1 \cap A_2 \cap A_3) + P(A_1^c \cap A_2 \cap A_3) + P(A_1 \cap A_2^c \cap A_3) + P(A_1^c \cap A_2^c \cap A_3) =$$

$$= P(A_1) P(A_2 | A_1) P(A_3 / A_1 \cap A_2) + P(A_1^c) P(A_2 / A_1^c) P(A_3 | A_1^c \cap A_2) +$$
$$+ P(A_1) P(A_2^c / A_1) P(A_3 / A_1 \cap A_2^c) + P(A_1^c) P(A_2^c / A_1^c) P(A_3 | A_1^c \cap A_2^c)$$

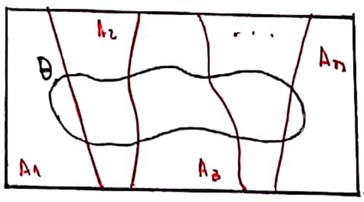
Fórmula da Probabilidade Total

Seja $A_1, \dots, A_n$ uma partição de Ω tal que $A_i \in \mathcal{F}$, $\forall i = 1, \dots, n$. ($A_i \cap A_j = \emptyset$, $i \neq j$, e $\bigcup_{i=1}^{n} A_i = \Omega$)

Seja $B \in \mathcal{F}$

$$P(B) = P(B \cap \Omega) = P(B \cap (\bigcup_{i=1}^{n} A_i)) =$$

$$= P(\bigcup_{i=1}^{n} (B \cap A_i)) = \sum_{i=1}^{n} P(B \cap A_i) \Rightarrow$$



$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B / A_i) P(A_i)$$

Teorema de Bayes. (Fórmula da Inversão)

$(\Omega, \mathcal{F}, \mathbb{P})$. Seja $A_1, \dots, A_n$, partição de Ω, com $A_i \in \mathcal{F}$, $i = 1, \dots, n$. Seja $B \in \mathcal{F}$. Então,

$$\mathbb{P}(A_i / B) = \frac{\mathbb{P}(B/A_i)\,\mathbb{P}(A_i)}{\sum_{j=1}^{n} \mathbb{P}(B/A_j)\,\mathbb{P}(A_j)}, \quad i = 1, \dots, n$$

$$\begin{array}{ccc} \mathbb{P}(A_1) & & \mathbb{P}'(A_1) \\ \mathbb{P}(A_2) & \text{Informação} & \mathbb{P}'(A_2) \\ \mathbb{P}(A_3) & \text{de Bayes} & \mathbb{P}'(A_3) \\ \vdots & & \vdots \\ \mathbb{P}(A_n) & & \mathbb{P}'(A_n) \end{array}$$

Exemplo.

5 bolas

? brancas
5-? verdes

A_i: a urna possui i bolas brancas, $i = 0, \dots, 5$.

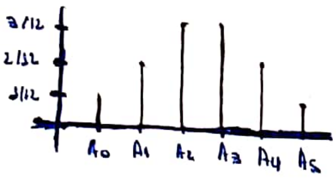
(0,5)
(1,4)
(2,3)
(3,2)
(4,1)
(5,0)

$\frac{1}{6}$

$\mathbb{P}(A_0) = \mathbb{P}(A_5) = 1/12$

$\mathbb{P}(A_1) = \mathbb{P}(A_4) = 2/12$

$\mathbb{P}(A_2) = \mathbb{P}(A_3) = 3/12$



Experimento que consiste em retirar 2 bolas da urna.

B_j: a amostra com j bolas brancas, $j = 0, 1, 2$.

$\mathbb{P}(A_i / B_2) = ?$

$$\mathbb{P}(B_2 / A_i) = \frac{\binom{i}{2}\binom{5-i}{0}}{\binom{5}{2}}$$

Pelo Teorema de Bayes

$$\mathbb{P}(A_i / B_2) = \frac{\mathbb{P}(B_2/A_i)\,\mathbb{P}(A_i)}{\sum_{j=0}^{5} \mathbb{P}(B_2/A_j)\,\mathbb{P}(A_j)}$$

$$\mathbb{P}(A_i / B_2) = \frac{\frac{\binom{0}{2}\binom{5}{0}}{\binom{5}{2}}\frac{1}{12}}{\frac{\binom{0}{2}\binom{5}{0}}{\binom{5}{2}}\frac{1}{12} + \frac{\binom{1}{2}\binom{4}{0}}{\binom{5}{2}}\frac{2}{12} + \frac{\binom{2}{2}\binom{3}{0}}{\binom{5}{2}}\frac{3}{12} + \frac{\binom{4}{2}\binom{1}{0}}{\binom{5}{2}}\frac{2}{12} + \frac{\binom{5}{2}\binom{0}{0}}{\binom{5}{2}}\frac{1}{12}} = S$$

$\mathbb{P}(A_0/B_2) = \mathbb{P}(A_1/B_2) = 0.$

$$\mathbb{P}(A_2/B_2) = \frac{\left[\binom{2}{2}\binom{5}{0}/\binom{5}{2}\right]\frac{3}{12}}{5} = \frac{3/120}{34/120} = \frac{3}{34}$$

$$\mathbb{P}(A_3/B_2) = \frac{9/120}{34/120} = \frac{9}{34}$$

$$\mathbb{P}(A_4/B_2) = \frac{12}{34} \quad e \quad \mathbb{P}(A_5/B_2) = \frac{10}{34}$$

Exemplo 2-

D: doente

$P(D) = 5\%$

E: resultado do exame é positivo

$\mathbb{P}(E/D) = 95\%$

$\mathbb{P}(E/D^c) = 1\%$

$$\mathbb{P}(D/E) = \frac{\mathbb{P}(E/D)\mathbb{P}(D)}{\mathbb{P}(E/D)\mathbb{P}(D) + \mathbb{P}(E/D^c)\mathbb{P}(D^c)} = \frac{95\% \cdot 5\%}{95\% \cdot 5\% + 1\% \cdot 95\%} \Rightarrow \mathbb{P}(D/E) = \frac{5}{6}$$

Quando

$\mathbb{P}(A/B) = \mathbb{P}(A)$, dizemos que o evento A é independente do evento B.

A é independente de B $\Rightarrow \mathbb{P}(A/B) = \mathbb{P}(A)$.

Vimos que

$$\mathbb{P}(A/B)\mathbb{P}(B) = \mathbb{P}(B/A)\mathbb{P}(A)$$

$$\Downarrow$$

$$\mathbb{P}(B/A) = \mathbb{P}(B)$$

$$\Downarrow$$

B é independente de A.

Quando A é independente de B (e, consequentemente, B é independente de A), dizemos que A e B são independentes.

Alternativamente,

A e B são independentes $\Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

Exemplo. Lançamento dado

$A = \{2,4,6\}$ $\quad \mathbb{P}(A) = 1/2$ $\quad \mathbb{P}(A\cap B) = 1/6 = 1/2 \cdot 1/3 = \mathbb{P}(A)\mathbb{P}(B)$

$B = \{5,6\}$ $\quad \mathbb{P}(B) = 1/3$ $\quad$ Logo, A e B são independentes

$C = \{1,3,4,5,6\}$ $\quad \mathbb{P}(C) = 5/6$

$\mathbb{P}(A\cap C) = \frac{2}{6} \neq \frac{1}{2}\frac{5}{6} = \mathbb{P}(A)\mathbb{P}(C)$. Logo A e C não são independentes.

Consequências:

A e B independentes $\Rightarrow$ A e B^c são independentes.

A^c e B " "

A^c e B^c " "

- A e B^c

$$\mathbb{P}(A\cap B^c) = \mathbb{P}(A) - \mathbb{P}(A\cap B) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)(1-\mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(B^c)$$

- A^c e B

$$\mathbb{P}(A^c\cap B) = \mathbb{P}(B) - \mathbb{P}(B\cap A) = \mathbb{P}(B) - \mathbb{P}(B)\mathbb{P}(A) = \mathbb{P}(B)(1-\mathbb{P}(A)) = \mathbb{P}(A^c)\mathbb{P}(B)$$

- A^c e B^c

$$\mathbb{P}(A^c\cap B^c) = 1-\mathbb{P}(A\cup B) = 1-\mathbb{P}(A)-\mathbb{P}(B)+\mathbb{P}(A\cap B) = 1-\mathbb{P}(A)-\mathbb{P}(B)+\mathbb{P}(A)\mathbb{P}(B) =$$
$$= \underbrace{1-\mathbb{P}(A)} - \mathbb{P}(B)(1-\mathbb{P}(A)) = \mathbb{P}(A^c) - \mathbb{P}(B)\mathbb{P}(A^c) = \mathbb{P}(A^c)(1-\mathbb{P}(B)) =$$
$$= \mathbb{P}(A^c)\mathbb{P}(B^c) \checkmark$$

Algumas Generalizações

Definição: $A_1, A_2, \dots, A_n$ são (coletivamente) independentes se

$$\forall B \subseteq \{1,\dots,n\},\ |B| \geq 2,\ \mathbb{P}\Big(\bigcap_{j\in B} A_j\Big) = \prod_{j\in B} \mathbb{P}(A_j)$$

número de igualdade a serem verificadas: $2^n - \binom{n}{1} - \binom{n}{0}$

DEFINIÇÃO: $A_1, \ldots, A_n$ são independentes dois a dois

$$\forall B \subseteq \{1,\ldots,n\} \text{ tal que } |B| = 2, \ \mathbb{P}\Big(\bigcap_{j \in B} A_j\Big) = \prod_{j \in B} \mathbb{P}(A_j)$$

$(\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j), \forall i = 1,\ldots,n, \forall j = 1,\ldots,n \text{ com } i \neq j)$

Exemplo: Lançamento de dois dados (um branco e outro vermelho)

$\Omega = \{1,\ldots,6\}^2 \qquad |\Omega| = 36$

$\mathcal{F} = \mathcal{P}(\Omega)$

$A \in \mathcal{F}, \ \mathbb{P}(A) = \frac{|A|}{36}$

A: Soma 7

B: Dado branco resulta em face par

C: Dado vermelho resulta valor maior que 4

D: Soma par

E: Dado vermelho resulta face par

F: Soma igual à 9

G: Dado branco resulta face maior que 3

A, B e C são independentes.

$\mathbb{P}(A) = 6/36 \qquad \mathbb{P}(A \cap B) = 3/36 \qquad \mathbb{P}(A \cap B \cap C) = 1/36$

$\mathbb{P}(B) = 18/36 \qquad \mathbb{P}(A \cap C) = 2/36$

$\mathbb{P}(C) = 12/36 \qquad \mathbb{P}(B \cap C) = 6/36$

Fazendo as checagens, conclui-se que A, B e C são independentes.

B, D, E

$\mathbb{P}(B) = 18/36 \qquad \mathbb{P}(B \cap D) = \mathbb{P}(B \cap E) = \mathbb{P}(D \cap E) = 1/4$

$\mathbb{P}(D) = 18/36 \qquad$ B, D e E são independentes 2 a 2

$\mathbb{P}(E) = 18/36 \qquad$ Mas $\mathbb{P}(B \cap D \cap E) = \frac{9}{36} \neq \frac{1}{8} = \mathbb{P}(B)\mathbb{P}(D)\mathbb{P}(E)$

B^c, G, F

$\mathbb{P}(B^c) = 18/36$

$\mathbb{P}(F) = 4/36$

$\mathbb{P}(G) = 18/36$

$\mathbb{P}(B^c \cap F \cap G) = 1/36 = \mathbb{P}(B^c)\,\mathbb{P}(F)\,\mathbb{P}(G)$

No entanto,

$\mathbb{P}(B^c \cap G) = 6/36 \neq 1/4 = \mathbb{P}(B^c)\,\mathbb{P}(G)$

Logo B^c, F e G NÃO são independentes.

Definição: Uma seqüência $(A_n)_{n \geq 1}$ é de eventos independentes se $A_1, A_2, \dots, A_n$ são independentes, $\forall n \geq 2$.

$(\Omega, \mathcal{F}, \mathbb{P})$

Seja $\mathcal{G} \subseteq \mathcal{F}$

$\mathcal{G}$ é uma classe de eventos independentes se

$$\forall \mathcal{H} \subseteq \mathcal{G},\ \mathcal{H} \text{ finito},\ \mathbb{P}\Big(\bigcap_{A \in \mathcal{H}} A\Big) = \prod_{A \in \mathcal{H}} \mathbb{P}(A)$$

Probabilidade e Inferência Estatística I

Aula 06 - **15**/03/2013 Luiz Gustavo Esteves

Optimal Statistical Decisions (M. H. DeGroot) Cap. 6.

Luiz Eduardo Monteya Belgado (1996 a 1998) C.A.B. Pereira (T.D.)

"Futebol, Bayes, Poisson, DeFinetti". Marcelo Leme Arruda (2000)

(D.M.) S. Wochsbu

www.chacedegol.com.br

Definição: $(\Omega, \mathcal{F}, \mathbb{P})$

$\mathcal{F}_1$, $\mathcal{F}_2$ duas σ-álgebras de subconjuntos de Ω, tais que $\mathcal{F}_i \subseteq \mathcal{F}$, $i=1,2$.

$\mathcal{F}_1$ e $\mathcal{F}_2$ são independentes se

$$\forall A_1 \in \mathcal{F}, \forall A_2 \in \mathcal{F}_2, \quad \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$$

Exemplo: $\Omega = \{1,2,3,4,5,6\}$

$\mathcal{F} = \mathcal{P}(\Omega)$

$\mathbb{P}(A) = \dfrac{|A|}{6}$

$\mathcal{F}_1 = \{\emptyset, \Omega, \{2,4,6\}, \{1,3,5\}\} = \sigma(\{2,4,6\}, \{1,3,5\})$

$\mathcal{F}_2 = \sigma(\{\{1,2\}, \{3,4\}, \{5,6\}\})$

$\mathcal{F}_3 = \sigma(\{\{1,2,3\}\})$

$A_1 = \{2,4,6\}$

$A_2 \in \mathcal{F}_2$ com $|A_2| = 2$

$$\mathbb{P}(A_1 \cap A_2) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = \mathbb{P}(A_1)\,\mathbb{P}(A_2)$$

$A_2 \in \mathcal{F}_2$ com $|A_2| = 4$,

$$\mathbb{P}(A_1 \cap A_2) = \frac{2}{6} = \frac{1}{2} \cdot \frac{2}{3}$$

Logo, $\mathcal{F}_1$ e $\mathcal{F}_2$ são independentes!

$\mathcal{F}_1$ e $\mathcal{F}_3$ não são independentes!

$A_1 = \{2,4,6\} \in \mathcal{F}_1$

$A_2 = \{1,2,3\} \in \mathcal{F}_3$

$$\mathbb{P}(A_1 \cap A_2) = \frac{1}{6} \neq \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A_1)\mathbb{P}(A_2).$$

Exemplo:

$\Omega = \mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$

$\mathcal{F} = \mathcal{P}(\Omega)$

Para $z \in \Omega$, $\mathbb{P}(\{z\}) = (1/2)^{|z|+1}$

$A \in \mathcal{F}$

$$\mathbb{P}(A) = \sum_{z \in A} \mathbb{P}(\{z\})$$

$\mathcal{F}_1 = \{\emptyset, \Omega, \mathbb{Z}_+^*, \mathbb{Z}_-^*\}$

$\mathcal{F}_2 = \sigma(\{\{-n,n\} : n \geq 1\}) = \{\bigcup_{n \in J} \{-n,n\} : J \subseteq \{1,2,3,\dots\}\}$

$$\mathbb{P}(\mathbb{Z}_+^*) = \sum_{z \in \mathbb{N}} \mathbb{P}(\{z\}) = \sum_{z \in \mathbb{N}} (1/2)^{z+1} = \frac{1}{2}$$

Seja $B \in \mathcal{F}_2$

$B = \bigcup_{n \in J} \{-n,n\}$, para algum $J \subseteq \{1,2,\dots\}$

$\mathbb{P}(B) = 2 \sum_{n \in J} \mathbb{P}(\{-n,n\}) = \sum_{n \in J} \left(\frac{1}{2}\right)^n$

$$\mathbb{P}(B) = \sum_{n \in J} \mathbb{P}(\{-n,n\}) = \sum_{n \in J} (1/2)^n$$

$$\mathbb{P}(A \cap B) = \mathbb{P}\left(\bigcup_{n \in J} \{n\}\right) = \sum_{n \in J} \mathbb{P}(\{n\}) = \sum_{n \in J} (1/2)^{n+1} = \frac{1}{2} \sum_{n \in J} (1/2)^n = \mathbb{P}(A)\mathbb{P}(B)$$

Logo, $\mathcal{F}_1$ e $\mathcal{F}_2$ são independentes.

Como $\forall A_1 \in \mathcal{F}_1$ e $\forall A_2 \in \mathcal{F}_2$

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$$

$\mathcal{F}_1$ e $\mathcal{F}_2$ são independentes!

Exemplo: $(\Omega, \mathcal{F}, \mathbb{P})$

$\mathcal{F}_0 = \{\emptyset, \Omega\}$

$\mathcal{F}_0$ é independente de qualquer σ-álgebra $\mathcal{F}' \subseteq \mathcal{F}$

$$\mathbb{P}(\emptyset \cap A_2) = \mathbb{P}(\emptyset)\mathbb{P}(A_2)$$
$$\mathbb{P}(\Omega \cap A_2) = \mathbb{P}(\Omega)\mathbb{P}(A_2)$$

Espaços - produtos

$(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$

$(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$

$\vdots$

$(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$

$$\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n$$

$\mathcal{C} = \{E_1 \times E_2 \times \dots \times E_n : E_i \in \mathcal{F}_i, i = 1, \dots, n\}$

$\mathcal{F} = \sigma(\mathcal{C})$

$\mathbb{P} : \mathcal{F} \to \mathbb{R}_+$ de modo que para $E_1 \times \dots \times E_n \in \mathcal{C}$,

$\mathbb{P}(E_1 \times \dots \times E_n) = \mathbb{P}(E_1) \dots \mathbb{P}(E_n)$

$((\Omega_n, \mathcal{F}_n, \mathbb{P}_n))_{n \geq 1}$

$\Omega = \Omega_1 \times \dots \times \Omega_n \times \dots$

$E_1 \times E_2 \times \dots \times E_n \times \Omega_{n+1} \times \Omega_{n+2} \times \dots \to$ cilindro n-dimensional de base $E_1 \times E_2 \times \dots \times E_{n-1} \times E_n$

$\mathcal{C} = \bigcup_{n=1}^{\infty} [E_1 \times E_2 \times \dots \times E_n \times \Omega_{n+1} \times \Omega_{n+2} \times \dots : E_i \in \mathcal{F}_i, i = 1, \dots, n\}$

$\mathbb{P}(E_1 \times E_2 \times \dots \times E_n \times \Omega_{n+1} \times \Omega_{n+2} \times \dots) = \mathbb{P}_1(E_1)\mathbb{P}(E_2) \dots \mathbb{P}(E_n)$

Exemplo: Lançamentos independentes de moedas

$\Omega_n = \{c, \bar{c}\}, \forall n \geq 1$

$\mathcal{F}_n = \mathcal{P}(\Omega_n), \forall n \geq 1$

$\mathbb{P}_n(\{c\}) = p_n, \ 0 < p_n < 1, \ n \geq 1$

1. Obter cara nos 3 primeiros lançamentos

$$\mathbb{P}(\{c\}\times\{c\}\times\{c\}\times\Omega_4\times\Omega_5\times\ldots)=$$
$$=\mathbb{P}_1(\{c\})\,\mathbb{P}_2(\{c\})\,\mathbb{P}_3(\{c\})=$$
$$=P_1P_2P_3.$$

2. Cara no 3º lançamento e coroa no 6º.

$$\mathbb{P}(\Omega_1\times\Omega_2\times\{c\}\times\Omega_4\times\Omega_5\times\{\bar{c}\}\times\Omega_7\times\Omega_8\times\ldots)=$$
$$=\mathbb{P}_1(\Omega_1)\,\mathbb{P}_2(\Omega_2)\,\mathbb{P}_3(\{c\})\,\mathbb{P}_4(\Omega_4)\,\mathbb{P}_5(\Omega_5)\,\mathbb{P}_6(\{\bar{c}\})=$$
$$=1\cdot1\cdot P_3\cdot1\cdot1\cdot(1-P_6)=P_3(1-P_6).$$

Variáveis Aleatórias

Exemplo: Eleições

$c_1,\ldots,c_K,c_{K+1}$ ↳ nulo ou branco

c_1 ↓ candidato 1

c_K ↓ candidato K

$$\Omega=\{c_1,c_2,\ldots,c_K,c_{K+1}\}^{2000}$$

$$X_j:\Omega\mapsto\mathbb{R}$$
$$(\omega_1,\ldots,\omega_{2000})\mapsto X_j((\omega_1,\ldots,\omega_{2000}))=\sum_{i=1}^{2000}\mathbb{I}_{\{c_j\}}(\omega_i),\ j=1,\ldots,K$$

Vetores Aleatórios

$(\Omega,\mathcal{F},\mathbb{P})$

Definição: A transformação $X:\Omega\to\mathbb{R}$ ($\mathbb{R}^k$) é uma variável aleatória (vetor) se

$$X^{-1}(B)=\{\omega\in\Omega: X(\omega)\in B\}\in\mathcal{F},\ \forall B\in\mathcal{B}(\mathbb{R})\quad(\forall B\in\mathcal{B}(\mathbb{R}^k))$$

Imagem inversa de B pela transformação X.

Resultado:

X é v.a. $\Leftrightarrow \forall t \in \mathbb{R}, X^{-1}((-\infty,t]) = \{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{F}$

($\Rightarrow$) imediato

($\Leftarrow$) seja $\mathcal{E} = \{(-\infty,t), t \in \mathbb{R}\}$

$\mathcal{A} = \{B \subseteq \mathbb{R} : X^{-1}(B) \in \mathcal{F}\}$

$\mathcal{A}$ é σ-álgebra de subconjuntos de $\mathbb{R}$

1. $\mathbb{R} \in \mathcal{A}$

$X^{-1}(\mathbb{R}) = \Omega \in \mathcal{F}$

2. $A \in \mathcal{A} \Rightarrow X^{-1}(A) \in \mathcal{F} \Rightarrow (X^{-1}(A))^c \in \mathcal{F}$

complemento da pré-imagem de A é a pré-imagem do completo

$\Rightarrow X^{-1}(A^c) \in \mathcal{F} \Rightarrow A^c \in \mathcal{A}$

$X(\omega) \in A^c \Rightarrow X(\omega) \notin A$

$\omega \in X^{-1}(A)$

$(X^{-1}(A))^c = X^{-1}(A^c)$

3. $(A_n)_{n \geq 1}, A_n \in \mathcal{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

$X^{-1}(A_n) \in \mathcal{F}, \forall n \geq 1$

$\bigcup_{n=1}^{\infty} (X^{-1}(A_n)) \in \mathcal{F}$

$X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

($\Leftarrow$) Voltando à prova

$(-\infty,t] \in \mathcal{A} \Rightarrow \mathcal{E} \subseteq \mathcal{A} \Rightarrow \underbrace{\sigma(\mathcal{E})}_{\mathcal{B}(\mathbb{R})} \subseteq \mathcal{A} \Rightarrow \mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}$

$\underbrace{\{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}}_{\text{todas as pré-imagens dos borelianos}} \subseteq \underbrace{\{X^{-1}(B) : B \subseteq \mathcal{A}\}}_{\text{todas as pré-imagens de } \mathcal{A}} \subseteq \mathcal{F}$

Então, por definição, X e v.a.

Exemplo: $\Omega = \{1,2,3,4,5,6\}$

1. $\mathcal{F} = \mathcal{P}(\Omega)$

$$X: \Omega \to \mathbb{R}$$
$$\omega \mapsto X(\omega) = \omega$$

$$Y: \Omega \to \mathbb{R}$$
$$\omega \mapsto Y(\omega) = \begin{cases} 1, & \omega \in \{2,4,6\} \\ 0, & \omega \in \{1,3,5\} \end{cases}$$

$$X^{-1}((-\infty,t]) = \begin{cases} \emptyset, & t < 1 \\ \{1\}, & 1 \leq t < 2 \\ \{1,2\}, & 2 \leq t < 3 \\ \{1,2,3\}, & 3 \leq t < 4 \\ \{1,2,3,4\}, & 4 \leq t < 5 \\ \{1,2,3,4,5\}, & 5 \leq t < 6 \\ \underbrace{\{1,2,3,4,5,6\}}_{= \Omega}, & t \geq 6 \end{cases}$$

Como $X^{-1}((-\infty,t]) \in \mathcal{F}$, X é v.a.

$$Y^{-1}((-\infty,t]) = \begin{cases} \emptyset, & t < 0 \\ \{1,3,5\}, & 0 \leq t < 1 \\ \Omega, & t \geq 1 \end{cases}$$

Como $Y^{-1}((-\infty,t]) \in \mathcal{F}, \forall t$, Y é variável aleatória.

2.
Agora, supondo

$$\mathcal{F}' = \{\emptyset, \Omega, \{1,3,5\}, \{2,4,6\}\}$$

Como $X^{-1}((-\infty, 1]) \notin \mathcal{F}'$, X não é variável aleatória.

Como $Y^{-1}((-\infty,t]) \in \mathcal{F}'$, $\forall t \in \mathbb{R}$, Y é variável aleatória.

Uma v.a. é uma função $\mathcal{F}$-mensurável (matemáticamente)

3. $\Omega = [-1, 1]$

$\mathcal{F} = \mathcal{B}([-1,1]) = \{A \cap [-1,1] : A \in \mathcal{B}(\mathbb{R})\}$

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

$\omega \in \Omega \mapsto X(\omega) = \omega^2$

$$X^{-1}((-\infty, t]) = \begin{cases} \emptyset, \ \forall t < 0 & \in \mathcal{F} \\ \{0\}, \ t = 0 & \in \mathcal{F} \qquad \{0\} = \{0\} \cap [-1,1] \\ [-\sqrt{t}, \sqrt{t}], \ 0 < t < 1 & \in \mathcal{F} \qquad [-\sqrt{t}, \sqrt{t}] \cap [-1,1] \\ \Omega, \ t \geq 1 & \in \mathcal{F} \end{cases}$$

$\forall t \in \mathbb{R}$, $X^{-1}((-\infty, t]) \in \mathcal{F}$. Logo, X é v.a.

Exemplo. 4. $(\Omega, \mathcal{F}, \mathbb{P})$. Seja $x_0 \in \mathbb{R}$

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

$\omega \in \Omega \to X(\omega) = x_0$

$$X^{-1}((-\infty, t]) = \begin{cases} \emptyset, \ t < x_0 \\ \Omega, \ t \geq x_0 \end{cases}$$

Logo, a função constante sempre é v.a., em qualquer espaço.

Logo, X é variável aleatória.

Probabilidade e Inferência Estatística I

Aula 07 - 19/03/2013 Luiz Gustavo Esteves

Variável Aleatória

$(\Omega, \mathcal{F}, \mathbb{P})$

$X : \Omega \to \mathbb{R}$ é v.a. se $\forall B \in \mathcal{B}(\mathbb{R})$,

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F} \Leftrightarrow$$

$$X^{-1}((-\infty, t]) \in \mathcal{F}, \forall t \in \mathbb{R}$$

$X: \Omega \to \mathbb{R}^k$ é vetor aleatório se $\forall B \in \mathcal{B}(\mathbb{R}^k)$,

$$X^{-1}(B) = \ldots \in \mathcal{F} \Leftrightarrow$$

$$\Leftrightarrow X^{-1}((-\infty, t_1] \times \ldots \times (-\infty, t_k]) \in \mathcal{F}, \forall (t_1, \ldots, t_k) \in \mathbb{R}^k$$

Seja X uma variável aleatória

Seja $\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$

Pode-se provar que $\sigma(X)$ é uma σ-álgebra de subconjuntos de $\mathcal{F}$ ($\sigma(X) \subseteq \mathcal{F}$).

A $\sigma(X)$ damos o nome de σ-álgebra gerada por X.
(induzida)

$$\sigma(X) = \sigma(\{X^{-1}((-\infty, t]) : t \in \mathbb{R}\})$$

Exemplo:

$\Omega = \{1, 2, 3, 4, 5, 6\}$

$\mathbb{P}(A) = \frac{|A|}{6}$

$\mathcal{F} = \mathcal{P}(\Omega)$

$X(\omega) = \omega$ $\quad\quad$ $Y(\omega) = \begin{cases} 1, & \omega \in \{2, 4, 6\} \\ 0, & \text{c.c.} \end{cases}$

$$X^{-1}((-\infty, t]) = \begin{cases} \emptyset, & t < 1 \\ \{1\}, & 1 \leq t < 2 \\ \{1, 2\}, & 2 \leq t < 3 \\ \vdots & \\ \{1, 2, 3, 4, 5, 6\}, & t \geq 6 \end{cases}$$

Nesse caso, $\sigma(X) = \mathcal{F} = \mathcal{P}(\Omega)$

$$Y^{-1}((-\infty,t]) = \begin{cases} \emptyset, & t<0 \\ \{1,3,5\}, & 0\leq t<1 \\ \Omega, & t\geq 1 \end{cases}$$

$$\sigma(Y) = \sigma(\{\emptyset, \{1,3,5\}, \Omega\}) = \{\emptyset, \Omega, \{1,3,5\}, \{2,4,6\}\}$$

$V(\omega) = 2$

$$V^{-1}((-\infty,t]) = \begin{cases} \emptyset, & t<2 \\ \Omega, & t\geq 2 \end{cases}$$

$$\sigma(V) = \{\emptyset, \Omega\}$$

Exemplo: $\Omega = [-1,1]$

$$\mathcal{F} = \mathcal{B}([-1,1])$$

$$\mathbb{P}(A) = \int_A \frac{1}{2}\, dx$$

$$\mathbb{P}(A) = \frac{\text{"comprimento de A"}}{2}, \text{ se A é intervalo}$$

$X(\omega) = \omega^2$

$$\{X^{-1}((-\infty,t]): t\in\mathbb{R}\} = \{[-t,t], t\in[0,1]\} \cup \{\emptyset\}$$

$$\therefore\ \sigma(X) = \sigma(\{[-t,t]\ t\in[0,1]\}\cup\{\emptyset\})$$

$$= \{A \cup -A : A \in \mathcal{B}([0,1])\}, \text{ em que}$$

$$-A = \{-\omega : \omega \in A\}$$

$$\{-1/2, 1/2\} \in \sigma(X)$$

$$(-1/2, -1/3] \cup [1/3, 1/2) \in \sigma(X)$$

$$[1/3, 1/2) \notin \sigma(X)$$

Exemplo. Eleição

$$\Omega = \{c_1, \dots, c_k, c_{k+1}\}^{2000}$$

$$\mathcal{F} = \mathcal{P}(\Omega)$$

$$X_1((\omega_1, \dots, \omega_{2000})) = \sum_{i=1}^{2000} \mathbb{I}_{\{c_1\}}(\omega_i)$$

Seja $A_i = \{\omega \in \Omega : X_1(\omega) = i\}$, $i = 0, \dots, 2000$

$$\sigma(X_1) = \sigma(\{A_0, \dots, A_{2000}\})$$

$$(\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{X} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$$

$$\mathbb{P}_X : \mathcal{B}(\mathbb{R}) \to \mathbb{R}_+$$

$$B \in \mathcal{B}(\mathbb{R}) \mapsto \mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\})$$

1.

$$\mathbb{P}_X(\mathbb{R}) = \mathbb{P}(X^{-1}(\mathbb{R})) = \mathbb{P}(\Omega) = 1$$

2.

$(B_n)_{n \geq 1}$, com $B_n \in \mathcal{B}(\mathbb{R})$, $\forall n \geq 1$ e $B_i \cap B_j = \emptyset$, $i \neq j$.

$$\mathbb{P}_X\left(\bigcup_{n=1}^{\infty} B_n\right) = \mathbb{P}\left(X^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right)\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} X^{-1}(B_n)\right) =$$

$$= \sum_{n=1}^{\infty} \mathbb{P}(X^{-1}(B_n)) = \sum_{n=1}^{\infty} \mathbb{P}_X(B_n)$$

Logo, $\mathbb{P}_X$ é, de fato, probabilidade em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

À medida $\mathbb{P}_X$, damos o nome de medida de probabilidade induzida pela variável aleatória X ou distribuição de X.

Voltando ao exemplo do dado:

$X(\omega) = \omega$

$B \in \mathcal{B}(\mathbb{R})$

$$\mathbb{P}_X(B) = \sum_{i=1}^{6} \frac{1}{6} \mathbb{I}_B(i)$$

$$Y(\omega) = \begin{cases} 1, & \omega \in \{2,4,6\} \\ 0, & \text{c.c.} \end{cases}$$

$$\mathbb{P}(Y^{-1}(\{0\})) = \mathbb{P}(\{1,3,5\}) = 1/2 = \mathbb{P}_Y(\{0\})$$

$$\mathbb{P}_Y(B) = \frac{1}{2}\mathbb{I}_B(0) + \frac{1}{2}\mathbb{I}_B(1)$$

Exemplo: Lançamento de duas moedas

$\Omega = \{C, \bar{C}\}^2$

$\mathcal{F} = \mathcal{P}(\Omega)$

$\mathbb{P}(A) = \frac{|A|}{4}$

$$X = \begin{cases} 1, & \text{moeda de R\$1 resulta } C \\ 0, & \text{c.c.} \end{cases}$$

Y: número de caras observadas

$$\mathbb{P}_X(\{1\}) = \mathbb{P}(X^{-1}(\{1\})) = \mathbb{P}(\{CC, C\bar{C}\}) = \frac{1}{2}$$

Do mesmo modo:

$$\mathbb{P}_X(\{0\}) = 1/2$$

$$\mathbb{P}_Y(\{0\}) = \mathbb{P}(Y^{-1}(\{0\})) = \mathbb{P}(\{\bar{C}\bar{C}\}) = 1/4$$

$$\mathbb{P}_Y(\{1\}) = \mathbb{P}(Y^{-1}(\{1\})) = \mathbb{P}(\{C\bar{C}, \bar{C}C\}) = 1/2$$

$$\mathbb{P}_X(B) = \frac{1}{2}\mathbb{I}_B(0) + \frac{1}{2}\mathbb{I}_B(1)$$

$$\mathbb{P}_Y(B) = \frac{1}{4}\mathbb{I}_B(0) + \frac{2}{4}\mathbb{I}_B(1) + \frac{1}{4}\mathbb{I}_B(2)$$

Para caracterizar X (ou, mais precisamente, $\mathbb{P}_X$) basta avaliarmos $\mathbb{P}_X$ para eventos em

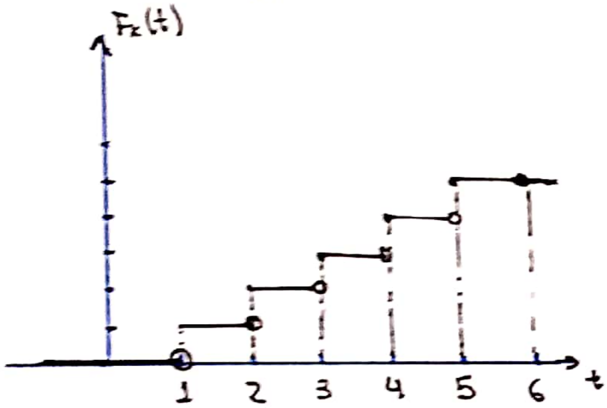
$$\mathcal{E} = \{(-\infty, t] : t \in \mathbb{R}\}$$

Definição: Seja X uma variável aleatória.

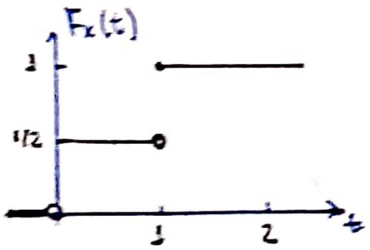
À função $F_X : \mathbb{R} \to [0,1]$ que associa a cada $t \in \mathbb{R}$, $F_X(t) = P_X((-\infty, t]) = P(\{\omega \in \Omega : X(\omega) \leq t\})$, damos o nome de função de distribuição da variável aleatória X.

Exemplo: Dado

$$P_X((-\infty,t]) = \sum_{i=1}^{6} \frac{1}{6} \mathbb{I}_{(-\infty,t]}(i)$$



$$P_Y((-\infty,t]) = \frac{1}{2} \mathbb{I}_{(-\infty,t]}(0) + \frac{1}{2}(1)$$



$\Omega = [0,1]^2$ $\quad$ $P(A)$: "Área (A)"

$\mathcal{F} = \mathcal{B}([0,1]^2)$

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

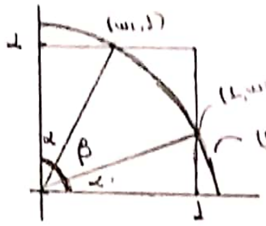
$(\omega_1, \omega_2) \mapsto X(\omega_1, \omega_2) = \sqrt{\omega_1^2 + \omega_2^2}$

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

$F_X(t) = P(\{(\omega_1, \omega_2) \in \Omega : X(\omega_1, \omega_2) \leq t\})$

$$F_X(t)=\begin{cases}0, & t<0\\ \pi t^2/4, & 0\le t<1 \\ \sqrt{t^2-1}+\frac{t^2}{2}\left(\frac{\pi}{2}-2\operatorname{arctg}(\sqrt{t^2-1})\right), & 1\le t\le\sqrt{2}\end{cases}$$

→ $A_0=\pi r^2$, Nesse caso, $r=t/2$



$\beta = 2\alpha-\frac{\pi}{2}$

$\omega_1^2+1=t^2 \Leftrightarrow \omega_1=\sqrt{t^2-1}$

$A_\Delta = \frac{1}{2}\,1.\sqrt{t^2-1} = A_\nabla$

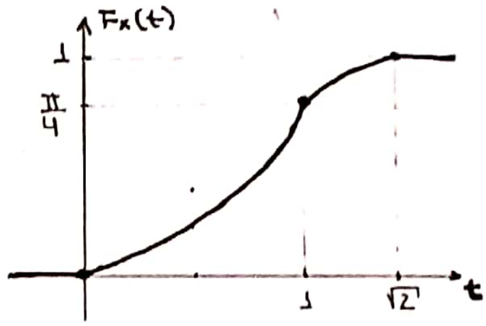
$A_\triangleleft = \frac{1}{2}\beta t^2.$

$tg(\alpha)=\sqrt{t^2-1}/1 \Rightarrow$

$\alpha = \operatorname{arctg}(\sqrt{t^2-1}) \Rightarrow$

$\beta = \frac{\pi}{2}-2\operatorname{arctg}(\sqrt{t^2-1})$,

$A = 2A_\Delta + A_\triangleleft,\ 1\le t\le\sqrt{2}$



Propriedades de F_X:

1. F_X é não-decrescente

$$t_1\le t_2 \Rightarrow (-\infty,t_1]\subseteq(-\infty,t_2]\Rightarrow$$
$$\Rightarrow \mathbb{P}_X((-\infty,t_1])\le\mathbb{P}_X((-\infty,t_2])\Rightarrow F_X(t_1)\le F_X(t_2).$$

2. F_X é contínua à direita

(ou seja, $t_n\downarrow t \Rightarrow F(t)=\lim_{n\to\infty}F(t_n)$)

$$F(t_n)\downarrow F(t)$$

Seja $(t_n)_{n\ge1}$ decrescente, t.q. $t_n\downarrow t \Rightarrow$

$\Rightarrow ((-\infty,t_n])_{n\ge1}$ é monótona decrescente e, portanto,

$$\lim_{n\to\infty} \mathbb{P}_X((-\infty,t_n]) = \mathbb{P}_X(\underbrace{\lim_{n\to\infty}(-\infty,t_n]}_{=(-\infty,t]}) \Rightarrow \lim_{n\to\infty} F_X(t_n) = F_X(t)$$

3. Analogamente,

$$\lim_{t\downarrow-\infty} F_X(t) = 0 \quad e \quad \lim_{t\uparrow\infty} F_X(t) = 1$$

Consequências.

1. $\mathbb{P}(X<t) = \mathbb{P}(\{\omega\in\Omega : X(\omega)<t\}) =$

$$= \lim_{x\to t^-} F_X(t) = F_X(t^-)$$

2. $\mathbb{P}(X=t) = F_X(t) - F_X(t^-)$

Comentários:

1. $\mathbb{P}_X : \mathcal{B}(\mathbb{R}) \mapsto \mathbb{R}_+$

$$\downarrow$$

$$F_X$$

A partir de uma função F, atendendo as propriedades (1), (2), (3), podemos construir uma medida $\mathbb{P}$ em $(\mathbb{R},\mathcal{B}(\mathbb{R}))$.

2. Outras definições de função distribuição.

$$F(t) = \mathbb{P}(X<t)$$

3. Duas v.a. totalmente distintas (nem mesmo definidas no mesmo espaço) podem existir a mesma medida induzida ou a mesma função de distribuição.

Exemplo: $(\{1,2,3,4,5,6\}, \mathcal{P}(\Omega_1), \mathbb{P}_1)$

$$Y(\omega) = \begin{cases} 1, & \omega \in \{2,4,6\} \\ 0, & \text{c.c} \end{cases}$$

$(\{CC, C\bar{C}, \bar{C}C, \bar{C}\bar{C}\}, \mathcal{P}(\Omega_2), \mathbb{P}_2)$

$$X(\omega) = \begin{cases} 1, & \text{cara na moeda de R\$ 1,00} \\ 0, & \text{c.c.} \end{cases}$$

$([0,1], \mathcal{B}([0,1]), \mathbb{P}_3)$

$$Z(\omega) = \begin{cases} 1, & \omega \leq 1/2 \\ 0, & \text{c.c.} \end{cases}$$

Note que

$\mathbb{P}_X = \mathbb{P}_Y = \mathbb{P}_Z$ (e, consequentemente)

$F_X = F_Y = F_Z$, dado por $\mathbb{P}_X(B) = \frac{1}{2}(\mathbb{I}_B(0) + \mathbb{I}_B(1))$

Além disso, $1-X$ e X induzem a mesma $\mathbb{P}_X$.

Nesse caso, dizemos que:

$X \overset{d}{=} 1-X$ ou

$X \overset{d}{=} Y \overset{d}{=} Z$

Probabilidade e Inferência Estatística I

21 de março de 2013 - Aula 08 Luiz Gustavo Esteves

Na aula passada:

A partir de uma v.a. construímos um novo espaço $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_x)$. Ou seja,

$$(\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{X} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$$

Definição: Seja $X: \Omega \to \mathbb{R}^k$ um vetor aleatório que induz a medida $\mathbb{P}_X: \mathcal{B}(\mathbb{R}^k) \to \mathbb{R}_+$.

À função $F_X: \mathbb{R}^k \to [0,1]$ associada a cada $(t_1, \dots, t_k) \in \mathbb{R}^k$, o número

$$\mathbb{P}_X\left((-\infty, t_1] \times (-\infty, t_2] \times \dots \times (-\infty, t_k]\right)$$

damos o nome de função distribuição (conjunta) de X $(= (X_1, \dots, X_n))$

Propriedades:

(1) F_X é não decrescente em cada componente

$t_{i1} \leq t_{i2}$

$$(-\infty, t_1] \times \dots \times (-\infty, t_{i1}] \times \dots \times (-\infty, t_k] \subseteq (-\infty, t_1] \times \dots \times (-\infty, t_{i2}] \times \dots \times (-\infty, t_k]$$

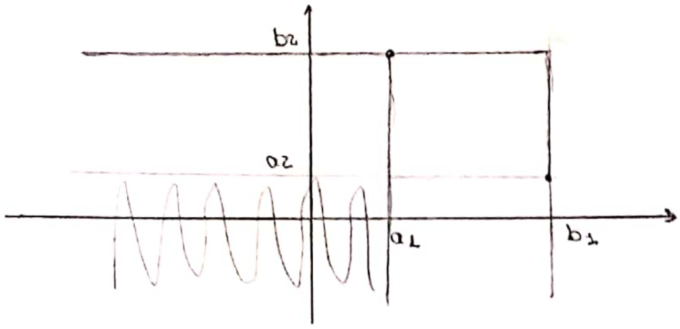
2. F_X é contínua à direita em cada componente

3. $\lim\limits_{t_i \to -\infty} F_X(t_1, \dots, t_k) = 0$ para algum i e $\lim\limits_{t_i \to \infty} F_X(t_1, \dots, t_k) = 1$ para TODO i

4. F_X deve ser tal $\mathbb{P}_X((a_1, b_1] \times (a_2, b_2] \times \dots \times (a_k, b_k])$ (expressa em função de F_X) deve ser ≥ 0.

A propriedade 4. se faz necessária para o caso multivariado para que a medida possa ser recuperada. No caso univariado, valem 1, 2 e 3.

Exemplo: Caso bivariado (Propriedade 4).



$$P_X((a_1,b_1]\times(a_2,b_2]) = F_X(b_1,b_2) - F_X(b_1,a_2) - F_X(a_1,b_2) + F_X(a_1,a_2) \geq 0$$

Exemplo 1.

$$F_X(t_1,t_2) = (1-e^{-t_1})(1-e^{-t_2}-t_2e^{-t_2})\,\mathbb{I}_{\mathbb{R}_+}(t_1)\,\mathbb{I}_{\mathbb{R}_+}(t_2)$$

$(a_1,b_1),(a_2,b_2)$

$$F_X(b_1,b_2) - F_X(b_1,a_2) - F_X(a_1,b_2) + F_X(a_1,a_2) =$$

$$= (1-e^{-b_1})[e^{-a_2}+a_2e^{-a_2}-e^{-b_2}-b_2e^{-b_2}] - (1-e^{-a_1})[e^{-a_2}+a_2e^{-a_2}-e^{-b_2}-b_2e^{-b_2}]$$

$$= [e^{-a_2}+a_2e^{-a_2}-e^{-b_2}-b_2e^{-b_2}](e^{-a_1}-e^{-b_1}) > 0.$$

Exemplo 2.

$$F_X(t_1,t_2) = (1-e^{-(t_1+t_2)})\,\mathbb{I}_{\mathbb{R}_+}(t_1)\,\mathbb{I}_{\mathbb{R}_+}(t_2)$$

$(0,1]\times(0,1]$

$$F_X(1,1) - F_X(1,0) - F(0,1) + F(0,0) =$$

$$= (1-e^{-2}) - (1-e^{-1}) - (1-e^{-1}) + 0 =$$

$$= -1 - e^{-2} + 2e^{-1} = -(1+e^{-2}-2e^{-1}) < 0$$

Logo, F não é função de distribuição de vetor aleatório algum.

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

Seja $X=(X_1,\dots,X_k):\Omega\to\mathbb{R}^k$ uma transformação.

X é vetor aleatório $\Leftrightarrow$ X_i é variável aleatória.

$(\Rightarrow)$ $\forall (t_1,\dots,t_k)\in\mathbb{R}^k$, $X^{-1}((-\infty,t_1]\times(-\infty,t_2]\times\dots\times(-\infty,t_k])\in\mathcal{F}$

$$X^{-1}((-\infty,t_1]\times\dots\times(-\infty,t_k]) = \{\omega\in\Omega : X(\omega) = (X_1(\omega), X_2(\omega), \dots, X_k(\omega))\in(-\infty,t_1]\times\dots\times(-\infty,t_k]\}$$

$$= \bigcap_{i=1}^{k}\{\omega : X_i(\omega)\in(-\infty,t_i]\}$$

Para todo $n\geq 1$

$$\bigcap_{j=2}^{k}\{\omega : X_j(\omega)\leq n\}\cap\{\omega : X_1(\omega)\leq t_1\}\in\mathcal{F}$$

Logo,

$$\bigcup_{n=1}^{\infty}\left[\bigcap_{j=2}^{k}\{\omega : X_j(\omega)\leq n\}\cap\{\omega : X_1(\omega)\leq t_1\}\right]\in\mathcal{F} =$$

$$=\{\omega : X_1(\omega)\leq t_1\}\cap\Big(\underbrace{\bigcup_{n\geq 1}\Big[\bigcap_{j=2}^{k}\{\omega : X_j(\omega)\leq n\}\Big]}_{\Omega}\Big)\in\mathcal{F}$$

$$\Rightarrow \{\omega : X_1(\omega)\leq t_1\}\in\mathcal{F}$$

Logo, X_1 é variável aleatória.

$(\Leftarrow)$ X_i é v.a., $\forall i = 1, 2, \dots, k$

$$X^{-1}((-\infty,t_1]\times\dots\times(-\infty,t_k]) = \bigcap_{i=1}^{k}\{\omega : X_i(\omega)\in(-\infty,t_i]\} = \bigcap_{i=1}^{k}X_i^{-1}(-\infty,t_i]\in\mathcal{F}$$

Logo X é vetor aleatório.

Seja $X = (X_1,\dots,X_k)$ vetor aleatório com função distribuição F_X.

Vamos avaliar:

$$\lim_{n\to\infty}F_X(t_1,n,n,\dots,n) = \lim_{n\to\infty}\mathbb{P}_X((-\infty,t_1]\times(-\infty,n]\times\dots\times(-\infty,n]) =$$

$$=\mathbb{P}_X(\lim_{n\to\infty}(-\infty,t_1]\times(-\infty,n]\times\dots\times(-\infty,n]) = \mathbb{P}_X\Big(\bigcup_{n=1}^{\infty}(-\infty,t_1]\times(-\infty,n]\times\dots\times(-\infty,n]\Big) =$$

$$=\mathbb{P}_X((-\infty,t_1]\times\mathbb{R}\times\mathbb{R}\times\dots\times\mathbb{R}) = \mathbb{P}_X(\omega\in\Omega : X(\omega)\in(-\infty,t_1]\times\underbrace{\mathbb{R}\times\dots\times\mathbb{R}}_{k-1\text{ vezes}}) =$$

$$=\mathbb{P}_X(\{\omega\in\Omega : X_1(\omega)\leq t_1\}\cap\Omega\cap\Omega\cap\dots\cap\Omega) =$$

$$\Rightarrow \underbrace{\lim_{n\to\infty} F_X(t_1,n,\dots,n)}_{F_X(t_1,\infty,\infty,\dots,\infty)} = P_{X_1}((-\infty,t_1]) = F_{X_1}(t_1)$$

Analogamente

$$F_{X_2}(t) = F_X(\infty,t,\infty,\dots,\infty)$$

Vale ainda que

$$F_{(X_1,X_2)}(t_1,t_2) = F_X(t_1,t_2,\infty,\infty,\dots,\infty)$$

Exemplo.

1. Sejam $F_1, F_2, F_3 : \mathbb{R} \to [0,1)$ função distribuição

Seja $F: \mathbb{R}^2 \to [0,1]$ dada por

$$F(t_1,t_2) = F_1(t_2)F_3(t_2) - F_3(t_2)(1-F_2(t_1))(F_1(t_2)-F_1(t_1))\,\mathbb{I}_{(-\infty,t_2]}(t_1)$$

$$F_{X_1}(t_1) = F_X(t_1,\infty) = 1-(1-F_2(t_1))(1-F_1(t_1))$$

$$F_{X_2}(t_2) = F_X(\infty,t_2) = F_1(t_2)F_3(t_2)$$ pois com $t_1\to\infty$, $\mathbb{I}_{(-\infty,t_2]}(t_1)=0$ já que $t_1 > t_2$.

2. $F_1, F_2, F_3, F_4 : \mathbb{R} \to [0,1)$ função distribuição

Seja $F_X : \mathbb{R}^3 \to \mathbb{R}_+$ $\quad X = (X_1,X_2,X_3)$

$$F_X(t_1,t_2,t_3) = F_1(\min\{t_1,t_2,t_3\})F_2(t_1)F_3(t_2)F_4(t_3)$$

$$F_{X_2}(t_2) = F_X(\infty,t_2,\infty) = F_1(t_2)F_3(t_2)$$

$$F_{(X_1,X_3)}(t_1,t_3) = F_X(t_1,\infty,t_3) = F_1(\min\{t_1,t_3\})F_2(t_1)F_4(t_3)$$

Tipos de Vetores Aleatórias

1. Vetores Discretos

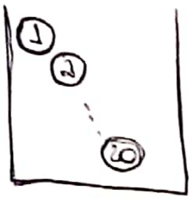
Definição: $X:\Omega \to \mathbb{R}^k$ é discreto se $\exists \{x_n\}_{n\geq 1}$, $x_n \in \mathbb{R}^k$, $\forall n \geq 1$, tal que

$$P_X(\{x_n\}_{n\geq 1}) = 1.$$

Nesse caso, denotando $P_X(\{x_n\}) = p_n$, $n \geq 1$, P_X fica "caracterizada" (representada) pela lista $\boxed{((x_n, p_n))_{n\geq 1}.}$ → função (densidade) de probabilidade de X

$$P_X(A) = \sum_{n \geq 1}^{\infty} p_n \mathbb{I}_A(x_n)$$

Exemplo:



Retirar 4 bolas

$\Omega = \{A \subseteq \{1, \dots, 20\} : |A| = 4\}$

X_1: menor valor observado $\quad X_1(\omega) = \min \omega$

X_2: maior valor observado $\quad X_2(\omega) = \max \omega$

$X = (X_1, X_2)$

$X(\Omega) = \{X(\omega) : \omega \in \Omega\} = \{(a,b) \in \{1, \dots, 20\}^2 : b \geq 4 \text{ e } a \leq b-3\}$

$$P_X(\{(a,b)\}) = \underbrace{P(X = (a,b))}_{\{\omega \in \Omega : X(\omega) = (a,b)\}} = P(X_1 = a, X_2 = b) = \frac{\binom{b-1-a}{2}}{\binom{20}{4}}, \quad (a,b) \in X(\Omega)$$

↳ $(a, b, \overset{x_1}{-}, \overset{x_2}{-})$

$a \leq x_1 \leq b$

$a \leq x_2 \leq b$

$x_1 \neq x_2$

$X_i(\Omega) = \{X_i(\omega) : \omega \in \Omega\} < \infty$, $i = 1, 2$.

$$P(X_1 = a) = \frac{\binom{20-a}{3}}{\binom{20}{4}} \qquad a = 1, \dots, 17$$

$$P(X_2 = b) = \frac{\binom{b-1}{3}}{\binom{20}{4}} \qquad b = 4, 5, \dots, 20.$$

PRINCIPAIS MODELOS DISCRETOS

1. Uniforme

Dizemos que a v.a. X é distribuída segundo o modelo UNIFORME em $\{x_1, \dots, x_k\}$ se

$$\mathbb{P}(X = x_i) = \frac{1}{k}, \text{ para } i = 1, \dots, k$$

Notação $X \sim \text{Uniforme}(\{x_1, \dots, x_k\})$

2. Bernoulli:

Dizemos que X é distribuído segundo o modelo BERNOULLI de parâmetro p, $0 < p < 1$.

$$\mathbb{P}(X=1) = 1 - \mathbb{P}(X=0) = p$$

$$\mathbb{P}(X=x) = p^x (1-p)^{1-x} \mathbb{I}_{\{0,1\}}(x)$$

Notação: $X \sim Ber(p)$.

3. Binomial

Dizemos que X é distribuída segundo o modelo BINOMIAL de parâmetros n e p, $n \in \mathbb{N} \setminus \{0\}$, $0 < p < 1$, se

$$\mathbb{P}(X=j) = \binom{n}{j} p^j (1-p)^{n-j} \mathbb{I}_{\{0,1,\dots,n\}}(j)$$

Notação: $X \sim Bin(n,p)$

4. Poisson

Dizemos que X é distribuída segundo o modelo POISSON de parâmetro λ, $\lambda > 0$, se

$$\mathbb{P}(X=j) = \frac{e^{-\lambda} \lambda^j}{j!} \mathbb{I}_{\mathbb{N}}(j) \qquad X \sim Poi(\lambda)$$

Bolinhas: ① ② ③ ④ ... Ⓝ

Distribuição dos eventos raros

Urnas: $\underset{1}{\sqcup}$ $\underset{2}{\sqcup}$ $\underset{3}{\sqcup}$ $\underset{4}{\sqcup}$ ... $\underset{N}{\sqcup}$

$X :=$ nº de bolas na urna 1 ou na urna 2.

$X \sim Bin(N, 2/N)$

$$\mathbb{P}(X=j) = \binom{N}{j} \left(\frac{2}{N}\right)^j \left(\frac{N-2}{N}\right)^{N-j} \xrightarrow[N \to \infty]{} \frac{e^{-2} 2^j}{j!}$$

5. Geométrico

Dizemos que a v.a. X é distribuída segundo o modelo GEOMÉTRICO de parâmetro p, $0 < p \leq 1$.

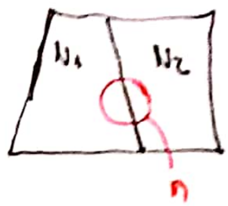
$$\mathbb{P}(X=j) = (1-p)^{j-1} \cdot p \, \mathbb{I}_{\{1,2,3,4,\dots\}}(j)$$

Notação: $X \sim Geo(p)$

6. Hipergeométrica

Dizemos que a variável aleatória X é distribuída segundo o modelo HIPERGEOMÉTRICO de parâmetros N_1, N_2 e n, $N_1, N_2, n \in \mathbb{N}\setminus\{0\}$ se

$$\mathbb{P}(X=j) = \frac{\binom{N_1}{j}\binom{N_2}{n-j}}{\binom{N_1+N_2}{n}} \mathbb{I}_{\mathbb{N}}(j), \quad \text{com } \binom{a}{b} = 0 \text{ se } a < b$$



- útil na amostragem: captura-recapturas, sem reposição.

7. Multinomial

Dizemos que o vetor aleatório $X = (X_1, X_2, \dots, X_k)$ é distribuído segundo o modelo MULTINOMIAL de parâmetros n e p, $n \in \mathbb{N}\setminus\{0\}$ e $p \in S = \{(p_1, \dots, p_k) \in [0,1]^k : p_1+p_2+\dots+p_k \leq 1\}$

$$\mathbb{P}(X_1 = x_1, \dots, X_k = x_k) = \frac{n!}{x_1! \, x_2! \dots x_k! \, x_{k+1}!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} p_{k+1}^{x_{k+1}} \mathbb{I}_A(x_1, \dots, x_k), \text{ onde}$$

$A = \{(x_1, \dots, x_k) \in \mathbb{N}^k : x_1 + x_2 + \dots + x_k \leq n\}$,

$x_{k+1} = n - (x_1 + \dots + x_k)$ e $p_{k+1} = 1 - (p_1 + \dots + p_k)$

$X \sim Multinomial(n, (p_1, \dots, p_k))$

Probabilidade e Inferência Estatística I

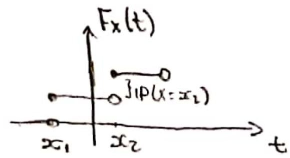
Aula 09 - 22/03/2013

Luiz Gustavo Esteves

Atendimento Terça (26) 16:00/18:00 (Devido à semana santa)

Discreto

$\mathbb{P}_X$, F_X, $((x_n, p_n))_{n \geq 1}$



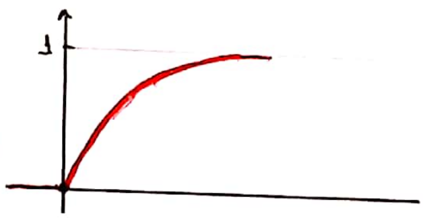
Variáveis Contínuas

Seja X um vetor aleatório com função distribuição F_X.

Definição: X é um vetor aleatório contínuo se F_X é contínua.

Exemplo 1. X v.a.

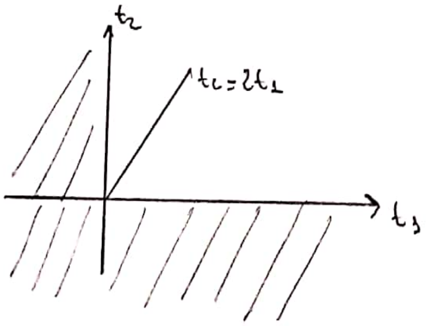
$$F_X(t) = (1 - e^{-t})\, \mathbb{I}_{\mathbb{R}_+}(t)$$



Como F_X é contínua, X é v.a. contínua.

Exemplo 2. X vetor aleatório

$$F_X(t_1, t_2) = (1 - e^{-t_1})(1 - e^{-t_2} - t_2 e^{-t_2})\, \mathbb{I}_{\mathbb{R}_+}(t_1)\, \mathbb{I}_{\mathbb{R}_+}(t_2)$$

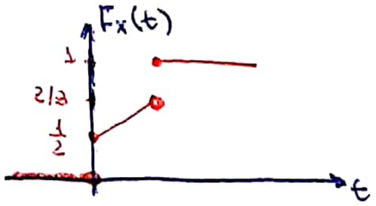


X é vetor aleatório contínuo

3. $F_X(t_1,t_2) = \left(1 - e^{-\min\{t_1, t_2/2\}}\right) \mathbb{I}_{\mathbb{R}_+}(t_1)\, \mathbb{I}_{\mathbb{R}_+}(t_2)$

X é contínuo.

4. $F_X(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{2} + \frac{t}{3}, & 0 \leq t < 1 \\ 1, & t \geq 1 \end{cases}$



Seja X um vetor aleatório com função distribuição F_X.

Definição: X é um vetor aleatório ABSOLUTAMENTE CONTÍNUO ("CONTÍNUO") se

$\exists\, f_X : \mathbb{R}^k \to \mathbb{R}_+$ tal que

$$F_X(t_1, \dots, t_k) = \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \cdots \int_{-\infty}^{t_k} f_X(x_1, \dots, x_k)\, dx$$

Nesse caso, vale que $\dfrac{\partial^{(k)} F_X(t_1, \dots, t_k)}{\partial t_1 \partial t_2 \dots \partial t_k} = f_X(t_1, \dots, t_k)$.

$f_X(x_1, \dots, x_k)$: função densidade de probabilidade (conjunta)

Exemplos:

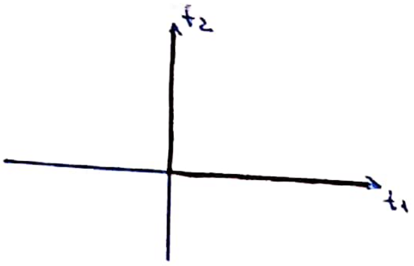
1. $F_X(t) = (1 - e^{-t})\mathbb{I}_{\mathbb{R}_+}(t)$

$$\frac{dF_X(t)}{dt} = \begin{cases} e^{-t}, & t > 0 \\ 0, & t < 0 \end{cases}$$

Nesse caso, X é v.a. absolutamente contínua.

$f_X(t) = e^{-t}\, \mathbb{I}_{\mathbb{R}_+}(t)$

2. $F_X(t_1,t_2) = (1-e^{-t_1})(1-e^{-t_2}-t_2e^{-t_2})\mathbb{I}_{\mathbb{R}_+}(t_1)\,\mathbb{I}_{\mathbb{R}_+}(t_1)$



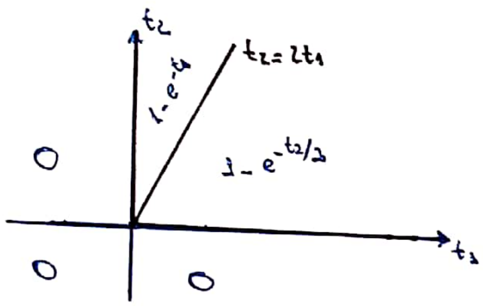
$$\frac{\partial^2 F_X(t_1,t_2)}{dt_1\,dt_2} = e^{-t_1}\cdot(e^{-t_2}-e^{-t_2}+t_2e^{-t_2}),\ t_i > 0$$

$$\frac{\partial^2 F_X(t_1,t_2)}{\partial t_1 \partial t_2} = (e^{-t_1}\cdot t_2 e^{-t_2})\mathbb{I}_{\mathbb{R}_+}(t_1)\,\mathbb{I}_{\mathbb{R}_+}(t_2)$$

L Problema
medida nula

É contínuo

3. $F_X(t_1,t_2) = (1-e^{-\min\{t_1,\,t_2/2\}})\,\mathbb{I}_{\mathbb{R}_+}(t_1)\,\mathbb{I}_{\mathbb{R}_+}(t_2)$



$$\frac{\partial^2 F_X(t_1,t_2)}{\partial t_1 \partial t_2} = 0,$$ para todo (t_1,t_2) em que F_X é diferenciável.

L problema

(F_X é singular: F_X é contínua, mas "$F'(t)=0,\ \forall t$")

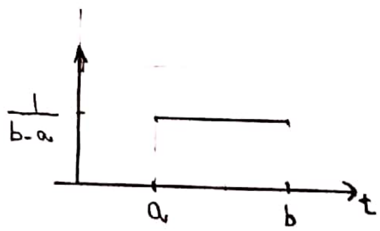
Ver Função de distribuição de Cantor (Barry James)

Alguns Modelos (ABS.) Contínuos

1. Uniforme

X é distribuído segundo o modelo Uniforme no intervalo (a,b), $a,b \in \mathbb{R}$, $a<b$, se sua densidade é dada por

$$f_X(t) = \frac{1}{b-a}\,\mathbb{I}_{(a,b)}(t)$$



Notação: $X \sim U(a,b)$

2. Beta

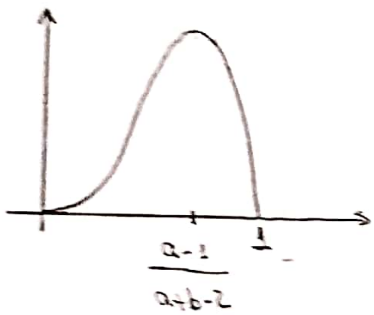
X é distribuído segundo o modelo BETA de parâmetros a e b, $a, b > 0$, se sua densidade é dada por

$$f_X(t) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} t^{a-1} (1-t)^{b-1} \mathbb{I}_{(0,1)}(t) \text{ , onde } \Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dt$$

Análise da Influência dos parâmetros:

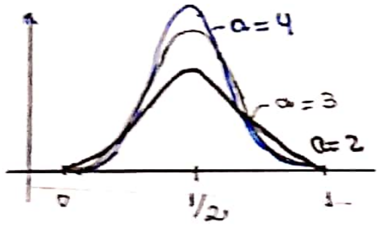
- Se $a = b = 1 \rightarrow \mathcal{U}(0,1)$

- Se $a, b > 1 \rightarrow$

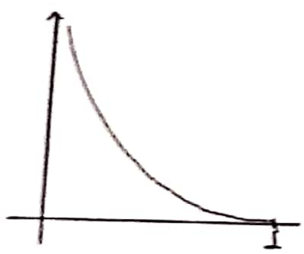


Se $a > b$, o ponto crítico fica mais próximo de 1 e, mais distante, se $b > 0$.

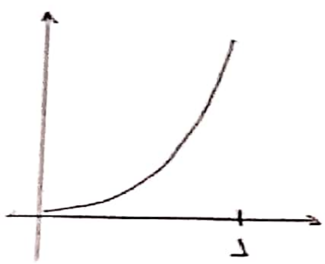
- Se $a = b$, temos um modelo simétrico em torno de 0.5. Quanto maior forem a e b, mais leptocúrtica será a curva.



- Se $a = 1/2$, $b > 1$



- Se $b = 1/2$, $a > 1$

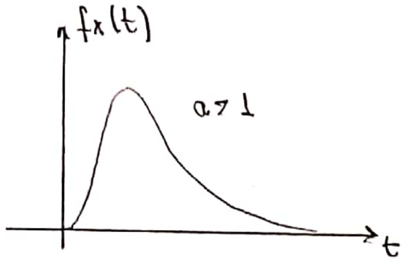


$X \sim \text{Beta}(a, b)$

3. Gama $(X \sim Gama(a,b))$

X é distribuído segundo o modelo GAMA de parâmetros a e b, $a,b>0$, se sua densidade é dada por

$$f_X(t) = \frac{b^a}{\Gamma(a)} t^{a-1} e^{-bt} \mathbb{I}_{\mathbb{R}_+}(t)$$



- Casos Particulares:

$a=1 \rightarrow f_X(t) = b e^{-bt} \mathbb{I}_{\mathbb{R}_+}(t)$ Modelo Exponencial de parâmetro b, $b>0$: $X \sim Exp(b)$

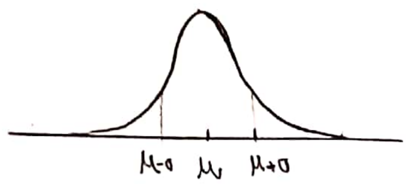
$b=\frac{1}{2} \rightarrow X \sim \chi^2_{2a} \rightarrow$ Modelo Qui-quadrado de parâmetro $2a$

$Gama\left(\frac{a}{2}, \frac{1}{2}\right) \sim \chi^2_a$.

4. NORMAL

Dizemos que X é distribuído segundo o modelo normal de parâmetros μ e σ^2, $\mu \in \mathbb{R}$, $\sigma^2>0$, se.

$$f_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$



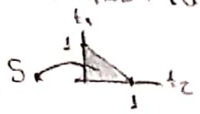
5. DIRICHLET

$X=(X_1, \ldots, X_k)$ é distribuído segundo o modelo DIRICHLET de parâmetros $a_1, a_2, \ldots, a_k$ e a_0, $a_i>0$, $i=0,1,\ldots,k$ se

$$f_X(t_1,\ldots,t_k) = \frac{\Gamma(a_0+a_1+\cdots+a_k)}{\Gamma(a_0)\Gamma(a_1)\Gamma(a_2)\ldots\Gamma(a_k)} t_1^{a_1-1} t_2^{a_2-1} \cdots t_k^{a_k-1} (1-t_1-t_2-\cdots-t_k)^{a_0-1} \mathbb{I}_{S_k}(t_1,\ldots,t_k)$$

onde $S_k = \{(t_1,\ldots,t_k) \in \mathbb{R}_+ : t_1+t_2+\cdots+t_k \leq 1\}$

- Se DIRICHLET (a_0, a_1, a_2)



- Se DIRICHLET (a_0, a_1, a_2, a_3)

Variáveis Aleatórias Mistas

Seja X uma v.a. com função de distribuição F_X.

Dizemos que X é uma v.a. mista se

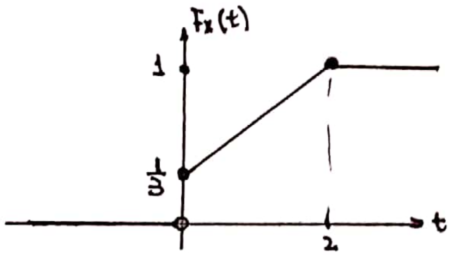
$$F_X(t) = \alpha F_c(t) + (1-\alpha) F_d(t), \quad \text{para algum } \alpha \in (0,1)$$

e F_c função de distribuição absolutamente contínua e

F_d função de distribuição discreta

Exemplo:

1. $F_X(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{3}(t+1), & 0 \leq t \leq 2 \\ 1, & t \geq 2 \end{cases}$



$$F_d(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

$$F_d(t) = \frac{1}{(1/3)} \begin{cases} 0, & t < 0 \\ 1/3, & t \geq 0 \end{cases}$$

$$F_c(t) = \begin{cases} 0, & t < 0 \\ \frac{t-0}{2-0}, & 0 \leq t \leq 2 \\ 1, & t \geq 2 \end{cases}$$

$\downarrow$

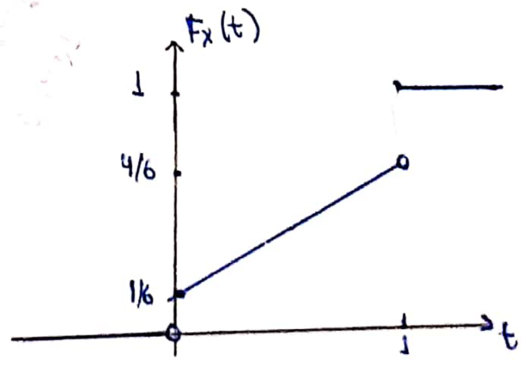
$U(0,2)$

$$F_c(t) = \frac{1}{(2/3)} \begin{cases} 0, & t < 0 \\ t, & 0 \leq t \leq 2 \end{cases}$$

Aqui, $F_X(t) = \frac{2}{3} F_c(t) + \frac{1}{3} F_d(t)$.

2.

$$F_X(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{6} + \frac{1}{2}t, & 0 \leq t \leq 1 \\ 1, & t \geq 1 \end{cases}$$



$$F_d(t) = \begin{cases} 0, & t < 0 \\ 1/3, & 0 \leq t < 1 \\ 1, & t \geq 1 \end{cases}$$

Função de Distribuição de uma Bernoulli (2/3)

$$F_c(t) = \begin{cases} 0, & t < 0 \\ t, & 0 \leq t < 1 \\ 1, & t \geq 1 \end{cases}$$

$$F_X(t) = \frac{1}{2} F_d(t) + \frac{1}{2} F_c(t)$$

$$\left(\frac{1}{6} + \frac{2}{6}\right) \qquad \left(\frac{6}{6} - \frac{1}{6}\right)$$

Próxima aula:

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

F_X

$F_{X_1}(t) = F_X(t, \infty, \infty, \ldots, \infty)$

$F_{X_3}(t) = F_X(\infty, \infty, t, \infty, \ldots, \infty)$

$F_{(X_1, X_3)}(t_1, t_2) = F_X(t_1, \infty, t_2, \infty, \ldots, \infty)$

Obtenção de f_{X_1}, $f_{(X_1, X_2)}$ a partir das funções densidades.

Probabilidade e Inferência Estatística I

Aula 10 - 02/04/2013

Luiz Gustavo Esteves

X

$P_X, F_X : \mathbb{R}^n \to [0,1]$ $\qquad F_{X_3}(t) = F_X(\infty, \infty, t, \infty, \ldots)$

caso discreto $\quad ((x_n, p_n))_{n \geq 1}$

Abs. contínuo $\quad f_X$

$X = (X_1, \dots, X_k)$

$X_i(\Omega) = \{X_i(\omega) : \omega \in \Omega\}, \ i = 1, \dots, k$

$\mathbb{P}(X_1 = x_1) = \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = x_1\})$

$= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = x_1, X_2(\omega) \in X_2(\Omega), \dots, X_k(\omega) \in X_k(\Omega)\})$

$$= \sum_{x_2 \in X_2(\Omega)} \dots \sum_{x_k \in X_k(\Omega)} \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k).$$

Exemplo 1.

20 bolas

$$\mathbb{P}(X_1 = a, X_2 = b) = \frac{\binom{b-a-1}{2}}{\binom{20}{4}} \underbrace{\mathbb{I}_{\{1,\dots,17\}}(a)\, \mathbb{I}_{\{a+3,\dots,20\}}(b)}_{\mathbb{I}_{\{4,\dots,20\}}(b)\, \mathbb{I}_{\{1,\dots,b-3\}}(a)}$$

X_1: menor valor

X_2: maior valor

$X_2(\Omega) = \{4, 5, \dots, 20\}$

$X_1(\Omega) = \{1, 2, \dots, 17\}$

$$\mathbb{P}(X_1 = a) = \sum_{x_2 \in X_2(\Omega)} \mathbb{P}(X_1 = a, X_2 = x_2) = \sum_{x_2=4}^{20} \frac{\binom{x_2-a-1}{2}}{\binom{20}{4}} \mathbb{I}_{\{1,\dots,17\}}(a)\, \mathbb{I}_{\{a+3,\dots,20\}}(x_2) =$$

$$\sum_{x_2=a+3}^{20} \frac{\binom{x_2-a-1}{2}}{\binom{20}{4}} \mathbb{I}_{\{1,\dots,17\}}(a) = \frac{\binom{20-a}{3}}{\binom{20}{4}} \mathbb{I}_{\{1,\dots,17\}}(a)$$

$$\sum_{x_2=a+3}^{20} \binom{x_2-a-1}{2} = \binom{2}{2} + \binom{3}{2} + \dots + \binom{19-a}{2} \overset{\text{T. Pascal}}{=} \binom{20-a}{3}$$

2. $X = (X_1, \dots, X_k) \sim \text{Mult}(n, (p_1, \dots, p_k))$

$$\mathbb{P}(X_1 = x_1, X_2 = x_2) = \sum_{x_3 \in X_3(\Omega)} \dots \sum_{x_k \in X_k(\Omega)} \mathbb{P}(X_1 = x_1, X_2 = x_2, X_3 = x_3, \dots, X_k = x_k) =$$

$$= \sum_{x_3=0}^{n} \cdots \sum_{x_k=0}^{n} \frac{n!}{x_1!\, x_2!\, x_3! \cdots x_k!\, x_{k+1}!}\, p_1^{x_1}\, p_2^{x_2}\, p_3^{x_3} \cdots p_k^{x_k}\, (p_{k+1})^{x_{k+1}}\, \mathbb{I}_{A_{k,n}}(x_1, x_2, \dots, x_k) =$$

onde

$x_{k+1} = n - x_1 - x_2 - \cdots - x_k$, $p_{k+1} = 1 - p_1 - p_2 - \cdots - p_k$ e

$A_{k,n} = \{(a_1, \dots, a_k) \in \mathbb{N} : a_1 + a_2 + \cdots + a_n \leq n\}$

$$= n!\, \frac{p_1^{x_1}\, p_2^{x_2}}{x_1!\, x_2!} \sum_{x_3=0}^{n} \cdots \sum_{x_k=0}^{n} \frac{p_3^{x_3} \cdots p_k^{x_k}\, p_{k+1}^{x_{k+1}}}{x_3! \cdots x_k!\, x_{k+1}!}\, \mathbb{I}_{A_{k-2,\, n-x_1-x_2}}(x_3, \dots, x_k) =$$

$$= \frac{(1-p_1-p_2)^{n-x_1-x_2}}{(n-x_1-x_2)!}\, n! \underbrace{\frac{p_1^{x_1}\, p_2^{x_2}}{x_1!\, x_2!} \sum_{x_3=0}^{n} \cdots \sum_{x_k=0}^{n} \frac{(n-x_1-x_2)!}{x_3! \cdots x_k!\, x_{k+1}!}\, \frac{p_3^{x_3} \cdots p_k^{x_k}\, p_{k+1}^{x_{k+1}}}{(1-p_1-p_2)^{n-x_1-x_2}}\, \mathbb{I}_{A_{k-2,\, n-x_1-x_2}}(x_3, \dots, x_k)}_{\mathbb{P}(Y_1 = x_3, \dots, Y_{k-2} = x_k)} =$$

$\mathbb{P}(Y_1 = x_3, \dots, Y_{k-2} = x_k)$ onde

$(Y_1, \dots, Y_{k-2}) \sim \text{Mult}\left(n - x_1 - x_2, \left(\frac{p_3}{1-p_1-p_2}, \frac{p_4}{1-p_1-p_2}, \dots, \frac{p_k}{1-p_1-p_2}\right)\right)$

$\Rightarrow$

$$P(X_1 = x_1, X_2 = x_2) = \frac{n!}{x_1!\, x_2!\, (n-x_1-x_2)!}\, p_1^{x_1}\, p_2^{x_2}\, (1-p_1-p_2)^{n-x_1-x_2}\, \mathbb{I}_{A_{2,n}}(x_1, x_2)$$

——— ⋅⋅ ———

I $X = (X_1, \dots, X_k)$ absolutamente contínuo

$$F_{X_1}(t) = F_X(t, \infty, \infty, \dots, \infty) = \int_{-\infty}^{t_1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_k)\, dx_k\, dx_{k-1} \dots dx_2\, dx_1$$

$$= \int_{-\infty}^{t_1} \underbrace{\left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_k)\, dx_k \dots dx_2 \right]}_{f_{X_1}(x_1)} dx_1$$

Rever Teorema de Fubini!

Exemplo:

$X = (X_1, X_2, \dots, X_k) \sim DIR(a_1, a_2, \dots, a_k, a_0)$

$$f_X(x_1, \dots, x_k) = \frac{\Gamma(a_0 + a_1 + \dots + a_k)}{\Gamma(a_0)\Gamma(a_1)\dots\Gamma(a_k)} x_1^{a_1 - 1} x_2^{a_2 - 1} \dots x_k^{a_k - 1} (1 - x_1 - \dots - x_k)^{a_0 - 1} \mathbb{I}_{S_k}(x_1, \dots, x_k)$$

$$f_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\Gamma(a_0 + \dots + a_k)}{\Gamma(a_0)\Gamma(a_1)\dots\Gamma(a_k)} x_1^{a_1 - 1} x_2^{a_2 - 1} x_3^{a_3 - 1} \dots x_k^{a_k - 1} (1 - x_1 - \dots - x_k)^{a_0 - 1} \mathbb{I}_{S_k}(x_1, \dots, x_k)\, dx_3 \dots dx_k$$

$$S_k = \{(t_1, \dots, t_k) \in \mathbb{R}_+^k : t_1 + \dots + t_k \leq 1\}$$

$$= \frac{\Gamma(a_0 + a_1 + \dots + a_k)}{\Gamma(a_1)\Gamma(a_2)} x_1^{a_1 - 1} x_2^{a_2 - 1} \mathbb{I}_{S_2}(x_1, x_2) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{x_3^{a_3 - 1} \dots x_k^{a_k - 1} (1 - x_1 - \dots - x_k)^{a_0 - 1}}{\Gamma(a_0)\Gamma(a_1)\dots\Gamma(a_k)} \cdot \mathbb{I}_{S'_{k-2}}(x_3, \dots, x_k)\, dx_3 \dots dx_k,$$

onde

$$S'_{k-2} = \{(t_1, t_2, \dots, t_{k-2}) \in \mathbb{R}_+^{k-2} ; t_1 + \dots + t_{k-2} \leq 1 - x_1 - x_2\}$$

Fazendo

$$t_1 = \frac{x_3}{1-x_1-x_2}$$

$$t_2 = \frac{x_4}{1-x_1-x_2}$$

$\vdots$

$$t_{k-2} = \frac{x_k}{1-x_1-x_2}$$

$$= \frac{\Gamma(a_0+\dots+a_k)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_0+a_3+a_4+\dots+a_k)} x_1^{a_1-1} x_2^{a_2-1} \mathbb{I}_{S_2}(x_1,x_2) \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \frac{\Gamma(a_0+a_3+a_4+\dots+a_k)}{\Gamma(a_0)\Gamma(a_3)\dots\Gamma(a_k)} t_1^{a_3-1} t_2^{a_4-1} \dots t_{k-2}^{a_k-1} (1-t_1-\dots-t_{k-2})^{a_0-1} \cdot$$

$$\cdot (1-x_1-x_2)^{a_0+a_3+\dots+a_k-1} \cdot \mathbb{I}_{S_{k-2}}(t_1,\dots,t_{k-2})\, dt_1 \dots dt_{k-2}$$

$\Longrightarrow$

$$f_{X_1,X_2}(x_1,x_2) = \frac{\Gamma(a_0+\dots+a_k)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_0+a_3+a_4+\dots+a_k)} x_1^{a_1-1} x_2^{a_2-1} (1-x_1-x_2)^{a_0+a_3+\dots+a_k-1} \mathbb{I}_{S_2}(x_1,x_2) \cdot$$

$$\cdot \underbrace{\left[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \frac{\Gamma(a_0+a_3+a_4+\dots+a_k)}{\Gamma(a_0)\Gamma(a_3)\dots\Gamma(a_k)} t_1^{a_3-1} t_2^{a_4-1} \dots t_{k-2}^{a_k-1} (1-t_1-\dots-t_{k-2})^{a_0-1} \mathbb{I}_{S_{k-2}}(t_1,\dots,t_{k-2})\, dt_1 \dots dt_{k-2}\right]}_{=1}$$

$(X_1, X_2) \sim DIR(a_1, a_2, a_0+a_3+\dots+a_k)$

INDEPENDÊNCIA

$X = (X_1, \dots, X_k)$ vetor aleatório

Definição: Dizemos que $X_1, \dots, X_k$ são independentes se

$$\mathbb{P}(X_1 \in B_1, X_2 \in B_2, \dots, X_k \in B_k) = \prod_{i=1}^{k} \mathbb{P}(X_i \in B_i) \;,\; \forall B_i \in \mathcal{B}(\mathbb{R}) \;,\; \forall i = 1, \dots, k$$

Alternativamente,

Definição: $X_1, \dots, X_k$ são independentes se $\sigma(X_1), \sigma(X_2), \dots, \sigma(X_k)$ são independentes.

$$\sigma(X_i) = X_i^{-1}(\mathcal{B}(\mathbb{R}))$$

Exemplo 1.

$$\Omega = \{1,2,3,4,5,6\}$$

$$X_1(\omega) = \mathbb{I}_{\{1,2\}}(\omega)$$

$$X_2(\omega) = \mathbb{I}_{\{2,4,6\}}(\omega)$$

$$X_3(\omega) = \mathbb{I}_{\{4,5,6\}}(\omega)$$

$$\{\omega : X_1(\omega) \leq t\} = \begin{cases} \emptyset \;,\; t < 0 \\ \{3,4,5,6\} \;,\; 0 \leq t < 1 \\ \Omega \;,\; t \geq 1 \end{cases}$$

$$\sigma(X_1) = \{\emptyset, \{1,2\}, \{3,4,5,6\}\}$$

$$\sigma(X_2) = \{\emptyset, \Omega, \{2,4,6\}, \{1,3,5\}\}$$

$$\sigma(X_3) = \{\emptyset, \Omega, \{1,2,3\}, \{4,5,6\}\}$$

Como $\sigma(X_1)$ e $\sigma(X_2)$ são independentes, dizemos que X_1 e X_2 são independentes.

