MAE 5753 - TEORIA DA DECISÃO

2º semestre/2015 Luís Gustavo Esteves 108 – B

PROGRAMA

1) Introdução: elementos de um problema de decisão

2) Probabilidade e utilidade: construção (coerência)

3) Maximização de utilidade esperada

4) Formas normal e extensiva de um problema de decisão (resultados importantes)

5) Exemplos em Inferencia Estatística

6)* Tópicos adicionais: Teoria da Decisão Coletiva, aplicações em Controle de Qualidade, testes simultâneos e decisões seqüenciais.

Const. prob.
Const. utilidade
Formas normal e extensiva

BIBLIOGRAFIA

* / Digit.

Livro Base → 1) DeGroot, M. H., (2004). Optimal Statistical Decisions. Wiley-Interscience. Cap. 6, 7 e 8

2) French, S., Insua, D. R., (2010). Statistical Decision Theory: Kendall´s Library of Statistics 9. Wiley. (Ñ mto detalhado - muitas referências) Ñ encont. digit.

* 3) Schervish, M. J. (1996). Theory of Statistics. Springer Series in Statistics. Mais sofisticado do ponto de vista matemático. (Mtos ex. e exemplos mais elaborados)

-> 4) Berger, J. O. (1985). Statistical Decision Theory: Foundations, Concepts and Methods. Springer-Verlag New York. Mais det. que 2, mas com menos referências.
Ñ enc. digit. Tem outro!

Dig
* 5) Parmigiani, G., Inoue, L. (2009). Decision Theory: Principles and Approaches. Wiley. Tbm um pouco de tudo → mais aplicações.

Digit. epub.
* 6) Lindley, D.V. (1991). Making Decisions. 2a edição. Wiley.
Teoria da Decisão: interpretações. Tudo com boli e urnas.

7) Lindley, D.V. (2006). Understanding Uncertainty. Wiley-Interscience.
Só sobre probabilidade. como est. prob., teo. da representação, ...

Simplicidade matemática muito grande.

* 8) Kadane, J. B. (2011) Principles of Uncertainty. Chapman and Hall/CRC.
Digit.

Textos mais complicados em questões filosóficas e matemáticos.

Dig.

* * 9) Savage, L. J., (1972). The foundations of Statistics. Dover Publications; 2 Revised edition.

* * 10) De Finetti, B. (1990) Theory of Probability, Volume 1: A Critical Introductory Treatment (Wiley Classics Library) Wiley. Digit.

* 11) De Finetti, B. (1974) Theory of Probability. A critical Introductory Treatment. Volume 2. (Probability & Mathematical Statistics) Wiley. ~~[illegible]~~ Ñ econti dig.

* 12) Blackwell, D., Girshick, M. A., (1974). Theory of Games and Statistical Decisions (Dover Books on Mathematics) Dover Publications. (Nomon freidig)

Pasta 04. Material interessante, listas de exercícios, ...

Teoria da Decisão

Aula 01 10/08/2015

- Como chegar à um procedimento ótimo de decisão (umas 4 semanas)
- Rebaixar paradigma de maximização de utilidade esperada
- Estudar três problemas (   ) pela teoria da decisão.
- Talvez cont. qualidade, teor. da dec. coletiva

Contexto:

Você deve escolher uma alternativa (ação) dentro de um conjunto de decisões disponíveis $\mathcal{D}$.

Exemplos:

1. ~~Fazer~~ Cursar Teoria da Decisão
2. Escolher as disciplinas do semestre
3. Decidir se vai ao jogo em Florianópolis
4. $\theta = (\mu, \sigma^2)$ Estimar (exibir uma estimativa para θ)
5. Testar hipóteses

$$H: \mu = 0$$
$$H': \mu = 0, \sigma^2 = 1$$
$$H'': \mu > 1$$

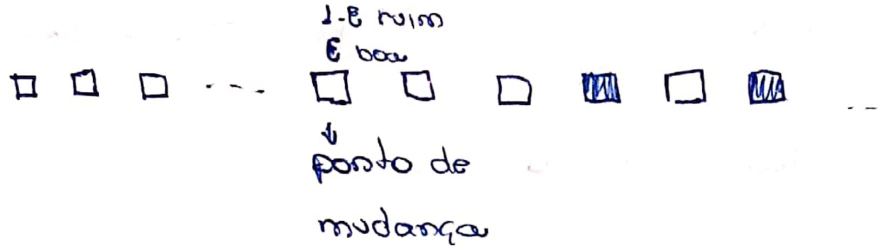
6. Estimar (por intervalo) μ

(A, B), A e B dependem dos dados

Seria desejável $B - A$ pequeno e A e B bem conectados aos dados

7. Controle de qualidade

- Máquina que produz peças, uma a uma;
- No início do proc. de prod., a máq. está bem regulada;
- Com o passar do tempo, em algum momento (pto de mudança), essa máquina começa a alternar entre peças boas e ruins.



No fim de 90, Taguchi propôs um ~~método~~ processo de monitoramento online.

- Tomo $m \in \mathbb{N}$ e averiguo as peças de m em m. Quando encontrar algo def., paro o processo.

intervalo entre inspeções

Problema: Qual deve ser m?

- Custos: inspeção, prod. defeituosa, manutenção

C_i $\quad C_d \quad C_r$

Até 50, apenas as Cartas de Controle. Taguchi foi original ao considerar custos operacionais.

8. Escolher a composição dos investimentos

Aspectos Importantes:

(1) Potenciais aplicações

(2) Aspectos matemáticos e relação com outras áreas (Análise convexa, Análise, Teoria dos Jogos, Pesquisa Operacional).

Teoria dos Conj., etc.)

(3) Relação com inferência estatística
(principalmente Inferência Bayesiana)

A incerteza, em geral, aparece em um problema de decisão.

Probabilidade → Representação numérica de incerteza (int. subjetiva)

Do ponto de vista mat., $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$ é uma prob. se

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

ii. $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$, se A_n's disj. 2 a 2.

Exemplo 1:

$\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$, $\omega_0 \in \Omega$; $\mathbb{P}(A) = \mathbb{I}_A(\omega_0)$ é uma med. de prob. (medida de prob. degenerada em ω_0).

$\mathbb{P}(\Omega) = \mathbb{I}_\Omega(\omega_0) = 1$, uma vez que $\omega_0 \in \Omega$; A_n's disj. $\mathbb{P}(\bigcup A_n) = \mathbb{I}_{\bigcup A_n}(\omega_0) = \sum_{n \geq 1} \mathbb{I}_{A_n}(\omega_0) = 1$

Exemplo 2:

$\mathbb{P}_1, \mathbb{P}_2 : \mathcal{F} \to [0,1]$ prob.; $\alpha \in [0,1]$.

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

$A \longmapsto \mathbb{P}'(A) = \alpha \mathbb{P}_1(A) + (1-\alpha)\mathbb{P}_2(A)$ (medida mista)

mistura de $\mathbb{P}_1$ e $\mathbb{P}_2$.

$$\mathbb{P}'\left(\bigcup_{n \geq 1}^{\infty} A_n\right) = \alpha \mathbb{P}_1\left(\bigcup_{n=1}^{\infty} A_n\right) + (1-\alpha)\mathbb{P}_2\left(\bigcup_{n \geq 1}^{\infty} A_n\right) \overset{*}{=}$$

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

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

$$= \sum_{n \geq 1} \mathbb{P}'(A_n)$$

* Uma vez que $\mathbb{P}(\bigcup A_n) \leq 1$, segue que $\sum \mathbb{P}(A_n)$ é conv. Daí, $\alpha \sum \mathbb{P}(A_n) = \sum \alpha \mathbb{P}(A_n)$

Corint pessimista.

V : $\quad V \precsim E$ (mais fé no empate do que na vitória)

E : $\quad V \precsim D$ (" " na derrota " " " ")

D. $\quad V \sim E$ (V é equivalente à E)

$\quad V \prec E$ (acredito mais em V do que em E)

Algumas suposições. $V \prec E$, então $V \cup D \prec E \cup D$

$\Downarrow$ Teorema.

$\exists\, \mathbb{P}$ t.q. $A \precsim B \Leftrightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$

Teoria da Decisão, Aula 01, Exercícios

1. Mostre que cada um das medidas abaixo é uma medida de probabilidade sob um dado espaço mensurável $(\Omega, \mathcal{F})$.

a) $\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$, $\omega_0 \in \Omega$

$A \mapsto \mathbb{P}(A) = \mathbb{I}_{\omega_0}(A)$

b) $\mathbb{P}_1, \mathbb{P}_2: \mathcal{F} \to [0,1]$ prob. e $\alpha \in [0,1]$. Defina

$$\mathbb{P}': \mathcal{F} \to \mathbb{R}_+$$
$$A \mapsto \mathbb{P}'(A) = \alpha \mathbb{P}_1(A) + (1-\alpha)\mathbb{P}_2(A)$$

Por definição, dado um espaço $(\Omega, \mathcal{F}, \mu)$, dizemos que μ é uma medida de probabilidade se

1. $\mu(\Omega) = 1$;
2. Dada uma seq. $(A_n)_{n \geq 1}$ de eventos disjuntos dois a dois de $\mathcal{F}$,
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Logo, basta verificar que $\mathbb{P}$ e $\mathbb{P}'$ satisfazem (1) e (2).

a) 1) $\mathbb{P}(\Omega) = \mathbb{I}_{\Omega}(\omega_0) = 1$, uma vez que $\omega_0 \in \Omega$

2) Dados $A_n \in \mathcal{F}, n \geq 1$ com $A_n \cap A_m = \emptyset$ p/ $n \neq m$,

Ou

$$\omega_0 \notin \bigcup_{n \geq 1} A_n \Rightarrow \mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = \mathbb{I}_{\bigcup_{n \geq 1} A_n}(\omega_0) = 0 \Rightarrow \omega_0 \in \bigcap_{n \geq 1} A_n^c$$

$$\Rightarrow \omega_0 \in A_n^c \ \forall n \geq 1 \Rightarrow \omega_0 \notin A_n \ \forall n \Rightarrow \mathbb{I}_{A_n}(\omega_0) = 0, \forall n \geq 1$$

$$\Rightarrow \sum_{n \geq 1} \mathbb{P}(A_n) = \sum_{n \geq 1} \mathbb{I}_{A_n}(\omega_0) = 0;$$

Ou $\omega_0 \in \bigcup_{n \geq 1} A_n \Rightarrow \mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = 1.$

* Além disso, $\exists\, n_0$ t.q. $\omega_0 \in A_{n_0}$ e $\omega_0 \notin A_n\ \forall n \neq n_0$. Logo

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

Logo, $\mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mathbb{P}(A_n)$.

* Sabemos que exi. ao menos um n t.q. $\omega_0 \in A_n$. É único, pois se $\omega_0 \in A_{n_0}$ e $\omega_0 \in A_{n_1}$ com $n_0 \neq n_1$, $\omega \in A_{n_0} \cap A_{n_1} = \emptyset$. ($\Rightarrow\Leftarrow$)

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

$= \alpha \times 1 + (1-\alpha) \times 1 = \alpha + 1 - \alpha = 1.$ ✓

- Dada $(A_n)_{n \geq 1}$ com $A_n \cap A_m = \emptyset\ \forall n \neq m$, $n, m \in \mathbb{N}$,

$$\mathbb{P}'\left(\bigcup_{n \geq 1} A_n\right) = \alpha \mathbb{P}_1\left(\bigcup_{n \geq 1} A_n\right) + (1-\alpha)\mathbb{P}_2\left(\bigcup_{n \geq 1} A_n\right)$$
$$\overset{*}{=} \alpha \sum_{n \geq 1} \mathbb{P}_1(A_n) + (1-\alpha) \sum_{n \geq 1} \mathbb{P}_2(A_n)$$
$$= \sum_{n \geq 1} \alpha \mathbb{P}_1(A_n) + (1-\alpha)\mathbb{P}_2(A_n)$$
$$= \sum_{n \geq 1} \mathbb{P}'(A_n) \quad ✓$$

* Como $\mathbb{P}_i\left(\bigcup_{n \geq 1} A_n\right) \leq 1$, $\sum_{n \geq 1} \mathbb{P}_i(A_n)$ é convergente, o que faz com que $\underset{\downarrow}{\alpha} \sum_{n \geq 1} \mathbb{P}_i(A_n) = \sum_{n \geq 1} \alpha \mathbb{P}_i(A_n)$, $i = 1, 2$.

pode ser 0 ($0 \times \infty$ dá probl.)

Teoria da Decisão

12/08/2015 Aula 02

"Construção da prob. a partir de ideias primitivas"

$\Omega \rightarrow$ conjunto não-vazio representando os possíveis realizações

$\mathcal{F} \rightarrow$ classe de subconjuntos de Ω que satisfaz

(1) $\phi \in \mathcal{F}$

(2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

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

$\Big\}$ σ-álgebra de subconjuntos de Ω.

(Roussas, 197?)

(Barry James, 198?)

Vamos definir $\precsim$ em $\mathcal{F} \times \mathcal{F}$.

$A \precsim B \Rightarrow$ "Acredito em B tanto quanto ou mais que A"

Ex: Corinthians

$\Omega = \{v, e, d\}$

$A = \{v\}$ $\{v\} \precsim \{e\}$

$B = \{e\}$ $\{e\} \precsim \{d\}$

$C = \{d\}$

Nosso objetivo é especificar condições sobre $\precsim$ de modo a representá-la por uma única medida de probabilidade.

Def: Dizemos que $f: \mathcal{F} \rightarrow \mathbb{R}$ "representa" ("coincide") com $\precsim$ se
$\forall A, B \in \mathcal{F}$, $A \precsim B \Leftrightarrow f(A) \leq f(B)$.

"Há várias construções desse tipo, inclusive uma em que se constrói simultaneamente prob. e utilidade, de onde tbm decorre o paradigma da otimiz. da utilidade esperada".

