

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

Lançamento de um dado

Espaço Amostral (Ω)

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

Campeonato paulista

$\{\text{PAL, SP, PR, COR, ...}\}$ (20 elementos)

Temperatura

$[10, 40]$, um dos conjuntos possíveis para Ω

Mega Sena

$$\{A \subseteq \{1, \dots, 60\} : |A|=6\}$$

Evento

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

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

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

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

Congunto: coleção de objetos distintos. Denotaremos 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: $w \in A$, $w \notin A$ Conjunto vazio: \emptyset $A^c = \{w \in \Omega : w \notin A\}$

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

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

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

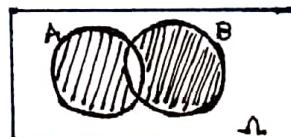
De modo mais geral, Λ conjunto de índices: $\bigcup_{\lambda \in \Lambda} A_\lambda$

$A \cap B$

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

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

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



$\square A \Delta B$

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

Diferença simétrica

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

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

Sequência de conjuntos: $A_1, A_2, A_3, A_4, \dots$

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

Exemplo:

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

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

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

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

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

$$5. E_n = \begin{cases} \{1, \dots, n\}, & n \text{ ímpar} \\ \{n\}, & n \text{ par} \end{cases}$$

- 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 $\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n$.
No exemplo 1., $\lim_{n \rightarrow \infty} A_n = \mathbb{N} / \{0\}$
- Quando $A_n \downarrow$, dizemos que existe $\lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$.
No exemplo 3., $\lim_{n \rightarrow \infty} A_n = \emptyset$.

Recordemos que

$$\underbrace{\inf \{a_1, a_2, \dots\}}_{b_n \uparrow} \leq a_n \leq \underbrace{\sup \{a_1, a_2, \dots\}}_{c_n \downarrow}$$

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

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

$$\underbrace{A_1 \cap A_2 \cap A_3 \cap \dots}_{B_n} \subseteq A_n \subseteq \underbrace{A_1 \cup A_2 \cup A_3 \cup \dots}_{C_n}$$

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

Como $(B_n)_{n \geq 1}$ é monótona, $\exists \lim_{n \rightarrow \infty} B_n = \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \liminf_{n \rightarrow \infty} A_n$.

Da mesma forma, $\exists \lim_{n \rightarrow \infty} C_n = \bigcap_{n=1}^{\infty} C_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \limsup_{n \rightarrow \infty} A_n$.

$$B_n \subseteq A_n \subseteq C_n$$

Se $\liminf_{n \rightarrow \infty} A_n = \limsup_{n \rightarrow \infty} A_n$, dizemos que existe

$$\lim_{n \rightarrow \infty} A_n = \liminf_{n \rightarrow \infty} A_n = \limsup_{n \rightarrow \infty} A_n$$

● Interpretações

$\liminf_{n \rightarrow \infty} A_n$: conjunto de pontos que estão em todos os A_n , exceto em um número finito deles.

$\limsup_{n \rightarrow \infty} A_n$: está em todos os A_n a partir de um ponto.

Exemplos:

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

$$\liminf_{n \rightarrow \infty} A_n = \mathbb{N} / \{0\}$$

$$\limsup_{n \rightarrow \infty} A_n = \mathbb{N} / \{0\}$$

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

$$\liminf_{n \rightarrow \infty} B_n = \{0\}$$

$$\limsup_{n \rightarrow \infty} B_n = \{0\}$$

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

$$\liminf_{n \rightarrow \infty} C_n = \emptyset$$

$$\limsup_{n \rightarrow \infty} C_n = \emptyset$$

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

$$\liminf_{n \rightarrow \infty} D_n = \emptyset$$

$$\limsup_{n \rightarrow \infty} D_n = \{0, 1\}$$

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

$$\liminf_{n \rightarrow \infty} E_n = \emptyset$$

$$\limsup_{n \rightarrow \infty} E_n = \mathbb{N} / \{0\}$$

Probabilidade e Inferência Estatística I

Aula 2 - 07 de março de 2013

Luis Gustavo Esteves

Primeira lista - já disponível

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

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

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

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

$$P(\Omega) = 1$$

$$P(A^c) = 1 - P(A)$$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2), \text{ para } A_1 \cap A_2 = \emptyset.$$

Agora, dado $\mathcal{F} \neq \emptyset$, 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. $\Omega = \mathbb{N}$

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

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:

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 contém $\{A\}$.

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

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

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

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

$\bigcup_{n=1}^{\infty} A_n$ e $(\bigcup_{n=1}^{\infty} A_n)^c$ NÃO são finitos, portanto $\bigcup_{n=1}^{\infty} A_n \notin \mathcal{F}$.
 $\begin{matrix} \uparrow \\ \text{pares} \end{matrix}$ $\begin{matrix} \uparrow \\ \text{ímpares} \end{matrix}$

Logo, \mathcal{F} não é σ -álgebra.

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

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

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

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

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

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

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

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.

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

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

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

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

Notemos que

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

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

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

. Seja $\mathcal{U} \neq \emptyset$ qualquer.

. Seja \mathcal{C} uma classe de subconjuntos qualquer de \mathcal{U} .

. Note que sempre existe ao menos uma σ-álgebra de subconjuntos de \mathcal{U} que contém \mathcal{C} .

. De fato, $\mathcal{P}(\mathcal{U})$ é σ-álgebra e $\mathcal{C} \subseteq \mathcal{P}(\mathcal{U})$.

Exemplo:

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

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

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

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

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

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

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

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

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

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

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

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

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

Considere

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

↓

é uma σ -álgebra que contém \mathcal{G}

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

Alternativamente, é chamada também de "álgebra gerada por \mathcal{G} ".

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

Exemplos:

1. $\Omega = \{1, 2, 3, 4, 5, 6\}$
 $\mathcal{G}_1 = \{\{1, 2, 3, 4\}, \{5, 6\}\} \quad \sigma(\mathcal{G}_1) = \{\emptyset, \Omega, \{1, 2, 3, 4\}, \{5, 6\}\}$
 $\mathcal{G}_2 = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\} \quad \sigma(\mathcal{G}_2) = \mathcal{P}(\Omega)$

2. Ω com ao menos 3 elementos

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

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

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

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

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

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

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

$\sigma(\mathcal{E})$

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

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

Os boreelianos podem ser gerados tanto por \mathcal{E} , \mathcal{E}_1 ou \mathcal{E}_2 .

Vejamos para \mathcal{E} e \mathcal{E}_1 a equivalência:

$$(-\infty, t) = \bigcup_{n=1}^{\infty} (-\infty, t - \frac{1}{n}] \Rightarrow (-\infty, t) \in \sigma(\mathcal{E}), \forall t \in \mathbb{R} \Rightarrow$$

↓
elemento
de \mathcal{E}_1

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

Analogamente

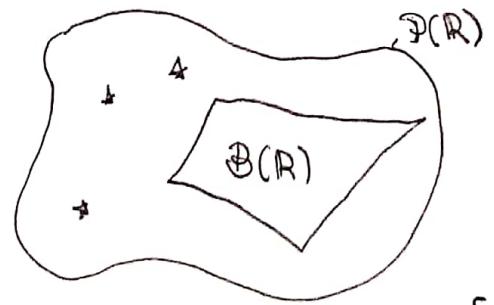
$$(-\infty, t] = \bigcap_{n=1}^{\infty} \left(-\infty, t + \frac{1}{n}\right) \Rightarrow (-\infty, t] \in \sigma(\mathcal{E}_1), \forall t \in \mathbb{R} \Rightarrow$$

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

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

• Ver livro How to gamble if you must.

How and How



$$[\alpha, \beta] = (-\infty, \beta] \cap (-\infty, \alpha]$$

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

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 (Ω, \mathcal{F}) , dá-se o nome de espaço mensurável.

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

$$1. P(\Omega) = 1$$

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

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

Exemplos:

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

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

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

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

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

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

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

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

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

$0 \notin \bigcup_{n=1}^{\infty} A_n \Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) = 0 \Rightarrow 0 \in \bigcap_{n=1}^{\infty} A_n^c \Rightarrow 0 \in A_1^c, \forall n \geq 1 \Rightarrow$
 $\Rightarrow 0 \notin A_n, \forall n \geq 1$ e, portanto, $P(A_n) = 0, \forall n \geq 1$.

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

Em geral, dado (Ω, \mathcal{F}) e $w_0 \in \Omega$

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

$$P(A) = \mathbb{I}_A(w_0), \text{ é probabilidade.}$$

(medida de probabilidade degenerada/concentrada em w_0).

2. (Ω, \mathcal{F})

Sejam $P_1, P_2: \mathcal{F} \rightarrow \mathbb{R}_+$, probabilidades e $\alpha \in (0, 1)$

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

$$A \in \mathcal{F} \mapsto P(A) = \alpha P_1(A) + (1-\alpha) P_2(A)$$

$$\begin{aligned} \text{(i)} \quad P(\Omega) &= \alpha P_1(\Omega) + (1-\alpha) P_2(\Omega) = \\ &= \alpha \cdot 1 + (1-\alpha) \cdot 1 = \\ &= 1 \end{aligned}$$

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

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= \alpha P_1\left(\bigcup_{n=1}^{\infty} A_n\right) + (1-\alpha) P_2\left(\bigcup_{n=1}^{\infty} A_n\right) = \\ &= \alpha \sum_{n=1}^{\infty} P_1(A_n) + (1-\alpha) \sum_{n=1}^{\infty} P_2(A_n) = \\ &= \sum_{n=1}^{\infty} \alpha P_1(A_n) + \sum_{n=1}^{\infty} (1-\alpha) P_2(A_n) = \\ &= \sum_{n=1}^{\infty} \alpha P_1(A_n) + (1-\alpha) P_2(A_n) = \\ &= \sum_{n=1}^{\infty} P(A_n) \end{aligned}$$

Como $P_1\left(\bigcup_{n=1}^{\infty} A_n\right)$ é um nº entre $0 \leq 1$, $\sum P(A_n)$ é convergente
Dai, $\alpha \sum P(A_n) = \sum \alpha P(A_n)$

3. Como no exemplo 2, defina:

$$\bar{P}(A) = P_1(A) + P_2(A)$$

i. $\bar{P}(\Omega) = P_1(\Omega) + P_2(\Omega) = 1 + 1 \neq 1 \times$

Note que \bar{P} atende a condição (ii).

4. Como no exemplo 2, defina

$$\bar{P}''(A) = \min\{P_1(A), P_2(A)\}$$

i. $\bar{P}''(\Omega) = \min\{P_1(\Omega), P_2(\Omega)\} = \min\{1, 1\} = 1 \checkmark$

ii. NÃO é verdade!

E.x: $\Omega = \mathbb{N}$ (Contradição)

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

$$P_1(A) = I_A(1)$$

$$P_2(A) = I_A(2)$$

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

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

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

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

5. (Ω, \mathcal{F})

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

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

Seja $P_B: \mathcal{F} \rightarrow \mathbb{R}_+$ dada por

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \quad (\text{medida de probabilidade condicional, dado } B)$$

i. $P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1 \checkmark$

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

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

Ào triô (Ω, \mathcal{F}, P) dá-se o nome de espaço de probabilidade, ou modelo probabilístico.

42 Exp. (Dados)

$$\begin{aligned} P(A_1) &\longrightarrow P(A_1 | \text{Dados}) \\ P(A_2) &\longrightarrow P(A_2 | \text{Dados}) \\ P(A_3) &\longrightarrow P(A_3 | \text{Dados}) \\ P(D) &\longrightarrow P(D | \text{Dados}) \end{aligned}$$

A medida de probabilidade condicional é útil na abordagem bayesiana.

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

Propriedades.

i. $P(\emptyset) = 0$

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

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

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

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

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

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

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

iii. $P(A^c) = 1 - P(A)$

Usando (ii), tomamos

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

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

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

$$P(B) = P((A \cap B) \cup (A^c \cap B)) = P(A \cap B) + P(A^c \cap B) \stackrel{\text{ASB}}{=} P(A) + P(A^c \cap B)$$

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

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

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

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

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

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

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

Definimos

$$B_1 = A_1$$

$$B_2 = A_1^c \cap A_2$$

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

:

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

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,

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

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

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

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

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

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

Assim,

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

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

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

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

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

$$P\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n) \quad (\text{continuidade da probabilidade})$$

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

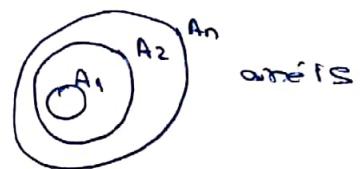
$$B_1 = A_1$$

$$B_2 = A_1^c \cap A_2$$

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

:

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



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

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

$$\begin{aligned} B_1 &= A_1, \quad A_2 = (A_1^c \cap A_2) \cup (A_1 \cap A_2) = \\ &= B_2 \cup A_1 \Rightarrow P(B_2) = P(A_2) - P(A_1) \end{aligned}$$

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

$$= \lim_{K \rightarrow \infty} P(A_K) = \lim_{n \rightarrow \infty} P(A_n) \blacksquare$$

Agora, se

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

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n^c\right) &= P\left(\lim_{n \rightarrow \infty} A_n^c\right) = \lim_{n \rightarrow \infty} P(A_n^c) = \lim_{n \rightarrow \infty} (1 - P(A_n)) = \\ &= 1 - \lim_{n \rightarrow \infty} P(A_n) \end{aligned}$$

Mas,

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

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

Logo,

$$P\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n)$$

o - Aditividade

Aditividade - Finita

$$\exists. P(\Omega) = 1$$

\Leftrightarrow

$$\exists. P(\Omega) = 1$$

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

2'. A_1, \dots, A_n disjuntos

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

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

3. $(A_n)_{n \geq 1}$, $A_n \neq \emptyset$, $\lim_{n \rightarrow \infty} P(A_n) = 0$

(continuidade para o varíó)

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 $A_i \cap A_j = \emptyset$, $i \neq j$.

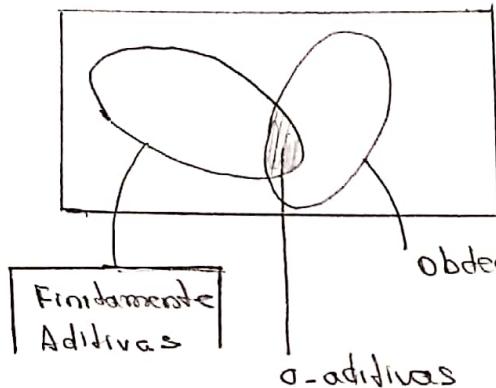
$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P(A_1 \cup A_2 \cup \dots \cup A_k \cup \left(\bigcup_{n=k+1}^{\infty} A_n\right)) \\ &= P(A_1) + P(A_2) + \dots + P(A_k) + P\left(\bigcup_{n=k+1}^{\infty} A_n\right) \end{aligned}$$

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= \underbrace{\sum_{i=1}^k P(A_i)}_{\text{a.s.}} + \underbrace{P\left(\bigcup_{n=k+1}^{\infty} A_n\right)}_{\text{b.u.}} \\ &\quad \downarrow \quad \uparrow \\ &\quad \text{seq. n}\bar{o}s \quad \text{seq. n}\bar{o}s \\ &\quad \text{decrescente} \quad \text{crescente} \end{aligned}$$

$$\begin{aligned} \Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) &= \left(\lim_{k \rightarrow \infty} \sum_{i=1}^k P(A_i)\right) + \lim_{k \rightarrow \infty} P\left(\bigcup_{n=k+1}^{\infty} A_n\right) = \begin{cases} \bigcup_{n=k+1}^{\infty} A_n \downarrow \emptyset \Rightarrow \\ \lim_{k \rightarrow \infty} P\left(\bigcup_{n=k+1}^{\infty} A_n\right) = 0 \end{cases} \\ &= \sum_{i=1}^{\infty} P(A_i) + \underset{\text{por } (3')}{0} \end{aligned}$$

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

Fixados Ω e \mathcal{F} :



Probabilidade e Inferência Estatística I

Aula 04 - 12/03/2013

Luiz Gustavo Esteves

• Modelos Discretos (Construção)

Ω : enumerável $\Omega = \{w_1, w_2, \dots\}$

$\mathcal{F}: \mathcal{P}(\Omega)$: 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 $A \in \mathcal{P}(\Omega)$

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

Verificação que P é, de fato, probabilidade.

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

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

$$\begin{aligned} P(A_1 \cup A_2) &= \sum_{i \geq 1} P_i \mathbb{I}_{A_1 \cup A_2}(w_i) = \sum_{i \geq 1} P_i [\mathbb{I}_{A_1}(w_i) + \mathbb{I}_{A_2}(w_i)] \\ &= \sum_{i \geq 1} [P_i \mathbb{I}_{A_1}(w_i) + P_i \mathbb{I}_{A_2}(w_i)] = \\ &= \sum_{i \geq 1} P_i \mathbb{I}_{A_1}(w_i) + \sum_{i \geq 1} P_i \mathbb{I}_{A_2}(w_i) = \\ &= P(A_1) + P(A_2). \end{aligned}$$

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

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

$$A_{n_K} \subseteq \{w_K, w_{K+1}, w_{K+2}, \dots\}$$

$$\text{Logo, } \forall K \geq 1 : P(A_{n_K}) \leq P(\{w_K, w_{K+1}, \dots\}) = \sum_{i \geq K} P_i \Rightarrow$$

$$\Rightarrow \limsup_K P(A_K) \leq \lim_{K \rightarrow \infty} P(\{w_K, w_{K+1}, \dots\})$$

$$= \lim_{K \rightarrow \infty} \sum_{i \geq K} P_i = 0 \quad (\text{pois a série } \{P_i\} \text{ é convergente})$$

Como $\liminf A_{n_K} = \limsup A_{n_K}$, então $\lim_{n \rightarrow \infty} P(A_n) = 0$.

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

Então P é probabilidade.

Exemplos

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

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

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

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

Logo, P é probabilidade.

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

$$\omega_i = 3i$$

$$P_i = \frac{2^{i-1}}{(i-1)!} \quad \sum P_i = e^2$$

$$P_i = \frac{P_i}{e^2} = \frac{e^{-2} 2^{i-1}}{(i-1)!}$$

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

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

$$p_i = i^2$$

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

● Quando Ω é finito com N elementos,

$$\text{tornando } p_i = \frac{1}{N}, \quad i=1, \dots, N$$

obtemos o Modelo Clássico.

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

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

Exemplo:

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

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

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

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

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

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

b) f é sobrejetora

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

B = f é sobrejetora

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

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

Mas,

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

e,

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

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

Em geral,

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

Logo

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

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

Assim, a probabilidade procurada é

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

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

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

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

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

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

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

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

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

$$P^*(B) = \inf \left\{ \sum_{i=1}^{\infty} P(A_i) \text{, onde } A_i \in \mathcal{L} \text{ e } B \subset \bigcup_{i=1}^{\infty} A_i \right\}$$

união finita de intervalos finitos

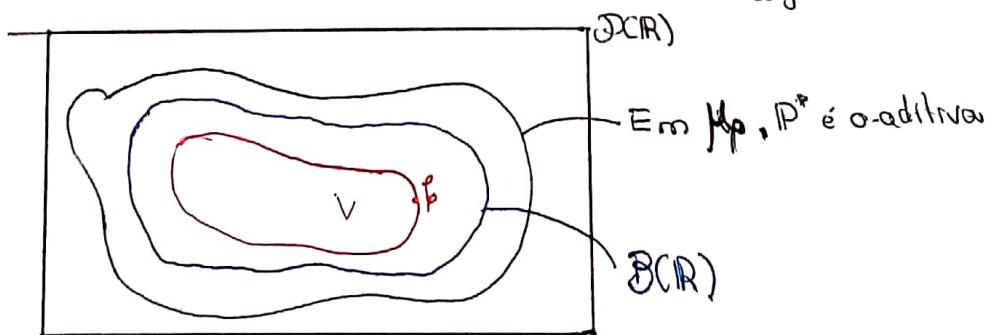
Mas assim,

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

Idéia: Restringir P^* ao conjunto

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

conjunto mensuráveis da reta.



Já disponível na 2ª lista.

Probabilidade e Inferência Estatística I

Aula 05 - 18/03/2013

Luiz Gustavo Esteves

Probabilidade Condicional

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

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

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

Exemplo:

1. Lançamento dado

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

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

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

$$F \in \mathcal{F} \mapsto P(F) = \frac{|F|}{6}$$

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

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

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

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

Da definição, temos

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

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

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

Exemplo. (Modelo de urna Pólya-Eggenberger)

10 brancos
6 verdes

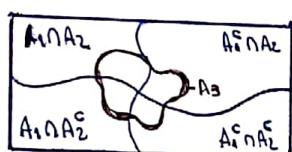
A_i : bola branca retirada na i -ésima extração, $i=1,2,3$.

a.

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

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

$$P(A_3) = ?$$



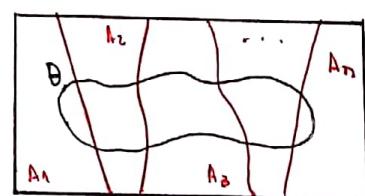
$$\begin{aligned} P(A_3) &= P(A_1 \cap A_2 \cap A_3) + P(A_1^c \cap A_2 \cap A_3) + P(A_1 \cap A_2^c \cap A_3) + P(A_1^c \cap A_2^c \cap A_3) = \\ &= P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2) + P(A_1^c) P(A_2 | A_1^c) P(A_3 | A_1^c \cap A_2) + \\ &\quad + P(A_1) P(A_2^c | A_1) P(A_3 | A_1 \cap A_2^c) + P(A_1^c) P(A_2^c | A_1^c) P(A_3 | A_1^c \cap A_2^c) \end{aligned}$$

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 $B \in \mathcal{F}$

$$\begin{aligned} P(B) &= P(B \cap \Omega) = P(B \cap (\bigcup_{i=1}^n A_i)) = \\ &= P\left(\bigcup_{i=1}^n (B \cap A_i)\right) = \sum_{i=1}^n P(B \cap A_i) \Rightarrow \end{aligned}$$



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

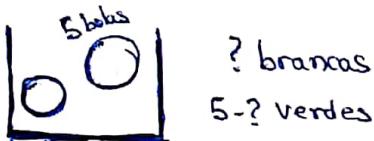
Teorema de Bayes. (Fórmula da Inversão)

(Ω, \mathcal{F}, 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,

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

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

Exemplo.

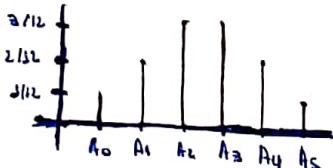


A_i : a urna possui i bolas brancas, $i=0, \dots, 5$.

$$P(A_0) = P(A_5) = 1/12$$

$$P(A_1) = P(A_4) = 2/12$$

$$P(A_2) = P(A_3) = 3/12$$



(Ω, \mathcal{F})
 (ω, ω)
 $(1, 1)$
 $(2, 3)$
 $(3, 2)$
 $(4, 1)$
 $(5, 0)$

Experimento que consiste em retirar 2 bolas da urna.

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

$$P(A_i | B_2) = ?$$

$$P(B_2 | A_i) = \frac{\binom{i}{2} \binom{s-i}{0}}{\binom{s}{2}}$$

Pelo Teorema de Bayes

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

$$\frac{\binom{0}{2} \binom{5}{0}}{\binom{5}{2}} \frac{1}{12}$$

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

$$P(A_0 | B_2) = P(A_1 | B_2) = 0.$$

$$P(A_2 | B_2) = \frac{\left[\binom{3}{2} \binom{5}{6} / \binom{5}{2} \right] \frac{3}{12}}{5} = \frac{3/120}{84/120} = \frac{3}{84}$$

$$P(A_3 | B_2) = \frac{9/120}{84/120} = \frac{9}{84}$$

$$P(A_4 | B_2) = \frac{12}{34} < P(A_5 | B_2) = \frac{10}{34}$$

Exemplo 2.

Doente

$$P(D) = 5\%$$

E: resultado do exame é positivo

$$P(E|D) = 95\%$$

$$P(E|D^c) = 1\%$$

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

Quando

$P(A|B) = P(A)$, dizemos que o evento A é independente do evento B.

A é independente de B $\Rightarrow P(A|B) = P(A)$.

Vimos que

$$P(A|B)P(B) = P(B|A)P(A)$$



$$P(B|A) = P(B)$$



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,

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

Exemplo. Lançamento dado

$$\begin{array}{lll} A = \{2, 4, 6\} & P(A) = \frac{1}{2} & P(A \cap B) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = P(A)P(B) \\ B = \{5, 6\} & P(B) = \frac{1}{3} & \\ C = \{1, 3, 4, 5, 6\} & P(C) = \frac{5}{6} & \text{Logo, } A \text{ e } B \text{ são independentes} \end{array}$$

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

Consequências:

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

$$A^c \in B \quad " \quad "$$

$$A^c \in B^c \quad " \quad "$$

EG

- $A \in B^c$

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

- $A^c \in B$

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

- $A^c \in B^c$

$$\begin{aligned} P(A^c \cap B^c) &= 1 - P(A \cup B) = 1 - P(A) - P(B) + P(A \cap B) = 1 - P(A) - P(B) + P(A)P(B) = \\ &= 1 - \underbrace{P(A) - P(B)(1 - P(A))}_{P(A^c)} = P(A^c) - P(B)P(A^c) = P(A^c)(1 - P(B)) = \\ &= P(A^c)P(B^c) \quad \checkmark \end{aligned}$$

Algumas Generalizações

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

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

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

DEFINIÇÃO: A_1, \dots, A_n são independentes dois a dois

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

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

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

$$\Omega = \{1, \dots, 6\}^2 \quad |\Omega| = 36$$

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

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

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.

$$P(A) = 6/36$$

$$P(A \cap B) = 3/36$$

$$P(A \cap B \cap C) = 1/36$$

$$P(B) = 18/36$$

$$P(A \cap C) = 2/36$$

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

$$P(C) = 12/36$$

$$P(B \cap C) = 6/36$$

B, D, E

$$P(B) = 18/36$$

$$P(B \cap D) = P(B \cap E) = P(D \cap E) = 1/4$$

$$P(D) = 18/36$$

B, D e E são independentes 2a2

$$P(E) = 18/36$$

$$\text{Mas } P(B \cap D \cap E) = \frac{9}{36} \neq \frac{1}{8} = P(B)P(D)P(E)$$

B^c, G, F

$$P(B^c) = 18/36$$

$$P(F) = 4/36$$

$$P(G) = 18/36$$

$$P(B^c \cap F \cap G) = 1/36 = P(B^c) P(F) P(G)$$

No entanto,

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

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

Definição: Uma sequência $(A_n)_{n \in \mathbb{N}}$ é de eventos independentes se A_1, A_2, \dots, A_n são independentes, $\forall n \in \mathbb{N}$.

(Ω, \mathcal{F}, P)

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

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

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

Probabilidade e Inferência Estatística I

Aula 06 - 15/03/2013

Luiz Gustavo Esteves

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

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

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

(D.M.) S. Wochshu

www.chacedegol.com.br

Definição: (Ω, \mathcal{F}, P)

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

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

$$\forall A_1 \in \mathcal{F}_1, \forall A_2 \in \mathcal{F}_2, P(A_1 \cap A_2) = P(A_1)P(A_2)$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Exemplo:

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

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

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

$$A \in \mathcal{F}$$

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

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

$$\mathcal{F}_2 = \sigma(\{[-n, n] : n \in \mathbb{N}\}) = \left\{ \bigcup_{n \in J} [-n, n] : J \subseteq \{1, 2, 3, \dots\} \right\}$$

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

$$\begin{array}{c} \uparrow \\ \# \end{array} \quad \begin{array}{c} \frac{1}{2} & \frac{1}{2} \\ \vdots & \vdots \\ \frac{1}{2} & \frac{1}{2} \end{array} \quad \frac{1}{2}^{n+1}$$

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

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

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

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

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

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

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

$$P(A_1 \cap A_2) = P(A_1)P(A_2)$$

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

Exemplo: (Ω, \mathcal{F}, P)

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

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

$$P(\emptyset \cap A_2) = P(\emptyset) P(A_2)$$

$$P(\Omega \cap A_2) = P(\Omega) P(A_2)$$

Espaços - produtos

$$(\Omega_1, \mathcal{F}_1, P_1)$$

$$(\Omega_2, \mathcal{F}_2, P_2)$$

:

$$(\Omega_n, \mathcal{F}_n, P_n)$$

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

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

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

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

$$P(E_1 \times \dots \times E_n) = P(E_1) \dots P(E_n)$$

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

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

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

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

$$P(E_1 \times E_2 \times \dots \times E_n \times \Omega_{n+1} \times \Omega_{n+2} \times \dots) = P_1(E_1) P_2(E_2) \dots P_n(E_n)$$

Exemplo: Lançamentos independentes de moedas

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

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

$$P_n(\{c\}) = p_n, 0 \leq p_n \leq 1, n \geq 1$$

1. Obter cara nos 3 primeiros lançamentos

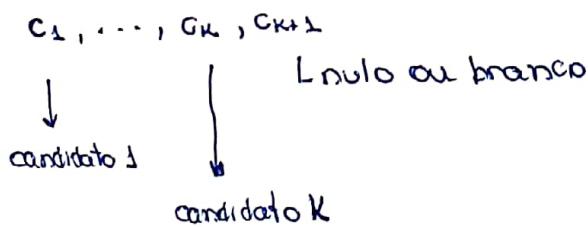
$$\begin{aligned} P(\{c\} \times \{c\} \times \{c\} \times \{\bar{c}\} \times \{\bar{c}\} \times \dots) &= \\ = P_1(\{c\}) P_2(\{c\}) P_3(\{c\}) &= \\ = P_1 P_2 P_3. \end{aligned}$$

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

$$\begin{aligned} P(\{\bar{c}\} \times \{\bar{c}\} \times \{c\} \times \{\bar{c}\} \times \{c\} \times \{\bar{c}\} \times \dots) &= \\ = P_1(\{\bar{c}\}) P_2(\{\bar{c}\}) P_3(\{c\}) P_4(\{\bar{c}\}) P_5(\{c\}) P_6(\{\bar{c}\}) &= \\ = 1 \cdot 1 \cdot P_3 \cdot 1 \cdot 1 \cdot (1 - P_6) &= P_3(1 - P_6). \end{aligned}$$

Variáveis Aleatórias

Exemplo: Eleições



$$\Omega = \{c_1, c_2, \dots, c_k, c_{k+1}\}^{2000}$$

$$X_j: \Omega \rightarrow \mathbb{R}$$

$$(w_1, \dots, w_{2000}) \mapsto X_j((w_1, \dots, w_{2000})) = \sum_{i=1}^{2000} \mathbb{I}_{\{c_i\}}(w_i), j = 1, \dots, k$$

Vetores Aleatórios

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

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

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

Imagem inversa de B pela transformação X .

Resultado:

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

(\Rightarrow) imediato

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

$$\mathcal{A}_0 = \{B \subseteq \mathbb{R} : X^{-1}(B) \in \mathcal{F}\}$$

\mathcal{A}_0 é σ -álgebra de subconjuntos de \mathbb{R}

1. $\mathbb{R} \in \mathcal{A}_0$

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

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

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

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

$$\begin{aligned} X(\omega) \in A^c &\Rightarrow X(\omega) \notin A \\ \omega \in X^{-1}(A) & \\ (X^{-1}(A))^c &= X^{-1}(A^c) \end{aligned}$$

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

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

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

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

(\Leftarrow) Voltando à prova

$$(-\infty, t] \in \mathcal{B} \Rightarrow \emptyset \subseteq \mathcal{B} \Rightarrow \sigma(\mathcal{B}) \subseteq \mathcal{B} \Rightarrow \mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}$$

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

todas as pré-imagens dos boreelianos

todas as pré-imagens de \mathcal{B}

Então, por definição, X é v.a.

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

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

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

$$\omega \mapsto X(\omega) = \omega$$

$$Y: \Omega \rightarrow \mathbb{R}$$

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

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

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

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

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 f -mensurável (matematicamente))

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

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

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

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

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

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

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

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

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

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

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

(Ω, \mathcal{F}, P)

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

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

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

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

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

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

Seja X uma variável aleatória

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

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:

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

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

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

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

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

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

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

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

$$V(\omega) = 2$$

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

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

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

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

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

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

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

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

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

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

$$-A = [-\omega : \omega \in A]$$

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

$$(-\frac{1}{2}, -\frac{1}{3}] \cup [\frac{1}{3}, \frac{1}{2}) \in \sigma(X)$$

$$[\frac{1}{3}, \frac{1}{2}) \notin \sigma(X)$$

Exemplo. Eleição

$$\Omega = \{c_1, \dots, c_n, c_{n+1}\}^{2000}$$

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

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

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

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

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

$$P_X : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}_+$$

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

1.

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

2.

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

$$\begin{aligned} P_X\left(\bigcup_{n=1}^{\infty} B_n\right) &= P(X^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right)) = P\left(\bigcup_{n=1}^{\infty} X^{-1}(B_n)\right) = \\ &= \sum_{n=1}^{\infty} P(X^{-1}(B_n)) = \sum_{n=1}^{\infty} P_X(B_n) \end{aligned}$$

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

À medida P_X , damos o nome de medida de probabilidade induzida pela variável aleatória X ou distribuição de X .

Voltando ao exemplo do dado:

$$X(\omega) = \omega$$

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

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

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

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

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

Exemplo: Lançamento de duas moedas

$$\Omega = \{\text{C, C}\}^2 \quad X = \begin{cases} 1, & \text{moeda de R\$1 resulta C} \\ 0, & \text{c.c.} \end{cases}$$

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

$$P(A) = \frac{|A|}{4} \quad Y: \text{número de caras observadas}$$

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

Do mesmo modo:

$$P_X(\{0\}) = 1/2$$

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

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

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

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

—————

Para caracterizar X (ω , mais precisamente, P_X) basta avaliarmos P_X para eventos em

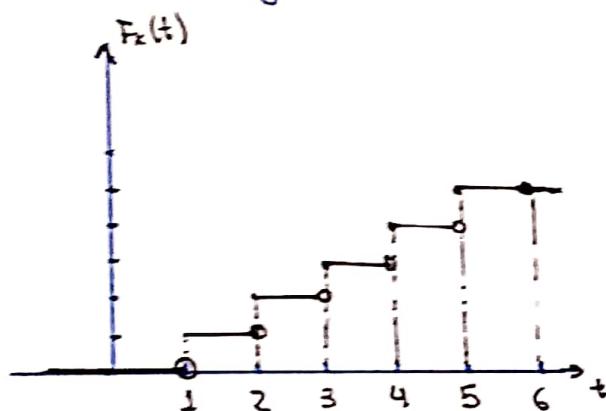
$$\mathcal{F} = \{(-\infty, t] : t \in \mathbb{R}\}$$

Definição: Seja X uma variável aleatória.

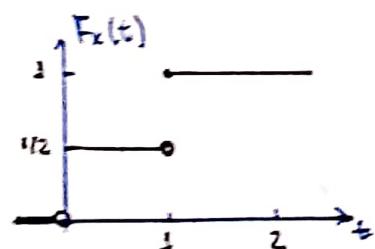
A função $F_X: \mathbb{R} \rightarrow [0,1]$ que associa a cada $t \in \mathbb{R}$, $F_X(t) = P_X((-\infty, t]) = P(\{\omega \in \Omega : X(\omega) \leq t\})$, dámos o nome de função de distribuição da variável aleatória X .

Exemplo: Dado

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



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



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

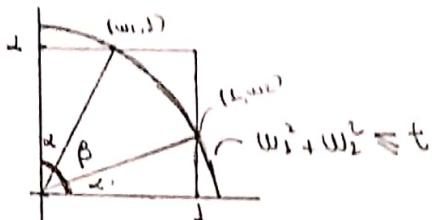
$$\mathcal{F} = \mathcal{B}([0,1]^2)$$

$$\begin{aligned} X: \Omega &\rightarrow \mathbb{R} \\ (\omega_1, \omega_2) \in \Omega &\mapsto X(\omega_1, \omega_2) = \sqrt{\omega_1^2 + \omega_2^2} \end{aligned}$$

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

$$F_X(t) = P(\{(x_1, x_2) \in \Omega : X(x_1, x_2) \leq t\})$$

$$F_x(t) = \begin{cases} 0, & t \leq 0 \\ \pi t^2/4, & 0 \leq t \leq 1 \rightarrow A_0 = \pi r^2, \text{ Nesse caso, } r = t/2. \\ \frac{\sqrt{t^2-1}}{2} + \frac{t^2}{2} \left(\frac{\pi}{2} - 2 \arctg(\sqrt{t^2-1}) \right), & 1 \leq t \leq \sqrt{2} \end{cases}$$

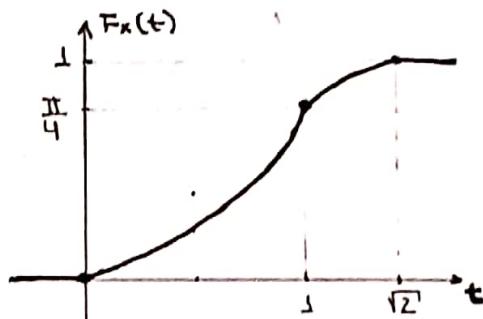


$$\beta = 2\alpha - \frac{\pi}{2},$$

$$w_1^2 + 1 = t^2 \Leftrightarrow w_1 = \sqrt{t^2 - 1}$$

$$A_A = \frac{1}{2} 1 \cdot \sqrt{t^2 - 1} = A \sqrt{t^2 - 1}$$

$$A_D = \frac{1}{2} \beta t^2$$



Propriedades de F_x :

1. F_x é não-decrescente

$$t_1 \leq t_2 \Rightarrow (-\infty, t_1] \subseteq (-\infty, t_2] \Rightarrow$$

$$\Rightarrow P_x((-\infty, t_1]) \leq P_x((-\infty, t_2]) \Rightarrow F_x(t_1) \leq F_x(t_2).$$

2. F_x é continua à direita

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

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

Seja $(t_n)_{n \geq 1}$ decrescente, t. q. $t_n \downarrow t \Rightarrow$

$\Rightarrow ((-\infty, t_n])_{n \geq 1}$ é monótona decrescente e, portanto,

$$\lim_{n \rightarrow \infty} P_X((-\infty, t_n]) = P_X\left(\underbrace{\lim_{n \rightarrow \infty} (-\infty, t_n]}\right) \Rightarrow \lim_{n \rightarrow \infty} F_X(t_n) = F_X(t)$$

$$= (-\infty, t]$$

3. Analogamente,

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

Consequências.

$$1. P(X < t) = P(\{\omega \in \Omega : X(\omega) < t\}) = \\ = \lim_{x \rightarrow t^-} F_X(x) = F_X(t^-)$$

$$2. P(X = t) = F_X(t) - F_X(t^-)$$

Comentários:

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



$$F_X$$

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

2. Outras definições de função distribuição.

$$F(t) = P(X < t)$$

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: $(\{1, 2, 3, 4, 5, 6\}, \mathcal{P}(\Omega_1), P_1)$

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

$(\{\text{CC, C}\bar{C}, \bar{C}C, \bar{C}\bar{C}\}, \mathcal{P}(\Omega_2), P_2)$

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

$([0, 1], \mathcal{B}([0, 1]), P_3)$

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

Note que

$$P_X = P_Y = P_Z \quad (\text{e, consequentemente})$$

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

Além disso, $1-X$ e X induzem a mesma P_X .

Nesse caso, diremos que:

$$X \stackrel{d}{=} 1-X \quad \text{ou}$$

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

Probabilidade e Inferência Estatística I

21 de março de 2013 - Aula 08

Luiz Gustavo Esteves

Nas aulas passadas:

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

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

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

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

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

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

Propriedades:

(1) F_x é não decrescente em cada componente

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

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

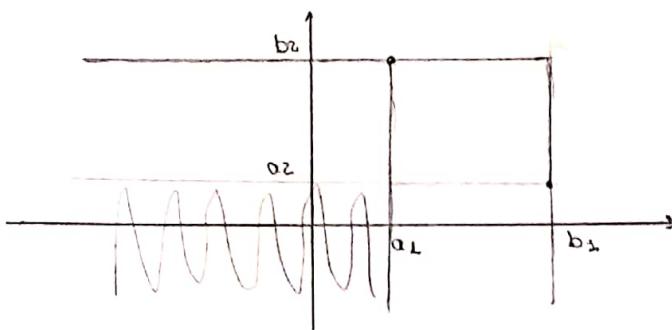
2. F_x é contínua à direita em cada componente

3. $\lim_{t_i \rightarrow -\infty} F_x(t_1, \dots, t_k) = 0$ e $\lim_{t_i \rightarrow \infty} F_x(t_1, \dots, t_k) = 1$
 para algum i para todos 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, basta m. 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.

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

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

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

$$\begin{aligned} &= (1 - e^{-b_1})[e^{-a_2} + a_2 e^{-a_2} - e^{-b_2} - b_2 e^{-b_2}] - (1 - e^{-a_1})[e^{-a_2} + a_2 e^{-a_2} - e^{-b_2} - b_2 e^{-b_2}] \\ &= [e^{-a_2} + a_2 e^{-a_2} - e^{-b_2} - b_2 e^{-b_2}] (e^{-a_1} - e^{-b_1}) \geq 0. \end{aligned}$$

Exemplo 2.

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

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

$$F_x(1, 1) - F_x(1, 0) - F_x(0, 1) + F_x(0, 0) =$$

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

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

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

(ii, 3, D)

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

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

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

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

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

Para todo $n \geq 1$

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

Logo,

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

$$= \{w : X_1(w) \leq t_1\} \cap \underbrace{\left(\bigcup_{n=1}^{\infty} \left[\bigcap_{j=2}^k \{w : X_j(w) \leq n\} \right] \right)}_{\mathcal{M}_2} \in \mathcal{F}$$

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

Logo, X_1 é variável aleatória.

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

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

Logo X é vetor aleatório.

Seja $X = (X_1, \dots, X_k)$ vetor aleatório com funções distributivas F_x .

Vamos avaliar:

$$\lim_{n \rightarrow \infty} F_x(t_1, n, n, \dots, n) = \lim_{n \rightarrow \infty} P_X((-\infty, t_1] \times (-\infty, t_2] \times \dots \times (-\infty, n]) =$$

$$= P_X(\lim_{n \rightarrow \infty} ((-\infty, t_1] \times (-\infty, n] \times \dots \times (-\infty, n])) = P_X(\bigcup_{n=1}^{\infty} ((-\infty, t_1] \times (-\infty, n] \times \dots \times (-\infty, n])) =$$

$$\Rightarrow P_X((-\infty, t_1] \times \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}) = P_X(w \in \Omega : X(w) \in (-\infty, t_1] \times \mathbb{R} \times \dots \times \mathbb{R}) =$$

$$= P_X(\{w \in \Omega : X_1(w) \leq t_1\} \cap \overbrace{\mathbb{R} \cap \mathbb{R} \cap \dots \cap \mathbb{R}}^{k-1 \text{ vezes}}) =$$

$$\Rightarrow \lim_{n \rightarrow \infty} F_X(t_1, n, \dots, n) = P_{X_L}([-\infty, t_1]) = F_{X_1}(t_1)$$

$F_X(t_1, \infty, \infty, \dots, \infty)$

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 $F_1, F_2, F_3: \mathbb{R} \rightarrow [0, 1]$ funções distribuição

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

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

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

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

2. $F_1, F_2, F_3, F_4: \mathbb{R} \rightarrow [0, 1]$ funções distribuição

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

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

$$F_{X_1}(t_1) = F_X(\infty, t_1, \infty) = F_1(t_1) F_3(t_1)$$

$$F_{(X_1, X_2)}(t_1, t_2) = F_X(t_1, \infty, t_2) = F_1(\min\{t_1, t_2\}) F_2(t_1) F_4(t_2)$$

Tipos de Vectors Aleatórios

1. Vectors Discretos

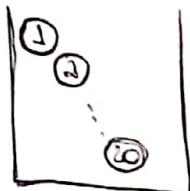
Definição: $X: \Omega \rightarrow \mathbb{R}^k$ é discreto se $\exists \{x_n\}_{n \geq 1}, x_n \in \mathbb{R}^k, \forall n \geq 1$, tal que

$$\mathbb{P}_X(\{x_n\}_{n \geq 1}) = 1.$$

Nesse caso, denotando $\mathbb{P}_X(\{x_n\}) = p_n, n \geq 1$, \mathbb{P}_X ficou "caracterizada" (representada) pela lista $((x_n, p_n))_{n \geq 1}$. \rightarrow função (densidade) de probabilidade de X

$$\mathbb{P}_X(A) = \sum_{n \geq 1} p_n \mathbb{I}_A(x_n)$$

Exemplo:



Retirar 4 bolas

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

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

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

$$X = (X_1, X_2)$$

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

$$\mathbb{P}_X(\{(a, b)\}) = \mathbb{P}(X = (a, b)) = \mathbb{P}(X_1 = a, X_2 = b) = \left[\frac{\binom{b-a}{2}}{\binom{20}{4}} \right], \quad (a, b) \in X(\Omega)$$

\downarrow

$$\{ \omega \in \Omega : X(\omega) = (a, b) \}$$

$a \leq a \leq b$
 $a \leq a \leq b$
 $a \neq a$

$$X_i(\Omega) = \{X_i(\omega) : \omega \in \Omega\} \subset \mathbb{N}, i=1, 2.$$

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

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

PRINCIPAIS MODELOS DISCRETOS

1. Uniforme

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

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

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

2. Bernoulli:

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

$$P(X=1) = p \quad P(X=0) = 1-p$$

$$P(X=x) = p^x (1-p)^{1-x} \prod_{j \in \{0,1\}} (x_j)$$

Notação: $X \sim \text{Ber}(p)$.

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

$$P(X=j) = \binom{n}{j} p^j (1-p)^{n-j} \prod_{j \in \{0,1,\dots,n\}} (j)$$

Notação: $X \sim \text{Bin}(n, p)$

4. Poisson

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

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

Bolinhas: ① ② ③ ④ ... ⑩

Distribuição dos eventos raros

Urnas: $\begin{smallmatrix} \sqcup \\ 1 \end{smallmatrix} \quad \begin{smallmatrix} \sqcup \\ 2 \end{smallmatrix} \quad \begin{smallmatrix} \sqcup \\ 3 \end{smallmatrix} \quad \begin{smallmatrix} \sqcup \\ 4 \end{smallmatrix} \quad \dots \quad \begin{smallmatrix} \sqcup \\ N \end{smallmatrix}$

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

$X \sim \text{Bin}(N, 2/N)$

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

5. Geométrico

Dizemos que a v.a. X é distribuída segundo o modelo GEOMÉTRICO de parâmetro p , se $p \in [0, 1]$.

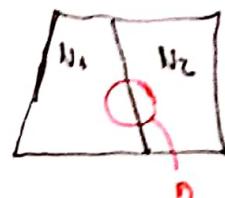
$$\mathbb{P}(X=j) = (1-p)^{j-1} \cdot p \mathbb{I}_{\{j=1,2,3,4,\dots\}}(j)$$

Notação. $X \sim \text{Geo}(p)$

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

$$\mathbb{P}(X=j) = \frac{\binom{N_1}{j} \binom{N_2}{n-j}}{\binom{N_1+N_2}{n}} \mathbb{I}_{\{0,1,2,\dots,n\}}(j), \quad \text{com } \begin{cases} 0 & \text{se } a < b \\ 1 & \text{se } a \geq b \end{cases}$$



- útil na amostragem: captura-recaptura, sorte repositório.

7. Multinomial

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

$$\mathbb{P}(X_1=x_1, \dots, X_n=x_n) = \frac{n!}{x_1! x_2! \dots x_n!} p_1^{x_1} p_2^{x_2} \dots p_n^{x_n} \mathbb{I}_A(x_1, \dots, x_n), \text{ onde}$$

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

$$x_{n+1} = n - (x_1 + \dots + x_n) \text{ e } p_{n+1} = 1 - (p_1 + \dots + p_n)$$

$$X \sim \text{Multinomial}(n, (p_1, \dots, p_n))$$

Probabilidade e Inferência Estatística I

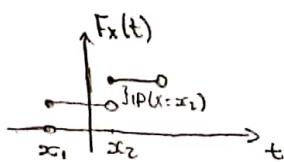
Aula 09 - 22/03/2013

Luiz Gustavo Esteves

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

Discreto

$$\mathbb{P}_x, F_x, ((\omega_n, p_n))_{n \in \mathbb{Z}}$$



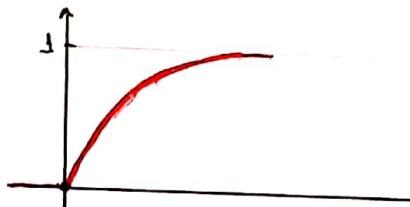
Variáveis Contínuas

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.

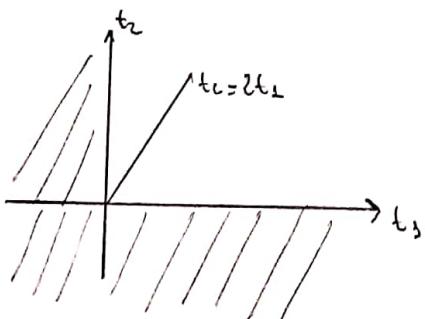
$$F_X(t) = (1 - e^{-t}) \mathbb{I}_{\mathbb{R}_+}(t)$$



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

Exemplo 2. X vetor aleatório

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

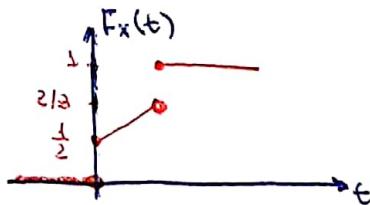


X é vetor aleatório contínuo

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

X é contínuo.

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



Seja X um vetor aleatório com função distribuição F_X .

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

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

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

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

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

Exemplos:

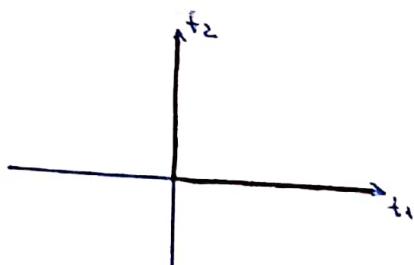
$$1. F_X(t) = (1 - e^{-t}) \mathbb{I}_{\mathbb{R}_+}(t)$$

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

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

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

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



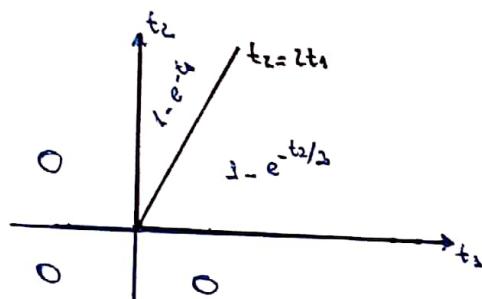
L Problema
mediante nula

$$\frac{\partial^2 F_x(t_1, t_2)}{\partial t_1 \partial t_2} = e^{-t_1} (e^{-t_2} - e^{-t_2} + t_2 e^{-t_2}), t_1 > 0$$

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

E continuo

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



L problema

$$\frac{\partial^2 F_x(t_1, t_2)}{\partial t_1 \partial t_2} = 0, \text{ para todo } (t_1, t_2) \text{ em que } F_x \text{ é diferen-} \\ \text{ciável.}$$

(F_x é singular: F_x é continua, mas $F'(t) = 0, \forall t$)

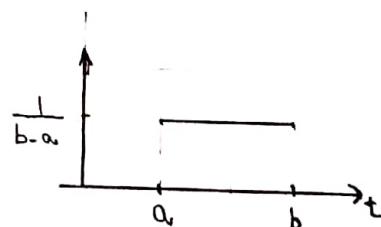
Ver Função de distribuição de Cantor (Barry James)

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

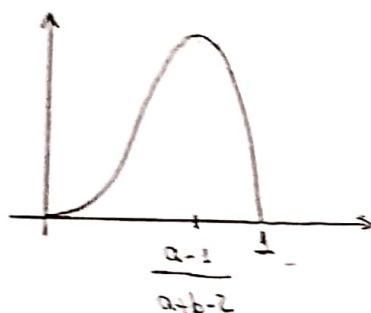
X é distribuído segundo o modelo BETA de parâmetros $a \neq b$, $a, b > 0$, se sua densidade é dada por

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

Análise da influência dos parâmetros:

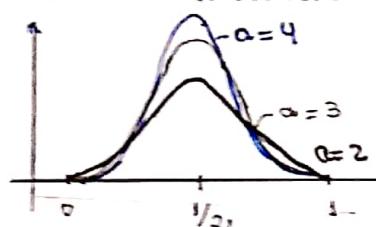
- Se $a=b=1 \rightarrow U(0,1)$

- Se $a, b \geq 1 \rightarrow$

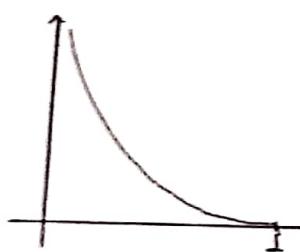


Se $a > b$, o ponto crítico fica mais próximo de 1 e, mais distante, se $b \geq 0$.

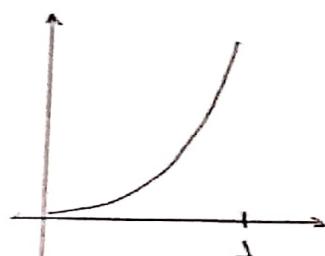
- Se $a=b$, temos um modelo simétrico em torno de 0,5. Quanto maior forem $a \neq b$, mais leptocúrtica será a curva.



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



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

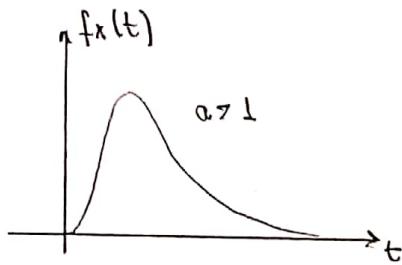


$X \sim \text{Beta}(a, b)$

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

X é distribuído segundo o modelo GAMA de parâmetros $a \in \mathbb{R}$, $a, b > 0$, 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:

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

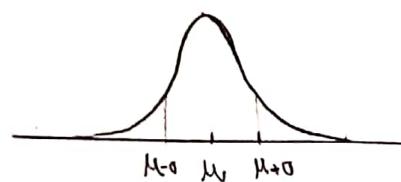
$$b=\frac{1}{2} \rightarrow X \sim \chi^2_{2a} \rightarrow \text{Modelo Qui-quadrado de parâmetro } 2a.$$

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

4. NORMAL

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

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



5. DIRICHLET

$X = (X_1, \dots, X_k)$ é distribuído segundo o modelo DIRICHLET, de parâmetros $\alpha_1, \alpha_2, \dots, \alpha_k$ e α_0 , $\alpha_i > 0$, $i = 0, 1, \dots, k$ se

$$f_X(t_1, \dots, t_k) = \frac{\Gamma(\alpha_0 + \alpha_1 + \dots + \alpha_k)}{\Gamma(\alpha_0)\Gamma(\alpha_1)\Gamma(\alpha_2)\dots\Gamma(\alpha_k)} t_1^{\alpha_1-1} t_2^{\alpha_2-1} \dots t_k^{\alpha_k-1} (1-t_1-t_2-\dots-t_k)^{\alpha_0-1} \prod_{i=1}^k \mathbb{I}_{S_k}(t_i, t_k)$$

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

Se DIRICHLET ($\alpha_0, \alpha_1, \alpha_2$)



Se DIRICHLET ($\alpha_0, \alpha_1, \alpha_2, \alpha_3$)



Variáveis Aleatórias Mistas

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

Dizemos que X é uma v.a. mista se

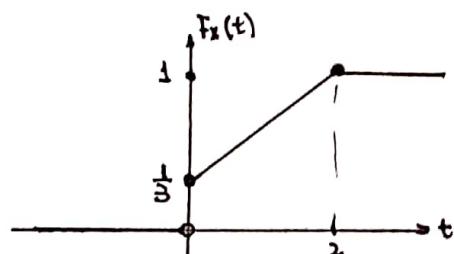
$$F_X(t) = \alpha F_c(t) + (1-\alpha) F_d(t), \text{ para algum } \alpha \in (0,1)$$

e F_c função de distribuição absolutamente contínua e

F_d função de distribuição discreta

Exemplo:

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



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

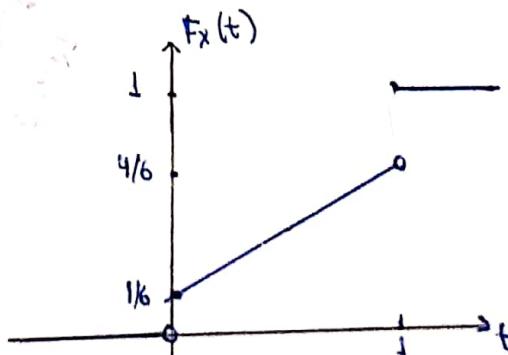
$$F_d(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{3}, & 0 \leq t \leq 2 \\ 1, & t > 2 \end{cases}$$

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

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

2.

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



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

Função de Distribuição de uma Bernoulli ($\frac{2}{3}/\frac{1}{3}$)

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

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

$$\left(\frac{1+2}{6} \right) \quad \left(\frac{0}{6} - \frac{1}{6} \right)$$

Próxima aula:

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

F_X

$$F_{x_1}(t) = F_X(t, \infty, \infty, \dots, \infty)$$

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

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

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

$$F_X, F_x: \mathbb{R}^n \rightarrow [0, 1] \quad F_{x_i}(t) = F_X(\infty, \infty, t, \infty, \dots)$$

caso discreto $((x_n, p_n))_{n \in \mathbb{N}}$

abs.
contínuo f_X

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

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

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

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

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

Exemplo. 1.

$$\begin{array}{|c|} \hline 20 \\ \hline \text{below} \\ \hline \end{array}$$

$$P(X_1=a, X_2=b) = \frac{\binom{b-a-1}{2}}{\binom{20}{4}}$$

$\overbrace{\quad}^{\text{II}(a)}$	$\overbrace{\quad}^{\text{II}(b)}$
$\{1, \dots, 17\}$	$\{a+3, \dots, 20\}$

$\overbrace{\quad}^{\text{II}(b)}$	$\overbrace{\quad}^{\text{II}(a)}$
$\{4, \dots, 20\}$	$\{1, \dots, b-3\}$

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

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

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

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

$$\sum_{x_1=a+2}^{20} \binom{x_2-a-1}{2} = \binom{2}{2} + \binom{3}{2} + \dots + \binom{19-a}{2} \stackrel{\text{T. Pascal}}{=} \binom{20-a}{3}$$

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

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

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

ende

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

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

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

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

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

\Rightarrow

$$P(X_1=x_1, X_2=x_2) = \frac{n!}{x_1! x_2! (n-x_1-x_2)!} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{n-x_1-x_2} \prod_{A_{2,n}} (x_1, x_2)$$

$X = (X_1, \dots, X_n)$ absolutamente continuo

$$F_{X_1}(t_1) = F_X(t_1, \infty, \infty, \dots, \infty) = \int_{-\infty}^{t_1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_n) dx_K dx_{K-1} \cdots dx_2 dx_1$$

$$= \int_{-\infty}^{t_1} \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_n) dx_K \cdots dx_2 \right] dx_1$$

$f_{X_1}(x_1)$

Rever Teorema de Fubini!

Exemplo:

$$X \sim (X_1, X_2, \dots, X_n) \sim \text{DIR}(a_1, a_2, \dots, a_k, a_0)$$

$$f_X(x_1, \dots, x_n) = \frac{\Gamma(a_0 + a_1 + \dots + a_n)}{\Gamma(a_0)\Gamma(a_1)\cdots\Gamma(a_n)} x_1^{a_1-1} x_2^{a_2-1} \cdots x_n^{a_n-1} (1-x_1-\dots-x_n)^{a_0-1} \mathbb{I}_{S_K}(x_1, \dots, x_n)$$

$$f_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\Gamma(a_0 + \dots + a_n)}{\Gamma(a_0)\Gamma(a_1)\cdots\Gamma(a_n)} x_1^{a_1-1} x_2^{a_2-1} x_3^{a_3-1} \cdots x_n^{a_n-1} (1-x_1-\dots-x_n)^{a_0-1} \mathbb{I}_{S_K}(x_1, \dots, x_n) dx_3 \cdots dx_n$$

$$S_K = \{(t_1, \dots, t_n) \in \mathbb{R}_+^n : t_1 + \dots + t_n \leq 1\}$$

$$= \frac{\Gamma(a_0 + a_1 + \dots + a_n)}{\Gamma(a_1)\Gamma(a_2)} x_1^{a_1-1} x_2^{a_2-1} \mathbb{I}_{S_2}(x_1, x_2) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{x_3^{a_3-1} \cdots x_n^{a_n-1} (1-x_1-\dots-x_n)^{a_0-1}}{\Gamma(a_3)\Gamma(a_4)\cdots\Gamma(a_n)}$$

onde

$$\mathbb{I}_{S_{K-2}}(x_3, \dots, x_n) dx_3 \cdots dx_n$$

$$S_{K-2} = \{(t_1 + t_2, \dots, t_{K-2}) \in \mathbb{R}_+^{K-2} : t_1 + \dots + t_{K-2} \leq 1 - x_1 - x_2\}$$

Fazendo

$$t_1 = \frac{x_3}{1-x_1-x_2}$$

$$t_2 = \frac{x_4}{1-x_1-x_2}$$

⋮

$$t_{k-2} = \frac{x_k}{1-x_1-x_2}$$

$$= \frac{\Gamma(a_0 + \dots + a_K)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_0+a_3+a_4+\dots+a_K)} x_1^{a_1-1} x_2^{a_2-1} \prod_{S_2}^{\infty} (x_1, x_2) \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{\Gamma(a_0+a_3+a_4+\dots+a_K)}{\Gamma(a_0)\Gamma(a_3)\dots\Gamma(a_K)} t_1^{a_3-1} t_2^{a_4-1} \dots t_{K-2}^{a_{K-1}} (1-t_1-\dots-t_{K-2})^{a_{K-1}}$$

$$\cdot \prod_{1-x_1-x_2}^{a_0+a_3+\dots+a_{K-1}} \prod_{S_{K-2}}^{(t_3, \dots, t_{K-2})} dt_3 \dots dt_{K-2}$$



$$f_{X_1, X_2}(x_1, x_2) = \frac{\Gamma(a_0 + \dots + a_K)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_0+a_3+a_4+\dots+a_K)} x_1^{a_1-1} x_2^{a_2-1} (1-x_1-x_2)^{a_0+a_3+\dots+a_{K-1}} \prod_{S_2}^{\infty} (x_1, x_2)$$

$$\left[\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{\Gamma(a_0+a_3+a_4+\dots+a_K)}{\Gamma(a_0)\Gamma(a_3)\dots\Gamma(a_K)} t_1^{a_3-1} t_2^{a_4-1} \dots t_{K-2}^{a_{K-1}} (1-t_1-\dots-t_{K-2})^{a_{K-1}} \prod_{S_{K-2}}^{(t_3, \dots, t_{K-2})} dt_3 \dots dt_{K-2} \right] = 1$$

$$(X_1, X_2) \sim \text{DIR}(a_1, a_2, a_0+a_3+\dots+a_K)$$

INDEPENDÊNCIA

$X = (X_1, \dots, X_K)$ vetor aleatório

Definição: Dizemos que X_1, \dots, X_K são independentes se

$$\mathbb{P}(X_1 \in B_1, X_2 \in B_2, \dots, X_K \in B_K) = \prod_{i=1}^K \mathbb{P}(X_i \in B_i), \quad \forall B_i \in \mathcal{B}(\mathbb{R}), \quad \forall i = 1, \dots, K$$

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.

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

$$X_1(\omega) = \begin{cases} \mathbb{I}_{\{1,2\}}(\omega) \\ \mathbb{I}_{\{3,4,5,6\}}(\omega) \end{cases}$$

$$X_2(\omega) = \begin{cases} \mathbb{I}_{\{2,4,6\}}(\omega) \\ \mathbb{I}_{\{1,3,5\}}(\omega) \end{cases}$$

$$X_3(\omega) = \begin{cases} \mathbb{I}_{\{4,5,6\}}(\omega) \\ \mathbb{I}_{\{1,2,3\}}(\omega) \end{cases}$$

$$\{\omega : X_1(\omega) \leq t\} = \begin{cases} \emptyset, & t < 0 \\ \{3, 4, 5, 6\}, & 0 \leq t < 1 \\ \Omega, & t \geq 1 \end{cases}$$

$$\sigma(X_1) = \{\emptyset, \{1, 2\}, \{3, 4, 5, 6\}\}$$

$$\sigma(X_2) = \{\emptyset, \Omega, \{2, 4, 6\}, \{1, 3, 5\}\}$$

$$\sigma(X_3) = \{\emptyset, \Omega, \{1, 2, 3\}, \{4, 5, 6\}\}$$

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 $\mathbb{P}(\{3, 2\} \cap \{1, 2, 3\}) \neq \mathbb{P}(\{1, 2\}) \cdot \mathbb{P}(\{1, 2, 3\})$),

dizemos que X_1 e X_3 não são independentes.

Nesse caso, dizemos que X_1 e X_3 são dependentes.

Verificar que X_2 e X_3 são dependentes!

Ex. 2:

$$\Omega = \mathbb{Z}_*$$

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

$$P(\{j\}) = \left(\frac{1}{2}\right)^{|j|+1}$$

$$X_1(\omega) = \begin{cases} 1, & \omega > 0 \\ 0, & \text{c.c.} \end{cases} \quad X_2(\omega) = |\omega| = \begin{cases} \omega, & \omega > 0 \\ -\omega, & \omega \leq 0 \end{cases}$$

Dá-se

$$\sigma(X_1) = \{\emptyset, \Omega, \mathbb{Z}_*, \mathbb{Z}_*^+\}$$

$$\{\omega : X_2(\omega) \leq t\} = \begin{cases} \emptyset, & t < 1 \\ \{-1, 1\}, & 1 \leq t < 2 \\ \{-2, -1, 1, 2\}, & 2 \leq t < 3 \\ \vdots \end{cases}$$

$$\sigma(X_2) = \sigma(\{\{-n, n\} : n \in \mathbb{N}\}) = \left\{ \bigcup_{n \in \mathbb{N}} \{-n, n\} : J \subseteq \{1, 2, 3, \dots\} \right\}$$

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 módulo do número, continuamos incertos quanto ao sinal do mesmo.

INDEPENDÊNCIA

Resultado: X_1, \dots, X_k são independentes $\Leftrightarrow F_X(t_1, \dots, t_k) = F_{X_1}(t_1) \dots F_{X_k}(t_k)$, para todo $(t_1, t_2, \dots, t_k) \in \mathbb{R}^k$.

$$\begin{array}{l|l} X = (X_1, \dots, X_K) & (\Rightarrow) B_j = (-\infty, t_j], j=1, \dots, K \\ P(X_1 \in B_1, \dots, X_K \in B_K) & (\Leftarrow)? \text{ Gimos} \\ = \prod_{i=1}^K P(X_i \in B_i) & \end{array}$$

1. $F_{(x_1, x_2)}(t_1, t_2) = (1 - e^{-t_1})(1 - e^{-t_2} - t_2 e^{-t_2}) \mathbb{I}_{\mathbb{R}_+}(t_1) \mathbb{I}_{\mathbb{R}_+}(t_2)$

$$F_{x_1}(t_1) = F_{(x_1, x_2)}(t_1, \infty) = (1 - e^{-t_1}) \mathbb{I}_{\mathbb{R}_+}(t_1)$$

$$F_{x_2}(t_2) = F_{(x_1, x_2)}(\infty, t_2) = (1 - e^{-t_2} - t_2 e^{-t_2}) \mathbb{I}_{\mathbb{R}_+}(t_2)$$

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. $F_{(x_1, x_2)}(t_1, t_2) = (1 - e^{-\min\{t_1, t_2/2\}}) \mathbb{I}_{\mathbb{R}_+}(t_1) \mathbb{I}_{\mathbb{R}_+}(t_2)$

$$F_{x_1}(t_1) = (1 - e^{-t_1}) \mathbb{I}_{\mathbb{R}_+}(t_1)$$

$$F_{x_2}(t_2) = (1 - e^{-t_2/2}) \mathbb{I}_{\mathbb{R}_+}(t_2)$$

$$F_{(x_1, x_2)}(1, 2) = 1 - e^{-1}$$

Logo, X_1 e X_2 são dependentes!

$$F_{x_1}(1) = 1 - e^{-1} \quad F_{x_2}(2) = 1 - e^{-1}$$

INDEPENDÊNCIA

Resultado: X_1, X_2, \dots, X_K são independentes \Leftrightarrow

$$P(X_1=x_1, \dots, X_K=x_K) = P(X_1=x_1) \dots P(X_K=x_K), \quad \forall x_i \in X_i(\omega), i=1, \dots, K$$

$X = (X_1, \dots, X_K)$ discreto.

$$\Rightarrow B_j = \{x_j\}, j=1, \dots, n$$

$$\Leftrightarrow F_X(t_1, \dots, t_n) = P(X_1 \leq t_1, X_2 \leq t_2, \dots, X_n \leq t_n) =$$

$$= P(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n), \text{ onde } B_i = (-\infty, t_i] \cap X_i(\omega)$$

\downarrow
prob. $(X_1, \dots, X_n) \in B_i$ ($i \in \mathbb{N}$) $\times B_n$ (prod. cartesiana) \downarrow
conjunto finito ou enumerável

$$= \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} \dots \sum_{x_n \in B_n} P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) =$$

$$= \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} \dots \sum_{x_n \in B_n} P(X_1 = \infty) P(X_2 = \infty) \dots P(X_n = \infty) =$$

$$= \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} \dots \sum_{x_{n-1} \in B_{n-1}} P(X_1 = \infty) \dots P(X_{n-1} = \infty) \underbrace{\sum_{x_n \in B_n} P(X_n = \infty)}_{\text{repetindo o processo de isolação}} =$$

$$= \dots = \prod_{i=1}^n \left(\underbrace{\sum_{x_i \in B_i} P(X_i = \infty)}_{F_{X_i}(t_i)} \right) \Rightarrow$$

$$F_X(t_1, \dots, t_n) = F_{X_1}(t_1) F_{X_2}(t_2) \dots F_{X_n}(t_n) \quad \text{e, pelo resultado anterior,}$$

X_1, \dots, X_n são independentes!

1. Urna.

$$P(X_1=a, X_2=b) = \frac{\binom{b-a-1}{2}}{\binom{20}{4}} \mathbb{I}(a) \mathbb{I}(b)$$

$\{1, \dots, 17\}$ $\{a+3, \dots, 20\}$

$$P(X_1=a) = \frac{\binom{20-a}{3}}{\binom{20}{4}} \mathbb{I}(a)$$

$\{1, \dots, 17\}$

$$P(X_2=b) = \frac{\binom{b-1}{3}}{\binom{20}{4}} \mathbb{I}(b)$$

$\{4, \dots, 20\}$

$$\begin{array}{l} a=5 \\ b=4 \end{array}$$

$$P(X_1=5, X_2=4) = 0$$

$$P(X_1=5) = \frac{\binom{15}{3}}{\binom{20}{4}} > 0 \quad e \quad P(X_2=4) = \frac{1}{\binom{20}{4}} > 0$$

Como $P(X_1=5, X_2=4) \neq P(X_1=5)P(X_2=4)$, então X_1 e X_2 são dependentes!

Exemplo 2

x_1	1	2	3	
0	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$P(X_1=0) = \frac{1}{3}$
1	$\frac{2}{12}$	$\frac{4}{12}$	$\frac{2}{12}$	$P(X_1=1) = \frac{2}{3}$
$P(X_2=1)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	

X_1 e X_2 são independentes!

Exemplo 3.

x_1	1	2	3	
0	$\frac{1}{12}$	0	$\frac{5}{12}$	$\frac{1}{3}$
1	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{3}{12}$	$\frac{2}{3}$
$P(X_2=2)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	

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!
 $(P(X_1=0, X_2=2) \neq P(X_1=0)P(X_2=2))$.

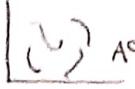
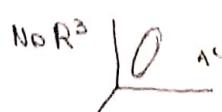
Probabilidade e Inferência Estatística I

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 "Volume_K(A^c) = 0".

No \mathbb{R}^2  No \mathbb{R}^3  A^c não pode ter volume

$$\begin{matrix} a=5 \\ b=4 \end{matrix}$$

$$P(X_1=5, X_2=4) = 0$$

$$P(X_1=5) = \frac{\binom{15}{3}}{\binom{20}{4}} > 0 \quad \in P(X_2=4) = \frac{1}{\binom{20}{4}} > 0$$

Como $P(X_1=5, X_2=4) \neq P(X_1=5)P(X_2=4)$, então X_1 e X_2 são dependentes!

Exemplo 2.

x_1	0	1	2	3	
x_2	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{1}{12}$		$P(X_1=0) = \frac{1}{3}$
	$\frac{2}{12}$	$\frac{4}{12}$	$\frac{2}{12}$		$P(X_1=1) = \frac{2}{3}$
	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$		$P(X_2=1) \quad P(X_2=2) \quad P(X_2=3)$

X_1 e X_2 são independentes!

Exemplo 3.

x_1	1	2	3	
0	$\frac{1}{12}$	0	$\frac{9}{12}$	$\frac{1}{3}$
1	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{3}{12}$	$\frac{2}{3}$
	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	

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!
 $(P(X_1=0, X_2=1) \neq P(X_1=0)P(X_2=1))$.

Probabilidade e Inferência Estatística I

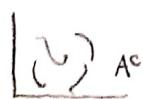
04/04/2013 - Aula 11

Luz Gustavo Esteves

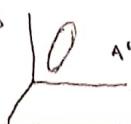
$X = (X_1, \dots, X_K)$ absolutamente contínuo

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 "Volume_K(A^c) = 0".

No \mathbb{R}^2



No \mathbb{R}^3



A^c não pode ter volume

A^c não pode ser \emptyset

(⇒)

$$F_X(t_1, \dots, t_n) = F_{X_1}(t_1) \dots F_{X_n}(t_n)$$

$$\frac{\partial^k F_X}{\partial t_1 \dots \partial t_n} = \downarrow$$

$$f_X(t_1, \dots, t_n) = f_{X_1}(t_1) \dots f_{X_n}(t_n)$$

(⇐)

$$\int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \dots \int_{-\infty}^{t_n} f_{X_1}(x_1) \dots f_{X_n}(x_n) dx_n \dots dx_1 = F_{X_1}(t_1) \dots F_{X_n}(t_n).$$

Transformações de Vetores Aleatórios

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

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^k \quad (x_1, \dots, x_n) \mapsto \left(\bar{x}, \frac{\sum (x_i - \bar{x})^2}{n} \right)$$

 X v.a. X^2 $\text{cov}(X)$

$$T: \mathbb{R} \rightarrow \mathbb{R}$$

$$x \mapsto T(x) = x^2$$

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

$x \mapsto x^2$

$Y(\omega) = T(X(\omega))$

$$\{\omega \in \Omega; Y(\omega) \leq t\} =$$

$$\{\omega \in \Omega; X^2(\omega) \leq t\} = \begin{cases} \emptyset, & t < 0 \in \mathbb{F} \\ \{\omega; X(\omega) \in [-\sqrt{t}, \sqrt{t}]\} = X^{-1}([- \sqrt{t}, \sqrt{t}]) \in \mathbb{F} \end{cases}$$

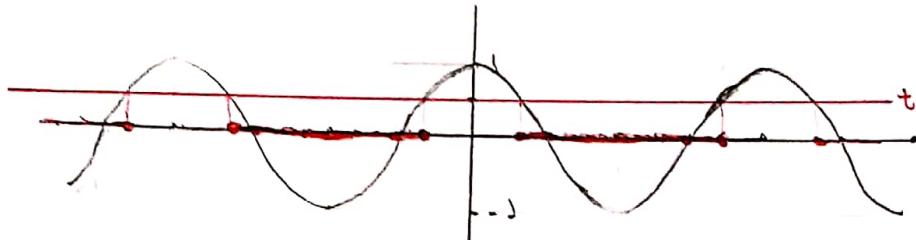
Logo, Y é v.a.

$$U = \cos(X)$$

$$\{w \in \Omega : U(w) \leq t\} =$$

$$\approx \{w \in \Omega : \cos(w) \leq t\} = \begin{cases} \emptyset, & t < -1 \\ \bigcup_{n \in \mathbb{Z}} \{w \in \Omega : X(w) \in I_n\} \text{cf} \\ \Omega, & t \geq 1 \text{ cf} \end{cases}$$

Logo, U é v.a.



—

$$(\Omega, \mathcal{F}, P) \xrightarrow{X} (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_x) \xrightarrow{T} (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), P_y)$$

$y = T \circ X$

$$X: \Omega \rightarrow \mathbb{R}^n$$

Se T é tal que $\forall B \in \mathcal{B}(\mathbb{R}^k)$

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^k$$

$T^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$ (T é $\mathcal{B}(\mathbb{R}^n)$ -mensurável)

$$Y: \Omega \rightarrow \mathbb{R}^k$$

Vale que, para todo $B \in \mathcal{B}(\mathbb{R}^k)$

$$Y = T \circ X$$

$$Y^{-1}(B) = (T \circ X)^{-1}(B) = X^{-1}(T^{-1}(B)) \in \mathcal{F} \in \mathcal{B}(\mathbb{R}^n)$$

$$X_1, X_2 \quad Y = \min\{X_1, X_2\}$$

$$W = X_1 + X_2$$

$$X_1, X_2$$

Logo, Y é v.a.

Nesse caso, as medidas induzidas por Y deve satisfazer:

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

$$P_Y(B) = P(Y^{-1}(B)) = P(X^{-1}(\underbrace{T^{-1}(B)}_A)) = P_X(T^{-1}(B))$$

↓

Ou seja, a distribuição da Y é "derivada" da distribuição de X .

CASO DISCRETO

$X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ discreto

Seja $A = \{(x_{1m}, x_{2m}, \dots, x_{nm}) : m \in \mathbb{N}\} \subseteq \mathbb{R}^n$ tal que $P_X(A) = 1$.

Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^k$ ($\mathcal{B}(\mathbb{R}^n)$ -mensurável)

Seja $T(A) = \{\underbrace{T(x)}_{\in \mathbb{R}^k} : x \in A\}$, $T(A)$ é enumerável (finito)

Seja $Y = T \circ X$

$$P_Y(T(A)) = P_X(T^{-1}(T(A))) \geq P_X(A) = 1$$

Para caracterizar P_Y , basta avaliar para todo $y \in T(A)$,

$$P_Y(\{y\}) = P_X(T^{-1}(\{y\})) = P(X \in T^{-1}(\{y\})) = \sum_{x \in A : T(x)=y} P(X=x)$$

Exemplo:

X Uniforme $\{0, 1, 2, 3, 4\}$

$$A = \{0, 1, 2, 3, 4\} \quad (P_X(A) = 1)$$

$T : \mathbb{R} \rightarrow \mathbb{R}$

$$x \in \mathbb{R} \rightarrow \text{sen} \frac{\pi x}{2}$$

$$T(A) = T(\{0, 1, 2, 3, 4\}) = \{0, 1, -1\}$$

Para cada $y \in T(A)$, determinamos $P_Y(\{y\}) = P(Y=y)$.

$$P_Y(\{-1\}) = P_X(T^{-1}(\{-1\})) = \sum_{x \in A : \text{sen} \frac{\pi x}{2} = -1} P(X=x) = P(X=3) = \frac{1}{5}$$

$$P_Y(\{0\}) = \sum_{x \in A : \text{sen} \frac{\pi x}{2} = 0} P(X=x) = \sum_{x \in \{0, 1, 4\}} P(X=x) = \frac{3}{5}$$

$$P_Y(\{1\}) = P(Y=1) = \frac{1}{5}$$

(2). $X = (X_1, X_2)$ tal que

X_1 e X_2 independentes, $X_i \sim \text{Poi}(\lambda_i)$, $\lambda_i > 0$, $i = 1, 2$.

$$A = \mathbb{N}^2 \rightarrow \text{conjunto enumerável} \quad P(X_1=x_1, X_2=x_2) = \frac{e^{-\lambda_1} \lambda_1^{x_1}}{x_1!} \cdot \frac{e^{-\lambda_2} \lambda_2^{x_2}}{x_2!} \quad \text{I.I. } (x_1, x_2) \in \mathbb{N}^2$$

$$T: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$(x_1, x_2) \mapsto T(x_1, x_2) = x_1 + x_2$$

$$T(A) = \mathbb{N} \quad Y = T(X)$$

Para $y \in \mathbb{N} = T(A)$, temos

$$P_Y(\{y\}) = P(Y=y) = \sum_{(x_1, x_2) \in A, x_1+x_2=y} P(X_1=x_1, X_2=x_2) =$$

$$= \sum_{(x_1, x_2) \in \{(0,y), (1,y-1), \dots, (y-1, 1), (y, 0)\}} P(X_1=x_1, X_2=x_2) = \sum_{j=0}^y P(X_1=j, X_2=y-j) =$$

$$= \sum_{j=0}^y \frac{e^{-\lambda_1} \lambda_1^j}{j!} \cdot \frac{e^{-\lambda_2} \lambda_2^{(y-j)}}{(y-j)!} \cdot \frac{y!}{y!} = \frac{e^{-(\lambda_1+\lambda_2)}}{y!} \underbrace{\sum_{j=0}^y \binom{y}{j} \lambda_1^j \lambda_2^{y-j}}_{(\lambda_1+\lambda_2)^y} =$$

$$\Rightarrow P(Y=y) = \frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1+\lambda_2)^y}{y!}$$

Logo, $Y \sim \text{Poi}(\lambda_1+\lambda_2)$.

Exemplo 3.

$X = (X_1, X_2, X_3)$ tal que

X_1, X_2 e X_3 são independentes,

$X_i \sim \text{Poi}(\lambda_i)$, $\lambda_i > 0$, $i = 1, 2, 3$.

$$A = \mathbb{N}^3$$

$$T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$$

$$(x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto T(x_1, x_2, x_3) = (x_1 + x_3, x_2 + x_3)$$

$$Y = (Y_1, Y_2) = T(X)$$

$$T(A) = \mathbb{N}^2$$

$$\text{Para } (y_1, y_2) \in \mathbb{N}^2,$$

$$P_Y(\{(y_1, y_2)\}) = P(Y = (y_1, y_2)) =$$

$$= \sum_{\substack{(x_1, x_2, x_3) \in \mathbb{N}^3 : \\ x_1 + x_3 = y_1 \\ x_2 + x_3 = y_2}} P(X_1 = x_1, X_2 = x_2, X_3 = x_3) =$$

$$= \sum_{\substack{(x_1, x_2, x_3) \in \mathbb{N}^3 \\ x_1 + x_3 = y_1 \\ x_2 + x_3 = y_2}} \frac{e^{-\lambda_1} \lambda_1^{x_1}}{x_1!} \frac{e^{-\lambda_2} \lambda_2^{x_2}}{x_2!} \frac{e^{-\lambda_3} \lambda_3^{x_3}}{x_3!}$$

$$= \sum_{j=0}^{\min\{y_1, y_2\}} \frac{e^{-\lambda_1} \lambda_1^{y_1-j}}{(y_1-j)!} \frac{e^{-\lambda_2} \lambda_2^{y_2-j}}{(y_2-j)!} \frac{e^{-\lambda_3} \lambda_3^j}{j!} \Rightarrow$$

$$\Rightarrow P(Y_1 = y_1, Y_2 = y_2) = e^{-(\lambda_1 + \lambda_2 + \lambda_3)} \sum_{j=0}^{\min\{y_1, y_2\}} \frac{\lambda_1^{y_1-j} \lambda_2^{y_2-j} \lambda_3^j}{(y_1-j)! (y_2-j)! j!}$$

$$(Y_1, Y_2) \sim \text{Holgate}(\lambda_1, \lambda_2, \lambda_3)$$

$$P(Y_1 = 0, Y_2 = 0) = e^{-(\lambda_1 + \lambda_2 + \lambda_3)} \neq P(Y_1 = 0) P(Y_2 = 0) = e^{-(\lambda_1 + \lambda_3)} e^{-(\lambda_2 + \lambda_3)}$$

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 $T: \mathbb{R} \rightarrow \mathbb{R}$ crescente (decreasing) e consideremos $Y = T(X)$

$$\stackrel{!}{=} T^{-1}: \mathbb{R} \rightarrow \mathbb{R}$$

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(T(X) \leq y) = P(T^{-1}(T(X)) \leq T^{-1}(y)) = \\ &= P(X \leq T^{-1}(y)) = F_X(T^{-1}(y)) \end{aligned}$$

Se T^{-1} é diferenciável, então

$$F'_X(T^{-1}(y)) = f_X(T^{-1}(y)) \frac{dT^{-1}(y)}{dy} = F'_Y(y) = f_Y(y)$$

• decaiente

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(T(X) \leq y) = P(T^{-1}(T(X)) \geq T^{-1}(y)) = \\ &= P(X \geq T^{-1}(y)) = 1 - F_X(T^{-1}(y)) \end{aligned}$$

• 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 decaiente

$$F_Y(y) = 1 - F_X(T^{-1}(y))$$

Se $T^{-1}(.)$ é diferenciável

$$f_Y(y) = -f_X(T^{-1}(y)) \frac{dT^{-1}(y)}{dy}$$

Exemplo 1:

$$X \sim N(\mu, \sigma^2)$$

$$T: \mathbb{R} \rightarrow \mathbb{R}$$

$$x \mapsto T(x) = \frac{x - \mu}{\sigma}$$

$$T^{-1}(y) = \sigma y + \mu$$

$$f_Y(y) = f_X(\sigma y + \mu) \cdot \sigma$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sigma y + \mu - \mu)^2}{2\sigma^2}} \cdot \sigma =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \therefore Y \sim N(0, 1)$$

Exemplo 2:

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

$$T: \mathbb{R}_+ \rightarrow \mathbb{R}$$

$$x \in \mathbb{R}_+ \rightarrow T(x) = -\frac{1}{\lambda} \log x$$

$$T^{-1}(y) = e^{-\lambda y}$$

$$Y = T(X)$$

$T(x)$ é decrescente

$$f_Y(y) = f_X(e^{-\lambda y}) (-\lambda e^{-\lambda y}) (1) =$$

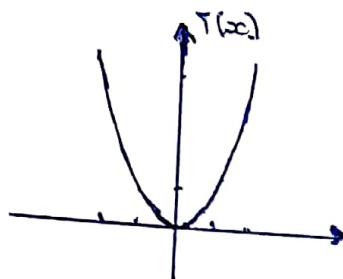
$$-\lambda e^{-\lambda y} \mathbb{I}_{(0,1)}(e^{-\lambda y})$$

$$\Rightarrow f_Y(y) = \lambda e^{-\lambda y} \mathbb{I}_{\mathbb{R}_+}(y)$$

Logo, $Y \sim \text{Exp}(\lambda)$.

(3) $X \sim N(0,1)$ $T: \mathbb{R} \rightarrow \mathbb{R}$

$x \in \mathbb{R} \rightarrow T(x) = x^2$



$Y = X^2$

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = \begin{cases} 0, & y < 0 \\ P(-\sqrt{y} \leq X \leq \sqrt{y}) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}), & y \geq 0 \end{cases}$$

Nesse caso, para $y \geq 0$

$$f_Y(y) = f_X(\sqrt{y}) \frac{1}{2\sqrt{y}} - f_X(-\sqrt{y}) \left(-\frac{1}{2\sqrt{y}} \right) \Rightarrow$$

$$f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}} = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{y})^2}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{y})^2}{2}}}{2\sqrt{y}} = \frac{\frac{1}{\sqrt{\pi}} y^{\frac{1}{2}-1} e^{-\frac{y}{2}}}{2\sqrt{y}}$$

Dai,

$$f_Y(y) = \begin{cases} 0, & y < 0 \\ \frac{(1/2)^{1/2}}{\sqrt{\pi}} y^{1/2-1} e^{-y/2}, & y \geq 0 \end{cases}$$

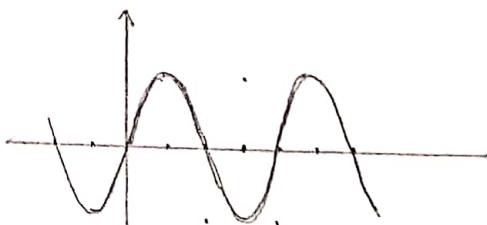
$\therefore Y \sim \text{Gamma}(1/2, 1/2) \sim \chi_1^2.$

(4) $X \sim U(0, 2\pi)$ $T: \mathbb{R} \rightarrow \mathbb{R}$

$x \in \mathbb{R} \mapsto T(x) = \sin x$

$$Y = T(X) = \sin X$$

$$P(\sin X \leq t)$$



Probabilidade e Inferência Estatística I

Aula 12 - 05 de Abril de 2013

Luiz Gustavo Esteves

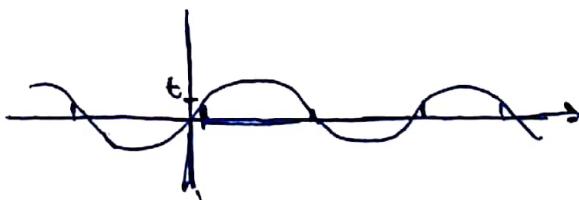
Transformação de V. A.

$$(2) X \sim U(0, 2\pi)$$

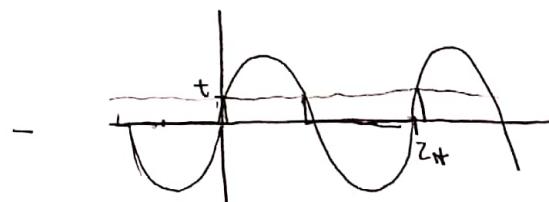
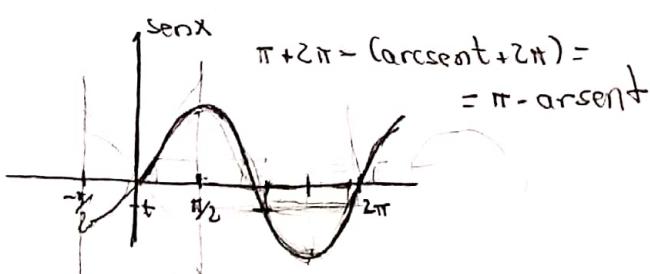
$$T: \mathbb{R} \mapsto \mathbb{R}$$

$$x \in \mathbb{R} \mapsto T(x) = \sin x$$

$$Y = T(X) = \sin X$$



$$F_Y(t) = P(Y \leq t) = P(\sin X \leq t) = P(\sin X \leq t, X \in (0, 2\pi)) = \begin{cases} 0, & t < -1 \\ P(X \in [\pi - \arcsent, \arcsent + 2\pi]), & -1 \leq t \leq 1 \\ 1, & t > 1 \end{cases}$$

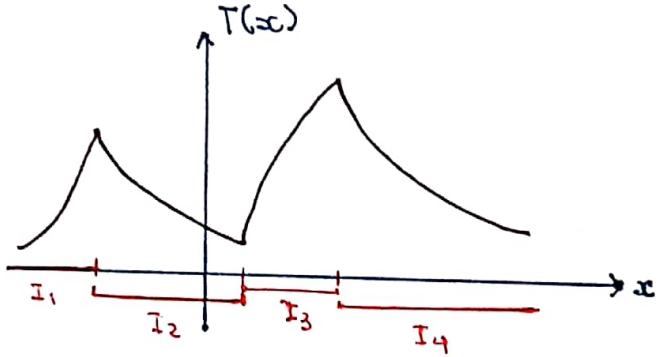


$$\Rightarrow F_Y(t) = \begin{cases} 0, & t \leq -1 \\ P(X \in [\pi - \arcsent, \arcsent + 2\pi]) = \frac{(2\pi + \arcsent) - (\pi - \arcsent)}{2\pi} = \frac{\pi + 2\arcsent}{2\pi}, & -1 < t \leq 0 \\ P(X \in [0, \arcsent] \cup [\pi - \arcsent, 2\pi]) = \frac{\arcsent + \pi + \arcsent}{2\pi}, & 0 < t \leq 1 \\ 1, & t > 1 \end{cases}$$

$$F_Y(t) = \begin{cases} 0, & t < -1 \\ \frac{\pi + 2 \arcsin t}{2\pi}, & -1 \leq t \leq 1 \\ 1, & t > 1 \end{cases}$$

$$f_Y(y) = \frac{1}{\pi \sqrt{1-y^2}} \mathbb{I}_{(-1,1)}(t)$$

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 $T_j : I_j \rightarrow B_j$

$$x \in I_j \Rightarrow T_j(x) = T(x)$$

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(T(X) \leq y) = P(T(X) \leq y, X \in A) = \\ &= P(T(X) \leq y, X \in \bigcup_{j=1}^r I_j) = P(X \in T^{-1}((-\infty, y]), X \in \bigcup_{j=1}^r I_j) = \\ &= P(X \in \bigcup_{j=1}^r I_j \cap T^{-1}((-\infty, y])) = \sum_{j=1}^r P(X \in I_j \cap T^{-1}((-\infty, y])) = \\ &= \sum_{j=1}^r \int_{I_j \cap T^{-1}((-\infty, y])} f_X(x) dx = \sum_{j=1}^r \int_{I_j \cap T^{-1}((-\infty, y])} f_X(x) dx = \sum_{j=1}^r \int_{I_j} \mathbb{I}_{T^{-1}((-\infty, y])}(x) f_X(x) dx = \end{aligned}$$

$$= \sum_{j=1}^r \int_{-\infty}^y \mathbb{I}_{B_j}(\omega) f_X(T_j^{-1}(\omega)) \left| \frac{d T_j^{-1}(\omega)}{d\omega} \right| d\omega \Rightarrow$$

$$\Rightarrow F_Y(y) = \int_{-\infty}^y \left\{ \sum_{j=1}^r \mathbb{I}_{B_j}(\omega) f_X(T_j^{-1}(\omega)) \left| \frac{d T_j^{-1}(\omega)}{d\omega} \right| \right\} d\omega$$

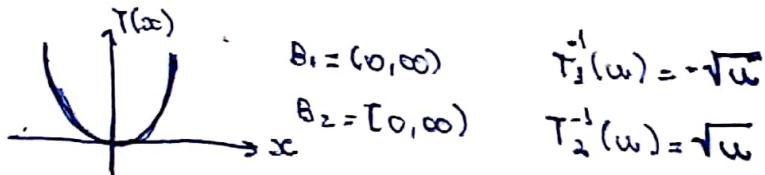
$\underbrace{\qquad\qquad\qquad}_{f_Y(\omega)}$

Voltando ao exemplo 3.

$$X \sim N(0,1)$$

$$T(\infty) = \infty$$

$$Y = T(X)$$



$$I_L = (-\infty, 0)$$

$$I_2 = [0, \infty)$$

$$f_Y(y) = \mathbb{I}_{R_+}(\omega) f_X(-\sqrt{\omega}) \left| \frac{d(-\sqrt{\omega})}{d\omega} \right| + \mathbb{I}_{R_+}(\omega) f_X(\sqrt{\omega}) \left| \frac{d\sqrt{\omega}}{d\omega} \right| =$$

$$= \mathbb{I}_{R_+}(\omega) \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{\omega})^2}{2}} \cdot \left| -\frac{1}{2\sqrt{\omega}} \right| + \mathbb{I}_{R_+}(\omega) \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{\omega})^2}{2}} \left| \frac{1}{2\sqrt{\omega}} \right| = \dots$$

Exemplo 2:

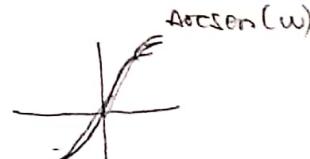
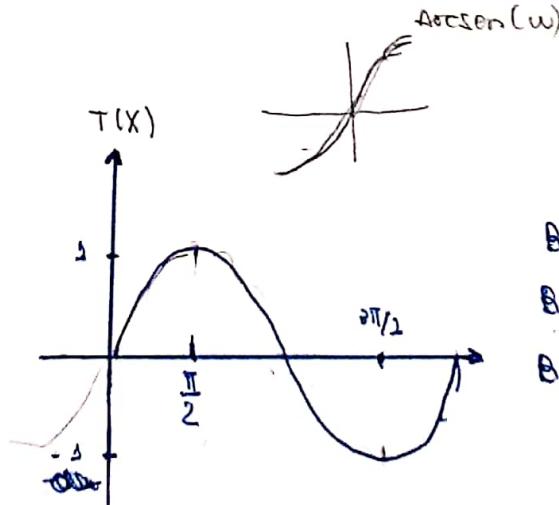
$$X \sim U(0, 2\pi)$$

$$T(X) = \sin X$$

$$I_1 = (0, \pi/2)$$

$$I_2 = (\pi/2, 3\pi/2)$$

$$I_3 = (3\pi/2, 2\pi)$$



$$B_1 = (0, \pi) \quad T_1^{-1}(w) = \arcsen w$$

$$B_2 = (-1, 1) \quad T_2^{-1}(w) = \pi - \arcsen w$$

$$B_3 = (-\pi, 0) \quad T_3^{-1}(w) = 2\pi + \arcsen w$$

Assim,

$$f_Y(y) = \mathbb{I}_{(0,1)}(y) \frac{1}{2\pi} \left| \frac{1}{\sqrt{1-y^2}} \right| + \mathbb{I}_{(-1,0)}(y) \frac{1}{2\pi} \left| -\frac{1}{\sqrt{1-y^2}} \right| + \mathbb{I}_{(-1,0)}(y) \frac{1}{2\pi} \left| \frac{1}{\sqrt{1-y^2}} \right| = \\ = \frac{1}{2\pi \sqrt{1-y^2}} 2\mathbb{I}_{(-1,0)}(y) = \frac{1}{\pi \sqrt{1-y^2}} \mathbb{I}_{(-1,0)}(y).$$

CASO MULTIVARIADO

$X = (X_1, \dots, X_K)$ absolutamente contínua com densidade f_X . e $\int_A f(x) dx = 1$.

$T: \mathbb{R}^K \rightarrow \mathbb{R}^k$ 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

$$T_j: I_j \rightarrow B_j, \quad j = 1, \dots, r \text{ com inversa} \\ x \in I_j \mapsto T_j(x) = T(x)$$

$T_j^{-1} = (\mu_1^{(j)}, \dots, \mu_K^{(j)})$ de modo que existam as derivadas parciais

$$\frac{\partial \mu_i^{(j)}}{\partial y_l} \text{ para todo } i, l = 1, \dots, K \text{ e } j = 1, \dots, r.$$

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

$$M^{(i)}(y_1, \dots, y_K) \text{ , onde } M_{i,k}^{(j)}(y_1, \dots, y_K) = \frac{d M_i^{(j)}}{dy_k}, i, k = 1, \dots, K$$

$\nwarrow k \times k$

Exemplo: $X = (X_1, X_2)$

$$X_i \sim \text{Gama}\left(\frac{\alpha_i}{2}, \frac{1}{2}\right) \sim \chi_{\alpha_i}, i=1,2, T: \mathbb{R}_+^2 \rightarrow \mathbb{R}_+^2$$

$$(x_1, x_2) \mapsto \left(\frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, x_2\right)$$

X_1 e X_2 independentes

$$\frac{X_1}{\alpha_1} \quad ?$$

$$\frac{X_2}{\alpha_2}$$

$$Y = T(X)$$

$$Y = (Y_1, Y_2) = \left(\frac{\alpha_1}{\alpha_2} \frac{X_1}{X_2}, X_2\right)$$

$$T^{-1}(y_1, y_2) = \left(\frac{\alpha_1}{\alpha_2} y_1 y_2, y_2\right)$$

$$J(y_1, y_2) = \begin{vmatrix} \frac{\alpha_1}{\alpha_2} y_2 & \frac{\alpha_1}{\alpha_2} y_1 \\ 0 & 1 \end{vmatrix} = \frac{\alpha_1}{\alpha_2} y_2.$$

$$T(\mathbb{R}_+^2) = \mathbb{R}_+^2$$

$$f_Y(y_1, y_2) = \frac{\alpha_1}{\alpha_2} J(y_2) \cdot f_X\left(\frac{\alpha_1}{\alpha_2} y_1 y_2, y_2\right) I_{\mathbb{R}_+^2}(y_1, y_2) =$$

$$= \frac{\alpha_1}{\alpha_2} y_2 \frac{\left(\frac{1}{2}\right)^{\frac{\alpha_1}{2}}}{\Gamma\left(\frac{\alpha_1}{2}\right)} \left(\frac{\alpha_1}{\alpha_2} y_1 y_2\right)^{\frac{\alpha_1}{2}-1} e^{-\frac{1}{2} \frac{\alpha_1}{\alpha_2} y_1 y_2} \frac{\left(\frac{1}{2}\right)^{\frac{\alpha_1}{2}}}{\Gamma\left(\frac{\alpha_2}{2}\right)} y_2^{\frac{\alpha_2}{2}-1} e^{-\frac{1}{2} y_2^2} I_{\mathbb{R}_+}(y_1) I_{\mathbb{R}_+}(y_2)$$

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_Y(y_1, y_2) dy_2 = \frac{\alpha_1}{\alpha_2} \frac{\left(\frac{1}{2}\right)^{\frac{\alpha_1+\alpha_2}{2}}}{\Gamma\left(\frac{\alpha_1}{2}\right)\Gamma\left(\frac{\alpha_2}{2}\right)} y_1^{\frac{\alpha_1}{2}-1} I_{\mathbb{R}_+}(y_1).$$

$$\cdot \int_{-\infty}^{\infty} y_2^{\frac{\alpha_1}{2}-1} e^{-y_2 \left\{ \frac{1}{2} \frac{\alpha_1}{\alpha_2} y_1 + \frac{1}{2} \right\}} dy_2 \Rightarrow$$

$$f_{Y_1}(y_1) = \left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{\alpha_1}{2}} \frac{\left(\frac{1}{2}\right)^{\frac{\alpha_1+\alpha_2}{2}}}{\Gamma\left(\frac{\alpha_1}{2}\right)\Gamma\left(\frac{\alpha_2}{2}\right)} y_1^{\frac{\alpha_1}{2}-1} \frac{\Gamma\left(\frac{\alpha_1+\alpha_2}{2}\right)}{\left[\frac{1}{2} + \frac{1}{2} \frac{\alpha_1}{\alpha_2} y_1\right]^{\frac{\alpha_1+\alpha_2}{2}}} I_{\Omega_1}(y_1) \quad Y_1 \sim F(\alpha_1, \alpha_2)$$

$$X \sim \text{Beta}\left(\frac{\alpha_1}{2}, \frac{\alpha_2}{2}\right)$$

$$\frac{X}{\frac{\alpha_1/2}{1-X}} \sim F(\alpha_1, \alpha_2)$$

Exemplo 2. (exercício)

$$X_1 \sim N(0,1)$$

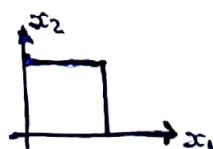
$$X_2 \sim \text{Gama}\left(\frac{n}{2}, \frac{1}{2}\right) \sim \chi^2_n$$

$$Y_1 = \frac{X_1}{\sqrt{\frac{X_2}{n}}} \sim t_n$$

Exemplo 3. X_1, X_2 independentes e identicamente distribuídas (v.a.i.i.d) segundo o modelo uniforme $(0,1)$.

$$Y_1 = X_1 + X_2$$

$$Y_2 = X_2$$



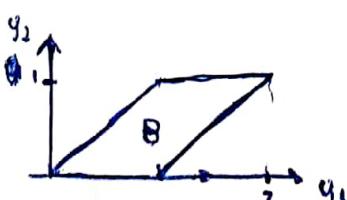
$$A = (0,1)^2$$

$$T(x_1, x_2) = (x_1 + x_2, x_2) \quad (\text{bijetora})$$

$$T^{-1}(y_1, y_2) = (y_1 - y_2, y_2)$$

$$u_1(y_1, y_2) \quad u_2(y_1, y_2)$$

$$B = T(A)$$



$$J(y_1, y_2) = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

$$f_Y(y) = 1 \cdot f_X(y_1 - y_2, y_2) I_{\mathbb{B}}(y_1, y_2) =$$

$$= f_{X_1}(y_1 - y_2) f_{X_2}(y_2) I_{\mathbb{B}}(y_1, y_2) =$$

$$= 1 \cdot 1 \cdot I_{\mathbb{B}}(y_1, y_2)$$

$$f_{Y_2}(y_2) = \int_{-\infty}^{\infty} f_Y(y_1, y_2) dy_1 =$$

$$= \int_{-\infty}^{\infty} I_{\mathbb{B}}(y_1, y_2) dy_1 = \begin{cases} \int_{-\infty}^{\infty} I(y_2) dy_1, & 0 \leq y_2 \leq 1 \\ \int_{-\infty}^{\infty} I(y_2) dy_1, & 1 \leq y_2 \leq 2 \end{cases} \Rightarrow$$

$$\Rightarrow f_{Y_2}(y_2) = \begin{cases} y_2 - 0 = y_2, & 0 \leq y_2 \leq 1 \\ 1 - (y_2 - 1) = 2 - y_2, & 1 \leq y_2 \leq 2 \end{cases}$$

$$I_{(0,1)}(y_1 - y_2) I_{(0,1)}(y_2) = 1$$

⇒

$$0 \leq y_1 - y_2 \leq 1 \quad -y_1 \leq -y_2 \leq 1 - y_1$$

$$0 \leq y_2 \leq 1 \quad 1 - y_1 \leq y_2 \leq y_1$$

$$\begin{array}{l} \downarrow \\ y_2 \leq y_1 \\ y_2 \geq 1 - y_1 \end{array}$$

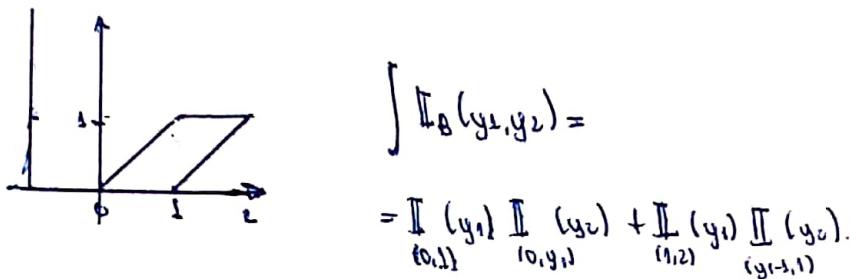
$$\max\{0, 1-y_1\} \leq y_2 \leq \min\{y_1, 1\}$$

Probabilidade e Inferência Estatística I

Aula 13 - 09/04/13

Luis Gustavo Esteves

Aula passada



— n — n — n —

Estatísticas de Ordem

$X = (X_1, \dots, X_n)$ absolutamente contínua com densidade de f_X .