Como $\sigma(X_1)$ e $\sigma(X_3)$ não são independentes (pois $\mathbb{P}(\{1,2\} \cap \{1,2,3\}) \neq \mathbb{P}(\{1,2\}) \cdot \mathbb{P}(\{1,2,3\})$), dizemos que X_1 e X_3 não são independentes.

Nesse caso, dizemos que X_1 e X_3 são <u>dependentes</u>.

Verificar que X_2 e X_3 são dependentes!

Ex. 2:

$\Omega = \mathbb{Z}_*$

$\mathcal{F} = \mathcal{P}(\Omega)$

$\mathbb{P}(\{j\}) = \left(\frac{1}{2}\right)^{|j|+1}$

$X_1(\omega) = \begin{cases} 1, & \omega > 0 \\ 0, & \text{c.c.} \end{cases}$ $\qquad X_2(\omega) = |\omega| = \begin{cases} \omega, & \omega > 0 \\ -\omega, & \omega < 0 \end{cases}$

Daí,

$\sigma(X_1) = \{\emptyset, \Omega, \mathbb{Z}_*^-, \mathbb{Z}_*^+\}$

$\{\omega : X_2(\omega) \leq t\} = \begin{cases} \emptyset, & t < 1 \\ \{-1, 1\}, & 1 \leq t < 2 \\ \{-2, -1, 1, 2\}, & 2 \leq t < 3 \\ \vdots \end{cases}$

$\sigma(X_2) = \sigma(\{\{-n, n\} : n \geq 1\}) = \left\{ \bigcup_{n \in J} \{-n, n\} : J \subseteq \{1, 2, 3, \ldots\} \right\}$

Como $\sigma(X_1)$ e $\sigma(X_2)$ são independentes, temos que as variáveis aleatórias X_1 e X_2 são independentes.

- Nessas condições, significa que dado o modulo do número, continuamos incertos quanto ao sinal do mesmo.

INDEPENDÊNCIA

Resultado: $X_1, \ldots, X_k$ são independentes $\Leftrightarrow F_X(t_1, \ldots, t_k) = F_{X_1}(t_1) \ldots F_{X_k}(t_k)$, para todo $(t_1, t_2, \ldots, t_k) \in \mathbb{R}^k$.

$X = (X_1, \dots, X_k)$

$\mathbb{P}(X_1 \in B_1, \dots, X_k \in B_k) = \prod_{i=1}^{k} \mathbb{P}(X_i \in B_i)$

$(\Rightarrow)$ $B_j = (-\infty, t_j]$, $j = 1, \dots, k$

$(\Leftarrow)$? Gimos

1. $F_{(X_1,X_2)}(t_1,t_2) = (1-e^{-t_1})(1-e^{-t_2}-t_2e^{-t_2})\mathbb{I}_{\mathbb{R}_+}(t_1)\mathbb{I}_{\mathbb{R}_+}(t_2)$

$F_{X_1}(t_1) = F_{(X_1,X_2)}(t_1,\infty) = (1-e^{-t_1})\mathbb{I}_{\mathbb{R}_+}(t_1)$

$F_{X_2}(t_2) = F_{(X_1,X_2)}(\infty,t_2) = (1-e^{-t_2}-t_2e^{-t_2})\mathbb{I}_{\mathbb{R}_+}(t_2)$

Logo, $F_{(X_1,X_2)}(t_1,t_2) = F_{X_1}(t_1)F_{X_2}(t_2)$, $\forall (t_1,t_2) \in \mathbb{R}^2$ e, portanto,

X_1 e X_2 são independentes.

2. $F_{(X_1,X_2)}(t_1,t_2) = \left(1-e^{-\min\{t_1, t_2/2\}}\right)\mathbb{I}_{\mathbb{R}_+}(t_1)\mathbb{I}_{\mathbb{R}_+}(t_2)$

$F_{X_1}(t_1) = (1-e^{-t_1})\mathbb{I}_{\mathbb{R}_+}(t_1)$

$F_{X_2}(t_2) = (1-e^{-t_2/2})\mathbb{I}_{\mathbb{R}_+}(t_2)$

$F_{(X_1,X_2)}(1,2) = 1-e^{-1}$

$F_{X_1}(1) = 1-e^{-1}$ $F_{X_2}(2) = 1-e^{-1}$

Logo, X_1 e X_2 são Dependentes!

INDEPENDÊNCIA

Resultado: $X_1, X_2, \dots, X_k$ são independentes $\Leftrightarrow$

$\mathbb{P}(X_1 = x_1, \dots, X_k = x_k) = \mathbb{P}(X_1 = x_1)\dots\mathbb{P}(X_k = x_k)$, $\forall x_i \in X_i(\Omega)$, $i = 1, \dots, k$

$X = (X_1, \dots, X_k)$ discreto.

$(\Rightarrow)$ $B_j = \{x_j\}$, $j=1,\dots,k$

$(\Leftarrow)$ $F_X(t_1,\dots,t_k) = \mathbb{P}(X_1 \leq t_1, X_2 \leq t_2, \dots, X_k \leq t_k) =$

$= \mathbb{P}(X_1 \in B_1, X_2 \in B_2, \dots, X_k \in B_k)$, onde $B_i = (-\infty, t_i] \cap X_i(\Omega)$

prob. $(X_1,\dots,X_k) \in B_1 \times \dots \times B_k$ (prod. cartesiano)

conjunto finito ou enumerável

$$= \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} \cdots \sum_{x_k \in B_k} \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) =$$

$$= \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} \cdots \sum_{x_k \in B_k} \mathbb{P}(X_1 = x_1)\,\mathbb{P}(X_2 = x_2) \cdots \mathbb{P}(X_k = x_k) =$$

$$= \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} \cdots \sum_{x_{k-1} \in B_{k-1}} \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_{k-1} = x_{k-1}) \left(\sum_{x_k \in B_k} \mathbb{P}(X_k = x_k) \right) =$$

repetindo o processo de isolação

$$= \dots = \prod_{i=1}^{k} \underbrace{\left(\sum_{x_i \in B_i} \mathbb{P}(X_i = x_i) \right)}_{F_{X_i}(t_i)} \Rightarrow$$

$$F_X(t_1,\dots,t_k) = F_{X_1}(t_1)\,F_{X_2}(t_2) \cdots F_{X_k}(t_k)$$

e, pelo resultado anterior, $X_1,\dots,X_k$ são independentes!

1. Urna.

$$\mathbb{P}(X_1 = a, X_2 = b) = \frac{\binom{b-a-1}{2}}{\binom{20}{4}} \mathbb{I}_{\{1,\dots,17\}}(a)\, \mathbb{I}_{\{a+3,\dots,20\}}(b)$$

$$\mathbb{P}(X_1 = a) = \frac{\binom{20-a}{3}}{\binom{20}{4}} \mathbb{I}_{\{1,\dots,17\}}(a)$$

$$\mathbb{P}(X_2 = b) = \frac{\binom{b-1}{3}}{\binom{20}{4}} \mathbb{I}_{\{4,\dots,20\}}(b)$$

$a=5$

$b=4$

$$\mathbb{P}(X_1=5, X_2=4)=0$$

$$\mathbb{P}(X_1=5)=\frac{\binom{15}{3}}{\binom{20}{4}} > 0 \quad e \quad \mathbb{P}(X_2=4)=\frac{1}{\binom{20}{4}} > 0$$

Como $\mathbb{P}(X_1=5, X_2=4) \neq \mathbb{P}(X_1=5)\mathbb{P}(X_2=4)$, então X_1 e X_2 são dependentes!

Exemplo 2

$x_1 \backslash x_2$	1	2	3	
0	1/12	2/12	1/12	$\mathbb{P}(X_1=0)=1/3$
1	2/12	4/12	2/12	$\mathbb{P}(X_1=1)=2/3$
	$\mathbb{P}(X_2=1) = \frac{1}{4}$	$\mathbb{P}(X_2=2) = \frac{2}{4}$	$\mathbb{P}(X_3=3) = \frac{1}{4}$	

X_1 e X_2 são Independentes!

Exemplo 3.

$x_1 \backslash x_2$	1	2	3	
0	1/12	0	3/12	1/3
1	2/12	3/12	3/12	2/3
	1/4	1/4	2/4	

Note que os eventos $\{X_1=0\}$ e $\{X_2=1\}$ são INDEPENDENTES mas as v.a. X_1 e X_2 NÃO são independentes!

$$(\mathbb{P}(X_1=0, X_2=2) \neq \mathbb{P}(X_1=0)\mathbb{P}(X_2=2)).$$

Probabilidade e Inferência Estatística I

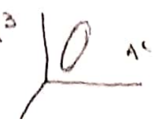
04/04/2013 - Aula 11 — Luiz Gustavo Esteves

$X=(X_1,\dots,X_k)$ absolutamente contínuo

$X_1,\dots,X_k$ são independentes $\Leftrightarrow f_X(x_1,\dots,x_k)=f_{X_1}(x_1)\dots f_{X_k}(x_k)$, para $(x_1,\dots,x_k)\in A \subseteq \mathbb{R}^k$, tal que "$\text{Volume}_k(A^c)=0$".

No $\mathbb{R}^2$ 

A^c não pode ser □

No $\mathbb{R}^3$ 

A^c não pode ter volume

$a=5$ $\quad$ $\mathbb{P}(X_1=5, X_2=4)=0$

$b=4$

$$\mathbb{P}(X_1=5)=\frac{\binom{15}{3}}{\binom{20}{4}}>0 \quad e \quad \mathbb{P}(X_2=4)=\frac{1}{\binom{20}{4}}>0$$

Como $\mathbb{P}(X_1=5, X_2=4) \neq \mathbb{P}(X_1=5)\,\mathbb{P}(X_2=4)$, então X_1 e X_2 são dependentes!

Exemplo 2.

$x_1 \backslash x_2$	1	2	3	
0	1/12	2/12	1/12	$P(X_1=0)=1/3$
1	2/12	4/12	2/12	$P(X_1=1)=2/3$
	$P(X_2=1) = \frac{1}{4}$	$P(X_2=2) = \frac{2}{4}$	$P(X_2=3) = \frac{1}{4}$	

X_1 e X_2 são Independentes!

Exemplo 3.

$x_1 \backslash x_2$	1	2	3	
0	1/12	0	3/12	1/3
1	2/12	3/12	3/12	2/3
	1/4	1/4	2/4	

Note que os eventos $\{X_1=0\}$ e $\{X_2=1\}$ são INDEPENDENTES mas as v.a. X_1 e X_2 NÃO são independentes!

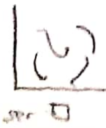
$(\mathbb{P}(X_1=0, X_2=2) \neq \mathbb{P}(X_1=0)\,\mathbb{P}(X_2=2))$.

Probabilidade e Inferência Estatística I

04/04/2013 - Aula 11 $\qquad$ Luiz Gustavo Esteves

$X=(X_1,\dots,X_k)$ absolutamente contínuo

$X_1,\dots,X_k$ são independentes $\Leftrightarrow f_X(x_1,\dots,x_k)=f_{X_1}(x_1)\cdots f_{X_k}(x_k)$, para $(x_1,\dots,x_k)\in A\subseteq \mathbb{R}^k$, tal que "$\text{Volume}_k$" $(A^c)=0$.

No $\mathbb{R}^2$  A^c $\qquad$ No $\mathbb{R}^3$  A^c

A^c não pode ser □ $\qquad$ A^c não pode ter volume

$(\Rightarrow)$

$$F_X(t_1,\dots,t_k) = F_{X_1}(t_1)\dots F_{X_k}(t_k)$$

$$\frac{\partial^k F_X}{\partial t_1 \dots \partial t_k} \downarrow$$

$$f_X(t_1,\dots,t_k) = f_{X_1}(t_1)\dots f_{X_k}(t_k)$$

$(\Leftarrow)$

$$\int_{-\infty}^{t_1}\int_{-\infty}^{t_2}\cdots\int_{-\infty}^{t_k} f_{X_1}(x_1)\dots f_{X_k}(x_k)\,dx_k\dots dx_1 = F_{X_1}(t_1)\dots F_{X_k}(t_k).$$

Transformações de Vetores Aleatórios

$$(\Omega,\mathcal{F},P) \xrightarrow{X} (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_X)$$

$T:\mathbb{R}^n \to \mathbb{R}^k$ $\qquad (X_1,\dots,X_n)\mapsto \left(\bar X, \frac{\sum(X_i-\bar X)^2}{n}\right)$

X v.a.

X^2

$\cos(X)$

$T:\mathbb{R}\to\mathbb{R}$

$x \to T(x) = x^2$

$$\begin{tikzcd}
(\Omega,\mathcal{F},P) \arrow[r, "X"] \arrow[rr, bend right, "Y = T\circ X"'] & (\mathbb{R},\mathcal{B}(\mathbb{R}),P_X) \arrow[r, "T"] & (\mathbb{R},\mathcal{B}(\mathbb{R}),\underset{?}{P_Y})
\end{tikzcd}$$

$Y(\omega) = T(X(\omega))$

$\{\omega\in\Omega : Y(\omega)\le t\} =$

$$= \{\omega\in\Omega : X^2(\omega)\le t\} = \begin{cases} \emptyset, & t<0 \in \mathcal{F} \\ \{\omega : X(\omega)\in[-\sqrt{t},\sqrt{t}]\} = X^{-1}([-\sqrt{t},\sqrt{t}]) \in \mathcal{F} \end{cases}$$

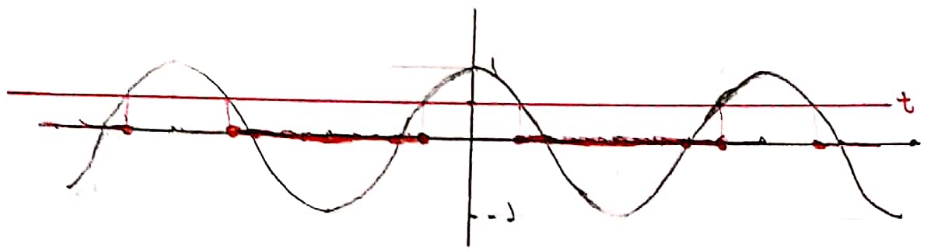
Logo, Y é v.a.

$U = \cos(X)$

$\{\omega \in \Omega : U(\omega) \leq t\} =$

$= \{\omega \in \Omega ; \cos(\omega) \leq t\} = \begin{cases} \emptyset^{\in \mathcal{F}}, & t < -1 \\ \bigcup_{n \in \mathbb{Z}} \{\omega \in \Omega : X(\omega) \in I_n\} \in \mathcal{F} \\ \Omega, & t > 1 \in \mathcal{F} \end{cases}$

Logo, U é v.a.,



$$\begin{tikzcd}(\Omega, \mathcal{F}, \mathbb{P}) \arrow[r, "X"] \arrow[rr, bend right, "Y = T \circ X"'] & (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P}_X) \arrow[r, "T"] & (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mathbb{P}_Y)\end{tikzcd}$$

$X : \Omega \to \mathbb{R}^n$

$T : \mathbb{R}^n \to \mathbb{R}^k$

$Y : \Omega \to \mathbb{R}^k$

$Y = T \circ X$

Se T é tal que $\forall B \in \mathcal{B}(\mathbb{R}^k)$

$T^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$ (T é $\mathcal{B}(\mathbb{R}^n)$-mensurável)

Vale que, para todo $B \in \mathcal{B}(\mathbb{R}^k)$

$$Y^{-1}(B) = (T \circ X)^{-1}(B) = X^{-1}(\underbrace{T^{-1}(B)}_{\in \mathcal{B}(\mathbb{R}^n)}) \in \mathcal{F}$$

$X_1, X_2 \quad Y = \min\{X_1, X_2\}$

$W = X_1 + X_2$

$X_1 \cdot X_2$

Logo, Y é v.a.

Nesse caso, a medida induzida por Y deve satisfazer:

$B \in \mathcal{B}(\mathbb{R}^k)$

$$\mathbb{P}_Y(B) = \mathbb{P}(Y^{-1}(B)) = \mathbb{P}(X^{-1}(\underbrace{T^{-1}(B)}_{A})) = \mathbb{P}_X(T^{-1}(B))$$

$\downarrow$

ou seja, a distribuição da Y é derivada da distribuição de X.

CASO DISCRETO

$X = (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$ discreto

Seja $A = \{(x_{1m}, x_{2m}, \dots, x_{nm}) : m \geq 1\} \subseteq \mathbb{R}^n$ tal que $\mathbb{P}_X(A) = 1$.

Seja $T: \mathbb{R}^n \to \mathbb{R}^k$ ($\mathcal{B}(\mathbb{R}^n)$-mensurável)

Seja $T(A) = \underbrace{\{T(x) : x \in A\}}_{\subseteq \mathbb{R}^k}$, $T(A)$ é enumerável (finito)

Seja $Y = T \circ X$

$$\mathbb{P}_Y(T(A)) = \mathbb{P}_X(T^{-1}(T(A))) \geq \mathbb{P}_X(A) = 1$$

com $T^{-1}(T(A)) \supseteq A$.

Para caracterizar $\mathbb{P}_Y$, basta avaliar para todo $y \in T(A)$,

$$\mathbb{P}_Y(\{y\}) = \mathbb{P}_X(T^{-1}(\{y\})) = \mathbb{P}(X \in T^{-1}(\{y\})) = \sum_{x \in A : T(x) = y} \mathbb{P}(X = x)$$

Exemplo:

$X \sim$ Uniforme $\{0, 1, 2, 3, 4\}$

$A = \{0, 1, 2, 3, 4\}$ $(\mathbb{P}_X(A) = 1)$

$T: \mathbb{R} \to \mathbb{R}$

$x \in \mathbb{R} \to \operatorname{sen} \frac{\pi x}{2}$

$T(A) = T(\{0, 1, 2, 3, 4\}) = \{0, 1, -1\}$

Para cada $y \in T(A)$, determinamos $\mathbb{P}_Y(\{y\}) = \mathbb{P}(Y = y)$.

$$\mathbb{P}_Y(\{-1\}) = \mathbb{P}_X(T^{-1}(\{-1\})) = \sum_{x \in A : \operatorname{sen} \frac{\pi x}{2} = -1} \mathbb{P}(X = x) = \mathbb{P}(X = 3) = \frac{1}{5}$$

$$\mathbb{P}_Y(\{0\}) = \sum_{x \in A : \operatorname{sen} \frac{\pi x}{2} = 0} \mathbb{P}(X = x) = \sum_{x \in \{0, 2, 4\}} \mathbb{P}(X = x) = \frac{3}{5}$$

$$\mathbb{P}_Y(\{1\}) = \mathbb{P}(Y = 1) = \frac{1}{5}$$

(2). $X = (X_1, X_2)$ tal que

X_1 e X_2 independentes, $X_i \sim Poi(\lambda_i)$, $\lambda_i > 0$, $i = 1, 2$.

$A = \mathbb{N}^2$ → conjunto enumerável $\quad \mathbb{P}(X_1 = x_1, X_2 = x_2) = \dfrac{e^{-\lambda_1} \lambda_1^{x_1}}{x_1!} \dfrac{e^{-\lambda_2} \lambda_2^{x_2}}{x_2!} \mathbb{I}_{\mathbb{N}^2}(x_1, x_2)$

$$T : \mathbb{R}^2 \to \mathbb{R}$$
$$(x_1, x_2) \mapsto T(x_1, x_2) = x_1 + x_2$$

$T(A) = \mathbb{N} \qquad Y = T(X)$

Para $y \in \mathbb{N} = T(A)$, temos

$$\mathbb{P}_Y(\{y\}) = \mathbb{P}(Y = y) = \sum_{(x_1,x_2) \in A,\ x_1 + x_2 = y} \mathbb{P}(X_1 = x_1, X_2 = x_2) =$$

$$= \sum_{(x_1,x_2) \in \{(0,y),(1,y-1),\dots,(y-1,1),(0,y)\}} \mathbb{P}(X_1 = x_1, X_2 = x_2) = \sum_{j=0}^{y} \mathbb{P}(X_1 = j, X_2 = y - j) =$$

$$= \sum_{j=0}^{y} \frac{e^{-\lambda_1} \lambda_1^{j}}{j!} \frac{e^{-\lambda_2} \lambda_2^{(y-j)}}{(y-j)!} \frac{y!}{y!} = \frac{e^{-(\lambda_1+\lambda_2)}}{y!} \underbrace{\sum_{j=0}^{y} \binom{y}{j} \lambda_1^{j} \lambda_2^{y-j}}_{(\lambda_1+\lambda_2)^y} \Rightarrow$$

$$\Rightarrow \mathbb{P}(Y = y) = \frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1+\lambda_2)^y}{y!}$$

Logo, $Y \sim Poi(\lambda_1 + \lambda_2)$.

Exemplo 3.

$X = (X_1, X_2, X_3)$ tal que

X_1, X_2 e X_3 são independentes,

$X_i \sim P(\lambda_i)$, $\lambda_i > 0$, $i = 1, 2, 3$

$A = \mathbb{N}^3$

$T: \mathbb{R}^3 \to \mathbb{R}^2$

$(x_1, x_2, x_3) \in \mathbb{R}^3 \longmapsto T(x_1, x_2, x_3) = (x_1 + x_3, x_2 + x_3)$

$Y = (Y_1, Y_2) = T(X)$

$T(A) = \mathbb{N}^2$

Para $(y_1, y_2) \in \mathbb{N}^2$,

$\mathbb{P}_Y(\{(y_1, y_2)\}) = \mathbb{P}(Y = (y_1, y_2)) =$

$$= \sum_{\substack{(x_1,x_2,x_3)\in\mathbb{N}^3 \,:\, x_1+x_3=y_1 \\ x_2+x_3=y_2}} \mathbb{P}(X_1 = x_1, X_2 = x_2, X_3 = x_3) =$$

$$= \sum_{\substack{(x_1,x_2,x_3)\in\mathbb{N}^3 \\ x_1+x_3=y_1 \\ x_2+x_3=y_2}} \frac{e^{-\lambda_1} \lambda_1^{x_1}}{x_1!} \frac{e^{-\lambda_2} \lambda_2^{x_2}}{x_2!} \frac{e^{-\lambda_3} \lambda_3^{x_3}}{x_3!}$$

$$= \sum_{j=0}^{\min(y_1,y_2)} \frac{e^{-\lambda_1} \lambda_1^{y_1-j}}{(y_1-j)!} \frac{e^{-\lambda_2} \lambda_2^{y_2-j}}{(y_2-j)!} \frac{e^{-\lambda_3} \lambda_3^{j}}{j!} \Rightarrow$$

$$\Rightarrow \mathbb{P}(Y_1 = y_1, Y_2 = y_2) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{j=0}^{\min\{y_1,y_2\}} \frac{\lambda_1^{y_1-j} \lambda_2^{y_2-j} \lambda_3^{j}}{(y_1-j)!\,(y_2-j)!\,j!}$$

$(Y_1, Y_2) \sim$ Holgate $(\lambda_1, \lambda_2, \lambda_3)$

$$\mathbb{P}(Y_1 = 0, Y_2 = 0) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \neq \mathbb{P}(Y_1 = 0)\,\mathbb{P}(Y_2 = 0) = e^{-(\lambda_1+\lambda_3)} e^{-(\lambda_2+\lambda_3)}$$

Y_1 e Y_2 são, portanto, dependentes!

CASO CONTÍNUO

X é uma v.a. abs. contínua com f.d. F_X e f.d.p. f_X.

Seja $T: \mathbb{R} \to \mathbb{R}$ crescente (decrescente) e consideremos $Y = T(X)$

$\exists^1 \, T^{-1}: \mathbb{R} \to \mathbb{R}$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(T(X) \leq y) = \mathbb{P}(T^{-1}(T(X)) \leq T^{-1}(y)) =$$

$$= \mathbb{P}(X \leq T^{-1}(y)) = F_X(T^{-1}(y))$$

Se T^{-1} é diferenciável, então

$$F_X'(T^{-1}(y)) = f_X(T^{-1}(y)) \frac{dT^{-1}(y)}{dy} = F_Y'(y) = f_Y(y)$$

- decrescente

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(T(X) \leq y) = \mathbb{P}(T^{-1}(T(X)) \geq T^{-1}(y)) =$$

$$= \mathbb{P}(X \geq T^{-1}(y)) = 1 - F_X(T^{-1}(y))$$

- Se T é crescente

$$F_Y(y) = F_X(T^{-1}(y))$$

Se $T^{-1}(.)$ é diferenciável,

$$f_Y(y) = f_X(T^{-1}(y)) \frac{dT^{-1}(y)}{dy}$$

- Se T é decrescente

$$F_Y(y) = 1 - F_X(T^{-1}(y))$$

Se $T^{-1}(.)$ é diferenciável

$$f_Y(y) = -f_X(T^{-1}(y)) \frac{dT^{-1}(y)}{dy}$$

Exemplo 1:

$X \sim N(\mu, \sigma^2)$

$T: \mathbb{R} \to \mathbb{R}$

$x \mapsto T(x) = \frac{x-\mu}{\sigma}$

$T^{-1}(y) = \sigma y + \mu$

$f_Y(y) = f_X(\sigma y + \mu) \cdot \sigma$

$= \frac{1}{\cancel{\sigma}\sqrt{2\pi}} e^{-\frac{(\sigma y + \cancel{\mu} - \cancel{\mu})^2}{2\sigma^2}} \cdot \cancel{\sigma} =$

$= \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \quad \therefore \; Y \sim N(0,1)$

Exemplo 2:

$X \sim U(0,1)$

$T: \mathbb{R}_+ \to \mathbb{R}$

$x \in \mathbb{R}_+ \to T(x) = -\frac{1}{\lambda} \log x$

$T^{-1}(y) = e^{-\lambda y}$

$Y = T(X)$

$T(x)$: decrescente

$f_Y(y) = f_X(e^{-\lambda y})(-\lambda e^{-\lambda y})(-1) =$

$= \lambda e^{-\lambda y} \, \mathbb{I}_{(0,1)}(e^{-\lambda y})$

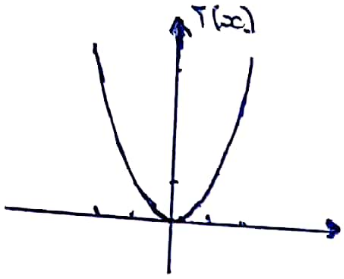
$\Rightarrow f_Y(y) = \lambda e^{-\lambda y} \, \mathbb{I}_{\mathbb{R}_+}(y)$

Logo, $Y \sim Exp(\lambda)$.

(3) $X \sim N(0,1)$

$T: \mathbb{R} \to \mathbb{R}$

$x \in \mathbb{R} \to T(x) = x^2$



$Y = X^2$

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = \begin{cases} 0, & y < 0 \\ = P(-\sqrt{y} \leq X \leq \sqrt{y}) = P(-\sqrt{y} < X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}), & y \geq 0 \end{cases}$$

Nesse caso, para $y > 0$

$$f_Y(y) = f_X(\sqrt{y})\,\frac{1}{2\sqrt{y}} - f_X(-\sqrt{y})\left(-\frac{1}{2\sqrt{y}}\right) \Rightarrow$$

$$f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}} = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{y})^2}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{y})^2}{2}}}{2\sqrt{y}} = \frac{\left(\frac{1}{2}\right)^{1/2}}{\sqrt{\pi}}\, y^{\frac{1}{2}-1}\, e^{-\frac{y}{2}}.$$

Daí,

$$f_Y(y) = \begin{cases} 0, & y < 0 \\ \dfrac{\left(\frac{1}{2}\right)^{1/2}}{\sqrt{\pi}}\, y^{\frac{1}{2}-1}\, e^{-y/2}, & y > 0 \end{cases}$$

$\therefore Y \sim \text{Gama}(1/2, 1/2) \sim \chi^2_1$.

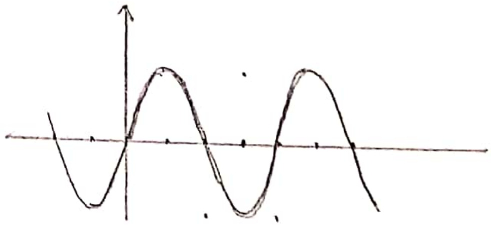
(4) $X \sim \mathcal{U}(0, 2\pi)$

$T: \mathbb{R} \to \mathbb{R}$

$x \in \mathbb{R} \mapsto T(x) = \operatorname{sen} x$

$Y = T(X) = \text{sen } X$

$\mathbb{P}(\text{sen } X \leq t)$



Probabilidade e Inferência Estatística I

Aula 12 - 05 de Abril de 2013

Luiz Gustavo Esteves

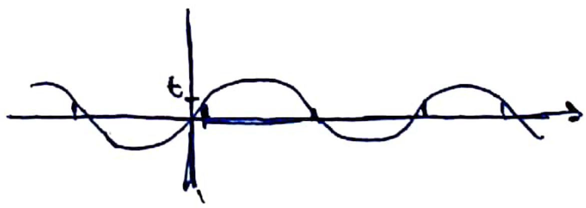
Transformação de V.A.

(2) $X \sim U(0, 2\pi)$

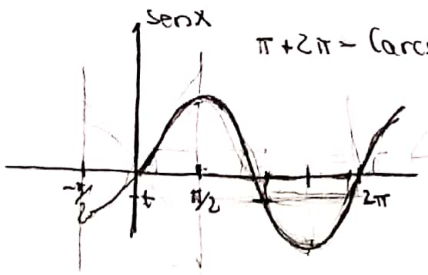
$T: \mathbb{R} \mapsto \mathbb{R}$

$x \in \mathbb{R} \mapsto T(x) = \text{sen } x$

$Y = T(X) = \text{sen } X$

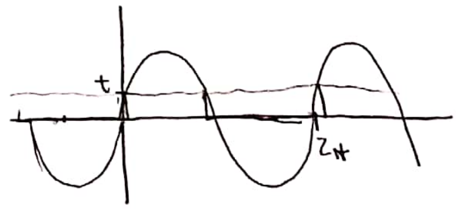


$$F_Y(t) = \mathbb{P}(Y \leq t) = \mathbb{P}(\text{sen } X \leq t) = \mathbb{P}(\text{sen } X \leq t,\ X \in (0, 2\pi)) = \begin{cases} 0, & t < -1 \\ \mathbb{P}(X \in [\pi - \text{arcsen} t, \text{arcsen} t + 2\pi]) & \\ 1, & t \geq 1 \end{cases}$$



$\pi + 2\pi - (\text{arcsen} t + 2\pi) =$

$= \pi - \text{arsen} t$

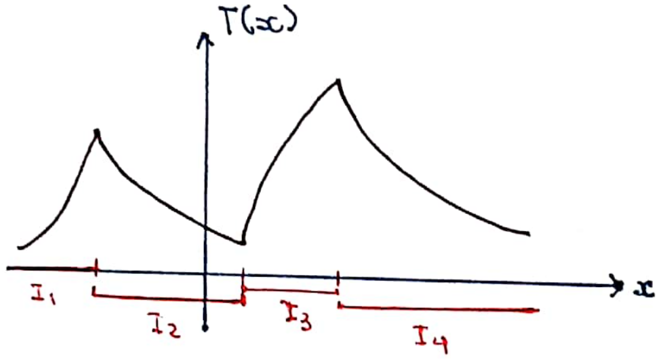


$$\Rightarrow F_Y(t) = \begin{cases} 0, & t < -1 \\ \mathbb{P}(X \in [\pi - \text{arcsen} t, \text{arcsen} t + 2\pi] = \dfrac{(2\pi + \text{arcsen} t) - (\pi - \text{arcsen} t)}{2\pi} = \dfrac{\pi + 2\,\text{arcsen} t}{2\pi}, & -1 < t < 0 \\ \mathbb{P}(X \in [0, \text{arcsen} t] \cup [\pi - \text{arcsen} t, 2\pi] = \dfrac{\text{arcsen} t + \pi + \text{arcsen} t}{2\pi}, & 0 < t < 1 \\ 1, & t \geq 1 \end{cases}$$

$$F_Y(t) = \begin{cases} 0, & t < -1 \\ \dfrac{\pi + 2\arcsin t}{2\pi}, & -1 \leq t \leq 1 \\ 1, & t > 1 \end{cases}$$

$$f_Y(y) = \frac{1}{\pi\sqrt{1-t^2}} \mathbb{I}_{(-1,1)}(t)$$

Em geral,



$A = \{x \in \mathbb{R} : f_X(x) > 0\}$ suporte de X

$I_1, \dots, I_r$ partição mensurável de A

Sejam $B_j = T(I_j)$, $j = 1, \dots, r$

Sejam $T_j : I_j \to B_j$

$x \in I_j \to T_j(x) = T(x)$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(T(X) \leq y) = \mathbb{P}(T(X) \leq y, X \in A) =$$

$$= \mathbb{P}(T(X) \leq y, X \in \bigcup_{j=1}^{r} I_j) = \mathbb{P}(X \in T^{-1}((-\infty, y]), X \in \bigcup_{j=1}^{r} I_j) =$$

$$= \mathbb{P}(X \in \bigcup_{j=1}^{r} I_j \cap T^{-1}((-\infty, y])) = \sum_{j=1}^{r} \mathbb{P}(X \in I_j \cap T^{-1}((-\infty, y])) =$$

$$= \sum_{j=1}^{r} \int_{I_j \cap T^{-1}((-\infty, y])} f_X(x)\,dx = \sum_{j=1}^{r} \int_{I_j \cap T_j^{-1}((-\infty, y])} f_X(x)\,dx = \sum_{j=1}^{r} \int_{T^{-1}((-\infty, y])} \mathbb{I}_{I_j}(x) f_X(x)\,dx =$$

$$= \sum_{J=1}^{r} \int_{-\infty}^{y} \mathbb{I}_{B_J}(w)\, f_X(T_J^{-1}(w)) \left| \frac{dT_J^{-1}(w)}{dw} \right| dw \Rightarrow$$

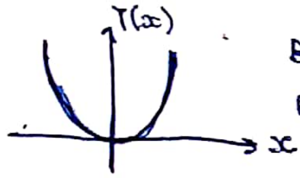
$$\Rightarrow F_Y(y) = \int_{-\infty}^{y} \underbrace{\left\{ \sum_{J=1}^{r} \mathbb{I}_{B_J}(w)\, f_X(T_J^{-1}(w)) \left| \frac{dT_J^{-1}(w)}{dw} \right| \right\}}_{f_Y(w)} dw$$

Voltando ao exemplo 1.

$X \sim N(0,1)$

$T(x) = x^2$

$Y = T(x)$



$B_1 = (0,\infty)$ $\quad T_1^{-1}(w) = -\sqrt{w}$

$B_2 = [0,\infty)$ $\quad T_2^{-1}(w) = \sqrt{w}$

$I_1 = (-\infty, 0)$

$I_2 = [0, \infty)$

$$f_Y(y) = \mathbb{I}_{\mathbb{R}_+}(w)\, f_X(-\sqrt{w}) \left| \frac{d(-\sqrt{w})}{dw} \right| + \mathbb{I}_{\mathbb{R}_+}(w)\, f_X(\sqrt{w}) \left| \frac{d\sqrt{w}}{dw} \right| =$$

$$= \mathbb{I}_{\mathbb{R}_+}(w) \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{w})^2}{2}} \cdot \left| -\frac{1}{2\sqrt{w}} \right| + \mathbb{I}_{\mathbb{R}_+}(w) \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{w})^2}{2}} \left| \frac{1}{2\sqrt{w}} \right| = \dots$$

Exemplo 2:

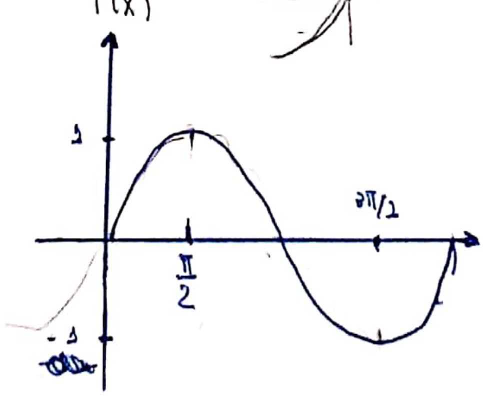
$X \sim U(0, 2\pi)$

$T(X) = \text{sen} X$

$I_1 = (0, \pi/2)$

$I_2 = (\pi/2, 3\pi/2)$

$I_3 = (3\pi/2, 2\pi)$



$B_1 = (0,1)$ $\quad T_1^{-1}(w) = \text{arcsen}\, w$

$B_2 = (-1,1)$ $\quad T_2^{-1}(w) = \pi - \text{arcsen}\, w$

$B_3 = (-1,0)$ $\quad T_3^{-1}(w) = 2\pi + \text{arcsen}\, w$

Assim,

$$f_Y(y) = \mathbb{I}_{(0,1)}(y)\,\frac{1}{2\pi}\left|\frac{1}{\sqrt{1-y^2}}\right| + \mathbb{I}_{(-1,1)}(y)\,\frac{1}{2\pi}\left|-\frac{1}{\sqrt{1-y^2}}\right| + \mathbb{I}_{(-1,0)}(y)\,\frac{1}{2\pi}\left|\frac{1}{\sqrt{1-y^2}}\right| =$$

$$= \frac{}{2\pi\sqrt{1-y^2}}\,2\,\mathbb{I}_{(-1,1)}(y) = \frac{1}{\pi\sqrt{1-y^2}}\,\mathbb{I}_{(-1,1)}(y).$$

CASO MULTIVARIADO

$X = (X_1, \dots, X_k)$ absolutamente contínua com densidade f_X. e $\int_A f(x)\,dx = 1$.

$T: \mathbb{R}^k \to \mathbb{R}^k$ mensurável

$Y = T(X)$

$Y_1 = (Y_1(X_1, \dots, X_k), Y_2(X_1, \dots, X_k), \dots, Y_k(X_1, \dots, X_k))$

Suponhamos que existam $I_1, \dots, I_r$ partição de A.

Seja $B_j = T(I_j)$ e definamos

$$\begin{aligned} T_j &: I_j \to B_j \\ x \in I_j &\mapsto T_j(x) = T(x) \end{aligned}, \quad j = 1, \dots, r \quad \text{com inversa}$$

$T_j^{-1} = (\mu_1^{(j)}, \dots, \mu_k^{(j)})$ de modo que existam as derivadas parciais

$\dfrac{\partial \mu_i^{(j)}}{\partial y_l}$ para todo $i, l = 1, \dots, k$ e $j = 1, \dots, r$.

Então,

$$f_Y(y_1, \dots, y_k) = \sum_{j=1}^{r} |J_j(y_1, \dots, y_k)|\, f_X(\mu_1^{(j)}(y_1, \dots, y_k), \mu_2^{(j)}(y_1, \dots, y_k), \dots, \mu_k^{(j)}(y_1, \dots, y_k))\,\mathbb{I}_{B_j}(y_1, \dots, y_k)$$

onde $J_j(y_1, \dots, y_k)$ é o determinante da matriz

$\underset{k\times k}{M^{(1)}}(y_1,\dots,y_k)$, onde $\quad M_{il}^{(1)}(y_1,\dots,y_k) = \dfrac{d\mu_i^{(1)}}{dy_l}$, $i,l=1,\dots,k$

Exemplo: $X=(X_1,X_2)$

$X_i \sim \text{Gama}(a_i/2, 1/2) \sim \chi_{a_i}$, $i=1,2$. $T: \mathbb{R}_+^2 \to \mathbb{R}_+^2$

$(x_1,x_2) \mapsto \left(\frac{a_2}{a_1}\frac{x_1}{x_2}, x_2\right)$

X_1 e X_2 independentes

$\dfrac{\frac{X_1}{a_1}}{\frac{X_2}{a_2}}$?

$Y = T(X)$

$Y = (Y_1,Y_2) = \left(\frac{a_2}{a_1}\frac{X_1}{X_2}, X_2\right)$

$T^{-1}(y_1,y_2) = \left(\frac{a_1}{a_2} y_1 y_2, y_2\right)$

$$J(y_1,y_2) = \begin{vmatrix} \frac{a_1}{a_2} y_2 & \frac{a_1}{a_2} y_1 \\ 0 & 1 \end{vmatrix} = \frac{a_1}{a_2} y_2 .$$

$T(\mathbb{R}_+^2) = \mathbb{R}_+^2$

$$f_Y(y_1,y_2) = \frac{a_1}{a_2}|y_2| \cdot f_X\left(\frac{a_1}{a_2} y_1 y_2, y_2\right) \mathbb{I}_{\mathbb{R}_+^2}(y_1,y_2) =$$

$$= \frac{a_1}{a_2} y_2 \frac{\left(\frac{1}{2}\right)^{\frac{a_1}{2}}}{\Gamma\left(\frac{a_1}{2}\right)} \left(\frac{a_1}{a_2} y_1 y_2\right)^{\frac{a_1}{2}-1} e^{-\frac{1}{2}\frac{a_1}{a_2} y_1 y_2} \frac{\left(\frac{1}{2}\right)^{\frac{a_2}{2}}}{\Gamma\left(\frac{a_2}{2}\right)} y_2^{\frac{a_2}{2}-1} e^{-\frac{1}{2} y_2} \mathbb{I}_{\mathbb{R}_+}(y_1)\, \mathbb{I}_{\mathbb{R}_+}(y_2)$$

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_Y(y_1,y_2)\, dy_2 = \frac{a_1}{a_2} \frac{\left(\frac{1}{2}\right)^{\frac{a_1+a_2}{2}}}{\Gamma\left(\frac{a_1}{2}\right)\Gamma\left(\frac{a_2}{2}\right)} y_1^{\frac{a_1}{2}-1} \mathbb{I}_{\mathbb{R}_+}(y_1) \,\cdot$$

$$\cdot \int_{-\infty}^{\infty} y_2^{\overbrace{\left(\frac{a_1}{2}+\frac{a_2}{2}\right)}^{a}-1} e^{-y_2 \overbrace{\left\{\frac{1}{2}\frac{a_1}{a_2} y_1 + \frac{1}{2}\right\}}^{b}} dy_2 \Rightarrow$$

$\Rightarrow$

$$f_{Y_1}(y_1) = \left(\frac{a_1}{a_2}\right)^{\frac{a_1}{2}} \frac{\left(\frac{1}{2}\right)^{\frac{a_1+a_2}{2}}}{\Gamma\left(\frac{a_1}{2}\right)\Gamma\left(\frac{a_2}{2}\right)} y_1^{\frac{a_1}{2}-1} \frac{\Gamma\left(\frac{a_1+a_2}{2}\right)}{\left[\frac{1}{2}+\frac{1}{2}\frac{a_1}{a_2}y_1\right]^{\frac{a_1+a_2}{2}}} \mathbb{I}_{\mathbb{R}_+}(y_1) \qquad Y_1 \sim F(a_1, a_2)$$

$$X \sim Beta\left(\frac{a_1}{2}, \frac{a_2}{2}\right)$$

$$\frac{\frac{X}{a_1/2}}{\frac{1-X}{a_2/2}} \sim F(a_1, a_2)$$

Exemplo 2. (exercício)

$X_1 \sim N(0,1)$

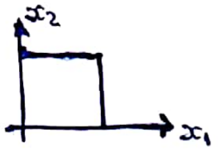
$X_2 \sim Gama\left(\frac{n}{2}, \frac{1}{2}\right) \sim \chi^2_n$

$$Y_1 = \frac{X_1}{\sqrt{\frac{X_2}{n}}} \sim t_n$$

Exemplo 3. X_1, X_2 independentes e identicamente distribuídas (v.a.i.i.d) segundo o modelo uniforme $(0,1)$.

$Y_1 = X_1 + X_2$

$Y_2 = X_2$

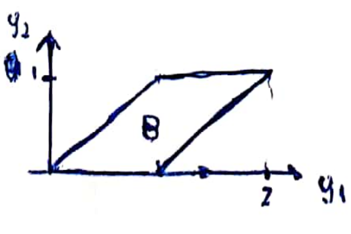


$A = (0,1)^2$

$T(x_1, x_2) = (x_1 + x_2, x_2)$ (bijetora)

$$T^{-1}(y_1, y_2) = (\underbrace{y_1 - y_2}_{w_1(y_1,y_2)}, \underbrace{y_2}_{w_2(y_1,y_2)})$$

$B = T(A)$



$$J(y_1,y_2) = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

$$f_Y(y) = |1| . f_{X}(y_1 - y_2, y_2)\, \mathbb{I}_B(y_1,y_2) =$$

$$= f_{X_1}(y_1-y_2) f_{X_2}(y_2)\, \mathbb{I}_B(y_1,y_2) =$$

$$= 1 . 1 . \mathbb{I}_B(y_1,y_2)$$

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_Y(y_1,y_2)\, dy_2 =$$

$$= \int_{-\infty}^{\infty} \mathbb{I}_B(y_1,y_2)\, dy_2 = \begin{cases} \int_{-\infty}^{\infty} \mathbb{I}_{(0,y_1)}(y_2)\, dy_2 & , 0 < y < 1 \\ \int_{-\infty}^{\infty} \mathbb{I}_{(y_1-1,1)}(y_2)\, dy_2 & , 1 < y < 2 \end{cases} \Rightarrow$$

$$\Rightarrow$$

$$f_{Y_1}(y_1) = \begin{cases} y_1 - 0 = y_1 , & 0 < y_1 < 1 \\ 1 - (y_1 - 1) = 2 - y_1 , & 1 < y_1 < 2 \end{cases}$$

$$\mathbb{I}_{(0,1)}(y_1-y_2)\, \mathbb{I}_{(0,1)}(y_2) = 1$$

$$\Downarrow$$

$$0 \leq y_1 - y_2 \leq 1 \qquad -y_1 \leq -y_2 \leq 1 - y_1$$

$$0 \leq y_2 \leq 1 \qquad 1 - y_1 \geq y_2 \geq y_1$$

$$\Downarrow$$

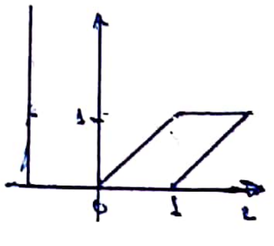
$$y_2 \leq y_1$$
$$y_2 \geq 1 - y_1$$

$$\max\{0, 1-y_1\} \leq y_2 \leq \min\{y_1, 1\}$$

Probabilidade e Inferênncia Estatística I

Aula 13 - 09/04/13 — Luis Gustavo Esteves

Aula passada



$$\int \mathbb{I}_B(y_1,y_2) =$$

$$= \mathbb{I}_{(0,1)}(y_1)\,\mathbb{I}_{(0,y_1)}(y_2) + \mathbb{I}_{(1,2)}(y_1)\,\mathbb{I}_{(y_1-1,1)}(y_2).$$

—— " —— " —— " ——

Estatísticas de Ordem

$X = (X_1, \dots, X_k)$ absolutamente contínua com densidade de f_X.

$X_{(j)}$ a j-ésima menor componente do vetor X

$$X_1(\omega) = \min\{X_1(\omega), \dots, X_k(\omega)\}$$

$$X_2(\omega) = \min\{X_1(\omega), \dots, X_k(\omega)\} \setminus \{X_{(1)}(\omega)\}$$

$$\vdots$$

$$X_n(\omega) = \text{máx}\{X_1(\omega), \dots, X_k(\omega)\}$$

$$T: \mathbb{R}^k \to \mathbb{R}^k$$

$$(x_1,\dots,x_k) \in \mathbb{R}^k \to T(x_1,\dots,x_k) = (x_{(1)}, x_{(2)}, \dots, x_{(k)})$$

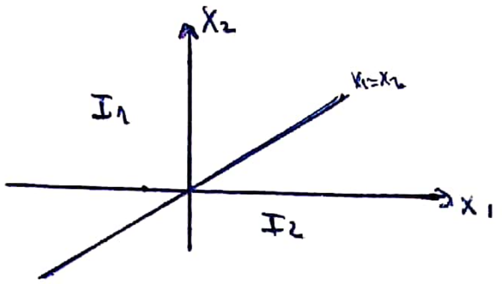
$$Y = T(X)$$

$k = 2$

$$(1,3) \to T(1,3) = (1,3)$$

$$(3,1) \to T(3,1) = (1,3)$$

$k=2$



$I_1=\{(a,b)\in\mathbb{R}^2 : a<b\}$, $T(I_1)=I_A=B=B_1$

$I_2=\{(a,b)\in\mathbb{R}^2 : a>b\}$, $T(I_2)=I_1=B=B_2$

$T_1 : I_1 \to I_1$

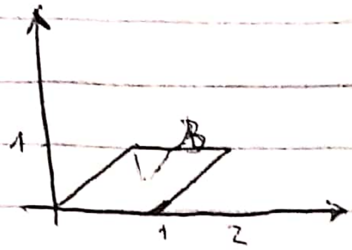
$(x_1,x_2)\longmapsto T_1(x_1,x_2)=T_1(x_1,x_2)=(x_1,x_2)$

$\Downarrow$

$T_1^{-1}(y_1,y_2)=(y_1,y_2)$

09/04

Aula Passada



$\mathbb{I}_B(y_1,y_2) =$

$= \mathbb{I}_{(0,1)}(y_1)\, \mathbb{I}_{(0,y_1)}(y_2) + \mathbb{I}_{(1,2)}(y_1)\, \mathbb{I}_{(y_1-1,1)}(y_2)$

Estatísticas de Ordem

$X = (X_1, \dots, X_K)$ absolutamente contínuo c/ dens. de f_X

$X_{(j)}$ a j-ésima menor componente do vetor X

$X_{(1)}(\omega) = \min\{X_1(\omega), X_2(\omega), \dots, X_K(\omega)\}$

$X_{(2)}(\omega) = \min\{X_1(\omega), \dots, X_K(\omega)\} \setminus \{X_{(1)}(\omega)\}$

$X_{(K)}(\omega) = \max\{X_1(\omega), \dots, X_K(\omega)\}$

$T: \mathbb{R}^K \to \mathbb{R}^K$

$(x_1, \dots, x_K) \in \mathbb{R}_K \longmapsto T(x_1, \dots, x_K) = (x_{(1)}, x_{(2)}, \dots, x_{(K)})$

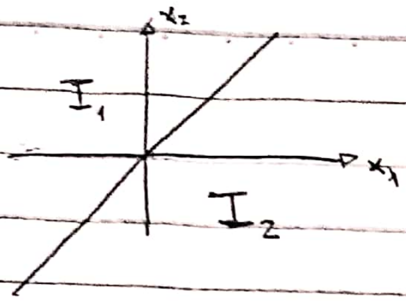
$Y = T(X)$

$K=2$

$(1,3) \longrightarrow T(1,3) = (1,3)$

$(3,1) \longrightarrow T(3,1) = (1,3)$

$k=2$



$I_1 = \{(a,b) \in \mathbb{R}^2 : a < b\}$ $\quad T(I_1) = I_1 = B = B_1$

$I_2 = \{(a,b) \in \mathbb{R}^2 : a > b\}$ $\quad T(I_2) = I_1 = B = B_2$

$T_1 : I_1 \to I_1$

$(x_1, x_2) \longmapsto T_1(x_1,x_2) = T(x_1,x_2) = (x_1, x_2)$

$\Downarrow$

$T_1^{-1}(y_1,y_2) = (y_1, y_2)$

$u_1^{(1)}(y_1;y_2) = y_1$

$u_2^{(1)}(y_1,y_2) = y_2$

$T_2 : I_2 \longrightarrow I_1$

$(x_1;x_2) \longmapsto T_2(x_1,x_2) = T(x_1,x_2) = (x_2, x_1)$

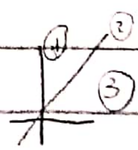
$T_1^{-1}(y_1, y_2) = (y_2, y_1)$

$u_1^{(2)}(y_1,y_2) = y_2 \qquad u_2^{(2)}(y_1, y_2) = y_1$

$$f_Y(y_1,y_2) = f_X(y_1,y_2)\left|J^{(1)}(y_1,y_2)\right| \mathbb{I}_{I_1}(y_1,y_2) + f_X(y_2,y_1)\left|J^{(2)}(y_1,y_2)\right| \mathbb{I}_{I_1}(y_1,y_2)$$

$\Downarrow$

$$\left.\begin{aligned} J^{(1)}(y_1,y_2) &= \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \\ J^{(2)}(y_1,y_2) &= \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1 \end{aligned}\right\} \Rightarrow f_Y(y_1,y_2) = \mathbb{I}_{I_1}(y_1,y_2)\{\overbrace{f_X(y_1,y_2)}^{f_{X_1}(y_1)f_{X_2}(y_2)} + f_X(y_2,y_1)\}$$



$2! + 1$

∇ Para $k=3$
$$\left.\begin{aligned} &x_1 < x_2 < x_3 \\ &x_1 < x_3 < x_2 \\ &\vdots \end{aligned}\right\} \text{6 parcelas}$$

Em geral, particionamos $\mathbb{R}^k$ em $K!+1$ regiões

Para cada permutação $\pi:\{1,\dots,k\}\to\{1,\dots,k\}$, Considere a região de $\mathbb{R}^k$

$$x_{\pi(1)} < x_{\pi(2)} < \dots < x_{\pi(k)}$$

Assim,

$$f_Y(y_1,\dots,y_k) = \sum_{\substack{\pi:\ \pi \text{ é} \\ \text{permutação de } \{1,\dots,k\}}} f_X(y_{\pi(1)},\dots,y_{\pi(k)})\, \mathbb{I}_B(y_1,\dots,y_k)$$

$$B=\{(y_1,y_2,\dots,y_k)\in\mathbb{R}^k : y_1<y_2<\dots<y_k\}$$

Se $X_1,\dots,X_k$ são v.a.i.i.d com densidade f_{X_1}, vale que

$$f_Y(y_1,\dots,y_k) = K!\, f_{X_1}(y_1)\cdots f_{X_1}(y_k)\, \mathbb{I}_B(y_1,\dots,y_k)$$

Ex: $X_1,X_2,\dots,X_k$ v.a.i.i.d $Exp(\lambda)$

$$Y=(\overset{Y_1}{X_{(1)}}, X_{(2)},\dots,\overset{Y_k}{X_{(k)}})$$

$$f_Y(y_1,\dots,y_k) = K!\prod_{i=1}^{k}\left\{\lambda e^{-\lambda y_i}\,\mathbb{I}_{\mathbb{R}_+}(y_i)\right\}\mathbb{I}_B(y_1,\dots,y_k) =$$

$$= K!\,\lambda^k e^{-\lambda\sum_{i=1}^{k} y_i}\,\mathbb{I}_{B'}(y_1,\dots,y_k) \qquad B'=\{(y_1,\dots,y_k)\in\mathbb{R}^k : 0<y_1<y_2<\dots<y_k\}$$

/ /

$(X_1, \dots, X_K)$ vetor aleatório com $X_1, \dots, X_K$ independentes

$$F_{X_{(1)}}(t) = \mathbb{P}(X_{(1)} \leq t) =$$

$$= 1 - \mathbb{P}(X_{(1)} > t) = 1 - \mathbb{P}\left(\bigcap_{i=1}^{K} (X_i > t)\right) =$$

$$= 1 - \prod_{i=1}^{K} \mathbb{P}(X_i > t) = 1 - \prod_{i=1}^{K} \left[1 - F_{X_i}(t)\right]$$

$$F_{X(n)}(t) = \mathbb{P}(X_{(n)} \leq t) = \mathbb{P}\left(\bigcap_{i=1}^{K} (X_i \leq t)\right) = \prod_{i=1}^{K} F_{X_i}(t).$$

Exemplo:

(1) $X_1, \dots, X_n$ v.a. independentes $X_i \sim Exp(\lambda_i)$, $i=1,\dots,K$

$\lambda_i > 0 \; ; i=1,\dots,K)$

$$F_{X_{(1)}}(t) = 1 - \prod_{i=1}^{K} \left[1 - \underbrace{\left(1 - e^{-\lambda_i t}\right)}_{e^{-t\sum_{i=1}^{K}\lambda}}\right] \mathbb{I}_{\mathbb{R}_+}(t) = 1 - e^{-(\sum_{i=1}^{K}\lambda_i)K}$$

Logo, $X_{(1)} \sim Exp\left(\sum_{i=1}^{K} \lambda_i\right)$

(2) $X_1, \dots, X_K$ v.a.i.i.d $\mathcal{U}(0,1)$

$$F_{X(n)}(t) = \prod_{i=1}^{K} F_{X_i}(t) = \left[F_{X_1}(t)\right]^n = \begin{cases} 0, & t \leq 0 \\ t^n, & 0 < t \leq 1 \\ 1, & t > 1 \end{cases}$$

$$1 - \left[1 - F_{X_1}(t)\right]^n = \begin{cases} 0, & t \leq 0 \\ 1-(1-t)^n, & 0 \leq t \leq 1 \\ 1, & t \geq 1 \end{cases}$$

$n(1-t)^{n-1} \mathbb{I}_{(0,1)}(t)$

$$f_{X(n)}(t) = \begin{cases} n t^{n-1}, & 0 < t \leq 1 \\ 0, & \text{c.c.} \end{cases}$$

$X_{(n)} \sim Beta(n, 1)$

$X_{(1)} \sim Beta(1, n)$
Exercício

Transformações Lineares

$(X_1, \ldots, X_k)$ absolutamente contínua com dens. f_X

$$Y_i = \sum_{j=1}^{k} a_{ij} X_j \quad , i = 1, \ldots, k$$

$(Y_1, \ldots, Y_k)$ vetor aleatório de modo que $Y' = A . X'$.

$$A = (a_{ij})_{k \times k}$$

Se A é invertível, com inversa A^{-1}, podemos escrever

$$X' = A^{-1} . Y'$$

Seja $A^{-1} = (c_{ij})_{k \times k}$

Assim

$$x_i = \sum_{j=1}^{k} c_{ij} y_j \quad , \quad i = 1, \ldots, k.$$

$u_i(y_1, \ldots, y_k)$

$$\begin{array}{|cccc|} c_{11} & c_{12} & \cdots & c_{1k} \\ & & & \\ c_{k1} & c_{k2} & \cdots & c_{kk} \end{array}$$

$$\frac{dx_i}{dy_j} = \frac{d \sum_{l=1}^{k} c_{il} \cdot y_l}{d y_j} = c_{ij} \qquad J(y_1 \ldots y_k) = \det A^{-1} = \frac{1}{\det A}$$

$$f_Y(y_1, \ldots, y_k) = \frac{1}{|\det A|} f_X \left(\sum_{j=1}^{k} c_{1j} y_j , \sum_{j=1}^{k} c_{2j} y_j , \ldots , \sum_{j=1}^{k} c_{kj} y_j \right)$$

Se A é ortogonal (isto é, $A' = A^{-1}$), então

$\det A = \pm 1$ e $c_{ij} = a_{ji}$

$\det A = \det A' = \pm 1$

$\det A'$

$\det A$

Portanto

$$f_Y(y_1,\dots,y_k) = f_X\left(\sum_{j=1}^{k} a_{j1} y_j, \sum_{j=1}^{k} a_{j2} y_j, \dots, \sum_{j=1}^{k} a_{jk} y_j\right)$$

tranformação Linear ortogonal

Resultado $X=(X_1,\dots,X_k)$ com $X_1,\dots,X_k$ v.a. i.i.d $N(0,1)$.

Seja $A_{k\times k}$ ortogonal e $Y=(Y_1,\dots,Y_k)$, $Y'=A.X'$,

$$f_Y(y_1,\dots,y_k) = f_X\left(\sum_{j=1}^{k} a_{j1} y_j, \dots, \sum_{j=1}^{k} a_{jk} y_j\right) =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\sum_{j=1}^{k} a_{j1} y_j\right)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\sum_{j=1}^{k} a_{j2} y_j\right)^2}{2}} \cdots \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\sum_{j=1}^{k} a_{jk} y_j\right)^2}{2}} \Rightarrow$$

$$\Rightarrow f_Y(y_1,\dots,y_k) = \left(\frac{1}{\sqrt{2\pi}}\right)^k \cdot e^{-\sum_{i=1}^{k}\left(\sum_{j=1}^{k} a_{ji} y_j\right)^2/2} =$$

$$\sum_{i=1}^{k} b_i^2 = b.b'$$

$y.A$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-(yA)(yA)'/2} =$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-(yA)(A'.y')/2} = \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-y.y'/2}$$

$$\Rightarrow f_Y(y_1, \dots, y_k) = \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-\sum_{i=1}^{k} y_i^2/2}$$

Logo, $Y_1, \dots, Y_k$ são i.i.d $N(0,1)$

Exemplo: $k=2$

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

(X_1, X_2) vetor de v.a. i.i.d $N(0,1)$

$\frac{X_1 + X_2}{\sqrt{2}}$ e $\frac{X_1 - X_2}{\sqrt{2}}$ são ind. $N(0,1)$

Comentário

Se $A_{k\times k}$ é ortogonal, $X_{1\times k}$ e $Y' = A.X'$

então $\sum_{i=1}^{k} Y_i^2 = \sum_{i=1}^{k} X_i^2$

$$\sum y_i^2 = y.y' = (x.A')(A.x') = x.(A'A)x' = x.x' = \sum_{i=1}^{k} x_i^2$$

Resultado

$X_1, X_2, \dots, X_k$ são ind.
$g_i : \mathbb{R} \to \mathbb{R}$ mensuráveis $\Bigg\} \Rightarrow g_1(X_1), g_2(X_2), \dots, g_k(X_k)$ são ind.

/ /

$$Y = (\overset{Y_1}{g_1(X_1)}, \dots, \overset{Y_k}{g_k(X_k)})$$

$$F_Y(t_1,\dots,t_k) = \mathbb{P}(g_1(X_1) \le t_1, \dots, g_k(X_k) \le t_k) = \mathbb{P}(X_1 \in g_1^{-1}((-\infty,t_1]),$$
$$\dots, X_k \in g_k^{-1}((-\infty,t_k])) = \mathbb{P}\left(\bigcap_{i=1}^{k}(X_i \in g_i^{-1}((-\infty,t_i]))\right) =$$

$$= \prod_{i=1}^{k} \mathbb{P}(X_i \in g_i^{-1}((-\infty,t_i])) = \prod_{i=1}^{k} \mathbb{P}(g_i(X_i) \in (-\infty, t_i]) =$$

$$= \prod_{i=1}^{k} F_{g_i(X_i)}(t_i) = \prod_{i=1}^{k} F_{Y_i}(t_i)$$

Esse resultado é mais geral

$X_1,\dots,X_k$ são ind., então funções de "pedaços" de $(X_1,\dots,X_k)$ são independentes

$K=5$ $\quad X_1, X_2, X_3, X_4, X_5$ ind.

$g_1 : \mathbb{R}^2 \to \mathbb{R}$ $\quad g_1(X_1, X_4), g_2(X_5), g_3(X_2, X_3)$

$g_2 : \mathbb{R} \to \mathbb{R}$

$g_3 : \mathbb{R}^2 \to \mathbb{R}$ $\quad X_1 X_4 \quad \text{arctg}\, X_5 \quad e^{X_2/X_3+1}$

são ind

—·—

X_1, X_2 iid $N(0,1)$

$$\overbrace{\frac{X_1+X_2}{\sqrt{2}}}^{Y_1} \text{ e } \overbrace{\frac{X_1-X_2}{\sqrt{2}}}^{Y_2} \quad \text{i.i.d } N(0,1)$$

$\sqrt{2}\,\bar{X}_2$ $\qquad X_1^2 + X_2^2 = Y_1^2 + Y_2^2$ $\qquad \sum_{i=1}^{2}(X_i - \bar{X}_2)$

$$Y_2^2 = X_1^2 + X_2^2 - Y_1^2 = X_1^2 + X_2^2 - 2\bar{X}_2^2$$

No caso k-variado

$$A = \begin{pmatrix} 1/\sqrt{k} & 1/\sqrt{k} & \dots & 1/\sqrt{k} \\ & & & \\ & & & \\ & & & \end{pmatrix}$$ De Groot

$Y_1 = \sqrt{k} \cdot \bar{X}_k$ $\qquad Y_2^2 + Y_3^2 + \dots + Y_3^2 = X_1^2 + X_2^2 + \dots + X_k^2 - Y_1^2$

$$Y_2^2 + \dots + Y_k^2 = \sum_{i=1}^{k} X_i^2 - k\bar{X}_k^2 =$$

$$= \frac{\sum (X_i - \bar{X}_k)^2}{k}$$

$X_1, \dots, X_k$ i.i.d $N(\mu, \sigma^2)$

$\frac{X_1 - \mu}{\sigma}, \frac{X_2 - \mu}{\sigma}, \dots, \frac{X_k - \mu}{\sigma}$ são i.i.d $N(0,1)$

$\sqrt{k}\left(\frac{\bar{X}_k - \mu}{\sigma}\right)$ e $\sum_{i=1}^{k}\left(\frac{X_i - \mu}{\sigma}, \frac{\bar{X}_k - \mu}{\sigma}\right)^2 = \sum_{i=1}^{k} (X_i - \bar{X}_k)^2$

$\bar{X}_k \qquad S_k^2$

Normal Multivariada

$X_1, X_2, \ldots, X_k$ v.a.i.i.d $N(0,1)$

$g_i(x_i) = \sigma_i x_i + \mu_i$, $\sigma_i > 0$, $\mu_i \in \mathbb{R}$, $i = 1, \ldots, k$

$Y_i = g_i(X_i)$

$(Y_1, \ldots, Y_k)$ de Normas independentes

$Y' = A.X' + \mu'$, $A = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ 0 & & \ddots & \\ & & & \sigma_k \end{bmatrix}$

$\mu = (\mu_1, \mu_2, \ldots, \mu_k)$

A mais geral:

$Y' = AX' + \mu' \Rightarrow Y' - \mu' = A.X'$

$X' = A^{-1}(Y' - \mu')$; $A^{-1} = (c_{ij})_{k \times k}$

$\mu_i(y_1, \ldots, y_k)$ $x_i = \sum_{j=1}^{k} c_{ij}(y_j - \mu_j)$

$$\frac{dx_i}{dy_j} = \frac{d \sum_{l=1}^{k} c_{il}(y_l - \mu_l)}{dy_j} = c_{ij}$$

Pelo resultado, $f_Y(y_1,\dots,y_K) = \frac{1}{|\det A|} f_X\left(\sum_{j=1}^{K} c_{1j}(y_j-\mu_j), \sum_{j=1}^{K} c_{2j}(y_j-\mu_j), \dots, \sum_{j=1}^{K} c_{Kj}(\dots\right.$

$$= \frac{1}{|\det A|}\left(\frac{1}{\sqrt{2\pi}}\right)^K e^{-\frac{1}{2}\sum_{i=1}^{K}\left\{\sum_{j=1}^{K} c_{ij}(y_j-\mu_j)\right\}^2}$$

$$= \frac{1}{|\det A|}\left(\frac{1}{\sqrt{2\pi}}\right)^K e^{-\frac{1}{2}(y-\mu)(A^{-1})(y-\mu)'} =$$

$$= \frac{1}{|\det A|}\left(\frac{1}{\sqrt{2\pi}}\right)^K \cdot e^{-\frac{1}{2}(y-\mu)(A')^{-1}A^{-1}(y-\mu)}$$

$$\Rightarrow f_Y(y) = \frac{1}{|\det A|}\left(\frac{1}{\sqrt{2\pi}}\right)^K e^{-\frac{1}{2}(y-\mu)(A\cdot A')^{-1}(y-\mu)'}$$

$$\boxed{A\cdot A' = \Sigma} \overset{\text{Binet}}{\Rightarrow} \det A \cdot \det A' = \det \Sigma$$

$$f_Y(y) = \frac{1}{\sqrt{\det \Sigma}}\left(\frac{1}{\sqrt{2\pi}}\right)^K \cdot e^{-1/2\cdot(y-\mu)\Sigma^{-1}(y-\mu)'}$$

$$Y_i = \sum_{j=1}^{K} a_{ij}X_j + \mu_i \sim N\left(\mu_i, \sum_{j=1}^{K} a_{ij}^2\right)$$

$$Y \sim N_K(\mu, \Sigma)$$

Obs: Transf. de um vetor discreto resulta noutro vetor aleatório discreto.
Já transformação de v.a abs. contínua nem sempre resulta noutra contínua

Ex: $X \sim Exp(\lambda)$

$Y = [X]$ maior inteiro maior ou igual a X

Note que $\mathbb{P}(Y \in \mathbb{N}) = 1$

Para $y \in \mathbb{N}$

$$\mathbb{P}(Y=y) = \mathbb{P}(y \leq X < y+1) \overset{X \text{ cont.}}{=}$$

$$= F_X(y+1) - F_X(y) = (1 - e^{-\lambda(y+1)}) - (1 - e^{-\lambda y}) =$$

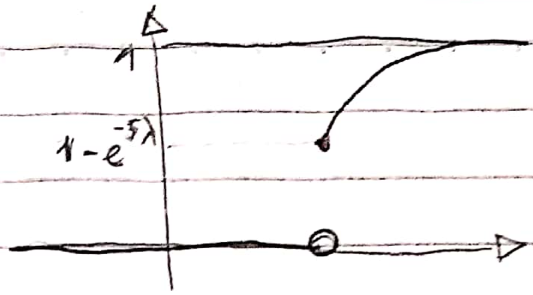
$$= e^{-\lambda y} - e^{-\lambda(y+1)} \Rightarrow \mathbb{P}(Y=y) = e^{-\lambda y}(1 - e^{-\lambda})$$

$$Y \sim Geo_{(0)}(1 - e^{-\lambda})$$

(2) $V = \max\{X, 5\}$

$$F_V(v) = \mathbb{P}(V \leq v) = \begin{cases} 0, & v < 5 \\ \mathbb{P}(X \leq v), & v \geq 5 \end{cases}$$

$$= \begin{cases} 0 & , v < 5 \\ 1 - e^{-\lambda v} & , v \geq 5 \end{cases}$$

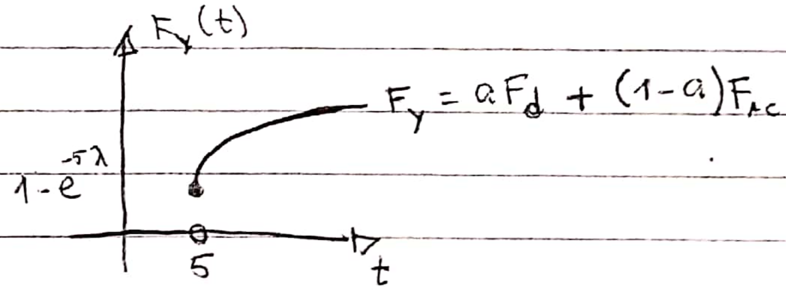


11/04

Esperança Matemática $X : \Omega \to \mathbb{R}$

X discreto $\to E(X) = \sum_{x} x \, \mathbb{P}(X = x)$

X abs contínua $\to E(X) = \int_{\mathbb{R}} x \, f_X(x) \, dx$



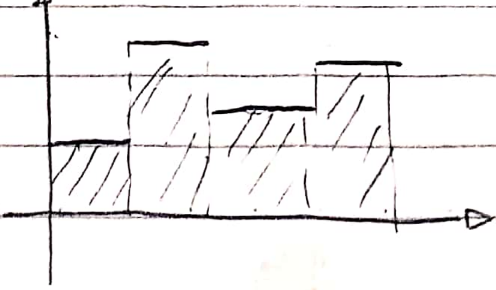
$\Omega = [0,1]$

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

$\mathbb{P}$: Comprimento

$$X(\omega) = \begin{cases} x_1 & , 0 \leq \omega \leq k_1 \\ x_2 & , k_1 < \omega \leq k_2 \\ x_3 & , k_2 < \omega \leq k_3 \\ x_4 & , k_3 \leq \omega \leq 1 \end{cases}$$

$$E(X) = \sum_{j=1}^{4} x_j \, \mathbb{P}(X = x_j) = \sum_{j=1}^{4} x_j (k_j - k_{j-1})$$



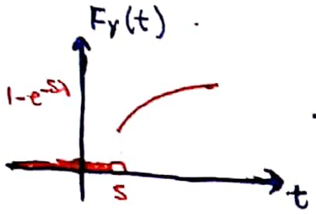
Probabilidade e Inferência Estatística

11/04/2013 - Aula 13 Luis Gustavo Esteves

Esperança Matemática

X discreta $\rightarrow E(X) = \sum_{x} x \, \mathbb{P}(X = x)$

X abs. contínua $\rightarrow E(X) = \int_{\mathbb{R}} x \, f_X(x) \, dx$



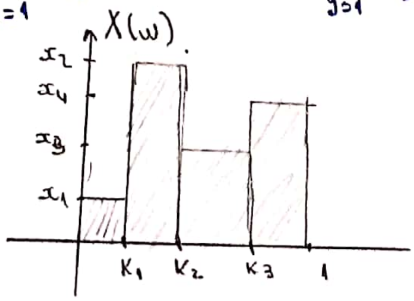
$- F_Y = a F_d + (1-a) F_{AC}$

$\Omega = [0,1]$ $\mathbb{P}$: comprimento

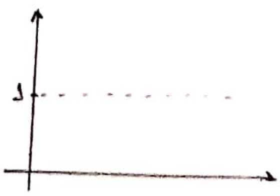
$X : \Omega \rightarrow \mathbb{R}$

$$X(\omega) = \begin{cases} x_1, & \overset{k_0}{0} \leq \omega \leq k_1 \\ x_2, & k_1 \leq \omega \leq k_2 \\ x_3, & k_2 \leq \omega \leq k_3 \\ x_4, & k_3 \leq \omega \leq 1 = k_4 \end{cases}$$

$$E(X) = \sum_{j=1}^{4} x_j \, \mathbb{P}(X = x_j) = \sum_{j=1}^{4} x_j (k_j - k_{j-1})$$



$$Y(\omega) = \begin{cases} 1, & \omega \in \mathbb{Q} \\ 0, & \omega \notin \mathbb{Q} \end{cases}$$



$E(Y) = 1 \cdot \mathbb{P}(Y = 1) + 0 \, \mathbb{P}(Y = 0)$

$= 1 \cdot 0 + 0 \, 1 = 0$ ✗

$(\Omega, \mathcal{F}, \mathbb{P})$

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

$\int_\Omega X \, d\mathbb{P} \to$ Integral de Lebesgue de X contra $\mathbb{P}$.

Definição: Uma função $X : \Omega \mapsto \mathbb{R}$ é chamada SIMPLES se assume apenas um número finito de valores ($X(\Omega) = \{X(\omega) : \omega \in \Omega\}$ é finito).

Em geral $X : \Omega \to \mathbb{R}$ simples pode ser representada da seguinte maneira:

$$X = \sum_{j=1}^{k} a_j \mathbb{I}_{A_j}, \text{ com} \qquad X = \sum_{j=1}^{k} a_j \mathbb{I}_{A_j} = a_1 \mathbb{I}_{A_1} + a_2 \mathbb{I}_{A_2} + a_3 \mathbb{I}_{A_3} + \dots + a_k \mathbb{I}_{A_k}$$

$A_1, A_2, \dots, A_k \in \mathcal{F}$ disjuntos e tais que $\Omega = A_1 \cup A_2 \cup \dots \cup A_k$.

Definição: Seja $X : \Omega \to \mathbb{R}_+$ simples, isto é, X é da forma $X = \sum_{j=1}^{k} a_j \mathbb{I}_{A_j}$, $a_j \geq 0$, $j = 1, \dots, k$ e $A_1, \dots, A_k$ partição de Ω.

A integral de X em relação (contra) $\mathbb{P}$ é definida por

$$\int_\Omega X \, d\mathbb{P} = \sum_{j=1}^{k} a_j \mathbb{P}(A_j).$$

Existem outras formas de definir a somatória:

$$X(\omega) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \mathbb{I}_{B_i \cap A_j} \qquad c_{ij} = b_i = a_j$$

$$\int_\Omega X \, d\mathbb{P} = \sum_{i=1}^{n} \sum_{j=1}^{n} \underset{b_i}{c_{ij}} \, \mathbb{P}(B_i \cap A_j).$$

$$\int_\Omega X \, d\mathbb{P} = \sum_{i=1}^{n} b_i \sum_{j=1}^{n} \mathbb{P}(B_i \cap A_j) = \sum_{i=1}^{n} b_i \mathbb{P}(B_i)$$

$$\int_{\Omega} X\,d\mathbb{P} = \sum_{i=1}^{n}\sum_{j=1}^{k} a_j\,\mathbb{P}(B_i\cap A_j) = \sum_{j=1}^{k} a_j \sum_{i=1}^{n}\mathbb{P}(B_i\cap A_j) = \sum_{j=1}^{k} a_j\,\mathbb{P}(A_j).$$

Propriedades

(1) $\int_{\Omega} X\,d\mathbb{P} \geq 0$

(2) $c>0$, $\int_{\Omega}(cX)\,d\mathbb{P} = \sum_{j=1}^{k}(ca_j)\,\mathbb{P}(A_j) = c\sum_{j=1}^{k} a_j\,\mathbb{P}(A_j) = c\int_{\Omega} X\,d\mathbb{P}$.

$cX = \sum_{j=1}^{k}(ca_j)\,\mathbb{I}_{A_j}$

(3) X, Y simples

$\int_{\Omega}(X+Y)\,d\mathbb{P}$

$X = \sum_{j=1}^{k} a_j\,\mathbb{I}_{A_j}$

$Y = \sum_{i=1}^{n} b_i\,\mathbb{I}_{B_i}$

$X+Y = \sum_{i=1}^{n}\sum_{j=1}^{k}(b_i+a_j)\,\mathbb{I}_{A_j\cap B_i}$

$$\int_{\Omega}(X+Y)\,d\mathbb{P} = \sum_{i=1}^{n}\sum_{j=1}^{k}(b_i+a_j)\,\mathbb{P}(A_j\cap B_i) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{k} b_i\,\mathbb{P}(A_j\cap B_i) + \sum_{i=1}^{n}\sum_{j=1}^{k} a_j\,\mathbb{P}(A_j\cap B_i)$$

$$= \sum_{i=1}^{n} b_i\,\mathbb{P}(B_i) + \sum_{j=1}^{k} a_j\,\mathbb{P}(A_j) =$$

$$= \int_{\Omega} Y\,d\mathbb{P} + \int_{\Omega} X\,d\mathbb{P}.$$

Funções Negativas

Seja $X:\Omega\to\mathbb{R}_+$

Seja $\mathcal{E}_X = \{ f:\Omega\to\mathbb{R}_+ : f \text{ é simples e } f\leq X\}$

Note que $\mathcal{E}_X \neq \emptyset$, pois $X_0(\omega)=0$, $\forall\,\omega\in\mathbb{R}$ é elemento de $\mathcal{E}_X$.

* ponder cada valor da v.a. pela medida do seu conjunto

Definição: Para $X: \Omega \to \mathbb{R}_+$, definimos

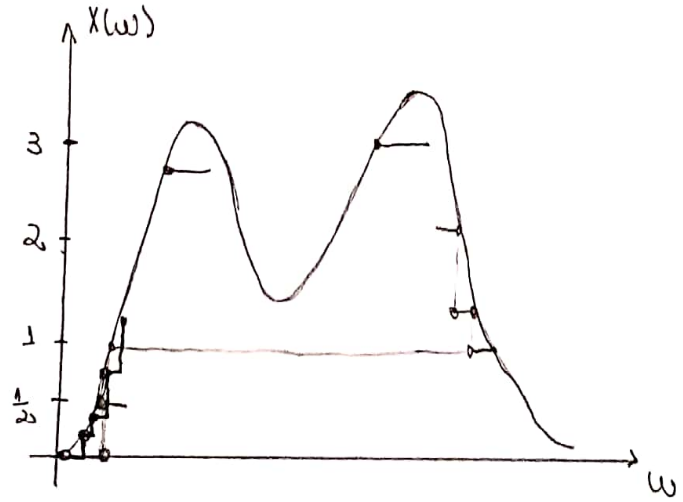
$$\int_\Omega X \, d\mathbb{P} = \sup\left\{ \int_\Omega f \, d\mathbb{P} : f \in \mathcal{E}_X \right\}$$

Alternativamente, podemos definir $\int_\Omega X \, d\mathbb{P}$ da seguinte maneira:

Seja $X_n: \Omega \to \mathbb{R}_+$ simples, $\forall n \geq 1$, com $X_n \leq X_{n+1}$, $\forall n \geq 1$ e tal que $X_n \uparrow X$ $(X_n(\omega) \uparrow X(\omega), \forall \omega \in \Omega)$. Então

$$\int_\Omega X \, d\mathbb{P} = \lim_{n \to \infty} \int_\Omega X_n \, d\mathbb{P}.$$

Mostrar que $X \leq Y$, então $\int_\Omega X \, d\mathbb{P} \leq \int_\Omega Y \, d\mathbb{P}$.



$$X_1(\omega) = \begin{cases} 0, & X(\omega) < 1/2 \\ 1/2, & X(\omega) \in [1/2, 1) \\ 1, & X(\omega) \geq 1 \end{cases}$$

$$X_2(\omega) = \sum_{i=0}^{7} \frac{i}{4} \mathbb{I}_{X^{-1}\left(\left[\frac{i}{2^2}, \frac{i+1}{2^2}\right)\right)}(\omega) + 2\,\mathbb{I}_{X^{-1}([2,\infty))}(\omega)$$

$$X_2(\omega) = \sum_{i=0}^{7} \frac{i}{4} \mathbb{I}_{X^{-1}\left(\left[\frac{i}{2^2}, \frac{i+1}{2^2}\right)\right)}(\omega) + 2\, \mathbb{I}_{X^{-1}([2,\infty))}(\omega)$$

$$X_n(\omega) = \sum_{i=0}^{n2^n-1} \frac{i}{2^n} \mathbb{I}_{X^{-1}\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right)}(\omega) + n\, \mathbb{I}_{X^{-1}([n,\infty))}(\omega)$$

Propriedades.

1. $\int_\Omega X \, d\mathbb{P} = \lim_{n\to\infty} \int_\Omega X_n \, d\mathbb{P} \geq 0$

2. $c > 0$, $X : \Omega \to \mathbb{R}_+$ $\qquad cX_n \uparrow cX$

$$\int_\Omega (cX)\, d\mathbb{P} = \lim_{n\to\infty} \int_\Omega (cX_n)\, d\mathbb{P} = \lim_{n\to\infty} c \int_\Omega X_n \, d\mathbb{P} = c \lim_{n\to\infty} \int X_n \, d\mathbb{P} = c\int_\Omega X \, d\mathbb{P}$$

3. $X, Y : \Omega \to \mathbb{R}_+$

$$\int_\Omega (X+Y)\, d\mathbb{P} = \qquad \begin{matrix} X_n \uparrow X \\ Y_n \uparrow Y \end{matrix} \qquad X_n + Y_n \uparrow X + Y$$

$$= \lim_{n\to\infty} \int_\Omega (X_n + Y_n)\, d\mathbb{P} = \lim_{n\to\infty} \left[\int_\Omega X_n \, d\mathbb{P} + \int_\Omega Y_n \, d\mathbb{P} \right] =$$

$$= \lim_{n\to\infty} \int_\Omega X_n \, d\mathbb{P} + \lim_{n\to\infty} \int_\Omega Y_n \, d\mathbb{P} = \int_\Omega X \, d\mathbb{P} + \int_\Omega Y \, d\mathbb{P}.$$

Resultado:

X v.a. assumindo valores em $\mathbb{N}$

$$E(X) = \int_\Omega X \, d\mathbb{P} = \sum_{n=1}^{\infty} \mathbb{P}(X \geq \omega)$$

$$X = \sum_{j=1}^{\infty} j\, \mathbb{I}_{A_j}, \quad A_j = X^{-1}(\{j\})$$

Seja $X_n = \sum_{j=1}^{n-1} j \, \mathbb{I}_{A_j} + n \, \mathbb{I}_{\bigcup_{j=n}^{\infty} A_j}$

$X_{n+1} = \sum_{j=1}^{n} j \, \mathbb{I}_{A_j} + (n+1) \mathbb{I}_{\bigcup_{j=n+1}^{\infty} A_j}$ $\qquad X_n \leq X_{n+1}$

$$\int_\Omega X_n \, d\mathbb{P} = \sum_{j=1}^{n-1} j \, \mathbb{P}(A_j) + n \, \mathbb{P}\Big(\bigcup_{j=n}^{\infty} A_j\Big)$$

$$= \sum_{j=1}^{n-1} \sum_{i=1}^{j} \mathbb{P}(A_j) + n \, . \, \mathbb{P}\Big(\bigcup_{j=n}^{\infty} A_j\Big)$$

Teorema de Tonelli
Teorema de Fubini: $\sum_{i}^{\infty} \sum_{j}^{\infty} a_{ij} = \sum_{j}^{\infty} \sum_{i}^{\infty} a_{ij}$

$$= \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} \mathbb{P}(A_j) + n \, \mathbb{P}(X \geq n) =$$

$$= \sum_{i=1}^{n-1} \mathbb{P}(i \leq X \leq n-1) + n \, \mathbb{P}(X \geq n) \overset{(*)}{=}$$

$$= \sum_{i=1}^{n} \mathbb{P}(X \geq i) = \int_\Omega X_n \, d\mathbb{P}$$

Assim,

$$\int_\Omega X \, d\mathbb{P} = \lim_{n \to \infty} \int_\Omega X_n \, d\mathbb{P} = \lim_{n \to \infty} \sum_{i=1}^{n} \mathbb{P}(X \geq i) = \sum_{i=1}^{\infty} \mathbb{P}(X \geq i).$$

(*) $\mathbb{P}(1 \leq X \leq n-1) + \mathbb{P}(X \geq n) \Rightarrow \mathbb{P}(X \geq 1)$

$\mathbb{P}(2 \leq X \leq n-1) + \mathbb{P}(X \geq n) \Rightarrow \mathbb{P}(X \geq 2)$

$\mathbb{P}(3 \leq X \leq n-1) + \mathbb{P}(X \geq n) \Rightarrow \mathbb{P}(X \geq 3)$

$\vdots$

$\mathbb{P}(n-1 \leq X \leq n-1) + \mathbb{P}(X \geq n) \Rightarrow \mathbb{P}(X \geq n-1)$

$\mathbb{P}(X \geq n) \Rightarrow \mathbb{P}(X \geq n)$

Ex. $X \sim Geo(p)$

$$E(X) = \int_\Omega X\, d\mathbb{P} = \sum_{i=1}^{\infty} \mathbb{P}(X \geq i) = *$$

$$\mathbb{P}(X \geq i) = \sum_{j=i}^{\infty} \mathbb{P}(X = j) = \sum_{j=i}^{\infty} (1-p)^{j-1} p = \frac{p(1-p)^{i-1}}{1-(1-p)} = (1-p)^{i-1}$$

$$* = \sum_{i=1}^{\infty} (1-p)^{i-1} = \frac{1}{1-(1-p)} = \frac{1}{p} \,\,_{\shortparallel}$$

Definição: Para $E \in \mathcal{F}$

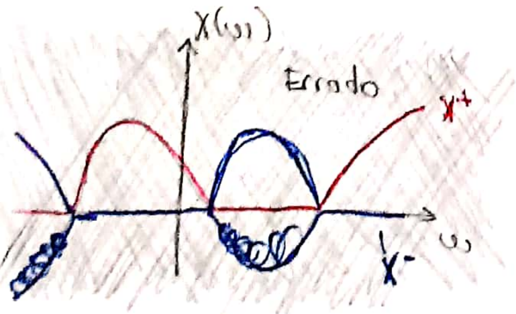
$$\int_E X\, d\mathbb{P} = \int_\Omega X \mathbb{I}_E \, d\mathbb{P}.$$

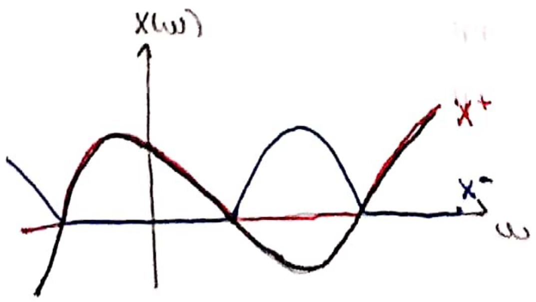
Caso Geral.