- $A \sim B \Leftrightarrow A \precsim B$ e $B \precsim A$
- $B \prec A$ representa a negação de $A \precsim B$

Suposições

SP1 $\forall A, B$, $A \precsim B$ ou $B \precsim A$. (comparabilidade ou rel. completa)

SP2 A_1, A_2, B_1, B_2 com $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$

$$A_i \precsim B_i,\ i = 1,2 \Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$$

Se adicionalmente, $A_i \prec B_i$ para algum $i \Rightarrow A_1 \cup A_2 \prec B_1 \cup B_2$

Ex: Corinthians.

$$\underbrace{\{(1,0)\}}^{A_1} \leq \underbrace{\{(0,1)\}}^{A_2}$$

$$\underbrace{\bigcup_{n=2}^{\infty} \{(n,0)\}}_{B_1} \precsim \underbrace{\bigcup_{n=2}^{\infty} \{(0,n)\}}_{B_2} \Rightarrow \bigcup_{n=1}^{\infty} \{(n,0)\} \precsim \bigcup_{n=1}^{\infty} \{(0,n)\}$$

Consequência:

$A_1, \dots, A_n, B_1, \dots, B_n$, $A_i \cap A_j = B_i \cap B_j = \emptyset$, $i \neq j$.

$$A_i \precsim B_i \Rightarrow \bigcup_{i=1}^{n} A_i \precsim \bigcup_{i=1}^{n} B_i$$

Lema. $A, B, C \in \mathcal{F}$ tais que $A \cap C = B \cap C = \emptyset$.

$$A \precsim B \Leftrightarrow A \cup C \precsim B \cup C$$

Dem:

$\Rightarrow$ Na sup. SP2, tomamos

$A_1 = A$, $B_1 = B$, $A_2 = B_2 = C$ $\Rightarrow A \cup C \precsim B \cup C$

$A \precsim B$

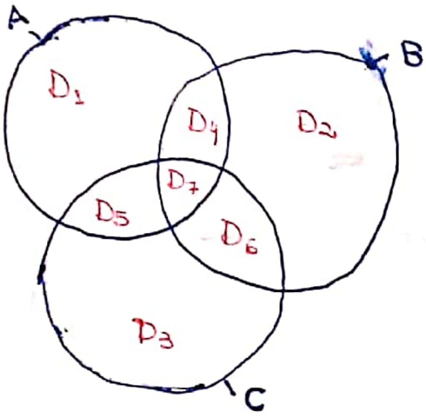
$C \precsim C$ (por SP1)

Suponha agora

$$B \prec A \overset{SP2}{\Rightarrow} B \cup C \prec A \cup C$$

Logo, segue o resultado ∎

Teorema. $\precsim$ é transitiva.
($A, B, C \in \mathcal{F}$ são tais que $A \precsim B$ e $B \precsim C$ então $A \precsim C$)



$$A \precsim B \Leftrightarrow D_1 \cup D_4 \cup D_5 \cup D_7 \precsim D_2 \cup D_4 \cup D_6 \cup D_7$$
$$\Leftrightarrow (D_1 \cup D_5) \cup (D_4 \cup D_7) \precsim (D_2 \cup D_6) \cup (D_4 \cup D_7) \overset{\text{Lema}}{\Leftrightarrow}$$
$$\Leftrightarrow (D_1 \cup D_5) \precsim (D_2 \cup D_6) \quad (\text{I})$$

$$B \precsim C \Leftrightarrow D_2 \cup D_4 \cup D_6 \cup D_7 \precsim D_3 \cup D_5 \cup D_6 \cup D_7$$
$$\overset{\text{lema}}{\Leftrightarrow} D_2 \cup D_4 \precsim D_3 \cup D_5 \quad (\text{II})$$

Pela SP2, segue ~~que~~ de I e II que

$$D_1 \cup \underline{D_2} \cup D_4 \cup D_5 \precsim \underline{D_2} \cup D_3 \cup D_5 \cup D_6 \overset{\text{Lema}}{\Leftrightarrow}$$
$$\Leftrightarrow D_1 \cup D_4 \cup D_5 \precsim D_3 \cup D_5 \cup D_6 \Leftrightarrow$$
$$\Leftrightarrow D_1 \cup D_4 \cup D_5 \cup D_7 \precsim D_3 \cup D_5 \cup D_6 \cup D_7 \Leftrightarrow$$

$$A \precsim C.$$

Noto que com as cond. até aqui, $\precsim$ admite uma representação.
no caso finito

Basta def. para cada $A \in \mathcal{F}$,

$$f(A) = |\{C \in \mathcal{F} : C \precsim A\}|$$

($A \precsim B \Leftrightarrow f(A) \leq f(B)$), mas f não é uma prob.

Resultado:

$A \precsim B \Leftrightarrow B^c \precsim A^c$

$A \precsim B \Leftrightarrow (A\cap B)\cup(A\cap B^c) \precsim (A\cap B)\cup(A^c\cap B) \Leftrightarrow$

$(A\cap B^c) \precsim A^c \cap B \Leftrightarrow$

$(A\cap B^c)\cup(A^c\cap B^c) \precsim (A^c\cap B)\cup(A^c\cap B^c) \Leftrightarrow$

$B^c \precsim A^c$ $\square$

p/ evitar caso constante

SP3. $\forall A \in \mathcal{F}, \emptyset \precsim A$. ($\emptyset \prec \Omega$)

Consequências:

1. $\forall A \in \mathcal{F}, A \precsim \Omega$.

Dem: De SP3, $\emptyset \precsim A^c$ ∴ ~~[illegible]~~ Como $\emptyset \cap A = \emptyset$ e $A^c \cap A = \emptyset$, segue do lema que $\emptyset \cup A \precsim A^c \cup A \Leftrightarrow A \precsim \Omega$ $\square$

2. $A, B \in \mathcal{F}$ e $A \subseteq B$, $A \precsim B$.

Dem: Como $A \subseteq B$, podemos escrever

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

Da sup. SP3, temos

$$\emptyset \precsim A^c\cap B \overset{\text{Lema}}{\Leftrightarrow} \emptyset \cup A \precsim (A^c\cap B)\cup A \Leftrightarrow A \precsim B.$$

SP4. $(A_n)_{n\geq 1}$ sequência de elementos de $\mathcal{F}$ com $A_n \supseteq A_{n+1}$.
Seja $B \in \mathcal{F}$

$$B \precsim A_n \Rightarrow B \precsim \bigcap_{n=1}^{\infty} A_n$$

Consequência:

$(A_n)_{n\geq 1}$ seq. com $A_n \subseteq A_{n+1}$, $A_n \precsim B$, $\forall n \geq 1$. Logo

$$\bigcup_{n=1}^{\infty} A_n \precsim B.$$

Dem: É fácil ver que $(A_n^c)_{n\geq 1}$ é tal que $A_n^c \supseteq A_{n+1}^c$. Além disso, $A_n \precsim B$, $\forall n \geq 1$, segue do resultado anterior que $B^c \precsim A_n^c$, $\forall n \geq 1$. Da SP4,

$$B^c \precsim \bigcap_{n\geq 1} A_n^c$$

Usando o último resultado por uma seg. vez, segue pelas Leis de De Morgan que $\left(\bigcap_{n\geq 1} A_n^c\right)^c \precsim (B^c)^c \Leftrightarrow$

$$\bigcup_{n\geq 1} A_n \precsim B$$

Resultado.

1. $(A_n)_{n\geq 1}$ e $(B_n)_{n\geq 1}$ sequências em $\mathcal{F}$ tais que $A_i \cap A_j = B_i \cap B_j = \emptyset$, $i \neq j$. Se $A_n \precsim B_n$, $\forall n \geq 1 \Rightarrow \bigcup_{n=1}^{\infty} A_n \precsim \bigcup_{n=1}^{\infty} B_n$

2. Se, adicionalmente, $\exists\, n_0 \in \mathbb{N}$ t.q.
$$A_{n_0} \prec B_{n_0},$$
então $\bigcup_{n=1}^{\infty} A_n \prec \bigcup_{n=1}^{\infty} B_n$.

Dem.

(1) Da extensão de SP2, segue que, $\forall n \geq 1$,

$$\bigcup_{i=1}^{n} A_i \precsim \bigcup_{i=1}^{n} B_i \precsim \bigcup_{n=1}^{\infty} B_n, \text{ pois } \forall n \geq 1 \ \bigcup_{i=1}^{n} B_i \subseteq \bigcup_{n=1}^{\infty} B_n.$$

Logo, pela transitividade, c/

$$\bigcup_{i=1}^{n} A_i \precsim \overbrace{\bigcup_{n=1}^{\infty} B_i}^{B}, \quad \forall n \geq 1.$$

Da conseq. da SP4, segue que

$$\bigcup_{n\geq 1} A_n \precsim \bigcup_{n\geq 1} B_n$$

(2) Suponhamos $A_{n_0} \prec B_{n_0}$, $n_0 \in \mathbb{N}^*$.
Da suposição SP2

$$\bigcup_{i=1}^{n_0-1} A_i \precsim \bigcup_{i=1}^{n_0-1} B_i \qquad \Rightarrow \bigcup_{i=1}^{n_0} A_i \prec \bigcup_{i=1}^{n_0} B_i$$

$$A_n \prec B_{n_0}$$

Da prim. parte do resultado, vale que $\bigcup_{n=n_0+1}^{\infty} A_n \precsim \bigcup_{n=n_0+1}^{\infty} B_n$. Então de SP2 novamente,

$$\left(\bigcup_{i=1}^{n_0} A_i\right) \cup \left(\bigcup_{i=n_0+1}^{\infty} A_i\right) \prec \left(\bigcup_{i=1}^{n_0} B_i\right) \cup \left(\bigcup_{i=n_0+1}^{\infty} B_i\right), \text{ i.e.},$$

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

Exemplos:

1: $\Omega \neq \emptyset$, $\omega_0 \in \Omega$, $\mathcal{F}$ σ-álg. de subconjuntos de Ω.
Para $A, B \in \mathcal{F}$

$$A \precsim B \Leftrightarrow \begin{cases} \omega_0 \in B, \text{ ou} \\ \omega_0 \notin B \text{ e } \omega_0 \notin A \end{cases}$$ (Atende SP1, 2, 3, 4)

2: $\Omega = \mathbb{N}$
$\mathcal{F} = \mathcal{P}(\mathbb{N})$
A e $B \in \mathcal{F}$

$$A \precsim B \Leftrightarrow \begin{cases} B \text{ infinito} \\ B \text{ finito}, A \text{ finito}, |A| \leq |B| \end{cases}$$

Teoria da Decisão, Aula 02, Exercícios

1. Construção da prob. a partir de ideias primitivos
2. Estabeleço Ω como cjto dos result. possíveis e $\mathcal{F}$ como uma coleção de eventos de Ω satisfazendo 3 condições (as que def. uma σ-álgebra).
3. A partir daí, defino uma relação $\lesssim$ em $\mathcal{F} \times \mathcal{F}$, tal que
$A \lesssim B$ indica maior crença em A do que em B.

4. Objetivo é especificar condições sobre $\lesssim$ de modo a representá-la por uma única medida de prob. Para isso, adianta-se a definição

Def: Dizemos que $f: \mathcal{F} \to \mathbb{R}$ "representa" (coincide) com $\lesssim$ se
$\forall A, B \in \mathcal{F}$, $A \lesssim B \Leftrightarrow f(A) \leq f(B)$

5. Suposições e Consequências

Antes, mais algumas notações

$$A \sim B \Leftrightarrow A \lesssim B \text{ e } B \lesssim A$$

$B < A$ repres. a negação de $A \lesssim B$.

SP1. $\forall A, B$, $A \lesssim B$ ou $B \lesssim A$ (comparabilidade ou rel. completa)

SP2. A_1, A_2, B_1, B_2 com $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$,

$$A_i \lesssim B_i,\ i = 1, 2 \Rightarrow A_1 \cup A_2 \lesssim B_1 \cup B_2$$

Se, adicionalmente, $A_i < B_i$ p/ algum i, $A_1 \cup A_2 < B_1 \cup B_2$.

Dúvida: Se $A_i \precsim B_i$, $i=1,2$, $A_1 \cup A_2 \precsim B_1 \cup B_2$.

$A_1 \cup A_2$ ocorre se ocorre A_1

A_2

$A_1 \cap A_2$

$B_1 \cup B_2$ ocorre se ocorre B_1

B_2

$B_1 \cap B_2$

$A_1 \precsim B_1$, $A_2 \precsim B_2$, $A_1 \cap A_2$ $B_1 \cap B_2$

Exemplo Lanço um dado. Def. $A \precsim B$ se $|A| \leq |B|$, $\forall A, B \in \mathcal{F}$.

$A_1 = \{1\}$, $A_2 = \{1, 2\}$

$B_1 = \{3, 6\}$, $B_2 = \{1, 4, 5\}$

$A_1 \precsim B_1$; $A_2 \precsim B_2 \Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$

$A_1 \cap A_2 = \{1\}$, $B_1 \cap B_2 = \emptyset$

Conseq. SP2. $A_1, \dots, A_n$; $B_1, \dots, B_n$, $A_i \cap A_j = B_i \cap B_j = \emptyset$, $i \neq j$,

$A_i \precsim B_i \Rightarrow \bigcup_{i=1}^{n} A_i \precsim \bigcup_{i=1}^{n} B_i$

Para $n=1$, vale pois $A_1 \precsim B_1$ e

$\bigcup_{i=1}^{1} A_i = A_1$ e $\bigcup_{i=1}^{1} B_i = B_1$

Para $n=2$, também, pois $A_i \precsim B_i$, $i=1,2$, que garante pela SP2 que

$A_1 \cup A_2 \precsim B_1 \cup B_2$

Supondo válido para $n=k$, segue que $\bigcup_{i=1}^{k} A_i \precsim \bigcup_{i=1}^{k} B_i$. Ademais, sabe-se que $A_{k+1} \precsim B_{k+1}$. Pela SP2, segue

$\bigcup_{i=1}^{k+1} A_i \precsim \bigcup_{i=1}^{k+1} B_i$. Por indução, segue o resultado.

Lema (~~[illegible]~~) $A, B, C \in \mathcal{F}$ t.q. $A \cap C = B \cap C = \emptyset$.

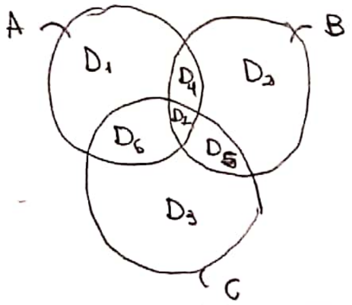
$$A \precsim B \Leftrightarrow A \cup C \precsim B \cup C.$$

Dem: Na SP2, tome $A_1 = A$, $B_1 = B$, $A_2 = B_2 = C$. Daí, segue o resultado, uma vez que por SP1 $C \precsim C$.