$X_{(j)}$ a j-ésima menor componente do vetor X

$$X_1(\omega) = \min\{X_1(\omega), \dots, X_n(\omega)\}$$

$$X_2(\omega) = \min\{X_1(\omega), \dots, X_n(\omega)\} \setminus \{X_{(1)}(\omega)\}$$

:

$$X_n(\omega) = \max\{X_1(\omega), \dots, X_n(\omega)\}$$

$$T: \mathbb{R}^k \rightarrow \mathbb{R}^k$$

$$(x_1, \dots, x_n) \in \mathbb{R}^n \rightarrow T(x_1, \dots, x_n) = (x_{(1)}, x_{(2)}, \dots, x_{(n)})$$

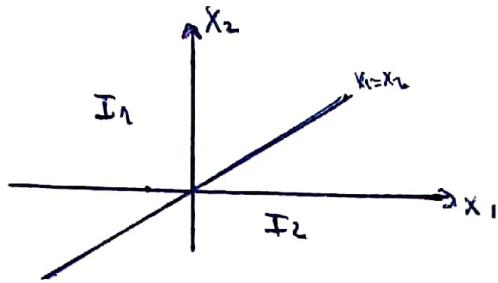
$$Y = T(X)$$

$$k=2$$

$$(1, 3) \rightarrow T(1, 3) = (1, 3)$$

$$(3, 1) \rightarrow T(3, 1) = (1, 3)$$

k=2



$$I_1 = \{(a, b) \in \mathbb{R}^2 : a < b\}, T(I_1) = I_A = B = B_1$$

$$I_2 = \{(a, b) \in \mathbb{R}^2 : a > b\}, T(I_2) = I_B = B = B_2$$

$$T_1 : I_1 \rightarrow I_1$$

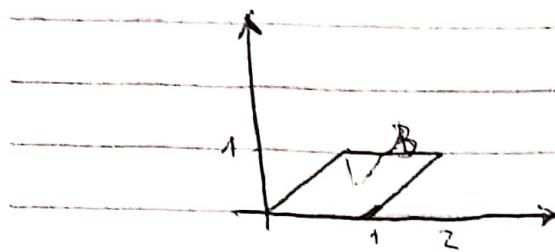
$$(x_1, x_2) \mapsto T_1(x_1, x_2) = T_1(x_1, x_2) = (x_1, x_2)$$



$$T_1^{-1}(y_1, y_2) = (y_1, y_2)$$

Aula Passada

09/10/4



$$\begin{aligned} II_B(y_1, y_2) &= \\ &= II(y_1) II(y_2) + II(y_1) II(y_2) \end{aligned}$$

$(y_1, 1)$ $(0, y_2)$ $(1, 2)$ $(y_1 - 1, y_2)$

Estatísticas de Ordem

$X = (X_1, \dots, X_k)$ absolutamente contínuo c/ dens. de f_X

$X_{(j)}$ a j-ésima menor componente de vetor X

$$X_{(n)}(\omega) = \min \{X_1(\omega), X_2(\omega), \dots, X_n(\omega)\}$$

$$X_{(1)}(\omega) = \min \{X_1(\omega), \dots, X_n(\omega)\} \setminus \{X_{(n)}(\omega)\}$$

$$X_{(n)}(\omega) = \max \{X_1(\omega), \dots, X_n(\omega)\}$$

$$T: \mathbb{R}^k \rightarrow \mathbb{R}^k$$

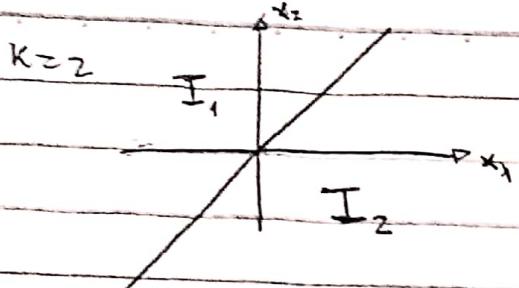
$$(x_1, \dots, x_k) \in \mathbb{R}_K \mapsto T(x_1, \dots, x_k) = (x_{(1)}, x_{(2)}, \dots, x_{(n)})$$

$$Y = T(X)$$

$$k=2$$

$$(1, 3) \rightarrow T(1, 3) = (1, 3)$$

$$(3, 1) \rightarrow T(3, 1) = (1, 3)$$



$$I_1 = \{(a, b) \in \mathbb{R}^2 : a < b\} \quad T(I_1) = I_1 = B = B_1$$

$$I_2 = \{(a, b) \in \mathbb{R}^2 : a > b\} \quad T(I_2) = I_2 = B = B_2$$

$$T_1 : I_1 \rightarrow I_1$$

$$(x_1, x_2) \mapsto T_1(x_1, x_2) = T(x_1, x_2) = (x_1, x_2)$$



$$T_1^{-1}(y_1, y_2) = (y_1, y_2)$$

$$u_1^{(1)}(y_1, y_2) = y_1$$

$$u_2^{(1)}(y_1, y_2) = y_2$$

$$T_2 : I_2 \rightarrow I_1$$

$$(x_1, x_2) \mapsto T_2(x_1, x_2) = T(x_1, x_2) = (x_2, x_1)$$

$$T_1^{-1}(y_1, y_2) = (y_2, y_1)$$

$$u_1^{(2)}(y_1, y_2) = y_2 \rightarrow u_2^{(2)}(y_1, y_2) = y_1$$

$$f_Y(y_1, y_2) = f_X(y_1, y_2) \left| J_{I_1}^{(1)}(y_1, y_2) \right| \underset{I_1}{\mathbb{I}}(y_1, y_2) + f_X(y_2, y_1) \left| J_{I_1}^{(2)}(y_1, y_2) \right| \underset{I_1}{\mathbb{I}}(y_1, y_2)$$



$$J^{(1)}(y_1, y_2) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \quad \left(f_Y(y_1, y_2) = \underset{I_1}{\mathbb{I}}(y_1, y_2) \left\{ f_X(y_1, y_2) + f_X(y_2, y_1) \right\} \right)$$

$$J^{(2)}(y_1, y_2) = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

$$\begin{matrix} \textcircled{1} & \textcircled{2} & z! + 1 \\ \diagup & \diagdown & \\ \textcircled{3} & & \end{matrix}$$

Para $k=3 \quad x_1 < x_2 < x_3$

$$x_1 < x_3 < x_2$$

6 parcelas

(/ /)

Em geral, particionamos \mathbb{R}^k em $k! + 1$ regiões

Para cada permutação $\pi: \{1, \dots, k\} \rightarrow \{1, \dots, k\}$,
 Considere a região de \mathbb{R}^k

$$x_{\pi(1)} < x_{\pi(2)} < \dots < x_{\pi(k)}$$

Assim,

$$f_Y(y_1, \dots, y_k) = \sum_{\substack{\pi \in \Pi \\ \text{permutação de } \{1, \dots, k\}}} f_X(y_{\pi(1)}, \dots, y_{\pi(k)}) \prod_B I_B(y_1, \dots, y_k)$$

$$B = \{y_1, y_2, \dots, y_k\} \in \mathbb{R}^k : y_1 < y_2 < \dots < y_k\}$$

Se X_1, \dots, X_k são v.a.i.i.d com densidade f_{X_i} , vale que

$$f_Y(y_1, \dots, y_k) = k! f_{X_1}(y_1) \dots f_{X_k}(y_k) \prod_B I_B(y_1, \dots, y_k)$$

Ex: X_1, X_2, \dots, X_k v.a.i.i.d $\text{Exp}(\lambda)$

$$Y = \left(\begin{smallmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{smallmatrix}, X_{(1)}, X_{(2)}, \dots, X_{(k)} \right)$$

$$f_Y(y_1, \dots, y_k) = k! \prod_{i=1}^k \left\{ \lambda e^{-\lambda y_i} I_{\mathbb{R}_+}(y_i) \right\} \prod_B I_B(y_1, \dots, y_k) =$$

$$= k! \lambda^k e^{-\lambda \sum_{i=1}^k y_i} \prod_B I_B(y_1, \dots, y_k)$$

$$B = \{y_1, y_2, \dots, y_k\} \in \mathbb{R}^k : 0 < y_1 < y_2 < \dots < y_k\}$$

(X_1, \dots, X_K) vetor aleatório com X_1, \dots, X_K independentes

$$F_{X_{(1)}}(t) = P(X_{(1)} \leq t) =$$

$$= 1 - P(X_{(1)} > t) = 1 - P\left(\bigcap_{i=1}^K (X_i > t)\right) =$$

$$= 1 - \prod_{i=1}^K P(X_i > t) = 1 - \prod_{i=1}^K [1 - F_{X_i}(t)]$$

$$F_{X_{(m)}}(t) = P(X_{(m)} \leq t) = P\left(\bigcap_{i=1}^m (X_i \leq t)\right) = \prod_{i=1}^m F_{X_i}(t).$$

Exemplo:

(1) X_1, \dots, X_m v.a. independentes $X_i \sim \text{Exp}(\lambda_i), i=1, \dots, K$

$$\lambda_i > 0, i=1, \dots, K$$

$$F_{X_{(1)}}(t) = 1 - \prod_{i=1}^K \left[1 - \underbrace{\left(1 - e^{-\lambda_i t}\right)}_{e^{-t \sum_{j=1}^i \lambda_j}}\right] = 1 - e^{-\left(\sum_{i=1}^K \lambda_i\right)t}$$

$$\text{Logo, } X_{(1)} \sim \text{Exp}\left(\sum_{i=1}^K \lambda_i\right)$$

(2) X_1, \dots, X_K v.a.i.i.d $\sim U(0, 1)$

$$F_{X_{(m)}}(t) = \prod_{i=1}^K F_{X_i}(t) = \left[F_{X_1}(t) \right]^n = \begin{cases} 0, & t \leq 0 \\ t^n, & 0 < t \leq 1 \\ 1, & t > 1 \end{cases}$$

$$= n(1-t)^{n-1} \Big|_{t=1}$$

$$f_{X_{(m)}}(t) = \begin{cases} nt^{n-1}, & 0 < t \leq 1 \\ 0, & \text{c.c.} \end{cases} \quad X_{(m)} \sim \text{Beta}_n(n, 1)$$

$$X_{(1)} \sim \text{Beta}_n(1, n)$$

Exercício

Transformações Lineares

(X_1, \dots, X_n) absolutamente contínua com dens. f_X

$$Y_i = \sum_{j=1}^k a_{ij} X_j, \quad i = 1, \dots, k$$

(Y_1, \dots, Y_k) vetor aleatório de modo que $\mathbf{Y}' = A \cdot \mathbf{X}'$,

$$A = (a_{ij})_{k \times k}$$

Se A é invertível, com inversa A^{-1} , podemos escrever

$$\mathbf{X}' = A^{-1} \cdot \mathbf{Y}'$$

Seja $A^{-1} = (c_{ij})_{k \times k}$

Assim

$$X_i = \sum_{j=1}^k c_{ij} Y_j, \quad i = 1, \dots, k$$

$|$

$u_i(y_1, \dots, y_k)$

$$\begin{vmatrix} c_{11} & c_{12} & \dots & c_{1k} \\ & & & \\ c_{k1} & c_{k2} & \dots & c_{kk} \end{vmatrix}$$

$$\frac{dx_i}{dy_j} = \frac{d}{dy_j} \sum_{j=1}^k c_{ij} y_j = c_{ij} \quad J(y_1, \dots, y_k) = \det A^{-1} = \frac{1}{\det A}$$

$$f_Y(y_1, \dots, y_k) = \frac{1}{|\det A|} f_X \left(\sum_{j=1}^k c_{1j} y_j, \sum_{j=1}^k c_{2j} y_j, \dots, \sum_{j=1}^k c_{kj} y_j \right)$$

Se A é ortogonal (*isto é*, $A^{-1} = A^\top$), então

$$\det A = \pm 1 \quad \text{e} \quad c_{ij} = a_{ji}$$

$$\left(\begin{array}{c} \det A = \det A^T = \pm 1 \\ \det A \\ \det A^T \end{array} \right)$$

Portanto

$$f_Y(y_1, \dots, y_k) = f_X\left(\sum_{j=1}^k a_{j1}y_j, \sum_{j=1}^k a_{j2}y_j, \dots, \sum_{j=1}^k a_{jk}y_j\right)$$

Transformação Linear Ortogonal

Resultado $X = (X_1, \dots, X_k)$ com X_1, \dots, X_k r.a.i.i.d $\mathcal{N}(0, 1)$.

Seja $A_{k \times k}$ ortogonal e $Y = (Y_1, \dots, Y_k)$, $Y = A \cdot X$.

$$\begin{aligned} f_Y(y_1, \dots, y_k) &= f_X\left(\sum_{j=1}^k a_{j1}y_j, \dots, \sum_{j=1}^k a_{jk}y_j\right) = \\ &= -\frac{\left(\sum_{j=1}^k a_{j1}y_j\right)^2}{2} - \frac{\left(\sum_{j=1}^k a_{j2}y_j\right)^2}{2} - \frac{\left(\sum_{j=1}^k a_{jk}y_j\right)^2}{2} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{\sum_{j=1}^k \left(\sum_{i=1}^k a_{ij}y_i\right)^2}{2}} \Rightarrow \end{aligned}$$

$$\Rightarrow f_Y(y_1, \dots, y_k) = \left(\frac{1}{\sqrt{2\pi}}\right)^k \cdot e^{-\sum_{i=1}^k \left(\sum_{j=1}^k a_{ij}y_j\right)^2 / 2} =$$

$$\left. \begin{aligned} \sum_{i=1}^k b_i^2 &= b \cdot b \\ y \cdot A & \end{aligned} \right\} = \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-\frac{(y \cdot A)(A^T \cdot y)}{2}} =$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-\frac{(y \cdot A)(A^T \cdot y)}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-\frac{y \cdot A^T \cdot y}{2}}$$

11

$$\Rightarrow f_Y(y_1, \dots, y_k) = \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-\sum_{i=1}^k y_i^2/2}$$

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 r.a. i.i.d $N(0,1)$

$$\frac{X_1 + X_2}{\sqrt{2}} \text{ e } \frac{X_1 - X_2}{\sqrt{2}} \text{ são ind. } N(0,1)$$

Comentário

Se $A_{n \times k}$ é ortogonal, $X_{1:n}$ e $\mathbf{y}' = A \cdot \mathbf{x}'$

$$\text{então } \sum_{i=1}^k y_i^2 = \sum_{i=1}^k x_i^2$$

$$\sum y_i^2 = \mathbf{y}' \cdot \mathbf{y}' = (\mathbf{x}' \cdot A') (A \cdot \mathbf{x}') = \mathbf{x}' \cdot (A' A) \mathbf{x}' = \mathbf{x}' \cdot \mathbf{x}' = \sum_{i=1}^k x_i^2$$

Resultado

X_1, X_2, \dots, X_n são ind. $\left| \Rightarrow g_1(x_1), g_2(x_2), \dots, g_n(x_n) \right.$
 $g_i: \mathbb{R} \rightarrow \mathbb{R}$ mensuráveis $\left. \right|$ são ind.

11

$$Y = (g_1(x_1), \dots, g_k(x_k))$$

$$\begin{aligned} F_Y(t_1, \dots, t_k) &= P(g_1(x_1) \leq t_1, \dots, g_k(x_k) \leq t_k) = P(x_i \in g_i^{-1}(-\infty, t_i]), \\ \dots, x_k \in g_k^{-1}(-\infty, t_k]) &= P\left(\bigcap_{i=1}^k (x_i \in g_i^{-1}(-\infty, t_i])\right) = \\ &= \prod_{i=1}^k P(x_i \in g_i^{-1}(-\infty, t_i]) = \prod_{i=1}^k P(g_i(x_i) \in (-\infty, t_i]) = \\ &= \prod_{i=1}^k F_{g_i(x_i)}(t_i) = \prod_{i=1}^k F_{Y_i}(t_i) \end{aligned}$$

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.

$$\begin{array}{ll} g_1: \mathbb{R}^2 \rightarrow \mathbb{R} & g_1(x_1, x_4), g_2(x_5), g_3(x_2, x_3) \\ g_2: \mathbb{R} \rightarrow \mathbb{R} & \\ g_3: \mathbb{R}^2 \rightarrow \mathbb{R} & x_1, x_4 \text{ arctg } x_5 \text{ e } x_2/x_3+1 \end{array}$$

são ind.

X_1, X_2 iid $N(0, 1)$

$$\frac{\overbrace{X_1 + X_2}^{Y_1}}{\sqrt{2}} \quad \frac{\overbrace{X_1 - X_2}^{Y_2}}{\sqrt{2}} \quad \text{i.i.d } N(0, 1)$$

$$\sqrt{2} \bar{X}_2$$

$$Y_1^2 + Y_2^2 = X_1^2 + X_2^2$$

$$Y_2^2 = X_1^2 + X_2^2 - Y_1^2 = X_1^2 + X_2^2 - 2X_1 X_2$$

$$\sum_{i=1}^n (X_i - \bar{X})^2$$

11

No caso k-variável

$$A = \begin{pmatrix} 1/\sqrt{k} & 1/\sqrt{k} & \dots & 1/\sqrt{k} \end{pmatrix} \quad D_{k, \text{Groot}}$$

$$Y_k = \sqrt{k} \cdot \bar{X}_k$$

$$Y_1^2 + Y_2^2 + \dots + Y_k^2 = X_1^2 + X_2^2 + \dots + X_k^2 - \bar{X}_k^2$$

$$Y_1^2 + \dots + Y_k^2 = \sum_{i=1}^k X_i^2 - k \bar{X}_k^2 =$$

$$= \frac{1}{k} \sum_{i=1}^k (X_i - \bar{X}_k)^2$$

$$X_1, \dots, X_k \text{ i.i.d } N(\mu, \sigma^2)$$

$$\frac{X_1 - \mu}{\sigma}, \frac{X_2 - \mu}{\sigma}, \dots, \frac{X_k - \mu}{\sigma} \text{ são i.i.d } N(0, 1)$$

$$\sqrt{k} \left(\frac{\bar{X}_k - \mu}{\sigma} \right) = \sqrt{k} \sum_{i=1}^k \left(\frac{X_i - \mu}{\sigma}, \frac{\bar{X}_k - \mu}{\sigma} \right)^2 = \sum_{i=1}^k (X_i - \bar{X}_k)^2$$

$$\bar{X}_k$$

$$S_k^2$$

Normal Multivariada

X_1, X_2, \dots, X_k v.a.i.i.d $N(0, 1)$

$$g_i(x_i) = \sigma_i x_i + \mu_i, \sigma_i > 0, \mu_i \in \mathbb{R}, i=1, \dots, k$$

$$Y_i = g_i(x_i)$$

(Y_1, \dots, Y_k) de Normais independentes

$$Y' = A \cdot X' + \mu', \quad A = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix}$$

$$\mu = (\mu_1, \mu_2, \dots, \mu_k)$$

A mais geral:

$$Y' = A X' + \mu' \Rightarrow Y' - \mu' = A \cdot X'$$

$$X' = A^{-1}(Y' - \mu') ; \quad A^{-1} = (c_{ij})_{k \times k}$$

$$\mu_i(y_1, \dots, y_k) \quad X_i = \sum_{j=1}^k c_{ij} (y_j - \mu_j)$$

$$\frac{dX_i}{dy_j} = \frac{d}{dy_j} \sum_{l=1}^k c_{il} (y_l - \mu_l) = c_{ij}$$

/ /

$$\text{Pelo resultado, } f_Y(y_1, \dots, y_k) = \frac{1}{|\det A|} f_X\left(\sum_{j=1}^k c_{1j}(y_j - \mu_j), \sum_{j=1}^k c_{2j}(y_j - \mu_j), \dots, \sum_{j=1}^k c_{nj}(y_j - \mu_j)\right)$$

$$= \frac{1}{|\det A|} \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-\frac{1}{2} \sum_{i=1}^k \left\{ \sum_{j=1}^k c_{ij}(y_j - \mu_j) \right\}^2}$$

$$= \frac{1}{|\det A|} \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-\frac{1}{2} (y - \mu)(A^{-1})(y - \mu)} =$$

$$= \frac{1}{|\det A|} \left(\frac{1}{\sqrt{2\pi}}\right)^k \cdot e^{-\frac{1}{2} (y - \mu)(A')^{-1} A^{-1} (y - \mu)}$$

$$\Rightarrow f_Y(y) = \frac{1}{|\det A|} \left(\frac{1}{\sqrt{2\pi}}\right)^k e^{-\frac{1}{2} (y - \mu)(A \cdot A')^{-1} (y - \mu)}$$

$$\boxed{A \cdot A' = \Sigma} \stackrel{\text{Binat}}{\Rightarrow} \det A \cdot \det A' = \det \Sigma$$

$$f_Y(y) = \frac{1}{\sqrt{\det \Sigma}} \left(\frac{1}{\sqrt{2\pi}}\right)^k \cdot e^{-\frac{1}{2} (y - \mu)' \Sigma^{-1} (y - \mu)}$$

$$Y_i = \sum_{j=1}^k a_{ij} X_j + \mu_i \sim N\left(\mu_i, \sum_{j=1}^k a_{ij}^2\right)$$

$$Y \sim N_k(\mu, \Sigma)$$

Obse Transt. de um vetor discreto resulta
noutro vetor aleatório discreto.

Já transformação de v.a abs. contínua nem
sempre resulta numa contínua

Ex: $X \sim \text{Exp}(\lambda)$

$Y = \lceil X \rceil$, maior inteiro maior ou igual a X

Note que $P(Y \in \mathbb{N}) = 1$

Para $y \in \mathbb{N}$

$$\begin{aligned} P(Y=y) &= P(y \leq X < y+1) \\ &= F_X(y+1) - F_X(y) = (1 - e^{-\lambda(y+1)}) - (1 - e^{-\lambda y}) = \\ &= e^{-\lambda y} - e^{-\lambda(y+1)} \Rightarrow P(Y=y) = e^{-\lambda y}(1 - e^{-\lambda}) \end{aligned}$$

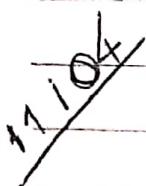
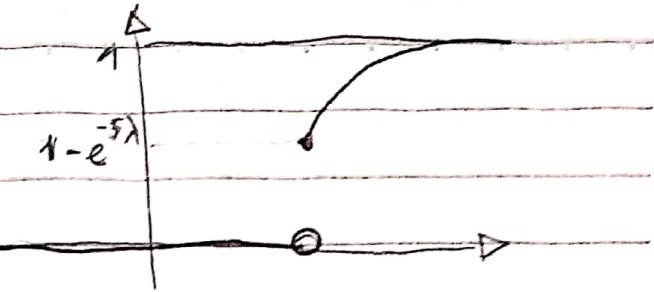
$$Y \sim \text{Geo}_{(1-e^{-\lambda})}$$

(2) $V = \max\{X, 5\}$

$$F_V(v) = P(V \leq v) = \begin{cases} 0, & v < 5 \\ \dots & \end{cases}$$

$$P(X \leq 5), V \geq 5$$

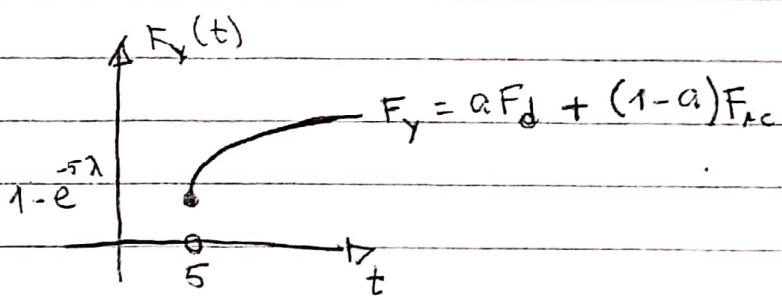
$$= \begin{cases} 0, & v < 5 \\ 1 - e^{-\lambda v}, & v \geq 5 \end{cases}$$



Esperança Matemática $X: \Omega \rightarrow \mathbb{R}$

$$X \text{ discreto} \rightarrow E(X) = \sum_x x P(X=x)$$

$$X \text{ abs contínuo} \rightarrow E(X) = \int_{\mathbb{R}} x f_X(x) dx$$



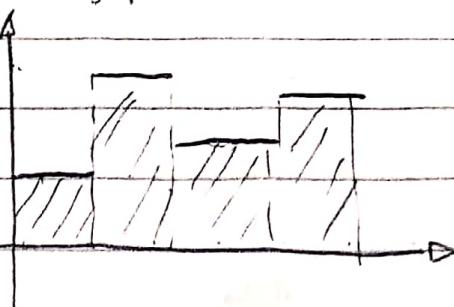
$$\Omega = [0, 1]$$

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

\bar{P} : Complemento

$$X(\omega) = \begin{cases} X_1, & 0 \leq \omega \leq K_1 \\ X_2, & K_1 < \omega \leq K_2 \\ X_3, & K_2 < \omega \leq K_3 \\ X_4, & K_3 < \omega \leq 1 \end{cases}$$

$$E(X) = \sum_{j=1}^4 X_j P(X=X_j) = \sum_{j=1}^4 X_j (K_j - K_{j-1})$$



Probabilidade e Inferência Estatística

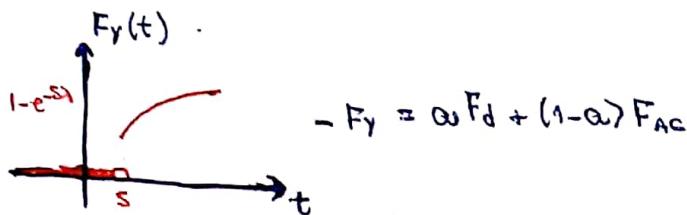
11/04/2013 - Aula 13

Luis Gustavo Esteves

Esperança Matemática

$$X \text{ discreta} \rightarrow E(X) = \sum_{\infty} \infty P(X=\infty)$$

$$X \text{ abs. contínua} \rightarrow E(X) = \int_{\mathbb{R}} \infty f_X(\infty) d\infty$$

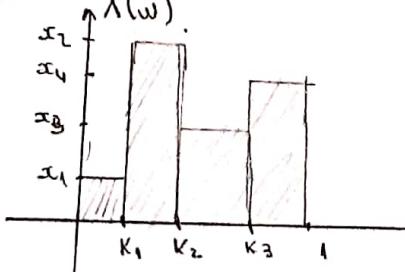


$$\Omega = [0,1] \quad P: \text{comprimento}$$

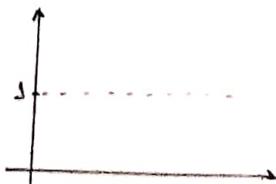
$$X: \Omega \rightarrow \mathbb{R}$$

$$X(\omega) = \begin{cases} x_1, & 0 \leq \omega \leq k_1 \\ x_2, & k_1 \leq \omega \leq k_2 \\ x_3, & k_2 \leq \omega \leq k_3 \\ x_4, & k_3 \leq \omega \leq 1 = k_4 \end{cases}$$

$$E(X) = \sum_{j=1}^4 x_j P(X=x_j) = \sum_{j=1}^4 x_j (k_j - k_{j-1})$$



$$Y(\omega) = \begin{cases} 1, & \omega \in Q \\ 0, & \omega \notin Q \end{cases}$$



$$\begin{aligned} E(Y) &= 1 \cdot \Phi(Y=1) + 0 \cdot \Phi(Y=0) \\ &= 1 \cdot 0 + 0 \cdot 1 = 0 \end{aligned}$$

(Ω, \mathcal{F}, P)

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

$\int_{\Omega} X dP \rightarrow$ Integral de Lebesgue de X contra P .

Definição: Uma função $X: \Omega \rightarrow \mathbb{R}$ é chamada SIMPLES se assume apenas um número finito de valores ($X(\Omega) = \{X(\omega) : \omega \in \Omega\}$ é finito).

Em geral $X: \Omega \rightarrow \mathbb{R}$ simples pode ser representada da seguinte maneira:

$$X = \sum_{j=1}^k a_j \mathbb{1}_{A_j}, \text{ com } X = \sum_{j=1}^k a_j \mathbb{1}_{A_j} = a_1 \mathbb{1}_{A_1} + a_2 \mathbb{1}_{A_2} + a_3 \mathbb{1}_{A_3} + \dots + a_k \mathbb{1}_{A_k}$$

$A_1, A_2, \dots, A_k \in \mathcal{F}$ disjuntos e tais que $\Omega = A_1 \cup A_2 \cup \dots \cup A_k$.

Definição: Seja $X: \Omega \rightarrow \mathbb{R}$ simples, isto é, X é da forma $X = \sum_{j=1}^k a_j \mathbb{1}_{A_j}$, $a_j \geq 0, j=1, \dots, k$ e A_1, \dots, A_k partição de Ω .

A integral de X em relação (contra) P é definida por

$$\int_{\Omega} X dP = \sum_{j=1}^k a_j P(A_j).$$

Existem outras formas de definir a soma acima:



$$X(\omega) = \sum_{i=1}^n \sum_{j=1}^n c_{ij} \mathbb{1}_{B_i \cap A_j} \quad c_{ij} = b_i = a_j$$

$$\int_{\Omega} X dP = \sum_{i=1}^n \sum_{j=1}^n a_j P(B_i \cap A_j).$$

$$\int_{\Omega} X dP = \sum_{i=1}^n b_i \sum_{j=1}^n P(B_i \cap A_j) = \sum_{i=1}^n b_i P(B_i)$$

$$\int_{\Omega} X dP = \sum_{i=1}^n \sum_{j=1}^k a_j P(B_i \cap A_{ij}) = \sum_{j=1}^k a_j \sum_{i=1}^n P(B_i \cap A_{ij}) = \sum_{j=1}^k a_j P(A_{ij}).$$

Propriedades

(1) $\int_{\Omega} X dP \geq 0$

(2) $c > 0$, $\int_{\Omega} (cX) dP = \sum_{j=1}^k (ca_j) P(A_{ij}) = c \sum_{j=1}^k a_j P(A_{ij}) = c \int_{\Omega} X dP.$

~~3~~ $cX = \sum_{j=1}^k (ca_j) \mathbb{I}_{A_{ij}}$

(3) X, Y simples

$$\int_{\Omega} (X+Y) dP$$

$$X+Y = \sum_{i=1}^n \sum_{j=1}^k (b_i + a_j) \mathbb{I}_{A_{ij} \cap B_i}$$

$$X = \sum_{j=1}^k a_j \mathbb{I}_{A_{ij}}$$

$$\int_{\Omega} (X+Y) dP = \sum_{i=1}^n \sum_{j=1}^k (b_i + a_j) P(A_{ij} \cap B_i) =$$

$$Y = \sum_{i=1}^n b_i \mathbb{I}_{B_i}$$

$$= \sum_{i=1}^n \sum_{j=1}^k b_i P(A_{ij} \cap B_i) + \sum_{i=1}^n \sum_{j=1}^k a_j P(A_{ij} \cap B_i)$$

$$= \sum_{i=1}^n b_i P(B_i) + \sum_{j=1}^k a_j P(A_j) =$$

$$= \int_{\Omega} Y dP + \int_{\Omega} X dP$$

Funções Negativas

Seja $X: \Omega \rightarrow \mathbb{R}_+$

Seja $\mathcal{L}_X = \{ f: \Omega \rightarrow \mathbb{R}_+ : f \text{ é simples e } f \leq X \}$

Note que $\mathcal{L}_X \neq \emptyset$, pois $X(w) = 0, \forall w \in \Omega$ é elemento de \mathcal{L}_X .

* ponder cada valor da v.a. pela medida do seu conjunto

Definição: Para $X: \Omega \rightarrow \mathbb{R}_+$, definimos

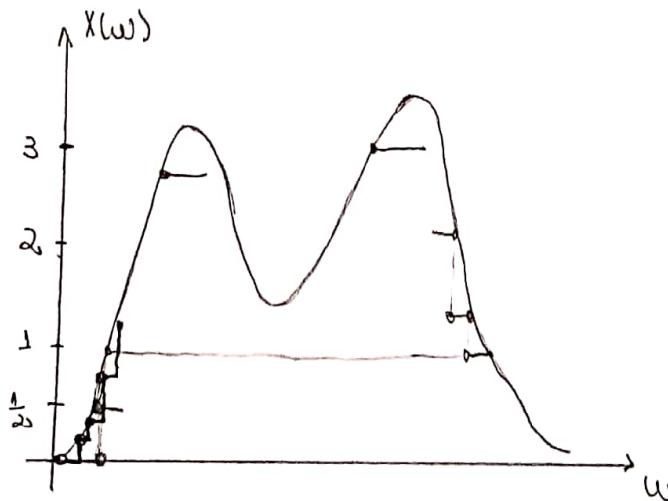
$$\int_{\Omega} X dP = \sup \left\{ \int_{\Omega} f dP : f \in \mathcal{L}_X \right\}$$

Alternativamente, podemos definir $\int_{\Omega} X dP$ da seguinte maneira:

Seja $X_n: \Omega \rightarrow \mathbb{R}_+$ 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

$$\int_{\Omega} X dP = \lim_{n \rightarrow \infty} \int_{\Omega} X_n dP.$$

Mostrar que $X \leq Y$, então $\int_{\Omega} X dP \leq \int_{\Omega} Y dP$.



$$X_1(\omega) = \begin{cases} 0, & X(\omega) < 1/2 \\ 1/2, & X(\omega) \in [1/2, 1] \\ 1, & X(\omega) > 1 \end{cases}$$

$$X_2(\omega) = \sum_{i=0}^2 \frac{1}{4} \mathbb{I}_{X^{-1}\left(\left[\frac{i}{2}, \frac{i+1}{2}\right]\right)}(\omega) + 2 \mathbb{I}_{X^{-1}([2, \infty))}(\omega)$$

$$X_2(\omega) = \sum_{i=0}^{\infty} \frac{1}{4^i} \mathbb{I}(\omega) + 2 \mathbb{I}(\omega) \\ X^{-1}\left(\left[\frac{i}{4^i}, \frac{i+1}{4^i}\right)\right) \quad X^{-1}([2, \infty))$$

$$X_n(\omega) = \sum_{i=0}^{n2^n-1} \frac{i}{2^n} \mathbb{I}(\omega) + n \mathbb{I}(\omega) \\ X^{-1}\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) \quad X^{-1}([n, \infty))$$

Propriedades.

$$1. \int_{\Omega} X dP = \lim_{n \rightarrow \infty} \int_{\Omega} X_n dP \geq 0$$

$$2. c > 0, X: \Omega \rightarrow \mathbb{R}_+ \quad cX_n \uparrow cX$$

$$\int_{\Omega} (cX) dP = \lim_{n \rightarrow \infty} \int_{\Omega} (cX_n) dP = \lim_{n \rightarrow \infty} c \int_{\Omega} X_n dP = c \lim_{n \rightarrow \infty} \int_{\Omega} X_n dP = c \int_{\Omega} X dP$$

$$3. X, Y: \Omega \rightarrow \mathbb{R}_+$$

$$\int_{\Omega} (X+Y) dP = \begin{matrix} X_n \uparrow X \\ Y_n \uparrow Y \end{matrix} \quad X_n + Y_n \uparrow X + Y$$

$$= \lim_{n \rightarrow \infty} \int_{\Omega} (X_n + Y_n) dP = \lim_{n \rightarrow \infty} \left[\int_{\Omega} X_n dP + \int_{\Omega} Y_n dP \right] =$$

$$= \lim_{n \rightarrow \infty} \int_{\Omega} X_n dP + \lim_{n \rightarrow \infty} \int_{\Omega} Y_n dP = \int_{\Omega} X dP + \int_{\Omega} Y dP.$$

Resultado:

X r.a. assumindo valores em \mathbb{N}

$$E(X) = \int_{\Omega} X dP = \sum_{n=1}^{\infty} P(X \geq n)$$

$$X = \sum_{j=1}^{\infty} j \mathbb{I}_{A_j}, \quad A_j = X^{-1}(\{j\})$$

$$\text{Seja } X_n = \sum_{j=1}^{n-1} j \mathbb{I}_{A_j} + n \mathbb{I}_{\bigcup_{j=n}^{\infty} A_j}$$

$$X_{n+1} = \sum_{j=1}^n j \mathbb{I}_{A_j} + (n+1) \mathbb{I}_{\bigcup_{j=n+1}^{\infty} A_j} \quad X_n \leq X_{n+1}$$

$$\int_n X_n dP = \sum_{j=1}^{n-1} j P(A_j) + n P\left(\bigcup_{j=n}^{\infty} A_j\right)$$

$$= \sum_{j=1}^{n-1} \sum_{i=j}^j P(A_j) + n \cdot P\left(\bigcup_{j=n}^{\infty} A_j\right)$$

$$= \sum_{i=1}^{n-1} \sum_{j=1}^i P(A_j) + n P(X \geq i) =$$

$$= \sum_{i=1}^{n-1} P(i \leq X \leq n-1) + n P(X \geq n) \stackrel{(*)}{=} \quad (*)$$

$$= \sum_{i=1}^n P(X \geq i) = \int_n X_n dP$$

Assim,

$$\int_n X dP = \lim_{n \rightarrow \infty} \int_n X_n dP = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(X \geq i) = \sum_{i=1}^{\infty} P(X \geq i).$$

$$(*) \quad P(1 \leq X \leq n-1) + P(X \geq n) \Rightarrow P(X \geq 1)$$

$$P(2 \leq X \leq n-1) + P(X \geq n) \Rightarrow P(X \geq 2)$$

$$P(3 \leq X \leq n-1) + P(X \geq n) \Rightarrow P(X \geq 3)$$

;

$$P(n-1 \leq X \leq n-1) + P(X \geq n) \Rightarrow P(X \geq n-1)$$

$$P(X \geq n) \Rightarrow P(X \geq n)$$

Teorema de Tonelli

Teorema de Fubini:

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}$$

Ex. $X \sim Geo(p)$

$$E(X) = \int_{\Omega} X dP = \sum_{i=1}^{\infty} P(X=i) - *$$

$$P(X=i) = \sum_{j=1}^{\infty} P(X=j) = \sum_{j=i}^{\infty} (1-p)^{j-1} p = \frac{p(1-p)^{i-1}}{1-(1-p)} = (1-p)^{i-1}$$

$$* = \sum_{i=1}^{\infty} (1-p)^{i-1} = \frac{1}{1-(1-p)} = \frac{1}{p},$$

Definição: Para $E \in \mathcal{F}$

$$\int_E X dP = \int_{\Omega} X \mathbb{I}_E dP.$$

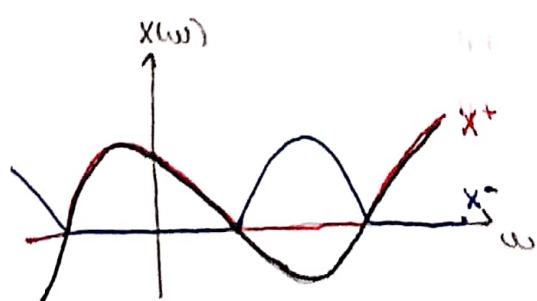
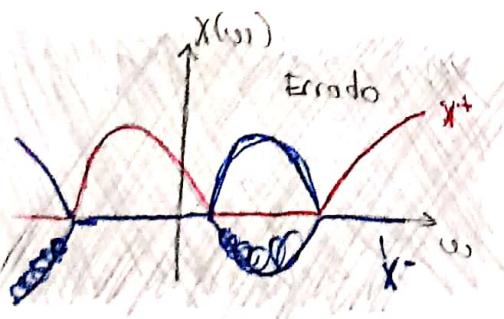
Caso Geral.

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

Sejam $X^+, X^- : \Omega \rightarrow \mathbb{R}_+$ 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 $X: \Omega \rightarrow \mathbb{R}$.

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

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

$$X^+ = \max\{X, 0\}$$

$$X^- = \max\{-X, 0\}$$

$$X = X^+ - X^-$$

$$|X| = X^+ + X^-$$

Definição: Se $\int_{\Omega} X^+ dP < \infty$ ou $\int_{\Omega} X^- dP < \infty$, então definimos

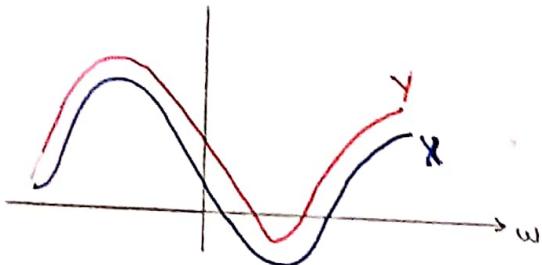
$$\int_{\Omega} X dP = \int_{\Omega} X^+ dP - \int_{\Omega} X^- dP$$

Comentário: Se $\int_{\Omega} X^+ dP < \infty$ e $\int_{\Omega} X^- dP < \infty$, dizemos que X é integrável.

Propriedades:

1. $X, Y: \Omega \rightarrow \mathbb{R}$, 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

$$X^+(\omega) \leq Y^+(\omega) \quad \text{e} \quad Y^-(\omega) \leq X^-(\omega)$$

$$\text{Logo, } \int_{\Omega} X dP = \int_{\Omega} X^+ dP - \int_{\Omega} X^- dP \leq \int_{\Omega} Y^+ dP - \int_{\Omega} Y^- dP = \int_{\Omega} Y dP$$

$X, Y: \Omega \rightarrow \mathbb{R}^+$, $X \leq Y$

$$\int_{\Omega} X dP = \sup \left\{ \int_{\Omega} f dP : f \in \mathcal{F}_X \right\}$$

$$\int_{\Omega} Y dP = \sup \left\{ \int_{\Omega} f dP : f \in \mathcal{F}_Y \right\}$$

($\mathcal{F}_X \subseteq \mathcal{F}_Y$, pois $X \leq Y$)

2. $X: \Omega \rightarrow \mathbb{R}$, $c \in \mathbb{R}$

cX

$$\int_{\Omega} (cX) dP = c \int_{\Omega} X dP$$

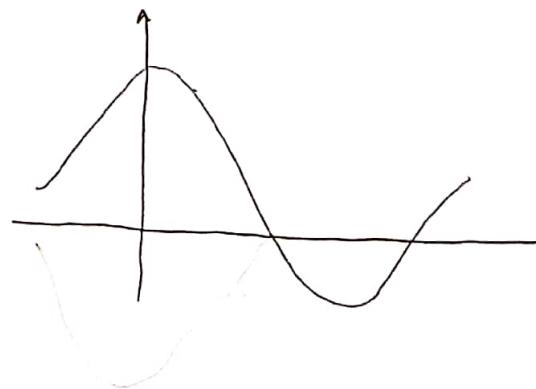
$$c > 0: (cX)^+ = cX^+$$

$$(cX)^- = cX^-$$

$$\int_{\Omega} (cX) dP = \int_{\Omega} (cX)^+ dP - \int_{\Omega} (cX)^- dP =$$

$$= \int_{\Omega} cX^+ dP - \int_{\Omega} cX^- dP =$$

$$= c \int_{\Omega} X^+ dP - c \int_{\Omega} X^- dP =$$



$$c < 0: (cX)^+ = (-c)X^-$$

$$(cX)^- = (-c)X^+$$

$$\int_{\Omega} (cX) dP = \int_{\Omega} (-c)X^- dP - \int_{\Omega} (-c)X^+ dP =$$

$$= (-c) \int_{\Omega} X^- dP + c \int_{\Omega} X^+ dP$$

$$3. \int_{\Omega} X+Y dP = \int_{\Omega} X dP + \int_{\Omega} Y dP$$

$$|x+y| \leq |x| + |y| \Rightarrow \int_{\Omega} |x+y| dP \leq \int_{\Omega} |x| dP + \int_{\Omega} |y| dP$$

$$\left. \begin{aligned} X+Y &= (X+Y)^+ - (X+Y)^- \\ X+Y &= X^+ - X^- + Y^+ - Y^- \end{aligned} \right\} \Rightarrow$$

$$\Rightarrow (X+Y)^+ - (X+Y)^- = X^+ - X^- + Y^+ - Y^- \Rightarrow$$

$$\Rightarrow (X+Y)^+ + X^- + Y^- = X^+ + Y^+ + (X+Y)^-$$

$$\Rightarrow \int_{\Omega} (X+Y)^+ dP + \int_{\Omega} X^- dP + \int_{\Omega} Y^- dP = \int_{\Omega} X^+ dP + \int_{\Omega} Y^+ dP + \int_{\Omega} (X+Y)^- dP \Rightarrow$$

$$\Rightarrow \underbrace{\int_{\Omega} (X+Y)^+ dP - \int_{\Omega} (X+Y)^- dP}_{\int_{\Omega} (X+Y) dP} = \underbrace{\int_{\Omega} X^+ dP - \int_{\Omega} X^- dP}_{\int_{\Omega} X dP} + \underbrace{\int_{\Omega} Y^+ dP - \int_{\Omega} Y^- dP}_{\int_{\Omega} Y dP}.$$

Exemplo:

X assume valores em $\{x_1, x_2, \dots\}$

$$X = \sum_{i=1}^{\infty} x_i I_{A_i}, \text{ onde } A_i = X^{-1}(\{x_i\}), i=1,2,\dots$$

$$X(\Omega) = \{X(\omega); \omega \in \Omega\} = \{x_1, \dots, x_n, \dots\}$$

$$X(\Omega)_+ = X(\Omega) \cap \mathbb{R}_+ = \{x_{k_1}, x_{k_2}, \dots\}$$

$$X(\Omega)_- = \cancel{X(\Omega)} - X(\Omega)_+ = \{x_{k_1}, x_{k_2}, \dots\}$$

X^+ assume valores em $X(\Omega)_+$.

$$X^+ = \sum_{i=1}^{\infty} x_{k_i} I_{A_{k_i}}$$

Se a soma acima for finita, X^+ é simples e $\int X^+ dP = \sum_{x \in X(\Omega)_+} x P(A_{x_0})$

Se for finito, definimos

$$X_n^+ = \sum_{i=1}^n x_{K_i} \mathbb{I}_{A_{K_i}} \uparrow X^+$$

Dai, $\int_{\Omega} X_n^+ dP = \sum_{i=1}^n x_{K_i} P(A_{K_i})$

$$\int_{\Omega} X^+ dP = \lim_{n \rightarrow \infty} \int_{\Omega} X_n^+ dP = \lim_{n \rightarrow \infty} \sum_{i=1}^n x_{K_i} P(A_{K_i}) =$$

$$= \sum_{i=1}^{\infty} x_{K_i} P(A_{K_i}) = \sum_{i: x_i > 0} x_{K_i} P(A_{K_i}) = \sum_{i: x_i > 0} x_i P(A_i) = \sum_{i: x_i > 0} x_i P(X=x_i)$$

Analogamente,

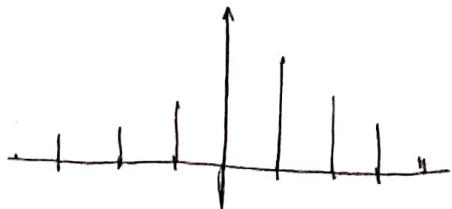
$$\int_{\Omega} X^- dP = \sum_{i: x_i < 0} (-x_i) P(A_i) = \sum_{i: x_i < 0} (-x_i) P(X=x_i)$$

¶ Se $\int_{\Omega} X^+ dP < \infty$ ou $\int_{\Omega} X^- dP < \infty$, vale que

$$\int_{\Omega} X dP = \sum_{i: x_i > 0} x_i P(X=x_i) - \sum_{i: x_i < 0} (-x_i) P(X=x_i)$$

Exemplo:

$$P(X=x) = \frac{1}{2} \left(\frac{2}{3}\right)^{x-1} \cdot \frac{1}{3} \mathbb{I}_{\mathbb{Z}_-^*}(x) + \frac{1}{2} \left(\frac{1}{2}\right)^x \mathbb{I}_{\mathbb{Z}_+^*}(x)$$



$\frac{1}{2}$ Geométrica $(\frac{2}{3}, x)$ para $x \in \mathbb{Z}_+^*$
 $\frac{1}{2}$ " " $(\frac{1}{2}, x)$ " " $x \in \mathbb{Z}_-^*$

$$\int_{\Omega} X^+ dP = \sum_{i=1}^{\infty} i \cdot \frac{1}{2} \left(\frac{1}{2}\right)^{i-1} = \frac{1}{2} \sum_{i=1}^{\infty} i \underbrace{\left(\frac{1}{2}\right)^{i-1}}_2 \cdot \frac{1}{2} = \frac{1}{2} \cdot 2 = 1$$

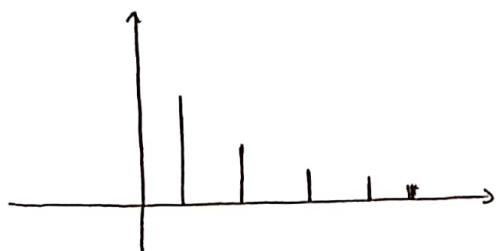
$$\int_{\Omega} X^- dP = \sum_{x \in \mathbb{Z}^+} (-i) P(X=i) = \sum_{i=1}^{\infty} i P(X=-i) =$$

$$= \sum_{i=1}^{\infty} i \cdot \frac{1}{2} \left(\frac{1}{3}\right)^{i-1} \cdot \frac{1}{3} = \frac{1}{2} \cdot \frac{1}{1/3} = \frac{3}{2}.$$

Logo

$$\int_{\Omega} X dP = 1 - \frac{3}{2} = -\frac{1}{2}$$

Exemplo 2. $P(X=\infty) = \lim_{x \rightarrow \infty} \frac{1}{x^2}$

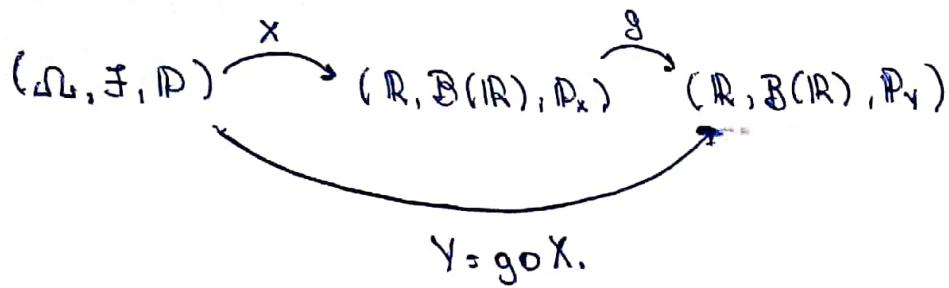


$$E(X^+) = \int_{\Omega} X^+ dP = \sum_{x=1}^{\infty} x \cdot \lim_{x \rightarrow \infty} \frac{1}{x^2} = \infty$$

$$E(X^-) = \int_{\Omega} X^- dP = \sum_{x \in \mathbb{Z}^+} (-x) \cdot \lim_{x \rightarrow \infty} \frac{1}{x^2} = \sum_{i=1}^{\infty} -i \cdot \lim_{x \rightarrow \infty} \frac{1}{x^2} = -\infty.$$

Nesse caso, $\int_{\Omega} X^+ dP = \infty$ e $\int_{\Omega} X^- dP = -\infty$.

Logo, dizemos que não existe $\int_{\Omega} X dP$.



Resultado:

$$\int_{\Omega} Y dP \stackrel{\text{def}}{=} \int_{\Omega} g \circ X dP = \int_{\Omega} g dP_x \quad (\text{Lei da Estatística Inconsciente})$$

Verificação:

$$g = \sum_{i=1}^r g_i \mathbb{I}_{B_i} \quad ; \quad g_1, g_2, \dots, g_r \geq 0$$

$$B_1, B_2, \dots, B_r \in \mathcal{B}(R)$$

$$\int_{\Omega} g \circ X dP = \int_{\Omega} g(X) dP = \int_{\Omega} \left(\sum_{i=1}^r g_i \mathbb{I}_{B_i}(X) \right) dP =$$

\downarrow
 $X(\omega) \in B_i$
 \Downarrow
 $\omega \in X^{-1}(B_i)$

$$= \int_{\Omega} \left(\sum_{i=1}^r g_i \mathbb{I}_{X^{-1}(B_i)} \right) dP = \sum_{i=1}^r g_i P(X^{-1}(B_i)) = \sum_{i=1}^r g_i P_X(B_i) = \int_{\Omega} g dP_x.$$

$$\sum_{i=1}^r g_i \mathbb{I}_{B_i}(X(\omega)) = \sum_{i=1}^r g_i \mathbb{I}_{X^{-1}(B_i)}(\omega)$$

• Se X discreto

$$\sum_{i=1}^r g_i \sum_{x \in B_i} P(X=x) = \sum_{i=1}^r g_i \sum_{x \in X(\Omega)} \mathbb{I}_{B_i}(x) P(X=x) =$$

$$= \sum_{i=1}^r \left\{ \sum_{x \in X(\Omega)} g_i \mathbb{I}_{B_i}(x) P(X=x) \right\} =$$

$$= \sum_{x \in X(\Omega)} \left\{ \sum_{i=1}^r g_i \mathbb{I}_{B_i}(x) \right\} P(X=x)$$

Verificações:

$$g \geq 0, \quad g_n \xrightarrow{\text{simples}} g$$

$$g_n(x) \uparrow g(x)$$

$$\int (g \circ X) dP = \lim_{n \rightarrow \infty} \underbrace{\int_R (g_n \circ X) dP}_{\int_R g_n dP_x} = \lim_{n \rightarrow \infty} \int_R g_n dP_x = \dots = \int_R g dP_x$$

X contínuo

$$\sum_{i=1}^r g_i \int_{B_i} f_X(t) dt = \sum_{i=1}^r g_i \int_{-\infty}^{\infty} I_{B_i}(t) f_X(t) dt = \dots = \int_{-\infty}^{\infty} \left(\sum_{i=1}^r g_i I_{B_i}(t) \right) f_X(t) dt$$

$$(g(\infty))^+ \cdot (g(\infty))^-.$$

Como consequência, temos

$$X \text{ discreto: } \int_R (g \circ X) dP = E[g(X)] = \sum_{\infty} g(\infty) P(X=\infty),$$

$$\text{desde } \sum_{\infty} g^+(\infty) P(X=\infty) < \infty \text{ e } \sum_{\infty} g^-(\infty) P(X=\infty) < \infty$$

$$X \text{ absolutamente contínua: } E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

Exemplo 1: $X \sim \text{Beta}(a, b)$, $a, b > 0$.

$$g(t) = t^n (1-t)^m, \quad n, m \geq 0$$

$$Y = g(X) = X^n (1-X)^m$$

$$\begin{aligned}
 E(Y) &= E(X^n(1-x)^m) = \int_{-\infty}^{\infty} x^n(1-x)^m f_X(x) dx = \\
 &= \int_0^1 x^n(1-x)^m \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} dx = \\
 &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+n)\Gamma(b+m)}{\Gamma(a+b+n+m)} \int_0^1 \underbrace{\frac{\Gamma(a+n+b+m)}{\Gamma(a+n)\Gamma(b+m)} x^{a+n-1} (1-x)^{b+m-1}}_{\downarrow} dx
 \end{aligned}$$

$$n=1, m=0 \Rightarrow E(X) = \frac{a}{a+b}$$

$$n=2, m=0 \Rightarrow E(X^2) = \frac{(a+1)a}{(a+b+1)(a+b)}$$

Ejemplo 2. $X \sim \text{Poisson}$

$$\begin{aligned}
 g: \mathbb{N} &\rightarrow \mathbb{N} \\
 n \in \mathbb{N} &\mapsto g(n) = n!
 \end{aligned}$$

$$Y = g(X) = X!$$

$$\begin{aligned}
 E(Y) &= E(X!) = \sum_{x=0}^{\infty} x! P(X=x) = \sum_{x=0}^{\infty} x! \frac{e^{-\lambda} \lambda^x}{x!} = \sum_{x=0}^{\infty} e^{-\lambda} \lambda^x = \\
 &= \begin{cases} \infty, & \lambda \geq 1 \\ \frac{e^{-\lambda}}{1-\lambda}, & 0 < \lambda < 1 \end{cases}
 \end{aligned}$$

Se X é vetor aleatório discreto (contínuo) a valores em \mathbb{R}^k e g. $\mathbb{R}^k \rightarrow \mathbb{R}$ monótona, vale que

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}^k} (g \circ X) dP = \sum_{\mathbf{x}} g(\mathbf{x}) P(X=\mathbf{x}) \left(\int_{\mathbb{R}^k} g(\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} \right)$$

↓
def. das. - f(x)

Exemplo.

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

$$g: \mathbb{R}^k \rightarrow \mathbb{R}$$

$$(x_1, \dots, x_k) \mapsto e^{t_1 x_1 + t_2 x_2 + \dots + t_k x_k}, \quad t_1, t_2, \dots, t_k \in \mathbb{R}$$

$$Y = g(X_1, \dots, X_k) = e^{t_1 X_1 + \dots + t_k X_k}$$

$$\mathbb{E}(Y) = ?$$

$$\mathbb{E}[Y] = \mathbb{E}[g(X_1, \dots, X_k)] = \sum_{(x_1, \dots, x_k) \in A_{k,n}} e^{t_1 x_1 + \dots + t_k x_k} \cdot P(X_1=x_1, \dots, X_k=x_k)$$

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

$$= \sum_{A_{k,n}} e^{t_1 a_1 + \dots + t_k a_k} \cdot \frac{n!}{a_1! a_2! \dots a_k! a_{k+1}!} \cdot p_1^{a_1} p_2^{a_2} \dots p_k^{a_k} (p_{k+1})^{a_{k+1}} =$$

$p_{k+1} = 1 - p_1 - \dots - p_k$

$\frac{1}{a_1 + a_2 + \dots + a_k}$

$$= \sum_{A_{K,n}} \frac{n!}{\prod_{i=1}^{K+1} x_i!} (p_1 e^{t_1})^{x_1} \cdots (p_K e^{t_K})^{x_K} \cdot p_{K+1}^{x_{K+1}} =$$

$$= (p_1 e^{t_1} + \cdots + p_K e^{t_K} + p_{K+1})^n \cdot \underbrace{\sum_{A_{K,n}} \frac{n!}{\prod_{i=1}^{K+1} x_i!} \frac{(p_1 e^{t_1})^{x_1} \cdots (p_K e^{t_K})^{x_K} \cdot p_{K+1}^{x_{K+1}}}{(p_1 e^{t_1} + \cdots + p_K e^{t_K} + p_{K+1})^n}}$$

≈ 1

Então,

$$\mathbb{E}[e^{t_1 X_1 + \cdots + t_K X_K}] = (p_1 e^{t_1} + p_2 e^{t_2} + \cdots + p_K e^{t_K} + 1 - p_1 - \cdots - p_K)^n$$

Exemplo 4.

 (X_1, X_2) contínuas

$$f_{X_1, X_2}(x_1, x_2) = \frac{1}{x_2} e^{-x_2} I_{(0, x_2)}(x_1) I_{R_+}(x_2)$$



Exemplo 5.

$$X = (X_1, \dots, X_K) \sim \text{DIR}(\alpha_1, \alpha_2, \dots, \alpha_K, \alpha_0) \quad (\text{Tarefa})$$

$$\mathbb{E}(X_1^{\alpha_1} X_2^{\alpha_2} \cdots X_K^{\alpha_K}).$$

Ex. 4.

$$\mathbb{E}(Y) = \mathbb{E}(X_1 X_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2 f_{(X_1, X_2)}(x_1, x_2) dx_1 dx_2 =$$

$$= \int_0^{\infty} \int_0^{\alpha_2} x_1 x_2 \frac{1}{x_2} e^{-x_2} dx_1 dx_2 =$$

$$= \int_0^\infty e^{-x_2} \left[\int_0^{x_2} x_1 dx_1 \right] dx_2 = \int_0^\infty e^{-x_2} \frac{x_2^2}{2} dx_2 =$$

$$= \frac{1}{2} \cdot \frac{\Gamma(3)}{1^3} \underbrace{\int_0^\infty t^{3-2} e^{-t} \frac{t^3}{\Gamma(3)} dt}_{\substack{\parallel \\ \downarrow}} = \frac{1}{2} \cdot 2! = 1$$

⇒ $E(X_1 X_2) = \frac{1}{2} \cdot 2! = 1.$

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. $X_N \sim \text{Bin}(n, p)$

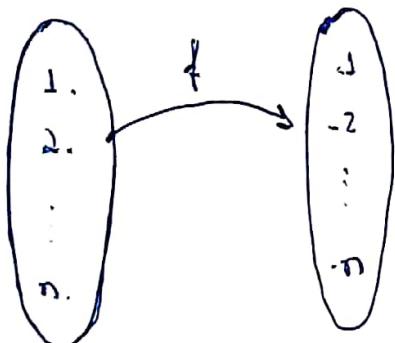
$$E(X) = \sum_{x=0}^n x P(X=x)$$

$$X \stackrel{d}{=} X_1 + \dots + X_n, \text{ onde}$$

X_1, \dots, X_n são v.a. i.i.d. de $\text{Ber}(p)$.

$$E(X) = E(X_1 + \dots + X_n) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n \underbrace{\{1P(X_i=1) + 0P(X_i=0)\}}_p \Rightarrow E(X) = np$$

Exemplo 2.



$$f: \{1, \dots, N\} \rightarrow \{1, \dots, N\}$$

$$\Omega_2 = \{f: \{1, \dots, n\} \rightarrow \{1, \dots, n\} : f \text{ é bijetora}\}$$

X_i : N° de pontos do domínio tais que $f(i) = i$.

$$E(X) = ?$$

$$X_i = \begin{cases} 1, & \text{se } f(i) = i \\ 0, & \text{o.c.} \end{cases}$$

$$X_1 + \dots + X_N = X$$

$$\begin{aligned} E(X) &= E(X_1 + \dots + X_N) = \sum_{i=1}^N \underbrace{E(X_i)}_{\substack{1 \text{ } P(X_i=1) + 0 \text{ } P(X_i=0) \\ \vdots \\ \frac{(N-1)!}{N!} = \frac{1}{N}}} = \sum_{i=1}^N \frac{1}{N} = 1. \end{aligned}$$

(Sheldon Ross)

Exemplo 3.

Lance uma moeda n vezes.

"Run" de tamanho k , $k \leq n$, uma sequência de k caras seguidas.

$$\begin{array}{l} n=6 \\ k=2 \\ \text{C C C } \bar{C} \text{ C C} \\ \text{"Runs" de tamanho 2.} \end{array}$$

X_i : N. de "runs" de tamanho k

$$P(X_i=j) = ?$$

X : N.º de runs de tamanho k

$$P(X=j) = ?$$

$$X_i = \begin{cases} 1, & \text{ocorrência de "run" de tamanho } k \text{ no } i\text{-ésimo lançamento} \\ 0, & \text{c.c.} \end{cases}$$

X_1, \dots, X_{n-k+1} variáveis:

$$X_i = \begin{cases} 1, & \text{ocorrência de "run" de tamanho } k \text{ começando no } i\text{-ésimo lançamento} \\ 0, & \text{c.c.} \end{cases}$$

$$X = X_1 + \dots + X_{n-k+1}$$

$$X = X_1 + \dots + X_{n-k+1}$$

$$E(X) = E(X_1 + \dots + X_{n-k+1}) = E(X_1) + \dots + E(X_{n-k+1}) =$$

$$P(X_i=1) = p^k, \quad i=1, \dots, n-k+1$$

$$E(X) = (n-k+1)p^k.$$

Resultado:

Sejam X_1, \dots, X_n v.a. independentes e $g_1, g_2, \dots, g_n: \mathbb{R} \rightarrow \mathbb{R}$ mensuráveis tal que $g_1(X_1), \dots, g_n(X_n)$ são integráveis.

Então,

$$\mathbb{E}[g_1(X_1)g_2(X_2)\dots g_n(X_n)] = \mathbb{E}[g_1(X_1)] \dots \mathbb{E}[g_n(X_n)].$$

Exemplo. $X_1 \sim \text{Beta}(a, b)$,
 $X_2 \sim \text{Poisson}(\lambda)$) independentes

$$W = X_1^2 + X_2 \quad \begin{aligned} g_1(t) &= t^2 \\ g_2(t) &= t \end{aligned}$$

$$\mathbb{E}(W) = \mathbb{E}(X_1^2 + X_2) = \mathbb{E}(X_1^2) \mathbb{E}(X_2) = \frac{a(a+1)}{(a+b)(a+b+1)} \cdot \lambda$$

MOMENTOS

$X: \Omega \rightarrow \mathbb{R}$ uma v.a.

$$g: \mathbb{R} \rightarrow \mathbb{R}$$

$$g(t) \rightarrow t^k$$

$\mathbb{E}[g(X)] = \mathbb{E}[X^k] \rightarrow k\text{-ésimo momento da v.a. } X$

$$g(t) = (t - \mathbb{E}(X))^k, \mathbb{E}(X) < \infty.$$

$\mathbb{E}[g(X)] = \mathbb{E}[(X - \mathbb{E}(X))^k] \rightarrow k\text{-ésimo momento central de } X$

$$K=1 \rightarrow E[(X-E(X))] = 0$$

$$K=2 \rightarrow E[(X-E(X))^2] = \text{Var}(X) . \text{ Variância da v.a. } X.$$

$$\text{Var}(X) = \dots = E(X^2) - (E(X))^2$$

Exemplo:

$$X \sim \text{Beta}(\alpha, \beta)$$

Vimos que

$$E(X) = \frac{\alpha}{\alpha+\beta} \quad \text{e} \quad E(X^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta} \right)^2 =$$

$$= \frac{\alpha(\alpha+1)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta)^2 \cdot (\alpha+\beta+1)} = \frac{\alpha b}{(\alpha+\beta)^2 (\alpha+\beta+1)}$$

Propriedades

$$1. \text{Var}(X) \geq 0$$

$$2. \text{Var}(X)=0 \Leftrightarrow P(X=E(X))=1.$$

$$Y \geq 0.$$

$$E(Y) = 0 \Rightarrow P(Y=0)=1$$

$$P(Y \geq 1/n) = E \left[\underbrace{\mathbb{I}_{\{w: Y(w) \geq 1/n\}}}_{A_n} \right] = \int_{\Omega} \mathbb{I}_{A_n} dP = n \int_{\Omega} \frac{1}{n} \mathbb{I}_{A_n} dP$$

$$\leq \int_{\Omega} Y dP = 0, \forall n \in \mathbb{N}.$$

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} \{Y \neq Y_n\}\right) = 0 \Rightarrow$$

$$\Rightarrow P(Y \neq 0) = 0 \stackrel{Y \sim 0}{\Rightarrow} P(Y = 0) = 1.$$

3. $\text{Var}(\alpha X + b) = \alpha^2 \text{Var}(X)$ (Exercício)

Propriedade: Se X possui momento de ordem k , então possui momento de ordem j , para todo $j < k$.

$$E(|X^j|) = \int_{\Omega} |X^j| dP = \int_{\Omega} |X^j| \mathbb{I}_{\{\omega : |X(\omega)| \leq j\}} dP + \int_{\Omega} |X^j| \mathbb{I}_{\{\omega : |X(\omega)| > j\}} dP \leq$$

$$\int_{\Omega} \mathbb{I}_{\{\omega : |X(\omega)| \leq j\}} dP + \int_{\Omega} |X^k| \mathbb{I}_{\{\omega : |X(\omega)| > j\}} dP \leq$$

$$1 + \int_{\Omega} |X^k| dP < \infty.$$

Seja X uma variável aleatória,

Seja $\psi_X: I \rightarrow \mathbb{R}_+$ a função que associa a cada $t \in I = (-\varepsilon, \varepsilon)$, para algum $\varepsilon > 0$, o número $\psi_X(t) = E[e^{tX}]$.

A função ~~função~~ ψ_X damos o nome de FUNÇÃO GERADORA DE MOMENTOS.

EX) $X \sim \text{Poisson}(\lambda)$

$$\psi_X(t) = E(e^{tX}) = \sum_{j=0}^{\infty} e^{ts} P(X=j) = \sum_{j=0}^{\infty} e^{ts} \frac{e^{-\lambda} \lambda^j}{j!} = e^{-\lambda} \sum_{j=0}^{\infty} \frac{(e^t \lambda)^j}{j!} =$$

$$= e^{-\lambda} e^{\lambda e^t} \underbrace{\sum_{j=0}^{\infty} \frac{(e^t \lambda)^j}{j!}}_{1} e^{-\lambda e^t} \Rightarrow \psi_X(t) = e^{\lambda e^t - \lambda}$$

Exemplo 2:

$X \sim \text{Gama}(\alpha, b)$

$$\varphi_x(t) = E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f_x(x) dx = \int_0^{\infty} e^{tx} \frac{b^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-bx} dx =$$

$$= \frac{b^\alpha}{(b-t)^\alpha} \int_0^{\infty} \frac{(b-t)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-(b-t)x} dx \stackrel{t < b}{\Rightarrow} \varphi_x(t) = \left(\frac{b}{b-t} \right)^\alpha.$$

Propriedades:

$$1. \varphi_x(0) = 1$$

$$2. \varphi_{ax+b}(t) = E[e^{t(ax+b)}] = E[e^{tb} e^{atx}] = e^{tb} E[e^{atx}] \Rightarrow \varphi_{ax+b}(t) = e^{tb} \varphi_x(at)$$

$$3. X, Y \text{ são independentes} \Rightarrow \varphi_{x+y}(t) = \varphi_x(t) \cdot \varphi_y(t).$$

$$\varphi_{ax+b}(t) = e^{tb} \varphi_x(at),$$

$$\varphi_{x+y}(t) = E(e^{t(x+y)}) = E(e^{tx} \cdot e^{ty}) = E(e^{tx}) E(e^{ty}) = \varphi_x(t) \varphi_y(t).$$

$$4. E(X^k) < \infty, \forall k \geq 1 \text{ e } \varphi_x \text{ existe e admite derivadas de toda ordem} \Rightarrow$$

$$\left. \frac{d^k \varphi_x(t)}{dt^k} \right|_{t=0} = E(X^k)$$

PS:

$$e^{tx} = \sum_{j=0}^{\infty} \frac{(tx)^j}{j!} \Rightarrow E(e^{tx}) = \int_0^{\infty} e^{tx} dx = \sum_j \int_0^{\infty} x^j dx$$

Exemplo:

$X \sim \text{Poisson}(\lambda)$

$$\varphi_x(t) = e^{\lambda e^t - \lambda}$$

$$\varphi'_x(t) = e^{\lambda e^t - \lambda} \cdot \lambda e^t$$

$$(\varphi'_x(0) = \lambda = E(X)).$$

$$\Psi_x''(t) = \lambda \left\{ e^{\lambda e^t - \lambda}, \lambda e^t, e^t + e^{\lambda e^t - \lambda} \cdot e^t \right\}$$

$$\Psi_x''(0) = \lambda \{\lambda + 1\} = \lambda^2 + \lambda = E(X^2)$$

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Exemplo:

Seja (X, Y) um vetor aleatório contínuo com densidade

$$f_{x,y}(x,y) = \frac{1}{y} e^{-\frac{x}{y}} e^{-y} \mathbb{I}_{\mathbb{R}_+}(x) \mathbb{I}_{\mathbb{R}_+}(y)$$

$$E(X^k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k f_{x,y}(x,y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k \frac{1}{y} e^{-\frac{x}{y}} e^{-y} dx dy =$$

$$= \int_0^{\infty} \frac{1}{y} e^{-y} \left[\int_0^{\infty} \frac{(1/y)^{k+1}}{\Gamma(k+1)} x^k e^{-x/y} dx \right] \frac{\Gamma(k+1)}{(1/y)^{k+1}} dy \Rightarrow$$

$$E(X^k) = \int_0^{\infty} \Gamma(k+1) y^k e^{-y} dy = \Gamma(k+1) \frac{\Gamma(k+1)}{\Gamma(k+1)} \int_0^{\infty} \frac{1^{k+1}}{\Gamma(k+1)} y^k e^{-y} dy \Rightarrow$$

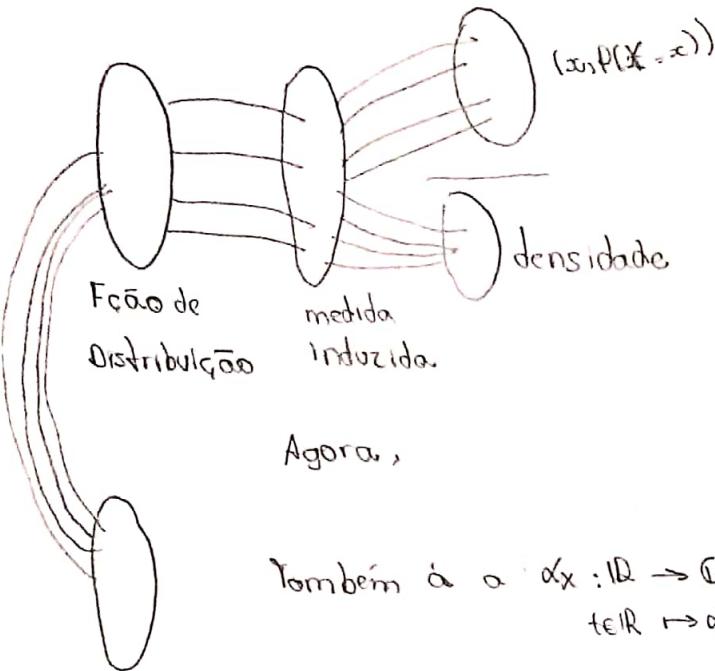
$$\Rightarrow E(X^k) = (k!)^2 < \infty.$$

$$\Psi_x(t) = E(e^{tx}) =$$

$$= \int_0^{\infty} \int_0^{\infty} e^{tx} \cdot \frac{1}{y} e^{-\frac{x}{y}} e^{-y} dx dy = \int_0^{\infty} \frac{1}{y} e^{-y} \left(\int_0^{\infty} e^{-(\frac{t}{y}-t)x} dx \right) dy$$

converge para $t \leq 0$.

Ramadas Brat - Livro legal.



bijection - se uma, então outra ...

Agora,

Também é $\alpha: \mathcal{X}: \mathbb{R} \rightarrow \mathbb{C}$

$$t \in \mathbb{R} \mapsto \alpha_X(t) = E[e^{itX}] . \text{ Essa, sempre existe!}$$

Funções
Geradoras de
Momentos

$$\psi_X(t) = E[e^{tx}] \rightarrow \text{transformada de Laplace}$$

~~$$E(aX + bY) = aE(X) + bE(Y)$$~~

Agora,

$$\text{Var}(X+Y) = E((X+Y)^2) - [E(X)+E(Y)]^2 =$$

$$= E(X^2) + 2E(XY) + E(Y^2) - (E(X))^2 - 2E(X)E(Y) - (E(Y))^2 =$$

$$= \text{Var}(X) + \text{Var}(Y) + 2\{E(XY) - E(X)E(Y)\}$$

$\text{Cov}(X, Y)$: covariância entre as variáveis
 X e Y

Poderemos ainda escrever:

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$$

Exemplo:

$$(X_1, X_2) \sim \text{Dir}(\alpha_1, \alpha_2, \alpha_0)$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

$$x_1 + x_2 \leq 1$$

$$X_1 \sim \text{Beta}(\alpha_1, \alpha_0 + \alpha_2)$$

$$X_2 \sim \text{Beta}(\alpha_2, \alpha_0 + \alpha_1)$$

$$E(X_1 X_2) = \int_{A_2} x_1 x_2 \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_0)}{\Gamma(\alpha_1) \Gamma(\alpha_2) \Gamma(\alpha_0)} x_1^{\alpha_1-1} x_2^{\alpha_2-1} (1-x_1-x_2)^{\alpha_0-1} dx_1 dx_2 =$$

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_0)}{\Gamma(\alpha_1) \Gamma(\alpha_2) \Gamma(\alpha_0)} \cdot \frac{\Gamma(\alpha_0) \Gamma(\alpha_1+1) \Gamma(\alpha_2+1)}{\Gamma(\alpha_0 + \alpha_1 + \alpha_2 + 2)} \left[\int_{A_2} \frac{\Gamma(\alpha_0 + \alpha_1 + \alpha_2 + 2)}{\Gamma(\alpha_0) \Gamma(\alpha_1+1) \Gamma(\alpha_2+1)} x_1^{\alpha_1+1-1} x_2^{\alpha_2+1-1} (1-x_1-x_2)^{\alpha_0-1} dx_1 dx_2 \right]$$

$$= \frac{\alpha_1 \cdot \alpha_2}{(\alpha_0 + \alpha_1 + \alpha_2)(\alpha_0 + \alpha_1 + \alpha_2 + 1)}.$$

Logo,

$$\text{Cov}(X_1, X_2) = \frac{\alpha_1 \cdot \alpha_2}{(\alpha_0 + \alpha_1 + \alpha_2)(\alpha_0 + \alpha_1 + \alpha_2 + 1)} - \frac{\alpha_1}{(\alpha_0 + \alpha_1 + \alpha_2)} \frac{\alpha_2}{(\alpha_0 + \alpha_1 + \alpha_2)} < 0.$$

Propriedades:

$$1. \text{Cov}(X, Y) = \text{Cov}(Y, X)$$

$$2. \text{Cov}(X, X) = \text{Var}(X)$$

$$3. \text{Cov}\left(\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov}(X_i, Y_j)$$

Consequência.

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \text{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^n X_j\right) = \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(X_i, X_j) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{i=1}^n \sum_{j \neq i} \text{Cov}(X_i, X_j)$$

Voltando ao exemplo dos pacamentos

X_i : N° de elementos de $\{1, \dots, N\}$ tais que $f(i) = i$.

$$X = \sum_{i=0}^n X_i$$

$$\text{Var}(X) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{i=1}^n \sum_{j \neq i} \text{Cov}(X_i, X_j)$$

Vemos que $X_i \sim \text{Ber}(1/N)$, $i=1, \dots, N$

$$\text{Var}(X_i) = E(X_i^2) - (E(X_i))^2 =$$

$$= \{1^2 \cdot P(X_i=1) + 0^2 \cdot P(X_i=0)\} - \left(\frac{1}{N}\right)^2 =$$

$$= \frac{1}{N} - \left(\frac{1}{N}\right)^2$$

$i \neq j$

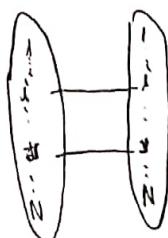
$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i) E(X_j)$$

$$= P(X_i=1, X_j=1) - \frac{1}{N} \frac{1}{N} =$$

(X_1, X_2) é um par de Bernoulli, então

$$X_1 X_2 \sim \text{Ber}(P(X_1=1, X_2=1))$$

$$= \frac{(N-2)!}{N!} - \frac{1}{N^2} =$$



$$= \frac{1}{N(N-1)} - \frac{1}{N^2}.$$

Logo,

$$\text{Var}(X) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{i=1}^n \sum_{j \neq i} \text{Cov}(X_i, X_j) = \sum_{i=1}^n \frac{1}{N} \left(\frac{N-1}{N} \right) + \sum_{i=1}^n \sum_{j \neq i} \frac{1}{N^2(N-1)} =$$

$$= \frac{N-1}{N} + (N^2 - N) \frac{1}{N^2(N-1)} = \frac{N-1}{N} + \frac{1}{N} = 1$$

Seja (X, Y) vetor aleatório.

DEFINIÇÃO: À quantidade

$\frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$ damos o nome de COEFICIENTE DE CORRELAÇÃO ENTRE X e Y.

Notação: CORR(X, Y).

DEFINIÇÃO: Ao número $\sqrt{\text{Var}(X)}$ damos o nome de DESVIO-PADRÃO de X.

NOTAÇÃO: DP(X) = $\sqrt{\text{Var}(X)}$.

Assim,

$$\text{CORR}(X, Y) = \frac{\text{Cov}(X, Y)}{\text{DP}(X) \text{DP}(Y)}$$

Propriedades:

(1) $\text{CORR}(X, Y) \in [-1, 1] \quad \forall t \in \mathbb{R}$

$$E\left[\left(t(X - E(X)) + (Y - E(Y))\right)^2\right] \geq 0.$$

$$t^2 E[(X - E(X))^2] + 2t E[(X - E(X))(Y - E(Y))] + E[(Y - E(Y))^2] \geq 0 \quad \forall t \in \mathbb{R}$$

$$\text{Var}(X)t^2 + 2\text{Cov}(X, Y)t + \text{Var}(Y) \geq 0$$



$$(2\text{Cov}(X, Y))^2 - 4\text{Var}(X)\text{Var}(Y) \leq 0$$

↔

$$\frac{4 \text{Cov}^2(X,Y)}{4 \text{Var}(X) \text{Var}(Y)} \leq 1 \Rightarrow (\text{CORR}(X,Y))^2 \leq 1 \Rightarrow -1 \leq \text{CORR}(X,Y) \leq 1.$$

(2) $\text{CORR}(ax, by) = \frac{ab}{|ab|} \text{CORR}(X,Y)$

(3) $\text{CORR}(X,Y) = 1 \Leftrightarrow \exists a > 0 \text{ tais que } P(Y = ax + b) = 1.$

(-1) ($a \neq 0$)

$\epsilon b \in \mathbb{R}$

Probabilidade e Inferência Estatística I

Aula 16 - 18/04/2013

Luz Gustavo Esteves

$$\text{CORR}(X, Y) = \frac{\text{COV}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

$$(1) \text{ CORR}(X, Y) \in [-1, 1]$$

$$(2) \text{ CORR}(aX, bY) = \frac{ab}{|ab|} \text{ CORR}(X, Y)$$

$$(3) \text{ CORR}(X, Y) = 1 \Leftrightarrow \begin{matrix} \exists a > 0 \\ \text{e} \\ b \in \mathbb{R} \end{matrix} \text{ tal que } D(Y = aX + b) = 1$$

Exemplo.

$$(X_1, X_2) \sim DIR(a_1, a_2, a_3)$$

Vemos que

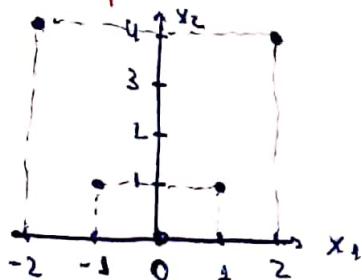
$$\begin{aligned} \text{Cov}(X_1, X_2) &= \frac{a_1 \cdot a_2}{(a_1 + a_2 + a_3)(a_1 + a_2 + a_3 + 1)} - \frac{a_1 a_2}{(a_1 + a_2 + a_3)^2} = \\ &= - \frac{a_1 a_2}{(a_0 + a_1 + a_2)^2 (a_0 + a_1 + a_2 + 1)} \end{aligned}$$

$$\text{Var}(X_1) = \frac{a_1 (a_0 + a_2)}{(a_0 + a_1 + a_2)^2 (a_0 + a_1 + a_2 + 1)}$$

$$\text{CORR}(X_1, X_2) = \frac{-\frac{\alpha_1 \alpha_2}{(\alpha_0 + \alpha_1 + \alpha_2)^2 (\alpha_0 + \alpha_1 + \alpha_2 + 1)}}{\sqrt{\frac{\alpha_1 (\alpha_0 + \alpha_2)}{\alpha_0 (\alpha_0 + \alpha_2)} \cdot \frac{\alpha_2 (\alpha_0 + \alpha_1)}{\alpha_1 (\alpha_0 + \alpha_1)}}}$$

$$\text{CORR}(X_1, X_2) = \frac{-\alpha_1 \alpha_2}{\sqrt{\alpha_1 \alpha_2 (\alpha_0 + \alpha_2)(\alpha_0 + \alpha_1)}}$$

Exemplo 2:



$$P(X_2 = X_1^2) = 1$$

$$\text{CORR}(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\text{Var}(X_1) \text{Var}(X_2)}$$

Note que $X_i \sim U\{-2, -1, 0, 1, 2\}$

$$\text{Cov}(X_1, X_2) = 0 \Rightarrow$$

$$\text{CORR}(X_1, X_2) = 0$$

Momentos Conjuntos

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

A quantidade $E(X_1^{n_1} X_2^{n_2} \cdots X_n^{n_n})$ damos o nome de ~~momento~~ momento conjunto

(n_1, n_2, \dots, n_n) - momento conjunto

Exemplo.

$$S_k = \{(b_1, \dots, b_n) \in \mathbb{R}_+^n : b_1 + b_2 + \dots + b_n \leq k\}$$

$(X_1, X_2, \dots, X_K) \sim \text{DIR}(\alpha_1, \dots, \alpha_K, \alpha_0)$

$$\begin{aligned} E(X_1^{n_1} \dots X_K^{n_K}) &= \int_{\Delta^k} x_1^{n_1} x_2^{n_2} \dots x_K^{n_K} \frac{\prod_{i=0}^k (\sum_{i=0}^k \alpha_i)}{\prod_{i=0}^k \prod_{i=1}^k \alpha_i} \prod_{i=1}^k x_i^{\alpha_i-1} \cdot (1 - x_1 - \dots - x_K)^{\alpha_0-1} dx \\ &= \Delta^k \int_{\Delta^k} \frac{\prod_{i=0}^k (\alpha_0 + \sum_{i=1}^k (\alpha_i + n_i))}{\prod_{i=0}^k \prod_{i=1}^k \alpha_i \cdot \prod_{i=1}^k (\alpha_i + n_i)} \prod_{i=1}^k x_i^{\alpha_i+n_i-1} \cdot (1 - x_1 - \dots - x_K)^{\alpha_0-1} dx \end{aligned}$$

$$\Rightarrow E(X_1^{n_1} \dots X_K^{n_K}) = \frac{\prod_{i=0}^k (\alpha_0 + \sum_{i=1}^k \alpha_i)}{\prod_{i=0}^k (\alpha_0 + \sum_{i=1}^k (\alpha_i + n_i))} \frac{\prod_{i=1}^k \prod_{i=1}^k (\alpha_i + n_i)}{\prod_{i=1}^k \prod_{i=1}^k \alpha_i}$$

$$X : \Omega \rightarrow \mathbb{R} \quad \psi_x(t) = E[e^{tx}]$$

t, x'

Para $X = (X_1, \dots, X_K) : \Omega \rightarrow \mathbb{R}^K$ vetor aleatório, definimos

$$\psi_x : I^4 \rightarrow \mathbb{R}$$

$$(t_1, \dots, t_K) \mapsto \psi_x(t_1, \dots, t_K) = E[e^{t_1 X_1 + t_2 X_2 + \dots + t_K X_K}] = E[e^{t \cdot x'}]$$

é chamada

FUNÇÃO GERADORA DE MOMENTOS CONJUNTA DE X .

$$I = (-\epsilon, \epsilon)^K$$

Propriedade.

$$\text{J. } E(X_1^{n_1} \cdots X_k^{n_k}) = \frac{\partial}{\partial t_1^{n_1} \partial t_2^{n_2} \cdots \partial t_k^{n_k}} \psi_x(t_1, \dots, t_k) \Big|_{(0,0, \dots, 0,0)}$$

Na última vez, calculamos

$$E(e^{t_1 X_1 + \cdots + t_k X_k}) \text{ para } X = (X_1, \dots, X_k) \sim \text{Mult}(n, p_1, \dots, p_k)$$

$$E(e^{t_1 X_1 + \cdots + t_k X_k}) = (p_1 e^{t_1} + \cdots + p_k e^{t_k} + 1 - p_1 - \cdots - p_k)^n$$

$$\psi_x(t_1, \dots, t_k)$$

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i) E(X_j)$$

$$= \frac{\partial^2 \psi_x(t_1, \dots, t_k)}{\partial t_i \partial t_j} \Big|_{(0,0, \dots, 0)} - \left(\frac{\partial \psi_x(t_1, \dots, t_k)}{\partial t_i} \Big|_{(0,0, \dots, 0)} \right) \left(\frac{\partial \psi_x(t_1, \dots, t_k)}{\partial t_j} \Big|_{(0,0, \dots, 0)} \right) =$$

$$= n(n-1)p_i p_j - n p_i \cdot n p_j \Rightarrow \text{Cov}(X_i, X_j) = -n p_i p_j$$

$$\frac{\partial \psi_x(t_1, \dots, t_k)}{\partial t_i} = n(p_1 e^{t_1} + \cdots + p_k e^{t_k} + 1 - p_1 - \cdots - p_k)^{n-1} p_i e^{t_i}$$

Analogamente,

$$\frac{\partial^2 \psi_x(t_1, \dots, t_k)}{\partial t_i \partial t_j} = n(n-1) (p_1 e^{t_1} + \cdots + p_k e^{t_k} + 1 - p_1 - \cdots - p_k)^{n-2} p_i p_j e^{t_i} e^{t_j}$$

Exercício:

$$X = (X_1, X_2) \sim \text{Holgate}(\lambda_1, \lambda_2, \lambda_3)$$

$$E(e^{t_1 X_1 + t_2 X_2}) = \varphi_X(t_1, t_2).$$

Algumas exemplos adicionais:

$$X \sim N(0,1)$$

$$\varphi_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx =$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2 - 2tx + t^2)} e^{tx} dx =$$

$$= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} dx \Rightarrow \varphi_X(t) = e^{t^2/2}.$$

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

$$Y \stackrel{d}{=} cX + \mu$$

$$\varphi_Y(t) = \varphi_{cX + \mu}(t) = e^{t\mu} \varphi_X(ct) = e^{t\mu} \cdot e^{\frac{c^2 t^2}{2}},$$

sejamos X_1, \dots, X_n v.a.i.i.d. $N(0,1)$

$$a_1, a_2, \dots, a_n \in \mathbb{R} \text{ e } \mu \in \mathbb{R}$$

$$Y = a_1 X_1 + a_2 X_2 + \dots + a_n X_n + \mu$$

$$\varphi_Y(t) = \varphi_{\sum_{i=1}^n a_i X_i + \mu}(t) = e^{t\mu} \varphi_{\sum_{i=1}^n a_i X_i}(t) = e^{t\mu} \prod_{i=1}^n \varphi_{a_i X_i}(t) = e^{t\mu} \prod_{i=1}^n \varphi_{a_i}(at) = e^{t\mu} \prod_{i=1}^n e^{\frac{a_i^2 t^2}{2}}$$

$$\Rightarrow \Psi_Y(t) = e^{t\mu} \cdot e^{\left(\frac{1}{2}\sum_{i=1}^k a_i^2\right)t^2/2}$$

$$Y \sim N(\mu, \sum_{i=1}^k a_i^2)$$

(4)

X_1, \dots, X_k são independentes com $X_i \sim \text{Gama}(a_i, b)$, $a_1, \dots, a_k, b > 0$.

$$Y = X_1 + \dots + X_k$$

$$\Psi_Y(t) = \Psi_{X_1 + \dots + X_k} = \prod_{i=1}^k \Psi_{X_i}(t) = \prod_{i=1}^k \left[\frac{b}{b-t} \right]^{a_i} = \left[\frac{b}{b-t} \right]^{\sum_{i=1}^k a_i}$$

Lego, $Y \sim \text{Gama}(\sum_{i=1}^k a_i, b)$.

Resultado: (Ω, \mathcal{F})

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

$P_1, P_2: \mathcal{F} \rightarrow [0, 1]$ probabilidades, $\alpha \in (0, 1)$

$$\int_{\Omega} X d(\alpha P_1 + (1-\alpha) P_2) = \alpha \int_{\Omega} X dP_1 + (1-\alpha) \int_{\Omega} X dP_2.$$

X simples

$$X = \sum_{i=1}^k a_i \mathbb{I}_{A_i}, \quad a_1, a_k \geq 0, A_1, A_2, \dots, A_k \in \mathcal{F}$$

$$\int_{\Omega} X d(\alpha P_1 + (1-\alpha) P_2) = \sum_{i=1}^k a_i (\alpha P_1(A_i) + (1-\alpha) P_2(A_i)) =$$

$$= \alpha \sum_{i=1}^k a_i P_1(A_i) + (1-\alpha) \sum_{i=1}^k a_i P_2(A_i) =$$

$$= \alpha \int_{\Omega} X dP_1 + (1-\alpha) \int_{\Omega} X dP_2.$$

$X \geq 0$

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 $X_n \uparrow X$.

$$\begin{aligned} \int_{\Omega} X d(\alpha P_1 + (1-\alpha) P_2) &= \lim_{n \rightarrow \infty} \int_{\Omega} X_n d(\alpha P_1 + (1-\alpha) P_2) = \\ &= \lim_{n \rightarrow \infty} \left[\alpha \int_{\Omega} X_n dP_1 + (1-\alpha) \int_{\Omega} X_n dP_2 \right] = \\ &= \alpha \int_{\Omega} X dP_1 + (1-\alpha) \int_{\Omega} X dP_2. \end{aligned}$$

Resultado: (Ω, \mathcal{F}) .

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

$P_1, P_2, f: \Omega \rightarrow [0, 1]$ probabilidades, $\alpha \in (0, 1)$

$$\int_{\Omega} X d(\alpha P_1 + (1-\alpha) P_2) = \alpha \int_{\Omega} X dP_1 + (1-\alpha) \int_{\Omega} X dP_2$$

Suponhamos que existe $A \in \mathcal{B}(\mathbb{R})$, A enumerável, tal que $P_1(X \in A) = 1$ e que existe

$$\phi_2: \mathbb{R} \rightarrow \mathbb{R}_+ \text{ tal que } P_2(X \in B) = \int_B \phi_2(t) dt,$$

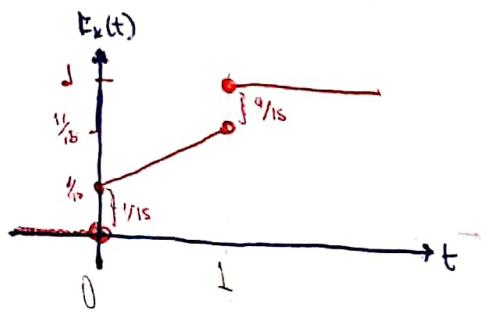
$B \in \mathcal{B}(\mathbb{R})$.

Nesse caso,

$$\mathbb{E}(X) = \int_{\Omega} X d(\alpha P_1 + (1-\alpha) P_2) = \alpha \sum_{x \in A} x P_1(X=x) + (1-\alpha) \int_{-\infty}^{\infty} x f_2(x) dx$$

Exemplo:

$$F_X(t) = \begin{cases} 0, & t \leq 0 \\ \frac{1}{15} + \frac{2}{3}t, & 0 \leq t \leq 1 \\ 1, & t \geq 1 \end{cases}$$



$$\frac{1}{3} \begin{cases} 0, & t \leq 0 \\ \frac{1}{3}, & 0 \leq t < 1 \\ \frac{2}{3}, & t \geq 1 \end{cases} + \begin{cases} 0, & t \leq 0 \\ t, & 0 \leq t \leq 1 \\ 1, & t \geq 1 \end{cases}$$

$$F_X(t) = \alpha F_d(t) + (1-\alpha) F_e(t)$$

$$\downarrow \quad \downarrow$$

$$P_1(X \in (-\infty, t]) \quad P_2(X \in (-\infty, t))$$

$$A = \{0, 1\} \quad P_i(X \in A) = 1$$

$$F_d(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{15}, & 0 \leq t < 1 \\ 1, & t \geq 1 \end{cases}$$

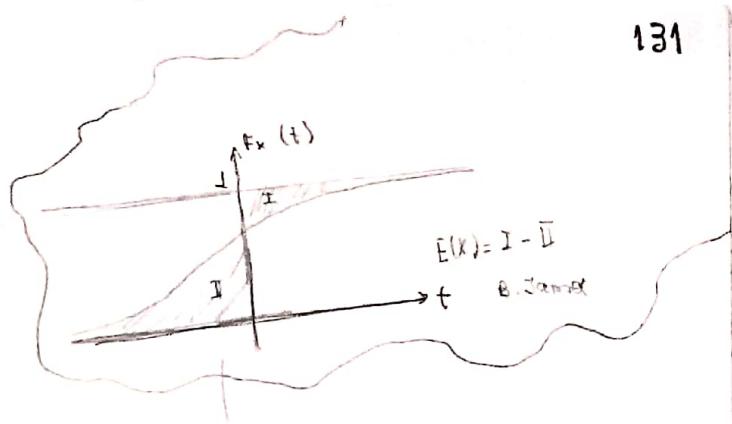
$$F_e(t) = P_2(X \in (-\infty, t)) = \begin{cases} 0, & t \leq 0 \\ t, & 0 \leq t \leq 1 \\ 1, & t \geq 1 \end{cases}$$

$$f_2(x) = \mathbb{I}_{(0,1)}(x)$$

No exemplo,

$$\mathbb{E}(X) = \alpha \sum_{x \in A} x P(X=x) + (1-\alpha) \int_{-\infty}^{\infty} x f_2(x) dx =$$

$$= \frac{1}{3} \left[0 \cdot \frac{1}{5} + 1 \cdot \frac{4}{5} \right] + \frac{2}{3} \int_0^1 x \cdot 1 \, dx = \\ = \frac{1}{3} \cdot \frac{4}{5} + \frac{2}{3} \cdot \frac{1}{2} = \frac{9}{15}.$$



Último comentário.

1. Quando $\text{Cov}(X, Y) = 0$ (ou $\text{CORR}(X, Y) = 0$), dizemos que X e Y são NÃO-CORRELACIONADOS
2. X_1 e X_2 são independentes \Rightarrow

$$E(X_1 X_2) = E(X_1) E(X_2) \Rightarrow$$

$$\Rightarrow E(X_1 X_2) - E(X_1) E(X_2) = 0 \Rightarrow X_1 \text{ e } X_2 \text{ são não-correlacionados.}$$

No entanto,

X_1 e X_2 não correlacionadas $\not\Rightarrow X_1$ e X_2 são independentes

$$P(X_1 = -1, X_2 = 1) = 1/4$$

$$P(X_1 = 1, X_2 = 1) = 1/4$$

$$P(X_1 = 0, X_2 = 1) = 1/2$$

$$P(X_1 = -1) = 1/4. \quad \text{Claro que}$$

$$P(X_2 = 1) = 1/2. \quad P(X_1 = -1, X_2 = 1) \neq P(X_1 = -1) P(X_2 = 1)$$

Sendo $X : \Omega \rightarrow \mathbb{R}$ uma v.a.

$$P_{X|B} : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$$

$$A \in \mathcal{B}(\mathbb{R}) \mapsto P_{X|B}(A) = \frac{P(X^{-1}(A) \cap B)}{P(B)} = \frac{P((X \in A) \cap B)}{P(B)}$$

Distribuição condicional de X dado B .

Vamos verificar que $P_{X|B}$ é, de fato, medida de probabilidade em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

$$(1) P_{X|B}(\mathbb{R}) = \frac{P(X^{-1}(\mathbb{R}) \cap B)}{P(B)} = \frac{P(\mathbb{R} \cap B)}{P(B)} = 1$$

(2) $(A_n)_{n \geq 1}$, tais que $A_n \in \mathcal{B}(\mathbb{R})$, $\forall n \in \mathbb{N}$, e $A_i \cap A_j = \emptyset$, $i \neq j$.

$$\begin{aligned} P_{X|B}\left(\bigcup_{n=1}^{\infty} A_n\right) &= \frac{P(X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) \cap B)}{P(B)} = \frac{P\left(\left(\bigcup_{n=1}^{\infty} X^{-1}(A_n)\right) \cap B\right)}{P(B)} = \\ &= \frac{P\left(\bigcup_{n=1}^{\infty} (B \cap X^{-1}(A_n))\right)}{P(B)} = \sum_{n=1}^{\infty} \frac{P(X^{-1}(A_n) \cap B)}{P(B)} = \sum_{n=1}^{\infty} P_{X|B}(A_n). \end{aligned}$$

De (1) e (2), $P_{X|B}$ é de fato probabilidade.

Analogamente, podemos

$$\begin{aligned} F_{X|B}: \mathbb{R} &\rightarrow [0,1] \\ t \in \mathbb{R} &\mapsto F_{X|B}((-\infty, t]) \end{aligned}$$

Função de Distribuição Condicional de X dado B.

Na proximavem...

$$Y: \Omega \rightarrow \mathbb{R}$$

$$Y \text{ discreta } P(Y \in \{y_1, y_2, \dots\}) = 1$$

$$B_i = \{Y = y_i\}$$

$$P_{X|B_i}$$

Probabilidade e Inferência Estatística I

19/04/2013. Aula 17

Luiz Gustavo Esteves

(Ω, \mathcal{F}, P)

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

$X | |X|=t, t \geq 0$

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

$P_{x|B}: \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$

$$A \in \mathcal{B}(\mathbb{R}) \mapsto P_{x|B}(A) = \frac{P(X^{-1}(A) \cap B)}{P(B)}$$

$F_{x|B}: \mathbb{R} \rightarrow [0, 1]$

$$t \in \mathbb{R} \mapsto F_{x|B}(t) = P_{x|B}((-\infty, t])$$

$Y: \Omega \rightarrow \mathbb{R}$ é outra v.a. discreta, tal que $P(Y \in \{y_1, y_2, \dots\}) = 1$ e

$$P(Y_i = y_i) > 0, i = 1, 2, \dots$$

$$B_i = \{Y_i = y_i\}$$

$P_{x|B_i}$: dist. condicional de X dado $Y = y_i$.

$F_{x|B_i}$: função de distribuição condicional de X dado $Y = y_i$.

$$\downarrow F_{X|Y=y_i}.$$

Se $X: \Omega \rightarrow \mathbb{R}$ é também discreto, (X, Y) é discreto.

Logo, $P_{X|Y=y_i}$ é tal que $\exists A \in \mathcal{B}(\mathbb{R})$, A enumerável, com $P_{X|Y=y_i}(A) = 1$.

Para cada possível valor ω_i de X , podemos avaliar

$$P_{X/Y=y_3}(\{\omega_i\}) = \frac{P(X^{-1}(\{\omega_i\}) \cap (Y=y_3))}{P(Y=y_3)} \Rightarrow$$

$$\Rightarrow P_{X/Y=y_3}(\{\omega_i\}) = \frac{P(X=\omega_i, Y=y_3)}{P(Y=y_3)}$$

Exemplo 1:

$$P(X=\omega, Y=y) = \frac{1}{2y+1} \binom{10}{y} \left(\frac{1}{2}\right)^y \underset{\{-y, \dots, 0, \dots, y\}}{\overset{10}{\text{II}}}(\omega) \underset{\{0, \dots, 10y\}}{\overset{10}{\text{II}}}(\gamma)$$

$$y \in \{0, \dots, 10\}$$

$$P(X=\omega | Y=y) = \frac{P(X=\omega, Y=y)}{P(Y=y)} = \frac{\frac{1}{2y+1} \binom{10}{y} \left(\frac{1}{2}\right)^y \underset{\{-y, \dots, 0, \dots, y\}}{\overset{10}{\text{II}}}(\omega) \underset{\{0, \dots, 10y\}}{\overset{10}{\text{II}}}(\gamma)}{P(Y=y)}$$

Mas

$$P(Y=y) = \sum_{\omega} P(X=\omega, Y=y) = \sum_{x=-y}^y \frac{1}{2y+1} \binom{10}{y} \left(\frac{1}{2}\right)^y \underset{\{0, \dots, 10y\}}{\overset{10}{\text{II}}}(\gamma) = \\ = \binom{10}{y} \left(\frac{1}{2}\right)^y \underset{\{0, \dots, 10y\}}{\overset{10}{\text{II}}}(\gamma) \quad \therefore Y \sim \text{Binomial}(10, 1/2).$$

Dai,

$$P(X=\omega | Y=y) = \frac{\frac{1}{2y+1} \binom{10}{y} \left(\frac{1}{2}\right)^y \underset{\{-y, \dots, 0, \dots, y\}}{\overset{10}{\text{II}}}(\omega) \underset{\{0, \dots, 10y\}}{\overset{10}{\text{II}}}(\gamma)}{\binom{10}{y} \left(\frac{1}{2}\right)^y \underset{\{0, \dots, 10y\}}{\overset{10}{\text{II}}}(\gamma)} = \frac{1}{2y+1} \underset{\{-y, \dots, 0, \dots, y\}}{\overset{10}{\text{II}}}(\omega)$$

$\therefore X | Y=y \sim \text{Unif}(\{-y, \dots, 0, \dots, y\})$

∞	$P(X=\infty Y=y)$
-y	$1/(2y+1)$
:	:
-1	
0	$y/(2y+1)$
1	
:	:
y	$1/(2y+1)$

Dai,

$$E(X|Y=y) = \sum_{x=-\infty}^{\infty} x P(X=x | Y=y) = 0$$

$$\text{Var}(X|Y=y) = \sum_{x=-\infty}^{\infty} (x - E(X|Y=y))^2 P(X=x | Y=y) =$$

$$= \sum_{x=-y}^y x^2 \frac{1}{2y+1} = \frac{1}{2y+1} 2 \sum_{i=1}^y i^2 =$$

$$= \frac{2}{2y+1} \frac{y(y+1)(2y+1)}{6} \Rightarrow$$

$$\text{Var}(X|Y=y) = \frac{y(y+1)}{3}$$

Exemplo 2:

$$P(X=x, Y=y) = \frac{e^{-\lambda_1} \lambda_1^x}{x!} \cdot \frac{e^{-\lambda_2} \lambda_2^{y-x}}{(y-x)!} \prod_{\substack{x \\ \{0, 1, \dots, y\}}}^N \prod_{\substack{y \\ \{N+1, \dots\}}}$$

$$P(Y=y) = \sum_{x=0}^y \frac{e^{-\lambda_1} \lambda_1^x}{x!} \frac{e^{-\lambda_2} \lambda_2^{y-x}}{(y-x)!} \prod_N^y =$$

$$= e^{-(\lambda_1 + \lambda_2)} \prod_N^y \frac{1}{y!} \sum_{x=0}^y \frac{y!}{x!(y-x)!} \frac{\lambda_1^x \lambda_2^{y-x}}{x!(y-x)!} \Rightarrow$$

$$\Rightarrow P(Y=y) = \frac{e^{-(\lambda_1 + \lambda_2)}}{y!} \frac{(\lambda_1 + \lambda_2)^y}{y!} \prod_N^y$$

$$\therefore Y \sim \text{Doi}(\lambda_1 + \lambda_2)$$

Dai,

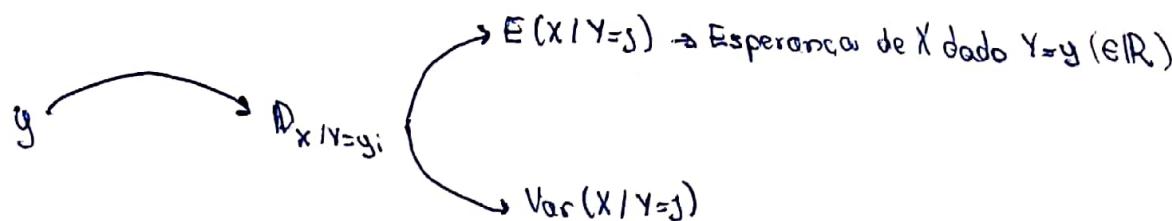
$y \in \mathbb{N}$

$$P(X=x | Y=y) = \frac{\frac{e^{-\lambda_1} \lambda_1^x}{x!} \frac{e^{-\lambda_2} \lambda_2^{y-x}}{(y-x)!} \prod_{i=0, \dots, y} \prod_{j=1}^N \frac{I_j(x)}{I_j(y)}}{\frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1+\lambda_2)^y}{y!} I_N(y)} \Rightarrow$$

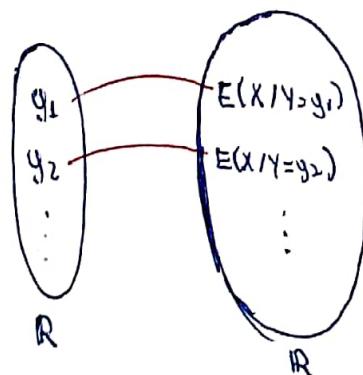
$$P(X=x | Y=y) = \binom{y}{x} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{y-x} \prod_{i=0, \dots, y} \frac{I_i(x)}{I_i(y)} \quad \therefore X|Y=y \sim \text{Bin}(y, \frac{\lambda_1}{\lambda_1+\lambda_2})$$

$$E(X|Y=y_i) = y_i \cdot \frac{\lambda_1}{\lambda_1+\lambda_2}.$$

$$\text{Var}(X|Y=y_i) = \sum_{\infty} (x - E(X|Y=y_i))^2 P(X=x | Y=y) = y_i \frac{\lambda_1}{\lambda_1+\lambda_2} \cdot \frac{\lambda_2}{\lambda_1+\lambda_2}.$$



Note que para cada $y \in \mathbb{R}$ tal que $P(Y=y) > 0$, posso associar o número $E(X|Y=y)$.



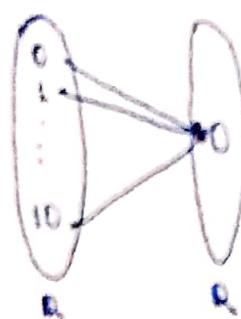
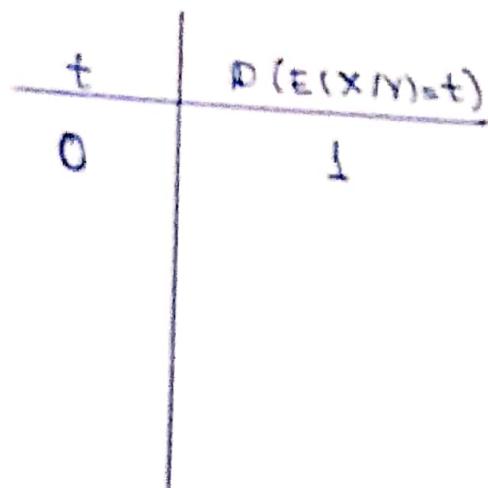
$$g: \mathbb{R} \rightarrow \mathbb{R}$$

$$y \in \mathbb{R} \rightarrow g(y) = E(X|Y=y)$$

$$g(Y) = E(X|Y)$$

Esperança condicional de X dado Y .

No exemplo 1.



No exemplo 2.

t	$P(E(X/Y)=t)$
0	$e^{-(\lambda_1 + \lambda_2)}$
$\frac{\lambda_1}{\lambda_1 + \lambda_2}$	$e^{\frac{(\lambda_1 + \lambda_2)}{\lambda_1} - (\lambda_1 + \lambda_2)}$
$\frac{2\lambda_1}{\lambda_1 + \lambda_2}$	$e^{\frac{-(\lambda_1 + \lambda_2)}{\lambda_1} - (\lambda_1 + \lambda_2)^2}$
\vdots	\vdots
$\frac{i\lambda_1}{\lambda_1 + \lambda_2}$	$e^{\frac{-(\lambda_1 + \lambda_2)}{\lambda_1} - (\lambda_1 + \lambda_2)^i}$

$$E(X/Y) = y \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

$$E(X/Y) = \frac{\lambda_1}{\lambda_1 + \lambda_2} Y$$

$$P(E(X/Y) = i\lambda_1 / (\lambda_1 + \lambda_2)) = P\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} Y = i \frac{\lambda_1}{\lambda_1 + \lambda_2}\right) =$$

$$= P(Y=i) = \frac{e^{-\lambda_1 - \lambda_2}}{i!} \frac{(\lambda_1 + \lambda_2)^i}{i!}$$

$f_y: \mathbb{R} \rightarrow \mathbb{R}$

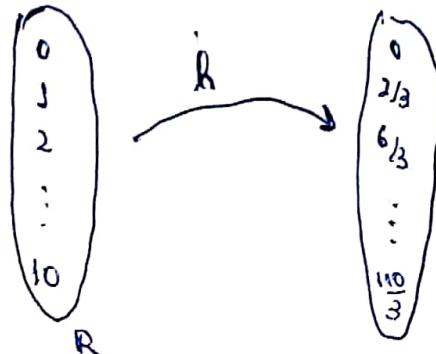
$y \in \mathbb{R} \mapsto h(y) = \text{Var}(X/Y=y) \Rightarrow$ Variância Condisional de X dado $Y=y$

$h(Y) = \text{VAR}(X/Y) \Rightarrow$ Variância Condisional de X dado Y

No exemplo 1

$$\text{Var}(X|Y=y) = \frac{y(y+1)}{3}$$

w	$P(\text{Var}(X Y=y) \neq w)$
0	$\binom{10}{0} \left(\frac{1}{2}\right)^{10}$
$\frac{2}{3}$	$\binom{10}{1} \left(\frac{1}{2}\right)^{10}$
$\frac{6}{3}$	$\binom{10}{2} \left(\frac{1}{2}\right)^{10}$
:	:
$\frac{10}{3}$	$\binom{10}{10} \left(\frac{1}{2}\right)^{10}$



No exemplo 1:

$$E[E(X|Y)] = 0 \cdot P(E(X|Y)=0) = 0 \cdot 1 = 0.$$

$$\text{Var}[E(X|Y)] = 0.$$

$$E[\underbrace{E(X|Y)}_{g(Y)}] = \sum_y E(X|Y=y) P(Y=y) =$$

$$= \sum_y \left\{ \sum_x x P(X=x|Y=y) P(Y=y) \right\} =$$

$$= \sum_y \left\{ \sum_x \overbrace{x P(X=x, Y=y)}^{I(x,y)=x} \right\} = \sum_x \left\{ \sum_y x P(X=x, Y=y) \right\} =$$

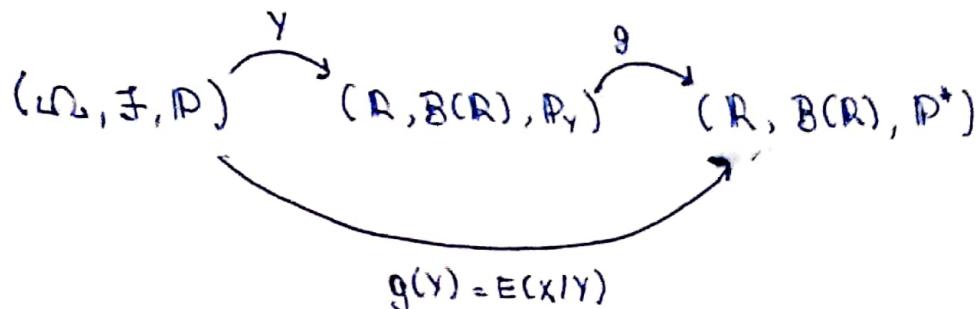
$$= \sum_x x P(X=x) = E(X)$$

Quando

BB

$$E(X) < \infty,$$

$$E(E(X|Y)) = E(X).$$



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

$$(g(Y))^{-1}(B) = Y^{-1}(g^{-1}(B)) \in \sigma(Y)$$

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

$$E(X_n)$$

$$E(X_n | X_{n-1} = c) = (2c) \cdot p + (c/2) \cdot (1-p) \Rightarrow$$

$$\Rightarrow E(X_n | X_{n-1} = c) = \frac{c}{2} + p \left\{ 2c - \frac{c}{2} \right\} = \boxed{\cancel{c}} = c \left\{ \frac{1}{2} + \frac{3}{2} p \right\} = c \left(\frac{3p+1}{2} \right)$$

$$E(X_n | X_{n-1}) = X_{n-1} \left(\frac{3p+1}{2} \right)$$

189

$$g(X_{n-1})$$

Como se trata de uma v.a. discreta,

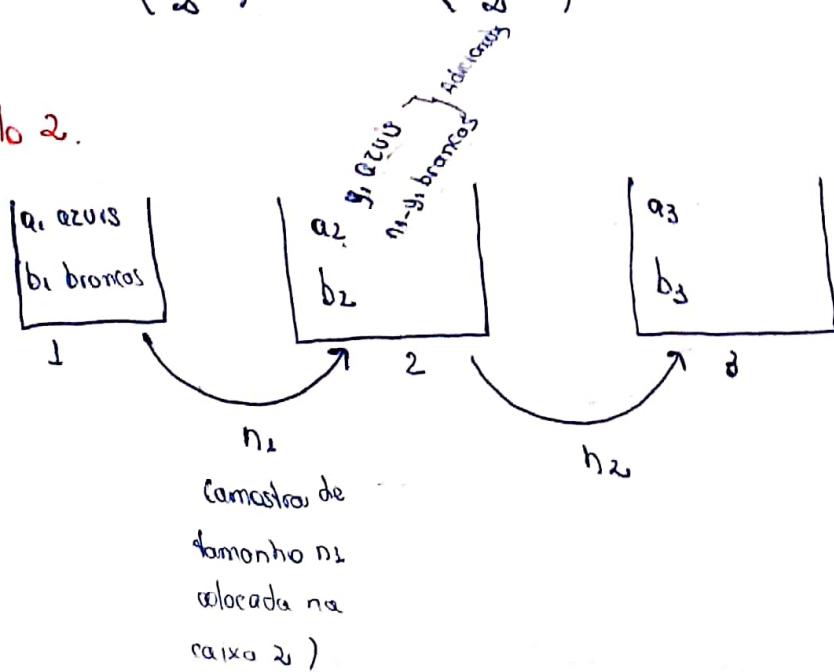
$$E(X_n) = E[E(X_n | X_{n-1})] \Rightarrow$$

$$E(X_n) = E\left(\frac{3p+1}{2} X_{n-1}\right) \Rightarrow$$

$$\Rightarrow \overbrace{E(X_n)}^{c_n} = \frac{\overbrace{c_{n-1}}^{c_{n-1}}}{2} E(X_{n-1})$$

$$\Rightarrow E(X_n) = \left(\frac{3p+1}{2}\right)^n E(X_0) = \left(\frac{3p+1}{2}\right)^n c_0$$

Exemplo 2.