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

Sejam $X^+, X^- : \Omega \to \mathbb{R}_+$ dadas por

$$X^+(\omega) = \max\{X(\omega), 0\}$$
$$X^-(\omega) = \max\{-X(\omega), 0\}$$





$$X(\omega) = X^+(\omega) - X^-(\omega)$$

$$X = X^+ - X^-$$

$$|X| = X^+ + X^-$$

Definição: Seja $X : \Omega \to \mathbb{R}$.

Se $\int_\Omega X^+ dP$ ou $\int_\Omega X^- dP$ for FINITO, então definimos

$$\int_\Omega X \, dP = \int_\Omega X^+ dP - \int_\Omega X^- dP.$$

Probabilidade e Inferência Estatística I

12/04/2013 - Aula 14 Luiz Gustavo Esteves

(1) Funções simples
(2) Não-negativas
(3) Caso geral

$X : \Omega \mapsto \mathbb{R}$

$X^+ = \max\{X, 0\}$

$X^- = \max\{-X, 0\}$

$X = X^+ - X^-$

$|X| = X^+ + X^-$

Definição: Se $\int_\Omega X^+ dP < \infty$ ou $\int_\Omega X^- dP < \infty$, então definimos

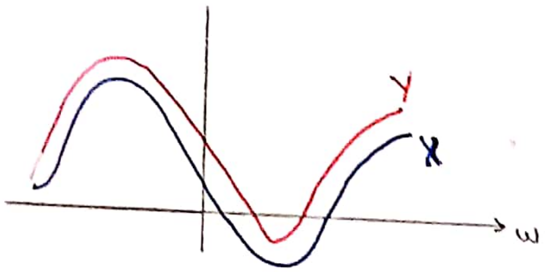
$$\int_\Omega X \, dP = \int_\Omega X^+ dP - \int_\Omega X^- dP$$

Comentário: Se $\int_\Omega X^+ dP < \infty$ e $\int_\Omega X^- dP < \infty$, dizemos que X é integrável.

Propriedades:

1. $X, Y: \Omega \to \mathbb{R}$, com $X \leq Y$ ($X(\omega) \leq Y(\omega), \forall\, \omega \in \Omega$)

   Se as integrais existem, $\int_\Omega X dP \leq \int_\Omega Y dP$.



É fácil verificar que

$$X^+(\omega) \leq Y^+(\omega) \quad e \quad Y^-(\omega) \leq X^-(\omega)$$

Logo, $\int_\Omega X dP = \int_\Omega X^+ dP - \int_\Omega X^- dP \leq \int_\Omega Y^+ dP - \int_\Omega Y^- dP = \int_\Omega Y dP$

$X, Y: \Omega \to \mathbb{R}^+$, $X \leq Y$

$$\int_\Omega X dP = \sup\left\{\int_\Omega f dP : f \in \mathcal{E}_X\right\}$$

$$\int_\Omega Y dP = \sup\left\{\int_\Omega f dP : f \in \mathcal{E}_Y\right\}$$

$$(\mathcal{E}_X \subseteq \mathcal{E}_Y, \text{ pois } X \leq Y)$$

2. $X: \Omega \to \mathbb{R}$, $c \in \mathbb{R}$

   cX

$$\int_\Omega (cX)\,d\mathbb{P} = c\int_\Omega X\,d\mathbb{P}$$

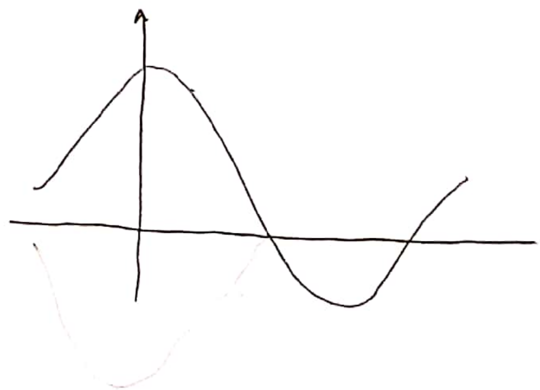
$c > 0$: $(cX)^+ = cX^+$

$(cX)^- = cX^-$

$$\int_\Omega (cX)\,d\mathbb{P} = \int_\Omega (cX)^+\,d\mathbb{P} - \int_\Omega (cX)^-\,d\mathbb{P} =$$

$$= \int_\Omega cX^+\,d\mathbb{P} - \int_\Omega cX^-\,d\mathbb{P} =$$

$$= c\int_\Omega X^+\,d\mathbb{P} - c\int_\Omega X^-\,d\mathbb{P} =$$



$\underline{c < 0}$: $(cX)^+ = (-c)X^-$

$(cX)^- = (-c)X^+$

$$\int_\Omega (cX)\,d\mathbb{P} = \int_\Omega (-c)X^-\,d\mathbb{P} - \int_\Omega (-c)X^+\,d\mathbb{P} =$$

$$= (-c)\int_\Omega X^-\,d\mathbb{P} + c\int_\Omega X^+\,d\mathbb{P}$$

3. $$\int_\Omega X+Y\,d\mathbb{P} = \int_\Omega \overset{<\infty}{X}\,d\mathbb{P} + \int_\Omega \overset{<\infty}{Y}\,d\mathbb{P}$$

$$|X+Y| \le |X|+|Y| \Rightarrow \int_\Omega |X+Y|\,d\mathbb{P} \le \int_\Omega \overset{<\infty}{|X|}\,d\mathbb{P} + \int_\Omega \overset{<\infty}{Y}\,d\mathbb{P}$$

$$\left.\begin{array}{l}(X+Y) = (X+Y)^+ - (X+Y)^- \\ (X+Y) = X^+ - X^- + Y^+ - Y^-\end{array}\right\} \Rightarrow$$

$$\Rightarrow (X+Y)^+ - (X+Y)^- = X^+ - X^- + Y^+ - Y^- \Rightarrow$$

$$\Rightarrow (X+Y)^+ + X^- + Y^- = X^+ + Y^+ + (X+Y)^-$$

$$\Rightarrow \int_\Omega (X+Y)^+ \, d\mathbb{P} + \int_\Omega X^- \, d\mathbb{P} + \int_\Omega Y^- \, d\mathbb{P} = \int_\Omega X^+ \, d\mathbb{P} + \int_\Omega Y^+ \, d\mathbb{P} + \int_\Omega (X+Y)^- \, d\mathbb{P} \Rightarrow$$

$$\Rightarrow \underbrace{\int_\Omega (X+Y)^+ \, d\mathbb{P} - \int_\Omega (X+Y)^- \, d\mathbb{P}}_{\int_\Omega (X+Y)\, d\mathbb{P}} = \underbrace{\int_\Omega X^+ \, d\mathbb{P} - \int_\Omega X^- \, d\mathbb{P}}_{\int_\Omega X \, d\mathbb{P}} + \underbrace{\int_\Omega Y^+ \, d\mathbb{P} - \int_\Omega Y^- \, d\mathbb{P}}_{\int_\Omega Y \, d\mathbb{P}}$$

$$\int_\Omega (X+Y)\, d\mathbb{P} = \int_\Omega X \, d\mathbb{P} + \int_\Omega Y \, d\mathbb{P}.$$

Exemplo:

X assume valores em $\{x_1, x_2, \dots\}$

$$X = \sum_{i=1}^{\infty} x_i \, \mathbb{I}_{A_i} \quad , \text{ onde } A_i = X^{-1}(\{x_i\}),\ i = 1, 2, \dots$$

$$X(\Omega) = \{X(\omega) : \omega \in \Omega\} = \{x_1, \dots, x_n, \dots\}$$

$$X(\Omega)_+ = X(\Omega) \cap \mathbb{R}_+ = \{x_{k_1}, x_{k_2}, \dots\}$$

$$X(\Omega)_- = X(\Omega) - X(\Omega)_+ = \{x_{k'_1}, x_{k'_2}, \dots\}$$

X^+ assume valores em $X(\Omega)_+$.

$$X^+ = \sum_{i=1}^{\infty} x_{k_i} \, \mathbb{I}_{A_{k_i}}$$

Se a soma acima for finita, X^+ é simples e $\int X^+ d\mathbb{P} = \sum_{x \in X(\Omega)_+} x \, \mathbb{P}(A_x)$

Se for finito, definimos

$$X_n^+ = \sum_{i=1}^{n} x_{k_i} \mathbb{I}_{A_{k_i}} \uparrow X^+$$

Daí, $\displaystyle\int_\Omega X_n^+ \, d\mathbb{P} = \sum_{i=1}^{n} x_{k_i} \mathbb{P}(A_{k_i})$ e

$$\int X^+ d\mathbb{P} = \lim_{n\to\infty} \int_\Omega X_n^+ d\mathbb{P} = \lim_{n\to\infty} \sum_{i=1}^{n} x_{k_i} \mathbb{P}(A_{k_i}) =$$

$$= \sum_{i=1}^{\infty} x_{k_i} \mathbb{P}(A_{k_i}) = \sum_{i: x_i \geq 0} x_{k_i} \mathbb{P}(A_{k_i}) = \sum_{i: x_i \geq 0} x_i \mathbb{P}(A_i) = \sum_{i: x_i \geq 0} x_i \mathbb{P}(X = x_i)$$

Analogamente,

$$\int_\Omega X^- d\mathbb{P} = \sum_{i: x_i < 0} (-x_i) \mathbb{P}(A_i) = \sum_{i: x_i < 0} (-x_i) \mathbb{P}(X = x_i)$$

∴ Se $\int_\Omega X^+ d\mathbb{P} < \infty$ ou $\int_\Omega X^- d\mathbb{P} < \infty$, vale que

$$\int_\Omega X \, d\mathbb{P} = \sum_{i: x_i \geq 0} x_i \mathbb{P}(X = x_i) - \sum_{i: x_i < 0} (-x_i) \mathbb{P}(X = x_i)$$

Exemplo 1:

$$\mathbb{P}(X = x) = \frac{1}{2} \left(\frac{2}{3}\right)^{-x-1} \cdot \frac{1}{3} \mathbb{I}_{\mathbb{Z}_-^*}(x) + \frac{1}{2} \left(\frac{1}{2}\right)^{x} \mathbb{I}_{\mathbb{Z}_+^*}(x)$$

$\frac{1}{2}$ Geométrica $(1/2, x)$ para $x \in \mathbb{Z}_+$

$\frac{1}{2}$ " $(1/3, x)$ " $x \in \mathbb{Z}_-$

$$\int_\Omega X^+ \, d\mathbb{P} = \sum_{i=1}^{\infty} x \frac{1}{2} \left(\frac{1}{2}\right)^x = \frac{1}{2} \underbrace{\sum_{i=1}^{\infty} i \,(1/2)^{i-1}}_{2} \cdot 1/2 = \frac{1}{2} \cdot 2 = 1$$

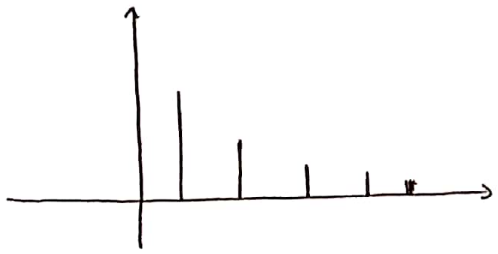
$$\int_\Omega X^- \, d\mathbb{P} = \sum_{i \in \mathbb{Z}^*_-} (-i)\, \mathbb{P}(X = i) = \sum_{i=1}^{\infty} i \,\mathbb{P}(X = -i) =$$

$$= \sum_{i=1}^{\infty} i \, \frac{1}{2} \left(\frac{2}{3}\right)^{i-1} \cdot \frac{1}{3} = \frac{1}{2} \cdot \frac{1}{1/3} = \frac{3}{2}.$$

Logo

$$\int_\Omega X \, d\mathbb{P} = 1 - \frac{3}{2} = -\frac{1}{2}$$

Exemplo 2. $\mathbb{P}(X = x) = C \frac{1}{x^2} \mathbb{I}_{\mathbb{Z}^*}(x)$



$$E(X^+) = \int_\Omega X^+ \, d\mathbb{P} = \sum_{x=1}^{\infty} x \, C \frac{1}{x^2} = \infty$$

$$E(X^-) = \int_\Omega X^- \, d\mathbb{P} = \sum_{x \in \mathbb{Z}^*_-} (-x)\, C \frac{1}{x^2} = \sum_{i=1}^{\infty} -x \, C \frac{1}{x^2} = -\infty.$$

Nesse caso, $\int_\Omega X^+ \, d\mathbb{P} = \infty$ e $\int_\Omega X^- \, d\mathbb{P} = -\infty$.

Logo, dizemos que não existe $\int_\Omega X \, d\mathbb{P}$.

$$
\begin{tikzcd}
(\Omega, \mathcal{F}, \mathbb{P}) \arrow[r, bend left, "X"] \arrow[rr, bend right, "Y = g \circ X."'] & (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X) \arrow[r, bend left, "g"] & (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_Y)
\end{tikzcd}
$$

Resultado:

$$\int_\Omega Y \, d\mathbb{P} \overset{\text{def}}{=} \int_\Omega g \circ X \, d\mathbb{P} = \int_{\mathbb{R}} g \, d\mathbb{P}_X$$ (Lei do Estatístico inconsciente)

Verificação:

$$g = \sum_{i=1}^{r} g_i \mathbb{I}_{B_i} \quad , \; g_1, g_2, \dots, g_r \geq 0$$

$$B_1, B_2, \dots, B_r \in \mathcal{B}(\mathbb{R})$$

$$\int_\Omega g \circ X \, d\mathbb{P} = \int_\Omega g(X) \, d\mathbb{P} = \int_\Omega \Big(\sum_{i=1}^{r} g_i \mathbb{I}_{B_i}(X) \Big) d\mathbb{P} =$$

$$X(\omega) \in B_i \;\Leftrightarrow\; \omega \in X^{-1}(B_i)$$

$$= \int_\Omega \Big(\sum_{i=1}^{r} g_i \mathbb{I}_{X^{-1}(B_i)} \Big) d\mathbb{P} = \sum_{i=1}^{r} g_i \mathbb{P}(X^{-1}(B_i)) = \sum_{i=1}^{r} g_i \mathbb{P}_X(B_i) = \int_{\mathbb{R}} g \, d\mathbb{P}_X .$$

$$\sum_{i=1}^{r} g_i \mathbb{I}_{B_i}(X(\omega)) = \sum_{i=1}^{r} g_i \mathbb{I}_{X^{-1}(B_i)}(\omega)$$

- Se X discreto

$$\sum_{i=1}^{r} g_i \sum_{x \in B_i} \mathbb{P}(X = x) = \sum_{i=1}^{r} g_i \sum_{x \in X(\Omega)} \mathbb{I}_{B_i}(x) \, \mathbb{P}(X = x) =$$

$$= \sum_{i=1}^{r} \Big\{ \sum_{x \in X(\Omega)} g_i \mathbb{I}_{B_i}(x) \, \mathbb{P}(X = x) \Big\} =$$

$$= \sum_{x \in X(\Omega)} \Big\{ \overbrace{\sum_{i=1}^{r} g_i \mathbb{I}_{B_i}(x)}^{g(x)} \Big\} \mathbb{P}(X = x)$$

Verificação:

$g \geq 0$, $g_n \uparrow g$, simples

$g_n(X) \uparrow g(X)$

$$\int (g \circ X)\, d\mathbb{P} = \lim_{n\to\infty} \underbrace{\int_{\Omega} (g_n \circ X)\, d\mathbb{P}}_{\int_{\mathbb{R}} g_n \, d\mathbb{P}_X} = \lim_{n\to\infty} \int_{\mathbb{R}} g_n \, d\mathbb{P}_X = \dots = \int_{\mathbb{R}} g \, d\mathbb{P}_X$$

- X contínuo

$$\sum_{i=1}^{r} g_i \int_{B_i} f_X(t)\, dt = \sum_{i=1}^{r} g_i \int_{-\infty}^{\infty} \mathbb{I}_{B_i}(t) f_X(t)\, dt = \dots = \int_{-\infty}^{\infty} \underbrace{\Big(\sum_{i=1}^{r} g_i \mathbb{I}_{B_i}(t)\Big)}_{g(t)} f_X(t)\, dt$$

$(g(x))^+ \cdot (g(x))^-$.

Como consequência, temos

X discreto: $\int_{\Omega} (g \circ X)\, d\mathbb{P} = E[g(X)] = \sum_x g(x) \mathbb{P}(X = x)$,

desde $\sum_x g^+(x) \mathbb{P}(X=x) < \infty$ e $\sum_x g^-(x) \mathbb{P}(X=x) < \infty$

X absolutamente contínua: $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$

Exemplo 1. $X \sim Beta(a,b)$, $a,b > 0$.

$g(t) = t^n (1-t)^m$, $n,m > 0$

$Y = g(X) = X^n (1-X)^m$

$$E(Y) = E(X^n(1-X)^m) = \int_{-\infty}^{\infty} x^n(1-x)^m f_X(x)\,dx =$$

$$= \int_0^1 x^n(1-x)^m \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\,dx =$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+n)\Gamma(b+m)}{\Gamma(a+b+n+m)} \underbrace{\int_0^1 \frac{\Gamma(a+n+b+m)}{\Gamma(a+n)\Gamma(b+m)} x^{a+n-1}(1-x)^{b+m-1}\,dx}_{1}$$

$$n=1,\ m=0 \Rightarrow E(X) = \frac{a}{a+b}$$

$$n=2,\ m=0 \Rightarrow E(X^2) = \frac{(a+1)\,a}{(a+b+1)(a+b)}$$

Exemplo 2. $X \sim$ Poisson

$$g: \mathbb{N} \to \mathbb{N}$$
$$n \in \mathbb{N} \mapsto g(n) = n!$$

$$Y = g(X) = X!$$

$$E(Y) = E(X!) = \sum_{x=0}^{\infty} x!\,\mathbb{P}(X=x) = \sum_{x=0}^{\infty} x!\,\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} e^{-\lambda}\lambda^x =$$

$$= \begin{cases} \infty, & \lambda \geq 1 \\ \dfrac{e^{-\lambda}}{1-\lambda}, & 0 < \lambda < 1 \end{cases}$$

Se X é vetor aleatório discreto (contínuo) a valores em $\mathbb{R}^k$ e $g: \mathbb{R}^k \to \mathbb{R}$ mensurável, vale que

$$E[g(X)] = \int_\Omega (g \circ X)\, dP = \sum_x g(x) P(X = x) \quad \left(\int_{\mathbb{R}^k} g(x) f_X(x)\, \underbrace{dx}_{dx_1 dx_2 \cdots dx_k} \right)$$

Exemplo.

$X = (X_1, \ldots, X_k) \sim \text{Multinomial}(n, (p_1, \ldots, p_k))$.

$g: \mathbb{R}^k \to \mathbb{R}$

$(x_1, \ldots, x_k) \to e^{t_1 x_1 + t_2 x_2 + \cdots + t_k x_k}$, $t_1, t_2, \ldots, t_k \in \mathbb{R}$

$Y = g(X_1, \ldots, X_k) = e^{t_1 X_1 + \cdots + t_k X_k}$

$E(Y) = ?$

$$E[Y] = E[g(X_1, \ldots, X_k)] = \sum_{(x_1, \ldots, x_k) \in A_{k,n}} e^{t_1 x_1 + \cdots + t_k x_k} \cdot P(X_1 = x_1, \ldots, X_k = x_k)$$

$A_{k,n} = \{(a_1, \ldots, a_k) \in \mathbb{N}^k : a_1 + a_2 + \cdots + a_k \leq n\}$

$p_{k+1} = 1 - p_1 - \cdots - p_k$

$$= \sum_{A_{k,n}} e^{t_1 x_1 + \cdots + t_k x_k} \cdot \frac{n!}{x_1! x_2! \cdots x_k! \underbrace{x_{k+1}!}_{n - x_1 - x_2 - \cdots - x_k}} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} (p_{k+1})^{x_{k+1}} =$$

$$= \sum_{A_{k,n}} \frac{n!}{\prod_{i=1}^{k+1} x_i!} (p_1 e^{t_1})^{x_1} \cdots (p_k e^{t_k})^{x_k} \cdot p_{k+1}^{x_{k+1}} =$$

$$= (p_1 e^{t_1} + \dots + p_k e^{t_k} + p_{k+1})^n \cdot \underbrace{\sum_{A_{k,n}} \frac{n!}{\prod_{i=1}^{k+1} x_i!} \frac{(p_1 e^{t_1})^{x_1} \cdots (p_k e^{t_k})^{x_k} \cdot p_{k+1}^{x_{k+1}}}{(p_1 e^{t_1} + \dots + p_k e^{t_k} + p_{k+1})^n}}_{= 1}$$

Então,

$$E[e^{t_1 X_1 + \dots + t_k X_k}] = (p_1 e^{t_1} + p_2 e^{t_2} + \dots + p_k e^{t_k} + 1 - p_1 - \dots - p_k)^n$$

Exemplo 4.

(X_1, X_2) contínuo

[redacted]

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{x_2} e^{-x_2} \mathbb{I}_{(0,x_2)}(x_1) \mathbb{I}_{\mathbb{R}_+}(x_2)$$

Exemplo 5.

$X = (X_1, \dots, X_k) \sim DIR(a_1, a_2, \dots, a_k, a_0)$ (Tarefa)

$E(X_1^{n_1} X_2^{n_2} \cdots X_k^{n_k})$.

Ex. 4.

$$E(Y) = E(X_1 X_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 \cdot x_2 \, f_{(X_1,X_2)}(x_1,x_2) \, dx_1 \, dx_2 =$$

$$= \int_0^{\infty} \int_0^{x_2} x_1 \cancel{x_2} \frac{1}{\cancel{x_2}} e^{-x_2} \, dx_1 \, dx_2 =$$

$$= \int_0^\infty e^{-x_2}\left[\int_0^{x_2} x_1\, dx_1\right] dx_2 = \int_0^\infty e^{-x_2}\frac{x_2^2}{2}\, dx_2 =$$

$$= \frac{1}{2}\,\frac{\Gamma(3)}{1^3}\int_0^\infty t^{3-1}e^{-t}\frac{1^3}{\Gamma(3)}\,dt = \frac{1}{2}\cdot 2! = 1$$

$$\int_0^\infty t^{3-1}e^{-t}\frac{1^3}{\Gamma(3)}\,dt = 1$$

$$\Rightarrow E(X_1 X_2) = \frac{1}{2}\,2! = 1.$$

Probabilidade e Inferência Estatística I

Aula 15 - 16/04/2013

Luiz Gustavo Esteves

$E(X_1+X_2) = E(X_1) + E(X_2)$

$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$

Ex 1. $X \sim Bin(n,p)$

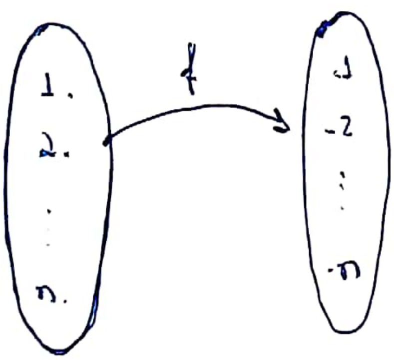
$$E(X) = \sum_{x=0}^{n} x\, P(X=x)$$

$X \overset{d}{=} X_1 + \dots + X_n$, onde

$X_1, \dots, X_n$ são v.a. i.i.d. de $Ber(p)$.

$$E(X) = E(X_1 + \dots + X_n) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \{\underbrace{1\,P(X_i=1) + 0\,P(X_i=0)}_{p}\} \Rightarrow E(X) = np$$

Exemplo 2.



$f: \{1,\dots,N\} \to \{1,\dots,N\}$

$\Omega = \{ f: \{1,\dots,n\} \to \{1,\dots,n\} : f \text{ é bijetora} \}$

X: Nº de pontos do domínio tais que $f(i) = i$.

$E(X) = ?$

$$X_i = \begin{cases} 1, & \text{se } f(i) = i \\ 0, & \text{c.c.} \end{cases}$$

$X_1 + \dots + X_N = X$

$$E(X) = E(X_1 + \dots + X_N) = \sum_{i=1}^{N} \underbrace{E(X_i)}_{1P(X_i=1)+0P(X_i=0) = \frac{(N-1)!}{N!} = \frac{1}{N}} = \sum_{i=1}^{N} \frac{1}{N} = 1.$$

(Sheldon Ross)

Exemplo 3.

Lance uma moeda n vezes.

"Run" de tamanho k, $k<n$, uma sequência de k caras seguidas.

$n=6$ $\quad$ $c\ c\ c\ \bar{c}\ c\ c$

$k=2$

3 "runs" de tamanho 2.

X: N. de "runs" de tamanho k

$P(X=j)=?$

$X_1, \dots, X_{n-k+1}$ variáveis:

$$X_i = \begin{cases} 1, & \text{ocorre um "run" de tamanho } k \text{ começando no } i\text{-ésimo lançamento} \\ 0, & \text{c.c.} \end{cases}$$

$X = X_1 + \dots + X_{n-k+1}$

X: Nº de runs de tamanho k

$P(X=j)=?$

$$X_i = \begin{cases} 1, & \text{ocorre "run" de tamanho } k \text{ i-ésimo} \\ 0, & \text{c.c.} \end{cases}$$

$X = X_1 + \dots + X_{n-k+1}$

$E(X) = E(X_1 + \dots + X_{n-k+1}) = E(X_1) + \dots + E(X_{n-k+1}) =$

$P(X_i = 1) = p^k \quad , i = 1, \dots, n-k+1$

$E(X) = (n-k+1)p^k.$

D

Resultado:

Sejam $X_1, \dots, X_k$ v.a. independentes e $g_1, g_2, \dots, g_k : \mathbb{R} \to \mathbb{R}$ mensuráveis tais que $g_1(X_1), \dots, g_k(X_k)$ são integráveis.

Então,

$$E[g_1(X_1)g_2(X_2)\dots g_k(X_k)] = E[g_1(X_1)] \dots E[g_k(X_k)].$$

Exemplo. $X_1 \sim Beta(a,b)$) Independentes
$X_2 \sim Poisson(\lambda)$

$$W = X_1^2 \cdot X_2 \qquad \begin{array}{l} g_1(t) = t^2 \\ g_2(t) = t \end{array}$$

$$E(W) = E(X_1^2 X_2) = E(X_1^2)E(X_2) = \frac{a(a+1)}{(a+b)(a+b+1)} \cdot \lambda$$

MOMENTOS

$X : \Omega \to \mathbb{R}$ uma v.a.

$g : \mathbb{R} \to \mathbb{R}$
$g(t) \to t^k$

$E[g(X)] = E[X^k] \to$ k-ésimo momento da v.a. X

$g(t) = (t - E(X))^k$, $E(X) < \infty$.

$E[g(X)] = E[(X - E(X))^k] \to$ k-ésimo momento central de X

$k=1 \rightarrow E[(X-E(X))]=0$

$k=2 \rightarrow E[(X-E(X))^2]=Var(X)$. Variância da v.a. X.

$Var(X)=\dots=E(X^2)-(E(X))^2$

Exemplo:

$X \sim Beta(a,b)$

Vimos que

$$E(X)=\frac{a}{a+b} \quad e \quad E(X^2)=\frac{a(a+1)}{(a+b)(a+b+1)}$$

$$Var(X)=E(X^2)-(E(X))^2=\frac{a(a+1)}{(a+b)(a+b+1)}-\left(\frac{a}{a+b}\right)^2=$$

$$=\frac{a(a+1)(a+b)-a^2(a+b+1)}{(a+b)^2.(a+b+1)}=\frac{ab}{(a+b)^2(a+b+1)}$$

Propriedades

1. $Var(X)\geq 0$

2. $Var(X)=0 \Leftrightarrow P(X=E(X))=1$.

$Y\geq 0$.

$E(Y)=0 \Rightarrow P(Y=0)=1$

$$P(Y\geq 1/n)=E\Big[\mathbb{I}_{\underbrace{\{w: Y(w)\geq 1/n\}}_{A_n}}\Big]=\int_\Omega \mathbb{I}_{A_n}\,dP=n\int_\Omega \frac{1}{n}\mathbb{I}_{A_n}\,dP$$

$$\leq \int_\Omega Y\,dP=0,\ \forall n\geq 1.$$

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{Y > 1/n\}\right) = 0 \Rightarrow$$

$$\Rightarrow \mathbb{P}(Y > 0) = 0 \overset{Y \geq 0}{\Rightarrow} \mathbb{P}(Y = 0) = 1.$$

5. $Var(aX+b) = a^2 Var(X)$ (Exercício)

Propriedade: Se X possui momento de ordem k, então possui momento de ordem j, para todo $j < k$.

$$E(|X^j|) = \int_\Omega |X^j| \, d\mathbb{P} = \int_\Omega |X^j| \, \mathbb{I}_{\{\omega : |X(\omega)| \leq 1\}} \, d\mathbb{P} + \int_\Omega |X^j| \, \mathbb{I}_{\{\omega : |X(\omega)| > 1\}} \, d\mathbb{P} \leq$$

$$\int_\Omega \mathbb{I}_{\{\omega : |X(\omega)| \leq 1\}} \, d\mathbb{P} + \int_\Omega |X^k| \, \mathbb{I}_{\{\omega : |X(\omega)| > 1\}} \, d\mathbb{P} \leq$$

$$1 + \int_\Omega |X^k| \, d\mathbb{P} < \infty.$$

Seja X uma variável aleatória,

Seja $\varphi_X : I \to \mathbb{R}_+$ a função que associa a cada $t \in I = (-\varepsilon, \varepsilon)$, para algum $\varepsilon > 0$, o número $\varphi_X(t) = E[e^{tX}]$.

À função φ_X damos o nome de FUNÇÃO GERADORA DE MOMENTOS.

EX 1. $X \sim Poisson(\lambda)$

$$\varphi_X(t) = E(e^{tX}) = \sum_{j=0}^{\infty} e^{tj} \, \mathbb{P}(X=j) = \sum_{j=0}^{\infty} e^{tj} \frac{e^{-\lambda}\lambda^j}{j!} = e^{-\lambda} \sum_{j=0}^{\infty} \frac{(e^t\lambda)^j}{j!} =$$

$$= e^{-\lambda} e^{\lambda e^t} \cdot \underbrace{\sum_{j=0}^{\infty} \frac{(e^t\lambda)^j}{j!} e^{-\lambda e^t}}_{1} \Rightarrow \varphi_X(t) = e^{\lambda e^t - \lambda}$$

Exemplo 2:

$X \sim$ Gama (a,b)

$$\varphi_X(t) = E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx = \int_0^{\infty} e^{tx} \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}\,dx =$$

$$= \frac{b^a}{(b-t)^a} \int_0^{\infty} \frac{(b-t)^a}{\Gamma(a)} x^{a-1} e^{-(b-t)x}\,dx \overset{t<b}{\Rightarrow} \varphi_X(t) = \left(\frac{b}{b-t}\right)^a.$$

Propriedades:

1. $\varphi_X(0) = 1$

2. $\varphi_{aX+b}(t) = E[e^{t(aX+b)}] = E[e^{tb} e^{(at)X}] = e^{tb} E[e^{(at)X}] \Rightarrow \varphi_{aX+b}(t) = e^{tb} \varphi_X(at)$

$\varphi_{aX+b}(t) = e^{tb} \varphi_X(at)$.

3. X, Y são independentes $\Rightarrow \varphi_{X+Y}(t) = \varphi_X(t) . \varphi_Y(t)$.

$$\varphi_{X+Y}(t) = E(e^{t(X+Y)}) = E(e^{tX} . e^{tY}) = E(e^{tX}) E(e^{tY}) = \varphi_X(t) \varphi_Y(t).$$

4. $E(X^k) < \infty$, $\forall k \geq 1$ e φ_X existe e admite derivadas de toda ordem $\Rightarrow$

$$\left.\frac{d^k \varphi_X(t)}{dt^k}\right|_{t=0} = E(X^k)$$

PS:

$$e^{tX} = \sum_{j=0}^{\infty} \frac{(tX)^j}{j!} \Rightarrow E(e^{tX}) = \int \Sigma\, dx \overset{\text{T. Fubini}}{=} \Sigma \int dx$$

Exemplo:

$X \sim$ Poisson (λ)

$\varphi_X(t) = e^{\lambda e^t - \lambda}$

$\varphi'_X(t) = e^{\lambda e^t - \lambda} . \lambda e^t$

$\varphi'_X(0) = \lambda = E(X)$.

$$\varphi''_X(t) = \lambda\left\{e^{\lambda e^t - \lambda} \cdot \lambda e^t \cdot e^t + e^{\lambda e^t - \lambda} \cdot e^t\right\}$$

$$\varphi''_X(0) = \lambda\{\lambda+1\} = \lambda^2 + \lambda = E(X^2)$$

$$Var(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Exemplo:

Seja (X,Y) um vetor aleatório contínuo com densidade

$$f_{X,Y}(x,y) = \frac{1}{y} e^{-\frac{x}{y}} e^{-y} \mathbb{I}_{\mathbb{R}_+}(x) \mathbb{I}_{\mathbb{R}_+}(y)$$

$$E(X^k) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k f_{X,Y}(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k \frac{1}{y} e^{-\frac{x}{y}} e^{-y}\,dx\,dy =$$

$$= \int_0^{\infty} \frac{1}{y} e^{-y}\left[\int_0^{\infty} \frac{(1/y)^{k+1}}{\Gamma(k+1)} x^k e^{-x/y}\,dx\right] \frac{\Gamma(k+1)}{(1/y)^{k+1}}\,dy \Rightarrow$$

$$E(X^k) = \int_0^{\infty} \Gamma(k+1) y^k e^{-y}\,dy = \Gamma(k+1)\frac{\Gamma(k+1)}{1}\int_0^{\infty} \frac{1^{k+1}}{\Gamma(k+1)} y^k e^{-y}\,dy \Rightarrow$$

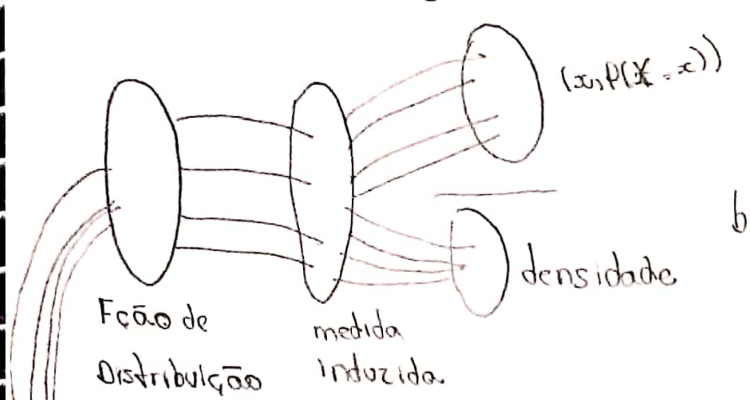
$$\Rightarrow E(X^k) = (k!)^2 < \infty.$$

$$\varphi_X(t) = E(e^{tX}) =$$

$$= \int_0^{\infty}\int_0^{\infty} e^{tx} \cdot \frac{1}{y} e^{-x/y} \cdot e^{-y}\,dx\,dy = \int_0^{\infty} \frac{1}{y} e^{-y}\left(\int_0^{\infty} e^{-\left(\frac{1}{y}-t\right)x}\,dx\right) dy$$

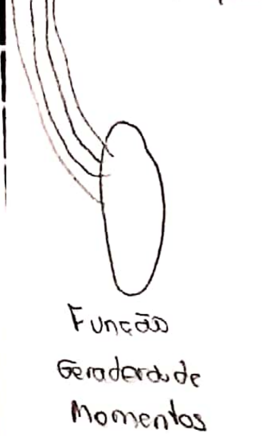
converge para $t \leq 0$.

Ramdas Bhat – Livro legal.



bijeção – se uma, então outra...

Agora,



Também á a $\alpha_X : \mathbb{R} \to \mathbb{C}$

$t \in \mathbb{R} \mapsto \alpha_X(t) = E[e^{itX}]$. Essa, sempre existe!

$\varphi_X(t) = E[e^{tX}] \to$ transformada de Laplace

$E(aX + bY) = aE(X) + bE(Y)$

Agora,

$$Var(X+Y) = E((X+Y)^2) - [E(X) + E(Y)]^2 =$$

$$= E(X^2) + 2E(XY) + E(Y^2) - (E(X))^2 - 2E(X)E(Y) - (E(Y))^2 =$$

$$= Var(X) + Var(Y) + 2\{\underbrace{E(XY) - E(X)E(Y)}_{}\}$$

$Cov(X,Y)$: covariância entre as variáveis X e Y

Podemos ainda escrever:

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$$

Exemplo:

$(X_1, X_2) \sim DIR(a_1, a_2, a_0)$ $\qquad x_1 \geq 0$, $x_2 \geq 0$, $x_1 + x_2 \leq 1$

$X_1 \sim Beta(a_1, a_0 + a_2)$

$X_2 \sim Beta(a_2, a_0 + a_1)$

$$E(X_1 X_2) = \int_{A_2} x_1 x_2 \frac{\Gamma(a_1+a_2+a_0)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_0)} x_1^{a_1-1} x_2^{a_2-1} (1-x_1-x_2)^{a_0-1} \, dx =$$

$$= \frac{\Gamma(a_1+a_2+a_0)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_0)} \cdot \frac{\Gamma(a_0)\Gamma(a_1+1)\Gamma(a_2+1)}{\Gamma(a_0+a_1+a_2+2)} \int_{A_2} \frac{\Gamma(a_0+a_1+a_2+2)}{\Gamma(a_0)\Gamma(a_1+1)\Gamma(a_2+1)} x_1^{(a_1+1)-1} x_2^{(a_2+1)-1} (1-x_1-x_2)^{a_0-1} \, dx$$

$$= \frac{a_1 \cdot a_2}{(a_0+a_1+a_2)(a_0+a_1+a_2+1)}.$$

Logo,

$$Cov(X_1, X_2) = \frac{a_1 \cdot a_2}{(a_0+a_1+a_2)(a_0+a_1+a_2+1)} - \frac{a_1}{(a_0+a_1+a_2)} \frac{a_2}{(a_0+a_1+a_2)} < 0.$$

Propriedades:

1. $Cov(X,Y) = Cov(Y,X)$
2. $Cov(X,X) = Var(X)$
3. $Cov\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j Cov(X_i, Y_j)$

Consequência.

$$Var\left(\sum_{i=1}^{n} X_i\right) = Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_i, X_j) = \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n} \sum_{j \neq i} Cov(X_i, X_j)$$

Voltando ao exemplo dos pareamentos

X_i: Nº de elementos de $\{1,\dots,N\}$ tais que $p(i)=i$.

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

$$Var(X) = \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n} \sum_{j \neq i} Cov(X_i, X_j)$$

Vimos que $X_i \sim Ber(1/N)$, $i=1,\dots,N$

$$Var(X_i) = E(X_i^2) - (E(X_i))^2 =$$

$$= \{1^2 . P(X_i=1) + 0^2 P(X_i=0)\} - \left(\frac{1}{N}\right)^2 =$$

$$= \frac{1}{N} - \left(\frac{1}{N}\right)^2$$

$i \neq j$

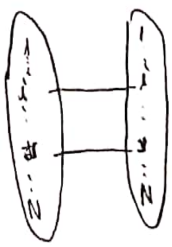
$$Cov(X_i, X_j) = E(X_i X_j) - E(X_i) E(X_j)$$

$$= P(X_i=1, X_j=1) - \frac{1}{N}\frac{1}{N} =$$

(X_1, X_2) é dem marginais Bernoulli, então

$X_1 X_2 \sim Ber(P(X_1=1, X_2=1))$

$$= \frac{(N-2)!}{N!} - \frac{1}{N^2} =$$



$$= \frac{1}{N(N-1)} - \frac{1}{N^2}.$$

Logo,

$$Var(X) = \sum_{i=1}^{N} Var(X_i) + \sum_{i=1}^{N} \sum_{j \neq i} Cov(X_i, X_j) = \sum_{i=1}^{N} \frac{1}{N}\left(\frac{N-1}{N}\right) + \sum_{i=1}^{N}\sum_{j\neq i} \frac{1}{N^2(N-1)} =$$

$$= \frac{N-1}{N} + (N^2-N)\frac{1}{N^2(N-1)} = \frac{N-1}{N} + \frac{1}{N} = 1$$

Seja (X,Y) vetor aleatório.

DEFINIÇÃO: À quantidade

$$\frac{Cov(X,Y)}{\sqrt{VAR(X)\,VAR(Y)}}$$ damos o nome de COEFICIENTE DE CORRELAÇÃO ENTRE X e Y.

Notação: $CORR(X,Y)$.

DEFINIÇÃO: Ao número $\sqrt{VAR(X)}$ damos o nome de DESVIO-PADRÃO de X.

NOTAÇÃO: $DP(X) = \sqrt{VAR(X)}$.

Assim,

$$CORR(X,Y) = \frac{COV(X,Y)}{DP(X)\,DP(Y)}$$

Propriedades:

(1) $CORR(X,Y) \in [-1,1]$ $\forall t \in \mathbb{R}$

$$E\left[\{t(X-E(X)) + (Y-E(Y))\}^2\right] \geq 0.$$

$$t^2 E[(X-E(X))^2] + 2t\,E[(X-E(X))(Y-E(Y))] + E[(Y-E(Y))^2] \geq 0 \qquad \forall t \in \mathbb{R}$$

$$Var(X)\,t^2 + 2\,Cov(X,Y)\,t + Var(Y) \geq 0$$

$$\Downarrow$$

$$(2\,Cov(X,Y))^2 - 4\,Var(X)\,Var(Y) \leq 0$$

$$\Downarrow$$

K.

$$\frac{4\,Cov^2(X,Y)}{4\,Var(X)\,Var(Y)} \leq 1 \Rightarrow (CORR(X,Y))^2 \leq 1 \Rightarrow -1 \leq CORR(X,Y) \leq 1.$$

(2) $CORR(aX, bY) = \frac{ab}{|ab|} CORR(X,Y)$

(3) $CORR(X,Y) = 1 \iff \exists\, a > 0$ tais que $P(Y = aX + b) = 1$.

(-1) $(a < 0)$

e $b \in \mathbb{R}$

Probabilidade e Inferência Estatística I

Aula 16 - 18/04/2013 Luiz Gustavo Esteves

$$CORR(X,Y) = \frac{COV(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$$

(1) $CORR(X,Y) \in [-1,1]$

(2) $CORR(aX,bY) = \frac{ab}{|ab|} CORR(X,Y)$

(3) $CORR(X,Y) = 1 \Leftrightarrow \exists\, a > 0$ e $b \in \mathbb{R}$ tais que $\mathbb{P}(Y = aX + b) = 1$

Exemplo.

$(X_1, X_2) \sim DIR(a_1, a_2, a_3)$

Vimos que

$$Cov(X_1, X_2) = \frac{a_1 \cdot a_2}{(a_1+a_2+a_0)(a_1+a_2+a_0+1)} - \frac{a_1 a_2}{(a_1+a_2+a_0)^2} =$$

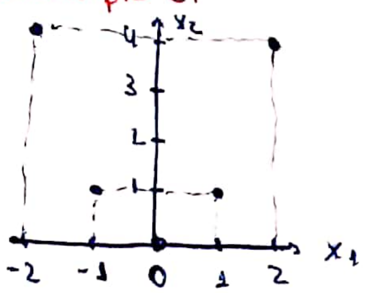
$$= -\frac{a_1 a_2}{(a_0+a_1+a_2)^2 (a_0+a_1+a_2+1)}$$

$$Var(X_1) = \frac{a_1 (a_0 + a_2)}{(a_0+a_1+a_2)^2 (a_0+a_1+a_2+1)}$$

$$CORR(X_1,X_2) = \frac{\dfrac{-a_1a_2}{(a_0+a_1+a_2)^2(a_0+a_1+a_2+1)}}{\sqrt{a_1(a_0+a_2) \cdot a_2(a_0+a_1)}} \Rightarrow$$

$$CORR(X_1,X_2) = \frac{-a_1a_2}{\sqrt{a_1a_2(a_0+a_2)(a_0+a_1)}}$$

Exemplo 2:



$$P(X_2 = X_1^2) = 1$$

$$CORR(X_1,X_2) = \frac{cov(X_1,X_2)}{\sqrt{Var(X_1)\,Var(X_2)}}.$$

Note que $X_1 \sim U(\{-2,-1,0,1,2\})$

$Cov(X_1,X_2) = 0 \Rightarrow$

$CORR(X_1,X_2) = 0$

Momentos Conjuntos

$X = (X_1, \dots, X_k)$

À quantidade $E(X_1^{n_1} X_2^{n_2} \cdots X_k^{n_k})$ damos o nome de

$(n_1, n_2, \dots, n_k)$ - momento conjunto

Exemplo.

$S_k = \{(b_1,\dots,b_k) \in \mathbb{R}_+^* : b_1+b_2+\dots+b_k < 1\}$

$(X_1, X_2, \dots, X_k) \sim DIR(a_1, \dots, a_k, a_0)$

$$E(X_1^{n_1} \dots X_k^{n_k}) = \int_{S_k} x_1^{n_1} \cdots x_k^{n_k} \frac{\Gamma\left(\sum_{i=0}^{k} a_i\right)}{\prod_{i=0}^{k} \Gamma(a_i)} \prod_{i=1}^{k} x_i^{a_i - 1} \cdot (1 - x_1 - \dots - x_k)^{a_0 - 1} dx$$

$$= \Delta^{-1} \int_{S_k} \frac{\Gamma\left(a_0 + \sum_{i=1}^{k}(a_i + n_i)\right)}{\Gamma(a_0) \cdot \prod_{i=1}^{k} \Gamma(a_i + n_i)} \prod_{i=1}^{k} x_i^{a_i + n_i - 1} (1 - x_1 - \dots - x_k)^{a_0 - 1} dx$$

$$\Rightarrow E(X_1^{n_1} \dots X_k^{n_k}) = \frac{\Gamma\left(a_0 + \sum_{i=1}^{k} a_i\right)}{\Gamma\left(a_0 + \sum_{i=1}^{k}(a_i + n_i)\right)} \frac{\prod_{i=1}^{k} \Gamma(a_i + n_i)}{\prod_{i=1}^{k} \Gamma(a_i)}$$

$X : \Omega \to \mathbb{R} \quad \varphi_X(t) = E[e^{tX}]$

$t.X'$

Para $X = (X_1, \dots, X_k) : \Omega \to \mathbb{R}^k$ vetor aleatório, definimos

$\varphi_X : I^{\subseteq \mathbb{R}^k} \to \mathbb{R}$

$(t_1, \dots, t_k) \mapsto \varphi_X(t_1, \dots, t_k) = E[e^{t_1X_1 + t_2X_2 + \dots + t_kX_k}] = E[e^{t.X'}]$ é chamada

FUNÇÃO GERADORA DE MOMENTOS CONJUNTA DE X.

$I = (-\varepsilon, \varepsilon)^k$

Propriedade.

1. $$E(X_1^{n_1} \cdots X_k^{n_k}) = \frac{\partial^{n_1+n_2+\cdots+n_k} \varphi_X(t_1,\dots,t_k)}{\partial t_1^{n_1} \partial t_2^{n_2} \cdots \partial t_k^{n_k}} \Bigg|_{(0,0,\dots,0,0)}$$

Na última vez, calculamos

$E(e^{t_1X_1+\cdots+t_kX_k})$ para $X=(X_1,\dots,X_k) \sim Mult(n, p_1,\dots,p_k)$

$$E(e^{t_1X_1+\cdots+t_kX_k}) = (p_1e^{t_1}+\cdots+p_ke^{t_k}+1-p_1-\cdots-p_k)^n$$

$$\Vert$$

$$\varphi_X(t_1,\dots,t_k)$$

$$Cov(X_i,X_j) = E(X_iX_j) - E(X_i)E(X_j)$$

$$= \frac{\partial^2 \varphi_X(t_1,\dots,t_k)}{\partial t_i \partial t_j}\Bigg|_{(0,\dots,0)} - \left(\frac{\partial \varphi_X(t_1,\dots,t_k)}{\partial t_i}\Bigg|_{(0,0,\dots,0)}\right)\left(\frac{\partial \varphi_X(t_1,\dots,t_k)}{\partial t_j}\Bigg|_{(0,\dots,0)}\right) =$$

$$= n(n-1)p_ip_j - np_i \cdot np_j \Rightarrow Cov(X_i,X_j) = -np_ip_j$$

$$\frac{\partial \varphi_X(t_1,\dots,t_k)}{\partial t_i} = n(p_1e^{t_1}+\cdots+p_ke^{t_k}+1-p_1-\cdots-p_k)^{n-1} p_ie^{t_i}$$

Analogamente,

$$\frac{\partial^2 \varphi_X(t_1,\dots,t_k)}{\partial t_i \partial t_j} = n(n-1)(p_1e^{t_1}+\cdots+p_ke^{t_k}+1-p_1-\cdots-p_k)^{n-2} p_ip_je^{t_i}e^{t_j}$$

Exercício:

$$X = (X_1, X_2) \sim \text{Holgate}(\lambda_1, \lambda_2, \lambda_3)$$

$$E(e^{t_1X_1 + t_2X_2}) = \varphi_X(t_1, t_2).$$

Alguns exemplos adicionais

$X \sim N(0,1)$

$$\varphi_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx =$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-1/2(x^2 - 2tx + t^2)} e^{t^2} \, dx =$$

$$= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-1/2(x-t)^2} \, dx \Rightarrow \varphi_X(t) = e^{t^2/2}.$$

$Y \sim N(0,1)$

$$Y \overset{d}{=} \sigma X + \mu$$

$$\varphi_Y(t) = \varphi_{\sigma X + \mu}(t) = e^{t\mu} \varphi_X(\sigma t) = e^{t\mu} \cdot e^{\frac{\sigma^2 t^2}{2}}.$$

Sejam $X_1, \dots, X_k$ v.a. i.i.d. $N(0,1)$

$a_1, a_2, \dots, a_k \in \mathbb{R}$ e $\mu \in \mathbb{R}$

$$Y = a_1X_1 + a_2X_2 + \dots + a_kX_k + \mu$$

$$\varphi_Y(t) = \varphi_{\sum_{i=1}^{k} a_iX_i + \mu}(t) = e^{t\mu} \varphi_{\sum_{i=1}^{k} a_iX_i}(t) = e^{t\mu} \prod_{i=1}^{k} \varphi_{a_iX_i}(t) = e^{t\mu} \prod_{i=1}^{k} \varphi_{X_i}(a_it) = e^{t\mu} \prod_{i=1}^{k} e^{\frac{a_i^2 t^2}{2}}$$

$$\Rightarrow \varphi_Y(t) = e^{t\mu} \cdot e^{\left(\sum_{i=1}^{k} a_i^2\right) \cdot t^2/2}$$

$$Y \sim N\left(\mu, \sum_{i=1}^{k} a_i^2\right)$$

(4)

$X_1, \dots, X_k$ são independentes com $X_i \sim \text{Gama}(a_i, b)$, $a_1, \dots, a_k, b > 0$.

$$Y = X_1 + \dots + X_k$$

$$\varphi_Y(t) = \varphi_{X_1 + \dots + X_k}(t) = \prod_{i=1}^{k} \varphi_{X_i}(t) = \prod_{i=1}^{k} \left[\frac{b}{b-t}\right]^{a_i} = \left[\frac{b}{b-t}\right]^{\sum_{i=1}^{k} a_i}$$

Logo, $Y \sim \text{Gama}\left(\sum_{i=1}^{k} a_i, b\right)$.

Resultado: $(\Omega, \mathcal{F})$

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

$\mathbb{P}_1, \mathbb{P}_2 : \mathcal{F} \to [0,1]$ probabilidades, $\alpha \in (0,1)$

$$\int_\Omega X \, d(\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) = \alpha \int_\Omega X \, d\mathbb{P}_1 + (1-\alpha) \int_\Omega X \, d\mathbb{P}_2 .$$

• X simples

$$X = \sum_{i=1}^{k} a_i \mathbb{I}_{A_i} , \quad a_1, a_k \geq 0 \quad A_1, A_2, \quad A_k \in \mathcal{F}$$

$$\int_\Omega X \, d(\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) = \sum_{i=1}^{k} a_i \left(\alpha \mathbb{P}_1(A_i) + (1-\alpha)\mathbb{P}_2(A_i)\right) =$$

$$= \alpha \sum_{i=1}^{k} a_i \mathbb{P}_1(A_i) + (1-\alpha) \sum_{i=1}^{k} a_i \mathbb{P}_2(A_i) =$$

$$= \alpha \int_\Omega X \, d\mathbb{P}_1 + (1-\alpha) \int_\Omega X \, d\mathbb{P}_2 .$$

$X \geq 0$

Seja $(X_n)_{n \geq 1}$, $X_n \leq X_{n+1}$, $\forall n \geq 1$, $X_n \geq 0$ simples, $\forall n \geq 1$, tal que $X_n \uparrow X$.

$$\int_\Omega X \, d(\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) = \lim_{n \to \infty} \int_\Omega X_n \, d(\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) =$$

$$= \lim_{n \to \infty} \left[\alpha \int_\Omega X_n \, d\mathbb{P}_1 + (1-\alpha) \int_\Omega X_n \, d\mathbb{P}_2 \right] =$$

$$= \alpha \int_\Omega X \, d\mathbb{P}_1 + (1-\alpha) \int_\Omega X \, d\mathbb{P}_2 .$$

Resultado: $(\Omega, \mathcal{F})$.

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

$\mathbb{P}_1, \mathbb{P}_2 : \mathcal{F} \to [0,1]$ probabilidades, $\alpha \in (0,1)$

$$\int_\Omega X \, d(\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) = \alpha \int_\Omega X \, d\mathbb{P}_1 + (1-\alpha) \int_\Omega X \, d\mathbb{P}_2$$

Suponhamos que exista $A \in \mathcal{B}(\mathbb{R})$, A enumerável, tal que $\mathbb{P}_1(X \in A) = 1$ e que existe

$$\varphi_2 : \mathbb{R} \to \mathbb{R}_+ \text{ tal que } \mathbb{P}_2(X \in B) = \int_B \varphi(t)\, dt ,$$

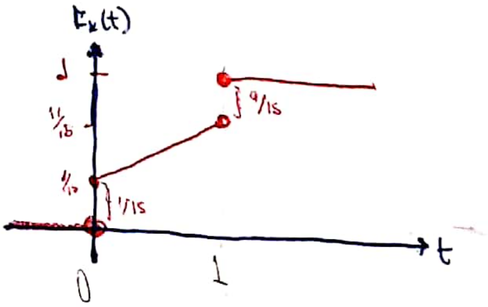
$B \in \mathcal{B}(\mathbb{R})$.

Nesse caso,

$$E(X) = \int_{\Omega} X \, d(\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) = \alpha \sum_{x \in A} x \, \mathbb{P}_1(X=x) + (1-\alpha)\int_{-\infty}^{\infty} x f_2(x)\, dx$$

Exemplo.

$$F_X(t) = \begin{cases} 0, & t<0 \\ \frac{1}{15} + \frac{2}{3}t, & 0 \le t < 1 \\ 1, & t \geq 1 \end{cases}$$



$$\frac{1}{3}\begin{cases} 0, & t<0 \\ 1/5, & 0 \le t < 1 \\ 1, & t \geq 1 \end{cases} + \frac{2}{3}\begin{cases} 0, & t<0 \\ t, & 0 \le t < 1 \\ 1, & t \geq 1 \end{cases}$$

$$F_X(t) = \underset{\overset{\shortparallel}{1/3}}{\alpha} F_d(t) + (1-\underset{\overset{\shortparallel}{2/3}}{\alpha}) F_c(t)$$

$$F_d(t) \to \mathbb{P}_1(X \in (-\infty, t]) \qquad F_c(t) \to \mathbb{P}_2(X \in (-\infty, t])$$

$A = \{0, 1\}$ $\quad \mathbb{P}_1(X \in A) = 1$

$$F_1(t) = \begin{cases} 0, & t<0 \\ 1/5, & 0 \le t < 1 \\ 1, & t \geq 1 \end{cases}$$

$$F_c(t) = \mathbb{P}_2(X \in (-\infty, t]) = \begin{cases} 0, & t<0 \\ t, & 0 \le t \le 1 \\ 1, & t \geq 1 \end{cases}$$

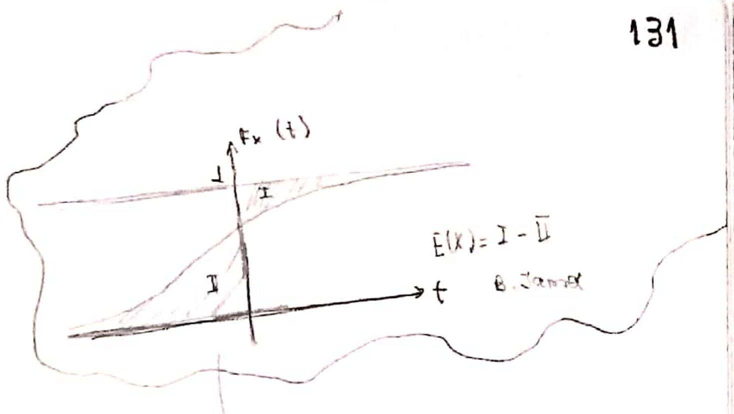
$$f_2(x) = \mathbb{I}_{(0,1)}(x)$$

No exemplo,

$$E(X) = \alpha \sum_{x \in A} x \, \mathbb{P}(X=x) + (1-\alpha)\int_{-\infty}^{\infty} x f_2(x)\, dx =$$

$$= \frac{1}{3}\left[0.\frac{1}{5} + 1.\frac{4}{5}\right] + \frac{2}{3}\int_0^1 x.1\, dx =$$

$$= \frac{1}{3}.\frac{4}{5} + \frac{2}{3}\frac{1}{2} = \frac{9}{15}.$$



Último comentário.

1. Quando $Cov(X,Y) = 0$ (ou $CORR(X,Y) = 0$), dizemos que X e Y são NÃO-CORRELACIONADOS

2. X_1 e X_2 são independentes $\Rightarrow$

$$E(X_1X_2) = E(X_1)E(X_2) \Rightarrow$$

$$\Rightarrow E(X_1X_2) - E(X_1)E(X_2) = 0 \Rightarrow X_1 \text{ e } X_2 \text{ são não-correlacionadas.}$$

No entanto,

X_1 e X_2 não correlacionadas $\not\Rightarrow$ X_1 e X_2 são independentes

$P(X_1 = -1, X_2 = 1) = 1/4$

$P(X_1 = 1, X_2 = 1) = 1/4$

$P(X_1 = 0, X_2 = \pi) = 1/2$

$P(X_1 = -1) = 1/4.$

$P(X_2 = 1) = 1/2$

Claro que

$P(X_1 = -1, X_2 = 1) \neq P(X_1 = -1)P(X_2 = 1)$

Seja $X: \Omega \to \mathbb{R}$ uma v.a.

$$P_{X|B}: \mathcal{B}(\mathbb{R}) \to [0,1]$$

$$A \in \mathcal{B}(\mathbb{R}) \mapsto P_{X|B}(A) = \frac{P(X^{-1}(A) \cap B)}{P(B)} = \frac{P((X \in A) \cap B)}{P(B)}$$

Distribuição condicional de X dado B.

Vamos verificar que $\mathbb{P}_{X|B}$ é, de fato, medida de probabilidade em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

(1) $\mathbb{P}_{X|B}(\mathbb{R}) = \dfrac{\mathbb{P}(X^{-1}(\mathbb{R}) \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = 1$

(2) $(A_n)_{n \geq 1}$, tais que $A_n \in \mathcal{B}(\mathbb{R})$, $\forall n \geq 1$, e $A_i \cap A_j = \emptyset$, $i \neq j$.

$$\mathbb{P}_{X|B}\left(\bigcup_{n=1}^{\infty} A_n\right) = \frac{\mathbb{P}\left(X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) \cap B\right)}{\mathbb{P}(B)} = \frac{\mathbb{P}\left(\left(\bigcup_{n=1}^{\infty} X^{-1}(A_n)\right) \cap B\right)}{\mathbb{P}(B)} =$$

$$= \frac{\mathbb{P}\left(\bigcup_{n=1}^{\infty} (B \cap X^{-1}(A_n))\right)}{\mathbb{P}(B)} = \sum_{n=1}^{\infty} \frac{\mathbb{P}(X^{-1}(A_n) \cap B)}{\mathbb{P}(B)} = \sum_{n=0}^{\infty} \mathbb{P}_{X|B}(A_n).$$

De (1) e (2), $\mathbb{P}_{X|B}$ é de fato probabilidade.

Analogamente, podemos

$$F_{X|B} : \mathbb{R} \longrightarrow [0,1]$$
$$t \in \mathbb{R} \longmapsto F_{X|B}((-\infty, t])$$

→ Função de Distribuição Condicional de X dado B.

Na próxima vez...

$Y : \Omega \to \mathbb{R}$

Y discreta $\mathbb{P}(Y \in \{y_1, y_2, \dots\}) = 1$ $\quad P(Y = y_i) > 0$

$B_i = \{Y = y_i\}$

$\mathbb{P}_{X|B_i}$

Probabilidade e Inferência Estatística I

19/04/2013 - Aula 17

Luiz Gustavo Esteves

$(\Omega, \mathcal{F}, \mathbb{P})$

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

$X \mid |X| = t,\ t > 0$

$B \in \mathcal{F},\ \mathbb{P}(B) > 0$

$\mathbb{P}_{X|B} : \mathcal{B}(\mathbb{R}) \to [0,1]$

$$A \in \mathcal{B}(\mathbb{R}) \mapsto \mathbb{P}_{X|B}(A) = \frac{\mathbb{P}(X^{-1}(A) \cap B)}{\mathbb{P}(B)}$$

$$F_{X|B} : \mathbb{R} \to [0,1]$$
$$t \in \mathbb{R} \longmapsto F_{X|B}(t) = \mathbb{P}_{X|B}((-\infty, t])$$

$Y : \Omega \to \mathbb{R}$ é outra v.a. discreta, tal que $\mathbb{P}(Y \in \{y_1, y_2, \ldots\}) = 1$ e

$\mathbb{P}(Y_i = y_i) > 0,\ i = 1, 2, \ldots$

$B_i = \{Y_i = y_i\}$

$\mathbb{P}_{X|B_i}$: dist. condicional de X dado $Y = y_i$.

$F_{X|B_i}$: função de distribuição condicional de X dado $Y = y_i$.

$F_{X|Y=y_i}$.

Se $X : \Omega \to \mathbb{R}$ é também discreta, (X, Y) é discreto.

Logo, $\mathbb{P}_{X|Y=y_i}$ é tal que $\exists\, A \in \mathcal{B}(\mathbb{R})$, A enumerável, com $\mathbb{P}_{X|Y=y_i}(A) = 1$.

mark@usp.br

Para cada possível valor x_i de X, podemos avaliar

$$P_{X/Y=y_j}(\{x_i\}) = \frac{P(X^{-1}(\{x_i\}) \cap (Y=y_j))}{P(Y=y_j)} \Rightarrow$$

$$\Rightarrow P_{X/Y=y_j}(\{x_i\}) = \frac{P(X=x_i, Y=y_j)}{P(Y=y_j)}$$

Exemplo 1:

$$P(X=x, Y=y) = \frac{1}{2y+1}\binom{10}{y}\left(\frac{1}{2}\right)^{10} \mathbb{I}_{\{-y,\dots,0,\dots,y\}}(x) \quad \mathbb{I}_{\{0,\dots,10\}}(y)$$

$y \in \{0,\dots,10\}$ $\quad \mathbb{I}_{\{\mathbb{N},\dots,10\}}(y) \quad \mathbb{I}_{\{0,\dots,10\}}(x)$

$$P(X=x/Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{\frac{1}{2y+1}\binom{10}{y}\left(\frac{1}{2}\right)^{10} \mathbb{I}_{\{-y,\dots,0,\dots,y\}}(x)\ \mathbb{I}_{\{0,\dots,10\}}(y)}{P(Y=y)}$$

Mas

$$P(Y=y) = \sum_x P(X=x, Y=y) = \sum_{x=-y}^{y}\left(\frac{1}{2y+1}\right)\binom{10}{y}\left(\frac{1}{2}\right)^{10} \mathbb{I}_{\{0,\dots,10\}}(y) =$$

$$= \binom{10}{y}\left(\frac{1}{2}\right)^{10} \mathbb{I}_{\{0,\dots,10\}}(y) \quad \therefore Y \sim \text{Binomial}(10, 1/2).$$

Daí,

$$P(X=x/Y=y) = \frac{\frac{1}{2y+1}\cancel{\binom{10}{y}}\cancel{\left(\frac{1}{2}\right)^{10}} \mathbb{I}_{\{-y,\dots,0,\dots,y\}}(x)\ \mathbb{I}_{\{0,\dots,10\}}(y)}{\cancel{\binom{10}{y}}\cancel{\left(\frac{1}{2}\right)^{10}}\cancel{\mathbb{I}_{\{0,\dots,10\}}(y)}} = \frac{1}{2y+1}\mathbb{I}_{\{-y,\dots,0,\dots,y\}}(x)$$

$\therefore X/Y=y \sim \text{Unif}(\{-y,\dots,0,\dots,y\})$

x	$P(X=x/Y=y)$
$-y$	$1/(2y+1)$
$\vdots$	$\vdots$
-1	
0	$1/(2y+1)$
1	
$\vdots$	$\vdots$
y	$1/(2y+1)$

Daí,

$$E(X/Y=y) = \sum_x x\, P(X=x/Y=y) = 0$$

$$Var(X/Y=y) = \sum_x (x - \underbrace{E(X/Y=y)}_{=0})^2\, P(X=x/Y=y) =$$

$$= \sum_{x=-y}^{y} x^2 \frac{1}{2y+1} = \frac{1}{2y+1} 2 \sum_{i=1}^{y} i^2 =$$

$$= \frac{2}{2y+1} \frac{y(y+1)(2y+1)}{6} \Rightarrow.$$

$$Var(X/Y=y) = \frac{y(y+1)}{3}.$$

Exemplo 2:

$$P(X=x, Y=y) = \frac{e^{-\lambda_1} \lambda_1^x}{x!} \cdot \frac{e^{-\lambda_2} \lambda_2^{y-x}}{(y-x)!} \mathbb{I}_{\{0,1,\dots,y\}}(x) \quad \mathbb{I}_{\mathbb{N}}(y)$$

$$P(Y=y) = \sum_{x=0}^{y} \frac{e^{-\lambda_1} \lambda_1^x}{x!} \frac{e^{-\lambda_2} \lambda_2^{y-x}}{(y-x)!} \mathbb{I}_{\mathbb{N}}(y) =$$

$$= e^{-(\lambda_1+\lambda_2)} \mathbb{I}_{\mathbb{N}}(y) \frac{1}{y!} \sum_{x=0}^{y} \frac{y!\, \lambda_1^x \lambda_2^{y-x}}{x!(y-x)!} \Rightarrow$$

$$\Rightarrow P(Y=y) = \frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1+\lambda_2)^y}{y!} \mathbb{I}_{\mathbb{N}}(y)$$

$\therefore Y \sim Poi(\lambda_1+\lambda_2)$

Daí,

$y \in \mathbb{N}$

$$\mathbb{P}(X=x \mid Y=y) = \frac{\dfrac{e^{-\lambda_1}\lambda_1^{x}}{x!}\;\dfrac{e^{-\lambda_2}\lambda_2^{y-x}}{(y-x)!}\;\mathbb{I}_{\{0,\dots,y\}}(x)\;\mathbb{I}_{\mathbb{N}}(y)}{\dfrac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^{y}}{y!}\;\mathbb{I}_{\mathbb{N}}(y)} \Rightarrow$$

$$\mathbb{P}(X=x \mid Y=y) = \binom{y}{x}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{x}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{y-x}\mathbb{I}_{\{0,1,\dots,y\}}(x) \quad \therefore\; X \mid Y=y \sim Bin\left(y, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$$

$$E(X \mid Y=y_i) = y \cdot \frac{\lambda_1}{\lambda_1+\lambda_2}.$$

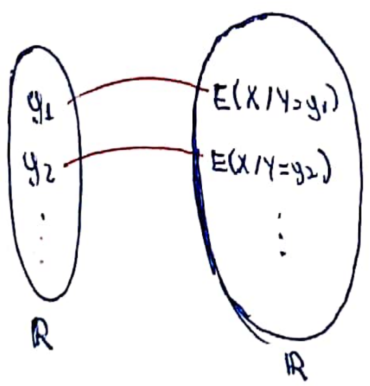
$$Var(X \mid Y=y_i) = \sum_{x}\left(x - E(X \mid Y=y_i)\right)^2 \mathbb{P}(X=x \mid Y=y) = y\,\frac{\lambda_1}{\lambda_1+\lambda_2}\cdot\frac{\lambda_2}{\lambda_1+\lambda_2}.$$

$y \rightarrow \mathbb{P}_{X \mid Y=y_i}$

$\rightarrow E(X \mid Y=y) \rightarrow$ Esperança de X dado $Y=y$ $(\in \mathbb{R})$

$\rightarrow Var(X \mid Y=y)$

Note que para cada $y \in \mathbb{R}$ tal que $\mathbb{P}(Y=y) > 0$, posso associar o número $E(X \mid Y=y_i)$.



$g: \mathbb{R} \rightarrow \mathbb{R}$

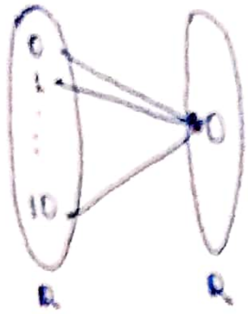
$y \in \mathbb{R} \rightarrow g(y) = E(X \mid Y=y)$

$g(Y) = E(X \mid Y)$

Esperança condicional de X dado Y.

No exemplo 1.

t	$P(E(X/Y)=t)$
0	1



No exemplo 2.

t	$P(E(X/Y)=t)$
0	$e^{-(\lambda_1+\lambda_2)}$
$\frac{\lambda_1}{\lambda_1+\lambda_2}$	$e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)$
$\frac{2\lambda_1}{\lambda_1+\lambda_2}$	$\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^2}{2!}$
$\vdots$	$\vdots$
$\frac{i\lambda_1}{\lambda_1+\lambda_2}$	$\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^i}{i}$

$$E(X/Y=y) = y\,\frac{\lambda_1}{\lambda_1+\lambda_2}$$

$$E(X/Y) = \frac{\lambda_1}{\lambda_1+\lambda_2}\,Y$$

$$P\left(E(X/Y)= i\lambda_1/(\lambda_1+\lambda_2)\right) = P\left(\frac{\lambda_1}{\lambda_1+\lambda_2}Y = \frac{i\lambda_1}{\lambda_1+\lambda_2}\right) =$$

$$= P(Y=i) = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^i}{i!}$$

$h: \mathbb{R} \longrightarrow \mathbb{R}$

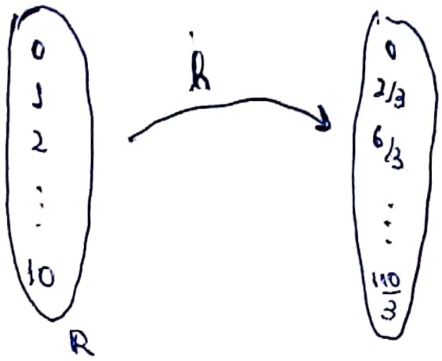
$y \in \mathbb{R} \longmapsto h(y) = Var(X/Y=y)$ → Variância condicional de X dado $Y=y$

$h(Y) = VAR(X/Y)$ → Variância Condicional de X dado Y

No exemplo 1

$$Var(X/Y=y) = \frac{y(y+1)}{3}$$

w	$P(Var(X/Y=y)=w)$
0	$\binom{10}{0}(1/2)^{10}$
$2/3$	$\binom{10}{1}(1/2)^{10}$
$6/3$	$\binom{10}{2}(1/2)^{10}$
$\vdots$	$\vdots$
$110/3$	$\binom{10}{10}(1/2)^{10}$



No exemplo 1:

$$E[E(X/Y)] = 0 \cdot P(E(X/Y)=0) = 0 \cdot 1 = 0.$$

$$Var[E(X/Y)] = 0.$$

$$E[\overbrace{E(X/Y)}^{g(Y)}] = \sum_y E(X/Y=y)\,P(Y=y) =$$

$$= \sum_y \left\{ \sum_x x\,P(X=x/Y=y)\,P(Y=y) \right\} =$$

$$= \sum_y \left\{ \sum_x \overbrace{x}^{f(x,y)=x} P(X=x, Y=y) \right\} = \sum_x \left\{ \sum_y x\,P(X=x, Y=y) \right\} =$$

$$= \sum_x x\,P(X=x) = E(X)$$

Quando

$$E(X) < \infty,$$

$$E(E(X/Y)) = E(X).$$

$$\begin{tikzcd}
(\Omega, \mathcal{F}, \mathbb{P}) \arrow[r, "Y"] \arrow[rr, bend right, "g(Y) = E(X/Y)"'] & (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_Y) \arrow[r, "g"] & (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}^*)
\end{tikzcd}$$

$B \in \mathcal{B}(\mathbb{R})$

$$(g(Y))^{-1}(B) = Y^{-1}(g^{-1}(B)) \in \sigma(Y)$$

X é v.a. $\Rightarrow$ X é $\mathcal{F}$-mensurável
$(X^{-1}(B) \in \mathcal{F})$

$E(X/Y)$ é v.a. $\Rightarrow$ $E(X/Y)$ é $\mathcal{F}$-mensurável
mas é também $\sigma(Y)$-mensurável.

Exemplo.

C_0: capital inicial

X_n: capital após o n-ésimo lançamento da moeda

$E(X_n)$

$$E(X_n / X_{n-1} = c) = (2c) \cdot p + (c/2)(1-p) \Rightarrow$$

$$\Rightarrow E(X_n / X_{n-1} = c) = \frac{c}{2} + p\left\{2c - \frac{c}{2}\right\} = c\left\{\frac{1}{2} + \frac{3}{2}p\right\} = c\left(\frac{3p+1}{2}\right)$$

$$E(X_n / X_{n-1}) = \underbrace{X_{n-1}\left(\frac{3p+1}{2}\right)}_{g(X_{n-1})}$$

Como se trata de uma v.a. discreta,

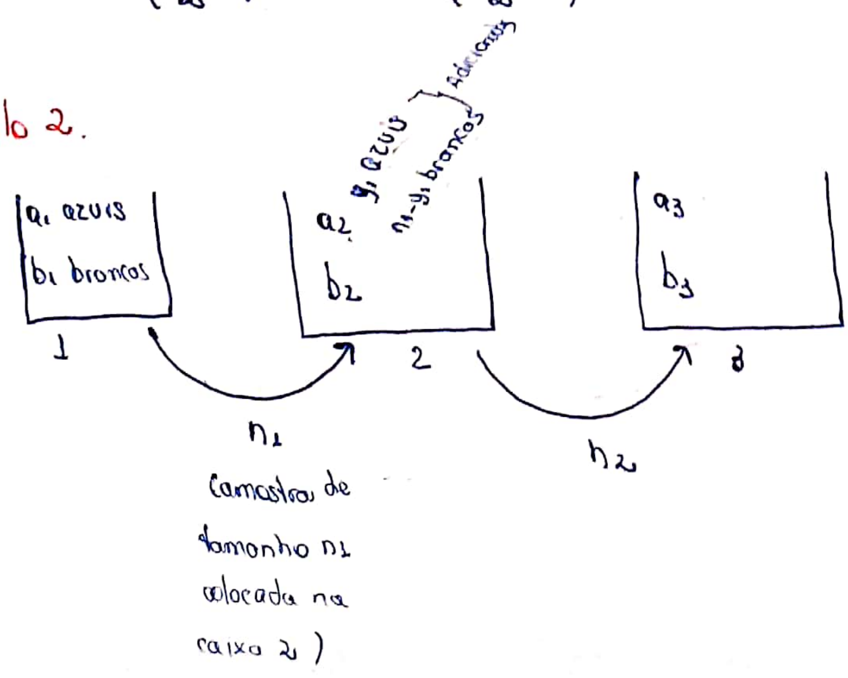
$$E(X_n) = E[E(X_n / X_{n-1})] \Rightarrow$$

$$E(X_n) = E\left(\frac{3p+1}{2} X_{n-1}\right) \Rightarrow$$

$$\Rightarrow \overbrace{E(X_n)}^{e_n} = \frac{3p+1}{2} \overbrace{E(X_{n-1})}^{e_{n-1}}$$

$$\Rightarrow E(X_n) = \left(\frac{3p+1}{2}\right)^n E(X_0) = \left(\frac{3p+1}{2}\right)^n C_0$$

Exemplo 2.



X: Número de bolas azuis em 3 depois das duas extrações

Y_i: Nº de azuis retiradas da urna i e colocadas na urna $i+1$, $i = 1, 2$.

$X = Y_2 + a_3$

$E(X) = E(Y_2) + a_3$

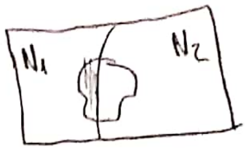
$Var(X) = Var(Y_2)$

Ex.

$X \sim HG(N_1, N_2, n)$

$E(X) = n \dfrac{N_1}{N_1+N_2}$

$Var(X) =$



$X = Y_1 + \dots + Y_{N_1}$

$Y_1 \sim HG(a_1, b_1, n_1)$

$Y_2 / Y_1 = y_1 \sim HG(a_2 + y_1, b_2 + n_1 - y_1, n_2)$

$Y_1 \sim HG$

$$E(Y_2 / Y_1 = y_1) = n_2 \frac{a_2 + y_1}{a_2 + b_2 + n_1}$$

$$E(Y_2) = E(E(Y_2 / Y_1)) =$$

$$= E\left(n_2 \cdot \frac{a_2 + Y_1}{a_2 + b_2 + n_1}\right) = \frac{n_2 a_2}{a_2 + b_2 + n_1} + \frac{n_2}{a_2 + b_2 + n_1} E(Y_1) \Rightarrow$$

$$\Rightarrow E(Y_2) = \frac{n_2 a_2}{a_2 + b_2 + n_1} + \frac{n_2}{a_2 + b_2 + n_1} n_1 \frac{a_1}{a_1 + b_1} + a_3$$

$$Var(X) = \sum_x (x - E(X))^2 P(X = x) =$$

$$= \sum_x (x - E(X))^2 \sum_y P(X = x, Y = y)$$

[redacted]

$$= \sum_x \sum_y (x - E(X/Y=y) + E(X/Y=y) - E(X))^2 P(X=x, Y=y)$$

$$\sum_y \sum_x (x - E(X/Y=y) + E(X/Y=y) - E(X))^2 \, P(X=x/Y=y)\, P(Y=y) =$$

$$= \sum_y P(Y=y) \sum_x \{(x - E(X/Y=y))^2 + (E(X/Y=y) - E(X))^2 + 2(x - E(X/Y=y))(E(X/Y=y) - E(X)\} P(X=x/Y=y)$$

$$= \sum_y P(Y=y)\{ Var(X/Y=y) + (E(X/Y=y) - E(X))^2 + 0\} \Rightarrow$$

$$\Rightarrow Var(X) = \sum_y P(Y=y) \left\{ \overbrace{Var(X/Y=y)}^{h(y)} + \underbrace{\left(E(X/Y=y) - E(X)\right)^2}_{(g(y) - E(g(y)))^2} \right\} =$$

$$= E[Var(X/Y)] + Var[E(X/Y)]$$

Voltando ao Exemplo 1 da aula.

$X/Y=y \sim$ Uniforme $(\{-y, \dots, 0, \dots, y\})$

$E(X/Y=y) = 0 \quad \rightarrow E(X/Y) = 0$

$Var(X/Y=y) = \dfrac{y(y+1)}{3} \quad \Rightarrow Var(X/Y) = \dfrac{Y(Y+1)}{3}$

$$Var(X) = E[Var(X/Y)] + Var[E(X/Y)] =$$

$$= E\left[\frac{Y(Y+1)}{3}\right] + \underbrace{Var(0)}_{0} \Rightarrow$$

$$Var(X) = \frac{1}{3} E(Y + Y^2) = \frac{1}{3} E(Y) + \frac{1}{3}\{(E(Y))^2 + Var(Y)\} =$$

$$= \frac{1}{3} \cdot 10 \cdot \frac{1}{2} + \frac{1}{3}\left(10 \cdot \frac{1}{2}\right)^2 + \frac{1}{3}\left(10 \cdot \frac{1}{2} \cdot \frac{1}{2}\right) =$$

$$= \frac{10}{6} + \frac{50}{6} + \frac{5}{6} = \frac{65}{6} \,/\!/$$

Probabilidade e Inferência Estatística I

23/04/2013 - Aula 18

Luis Gustavo Esteves

$X \sim N(0,1)$

$|X| = 3 \begin{cases} X = 3 \\ X = -3 \end{cases}$

$X / |X| = 3$

Na aula passada, (X,Y) discreto

$$F_{X,Y}(x,y) = \sum_{y_j \leq y} \sum_{x_i \leq x} P(X = x_i, Y = y_j) = \sum_{y_j \leq y} \underbrace{\left\{ \sum_{x_i \leq x} P(X = x_i / Y = y_j) \right\}}_{\substack{P(X \leq x / Y = y_j) \\ \| \\ F_{X/Y=y_j}(x)}} P(Y = y_j)$$

"A soma acima é válida para todo caso, no entanto precisamos de um resultado (teorema) da teoria de medida.

No fundo, vamos generalizar o caso imediatamente acima"

Teorema

Em geral, se (X,Y) é vetor aleatório, podemos escrever

$\forall A \in \mathcal{B}(\mathbb{R}), \forall B \in \mathcal{B}(\mathbb{R}),$

distribuição condicional de X dado Y=t.

$$P(X \in A, Y \in B) = \int_B h(A,y)\, dP_Y$$

" $P(X \in A / Y = y)$ "

Essa função h é tal que

- fixado $y \in \mathbb{R}$, $h(\cdot, y) : \mathcal{B}(\mathbb{R}) \to [0,1]$ é medida de probabilidade.
- fixado $A \in \mathcal{B}(\mathbb{R})$, $h(A, \cdot)$ é função $\mathcal{B}(\mathbb{R})$-mensurável.

Para $A = (-\infty, x]$ e $B = (-\infty, y]$, $x, y \in \mathbb{R}$, temos

$$\mathbb{P}(X \in A, Y \in B) = F_{X,Y}(x,y) = \int_{(-\infty,y]} \overbrace{h((-\infty,x],t)}^{\text{Função de Distribuição Condicional de } X \text{ dado } Y} d\mathbb{P}_Y$$

em geral, é difícil achar h dessa forma.

CASOS FÁCEIS

(1) (X,Y) discreto

(2) (X,Y) é absolutamente contínuo

$$F_{(X,Y)}(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{y} \int_{-\infty}^{x} \frac{f_{X,Y}(u,v)}{f_Y(v)}\, du\, f_Y(v)\, dv =$$

função densidade de probabilidade condicional de X dado $Y=v$: $f_{X|Y=v}(u)$.