Se $B \prec A$, pela SP2, $B \cup C \prec A \cup C$.

Teorema. $\precsim$ é transitiva.

A, B e $C \in \mathcal{F}$ t.q. $A \precsim B$, $B \precsim C$. Então $A \precsim C$.



$D_1 = A \cap B^c \cap C^c$

$D_2 = B \cap A^c \cap C^c$

$D_3 = C \cap A^c \cap B^c$

$D_4 = A \cap B \cap C^c$

$D_5 = B \cap C \cap A^c$

$D_6 = A \cap C \cap B^c$

$D_7 = A \cap B \cap C$

$A \precsim B \Rightarrow (D_1 \cup D_5) \precsim (D_2 \cup D_6)$

$B \precsim C \Rightarrow D_2 \cup D_4 \precsim D_3 \cup D_5$

$\downarrow$ SP2 (adiciono / retiro quantos conjuntos até obter o desejado)

Res: $A \precsim B \Leftrightarrow B^c \precsim A^c$.

SP3. $\forall A \in \mathcal{F}$, $\emptyset \precsim A$ $(\emptyset \prec \Omega)$

Conseq.

1. $\forall A \in \mathcal{F}$, $A \precsim \Omega$. $(\emptyset \overset{SP3}{\precsim} A^c \overset{Lema}{\Rightarrow} A \precsim A^c \cup A \Rightarrow A \precsim \Omega)$
2. $A, B \in \mathcal{F}$, $A \subseteq B \Rightarrow A \precsim B$. $(\emptyset \overset{SP3}{\precsim} A^c \cap B \overset{Lema}{\Rightarrow} A \precsim \overbrace{A \cup (A^c \cap B)}^{B})$

SP4. $(A_n)_{n \geq 1}$ seq. el. de $\mathcal{F}$ com $A_n \supseteq A_{n+1}$. Seja $B \in \mathcal{F}$

$$B \precsim A_n \Rightarrow B \precsim \bigcap_{n=1}^{\infty} A_n$$

Conseq.

$(A_n)_{n\geq 1}$ seq com $A_n \subseteq A_{n+1}$, $A_n \precsim B$ $\forall n \geq 1$. Segue que $\bigcup_{n\geq 1} A_n \precsim B$.

Dem: $(A_n^c)_{n\geq 1}$ é t.q. $A_n^c \supseteq A_{n+1}^c$. Ademais, $A_n \precsim B$, $\forall n \geq 1$. Segue do Resultado ant. que $B^c \precsim A_n^c$ $\forall n \geq 1$. Da SP4,

$$B^c \precsim \bigcap_{n\geq 1} A_n^c$$

Usando o ult. res. mais uma vez, segue $\left(\bigcap_{n\geq 1} A_n^c\right)^c \precsim (B^c)^c \Leftrightarrow \bigcup_{n\geq 1} A_n \precsim B$.

C...

Exemplos:

1.

SP1. $A \precsim B$ ou $B \precsim A$

SP2 $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$ com $A_i \precsim B_i$. $\Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$

SP3. $\emptyset \precsim A$ $(\emptyset \prec \Omega)$

SP4. $(A_n)_{n\geq 1}$ em $\mathcal{F}$ com $A_n \supseteq A_{n+1}$ e $B \in \mathcal{F}$.

$$B \precsim A_n \ \forall n \Rightarrow B \precsim \bigcap_{n=1}^{\infty} A_n$$

SP1. Dados A e $B \in \mathcal{F}$,

Se $w_0 \in B$, $A \precsim B$

Se $w_0 \notin B$, $B \precsim A$

SP2. Dados A_i, B_i, $i = 1, 2$, com $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$ e $A_i \precsim B_i$,

ou $w_0 \in B_1 \cup B_2 \Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$;

ou $w_0 \notin B_1 \cup B_2 \Rightarrow w_0 \notin B_1$ e $w_0 \notin B_2 \Rightarrow w_0 \notin A_1$ e $w_0 \notin A_2$,

uma vez que $A_i \precsim B_i, i = 1, 2 \Rightarrow w_0 \notin A_1 \cup A_2 \Rightarrow$

$A_1 \cup A_2 \precsim B_1 \cup B_2$

SP3. Como $\omega_0 \notin \emptyset$, $\forall A \in \mathcal{F}$, segue da def. da relação que

$$\emptyset \precsim A \qquad (*)$$

pois se $\omega_0 \in A$, vale $\emptyset \precsim A$ e se $\omega_0 \in A$, temos $\omega_0 \in \emptyset$, o que ainda garante $(*)$.

SP4. $(A_n)_{n \geq 1}$ com $A_n \supseteq A_{n+1}$ e $B \in \mathcal{F}$

$$B \precsim A_n \ \forall n \Rightarrow B \precsim \bigcap_{n=1}^{\infty} A_n$$

Ou $\omega_0 \notin B \Rightarrow B \precsim \bigcap_{n=1}^{\infty} A_n$;

Ou $\omega_0 \notin B \Rightarrow \omega_0 \in A_n, \forall n \geq 1 \Rightarrow \omega_0 \in \bigcap_{n\geq 1} A_n \Rightarrow B \precsim \bigcap_{n \geq 1} A_n$.

2. $\Omega = \mathbb{N}$, $\mathcal{F} = \mathcal{P}(\mathbb{N})$, A e $B \in \mathcal{F}$ $\quad A \precsim B \Leftrightarrow \begin{cases} B \text{ inf.} \\ B \text{ fin.}, A \text{ fin.}, |A| \leq |B| \end{cases}$

SP1. Dados $A, B \in \mathcal{P}(\mathbb{N})$, ou

- A, B finitos $\Rightarrow \begin{cases} A \precsim B \text{ se } |A| \leq |B| \\ B \precsim A \text{ se } |A| \geq |B| \end{cases}$

- B infinito $\Rightarrow A \precsim B$
- A infinito $\Rightarrow B \precsim A$

SP2. Dados $A_i \precsim B_i$, $i=1,2$, com $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$,

- $B_1 \cup B_2$ infinito $\Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$ ✓
- $B_1 \cup B_2$ finito $\Rightarrow B_1$ e B_2 são finitos $\overset{\text{hipót}}{\Rightarrow}$ A_1 e A_2 são finitos tais que $|A_1| \leq |B_1|$ e $|A_2| \leq |B_2| \Rightarrow$

$$|A_1| + |A_2| \leq |B_1| + |B_2| \underset{\text{disj.}}{\Rightarrow} |A_1 \cup A_2| \leq |B_1 \cup B_2|$$

$$\Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$$

SP3. Uma vez que $|\emptyset| = 0$, segue que

$$|\emptyset| \leq |A|, \forall A \in \mathcal{F} \Rightarrow \emptyset \precsim A, \forall A \in \mathcal{F}.$$

SP4. $(A_n)_{n\geq 1}$ com $A_n \supseteq A_{n+1}$ e B em $\mathcal{F}$.

$$B \mp A_n \ \forall n \Rightarrow B \mp \bigcap_{n\geq 1} A_n$$

Seja $B = \overline{\mathbb{N}}$ e $A_n = \overline{\mathbb{N}} / \{0, \ldots, n-1\}$. Nessas condições, temos

$$B \mp A_n \ , \ \forall n \geq 1$$

Mas $\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} \{n, \ldots\} \overset{*}{=} \{+\infty\}$ Daí, resulta

$B \not\mp \{+\infty\}$, o que contradiz SP4.

Teoria da Decisão

Aula 03 17/08/15

SP1. $A, B \in \mathcal{F}$, $A \precsim B$ ou $B \precsim A$

SP2. $A_1, A_2, B_1, B_2 \in \mathcal{F}$ t.q. $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$ e $A_i \precsim B_i$, $i = 1, 2$,

$$A_1 \cup A_2 \precsim B_1 \cup B_2$$

SP3. $\emptyset \precsim A$ $(\emptyset \precsim \Omega)$

SP4. $(A_n)_{n \geq 1} \subseteq \mathcal{F}$ com $A_n \supseteq A_{n+1}$, $\forall n \geq 1$ e $B \in \mathcal{F}$ t.q.

$$B \precsim A_n, \forall n \geq 1 \Rightarrow B \precsim \bigcap_{n=1}^{\infty} A_n$$

Ex.

$\Omega \neq \emptyset$, $w_0 \in \Omega$

$$A \precsim B \Leftrightarrow \begin{cases} w_0 \in B, \text{ ou} \\ w_0 \notin B, w_0 \notin B. \end{cases}$$

• SP1.

Se $w_0 \in A \cup B \Rightarrow A \precsim B$ ou $B \precsim A$

$w_0 \notin A \cup B \Rightarrow A \precsim B$ e $B \precsim A$

• SP2

(1) $\exists\, i = 1, 2$ tal que $w_0 \in B_i \Rightarrow w_0 \in B_1 \cup B_2 \Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$

$w_0 \in B_i, i = 1, 2, \Rightarrow w_0 \notin B_1$, $w_0 \notin B_2$, $w_0 \notin A_1$ e $w_0 \notin A_2$

$\Rightarrow w_0 \notin B_1 \cup B_2$ e $w_0 \notin A_1 \cup A_2$

$\Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$

SP3

Se $\omega_0 \in A \Rightarrow \emptyset \precsim A$; e se $\omega_0 \in A$, segue que $\emptyset \precsim A$ uma vez que $\omega_0 \notin \emptyset$. (por const.)

SP4.

- $\omega_0 \in A_n, \forall n \geq 1 \Rightarrow \omega_0 \in \bigcap_{n\geq 1} A_n \Rightarrow B \precsim \bigcap_{n\geq 1} A_n$
- $\exists n_0 \in \mathbb{N}$ t.q. $\omega_0 \notin A_{n_0} \Rightarrow \omega_0 \notin \bigcap_{n\geq 1} A_n$ e $\omega_0 \notin B \Rightarrow B \precsim \bigcap_{n\geq 1} A_n$

Ex.2.

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

$$A \precsim B \Leftrightarrow \begin{cases} B \text{ infinito} \\ B \text{ finito, } A \text{ finito e } |A| \leq |B|. \end{cases}$$

SP1. ✓

(1) Se A ou B é infinito $\Rightarrow A \precsim B$ ou $B \precsim A$.

(2) A e B finitos $\Rightarrow |A| \leq |B|$ ou $|B| \leq |A| \Rightarrow A \precsim B$ ou $B \precsim A$.

SP2. ✓

(1) $B_1 \cup B_2$ infinito $\Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$;

(2) B_1 e B_2 finitos $\Rightarrow \begin{cases} B_1 \cup B_2 \text{ é finito,} \\ A_1 \text{ é finito, } |A_1| \leq |B_1| \\ A_2 \text{ é finito, } |A_2| \leq |B_2| \end{cases} \Rightarrow$

$\Rightarrow B_1 \cup B_2$ é finito, $A_1 \cup A_2$ é finito e $|A_1 \cup A_2| \leq |B_1 \cup B_2|$

$\Rightarrow A_1 \cup A_2 \precsim B_1 \cup B_2$.

SP3. A é infinito $\Rightarrow \emptyset \precsim A$

A é finito $\Rightarrow 0 \leq |A| < \infty \Rightarrow \emptyset \precsim A$.

SP4.

$B = \{2\}$

$A_n = \{n, n+1, n+2, \dots\}$

Claro que $A_n \supseteq A_{n+1}$, $\forall n \geq 1$.

Como A_n é infinito para todo $n \geq 1$, segue que $B \precsim A_n$, $\forall n \geq 1$

Mas $\bigcap_{n \geq 1} A_n = \phi$.

Note que $\bigcap_{n=1}^{\infty} A_n$ é finito, B é finito e $|\bigcap_{n=1}^{\infty} A_n| = 0 < 1 = |B|$

Como $\bigcap_{n=1}^{\infty} A_n < B$, não vale SP4!

Pergunta. Se valem SP1 e SP2, então

$$A \precsim B \Leftrightarrow B^c \precsim A^c, \quad \text{isto é,}$$

$$A \precsim B \Rightarrow B^c \precsim A^c.$$

Tomando $B = \mathbb{N}$ e $A = \{2n+1, n \geq 0\}$, segue

$$B \precsim A \Rightarrow A^c \precsim B^c \Rightarrow \{2n, n \geq 0\} \precsim \phi \ (?)$$

(Se aceito SP1, SP2, devo aceitar SP3?)

SP5

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

(1) $X(\omega) \in [0,1]$, $\forall \omega \in \Omega$

(2) I_1 e I_2 são intervalos subconjuntos de $[0,1]$,

$$\{X \in I_1\} \precsim \{X \in I_2\} \Leftrightarrow \lambda(I_1) \leq \lambda(I_2)$$

Tal transformação X é denominada UNIFORME.

Observação.

(1) $\{X \in [a_1, b_1]\} \precsim \{X \in [a_2, b_2]\} \Rightarrow b_1 - a_1 \leq b_2 - a_2$

(2) $a, b \in [0,1]$ com $a < b$.

$$\{X \in [a,b]\} \sim \{X \in (a,b]\} \sim \{X \in [a,b)\} \sim \{X \in (a,b)\}$$

Teorema. Suponha que $\precsim$ atende SP1 à SP5. Então, para todo $A \in \mathcal{F}$, existe um único $a^* \in [0,1]$ t.q.

$$A \sim \{X \in [0, a^*]\}$$

Dem: Seja $A \in \mathcal{F}$. Definamos $\mathcal{U}(A) = \{a \in [0,1] . A \precsim \{X \in [0,a]\}\}$. Note que $\mathcal{U}(A) \neq \emptyset$, pois $1 \in \mathcal{U}(A)$ uma vez que $A \precsim \Omega = \{X \in [0,1]\}$.

Seja $a^* = \inf \mathcal{U}(A)$. Por um lado, Existe $(a_n)_{n \geq 1}$ tal que $a_n \geq a_{n+1}$, $\forall n \geq 1$ e $a_n \in \mathcal{U}(A)$, $\forall n \geq 1$ e $\lim_{n \to \infty} a_n = a^*$ ($a_n \downarrow a^*$).