X_i : Número de bolas azuis em \hat{z} depois das duas extrações

y_i : N° de azuis retiradas da urna i e colocadas na urna $i+2$, $i=1, 2$.

$$X = Y_2 + \alpha_3$$

$$E(X) = E(Y_2) + \alpha_3$$

$$\text{Var}(X) = \text{Var}(Y_2)$$

$$Y_1 \sim HG(a_1, b_1, n_1)$$

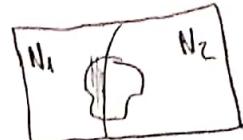
$$Y_2 | Y_1 = y_1 \sim HG(a_2 + y_1, b_2 + n_1 - y_1, n_2)$$

Ex.

$$X \sim HG(n_1, n_2, n)$$

$$E(X) = n \frac{n_1}{n_1 + n_2}$$

$$\text{Var}(X) = \dots$$



$$X = Y_1 + \dots + Y_{N_2}$$

$$Y_1 \sim HG(a_1, b_1, n_1)$$

$$E(Y_2 | Y_1 = y_1) = n_2 \frac{a_2 + y_1}{a_2 + b_2 + n_2}$$

$$E(Y_2) = E(E(Y_2 | Y_1)) =$$

$$= E\left(n_2 \cdot \frac{a_2 + Y_1}{a_2 + b_2 + n_1}\right) = \frac{n_2 a_2}{a_2 + b_2 + n_1} + \frac{n_2}{a_2 + b_2 + n_1} E(Y_1) \Rightarrow$$

$$\Rightarrow E(Y_2) = \frac{n_2 a_2}{a_2 + b_2 + n_1} + \frac{n_2}{a_2 + b_2 + n_1} n_1 \frac{a_1}{a_1 + b_1} \times a_3$$

$$\text{Var}(X) = \sum_x (\infty - E(\infty))^2 P(X=\infty) =$$

$$= \sum_x (\infty - E(\infty))^2 \sum_y P(X=x, Y=y)$$

~~Summation over all x and y~~

$$= \sum_x \sum_y (\infty - E(X|Y=y) + E(X|Y=y) - E(x))^2 P(X=x, Y=y)$$

$$\begin{aligned}
 & \sum_y \sum_x (x - E(X|Y=y))^2 P(X=x \cap Y=y) P(Y=y) = \\
 &= \sum_y P(Y=y) \sum_x \left\{ (x - E(X|Y=y))^2 + (E(X|Y=y) - E(X))^2 + 2(x - E(X|Y=y))(E(X|Y=y) - E(X)) \right\} P(X=x \cap Y=y) = \\
 &= \sum_y P(Y=y) \left\{ \text{Var}(X|Y=y) + (E(X|Y=y) - E(X))^2 + 0 \right\} \Rightarrow \\
 &\Rightarrow \text{Var}(X) = \sum_y P(Y=y) \left\{ \overbrace{\text{Var}(X|Y=y)}^{h(y)} + (E(X|Y=y) - E(X))^2 \right\} = \\
 &\qquad\qquad\qquad \downarrow \\
 &\qquad\qquad\qquad (g(y) - E(g(y)))^2 \\
 &= E[\text{Var}(X|Y)] + \text{Var}[E(X|Y)]
 \end{aligned}$$

Voltando ao Exemplo 1 da aula.

$X|Y=y$ N Uniforme ($\{-y, \dots, 0, \dots, y\}$)

$$E(X|Y=y) = 0 \rightarrow E(X|Y) = 0$$

$$\text{Var}(X|Y=y) = \frac{y(y+1)}{3} \rightarrow \text{Var}(X|Y) = \frac{y(y+1)}{3}$$

$$\text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}[E(X|Y)] =$$

$$\frac{1}{3} E \left[\frac{y(y+1)}{3} \right] + \underbrace{\text{Var}(0)}_0 \Leftarrow$$

$$\text{Var}(X) = \frac{1}{3} E(Y+Y^2) = \frac{1}{3} E(Y) + \frac{1}{3} \left\{ (E(Y))^2 + \text{Var}(Y) \right\} =$$

$$= \frac{1}{3} \cdot 10 \cdot \frac{1}{2} + \frac{1}{3} \left(10 \cdot \frac{1}{2} \right)^2 + \frac{1}{3} \left(10 \cdot \frac{1}{2} \cdot \frac{1}{2} \right) =$$

$$= \frac{10}{6} + \frac{50}{6} + \frac{5}{6} = \frac{65}{6} \text{ "}$$

Probabilidade e Inferência Estatística I

23/04/2013 - Aula 18

Luis Gustavo Esteves

$$X \sim N(0,1)$$

$$|x|=3 \quad \begin{cases} x=3 \\ x=-3 \end{cases}$$

$$X / |x|=3$$

Na aula passada, (X,Y) discreto

$$F_{X,Y}(x,y) = \sum_{y_j \leq y} \sum_{x_i \leq x} P(X=x_i, Y=y_j) = \sum_{y_j \leq y} \left\{ \underbrace{\sum_{x_i \leq x} P(X=x_i | Y=y_j)}_{P(X \leq x | Y=y_j)} \right\} P(Y=y_j)$$

$$P(X \leq x | Y=y_j)$$

$$F_{X/Y=y_j}(x)$$

"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

$$\forall A \in \mathcal{B}(\mathbb{R}), \forall B \in \mathcal{B}(\mathbb{R}),$$

$$P(X \in A, Y \in B) = \int_B \underbrace{h(A,y)}_{\text{distribuição condicional de } X \text{ dado } Y=t} dP_Y$$

" $P(X \in A | Y=y)$

distribuição condicional de X dado $Y=t$.

Essa função é tal que

- fixado $y \in \mathbb{R}$, $\lambda(\cdot, y) : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ é medida de probabilidade.
- fixado $A \in \mathcal{B}(\mathbb{R})$, $\lambda(A, \cdot)$ é função $\mathcal{B}(\mathbb{R})$ -mensurável.

Para $A = (-\infty, x]$ e $B = (-\infty, y]$, $x, y \in \mathbb{R}$, temos

Função de Distribuição Condicional de X dado Y .

$$\mathbb{P}(X \in A, Y \in B) = F_{X,Y}(x, y) = \int_{(-\infty, y]} \lambda((- \infty, x], t) dP_y$$

↓

em geral, é difícil achar
h dessa forma.

CASOS FÁCEIS

(1) (X, Y) discreto

(2) (X, Y) é absolutamente contínuo

$$F_{(X,Y)}(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^y \int_{-\infty}^x \frac{f_{X,Y}(w, v)}{f_Y(v)} dw f_Y(v) dv =$$

$$= \int_{-\infty}^y \left[\int_{-\infty}^x \frac{f_{X,Y}(w, v)}{f_Y(v)} dw \right] f_Y(v) dv$$

→ função densidade de probabilidade condicional de X dado $Y=v$:
 $f_{X|Y=v}(w)$.
 $X|Y=v$ é abs. contínua

Função De Distribuição

Condisional de X Dado $Y=v$: $F_{X|Y=v}$

Dai,

$$E(X|Y=y) = \int_{-\infty}^{\infty} x \cdot f_{X|Y=y}(x) dx$$

$$E[g(x)|Y=y] = \int_{-\infty}^{\infty} g(x) f_{X|Y=y}(x) dx$$

Exemplos:

(i) $f_{X,Y}(x,y) = \frac{1}{y} e^{-y} \mathbb{I}_{(0,y)}(x) \mathbb{I}_{\mathbb{R}_+}(y)$. $X|Y=y \sim ?$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \int_{-\infty}^{\infty} \frac{1}{y} e^{-y} \mathbb{I}_{(0,y)}(x) \mathbb{I}_{\mathbb{R}_+}(y) dx =$$

$$= \frac{1}{y} e^{-y} \mathbb{I}_{\mathbb{R}_+}(y) \int_0^y 1 dx = \frac{1}{y} e^{-y} \mathbb{I}_{\mathbb{R}_+}(y) \cdot y \Rightarrow f_Y(y) = e^{-y} \mathbb{I}_{\mathbb{R}_+}(y)$$

$\gamma_N \exp(1).$

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{1}{y} e^{-y} \mathbb{I}_{(0,y)}(x)}{e^{-y}} = \frac{1}{y} \mathbb{I}_{(0,y)}(x)$$

$$X|Y=y \sim U(0,y)$$

$$E(X|Y=y) = \frac{y}{2}$$

$$\text{Var}(X|Y=y) = \frac{y^2}{12}$$

Agora,

$$E(X|Y) = Y/2 . \text{ Da}(\ E(E(X|Y))) = E(Y/2) = \frac{1}{2} E(Y) = \frac{1}{2} .$$

Exercício: Calcular $E(X)$ e verificar que $E(X) = 1/2$.

$$E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{x,y}(x,y) dx dy = \int_0^{\infty} \int_0^y x + e^{-y} dx dy = \int_0^{\infty} \frac{1}{2} e^{-y} \cdot \frac{y^2}{2} dy = \int_0^{\infty} \frac{ye^{-y}}{2} dy = \frac{1}{2} \int_0^{\infty} y e^{-y} dy = \frac{1}{2}$$

Note que $E(E(X,Y)) = E(X)$.

Exemplo 2.

$$f_{x,y}(x,y) = \frac{1}{2\pi} |y| e^{-\frac{|y|^2}{2}(1+x^2)}$$

$$f_Y(y) = \int_{-\infty}^{\infty} \frac{1}{2\pi} |y| e^{-\frac{|y|^2}{2}(1+x^2)} dx = \frac{|y|}{2\pi} e^{-\frac{|y|^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{|y|^2}{2}x^2} dx =$$

$$= \frac{|y|}{2\pi} e^{-\frac{|y|^2}{2}} \sqrt{\frac{2\pi}{|y|^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\frac{2\pi}{|y|^2}}} e^{-\frac{x^2}{\frac{2\pi}{|y|^2}}} dx \Rightarrow f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{|y|^2}{2}}$$

$\therefore Y \sim N(0,1)$

$$f_{X/Y=y}(x) = \frac{\frac{1}{2\pi} |y| e^{-\frac{|y|^2}{2}} \cdot e^{\frac{-x^2 y^2}{2}}}{\frac{1}{\sqrt{2\pi}} e^{-\frac{|y|^2}{2}}} \Rightarrow$$

$$\Rightarrow f_{X/Y=y}(x) = \frac{|y|}{\sqrt{2\pi}} e^{-\frac{x^2 y^2}{2}} \Rightarrow f_{X/Y=y}(x) = \frac{1}{\sqrt{2\pi \cdot 1/y^2}} e^{-\frac{x^2}{2 \cdot \frac{1}{y^2}}}$$

$\therefore X/Y = y \sim N(0, 1/y^2), y \neq 0$.

$$E(X|Y=y) = 0$$

$$\text{Var}(X|Y=y) = \frac{1}{y^2}$$

$E(X|Y)$ é degenerada $\Rightarrow P(E(X|Y)=0)=1 \Rightarrow E(E(X|Y))=0.$

$$f_X(x) = \int_{-\infty}^{\infty} f_{x,y}(x,y) dy = \int_{-\infty}^{\infty} \frac{1}{2\pi} |y| e^{-\frac{|y|^2(x^2+1)}{2}} dy =$$

$$= \int_{-\infty}^{0} \frac{1}{2\pi} (-y) e^{-\frac{(-y)^2(x^2+1)}{2}} dy + \int_0^{\infty} y \frac{1}{2\pi} e^{-\frac{y^2(x^2+1)}{2}} dy =$$

$$\frac{1}{2\pi} \frac{1}{x^2+1} e^{-\frac{x^2+1}{2}} \left[-\frac{1}{2\pi} \frac{1}{x^2+1} e^{-\frac{y^2(x^2+1)}{2}} \right]_0^{\infty} =$$

$$= \frac{1}{2\pi(x^2+1)} + \frac{1}{2\pi(x^2+1)} \Rightarrow f_X(x) = \frac{1}{\pi(1+x^2)} \quad X \sim \text{Cauchy Padrão}$$

$$E(X^+) = \int_0^{\infty} x \frac{1}{\pi(1+x^2)} dx = \frac{1}{2\pi} \log(1+x^2) \Big|_0^{\infty} = \infty.$$

$$E(X^-) = \int_{-\infty}^0 (-x) \frac{1}{\pi(1+x^2)} dx = -\frac{1}{2\pi} \log(1+x^2) \Big|_{-\infty}^0 = \infty$$

\therefore Não existe $E(X)!!$, mas $E(E(X|Y))=0!$

Exemplo 3.

$$f_{X,Y}(x,y) = y e^{-(x+1)y} \mathbb{I}_{D_Y}(x) \mathbb{I}_{R_Y}(y)$$

$$X|Y=y \sim N(\mu_y, \sigma^2_y)$$

$$E(X|Y=y) = \frac{1}{y} \quad E(X) = \infty$$

$$E(E(X|Y)) = E(1/y) = \infty$$

Outro caso "simples".

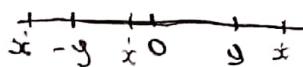
X v.a. absolutamente contínua com densidade f_X . Seja $Y=|X|$.

(X, Y)

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y) = P(X \leq x, |X| \leq y) =$$

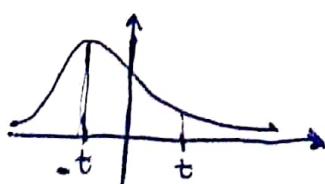
$$= P(X \leq x, -y \leq X \leq y) \Rightarrow$$

$$\Leftrightarrow F_{X,Y}(x,y) = \begin{cases} 0, & y < 0 \\ 0, & y \geq 0 \text{ e } x < -y \\ F_X(x) - F_X(-y), & y \geq 0, -y < x < y \\ F_X(y) - F_X(-y), & y \geq 0, x \geq y \end{cases}$$



$$X / |X|=t$$

Candidato à distribuição condicional de $X|Y=t$, $t > 0$.



$$P(X=t|Y=t) = \frac{f_X(t)}{f_X(-t)+f_X(t)}$$

$$P(X=-t|Y=t) = \frac{f_X(t)}{f_X(-t)+f_X(t)}$$

Nesse caso,

$$F_{x|y=t}(x) = \begin{cases} 0, & x < -t \\ \frac{f_x(-t)}{f_x(t) + f_x(-t)}, & -t \leq x \leq t \\ 1, & x > t \end{cases}$$

Temos que verificar que

$$F_{x,y}(x,y) = \int_{(-\infty, y]} F_{x|y=t}(x) dP_y(t) = \int_{-\infty}^y F_{x|y=t}(x) f_y(t) dt =$$

$$\left(\text{Lembramos que } f_y(t) = (f_x(t) + f_x(-t)) \mathbb{I}_{R_+}(t) \right)$$

$\sum_{|X|}$

$$\cdot y < 0$$

$$\int_{-\infty}^y F_{x|y=t}(x) f_y(t) dt = \int_{-\infty}^y F_{x|y=t}(x) \cdot 0 dt = 0 \quad \checkmark$$

$$\cdot y > 0 \text{ e } x < -y$$

$$\text{Se } y > 0$$

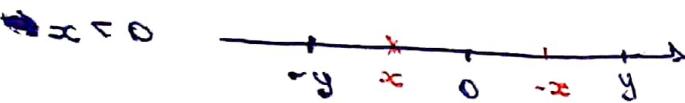
$$\int_{-\infty}^y F_{x|y=t}(x) f_y(t) dt = \int_0^y F_{x|y=t}(x) (f_x(-t) + f_x(t)) dt$$

$$\cdot y > 0 \text{ e } x < -y$$

$$0 < t < y \Rightarrow -y < -t < 0 \Rightarrow x < -y < -t < 0 \Rightarrow F_{x|y=t}(x) = 0.$$

$$\therefore \int_0^y F_{x|y=t}(x) (f_x(-t) + f_x(t)) dt = 0 \quad \checkmark$$

• $y > 0, -y \leq x \leq y$



$$t < |x| \Rightarrow F_{X|Y=t}(x) = 0$$

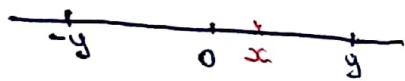
$$t > |x| \Rightarrow F_{X|Y=t}(x) = \frac{f_x(-t)}{f_x(-t) + f_x(t)}$$

$$\int_{-\infty}^y F_{X|Y=t}(x) f_Y(t) dt = \int_0^y F_{X|Y=t}(x) \cdot (f_x(-t) + f_x(t)) dt =$$

$$= \int_0^{|x|} 0 \cdot (f_x(t) + f_x(-t)) dt + \int_{|x|}^y \frac{f_x(-t)}{f_x(t) + f_x(t)} (f_x(-t) + f_x(t)) dt =$$

$$= F_x(-|x|) - F_x(-y) = F_x(x) - F_x(-y)$$

$\rightarrow x > 0$



$$x > t \Rightarrow F_{X|Y=t}(x) = 1$$

$$x \leq t \Rightarrow F_{X|Y=t}(x) = \frac{f_x(-t)}{f_x(-t) + f_x(t)}$$

$$\int_0^y F_{X|Y=t}(x) (f_x(t) + f_x(-t)) dt =$$

$$= \int_0^{\infty} \underbrace{1 \cdot (f_x(t) + f_x(-t))}_{f_Y(t)} dt + \int_{\infty}^y \underbrace{\frac{f_x(-t)}{f_x(-t) + f_x(t)}}_{F_x(-\infty) - F_x(-y)} (f_x(-t) + f_x(t)) dt =$$

$$= F_X(x) - F_X(-\infty) + F_X(-x) - F_X(-y) = F_X(x) - F_X(-y) \quad \checkmark$$

• $x > y \geq 0$

$0 < t < y < \infty$

$$F_{X,Y=t}(x) = 1$$

Logo

$$\int_0^y F_{X,Y=t}(x) (f_X(t) + f_X(-t)) dt =$$

$$= \int_0^y 1 \underbrace{(f_X(t) - f_X(-t))}_{f_Y(y)} dt \quad P(Y \leq y) =$$

$$P(|X| \leq y) = P(-y \leq X \leq y) =$$

$$= F_X(y) - F_X(-y)$$

$$= F_X(y) - F_X(-y) \quad \checkmark$$

Esperança condicional dado uma σ -álgebra.

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

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

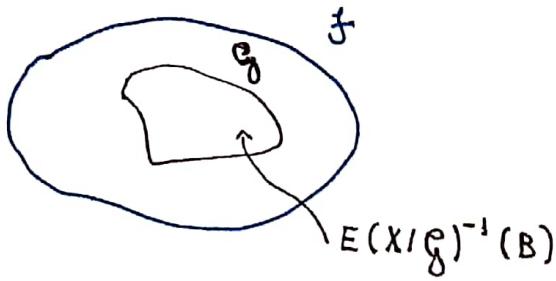
$\mathcal{G} \subseteq \mathcal{F}$ uma σ -álgebra de Ω .

Definição: A ~~função~~ 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) $\int_A X dP = \int_A E(X|\mathcal{G}) dP, \forall A \in \mathcal{G}$

$E(X|\mathcal{G}): \Omega \rightarrow \mathbb{R}$



Exemplo 1.

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

$$E(X|\mathcal{F}_0)(\omega) = \text{cte}, \quad \forall \omega \in \Omega$$

$$X(\omega) = K$$

$$K^{-1}((-\infty, t)) \in \{\emptyset, \Omega\}$$

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,

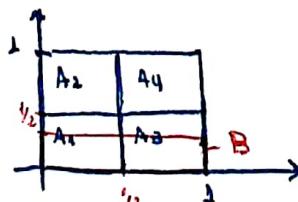
$$E(X|g) = E(X) \text{ com probabilidade } 1.$$

Exemplo 2.

$$\Omega = [0,1]^2$$

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

$$X(\omega) = \sum_{i=1}^4 i \mathbb{I}_{A_i}(\omega)$$



$$B = [0,1] \times [0,1/3]$$

$$\mathcal{G} = \sigma(\{B\}) = \{\emptyset, \Omega, B, B^c\}$$

$$E(X|g)$$

$$E(X|g) = \begin{cases} u, & w \in B \\ v, & w \notin B \end{cases}$$

Para achar $w \in \omega$, usamos

$$\int_B X dP = \int_B E(X|g) dP$$

$$\int_{B^c} X dP = \int_{B^c} E(X|g) dP$$

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}

$X: \Omega \rightarrow \mathbb{R}$ 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). $\int_A X d\mathbb{P} = \int_A E(X|\mathcal{G}) d\mathbb{P}, \forall A \in \mathcal{G}$.

(3) $\mathcal{G} = \mathcal{F}_0 = \{\emptyset, \Omega\}$

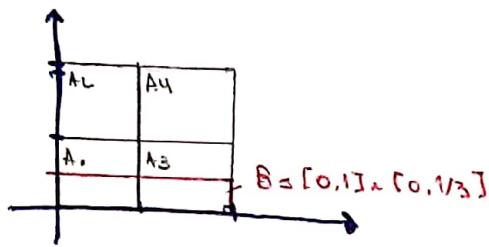
$E(X|\mathcal{G}) = E(X) \quad \mathbb{P}\text{-q.c.}$

$(\mathbb{P}(E(X|\mathcal{G}) = E(X)) = 1)$

Comentário: Se \mathbb{A} é outra função tal que \mathbb{A} atende (1) e (2), então

$\mathbb{A} = E(X|\mathcal{G}) \quad \mathbb{P}\text{-q.c.}$

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



$$X(\omega) = \sum_{i=1}^4 i \mathbb{I}_{A_i}(\omega) \quad \mathcal{G} = \{\emptyset, \Omega, \mathcal{B}, \mathcal{B}^c\}$$

$$E(X|\mathcal{G})(\omega) = \begin{cases} u, & \omega \in \mathcal{B} \\ v, & \omega \in \mathcal{B}^c \end{cases} = \mu \mathbb{I}_{\mathcal{B}}(\omega) + v \mathbb{I}_{\mathcal{B}^c}(\omega).$$

$$\int_B X dP = \int_B E(X|\mathcal{G}) dP \Rightarrow \int_B X \mathbb{I}_{\mathcal{B}} dP = \int_B E(X|\mathcal{G}) \mathbb{I}_{\mathcal{B}} dP \Rightarrow$$

$$\int_B \mathbb{I}_{\mathcal{B}} \sum_{i=1}^4 i \mathbb{I}_{A_i} dP = \int_B \mathbb{I}_{\mathcal{B}} (u \mathbb{I}_{\mathcal{B}} + v \mathbb{I}_{\mathcal{B}^c}) dP \Rightarrow$$

$$\int_B \sum_{i=1}^4 i \mathbb{I}_{A_i \cap \mathcal{B}} dP = \int_B u \mathbb{I}_{\mathcal{B}} dP \Rightarrow \frac{1}{6} + 2 \cdot 0 + 3 \cdot \frac{1}{6} + 4 \cdot 0 = u \cdot \frac{1}{3} \Rightarrow$$

$$\frac{u}{3} = \frac{2}{3} \Rightarrow u = 2$$

e deve satisfazer também

$$\int_{\mathcal{B}^c} X dP = \int_{\mathcal{B}^c} E(X|\mathcal{G}) dP \Rightarrow \int_B \sum_{i=1}^4 i \mathbb{I}_{A_i \cap \mathcal{B}^c} dP = \int_B \mathbb{I}_{\mathcal{B}^c} (\mu \mathbb{I}_{\mathcal{B}} + v \mathbb{I}_{\mathcal{B}^c}) dP \Rightarrow$$

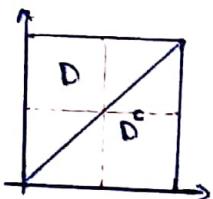
$$\Rightarrow 1 \cdot \frac{1}{12} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{12} + 4 \cdot \frac{1}{4} = v \cdot \frac{1}{3} \Rightarrow \frac{22}{12} = \frac{8}{12} v \Rightarrow v = \frac{11}{4} //$$

Logo,

$$E(X|g)(\omega) = \begin{cases} 2, & \omega \in B \\ \frac{11}{4}, & \omega \in B^c \end{cases}$$

$$E(E(X|G)) = 2 \cdot \frac{1}{3} + \frac{11}{4} \cdot \frac{2}{3} = \frac{30}{12} = \frac{5}{2}$$

Exemplo 3.



$$\mathcal{G}_1 = \{\emptyset, \Omega, D, D^c\}$$

$$E(X|\mathcal{G}_1)(\omega) = \begin{cases} a, & \omega \in D \\ b, & \omega \in D^c \end{cases}$$

$$\int_D X dP = \int_D E(X|\mathcal{G}_1) dP =$$

$$= \int_D X \mathbb{I}_D dP = \int_D \underbrace{E(X|\mathcal{G}_1) \mathbb{I}_D}_{a\mathbb{I}_D + b\mathbb{I}_{D^c}} dP \Rightarrow \int_D \sum_{i=1}^4 i \mathbb{I}_{A_i \cap D} dP = \int_D a \mathbb{I}_D dP \Rightarrow$$

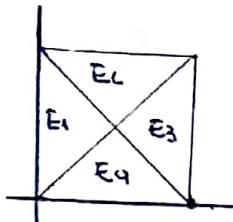
$$\Rightarrow 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{4} + 3 \cdot 0 + 4 \cdot \frac{1}{8} = a \cdot \frac{1}{2} \Rightarrow a = \frac{9}{4}.$$

$$\int_{D^c} X dP = \int E(X|\mathcal{G}_1) dP = \int_D \sum_{i=1}^4 i \mathbb{I}_{A_i \cap D^c} dP = \int_D b \mathbb{I}_{D^c} dP \Rightarrow$$

$$\Rightarrow 1 \cdot \frac{1}{8} + 2 \cdot 0 + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} = b \cdot \frac{1}{2} \Rightarrow b = \frac{11}{4}.$$

$$E(X|g_1)(\omega) = \begin{cases} \frac{9}{4}, & \omega \in D, \\ \frac{11}{4}, & \omega \in D^c. \end{cases}$$

Exercício:



$$\mathcal{G}_2 = \sigma(\{E_1, E_L, E_3, E_4\})$$

$$E(X|\mathcal{G}_2)(\omega) = \sum_{i=1}^4 a_i \mathbb{I}_{E_i}(\omega)$$

$$\int_{E_i} X dP = \int_{E_i} E(X|\mathcal{G}_2) dP \quad \left| \begin{array}{l} \text{onde } a_1 = 3/2 \\ a_2 = 2 \\ a_3 = 7/2 \\ a_4 = 3 \end{array} \right.$$

$i=1, 2, 3, 4$

$$E(X|g_1)$$

$$\frac{E(E(X|g_1)|g_2)}{X}$$

$$\int_A E(X|g_1) dP = \int_A E(E(X|g_1)|g_2) dP, \quad \forall A \in \mathcal{G}_2.$$

Em geral, se A_1, \dots, A_n é uma partição de Ω ($A_i \in \mathcal{F}$), e

$$g = \sigma(\{A_1, A_2, \dots, A_n\})$$

então uma função g -mensurável assume no máximo n valores.

Desse modo

$$E(X|g) \text{ é dada por } E(X|g)(\omega) = \sum_{i=1}^n a_i \mathbb{I}_{A_i},$$

onde a_1, \dots, a_n são tais que

$$\int_{A_i} X dP = \int_{A_i} E(X|g) dP \Rightarrow$$

$$\Rightarrow \int_{\Omega} X dP = \int_{\Omega} I_{A_i} E(X|g) dP \Rightarrow$$

$$\Rightarrow a_i P(A_i) = \int_{A_i} X dP \Rightarrow$$

$$a_i = \frac{\int_{A_i} X dP}{P(A_i)}$$

Algumas Propriedades

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

$$E(X|\mathcal{F}_0) = E(X) \quad P\text{-q.c.}$$

$$\int_A X dP = \int_A E(X|\mathcal{F}_0) dP \quad \begin{array}{l} \cdot \text{Se } A = \emptyset \Rightarrow 0 = 0 \\ \cdot \text{Se } A = \Omega \Rightarrow \int_A X dP = E(X) \end{array}$$

2. $g = f \quad (\Omega, \mathcal{F}, P)$

$$E(X|f) = X, \quad P\text{-q.c.}$$

$$\int_A E(X|f) dP = \int_A X dP, \quad A \in \mathcal{F}$$

3. $X, Y: \Omega \rightarrow \mathbb{R}$

$$a, b \in \mathbb{R}$$

$$E(ax + by | g) = a E(X|g) + b E(Y|g)$$

para $A \in \mathcal{G}_1$,

$$\int_A (\alpha X + b Y) dP = \int_{\Omega} (\alpha X + b Y) \mathbb{I}_A dP = \int_{\Omega} \alpha X \mathbb{I}_A dP + \int_{\Omega} b Y \mathbb{I}_A dP =$$

$$= \alpha \int_{\Omega} X \mathbb{I}_A dP + b \int_{\Omega} Y \mathbb{I}_A dP = \alpha \int_A X dP + b \int_A Y dP =$$

$$= \alpha \int_A E(X|g) dP + b \int_A E(Y|g) dP = \int_A \alpha E(X|g) dP + \int_A b E(Y|g) dP =$$

$$= \int_{\Omega} \alpha E(X|g) \mathbb{I}_A dP + \int_{\Omega} b E(Y|g) \mathbb{I}_A dP = \int_A (\alpha E(X|g) + b E(Y|g)) dP$$

Portanto,

$$E(\alpha X + b Y | g) = \alpha E(X|g) + b E(Y|g).$$

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 $A \in \mathcal{G}_1$,

$$\int_A E(X|\mathcal{G}_2) dP = \int_A X dP = \int_A E(X|\mathcal{G}_1) dP \Rightarrow \forall A \in \mathcal{G}_1,$$

$$\int_A \underbrace{E(X|\mathcal{G}_2)}_Y dP = \int_A E(X|\mathcal{G}_1) dP$$

$$\int_A E(E(X|\mathcal{G}_2)|\mathcal{G}_1) dP$$

Como $\int_A E(X|g_1) dP = \int_A E(X|g_2) dP$, $\forall A \in \mathcal{G}_1$ e $E(X|g_1)$ é \mathcal{G}_1 -mensurável, então

$$E(E(X|g_2)|\mathcal{G}_1) = E(X|\mathcal{G}_1) \quad P\text{-q.c.}$$

Comentário.

$$\text{Dado } P(X=x) = P(\cdot | X=x) \quad \begin{matrix} f_X & f_X(y) = \\ x & y | X=x \end{matrix}$$

$$f_{(x,y)}(x,y) = P(X=x) P(\cdot | X=x). \quad f_{X,Y}(x,y) = f_X(x) f_Y(y).$$

CONVERGÊNCIA QUASE-CERTA

(Ω, \mathcal{F}, P)

$X_1, X_2, X_3, X_4, \dots$

$X_n : \Omega \rightarrow \mathbb{R}$ variável aleatória, $n \geq 1$

$(X_n)_{n \geq 1}$: sequência de v.a.

Exemplo: $\Omega = \mathbb{R}$

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

$$X_1(\omega) = \omega \quad (X_n(1))_{n \geq 1} = (1, 2, 3, 4, \dots)$$

$$X_2(\omega) = 2\omega \quad (X_n(2))_{n \geq 2} = (2, 4, 6, 8, \dots)$$

$(X_n(\omega))$

"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 $(X_n)_{n \in \mathbb{N}}$ converge quase certamente para $X: \Omega \rightarrow \mathbb{R}$ se

$$P(\{\omega \in \Omega : X_n(\omega) \rightarrow X(\omega)\}) = 1$$

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$$

Notação:

$$X_n \xrightarrow{q.c.} X.$$

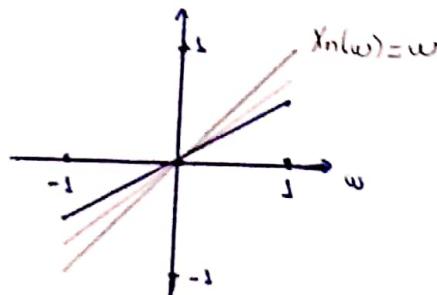
Exemplo:

$$\Omega = [-1, 1] \quad P(A) = \frac{\text{"Comprimento A"}}{2}$$

$$X_n: \Omega \rightarrow \mathbb{R}$$

$$\omega \in \Omega \mapsto X_n(\omega) = \frac{n}{n+1} \omega$$

$$X_n(\omega) = \frac{n}{n+1} \cdot \omega$$



$$\{\omega \in \Omega : X_n(\omega) \text{ converge}\} = [-1, 1]$$

$$\omega \in [-1, 1]$$

$$X_n(\omega) = \frac{n}{n+1} \omega \rightarrow \omega$$

$$[-1, 1]$$

$$P(\{\omega \in \Omega : X_n(\omega) \rightarrow X(\omega) = \omega\})$$

Nesse caso, dizemos que

$$X_n \xrightarrow{q.c.} X,$$

onde $X(\omega) = \omega$.

Exemplo 2.

$$Y_n(\omega) = \omega^n$$

$$\omega = 1 \quad Y_n(1) \rightarrow 1$$

$$\omega = 0 \quad Y_n(0) \rightarrow 0$$

$$1 > \omega > 0 \quad Y_n(\omega) \rightarrow 0$$

$$-1 < \omega < 0 \quad Y_n(\omega) \rightarrow 0$$

$$\omega = -1 \quad Y_n(-1) = \begin{cases} -1 & n \text{ ímpar} \\ 1 & n \text{ par} \end{cases} \quad \cancel{\rightarrow}$$

não converge

$$\{\omega \in \Omega : Y_n(\omega) \text{ converge}\} = \Omega \setminus \{-1\} = (-1, 1]$$



$$\mathbb{P}(\{\omega \in \Omega : Y_n(\omega) \rightarrow Y(\omega)\}) = \mathbb{P}((-1, 1]) = 1, \text{ onde}$$

$$Y(\omega) = \begin{cases} 1, & \omega = 1 \\ 0, & \omega \neq 1 \end{cases}$$

Nesse caso, dizemos que

$$Y_n \xrightarrow{q.c.} Y,$$

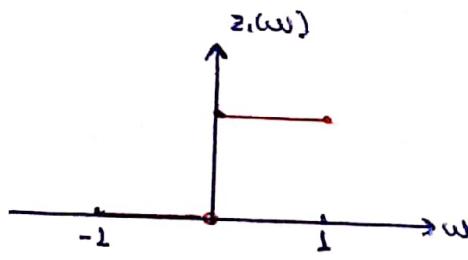
Note que

$$\gamma_n \xrightarrow{q.c} \gamma', \text{ onde}$$

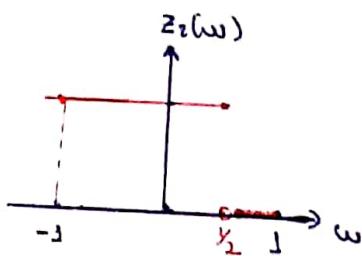
$$\gamma'(\omega) = \begin{cases} 1, & \omega=1 \\ 0, & -1 < \omega < 1 \\ \pi^2 / \arctg 12, & \omega=1 \end{cases}$$

Exemplo 3:

$$z_n(\omega) = \begin{cases} \frac{\pi}{[-1, 1/2]}, & n \text{ par} \\ \frac{\pi}{[0, 1]}, & n \text{ ímpar} \end{cases}$$



n ímpar



$\omega < 0$

$$(z_n(\omega))_{n \geq 1} = (0, 1, 0, 1, 0, \dots) \not\rightarrow ?$$

$0 \leq \omega \leq 1/2$

$$(z_n(\omega))_{n \geq 1} = (1, 1, 1, \dots) \rightarrow 1$$

$\omega \in (1/2, 1]$

$$(z_n(\omega))_{n \geq 1} = (1, 0, 1, 0, \dots) \not\rightarrow ?$$

$\{\omega \in \Omega : (z_n(\omega))_{n \geq 1} \text{ converge}\} = [0, 1/2]$

$$P(\{\omega \in \Omega : (z_n(\omega))_{n \geq 1} \text{ é convergente}\}) = P([0, 1/2]) = \frac{1/2}{2} = \frac{1}{4} < 1$$

d)

$$W_n(\omega) = \sum_{j=0}^n |\omega|^j$$

$$(W_n(1/2)) = (1, 3/2, 7/4, \dots) \rightarrow 2$$

$$\omega \in (-1, 1) \Rightarrow W_n(\omega) = \sum_{j=0}^n |\omega|^j = \sum_{j=0}^{\infty} |\omega|^j = \frac{1}{1 - |\omega|}$$

$$\omega \in \{-1, 1\} \Rightarrow W_n(\omega) = (1, 2, 3, \dots) \rightarrow \infty?$$

$\{\omega \in \Omega : (W_n(\omega))_{n \geq 1} \text{ é convergente}\} =$

$$= (-1, 1) = \Omega \setminus \{-1, 1\}$$

$$W(\omega) = \begin{cases} 0, & \omega \in \{-1, 1\} \\ \frac{1}{1 - |\omega|}, & \omega \in (-1, 1) \end{cases}$$

$$P(\{\omega \in \Omega : W_n(\omega) \rightarrow W(\omega)\}) =$$

$$= P((-1, 1)) = \frac{2}{2} = 1$$

Assim,

$$W_n \xrightarrow{q.s.} W.$$

"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 a 1, temos a convergência quase certa."

$\{w \in \Omega : (z_n(w))_{n \geq 1} \text{ converge}\} = [0, 1/2]$

$$P(\{w \in \Omega : (z_n(w))_{n \geq 1} \text{ é convergente}\}) = P([0, 1/2]) = \frac{1/2}{2} = \frac{1}{4} < 1$$

d)

$$W_n(w) = \sum_{j=0}^n |w|^j$$

$$(W_n(1/2)) = (1, 3/2, 7/4, \dots) \rightarrow 2$$

$$w \in (-1, 1) \Rightarrow W_n(w) = \sum_{j=0}^n |w|^j = \sum_{j=0}^{\infty} |w|^j = \frac{1}{1-|w|}$$

$$w \in \{-1, 1\} \Rightarrow W_n(w) = (1, 2, 3, \dots) \rightarrow \infty?$$

$\{w \in \Omega : (W_n(w))_{n \geq 1} \text{ é convergente}\} =$

$$= (-1, 1) = \Omega \setminus \{-1, 1\}$$

$$W(w) = \begin{cases} 0, w \in \{-1, 1\} \\ \frac{1}{1-|w|}, w \in (-1, 1) \end{cases}$$

$$P(\{w \in \Omega : W_n(w) \rightarrow W(w)\}) =$$

$$= P((-1, 1)) = \frac{2}{2} = 1$$

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 de primeiro pacote é igual a 1, temos a convergência quase certa."

$$V_n(w) = \sum_{s=0}^n (2w)^s$$

$$w \in (-\frac{1}{2}, \frac{1}{2}), V_n(w) \rightarrow \frac{1}{1-2w}$$

$w \notin (-\frac{1}{2}, \frac{1}{2})$, $V_n(w) \not\rightarrow ?$

$P(\{w \in \Omega : (V_n(w))_{n \geq 1} \text{ converge}\}) = P(\{w \in [-\frac{1}{2}, \frac{1}{2}]\}) = \frac{1}{2} \Rightarrow$ não converge quase certamente.

Aula 20 - Luis Gustavo Esteves

26/04/13

$$\hat{P}(\{\omega \in \Omega : w_n(\omega) \rightarrow w(\omega)\}) = \hat{P}((-1, 1)) = \frac{2}{2} = 1$$

Assim

$$w_n \xrightarrow{q.c.} w$$

Para Pensar !

$$V_n(\omega) = \sum_{j=0}^n |2\omega|^j$$

b)

Review

$$X_n \xrightarrow{q.c.} X \Leftrightarrow P(\{\omega \in \Omega : X_n(\omega) \xrightarrow{n \rightarrow \infty} X(\omega)\}) = 1$$

Comentários

(1) No exemplo (2):

$$Y_n \rightarrow Y, \quad Y(\omega) = \begin{cases} 1, & \omega = 1 \\ 0, & \omega \neq 1 \end{cases}$$

$$Y_n \rightarrow Y'', \quad Y''(\omega) = 0, \quad \forall \omega \in \Omega$$

$$X_n(\omega) = \frac{n}{n+1}\omega \xrightarrow{q.c.} \omega = X(\omega) \rightarrow X \sim U(-1, 1)$$

$$Y_n(\omega) \xrightarrow{q.c.} Y \quad P(Y \approx 0) \approx 1$$

$$Y_n \xrightarrow{q.c.} C$$

$$\{w \in \Omega : X_n(w) \rightarrow X(w)\} = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \bigcap_{k=n}^{\infty} \{w : |X_k(w) - X(w)| < \epsilon\}$$

$$X_n(w_0) \rightarrow X(w_0) \Leftrightarrow \forall \epsilon > 0, \exists n \in \mathbb{N} \text{ s.t. } \forall k \geq n \quad |X_k(w_0) - X(w_0)| < \epsilon$$

$$\Rightarrow \forall m \geq 1, w_0 \in \liminf_{n \rightarrow \infty} \{w : |X_n(w) - X(w)| < 1/m\}$$

$$w_0 \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{w : |X_k(w) - X(w)| < 1/m\} \Rightarrow$$

$$\Rightarrow w_0 \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{w : |X_k(w) - X(w)| < 1/m\}$$

Propriedades

(Foto) $(A_n)_{n \geq 1}$ tais que $P(A_n) = 1, \forall n \geq 1 \Rightarrow P(\bigcap_{n \geq 1} A_n) = 1$

$$(1) \quad X_n \xrightarrow{q.c} X \quad \left\{ \begin{array}{l} X_n + Y_n \xrightarrow{q.c} X + Y \\ X_n \xrightarrow{q.c} Y \end{array} \right.$$

$$\text{Def: } A = \{w : X_n(w) \rightarrow X(w)\}$$

$$B = \{w \in \Omega : Y_n(w) \rightarrow Y(w)\}$$

$$w_0 \in A \cap B \Rightarrow \left\{ \begin{array}{l} X_n(w_0) \rightarrow X(w_0) \Rightarrow X_n(w) + Y_n(w_0) \rightarrow \\ Y_n(w_0) \rightarrow Y(w_0) \end{array} \right.$$

$$\Rightarrow X(w_0) + Y(w_0) \Rightarrow w_0 \in \{w \in \Omega : (X_n + Y_n)(w) \rightarrow (X + Y)(w)\} \Rightarrow$$

$$\text{Def: } X_n(w) + Y_n(w)$$

$$\Rightarrow X_n + Y_n \xrightarrow{q.c} X + Y$$

$$(2) \begin{cases} X_n \xrightarrow{qc} X \\ X_n \xrightarrow{qc} Y \end{cases} \Rightarrow P(X=Y)=1$$

$$A = \{\omega : X_n(\omega) \rightarrow X(\omega)\}$$

$$B = \{\omega : X_n(\omega) \rightarrow Y(\omega)\}$$

$$\text{se } w_0 \in A \cap B \Rightarrow \begin{cases} X_n(w_0) \rightarrow X(w_0) \Rightarrow X(w_0) = Y(w_0) \Rightarrow \\ X_n(w_0) \rightarrow Y(w_0) \end{cases} \text{ se } w_0 \in A \cap B \Rightarrow Y(w_0)$$

$$\Rightarrow P(X=Y)=1$$

Ese

$$(3) X_n \xrightarrow{qc} X$$

$$\begin{aligned} g: \mathbb{R} \rightarrow \mathbb{R} & \text{ continua} \\ g & \text{ mensurável} \end{aligned} \Rightarrow g(X_n) \xrightarrow{qc} g(X)$$

Ese

$$(4) X_n \xrightarrow{qc} X \Rightarrow \bar{X}_n = \frac{X_1 + \dots + X_n}{n} \xrightarrow{qc} X$$

Resultado:

$$X_n \xrightarrow{qc} X$$

↑
↓

$$P\left(\limsup \{|X_n - X| > \epsilon\}\right) = 0, \forall \epsilon > 0$$

$|X_n - X| > \epsilon$ i.v.

Dem

$$\underline{X_n \xrightarrow{q.c} X \iff P\left(\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{ |X_k - x| < 1/m \}\right) = 1 \iff}$$

$$\iff P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{ |X_k - x| \geq 1/m \}\right) = 0 \iff$$

$$\iff \forall m \geq 1, P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{ |X_k - x| \geq 1/m \}\right) = 0$$

Resultado

 (Ω, \mathcal{F}, P) $(A_n)_{n \in \mathbb{N}}$ seq. de eventos de \mathcal{F}

$$(1) \sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(\limsup A_n) = 0$$

$$(2) \sum_{n=1}^{\infty} P(A_n) = \infty \text{ e os } A_n \text{'s são independentes}$$



$$P(\limsup_{n \rightarrow \infty} A_n) = 1$$

Dem:

$$(1) P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) \leq P\left(\bigcup_{k=n}^{\infty} A_k\right)$$

$$\leq \sum_{k=n}^{\infty} P(A_k)$$

$$\forall m \geq 1, \exists n = n(m) \text{ t.q. } n > n(\omega) \Rightarrow \sum_{k=n}^{\infty} P(A_k) < 1/m$$

$$\text{Logo, } P(\limsup A_n) < 1/m, \forall m \geq 1$$



$$P(\limsup A_n) = 0$$

(2) $P\left(\bigcap_{k=n}^{\infty} A_k\right)$ Basta verificar $P\left(\bigcup_{k=n}^{\infty} A_k\right) = 1, \forall n \geq 1$

$\forall m \in \mathbb{N}^*$,

$$1-t \leq e^{-t}$$

$$P\left(\bigcup_{k=n}^{\infty} A_k\right) \geq P\left(\bigcup_{k=n}^{n+m} A_k\right) \Rightarrow P\left(\left(\bigcup_{k=n}^{\infty} A_k\right)^c\right) \leq P\left(\bigcap_{k=n}^{n+m} A_k^c\right)$$

$$= \prod_{k=n}^{n+m} P(A_k^c) = \prod_{k=n}^{n+m} (1 - P(A_k)) \leq \prod_{k=n}^{n+m} e^{-P(A_k)} = e^{-\sum_{k=n}^{n+m} P(A_k)}$$

$$\forall m \geq 1, P\left(\left(\bigcup_{k=n}^{\infty} A_k\right)^c\right) \leq \left(e^{-\sum_{k=n}^{n+m} P(A_k)}\right)^{1-a_m} \Rightarrow$$

$$\Rightarrow P\left(\left(\bigcup_{k=n}^{\infty} A_k\right)^c\right) = 0 \Rightarrow P\left(\bigcup_{k=n}^{\infty} A_k\right) = 1$$

Ex:

$(X_n)_{n \geq 1}$ v.a independentes tais que

$$X_n \sim U(0, \frac{1}{n^2})$$

Considerar $Y_n = \mathbb{I}_{(0,1)}(X_n) \xrightarrow{q.c.} 0$?

$$P(|Y_n - 0| > \epsilon) = \begin{cases} 0 & \epsilon > 1 \\ P(Y_n = 1), \epsilon < 1 & \end{cases} \quad \leftarrow P(Y_n = 1) \parallel$$

$$P(X_n \in (0,1)) = \frac{1}{n^2}$$

Logo a série $\sum_{n=1}^{\infty} P(|Y_n - 0| > \epsilon) \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty \Rightarrow$

$$\Rightarrow P\left(\limsup_{n \rightarrow \infty} |Y_n - 0| > \epsilon\right) = 0 \Rightarrow Y_n \xrightarrow{q.c.} 0$$

172

Ex 2

 $(X_n)_{n \geq 1}$ v.a ind. tais que $X_n \sim U(0, n)$

$$Y_n = \prod_{\{x_i=1\}} (x_i) \xrightarrow{\text{a.c.}} ?$$

• $P(Y_n = 0) = P(X_n \in (1, n)) = \frac{n-1}{n} \Rightarrow$

$$\sum_{n=1}^{\infty} P(Y_n = 0) = \sum_{n=1}^{\infty} \frac{n-1}{n} = \infty. \text{ Como os eventos}$$

$\{Y_n = 0\}_{n \geq 1}$, são independentes entçõ

$$P(\limsup_{n \rightarrow \infty} \{Y_n = 0\}) = 1 \quad (\text{B-C}(z))$$

• $P(Y_n = 1) = P(X_n \in (0, 1)) = \frac{1}{n} \Rightarrow \sum_{n=1}^{\infty} P(Y_n = 1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$

Logo B-C(z), $P(\limsup_{n \rightarrow \infty} \{Y_n = 1\}) = 1$.

$\Rightarrow \limsup_{n \rightarrow \infty} \{Y_n = 0\} \cap \limsup_{n \rightarrow \infty} \{Y_n = 1\}$

$$w_0 \in A \cap B \Rightarrow \begin{cases} Y_n(w_0) = 0, \text{ p/ inf. valores de } n \\ Y_n(w_0) = 1, \text{ p/ inf. valores de } n \end{cases} \Rightarrow$$

$\Rightarrow Y_n(w_0)$ não converge $\Rightarrow w_0 \in \{w \in \Omega : (Y_n(w))_{n \geq 1} \text{ não converge}\}$

$\therefore Y_n$ ~~converge~~

11

 $(\Omega, \mathcal{F}, \mathbb{P})$ $X_1, X_2, \dots : \Omega \rightarrow \mathbb{R}$ $X : \Omega \rightarrow \mathbb{R}$

Dizemos que $(X_n)_{n \in \mathbb{N}}$ converge em prob. p/ $X, X_n \xrightarrow{\mathbb{P}} X$

Se $\forall \varepsilon > 0, (\exists m \in \mathbb{N}^*)$

$$\lim_{n \rightarrow \infty} \mathbb{P}(\{\omega \in \Omega : |X_m(\omega) - X(\omega)| > \varepsilon\}) = 0, \forall \varepsilon > 0$$

Exemplos:

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

$\mathbb{P} = \frac{\text{comp.}}{2}$

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

$$(1) X_n(\omega) = \frac{n}{n+1} \omega \quad X(\omega) = \omega$$

$$\mathbb{P}(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) = \mathbb{P}\left(\{\omega \in \Omega : \left|\frac{n\omega}{n+1} - \omega\right| > \varepsilon\}\right)$$

$$= \mathbb{P}\left(\{\omega \in \Omega : \left|\frac{\omega}{n+1}\right| > \varepsilon\}\right) = \mathbb{P}\left(\{\omega \in \Omega : |\omega| > \varepsilon(n+1)\}\right)$$

$$\stackrel{n \gg \varepsilon}{=} \mathbb{P}(\emptyset) = 0. \quad \forall n \gg n \approx \frac{2}{\varepsilon}$$

Lego

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0, \forall \varepsilon > 0$$

$$\therefore X_n \xrightarrow{\mathbb{P}} X$$

$$(2) Y_n(\omega) = \omega^n$$

$$\forall \epsilon > 0 \quad Y(\omega) = 0, \forall \omega \in \Omega \quad (Y_n \xrightarrow{q.c} 0)$$

$$P\left(\{\omega \in \Omega : |Y_n(\omega) - 0| > \epsilon\}\right) = P\left(\{\omega \in \Omega : |\omega^n| > \epsilon^{\frac{1}{n}}\}\right) =$$

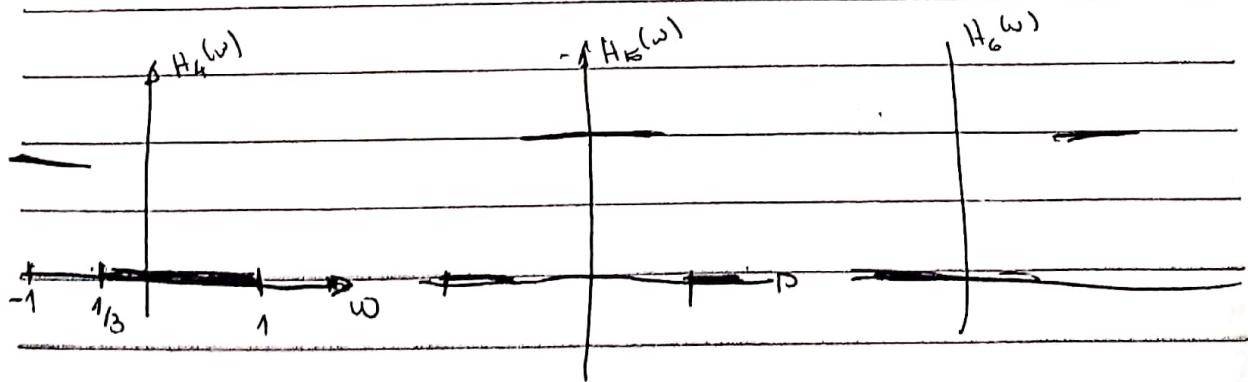
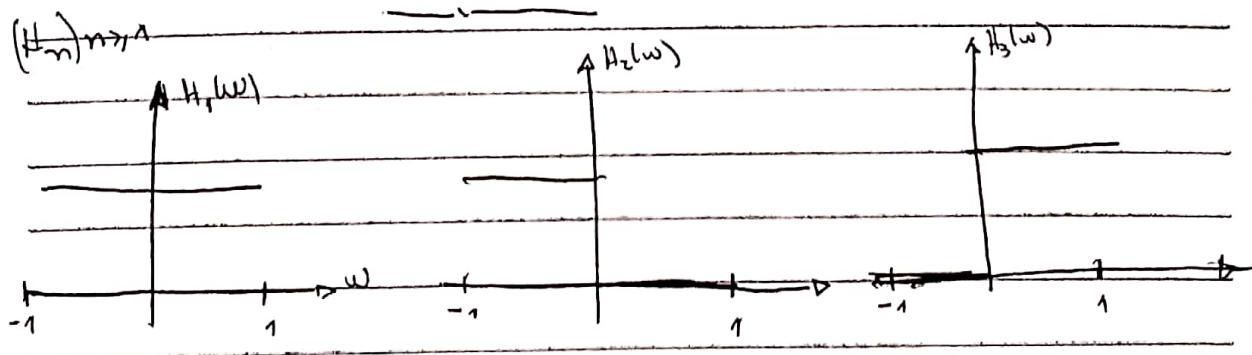
$$= P\left(\{\omega \in \Omega : |\omega| > \epsilon^{1/n}\}\right) = P\left(\{\omega \in \Omega : \omega < -\epsilon^{1/n} \text{ ou } \omega > \epsilon^{1/n}\}\right)$$

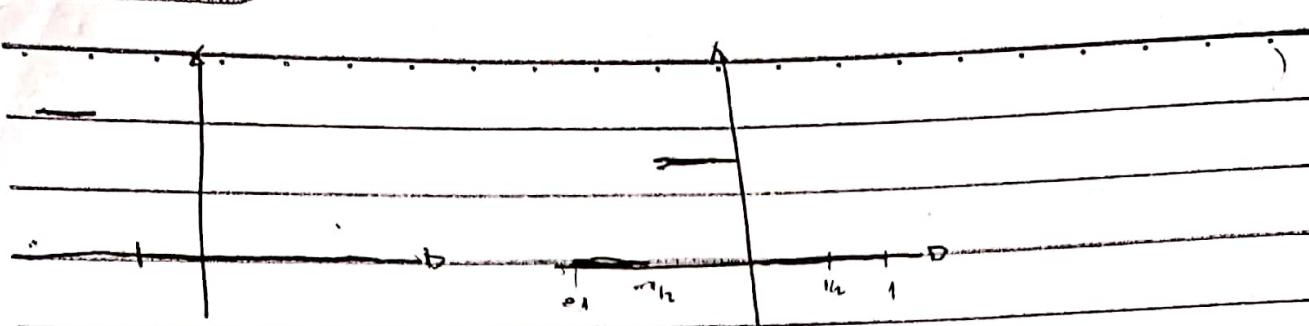
$$= \begin{cases} P(\emptyset) = 0 & ; \epsilon \geq 1 \\ \end{cases}$$

$$\frac{-\epsilon^{1/n} + 1}{2} + \frac{1 - \epsilon^{1/n}}{2} \xrightarrow{n \rightarrow \infty} 0, \quad 0 < \epsilon < 1$$

Portanto, $\forall \epsilon > 0$,

$$\lim_{n \rightarrow \infty} P(|Y_n - 0| > \epsilon) = 0 \quad \text{e, portanto, } Y_n \xrightarrow{P} 0$$





$$\text{TP}(\{\omega \in \Omega : |H_n(\omega) - 0| > \epsilon\}) \xrightarrow{n \leftarrow \infty} (1, 1/2, 1/2, 1/3, 1/3, 1/5, \dots)^{\rightarrow 0}$$

Logo

$$\lim_{n \rightarrow \infty} \text{TP}(\{\omega : H_n(\omega) - 0| > \epsilon\}) = 0, \forall \epsilon > 0$$

Note que a sequência $(A_n(\omega))_{n \in \mathbb{N}}$ "possui"

infinitos zeros e infinitos 1's, para todo $\omega \in \Omega$.

Logo, $(A_n(\omega))_{n \in \mathbb{N}}$ não converge, $\forall \omega \in \Omega$

consequentemente,

$$H_n \not\xrightarrow{a.s.}$$

Probabilidade e Inferência Estatística I

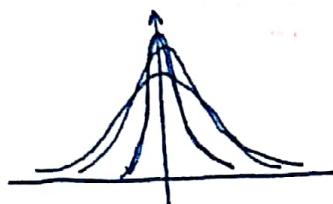
Aula 21 - 30/04/2013

Luis Gustavo Esteves

$$X_n \xrightarrow{P} X \Leftrightarrow \forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| \geq \varepsilon) = 0.$$

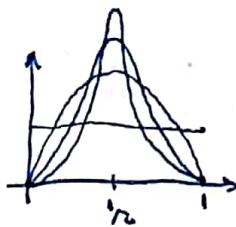
$$(4) X_n \sim N(0, \gamma_n)$$

$$X_n \xrightarrow{D} 0?$$



$$X_n \sim \text{Beta}(n, n)$$

$$X_n \xrightarrow{D} \frac{1}{2}$$



Desigualdade de Markov

Seja X v.a. não negativa. Então, para todo $a \geq 0$,

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

$$X = X \mathbb{I}_{X^{-1}([a, \infty))} + X \mathbb{I}_{X^{-1}([0, a))}$$

$$X \geq X \mathbb{I}_{X^{-1}([a, \infty))} \geq a \mathbb{I}_{X^{-1}([a, \infty))} \Rightarrow E(X) \geq a P(X^{-1}([a, \infty))) \Rightarrow E(X) \geq a P(X \geq a).$$

Desigualdade de Tchebyshev.

Seja X v.a. tal que $E(X) < \infty$.

$$P(|X - E(X)| \geq a) \leq \frac{\text{Var}(X)}{a^2}.$$

Basta tomar $Y = (X - E(X))^2$.

Aplicando a Des. Markov,

$$\begin{aligned} P(|X - E(X)| \geq a) &= P(|X - E(X)|^2 \geq a^2) = P((X - E(X))^2 \geq a^2) \stackrel{\text{D.M.}}{\leq} \\ &= \frac{E(X - E(X))^2}{a^2} = \frac{\text{Var}(X)}{a^2}. \end{aligned}$$

Exemplo 5.

$X_n \sim \text{Beta}(n, n)$

$E \geq 0$

$$P\left(\left|X_n - \frac{1}{2}\right| \geq \frac{\epsilon}{2}\right) \leq \frac{\text{Var}(X_n)}{\epsilon^2} = \frac{1}{\epsilon^2} \frac{n \cdot n}{(n+n)^2 (2n+1)} \xrightarrow{n \rightarrow \infty} 0, \quad \forall \epsilon > 0.$$

Logo, $X_n \xrightarrow{P} \frac{1}{2}$.

Exemplo 6.

$(X_n)_{n \geq 1}$ seq. v.a. tais que $E(X_n) = \mu, \forall n \geq 1$.

Seja $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$

Des. Markov

$$P\left(\left|\bar{X}_n - \mu\right| \geq \epsilon\right) \leq \frac{\text{Var}(\bar{X}_n)}{\epsilon^2} = \frac{1}{n^2 \epsilon^2} \text{Var}(X_1 + \dots + X_n) = \frac{1}{n^2 \epsilon^2} \sum_{i=1}^n \text{Var}(X_i) \leq \frac{C}{n^2 \epsilon^2} \xrightarrow{n \rightarrow \infty} 0.$$

$(X_n)_{n \geq 1}$ são não-correlacionadas

$\text{Var}(X_n) \leq C, \forall n \geq 1$

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)

$$\frac{(X_1 + \dots + X_n) - E(X_1 + \dots + X_n)}{n} \xrightarrow{\text{P}} 0 \quad (\xrightarrow{\text{q.c.}})$$

Exemplo 7. $(X_n)_{n \geq 1}$ v.a. independentes com $X_n \sim U(0,1)$

$$Y_n = \mathbb{I}_{(0,1)}(X_n)$$

Vimos que $\cancel{Y_n \xrightarrow{\text{q.c.}}}$

$$P(|Y_n - 0| > \epsilon) = \begin{cases} 0, & \epsilon \geq 1 \\ P(Y_n = 1) = P(X_n \in (0,1)) = 1/n, & 0 < \epsilon < 1 \end{cases}$$

Assim, para todo $\epsilon > 0$,

$$P(|Y_n - 0| > \epsilon) \leq \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Logo, $Y_n \xrightarrow{\text{P}} 0$.

Resultado:

$$X_n \xrightarrow{\text{q.c.}} X \Rightarrow Y_n \xrightarrow{\text{P}} X.$$

$$X_n \xrightarrow{\text{q.c.}} X \Rightarrow \forall \epsilon > 0, P(\limsup_{n \rightarrow \infty} \{|X_n - X| > \epsilon\}) = 0 \Rightarrow \forall \epsilon > 0, P(\bigcap_{k=1}^{\infty} \{ |X_k - X| \leq \epsilon \}) = 1$$

$$\Rightarrow \forall \epsilon > 0, \lim_{n \rightarrow \infty} P\left(\bigcup_{k=n}^{\infty} \{ |X_k - X| > \epsilon \}\right) = 0. \Rightarrow$$

$$\Rightarrow \forall \varepsilon > 0, 0 = \lim_{n \rightarrow \infty} P\left(\bigcup_{k=n}^{\infty} \{ |X_k - X| > \varepsilon \}\right) \geq \limsup_{n \rightarrow \infty} P(\{|X_n - X| > \varepsilon\}) \geq \liminf_{n \rightarrow \infty} P(\{|X_n - X| > \varepsilon\}) \geq 0 \Rightarrow$$

$\rightarrow \forall \varepsilon > 0,$

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0 \quad \text{e},$$

portanto, $X_n \xrightarrow{P} X.$

Propriedades:

$$(1) \begin{cases} X_n \xrightarrow{P} X \\ Y_n \xrightarrow{P} Y \end{cases} \quad X_n + Y_n \xrightarrow{P} X + Y$$

$$(\alpha X_n \xrightarrow{P} \alpha X) \quad (\text{exercício})$$

$$|(X_n + Y_n) - (X + Y)| = |(X_n - X) + (Y_n - Y)| \leq |X_n - X| + |Y_n - Y|$$

$$\{\omega : |(X_n(\omega) + Y_n(\omega)) - (X(\omega) + Y(\omega))| < \varepsilon\} \supseteq \{\omega : |X_n(\omega) - X(\omega)| < \varepsilon/2\} \cap \{\omega : |Y_n(\omega) - Y(\omega)| < \varepsilon/2\} \Rightarrow$$

$$\{\omega : |(X_n(\omega) + Y_n(\omega)) - (X(\omega) + Y(\omega))| \geq \varepsilon\} \subseteq \{\omega : |X_n(\omega) - X(\omega)| \geq \varepsilon/2\} \cup \{\omega : |Y_n(\omega) - Y(\omega)| \geq \varepsilon/2\} \Rightarrow$$

$$\Rightarrow P(|(X_n + Y_n) - (X + Y)| \geq \varepsilon) \leq P(|X_n - X| \geq \varepsilon/2) + P(|Y_n - Y| \geq \varepsilon/2) \Rightarrow$$

$$\Rightarrow \lim_{n \rightarrow \infty} P(|(X_n + Y_n) - (X + Y)| \geq \varepsilon) = 0, \forall \varepsilon > 0. \text{ Logo, } X_n + Y_n \xrightarrow{P} X + Y.$$

$$(2) \begin{cases} X_n \xrightarrow{P} X \\ Y_n \xrightarrow{P} Y \end{cases} \Rightarrow P(X=Y)=1$$

$$\left| X - Y \right| = |X - X_n + X_n - Y| \leq |X_n - X| + |X_n - Y| \xrightarrow[\sim]{\frac{1}{2m}} \frac{1}{2m}, \forall m \in \mathbb{N}$$

mesma ideia da anterior.

$$(3) X_n \xrightarrow{P} X \Rightarrow g(X_n) \xrightarrow{P} g(X)$$

$g: \mathbb{R} \rightarrow \mathbb{R}$ contínua.

Convergência em Distribuição

Convergência da sequência de funções de distribuição para uma função. As variáveis aleatórias envolvidas não precisam estar definidas em um mesmo espaço de probabilidade".

$$\begin{array}{ccc} (\Omega_1, \mathcal{F}_1, P_1) & \xrightarrow{X_1} & (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_{X_1}) \\ (\Omega_2, \mathcal{F}_2, P_2) & \xrightarrow{X_2} & (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_{X_2}) \end{array}$$

Sejam X_1, X_2, \dots v.a. com funções de distribuição

F_{X_1}, F_{X_2}, \dots respectivamente.

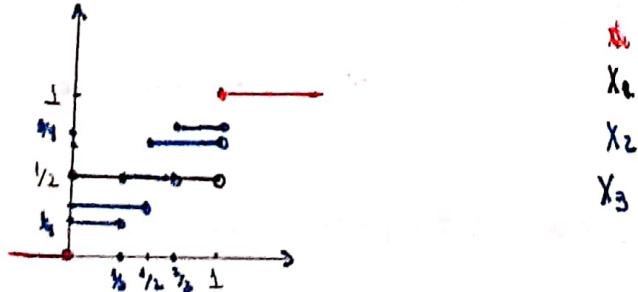
Seja X uma v.a. com função de distribuição F_X .

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

$$F_{X_n}(x) \xrightarrow[n \rightarrow \infty]{} F_X(x), \forall x \in C_X, \text{ onde } C_X \text{ é o conjunto dos pontos de continuidade de } F_X.$$

Exemplo 1.

$$X_n \sim U(\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}) \quad P(X_n = s/n) = \frac{1}{n+1}, s \in \{0, 1, \dots, n\}$$



$$F_{X_n}(x) = \begin{cases} 0, & x < 0 \\ \frac{\sum_{j=0}^n I_{(-\infty, x]}(j/n)}{n+1} = \frac{[nx] + 1}{n+1}, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = \begin{cases} 0, & x < 0 \\ \infty, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

Para $x \in [0, 1]$,

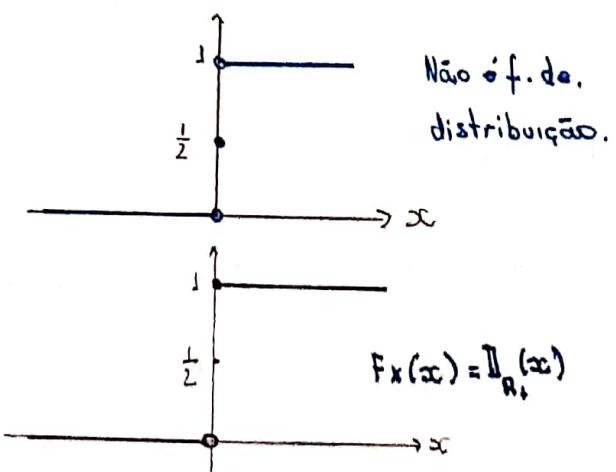
$$\frac{nx}{n+1} \leq \frac{[nx] + 1}{n+1} \leq \frac{nx + 1}{n+1} \quad \downarrow \quad \downarrow$$

$$\frac{x}{x}$$

Exemplo 2: $X_n \sim N(0, 1/n)$.

$$F_{X_n}(x) = P(X_n \leq x) = P\left(\frac{X_n}{\sqrt{n}} \leq \frac{x}{\sqrt{n}}\right) =$$

$$= P(Z \leq x\sqrt{n}) \xrightarrow{n \rightarrow \infty} \begin{cases} 0, & x < 0 \\ 1/2, & x = 0 \\ 1, & x > 0 \end{cases}$$



Note que $F_{X_n}(x) \rightarrow F_X(x)$, $\forall x \neq 0$ (\neq ponto de continuidade de F_X).

Assim, $X_n \xrightarrow{d} X$,
 $X_n \xrightarrow{d} 0$.

Exemplo 3. $X_n \sim U(n, n+1)$

$$F_{X_n}(x) = \begin{cases} 0, & x < n \\ \frac{x-n}{(n+1)-n} = \frac{x-n}{1}, & n \leq x \leq n+1 \\ 1, & x > n+1 \end{cases}$$

"tight"

Se $x < 0$, então

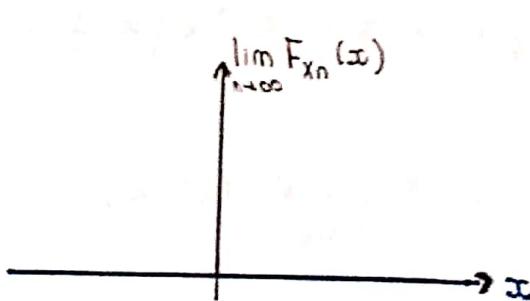
$$\lim_{n \rightarrow \infty} F_{X_n}(x) = 0$$

Para $x > 0$, existe $n_0 = n_0(x)$ tal que

$$x < n_0 ([x] + 1)$$

Então, para $n \geq n_0$, $F_{X_n}(x) = 0$.

Logo, $\lim_{n \rightarrow \infty} F_{X_n}(x) = 0, \forall x > 0$.



Note que $(F_{X_n}(x))_{n \geq 1}$ converge para todo x , mas a função limite NÃO é função de distribuição.

Nesse caso,

~~$X_n \xrightarrow{d} 0$~~

Resultado:

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

Seja $x \in \mathbb{R}$ ponto de continuidade de F_x . $\forall m \in \mathbb{N}^*$

$$\{X_n \leq x\} = \{X_n \leq x, |X_n - x| \leq \frac{1}{m}\} \cup \{X_n \leq x, |X_n - x| > \frac{1}{m}\} \Rightarrow \begin{cases} X_n \leq x \\ X_n - \varepsilon \leq x \leq X_n + \varepsilon \end{cases} \Rightarrow x \leq x + \varepsilon$$

$$\{X_n \leq x\} \subseteq \{X \leq x + \frac{1}{m}\} \cup \{|X_n - x| > \frac{1}{m}\} \Rightarrow$$

$$P(X_n \leq x) \leq P(X \leq x + \frac{1}{m}) + P(|X_n - x| > \frac{1}{m}) \quad (\text{I})$$

Analogamente

$$\{X \leq x - \frac{1}{m}\} = \{X \leq x - \frac{1}{m}, |X_n - x| \leq \frac{1}{m}\} \cup \{X \leq x - \frac{1}{m}, |X_n - x| > \frac{1}{m}\}$$

$$\{X \leq x - \frac{1}{m}\} \subseteq \{X_n \leq x\} \cup \{|X_n - x| > \frac{1}{m}\} \Rightarrow$$

$$P(X \leq x - \frac{1}{m}) \leq P(X_n \leq x) + P(|X_n - x| > \frac{1}{m}) \quad (\text{II})$$

Aplicando limite em n (em I e II)

$$\left. \begin{array}{l} \limsup P(X_n \leq x) \leq F_x(x + \frac{1}{m}) \\ F_x(x - \frac{1}{m}) \leq \liminf P(X_n \leq x) \end{array} \right\} \xrightarrow[m \rightarrow \infty]{n \rightarrow \infty} \underbrace{\lim P(X_n \leq x)}_{F_{X_n}(x)} = F_x(x)$$

Logo, $X_n \xrightarrow{d} X$. (pois $F_{X_n}(x) \xrightarrow{n \rightarrow \infty} F_x(x)$ pt todo ponto de continuidade de F_x)

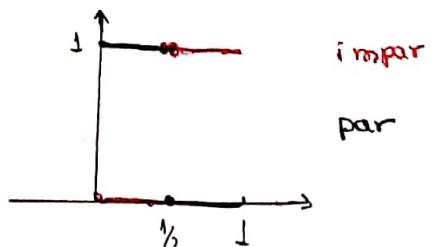
Exemplo:

$$\Omega = [0,1]$$

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

P: "comprimento"

$$X_n(\omega) = \begin{cases} \mathbb{I}_{[0, \frac{1}{2}]}(\omega), & n \text{ par} \\ \mathbb{I}_{[\frac{1}{2}, 1]}(\omega), & n \text{ ímpar} \end{cases}$$



$$X(\omega) = \begin{cases} \mathbb{I}_{(0,1/4)}(\omega) \\ \mathbb{I}_{(1/4,3/4]}(\omega) \end{cases}$$

$X_n \sim \text{Ber}(1/2)$, $\forall n \in \mathbb{N}$

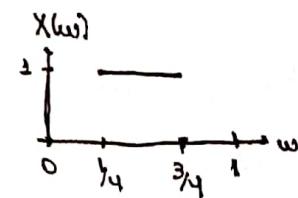
$$\text{Logo, } F_{X_n}(x) = \sum_{\omega \in \Omega} \mathbb{I}_{(0,1)}(\omega) + \mathbb{I}_{[1,\infty)}(\omega)$$

$X \sim \text{Ber}(1/2)$

Logo $X_n \xrightarrow{d} X$.

$$0 < \varepsilon < 1$$

$$\begin{aligned} P(|X_n - X| > \varepsilon) &= P(X_n = 1, X = 0) + P(X_n = 0, X = 1) = \\ &= \begin{cases} P((0, 1/4)) + P((1/2, 3/4)) & , n \text{ par} \\ " & " \\ P((3/4, 1)) + P((1/4, 1/2)) & , n \text{ impar} \\ " & " \end{cases} \end{aligned}$$



$$= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \quad \forall n \in \mathbb{N}.$$

$$P(|X_n - X| > \varepsilon) = 1/2, \quad \forall n \in \mathbb{N}.$$

$$\text{Logo, } \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = \frac{1}{2} \neq 0.$$

Portanto, $X_n \cancel{\xrightarrow{P}} X$.

Resultado.

Sejam X, X_1, X_2, \dots v.a. discretas assumindo valores em \mathbb{N}

Sejam $(p_1(k))_{k \in \mathbb{N}}, (p_2(k))_{k \in \mathbb{N}}, \dots$ as funções das probabilidades de X_1, X_2, \dots , respectivamente.

Seja $(p(k))_{k \in \mathbb{N}}$

|

$P(X=k)$

Resultado.

$$X_n \xrightarrow{d} X \Leftrightarrow p_n(k) \xrightarrow{n \rightarrow \infty} p(k), \forall k \in \mathbb{N}$$

Verificação:

(\Rightarrow)

$$\text{Note que } p_n(k) = P(X_n=k) = P(k - 1/2 \leq X_n \leq k + 1/2) =$$

pois $k - \frac{1}{2} \leq k + \frac{1}{2}$ são pontos de continuidade de F_X

$$= F_{X_n}(k + 1/2) - F_{X_n}(k - 1/2) \xrightarrow{n \rightarrow \infty} F_X(k + 1/2) - F_X(k - 1/2) = P(X=k) = p(k), \forall k \in \mathbb{N}$$

(\Leftarrow)

$$F_{X_n}(x) = \sum_{k=0}^{\infty} \underbrace{P(X_n=k)}_{p_n(k)} \xrightarrow{n \rightarrow \infty} \sum_{k=0}^{\infty} \underbrace{P(X=k)}_{p(k)} = F_X(x).$$

Exemplo.

$$X_n \sim \text{Bin}(n, p_n), \quad n p_n \xrightarrow{n \rightarrow \infty} \lambda, \quad \lambda > 0.$$

$$K \in \mathbb{N} \quad (n \geq K)$$

$$P(X_n=k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \\ D_n(k)$$

$$= \frac{n(n-1)\dots(n-k+1)(n-k)!}{k! (n-k)!} p_n^k (1-p_n)^{n-k} =$$

$$= \frac{1}{k!} \left\{ \prod_{j=0}^{k-1} (n-j) p_n \right\} \left(1 - \frac{np_n}{n} \right)^{n-k}$$

$$= \frac{1}{k!} \left\{ \prod_{j=0}^{k-1} (n-j) p_n \right\} \left(1 - \frac{np_n}{n} \right)^n \left(1 - \frac{np_n}{n} \right)^{-k}$$

$$p_n \rightarrow 0$$

$$\alpha_n = np_n \rightarrow \lambda$$

$$\xrightarrow[n \rightarrow \infty]{} \frac{1}{k!} \lambda^k e^{-\lambda} = p(k), \forall k \in \mathbb{N}.$$

Logo $X_n \xrightarrow{d} X$, onde $X \sim \text{Poisson}(\lambda)$.

Exercício.

$$X_n \sim \text{HG}(np, n(1-p))_k, \quad p \in (0,1)$$

$$X_n \xrightarrow{d} X, \quad \text{onde } X \sim \text{Binomial}(k, p)$$

$$k \in \mathbb{N},$$

$$p_n(k) = P(X_n=k) = \binom{np}{k} \left(\begin{matrix} n \\ k \end{matrix} \right) (1-p)^{n-k}$$

111

$$\text{O2} \setminus \text{O5} / X_n \xrightarrow{\varphi} X \Leftrightarrow F_{X_n}(t) \rightarrow F_X(t), \forall t \in C_X$$

PI Verificare: $X_n \xrightarrow{d} c \Rightarrow X_n \xrightarrow{P} c$

No exemplo $x_n \sim HG(k_{1,n}, k_{2,n}, m)$

$$K_1, K_2 \in \mathbb{N}^*$$

$$X_n \xrightarrow{d} X \sim \text{Bin}\left(m, \frac{k_1}{k_1 + k_2}\right)$$

Propriedades (Teo de Slutsky)

$$(1) \quad \begin{array}{c} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{d} C \end{array} \left\{ \begin{array}{l} \Rightarrow X_n + Y_n \xrightarrow{d} X + C \\ \mathbb{P}(X_n + Y_n \leq t) \rightarrow \mathbb{P}(X + C \leq t) \end{array} \right.$$

$$(z) \quad \left. \begin{array}{l} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{d} c \end{array} \right\} \Rightarrow X_n - Y_n \xrightarrow{d} cX$$

Idea

$$\overline{P}(X_n + Y_n \leq t) = \overline{P}(X_n + Y_n \leq t, Y_n < c - \frac{1}{m}) +$$

↑

$$+ \overline{P}(X_n + Y_n \leq t, Y_n \geq c - \frac{1}{m})$$

$$\overline{P}(Y_n < c - \frac{1}{m}) \quad \quad \quad \overline{P}(X_n \leq -)$$

$$\overline{P}(X_n \leq t - c - Y_m) = \overline{P}\left(X_n \leq t - c - \frac{1}{m}, X_n + Y_n \leq t\right) + \overline{P}\left(X_n \leq t - c - \frac{1}{m}, X_n + Y_n > t\right)$$

Ex: $\Omega = [0, 1]$

$$\mathcal{G} = \mathcal{B}([0,1])$$

TP = "comprimento"^{v1}

$$X_n(\omega) = \begin{cases} \mathbb{I}_{[0,1/2]}(\omega), & n \text{ par} \\ \mathbb{I}_{[1/2,1]}(\omega), & n \text{ ímpar} \end{cases}$$

$$Y_n(\omega) = \begin{cases} \mathbb{I}_{[0,1/2]}(\omega), & n \text{ par} \\ \mathbb{I}_{[1/2,1]}(\omega), & n \text{ ímpar} \end{cases}$$

$$X(\omega) = \mathbb{I}_{[1/4,3/4]}(\omega)$$

$$X_n \sim \text{Ber}(1/2), \forall n \geq 1$$

$$Y_n \sim \text{Ber}(1/2), \forall n \geq 1$$

$$X_n \xrightarrow{d} X \quad Y_n \xrightarrow{d} X \quad | \quad X_n + Y_n \xrightarrow{d} 2X$$

No entanto,

$$X_n(\omega) + Y_n(\omega) = 1, \forall \omega \in \Omega$$

$$\text{Logo, } \mathbb{P}(X_n + Y_n = 1) = 1, \forall n \geq 1$$

$$X_n + Y_n \xrightarrow{d} 1$$

$$Z = X + X' \text{ tal que } \mathbb{P}(Z=2) = \mathbb{P}(Z=0) = 1/2$$

Ideia \rightarrow Prox

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sqrt{n\sigma^2}} \rightarrow N(0,1)$$

$$\mathbb{P}(Z_n \leq t) \rightarrow \mathbb{P}(Z \leq t)$$

Resultado

$(x_n)_{n \geq 1}$ seq. de v.a com f.s geradora φ_{x_n}

x v.a com f.s geradora φ_x

$$\varphi_{x_n}(t) \rightarrow \varphi_x(t), \forall t \in I \Rightarrow x_n \xrightarrow{d} x$$

De um modo mais geral, se $\varphi_{x_n}(t) \rightarrow \varphi$ e φ é contínua no zero então φ é f.s geradora de algumas v.a x e $x_n \xrightarrow{d} x$

Exemplos

$$P(x_0 = 0) = 1$$

$$(1) x_n \sim N(0, 1/n)$$

$$\varphi_{x_n}(t) = e^{t \cdot 0} \cdot e^{\frac{(1-t^2)/2}{1/n}} = e^{t^2/2n} \rightarrow 1, \forall t \in \mathbb{R}$$

$$\varphi_{x_0}(t) = e^{t \cdot 0} P(x_0 = 0) = 1, \forall t \in \mathbb{R}$$

Pelo Resultado $x_n \xrightarrow{d} x_0 \equiv 0$

Ex 2.

$$x_n \sim \text{Bin}(n, p_n)$$

$$np_n \xrightarrow{n \rightarrow \infty} \lambda, \lambda > 0$$

$$\begin{aligned} \varphi_{x_n}(t) &= (p_n e^t + 1 - p_n)^n \\ z &\sim \text{Bin}(n, p) \end{aligned}$$

$$\varphi_{x_n}(t) = (p_n e^t + 1 - p_n)^n = (1 + p_n(e^t - 1))^n =$$

$$= \left(1 + \frac{np_n(e^t - 1)}{n}\right)^n \xrightarrow{n \rightarrow \infty} e^{\lambda(e^t - 1)} = \varphi_x(t), x \sim P_D(\lambda)$$

Logo, $x_n \xrightarrow{d} x$

11

(3) $X_n \sim N(0, n)$

$$\varphi_{X_n}(t) = e^{t \cdot 0} \cdot e^{-nt^2/2} = e^{-nt^2/2} \rightarrow \begin{cases} 1, & t = 0 \\ \infty, & t \neq 0 \end{cases}$$

(4) $X_n \sim \text{Exp}(1/n)$?

Fazer

Teo do Limite Central

$$(X_n)_{n \geq 1} \text{ v.a. i.i.d com } E(X_1) = \mu \text{ e } \text{VAR}(X_1) = \sigma^2 < \infty$$

Então,

$$\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} Z \sim N(0, 1)$$

Verificação:

$$\varphi_{\frac{x_1 + \dots + x_n - n\mu}{\sqrt{n\sigma^2}}}(t) = \varphi_{\frac{\sum_{i=1}^n x_i - \mu}{\sqrt{n\sigma^2}}}(t) = \varphi_{\frac{\sum_{i=1}^n x_i - \mu}{\sqrt{n\sigma^2}}} \left(\frac{t}{\sqrt{n\sigma^2}} \right) =$$

$$= \prod_{i=1}^n \varphi_{x_i - \mu} \left(\frac{t}{\sqrt{n\sigma^2}} \right) \stackrel{i.i.d.}{=} \left(\varphi_{x_1 - \mu} \left(\frac{t}{\sqrt{n\sigma^2}} \right) \right)^n$$

$$\varphi_{x_1 - \mu} \left(\frac{t}{\sqrt{n\sigma^2}} \right) = \varphi_{x_1 - \mu}(0) + \frac{\varphi'(0)}{\sigma\sqrt{n}} \left(\frac{t}{\sqrt{n\sigma^2}} - 0 \right) +$$

$$+ \frac{1}{2!} \varphi''(\bar{E}_n) \left(\frac{t}{\sqrt{n\sigma^2}} - 0 \right)^2, \text{ onde } 0 < \bar{E}_n < \frac{t}{\sigma\sqrt{n}}$$

$$\text{Assim, } \varphi_{x_1 - \mu} \left(\frac{t}{\sqrt{n\sigma^2}} \right) = 1 + \frac{\varphi''(\bar{E}_n)t^2}{\sigma^2 n}$$

$$\text{Logo, } \varphi''(0) = E((x_1 - \mu)^2)$$

$$\varphi_{\frac{x_1 + \dots + x_n - n\mu}{\sqrt{n\sigma^2}}}(t) = \left(1 + \frac{\frac{\varphi''(\bar{E}_n)t^2}{\sigma^2 n}}{n} \right)^n \xrightarrow[n \rightarrow \infty]{\approx} e^{\frac{t^2}{2\sigma^2}} = \varphi_e(t)$$

$$\text{Logo, } \frac{x_1 + \dots + x_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0, 1)$$

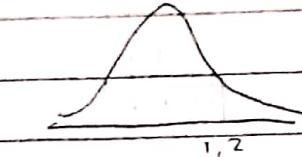
$$\text{P/ } n \text{ grande } \sum x_i \stackrel{a}{\sim} N(n\mu, n\sigma^2)$$

Exemplos

X_1, \dots, X_{100} iid Beta(4,4)

$$P(X_1 + \dots + X_{100} \leq 52) = P\left(\frac{X_1 + \dots + X_{100} - 100}{\sqrt{\frac{100 \cdot 4 \cdot 4}{889}}} \leq \frac{52 - 100 \cdot \frac{1}{2}}{\sqrt{\frac{100 \cdot 4 \cdot 4}{889}}}\right) =$$

$$= P\left(Z \leq \frac{52 - 50}{\sqrt{\frac{25}{9}}}\right) = P(Z \leq 1,2)$$



Ex 2: 200 lançamentos de um dado

X_i : valor observado no i -ésimo lançamento,
 $i = 1, \dots, 200$

$$P(X_1 + \dots + X_{200} > 670) = 1 - P(X_1 + \dots + X_{200} \leq 670) =$$

$$= 1 - P\left(\frac{(X_1 + \dots + X_{200}) - 200 \cdot \frac{7}{2}}{\sqrt{\frac{200 \cdot 35}{12}}} \leq \frac{670 - 200 \cdot \frac{7}{2}}{\sqrt{\frac{200 \cdot 35}{12}}}\right) \approx 1 - P\left(Z \leq \frac{-30}{\sqrt{\frac{35}{6}}}\right) \approx 1 - P\left(Z \leq \frac{-3\sqrt{10}}{\sqrt{35}}\right)$$

$$X_i \sim U\{1, 2, 3, 4, 5, 6\}$$

$$E(X_i) = \frac{7}{2} \quad \text{Var}(X_i) = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} - \left(\frac{7}{2}\right)^2 =$$

$$= \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

11

$$(3) \lim_{n \rightarrow \infty} \int_0^n \frac{t^{n-1}}{(n-1)!} e^{-t} dt ?$$

$\underbrace{\qquad\qquad\qquad}_{\text{Gamma}(n, 1)}$

$= \lim_{n \rightarrow \infty} \mathbb{P}(X_1 + \dots + X_n \leq n)$, onde $(X_n)_{n \geq 1}$ são

v.a.i.i.d $\text{Exp}(1)$

$$= \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{X_1 + \dots + X_n - n}{\sqrt{n}} \leq \frac{n-n}{\sqrt{n}}\right) = \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{X_1 + \dots + X_n - n}{\sqrt{n}} \leq 0\right)$$

$$\mathbb{P}(Z \leq 0) = \frac{1}{2}$$

Lei Forte dos Grandes Números (LFGN)

$(X_n)_{n \geq 1}$ v.a.i.i.d tais que $E(X_i) = \mu < \infty$

$$\text{Então } \bar{X}_n = \frac{X_1 + \dots + X_n}{n} \xrightarrow{q.c.} \mu$$

Comentário

0	1
0	0

$$\begin{aligned} \frac{X_1 + \dots + X_n}{n} &\rightarrow \text{Beta}(2, 1) \\ \frac{X_1 + \dots + X_n}{n} &\xrightarrow{2} \frac{2}{3} \end{aligned}$$

11

Exemplo

(1) $(X_n)_{n \geq 1}$ v.a.i.i.d Poisson(λ), $\lambda > 0$

Pela LFGN

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow{q.c} E(X_1) = \lambda$$

$$P(X_1^2 \leq t) = P(X_1^2 \leq t)$$

$$\frac{X_1^2 + \dots + X_n^2}{n} \xrightarrow{q.c} E(X_1^2) = \lambda + \lambda^2$$

$$E(X_1^2) = \text{Var}(X_1) + E^2(X_1) = \lambda + \lambda^2 < \infty$$

(2) $(X_n)_{n \geq 1}$ v.a.i.i.d Gama(a, b), $a, b > 0$

$$\frac{X_1 + \dots + X_n}{\sqrt{n(X_1^2 + \dots + X_n^2)}} = \frac{\frac{X_1 + \dots + X_n}{n}}{\sqrt{\frac{X_1^2 + \dots + X_n^2}{n}}} \xrightarrow{q.c} \frac{a/b}{\sqrt{\frac{a^2 + b^2}{b^2}}} = \sqrt{\frac{a^2 + b^2}{b^2}}$$

(3) (X_n) v.a ind. tais que $X_{2n-1} \sim \text{Beta}(2, 1)$, $n \geq 1$

$$X_{2n} \sim U(0, 1), n \geq 1$$

$$\bar{X}_n = \frac{x_1 + \dots + x_n}{n}$$

~~termo~~

$$\bar{X}_{2n} = \frac{x_1 + x_2 + \dots + x_{2n}}{2n} =$$

$$= \frac{x_1 + x_3 + \dots + x_{2n-1}}{2n} + \frac{x_2 + x_4 + \dots + x_{2n}}{2n} =$$

$$= \frac{1}{2} \underbrace{\frac{x_1 + x_3 + \dots + x_{2n-1}}{n}}_{\frac{1}{2} E(x_1)} + \frac{1}{2} \underbrace{\frac{x_2 + x_4 + \dots + x_{2n}}{n}}_{\frac{1}{2} E(x_2)}$$

$$\xrightarrow{ac} \frac{1}{2} E(x_1) + \frac{1}{2} E(x_2) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{4}$$

$$\bar{X}_{2n} \xrightarrow{ac} \frac{7}{12}$$

$$\bar{X}_{2n-1} = \frac{x_1 + \dots + x_{2n-1}}{2n-1} = \frac{x_1 + x_3 + \dots + x_{2n-1}}{2n-1} + \frac{x_2 + x_4 + \dots + x_{2n-2}}{2n-1}$$

$$= \frac{n}{2n-1} \underbrace{\frac{x_1 + x_3 + \dots + x_{2n-1}}{n}}_{\frac{n-1}{2n-1} E(x_1)} + \frac{n-1}{2n-1} \underbrace{\frac{x_2 + x_4 + \dots + x_{2n-2}}{n-1}}_{E(x_2)} \xrightarrow{ac} \frac{7}{12}$$

$$\text{Como } \bar{X}_{2n} \xrightarrow{ac} \frac{7}{12} \quad \bar{X}_{2n-1} \xrightarrow{ac} \frac{7}{12}$$



$$A = \{w : \bar{X}_{2n}(w) \rightarrow \frac{7}{12}\} \quad B = \{w : \bar{X}_{2n-1}(w) \rightarrow \frac{7}{12}\}$$

$$w_0 \in A \cap B \Rightarrow$$

$$\Rightarrow \bar{X}_{2n}(w_0) \rightarrow \frac{7}{12} \Rightarrow \bar{X}_n(w_0) \rightarrow \frac{7}{12} \Rightarrow$$

$$\bar{X}_{2n-1}(w_0) \rightarrow \frac{7}{12}$$

$$\Rightarrow w_0 \in \{w : \bar{X}_n(w) \rightarrow \frac{7}{12}\}$$

$$A \cap B \subseteq$$

$$\Rightarrow \text{TP}(\bar{X}_n \rightarrow \frac{7}{12}) = 1 \quad \text{e., i., } \bar{X}_n = \underbrace{x_1 + \dots + x_n}_{n} \xrightarrow{\text{av}} \frac{7}{12}$$

(4) $(X_n)_{n \geq 1}$ v.a.i.id $\text{Ber}(p)$ $p \in (0,1)$

$$Y_1 = X_1 \cdot X_2 \cdot X_3$$

$$Y_2 = X_2 \cdot X_3 \cdot X_4 \sim \underbrace{Y_1 + Y_2 + \dots + Y_n}_{n} \xrightarrow{D} p^3$$

$$Y_n = X_n \cdot X_{n+1} \cdot X_{n+2} \quad \text{Verif. que } Y_1 \text{ e } Y_2 \text{ sao ind.}$$

Ideia

$$\bar{Y}_{3n} = \frac{Y_1 + \dots + Y_{3n}}{3n}$$

$$\bar{Y}_{3n-1} \quad Y_2 \quad Y_5 \quad Y_8$$

$$\bar{Y}_{3n-2} \quad Y_3 \quad Y_6 \quad Y_9$$

Módulo II

Aula 1 - 03/05/2013

Luis Gustavo Esteves

Referências

M. H. De Groot (M. J. Schervish)

M. J. Schervish

G. G. Roussas

"Theory of Statistics"

Casella , Berger

Inferência Estatística

"Fazer um procedimento que faça com que você aprenda ou tome decisão a partir dos dados".

Exemplo (1).



parâmetro

0: número de bolas brancas na urna

espaço paramétrico

$$\Theta = \{0, 1, 2, 3, 4, 5\}$$

0 brancas

Contar o número de bolas brancas na urna.

5-0 verdes

$$X = (X_1, X_L) \rightarrow \text{Amostra}$$

$$X = \{0, 1\}^L \rightarrow \text{espaço amostral}$$

- Encontrar alguma relação funcional entre aquilo que queremos descobrir (θ) e X , a amostra extraída.

(2) (MODELO DE TAGUCHI PARA ATRIBUTOS)

*Notícias:
Ano (estimação de tempo parado, etc.)
Legis (2008)*

Processo que funciona direitinho até um determinado momento (desintento).

θ : 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

$$\Theta = \mathbb{N}^*$$

$$X = \mathbb{N}^*$$

(3) N bolas, N conhecida

θ_1 cor 1
θ_2 cor 2
\vdots
θ_K cor K

K = cores

$$\Theta = (\theta_1, \theta_2, \dots, \theta_K)$$

$$\mathbb{H} = \{(a_1, a_2, \dots, a_K) \in \mathbb{N}^k : a_1 + a_2 + \dots + a_K \leq N\}$$

$X = (X_1, \dots, X_K)$, onde X_i é o número de bolas da cor i na amostra, $i=1, \dots, K$

$$X = \{(a_1, \dots, a_K) \in \mathbb{N}^k : a_1 + a_2 + \dots + a_K \leq n\}$$

(4)

θ : tempo médio de vida de lâmpadas produzidas

$$\mathbb{H} = \mathbb{R}_+$$

$$X = (X_1, X_2, X_3, X_4)$$

$$X = \mathbb{R}_+^4$$

No exemplo (Relação funcional)

1.

$$P(X_1=1, X_2=1 | \Theta=j) = \frac{1}{5} \frac{(j-1)}{4}$$

2.

$$P(X=j | \Theta) = \prod_{\theta \in \Theta} m_j(\theta) / \prod_{\theta \in \Theta} m_{j-1}(\theta)$$



3.

$$P(X_1=x_1, \dots, X_N=x_N | \Theta=(\theta_1, \dots, \theta_N)) = \frac{\binom{\theta_1}{x_1} \binom{\theta_2}{x_2} \dots \binom{\theta_N}{x_N} \binom{N-\theta_1-\dots-\theta_N}{n-x_1-\dots-x_N}}{\binom{N}{n}}$$

"Distribuição hipergeométrica de variável".

4.

X_1, X_2, X_3, X_4 , dado Θ , condicionalmente e i.d. $\text{Exp}(1/\Theta)$.

↳ pois se fossem independentes condicionalmente de Θ , então X_1, \dots, X_4 seriam ind. de Θ , o que não faria sentido, pois não trariam informação sobre o mesmo.

$$f(x_1, x_2, x_3, x_4 | \Theta) = \prod_{i=1}^4 f(x_i | \Theta)$$

$\Theta = ?$

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 no Bayesiano:

$f(x, \theta)$ 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)

$$\Theta \sim W(60, 1, 2, 3, 4, 5)$$

$x_1, x_2 | \theta$ já conhecidos

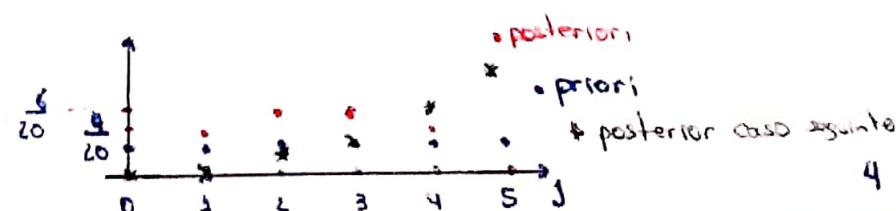
Experimento

$$x_1=1, x_2=0$$

\Rightarrow

$$P(\theta=3 | x_1=1, x_2=0) = \frac{P(x_1=1, x_2=0 | \theta=3) P(\theta=3)}{\sum_{i=0}^5 P(x_1=1, x_2=0 | \theta=i) P(\theta=i)} =$$

$$= \frac{\frac{1}{5} \cdot \frac{(5-3)}{4} \cdot \frac{1}{6}}{\sum_{i=0}^5 \frac{i}{5} \cdot \frac{(5-i)}{4} \cdot \frac{1}{6}} = \frac{\frac{1}{5} (5-3)}{\sum_{i=0}^5 i (5-i)}$$



E. Se $X_1=1 \circ X_2=1$,

$$\cdot P(\theta=3 | X_1=1, X_2=1) = \frac{P(X_1=1, X_2=1 | \theta=3) P(\theta=3)}{\sum_{i=0}^5 P(X_1=1, X_2=1 | \theta=i) P(\theta=i)}$$

$$\cdot \frac{\frac{1}{5} \frac{(3-1)}{4} \frac{1}{6}}{\sum_{i=0}^5 \frac{1}{5} \frac{(i-1)}{4} \frac{1}{6}} = \frac{3(3-1)}{\sum_{i=0}^5 (i-1)} \quad * \text{ posterior no gráfico anterior}$$

"Posterior de hoje é a priori de amanhã". Não há problemas em construir a posterior, dando informações passo a passo, da seguinte forma.

2)

On Geo ($\frac{1}{100}$)

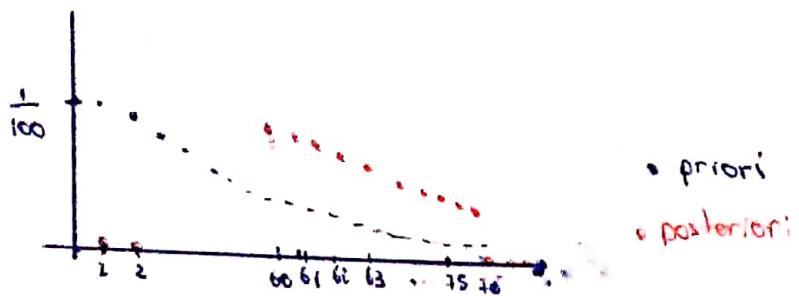
$$P(X=3 | \theta) = \prod_{\{0, 1, 2, \dots, 3, \dots, 99\}} (m_j)$$

$$\begin{aligned} m_1 &= 15 & P(\theta=3 | X=3) &= \frac{P(X=3 | \theta=3) P(\theta=3)}{\sum_{i=0}^{99} P(X=3 | \theta=i) P(\theta=i)} \\ X=3 & & & = \end{aligned}$$

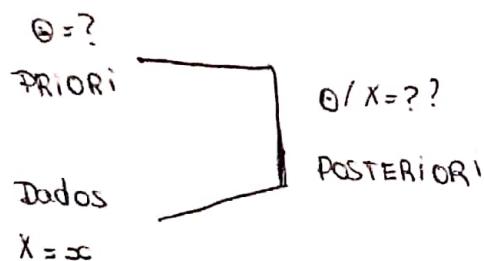
$$75 \pi_3 \in 60 \pi_{3-1}$$

$$\begin{aligned} &= \frac{\prod_{\{1, 2, \dots, 3, \dots, 99\}} (75) \prod_{\{1, \dots, 13-3\}} (60) \left(\frac{99}{100}\right)^{3-1} \frac{1}{100}}{\sum_{i=0}^{99} \prod_{\{1, 2, \dots, 3, \dots, 99\}} (75) \prod_{\{1, \dots, 13-3\}} (60) \left(\frac{99}{100}\right)^{i-1} \frac{1}{100}} = \frac{\prod_{\{1, \dots, 75\}} (j) \left(\frac{99}{100}\right)^{j-1}}{\sum_{i=0}^{\infty} \prod_{\{1, \dots, 75\}} (j) \left(\frac{99}{100}\right)^{j-1}} \rightarrow \end{aligned}$$

$$\Rightarrow P(\theta=3 | X=3) = \frac{\left(\frac{99}{100}\right)^{j-1}}{\sum_{i=0}^{99} \left(\frac{99}{100}\right)^{i-1}} \mathbb{I}_{\{1, \dots, 75\}}(j)$$



Operação Bayesiana



Vimos que existe uma função que relaciona o resultado do experimento (x) com parâmetro (θ).

$$f(x|\theta)$$

$$\mathbb{P}(x=x|\theta)$$

(1) Quando fixamos $\theta \in \Theta$,

$f(\cdot|\theta)$ é uma distribuição de probabilidade em \mathcal{X} . (Distribuição Amostral)

(2) Quando fixamos $x \in \mathcal{X}$,

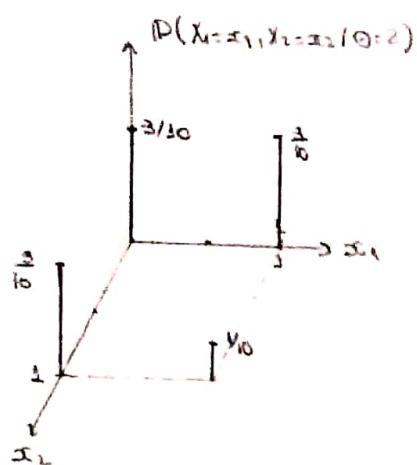
$f(x|\cdot)$ é uma função não-negativa de θ

$$V_x(\cdot) \quad (\text{Função de Verossimilhança para } \theta \text{ gerada por } x)$$

$$L_x(\cdot)$$

Voltando ao Exemplo 1.

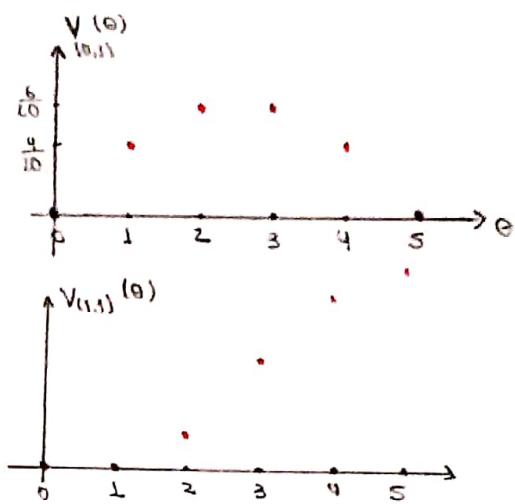
$$\Theta = 2$$



Tantos quantos os elementos de Θ .

$$x = (0,1) \in \mathcal{X}$$

$$P(X=(0,1) | \Theta) = P(X_1=0, X_2=1 | \Theta) = \frac{(5-\Theta)}{5} \cdot \frac{\Theta}{4}$$



$$V_{(1,1)}(\Theta) = \frac{\Theta(\Theta-1)}{10}$$

$$\sum V_{(1,1)}(\Theta) \neq 1$$

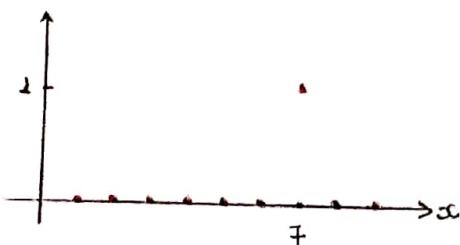
Tantos quantos os elementos em \mathcal{X} .

No exemplo 2.

$$\Theta = 100$$

$$m \approx 15$$

$$P(X=7 | \Theta=100) = 1$$



Tantos gatos o valores de Θ .

Um conjunto infinito enumerável de funções

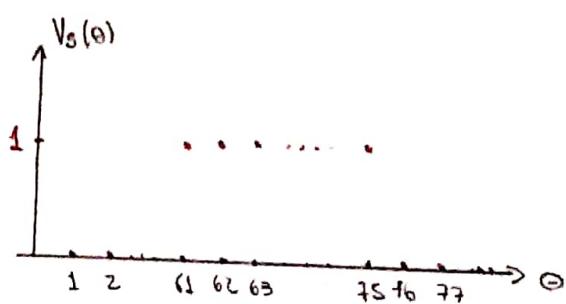
Note que $\Theta = 93, 101, \dots, 105$ levam a mesma $P(X=7 | \Theta = \dots)$.

Nesse caso, os pontos 93, ..., 105 são ditos observationalmente equivalentes, pois afirmam a mesma informação.

$$Y=5$$

$$V_5(\Theta) = P(X=5 | \Theta) = \mathbb{I}_{\{0\}}(\Theta)$$

$$\{0, \dots, 75\}$$



Probabilidade e Inferência Estatística I

Aula 2 - 07/05/2013

Luis Gustavo Esteves

Operações Bayesiana.

Verossimilhança gerada
para Θ

$$P(\Theta | \text{Dados}) = \frac{P(\text{Dados} | \Theta) P(\Theta)}{\sum_i P(\text{Dados} | i) P(i)} \rightarrow \text{Dist. a priori}$$

$$\downarrow \quad \sum_i P(\text{Dados} | i) P(i)$$

Distribuição a
posteriori

(1) Temporal Coherence Michael Goldstein (1985)

Bayesian Statistics (Valência)

(2) "Imprevistos e suas consequências". Rosângela H. Loschi (1998)

(1) Urna
(2) Qualidade } Exemplos

$$P(X=x | \Theta)$$

(1) Fixado $B \in \Theta$, $P(X=x | \Theta)$ é uma distribuição de probabilidade em \mathcal{X} .

(2) Fixado $x \in \mathcal{X}$, $P(X=x | \cdot)$ é chamada função de verossimilhança gerada por x .

$$V_x: \Theta \rightarrow \mathbb{R}$$

$$\Theta \ni \theta \mapsto V_x(\theta) = P(X=x | \theta)$$

Objetivo

$$f(\theta | x) = \frac{f(x | \theta) f(\theta)}{f(x)}$$

$$\text{com } f(x) = \int_{\Theta} f(\theta, x) d\theta$$

Resultado

X_1, X_2, \dots, X_n , dado Θ , são i.i.d. $Ber(\theta)$.

$$\Theta = (0,1) \quad \mathcal{X} = \{0,1\}^n$$

$\Theta \sim Beta(a, b) \Rightarrow a, b > 0$

$$\Theta | X_1=x_1, X_2=x_2, \dots, X_n=x_n \sim Beta(a + \sum_{i=1}^n x_i, b+n - \sum_{i=1}^n x_i)$$

$$f(\theta | X_1=x_1, \dots, X_n=x_n) = \frac{P(X_1=x_1, \dots, X_n=x_n | \theta)}{P(X_1=x_1, \dots, X_n=x_n)} f(\theta) =$$

$$= \frac{P(X_1=x_1, \dots, X_n=x_n | \theta) f(\theta)}{\int_0^1 P(X_1=x_1, \dots, X_n=x_n | \theta) f(\theta) d\theta} =$$

$$= \frac{\prod_{i=1}^n P(X_i=x_i | \theta) f(\theta)}{\int_0^1 \left(\prod_{i=1}^n P(X_i=x_i | \theta) \right) f(\theta) d\theta} = \frac{\prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} \mathbb{I}_{(0,1)}(\theta)}{\int_0^1 \prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} d\theta} =$$

$$= \frac{\theta^{\sum x_i} (1-\theta)^{n-\sum x_i} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} \mathbb{I}_{(0,1)}(\theta)}{\int_0^1 \theta^{\sum x_i} (1-\theta)^{n-\sum x_i} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} d\theta} =$$

~~Integrando~~

$$= \frac{\theta^{a+\sum_{i=1}^n x_i - 1} (1-\theta)^{b+n-\sum_{i=1}^n x_i - 1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mathbb{I}_{(0,1)}(\theta)}{\int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a+\sum x_i - 1} (1-\theta)^{b+n-\sum x_i - 1} d\theta} =$$

$$= \frac{\theta^{a+\sum_{i=1}^n x_i - 1} (1-\theta)^{b+n-\sum_{i=1}^n x_i - 1} \mathbb{I}_{(0,1)}(\theta)}{\int_0^1 \frac{\Gamma(a+\sum x_i + b+n-\sum x_i)}{\Gamma(a+\sum x_i)\Gamma(b+n-\sum x_i)} \theta^{a+\sum x_i - 1} (1-\theta)^{b+n-\sum x_i - 1} d\theta} \Rightarrow$$

$$f(\theta | X_1=x_1, \dots, X_n=x_n) = \frac{\Gamma(a+b+n)}{\Gamma(a+\sum x_i) \Gamma(b+n-\sum x_i)} \theta^{a+\sum x_i} (1-\theta)^{b+n-\sum x_i-1} \prod_{i=1}^n \frac{1}{x_i!}$$

Comentário

$\Theta \sim \text{Beta}(1,4)$

$$\text{Var}(\theta) = \frac{1.4}{5.5.6} = \frac{4}{150} = \frac{2}{75} > 2\%$$

$n=?$ $\text{Var}(\theta | X_1=x_1, \dots, X_n=x_n) \leq 1/1000$, $\forall (x_1, \dots, x_n) \in [0,1]^n$

$$\text{Var}(\theta | X_1=x_1, \dots, X_n=x_n) = \frac{(a+\sum x_i)(b+n-\sum x_i)}{(a+b+n)(a+b+n+1)} \leq \frac{1/4}{a+b+n+1}$$

$$\leq \frac{1}{1000} \Rightarrow n \geq 250 - a - b - 1 \quad (\text{Raiffa Schallalifer})$$

$$E(\theta) = \frac{a}{a+b}$$

$$E(\theta | X_1=x_1, \dots, X_n=x_n) = \frac{a+\sum x_i}{a+b+n} = \frac{a+n\bar{x}}{a+b+n} =$$

$$= \frac{(a+b)}{a+b+n} \cdot \frac{a}{(a+b)} + \frac{n\bar{x}}{a+b+n} = \frac{(a+b)}{a+b+n} \cdot \frac{a}{a+b} + \cancel{\frac{n}{a+b+n}} \stackrel{n \rightarrow \infty}{\rightarrow} \bar{x}$$

— n — " —

Espaço Parâmetrônico $\Theta = \mathbb{R}_+$ $\mathcal{X} = \mathbb{R}_+^n$ relativo a X_n

X_1, X_2, \dots, X_n dado θ são i.i.d. $\text{Exp}(\theta)$.