$X|Y=v$ é abs. contínua

$$= \int_{-\infty}^{y} \underbrace{\left[\int_{-\infty}^{x} \frac{f_{X,Y}(u,v)}{f_Y(v)}\, du \right]}_{\text{Função De Distribuição Condicional de } X \text{ Dado } Y=v \,:\, F_{X|Y=v}} f_Y(v)\, dv$$

Daí,

$$E(X/Y=v) = \int_{-\infty}^{\infty} x \cdot f_{X/Y=v}(x)\, dx$$

$$E[g(X)/Y=v) = \int_{-\infty}^{\infty} g(x)\, f_{X/Y=v}(x)\, dx$$

Exemplos:

(i) $f_{X,Y}(x,y) = \frac{1}{y} e^{-y}\, \mathbb{I}_{(0,y)}(x)\, \mathbb{I}_{\mathbb{R}_+}(y)$. $X/Y=y \sim$?

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx = \int_{-\infty}^{\infty} \frac{1}{y} e^{-y}\, \mathbb{I}_{(0,y)}(x)\, \mathbb{I}_{\mathbb{R}_+}(y)\, dx =$$

$$= \frac{1}{y} e^{-y}\, \mathbb{I}_{\mathbb{R}_+}(y) \int_0^y 1\, dx = \frac{1}{y} e^{-y}\, \mathbb{I}_{\mathbb{R}_+}(y) \cdot y \Rightarrow f_Y(y) = e^{-y}\, \mathbb{I}_{\mathbb{R}_+}(y)$$

$Y \sim Exp(1)$.

$$f_{X/Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1/y\, e^{-y}\, \mathbb{I}_{(0,y)}(x)}{e^{-y}} = \frac{1}{y}\, \mathbb{I}_{(0,y)}(x)$$

$X/Y=y \sim U(0,y)$

$$E(X/Y=y) = \frac{y}{2}$$

$$Var(X/Y=y) = \frac{y^2}{12}$$

Agora,

$E(X/Y) = Y/2$. Daí $E(E(X/Y)) = E(Y/2) = \frac{1}{2} E(Y) = \frac{1}{2}$.

Exercício: Calcular $E(X)$ e verificar que $E(X)=1/2$.

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{X,Y}(x,y)\,dx\,dy=\int_0^{\infty}\int_0^{y} x\frac{1}{y}e^{-y}\,dx\,dy=\int_0^{\infty}\frac{1}{y}e^{-y}\cdot\frac{y^2}{2}\,dy=\int_0^{\infty}\frac{ye^{-y}}{2}\,dy$$

$$=\frac{1}{2}\underbrace{\int_0^{\infty} ye^{-y}\,dy}_{1}=\frac{1}{2}$$

Note que $E(E(X/Y))=E(X)$.

Exemplo 2.

$$f_{X,Y}(x,y)=\frac{1}{2\pi}|y|\,e^{-\frac{y^2}{2}(1+x^2)}$$

$$f_Y(y)=\int_{-\infty}^{\infty}\frac{1}{2\pi}|y|\,e^{-\frac{y^2}{2}(1+x^2)}\,dx=\frac{|y|}{2\pi}e^{-y^2/2}\int_{-\infty}^{\infty}e^{-\frac{y^2x^2}{2}}\,dx=$$

$$=\frac{|y|}{2\pi}e^{-y^2/2}\sqrt{2\pi/y^2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi/y^2}}e^{-\frac{x^2}{2\cdot\frac{1}{y^2}}}\,dx \Rightarrow f_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-y^2/2}$$

$\therefore Y \sim N(0,1)$

$$f_{X|Y=y}(x)=\frac{\frac{1}{2\pi}\cdot|y|\cdot e^{-y^2/2}\cdot e^{\frac{x^2y^2}{2}}}{\frac{1}{\sqrt{2\pi}}e^{-y^2/2}}\Rightarrow$$

$$\Rightarrow f_{X|Y=y}(x)=\frac{|y|}{\sqrt{2\pi}}e^{-\frac{x^2y^2}{2}}\Rightarrow f_{X|Y=y}(x)=\frac{1}{\sqrt{2\pi\cdot 1/y^2}}e^{-\frac{x^2}{2\cdot\frac{1}{y^2}}}.$$

$\therefore X/Y=y \sim N(0, 1/y^2),\ y\neq 0.$

$E(X/Y=y)=0$

$Var(X/Y=y)=\frac{1}{y^2}$

$E(X/Y)$ é degenerada $\Rightarrow P(E(X/Y)=0)=1 \Rightarrow E(E(X/Y))=0.$

$$f_X(x)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy=\int_{-\infty}^{\infty}\frac{1}{2\pi}|y|\,e^{-y^2\frac{(x^2+1)}{2}}\,dy=$$

$$=\int_{-\infty}^{0}\frac{1}{2\pi}(-y)\,e^{-y^2\frac{(x^2+1)}{2}}\,dy+\int_{0}^{\infty}y\,\frac{1}{2\pi}\,e^{-y^2\frac{(x^2+1)}{2}}\,dy=$$

$$\frac{1}{2\pi}\,\frac{1}{x^2+1}\,e^{-y^2\frac{(x^2+1)}{2}}\Big|_{-\infty}^{0}+\frac{1}{2\pi}\,\frac{-1}{x^2+1}\,e^{-y^2\frac{(x^2+1)}{2}}\Big|_{0}^{\infty}=$$

$$=\frac{1}{2\pi(x^2+1)}+\frac{1}{2\pi(x^2+1)}\Rightarrow f_X(x)=\frac{1}{\pi(1+x^2)}$$

$X\sim$ Cauchy Padrão

$$E(X^+)=\int_0^{\infty}x\,\frac{1}{\pi(1+x^2)}\,dx=\frac{1}{2\pi}\log(1+x^2)\Big|_0^{\infty}=\infty.$$

$$E(X^-)=\int_{-\infty}^{0}(-x).\frac{1}{\pi(1+x^2)}\,dx=-\frac{1}{2\pi}\log(1+x^2)\Big|_{-\infty}^{0}=\infty$$

$\therefore$ Não existe $E(X)$!!, mas $E(E(X/Y))=0$!

v' i

Exemplo 3.

$$f_{X,Y}(x,y) = y\, e^{-(x+1)y}\, \mathbb{I}_{\mathbb{R}_+}(x)\, \mathbb{I}_{\mathbb{R}_+}(y)$$

$$X \mid Y=y \sim Exp(y)$$

$$E(X \mid Y=y) = \frac{1}{y} \qquad E(X) = \infty$$

$$E(E(X/Y)) = E(1/y) = \infty$$

Outro caso "simples".

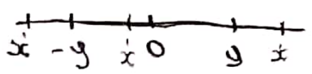
X v.a. absolutamente contínua com densidade f_X. Seja $Y = |X|$.

(X,Y)

$$F_{X,Y}(x,y) = P(X \le x, Y \le y) = P(X \le x, |X| \le y) =$$

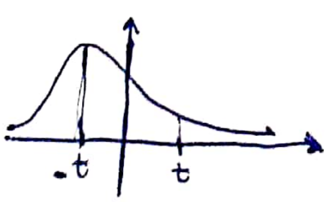
$$= P(X \le x, -y \le X \le y) \Rightarrow$$

$$\Rightarrow F_{X,Y}(x,y) = \begin{cases} 0, & y < 0 \ \checkmark \\ 0, & y \ge 0 \text{ e } x < -y \ \checkmark \\ F_X(x) - F_X(-y), & y > 0,\ -y < x < y \ \checkmark \\ F_X(y) - F_X(-y), & y > 0,\ x > y \ \checkmark \end{cases}$$



$$\overset{Y}{X \mid |X| = t}$$

Candidato à distribuição condicional de $X \mid Y=t$, $t>0$.



$$P(X=t \mid Y=t) = \frac{f_X(t)}{f_X(-t)+f_X(t)}$$

$$P(X=-t \mid Y=t) = \frac{f_X(t)}{f_X(-t)+f_X(t)}$$

Nesse caso,

$$F_{X|Y=t}(x) = \begin{cases} 0, & x < -t \\ \dfrac{f_X(-t)}{f_X(t)+f_X(-t)}, & -t \leq x \leq t \\ 1, & x \geq t \end{cases}$$

Temos que verificar que

$$F_{X,Y}(x,y) = \int_{(-\infty,y]} F_{X|Y=t}(x)\, d\mathbb{P}_Y(t) = \int_{-\infty}^{y} F_{X|Y=t}(x)\, f_Y(t)\, dt$$

(Lembramos que $f_Y(t) = (f_X(t)+f_X(-t))\mathbb{I}_{\mathbb{R}_+}(t)$)

$\underbrace{}_{|X|}$

- $y < 0$

$$\int_{-\infty}^{y} F_{X|Y=t}(x)\, f_Y(t)\, dt = \int_{-\infty}^{y} F_{X|Y=t}(x) \cdot 0\, dt = 0$$

- $y > 0$ e $x < -y$

Se $y > 0$

$$\int_{-\infty}^{y} F_{X|Y=t}(x)\, f_Y(t)\, dt = \int_{0}^{y} F_{X|Y=t}(x)\, (f_X(-t)+f_X(t))\, dt$$

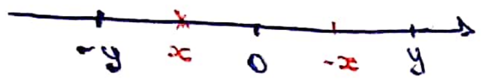
- $y > 0$ e $x < -y$

$0 < t < y \Rightarrow -y < -t < 0 \Rightarrow x < -y < -t < 0 \Rightarrow F_{X|Y=t}(x) = 0.$

$\therefore \int_{0}^{y} F_{X|Y=t}(x)\, (f_X(-t)+f_X(t))\, dt = 0$

• $y>0$, $-y<x<y$

➔ $x<0$



$$t<|x| \Rightarrow F_{X|Y=t}(x)=0$$

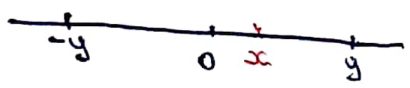
$$t>|x| \Rightarrow F_{X|Y=t}(x)=\frac{f_X(-t)}{f_X(-t)+f_X(t)}$$

$$\int_{-\infty}^{y} F_{X|Y=t}(x)\, f_Y(t)\, dt = \int_0^y F_{X|Y=t}(x) \,.\left(f_X(-t)+f_X(t)\right).dt =$$

$$=\int_0^{|x|} 0.\left(f_X(t)+f_X(-t)\right)dt + \int_{|x|}^{y} \frac{f_X(-t)}{f_X(-t)+f_X(t)} \left(f_X(-t)+f_X(t)\right) dt =$$

$$= F_X(-|x|) - F_X(-y) = F_X(x) - F_X(-y)$$

➔ $x>0$



$$x>t \Rightarrow F_{X|Y=t}(x)=1$$

$$x<t \Rightarrow F_{X|Y=t}(x)=\frac{f_X(-t)}{f_X(-t)+f_X(t)}$$

$$\int_0^y F_{X|Y=t}(x) \left(f_X(t)+f_X(-t)\right) dt =$$

$$=\int_0^x 1.\underbrace{\left(f_X(t)+f_X(-t)\right)}_{f_Y(t)} dt + \overbrace{\int_x^y \frac{f_X(-t)}{f_X(-t)+f_X(t)} \left(f_X(-t)+f_X(t)\right) dt}^{F_X(-x)-F_X(-y)} =$$

$= F_X(x) - F_X(\cancel{x}) + F_X(-x) - F_X(-y) = F_X(x) - F_X(-y)$

- $x > y > 0$

$0 < t < y < x$

$F_{X|Y=t}(x) = 1$

Logo

$$\int_0^y F_{X|Y=t}(x)\,(f_X(t) + f_X(-t))\,dt =$$

$$= \int_0^y 1\,\underbrace{(f_X(t) - f_X(-t))}_{f_Y(y)}\,dt$$

$$= F_X(y) - F_X(-y)$$

$\mathbb{P}(Y \leq y) =$

$\mathbb{P}(|X| \leq y) = \mathbb{P}(-y \leq X \leq y) =$

$= F_X(y) - F_X(-y)$

Esperança Condicional dado uma σ-álgebra.

$(\Omega, \mathcal{F}, \mathbb{P})$

$X : \Omega \mapsto \mathbb{R}$

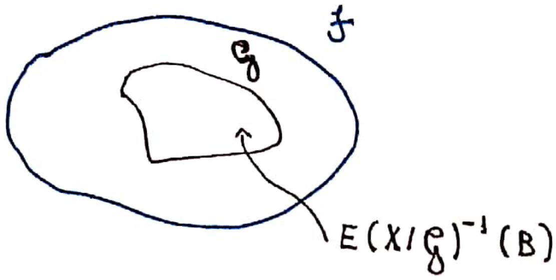
$\mathcal{G} \subseteq \mathcal{F}$ uma σ-álgebra de Ω.

Definição: A [redacted] esperança condicional de X dado $\mathcal{G}$, $E(X|\mathcal{G})$, é uma v.a. tal que

(1) $E(X|\mathcal{G})$ é $\mathcal{G}$-mensurável

(2) $\int_A X\,d\mathbb{P} = \int_A E(X|\mathcal{G})\,d\mathbb{P}$, $\forall A \in \mathcal{G}$

. $E(X|\mathcal{G}) : \Omega \mapsto \mathbb{R}$



Exemplo 1.

$\mathcal{G} = \mathcal{F}_0 = \{\emptyset, \Omega\} \subseteq \mathcal{F}$

$E(X/\mathcal{F}_0)(\omega) = cte, \ \forall \omega \in \Omega$

$X(\omega) = K$

$K^{-1}((-\infty, t)) \in \{\emptyset, \Omega\}$

Nesse caso, tomando $A = \Omega$,

$$\int_\Omega X \, dP = \int_\Omega \underbrace{E(X/\mathcal{F}_0)}_{K} \, dP = K \cdot P(\Omega) = K \Rightarrow K = E(X) < \infty.$$

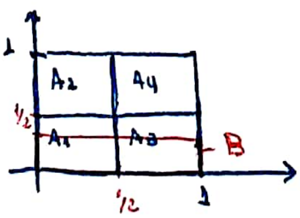
Logo,

$E(X/\mathcal{G}) = E(X)$ com probabilidade 1.

Exemplo 2.

$\Omega = [0,1]^2$

$\mathcal{F} = \mathcal{B}([0,1]^2)$



$X(\omega) = \sum_{i=1}^{4} i \, \mathbb{I}_{A_i}(\omega)$

$B = [0,1] \times [0, 1/3]$

$\mathcal{G} = \sigma(\{B\}) = \{\emptyset, \Omega, B, B^c\}$

$E(X|\mathcal{G})$

$$E(X/\mathcal{G}) = \begin{cases} u, & \omega \in B \\ v, & \omega \notin B \end{cases}.$$

Para achar u e v, usamos

$$\int_B X \, d\mathbb{P} = \int_B E(X/\mathcal{G}) \, d\mathbb{P}$$

$$\int_{B^c} X \, d\mathbb{P} = \int_{B^c} E(X/\mathcal{G}) \, d\mathbb{P}$$

Probabilidade e Inferência Estatística I

26/04/2013 - Aula 19

Luis Gustavo Esteves

Avaliação 1. 09 de maio de 2013. (Quinta-feira)

Lista 06 - Na pasta 13.

$(\Omega, \mathcal{F}, \mathbb{P})$

$\mathcal{G} \subseteq \mathcal{F}$ sub-σ-álgebra de $\mathcal{F}$

$X : \Omega \to \mathbb{R}$ v.a.

A esperança condicional de X dado $\mathcal{G}$, $E(X/\mathcal{G})$, é uma v.a. tal que:

(1). $E(X/\mathcal{G})$ é $\mathcal{G}$-mensurável

(2). $\int_A X \, d\mathbb{P} = \int_A E(X/\mathcal{G}) \, d\mathbb{P}$, $\forall A \in \mathcal{G}$.

(1) $\mathcal{G} = \mathcal{F}_0 = \{\emptyset, \Omega\}$

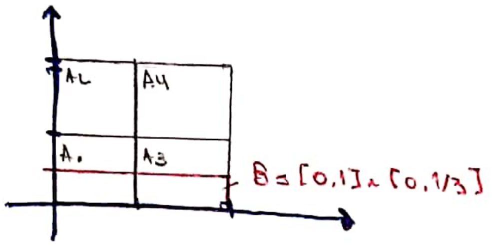
$E(X/\mathcal{G}) = E(X)$ $\mathbb{P}$-q.c.

$(\mathbb{P}(E(X/\mathcal{G}) = E(X)) = 1)$

Comentário: Se h é outra função tal que h atende (1) e (2), então

$$h = E(X/\mathcal{G}) \quad \mathbb{P}\text{- q.c.}$$

(2) $\Omega = [0,1]^2 \quad \mathcal{F} = \mathcal{B}([0,1]^2)$



$$X(\omega) = \sum_{i=1}^{4} i \, \mathbb{I}_{A_i}(\omega) \qquad \mathcal{G} = \{\emptyset, \Omega, B, B^c\}$$

$$E(X/\mathcal{G})(\omega) = \begin{cases} u, & \omega \in B \\ v, & \omega \in B^c \end{cases} \Rightarrow u \, \mathbb{I}_B(\omega) + v \, \mathbb{I}_B(\omega).$$

$$\int_B X \, dP = \int_B E(X/\mathcal{G}) \, dP \Rightarrow \int_\Omega X \, \mathbb{I}_B \, dP = \int_\Omega E(X/\mathcal{G}) \mathbb{I}_B \, dP \Rightarrow$$

$$\int_\Omega \mathbb{I}_B \sum_{i=1}^{4} i \, \mathbb{I}_{A_i} \, dP = \int_\Omega \mathbb{I}_B (u \, \mathbb{I}_B + v \, \mathbb{I}_{B^c}) \, dP \Rightarrow$$

$$\int_\Omega \sum_{i=1}^{4} i \, \mathbb{I}_{A_i \cap B} \, dP = \int_\Omega u \, \mathbb{I}_B \, dP \Rightarrow 1 \cdot \frac{1}{6} + 2 \cdot 0 + 3 \cdot \frac{1}{6} + 4 \cdot 0 = u \cdot \frac{1}{3} \Rightarrow$$

$$\frac{u}{3} = \frac{2}{3} \Rightarrow u = 2$$

e deve satisfazer também

$$\int_{B^c} X \, dP = \int_{B^c} E(X/\mathcal{G}) \, dP \Rightarrow \int_\Omega \sum_{i=1}^{4} i \, \mathbb{I}_{A_i \cap B^c} \, dP = \int_\Omega \mathbb{I}_{B^c} (u \, \mathbb{I}_B + v \, \mathbb{I}_{B^c}) \, dP \Rightarrow$$

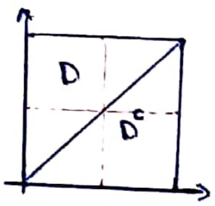
$$\Rightarrow 1 \cdot \frac{1}{12} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{12} + 4 \cdot \frac{1}{4} = v \cdot \frac{2}{3} \Rightarrow \frac{22}{12} = \frac{8v}{12} \Rightarrow v = \frac{11}{4}$$

Logo,

$$E(X/\mathcal{G})(\omega) = \begin{cases} 2, & \omega \in B \\ 11/4, & \omega \in B^c \end{cases}$$

$$E(E(X/\mathcal{G})) = 2 \cdot \frac{1}{3} + \frac{11}{4} \cdot \frac{2}{3} = \frac{30}{12} = \frac{5}{2}$$

Exemplo 3.



$$\mathcal{G}_1 = \{\emptyset, \Omega, D, D^c\}$$

$$E(X/\mathcal{G}_1)(\omega) = \begin{cases} a, & \omega \in D \\ b, & \omega \in D^c \end{cases}$$

$$\int_D X \, dP = \int_D E(X/\mathcal{G}_1) \, dP =$$

$$= \int_\Omega X \mathbb{I}_D \, dP = \int_\Omega \underbrace{E(X/\mathcal{G}_1)}_{a\mathbb{I}_D + b\mathbb{I}_{D^c}} \mathbb{I}_D \, dP \Rightarrow \int_\Omega \sum_{i=1}^{4} i \mathbb{I}_{A_i \cap D} \, dP = \int_\Omega a \, \mathbb{I}_D \, dP \Rightarrow$$

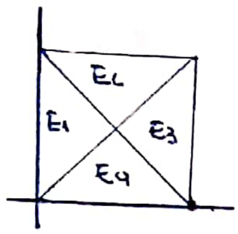
$$\Rightarrow 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{4} + 3 \cdot 0 + 4 \cdot \frac{1}{8} = a \cdot \frac{1}{2} \Rightarrow a = \frac{9}{4}.$$

$$\int_{D^c} X \, dP = \int E(X/\mathcal{G}_1) dP = \int_\Omega \sum_{i=1}^{4} i \, \mathbb{I}_{A_i \cap D^c} \, dP = \int_\Omega b \, \mathbb{I}_{D^c} \, dP \Rightarrow$$

$$\Rightarrow 1 \cdot \frac{1}{8} + 2 \cdot 0 + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} = b \cdot \frac{1}{2} \Rightarrow b = \frac{11}{4}.$$

$$E(X/\mathcal{G}_1)(\omega) = \begin{cases} 9/4, & \omega \in D. \\ 11/4, & \omega \in D^c. \end{cases}$$

Exercício:



$$\mathcal{G}_2 = \sigma(\{E_1, E_2, E_3, E_4\})$$

$$E(X/\mathcal{G}_2)(\omega) = \sum_{i=1}^{4} a_i \mathbb{I}_{E_i}(\omega)$$

$$\int_{E_i} X \, d\mathbb{P} = \int_{E_i} E(X/\mathcal{G}_2) \, d\mathbb{P}$$

$i = 1, 2, 3, 4$

onde $a_1 = 3/2$

$a_2 = 2$

$a_3 = 7/2$

$a_4 = 3$

$E(X/\mathcal{G}_1)$

$$E(\underbrace{E(X/\mathcal{G}_1)}_{X'} / \mathcal{G}_2)$$

$$\int_A E(X/\mathcal{G}_1) \, d\mathbb{P} = \int_A E(E(X/\mathcal{G}_1)/\mathcal{G}_2) \, d\mathbb{P}, \quad \forall A \in \mathcal{G}_2.$$

Em geral, se $A_1, \dots, A_n$ é uma partição de Ω ($A_i \in \mathcal{F}$), e

$$\mathcal{G} = \sigma(\{A_1, A_2, \dots, A_n\})$$

então uma função $\mathcal{G}$-mensurável assume no máximo n valores.

Desse modo

$E(X/\mathcal{G})$ é dada por $E(X/\mathcal{G})(\omega) = \sum_{i=1}^{n} a_i \mathbb{I}_{A_i}$.

onde $a_1, \dots, a_n$ são tais que

$$\int_{A_i} X \, d\mathbb{P} = \int_{A_i} E(X/\mathcal{G}) \, d\mathbb{P} \Rightarrow$$

$$\Rightarrow \int_{A_i} X \, d\mathbb{P} = \int_{\Omega} \mathbb{I}_{A_i} E(X/\mathcal{G}) \, d\mathbb{P} \Rightarrow$$

$$\Rightarrow a_i \, \mathbb{P}(A_i) = \int_{A_i} X \, d\mathbb{P} \Rightarrow$$

$$a_i = \frac{\int_{A_i} X \, d\mathbb{P}}{\mathbb{P}(A_i)}$$

Algumas Propriedades

1. $\mathcal{F}_0 = \{\emptyset, \Omega\}$

$$E(X/\mathcal{F}_0) = E(X) \quad \mathbb{P}\text{-q.c}$$

$$\int_A X \, d\mathbb{P} = \int_A E(X/\mathcal{F}_0) \, d\mathbb{P}$$

- Se $A = \emptyset \Rightarrow 0 = 0$
- Se $A = \Omega \Rightarrow \int X \, d\mathbb{P} = E(X)$

2. $\mathcal{G} = \mathcal{F}$ $\quad (\Omega, \mathcal{F}, \mathbb{P})$

$$E(X/\mathcal{G}) = X. \quad \mathbb{P}\text{-q.c.}$$

$$\int_A E(X/\mathcal{G}) \, d\mathbb{P} = \int_A X \, d\mathbb{P}, \ A \in \mathcal{F}$$

3. $X, Y : \Omega \to \mathbb{R}$

$a, b \in \mathbb{R}$

$$E(aX + bY/\mathcal{G}) = a E(X/\mathcal{G}) + b E(Y/\mathcal{G})$$

Para $A \in \mathcal{G}$,

$$\int_A (aX+bY)\,dP = \int_\Omega (aX+bY)\mathbb{I}_A\,dP = \int_\Omega aX\mathbb{I}_A\,dP + \int_\Omega bY\mathbb{I}_A\,dP =$$

$$= a\int_\Omega X\mathbb{I}_A\,dP + b\int_\Omega Y\mathbb{I}_A\,dP = a\int_A X\,dP + b\int_A Y\,dP =$$

$$= a\int_A E(X/\mathcal{G})\,dP + b\int_A E(Y/\mathcal{G})\,dP = \int_A aE(X/\mathcal{G})\,dP + \int_A bE(Y/\mathcal{G})\,dP =$$

$$= \int_\Omega aE(X/\mathcal{G})\mathbb{I}_A\,dP + \int_\Omega bE(Y/\mathcal{G})\mathbb{I}_A\,dP = \int_A (aE(X/\mathcal{G}) + bE(Y/\mathcal{G}))\,dP$$

Portanto,

$$E(aX+bY/\mathcal{G}) = aE(X/\mathcal{G}) + bE(Y/\mathcal{G}).$$

4. $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$

$$E(E(X/\mathcal{G}_2)/\mathcal{G}_1) = E(X/\mathcal{G}_1)$$

Para $A \in \mathcal{G}_1$,

$$\int_A E(X/\mathcal{G}_1)\,dP = \int_A X\,dP = \int_A E(X/\mathcal{G}_2)\,dP \Rightarrow \forall A \in \mathcal{G}_1,$$

$$\int_A E(\underbrace{X/\mathcal{G}_2}_{Y})\,dP = \int_A E(X/\mathcal{G}_1)\,dP$$

$$\|$$

$$\int_A E(\overbrace{E(X/\mathcal{G}_2)}^{Y}/\mathcal{G}_1)\,dP$$

Como $\int_A E(X/\mathcal{G}_1)\,d\mathbb{P} = \int_A E(X/\mathcal{G}_2)\,d\mathbb{P}$, $\forall A \in \mathcal{G}_1$ e $E(X/\mathcal{G}_1)$ é $\mathcal{G}_1$-mensurável, então

$$\mathbb{E}(E(X/\mathcal{G}_2)\backslash \mathcal{G}_1) = E(X/\mathcal{G}_1) \quad \mathbb{P}\text{-q.c.}$$

Comentário.

Dado $\mathbb{P}(X=x) = \mathbb{P}(\cdot \mid X=x)$ $\qquad f_X \quad f_x(y) =$

$X \qquad Y/X=x$

$f_{X,Y}(x,y) = f_X\, f_x(y).$

$\mathbb{P}(X=x, Y\in\cdot) = \mathbb{P}(X=x)\,\mathbb{P}(\cdot \mid X=x).$

CONVERGÊNCIA QUASE-CERTA

$(\Omega, \mathcal{F}, \mathbb{P})$

$X_1, X_2, X_3, X_4, \dots$

$X_n : \Omega \to \mathbb{R}$ variável aleatória, $n \geq 1$

$(X_n)_{n\geq 1}$: sequência de v.a.

Exemplo: $\Omega = \mathbb{R}$

$\mathcal{F} = \mathcal{B}(\mathbb{R})$

$X_1(\omega) = \omega \qquad (X_n(1))_{n\geq 1} = (1, 2, 3, 4, \dots)$

$X_2(\omega) = 2\omega \qquad (X_n(2))_{n\geq 1} = (2, 4, 6, 8, \dots)$

$(X_n(\omega))$

"Dividir o espaço amostral em dois pacotes, Se o pacote onde a convergência ocorre tem probabilidade 1, então X_n converge q.c. para X. Caso contrário, não"

Definição:

Dizemos que a sequência $(X_n)_{n\geq 1}$ converge quase certamente para $X: \Omega \to \mathbb{R}$ se

$$P(\{\omega \in \Omega : X_n(\omega) \to X(\omega)\}) = 1$$

$$\lim_{n\to\infty} X_n(\omega) = X(\omega)$$

Notação:

$$X_n \xrightarrow{q.c.} X.$$

Exemplo:

$\Omega = [-1,1]$

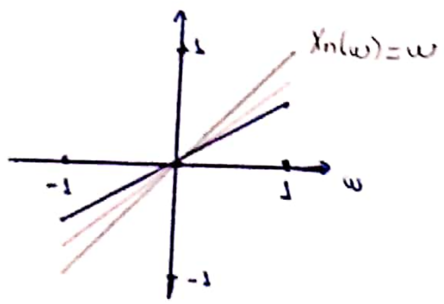
$\mathcal{F} = \mathcal{B}([-1,1])$

$$P(A) = \frac{\text{"Comprimento A"}}{2}$$

$X_n : \Omega \longrightarrow \mathbb{R}$

$\omega \in \Omega \mapsto X_n(\omega) = \frac{n}{n+1}\omega$

$$X_n(\omega) = \frac{n}{n+1} \cdot \omega$$



$\{\omega \in \Omega : X_n(\omega) \text{ converge}\} = [-1,1]$

$\omega \in [-1,1]$

$X_n(\omega) = \frac{n}{n+1}\omega \to \omega$

$\boxed{[-1,1]}$

$P(\{\omega \in \Omega : X_n(\omega) \to X(\omega) = \omega\}$

Nesse caso, dizemos que

$$X_n \xrightarrow{q.c.} X,$$

onde $X(\omega) = \omega$.

Exemplo 2.

$$Y_n(\omega) = \omega^n$$

$\omega = 1$ $\quad Y_n(1) \to 1$

$\omega = 0$ $\quad Y_n(0) \to 0$

$1 > \omega > 0$ $\quad Y_n(\omega) \to 0$

$-1 < \omega < 0$ $\quad Y_n(\omega) \to 0$

$\omega = -1$ $\quad Y_n(-1) = \begin{cases} -1, & n \text{ ímpar} \\ 1, & n \text{ par} \end{cases} \not\to$

não converge

$$\{\omega \in \Omega : Y_n(\omega) \text{ converge}\} = \Omega \setminus \{-1\} = (-1, 1]$$

$$\mathbb{P}(\{\omega \in \Omega : Y_n(\omega) \to Y(\omega)\}) = \mathbb{P}((-1,1]) = 1, \text{ onde}$$

$$Y(\omega) = \begin{cases} 1, & \omega = 1 \\ 0, & \omega \neq 1 \end{cases}$$

Nesse caso, dizemos que

$$Y_n \xrightarrow{q.c.} Y.$$

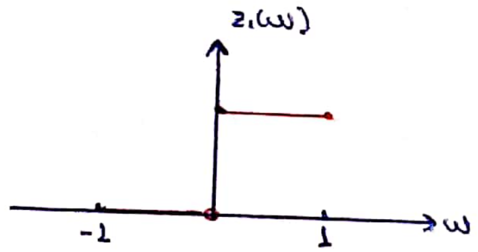
Note que

$$Y_n \xrightarrow{q.c} Y', \text{ onde}$$

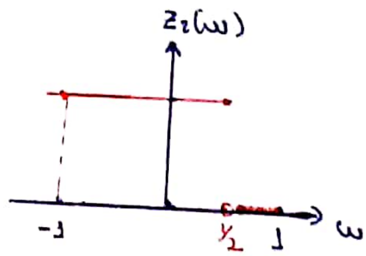
$$Y'(\omega) = \begin{cases} 1, & \omega = 1 \\ 0, & -1 < \omega < 1 \\ \pi^2/\operatorname{arctg} 12, & \omega = 1 \end{cases}$$

Exemplo 3:

$$Z_n(\omega) = \begin{cases} \mathbb{I}_{[-1,1/2]}(\omega), & n \text{ par} \\ \mathbb{I}_{[0,1]}(\omega), & n \text{ ímpar} \end{cases}$$



n ímpar



$\omega < 0$

$$(Z_n(\omega))_{n \geq 1} = (0, 1, 0, 1, 0, \dots) \not\to ?$$

$0 \leq \omega \leq 1/2$

$$(Z_n(\omega))_{n \geq 1} = (1, 1, 1, \dots) \to 1$$

$\omega \in (1/2, 1]$

$$(Z_n(\omega))_{n \geq 1} = (1, 0, 1, 0, \dots) \not\to ?$$

$\{\omega \in \Omega : (Z_n(\omega))_{n\geq 1}$, converge$\} = [0, 1/2]$

$$P(\{\omega \in \Omega : (Z_n(\omega))_{n\geq 1} \text{ é convergente}\}) = P([0,1/2]) = \frac{1/2}{2} = \frac{1}{4} < 1$$

d)

$$W_n(\omega) = \sum_{j=0}^{n} |\omega|^j$$

$$(W_n(1/2)) = (1, 3/2, 7/4, \dots) \to 2$$

$$\omega \in (-1,1) \Rightarrow W_n(\omega) = \sum_{j=0}^{n} |\omega|^j = \sum_{j=0}^{\infty} |\omega|^j = \frac{1}{1-|\omega|}$$

$$\omega \in \{-1,1\} \Rightarrow W_n(\omega) = (1, 2, 3, \dots) \to \infty ?$$

$$\{\omega \in \Omega : (W_n(\omega))_{n\geq 1} \text{ é convergente}\} =$$

$$= (-1,1) = \Omega \setminus \{-1,1\}$$

$$W(\omega) = \begin{cases} 0 & , \omega \in \{-1,1\} \\ \frac{1}{1-|\omega|} & , \omega \in (-1,1) \end{cases}$$

$$P(\{\omega \in \Omega : W_n(\omega) \to W(\omega)\}) =$$

$$= P((-1,1)) = \frac{2}{2} = 1$$

Assim,

$$W_n \xrightarrow{q.c} W.$$

" Dividir Ω em dois pacotes. Um com os pontos onde a convergência ocorre e outro onde a convergência não ocorre. Se a probabilidade do primeiro pacote é igual à 1, temos a convergência quase certa."

$\{\omega \in \Omega : (Z_n(\omega))_{n \geq 1}$, converge $\} = [0, 1/2]$

$$\mathbb{P}(\{\omega \in \Omega : (Z_n(\omega))_{n \geq 1} \text{ é convergente}\}) = \mathbb{P}([0, 1/2]) = \frac{1/2}{2} = \frac{1}{4} < 1$$

d)

$$W_n(\omega) = \sum_{j=0}^{n} |\omega|^j$$

$$(W_n(1/2)) = (1, 3/2, 7/4, \dots) \to 2$$

$$\omega \in (-1,1) \Rightarrow W_n(\omega) = \sum_{j=0}^{n} |\omega|^j = \sum_{j=0}^{\infty} |\omega|^j = \frac{1}{1-|\omega|}$$

$$\omega \in \{-1, 1\} \Rightarrow W_n(\omega) = (1, 2, 3, \dots) \to \infty ?$$

$\{\omega \in \Omega : (W_n(\omega))_{n \geq 1}$ é convergente $\} =$

$= (-1, 1) = \Omega \setminus \{-1, 1\}$

$$W(\omega) = \begin{cases} 0 & , \omega \in \{-1, 1\} \\ \frac{1}{1-|\omega|} & , \omega \in (-1, 1) \end{cases}$$

$\mathbb{P}(\{\omega \in \Omega : W_n(\omega) \to W(\omega)\}) =$

$= \mathbb{P}((-1, 1)) = \frac{2}{2} = 1$

Assim,

$W_n \xrightarrow{q.c} W$.

" Dividir Ω em dois pacotes. Um com os pontos onde a convergência ocorre e outro onde a convergência não ocorre. Se a probabilidade do primeiro pacote é igual à 1, temos a convergência quase certa."

$$V_n(\omega) = \sum_{j=0}^{n} |2\omega|^j$$

$\omega \in (-1/2, 1/2)$, $V_n(\omega) \to \dfrac{1}{1-|2\omega|}$

$\omega \notin (-1/2, 1/2)$, $V_n(\omega) \not\to$?

$\mathbb{P}(\{\omega \in \Omega : (V_n(\omega))_{n \geq 1} \text{ converge}\}) = \mathbb{P}((-1/2, 1/2)) = \dfrac{1}{2} \Rightarrow$ não converge quase certamente.

26/04/13

Aula 20 - Luis Gustavo Esteves

$$\mathbb{P}(\{w \in \Omega : W_n(w) \longrightarrow w(w)\}) = \mathbb{P}((-1,1)) = \frac{2}{2} = 1$$

Assim

$$W_n \xrightarrow{q.c.} W$$

Para Pensar !

$$V_n(w) = \sum_{j=0}^{n} |2w|^j$$

—— \\ ——

26/04

Review

$$X_n \xrightarrow{q.c} X \Longleftrightarrow \mathbb{P}(\{w \in \Omega : X_n(w) \xrightarrow{n \to \infty} X(w)\}) = 1$$

Comentário

(1) No exemplo (2):

$$Y_n \longrightarrow Y, \quad Y(w) = \begin{cases} 1, & w = 1 \\ 0, & w \neq 1 \end{cases}$$

$$Y_n \rightarrow Y'', \quad Y''(w) = 0, \forall w \in \Omega$$

$$X_n(w) = \frac{n}{n+1} w \xrightarrow{q.c} w = X(w) \longrightarrow X \sim U(-1,1)$$

$$Y_n(w) \xrightarrow{q.c} Y \quad \mathbb{P}(Y=0) = 1$$

$$Y_n \xrightarrow{q.c} c$$

$\{w \in \Omega : X_n(w) \longrightarrow X(w)\} = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{w : |X_k(w) - X(w)| < 1/m\}$

$\forall m \in \mathbb{N}^*$

$X_n(w_0) \longrightarrow X(w_0) \iff \forall \varepsilon > 0, \exists n \in \mathbb{N}$ tq

$k \geq n \quad |X_k(w_0) - X(w_0)| < \varepsilon$

$< 1/m$

$\Rightarrow \forall m \geq 1, \ w_0 \in \liminf_{n \to \infty} \{w : |X_n(w) - X(w)| < 1/m\}$

$w_0 \in \bigcup_{n=1}^{\infty} \bigcap_{k=n} \{w : |X_n(w) - X(w)| < 1/m\} \Rightarrow$

$\Rightarrow w_0 \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n} \{w : |X_k(w) - X(w)| < 1/m$

Propriedades

(Fato $(A_n)_{n \geq 1}$ tais que $\mathbb{P}(A_n) = 1, \forall n \geq 1 \Rightarrow \mathbb{P}(\bigcap_{n=1}^{\infty} A_n) = 1$)

(1) $\left.\begin{array}{l} X_n \xrightarrow{q.c} X \\ Y_n \xrightarrow{q.c} Y \end{array}\right\} \Rightarrow X_n \pm Y_n \xrightarrow{q.c} X \pm Y$

Dem: $A = \{w : X_n(w) \longrightarrow X(w)\}$

$B = \{w \in \Omega : Y_n(w) \longrightarrow Y(w)\}$

$w_0 \in A \cap B \Rightarrow \left\{\begin{array}{l} X_n(w_0) \to X(w_0) \\ Y_n(w_0) \to Y(w_0) \end{array}\right. \Rightarrow X_n(w) + Y_n(w_0) \longrightarrow$

$\longrightarrow X(w_0) + Y(w_0) \Rightarrow w_0 \in \{w \in \Omega : \underbrace{(X_n + Y_n)(w)}_{X_n(w) + Y_n(w)} \to (X+Y)(w)\} \Rightarrow$

$\Rightarrow X_n + Y_n \xrightarrow{q.c} X + Y$

/ /

$$(2)\quad \left.\begin{array}{l} X_n \xrightarrow{qc} X \\ X_n \xrightarrow{qc} Y \end{array}\right\} \Rightarrow \mathbb{P}(X=Y)=1$$

$A = \{w : X_n(w) \longrightarrow X(w)\}$

$B = \{w : X_n(w) \longrightarrow Y(w)\}$

$$w_0 \in A\cap B \Rightarrow \left\{\begin{array}{l} X_n(w_0) \longrightarrow X(w_0) \\ X_n(w_0) \longrightarrow Y(w_0) \end{array}\right. \Rightarrow X(w_0) = Y(w_0) \Rightarrow$$

$w_0 \in \{w \in \Omega : X(w) = Y(w)\}$

$\Rightarrow \mathbb{P}(X=Y)=1$

Exe

$$(3)\quad \left.\begin{array}{l} X_n \xrightarrow{qc} X \\ g:\mathbb{R}\rightarrow\mathbb{R} \text{ contínua} \\ g \text{ mensurável} \end{array}\right\} \Rightarrow g(X_n) \xrightarrow{qc} g(X)$$

Exe

$$(4)\quad X_n \xrightarrow{qc} X \Rightarrow \overline{X}_n = \frac{X_1+\dots+X_n}{n} \xrightarrow{qc} X$$

Resultado:

$$X_n \xrightarrow{qc} X$$

$$\Updownarrow$$

$$\mathbb{P}\left(\limsup \{|X_n - X| \geq \varepsilon\}\right) = 0,\ \forall \varepsilon > 0$$

$|X_n - X| > \varepsilon$ i.v

Dem

$$X_n \xrightarrow{q.c.} X \iff \mathbb{P}\left(\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{|X_k - X| < 1/m\}\right) = 1 \iff$$

$$\iff \mathbb{P}\left(\bigcup_{m=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{|X_k - X| \geq 1/m\}\right) = 0 \iff$$

$$\iff \forall m \geq 1,\ \mathbb{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{|X_k - X| \geq 1/m\}\right) = 0$$

Resultado

$(\Omega, \mathcal{F}, \mathbb{P})$ $(A_n)_{n\geq 1}$ seq. de eventos de $\mathcal{F}$

(1) $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \Rightarrow \mathbb{P}(\limsup A_n) = 0$

(2) $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ e ~~os~~ A_n's são independentes

$$\Downarrow$$

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 1$$

Dem:

$$(1)\ \mathbb{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k\right) \leq \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right)$$

$$\leq \sum_{k=n}^{\infty} \mathbb{P}(A_k)$$

$\forall m \geq 1, \exists\ n = n(m)$ tq $n \geq n(\omega) \Rightarrow \sum_{k=n}^{\infty} \mathbb{P}(A_k) < 1/m$

Logo, $\mathbb{P}(\limsup A_n) < 1/m$, $\forall m \geq 1$

$$\Downarrow$$

$$\mathbb{P}(\limsup A_n) = 0$$

(2) $\mathbb{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k\right)$: Basta verificar $\mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right) = 1, \forall n \geq 1$

$\forall m \in \mathbb{N}^*$, $\boxed{1-t \leq e^{-t}}$

$$\mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right) \geq \mathbb{P}\left(\bigcup_{k=n}^{n+m} A_k\right) \Rightarrow \mathbb{P}\left(\left(\bigcup_{k=n}^{\infty} A_k\right)^c\right) \leq \mathbb{P}\left(\bigcap_{k=n}^{n+m} A_k^c\right)$$

$$= \prod_{k=n}^{n+m} \mathbb{P}(A_k^c) = \prod_{k=n}^{n+m} (1-\mathbb{P}(A_k)) \leq \prod_{k=n}^{n+m} e^{-\mathbb{P}(A_k)} = e^{-\sum_{k=n}^{n+m} \mathbb{P}(A_k)}$$

$$\forall m \geq 1, \mathbb{P}\left(\left(\bigcup_{k=n}^{\infty} A_k\right)^c\right) \leq e^{-\sum_{k=n}^{n+m} \mathbb{P}(A_k)} - a_m \Rightarrow$$

$$\Rightarrow \mathbb{P}\left(\left(\bigcup_{k=n}^{\infty} A_k\right)^c\right) = 0 \Rightarrow \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right) = 1$$

!!

Ex:

$(X_n)_{n\geq 1}$ v.a independentes tais que $X_n \sim U(0, n^2)$

Considere $Y_n = \mathbb{I}_{(0,1)}(X_n) \xrightarrow{q.c} 0$?

$$\mathbb{P}(|Y_n - 0| > \varepsilon) = \begin{cases} 0 & \varepsilon > 1 \\ \mathbb{P}(Y_n = 1), & \varepsilon < 1 \end{cases} \leq \mathbb{P}(Y_n = 1) = \mathbb{P}(X_n \in (0,1)) = \frac{1}{n^2}$$

Logo a série $\sum_{n=1}^{\infty} \underbrace{\mathbb{P}(|Y_n - 0| > \varepsilon)}_{A_n} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty \Rightarrow$

$\overset{(1)}{\Rightarrow} \mathbb{P}\left(\limsup_{n\to\infty} \{|Y_n - 0| > \varepsilon\}\right) = 0 \Rightarrow Y_n \xrightarrow{q.c} 0$ (critério)

Ex 2

$(X_n)_{n\geq 1}$ v.a ind. tais que $X_n \sim U(0,n)$

$$Y_n = \mathbb{I}_{(0,1)}(X_n) \xrightarrow{q.c} ?$$

- $\mathbb{P}(Y_n = 0) = \mathbb{P}(X_n \in (1,n)) = \frac{n-1}{n} \Rightarrow$

$\sum_1^\infty \mathbb{P}(Y_n = 0) = \sum^\infty \frac{n-1}{n} = \infty$. Como os eventos $\{Y_n = 0\}$, $n \geq 1$, são independentes então

$\mathbb{P}(\limsup_{n\to\infty} \{Y_n = 0\}) = 1$ (B-C(2))

- $\mathbb{P}(Y_n = 1) = \mathbb{P}(X_n \in (0,1)) = \frac{1}{n} \Rightarrow \sum_{n=1}^\infty \mathbb{P}(Y_n = 1) = \sum_{n=1}^\infty \frac{1}{n} = \infty$

Pelo B-C(2), $\mathbb{P}(\limsup_{n\to\infty} \{Y_n = 1\}) = 1$.

$$\Rightarrow \overbrace{\limsup_{n\to\infty} \{Y_n = 0\}}^{A} \cap \overbrace{\limsup_{n\to\infty} \{Y_n = 1\}}^{B}$$

$w_0 \in A \cap B \Rightarrow \begin{cases} Y_n(w_0) = 0, \text{ p/ inf. valores } n \\ Y_n(w_0) = 1, \text{ p/ inf. valores de } n \end{cases} \Rightarrow$

$\Rightarrow Y_n(w_0)$ não converge $\Rightarrow w_0 \in \{w \in \Omega : (Y_n(w))_{n\geq 1} \text{ ñ converge}\}$

$\therefore Y_n \not\xrightarrow{q.c}$

/ /

$(\Omega, \mathcal{F}, \mathbb{P})$

$X_1, X_2, \ldots : \Omega \longrightarrow \mathbb{R}$

$X : \Omega \longrightarrow \mathbb{R}$

Dizemos que $(X_n)_{n\geq 1}$ converge em prob. p/ X, $X_n \xrightarrow{\mathbb{P}} X$

se $\forall \varepsilon > 0$, $(\forall m \in \mathbb{N}^*)$

$$\lim_{n\to\infty} \mathbb{P}(\{w \in \Omega : |X_n(w) - X(w)| > \varepsilon) = 0, \ \forall \varepsilon > 0$$

$> 1/m$

Exemplos:

$\Omega = [-1, 1]$ $\qquad \mathbb{P} = \dfrac{\text{comp.}}{2}$

$\mathcal{F} = \mathcal{B}([-1,1])$

(1) $X_n(w) = \dfrac{n}{n+1} w \qquad X(w) = w$

$$\mathbb{P}(\{w : |X_n(w) - X(w)| > \varepsilon\}) = \mathbb{P}\left(\left\{w \in \Omega : \left|\frac{nw}{n+1} - w\right| > \varepsilon\right\}\right)$$

$$= \mathbb{P}\left(\left\{w \in \Omega : \left|\frac{w}{n+1}\right| > \varepsilon\right\}\right) = \mathbb{P}(\{w \in \Omega : |w| > \varepsilon(n+1)\})$$

$$\overset{\text{"}n > 2/\varepsilon\text{"}}{=} \mathbb{P}(\phi) = 0 \ . \ \forall n \geq n^* \simeq \frac{2}{\varepsilon}$$

Logo

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0, \ \forall \varepsilon > 0$$

$$\therefore X_n \xrightarrow{\mathbb{P}} X$$

(2) $Y_n(w) = w^n$

$Y_n \xrightarrow{q.c} Y$

$\varepsilon > 0$ $\quad Y(w) = 0, \forall w \in \Omega \qquad (Y_n \xrightarrow{qc} 0$

$$\mathbb{P}(\{w \in \Omega : |Y_n(w) - 0| > \varepsilon\}) = \mathbb{P}(\{w \in \Omega : \underbrace{|w^n|}_{|w|^n} > \varepsilon\}) =$$

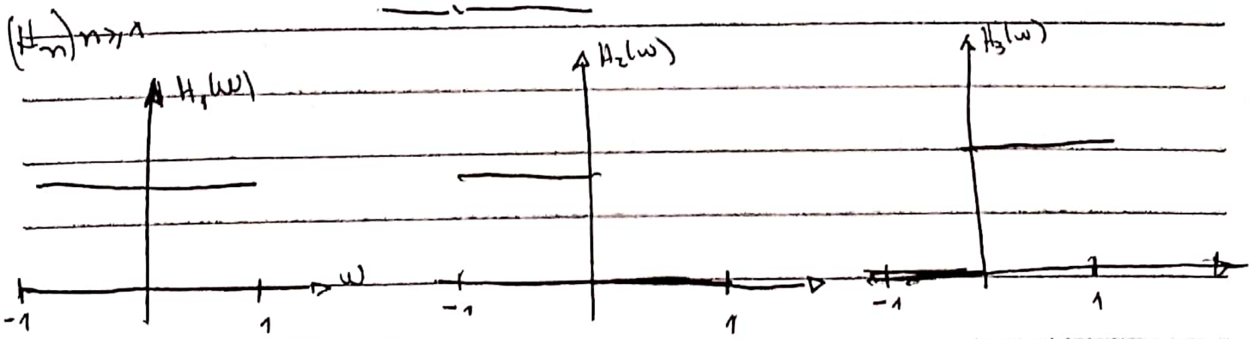
$$= \mathbb{P}(\{w \in \Omega : |w| > \varepsilon^{1/n}\}) = \mathbb{P}(\{w \in \Omega : w < -\varepsilon^{1/n} \text{ ou } w > \varepsilon^{1/n}\})$$

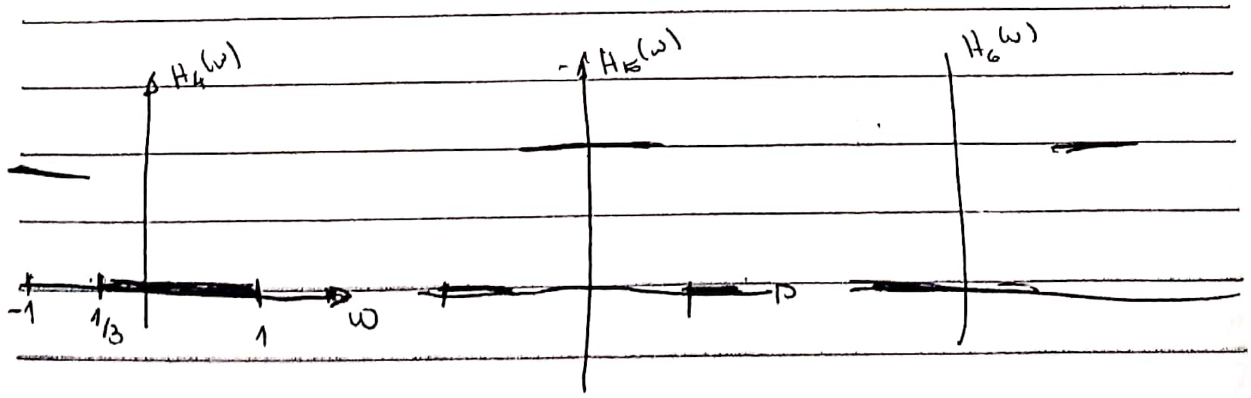
$$= \begin{cases} \mathbb{P}(\emptyset) = 0 & ; \varepsilon \geq 1 \\ \dfrac{-\varepsilon^{1/n} + 1}{2} + \dfrac{1 - \varepsilon^{1/\varepsilon}}{2} \xrightarrow[n \to \infty]{} 0, & 0 < \varepsilon < 1 \end{cases}$$

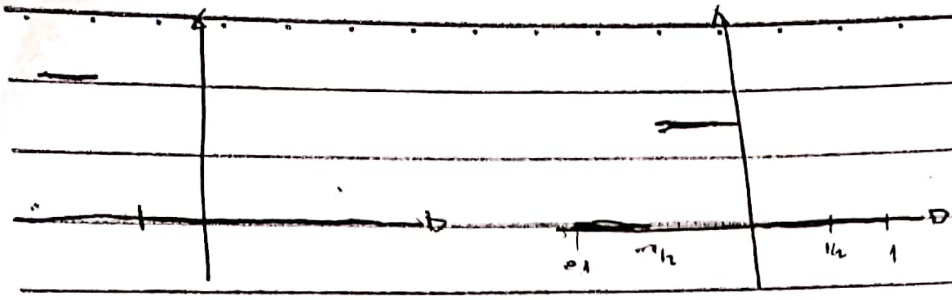
Portanto, $\forall \varepsilon > 0$,

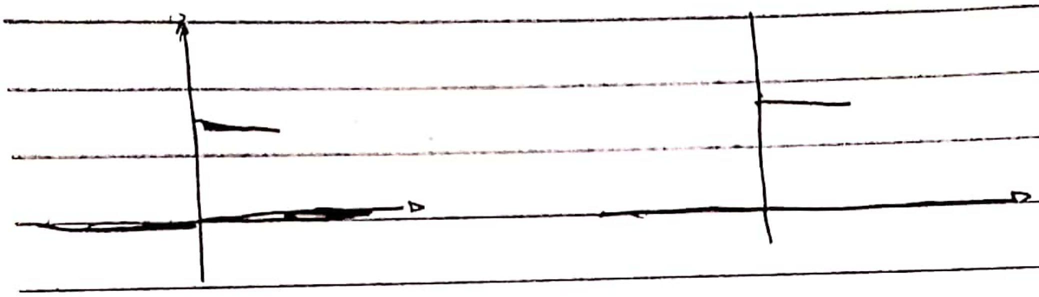
$\lim_{n \to \infty} \mathbb{P}(|Y_n - 0| > \varepsilon) = 0$ e, portanto, $Y_n \xrightarrow{\mathbb{P}} 0$

$(H_n)_{n \geq 1}$









$$\mathbb{P}(\{w\in\Omega : |H_n(w)-0|>\varepsilon\}) \overset{(0<\varepsilon<1)}{=} \left(1, 1/2, 1/2, 1/3, 1/3, 1/3, \dots\right) \to 0$$

Logo

$$\lim_{n\to\infty} \mathbb{P}(\{w: |H_n(w)-0|>\varepsilon\}) = 0, \quad \forall \varepsilon>0$$

Note que a sequência $(H_n(w))_{n\geq 1}$ "possui" infinitos zeros e infinitos 1's, para todo $w\in\Omega$.

Logo, $(H_n(w))_{n\geq 1}$ não converge, $\forall w\in\Omega$

consequentemente, $H_n \cancel{\xrightarrow{q.c.}}$.

Probabilidade e Inferência Estatística I

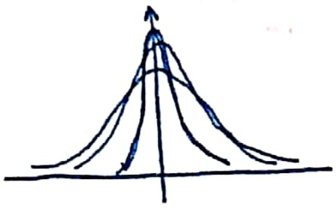
Aula 21- 30/04/2013

Luis Gustavo Esteves

$X_n \xrightarrow{P} X \iff \forall \varepsilon > 0, \lim_{n \to \infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0.$

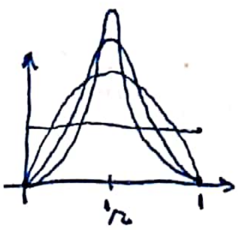
(4) $X_n \sim N(0, 1/n)$

$X_n \xrightarrow{D} 0$?



$X_n \sim Beta(n,n)$

$X_n \xrightarrow{D} 1/2$



Desigualdade de Markov

Seja X v.a. não negativa. Então, para todo $a > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{E(X)}{a}.$$

$$X = X \mathbb{I}_{X^{-1}([a,\infty))} + X \mathbb{I}_{X^{-1}([0,a))}$$

$$X \geq X \mathbb{I}_{X^{-1}([a,\infty))} \geq a \mathbb{I}_{X^{-1}([a,\infty))} \Rightarrow E(X) \geq a\, \mathbb{P}(X^{-1}([a,\infty))) \Rightarrow E(X) \geq a\, \mathbb{P}(X \geq a).$$

Desigualdade de Tchebyschev.

Seja X v.a. tal que $E(X) < \infty$.

$$P(|X - E(X)| \geq a) \leq \frac{Var(X)}{a^2}.$$

Basta tomar $Y = (X - E(X))^2$.

Aplicando a Des. Markov,

$$P(|X - E(X)| \geq a) = P(|X - E(X)|^2 \geq a^2) = P((X - E(X))^2 \geq a^2) \overset{D.M.}{\leq}$$

$$= \frac{E(X - E(X))^2}{a^2} = \frac{Var(X)}{a^2}.$$

Exemplo 5.

$X_n \sim Beta(n,n)$

$\varepsilon > 0$

$$P(|X_n - \underbrace{1/2}_{E(X_n)}| \geq \varepsilon) \leq \frac{Var(X_n)}{\varepsilon^2} = \frac{1}{\varepsilon^2} \frac{n \cdot n}{(n+n)^2 (2n+1)} \xrightarrow[n \to \infty]{} 0, \quad \forall \varepsilon > 0.$$

Logo, $X_n \xrightarrow{P} 1/2$.

Exemplo 6.

$(X_n)_{n \geq 1}$ seq. v.a. tais que $E(X_n) = \mu, \forall n \geq 1$.

$(X_n)_{n\geq 1}$ são não-correlacionadas

Seja $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$

$Var(X_n) \leq c, \forall n \geq 1$

$$P(|\bar{X}_n - \underbrace{\mu}_{E(\bar{X})}| \geq \varepsilon) \overset{\text{Des. Markov}}{\leq} \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{1}{n^2 \varepsilon^2} Var(X_1 + \dots + X_n) = \frac{1}{n^2 \varepsilon^2} \sum_{i=1}^{n} Var(X_i) \leq \frac{nc}{n^2 \varepsilon^2} \xrightarrow[n \to \infty]{} 0.$$

Logo, $\bar{X}_n \xrightarrow{P} \mu$.

Em geral, dizemos que a sequência

$(X_n)_{n\geq 1}$ satisfaz a Lei Fraca dos Grandes Números se
(Forte)

$$\frac{(X_1+\dots+X_n)-E(X_1+\dots+X_n)}{n} \xrightarrow{P} 0 \quad \left(\xrightarrow{q.c.}\right)$$

Exemplo 7. $(X_n)_{n\geq 1}$ v.a. independentes com $X_n \sim U(0,n)$

$$Y_n = \mathbb{I}_{(0,1)}(X_n)$$

Vimos que $\cancel{Y_n \xrightarrow{q.c}}$

$$P(|Y_n-0|>\varepsilon) = \begin{cases} 0 & , \varepsilon > 1 \\ P(Y_n=1) = P(X_n \in (0,1)) = 1/n & , 0<\varepsilon<1 \end{cases}$$

Assim, para todo $\varepsilon>0$,

$$P(|Y_n-0|>\varepsilon) \leq \frac{1}{n} \xrightarrow[n\to\infty]{} 0.$$

Logo, $Y_n \xrightarrow{P} 0$.

Resultado:

$$X_n \xrightarrow{q.c.} X \Rightarrow X_n \xrightarrow{P} X.$$

$$X_n \xrightarrow{q.c.} X \Rightarrow \forall \varepsilon>0, P(\limsup_{n\to\infty}\{|X_n-X|>\varepsilon\})=0 \Rightarrow \forall \varepsilon>0, P(\bigcap_{n=1}^{\infty}\underbrace{\bigcup_{k=n}^{\infty}\{|X_k-X|>\varepsilon\}}_{B_n})=0$$

$$\Rightarrow \forall \varepsilon>0, \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty}\{|X_k-X|>\varepsilon\}\right)=0. \Rightarrow$$

$\Rightarrow \forall\, \varepsilon > 0,\ 0 = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k=n}^{\infty} \{|X_k - X| > \varepsilon\} \right) \geq \limsup_{n\to\infty} \mathbb{P}(\{|X_n - X| > \varepsilon\}) \geq \liminf_{n\to\infty} \mathbb{P}(\{|X_n - X| > \varepsilon\}) \geq 0 \Rightarrow$

$\Rightarrow \forall\, \varepsilon > 0,$

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0 \quad e,$$

portanto, $X_n \xrightarrow{\mathbb{P}} X$.

Propriedades:

(1) $\left.\begin{array}{l} X_n \xrightarrow{\mathbb{P}} X \\ Y_n \xrightarrow{\mathbb{P}} Y \end{array}\right\} X_n \pm Y_n \xrightarrow{\mathbb{P}} X \pm Y$

$(aX_n \xrightarrow{\mathbb{P}} aX)$ (exercício)

$$|(X_n + Y_n) - (X+Y)| = |(X_n - X) + (Y_n - Y)| \leq |X_n - X| + |Y_n - Y|$$

$\{\omega : |(X_n(\omega) + Y_n(\omega)) - (X(\omega) + Y(\omega))| < \varepsilon\} \supseteq \{\omega : |X_n(\omega) - X(\omega)| < \varepsilon/2\} \cap \{\omega : |Y_n(\omega) - Y(\omega)| < \varepsilon/2\} \Rightarrow$

$\{\omega : |(X_n(\omega) + Y_n(\omega)) - (X(\omega) + Y(\omega))| \geq \varepsilon\} \subseteq \{\omega : |X_n(\omega) - X(\omega)| \geq \varepsilon/2\} \cup \{\omega : |Y_n(\omega) - Y(\omega)| \geq \varepsilon/2\} \Rightarrow$

$\Rightarrow \mathbb{P}(|(X_n + Y_n) - (X+Y)| \geq \varepsilon) \leq \mathbb{P}(|X_n - X| \geq \varepsilon/2) + \mathbb{P}(|Y_n - Y| \geq \varepsilon/2) \Rightarrow$

$\Rightarrow \lim_{n\to\infty} \mathbb{P}(|(X_n + Y_n) - (X+Y)| \geq \varepsilon) = 0, \forall\, \varepsilon > 0$. Logo, $X_n + Y_n \xrightarrow{\mathbb{P}} X + Y$.

(2) $\left.\begin{array}{l} X_n \xrightarrow{\mathbb{P}} X \\ X_n \xrightarrow{\mathbb{P}} Y \end{array}\right\} \Rightarrow \mathbb{P}(X = Y) = 1$

$|X - Y| = |X - X_n + X_n - Y| \leq \underbrace{|X_n - X|}_{< \frac{1}{2m}} + \underbrace{|X_n - Y|}_{< \frac{1}{2m}}, \forall m \in \mathbb{N}^*$

mesma idéia da anterior.

(3) $X_n \xrightarrow{P} X \Rightarrow g(X_n) \xrightarrow{P} g(X)$

$g: \mathbb{R} \to \mathbb{R}$ contínua

Convergência em Distribuição

"Convergência da sequência de funções de distribuições para uma função. As variáveis aleatórias envolvidas não precisam estar definidas em um mesmo espaço de probabilidade".

$(\Omega_1, \mathcal{F}_1, \mathbb{P}_1) \xrightarrow{X_1} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_{X_1}) \nearrow F_{X_1}$

$(\Omega_2, \mathcal{F}_2, \mathbb{P}_2) \xrightarrow{X_2} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_{X_2}) \to F_{X_2}$

Sejam $X_1, X_2, \ldots$ v.a. com funções de distribuição

$F_{X_1}, F_{X_2}, \ldots$ respectivamente.

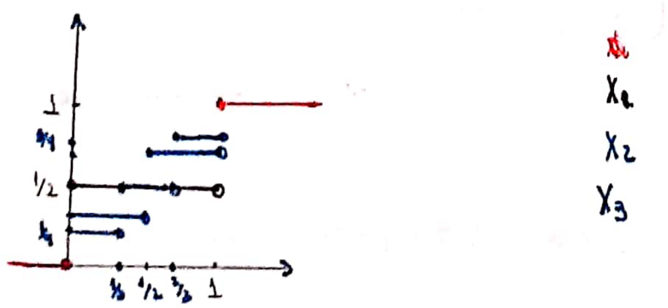
Seja X uma v.a. com função de distribuição F_X.

Dizemos que $(X_n)_{n\geq 1}$ converge em distribuição para X (ou $(F_{X_n})_{n\geq 1}$ converge fracamente para F_X) se

$F_{X_n}(x) \xrightarrow[n\to\infty]{} F_X(x)$, $\forall x \in C_X$, onde C_X é o conjunto dos pontos de continuidade de F_X.

Exemplo 1.

$X_n \sim U(\{0, 1/n, 2/n, \ldots, n-1/n, 1\})$ $\quad \mathbb{P}(X_n = j/n) = \frac{1}{n+1}$, $j \in \{0, 1, \ldots, n\}$



$$F_{X_n}(x) = \begin{cases} 0 \ , & x < 0 \\ \dfrac{\sum_{j=0}^{n} \mathbb{I}_{(-\infty,x]}(j/n)}{n+1} = \dfrac{[nx]+1}{n+1} \ , & 0 \leq x < 1 \\ 1 \ , & x \geq 1. \end{cases}$$

$$\lim_{n\to\infty} F_{X_n}(x) = \begin{cases} 0 \ , & x < 0 \\ x \ , & 0 \leq x < 1 \\ 1 \ , & x \geq 1 \end{cases}$$

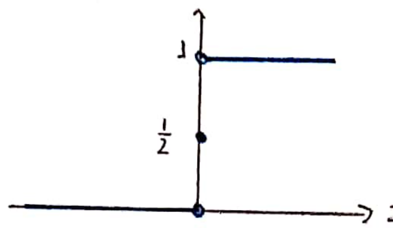
Para $x \in [0,1)$,

$$\underset{\downarrow \atop x}{\frac{nx}{n+1}} < \frac{[nx]+1}{n+1} \leq \underset{\downarrow \atop x}{\frac{nx+1}{n+1}}$$

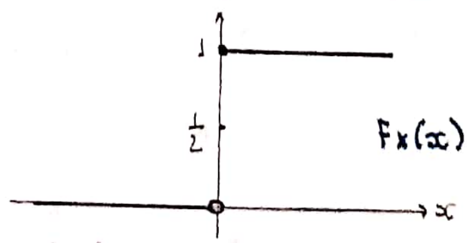
Exemplo 2: $X_n \sim N(0, 1/n)$.

$$F_{X_n}(x) = \mathbb{P}(X_n \leq x) = \mathbb{P}\left(\frac{X_n}{\sqrt{1/n}} \leq \frac{x}{\sqrt{1/n}}\right) =$$

$$= \mathbb{P}(Z \leq x\sqrt{n}) \xrightarrow{n\to\infty} \begin{cases} 0 \ , & x < 0 \\ 1/2 \ , & x = 0 \\ 1 \ , & x > 0 \end{cases}$$



Não é f. de distribuição.



$F_X(x) = \mathbb{I}_{\mathbb{R}_+}(x)$

Note que $F_{X_n}(x) \to F_X(x)$, $\forall x \neq 0$ ($\forall x$ ponto de continuidade de F_X).

Assim, $X_n \xrightarrow{d} X$:

$$X_n \xrightarrow{d} 0.$$

Exemplo 3. $X_n \sim \mathcal{U}(n, n+1)$

$$F_{X_n}(x) = \begin{cases} 0, & x < n \\ \dfrac{x-n}{(n+1)-n} = x-n, & n \leq x < n+1 \\ 1, & x \geq n+1 \end{cases}$$

"tight"

Se $x < 0$, então

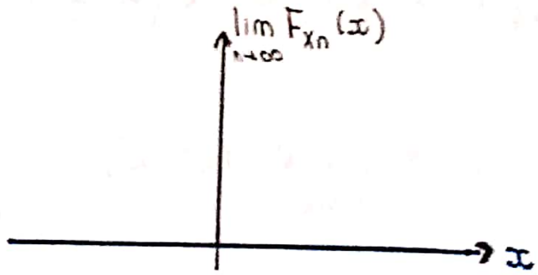
$$\lim_{n\to\infty} F_{X_n}(x) = 0$$

Para $x > 0$, existe $n_0 = n_0(x)$ tal que

$$x < n_0 \, (\overbrace{[x]+1}^{n_0})$$

Então, para $n \geq n_0$, $F_{X_n}(x) = 0$.

Logo, $\lim\limits_{n\to\infty} F_{X_n}(x) = 0, \forall x > 0$.



Note que $(F_{X_n}(x))_{n \geq 1}$ converge para todo x, mas a função limite NÃO é função de distribuição.

Nesse caso,

$$\cancel{X_n \xrightarrow{d}}$$

Resultado:

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

Seja $x \in \mathbb{R}$ ponto de continuidade de F_X. $\forall\, m \in \mathbb{N}^*$

$$\{X_n \leq x\} = \{X_n \leq x, |X_n - X| \leq \tfrac{1}{m}\} \cup \{X_n \leq x, |X_n - X| > 1/m\} \Rightarrow \left.\begin{matrix} X_n \leq x \\ X_n - \varepsilon \leq X \leq X_n + \varepsilon \end{matrix}\right\} \to$$

$$X \leq x + \varepsilon$$

$$\{X_n \leq x\} \subseteq \{X \leq x + 1/m\} \cup \{|X_n - X| > 1/m\} \Rightarrow$$

$$\mathbb{P}(X_n \leq x) \leq \mathbb{P}(X \leq x + 1/m) + \mathbb{P}(|X_n - X| > 1/m) \qquad (\mathrm{I})$$

Analogamente

$$\{X \leq x - 1/m\} = \{X \leq x - 1/m, |X_n - X| \leq 1/m\} \cup \{X \leq x - 1/m, |X_n - X| > 1/m\}$$

$$\{X \leq x - 1/m\} \subseteq \{X_n \leq x\} \cup \{|X_n - X| \geq 1/m\} \Rightarrow$$

$$\mathbb{P}(X \leq x - 1/m) \leq \mathbb{P}(X_n \leq x) + \mathbb{P}(|X_n - X| \geq 1/m) \quad (\mathrm{II})$$

Aplicando limite em n (em I e II)

$$\left.\begin{matrix} \limsup \mathbb{P}(X_n \leq x) \leq F_X\left(x + \frac{1}{m}\right) \\ \\ F_X(x - 1/m) \leq \lim\limits_{n\to\infty} \mathbb{P}(X_n \leq x) \end{matrix}\right\} \overset{m\to\infty}{\Rightarrow} \lim_{n\to\infty} \underbrace{\mathbb{P}(X_n \leq x)}_{F_{X_n}(x)} = F_X(x)$$

Logo, $X_n \xrightarrow{d} X$. (Pois $F_{X_n}(x) \xrightarrow{n\to\infty} F_X(x)$ p/ todo ponto de continuidade de F_X)

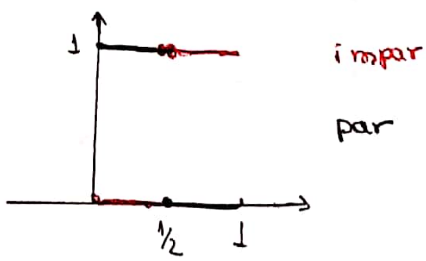
Exemplo:

$\Omega = [0,1]$

$\mathcal{F} = \mathcal{B}([0,1])$

$\mathbb{P}$: "comprimento"

$$X_n(\omega) = \begin{cases} \mathbb{I}_{[0,1/2]}(\omega), & n \text{ par} \\ \\ \mathbb{I}_{(1/2,1]}(\omega), & n \text{ ímpar} \end{cases}$$



$$X(\omega) = \mathbb{I}_{[1/4,3/4]}(\omega)$$

$X_n \sim Ber(1/2), \forall n \geq 1$

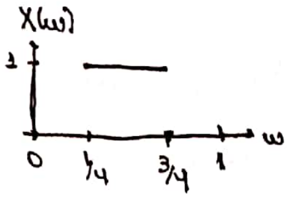
Logo, $F_{X_n}(x) = \frac{1}{2}\mathbb{I}_{[0,1)}(x) + \mathbb{I}_{[1,\infty)}(x)$

$X \sim Ber(1/2)$

Logo $X_n \xrightarrow{d} X$.

$0 < \varepsilon < 1$

$$P(|X_n - X| > \varepsilon) = P(X_n = 1, X = 0) + P(X_n = 0, X = 1) =$$



$$= \begin{cases} \underset{1/4}{\underset{\shortparallel}{P((0,1/4))}} + \underset{1/4}{\underset{\shortparallel}{P((1/2,3/4))}}, & n \text{ par} \\ \underset{1/4}{\underset{\shortparallel}{P((3/4,1))}} + \underset{1/4}{\underset{\shortparallel}{P((1/4,1/2))}}, & n \text{ ímpar} \end{cases}$$

$$= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \forall n \geq 1.$$

$P(|X_n - X| > \varepsilon) = 1/2, \forall n \geq 1.$

Logo, $\lim_{n \to \infty} P(|X_n - X| > \varepsilon) = \frac{1}{2} \neq 0.$

Portanto, $X_n \not\xrightarrow{P} X$.

Resultado.

Sejam $X, X_1, X_2, \ldots$ v.a. discretas assumindo valores em $\mathbb{N}$

Sejam $(\underset{P(X_1=k)}{p_1(k)})_{k\in\mathbb{N}}, (p_2(k))_{k\in\mathbb{N}}, \ldots$ as funções das probabilidades de $X_1, X_2, \ldots,$ respectivamente.

Seja $(\underset{\mathbb{P}(X=k)}{p(k)})_{k\in\mathbb{N}}$

Resultado.

$$X_n \xrightarrow{d} X \Leftrightarrow p_n(k) \xrightarrow[n\to\infty]{} p(k), \forall k\in\mathbb{N}$$

Verificação:

$(\Rightarrow)$

Note que $p_n(k) = \mathbb{P}(X_n=k) = \mathbb{P}(k-1/2 \leq X_n \leq k+1/2) =$

pois $k-\frac{1}{2}$ e $k+\frac{1}{2}$ são pontos de continuidade de F_X

$$= F_{X_n}(k+1/2) - F_{X_n}(k-1/2) \xrightarrow{n\to\infty} F_X(k+1/2) - F_X(k-1/2) = \mathbb{P}(X=k) = p(k), \forall \in\mathbb{N}$$

$(\Leftarrow)$

$$F_{X_n}(x) = \sum_{k=0}^{[x]} \underbrace{\mathbb{P}(X_n=k)}_{p_n(k)} \xrightarrow[n\to\infty]{} \sum_{k=0}^{[x]} \underbrace{\mathbb{P}(X=k)}_{p(k)} = F_X(x).$$

Exemplo.

$$X \sim Bin(n, p_n), \quad np_n \xrightarrow{n\to\infty} \lambda, \lambda > 0.$$

$k\in\mathbb{N}$ $(n\geq k)$

$$\underset{p_n(k)}{\mathbb{P}(X_n=k)} \overset{(n\geq k)}{=} \binom{n}{k} p_n^k (1-p_n)^{n-k} =$$

$$= \frac{n(n-1)\dots(n-k+1)\cancel{(n-k)!}}{k!\,\cancel{(n-k)!}}\, p_n^k\,(1-p_n)^{n-k} =$$

$$= \frac{1}{k!}\left\{\prod_{j=0}^{k-1}(n-j)\,p_n\right\}\left(1-{}^{n p_n}/_{n}\right)^{n-k}$$

$$= \frac{1}{k!}\left\{\prod_{j=0}^{k-1}(n-j)\,p_n\right\}\left(1-\frac{n p_n}{n}\right)^{n}\left(1-\frac{n p_n}{n}\right)^{-k}$$

$p_n \to 0$

$a_n = n p_n \to \lambda$

$$\xrightarrow[n\to\infty]{} \frac{1}{k!}\lambda^k e^{-\lambda} = p(k),\ \forall k \in \mathbb{N}.$$

Logo $X_n \xrightarrow{d} X$, onde $X \sim \text{Poisson}(\lambda)$.

Exercício.

$X_n \sim HG(np, n(1-p), k)$, $p \in (0,1)$

$X_n \xrightarrow{d} X$, onde $X \sim \text{Binomial}(k, p)$

$k \in \mathbb{N}$,

$p_n(k) = \mathbb{P}(X_n = k) = \binom{np}{k}\Big(^{n(1-p)}$

02/05 $X_n \xrightarrow{d} X \iff F_{X_n}(t) \to F_X(t)$, $\forall t \in C_X$

P/ Verificar: $X_n \xrightarrow{d} c \Rightarrow X_n \xrightarrow{\mathbb{P}} c$

No exemplo $X_n \sim HG(k_1 n, k_2 n, m)$

$k_1, k_2 \in \mathbb{N}^*$

$X_n \xrightarrow{d} X \sim Bin\left(m, \frac{k_1}{k_1+k_2}\right)$

Propriedades (Teo de Slutsky)

(1) $\left.\begin{array}{l} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{d} c \end{array}\right| \Rightarrow X_n + Y_n \xrightarrow{d} X + c$

$\mathbb{P}(X_n + Y_n \leq t) \to \mathbb{P}(X + c \leq t)$
$= \mathbb{P}(X \leq t - c)$

(2) $\left.\begin{array}{l} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{d} c \end{array}\right| \Rightarrow X_n Y_n \xrightarrow{d} cX$

Ideia

$\mathbb{P}(X_n + Y_n \leq t) = \mathbb{P}(X_n + Y_n \leq t, Y_n < c - \frac{1}{m}) +$
$+ \mathbb{P}(X_n + Y_n \leq t, Y_n \geq c - 1/m)$

$\leq \mathbb{P}(Y_n < c - 1/m)$ $\quad$ $\mathbb{P}(X_n \leq -)$

$\mathbb{P}(X_n \leq t - c - 1/m) = \mathbb{P}(X_n \leq t - c - \frac{1}{m}, X_n + Y_n \leq t) + \mathbb{P}(X_n \leq t - c - 1/m, X_n + Y_n > t)$

Ex: $\Omega = [0,1]$
$\mathcal{F} = \mathcal{B}([0,1])$
$\mathbb{P} =$ "comprimento"

$$X_n(\omega) = \begin{cases} \mathbb{I}_{[0,1/2]}(\omega) & , n \text{ par} \\ \mathbb{I}_{[1/2,1]}(\omega) & , n \text{ ímpar} \end{cases} \qquad Y_n(\omega) = \begin{cases} \mathbb{I}_{[1/2,1]}(\omega) & , n \text{ par} \\ \mathbb{I}_{[0,1/2]} & , n \text{ ímpar} \end{cases}$$

$$X(\omega) = \mathbb{I}_{[1/4,3/4]}(\omega)$$

$X_n \sim Ber(1/2), \forall n \geq 1$
$Y_n \sim Ber(1/2), n \geq 1$

$$X_n \xrightarrow{d} X \qquad Y_n \xrightarrow{d} X \quad \Big| \quad \cancel{X_n + Y_n \xrightarrow{d} 2X}$$

No entanto,

$$X_n(\omega) + Y_n(\omega) = 1, \forall \omega \in \Omega$$

Logo, $\mathbb{P}(X_n + Y_n = 1) = 1, \forall n \geq 1$

$$X_n + Y_n \xrightarrow{d} 1$$

$Z = X + X$ é tal que $\mathbb{P}(Z = 2) = \mathbb{P}(Z = 0) = 1/2$

Ideia da Prox

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sqrt{n\sigma^2}} \longrightarrow N(0,1)$$

$$\mathbb{P}(Z_n \leq t) \longrightarrow \mathbb{P}(Z \leq t)$$

/ /

Resultado

$(X_n)_{n\geq 1}$ seq. de v.a com fç geradoras $(\varphi_{X_n})_{n\geq 1}$

X v.a com fç geradora φ_X

$$\varphi_{X_n}(t) \to \varphi_X(t), \forall t \in I \Rightarrow X_n \xrightarrow{d} X$$

De um modo mais geral, se $\varphi_{X_n}(t) \to \varphi$ e φ é contínua no zero então φ é fç geradora de algumas v.a X e $X_n \xrightarrow{d} X$

Exemplos $\quad \mathbb{P}(X_0=0)=1$.

(1) $X_n \sim N(0, 1/n)$

$$\varphi_{X_n}(t) = e^{t.0} . e^{(\frac{1}{n}t^2)/2} = e^{t^2/2n} \to 1, \forall t \in \mathbb{R}$$

$$\varphi_{X_0}(t) = e^{t.0}\mathbb{P}(X_0=0) = 1, \forall t \in \mathbb{R}$$

Pelo Resultado $X_n \xrightarrow{d} X_0 \equiv 0$

Ex 2.

$$\boxed{\begin{array}{c}\varphi_{Z}(t) = (pe^t + 1-p)^n \\ Z \sim Bin(n,p)\end{array}}$$

$X_n \sim Bin(n, p_n)$

$np_n \xrightarrow[n\to\infty]{} \lambda, \lambda > 0$

$$\varphi_{X_n}(t) = (p_n e^t + 1 - p_n)^n = (1 + p_n(e^t-1))^n =$$

$$= \left(1 + \frac{np_n(e^t-1)}{n}\right)^n \xrightarrow[n\to\infty]{} e^{\lambda(e^t-1)} = \varphi_X(t), X \sim Po(\lambda)$$

Logo, $X_n \xrightarrow{d} X$

(/ /)

(3) $X_n \sim N(0,n)$

$$\varphi_{X_n}(t) = e^{t.0} \cdot e^{nt^2/2} = e^{nt^2/2} \rightarrow \begin{cases} 1, & t = 0 \\ \infty, & t \neq 0 \end{cases}$$

(4) $X_n \sim Exp(1/n)$?

Fazer

Teo do Limite Central

$(X_n)_{n \geq 1}$ v.a i.i.d com $E(X_1) = \mu$ e
$VAR(X_1) = \sigma^2 < \infty$

Então,

$$\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} Z \sim N(0,1)$$

Verificação

$$\varphi_{\frac{X_1+\dots+X_n-n\mu}{\sqrt{n\sigma^2}}}(t) = \varphi_{\frac{1}{\sqrt{n\sigma^2}}\sum_{i=1}^{n} X_i-\mu}(t) = \varphi_{\sum_{i=1}^{n} X_i-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right) \overset{ind}{=}$$

$$= \prod_{i=1}^{n} \varphi_{X_i-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right) \overset{i.d}{=} \Big(\varphi_{\underset{\text{qq um}}{X_{17}}-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right)\Big)^n$$

$$\varphi_{X_{17}-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right) = \varphi_{X_{17}-\mu}(0) + \overbrace{\varphi'_{X_{17}-\mu}(0)}^{E(X_{17}-\mu)=0}\left(\frac{t}{\sigma\sqrt{n}}-0\right) +$$

$$+\frac{1}{2!}\varphi''(\Xi_n)\left(\frac{t}{\sigma\sqrt{n}}-0\right)^2, \text{ onde } \quad 0 \leq \Xi_n < \frac{t}{\sigma\sqrt{n}}$$

Assim, $\varphi_{X_{17}-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 + \frac{\varphi''(\Xi_n)t^2}{\sigma^2 n}$

Logo, $\varphi''(0) = E\big((X_{17}-\mu)^2\big)$

$$\varphi_{\frac{X_1+\dots+X_n-n\mu}{\sqrt{n\sigma^2}}}(t) = \left(1 + \frac{\overbrace{\frac{\varphi''(\Xi_n)t^2}{2\sigma^2}}^{=a_n}}{n}\right)^n \xrightarrow[n\to\infty]{} e^{t^2/2} = \varphi_Z(t)$$

Logo, $\frac{X_1+\dots+X_n-n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0,1)$

p/ n grande $\sum_i X_i \overset{a}{\approx} N(n\mu, n\sigma^2)$

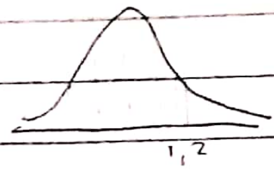
Exemplos

$X_1, \dots, X_{100}$ iid Beta$(4,4)$

média da Beta

$$\mathbb{P}(X_1 + \dots + X_{100} \le 52) = \mathbb{P}\left(\frac{X_1 + \dots + X_{100} - 100 \cdot \frac{1}{2}}{\sqrt{100 \cdot \frac{4 \cdot 4}{8 \cdot 8 \cdot 9}}} \le \frac{52 \cdot 100 \cdot 1/2}{\sqrt{100 \cdot \frac{4 \cdot 4}{8 \cdot 8 \cdot 9}}}\right) =$$

$$\simeq \mathbb{P}\left(Z \le \frac{52-50}{\sqrt{\frac{25}{9}}}\right) = \mathbb{P}(Z \le 1,2)$$



Ex 2 : 200 lançamentos de um dado

X_i : valor observado no i-ésimo lançamento
$i = 1, \dots, 200$

$$\mathbb{P}(X_1 + \dots + X_{200} > 670) = 1 - \mathbb{P}(X_1 + \dots + X_{200} \le 670) =$$

$$= 1 - \mathbb{P}\left(\frac{(X_1 + \dots + X_{200}) - 200(7/2)}{\sqrt{200 \cdot \frac{35}{12}}} \le \frac{670 - 200(7/2)}{\sqrt{200 \cdot \frac{35}{12}}}\right) \simeq 1 - \mathbb{P}\left(Z \le \frac{-30}{10\sqrt{\frac{35}{6}}}\right)$$

$$\simeq 1 - \mathbb{P}\left(Z \le \frac{-3\sqrt{6}}{\sqrt{35}}\right)$$

$X_i \sim \mathcal{U}(\{1,2,3,4,5,6\})$

$E(X_i) = 7/2 \qquad Var(X_i) = \dfrac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} - \left(\dfrac{7}{2}\right)^2 =$

$$= \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

(3) $\lim_{n\to\infty} \underbrace{\int_0^n \frac{1}{(n-1)!} t^{n-1} e^{-t} \, dt}_{Gama(n,1)}$?