Se $a_n \in \mathcal{U}(A)$, então $A \precsim X \in [0, a_n]$.
Logo, $\forall n \geq 1$, $A \precsim X \in [0, a_n]$ e, além disso,

$$\{X \in [0, a_n]\} \supseteq \{X \in [0, a_{n+1}]\}$$

Pela SP4,

$$A \precsim \bigcap_{n=1}^{\infty} \{X \in [0, a_n]\} \Rightarrow$$

$$A \precsim \{X \in \bigcap_{n=1}^{\infty} [0, a_n]\} \text{ , ou ainda}$$

$$A \precsim \{X \in [0, a^*]\}. \qquad (I)$$

Supon. $a^* > 0$

Por outro lado, seja $(b_n)_{n \geq 1}$ seq tal que $b_n \leq b_{n+1}$, $\forall n \geq 1$, $b_n < a^*$ $\forall n \geq 1$

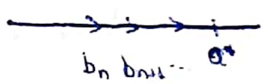
e $\lim_{n\to\infty} b_n = a^*$.

$$\{x\in[0,b_n]\} \top A, \ \forall n\geq 1$$
$$\{x\in[0,b_n]\} \tilde{\top} A, \ \forall n\geq 1$$

$\{x\in[0,b_n]\} \subseteq \{x\in[0,b_{n+1}]\}$.

Da conseq. da SP4 para seq. "crescentes", segue que

$$\bigcup_{n=1}^{\infty} \{x\in[0,b_n]\} \tilde{\top} A \Rightarrow$$



$$\Rightarrow \{x\in \bigcup_{n=1}^{\infty} [0,b_n]\} \tilde{\top} A \Rightarrow$$

$$\Rightarrow \{x\in[0,a^*)\} \tilde{\top} A \underset{\text{Uniformidade}}{\Rightarrow} \{x\in[0,a^*]\} \tilde{\top} A \qquad (\text{II})$$

De (I) e (II), $A \sim \{x\in[0,a^*]\}$.

Unicidade de a^*. Sejam a_1 e a_2 tais que

$$A\sim\{x\in[0,a_1]\} \text{ e } A\sim\{x\in[0,a_2]\} \Rightarrow$$
$$\{x\in[0,a_1]\}\sim\{x\in[0,a_2]\} \Leftrightarrow$$
$$a_1 - 0 \leq a_2 - 0 \text{ e } a_1 - 0 \geq a_2 - 0 \Rightarrow a_1 = a_2. \qquad \square$$

Vamos, finalmente, definir uma transformação

$$P: \mathcal{F} \to \mathbb{R}_+ \text{ que "representa" (coincide com) } \tilde{\top}$$

$P: \mathcal{F} \to \mathbb{R}_+$
$A\in\mathcal{F} \mapsto P(A) = a^*$, onde a^* é tal que $A\sim\{x\in[0,a^*]\}$.

Desse modo,

$$A \sim \{X \in [0, \mathbb{P}(A)]\}.$$

Teorema. $\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$ definida acima representa $\precsim$.

Para $A, B \in \mathcal{F}$, $\exists\, \mathbb{P}(A), \mathbb{P}(B) \geq 0$ tais que

$$A \sim \{X \in [0, \mathbb{P}(A)]\}$$
$$B \sim \{X \in [0, \mathbb{P}(B)]\}$$

$$A \precsim B \Leftrightarrow \{X \in [0, \mathbb{P}(A)]\} \precsim \{X \in [0, \mathbb{P}(B)]\} \Leftrightarrow \mathbb{P}(A) \leq \mathbb{P}(B).$$

$\square$

Teoria da Decisão

Aula 04 19/08/15

SP2 $A_1, A_2, B_1, B_2 \in \mathcal{F}$ com $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$

$A_i \precsim B_i, i = 1,2$, $A_1 \cup A_2 \precsim B_1 \cup B_2$

Caso $A_1 \prec B_1 \Rightarrow A_1 \cup A_2 \prec B_1 \cup B_2$

Resultado

$\forall A \in \mathcal{F}, \exists! \, a^* \in [0,1]$ tal que

$$A \sim \{X \in [0, a^*]\}$$

Vimos que $A \precsim B \Leftrightarrow P(A) \leq P(B)$.

1) $A \precsim B \Leftrightarrow \begin{cases} w_0 \in B \\ w_0 \notin B \text{ e } w_0 \notin A \end{cases}$ 2 $\Omega = \mathbb{N}, \mathcal{F} = \mathcal{P}(\mathbb{N})$

$A \precsim B \Leftrightarrow \begin{cases} B \text{ inf.} \\ B \text{ fini.}, A \text{ finito}, |A| \leq |B| \end{cases}$

1. $A_2 \prec B_2 \overset{Def}{\Rightarrow} w_0 \notin A_2 , w_0 \in B_2 \underset{B_1 \cap B_2 = \emptyset}{\Rightarrow} w_0 \notin B_1 \underset{A_1 \precsim B_1}{\Rightarrow} w_0 \notin A_1$

$\Rightarrow w_0 \notin A_1 \cup A_2 \quad \Rightarrow A_1 \cup A_2 \prec B_1 \cup B_2$.

$w_0 \in B_1 \cup A_2$

$A_1 \cup A_2 \precsim B_1 \cup B_2$ mas não $B_1 \cup B_2 \precsim A_1 \cup A_2$

2. No exemplo 2, prov. que vale a primeira parte de SP2. No entanto, não vale a parte (2)

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

$B_1 = \{1,3,5,\dots\}$ $\qquad A_1 \not\top B_1$ $\qquad$ e $\qquad A_1 \cup A_2 \not\top B_1 \cup B_2$

$A_2 = \{1\}$ $\qquad A_2 \not\top B_2$ $\qquad B_1 \cup B_2 \not\top A_1 \cup A_2$

$B_2 = \{2,4,6,\dots\}$

$A_1 \cup A_2 \sim B_1 \cup B_2$

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

$$A \in \mathcal{F} \mapsto \mathbb{P}(A) = a^* \text{ tal que } A \sim \{x \in [0, a^*]\}$$

é uma medida de probabilidade.

Demonstração:

Por construção, $\mathbb{P}(A) \geq 0$. Além disso, $\{x \in [0,1]\} = \Omega$.

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

. Aditividade finita de $\mathbb{P}$

$A_1, A_2 \in \mathcal{F}$ com $A_1 \cap A_2 = \emptyset$. Da definição de $\mathbb{P}$, temos:

$A_1 \sim \{x \in [0, \mathbb{P}(A_1)]\}$, $A_2 \sim \{x \in [0, \mathbb{P}(A_2)]\}$

$\sim \{x \in [0, \mathbb{P}(A))\}$

$A_1 \cup A_2 \sim \{x \in [0, \mathbb{P}(A_1 \cup A_2)]\}$

Vamos verificar que $A_2 \sim \{x \in [\mathbb{P}(A_1), \mathbb{P}(A_1 \cup A_2)]\}$. Suponhamos,

$A_2 \top \{x \in [\mathbb{P}(A_1), \mathbb{P}(A_1 \cup A_2)]\} \xrightarrow{SP2} A_1 \cup A_2 \top \{x \in [0, \mathbb{P}(A_1 \cup A_2)]\}$

$A_1 \not\top \{x \in [0, \mathbb{P}(A_1))\}$ $(\Rightarrow \Leftarrow)$

∴ Fazer sup. $A_2 \top \{x \in [\mathbb{P}(A_1), \mathbb{P}(A_1 \cup A_2)]\}$.

Logo, $A_2 \sim \{x \in [\mathbb{P}(A_1), \mathbb{P}(A_1 \cup A_2)]\}$

Logo,

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

Exercício.

$A_1, \dots, A_n \in \mathcal{F}$, $A_i \cap A_j = \emptyset$, $i \neq j$

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

Continuidade 'no' vazio.

$(A_n)_{n \geq 1}$, com $A_n \in \mathcal{F}$.

$A_n \supseteq A_{n+1}$, $\forall n \geq 1$ tal que

$$\lim_{n \to +\infty} A_n = \bigcap_{n=1}^{\infty} A_n = \emptyset$$

$A_{n+1} \subseteq A_n \Rightarrow A_{n+1} \subseteq A_n \Leftrightarrow \mathbb{P}(A_{n+1}) \leq \mathbb{P}(A_n)$.

A seq. $(\mathbb{P}(A_n))_{n \geq 1}$ é não crescente ~~(decrescente)~~ e limitada inferiormente.

Logo,

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

Claro que $\mathbb{P}(A_n) \geq c$, $\forall n \geq 1$.

Mas $\mathbb{P}(A_n) \geq c \Rightarrow \{x \in [0, c]\} \subseteq A_n$, $\forall n \geq 1$

Segue de SP4 que

$$\{x \in [0, c]\} \subseteq \bigcap_{n=1}^{\infty} A_n = \emptyset \Rightarrow c \leq 0$$

Uma vez que $c \geq 0$, segue que

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

Da aditividade finita e continuidade no vazio, segue que $\mathbb{P}$ é σ-aditiva, isto é,

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

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

Vimos até agora que existe $\mathbb{P}$ que representa $\precsim$ ($\precsim$ satisf. SP1 a SP5), onde $\mathbb{P}$ é uma med. de prob.

Seja $\mathbb{P}' : \mathcal{F} \to \mathbb{R}_+$ prob. que representa $\precsim$.

Como $\mathbb{P}'$ é med. de prob.,

$$\mathbb{P}'(\Omega) = \mathbb{P}'(\{X \in [0,1]\}) = 1.$$

e

$$\mathbb{P}'(\emptyset) = \mathbb{P}'(\{X \in [0,0]\}) = 0.$$

Vamos verificar que $\forall a \in [0,1], \; \mathbb{P}'(\{X \in [0,a]\}) = a$.

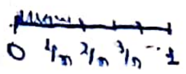
Obs:

$\forall n \geq 1, \; \forall j = 1 \ldots n,$

$\{X \in [\frac{j-1}{n}, \frac{j}{n})\} \sim \{X \in [0, \frac{1}{n})\}$. Disso, decorre que

$$\mathbb{P}'(\{X \in (\tfrac{j-1}{n}, \tfrac{j}{n})\}) = 1/n.$$

↳ segue da uniformidade de $\mathbb{P}'$ em $[0,1]$



Por construção deve valer que $\mathbb{P}(\;\;) = 1/n$.

Pela aditividade de $\mathbb{P}'$,

$$\mathbb{P}'\left(\left\{X \in \left[0, \tfrac{j}{n}\right)\right\}\right) = \frac{j}{n}, \; \forall n \geq 1, \; j = 1, \ldots, n.$$

$\forall n \geq 1$

$$\lfloor an \rfloor \leq an \leq \lfloor an \rfloor + 1$$

$\lfloor \ \rfloor$ "função piso" "menor parte inteira"

$$\frac{\lfloor an \rfloor}{n} \leq a \leq \frac{\lfloor an \rfloor}{n} + \frac{1}{n}$$

Conseq.

$$\{x \in [0, \lfloor an \rfloor / n]\} \subseteq \{x \in [0, a]\} \subseteq \{x \in [0, \lfloor an \rfloor / n + 1/n]\}$$

Conseq.

$$P'(\{x \in [0, \lfloor an \rfloor / n]\}) \leq P'(\{x \in [0,a]\}) \leq P'(\{x \in [0, \lfloor an \rfloor / n + 1/n]\}) \Rightarrow$$

$$\frac{\lfloor an \rfloor}{n} \leq P'(\{x \in [0,a]\}) \leq \frac{\lfloor an \rfloor + 1}{n}$$

Tomando limites ($n \to \infty$), resulta que

$$P'(\{x \in [0,a]\}) = a.$$

Finalmente, $\forall A \in \mathcal{F}$, temos

$$A \sim \{x \in [0, P(A)]\}$$

$P'(A) = P'(\{x \in [0, P(A)]\}) = P(A)$. Logo, $P'(\cdot) = P(\cdot)$ e, port., P é única.

D, in—

$A, B, D \in \mathcal{F}$.

$A, B, D \in \mathcal{F}$

$$A \mathbin{\mathcal{I}^D} B.$$

Dados

$A = \{3, 4, 5, 6\}$ $\qquad B \mathbin{\mathcal{I}} A$

$B = \{1\}$ $\qquad A \mathbin{\mathcal{I}^D} B$

$D = \{1, 2\}$

SP6

$$A \mathbin{\mathcal{I}^D} B \Leftrightarrow A \cap D \mathbin{\mathcal{I}} B \cap D$$

Ex. (se $D = \emptyset$? = tudo equivalente)

Consequências.

(1) $A \mathbin{\mathcal{I}^D} B$ ou $B \mathbin{\mathcal{I}^D} A$

(2) $\mathcal{I}^D$ é transitiva

$$\left.\begin{array}{l} A \mathbin{\mathcal{I}^D} B \Rightarrow A \cap D \mathbin{\mathcal{I}} B \cap D \\ B \mathbin{\mathcal{I}^D} C \Rightarrow B \cap D \mathbin{\mathcal{I}} C \cap D \end{array}\right. \Rightarrow A \cap D \mathbin{\mathcal{I}} C \cap D \Rightarrow A \mathbin{\mathcal{I}^D} C.$$

(3) $A \mathbin{\mathcal{I}^D} B \Leftrightarrow A \cap D \mathbin{\mathcal{I}} B \cap D \Leftrightarrow$

$$(A \cap B \cap D) \cup (A \cap B^c \cap D) \mathbin{\mathcal{I}} (A \cap B \cap D) \cup (A^c \cap B \cap D) \Leftrightarrow$$

$$A \cap B^c \cap D \mathbin{\mathcal{I}} A^c \cap B \cap D \Leftrightarrow$$

$$(A^c \cap B^c \cap D) \cup (A \cap B^c \cap D) \mathbin{\mathcal{I}} (A^c \cap B \cap D) \cup (A^c \cap B^c \cap D) \Leftrightarrow$$

$$B^c \cap D \mathbin{\mathcal{I}} A^c \cap D \Leftrightarrow$$

$$B^c \mathbin{\mathcal{I}^D} A^c.$$

Exercícios:

4) $(A_n)_{n \geq 1}$, [redacted] $A_n \in \mathcal{F}$, $A_n \supseteq A_{n+1}$, $\forall n \geq 1$ e $B \in \mathcal{F}$ t.q.

$$B \mathbin{\mathcal{I}^D} A_n, \forall n \geq 1 \overset{\text{Exercício}}{\overset{?}{\Rightarrow}} B \mathbin{\mathcal{I}^D} \bigcap_{n=1}^{\infty} A_n.$$

Note que

$$P_D : \mathcal{F} \to \mathbb{R}$$
$$A \in \mathcal{F} \mapsto P_D(A) = \frac{P(A \cap D)}{P(D)}$$

representa $\precsim^D$, se $P(D) > 0$.