$\theta \sim \text{Gama}(a, b)$, $a, b > 0$

$\theta | X_1=x_1, X_2=x_2, \dots, X_n=x_n \sim \text{Gama}(a+n, b+\sum_{i=1}^n x_i)$

$$f(\theta/x_1=x_1, \dots, x_n=x_n) = \frac{f(x_1, \dots, x_n|\theta) f(\theta)}{f(x_1, \dots, x_n)} \propto f(x_1, \dots, x_n|\theta) f(\theta) \propto$$

$$\propto \prod_{i=1}^n f(x_i|\theta) f(\theta) \propto \prod_{i=1}^n \theta e^{-\theta x_i} I_{R^+}(x_i) \cdot \frac{\theta^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} I_{R^+}(\theta) \propto$$

$$\propto \theta^n e^{-\theta \sum x_i} I_{R^+}(\theta) \theta^{a-1} e^{-b\theta} \Rightarrow$$

$$f(\theta/x_1, \dots, x_n) \propto \theta^{n+a-2} e^{-(\theta + \sum x_i)\theta} I_{R^+}(\theta)$$

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

$\theta \sim N(\mu_0, \sigma_0^2)$
conhecidos

$\theta / x_1=x_1, x_2=x_2, \dots, x_n=x_n \sim$

$$f(\theta/x_1, \dots, x_n) = \frac{f(x_1, \dots, x_n|\theta) f(\theta)}{f(x_1, \dots, x_n)} \propto f(x_1, \dots, x_n|\theta) f(\theta) \propto$$

$$\propto \prod_{i=1}^n f(x_i|\theta) f(\theta) \propto \prod_{i=1}^n \frac{1}{\sigma_0 \sqrt{2\pi}} e^{-\frac{(x_i-\theta)^2}{2\sigma_0^2}} \cdot \frac{1}{b_0 \sqrt{2\pi}} e^{-\frac{(\theta-\mu_0)^2}{2b_0^2}} =$$

$$= \left(\frac{1}{\sigma_0 \sqrt{2\pi}} \right)^n e^{-\sum_{i=1}^n \frac{(x_i-\theta)^2}{2\sigma_0^2}} \cdot \frac{1}{b_0 \sqrt{2\pi}} e^{-\frac{(\theta-\mu_0)^2}{2b_0^2}} \propto$$

$$\propto e^{-\frac{1}{2} \left\{ \frac{\sum_{i=1}^n (x_i^2 - 2x_i\theta + \theta^2)}{\sigma_0^2} + \frac{\theta^2 - 2\mu_0\theta + \mu_0^2}{b_0^2} \right\}} \propto$$

$$\propto e^{-\frac{1}{2} \left[\frac{\sum x_i^2 - 2\theta \sum x_i + \theta^2}{\sigma_0^2} + \frac{\theta^2 - 2\theta_0 \theta + \theta_0^2}{b\sigma^2} \right]} \propto$$

$$\propto e^{-\frac{1}{2} \left\{ \theta^2 \left(\frac{n}{\sigma_0^2} + \frac{1}{b\sigma^2} \right) - 2\theta \left(\frac{\sum x_i}{\sigma_0^2} + \frac{\theta_0}{b\sigma^2} \right) + \frac{\sum x_i^2 + \theta_0^2}{\sigma_0^2} \right\}} \propto$$

$$\propto e^{-\frac{1}{2} \left(\frac{n}{\sigma_0^2} + \frac{1}{b\sigma^2} \right) \left\{ \theta^2 - 2\theta \frac{\left(\frac{\sum x_i}{\sigma_0^2} + \frac{\theta_0}{b\sigma^2} \right)}{\left(\frac{n}{\sigma_0^2} + \frac{1}{b\sigma^2} \right)} + \left[\frac{\left(\frac{\sum x_i}{\sigma_0^2} + \frac{\theta_0}{b\sigma^2} \right)^2}{\left(\frac{n}{\sigma_0^2} + \frac{1}{b\sigma^2} \right)} \right] \right\}} \Rightarrow$$

$$= \frac{1}{2 \left(\frac{n}{\sigma_0^2} + \frac{1}{b\sigma^2} \right)^{-1}} \left(\theta - \frac{b\sigma^2 \sum x_i + \theta_0^2 \sigma_0^2}{nb\sigma^2 + \sigma_0^2} \right)^2$$

$$f(\theta / x_1 = x_1, \dots, x_n = x_n) \propto e^{-\frac{1}{2} \left(\frac{n}{\sigma_0^2} + \frac{1}{b\sigma^2} \right)^{-1} \left(\theta - \frac{b\sigma^2 \sum x_i + \theta_0^2 \sigma_0^2}{nb\sigma^2 + \sigma_0^2} \right)^2}$$

$$\theta / x_1 = x_1, \dots, x_n = x_n \sim N \left(\frac{b\sigma^2 n \bar{x} + \theta_0^2 \sigma_0^2}{nb\sigma^2 + \sigma_0^2}, \left(\frac{n}{\sigma_0^2} + \frac{1}{b\sigma^2} \right)^{-1} \right)$$

$$\Theta = (\theta_1, \theta_2) \quad \mathcal{X} = \mathbb{R}^n \quad \Theta \subseteq \mathbb{R} \times \mathbb{R}_+$$

$$x_1, \dots, x_n \text{ dado } \theta : \text{ i.i.d. } N(\theta_1, \theta_2)$$

$$\theta_2 \sim \text{Gama}(a, b), \quad a, b > 0$$

$$\theta_1 / \theta_2 \sim N(c, d/\theta_2), \quad c \in \mathbb{R}, \quad d > 0$$

$$f(\theta_1, \theta_2) = f(\theta_2) f(\theta_1 / \theta_2)$$

$$(\theta_1, \theta_2) / x_1 = x_1, \dots, x_n = x_n \sim$$

$$f(\theta_1, \theta_2 / x_1, \dots, x_n) \propto f(x_1, \dots, x_n / \theta_1, \theta_2) f(\theta_1, \theta_2) \propto$$

$$\propto \prod_{i=1}^n f(x_i / \theta_1, \theta_2) f(\theta_1, \theta_2) f(\theta_2) \propto$$

$$\propto \prod_{i=1}^n \frac{1}{\sqrt{\frac{1}{\theta_2} + \frac{1}{2\pi}}} e^{-\frac{(x_i - \theta_1)^2}{2\frac{1}{\theta_2}}} \cdot \frac{1}{\sqrt{\frac{d}{\theta_2} + 2\pi}} e^{-\frac{(d - \theta_2)^2}{2\frac{d}{\theta_2}}} \cdot \frac{b}{\Gamma(a)} \theta_2^{a-1} e^{-b\theta_2} I_{\theta_2}^a(\theta_2)$$

$$\propto \left(\prod_{i=1}^n \theta_2^{b_i} \right) e^{-\frac{\theta_2}{2} \sum_{i=1}^n (x_i - \theta_1)^2} \cdot \theta_2^{a+1} e^{-\frac{\theta_2(b_0 - c)}{2d}} \theta_2^{a+L} e^{-b\theta_2} \propto$$

$$\propto \theta_2^{a+\frac{n+1}{2}-1} e^{-b\theta_2} \cdot e^{-\frac{\theta_2}{2} \left\{ \frac{1}{m} \sum_{i=1}^n x_i^2 - 2\theta_1 \sum_{i=1}^n x_i + n\theta_1^2 \right\}} \cdot e^{-\left[\frac{\theta_2 b_0^2}{2d} - \frac{2\theta_1 b_0 c}{2d} + \frac{\theta_2 c^2}{2d} \right]}$$

$$\propto \theta_2^{a+\frac{n+1}{2}-1} e^{-\theta_2 \left(b + \frac{\sum x_i^2}{2} + \frac{c^2}{2d} \right)} e^{-\frac{\theta_2}{2} \left\{ \theta_1 \left(n + \frac{1}{d} \right) - 2\theta_1 \left(\sum x_i + \frac{c}{d} \right) \right\}}$$

$$\propto \theta_2^{a+\frac{n+1}{2}-1} e^{-\theta_2 \left(b + \frac{\sum x_i^2}{2} + \frac{c^2}{2d} \right)} e^{-\frac{\theta_2}{2} \left(n + \frac{1}{d} \right)} \left\{ \theta_1^2 - \frac{2\theta_1 (\sum x_i + \frac{c}{d})}{(n+1/d)} + \left[\frac{\sum x_i + c/d}{n+1/d} \right]^2 \right\} \times$$

$$\times e^{\frac{\theta_2}{2} \left(n + \frac{1}{d} \right) \left[\frac{\sum x_i + c/d}{n+1/d} \right]^2}$$

$$\propto \underbrace{\theta_2^{a+L} e^{-\theta_2 b}}_{f(\theta_2 | x_1, \dots, x_n)} \underbrace{e^{-\frac{1}{2[\theta_2(n+1/d)]^{-1}} (\theta_1 - c')^2}}_{f(\theta_1 | \theta_2, x_1, \dots, x_n)}$$

Prob. e Inf. Est. I

10/05/2013 - Aula 3

Luis Gustavo Esteves

No 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 Statistical Decisions (CAP-11) → ver sobre conjugações

$A \leq B$ (relação de ordem dizendo que acredito mais em B do que A)

$\begin{cases} A \leq B \\ A \leq C \end{cases} \Rightarrow B \leq C$ (Capítulo 6 do Optimal Statistical Decision. Construção de medida de subjetividade.

→ Rosângela H. Loschi (1992)

→ S. French, D. Insua Rios

→ Cadani

$\Theta \sim \text{Beta}(\alpha, \beta)$

X_1, \dots, X_n dado θ e i.i.d. $\text{Ber}(\theta)$

$\theta | X_1 = x_1, \dots, X_n = \text{Bin} n \text{ Beta}(\alpha + \sum_{i=1}^n x_i, \beta + n - \sum_{i=1}^n x_i)$

"Preserva a qualidade da inferência, tudo que vc precisa da amostra suficiente"

$X((x_1, \dots, x_n))$ amostra

Def. Qualquer transformação $T: \mathcal{X} \rightarrow \mathbb{R}$ mensurável é chamada Estatística.

Exemplo: $X = (x_1, \dots, x_n)$, x_i a valores reais, $i=1, \dots, n$ ($\mathcal{X} = \mathbb{R}^n$).

$$T_1(X) = \frac{x_1 + \dots + x_n}{n} = \bar{x}_n$$

$$T_2(X) = X_{(n)} = \max\{x_1, \dots, x_n\}$$

$$T: \mathcal{X} \rightarrow \mathbb{R}_+$$

$$(x_1, \dots, x_n) \mapsto T(x_1, \dots, x_n) = \left(\bar{x}, \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \right)$$

$$T_3(x_1, \dots, x_n) = (x_1, \dots, x_n)$$

$$\text{seja } \mathcal{U} = \{f: \Theta \rightarrow \mathbb{R}\}$$

$$H: X \rightarrow \mathcal{U}$$

$$x \in \mathcal{X} \rightarrow H(x) = V_{x \in (\cdot)}$$

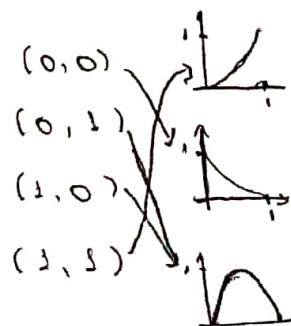
Exemplo.

X_1, X_2 , dado Θ , e i.i.d. $\text{Ber}(\theta)$.

$$\mathcal{X} = \{0,1\}^2$$

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: $X((x_1, \dots, x_n))$. $T: \mathcal{X} \rightarrow \mathcal{T}$ é uma estatística suficiente para Θ se para toda priori para Θ ,

$$\Theta | X=x \stackrel{d}{=} \Theta | T(x)=t(x), \forall x \in \mathcal{X}.$$

Exemplo 1:

X_1, \dots, X_n dado Θ i.i.d. $\text{Ber}(\theta)$.

$$\mathcal{X} = \{0, 1\}^n \quad \Theta = (0, 1)$$

$$f(\Theta | x_1, \dots, x_n) \propto f(x_1, \dots, x_n | \theta) f(\theta)$$

$$\propto \prod_{i=1}^n P(X_i=x_i | \theta) p(\theta) \propto \theta^{\sum x_i} (1-\theta)^{n-\sum x_i} f(\theta)$$

$$f(\Theta | \sum_{i=1}^n X_i = \sum_{i=1}^n x_i) \propto P\left(\sum_{i=1}^n X_i = \sum_{i=1}^n x_i | \theta\right) f(\theta)$$

$$\propto \binom{n}{\sum x_i} \theta^{\sum x_i} (1-\theta)^{n-\sum x_i} f(\theta)$$

$$\forall x \in \mathcal{X}$$

$$\Theta | X_1=x_1, \dots, X_n=x_n \stackrel{d}{=} \Theta | \sum_{i=1}^n X_i = \sum_{i=1}^n x_i. \text{ Logo } \sum_{i=1}^n X_i \text{ é suficiente para } \Theta.$$

Exemplo 2.

X_1, \dots, X_n dado Θ com i.i.d. $\text{Exp}(\theta)$

$$\mathcal{X} = \mathbb{R}_+^n \quad \Theta = \mathbb{R}_+$$

$$f(\theta | x_1, \dots, x_n) \propto f(x_1, \dots, x_n | \theta) f(\theta) \propto$$

$$\propto \left(\prod_{i=1}^n f(x_i | \theta) \right) \cdot f(\theta) \propto \left(\prod_{i=1}^n \theta e^{-\theta x_i} I_{R_+}(x_i) \right) \cdot f(\theta) \Rightarrow$$

$$\Rightarrow f(\theta | x_1, \dots, x_n) \propto \theta^n e^{-\theta \sum_{i=1}^n x_i} f(\theta).$$

$$T(X) = \sum_{i=1}^n X_i$$

$$f(\theta | T(X) = T(x_1, \dots, x_n)) = f(\theta | \sum_{i=1}^n X_i = \sum_{i=1}^n x_i) \propto$$

$$\propto f_{\sum_{i=1}^n X_i}(\sum_{i=1}^n x_i | \theta) \cdot f(\theta) \propto \frac{\theta^n}{\Gamma(n)} (\sum_{i=1}^n x_i)^{n-1} e^{-\theta \sum_{i=1}^n x_i} I_{R_+}(\sum_{i=1}^n x_i) f(\theta)$$

$$\Rightarrow f(\theta | \sum_{i=1}^n x_i) \propto \theta^n e^{-\theta \sum_{i=1}^n x_i} f(\theta)$$

Como $\theta / X_i = x_1, \dots, X_n = x_n \stackrel{d}{=} \theta / \sum_{i=1}^n X_i = \sum_{i=1}^n x_i$, segue que $\sum_{i=1}^n X_i$ é suficiente para θ .

Exemplo 3.

X_1, \dots, X_m dado θ , i.i.d. $U(0, \theta)$, $\theta > 0$

$$\mathcal{X} = \mathbb{R}_+^n \quad \Theta = \mathbb{R}_+$$

$$f(\theta | x_1, \dots, x_n) \propto f(x_1, \dots, x_n | \theta) f(\theta) \propto$$

$$\propto \left(\prod_{i=1}^n f(x_i | \theta) \right) f(\theta) \propto \left(\prod_{i=1}^n \frac{1}{\theta} I_{(0, \theta)}(x_i) \right) f(\theta) \propto$$

$$\propto \frac{1}{\theta^n} \prod_{i=1}^n I_{(0, \theta)}(x_i) \prod_{i=1}^n I_{(0, \theta)}(x_{i,m}) f(\theta) \Rightarrow f(\theta | x_1, \dots, x_n) \propto \frac{1}{\theta^n} I_{(0, \theta)}(x_{i,m}) f(\theta)$$

$$\prod_{i=1}^n \mathbb{I}_{(0,\theta)}(x_i) = 1 \Leftrightarrow \mathbb{I}_{(0,\theta)}(x_i) = 1, \forall i=1, \dots, n \Leftrightarrow$$

$$0 < x_i < \theta, \forall i=1, \dots, n \Leftrightarrow 0 < x_{(n)} < x_{(1)} < \theta.$$

$$\Leftrightarrow \mathbb{I}_{(0,x_{(n)})}(x_{(n)}) \mathbb{I}_{(0,\theta)}(x_{(1)}) = 1$$

$$X_{(n)} = \max\{X_1, \dots, X_n\}$$

$$f(\theta/x_{(n)}) \propto f(x_{(n)}/\theta) f(\theta)$$

$$\propto \frac{n(x_{(n)})^{n-1}}{\theta^n} \mathbb{I}_{(0,\theta)}(x_{(n)}) f(\theta)$$

$$\propto \frac{1}{\theta^n} \mathbb{I}_{(0,\theta)}(x_{(n)}) f(\theta)$$

$\forall x \in \mathcal{X}$

$$\text{Logo, } \theta/x_1=x_1, \dots, x_n=x_n \stackrel{d}{=} \theta/x_{(n)}=x_{(n)}$$

Assim, $T(X) = X_{(n)}$ é suficiente para Θ .

$$P(\theta=j | X=x) = P(\theta=j | T(x)=T(x)) \Leftrightarrow$$

$$\frac{P(X=x | \theta=j) P(\theta=j)}{P(X=x)} = \frac{P(T(X)=T(x) | \theta=j) P(\theta=j)}{P(T(X)=T(x))} \Leftrightarrow$$

$$\frac{P(X=x | \theta=j)}{P(X=x)} = \frac{P(T(X)=T(x) | \theta=j)}{P(T(X)=T(x))} \Leftrightarrow$$

$$\frac{P(X=x | \theta=j)}{P(T(X)=T(x) | \theta=j)} = \frac{P(X=x)}{P(T(X)=T(x))} \Leftrightarrow$$

$$\frac{P(X=x, T(x)=T(x) / \Theta=j)}{P(T(x)=T(x) / \Theta=j)} = \frac{P(X=x)}{P(T(x)=T(x))} \Leftrightarrow$$

$$P(X=x / T(x)=T(x), \Theta=j) = \frac{P(X=x)}{P(T(x)=T(x))} \rightarrow \text{Não depende de } \Theta$$

→ Mostrou que a definição de suficiência bayesiana equivale à da clássica por meio do caso discreto.

Definição. $T: \mathcal{X} - \mathbb{J}$ 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

Pro casos contínuos e discretos, a volta é válida. Entretanto, existem casos em que a volta não é válida. (Slr, ...)

Exemplo. X_1, X_2 dado Θ , i.i.d. $Ber(\theta)$

$$T(X) = X_1 + X_2$$

$$P(X_1=x_1, X_2=x_2 / T(X)=t, \Theta) =$$

$$\frac{P(X_1=x_1, X_2=x_2, T(X)=t / \Theta)}{P(T(X)=t / \Theta)} = \begin{cases} 0, t \neq x_1+x_2 \\ \frac{\theta^{x_1+x_2} (1-\theta)^{2-x_1-x_2}}{\binom{2}{x_1+x_2} \theta^{x_1+x_2} (1-\theta)^{2-x_1-x_2}} = \frac{1}{\binom{2}{x_1+x_2}} \cdot \frac{1}{\binom{2}{t}} \end{cases}$$

Logo, $T(X) = X_1 + X_2$ é suficiente para Θ .

$$S(x) = X_1 \cdot X_2$$

$$S(x)/\theta \sim \text{Ber}(\theta^2)$$

$$\begin{aligned} P(X_1=x_1, X_2=x_2 | S(x)=0, \theta) &= \\ \frac{P(X_1=x_1, X_2=x_2, S(x)=0 | \theta)}{P(S(x)=0 | \theta)} &= \begin{cases} 0, x_1 \cdot x_2 \neq 0 \\ \frac{(1-\theta)^2}{1-\theta^2} = \frac{(1-\theta)(1-\theta)}{(1-\theta)(1+\theta)} = \frac{(1-\theta)}{1+\theta}, x_1 \cdot x_2 = 0 \end{cases} \\ &\downarrow \quad \forall x_1=0, x_2=0 \\ &\text{que depende} \\ &\text{de } \theta. \end{aligned}$$

Logo, $S(x)$ NÃO é suficiente para θ .

$$\theta / X_1=0, X_2=0$$

$$f(\theta / X_1=0, X_2=0) \propto P(X_1=0, X_2=0 | \theta) f(\theta) \propto (1-\theta)^2 f(\theta)$$

~~H~~

$$\theta / S(x)=0$$

$$f(\theta / S(x)=0) \propto P(S(x)=0 | \theta) f(\theta) \propto (1-\theta^2) f(\theta)$$

Critério Da Fatoração (Fisher-Neyman)

~~O maior para~~ • No livre de Roussas, uma def. mais geral.

$X((x_1, \dots, x_n))$ amostrou

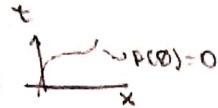
A estatística $T: \mathcal{X} \rightarrow J$ é suficiente para θ se e só se $\exists \mu: \mathcal{X} \rightarrow \mathbb{R}_+$ e $v: J \times \Theta \rightarrow \mathbb{R}_+$ tais que

$$f(x/\theta) = \mu(x) v(T(x), \theta), \forall x, \forall \theta.$$

(\Rightarrow)

$$P(X=x | \Theta) = P(X=x, T(X)=T(x) | \Theta) = \underbrace{P(X=x | T(X)=T(x), \Theta)}_{\mu(x)} \underbrace{P(T(X)=T(x), \Theta)}_{v(T(x), \Theta)}$$

A idéia pro caso contínuo



(\Leftarrow)

$$P(X=x | T(X)=t, \Theta) = \frac{P(X=x, T(X)=t | \Theta)}{P(T(X)=t | \Theta)} = \begin{cases} 0, & t \neq T(x) \\ \frac{P(X=x | \Theta)}{\sum_{y: T(y)=t} P(X=y | \Theta)}, & t = T(x) \end{cases}$$

$$\frac{P(X=x | \Theta)}{\sum_{y: T(y)=T(x)} P(X=y | \Theta)} = \frac{\mu(x) v(T(x), \Theta)}{\sum_{y: T(y)=T(x)} \mu(y) v(T(y), \Theta)} = \frac{\mu(x)}{\sum_{y: T(y)=t} \mu(y)} \text{ que não depende de } \Theta$$

Logo, T é suficiente!

Exemplos:

1. X_1, \dots, X_n , dado Θ , i.i.d. $N(\Theta, 1)$, $\Theta \in \mathbb{R}$

$$f(x_1, \dots, x_n | \Theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i - \Theta)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\sum_{i=1}^n \frac{(x_i - \Theta)^2}{2}} =$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum x_i^2 - 2\Theta \sum x_i + n\Theta^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum x_i^2}{2}} e^{\frac{+2\Theta \sum x_i - n\Theta^2}{2}}$$

$\mu(\omega)$ $v(2\sum x_i, \Theta)$

Pelo critério da fatoração

$T(X) = X_1 + \dots + X_n$ é suficiente para Θ .

Ex 2: x_1, \dots, x_n , dado Θ , i.i.d. $N(\mu, \sigma^2)$, $\theta \in \mathbb{R}$

$$f(x_1, \dots, x_n | \theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \underbrace{\left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i-\mu)^2}}_{v((\sum_{i=1}^n (x_i-\mu)^2), \sigma^2)} \cdot \frac{1}{\mu(x)}$$

Pelo critério da Fatoração, $S(X) = \sum_{i=1}^n (x_i - \mu)^2$ é suficiente para θ .

Exemplo 3:

x_1, \dots, x_n , dado $\Theta = (\theta_1, \theta_2)$, i.i.d. $N(\theta_1, \theta_2)$, $\theta_1 \in \mathbb{R}$, $\theta_2 > 0$

$$\mathcal{X} = \mathbb{R}^n \quad \Theta = \mathbb{R} \times \mathbb{R}_+$$

$$f(x_1, \dots, x_n | \theta) = \left(\frac{1}{\sqrt{2\pi\theta_2}} \right)^n e^{-\frac{\sum_{i=1}^n (x_i - \theta_1)^2}{2\theta_2}} =$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi\theta_2}} \right)^n e^{-\frac{\sum x_i^2 - 2\theta_1 \sum x_i + \theta_1^2 n}{2\theta_2}}}_{v((\sum x_i, \sum x_i^2), \theta_2)} \cdot \frac{1}{\mu(x)}$$

$$v((\sum x_i, \sum x_i^2), \theta_2)$$

Pelo critério de fatoração,

$T(X) = (\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2)$ é suficiente para θ .

↓

(x_1, \dots, x_n)

Exemplo 4.

x_1, \dots, x_n , dado θ_2 condicionalmente independentes tais que

$$x_i | \theta \sim N(\mu, \theta^2), i = 1, \dots, n$$

$$f(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i | \theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x_i-\mu)^2}{2\theta}} =$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \frac{1}{\sqrt{n!}}}_{\mathcal{V}\left(\sum_{i=1}^n \frac{(x_i-\mu)^2}{2}, \theta\right)} e^{-\frac{1}{2\theta} \sum_{i=1}^n \frac{(x_i-\mu)^2}{\lambda}} \cdot \underbrace{\frac{1}{\lambda^n}}_{\text{M}(x)}$$

Logo, pelo Criterio de Fatoração

$$T(X) = \sum_{i=1}^n \frac{(x_i-\mu)^2}{\lambda} \text{ é suficiente para } \Theta.$$

Aula 04 - 14/05

- Comparative Statistics Inference (Vic. Barnett)

24/05 às 11:00hs

Sala 252

Seminário de fundamentos

$$\Theta/X = x \stackrel{d}{=} \Theta/T(x) = T(x), \forall x$$

Critério da Fatoração.

Seguindo com os exemplos...

5) $X(x_1, \dots, x_n)$ amostrou

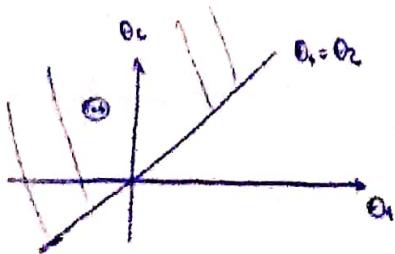
$$f(\underline{x}|\theta) = \underbrace{f(x_1|\theta)}_{v(x,\theta)} \cdot \underbrace{\dots}_{\vdots} \cdot \underbrace{f(x_n|\theta)}_{u(x)}$$

$T: \mathcal{X} \rightarrow \mathcal{X}$ dada por

$T(X) = X$ é suficiente para θ .

6) x_1, \dots, x_n dado $\theta = (\theta_1, \theta_2)$, i.i.d. $\mathcal{U}(\theta_1, \theta_2)$

$$\mathcal{C} = \{(a, b) \in \mathbb{R}^2 : a < b\} \quad \mathcal{X} = \mathbb{R}^n$$



$$(x_1, \dots, x_n) \in \mathcal{X}$$

$$f(x_1, \dots, x_n|\theta) = \prod_{i=1}^n f(x_i|\theta) = \prod_{i=1}^n \frac{1}{\theta_2 - \theta_1} \mathbb{I}_{(\theta_1, \theta_2)}(x_i) =$$

$$= \underbrace{\left(\frac{1}{\theta_2 - \theta_1}\right)^n}_{v((x_1, \dots, x_n), \theta)} \underbrace{\prod_{i=1}^n \mathbb{I}_{(\theta_1, \theta_2)}(x_i)}_{(\theta_1, \theta_2)} = \underbrace{\left(\frac{1}{\theta_2 - \theta_1}\right)^n}_{v((x_1, \dots, x_n), \theta)} \underbrace{\mathbb{I}_{(\theta_1, \theta_2)}}_{(\theta_1, \theta_2)} \underbrace{\prod_{i=1}^n \mathbb{I}_{(x_i, x_{i+1})}}_{\mathcal{L}(x)}$$

Pelo Critério da Fatoração,

$T(x) = (X_{(1)}, X_{(n)})$ é suficiente para Θ .

Resultado

Seja $T: \mathcal{X} \rightarrow J$ suficiente para Θ .

Seja $S: \mathcal{X} \rightarrow \mathcal{P}$ outra estatística tal que $S = g \circ T$, onde $g: J \rightarrow \mathcal{P}$ é bijetora (um a um). Então, S é suficiente para Θ .

$$S = g(T)$$

Existe $g^{-1}: \mathcal{P} \rightarrow J$

Como T é suficiente, então

$$f(x|\theta) = \mu(x) \cdot v(T(x), \theta) = \mu(x) \cdot v(g^{-1}(S(x)), \theta).$$

Definindo $v': \mathcal{P} \times \Theta \rightarrow \mathbb{R}_+$ por $v'(s, \theta) = v(g^{-1}(s), \theta)$, resulta que.

Assim,

$f(x|\theta) = \mu(x) \cdot v'(S(x), \theta)$. Logo, pelo Critério da Fatoração, S é suficiente para Θ .

Exemplo 1.

X_1, \dots, 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 $g: \mathbb{R} \rightarrow \mathbb{R}$ 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 .

$$S(X) = \left(\bar{X}_n, \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} \right) \quad \left[\frac{\sum_{i=1}^n X_i^2 - n(\bar{X}_n)^2}{n} \right]$$

Seja $g: \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R} \times \mathbb{R}_+$ dada por

$$g(a, b) = \left(\frac{a}{n}, \frac{b - n\left(\frac{a}{n}\right)^2}{n} \right)$$

g é bijetora. Logo,

$S(X) = \left(\bar{X}, \frac{\sum (X_i - \bar{X}_n)^2}{n} \right) = g \left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2 \right)$ é suficiente para Θ .

Exercício.

$$(X_{(1)}, X_{(n)}) \quad \left(\frac{X_{(1)} + X_{(n)}}{2}, X_{(n)} - X_{(1)} \right) \quad S(X) = \left(\frac{X_{(1)} + X_{(n)}}{2}, X_{(n)} - X_{(1)} \right) = g(X_{(1)}, X_{(n)})$$

$$\text{com } g: \mathbb{R}^2 \rightarrow \mathbb{R}^2 \text{ com } g(a, b) = \left(\frac{a+b}{2}, b-a \right)$$

Estatística Suf. Mínima = "Aquele que faz a 'maior redução' da amostra de forma a preservar a suficiência". "Se tentar reduzir ainda mais a E.S.M., perde-se a suficiência".

$T: \mathcal{X} \rightarrow \mathcal{S}$ é suficiente

Seja $T': \mathcal{X} \rightarrow \mathcal{S}'$ uma outra estatística suficiente.

Dizemos que $T: \mathcal{X} \rightarrow \mathcal{S}$ 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 $g: \mathcal{S}' \rightarrow \mathcal{S}$ tal que $T = g \circ T'$.

De 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 $T: \mathcal{X} \rightarrow \mathcal{S}$ uma estatística.

Se $\forall x_1, x_2 \in \mathcal{X}$,

$\frac{f(x_1 | \theta)}{f(x_2 | \theta)}$ não depende de $\theta \Leftrightarrow T(x_1) = T(x_2)$, então

T é suficiente mínima.

1. T é suficiente

$$P(X=x | T(X)=t, \theta) = \frac{\underset{t=T(x)}{P(X=x, T(X)=T(x)|\theta)}}{P(T(X)=T(x)|\theta)} = \sum_{\substack{y \in \mathcal{X}: \\ T(y)=T(x)}} P(X=y|\theta) =$$

$$= \frac{1}{\sum_{y \in \mathcal{X}: T(y)=T(x)} P(X=y|\theta)} \text{ não depende de } \theta!$$

pois $T(x_1) = T(x_2) \Rightarrow \frac{f(x_1|\theta)}{f(x_2|\theta)}$ não depende de θ

Como o dist. de X dado $T(X)=t$ é Θ não depende de Θ , $\forall t$, então
 T é suficiente para Θ .

2. Seja $T': \mathcal{X} \rightarrow \mathcal{S}$ uma estatística suficiente.

Sejam $x_1, x_2 \in \mathcal{X}$.

$$T'(x_1) = T'(x_2) \Rightarrow$$

$$\frac{P(X_1=x_1|\theta)}{P(X_2=x_2|\theta)} = \frac{\mu(x_1)\pi(T'(x_1),\theta)}{\mu(x_2)\pi(T'(x_2),\theta)} = \frac{\mu(x_1)}{\mu(x_2)} \Rightarrow$$

$$\frac{P(X_1 \leq x_1|\theta)}{P(X_2 \leq x_2|\theta)} \text{ não depende de } \theta \Rightarrow T(x_1) = T(x_2).$$

Logo, T é suficiente mínima.

Exemplos:

(J) X_1, \dots, X_n , dado, c.i.i.d. Beta($\theta, 1$)
 condicionalmente

$$\theta > 0, \mathcal{X} = (0,1)^n, \Theta = \mathbb{R}_+$$

Sejam $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathcal{X}$

$$\frac{f(x_1, \dots, x_n|\theta)}{f(y_1, \dots, y_n|\theta)} = \frac{\prod_{i=1}^n f(x_i|\theta)}{\prod_{i=1}^n f(y_i|\theta)} = \frac{\prod_{i=1}^n \theta x_i^{\theta-1} I_{(0,1)}(x_i)}{\prod_{i=1}^n \left\{ \theta y_i^{\theta-1} I_{(0,1)}(y_i) \right\}} =$$

$$\frac{\theta^n \left(\prod_{i=1}^n x_i \right)^{\theta-1} \prod_{i=1}^n I_{(0,1)}(x_i)}{\theta^n \left(\prod_{i=1}^n y_i \right)^{\theta-1} \prod_{i=1}^n I_{(0,1)}(y_i)} = \left(\frac{\prod_{i=1}^n x_i}{\prod_{i=1}^n y_i} \right)^{\theta-1} \frac{\prod_{i=1}^n I_{(0,1)}(x_i)}{\prod_{i=1}^n I_{(0,1)}(y_i)} \Leftrightarrow$$

$$\prod_{i=1}^n x_i = \prod_{i=1}^n y_i$$



$$T(x_1, \dots, x_n) = T(y_1, \dots, y_n), \text{ onde } T(x_1, \dots, x_n) = \prod_{i=1}^n x_i.$$

Logo, $T(X) = \prod_{i=1}^n X_i$ é suficiente mínima para Θ .

Exemplo 2. X_1, \dots, X_n dado $\Theta = (\theta_1, \theta_2)$, c.i.i.d. $N(\theta_1, \theta_2)$. $\mathcal{X} = \mathbb{R}^n$, $\Theta = \mathbb{R} \times \mathbb{R}_+$

$(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathcal{X}$.

$$\begin{aligned} \frac{f(x_1, \dots, x_n | \theta)}{f(y_1, \dots, y_n | \theta)} &= \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_i - \theta_1)^2}{2\theta_2}}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(y_i - \theta_1)^2}{2\theta_2}}} = \\ &= \frac{\left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n e^{-\frac{\sum (x_i - \theta_1)^2}{2\theta_2}}}{\left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n e^{-\frac{\sum (y_i - \theta_1)^2}{2\theta_2}}} = e^{\frac{\sum y_i^2 - 2\theta_1 \sum y_i + n\theta_1^2 - \sum x_i^2 + 2\theta_1 \sum x_i - n\theta_1^2}{2\theta_2}} \\ &= e^{\frac{2\theta_1 (\sum x_i - \sum y_i) + (\sum y_i^2 - \sum x_i^2)}{2\theta_2}} \quad \text{não depende de } \theta \end{aligned}$$



$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i = 0 \quad \text{e} \quad \sum_{i=1}^n y_i^2 - \sum_{i=1}^n x_i^2 = 0,$$



$$T(x_1, \dots, x_n) = T(y_1, \dots, y_n), \text{ onde}$$

$$T(x_1, \dots, x_n) = \left(\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2 \right)$$

Dado resultado,

$$T(X) = \left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2 \right) \in \text{sufliente m\'inima para } \Theta.$$

ESTIMAÇÃO

Θ brancas
5- Θ verdes

$$\Theta = \{0, 1, 2, 3, 4, 5\}$$

$$X = (x_1, x_2)$$

$$\mathcal{E} = \{0, 1\}^2$$

Uma estatística $T: \mathcal{E} \rightarrow \Theta$ é chamada Estimador para Θ .

$$\delta_1: \mathcal{E} \rightarrow \Theta$$

$$\delta_1(x_1, x_2) = \begin{cases} 0, & x_1 = x_2 = 0 \\ 2, & x_1 = 1, x_2 = 0 \\ 3, & x_1 = 0, x_2 = 1 \\ 4, & x_1 = 1, x_2 = 1 \end{cases}$$

$$\delta_2: \mathcal{E} \rightarrow \Theta$$

$$\delta_2(x_1, x_2) = \begin{cases} 1, & x_1 = x_2 = 0 \\ 3, & x_1 = 1, x_2 = 0 \\ 3, & x_1 = 0, x_2 = 1 \\ 5, & x_1 = x_2 = 1 \end{cases}$$

$$\delta_3(x_1, x_2) = 2, \forall (x_1, x_2) \in \mathcal{E}$$

$$\delta_4(x_1, x_2) = 0, \forall (x_1, x_2) \in \mathcal{E}$$

$\delta_i, i=1, 2, 3, 4$ são estimadores!

Dado

$$(0, 1), \delta_1(0, 1) = 3 \text{ é a estimativa para } \Theta$$

Observe,

$\delta(X)$ - estimador.

$\delta(x)$ - estimativa.

Em geral, $\Theta \subseteq \mathbb{R}^k$.

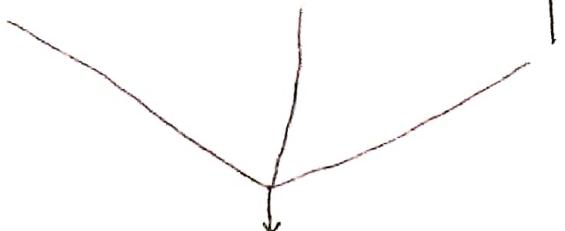
$$\Delta = \left\{ \delta: \{0,1\}^2 \rightarrow \{0,1,2,3,4,5\} \right\}$$

Δ , conjunto dos estimadores para Θ (nesse exemplo), possui 6^4 elementos.
1296

Como escolher o estimador nesse caso?

Devemos estabelecer critérios de ~~optimalidade~~ optimalidade.

$$|\delta(x) - \theta| \text{ ou } |\delta(x) - \theta|^2 \text{ ou } \left| \frac{\delta(x) - \theta}{\theta} \right|$$



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 optimalidade (penalidade), devemos escolher $\delta \in \Delta$ que minimize a penalidade média.

No nosso exemplo para $\delta \in \Delta$, vamos avaliar

$$p(\delta) = E[L(\delta(x), \theta)] = \sum_{x \in \mathcal{X}} \sum_{j=0}^5 L(\delta(x), j) P(x=x, \theta=j)$$

↓
risco

$$L: \Theta \times \Theta \rightarrow \mathbb{R}_+$$

$(\delta(x), \theta) \mapsto L(\delta(x), \theta)$: penalidade (perda) que o decisor incorre ao estimar $\delta(x)$ ao observar $x \in \mathcal{X}$.

Ideia: Escolher $\delta^* \in \Delta$ tal que

$$\rho(\delta^*) = \min \{ \rho(\delta) : \delta \in \Delta \}$$

Podemos reescrever $\rho(\delta)$ da seguinte maneira

$$\begin{aligned} \rho(\delta) &= \sum_{x \in \mathcal{X}} \sum_{j=0}^5 L(\delta(x), \theta_j) P(\theta=j | X=x) P(X=x) = \\ &= \sum_{x \in \mathcal{X}} P(X=x) \left\{ \sum_{j=0}^5 L(\delta(x), \theta_j) P(\theta=j | X=x) \right\} \end{aligned}$$

Ideia:

Para cada $x \in \mathcal{X}$, tomamos $\delta^*(x) \in \Theta$ que minimize (em d)

$$\sum_{j=0}^5 L(d_j, \theta_j) P(\theta=j | X=x)$$

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 .

Poderemos escrever $\rho(\delta)$ da seguinte maneira:

$$\begin{aligned}\rho(\delta) &= \sum_{\theta \in \Theta} \sum_{j=0}^s L(\delta(\theta), \theta) P(\theta=j | X=x) P(X=x) = \\ &= \sum_{x \in \mathcal{X}} P(X=x) \left\{ \sum_{j=0}^s L(\delta(\theta_j), \theta_j) P(\theta=j | X=x) \right\} \quad (\text{4 posteriores} \\ &\quad \text{cond.})\end{aligned}$$

Estatística: Para cada $x \in \mathcal{X}$, tomamos $\delta^*(x) \in \Theta$ que minimiza

$$\sum_{j=0}^s L(\delta_j, \theta_j) P(\theta=j | X=x).$$

~~P~~

Ao estatístico que associa a cada $x \in \mathcal{X}$ $\delta^*(x)$ conforme especificado acima, damos o nome de Estatístico de Bayes para θ considerando a função L .

Aula 5 - 15/05/2013

Revisão: última aula $\mathcal{X}, \Theta, (\theta, x), \Delta = \{\delta : \mathcal{X} \rightarrow \Theta\}$

$\delta : \mathcal{X} \rightarrow \Theta$

Para cada $S \in \Delta$, avaliamos

$$\rho(\delta) = E[L(\delta(\theta), \theta)] = \sum_{x \in \mathcal{X}} \sum_{\theta \in \Theta} L(\delta(\theta), \theta) P(X=x, \theta)$$

$$\Rightarrow \rho(\delta) = \sum_{x \in \mathcal{X}} P(X=x) \left\{ \sum_{\theta \in \Theta} L(\delta(\theta), \theta) P(\theta | X=x) \right\}$$

Risco de Estatística

Considere $\delta^* : \mathcal{X} \rightarrow \Theta$ que associa a cada $x \in \mathcal{X}$, o valor $\hat{\theta} \in \Theta$ tal que:

$$\sum_{j \in \Theta} L(\hat{\theta}, j) P(\theta=j | X=x) = \min_{\hat{\theta} \in \Theta} \sum_{j \in \Theta} L(\hat{\theta}, j) P(\theta=j | X=x)$$

Isto é, $\delta^*(x) = \hat{\theta}^*(x)$ depende da posteriori $E[L(\hat{\theta}, \theta) | X=x]$

É fácil ver que $p(\delta^*) \leq p(\delta)$, $\forall \delta \in \Delta$ é chamado ESTIMADOR DE BAYES e seu risco é chamado risco de Bayes.

Com efeito,

$$\begin{aligned} p(\delta^*) &= \sum_{x \in \mathcal{X}} P(X=x) \left\{ \sum_{j \in \Theta} L(\delta^*(x), j) P(\theta=j | X=x) \right\} \\ &\leq \sum_{x \in \mathcal{X}} P(X=x) \left\{ \sum_{j \in \Theta} L(\delta(x), j) P(\theta=j | X=x) \right\} = p(\delta), \forall \delta \in \Delta \end{aligned}$$

"Para cada $x \in \mathcal{X}$ considero x , dist. posteriori e encontromos o valor de $\hat{\theta}$ que minimiza os riscos a posteriori"

Principais funções de perda:

$$(1) \text{ Perda Quadrática: } L(d, \theta) = (d - \theta)^2$$

$$(2) \text{ Perda Absoluta: } L(d, \theta) = |d - \theta|$$

$$(3) \text{ Perda 0-1}$$

$$(1) L(\hat{d}, \theta) = (\hat{d} - \theta)^2$$

Para cada $x \in \mathcal{X}$, devemos obter \hat{d}^* tq.

$$E[L(\hat{d}^*, \theta) | X=x] = \min_{\hat{d} \in \Theta} E[L(\hat{d}, \theta) | X=x]$$

No caso de perda quadrática, temos:

$$\begin{aligned} E[L(\hat{d}, \theta) | X=x] &= E[(\hat{d} - \theta)^2 | X=x] = E[(\hat{d} - E(\theta | X=x))^2 + E(\theta | X=x) - \\ &\quad \theta)^2 | X=x] = E[(\hat{d} - E(\theta | X=x))^2 | X=x] + 2E[(\hat{d} - E(\theta | X=x))(E(\theta | X=x) - \theta)] | X=x \\ &\quad + E[(\theta - E(\theta | X=x))^2 | X=x] \\ &= (\hat{d} - E(\theta | X=x))^2 + \text{Var}(\theta | X=x), \text{ que é minimizado por } E(\theta | X=x). \end{aligned}$$

Logo, $\hat{d}^* = E(\theta | X=x)$.

Assim, se $E(\theta | X=x) \in \Theta$, tomamos $\hat{\delta}^*(x) = E(\theta | X=x)$.

Caso contrário, tomamos $\hat{d}^* \in \Theta$ tal que:

$$|\hat{d}^* - E(\theta | X=x)| = \min_{\hat{d} \in \Theta} |\hat{d} - E(\theta | X=x)|$$

No exemplo da urna,

$$\hat{\delta}^*, \{0,1\}^2 \rightarrow \{0,1,2,3,4,5\}$$

Vimos que,

$$P(\Theta=j | X=(0,0))$$

11

j	$P(\Theta=j X=(0,0))$	j	$P(\Theta=j X=(1,0))$
0	$\frac{29}{40}$	0	0
1	$\frac{12}{40}$	1	$\frac{4}{20}$
2	$\frac{6}{40}$	2	$\frac{6}{20}$
3	$\frac{2}{40}$	3	$\frac{4}{20}$
4	0	4	$\frac{4}{20}$
5	0	5	0
			1

$$(*) E[\Theta | X=(0,0)] = \frac{39}{40} = \frac{3}{4} \quad S^*(0,0) = 1 \quad (***) E[\Theta | X=(1,0)] = 2.5 \Rightarrow S^*((1,0)) = 2$$

$$j \quad P(\Theta=j | X=(1,1))***$$

0	0
1	0
2	$\frac{4}{40}$
3	$\frac{6}{40}$
4	$\frac{12}{40}$
5	$\frac{20}{40}$

$$(***) E[\Theta | X=(1,1)] = \frac{17}{4} \Rightarrow S^*((1,1)) = 4 \text{ (Estimador de Bayes)}$$

Logo, um estimador de Bayes para Θ é com relação a pergunta qualitativa,

$$\hat{\theta}^*(x) = \begin{cases} 1, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } x=(1,0) \\ 4, & x=(1,1) \end{cases}$$

$$\hat{\theta}_2^*(x) = \begin{cases} 1, & x=(0,0) \\ 2, & x=(0,1) \\ 3, & x=(1,0) \\ 4, & x=(1,1) \end{cases}$$

Existem \oplus 2 estimadores de Bayes
 $\hat{\theta}^*(x) \in \hat{\theta}_2^*(x)$

$$\text{Ex 2: } \Theta = (0,1) \quad \mathcal{X} = \{0,1\}^n$$

X_1, X_2, \dots, X_n , dado Θ , i.i.d $\text{Ber}(\Theta)$

$\Theta \sim \text{Ber}(a, b)$, $a, b > 0$.

(?)

Vemos que para cada $x = (x_1, \dots, x_n) \in \mathcal{X}$,

$$\Theta | X=x \sim \text{Beta}\left(a + \sum_{i=1}^n x_i, b + n - \sum_{i=1}^n x_i\right)$$

$$E(\Theta | X=x) = \frac{a + \sum_{i=1}^n x_i}{a+b+n} \in \Theta, \forall x \in \mathcal{X}$$

é binomial, neste caso

Assim, o estimador de Bayes com relação à perda quadrática é dado por

$$\hat{\theta}^*(X_1, \dots, X_n) = \frac{a + \sum_{i=1}^n x_i}{a+b+n} \quad \left(\frac{a+n\bar{x}}{a+b+n} = \frac{a+b}{a+b+n} \cdot \frac{a}{a+b} + \frac{n}{a+b+n} \bar{x}_n \right)$$

comendo

$$\text{Ex 3: } X_1, X_2, \dots, X_n \text{, dado } \Theta, \text{ i.i.d. } N(\Theta, \sigma_0^2)$$

$$\Theta \sim N(a_0, b_0^2), \quad \mathcal{X} = \mathbb{R}^n, \quad \Theta = \mathbb{R}$$

$$\Theta | X_1 = x_1, \dots, X_n = x_n \sim N\left(\frac{b_0^2 \sum_{i=1}^n x_i + a_0^2}{n b_0^2 + a_0^2}, \frac{1}{n b_0^2 + a_0^2}\right)$$

$$E(\Theta | X_1 = x_1, \dots, X_n = x_n) = \frac{b_0^2 \sum_{i=1}^n x_i + a_0^2}{n b_0^2 + a_0^2}.$$

Logo, o estimador de Bayes para Θ com relação à perda quadrática é:

$$\hat{\theta}^*(X_1, \dots, X_n) = \frac{b_0^2 \sum_{i=1}^n x_i + a_0^2}{n b_0^2 + a_0^2}$$

(2) Perda Absoluta: $L(d, \theta) = |d - \theta|$

Dizemos que m é uma Mediana da var X (da distribuição da var X) se

$$P(X \leq m) \geq \frac{1}{2} \quad \in \quad P(X \geq m) \geq \frac{1}{2}$$

Para $x \in \mathcal{X}$, seja m uma mediana da distribuição condicional de Θ dado $X=x$.

Vamos avaliar para $d > m$,

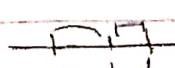
$$E[L(d, \theta) | X=x] - E[L(m, \theta) | X=x]$$

$$= E[|d - \theta| | X=x] - E|m - \theta| | X=x] = (\text{Suposição } E(|\theta|) < \infty)$$

$$= E[|d - \theta| - |m - \theta| | X=x] = \sum_{\theta \in \Theta} (|d - \theta| - |m - \theta|) \cdot P(\theta=j | X=x).$$

$$\overbrace{\hspace{10em}}^{\substack{m \\ d}} = \sum_{j \leq m} (d - j) P(\theta=j | X=x) + \sum_{m < j \leq d} [(d - j) - (j - m)] P(\theta=j | X=x)$$

$$+ \sum_{j > d} (-1)(d - m) P(\theta=j | X=x)$$



$$\geq \sum_{j \leq m} (d - m) P(\theta=j | X=x) + \sum_{m < j \leq d} (-1)(d - m) P(\theta=j | X=x) - (d - m) \sum_{j > d} P(\theta=j | X=x)$$

$$= (d - m) \left\{ P(\theta \leq m | X=x) - P(m < \theta \leq d | X=x) - P(\theta > d | X=x) \right\}$$

$$= (d - m) \left\{ P(\theta \leq m | X=x) - P(\theta \geq m | X=x) \right\}$$

$$\geq \underbrace{(d - m)}_{\geq \frac{1}{2}} \left\{ 2P(\theta \leq m | X=x) - 1 \right\} \geq 0$$

Assim, para $d > m$.

$$E[L(d, \theta) | X=x] \geq E[L(m, \theta) | X=x].$$

Analogamente para $d \leq m$!!!

Assim, tomemos, em geral, $\begin{cases} \hat{\theta}(X) \leftarrow \text{mediana } (\theta|X) \\ \leftarrow \text{Médiana } (\theta|X-x) \end{cases}$

No exemplo da urna,

j	$P(\theta=j X=(0,0))$
0	$2/40$
1	$12/40$
2	$6/40$
3	$24/40$
4	0
5	0

$$\text{Mediana } (\theta | X=(0,0)) = 0 \text{ (ou 1)}$$

$$\text{Mediana } (\theta | X=(1,1)) = 4 \text{ (ou 5)}.$$

$$\text{Mediana } (\theta | X=(5,0)) = 2 \text{ (ou 3)}.$$

j	$P(\theta=j X=(5,0))$
0	0
1	$7/20$
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:

$$S^*(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } (1,0) \\ 4, & x=(1,1) \end{cases}$$

No exemplo 2,

$$\Theta | X_1=x_1, \dots, X_n=x_n \sim N\left(\frac{b_0^2 \sum_{i=1}^n x_i + \sigma_0^2 a_0}{b_0^2 n + \sigma_0^2}, \frac{\sigma_0^2}{b_0^2 n + \sigma_0^2}\right)$$

$$\text{Mediana } (\Theta | X_1=x_1, \dots, X_n=x_n) = \frac{b_0^2 \sum_{i=1}^n x_i + \sigma_0^2 a_0}{b_0^2 n + \sigma_0^2}.$$

O estimador de Bayes com relação à perda absoluta é

$$S^*(x) = \text{Mediana } (\Theta | x) = \frac{b_0^2 \sum_{i=1}^n x_i + \sigma_0^2 a_0}{n b_0^2 + \sigma_0^2}.$$

(3) Função de perda 0-1

$$L(d, \theta) = \begin{cases} 0, & d=\theta \\ 1, & d \neq \theta \end{cases}$$

Para $x \in \mathcal{X}$,

$$\begin{aligned} E[L(d, \theta) | X=x] &= \sum_{j \in \Theta} L(d, j) P(\theta=j | X=x) \\ &= \sum_{j \neq d} 1 \cdot P(\theta=j | X=x) + 0 \cdot P(\theta=d | X=x) \\ &= P(\theta \neq d | X=x) = 1 - P(\theta=d | X=x) \end{aligned}$$

Assim, devemos tomar uma mede da dist a posterior de Θ dada $X=x$ como estimativa

X discreto

Lembrete Dizemos que m é a mede da dist de X se $P(X=m) \geq P(X=x), \forall x$

No exemplo da urna, um estandar de Bayes
Lembrete 4:

$$\delta^*(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } x=(1,0) \text{ (ou 3)} \\ 5, & x=(1,1) \end{cases}$$

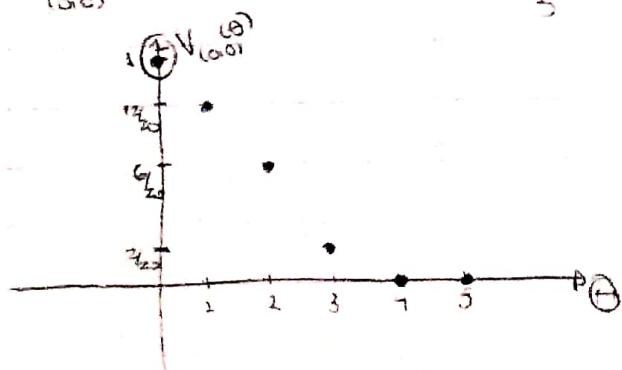
ESTIMADORES DE MÁXIMA VEROSIMILHANÇA

As estimadoras $\hat{\delta}_{MV}: \mathcal{X} \rightarrow \Theta$ que associa a cada $x \in \mathcal{X}$,

$\hat{\delta}_{MV}(x) = \arg \max_{\theta \in \Theta} V_x(\theta)$ devemos o nome de estimador de máxima verosimilhança para Θ .

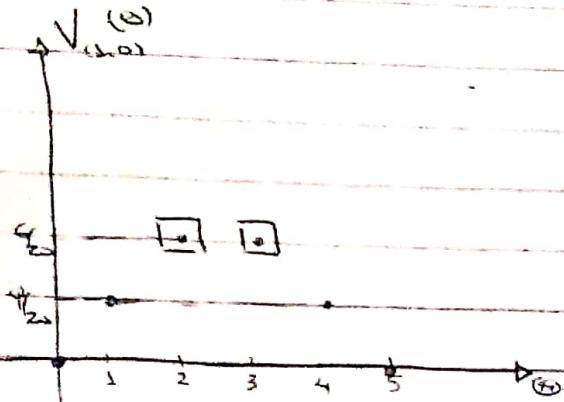
(1) URNA:

$$V_{(0,0)}(\theta) = P(X=(0,0)|\theta) = \frac{5-\theta}{5} \cdot \frac{4-\theta}{4}$$

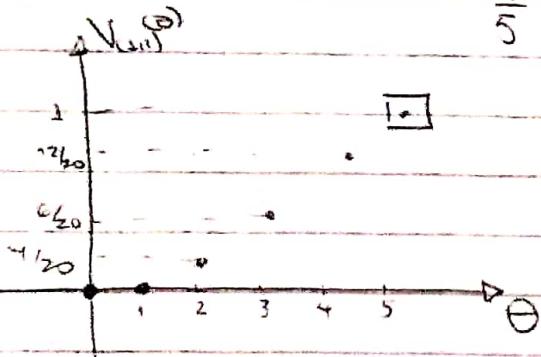


$$\hat{\delta}_{MV}(0,0) = 0.$$

$$V_{(1,0)}(\theta) = P(X=(1,0)|\theta) = \frac{\theta}{5} \quad \frac{(5-\theta)}{4}$$



$$V_{(1,1)}(\theta) = P(X=(1,1)|\theta) = \frac{\theta}{5} \quad \frac{(\theta-1)}{4}$$



Assim,

$$\sum_{nn}^{(1)}(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \text{ or } x=(1,0) \\ 5, & x=(1,1) \end{cases}$$

Or

$$\sum_{nn}^{(2)}(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \\ 3, & x=(1,0) \\ 5, & x=(1,1) \end{cases}$$

Ex 2: X_1, X_2, \dots, X_n i.i.d. $\text{Ber}(\theta)$

$$\mathcal{X} = \{0,1\}^n, \quad X = (X_1, X_2, \dots, X_n)$$

$$\Theta = [0,1]$$

$$x = (x_1, \dots, x_n) \in \mathcal{X},$$

$$V_x(\theta) = P(X_1=x_1, \dots, X_n=x_n | \theta) = \prod_{i=1}^n P(X_i=x_i | \theta) = \theta^{x_1} (1-\theta)^{1-x_1}$$

$$\Rightarrow V_x(\theta) = \theta^{\sum x_i} (1-\theta)^{n-\sum x_i}$$

$$x = (1, 1, \dots, 1) \Rightarrow V_x(\theta) = \theta^n = \delta_{\text{min}}(1), \quad D=1 \Rightarrow \frac{\partial}{\partial \theta} V_x(\theta) = 0$$

$$x = (0, 0, \dots, 0) \Rightarrow V_x(\theta) = (1-\theta)^n \quad \sim$$



$$\frac{\partial V_x(\theta)}{\partial \theta} = (\sum x_i) \theta^{\sum x_i - 1} (1-\theta)^{n-\sum x_i} + \sum x_i (n - \sum x_i) (1-\theta)^{n-\sum x_i - 1} (-1)$$

$$\Rightarrow \frac{\partial V_x(\theta)}{\partial \theta} = \theta^{\sum x_i - 1} (1-\theta)^{n-\sum x_i} \left\{ \sum x_i (1-\theta) - (n - \sum x_i) \theta \right\}$$

$$\Rightarrow \frac{\partial V_x(\theta)}{\partial \theta} = \theta^{\sum x_i - 1} (1-\theta)^{n-\sum x_i} \left\{ \sum x_i - n\theta \right\} > 0$$

$$\Leftrightarrow \sum_{i=1}^n x_i - n\theta > 0 \Leftrightarrow \theta < \frac{\sum x_i}{n}$$

$$(\Leftarrow) \Leftrightarrow \theta > \frac{\sum x_i}{n}$$



$\log \cdot$

$$f(x) = \sum_{n=1}^{\infty} (a_n x^n - b_n) \sum_{k=1}^n x^k$$

Not 04 Not 06

hence $x \in \mathbb{R} \rightarrow$ either

Excluded

$x \in \mathbb{R} \rightarrow V_1(\cdot)$

$\delta_x \ni \mathbb{R} \rightarrow \Theta$

where $V_1(\delta_x(\omega)) = \min_{\Theta(\omega)} V_1(\omega)$

Example 3) $X = \mathbb{R}$ able and Beta(Θ)

$X = (\text{out})^*$

$\Theta = \mathbb{R}$

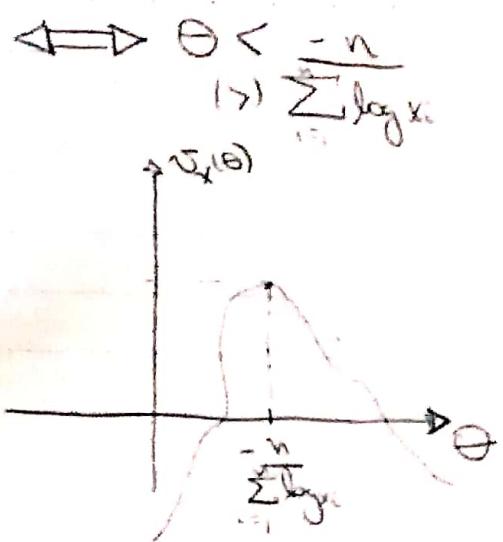
$\omega = (\omega_1, \omega_2, \dots) \in X$

$V_1(\omega) = f(\omega_1, \omega_2, \dots) = \prod_{i=1}^n f_i(\omega_i) = \prod_{i=1}^n e^{-\alpha_i \omega_i + \mu_i}$

$\mathcal{F}(\frac{d}{d\omega_i} V_1(\omega))$

$V_1(\omega) = \log V_1(\omega) = \log \left(\prod_{i=1}^n e^{-\alpha_i \omega_i + \mu_i} \right) = \log \omega_1 + \dots + \log \omega_n + \mu$

$\frac{dV_1(\omega)}{d\omega_i} = \frac{1}{\omega_i} + \sum_{j \neq i} \log \omega_j > 0 \Leftrightarrow \omega_i > -\sum_{j \neq i} \log \omega_j$



$$\text{log. } S_{nn}(x) = \frac{-n}{\sum_{i=1}^n \log x_i}$$

Ex 4 X_1, X_2, \dots, X_n dada $\Theta = (\Theta_1, \Theta_2)$, com $N(\Theta_1, \Theta_2)$

$$\mathcal{X} = \mathbb{R}^n \quad \Theta = \mathbb{R} + \mathbb{R}_+$$

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

$$\begin{aligned} V_x(\Theta) &= f(x_1, \dots, x_n | \Theta) = \prod_{i=1}^n f(x_i | \Theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\Theta_2}} e^{-\frac{(x_i - \Theta_1)^2}{2\Theta_2}} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n \Theta_2^{-n/2} e^{-\frac{\sum (x_i - \Theta_1)^2}{2\Theta_2}} \leq \left(\frac{1}{\sqrt{2\pi}}\right)^n \Theta_2^{-n/2} e^{-\frac{\sum (x_i - \bar{x})^2}{2\Theta_2}} \end{aligned}$$

$$W_x(\Theta_2) = \Theta_2^{-n/2} e^{-\frac{\sum (x_i - \bar{x})^2}{2\Theta_2}}$$

$$w_x(\Theta_2) = \log W_x(\Theta_2) = -\frac{n}{2} \log \Theta_2 - \frac{\sum (x_i - \bar{x})^2}{2\Theta_2}$$

$$\frac{dW_x(\theta)}{d\theta_2} = -n + \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{2\theta_2^2}$$

$$= \frac{1}{2\theta_2} \left\{ -n + \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\theta_2} \right\} > 0$$

$$\Leftrightarrow -n + \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\theta_2} > 0 \Leftrightarrow \theta_2 < \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

$$\text{Log. } \delta_{\text{MV}}(x) = \left(\bar{x}_m, \frac{\sum (x_i - \bar{x})^2}{n} \right)$$

(5) x_1, \dots, x_n data Θ , std cond $U(0, \theta)$

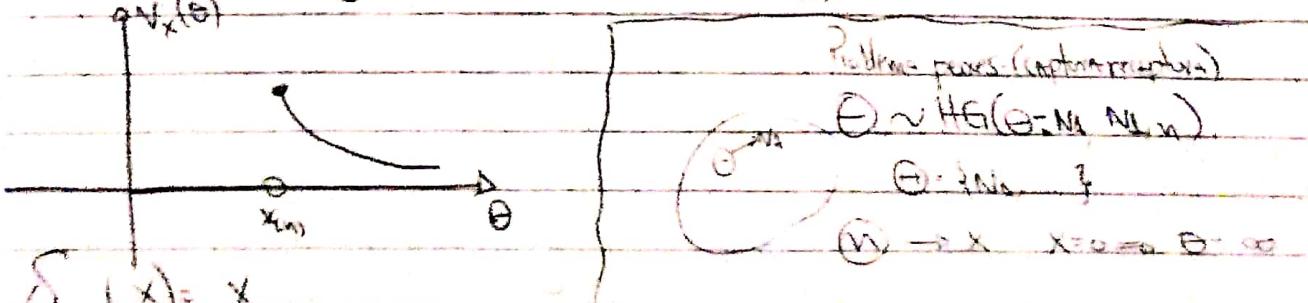
$$\Theta = \mathbb{R}_+$$

$$\mathcal{X} = \mathbb{R}^n$$

$$x = (x_1, \dots, x_n) \in \mathcal{X}$$

$$V_x(\theta) = f(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i | \theta) = \prod_{i=1}^n \frac{1}{\theta} \frac{\#(x_i)}{[0, \theta]} = \frac{1}{\theta^n} \frac{\#(x_m)}{[0, x_m]} \frac{\#(x_n)}{[0, \theta]}$$

$$V_x(\theta) = K(x) \cdot \frac{1}{\theta^n} \frac{\#(\theta)}{[x_m, \infty)}, \quad K(x) = \frac{\#(x_m)}{[0, x_m]} \frac{\#(x_n)}{[0, \infty)}$$



$$\delta_{\text{MV}}(x) = x_m$$

Posterior prior (soft thresholding)

$$\theta \sim H_G(\theta; M, N, n)$$

$$\theta \sim \text{Inv. } \Gamma$$

$$(N \rightarrow \infty, N = \infty)$$

\rightarrow Olga S. Yoshida (Mitsubishi)

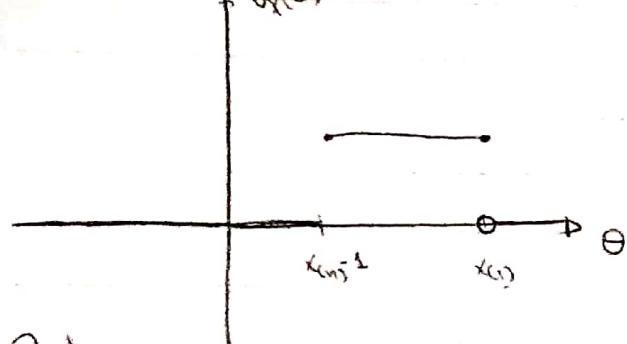
Jesse Silverman

(6) X_1, \dots, X_n dada Θ c.c.d. $U(\theta, \theta+1)$

$$\mathcal{X} = \mathbb{R}^n, \quad \Theta = \mathbb{R}$$

$$X = (X_1, \dots, X_n) \in \mathcal{X},$$

$$V_X(\theta) = \prod_{i=1}^n f(x_i, \theta) = \prod_{i=1}^n \frac{1}{\theta+1-\theta} \cdot \frac{\#(x_i)}{[\theta, \theta+1]} = \frac{\#(x_1)}{[\theta, \theta+1]} \cdots \frac{\#(x_n)}{[\theta, \theta+1]}$$



$$\begin{aligned} \theta &\xrightarrow{\text{Zero}} \theta+1 \\ \theta \leq x_{(1)} &\leq \theta+1 \\ \theta \leq x_{(n)} &\leq \theta+1 \\ \theta-1 \leq x_{(1)}-1 &\leq \theta \leq x_{(n)}-1 \end{aligned}$$

Portanto, para todo $\alpha \in (0,1)$

$$S_{MV}^{(\alpha)}(x) = \alpha X_{(1)} + (1-\alpha)(X_{(n)} - 1) \text{ é um EMV para } \theta.$$

Alguns comentários.

(1) T é uma estatística suficiente S_{MV} é único $\Rightarrow S_{MV}$ é função de T.
Quando S_{MV} não é único, o S_{MV} pode não ser função de alguma estatística suficiente.

$$S_{MV}^{(\alpha)} = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \\ 3, & x=(1,0) \\ 5, & x=(1,1) \end{cases} \quad T(x) = X_1 + X_2.$$

Por outro lado, necessário, existe S_{MV} que é função de T

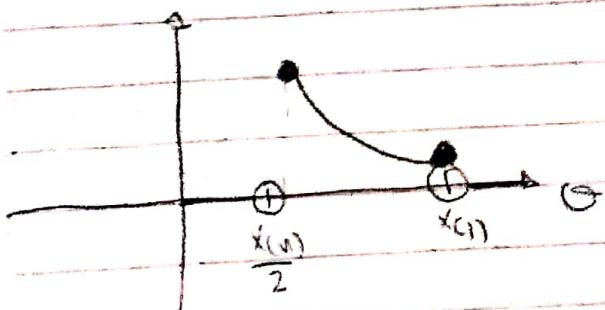
$$\sum_{MN}(x) = \begin{cases} 0, & x=(0,0) \\ 2, & x=(0,1) \text{ ou } x=(1,0) \\ 5, & x=(1,1) \end{cases}$$

(2) Quando \sum_{MN} é uma estatística suficiente e única $\Rightarrow \sum_{MN}$ é uma estatística suficiente mínima.

Ex. X_1, X_2, \dots, X_n dada Θ c. d. d. $U(\theta, 2\theta)$

$$\Theta = \mathbb{R}_+ \quad X = \mathbb{R}_+$$

$$\sum_X(\theta) = \prod_{i=1}^n \frac{1}{(2\theta - x_i)} \cdot \frac{x_i}{[\theta, 2\theta]} = \frac{1}{\theta^n} \cdot \frac{\#(x_{(1)})}{[\theta, x_{(n-1)}]} \cdot \frac{\#(x_{(n)})}{[\theta, 2\theta]}$$



$$\sum_{MN}(x) = \frac{x_{(n)}}{2}. \quad (\text{Não é uma estatística suficiente})$$

(3) Transformações dos EMN

Seja $g: \Theta \rightarrow \mathcal{G} (= g(\Theta))$, g qualquer.

δ_{MV} é o EMV para $\Theta \rightarrow g(\delta_{\text{MV}})$ é o EMV para $g(\Theta)$.

$$g: \Theta \rightarrow \mathcal{T}$$

$$\Gamma = g(\Theta)$$

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 Γ .

Nesse caso,

$$\tilde{\gamma}_{\text{MV}}: \mathcal{X} \rightarrow \mathcal{T} \text{ é tq}$$

$$L'_x(\tilde{\gamma}_{\text{MV}}(x)) = \max_{\Theta \in \mathcal{T}} L'_x(\Theta).$$

Seja $\tau^* = g(\delta_{\text{MV}})$. Para $x \in \mathcal{X}$,

$$\tilde{\gamma}^*(x) = g(\delta_{\text{MV}}(x))$$

$$L'_x(\tilde{\gamma}^*(x)) \leq \max_{\Theta \in \mathcal{T}} L'_x(\Theta) = L'_x(\tilde{\gamma}_{\text{MV}}(x)) \rightarrow L'_x(\tilde{\gamma}^*(x)) \leq L'_x(\tilde{\gamma}_{\text{MV}}(x))$$

$$L'_x(\tilde{\gamma}_{\text{MV}}(x)) = \max_{\Theta \in \mathcal{T}} L'_x(\Theta) = \max_{\Theta \in \mathcal{T}} \left\{ \max_{\Theta' \in g^{-1}(\{\Theta\})} L_x(\Theta') \right\} \leq \max_{\Theta \in \mathcal{T}} \left\{ L_x(\tilde{\gamma}_{\text{MV}}(x)) \right\}$$

$$\leq \max_{\Theta \in \mathcal{T}} \left\{ \max_{\Theta' \in g^{-1}(\{g(\delta_{\text{MV}}(x))\})} L_x(\Theta') \right\} \leq \max_{\Theta \in \mathcal{T}} L_x(g(\delta_{\text{MV}}(x)))$$

$$\Rightarrow L'_x(\tilde{\gamma}_{\text{MV}}(x)) \leq L'_x(\tilde{\gamma}^*(x))$$

Ex 1: X_1, \dots, X_n dado Θ c.c.i.d. $N(\theta, \sigma^2)$

$$g(\theta) = \text{Var}(X_i | \theta) = \frac{\sigma^2}{12}$$

Vemos que é EMV para θ é $\hat{\delta}_{\text{MV}}(x) = \bar{X}_{\text{env}}$

Pela invariância dos EMV, segue

$$\hat{g}(\theta)_{\text{MV}} = \frac{\sum x_i^2}{12} = \frac{\bar{x}_{\text{env}}^2}{12}$$

(2) X_1, \dots, X_n dado Θ c.c.i.d. $N(\theta, 1)$

$$P_{0.95} = g(\theta)$$

$$P(X \leq g(\theta) | \theta) = 95\%$$

$$P\left(\frac{X-\theta}{\sqrt{1}} \leq \frac{g(\theta)-\theta}{\sqrt{1}} | \theta\right) = 95\%$$

$$\Rightarrow P(Z \leq \frac{g(\theta)-\theta}{\sqrt{1}} | \theta) = 95\% \Rightarrow \frac{g(\theta)-\theta}{\sqrt{1}} = 1.64 \Rightarrow g(\theta) = \theta + 1.64.$$

Vemos que é EMV para θ é $\hat{\delta}_{\text{MV}}(x) = \bar{X}_n$. Lágo, pela invariância

dos EMV, segue que:

$$\hat{T}_{\text{MV}}(x) = \hat{g}(\hat{\theta})_{\text{MV}}(x) = \hat{\delta}_{\text{MV}}(x) + 1.64 = \bar{X}_n + 1.64$$

Cont. 2 $X_1, \dots, X_n | \theta \sim N(\theta_1, \theta_2)$, onde $\theta = (\theta_1, \theta_2)$

$$P(X_1 \leq g(\theta) | \theta) = 0.95 \iff P\left(\frac{X_1 - \theta_1}{\sqrt{\theta_2}} \leq \frac{g(\theta) - \theta_1}{\sqrt{\theta_2}} | \theta\right) = 0.95 \iff$$

$$\frac{g(\theta) - \theta_1}{\sqrt{\theta_2}} = 1.64 \iff g(\theta) = \theta_1 + 1.64\sqrt{\theta_2}$$

Vemos que o EMV para $\theta = (\theta_1, \theta_2)$ é

$$\delta_{MV}(x) = \left(\bar{X}_n, \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}} \right) \text{ Ligo, pela invariância dos EMV, temos.}$$

$$\hat{T}_{MV}(x) = \hat{g}(\theta) = g(\delta_{MV}(x)) = \bar{X}_n + 1.64 \sqrt{\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n}}$$

(4) X_1, \dots, X_n têm θ c.c.d. $Exp(\theta)$.

$g(\theta) = \text{Mediana}(\theta|x)$

$$P(X_1 \leq g(\theta) | \theta) = \frac{1}{2} \Rightarrow 1 - e^{-\theta g(\theta)} = \frac{1}{2} \Rightarrow e^{-\theta g(\theta)} = \frac{1}{2}$$

$$\Leftrightarrow -\theta g(\theta) = \log \frac{1}{2} \Rightarrow g(\theta) = \frac{\log 2}{\theta}$$

$$\text{Ex. } \delta_{MV} = \frac{1}{\bar{X}_n}$$

Pela invariância,

$$\hat{g}(\theta)_{MV}(x) = \frac{\log 2}{\bar{X}_n} = \frac{\bar{X}_n \cdot \log 2}{1}$$

Método dos Momentos

X_1, \dots, X_n dada $\Theta = (\theta_1, \dots, \theta_r)$ são c.i.d. com densidade $f(\cdot | \Theta)$, tais que $E(X_i^r | \Theta) < \infty$.

Em geral, para j=1, ..., r,

$$E(X_j^j | \Theta) = h_j(\theta_1, \dots, \theta_r).$$

Seja $M_j = \frac{1}{n} \sum_{i=1}^n X_i^j$ o j-ésimo momento amostral,
 $j=1, \dots, r$. Considere o sistema,

$$E(X_i^j | \Theta) = M_j, \quad j=1, \dots, r.$$

Uma solução deste sistema produz um estimador para $\Theta = (\theta_1, \dots, \theta_r)$
Chamado ESTIMADOR DE MOMENTOS

Exemplos (1) X_1, \dots, X_n dada Θ c.i.d. $Ber(\theta)$, ($r=1$)

$$E(X_i | \Theta) = \frac{X_1 + \dots + X_n}{n} \Rightarrow \Theta = \frac{X_1 + \dots + X_n}{n} \Rightarrow \hat{\Theta} = \bar{X}$$

(2) X_1, \dots, X_n dada $\Theta = (\theta_1, \theta_2)$ c.i.d. $N(\theta_1, \theta_2)$ ($r=2$)

$$\begin{cases} E[X_i | \Theta] = \frac{X_1 + \dots + X_n}{n} \\ E[X_i^2 | \Theta] = \frac{X_1^2 + \dots + X_n^2}{n} \end{cases} \Leftrightarrow \begin{cases} \theta_1 = \bar{X}_n \\ \theta_2 + \theta_1^2 = \sum_{i=1}^n \frac{X_i^2}{n} \end{cases}$$

$$\Rightarrow \begin{cases} \Theta_1 = \bar{x}_n \\ \Theta_2 = \sum \frac{x_i^2}{n} - \bar{x}_n^2 = \sum \frac{(x_i - \bar{x})^2}{n} \end{cases}$$

$$\hat{\Theta}_{\text{mam}}(x) = (\bar{x}_n, \sum \frac{(x_i - \bar{x})^2}{n})$$

(3) x_1, \dots, x_n , dado Θ_1 é iid. $U(0, \theta)$. ($n=2$)

$$E(X_{1|2}) = \frac{x_1 + x_2}{2} \Rightarrow \Theta = \frac{x_1 + x_2}{2} \Rightarrow \Theta = 2\bar{x}_n$$

$$\text{Logo, } \hat{\Theta}_{\text{mam}}(x) = 2\bar{x}_n$$

por exemplo, se $n=4$ e $(1, 2, 2, 15)$ $\hat{\Theta}_{\text{mam}}=6$??? $\left(\begin{array}{l} \hat{\Theta}_{\text{mam}}(1, 2, 2, 15)=10 \\ \hat{\Theta}_{\text{mam}}(10, 12, 13, 1)=15 \end{array} \right)$
na verdade
que não é suficiente!

Princípio da Suficiência

"A inferência deve ser baseada em x apenas através de $t(x)$ se t é uma estatística suficiente".

$(x, y \in \mathcal{X} \text{ t.q } t(x) = t(y), \text{ então as conclusões (inferências) sobre } \Theta \text{ a partir de } (y \text{ devem coincidir})$

Referências: De Groot, Casella & Berger, "the likelihood principle"
Berger e Wolpert (1988)

Prob. e Inferência Estatística I

23/05/2013 - Aula 7

Luis Gustavo Esteres

Estimador de Bayes

EMV

E. Momentos

Exercício.

X_1, \dots, X_n dado $\Theta = (\theta_1, \theta_2)$, c.i.i.d. Gamma(θ_1, θ_2) ($r=2$)

$\hat{\theta}_{\text{mom}} = ?$

Em geral,

$$g(\theta) = \mu(E(X_1|\theta), E(X_1^2|\theta), \dots, E(X_r^r|\theta))$$

$$\hat{g}(\theta)_{\text{mom}} = \mu(M_1, M_2, \dots, M_r)$$

Estimadores de Mínimos Quadrados

$X = (X_1, \dots, X_n)$ que dado θ , possui "densidade" $f(\cdot|\theta)$.

Idéia: Escolher para cada $x \in \mathbb{X}$, como estimativa o valor que minimiza a "diferença" entre valor observado e esperado.

Para cada $x \in \mathbb{X}$, tomamos $\delta_{\text{MQ}}(x)$ tal que

$$\sum_{i=1}^n (x_i - E(x_i | \delta_{\text{MQ}}(x)))^2 = \min_{\theta \in \Theta} \sum_{i=1}^n (x_i - E(x_i | \theta))^2.$$

$\underbrace{\quad}_{S(x, \delta_{\text{MQ}}(x))} \qquad \qquad \underbrace{\quad}_{S(x, \theta)}$

Exemplo 1. x_1, \dots, x_n , dado θ i.i.d. $N(\theta, s)$

$$E(x_i | \theta) = \theta. \quad \mathcal{X} = \mathbb{R}^n \quad \Theta = \mathbb{R}$$

Para $\bar{x} = (x_1, \dots, x_n) \in \mathcal{X}$,

$$S(\bar{x}, \theta) = \sum_{i=1}^n (x_i - E(x_i | \theta))^2 = \sum_{i=1}^n (x_i - \theta)^2, \text{ donde}$$

$$S(\bar{x}, \bar{\theta}) = \min_{\theta \in \Theta} S(\bar{x}, \theta). \text{ Assim, } \hat{\theta}_{\text{ML}}(x) = \bar{x}_n.$$

Exemplo 2.

x_1, \dots, x_n dado $\theta = (\theta_0, \theta_1)$, são condicionalmente independentes tais que $x_i | \theta \stackrel{d}{=} \theta_0 + \theta_1 t_i + z_i$, onde z_1, \dots, z_n são i.i.d. $N(0, 1)$.

$t_1, t_2, \dots, t_n \in \mathbb{R}$, 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 $\bar{x} = (x_1, \dots, x_n) \in \mathcal{X}$, temos

$$S(\bar{x}, \theta) = \sum_{i=1}^n (x_i - E(x_i | \theta))^2 =$$

$$= \sum_{i=1}^n (x_i - (\theta_0 + \theta_1 t_i))^2 =$$

$$\frac{\partial S(\bar{x}, \theta)}{\partial \theta_0} = 2 \sum_{i=1}^n [x_i - (\theta_0 + \theta_1 t_i)] (-1) \quad e \quad \frac{\partial S(\bar{x}, \theta)}{\partial \theta_1} = 2 \sum_{i=1}^n [x_i - (\theta_0 + \theta_1 t_i)] (-t_i)$$

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

$$\frac{\partial S(x, \theta)}{\partial \theta_1} = 0 \Leftrightarrow \sum_{i=1}^n x_i t_i - \theta_0 t_i - \theta_1 t_i^2 = 0 \Leftrightarrow \sum_{i=1}^n x_i t_i - \theta_0 \sum_{i=1}^n t_i - \theta_1 \sum_{i=1}^n t_i^2 = 0$$

Em resumo,

$$\begin{cases} \sum_{i=1}^n x_i - n\theta_0 - \theta_1 \sum_{i=1}^n t_i = 0 \\ \sum_{i=1}^n x_i t_i - \theta_0 \sum_{i=1}^n t_i - \theta_1 \sum_{i=1}^n t_i^2 = 0 \end{cases} \quad (\text{Equações Normais})$$

Exemplo 3.

x_1, \dots, x_n dado $\theta = (\theta_0, \theta_1, \theta_2)$, são condicionalmente independentes

faís que

$$x_i | \theta \stackrel{d}{=} \theta_0 + \theta_1 t_i + \theta_2 t_i^2 + z_i, \text{ onde } z_1, \dots, z_n \text{ são i.i.d. } N(0, 1)$$

e $t_1, \dots, t_n \in \mathbb{R}$, constantes conhecidas.

Nesse caso, $\mathcal{X} = \mathbb{R}^n \subset \Theta = \mathbb{R}^3$ 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

$$\frac{\partial S(x, \theta)}{\partial \theta_0} = 2 \sum_{i=1}^n (x_i - (\theta_0 + \theta_1 t_i + \theta_2 t_i^2)) (-1)$$

$$\frac{\partial S(x, \theta)}{\partial \theta_1} = 2 \sum_{i=1}^n (x_i - (\theta_0 + \theta_1 t_i + \theta_2 t_i^2)) (-t_i)$$

$$\frac{\partial S(x, \theta)}{\partial \theta_2} = 2 \sum_{i=1}^n (x_i - (\theta_0 + \theta_1 t_i + \theta_2 t_i^2)) (-t_i^2)$$

Então temos 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. Normal"

Só resolver o sistema.

Comparação (Avaliação) De Estimadores

$$\Delta = \{\delta : \mathcal{X} \rightarrow \Theta\}$$

Idéia: Escolher $\delta^* \in \Delta$ que minimiza $E[L(\delta(x), \theta)]$ em Δ

$$p(\delta^*) = \min_{\delta \in \Delta} p(\delta)$$

Se $h(\theta, x)$ é aleatório usamos a ideia da bayesiana.

Na clássica, θ é fixo!

→ Em paralelo

Modelo Bayesiano

$$(\mathcal{X} \times \Theta, \sigma, P)$$

Modelo Clássico

$$(\mathcal{X}, \sigma, \{P_x : x \in \Theta\})$$

Seja $\delta \in \Delta$.

A quantidade $\sum_{x \in \mathcal{X}} L(\delta(x), \theta) P(x=x|\theta) \left(\int_{\mathcal{X}} L(\delta(x), \theta) f(x|\theta) dx \right)$ chamamos

de Risco (frequentista) do estimador δ para estimar o parâmetro quando esse vale θ .

$$\text{NOTAÇÃO. } R(\delta, \theta) = E[L(\delta(x), \theta) | \theta] = E_{\theta}[L(\delta(x), \theta)]$$

↓
Aleatório

Exemplo: X_1, \dots, X_n dado θ são c. i.i.d. $\text{Ber}(\theta)$

$$\Theta = [0, 1] \quad \mathcal{X} = \{0, 1\}^n$$

$$L(d, \theta) = (d - \theta)^4 \quad X = (X_1, \dots, X_n)$$

$$\delta_1(x) = x_L$$

$$\delta_2(x) = \bar{x}_n \quad \delta_3(x) = \frac{a + \sum_{i=1}^n x_i}{a+b+n} \quad \delta_4(x) = \frac{\pi}{4}$$

Calculando os riscos dos estimadores...

$$R(\delta_1, \theta) = E_\theta [L(\delta_1, \theta)] = E_\theta [(\delta_1(x) - \theta)^2] = \\ = E_\theta ((x_L - \theta)^2) = \text{Var}_\theta (x_L) = \theta(1-\theta)$$

$$R(\delta_2, \theta) = E_\theta [L(\delta_2(x), \theta)] = E_\theta [(\delta_2(x) - \theta)^2] = \\ = E_\theta ((\bar{x}_n - \theta)^2) = \text{Var}_\theta (\bar{x}_n) = \frac{\theta(1-\theta)}{n}$$

$$R(\delta_3, \theta) = E_\theta [L(\delta_3(x), \theta)] = E_\theta [(\delta_3(x) - \theta)^2] = \\ = E_\theta \left[\left(\frac{a + \sum_{i=1}^n x_i}{a+b+n} - \theta \right)^2 \right] = E_\theta \left[\left(\frac{a+n\bar{x}_n}{a+b+n} - \theta \right)^2 \right] =$$

$$= \text{Var}_\theta \left(\frac{a+n\bar{x}_n}{a+b+n} - \theta \right) + \left\{ E_\theta \left(\frac{a+n\bar{x}_n}{a+b+n} - \theta \right) \right\}^2 =$$

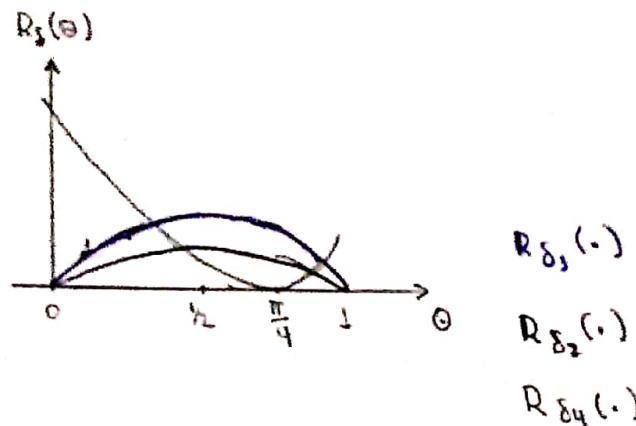
$$= \frac{n^2 \text{Var}_\theta (\bar{x}_n)}{(a+b+n)^2} + \left\{ \frac{a+n\theta}{a+b+n} - \theta \right\}^2 = \frac{n\theta(1-\theta)}{(a+b+n)^2} + \left(\frac{a+n\theta}{a+b+n} - \theta \right)^2$$

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 $R_\delta : \Theta \rightarrow \mathbb{R}_+$ (\mathbb{R})
 $\Leftrightarrow \theta \mapsto R_\delta(\theta) = R(\delta, \theta)$

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 .

$$\text{EQM}_\theta(\delta) = E_\theta[(\delta(x) - \theta)^2]$$

↓

$$\text{EQM}(\delta, \theta)$$

$$\text{EQM}(\delta | \theta)$$

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_2 é melhor que δ_1 mas não é melhor que δ_3 , dizemos que δ_2 "domina" δ_1 . Nesse caso, dizemos que δ_2 é inadmissível.

Seja $\Delta^* \subseteq \Delta$

Dizemos que $\delta^* \in \Delta$ é melhor em Δ^* se $\delta^* \in \Delta^*$ é melhor que δ , $\forall \delta \in \Delta^*$.

$E_{\theta}(\delta(x))$ é finito para todos os $\delta(x)$ medivel e em outro caso

$$EQM_{\theta}(\delta) = E_{\theta}[(\delta(x) - \theta)^2] = E_{\theta}[(\delta(x) - E_{\theta}(\delta(x))) + E_{\theta}(\delta(x)) - \theta]^2 =$$

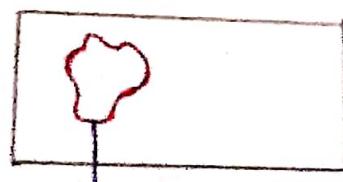
$$= E_{\theta}[(\delta(x) - E_{\theta}(\delta(x)))^2] + E_{\theta}[(E_{\theta}(\delta(x)) - \theta)^2] + 2E_{\theta}[(\delta(x) - E_{\theta}(\delta(x)))(E_{\theta}(\delta(x)) - \theta)] = \\ = Var_{\theta}(\delta(x)) + (E_{\theta}(\delta(x)) - \theta)^2$$

$$B_{\theta}(\delta) = E_{\theta}(\delta(x)) - \theta \quad \text{vício do estimador de } \delta.$$

Assim,

$$EQM_{\theta}(\delta) = Var_{\theta}(\delta) + B_{\theta}^2(\delta)$$

Vamos falar nos δ 's tais que $B_{\theta}(\delta) = 0$.



$$\Delta' = \{\delta \in \Delta : B_{\theta}(\delta) = 0, \forall \theta \in \Theta\}$$

Vamos restringir a busca à classe

$$\Delta' = \{\delta \in \Delta : E_{\theta}[\delta(x)] = 0, \forall \theta \in \Theta\}$$

Definição: Um estimador $\delta: \mathcal{X} \rightarrow \Theta$ para $\theta \in \Theta$ é dito NÃO-VIESADO para θ se

$$E_{\theta}[\delta(x)] = \theta, \quad \forall \theta \in \Theta$$

$y(\theta)$ $y(\theta)$

Exemplos:

(1) X_1, \dots, X_n , c.i.i.d. $\text{Ber}(\theta)$

$$E_{\theta}(\delta_1(x)) = E_{\theta}(X_1) = \theta, \quad \forall \theta \in \Theta. \text{ Logo, } \delta_1 \text{ é NÃO-VIESADO para } \theta.$$

$$E_{\theta}(\delta_2(x)) = E_{\theta}(\bar{X}_n) = \theta, \quad \forall \theta \in \Theta. \text{ Logo, } \delta_2 \text{ é NÃO-VIESADO para } \theta.$$

$$E_{\theta}(\delta_3(x)) = E_{\theta}\left(\frac{a+2X_1}{a+b+n}\right) = \frac{a}{a+b+n} + \frac{E_{\theta}(2X_1)}{a+b+n} = \frac{a+n\theta}{a+b+n} \neq \theta. \text{ Logo, } \delta_3 \text{ é viesado para } \theta. \text{ E,}$$

$$B_{\theta}(\delta_3) = E_{\theta}(\delta_3(x)) - \theta =$$

$$= \frac{a+n\theta}{a+b+n} - \theta = \frac{a+n\theta - a\theta - b\theta - n\theta}{a+b+n} \Rightarrow$$

$$B_{\theta}(\delta_3) = \frac{a - a\theta - b\theta}{a+b+n}.$$

No mesmo exemplo, seja $g(\theta) = \theta^2$.

$$\delta_0(x) = X_1 \cdot X_2$$

$$E_{\theta}[\delta_0(x)] = E_{\theta}(X_1 \cdot X_2) = E_{\theta}(X_1) E_{\theta}(X_2) \Rightarrow$$

$$E_{\theta}(\delta_0(x)) = \theta^2, \quad \forall \theta \in \Theta$$

Logo $\delta_0(x) = X_1 \cdot X_2$ é não-viesado para θ^2 .

Exemplo 2. X_1, \dots, X_n , dado Θ , c.i.i.d. $U(0, \theta)$

$$E_\theta(X_1) = \frac{\theta}{2}$$

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-VIÉSADO para θ .

$$\delta_2(x) = X_{(n)}$$

$$f_{X(n)|\theta} = \frac{n t^{n-1}}{\theta^n} I_{(0,\theta)}(t)$$

$$E_\theta(\delta_2(x)) = E_\theta(X_{(n)}) = \int_0^\theta t \cdot \frac{n t^{n-1}}{\theta^n} dt = \frac{n}{\theta^n} \frac{t^{n+1}}{n+1} \Big|_0^\theta = \frac{n \theta^{n+1}}{(n+1)\theta^n} = \frac{n}{n+1} \theta, \text{ com}$$

$$B_\theta(\delta_2) = E_\theta(\delta_2(x)) - \theta = \frac{n}{n+1} \theta - \theta = -\frac{\theta}{n+1}. \text{ Logo, } \delta_2 \text{ é viésado para } \theta.$$

Definindo $\delta_3(x) = \frac{1}{n} \sum_{i=1}^n X_{(n)}$, temos

$$E_\theta(\delta_3(x)) = E_\theta\left(\frac{1}{n} \sum_{i=1}^n X_{(n)}\right) = \frac{1}{n} \sum_{i=1}^n \frac{n}{n+1} \cdot \theta = \frac{n}{n+1} \theta, \forall \theta \in \Theta$$

Exemplo 3. X_1, \dots, X_n dado Θ , c.i.i.d. $\text{Exp}(\theta)$.

$$g(\theta) = e^{-\theta}$$

$$\delta_4(x) = \prod_{i=1}^n (x_i)$$

$$E_\theta(\delta_4(x)) = 1 \cdot P(\delta_4(x)=1|\theta) + 0 \cdot P(\delta_4(x)=0|\theta) =$$

$$P(1, \neq 1|\theta) = e^{-\theta+1} = e^{-\theta}, \forall \theta \in \Theta.$$

Logo, $\delta_4(x)$ é não-viésado 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

$$\mathcal{U} = \{U : \mathcal{X} \rightarrow \mathbb{R}, E_\theta(U(X)) = 0, \forall \theta \in \Theta\}$$

$$\Delta'' = \Delta'.$$

$$\begin{aligned} \delta \in \Delta'' \Rightarrow \delta &= \delta_0 + U \Rightarrow E_\theta(\delta_0 + U) = E_\theta(\delta_0) + E_\theta(U) = \\ &= g(\theta) + 0, \forall \theta \in \Theta \end{aligned}$$

Logo, $\delta \in \Delta'$, $\Delta'' \subseteq \Delta'$.

$$\delta \in \Delta' \Rightarrow \delta = \delta + \delta_0 - \delta_0 = \delta_0 + \underbrace{\delta - \delta_0}_{U} \Rightarrow \delta = \delta_0 + U, U \in \mathcal{U} \Rightarrow \delta \in \Delta''.$$

$U, \text{ onde } E_\theta(U) = 0, \forall \theta \in \Theta$

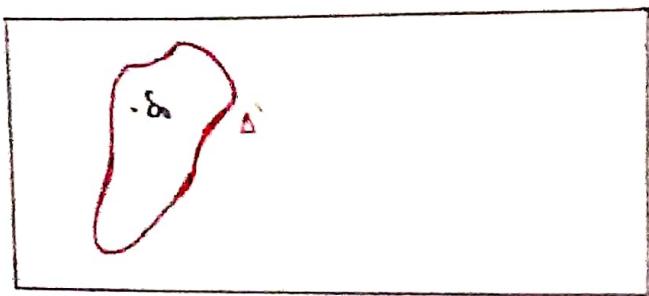
No nosso caso, vamos restringir a busca de um estimador ótimo a Δ' .

Isso é equivalente a buscar um estimador não-viesado δ^* tal que

$$\text{Var}_\theta(\delta^*) \leq \text{Var}_\theta(\delta), \forall \theta \in \Theta, \forall \delta \in \Delta'.$$

Definição: Dizemos que um estimador $\delta^* \in \Delta'$ é não-viesado de variância uniformemente mínima (ENVUM) se $\forall \delta \in \Delta'$

$$\text{Var}_\theta(\delta^*) \leq \text{Var}_\theta(\delta), \forall \theta \in \Theta.$$

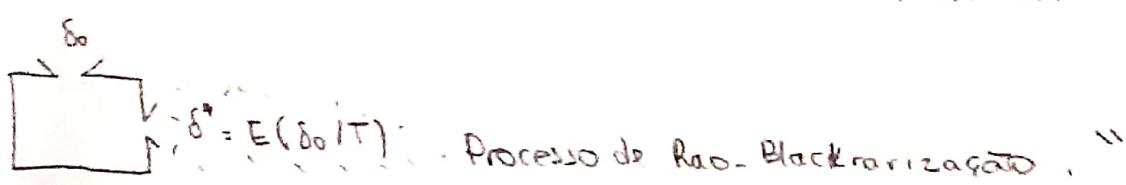


"Conseguimos aprimorar um estimador qualquer $\delta_0 \in \Delta$ por meio da operação

$$E(\delta_0 | T) = g(T)$$

depende de T . Como T é suficiente, segue que $E(\delta_0 | T)$ não depende de θ .

E, tem valor menor que δ_0



Teorema de Rao-Blackwell.

Seja $\delta_0: \mathcal{X} \rightarrow \Theta$ um estimador não-viesado para θ ($E(\delta_0) = \theta$). 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) $\text{Var}_\theta(\delta^*) \leq \text{Var}_\theta(\delta_0), \forall \theta \in \Theta$

$$E_\theta(\delta^*(X)) = E_\theta(E(\delta_\theta(X)|T)) = E_\theta(\delta_\theta(X)) = \Theta(g(\theta)), \forall \theta \in \Theta (g \in \Theta).$$

Logo, $\delta^* \in \Delta'$.

$$\begin{aligned} \text{Var}(\delta^*(X)) &= \text{Var}_\theta(E_\theta(\delta_\theta(X)|T)) = \\ &= \text{Var}_\theta(\delta_\theta(X)) - E_\theta(\text{Var}_\theta(\delta_\theta(X)|T)) \leq \\ &\leq \text{Var}_\theta(\delta_\theta(X)), \forall \theta \in \Theta. \end{aligned}$$

Exemplo 1. X_1, \dots, X_n dados θ e.i.d. $\text{Ber}(\theta)$.

$$\text{Seja } \delta_\theta(X) = X_1$$

δ_θ é não-viesado pois

$$E_\theta(\delta(X)) = E_\theta(X_1) = \theta, \forall \theta \in \Theta$$

$$\text{Seja } T = \sum_{i=1}^n X_i \text{ (suficiente para } \Theta)$$

Para $t \in \{0, 1, \dots, n\}$, avaliamos

$$E_\theta(Y_1 | \sum_{i=1}^n X_i = t) = 1 \cdot P_\theta(X_1=1 | \sum_{i=1}^n X_i=t) + 0 \cdot P_\theta(X_1=0 | \sum_{i=1}^n X_i=t) =$$

$$= \frac{P_\theta(Y_1=1 | \sum_{i=1}^n X_i=t)}{P_\theta(\sum_{i=1}^n X_i=t)} =$$

$$= \begin{cases} 0, & t=0 \\ \frac{P_\theta(X_1=1, X_2+\dots+X_n=t-1)}{P_\theta(\sum_{i=1}^n X_i=t)} & = \frac{P_\theta(X_1=1)P_\theta(X_2+\dots+X_n=t-1)}{P_\theta(Y_1+\dots+Y_n=t)} = \frac{\theta \binom{n-1}{t-1} \theta^{t-1} (1-\theta)^{n-t}}{\theta^t (1-\theta)^{n-t}} = \\ & = \frac{\binom{n-1}{t-1}}{\binom{n}{t}} = \frac{t}{n} \end{cases}$$

Logo, $\forall t \in \{0, 1, \dots, n\}$.

$$E(X_L | \sum_{i=1}^n X_i = t) = \frac{t}{n}$$

$$E(X_L | \sum_{i=1}^n X_i) = \frac{\sum_{i=1}^n X_i}{n} = \delta^*(X)$$

Exercício: Verificar tbm pl $\delta' = \frac{X_L + X_R}{2}$.

Exemplo 2.

X_1, \dots, X_n dado Θ , c.i.i.d. Poisson (θ)

$$g(\theta) = \theta e^{-\theta} = \frac{e^{-\theta} \theta^1}{1!} = P(X_L=1 | \theta)$$

$$\delta_0(X) = \prod_{i=1}^n (X_i)$$

$$E_\theta(\delta_0(X)) = 1 \cdot P_\theta(X_L=1) + 0 \cdot P_\theta(X_L=0) = \frac{e^{-\theta} \cdot \theta^1}{1!} = \theta e^{-\theta} = g(\theta), \forall \theta \in \Theta.$$

Ass. Blackwellizando δ_0 , temos

$$\delta^*(X) = E(\delta_0(X) | \sum_{i=1}^n X_i)$$

Para $t \in \mathbb{N}$,

$$E[\delta_0(X) | \sum_{i=1}^n X_i = t] = E_\theta[\prod_{i=1}^n (X_i) | \sum_{i=1}^n X_i = t] = P_\theta(X_L=1 | \sum_{i=1}^n X_i = t) =$$

$$= \frac{P_\theta(X_L=1, \sum_{i=1}^n X_i = t)}{P_\theta(\sum_{i=1}^n X_i = t)} = \begin{cases} 0, & t=0 \\ ?, & t \in \mathbb{N}^* \end{cases}$$

$$\hat{\delta} = \frac{P_0(X_1=1, X_2+\dots+X_n=t-1)}{P_0(X_1+\dots+X_n=t)} \cdot \frac{P_0(Y_{t+1}) P_0(Y_{t+2}+\dots+Y_m=t-1)}{P_0(Y_{t+1}+\dots+Y_m=t)}$$

$$= \frac{\theta e^{-\theta} \cdot \frac{e^{-(n-1)\theta}}{(t-1)!}}{\frac{e^{-n\theta} (n\theta)^t}{t!}} = \frac{(n-1)^{t-1}}{n^t} \cdot t$$

$\forall t \in \mathbb{N}$,

$$E(\delta_0(x) / \sum_{i=1}^n x_i = t) = \frac{(n-1)^{t-1}}{n^t} \cdot t$$

$$\delta^*(x) = E(\delta_0(x) / \sum_{i=1}^n x_i) = \frac{(n-1)^{n-1}}{n^n} \cdot \sum_{i=1}^n x_i$$

Caso:

$$n=2$$

$$X_1=1, X_2=0 \quad \delta_0(1,0)=1,$$

$$\delta^*(1,0) = \frac{1}{2}.$$

Esse estimador ainda não é muito bom, pois $\delta^*(1,0) = 1/2 \approx \text{mix}(g(\theta)) = \frac{1}{e}$

$$\text{p.e. } g(\Theta) = (0, e^{-1})$$

No nosso definição, não é um estimador. Alguns livros o referenciam como estimador não-viesado ou estatística não-viesada.

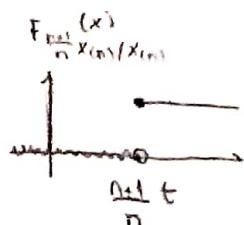
Exemplo 3. X_1, \dots, X_n dado θ , c.i.i.d., $U(0, \theta)$

$\delta_1(x) = 2X_1$ é não-viesado para θ .

$\delta_2(x) = \frac{n+1}{n} X_{(n)}$ é não-viesado para θ .

$\delta_3(x) = 2\bar{X}_n$

$$\text{Obs: } \delta_1^*(x) = E(\delta_1(x) / X_{(n)}) = \frac{n+1}{n} X_{(n)} = \delta_2(x)$$

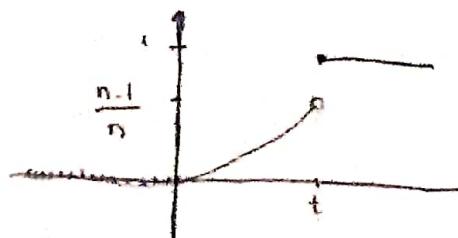


$$E(2X_1 / X_{(n)} = t)$$

$$F_{X_1 / X_{(n)} = t}(x) = \begin{cases} \frac{n-1}{n} \frac{F_{X_1}(x)}{F_{X_1}(t)}, & x < t \\ 1, & x \geq t \end{cases}$$

No nosso exemplo.

$$F_\theta(x / X_{(n)} = t) = \begin{cases} 0 & x < 0 \\ \frac{n-1}{n} \frac{x/\theta}{t/\theta}, & 0 \leq x \leq t \\ 1 & x > t \end{cases}$$



$t > 0$

$$\mathbb{E}(2X_L | X_{(n)} = t) = 2\mathbb{E}(X_L | X_{(n)} = t) =$$

$$= 2 \left\{ \int_0^t x \cdot \frac{n-1}{n} \cdot \frac{1}{t} dx + t \cdot \frac{1}{n} \right\} = 2 \left\{ \frac{n-1}{n} \frac{t^2}{2} + \frac{t}{n} \right\} =$$

$$= 2 \left(\frac{(n-1)t + 2t}{2n} \right) \Rightarrow$$

$$\Rightarrow \mathbb{E}(2X_L | X_{(n)} = t) = \frac{(n+1)}{n} t \quad . \text{ Logo } S_1 = \mathbb{E}(2X_L | X_{(n)}) = \frac{n+1}{n} X_{(n)} = S_2(X)$$

Exemplo 4.

6 brancas
5-6 verdes

$$X_1 = \begin{cases} 1, & \text{1-ésima bola extraída é branca} \\ 0, & \text{c. c.} \end{cases}$$

$X = (x_1, X_2)$

$$X_L | \Theta \sim \text{Ber}\left(\frac{\theta}{5}\right)$$

$$\mathbb{E}(X_L | \Theta) = \frac{\theta}{5} \quad S_0(X_1, X_2) = 5X_1$$

Exercício.

(1) Verifique que $T(X) = X_1 + X_2$ é suficiente

$$(2) \mathbb{E}[S_0 | X_1 + X_2] \stackrel{?}{=} \frac{5}{2} (X_1 + X_2)$$

Exemplo. X_1, \dots, X_n , dado Θ , c. i. i. d. $\mathcal{U}(\theta, 2\theta)$

$$\Theta = \mathbb{R}_+ \quad \mathfrak{X} = \mathbb{R}_+^n$$

Verificar que $T(X) = (X_{(1)}, X_{(n)})$ é suficiente para Θ .

$$U_i \stackrel{d}{=} \frac{X_{(i)} - \theta}{\theta} \quad | \quad \theta \sim \text{U}(0, L)$$

$$E_\theta\left(\frac{X_{(n)} - \theta}{\theta}\right) = \frac{n}{n+2} \quad E_\theta\left(\frac{X_{(1)} - \theta}{\theta}\right) = \frac{1}{n+1}$$

$U_{(n)} \sim \text{Beta}(1, n)$

$U_{(1)} \sim \text{Beta}(n, 1)$

$$\text{Var}_\theta\left(\frac{X_{(n)} - \theta}{\theta}\right) = \frac{n}{(n+1)^2(n+2)} = \text{Var}_\theta\left(\frac{X_{(1)} - \theta}{\theta}\right)$$

$$E_\theta(X_{(n)}) = \theta + \theta \frac{n}{n+1} = \frac{2n+1}{n+1} \theta$$

$$E_\theta(X_{(1)}) = \theta + \theta \frac{1}{n+1} = \frac{n+2}{n+1} \theta$$

$$\text{Var}_\theta(X_{(n)}) = \frac{n\theta^2}{(n+1)^2(n+2)} = \text{Var}_\theta(X_{(1)})$$

Agora, podemos propor

$$\delta_1(x) = \frac{n+1}{2n+1} X_{(n)} \quad \text{é não-viesado}$$

$$\delta_2(x) = \frac{n+1}{n+2} X_{(1)} \quad \text{é não-viesado}$$

Lembrete:

$$E(X_{(1)} / \theta) = \frac{3\theta}{2}$$

$$\delta_3(x) = \frac{2}{3} X_{(1)} \quad \text{é não-viesado}$$

Agora

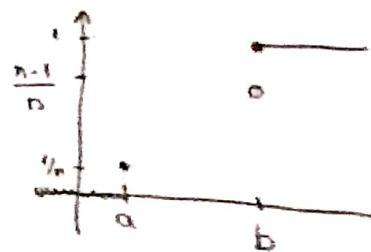
$$\delta_1^* = E[\delta_1 / X_{(1)}, X_{(n)}] = \delta_1(x)$$

$$\delta_2^* = E[\delta_2 / X_{(1)}, X_{(n)}] = \delta_2(x)$$

$$\delta_3^* = E[\delta_3 / X_{(1)}, X_{(n)}] = \frac{X_{(1)} + X_{(n)}}{3}$$

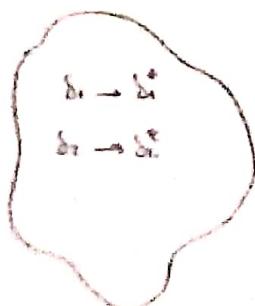
Ideia.

$$F_{\delta_3^*}(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{n}, & x = a \\ \frac{1}{n} + \frac{n-2}{n} \frac{F(x) - F(a)}{F(b) - F(a)}, & a \leq x \leq b \\ 1, & x \geq b \end{cases}$$



$$\delta_4^*(x) = \frac{n+1}{5n+4} (2X_{(n)} + X_{(1)})$$

$$\text{Var}_0 \delta_4^* \leq \text{Var}_0 \delta_3^* \leq \text{Var}_0 \delta_2^* \leq \text{Var}_0 \delta_1^*$$



Melhoramos os estimadores, mas não chegamos a um único δ^* .

(família de distribuição de \mathcal{D}_T)

Dizemos que a estatística $T: \mathcal{X} \rightarrow \mathbb{R}^k$ é completa se $\forall f: \mathbb{R}^k \rightarrow \mathbb{R}$,

$$E_\theta(f(T(x))) = 0, \forall \theta \in \Theta \Rightarrow f = 0.$$

$$(\mathcal{X}, \sigma, \{\mathcal{D}_\theta, \theta \in \Theta\})$$

$$\mathcal{P} = \{f_T(-10); \theta \in \Theta\}$$

Exemplo: X_1, \dots, X_n , dados Θ , c.i.i.d $Ber(\theta)$

$$\Theta = [0, 1]$$

$$T(X) = \sum_{i=1}^n X_i$$

Seja $h: \mathbb{R} \rightarrow \mathbb{R}$

$$E_\theta(h(T(X))) = 0, \forall \theta \in \Theta \Rightarrow$$

$$E_\theta(h(\sum_{i=1}^n X_i)) = 0, \forall \theta \in \Theta \Rightarrow \sum_{i=0}^n h(i) \binom{n}{i} \theta^i (1-\theta)^{n-i} = 0, \forall \theta \in (0, 1) \Rightarrow$$

$$(1-\theta)^n \sum_{i=0}^n h(i) \binom{n}{i} \left(\frac{\theta}{1-\theta}\right)^i = 0, \forall \theta \in (0, 1) \Rightarrow$$

$$\Rightarrow \sum_{i=0}^n h(i) \binom{n}{i} \left(\frac{\theta}{1-\theta}\right)^i = 0, \forall \theta \in \Theta$$

3'

$$\Rightarrow h(i) \binom{n}{i} = 0, \forall i = 0, 1, \dots, n \Rightarrow h(i) = 0, \forall i = 0, 1, \dots, n$$

$$\Rightarrow h = 0 \quad P_0 - q.s.$$

Logo $T(X) = \sum_{i=1}^n X_i$ é completa.

Exemplo 2. X_1, \dots, X_n , dados Θ , c.i.i.d $U(0, \theta)$

$$T(X) = X_{(n)}, \quad h: \mathbb{R}_+ \rightarrow \mathbb{R}_+$$

$$E_\theta(h(T(X))) = 0, \forall \theta \in \Theta \Rightarrow E_\theta(h(X_{(n)})) = 0, \forall \theta \in \Theta,$$

$$\int_{-\infty}^{\theta} \lambda(t) f(t|\theta) dt = 0, \forall \theta \in \Theta \Rightarrow$$

$$\Rightarrow \int_0^\theta \lambda(t) \cdot \frac{n}{\theta^n} t^{n-1} dt = 0, \forall \theta > 0 \Rightarrow \frac{1}{\theta^n} \int_0^\theta \lambda(t) n t^{n-1} dt = 0, \forall \theta > 0 \Rightarrow$$

$$\int_0^\theta \lambda(t) n t^{n-1} dt = 0, \forall \theta > 0 \Rightarrow$$

$\sigma(\theta)$ é integrável, e portanto, diferencável (em θ) $\Rightarrow 0 \forall \theta \in \Theta$ segue que

$$\frac{d}{dt} \int_0^\theta \lambda(t) dt = 0 \quad \forall \theta \in \Theta \Rightarrow$$

$$\Rightarrow \lambda(\theta) n \theta^{n-1} = 0, \forall \theta > 0 \Rightarrow \lambda(t) = 0, \forall t \geq 0.$$

Logo, $\lambda \equiv 0$ D.o.g.e.

Logo, T é completa!