$= \lim_{n\to\infty} \mathbb{P}(X_1+\dots+X_n \le n)$, onde $(X_n)_{n\geq 1}$ são v.a.i.i.d $Exp(1)$

$$= \lim_{n\to n} \mathbb{P}\left(\frac{X_1+\dots+X_n - n}{\sqrt{n}} \le \frac{n-n}{\sqrt{n}}\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{X_1+\dots+X_n-n}{\sqrt{n}} \le 0\right)$$

$$\parallel$$

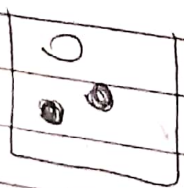
$$\mathbb{P}(Z \le 0) = \frac{1}{2}$$

Lei Forte dos Grandes Números (LFGN)

$(X_n)_{n\geq 1}$ v.a.i.i.d tais que $E(X_1) = \mu < \infty$

Então $\overline{X}_n = \frac{X_1+\dots+X_n}{n} \xrightarrow{q.c} \mu$

Comentário



$\frac{X_1 + \quad + X_n}{n} \to Beta(2,1)$

$\frac{X_1 + \quad + X_n}{n} \to \frac{2}{3}$

/ /

Exemplo

(1) $(X_n)_{n\geq 1}$ v.a. iid Poisson (λ), $\lambda > 0$

Pela LFGN

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow{qc} E(X_1) = \lambda$$

$$\mathbb{P}(X_1^2 \leq t) = \mathbb{P}(X_j^2 \leq t)$$

$$\frac{X_1^2 + \dots + X_n^2}{n} \xrightarrow{qc} E(X_1^2) = \lambda + \lambda^2$$

$$E(X_1^2) = Var(X_1) + E^2(X_1) = \lambda + \lambda^2 < \infty$$

(2) $(X_n)_{n\geq 1}$ v.a.i.id Gama(a,b), $a,b > 0$

$$\frac{X_1 + \dots + X_n}{\sqrt{n(X_1^2 + \dots + X_n^2)}} = \frac{\frac{X_1 + \dots + X_n}{n}}{\sqrt{\frac{X_1^2 + \dots + X_n^2}{n}}} \xrightarrow{qc} \frac{a/b}{\sqrt{\frac{a}{b^2} + \frac{a^2}{b^2}}}$$

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow{qc} E(X_1) = a/b$$

$$\frac{X_1^2 + \dots + X_n^2}{n} \to \frac{a}{b^2} + \frac{a^2}{b^2}$$

(3) (X_n) v.a ind. tais que $X_{2n-1} \sim Beta(2,1)$, $n \geq 1$

$X_{2n} \sim U(0,1)$, $n \geq 1$

$$\overline{X}_n = \frac{X_1 + \dots + X_n}{n}$$

$$\overline{X}_{2n} = \frac{X_1 + X_2 + \dots + X_{2n}}{2n} =$$

$$= \frac{X_1 + X_3 + \dots + X_{2n-1}}{2n} + \frac{X_2 + X_4 + \dots + X_{2n}}{2n} =$$

$$= \frac{1}{2} \frac{X_1 + X_3 + \dots + X_{2n-1}}{n} + \frac{1}{2} \frac{X_2 + X_4 + \dots + X_{2n}}{n} \xrightarrow{qc}$$

$$\frac{1}{2} E(X_1) \qquad \frac{1}{2} E(X_2)$$

$$\xrightarrow{qc} \frac{1}{2} E(X_1) + \frac{1}{2} E(X_2) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{4}$$

$$\overline{X}_{2n} \xrightarrow{qc} 7/12$$

$$\overline{X}_{2n-1} = \frac{X_1 + \dots + X_{2n-1}}{2n-1} = \frac{X_1 + X_3 + \dots + X_{2n-1}}{2n-1} + \frac{X_2 + X_4 + \dots + X_{2n-2}}{2n-1}$$

$$= \frac{n}{2n-1} \frac{X_1 + X_3 + \dots + X_{2n-1}}{n} + \frac{n-1}{2n-1} \frac{X_2 + X_4 + \dots + X_{2n-2}}{n-1} \xrightarrow{qc} \frac{7}{12}$$

Como $\overline{X}_{2n} \xrightarrow{qc} 7/12$ $\quad \overline{X}_{2n-1} \xrightarrow{qc} 7/12$

$A = \{ w : \overline{X}_{2n}(w) \to 7/12 \}$ $\quad$ $B = \{ w : \overline{X}_{2n-1}(w) \to 7/12 \}$

$w_0 \in A \cap B \Rightarrow$

$\Rightarrow \overline{X}_{2n}(w_0) \to 7/12 \Rightarrow \overline{X_n}(w_0) \to 7/12 \Rightarrow$

$\overline{X}_{2n-1}(w_0) \to 7/12$

$\Rightarrow w_0 \in \{ w . \overline{X}_n(w) \to 7/12 \}$

$A \cap B \subseteq$

$\Rightarrow \mathbb{P}(\overline{X}_n \to 7/12) = 1$ e, i.e., $\overline{X}_n = \dfrac{X_1 + \dots + X_n}{n} \xrightarrow{qc} \dfrac{7}{12}$

(4) $(X_n)_{n \geq 1}$ v.a.i.i.d Ber(p) $p \in (0,1)$

$Y_1 = X_1 \cdot X_2 X_3$

$Y_2 = X_2 X_3 \cdot X_4$ $\quad \rightsquigarrow \dfrac{Y_1 + Y_2 + \dots + Y_n}{n} \xrightarrow{?} p^3$

$\vdots$

$Y_n = X_n X_{n+1} X_{n+2}$ $\quad$ Verifique que Y_1 e Y_2 não são ind.

Ideia

$\overline{Y}_{3n}$ $\quad Y_1 \; Y_4 \; Y_7 \; \dots$ $\quad \overline{Y}_{3n} = \dfrac{Y_1 + \dots + Y_{3n}}{3n}$

$\overline{Y}_{3n-1}$ $\quad Y_2 \; Y_5 \; Y_8$

$\overline{Y}_{3n-2}$ $\quad Y_3 \; Y_6 \; Y_9$

Módulo II

Aula 1 - 03/05/2013

Luis Gustavo Esteves

Referências

M. H. De Groot (M. J. Schervish)

G. G. Roussas

Casella, Berger

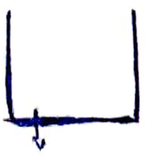
M. J. Schervish

"Theory of Statistics"

Inferência Estatística

"Fazer um procedimento que faça com que você aprenda ou tome decisão a partir dos dados".

Exemplo (1).



parâmetro

θ? número de bolas brancas na urna

espaço paramétrico

$\Theta = \{0, 1, 2, 3, 4, 5\}$

θ brancas

$5-\theta$ verdes

Conhecer o número de bolas brancas na caixa

$X = (X_1, X_2) \to$ amostra

$\mathcal{X} = \{0, 1\}^2 \to$ espaço amostral

- Encontrar alguma relação funcional entre aquilo que queremos descobrir (θ) e $\mathcal{X}$, a amostra extraída.

(2) (MODELO DE TAGUCHI PARA ATRIBUTOS) { Natália
Ana Cristina (Prof. [illegible])
Leyla (2008)

Processo que funciona direitinho até um determinado momento θ desconhecido

$\theta \overset{?}{=}$ momento em que o processo deteriora.

A cada $m \geq 1$ peças é feita uma inspeção do processo.

X: Número de inspeções num ciclo do processo

$\Theta = \mathbb{N}^*$

$\mathcal{X} = \mathbb{N}^*$

(3) N bolas, N conhecido

θ_1 cor 1
θ_2 cor 2
$\vdots$
θ_k cor k

$k \geq 1$ cores

$\theta = (\theta_1, \theta_2, \dots, \theta_k)$

$\Theta = \{(a_1, a_2, \dots, a_k) \in \mathbb{N}^k : a_1 + a_2 + \dots + a_k = N\}$

$X = (X_1, \dots, X_k)$, onde X_i é o número de bola da cor i na amostra $i = 1, \dots, k$

$\mathcal{X} = \{(a_1, \dots, a_k) \in \mathbb{N}^k : a_1 + a_2 + \dots + a_k \leq n\}$

(4)

θ: tempo médio de vida de lâmpadas produzidas

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

$X = (X_1, X_2, X_3, X_4)$

$\mathcal{X} = \mathbb{R}_+^4$

No exemplo (Relação funcional)

1.

$$\mathbb{P}(X_1=1, X_2=1 \mid \theta=j) = \frac{1}{5}\,\frac{(j-1)}{4}$$

2.

$$\mathbb{P}(X=j \mid \theta) = \mathbb{I}_{\{\theta,\theta+1,\ldots\}}(mj)\;\mathbb{I}_{\{1,\ldots,\theta-1\}}(m(j-1))$$

3.

$$\mathbb{P}(X_1=x_1,\ldots,X_k=x_k \mid \theta=(\theta_1,\ldots,\theta_k)) = \frac{\binom{\theta_1}{x_1}\binom{\theta_2}{x_2}\cdots\binom{\theta_k}{x_k}\binom{N-\theta_1-\ldots-\theta_k}{n-x_1-\ldots-x_k}}{\binom{N}{n}}$$

"Distribuição hipergeométrica de variável".

4.

X_1, X_2, X_3, X_4, dado θ, condicionalmente e i d. $Exp(1/\theta)$.

↳ pois se fossem independentes incondicionalmente de θ, então $X_1,\ldots,X_4$ seriam ind. de θ, o que não faria sentido, pois não trariam informação sobre o mesmo.

$$f(x_1,x_2,x_3,x_4 \mid \theta) = \prod_{i=1}^{4} f(x_i \mid \theta)$$

$\theta = ?$

Idealizar um experimento cuja realização deve ter conexão com θ e especificar uma relação funcional entre o registro do experimento e θ.

(X,Θ) desconhecido. Devemos especificar prob.

$$f(x,\theta) = f(x|\theta) f(\theta)$$

$$f(\theta|x) = \frac{f(\theta, x)}{f(x)} = \frac{f(x|\theta) f(\theta)}{f(x)}$$

"A incerteza revista pra θ à luz da observação X".

suposição forte na Bayesiana

$f(X,\theta)$ não teve oscilações entre a suposição e o resultado final

$f(X,\theta)$ se preserva ao longo do tempo, q nem sempre é verdade

Ou seja, distrib. pré-amostragem é a mesma da pós-amostragem

"Interferência de coerência amostral".

No exemplo

1)

$\Theta \sim U(\{0,1,2,3,4,5\})$

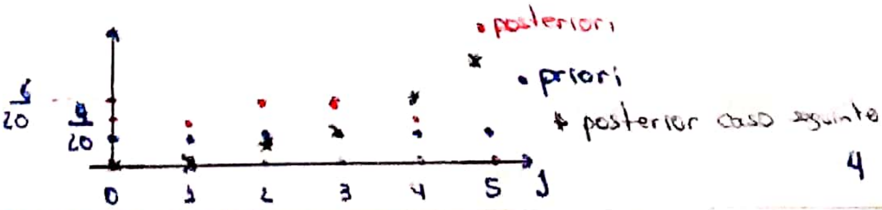
$X_1, X_2 | \Theta$ já conhecida

Experimento

$X_1 = 1,\ X_2 = 0$

$$\Rightarrow \quad P(\Theta = j | X_1 = 1, X_2 = 0) = \frac{P(X_1=1, X_2=0 | \Theta = j) P(\Theta = j)}{\sum_{i=0}^{5} P(X_1 = 1, X_2 = 0 | \Theta = i) P(\Theta = i)} =$$

$$= \frac{\frac{j}{5} \frac{(5-j)}{4} \cdot \frac{1}{6}}{\sum_{i=0}^{5} \frac{i}{5} \frac{(5-i)}{4} \frac{1}{6}} = \frac{j(5-j)}{\sum_{i=0}^{5} i(5-i)}$$



E se $X_1=1$ e $X_2=1$.

$$P(\theta=j/X_1=1, X_2=1) = \frac{P(X_1=1, X_2=1/\theta=j)\,P(\theta=j)}{\sum_{i=0}^{5} P(X_1=1, X_2=1/\theta=i)\,P(\theta=i)}$$

$$= \frac{\frac{j}{5}\frac{(j-1)}{4}\cdot\frac{1}{6}}{\sum_{i=0}^{5}\frac{i}{5}\frac{(i-1)}{4}\frac{1}{6}} = \frac{j(j-1)}{\sum_{i=0}^{5} i(i-1)}$$

* posteriori no gráfico anterior

"posterior de hoje é a priori de amanhã". Não há problemas em construir a posterior dando informações passo a passo, dá resultado o mesmo.

2)

$\theta \sim Geo\left(\frac{1}{100}\right)$

$$P(X=j/\theta) = \mathbb{I}_{\{\theta,\theta+1,\dots\}}(m_j)\,\mathbb{I}_{\{1,\dots,\theta-1\}}(m(j-1))$$

$m=15$

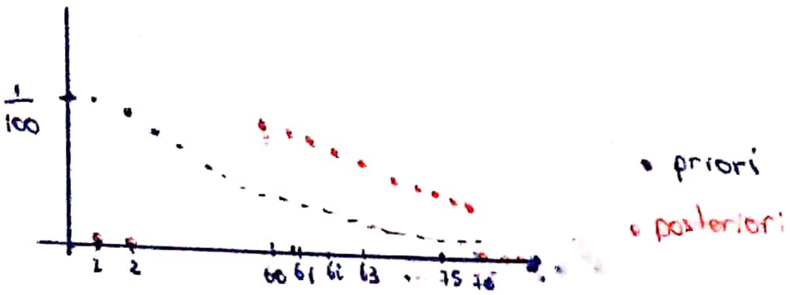
$X=5$

$$P(\theta=j/X=5) = \frac{P(X=5/\theta=j)\,P(\theta=j)}{\sum_{i=1}^{\infty} P(X=5/\theta=i)\,P(\theta=i)} =$$

$75 \geq j$ e $60 \leq j-1$

$$= \frac{\mathbb{I}_{\{j,j+1,\dots\}}(75)\;\mathbb{I}_{\{1,\dots,j-1\}}(60)\left(\frac{99}{100}\right)^{j-1}\frac{1}{100}}{\sum_{i=1}^{\infty}\mathbb{I}_{\{i,i+1,\dots\}}(75)\;\mathbb{I}_{\{1,\dots,i-1\}}(60)\left(\frac{99}{100}\right)^{i-1}\frac{1}{100}} = \frac{\mathbb{I}_{\{61,\dots,75\}}(j)\left(\frac{99}{100}\right)^{j-1}}{\sum_{i=0}^{\infty}\mathbb{I}_{\{61,\dots,75\}}\left(\frac{99}{100}\right)^{i-1}} \Rightarrow$$

$$\Rightarrow P(\theta=j/X=5) = \frac{\left(\frac{99}{100}\right)^{j-1}}{\sum_{i=61}^{75}\left(\frac{99}{100}\right)^{i-1}}\,\mathbb{I}_{\{61,\dots,75\}}(j)$$



Operação Bayesiana

$\Theta = ?$

PRIORI

$\theta / X = ??$

POSTERIORI

Dados

$X = x$

Vimos que existe uma função que relaciona o resultado do experimento (x) com parâmetro (θ).

$f(x|\theta)$

$P(X=x|\theta)$

(1) Quando fixamos $\theta \in \Theta$,

$f(\cdot|\theta)$ é uma distribuição de probabilidade em $\mathcal{X}$. (Distribuição Amostral)

(2) Quando fixamos $x \in \mathcal{X}$,

$f(x|\cdot)$ é uma função não-negativa de θ

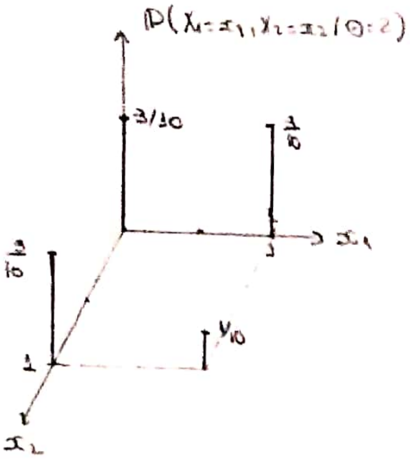
(Função de Verossimilhança Para θ Gerada por x)

$V_x(\cdot)$

$L_x(\cdot)$

Voltando ao Exemplo 1.

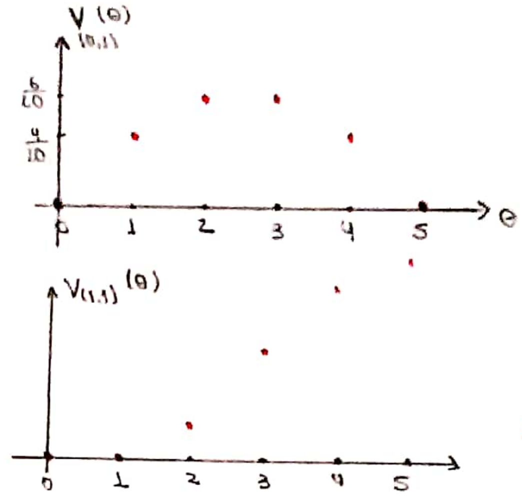
$\Theta = 2$



Tantas quantas os elementos de Θ.

$x = (0,1) \in \mathcal{X}$

$P(X=(0,1)/\theta) = P(X_1=0, X_2=1/\theta) = \frac{(5-\theta)}{5} \cdot \frac{\theta}{4}$



$V_{(1,1)}(\theta) = \frac{\theta(\theta-1)}{20}$

$\sum V_{(1,1)}(\theta) \neq 1$

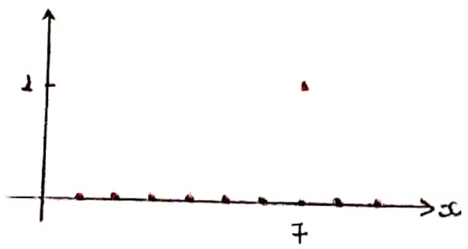
Tantas quantas os elementos em $\mathcal{X}$.

No exemplo 2.

$\Theta = 100$

$n = 15$

$P(X=7 / \Theta=100) = 1$



Tantas qtos o valores de Θ.

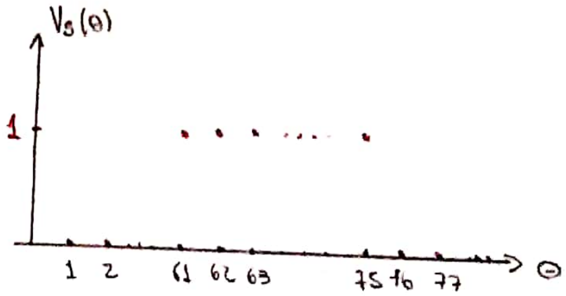
Um conjunto infinito enumerável de funções

Note que $\Theta = 91, 101, \ldots, 105$ levam a mesma $P(X=7/\Theta= \quad)$.

Nesse caso, os pontos $91, \ldots, 105$ são ditos observacionalmente equivalentes, pois trazem a mesma informação.

$x = 5$

$V_5(\theta) = P(X=5/\theta) = \mathbb{I}_{\{61,\ldots,75\}}(\theta)$



Aula 2 - 07/05/2013

Luis Gustavo Esteves

Operação Bayesiana

Verossimilhança gerada para Θ

$$\mathbb{P}(\Theta / \text{Dados}) = \frac{\mathbb{P}(\text{Dados}/\Theta)\,\mathbb{P}(\Theta)}{\sum_i \mathbb{P}(\text{Dados}/i)\,\mathbb{P}(i)} \rightarrow \text{Dist. à priori}$$

↓

Distribuição à posteriori

(1) Temporal Coherence Michael Goldstein (1985)
Bayesian Statistic (Valência)

(2) "Imprevistos e suas consequências". Rosângela H. Loschi (1998)

(1) Urna
(2) Qualidade
} Exemplos

$\mathbb{P}(X=x/\theta)$

(1) Fixado $B \in \Theta$, $\mathbb{P}(X=\cdot/\theta)$ é uma distribuição de probabilidade em $\mathcal{X}$.

(2) Fixado $x \in \mathcal{X}$, $\mathbb{P}(X=x/\cdot)$ é chamada função de verossimilhança gerada por x.

$$V_x : \Theta \rightarrow \mathbb{R}$$
$$\theta \in \Theta \mapsto V_x(\theta) = \mathbb{P}(X=x/\theta)$$

Objetivo

$$f(\theta/x) = \frac{f(x/\theta)f(\theta)}{f(x)} \qquad \text{com } f(x) = \int_\Theta f(\theta,x)\,d\theta$$

Resultado

$X_1, X_2, \dots, X_n$, dado Θ, são i.i.d. $Ber(\theta)$.

$$\Theta = (0,1) \qquad \mathcal{X} = \{0,1\}^n$$

$\Theta \sim Beta(a,b)$, $a,b > 0$

$\Theta / X_1 = x_1, X_2 = x_2, \dots, X_n = x_n \sim ?$ $Beta\left(a + \sum_{i=1}^{n} x_i, b + n - \sum_{i=1}^{n} x_i\right)$

$$f(\theta / X_1 = x_1, \dots, X_n = x_n) = \frac{\mathbb{P}(X_1 = x_1, \dots, X_n = x_n / \theta)}{\mathbb{P}(X_1 = x_1, \dots, X_n = x_n)} f(\theta) =$$

$$= \frac{\mathbb{P}(X_1 = x_1, \dots, X_n = x_n / \theta)\, f(\theta)}{\int_0^1 \mathbb{P}(X_1 = x_1, \dots, X_n = x_n / \theta)\, f(\theta)\, d\theta} =$$

$$= \frac{\prod_{i=1}^{n} \mathbb{P}(X_i = x_i / \theta)\, f(\theta)}{\int_0^1 \left(\prod_{i=1}^{n} \mathbb{P}(X_i = x_i / \theta)\right) f(\theta)\, d\theta} = \frac{\prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1}\, \mathbb{I}_{(0,1)}(\theta)}{\int_0^1 \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1}\, d\theta} =$$

$$= \frac{\theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n - \sum_{i=1}^{n} x_i} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1}\, \mathbb{I}_{(0,1)}(\theta)}{\int_0^1 \theta^{\sum x_i} (1-\theta)^{n - \sum x_i} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1}\, d\theta} =$$

$$= \frac{\theta^{a + \sum_{i=1}^{n} x_i - 1} (1-\theta)^{b + n - \sum_{i=1}^{n} x_i - 1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, \mathbb{I}_{(0,1)}(\theta)}{\int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a + \sum x_i - 1} (1-\theta)^{b + n - \sum_{i=1}^{n} x_i - 1}\, d\theta} =$$

$$= \frac{\theta^{a + \sum_{i=1}^{n} x_i - 1} (1-\theta)^{b + n - \sum_{i=1}^{n} x_i - 1}\, \mathbb{I}_{(0,1)}(\theta)}{\int_0^1 \frac{\Gamma(a + \sum x_i + b + n - \sum x_i)}{\Gamma(a + \sum x_i)\Gamma(b + n - \sum x_i)} \theta^{a + \sum x_i - 1} (1-\theta)^{b + n - \sum x_i - 1}} \Rightarrow$$

$$f(\theta / X_1 = x_1, \dots, X_n = x_n) = \frac{\Gamma(a+b+n)}{\Gamma(a+\sum x_i)\Gamma(b+n-\sum x_i)} \theta^{a+\sum x_i} (1-\theta)^{b+n-\sum x_i - 1} \mathbb{I}_{(0,1)}(\theta)$$

Comentário

$\theta \sim Beta(1,4)$

$$Var(\theta) = \frac{1.4}{5.5.6} = \frac{4}{150} = \frac{2}{75} > 2\%$$

$n = ?$ $Var(\theta / X_1 = x_1, \dots, X_n = x_n) \leq 1/1000$, $\forall (x_1, \dots, x_n) \in \{0,1\}^n$

$$Var(\theta / X_1 = x_1, \dots, X_n = x_n) = \frac{\overbrace{(a+\sum x_i)(b+n-\sum x_i)}^{p(1-p) \quad p \in (0,1)}}{(a+b+n)(a+b+n)(a+b+n+1)} \leq \frac{1/4}{a+b+n+1}$$

$\leq \frac{1}{1000} \Rightarrow n \geq 250 - a - b - 1$ (Raiffa Schallaifer)

$$E(\theta) = \frac{a}{a+b}$$

$$E(\theta / X_1 = x_1, \dots, X_n = x_n) = \frac{a + \sum x_i}{a+b+n} = \frac{a + n\bar{x}}{a+b+n} =$$

$$= \frac{(a+b)}{a+b+n} \cdot \frac{a}{(a+b)} + \frac{n\bar{x}}{a+b+n} = \underbrace{\frac{(a+b)}{a+b+n}}_{0 \; n \to \infty} \cdot \frac{a}{a+b} + \underbrace{\frac{n}{a+b+n}}_{1 \text{ qdo } n \to \infty} \bar{x}$$

——— " ——— " ———

Espaço Paramétrico $\Theta = \mathbb{R}_+$ $\qquad \mathcal{X} = \mathbb{R}_+^n$ relativo a X_n

$X_1, X_2, \dots, X_n$ dado θ são c.i.i.d. $Exp(\theta)$.

$\theta \sim Gama(a,b)$, $a, b > 0$

$\theta / X_1 = x_1, X_2 = x_2, \dots, X_n = x_n \sim Gama(a+n, b + \sum_{i=1}^{n} x_i)$

$$f(\theta / X_1 = x_1, \dots, X_n = x_n) = \frac{f(x_1, \dots, x_n / \theta) f(\theta)}{f(x_1, \dots, x_n)} \propto f(x_1, \dots, x_n / \theta) f(\theta) \propto$$

$$\propto \prod_{i=1}^{n} f(x_i / \theta) f(\theta) \propto \prod_{i=1}^{n} \theta e^{-\theta x_i} \mathbb{I}_{\mathbb{R}_+}(x_i) \cdot \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \mathbb{I}_{\mathbb{R}_+}(\theta) \propto$$

$$\propto \theta^n e^{-\theta \sum x_i} \mathbb{I}_{\mathbb{R}_+}(\theta) \theta^{a-1} e^{-b\theta} \Rightarrow$$

$$f(\theta / x_1, \dots, x_n) \propto \theta^{n+a-1} e^{-(b + \sum x_i)\theta} \mathbb{I}_{\mathbb{R}_+}(\theta)$$

Resultado $\Theta = \mathbb{R}$ $\mathcal{X} = \mathbb{R}^n$

$X_1, X_2, \dots, X_n$, dado θ, são i.i.d. $N(\theta, \sigma_0^2)$
↳ conhecido

$$\Theta \sim N(a_0, b_0^2)$$
conhecidos

$\theta / X_1 = x_1, X_2 = x_2, \dots, X_n = x_n \sim$

$$f(\theta / x_1, \dots, x_n) = \frac{f(x_1, \dots, x_n / \theta) f(\theta)}{f(x_1, \dots, x_n)} \propto f(x_1, \dots, x_n / \theta) f(\theta) \propto$$

$$\propto \prod_{i=1}^{n} f(x_i / \theta) f(\theta) \propto \prod_{i=1}^{n} \frac{1}{\sigma_0 \sqrt{2\pi}} e^{-\frac{(x_i - \theta)^2}{2\sigma_0^2}} \cdot \frac{1}{b_0 \sqrt{2\pi}} e^{-\frac{(\theta - a_0)^2}{2b_0^2}} =$$

$$= \left(\frac{1}{\sigma_0 \sqrt{2\pi}} \right)^n e^{-\frac{\sum_{i=1}^{n} (x_i - \theta)^2}{2\sigma_0^2}} \cdot \frac{1}{b_0 \sqrt{2\pi}} e^{-\frac{(\theta - a_0)^2}{2b_0^2}} \propto$$

$$\propto e^{-\frac{1}{2} \left\{ \frac{\sum_{i=1}^{n} (x_i^2 - 2x_i\theta + \theta^2)}{\sigma_0^2} + \frac{\theta^2 - 2a_0\theta + a_0^2}{b_0^2} \right\}} \propto$$

$$\propto e^{-\frac{1}{2}\left[\frac{\sum x_i^2 - 2\theta\sum x_i + n\theta^2}{\sigma_0^2} + \frac{\theta^2 - 2a_0\theta + a_0^2}{b_0^2}\right]} \propto$$

$$\propto e^{-\frac{1}{2}\left\{\theta^2\left(\frac{n}{\sigma_0^2}+\frac{1}{b_0^2}\right) - 2\theta\left(\frac{\sum x_i}{\sigma_0^2} + \frac{a_0}{b_0^2}\right) + \frac{\sum x_i^2}{\sigma_0^2} + \frac{a_0^2}{b_0^2}\right\}} \propto$$

$$\propto e^{-\frac{1}{2}\left(\frac{n}{\sigma_0^2}+\frac{1}{b_0^2}\right)\left\{\theta^2 - 2\theta\frac{\left(\frac{\sum x_i}{\sigma_0^2}+\frac{a_0}{b_0^2}\right)}{\left(\frac{n}{\sigma_0^2}+\frac{1}{b_0^2}\right)} + \left[\frac{\left(\frac{\sum x_i}{\sigma_0^2}+\frac{a_0}{b_0^2}\right)}{\left(\frac{n}{\sigma_0^2}+\frac{1}{b_0^2}\right)}\right]^2\right\}} \Rightarrow$$

$$f(\theta / X_1 = x_1, \dots, X_n = x_n) \propto e^{-\frac{1}{2\left(\frac{n}{\sigma_0^2}+\frac{1}{b_0^2}\right)^{-1}}\left(\theta - \frac{b_0^2\sum x_i + \sigma_0^2 a_0}{nb_0^2+\sigma_0^2}\right)^2}$$

$$\theta / X_1 = x_1, \dots, X_n = x_n \sim N\left(\frac{b_0^2 n\bar{x} + \sigma_0^2 a_0}{nb_0^2+\sigma_0^2}, \left(\frac{n}{\sigma_0^2}+\frac{1}{b_0^2}\right)^{-1}\right)$$

——— " ——— " ———

$\theta = (\theta_1, \theta_2)$ $\quad \mathcal{X} = \mathbb{R}^n$ $\quad \Theta = \mathbb{R}\times\mathbb{R}_+$

$X_1, \dots, X_n$ dado θ, i.i.d. $N(\theta_1, 1/\theta_2)$

$\theta_2 \sim$ Gama (a,b), $a,b>0$

$\theta_1/\theta_2 \sim N(c, d/\theta_2)$, $c\in\mathbb{R}$, $d>0$

$f(\theta_1,\theta_2) = f(\theta_2) f(\theta_1/\theta_2)$

$(\theta_1,\theta_2) / X_1 = x_1, \dots, X_n = x_n \sim$

$$f(\theta_1,\theta_2/x_1,\dots,x_n) \propto f(x_1,\dots,x_n/\theta_1,\theta_2) f(\theta_1,\theta_2) \propto$$

$$\propto \prod_{i=1}^n f(x_i/\theta_1,\theta_2) f(\theta_1/\theta_2) f(\theta_2) \propto$$

$$\propto \prod_{i=1}^n \frac{1}{\sqrt{1/\theta_2}\sqrt{2\pi}} e^{-\frac{(x_i-\theta_1)^2}{2\frac{1}{\theta_2}}} \cdot \frac{1}{\sqrt{\frac{d}{\theta_2}}\sqrt{2\pi}} e^{-\frac{(\theta_1-c)^2}{2\frac{d}{\theta_2}}} \cdot \frac{b^a}{\Gamma(a)} \theta_2^{a-1} e^{-b\theta_2} \mathbb{I}_{\mathbb{R}_+}(\theta_2)$$

$$\propto \left(\prod_{i=1}^{n} \theta_2^{1/2}\right) e^{-\frac{\theta_2}{2}\sum_{i=1}^{n}(x_i-\theta_1)^2} \cdot \theta_2^{1/2}\, e^{-\frac{\theta_2(\theta_1-c)^2}{2d}}\, \theta_2^{a-1} e^{-b\theta_2} \propto$$

$$\propto \theta_2^{a+\frac{n+1}{2}-1} e^{-b\theta_2} \cdot e^{-\frac{\theta_2}{2}\left\{\sum_{i=1}^{n} x_i^2 - 2\theta_1 \sum x_i + n\theta_1^2\right\}} \cdot e^{-\left[\frac{\theta_2\theta_1^2}{2d} - \frac{2\theta_1\theta_2 c}{2d} + \frac{\theta_2 c^2}{2d}\right]}$$

$$\propto \theta_2^{a+\frac{n+1}{2}-1} e^{-\theta_2\left(b+\frac{\sum x_i^2}{2}+\frac{c^2}{2d}\right)} e^{-\frac{\theta_2}{2}\left\{\theta_1^2\left(n+\frac{1}{d}\right) - 2\theta_1\left(\sum x_i + \frac{c}{d}\right)\right\}}$$

$$\propto \theta_2^{a+\frac{n+1}{2}-1} e^{-\theta_2\left(b+\frac{\sum x_i^2}{2}+\frac{c^2}{2d}\right)} e^{-\frac{\theta_2}{2}\left(n+\frac{1}{d}\right)\left\{\theta_1^2 - 2\theta_1 \overbrace{\frac{(\sum x_i + c/d)}{(n+1/d)}}^{c'} + \left[\frac{\sum x_i + c/d}{n+1/d}\right]^2\right\}} \times$$

$$\times e^{\frac{\theta_2}{2}(n+1/d)\left[\frac{\sum x_i + c/d}{n+1/d}\right]^2}$$

$$\propto \underbrace{\theta_2^{a'-1} e^{-\theta_2 b'}}_{f(\theta_2/x_1,\ldots,x_n)} \underbrace{e^{-\frac{1}{2[\theta_2(n+1/d)]^{-1}}(\theta_1-c')^2}}_{f(\theta_1|\theta_2,x_1,\ldots,x_n)}$$

Prob. e Inf. Est. I

Luis Gustavo Esteves

10/05/2013 - Aula 3

Na aula passada.

$$f(\theta/x) = \frac{f(x/\theta)f(\theta)}{f(x)}$$

$f(\theta, x)$

Modelo Estatístico: $(\mathcal{X} \times \Theta, \sigma(\mathcal{X} \times \Theta), P)$

Optimal Statiscal Decisions (CAP-11) → ver sobre conjugações

$A \precsim B$ (relação de ordem dizendo que acredito mais em B do que A)

$\left.\begin{matrix} A \precsim B \\ A \precsim C \end{matrix}\right\} \Rightarrow \exists P$ (Capítulo 6 do Optimal Statistical Decision. Construção de medida de subjetividade.

→ Rosângela H. Loschi (1992)

→ S. French, D Insua Rios

→ Cadani

$\theta \sim Beta(a, b)$

$X_1, \ldots, X_n$ dado θ e i.i.d. $Ber(\theta)$

$\theta / X_1 = x_1, \ldots, X_n = x_n \sim Beta(a + \sum_{i=1}^{n} x_i, b + n - \sum_{i=1}^{n} x_i)$
↓

"Preserva a qualidade da inferência, tudo que vc precisa da amostra"
Suficiente

$X((X_1,\dots,X_n))$ amostra

Def. Qualquer transformação $T: \mathcal{X} \to \mathcal{T}$ mensurável é chamada Estatística.

Exemplo: $X=(X_1,\dots,X_n)$, X_i a valores reais, $i=1,\dots,n$ $(\mathcal{X}=\mathbb{R}^n)$.

$$T_1(X) = \frac{X_1+\dots+X_n}{n} = \bar{X}_n$$

$$T_2(X) = X_{(n)} = \max\{X_1,\dots,X_n\}$$

$$T: \mathcal{X} \to \mathbb{R}_+^2$$

$$(x_1,\dots,x_n) \mapsto T(x_1,\dots,x_n) = \left(\bar{x}, \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}\right)$$

$$T_3(x_1,\dots,x_n) = (x_1,\dots,x_n)$$

Seja $\mathcal{V} = \{ f: \Theta \to \mathbb{R} \}$

$$H: X \to \mathcal{V}$$

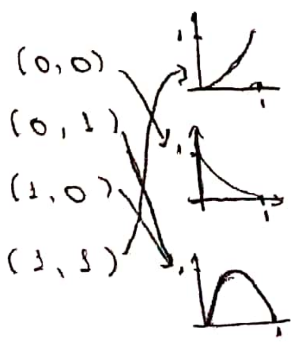
$$x \in \mathcal{X} \to H(x) = V_x(\cdot)$$

Exemplo.

X_1, X_2, dado Θ, e i.i.d. $Ber(\Theta)$.

$\mathcal{X} = \{0,1\}^2$

Para $(x_1,x_2) \in \mathcal{X}$



$$V_{(x_1,x_2)}(\Theta) = \Theta^{x_1+x_2}(1-\Theta)^{2-x_1-x_2}$$

Uma estatística suficiente é uma estatística que encerra toda a "informação" relevante da amostra para fazer inferência sobre θ !!

Def: $X((X_1,\dots,X_n))$. $T: \mathcal{X} \to \mathcal{J}$ é uma estatística suficiente para θ se para toda priori para θ,

$$\theta / X = x \overset{d}{=} \theta / T(X) = T(x) \quad , \forall x \in \mathcal{X}.$$

Exemplo 1:

$X_1,\dots,X_n$ dado θ e i.i.d. $Ber(\theta)$.

$\mathcal{X} = \{0,1\}^n \qquad \Theta = (0,1)$

$$f(\theta / x_1,\dots,x_n) \propto f(x_1,\dots,x_n / \theta) f(\theta)$$
$$\propto \prod_{i=1}^{n} P(X_i = x_i / \theta)\, f(\theta) \propto \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i} f(\theta)$$

$$f(\theta / \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} x_i) \propto P(\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} x_i / \theta)\, f(\theta)$$
$$\propto \binom{n}{\sum_{i=1}^{n} x_i} \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i} f(\theta)$$

$\forall x \in \mathcal{X}$

$\theta / X_1 = x_1, \dots, X_n = x_n \overset{d}{=} \theta / \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} x_i$. Logo $\sum_{i=1}^{n} X_i$ é suficiente para θ.

Exemplo 2.

$X_1,\dots,X_n$ dado θ com i.i.d. $Exp(\theta)$

$\mathcal{X} = \mathbb{R}_+^n \qquad \Theta = \mathbb{R}_+$

$$f(\theta/x_1,\dots,x_n) \propto f(x_1,\dots,x_n/\theta)f(\theta) \propto$$

$$\propto \left(\prod_{i=1}^{n} f(x_i/\theta)\right)\cdot f(\theta) \propto \left(\prod_{i=1}^{n} \theta e^{-\theta x_i} \mathbb{I}_{\mathbb{R}_+}(x_i)\right)\cdot f(\theta) \Rightarrow$$

$$\Rightarrow f(\theta/x_1,\dots,x_n) \propto \theta^n e^{-\theta\sum_{i=1}^{n} x_i} f(\theta).$$

$$T(X) = \sum_{i=1}^{n} X_i$$

$$f(\theta / T(X) = T(x_1,\dots,x_n)) = f(\theta / \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} x_i) \propto$$

$$\propto f_{\sum_{i=1}^{n} X_i}\left(\sum_{i=1}^{n} x_i / \theta\right).\ f(\theta) \propto \frac{\theta^n}{\Gamma(n)} \left(\sum_{i=1}^{n} x_i\right)^{n-1} e^{-\theta\sum_{i=1}^{n} x_i}\ \mathbb{I}_{\mathbb{R}_+}\left(\sum_{i=1}^{n} x_i\right) f(\theta)$$

$$\Rightarrow f(\theta/\sum_{i=1}^{n} x_i) \propto \theta^n e^{-\theta\sum_{i=1}^{n} x_i} f(\theta)$$

Como $\theta / X_1 = x_1,\dots,X_n = x_n \overset{d}{=} \theta / \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} x_i$, segue que $\sum_{i=1}^{n} X_i$ é suficiente para θ.

Exemplo 3.

$X_1,\dots,X_n$ dado θ, i.i.d. $\mathcal{U}(0,\theta)$, $\theta > 0$

$\mathcal{X} = \mathbb{R}_+^n \quad \Theta = \mathbb{R}_+$

$$f(\theta/x_1,\dots,x_n) \propto f(x_1,\dots,x_n/\theta)\ f(\theta) \propto$$

$$\propto \left(\prod_{i=1}^{n} f(x_i/\theta)\right) f(\theta) \propto \left(\prod_{i=1}^{n} \frac{1}{\theta} \mathbb{I}_{(0,\theta)}(x_i)\right) f(\theta) \propto$$

$$\propto \frac{1}{\theta^n} \mathbb{I}_{(0,x_{(n)})}(x_{(1)})\ \mathbb{I}_{(0,\theta)}(x_{(n)})\ f(\theta) \Rightarrow f(\theta/x_1,\dots,x_n) \propto \frac{1}{\theta^n} \mathbb{I}_{(0,\theta)}(x_{(n)})\ f(\theta)$$

$$\prod_{i=1}^{n} \mathbb{I}_{(0,\theta)}(x_i) = 1 \iff \mathbb{I}_{(0,\theta)}(x_i) = 1, \ \forall i = 1, \dots, n \iff$$

$$0 < x_i < \theta, \ \forall i = 1, \dots, n \iff 0 < x_{(1)} < x_{(n)} < \theta.$$

$$\iff \mathbb{I}_{(0,x_{(n)})}(x_{(1)}) \, \mathbb{I}_{(0,\theta)}(x_{(n)}) = 1$$

$X_{(n)} = \max\{X_1, \dots, X_n\}$

$$f(\theta / x_{(n)}) \propto f(x_{(n)} / \theta) f(\theta)$$

$$\propto \frac{n (x_{(n)})^{n-1}}{\theta^n} \mathbb{I}_{(0,\theta)}(x_{(n)}) \, f(\theta)$$

$$\propto \frac{1}{\theta^n} \mathbb{I}_{(0,\theta)}(x_{(n)}) \, f(\theta)$$

$\forall x \in \mathcal{X}$

Logo, $\theta / X_1 = x_1, \dots, X_n = x_n \overset{d}{=} \theta / X_{(n)} = x_{(n)}$

Assim, $T(X) = X_{(n)}$ é suficiente para θ.

$$P(\theta = s / X = x) = P(\theta = s \mid T(X) = T(x)) \iff$$

$$\frac{P(X = x / \theta = s) \cancel{P(\theta = s)}}{P(X = x)} = \frac{P(T(X) = T(x) / \theta = s) \cancel{P(\theta = s)}}{P(T(X) = T(x))} \iff$$

$$\frac{P(X = x / \theta = s)}{P(X = x)} = \frac{P(T(X) = T(x) / \theta = s)}{P(T(X) = T(x))} \iff$$

$$\frac{P(X = x / \theta = s)}{P(T(X) = T(x) / \theta = s)} = \frac{P(X = x)}{P(T(X) = T(x))} \iff$$

$$\frac{\mathbb{P}(X=x, T(X)=T(x) / \Theta=\jmath)}{\mathbb{P}(T(X)=T(x) / \Theta=\jmath)} = \frac{\mathbb{P}(X=x)}{\mathbb{P}(T(X)=T(x))} \Leftrightarrow$$

$$\mathbb{P}(X=x / T(X)=T(x), \Theta=\jmath) = \frac{\mathbb{P}(X=x)}{\mathbb{P}(T(X)=T(x))}$$ → Não depende de Θ

* Mostrou que a definição de suficiência bayesiana equivale à da clássica por meio do caso discreto.

Definição. $T: \mathcal{X} \to \mathcal{J}$ é suficiente para Θ se a distribuição de $X / T(X) = t, \Theta$ não depende de Θ!! para todo $t \in T(\mathcal{X})$ e $\forall \Theta \in \Theta$.

Suficiência Clássica $\Rightarrow$ Suficiência Bayesiana

Pra casos contínuos e discretos, a volta é válida. Entretanto, existem casos em que a volta não é válida. (Sh...)

Exemplo. X_1, X_2 dado Θ, i.i.d. Ber(Θ)

$$T(X) = X_1 + X_2$$

$$\mathbb{P}(X_1 = x_1, X_2 = x_2 / T(X) = t, \Theta) =$$

$$\frac{\mathbb{P}(X_1 = x_1, X_2 = x_2, T(X) = t / \Theta)}{\mathbb{P}(T(X) = t / \Theta)} = \begin{cases} 0, & t \neq x_1 + x_2 \\ \dfrac{\Theta^{x_1+x_2}(1-\Theta)^{2-x_1-x_2}}{\binom{2}{x_1+x_2}\Theta^{x_1+x_2}(1-\Theta)^{2-x_1-x_2}} = \dfrac{1}{\binom{2}{x_1+x_2}} = \dfrac{1}{\binom{2}{t}} \end{cases}$$

Logo, $T(X) = X_1 + X_2$ é suficiente para Θ.

$S(X) = X_1 \cdot X_2$

$S(X) / \Theta \sim Ber(\theta^2)$

$$\mathbb{P}(X_1 = x_1, X_2 = x_2 / S(X) = 0, \Theta) = \frac{\mathbb{P}(X_1 = x_1, X_2 = x_2, S(X) = 0 / \Theta)}{\mathbb{P}(S(X) = 0 / \Theta)} = \begin{cases} 0, & x_1 \cdot x_2 \neq 0 \\ \dfrac{(1-\Theta)^2}{1-\Theta^2} = \dfrac{(1-\Theta)(1-\Theta)}{(1-\Theta)(1+\Theta)} = \dfrac{(1-\Theta)}{1+\Theta}, & x_1 \cdot x_2 = 0 \end{cases}$$

$\forall\, x_1 = 0,\ x_2 = 0$

que depende de Θ.

Logo, $S(X)$ NÃO é suficiente para θ.

$\Theta / X_1 = 0, X_2 = 0$

$$f(\Theta / X_1 = 0, X_2 = 0) \propto P(X_1 = 0, X_2 = 0 / \Theta) f(\Theta) \propto (1-\Theta)^2 f(\Theta)$$

$$\#$$

$\Theta / S(X) = 0$

$$f(\Theta / S(X) = 0) \propto P(S(X) = 0 / \Theta) f(\Theta) \propto (1-\Theta^2) f(\Theta)$$

Critério Da Fatoração (Fisher-Neyman)

No livro do Roussas, uma dem. mais geral.

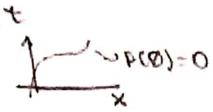
$X((X_1, \dots, X_n))$ amostra

A estatística $T: \mathcal{X} \to \mathcal{J}$ é suficiente para Θ se e só se $\exists\, \mu: \mathcal{X} \to \mathbb{R}_+$ e $v: \mathcal{J} \times \Theta \to \mathbb{R}_+$ tais que

$$f(x / \Theta) = \mu(x)\, v(T(x), \Theta), \ \forall x, \forall \Theta.$$

$(\Rightarrow)$

$$P(X=x/\theta) = P(X=x, T(X)=T(x)/\theta) = \underbrace{P(X=x \mid T(X)=T(x),\theta)}_{\mu(x)}\ \underbrace{P(T(X)=T(x),\theta)}_{v(T(x),\theta)}$$

A ida pro caso contínuo 

$(\Leftarrow)$

$$P(X=x \mid T(X)=t,\theta) = \frac{P(X=x, T(X)=t/\theta)}{P(T(X)=t/\theta)} = \begin{cases} 0, & t \neq T(x) \\ \dfrac{P(X=x/\theta)}{\sum\limits_{y \in \mathcal{X}:\,T(y)=t} P(X=y/\theta)}, & t = T(x) \ * \end{cases}$$

$$* \ \frac{P(X=x/\theta)}{\sum\limits_{y:T(y)=T(x)} P(X=y/\theta)} = \frac{\mu(x)\,v(T(x),\theta)}{\sum\limits_{y:T(y)=T(x)} \mu(y)\,v(T(x),\theta)} = \frac{\mu(x)}{\sum\limits_{y:T(y)=t} \mu(y)}$$ que não depende de θ

Logo, T é suficiente!

Exemplos:

1. $X_1,\dots,X_n$, dado θ, i.i.d. $N(\theta,1)$, $\theta \in \mathbb{R}$

$$f(x_1,\dots,x_n/\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\theta)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\theta)^2}{2}} =$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum x_i^2 - 2\theta\sum x_i + n\theta^2}{2}} = \underbrace{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^{n} x_i^2}{2}}}_{\mu(x)}\ \underbrace{e^{\frac{+2\theta\sum x_i - n\theta^2}{2}}}_{v(\sum x_i,\theta)}$$

Pelo critério da fatoração

$T(X) = X_1 + \dots + X_n$ é suficiente para θ.

Ex 2: $X_1, \ldots, X_n$, dado θ, i.i.d $N(1, \theta)$, $\theta > 0$

$$f(x_1, \ldots, x_n / \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x_i-1)^2}{2\theta}} = \underbrace{\left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \cdot e^{-\frac{1}{2\theta}\sum_{i=1}^{n}(x_i-1)^2}}_{v\left(\sum_{i=1}^{n}(x_i-1)^2, \theta\right)} \cdot \underbrace{1}_{h(x)}$$

Pelo critério da Fatoração, $S(X) = \sum_{i=1}^{n}(X_i-1)^2$ é suficiente para θ.

Exemplo 3:

$X_1, \ldots, X_n$, dado $\theta = (\theta_1, \theta_2)$, i.i.d. $N(\theta_1, \theta_2)$, $\theta_1 \in \mathbb{R}$; $\theta_2 > 0$

$\mathcal{X} = \mathbb{R}^n$ $\quad \Theta = \mathbb{R} \times \mathbb{R}$

$$f(x_1, \ldots, x_n / \theta) = \left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\theta_1)^2}{2\theta_2}} =$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n e^{-\frac{\sum x_i^2 - 2\theta_1 \sum x_i + n\theta_1^2}{2\theta_2}}}_{v\left((\sum x_i, \sum x_i^2), \theta\right)} \cdot \underbrace{1}_{h(x)}$$

Pelo critério de fatoração,

$T(X) = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$ é suficiente para θ.

$\downarrow$

$(X_1, \ldots, X_n)$

Exemplo 4.

$X_1, \ldots, X_n$, dado θ, condicionalmente independentes tais que

$X_i / \theta \sim N(2, i\theta)$, $i = 1, \ldots, n$

$$f(x_1,\dots,x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi i \theta}}\, e^{-\frac{(x_i-2)^2}{2i\theta}} =$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \frac{1}{\sqrt{n!}}\, e^{-\frac{1}{2\theta}\sum_{i=1}^{n}\frac{(x_i-2)^2}{i}}}_{g\left(\sum_{i=1}^{n}\frac{(x_i-2)^2}{i},\ \theta\right)} \cdot \underbrace{1}_{h(x)}$$

Logo, pelo Critério de Fatoração

$$T(X) = \sum_{i=1}^{n} \frac{(X_i-2)^2}{i}$$ é suficiente para θ.

Aula 04 - 14/05

- Comparative Statistics Inference (Vic. Barnett)

24/05 às 11:00hs
Sala 252
Seminário de fundamentos

$\Theta / X = x \overset{d}{=} \Theta / T(X) = T(x), \forall x$

<u>Critério da Fatoração.</u>

Seguindo com os exemplos...

5) $X(X_1, \ldots, X_n)$ amostra

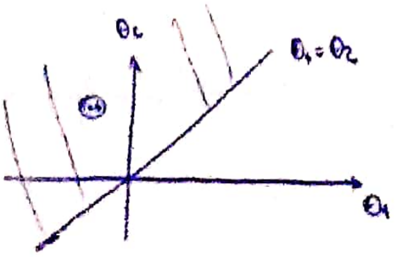
$$f(x|\theta) = \underbrace{f(x|\theta)}_{v(x,\theta)} \cdot \underbrace{1}_{\mu(x)}$$

$T: \mathcal{X} \to \mathcal{X}$ dada por

$T(X) = X$ é suficiente para Θ.

6) $X_1, \ldots, X_n$ dado $\Theta = (\Theta_1, \Theta_2)$, i.i.d. $\mathcal{U}(\Theta_1, \Theta_2)$

$\Theta = \{(a,b) \in \mathbb{R}^2 : a < b\}$ $\mathcal{X} = \mathbb{R}^n$



$(x_1, \ldots, x_n) \in \mathcal{X}$

$$f(x_1, \ldots, x_n | \theta) = \prod_{i=1}^{n} f(x_i | \theta) = \prod_{i=1}^{n} \frac{1}{\theta_2 - \theta_1} \mathbb{I}_{(\theta_1, \theta_2)}(x_i) =$$

$$= \underbrace{\left(\frac{1}{\theta_2 - \theta_1}\right)^n \prod_{i=1}^{n} \mathbb{I}_{(\theta_1, \theta_2)}(x_i) = \left(\frac{1}{\theta_2 - \theta_1}\right)^n \mathbb{I}_{(\theta_1, x_{(n)})}(x_{(1)}) \, \mathbb{I}_{(\theta_1, \theta_2)}(x_{(n)})}_{v((x_{(1)}, x_{(n)}), \Theta) \; (\theta_1, \theta_2)} \cdot \underbrace{1}_{\mu(x)}$$

Pelo Critério da Fatoração,

$$T(X) = (X_{(1)}, X_{(n)}) \text{ é suficiente para } \theta.$$

Resultado

Seja $T: \mathcal{X} \to \mathcal{J}$ suficiente para Θ

Seja $S: \mathcal{X} \to \mathcal{S}$ outra estatística tal que $S = g \circ T$, onde $g: \mathcal{J} \to \mathcal{S}$ é bijetora (um a um). Então, S é suficiente para Θ.

$S = g(T)$

Existe $g^{-1}: \mathcal{S} \to \mathcal{J}$

Como T é suficiente, então

$$f(x|\theta) = \mu(x).v(T(x),\theta) = \mu(x).v(g^{-1}(S(x)),\theta).$$

Definindo $v': \mathcal{S} \times \Theta \to \mathbb{R}_+$ por $v'(s,\theta) = v(g^{-1}(s),\theta)$, resulta que.

Assim,

$f(x/\theta) = \mu(x).v'(S(x),\theta)$. Logo, pelo Critério da Fatoração, S é suficiente para Θ.

Exemplo 1.

$X_1, \ldots, X_n$ dado θ, c.i.i.d. $N(\theta,1)$

Vimos que $T(X) = \sum_{i=1}^{n} X_i$ é suficiente para θ.

Definindo $S(X) = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$, é fácil ver que $S = g \circ T$, onde $g: \mathbb{R} \to \mathbb{R}$ é dada por $g(u) = \frac{u}{n}$ (inversível)

Logo, $S(X) = \bar{X}_n$ é suficiente para θ.

Exemplo 2:

$X_1, \dots, X_n$ dado $\Theta = (\Theta_1, \Theta_2)$, c.i.i.d.

$N(\Theta_1, \Theta_2)$, $(\Theta = \mathbb{R} \times \mathbb{R}_+)$

$T(X) = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2 \right)$ é suficiente para Θ.

$S(X) = \left(\bar{X}_n, \dfrac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n} \right)$ $\to \dfrac{\sum_{i=1}^{n} X_i^2 - n(\bar{X}_n)^2}{n}$

Seja $g: \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R} \times \mathbb{R}_+$ dada por

$$g(a,b) = \left(\frac{a}{n}, \frac{b - n\left(\frac{a}{n}\right)^2}{n} \right)$$

g é bijetora. Logo,

$$S(X) = \left(\bar{X}, \frac{\sum (X_i - \bar{X}_n)^2}{n} \right) = g\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2 \right) \text{ é suficiente para } \Theta.$$

Exercício.

$(X_{(1)}, X_{(n)})$ $\quad \left(\dfrac{X_{(1)} + X_{(n)}}{2}, X_{(n)} - X_{(1)} \right)$ $\quad S(X) = \left(\dfrac{X_{(1)} + X_{(n)}}{2}, X_{(n)} - X_{(1)} \right) = g(X_{(1)}, X_{(n)})$

ond $g: \mathbb{R} \to \mathbb{R}$ com $g(a,b) = \left(\dfrac{a+b}{2}, b-a \right)$

Estatística Suf. mínima → "Aquela que faz a maior redução da amostra de modo a preservar a suficiência". "Se tentar reduzir ainda mais a E.S.M., perdemos a suficiência".

$T: \mathcal{X} \to \mathcal{J}$ é suficiente

Seja $T': \mathcal{X} \to \mathcal{J}'$ uma outra estatística suficiente.

Dizemos que $T: \mathcal{X} \to \mathcal{J}$ é suficiente mínima para θ se T é função de qualquer outra estatística suficiente para θ.

Em outras palavras, T é suficiente mínima se para todo T' existe $g: \mathcal{J}' \to \mathcal{J}$ tal que $T = g \circ T'$.

Ou ainda, T é suficiente mínima se para toda T' suficiente vale que

$\forall x_1, x_2 \in \mathcal{X}$,

$$T'(x_1) = T'(x_2) \Rightarrow T(x_1) = T(x_2)$$

Resultado:

Seja $T: \mathcal{X} \to \mathcal{J}$ uma estatística.

Se $\forall x_1, x_2 \in \mathcal{X}$,

$$\frac{f(x_1 | \theta)}{f(x_2 | \theta)} \text{ não depende de } \theta \Leftrightarrow T(x_1) = T(x_2), \text{ então}$$

T é suficiente mínima.

1. T é suficiente

$$P(X = x | T(X) = t, \theta) \overset{t = T(x)}{=} \frac{\mathbb{P}(X = x, T(X) = T(x) | \theta)}{\mathbb{P}(T(X) = T(x) | \theta)} = \frac{\mathbb{P}(X = x | \theta)}{\sum_{y \in \mathcal{X}: T(y) = T(x)} \mathbb{P}(X = x | \theta)} =$$

$$= \frac{1}{\sum_{y \in \mathcal{X}: T(x) = T(y)} \frac{\mathbb{P}(X = y | \theta)}{\mathbb{P}(X = x | \theta)}} \text{ não depende de } \theta!$$

Pois $T(x_1) = T(y) \Rightarrow \frac{f(x_1 | \theta)}{f(x_2 | \theta)}$ ñ dep de θ

Como a dist. de X dado $T(X)=t$ e Θ não depende do Θ, $\forall t$, então

T é suficiente para Θ.

2. Seja $T': \mathcal{X} \to \mathcal{J}$ uma estatística suficiente.

Sejam $x_1, x_2 \in \mathcal{X}$.

$T'(x_1) = T'(x_2) \Rightarrow$

$$\frac{\mathbb{P}(X_1 = x_1 | \theta)}{\mathbb{P}(X_2 = x_2 | \theta)} = \frac{\mu(x_1)\, v(T'(x_1), \theta)}{\mu(x_2)\, v(T'(x_2), \theta)} = \frac{\mu(x_1)}{\mu(x_2)} \Rightarrow$$

$$\frac{\mathbb{P}(X_1 = x_1 | \theta)}{\mathbb{P}(X_2 = x_2 | \theta)} \text{ não depende de } \Theta \Rightarrow T(x_1) = T(x_2).$$

Logo, T é suficiente mínima.

Exemplos:

(1) $X_1, \dots, X_n$, dado Θ, c.i.i.d. $Beta(\theta, 1)$
condicionalmente

$\theta > 0$. $\mathcal{X} = (0,1)^n$, $\Theta = \mathbb{R}_+$

Sejam $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathcal{X}$

$$\frac{f(x_1, \dots, x_n | \theta)}{f(y_1, \dots, y_n | \theta)} = \frac{\prod_{i=1}^n f(x_i | \theta)}{\prod_{i=1}^n f(y_i | \theta)} = \frac{\prod_{i=1}^n \theta x_i^{\theta-1} \mathbb{I}_{(0,1)}(x_i)}{\prod_{i=1}^n \left\{ \theta y_i^{\theta-1} \mathbb{I}_{(0,1)}(y_i) \right\}} =$$

$$= \frac{\theta^n \left(\prod_{i=1}^n x_i\right)^{\theta-1} \prod_{i=1}^n \mathbb{I}_{(0,1)}(x_i)}{\theta^n \left(\prod_{i=1}^n y_i\right)^{\theta-1} \prod_{i=1}^n \mathbb{I}_{(0,1)}(y_i)} = \left(\frac{\prod_{i=1}^n x_i}{\prod_{i=1}^n y_i}\right)^{\theta-1} \frac{\prod_{i=1}^n \mathbb{I}_{(0,1)}(x_i)}{\prod_{i=1}^n \mathbb{I}_{(0,1)}(y_i)} \Leftrightarrow$$

$$\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} y_i$$

$$\Updownarrow$$

$T(x_1, \ldots, x_n) = T(y_1, \ldots, y_n)$, onde $T(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i$.

Logo, $T(X) = \prod_{i=1}^{n} X_i$ é suficiente mínima para θ.

Exemplo 2. $X_1, \ldots, X_n$ dado $\theta = (\theta_1, \theta_2)$, c.i.i.d. $N(\theta_1, \theta_2)$. $\mathcal{X} = \mathbb{R}^n$, $\Theta = \mathbb{R} \times \mathbb{R}_+$

$(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathcal{X}$.

$$\frac{f(x_1, \ldots, x_n / \theta)}{f(y_1, \ldots, y_n / \theta)} = \frac{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_i - \theta_1)^2}{2\theta_2}}}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(y_i - \theta_1)^2}{2\theta_2}}} =$$

$$= \frac{\left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i - \theta_1)^2}{2\theta_2}}}{\left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n e^{-\frac{\sum_{i=1}^{n}(y_i - \theta_1)^2}{2\theta_2}}} = e^{\frac{\sum_{i=1}^{n} y_i^2 - 2\theta_1 \sum_{i=1}^{n} y_i + n\theta_1^2 - \sum_{i=1}^{n} x_i^2 + 2\theta_1 \sum_{i=1}^{n} x_i - n\theta_1^2}{2\theta_2}}$$

$$= e^{\frac{2\theta_1 \left(\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i\right) + \left(\sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} x_i^2\right)}{2\theta_2}}$$ não depende de θ

$$\Updownarrow$$

$$\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i = 0 \quad \text{e} \quad \sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} x_i^2 = 0,$$

$$\Updownarrow$$

$T(x_1, \ldots, x_n) = T(y_1, \ldots, y_n)$, onde

$$T(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^2\right)$$

Pelo resultado,

$$T(X) = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right) \text{ é suficiente mínima para } \theta.$$

ESTIMAÇÃO

θ brancas
$5-\theta$ verdes

$\Theta = \{0,1,2,3,4,5\}$

$X = (X_1, X_2)$

$\mathcal{X} = \{0,1\}^2$

Uma estatística $T: \mathcal{X} \to \Theta$ é chamado Estimador para Θ.

$\delta_1 : \mathcal{X} \to \Theta$

$$\delta_1(x_1,x_2) = \begin{cases} 0, & x_1 = x_2 = 0 \\ 2, & x_1 = 1, x_2 = 0 \\ 3, & x_1 = 0, x_2 = 1 \\ 4, & x_1 = 1, x_2 = 1 \end{cases}$$

$\delta_2 : \mathcal{X} \to \Theta$

$$\delta_2(x_1,x_2) = \begin{cases} 1, & x_1 = x_2 = 0 \\ 3, & x_1 = 1, x_2 = 0 \\ 3, & x_1 = 0, x_2 = 1 \\ 5, & x_1 = x_2 = 1 \end{cases}$$

$\delta_3(x_1,x_2) = 2, \ \forall (x_1,x_2) \in \mathcal{X}$

$\delta_4(x_1,x_2) = 0, \ \forall (x_1,x_2) \in \mathcal{X}$

$\delta_i, i = 1,2,3,4$ são ~~os~~ estimadores!

Dado
$(0,1)$, $\delta_1(0,1) = 3$ é a estimativa para Θ

Usualmente,

$\delta(X)$ - estimador.

$\delta(x)$ - estimativa.

Em geral, $\Theta \subseteq \mathbb{R}^k$.

$$\Delta = \Big\{ \delta : \overset{\mathfrak{X}}{\{0,1\}^2} \rightarrow \overset{\Theta}{\{0,1,2,3,4,5\}} \Big\}$$

Δ, conjunto dos estimadores para Θ (nesse exemplo), possui $\underset{1296}{\underset{\parallel}{6^4}}$ elementos.

Como escolher o estimador nesse caso?

Devemos estabelecer critérios de otimalidade.

$$|\delta(x) - \theta| \quad \text{ou} \quad |\delta(x) - \theta|^2 \quad \text{ou} \quad \left| \frac{\delta(x) - \theta}{\theta} \right|$$

dependem de θ!! Logo, não podemos concluir nada, por enquanto.

Ideia: Escolher um estimador que minimize os critérios acima, em média.

Defina o critério de otimalidade (penalidade), devemos escolher $\delta \in \Delta$ que minimize a penalidade média.

No nosso exemplo para $\delta \in \Delta$, vamos avaliar

$$\underset{\substack{\downarrow \\ \text{risco}}}{\rho(\delta)} = E\big[L(\delta(X), \Theta)\big] = \sum_{x \in \mathfrak{X}} \sum_{j=0}^{5} L(\delta(x), j)\, P(X = x, \Theta = j)$$

$L: \Theta \times \Theta \rightarrow \mathbb{R}_+$

$(\delta(x), \theta) \longmapsto L(\delta(x), \theta)$: penalidade (perda) que o decisor incorre ao estimar $\delta(x)$ ao observar $x \in \mathfrak{X}$.

Ideia. Escolher $\delta^* \in \Delta$ tal que

$$\rho(\delta^*) = \min\{\rho(\delta) : \delta \in \Delta\}$$

Podemos reescrever $\rho(\delta)$ da seguinte maneira

$$\rho(\delta) = \sum_{x \in \mathcal{X}} \sum_{j=0}^{5} L(\delta(x), \theta)\, \mathbb{P}(\theta = j / X = x)\, \mathbb{P}(X = x) =$$

$$= \sum_{x \in \mathcal{X}} \mathbb{P}(X = x) \left\{ \left[\sum_{j=0}^{5} L(\delta(x), \theta)\, \mathbb{P}(\theta = j / X = x) \right] \right\}$$

Idéia:

Para cada $x \in \mathcal{X}$, tomamos $\delta^*(x) \in \Theta$ que minimize (em d)

$$\sum_{j=0}^{5} L(d, j)\, \mathbb{P}(\theta = j / X = x)$$

Ao estimador que associa a cada $x \in \mathcal{X} \cdot \delta^*(x)$ conforme especificado ao lado, damos o nome de ESTIMADOR DE BAYES para θ considerando a perda L.

Podemos reescrever $\rho(\delta)$ da seguinte maneira

$$\rho(\delta) = \sum_{x \in \mathcal{X}} \sum_{j=0}^{5} L(\delta(x), \theta)\, P(\theta = j \mid X = x)\, P(X = x) =$$

$$= \sum_{x \in \mathcal{X}} P(X=x) \left\{ \sum_{j=0}^{5} L(\delta(x), \theta)\, P(\theta = j \mid X = x) \right\}$$ (+ posteriori (média))

Ideia: Para cada $x \in \mathcal{X}$, tomamos $\delta^*(x) \in \Theta$ que minimize

$$\sum_{j=0}^{5} L(d, j)\, \mathbb{P}(\theta = j \mid X = x).$$

Ao estimador que associa a cada $x \in \mathcal{X}$ $\delta^*(x)$ conforme especificado acima, damos o nome de Estimador de Bayes para θ considerando a perda L.

Aula 5 - 15/05/2013

Revisão última aula: $\mathcal{X}$, Θ, (θ, x), $\Delta = \{\delta : \mathcal{X} \to \Theta\}$

$\delta : \mathcal{X} \to \Theta$

Para cada $\delta \in \Delta$, avaliamos

$$\rho(\delta) = E[L(\delta(X), \theta)] = \sum_{x \in \mathcal{X}} \sum_{j \in \Theta} L(\delta(x), j)\, \mathbb{P}(X = x, \theta = j)$$

$$\Rightarrow \underbrace{\rho(\delta)}_{\text{Risco do Estimador}} = \sum_{x \in \mathcal{X}} \mathbb{P}(X=x) \left\{ \sum_{j \in \Theta} L(\delta(x), j)\, \mathbb{P}(\theta = j \mid X = x) \right\}$$

Considere $\delta^*: \mathcal{X} \to \Theta$ que associa a cada $x \in \mathcal{X}$, o valor $d^* \in \Theta$ tal que

$$\sum_{j \in \Theta} L(d^*, j)\, P(\Theta = j \mid X = x) = \min_{d \in \Theta} \underbrace{\sum_{j \in \Theta} \overset{h_d(j)}{\overbrace{L(d, j)}}\, P(\Theta = j \mid X = x)}_{E[L(d, \Theta) \mid X = x]}$$

isto é, $\delta^*(x) = d^*(d^*_x)$ → depende da posteriori

É fácil ver que $\rho(\delta^*) \leq \rho(\delta), \forall \delta \in \Delta$ e é chamado ESTIMADOR DE BAYES e seu risco é chamado risco de Bayes.

Com efeito,

$$\rho(\delta^*) = \sum_{x \in \mathcal{X}} P(X = x) \left\{ \sum_{j \in \Theta} L(\overset{d^*}{\delta^*(x)}, j)\, P(\Theta = j \mid X = x) \right\}$$

$$\leq \sum_{x \in \mathcal{X}} P(X = x) \left\{ \sum_{j \in \Theta} L(d(x), j)\, P(\Theta = j \mid X = x) \right\} = \rho(\delta), \forall \delta \in \Delta.$$

"Para cada $x \in \mathcal{X}$ construí-se a dist. a posteriori e encontramos o valor de d que minimiza os riscos a posteriori"

Principais funções de perda:

(1) Perda Quadrática: $L(d, \Theta) = (\overset{\delta(x)}{d} - \Theta)^2$

(2) Perda Absoluta: $L(d, \Theta) = |d - \Theta|$

(3) Perda 0-1

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

Para cada $x\in\mathcal{X}$, devemos achar d^* tq.

$$E[L(d^*,\theta)|X=x]=\min_{d\in\Theta} E[L(d,\theta)|X=x]$$

No caso de perda quadrática, temos

$$E[L(d,\theta)|X=x]=E[(d-\theta)^2|X=x]=E[(d-E(\theta|X=x)+E(\theta|X=x)-$$

$$\theta)^2|X=x]=E[(d-E(\theta|X=x))^2|X=x]+2E[(d-E(\theta|X=x))(E(\theta|X=x)-\theta)|X=x]$$

$$+E[(\theta-E(\theta|X=x))^2|X=x]$$

$=(d-E(\theta|X=x))^2+Var(\theta|X=x)$, que é minimizado por $E(\theta|X=x)$.

Logo, $d^*=E(\theta|X=x)$.

Assim, se $E(\theta|X=x)\in\Theta$, tomamos $\delta^*(x)=E[\theta|X=x]$.

Caso contrário, tomamos $d^*\in\Theta$ tal que:

$$|d^*-E(\theta|X=x)|=\min_{d\in\Theta}|d-E(\theta|X=x)|$$

No exemplo da urna,

$\delta^*:\{0,1\}^2\longrightarrow\{0,1,2,3,4,5\}$

Vimos que, $P(\Theta = j \mid X = (0,1))$ $=$

j	$P(\Theta=j \mid X=(0,0))$ *	j	$P(\Theta=j \mid X=(1,0))$ **
0	20/40	0	0
1	12/40	1	4/20
2	6/40	2	6/20
3	2/40	3	4/20
4	0	4	4/20
5	0	5	0
	1		1

(*) $E[\Theta \mid X=(0,0)] = 30/40 = 3/4 \therefore \delta^*(0,0) = 1$ (**) $E[\Theta \mid X=(1,0)] = 2{,}5 \Rightarrow \delta^*((1,0)) = 2$

j	$P(\Theta=j \mid X=(1,1))$ ***
0	0
1	0
2	2/40
3	6/40
4	12/40
5	20/40

(***) $E(\Theta \mid X=(1,1)) = \frac{17}{4} \Rightarrow \delta^*((1,1)) = 4$ (Estimativa de Bayes)

Logo, um estimador de Bayes para Θ é, com relação a perda quadrática,

$$\delta_1^*(x) = \begin{cases} 1, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } x=(1,0) \\ 4, & x=(1,1) \end{cases}$$

$$\delta_2^*(x) = \begin{cases} 1, & x=(0,0) \\ 2, & x=(0,1) \\ 3, & x=(1,0) \\ 4, & x=(1,1) \end{cases}$$

Existem + 2 estimadores de Bayes $\delta_1^*(x)$ e $\delta_2^*(x)$

Ex 2 $\Theta = (0,1)$ $\quad \mathcal{X} = \{0,1\}^n$

$X_1, \ldots, X_n$, dado θ, c.i.i.d $Ber(\theta)$ (?)

$\theta \sim Ber(a,b)$, $a,b > 0$

Vimos que para cada $x = (x_1, \ldots, x_n) \in \mathcal{X}$,

$$\theta | X = x \sim Beta\left(a + \sum_{i=1}^{n} x_i,\ b + n - \sum_{i=1}^{n} x_i\right)$$

$$E(\theta | X = x) = \frac{a + \sum_{i=1}^{n} x_i}{a+b+n} \in \Theta,\ \forall x \in \mathcal{X}$$

é único, neste caso

Assim, o estimador de Bayes com relação à perda quadrática é dado por

$$\delta^*(X_1, \ldots, X_n) = \frac{a + \sum_{i=1}^{n} X_i}{a+b+n}. \qquad \left(\frac{a+n\bar{x}}{a+b+n} = \frac{a+b}{a+b+n}\frac{a}{a+b} + \frac{n}{a+b+n}\bar{x}_n\right)$$

Contínuo

Ex 3 $X_1, \ldots, X_n$, dado θ, c.i.i.d $N(\theta, \sigma_0^2)$

$\theta \sim N(a_0, b_0^2)$, $\quad \mathcal{X} = \mathbb{R}^n$, $\quad \Theta = \mathbb{R}$

$$\theta | X_1 = x_1, \ldots, X_n = x_n \sim N\left(\frac{b_0^2 \sum_{i=1}^{n} x_i + \sigma_0^2 a_0}{n b_0^2 + \sigma_0^2},\ ?\right)$$

$$E(\theta | X_1 = x_1, \ldots, X_n = x_n) = \frac{b_0^2 \sum_{i=1}^{n} x_i + \sigma_0^2 a_0}{n b_0^2 + \sigma_0^2}.$$

Logo, o estimador de Bayes para θ com relação à perda quadrática é:

$$\delta^*(X_1, \ldots, X_n) = \frac{b_0^2 \sum_{i=1}^{n} X_i + \sigma_0^2 a_0}{n b_0^2 + \sigma_0^2}$$

(2) Perda Absoluta: $L(d,\Theta) = |d-\Theta|$

Dizemos que m é uma Mediana da v.a X (da distribuição da v.a X) se

$$P(X \leq m) \geq 1/2 \quad e \quad P(X \geq m) \geq 1/2.$$

Para $x \in \mathcal{X}$, seja m uma mediana da distribuição condicional de Θ dado $X = x$.

Vamos avaliar para $d > m$.

$$E[L(d,\Theta)|X=x] - E[L(m,\Theta)|X=x]$$

$$= E[|d-\Theta| \,|X=x] - E[|m-\Theta| \,|X=x] = \left(\text{Suposição } E(|\Theta|) < \infty\right)$$

$$= E[|d-\Theta| - |m-\Theta| \,|X=x] = \sum_{\theta \in \Theta} (|d-j| - |m-j|) \cdot P(\Theta = j|X=x).$$

$$= \sum_{j \leq m} (d-m) P(\Theta = j|X=x) + \sum_{m<j\leq d} [(d-j)-(j-m)] P(\Theta=j|X=x)$$

$$+ \sum_{j>d} (-1)(d-m) P(\Theta=j|X=x)$$

$$\geq \sum_{j \leq m} (d-m) P(\Theta=j|X=x) + \sum_{m<j\leq d} (-1)(d-m) P(\Theta=j|X=x) - (d-m)\sum_{j>d} P(\Theta=j|X=x)$$

$$= (d-m)\{P(\Theta \leq m|X=x) - P(m<\Theta\leq d|X=x) - P(\Theta > d|X=x)\}$$

$$= (d-m)\{P(\Theta \leq m|X=x) - P(\Theta > m|X=x)\}$$

$$= (d-m)\{2\underbrace{P(\Theta \leq m|X=x)}_{\geq 1/2} - 1\} \geq 0$$

Assim, para $d > m$.

$$E[L(d,\theta)|X=x] \geq E[L(m,\theta)|X=x]$$

Analogamente para $d < m$!!!

Assim, tomamos, em geral, $\delta^*(X) = \text{mediana}(\theta|x)$
$x \longmapsto \text{Mediana}(\theta|X=x)$

No exemplo da urna,

j	$P(\theta=j \mid X=(0,0))$
0	20/40
1	12/40
2	6/40
3	2/40
4	0
5	0

$\text{Mediana}(\theta|X=(0,0)) = 0$ (ou 1)

$\text{Mediana}(\theta|X=(1,1)) = 4$ (ou 5).

$\text{Mediana}(\theta|X=(1,0)) = 2$ (ou 3).

j	$P(\theta=j \mid X=(1,0))$
0	0
1	4/20
2	4/20
3	6/20
4	4/20
5	0

Há 16 est de Bayes, neste caso!

UM estimador de Bayes é dado por:

$$\delta_1^*(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } (1,0) \\ 4, & x=(1,1) \end{cases}$$

No exemplo 2,

$$\Theta | X_1=x_1, \dots, X_n=x_n \sim N\left(\frac{b_0^2 \sum_{i=1}^{n} x_i + \sigma_0^2 a_0}{b_0^2 n + \sigma_0^2}, \ ? \right)$$

$$\text{Mediana}(\Theta | X_1=x_1, \dots, X_n=x_n) = \frac{b_0^2 \sum_{i=1}^{n} x_i + \sigma_0^2 a_0}{b_0^2 n + \sigma_0^2}.$$

O estimador de Bayes com relação à perda absoluta é

$$\delta^*(X) = \text{Mediana}(\Theta|X) = \frac{b_0^2 \sum_{i=1}^{n} X_i + \sigma_0^2 a_0}{n b_0^2 + \sigma_0^2}.$$

(3) Função de perda 0-1

$$L(d,\theta) = \begin{cases} 0, & d=\theta \\ 1, & d\neq\theta \end{cases}$$

Para $x \in \mathcal{X}$,

$$E[L(d,\theta)|X=x] = \sum_{j\in\Theta} L(d,j) P(\Theta=j|X=x)$$

$$= \sum_{j\neq d} 1 \cdot P(\Theta=j|X=x) + \underbrace{0}_{L(d,d)} \cdot P(\Theta=d|X=x)$$

$$= P(\Theta\neq d|X=x) = 1 - P(\Theta=d|X=x)$$

Assim, devemos tomar uma moda da dist a posteriori de Θ dado $X=x$ como estimativa

X discreto
Lembrete: Dizemos que m_0 é a moda da dist de X se $P(X=m_0) \geq P(X=x), \forall x$

No exemplo da urna, um estimador de Bayes é
↳ existem 4!

$$\delta^*(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } x=(1,0) \text{ (ou 3)} \\ 5, & x=(1,1) \end{cases}$$

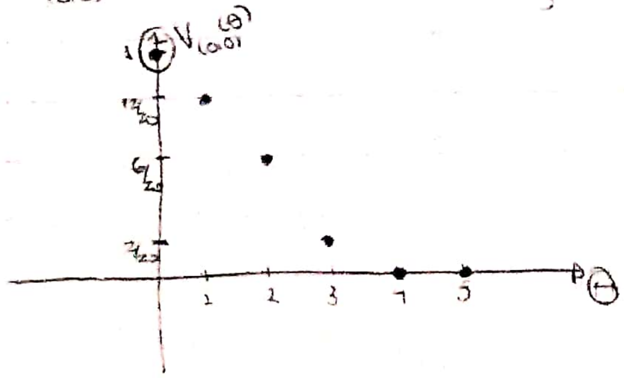
ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA

Ao estimador $\delta_{MV}: \mathcal{X} \to \Theta$ que associa a cada $x \in \mathcal{X}$,

$\delta_{MV}(x) = \underset{\theta \in \Theta}{\arg\max}\, V_x(\theta)$ damos o nome de estimador de máxima verossimilhança para θ.

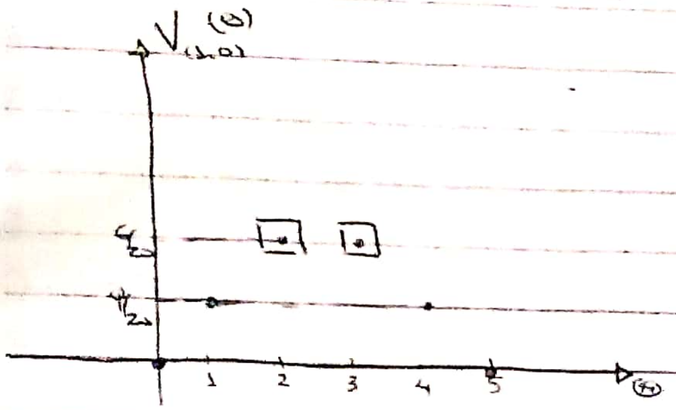
(1) URNA:

$$V_{(0,0)}(\theta) = P(X=(0,0)|\theta) = \frac{(5-\theta)}{5}\frac{(4-\theta)}{4}$$

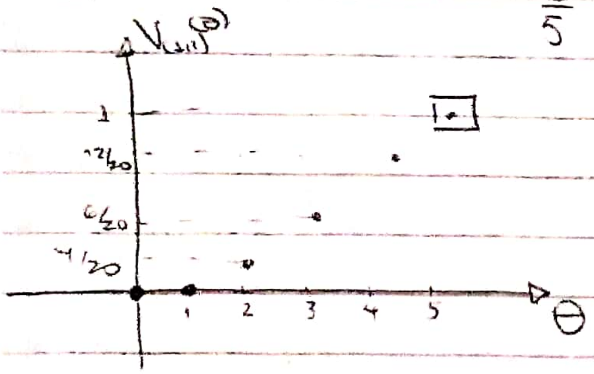


$\delta_{MV}(0,0) = 0$.

$$V_{(1,0)}(\theta) = P(X=(1,0)|\theta) = \frac{\theta}{5}\,\frac{(5-\theta)}{4}$$



$$V_{(1,1)}(\theta) = P(X=(1,1)|\theta) = \frac{\theta}{5}\,\frac{(\theta-1)}{4}$$



Assim,

$$\delta_{MV}^{(1)}(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } x=(1,0) \\ 5, & x=(1,1) \end{cases}$$

ou

$$\delta_{MV}^{(2)}(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \\ 3, & x=(1,0) \\ 5, & x=(1,1) \end{cases}$$

Ex 2 $X_1, \dots, X_n$ dado θ, i.i.d Ber(θ)

$\mathcal{X} = \{0,1\}^n$, $\quad X = (X_1, X_2, \dots, X_n)$

$\Theta = [0,1]$

$x = (x_1, \dots, x_n) \in \mathcal{X}$,

$$V_x(\theta) = P(X_1 = x_1, \dots, X_n = x_n | \theta) = \prod_{i=1}^{n} P(X_i = x_i | \theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

$$\Longrightarrow V_x(\theta) = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i}$$

$x = (1, 1, \dots, 1) \Longrightarrow V_x(\theta) = \theta^n \Longrightarrow \hat{\delta}_{MV}(1,1,\dots,1) = 1 \sim$

$x = (0, 0, \dots, 0) \Longrightarrow V_x(\theta) = (1-\theta)^n \sim$

$1 \leq \sum_{i=1}^n x_i \leq n-1$

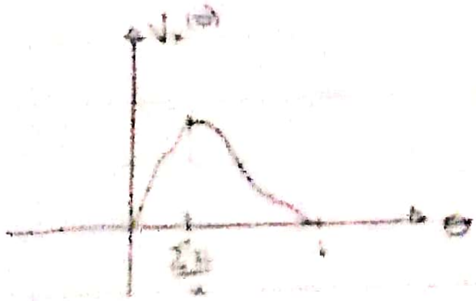
$$\frac{dV_x(\theta)}{d\theta} = \left(\sum_{i=1}^n x_i\right) \theta^{\sum x_i - 1} (1-\theta)^{n - \sum x_i} + \theta^{\sum x_i} \left(n - \sum_{i=1}^n x_i\right)(1-\theta)^{n - \sum x_i - 1}(-1)$$

$$\Longrightarrow \frac{dV_x(\theta)}{d\theta} = \theta^{\sum x_i - 1} (1-\theta)^{n - \sum x_i - 1} \left\{ \sum_{i=1}^n x_i (1-\theta) - \left(n - \sum_{i=1}^n x_i\right)\theta \right\}$$

$$\Longrightarrow \frac{dV_x(\theta)}{d\theta} = \theta^{\sum x_i - 1} (1-\theta)^{n - \sum x_i - 1} \left\{ \sum_{i=1}^n x_i - n\theta \right\} > 0$$

$$\Longleftrightarrow \sum_{i=1}^n x_i - n\theta > 0 \Longleftrightarrow \theta < \frac{\sum_{i=1}^n x_i}{n}$$

$$(<) 0 \Longleftrightarrow \theta > \frac{\sum_{i=1}^n x_i}{n}$$



Logo,

$\delta_{MV}(X) = \delta_{MV}(X_1, \dots, X_n) = \frac{\sum X_i}{n}$

17/04 Aula 06

Revisão: $x \in \mathcal{X} \to \theta | X = x$

$E[L(d,\theta) | X = x]$

$x \in \mathcal{X} \to V_x(\cdot)$

$\delta_{MV}: \mathcal{X} \to \Theta$

$\forall x \in \mathcal{X}, \; V_x(\delta_{MV}(x)) = \max_{\theta \in \Theta} V_x(\theta)$

Exemplo (3) $X_1, \dots, X_n | \theta$ i.i.d Beta$(\theta, 1)$

$\mathcal{X} = (0,1)^n$

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

$x = (x_1, x_2, \dots, x_n) \in \mathcal{X}$.