Com efeito, seja D t.q. $P(D) > 0$, $A, B \in \mathcal{F}$.

$$A \precsim^D B \Leftrightarrow A \cap D \precsim B \cap D \Leftrightarrow$$
$$P(A \cap D) \leq P(B \cap D) \Leftrightarrow$$
$$\frac{P(A \cap D)}{P(D)} \leq \frac{P(B \cap D)}{P(D)} \Leftrightarrow$$
$$P_D(A) \leq P_D(B)$$ ↪ é medida de probabilidade.

"Chamamos isso de Construção de Probabilidade Subjetiva". Partimos de alguns axiomas e chegamos a uma representação.

¡Rever conjugadas!

P_1 e P_2 são duas expressões de incerteza sobre ~~[illegible]~~ a recompensa que se têm a partir de um certo exp.

A partir daí, escolher a melhor a partir da utilidade (a ser construída).

Teoria da Decisão

Aula 05 24/08/15

$A \precsim B$. Condições sobre $\precsim$ $\Rightarrow$ 1) $\exists \mathbb{P}$ med. de prob. que representa (coincide) com $\precsim$ e

2) $\forall D \in \mathcal{F}$ com $P(D) > 0$,

$$P_D(\cdot) = \frac{P(\cdot \cap D)}{P(D)} \text{ rep. } \precsim^D.$$

$A_1, \dots, A_k$.

$$\mathbb{P}(A_i | D) = \frac{\mathbb{P}(D | A_i)\mathbb{P}(A_i)}{\mathbb{P}(D)}$$

$A_i \rightarrow$ evento associado à qtde de interesse "Θ"

$D \rightarrow$ dados

$$\mathbb{P}(A_i | \text{Dados}) \propto \mathbb{P}(\text{Dados} / A_i)\mathbb{P}(A_i)$$

$$f(\theta / x) \propto f(x | \theta)\pi(\theta)$$

Ex: $X_1, \dots, X_n$, dado Θ, são c. i.i.d. $Ber(\theta)$

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | \theta) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i | \theta)$$

$$= \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}$$

$\Theta \sim Beta(a,b)$

$f(\theta | x_1 \dots x_n)$

Permutabilidade

Dizemos que as variáveis $X_1, \dots, X_k$ são permutáveis se, $i = 1, \dots, k$,

$$P(X_1 = x_1, \dots, X_k = x_k) = P(X_{\pi_1} = x_{\pi_1}, \dots, X_k = \pi_k),$$

$\forall \pi : \{1, \dots, k\} \to \{1, \dots, k\}$ bijetora

Consequências.

1. $X_1, \dots, X_k$ são ident. dist.

$$P(X_i = x) = \sum_{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k} P(X_1 = x_1, \underline{X_2 = x_2}, X_3 = x_3, \dots, X_{i-1} = x_{i-1}, X_i = x, X_{i+1} = x_{i+1}, \dots, X_k)$$

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

$$= \sum_{x_1, x_3, \dots, x_k} P(X_1 = x, X_2 = x_1, X_3 = x_3, \dots, X_k = x_k) = P(X_1 = x).$$

Ex $k = 2$

$X_1 \backslash X_2$	x_1	x_2	x_3
x_1		p_{12}	p_{13}
x_2	p_{21}		p_{23}
x_3	p_{31}	p_{32}	

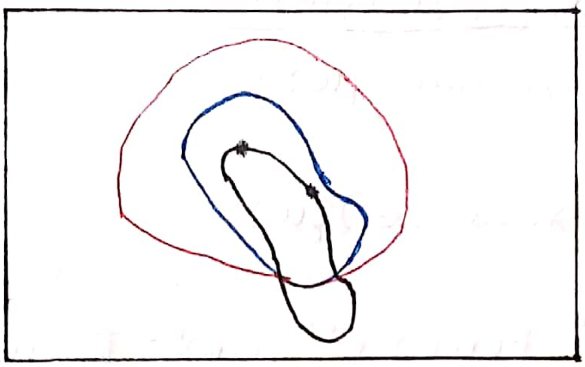
Permutabilidade $\not\Rightarrow$ Independência

$X_1 \backslash X_2$	0	1	
0	0	1/3	1/3
1	1/3	1/3	2/3
	1/3	2/3	1

$X_1 \overset{d}{=} X_2 \sim Ber(2/3)$

$P(X_1 = 1, X_2 = 1) = 1/3$ e

$P(X_1 = 1)P(X_2 = 1) = \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$



Definição. $(X_n)_{n\geq 1}$ é uma sequência de v.a. permutáveis se $\forall k \in \mathbb{N}$, $X_1, \dots, X_k$ são permutáveis.

Teorema da Representação de Definetti.

Seja $(X_n)_{n\geq 1}$ uma sequência **permutável** de v.a. de Bernoulli. Seja $\mathbb{P}$ a medida de probabilidade do processo. Então, existe $\mu : \mathcal{B}([0,1]) \to \mathbb{R}_+$ medida de probabilidade tal que

$\forall k \in \mathbb{N}^*$ e para $x_1, \dots, x_k \in \{0,1\}$, $i = 1, \dots, k$,

$$\mathbb{P}(X_1 = x_1, \dots, X_k = x_k) = \int_0^1 \theta^{\sum_{i=1}^{k} x_i} (1-\theta)^{k - \sum_{i=1}^{k} x_i} \, d\mu(\theta)$$

Além disso,

$$\bar{X}_k = \frac{X_1 + \dots + X_k}{k} \xrightarrow{q.c.} \Theta, \text{ onde } \Theta \text{ é distribuído segundo } \mu.$$

Alguns resultados conhecidos

1) $X_1, \dots, X_n$ dado Θ, são c.i.i.d. $Ber(\Theta)$.

$\Theta \sim Beta(a,b)$, $a, b > 0$.

$$f(\theta / x_1, \dots, x_n) = \frac{P(X_1 = x_1, \dots, X_n = x_n / \theta) f(\theta)}{P(X_1 = x_1, \dots, X_n = x_n)} = *$$

$$* \propto \frac{\mathbb{P}(X_1=x_1,\dots,X_n=x_n \mid \theta)\, f(\theta)}{\int_0^1 \mathbb{P}(X_1=x_1,\dots,X_n=x_n \mid \theta')\, f(\theta')\, d\theta'}$$

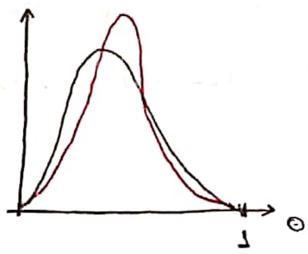
$$\propto \mathbb{P}(X_1=x_1,\dots,X_n=x_n \mid \theta)\, f(\theta)$$

$$\propto \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, \theta^{a-1}(1-\theta)^{b-1}\, \mathbb{I}_{(0,1)}(\theta)$$

$$\Rightarrow f(\theta \mid x_1,\dots,x_n) \propto \theta^{a+\sum x_i-1}(1-\theta)^{b+n-\sum x_i-1}\, \mathbb{I}_{(0,1)}(\theta)$$

Logo,

$$\theta \mid x_1,\dots,x_n \sim \text{Beta}\left(a+\sum_{i=1}^{n} x_i,\ b+n-\sum_{i=1}^{n} x_i\right)$$



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

$$\text{Var}(\theta) = \frac{ab}{(a+b)^2(a+b+1)}$$

"A família de distribuições beta é conjugada pelo modelo de Bernoulli"

2) $X_1,\dots,X_n$ dado θ, são c.i.i.d. Poisson(θ)

$\theta \sim$ Gama(a,b)

$$f(\theta \mid x_1,\dots,x_n) = \frac{\mathbb{P}(X_1=x_1,\dots,X_n=x_n \mid \theta)\, f(\theta)}{\mathbb{P}(X_1=x_1,\dots,X_n=x_n)}$$

$$\propto \mathbb{P}(X_1=x_1,\dots,X_n=x_n \mid \theta)\, f(\theta)$$

$$\propto \left(\prod_{i=1}^{n} \mathbb{P}(X_i=x_i \mid \theta)\right) f(\theta) \propto \left(\prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!}\right) \frac{b^a}{\Gamma(a)}\, \theta^{a-1} e^{-b\theta}\, \mathbb{I}_{\mathbb{R}_+}(\theta) \overset{de}{\prod_{i}} \mathbb{I}_{\mathbb{N}}(x_i)$$

$$\propto e^{-n\theta}\ \theta^{\sum_{i=1}^{n} x_i}\ \theta^{a-1}\ e^{-b\theta}\ \mathbb{I}_{\mathbb{R}_+}(\theta)$$

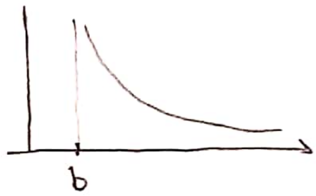
$$\Rightarrow f(\theta|x_1,\dots,x_n) \propto \theta^{a+\sum_{i=1}^{n} x_i - 1}\ e^{-(n+b)\theta}\ \mathbb{I}_{\mathbb{R}_+}(\theta)$$

$$\theta / x_1,\dots,x_n \sim \text{Gama}\left(a+\sum_{i=1}^{n} x_i,\ b+n\right)$$

a) $X_1,\dots,X_n$, dado θ, c.i.i.d. Uniforme $(0,\theta)$

$\theta \sim$ Pareto (a,b)

$$f(\theta) = \begin{cases} C(a,b)\ {}^{1}/_{\theta^{a+1}}\ \mathbb{I}_{(b,\infty)}(\theta) \\ 0,\ \text{c.c.} \end{cases}$$



$$f(\theta/x_1,\dots,x_n) \propto f(x_1,\dots,x_n|\theta)\, f(\theta)$$
$$= \prod_{i=1}^{n} f(x_i|\theta)\cdot f(\theta)$$

$$\propto \prod_{i=1}^{n} \frac{1}{\theta}\, \mathbb{I}_{(0,\theta)}(x_i)\ C(a,b)\, \frac{1}{\theta^{a+1}}\ \mathbb{I}_{(b,\infty)}(\theta)$$

$$= \frac{1}{\theta^{n}}\ \mathbb{I}_{(0,x_{(n)})}(x_{(1)})\ \mathbb{I}_{(0,\theta)}(x_{(n)})\ C(a,b)\ \frac{1}{\theta^{a+1}}\ \mathbb{I}_{(b,\infty)}(\theta)$$

$$\propto \frac{1}{\theta^{a+n+1}}\ \mathbb{I}_{(0,\infty)}(x_{(n)})\ \mathbb{I}_{(\max\{x_{(n)},b\},\infty)}(\theta) \Rightarrow$$

$$f(\theta|x_1,\dots,x_n) \propto \frac{1}{\theta^{a+n+1}}\ \mathbb{I}_{\{\max\{x_{(n)},b\},\infty\}}(\theta)$$

$\therefore$

$$\theta / x_1,\dots,x_n \sim \text{Pareto}\left(a+n,\ \max\{x_{(n)},b\}\right)$$

"A fam. de dist. da Pareto é conjug. pelo modelo uniforme."

Teoria da Decisão

Aula 06 26/08

Ex. 4.

$X_1, \dots, X_n$ dado $\Theta = (\Theta_1, \Theta_2)$ são i.i.d. $N(\Theta_1, k/\Theta_2)$, $k > 0$.

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

$\Theta_2 \sim Gama(a, b)$, $a, b > 0$ $\qquad \Theta_1 | \Theta_2 \sim N(c, d/\Theta_2)$

$$f(\Theta_1, \Theta_2) = f(\Theta_2) f(\Theta_1 | \Theta_2)$$

$$f(\Theta | x_1, \dots x_n) \propto f(x_1, \dots x_n | (\Theta_1, \Theta_2)) \underbrace{f(\Theta_1, \Theta_2)}_{f(\Theta_2) f(\Theta_1 | \Theta_2)}$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\, k/\Theta_2}} e^{-\frac{(x_i - \Theta_1)^2}{2k/\Theta_2}} \frac{b^a}{\Gamma(a)} \Theta_2^{a-1} e^{-b\Theta_2} \mathbb{I}_{\mathbb{R}_+}(\Theta_2) \frac{1}{\sqrt{2\pi\, d/\Theta_2}} e^{-\frac{(\Theta_1 - c)^2}{2d/\Theta_2}}$$

$$\propto \Theta_2^{n/2} e^{-\Theta_2/2k \sum_{i=1}^{n} (x_i - \Theta_1)^2} \Theta_2^{a-1} e^{-b\Theta_2} \Theta_2^{1/2} e^{-\Theta_2/2d (\Theta_1 - c)^2} \mathbb{I}_{\mathbb{R}_+}(\Theta_2)$$

$-\frac{2\Theta_1 n \bar{x}}{k}$ $\qquad -\frac{2\Theta_1 c}{d}$

$$\propto \Theta_2^{\frac{n+1}{2} + a - 1} e^{-b\Theta_2} \mathbb{I}_{\mathbb{R}_+}(\Theta_2) e^{-\frac{\Theta_2}{2} \left\{ \Theta_1^2 \left[\frac{1}{d} + \frac{n}{k} \right] - 2\Theta_1 \left[\frac{c}{d} + \frac{n\bar{x}}{k} \right] \right\}} e^{-\frac{\Theta_2}{2k} \sum^{n} x_i^2 - \frac{\Theta_2}{2d} c^2}$$

$$\propto \Theta_2^{\frac{n+1}{2} + a - 1} e^{-\Theta_2 \left[b + \sum x_i^2 / 2k + c^2/2d \right]} \mathbb{I}_{\mathbb{R}_+}(\Theta_2) e^{-\frac{\Theta_2}{2} \left(\frac{k+nd}{dk} \right) \left(\Theta_1 - \frac{ck + dn\bar{x}}{k+dn} \right)^2} \times$$

$$e^{\frac{\Theta_2}{2} \left(\frac{k+nd}{dk} \right) \left(\frac{ck + dn\bar{x}}{k+dn} \right)^2}$$

$$\propto \underbrace{\Theta_2^{\overbrace{\frac{n}{2} + a - 1}^{a'}} e^{-\Theta_2 \left[\overbrace{b + \frac{\sum x_i^2}{2k} + \frac{c^2}{2d} - \frac{1}{2} \left(\frac{k+nd}{dk} \right) \left(\frac{ck + dn\bar{x}}{k+dn} \right)^2}^{b'} \right]} \cdot \mathbb{I}_{\mathbb{R}_+}(\Theta_2)}_{*}$$

$$\times\, e^{-\frac{1}{2 \left(\underbrace{\frac{dk}{k+nd}}_{d'} \right) \frac{1}{\Theta_2}} \left(\Theta_1 - \underbrace{\frac{ck + dn\bar{x}}{k+dn}}_{c'} \right)^2} \times \Theta_2^{1/2}$$

$\theta_2 \mid x_1, \ldots, x_n \sim \text{Gama}(a', b')$

$\theta_1 \mid \theta_2, x_1, \ldots, x_n \sim \text{Normal}\left(c', \dfrac{d'}{\theta_2}\right)$

Livro De Groot, Cap. 9 → mais conjugadas

Populações Finitas

$P = \{1, \ldots, N\}$ uma população.