Exemplo 3. X_1, \dots, X_n , dados d. Poisson(θ)

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

$$\lambda: \mathbb{N} \rightarrow \mathbb{R}$$

$$E_\theta(\lambda(T(x))) = 0, \forall \theta > 0 \Rightarrow E_\theta(\lambda(\sum_{i=1}^n X_i)) = 0, \forall \theta > 0 \Rightarrow$$

$$\sum_{i=0}^n \frac{\lambda(i) e^{-\theta \theta} (\theta \theta)^i}{i!} = 0, \forall \theta > 0 \Rightarrow \sum_{i=0}^n \frac{\lambda(i) i^i}{i!} = 0, \forall \theta > 0 \Rightarrow$$

$$\lambda(i) i^i = 0, \forall i \in \mathbb{N} \Rightarrow \lambda(i) = 0, \forall i \in \mathbb{N} \Rightarrow$$

$\lambda \equiv 0$.

Exemplo 4. X_1, \dots, X_n , dado θ , c.i.i.d., $U(\theta, 2\theta)$

$$E_\theta(X_{(1)}) = \frac{n+2}{n+1} \theta$$

$$T(X) = (X_{(1)}, X_{(n)})$$

$$E_\theta(X_{(n)}) = \frac{2n+1}{n+1} \theta$$

$$\delta: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$(u, v) \mapsto \delta(u, v) = \frac{n+1}{n+2} u - \frac{n+1}{2n+1} v$$

$$E_\theta(\delta(T(X))) = E_\theta(\delta(X_{(1)}, X_{(n)})) = E_\theta\left(\frac{n+1}{n+2} X_{(1)} - \frac{n+1}{2n+1} X_{(n)}\right)$$

$$= \frac{n+1}{n+2} E_\theta(X_{(1)}) - \frac{n+1}{2n+1} E_\theta(X_{(n)}) = \frac{n+1}{n+2} \frac{n+2}{n+1} \theta - \frac{n+1}{2n+1} \frac{2n+1}{n+1} \theta = 0, \forall \theta \in \Theta$$

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

• NÃO Teoria da Decisão

$T: \mathcal{E} \rightarrow \mathbb{R}^k$ é completa se

$\forall h: \mathbb{R}^k \rightarrow \mathbb{R}$

$$E_{\theta}(h(T(x))) = 0, \forall \theta \in \Theta \Rightarrow h \equiv 0 \text{ P}_{\theta}-\text{q.c. } \forall \theta \in \Theta.$$

Exemplo 5.0.

$$\Theta = \{1, 2, 3, 4, \dots\} = \mathbb{N}^*$$

$$X | \theta \sim U\{1, \dots, \theta\}$$

$$P(X=x|\theta) = \frac{1}{\theta} \underset{\{1, \dots, \theta\}}{\mathbb{I}}(x)$$

$$h: \mathbb{N} \rightarrow \mathbb{R}$$

$$E_{\theta}(h(X)) = 0, \forall \theta \in \Theta \Rightarrow$$

$$E_1(h(X)) = h(1) \underbrace{P(X=1|\theta=1)}_{1} = 0 \Rightarrow h(1) = 0$$

$$E_2(h(X)) = h(1) \frac{1}{2} + h(2) \frac{1}{2} = 0 \Rightarrow h(2) = 0$$

$$E_3(h(X)) = h(1) \frac{1}{3} + h(2) \frac{1}{3} + h(3) \frac{1}{3} = 0 \Rightarrow h(3) = 0$$

$\Rightarrow h(j) = 0, \forall j \in \mathbb{N}^*$, portanto,

$$h = 0 \text{ a.s. q.c. } \forall \theta \in \mathbb{N}^*$$

Logo, $T(x) = x$ é completa!

5b) $n_0 \in \mathbb{N}^*$

$$\Theta = \mathbb{N}^* - \{n_0\}$$

$$h(t) = \begin{cases} a, & t = n_0 \\ -a, & t = n_0 + 1 \\ 0, & \text{c.c.} \end{cases}$$

$$E_\theta(h(T)) = \begin{cases} \sum_{t=0}^{\theta} \frac{1}{\theta} h(t) = 0 & , \theta \leq n_0 - 1 \\ \sum_{t=1}^{\theta} \frac{1}{\theta} h(t) = \sum_{t=1}^{n_0-1} \frac{1}{\theta} \cdot 0 + \frac{1}{\theta} (a) + \cancel{\frac{1}{\theta} (-a)} + \frac{1}{\theta} (-a) + \sum_{t=n_0+2}^{\theta} \frac{1}{\theta} \cdot 0 = 0 \\ \sum_{t=1}^{n_0-1} \frac{1}{\theta} \cdot 0 + \frac{1}{\theta} (a) + \frac{1}{\theta} (-a) = 0 \end{cases}$$

$\exists h: \mathbb{N} \rightarrow \mathbb{R}$, h não identicamente nula tal que $E_\theta(h(x)) = 0, \forall \theta \in \Theta$.

Logo, X não é completa!

$P' = \{f(\cdot / \theta) : \theta \in \mathbb{N}^* - \{n_0\}\}$ NÃO é completa!

Em geral, sejam $\Theta \subseteq \Theta'$. $T: \mathfrak{X} \rightarrow \mathbb{R}^k$.

Seja $A_{\Theta} = \{ h: \mathbb{R}^k \rightarrow \mathbb{R} : E_{\theta}(h(T(X))) = 0, \forall \theta \in \Theta \}$.

Note que $h_0 \equiv 0 \in A_{\Theta}$. ($h_0 \in A_{\Theta}$)

E

$A_{\Theta} \supseteq A_{\Theta'}$ onde $\Theta \subseteq \Theta'$

S. Stigler - Completeness and Unbiasedness. The American Statistician.

Teorema de Lehmann-Scheffé.

Seja $T: \mathfrak{X} \rightarrow \mathbb{R}^k$ 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 $\delta_2 = \delta_2(T) \in \Delta'$

$$E_{\theta} [\delta_1(T(X)) - \delta_2(T(X))] =$$

$$= \theta - \theta = 0, \forall \theta \in \Theta$$

↓

$$\delta_1(T(X)) = \delta_2(T(X)), P_{\theta}-\text{q.c. } \forall \theta \in \Theta$$

$\delta', \delta'' \in \Delta'$

$$E(\delta' | T),$$

$$E(\delta'' | T)$$

Exemplo:

(1) X_1, \dots, X_n , dado θ , c.i.i.d. $Ber(\theta)$

$$\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n} = g\left(\sum_{i=1}^n X_i\right)$$

$$\bar{X}_n \in A'$$

Pelo Teorema, \bar{X}_n é o ENVVUM para θ .

Exemplo: X_1, \dots, X_n , dado θ , c.i.i.d. $U(0, \theta)$

$$\delta^*(x) = \frac{n+1}{n} X_{(n)} = g(X_{(n)})$$

δ^* é função de $X_{(n)}$, estatística suficiente e completa.

Logo, δ^* é ENVVUM para θ .

Exemplo 3: X_1, \dots, X_n , dado θ , c.i.i.d. Poisson (θ)

$$g(\theta) = \theta e^{-\theta}$$

$$\delta_0(x) = I_{\{X_1 = 1\}} \rightarrow \delta^*(x) = E[\delta_0(x) | \sum_{i=1}^n X_i] = \frac{\sum_{i=1}^{n-1} X_i - 1}{n} \left(\frac{n-1}{n} \right)^{\sum_{i=1}^{n-1} X_i - 1}$$

Pelo Teorema, δ^* é o ENVVUM para $\theta e^{-\theta}$.

Exemplo 2

$$\delta^*(x) = \frac{n+1}{n} X_{(n)}, \text{ é o ENVVUM para } \theta$$



$$\delta^*(x) = \frac{n+1}{n} X_{(n)} = E_\theta[(\alpha X_{(n)} - \theta)^2] \rightarrow \alpha = \frac{n+1}{n+2}$$

\downarrow
mínimo

Ex. 3: X_1, \dots, X_n , dados em (μ, σ) , $\epsilon \in \mathbb{R}$ s.t. $\mathbb{E}(\mu + \epsilon)$

$$S_n^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \quad \text{é o ENVOLVIMENTO para } S^2 \text{ de } \sigma^2.$$

$$\mathbb{E}_0 \left[\left(\epsilon \sum_{i=1}^n (X_i - \bar{X})^2 - \sigma^2 \right)^2 \right] \xrightarrow{n \rightarrow \infty} \frac{1}{n-1} \sigma^4,$$

Resumindo:

δ^* é o ENVOLVIMENTO para $\mathbb{E}(\text{var}_0(S(X)))$ se e só se

$$\text{Cov}_0(\delta^*(X), U(X)) = 0, \forall \theta \in \Theta, \forall u \in U,$$

onde $U = \{U: \mathcal{X} \rightarrow \mathbb{R}: \mathbb{E}_0(U(X)) = 0, \forall \theta \in \Theta\}$.

Ex. X_1, \dots, X_n , dados em (μ, σ) , $\theta \in \Theta$.

$$\delta_1(X) = \frac{\sum_{i=1}^n X_{i+1}}{n+1} \approx \delta_1^*(X)$$

$$\delta_2(X) = \frac{\sum_{i=1}^n f_{i+1}}{n+1} \approx \delta_2^*(X)$$

$$\delta_3(X) = \frac{2X_1}{3} \Rightarrow \delta_3^* = \frac{2X_1 + k_m}{3}$$

(c.d.)

$$\text{Cov}_0 \left(\frac{X_{m+1} + k_m}{3}, \underbrace{\frac{\text{var}_0}{n+1} X_{m+1} + \frac{\text{var}_0}{n+1} k_m}_{f_{m+1}} \right)$$

$$= \frac{1}{3} \frac{\text{var}_0}{2(n+1)} \text{Cov}_0(X_{m+1}, X_{m+1}) + \frac{1}{3} \frac{\text{var}_0}{n+1} \text{Cov}_0(X_{m+1}, k_m) + \frac{1}{3} \frac{\text{var}_0}{n+1} \text{Cov}_0(k_m, k_m)$$

$$= \frac{1}{3} \text{Cov}_\theta(X_{(n)}, X_{(n)}) \left(\frac{n+1}{2n+1} - \frac{n+1}{n+2} \right) + \frac{1}{3} \text{Var}_\theta(X_{(n)}) \left(\frac{n+1}{2n+1} - \frac{n+1}{n+2} \right) =$$

$$= \frac{1}{3} \left(\frac{n+1}{2n+1} - \frac{n+1}{n+2} \right) \left\{ \text{Cov}_\theta(X_{(n)}, X_{(n)}) + \text{Var}_\theta(X_{(n)}) \right\} \leq 0, \quad \forall \theta \in \Theta.$$

$$\frac{\theta^2}{(n+2)(n+1)^2} \quad \frac{n\theta^2}{(n+2)(n+1)^2}$$

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

$\text{Var}_\theta(\delta(x)) \geq \frac{(g'(\theta))^2}{I_x(\theta)}$, onde $I_x(\theta)$ é chamada Informação de Fisher para θ contida na amostra,

$$I_x(\theta) = E_\theta[(\lambda(x/\theta))^2], \text{ onde}$$

$$\lambda(x/\theta) = \log f(x/\theta)$$

$$(\mathbb{X}, \sigma, \mathcal{P}) \quad \mathcal{P} = \{f(\cdot | \theta) : \theta \in \Theta\}$$

$$\lambda(x/\theta) = \log f(x/\theta)$$

$$\lambda'(x/\theta) = \frac{d}{d\theta} \log(f(x/\theta)) = \frac{1}{f(x/\theta)} \lambda(x/\theta) \rightarrow \text{FUNÇÃO ESCORE}.$$

Existem 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).

$$\underline{x} = (x_1, \dots, x_n) \in \mathbb{N}^n$$

$$f(\underline{x}/\theta) = P(X_i = x_i/\theta) =$$

$$= \prod_{i=1}^n \frac{e^{-\theta} \theta^{x_i}}{x_i!} = \frac{e^{-n\theta} \theta^{\sum x_i}}{\prod_{i=1}^n x_i!}$$

$$\lambda'(\underline{x}/\theta) = \log f(\underline{x}/\theta) = -n\theta + \sum_{i=1}^n x_i \log \theta - \log \prod_{i=1}^n x_i!$$

$$\lambda'(\underline{x}/\theta) = -n + \frac{\sum x_i}{\theta}.$$

$$\lambda'(\underline{x}/\theta) = -n + \frac{\sum x_i}{\theta}$$

$$I_x = E_\theta[(\lambda'(x/\theta))^2] = E_\theta\left(\left(\frac{\sum x_i}{\theta} - n\right)^2\right) = \frac{1}{\theta^2} E_\theta\left[\left(\sum_{i=1}^n x_i - n\theta\right)^2\right] = \frac{1}{\theta^2} \text{Var}_\theta\left(\sum_{i=1}^n x_i\right) =$$

$$= \frac{n\theta}{\theta^2} \Rightarrow I_x(\theta) = \frac{n}{\theta}.$$

Exemplo: X_1, \dots, X_n dado θ , c.i.i.d. Ber(θ)

$$f(\underline{x}/\theta) = \theta^{\sum x_i} (1-\theta)^{n-\sum x_i}$$

$$\lambda'(\underline{x}/\theta) = \log f(\underline{x}/\theta) = \sum x_i \log \theta + (n - \sum x_i) \log (1-\theta)$$

$$X(\bar{x}/\theta) = \frac{\sum x_i}{\theta} - \frac{(n-\sum x_i)}{1-\theta} \Rightarrow$$

$$X(\bar{x}/\theta) = \frac{\sum x_i(1-\theta) - \theta(n-\sum x_i)}{\theta(1-\theta)} = \frac{\sum x_i - n\theta}{\theta(1-\theta)} \Rightarrow$$

$$X(\bar{x}/\theta) = \frac{\sum_{i=1}^n x_i - n\theta}{\theta(1-\theta)}$$

$$I_x(\theta) = E_\theta[(X(\bar{x}/\theta))^2] = E_\theta\left[\left(\frac{\sum x_i - n\theta}{\theta(1-\theta)}\right)^2\right] = \frac{Var_\theta(\sum x_i)}{\theta^2(1-\theta)^2} = \frac{n}{\theta(1-\theta)}$$

Em geral, X_1, \dots, X_n dado θ , são c.i.i.d.

$$I_x(\theta) = n I_{X_1}(\theta)$$

~~Integrando~~

$$\int_x \delta(x) f(x/\theta) dx = g(\theta), \forall \theta \in \Theta \Rightarrow$$

$$\Rightarrow \frac{d}{d\theta} \int_x \delta(x) f(x/\theta) dx = g'(\theta), \forall \theta \in \Theta$$

$$\Rightarrow \int_x s(x) \frac{df(x/\theta)}{d\theta} f(x/\theta) dx = g'(\theta) \Rightarrow$$

$$\int_x \underbrace{\delta(x) \lambda'(x/\theta)}_{\text{f}(x/\theta)} \cdot f(x/\theta) dx = g'(\theta)$$

$$\Rightarrow E_\theta(\delta(x) \lambda'(x/\theta)) = g'(\theta). \quad (\text{I})$$

$$\int_x f(x/\theta) dp_\theta = 1, \forall \theta \in \Theta$$

$$\Rightarrow \frac{d}{d\theta} \int_x f(x/\theta) dx = 0, \forall \theta \in \Theta$$

$$\Rightarrow \int_x \underbrace{\frac{d}{d\theta} f(x/\theta)}_{f'(x/\theta)} f(x/\theta) dx = 0.$$

$$\Rightarrow E_\theta(\lambda'(x/\theta)) = 0.$$

$$\text{Logo: } \text{Var}_\theta(\lambda'(x/\theta)) = E_\theta((\lambda'(x/\theta))^2) = I_x(\theta).$$

Dar,

$$\text{Cov}_\theta(\delta(x), \lambda'(x/\theta)) = E_\theta(\delta(x) \lambda'(x/\theta)) - 0 \cdot g'(\theta)$$

$$[\text{CORR}_\theta(\delta(x), \lambda'(x/\theta))]^2 \leq 1 \Rightarrow$$

$$\frac{(\text{Cov}(\delta(x), \lambda'(x/\theta)))^2}{\text{Var}_\theta(\delta(x)) \text{Var}_\theta(\lambda'(x/\theta))} \leq 1 \Rightarrow \text{Var}_\theta(\delta(x)) \geq \frac{(g'(\theta))^2}{I_x(\theta)},$$

Exemplo: X_1, \dots, X_n dado θ , c.i.i.d. $\text{Ber}(\theta)$.

$$g(\theta) = \theta$$

$$\text{LI}_0 = \frac{\frac{1}{n}}{\frac{n}{\theta(1-\theta)}} = \frac{\theta(1-\theta)}{n}$$

$$\delta^*(x) = \bar{x}_n \Rightarrow \text{Var}(\bar{x}_n) = n \frac{\theta(1-\theta)}{n^2} = \frac{\theta(1-\theta)}{n} = \text{LI}_0$$

2. A media, no caso das Bernoullis, atinge o Limite Inferior de Cramér-Rao.

Exemplo 2:

x_1, \dots, x_n , dado θ , c.i.d. Poisson(θ)

$$g(\theta) = e^{-\theta}$$

Exercício:

$$S_0(\theta), \overline{I}_0(x) \xrightarrow{\text{Raz-Mackinellha}} E[S_0(\theta)|x_i] = \left(\frac{n-1}{n}\right)^{\sum x_i}$$

Pelo Teorema de Lehmann-Scheffé, $\delta^*(x) = \left(\frac{n-1}{n}\right)^{\sum x_i}$ é o ENVVUM para $e^{-\theta}$.

$$\text{Var}(\delta^*(x)) = E_\theta((\delta^*(x))^2) - (E_\theta(\delta^*(x)))^2 =$$

$$= E_\theta\left(\left(\frac{n-1}{n}\right)^{\sum x_i}\right) - e^{-2\theta}$$

$$= \left\{ \sum_{j=0}^{\infty} \frac{e^{-n\theta} (n\theta)^j}{j!} \left(\frac{n-1}{n}\right)^{2j} \right\} - e^{-2\theta}$$

$$= e^{-n\theta} \sum_{j=0}^{\infty} \frac{\left(\frac{n(n-1)^2}{n^2}\right)^j}{j!} - e^{-2\theta} = e^{\frac{\theta(n-1)^2}{n} - n\theta} - e^{-2\theta} =$$

$$= e^{\frac{\theta(-2n+1)}{n}} - e^{-2\theta} \Rightarrow \text{Var}_\theta(\delta^*(x)) = e^{-2\theta}(e^{\frac{\theta}{n}} - 1)$$

$$LI_0 = \frac{(-e^{-\theta})^2}{n/\theta} = e^{-2\theta} \cdot \frac{\theta}{n}$$

$$e^{-2} \approx 0.135$$

Logo, $V_{\theta_0}(S^*(X)) \geq LI_0$.

$$\theta = ? \quad \theta \in (0,1)$$

X: Número de sucessos em 10 tentativas

$$X/10 \sim \text{Bin}(10, \theta)$$

$$\frac{X}{10} = S_1(X) \quad (\text{suficiente})$$

Y: Entrevista indivíduos até obter o 1º sucesso.

$$Y/\theta \sim \text{BINNEG}(1, \theta)$$

$$S_2(Y) = \frac{2-1}{Y-1}$$

• Se $X \sim \text{BINNEG}(1, \theta)$

$$E\left(\frac{K-1}{X-1} \mid \theta\right) = \theta$$

$$\sum_{k=0}^{\infty} \frac{k-1}{k-1} \binom{k-1}{k-1} \theta^k (1-\theta)^{1-k} = \dots = \theta$$

NNNNNNNNNN

$$S_1(x) = \frac{2}{10} \quad (x=2) \quad \text{e} \quad S_2(x) = \frac{1}{9}$$

$$E_\theta(S(X)) = \theta$$

PRINCÍPIO DA VEROSIMILHANÇA.

"Toda a inferência deve ser baseada apenas no ponto amostral que é efetivamente observado".

$$\text{Exp. L } V_x^{(1)}(\theta) = k(x,y) V_y^2(\theta), \forall \theta \in \Theta$$

↓

as conclusões a partir de x e y devem coincidir.

Exp. 2.

- Dissertação de mestrado (1995)

Lourdes Yoshioka T. Inoue

- Princípio da Verossimilhança

"The Likelihood Principle". Berger & 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. $\text{Geo}(\theta)$.

$$g(\theta) = \frac{1}{\theta}$$

$$\delta^*(x) = \bar{X}_n$$

$$\text{Var}_{\theta}(\delta^*(x)) = \text{Var}_{\theta}(\bar{X}_n) = \frac{1}{n} \text{Var}(X_1) = \frac{1-\theta}{n\theta^2}$$

$$f(x/\theta) = \prod_{i=1}^n (1-\theta)^{x_i-1} \cdot \theta = (1-\theta)^{\sum_{i=1}^n x_i - n} \theta^n$$

$$\lambda(x/\theta) = (\sum_{i=1}^n x_i - n) \log(1-\theta) + n \log \theta$$

$$\lambda'(x/\theta) = -\frac{\sum_{i=1}^n x_i - n}{1-\theta} + \frac{n}{\theta}$$

$$I_x(\theta) = E_{\theta} \left[\left(\frac{-\sum_{i=1}^n x_i + n}{1-\theta} + \frac{n}{\theta} \right)^2 \right] = \frac{1}{\theta^2(1-\theta)^2} E_{\theta} \left[(-\theta \sum_{i=1}^n x_i + n\theta + n - n\theta)^2 \right] =$$

$$= \frac{1}{\theta^2(1-\theta)^2} E_{\theta}((\theta \sum_{i=1}^n x_i - n)^2) = \frac{\theta^2}{\theta^2(1-\theta)^2} E_{\theta} \left(\left(\sum_{i=1}^n x_i - \frac{n}{\theta} \right)^2 \right) = \frac{1}{(1-\theta)^2} n \frac{(1-\theta)}{\theta^2} =$$

$$= \frac{n}{\theta^2(1-\theta)}$$

$$LI(\theta) = \frac{(g'(\theta))^2}{I_x(\theta)} = \frac{(-\gamma \theta^2)^2}{\frac{n}{\theta^2(1-\theta)}} = \frac{1-\theta}{n\theta^2} = \text{Var}_{\theta}(\delta^*(x))$$

Logo, $\delta^*(x)$ é o ENVVUM para θ .

* forma alternativa

$$E_{\theta}(\lambda'(x)) = 0 \Rightarrow \int_{-\infty}^{\infty} \lambda'(x/\theta) f(x/\theta) dx = 0 \Leftrightarrow \int_{-\infty}^{\infty} \frac{d(\lambda'(x/\theta) f(x/\theta))}{d\theta} dx = 0 \Rightarrow$$
$$\int_{-\infty}^{\infty} \left\{ \lambda''(x/\theta) f(x/\theta) + \lambda'(x/\theta) \frac{\frac{df}{d\theta}}{f(x/\theta)} f(x/\theta) \right\} dx = 0 \Rightarrow$$
$$\Rightarrow E_{\theta}(\lambda''(x/\theta)) + \underbrace{E_{\theta}((\lambda'(x/\theta))^2)}_{I_x(\theta)} = 0 \Rightarrow I_x(\theta) = -E_{\theta}[\lambda''(x/\theta)].$$

Propriedades Assintóticas de Estimadores

Ter em mente sequência de v.a. $(X_n)_{n \geq 1}$.

$$X = (X_n)_{n \geq 1}$$

$$\delta_n(X) = \delta_n(x_1, \dots, x_n)$$

$(\delta_n)_{n \geq 1}$ sequência de estimadores

Definição: $(\delta_n)_{n \geq 1}$ é uma sequência consistente para $\theta(g(\theta))$ se

$$\delta_n(X) \xrightarrow{P_{\theta}} \theta(g(\theta)), \quad \forall \theta \in \Theta$$

Exemplo: $(X_n)_{n \geq 1}$, dado θ c.i.i.d. $U(0, \theta)$.
 (x_1, \dots, x_n)

$$\delta_n^*(X) = X_{(n)}$$

$$\delta_n(X) = \frac{n+1}{n} X_{(n)}$$

para $\epsilon > 0$,

$$P_{\theta}(|\delta_n(X) - \theta| \geq \epsilon) = P_{\theta}(|X_{(n)} - \theta| \geq \epsilon) = P_{\theta}(X_{(n)} \leq \theta - \epsilon) + P_{\theta}(X_{(n)} \geq \theta + \epsilon) =$$

$$= P_{\theta} (X_1 \in \Theta - \delta, \dots, X_n \in \Theta - \delta)$$

$$\Theta \subset \mathbb{R} \subset \Theta$$

$$= \left(\frac{\Theta - \delta}{\Theta} \right)^n \xrightarrow{n \rightarrow \infty} 0$$

Logo, $(\delta_n^*)_{n \geq 1}$ é uma sequência consistente para Θ .

$$\delta'_n(x) = \frac{n+1}{n} \delta_n^*(x) \xrightarrow{\substack{P_{\theta} \\ \text{L} \rightarrow \text{seq. conv. para } \Theta}} 1 \cdot \Theta, \forall \theta \in \Theta$$

\hookrightarrow seq. converge para 1

Logo, $(\delta'_n)_{n \geq 1}$ também é consistente para Θ .

Exemplo 2. $(X_n)_{n \geq 1}$, dado θ , c.i.i.d. $Ber(\theta)$

$$\bar{X}_n \xrightarrow{P_{\theta}} \theta, \forall \theta \in \Theta$$

$$1 - \bar{X}_n \xrightarrow{P_{\theta}} 1 - \theta, \forall \theta \in \Theta$$

$$\bar{X}_n(1 - \bar{X}_n) \xrightarrow{P_{\theta}} \theta(1 - \theta), \forall \theta \in \Theta$$

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

$$\sqrt{n} (\delta_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta)), \forall \theta \in \Theta$$

(para n "grande" $\delta_n \sim N\left(\theta, \frac{\sigma^2(\theta)}{n}\right)$)

Exemplo 1. X_1, \dots, X_n , dado θ , c.i.i.d. $Ber(\theta)$.

$$(X_n)_{n \geq 1}$$

$$\frac{\sum_{i=1}^n X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \xrightarrow{d} N(0,1)$$

"

$$\frac{n\bar{X}_n - n\theta}{\sqrt{n\theta(1-\theta)}}$$

"

$$\frac{\sqrt{n}(\bar{X}_n - \theta)}{\sqrt{\theta(1-\theta)}} \xrightarrow{d} N(0,1) \Rightarrow \sqrt{n}(\bar{X}_n - \theta) \xrightarrow{d} N(0, \theta(1-\theta))$$

Assim, $(\bar{X}_n)_{n \geq 1}$ é assintoticamente Normal $(0, \theta(1-\theta))$.

Exemplo. X_1, \dots, X_n , dados θ , c.i.i.d. tais que $E(X_1 | \theta) = \theta < \text{Var}(X_1 | \theta) = \sigma_\theta^2$, σ_θ^2 conhecido!

$$\frac{\bar{X}_n - \theta}{\sigma_\theta / \sqrt{n}} \xrightarrow{d} N(0,1) \Rightarrow \frac{\sqrt{n}(\bar{X}_n - \theta)}{\sigma_\theta} \xrightarrow{d} N(0,1) \Rightarrow$$

$$\sqrt{n}(\bar{X}_n - \theta) \xrightarrow{d} N(0, \sigma_\theta^2)$$

Assim, $(\bar{X}_n)_{n \geq 1}$ é Assintoticamente Normal $(0, \sigma_\theta^2)$.

Exemplo...

$$\sigma_\theta^2 = 1$$

Para cada $n \in \mathbb{N}$

Seja $M_{dn}(x) = \text{mediana}\{X_1, \dots, X_n\}$

$(M_{dn}(x))_{n \geq 1}$ é assintoticamente NORMAL $(0, \pi/2)$.

$$S_n = \sum_{1 \leq i \leq n, i \neq dn} I(X_i)$$

$$P(\sqrt{n}(M_{dn}(x) - \theta) \leq t) = P(M_{dn}(x) \leq \theta + t/\sqrt{n}) = P(S_n \leq \lfloor \frac{n}{2} \rfloor)$$

$$S_n / \theta \approx \text{Bin}(n, P_\theta(X_1 > \theta + t/\sqrt{n}))$$

$$P_\theta(S_n \leq \theta/2) = P_\theta\left(\frac{S_n - nP_\theta(X_1 > \theta + t/\sqrt{n})}{\sqrt{n}P_\theta(X_1 > \theta + t/\sqrt{n})(1 - P_\theta(X_1 > \theta + t/\sqrt{n}))}\right) \approx \frac{\frac{n}{2} - nP_\theta(X_1 > \theta + t/\sqrt{n})}{\sqrt{n}P_\theta(X_1 > \theta + t/\sqrt{n})(1 - P_\theta(X_1 > \theta + t/\sqrt{n}))}$$

$$\approx P\left(\frac{\frac{n}{2} - P_\theta(X_1 > \theta + t/\sqrt{n})}{\sqrt{\frac{P_\theta(X_1 > \theta + t/\sqrt{n})(1 - P_\theta(X_1 > \theta + t/\sqrt{n}))}{n}}}\right) \Rightarrow$$

↓

2º (θ). t. ordem

f é a densidade de $N(\theta, 1)$

$$P(\sqrt{n}(M_d(x) - \theta) \leq t) \xrightarrow[n \rightarrow \infty]{} P(\zeta \leq t \cdot 2f(\theta))$$

ou seja,

$$\sqrt{n}(M_d(x) - \theta) \xrightarrow{d} N\left(0, \left(\frac{1}{2f(\theta)}\right)^2\right) \stackrel{d}{=} N\left(0, \left(\frac{1}{2\sqrt{2\pi}}\right)^2\right) \stackrel{d}{=} N\left(0, \frac{\pi}{2}\right)$$

$$f(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta-\mu)^2}{2}} = \frac{1}{\sqrt{2\pi}} \quad \text{Logo, } (M_d(x))_{n \geq 1} \text{ é ass. Normal}(0, \frac{\pi}{2})$$

Resultado: $(\delta_n)_{n \geq 1}$ é assintoticamente Normal $(0, \sigma^2(\theta)) \Rightarrow (\delta_n)_{n \geq 1}$ é consistente para θ

Para $\epsilon > 0$,

$$P_\theta(|\delta_n(x) - \theta| > \epsilon) = P_\theta(|\sqrt{n}(\delta_n(x) - \theta)| > \epsilon\sqrt{n})$$

Para todo $M \in \mathbb{N}^*$, $\exists n_0 = n_0(M)$ tal que

$$n \geq n_0(M) \Rightarrow \epsilon\sqrt{n} > M. \quad \text{Para } n \geq n_0,$$

$$P_\theta(|\sqrt{n}(\delta_n(x) - \theta)| > \epsilon\sqrt{n}) \leq P_\theta(|\sqrt{n}(\delta_n(x) - \theta)| > M)$$

Então,

$$\limsup_{n \rightarrow \infty} P_\theta(|\sqrt{n}(\delta_n(x) - \theta)| > \epsilon\sqrt{n}) \leq \underbrace{\lim_{n \rightarrow \infty} P_\theta(|\sqrt{n}(\delta_n(x) - \theta)| > M)}_{P_\theta(N(0, \sigma^2(\theta)) > M)}, \forall M \in \mathbb{N}$$

Então,

$$\limsup_{n \rightarrow \infty} P_\theta(|\sqrt{n}(\delta_n(x) - \theta)| > \epsilon\sqrt{n}) \leq \underbrace{\lim_{M \rightarrow \infty} P_\theta(N(0, \sigma^2(\theta)) > M)}_{0}$$

Assim,

$$\lim_{n \rightarrow \infty} P(|\delta_n(x) - \theta| > \epsilon) = 0, \forall \epsilon > 0, \forall \theta \in \Theta$$

Logo, $\delta_n \xrightarrow{P_\theta} \theta, \forall \theta \in \Theta$ e

$(\delta_n)_{n \geq 1}$ é consistente para θ .

Algumas outras definições:

Definição: $(\delta_n^{(1)})_{n \geq 1}$ seq. assintoticamente normal $(0, \sigma_1^2(\theta)), i=1, 2,$

Dizemos que $(\delta_n^{(1)})_{n \geq 1}$ é assintoticamente melhor que $(\delta_n^{(2)})_{n \geq 1}$, se $\sigma_1^2(\theta) \leq \sigma_2^2(\theta) \quad \forall \theta \in \Theta$.

No exemplo, x_1, \dots, x_n dado θ , i.i.d. $N(\theta, 1)$

$\bar{x}_n(\bar{x}_n - \theta) \xrightarrow{d} N(0, 1) \quad (\sigma_1^2(\theta) = 1, \forall \theta)$ Logo, (\bar{x}_n) é assintoticamente melhor que $(M_{\bar{x}_n}(x))_{n \geq 1}$.

Definição: $(\delta_n^+)_n \in \mathbb{N}$ é "best asymptotically normal" (BAN) se

(1) $\sqrt{n}(\delta_n^+ - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$

(2) $(\delta_n^+)_n$ é melhor assintoticamente que qualquer outra $(\delta_n)_n$ quanto tecnicamente normal.

"Como se fosse o ENVVUM para os ass. normais"

Definição: $(\delta_n)_n \in \mathbb{N}$ e $(\delta'_n)_n \in \mathbb{N}$ são assintoticamente equivalentes se

$$\sqrt{n}(\delta_n - \delta'_n) \xrightarrow{P} 0, \forall \epsilon \in \mathbb{R}$$

Ex. 1: X_1, \dots, X_n , dado θ , c.i.i.d. $\mathcal{U}(0, \theta)$

$(X_{(n)})_n \in \left(\frac{n+1}{n} X_{(n)}\right)_n$ são equivalentes

$\forall \epsilon > 0$,

$$P_\theta(|\sqrt{n}(X_{(n)} - \frac{n+1}{n} X_{(n)})| > \epsilon) = P_\theta\left(|\sqrt{n}\left(-\frac{1}{n} X_{(n)}\right)| > \epsilon\right) =$$

$$= P_\theta\left(\left|\frac{X_{(n)}}{\sqrt{n}}\right| > \epsilon\right) = P_\theta(X_{(n)} > \epsilon\sqrt{n}) = 0, \forall n \geq n_0(\epsilon), \text{ onde}$$

$$n_0(\epsilon) = \min\{n \in \mathbb{N}: \epsilon\sqrt{n} > \theta\}. \text{ Logo, } \sqrt{n}(X_{(n)} - \frac{n+1}{n} X_{(n)}) \xrightarrow{P} 0, \forall \theta \in \mathbb{R}$$

Assim, $(X_{(n)})_n \in \left(\frac{n+1}{n} X_{(n)}\right)_n$ são assintoticamente equivalentes.

Exercício. X_1, \dots, X_n dado Θ , c.i.i.d. $\text{Ber}(\theta)$

$$\delta_n(x) = \bar{x}_n$$

$$\delta'_n(x) = \frac{a + \sum x_i}{a + b + n}$$

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,

$$\sqrt{n} (\delta_{MV}^{(n)} - \theta) \xrightarrow{d} N\left(0, \frac{1}{I_{X_n}(\theta)}\right), \text{ no caso } (X_n)_{n \geq 1}, \text{ dado } \theta, \text{ c.i.i.d.}$$

$$(\delta_{MV}^{(n)}(x)) \sim N\left(\theta, \frac{1}{n I_{X_n}(\theta)}\right)$$

$$z = (z_1, \dots, z_n)$$

$$\lambda(z/\theta) = \log f(z/\theta) = \log \prod_{i=1}^n f(z_i/\theta) = \sum_{i=1}^n \log(f(z_i/\theta)) +$$

$$\lambda'(z/\theta) = \underbrace{\sum_{i=1}^n \frac{d \log f(z_i/\theta)}{d\theta}}_{l_z(\theta)} \rightarrow \lambda'(z/\theta)$$

$$l_z(\theta)$$

No ponto $\theta_0 \in \Theta$

$$l_z(\theta) = l_z(\theta_0) + l_z'(\theta_0)(\theta - \theta_0) + \dots \approx l_z(\theta_0) + l_z'(\theta_0)(\theta - \theta_0) \quad (I)$$

Aplicando (I) no ponto $\theta = \delta_{MV}(z)$ temos

$$\underbrace{l_z(\delta_{MV}(z))}_{\delta_z} = l_z(\theta_0) + l_z'(\theta_0)(\delta_{MV}(z) - \theta_0) \Rightarrow$$

$$\delta_{\text{MV}}(x) - \theta_0 = -\frac{\lambda(x|\theta_0)}{I_x(\theta_0)}$$

Assim,

$$\begin{aligned} \delta_{\text{MV}}(x) - \theta_0 &= -\frac{\sum_{i=1}^n \lambda_i^*(x_i|\theta_0)}{\sum_{i=1}^n \lambda_i^*(x_i|\theta)} \Rightarrow \sqrt{n} (\delta_{\text{MV}}(x) - \theta_0) = \\ &= -\sqrt{n} \frac{\sum_{i=1}^n \lambda_i^*(x_i|\theta)}{\sum_{i=1}^n \lambda_i^*(x_i|\theta)} \\ &\stackrel{d}{\longrightarrow} N(0, I_{X_n}(\theta_0)) \end{aligned}$$

$\sqrt{n} \left[\frac{\sum \lambda_i^*(x_i|\theta_0) - n E[\lambda^*(x_i|\theta)]}{\sqrt{n} [I_{X_n}(\theta)]^{1/2}} \right] \xrightarrow{d} N(0, I_{X_n}(\theta_0))$
 $\xrightarrow{q.c.} E_\theta(-\lambda^*(x_i|\theta)) = J_{X_n}(\theta_0)$

$$\# \text{Var}(\lambda^*(x_i|\theta)) = \sqrt{J_{X_n}(\theta)}$$

$$\xrightarrow{d} N(0, I_{X_n}(\theta_0)) \xrightarrow{d} N\left(0, \frac{1}{J_{X_n}(\theta_0)}\right).$$

Exemplo: $\mathfrak{X} = \mathbb{R}_+$, $\Theta = \mathbb{R}_+$

X_1, \dots, X_n , dado θ , c.i.i.d. Gama($\theta, 1$)

$$f(x_i|\theta) = \frac{1}{\Gamma(\theta)} x_i^{\theta-1} e^{-x_i} I_{\mathbb{R}_+}(x_i)$$

$$x = (x_1, \dots, x_n) \in \mathbb{R}_+^n$$

$$f(x|\theta) = \prod_{i=1}^n f(x_i|\theta) = \prod_{i=1}^n \frac{x_i^{\theta-1} e^{-x_i}}{\Gamma(\theta)} \Rightarrow f(x|\theta) = \frac{\left(\prod_{i=1}^n x_i\right)^{\theta-1} e^{-\sum_{i=1}^n x_i}}{(\Gamma(\theta))^n}$$

$$\lambda(x|\theta) = \log f(x|\theta) = (\theta - 1) \sum_{i=1}^n \log x_i - \sum_{i=1}^n x_i - n \log \Gamma(\theta) \Rightarrow$$

$$I_x(\theta) = \frac{n \Gamma''(\theta) \Gamma(\theta) - (\Gamma'(\theta))^2}{(\Gamma(\theta))^2}$$



$$\chi^2(x/\theta) = \sum_{i=1}^n \log x_i - n \frac{\pi'(\theta)}{\pi(\theta)} \stackrel{?}{=} 0$$

$$\chi''(x/\theta) = \left\{ n \frac{\pi''(\theta)\pi(\theta) - \pi'(\theta)\pi'(\theta)}{(\pi(\theta))^2} \right\}$$

$$E_\theta(-\chi''(x/\theta)) = -E_\theta(\cdot)$$

\Rightarrow

$$I_x(\theta) = n \frac{\pi''(\theta)\pi(\theta) - (\pi'(\theta))^2}{(\pi(\theta))^2}$$

$$\text{#} \quad \sqrt{n}(\delta_n(x) - \theta) \xrightarrow{D} N\left(0, \frac{1}{I_{xx}(\theta)}\right) \stackrel{D}{=} N\left(0, \frac{\pi(\theta)^2}{(\pi'(\theta)\pi(\theta)) - (\pi'(\theta))^2}\right)$$

$$\lambda''(x|\theta) = - \frac{[+n \cdot r''(\theta)r'(\theta) - r'(\theta)r'(\theta)]}{(r(\theta))^2}$$

$$E_{\theta}(-\lambda''(x|\theta)) = -E_{\theta}(\dots)$$

$$\Rightarrow \pm_x(\theta) = n \cdot \frac{r''(\theta)r(\theta) - (r'(\theta))^2}{(r(\theta))^2}$$

$$\sqrt{n}(\delta_n(x) - \theta) \xrightarrow{\text{d.f.}} N(0, \frac{1}{\pm_{x,\theta}}) \stackrel{\text{d.f.}}{=} N(0, \frac{(r(\theta))^2}{(r''(\theta)r(\theta)) - (r'(\theta))^2})$$

de where only:

$$\text{Sob certas condições, } \sqrt{n}(\delta_n - \theta) \xrightarrow{\text{d.f.}} N(0, \frac{1}{\pm_{x,\theta}})$$

Resumo (Método Delta): $(\delta_n)_{n \geq 1}$ tal que $\sqrt{n}(\delta_n - \theta) \xrightarrow{\text{d.f.}} N(0, \sigma^2(\theta))$.

Seja $g: \mathbb{R} \rightarrow \mathbb{R}$ diferenciável. Então,

$$\sqrt{n}(g(\delta_n) - g(\theta)) \xrightarrow{\text{d.f.}} N(0, \sigma^2(\theta) (g'(\theta))^2).$$

$[(g(\delta_n))_{n \geq 1} \text{ é assimótica, é Normal}]$

$\sqrt{n}(\delta_n - \theta) \xrightarrow{\text{d.f.}} N(0, \sigma^2(\theta)) \Rightarrow (\delta_n)_{n \geq 1} \text{ é consistente, i.e., } \delta_n \xrightarrow{\text{P.s.}} \theta$

$$\frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} \xrightarrow{\text{P.s.}} g'(\theta). \quad (*)$$

1 / 1

Para todo $\epsilon > 0$, $\exists \delta > 0$,

$$|\delta_n - \theta| < \delta \Rightarrow \left| \frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} - g'(\theta) \right| < \epsilon$$

Assim,

$$\mathbb{P}(|\delta_n - \theta| < \delta) \leq \mathbb{P}\left(\left|\frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} - g'(\theta)\right| < \epsilon\right)$$

Como $\delta_n \xrightarrow{P} \theta$, temos

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} - g'(\theta)\right| < \epsilon\right) = 1.$$

Logo, vale ④

Portanto, $\sqrt{n}(g(\delta_n) - g(\theta)) = \frac{g(\delta_n) - g(\theta)}{\delta_n - \theta} \cdot \sqrt{n}(\delta_n - \theta)$

\xrightarrow{D}

$$N(0, \sigma^2(\theta) \cdot (g'(\theta))^2)$$

Aplicando $\sqrt{n}(\delta_{MV}^{(n)} - \theta) \xrightarrow{D} N(0, \frac{1}{f_x(\theta)})$

$(g(\delta_{MV}^{(n)}))_{n \geq 1}$. Do resultado,

$$\underbrace{\sqrt{n}(g(\delta_{MV}^{(n)}) - g(\theta))}_{g'(\theta)} \xrightarrow{D} N(0, \frac{(g'(\theta))^2}{f_x(\theta)})$$

Encontrar M.L.E. da dada θ , com $\text{Exp}(\theta)$

$$\hat{\delta}_{\text{ML}}(\theta) = \frac{1}{\bar{x}_n} \cdot (\text{Utilizando a quantização } \underline{\theta} \text{ e } \overline{\theta})$$

Então

$$\ln(\hat{\delta}_{\text{ML}}(\theta) - \theta) \xrightarrow{D} N\left(0, \frac{1}{\hat{f}_\theta(\theta)}\right).$$

(calcular $\hat{f}'_\theta(\theta)$, temos

$$Y(x|\theta) = \log(\theta) - \theta x$$

$$Y'(x|\theta) = \frac{1}{\theta} - x$$

$$Y''(x|\theta) = -\frac{1}{\theta^2}$$

$$\hat{f}'_\theta(\theta) = -E_{\theta} [Y''(x|\theta)] = \frac{1}{\theta^2}$$

$\log_x \propto \ln(\hat{\delta}_{\text{ML}}(\theta)) \xrightarrow{D} N(0, \theta^2)$

Para $g(\theta) = 1/\theta$, utilizando o método Delta:

$$\sqrt{n} \left(\bar{x}_n - \frac{1}{\theta} \right) \xrightarrow{D} N\left(0, \left(\frac{-1}{\theta^2}\right) \sigma^2\right) \stackrel{d}{=} N\left(0, \frac{1}{\theta^2}\right)$$

$$E_h h(\theta) = \frac{1}{\theta^2} = \text{Var}_\theta(X_i)$$

$$\widehat{h(\theta)}_{\text{ML}} = h(\hat{\delta}_{\text{ML}}) = \frac{1}{\left(\frac{1}{\bar{x}_n}\right)^2} = \bar{x}_n^2$$

/ /

Pelo método Delta,

$$\frac{\sqrt{n}(\hat{h}(\theta) - \frac{1}{\theta^2})}{\theta^2} \xrightarrow{\text{D}} N\left(0, \left(\frac{-2}{\theta^3}\right)^2 \cdot \theta^2\right)$$

\bar{x}_n^2

θ^2

$$N(0, 4).$$

INTERVALOS DE CONFIANÇA

$$\begin{cases} \delta: \mathcal{X} \rightarrow \Theta \\ \theta \rightarrow \{[a, b]; a < b\} \end{cases}$$

$$(\mathcal{X}, \mathcal{A}, \mathcal{P}) \quad \mathcal{P} = \{f(\cdot | \theta) : \theta \in \Theta\}$$

$$g(\Theta) \subseteq \mathbb{R} (\mathbb{R}^k)$$

Sejam $A, B: \mathcal{X} \rightarrow \mathbb{R}$ tais que $A(x) \leq B(x), \forall x \in \mathcal{X}$.

$$\mathcal{I} = \{[a, b] : a \leq b\}$$

Def.: A transformação $T: \mathcal{X} \rightarrow \mathcal{I}$ que a cada $x \in \mathcal{X}$ associa um intervalo em \mathcal{I} (isto é, $T(x) = [a(x), b(x)]$) denos o nome de INTERVALO ALEATÓRIO.

Exemplo: X_1, \dots, X_n dada θ , c.iid $N(\theta, 1)$ $X = (X_1, \dots, X_n)$

$$\text{Seja } A(X) = \bar{X} - 1$$

$$B(X) = \bar{X} + 2$$

$[A(X), B(X)]$ é um intervalo aleatório.

DEF: O intervalo aberto $[A(\bar{x}), B(\bar{x})]$ é um intervalo de confiança para $G(g(\theta))$ com coeficiente de confiança $1-\alpha$, $0 < \alpha < 1$, se

$$P(A(\bar{x}) \leq g^{-1}(B(\bar{x})) | \theta) \geq 1-\alpha, \forall \theta \in \Theta.$$

$\hookrightarrow P(\{\bar{x} \in \mathbb{R} : A(\bar{x}) \leq g^{-1}(B(\bar{x}))\} | \theta)$



$$[A(\bar{x}), B(\bar{x})]$$

$$[A(\bar{x}_1), B(\bar{x}_2)]$$

$$A(\bar{x}) \quad B(\bar{x})$$

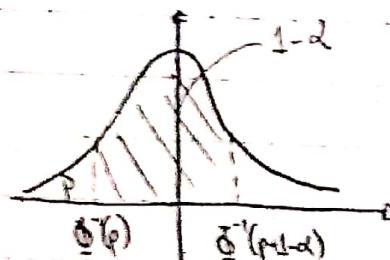
$$A(\bar{x}_1) \quad B(\bar{x}_2)$$

$$A(\bar{x}_1) \quad B(\bar{x}_2)$$

$$\bar{x} \quad R$$

Ex1. X_1, \dots, X_n , dada Θ , c.i.d $N(\theta, \sigma^2_0)$ (verificá)

$$\frac{\bar{X}_n - \theta}{\sigma/\sqrt{n}} \mid \theta \sim N(0, 1).$$



Para de $\Phi(\mu) = p$, $\mu \in (0, 1)$,

$$P\left(\Phi^{-1}(p) \leq \frac{\bar{X}_n - \theta}{\sigma/\sqrt{n}} \leq \Phi^{-1}(1-\alpha) \mid \theta\right) = 1-\alpha, \forall \theta \in \Theta, \text{ onde}$$

$\Phi^{-1}(\mu) \in \mathbb{R} \Leftrightarrow \mathbb{P}(Z \leq \Phi^{-1}(\mu)) = \mu, \mu \in (0, 1)$.

$$\Rightarrow P\left(-\Phi^{-1}(1-\alpha) \leq \frac{\bar{X}_n - \theta}{\sigma/\sqrt{n}} \leq \Phi^{-1}(p) \mid \theta\right) \geq 1-\alpha, \forall \theta \in \Theta.$$

$$\Rightarrow P\left(\underbrace{\bar{X}_n - \Phi^{-1}(1-\alpha) \frac{\sigma_0}{\sqrt{n}}} \leq \theta \leq \underbrace{\bar{X}_n - \Phi^{-1}(p) \frac{\sigma_0}{\sqrt{n}}} \mid \theta\right) \geq 1-\alpha, \forall \theta \in \Theta$$

$A(\bar{x}) \quad B(\bar{x})$

1 /

Logo, $\left[\bar{X}_n - \Phi^{-1}(p) \frac{\sqrt{\theta}}{\sqrt{n}}, \bar{X}_n - \Phi^{-1}(p) \frac{\sqrt{\theta}}{\sqrt{n}} \right]$ 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 !"

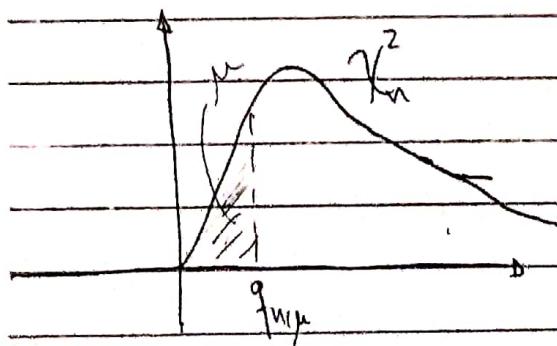
Ex. 2: X_1, \dots, X_n , dada θ , c.i.i.d $N(\mu_0, \theta)$ $T(X) = \sum_{i=1}^n (X_i - \mu_0)$

$$\frac{X_i - \mu_0}{\sqrt{\theta}} \sim N(0, 1)$$

$$\frac{(X_i - \mu_0)^2}{\theta} | \theta \sim \chi^2_1 \Rightarrow \sum_{i=1}^n \frac{(X_i - \mu_0)^2}{\theta} | \theta \sim \chi^2_n$$

$$P\left(q_{n,p} \leq \sum_{i=1}^n \frac{(X_i - \mu_0)^2}{\theta} \leq q_{n,p+1-\alpha}\right) = 1 - \alpha, \text{ V.G.E.R.}$$

$q_{n,\alpha}$ é o percentil de ordem α , $\alpha \in (0,1)$, da dist. χ^2_n .



$$\Rightarrow P\left(\frac{1}{q_{n,p+1-\alpha}} \leq \frac{\theta}{\sum_{i=1}^n (X_i - \mu_0)^2} \leq \frac{1}{q_{n,p}} | \theta\right) = 1 - \alpha, \text{ V.G.E.R.}$$

$$\Rightarrow P\left(\frac{\sum_{i=1}^n (X_i - \mu_0)^2}{q_{n,p+1-k}} < \theta \right) = \sum_{i=1}^n \frac{(X_i - \mu_0)^2}{q_{n,p}} | \theta = 1 - \alpha, \forall \theta \in \mathbb{R}_+$$

Logo, $[A(x), B(x)]$ é um I.C para Θ com coeficiente de confiança 1- α .

Obs. Para encontrar $\hat{\mu}$ que minimize o comp do IC, devemos minimizar $E \left\{ \sum_{i=1}^n (x_i - \mu)^2 \left[\frac{1}{f_{\text{trip}}} - \frac{1}{f_{\text{trips-}\infty}} \right] \right\}$

Ex 3: X_1, \dots, X_n , i.i.d. $\Theta = (\theta_1, \theta_2)$, c.i.d. $N(\theta_1, \theta_2)$

$$g(\theta) = \theta_1.$$

$$\frac{X_n - \theta_1}{\sqrt{\theta_2}} \sim N(0,1).$$

$$\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\Theta_2} | \Theta \sim \chi^2_{n-1}$$

$$\frac{X_n - \theta_1}{\sqrt{\sigma_n^2}} \mid \theta_1 \sim t_{n-1}.$$

$$\sqrt{\frac{\sum_{i=1}^n (\lambda_i - x)^2}{\Theta_2}} = \sqrt{(n-1)}$$

$$\frac{\sqrt{n}(\bar{X}_n - \theta_0)}{\sqrt{s^2}} \stackrel{D}{\rightarrow} t_{n-1}$$

/ /

$$\mathbb{P} \left(\frac{\bar{x} - \theta_1}{\frac{s_x}{\sqrt{n}}} \leq \frac{\bar{x}_n - \theta_1}{\frac{s_x}{\sqrt{n}}} \leq \frac{\bar{x} + t_{n-1, \alpha/2}}{\frac{s_x}{\sqrt{n}}} \mid \theta \right) = 1 - \alpha, \forall \theta \in \Theta,$$

onde $t_{n-1, \alpha/2}$ é o percentil de ordem $1 - \alpha/2$ da distribuição

t-Student de parâmetros $n-1$.

$$\Rightarrow \mathbb{P} \left(\frac{\bar{x}_n - \theta_1}{\frac{s_x}{\sqrt{n}}} \leq \frac{\theta_1 - \bar{x}_n}{\frac{s_x}{\sqrt{n}}} \leq \frac{\bar{x}_n + t_{n-1, \alpha/2}}{\frac{s_x}{\sqrt{n}}} \mid \theta \right) = 1 - \alpha, \forall \theta \in \Theta.$$

$$\Rightarrow \mathbb{P} \left(\underbrace{\bar{x}_n - t_{n-1, \alpha/2} \cdot \frac{s_x}{\sqrt{n}}}_{A(\bar{x})} \leq \theta_1 \leq \underbrace{\bar{x}_n + t_{n-1, \alpha/2} \cdot \frac{s_x}{\sqrt{n}}}_{B(\bar{x})} \mid \theta \right) = 1 - \alpha, \forall \theta \in \Theta.$$

Logo, $[A(\bar{x}), B(\bar{x})]$ é um t.c. para $\theta_1 = g(\theta)$ com coef. de confiança $1 - \alpha$.

Ex. 4: $X_1, \dots, X_n, Y_1, \dots, Y_m$, dada $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$, condicionalmente independentes tais que:

$$X_i | \theta \sim N(\mu_1, \sigma_1^2), \quad i=1, \dots, n \text{ e}$$

$$Y_j | \theta \sim N(\mu_2, \sigma_2^2), \quad j=1, \dots, m.$$

$$g(\theta) = \sigma_2^2.$$

$$\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma_1^2} | \theta \sim \chi^2_{n-1}.$$

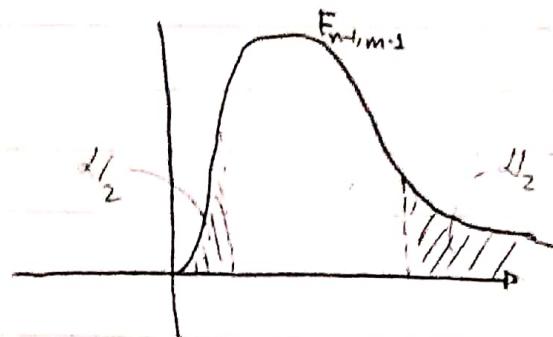
$$\sum_{j=1}^m \frac{(Y_j - \bar{Y})^2}{\sigma_2^2} | \theta \sim \chi^2_{m-1}.$$

$$\begin{aligned} & \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\sigma_x^2} \sim \chi^2_{n-1} \\ & \sum_{j=1}^m \frac{(y_j - \bar{y})^2}{\sigma_y^2} \sim \chi^2_{m-1} \end{aligned}$$

isto é,

$$\frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{s_x^2}{s_y^2} \Big| \theta \sim F_{n-1, m-1}$$

$$|\theta \sim F_{n-1, m-1}$$



$$P(F_{n-1, m-1, d_2} \leq \frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{s_x^2}{s_y^2}) \leq P_{n-1, m-1, 1-\alpha}(\theta) = 1-\alpha, \forall \theta \in \Theta.$$

$$\Rightarrow P(F_{n-1, m-1, d_2} \cdot \frac{s_y^2}{s_x^2} \leq \frac{\sigma_2^2}{\sigma_1^2} \leq F_{n-1, m-1, 1-\alpha} \cdot \frac{s_y^2}{s_x^2} | \theta) = 1-\alpha, \forall \theta \in \Theta.$$

$A(X, Y)$ $B(X, Y)$

Logo, $[A(X, Y), B(X, Y)]$ é um I.C. para $\frac{\sigma_2^2}{\sigma_1^2}$ com coeficiente de confiança $1-\alpha$.

Ex. 5: X_1, \dots, X_n , dada θ , c.c.d $N(0|\theta)$.

Vimos que $f_{X(n)}(t|\theta) = \frac{n!}{\theta^n} \cdot \frac{t^{n-1}}{(t/\theta)^n} \cdot e^{-t/\theta}$

$$(F_{X(n)}(t); ?)$$

$$\frac{X_{(n)}|\theta}{\theta} \sim \text{Beta}(n, 1)$$

/ /

$$F_{\frac{X_m}{d/2}}(t) = t^n, \quad t \in (0,1)$$

$$t^n = d/2 \Rightarrow t = (d/2)^{1/n}$$

$$t^n = 1 - \frac{d}{2} \Rightarrow t = \left(1 - \frac{d}{2}\right)^{1/n}$$

$$\mathbb{P}\left(\left(\frac{d}{2}\right)^{1/n} \leq \underline{X}_{(n)} \leq \left(1 - \frac{d}{2}\right)^{1/n} \mid \Theta\right) = 1 - d, \forall \theta > 0.$$

$$\Rightarrow \mathbb{P}\left(\frac{1}{\left(1 - \frac{d}{2}\right)^{1/n}} \leq \Theta \leq \frac{1}{\left(\frac{d}{2}\right)^{1/n}} \mid \Theta\right) = 1 - d, \forall \theta > 0.$$

$$\Rightarrow \mathbb{P}\left(\frac{\underline{X}_{(n)}}{\left(1 - \frac{d}{2}\right)^{1/n}} \leq \Theta \leq \frac{\overline{X}_{(n)}}{\left(\frac{d}{2}\right)^{1/n}} \mid \Theta\right) = 1 - d, \forall \theta > 0.$$

$\underline{A}(x) \qquad \qquad \qquad \overline{B}(x)$

Ex. 6: X_1, \dots, X_n dada Θ , com $\mathcal{U}\left(\Theta - \frac{1}{2}, \Theta + \frac{1}{2}\right)$. Seja $\pm C =$

$[\underline{X}_{(1)}, \overline{X}_{(n)}]$. Encontre d .

$$\mathbb{P}(\underline{X}_{(1)} \leq \Theta \leq \overline{X}_{(n)} \mid \Theta)$$



$$1 - \mathbb{P}(\underline{X}_{(1)} > \Theta) - \mathbb{P}(\overline{X}_{(n)} < \Theta)$$

$$= 1 - \mathbb{P}\left(\bigcap_{i=1}^n \{X_i > \Theta\} \mid \Theta\right) - \mathbb{P}\left(\bigcap_{i=1}^n \{X_i < \Theta\} \mid \Theta\right)$$

$$= 1 - \left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^n = 1 - \left(\frac{1}{2}\right)^{n-1}.$$

Logo, $[\underline{X}_{(1)}, \overline{X}_{(n)}]$ é um $\pm C$ para Θ com coeficientes de confiança

$$Y = 1 - \left(\frac{1}{2}\right)^{n-1}.$$

Exercício: Verificar que $[X_{(2)}, X_{(n)}]$ é t.c.p. θ ?

$$n=3 \quad X_1 = 5,13$$

$$X_2 = 5,001$$

$$X_3 = 5,999$$

$[5,001; 5,999]$ é estimativa por intervalo para θ com coeficiente de confiança $1 - \left(\frac{1}{2}\right)^{3-1} = \frac{3}{4}$

$$\begin{aligned} \theta - \frac{1}{2} &\leq 5,001 \\ 5,999 &\leq \theta + \frac{1}{2} \end{aligned} \Rightarrow \theta \in [5,499; 5,501]$$

$$[(-\infty, B(x))] \quad [\text{limite superior}]$$

DEF.: $[A(x), \infty)$ é um limite inferior de confiança com coeficiente de confiança $1-\alpha$, $0 < \alpha < 1$, se

$$P(A(x) \leq \theta < \infty | \theta) \geq 1-\alpha, \forall \theta \in \Theta.$$

$$(P(-\infty < \theta \leq B(x) | \theta) \geq 1-\alpha, \forall \theta \in \Theta)$$

Ex 1: X_1, \dots, X_n , dado θ , c.i.d $N(\theta, \sigma_0^2)$ ^{conhecida!}

$$\frac{\bar{X} - \theta}{\sigma_0 / \sqrt{n}} \mid \theta \sim N(0, 1)$$

$$P\left(\frac{\bar{X} - \theta}{\sigma_0 / \sqrt{n}} > \Phi^{-1}(\alpha) \mid \theta\right) = 1-\alpha, \forall \theta \in \Theta.$$

$$\Rightarrow P\left(\frac{\theta - \bar{X}}{\sigma_0 / \sqrt{n}} < -\Phi^{-1}(\alpha) \mid \theta\right) = 1-\alpha, \forall \theta \in \Theta.$$

$$\Rightarrow P\left(\theta < \bar{X} - \Phi^{-1}(\alpha) \frac{\sigma_0}{\sqrt{n}} \mid \theta\right) = 1-\alpha, \forall \theta \in \Theta.$$

Assim,

$(-\infty, \bar{X} - \Phi^{-1}(1-\alpha) \frac{\sigma_0}{\sqrt{n}}]$ é um limite superior de confiança p/ θ

com confiança $1-\alpha$.

06/06

Aula 12

$$\Theta \subseteq \mathbb{R}^k$$

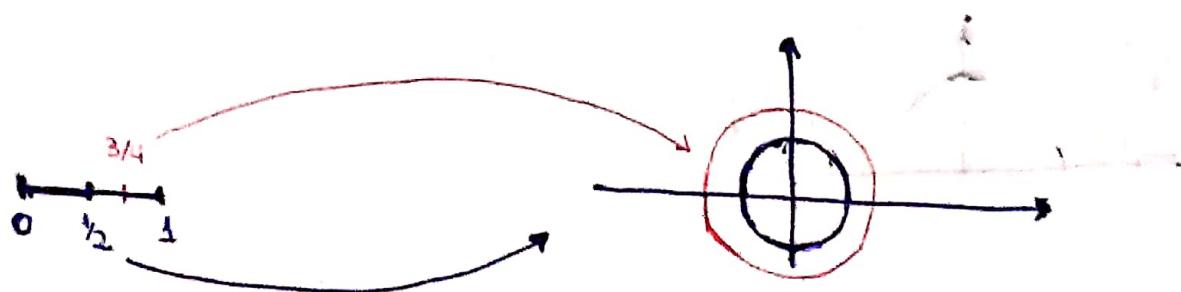
Definição:

Seja A um conjunto de subconjuntos de \mathbb{R}^n .

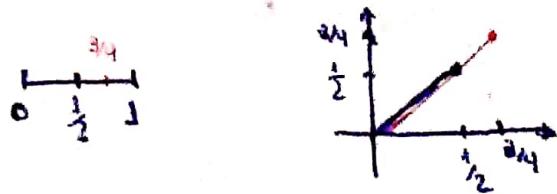
Definição: A transformação $R: \mathcal{X} \rightarrow A$ que associa a cada $x \in \mathcal{X}$ uma região (ou subconjunto) $R(x) \in A$, damos o nome de Região (conjunto) aleatório.

Exemplo: $X \sim U(0,1)$, $A = \mathcal{B}(\mathbb{R}^2)$

$$R(x) = \{(a,b) \in \mathbb{R}^2 : a^2 + b^2 \leq x^2\}$$



$$R_2(x) = \{\lambda(0,0) + (1-\lambda)(x,x) : \lambda \in [0,1]\}$$



$$g(\Theta) \subseteq \mathbb{R}^k$$

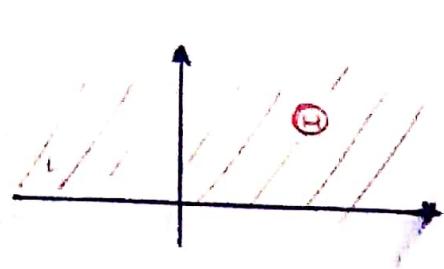
Definição: A região (conjunto) aleatório (Θ) R é uma região (conjunto) de confiança para $g(\Theta)$ com coeficiente de confiança $1-\alpha$, $0 < \alpha < 1$; se

$$P(\Theta \in R(x)|\theta) \geq 1-\alpha, \forall \theta \in \Theta$$

Exemplo 1. X_1, \dots, X_n dado $\Theta = (\theta_1, \theta_2)$, c.i.i.d. $N(\theta_1, \theta_2)$

$$\Theta = (\theta_1, \theta_2)$$

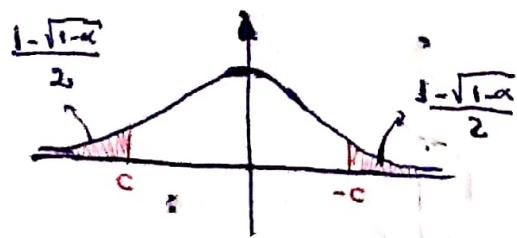
$$\Theta = \mathbb{R} \times \mathbb{R}_+$$



$$\left. \begin{aligned} & \frac{\bar{X} - \theta_1}{\sqrt{\theta_2/n}} \mid \theta \sim N(0,1) \\ & \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\theta_2} \mid \theta \sim \chi^2_{n-1} \end{aligned} \right\} \text{independentes (dado } \Theta \text{)}$$

Para $\alpha \in (0,1)$

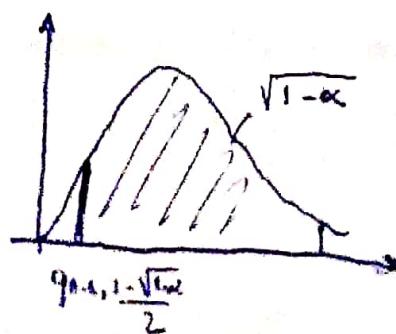
$$P\left(\Phi^{-1}\left(\frac{1-\sqrt{1-\alpha}}{2}\right) \leq \frac{\bar{X} - \theta_1}{\sqrt{\theta_2/n}} \leq -c \mid \theta\right) = \sqrt{1-\alpha}$$



Além disso,

$$P\left(q_{n-2, 1 - \frac{1-\sqrt{1-\alpha}}{2}} \leq \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\theta_2} \leq q_{n-2, 2 - \left(1 - \frac{1-\sqrt{1-\alpha}}{2}\right)} \mid \theta\right) =$$

$= \sqrt{1-\alpha}$, onde $q_{n-2, p}$ é o percentil de ordem p da χ^2_{n-2} .



Assim,

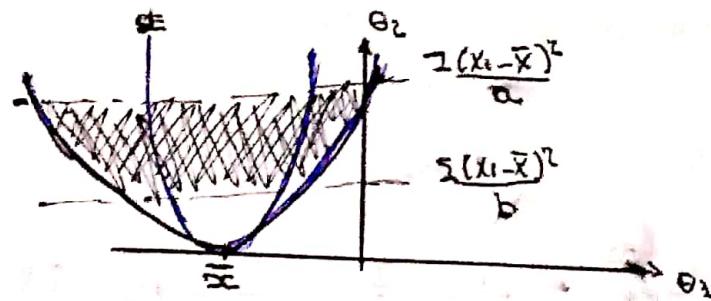
$$P\left(c \leq \frac{\bar{x} - \theta_2}{\sqrt{\theta_2/n}} \leq -c, a \leq \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\theta_2} \leq b \mid \theta\right) =$$

$$P\left(c \leq \frac{\bar{x} - \theta_2}{\sqrt{\theta_2/n}} \leq -c \mid \theta\right) P\left(a \leq \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\theta_2} \leq b \mid \theta\right) =$$

$$= \sqrt{1-\alpha} \cdot \sqrt{1-\alpha} = 1-\alpha, \forall \theta \in \Theta$$

$$\Rightarrow P\left(|\bar{x} - \theta_2| \leq |c|\sqrt{\theta_2/n}, \frac{\sum (x_i - \bar{x})^2}{b} \leq \theta_2 \leq \frac{\sum (x_i - \bar{x})^2}{a} \mid \theta\right) = 1-\alpha, \forall \theta \in \Theta$$

$$= P\left(\frac{\sum (x_i - \bar{x})^2}{b} \leq \theta_2 \leq \frac{\sum (x_i - \bar{x})^2}{a}, \theta_2 \geq \frac{(\theta_2 - \bar{x})^2}{c^2} \mid \theta\right) = 1-\alpha, \forall \theta \in \Theta$$



Intervalos Aproximados

Exemplo: X_1, \dots, X_n , dado θ , c.i.i.d. $Ber(\theta)$

$$\sqrt{n} (\bar{x} - \theta) \xrightarrow{d} N(0, \theta(1-\theta))$$

$$\sqrt{\bar{x}(1-\bar{x})} \xrightarrow{D} \sqrt{\theta(1-\theta)}$$

$$\frac{\sqrt{n}(\bar{X} - \theta)}{\sqrt{\bar{X}(1-\bar{X})}} \xrightarrow{d} \frac{1}{\sqrt{\theta(1-\theta)}} N(0, \theta(1-\theta)) \stackrel{d}{=} N(0, 1)$$

Para n grande

$$P\left(\Phi^{-1}\left(\frac{\alpha}{2}\right) \leq \frac{\sqrt{n}(\bar{X} - \theta)}{\sqrt{\bar{X}(1-\bar{X})}} \leq \Phi^{-1}(1-\alpha/2) \mid \theta\right) \approx 1-\alpha, \forall \theta \in \Theta.$$

$$P\left(\bar{X} - \Phi^{-1}(1-\frac{\alpha}{2}) \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \leq \theta \leq \bar{X} + \Phi^{-1}(\alpha/2) \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \mid \theta\right) \approx 1-\alpha, \forall \theta.$$

$A(\bar{X})$ $B(\bar{X})$

Logo, $[A(\bar{X}), B(\bar{X})]$ é um intervalo de confiança aproximado para θ de confiança $1-\alpha$.

Exemplo 2: X_1, \dots, X_n , dado $\theta = (\theta_1, \theta_2)$, c.i.i.d. $N(\theta_1, \theta_2)$,

$$\sqrt{n}(\bar{X} - \theta_1) / \theta_2 \sim N(0, 1)$$

$$\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1} = \frac{n}{n-1} \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n} = \frac{n}{n-1} \frac{\sum X_i^2 - n\bar{X}^2}{n} = \frac{n}{n-1} \left\{ \frac{\sum X_i^2}{n} - \bar{X}^2 \right\} \xrightarrow{P_0} \theta_2$$

$\downarrow P_0$ $\downarrow P_0$
 $E(X_i)$ $\theta_1^2 + \theta_2^2$

$$S_n^2 \xrightarrow{P_0} \theta_2$$

$$\sqrt{n}(\bar{X}_n - \theta_1) \xrightarrow{d} N(0, \theta_2)$$

$$\frac{\sqrt{n}(\bar{X}_n - \theta_1)}{\sqrt{s_n^2}} \xrightarrow{d} \frac{1}{\sqrt{\theta_2}} N(0, \theta_2) \stackrel{d}{=} N(0, 1)$$

$$P(\Phi^{-1}(\alpha/2) \leq \frac{\bar{X}_n - \theta}{S_n} \leq \Phi^{-1}(1-\frac{\alpha}{2}) | \theta) \approx 1-\alpha \quad \forall \theta \Rightarrow$$

$$\Rightarrow P\left(\bar{X}_n - \Phi^{-1}(1-\frac{\alpha}{2}) \frac{S_n}{\sqrt{n}} \leq \theta_1 \leq \bar{X}_n - \Phi^{-1}(\alpha/2) S_n / \sqrt{n} \mid \theta\right) \approx 1-\alpha$$

$A(x)$ $B(x)$

Exemplo 3. X_1, \dots, X_n , dado θ , c.i.i.d. Poisson(θ)

$$g(\theta) = e^{-\theta}$$

* J.C.(B) a Chang Yu Fang
Diss. Mestrado (2008)
Dr. Carlos Afonso

\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,

$$\sqrt{n}(e^{-\bar{X}_n} - e^{-\theta}) \xrightarrow{d} N(0, (-e^{-\theta})^2 \cdot \theta) \stackrel{d}{=} N(0, \theta e^{-2\theta})$$

$$\sqrt{n}(e^{-\bar{X}_n} - e^{-\theta}) \xrightarrow{d} N(0, \theta e^{-2\theta})$$

$$\bar{X}_n \xrightarrow{P_\theta} \theta$$

$$e^{-2\bar{X}_n} \xrightarrow{P_\theta} e^{-2\theta}$$

$$\bar{X}_n \cdot e^{-2\bar{X}_n} \xrightarrow{P_\theta} \theta e^{-2\theta}$$

$$\sqrt{\bar{X}_n \cdot e^{-2\bar{X}_n}} \xrightarrow{P_\theta} \sqrt{\theta e^{-2\theta}}$$

Assim,

$$\frac{\sqrt{n}(e^{-\bar{X}_n} - e^{-\theta})}{\sqrt{\bar{X}_n \cdot e^{-2\bar{X}_n}}} \xrightarrow{d} \frac{1}{\sqrt{\theta e^{-2\theta}}} N(0, \theta e^{-2\theta}) \stackrel{d}{=} N(0, 1)$$

para n grande

$$P\left(\Phi^{-1}(\alpha/2) \leq \frac{\sqrt{n}(e^{-\bar{x}_n} - e^{-\theta})}{\sqrt{\bar{x}_n e^{-2\bar{x}_n}}} \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \mid \theta\right) \approx 1 - \alpha, \forall \theta \in \Theta.$$

$$\Rightarrow P\left(\underbrace{e^{-\bar{x}_n} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}_{A(x)}, \underbrace{\sqrt{\frac{\bar{x}_n e^{-2\bar{x}_n}}{n}}}_{B(x)} \leq e^{-\theta} \leq e^{-\bar{x}_n} - \Phi^{-1}(\alpha/2), \sqrt{\frac{\bar{x}_n e^{-2\bar{x}_n}}{n}} \mid \theta\right) \approx$$

$$\approx 1 - \alpha, \forall \theta \in \Theta.$$

x_1, \dots, x_n , dado θ c.i.i.d.

No caso geral (sob condições de regularidade)

$$\sqrt{n}(\delta_{MV} - \theta) \xrightarrow{d} N\left(0, \frac{1}{I_{x_L}(\theta)}\right)$$

Se g é diferenciável:

$$\sqrt{n}(g(\delta_{MV}) - g(\theta)) \xrightarrow{d} N\left(0, (g'(\theta))^2 \cdot \frac{1}{I_{x_L}(\theta)}\right), \text{ pelo método delta}$$

$$\delta_{MV} \xrightarrow{P_\theta} \theta$$

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 $I_{x_L}(\cdot)$ é contínua,

$$I_{x_L}(\delta_{MV}) \xrightarrow{P_\theta} I_{x_L}(\theta)$$

Logo

$$\frac{(g'(\delta_{MV}))^2}{I_{x_L}(\delta_{MV})} \xrightarrow{P_\theta} \frac{(g'(\theta))^2}{I_{x_L}(\theta)}$$

Assim,

$$\frac{\ln(g(\hat{\theta}_{\text{MV}}) - g(\theta))}{\sqrt{\frac{(g'(\hat{\theta}_{\text{MV}}))^2}{J_{X_L}(\hat{\theta}_{\text{MV}})}}} \xrightarrow{d} \frac{1}{\sqrt{\frac{(g'(\theta))^2}{J_{X_L}(\theta)}}} N(0, \frac{(g'(\theta))^2}{J_{X_L}(\theta)}) \stackrel{d}{=} N(0, 1)$$

Para "n grande"

$$P\left(\Phi^{-1}\left(\frac{\alpha}{2}\right) \leq \frac{\ln(g(\hat{\theta}_{\text{MV}}) - g(\theta))}{\sqrt{\frac{(g'(\hat{\theta}_{\text{MV}}))^2}{J_{X_L}(\hat{\theta}_{\text{MV}})}}} \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \mid \theta\right) \approx 1 - \alpha, \forall \theta \in \mathbb{R}$$

Intervalo de probabilidade (ou de credibilidade)

Definição: $[A(x), B(x)]$ é um intervalo de probabilidade (ou de credibilidade) para $\theta = g(\theta)$ com probabilidade $1 - \alpha$, $0 < \alpha < 1$, se

$$P(A(x) \leq \theta \leq B(x) \mid X = x) \geq 1 - \alpha, \forall x \in \mathcal{X}$$

Exemplo: x_1, \dots, x_n , dado θ , c.i.i.d. $N(\theta, \sigma^2_\theta)$

Vimos que

$$\sigma^2_n = \sigma^2_n(a^2, b)$$

$$\theta \mid x_1 = x_1, \dots, x_n = \text{se} \underbrace{N(\bar{x}_n + (1-\alpha)\frac{\sigma^2_n}{a^2}, \sigma^2_n)}_{\text{contendo!}}$$

$$\frac{\theta - ax}{\sqrt{\sigma^2_n}} \mid X = x \sim N(0, 1)$$

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

$$P(\Phi^{-1}(\alpha_2) \leq \frac{\theta - \alpha}{\sqrt{\sigma_n^2}} \leq \Phi^{-1}(1-\alpha_2) \mid X=\alpha) = 1-\alpha \Rightarrow$$

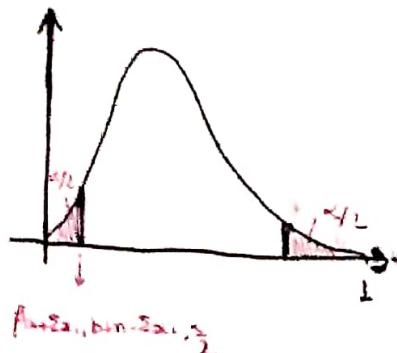
⇒

$$P(\alpha + \Phi^{-1}(\alpha_2)\sqrt{\sigma_n^2} \leq \theta \leq \alpha + \Phi^{-1}(1-\alpha_2)\sqrt{\sigma_n^2} \mid X=\alpha) = 1-\alpha.$$

Exemplo: X_1, \dots, X_n , dado Θ , c.i.i.d. $Ber(\theta)$

$\Theta \sim Beta(a, b)$

$$\Theta \mid X_1=x_1, \dots, X_n=x_n \sim Beta(a + \sum_{i=1}^n x_i, b+n - \sum_{i=1}^n x_i)$$



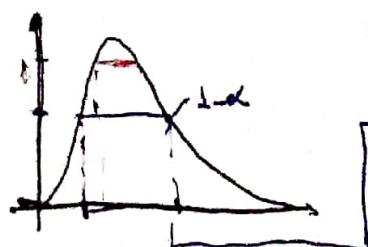
$$P\left(\beta_{a+\sum x_i, b+n-\sum x_i, \frac{1}{2}} \leq \theta \leq \beta_{a+\sum x_i, b+n-\sum x_i, 1-\alpha/2} \mid X=\alpha\right) = 1-\alpha$$

onde $\beta_{a,b,p}$ é o percentil de ordem p da Beta(a, b)

$$\left[\beta_{a+\sum x_i, b+n-\sum x_i, \alpha/2}, \beta_{a+\sum x_i, b+n-\sum x_i, 1-\alpha/2} \right] \text{ é um intervalo de credibilidade}$$

Probabilidade de $1-\alpha$.

Ideia: Elevador



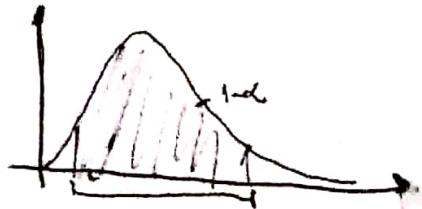
$$1-\alpha = 95\%$$

→ Highest Posterior Density Interval
Intervalo HPD.

Exemplo 3. X_1, \dots, X_n dado θ c.i.i.d. Poisson(θ)

$\Theta \sim \text{Gama}(a, b)$

$\Theta | X_1 = x_1, \dots, X_n = x_n \sim \text{Gama}\left(a + \sum_{i=1}^n x_i, b + n\right)$

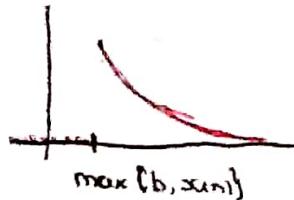


Exemplo 4. (Exercício),

X_1, \dots, X_n dado θ , c.i.i.d. $U(0, \theta)$

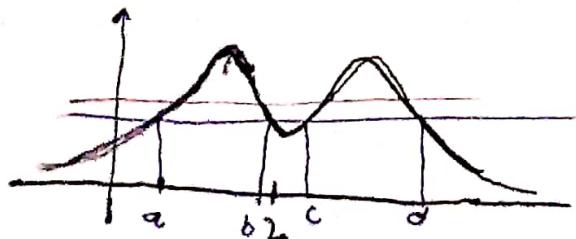
$\Theta \sim \text{Pareto}(a, b)$

$\Theta | X_1 = x_1, \dots, X_n = x_n \sim \text{Pareto}(a+n, \max\{b, x_{(n)}\})$



$$\Theta \sim \frac{1}{2\theta} N(0, 1) + \frac{1}{2\theta} N(4, 1)$$

$$f(\theta) = \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} + \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-(\theta-4)^2/2}$$



$(a, b) \cup (c, d)$

$$X \mid \theta \sim N(\theta, 1)$$

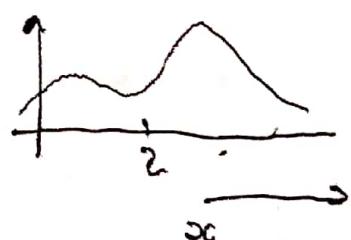
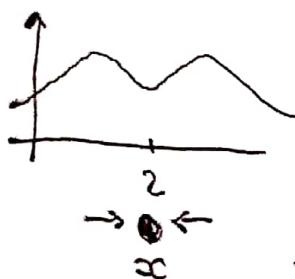
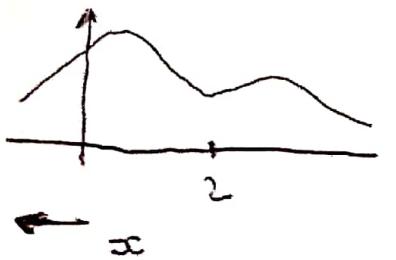
$$f(\theta|x) \propto f(x|\theta) f(\theta)$$

$$\propto \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} \left[\frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} + \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta-4)^2}{2}} \right]$$

$$\propto e^{-\frac{\theta^2 + (x-\theta)^2}{2}} + e^{-\frac{(\theta-4)^2 + (x-\theta)^2}{2}}$$

$\Rightarrow \dots$

$$\Rightarrow f(\theta|x) = \frac{e^{-x^2/4}}{e^{-x^2/4} + e^{-\frac{(x-4)^2}{4}}} \cdot \frac{1}{\sqrt{2\pi} \cdot \frac{1}{2}} e^{-\frac{(\theta-2)^2}{2 \cdot \frac{1}{2}}} + \frac{e^{-\frac{(x-4)^2}{4}}}{e^{-\frac{x^2}{4}} + e^{-\frac{(x-4)^2}{4}}} \cdot \frac{1}{\sqrt{2\pi} \cdot \frac{1}{2}} e^{-\frac{(\theta-4)^2}{2 \cdot \frac{1}{2}}}$$



Fenômeno: $\prod_{R^c} f(\theta) + \text{Complementar}(R) \rightarrow \text{otimizando termos o A.P.D.}$

Aula 13

07/06 -

"Statistical Inference: A Bayesian Perspective". Rafael B. Stern, Carlos A. B. Pereira

Dissertação Chang Yu Fang (2008)

Garthwaite, Jolliffe e Jones (IC para Poisson(θ)) Pag. 131.

Lista 7

Teste de Hipótese

④

$\Theta_0, \Theta_1 \subseteq \Theta$ tais que $\Theta = \Theta_0 \cup \Theta_1$ e $\Theta_0 \cap \Theta_1 = \emptyset$.

$(\theta_1, \theta_2, \dots, \theta_n)$

$$\theta_i > \frac{1}{2} \quad \bigcup_{i=1}^n \{\theta_i > \frac{1}{2}\}$$

$$\leq \frac{1}{2} \quad \bigcap_{i=1}^n \{\theta_i \leq \frac{1}{2}\}$$

Teste \rightarrow problema de decisão com duas alternativas

$$H_0: \theta \in \Theta_0$$

$$H_1: \theta \in \Theta_1$$

$$H_0: \theta = \theta_0$$

Exemplo:

B brancas

$\Theta = \{0, 1, 2, 3, 4, 5\}$ $X = (X_1, X_2)$, onde $X_i = \begin{cases} 1, & \text{a } i\text{-ésima bola retirada é branca} \\ 0, & \text{c.c.} \end{cases}$

C-verdes

$$\Theta_0 = \{3, 4, 5\}$$

$$\Theta_1 = \{0, 1, 2\}$$

$$i=1, 2$$

$$\mathcal{X} = \{(0,0), (0,1), (1,0), (1,1)\} = \{0,1\}^2$$

$$H_0: \theta \in \Theta_0 (\theta \geq 3)$$

$$H_1: \theta \in \Theta_1 (\theta < 3)$$

Definição: Uma função de teste (teste) é uma função $\psi: \mathcal{X} \rightarrow \{0,1\}$ para testar $H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_1$, que associa a cada $x \in \mathcal{X}$ $\psi(x)=0$ no caso de decisão* por H_0 ou partir de x ou $\psi(x)=1$ no caso de decisão** por H_1 .

$H_0: \theta \in \Theta_0 \rightarrow$ Hipótese Nula

$H_1: \theta \in \Theta_1 \rightarrow$ Hipótese Alternativa

* aceitação H_0 / Não rejeição de H_0

** rejeição de H_0

$$\psi(x_1, x_2) = \begin{cases} 0, & x_1 + x_2 \leq 1 \\ 1, & x_1 + x_2 > 1 \end{cases}$$

$$\psi_2(x_1, x_2) = \begin{cases} 0, & x_1 + x_2 \leq 2 \\ 1, & x_1 + x_2 > 2 \end{cases}$$

$$\psi_3(x_1, x_2) = \begin{cases} 0, & x_1 + x_2 = 0 \\ 1, & x_1 + x_2 \neq 0 \end{cases}$$

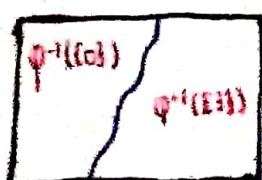
$$\psi_4(x_1, x_2) = 0, \forall x_1, x_2$$

Nesse caso, temos $2^4 (2^{|\mathcal{X}|})$ testes.

Ao conjunto $\psi^{-1}(1) = \{x \in \mathcal{X}: \psi(x)=1\}$ damos o nome de Região Crítica ou Região de Rejeição (de H_0)

" " " $\psi^{-1}(0) = \{x \in \mathcal{X}: \psi(x)=0\}$ " " " Região de Aceitação de H_0

$$\Psi = \{\psi: \mathcal{X} \rightarrow \{0,1\} \text{ tal que } \psi \text{ é mensurável}\}$$



Como escolher $\psi \in \Psi$ de maneira "ótima"?

Num problema de decisão com duas alternativas, você pode incorrer em dois possíveis erros:

(1) Você decidir por H_2 ($\psi(x)=3$) quando H_0 é verdadeira.

(rejeitar H_0) \rightarrow erro de tipo I

(2) Você decidir por H_0 ($\psi(x)=0$) quando H_2 é verdadeira.

(não rejeitar H_0) \rightarrow erro de tipo II

Suponhamos

$$\Theta = \{2, 4\} \quad \Theta_0 = \{2\} \quad \Theta_1 = \{4\}$$

$$H_0: \theta = 2 \quad \psi_1(x_1, x_2) = \sum_{i=1,2} (x_i + x_2)$$

$$H_1: \theta = 4$$

$$\psi_2(x_1, x_2) = \sum_{i=1,2} (x_1 + x_i)$$

$$\psi_3(x) = 0, \forall x \in \mathbb{R}$$

Assim, para o teste ψ_1

$$P(\text{Erro tipo I}) = P(\text{Rejeitar } H_0 \text{ quando } H_0 \text{ é verdadeira}) =$$

$$= P(\psi_1(x) = 1 | H_0 \text{ é verdadeira}) = P(\psi_1(x) = 1 | \theta = 2) =$$

$$= P(X_1 + X_2 = 2 | \theta = 2) = \frac{\binom{2}{2} \binom{3}{0}}{\binom{5}{2}} = \frac{1}{10} \rightarrow \text{Aqui temos A PROBABILIDADE DE ERRO DO TIPO I.}$$

$$X_1 + X_2 / 2 \sim N(0, 5, 0, 2)$$

e

$$P(\text{Erro tipo II}) = P(\text{Não rejeitar } H_0 | H_1 \text{ é falsa}) =$$

$$= P(\psi_1(x) = 0 | \theta = 4) = P(X_1 + X_2 \leq 1 | \theta = 4) = 1 - P(X_1 + X_2 = 2 | \theta = 4) = 1 - \frac{\binom{2}{2} \binom{3}{0}}{\binom{5}{2}} = \frac{4}{10}$$

Teste φ_2 :

$$\begin{aligned} P(\text{Erro tipo I}) &= P(\text{Rejeitar } H_0 \mid H_0 \text{ é verdadeira}) = \\ &= P(\varphi_2(X) = 1 \mid \theta = 2) = P(X_1 + X_2 \geq 1 \mid \theta = 2) = \\ &= 1 - P(X_1 + X_2 = 0 \mid \theta = 2) = 1 - \frac{\binom{2}{0}\binom{3}{2}}{\binom{5}{2}} = \frac{7}{10} \end{aligned}$$

$P(\text{Erro tipo II}) = P(\text{Não rejeitar } H_0 \mid H_0 \text{ é falsa})$:

$$= P(\varphi_2(X) = 0 \mid \theta = 4) = P(X_1 + X_2 = 0 \mid \theta = 4) = \frac{\binom{4}{0}\binom{1}{2}}{\binom{5}{2}} = 0$$

Teste φ_3 :

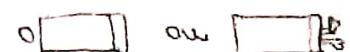
$$P(\text{Erro tipo I}) = P(\varphi_3(X) = 1 \mid \theta = 2) = 0$$

$$P(\text{Erro tipo II}) = P(\varphi_3(X) = 0 \mid \theta = 4) = 1$$

Logo, temos

Teste	$P(\text{Erro I})$	$P(\text{Erro II})$
φ_1	$3/10$	$4/10$
φ_2	$7/10$	0
φ_3	0	1

Analogia do cobertor curto
pequeno



Sobre a cabeça e não os pés ou vice-versa.

Não dá pra minimizar ambos!

Voltando ao problema original.

$$\Theta = \{0, 1, 2, 3, 4, 5\}$$

$$H_0: \theta \geq 3$$

$$H_1: \theta < 3$$

$$\Psi_1(x) = \mathbb{I}_{\{x_1+x_2\}} \quad \Psi_2(x) = \mathbb{I}_{\{x_1+x_2\leq 0\}}$$

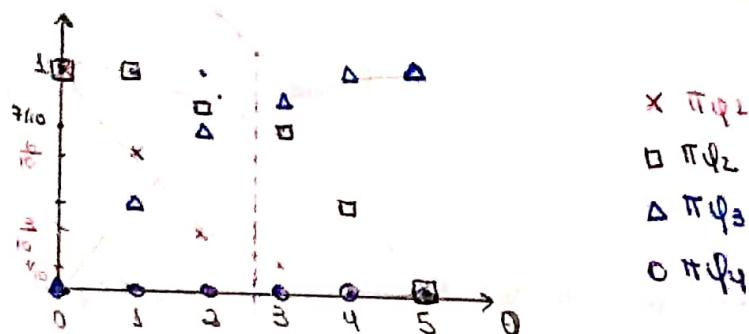
$$\Psi_3(x) = \mathbb{I}_{\{x_1+x_2>1\}} \quad \Psi_4(x) = 0, \forall x$$

Seja $\Psi: \mathcal{X} \rightarrow \{0,1\}$ um teste para $H_0: \theta \in \Theta_0$. O número

$$\pi_{\Psi}(\theta) = P(\Psi(X)=1|\theta)$$

damos o nome de FUNÇÃO PODER DO TESTE Ψ .

($\beta(\theta) = 1 - \pi_{\Psi}(\theta) \Rightarrow$ curva característica de operações).



$$\pi_{\Psi_1}(\theta) = P(\Psi_1(X)=1|\theta) = P(X_1+X_2=0|\theta) = \frac{\binom{\theta}{0} \binom{5-\theta}{2}}{\binom{5}{2}} = \frac{(5-\theta)(4-\theta)}{20}$$

$$\pi_{\Psi_2}(\theta) = P(\Psi_2(X)=1|\theta) = P(X_1+X_2 \leq 0|\theta) = 1 - P(X_1+X_2 > 0|\theta) = 1 - \frac{\binom{\theta}{1} \binom{5-\theta}{1}}{\binom{5}{2}} = 1 - \frac{\theta(\theta-1)}{20}$$

$$\pi_{\Psi_3}(\theta) = P(\Psi_3(X)=1|\theta) = P(X_1+X_2 \geq 1|\theta) = 1 - P(X_1+X_2=0|\theta) = 1 - \frac{(5-\theta)(4-\theta)}{20}$$

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 $|\Theta_i| \geq 1$.

Exemplo 3.

$H_0: \theta \geq 3 \rightarrow$ composta

$H_0: \theta \leq 3 \rightarrow$ composta

$H_0: \theta = 2 \rightarrow$ simples

$H_0: \theta \in \{2,4\} \rightarrow$ composta

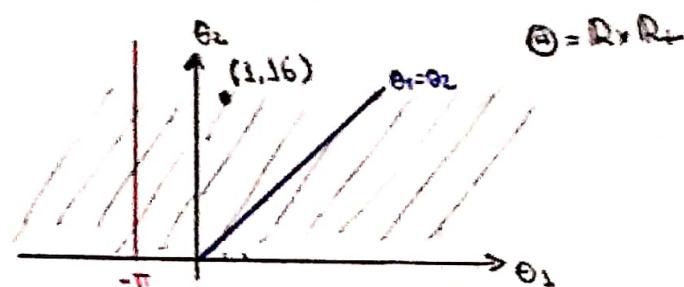
Exemplo 2.

$$X| \theta = (\theta_1, \theta_2) \sim N(\theta_1, \theta_2)$$

$H_0: \theta = (1, 16) \rightarrow$ simples

$H_1: \theta_1 = -\pi \rightarrow$ composta

$H_2: \theta_1 = \theta_2 \rightarrow$ composta



Seja $\psi: \mathcal{E} \rightarrow \{0,1\}$ um teste para $H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_1$.

Ao número $\sup_{\theta \in \Theta_0} \pi_\psi(\theta)$ damos o nome de Tamanho do Teste ψ .

Notação.

$$\alpha_{\psi_L}^* = \sup_{\theta \in \Theta_0} \pi_{\psi_L}(\theta)$$

No exemplo da urna:

$$\alpha_{\psi_L}^* = \sup_{\theta \in \Theta_0} \pi_{\psi_L}(\theta) = \sup_{\theta \in \{3,4,5\}} \pi_{\psi_L}(\theta) = \frac{1}{10}$$

$$\alpha_{\psi_R}^* = \sup_{\theta \in \Theta_0} \pi_{\psi_R}(\theta) = \frac{1}{10}$$

$$\alpha_{\psi_H}^* = \sup_{\theta \in \Theta_0} \pi_{\psi_H}(\theta) = 1$$

Analogamente, $\alpha_{\psi_H}^* = 0$.

Definición: Un teste de α es un nivel α , $0 < \alpha \leq 1$, se

$$\alpha_q \leq \alpha.$$

Significancia (sign.)

Teste Para Hipóteses Simples - Aula 14

$$\Theta = \{\theta_0, \theta_1\} \quad \Theta_0 = \{\theta_0\}, \quad \Theta_1 = \{\theta_1\}$$

$$H_0: \theta = \theta_0 \quad (\theta \in \Theta_0)$$

$$H_1: \theta = \theta_1 \quad (\theta \in \Theta_1)$$

$\Psi = \{ \psi: \mathcal{X} \rightarrow \{0,1\} \}$ onde $\psi \in \Psi$ é um teste para H_0 versus H_1 .

Para $\psi \in \Psi$,

$$\pi_{\psi}(\theta_0) = P(\psi(x) = 1 | \theta_0), \text{ prob. erro tipo I}$$

$$1 - \pi_{\psi}(\theta_1) = P(\psi(x) = 0 | \theta_1), \text{ prob. erro tipo II.}$$

Resultado:

Seja $\psi^*: \mathcal{X} \rightarrow \{0,1\}$ dada por

$$\psi^*(x) = \begin{cases} 1, & a\psi(x|\theta_0) < b\psi(x|\theta_1) \\ 0, & \text{c.c.} \end{cases} \quad \left(\frac{f(x|\theta_1)}{f(x|\theta_0)} > \frac{a}{b} \right)$$

$a, b > 0$ constantes.

Então, $\forall \psi \in \Psi$

$$a\pi_{\psi^*}(\theta_0) + b(1 - \pi_{\psi^*}(\theta_1)) \leq a\pi_{\psi}(\theta_0) + b(1 - \pi_{\psi}(\theta_1))$$

Para $\psi \in \Psi$,

$$a\pi_{\psi}(\theta_0) + b(1 - \pi_{\psi}(\theta_1)) = aP(\psi(x) = 1 | \theta_0) + b(1 - P(\psi(x) = 1 | \theta_1)) =$$

$$= a \int_{\psi^{-1}(1)} f(x|\theta_0) dx + b \left(1 - \int_{\psi^{-1}(1)} f(x|\theta_1) dx \right) =$$

$$= \int_{\psi^{-1}(\{1\})} a f(x/\theta_0) dx + b - \int_{\psi^{-1}(\{1\})} b f(x/\theta_1) dx =$$

$$= b + \int_{\psi^{-1}(\{1\})} [a f(x/\theta_0) - b f(x/\theta_1)] dx \geq b + \int_{\psi^{-1}(\{1\})} [a f(x/\theta_0) - b f(x/\theta_1)] dx$$

$$\Rightarrow a \pi_{\psi}(\theta_0) + b(1 - \pi_{\psi}(\theta_1)) \geq a \pi_{\varphi^*}(\theta_0) + b(1 - \pi_{\varphi^*}(\theta_1))$$

Exemplo:

$$\mathcal{X} = \{0,1\}^2$$

$$H_0: \theta = 2$$

$$H_1: \theta = 4$$

	$\alpha = 1$, $b = 2$			
x	(0,0)	(0,1)	(1,0)	(1,1)
$\psi(x/\theta=2)$	$3/10$	$3/10$	$3/10$	$4/10$
$\psi(x/\theta=4)$	0	$2/10$	$2/10$	$6/10$
$\psi(x/\theta=4)$	0	$2/3$	$2/3$	6
$\psi(x/\theta=2)$				

$$\psi^*(x) = \begin{cases} 1, & \frac{1}{\psi(x/\theta=4)} \rightarrow 1/2 \\ & \frac{1}{\psi(x/\theta=2)} \\ 0, & \text{c.c.} \end{cases}$$

Assim,

$$\psi^*(x) = \begin{cases} 1, & x_1 + x_2 \geq 1 \\ 0, & \text{G.C.} \end{cases} = \psi_2(x)$$

Fazendo

$$a' = b' = 1 \quad (\text{erros tipo I e II com mesmo peso})$$

L mais resistente à rejeitar

$$\psi^*(x) = \begin{cases} 1, & x_1 + x_2 = 2 \\ 0, & \text{o.c.} \end{cases} = \psi_L(x)$$

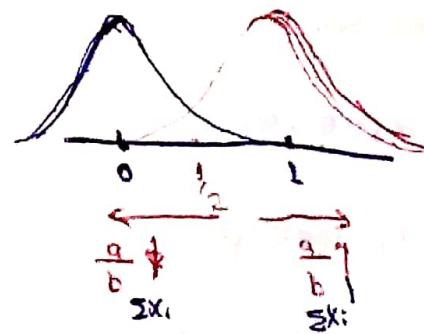
$$a'' = 7, \quad b'' = 1$$

$$f^*(x) = 0, \quad \forall x \in \mathbb{R}$$

Exemplo 2: x_1, \dots, x_n , dado θ , c.i.i.d. $N(\theta, 1)$.

$$H_0: \theta = 0 \quad a = 2, \quad b = 1$$

$$H_1: \theta = 1 \quad \mathbb{X} = \mathbb{R}^n$$



$$\frac{\psi(x|\theta=1)}{\psi(x|\theta=0)} = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-1)^2}{2}}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}} = \frac{e^{-\frac{\sum(x_i-1)^2}{2}}}{e^{-\frac{\sum x_i^2}{2}}} = e^{\frac{\sum x_i - n}{2}} = e^{\frac{\sum x_i - n}{2}}$$

$$\psi^*(x) = 1 \Leftrightarrow \frac{\psi(x|\theta=1)}{\psi(x|\theta=0)} > 2 \Leftrightarrow e^{\sum x_i - n/2} > 2 \Leftrightarrow$$

$$\Leftrightarrow \sum x_i - \frac{n}{2} > \log 2 \Leftrightarrow \bar{x}_n > \frac{1}{2} + \frac{\log 2}{n}.$$

$$\Psi^*(x) = \begin{cases} 1, & \bar{x}_n > 1/2 + \log^2/n \\ 0, & \text{c.c.} \end{cases}$$

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 , $0 \leq \alpha \leq 1$, ψ^* é um TESTE MAIS PODEROSO (MP) de nível α para $H_0: \theta \in \Theta_0$ versus $H_1: \theta = \theta_1$ se

$$\pi_{\psi^*}(\theta_1) \geq \pi_\psi(\theta_1), \quad \forall \psi \text{ de nível } \alpha.$$

Lema de Neyman-Pearson

$H_0: \theta = \theta_0$ versus $H_1: \theta = \theta_1$

seja $\psi^*: \mathcal{X} \rightarrow \{0, 1\}$ dada por

$$\psi^*(x) = \begin{cases} 1, & f(x|\theta_0) \leq K f(x|\theta_1), \quad K > 0 \\ 0, & \text{c.c.} \end{cases}$$

Então,

$$\pi_{\psi^*}(\theta_1) \geq \pi_\psi(\theta_1), \quad \forall \psi \in \Psi^* = \{\psi \in \Psi: \alpha_\psi \leq \pi_\psi(\theta_0)\}.$$

Isto é, nessas condições, ψ^* é MP de nível $\pi_{\psi^*}(\theta_0)$.

Então é, nessas condições,

$$\psi^* \in \text{MD de nível } \pi_{\psi^*}(\theta_0).$$

Verificação:

Pelo resultado anterior,

$$\left. \begin{array}{l} \forall \varphi \in \Psi : \pi_{\varphi}(\theta_0) + k(1 - \pi_{\varphi}(\theta_1)) \geq \pi_{\varphi^*}(\theta_0) + k(1 - \pi_{\varphi^*}(\theta_1)) \\ \forall \varphi \in \Psi^*, \quad \pi_{\varphi^*}(\theta_0) \geq \pi_{\varphi}(\theta_0) \end{array} \right\} \Rightarrow$$

$$\Rightarrow \forall \varphi \in \Psi^*, \quad k(1 - \pi_{\varphi}(\theta_1)) \geq k(1 - \pi_{\varphi^*}(\theta_1)) \Rightarrow$$

$$\Rightarrow \forall \varphi \in \Psi^*, \quad \pi_{\varphi^*}(\theta_1) \geq \pi_{\varphi}(\theta_1).$$

Definição: ψ^*

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. $N(\theta, 1)$

$$H_0: \theta = \theta_0$$

$$(\theta_0 < \theta_1)$$

$$H_1: \theta = \theta_1$$

$$\begin{aligned} \psi^*(x) = 1 &\Leftrightarrow \frac{f(x|\theta_1)}{f(x|\theta_0)} > \frac{1}{K} \Leftrightarrow \frac{\left(\frac{x}{\sqrt{2\pi}}\right)^n e^{-\frac{(x-\theta_1)^2}{2}}}{\left(\frac{x}{\sqrt{2\pi}}\right)^n e^{-\frac{(x-\theta_0)^2}{2}}} > K^{-1} \Leftrightarrow \\ &\Leftrightarrow \frac{e^{-\frac{(x-\theta_1)^2}{2}}}{e^{-\frac{(x-\theta_0)^2}{2}}} > K^{-1} \Leftrightarrow \end{aligned}$$

$$\frac{2x^2 - 2\theta_1 x + 2\theta_0^2 - 2x^2 + 2\theta_1 x - 2\theta_0^2}{2} > K^{-1}$$

$$\Leftrightarrow K^{-1}$$

$$e^{\frac{(x_i - \theta_0)^2}{2}} \geq 1 + K^{-1} e^{\frac{n(\theta_1^2 - n\theta_0^2)}{2}}$$

$$(x_i - \theta_0)^2 \geq c + \log K^{-1} e^{\frac{n(\theta_1^2 - n\theta_0^2)}{2}}$$

$$\sum_{i=1}^n x_i^2 \geq c, \text{ onde } c = \frac{\log K^{-1} e^{\frac{n(\theta_1^2 - n\theta_0^2)}{2}}}{\theta_1 - \theta_0}$$

Assim, escolhemos $c \in \mathbb{R}$ de modo $\Pi_{\psi^*}(\theta_0) = \infty$.

$$\Pi_{\psi^*}(\theta_0) = P(\psi^*(X) = 1 | \theta_0) = P\left(\sum_{i=1}^n x_i \geq c | \theta_0\right) =$$

$$= P\left(\frac{\sum x_i - n\theta_0}{\sqrt{n}} \geq \frac{c - n\theta_0}{\sqrt{n}} \mid \theta_0\right) = \infty \Rightarrow$$

$$\Rightarrow \frac{c - n\theta_0}{\sqrt{n}} = \Phi^{-1}(1-\alpha) \Rightarrow c = n\theta_0 + \sqrt{n}\Phi^{-1}(1-\alpha)$$

Assim,

$$\psi^*(x) = \begin{cases} 1, & \sum x_i \geq n\theta_0 + \sqrt{n}\Phi^{-1}(1-\alpha) \\ 0, & \text{c.c.} \end{cases}$$

* Em geral, fixa-se um valor ϵ máximo para os tamanhos das bolas. Escolhemos, se possível, ψ^* cujo tamanho, δ_{ψ^*} , atinge o valor ϵ .