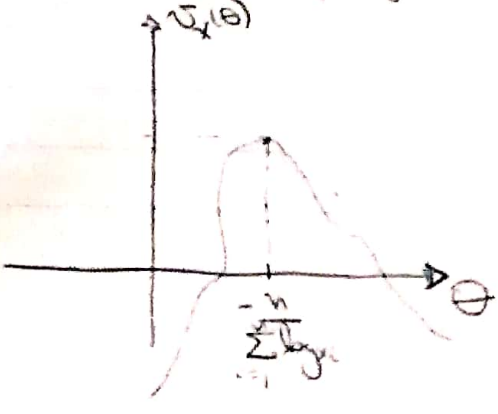
$V_x(\theta) = f(x_1, \dots, x_n | \theta) = \prod_{i=1}^{n} f(x_i | \theta) = \prod \theta x_i^{\theta - 1} \mathbb{1}_{(0,1)}(x_i)$

$= \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1} \mathbb{1}_{(0,1)}(x)$;

$\mathcal{V}_x(\theta) = \log V_x(\theta) = \log\left(\theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}\right) = n \log \theta + (\theta - 1) \log \prod_{i=1}^{n} x_i$

$\frac{d \mathcal{V}_x(\theta)}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i \underset{(<)}{>} 0 \iff \frac{n}{\theta} \underset{(<)}{>} -\sum_{i=1}^{n} \log x_i$

$$\Longleftrightarrow \theta \underset{(>)}{<} \frac{-n}{\sum_{i=1}^{n} \log x_i}$$



Logo, $\delta_{MV}(X) = \dfrac{-n}{\sum_{i=1}^{n} \log x_i}$

<u>Ex 4</u> $X_1, X_2, \ldots, X_n$ dado $\theta = (\theta_1, \theta_2)$, c.i.i.d. $N(\theta_1, \theta_2)$

$\mathcal{X} = \mathbb{R}^n \qquad \Theta = \mathbb{R} \times \mathbb{R}_+$

$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$V_x(\theta) = f(x_1, \ldots, x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_i - \theta_1)^2}{2\theta_2}}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \theta_2^{-n/2} e^{-\sum_{i=1}^{n}(x_i - \theta_1)^2/2\theta_2} \leq \left(\frac{1}{\sqrt{2\pi}}\right)^n \underbrace{\theta_2^{-n/2} e^{-\sum_{i=1}^{n}(x_i - \bar{x})^2/2\theta_2}}_{W_x(\theta_2)}$$

$$W_x(\theta_2) = \theta_2^{-n/2} e^{-\sum_{i=1}^{n}(x_i - \bar{x})^2/2\theta_2}$$

$$w_x(\theta_2) = \log W_x(\theta_2) = -\frac{n}{2} \log \theta_2 - \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{2\theta_2}$$

$$\frac{d\,\omega_x(\theta)}{d\theta_2} = \frac{-n}{2\theta_2} + \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{2\theta_2^2}$$

$$= \frac{1}{2\theta_2}\left\{-n + \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\theta_2}\right\} > 0$$

$$\iff -n + \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\theta_2} > 0 \iff \theta_2 < \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

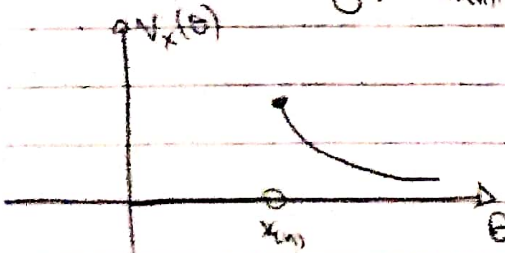
Logo, $\delta_{MV}(x) = \left(\bar{X}_n, \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}\right)$

(5) $X_1, \dots, X_n$, dado θ, são c.i.i.d $\mathcal{U}(0,\theta)$

$\Theta = \mathbb{R}_+$
$\mathcal{X} = \mathbb{R}_+^n$
$x = (x_1, \dots, x_n) \in \mathcal{X}$

$$V_x(\theta) = f(x_1 \dots x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{1}{\theta}\,\mathbb{I}_{[0,\theta]}(x_i) = \frac{1}{\theta^n}\,\mathbb{I}_{[0,x_{(n)}]}(x_{(1)})\,\mathbb{I}_{[0,\theta]}(x_{(n)})$$

$$V_x(\theta) = K(x)\,\frac{1}{\theta^n}\,\mathbb{I}_{[x_{(n)},\infty)}(\theta), \quad K(x) = \mathbb{I}_{[0,x_{(n)}]}(x_{(1)})\,\mathbb{I}_{[0,\infty)}(x_{(n)})$$



$\delta_{MV}(x) = x_{(n)}$

Problema peixes (captura-recaptura)

$\Theta \sim HG(\Theta; N_1, N_1, n)$

$\Theta = \{N_1, \dots\}$

$(n) \to x \quad x = 0 \Rightarrow \theta = \infty$

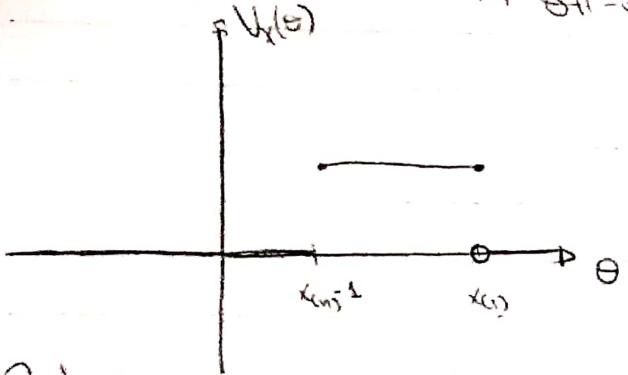
→ Olga S. Yoshida (Mestrado)
José Galvão Leite

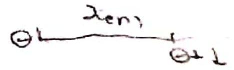
(6) $X_1, \dots X_n$ dado Θ c.i.i.d. $U(\Theta, \Theta+1)$

$\mathcal{X} = \mathbb{R}^n$; $\Theta = \mathbb{R}$

$x = (x_1, \dots, x_n) \in \mathcal{X}$,

$$V_x(\theta) = \prod_{i=1}^{n} f(x_1, \dots, x_n | \theta) = \prod_{i=1}^{n} \frac{1}{\theta+1-\theta} \mathbb{1}_{[\theta, \theta+1]}(x_i) = \mathbb{1}_{(\theta, x_{(n)})}(x_{(1)}) \, \mathbb{1}_{(\theta, \theta+1)}(x_{(n)})$$





$\theta \leq x_{(n)} \leq \theta+1$

$\theta \leq x_{(1)} \leq x_{(n)}$

$\theta - 1 \leq x_{(n)} - 1 \leq \theta \leq x_{(1)} \leq x_{(n)}$

Portanto, para todo $\alpha \in (0,1)$

$$\delta_{MV}^{(\alpha)}(x) = \alpha X_{(1)} + (1-\alpha)(X_{(n)} - 1) \text{ é um EMV para } \theta.$$

Alguns comentários.

(1) T é uma estatística suficiente δ_{MV} é único $\rightarrow$ δ_{MV} é função de T

Quando δ_{MV} não é único, o δ_{MV} pode não ser função de alguma estatística suficiente.

$$\delta_{MV}^{(2)}(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \\ 3, & x=(1,0) \\ 5, & x=(1,1) \end{cases} \qquad T(\underline{x}) = X_1 + X_2.$$

Por outro lado, nesse caso, existe δ_{MV} que é função de T

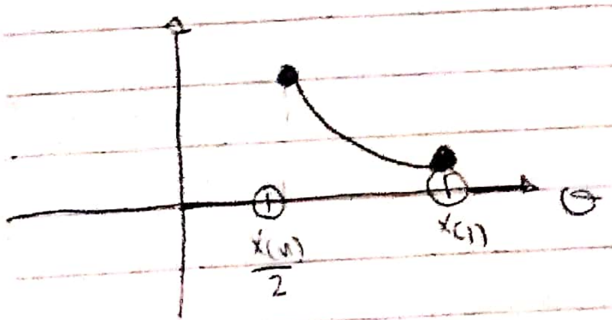
$$\delta_{MV}^{(1)}(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } x=(1,0) \\ 5, & x=(1,1) \end{cases}$$

(2) Quando δ_{MV} é uma estatística suficiente e única $\Rightarrow$ δ_{MV} é uma estatística suficiente mínima.

Ex $X_1, X_2, \ldots, X_n$ dado θ i.i.d. $U(\theta, 2\theta)$

$\Theta = \mathbb{R}_+^n \qquad \mathcal{X} = \mathbb{R}_+^n$

$$V_x(\theta) = \prod_{i=1}^{n} \frac{1}{(2\theta-\theta)} \mathbb{I}_{[\theta,2\theta]}(x_i) = \frac{1}{\theta^n} \cdot \mathbb{I}_{[\theta, x_{(n)}]}(x_{(1)}) \cdot \mathbb{I}_{[\theta,2\theta]}(x_{(n)})$$



$\delta_{MV}(X) = \frac{X_{(n)}}{2}$. (Não é uma estatística suficiente)

(3) Invariância dos EMV

Seja $g: \Theta \to \Gamma \ (= g(\Theta))$, g qualquer

δ_{MV} é o EMV para θ $\rightarrow$ $g(\delta_{MV})$ é o EMV para $g(\theta)$.

$g: \Theta \rightarrow \mathcal{T}$

$\tau = g(\theta)$

Para cada $\tau_0 \in \mathcal{T}$, seja $L'_x(\tau_0) = \max_{\theta \in g^{-1}(\{\tau_0\})} L_x(\theta)$

$L'_x: \mathcal{T} \rightarrow \mathbb{R}_+$ função de verossimilhança induzida para τ.

Nesse caso,

$\tau_{MV}: \mathcal{X} \rightarrow \mathcal{T}$ é tq

$L_x(\tau_{MV}(x)) = \max_{\tau \in \mathcal{T}} L'_x(\tau)$.

Seja $\tau^* = g(\delta_{MV})$. Para $x \in \mathcal{X}$,

$\tau^*(x) = g(\delta_{MV}(x))$

$L'_x(\tau^*(x)) \leq \max_{\tau \in \mathcal{T}} L'_x(\tau) = L'_x(\tau_{MV}(x)) \rightarrow L'_x(\tau^*(x)) \leq L'_x(\tau_{MV}(x))$

$L'_x(\tau_{MV}(x)) = \max_{\tau \in \mathcal{T}} L'_x(\tau) = \max_{\tau \in \mathcal{T}} \left\{ \max_{\theta \in g^{-1}(\{\tau\})} L_x(\theta) \right\} \leq \max_{\tau \in \mathcal{T}} \{ L_x(\delta_{MV}(x)) \}$

$\leq \max_{\tau \in \mathcal{T}} \left\{ \max_{\theta \in g^{-1}(\{g(\delta_{MV}(x))\})} L_x(\theta) \right\} \overset{\text{def}}{=} \max_{\tau \in \mathcal{T}} L'_x(\underbrace{g(\delta_{MV}(x))}_{\tau^*(x)})$

$\Rightarrow L'_x(\tau_{MV}(x)) \leq L_x(\tau^*(x))$

Ex 1. $X_1, \ldots, X_n$ dado Θ c.i.i.d. $U(0,\Theta)$

$g(\Theta) = Var(X_1|\Theta) = \frac{\Theta^2}{12}$

Vimos que o EMV para Θ é $\delta_{MV}(\mathbf{X}) = X_{(n)}$

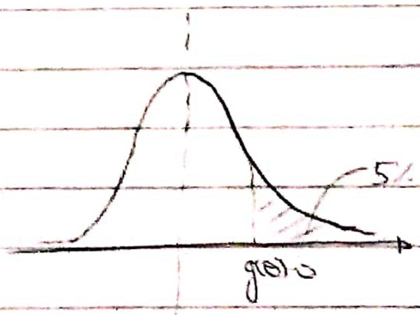
Pela invariância dos EMV, segue

$$\widehat{g(\Theta)}_{MV} = \frac{\delta_{MV}^2}{12} = \frac{X_{(n)}^2}{12}$$

(2) $X_1, \ldots, X_n$ dado Θ c.i.i.d. $N(\Theta, 1)$

$P_{95\%} = g(\Theta)$

$P(X \leq g(\Theta)|\Theta) = 95\%$



$P\left(\frac{X-\Theta}{1} \leq \frac{g(\Theta)-\Theta}{1} \,\middle|\, \Theta\right) = 95\%$

$\Rightarrow P(Z \leq g(\Theta)-\Theta|\Theta) = 95\% \Rightarrow g(\Theta)-\Theta = 1{,}64 \Rightarrow g(\Theta) = \Theta + 1{,}64.$

Vimos que o EMV para Θ é $\delta_{MV}(x) = \bar{X}_n$. Logo, pela invariância dos EMV, segue que:

$$T_{MV}(x) = \widehat{g(\Theta)}_{MV}(x) = \delta_{MV}(x) + 1{,}64 = \bar{X}_n + 1{,}64$$

Cont. 2 $X_1, \ldots, X_n|\theta \sim N(\theta_1, \theta_2)$, onde $\theta = (\theta_1, \theta_2)$

$$P(X_1 \leq g(\theta)|\theta) = 0{,}95 \iff P\left(\frac{X_1 - \theta_1}{\sqrt{\theta_2}} \leq \frac{g(\theta) - \theta_1}{\sqrt{\theta_2}} \middle| \theta\right) = 0{,}95 \iff$$

$$\iff \frac{g(\theta) - \theta_1}{\sqrt{\theta_2}} = 1{,}64 \iff g(\theta) = \theta_1 + 1{,}64\sqrt{\theta_2}$$

Vimos que o EMV para $\theta = (\theta_1, \theta_2)$ é

$$\delta_{MV}(X) = \left(\bar{X}_n, \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}\right)$$ Logo, pela invariância dos EMV, temos.

$$\tau_{MV}(x) = \hat{g(\theta)}_{MV} = g(\theta_{MV}(x)) = \bar{X}_n + 1{,}64\sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}}.$$

(4) $X_1, \ldots, X_n$ dado θ i.i.d. $\mathrm{Exp}(\theta)$.

$g(\theta) = \text{Mediana}(\theta|X)$

$$P(X_1 \leq g(\theta)|\theta) = 1/2 \Rightarrow 1 - e^{-\theta g(\theta)} = 1/2 \Rightarrow e^{-\theta g(\theta)} = \frac{1}{2}$$

$$\iff -\theta g(\theta) = \log 1/2 \Rightarrow g(\theta) = \frac{\log 2}{\theta}.$$

Ex. $\delta_{MV} = \dfrac{1}{\bar{X}_n}$.

Pela invariância,

$$\hat{g(\theta)}_{MV}(x) = \frac{\log 2}{1/\bar{X}_n} = \bar{X}_n \cdot \log 2$$

MÉTODO DOS MOMENTOS

$X_1, \ldots, X_n$ dado $\Theta = (\theta_1, \ldots, \theta_r)$ são c.i.i.d. com densidade $f(\cdot|\theta)$, tais que $E(X_1^r|\theta) < \infty$.

Em geral, para $j=1, \ldots, r$,

$$E(X_1^j|\theta) = h_j(\theta_1, \ldots, \theta_r)$$

Seja $M_j = \dfrac{X_1^j + \ldots + X_n^j}{n}$ o j-ésimo momento amostral, $j=1, \ldots, r$. Considere o sistema,

$$E(X_1^j|\theta) = M_j, \; j=1, \ldots, r.$$

Uma solução deste sistema produz um estimador para $\Theta = (\theta_1, \ldots, \theta_r)$ chamado ESTIMADOR DE MOMENTOS

Exemplos (1) $X_1, \ldots, X_n$ dado θ c.i.i.d. Ber(θ), $(r=1)$

$$E(X_1|\theta) = \frac{X_1 + \ldots + X_n}{n} \Rightarrow \theta = \frac{X_1 + \ldots + X_n}{n} \Rightarrow \hat{\theta}_{mom}(x) = \bar{X}$$

(2) $X_1, \ldots, X_n$ dado $\Theta = (\theta_1, \theta_2)$ c.i.i.d. $N(\theta_1, \theta_2)$ $(r=2)$

$$\begin{cases} E[X_1|\theta] = \dfrac{X_1 + \ldots + X_n}{n} \\ E[X_1^2|\theta] = \dfrac{X_1^2 + \ldots + X_n^2}{n} \end{cases} \Longleftrightarrow \begin{cases} \theta_1 = \bar{X}_n \\ \theta_2 + \theta_1^2 = \dfrac{\sum_{i=1}^n X_i^2}{n} \end{cases}$$

$$\Longrightarrow \begin{cases} \theta_1 = \bar{X}_n \\ \theta_2 = \dfrac{\sum X_i^2}{n} - \bar{X}_n^2 = \dfrac{\sum_{i=1}^{n} (X_i - \bar{x})^2}{n} \end{cases}$$

$$\hat{\theta}_{mom}(x) = \left(\bar{X}_n, \frac{\sum_{i=1}^{n} (X_i - \bar{x})^2}{n}\right)$$

---

(3) $X_1, \ldots, X_n$, dado Θ, c.i.i.d. $U(0, \theta)$. $(r=1)$

$$E(X_1 | \theta) = \frac{X_1 + \ldots + X_n}{2} \Rightarrow \frac{\theta}{2} = \frac{X_1 + \ldots + X_n}{n} \Rightarrow \theta = 2\bar{X}_n$$

Logo, $\hat{\theta}_{mom}(x) = 2\bar{X}_n$

Por exemplo, se $n=4$ e $(1, 2, 2, 15)$ $\hat{\theta}_{mom} = 10$!!? $\left(\begin{array}{l}\hat{\theta}_{mom}(1,2,2,15) = 10 \\ \hat{\theta}_{mom}(10,12,13,15) = 25\end{array}\right)$ não satisfaz o princípio da suficiência!

Princípio da Suficiência

"A inferência deve ser baseada em x apenas através de $t(x)$ se t é uma estatística suficiente".

($x, y \in \mathcal{X}$ tq $t(x) = t(y)$, então as conclusões (inferências) sobre θ a partir de y devem coincidir)

Referências: De Groot, Casella & Berger, "the likelihood principle" Berger e Wolpert (1988)

Prob. e Inferência Estatística I

21/05/2013 - Aula 7 Luis Gustavo Esteves

Estimador de Bayes

EMV

E. Momentos

Exercício.

$X_1, \ldots, X_n$ dado $\Theta = (\Theta_1, \Theta_2)$, c.i.i.d. Gama$(\Theta_1, \Theta_2)$ $(r=2)$

$\hat{\Theta}_{MOM} = ?$

Em geral,

$$g(\theta) = \mu\left(E(X_1|\theta), E(X_1^2|\theta), \ldots, E(X_1^r|\theta)\right)$$

$$\widehat{g(\theta)}_{MOM} = \mu(M_1, M_2, \ldots, M_r)$$

Estimadores de Mínimos Quadrados

$X = (X_1, \ldots, X_n)$ que dado Θ, possui "densidade" $f(\cdot|\theta)$.

Idéia: Escolher para cada $x \in \mathfrak{X}$, como estimativa o valor que minimiza a "diferença" entre valor observado e esperado.

Para cada $x \in \mathfrak{X}$, tomamos $\delta_{MQ}(x)$ tal que

$$\underbrace{\sum_{i=1}^{n} (x_i - E(X_i | \delta_{MQ}(x)))^2}_{S(x, \delta_{MQ}(x))} = \min_{\theta \in \Theta} \underbrace{\sum_{i=1}^{n} (x_i - E(X_i|\theta))^2}_{S(x, \theta)}.$$

Exemplo 1. $X_1, \ldots, X_n$, dado θ, c.i.i.d. $N(\theta, 1)$

$E(X_1|\theta) = \theta$. $\quad \mathcal{X} = \mathbb{R}^n \quad \Theta = \mathbb{R}$

Para $x = (x_1, \ldots, x_n) \in \mathcal{X}$,

$$S(x, \theta) = \sum_{i=1}^{n} (x_i - E(X_i|\theta))^2 = \sum_{i=1}^{n} (x_i - \theta)^2, \text{ donde}$$

$$S(x, \bar{x}) = \min_{\theta \in \Theta} S(x, \theta). \text{ Assim, } \delta_{MQ}(X) = \bar{X}_n.$$

Exemplo 2.

$X_1, \ldots, X_n$ dado $\theta = (\theta_0, \theta_1)$, são condicionalmente independentes tais que
$X_i | \theta \overset{d}{=} \theta_0 + \theta_1 t_i + Z_i$, onde $Z_1, \ldots, Z_n$ são i.i.d. $N(0,1)$ e

$t_1, t_2, \ldots, t_n \in \mathbb{R}$, constantes conhecidas.

$$E(X_i|\theta) = E(\theta_0 + \theta_1 t_i + Z_i) = \theta_0 + \theta_1 t_i$$

Aqui, $\mathcal{X} = \mathbb{R}^n$, $\Theta = \mathbb{R}^2$.

Para $x = (x_1, \ldots, x_n) \in \mathcal{X}$, temos

$$S(x, \theta) = \sum_{i=1}^{n} (x_i - E(X_i|\theta))^2 =$$

$$= \sum_{i=1}^{n} (x_i - (\theta_0 + \theta_1 t_i))^2$$

$$\frac{\partial S(x,\theta)}{\partial \theta_0} = 2 \sum_{i=1}^{n} [x_i - (\theta_0 + \theta_1 t_i)](-1) \quad \text{e} \quad \frac{\partial S(x,\theta)}{\partial \theta_1} = 2 \sum_{i=1}^{n} [x_i - (\theta_0 + \theta_1 t_i)](-t_i)$$

Agora,

$$\frac{\partial S(x,\theta)}{\partial \theta_0} = 0 \Leftrightarrow \sum_{i=1}^{n}(x_i - (\theta_0 + \theta_1 t_i)) = 0 \Leftrightarrow \sum_{i=1}^{n} x_i - n\theta_0 - \theta_1 \sum_{i=1}^{n} t_i = 0$$

e

$$\frac{\partial S(x,\theta)}{\partial \theta_1} = 0 \Leftrightarrow \sum_{i=1}^{n} x_i t_i - \theta_0 t_i - \theta_1 t_i^2 = 0 \Leftrightarrow \sum_{i=1}^{n} x_i t_i - \theta \sum_{i=1}^{n} t_i - \theta_1 \sum_{i=1}^{n} t_i^2 = 0$$

Em resumo,

$$\begin{cases} \sum_{i=1}^{n} x_i - n\theta_0 + \theta_1 \sum_{i=1}^{n} t_i = 0 \\ \sum_{i=1}^{n} x_i t_i - \theta \sum_{i=1}^{n} t_i - \theta_1 \sum_{i=1}^{n} t_i^2 = 0 \end{cases} \quad \text{(Equações Normais)}$$

Exemplo 3.

$X_1, \ldots, X_n$ dado $\theta = (\theta_0, \theta_1, \theta_2)$, são condicionalmente independentes tais que

$$X_i / \theta \overset{d}{=} \theta_0 + \theta_1 t_i + \theta_2 t_i^2 + Z_i, \text{ onde } Z_1, \ldots, Z_n \text{ são i.i.d. } N(0,1)$$

e $t_1, \ldots, t_n \in \mathbb{R}$, constantes conhecidas.

Nesse caso, $\mathfrak{X} = \mathbb{R}^n$ e $\Theta = \mathbb{R}^3$ e,

$$E(X_i/\theta) = E(\theta_0 + \theta_1 t_i + \theta_2 t_i^2 + Z_i) = \theta_0 + \theta_1 t_i + \theta_2 t_i^2$$

$$S(x,\theta) = \sum_{i=1}^{n} (x_i - E(X_i/\theta))^2 = \sum_{i=1}^{n} (x_i - (\theta_0 + \theta_1 t_i + \theta_2 t_i^2))^2$$

Logo, segue que

$$\frac{\partial S(x,\theta)}{\partial \theta_0} = 2\sum_{i=1}^{n}\left(x_i - (\theta_0 + \theta_1 t_i + \theta_2 t_i^2)\right)(-1)$$

$$\frac{\partial S(x,\theta)}{\partial \theta_1} = 2\sum_{i=1}^{n}\left(x_i - (\theta_0 + \theta_1 t_i + \theta_2 t_i^2)\right)(-t_i)$$

$$\frac{\partial S(x,\theta)}{\partial \theta_2} = 2\sum_{i=1}^{n}\left(x_i - (\theta_0 + \theta_1 t_i + \theta_2 t_i^2)\right)(-t_i^2)$$

~~Em nesse~~ Assim,

$$\frac{\partial S(x,\theta)}{\partial \theta_0} = 0 \Leftrightarrow \sum_{i=1}^{n} x_i - n\theta_0 - \theta_1\sum_{i=1}^{n} t_i - \theta_2\sum_{i=1}^{n} t_i^2 = 0$$

$$\frac{\partial S(x,\theta)}{\partial \theta_1} = 0 \Leftrightarrow \sum_{i=1}^{n} x_i t_i - \theta_0\sum_{i=1}^{n} t_i - \theta_1\sum_{i=1}^{n} t_i^2 - \theta_2\sum_{i=1}^{n} t_i^3 = 0$$

$$\frac{\partial S(x,\theta)}{\partial \theta_2} = 0 \Leftrightarrow \sum_{i=1}^{n} x_i t_i^2 - \theta_0\sum_{i=1}^{n} t_i^2 - \theta_1\sum_{i=1}^{n} t_i^3 - \theta_2\sum_{i=1}^{n} t_i^4 = 0$$

Em resumo, segue que.

$$\begin{cases} \sum_{i=1}^{n} x_i - n\theta_0 - \theta_1\sum_{i=1}^{n} t_i - \theta_2\sum_{i=1}^{n} t_i^2 = 0 \\ \sum_{i=1}^{n} x_i t_i - \theta_0\sum_{i=1}^{n} t_i - \theta_1\sum_{i=1}^{n} t_i^2 - \theta_2\sum_{i=1}^{n} t_i^3 = 0 \\ \sum_{i=1}^{n} x_i t_i^2 - \theta_0\sum_{i=1}^{n} t_i^2 - \theta_1\sum_{i=1}^{n} t_i^3 - \theta_2\sum_{i=1}^{n} t_i^4 = 0 \end{cases}$$

"Eq. Normais"

Só resolver o sistema.

Comparação (Avaliação) De Estimadores

$\Delta = \{\delta : \mathcal{X} \to \Theta\}$

Idéia: Escolher $\delta^* \in \Delta$ que minimiza $E[L(\delta(X), \theta)]$ em Δ

$$\rho(\delta^*) = \min_{\delta \in \Delta} \rho(\delta)$$

Se $h(\overset{\text{aleatório}}{\Theta}, X)$ usamos a ideia da bayesiana.

Na clássica, Θ é fixo!

→ Em paralelo

Modelo Bayesiano

$(\mathcal{X} \times \Theta, \sigma, \mathbb{P})$

Modelo Clássico

$(\mathcal{X}, \sigma, \{\mathbb{P}_\theta : \theta \in \Theta\})$

Seja $\delta \in \Delta$.

À quantidade $\sum_{x \in \mathcal{X}} L(\delta(x), \theta)\, \mathbb{P}(X = x / \theta)$ $\left(\int_{\mathcal{X}} L(\delta(x), \theta) f(x/\theta)\, dx \right)$ chamamos de Risco (frequentista) do estimador δ para estimar o parâmetro quando esse vale θ.

NOTAÇÃO. $R(\delta, \theta) = E[L(\underbrace{\delta(X)}_{\text{Aleatório}}, \theta) | \theta] = E_\theta[L(\delta(X), \theta)]$

Exemplo: $X_1, \dots, X_n$ dado θ são c.i.i.d. $Ber(\theta)$

$\Theta = [0,1]$ $\quad \mathcal{X} = \{0,1\}^n$

$L(d, \theta) = (d - \theta)^2$ $\quad X = (X_1, \dots, X_n)$

$\delta_1(X) = X_1$

$\delta_2(X) = \bar{X}_n$ $\qquad \delta_3(X) = \dfrac{a + \sum_{i=1}^{n} X_i}{a+b+n}$ $\qquad \delta_4(X) = \dfrac{\pi}{4}$

Calculando os riscos dos estimadores...

$$R(\delta_1,\theta) = E_\theta[L(\delta_1(X),\theta)] = E_\theta[(\delta_1(X)-\theta)^2] =$$

$$= E_\theta((X_1-\theta)^2) = Var_\theta(X_1) = \theta(1-\theta)$$

$$R(\delta_2,\theta) = E_\theta[L(\delta_2(X),\theta)] = E_\theta[(\delta_2(X)-\theta)^2] =$$

$$= E_\theta((\bar{X}_n-\theta)^2) = Var_\theta(\bar{X}_n) = \frac{\theta(1-\theta)}{n}$$

$$R(\delta_3,\theta) = E_\theta[L(\delta_3(X),\theta)] = E_\theta[(\delta_3(X)-\theta)^2] =$$

$$= E_\theta\left[\left(\frac{a + \sum^{n} X_i}{a+b+n} - \theta\right)^2\right] = E_\theta\left[\left(\frac{a+n\bar{X}_n}{a+b+n} - \theta\right)^2\right] =$$

$$= Var_\theta\left(\frac{a+n\bar{X}_n}{a+b+n} - \theta\right) + \left\{E_\theta\left(\frac{a+n\bar{X}_n}{a+b+n} - \theta\right)\right\}^2 =$$

$$= \frac{n^2 Var_\theta(\bar{X}_n)}{(a+b+n)^2} + \left[\frac{a+n\theta}{a+b+n} - \theta\right]^2 = \frac{n\theta(1-\theta)}{(a+b+n)^2} + \left(\frac{a+n\theta}{a+b+n} - \theta\right)^2,$$

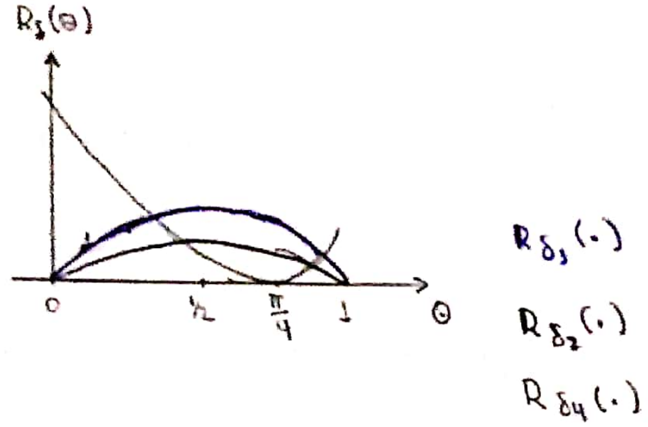
e, finalmente,

$$R(\delta_4,\theta) = E_\theta[L(\delta_4(X),\theta)] = E_\theta[(\delta_4(X)-\theta)^2] = \left(\frac{\pi}{4} - \theta\right)^2 \;//$$

Para $\delta \in \Delta$, seja $R_\delta : \Theta \to \mathbb{R}_+(\mathbb{R})$

$$\theta \in \Theta \mapsto R_\delta(\theta) = R(\delta,\theta)$$

R_δ é chamada FUNÇÃO DE RISCO do estimador δ.



Em geral, não se pode encontrar um estimador que seja melhor que o outro para todo $\theta \in \Theta$.

Quando $L(d,\theta) = (d-\theta)^2$, $R(\delta,\theta) = E_\theta((\delta(x)-\theta)^2)$ é chamado Erro Quadrático Médio de δ e θ.

$$EQM_\theta(\delta) = E_\theta[(\delta(X)-\theta)^2]$$

$EQM(\delta,\theta)$

$EQM(\delta|\theta)$

Dizemos que δ_1 é melhor que δ_2 se

$$R_{\delta_1}(\theta) \leq R_{\delta_2}(\theta), \forall \theta \in \Theta.$$

Dizemos que δ_1 e δ_2 são equivalentes se δ_1 é melhor que δ_2 e δ_2 é melhor que δ_1.

Quando δ_1 é melhor que δ_2 mas não é melhor que δ_1, dizemos que δ_1 "domina" δ_2. Nesse caso, dizemos que δ_2 é inadmissível.

Seja $\Delta^* \subseteq \Delta$

Dizemos que δ^* é o melhor em Δ^* se δ^* é melhor que δ, $\forall \delta \in \Delta^*$.

$E_\theta(\delta(x))$ é finita pois teríamos $\delta(x)$ inadmissível em outro caso

$$EQM_\theta(\delta) = E_\theta[(\delta(x)-\theta)^2] = E_\theta[(\delta(x) - E_\theta(\delta(x)) + E_\theta(\delta(x)) - \theta)^2] =$$

$= 0$

$$= E_\theta[(\delta(x) - E_\theta(\delta(x)))^2] + E_\theta[(E_\theta(\delta(x)) - \theta)^2] + 2E_\theta\left[(\delta(x) - E_\theta(\delta(x)))(E_\theta(\delta(x)) - \theta)\right]$$

$$= Var_\theta(\delta(x)) + (E_\theta(\delta(x)) - \theta)^2$$

$B_\theta(\delta) = E_\theta(\delta(x)) - \theta$. vício do estimador de δ

Assim,

$$EQM_\theta(\delta) = Var_\theta(\delta) + B^2(\delta)$$

Vamos focar nos δ's tais que $B(\delta) = 0$.

$\Delta' = \{\delta \in \Delta : B_\theta(\delta) = 0, \forall \theta \in \Theta\}$

Vamos restringir a busca à classe

$$\Delta' = \{\delta \in \Delta : E_\theta[\delta(x)] = \theta, \forall \theta \in \Theta\}$$

V

Definição: Um estimador $\delta: \mathfrak{X} \to \underset{g(\Theta)}{\Theta}$ para $\underset{g(\theta)}{\theta}$ é dito NÃO-VIESADO para θ se

$$\underset{g(\theta)}{E_\theta[\delta(X)]} = \underset{g(\theta)}{\theta}, \quad \forall \theta \in \Theta$$

Exemplos:

(1) $X_1, \ldots, X_n$, c.i.i.d. Ber(θ)

$E_\theta(\delta_1(X)) = E_\theta(X_1) = \theta$, $\forall \theta \in \Theta$. Logo, δ_1 é NÃO-VIESADO para θ.

$E_\theta(\delta_2(X)) = E_\theta(\bar{X}_n) = \theta$, $\forall \theta \in \Theta$. Logo, δ_2 é NÃO-VIESADO para θ.

$$E_\theta(\delta_3(X)) = E_\theta\left(\frac{a + \sum X_i}{a+b+n}\right) = \frac{a}{a+b+n} + \frac{E_\theta(\sum X_i)}{a+b+n} = \frac{a+n\theta}{a+b+n} \neq \theta.$$ Logo, δ_3 é viesado para θ. E,

$$B_\theta(\delta_3) = E_\theta(\delta_3(X)) - \theta =$$

$$= \frac{a+n\theta}{a+b+n} - \theta = \frac{a + \cancel{n\theta} - a\theta - b\theta - \cancel{n\theta}}{a+b+n} \Rightarrow$$

$$B_\theta(\delta_3) = \frac{a - a\theta - b\theta}{a+b+n}.$$

No mesmo exemplo, seja $g(\theta) = \theta^2$.

$$\delta_0(X) = X_1 . X_2$$

$$E_\theta[\delta_0(X)] = E_\theta(X_1 . X_2) = E_\theta(X_1)\, E_\theta(X_2) \Rightarrow$$

$$E_\theta(\delta_0(X)) = \theta^2, \ \forall \theta \in \Theta$$

Logo $\delta_0(X) = X_1 . X_2$ é não viesado para θ^2.

Exemplo 2. $X_1, \dots, X_n$, dado θ, c.i.i.d. $\mathcal{U}(0,\theta)$

$$E_\theta(X_1) = \frac{\theta}{2}$$

Definindo $\delta_1(X) = 2X_1$

$$E_\theta(\delta_1(X)) = E_\theta(2X_1) = 2E_\theta(X_1) = 2\cdot\frac{\theta}{2} = \theta, \forall \theta \in \Theta.$$

Logo, δ_1 é NÃO-VIESADO para θ.

$$\delta_2(X) = X_{(n)}$$

$$f_{X_{(n)}|\theta}(t) = \frac{n t^{n-1}}{\theta^n} \mathbb{I}_{(0,\theta)}(t)$$

$$E_\theta(\delta_2(X)) = E_\theta(X_{(n)}) = \int_0^\theta t\cdot\frac{n t^{n-1}}{\theta^n}\,dt = \frac{n}{\theta^n}\frac{t^{n+1}}{n+1}\Big|_0^\theta = \frac{n\,\theta^{n+1}}{(n+1)\theta^n} = \frac{n}{n+1}\theta, \text{ com}$$

$$B_\theta(\delta_2) = E_\theta(\delta_2(X)) - \theta = \frac{n}{n+1}\theta - \theta = -\frac{\theta}{n+1}$$ Logo, δ_2 é viesado para θ

Definindo $\delta_3(X) = \frac{n+1}{n} X_{(n)}$, temos

$$E_\theta(\delta_3(X)) = E_\theta\left(\frac{n+1}{n} X_{(n)}\right) = \frac{n+1}{n}\,\frac{n}{n+1}\cdot\theta, \forall \theta \in \Theta$$

Exemplo 3. $X_1, \dots, X_n$ dado θ, c.i.i.d $Exp(\theta)$

$$g(\theta) = e^{-\theta}$$

$$\delta_1(x) = \mathbb{I}_{(1,\infty)}(x_1)$$

$$E_\theta(\delta_1(X)) = 1\cdot P(\delta_1(X) = 1|\theta) + 0\cdot P(\delta_1(X) = 0|\theta) =$$

$$= P(X_1 > 1|\theta) = e^{-\theta\cdot 1} = e^{-\theta}, \forall \theta \in \Theta$$

Logo, $\delta_1(X)$ é não-viesado para $g(\theta) = e^{-\theta}$

Probabilidade e Inferência Estatística I

Luis Gustavo Esteves

23/05/2013 - Aula 08

Resultado:

Seja δ_0 um estimador não-viesado para $\theta(g(\theta))$

Seja $\Delta'' = \{\delta_0 + U : U \in \mathcal{U}\}$ onde

$$\mathcal{U} = \{U : \mathcal{X} \to \mathbb{R} : E_\theta(U(X)) = 0,\ \forall \theta \in \Theta\}$$

$\Delta'' = \Delta'$.

$$\delta \in \Delta'' \Rightarrow \delta = \delta_0 + U \Rightarrow E_\theta(\delta_0 + U) = E_\theta(\delta_0) + E_\theta(U) =$$
$$= g(\theta) + 0,\ \forall \theta \in \Theta$$

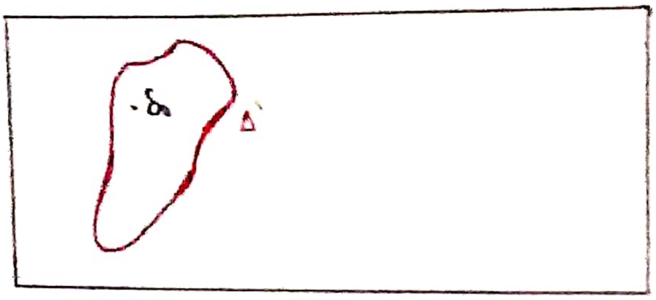
Logo, $\delta \in \Delta'$, $\Delta'' \subseteq \Delta'$.

$$\delta \in \Delta' \Rightarrow \delta = \delta + \delta_0 - \delta_0 = \delta_0 + \underbrace{\delta - \delta_0}_{U,\ \text{onde } E_\theta(U) = 0,\ \forall \theta \in \Theta} \Rightarrow \delta = \delta_0 + U,\ U \in \mathcal{U} \Rightarrow \delta \in \Delta''.$$

No nosso caso, vamos restringir a busca de um estimador ótimo a Δ'. Isso é equivalente a buscar um estimador não-viesado δ^+ tal que $Var_\theta(\delta^+) \leq Var_\theta(\delta)$, $\forall \theta \in \Theta$, $\forall \delta \in \Delta'$.

Definição: Dizemos que um estimador $\delta^* \in \Delta'$ é não-viesado de variância uniformemente mínima (ENVVUM) se $\forall \delta \in \Delta'$

$$Var_\theta(\delta^*) \leq Var_\theta(\delta),\ \forall \theta \in \Theta.$$



"Conseguimos aprimorar um estimador qualquer $\delta_0 \in \Delta'$ por meio da operação

$$\overset{\Delta}{E(\delta_0 | T)} = g(T)$$

depende de T. Como T é suficiente, segue que $E(\delta_0|T)$ não depende de θ.

E, tem var menor que δ_0

δ_0

$\delta^* = E(\delta_0|T)$. Processo de Rao-Blackwellização."

Teorema de Rao-Blackwell.

Seja $\delta_0 : \mathcal{X} \to \Theta$ um estimador não-viesado para θ ($g(\theta)$). Seja $g(\Theta)$

T uma estatística suficiente para θ.

Seja $\delta^* = E(\delta_0 | T)$.

Então

(1) δ^* depende de T

(2) $\delta^* \in \Delta'$

(3) $Var_\theta(\delta^*) \leq Var_\theta(\delta_0)$, $\forall \theta \in \Theta$

$$E_\theta(\delta^*(X)) = E_\theta(E(\delta_0(X) \mid T)) = E_\theta(\delta_0(X)) = \theta\ (g(\theta)),\ \forall \theta \in \Theta\ (g(\Theta)).$$

Logo, $\delta^* \in \Delta'$.

$$Var(\delta^*(X)) = Var_\theta(E_\theta(\delta_0(X) \mid T)) =$$

$$= Var_\theta(\delta_0(X)) - E_\theta(Var_\theta(\delta_0(X) \mid T)) \leq$$

$$\leq Var_\theta(\delta_0(X)),\ \forall \theta \in \Theta.$$

Exemplo 1. $X_1, \dots, X_n$ dado θ c.i.i.d. $Ber(\theta)$.

Seja $\delta_0(X) = X_1$

δ_0 é não-viesado pois

$$E_\theta(\delta(X)) = E_\theta(X_1) = \theta,\ \forall \theta \in \Theta$$

Seja $T = \sum_{i=1}^{n} X_i$ (suficiente para θ)

Para $t \in \{0, 1, \dots, n\}$, avaliamos

$$E_\theta(X_1 \mid \sum_{i=1}^{n} X_i = t) = 1 \cdot P_\theta(X_1 = 1 \mid \sum_{i=1}^{n} X_i = t) + 0 \cdot P_\theta(X_1 = 0 \mid \sum X_i = t) =$$

$$= \frac{P_\theta(X_1 = 1, \sum_{i=1}^{n} X_i = t)}{P_\theta(\sum_{i=1}^{n} X_i = t)} =$$

$$= \begin{cases} 0, & t = 0 \\ \dfrac{P_\theta(X_1 = 1, X_2 + \dots + X_n = t-1)}{P_\theta(\sum_{i=1}^{n} X_i = t)} = \dfrac{P_\theta(X_1 = 1) P_\theta(X_2 + \dots + X_n = t-1)}{P_\theta(X_1 + \dots + X_n = t)} = \dfrac{\theta \cdot \binom{n-1}{t-1} \theta^{t-1} (1-\theta)^{n-t}}{\theta^t (1-\theta)^{n-t}} = \end{cases}$$

$$= \frac{\binom{n-1}{t-1}}{\binom{n}{t}} = \frac{t}{n}$$

Logo, $\forall t \in \{0, 1, \ldots, n\}$.

$$E(X_1 / \sum_{i=1}^{n} X_i = t) = \frac{t}{n}$$

$$E(X_1 / \sum_{i=1}^{n} X_i) = \frac{\sum_{i=1}^{n} X_i}{n} = \delta^*(X)$$

Exercício: Verificar tbm p/ $\delta' = \frac{X_1 + X_2}{2}$.

Exemplo 2.

$X_1, \ldots, X_n$ dado θ, c.i.i.d. Poisson (θ)

$$g(\theta) = \theta e^{-\theta} = \frac{e^{-\theta}\theta^1}{1!} = P(X_1 = 1 / \theta)$$

$$\delta_0(X) = \mathbb{I}_{\{1\}}(X_1)$$

$$E_\theta(\delta_0(X)) = 1 \cdot P_\theta(X_1 = 1) + 0 \cdot P_\theta(X_1 = 0) = \frac{e^{-\theta} \cdot \theta^1}{1!} = \theta e^{-\theta} = g(\theta), \ \forall \theta \in \Theta.$$

Rao-Blackwellizando δ_0, temos

$$\delta^*(X) = E(\delta_0(X) / \sum_{i=1}^{n} X_i)$$

Para $t \in \mathbb{N}$,

$$E[\delta_0(X) / \sum_{i=1}^{n} X_i = t] = E_\theta[\mathbb{I}_{\{1\}}(X_1) / \sum_{i=1}^{n} X_i = t] = P_\theta(X_1 = 1 / \sum_{i=1}^{n} X_i = t) =$$

$$= \frac{P_\theta(X_1 = 1, \sum_{i=1}^{n} X_i = t)}{P_\theta(\sum_{i=1}^{n} X_i = t)} = \begin{cases} 0, & t = 0 \\ ?, & t \in \mathbb{N}^* \end{cases}$$

$$? = \frac{\mathbb{P}_\theta(X_1=1, X_2+\dots+X_n=t-1)}{\mathbb{P}_\theta(X_1+\dots+X_n=t)} = \frac{\mathbb{P}_\theta(X_1=1)\,\mathbb{P}_\theta(X_2+\dots+X_n=t-1)}{\mathbb{P}_\theta(X_1+\dots+X_n=t)} =$$

$$= \frac{\theta e^{-\theta} \cdot e^{-(n-1)\theta} \dfrac{(\theta(n-1))^{t-1}}{(t-1)!}}{\dfrac{e^{-n\theta} (n\theta)^t}{t!}} = \frac{(n-1)^{t-1}}{n^t} \cdot t$$

$\forall t \in \mathbb{N}$,

$$E(\delta_0(X) / \sum_{i=1}^{n} X_i = t) = \frac{(n-1)^{t-1}}{n^t} \cdot t$$

$$\delta^*(X) = E(\delta_0(X) / \sum_{i=1}^{n} X_i) = \frac{(n-1)^{\sum_{i=1}^{n} X_i - 1}}{n^{\sum_{i=1}^{n} X_i}} \cdot \sum_{i=1}^{n} X_i$$

Caso:

$n=2$

$X_1=1, X_2=0$ $\quad\quad \delta_0(1,0)=1$.

$\quad\quad\quad\quad\quad\quad\quad \delta^*(1,0) = \frac{1}{2}$.

Esse estimador ainda não é muito bom, pois $\delta^*(1,0) = 1/2 > \max_{\theta \in \Theta}(g(\theta)) = \frac{1}{e}$

pois $g(\Theta) = (0, e^{-1})$.

Na nossa definição, não é um estimador. Alguns livros o referenciam como estimador não-viesado ou estatística não viesada

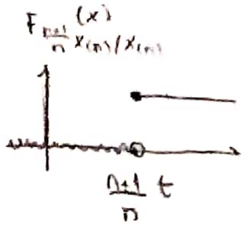
Exemplo 3. $X_1, \ldots, X_n$ dado θ, c.i.i.d. $\mathcal{U}(0,\theta)$

$\delta_1(X) = 2X_1$ é não-viesado para θ.

$\delta_2(X) = \frac{n+1}{n} X_{(n)}$ é não-viesado para θ.

$\delta_3(X) = 2\bar{X}_n$

Obs. $\delta_2^*(x) = E(\delta_2(X) / X_{(n)}) = \frac{n+1}{n} X_{(n)} = \delta_2(X)$

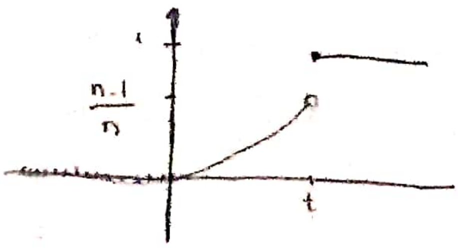


$E(2X_1 / X_{(n)} = t)$

$$F_{X_1/X_{(n)}=t}(x) = \begin{cases} \frac{n-1}{n}\frac{F_{X_1}(x)}{F_{X_1}(t)}, & x < t \\ 1, & x \geq t \end{cases}$$

No nosso exemplo.

$$F_\theta(x / X_{(n)} = t) = \begin{cases} 0 & x < 0 \\ \frac{n-1}{n}\frac{x/\theta}{t/\theta}, & 0 \leq x < t \\ 1, & x \geq t \end{cases}$$



$t>0$

$E(2X_1 / X_{(n)}=t) = 2E(X_1 | X_{(n)}=t) =$

$$= 2\left\{\left[\int_0^t x \cdot \frac{n-1}{n} \cdot \frac{1}{t}\,dx + t \cdot \frac{1}{n}\right]\right\} = 2\left\{\frac{n-1}{nt}\,\frac{t^2}{2} + \frac{t}{n}\right\} =$$

$$= 2\left(\frac{(n-1)t+2t}{2n}\right) \Rightarrow$$

$\Rightarrow E(2X_1 / X_{(n)}=t) = \frac{(n+1)}{n}t$. Logo $\delta^*_1 = E(2X_1 | X_{(n)}) = \frac{n+1}{n} X_{(n)} = \delta_2(X)$.

Exemplo 4.

θ brancas
$5-\theta$ verdes

$$X_i = \begin{cases} 1, & i\text{-ésima bola extraída é branca} \\ 0, & \text{c.c.} \end{cases}$$

$X = (X_1, X_2)$

$X_1/\theta \sim Ber\left(\frac{\theta}{5}\right)$

$E(X_1/\theta) = \frac{\theta}{5}$ $\qquad \delta_0(X_1,X_2) = 5X_1$

Exercício.

(1) Verifique que $T(X) = X_1+X_2$ é suficiente

(2) $E[\delta_0 / X_1+X_2] \overset{?}{=} \frac{5}{2}(X_1+X_2)$

Exemplo. $X_1,\dots,X_n$, dado θ, c.i.i.d. $\mathcal{U}(\theta,2\theta)$

$\Theta = \mathbb{R}_+ \quad \mathcal{X} = \mathbb{R}_+^n$

Verificar que $T(X) = (X_{(1)}, X_{(n)})$ é suficiente para θ.

$$U_i \overset{d}{=} \frac{X_i - \theta}{\theta} \,\Big|\, \theta \sim U(0,1)$$

$$E_\theta\left(\frac{X_{(n)} - \theta}{\theta}\right) = \frac{n}{n+1} \qquad E_\theta\left(\frac{X_{(1)} - \theta}{\theta}\right) = \frac{1}{n+1}$$

$U_{(n)} \sim Beta(1, n)$ $\qquad$ $U_{(1)} \sim Beta(n, 1)$

$$Var_\theta\left(\frac{X_{(1)} - \theta}{\theta}\right) = \frac{n}{(n+1)^2(n+2)} = Var_\theta\left(\frac{X_{(n)} - \theta}{\theta}\right)$$

$$E_\theta(X_{(n)}) = \theta + \theta\frac{n}{n+1} = \frac{2n+1}{n+1}\theta$$

$$E_\theta(X_{(1)}) = \theta + \theta\frac{1}{n+1} = \frac{n+2}{n+1}\theta$$

$$Var_\theta(X_{(1)}) = \frac{n\theta^2}{(n+1)^2(n+2)} = Var_\theta(X_{(n)})$$

Agora, podemos propor

$\delta_1(X) = \frac{n+1}{2n+1} X_{(n)}$ é não-viesado

$\delta_2(X) = \frac{n+1}{n+2} X_{(1)}$ é não-viesado

[illegible].

$E(X_1 / \theta) = \frac{3\theta}{2}$

$\delta_3(X) = \frac{2}{3} X_1$ é não-viesado

Agora

$$\delta_1^* = E[\delta_1 / X_{(1)}, X_{(n)}] = \delta_1(X)$$

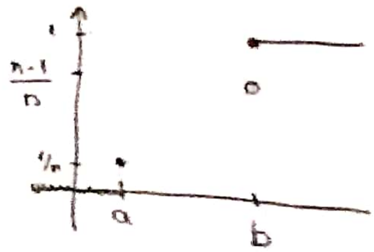
$$\delta_2^* = E[\delta_2 / X_{(1)}, X_{(n)}] = \delta_2(X)$$

$$\delta_3^* = E[\delta_3 / X_{(1)}, X_{(n)}] = \frac{X_{(1)} + X_{(n)}}{3}$$

$\downarrow$

Ideia.

$$F_{X_1 | X_{(1)} = a, X_{(n)} = b}(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{n}, & x = a \\ \frac{1}{n} + \frac{n-2}{n} \frac{F(x) - F(a)}{F(b) - F(a)}, & a < x < b \\ 1, & x \geq b \end{cases}$$



$$\delta_4^*(x) = \frac{n+1}{5n+4}(2X_{(n)} + X_{(1)})$$

$$Var_\theta \delta_4^* \leq Var_\theta \delta_3^* \leq Var_\theta \delta_1^* \leq Var_\theta \delta_2^*$$

$\delta_1 \to \delta_1^*$

$\delta_2 \to \delta_2^*$

Melhoramos os estimadores, mas não chegamos à um único δ^*

(família de distribuição de P_T)

Dizemos que a estatística $T: \mathfrak{X} \to \mathbb{R}^k$ é completa se $\forall h: \mathbb{R}^k \to \mathbb{R}$,

$$E_\theta(h(T(X))) = 0, \forall \theta \in \Theta \Rightarrow h \equiv 0.$$

$(\mathfrak{X}, \sigma, \{P_\theta \ \theta \in \Theta\})$ $\quad$ $\mathcal{P} = \{f_T(\cdot|\theta) : \theta \in \Theta\}$

$f(\cdot|\theta)$

$f_T(\cdot|\theta)$

Exemplo: $X_1, \dots, X_n$, dado θ, c.i.i.d $Ber(\theta)$

$\Theta = [0,1]$

$T(X) = \sum_{i=1}^{n} X_i$

Seja $h: \mathbb{R} \to \mathbb{R}$

$E_\theta(h(T(X))) = 0, \forall \theta \in \Theta \Rightarrow$

$E_\theta(h(\sum_{i=1}^{n} X_i)) = 0, \forall \theta \in \Theta \Rightarrow \sum_{i=0}^{n} h(i) \binom{n}{i} \theta^i (1-\theta)^{n-i} = 0, \forall \theta \in (0,1) \Rightarrow$

$$(1-\theta)^n \sum_{i=0}^{n} h(i) \binom{n}{i} \left(\frac{\theta}{1-\theta}\right)^i = 0, \forall \theta \in (0,1) \Rightarrow$$

$$\Rightarrow \sum_{i=0}^{n} h(i) \binom{n}{i} \left(\frac{\theta}{1-\theta}\right)^i = 0, \forall \theta \in \Theta$$

3^i

$\Rightarrow h(i) \binom{n}{i} = 0, \forall i = 0,1,\dots,n \Rightarrow h(i) = 0, \forall i = 0,1,\dots,n$

$\Rightarrow h \equiv 0 \; P_\theta$ - q.c.

Logo $T(X) = \sum_{i=1}^{n} X_i$ é completa.

Exemplo 2. $X_1, \dots, X_n$, dado θ, c.i.i.d. $\mathcal{U}(0,\theta)$

$T(X) = X_{(n)}$ $\quad h: \mathbb{R}_+ \to \mathbb{R}_+$

$E_\theta(h(T(X))) = 0, \forall \theta \in \Theta \Rightarrow E_\theta(h(X_{(n)})) = 0, \forall \theta > 0,$

$$\int_{-\infty}^{\infty} h(t)\, f_{X_{(n)}}(t|\theta)\, dt = 0, \ \forall \theta \in \Theta \Rightarrow$$

$$\Rightarrow \int_0^\theta h(t) \cdot \frac{n t^{n-1}}{\theta^n}\, dt = 0, \ \forall \theta > 0 \Rightarrow \frac{1}{\theta^n} \int_0^\theta h(t)\, n t^{n-1}\, dt = 0, \ \forall \theta > 0 \Rightarrow$$

$$\underbrace{\int_0^\theta h(t)\, n t^{n-1}}\, dt = 0, \ \forall \theta > 0 \Rightarrow$$

$v(\theta)$ é integrável, e portanto, diferenciável. Com $\int_0^\theta v(\theta)\, dt = 0 \ \forall \theta \in \Theta$ segue que

$$\frac{d}{dt} \int_0^\theta v(\theta)\, dt = 0 \quad \forall \theta \in \Theta \Rightarrow$$

$$\Rightarrow h(\theta)\, n\, \theta^{n-1} = 0, \ \forall \theta > 0 \Rightarrow h(t) = 0, \ \forall t > 0.$$

Logo, $h \equiv 0$ P_θ-q.c.

Logo, T é completa!

Exemplo 3. $X_1, \dots, X_n$, dado θ, c.i.i.d. Poisson(θ)

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

$h: \mathbb{N} \to \mathbb{R}$

$$E_\theta(h(T(X))) = 0, \ \forall \theta > 0 \Rightarrow E_\theta\left(h\left(\sum_{i=1}^n X_i\right)\right) = 0, \ \forall \theta > 0 \Rightarrow$$

$$\sum_{i=0}^{\infty} \frac{h(i)\, e^{-n\theta} (n\theta)^i}{i!} = 0, \ \forall \theta > 0 \Rightarrow \sum_{i=0}^{\infty} \frac{h(i)\, n^i}{i!} = 0, \ \forall \theta > 0 \Rightarrow$$

$$\frac{h(i)\, n^i}{i!} = 0, \ \forall i \in \mathbb{N} \Rightarrow h(i) = 0, \ \forall i \in \mathbb{N} \Rightarrow$$

$h \equiv 0$.

Exemplo 4. $X_1, \dots, X_n$, dado θ, c.i.i.d. $\mathcal{U}(\theta, 2\theta)$

$$E_\theta(X_{(1)}) = \frac{n+2}{n+1}\theta$$

$$T(X) = (X_{(1)}, X_{(n)})$$

$$E_\theta(X_{(n)}) = \frac{2n+1}{n+1}\theta$$

$$h: \mathbb{R}^2 \to \mathbb{R}$$

$$(u,v) \mapsto h(u,v) = \frac{n+1}{n+2}u - \frac{n+1}{2n+1}v$$

$$E_\theta(h(T(X))) = E_\theta(h(X_{(1)}, X_{(n)})) = E_\theta\left(\frac{n+1}{n+2}X_{(1)} - \frac{n+1}{2n+1}X_{(n)}\right)$$

$$= \frac{n+1}{n+2}E_\theta(X_{(1)}) - \frac{n+1}{2n+1}E_\theta(X_{(n)}) = \frac{n+1}{n+2}\cdot\frac{n+2}{n+1}\theta - \frac{n+1}{2n+1}\cdot\frac{2n+1}{n+1}\theta = 0, \forall \theta \in \Theta$$

Logo, $T(X) = (X_{(1)}, X_{(n)})$ não é completa!

Prob. E. Inf. Est. I

Aula 09 - 24/05/2013

Luis Gustavo Esteves

No 2º semestre

Disciplina: Probabilidade e Inferência Estatística

e NÃO Teoria da Decisão

$T: \mathfrak{X} \to \mathbb{R}^k$ é completa se

$\forall\, h: \mathbb{R}^k \to \mathbb{R}$

$E_\theta(h(T(X))) = 0, \forall \theta \in \Theta \Rightarrow h \equiv 0 \;\; P_\theta\text{-q.c.} \; \forall \theta \in \Theta.$

Exemplo 5.a.

$\Theta = \{1, 2, 3, 4, \ldots\} = \mathbb{N}^*$

$X|\theta \sim U(\{1, \ldots, \theta\})$

$P(X = x|\theta) = \frac{1}{\theta} \mathbb{I}_{\{1,\ldots,\theta\}}(x)$

$h: \mathbb{N} \to \mathbb{R}$

$E_\theta(h(X)) = 0, \forall \theta \in \Theta \Rightarrow$

$E_1(h(X)) = h(1) \overbrace{P(X=1|\theta=1)}^{1} = 0 \Rightarrow h(1) = 0,$

$E_2(h(X)) = h(1)\frac{1}{2} + h(2)\cdot\frac{1}{2} = 0 \Rightarrow h(2) = 0,$

$E_3(h(X)) = h(1)\cdot\frac{1}{3} + h(2)\frac{1}{3} + h(3)\cdot\frac{1}{3} = \Rightarrow h(3) = 0,$

$\vdots$

$\Rightarrow h(j)=0, \forall j \in \mathbb{N}^*$ e, portanto,

$$h \equiv 0 \quad P_\theta\text{-q.c. } \forall \theta \in \mathbb{N}^*$$

Logo, $T(X)=X$ é completa!

5b) $n_0 \in \mathbb{N}^*$

$$\Theta = \mathbb{N}^* - \{n_0\}$$

$$h(t) = \begin{cases} a, & t = n_0 \\ -a, & t = n_0+1 \\ 0, & \text{c.c.} \end{cases}$$

$$E_\theta(h(T)) = \begin{cases} \sum_{t=0}^{\theta} \frac{1}{\theta} \underbrace{h(t)}_{0} = 0, & \theta \leq n_0 - 1 \\ \sum_{t=1}^{\theta} \frac{1}{\theta} h(t) = \sum_{t=1}^{\theta-1} \frac{1}{\theta} \cdot 0 + \frac{1}{\theta}(+a) + \frac{1}{\theta}(-a) + \sum_{t=n_0+2}^{\theta} \frac{1}{\theta} \cdot 0 = 0 \\ \sum_{t=1}^{n-1} \frac{1}{\theta} \cdot 0 + \frac{1}{\theta}(a) + \frac{1}{\theta}(-a) = 0 \end{cases}$$

$\exists\, h: \mathbb{N} \to \mathbb{R}$, h não identicamente nula tal que $E_\theta(h(X))=0, \forall \theta \in \Theta$.

Logo, X não é completa!

$\mathcal{P}' = \{ f(\cdot/\theta) : \theta \in \mathbb{N}^* - \{n_0\}\}$ NÃO é completa!

Em geral, sejam Θ e Θ'. $T: \mathfrak{X} \to \mathbb{R}^k$.

Seja $A_\Theta = \{ l : \mathbb{R}^k \to \mathbb{R} : E_\theta(l(T(X))) = 0, \forall \theta \in \Theta \}$.

Note que $l_0 \equiv 0 \in A_\Theta$. ($l_0 \in A_{\Theta'}$)

E

$$A_\Theta \supseteq A_{\Theta'} \quad \text{onde} \quad \Theta \subseteq \Theta'$$

S. Stigler - Completeness and Unbiasedness. The American Statistician.

Teorema de Lehmann-Scheffé.

Seja $T: \mathfrak{X} \to \mathbb{R}^k$ uma estatística suficiente e completa. Então todos os estimadores não-viesados de θ ($g(\theta)$) que são função de T (dependem de X através de T) são iguais P_θ-q.c. $\forall \theta \in \Theta$.

Assim, se $\delta' \in \Delta'$ é função de uma estatística suficiente e completa, δ' é o ENVVUM para θ.

Sejam $\delta_1 = \delta_1(T) \in \Delta'$

e $\delta_2 = \delta_2(T) \in \Delta'$

$E_\theta(\delta_1(T(X)) - \delta_2(T(X))) =$

$= \theta - \theta = 0, \forall \theta \in \Theta$

$\Downarrow$

$\delta_1(T(X)) = \delta_2(T(X))$, P_θ-q.c. $\forall \theta \in \Theta$

$\delta', \delta'' \in \Delta$

$E(\delta' | T)$,

$E(\delta'' | T)$) =

Exemplo:

(1) $X_1, \ldots, X_n$, dado θ, c.i.i.d. $Ber(\theta)$

$$\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n} = g\left(\sum_{i=1}^{n} X_i\right)$$

$$\bar{X}_n \in \Delta'$$

Pelo Teorema, $\bar{X}_n$ é o ENVVUM para θ.

Exemplo: $X_1, \ldots, X_n$, dado θ, c.i.i.d. $U(0,\theta)$

$$\delta^*(X) = \frac{n+1}{n} X_{(n)} = g(X_{(n)})$$

$\delta^* \in \Delta'$ e é função de $X_{(n)}$, estatística suficiente e completa.

Logo, δ^* é ENVVUM para θ.

Exemplo 3: $X_1, \ldots, X_n$, dado θ, c.i.i.d. Poisson (θ)

$$g(\theta) = \theta e^{-\theta}$$

$$\delta_0(X) = I_{\{1\}}(X_1) \rightarrow \delta^*(X) = E\left[\delta_0(X) \mid \sum_{i=1}^{n} X_i\right] = \frac{\sum_{i=1}^{n} X_i}{n} \left(\frac{n-1}{n}\right)^{\sum X_i - 1}$$

Pelo Teorema, δ^* é o ENVVUM para $\theta e^{-\theta}$.

Exemplo 2

$$\delta^*(X) = \frac{n+1}{n} X_{(n)} \text{ é o ENVVUM para } \theta$$

$$\delta''(X) = \frac{n+2}{n+1} X_{(n)} \qquad E_\theta\left[(\alpha X_{(n)} - \theta)^2\right] \rightarrow \alpha = \frac{n+2}{n+1}$$

$\downarrow$

~~max~~ minimizar isso

Ex. 3: $X_1, \ldots, X_n$, dado $\theta = (\mu, \sigma^2)$, i.i.d. $N(\mu, \sigma^2)$

$$S_n^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1} \quad \text{é o ENVVUM para } g(\theta) = \sigma^2$$

$$E_\theta\left[\left(\alpha \sum_{i=1}^{n} (X_i - \bar{X})^2 - \sigma^2\right)^2\right] \to \alpha = \frac{1}{n+1}.$$

Resultado:

δ^* é o ENVVUM para θ (ou $E_\theta(\delta^*(X))$) se e só se

$$Cov_\theta(\delta^*(X), U(X)) = 0, \ \forall \theta \in \Theta, \ \forall U \in \mathcal{U},$$

onde $\mathcal{U} = \{U : \mathcal{X} \to \mathbb{R} : E_\theta(U(X)) = 0, \forall \theta \in \Theta\}$.

Ex. $X_1, \ldots, X_n$, dado θ, i.i.d. $U(\theta, 2\theta)$, $\theta > 0$.

$$\delta_1(X) = \frac{n+1}{2n+1} X_{(n)} \to \delta_1^*(X)$$

$$\delta_2(X) = \frac{n+1}{n+2} X_{(1)} \to \delta_2^*(X)$$

$$\delta_3(X) = \frac{2X_1}{3} \to \delta_3^* = \frac{X_{(1)} + X_{(n)}}{3}$$

$$Cov_\theta\left(\frac{X_{(1)} + X_{(n)}}{3}, \overbrace{\frac{n+1}{2n+1} X_{(n)} - \frac{n+1}{n+2} X_{(1)}}^{U \in \mathcal{U}}\right)$$

$$= \frac{1}{3}\frac{n+1}{2n+1} Cov_\theta(X_{(1)}, X_{(n)}) - \frac{1}{3}\frac{n+1}{n+2} Cov_\theta(X_{(1)}, X_{(1)}) + \frac{1}{3}\frac{n+1}{2n+1} Cov_\theta(X_{(n)}, X_{(n)}) - \frac{1}{3}\frac{n+1}{n+2} Cov_\theta(X_{(n)}, X_{(1)})$$

$$= \frac{1}{3} Cov_\theta(X_{(1)}, X_{(n)}) \left(\frac{n+1}{2n+1} - \frac{n+1}{n+2} \right) + \frac{1}{3} VAR_\theta(X_{(n)}) \left(\frac{n+1}{2n+1} - \frac{n+1}{n+2} \right) =$$

$$= \frac{1}{3} \left(\frac{n+1}{2n+1} - \frac{n+1}{n+2} \right) \left\{ \underbrace{Cov_\theta(X_{(1)}, X_{(n)})}_{= \frac{\theta^2}{(n+2)(n+1)^2}} + \underbrace{Var_\theta(X_{(n)})}_{= \frac{n\theta^2}{(n+2)(n+1)^2}} \right\} < 0, \ \forall \theta \in \Theta.$$

Logo, $\delta_3^*(X) = \frac{X_{(1)} + X_{(n)}}{3}$ não é ENVVUM!

Desigualdade da Informação (Cramer-Rao)

Seja δ um estimador não-viesado para $g(\theta)$.

Sob certas condições (condições de regularidade), vale que

$VAR_\theta(\delta(X)) \geq \frac{(g'(\theta))^2}{I_X(\theta)}$, onde $I_X(\theta)$ é chamada Informação de Fisher para θ contida na amostra,

$I_X(\theta) = E_\theta[(\lambda'(X/\theta))^2]$, onde

$\lambda(x/\theta) = \log f(x/\theta)$

$(\mathfrak{X}, \sigma, \mathcal{P}) \quad \mathcal{P} = \{ f(\cdot|\theta) : \theta \in \Theta \}$

$\lambda(x/\theta) = \log f(x/\theta)$

$\lambda'(x/\theta) = \frac{d}{d\theta} \log(f(x/\theta)) = \frac{d}{d\theta} \lambda(x|\theta) \to$ FUNÇÃO ESCORE.

Existe também $\lambda''(x/\theta)$

Informação de Fisher

$$E_\theta[(\lambda'(X/\theta))^2].$$

Exemplo: $X_1,\dots,X_n$ dado θ c.i.i.d. Poisson(θ).

$$f(x/\theta) = \mathbb{P}(X = \underset{\substack{|\\ x=(x_1,\dots,x_n)\in\mathbb{N}^n}}{x}/\theta) =$$

$$= \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\,\theta^{\sum x_i}}{\prod_{i=1}^{n} x_i!}$$

$$\lambda(x/\theta) = \log f(x/\theta) = -n\theta + \sum_{i=1}^{n} x_i \log\theta - \log\prod_{i=1}^{n} x_i!$$

$$\lambda'(x/\theta) = -n + \frac{\sum_{i=1}^{n} x_i}{\theta}.$$

$\Downarrow$

$$\lambda'(X/\theta) = -n + \frac{\sum_{i=1}^{n} X_i}{\theta}$$

$$I_X = E_\theta[(\lambda'(X/\theta))^2] = E_\theta\left(\left(\frac{\sum X_i}{\theta} - n\right)^2\right) = \frac{1}{\theta^2}E_\theta\left[\left(\sum_{i=1}^{n} X_i - n\theta\right)^2\right] = \frac{1}{\theta^2}\mathrm{Var}_\theta\left(\sum_{i=1}^{n} X_i\right) =$$

$$= \frac{n\theta}{\theta^2} \Rightarrow I_X(\theta) = \frac{n}{\theta}.$$

Exemplo: $X_1,\dots,X_n$, dado θ, c.i.i.d. Ber(θ)

$$f(x/\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$$

$$\lambda(x/\theta) = \log f(x/\theta) = \sum x_i \log\theta + (n - \sum x_i)\log(1-\theta)$$

$$\lambda'(x/\theta) = \frac{\sum x_i}{\theta} - \frac{(n-\sum x_i)}{1-\theta} \Rightarrow$$

$$\lambda'(x/\theta) = \frac{\sum x_i (1-\theta) - \theta(n-\sum x_i)}{\theta(1-\theta)} = \frac{\sum x_i - n\theta}{\theta(1-\theta)} \Rightarrow$$

$$\lambda'(X/\theta) = \frac{\sum_{i=1}^{n} X_i - n\theta}{\theta(1-\theta)}$$

$$I_X(\theta) = E_\theta\left[(\lambda'(X/\theta))^2\right] = E_\theta\left[\left(\frac{\sum X_i - n\theta}{\theta(1-\theta)}\right)^2\right] = \frac{Var_\theta(\sum X_i)}{\theta^2(1-\theta)^2} = \frac{n}{\theta(1-\theta)}$$

Em geral, $X_1, \dots, X_n$ dado θ, são c.i.i.d

$$I_X(\theta) = n I_{X_1}(\theta)$$

[redacted]

$$\int_X \delta(x) f(x/\theta)\, dx = g(\theta), \ \forall \theta \in \Theta \Rightarrow$$

$$\Rightarrow \frac{d}{d\theta}\int_X \delta(x) f(x/\theta)\, dx = g'(\theta), \ \forall \theta \in \Theta$$

$\Rightarrow$ [redacted] $$\int_X \delta(x) \frac{\frac{df(x|\theta)}{d\theta}}{f(x|\theta)} f(x/\theta)\, dx = g'(\theta) \Rightarrow$$

$$\int_{\mathcal{X}} \underbrace{\delta(x)\, \lambda'(x/\theta)}\cdot f(x/\theta)\, dx = g'(\theta)$$

$$\Rightarrow E_\theta(\delta(x)\,\lambda'(x/\theta)) = g'(\theta). \qquad (I)$$

$$\int_{\mathcal{X}} f(x/\theta)\, dx = 1 \ , \forall \theta \in \Theta$$

$$\Rightarrow \frac{d}{d\theta}\int_{\mathcal{X}} f(x|\theta)\, dx = 0 \ , \ \forall \theta \in \Theta$$

$$\Rightarrow \int_{\mathcal{X}} \frac{\frac{d}{d\theta} f(x/\theta)}{f(x/\theta)}\, f(x/\theta)\, dx = 0 .$$

$$\Rightarrow E_\theta(\lambda'(x/\theta)) = 0 .$$

Logo, $Var_\theta(\lambda'(X/\theta)) = E_\theta((\lambda'(X/\theta))^2) = I_X(\theta)$.

Daí,

$$Cov_\theta(\delta(X), \lambda'(X/\theta)) = E_\theta(\delta(X)\,\lambda'(X/\theta)) - 0 \cdot g(\theta)$$

$$[CORR_\theta(\delta(X), \lambda'(x/\theta)]^2 \leq 1 \Rightarrow$$

$$\frac{(COV(\delta(X), \lambda'(X/\theta)))^2}{VAR_\theta(\delta(X))\, VAR_\theta(\lambda'(X/\theta))} \leq 1 \Rightarrow Var_\theta(\delta(X)) \geq \frac{(g'(\theta))^2}{I_X(\theta)} .$$

Exemplo: $X_1, \dots, X_n$ dado θ, c.i.i.d $Ber(\theta)$.

$g(\theta) = \theta$

$$LI_\theta = \frac{1^2}{\frac{n}{\theta(1-\theta)}} = \frac{\theta(1-\theta)}{n}$$

$$\delta^*(X) = \bar{X}_n \Rightarrow Var(\bar{X}_n) = n\frac{\theta(1-\theta)}{n^2} = \frac{\theta(1-\theta)}{n} = LI_\theta$$

$\therefore$ A média, no caso das Bernoullis, atinge o Limite Inferior de Cramer-Rao.

Exemplo 2:

$X_1, \ldots, X_n$, dado θ, c.i.i.d. Poisson(θ)

$g(\theta) = e^{-\theta}$

Exercício

$$\delta_0(X) = \mathbb{I}_{\{0\}}(X) \xrightarrow{\text{Rao-Blackwelliza}} \overbrace{E[\delta_0 \mid \Sigma X_i]}^{\delta^*} = \left(\frac{n-1}{n}\right)^{\sum_{i=1}^n X_i}$$

Pelo Teorema de Lehmann-Scheffé, $\delta^*(X) = \left(\frac{n-1}{n}\right)^{\sum_{i=1}^n X_i}$ é o ENVVUM para $e^{-\theta}$.

$$Var(\delta^*(X)) = E_\theta((\delta^*(X))^2) - (E_\theta(\delta^*(X)))^2 =$$

$$= E_\theta\left(\left(\frac{n-1}{n}\right)^{2\Sigma X_i}\right) - e^{-2\theta}$$

$$= \left[\sum_{j=0}^{\infty} \frac{e^{-n\theta}(n\theta)^j}{j!}\left(\frac{n-1}{n}\right)^{2j}\right] - e^{-2\theta}$$

$$= e^{-n\theta}\sum_{j=0}^{\infty}\frac{\left(\frac{n\theta(n-1)^2}{n^2}\right)^j}{j!} - e^{-2\theta} = e^{\frac{\theta(n-1)^2}{n} - n\theta} - e^{-2\theta} =$$

$$= e^{\frac{\theta(-2n+1)}{n}} - e^{-2\theta} \Rightarrow Var_\theta(\delta^*(X)) = e^{-2\theta}(e^{\theta/n} - 1)$$

$$LI_\theta = \frac{(-e^{-\theta})^2}{n/\theta} = e^{-2\theta} \cdot \frac{\theta}{n}$$

$e^x - 1 > x$

Logo, $Var_\theta(\delta^*(X)) > LI_\theta$.

θ: ? $\theta \in (0,1)$

X: Numero de sucessos em 10 tentativas

$X|\theta \sim Bin(10,\theta)$ $\qquad \frac{X}{10} = \delta_1(x)$ (

Y: Entrevista individuos até obter o 2º sucesso.

$Y/\theta \sim BINNEG(2,\theta)$ $\qquad \delta_2(Y) = \frac{2-1}{Y-1}$

- Se $X \sim BINNEG(k,\theta)$

$$E\left(\frac{k-1}{X-1} \middle| \theta\right) = \theta$$

$$\sum_{j=k}^{\infty} \frac{k-1}{j-1} \binom{j-1}{k-1} \theta^k (1-\theta)^{j-k} = \qquad = \theta$$

N N S N N N N N N S

$$\delta_1(x) = \frac{2}{10} \; (x=2) \quad e \quad \delta_2(x) = \frac{1}{9}$$

$E_\theta[\delta(X)] = \theta$

PRINCIPIO DA VEROSSIMILHANÇA.

"Toda a inferência deve ser baseada apenas no ponto amostral $x \in \mathfrak{X}$ efetivamente observado".

Exp. 1 $V_x^{(1)}(\theta) = k(x,y)\, V_y^{2}(\theta), \forall \theta \in \Theta$

$\Downarrow$

as conclusões a partir de x e y devem coincidir.

Exp. 2.

- Dissertação de mestrado (1995)
  Lourdes Yoshiko T. Inoue

- Princípio da Verossimilhança
  "The Likelihood Principle". Berger e Wolpert (1988)

- "Breakthroughs in Statistics"
  Kotz, Johnson.
  Birnbaum (1962)

Aula 10

28/05/2013 .

Ex 3. $X_1, \dots, X_n$ dado θ, são c.i.i.d. Geo(θ).

$$g(\theta) = \frac{1}{\theta}$$

$$\delta^*(x) = \bar{X}_n$$

$$Var_\theta(\delta^*(X)) = Var_\theta(\bar{X}_n) = \frac{1}{n} Var(X_1) = \frac{1-\theta}{n\theta^2}$$

$$f(x/\theta) = \prod_{i=1}^{n} (1-\theta)^{x_i - 1} \cdot \theta = (1-\theta)^{\sum_{i=1}^{n} x_i - n} \, \theta^n$$

$$\lambda(x/\theta) = (\Sigma x_i - n)\log(1-\theta) + n\log\theta$$

$$\lambda'(x/\theta) = -\frac{\sum_{i=1}^{n} x_i - n}{1-\theta} + \frac{n}{\theta}$$

$$I_X(\theta) = E_\theta\left[\left(\frac{-\Sigma X_i + n}{1-\theta} + \frac{n}{\theta}\right)^2\right] = \frac{1}{\theta^2(1-\theta)^2} E_\theta\left[\left(-\theta\Sigma X_i + \cancel{n\theta} + n - \cancel{n\theta}\right)^2\right] =$$

$$= \frac{1}{\theta^2(1-\theta)^2} E_\theta((\theta\Sigma X_i - n)^2) = \frac{\theta^2}{\theta^2(1-\theta)^2} E_\theta\left(\left(\Sigma X_i - \frac{n}{\theta}\right)^2\right) = \frac{1}{(1-\theta)^2} n \frac{(1-\theta)}{\theta^2} =$$

$$= \frac{n}{\theta^2(1-\theta)}$$

$$LI(\theta) = \frac{(g'(\theta))^2}{I_x(\theta)} = \frac{(-1/\theta^2)^2}{\frac{n}{\theta^2(1-\theta)}} = \frac{1-\theta}{n\theta^2} = Var_\theta(\delta^*(x))$$

Logo, $\delta^*(x)$ é o ENVVUM para $1/\theta$.

* forma alternativa

$$E_\theta(\lambda'(X))=0 \Rightarrow \int_{\mathcal{X}} \lambda'(x/\theta) f(x/\theta)\,dx = 0 \Rightarrow \int_{\mathcal{X}} \frac{d(\lambda'(x/\theta) f(x/\theta))}{d\theta}\,dx = 0 \Rightarrow$$

$\lambda'(x/\theta)$

$$\int_{\mathcal{X}} \left\{ \lambda''(x/\theta) f(x/\theta) + \lambda'(x/\theta) \frac{\frac{df(x/\theta)}{d\theta}}{f(x/\theta)} f(x/\theta)\,dx = 0 \Rightarrow \right.$$

$$\Rightarrow E_\theta(\lambda''(x/\theta)) + \underbrace{E_\theta((\lambda'(x/\theta))^2)}_{I_x(\theta)} = 0 \Rightarrow I_X(\theta) = -E_\theta[\lambda''(X/\theta)].$$

Propriedades Assintóticas de Estimadores

Ter em mente sequência de v.a. $(X_n)_{n\geq 1}$.

$X = (X_n)_{n\geq 1}$

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

$(\delta_n)_{n\geq 1}$ sequência de estimadores

Definição: $(\delta_n)_{n\geq 1}$ é uma sequência consistente para θ $(g(\theta))$ se

$$\delta_n(X) \xrightarrow{P_\theta} \theta\ (g(\theta)),\ \forall \theta \in \Theta$$

Exemplo: $(X_n)_{n\geq 1}$, dado θ c.i.i.d. $\mathcal{U}(0,\theta)$.
$(x_1, \ldots, x_n)$

$\delta^*_n(X) = X_{(n)}$

$\delta_n(x) = \frac{n+1}{n} X_{(n)}$

Para $\varepsilon > 0$,

$$P_\theta(|\delta_n(X) - \theta| > \varepsilon) = P_\theta(|X_{(n)} - \theta| > \varepsilon) = P_\theta(X_{(n)} < \theta - \varepsilon) + \underbrace{P_\theta(X_{(n)} > \theta + \varepsilon)}_{0} =$$

$= \mathbb{P}_\theta(X_1 < \theta - \varepsilon, \ldots, X_n < \theta - \varepsilon)$

$0 < \varepsilon < \theta$

$= \left(\frac{\theta - \varepsilon}{\theta}\right)^n \xrightarrow[n \to \infty]{} 0$

Logo $(\delta_n^*)_{n \geq 1}$ é uma sequência consistente para θ.

$$\underbrace{\delta_n'(x) = \underbrace{\frac{n+1}{n}}_{\text{seq. converge para } 1} \underbrace{\delta_n^*(x)}_{\text{seq. conv. para } \theta}} \xrightarrow{\mathbb{P}_\theta} 1 \cdot \theta, \ \forall \theta \in \Theta$$

Logo, $(\delta_n')_{n \geq 1}$ também é consistente para θ.

Exemplo 2. $(X_n)_{n \geq 1}$, dado θ, c.i.i.d. $Ber(\theta)$

$$\bar{X}_n \xrightarrow{\mathbb{P}_\theta} \theta, \forall \theta \in \Theta$$

$$1 - \bar{X}_n \xrightarrow{\mathbb{P}_\theta} 1 - \theta, \forall \theta \in \Theta$$

$$\bar{X}_n(1 - \bar{X}_n) \xrightarrow{\mathbb{P}_\theta} \theta(1-\theta), \forall \theta \in \Theta$$

Assim,

$(\bar{X}_n(1-\bar{X}_n))_{n \geq 1}$ é consistente para $\theta(1-\theta)$.