Seja $X = (X_1, \ldots, X_N)$ um vetor de caract. populacionais, ou melhor vetor populacional de características.

$$T(X) = \frac{\sum_{i=1}^{N} X_i}{N}, \quad \mathbb{P}\left(T(X) = j/N\right) = \sum_{\substack{x_1 \ldots x_N \\ \sum_{i=1}^{N} x_i = j}} \mathbb{P}(X_1 = x_1, \ldots, X_N = x_N)$$

Amostra $X_1, \ldots, X_n$, $n < N$.

$$\mathbb{P}(T(X) = j/N \mid X_1 = x_1, \ldots, X_n = x_n).$$

Referências

1. Wechsler, S (1993). Exchangeability and Prediction

2. Iglesias, Loschi, Pereira, Wechsler (2009) BJPS, 23(2), 216-226

3. Dissertação. Brian A. R. Melo (2013)

4. Barlow, R (2003). "Engineering e Reliability"

Preferência (Utilidade)

$\mathcal{R} \rightarrow$ conjunto de recompensas (consequências)

1. Vendedor de pamonhas

Preço: v $\qquad\qquad$ $\mathcal{R}'$ $\{-c, v-c, \dots, 30v-c, \dots\}$

Custo do lote: c

$$iv - c \succsim' jv - c \quad i \geq j.$$

2. Jogo do Corinthians

$$\mathcal{R} = \{ 😢, 😐, 😊, U \} \qquad 😢 \precsim' 😐$$

Para cada $r \in \mathcal{R}$,

$$A(r) = \{ r' \in \mathcal{R} : r' \precsim^* r \}$$

$$U : \mathcal{R} \rightarrow \mathbb{R}_+$$
$$r \in \mathcal{R} \mapsto U(r) = |A(r)|$$

Sob certas condições, para $r_1, r_2 \in \mathcal{R}$,

$$r_1 \succsim^* r_2 \Leftrightarrow U(r_1) \geq U(r_2).$$

30 pamonhas

$$\mathbb{P}_{30}(\{jv - 30c\}) = \mathbb{P}(\text{"demanda } j\text{"}) \quad j = 0, \dots, 30$$

$$\mathbb{P}_{10}(\{jv - 10c\}) = \mathbb{P}(\text{"demanda } j\text{"}), \; j \leq 10$$

$$\mathbb{P}_{10}(\{10(v-c)\}) = \mathbb{P}(\text{"demanda} > 10\text{"}).$$

Cenário de Incerteza

1. Parranha

Produção de "i" parranhas, $i=0,\dots,30$

$P_i(\{v_j - ic\}) = P(\text{"demanda } j\text{"}), j \leq i$

$P_i(\{iv - c\}) = P(\text{"demanda} > i\text{"})$

$R = \{(v_j - c_i), i, j \in \{0, 1, \dots, 30\}$

$P_1 \overset{*}{\precsim} P_2$

Formulação.

$R \to$ conjunto de recompensas (consequências)

$A \to$ σ-álgebra de subconjuntos de R.

$\mathcal{P} \to$ conjunto de medidas de prob. em (R, A).

Vamos considerar $\precsim^{a}$ em $R \times R$, $r_1, r_2 \in R$

$r_1 \precsim^{*} r_2 \quad r_1 \sim^{*} r_2 \quad r_1 \prec^{*} r_2$

Para $r_1, r_2 \in R$

$[r_1, r_2] = \{r \in R : r_1 \precsim^{*} r \precsim^{*} r_2\}$

Para todo $r_1, r_2 \in R$, vamos supor que

$[r_1, r_2] \in A$.

Dizemos que $P \in \mathcal{P}$ é limitada se existem $r_1, r_2 \in R$ tais que

$P([r_1, r_2]) = 1$

Seja $\mathcal{P}_L = \{P \in \mathcal{P} : P \text{ é limitada}\}$.

Vamos estender $\succsim^*$ para $\mathcal{P}_L \times \mathcal{P}_L$.
Supomos $\succsim^*$ completa e transitiva, isto é,

$$P_1 \succsim^* P_2 \text{ ou } P_2 \succsim^* P_1 \quad e$$

$\forall P_1, P_2, P_3 \in \mathcal{P}_L$,

$$P_1 \succsim^* P_2 \text{ e } P_2 \succsim^* P_3 \Rightarrow P_1 \succsim^* P_3$$

Do mesmo modo que $\succsim^*$ restrita à $\mathbb{R} \times \mathbb{R}$,

$P_1 \sim^* P_2 \Leftrightarrow P_1 \succsim^* P_2$ e $P_2 \succsim^* P_1$

$P_1 \succ^* P_2$

Para cada $r \in \mathbb{R}$, vamos considerar r equivalente a $P_{(r)} \in \mathcal{P}_L$ onde

$$P_{(r)}(\{r\}) = 1.$$

Se P_1 e $P_2 \in \mathcal{P}_L$ e $\alpha \in [0,1]$, então $\alpha P_1 + (1-\alpha) P_2 \in \mathcal{P}_L$.

$\alpha P_1(A) + (1-\alpha) P_2(A)$.

Em particular, se $r_1, r_2 \in \mathbb{R}$,

$$\alpha r_1 + (1-\alpha) r_2 .$$

Supomos que existem $s_0, t_0 \in \mathbb{R}$ tais que $s_0 \succ^* t_0$.

Suposições Adicionais

SU_1: Sejam $\mathbb{P}_1, \mathbb{P}_2, \mathbb{P} \in \mathcal{P}_L$ e $\alpha \in (0,1)$.

$$\mathbb{P}_1 \succsim^* \mathbb{P}_2 \Leftrightarrow \alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P} \succsim^* \alpha \mathbb{P}_2 + (1-\alpha)\mathbb{P}$$

Teoria da Decisão

Aula 07 31/08

Suposições Adicionais.

SU1: $\mathbb{P}_1, \mathbb{P}_2$ e $\mathbb{P} \in \mathcal{P}_L$ e $\alpha \in (0,1)$.

$$\mathbb{P}_1 \precsim^* \mathbb{P}_2 \Leftrightarrow \alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P} \precsim \alpha \mathbb{P}_2 + (1-\alpha)\mathbb{P}$$

Implicações:

Lema 1. $\mathbb{P}_1, \mathbb{P}_2, \mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{P}_L$ tais que

$$\mathbb{P}_i \precsim \mathbb{Q}_i, \ i=1,2;\ \alpha \in (0,1)$$

Então, por SU1,

$$\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2 \precsim \alpha \mathbb{Q}_1 + (1-\alpha)\mathbb{Q}_2$$

Demonstração: Por SU1,

$$\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2 \precsim^* \alpha \mathbb{Q}_1 + (1-\alpha)\mathbb{P}_2 \precsim^* \alpha \mathbb{Q}_1 + (1-\alpha)\mathbb{Q}_2$$

Lema 2: $r_1, r_2 \in \mathbb{R}$ tais que $r_1 \prec^* r_2$, $\alpha \in (0,1)$.

$$r_1 \overset{*}{\sim} \alpha r_1 + (1-\alpha) r_1 \prec^* \alpha r_2 + (1-\alpha) r_1 \prec^* \alpha r_2 + (1-\alpha) r_2 \overset{*}{\sim} r_2$$

Logo, $r_1 \prec^* \alpha r_2 + (1-\alpha) r_1 \prec^* r_2$.

Lema 3. $r_1, r_2 \in \mathbb{R}$ tais que

$$r_1 \prec^* r_2$$

Sejam $\alpha, \beta \in [0,1]$.

$$\alpha r_2 + (1-\alpha) r_1 \prec^* \beta r_2 + (1-\beta) r_1 \Leftrightarrow \alpha < \beta.$$

Demonstração: Supondo que $\underline{\alpha < \beta}$ $(0 < \alpha < \beta < 1)$

$$\beta r_2 + (1-\beta) r_1 \sim \frac{1-\beta}{1-\alpha}\left[\alpha r_2 + (1-\alpha) r_1\right] + \left(1 - \frac{1-\beta}{1-\alpha}\right) r_2$$

$$\alpha r_2 + (1-\alpha) r_1 \sim \frac{1-\beta}{1-\alpha}\left[\alpha r_2 + (1-\alpha) r_1\right] \qquad \prec^*$$

$$+ \left(1 - \frac{1-\beta}{1-\alpha}\right)\underbrace{\left[\alpha r_2 + (1-\alpha) r_1\right]}_{\prec^* r_2}$$

Lema 2

SU1

$$\Rightarrow \prec^* \underbrace{\frac{1-\beta}{1-\alpha}\left[\alpha r_2 + (1-\alpha) r_1\right] + \left(1 - \frac{1-\beta}{1-\alpha}\right) r_2}_{\sim^* \beta r_2 + (1-\beta) r_1}$$

Então,

$$\alpha < \beta \Rightarrow \alpha r_2 + (1-\alpha) r_1 \prec^* \beta r_2 + (1-\beta) r_1$$

relação de preferência r_i pref. sob que conseq. quero

- $\underline{\alpha >_r \beta}$

$\alpha > \beta \Rightarrow \beta r_2 + (1-\beta) r_1 \prec^* \alpha r_2 + (1-\alpha) r_1$

$\alpha = \beta$

~~Sim~~:

SU2: Sejam $P_1, P_2, P \in \mathcal{D}_L$ tais que

$$P_1 \prec^* P \prec^* P_2,$$

$\exists\ \alpha, \beta \in (0,1)$ t.q. $\alpha P_1 + (1-\alpha) P_2 \prec^* P$ e $P \prec^* \beta P_1 + (1-\beta) P_2$

Teorema 1. Sejam $r_1, r, r_2 \in \mathbb{R}$ tais que

$$r_1 \prec^* r_2 \quad \text{e} \quad r_1 \preceq^* r \preceq^* r_2,$$

algum jogo

existe um $\check{\nu} \in [0,1]$ tal que

$$r \sim^* \check{\nu} r_2 + (1-\check{\nu}) r_1$$

Caso trivial. $r \sim^* r_2 \Rightarrow \check{\nu} = 1$

$r \sim^* r_1 \Rightarrow \check{\nu} = 0$

Suponha agora $r_1 \prec^* r \prec^* r_2$. Sejam

$$U = \{\alpha \in [0,1] : r \prec^* \alpha r_2 + (1-\alpha) r_1\}$$
$$D = \{\alpha \in [0,1] : \alpha r_2 + (1-\alpha) r_1 \prec^* r\}$$

Observamos que $1 \in U$ e $0 \in D$ ($U \neq \emptyset$ e $D \neq \emptyset$).
Sejam $\mu^* = \inf\{U\}$ e $d^* = \sup\{D\}$.

Se $\mu \in U$ e $\mu' > \mu$, então $\mu' \in U$. (I)

$$r \prec^* \mu r_2 + (1-\mu) r_1 \underset{\text{Lema 1}}{\prec} \mu' r_2 + (1-\mu') r_1 \Rightarrow$$

$$r \prec^* \mu' r_2 + (1-\mu) r_1.$$

Ademais, $\mu^* \notin U$, (II).

Se $\mu^* \in U \Rightarrow r \prec^* \mu^* r_2 + (1-\mu^*) r_1$ e sabemos que $r_1 \prec^* r$

Por SU2, há $\gamma \in (0,1)$ t.q.:

$$r \prec^* \gamma[\mu^* r_2 + (1-\mu^*) r_1] + (1-\gamma) r_1 \sim \gamma \mu^* r_2 + (1-\gamma\mu^*) r_1$$

$$\Rightarrow \gamma\mu^* \in U.$$

De (I) e (II), segue que: $U = (\mu^*, 1]$

Do mesmo modo, verificamos que

$$d \in D \Rightarrow \forall\, d' < d,\ d' \in D \text{ e que } d^* \notin D.$$ (Exercício!!!)

Assim $D = [0, d^*)$. Como $U \cap D = \emptyset$, temos que $d^* \leq \mu^*$.

Seja $\check{\nu} \in [d^*, \mu^*]$. $\check{\nu} \notin U$ e $\check{\nu} \notin D$. Portanto,

$$\left.\begin{array}{l} \check{\nu} \notin U \Rightarrow \check{\nu} r_2 + (1-\check{\nu}) r_1 \precsim^* r \\ \check{\nu} \notin D \Rightarrow r \precsim^* \check{\nu} r_2 + (1-\check{\nu}) r_1 \end{array}\right] \Rightarrow r \sim \check{\nu} r_2 + (1-\check{\nu}) r_1$$

Ademais, $\check{\nu}$ é único. Seja $\check{\nu}' \in [0,1)$ t.q. $\check{\nu}' r_2 + (1-\check{\nu}') r_1 \sim r$.
Assim, $\check{\nu} r_2 + (1-\check{\nu}) r_1 \sim \check{\nu}' r_2 + (1-\check{\nu}') r_1$. Do Lema 3, segue que $\check{\nu}' = \check{\nu}$!