Exemplo 2. X_1, \dots, X_n dada θ , i.i.d. $\text{Exp}(\theta)$

$$\text{No. } \theta > \theta_0 \quad (\theta_0 \in \Theta_1) \Leftrightarrow x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

$$\delta_{\psi^*} = \theta_0$$

$$\Psi^*(x) = 1 \Leftrightarrow \frac{f(x|\theta_1)}{f(x|\theta_0)} > 1 \Leftrightarrow \frac{\prod_{i=1}^n f(x_i|\theta_1)}{\prod_{i=1}^n f(x_i|\theta_0)} > k^{-1} \Leftrightarrow \left(\frac{\prod_{i=1}^n \theta_1 e^{-\theta_1 x_i}}{\prod_{i=1}^n \theta_0 e^{-\theta_0 x_i}} \right) > k^{-1}$$

$$\left(\frac{\theta_1}{\theta_0}\right)^n e^{\sum_{i=1}^n (\theta_0 - \theta_1)} > \frac{1}{k} \Leftrightarrow \sum_{i=1}^n x_i (\theta_0 - \theta_1) > \log k^{-1} \left(\frac{\theta_0}{\theta_1}\right)^n \Leftrightarrow$$

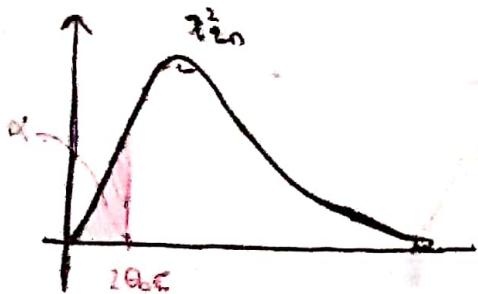
$$\sum_{i=1}^n x_i < \frac{\log k^{-1} \left(\frac{\theta_0}{\theta_1}\right)^n}{\theta_0 - \theta_1} = c$$

Assim,

$$\pi_{\Psi^*}(\theta_0) = \alpha \quad \sum x_i / \theta_0 \sim \text{Gamma}(n, \theta_0)$$

$$\begin{aligned} \pi_{\Psi^*}(\theta_0) &= P(\Psi^*(x) = 1 | \theta_0) = P\left(\sum_{i=1}^n x_i < c | \theta_0\right) = \\ &= P\left(2\theta_0 \sum_{i=1}^n x_i < 2\theta_0 c | \theta_0\right) = \alpha \end{aligned}$$

$$2\theta_0 \sum x_i / \theta_0 \sim \text{Gamma}\left(n, \frac{\theta_0}{2\theta_0}\right) = \chi^2_{2n}$$



$$\text{Logo, } 2\theta_0 c = q_{2n, \alpha} \Rightarrow c = \frac{q_{2n, \alpha}}{2\theta_0}.$$

$$\Psi^*(x) = \begin{cases} 1, & \sum x_i \leq q_{2n, \alpha} / 2\theta_0 \\ 0, & \text{c.c.} \end{cases}$$

Exemplo 3: $H_0: \theta = \theta_0 \times H_1: \theta = \theta_1 \quad \alpha = \{1, 2, 3, 4, 5, 6\}$

x	1	2	3	4	5	6	$\alpha = 20\%$
$f(x \theta_0)$	0.2	0.2	0.2	0.1	0.1	0.1	$P(X=1 \theta_0)=10\%$
$f(x \theta_1)$	0.3	0.2	0.1	0.1	0.1	0.1	$P\{X \in \{1,4\} \theta_1\}=20\%$
$\frac{f(x \theta_1)}{f(x \theta_0)}$	3	1	1/2	2	1/2	1/2	$\Psi^*(x) = \begin{cases} 1, & x \in \{1,4\} \text{ é o teste MP de} \\ 0, & \text{c.c.} \end{cases}$ nível 20% 139

Analogamente

$$\psi^*(x) = \begin{cases} 1, & x \in \{1, 2, 4\} \\ 0, & \text{c.c.} \end{cases} \quad \text{é o teste MP de nível } 40\%.$$

$$\alpha = 25\%.$$

$$\psi^*(x) = \begin{cases} 1, & x \in \{1, 4\} \\ \gamma, & x = 2 \\ 0, & \text{c.c.} \end{cases}$$

Vou obter γ de modo que

$$E[\psi^*(x)|\theta_0] = \alpha = 25\%.$$

$$\begin{aligned} E[\psi^*(x)|\theta_0] &= 1 \cdot P(\psi^*(x)=1|\theta_0) + \gamma \cdot P(\psi^*(x)=\gamma|\theta_0) + 0 \cdot P(\psi^*(x)=0|\theta_0) \\ &= 1 \cdot P(x \in \{1, 4\}|\theta_0) + \gamma \cdot P(x=2|\theta_0) = 1 \cdot (0,2) + \gamma \cdot (0,2) = 0,25 \Rightarrow \end{aligned}$$

$$\psi^*(x) = \begin{cases} 1, & x \in \{1, 4\} \\ \frac{\gamma}{4}, & x=2 \\ 0, & \text{c.c.} \end{cases} \quad \Rightarrow \gamma = \frac{1}{4}$$

Exemplo 4 $H_0: \Theta_2 \Theta_0 \times H_1: \Theta \neq \Theta_0$

x	1	2	3	4	5
$\psi(x \theta_0)$	0,02	0,02	0,06	0,4	0,5
$\psi(x \theta_1)$	0,6	0,12	0	0	0
χ/θ_0	40	10	0	0	0
χ/θ_1					

$$\alpha = 5\%.$$

$$P(X \leq 1|\theta_0) = 0,7.$$

$$P(X \leq 1|\theta_1) = 0,1.$$

$$P(X \leq 3|\theta_0) = 10^{-1}$$

~~α~~

$$\Psi_A^*(x) = \begin{cases} 1, & x \leq 2 \\ \frac{1}{6}, & x = 3 \\ 0, & x \geq 3 \end{cases} \quad \pi_{\Psi_A^*}(\theta_0) = 0,05$$

$$\pi_{\Psi_A^*}(\theta_1) = 1$$

$$\Psi^*(x) = \begin{cases} 1, & x \leq 2 \\ 0, & x \geq 2 \end{cases} - \pi_{\Psi}(\theta_0) = 0,04$$

$$\pi_{\Psi^*}(\theta_1) = 1$$

Exemplo: X_1, \dots, X_{10} dados Θ , c.i.i.d. $Ber(\theta)$

$$H_0: \Theta = 1/2 \quad x = (x_1, \dots, x_{10}) \in \mathfrak{X} = \{0,1\}^{10}$$

$$H_1: \Theta = 8/10$$

$$\Psi^*(x) = 1 \Leftrightarrow \frac{\Psi(x/\Theta = 8/10)}{\Psi(x/\Theta = 1/2)} = \frac{\left(\frac{8}{10}\right)^{\sum x_i} \left(\frac{2}{10}\right)^{n-\sum x_i}}{\left(\frac{1}{2}\right)^{\sum x_i} \left(\frac{1}{2}\right)^{n-\sum x_i}} > k^{-1} \Leftrightarrow$$

$$\Leftrightarrow \frac{2^n}{10^n} 8^{\sum x_i} 2^{n-\sum x_i} > \frac{1}{k} \Leftrightarrow \left(\frac{4}{10}\right)^n 4^{\sum x_i} > \frac{1}{k} \Leftrightarrow 4^{\sum x_i} > k^{-1} \left(\frac{10}{4}\right)^n \Leftrightarrow$$

$$\Leftrightarrow \sum x_i > \frac{\log k^{-1} \left(\frac{10}{4}\right)^n}{\log 4} = 6$$

$$\alpha = 5\%$$

$$P(X \geq 3 | \Theta = 1/2) = (1/2)^{10} = \frac{1}{1024} < 5\%$$

$$P(X \geq 9 | \Theta = 1/2) = \binom{10}{9} \left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right) = \frac{11}{1024} < 5\%$$

$$P(X \geq 8 | \Theta = 1/2) = \binom{10}{8} \left(\frac{1}{2}\right)^{10} + \binom{10}{9} \left(\frac{1}{2}\right)^{10} + \frac{1}{1024} = \frac{56}{1024} > 5\%$$

Assim,

$$\varphi^*(x) = \begin{cases} 1, & x \in \{9, 10\} \\ 0, & \text{c.c.} \end{cases} \quad \text{é o teste M.P de nível } \frac{11}{1024}$$

$$\varphi^{**}(x) = \begin{cases} 1, & x \in \{8, 9, 10\} \\ 0, & \text{c.c.} \end{cases} \quad \text{é o teste MR de nível } \frac{56}{1024}.$$

Avaliação 2 -

Avaliação 3 - 02/07

$$\Theta = \{\theta_0, \theta_1\} \quad \Theta_0 = \{0, 1, \dots, 0, 1\}$$

$$H_0: \theta = \theta_0 \quad (\theta \in \Theta_0)$$

$$H_1: \theta = \theta_1 \quad (\theta \in \Theta_1)$$

P*

$$P(\theta = \theta_1) = p_1$$

$$0 \leq p_1 \leq 1.$$

$$P(\theta = \theta_0) = p_0 = 1 - p_1$$

		$\theta = \theta_0$	$\theta = \theta_1$
Decisão	$\theta = \theta_0$	0	a_2
	$\theta = \theta_1$	$(L(0, \theta_0))$	$(L(0, \theta_1))$
Não Rejeitar H_0 (C)	a_1	0	
	$(L(1, \theta_0))$	$(L(1, \theta_1))$	

$a_1, a_2 \geq 0.$

$\Psi = \{\psi: \mathcal{X} \rightarrow \{0, 1\}\}$. Devemos escolher $\psi^* \in \Psi$ que minimiza

$$\rho(\psi) = E[L(\psi(x), \theta)]$$

Para $\psi \in \Psi$,

$$\rho(\psi) = \underbrace{E[L(\psi(x), \theta)]}_{g(x, \theta)} = \sum_{i=0}^1 \sum_{x \in \mathcal{X}} L(\psi(x), \theta_i) P(\theta = \theta_i, i=x) =$$

$$= \sum_{i=0}^1 \left\{ \sum_{x \in \mathcal{X}} L(\psi(x), \theta_i) P(i=x | \theta = \theta_i) \right\} \overbrace{P(\theta = \theta_i)}^{p_i} =$$

$$\begin{aligned}
 &= p_0 \left\{ \sum_{x \in \Omega(\theta_0)} L(q(x), \theta_0) P(X=x|\theta_0) + \sum_{x \in \Omega^c(\theta_0)} L(q(x), \theta_0) P(X=x|\theta_0) \right\} + \\
 &+ p_1 \left\{ \sum_{x \in \Omega(\theta_1)} L(q(x), \theta_1) P(X=x|\theta_1) + \sum_{x \in \Omega^c(\theta_1)} L(q(x), \theta_1) P(X=x|\theta_1) \right\} = \\
 &= p_0 \left\{ \sum_{x \in \Omega(\theta_0)} \alpha_1 P(X=x|\theta_0) \right\} + p_1 \left\{ \sum_{x \in \Omega(\theta_1)} \alpha_2 P(X=x|\theta_1) \right\} =
 \end{aligned}$$

$$= p_0 \alpha_1 P(q(x)=1|\theta_0) + p_1 \alpha_2 P(q(x)=0|\theta_1) \Leftrightarrow$$

$$\Rightarrow \rho(q) = p_0 \cdot \overset{\Delta}{\alpha_1} \pi_q(\theta_0) + p_1 \overset{\ominus}{\alpha_2} (1 - \pi_q(\theta_1)) \Rightarrow$$

$$\Rightarrow p_0 \alpha_1 \pi_q(\theta_0) - p_1 \alpha_2 (1 - \pi_q(\theta_1)) = \rho(q^*),$$

onde $q^* \in \Psi$ é dado por

$$q^*(x) = \begin{cases} 1, & p_0 \alpha_1 P(X=x|\theta_0) \geq p_1 \alpha_2 P(X=x|\theta_1) \\ 0, & \text{o.c.} \end{cases}$$

q^* é chamado, nesse caso, de teste de Bayes.

Notar que

$$q^*(x)=1 \Leftrightarrow \alpha_1 P(X=x|\theta_0) P(G=G_0) \geq \alpha_2 P(X=x|\theta_1) P(G=G_1) \Leftrightarrow$$

$$\Leftrightarrow \frac{\alpha_1 P(X=x|\theta_0) P(G=G_0)}{P(X=x)} \leq \frac{\alpha_2 P(X=x|\theta_1) P(G=G_1)}{P(X=x)} \Leftrightarrow$$

$$\Leftrightarrow \frac{P(G_0, \theta_0 | X=x)}{P(G_1, \theta_1 | X=x)} \leq \frac{P(G_1, \theta_1 | X=x)}{P(G_0, \theta_0 | X=x)} \Rightarrow$$

$$\psi^*(x) = 1 \Leftrightarrow P(\theta = \theta_1 / X = x) > \frac{a_1}{a_1 + a_2}$$

Exemplos:

(1) X_1, \dots, X_n dado θ , c.i.i.d. $N(\theta, 1)$

$$H_0: \theta = 0$$

$$H_1: \theta = 1$$

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

$$P(\theta = 1 / X = x) = \frac{\psi(x | \theta = 1) P(\theta = 1)}{\psi(x | \theta = 1) P(\theta = 1) + \psi(x | \theta = 0) P(\theta = 0)} =$$

$$= \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^n (x_i - 1)^2}{2}} \cdot p_1}{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^n (x_i - 0)^2}{2}} \cdot p_0 + \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^n x_i^2}{2}} \cdot p_0} =$$

$$= \frac{p_1}{p_1 + p_0 e^{\frac{2(x_{i=1})^2 - 2x_i^2}{2}}} = \frac{p_1}{p_1 + p_0 e^{\frac{-2\sum x_i + n}{2}}} \Rightarrow \frac{a_1}{a_1 + a_2} \Leftrightarrow$$

$$\Leftrightarrow p_1 a_2 > p_0 a_1 e^{\frac{n}{2} - \sum x_i} \Leftrightarrow \log \frac{p_1 a_2}{p_0 a_1} > \frac{n}{2} - \sum x_i \Leftrightarrow$$

$$\sum_{i=1}^n x_i > \frac{n}{2} - \log \left(\frac{p_1 a_2}{p_0 a_1} \right)$$

• 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. $\text{Exp}(\theta)$.

$$H_0: \theta = \theta_0$$

$$H_1: \theta = \theta_1$$

$$P(\theta = \theta_1) = p_1$$

$$\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{X}$$

$$P(\theta = \theta_1 | \mathbf{x} = \mathbf{x}) = \frac{f(\mathbf{x} | \theta_1) p_1}{f(\mathbf{x} | \theta_1) p_1 + f(\mathbf{x} | \theta_0) p_0} = \\ = \frac{\theta_1^n e^{-\theta_1 \sum x_i} p_1}{\theta_1^n e^{-\theta_1 \sum x_i} p_1 + \theta_0^n e^{-\theta_0 \sum x_i} p_0} =$$

$$= \frac{p_1 \theta_1^n}{p_1 \theta_1^n + p_0 \theta_0^n e^{(\theta_1 - \theta_0) \sum x_i}} \cdot \frac{a_1}{a_1 + a_2} \Leftrightarrow$$

$$\Leftrightarrow a_2 p_1 \theta_1^n > a_1 p_0 \theta_0^n e^{(\theta_1 - \theta_0) \sum x_i} \Leftrightarrow$$

$$\Leftrightarrow \sum_{i=1}^n x_i < \boxed{\frac{\log \left(\frac{a_2 p_1 \theta_1^n}{a_1 p_0 \theta_0^n} \right)}{(\theta_1 - \theta_0)}}$$

$$\psi^*(\mathbf{x}) = \begin{cases} 1, & \sum_{i=1}^n x_i \square \\ 0, & \text{c.c.} \end{cases}$$

- Rafael Izbricki (2010)
└ dualidade de hipóteses
- Gustavo Miranda da Silva (2010)
- Victor Fossaluza (2008)

Exemplo 3: $\lambda_1, \dots, \lambda_n$, dado $\Theta = \{0, 1\}$, a.s.d. $Ber(\Theta)$

$$\begin{aligned} H_0: \Theta &= \Theta_0 & (0 < \theta_0 < \theta_1 < 1) \\ H_1: \Theta &= \Theta_1 \end{aligned}$$

$$P(\Theta = \Theta_i) = p_i, i = 0, 1 \quad \mathcal{X} = \{0, 1\}^n$$

$$x = (x_1, \dots, x_n) \in \mathcal{X}$$

$$\begin{aligned} P(\Theta = \Theta_1 | X=x) &= \frac{P(X=x | \Theta_1) p_1}{P(X=x | \Theta_1) p_1 + P(X=x | \Theta_0) p_0} = \frac{\theta_1^{2x_1} (1-\theta_1)^{n-2x_1} p_1}{\theta_1^{2x_1} (1-\theta_1)^{n-2x_1} p_1 + \theta_0^{2x_1} (1-\theta_0)^{n-2x_1} p_0} = \\ &= \frac{p_1 \left(\frac{\theta_1}{1-\theta_1} \right)^{2x_1} (1-\theta_1)^n}{p_1 \left(\frac{\theta_1}{1-\theta_1} \right)^{2x_1} (1-\theta_1)^n + p_0 \left(\frac{\theta_0}{1-\theta_0} \right)^{2x_1} (1-\theta_0)^n} = \frac{p_1 (1-\theta_1)^n}{p_1 (1-\theta_1)^n + p_0 (1-\theta_0)^n \left(\frac{\theta_0 (1-\theta_1)}{p_1 (1-\theta_1)} \right)^{2x_1}} \xrightarrow{\theta_1} \end{aligned}$$

\Leftrightarrow

$$\alpha_2 p_1 (1-\theta_1)^n > \alpha_1 p_0 (1-\theta_0)^n \left(\frac{\theta_0 (1-\theta_1)}{p_1 (1-\theta_1)} \right)^{2x_1} \Leftrightarrow$$

$$\sum_{i=1}^n x_i > \frac{\log \left(\frac{\alpha_2 p_1 (1-\theta_1)^n}{\alpha_1 p_0 (1-\theta_0)^n} \right)}{\log \left(\frac{\theta_0 (1-\theta_1)}{\theta_1 (1-\theta_0)} \right)}$$

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

$$P(\Theta = \Theta_0) = p_0$$

$$P(\Theta = \Theta_1 | X=x) = \frac{P(X=x | \Theta = \Theta_1) p_1}{P(X=x | \Theta = \Theta_1) p_1 + P(X=x | \Theta = \Theta_0) p_0}$$

$$= \begin{cases} 0, & x = 3, 4, 5 \\ \frac{0.8 p_2}{0.8 p_2 + 0.02 p_0} = \frac{40 p_1}{40 p_1 + p_0}, & x = 1 \\ \frac{0.2 p_1}{0.2 p_2 + 0.02 p_0} = \frac{10 p_1}{10 p_1 + p_0}, & x = 2 \end{cases}$$

Considerando

$$p_2 = \frac{1}{3}, \quad p_0 = \frac{2}{3}$$

$$\alpha_1 = 9, \alpha_2 = 2 \rightarrow \frac{\alpha_1}{\alpha_1 + \alpha_2} = \frac{9}{10}$$

$$P(\Theta=1 | X=x) = \begin{cases} 0, & x = 3, 4, 5 \\ \frac{40}{42}, & x = 1 \\ \frac{10}{32}, & x = 2 \end{cases} \rightarrow \psi^*(x) = \begin{cases} 1, & x = 1 \\ 0, & \text{c.c.} \end{cases}$$

$$\alpha_1 = 4, \alpha_2 = 1$$

$$\frac{\alpha_1}{\alpha_1 + \alpha_2} = \frac{4}{5} = 80\%. \quad \psi^*(x) = \begin{cases} 1, & x = 1, 2 \\ 0, & \text{c.c.} \end{cases}$$

Problema passado

$$P(\Theta=\Theta_0) = p_0 \quad P(\Theta=\Theta_1) = p_1$$

Hipóteses Compostas

$$\Theta = \Theta_0 \cup \Theta_1, \quad \Theta_0 \cap \Theta_1 = \emptyset$$

$$H_0: \Theta \in \Theta_0$$

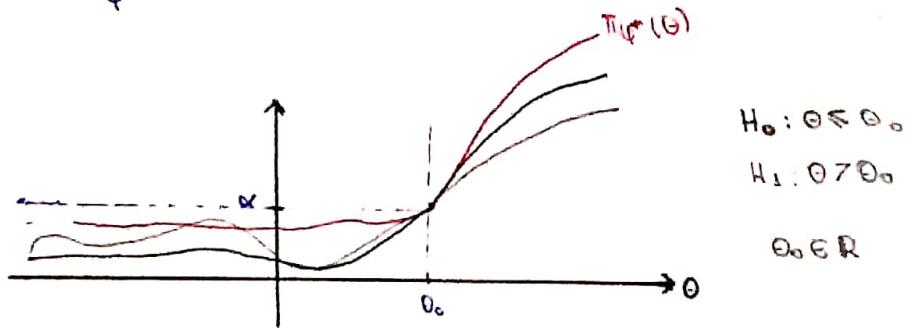
$$H_1: \Theta \in \Theta_1$$

$$\Psi = \{\psi: \mathfrak{X} \rightarrow [0,1]\}$$

Definição: O teste $\psi^* \in \Psi$ 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 $\forall \psi \in \Psi$ tal que $\alpha_{\psi} \leq \alpha$,

$$\Pi_{\psi^*}(\theta) \geq \Pi_{\psi}(\theta), \quad \forall \theta \in \Theta_1$$



Em geral, é difícil (não existe) teste UMP!!

Vamos olhar um caso importante onde existe teste UMP!

E

$$\Theta \subseteq \mathbb{R} \quad \mathcal{P} = \{f(\cdot | \theta) : \theta \in \Theta\}$$

$$\text{Seja } T: \mathfrak{X} \rightarrow \mathbb{R}$$

Definição: A família $\mathcal{P} = \{f(\cdot | \theta) : \theta \in \Theta\}$ possui razão de verossimilhança monótona crescente (RVM)

se $\frac{f(x|\theta_1)}{f(x|\theta_2)}$ é função crescente (decrecidente) de $T(x)$.

$$\forall \theta_1, \theta_2 \in \Theta, \text{ com } \theta_1 < \theta_2,$$

$\frac{f(x|\theta_2)}{f(x|\theta_1)}$ é função crescente (decrecidente) 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 $\theta_1 < \theta_2$.

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum(x_i-\theta_2)^2}{2}}}{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum(x_i-\theta_1)^2}{2}}} = e^{\frac{\sum(x_i-\theta_2)^2 - \sum(x_i-\theta_1)^2}{2}} = e^{\frac{\sum(x_i-\theta_2)^2 - 2\theta_1\sum x_i + n\theta_1^2 - \sum x_i^2 + 2\theta_2\sum x_i - n\theta_2^2}{2}}$$

$$= e^{\frac{n(\theta_1^2 - \theta_2^2)}{2}} \cdot e^{\frac{2(\theta_2 - \theta_1)}{2} \sum x_i} \text{ é crescente em } T(x) = \sum_{i=1}^n x_i.$$

Então, diremos que 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. $\text{Exp}(\theta)$

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

Sejam $\theta_1, \theta_2 > 0$, com $\theta_1 < \theta_2$

$$x \in \mathbb{R} = \mathbb{R}^n$$

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{\theta_2^n e^{-\theta_2 \sum x_i}}{\theta_1^n e^{-\theta_1 \sum x_i}} = \left(\frac{\theta_2}{\theta_1}\right)^n e^{(\theta_2 - \theta_1) \sum x_i} \text{ é decrescente em } \sum_{i=1}^n x_i.$$

Logo, P possui RVM decrescente em $T(X) = \sum_{i=1}^n x_i$

(crescente em $-\sum x_i$). Logo, P possui RVM crescente em $-\sum_{i=1}^n x_i$.

Scheruish

$\mathcal{D} = \{f(x|\theta) : \theta \in \Theta\}$ 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. $X | \theta \sim U(0, \theta)$

$$\Theta = \mathbb{R}_+ \subseteq \mathbb{R} \quad T(X) = X \geq 0$$

$\theta_1, \theta_2 > 0$ com $\theta_1 < \theta_2$

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{\frac{1}{\theta_2} I_{[0,\theta_2]}(x)}{\frac{1}{\theta_1} I_{[0,\theta_1]}(x)} = \frac{\theta_1}{\theta_2} \frac{I_{[0,\theta_2]}(x)}{I_{[0,\theta_1]}(x)} = \begin{cases} \frac{\theta_1}{\theta_2}, & 0 \leq x \leq \theta_2, \\ \infty, & \theta_1 < x \leq \theta_2, \\ ?, & x > \theta_2 \end{cases}$$

Exemplo 4. $X | \theta \sim \text{Cauchy}(\theta) ((\theta, \beta))$

$$f(x|\theta) = \frac{1}{\pi [1 + (x-\theta)^2]}, \quad x \in \mathbb{R}$$

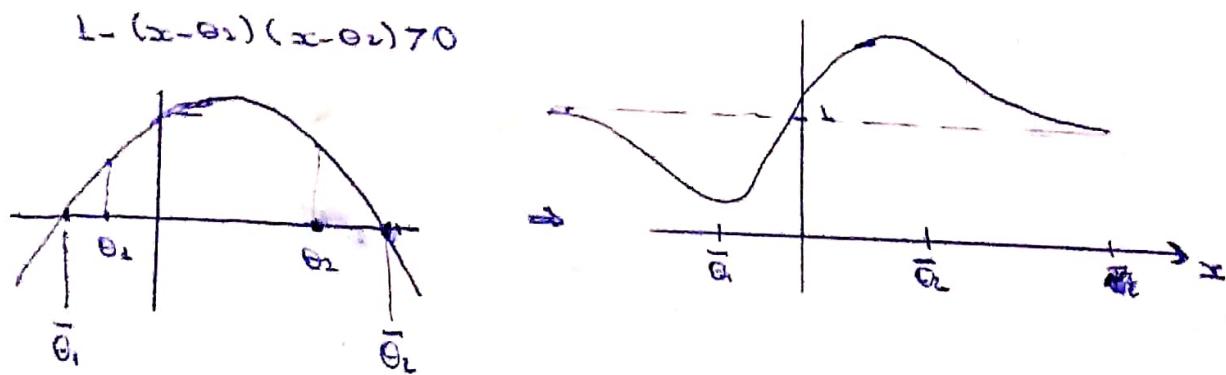
$$\Theta = \mathbb{R}$$

$\theta_1, \theta_2 \in \mathbb{R}, \theta_1 < \theta_2$

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{\frac{1}{\pi [1 + (x-\theta_2)^2]}}{\frac{1}{\pi [1 + (x-\theta_1)^2]}} = \frac{1 + (x-\theta_1)^2}{1 + (x-\theta_2)^2} = \mu(x)$$

$$\begin{aligned}
 M'(x) &= \frac{2(x-\theta_3)[1+(x-\theta_2)^2] - [1+(x-\theta_3)^2]2(x-\theta_2)}{[1+(x-\theta_2)^2]^2} = \\
 &= \frac{2}{[1+(x-\theta_2)^2]^2} \left\{ (x-\theta_3) + (x-\theta_3)(x-\theta_2)^2 - (x-\theta_2) - (x-\theta_2)(x-\theta_3)^2 \right\} = \\
 &= A(x) \cdot \left\{ (\theta_2 - \theta_3) + (x-\theta_2)(x-\theta_2) \underbrace{[(x-\theta_2) - (x-\theta_3)]}_{-(\theta_2 - \theta_3)} \right\} = \\
 &= \underbrace{A(x)(\theta_2 - \theta_3)}_{>0} \{ 1 - (x-\theta_2)(x-\theta_2) \}
 \end{aligned}$$

M' é crescente se, e só se, $M'(x) > 0 \Leftrightarrow$



Teorema (Karlin-Rubin)

$\theta \in \mathbb{R}$, $P, T: \mathcal{X} \rightarrow \mathbb{R}$, Ψ

Suponhamos que P possui RVM crescente em T . O teste $\Psi^*: \mathcal{X} \rightarrow \{0,1\}$ dada

por

$$\Psi^*(x) = \begin{cases} 1, & T(x) > c \quad (T(x) \leq c) \\ 0, & \text{c.c.} \end{cases}$$

para as hipóteses $H_0: \theta \leq \theta_0$ versus $H_1: \theta > \theta_0$, $\theta_0 \in \Theta$, é UMP de nível $\pi_{\psi^*}(\theta_0)$.

ψ^*

$\pi_{\psi^*}: \Theta \rightarrow [0,1] \in$ não-decrescente!

$\theta_1 < \theta_2 \Rightarrow$

$$\pi_{\psi^*}(\theta) = P(\psi^*(x) = 1 | \theta) = P(T(x) > c | \theta).$$

$$T(x) > c \Leftrightarrow \frac{f(x|\theta_1)}{f(x|\theta_2)} > d$$

Vamos verificar que $\forall t \in \mathbb{R}$, $P(T(x) > t | \theta_2) \leq P(T(x) > t | \theta_1)$

$t' \geq t$

$$\begin{aligned} P(T(x) > t | \theta_2) &= \int_{\{x: T(x) > t\}} f(x|\theta_2) dx \leq \int_{\{x: T(x) > t\}} \frac{1}{t'} f(x|\theta_2) dx = \frac{1}{t'} \int_{\{x: T(x) > t\}} f(x|\theta_2) dx \leq \\ &\leq \int_{\{x: T(x) > t\}} f(x|\theta_1) dx = P(T(x) > t | \theta_1) \end{aligned}$$

$t' \leq t$

$$\begin{aligned} P(T(x) \leq t | \theta_2) &= \int_{\{x: T(x) \leq t\}} f(x|\theta_2) dx \leq \int_{\{x: T(x) \leq t\}} t' \cdot f(x|\theta_1) dx = P(T(x) \leq t | \theta_1) = \\ &\Rightarrow P(T(x) > t | \theta_1) \leq P(T(x) > t | \theta_2). \end{aligned}$$

Assim, em particular para $t = c$, segue que $\pi_{\psi^*}(\theta_1) \leq \pi_{\psi^*}(\theta_2)$

Para $\theta \leq \theta_0$, segue que $\pi_{\psi^*}(\theta) \leq \pi_{\psi^*}(\theta_0)$. Logo, ψ^* tem tamanho $\pi_{\psi^*}(\theta_0)$, ou ainda,

$$\Psi^* \in \Psi^* = \{\psi \in \Psi : \alpha_\psi^* \leq \pi_{\psi^*}(\theta_0)\}$$

$\theta \in \Theta_1$

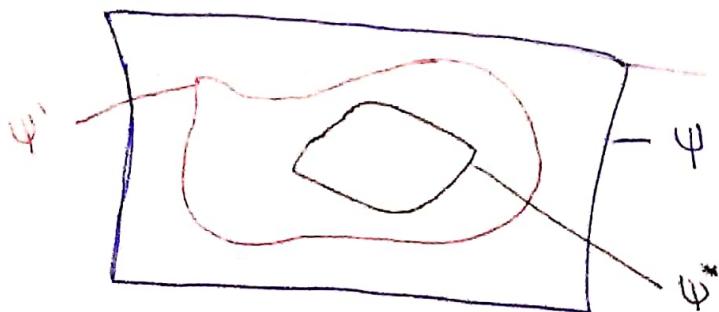
Vamos verificar que para todo $\theta > \theta_0$, $\pi_{\psi^*}(\theta) \geq \pi_\psi(\theta)$, $\forall \psi \in \Psi^*$.

Para $\theta_2 > \theta_0$,

$$\pi_{\psi^*}(\theta_2) = P(\psi^*(x) = 1 | \theta_2) = P(T(x) \geq c | \theta_2) = P\left(\frac{f(x|\theta_2)}{f(x|\theta_1)} \geq d | \theta\right)$$

$$\bar{\psi}(x) = \begin{cases} 1, & \frac{f(x|\theta_2)}{f(x|\theta_1)} \geq d \\ 0, & \text{c.c.} \end{cases}$$

Pelo Lema de Neyman-Pearson, $\pi_{\psi^*}(\theta_2) \geq \pi_\psi(\theta_2)$, $\forall \psi \in \Psi' = \{\psi \in \Psi : \pi_\psi(\theta) \leq \pi_{\psi^*}(\theta_0)\}$



Logo, $\forall \psi \in \Psi^*$, $\pi_{\psi^*}(\theta_2) \geq \pi_\psi(\theta_2)$, $\forall \theta_2 > \theta_0$

Logo, Ψ^* é UMP em Ψ^* . ($\forall \theta \in \Theta_1$).

Exemplo 1. X_1, \dots, X_n dado θ , c.i.i.d. $N(\theta, 1)$

$H_0: \theta \leq \theta_0$ $\theta_0 \in \mathbb{R}$ (fixado)

$H_1: \theta > \theta_0$

Vamos na ultima vez que P possua RVM crescente em $T(x) = \sum_{i=1}^n X_i$.

Para $\psi \in \Psi(0, 1)$

$$\psi^*(x) = \begin{cases} 1, & \sum_{i=1}^n x_i > c = c(\alpha) \\ 0, & \text{c.c.} \end{cases}$$

é UMP de nível α para H_0 versus H_1 .

Assim, $c \in \mathbb{R}$ é tal que

$$\pi_{\psi^*}(0_0) = \alpha \Rightarrow P(\psi^*(x) = 1 | 0_0) = \alpha \Rightarrow$$

$$\Rightarrow P\left(\sum_{i=1}^n x_i > c | 0_0\right) = \alpha \Rightarrow P\left(\frac{\sum_{i=1}^n x_i - n\theta_0}{\sqrt{n}} > \frac{c - n\theta_0}{\sqrt{n}} \mid 0_0\right) = \alpha$$

$$\Rightarrow P\left(Z > \frac{c - n\theta_0}{\sqrt{n}}\right) = \alpha \Rightarrow \frac{c - n\theta_0}{\sqrt{n}} = \Phi^{-1}(1-\alpha) \Rightarrow c = n\theta_0 + \sqrt{n} \Phi^{-1}(1-\alpha)$$

Daí,

$$\psi^*(x) = \begin{cases} 1, & \sum_{i=1}^n x_i > n\theta_0 + \sqrt{n} \Phi^{-1}(1-\alpha) \\ 0, & \text{c.c.} \end{cases}$$

Exemplo 2. X_1, \dots, X_n , dado θ , c.i.i.d. $\mathcal{U}(0, \theta)$

$$H_0: \theta \leq \theta_0 \quad \theta_0 > 0 \quad (\text{fixado})$$

$$H_1: \theta > \theta_0$$

Vimos na última vez que P possui fmm crescente em $T(x) = X_{(n)} = \max\{X_1, \dots, X_n\}$ para $x \in (0, \theta)$

$$\psi^*(x) = \begin{cases} 1, & x_{(n)} > c = c(\alpha) \\ 0, & \text{c.c.} \end{cases}$$

é UMA de nível α para H_0 vs. H_1 .

Assim, c7o é tal que

$$\Pi_{\varphi^*}(\theta_0) = \alpha \Rightarrow P(\psi(x)=1|\theta_0) = \alpha \Rightarrow$$

$$\Rightarrow P(x_m > c|\theta_0) = \alpha \Rightarrow P(x_m \leq c|\theta_0) = 1 - \alpha \Rightarrow$$

$$c \in (0, \theta_0)$$

$$\Rightarrow \left(\frac{c}{\theta_0}\right)^n = 1 - \alpha \Rightarrow c = \theta_0 \cdot (1 - \alpha)^{1/n}.$$

Exemplo 4. X_1, \dots, X_n , dado Θ , c.i.d. $\text{Exp}(\theta)$

$$H_0: \theta \leq \theta_0$$

$$H_1: \theta > \theta_0$$

$$\theta_0 > 0 \text{ (fixado)} \quad \Theta = \mathbb{R}_+$$

Vimos na última vez que \mathcal{T} possui RVM decrescente em $\sum_{i=1}^n X_i$ e, consequentemente, possui RVM crescente em $T(\lambda) = -\sum_{i=1}^n X_i$.

Para $k \in (0, 1)$

$$\psi(x) = \begin{cases} 1, & \sum_{i=1}^n X_i \leq k \\ 0, & \text{c.c.} \end{cases}$$

é UMP de nível α para testar H_0 vs H_1 .

$\sum_{i=1}^n X_i | \theta_0 \sim \text{Gamma}(n, \theta_0)$

Logo, $\mathcal{R} \neq \emptyset$ é tal que

$$\Pi_{\varphi}(\theta_0) = \alpha \Rightarrow P(\psi(x)=1|\theta_0) = \alpha \Rightarrow P\left(\sum_{i=1}^n X_i \leq k | \theta_0\right) = \alpha \Rightarrow$$

$$\Rightarrow P(2\theta_0 \sum_{i=1}^n X_i \leq 2\theta_0 k | \theta_0) = \alpha \Rightarrow$$

$\Rightarrow P(U \leq 2\theta_0) = \alpha$ onde $U \sim \chi^2_{2n}$

$$\Rightarrow 2\theta_0 K = q_{2n,\alpha} \Rightarrow K = \frac{q_{2n,\alpha}}{2\theta_0}$$

Dai,

para $x \in (0,1)$,

$$\psi^*(x) = \begin{cases} 1, & \sum_{i=1}^n x_i \leq q_{2n,\alpha}/2\theta_0 \\ 0, & \text{c.c.} \end{cases}$$

é UMP de nível α para testar H_0 versus H_2 .

Teorema Karlin --

O.S.R. $H_0: \theta \geq \theta_0$ } Tarefa!
 $H_1: \theta < \theta_0$

Resultado: X_1, \dots, X_n dado θ_1 , c.i.i.d. tal que $f(x|\theta) = c(\theta) e^{g(\theta) \cdot T(x)}$.
 $g(\cdot)$ estritamente crescente.

Considere as hipóteses

$$H_0: \theta \leq \theta_1 \cup \theta \geq \theta_2, \quad \theta_1 < \theta_2$$

$$H_1: \theta_1 < \theta < \theta_2$$

O teste $\psi^*: \mathcal{X} \rightarrow \{0,1\}$ dado por

$$\psi^*(x) = \begin{cases} 1, & c_1 \leq \sum_{i=1}^n T(x_i) \leq c_2 \\ 0, & \text{c.c.} \end{cases}$$

é UMP de nível $\pi_{\psi^*}(\theta_1) + \pi_{\psi^*}(\theta_2) = 1$ (c_1 e c_2 são tomados de modo a satisfaçõe (1)).

Exemplo 1. $X \mid \theta \sim \text{Exp}(\theta)$

$H_0: \theta \leq 1$ ou $\theta \geq 2$ $H_1: 1 < \theta < 2$.

$$\psi(\theta|x) = \theta e^{-\theta x} I_{(0, \infty)}(x) \quad c(\theta) = \theta \quad l(x) = I_{(0, \infty)}(x)$$

$$g(\theta) = \theta \quad T(x) = x$$

$$\psi^*(x) = \begin{cases} 1, & k_1 \leq x \leq k_2 \\ 0, & \text{c.c.} \end{cases}$$

Ou ainda,

$$\psi^*(x) = \begin{cases} 1, & k_1 \leq x \leq k_2 \\ 0, & \text{c.c.} \end{cases}$$

Fixado α , formamos $k_1, k_2 \in \mathbb{R}$ de modo que $\pi_{\psi}(1) = \pi_{\psi}(2) = \alpha$.

$$\pi_{\psi}(1) = P(\psi^*(X)=1|\theta) = P(k_1 \leq X \leq k_2/\theta) = e^{-k_1\theta} - e^{-k_2\theta}$$

Assim, devemos ter,

$$\left. \begin{array}{l} \pi_{\psi}(1) = e^{-k_1} - e^{-k_2} = \alpha \\ \pi_{\psi}(2) = e^{-2k_1} - e^{-2k_2} = \alpha \end{array} \right\} \Leftrightarrow \left. \begin{array}{l} e^{-k_1} - e^{-k_2} = \alpha \\ (e^{-k_1} - e^{-k_2})(e^{-k_1} + e^{-k_2}) = \alpha \end{array} \right.$$

$$e^{-k_2} = e^{-k_1} - \alpha = \frac{1-\alpha}{2} - \alpha = 1 - \frac{\alpha}{2} \Rightarrow k_2 = -\log\left(\frac{1-\alpha}{2}\right)$$

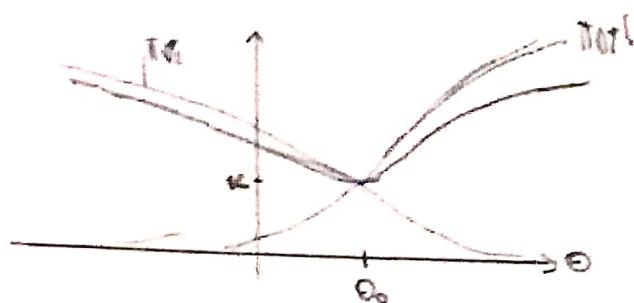
Assim,

$$\psi^*(x) = \begin{cases} 1, & -\log\left(\frac{1-\alpha}{2}\right) \leq x \leq -\log\left(\frac{1-\alpha}{2}\right) \\ 0, & \text{c.c.} \end{cases}$$

Obs.

$$H_0 \cdot \theta = 0$$

$$H_1 \cdot \theta \neq 0$$



Funcionamiento más eficiente

Volviendo al Ejemplo 3.

$$\chi/\theta \text{ resp}(\theta)$$

$$H_0 : \theta \leq L \quad n=1$$

$$H_1 : \theta > L \quad \theta \geq L$$

$$\Psi(x) = \begin{cases} 1, & x \in [-\log(1-\alpha), 0,0513] \Rightarrow \theta \leq L \\ 0, & \text{o.s.} \end{cases}$$

$$\chi = 0,7$$

$$\frac{\partial}{\partial \theta} \left[\frac{\theta}{1+\theta} - \frac{0,0513}{1+\theta} \right]$$

$$1 + \theta^2 \quad 1, 0,644 < x \leq 0,744$$

$$0, \text{o.s.}$$

$$1 > \theta \geq 2$$

X_1, \dots, X_n , dado Θ , c.i.i.d. $N(\theta, \sigma^2)$

$$\begin{array}{ll} H_0: \theta \leq \theta_0 & H_0'': \theta \geq \theta_0 \\ H_1: \theta > \theta_0 & H_1'': \theta < \theta_0 \end{array} \quad \varphi(x) = \begin{cases} 1, & T(x) > c \\ 0, & \text{c.c.} \end{cases} \quad \text{Teste unilaterais}$$

$$\begin{array}{ll} H_0: \theta = \theta_0 & (\theta_1 \leq \theta \leq \theta_2) \\ H_1: \theta \neq \theta_0 & (\theta < \theta_1 \text{ ou } \theta > \theta_2) \end{array} \quad \text{Testes bilaterais}$$

Um teste intuitivo para $H_0: \theta = \theta_0 \times H_1: \theta \neq \theta_0$ de nível α , é dado por:

$$\varphi(x) = \begin{cases} 1, & |Z - \theta_0| > c = c(\alpha) \\ 0, & \text{c.c.} \end{cases}$$

$$\pi_{\varphi}(\theta_0) = \alpha \Rightarrow P(\varphi(x)=1 | \theta_0) = \alpha \Rightarrow$$

$$\Rightarrow P(|Z - \theta_0| > c | \theta_0) = \alpha \Rightarrow$$

$$|Z|, Z \sim N(0, 1)$$

$$P\left(\frac{|Z - \theta_0|}{\sqrt{\frac{1}{n}} \sigma} > \frac{c}{\sqrt{n}} \mid \theta_0\right) = \alpha \Rightarrow$$

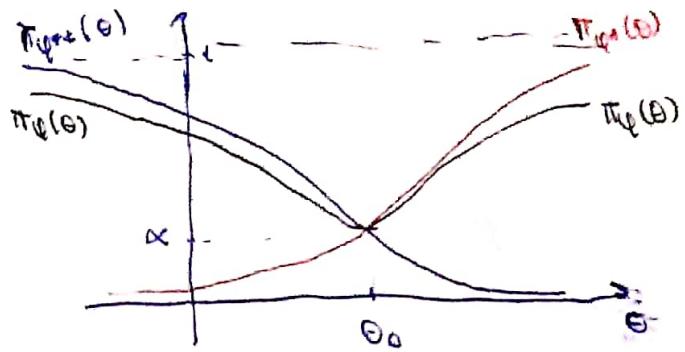
$$\Rightarrow \frac{c}{\sqrt{n}} = \Phi^{-1}(1 - \alpha/2)$$

$$\varphi(x) = \begin{cases} 1, & |Z - \theta_0| > \frac{\Phi^{-1}(1 - \alpha/2)}{\sqrt{n}} \\ 0, & \text{c.c.} \end{cases}$$

Isso não é UMP

$$\Psi^*(x) = \begin{cases} 1, & \bar{x} > \theta_0 + \frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}} \\ 0, & \text{c.c.} \end{cases}$$

$$\Psi^{**}(x) = \begin{cases} 1, & \bar{x} < \theta_0 - \frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}} \\ 0, & \text{c.c.} \end{cases}$$



Em geral, não conseguimos obter testes UMP como no caso mencionado.

Alternativa: Restringir comparações entre testes.

Def: Um teste $\Psi: \mathcal{X} \rightarrow \{0,1\}$ para $H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_1$ é não-viesado de nível α , $0 < \alpha < 1$, se

$$1. \quad \text{Type I}(\theta) \leq \alpha, \quad \forall \theta \in \Theta_0$$

$$2. \quad \text{Type II}(\theta) \geq \alpha, \quad \forall \theta \in \Theta_1$$

Def: Um teste $\Psi^*: \mathcal{X} \rightarrow \{0,1\}$ é dito NÃO-VIESADO UNIFORMEMENTE MAIS PODEROSO DE NÍVEL α (UMPU), $0 < \alpha < 1$ se Ψ^* é UMP dentro dos não-viesados de nível α .

Resultado: X_1, \dots, X_n , dado θ , c.i.i.d. com densidade

$$f(x|\theta) = c(\theta) e^{\theta \cdot T(x)} l(x) \quad \theta \in \Theta \subseteq \mathbb{R}.$$

Então,

$$\Psi^*(x) = \begin{cases} 1, & \sum_{i=1}^n T(x_i) \leq c_1 \text{ ou } \sum_{i=1}^n T(x_i) \geq c_2 \\ 0, & \text{c.c.} \end{cases}$$

$(H_0: \theta = \theta_0)$ $(H_1: \theta \neq \theta_0)$

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 $\pi_{\psi^*}(\theta_1) = \pi_{\psi^*}(\theta_2)$ ($\theta_1 < \theta_2$)

 $(\pi_{\psi^*}(\theta_0))$

Exemplo 1. X_1, \dots, X_n , dado θ , c.i.i.d., $N(\theta, 1)$

$$f(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} = \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}} e^{-\frac{x^2}{2}}}_{c(\theta)} e^{\frac{\theta x}{2}}, \quad T(x) = x$$

Logo, o teste

$$\psi^*(x) = \begin{cases} 1, & \sum_{i=1}^n x_i \leq c_1 \text{ ou } \sum_{i=1}^n x_i \geq c_2 \\ 0, & \text{c.c.} \end{cases}$$

é UMPU para $H_0: \theta = \theta_0 \times H_1: \theta \neq \theta_0$
de nível $\pi_{\psi^*}(\theta_0)$.

Assim, para $\alpha \in (0, 1)$, damos $c_1, c_2 \in \mathbb{R}$ tais que

$$\pi_{\psi^*}(\theta_0) = \alpha \Rightarrow P\left(\sum_{i=1}^n X_i \leq c_1 \cup \sum_{i=1}^n X_i \geq c_2 \mid \theta_0\right) = \alpha \Rightarrow$$

$$\Rightarrow \psi^*(x) = \begin{cases} 1, & \sum_{i=1}^n X_i - \theta_0 \leq \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}} \\ 0, & \text{c.c.} \end{cases}$$

Agora, segam $\theta_1, \theta_2 \in \mathbb{R}$ com $\theta_1 < \theta_2$.

$$H_0: \theta_1 \leq \theta \leq \theta_2$$

$$H_1: \theta < \theta_1 \cup \theta > \theta_2,$$

$$\psi^*(x) = \begin{cases} 1, & \sum_{i=1}^n X_i < c_1 \text{ ou } \sum_{i=1}^n X_i > c_2 \\ 0, & \text{c.c.} \end{cases}$$

$$\Pi_{\psi^*}(\theta_2) \sim \Pi_{\psi^*}(\theta_1) = \alpha$$

$$\Pi_{\psi^*}(\theta_1) = P(\sum x_i < c_1 \cup \sum x_i > c_2 | \theta_1) = P(X < k_1 \cup X > k_2 | \theta_1) =$$

$$= P\left(\frac{\bar{X} - \theta_1}{\sqrt{n}} < \frac{k_1 - \theta_1}{\sqrt{n}} \mid \theta_1\right) + P\left(\frac{\bar{X} - \theta_1}{\sqrt{n}} > \frac{k_2 - \theta_1}{\sqrt{n}} \mid \theta_1\right) = \alpha \Rightarrow$$

$$\Rightarrow \Phi\left(\frac{k_1 - \theta_1}{1/\sqrt{n}}\right) + 1 - \Phi\left(\frac{k_2 - \theta_1}{1/\sqrt{n}}\right) = \alpha \Rightarrow \Phi\left(\frac{k_2 - \theta_1}{1/\sqrt{n}}\right) - \Phi\left(\frac{k_1 - \theta_1}{1/\sqrt{n}}\right) = 1 - \alpha \quad (\text{I})$$

$$\Pi_{\psi^*}(\theta_2) = \alpha \Rightarrow$$

$$\Rightarrow \Phi\left(\frac{k_2 - \theta_2}{1/\sqrt{n}}\right) - \Phi\left(\frac{k_1 - \theta_2}{1/\sqrt{n}}\right) = 1 - \alpha, \quad (\text{II})$$



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)

$$H_0: \Theta = \Theta_0$$

$$H_1: \Theta \neq \Theta_0$$

$$\Psi^*(x) = \begin{cases} 1, & \frac{f(x|\Theta_0)}{f(x|\Theta_1)} > c \\ 0, & \frac{f(x|\Theta_0)}{f(x|\Theta_1)} \leq c \end{cases}$$

Suponha agora:

$$H_0: \Theta \in \Theta_0$$

$$H_1: \Theta \in \Theta_1$$

teste intuitivo

$$\Psi^*(x) = \begin{cases} 1, & \frac{\int_{\Theta_0} f(x|\theta) d\theta}{\int_{\Theta_1} f(x|\theta) d\theta} < c \text{ "testes verossimilhancistas"} \\ 0, & \text{c.c.} \end{cases}$$

ou

$$\Psi^*(x) = \begin{cases} 1, & \frac{\sup_{\Theta \in \Theta_0} f(x|\theta)}{\sup_{\Theta \in \Theta_1} f(x|\theta)} < c \\ 0, & \text{c.c.} \end{cases}$$

$$\downarrow$$

$$\text{TRVG} \Rightarrow \Psi^*(x) = \begin{cases} 1, & \frac{\sup_{\Theta \in \Theta_0} f(x|\theta)}{\sup_{\Theta \in \Theta_1} f(x|\theta)} < c \\ 0, & \text{c.c.} \end{cases}$$

$\forall \theta \in \Theta \in [0,1]$

$\sup_{\Theta \in \Theta_0} f(x|\theta)$

$\sup_{\Theta \in \Theta_1} f(x|\theta)$

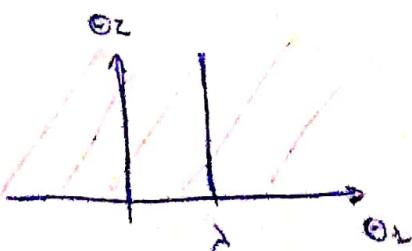
$\sqrt{-2 \log \lambda} \sim \chi^2_{n-1}$

Exemplo: x_1, \dots, x_n , dado $\Theta = (\Theta_1, \Theta_2)$, c.i.i.d. $N(\Theta_1, \Theta_2)$

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

$$H_0: \Theta_2 = \lambda$$

$$H_1: \Theta_2 \neq \lambda, \lambda \in \mathbb{R} \text{ fixado (conhecido)}$$



$$RV(x) = \frac{\sup_{\theta \in \Theta_0} f(x|\theta)}{\sup_{\theta \in \Theta} f(x|\theta)}$$

$$\text{Em } \Theta, \sup_{\theta \in \Theta} f(x|\theta) = f\left(x, \bar{x}, \sum_{i=1}^n (x_i - \bar{x})^2/n\right)$$

Em Θ_0 ,

$$f(x|\theta) = \left(\frac{1}{2\pi \sigma^2}\right)^n e^{-\frac{\sum_{i=1}^n (x_i - \theta)^2}{2\sigma^2}}, \text{ portanto,}$$

$$\sup_{\theta \in \Theta_0} f(x|\theta) = f\left(x \mid \left(\bar{x}, \sum_{i=1}^n (x_i - \bar{x})^2/n\right)\right)$$

$$RV(x) = \frac{\sup_{\theta \in \Theta_0} f(x|\theta)}{\sup_{\theta \in \Theta} f(x|\theta)} = \frac{\left(\frac{1}{\sqrt{\frac{2\pi \sum_{i=1}^n (x_i - \bar{x})^2}{n}}}\right)^n e^{-\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{2\sum_{i=1}^n (x_i - \bar{x})^2/n}}}{\left(\frac{1}{\sqrt{\frac{2\pi \sum_{i=1}^n (x_i - \bar{x})^2}{n}}}\right)^n e^{-\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{2\sum_{i=1}^n (x_i - \bar{x})^2/n}}} =$$

$$= \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \lambda)^2} \right)^{n/2} < c \Leftrightarrow \frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \lambda)^2} > \frac{(1/c)^2/n}{1} \Leftrightarrow$$

$$\frac{\sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \lambda)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} > c \Leftrightarrow \frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \lambda)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} > c \Leftrightarrow$$

$$1 + \frac{n(\bar{x} - \lambda)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} > c \Leftrightarrow \frac{n(\bar{x} - \lambda)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} > c - 1 \Leftrightarrow \frac{n(\bar{x} - \lambda)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} > \frac{c(n-1)}{n-2} \Leftrightarrow$$

$$\Rightarrow \left| \frac{(\bar{x} - \lambda) \bar{x}_n}{\sqrt{\frac{\sum_{i=1}^n (\bar{x}_i - \bar{x})^2}{n-1}}} \right| > K$$

$\downarrow S^2$

Sob H_0 , $\frac{\sqrt{n}(\bar{x} - \lambda)}{\sqrt{S^2/n-1}} \sim t_{n-1}$.

Logo, o teste RVG para $H_0: \theta_1 = \lambda$ versus $H_1: \theta \neq \lambda$ é dado por

$$\psi^*(x) = \begin{cases} 1, & \left| \frac{\sqrt{n}(\bar{x} - \lambda)}{\sqrt{S^2/n-1}} \right| > K^* = K^*(\alpha) \\ 0, & \text{c.c.} \end{cases}$$

$$\Rightarrow \theta_1 \leq \lambda,$$

$$|\bar{x} - \lambda| > 0$$

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 $X_{ij} | \Theta \sim N(\theta_i, \theta_{K+1}), i=1, \dots, K$
 $j=1, \dots, n_i$

$n_1 = \text{trat. 1}$

$n_2 = \text{trat. 2}$

\vdots

$n_K = \text{trat. 3}$

Imaginemos que as variâncias são iguais entre os grupos.

As médias diferem?

$$\textcircled{4} = \mathbb{R}^k \times \mathbb{R}_+$$

$$H_0: \theta_1 = \theta_2 = \dots = \theta_K$$

H_1 : caso contrário

$$\mathfrak{X} = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \dots \times \mathbb{R}^{n_K}$$

$$\mathbb{R}^{n_1 + n_2 + \dots + n_K}$$

$x \in \mathfrak{X}$

$$f(x|\theta) = \prod_{i=1}^K \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\theta_{ij}}} e^{-\frac{(x_{ij}-\theta_j)^2}{2\theta_{ij}}} = \left(\frac{1}{\sqrt{2\pi\theta_{\bar{x}}}} \right)^{\sum_{i=1}^K n_i} e^{-\frac{\sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij}-\theta_i)^2}{2\theta_{\bar{x}}}}$$

Em Θ_0 ,

$$\sup_{\theta \in \Theta_0} f(x|\theta) = f\left(x | (\bar{x}, \bar{x}, \dots, \bar{x}, \frac{\sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij}-\bar{x})^2}{n})\right), \text{ onde } \bar{x} = \frac{\sum_{i=1}^K \sum_{j=1}^{n_i} x_{ij}}{n_1 + \dots + n_K}, n = n_1 + \dots + n_K$$

Em $\textcircled{4}$,

$$\sup_{\theta \in \Theta} f(x|\theta) = f\left(x | (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_K, \frac{\sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij}-\bar{x}_i)^2}{n})\right), \text{ onde } \bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$$

$$RV(x_0) = \frac{\sup_{\theta \in \Theta} f(x|\theta)}{\sup_{\theta \in \Theta} f(x|\theta)} = \frac{\left(\frac{1}{\sqrt{2\pi \sum_{j=1}^{n_2} \frac{(x_{1j}-\bar{x})^2}{n}}} \right)^{n_2} e^{-\frac{\sum_{j=1}^{n_2} (x_{1j}-\bar{x})^2}{2\sum_{j=1}^{n_2} (x_{1j}-\bar{x})^2}}}{\left(\frac{1}{\sqrt{2\pi \sum_{j=1}^{n_2} \frac{(x_{1j}-\bar{x}_1)^2}{n}}} \right)^{n_2} e^{-\frac{\sum_{j=1}^{n_2} (x_{1j}-\bar{x}_1)^2}{2\sum_{j=1}^{n_2} (x_{1j}-\bar{x}_1)^2}}} =$$

$$= \left(\frac{\sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij}-\bar{x}_i)^2}{\sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij}-\bar{x})^2} \right)^{\frac{n}{2}} \subset C \Leftrightarrow \left(\frac{\sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij}-\bar{x}_i)^2}{\sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij}-\bar{x}_1, \bar{x}_2, \dots, \bar{x}_K)^2} \right)^{\frac{n}{2}} \subset C$$

$$\frac{\sum \sum (x_{ij} - \bar{x}_{ij})^2 + \sum \sum (x_{is} - \bar{x}_{is})^2}{\sum \sum (x_{ij} - \bar{x}_{ij})^2} \sim \chi^2_{n-2}$$

$$\frac{\sum \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{x}_{ij} - \bar{x})^2}{\sum \sum_{i=1}^k (x_{ij} - \bar{x}_{ij})^2} \rightarrow k \Leftrightarrow \frac{SQT \text{ trat}}{SQR \text{ Residuo}} \rightarrow k \Leftrightarrow \frac{SQT \text{ trat}}{\frac{k-1}{n-k}} \rightarrow k^*$$

$$\begin{array}{lll} SQR_{\text{Res}} & QM_{\text{Res}} & QM_{\text{Res}} + F_{n-2, n-k} \\ SQR_{\text{Res}} & QM_{\text{Res}} & QM_{\text{Res}} \end{array}$$

$$SQT = 0$$

Sob H_0 ,

$$\frac{SQT \text{ tratamentos}}{SQR \text{ Residuos}} \sim F_{k-1, n-k}$$

Aula 18

20/06/2013

Avaliações 2

Teste RVG

$$H_0: \Theta \in \Theta_0$$

$$H_1: \Theta \in \Theta_L$$

RV: $\mathcal{X} \rightarrow \{0,1\}$

$$x \in \mathcal{X} \Rightarrow RV(x) = \frac{\sup_{\theta \in \Theta_0} f(x|\theta)}{\sup_{\theta \in \Theta} f(x|\theta)}$$

$$\Psi(x) = \begin{cases} 1, & RV(x) < c, \\ 0, & \text{c.c.} \end{cases}$$

Exemplo 3. $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$, dado $\Theta = (\Theta_1, \Theta_2)$, c.i. talis que

$$X_i | \Theta \sim \text{Exp}(\Theta_1) \quad i=1, \dots, n_1 \in$$

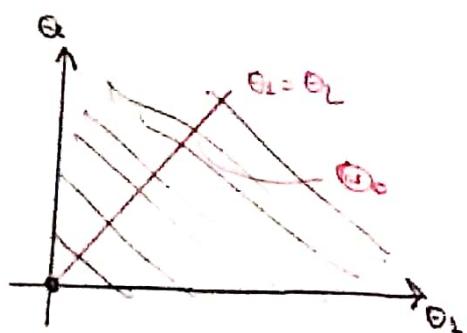
$$Y_j | \Theta \sim \text{Exp}(\Theta_2) \quad j=1, \dots, n_2$$

$$H_0: \Theta_1 = \Theta_2$$

$$\Theta = \mathbb{R}_+^2$$

$$H_1: \Theta_1 \neq \Theta_2$$

$$\mathcal{X} = \mathbb{R}_+^{n_1+n_2} \quad (\mathbb{R}_+^{n_1} \times \mathbb{R}_+^{n_2})$$



$$\mathbf{x} = (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \in \mathcal{X}$$

$$f(\mathbf{x}|\Theta) = \prod_{i=1}^{n_1} \Theta_1 e^{-\Theta_1 x_i} \prod_{i=1}^{n_1} (x_i) \cdot \prod_{j=1}^{n_2} \Theta_2 e^{-\Theta_2 y_j} \prod_{j=1}^{n_2} (y_j)$$

Em Θ ,

$$\sup_{\theta \in \Theta} f(\mathbf{x}|\theta) = f\left(\mathbf{x} \mid \left(\frac{n_1}{\sum x_i}, \frac{n_2}{\sum y_j}\right)\right)$$

Em $\Theta_0: \Theta_1 = \Theta_2 = \Theta$

$$\sup_{\theta \in \Theta_0} f(\mathbf{x}|\theta) = f\left(\mathbf{x} \mid \left(\frac{n_1+n_2}{\sum x_i + \sum y_j}, \frac{n_1+n_2}{\sum x_i + \sum y_j}\right)\right)$$

$$AV(x) = \frac{\sup_{x \in \Omega} f(x/\epsilon)}{\sup_{x \in \Omega} f(x/\delta)} = \frac{\left(\frac{x_1+x_2}{\delta x + \delta y}\right)^{n_1} e^{-\frac{n_1}{\delta x + \delta y}(2x_1+2y_1)}}{\left(\frac{x_1}{\delta x}\right)^{n_1} e^{\frac{n_1}{\delta x}2x} \left(\frac{x_2}{\delta y}\right)^{n_2} e^{-\frac{n_2}{\delta y}2y}}$$

$$= \frac{(x_1+x_2)^{n_1+n_2}}{x_1^{n_1} x_2^{n_2}} \cdot \frac{(\delta x)^{n_1} (\delta y)^{n_2}}{(\delta x + \delta y)^{n_1+n_2}} < c \Leftrightarrow$$

$$\frac{(\sum x_i)^{n_1} (\sum y_i)^{n_2}}{(\sum x_i + \sum y_i)^{n_1+n_2}} < c/c \Leftrightarrow$$

$$\frac{(\sum x_i)^{n_1} (\sum y_i)^{n_2}}{(\sum x_i + \sum y_i)^{n_1+n_2}} < c/c \Leftrightarrow$$

$$\left(\frac{\sum x_i}{\sum x_i + \sum y_i} \right)^{n_1} \left(\frac{\sum y_i}{\sum x_i + \sum y_i} \right)^{n_2} < k \Leftrightarrow$$



$$\Rightarrow \frac{\sum x_i}{\sum x_i + \sum y_i} < \delta_L \text{ on } \frac{\sum x_i}{\sum x_i + \sum y_i} > \delta_U \Leftrightarrow$$

$$\frac{\sum x_i + \sum y_i}{\sum x_i} > \delta_L \text{ on } \frac{\sum x_i + \sum y_i}{\sum x_i} < \delta_U \Leftrightarrow$$

$$\frac{\sum y_i}{\sum x_i} > \frac{1-\delta_L}{\delta_L} \text{ on } \frac{\sum y_i}{\sum x_i} < \frac{1-\delta_U}{\delta_U}$$

$$\Psi(x) = \begin{cases} 1, & \frac{\sum x_i}{\sum y_j} > \frac{1-\delta_1}{\delta_1} \text{ ou } \frac{\sum x_i}{\sum y_j} < \frac{1-\delta_2}{\delta_2} \\ 0, & \text{c.c.} \end{cases}$$

Impomos que, sob H_0 , $\pi_{\Psi}(e) \leq \alpha$.

Para todo $(G_1, G_2) \in \Theta_0$, temos que

$$\frac{\frac{\sum x_i}{n_1}}{\frac{\sum y_j}{n_2}} = \frac{20_1 \sum_{i=1}^{n_1} x_i}{20_2 \sum_{j=1}^{n_2} y_j} \text{ para } \theta \in \Theta_0,$$

$$\frac{\frac{\sum x_i}{n_1}}{\frac{\sum y_j}{n_2}} \rightarrow \begin{cases} 0 & \sim F_{2n_1, 2n_2} \end{cases}$$

Sejam $f_{\alpha/2}$ e $f_{1-\alpha/2}$ percentis de ordem $\frac{\alpha}{2}$ e $1 - \frac{\alpha}{2}$ do $F_{2n_1, 2n_2}$,

$$P\left(f_{\alpha/2} \leq \frac{\sum_{i=1}^{n_1} x_i}{\sum_{j=1}^{n_2} y_j} \frac{n_2}{n_1} \leq f_{1-\alpha/2} / \theta\right) = 1 - \alpha, \quad \forall \theta \in \Theta_0.$$

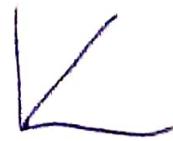
Assim,

$$\Psi(x) = \begin{cases} 1, & \frac{\sum x_i}{\sum y_j} > \frac{n_1}{n_2} f_{1-\alpha/2} \text{ ou } \frac{\sum x_i}{\sum y_j} < \frac{n_1}{n_2} f_{\alpha/2} \\ 0, & \text{c.c.} \end{cases}$$

$\theta \in \Theta_0$

$$\lambda(x) = -2 \log L(x) / \theta \stackrel{\text{aprox}}{\sim} \chi^2_{\mu(\theta_0, \theta)}$$

$Z = L$



$$D(\theta_0) = 1, D(\theta) = 2$$

Víctor Fossaluza (2008).

Testes Bayesianos

		$\theta = \theta_0$	$\theta = \theta_1$
$\theta, \theta \in \Theta_0$	decisão	c	a_2
	α (rejeitar)	a_1	0
$\theta, \theta \in \Theta_1$			

$$\Psi = \{\psi: \mathcal{X} \rightarrow \{0,1\}\}.$$

Vimos que o teste de Bayes é dado por:

$$\psi^*(x) = \begin{cases} 1, & P(\theta = \theta_1 | X=x) > \alpha_{\text{alar}} \\ 0, & \end{cases}$$

Agora, vamos supor que temos hipóteses compostas!

decisão	$\theta \in \Theta_0$	$\theta \in \Theta_1$
0 (não rejeitar)	0	α_2
1 (rejeitar)	α_1	0

Se constrói da mesma forma, mas agora avaliando a probab. a posteriori em toda a região crítica:

No caso geral,

$$H_0: \theta \in \Theta_0 \text{ versus } H_1: \theta \in \Theta_1,$$

Considerando a probabilidade do tipo $P(\theta \in \Theta_1 | x=x)$,

$$\psi^*(x) = \begin{cases} 1, & P(\theta \in \Theta_1 | x=x) > \alpha_1 / \alpha_1 + \alpha_2 \\ 0, & \text{c.c.} \end{cases}$$

Problema: sob modelos abs. contínuos com

um espaço com medida de Lebesgue nula, então $P(\theta \in \Theta_1 | x=x) = 1$ forall.

Solução: considerar outras funções de perda

- Cartinhas S Júlio Stern EBCD
- Inf. Jeffrey
- Atribuir densidade à Θ_0 e outra à Θ_1 e aplicar o teorema de Bayes.

$$P^2 = \varphi(1-p) - (1-p)^2 \rightarrow \text{equilibrium de Mordell}$$

Exemplo 1. $X \sim \text{Unif}(0, \theta)$

$\Theta \sim \text{Unif}(0, 30)$

$$H_0: \theta \leq 20$$

$$H_1: \theta > 20$$

$$x > 0$$

$$f(\theta|x) = \frac{f(x|\theta) f(\theta)}{f(x)} \propto f(x|\theta) f(\theta) =$$

$$= \frac{1}{\theta} \int_{0}^{x} \frac{1}{\theta} \mathbb{I}(x) \frac{1}{30} \mathbb{I}(\theta) \propto \frac{1}{\theta} \int_{0}^{x} \frac{1}{\theta} \mathbb{I}(x) \frac{1}{30} \mathbb{I}(\theta) = \frac{1}{\theta} \int_{0}^{x} \frac{1}{\theta} \mathbb{I}(\theta) \frac{1}{30} \mathbb{I}(\theta) =$$

$$= \frac{1}{\theta} \mathbb{I}_{(x, 30)}(\theta) \Rightarrow f(\theta|x) = \frac{c}{\theta} \mathbb{I}_{(x, 30)}(\theta)$$

$$\int_{-\infty}^{\infty} f(\theta|x) d\theta = 1 \rightarrow c \int_{0}^{30} \frac{1}{\theta} d\theta = 1 \rightarrow c \cdot \log(\theta) \Big|_{0}^{30} = 1 \Rightarrow$$

$$c \log\left(\frac{30}{x}\right) = 1 \Rightarrow c = \frac{1}{\log\left(\frac{30}{x}\right)}$$

Logo,

$$f(\theta|x) = \frac{1}{\log\left(\frac{30}{x}\right)} \frac{1}{\theta} \cdot \mathbb{I}_{(x, 30)}(\theta)$$

$$P(\epsilon \in \mathbb{G}_1 | X=x) = f(\theta > 20 | X=x) = \int_x^\infty f(\theta|x) d\theta =$$

$$= \int_x^\infty \frac{1}{\log(\frac{x}{\theta})} \frac{1}{\theta} I(\theta) d\theta = \begin{cases} \int_x^{20} \frac{1}{\log(\frac{x}{\theta})} \frac{1}{\theta} d\theta = 2 & , x > 20 \\ \int_x^{20} \frac{1}{\log(\frac{x}{\theta})} \frac{1}{\theta} d\theta = \frac{\log(\frac{x}{20})}{\log(\frac{x}{x})} & , x \leq 20 \end{cases}$$

Dados $a_1, a_2 \geq 0$,

para $x > 20$,

$$P(\theta > 20 | X=x) \geq \frac{a_1}{a_1 + a_2} \Leftrightarrow \frac{\log 20 - a_2 \log \frac{20}{x}}{a_1} \geq 0$$

$$\log 20 - a_2 \log \frac{20}{x} \geq 0 \Leftrightarrow a_2 \log 20 - (a_1 + a_2) \log 20 \geq -a_1 \log x \Leftrightarrow$$

$$\frac{\log 20 - \log x}{\log 20 - \log 20} \geq \frac{a_1}{a_1 + a_2} \Leftrightarrow \frac{a_2 \log 20 - (a_1 + a_2) \log 20}{a_2} \geq -a_1 \Leftrightarrow$$

$$\log x \geq \frac{a_2 \log 20 - (a_1 + a_2) \log 20}{-a_1} \Leftrightarrow$$

$$x \geq e^{\frac{(a_1 + a_2) \log 20 - a_2 \log 20}{a_1}}$$

Assim, o teste de Bayes é dado por

$$\psi^*(x) = \begin{cases} 1, & x \geq \min\{20, e^{\frac{(a_1 + a_2) \log 20 - a_2 \log 20}{a_1}}\} \\ 0, & \text{o.c.} \end{cases}$$

Exemplo 2. x_1, \dots, x_n , dado θ , com d. Exp(6)

$\epsilon \sim \text{Gamma}(a, b)$

$x_0 \in \mathbb{R}^3 \quad x = \mathbb{R}^n \quad \alpha = (x_1, \dots, x_n) \in \mathbb{R}^n$

$\theta \in \mathbb{R}^3 \quad f(x_1, \dots, x_n) \propto f(x_1, \dots, x_n | \theta) =$

$$= \prod_{i=1}^n f(x_i | \theta) f(\theta) \propto \theta^a e^{-bx_1} e^{a-1} e^{-b\theta} I_{[0, \infty)}(\theta)$$

$$\Rightarrow f(\theta | x) \propto \theta^{a+n-1} e^{-(b+2x_1)\theta} I_{[0, \infty)}(\theta)$$

$\theta / x = x_1, \dots, x_n \sim \text{Gamma}(a+n, b+2x_1)$

Logo,

$$f(\theta | x) = \begin{cases} 1, & \theta \geq 0, \frac{(b+2x_1)^{a+n}}{\Gamma(a+n)} \theta^{a+n-1} e^{-(b+2x_1)\theta} \text{ se } \theta \geq \frac{x_1}{a+n-2} \\ 0, & \text{o.c.} \end{cases}$$

Testar H_0 -hipóteses

de invés de $\Theta = \Theta_0 \cup \Theta_1$, $\Theta_0 \cap \Theta_1 = \emptyset$, formadas

$$\Theta_0 = \Theta_0 \cup \Theta_1 \cup \dots \cup \Theta_n, \quad \Theta_0 \cap \Theta_1 = \emptyset, i \neq j$$

$$\Rightarrow H_0: \theta \leq 1/3$$

$$H_1: 1/4 \leq \theta \leq 9/10$$

$$H_2: \theta \geq 9/10$$

$$H_3: \theta \geq 7/10$$

Agora

$$\Psi = \{\varphi : \Xi \rightarrow \{1, \dots, k\}\}$$

para $\varphi \in \Psi$, $\varphi(x) = j$ denota por decisão $h_j: \Theta \in \Theta_j$

Decisão	$\theta \in \Theta_1$	$\theta \in \Theta_2$...	$\theta \in \Theta_n$
1	0	h_{12}	...	h_{1n}
2	h_{21}	0	...	h_{2n}
3	h_{31}	h_{32}	...	h_{3n}
...
k	h_{k1}	h_{k2}	...	0

Para $\varphi \in \Psi$, vamos avaliar

$$p(\varphi) = E[L(\varphi(x), \theta)] = \sum_{x \in \Xi} \sum_{\theta \in \Theta} L(\varphi(x), \theta) P(\theta = \theta^*, x = x) =$$

$$= \sum_{x \in \Xi} \sum_{\theta \in \Theta} \left[\sum_{i=1}^k \sum_{m=1}^k h_{ij} I(\theta) I(x) \right] P(\theta = \theta^* / x = x) P(x = x) =$$

$$= \sum_{x \in \Xi} P(x = x) \sum_{i=1}^k \sum_{j=1}^k \sum_{\theta \in \Theta} h_{ij} I(\theta) I(x) P(\theta = \theta^* / x = x)$$

$$= \sum_{x \in \Xi} P(x = x) \left\{ \sum_{i=1}^k I(x) \left[\sum_{j=1}^k h_{ij} \sum_{\theta \in \Theta} I(\theta) P(\theta = \theta^* / x = x) \right] \right\} \Rightarrow$$

$$\Rightarrow p(\varphi) = \sum_{x \in \mathcal{X}} P(X=x) \left\{ \sum_{i=1}^k \mathbb{I}(\omega)_{\varphi^{-1}(i)} \left\{ \sum_{j=1}^k l_{ij} P(\theta \in \Theta_j | X=x) \right\} \right\}$$

↓

Assim, para minimizar p em Ψ , devo tomar $\varphi^*: \mathcal{X} \rightarrow \{1, \dots, k\}$ que a cada $x \in \mathcal{X}$ associa $i^* \in \{1, \dots, k\}$ tal que

$$\sum_{j=1}^k l_{ij^*} P(\theta \in \Theta_j | X=x) = \min_{i \in \{1, \dots, k\}} \sum_{j=1}^k l_{ij} P(\theta \in \Theta_j | X=x)$$

Exemplo 1:

$$\Theta = \{1/4, 1/2, 3/4\} \quad X \sim \text{Geof}(0)$$

$$P(\theta = \frac{1}{4}) = P(\theta = \frac{1}{2}) = \frac{1}{4} \Rightarrow P(\theta = \frac{3}{4}) = \frac{1}{4}.$$

$$\#_1: \theta = 1/4$$

~~1/2, 3/4~~

$$\#_2: \theta = 1/2$$

$$\#_3: \theta = 3/4$$

$$\Psi = \{\varphi: \mathcal{X} \rightarrow \{1, 2, 3\}\} \quad \mathcal{X} = \mathbb{N}^*$$

		Decisão		
		$\theta = 1/4$	$\theta = 1/2$	$\theta = 3/4$
1	1	0	2	2
	2	$3/2$	0	$3/2$
3	2	2	0	

para cada $i \in \{1, 2, 3\}$

$$\sum_{s=1}^3 l_{is} P(\theta = \frac{s}{4} | X=x)$$

$$i=1 \rightarrow l_{11}^{\textcircled{1}} P\left(\theta = \frac{1}{4} | X=x\right) + l_{12}^{\textcircled{2}} P\left(\theta = \frac{2}{4} | X=x\right) + l_{13}^{\textcircled{3}} P\left(\theta = \frac{3}{4} | X=x\right)$$

$$i=2 \rightarrow l_{21}^{\textcircled{1}} P\left(\theta = \frac{1}{4} | X=x\right) + l_{22}^{\textcircled{2}} P\left(\theta = \frac{2}{4} | X=x\right) + l_{23}^{\textcircled{3}} P\left(\theta = \frac{3}{4} | X=x\right)$$

$$i=3 \rightarrow l_{31}^{\textcircled{1}} P\left(\theta = \frac{1}{4} | X=x\right) + l_{32}^{\textcircled{2}} P\left(\theta = \frac{2}{4} | X=x\right) + l_{33}^{\textcircled{3}} P\left(\theta = \frac{3}{4} | X=x\right)$$

MAE 5702 - Probabilidade e Inferência Estatística I	1º semestre/2013
Prof. Luis 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

Dizemos que a estatística $T: \mathcal{E} \rightarrow \mathbb{R}^k$ é completa se $\forall B: \mathbb{R}^k \rightarrow \mathbb{R}$, $E_B(\chi_{\{T(x) \in B\}}) = 0, \forall \theta \in \Theta \Rightarrow h \equiv 0$.

Teorema de Lehmann-Scheffé: Seja $T: \mathcal{E} \rightarrow \mathbb{R}^k$ uma estatística suficiente e completa. Então todos os estimadores não-viesados de $\theta(g(\theta))$ que são funções de T (dependem de x através de T) são R-que $\forall \theta \in \Theta$.³ Assim, se $S^1 \in \Delta^k$ é função de uma est. suf. e completa, S^1 é o ENUVUM para θ .

Desigualdade da Informação (Cramér-Rao). Seja δ um est. não-vies. pl $g(\theta)$. Sob certas condições, vale que $\text{Var}_\theta(\delta(x)) \geq (g'(\theta))^2 / I_x(\theta)$, onde $I_x(\theta)$ é chamada Inf. de Fisher para θ contida na amostra. $I_x(\theta) = E_\theta[(x'(x/\theta))^2]$, $\lambda(x/\theta) = \log f(x/\theta)$ "Toda inf. deve ser baseada apenas naquela amostra $x \in \mathcal{E}$!".

Def. $(\delta_n)_{n \geq 1}$ é uma seq. consistente para $\theta(g(\theta))$ se $\delta_n \xrightarrow{P_\theta} \theta(g(\theta)), \forall \theta \in \Theta$. (P /migrante)

Def. $(\delta_n)_{n \geq 1}$ é Assintoticamente Normal ($0, \sigma^2(\theta)$) se $\sqrt{n}(\delta_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta)), \forall \theta \in \Theta$. ($N(0, \sigma^2(\theta)/n)$)

Res. $(\delta_n)_{n \geq 1}$ é assint. Normal ($0, \sigma^2(\theta)$) $\Rightarrow (\delta_n)_{n \geq 1}$ é consistente para θ .

Def. $(\delta_n^*)_{n \geq 1}$ 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 $(\delta_n)_{n \geq 1}$ assint. normal. $\frac{\lambda(x)}{\theta} | \Theta \sim \text{Beta}(n, 1)$

Def. $(\delta_n)_{n \geq 1}$ e $(\delta_n^*)_{n \geq 1}$ são assintoticamente equivalentes se $\sqrt{n}(\delta_n - \delta_n^*) \xrightarrow{P_\theta} 0, \forall \theta \in \Theta$.

Res. $(S_{MV})_{n \geq 1}$ seq. de EMV pl θ . Sob cond. de regularidade, $\sqrt{n}(S_{MV}^{(n)} - \theta) \xrightarrow{d} N(0, \frac{1}{I_{X_1}(\theta)})$, n.o caso

$(X_{n1})_{n \geq 1}$, dado θ , c.i.i.d. $(S_{MV}^{(n)}) \approx N(\theta, \frac{1}{n I_{X_1}(\theta)})$. $\frac{\sum (x_i - \mu_0)^2}{\theta} | \Theta \sim \chi_{n-1}^2 | \Theta \sim \text{Beta}(n-1, 2)$

Res. (Método Delta). $(\delta_n)_{n \geq 1}$ t.q. $\sqrt{n}(\delta_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$. Seja $f: \mathbb{R} \rightarrow \mathbb{R}$ diferenciável. Então

$\sqrt{n}(g(\delta_n) - g(\theta)) \xrightarrow{d} N(0, \sigma^2(\theta)(g'(\theta))^2)$. [$g(\delta_n)_{n \geq 1}$ é ass. Normal].

Intervalos Aproximados: Sob cond. de reg. $\sqrt{n}(S_{MV} - \theta) \xrightarrow{d} N(0, \frac{1}{I_{X_1}(\theta)})$. Se g. dif. $\sqrt{n}(g(S_{MV}) - g(\theta)) \xrightarrow{d} N(0, (g'(\theta))^2 / I_{X_1}(\theta))$

para "n grande" $P(\Phi^{-1}(\alpha/2) \leq \frac{\sqrt{n}(g(S_{MV}) - \theta)}{(g'(S_{MV}))^2 / I_{X_1}(S_{MV})} \leq \Phi^{-1}(1-\alpha/2) | \theta) = 1-\alpha, \forall \theta \in \Theta$.

Def. Uma função de teste (teste) é uma f. $\Psi: \mathcal{E} \rightarrow \{0, 1\}$ para testar $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_1$ que associa a cada $x \in \mathcal{E}$ $\Psi(x) = 0$ no caso de decisão por H_0 a partir de x ou $\Psi(x) = 1$ no caso de decisão por H_1 . Ao c.g. $\Psi^*(f(x))$

Região Crítica ou R. Rejeição de H_0 e ao c.g. $\Psi^*(f(x))$ Região de Aceitação de H_0 . $\mathcal{U} = \{\Psi: \mathcal{E} \rightarrow \{0, 1\}\}$ é f. m.m.s.

1. decidir por H_2 ($\Psi(x) = 1$) qdo H_0 é vdd + 2. decidir por H_0 ($\Psi(x) = 0$) qdo H_2 é vdd.

Seja $\Psi: \mathcal{E} \rightarrow \{0, 1\}$ um teste para $H_0: \theta \in \Theta_0$, ao número $\pi_{\Psi}(\theta) = P(\Psi(x) = 1 | \theta)$ → Função Poder do Teste!

Ao número $\sup_{\theta \in \Theta_0} \pi_{\Psi}(\theta)$ damos o nome de Tamanho do teste Ψ . ($\alpha_{\Psi} = \sup_{\theta \in \Theta_0} \pi_{\Psi}(\theta)$). Um teste

é d. nível α , se $\alpha_{\Psi} \leq \alpha$. $\{1, \alpha_{\Psi}(x/\theta_0) < b_{\Psi}(x/\theta_0)\}, a, b$ b.c.d.s. Então, $\Psi \in \mathcal{U}$.

Alfaiômetros Simples: Result. Seja $\Psi^*: \mathcal{E} \rightarrow \{0, 1\}$ dada por $\Psi^*(x) = \begin{cases} 1, & a(x/\theta_0) < b(x/\theta_0), \\ 0, & \text{c.c.} \end{cases}$

Def. Um teste d. nível α , $0 < \alpha < 1$, Ψ^* é um teste MAIS PODEROSO (M.P) de nível α para H_0 se $\pi_{\Psi^*}(\theta_1) \geq \pi_{\Psi}(\theta_1), \forall \theta \in \Theta_1$.

Lema de Neyman-Pearson: $H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$. Seja $\Psi^*: \mathcal{E} \rightarrow \{0, 1\}$ dado por $\Psi^*(x) = \begin{cases} 1, & f(x/\theta_0) < k f(x/\theta_1), \\ 0, & \text{c.c.} \end{cases}$

Então, $\pi_{\Psi^*}(\theta_0) \neq \pi_{\Psi^*}(\theta_1), \forall \theta \in \Theta_1$, $\forall \Psi \in \mathcal{U}$. $\alpha_{\Psi^*} = \inf_{\theta \in \Theta_1} \pi_{\Psi^*}(\theta) = \inf_{\theta \in \Theta_1} P(\Psi^*(x) = 1 | \theta) = \inf_{\theta \in \Theta_1} P(\theta = \theta_1 | x = x)$

$\alpha_{\Psi^*}(x) = \begin{cases} 1, & \text{para } x: P(x = x | \theta_0) \leq p_1 \text{ e } P(x = x | \theta_1) \leq p_2 \\ 0, & \text{caso contrário.} \end{cases}$

Definição: O teste $\Psi^* \in \mathcal{U}$ d. nível α , $0 < \alpha < 1$, é um teste UNIFORMEMENTE mais poderoso (UMP) de nível α .

H. Compatíveis: Definição: O teste $\Psi^* \in \mathcal{U}$ d. nível α , $0 < \alpha < 1$, se $\forall \Psi \in \mathcal{U}$ t.q. $\alpha_{\Psi} \leq \alpha$, $\pi_{\Psi^*}(\theta) \geq \pi_{\Psi}(\theta), \forall \theta \in \Theta_1$. "Em geral, é teste UMP".

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ possui razão de verossimilhança monótona crescente (decrescente) em T se $\forall \theta_1, \theta_2 \in \Theta_1$,

e para $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_1$, se $\forall \Psi \in \mathcal{U}$ t.q. $\alpha_{\Psi} \leq \alpha$, $\pi_{\Psi^*}(\theta) \geq \pi_{\Psi}(\theta), \forall \theta \in \Theta_1$.

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cresce (decresce) de $T(x)$ nos possíveis valores de T sob Θ_1 ou Θ_2 .

Def. A família $\mathcal{D} = \{f(x/\theta_1): \theta \in \Theta_1\}$ é f. cres

$$\text{Malcor Aparecido Pinheiro} \quad \frac{b^a}{\Gamma(a)} 0^{a-1} e^{-bx} \prod_{i=1}^n f_{X_i}(x_i) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} (1-x)^{a-1} x^{b-1}$$

Beta(θ), Beta($a + \sum x_i, b + n - \sum x_i$); Exp(θ), Gamma(a, b), Gamma($a+n, b+\sum x_i$).

Definição: O oper transformação $T: \mathcal{X} \rightarrow \mathcal{J}$ mensurável é chamada Estatística.

Definição: $X((x_1, \dots, x_n))$, $T: \mathcal{X} \rightarrow \mathcal{J}$ é uma est. suf. para θ se para toda priori para θ ; $\theta / X = \theta / T(X) \forall \theta \in \Theta$.

Definição: $T: \mathcal{X} \rightarrow \mathcal{J}$ é suf. para θ se a dist. de $X / T(X) = T(x)$, θ não depende de θ para todo $t \in T(\mathcal{X})$ e $\forall \theta \in \Theta$.

Critério da fatoração (Fisher-Neyman) $X((x_1, \dots, x_n))$ amostra.

A estatística $T: \mathcal{X} \rightarrow \mathcal{J}$ é suficiente para θ se, e só se, $\exists g: \mathcal{J} \rightarrow \mathbb{R}_+$ e $\forall t \in \mathcal{J}, \forall \theta \in \Theta$ tal que

$$f(x|\theta) = g(t) \cdot r(T(x), \theta), \forall x, \forall \theta.$$

Resultado: Seja $T: \mathcal{X} \rightarrow \mathcal{J}$ suf. para θ . Seja $S: \mathcal{X} \rightarrow \mathcal{J}'$ outra est. t.q. $S = g \circ T$, onde $g: \mathcal{J} \rightarrow \mathcal{J}'$ é bijetora (um-a-um). Então, S é suficiente para θ .

$T: \mathcal{X} \rightarrow \mathcal{J}$ é suf. Seja $T': \mathcal{X} \rightarrow \mathcal{J}'$ uma outra est. suf. Dizemos que $T: \mathcal{X} \rightarrow \mathcal{J}$ é suf. mínima p/ θ se T' é função de oper. outra estatística suficiente para θ .

Em outras palavras, T é suf. mínima se para todo T' existe $g: \mathcal{J} \rightarrow \mathcal{J}'$ t.q. $T' = g \circ T$. Quando, T é suf. mínima para todo T' suf. 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 $T: \mathcal{X} \rightarrow \mathcal{J}$ uma estatística. Se $\forall x_1, x_2 \in \mathcal{X}, f(x_1|\theta) \text{ não depende de } \theta \Leftrightarrow T(x_1) = T(x_2)$,

então T é suf. mínima.

Uma estat. $T: \mathcal{X} \rightarrow \Theta$ é chamada Estimador p/ θ . [Como escolher? Devermos estabelecer critério de penalidades]

Idéia: Escolher um estimador que minimize os critérios de média. $p(\delta) := E[L(\delta(x), \theta)] = \sum_{x \in \mathcal{X}} \sum_{\theta \in \Theta} L(\delta(x), \theta) p(x|\theta)$

$L: \Theta \times \Theta \rightarrow \mathbb{R}_+$ penalidade (perda) q/ o decisior incorre ao estimar θ por $\delta(x)$ no obs $x \in \mathcal{X}$.

$(x_1, \theta) \rightarrow L(\delta(x_1, \theta)) \Rightarrow$ Idéia: Escolher $\delta^* \in \Delta$ t.q. $p(\delta^*) = \min \{p(\delta) : \delta \in \Delta\}$

fazemos recreser $p(\delta) = \sum_{x \in \mathcal{X}} \sum_{\theta \in \Theta} L(\delta(x), \theta) P(\theta=x|X=x) P(x=x) = \sum_{x \in \mathcal{X}} P(x=x) \{ \sum_{\theta \in \Theta} L(\delta(x), \theta) P(\theta=x|X=x) \}$. Idéia: Para

cada $x \in \mathcal{X}$, formamos $\delta^*(x) \in \Delta$ que minimiza $\sum_{\theta \in \Theta} L(\delta(x), \theta) P(\theta=x|X=x)$ em(d) $\delta^*(x)$ - estimador de Bayes.

"P/ cada $x \in \mathcal{X}$ constrói-se o. dist. a posteriori e encontramos o valor de δ que minimiza os riscos a post."

1. $L(d, \theta) = (d-\theta)^2$, $d^* = E(\theta|x=x)$ ou t.q. $d^* = \min_{\theta \in \Theta} |d - E(\theta|x=x)|$.

2. $L(d, \theta) = |d - \theta|$, $\delta^*(x)$ mediana($\theta|x$) ; 3. Perda $\theta - \delta$, $\delta^*(x) = \text{meda}(\theta|x)$

Est. M. Verossi: $V_d(\cdot, \cdot) = R$

$\delta_{MV}(\theta) = \arg \min_{\theta \in \Theta} V_d(\theta)$

1. T é uma est. suf. c/única e única $\Rightarrow \delta_{MV}$ é função de T .

2. O do δ_{MV} é uma est. suf. c/única $\Rightarrow \delta_{MV}$ é suf. mínima.

3. Invariância dos EMV: Seja $g: \Theta \rightarrow \mathcal{J}$ ($-g(\theta)$), $g \circ \delta_{MV}$ é o EMV para $g(\theta)$.

δ_{MV} é o EMV para $\theta \Rightarrow g(\delta_{MV})$ é o EMV para $g(\theta)$.

Método mnm. $E(Y_1/\theta) = Y_1 + \dots + Y_n/n$

$E(Y_1^2/\theta) = Y_1^2 + \dots + Y_n^2/n$

Ed. Mínimos Quadrados. Idéia: Escolhe p/ cada

$x \in \mathcal{X}$, como estimativa o valor que minimiza a

"diferença" entre o valor obs. e esperado. P/ cada

$x \in \mathcal{X}$, formamos $\delta_{MQ}(x)$ t.q.

$$\sum_{i=1}^n (x_i - E(x_i|\delta_{MQ}(x)))^2 = \min_{\theta \in \Theta} \sum_{i=1}^n (x_i - E(x_i|\theta))^2$$

$E(x_i|\theta)$

Comp. (Av.) do Estimadores: Seja $\delta \in \Delta$. Aqde $\sum_{x \in \mathcal{X}} L(\delta(x), \theta) P(X=x|\theta) (\sum_{x \in \mathcal{X}} L(\delta(x), \theta) f(x|\theta) dx)$ chamada de

func. Frequentista do estimador δ para estimar o parâmetro qdo esse vale θ . Not: $R(\delta, \theta) = E[L(\delta(x), \theta)|\theta] =$

$E_\theta [L(\delta(x), \theta)]$, é p.m. $\delta \in \Delta$, $R_2: \Theta \rightarrow \mathbb{R}$ é chamada FUNÇÃO DE RISCO do estimador δ . Qdo $\theta \in \Theta \Rightarrow R_2 = R(\delta, \theta)$

$L(d, \theta) = (d-\theta)^2$, $R(\delta, \theta) = E_\theta [(E(x)-\theta)^2]$ é chamada ERRO QUADRÁTICO MÉDIO de δ -o.

Def. Dizemos que um estimador $\delta^* \in \Delta$ qd $\delta \in \Delta$: $E_\theta [\delta(x)] = \theta, \forall \theta \in \Theta$ é não-viésado de var. uniformemente

menor (ENVM) se $\delta^* \in \Delta$: $\text{Var}_\theta(\delta^*) \leq \text{Var}_\theta(\delta), \forall \theta \in \Theta$.

Teorema de Fins-Steckatti: Seja $\delta: \mathcal{X} \rightarrow \Theta$ um est. fi-mesado p/ $\theta(g(\theta))$. Seja T uma est. suf. p/ θ . Seja $\delta^* = E(\delta|T)$ entõo 1. δ^* depende de T ; 2. $\delta^* \in \Delta$; 3. $\text{Var}_\theta(\delta^*) \leq \text{Var}_\theta(\delta_\theta), \forall \theta \in \Theta$

Def. Uma função $X: \Omega \rightarrow \mathbb{R}$ é chamada SIMPLÉ se assume apenas um nº finito de valores ($X(\omega) = \{X(\omega); \omega \in \Omega\}$ é finito) $X = \sum_{j=1}^k q_j A_j$
 $\int_X X dP = \sum_{j=1}^k q_j P(A_j)$

Def. Uma classe \mathcal{F} de subconjuntos de Ω ($\Omega \neq \emptyset$) é chamada σ-algebra se

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

$$2. A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$$

$$3. (A_n)_{n \geq 1}, \text{ t. q. } A_n \in \mathcal{F}, \forall n \geq 1 \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$$

$$\sigma(\mathcal{F}) = \bigcap_{S \in \mathcal{F}(\mathcal{F})} S$$

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

$$1. P(\Omega) = 1$$

$$2. (A_n)_{n \geq 1}, \text{ t. q. } A_n \in \mathcal{F}, \forall n \geq 1, A_i \cap A_j = \emptyset, i \neq j, P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

(An) é quaisquer

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

Def. A_1, \dots, A_n são (coletivamente) independentes se

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

A_1, A_2 ind. se $A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2, P(A_1 \cap A_2) = P(A_1)P(A_2)$

Def. A trans. $X: \Omega \rightarrow \mathbb{R}$ é uma v.a. se $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}, \forall B \in \mathcal{B}(\mathbb{R})$ ($\mathcal{B}(\mathbb{R}^+)$)

DIRECHILET (a_0, a_1, \dots, a_k)

$$f_X(t_1, \dots, t_k) = \frac{\Gamma(a_0 + \dots + a_k)}{\Gamma(a_0)\Gamma(a_1) \dots \Gamma(a_k)} t_1^{a_1-1} \dots t_k^{a_k-1} (1-t_1 - \dots - t_k)^{a_0-1} \prod_{i=1}^k t_i$$

indep. v.a.

$$P(Y_i \in B_i, \dots, Y_k \in B_k) = \prod_{i=1}^k P(Y_i \in B_i)$$

$B_i \in \mathcal{B}(\mathbb{R}), \forall i = 1, \dots, k$

Sendo $X: \Omega \rightarrow \mathbb{R}$ uma v.a.

$$P_{X|B}: \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$$

$$A \in \mathcal{B}(\mathbb{R}) \Rightarrow P_{X|B}(A) = \frac{P(X^{-1}(A) \cap B)}{P(B)} = \frac{P((X \in A) \cap B)}{P(B)}$$

Dist. cond. de X dado B

$$\text{Var}(Y) = E(\text{Var}(Y|B)) + \text{Var}(E(Y|B))$$

Analogamente

$$F_{X|B}: \mathbb{R} \rightarrow [0, 1]$$

$$t \in \mathbb{R} \Rightarrow P_{X|B}((-\infty, t])$$

F.v. dis. condicional de X dado B

Def. A esperança condicional de X dado B , $E(X|B)$ é uma v.a. t. q.

$$1. E(X|B)$$
 é B -mensurável

$$2. \int_A X dP = \int_A E(X|B) dP, \forall A \in \mathcal{F}$$

Em geral, se A_1, \dots, A_n é uma partição de Ω ($A_i \in \mathcal{F}$)
 $\in \mathcal{G} = \sigma(A_1, \dots, A_n)$

então uma f. \mathcal{G} -mensurável assumindo no máximo n valores.

$$\text{Desse modo, } E(X|B) = \sum_{i=1}^n q_i \mathbb{I}_{A_i}, q_i = \frac{\int_A X dP}{P(A_i)}$$

Definição. Dizemos que a sequência $(X_n)_{n \geq 1}$ converge quase certamente para $X: \Omega \rightarrow \mathbb{R}$ se $\mathbb{P}(\{\omega \in \Omega : X_n(\omega) \rightarrow X(\omega)\}) = 1$.

$$\{\omega \in \Omega : X_n(\omega) \rightarrow X(\omega)\} = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{\omega : |X_k(\omega) - X(\omega)| < \frac{1}{m}\} \text{ pelo}$$

$$X_n(\omega_0) \rightarrow X(\omega_0) \Leftrightarrow \forall \epsilon > 0, \exists n \in \mathbb{N} \text{ t.q. } k \geq n \quad |X_k(\omega_0) - X(\omega_0)| < \epsilon \Rightarrow \frac{1}{m}$$

$$\forall m \geq 1, \omega_0 \in \liminf_{n \rightarrow \infty} \{\omega : |X_n(\omega) - X(\omega)| < \frac{1}{m}\}$$

$$\omega_0 \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{\omega : |X_n(\omega) - X(\omega)| < \frac{1}{m}\} \Rightarrow \omega_0 \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{\omega : |X_k(\omega) - X(\omega)| < \frac{1}{m}\}$$

$$\begin{aligned} & X_n \xrightarrow{q.c.} X \\ & g: \Omega \rightarrow \mathbb{R} \text{ contínua} \quad \left\{ \begin{aligned} & \Rightarrow g(X_n) \xrightarrow{q.c.} g(X) \\ & \text{g mensurável} \end{aligned} \right. \end{aligned}$$

Resultado:

$$X_n \xrightarrow{q.c.} X \Leftrightarrow \mathbb{P}(\limsup \{|X_n - X| \geq \epsilon\}) = 0, \forall \epsilon > 0$$

Res.

$(\Omega, \mathcal{F}, \mathbb{P})$ (An) $n \geq 1$ seq de eventos de \mathcal{F}

$$(1) \sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \Rightarrow \mathbb{P}(\limsup A_n) = 0$$

$$(2) \sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \text{ e os } A_n's \text{ são independentes} \Rightarrow \mathbb{P}(\limsup A_n) = 1$$

$X_1, X_2, \dots : \Omega \rightarrow \mathbb{R}$ dizemos que $(X_n)_{n \geq 1}$ converge em probabi. pl. $X, X_n \xrightarrow{P} X$ se
 $X: \Omega \rightarrow \mathbb{R}$ $\forall \epsilon > 0, (\forall m \in \mathbb{N}^*) \lim_{n \rightarrow \infty} \mathbb{P}(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}) = 0, \forall \epsilon > 0, \frac{1}{m}$

Desigualdade de Markov

X v.a. n negativa. Então pl todos $a \geq 0, \mathbb{P}(X \geq a) \leq E(X)/a$

Tchebychev

X v.a. tal que $E(X) < \infty, \mathbb{P}(|X - E(X)| \geq a) \leq \text{Var}(X)/a^2$

X_1, X_2, \dots v.a com f.d. F_{X_1}, F_{X_2}, \dots , respectivamente

Seja X v.a v.a com f.d. F_X . Dizemos que $(X_n)_{n \geq 1}$ converge em dist. para X (ou $(F_{X_n})_{n \geq 1}$ é conv. frac. para F_X) se

$$F_{X_n}(x) \xrightarrow{n \rightarrow \infty} F_X(x), \forall x \in C_X, \text{ onde } C_X \text{ é}$$

o conjunto dos pontos de continuidade de F_X

Resultado. Sejam X_1, X_2, X_3, \dots v.a. discretas assumindo valores em \mathbb{N} .

Sejam $(p_{1(k)})_{k \in \mathbb{N}}, (p_{2(k)})_{k \in \mathbb{N}}, \dots$ as f.d. de probabilidades de X_1, X_2, \dots , resp

$$X_n \xrightarrow{d} X \Leftrightarrow p_{n(k)} \xrightarrow{n \rightarrow \infty} p(k), \forall k \in \mathbb{N}$$

Aula 1

Conjunto: coleção de objetos distintos. Denotaremos por letras maiúsculas A, B, C, ..., com exceção do espaço armazém, que será denotado por \mathbb{N} (também denotado por S por alguns autores).

Relação de pertinência: $w \in A$, $w \notin A$ conjunto vazio: \emptyset $A^c = \{w \in \Omega : w \notin A\}$

$\bigcup_{x \in \Delta} A_x$

$\bigcap_{x \in \Lambda} A_x$

$$A \Delta B = (A \cap B^c) \cup (A^c \cap B) = \{x \in \Omega : x \in A \cup B \text{ e } x \notin A \cap B\}$$

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

Seq. de conjuntos: A_1, A_2, \dots

$(A_n)_{n \in \mathbb{N}}$

monótona se $A_n \subseteq A_{n+1}$ ou $A_n \supseteq A_{n+1}$, $\forall n \in \mathbb{N}$

Se $A_n \subseteq A_{n+1}, \forall n \in \mathbb{N}$, $A^\uparrow \rightarrow \exists \lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n$

Se $A_n \supseteq A_{n+1}, \forall n \in \mathbb{N}$, $A^\downarrow \rightarrow \exists \lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$

$$\liminf_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

$$\circ \limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

Aula 2

Definição. Dizemos que uma classe \mathcal{F} de sub. de Ω é uma álgebra se

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

Consequências

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

Def. 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 $A_n \in \mathcal{F}, \forall n \geq 1 \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Comentário: $A_1, A_2, \dots, A_n, \dots \in \mathcal{F} \Rightarrow \bigcap_{n=1}^{\infty} A_n^c \in \mathcal{F}$

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

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

\mathcal{F}_1 e \mathcal{F}_2 duas σ-álg. de sub. de Ω , $\mathcal{F}_1 \cap \mathcal{F}_2$ é σ-álg. de sub. de Ω .

Para uma dada classe \mathcal{G} de sub. de Ω , seja

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

$\mathbb{F}(\mathcal{G}) \neq \emptyset$, pois $\mathcal{D}(\Omega) \in \mathbb{F}(\mathcal{G})$

$\bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{G})} \mathcal{F} \subseteq \mathcal{G}$, $\forall \mathcal{F} \in \mathbb{F}(\mathcal{G}) \Rightarrow \bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{G})} \mathcal{F}$ é a menor σ-álgebra que contém \mathcal{G}

$$\sigma(\mathcal{G}) = \bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{G})} \mathcal{F}$$

Límite superior e inferior de uma sequência de conjuntos

Em algumas situações, sobretudo na teoria da medida, é conveniente definir os conceitos de limite superior e inferior para uma sequência de conjuntos. Se E_n é uma sequência de conjuntos, então define-se:

- O limite superior é o conjunto formado por todos os elementos que pertencem a uma infinitade de conjuntos E_n .
- O limite inferior é o conjunto formado por todos os elementos que pertencem a cada um dos E_n exceto por um número finito deles.

$$\limsup_{n \rightarrow \infty} E_n = \overline{\lim}_{n \rightarrow \infty} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k \quad \{A_n \text{ infinitas vezes}\}$$

$$\liminf_{n \rightarrow \infty} E_n = \underline{\lim}_{n \rightarrow \infty} E_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k \rightarrow$$

$\hookrightarrow \{A_n \text{ pt todo } n \text{ suf. grande}\}$

Uma vez que

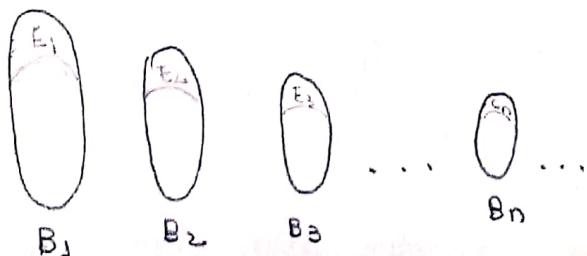
$w \in \liminf_{n \rightarrow \infty} E_n$

pertence a pelo menos uma das intersecções. Uma vez encontra a primeira intersecção que o contém, as posteriores também o conterão. Mas pertencer às intersecções é equivalente a pertencer a todos os A_n 's com exceção de um número finito elas, número esse equivalente ao número dos A_i 's desconsiderados nesse prenúcleo ou i^{a} intersecção que constitui $\limsup_{n \rightarrow \infty} E_n$.

$$\limsup_{n \rightarrow \infty} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k \Rightarrow \text{conjunto que pertence a todas as uniões} \Rightarrow$$

$$w \in \limsup_{n \rightarrow \infty} A_n \Leftrightarrow \forall n \in \mathbb{N}, \exists k \geq n \text{ t.q. } w \in A_k \\ \Leftrightarrow |\{n : w \in A_n\}| = \infty$$

Seja $B_i = \bigcup_{k=i}^{\infty} E_k$. Ilustrando os termos:



$w \in B_i \Rightarrow w \in E_i$, para algum $i \in \mathbb{N}^*$. Encontrando o i^{a} E_k , t.q. $w \in E_k$, notemos que:
 $w \in E_{k+1} \Rightarrow w \in E_k$, para algum $i = k+1, \dots \Rightarrow w \in E_i \neq E_k$. Nisso encontramos $k > k$ t.q. $w \in E_k$. E assim sucessivamente!

Considere $\{a_n\}_{n \geq 1}$ uma sequência de números reais. Defina a seq. auxiliar:

$$A_N = \sup_{n \geq N} a_n$$

OU-seja,

$$A_1 = \sup_{n \geq 1} a_n = \sup \{a_1, a_2, \dots\}$$

$$A_2 = \sup_{n \geq 2} a_n = \sup \{a_2, a_3, \dots\}$$

$$\vdots$$

$$A_N = \sup_{n \geq N} a_n = \sup \{a_{N+1}, a_{N+2}, \dots\}$$

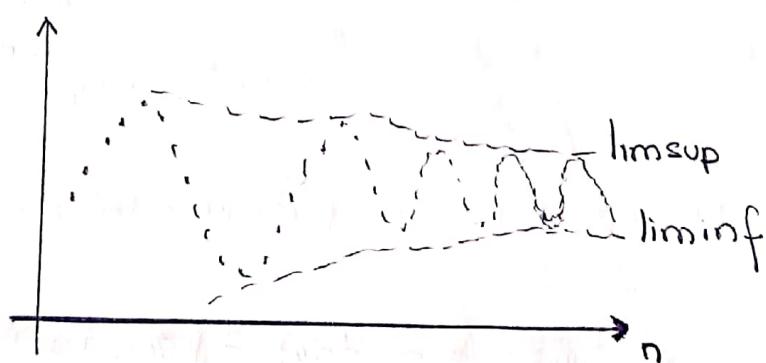
Um elemento $\in S$ é dito supremo de S se for o menor dos majorantes.

A sequência A_n é claramente não-crescente, pois é o supremo de uma família cada vez menor de números reais. Por ser uma sequência monótona, seu limite existe (podendo ser infinito se cada A_n for infinito) e o ínfimo da sequência

$$\limsup_{n \rightarrow \infty} a_n = \underline{\lim}_{n \rightarrow \infty} a_n = \inf_{n=1}^{\infty} \sup_{n \geq N} a_n = \lim_{N \rightarrow \infty} \sup_{n \geq N} a_n$$

E, de forma perfeitamente análoga, se define o limite superior.

$$\liminf_{n \rightarrow \infty} a_n = \underline{\lim}_{n \rightarrow \infty} a_n = \sup_{n=1}^{\infty} \inf_{n \geq N} a_n = \lim_{N \rightarrow \infty} \inf_{n \geq N} a_n$$



Def. Dizemos que uma classe \mathcal{F} de subconjuntos de Ω é uma álgebra se

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

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

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

$\mathbb{F}(\mathcal{F})$: conjunto de todos os σ -álgebras de Ω que contém \mathcal{F} , com \mathcal{F} um conjunto de subconjuntos de Ω .

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

$\sigma(\mathcal{F}) = \bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{F})} \mathcal{F}$: σ -álgebra gerada por \mathcal{F} ou menor σ -álgebra gerada por \mathcal{F} .

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

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

1. $P(\Omega) = 1$

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

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

(Ω, \mathcal{F}, P) - espaço de probabilidade

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

Def. (Ω, \mathcal{F}, P) BGS, $P(B) \geq 0$.

↳ A prob. condicional de A dado B é definida por

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

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

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

Fórmula da Probabilidade Total

seja A_1, \dots, A_n uma partição de Ω t.q. $A_i \in \mathcal{F}, \forall i = 1, \dots, n$. ($A_i \cap A_j = \emptyset, \forall i \neq j$).

seja $B \in \mathcal{F}$

$$\begin{aligned} P(B) &= P(B \cap \Omega) = P(B \cap \bigcup_{i=1}^n A_i) = P\left(\bigcup_{i=1}^n (B \cap A_i)\right) = \sum_{i=1}^n P(B \cap A_i) \Rightarrow \\ P(B) &= \sum_{i=1}^n P(B|A_i) P(A_i) \end{aligned}$$

A e B são ind. $\Leftrightarrow P(A \cap B) = P(A)P(B)$

Teo. de Bayes. (Form. da Inversão)

(Ω, \mathcal{F}, 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,

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

Def. A_1, \dots, A_n são (coletivamente) independentes se

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

Def. A_1, \dots, A_n são independentes dois a dois se

$$\forall B \subseteq \{1, \dots, n\} \text{ t. q. } |B|=2, P(\bigcap_{j \in B} A_j) = \prod_{j \in B} P(A_j)$$

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

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

(Ω, \mathcal{F}, P)

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

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

Def. (Ω, \mathcal{F}, 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

$$\forall A_1 \in \mathcal{F}_1, \forall A_2 \in \mathcal{F}_2, P(A_1 \cap A_2) = P(A_1)P(A_2)$$

Def. A transformação $X: \Omega \xrightarrow[\mathbb{R}^k]{} \mathbb{R}$ é uma variável aleatória se

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

Imagem inversa de B pela transf. X .

Def. Seja X uma v.a.

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

Propriedades de F_X :

1. F_X é não-decrescente
 $t_1 \leq t_2 \Rightarrow (-\infty, t_1] \subseteq (-\infty, t_2] \Rightarrow P_X((-\infty, t_1]) \leq P_X((-\infty, t_2]) \Rightarrow F_X(t_1) \leq F_X(t_2)$

2. F_X é contínua à direita

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

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

Seja $(t_n)_{n \geq 1}$ decrescente, t.q. $t_n \downarrow t \Rightarrow$

$\Rightarrow ((-\infty, t_n])_{n \geq 1}$ é monótona decrescente e, portanto,

$$\lim_{n \rightarrow \infty} P_X((-\infty, t_n]) = P_X(\lim_{n \rightarrow \infty} (-\infty, t_n]) \xrightarrow{\text{definição}} \lim_{n \rightarrow \infty} F_X(t_n) \leftarrow F_X(t) \\ = (-\infty, t]$$

3. Analogamente,

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

Def. Seja $X: \Omega \rightarrow \mathbb{R}^k$ um vetor aleatório que induz a medida $P_X: B(\mathbb{R}^k) \rightarrow \mathbb{R}_+$.

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

$$P_X((-\infty, t_1] \times \dots \times (-\infty, t_k])$$

damos o nome de função dist. (cota) de $X = (X_1, \dots, X_n)$

prop.

1. F_X é não-decrescente em cada componente

2. F_X é contínua à direita em cada componente

3. $\lim_{t_i \rightarrow -\infty} F_X(t_1, \dots, t_k) = 0$ e $\lim_{t_i \rightarrow \infty} F_X(t_1, \dots, t_k) = 1$
 para algum i para todos

4. F_X deve ser tal que $P_X((a_1, b_1] \times (a_2, b_2] \times \dots \times (a_k, b_k])$ (expressão em funções de F_X) deve ser ≥ 0 .

I. Vectors Disc.

Def. $X: \Omega \rightarrow \mathbb{R}^K$ é discreto se $\exists \{x_n\}_{n \in \mathbb{N}}, x_n \in \mathbb{R}^K, \forall n \geq 1$, t.q.

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

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

$$P_X(A) = \sum_{n=1}^{\infty} p_n I_A(x_n)$$

Def. X é v.a. contínua se F_X é cont.

Def. X é um vetor aleatório ABSOLUTAMENTE CONTÍNUO se

$\exists f_X: \mathbb{R}^K \rightarrow \mathbb{R}$ tal que

$$F_X(t_1, \dots, t_K) = \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \dots \int_{-\infty}^{t_K} f_X(x_1, \dots, x_K) dx_1 dx_2 \dots dx_K$$

Nesse caso, vale que $\frac{\partial^K F_X(t_1, \dots, t_K)}{\partial t_1 \partial t_2 \dots \partial t_K} = f_X(t_1, \dots, t_K)$.

Def. Dizemos que X_1, \dots, X_K são independentes se

$$P(X_1 \in B_1, X_2 \in B_2, \dots, X_K \in B_K) = \prod_{i=1}^K P(X_i \in B_i), \forall B_i \in \mathcal{B}(\mathbb{R}), \forall i = 1, \dots, K$$

Res: X_1, \dots, X_K são independentes $\Leftrightarrow F_X(t_1, \dots, t_K) = F_{X_1}(t_1) \dots F_{X_K}(t_K), \forall (t_1, \dots, t_K) \in \mathbb{R}^K$.

Res: X_1, \dots, X_K são independentes $\Leftrightarrow P(X_1 = x_1, \dots, X_K = x_K) = P(X_1 = x_1) \dots P(X_K = x_K), \forall x_i \in X_i(\mathbb{R}), \forall i = 1, \dots, K$

Aula 3

(Ω, \mathcal{F}) dá-se o nome espaço mensurável.

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

1. $P(\Omega) = 1$

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

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

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

$$P: \mathcal{F} \rightarrow \mathbb{R}, \text{ dada por}$$

$P(A) = \mathbb{I}_A(\omega_0)$ é prob. (med. de prob. degenerada/concentrada em ω_0)

(Ω, \mathcal{F}, P) dá-se o nome de espaço de probabilidade.