DEFINIÇÃO: $(\delta_n)_{n \geq 1}$ é ASSINTOTICAMENTE NORMAL $(0, \sigma^2(\theta))$ se

$$\sqrt{n}(\delta_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta)), \forall \theta \in \Theta$$

$$\left(\text{para } n \text{ "grande" } \delta_n \overset{a}{\approx} N\left(\theta, \frac{\sigma^2(\theta)}{n}\right)\right)$$

Exemplo 1. $X_1, \ldots, X_n$, dado θ, c.i.i.d. $Ber(\theta)$.
$(X_n)_{n \geq 1}$

$$\frac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \xrightarrow{d} N(0,1)$$

$$\|$$

$$\frac{n\bar{X}_n - n\theta}{\sqrt{n}\sqrt{\theta(1-\theta)}}$$

$$\|$$

$$\frac{\sqrt{n}(\bar{X}_n-\theta)}{\sqrt{\theta(1-\theta)}} \xrightarrow{d} N(0,1) \Rightarrow \sqrt{n}(\bar{X}_n-\theta) \xrightarrow{d} N(0,\theta(1-\theta))$$

Assim, $(\bar{X}_n)_{n\geq 1}$ é assintoticamente Normal $(0, \theta(1-\theta))$.

Exemplo. $X_1, \dots, X_n$, dado θ, c.i.i.d. tais que $E(X_1|\theta)=\theta$ e $Var(X_1/\theta)=\sigma_0^2$, σ_0^2 conhecido!

$$\frac{\bar{X}_n-\theta}{\sigma_0/\sqrt{n}} \xrightarrow{d} N(0,1) \Rightarrow \frac{\sqrt{n}(\bar{X}_n-\theta)}{\sigma_0} \xrightarrow{d} N(0,1) \Rightarrow$$

$$\sqrt{n}(\bar{X}_n-\theta) \xrightarrow{d} N(0,\sigma_0^2)$$

Assim, $(\bar{X}_n)_{n\geq 1}$ é Assintoticamente Normal $(0,\sigma_0^2)$.

Exemplo...

$\sigma_0^2 = 1$

Para cada $n \in \mathbb{N}$

seja $Md_n(X) = \text{mediana}\{X_1, \dots, X_n\}$

$(Md_n(X))_{n\geq 1}$ é assintoticamente NORMAL $(0, \pi/2)$.

$$S_n = \sum \mathbb{I}_{\{\theta+t/\sqrt{n}, \infty\}}(X_i)$$

$$P(\sqrt{n}(Md(X)-\theta) \leq t) = P(Md_n(X) \leq \theta + t/\sqrt{n}) = P(S_n \leq [\tfrac{n}{2}])$$

$$S_n/\theta \sim Bin(n, P_\theta(X_1 > \theta + t/\sqrt{n}))$$

$$P_\theta(S_n \leq n/2) = P_\theta\left(\frac{S_n - nP_\theta(X_1 > \theta + t/\sqrt{n})}{\sqrt{nP_\theta(X_1 > \theta + t/\sqrt{n})(1-P_\theta(X_1 > \theta+t/\sqrt{n}))}} \leq \frac{n}{2} - nP_\theta(X_1 > \theta + t/\sqrt{n})\right)$$

$$\approx P\left(Z \leq \frac{\frac{1}{2} - P_\theta(X_1 > \theta + t/\sqrt{n})}{\sqrt{\frac{P_\theta(X_1 > \theta+t/\sqrt{n})(1-P_\theta(X_1 > \theta+t/\sqrt{n}))}{n}}}\right) \Rightarrow$$

$\downarrow$

$2f(\theta).t$, onde

f é a densidade de $N(0,1)$

$$P(\sqrt{n}(Md(X) - \theta) \leq t) \xrightarrow[n \to \infty]{} P(Z \leq t.2f(\theta))$$

ou seja,

$$\sqrt{n}(Md(X) - \theta) \xrightarrow{d} N\left(0, \left(\frac{1}{2f(\theta)}\right)^2\right) \stackrel{d}{=} N\left(0, \left(\frac{1}{2.\frac{1}{\sqrt{2\pi}}}\right)^2\right) \stackrel{d}{=} N\left(0, \frac{\pi}{2}\right)$$

$$f(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta-\theta)^2}{2}} = \frac{1}{\sqrt{2\pi}}$$ Logo, $(Md(X))_{n \geq 1}$ é ass. Normal $(0, \pi/2)$

Resultado: $(\delta_n)_{n\geq 1}$ é assintoticamente Normal $(0, \sigma^2(\theta)) \Rightarrow (\delta_n)_{n \geq 1}$ é consistente para θ

Para $\varepsilon > 0$,

$$P_\theta(|\delta_n(X) - \theta| > \varepsilon) = P_\theta(|\sqrt{n}(\delta_n(X) - \theta)| > \varepsilon\sqrt{n})$$

Para todo $M \in \mathbb{N}^*$, $\exists\, n_0 = n_0(M)$ tal que

$n \geq n_0(M) \Rightarrow \varepsilon\sqrt{n} > M$. Para $n \geq n_0$,

$$P_\theta(|\sqrt{n}(\delta_n(X)-\theta)|>\varepsilon\sqrt{n}) \leq P_\theta(|\sqrt{n}(\delta_n(X)-\theta)|>M)$$

Então,

$$\limsup_{n\to\infty} P_\theta(|\sqrt{n}(\delta_n(X)-\theta)|>\varepsilon\sqrt{n}) \leq \underbrace{\lim_{n\to\infty} P_\theta(|\sqrt{n}(\delta_n(X)-\theta)|>M)}_{P_\theta(|N(0,\sigma^2(\theta))|>M)}, \forall M\in\mathbb{N}^*$$

Então,

$$\limsup_{n\to\infty} P_\theta(|\sqrt{n}(\delta_n(X)-\theta)|>\varepsilon\sqrt{n}) \leq \underbrace{\lim_{M\to\infty} P_\theta(|N(0,\sigma^2(\theta))|>M)}_{\substack{\shortparallel\\ 0}}$$

Assim,

$$\lim_{n\to\infty} P(|\delta_n(X)-\theta|>\varepsilon)=0, \forall \varepsilon>0, \forall \theta\in\Theta$$

Logo, $\delta_n \xrightarrow{P_\theta} \theta, \forall\theta\in\Theta$ e

$(\delta_n)_{n\geq1}$ é consistente para θ.

Algumas outras definições:

Definição: $(\delta_n^{(i)})_{n\geq1}$ seq. assintoticamente normal $(0,\sigma_i^2(\theta))$, $i=1,2$.

Dizemos que $(\delta_n^{(1)})_{n\geq1}$ é assintoticamente melhor que $(\delta_n^{(2)})_{n\geq1}$, se $\sigma_1^2(\theta)\leq\sigma_2^2(\theta)$ $\forall\theta\in\Theta$.

No exemplo, $X_1,\dots,X_n$, dado θ, i.i.d. $N(\theta,1)$

$\sqrt{n}(\bar{X}_n-\theta)\xrightarrow{d}N(0,1)$ $(\sigma_1^2(\theta)=1,\forall\theta)$ Logo, $(\bar{X}_n)$ é assintoticamente

$\sqrt{n}(Md(X)-\theta)\xrightarrow{d}N(0,\pi/2)$ $(\sigma_2^2(\theta)=\frac{\pi}{2},\forall\theta)$ melhor que $(Md_n(X))_{n\geq1}$.

Definição: $(\delta_n^+)_{n\geq 1}$ é "best asymptotically normal" (BAN) se

(1) $\sqrt{n}\,(\delta_n^+ - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$

(2) $(\delta_n^+)_{n\geq 1}$ é melhor assintoticamente que qualquer outra $(\delta_n)_{n\geq 1}$ assintoticamente normal.

"como se fosse o ENVVUM para os ass. normais"

Definição: $(\delta_n)_{n\geq 1}$ e $(\delta'_n)_{n\geq 1}$ são assintoticamente equivalentes se

$$\sqrt{n}\,(\delta_n - \delta'_n) \xrightarrow{P_\theta} 0, \ \forall \theta \in \Theta$$

Ex. 1: $X_1, \dots, X_n$, dado θ, c.i.i.d. $\mathcal{U}(0,\theta)$

$$(X_{(n)})_{n\geq 1} \text{ e } \left(\frac{n+1}{n} X_{(n)}\right)_{n\geq 1} \text{ são equivalentes}$$

$\forall \varepsilon > 0$,

$$P_\theta\left(\left|\sqrt{n}\left(X_{(n)} - \frac{n+1}{n}X_{(n)}\right) - 0\right| > \varepsilon\right) = P_\theta\left(\left|\sqrt{n}\left(-\frac{1}{n}X_{(n)}\right) - 0\right| > \varepsilon\right) =$$

$$= P_\theta\left(\left|\frac{X_{(n)}}{\sqrt{n}}\right| > \varepsilon\right) = P_\theta\left(X_{(n)} > \varepsilon\sqrt{n}\right) = 0 \ , \ \forall n \geq n_0(\theta), \text{ onde}$$

$$n_0(\theta) = \min\{n \geq 1 : \varepsilon\sqrt{n} > \theta\}. \text{ Logo, } \sqrt{n}\left(X_{(n)} - \frac{n+1}{n}X_{(n)}\right) \xrightarrow{P} 0, \forall \theta \in \Theta$$

Assim, $(X_{(n)})_{n\geq 1}$ e $\left(\frac{n+1}{n}X_{(n)}\right)_{n\geq 1}$ são assintoticamente equivalentes.

Exercício. $X_1, \ldots, X_n$ dado θ, c.i.i.d. $Ber(\theta)$

$\delta_n(X) = \bar{X}_n$

$\delta'_n(X) = \dfrac{a + \sum_{i=1}^{n} X_i}{a+b+n}$ Verificar que $(\delta_n)_{n\geq 1}$ e $(\delta'_n)_{n\geq 1}$ são assintoticamente equivalentes.

Resultado:

$(\delta_{mv}^{(n)})_{n\geq 1}$ sequência de EMV para θ.

sob condições de regularidade,

$\sqrt{n}\left(\delta_{mv}^{(n)} - \theta\right) \xrightarrow{d} N\left(0, \dfrac{1}{I_{X_1}(\theta)}\right)$, no caso $(X_n)_{n\geq 1}$, dado θ, c.i.i.d.

$\left(\delta_{mv}^{(n)}(X) \overset{a}{\approx} N\left(\theta, \dfrac{1}{n I_{X_1}(\theta)}\right)\right)$

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

$\lambda(x/\theta) = \log f(x|\theta) = \log \prod_{i=1}^{n} f(x_i|\theta) = \sum_{i=1}^{n} \log(f(x_i|\theta))$

$\lambda'(x|\theta) = \sum_{i=1}^{n} \dfrac{d \log f(x_i|\theta)}{d\theta}$ — $\lambda'_i(x_i|\theta)$

$l_x(\theta)$

No ponto $\theta_0 \in \Theta$

$l_x(\theta) = l_x(\theta_0) + l'_x(\theta_0)(\theta - \theta_0) + \ldots \approx l_x(\theta_0) + l'_x(\theta_0)(\theta - \theta_0)$ (I)

Aplicando (I) no ponto $\theta = \delta_{mv}(x)$, temos

$\underbrace{l_x(\delta_{mv}(x))}_{= 0} = l_x(\theta_0) + l'_x(\theta_0)(\delta_{mv}(x) - \theta_0) \Rightarrow$

$$\delta_{mv}(x) - \theta_0 = \frac{-\lambda_x(\theta_0)}{\lambda'_x(\theta_0)}$$

Assim,

$$\delta_{mv}(X) - \theta_0 = \frac{-\sum_{i=1}^{n} \lambda_i'(X_i/\theta_0)}{\sum_{i=1}^{n} \lambda_i''(X_i/\theta)} \Rightarrow \sqrt{n}\,(\delta_{mv}(X) - \theta_0) =$$

$$= \frac{-\sqrt{n}\sum_{i=1}^{n} \lambda_i'(X_i/\theta)}{\sum_{i=1}^{n} \lambda_i''(X_i/\theta)} = \frac{\underbrace{\sqrt{I_{X_1}(\theta)}\,\frac{\sum \lambda_i'(X_i/\theta_0) - n\,E[\lambda'(X_1/\theta)]}{\sqrt{n}\,\sqrt{I_{X_1}(\theta)}^{*}}}_{\xrightarrow{d} N(0,\, I_{X_1}(\theta_0))}}{\sum_{i=1}^{n} \frac{-\lambda_i''(X_i/\theta_0)}{n} \xrightarrow{q.c.} E_\theta(-\lambda''(X_1/\theta)) = I_{X_1}(\theta_0)}$$

$*$ $Var(\lambda'(X_1/\theta)) = \sqrt{I_{X_1}(\theta)}$

$$\xrightarrow{d} N\left(0,\, I_{X_1}(\theta_0)\right)\frac{1}{I_{X_1}(\theta_0)} \overset{d}{=} N\left(0,\, \frac{1}{I_{X_1}(\theta_0)}\right).$$

Exemplo: $\mathfrak{X} = \mathbb{R}_+^n$, $\Theta = \mathbb{R}_+$

$X_1, \ldots, X_n$, dado θ, c.i.i.d. Gama$(\theta, 1)$

$$f(x_1/\theta) = \frac{1^{\theta}}{\Gamma(\theta)}\, x_1^{\theta-1}\, e^{-x_1}\, \mathbb{I}_{\mathbb{R}_+}(x_1)$$

$x = (x_1, \ldots, x_n) \in \mathbb{R}_+^n$

$$f(x/\theta) = \prod_{i=1}^{n} f(x_i/\theta) = \prod_{i=1}^{n} \frac{x_i^{\theta-1}\, e^{-x_i}}{\Gamma(\theta)} \Rightarrow f(x/\theta) = \frac{\left(\prod_{i=1}^{n} x_i\right)^{\theta-1} e^{-\sum_{i=1}^{n} x_i}}{(\Gamma(\theta))^n}$$

$$\lambda(x/\theta) = \log f(x/\theta) = (\theta-1)\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n} x_i - n\log\Gamma(\theta) \Rightarrow$$

$$I_X(\theta) = n\,\frac{\Gamma''(\theta)\,\Gamma(\theta) - (\Gamma'(\theta))^2}{(\Gamma(\theta))^2}$$

$$\lambda'(x/\theta) = \sum_{i=1}^{n} \log x_i - n\frac{\Gamma'(\theta)}{\Gamma(\theta)} \overset{?}{=} 0$$

$$\lambda''(x/\theta) = -\left\{ n \frac{\Gamma''(\theta)\Gamma(\theta) - \Gamma'(\theta)\Gamma'(\theta)}{(\Gamma(\theta))^2} \right\}$$

$$E_\theta(-\lambda''(X/\theta)) = -E_\theta(\ \ )$$

$\Rightarrow$

$$I_X(\theta) = n\frac{\Gamma''(\theta)\Gamma(\theta) - (\Gamma'(\theta))^2}{(\Gamma(\theta))^2}$$

$$\sqrt{n}\,(\delta_n(X) - \theta) \xrightarrow{\mathcal{D}} N\left(0, \frac{1}{I_{X_1}(\theta)}\right) \overset{\mathcal{D}}{=} N\left(0, \frac{\Gamma(\theta)^2}{(\Gamma''(\theta)\Gamma(\theta)) - (\Gamma'(\theta))^2}\right)$$

$$\lambda''(x|\theta) = -\left[\frac{+n\,\Gamma''(\theta)\Gamma(\theta) - \Gamma'(\theta)\,\Gamma'(\theta)}{(\Gamma(\theta))^2}\right]$$

$$E_\theta\left(-\lambda''(x|\theta)\right) = -E_\theta\left(\quad\right)$$

$$\Rightarrow I_X(\theta) = n \cdot \frac{\Gamma''(\theta)\Gamma(\theta) - (\Gamma'(\theta))^2}{(\Gamma(\theta))^2}.$$

$$\sqrt{n}\left(\delta_n(X) - \theta\right) \xrightarrow{\mathcal{D}} N\left(0, \frac{1}{I_{X_1}(\theta)}\right) \overset{\mathcal{D}}{=} N\left(0, \frac{(\Gamma(\theta))^2}{(\Gamma''(\theta)\Gamma(\theta)) - (\Gamma'(\theta))^2}\right).$$

Na última aula:

Sob certas condições, $\sqrt{n}\,(\delta_{MV} - \theta) \xrightarrow{\mathcal{D}} N\left(0, \frac{1}{I_{X_1}(\theta)}\right)$

Aula 11 04/06

Resultado (Método Delta): $(\delta_n)_{n\geq 1}$ tal que $\sqrt{n}\,(\delta_n - \theta) \xrightarrow{\mathcal{D}} N(0, \sigma^2(\theta))$.

Seja $g: \mathbb{R} \to \mathbb{R}$ diferenciável. Então,

$$\sqrt{n}\cdot\left(g(\delta_n) - g(\theta)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2(\theta)\,(g'(\theta))^2\right).$$

[$(g(\delta_n))_{n\geq 1}$ é assintoticamente Normal]

$\sqrt{n}\cdot(\delta_n - \theta) \xrightarrow{\mathcal{D}} N(0, \sigma^2(\theta)) \Rightarrow (\delta_n)_{n\geq 1}$ é consistente, i.e., $\delta_n \xrightarrow{P_\theta} \theta$.

$$\frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} \xrightarrow{P_\theta} g'(\theta).\ (*)$$

/ /

Para todo $\varepsilon > 0$, $\exists \delta > 0$,

$$|\delta_n - \theta| < \delta \Rightarrow \left| \frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} - g'(\theta) \right| < \varepsilon$$

Assim,

$$\underbrace{\mathbb{P}(|\delta_n - \theta| < \delta)}_{1} \leq \mathbb{P}\left(\left| \frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} - g'(\theta) \right| < \varepsilon \right)$$

Como $\delta_n \xrightarrow{\mathbb{P}} \theta$, temos

$$\lim_{n \to \infty} \mathbb{P}\left(\left| \frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} - g'(\theta) \right| < \varepsilon \right) = 1.$$

Logo, vale $(*)$

Portanto,

$$\sqrt{n}\,(g(\delta_n) - g(\theta)) = \underbrace{\frac{g(\delta_n) - g(\theta)}{\delta_n - \theta}}_{\xrightarrow{\mathbb{P}} g'(\theta)} \cdot \underbrace{\sqrt{n}\,(\delta_n - \theta)}_{\xrightarrow{\mathcal{D}} N(0, \sigma^2(\theta))}$$

$$\Big\downarrow \mathcal{D}$$

$$N(0, \sigma^2(\theta) \cdot (g'(\theta))^2).$$

Aplicação: $\sqrt{n}\,(\delta_{MV}^{(n)} - \theta) \xrightarrow{\mathcal{D}} N\left(0, \frac{1}{I_{X_1}(\theta)}\right)$.

$(g(\delta_{MV}^{(n)}))_{n \geq 1}$. Do resultado,

$$\sqrt{n}\,(\underbrace{g(\delta_{MV}^{(n)})}_{\widehat{g(\theta)}} - g(\theta)) \xrightarrow{\mathcal{D}} N\left(0, \frac{(g'(\theta))^2}{I_{X_1}(\theta)}\right)$$

Exemplo: $X_1, \ldots, X_n$ dado θ, c.i.i.d. $Exp(\theta)$

$\delta_{MV}(X) = \frac{1}{\bar{X}_n}$. (Utilizando a parametrização $\theta e^{-\theta x}$)

Então

$$\sqrt{n}(\delta_{MV}(X) - \theta) \xrightarrow{\mathcal{D}} N\left(0, \frac{1}{I_1(\theta)}\right).$$

Calculando $I_1(\theta)$, temos

$$\lambda(x|\theta) = \log\theta - \theta x$$

$$\lambda'(x|\theta) = \frac{1}{\theta} - x$$

$$\lambda''(x|\theta) = \frac{-1}{\theta^2}$$

$$I_1(\theta) = -E_\theta[\lambda''(X|\theta)] = \frac{1}{\theta^2}$$

Logo, se $\sqrt{n}(\delta_{MV} - \theta) \xrightarrow{\mathcal{D}} N(0, \theta^2)$

Para $g(\theta) = 1/\theta$, utilizando o método Delta:

$$\sqrt{n}\left(\bar{X}_n - \frac{1}{\theta}\right) \xrightarrow{\mathcal{D}} N\left(0, \left(\frac{-1}{\theta^2}\right)^2 \theta^2\right) \stackrel{d}{=} N\left(0, \frac{1}{\theta^2}\right)$$

E, $h(\theta) = \frac{1}{\theta^2} = Var_\theta(X_1)$

$$\widehat{h(\theta)}_{MV} = h(\delta_{MV}) = \frac{1}{\left(\frac{1}{\bar{X}_n}\right)^2} = \bar{X}_n^2$$

/ /

Pelo método Delta,

$$\sqrt{n}\Big(\underbrace{\widehat{h(\theta)}_{MV}}_{\bar{X}_n^2} - \frac{1}{\theta^2}\Big) \xrightarrow{\mathcal{D}} N\left(0, \left(-\frac{2}{\theta^3}\right)^2 \cdot \theta^2\right)$$

$$\Vert\Vert$$

$$N\left(0, \frac{4}{\theta^4}\right).$$

INTERVALOS DE CONFIANÇA

$$\left(\begin{array}{l} \delta: \mathcal{X} \to \Theta \\ \mathcal{X} \to \{[a,b]; a \leq b\} \end{array}\right)$$

$$(\mathcal{X}, \mathcal{A}, \mathcal{P}) \qquad \mathcal{P} = \{f(\cdot|\theta) : \theta \in \Theta\}$$

$$g(\Theta) \subseteq \mathbb{R} \ (\mathbb{R}^K).$$

Sejam $A, B: \mathcal{X} \to \mathbb{R}$ tais que $A(x) \leq B(x), \forall x \in \mathcal{X}$.

$$\mathcal{J} = \{[a,b] . a \leq b\}$$

Def.: A transformação $T: \mathcal{X} \to \mathcal{J}$ que a cada $x \in \mathcal{X}$ associa um intervalo em $\mathcal{J}$ (isto é, $T(x) = [a(x), b(x)]$) damos o nome de INTERVALO ALEATÓRIO.

Exemplo: $X_1, \dots, X_n$ dado θ, c.i.i.d $N(\theta, 1)$ $X = (X_1, \dots, X_n)$

Seja $A(X) = \bar{X} - 1$

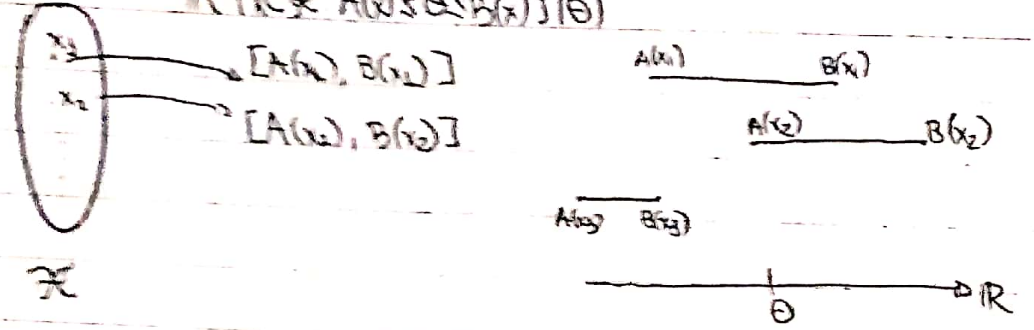
$B(X) = \bar{X} + 2$

$[A(X), B(X)]$ é um intervalo aleatório.

DEF: O intervalo aleatório $[A(X), B(X)]$ é um intervalo de confiança (I.C.) para θ $(g(\theta))$ com coeficiente de confiança $1-\alpha$, $0<\alpha<1$, se

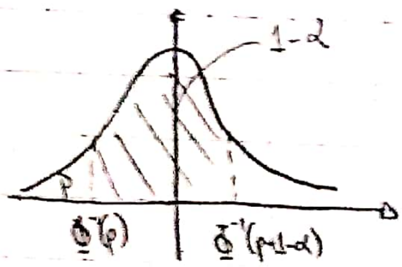
$$\mathbb{P}\left(A(X) \le \overset{(g(\theta))}{\theta} \le B(X) \,\middle|\, \theta\right) \ge 1-\alpha, \ \forall \theta \in \Theta.$$

$\hookrightarrow \mathbb{P}(\{x \in \mathfrak{X} : A(x) \le \theta \le B(x)\} | \theta)$



Ex 1. $X_1, \dots, X_n$, dado θ, c.i.i.d $N(\theta, \sigma_0^2)$ ← conhecido

$$\frac{\bar{X}_n - \theta}{\sigma_0/\sqrt{n}} \,\Big|\, \theta \sim N(0,1).$$



Para $\alpha \in (0,1)$ e $p \in (0,\alpha)$,

$$\mathbb{P}\left(\Phi^{-1}(p) \le \frac{\bar{X}_n - \theta}{\sigma_0/\sqrt{n}} \le \Phi^{-1}(p+1-\alpha) \,\middle|\, \theta\right) = 1-\alpha, \ \forall \theta \in \Theta, \text{ onde}$$

$\Phi^{-1}(u) \in \mathbb{R}$ é tq $\mathbb{P}(Z \le \Phi^{-1}(u)) = u$, $u \in (0,1)$.

$$\Rightarrow \mathbb{P}\left(-\Phi^{-1}(p+1-\alpha) \le \frac{\theta - \bar{X}_n}{\sigma_0/\sqrt{n}} \le -\Phi^{-1}(p) \,\middle|\, \theta\right) \ge 1-\alpha, \ \forall \theta \in \mathbb{R}.$$

$$\Rightarrow \mathbb{P}\left(\underbrace{\bar{X}_n - \Phi^{-1}(p+1-\alpha)\frac{\sigma_0}{\sqrt{n}}}_{A(X)} \le \theta \le \underbrace{\bar{X}_n - \Phi^{-1}(p)\cdot\frac{\sigma_0}{\sqrt{n}}}_{B(X)} \,\middle|\, \theta\right) \ge 1-\alpha, \ \forall \theta \in \Theta \text{ "}\mathbb{R}\text{"}$$

/ /

Logo, $\left[\bar{X}_n - \Phi^{-1}(p+1-\alpha)\cdot\frac{\sigma_0}{\sqrt{n}},\ \bar{X}_n - \Phi^{-1}(p)\frac{\sigma_0}{\sqrt{n}}\right]$ é um I.C. para θ com coeficiente de confiança $1-\alpha$.

"Em geral, procura-se $T(X,\theta)|\theta$ cuja distribuição não depende de θ!"

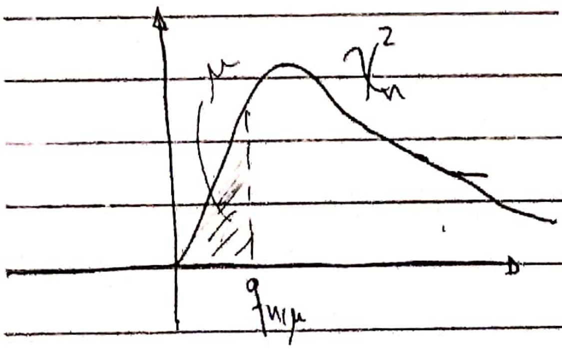
Ex. 2: $X_1, \dots, X_n$, dado θ, c.i.i.d. $N(\overset{\text{CONHECIDO!}}{\mu_0}, \theta)$ "$T(X) = \sum_{i=1}^{n}(X_i-\mu_0)^2$"

$$\frac{X_i-\mu_0}{\sqrt{\theta}}\Big|\theta \sim N(0,1)$$

$$\frac{(X_i-\mu_0)^2}{\theta}\Big|\theta \sim \chi^2_1 \implies \sum_{i=1}^{n}\frac{(X_i-\mu_0)^2}{\theta}\Big|\theta \sim \chi^2_n$$

$$\mathbb{P}\left(q_{n,p} \leq \sum_{i=1}^{n}\frac{(X_i-\mu_0)^2}{\theta} \leq q_{n,p+1-\alpha}\right) = 1-\alpha,\ \forall\theta\in\mathbb{R}_+, \text{ onde:}$$

$q_{n,\mu}$ é o percentil de ordem μ, $\mu\in(0,1)$, da dist. χ^2_n.



$$\Rightarrow \mathbb{P}\left(\frac{1}{q_{n,p+1-\alpha}} \leq \frac{\theta}{\sum_{i=1}^{n}(X_i-\mu_0)^2} \leq \frac{1}{q_{n,p}}\Big|\theta\right) = 1-\alpha,\ \forall\theta\in\mathbb{R}_+$$

$$\Longrightarrow \mathbb{P}\left(\underbrace{\frac{\sum_{i=1}^{n}(X_i-\mu_0)^2}{q_{n,p+1-\alpha}}}_{A(X)} \leq \Theta \leq \underbrace{\frac{\sum_{i=1}^{n}(X_i-\mu_0)^2}{q_{n,p}}}_{B(X)} \,\middle|\, \Theta\right) = 1-\alpha,\ \forall \Theta \in \mathbb{R}_+ .$$

Logo, $[A(X), B(X)]$ é um I.C para Θ com coeficiente de confiança $1-\alpha$.

Obs. Para encontrar p que minimize o comp do IC, devemos minimizar $E\left\{\sum_{i=1}^{n}(X_i-\mu_0)^2\left[\frac{1}{q_{n,p}} - \frac{1}{q_{n,p+1-\alpha}}\right]\right\}$

Ex 3: $X_1, \dots, X_n$, dado $\Theta = (\Theta_1, \Theta_2)$, c.i.i.d $N(\Theta_1, \Theta_2)$

$g(\Theta) = \Theta_1$.

$$\frac{\bar{X}_n - \Theta_1}{\sqrt{\Theta_2/n}} \mid \Theta \sim N(0,1).$$

$$\sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{\Theta_2} \mid \Theta \sim \chi^2_{n-1}.$$

$$\frac{\dfrac{\bar{X}_n - \Theta_1}{\sqrt{\Theta_2/n}}}{\sqrt{\dfrac{\sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{\Theta_2}}{(n-1)}}} \mid \Theta \sim t_{n-1}.$$

$$\overset{\text{lld}}{\frac{\sqrt{n}\,(\bar{X}_n - \Theta_1)}{\sqrt{S^2_x}}} \mid \Theta \sim t_{n-1}.$$

/ /

$$\mathbb{P}\left(-t_{n-1,\alpha/2} \leq \frac{\sqrt{n}(\bar{X}-\theta_1)}{S_X} \leq t_{n-1,\alpha/2} \,\middle|\, \theta\right) = 1-\alpha, \forall \theta \in \Theta,$$

onde $t_{n-1,\alpha/2}$ é o percentil de ordem $1-\alpha/2$ da distribuição t-Student de parâmetro $n-1$.

$$\Rightarrow \mathbb{P}\left(-t_{n-1,\alpha/2} \leq \frac{\theta_1 - \bar{X}_n}{S_X/\sqrt{n}} \leq t_{n-1,\alpha/2} \,\middle|\, \theta\right) = 1-\alpha, \forall \theta \in \Theta$$

$$\Rightarrow \mathbb{P}\left(\underbrace{\bar{X}_n - t_{n-1,\alpha/2} \cdot \frac{S_X}{\sqrt{n}}}_{A(X)} \leq \theta \leq \underbrace{\bar{X}_n + t_{n-1,\alpha/2} \cdot \frac{S_X}{\sqrt{n}}}_{B(X)} \,\middle|\, \theta\right) = 1-\alpha, \forall \theta \in \Theta.$$

Logo, $[A(X), B(X)]$ é um I.C. para $\theta_1 = g(\theta)$ com coef. de confiança $1-\alpha$.

Ex. 4: $X_1, \ldots, X_n, Y_1, \ldots, Y_m$, dado $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$, condicionalmente independentes tais que:

$$X_i | \theta \sim N(\mu_1, \sigma_1^2), \quad i=1,\ldots,n \quad e$$

$$Y_j | \theta \sim N(\mu_2, \sigma_2^2), \quad j=1,\ldots,m.$$

$$g(\theta) = \frac{\sigma_2^2}{\sigma_1^2}.$$

$$\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma_1^2} \Big| \theta \sim \chi^2_{n-1}.$$

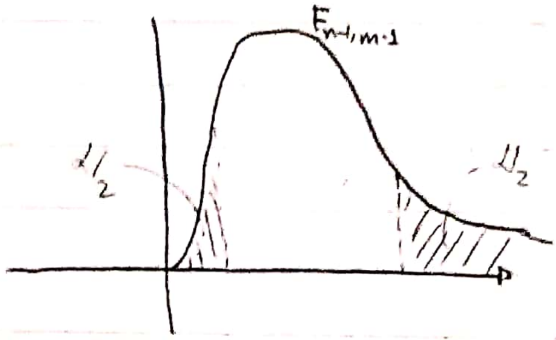
$$\sum_{j=1}^{m} \frac{(Y_j - \bar{Y})}{\sigma_2^2} \Big| \theta \sim \chi^2_{m-1}.$$

$$\frac{\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2 \big/ \sigma_1^2}{n-1}}{\dfrac{\sum_{j=1}^{m}(Y_j-\bar{Y})^2 \big/ \sigma_2^2}{m-1}} \Bigg| \theta \sim F_{n-1,m-1}$$

S_x^2 $\qquad$ S_y^2

isto é,

$$\frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{S_x^2}{S_y^2} \Bigg| \theta \sim F_{n-1,m-1}.$$



$$\mathbb{P}\left(F_{n-1,m-1,\alpha/2} \leq \frac{\sigma_2^2}{\sigma_1^2}\,\frac{S_x^2}{S_y^2} \leq F_{n-1,m-1,1-\alpha/2} \Bigg| \theta\right) = 1-\alpha, \ \forall \theta \in \Theta.$$

$$\Longrightarrow \mathbb{P}\left(\underbrace{F_{n-1,m-1,\alpha/2} \cdot \frac{S_y^2}{S_x^2}}_{A(X,Y)} \leq \frac{\sigma_2^2}{\sigma_1^2} \leq \underbrace{F_{n-1,m-1,1-\alpha/2}\,\frac{S_y^2}{S_x^2}}_{B(X,Y)} \Bigg| \theta\right) = 1-\alpha, \ \forall \theta \in \Theta.$$

Logo, $[A(X,Y), B(X,Y)]$ é um I.C. para $\frac{\sigma_2^2}{\sigma_1^2}$ c/ coeficiente de confiança $1-\alpha$.

Ex. 5: $X_1, \ldots, X_n$, dado θ, c.i.i.d. $\mathcal{U}(0,\theta)$.

Vimos que $f_{X_{(n)}}(t|\theta) = \frac{n t^{n-1}}{\theta^n} \cdot \mathbb{I}_{(0,\theta)}(t)$

$(F_{X_{(n)}}(t) = ?)$

$$\frac{X_{(n)}}{\theta} \Big| \theta \sim Beta(n,1)$$

$$F_{\frac{X_{(n)}}{\theta}|\theta}(t) = t^n, \quad t \in (0,1)$$

$$t^n = \alpha/2 \Rightarrow t = (\alpha/2)^{1/n}$$

$$t^n = 1 - \frac{\alpha}{2} \Rightarrow t = \left(1 - \frac{\alpha}{2}\right)^{1/n}$$

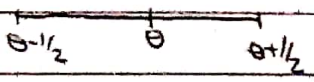
$$\mathbb{P}\left(\left(\frac{\alpha}{2}\right)^{1/n} \leq \frac{X_{(n)}}{\theta} \leq \left(1 - \frac{\alpha}{2}\right)^{1/n} \middle| \theta\right) = 1 - \alpha, \ \forall \theta > 0$$

$$\Rightarrow \mathbb{P}\left(\frac{1}{(1-\alpha/2)^{1/n}} \leq \frac{\theta}{X_{(n)}} \leq \frac{1}{(\alpha/2)^{1/n}} \middle| \theta\right) = 1 - \alpha, \ \forall \theta > 0.$$

$$\Rightarrow \mathbb{P}\left(\underbrace{\frac{X_{(n)}}{(1-\alpha/2)^{1/n}}}_{A(X)} \leq \theta \leq \underbrace{\frac{X_{(n)}}{(\alpha/2)^{1/n}}}_{B(X)} \middle| \theta\right) = 1 - \alpha, \ \forall \theta > 0.$$

Ex. 6: $X_1, \ldots, X_n$, dado θ, c.i.i.d. $\mathcal{U}\left(\theta - \frac{1}{2}, \theta + \frac{1}{2}\right)$. Seja o IC $=$ $[X_{(1)}, X_{(n)}]$. Encontre α.

$$\mathbb{P}(X_{(1)} \leq \theta \leq X_{(n)} | \theta)$$



$$= 1 - \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i > \theta\} \middle| \theta\right) - \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i < \theta\} \middle| \theta\right)$$

$$= 1 - \left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^n = 1 - \left(\frac{1}{2}\right)^{n-1}.$$

Logo, $[X_{(1)}, X_{(n)}]$ é um IC para θ com coeficiente de confiança $\gamma = 1 - \left(\frac{1}{2}\right)^{n-1}$.

Exercício: Verificar que $[X_{(2)}, X_{(n)}]$ é I.C. p/ $\gamma = ?$

n=3 $\quad X_1 = 5{,}3$
$\quad X_2 = 5{,}001$
$\quad X_3 = 5{,}999$

$[5{,}001;\ 5{,}999]$ é estimativa por intervalo para θ com coeficiente de confiança $1 - \left(\frac{1}{2}\right)^{3-1} = \frac{3}{4}$

$$\left.\begin{array}{l} \theta - 1/2 \leq 5{,}001 \\ 5{,}999 \leq \theta + 1/2 \end{array}\right\} \Rightarrow \theta \in [5{,}499;\ 5{,}501].$$

$[(-\infty, B(X))]$ [limite superior]

DEF.: $[A(X), \infty)$ é um limite inferior de confiança com coeficiente de confiança $1-\alpha$, $0<\alpha<1$, se

$$\mathbb{P}(A(X) \leq \theta < \infty \mid \theta) \geq 1-\alpha,\ \forall \theta \in \Theta.$$

$$(\mathbb{P}(-\infty < \theta \leq B(X) \mid \theta) \geq 1-\alpha,\ \forall \theta \in \Theta)$$

Ex 1: $X_1, \dots, X_n$, dado θ, c.i.i.d $N(\theta, \sigma_0^2)$ conhecido!

$$\left.\frac{\bar{X}-\theta}{\sigma_0/\sqrt{n}}\right| \theta \sim N(0,1)$$

$$\mathbb{P}\left(\frac{\bar{X}-\theta}{\sigma_0/\sqrt{n}} \geq \Phi^{-1}(\alpha) \,\middle|\, \theta\right) = 1-\alpha,\ \forall \theta \in \Theta$$

$$\Rightarrow \mathbb{P}\left(\frac{\theta-\bar{X}}{\sigma_0/\sqrt{n}} \leq -\Phi^{-1}(\alpha) \,\middle|\, \theta\right) = 1-\alpha,\ \forall \theta \in \Theta$$

$$\Rightarrow \mathbb{P}\left(\theta \leq \bar{X} - \Phi^{-1}(\alpha)\frac{\sigma_0}{\sqrt{n}} \,\middle|\, \theta\right) = 1-\alpha,\ \forall \theta \in \Theta.$$

Assim,

$\left(-\infty, \bar{X} - \Phi^{-1}(1-\alpha)\frac{\sigma_0}{\sqrt{n}}\right]$ é um limite superior de confiança p/ θ com confiança $1-\alpha$.

06/06 Aula 12

$\Theta \subseteq \mathbb{R}^k$

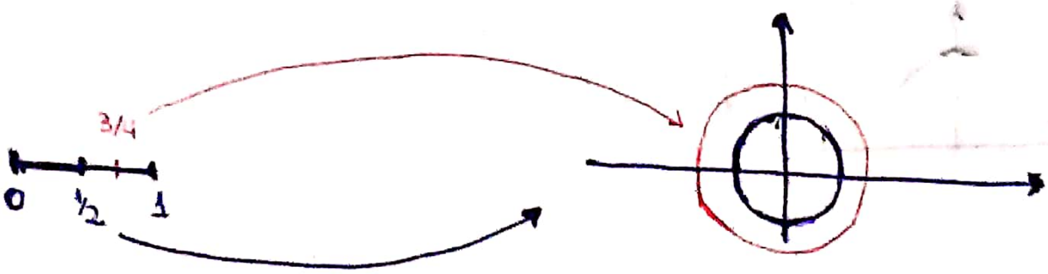
Definição:

Seja A um conjunto de subconjuntos de $\mathbb{R}^k$.

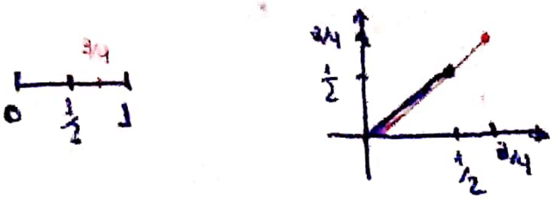
Definição: A transformação $R: \mathcal{X} \to A$ que associa a cada $x \in \mathcal{X}$ uma região (ou subconjunto) $R(x) \in A$, damos o nome de Região (conjunto) aleatório.

Exemplo: $X \sim U(0,1)$, $A = \mathcal{B}(\mathbb{R}^2)$

$$R(x) = \{(a,b) \in \mathbb{R}^2 : a^2+b^2 \leq x^2\}$$



$R_2(x) = \{\lambda(0,0) + (1-\lambda)(x,x) \quad \lambda \in [0,1]\}$



$g(\Theta) \subseteq \mathbb{R}^k$

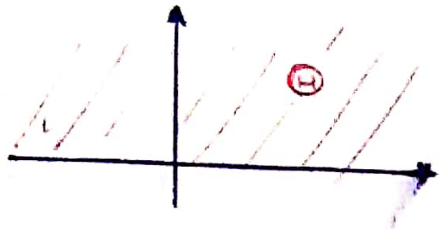
Definição: A região (conjunto) aleatória (Θ) R é uma região (conjunto) de confiança para $g(\theta)$ com coeficiente de confiança $1-\alpha$, $0<\alpha<1$; se

$$\mathbb{P}(\theta \in R(X) | \theta) \geq 1-\alpha, \ \forall \theta \in \Theta$$

Exemplo 1. $X_1, \ldots, X_n$ dado $\theta = (\theta_1, \theta_2)$, c.i.i.d. $N(\theta_1, \theta_2)$

$\theta = (\theta_1, \theta_2)$

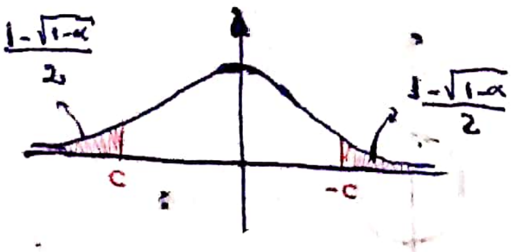
$\Theta = \mathbb{R} \times \mathbb{R}_+$



$$\left.\begin{array}{l} \dfrac{\bar{X} - \theta_1}{\sqrt{\theta_2/n}} \Big| \theta \sim N(0,1) \\[2ex] \displaystyle\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\theta_2} \Big| \theta \sim \chi^2_{n-1} \end{array}\right\} \text{independentes (dado } \theta)$$

Para $\alpha \in (0,1)$

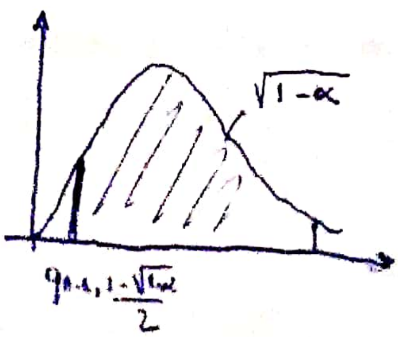
$$P\left(\Phi^{-1}\left(\frac{1-\sqrt{1-\alpha}}{2}\right) \leq \frac{\bar{X} - \theta_1}{\sqrt{\theta_2/n}} \leq -c \,\Big|\, \theta \right) = \sqrt{1-\alpha}$$



Além disso,

$$P\left(\underbrace{q_{n-1, \frac{1-\sqrt{1-\alpha}}{2}}}_{a} \leq \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{\theta_2} \leq \underbrace{q_{n-1, 1-\left(\frac{1-\sqrt{1-\alpha}}{2}\right)}}_{b} \,\Big|\, \theta \right) =$$

$= \sqrt{1-\alpha}$, onde $q_{n-1,p}$ é o percentil de ordem p da χ^2_{n-1}.



Assim,

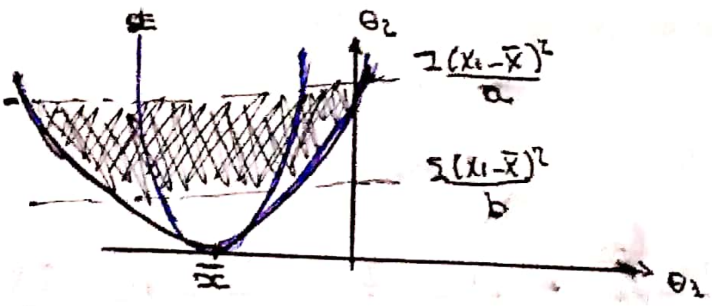
$$\mathbb{P}\left(c \leq \frac{\bar{X} - \theta_1}{\sqrt{\theta_2/n}} \leq -c ,\ a \leq \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\theta_2} \leq b \,\middle|\, \theta \right) =$$

$$\mathbb{P}\left(c \leq \frac{\bar{X} - \theta_1}{\sqrt{\theta_2/n}} \leq -c \,\middle|\, \theta \right) \mathbb{P}\left(a \leq \frac{\sum (X_i - \bar{X})^2}{\theta_2} \leq b \,\middle|\, \theta \right) =$$

$$= \sqrt{1-\alpha} \cdot \sqrt{1-\alpha} = 1-\alpha ,\ \forall \theta \in \Theta$$

$$\Rightarrow \mathbb{P}\left(|\bar{X} - \theta_1| \leq |c| \sqrt{\theta_2/n} ,\ \frac{\sum (X_i - \bar{X})^2}{b} \leq \theta_2 \leq \frac{\sum (X_i - \bar{X})^2}{a} \,\middle|\, \theta \right) = 1-\alpha, \forall \theta \in \Theta$$

$$= \mathbb{P}\left(\frac{\sum (X_i - \bar{X})^2}{b} \leq \theta_2 \leq \frac{\sum (X_i - \bar{X})^2}{a} ,\ \theta_2 \geq \frac{n (\theta_1 - \bar{x})^2}{c^2} \,\middle|\, \theta \right) = 1-\alpha ,\ \forall \theta \in \Theta$$



Intervalos Aproximados

Exemplo: $X_1, \dots, X_n$, dado θ, c.i.i.d. Ber(θ)

$$\sqrt{n} \, (\bar{X} - \theta) \xrightarrow{d} N(0, \theta(1-\theta))$$

$$\sqrt{\bar{X}(1-\bar{X})} \xrightarrow{\mathbb{P}_\theta} \sqrt{\theta(1-\theta)}$$

$$\frac{\sqrt{n}(\bar{X}-\theta)}{\sqrt{\bar{X}(1-\bar{X})}} \xrightarrow{d} \frac{1}{\sqrt{\theta(1-\theta)}} N(0,\theta(1-\theta)) \stackrel{d}{=} N(0,1)$$

Para n grande

$$\mathbb{P}\left(\Phi^{-1}\left(\frac{\alpha}{2}\right) \leq \frac{\sqrt{n}(\bar{X}-\theta)}{\sqrt{\bar{X}(1-\bar{X})}} \leq \Phi^{-1}(1-\alpha/2) \,\middle|\, \theta\right) \approx 1-\alpha, \forall \theta \in \Theta.$$

$$\mathbb{P}\left(\underbrace{\bar{X} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}}_{A(X)} \leq \theta \leq \underbrace{\bar{X} - \Phi^{-1}(\alpha/2)\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}}_{B(X)} \,\middle|\, \theta\right) \approx 1-\alpha, \forall \theta.$$

Logo, $[A(X), B(X)]$ é um IC aproximado para θ de confiança $1-\alpha$.

Exemplo 2: $X_1, \dots, X_n$, dado $\theta = (\theta_1, \theta_2)$, c.i.i.d. $N(\theta_1, \theta_2)$,

$$\sqrt{n}(\bar{X}-\theta_1)|\theta \sim N(0,\theta_2)$$

$$\sum_{i=1}^{n} \frac{(X_i-\bar{X})^2}{n-1} = \frac{n}{n-1}\sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{n} = \frac{n}{n-1}\,\frac{\sum X_i^2 - n\bar{X}^2}{n} = \underbrace{\frac{n}{n-1}}_{\downarrow\, 1}\left\{\underbrace{\frac{\sum X_i^2}{n}}_{\downarrow P_\theta\; E(X_1^2) = \theta_1^2+\theta_2} - \underbrace{\bar{X}^2}_{\downarrow P_\theta\; \theta_1^2}\right\} \xrightarrow{P_\theta} \theta_2$$

$$S_n^2 \xrightarrow{P_\theta} \theta_2$$

$$\sqrt{n}(\bar{X}_n-\theta_1) \xrightarrow{d} N(0,\theta_2)$$

$$\frac{\sqrt{n}(\bar{X}_n-\theta_1)}{\sqrt{S_n^2}} \xrightarrow{d} \frac{1}{\sqrt{\theta_2}} N(0,\theta_2) \stackrel{d}{=} N(0,1)$$

$$\mathbb{P}\left(\Phi^{-1}(\alpha/2) \leq \frac{\sqrt{n}(\bar{X}-\theta)}{S_n} \leq \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \Big| \theta\right) \approx 1-\alpha \ \forall \theta \Rightarrow$$

$$\Rightarrow \mathbb{P}\Big(\underbrace{\bar{X}_n - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\frac{S_n}{\sqrt{n}}}_{A(x)} \leq \theta_1 \leq \underbrace{\bar{X}_n - \Phi^{-1}(\alpha/2)\, S_n/\sqrt{n}}_{B(x)} \Big| \theta\Big) \approx 1-\alpha$$

Exemplo 3. $X_1, \ldots, X_n$, dado θ, c.i.i.d. Poisson(θ)

* IC(θ) → Chang Yu Fang
Diss Mestrado (2008)
Dr. Carlos Pereira

$g(\theta) = e^{-\theta}$

$\bar{X}_n$ é o EMV para θ. Sob certas condições,

$$\sqrt{n}(\bar{X}_n - \theta) \xrightarrow{d} N(0, \theta)$$

Pelo Método Delta,

$$\sqrt{n}\left(e^{-\bar{X}_n} - e^{-\theta}\right) \xrightarrow{d} N\left(0, (-e^{-\theta})^2 \cdot \theta\right) \stackrel{d}{=} N(0, \theta e^{-2\theta})$$

$$\sqrt{n}\left(e^{-\bar{X}_n} - e^{-\theta}\right) \xrightarrow{d} N(0, \theta e^{-2\theta})$$

$$\bar{X}_n \xrightarrow{P_\theta} \theta$$

$$e^{-2\bar{X}_n} \xrightarrow{P_\theta} e^{-2\theta}$$

$$\bar{X}_n\, e^{-2\bar{X}_n} \xrightarrow{P_\theta} \theta e^{-2\theta}$$

$$\sqrt{\bar{X}_n e^{-2\bar{X}_n}} \xrightarrow{P_\theta} \sqrt{\theta e^{-2\theta}}$$

Assim,

$$\frac{\sqrt{n}\left(e^{-\bar{X}_n} - e^{-\theta}\right)}{\sqrt{\bar{X}_n \cdot e^{-2\bar{X}_n}}} \xrightarrow{d} \frac{1}{\sqrt{\theta e^{-2\theta}}} N(0, \theta e^{-2\theta}) \stackrel{d}{=} N(0,1)$$

Para n grande

$$\mathbb{P}\left(\Phi^{-1}(\alpha/2) \leq \frac{\sqrt{n}\,(e^{-\bar{X}_n} - e^{-\theta})}{\sqrt{\bar{X}_n e^{-2\bar{X}_n}}} \leq \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \Big| \theta \right) \approx 1-\alpha, \ \forall \theta \in \Theta.$$

$$\Rightarrow \mathbb{P}\left(\underbrace{e^{-\bar{X}_n} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{X}_n e^{-2\bar{X}_n}}{n}}}_{A(X)} \leq e^{-\theta} \leq \underbrace{e^{-\bar{X}_n} - \Phi^{-1}(\alpha/2)\sqrt{\frac{\bar{X}_n e^{-2\bar{X}_n}}{n}}}_{B(X)} \Big| \theta \right) \approx$$

$\approx 1-\alpha, \ \forall \theta > 0.$

$X_1, \dots, X_n$, dado θ c.i.i.d.

No caso geral (sob condições de regularidade)

$$\sqrt{n}\,(\delta_{MV} - \theta) \xrightarrow{d} N\left(0, \frac{1}{I_{X_1}(\theta)}\right)$$

Se g é diferenciável:

$$\sqrt{n}\,(g(\delta_{MV}) - g(\theta)) \xrightarrow{d} N\left(0, (g'(\theta))^2 \cdot 1/I_{X_1}(\theta)\right), \text{ pelo método delta}$$

$\delta_{MV} \xrightarrow{P_\theta} \theta$

Se g' é contínua, então,

$$g'(\delta_{MV}) \xrightarrow{P_\theta} g'(\theta) \Rightarrow (g'(\delta_{MV}))^2 \xrightarrow{P_\theta} (g'(\theta))^2.$$

$\delta_{MV} \xrightarrow{P_\theta} \theta$

Se $I_{X_1}(\cdot)$ é contínua,

$$I_{X_1}(\delta_{MV}) \xrightarrow{P_\theta} I_{X_1}(\theta)$$

Logo

$$\frac{(g'(\delta_{MV}))^2}{I_{X_1}(\delta_{MV})} \xrightarrow{P_\theta} \frac{(g'(\theta))^2}{I_{X_1}(\theta)}$$

Assim,

$$\frac{\sqrt{n}\,(g(\delta_{mv}) - g(\theta))}{\sqrt{\frac{(g'(\delta_{mv}))^2}{I_{X_1}(\delta_{mv})}}} \xrightarrow{d} \frac{1}{\sqrt{\frac{(g'(\theta))^2}{I_{X_1}(\theta)}}} N\left(0, \frac{(g'(\theta))^2}{I_{X_1}(\theta)}\right) \stackrel{d}{=} N(0,1)$$

Para "n grande"

$$\mathbb{P}\left(\Phi^{-1}\left(\frac{\alpha}{2}\right) \leq \frac{\sqrt{n}\,(g(\delta_{mv}) - g(\theta))}{\sqrt{\frac{(g'(\delta_{mv}))^2}{I_{X_1}(\delta_{mv})}}} \leq \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \,\middle|\, \theta\right) \approx 1-\alpha, \forall \theta \in \Theta$$

Intervalo de probabilidade (ou de credibilidade)

Definição: $[A(X), B(X)]$ é um intervalo de probabilidade (ou de credibilidade) para θ ($g(\theta)$) com probabilidade $1-\alpha$, $0<\alpha<1$, se

$$\mathbb{P}(A(X) \leq \theta \leq B(X) \,/\, X = x) \geq 1-\alpha, \forall x \in \mathcal{X}$$

contrário!

Exemplo 1: $X_1, \dots, X_n$, dado θ, c.i.i.d. $N(\theta, \sigma_0^2)$

Vimos que

$\sigma'_n = \sigma^2_n(\sigma_0^2, h)$

$$\theta / X_1 = x_1, \dots, X_n = x_n \sim N(\underbrace{\alpha_n \bar{x}_n + (1-\alpha)a}_{a_x}; \sigma^2_n)$$

$$\frac{\theta - a_x}{\sqrt{\sigma^2_n}} \Big| X = x \sim N(0,1)$$

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

$$\mathbb{P}\left(\Phi^{-1}(\alpha/2) \leq \frac{\theta - a x}{\sqrt{\sigma_n^2}} \leq \Phi^{-1}(1-\alpha/2) \,\Big|\, X = x\right) = 1 - \alpha \Rightarrow$$

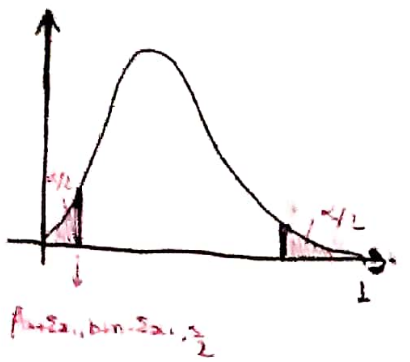
$\Rightarrow$

$$\mathbb{P}\left(ax + \Phi^{-1}(\alpha/2)\sqrt{\sigma_n^2} \leq \theta \leq ax + \Phi^{-1}(1-\alpha/2)\sqrt{\sigma_n^2} \,\Big|\, X = x\right) = 1-\alpha.$$

Exemplo: $X_1, \ldots, X_n$, dado θ, c.i.i.d. $Ber(\theta)$

$\theta \sim Beta(a,b)$

$\theta / X_1 = x_1, \ldots, X_n = x_n \sim Beta\left(a + \sum_{i=1}^n x_i,\ b+n - \sum_{i=1}^n x_i\right)$



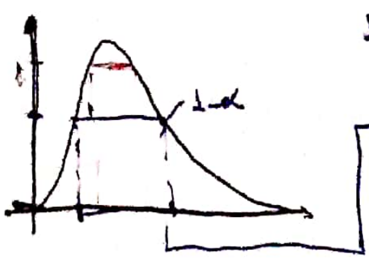
$$\mathbb{P}\left(\beta_{a+\Sigma x_i,\, b+n-\Sigma x_i,\, \frac{\alpha}{2}} \leq \theta \leq \beta_{a+\Sigma x_i,\, b+n-\Sigma x_i,\, 1-\frac{\alpha}{2}} \,\Big|\, X = x\right) = 1-\alpha$$

onde $\beta_{a,b,p}$ é o percentil de ordem p da $Beta(a,b)$

$$\Big[\underbrace{\beta_{a+\sum_{i=1}^n x_i,\ b+n-\sum_{i=1}^n x_i,\ \alpha/2}}_{A(X)},\ \underbrace{\beta_{a+\sum_{i=1}^n x_i,\ b+n-\sum_{i=1}^n x_i,\ 1-\alpha/2}}_{B(X)}\Big]$$ é um intervalo de credibi_

lidade de $1-\alpha$.

Ideia: Elevador



$1-\alpha = 95\%$

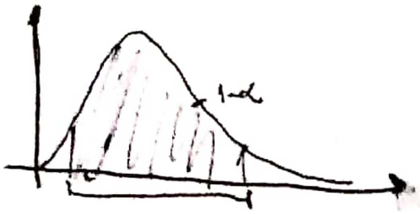
$\rightarrow$ Highest Posterior Density Interval

Intervalo HPD.

Exemplo 3. $X_1, \ldots, X_n$ dado θ c.i.i.d. Poisson(θ)

$\theta \sim$ Gama(a, b)

$\theta / X_1 = x_1, \ldots, X_n = x_n \sim$ Gama$(a + \sum_{i=1}^{n} x_i, b + n)$

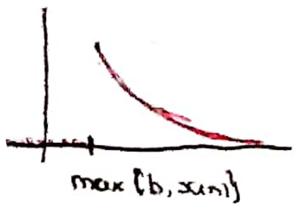


Exemplo 4. (Exercício).

$X_1, \ldots, X_n$ dado θ, c.i.i.d. $U(0, \theta)$

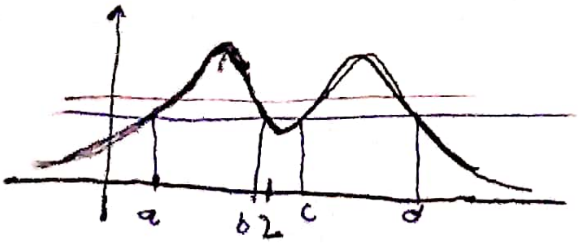
$\theta \sim$ Pareto(a, b)

$\theta / X_1 = x_1, \ldots, X_n = x_n \sim$ Pareto$(a + n, \max\{b, x_{(n)}\})$.



$\theta \sim \frac{1}{2} N(0, 1) + \frac{1}{2} N(4, 1)$

$$f(\theta) = \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} + \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta-4)^2}{2}}$$



$(a, b) \cup (c, d)$

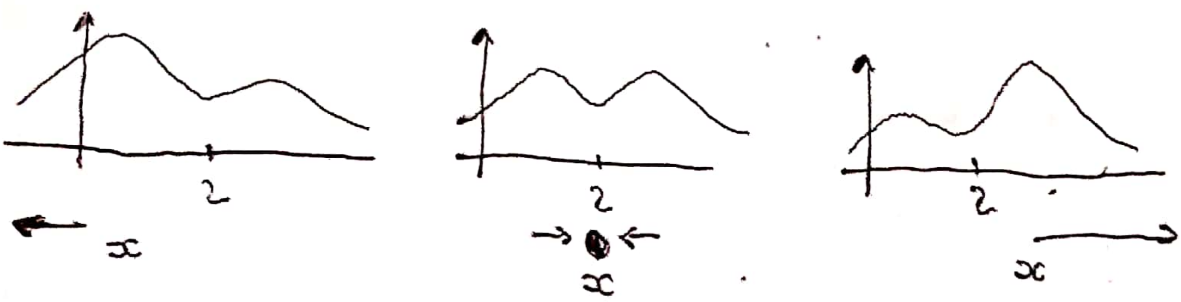
$X/\Theta \sim N(\theta, 1)$

$f(\theta/x) \propto f(x|\theta)\, f(\theta)$

$$\propto \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} \left[\frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} + \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta-4)^2}{2}} \right]$$

$$\propto e^{-\frac{\theta^2 + (x-\theta)^2}{2}} + e^{-\frac{(\theta-4)^2 + (x-\theta)^2}{2}} \Rightarrow \dots$$

$$\Rightarrow f(\theta/x) = \frac{e^{-x^2/4}}{e^{-x^2/4} + e^{-\frac{(x-4)^2}{4}}} \cdot \frac{1}{\sqrt{2\pi \cdot \frac{1}{2}}} e^{-\frac{(\theta - x/2)^2}{2 \cdot 1/2}} + \frac{e^{-\frac{(x-4)^2}{4}}}{e^{-\frac{x^2}{4}} + e^{-\frac{(x-4)^2}{4}}} \frac{1}{\sqrt{2\pi \cdot \frac{1}{2}}} e^{-\frac{\left(\theta - \frac{(x+4)}{2}\right)^2}{2 \cdot \frac{1}{2}}}$$



Penalidade : $\mathbb{I}_{R^c}(\theta)$ + Complementar (R) → otimizando temos o HPD.

07/05 - # Aula 13

"Statistical Information: A Bayesian Perspective". Rafael B. Stern, Carlos A. B. Pereira

Dissertação Chang Yu Fang (2008)

Garthwaite, Jolliffe e Jones (IC para Poisson(θ)) Pag. 111.

Lista 7

Teste de Hipótese

Θ

$\Theta_0, \Theta_1 \subseteq \Theta$ tais que $\Theta = \Theta_0 \cup \Theta_1$ e $\Theta_0 \cap \Theta_1 = \emptyset$.

$(\theta_1, \theta_2, \dots, \theta_k)$

$\theta_2 > 1/2$ $\quad \bigcup_{i=1}^{k} \{\theta_i > 1/2\}$

$\leq 1/2$ $\quad \bigcap_{i=1}^{k} \{\theta_i \leq 1/2\}$

Teste $\rightarrow$ problema de decisão com duas alternativas

$H_0: \theta \in \Theta_0$

$H_1: \theta \in \Theta_1$

$H_0: \theta = 10$

Exemplo:

θ brancas
$5-\theta$ verdes

$\Theta = \{0,1,2,3,4,5\}$ $\quad X = (X_1, X_2)$, onde $X_i = \begin{cases} 1, & \text{a i-ésima bola retirada é branca} \\ 0, & \text{c.c.} \end{cases}$

$\Theta_0 = \{3,4,5\}$

$\Theta_1 = \{0,1,2\}$ $\quad i = 1,2$

$$\mathcal{X} = \{(0,0),(0,1),(1,0),(1,1)\} = \{0,1\}^2$$

$H_0: \theta \in \Theta_0 \ (\theta \geq 3)$

$H_1: \theta \in \Theta_1 \ (\theta < 3)$

Definição: Uma função de teste (teste) é uma função $\varphi : \mathfrak{X} \to \{0,1\}$ para testar os $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$ que associa a cada $x \in \mathfrak{X}$ $\varphi(x) = 0$ no caso de decisão* por H_0 a partir de x ou $\varphi(x) = 1$ no caso de decisão** por H_1.

$H_0 : \theta \in \Theta_0 \to$ Hipótese Nula

$H_1 : \theta \in \Theta_1 \to$ Hipótese Alternativa

* aceitação H_0 / Não rejeição de H_0

** rejeição de H_0

$$\varphi_1(x_1, x_2) = \begin{cases} 0, & x_1 + x_2 \geq 1 \\ 1, & x_1 + x_2 = 0 \end{cases}$$

$$\varphi_2(x_1, x_2) = \begin{cases} 0, & x_1 + x_2 = 2 \\ 1, & x_1 + x_2 < 2 \end{cases}$$

$$\varphi_3(x_1, x_2) = \begin{cases} 0, & x_1 + x_2 = 0 \\ 1, & x_1 + x_2 \geq 1 \end{cases}$$

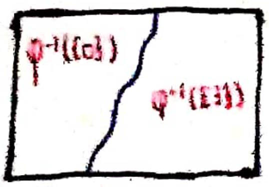
$$\varphi_4(x_1, x_2) = 0, \forall x_1, x_2$$

Nesse caso, temos 2^4 $(2^{|\mathfrak{X}|})$ testes.

Ao conjunto $\varphi^{-1}(\{1\}) = \{x \in \mathfrak{X} : \varphi(x) = 1\}$ damos o nome de Região Crítica ou Região de Rejeição (de H_0)

" " $\varphi^{-1}(\{0\}) = \{x \in \mathfrak{X} : \varphi(x) = 0\}$ " " " " Região de Aceitação de H_0

$\Psi = \{\varphi : \mathfrak{X} \to \{0,1\}$ tal que φ é mensurável$\}$



Como escolher $\varphi \in \Psi$ de maneira "ótima"?

Num problema de decisão com duas alternativas, você pode incorrer em dois possíveis erros:

(1) Você decidir por H_1 $(\varphi(x)=1)$ quando H_0 é verdadeira.

(rejeitar H_0) $\rightarrow$ erro de tipo I

(2) Você decidir por H_0 $(\varphi(x)=0)$ quando H_1 é verdadeira

(não rejeitar H_0) $\rightarrow$ erro de tipo II

Suponhamos

$\Theta' = \{2,4\}$ $\Theta_0 = \{2\}$ $\Theta_1 = \{4\}$

$H_0' : \theta = 2$ $\qquad \varphi_1'(x_1,x_2) = \mathbb{I}_{\{2\}}(x_1+x_2)$

$H_1' : \theta = 4$

$\varphi_2'(x_1,x_2) = \mathbb{I}_{\{1,2\}}(x_1+x_2)$

$\varphi_3'(x) = 0, \ \forall x \in \mathfrak{X}$

Assim, para o teste φ_1'

$\mathbb{P}(\text{Erro tipo I}) = \mathbb{P}(\text{Rejeitar } H_0 \text{ quando } H_0 \text{ é verdadeira}) =$

$= \mathbb{P}(\varphi_1'(X) = 1 / H_0 \text{ é verdadeira}) = \mathbb{P}(\varphi_1'(X) = 1 / \theta = 2) =$

$= \mathbb{P}(X_1 + X_2 = 2 / \theta = 2) = \dfrac{\binom{2}{2}\binom{3}{0}}{\binom{5}{2}} = \dfrac{1}{10}$ $\rightarrow$ Aqui temos A PROBABILIDADE DE ERRO DO TIPO I.

$X_1 + X_2 / \theta \sim HG(\theta, 5-\theta, 2)$

e

$\mathbb{P}(\text{Erro tipo II}) = \mathbb{P}(\text{Não rejeitar } H_0' \mid H_0' \text{ é falsa}) =$

$= \mathbb{P}(\varphi_1'(X) = 0 / \theta = 4) = \mathbb{P}(X_1 + X_2 \neq 2 / \theta = 4) = 1 - \mathbb{P}(X_1 + X_2 = 2 / \theta = 4) = 1 - \dfrac{\binom{4}{2}\binom{1}{0}}{\binom{5}{2}} = \dfrac{4}{10}$

Teste φ_2':

$$\mathbb{P}(\text{Erro tipo I}) = \mathbb{P}(\text{Rejeitar } H_0' \mid H_0' \text{ é verdadeira}) =$$

$$= \mathbb{P}(\varphi_2'(X) = 1 / \theta = 2) = \mathbb{P}(X_1 + X_2 \geq 1 / \theta = 2) =$$

$$= 1 - \mathbb{P}(X_1 + X_2 = 0 / \theta = 2) = 1 - \frac{\binom{2}{0}\binom{3}{2}}{\binom{5}{2}} = \frac{7}{10}$$

$$\mathbb{P}(\text{Erro tipo II}) = \mathbb{P}(\text{Não rejeitar } H_0' \mid H_0' \text{ é falsa}) =$$

$$= \mathbb{P}(\varphi_2'(X) = 0 / \theta = 4) = \mathbb{P}(X_1 + X_2 = 0 / \theta = 4) = \frac{\binom{4}{0}\binom{1}{2}}{\binom{5}{2}} = 0$$

Teste φ_3':

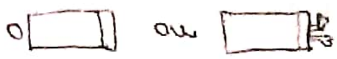
$$\mathbb{P}(\text{Erro tipo I}) = \mathbb{P}(\varphi_3'(X) = 1 / \theta = 2) = 0$$

$$\mathbb{P}(\text{Erro tipo II}) = \mathbb{P}(\varphi_3'(X) = 0 / \theta = 4) = 1$$

Logo, temos

Teste	$\mathbb{P}$(Erro I)	$\mathbb{P}$(Erro II)
φ_1'	1/10	4/10
φ_2'	7/10	0
φ_3'	0	1

Analogia do cobertor curto
pequeno



Cobre a cabeça e não os pés ou vice-versa.
Não dá pra minimizar ambos!

Voltando ao problema original.

$\Theta = \{0, 1, 2, 3, 4, 5\}$

$H_0: \theta \geq 3$.

$H_1: \theta < 3$.

$\varphi_1(x) = \mathbb{I}_{\{0\}}(x_1+x_2)$ $\quad \varphi_2(x) = \mathbb{I}_{\{0,1\}}(x_1+x_2)$

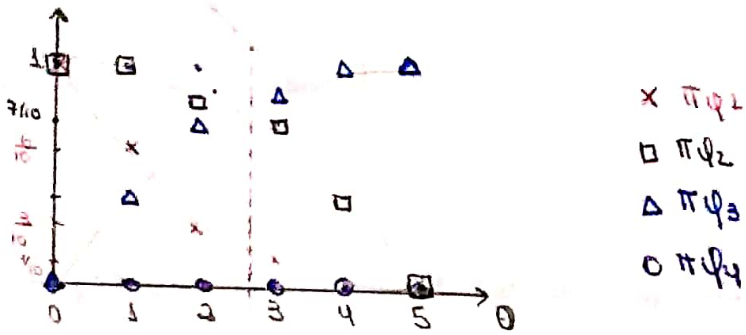
$\varphi_3(x) = \mathbb{I}_{\{1,2\}}(x_1+x_2)$ $\quad \varphi_4(x) = 0, \forall x$

Seja $\varphi: \mathcal{X} \to \{0,1\}$ um teste para $H_0: \theta \in \Theta_0$ o número

$$\pi_\varphi(\theta) = \mathbb{P}(\varphi(X)=1|\theta)$$

damos o nome de FUNÇÃO PODER DO TESTE φ.

($\beta_\varphi(\theta) = 1 - \pi_\varphi(\theta) \Rightarrow$ curva característica de operação).



$$\pi_{\varphi_1}(\theta) = \mathbb{P}(\varphi_1(X)=1|\theta) = P(X_1+X_2=0|\theta) = \frac{\binom{\theta}{0}\binom{5-\theta}{2}}{\binom{5}{2}} = \frac{(5-\theta)(4-\theta)}{20}$$

$$\pi_{\varphi_2}(\theta) = \mathbb{P}(\varphi_2(X)=1|\theta) = \mathbb{P}(X_1+X_2 \leq 0|\theta) = 1-\mathbb{P}(X_1+X_2=2|\theta) = 1 - \frac{\binom{\theta}{2}\binom{5-\theta}{0}}{\binom{5}{2}} = 1 - \frac{\theta(\theta-1)}{20}$$

$$\pi_{\varphi_3}(\theta) = \mathbb{P}(\varphi_3(X)=1|\theta) = \mathbb{P}(X_1+X_2 \geq 1|\theta) = 1-\mathbb{P}(X_1+X_2=0|\theta) = 1 - \frac{(5-\theta)(4-\theta)}{20}$$

Dizemos que uma hipótese

$$H_i: \theta \in \Theta_i \text{ é SIMPLES se } |\Theta_i| = 1$$

Dizemos que $H_i: \theta \in \Theta_i$ é composta se $|\Theta_i| > 1$.

Exemplo 1.

$H_0: \theta \geq 3 \rightarrow$ composta

$H_0: \theta < 3 \rightarrow$ composta

$H_0'': \theta = 2 \rightarrow$ simples

$H_0''': \theta \in \{2, 4\} \rightarrow$ composta

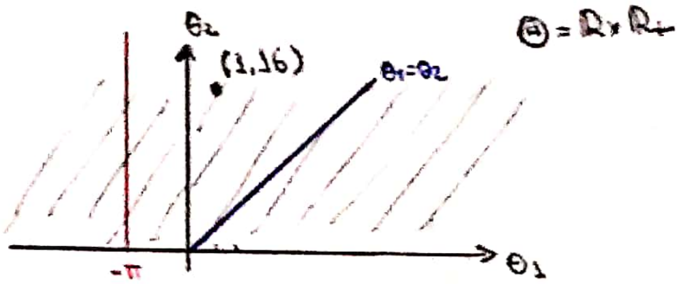
Exemplo 2.

$X \mid \theta = (\theta_1, \theta_2) \sim N(\theta_1, \theta_2)$

$H_0: \theta = (1, 16) \rightarrow$ simples

$H_1: \theta_1 = -\pi \rightarrow$ composta

$H_2: \theta_1 = \theta_2 \rightarrow$ composta



Seja $\varphi: \mathfrak{X} \rightarrow \{0,1\}$ um teste para $H_0: \theta \in \Theta_0$ versus $\Theta_1: \theta \in \Theta_1$.

Ao número $\sup_{\theta \in \Theta_0} \pi_\varphi(\theta)$ damos o nome de Tamanho do Teste φ.

Notação.

$$\alpha^*_\varphi = \sup_{\theta \in \Theta_0} \pi_\varphi(\theta)$$

No exemplo da urna:

$$\alpha^*_{\varphi_1} = \sup_{\theta \in \Theta_0} \pi_{\varphi_1}(\theta) = \sup_{\theta \in \{3,4,5\}} \pi_{\varphi_1}(\theta) = \frac{1}{10}$$

$$\alpha^*_{\varphi_2} = \sup_{\theta \in \Theta_0} \pi_{\varphi_2}(\theta) = \frac{7}{10}$$

$$\alpha^*_{\varphi_3} = \sup_{\theta \in \Theta_0} \pi_{\varphi_3}(\theta) = 1$$

Analogamente, $\alpha^*_{\varphi_4} = 0$.

Definição: Um teste ϕ é de nível α, $0 \leq \alpha \leq 1$, se

$$\alpha^*_\phi \leq \alpha .$$

[illegible]

Teste Para Hipóteses Simples - Aula 14

$\Theta = \{\theta_0, \theta_1\}$ $\quad \Theta_0 = \{\theta_0\}$, $\Theta_1 = \{\theta_1\}$

$H_0: \theta = \theta_0 \quad (\theta \in \Theta_0)$

$H_1: \theta = \theta_1 \quad (\theta \in \Theta_1)$

$\Psi = \{\varphi: \mathcal{X} \to \{0,1\}\}$ onde $\varphi \in \Psi$ é um teste para Ho versus H_1.

Para $\varphi \in \Psi$,

$\pi_\varphi(\theta_0) = P(\varphi(X) = 1 \mid \theta_0)$: prob. erro tipo I

$1 - \pi_\varphi(\theta_1) = P(\varphi(X) = 0 \mid \theta_1)$: prob. erro tipo II.

Resultado:

Seja $\varphi^*: \mathcal{X} \to \{0,1\}$ dada por

$$\varphi^*(x) = \begin{cases} 1, & a f(x|\theta_0) < b f(x|\theta_1) \\ 0, & \text{c.c.} \end{cases} \quad \left(\frac{f(x|\theta_1)}{f(x|\theta_0)} > \frac{a}{b} \right)$$

$a, b > 0$ constantes

Então, $\forall \varphi \in \Psi$

$$a \pi_{\varphi^*}(\theta_0) + b(1 - \pi_{\varphi^*}(\theta_1)) \leq a \pi_\varphi(\theta_0) + b(1 - \pi_\varphi(\theta_1))$$

Para $\varphi \in \Psi$,

$$a \pi_\varphi(\theta_0) + b(1 - \pi_\varphi(\theta_1)) = a P(\varphi(X) = 1 \mid \theta_0) + b(1 - P(\varphi(X) = 1 \mid \theta_1)) =$$

$$= a \int_{\varphi^{-1}\{1\}} f(x|\theta_0)\,dx + b\left(1 - \int_{\varphi^{-1}(\{1\})} f(x|\theta_1)\,dx\right) =$$

$$= \int_{\varphi^{-1}(\{1\})} a f(x/\theta_0)\,dx + b - \int_{\varphi^{-1}(\{1\})} b f(x/\theta_1)\,dx =$$

$$= b + \int_{\varphi^{-1}(\{1\})} [a f(x/\theta_0) - b f(x/\theta_1)]\,dx \geq b + \int_{\varphi^{*-1}(\{1\})} [a f(x/\theta_0) - b f(x/\theta_1)]\,dx$$

$$\Rightarrow a\,\pi_\varphi(\theta_0) + b(1 - \pi_\varphi(\theta_1)) \geq a\,\pi_{\varphi^*}(\theta_0) + b(1 - \pi_{\varphi^*}(\theta_1))$$

Exemplo:

$\mathcal{X} = \{0,1\}^2$ $\quad a = 1,\ b = 2$

$H_0: \theta = 2$

$H_1: \theta = 4$

x	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$f(x/\theta=2)$	$3/10$	$3/10$	$3/10$	$1/10$
$f(x/\theta=4)$	0	$2/10$	$2/10$	$6/10$
$\frac{f(x/\theta=4)}{f(x/\theta=2)}$	0	$2/3$	$2/3$	6

$$\varphi^*(x) = \begin{cases} 1, & \frac{f(x/\theta=4)}{f(x/\theta=2)} > 1/2 \\ 0, & \text{c.c.} \end{cases}$$

Assim,

$$\varphi^*(x) = \begin{cases} 1, & x_1 + x_2 \geq 1 \\ 0, & \text{c.c.} \end{cases} = \varphi_2'(x)$$

Fazendo

$a' = b' = 1$ (erros tipo I e II com mesmo peso)

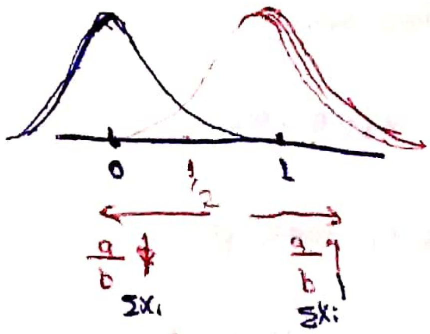
(mais resistente à rejeitar

$$\varphi^{*'}(x) = \begin{cases} 1, & x_1 + x_2 = 2 \\ 0, & \text{c.c.} \end{cases} = \varphi_1'(x)$$

$a'' = 7, \quad b'' = 1$

$\varphi^*(x) = 0, \ \forall x \in \mathcal{X}$

Exemplo 2: $X_1, \dots, X_n$, dado θ, c.i.i.d. $N(\theta, 1)$.



$H_0: \theta = 0 \qquad a = 2, \ b = 1$

$H_1: \theta = 1 \qquad \mathcal{X} = \mathbb{R}^n$

$$\frac{f(x|\theta=1)}{f(x|\theta=0)} = \frac{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-1)^2}{2}}}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}} = \frac{e^{-\frac{\sum(x_i-1)^2}{2}}}{e^{-\frac{\sum x_i^2}{2}}} = e^{\frac{2\sum x_i - n}{2}} = e^{\sum x_i - \frac{n}{2}}$$

$$\varphi^*(x) = 1 \Leftrightarrow \frac{f(x|\theta=1)}{f(x|\theta=0)} > \frac{2}{1} \Leftrightarrow e^{\sum x_i - n/2} > 2 \Leftrightarrow$$

$$\Leftrightarrow \sum x_i - \frac{n}{2} > \log 2 \Leftrightarrow \bar{x}_n > \frac{1}{2} + \frac{\log 2}{n}.$$

$$\varphi^*(x) = \begin{cases} 1, & \bar{x}_n > 1/2 + \log 2/n \\ 0, & \text{c.c.} \end{cases}$$

1. $H_0: \theta \in \Theta_0$

$H_1: \theta \in \Theta_1$

$\Theta_1 = \{\theta_1\}$ (Θ_1 é hipótese simples)

Definição: Um teste de nível α, $0 \leq \alpha \leq 1$, φ^* é um TESTE MAIS PODEROSO (MP) de nível α para $H_0: \theta \in \Theta_0$ versus $\Theta_1: \theta = \theta_1$ se

$$\pi_{\varphi^*}(\theta_1) \geq \pi_{\varphi}(\theta_1), \ \forall \varphi \text{ de nível } \alpha.$$

Lema de Neyman-Pearson

$H_0: \theta = \theta_0$ versus $H_1: \theta = \theta_1$

Seja $\varphi^*: \mathfrak{X} \to \{0, 1\}$ dada por

$$\varphi^*(x) = \begin{cases} 1, & f(x|\theta_0) < k f(x|\theta_1) \\ 0, & \text{c.c.} \end{cases}, \quad k > 0$$

Então,

$$\pi_{\varphi^*}(\theta_1) \geq \pi_{\varphi}(\theta_1), \ \forall \varphi \in \Psi^* = \{\varphi \in \Psi : \alpha_{\varphi} \leq \pi_{\varphi^*}(\theta_0)\}.$$

Isto é, nessas condições, φ^* é MP de nível $\pi_{\varphi^*}(\theta_0)$.

Isto é, nessas condições

ψ^* é MD de nível $\pi_{\psi^*}(\theta_0)$.