Definição da função U.

Lembre que $s_0, t_0 \in \mathbb{R}$ t.q.

$$s_0 \prec^* t_0$$

Seja $r \in \mathcal{R}$.

1. Se $s_0 \precsim^* r \precsim^* t_0 \overset{\text{Teo. 1}}{\Longrightarrow} \exists\, \nu \in [0,1]$ t.q.

$$r \stackrel{*}{\sim} \nu t_0 + (1-\nu) s_0$$

Seja $U: R \to \mathbb{R}$. Definimos $U(r) = \nu$,

Assim, $r \stackrel{*}{\sim} U(r) t_0 + (1 - U(r)) s_0$.

Consequência: $U(t_0) = 1$ e $U(s_0) = 0$.

2. $r \prec^* s_0$. Assim, $r \prec^* s_0 \prec^* t_0$. Daí, $\exists\, \gamma^* \in (0,1)$ t.q.

$$s_0 \stackrel{*}{\sim} \gamma^* t_0 + (1-\gamma^*) r$$

Nesse caso, vamos definir $U(r) = \dfrac{-\gamma^*}{1-\gamma^*}$.

Finalmente,

3. $t_0 \prec^* r$.

Neste caso, $s_0 \prec^* t_0 \prec^* r$. Pelo Teorema X, $\exists\, \beta \in (0,1)$ t.q

$$t_0 \stackrel{*}{\sim} \beta r + (1-\beta) s_0 \quad (U(t_0) = \beta U(r) + (1-\beta) U(s_0))$$

Definimos nesse caso, $U(r) = \dfrac{1}{\beta}$.

Notamos que:

i) $r \in [s_0, t_0]$, $r \overset{\circ}{\sim} U(r) t_0 + (1 - U(r)) s_0$

ii) $r \prec^{\circ} s_0$, $s_0 \overset{\circ}{\sim} \dfrac{-U(r)}{1 - U(r)} t_0 + \dfrac{1}{1 - U(r)} r$.

iii) $t_0 \prec r$, $t_0 \overset{\circ}{\sim} \dfrac{1}{U(r)} r + \left(1 - \dfrac{1}{U(r)}\right) s_0$.

Teorema 2. Sejam $r_1, r_2, r_3 \in \mathcal{R}$ t.q.

$r_2 \overset{\circ}{\sim} \alpha r_3 + (1 - \alpha) r_1$ para algum $\alpha \in [0,1]$. Então

$U(r_2) = \alpha U(r_3) + (1 - \alpha) U(r_1)$.

Teoria da Decisão

Aula 08 02/09

$\forall r \in [r_1, r_2]$, onde $r_1, r_2 \in \mathcal{R}$ são tais que $r_1 \precsim^{\circ} r_2$, $\exists \alpha = \alpha(r) \in [0,1]$ tal que

$$r \sim^{\circ} \alpha(r) r_2 + (1-\alpha(r)) r_1 .$$

Definimos $U: \mathcal{R} \to \mathbb{R}$ t.q.

1) $r \in [s_0, t_0]$

$$r \sim U(r) t_0 + (1 - U(r)) s_0$$

2) $r \precsim^{\circ} s_0$

$$s_0 \sim^{\circ} \frac{-U(r)}{1-U(r)} t_0 + \frac{1}{1-U(r)} r$$

3) $t_0 \precsim^{\circ} \frac{1}{U(r)} r + \left(1 - \frac{1}{U(r)}\right) s_0$

Teorema 2. $r_1, r_2, r_3 \in \mathcal{R}$ tais que

$$r_2 \sim^{\circ} \alpha r_3 + (1-\alpha) r_1, \text{ para algum } \alpha \in [0,1]$$

Então

$$U(r_2) = \alpha U(r_3) + (1-\alpha) U(r_1).$$

$\exists\, s_1, t_1 \in \mathbb{R}$ tais que

$s_1 \succ^* t_1$ e

$s_1 \succeq^* r \succeq^* t_1, \; \forall r \in \mathbb{R}' = \{r_1, r_2, r_3, s_0, t_0\}$

$\forall r \in \mathbb{R}', \exists\, \gamma \in [0,1]$ tal que

$$r \sim^* \gamma t_1 + (1-\gamma) s_1$$

Definimos U^* de modo que

$$r \sim^* U^*(r) t_1 + (1-U(r)) s_1 \qquad (ii)$$

Então,

$$r_2 \sim^* U^*(r_2) t_1 + (1-U(r_2)) s_1 \qquad (*)$$

Por outro lado,

$$r_2 \sim^* \alpha r_3 + (1-\alpha) r_1 \sim^* \alpha \left[U^*(r_3) t_1 + (1-U^*(r_3)) s_1 \right] + (1-\alpha)\left[U^*(r_1) t_1 + (1-U^*(r_1)) s_1 \right] \Rightarrow$$

$$r_2 \sim^* (\alpha U^*(r_3) + (1-\alpha) U^*(r_1)) t_1 + [1 - (\alpha U^*(r_3) + (1-\alpha) U^*(r_1))] s_1 \quad (**)$$

Do Lema 3,

$$U^*(r_2) = \alpha U^*(r_3) + (1-\alpha) U^*(r_1) \quad . \; r \in [s_0, t_0]$$

$r \sim^* U(r) t_0 + (1-U(r)) s_0 \Rightarrow$

$r \sim^* U(r) [U^*(t_0) t_1 + (1-U^*(t_0)) s_1] + (1-U(r)) [U^*(s_0) t_1 + (1-U^*(s_0)) s_1] \Rightarrow$

$r \sim^* \underline{(U(r) U^*(t_0) + (1-U(r)) U^*(s_0))} t_1 + (1 - \quad) . s_1 \qquad (i)$

De (i) e (ii)

$$U^*(r) = U(r)U^*(t_0) + (1-U(r))U^*(s_0) \Rightarrow U^*(r) = U(r)[U^*(t_0) - U^*(s_0)] + U^*(s_0)$$ (□)

2) $r \succsim^* s_0$

$$s_0 \sim^* \frac{-U(r)}{1-U(r)} t_0 + \frac{1}{1-U(r)} r \Rightarrow$$

$$\Rightarrow s_0 \sim^* \frac{-U(r)}{1-U(r)} \left[U^*(t_0)t_1 + (1-U^*(t_0))s_1\right] + \frac{1}{1-U(r)} \left[U^*(r)t_1 + (1-U^*(r))s_1\right]$$ (⇒)

~~3) $t_0 \precsim^* r$~~

$\cancel{t_0 \sim^* \frac{1}{U(r)} r + \left(1 - \frac{1}{U(r)}\right) s_0}$

$$(\Rightarrow)\ s_0 \sim^* \left[\underbrace{-\frac{U(r)}{1-U(r)} U^*(t_0) + \frac{1}{1-U(r)} U^*(r)}\right] t_1 + (1 - \quad) s_1$$ (o)

Por outro lado,

$$s_0 \sim^* U^*(s_0) + (1-U^*(s_0))s_1$$ (oo)

De (°) e (°°)

$$U^*(s_0) = \frac{-U(r)}{1-U(r)}\, U^*(t_0) + \frac{1}{1-U(r)}\, U^*(r) \Rightarrow$$

$$\Rightarrow U^*(r) = U^*(s_0) + U(r)\left[U^*(t_0) - U^*(s_0)\right]. \quad (\square\square)$$

3) $t_0 \; \tau^* \; r$

$$t_0 \sim^* \frac{1}{U(r)}\, r + \left(1 - 1/U(r)\right) s_0 \sim^*$$

$$\sim^* \frac{1}{U(r)}\left[U^*(r)t_1 + (1-U^*(r))s_1\right] + \left(1 - \frac{1}{U(r)}\right)\left[U^*(s_0)t_1 + (1-U^*(s_0))s_1\right] \Rightarrow$$

$$\Rightarrow t_0 \sim^* \left(\frac{U^*(r)}{U(r)} + \left(\frac{U(r)-1}{U(r)}\right) U^*(s_0)\right) U^*(s_0)\Bigg) t_1 + (1 - \quad) s_1 \quad (\Delta)$$

Por outro lado, $t_0 \sim^* U^*(t_0)t_1 + (1-U^*(t_0))s_1 \quad (\Delta\Delta)$

De (Δ) e $(\Delta\Delta)$, temos

$$U^*(t_0) = \frac{U^*(r)}{U(r)} + \left(\frac{U(r)-1}{U(r)}\right) U^*(s_0) \Rightarrow$$

$$\Rightarrow U^*(r) = U^*(s_0) + U(r)\left(U^*(t_0) - U^*(s_0)\right) \quad (\square\square\square)$$

Assim, de $(\square)$, $(\square\square)$ e $(\square\square\square)$, $\forall\, r \in R'$,

$$U^*(r) = U^*(s_0) + U(r)\left(U^*(t_0) - U^*(s_0)\right)$$

Como

$$U^*(r_2) = \alpha U^*(r_3) + (1-\alpha) U^*(r_1)$$

$$U^*(s_0) + U(r_2)(U^*(t_0) - U^*(s_0)) =$$

$$= \alpha [U^*(s_0) + U(r_3)(U^*(t_0) - U^*(s_0))] + (1-\alpha)[U^*(s_0) + U(r_1)(U^*(t_0) - U^*(s_0))] \Rightarrow$$

$$\Rightarrow U(r_2) = \alpha U(r_3) + (1-\alpha) U(r_1)$$

$$\Rightarrow P_1 \precsim' P_2 \Rightarrow \int U dP_1 \leq \int U dP_2 .$$

SU 3: $\forall r_1, r_2, r_3 \in R$, $\forall \alpha, \beta \in [0,1]$

$$\{r \in R : \alpha r + (1-\alpha) r_1 \precsim^* \beta r_2 + (1-\beta) r_3\} \in \mathcal{A}_0 .$$

Resultado. $U : R \to \mathbb{R}$ como def. na aula anterior é uma v.a. ($\mathcal{A}$-mensurável)
Vamos verificar que $\forall x \in \mathbb{R}$,

$$\{r \in R : U(r) \leq x\} \in \mathcal{A} .$$

(1) $x < 0 \to r \prec^* s_0$

Logo

$$s_0 \sim^* \frac{-U(r)}{1-U(r)} t_0 + \frac{1}{1-U(r)} r .$$

$$U(r) \leq x \Leftrightarrow \frac{-x}{1-x} \leq \frac{-U(r)}{1-U(r)} \overset{\text{Lema 3}}{\Longleftrightarrow}$$

$$-\frac{x}{1-x}t_0 + \frac{1}{1-x} r \precsim^* \underbrace{\frac{-U(r)}{1-U(r)} t_0 + \frac{1}{1-U(r)} r}_{s_0}$$

Logo,

$$U(r) \leq x \Leftrightarrow \frac{1}{1-x} r + \frac{-x}{1-x} t_0 \precsim^* s_0$$

Logo,

$$\{r \in \mathbb{R} : U(r) \leq x\} \in \mathcal{A}.$$

1). $x \in [0,1]$

$$\{r \in \mathbb{R} : U(r) \leq x\} = \underbrace{\{r \in \mathbb{R} : U(r) < 0\}}_{\substack{\in \mathcal{A},\ P_{0,1} \\ \| \\ \bigcup_{n=1}^{\infty} \{r \in \mathbb{R} : U(r) \leq -1/n\}}} \cup \{r \in \mathbb{R} : 0 \leq U(r) \leq x\}$$

$x \in [0,1)$

$\{r \in \mathbb{R} : 0 \leq U(r) \leq x\}$

Se $0 \leq U(r) \leq x$, então

$$r \sim_* U(r) t_0 + (1-U(r)) s_0$$

$$r \precsim^* x t_0 + (1-x) s_0$$

Assim, $0 \leq U(r) \leq x \Leftrightarrow r \precsim^* x t_0 + (1-x) s_0$.

Logo, $\{r \in \mathbb{R} : 0 \leq U(r) \leq x\} \in \mathcal{A}$.

3) $x > 1$

$$\{r \in R : U(r) \leq x\} = \{r \in R : U(r) \leq 1\} \cup \{r \in R : 1 < U(r) \leq x\}$$

$$\underset{\in \mathcal{A}.}{\uparrow}$$

Se $1 < U(r) \leq x$, então $t_0 \, \tau' \, r$ e, assim,

$$t_0 \sim \frac{1}{U(r)} r + \left(1 - \frac{1}{U(r)}\right) s_0$$

Como $U(r) \leq x \Leftrightarrow 1/x \leq 1/U(r)$, vale que

$$\frac{1}{x} r + \left(1 - \frac{1}{x}\right) s_0 \ \tau^* \ \underbrace{\frac{1}{U(r)} r + \left(1 - \frac{1}{U(r)}\right) s_0}_{t_0}.$$

Assim,

$$1 < U(r) \leq x \Leftrightarrow \frac{1}{x} r + \left(1 - \frac{1}{x}\right) s_0 \ \tau^* \ t_0$$

Logo

$$\{r \in R : 1 < U(r) \leq x\} \in \mathcal{A}$$

Logo

$$\{r \in R : U(r) \leq x\} \in \mathcal{A}$$

Assim, $\forall x \in \mathbb{R}$, $\{r \in R : U(r) \leq x\} \in \mathcal{A}$ e, portanto, U é variável aleatória.

Suponha $r_1, r_2 \in \mathbb{R}$ com $r_1 \precsim^* r_2$. Para $r \in [r_1, r_2]$, $\exists\, \alpha = \alpha(r) \in [0,1]$ tal que

$$r \sim^* \alpha(r) r_2 + (1-\underbrace{\alpha}_{\alpha(r)}) r_1. \qquad (\text{Teorema } 1)$$

Do Teorema 2, segue que

$$U(r) = \alpha(r)\, U(r_2) + (1-\alpha(r))\, U(r_1) \Rightarrow$$

$$\Rightarrow$$

$$\alpha(r) = \frac{U(r) - U(r_1)}{U(r_2) - U(r_1)}$$

Seja $\mathbb{P} \in \mathcal{P}_L$ ($\exists\, r'$ e $r'' \in \mathbb{R}$ t.q. $\mathbb{P}([r', r'']) = 1$).