Propriedades

- i. $P(\emptyset) = 0$

$$\Omega = \omega_0 \cup \emptyset \cup \omega_0 \cup \dots \quad P(\omega_0 \cup \emptyset \cup \omega_0 \cup \dots) = P(\emptyset) + \sum_{n=2}^{\infty} P(\emptyset) \Rightarrow P(\emptyset) = 0.$$

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

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

$$B_i = \emptyset, i = n+1, \dots, \quad P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) \Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

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

- iii. $P(A^c) = 1 - P(A)$.

$$P(A \cup A^c) = P(A) + P(A^c) \Rightarrow P(A^c) = 1 - P(A).$$

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

$$P(B) = P((A \cap B) \cup (A^c \cap B)) = P(A \cap B) + P(A^c \cap B) \Rightarrow$$

$$P(B) = P(A) + P(A^c \cap B) \Rightarrow P(B) \geq P(A)$$

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

Vamos provar.

$$(\emptyset \subseteq A \subseteq \Omega) \quad P(\emptyset) \geq P(A) \geq P(\Omega) \Rightarrow 1 \geq P(A) \geq 0 \Leftrightarrow 0 \leq P(A) \leq 1.$$

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

Def.

$$B_1 = A_1$$

$$B_2 = A_1^c \cap A_2$$

:

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

Note

$$B_i \cap B_j = \emptyset, i \neq j \quad \cup B_n = \cup A_n$$

$$\begin{aligned} \text{Assim} \quad P(\cup A_n) &= P(\cup B_n) = \sum P(B_n) = P(A_1) + \overbrace{\sum_{n=2}^{\infty} P(A_1^c \cap \dots \cap A_{n-1}^c \cap A_n)} \leq \\ &\leq P(A_n) \end{aligned}$$

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

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

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

$$\begin{aligned} P(A_1 \cup A_2) &= P(A_1) + P(A_1^c \cap A_2) \\ &\quad \vdots \\ &P(A_2) = P(A_1 \cap A_2) \end{aligned}$$

Em geral

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

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

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

$$P\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n) \quad (\text{cont. da probabilidade})$$

$$\lim_{n \rightarrow \infty} = \bigcup A_n \quad \begin{aligned} B_1 &= A_1 \\ B_2 &= A_1^c \cap A_2 \end{aligned}$$

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

$$P\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n)$$

$$\begin{aligned} P\left(\lim_{n \rightarrow \infty} A_n\right) &= P(\cup A_n) = P(\cup B_n) = \sum P(B_n) = \lim_{K \rightarrow \infty} \sum_{n=1}^K P(B_n) = \\ &= \lim_{K \rightarrow \infty} [P(B_1), P(B_2), \dots, P(B_K)] = \lim_{K \rightarrow \infty} P(A_K) \end{aligned}$$

Aula 4

Modelos discretos e contínuos

Aula 5

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

Definição. A prob. condicional de A dado B é definida por

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

Da definição,

$$P(A \cap B) = P(B)P(A|B)$$

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

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

Fórmula da Probabilidade Total.

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

seja $B \in \mathcal{F}$

$$\begin{aligned} P(B) &= P(B \cap \Omega) = P(B \cap (\bigcup A_i)) = \\ &= P(\bigcup (B \cap A_i)) = \sum_{i=1}^n P(B \cap A_i) \Rightarrow \\ P(B) &= \sum_{i=1}^n P(B|A_i)P(A_i) \end{aligned}$$

Téorema de Bayes.

(Ω, \mathcal{F}, 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

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

Qdo

$P(A|B) = P(A)$, dizemos que o ev. A é indep. do ev. B.

Vimos que

$$P(A \cup B) = P(B|A)P(A)$$

$\Leftrightarrow A \text{ ind. de } B$

$$P(A)P(B) = P(B|A)P(A) \Rightarrow$$

$$P(B|A) = P(B) \Rightarrow B \text{ ind. de } A$$

Qdo A é ind. de B (e, cons., B é ind. de A), dizemos que A e B são independentes

Alternativamente,

$$A \text{ e } B \text{ são ind.} \Leftrightarrow P(A \cap B) = P(A)P(B)$$

consequências

$A \text{ e } B$ independentes $\Rightarrow A \text{ e } B^c$ são independentes

$A^c \text{ e } B$ são independentes

$A^c \text{ e } B^c$ são ind.

$$P[(B \cap A^c) \cap D] = P(B \cap D) - P(A \cap D)$$
$$= P(B)P(D) - P(A)P(D) = P(B)(P(D) - P(A))$$

, $A \in B^c$

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

, $A^c \in B$

$$P(A^c \cap B) = P(B) - P(B \cap A) = P(B) - P(B)P(A) = P(B)(1 - P(A)),$$

, $A^c \in B^c$

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - P(A) - P(B) + P(A \cap B) = 1 - P(A) - P(B) + P(A)P(B) =$$

$$= 1 - P(A) - P(B)(1 - P(A)) = (1 - P(A))(1 - P(B)) = P(A^c)P(B^c)$$

Agulhamos gen.

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

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

no de igualdades a serem verif.: $2^n - \binom{n}{1} \cdot \binom{n}{0}$

Def. A_1, \dots, A_n são independentes dois a dois

$$\forall B \subseteq \{1, \dots, n\} \text{ t. q. } |B|=2, P(\bigcap_{j \in B} A_j) = \prod_{j \in B} P(A_j)$$

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

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

(Ω, \mathcal{F}, P)

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

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

$$\forall \mathcal{K} \subseteq \mathcal{G}, \mathcal{K} \text{ finito}, P\left(\bigcap_{A \in \mathcal{K}} A\right) = \prod_{A \in \mathcal{K}} P(A)$$

Definição: (Ω, \mathcal{F}, P)

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

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

$$\forall A_1 \in \mathcal{F}_1, \forall A_2 \in \mathcal{F}_2, P(A_1 \cap A_2) = P(A_1)P(A_2)$$

Espaços Produtos

$(\Omega_1, \mathcal{F}_1, P_1)$

$(\Omega_2, \mathcal{F}_2, P_2)$

:

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

$(\Omega_n, \mathcal{F}_n, P_n)$ $\mathcal{E} = \{E_1 \times E_2 \times \dots \times E_n : E_i \in \mathcal{F}_i, i=1, \dots, n\}$

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

$P : \mathcal{F} \rightarrow \mathbb{R}_+$ de modo que $\forall E_1 \times \dots \times E_n \in \mathcal{E}$,

$$P(E_1 \times \dots \times E_n) = P(E_1) \cdots P(E_n)$$

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

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

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

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

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

Definição. A transformação $X: \Omega \rightarrow \mathbb{R}^k$ é uma variável aleatória se e somente se

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

Imagem inversa de B pela transformação X .

Resultado:

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

(\Rightarrow) imediatamente, $B = (-\infty, t] \Rightarrow X^{-1}(B) \in \mathcal{F}$

(\Leftarrow) Seja $\mathcal{F} = \{(-\infty, t], t \in \mathbb{R}\}$

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

\mathcal{F} é σ -alg. de subconjuntos de \mathbb{R}

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

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

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

complemento da pré-imagem é a pré-imagem do complemento

$$\Rightarrow X^{-1}(A^c) \in \mathcal{F} \Rightarrow A^c \in \mathcal{B}(\mathbb{R}) \quad X(\omega) \in A^c \Rightarrow X(\omega) \notin A$$

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

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

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

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

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

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

(\Leftarrow)

$$(-\infty, t] \in \mathcal{F} \Rightarrow \mathcal{F} \subseteq \mathcal{F} \Rightarrow \sigma(\mathcal{F}) \subseteq \mathcal{F} \Rightarrow \mathcal{B}(\mathbb{R}) \subseteq \mathcal{F}$$

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

todas as pré-imagens
dos boreianos

todas as pré-imagens
de \mathcal{F}

Então, por def., X é v.a.

Aula 7

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

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

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

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

$$X^{-1}((-\infty, t_1], \dots, (-\infty, t_K]) \in \mathcal{F}, \forall t_1, \dots, t_K \in \mathbb{R}^K$$

X uma v.a.

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

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

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

Ex.

$$V(\omega) = \zeta$$

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

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

$$D_X: \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}_+$$

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

$$1. D_X(\mathbb{R}) = P_X(\{X \in \mathbb{R}\}) = P(\Omega) = 1$$

$$2. \{B_n\}_{n \geq 1} \text{ com } B_n \in \mathcal{B}(\mathbb{R}), \forall n \neq l \in \mathbb{N} \cap B_l \cap B_j = \emptyset, \forall l \neq j$$

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

Logo P_X é prob em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Definição. Seja X uma v.a.

A função $F_X: \mathbb{R} \rightarrow [0, 1]$ 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 v.a. X .

Propriedades de F_x :

1. F_x é não decrescente

$$t_1 \leq t_2 \Rightarrow (-\infty, t_1] \subseteq (-\infty, t_2] \Rightarrow \\ \Rightarrow P_x((-\infty, t_1]) \leq P_x((-\infty, t_2]) \Rightarrow F_x(t_1) \leq F_x(t_2).$$

2. F_x é contínua à direita

$$(\text{ou seja} \forall t \in \mathbb{R} \Rightarrow F(t) = \lim_{n \rightarrow \infty} F(t_n))$$

$$F(t_n) \nearrow F(t)$$

Seja $(t_n)_{n \geq 1}$ decrescente, t.g. $t_n \downarrow t \Rightarrow$

$\Rightarrow ((-\infty, t_n])_{n \geq 1}$ é monótona decrescente e, portanto,

$$\lim_{n \rightarrow \infty} P_x((-\infty, t_n]) = P_x(\lim_{n \rightarrow \infty} (-\infty, t_n]) \Rightarrow \lim_{n \rightarrow \infty} F_x(t_n) = F_x(t)$$

Definição: Seja $X: \Omega \rightarrow \mathbb{R}^k$ um vetor aleatório que induz a medida $P_x: \mathcal{B}(\mathbb{R}^k) \rightarrow \mathbb{R}_+$. A função $F_x: \mathbb{R}^k \rightarrow [0, 1]$ associa a cada $(t_1, \dots, t_k) \in \mathbb{R}^k$,

o n.º

$$P_x((-\infty, t_1] \times (-\infty, t_2] \times \dots \times (-\infty, t_k])$$

damos o nome de função distribuição conjunta de X .

Propriedades:

1. F_x é não decr. em cada comp.

2. F_x é contínua à direita em cada vmp.

3. $\lim_{t_i \rightarrow -\infty} F_x(t_1, \dots, t_k) = 0$ e $\lim_{t_i \rightarrow \infty} F_x(t_1, \dots, t_k) = 1$
para todo;

4. F_x deve ser tal que $P_x((a_1, b_1] \times (a_2, b_2] \times \dots \times (a_k, b_k]) \geq 0$.

(Ω, F, P)

$X = (X_1, \dots, X_K) : \Omega \rightarrow \mathbb{R}^K$ uma transformação.

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

(\Rightarrow)

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

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

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

Para todo $n \geq 1$

$$\bigcap_{j=1}^K \{\omega : X_{0j}(\omega) \leq n\} \cap \{\omega : X_1(\omega) \leq t_1\} \in \mathcal{F}$$

Logo

$$\bigcup_{n=1}^{\infty} \left[\bigcap_{j=1}^K \{\omega : X_{0j}(\omega) \leq n\} \cap \{\omega : X_1(\omega) \leq t_1\} \right] \in \mathcal{F}$$

$$= \{\omega : X_1(\omega) \leq t_1\} \cap \underbrace{\left(\bigcup_{n=1}^{\infty} \left[\bigcap_{j=1}^K \{\omega : X_{0j}(\omega) \leq n\} \right] \right)}_{\mathcal{N}} \in \mathcal{F}$$

$\Rightarrow \{\omega : X_1(\omega) \leq t_1\} \in \mathcal{F} \rightarrow X_1 \text{ é v.a.}$

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

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

Logo X é vetor aleatório

$X = (X_1, \dots, X_n)$ com f.d. F_X

$$\lim_{n \rightarrow \infty} F_X(t_1, n, m, n, \dots, m) = \lim_{n \rightarrow \infty} P_X((-\infty, t_1] \times (-\infty, t_2] \times \dots \times (-\infty, t_n]) =$$

$$= P_X(\lim_{n \rightarrow \infty} ((-\infty, t_1] \times (-\infty, t_2] \times \dots \times (-\infty, t_n])) =$$

$$= P_X(\bigcup_{n=1}^{\infty} ((-\infty, t_1] \times (-\infty, t_2] \times \dots \times (-\infty, t_n))) =$$

$$= P_X((-\infty, t_1] \times \mathbb{R} \times \dots \times \mathbb{R}) = P_X(\{\omega \in \Omega : X(\omega) \in (-\infty, t_1] \times \mathbb{R} \times \dots \times \mathbb{R}\}) =$$
$$= P(\{\omega \in \Omega : X_1(\omega) \leq t_1\} \cap \Omega \cap \Omega \cap \dots \cap \Omega) = P_{X_1}((- \infty, t_1])$$

Def. X é um vetor aleatório abs. contínuo se

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

$$F_X(t_1, \dots, t_k) = \int_{-\infty}^{t_1} \dots \int_{-\infty}^{t_k} f_X(t_1, \dots, t_k) dt_1 dt_2 \dots dt_k$$

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

Aula 01

"Fazer ~~inférença~~ um procedimento que faça com que você aprenda o que tem de só a partir dos dados".

Encontrar alguma relação funcional entre aquilo que queremos descobrir (θ) e X , a amostra extraída.

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 esp.

$$f(x, \theta) = f(x|\theta) f(\theta)$$

$$f(\theta|x) = \frac{f(x, \theta)}{f(x)} = \frac{f(x|\theta) f(\theta)}{f(x)}$$

"A incerteza revista para θ à luz da obs. 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, iq nem é sempre vld.

Ou seja, distribuição pré-amostragem é a mesma da pós-amostragem. "Inferência de coerência amostral"

Vimos que $\exists f$ que relaciona o res. de exp. (x) com o parâmetro (θ).

$$f(x|\theta) \text{ ou } P(x=x|\theta)$$

• Qdo fixamos $\theta \in \Theta$,

$f(\cdot|\theta)$ é uma dist. de prob. em \mathcal{X} (Distr. Amostral)

• Qdo fixarmos $x \in \mathcal{X}$,

$f(x|\cdot)$ é um f. n-negativa de θ .

$\hookrightarrow V_x(\cdot) \Rightarrow$ Função de Verossimilhança para θ gerada por x .

Aula 2

$$V_k : \Theta \rightarrow \mathbb{R}$$

$$\theta \in \Theta \mapsto V_k(\theta) = P(X=x|\theta)$$

Obs.

$$f(\theta|x) = \frac{f(x|\theta) f(\theta)}{f(x)} \text{ com } f(x) = \int_{\Theta} f(x|\theta) d\theta$$

Ber(θ), Beta(a, b), Beta($a + \sum x_i, b + n - \sum x_i$)

Exp(θ), Gama(a, b), Gama($a+n, b+\sum x_i$)

$$N(\theta, \sigma_0^2), N(a_0, b_0), N\left(\frac{b_0^{-1}n\bar{x} + a_0^{-1}a_0}{n b_0^{-1} + a_0^{-1}}, \left(\frac{a_0^{-1}}{n b_0^{-1} + a_0^{-1}} + \frac{1}{b_0}\right)^{-1}\right)$$

(contínuo)

Aula 3

Definição. Dizer transformação $T: \mathcal{X} \rightarrow \mathcal{Y}$ mensurável é chamada Estatística.

Definição. $x = (x_1, \dots, x_n)$, $T: \mathcal{X} \rightarrow \mathcal{Y}$ é uma estatística suficiente para θ se para toda priori $p(\theta)$,

$$\theta/x = \theta/T(x) \stackrel{d}{=} T(x), \quad \forall x \in \mathcal{X}$$

ou seja, a distribuição condicional de θ dado x é a mesma que a distribuição condicional de θ dado $T(x)$.

$$f(x_m | \theta) = \frac{n(x_m)^{n-1}}{\theta^n}$$

Def. $T: \mathcal{X} \rightarrow \mathcal{Y}$ é suficiente para θ se a dist. de $x | T(x) = t(x)$, θ não depende de $\theta | t(x)$ para todo $t \in T(\mathcal{X})$ e $\forall \theta \in \Theta$.

Suf. Clássica \Rightarrow Suf. Bayesiana.

Critério da Fatoração (Fisher-Neyman)

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

A estat. $T: \mathcal{X} \rightarrow \mathcal{Y}$ é suf. para θ se e só se $\exists \mu: \mathcal{X} \rightarrow \mathbb{R}^+$ e $v: \mathcal{Y} \times \Theta \rightarrow \mathbb{R}^+$ tais que

$$f(x | \theta) = \mu(x)v(T(x), \theta), \quad \forall x, \forall \theta.$$

Aula 4.

Resultado: Seja $T: \mathcal{X} \rightarrow \mathcal{Y}$ suf. para θ .

Seja $S: \mathcal{X} \rightarrow \mathcal{P}$ outra estatística t.q. $S = g \circ T$, onde $g: \mathcal{Y} \rightarrow \mathcal{P}$ é bijetora (um a um). Então S é suficiente para θ .

$T: \mathcal{X} \rightarrow \mathcal{Y}$ é suf.

Seja $T': \mathcal{X} \rightarrow \mathcal{Y}'$ um outra estat. suficiente. Dizemos que $T: \mathcal{X} \rightarrow \mathcal{Y}$ é suf. mínima para θ se T é função de que r outra estatística suficiente para θ .

Em outras palavras, T é suf. mínima se para todo T' existe $g: \mathcal{Y}' \rightarrow \mathcal{Y}$ tal que $T = g \circ T'$.

Ou ainda, T é suf. mín. se para toda T' suf. 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 $T: \mathcal{X} \rightarrow \mathcal{Y}$ uma estatística

Se $\forall x_1, x_2 \in \mathcal{X}$,

$f(x_1 | \theta)$ não depende de $\theta \Leftrightarrow T(x_1) = T(x_2)$, então T é suf. mínima.

Estimação

Uma Estatística $T: \mathcal{X} \rightarrow \Theta$ é chamada Estimador para θ .

Usualmente $S(x)$ -estimador e $\delta(x)$ estimativa, $\Theta \subseteq \mathbb{R}^k$.

Como escolher o estimado? Deveremos est. com critérios de otimalidade.

$|S(x) - \theta|$ ou $(\delta(x) - \theta)^2$ ou $\left|\frac{\delta(x) - \theta}{\theta}\right|$ → depende de θ . Logo, não podemos concluir nada.

Ideia: Escolher um estimador que minimize os critérios acima, em média.

$$\rho(\delta) = E[L(S(x), \theta)] = \sum_{x \in \mathcal{X}} \sum_{\theta \in \Theta} L(\delta(x), \theta) P(X=x, \theta=\cdot)$$

L($\delta(x), \theta$)

$L(\theta, \delta) \rightarrow \mathbb{R}_+$ penalidade (perda) que o decisar incurre ao estimar θ por $\delta(x)$ no obs. $x \in \mathcal{X}$.
 $L(\theta, \delta) \sim L(\delta(x), \theta)$

Ideia: Escolher $\delta^* \in \Delta$ t.q.

$$\rho(\delta^*) = \min_{\delta \in \Delta} \rho(\delta).$$

Poderemos reescrever

$$\rho(\delta) = \sum_{x \in \mathcal{X}} \sum_{j=0}^k L(\delta(x), \theta_j) P(\theta=j | X=x) F(x=x) = \sum_{x \in \mathcal{X}} P(X=x) \left\{ \sum_{j=0}^k L(\delta(x), \theta_j) P(\theta=j | X=x) \right\}$$

Ideia: Para cada $x \in \mathcal{X}$, tomamos $\delta^*(x) \in \Theta$ que minimize $\sum_{j=0}^k L(d_j, j) P(\theta=j | X=x)$ em d_j .

Ideia: Para cada $x \in \mathcal{X}$, tomamos $\delta^*(x) \in \Theta$ que minimize $E[L(\delta(x), \theta) | X=x]$.
 Ao estimador que associa a cada $x \in \mathcal{X}$ $\delta^*(x)$ t.o. especificado o lado, damos o nome de Estimador de Bayes para θ considerando a perda L .

Aula 5.

Cons. $\delta^*: \mathcal{X} \rightarrow \Theta$ que associa a cada $x \in \mathcal{X}$, o valor $d^* \in \Theta$ tal que:

$$\sum_{\theta \in \Theta} L(d^*, \theta) P(\theta=j | X=x) = \min_{d \in \Theta} \sum_{\theta \in \Theta} L(d, \theta) P(\theta=j | X=x)$$

$E[L(d, \theta) | X=x]$

Isto é, $\delta^*(x) = d^*(x)$ depende da posteriori.

δ^* - estimador de Bayes e $\rho(\delta^*)$ - Risco de Bayes.

"Para cada $x \in \mathcal{X}$ constrói-se a distribuição a posteriori e encontramos o valor de d que minimiza os riscos a posteriori."

P. f. depende:

$$L(d, \theta) = (d - \theta)^2 \Rightarrow d^* = E(\theta | X=x) \text{ a.t.q. } d^* = \min_{d \in \Theta} |d - E_\theta(\theta | X=x)|$$

1. Perda Quadrática: $L(d, \theta) = (d - \theta)^2 \Rightarrow \delta^*(x) = \text{mediana}(\theta | X=x)$

2. Perda Absoluta: $L(d, \theta) = |d - \theta| \Rightarrow \delta^*(x) = \text{moda}(\theta | X=x)$

3. Perda 0-1

Est. & Maxima Verossimilhança

$$\delta_{\text{MLR}}(x) = \arg \max_{\theta \in \Theta} V_x(\theta)$$

Aula 6

Comments

1. T é uma est. suficiente δ_{MV} é unica $\Rightarrow \delta_{MV}$ é função de T .
 2. Qdo δ_{MV} é uma est. suficiente e única $\Rightarrow \delta_{MV}$ é suficiente minim.
 3. Invariância dos EMV:

3. Invariância dos ENV:

Seja $g: \Theta \rightarrow T$ ($\Rightarrow g(\Theta)$) é que

$\text{g}_{\text{Mv}} \circ \text{ENV} \text{ para } \theta \Rightarrow g(\text{Mv}) \in \text{EMV} \text{ por } g(\theta).$

Método dos Momentos

$$E(Y_1 | \Theta) = x_n, \dots, y_n$$

$$E(X_1^2 | \mathcal{G}) = x_1^2 + \dots + x_n^2 / n$$

$$f(x) = \min(x, 1-x)$$

Aula 7. Estimadores de mínimos cuadrados

Ideia: Escolher para cada \hat{x}_i , como estimativa o valor que minimiza a "diferença" entre o val. obs. e esperado.

para cada $x \in \mathbb{X}$, tomamos $\delta_{\text{na}}(x)$ t.q. \mathbb{B}

$$\sum_{i=1}^n (x_i - E(x_i | \delta_{\text{ML}}(\boldsymbol{x})))^2 = \min_{\theta \in \Theta} \sum_{i=1}^n (x_i - E(x_i | \theta))^2$$

$\underbrace{s(\boldsymbol{x}, \delta_{\text{ML}}(\boldsymbol{x}))}_{\text{ML}}$ $\underbrace{s(\boldsymbol{x}, \theta)}_{\text{General}}$

Comp. F. Avaliação) De Estimadores

Seja $\delta \in \Delta$. A quantidade $\sum_{x \in \mathcal{X}} L(\delta(x), \theta) p(x=x/\Theta) \left(\int_{\mathcal{X}} L(\delta(x), \theta) f(x/\Theta) dx \right)$ chamamos de risco frequentista do estimador δ para estimar o parâmetro que esse vale θ .

Notação: $R(\delta, \theta) = E[L(\delta(x), \theta) | \theta] = E_{\theta}[L(\delta(x), \theta)]$

foro $\delta \in \Delta$, $R_\delta : \Theta \rightarrow \mathbb{R}_+$

$$\text{SEQR} \hookrightarrow \mathbb{R}_\delta(G) = \mathbb{R}(\delta, \theta)$$

e "chamada FUNÇÃO PERÍODICA do estimador \hat{g} ".

$$\text{QdO } L(d, \theta) = (d - \theta)^L, R(s, \theta) = Es^L$$

Dizemos que δ_1 é melhor que δ_2 se

$$\delta_1(\theta) \in R_{\delta_2}(\theta), \forall \theta \in \Theta.$$

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 .

Só o δ_1 é melhor que δ_2 mas δ_2 não é melhor que δ_1 , dizemos que δ_1 "domina" δ_2 . Nesse caso, dizemos que δ_2 é inadmissível.

$$EQM_\theta(\delta) = Var_\theta(\delta) + B^2(\delta)$$

$$L = E_\theta(\delta(x)) - \theta \cdot \text{vício de estimador de } \theta.$$

Restringir a busca à classe

$$\Delta' = \{\delta \in \Delta : E_\theta[\delta(x)] = 0, \forall \theta \in \Theta\}$$

Definição: Um estimador $\delta: \mathcal{X} \rightarrow \mathbb{R}$ p/ $\theta \in \Theta$ é dito NÃO-VIESADO para θ se

$$E_\theta[\delta(x)] = \theta, \forall \theta \in \Theta$$

Aula 08

Se δ é um estimador não-viesado p/ $\theta = g(\theta)$.

Seja $\Delta'' = \{\delta_0 + U : U \in \mathcal{U}\}$ onde $\mathcal{U} = \{U : \mathcal{X} \rightarrow \mathbb{R} : E_\theta(U(x)) = 0, \forall \theta \in \Theta\}$ $\Delta'' = \Delta'$.

$$\Delta'' = \Delta'.$$

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 $Var_\theta(\delta^*) \leq Var_\theta(\delta), \forall \theta \in \Theta, \forall \delta \in \Delta$.

Def. Dizemos que um estimador $\delta^* \in \Delta'$ é não-viesado de variância uniformemente mínima (ENVVUM) se $\forall \delta \in \Delta'$

$$Var_\theta(\delta^*) \leq Var_\theta(\delta), \forall \theta \in \Theta$$

Teorema de Rao-Blackwell.

Seja $\delta_0: \mathcal{X} \rightarrow \mathbb{R}$ um estimador não-viesado para $\theta = g(\theta)$. Seja T uma estatística suficiente para θ . Seja

$$\delta^* = E(\delta_0 | T)$$

então:

1. δ^* depende de T

2. $\delta^* \in \Delta'$

3. $Var_\theta(\delta^*) \leq Var_\theta(\delta_0), \forall \theta \in \Theta$

Dizemos que a estatística $T: \mathcal{X} \rightarrow \mathbb{R}^K$ é completa se $\forall h: \mathbb{R}^K \rightarrow \mathbb{R}$,

$$E_\theta(h(T(x))) = 0, \forall \theta \in \Theta \Rightarrow h \equiv 0$$

Aula 9

$T: \mathcal{X} \rightarrow \mathbb{R}^k$ é completa se $\forall h: \mathbb{R}^k \rightarrow \mathbb{R}$ $E_\theta(h(T(x))) = 0, \forall \theta \in \Theta \Rightarrow h \equiv 0$. P. q.c. H o c.

Teorema de Lehmann-Scheffé

Seja $T: \mathcal{X} \rightarrow \mathbb{R}^k$ uma estatística suficiente e completa. Então todos os estimadores não-viesados de $E(g(\theta))$ que são função de T (dependem de x através de T) são P_θ -q.c. H o c.

Assim, se $\delta^* \in A^*$ é função de uma estatística suficiente e completa, δ^* é o ENVVUM para θ .

Resultado.

δ^* é o ENVVUM para θ (ou $E_\theta(\delta^*(x))$) se e só se

$$\text{Cov}_\theta(\delta^*(x), U(x)) = 0, \forall \theta \in \Theta, \forall u \in \mathcal{U},$$

onde $\mathcal{U} = \{U: \mathcal{X} \rightarrow \mathbb{R} : E_\theta(U(x)) = 0, \forall \theta \in \Theta\}$.

Desigualdade da Informação (Cramer-Rao)

Seja δ um estimador não-viesado para $g(\theta)$. Sob certas condições (condições de regularidade), vale que

$$\text{VAR}_\theta(\delta(x)) \geq \frac{(g'(\theta))^2}{I_x(\theta)},$$

onde $I_x(\theta)$ é chamada Informação de Fisher para θ contida na amostra,

$$I_x(\theta) = E_\theta[(\lambda'(x/\theta))^2], \text{ onde}$$

$$\lambda(x/\theta) = \log f(x/\theta).$$

$\frac{d}{d\theta} \lambda(x/\theta) \rightarrow$ FUNÇÃO ESCORE

"Toda a inf. deve ser baseada apenas no pto amostral se é efetivamente observado".

$$I_x(\theta) = -E_\theta[\lambda''(x/\theta)]$$

Prop. Assintóticas.

Seq. de v.a. $(X_n)_{n \geq 1}$

$$X = (X_n)_{n \geq 1}, \quad S_n(x) = \delta_n(x_1, \dots, x_n)$$

$(\delta_n)_{n \geq 1}$ seq. de estimadores

Def. $(\delta_n)_{n \geq 1}$ é uma seqüência consistente para $\theta(g(\theta))$ se $S_n(x) \xrightarrow{P_\theta} \theta(g(\theta))$, H o c.

Def. $(\delta_n)_{n \geq 1}$ é Assintoticamente Normal $(0, \sigma^2(\theta))$ se

$$\sqrt{n}(\delta_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta)), \forall \theta \in \Theta \quad (\text{plano grande}) \quad \ln \approx N(0, \sigma^2(\theta)/n)$$

Resultado: $(\delta_n)_{n \geq 1}$ é assintoticamente Normal $(0, \sigma^2(\theta)) \Rightarrow$

$(\delta_n)_{n \geq 1}$ é consistente para θ .

Def. $(\delta_n^{(i)})_{n \geq 1}$ seq. assintoticamente Normal $(0, \sigma_i^2(\theta))$, $i=1, 2$. Dizemos que $(\delta_n^{(1)})_{n \geq 1}$ é assintoticamente melhor que $(\delta_n^{(2)})_{n \geq 1}$ se $\sigma_1^2(\theta) \leq \sigma_2^2(\theta) \quad \forall \theta \in \Theta$.

Def. $(\delta_n^*)_{n \geq 1}$ 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.

Def. $(\delta_n)_{n \geq 1}$ e $(\delta'_n)_{n \geq 1}$ são assintoticamente equivalentes se

$$\sqrt{n}(\delta_n - \delta'_n) \xrightarrow{D_\theta} 0, \forall \theta \in \Theta$$

Resultado: $(\delta_{MV}^{(n)})_{n \geq 1}$ sequência de EMV para θ . Sob condições de regularidade,

$$\sqrt{n}(\delta_{MV}^{(n)} - \theta) \xrightarrow{d} N\left(0, \frac{1}{nJ_{ML}(\theta)}\right)$$

no caso $(X_n)_{n \geq 1}$, dado θ , c.i.i.l.

$$\delta_{MV}^{(n)}(x) \approx N\left(\theta, \frac{1}{nJ_{ML}(\theta)}\right)$$

Aula 31

Resultado (Método Delta): $(\delta_n)_{n \geq 1}$ tal que

~~$$\sqrt{n}(\delta_n - \theta) \xrightarrow{D} N(0, \sigma^2(\theta))$$~~

Seja $g: \mathbb{R} \rightarrow \mathbb{R}$ diferenciável. Então,

$$\sqrt{n}(g(\delta_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2(\theta)(g'(\theta))^2)$$

$[g(\delta_n)]_{n \geq 1}$ é ass. Normal].

Intervalo de Confiança

Sejam $A, B: \mathcal{X} \rightarrow \mathbb{R}$ t.q. $A(x) \leq B(x), \forall x \in \mathcal{X}$

$$\{[a, b] : a \leq b\}$$

Def. A transf. $T: \mathbb{R} \rightarrow \mathbb{I}$ que a cada $x \in \mathbb{R}$ associa um intervalo em \mathbb{I} (i.e., $T(x) = [a(x), b(x)]$) damos o nome de Intervalo Aleatório.

Def. O intervalo aleatório $[A(x), B(x)]$ é um intervalo de confiança (I.C.) para $\theta(g(\Theta))$ com coeficiente de confiança $1-\alpha$, $0 \leq \alpha \leq 1$, se

$$P(A(x) \leq \theta(g(\Theta)) \leq B(x) | \theta) \geq 1-\alpha, \forall \theta \in \Theta$$

$$\hookrightarrow P(\{\omega \in \Omega : A(x) \leq \theta(g(\omega)) \leq B(x)\}) \geq 1-\alpha$$

"Em geral, procura-se $T(k, \theta)/\theta$ (uma dist. não depende de θ).

$$[(-\infty, B(x))] \text{ (l. sup.)}$$

Def. $[A(x), \infty)$ é um limite inferior de confiança com coeficiente de confiança $1-\alpha$, $0 \leq \alpha \leq 1$ se

$$P(A(x) \leq \theta \leq \infty | \theta) \geq 1-\alpha, \forall \theta \in \Theta$$

$$P(-\infty \leq \theta \leq B(x) | \theta) \geq 1-\alpha, \forall \theta \in \Theta$$

Aula 12

Definição: A transf. $R: \mathbb{R} \rightarrow \mathbb{A}$ que associa a cada $x \in \mathbb{R}$ uma região (ou sub-conjunto) $R(x) \in \mathbb{A}$, damos o nome de Região (cpto) aleatório.

Definição: A região (cpto) aleatória (ϵ) R é uma região (cpto) de confiança pl $g(\epsilon)$ com coeficiente de confiança $1-\alpha$ ($0 \leq \alpha \leq 1$), se

$$P(G \in R(\epsilon) | \theta) \geq 1-\alpha, \forall \theta \in \Theta$$

Intervalos Aproximados

No caso geral, (sob cond. de regularidade)

$$\sqrt{n}(\delta_{mu} - \theta) \xrightarrow{d} N(0, \frac{1}{J_{\theta}(g)})$$

Se $g \in d.f.$
 $\sqrt{n}(g(\delta_{mu}) - g(\theta)) \xrightarrow{d} N(0, (g'(\theta))^2 / J_{\theta}(g))$, pelo método Det G

Sist:

para "n grande"

$$P\left(\Phi^{-1}(\gamma_2) \leq \frac{\sqrt{n}(g(\delta_{mu}) - g(\theta))}{\sqrt{(g'(\delta_{mu}))^2 / J_{\theta}(\delta_{mu})}} \leq \Phi^{-1}(1-\gamma_2) | \theta\right) \geq 1-\alpha, \forall \theta \in \Theta$$

Int. de Probabilidade (ou de Credibilidade)

Def. $[A(x), B(x)]$ é um intervalo de probabilidade (ou de credibilidade) para o θ ($\theta(\omega)$) com probabilidade $1-\alpha$, $0 \leq \alpha \leq 1$ se

$$P(A(x) \leq \theta \leq B(x) | X=x) \geq 1-\alpha, \forall x \in \mathbb{R}$$

Aula 13

Teste de Hipótese

Def. Uma função de teste (teste) é uma função $\psi: \mathcal{X} \rightarrow \{0,1\}$ para testar $H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_1$ que associa a cada $x \in \mathcal{X}$ $\psi(x)=1$ no caso de decisão por H_1 . H_0 a partir de x ou $\psi(x)=0$ no caso de decisão por H_1 .

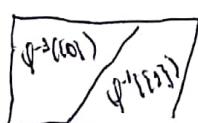
$$H_0: \theta \in \Theta_0 \rightarrow \text{Hip. Nula}$$

$$H_1: \theta \in \Theta_1 \rightarrow \text{Hip. Alternativa.}$$

As regiões $\psi^{-1}(\{1\}) = \{x \in \mathcal{X} : \psi(x)=1\}$ dão o nome de Região Crítica ou Região de Rejeição (de H_0).

A região $\psi^{-1}(\{0\}) = \{x \in \mathcal{X} : \psi(x)=0\}$ d... "a Região de Aceitação de H_0 ".

$\psi: \{\psi: \mathcal{X} \rightarrow \{0,1\}\} \ni \psi \text{ é mensurável}$



Como escolher ψ de maneira otima?

No prob. decisao com 2 alt., pode-se incorrer 2 possiveis erros.

1. Decidir por H_1 ($\psi(x)=1$) qdo H_0 é vdd. - Erro tipo I

2. Decidir "a" H_0 ($\psi(x)=0$) qdo H_1 - Erro tipo II

Seja $\psi: \mathcal{X} \rightarrow \{0,1\}$ um teste para $H_0: \theta \in \Theta_0$, ao número

$\pi_\psi(\theta) = P(\psi(x)=1 | \theta)$ damos o nome de função Poder de Teste,

Dizemos que uma hipótese $H_1: \theta \in \Theta_1$ é SIMPLES se $|\Theta_1| = 1$.

" " " $H_1: \theta \in \Theta_1$ é COMPOSTA se $|\Theta_1| > 1$.

Seja $\psi: \mathcal{X} \rightarrow \{0,1\}$ um teste para $H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_1$. Ao número $\sup_{\theta \in \Theta_1} \pi_\psi(\theta)$ damos o nome de Tamanho do teste ψ . ($\alpha_\psi^1 = \sup_{\theta \in \Theta_1} \pi_\psi(\theta)$)

Def. Um teste Ψ é de nível α , $0 \leq \alpha \leq 1$ se

$$\alpha_{\Psi} \leq \alpha.$$

Aula 14

Teste Para Hipótese Simples

$$\Theta = \{\theta_0, \theta_1\} \quad H_0 = \{\theta_0\} \quad H_1 = \{\theta_1\}$$

$$H_0: \theta = \theta_0 \quad (\theta \in \Theta_0)$$

$$H_1: \theta = \theta_1 \quad (\theta \in \Theta_1)$$

$\Psi = \{\Psi: \mathcal{X} \rightarrow \{0,1\}\}$ e $\Psi \in \Psi$ é teste para H_0 vs H_1 .

Para $\theta \in \Psi$

$$\pi_{\Psi}(\theta_0) = P(\Psi(x) = 1 / \theta_0) = \text{prob. erro tipo I}$$

$$1 - \pi_{\Psi}(\theta_1) = P(\Psi(x) = 0 / \theta_1) = \text{prob. erro tipo II}$$

Resultado:

Seja $\Psi^*: \mathcal{X} \rightarrow \{0,1\}$ dada por

$$\Psi^*(x) = \begin{cases} 1, & af(x/\theta_0) \leq bf(x/\theta_1) \quad \left(\frac{f(x/\theta_1)}{f(x/\theta_0)} \geq \frac{a}{b} \right) \\ 0, & \text{c.c.} \end{cases} \quad a, b > 0 \text{ cte}$$

Então, $\forall \Psi \in \Psi$,

$$a\pi_{\Psi}(\theta_0) + b(1 - \pi_{\Psi^*}(\theta_1)) \leq a\pi_{\Psi}(\theta_0) + b(1 - \pi_{\Psi}(\theta_1)).$$

Definição: Um teste de nível α , $0 \leq \alpha \leq 1$, Ψ^* , é um Teste Mais Poderoso (MP) de nível α para $H_0: \theta \in \Theta_0$ versus $H_1: \theta = \theta_1$ se

$$\pi_{\Psi^*}(\theta_1) \geq \pi_{\Psi}(\theta_1), \forall \Psi \text{ de nível } \alpha.$$

Lema de Neyman-Pearson

$$H_0: \theta = \theta_0 \text{ vs } H_1: \theta = \theta_1$$

Seja $\Psi^*: \mathcal{X} \rightarrow \{0,1\}$ dada por

$$\Psi^*(x) = \begin{cases} 1, & f(x/\theta_0) \leq Kf(x/\theta_1), K > 0 \\ 0, & \text{c.c.} \end{cases}$$

Então

$$\pi_{\Psi^*}(\theta_1) \geq \pi_{\Psi}(\theta_1), \forall \Psi \in \Psi^* = \{\Psi \in \Psi : \alpha_{\Psi} \leq \pi_{\Psi^*}(\theta_0)\}$$

Isso é, nessas condições, Ψ^* é MP de nível $\pi_{\Psi^*}(\theta_0)$.

Aula 15

$$P(\theta = \theta_1) = p_1$$

$$0 < p_1 \leq 1$$

$$P(\theta = \theta_0) = 1 - p_1$$

	Decisão		$\theta > \theta_0$	$\theta = \theta_0$
	Nao R. H_0	R. H_0	α_1 $(L(\theta_1, \theta_0))$	α_2 $(L(\theta_0, \theta_1))$
$P(\theta = \theta_1) = p_1$	α_1 $(L(\theta_1, \theta_0))$	α_2 $(L(\theta_0, \theta_1))$	$\alpha_1, \alpha_2 > 0$	
$P(\theta = \theta_0) = 1 - p_1$				

$\Psi^* = \{\psi: \mathcal{X} \rightarrow \{0,1\}\}$. Devemos escolher $\psi \in \Psi^*$ que minimiza

$$P(\psi) = E[L(\psi(X))]$$

$$\psi^*(x) = \begin{cases} 1 & \text{para } P(X=x|\theta_0) \leq p_{\text{alr}} \\ 0 & \text{caso contrário} \end{cases}$$

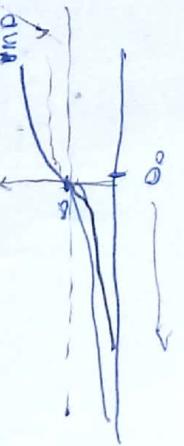
ψ^* é chamada mesmo caso de teste de Bayes

Hipóteses Compostas

Definição: O teste ψ^* de nível α , ou seja, é um teste uniformemente mais poderoso (UMP) de nível α para $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_1$, se $\psi \in \Psi$ tal que

$$\alpha^* \leq \alpha,$$

$$\pi_{\psi^*}(\theta) \geq \pi_{\psi}(\theta), \forall \theta \in \Theta_1$$



Em geral, é teste UMP:

Namorar olhar um caso importante onde é teste UMP!

$$\exists c \in \mathbb{R}, \mathcal{D} = \{f(.|\theta): \theta \in \Theta\}$$

$$\text{Seja } T: \mathcal{X} \rightarrow \mathbb{R}.$$

Def. A família $\mathcal{D} = \{f(.|\theta): \theta \in \Theta\}$ possui razão de verossimilhança

monótona crescente (decrece) em T se $\forall \theta_1, \theta_2 \in \Theta$, com $\theta_1 < \theta_2$,

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} \text{ é f. crescente (dec.) de } T(x)$$

nos possíveis valores (int. devalores) de T sob θ_1 ou θ_2 .

Aula 26

Teorema (Karlin-Rubin).

$\Theta \subseteq \mathbb{R}$, $D, T: \mathcal{X} \rightarrow \mathbb{R}$, Ψ supõe que \mathcal{D} possui BVM crescente (dec) em T . O teste $\psi^*: \mathcal{X} \rightarrow \{0,1\}$ dado por

$$\psi^*(x) = \begin{cases} 1 & T(x) \geq c \\ 0 & \text{caso contrário} \end{cases}$$

para as hip. $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_1$, é UMP do nível $\pi_{\psi^*}(\theta_0)$

idem.

Resumo: $\psi^* = \psi^*(x)$, dado o c.i. d. tais que $f(x|\theta) = c(\theta) e^{q(T(x))}$, $\psi^*(x)$ é estrit. crescente

$\theta_0: \theta \in \Theta_0$, $\theta_1: \theta \in \Theta_1$, $\psi^*(x) = \begin{cases} 1 & x \in \mathcal{Z}_{T(x)} \\ 0 & \text{caso contrário} \end{cases}$ é UMP de nível $\pi_{\psi^*}(\theta_0) = \pi_{\psi^*}(0)$

AVALIAÇÃO 1

MAE 5702 - 09/05/2013

Escreva seu NOME e NÚMERO USP no caderno de respostas.

Desligue seu telefone celular ou qualquer aparelho que possa produzir barulho.

Você pode consultar um resumo de no máximo 2 páginas do conteúdo da disciplina.

SEMPRE EXPLIQUE SEU RACIOCÍNIO!!!!

(1) SEJA (X, Y) UM VETOR ALEATÓRIO ABSOLUTAMENTE CONTÍNUO COM FUNÇÃO DENSIDADE DE PROBABILIDADE DADA POR

$$f_{(x,y)}(x,y) = 2 \cdot I_A, \text{ ONDE } A = \{(a,b) \in \mathbb{R}^2 : 0 \leq a \leq 1, a \leq b \leq 1\}.$$

(a) OBTENHA $P(X \geq 1/2, Y \leq 1/2)$. X E Y SÃO INDEPENDENTES? JUSTIFIQUE.

(b) SEJAM $U = X$ E $V = \frac{X+Y}{2}$. OBTENHA A FUNÇÃO DENSIDADE DO VETOR (U, V) .

(c) OBTENHA A FUNÇÃO DENSIDADE DE PROBABILIDADE CONDICIONAL DE U DADO $V=v$, $v \in (0,1)$.

(PARA CADA $v \in (0,1)$). QUAL É A ESPERANÇA CONDICIONAL DE U DADO $V=v$, $v \in (0,1)$.

(2) SEJA (X, Y) UM VETOR ALEATÓRIO DISCRETO TAL QUE $P(X=j) = p_j$, $j=1,2,3$, COM

$p_j \in (0,1)$, $j=1,2,3$, E $P_1 + P_2 + P_3 = 1$, E TAL QUE A DISTRIBUIÇÃO CONDICIONAL DE Y

DADO $X=j$ É DESCRITA PELO MODELO GEOMÉTRICO DE PARÂMETRO $\min\left\{\frac{j}{4}, 1 - \frac{j}{4}\right\}$, $j=1,2,3$.

(a) PARA CADA $n \in \mathbb{N}^*$, OBTENHA A DISTRIBUIÇÃO CONDICIONAL DE X DADO $Y=n$.

(b) PARA CADA $n \in \mathbb{N}^*$, SEJA W_n UMA VARIÁVEL ALEATÓRIA CUJA DISTRIBUIÇÃO COINCIDE COM A

DISTRIBUIÇÃO CONDICIONAL DE X DADO $Y=n$ OBTIDA EM (a). A SÉQUENCIA $(W_n)_{n \geq 1}$ CONVERGE EM DISTRIBUIÇÃO? QUAL É O LIMITE? (INTERPRETE O RESULTADO).

(3) PARA CADA $n \in \mathbb{N}^*$, SEJA $I_n = \{1, \dots, n\}$. SEJA TF O CONJUNTO DAS FUNÇÕES COM DOMÍNIO I_n E CONTRA-DOMÍNIO I_m , $m, n \in \mathbb{N}^*$. CONSIDERE O EXPERIMENTO QUE CONSISTE EM "ESCOLHER AO ACASO" UMA FUNÇÃO EM TF E REGISTRAR X : O NÚMERO DE ELEMENTOS QUE COMPõEM O CONJUNTO IMAGEM DA FUNÇÃO ESCOLHIDA (ISTO É, PARA $f \in \text{TF}$, SEJA $X(f) = |f(I_n)| = |\{f(i) : i \in I_n\}|$). POR EXEMPLO, PARA $n=5 < m=4$

$$X\left(\underbrace{\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array}}_{m^n}\right) = 3 \quad \text{E} \quad X\left(\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array}\right) = 2.$$

OBTENHA A ESPERANÇA E A VARIÂNCIA DE X .

(4) CONSIDERE O ESPAÇO (Ω, \mathcal{F}, P) . SEJAM $(X_n)_{n \geq 1}$ E $(Y_n)_{n \geq 1}$ SEQUÊNCIAS DE VARIÁVEIS ALEATÓRIAS E X UMA VARIÁVEL ALEATÓRIA DEFINIDAS EM Ω , TALIS QUE $X_n \xrightarrow{q.c.} X$ E $Y_n \xrightarrow{q.c.} X$. PARA CADA $n \in \mathbb{N}^*$, SEJA $A_n = \{w \in \Omega : X_n(w) = Y_n(w)\} = \{X_n = Y_n\}$.

(a) SE $X_n(\Omega) \subseteq \mathbb{Z}$, $\forall n \geq 1$, E $Y_n(\Omega) \subseteq \mathbb{Z}$, $\forall n \geq 1$, ENTÃO $P(\liminf_{n \rightarrow \infty} A_n) = 1$.
VERDADEIRO OU FALSO? JUSTIFIQUE.

$$\begin{array}{l} \xrightarrow{q.c.} X_n \rightarrow X \\ \xrightarrow{q.c.} Y_n \rightarrow X \end{array}$$

(b) SE $X_n(\Omega) \subseteq \mathbb{R}_+$, $\forall n \geq 1$, E $Y_n(\Omega) \subseteq \mathbb{R}_+$, $\forall n \geq 1$, ENTÃO $P(\liminf_{n \rightarrow \infty} A_n) = 1$.
VERDADEIRO OU FALSO? JUSTIFIQUE.

Avaliação 1 (Esboço das soluções)

QUESTÃO 1

$$(a) P(X \geq 1/2, Y \leq 1/2) = \int_{-\infty}^{\infty} \int_{-\infty}^{1/2} f_{(X,Y)}(x,y) dx dy = \int_{1/2}^{\infty} \int_{-\infty}^{1/2} 0 \cdot dy dx = 0.$$

$$P(X \geq 1/2) = \int_{-\infty}^{\infty} \int_{1/2}^{\infty} f_{(X,Y)}(x,y) dx dy = \int_{1/2}^{\infty} \int_x^{\infty} 2 dy dx > 0.$$

$$P(Y \leq 1/2) = \int_{-\infty}^{\infty} \int_{-\infty}^{1/2} f_{(X,Y)}(x,y) dy dx = \int_0^{1/2} \int_x^{1/2} 2 dy dx > 0.$$

Como $P(X \geq 1/2) \cdot P(Y \leq 1/2) > 0$ e $P(X \geq 1/2, Y \leq 1/2) = 0$, segue que os eventos

$\{X \geq 1/2\}$ e $\{Y \leq 1/2\}$ NÃO são independentes, consequentemente, as variáveis X e Y não são independentes.

(b) Como a transformação é bijetora, com derivadas parciais (da inversa) bem definidas e contínuas, segue que

$$f_{U,V}(\mu, v) = f_{(X,Y)}(x(\mu, v), y(\mu, v)) \cdot |J|.$$

Como $U = X + V$ e $V = \frac{(X+Y)}{2} \Rightarrow X = \mu$ e $Y = 2v - \mu$. Assim,

$$J = \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = 2.$$

Logo,

$$f_{U,V}(\mu, v) = 2 \cdot \underset{(0,1)}{\mathbb{I}(\mu)} \underset{(\mu,1)}{\mathbb{I}(2v-\mu)} \cdot 2$$

PASTA N°:	13
QTDE.FLS.:	06

(c) Para obtermos a densidade condicional de U dado $V=v$, devemos obter, inicialmente, a densidade marginal de V :

$$f_V(v) = \int_{-\infty}^{\infty} f_{U,V}(u,v) du = \int_{-\infty}^{\infty} 4 \mathbb{I}_{(0,1)}(u) \mathbb{I}_{(v,1)}(2v-u) du = \int_{-\infty}^{\infty} 4 \mathbb{I}_{\left(\max\{2v-1, 0\}, \min\{1, v\}\right)}(u) du =$$

$$\Rightarrow f_V(v) = 4 \left(\min\{1, v\} - \max\{2v-1, 0\} \right).$$

Assim,

$$f_V(v) = \begin{cases} 4v, & v \in (0, \frac{1}{2}) \\ 4(1-v), & v \in (\frac{1}{2}, 1). \end{cases}$$

$$\text{Logo, } f_{U|V=v}(u) = \begin{cases} \frac{f_{U,V}(u,v)}{f_V(v)} = \frac{4 \mathbb{I}_{(0,v)}(u)}{4v} = \frac{1}{v} \mathbb{I}_{(0,v)}(u), & v < \frac{1}{2} \\ \frac{4 \mathbb{I}_{(2v-1,v)}(u)}{4(1-v)} = \frac{1}{(1-v)} \mathbb{I}_{(2v-1,v)}(u), & v > \frac{1}{2}. \end{cases}$$

Assim,

$$U|V=v \sim \text{Uniforme}(0, v), \quad v < \frac{1}{2}$$

$$U|V=v \sim \text{Uniforme}(2v-1, v), \quad v > \frac{1}{2}.$$

Assim,

$$E(U|V=v) = \begin{cases} \frac{v}{2}, & 0 < v < \frac{1}{2} \\ \frac{3v-1}{2}, & v > \frac{1}{2} \end{cases}$$

ESTÃO 2

Do enunciado, $P(X=j) = P_j$, $j=1, 2, 3$. e $Y|X=j \sim Geo\left(\min\left\{\frac{j}{4}, 1-\frac{j}{4}\right\}\right)$.

Pelo Teorema de Bayes, dado $Y=n$, $n \in \mathbb{N}^*$, temos

$$P(X=j|Y=n) = \frac{P(Y=n|X=j)P(X=j)}{\sum_{i=1}^3 P(Y=n|X=i)P(X=i)} . \quad \text{Para } j=1,$$

$$P(X=1|Y=n) = \frac{\left(\frac{3}{4}\right)^{n-1} \cdot \frac{1}{4} \cdot P_1}{\left(\frac{3}{4}\right)^{n-1} \cdot \frac{1}{4} \cdot P_1 + \left(\frac{2}{4}\right)^{n-1} \cdot \frac{2}{4} \cdot P_2 + \left(\frac{3}{4}\right)^{n-1} \cdot \frac{1}{4} \cdot P_3} \Rightarrow (\because 4^n)$$

$$\Rightarrow P(X=1|Y=n) = \frac{3^{n-1} \cdot P_1}{3^{n-1} \cdot P_1 + 2^n \cdot P_2 + 3^{n-1} \cdot P_3}$$

Analogamente

$$P(X=2|Y=n) = \frac{2^n \cdot P_2}{3^{n-1} \cdot P_1 + 2^n \cdot P_2 + 3^{n-1} \cdot P_3} \quad \text{e} \quad P(X=3|Y=n) = \frac{3^{n-1} \cdot P_3}{3^{n-1} \cdot P_1 + 2^n \cdot P_2 + 3^{n-1} \cdot P_3}$$

Assim,

j	$P(X=j Y=n)$
1	$\frac{3^{n-1} \cdot P_1}{3^{n-1} \cdot P_1 + 2^n \cdot P_2 + 3^{n-1} \cdot P_3}$
2	$\frac{2^n \cdot P_2}{3^{n-1} \cdot P_1 + 2^n \cdot P_2 + 3^{n-1} \cdot P_3}$
3	$\frac{3^{n-1} \cdot P_3}{3^{n-1} \cdot P_1 + 2^n \cdot P_2 + 3^{n-1} \cdot P_3}$

(b) W_n é v.a. tal que $W_n \stackrel{d}{=} X | Y = n$. Assim, W_n assume valores em $\{1, 2, 3\}$, $\forall n \geq 1$. Vamos avaliar $\lim_{n \rightarrow \infty} P(W_n = j)$, $j = 1, 2, 3$.

$$P(W_n = 1) = \frac{3^{n-1} \cdot P_1}{3^{n-1} \cdot P_1 + 2^n \cdot P_2 + 3^{n-1} \cdot P_3} \stackrel{(\div 3^{n-1})}{=} \frac{P_1}{P_1 + \left(\frac{2}{3}\right)^{n-1} \cdot 2P_2 + P_3} \xrightarrow{n \rightarrow \infty} \frac{P_1}{P_1 + P_3}.$$

Analogamente,

$$P(W_n = 2) = \frac{\left(\frac{2}{3}\right)^{n-1} \cdot 2P_2}{P_1 + \left(\frac{2}{3}\right)^{n-1} \cdot 2P_2 + P_3} \xrightarrow{n \rightarrow \infty} 0$$

$$P(W_n = 3) = \frac{P_3}{P_1 + \left(\frac{2}{3}\right)^{n-1} \cdot 2P_2 + P_3} \xrightarrow{n \rightarrow \infty} \frac{P_3}{P_2 + P_3}.$$

Como $P(W_n = j) \xrightarrow{n \rightarrow \infty} P(W = j)$, onde $P(W = j) = \frac{P_j}{P_1 + P_3}$, $j = 1, P(W = j) = 0$, $j = 2$ e

$P(W = j) = \frac{P_3}{P_1 + P_3}$, $j = 3$, segue, pelo resultado visto em aula, que

$W_n \xrightarrow{d} W$ onde W é tal que (distribuído segundo)

j	$P(W = j)$
1	$\frac{P_1}{P_1 + P_3}$
3	$\frac{P_3}{P_1 + P_3}$

ESTÃO 3

Para $f \in \text{IF}$, seja $X_i(f) = \begin{cases} 1, & \text{se } i \text{ pertence à imagem de } f \\ 0, & \text{caso contrário} \end{cases}, i=1, \dots, m.$

Assim, para todo $f \in \text{IF}$,

$$X(f) = X_1(f) + \dots + X_m(f), \text{ isto é, } X = \sum_{i=1}^m X_i$$

$$\begin{aligned} \text{Logo, } E(X) &= E\left(\sum_{i=1}^m X_i\right) = \sum_{i=1}^m E(X_i) = \sum_{i=1}^m \left\{ 1 \cdot P(X_i=1) + 0 \cdot P(X_i=0) \right\} = \sum_{i=1}^m P(X_i=1) = \\ &= \sum_{i=1}^m \left(1 - \left(\frac{m-1}{m}\right)^n \right) \Rightarrow E(X) = m \left[1 - \left(\frac{m-1}{m}\right)^n \right]. \end{aligned}$$

Note que $X_i \sim \text{Ber}\left(1 - \left(\frac{m-1}{m}\right)^n\right)$

$$\begin{aligned} \text{VAR}(X) &= \text{VAR}\left(\sum_{i=1}^m X_i\right) = \sum_{i=1}^m \text{Var}(X_i) + \sum_{i=1}^m \sum_{j \neq i} \text{Cov}(X_i, X_j) = \\ &= \sum_{i=1}^m \left(\frac{m-1}{m}\right)^n \left[1 - \left(\frac{m-1}{m}\right)^n \right] + \sum_{i=1}^m \sum_{j \neq i} \left\{ E(X_i X_j) - E(X_i)E(X_j) \right\}, \quad (*). \end{aligned}$$

$$\begin{aligned} \text{Mas, } E(X_i X_j) &= P(X_i=1, X_j=1) = 1 - P(X_i=0) - P(X_j=0) + P(X_i=0, X_j=0) = \\ &= 1 - \left(\frac{m-1}{m}\right)^n - \left(\frac{m-1}{m}\right)^n + \left(\frac{m-2}{m}\right)^n \quad (\text{por que?}). \end{aligned}$$

$$\text{Logo, } \text{Cov}(X_i, X_j) = 1 - 2\left(\frac{m-1}{m}\right)^n + \left(\frac{m-2}{m}\right)^n - \left[1 - \left(\frac{m-1}{m}\right)^n \right]^2$$

Assim,

$$\text{VAR}(X) = m \left(\frac{m-1}{m}\right)^n \left[1 - \left(\frac{m-1}{m}\right)^n \right] + m(m-1) \left\{ 1 - 2\left(\frac{m-1}{m}\right)^n + \left(\frac{m-2}{m}\right)^n - \left[1 - \left(\frac{m-1}{m}\right)^n \right]^2 \right\}$$

QUESTÃO 4

$$(a) \begin{array}{c} X_n \xrightarrow{\text{q.c.}} X \\ Y_n \xrightarrow{\text{q.c.}} X \end{array} \Rightarrow X_n - Y_n \xrightarrow{\text{q.c.}} X - X = 0.$$

Sug: $A = \{w \in \Omega : X_n(w) - Y_n(w) \rightarrow 0\}$.

$$w_0 \in A \Rightarrow X_n(w_0) - Y_n(w_0) \rightarrow 0 \Rightarrow \exists k \in \mathbb{N} \text{ tal que } n \geq k \Rightarrow |(X_n(w_0) - Y_n(w_0)) - 0| < \frac{1}{100} \Rightarrow$$

$$\Rightarrow \exists k \in \mathbb{N}, n \geq k \mid X_n(w_0) - Y_n(w_0) \mid < \frac{1}{100} \Rightarrow \exists k \in \mathbb{N} \text{ tal que } n \geq k \Rightarrow X_n(w_0) = Y_n(w_0), \text{ pois}$$

$$X_n \subset Y_n \text{ armazém valores inteiros} \Rightarrow \exists k \in \mathbb{N} \text{ tal que } n \geq k \Rightarrow X_n(w_0) = Y_n(w_0) \Rightarrow$$

$$\Rightarrow w_0 \in \left\{ w \in \Omega : \exists k \in \mathbb{N}, n \geq k \Rightarrow X_n(w) = Y_n(w) \right\} \Rightarrow w_0 \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{X_n(w) = Y_n(w)\} \Rightarrow$$

$$\Rightarrow w_0 \in \liminf_{n \rightarrow \infty} A_n. \text{ Logo, } A \subseteq \liminf A_n \Rightarrow P(\liminf A_n) = 1 \text{ (verdadeiro!)}$$

$$(b) X_n(\Omega) \subseteq \mathbb{R}_+$$

$$Y_n(\Omega) \subseteq \mathbb{R}_+.$$

$$\text{Sug: } \Omega = [0, 1]$$

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

P: 'comprimento'. Sug: sim,

$$X_n(w) = w + \frac{1}{n} \quad \text{e} \quad Y_n(w) = w + \frac{2}{n}.$$

Claro que $X_n \xrightarrow{\text{q.c.}} X$ e $Y_n \xrightarrow{\text{q.c.}} X$, onde $X(w) = w, \forall w \in [0, 1]$.

$$\left(\{w : X_n(w) \rightarrow X(w)\} = \{w : Y_n(w) \rightarrow X(w)\} = [0, 1] \right).$$

$$\text{Mas, } \forall n \geq 1, \{w \in \Omega : X_n(w) = Y_n(w)\} = \emptyset \quad \left(X_n(w) = Y_n(w) = \frac{1}{n}, \forall w \right).$$

Logo, $A_n = \emptyset, \forall n \geq 1$, portanto, $\liminf A_n = \emptyset$, donde $P(\liminf A_n) = 0$.
(falso!).

- ESCREVA SE NOME E NÚMERO USP NO CADERNO DE RESPOSTAS.

- VOCÊ PODE CONSULTAR UM RESUMO DE ATÉ 2 PÁGINAS DO CONTEÚDO DA DISCIPLINA.

- SEMPRE EXPLIQUE SEU RACIOCÍNIO !!!!!

(1) SEJA X UMA (ÚNICA) VARIÁVEL ALEATÓRIA QUE, DADO θ , $\theta > 0$, POSSUI FUNÇÃO DENSIDADE DE PROBABILIDADE DADA POR

$$f(x|\theta) = \frac{1}{3\theta} \mathbb{I}_{[-\theta, 2\theta]}(x) \quad (\text{ISTO É, } X|\theta \sim \text{Uniforme}(-\theta, 2\theta)).$$

(a) OBTENHA O ESTIMADOR DE MÁXIMA VEROSIMILHANÇA PARA θ , δ_{MV} . QUAL É A ESTIMATIVA DE MV SE $X = -0,9$? E SE $X = 1,6$? PARA QUE VALOR DE $C \in \mathbb{R}$ O ESTIMADOR $\delta_C(x) = C \cdot \delta_{MV}(x)$ É NÃO-VIESADO PARA θ ?

(b) SUPONHA QUE, A PRIORI, $\theta \sim \text{GAMA}(2,2)$. DETERMINE A DISTRIBUIÇÃO A POSTERIORI DE θ DADO $X=x$, $x \in \mathbb{R}$. QUAL É A ESTIMATIVA DE BAYES COM RELAÇÃO À PERDA QUADRÁTICA SE $X = -0,9$?

(2) SEJAM X_1, \dots, X_n VARIÁVEIS ALEATÓRIAS QUE, DADO θ , POSSUEM FUNÇÃO DE PROBABILIDADE (CONJUNTA) DADA POR

$$\Pr(X_1=x_1, X_2=x_2, \dots, X_n=x_n | \theta) = \frac{\theta \Gamma(\theta + \sum_{i=1}^n x_i) \Gamma(n+1 - \sum_{i=1}^n x_i)}{\Gamma(n+1+\theta)} \cdot \prod_{i=1}^n \mathbb{I}_{\{0,1\}}(x_i).$$

(a) Especifique o espaço amostral, \mathcal{X} , e obtenha uma estatística $T: \mathcal{X} \rightarrow \mathbb{N}$ suficiente para Θ .

(b) Para $n = 20$, obtenha o teste mais poderoso de nível $\alpha = \frac{2}{21}$ para testar as

Hipóteses $H_0: \Theta = 1$ versus $H_1: \Theta = 5$. Dica: Mostre que

$$\sum_{i=1}^{20} X_i \mid \Theta = 1 \sim \text{Uniforme}\{0, 1, 2, \dots, 20\}.$$

(c) Qual é a probabilidade de erro do tipo II do teste MP obtido em (b)?

(3) Seja $\mathbb{H} \subseteq \mathbb{R}$. Sejam $A, B: \mathcal{X} \rightarrow \mathbb{R}$ tais que $A(x) \leq B(x), \forall x \in \mathcal{X}$.

(a) Mostre que se $[A(x), B(x)]$ é um intervalo de confiança para Θ com coeficiente de confiança $1-\alpha$, $0 < \alpha < 1$, e $g: \mathbb{H} \rightarrow \mathbb{R}$ é não-decrescente, então $[g(A(x)), g(B(x))]$ é um intervalo de confiança para $g(\Theta)$ com confiança $1-\alpha$. Enuncie resultado similar (equivalente) para $g: \mathbb{H} \rightarrow \mathbb{R}$ não-crescente.

(b) Sejam X_1, \dots, X_n , variáveis aleatórias que, dado Θ , $\Theta > 0$, são condicionalmente independentes e identicamente distribuídas segundo o modelo beta de parâmetros Θ e 1. Construa um intervalo de confiança para $E(X_1 | \Theta)$ com coeficiente de confiança $1-\alpha$, $0 < \alpha < 1$, baseado em uma estatística suficiente para Θ .

(4) Sejam X_1, \dots, X_n , $n \geq 3$, variáveis aleatórias que, dado Θ , $0 < \Theta < 1$, são condicionalmente independentes e identicamente distribuídas segundo o modelo Bernoulli de parâmetro Θ .

(a) Obtenha, se possível, o enunciado para $\hat{\Theta}^r$, $r \in \{2, \dots, n-1\}$.

(b) Determine um limitante inferior (maior que zero !!) para a variância do estimador obtido em (a).

(c) Considere, para cada $n \geq 3$, o estimador $\hat{\delta}_n$ obtido em (a). Mostre que a sequência $(\hat{\delta}_n)_{n \geq 3}$ é consistente para $\hat{\Theta}$.

OBSERVAÇÕES :

- 1) Desligue IMEDIATAMENTE seu telefone celular ou qualquer aparelho que possa produzir barulho.
- 2) A avaliação tem duração de 2 horas e 30 minutos.
- 3) SEMPRE JUSTIFIQUE SUAS RESPOSTAS.

1. Seja (X, Y) um vetor aleatório com distribuição de probabilidade dada por:

$$P(X = x, Y = y) = \int_0^1 \theta^{x+y} (1-\theta)^{2-x-y} f(\theta) d\theta I_{\{0,1\}}(x) I_{\{0,1\}}(y),$$

onde f é uma função densidade de probabilidade tal que $\int_0^1 f(\theta) d\theta = 1$. Mostre que $COV(X, Y) > 0$. (1,5)

2. Mostre ou exiba um contra-exemplo.

(a) Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade. Sejam $A, B, D \in \mathcal{F}$, $A \subset B$. Se A é independente de D e B é independente de D , então $B \cap A^c$ é independente de D . (1,0)

(b) Seja $F : \mathbb{R} \rightarrow \mathbb{R}_+$ uma função de distribuição. A função $G : \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$G(t) = \frac{F(t) + 1 - F(-t)}{2}$$

$P(X=t) = P(X \leq t) = F(t)$

$P(X=-t) = P(X \leq -t) = 1 - F(-t)$

$\frac{1}{2}(F(t) + 1 - F(-t)) + \frac{1}{2}$

$\frac{1}{2}(F(t) - F(-t)) + \frac{1}{2}$

$\frac{1}{2}(F(t) - F(-t)) + \frac{1}{2}$

é função de distribuição. (1,0)

3. Determine

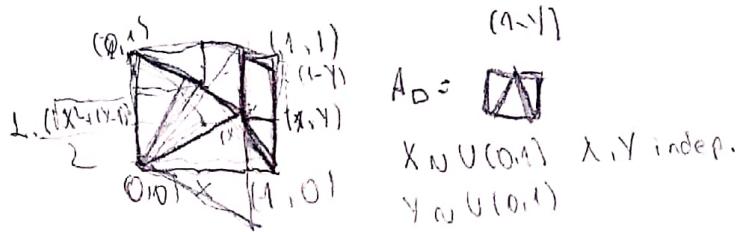
$$\lim_{n \rightarrow \infty} \int_n^\infty \frac{1}{2^n(n-1)!} t^{n-1} e^{-t/2} dt. \quad B \cap A = A$$

(1,5)

$$\begin{aligned} & \text{Let } u = t/2, \quad du = dt \\ & \int_n^\infty \frac{1}{2^n(n-1)!} t^{n-1} e^{-t/2} dt = \int_n^\infty \frac{1}{2^n(n-1)!} (2u)^{n-1} e^{-u} 2 du \\ & = \frac{1}{2^n(n-1)!} \int_n^\infty (2u)^{n-1} e^{-u} du \\ & = \frac{1}{2^n(n-1)!} \Gamma(n) \end{aligned}$$

$$\begin{aligned} B &= B \cap A \cup B \cap A^c \\ B &= A \cup B \cap A^c \\ P(B \cap A) &= P(A \cup B \cap A^c) \cap D = \\ &= P[(A \cap D) \cup (B \cap A^c \cap D)] = \end{aligned}$$

$$\begin{aligned} P(B) P(D) &= P(A) P(D) + P(B \cap A^c \cap D) \\ &= P(D)(P(B) - P(A)) = \\ &= P(D)(P(B \cap A^c)) \end{aligned}$$



4. Um ponto é escolhido ao acaso no interior do quadrado $[0, 1]^2$. Sejam X e Y as coordenadas do ponto selecionado. Sejam U : a área do triângulo com vértices $(0, 0)$, $(1, 0)$ e (X, Y) e V : a área do quadrilátero com vértices (X, Y) , $(X, 1)$, $(1, 0)$ e $(1, 1)$.

- (a) Determine a função densidade de probabilidade do vetor aleatório (U, V) . (1,0)
- (b) Determine a distribuição condicional de V dado $U = u$, $u \in (0, 1/2)$. (1,0)
- (c) Determine a variância de V . (1,0)

5. Uma urna contém n bolas numeradas de 1 a n , $n \geq 3$. Uma amostra de k bolas, $k \in \mathbb{N}$, $2 \leq k < n$, é retirada "ao acaso" da urna (isto é, supõe-se que todos os subconjuntos de k bolas são equiprováveis) e são registrados os números das bolas selecionadas. Seja X_n o número de pares de números consecutivos registrados (por exemplo, se $n = 10$, $k = 6$ e se são retiradas as bolas 2, 3, 4, 6, 8 e 9, então $X_n(\{2, 3, 4, 6, 8, 9\}) = 3$).

- (a) Determine a esperança e a variância de X_n . (1,5)
- (b) Mostre que $X_n \xrightarrow{P} 0$ quando $n \rightarrow \infty$ (com k fixo). (1,0)

List 1

Exercício 1: Sejam $a, b \in \mathbb{R}$, $a < b$. Mostre que

- a) $(a, b) = \bigcup_{n=1}^{\infty} \left(a, b - \frac{1}{n}\right] = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b\right)$
- b) $(a, b] = \bigcap_{n=1}^{\infty} \left(a, b + \frac{1}{n}\right)$

Exercício 2: Mostre que os conjuntos \overline{A} , \underline{A} e $\lim_{n \rightarrow \infty} A_n$, sempre que existirem, pertencem a \mathcal{F} , usando o fato de que $A_n \in \mathcal{F}$, $n \geq 1$.

Exercício 3: Sejam A e B subconjuntos de Ω e para $n = 1, 2, \dots$, defina os conjuntos A_n da seguinte forma: $A_{2n-1} = A$, $A_{2n} = B$. Sabendo disso, mostre que

$$\liminf_{n \rightarrow \infty} A_n = A \cap B, \quad \limsup_{n \rightarrow \infty} A_n = A \cup B.$$

Exercício 4: Seja $\Omega = \mathbb{R}$ e defina os subconjuntos A_n, B_n , $n = 1, 2, \dots$ de Ω como segue:

$$A_n = \left\{ x \in \mathbb{R} : -5 + \frac{1}{n} < x < 20 - \frac{1}{n} \right\}, \quad B_n = \left\{ x \in \mathbb{R} : 0 < x < 7 + \frac{3}{n} \right\}.$$

Mostre que A_n é não decrescente e B_n é não crescente. Encontre os limites A e B dessas sequências.
(Nesse caso, denotamos $A_n \uparrow A$ e $B_n \downarrow B$).

Exercício 5: Seja $\Omega = \mathbb{R}^2$ e defina os subconjuntos A_n, B_n , $n = 1, 2, \dots$ de Ω como segue:

$$A_n = \left\{ (x, y) \in \mathbb{R}^2 : 3 + \frac{1}{n} \leq x < 6 - \frac{2}{n}, 0 \leq y \leq 2 - \frac{1}{n^2} \right\},$$

$$B_n = \left\{ (x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq \frac{1}{n^3} \right\}.$$

Mostre que $A_n \uparrow A$ e $B_n \downarrow B$ e identifique A e B .

Exercício 6: Determine os limites inferiores e superiores das sequências:

a) $A_n = \{(-1)^n\}$

b) $A_n = \begin{cases} [0, n], & n \text{ par} \\ [-n, 0], & n \text{ ímpar} \end{cases}$

c) $A_n = \left\{ (x, y) \in \mathbb{R}^2 : \left(x - \frac{(-1)^n}{n} \right)^2 + y^2 \leq 1 \right\}$

Exercício 7: Sejam A_1, a_2, \dots eventos aleatórios. Mostre que:

a. $P\left(\bigcap_{k=1}^n A_k\right) \geq 1 - \sum_{k=1}^n P(A_k^c)$.

b. Se $P(A_k) \geq 1 - \varepsilon$, para $k = 1, 2, \dots, n$, então $P\left(\bigcap_{k=1}^n A_k\right) \geq 1 - n\varepsilon$.

c. $P\left(\bigcap_{k=1}^{\infty} A_k\right) \geq 1 - \sum_{k=1}^{\infty} P(A_k^c)$.

Exercício 8: Demonstre as seguintes propriedades:

a. Se $P(A_n) = 0$ para $n = 1, 2, \dots$, então $P\left(\bigcup_{n=1}^{\infty} A_n\right) = 0$;

b. Se $P(A_n) = 1$ para $n = 1, 2, \dots$, então $P\left(\bigcap_{n=1}^{\infty} A_n\right) = 1$.

Exercício 9: Demonstre: se A_1, A_2, \dots e B_1, B_2, \dots são eventos aleatórios do mesmo espaço de probabilidade tais que $P(A_n) \rightarrow 1$ e $P(B_n) \rightarrow p$, quando $n \rightarrow \infty$, então $P(A_n \cap B_n) \rightarrow p$.

Exercício 10: Considere $\Omega = \{1, 2, 3, 4\}$. Verifique se \mathcal{A} e \mathcal{B} são σ -álgebras.

a) $\mathcal{A} = \{\emptyset, \Omega, \{1\}, \{2\}, \{1, 2\}, \{3, 4\}, \{1, 3, 4\}, \{2, 3, 4\}\}$

b) $\mathcal{B} = \{\emptyset, \Omega, \{1\}, \{2, 3, 4\}, \{2\}, \{1, 3, 4\}, \{3\}, \{1, 2, 4\}\}$

Exercício 11: Sejam A_1, A_2, \dots, A_n eventos em $(\Omega, \mathcal{F}, \mathcal{P})$ com $p_n = P(A_n) \geq c > 0$, para todo $n \geq 1$. Mostre que $P(\limsup A_n) \geq c$.

Exercício 12: Seja Ω um conjunto não enumerável e defina

$$\mathcal{F} = \{A \subset \Omega : A \text{ é enumerável ou } A^c \text{ é enumerável}\}.$$

Mostre que \mathcal{F} é uma σ -álgebra de subconjuntos de Ω .

Exercício 13: Seja \mathcal{F} uma σ -álgebra de conjuntos de Ω . Fixamos $B \in \mathcal{F}$ e definimos uma nova coleção de subconjuntos

$$\mathcal{F} = \{C : C = A \cap B, \text{ para algum } A \in \mathcal{F}\}.$$

Mostre que \mathcal{F} é uma σ -álgebra de subconjuntos de B .

Exercício 14: Considere o espaço de probabilidade $(\Omega, \mathcal{F}, \mathcal{P})$. Seja

$$\mathcal{F}_1 = \{A : A \in \mathcal{F} \text{ e } P(A) = 1\}.$$

Mostre que \mathcal{F}_1 é uma σ -álgebra de conjuntos de Ω .

Exercício 15: Seja $(\Omega, \mathcal{F}, \mathcal{P})$ um espaço de probabilidade e para cada $B \in \mathcal{F}$ defina a seguinte classe de subconjuntos:

$$\mathcal{F}_B = \{A : A \in \mathcal{F} \text{ e } P(A \cap B) = 0 \text{ ou } 1\}.$$

Mostre que \mathcal{F}_B é uma σ -álgebra de conjuntos de Ω .

Exercício 16: Considere $\Omega = \{1, 2, 3, 4, 5, 6\}$. Determine a σ -álgebra gerada por

- a) $\mathcal{C}_1 = \{\{1, 2\}, \{3, 4, 5, 6\}\}$
- b) $\mathcal{C}_2 = \{\{1\}, \{2\}, \{3, 4, 5, 6\}\}$
- c) $\mathcal{C}_3 = \{\{1, 2\}, \{2, 4, 5, 6, 3\}\}$

Exercício 17: Considere (A_1, A_2, \dots, A_n) uma partição de Ω (isto é, $A_i \cap A_j = \emptyset$, se $i \neq j$, e $\Omega = A_1 \cup A_2 \cup \dots \cup A_n$). Determine $\sigma(\{A_1, \dots, A_n\})$.

Exercício 18: Seja $\Omega = \mathbb{R}$. Mostre que a σ -álgebra gerada por $\mathcal{C} = \{(-\infty, x] : x \in \mathbb{R}\}$ concide com a σ -álgebra gerada por $\mathcal{D} = \{[x, y) : x, y \in \mathbb{R}, x < y\}$.

Exercício 19: Considere $\Omega = \mathbb{N}$ e $\mathcal{F} = \mathcal{P}(\mathbb{N})$. Seja $P : \mathcal{F} \rightarrow [0, 1]$ dada por $P(A) = 1$, se A é infinito, e $P(A) = 0$, se A é finito. P é medida de probabilidade?

Exercício 20: Seja Ω não-enumerável e $\mathcal{F} = \{A \subseteq \Omega : A \text{ ou } A^c \text{ é enumerável}\}$. Seja $P : \mathcal{F} \rightarrow [0, 1]$ dada por $P(A) = 1$, se A é não-enumerável, e $P(A) = 0$, se A é enumerável. P é medida de probabilidade?

Exercício 21: Seja $(P_n)_{n \geq 1}$ sequência de medidas de probabilidades em (Ω, \mathcal{F}) . $P : \mathcal{F} \rightarrow [0, 1]$ dada por

$$P(A) = \sum_{n=1}^{\infty} \alpha_n P_n(A), \quad \alpha_n \geq 0, \quad \forall n \geq 1, \quad \text{com} \quad \sum_{n=1}^{\infty} \alpha_n = 1, \quad \text{é medida de probabilidade?}$$

Exercício 22: Considere (Ω, \mathcal{F}, P) e $(A_n)_{n \geq 1}$ uma sequência de eventos ($A_n \in \mathcal{F}, \forall n \geq 1$). Mostre que $P(\liminf A_n) \leq \liminf P(A_n)$ e $\limsup P(A_n) \leq P(\limsup A_n)$.

List 2

(George Roussas - Cap 2)

Exercício 1: Suponha que os eventos $A_j, j = 1, \dots$ são tais que

$$\sum_{j=1}^{\infty} P(A_j) < \infty.$$

Use a definição 3 do capítulo 1 e o teorema 2 do capítulo 2 para mostrar que $P(\overline{A}) = 0$.

Exercício 2: Seja $\{A_j, j = 1, \dots, 5\}$ uma partição de Ω e suponha que $P(A_j) = j/15$ e $P(A|A_j) = (5-j)/15, j = 1, \dots, 5$. Calcule as probabilidades $P(A_j|A), j = 1, \dots, 5$.

Exercício 3: Se $P(A|B) > P(A)$, então mostre que $P(B|A) > P(B)$.

Exercício 4: Mostre que:

i) $P(A^c|B) = 1 - P(A|B)$;

ii) $P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$.

Mostram também, por meio de contra-exemplos, que as seguintes equações não precisam ser verdadeiras:

iii) $P(A|B^c) = 1 - P(A|B)$;

iv) $P(C|A + B) = P(C|A) + P(C|B)$.

Exercício 5: Considere três urnas $U_j, j = 1, 2, 3$ de modo que a urna U_j contém m_j bolas brancas e n_j bolas pretas. Uma urna é escolhida aleatoriamente e uma bola é retirada também aleatoriamente. Se a bola retirada foi branca, qual é a probabilidade que a urna sorteada foi U_1 ou U_2 .

Exercício 6: Considere k urnas U_j , $j = 1, \dots, k$ cada uma contendo m bolas brancas e n bolas pretas. Uma bola é retirada aleatoriamente da urna U_1 e é colocada na urna U_2 . Depois uma bola é retirada aleatoriamente da urna U_2 e é colocada na urna U_3 etc. Finalmente, uma bola é retirada aleatoriamente da urna U_{k-1} e é colocada na urna U_k . Calcule a probabilidade da última bola ser preta.

Exercício 7: Seja Ω um conjunto de todas n^3 -letras de uma língua e seja P a medida de probabilidade igualmente provável sobre os eventos de Ω . Defina os eventos, A , B e C da seguinte forma:

$$A = \{s \in \Omega; s \text{ começa com uma letra especificada}\}$$

$$B = \{s \in \Omega; s \text{ tem a letra específica (mencionada na definição de } A) \text{ no meio}\}$$

$$C = \{s \in \Omega; s \text{ tem exatamente duas letras iguais}\}$$

Em seguida, mostre que:

$$P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C),$$

$$P(B \cap C) = P(B)P(C), \quad P(A \cap B \cap C) \neq P(A)P(B)P(C).$$

Assim os eventos A , B e C são dois-a-dois independentes mas não mutuamente independentes.

(Barry James - Cap 1)

Exercício 8: (Problema dos casamentos.)

a) No exercício 1 (do livro), se você identificasse ao acaso as equações e frases, qual a probabilidade de que você acertaria pelo menos um casamento?

b) Resolva o item a) para p caso em que há n equações e n frases para serem identificadas. (Sugestão. Use o exercício 13 b) (do livro).)

c) Mostra que a probabilidade considerada no item b) converge para $p = \frac{1}{e}$ quando $n \rightarrow \infty$.

Exercício 9: Seja (Ω, \mathcal{F}, P) um espaço de probabilidade e suponha que todos os conjuntos abaixo pertençam a \mathcal{F} . Prove:

a) Se os A_n são disjuntos e $P(B|A_n) \geq c$ para todo n , então $P(B| \cup A_n) \geq c$ (pode supor $P(A_n) > 0$ para todo n).

- b)** O item a) com " $=$ " no lugar de " \geq ".
- c)** Se $A_n \supset A_{n+1}$ e $P(A_{n+1}|A_n) \leq \frac{1}{2}$ para todo n , então $P(A_n) \rightarrow 0$ quando $n \rightarrow \infty$.
- d)** Se os A_n são disjuntos e $P(B|A_n) = P(C|A_n) \forall n$, então

$$P(B| \cup A_n) = P(C| \cup A_n).$$

- e)** Se A_1, A_2, \dots são disjuntos e $\cup A_n = \Omega$, então

$$P(B|C) = \sum_n P(A_n|C)P(B|A_n \cap C).$$

Exercício 10: Suponha que a ocorrência ou não de chuva dependa das condições do tempo no dia imediatamente anterior. Admita-se que se chove hoje, choverá amanhã com probabilidade 0,7 e que se não chove hoje choverá amanhã com probabilidade 0,4; Sabendo-se que choveu hoje, calcule a probabilidade de que choverá depois de amanhã.

Exercício 11: Em teste de múltipla escolha, a probabilidade do aluno saber a resposta é p . Havendo m escolhas, se ele sabe a resposta ele responde corretamente com probabilidade 1, se não sabe ele responde corretamente com probabilidade $\frac{1}{m}$. Qual a probabilidade que ele sabia a resposta dado que a pergunta foi respondida corretamente?

Exercício 12: Sejam A_1, \dots, A_n eventos aleatórios independentes, com $p_k = P(A_k)$, $k = 1, \dots, n$. Obtenha a probabilidade de ocorrência dos seguintes eventos, em termos das probabilidades p_k :

- a)** A ocorrência de nenhum dos A_k . **b)** A ocorrência de pelo menos um dos A_k . **c)** A ocorrência de exatamente um dos A_k . **d)** A ocorrência de exatamente dois dos A_k . **e)** A ocorrência de todos os A_k . **f)** A ocorrência de no máximo, $n - 1$ dos A_k .

(Sheldon Ross - Cap 3)

Exercício 13: Prove ou dê contra exemplos das seguintes afirmações

- a)** Se E é independente de F e E é independente de G então E é independente de $F \cup G$.
- b)** Se E é independente de F , E é independente de G e $F \cap G = \emptyset$ então E é independente de $F \cup G$.
- c)** Se E é independente de F , F é independente de G e E é independente de $F \cap G$ então G é independente de $E \cap F$.

Exercício 14: Suponha que $(E_n)_{n \geq 1}$ e $(F_n)_{n \geq 1}$ são sequências crescentes de eventos com limites E e F , respectivamente. Mostre que se E_n é independente de F_n para todo n , então E é independete de F .

Exercício 15: Se $0 \leq a_i \leq 1$, $i = 1, 2, \dots$, mostre que

$$\sum_{i=1}^{\infty} \left[a_i \prod_{j=1}^{i-1} (1 - a_j) \right] + \prod_{i=1}^{\infty} (1 - a_i) = 1$$

DICA: Suponha que um número infinito de moedas são jogadas. Seja a_i a probabilidade de cara da i -ésima moeda e considere a primeira ocorrência de cara.

Exercício 16: Ensaios independentes com probabilidade de sucesso p e probabilidade de falha $1 - p$ são chamados de ensaios de Bernoulli. Seja P_n denotar a probabilidade que em n ensaios de Bernoulli o resultado é um número par de sucessos (0 é considerado como um número par). Mostre que

$$P_n = p(1 - P_{n-1}) + (1 - p)P_{n-1} \quad n \geq 1$$

e use isso para orivar (por indução) que

$$P_n = \frac{1 + (1 - 2p)^n}{2}$$

Exercício 17: Considere duas urnas, ambas contêm bolas brancas e pretas. A probabilidade de retirar bolas brancas da primeira e segunda urna são, respectivamente, p e p' . As bolas são sequencialmente selecionadas com reposição da seguinte forma: Com probabilidade α a bola é inicialmente escolhida da primeira urna, e com probabilidade $1 - \alpha$ é escolhida da segunda urna. As seguinte extrações são feitas da seguinte forma: sempre que uma bola branca for retirada (e devolvida) a bola seguinte será retirada da mesma urna, caso contrário, será retirada da outra urna. Seja α_n denotar a probabilidade que a n -ésima bola é retirada da primeira urna. Mostre que

$$\alpha_{n+1} = \alpha_n(p + p' - 1) + 1 - p' \quad n \geq 1$$

e use isso para provar que

$$\alpha_n = \frac{1 - p'}{2 - p - p'} + \left(\alpha - \frac{1 - p'}{2 - p - p'} \right) (p + p' - 1)^{n-1}$$

Seja P_n denota a probabilidade que a n -ésima bola selecionada é branca. Encontre P_n . Também calcule $\lim_{n \rightarrow \infty} \alpha_n$ e $\lim_{n \rightarrow \infty} P_n$.

Exercício 18: Como um modelo simplificado para previsão do tempo, suponha que o clima de amanhã (chuva ou sol) será o mesmo clima de hoje com probabilidade p . Se 1 de janeiro for sol, mostre que P_n , a probabilidade de sol em n dias depois satisfaz

$$P_n = (2p - 1)P_{n-1} + (1 - p) \quad n \geq 1$$

$$P_0 = 1$$

Prove que

$$P_n = \frac{1}{2} + \frac{1}{2}(2p - 1)^n \quad n \geq 0$$

Exercício 19: Provar ou dar um contra-exemplo. Se E_1 e E_2 são independentes, então eles são condicionalmente independentes dado F .

List 3

(1-6. Marcos Magalhães - Cap 2.5)

Exercício 1: Sendo c uma constante real qualquer, verifique que $X(\omega) = c$, para todo $\omega \in \Omega$, é variável aleatória em qualquer σ -álgebra de conjuntos de Ω .

Exercício 2: Considere um espaço de probabilidade $(\Omega, \mathcal{F}, \mathcal{P})$ e sejam E e $F \subset \Omega$ com $E \in \mathcal{F}$ mas $F \notin \mathcal{F}$. Mostre que a função indicadora de E é variável aleatória mas a de F não é.

Exercício 3: Sejam A_1, A_2, \dots, A_n eventos em $(\Omega, \mathcal{F}, \mathcal{P})$, verifique se $X = \sum_{i=1}^n I_{A_i}$ e $Y = \prod_{i=1}^n I_{A_i}$ são variáveis aleatórias.

Exercício 4: Seja $(\Omega, \mathcal{F}, \mathcal{P})$ um espaço de probabilidade com os eventos A, B e C . Mostre que $X = I_{A \cap B} + I_C$ é variável aleatória nesse espaço.

Exercício 5: Seja X uma variável aleatória com distribuição F . Mostre que, para todo número real c , $P(X = c) = F(c) - F(c^-)$.

Exercício 6: Mostre que se F e G forem funções de distribuição então, para $0 \leq \alpha \leq 1$, $\alpha F + (1 - \alpha)G$ também é função de distribuição.

Exercício 7: Sejam $(\Omega, \mathcal{F}, \mathcal{P})$ e $X : \Omega \rightarrow \mathbb{R}$ uma variável aleatória. Mostre que $\{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$ é uma σ -álgebra de subconjuntos de Ω .

Exercício 8: Sejam $F_1, F_2, \dots, F_n : \mathbb{R} \rightarrow [0, 1]$ funções de distribuição. Verifique se F, G e H definidas abaixo são funções de distribuição.

a) $F = F_1 F_2 \cdots F_n$ ($F(x) = F_1(x) F_2(x) \cdots F_n(x)$, $x \in \mathbb{R}$)

b) $G = F_1 + F_2 + \cdots + F_n$

c) $H = 1 - (1 - F_1)(1 - F_2) \cdots (1 - F_n)$

(9-27. Barry James - Cap 2)

Exercício 9: Seja X uma variável aleatória com distribuição de Poisson, parâmetro $\lambda > 0$. Mostre que a função de distribuição de X é

$$F(x) = A_n = \begin{cases} \frac{1}{n!} \int_{\lambda}^{\infty} e^{-t} t^n dt & \text{se } n \leq x \leq n+1, n = 0, 1, 2, \dots \\ 0 & \text{se } x < 0. \end{cases}$$

Exercício 10: Suponha que a vida útil de certo tipo de lâmpada tenha distribuição exponencial com parâmetro λ .

a) (*Falta de memória* da distribuição exponencial. Compare com o exercício 27c) do Capítulo 1.) Seja T a vida de uma lâmpada desse tipo. Mostre que

$$P(T > t + s | T > t) = P(T > s) \forall s, t > 0.$$

b) Suponha que $\lambda = 3$ quando a vida é expressa em dias. Uma lâmpada solitária é ligada em uma sala no instante $t = 0$. Um dia depois, você entra na sala e fica ali durante 8 horas, saindo no final desse período.

i. Qual a probabilidade de que você entre na sala quando já está escura?

ii. Qual a probabilidade de você entrar na sala com a lâmpada ainda acesa e sair da sala depois da lâmpada queimar?

Exercício 11: Seja X uma variável aleatória com densidade

$$f(x) = \begin{cases} \frac{1}{(1+x)^2}, & \text{se } x > 0 \\ 0, & \text{se caso contrário.} \end{cases}$$

Seja $Y = \max(X, c)$, onde c é uma constante > 0 .

a) Ache a função de distribuição de Y .

b) Decomponha F_Y em suas partes discreta, absolutamente contínua e singular.

Exercício 12: Uma urna contém três bolas numeradas 1, 2, e 3. Duas bolas são tiradas sucessivamente da urna, ao acaso e *sem* reposição. Seja X o número da primeira bola tirada e Y o número da segunda.

a) Descreva a distribuição conjunta de X e Y .

b) Calcule $P(X < Y)$.

Exercício 13: Seleciona-se, ao acaso, um ponto do círculo unitário $\{(x, y) : x^2 + y^2 \leq 1\}$. Sejam X e Y as coordenadas do ponto selecionado.

a) Qual a densidade conjunta de X e Y ?

b) Determine $P(X < Y)$, $P(X > Y)$ e $P(X = Y)$.

Exercício 14: Seleciona-se, ao acaso, um ponto do quadrado unitário $\{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$. Sejam X e Y as coordenadas do ponto selecionado.

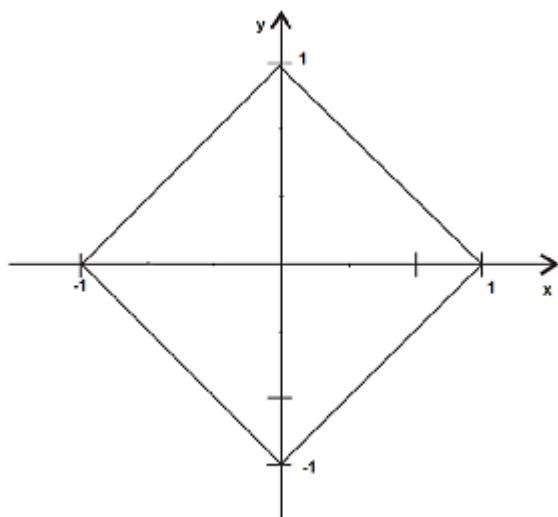
a) Qual a densidade conjunta de X e Y ?

b) Calcule $P(|Y/X - 1| \leq 1/2)$.

c) Calcule $P(Y \geq X | Y \geq 1/2)$.

Exercício 15: Demonstre: se a variável aleatória X é independente de si mesma, então X é constante com probabilidade 1 (i.e.m existe uma constante c tal que $P(X = c) = 1$).

Exercício 16: Um ponto é selecionado, ao acaso (i.e., conforme a distribuição uniforme), do seguinte quadrado:



Sejam X e Y as coordenadas do ponto selecionado.

- a)** Qual a densidade conjunta de X e Y ?
- b)** Obtenha a densidade marginal de X .
- c)** X e Y são independentes?

Exercício 17: Sejam X e Y variáveis aleatórias independentes com distribuição uniforme em $[\theta - 1/2, \theta + 1/2]$, onde $\theta \in \mathbb{R}$. Prove que a distribuição de $X - Y$ não depende de θ , achando sua densidade.

Exercício 18: Se X e Y são coordenadas de um ponto selecionado, ao acaso, do círculo unitário $\{(x, y) : x^2 + y^2 \leq 1\}$, qual a distribuição da variável aleatória $Z = X^2 + Y^2$?

Exercício 19: Sejam X e Y variáveis aleatórias independentes, tendo distribuição comum $U[0, 1]$.

- a)** Qual a densidade da variável aleatória $Z = X + Y$?
- b)** Ache a probabilidade da equação quadrática $Xt^2 + Yt + Z = 0$ ter raízes reais.

Exercício 20: Sejam X e Y variáveis aleatórias independentes, X tendo distribuição de Poisson com parâmetro $\lambda = 5$, e Y tendo distribuição uniforme em $[0, 1]$. Ache a densidade de $Z = X + Y$.

Exercício 21: Sejam X e Y variáveis aleatórias independentes com distribuição comum $\exp(\lambda)$. Prove que $Z = \frac{X}{X+Y} \sim U[0, 1]$.

Exercício 22:

Se X possui densidade $f(x)$, qual a densidade de $Y = \cos X$? (Sugestão: Use o exercício 54.)

Exercício 23: Sejam X e Y variáveis aleatórias independentes, tendo distribuição comum $U[0, 1]$, e sejam $R = \sqrt{2 \log(1/(1-x))}$ e $\Theta = \pi(2Y - 1)$.

a) Mostre que $\Theta \sim U[-\pi, \pi]$ e que R tem distribuição de Rayleigh com densidade

$$f(r) = \begin{cases} r e^{-r^2/2} & , \quad r > 0 \\ 0 & , \quad r \leq 0. \end{cases}$$

b) Mostre que Z e W , definidas por $R = \cos \Theta$ e $W = R \sin \Theta$, são independentes com distribuição comum $N(0, 1)$. (Observação. Este resultado é de interesse na simulação de variáveis aleatórias independentes e normais, pois indica como transformar números “pseudo-aleatórios” (simulações de variáveis aleatórias independentes e $U[0, 1]$) gerados por computador.)

Exercício 24: Suponha que X e Y sejam independentes com $X \sim \Gamma(\alpha_1, 1)$ e $Y \sim \Gamma(\alpha_2, 1)$, onde $\alpha_1 > 0$ e $\alpha_2 > 0$. Mostre que $X + Y$ e X/Y são independentes e ache as suas distribuições.

Exercício 25: Suponha que X_1, \dots, X_n sejam independentes e identicamente distribuídas com densidade comum f . Mostre que a densidade conjunta de $U = \min_{1 \leq i \leq n} X_i$ e $V = \max_{1 \leq i \leq n} X_i$ é

$$f_{U,V}(u, v) = \begin{cases} n(n-1)[F(v) - F(u)]^{n-2} & , \text{ se } u < v \\ 0 & , \text{ se } u \geq v. \end{cases}$$

(Sugestão. Primeiro obtenha $P(u \leq U, V \leq v)$. Depois, calcule a derivada da função de distribuição conjunta.)

Exercício 26: Sejam X_1, \dots, X_n variáveis aleatórias e identicamente distribuídas, com distribuição uniforme em $[0, \theta]$, onde $\theta > 0$. Sejam

$$U = \min_{1 \leq i \leq n} X_i, \quad V = \max_{1 \leq i \leq n} X_i.$$

a) Prove que a densidade conjunta de (U, V) é

$$\begin{cases} n(n-1)(v-u)^{n-2}/\theta^n & , \text{ se } 0 \leq u < v \leq \theta \\ 0 & , \text{ se caso contrário.} \end{cases}$$

(Sugestão. Exercício 60 do livro, aqui o exercício 25.)

Exercício 27: Se X_1, \dots, X_n são independentes com distribuição comum $U[0, 1]$, mostre que

$$-2n \log Y \sim \chi^2(2n),$$

onde Y é a média geométrica das X_i , definida por

$$Y = \left(\prod_{i=1}^n X_i \right)^{1/2}.$$

(28-30. George Roussas - Cap 7)

Exercício 28: Seja X_1 e X_2 duas variáveis aleatórias com função densidade de probabilidade f dada por

$$f(x_1, x_2) = e^{-x_1 - \rho x_1 x_2 - x_2} I_{(0, \infty) \times (0, \infty)}(x_1, x_2), \quad \rho \geq 0 \text{ (constante).}$$

Mostre que X_1, X_2 são independentes se e somente se $\rho = 0$.

Exercício 29: Sejam as variáveis aleatórias X_1, X_2, x_3 conjuntamente distribuídas com função de densidade de probabilidade f dada por

$$f(x_1, x_2, x_3) = \frac{1}{4} I_A(x_1, x_2, x_3),$$

com

$$A = \{(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 1)\}.$$

Então mostre que:

- a) $X_i, X_j, i \neq j$ são independentes.
- b) X_1, X_2, X_3 são dependentes.

Exercício 30: Sejam $X_j, j = 1, \dots, n$ variáveis aleatórias i.i.d. com função densidade de probabilidade f e seja \mathcal{F} o conjunto de Borel em \mathbb{R} .

- a) Em termos de f , expresse a probabilidade que pelo menos k dos X 's encontram-se em \mathcal{F} para k algum fixado ($1 \leq k \leq n$).
- b) Simplifique essa expressão para f exponencial negativa com parâmetro λ e $B = (1/\lambda, \infty)$.
- c) Encontre uma resposta numérica para $n = 10, k = 5, \lambda = \frac{1}{2}$.

(31-35. George Roussas - Cap 9)

Exercício 31: ex 3 - cap 9

Exercício 32: Seja X uma variável aleatória contínua e o conjunto $Y = \sum_{j=1}^n c_j I_{B_j}(X)$, onde $B_j, j = 1, \dots, n$ são conjuntos de Borel disjuntos 2 a 2 e $c_j, j = 1, \dots, n$ são contantes.

- a) Expresse a função densidade de probabilidade de Y em termos de X , e note que Y é uma variável aleatória discreta enquanto que X é uma variável aleatória contínua.
- b) Se $n = 3, X \sim N(99, 5)$ e $B_1 = (95, 105), B_2 = (92, 95) + (105, 107), B_3 = (-\infty, 92] + [107, \infty)$, determine a distribuição da variável aleatória Y definida acima.
- c) Se X é interpretado como uma medição especificada tomada sobre cada item de um produto feito por um determinado processo de fabricação e $c_j, j = 1, 2, 3$ como o lucro (em dólares) obtido pela venda de um item sob a condição de $X \in B_j, j = 1, 2, 3$, respectivamente, encontre o lucro esperado de venda de um item.

Exercício 33: Sejam X_1, X_2 variáveis aleatórias independentes $N(0, \sigma^2)$. Encontre a função densidade de probabilidade da variáveis aleatórias $X_1^2 + X_2^2, X_1/X_2$ e mostre que essas variáveis aleatórias são independentes.

Exercício 34: Sejam X_1, X_2 variáveis aleatórias independentes $U(\alpha, \alpha + 1)$. Então:

- a) Encontre as funções densidade de probabilidade das seguintes variáveis aleatórias: $X_1 + X_2, X_1 - X_2, X_1 X_2, X_1^2 + X_2^2$.
- b) Examine as variáveis aleatórias $X_1 + X_2$ e $X_1 - X_2$ do ponto de vista de independencia.

Exercício 35: Sejam X_1, X_2 variáveis aleatórias independentes com função densidade de probabilidade dada por

$$f(x) = \frac{1}{x^2} I_{(1, \infty)}(x).$$

Determine a distribuição da variável aleatória $X = X_1/X_2$.

(36-39. Sheldon Ross - Cap 4)

Exercício 36: Existem N tipos distintos de cupons de desconto, e cada vez que um deles é recolhido, ele tem probabilidade $P_i, i = 1, \dots, N$, de ser de qualquer tipo, independentemente das escolhas anteriores. Suponha que T represente o número de cupons que alguém precise selecionar para obter pelo menos um cupom de cada tipo. Calcule $P(T = n)$.

DICA: Probabilidade da união de eventos.

Exercício 37: Provar ou dar um contra-exemplo:

$$P(X < b) = \lim_{b_n \rightarrow b} P(X < b_n).$$

Exercício 38: Um pote contém $m + n$ pedras, numeradas $1, 2, \dots, n + m$. Um conjunto de tamanho n é retirado do pote. Se X representar o número retirado de pedras cujo número é maior do que cada um dos números das pedras restantes, calcule a função de probabilidade de X .

Exercício 39: Um pote contém n pedras. Suponha que um garoto tire uma pedra do pote sucessivamente, recolocando a pedra retirada antes de tirar a próxima. O processo continua até que o garoto retire pela segunda vez uma pedra já retirada. Suponha que X represente o número de vezes que o menino retira uma pedra, e calcule a sua função de probabilidade.

(40-44. Sheldon Ross - Cap 6)**Exercício 40:**

a) Se X tem distribuição gama com parâmetros (t, λ) , qual é a distribuição de cX , $c > 0$?

b) Mostre que

$$\frac{1}{2\lambda} \chi_{2n}^2$$

tem distribuição gama com parâmetros n, λ quando n é um inteiro positivo e χ_{2n}^2 é uma variável aleatória qui-quadrado com $2n$ graus de liberdade.

Exercício 41: Sejam X_1, \dots, X_n variáveis aleatórias exponenciais independentes com parâmetro comum λ . Determine a distribuição de $\min(X_1, \dots, X_n)$.

Exercício 42: Mostre que a mediana de uma amostra de tamanho $2n + 1$ de uma distribuição uniforme ao longo do intervalo $(0, 1)$ tem distribuição beta com parâmetros $(n + 1, n + 1)$.

Exercício 43: Suponha que X_1, \dots, X_n sejam variáveis aleatórias independentes e identicamente distribuídas com função distribuição F e densidade f . A grandeza $M = [X_{(1)} + X_{(n)}]/2$, definida como a média do menor e do maior valor em X_1, \dots, X_n é chamada de *alcance central* da sequência. Mostre que sua função distribuição é

$$F_M(m) = n \int_{-\infty}^m [F(2m - x) - F(x)]^{n-1} f(x) dx$$

Exercício 44: Sejam X_1, \dots, X_n variáveis aleatórias independentes e uniformes em $(0, 1)$. Faça com que $R = X_{(n)} - X_{(1)}$ represente o alcance e $M = [X_{(1)} + X_{(n)}]/2$ represente o alcance central de X_1, \dots, X_n . Calcule a função densidade conjunta de R e M .

(45-49. Sheldon Ross - Cap 6)

Exercício 45: Numa caixa com 5 transístores sabe-se que 2 estão com defeito. Os transístores são testados, um de cada vez, até que os defeituosos sejam identificados. Denote por N_1 o número de testes feito até o primeiro defeituoso seja encontrado e por N_2 o número de testes adicionais até o segundo defeituoso seja encontrado. Encontre a função de probabilidade conjunta de N_1 e N_2 .

Exercício 46: Quando uma corrente I (em amperes) flui através de uma resistência T (em ohms), a energia gerada é dada por $W = I^2 R$ (em watts). Suponha que I e R são variáveis aleatórias independentes com densidades

$$f_I(x) = 6x(1-x) \quad 0 \leq x \leq 1$$

$$f_R(x) = 2x \quad 0 \leq x \leq 1$$

Determine a função densidade de W .

Exercício 47: Se X_1, X_2, X_3 são variáveis aleatórias independentes $U(a, b)$, calcule a probabilidade da maior das três ser maior do que a soma das outras duas.

Exercício 48: Seja X e Y denota as coordenadas do ponto escolhido aleatoriamente do círculo unitário (centrado na origem). Isso é, com densidade conjunta dada por

$$f(x, y) = \frac{1}{\pi}, \quad x^2 + y^2 \leq 1$$

Encontre a função de densidade conjunta das coordenadas polares $R = (X^2 + Y^2)^{1/2}$ e $\Theta = \tan^{-1} Y/X$.

Exercício 49: Se X e Y possuem densidade conjunta

$$f(x, y) = \frac{1}{x^2 y^2} \quad x \geq 1, y \geq 1$$

a) Calcule a função de densidade conjuta de $U = XY, V = X/Y$.

b) Qual são as densidades marginais?

List 4

(1-14. Barry James - Cap 3)

Exercício 1: Seja X uma variável aleatória tendo *distribuição logística* com densidade

$$f(x) = \frac{e^{-x}}{(1 + e^{-x})^2}, \quad x \in \mathbb{R}.$$

- a) Prove que a distribuição de X é simétrica em torno de zero (veja o exercício 1 do livro).
- b) Determine se X tem esperança finita. Se tiver, ache o valor.
- c) Obtenha a densidade de $Y = e^X$ e ache $E(Y)$.

Exercício 2: Calcule $E(X)$.

$$f(x) = \frac{1}{(1+x)^2}, \text{ se } x > 0; \quad f(x) = 0, \text{ se } x \leq 0.$$

Exercício 3:

- a) Se $X \sim \Gamma(\alpha, \beta)$, onde $\alpha > 0$ e β , qual é $E(X)$?
- b) Dizemos que X tem distribuição de Weibull se possui densidade

$$f(x) = \begin{cases} \lambda\alpha x^{\alpha-1} e^{\lambda x^\alpha} & , \quad x > 0 \\ 0 & , \quad x \leq 0. \end{cases}$$

Calcule $E(X)$ neste caso.

Exercício 4: Sejam X e Y variáveis aleatórias independentes com distribuição comum $U[0, 1]$. Calcule $E(Z)$ e $E(W)$, onde $Z = \min(X, Y)$ e $W = \max(X, Y)$.

Exercício 5: Um jogador vai lançando uma moeda honesta. Ele para depois de lançar ou duas caras sucessivas ou duas coroas sucessivas. Qual a esperança do número de lançamentos?

Exercício 6: Uma urna contém n bolas numeradas $1, 2, \dots, n$. Uma pessoa tira uma bola e a devolve, tira outra e a devolve, continuando até tirar uma bola pela segunda vez. Seja X o número total de retiradas necessárias para obter essa perfeição.

a) Ache a distribuição de X . (Sugestão. Ache $P(X > k)$.)

b) Mostre que

$$E(X) = 2 + \left(1 - \frac{1}{n}\right) + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \cdots + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{n-1}{n}\right).$$

Exercício 7: Dois jogadores lançam moedas simultaneamente até obterem o primeiro casamento (i.e., ou duas caras ou duas coroas). Se os dois lançam “cara” simultaneamente, ganha o jogador I ; se ambos lançam “coroa”, ganha o jogador II . Por exemplo, se os dois obtêm “cara”, no primeiro lançamento, então o jogo termina e o jogador I ganha o jogo. Suponha que a moeda do jogador I seja honesta, mas que a moeda do outro não necessariamente o seja, tendo probabilidade p de “cara”, $0 \leq p \leq 1$.

a) Calcule a esperança do número de lançamentos (i.e. o número de vezes que o jogador I lança a moeda até terminar o jogo).

b) Ache a probabilidade do jogador I ganhar o jogo (mais cedo ou mais tarde).

Exercício 8: Calcule $E(e^X)$, onde X tem densidade logística $f(x) = \frac{e^{-x}}{(1 + e^{-x})^2}$, $x \in \mathbb{R}$, utilizando o Teorema 3.1. Compare com o resultado obtido no exercício 2(c) do livro.

Exercício 9:

a) Prove: se a variável aleatória X é limitada, então tem momentos finitos de toda ordem.

b) Seja A um evento aleatório. Calcule todos os momentos absolutos da variável aleatória $X + I_A$.

c) Demonstre: Se $X \sim N(\mu, \sigma^2)$, então todos os momentos absolutos de X são finitos.

d) Seja $X \sim Cauchy(0, 1)$. Quais são os momentos absolutos finitos de X ?

Exercício 10: Sejam X_1, \dots, X_n variáveis aleatórias independentes com $E(X_j) = \mu_j$ e $Var(X_j) = \sigma_j^2$. Considere combinações lineares $Y = \sum_{j=1}^n p_j X_j$, onde $p_j \geq 0$ e $\sum_{j=1}^n p_j = 1$. Prove que $Var(Y)$ é minimizada pela escolha de

$$p_j = \frac{\sigma_j^{-2}}{\sum_{i=1}^n \sigma_i^{-2}}, \quad j = 1, \dots, n.$$

Exercício 11: Sejam X e Y variáveis aleatórias com variâncias finitas. Mostre que se $Var(X) \neq Var(Y)$, então $X + Y$ e $X - Y$ não são independentes.

Exercício 12: Demostre que a covariância é *bilinear*:

$$Cov\left(\sum_{i=1}^m a_i X_i, \sum_{j=1}^n b_j Y_j\right) = \sum_{j=1}^n \sum_{i=1}^m a_i b_j Cov(X_i, Y_j),$$

onde os a_i e b_j são números reais. (Suponha que as X_i e Y_j possuam variâncias finitas.)

Exercício 13: Seja $(X_1, \dots, X_m, Y_1, \dots, Y_n)$ um vetor aleatório $(m+n)$ -dimensional tal que $Var(X_i) = Var(Y_j) = 1$, $\rho(X_i, X_j) = \rho_1$ e $\rho(Y_i, Y_j) = \rho_2$, $\forall i \neq j$, e $\rho(X_i, Y_j) = \rho_3$, $\forall i, j$. Se $U = X_1 + \dots + X_m$ e $V = Y_1 + \dots + Y_n$, então prove que o coeficiente de correlação entre U e V , $\rho(U, V)$, é igual a

$$\frac{\sqrt{mn\rho_3}}{\sqrt{1 + (m-1)\rho_1}\sqrt{1 + (n-1)\rho_2}}$$

Exercício 14: Sejam X e Y variáveis aleatórias independentes com distribuição comum $U[0, 1]$, e sejam $X_{(1)} = \min(X, Y)$, $X_{(2)} = \max(X, Y)$. Calcule o coeficiente de correlação $\rho(X_{(1)}, X_{(2)})$.