Verificação:

Pelo resultado anterior,

$$\forall \psi \in \Psi, \quad \pi_\psi(\theta_0) + K(1-\pi_\psi(\theta_1)) \geq \pi_{\psi^*}(\theta_0) + K(1-\pi_{\psi^*}(\theta_1))$$

$$\forall \psi \in \Psi^*, \quad \pi_{\psi^*}(\theta_0) \geq \pi_\psi(\theta_0)$$

$\Rightarrow$

$$\Rightarrow \forall \psi \in \Psi^*, \quad K(1-\pi_\psi(\theta_1)) \geq K(1-\pi_{\psi^*}(\theta_1)) \Rightarrow$$

$$\Rightarrow \forall \psi \in \Psi^*, \quad \pi_{\psi^*}(\theta_1) \geq \pi_\psi(\theta_1).$$

Em geral, fixa-se um valor α máximo para os tamanhos dos testes. Escolhemos, se possível, ψ^* cujo tamanho, α_{ψ^*}, atinge o valor α

Exemplo: $X_1, \dots, X_n$, dado θ, c.i.i.d. $N(\theta, 1)$

$H_0: \theta = \theta_0$
$H_1: \theta = \theta_1$ $\quad (\theta_0 < \theta_1)$

$$\psi^*(x) = 1 \Leftrightarrow \frac{f(x|\theta_1)}{f(x|\theta_0)} > \frac{1}{K} \Leftrightarrow \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum(x_i-\theta_1)^2}{2}}}{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum(x_i-\theta_0)^2}{2}}} > K^{-1} \Leftrightarrow$$

$$e^{\frac{\sum x_i^2 - 2\theta_0\sum x_i + n\theta_0^2 - \sum x_i^2 + 2\theta_1 \sum x_i - n\theta_1^2}{2}} > K^{-1} \Leftrightarrow$$

$$e^{(\theta_1-\theta_0)\sum_{i=1}^{n} x_i} > k^{-1} e^{\frac{n\theta_1^2 - n\theta_0^2}{2}} \Leftrightarrow$$

$$(\theta_1-\theta_0)\sum_{i=1}^{n} x_i > \log k^{-1} e^{\frac{n\theta_1^2 - n\theta_0^2}{2}} \Leftrightarrow$$

$$\sum_{i=1}^{n} x_i > C \text{ , onde } C = \frac{\log k^{-1} e^{\frac{n\theta_1^2 - n\theta_0^2}{2}}}{\theta_1 - \theta_0}$$

Assim, escolhemos $C \in \mathbb{R}$ de modo $\pi_{\varphi^*}(\theta_0) = \alpha$.

$$\pi_{\varphi^*}(\theta_0) = P(\varphi^*(X) = 1 \mid \theta_0) = P\left(\sum_{i=1}^{n} X_i > C \mid \theta_0\right) =$$

$$= P\left(\frac{\sum X_i - n\theta}{\sqrt{n}} > \frac{C - n\theta_0}{\sqrt{n}} \middle| \theta_0\right) = \alpha \Rightarrow$$

$$\Rightarrow \frac{C - n\theta_0}{\sqrt{n}} = \Phi^{-1}(1-\alpha) \Rightarrow C = n\theta_0 + \sqrt{n}\,\Phi^{-1}(1-\alpha)$$

Assim,

$$\varphi^*(x) = \begin{cases} 1, & \sum x_i > n\theta_0 + \sqrt{n}\,\Phi^{-1}(1-\alpha) \\ 0, & \text{c.c.} \end{cases}$$

* Em geral, fixa-se um valor α máximo para os tamanhos dos testes. Escolhemos, se possível, φ^* cujo tamanho, α_{φ^*}, atinge o valor α

Exemplo 2. $X_1, \ldots, X_n$, dado θ, c.i.i.d. $Exp(\theta)$

$H_0: \theta = \theta_0 \qquad (\theta_0 < \theta_1) \qquad x = (x_1, \ldots, x_n) \in \mathbb{R}_+^n$

$H_1: \theta = \theta_1$

$$\varphi^*(x)=1 \Leftrightarrow \frac{\varphi(x|\theta_1)}{\varphi(x|\theta_0)} > \frac{1}{k} \Leftrightarrow \frac{\prod_{i=1}^{n} f(x_i|\theta_1)}{\prod_{i=1}^{n} f(x_i|\theta_0)} > k^{-1} \Leftrightarrow \left(\frac{\prod_{i=1}^{n} \theta_1 e^{-\theta_1 x_i}}{\prod_{i=1}^{n} \theta_0 e^{-\theta_0 x_i}}\right) > k^{-1} \Leftrightarrow$$

$$\left(\frac{\theta_1}{\theta_0}\right)^n e^{\sum x_i(\theta_0-\theta_1)} > \frac{1}{k} \Leftrightarrow \sum_{i=1}^{n} x_i(\theta_0-\theta_1) > \log k^{-1}\left(\frac{\theta_0}{\theta_1}\right)^n \Leftrightarrow$$

$$\sum_{i=1}^{n} x_i < \frac{\log k^{-1}\left(\frac{\theta_0}{\theta_1}\right)^n}{\theta_0-\theta_1} = c$$

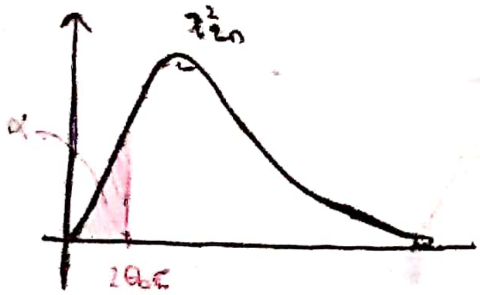
Assim,

$\pi_{\varphi^*}(\theta_0)=\alpha$ $\qquad \sum X_i|\theta_0 \sim Gama(n,\theta_0)$

$\pi_{\varphi^*}(\theta_0) = P(\varphi^*(X)=1|\theta_0) = P(\sum_{i=1}^{n} X_i < c|\theta_0) =$

$= P(2\theta_0\sum_{i=1}^{n} X_i < 2\theta_0 c|\theta_0) = \alpha$

$2\theta_0\sum X_i|\theta_0 \sim Gama\left(n, \frac{\theta_0}{2\theta_0}\right) = \chi^2_{2n}$



Logo, $2\theta_0 c = q_{2n,\alpha} \Rightarrow c = \frac{q_{2n,\alpha}}{2\theta_0}$.

$$\varphi^*(x) = \begin{cases} 1, & \sum x_i < q_{2n,\alpha}/2\theta_0 \\ 0, & c.c. \end{cases}$$

Exemplo 3: $H_0: \theta=\theta_0 \times H_1: \theta=\theta_1$ $\qquad \mathcal{X}=\{1,2,3,4,5,6\}$

x	1	2	3	4	5	6
$f(x\|\theta_0)$	0,1	0,2	0,2	0,1	0,2	0,2
$f(x\|\theta_1)$	0,3	0,2	0,1	0,2	0,1	0,1
$\frac{f(x\|\theta_1)}{f(x\|\theta_0)}$	3	1	1/2	2	1/2	1/2

$\alpha = 20\%$

$P(X=1|\theta_0)=10\%$

$P\{x\in\{1,4\}|\theta_0\}=20\%$

$\varphi^*(x) = \begin{cases} 1, & x\in\{1,4\} \\ 0, & c.c. \end{cases}$ é o teste MP de nível 20%

Analogamente

$$\varphi^{**}(x) = \begin{cases} 1, & x \in \{1,2,4\} \\ 0, & c.c. \end{cases}$$ é o teste MP de nível 40%

$\alpha = 25\%$

$$\varphi^{*}(x) = \begin{cases} 1, & x \in \{1,4\} \\ \gamma, & x = 2 \\ 0, & c.c. \end{cases}$$

Você obtém γ de modo que

$$E[\varphi_A^*(X) \mid \theta_0] = \alpha = 25\%.$$

$$E[\varphi_A^*(x) \mid \theta_0] = 1 . P(\varphi_A^*(x) = 1 \mid \theta) + \gamma\, P(\varphi_A^*(x) = \gamma \mid \theta_0) + 0 . P(\varphi_A^*(x) = 0 \mid \theta_0) =$$

$$= 1 . P(X \in \{1,4\} \mid \theta_0) + \gamma . P(X = 2 \mid \theta_0) = 1 . (0,2) + \gamma\,(0,2) = 0,25 \Rightarrow$$

$$\varphi_A^*(x) = \begin{cases} 1, & x \in \{1,4\} \\ 1/4, & x = 2 \\ 0, & c.c. \end{cases} \qquad \Rightarrow \gamma = \frac{1}{4}$$

Exemplo 4 $H_0: \theta = \theta_0 \times H_1: \theta = \theta_1$

x	1	2	3	4	5
$f(x/\theta_0)$	0,02	0,02	0,06	0,4	0,5
$f(x/\theta_1)$	0,6	0,2	0	0	0
$\frac{f(x/\theta_1)}{f(x/\theta_0)}$	40	10	0	0	0

$\alpha = 5\%$

$P(X = 1 \mid \theta_0) = 2\%$

$P(X \leq 2 \mid \theta_0) = 4\%$

$P(X \leq 3 \mid \theta_0) = 10\%$

~~$\alpha = 5\%$~~

$$\varphi_A^*(x) = \begin{cases} 1, & x \leq 2 \\ 1/6, & x = 3 \\ 0, & x > 3 \end{cases} \qquad \pi_{\varphi_A^*}(\theta_0) = 0{,}05 \qquad \pi_{\varphi_A^*}(\theta_1) = 1$$

$$\varphi^*(x) = \begin{cases} 1, & x \leq 2 \\ 0, & x > 2 \end{cases} - \quad \pi_{\varphi^*}(\theta_0) = 0{,}04 \qquad \pi_{\varphi^*}(\theta_1) = 1$$

Exemplo: $X_1, \ldots, X_{10}$ dado θ, c.i.i.d. $Ber(\theta)$

$H_0: \theta = 1/2$ $\qquad x = (x_1, \ldots, x_{10}) \in \mathcal{X} = \{0,1\}^{10}$

$H_1: \theta = 8/10$

$$\varphi^*(x) = 1 \Leftrightarrow \frac{f(x/\theta = 8/10)}{f(x/\theta = 1/2)} = \frac{\left(\frac{8}{10}\right)^{\Sigma x_i} \left(\frac{2}{10}\right)^{n - \Sigma x_i}}{\left(1/2\right)^{\Sigma x_i} \left(1/2\right)^{n - \Sigma x_i}} > k^{-1} \Leftrightarrow$$

$$\Leftrightarrow \frac{2^n}{10^n} 8^{\Sigma x_i} 2^{n - \Sigma x_i} > \frac{1}{k} \Leftrightarrow \left(\frac{4}{10}\right)^n 4^{\Sigma x_i} > \frac{1}{k} \Leftrightarrow 4^{\Sigma x_i} > k^{-1} \left(\frac{10}{4}\right)^n \Leftrightarrow$$

$$\Leftrightarrow \Sigma x_i > \frac{\log k^{-1} \left(\frac{10}{4}\right)^n}{\log 4} = c$$

$\alpha = 5\%$

$$P(X \geq 10 \mid \theta = 1/2) = (1/2)^{10} = \frac{1}{1024} \leq 5\%$$

$$P(X \geq 9 \mid \theta = 1/2) = \binom{10}{9}\left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right) = \frac{11}{1024} < 5\%$$

$$P(X \geq 8 \mid \theta = 1/2) = \binom{10}{8}\left(\frac{1}{2}\right)^{10} + \binom{10}{9}\left(\frac{1}{2}\right)^{10} + \frac{1}{1024} = \frac{56}{1024} > 5\%$$

Assim,

$$\varphi^*(x) = \begin{cases} 1, & x \in \{9, 10\} \\ 0, & \text{c.c.} \end{cases}$$ é o teste M.P de nível $\frac{11}{1024}$

$$\varphi^{**}(x) = \begin{cases} 1, & x \in \{8, 9, 10\} \\ 0, & \text{c.c.} \end{cases}$$ é o teste MR de nível $\frac{56}{1024}$.

Aula 15

Lista 8

Avaliação 2 -

Avaliação 3 - 02/07

$\Theta = \{\theta_0, \theta_1\}$ $\Theta_i = \{\theta_i\}$, $i = 0, 1$.

$H_0 : \theta = \theta_0 \; (\theta \in \Theta_0)$

$H_1 : \theta = \theta_1 \; (\theta \in \Theta_1)$

$\mathbb{P}(\theta = \theta_1) = p_1$

$0 < p_1 < 1$.

$\mathbb{P}(\theta = \theta_0) = p_0 = 1 - p_1$

Decisão	$\theta = \theta_0$	$\theta = \theta_1$
Não Rejeitar H_0 (0)	0 ($L(0, \theta_0)$)	a_2 ($L(0, \theta_1)$)
Rejeitar H_0 (1)	a_1 ($L(1, \theta_0)$)	0 ($L(1, \theta_1)$)

$a_1, a_2 > 0$.

$\Psi = \{\varphi : \mathcal{X} \to \{0, 1\}\}$. Devemos escolher $\varphi^* \in \Psi$ que minimiza

$$\rho(\varphi) = E[L(\varphi(X), \theta)]$$

Para $\varphi \in \Psi$,

$$\rho(\varphi) = E[\underbrace{L(\varphi(X), \theta)}_{g(X, \theta)}] = \sum_{i=0}^{1} \sum_{x \in \mathcal{X}} L(\varphi(x), \theta_i) P(\theta = \theta_i, X = x) =$$

$$= \sum_{i=0}^{1} \left\{ \sum_{x \in \mathcal{X}} L(\varphi(x), \theta_i) P(X = x / \theta = \theta_i) \right\} \overbrace{P(\theta = \theta_i)}^{p_i} =$$

$$= p_0 \left\{ \overbrace{\sum_{x \in \varphi^{-1}(1)} L(\overset{1}{\varphi(x)}, \theta_0) P(X=x/\theta_0)} + \overbrace{\sum_{x \in \varphi^{-1}(0)} L(\overset{0}{\varphi(x)}, \theta_0) P(X=x/\theta_0)}^{0} \right\} +$$

$$+ p_1 \left\{ \underbrace{\sum_{x \in \varphi^{-1}(1)} L(\overset{1}{\varphi(x)}, \theta_1) P(X=x/\theta_1)}_{0} + \underbrace{\sum_{x \in \varphi^{-1}(0)} L(\overset{0}{\varphi(x)}, \theta_1) P(X=x/\theta_1)} \right\} =$$

$$= p_0 \left\{ \sum_{x \in \varphi^{-1}(1)} a_1 P(X=x/\theta_0) \right\} + p_1 \left\{ \sum_{x \in \varphi^{-1}(0)} a_2 P(X=x/\theta_1) \right\} =$$

$$= p_0 a_1 P(\varphi(X)=1/\theta_0) + p_1 a_2 P(\varphi(X)=0/\theta_1) \Rightarrow$$

$$\Rightarrow \quad \rho(\varphi) = \overset{A}{\widehat{p_0 \, a_1}} \, \pi_\varphi(\theta_0) + \overset{B}{\widehat{p_1 a_2}} \, (1 - \pi_\varphi(\theta_1)) \geq$$

$$\geq p_0 a_1 \pi_{\varphi^*}(\theta_0) + p_1 a_2 (1 - \pi_{\varphi^*}(\theta_1)) = \rho(\varphi^*),$$

onde $\varphi^* \in \Psi$ é dado por

$$\varphi^*(x) = \begin{cases} 1, & p_0 a_1 P(X=x/\theta_0) < p_1 a_2 P(X=x/\theta_1) \\ 0, & \text{c.c.} \end{cases}$$

φ^* é chamado, nesse caso, de <u>teste de Bayes.</u>

Note que

$$\varphi^*(x) = 1 \Leftrightarrow a_1 P(X=x/\theta_0) P(\theta=\theta_0) < a_2 P(X=x/\theta_1) P(\theta=\theta_1) \Leftrightarrow$$

$$\Leftrightarrow a_1 \frac{P(X=x/\theta_0) P(\theta=\theta_0)}{P(X=x)} < a_2 \frac{P(X=x/\theta_1) P(\theta=\theta_1)}{P(X=x)} \Leftrightarrow$$

$$\Leftrightarrow a_1 \overbrace{P(\theta=\theta_0/X=x)}^{1 - P(\theta=\theta_1/X=x)} < a_2 P(\theta=\theta_1/X=x) \Rightarrow$$

$$\varphi^*(x) = 1 \Leftrightarrow P(\Theta = \Theta_1 / X = x) > \frac{a_1}{a_1 + a_2}$$

Exemplos:

(1) $X_1, \dots, X_n$ dado Θ, c.i.i.d. $N(\Theta, 1)$

$H_0: \Theta = 0$

$H_1: \Theta = 1$

$x = (x_1, \dots, x_n) \in \mathcal{X}$

$$P(\Theta = 1 / X = x) = \frac{\varphi(x \mid \Theta = 1)\, P(\Theta = 1)}{\varphi(x \mid \Theta = 1) P(\Theta = 1) + \varphi(x \mid \Theta = 0)\, P(\Theta = 0)} =$$

$$= \frac{\cancel{\left(\frac{1}{\sqrt{2\pi}}\right)^n} e^{-\frac{\sum_{i=1}^n (x_i - 1)^2}{2}} \cdot p_1}{\cancel{\left(\frac{1}{\sqrt{2\pi}}\right)^n} e^{-\frac{\sum_{i=1}^n (x_i - 1)^2}{2}} \cdot p_1 + \cancel{\left(\frac{1}{\sqrt{2\pi}}\right)^n} e^{-\frac{\sum_{i=1}^n x_i^2}{2}} \cdot p_0} =$$

$$= \frac{p_1}{p_1 + p_0\, e^{\frac{\sum (x_i - 1)^2}{2} - \frac{\sum x_i^2}{2}}} = \frac{p_1}{p_1 + p_0\, e^{\frac{-2\sum x_i + n}{2}}} > \frac{a_1}{a_1 + a_2} \Leftrightarrow$$

$$\Leftrightarrow p_1 a_2 > p_0 a_1\, e^{\frac{n}{2} - \sum x_i} \Leftrightarrow \log \frac{p_1 a_2}{p_0 a_1} > \frac{n}{2} - \sum x_i \Leftrightarrow$$

$$\sum_{i=1}^n x_i > \frac{n}{2} - \log\left(\frac{p_1 a_2}{p_0 a_1}\right)$$

- Refazer para

$H_0: \Theta = \Theta_0$

$H_1: \Theta = \Theta_1$

Exemplo.

(1) $X_1, \dots, X_n$, dado θ, c.i.i.d. $Exp(\theta)$.

$H_0: \Theta = \Theta_0$

$H_1: \Theta = \Theta_1$

$P(\Theta = 1) = p_1$

$x = (x_1, \dots, x_n) \in \mathcal{X}$

$$P(\Theta = 1 / X = x) = \frac{f(x/\Theta_1)\, p_1}{f(x/\Theta_1)\, p_1 + f(x/\Theta_0)\, p_0} =$$

$$= \frac{\Theta_1^n e^{-\Theta_1 \Sigma x_i} p_1}{\Theta_1^n e^{-\Theta_1 \Sigma x_i} p_1 + \Theta_0^n e^{-\Theta_0 \Sigma x_i} \cdot p_0} =$$

$$= \frac{p_1 \Theta_1^n}{p_1 \Theta_1^n + p_0 \Theta_0^n e^{(\Theta_1 - \Theta_0)\Sigma x_i}} > \frac{a_1}{a_1 + a_2} \Leftrightarrow$$

$$\Leftrightarrow a_2 p_1 \Theta_1^n > a_1 p_0 \Theta_0^n e^{(\Theta_1 - \Theta_0)\Sigma x_i} \Leftrightarrow$$

$$\Leftrightarrow \sum_{i=1}^n x_i < \boxed{\frac{\log\left(\frac{a_2 p_1 \Theta_1^n}{a_1 p_0 \Theta_0^n}\right)}{(\Theta_1 - \Theta_0)}}$$

$$\varphi^*(x) = \begin{cases} 1, & \sum_{i=1}^n x_i < \square \\ 0, & \text{c.c} \end{cases}$$

- Rafael Izbicki (2010)
  L dualidade de hipóteses
- Gustavo Miranda da Silva (2010)
- Victor Fossaluza (2008)

Exemplo 3: $X_1, \dots, X_n$, dado Θ, c.i.i.d. $Ber(\Theta)$

$$H_0: \Theta = \Theta_0 \qquad (0 < \Theta_0 < \Theta_1 < 1)$$
$$H_1: \Theta = \Theta_1$$

$P(\Theta = \Theta_i) = p_i, \; i = 0, 1 \qquad \mathfrak{X} = \{0, 1\}^n$

$x = (x_1, \dots, x_n) \in \mathfrak{X}$

$$P(\Theta = \Theta_1 / X = x) = \frac{P(X = x / \Theta_1)\, p_1}{P(X = x / \Theta_1)\, p_1 + P(X = x / \Theta_0)\, p_0} = \frac{\Theta_1^{\sum x_i} (1 - \Theta_1)^{n - \sum x_i}\, p_1}{\Theta_1^{\sum x_i} (1 - \Theta_1)^{n - \sum x_i}\, p_1 + \Theta_0^{\sum x_i} (1 - \Theta_0)^{n - \sum x_i}\, p_0} =$$

$$= \frac{p_1 \left(\frac{\Theta_1}{1 - \Theta_1}\right)^{\sum x_i} (1 - \Theta_1)^n}{p_1 \left(\frac{\Theta_1}{1 - \Theta_1}\right)^{\sum x_i} (1 - \Theta_1)^n + p_0 \left(\frac{\Theta_0}{1 - \Theta_0}\right)^{\sum x_i} (1 - \Theta_0)^n} = \frac{p_1 (1 - \Theta_1)^n}{p_1 (1 - \Theta_1)^n + p_0 (1 - \Theta_0)^n \left(\frac{\Theta_0 (1 - \Theta_1)}{(1 - \Theta_0)\Theta_1}\right)^{\sum x_i}} > \frac{a_1}{a_1 + a_2}$$

$\Leftrightarrow$

$$a_2\, p_1 (1 - \Theta_1)^n > a_1\, p_0 (1 - \Theta_0)^n \left(\frac{\Theta_0 (1 - \Theta_1)}{\Theta_1 (1 - \Theta_0)}\right)^{\sum x_i} \Leftrightarrow$$

$$\sum_{i=1}^{n} x_i > \frac{\log\left(\frac{a_2\, p_1 (1 - \Theta_1)^n}{a_1\, p_0 (1 - \Theta_0)^n}\right)}{\log\left(\frac{\Theta_0 (1 - \Theta_1)}{\Theta_1 (1 - \Theta_0)}\right)}.$$

Exemplo 4.

x	1	2	3	4	5
$P(X = x / \Theta_0)$	0,02	0,02	0,06	0,4	0.5
$P(X = x / \Theta_1)$	0,8	0,2	0	0	0

$P(\Theta = \Theta_0) = p_0$

$$P(\Theta = \Theta_1 / X = x) = \frac{P(X = x / \Theta = \Theta_1)\, p_1}{P(X = x / \Theta = \Theta_1)\, p_1 + P(X = x / \Theta = \Theta_0)\, p_0} =$$

$$= \begin{cases} 0 & , x = 3,4,5 \\ \dfrac{0{,}8\, p_1}{0{,}8\, p_1 + 0{,}02\, p_0} = \dfrac{40\, p_1}{40\, p_1 + p_0} & , x = 1 \\ \dfrac{0{,}2\, p_1}{0{,}2\, p_1 + 0{,}02\, p_0} = \dfrac{10\, p_1}{10\, p_1 + p_0} & , x = 2 \end{cases}$$

Considerando

$p_1 = \frac{1}{3}$, $p_0 = \frac{2}{3}$

$a_1 = 9$, $a_2 = 2$ $\rightarrow \dfrac{a_1}{a_1+a_2} = \dfrac{9}{10}$

$$\mathbb{P}(\theta = 1 / X = x) = \begin{cases} 0 & , x = 3,4,5 \\ \frac{40}{42} & , x = 1 \\ 10/12 & , x = 2 \end{cases} \rightarrow \varphi^*(x) = \begin{cases} 1 & , x = 1 \\ 0 & , \text{c.c.} \end{cases}$$

$a_1 = 4$, $a_2 = 1$

$\dfrac{a_1}{a_1+a_2} = \dfrac{4}{5} = 80\%$ $\qquad \varphi^*(x) = \begin{cases} 1 & , x = 1,2 \\ 0 & , \text{c.c.} \end{cases}$

Problema Passado

$\mathbb{P}(\theta = \theta_0) = p_0 \qquad \mathbb{P}(\theta = \theta_1) = p_1$

Hipóteses Compostas

$\Theta = \Theta_0 \cup \Theta_1, \quad \Theta_0 \cap \Theta_1 = \emptyset$

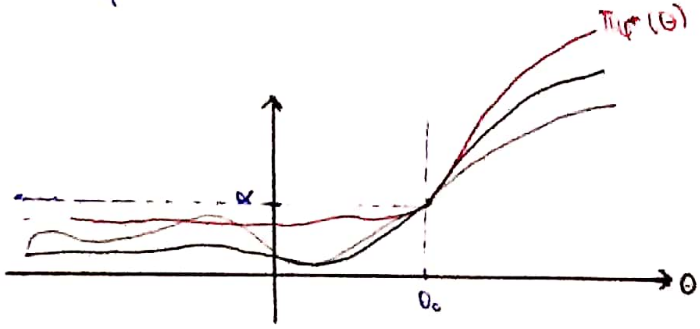
$H_0 : \theta \in \Theta_0$

$H_1 : \theta \in \Theta_1$

$\Psi = \{\varphi : \mathfrak{X} \to [0,1]\}$

Definição: O teste $\varphi^* \in \Psi$ de nível α, $0 < \alpha < 1$, é um teste uniformemente mais poderoso (UMP) de nível α para $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$ se $\forall \varphi \in \Psi$ tal que $\alpha_\varphi \leq \alpha$,

$$\Pi_{\varphi^*}(\theta) \geq \Pi_\varphi(\theta), \quad \forall \theta \in \Theta_1$$



$H_0 : \theta \leq \theta_0$

$H_1 : \theta > \theta_0$

$\theta_0 \in \mathbb{R}$

Em geral, é difícil (não existe) teste UMP!!

Vamos olhar um caso importante onde existe teste UMP!

$\mathfrak{X}$

$\Theta \subseteq \mathbb{R} \quad \mathcal{P} = \{f(\cdot|\theta) : \theta \in \Theta\}$

Seja $T : \mathfrak{X} \to \mathbb{R}$

Definição: A família $\mathcal{P} = \{f(\cdot|\theta) : \theta \in \Theta\}$ possui razão de verossimilhança monótona crescente (RVM) (decrescente) em T se

$\forall \theta_1, \theta_2 \in \Theta$, com $\theta_1 < \theta_2$,

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} \text{ é função crescente (decrescente) de } T(x).$$

nos possíveis valores (intervalo de valores) de T sob θ_1 ou θ_2.

Exemplo 1: $X_1, \dots, X_n$, dado θ, c.i.i.d. $N(\theta, 1)$

$\Theta = \mathbb{R}$

Sejam $\theta_1, \theta_2 \in \mathbb{R}$, com $\theta_1 < \theta_2$.

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum(x_i-\theta_2)^2}{2}}}{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum(x_i-\theta_1)^2}{2}}} = e^{\frac{\sum x_i^2 - 2\theta_1 \sum x_i + n\theta_1^2 - \sum x_i^2 + 2\theta_2 \sum x_i - n\theta_2^2}{2}} =$$

$$= e^{\frac{n(\theta_1^2 - \theta_2^2)}{2}} \cdot e^{\frac{2(\theta_2 - \theta_1)}{2} \sum_{i=1}^n x_i} \text{ é crescente em } T(x) = \sum_{i=1}^n x_i.$$

Então, dizemos que $\mathcal{P}$ possui RVM em $T(X) = \sum_{i=1}^n X_i$.

Exemplo 2: $X_1, \dots, X_n$, dado θ, c.i.i.d. $Exp(\theta)$

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

Sejam $\theta_1, \theta_2 > 0$, com $\theta_1 < \theta_2$

$x \in \mathfrak{X} = \mathbb{R}_+^n$

$$\frac{f(x/\theta_2)}{f(x/\theta_1)} = \frac{\theta_2^n e^{-\theta_2 \sum_{i=1}^n x_i}}{\theta_1^n e^{-\theta_1 \sum x_i}} = \left(\frac{\theta_2}{\theta_1}\right)^n e^{(\theta_1 - \theta_2)\sum_{i=1}^n x_i} \text{ é decrescente em } \sum_{i=1}^n x_i.$$

Logo, $\mathcal{P}$ possui RVM decrescente em $T(X) = \sum_{i=1}^n X_i$

(crescente em $-\sum_{i=1}^n X_i$. Logo, $\mathcal{P}$ possui RVM crescente em $-\sum_{i=1}^n X_i$)

Aula 16 - 14/06

Schervish

$\mathcal{P} = \{ f(\cdot|\theta) : \theta \in \Theta \}$ possui RVM crescente em T se $\forall \theta_1, \theta_2 \in \Theta \subseteq \mathbb{R}$,

com $\theta_1 < \theta_2$

$$\frac{f(x|\theta_2)}{f(x|\theta_1)}$$ é crescente em $T(x)$ nos valores de T sob θ_1 ou θ_2.

Exemplo 3. $X|\theta \sim \mathcal{U}(0,\theta)$

$\Theta = \mathbb{R}_+ \subseteq \mathbb{R}$ $\quad T(X) = X \geq 0$

$\theta_1, \theta_2 > 0$ com $\theta_1 < \theta_2$

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{\frac{1}{\theta_2} \mathbb{I}_{[0,\theta_2]}(x)}{\frac{1}{\theta_1} \mathbb{I}_{[0,\theta_1]}(x)} = \frac{\theta_1}{\theta_2} \frac{\mathbb{I}_{[0,\theta_2]}(x)}{\mathbb{I}_{[0,\theta_1]}(x)} = \begin{cases} \frac{\theta_1}{\theta_2}, & 0 \leq x \leq \theta_1, \\ \infty, & \theta_1 < x \leq \theta_2, \\ ?, & x > \theta_2 \end{cases}$$

Exemplo 4. $X|\theta \sim \text{Cauchy}(\theta)\ ((\theta,1))$

$$f(x|\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, \quad x \in \mathbb{R}$$

$\Theta = \mathbb{R}$

$\theta_1, \theta_2 \in \mathbb{R}$, $\theta_1 < \theta_2$

$$\frac{f(x|\theta_2)}{f(x_1)} = \frac{\frac{1}{\pi[1+(x-\theta_2)^2]}}{\frac{1}{\pi[1+(x-\theta_1)^2]}} = \frac{1+(x-\theta_1)^2}{1+(x-\theta_2)^2} = \mu(x)$$

$$\mu'(x) = \frac{2(x-\theta_1)[1+(x-\theta_2)^2] - [1+(x-\theta_1)^2]\,2(x-\theta_2)}{[1+(x-\theta_2)^2]^2} = ,$$

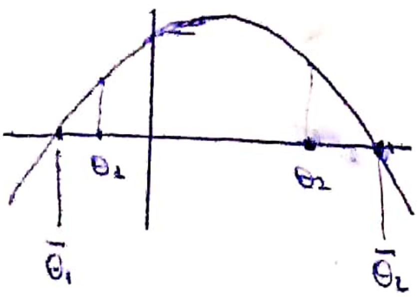
$$= \frac{2}{[1+(x-\theta_2)^2]^2}\left\{(x-\theta_1) + (x-\theta_1)(x-\theta_2)^2 - (x-\theta_2) - (x-\theta_2)(x-\theta_1)^2\right\} =$$

$$= A(x)\left\{(\theta_2-\theta_1) + (x-\theta_1)(x-\theta_2)\Big[\underbrace{(x-\theta_2)-(x-\theta_1)}_{-(\theta_2-\theta_1)}\Big]\right\} =$$

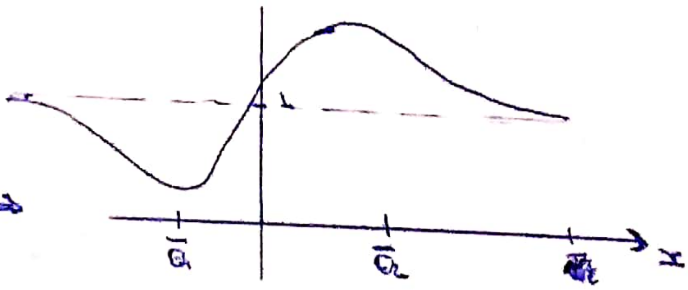
$$= \underbrace{A(x)}_{>0}(\theta_2-\theta_1)\{1-(x-\theta_1)(x-\theta_2)\}$$

μ é crescente se, e só se, $\mu'(x) > 0 \Leftrightarrow$

$1-(x-\theta_1)(x-\theta_2) > 0$



$\Rightarrow$



Teorema (Karlin-Rubin)

$\Theta \subseteq \mathbb{R},\ \mathcal{P},\ T: \mathcal{X} \to \mathbb{R},\ \psi$

Suponhamos que $\mathcal{P}$ possui RVM crescente (decrescente) em T. O teste $\psi^*: \mathcal{X} \to \{0,1\}$ dada por

$$\psi^*(x) = \begin{cases} 1, & T(x) > C \quad (T(x) < C) \\ 0, & \text{c.c.} \end{cases}$$

para as hipóteses $H_0: \theta \leq \theta_0$ versus $H_1: \theta > \theta_0$, $\theta_0 \in \Theta$, é UMP de nível $\pi_{\varphi^*}(\theta_0)$.

$\alpha^*_{\varphi^*}$

$\pi_{\varphi^*}: \Theta \to [0,1]$ é não-decrescente!

$\theta_1 < \theta_2 \Rightarrow$

$$\pi_{\varphi^*}(\theta) = \mathbb{P}(\varphi^*(X) = 1 / \theta) = P(T(X) > c / \theta).$$

$$T(X) > c \Leftrightarrow \frac{f(x/\theta_2)}{f(x/\theta_1)} > d$$

Vamos verificar que $\forall t \in \mathbb{R}$, $P(T(X) > t / \theta_1) \leq P(T(X) > t / \theta_2)$

$t' \geq 1$

$$P(T(X) > t / \theta_1) = \int_{\{x: T(x) > t\}} f(x/\theta_1) dx \leq \int_{\{x: T(x) > t\}} \frac{1}{t'} f(x/\theta_2) dx = \frac{1}{t'} \int_{\{x: T(x) > t\}} f(x/\theta_2) dx \leq$$

$$\leq \int_{\{x: T(x) > t\}} f(x/\theta_2) dx = P(T(x) > t / \theta_2)$$

$t' < 1$

$$P(T(X) \leq t / \theta_2) = \int_{\{x: T(x) \leq t\}} f(x/\theta_2) dx \leq \int_{\{x: T(x) \leq t\}} t' . f(x/\theta_1) dx = t' \int_{\{x \in \mathcal{X}: T(x) \leq t\}} f(x/\theta_1) dx = P(T(X) \leq t / \theta_1) \Rightarrow$$

$$\Rightarrow P(T(X) > t / \theta_1) \leq P(T(X) > t / \theta_2).$$

Assim, em particular para $t = c$, segue que $\pi_{\varphi^*}(\theta_1) \leq \pi_{\varphi^*}(\theta_2)$

Para $\theta \leq \theta_0$, segue que $\pi_{\varphi^*}(\theta) \leq \pi_{\varphi^*}(\theta_0)$. Logo, φ^* tem tamanho $\pi_{\varphi^*}(\theta_0)$, ou ainda,

$$\varphi^* \in \Psi^* = \{\varphi \in \Psi : \alpha^*_\varphi \leq \pi_\varphi(\Theta_0)\}$$

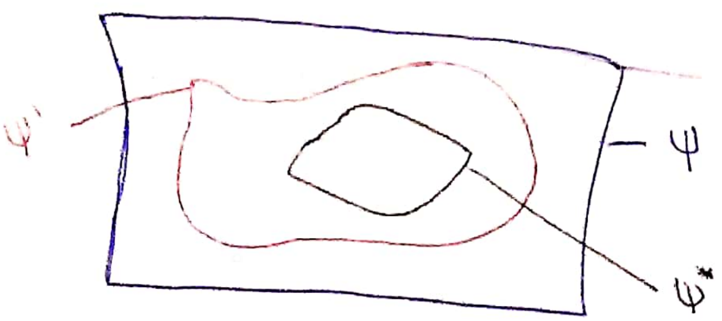
Vamos verificar que para todo $\overbrace{\theta > \theta_0}^{\theta \in \Theta_1}$, $\pi_{\varphi^*}(\theta) \geq \pi_\varphi(\theta)$, $\forall \varphi \in \Psi^*$.

Para $\theta_1 > \theta_0$,

$$\pi_{\varphi^*}(\theta_1) = \mathbb{P}(\varphi^*(X) = 1 \mid \theta_1) = \mathbb{P}(T(X) > c \mid \theta_1) = \mathbb{P}\left(\frac{f(x|\theta_1)}{f(x|\theta_2)} > d \mid \theta\right)$$

$$\bar{\varphi}(x) = \begin{cases} 1, & \frac{f(x|\theta_1)}{f(x|\theta_2)} > d \\ 0, & \text{c.c.} \end{cases}$$

Pelo Lema de Neyman-Pearson, $\pi_{\varphi^*}(\theta_1) \geq \pi_{\varphi'}(\theta_1)$, $\forall \varphi' \in \Psi' = \{\varphi \in \Psi \; \pi_\varphi(\theta) \leq \pi_{\varphi^*}(\theta_0)\}$



Logo, $\forall \varphi \in \Psi^*$, $\pi_{\varphi^*}(\theta_1) \geq \pi_\varphi(\theta_1)$, $\forall \theta_1 > \theta_0$

Logo, φ^* é UMP em Ψ^*. $(\forall \theta_1 \in \Theta_1)$.

Exemplo 1. $X_1, \ldots, X_n$ dado θ, c.i.i.d. $N(\theta, 1)$

$H_0: \theta \leq \theta_0$ $\qquad \theta_0 \in \mathbb{R}$ (fixado)

$H_1: \theta > \theta_0$

Vimos na última vez que $\mathcal{P}$ possui RVM crescente em $T(X) = \sum_{i=1}^n X_i$.

Para $\alpha \in (0,1)$

$$\varphi^*(x) = \begin{cases} 1, & \sum_{i=1}^{n} x_i > c = c(\alpha) \\ 0, & \text{c.c.} \end{cases}$$

é UMP de nível α para H_0 versus H_1.

Assim, $c \in \mathbb{R}$ é tal que

$\pi_{\varphi^*}(\theta_0) = \alpha \Rightarrow P(\varphi^*(X) = 1 \mid \theta_0) = \alpha \Rightarrow$

$\Rightarrow P\left(\sum_{i=1}^{n} X_i > c \mid \theta_0\right) = \alpha \Rightarrow P\left(\frac{\sum X_i - n\theta_0}{\sqrt{n}} > \frac{c - n\theta_0}{\sqrt{n}} \,\middle|\, \theta_0\right) = \alpha$

$\Rightarrow P\left(Z > \frac{c - n\theta_0}{\sqrt{n}}\right) = \alpha \Rightarrow \frac{c - n\theta_0}{\sqrt{n}} = \Phi^{-1}(1-\alpha) \Rightarrow c = n\theta_0 + \sqrt{n}\,\Phi^{-1}(1-\alpha)$

Daí,

$$\varphi^*(x) = \begin{cases} 1, & \sum_{i=1}^{n} x_i > n\theta_0 + \sqrt{n}\,\Phi^{-1}(1-\alpha) \\ 0, & \text{c.c.} \end{cases}$$

Exemplo 2. $X_1, \dots, X_n$, dado θ, c.i.i.d. $U(0, \theta)$.

$H_0: \theta \leq \theta_0$ $\quad\quad \theta_0 > 0$ (fixado)

$H_1: \theta > \theta_1$

Vimos na última vez que $\mathcal{P}$ possui RVM crescente em $T(x) = X_{(n)} = \max\{X_1, \dots, X_n\}$

Para $\alpha \in (0, 1)$

$$\varphi^*(x) = \begin{cases} 1, & x_{(n)} > c = c(\alpha) \\ 0, & \text{c.c.} \end{cases}$$

é UMP de nível α para H_0 vs. H_1.

Assim, $c>0$ é tal que

$$\pi_{\varphi^*}(\theta_0)=\alpha \Rightarrow P(\varphi^*(X)=1\,|\,\theta_0)=\alpha \Rightarrow$$

$$\Rightarrow P(X_{(n)}>c\,|\,\theta_0)=\alpha \Rightarrow P(X_{(n)}\leq c\,|\,\theta_0)=1-\alpha \Rightarrow$$

$c\in(0,\theta_0)$

$$\Rightarrow \left(\frac{c}{\theta_0}\right)^n=1-\alpha \Rightarrow c=\theta_0\cdot(1-\alpha)^{1/n}.$$

Exemplo 4. $X_1,\dots,X_n$, dado θ, c.i.i.d. $Exp(\theta)$

$H_0: \theta\leq\theta_0$
$H_1: \theta>\theta_0$ $\qquad \theta_0>0$ (fixado) $\Theta=\mathbb{R}_+$

Vimos na última vez que $\mathcal{P}$ possui RVM decrescente em $\sum_{i=1}^n X_i$ e, consequentemente, possui RVM crescente em $T(x)=-\sum_{i=1}^n X_i$.

Para $\alpha\in(0,1)$

$$\varphi^*(x)=\begin{cases}1, & \sum_{i=1}^n x_i<k\\ 0, & \text{c.c.}\end{cases}$$

é UMP de nível α para testar H_0 vs H_1.

$\sum X_i\,|\,\theta_0 \sim Gama(n,\theta_0)$

Logo, $k>0$ é tal que

$$\pi_{\varphi}(\theta_0)=\alpha \Rightarrow P(\varphi^*(x)=1\,|\,\theta_0)=\alpha \Rightarrow P\left(\sum_{i=1}^n X_i<k\,|\,\theta_0\right)=\alpha \Rightarrow$$

$$\Rightarrow P\left(2\theta_0\sum_{i=1}^n X_i\leq 2\theta_0 k\,|\,\theta_0\right)=\alpha \Rightarrow$$

$\Rightarrow P(U \leq 2\theta_0 k) = \alpha$ onde $U \sim \chi^2_{2n}$

$$\Rightarrow 2\theta_0 K = q_{2n,\alpha} \Rightarrow k = \frac{q_{2n,\alpha}}{2\theta_0}$$

Daí,

Para $\alpha \in (0,1)$,

$$\varphi^*(x) = \begin{cases} 1, & \sum_{i=1}^{n} x_i < q_{2n,\alpha}/2\theta_0 \\ 0, & \text{c.c.} \end{cases}$$

é UMP de nível α para testar H_0 versus H_1.

Teorema Karlin...

OBS,

$$\left.\begin{array}{l} H_0: \theta \geq \theta_0 \\ H_1: \theta < \theta_0 \end{array}\right\} \text{Tarefa!}$$

Resultado: $X_1, \ldots, X_n$ dado θ, c.i.i.d. tais que $f(x|\theta) = c(\theta) e^{g(\theta).T(x)} . h(x)$, $g(\cdot)$ estritamente crescente.

Considere as hipóteses

$$\begin{array}{l} H_0: \theta \leq \theta_1 \cup \theta \geq \theta_2 \\ H_1: \theta_1 < \theta < \theta_2 \end{array}, \quad \theta_1 < \theta_2$$

O teste $\varphi^*: \mathfrak{X} \to \{0,1\}$ dado por

$$\varphi^*(x) = \begin{cases} 1, & c_1 < \sum_{i=1}^{n} T(x_i) < c_2 \\ 0, & \text{c.c.} \end{cases}$$

é UMP de nível $\pi_{\varphi^*}(\theta_1) = \pi_{\varphi^*}(\theta_2)$ (1) (c_1 e c_2 são tomados de modo a satisfazer (1)).

Exemplo 1. $X|\theta \sim Exp(\theta)$

$H_0: \theta \leq 1$ ou $\theta \geq 2$ $\quad H_1: 1 < \theta < 2$.

$$f(x|\theta) = \theta e^{-\theta x} \mathbb{I}_{(0,\infty)}(x) \qquad c(\theta) = \theta \quad h(x) = \mathbb{I}_{\mathbb{R}_+}(x)$$

$$q(\theta) = \theta \quad T(x) = -x$$

$$\varphi^*(x) = \begin{cases} 1, & c_1 < -x < c_2 \\ 0, & \text{c.c.} \end{cases}$$

ou ainda,

$$\varphi^*(x) = \begin{cases} 1, & k_1 < x < k_2 \\ 0, & \text{c.c.} \end{cases}$$

Fixado α, tomamos $k_1, k_2 \in \mathbb{R}$ de modo que $\pi_{\varphi^*}(1) = \pi_{\varphi^*}(2) = \alpha$.

$$\pi_{\varphi^*}(\theta) = P(\varphi^*(X) = 1|\theta) = P(k_1 < X < k_2|\theta) = e^{-k_1\theta} - e^{-k_2\theta}$$

Assim, devemos ter,

$$\left.\begin{array}{l} \pi_{\varphi^*}(1) = e^{-k_1} - e^{-k_2} = \alpha \\ \pi_{\varphi^*}(2) = e^{-2k_1} - e^{-2k_2} = \alpha \end{array}\right\} \Leftrightarrow \begin{cases} e^{-k_1} - e^{-k_2} = \alpha \\ (e^{-k_1} - e^{-k_2})(e^{-k_1} + e^{-k_2}) = \alpha \end{cases}$$

$$e^{-k_2} = e^{-k_1} - \alpha = \frac{1+\alpha}{2} - \alpha = \frac{1-\alpha}{2} \Rightarrow k_2 = -\log\left(\frac{1-\alpha}{2}\right)$$

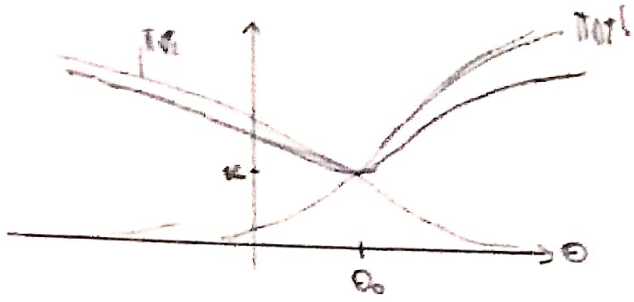
Assim,

$$\varphi^*(x) = \begin{cases} 1, & -\log\left(\frac{1+\alpha}{2}\right) < x < -\log\left(\frac{1-\alpha}{2}\right) \\ 0, & \text{c.c.} \end{cases}$$

Obs.

$H_0: \theta = \theta_0$

$H_1: \theta \neq \theta_0$



Rendimento não viesado

Voltando ao Exemplo 3.

$X/\theta \sim \exp(\theta)$

$H_0: \theta \leq 1$ $\quad n=1$

$H_1: \theta > 1$ $\quad \theta_0 = 1$

$$\varphi^*(x) = \begin{cases} 1, & x \leq -\log(1-\alpha) = 0{,}0513 \\ 0, & c.c. \end{cases} \Rightarrow \theta \leq 1$$

$X = 0{,}7$

$H_0': \begin{cases} \theta \leq 1 \cup \theta \geq 2 \\ 1 < \theta < 2 \end{cases}$ $\qquad \varphi' \begin{cases} 1, & 0{,}644 < x < 0{,}744 \\ 0, & c.c. \end{cases}$

$1 < \theta < 2$

Aula 17

18/06

$X_1, \dots, X_n$, dado Θ, c.i.i.d. $N(\Theta, 1)$

$H_0': \Theta \leq \Theta_0$ $\quad H_0'': \Theta \geq \Theta_0$ $\quad \varphi(x) = \begin{cases} 1, & T(x) > c \\ 0, & c.c. \end{cases}$ Teste unilaterais

$H_1': \Theta > \Theta_0$ $\quad H_1'': \Theta < \Theta_0$

$H_0: \Theta = \Theta_0 \quad (\Theta_1 \leq \Theta \leq \Theta_2)$

$H_1: \Theta \neq \Theta_0 \quad (\Theta < \Theta_1 \text{ ou } \Theta > \Theta_2)$ Testes bilaterais

Um teste intuitivo para $H_0: \Theta = \Theta_0 \times H_1: \Theta \neq \Theta_0$ de nível α, é dado por:

$$\varphi(x) = \begin{cases} 1, & |\bar{x} - \Theta_0| > c = c(\alpha) \\ 0, & c.c. \end{cases}$$

$\pi_\varphi(\Theta_0) = \alpha \Rightarrow P(\varphi(X) = 1 | \Theta_0) = \alpha \Rightarrow$

$\Rightarrow P(|\bar{X} - \Theta_0| > c | \Theta_0) = \alpha \Rightarrow$

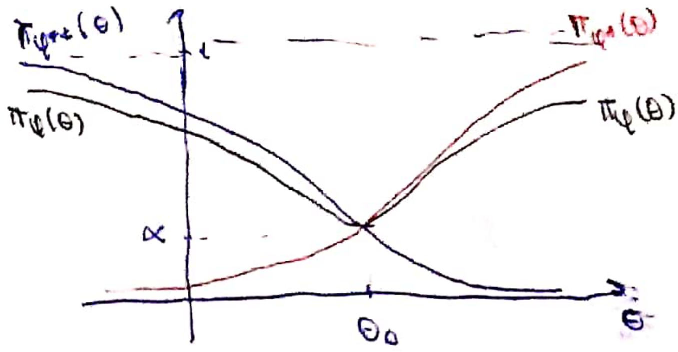
$$P\Big(\underbrace{\frac{|\bar{X} - \Theta_0|}{\sqrt{1/n}}}_{|Z|,\ Z \sim N(0,1)} > \frac{c}{1/\sqrt{n}} \Big| \Theta_0\Big) = \alpha \Rightarrow$$

$$\Rightarrow \frac{c}{\sqrt{\frac{1}{n}}} = \Phi(1 - \alpha/2)$$

$$\varphi(x) = \begin{cases} 1, & |\bar{x} - \Theta_0| > \dfrac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}} \\ 0, & c.c. \end{cases}$$

↳ ela não é UMP

$$\varphi^{*}(x) = \begin{cases} 1, & \bar{x} > \theta_0 + \dfrac{\Phi^{-1}(1-\alpha)}{\sqrt{n}} \\ 0, & \text{c.c.} \end{cases}$$



$$\varphi^{**}(x) = \begin{cases} 1, & \bar{x} < \theta_0 - \dfrac{\Phi^{-1}(1-\alpha)}{\sqrt{n}} \\ 0, & \text{c.c.} \end{cases}$$

Em geral, não conseguimos obter testes UMP como no caso mencionado.

Alternativa: Restringir comparação entre testes.

Def: Um teste $\varphi : \mathcal{X} \to \{0,1\}$ para $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$ é não-viesado de nível α, $0 < \alpha < 1$, se

1. $\pi_{\varphi}(\theta) \leq \alpha, \ \forall \theta \in \Theta_0$
2. $\pi_{\varphi}(\theta) \geq \alpha, \ \forall \theta \in \Theta_1$

Def: Um teste $\varphi^{*} : \mathcal{X} \to \{0,1\}$ é dito NÃO-VIESADO UNIFORMEMENTE MAIS PODEROSO DE NÍVEL α (UMPU), $0 < \alpha < 1$ se φ^{*} é UMP dentre os não-viesados de nível α.

Resultado: $X_1, \ldots, X_n$, dado θ, c.i.i.d. com densidade

$$f(x|\theta) = c(\theta)\, e^{\theta . T(x)}\, h(x) \qquad \theta \in \Theta \subseteq \mathbb{R}.$$

Então,

$$\varphi^{*}(x) = \begin{cases} 1, & \sum_{i=1}^{n} T(x_i) < c_1 \text{ ou } \sum_{i=1}^{n} T(x_i) > c_2 \\ 0, & \text{c.c.} \end{cases}$$

$(H_0: \theta = \theta_0)$ $(H_1: \theta \neq \theta_0)$

é UMPU para testar $H_0: \theta_1 \leq \theta \leq \theta_2$ versus $H_1: (\theta < \theta_1) \cup (\theta > \theta_2)$ dentre os teste nível $\pi_{\varphi^*}(\theta_1) = \pi_{\varphi^*}(\theta_2)$ $(\theta_1 \neq \theta_2)$

$(\pi_{\varphi^*}(\theta_0))$

Exemplo 1. $X_1, \dots, X_n$, dado θ, c.i.i.d., $N(\theta, 1)$

$$f(x_1|\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} = \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\theta^2/2}}_{c(\theta)} \underbrace{e^{-x^2/2}}_{h(x)} e^{\boxed{\theta x}} \quad , T(x) = x$$

Logo, o teste

$$\varphi^*(x) = \begin{cases} 1, & \sum_{i=1}^n x_i < c_1 \text{ ou } \sum_{i=1}^n x_i > c_2 \\ 0, & \text{c.c.} \end{cases}$$

é UMPU para $H_0: \theta = \theta_0 \times H_1: \theta \neq \theta_0$ de nível $\pi_{\varphi^*}(\theta_0)$

Assim, para $\alpha \in (0,1)$, tomamos $c_1, c_2 \in \mathbb{R}$ tais que

$$\pi_{\varphi^*}(\theta_0) = \alpha \Rightarrow P\left(\sum_{i=1}^n X_i < c_1 \cup \sum_{i=1}^n X_i > c_2 \,\Big|\, \theta_0\right) = \alpha \Rightarrow$$

$$\Rightarrow \varphi^*(x) = \begin{cases} 1, & |\bar{x} - \theta_0| > \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}} \\ 0, & \text{c.c.} \end{cases}$$

Agora, sejam $\theta_1, \theta_2 \in \mathbb{R}$ com $\theta_1 < \theta_2$.

$H_0: \theta_1 \leq \theta \leq \theta_2$

$H_1: \theta < \theta_1 \cup \theta > \theta_2$,

$$\varphi^*(x) = \begin{cases} 1, & \sum X_i < c_1 \text{ ou } \sum X_i > c_2 \\ 0, & \text{c.c.} \end{cases}$$

$$\pi_{\varphi^*}(\theta_1) = \pi_{\varphi^*}(\theta_2) = \alpha$$

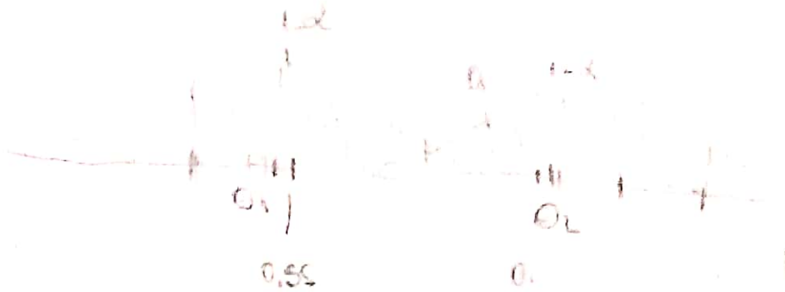
$$\pi_{\varphi^*}(\theta_1) = P\left(\sum X_i < c_1 \cup \sum X_i > c_2 \mid \theta_1\right) = P(\bar{X} < k_1 \cup \bar{X} > k_2 \mid \theta_1) =$$

$$= P\left(\frac{\bar{X} - \theta_1}{\sqrt{1/n}} < \frac{k_1 - \theta_1}{\sqrt{1/n}} \middle| \theta_1\right) + P\left(\frac{\bar{X} - \theta_1}{\sqrt{1/n}} > \frac{k_2 - \theta_1}{\sqrt{1/n}} \middle| \theta_1\right) = \alpha \Rightarrow$$

$$\Rightarrow \Phi\left(\frac{k_1 - \theta_1}{1/\sqrt{n}}\right) + 1 - \Phi\left(\frac{k_2 - \theta_1}{1/\sqrt{n}}\right) = \alpha \Rightarrow \Phi\left(\frac{k_2 - \theta_1}{1/\sqrt{n}}\right) - \Phi\left(\frac{k_1 - \theta_1}{1/\sqrt{n}}\right) = 1 - \alpha \quad (I)$$

$$\pi_{\varphi^*}(\theta_2) = \alpha \Rightarrow$$

$$\Rightarrow \Phi\left(\frac{k_2 - \theta_2}{1/\sqrt{n}}\right) - \Phi\left(\frac{k_1 - \theta_2}{1/\sqrt{n}}\right) = 1 - \alpha, \quad (II)$$



Resultado (I) e (II) (em k_1 e k_2):

$$P(N(\theta_1, 1/n) \in (k_1, k_2)) = 1 - \alpha$$

$$P(N(\theta_2, 1/n) \in (k_1, k_2)) = 1 - \alpha$$

Testes Da Razão de Verossimilhança Generalizada (TRVG)

$H_0: \theta = \theta_0$

$H_1: \theta = \theta_1$

$$\varphi^*(x) = \begin{cases} 1, & \frac{f(x/\theta_1)}{f(x/\theta_0)} > c \\ 0, & \end{cases} \quad \left(\frac{f(x/\theta_0)}{f(x/\theta_1)} < c' \right)$$

Suponha agora

$H_0: \theta \in \Theta_0$

$H_1: \theta \in \Theta_1$

$\xrightarrow{\text{teste intuitivo}}$

$$\varphi^*(x) = \begin{cases} 1, & \frac{\int_{\Theta_0} f(x/\theta)\, d\theta}{\int_{\Theta_1} f(x/\theta)\, d\theta} < c \\ 0, & \text{c.c.} \end{cases}$$

"testes verossimilhancistas"

ou

$$\varphi^*(x) = \begin{cases} 1, & \frac{\sup_{\theta \in \Theta_0} f(x/\theta)}{\sup_{\theta \in \Theta_1} f(x/\theta)} < c \\ 0, & \text{c.c.} \end{cases}$$

$\Downarrow$

$\in [0,1]$

$$\text{TRVG} \Rightarrow \varphi^*(x) = \begin{cases} 1, & \frac{\sup_{\theta \in \Theta_0} f(x/\theta)}{\sup_{\theta \in \Theta} f(x/\theta)} < c \\ 0, & \text{c.c.} \end{cases}$$

Vtg RV $\in [0,1]$

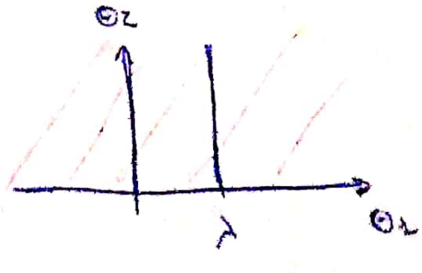
$-2\log \square \sim \chi^2_{n-1}$

Exemplo: $X_1, \dots, X_n$, dado $\theta = (\theta_1, \theta_2)$, c.i.i.d. $N(\theta_1, \theta_2)$

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

$H_0: \theta_1 = \lambda$

$H_1: \theta_1 \neq \lambda$, $\lambda \in \mathbb{R}$ fixado (conhecido)



$$RV(x) = \frac{\sup_{\theta \in \Theta_0} f(x/\theta)}{\sup_{\theta \in \Theta} f(x/\theta)}$$

Em Θ, $\sup_{\theta \in \Theta} f(x/\theta) = f\left(x \,|\, \left(\bar{x}, \sum_{i=1}^{n} (x_i - \bar{x})^2 / n\right)\right)$

Em Θ_0,

$$f(x/\theta) = \left(\frac{1}{\sqrt{2\pi\,\theta_2}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\lambda)^2}{2\theta_2}}$$, portanto,

$$\sup_{\theta \in \Theta_0} f(x/\theta) = f\left(x \,\middle|\, \left(\lambda, \frac{\sum_{i=1}^{n}(x_i-\lambda)^2}{n}\right)\right)$$

$$RV(x) = \frac{\sup_{\theta \in \Theta_0} f(x/\theta)}{\sup_{\theta \in \Theta} f(x/\theta)} = \frac{\left(\frac{1}{\sqrt{2\pi \frac{\sum(x_i-\lambda)^2}{n}}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\lambda)^2}{2\frac{\sum_{i=1}^{n}(x_i-\lambda)^2}{n}}}}{\left(\frac{1}{\sqrt{2\pi \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{2\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}}} =$$

$$= \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\lambda)^2}\right)^{n/2} < c \iff \frac{\sum(x_i-\lambda)^2}{\sum(x_i-\bar{x})^2} > \underbrace{\left(\frac{1}{c}\right)^{2/n}}_{c'} \iff$$

$$\frac{\sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\lambda)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} > c' \iff \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\lambda)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} > c' \iff$$

$$1 + \frac{n(\bar{x}-\lambda)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} > c' \iff \frac{n(\bar{x}-\lambda)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} > \underbrace{c'-1}_{c''} \iff \frac{n(\bar{x}-\lambda)^2}{\frac{\sum(x_i-\bar{x})^2}{n-1}} > c''' \quad {}^{= c''(n-1)}$$

$$\Rightarrow \left| \frac{(\bar{x} - \lambda)\sqrt{n}}{\sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}} \right| > k$$

S^2

Sob H_0, $\dfrac{\sqrt{n}(\bar{X} - \lambda)}{\sqrt{S^2/n-1}} \sim t_{n-1}$.

Logo o teste RVG para $H_0: \theta_1 = \lambda$ versus $H_1: \theta \neq \lambda$ é dado por

$$\varphi^*(x) = \begin{cases} 1, & \left| \dfrac{\sqrt{n}(\bar{x} - \lambda)}{\sqrt{S^2/n-1}} \right| > k^* = k^*(\alpha) \\ 0, & \text{c.c.} \end{cases}$$

se $\theta_2 = 1$,

$|\bar{x} - \lambda| > c$

Exemplo 2:

$X_{11}, \dots, X_{1n_1}, X_{21}, X_{22}, \dots, X_{2n_2}, \dots, X_{k1}, X_{k2}, \dots, X_{kn_k}$ que, dado $\theta = (\theta_1, \theta_2, \dots, \theta_k, \theta_{k+1})$, são cond. independentes tais que $X_{ij} | \theta \sim N(\theta_i, \theta_{k+1})$, $i = 1, \dots, k$
$j = 1, \dots, n_i$

n_1 - trat. 1
n_2 - trat. 2
$\vdots$
n_k - trat. 3

Imaginamos que as variâncias são iguais entre os grupos

As médias diferem?

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

$H_0: \theta_1 = \theta_2 = \ldots = \theta_k$

$H_1:$ caso contrário

$\mathfrak{X} = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \ldots \times \mathbb{R}^{n_k}$

$\mathbb{R}^{n_1 + n_2 + \ldots + n_k}$

$x \in \mathfrak{X}$

$$f(x|\theta) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\theta_{k+1}}} e^{-\frac{(x_{ij}-\theta_i)^2}{2\theta_{k+1}}} = \left(\frac{1}{2\pi\theta_{k+1}}\right)^{\sum_{i=1}^{k} n_i} e^{-\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\theta_i)^2}{2\theta_{k+1}}}$$

Em Θ_0,

$$\sup_{\theta \in \Theta_0} f(x|\theta) = f\left(x \,\Big|\, \left(\bar{x}, \bar{x}, \ldots, \bar{x}, \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x})^2}{n}\right)\right), \text{ onde } \bar{x} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}}{n_1 + \ldots + n_k}, \; n = n_1 + \ldots + n_k$$

Em Θ,

$$\sup_{\theta \in \Theta} f(x|\theta) = f\left(x \,\Big|\, \left(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k, \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_i-\bar{x}_i)^2}{n}\right)\right), \text{ onde } \bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}.$$

$$RV(x) = \frac{\sup_{\theta\in\Theta_0} f(x|\theta)}{\sup_{\theta\in\Theta} f(x|\theta)} = \frac{\left(\frac{1}{2\pi \frac{\sum_i\sum_j (x_{ij}-\bar{x})}{n}}\right)^{n/2} \cdot e^{-\frac{\sum_i\sum_j (x_{ij}-\bar{x})^2}{2\frac{\sum_i\sum_j (x_{ij}-\bar{x})^2}{n}}}}{\left(\frac{1}{2\pi \frac{\sum_i\sum_j (x_{ij}-\bar{x}_i)^2}{n}}\right)^{n/2} e^{-\frac{\sum_i\sum_j (x_{ij}-\bar{x}_i)^2}{2\frac{\sum_i\sum_j (x_{ij}-\bar{x}_i)^2}{n}}}} =$$

$$= \left(\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x})^2}\right)^{n/2} < c \Leftrightarrow \left(\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i+\bar{x}_i-\bar{x})^2}\right) < c^{2/n} \Leftrightarrow$$

$$\Leftrightarrow \frac{\sum\sum (x_{ij}-\bar{x}_{i.})^2 + \sum_i \sum_j (x_{ij}-\bar{x})^2}{\sum\sum (x_{ij}-\bar{x}_i)^2} \geq c^{-2/n}$$

$$\Leftrightarrow \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} (\bar{x}_{i.}-\bar{x})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i} (x_{ij}-\bar{x}_i)^2} > k \Leftrightarrow \frac{\overset{\text{sob } H_0,\ \sim \chi^2_{k-1}}{SQTratamento}}{\underset{\sim \chi^2_{\sum n_i - k}}{SQResíduos}} > k \Leftrightarrow \frac{\frac{SQTrat}{k-1}}{\frac{SQResíduo}{n-k}} > k^*$$

SQReg	QMReg	$\frac{QMReg}{QMRes}$	$F_{k-1,\,n-1}$
SQRes	QMRes		
SQT	Q		

Sob Ho,

$$\frac{\frac{SQTratamentos}{k-1}}{\frac{SQResíduos}{n-k}} \sim F_{k-1,\,n-k}$$

Aula 18 20/06/2013

Avaliação 2
28/06
14:00/18:00
SALA B-2

Avaliação 3
02/07
14:00/18:00
SALA B-5

Teste RVG

$H_0: \theta \in \Theta_0$

$H_1: \theta \in \Theta_1$

$RV: \mathcal{X} \to \{0,1\}$

$$x \in \mathcal{X} \to RV(x) = \frac{\sup_{\theta \in \Theta_0} f(x/\theta)}{\sup_{\theta \in \Theta} f(x/\theta)} \leq c$$

$$\varphi(x) = \begin{cases} 1, & RV(x) \leq c. \\ 0, & c.c \end{cases}$$

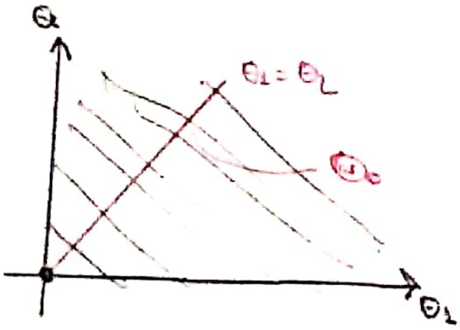
Exemplo 3. $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$, dado $\theta = (\theta_1, \theta_2)$, c.i. tais que

$X_i/\theta \sim Exp(\theta_1)$, $i = 1, \dots, n_1$ e

$Y_j/\theta \sim Exp(\theta_2)$, $j = 1, \dots, n_2$

$H_0: \theta_1 = \theta_2$ $\Theta = \mathbb{R}_+^2$

$H_1: \theta_1 \neq \theta_2$ $\mathcal{X} = \mathbb{R}_+^{n_1+n_2}$ $(\mathbb{R}_+^{n_1} \times \mathbb{R}_+^{n_2})$.



$x = (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \in \mathcal{X}$

$$f(x/\theta) = \prod_{i=1}^{n_1} \theta_1 e^{-\theta_1 x_i} \mathbb{I}_{\mathbb{R}_+}(x_i) \cdot \prod_{j=1}^{n_2} \theta_2 e^{-\theta_2 y_j} \mathbb{I}_{\mathbb{R}_+}(y_j)$$

Em Θ,

$$\sup_{\theta \in \Theta_0} f(x/\theta) = f\left(x \,\middle|\, \left(\frac{n_1}{\sum_{i=1} x_i}, \frac{n_2}{\sum_{j=1} y_j}\right)\right)$$

Em Θ_0: $\theta_1 = \theta_2 = \theta$

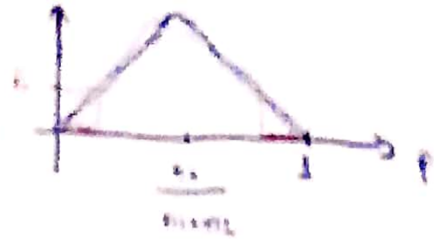
$$\sup_{\theta \in \Theta_0} f(x/\theta) = f\left(x \,\middle|\, \left(\frac{n_1+n_2}{\sum x_i + \sum y_j}, \frac{n_1+n_2}{\sum x_i + \sum y_j}\right)\right)$$

$$RV(x) = \frac{\sup_{\theta \in \Theta_0} f(x/\theta)}{\sup_{\theta \in \Theta} f(x/\theta)} = \frac{\left(\frac{n_1+n_2}{\sum x_i + \sum y_j}\right)^{n_1+n_2} e^{-\frac{n_1+n_2}{\sum x_i + \sum y_j}(\sum x_i + \sum y_j)}}{\left(\frac{n_1}{\sum x_i}\right)^{n_1} e^{-\frac{n_1}{\sum x_i}\sum x_i} \left(\frac{n_2}{\sum y_j}\right)^{n_2} e^{-\frac{n_2}{\sum y_j}\sum y_j}}$$

$$= \frac{(n_1+n_2)^{n_1+n_2}}{n_1^{n_1}\, n_2^{n_2}} \frac{\left(\sum_{i=1}^{n_1} x_i\right)^{n_1} \left(\sum_{j=1}^{n_2} y_j\right)^{n_2}}{\left(\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j\right)^{n_1+n_2}} < c \iff$$

$$\frac{\left(\sum_{i=1}^{n_1} x_i\right)^{n_1} \left(\sum_{j=1}^{n_2} y_j\right)^{n_2}}{\left(\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j\right)^{n_1+n_2}} < c/c' \iff$$

$$\left(\frac{\sum_{i=1}^{n_1} x_i}{\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j}\right)^{n_1} \left(\frac{\sum_{j=1}^{n_2} y_j}{\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j}\right)^{n_2} < k \iff$$

$p^{n_1}(1-p)^{n_2} < k \longrightarrow$ 

$$\Rightarrow \frac{\sum_{i=1}^{n_1} x_i}{\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j} < \delta_1 \text{ ou } \frac{\sum_{i=1}^{n_1} x_i}{\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j} > \delta_2 \iff$$

$$\frac{\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j}{\sum_{i=1}^{n_1} x_i} > \delta_1 \text{ ou } \frac{\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j}{\sum_{j=1}^{n_2} y_j} < \delta_2 \iff$$

$$\frac{\sum_{j=1}^{n_2} y_j}{\sum_{i=1}^{n_1} x_i} > \frac{1-\delta_1}{\delta_1} \text{ ou } \frac{\sum_{j=1}^{n_2} y_j}{\sum_{i=1}^{n_1} x_i} < \frac{1-\delta_2}{\delta_2}$$

$$\psi(x) = \begin{cases} 1, & \dfrac{\Sigma x_i}{\Sigma y_j} > \dfrac{1-\delta_1}{\delta_1} \text{ ou } \dfrac{\Sigma x_i}{\Sigma y_j} < \dfrac{1-\delta_2}{\delta_2} \\ 0, & \text{c.c.} \end{cases}$$

Impomos que, sob H_0, $\pi_\psi(\theta) \leq \alpha$.

Para todo $(\theta_1, \theta_2) \in \Theta_0$, temos que

$$\frac{\sum_{i=1}^{n_1} X_i}{\sum_{j=1}^{n_2} Y_j} = \frac{2\theta_1 \sum_{i=1}^{n_1} X_i}{2\theta_1 \sum_{j=1}^{n_2} Y_j} \text{ para } \theta \in \Theta_0,$$

$$\left. \frac{\frac{\sum_{i=1}^{n_1} X_i}{2n_1}}{\frac{\sum_{j=1}^{n_2} Y_j}{2n_2}} \right| \theta \sim F_{2n_1, 2n_2}$$

Sejam $f_{\alpha/2}^{2n_1, 2n_2}$ e $f_{1-\alpha/2}^{2n_1, 2n_2}$ percentis de ordem $\frac{\alpha}{2}$ e $1-\frac{\alpha}{2}$ do $F_{2n_1, 2n_2}$;

$$P\left(f_{\alpha/2} \leq \frac{\sum_{i=1}^{n_1} X_i}{\sum_{j=1}^{n_2} Y_j} \frac{n_2}{n_1} \leq f_{1-\alpha/2} \,\Big|\, \theta \right) = 1 - \alpha, \ \forall \theta \in \Theta_0.$$

Assim,

$$\psi(x) = \begin{cases} 1, & \dfrac{\sum^{n_1} x_i}{\sum^{n_2} y_j} > \dfrac{n_1}{n_2} f_{1-\alpha/2} \text{ ou } \dfrac{\sum^{n_1} x_i}{\sum^{n_2} y_j} < \dfrac{n_1}{n_2} f_{\alpha/2} \\ 0, & \text{c.c.} \end{cases}$$

$\theta \in \Theta_0$

$\lambda(x) = -2\log \lambda(x) \,|\, \theta \overset{a}{\sim} \chi^2_{\mu(\Theta_0,\Theta)}$

$2-1$

$D(\Theta_0)=1, \quad D(\Theta)=2$

Victor Fossaluza (2008).

Testes Bayesianos

$H_0: \theta \in \Theta_0$

$H_1: \theta \in \Theta_1$

decisão	$\theta = \Theta_0$	$\theta = \Theta_1$
0 (n rejeitar)	0	a_2
1 (rejeitar)	a_1	0

$\Psi = \{\varphi : \mathcal{X} \to \{0,1\}\}.$

Vimos que o teste de Bayes é dado por:

$$\varphi^*(x) = \begin{cases} 1, & P(\theta = \Theta_1 \mid X = x) > a_1/(a_1+a_2) \\ 0, & \end{cases}$$

Agora, vamos supor que tenhamos hipóteses compostas!

decisão	$\theta \in \Theta_0$	$\theta \in \Theta_1$
0 (ñ rejeitar)	0	a_2
1 (rejeitar)	a_1	0

"Se constrói da mesma forma, mas agora avaliamos a probab. à posteriori em toda a região crítica".

No caso geral,

$H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_1$,

Considerando a probabilidade do tipo $\overset{(0-1-c)}{0-a_1-a_2}$,

$$\varphi^*(x) = \begin{cases} 1, & P(\theta \in \Theta_1 / X = x) > a_1/(a_1+a_2) \\ 0, & \text{c.c.} \end{cases}$$

Problema sob modelos abs. contínua com Θ_0 um cjto com medida de Lebesgue nula, então $P(\theta \in \Theta_1 / X = x) = 1 \; \forall x \in \theta$,

Solução: Considerar outras funções de perda!

- Carlinhos & Júlio Stern FBST
- Inf. Jeffrey
- Atribuir densidade a Θ_0 e outra à Θ_1 e atualizar pelo teorema de Bayes.

$p^2 \quad 2p(1-p) \quad (1-p)^2 \rightarrow$ equilíbrio de Hardy

Exemplo 1. $X/\theta \sim U(0,\theta)$

$\theta \sim U(0,30)$

$H_0: \theta \leq 20$

$H_1: \theta > 20$

$x > 0$

$$f(\theta/x) = \frac{f(x/\theta) f(\theta)}{f(x)} \propto f(x/\theta) f(\theta) =$$

$$= \frac{1}{\theta} \mathbb{I}_{(0,\theta)}(x) \frac{1}{30} \mathbb{I}_{(0,30)}(\theta) \propto \frac{1}{\theta} \mathbb{I}_{(0,\theta)}(x) \mathbb{I}_{(0,30)}(\theta) \cdot \frac{1}{\theta} \mathbb{I}_{(x,\infty)}(\theta) \mathbb{I}_{(0,30)}(\theta) =$$

$$= \frac{1}{\theta} \mathbb{I}_{(x,30)}(\theta) \Rightarrow f(\theta/x) = \frac{c}{\theta} \mathbb{I}_{(x,30)}(\theta)$$

$$\int_{-\infty}^{\infty} f(\theta/x)\, d\theta = 1 \rightarrow c \int_x^{30} \frac{1}{\theta}\, d\theta = 1 \Rightarrow c \cdot \log(\theta)\Big|_x^{30} = 1 \Rightarrow$$

$$c \log\left(\frac{30}{x}\right) = 1 \Rightarrow c = \frac{1}{\log\left(\frac{30}{x}\right)}$$

logo,

$$f(\theta/x) = \frac{1}{\log\left(\frac{30}{x}\right)} \frac{1}{\theta} \cdot \mathbb{I}_{(x,30)}(\theta)$$

$$P(\theta \in \Theta_1 \mid X=x) = P(\theta > 20 \mid X=x) = \int_{20}^{\infty} f(\theta \mid x)\, d\theta =$$

$$= \int_{20}^{\infty} \frac{1}{\log\left(\frac{30}{x}\right)} \frac{1}{\theta} \mathbb{I}_{(x,30)}(\theta)\, d\theta = \begin{cases} \displaystyle\int_{x}^{30} \frac{1}{\log\frac{30}{x}} \frac{1}{\theta}\, d\theta = 1 & , x > 20 \\ \displaystyle\int_{20}^{30} \frac{1}{\log\frac{30}{x}} \frac{1}{\theta}\, d\theta = \frac{\log\left(\frac{30}{20}\right)}{\log\left(\frac{30}{x}\right)} & , x < 20 \end{cases}$$

Fixados $a_1, a_2 > 0$.

Para $x < 20$,

$$P(\theta > 20 \mid X=x) > \frac{a_1}{a_1+a_2} \iff$$

$$\frac{\log 30 - \log 20}{\log 30 - \log x} > \frac{a_1}{a_1+a_2} \iff a_2 \log 30 - (a_1+a_2)\log 20 > -a_1 \log x \iff$$

$$\log x > \frac{a_2 \log 30 - (a_1+a_2)\log 20}{-a_1} \iff$$

$$x > e^{\frac{(a_1+a_2)\log 20 - a_2 \log 30}{a_1}}$$

Assim, o teste de Bayes é dado por

$$\varphi^*(x) = \begin{cases} 1 & , x > \min\left\{20, e^{\frac{(a_1+a_2)\log 20 - a_2\log 30}{a_1}}\right\} \\ 0 & , \text{c.c.} \end{cases}$$

Exemplo 2. $X_1, \dots, X_n$ dado θ, c.i.i.d. $Exp(\theta)$

$\theta \sim Gama(a, b)$

$H_0: \theta \leq 3$ $\qquad \mathcal{X} = \mathbb{R}_+^n \quad x = (x_1, \dots, x_n) \in \mathcal{X}$

$H_1: \theta > 3$

$$f(\theta / x_1, \dots, x_n) \propto f(x_1, \dots, x_n / \theta)\, f(\theta) =$$

$$= \prod_{i=1}^{n} f(x_i / \theta)\, f(\theta) \propto \theta^n e^{-\theta \sum x_i}\, \theta^{a-1} e^{-b\theta}\, \mathbb{I}_{\mathbb{R}_+}(\theta)$$

$$\Rightarrow f(\theta / x) \propto \theta^{a+n-1} e^{-(b+\sum x_i)\theta}\, \mathbb{I}_{\mathbb{R}_+}(\theta)$$

$\theta / X_1 = x_1, \dots, X_n = x_n \sim Gama(a+n, b + \sum x_i)$

Logo,

$$\varphi^*(x) = \begin{cases} 1, & \int_3^{\infty} \frac{(b+\sum x_i)^{a+n}}{\Gamma(a+n)} \theta^{a+n-1} e^{-(b+\sum x_i)\theta} d\theta > \frac{a_1}{a_1 + a_2} \\ 0, & c.c \end{cases}$$

Testar k-hipóteses

Ao invés de $\Theta = \Theta_0 \cup \Theta_1$, $\Theta_0 \cap \Theta_1 = \emptyset$, tomamos

$$\Theta = \Theta_0 \cup \Theta_1 \cup \dots \cup \Theta_k, \quad e \quad \Theta_i \cap \Theta_j = \emptyset,\ i \neq j$$

Ex $H_1: \theta \leq 1/4$

$H_2: 1/4 < \theta \leq 1/2$

$H_3: 1/2 < \theta \leq 9/10$

$H_4: \theta > 9/10$

Agora

$\Psi = \{ \varphi : \mathfrak{X} \to \{1, \dots, k\} \}$

Para $\varphi \in \Psi$, $\varphi(x) = j$ denota por decisão $H_j : \theta \in \Theta_j$

Decisão	$\theta \in \Theta_1$	$\theta \in \Theta_2$	$\dots$	$\theta \in \Theta_k$
1	0	l_{12}	$\cdot$	l_{1k}
2	l_{21}	0		l_{2k}
3	l_{31}	l_{32}	$\cdot$	l_{3k}
$\vdots$	$\vdots$	$\vdots$	$\cdot$	$\cdot$
k	l_{k1}	l_{k2}	$\dots$	0

Para $\varphi \in \Psi$, vamos avaliar

$$\rho(\varphi) = E[L(\varphi(x), \theta)] = \sum_{x \in \mathfrak{X}} \sum_{\theta' \in \Theta} L(\varphi(x), \theta') \mathbb{P}(\theta = \theta', X = x) =$$

$$= \sum_{x \in \mathfrak{X}} \sum_{\theta' \in \Theta} \left\{ \sum_{i=1}^{k} \sum_{j=1}^{k} l_{ij} \, \mathbb{I}_{\Theta_j}(\theta') \, \mathbb{I}_{\varphi^{-1}(\{i\})}(x) \right\} \mathbb{P}(\theta = \theta' / X = x) \, \mathbb{P}(X = x) =$$

$$= \sum_{x \in \mathfrak{X}} \mathbb{P}(X = x) \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{\theta' \in \Theta} l_{ij} \, \mathbb{I}_{\Theta_j}(\theta') \, \mathbb{I}_{\varphi^{-1}(\{i\})}(x) \, \mathbb{P}(\theta = \theta' / X = x) =$$

$$= \sum_{x \in \mathfrak{X}} \mathbb{P}(X = x) \left\{ \sum_{i=1}^{k} \mathbb{I}_{\varphi^{-1}(\{i\})}(x) \left[\sum_{j=1}^{k} l_{ij} \sum_{\theta' \in \Theta} \mathbb{I}_{\Theta_j}(\theta') \, \mathbb{P}(\theta = \theta' / X = x) \right] \right\} \Rightarrow$$

$$\Rightarrow \rho(\varphi) = \sum_{x \in \mathfrak{X}} P(X=x) \left\{ \sum_{i=1}^{k} \mathbb{I}_{\varphi^{-1}(\{i\})}(x) \left\{ \sum_{j=1}^{k} l_{ij} P(\Theta \in \Theta_j / X=x) \right\} \right\}$$

$\downarrow$

Assim, para minimizar ρ em Ψ, devo tomar $\varphi^*: \mathfrak{X} \to \{1, \dots, k\}$ que a cada $x \in \mathfrak{X}$ associa $i^* \in \{1, \dots, k\}$ tal que

$$\sum_{j=1}^{k} l_{i^*j} P(\Theta \in \Theta_j / X=x) = \min_{i \in \{1,\dots,k\}} \sum_{j=1}^{k} l_{ij} P(\Theta \in \Theta_j / X=x)$$

Exemplo 1:

$\Theta = \{1/4, 1/2, 3/4\}$ $\qquad X/\theta \sim Geo(\theta)$

$P(\theta = \frac{1}{4}) = P(\theta = \frac{2}{4}) = \frac{1}{4}$ e $P(\theta = \frac{3}{4}) = \frac{2}{4}$.

$H_1: \theta = 1/4$

$H_2: \theta = 2/4$

$H_3: \theta = 3/4$

$\Psi = \{\varphi: \mathfrak{X} \to \{1, 2, 3\}\}$ $\qquad \mathfrak{X} = \mathbb{N}^*$

Decisão	$\theta = 1/4$	$\theta = 1/2$	$\theta = 3/4$
1	0	2	2
2	3/2	0	3/2
3	2	2	0

Para cada $i \in \{1,2,3\}$

$$\sum_{j=1}^{3} l_{ij}\, \mathbb{P}\left(\Theta = \frac{j}{4} \,/\, X = x\right)$$

$$i = 1 \rightarrow \overset{0}{l_{11}}\, \mathbb{P}\left(\Theta = \frac{1}{4} \,/\, X = x\right) + \overset{2}{l_{12}}\, \mathbb{P}\left(\Theta = \frac{2}{4} \,\middle|\, X = x\right) + \overset{2}{l_{13}}\, \mathbb{P}\left(\Theta = \frac{3}{4} \,\middle|\, X = x\right)$$

$$i = 2 \rightarrow \overset{3/2}{l_{21}}\, \mathbb{P}\left(\Theta = \frac{1}{4} \,/\, X = x\right) + \overset{0}{l_{22}}\, \mathbb{P}\left(\Theta = \frac{2}{4} \,/\, X = x\right) + \overset{1/2}{l_{23}}\, \mathbb{P}\left(\Theta = \frac{3}{4} \,/\, X = x\right)$$

$$i = 3 \rightarrow \overset{2}{l_{31}}\, \mathbb{P}\left(\Theta = \frac{1}{4} \,/\, X = x\right) + \overset{2}{l_{32}}\, \mathbb{P}\left(\Theta = \frac{2}{4} \,/\, X = x\right) + \overset{0}{l_{33}}\, \mathbb{P}\left(\Theta = \frac{3}{4} \,\middle|\, X = x\right)$$

MAE 5702 - Probabilidade e Inferência Estatística I 1º semestre/2013

Prof. Luís Gustavo Esteves (sala B - 108)

Aulas: 3ª feira – 16:00, 5ª feira – 16:00, 6ª feira – 14:00 Sala B – 02

Atendimento: 2ª feira e 4ª feira – 18:00 às 19:00 (Julio) Sala B – 169

3ª feira e 5ª feira – 18:00 às 19:00 (Luís) Sala B – 108

PROGRAMA

Módulo I : Probabilidade

1. **Probabilidade: definição, propriedades elementares, probabilidade condicional, Teorema de Bayes e Independência.**
2. **Vetores aleatórios: definição, caracterização, transformações de vetores aleatórios.**
3. **Esperança: definição, propriedades, momentos e esperança condicional.**
4. Convergência de variáveis aleatórias: tipos de convergência, a Lei Forte dos Grandes Números e o Teorema do Limite Central.

Módulo II : Inferência Estatística

1. **Conceitos básicos: espaço amostral, espaço paramétrico, modelo estatístico e suficiência.**
2. **Elementos de Inferência Bayesiana: distribuições a priori e a posteriori, classes conjugadas e estimadores de Bayes.**
3. **Métodos de estimação: momentos, máxima verossimilhança e mínimos quadrados.**
4. **Critérios para avaliação de estimadores: estimadores de mínima variância, a desigualdade de Cramer-Rao, eficiência e eficiência assintótica.**
5. **Intervalos de confiança: conceituação, interpretação e construção.**
6. **Testes de hipóteses: conceitos básicos, o lema de Neyman-Pearson, hipóteses compostas, a função poder, testes da razão de verossimilhança e testes bayesianos.**

REFERÊNCIAS

1. DeGroot, M.H., Schervish, M. J. (2011). Probability and Statistics. Pearson.
2. Roussas, G.G. (1973). A first course in mathematical statistics. Addison Wesley.
3. James, B.R. (1981). Probabilidade: um curso em nível intermediário. CNPq-IMPA Projeto Euclides.
4. Ross, S. (2012). A first course in probability. Pearson.
5. Casella, G., Berger, R.L. (2001). Statistical Inference. Duxbury Press.

Referências Complementares

6. Magalhães, M. N. (2011). Probabilidade e Variáveis Aleatórias. Edusp.
7. Bickel, P.J., Doksum, K.A. (1977). Mathematical Statistics. Holden-Day, Inc.

Datas prováveis das avaliações:

Prova 1 (módulo I) - 07/05

Prova 2 (módulo II) - 27/06

Prova 3 (módulos I e II) - 02/07

Critério: Média Aritmética