(15-22. Morris DeGroot - Cap 4)

Exercício 15: Suponha que as variáveis X_1, \dots, X_n formam uma amostra aleatória de tamanho n de uma dada distribuição contínua na reta f . Encontrar a expectativa do número de observações na amostra que estão dentro de um intervalo especificado $a \leq x \leq b$.

Exercício 16: Suponha que em cada partida de um determinado jogo seja igualmente provável o jogador ganhar ou perder. Suponha que quando ele ganha, sua fortuna é o dobro e quando ele perde, sua fortuna é dividida ao meio. Se ele começa o jogo com uma dada fortuna c , qual é a fortuna esperada dele depois de n partidas independentes do jogo?

Exercício 17: Suponha que uma classe contém 10 meninos e 15 meninas, e suponha que 8 estudantes são selecionados aleatoriamente da classe sem reposição. Seja X o número de meninos que foram selecionados, e seja Y o número de meninas que foram selecionados. Encontre $E(X|Y)$.

Exercício 18: Suponha que X_1, \dots, X_n é uma amostra aleatória de tamanho n da seguinte distribuição:

$$f(x) = \begin{cases} 2x & , \quad 0 < x < 1, \\ 0 & , \quad \text{caso contrário.} \end{cases}$$

Seja $Y_n = \max(X_1, \dots, X_n)$. Avalie $E(Y_n)$.

Exercício 19: Suponha que você vai vender refrigerante em um jogo de futebol e deve decidir com antecedência quanto levar de refrigerante. Suponha que a demanda por refrigerante no jogo, em litros tem uma distribuição contínua. Suponha que você tem um lucro de g centavos em cada litro que você vender no jogo e que você tem uma perda de centavos de c em cada litro que você não vender. Qual é a quantidade ideal de refrigerante para você maximizar o seu ganho líquido esperado?

Exercício 20: Suponha que X é o número de horas que uma máquina opera antes de falhar, distribuição contínua com função densidade $f(x)$. Suponha que no momento em que a máquina começa a operar, você deve decidir quando você voltará a inspecioná-la. Se você inspecionar a máquina antes de quebra você terá um gasto de b dólares (pela inspeção). Se você retornar a inspecionar depois que a máquina quebrou, você gastará c dólares por cada hora que a máquina ficar sem operar. Qual é o número ótimo de horas que você deve esperar para retornar a fazer a inspeção de forma que minimize seu gasto?

Exercício 21: Suponha que X_0, X_1, \dots, X_n são variáveis aleatórias independentes, cada uma tem a mesma variância σ^2 . Seja $Y_j = X_j - X_{j-1}$ para $j = 1, \dots, n$, e seja $\bar{Y}_n = \frac{1}{n} \sum_{j=1}^n Y_j$. Determine a $Var(\bar{Y}_n)$.

Exercício 22: Suponha que X_1, \dots, X_n são variáveis aleatórias e que $Var(X_i) = \sigma^2$ para $i = 1, \dots, n$ e $\rho(X_i, X_j) = \rho$ para $i \neq j$. Prove que $\rho \geq -\frac{1}{1-n}$.

(23-35. Sheldon Ross - Cap 7)

Exercício 23: Mostre que a esperança $E[(X - a)^2]$ é minimizada quando $a = E[X]$.

Exercício 24: Considere n tentativas independentes, cada uma levando a um dos r resultados possíveis com probabilidades P_1, P_2, \dots, P_r . Suponha que X represente o número de resultados que nunca ocorrem em qualquer uma da tentativas. Determine $E[X]$ e mostre que, entre todos os vetores probabilidade $P_1, P_2, \dots, P_r, E[X]$ é minimizada quando $P_r = 1/r, i = 1, \dots, r$.

Exercício 25: Mostre que, se X e Y são identicamente distribuídas e não necessariamente independentes, então

$$Cov(X, Y) = 0.$$

Exercício 26: Suponha que $X_{(i)}, i = 1, \dots, n$, represente as estatísticas de ordem de um conjunto de n variáveis aleatórias uniformes no intervalo $(0, 1)$, e observe que a função densidade de $X_{(i)}$ é dada por

$$f(x) = \frac{n!}{(i-1)!(n-i)!} x^{i-1} (1-x)^{n-i}, \quad 0 < x < 1.$$

- a) Calcule $Var(X_{(i)}), i = 1, \dots, n$.
- b) Que valores de i minimizam e maximizam $Var(X_{(i)})$, respectivamente?

Exercício 27: Mostre que, se Z é uma variável aleatória normal padrão e Y é definida como $Y = a + bZ + cZ^2$, então

$$\rho(Y, Z) = \frac{b}{\sqrt{b^2 + 2c^2}}.$$

Exercício 28: Demonstre a desigualdade de Cauchy-Schwartz:

$$(E[XY])^2 \leq E[x^2] E[Y^2].$$

Dica: A menos que $Y = -tX$ para alguma constante, caso no qual a desigualdade se torna uma igualdade, tem-se que, para todo t ,

$$0 < E[(tX + Y)^2] = E[X^2]t^2 + 2E[XY]t + E[Y^2]$$

Portanto, as raízes da equação quadrática devem ser imaginárias, o que implica o discriminante desta equação ser negativo.

Exercício 29: Se Z é uma variável aleatória normal padrão, calcule $Cov(Z, Z^2)$?

Exercício 30: Seja X_1, X_2, \dots uma sequência de variáveis aleatórias independentes e indenidicamente distribuídas contínua. Seja $N \geq 2$ tal que

$$X_1 \geq X_2 \geq \dots \geq X_{N-1} < X_N$$

Isso é, N é o ponto em que a sequência para de decrescer. Mostre que $E[N] = e$. Dica: Primeiro encontre $P\{N \geq n\}$.

Exercício 31: Seja X ser o número de 1's e Y ser o número de 2's que ocorre em n lançamentos de um dado honesto. Calcule $Cov(X, Y)$.

Exercício 32: Sejam X_1, \dots, X_n independentes com média μ e variância σ^2 comum, e $Y_n = X_n + X_{n+1} + X_{n+2}$. Encontre para $j \geq 0$, $Cov(Y_n, Y_{n+j})$.

Exercício 33: A densidade conjunta de X e Y é dada por

$$f(x, y) = \frac{1}{y} e^{-(y+x/y)}, x > 0, y > 0$$

Encontre $E[X]$, $E[Y]$, e mostre que $Cov(X, Y) = 1$.

Exercício 34: Um grupo de 20 pessoas (10 homens e 10 mulheres) são alocados aleatoriamente em 10 pares. Calcule a média e variância do número de pares que é formado por um homem e uma mulher. Agora suponha que as 20 pessoas são 10 casais. Calcule a média e variância do número de casais.

Exercício 35: Sejam X_1, X_2, \dots, X_n variáveis aleatórias independentes com função de distribuição contínua F desconhecida, e sejam Y_1, Y_2, \dots, Y_m variáveis aleatórias independentes com função de distribuição contínua G desconhecida. Agora ordene essas $n + m$ variáveis e seja

$$I_i = \begin{cases} 1 & \text{se a } i\text{-ésima menor das } n+m \text{ variáveis da amostra é de } X \\ 0 & \text{caso contrário.} \end{cases}$$

A variável aleatória $R = \sum_{i=1}^{n+m} i I_i$, é a soma dos postos da amostra X e é a base de um procedimento estatístico padrão (Chamado teste de Wilcoxon) para testar se F e G são indenticamente distribuído. Esse teste aceita a hipótese que $F = G$ quando R não é nem grande e nem pequeno. Assumindo que a hipótese da igualdade é de fato correta, calcule a média e variância de R .

List 5

(1-13. Barry James - Cap 4)

Exercício 1: Sejam X_1 e X_2 variáveis aleatórias independentes, cada uma com distribuição geométrica definida por

$$P(X_i = n) = p(1-p)^n, \quad n = 0, 1, 2, \dots; \quad i = 1, 2;$$

onde $0 < p < 1$. (Observação: esta versão da distribuição geométrica corresponde à distribuição do número de fracassos antes do primeiro sucesso em uma sequência de ensaios de Bernoulli.)

- a) Calcule $P(X_1 = X_2)$ e $P(X_1 < X_2)$.
- b) Determine a distribuição condicional de X_1 dada $X_1 + X_2$.

Exercício 2: Uma certa lâmpada tem uma vida, em *horas*, tendo distribuição exponencial de parâmetro 1. Um jogador acende a lâmpada e, enquanto a lâmpada ainda estiver acesa, lança um dado equilibrado de quze em quize *segundos*. Qual o número esperado de 3's lançamentos pelo jogador até a lâmpada se apagar?

Exercício 3: Suponha que o número de passa num bolo inglês tenha distribuição de Poisson de parâmetro 60. Um jogador compra um bolo, tira todas as passas uma por uma e reparte as passas entre ele e você da seguinte maneira: depois da extração de cada passa ele joga uma moeda equilibrada, dando a passa para você se der cara, comendo ele mesmo a passa se der coroa. Qual a distribuição de número de passa que você recebe? A esperança?

Exercício 4: Observam-se duas lâmpadas durante suas vidas úteis. Suponha as vidas independentes e exponenciais de parâmetros λ . Sejam X o tempo de quima da primeira lâmpada a queimar e Y o tempo de queima da segunda a queimar ($X \leq Y$).

- a) Qual a distribuição condicional de X dado Y ?
- b) Qual a distribuição de Y dada X ?

Exercício 5: Sejam X e Y independentes, cada uma com distribuição $N(0, \sigma^2)$. Qual a distribuição condicional de X dada $\sqrt{X^2 + Y^2}$? (Sugestão. Exemplo 10 e o princípio da substituição.) Determine $E(X|\sqrt{X^2 + Y^2})$.

Exercício 6: Sejam X_1, \dots, X_n variáveis aleatórias independentes e identicamente distribuídas, com distribuição contínua F . Seja $X = \max_{1 \leq i \leq n} X_i$.

a) Mostre que para todo $k = 1, 2, \dots, n$,

$$P(X_k \leq x | X = t) = \begin{cases} \frac{(n-1)F(x)}{nF(t)} & , \quad x < t \\ 1 & , \quad x \geq t. \end{cases}$$

(Sugestão. $x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + \dots + xy^{n-2} + y^{n-1})$.)

b) Suponha que F diferenciável. Existe densidade condicional no item (a)?

Exercício 7: Seja X uma variável aleatória Cauchy-padrão.

a) Conforme a Definição 4.3 (do livro), qual é $E(X| |X| = y)$?

b) Deduza que $E[E(X| |X|)] = 0$.

c) A propriedade básica $E[E(X|Y)] = E[X]$ não vale nesse caso. Existe uma contradição nisso?

Exercício 8: Calcule $E(X_{(2)}|X_{(1)}, X_{(3)})$, onde $X_{(1)}, X_{(2)}, X_{(3)}$ são as estatísticas de ordem de uma amostra aleatória da $U[0, 1]$. (Veja o exercício 24(b) do livro.)

Exercício 9: Seja X uma variável aleatória com distribuição exponencial de parâmetro 1. Seja $t > 0$ fixo. Ache $E(X| \max(X, t))$ e $E(X| \min(X, t))$.

Exercício 10: Mostre que se X e Y são variáveis aleatórias, então

$$F_{X,Y}(x, y) \leq \sqrt{F_X(x)F_Y(y)}.$$

(Sugestão. Use indicadores e Cauchy-Schwarz.)

Exercício 11: Seja (X, Y) um vetor aleatório bidimensional. Suponha que (i) X tem distribuição exponencial com parâmetro $1/2$ e (ii) para cada $x > 0$, a distribuição condicional de Y dado que $X = x$ é $U[0, x^2]$. Em outras palavras, $X \sim \exp\left(\frac{1}{2}\right)$ e $Y|X=x \sim U[0, x^2]$.

- a) Qual é a distribuição de $Z = Y/X^2$?
- b) Calcule $E(X)$, $E(Y)$ e $Cov(X, Y)$.

Exercício 12: Suponha que em um temporal, o número X de gotas de chuva que caem no IMPA durante um segundo tenha distribuição de Poisson com parâmetro $\lambda > 0$, onde λ representa a intensidade de chuva. Suponha que o parâmetro λ seja uma variável aleatória que tenha distribuição gama com parâmetros $\alpha > 0$ e $\beta = 1$, i.e., que sua densidade seja dada por

$$f(\lambda) = \begin{cases} \frac{1}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\lambda} & , \quad \lambda \geq 0 \\ 0 & , \quad \lambda < 0. \end{cases}$$

- a) Mostr que

$$P(X = k) = \frac{\Gamma(k + \alpha)}{\Gamma(\alpha)\Gamma(k + 1)} \left(\frac{1}{2}\right)^{k+\alpha}, \quad k = 0, 1, 2, \dots$$

- b) Usando métodos probabilísticos, demonstre que

$$\sum_{k=1}^{\infty} C_n^{k+n-1} \frac{1}{2^k} = 2^n, \quad \text{para } n = 1, 2, \dots$$

(Sugestão. Calcule a esperança de X de duas maneiras diferentes.)

Exercício 13: Seleciona-se ao acaso (i.e., conforme a distribuição uniforme) um número entre 0 e 1. Se x é o número selecionado, lança-se n vezes (independentemente) uma moeda com probabilidade x de dar “cara”. Seja Y a variável aleatória que representa o número de caras lançadas.

- a) Y é variável aleatória discreta ou contínua?
- b) Calcule a esperança e avariância de Y .

(14-16. George Roussas - Cap 4)

Exercício 14: Se a f.d.p. conjunta f das v.a.'s $X_j, j = 1, 2, 3$ é dada por

$$f(x_1, x_2, x_3) = c^3 e^{-c(x_1+x_2+x_3)} I_A(x_1, x_2, x_3),$$

onde

$$A = (0, \infty) \times (0, \infty) \times (0, \infty),$$

- a)** Determine a constante c .
- b)** Encontre a marginal de cada uma das v.a.'s $X_j, j = 1, 2, 3$.
- c)** Encontre a condicional de $X_1, X_2 | X_3$, e a condicional de $X_1 | X_2, X_3$.
- d)** Encontre as funções de distribuição das condicionais de **c**).

Exercício 15: Seja Y uma v.a distribuída como $P(\lambda)$ e suponha que a distribuição condicional da v.a. $X|Y = n$ é uma $B(n, p)$. Determine a f.d.p. de X e f.d.p condicional $Y|X = x$.

Exercício 16: Considere a função f definida a seguir

$$f(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right) + \frac{1}{4\pi e} x_1^3 x_2^3 I_{[-1,1] \times [-1,1]}(x_1, x_2)$$

e mostre que:

- a)** f é uma f.d.p. non-Normal bivariada.
- b)** Ambas marginais

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2$$

e

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1$$

são f.d.p.'s Normal.

(17-19. Morris DeGroot - Cap 4.7)

Exercício 17: Suponha que um ponto X_1 é escolhido de uma distribuição uniforme no intervalo $(0, 1)$; e depois que o valor de $X_1 = x_1$ é observado, um ponto X_2 é escolhido de uma distribuição uniforme $(x_1, 1)$. Suponha mais adiante que outras adicionais variáveis X_3, X_4, \dots são geradas da mesma forma. Em geral, para $j = 1, 2, \dots$, depois do valor $X_j = x_j$ ter sido observado, X_{j+1} é escolhido de uma distribuição uniforme no intervalo $(x_j, 1)$. Encontre o valor da $E(X_n)$.

Exercício 18: Suponha que X e Y tem distribuição conjunta contínua dada por:

$$f(x, y) = \begin{cases} x + y & , \quad 0 \leq x, y \leq 1 \\ 0 & , \quad \text{caso contrário.} \end{cases}$$

Encontre $E(Y|X)$ e $Var(Y|X)$.

Exercício 19: Suponha que X e Y são variáveis aleatórias tal que $E(Y|X) = aX + b$. Assumino que $Cov(X, Y)$ existe e que $0 < Var(X) < \infty$, determine a expressão de a e b em termos de $E(X)$, $E(Y)$, $Var(X)$, e $Cov(X, Y)$.

(20-25. Morris DeGroot - Cap 4.10)

Exercício 20: Suponha que X é o número de horas que uma máquina opera antes de falhar, distribuição contínua com função densidade $f(x)$. Suponha que no momento em que a máquina começa a operar, você deve decidir quando você voltará a inspecioná-la. Se você inspecionar a máquina antes de quebra você terá um gasto de b dólares (pela inspeção). Se você retornar a inspecionar depois que a máquina quebrou, você gastará c dólares por cada hora que a máquina ficar sem operar. Qual é o número ótimo de horas que você deve esperar para retornar a fazer a inspeção de forma que minimize seu gasto? (EXERCÍCIO REPETIDO DA LISTA ANTERIOR, 20)

Exercício 21: Suponha que X_0, X_1, \dots, X_n são variáveis aleatórias independentes, cada uma tem a mesma variância σ^2 . Seja $Y_j = X_j - X_{j-1}$ para $j = 1, 2, \dots, n$, e seja $\bar{Y} = \frac{1}{n} \sum_{j=1}^n Y_j$. Determine a $Var(\bar{Y}_n)$. (EXERCÍCIO REPETIDO DA LISTA ANTERIOR, 21)

Exercício 22: Suponha que X_1, \dots, X_n são variáveis aleatórias com $Var(X_i) = \sigma^2$ para $i = 1, \dots, n$ e $\rho(X_i, X_j) = \rho$ para todo $i \neq j$. Prove que $\rho \geq -\frac{1}{n-1}$. (EXERCÍCIO REPETIDO DA LISTA ANTERIOR, 22)

Exercício 23: Suponha que X e Y são variáveis aleatórias tal que $E(Y|X) = 7 - (1/4)X$ e $E(X|Y) = 10 - Y$. Determine a correlação entre X e Y .

Exercício 24: Suponha que um bastão tem comprimento de 3 pés e é quebrado em 2 pedaços, e que esse ponto é escolhido de uma f.d.p $f(x)$. Qual é a correlação entre o comprimento do maior pedaço e o comprimento do menor pedaço?

Exercício 25: Suponha que X e Y tem distribuição conjunta com correlação $\rho > 1/2$ e que $Var(X) = Var(Y) = 1$. Mostre que $b = -\frac{1}{2\rho}$ é o único valor de b tal que a correlação de X e $X + bY$ é também ρ .

Lista 6

(1-12. Barry James - Cap 5)

Exercício 1: Sejam X_1, X_2, \dots variáveis independentes com distribuição comum Poisson (λ). Qual o limite em probabilidade da sequência $(Y_n)_{n \geq 1}$, onde $Y_n = \frac{X_1^2 + \dots + X_n^2}{n}$?

Exercício 2: Seja $(X_n)_{n \geq 1}$ uma sequência de variáveis aleatórias. Prove que se $E(X_n) \rightarrow \alpha$ e $Var(X_n) \rightarrow 0$, então $X_n \xrightarrow{p} \alpha$.

Exercício 3: Sejam X_1, X_2, \dots variáveis aleatórias independentes tais que $X_1 = 0$ e para $j \geq 2$, X_j é variável aleatória discreta satisfazendo

$$P(X_j = k) = \begin{cases} \frac{1}{j^3} & , \quad k = \pm 1, \pm 2, \dots, \pm j \\ 1 - \frac{2}{j^2} & , \quad k = 0. \end{cases}$$

Prove que

$$\frac{\sum_{j=1}^n X_j}{n^\alpha} \xrightarrow{p} 0$$

quando $n \rightarrow \infty$, se $\alpha > \frac{1}{2}$. (Dado: $\sum_{k=1}^j k^2 = \frac{j(j+1)(j+2)}{6}$.)

Exercício 4: Sejam X_1, X_2, \dots variáveis aleatórias independentes tais que X_n tem distribuição $U[0, a_n]$ onde $a_n > 0$. Mostre:

a) Se $a_n = n^2$, então com probabilidade 1, somente um número finito das X_n toma valores menores que 1.

b) Se $a_n = n$, então com probabilidade 1, um número infinito das X_n toma valores menores que 1.

Exercício 5: Sejam X_1, X_2, \dots variáveis aleatórias independentes tais que $P(X_n = 1) = 1/n$, $P(X_n = 0) = 1 - 1/n$. Mostre que $X_n \xrightarrow{P} 0$ mas $P(X_n \rightarrow 0) = 0$.

Exercício 6: Sejam X_1, X_2, \dots variáveis aleatórias.

a) Demonstre: Se $\sum_{n=1}^{\infty} P(|X_n| > n) < \infty$, então $\lim_{n \rightarrow \infty} \frac{|X_n|}{n} < 1$ quase certamente.

b) Se as X_n são identicamente distribuídas e integráveis, demonstre que

$$\lim_{n \rightarrow \infty} \sup \frac{|X_n|}{n} \leq 1 \quad \text{quase certamente.}$$

Exercício 7: Sejam X_1, X_2, \dots variáveis aleatórias independentes e identicamente distribuídas tais que $X_1 \sim U[0, 1]$. Prove que $n^{-X_n} \rightarrow 0$ em probabilidade, mas n^{-X_n} não converge quase certamente para 0. (Sugestão para a parte quase certa: Prove que $P(n^{-X_n} \rightarrow 0) = 0$.)

Exercício 8: Sejam X_1, X_2, \dots independentes e identicamente distribuídas, com $X_1 \sim U[0, 1]$. Ache o limite quase certo da média geométrica $\left(\prod_{k=1}^n X_k \right)^{1/n}$. (Sugestão. Tome logaritmos.)

Exercício 9: Demonstre: se X_1, X_2, \dots são independentes e identicamente distribuídas, com $E[X_1] = 1 = Var[X_1]$, então

$$\frac{\sum_{i=1}^n X_i}{\sqrt{n \sum_{i=1}^n X_i^2}} \xrightarrow{D} \frac{1}{\sqrt{2}} \text{ quase certamente.}$$

Exercício 10: Sejam X_1, X_2, \dots variáveis aleatórias independentes com densidade comum

$$f(x) = \begin{cases} e^{-(x+1/2)} & , \quad x \geq -1/2 \\ 0 & , \quad x < -1/2. \end{cases}$$

Demonstre que $S_n \rightarrow +\infty$ quase certamente, onde $S_n = X_1 + \dots + X_n$.

Exercício 11: Sejam X_1, X_2, \dots independentes com distribuição comum $N(0, 1)$. Qual o limite quase certo de

$$\frac{X_1^2 + \dots + X_n^2}{(X_1 - 1)^2 + \dots + (X_n - 1)^2} ?$$

Exercício 12: Sejam X_1, X_2, \dots variáveis aleatórias independentes tais que $X_n \sim [0, n]$, $n = 1, 2, \dots$. Chame o n -ésimo ensaio de *sucesso* se $X_{2n} > X_{2n-1}$, *fracasso* se $X_{2n} \leq X_{2n-1}$, para $n = 1, 2, \dots$. Determine a probabilidade de haver sucesso no n -ésimo ensaio e ache o limite (se existir) de S_n/n , onde S_n = número de sucessos nos primeiros n ensaios. Esse limite é limite em probabilidade e/ou quase certo?

(13-17. Barry James - Cap 6)

Exercício 13:

a) Suponha que $X_n \xrightarrow{D} N(0, 1)$, $Y_n \xrightarrow{D} N(0, 1)$ e, para todo n , X_n seja independente de Y_n . Mostre que $X_n + Y_n \xrightarrow{D} N(0, 2)$.

b) Generalize (a), probando que se $X_n \xrightarrow{D} F$ e $Y_n \xrightarrow{D} G$, onde X_n e Y_n são independentes $\forall n$, então $X_n + Y_n \xrightarrow{D} F * G$. Aqui, F e G são funções de distribuição e $F * G$ é sua convolução.

Exercício 14: Mostre que é possível para uma sequência de funções de distribuição convergir em todo ponto sem o limite ser uma função de distribuição. (Sugestão. Considere as variáveis aleatórias constantes $X_n = n$.)

Exercício 15: Prove: Se $F_n \rightarrow F$ fracamente e F é contínua, então $F_n(x)$ converge para $F(x)$ uniformemente na reta.

Exercício 16: Utilize funções características para provar: Se $X_n \xrightarrow{D} N(0, 1)$ e $(a_n)_{n \geq 1}$ é uma sequência de números reais tal que $a_n \rightarrow a$ finito, então $X + a_n \xrightarrow{D} N(a, 1)$.

Exercício 17: Sejam X_1, X_2, \dots independentes e identicamente distribuídas, com $X_n \sim U[0, 1]$, e sejam $Y_n = \min(X_1, \dots, X_n)$, $Z_n = \max(X_1, \dots, X_n)$, $U_n = nY_n$, $V_n = n(1 - Z_n)$. Mostre que, quando $n \rightarrow \infty$:

- a) $Y_n \xrightarrow{P} 0$ e $Z_n \xrightarrow{P} 1$.
- b) $U_n \xrightarrow{D} W$ e $V_n \xrightarrow{D} W$, onde W tem distribuição exponencial de parâmetro 1.

(18-25. Marcos Magalhães - Cap 6)

Exercício 18: Sejam $\Omega = [0, 1]$, $\mathcal{B}([0, 1])$ a correspondente σ -álegra de Borel e \mathcal{P} a função de probabilidade que atribui, a cada intervalo, seu comprimento. Defina as variáveis $X_n, n \geq 1$, tais que, com $j = 0, 1, \dots, 2m$ e $m = 1, 2, \dots$,

$$X_{m^2+j}(\omega) = \begin{cases} m & , \quad \frac{j}{2m+1} \leq \omega \leq \frac{j+1}{2m+1} \\ 0 & , \quad \text{caso contrário.} \end{cases}$$

Mostre que:

- a) $X_n \xrightarrow{P} 0$.
- b) $X_n \not\rightarrow 0$, para qualquer $\omega \in \Omega$.
- c) $X_{n^2} \rightarrow 0, \forall \omega \in \Omega$.
- d) $E[X_n(\omega)] \not\rightarrow 0$.

Exercício 19: Se $X_n \xrightarrow{d} X$, então $aX_n + b \xrightarrow{d} aX + b$, para a e b constantes reais.

Exercício 20: As variáveis X_n , $n \geq 1$, são independentes e todas têm distribuição Exponencial de parâmetro λ . Mostre que a sequência $\{X_n^2 : n \geq 1\}$ satisfaz a Lei Forte dos Grande Números.

Exercício 21: Sejam X_1, X_2, \dots, X_n variáveis aleatórias i.i.d. com densidade Uniforme Contínua em $[-1, 1]$. Defoma a variável $Y_n = \min(|X_1|, |X_2|, \dots, |X_n|)$. Mostre que $Y_n \rightarrow 0$ em probabilidade.

Exercício 22: Suponha que X_n , para $n \geq 1$, tem a seguinte função de distribuição:

$$F_n(x) = \begin{cases} 0 & , \quad x \leq 0; \\ x^n & , \quad 0 < x \leq 1; \\ 1 & , \quad x > 1. \end{cases}$$

Determine o limite em distribuição para a sequência $\{X_n : n \geq 1\}$.

Exercício 23: Sejam X_n , $n \geq 1$, variáveis independentes tais que $X_n \sim Bernoulli(p)$. Considere, para $n \geq 1$, $Y_n = \left(\prod_{k=1}^n X_k\right)^{1/n}$.

a) Calcule $P(Y_n > 0)$ e $E(Y_n)$.

b) Verifique se Y_n converge em probabilidade.

Exercício 24: Considere uma função $f : \mathbb{R} \rightarrow \mathbb{R}$, tal que

$$f(t) = \begin{cases} t - 1 & , \quad t < 0; \\ t + 1 & , \quad t \geq 0. \end{cases}$$

Sejam $X_n, n \geq 1$, variáveis aleatórias com $P(X_n = -1/n) = 1$. Seja X igual a 0 com probabilidade 1. Mostre que $X_n \xrightarrow{P} X$, mas $f(X_n)$ não converge em probabilidade para $f(x)$.

Exercício 25: Seja X uma variável aleatória finita e defina X_n como o trucamento de X da seguinte forma: $X_n = XI_{\{|X| \leq n\}}$. Mostre que $Z_n \xrightarrow{qc} X$.

List 1 - Inferência

(1-5. Morris DeGroot - Cap 6.2)

Exercício 1: Suponha que o número de defeitos em um rolo de fita de gravação magnética tem uma distribuição Poisson com média λ 1.0 ou 1.5, e que a priori para λ é a seguinte:

$$\xi(1.0) = 0.4 \text{ e } \xi(1.5) = 0.6.$$

Se um rolo de fita é selecionado aleatoriamente e é encontrado três defeitos, qual é a posteriori para λ ?

Exercício 2: Suponha que a proporção θ de itens defeituosos em um lote fabricado é desconhecido, e que a distribuição a priori para θ é uniformemente distribuída no intervalo $(0, 1)$. Oito itens são selecionados do lote dos quais três são defeituosos. Determine a distribuição a posteriori de θ .

Exercício 3: Suponha que X_1, \dots, X_n é uma amostra aleatória de $f(x|\theta)$, com θ desconhecido, e que a priori de θ é $\xi(\theta)$. Mostre que a posterior $\xi(\theta|x)$ é a mesma independentemente se é calculada usando $\xi(\theta|x) \propto f(x|\theta)\xi(\theta)$ ou se é calculada sequencialmente usando $\xi(\theta|x_1) \propto f(x_1|\theta)\xi(\theta)$, $\xi(\theta|x_1, x_2) \propto f(x_2|\theta)\xi(\theta|x_1)$, \dots , $\xi(\theta|x) \propto f(x_n|\theta)\xi(\theta|x_1, \dots, x_{n-1})$.

Exercício 4: Considere novamente o problema descrito no exercício 2, e assuma a mesma distribuição a priori para θ . Suponha agora, no entanto, que em vez de selecionar amostras aleatórias de oito itens do lote, realizamos o seguinte experimento: Os itens do lote são selecionados aleatoriamente um a um até que exatamente três defeitos tenham sido encontrados. Se acharmos que devemos selecionar um total de oito itens nesta experiência, qual é a distribuição posterior de θ no final do experimento?

Exercício 5: Suponha que X é observado de uma distribuição $Uniforme(\theta - 1/2, \theta + 1/2)$ e que $\theta \sim Uniforme(10, 20)$. Seis observações são selecionadas $(11.0, 11.5, 11.7, 11.1, 11.4, 10.9)$. Determine a distribuição a posteriori de θ .

(6-12. Morris DeGroot - Cap 6.3)

Exercício 6: Seja θ o número médio de defeitos por 100 pés de um certo tipo de fita magnética. Suponha que o valor de θ é desconhecido, e que a distribuição a priori para θ é $Gama(\alpha = 2, \beta = 10)$. Quando um rolo de 1200 pés desta fita é inspecionado, exatamente quatro defeitos são encontrados. Determine a distribuição a posteriori de θ .

Exercício 7: Suponha que a altura de um indivíduo de uma certa população tem distribuição $N(\theta, 4)$. Suponha também que $\theta \sim N(68, 1)$. 10 pessoas são selecionadas e a média da amostra foi 69.5 polegadas.

- a) Qual intervalo de uma polega que tem a maior probabilidade de conter o valor de θ a priori?
- b) Qual intervalo de uma polega que tem a maior probabilidade de conter o valor de θ a posteriori?
- c) Encontre as probabilidades de a) e b).

Exercício 8: Suponha que uma amostra aleatória é retirada de uma distribuição $N(\theta, 4)$, e que a distribuição a priori é normal com desvio padrão igual a 1. Qual é o menor número de observações na amostra para o desvio padrão da distribuição a posteriori de θ ser 0.01?

Exercício 9: Mostre que a família de distribuição beta para priori é conjugada com a família de distribuição binomial negativa para a amostra com parâmetro r conhecido e com parâmetro p desconhecido ($0 < p < 1$).

Exercício 10: Seja $\xi(\theta)$ uma f.d.p. para $\alpha, \beta > 0$ definida por:

$$\xi(\theta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} e^{-\beta/\theta} & , \quad \theta > 0, \\ 0 & , \quad \theta \leq 0. \end{cases}$$

- a) Verifique que $\xi(\theta)$ é de fato uma densidade.
- b) Mostre que a família de ξ para a priori é conjugada com a família de distribuição $N(\mu, \theta)$ para a amostra com μ conhecido e θ desconhecido.

Exercício 11: Mostre que a família de distribuição Pareto para a priori é conjugada com a família a família de distribuição $Uniforme(0, \theta)$ para a amostra com θ desconhecido.

Exercício 12: Suponha que X_1, \dots, X_n é uma amostra aleatória da seguinte distribuição:

$$f(x|\theta) = \begin{cases} \theta x^{\theta-1} & , \quad 0 < x < 1, \\ 0 & , \quad c. c. \end{cases}$$

Suponha também que θ é desconhecido e que a distribuição para θ é $gama(\alpha, \beta)$. Determine a média e a variância da posterior.

(13-17. Morris DeGroot - Cap 6.7)

Exercício 13: Seja X_1, \dots, X_n uma amostra aleatória da distribuição binomial negativa com parâmetros r conhecido e p desconhecido, mostre que $T = \sum_{i=1}^n X_i$ é uma estatística suficiente para p .

Exercício 14: Seja X_1, \dots, X_n uma amostra aleatória da distribuição beta com parâmetros α desconhecido e β conhecido, mostre que $T = \prod_{i=1}^n X_i$ é uma estatística suficiente para α .

Exercício 15: Seja X_1, \dots, X_n uma amostra aleatória da distribuição uniforme discreta nos inteiros $1, 2, \dots, \theta$, em que θ é desconhecido ($\theta = 1, 2, \dots$), mostre que $T = \max\{X_1, \dots, X_n\}$ é uma estatística suficiente para θ .

Exercício 16: Considere uma distribuição com f.p. ou f.d.p. $f(x|\theta)$, em que $\theta \in \Theta$. Diz-se que a família de distribuição obtida por diferentes $\theta \in \Theta$ é uma família exponencial, ou uma família Koopman-Darmois, se $f(x|\theta)$ pode ser escrita da seguinte forma:

$$f(x|\theta) = a(\theta)b(x)\exp[c(\theta)d(x)] \quad \forall x.$$

$a(\theta)$ e $c(\theta)$ são funções arbitrárias de θ , $b(x)$ e $d(x)$ são funções arbitrárias de x . Assumino que as variáveis aleatórias X_1, \dots, X_n é uma amostra aleatória de uma distribuição pertencente a família exponencial, mostre que $T = \sum_{i=1}^n d(X_i)$ é uma estatística suficiente para θ .

Exercício 17: Mostre que cada uma das seguintes distribuições pertencem a família exponencial.

- a)** A família de distribuição Poisson com média desconhecida.
- b)** A família de distribuição Normal com variância desconhecida e média conhecida.
- c)** A família de distribuição beta com α conhecido e com β desconhecido.

(18-22. Morris DeGroot - Cap 6.8)

Exercício 18: Uma distribuição beta com α e β desconhecidos. Ache $T_1 = \prod_{i=1}^n X_i$ e $T_2 = \prod_{i=1}^n (1 - X_i)$.

Exercício 19: Uma distribuição Pareto com x_0 e α desconhecidos. Ache $T_1 = \min(X_1, \dots, X_n)$ e $T_2 = \prod_{i=1}^n X_i$.

Exercício 20: Suponha que os vetores $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ forma uma amostra aleatória bi-dimensional da normal bivariada com médias, variâncias e correlação desconhecidas. Mostre que as seguintes estatísticas são conjutamente suficientes:

$$\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i, \sum_{i=1}^n X_i^2, \sum_{i=1}^n Y_i^2 \text{ e } \sum_{i=1}^n X_i Y_i.$$

Exercício 21: Considere uma distribuição com p.d.p. ou f.p. $f(x|\theta)$, em que o parâmetro θ é um vetor k-dimensional e desconhecido no espaço paramétrico Θ . Diz-se que a família de distribuição indexada por $\theta \in \Theta$ é uma família exponencial k-paramétrica, ou família Koopman-Darmols k-paramétrica se $f(x|\theta)$ pode ser escrita da seguinte forma:

$$f(x|\theta) = a(\theta)b(x) \exp \left[\sum_{i=1}^n c_i(\theta)d_i(x) \right].$$

Aqui a e c_1, \dots, c_k são funções arbitrárias de θ , e b e d_1, \dots, d_k são funções arbitrárias de x . Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição pertencente a família exponencial k-paramétrica, e defina k estatísticas T_1, \dots, T_k com a seguir:

$$T_i = \sum_{j=1}^k d_i(X_j) \text{ para } i = 1, \dots, k.$$

Mostre que as estatísticas T_1, \dots, T_k são conjuntamente suficientes para θ .

Exercício 22: Mostre que cada uma das famílias de distribuição a seguir pertence a família exponencial bi-paramétrica.

- a) A família de todas distribuições normais com média e variância desconhecidas.
- b) A família de todas distribuições gama com α e β desconhecidos.
- c) A família de todas distribuições beta com α e β desconhecidos.

(23-27. George Roussas - Cap 11.1)

Exercício 23: (Poisson truncada) Seja X_1, X_2 variáveis aleatórias i.i.d. com p.d.p. $f(\cdot|\theta)$ dado por:

$$f(0|\theta) = e^{-\theta}, f(1|\theta) = \theta e^{-\theta}, f(2|\theta) = 1 - \theta e^{-\theta} - e^{-\theta}, f(x|\theta) = 0, \quad x \neq 1, 2, 3,$$

quando $\theta > 0$. Mostre que $X_1 + X_2$ não é suficiente para θ .

Exercício 24: Sejam X_1, \dots, X_n variáveis aleatórias i.i.d. com p.d.p. exponencial dupla com $f(\cdot|\theta)$ dado no exercício 9(iii) do capítulo 3 do livro. Mostre que $\sum_{j=1}^n |X_j|$ é uma estatística suficiente para θ .

Exercício 25: Se $X_j = (X_{1j}, X_{2j})'$, $j = 1, \dots, n$ é uma amostra aleatória de tamanho n da normal bivariada, use o teorema (se T é suficiente e g é uma função um-a-um, então $g(T)$ é suficiente) para mostrar que

$$\left(\bar{X}_1, \bar{X}_2, \sum_{j=1}^n X_{1j}^2, \sum_{j=1}^n X_{2j}^2, \sum_{j=1}^n X_{1j} X_{2j} \right)'$$

é uma estatística suficiente para θ .

Exercício 26: Se X_1, \dots, X_n é uma amostra aleatória de tamanho n de uma $U(-\theta, \theta)$, $\theta \in (0, \infty)$, mostre que $(X_{(1)}, X_{(n)})$ é uma estatística suficiente para θ . Mostre que a estatística $T = \max(|X_1|, \dots, |X_n|)$ é suficiente porém não é minimal.

Exercício 27: Se X_1, \dots, X_n é uma amostral aleatória de tamanho n da $N(\theta, \theta^2)$, $\theta \in R$. Mostre que

$$\left(\sum_{j=1}^n X_j, \sum_{j=1}^n X_j^2 \right)' \text{ ou } \left(\bar{X}, \sum_{j=1}^n X_j^2 \right)'$$

é uma estatística suficiente para θ .

(28-33. Casella e Berger - Cap 6)

Exercício 28: Seja X uma observação da $N(0, \sigma^2)$. $|X|$ é uma estatística suficiente?

Exercício 29: Sejam X_1, \dots, X_n variáveis aleatórias com densidade

$$f_{X_i}(x|\theta) = \begin{cases} e^{i\theta-x}, & x \geq i\theta, \\ 0, & x < i\theta. \end{cases}$$

Prove que $T = \max_i (X_i/i)$ é uma estatística suficiente para θ .

Exercício 30: Seja X_1, \dots, X_n uma amostra aleatória da f.d.p.

$$f(x|\mu, \sigma) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma}, \quad \mu < x < \infty, \quad 0 < \sigma < \infty.$$

Encontre uma estatística suficiente bi-dimensional para (μ, σ) .

Exercício 31: Sejam X_1, \dots, X_n variáveis aleatórias independentes com f.d.p

$$f(x_1|\theta) = \begin{cases} \frac{1}{2i\theta}, & -i(\theta-1) < x_1 < i(\theta+1), \\ 0, & c.c. \end{cases}$$

em que $\theta > 0$. Encontre uma estatística suficiente bi-dimensional para θ .

Exercício 32: Sejam X_1, \dots, X_n uma amostra aleatória de uma gamma(α, β). Encontre uma estatística suficiente bi-dimensional para (α, β) .

Exercício 33: Seja $f(x, y|\theta_1, \theta_2, \theta_3, \theta_4)$ uma f.d.p. bivariada uniformemente distribuída no retângulo de base (θ_1, θ_2) e altura (θ_3, θ_4) no \mathbb{R}^2 . Os parâmetros safisfaz $\theta_1 < \theta_3$ e $\theta_2 < \theta_4$. Seja $(X_1, Y_1), \dots, (X_n, Y_n)$ uma amostra aleatória dessa f.d.p. Encontre uma estatística suficiente 4-dimensional para $(\theta_1, \theta_2, \theta_3, \theta_4)$.

(34-35. Casella e Berger - Cap 7)

Exercício 34: Seja X_1, \dots, X_n uma amostra aleatória da f.d.p.

$$f(x|\theta) = \theta x^{-2}, \quad 0 < \theta \leq x < \infty.$$

Qual é a estatística suficiente para θ ?

Exercício 35: As variáveis aleatórias X_1, \dots, X_n têm distribuição comum

$$P(X_i < x | \alpha, \beta) = \begin{cases} 0 & , \quad x < 0 \\ (x/\beta)^\alpha & , \quad 0 \leq x \leq \beta \\ 1 & , \quad x > \beta, \end{cases}$$

em que os parâmetros α e β são positivos. Encontre uma estatística suficiente bi-dimensional para (α, β) .

Exercício 36: $\Theta = [0, 1]^2$. $\theta = (\theta_1, \theta_2)$. θ_1 e θ_2 independentes com $\theta_i \sim Beta(a_i, b_i)$, $a_i, b_i > 0$, $i = 1, 2$. $\mathfrak{X} = \{0, 1\}^n$. $X_1 | \theta \sim Ber(\theta_1)$. Para $j = 2, \dots, n$, $X_j | \theta, X_1 = x_1, \dots, X_{j-1} = 0 \sim Ber(\theta_2)$ e $X_j | \theta, X_1 = x_1, \dots, X_{j-1} = 1 \sim Ber(3/4)$.

Obtenha a posteriori de θ dado $X_1 = x_1, \dots, X_n = x_n$. Mostre que $T(X_1, \dots, X_n) = (X_1, \sum_{i=1}^{n-1} X_i, \sum_{i=1}^{n-1} (1 - X_i)X_{i+1})$ é suficiente para θ .

Exercício 37: $\Theta = [0, 1]^3$. $\theta = (\theta_1, \theta_2, \theta_3)$. θ_1, θ_2 e θ_3 independentes com $\theta_i \sim Beta(a_i, b_i)$, $a_i, b_i > 0$, $i = 1, 2, 3$. $\mathfrak{X} = \{0, 1\}^n$. $X_1 | \theta \sim Ber(\theta_1)$. Para $j = 2, \dots, n$, $X_j | \theta, X_1 = x_1, \dots, X_{j-1} = 0 \sim Ber(\theta_2)$ e $X_j | \theta, X_1 = x_1, \dots, X_{j-1} = 1 \sim Ber(\theta_3)$. Obtenha a posteriori de θ dado $X_1 = x_1, \dots, X_n = x_n$. Obtenha $T : \mathfrak{X} \rightarrow \mathbb{N}^4$ suficiente para θ .

Exercício 38: $\Theta = \mathbb{R}_+$. $\theta \sim Gama(a, b)$, $a, b > 0$. $\mathfrak{X} = \mathbb{N}^n$. $X_1 | \theta \sim Geo(1 - e^{-\theta})$. Para $j = 2, \dots, n$, $X_j | \theta, X_1 = x_1, \dots, X_{j-1} = x_{j-1} \sim Poisson(\theta(1 + x_{j-1}))$. Obtenha a posteriori de θ dado $X_1 = x_1, \dots, X_n = x_n$.

Exercício 39: $\Theta = \mathbb{R}_+$. $\theta \sim Gama(a, b)$, $a, b > 0$. $\mathfrak{X} = \mathbb{R}^n$. $X_1, \dots, X_n | \theta$, condicionalmente independentes tais que $X_i | \theta \sim N\left(1, \frac{i}{\theta}\right)$, $i = 1, \dots, n$. Obtenha a posteriori de θ dado $X_1 = x_1, \dots, X_n = x_n$. Mostre, pela definição de suficiência bayesiana, que $\sum_{i=1}^{n-1} \frac{(X_i - 1)^2}{i}$ é suficiente para θ .

Exercício 40: $\Theta = \{(a, b) \in \mathbb{R}^2; a < b\}$. X_1, \dots, X_n , dado $\theta = (\theta_1, \theta_2)$, condicionalmente independentes e identicamente distribuídas tais que $X_1|\theta \sim Uniforme(\theta_1, \theta_2)$. Mostre que $(X_{(1)}, X_{(n)})$ é, pela definição bayesiana de suficiência, uma estatística suficiente para θ , onde $X_{(1)} = \min \{X_1, \dots, X_n\}$ e $X_{(n)} = \max \{X_1, \dots, X_n\}$.

Exercício 41: Seja $\mathfrak{X} = \{0, 1, \dots, 2M + 1\}$, $M \in \mathbb{N}^*$. Suponha que para todo $\theta \in \Theta$ e para todo $x \in \mathfrak{X}$,

$$P(X = x|\theta) = P(X = 2M + 1 - x|\theta).$$

Mostre que $T(X) = \min \{X, 2M + 1 - X\}$ é suficiente para θ .

Exercício 42: $\Theta = [0, 1]$.

		$X_1 \setminus X_2$	
		0	1
0	0	$\frac{\theta}{2}$	$\frac{\theta}{2}$
	1	$\frac{(1-\theta)}{2}$	$\frac{(1-\theta)}{2}$

Verifique se X_1 é suficiente para θ . Faça o mesmo para X_2 .

Exercício 43: $\Theta = [0, 1]$.

		$X_1 \setminus X_2$		
		0	1	2
0	0	$\frac{\theta^2}{2}$	$\frac{(1-\theta)^2}{2}$	
	1	$\frac{\theta^2}{12}$	$\frac{\theta(1-\theta)}{2}$	$\frac{\theta^3}{3}$
	2	$\frac{(1-\theta)^2}{2}$	$\frac{\theta^2}{12}$	$\frac{\theta(1-\theta)}{2}$

Verifique se $T(X_1, X_2) = |X_1 - X_2|$ é suficiente para θ .

List 2 - Inferência

(1-5. Morris Degroot - Cap 6.4)

Exercício 1: Suponha que a proporção θ de itens defeituosos em uma grande remessa é desconhecido, e que a distribuição a priori de θ é uma $Beta(5, 10)$. Suponha também que 20 itens são selecionados de forma aleatória e 1 item é defeituoso. Considerando a perda quadrática, qual é a estimativa de Bayes para θ ?

Exercício 2: Suponha que o tempo em minutos necessário para atender um cliente tenha um distribuição exponencial com parâmetro θ desconhecido; a distribuição a priori para θ é uma distribuição gama com média 0.2 e desvio-padrão 1; e que o tempo médio observado em uma amostra aleatória de 20 clientes foi 3.8 minutos. Usando a perda quadrática qual será a estimativa de θ ?

Exercício 3: Suponha que uma amostra aleatória é retirada da $N(\theta, 4)$; a distribuição a priori para θ é $N(\cdot, 1)$; e que θ deve ser estimado utilizando a perda quadrática. Qual é o tamanho da amostra aleatória para que o erro quadrático médio do estimador de Bayes de θ seja ≤ 0.01 ?

Exercício 4: Seja θ a proporção de votos registrados numa cidade a favor de uma certa política. Suponha que o valor de θ é desconhecido, e que dois estatísticos A e B atribui diferentes prioris para θ .

$$\xi_A(\theta) = 2\theta, \quad 0 < \theta < 1,$$

$$\xi_B(\theta) = 4\theta^3, \quad 0 < \theta < 1.$$

Numa amostra aleatória de 1000 votos, verifica-se que 710 são a favor da política.

a) Encontre a distribuição a posterior de cada estatístico.

b) Encontre a estimativa de Bayes de cada estatístico com base na perda quadrática.

c) Mostre que após a opinião dos 1000 eleitores na amostra aleatória foram obtidas, a estimativa de Bayes para os dois estatísticos não poderia diferir por mais de 0.002, independentemente do números de votos favoráveis a política.

Exercício 5: Suponha que X_1, \dots, X_n forma uma amostra aleatória da $U(0, \theta)$, com θ desconhecido. Suponha também que a distribuição a priori para θ é uma $Pareto(x_0, \alpha)$. Se o valor de θ é estimado usando a perda quadrática, qual é o estimador de Bayes para θ ?

(6-11. Morris Degroot - Cap 6.5)

Exercício 6.0: Não se sabe em que proporção p a compra de uma determinada marca de cereal de café da manhã é feitas por mulheres e qual é a proporção feitas por homens. Em uma amostra aleatória de 70 compras deste cereal, foi encontrado que 58 foram feitas por mulheres e que 12 foram feitas por homens. Encontre o EMV de p .

Exercício 6: Considere novamente as condições do exercício anterior, mas agora suponha que $1/2 \leq p \leq 2/3$. Qual é o *EMV* de p ?

Exercício 7: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição exponencial com parâmetro β desconhecida. Encontre o *EMV* de β .

Exercício 8: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição com f.d.p.

$$f(x|\theta) = \begin{cases} e^{\theta-x} & , \quad x > \theta, \\ 0 & , \quad x \leq \theta. \end{cases}$$

Também, suponha que o valor de θ é desconhecido ($-\infty < \theta < \infty$).

a) Mostre que o *EMV* de θ não existe.

b) Determine outra versão da f.d.p. desta mesma distribuição para que o *EMV* de θ exista, e encontre este estimador.

Exercício 9: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição com f.d.p. dada por:

$$f(x|\theta) = \begin{cases} \theta x^{\theta-1} & , \quad 0 < x < 1, \\ 0 & , \quad \text{caso contrário.} \end{cases}$$

Suponha também que $\theta > 0$ é desconhecido. Encontre o *EMV* de θ .

Exercício 10: Suponha que X_1, \dots, X_n forma uma amostra aleatória da $U(\theta_1, \theta_2)$, em que θ_1 e θ_2 são desconhecidos ($-\infty < \theta_1 < \theta_2 < \infty$). Encontre os *EMV*'s de θ_1 e θ_2 .

Exercício 11: Suponha que uma certa população contém k diferentes tipos de indivíduos ($k \geq 2$); e seja θ_i a proporção de indivíduos do tipo i , para $i = 1, \dots, k$. $0 \leq \theta_i \leq 1$ e $\theta_1 + \dots + \theta_k = 1$. Suponha também que uma amostra aleatória de n indivíduos é retirada da população, com exatamente n_i indivíduos do tipo i , em que $n_1 + \dots + n_k = n$. Encontre os *EMV*'s de $\theta_1, \dots, \theta_k$.

(12-18. Morris Degroot - Cap 6.6)

Exercício 12: Suponha que as variáveis aleatórias X_1, \dots, X_n forma uma amostra aleatória da distribuição Poisson com média desconhecida. Determine o *EMV* do desvio-padrão da distribuição.

Exercício 13: Suponha que o tempo de vida de um certo tipo de lâmpada tem distribuição exponencial de parâmetro β desconhecido. Uma amostra aleatória de n lâmpadas são testadas por um período de T horas e X é o número de lâmpadas que queimaram durante o período de observação; mas o tempo em que as falhas ocorre não é anotado. Determine o *EMV* de β baseado em X .

Exercício 14.0: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição normal com média e variância desconhecidas. Encontre o *EMV* do quantil 0.95 da distribuição, ou seja, o ponto θ tal que $P(X < \theta) = 0.95$.

Exercício 14: Para as condições do exercício anterior, encontre o *EMV* de $v = P(X > 2)$.

Exercício 15: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição gama com α e β desconhecido. Encontre o *EMV* de α/β .

Exercício 16: Suponha que 21 observações são retiradas aleatoriamente da distribuição exponencial com média μ desconhecida; a média de 20 destas observações é de 6; e que, embora o valor exato da outra observação não pudesse ser determinado, sabe-se que é maior que 15. Determine o *EMV* de μ .

Exercício 17: Suponha que cada um de dois estatísticos A e B deve estimar um certo parâmetro θ que é desconhecido ($\theta > 0$). O estatístico A pode observar o valor de uma variável aleatória $X \sim Gama(3, \theta)$; e o estatístico B pode observar o valor de uma variável aleatória $Y \sim Poisson(2\theta)$. Suponha que o valor observado pelo estatístico A foi $X = 2$ e o valor observado pelo estatístico B foi $Y = 3$. Mostre que a função de verossimilhança determinadas por essas observações são proporcionais, e encontre o valor comum do *EMV* de θ obtido por cada estatístico.

Exercício 18: Suponha que cada um de dois estatísticos A e B deve estimar um certo parâmetro p que é desconhecido ($0 < p < 1$). O estatístico A pode observar o valor de uma variável aleatória $X \sim Binomial(10, p)$; e o estatístico B pode observar o valor de uma variável aleatória $Y \sim BinomialNegativa(4, p)$. Suponha que o valor observado pelo estatístico A foi $X = 4$ e o valor observado pelo estatístico B foi $Y = 6$. Mostre que a função de verossimilhança determinadas por essas observações são proporcionais, e encontre o valor comum do *EMV* de p obtido por cada estatístico.

(19-27. Morris Degroot - Cap 6.10)

Exercício 19: Suponha que X_1 e X_2 são variáveis aleatórias independentes, e que X_i tem distribuição $N(b_i\mu, \sigma_i^2)$ para $i = 1, 2$. Suponha também que b_1, b_2, σ_1^2 e σ_2^2 são constantes positivas conhecidas, e que μ é um parâmetro desconhecido. Determine o *EMV* de μ baseado em X_1 e X_2 .

Exercício 20: Suponha que uma lâmpada regular, uma lâmpada de longa vida e uma lâmpada super longa vida estão sendo testadas. O tempo de vida da lâmpada regular é $X_1 \sim Exp(1/\theta)$, o tempo de vida da lâmpada longa vida é $X_2 \sim Exp(1/2\theta)$ e o tempo de vida da lâmpada super longa vida é $X_3 \sim Exp(1/3\theta)$.

a) Determine o *EMV* de θ baseado nas observações de X_1, X_2 e X_3 .

b) Seja $\psi = 1/\theta$, e suponha que a distribuição a priori de $\psi \sim Gama(\alpha, \beta)$. Determine a distribuição a posteriori de $\psi|X_1, X_2, X_3$.

Exercício 21: Suponha que uma observação de X é obtida da distribuição cuja f.d.p. é:

$$f(x|\theta) = \begin{cases} \frac{1}{\theta} & , \quad 0 < x < \theta, \\ 0 & , \quad c. c. \end{cases}$$

Também, suponha que a f.d.p. da priori é

$$\xi(\theta) = \begin{cases} \theta e^{-\theta} & , \quad \theta > 0, \\ 0 & , \quad c. c. \end{cases}$$

Determine o estimador de Bayes de θ com respeito a perda quadrática e perda absoluta.

Exercício 22: O método de *randomized response* é às vezes usada para realizar pesquisas em tópicos sensíveis. Uma simples versão do método pode ser descrita a seguir: Uma amostra aleatória de n pessoas é retirada de uma população. Para cada pessoa na amostra há uma probabilidade de $1/2$ de que a pessoa responda uma questão padrão e probabilidade de $1/2$ de que a pessoa responda uma questão sensível. Além disso, esta seleção de questão padrão ou sensível é feita de forma independente de pessoa para pessoa. Se uma pessoa responde uma questão padrão, então há probabilidade de $1/2$ de que ela dará uma resposta positiva; mas se ela responde uma questão sensível, então há uma probabilidade p de que ela dará uma resposta positiva. Um estatístico só pode observar o número total X de repostas positivas que foi dada pelas n pessoas na amostra. Ele não pode observar quais pessoas responderam as questões sensíveis ou quantas pessoas da amostra responderam as questões sensíveis. Determine o *EMV* de p baseado na observação X .

Exercício 23.0: Suponha que uma amostra aleatória de 4 observações são extraídas de uma distribuição $Uniforme(0, \theta)$, e que a distribuição a priori tem a seguinte f.d.p:

$$f(x|\theta) = \begin{cases} \frac{1}{\theta^2} & , \quad \theta \geq 1, \\ 0 & , \quad c. c. \end{cases}$$

Suponha que os valores observados foram $0.6, 0.4, 0.8$ e 0.9 . Determine a estimativa de Bayes de θ com respeito a função de perda quadrática.

Exercício 23: Para as condições do exercício anterior, determine a estimativa de Bayes de θ com respeito a função de perda absoluta.

Exercício 24: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição com f.d.p.:

$$f(x|\beta, \theta) = \begin{cases} \beta e^{-\beta(x-\theta)} & , \quad x \geq \theta, \\ 0 & , \quad c. c. \end{cases}$$

em que β e θ são desconhecidos ($\beta > 0, -\infty < \theta < \infty$). Determine o par de estatísticas conjuntamente suficiente.

Exercício 25: Suponha que a variável aleatória X tem distribuição binomial com n desconhecido e p conhecido. Determine o *EMV* de n baseado em uma observação X .

Dica: Considere a razão

$$\frac{f(x|n+1, p)}{f(x|n, p)}.$$

Exercício 26: Suponha que duas observações X_1 e X_2 são retiradas aleatoriamente de uma distribuição uniforme com f.d.p.

$$f(x|\theta) = \begin{cases} \frac{1}{2\theta} & , \quad 0 \leq x \leq \theta \text{ ou } 2\theta \leq x \leq 3\theta, \\ 0 & , \quad \text{c. c.} \end{cases}$$

em que o valor de $\theta > 0$ é desconhecido. Determine o EMV de θ para cada um dos seguintes pares de observações de X_1 e X_2 .

- a) $X_1 = 7$ e $X_2 = 9$.
- b) $X_1 = 4$ e $X_2 = 9$.
- c) $X_1 = 5$ e $X_2 = 9$.

Exercício 27: Suponha que uma amostra aleatória X_1, \dots, X_n retirada de uma $N(\theta, 100)$, θ desconhecido; e que a distribuição a priori para θ é uma $N(\mu, 25)$, com μ especificado. Suponha que θ foi estimado usando a perda quadrática, e que o custo da amostragem de cada observação é 0.25 (unidades). Se o custo total do processo de estimação é igual à perda esperada do estimador de Bayes mais o custo de amostragem $(0.25)n$, qual é o tamanho da amostra (n) que minimiza o custo total?

(28-32. George Roussas - 12)

Exercício 28: Suponha que certas partículas são emitidas de uma fonte radioativa de acordo com uma distribuição Poisson com parâmetro θ durante uma unidade de tempo. A fonte em questão é observada em n unidades de tempo, e seja X a variável aleatória que denota o número de vezes que nenhuma partícula foi emitida. Encontre o EMV de θ em termos de X .

Exercício 29: Sejam X_1, \dots, X_n variáveis aleatórias i.i.d. com f.d.p dada por

$$f(x|\theta_1, \theta_2) = \frac{1}{\theta_2} \exp \left[-\frac{(x - \theta_1)}{\theta_2} \right], \quad x > \theta_1, \quad \theta = (\theta_1, \theta_2)' \in \Theta = \mathbb{R} \times (0, \infty).$$

Encontre os EMV 's de θ_1, θ_2 .

Exercício 30: Sejam X_1, \dots, X_n variáveis aleatórias i.i.d. da $U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, $\theta \in \Theta = \mathbb{R}$, e seja

$$\hat{\theta} = \hat{\theta}(X_1, \dots, X_n) = (X_{(n)} - \frac{1}{2}) + (\cos^2 X_1)(X_{(1)} - X_{(n)} + 1).$$

Então mostre que $\hat{\theta}$ é o *EMV* de θ mas não é uma função apenas da estatística suficiente $(X_{(1)}, X_{(n)})'$.

Exercício 31: Em muitos casos, tal como Binomial, Poisson, Normal, etc, acontece que o *EMV* da média é a média da amostra. Contudo, isso não acontece sempre (necessariamente). Considere o seguinte caso: Sejam X_1, \dots, X_n variáveis aleatórias independentes $N(\theta, \theta)$, $\theta \in \Theta = (0, \infty)$. Então mostre que *EMV* de θ é dado por

$$\hat{\theta}_n = \frac{1}{2} \left[\left(1 + \frac{4}{n} \sum_{j=1}^n X_j^2 \right)^{1/2} - 1 \right].$$

Exercício 32: Sejam X_1, \dots, X_n variáveis aleatórias i.i.d. da distribuição $Beta(\alpha, \beta)$, e encontre os estimadores de momentos de α e β .

(33-39. Casella e Berger - Cap 7)

Exercício 33: Uma observação é retirada de uma variável aleatória discreta X com f.p. $f(x|\theta)$, em que $\theta \in \{1, 2, 3\}$. Encontre o Estimador de Máxima Verossimilhança (*EMV*) de θ .

x	$f(x 1)$	$f(x 2)$	$f(x 3)$
0	$\frac{1}{3}$	$\frac{1}{4}$	0
1	$\frac{1}{3}$	$\frac{1}{4}$	0
2	0	$\frac{1}{4}$	$\frac{1}{4}$
3	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{2}$
4	$\frac{1}{6}$	0	$\frac{1}{4}$

Exercício 34: Seja X_1, \dots, X_n uma amostra aleatória de uma f.d.p.

$$f(x|\theta) = \theta x^{-2}, \quad 0 < \theta \leq x < \infty.$$

a) Encontre o *EMV* de θ .

b) Encontre o estimador pelo método de momentos θ .

Exercício 35: Sejam X_1, \dots, X_n i.i.d. com uma das duas f.d.p. Se $\theta = 0$, então

$$f(x|\theta) = \begin{cases} 1 & , \quad 0 < x < 1 \\ 0 & , \quad x < c.c. , \end{cases}$$

enquanto se $\theta = 1$, então

$$f(x|\theta) = \begin{cases} 1/(2\sqrt{x}) & , \quad 0 < x < 1 \\ 0 & , \quad c.c. \end{cases}$$

Encontre o EMV de θ .

Exercício 36: Sejam X_1, \dots, X_n i.i.d. com f.d.p

$$f(x|\theta) = \theta x^{\theta-1}, \quad 0 \leq x \leq 1, \quad 0 < \theta < \infty.$$

a) Encontre o EMV de θ , e mostre que sua variância $\rightarrow 0$ quando $n \rightarrow \infty$.

b) Encontre o estimador pelo método de momentos de θ .

Exercício 37: Seja X_1, \dots, X_n uma amostra da distribuição expoencial dupla, com f.d.p

$$f(x|\theta) = \frac{1}{2}e^{-|x-\theta|}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty.$$

Encontre o EMV de θ . *Dica:* Considere o caso de n par separado do caso de n ímpar, e encontre o EMV em termos das estatísticas de ordem.

Exercício 38: Seja X e Y variáveis aleatórias exponencial independentes, com

$$f(x|\lambda) = \frac{1}{\lambda}e^{-x/\lambda}, \quad x > 0, \quad f(y|\mu) = \frac{1}{\mu}e^{-y/\mu}, \quad y > 0.$$

Nós observamos Z e W com

$$Z = \min \{X, Y\} \text{ e } W = \begin{cases} 1 & , \quad Z = X, \\ 0 & , \quad Z = Y. \end{cases}$$

Agora assuma que $(Z_i, W_i), i = 1, \dots, n$, são n observações. Encontre os EMV 's de λ e μ .

Exercício 39: Seja X_1, \dots, X_n uma amostra aleatória da *normal inversa* com f.d.p.,

$$f(x|\mu, \lambda) = \left(\frac{\lambda}{2\pi x^3} \right)^{1/2} \exp \left\{ -\lambda(x-\mu)^2 / (2\mu^2 x) \right\}, \quad x > 0.$$

Mostre que os *EMV*'s de μ e λ são

$$\hat{\mu}_n = \bar{X} \text{ e } \hat{\lambda}_n = \frac{n}{\sum_{i=1}^n \left(\frac{1}{X_i} - \frac{1}{\bar{X}} \right)}.$$

List 3 - Inferência

(1. George Roussas - Cap 11)

Exercício 1: Se \mathcal{F} é a família de todas $U(-\theta, \theta), \theta > 0$, então mostre que \mathcal{F} não é completa.

(2-9. George Roussas - Cap 12)

Exercício 2: Sejam X_1, \dots, X_n variáveis aleatórias independentes $U(0, \theta), \theta \in \Theta = (0, \infty)$. Encontre estimadores não viesados da média e variância de X dependendo apenas da estatística suficiente para θ .

Exercício 3: Sejam X_1, \dots, X_n variáveis aleatórias independentes $U(\theta_1, \theta_2), \theta_1 < \theta_2$ e encontre estimadores não viesados para média $(\theta_1 + \theta_2)/2$ e amplitude $(\theta_2 - \theta_1)$ de X dependendo apenas da estatística suficiente para $(\theta_1, \theta_2)'$.

Exercício 4: Sejam X_1, \dots, X_n variáveis aleatórias i.i.d. $U(\theta, 2\theta), \theta \in \Theta = (0, \infty)$ e defina

$$U_1 = \frac{n+1}{2n+1} X_{(n)} \text{ e } U_2 = \frac{n+1}{5n+4} [2X_{(n)} + X_{(1)}].$$

Então mostre que U_1 e U_2 são estimadores não viesados de θ e que U_2 é uniformemente melhor que U_1 (em relação a variância).

Exercício 5: Seja X uma variável aleatória com distribuição geométrica; isto é,

$$f(x|\theta) = \theta(1-\theta)^x, \quad x = 0, 1, \dots, \quad \theta \in \Theta = (0, 1),$$

e seja $U(X)$ definida a seguir: $U(X) = 1$ se $X = 0$ e $U(X) = 0$ se $X \neq 0$. Usando o teorema 1 (do livro), mostre que $U(X)$ é um ENVVUM de θ e conclua que isso não é razoável.

Exercício 6: Seja X uma variável aleatória que denota o número de chamadas telefônicas que chega a uma determinada central telefônica, e suponha que $X \sim Poisson(\theta)$, em que $\theta > 0$ é o número de chamadas que chegam à central telefônica num período de 15 minutos. Então o número de chagamas tefefônicas que chega a uma determinada central tefefônica em 30 minutos é uma variável aleatória

$Y \sim Poisson(2\theta)$, como pode-se demonstrar. Portanto $P(Y=0) = e^{-2\theta} = g(\theta)$. Defina $U(X) = (-1)^x$. Então mostre que $U(X)$ é o ENVVUM de $g(\theta)$ e conclua que é um estimador inteiramente irracional. (Dica: Use o teorema 1 do livro).

Exercício 7: Seja $X \sim Binomial(n, \theta)$. Mostre que não há um ENV para $g(\theta) = 1/\theta$ baseado em X .

Exercício 8: Seja X uma variável aleatória que denota o tempo de vida de um equipamento. Então a confiabilidade do equipamento no tempo x , $R(x)$, é definido como $P(X > x)$. Se $X \sim Exp(\theta)$, encontre o ENVVUM da confiabilidade $R(x|\theta)$ baseado em n observações de X .

Exercício 9: Sejam X_1, \dots, X_n e Y_1, \dots, Y_n duas amostra aleatórias independentes com mesma média θ e variâncias σ_1^2 e σ_2^2 conhecidas. Então mostre que para todo $c \in [0, 1]$, $U = c\bar{X} + (1 - c)\bar{Y}$ é um *ENV* de θ . Também encontre o valor de c para que a variância de U seja minimizada.

(2-9. Morris DeGroot - Cap ?)

Exercício 10: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição cuja f.d.p. é $f(x|\theta)$, em que $\theta \in \Theta$; seja $\hat{\theta}$ o *EMV* de θ . Suponha também que T é uma estatística suficiente para θ ; e seja δ_0 o estimador definido $\delta_0 = E(\hat{\theta}|T)$. Compara os estimadores $\hat{\theta}$ e δ_0 .

Exercício 11: Suponha que X_1, \dots, X_n forma uma amostra aleatória $Poisson(\theta), \theta > 0$. Seja $T = \sum_{i=1}^n X_i$; e para $i = 1, \dots, n$, seja Y_i a estatística definida como:

$$Y_i = 1 \text{ se } X_i = 0,$$

$$Y_i = 0 \text{ se } X_i > 0.$$

Determine a forma do estimador $E(Y_i|T)$.

Exercício 12: Suponha uma amostra aleatória de n observações da distribuição $Poisson(\theta), \theta > 0$, e seja $\beta = e^{-\theta}$, e suponha também as condições do exercício anterior. Determine a forma do estimador $E(\hat{\beta}|T)$.

(13-17. Casella e Berger - Cap 7)

Exercício 13: Seja X_1, X_2 e X_3 uma amostra aleatória de tamanho três da distribuição $Uniform(\theta, 2\theta)$, em que $\theta > 0$.

a) Encontre o estimador de θ pelo método de momentos.

b) Encontre *EMV* de θ , e encontre a constante k tal que $E_\theta(k\hat{\theta}) = \theta$.

c) Qual dos dois estimadores pode ser melhorado usando a suficiência? Como?

d) Encontre o estimador pelo método de momentos e o *EMV* de θ baseado nos dados 1.29, 0.86, 1.33, três observações de tamanhos médios (em centímetros) de uvas para vinho.

Exercício 14: Suponha que $X_i, i = 1, \dots, n$, são $Bernoulli(p)$ i.i.d.

a) Mostre que a variância do *EMV* de p atinge o limite inferior de Cramér-Rao.

b) Para $n \geq 4$, mostre o produto $X_1 X_2 X_3 X_4$ é um estimador não viesado de p^4 , e use essa fato para encontrar o melhor estimador não viesado de p^4 .

Exercício 15: Sejam X_1, \dots, X_n i.i.d. $N(\theta, \theta^2)$, $\theta > 0$. Para este modelos \bar{X} e cS são ENV de θ , em que $c = \frac{\sqrt{n-1}\Gamma((n-1)/2)}{\sqrt{2}\Gamma(n/2)}$.

- a) Prove que para qualquer α o estimador $\alpha\bar{X} + (1-\alpha)cS$ é um ENV de θ .
- b) Encontre o valor de α que leva ao estimador de menor variância.
- c) Mostre que (\bar{X}, S^2) é uma estatística suficiente para θ mas não são estatística suficiente completa.

Exercício 16: Sejam X_1, \dots, X_{n+1} i.i.d. $Bernoulli(p)$, e defina a função $h(p)$ por

$$h(p) = P\left(\sum_{i=1}^n X_i > X_{n+1} | p\right),$$

a probabilidade de que as n primeiras observações ultrapassem os $(n+1)^\circ$.

- a) Mostre que

$$T(X_1, \dots, X_{n+1}) = \begin{cases} 1 & , \quad \sum_{i=1}^n X_i > X_{n+1} \\ 0 & , \quad c.c. \end{cases}$$

é um ENV de $h(p)$.

- b) Encontre o melhor ENV de $h(p)$.

Exercício 17: Seja X uma observação de f.d.p.

$$f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}. \quad x = -1, 0, 1; \quad 0 \leq \theta \leq 1.$$

- a) Encontre o EMV de θ .

- b) Defina o estimador $T(X)$ por

$$T(X) = \begin{cases} 2 & , \quad x = 1 \\ 0 & , \quad c.c. \end{cases}$$

Mostre que $T(X)$ é um ENV de θ .

- c) Encontre um estimador melhor que $T(X)$ e prove que é melhor.

Exercício 18: X_1, \dots, X_n , dado θ , c.i.i.d. $Poisson(\theta)$. Obtenha, se possível, o ENVVUM para

- a) $g_1(\theta) = e^{-\theta}\theta^r, r \in \mathbb{N}$.
- b) $g_2(\theta) = e^{-k\theta}, k \in \mathbb{N}, k < n$.
- c) $g_3(\theta) = e^{-2\theta}\theta^2$.

Exercício 19: $X_1 \dots, X_n$, dado θ , c.i.i.d. tais que

$$P(X = j|\theta) = \left(\frac{1}{2}\right)^{j+1-\theta} \mathbb{I}_{\theta, \theta+1, \dots}(j). \quad \Theta = \mathbb{Z}.$$

Obtenha, se possível, o ENVVUM de θ .

Exercício 20: X, \dots, X_n , dado θ , c.i.i.d. $N(\theta, 1)$. Seja $a_0 \in \mathbb{R}$ (conhecido). Obtenha, se possível, o EMVVUM para $g(\theta) = P(X_1 \leq a|\theta)$. Suponha \bar{X} é competente e que $X_1|\bar{X} = \bar{x} \sim N(\bar{x}, \frac{n-1}{n})$.

Exercício 21: X_1, X_2, X_3 tais que $X_1 \sim Geo(1/2)$, $X_2|X_1 = x_1 \sim Poisson(\theta x_1)$ e $X_3|X_1 = x_1, X_2 = x_2 \sim Poisson(\theta x_2), \theta > 0$.

- a) Obtenha uma estatística suficiente para θ assumindo valores em \mathbb{N}^2 .
- b) Obtenha $c \in \mathbb{R}$ de modo que $\delta(x) = cX_2$ seja não-viesado para θ .
- c) “Rao-Blackwellize” δ encontrado em b).

List 4 - Inferência

(1-3. George Roussas - Cap 12)

Exercício 1: Seja X_1, \dots, X_n variáveis aleatórias i.i.d. com f.d.p. $f(\cdot|\theta)$, $\theta \in \Theta \subset \mathbb{R}$. Para um estimador $V = V(X_1, \dots, X_n)$ de θ com $E_\theta(V)$ é finito, escreva $E_\theta(V) = \theta + b(\theta)$. Então $b(\theta)$ é chamado de vício de V . Mostre que, sobre as condições de regularidade (teorema 2 do livro)

$$\sigma_\theta^2(V) \geq \frac{[1 + b'(\theta)]^2}{nE_\theta[(\partial/\partial\theta)\log f(X|\theta)]}, \quad \theta \in \Theta.$$

Aqui X é uma variável aleatória com f.d.p. $f(\cdot|\theta)$ e $b'(\theta) = db(\theta)/d\theta$.

Exercício 2: Seja X_1, \dots, X_n variáveis aleatórias de uma distribuição $gama(\alpha, \theta)$, α conhecido e $\theta \in (0, \infty)$ desconhecido. Então mostre que ENVVUM de θ é

$$U(X_1, \dots, X_n) = \frac{1}{n\alpha} \sum_{j=1}^n X_j$$

e a variância atinge o limite de Cramér-Rao.

Exercício 3: Seja X_1, \dots, X_n variáveis aleatórias independentes com distribuição $N(\theta, 1)$. Mostre que $\overline{X}^2 - (1/n)$ é o ENVVUM de $g(\theta) = \theta^2$. Também mostre que o limite de Cramér-Rao não é atingido.

(4-9. Morris DeGroot - Cap 7.7)

Exercício 4: Suponha que a variável aleatória X tem distribuição Poisson com média λ desconhecida. Encontre a estatística $\delta(X)$ que vai ser um estimador não viesado de e^λ .

Dica: Se $E[\delta(X)] = e^\lambda$, então

$$\sum_{x=0}^{\infty} \frac{\delta(x)e^{-\lambda}\lambda^x}{x!} = e^\lambda.$$

Multiplique ambos os lados da equação por e^λ ; expandir o lado direto como uma série de potência em λ ; e iguale os coeficientes de λ^x em ambos os lados da equação para $x = 0, 1, \dots$.

Exercício 5: Suponha X_1, \dots, X_n forma uma amostra aleatória de uma distribuição $N(\mu, \sigma^2)$ com ambos parâmetros desconhecidos. Sejam S_0^2 e S_1^2 os dois estimadores de σ^2 que são definidos a seguir:

$$S_0^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \text{ e } S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Mostre que o E.Q.M. do S_0^2 é menor que o E.Q.M. do S_1^2 para todos possíveis valores de μ e σ^2 .

Exercício 6: Suponha que uma variável aleatória X tem distribuição Poisson com média λ desconhecida. Mostre que o único estimador não viesado de $e^{-2\lambda}$ é o estimador $\delta(X)$ tal que $\delta(X) = 1$ se X é par e $\delta(X) = -1$ se X é ímpar.

Exercício 7: Considere uma sequência infinita de ensaios de Bernoulli com parâmetro p desconhecido, e suponha que a amostragem continua até obter exatamente k sucessos, em que k é um inteiro fixo ($k \geq 2$). Seja N denotar o número total de ensaios que são necessários para obter k sucessos. Mostre que o estimador $(k-1)/(N-1)$ é um estimador não viesado de p .

Exercício 8: Suponha que uma certa droga dada a dois diferentes tipos de animais A e B. E sabe-se que a resposta média dos animais do tipo A é a mesma média dos animais do tipo B; mas que o valor comum θ é a média desconhecida e deve ser estimada. Também se sabe que a variância da resposta dos animais do tipo A é quatro vezes a variância da resposta dos animais do tipo B. Seja X_1, \dots, X_m denotar uma amostra aleatória da resposta de m animais do tipo A, e Y_1, \dots, Y_n denotar uma amostra aleatória da resposta de n animais do tipo B. Finalmente, considere o estimador $\hat{\theta} = \alpha \bar{X}_m + (1-\alpha) \bar{Y}_n$.

a) Para qual valor de α , m e n , $\hat{\theta}$ é um estimador não viesado de θ ?

b) Para valores fixo de m e n , qual valor de α produzirá o estimador não viesado com mínima variância?

Exercício 9: Suponha que um certa população de indivíduos é composta de k diferentes estratos ($k \geq 2$); e que para $i = 1, \dots, k$, a proporção de indivíduos na população do estrato i é p_i , em que $p_i > 0$ e $\sum_{i=1}^k p_i = 1$. Nossa interesse é estimar a média μ de certa característica da população. Entre os indivíduos do estrato i que tem essa característica tem média μ_i e variância σ_i^2 , em que o valor de μ_i é desconhecido e σ_i^2 é conhecido. Suponha que uma amostra estratificada é retirada da população da seguinte forma: De cada estrato i , uma amostra aleatória de n_i indivíduos são retirados. As amostras dos k estratos são feitas de forma independentes. Seja \bar{X}_i a média das n_i medidas da amostra do estrato i .

a) Mostre que $\mu = \sum_{i=1}^k p_i \mu_i$, e mostre também que $\hat{\mu} = \sum_{i=1}^k p_i \bar{X}_i$ é um estimador não viesado de μ .

b) Seja $n = \sum_{i=1}^k n_i$ o número total de observações nas k amostras. Para um n fixo, encontre os n_1, \dots, n_k para que a variância de $\hat{\mu}$ seja mínima.

(10-13. Morris DeGroot - Cap 7.8)

Exercício 10: Suponha que a variável aleatória tem uma distribuição $N(0, \sigma^2)$, $\sigma^2 > 0$ desconhecido. Encontre a informação de Fisher $I(\sigma)$ de X .

Exercício 11: Suponha que uma única observação X é retirada de uma $N(0, \sigma^2)$, σ^2 desconhecido. Encontre o estimador não viesado de σ , determine sua variância, e mostre que essa variância é maior que $1/I(\sigma)$ para todo valor de $\sigma > 0$. Note que o valor de $I(\sigma)$ foi encontrado no exercício anterior.

Exercício 12: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma $N(0, \sigma^2)$, $\sigma > 0$ desconhecido. Encontre o limite inferior especificado pela desigualdade para a variância de um qualquer estimador não viesado de $\log \sigma$.

Exercício 13: Determine qual é o erro no seguinte argumento: Suponha que a variável aleatória X tem distribuição uniforme no intervalo $(0, \theta)$, em que o valor de $\theta > 0$ é desconhecido. Então $f(x|\theta) = 1/\theta$, $\lambda(x|\theta) = -\log \theta$, e $\lambda'(x|\theta) = -(1/\theta)$. Portanto,

$$I(\theta) = E_\theta\{[\lambda'(X|\theta)]^2\} = \frac{1}{\theta^2}.$$

$2X$ é um estimador não viesado de θ , a desigualdade de afirma que

$$\text{Var}(2X) \geq \theta^2.$$

Mas

$$\text{Vas}(2X) = 4\text{Var}(X) = 4 \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3} < \theta^2.$$

Daí, a desigualdade não é correta.

(14-21. Morris DeGroot - Cap 7.9)

Exercício 14: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição de probabilidade P desconhecida na reta. Seja A um subconjunto da reta, e seja $\theta = P(A)$. Considere um estimador não viesado de θ , e especifique sua variância.

Exercício 15: Suponha que X_1, \dots, X_m forma uma amostra aleatória $N(\mu_1, \sigma^2)$; e que independentemente Y_1, \dots, Y_n forma uma amostra aleatória $N(\mu_2, 2\sigma^2)$. Seja $S_X^2 = \sum_{i=1}^m (X_i - \bar{X}_m)^2$ e $S_Y^2 = \sum_{i=1}^n (Y_i - \bar{Y}_n)^2$.

a) Para quais valores de α e β o estimador $\alpha S_X^2 + \beta S_Y^2$ é um estimador não viesado de σ^2 ?

b) Determine os valores de α e β para que $\alpha S_X^2 + \beta S_Y^2$ seja um estimador não viesado com variância mínima.

Exercício 16: Suponha que X_1, \dots, X_n forma uma amostra aleatória Poisson com média θ desconhecida, e seja $Y = \sum_{i=1}^n X_i$.

a) Determine o valor da constante c tal que o estimador e^{-cY} é um estimador não viesado de $e^{-\theta}$.

b) Use a desigualdade para obter o limite inferior da variância do estimador não viesado encontrado em a).

Exercício 17: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição com f.d.p.

$$f(x|\theta) = \begin{cases} \theta x^{\theta-1} & , \quad 0 < x < 1, \\ 0 & , \quad \text{c. c.} \end{cases}$$

em que o valor $\theta > 0$ é desconhecido. Determine a distribuição assintótica do EMV de θ .

Exercício 18: Suponha que uma variável aleatória X tem uma distribuição exponencial com parâmetro $\beta > 0$ desconhecido. Encontre a informação de Fisher $I(\beta)$ de X .

Exercício 19: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição Bernoulli com parâmetro p desconhecido. Mostre que a variância de qualquer estimador não viesado de $(1-p)^2$ deve ser menor que $4(1-p)^3/n$.

Exercício 20: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição exponencial com parâmetro β desconhecido. Mostre que se n é grande, a distribuição do EMV de β vai se aproximar de uma distribuição $N(\beta, \beta^2/n)$.

Exercício 21: Considere novamente as condições do exercício anterior, e seja $\hat{\beta}_n$ denota o EMV de β .

a) Use o método delta para determinar a distribuição assintótica de $1/\hat{\beta}_n$.

b) Mostre que $1/\hat{\beta}_n = \bar{X}_n$, use o teorema central do limite para determinar a distribuição assintótica de $1/\hat{\beta}_n$.

List 5 - Inferência

(1-3. Morris DeGroot - Cap 7.5)

Exercício 1: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição $N(\mu, \sigma^2)$ com ambos parâmetros desconhecidos; e seja a variável aleatória L que é o comprimento do intervalo de confiança mais curto para μ que pode ser construído com os valores observados na amostra. Encontre $E(L^2)$ para os seguintes valores de n e coeficiente de confiança γ .

- a) $n = 5, \gamma = 0.95$.
- b) $n = 10, \gamma = 0.95$.
- c) $n = 30, \gamma = 0.95$.
- d) $n = 8, \gamma = 0.90$.
- e) $n = 8, \gamma = 0.95$.
- f) $n = 8, \gamma = 0.99$.

Exercício 2: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição $N(\mu, \sigma^2)$ com ambos parâmetros desconhecidos. Qual tamanho amostral (n) necessário para que o intervalo de confiança de μ com coeficiente de confiança 0.95 seja menor que 0.01σ .

Exercício 3: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição exponencial com média μ desconhecida. Descreva o método de construção do intervalo de confiança de μ com coeficiente de confiança $\gamma \in (0, 1)$.

Dica: Determine as constantes c_1 e c_2 tal que $P[c_1 < (1/\mu) \sum_{i=1}^n X_i < c_2] = \gamma$.

(4-12. George Roussas - Cap ?)

Exercício 4: Sejam X_1, \dots, X_n variáveis aleatórias independentes exponenciais com parâmetro θ . Construa um intervalo de confiança para θ com coeficiente de confiança $\gamma \in (0, 1)$.

Exercício 5: Sejam X_1, \dots, X_n variáveis aleatórias i.i.d. com f.d.p. dada por

$$f(x|\theta) = e^{-(x-\theta)} I_{(\theta, \infty)}(x), \quad \theta \in \Theta = \mathbb{R},$$

e seja $Y_1 = X_{(1)}$. Então mostre que:

- a) A f.d.p. g de Y_1 é dada por $g(y) = n e^{-n(y-\theta)I_{(\theta, \infty)}(y)}$.
- b) A variável aleatória $T_n(\theta) = 2n(Y_1 - \theta)$ é distribuída como χ_2^2 .

- c) O intervalo de confiança de θ , baseado em $T_n(\theta)$, com intervalo de confiança $\gamma = 1 - \alpha \in (0, 1)$ é da forma $[Y_1 - (b/2n), Y_1 - (a/2n)]$.

- d) O menor intervalo de confiança da forma dada em c) é fornecido por

$$\left[Y_1 - \frac{\chi_{2;\alpha}^2}{2n}, Y_1 \right],$$

em que $\chi_{2;\alpha}^2$ é o quantil de ordem α da distribuição χ_2^2 .

Exercício 6: Sejam X_1, \dots, X_n variáveis aleatórias independentes da distribuição Weibull. Mostre que:

- a) A variável aleatória $T_n(\theta) = 2Y/\theta$ tem distribuição χ_{2n}^2 , em que $Y = \sum_{j=1}^n X_j$.
- b) O intervalo de confiança de θ , baseado em $T_n(\theta)$, com coeficiente de confiança $\gamma = 1 - \alpha$ é $[2Y/b, 2Y/a]$.
- c) O menor intervalo de confiança da forma dado em b) é tomado a e b que satisfaz a equação

$$\int_a^b g_{2n}(t)dt = \gamma \text{ e } a^2 g_{2n}(a) = b^2 g_{2n}(b).$$

Em que g_{2n} é a f.d.p. da distribuição χ_{2n}^2 .

Exercício 7: Sejam X_1, \dots, X_n variáveis aleatórias i.i.d. com f.d.p. dada por

$$f(x|\theta) = \frac{1}{2\theta} e^{-|x|/\theta}, \quad \theta \in \Theta = (0, \infty).$$

Mostre que:

- a) A variável aleatória $T_n(\theta) = \frac{2Y}{\theta}$ segue uma distribuição χ_{2n}^2 , em que $Y = \sum_{j=1}^n |X_j|$.

b) e c) como do exercício anterior.

Exercício 8: Considere X_1, \dots, X_m variáveis aleatórias da distribuição $N(\mu_1, \sigma_1^2)$ e independentemente Y_1, \dots, X_n variáveis aleatórias independentes da distribuição $N(\mu_2, \sigma_2^2)$, em que σ_1, σ_2 são conhecidos e μ_1, μ_2 são desconhecidos, e seja a variável aleatória $T_{m,n}(\mu_1 - \mu_2)$ definida por

$$T_{m,n}(\mu_1 - \mu_2) = \frac{(\bar{X}_m - \bar{Y}_m) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/m) + (\sigma_2^2/n)}}.$$

Mostre que:

a) O intervalo de confiança para $\mu_1 - \mu_2$, baseado em $T_{m,n}(\mu_1 - \mu_2)$, com coeficiente de confiança $\gamma = 1 - \alpha$ é dado por

$$\left[(\bar{X}_m - \bar{Y}_m) - b\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}, (\bar{X}_m - \bar{Y}_m) - a\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} \right],$$

em que a e b são tal que $\Phi(b) - \Phi(a) = \gamma$.

b) O menor intervalo de confiança acima é obtido quando $-a = b = z_{\alpha/2}$.

Exercício 9: Referente ao exercício anterior mas agora suponha que μ_1 e μ_2 são conhecidos e σ_1 e σ_2 são desconhecido. Considere a variável aleatória

$$\bar{T}_{m,n} \left(\frac{\sigma_1}{\sigma_2} \right) = \frac{\sigma_1^2}{\sigma_2^2} \frac{S_n^2}{S_m^2}$$

e mostre que o intervalo de confiança de σ_1^2/σ_2^2 , baseado em $\bar{T}_{m,n}(\sigma_1/\sigma_2)$, com coeficiente de confiança $\gamma = 1 - \alpha$ é dado por

$$\left[a \frac{S_m^2}{S_n^2}, b \frac{S_m^2}{S_n^2} \right],$$

em que $0 < a < b$ são tal que $P(a \leq F_{n,m} \leq b) = \gamma$.

Exercício 10: Sejam X_1, \dots, X_m e Y_1, \dots, Y_n variáveis aleatórias independentes de uma distribuição exponencial com parâmetro θ_1 e θ_2 , respectivamente. Seja $\theta_1/\theta_2 = \Delta$ e derive o intervalo de confiança de Δ com coeficiente de confiança $\gamma = 1 - \alpha$.

Exercício 11: Sejam X_1, \dots, X_n variáveis aleatórias i.i.d. com média μ (finita) desconhecida e variância σ^2 (finita) desconhecida, e suponha que n é grande. Então um intervalo de confiança para μ com coeficiente de confiança $\gamma = 1 - \alpha$ pode ser aproximado.

a) Qual é esse intervalo para $n = 100$ e $\alpha = 0.05$?

b) Mostre que o comprimento desse intervalo de confiança tende a 0 em probabilidade (e também quase certamente) quando $n \rightarrow \infty$.

c) Discuta o item a) para o caso que a distribuição $B(1, \theta)$, $\theta \in \Theta = (0, 1)$ ou $Poisson(\theta)$, $\theta \in \Theta = (0, \infty)$.

Exercício 12: Seja X_1, \dots, X_n variáveis aleatórias independentes com distribuição exponencial de parâmetro $\theta \in \Theta = (0, \infty)$ e suponha que n é grande. Use o *TCL* para construir um intervalo de confiança para θ com coeficiente de confiança $\gamma = 1 - \alpha$. Compare esse intervalo com o construído no Exercício 4.

Exercício 13: Suponha, a priori, $\theta \sim Beta(2, 1)$. Suponha $X|\theta \sim Ber(\max\{\theta, 1 - \theta\})$.

a) Obtenha a distribuição de $\theta|X = 1$.

b) Obter intervalo (região) *HDP* de probabilidade $1 - \alpha, 0 < \alpha < 1$.

Exercício 14: Suponha que, a priori, θ possui densidade

$$f(\theta) = 4 \left| \theta - \frac{1}{2} \right| \mathbb{I}_{(0,1)}(\theta).$$

Suponha que $X|\theta \sim Bernoulli(\theta)$

a) Obtenha a distribuição a posteriori de $\theta|X = 1$.

b) Obter o intervalo (região) *HPD* de probabilidade $1 - \alpha, 0 < \alpha < 1$.

List 6 - Inferência

(1-5. George Roussas - Cap 13)

Exercício 1: Sejam X_1, \dots, X_n variáveis aleatórias independentes com distribuição $\text{Gamma}(10, \beta)$. Construa um teste MP para as hipóteses $H : \beta = 2 \times A : \beta = 3$ com nível de significância 0.05.

Exercício 2: Sejam X_1, \dots, X_n variáveis aleatórias independentes com distribuição $N(\mu, \sigma^2)$, com média desconhecida e variância conhecida. Mostre que o tamanho da amostra n pode ser determinado quando testamos as hipóteses $H : \mu = 0 \times A : \mu = 1$, com α e β prefixado. Qual é o valor de n para $\alpha = 0.5$, $\beta = 0.9$ e $\sigma = 1$?

Exercício 3: Sejam X_1, \dots, X_n variáveis aleatórias independentes com distribuição $N(\mu, \sigma^2)$, com média desconhecida e variância conhecida. Para testar as hipóteses $H : \mu = \mu_1 \times A : \mu = \mu_2$, mostre que α pode ser arbitrariamente pequeno e β arbitrariamente grande para n suficientemente grande.

Exercício 4: Sejam X_1, \dots, X_{100} variáveis aleatórias independentes com distribuição $N(\mu, \sigma^2)$. Se $\bar{x} = 3.2$, construa o teste MP para as hipóteses $H : \mu = 3, \sigma^2 = 4 \times A : \mu = 3.5, \sigma^2 = 4$ com nível de significância 0.01.

Exercício 5: Seja X uma variável aleatória cuja f.d.p. é $U(0, 1)$, denotada por f_0 , ou é triangular no intervalo $[0, 1]$, denotada por f_1 (isso é, $f_1(x) = 4x, 0 \leq x < 1/2$ e $f_1(x) = 4 - 4x, 1/2 \leq x < 1$). Há uma única observação de X , construa o teste MP para as hipóteses $H : f = f_0 \times A : f = f_1$ com nível de significância 0.05.

(6-10. Morris DeGroot - Cap 8.1)

Exercício 6: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição $U(0, \theta)$, e que as seguintes hipóteses serão testadas: $H : \theta \geq 2 \times A : \theta < 2$.

Seja $Y_n = \max \{X_1, \dots, X_n\}$, e considere o teste tal que a região crítica é $Y_n \leq 1.5$.

a) Determine a função poder do teste.

b) Determine o tamanho do teste.

Exercício 7: Suponha que a proporção p de itens defeituosos de uma grande população de itens é desconhecida, e que é de interesse testar as seguintes hipóteses: $H : p = 0.2 \times A : p \neq 0.2$.

Suponha também que uma amostra aleatória de 20 itens são retirados dessa população. Seja Y denotar o número de itens defeitos na amostra, e considere o teste tal que a região crítica é $Y_n \geq 7$ ou $Y_n \leq 1$.

a) Determine o valor da função poder $\pi(p)$ nos pontos $p = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ e 1.0 , e esboce a função poder.

b) Determine o tamanho do teste.

Exercício 8: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição $N(\mu, 1)$, com média desconhecida. Suponha também que μ_0 é um certo número, e que as seguintes hipóteses serão testadas: $H : \mu = \mu_0 \times A : \mu_0 \neq \mu_0$.

Finalmente, suponha que $n = 25$, e considere o teste tal que H_0 não é rejeitado se $|\bar{X}_n - \mu_0| < c$. Determine o valor de c tal que o tamanho do teste seja 0.05.

Exercício 9: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição $N(\mu, \sigma^2)$, com média e variância desconhecida. Classifique as seguintes hipóteses como simple ou composta:

- a)** $H : \mu = 0$ e $\sigma = 1$.
- b)** $H : \mu = 3$ e $\sigma < 1$.
- c)** $H : \mu = -2$ e $\sigma^2 < 5$.
- d)** $H : \mu = 0$.

Exercício 10: Suponha que uma única observação X é obtida de uma distribuição uniforme no intervalo $\left(\theta - \frac{1}{2}, \theta + \frac{1}{2}\right)$, e suponha que a seguintes hipóteses serão testadas: $H : \theta < 3 \times A : \theta \geq 4$.

Construa um teste que a função poder tenha os seguintes valores: $\pi(\theta) = 0, \theta \leq 3$ e $\pi(\theta) = 1, \theta \geq 4$.

(11-24. Morris DeGroot - Cap 8.2)

Exercício 11: Consdere duas f.d.p. $f_0(x)$ e $f_1(x)$ que são definidas:

$$f_0(x) = \begin{cases} 1 & , \quad 0 \leq x \leq 1, \\ 0 & , \quad c. c, \end{cases}$$

e

$$f_1(x) = \begin{cases} 2x & , \quad 0 \leq x \leq 1, \\ 0 & , \quad c. c. \end{cases}$$

Suponha uma única observação de X com f.d.p $f(x)$ que é $f_0(x)$ ou $f_1(x)$, e que as seguintes hipóteses serão testadas: $H : f(x) = f_0(x) \times A : f(x) = f_0(x)$.

- a)** Descreve o procedimento de encontrar o teste que minimiza $3\alpha(\delta) + \beta(\delta)$.
- b)** Determine o valor mínimo de $3\alpha(\delta) + \beta(\delta)$.

Exercício 12: Considere as condições do exercício anterior, mas suponha agora que o interesse é encontrar o teste que $\alpha(\delta) \leq 0.1$ e $\beta(\delta)$ é mínimo.

a) Descreve o procedimento.

b) Determine o valor mínimo de $\beta(\delta)$ que atende o procedimento.

Exercício 13: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição $N(\theta, 1)$, com média desconhecida; e que as seguintes hipóteses serão testadas: $H : \theta = 3.5 \times A : \theta = 5.0$

a) Dentre todos os teste que $\beta(\delta) \leq 0.05$, descreve o que $\alpha(\delta)$ é mínimo.

b) Para $n = 4$, encontre o valor mínimo de $\alpha(\delta)$ que atende a).

Exercício 14: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição Bernoulli com parâmetro p desconhecido. Sejam p_0 e p_1 os valores especificados tal que $0 < p_0 < p_1 < 1$, e suponha que as seguintes hipóteses serão testadas: $H : p = p_0 \times A : p = p_1$.

a) Mostre que o teste que minimiza $\alpha(\delta) + \beta(\delta)$ é tal que rejeita H quando $\bar{X}_n < c$.

b) Encontre o valor da constante c .

Exercício 15: Suponha que X_1, \dots, X_n forma uma amostra aleatória da distribuição $N(\theta, \sigma^2)$, com média conhecida e variância desconhecida; e que as seguintes hipóteses serão testadas: $H : \sigma^2 = 2 \times A : \sigma^2 = 3$

a) Mostre que dentre todos os teste que $\alpha(\delta) \leq 0.05$, o valor de $\delta(\delta)$ é mínimo quando rejeita H quando $\sum_{i=1}^n (X_i - \mu)^2 > c$.

b) Para $n = 8$, encontre o valor da constante c de a).

Exercício 16: Suponha que uma única observação de X é retirada de uma distribuição uniforme no intervalo $(0, \theta)$, em que θ é desconhecido; e que as seguintes hipóteses simples serão testadas: $H : \theta = 1 \times A : \theta = 2$.

a) Mostre que existe um teste que $\alpha(\delta) = 0$ e $\beta(\delta) < 1$.

b) Dentre todos os teste que $\alpha(\delta) = 0$, encontre o que $\beta(\delta)$ é mínimo.

Exercício 17: Suponha que uma amostra aleatória X_1, \dots, X_n é retirada de uma distribuição uniforma no intervalo $(0, \theta)$; e considere novamente o teste descrito no exercício anterior. Encontre o valor mínimo de $\beta(\delta)$ que pode ser obtido dentre todos os teste que $\alpha(\delta) = 0$.

Exercício 18: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição Poisson com média λ desconhecida. Sejam λ_0 e λ_1 valores especificados tal que $\lambda_1 > \lambda_0 > 0$, e suponha que o interesse é testar as seguintes hipóteses simples:

$H : \lambda = \lambda_0 \times A : \lambda = \lambda_1$.

a) Mostre que $\alpha(\delta) + \beta(\delta)$ é minimizado quando o teste rejeita H quando $\bar{X}_n > c$.

b) Encontre o valor de c .

c) Para $\lambda_0 = 1/4$, $\lambda_1 = 1/2$ e $n = 20$, determine o valor de $\alpha(\delta) + \beta(\delta)$ que pode ser obtido.

Exercício 19: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição $N(\mu, 4)$, com média desconhecida; e que as seguintes hipóteses serão testadas: $H : \mu = -1 \times A : \mu = 1$.

Determine o valor mínimo de $\alpha(\delta) + \beta(\delta)$ que pode ser obtido para cada um dos seguintes n .

a) $n = 1$.

b) $n = 4$.

c) $n = 16$.

d) $n = 36$.

Exercício 20: Uma única observação X é retirada de uma distribuição contínua com f.d.p. f_0 ou f_1 , em que

$$f_0(x) = \begin{cases} 1 & , \quad 0 < x < 1, \\ 0 & , \quad \text{c.c.} \end{cases}$$

e

$$f_1(x) = \begin{cases} 4x^3 & , \quad 0 < x < 1, \\ 0 & , \quad \text{c.c.} \end{cases}$$

Com base na observação X , deve ser decidido se a correta f.d.p. é f_0 ou f_1 . Suponha a priori que a probabilidade de f_0 ser correta é $2/3$ e f_1 é $1/3$. Supona também que a perda de tomar a decisão certa é 0; e a perda de decidir f_1 quando f_0 é a correta é 1 unidade; e a perda de decidir f_0 quando f_1 é a correta é 4 unidades. Se a perda esperada é minimizada, para qual valor de X deve-se decidir que f_0 é correta?

Exercício 21: Suponha que um certo processo industrial pode estar em controle ou fora de controle; e que em algum tempo especificado a priori a probabilidade do processo estar em controle é 0.9 e fora de controle 0.1. Uma única observação X é retirada no final do processo, e deve ser decidido imediatamente se o processo está ou não em controle. Se o processo está em controle, então $X \sim N(50, 1)$. Se o processo está fora de controle, então $X \sim N(52, 1)$. Se a decisão é que o processo está fora de controle quando de fato está em controle teremos um custo de 1000R\$ por ter parado o processo. Se a decisão é que o processo está em controle quando de fato está fora de controle teremos um custo de 18000R\$ por continuar o processo. Se a decisão é tomada, então não há custos. E de interesse encontrar o teste que a perda esperada seja minimizada. Para qual valor de X deve-se decidir se o processo está em controle ou fora de controle?

Exercício 22: Suponha que a proporção p de itens defeituosos em um lote é desconhecido, e que o interesse é testar as seguintes hipóteses simples: $H : p = 0.3 \times A : p = 0.4$.

Suponha que a priori a probabilidade de $p = 0.3$ é $1/4$ e que $p = 0.4$ é $3/4$; e também que a perda de escolher errado é 1 unidade e de escolher certo é 0. Suponha que uma amostra aleatória de n itens são extraídas do lote. Mostre que o teste de Bayes rejeita H se e somente se a proporção de itens defeituosos na amostra é maior que

$$\frac{\log \frac{7}{6} + \log \frac{1}{3}}{\log \frac{14}{9}}.$$

(23-28. Morris DeGroot - Cap 8.4)

Exercício 23: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição que pertence a família exponencial, e que a f.d.p. ou f.d. é $f(x|\theta)$. Supona também que $c(\theta)$ é uma função crescente em θ . Mostre que a f.d.p. ou f.d conjunta de X_1, \dots, X_n tem razão de verossimilhanças monótona na estatística $\sum_{i=1}^n d(X_i)$.

Exercício 24.0: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição $N(0, \sigma^2)$, com variância desconhecida, e suponha que o interesse é testar as seguintes hipóteses: $H : \sigma^2 \leq 2 \times A : \sigma^2 > 2$.

Mostre que existe um teste UMP das hipóteses para qualquer nível de significância α .

Exercício 24: Mostre que o teste UMP do exercício anterior rejeita H quando $\sum_{i=1}^n x_i^2 \geq c$, e determine o valor de c quando $n = 10$ e $\alpha = 0.05$.

Exercício 25: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição Bernoulli com p desconhecido, e suponha que o interesse é testar as seguintes hipóteses: $H : p \leq \frac{1}{2} \times A : p > \frac{1}{2}$.

Mostre que se $n = 20$, então existe um teste UMP não aleatorizado com nível de significância $\alpha = 0.0577$ e $\alpha = 0.0207$.

Exercício 26: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição Poisson com média λ desconhecida, e suponha que o interesse é testar as seguintes hipóteses: $H : \lambda \leq 1 \times A : \lambda > 1$.

Mostre que se $n = 10$, então existe um teste UMP não aleatorizado com nível de significância $\alpha = 0.0143$.

Exercício 27: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição Poisson com média λ desconhecida, e suponha que o interesse é testar as seguintes hipóteses: $H : \lambda \geq 1 \times A : \lambda < 1$.

Suponha também que $n = 10$. Com nível de significância no intervalo $(0, 0.03)$ existe um teste UMP não aleatorizado?

Exercício 28: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição exponencial com parâmetro β desconhecido, e suponha que o interesse é testar as seguintes hipóteses: $H : \beta \geq \frac{1}{2} \times A : \beta < \frac{1}{2}$.

Mostre que para qualquer nível de significância α , existe um teste *UMP* que rejeita H quando $\bar{X}_n > c$.

(29-35. Morris DeGroot - Cap 8.10)

Exercício 29: Suponha que uma única observação X é retirada da seguinte f.d.p.

$$f(x|\theta) = \begin{cases} 2(1-\theta)x + \theta & , \quad 0 \leq x \leq 1, \\ 0 & , \quad \text{c.c}, \end{cases}$$

em que o valor $\theta \in [0, 2]$ é desconhecido. Suponha também que o interesse é testar as seguintes hipóteses: $H : \theta = 2 \times A : \theta = 0$.

Determine o teste δ que $\alpha(\delta) + 2\beta(\delta)$ é minimizado, e calcule o valor mínimo.

Exercício 30: Considere as condições do exercício anterior, e suponha que o valor de α é dado. Determine o teste δ que $\beta(\delta)$ vai ser minimizado, e calcule o valor mínimo.

Exercício 31: Considere as condições do exercício 29, mas agora suponha que as seguintes hipóteses serão testadas:

$H : \theta \geq 1 \times A : \theta < 1$.

a) Determine a função poder do teste δ que rejeita H quando $X > 9$.

b) Quê é o tamanho do teste δ ?

Exercício 32: Suponha que uma caixa contenha um grande número de chips de diferentes cores, vermelho, marrom, e azul; e que o interesse é testar a hipótese nula H que a proporção das três cores são iguais versus a hipótese alternativa A que a proporção das três cores são diferentes. Suponha que três chips são retirados aleatoriamente da caixa, e que H é rejeitada se e somente se pelo menos dois chips tem a mesma cor.

a) Determine o tamanho do teste.

b) Determine o poder do teste se $1/7$ dos chips são vermelhos, $2/7$ são marrons, e $4/7$ são azuis.

Exercício 33: Suponha que uma única observação X é retirada da distribuição P , e que as seguintes hipóteses simples serão testadas: $H : P$ é uma distribuição uniforme no intervalo $(0, 1) \times A : P$ é uma distribuição $N(0, 1)$.

Determine o teste mais poderoso de tamanho 0.01 e calcule o poder do teste quando A é verdade.

Exercício 34: Suponha que oito observações X_1, \dots, X_8 são retiradas aleatoriamente de uma distribuição com a seguinte f.d.p.:

$$f(x|\theta) = \begin{cases} \theta x^{\theta-1} & , \quad 0 < x < 1, \\ 0 & , \quad \text{c.c.} \end{cases}$$

Suponha também que θ é desconhecido, e que o interesse é testar as seguintes hipóteses: $H : \theta \leq 1 \times A : \theta > 1$.

Mostre que o teste UMP de nível de significância $\alpha = 0.05$ rejeita H se $\sum_{i=1}^8 \log X_i \geq -3.981$.

Exercício 35: Suponha que X_1, \dots, X_n forma uma amostra aleatória de uma distribuição χ^2 com graus de liberdade $\theta \in \mathbb{N}$ é desconhecido, e que o interesse é testar as seguinte hipóteses com nível de significância α : $H : \theta \leq 8 \times A : \theta \geq 9$.

Mostre que existe um teste UMP , e que esse teste rejeita H se $\sum_{i=1}^n \log X_i \geq k$.

Exercício 36: $H_0 : \theta = \theta_0 \times H_1 : \theta = \theta_1$

x	1	2	3	4	5	6	7
$f(x \theta_0)$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{5}{100}$
$f(x \theta_1)$	$\frac{6}{100}$	$\frac{5}{100}$	$\frac{4}{100}$	$\frac{1}{100}$	$\frac{2}{100}$	$\frac{80}{100}$	$\frac{2}{100}$

Determine o teste MP para H_0 versus H_1 de nível 4%, ψ^* . Calcule $\pi_{\psi^*}(\theta_1)$. Calcule a correspondente probabilidade de erro de tipo II .

Exercício 37: $X|\theta \sim Binomial\left(\theta, \frac{1}{2}\right)$. Obtenha o teste MP de nível $\alpha = \frac{9}{256}$ para testar $H_0 : \theta = 8$ versus $H_1 : \theta = 10$, ψ^* .

Calcule $\pi_{\psi^*}(10)$.

Exercício 38: $X|\theta \sim Hipergeométrica(\theta, 5, 3)$. Obtenha o teste MP para testar $H_0 : \theta = 5$ versus $H_1 : \theta = 8$, ψ^* . Avalie ψ^* em função de $\alpha \in [0, 1]$. Obtenha ψ^* .

Exercício 39: (X_1, X_2) vetor aleatório que, dado $\theta = (\theta_1, \theta_2)$, possui densidade dada por

$$f(x_1, x_2|\theta) = \frac{\Gamma(\theta_1 + \theta_2 + 1)}{\Gamma(\theta_1)\Gamma(\theta_2)\Gamma(1)} x_1^{\theta_1-1} x_2^{\theta_2-1} \mathbb{I}(x_1, x_2), \text{ em que } A = \{(a, b) \in \mathbb{R}_+^2 : a + b \leq 1\}.$$

Obtenha o teste MP de nível $\alpha \in (0, 1)$, para $H_0 : \theta = (2, 1)$ versus $H_1 : \theta = (1, 2)$. Determine a probabilidade de erro tipo II de tal teste. Para $\alpha = \frac{9}{100}$, qual seria sua decisão se $X_1 = 0.24$ e $X_2 = 0.52$?

Exercício 40: $X|\theta$ com densidade dada por $f(x|\theta) = \frac{e^{(x-\theta)}}{1+e^{(x-\theta)}}, x \in \mathbb{R}, \theta \in \mathbb{R}$. Determine o teste UMP de nível $\alpha \in (0, 1)$, para $H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$.

Exercício 41: $X|\theta \sim Beta(\theta, \theta), \theta > 0$, Obtenha o teste UMP de nível $\alpha \in (0, 1)$, para as hipóteses $H_0 : \theta \leq 1$ versus $H_1 : \theta > 1$.

Exercício 42: $X|\theta$ com densidade dada por

$$f(x|\theta) = \begin{cases} \frac{1}{2\theta} & , \quad 0 < x \leq \theta \\ \frac{\theta}{2x^2} & , \quad x > \theta \end{cases}, \theta > 0.$$

Obtenha o teste UMP de nível $\alpha \in (0, 1/2)$, para as hipóteses $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0, \theta_0 \in \mathbb{R}_+$.

Exercício 43: $\mathfrak{X} = \{1, 2, 3, 4\}$. $\Theta = \{1, 2, 3\}$. $\Psi = \{\psi : \mathfrak{X} \rightarrow \{0, 1\}\}$. $H_0 = \theta \leq 2 \times H_1 : \theta = 3$.

	1	2	3	4
$\mathbb{P}(x \theta_1)$	$\frac{2}{13}$	$\frac{4}{13}$	$\frac{3}{13}$	$\frac{4}{13}$
$\mathbb{P}(x \theta_2)$	$\frac{4}{13}$	$\frac{2}{13}$	$\frac{1}{13}$	$\frac{6}{13}$
$\mathbb{P}(x \theta_3)$	$\frac{4}{13}$	$\frac{3}{13}$	$\frac{2}{13}$	$\frac{4}{13}$

a) Para $\alpha = \frac{5}{13}$, especifique $\Psi_\alpha = \{\psi \in \Psi : \alpha_\psi^* \leq \alpha\}$.

b) Obtenha o teste MP de nível $\alpha = \frac{5}{13}$ para H_0 versus H_1 .

c) Refaca a) e b) com $\alpha = \frac{6}{13}$.

Exercício 44: $X|\theta$ com densidade dada por $f(x|\theta) = (\theta + 1)2^\theta (\min\{x, 1-x\})^\theta \mathbb{I}_{[0,1]}(x)$. Construa o teste UMP de nível $\alpha \in (0, 1/2)$, para testar $H_0 : \theta \geq 1$ versus $H_1 : \theta < 1$. Para $\alpha = \frac{9}{100}$, qual é a sua decisão a partir da observação $X = 0.9$.