$\forall r \in [r', r'']$, $\exists\, \alpha = \alpha(r) \in [0,1]$

$$r \sim^* \alpha(r) r'' + (1-\alpha(r)) r', \quad \text{onde} \quad \alpha(r) = \frac{U(r) - U(r')}{U(r'') - U(r')}$$

Seja $\beta = \int_{[r', r'']} \alpha(r)\, d\mathbb{P}(r)$ (2)

$$\sum_{r \in [r', r'']} \alpha(r)\, \mathbb{P}(R = r)$$

SU4. Seja $\mathbb{P} \in \mathcal{P}_L$ de modo que $\mathbb{P}([r_1', r_2']) = 1$. Para cada $r \in [r', r'']$, seja $\alpha(r)$ como dado em (1). Seja β como dado em (2). Então

$$\mathbb{P} \sim^* \beta r'' + (1-\beta) r'$$

Teorema. Sejam $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}_L$. Seja $\precsim^*$ definida sobre $\mathcal{P}_L \times \mathcal{P}_L$ atende SU1 a SU4. Então existe $U: \mathbb{R} \to \mathbb{R}$ tal que

$$\mathbb{P}_1 \precsim^* \mathbb{P}_2 \Leftrightarrow E(U \mid \mathbb{P}_1) \leq E(U \mid \mathbb{P}_2)$$

$$\left(E(U|\mathbb{P}_i) = \int_{\mathcal{R}} U \, d\mathbb{P}_i = \int_{\mathcal{R}} U(r) \, d\mathbb{P}_i(r) \, , \ i = 1, 2 \right)$$

Prova:

Como $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}_L$, existem $r_*, r^* \in \mathcal{R}$ tais que

$$\mathbb{P}_1([r_*, r^*]) = \mathbb{P}_2([r_*, r^*]) = 1.$$

$$\beta_1 = \int_{[r_*, r^*]} \alpha(r) \, d\mathbb{P}_1 \quad e \quad \beta_2 = \int_{[r_*, r^*]} \alpha(r) \, d\mathbb{P}_2 .$$

Da SU4,

$$\mathbb{P}_1 \sim^* \beta_1 r^* + (1-\beta_1) r_* \quad e \quad \mathbb{P}_2 \sim^* \beta_2 r^* + (1-\beta_2) r_*$$

Do Lema 3

$$\mathbb{P}_1 \precsim^* \mathbb{P}_2 \overset{\text{Lema 3}}{\Leftrightarrow} \beta_1 \leq \beta_2 \Leftrightarrow \int_{[r_*, r^*]} \alpha(r) \, d\mathbb{P}_1 \leq \int_{[r_*, r^*]} \alpha(r) \, d\mathbb{P}_2$$

$$\Leftrightarrow \int_{[r_*, r^*]} \frac{U(r) - U(r_*)}{U(r^*) - U(r_*)} \, d\mathbb{P}_1 \leq \int_{[r_*, r^*]} \frac{U(r) - U(r_*)}{U(r^*) - U(r_*)} \, d\mathbb{P}_2$$

$$\Leftrightarrow \frac{\int_{[r_*, r^*]} U(r) \, d\mathbb{P}_1 - U(r_*)}{U(r^*) - U(r_*)} \leq \frac{\int_{[r_*, r^*]} U(r) \, d\mathbb{P}_2(r) - U(r_*)}{U(r^*) - U(r_*)} \Rightarrow$$

$$\mathbb{P}_1 \precsim^* \mathbb{P}_2 \Leftrightarrow E(U|\mathbb{P}_1) \leq E(U|\mathbb{P}_2).$$

Teoria da Decisão

Aula 09

14/09

Teorema 3.

$\precsim^*$ em $\mathcal{P}_L \times \mathcal{P}_L$

$\precsim^*$ atende SU_2 a SU_4.

$\exists\, U: \mathcal{R} \to \mathbb{R}$ tal que $P_1 \precsim^* P_2 \Leftrightarrow E(U \mid P_1) \leq E(U \mid P_2)$

$r_1, r_2 \in \mathcal{R}$

$r_1 \precsim^* r_2 \quad \Leftrightarrow 1.U(r_1) \leq 1.U(r_2).$

$(P_{\{r_1\}} \precsim^* P_{\{r_2\}})$

Resultado. Sejam U_1 e U_2 funções de $\mathbb{R} \to \mathbb{R}$ satisfazendo as condições do Teorema 3. Então, $\exists\, a > 0$ e $b \in \mathbb{R}$ tais que $\forall r \in \mathcal{R}$,

$$U_2(r) = a\, U_1(r) + b.$$

Dem:

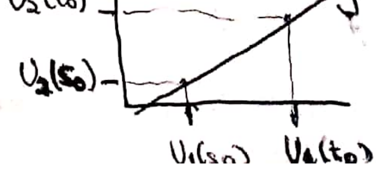
Sejam $U_1, U_2 : \mathcal{R} \to \mathbb{R}$ tais que $\forall\, P_1, P_2 \in \mathcal{P}_L$,

$$P_1 \precsim^* P_2 \Leftrightarrow E(U_i \mid P_1) \leq E(U_i \mid P_2),\ i = 1, 2.$$

Como existem $s_0, t_0 \in \mathcal{R}$ com $s_0 \prec^* t_0$, vale que

$$U_1(s_0) < U_1(t_0) \text{ e } U_2(s_0) < U_2(t_0)$$

$$y = U_2(s_0) + \frac{U_2(t_0) - U_2(s_0)}{U_1(t_0) - U_1(s_0)}\,(x - U_1(s_0))$$



- Para $r \in \{s_0, t_0\}$, a relação linear é verificada.

- Para $r \in \mathcal{R}$ tal que $s_0 \preceq^* r \preceq^* t_0$, $\exists\, \gamma \in (0,1)$ t.q.

$$r \sim^* \gamma t_0 + (1-\gamma) s_0 \text{ . Assim,}$$

$$U_1(r) = \gamma U_1(t_0) + (1-\gamma) U(s_0) \quad \text{e}$$

$$U_2(r) = \gamma U_2(t_0) + (1-\gamma) U_2(s_0)$$

$$U_2(s_0) + \frac{U_2(t_0) - U_2(s_0)}{U_1(t_0) - U_1(s_0)} \left[\gamma U_1(t_0) + (1-\gamma) U_1(s_0) - U_1(s_0)\right] =$$

$$= U_2(s_0) + \frac{U_2(t_0) - U_2(s_0)}{\cancel{U_1(t_0) - U_1(s_0)}} \cdot \gamma(\cancel{U_1(t_0) - U_1(s_0)})$$

$$= \gamma U_2(t_0) + (1-\gamma) U_2(s_0) = U_2(r).$$

- Para $r \prec^* s_0$, $\exists\, \gamma \in (0,1)$ t.q.

$$s_0 \sim^* \gamma t_0 + (1-\gamma) r \quad \text{. Assim,}$$

$$U_1(s_0) = \gamma U_1(t_0) + (1-\gamma) U_1(r) \quad \text{e}$$
$$U_2(s_0) = \gamma U_2(t_0) + (1-\gamma) U_2(r) \quad \text{, isto é,}$$

$$U_1(r) = \frac{U_1(s_0) - \gamma U_1(t_0)}{1-\gamma} \quad \text{e} \quad U_2(r) = \frac{U_2(s_0) - \gamma U_2(t_0)}{1-\gamma}.$$

$$U_2(s_0) + \frac{U_2(t_0) - U_2(s_0)}{U_1(t_0) - U_1(s_0)} \left[\frac{U_1(s_0) - \gamma U_1(t_0)}{1-\gamma} - U_1(s_0) \right] =$$

$$= U_2(s_0) + \frac{U_2(t_0) - U_2(s_0)}{U_1(t_0) - U_1(s_0)} \frac{(-\gamma)}{1-\gamma} \left[U_1(t_0) - U_1(s_0) \right]$$

$$= \frac{U_2(s_0) - \gamma U_2(t_0)}{1-\gamma} = U_2(r).$$

. Exercício

- (Terceiro caso). $r \succsim t_0$. Procedendo do mesmo modo, obtemos para $r \in \mathcal{R}$ t.q. $t_0 \precsim r$, que

$$U_2(r) = U_2(s_0) + \frac{U_2(t_0) - U_2(s_0)}{U_1(t_0) - U_1(s_0)} (U_1(r) - U_1(s_0)). \quad (*)$$

Assim, (*) vale para todo $r \in \mathcal{R}$ e basta tomar

$$a = \frac{U_2(t_0) - U_2(s_0)}{U_1(t_0) - U_1(s_0)} \text{ e } b = U_2(s_0) - U_1(s_0) \frac{U_2(t_0) - U_2(s_0)}{U_1(t_0) - U_1(s_0)}$$

Comentários.

(1) O desenvolvimento feito é restrito a medidas de probabilidade limitadas. Extensões podem ser obtidas (seção 7.10 De Groot discute a repres. para medidas de probabilidade para as quais $U: \mathcal{R} \to \mathbb{R}$ é integrável).

"A weaker system of axioms"
(H. Rubin)

(2) Construção dá sustentação ao paradigma de maximização de utilidade.

esperada.

(3) Representação de preferência por números (ou existência de função utilidade).

Paradoxo de ALLAIS (1953)

Problema 1:

P_1 : Prêmio U\$ 500 000,00 com probabilidade 1

P_2 : $\begin{cases} \text{U\$ } ~~\text{[illegible]}~~ \text{ 2 500 000,00}, & p = 0{,}1 \\ \text{U\$ 500 000,00}, & p = 0{,}89 \\ \text{U\$ 0}, & p = 0{,}01 \end{cases}$

Problema 2:

P_3 : $\begin{cases} \text{U\$ 500 000,00}, & p = 0{,}11 \\ \text{U\$ 0}, & p = 0{,}89 \end{cases}$

P_4 = $\begin{cases} \text{U\$ 2 500 000,00}, & p = 0{,}1 \\ \text{U\$ 0}, & p = 0{,}9 \end{cases}$

Prob 2 / Prob 1	P_3	P_4	
P_1	0	7	7
P_2	0	10	10
	0	17	17

$P_2 \precsim P_1 \Leftrightarrow E(U|P_2) \leq E(U|P_1) \Leftrightarrow$

$U(2500000)\frac{1}{10} + U(5000000)\frac{89}{1000}$

$+ U(0)\,{}^1/_{100} < U(500000)$

$\Leftrightarrow \frac{11}{100} U(500000) > \frac{10}{100} U(2\,500\,000) + \frac{1}{1000} U(0)$ (★)

$P_3 \prec^* P_4 \Leftrightarrow E(U|P_3) < E(U|P_4) \Leftrightarrow$

$$\Leftrightarrow U(500000)\frac{11}{100} + U(0)\frac{89}{1000} < U(2500000)\frac{10}{100} + U(0)\frac{90}{1000}$$

$$\Leftrightarrow \frac{11}{100}U(500000) < \frac{10}{100}U(2500000) + \frac{1}{100}U(0) \quad (**)$$

(4) Equivalência entre transformações lineares de utilidade. (último resultado visto)

(5) Supontamos que $R = \{r_1, r_2, r_3\}$ onde $r_1 \prec^* r_2 \prec^* r_3$.

$U'(r_1) \quad U'(r_2) \quad U'(r_3) \qquad U(r) = \dfrac{U'(r) - U'(r_1)}{U'(r_3) - U'(r_1)}$

$U(r_1) = 0$

$U(r_3) = 1$ Sejam P_1 e P_2 medidas sobre R.

$U(r_2) \in (0,1)$ $P_1 \precsim^* P_2 \Leftrightarrow E(U|P_1) \leq E(U|P_2)$, $P_1(r_i) = p_i, i=1,2,3$

$\Leftrightarrow 0p_1 + U(r_2)p_2 + p_3 \leq 0q_1 + U(r_2)q_2 + q_3$. $P_2(r_i) = q_i, i=1,2,3$

Então

$\Rightarrow P_1 \precsim^* P_2 \Leftrightarrow (p_3 - q_3) + U(r_2)(p_2 - q_2) \leq 0$.

No paradoxo de Allais(1953),

$R = \{0, 500000, 2500000\}$

No prob. 1,

$P_1, P_2 \rightarrow p_3 - q_3 = (0 - 0{,}1) = -0{,}1$

$p_2 - q_2 = 1 - 0{,}89 = 0{,}11$

No prob. 2

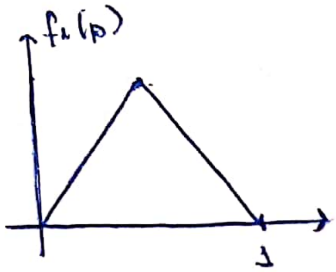
$P_3, P_4 \rightarrow p_3' - q_3' = (0 - 0{,}1) = -0{,}1$

$p_2' - q_2' = (0{,}11 - 0) = 0{,}11$

Versão contínua do Paradoxo de Allais

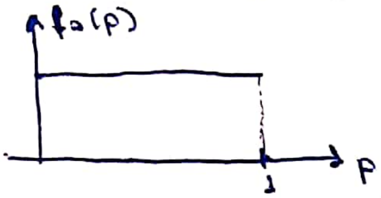
• Problema 1

P_1



$$f_1(p) = \begin{cases} 4p, & p \in [0, 1/2] \\ 4(1-p), & p \in (1/2, 1) \end{cases}$$

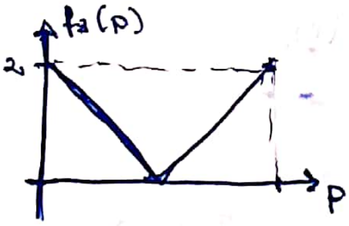
P_2



$f_2(p) = 1, \ 0 \le p \le 1.$

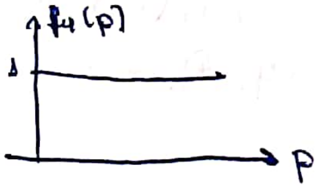
• Problema 2.

P_3



$f_3(p) = 4\,|p - 1/2|, \ p \in (0,1)$

P_4



$f_4(p) = 1, \ 0 \le p \le 1.$

Note que $P_1 \preceq' P_2 \Leftrightarrow P_4 \preceq' P_3$ (Supondo $U: (0,1) \to \mathbb{R}$ limitada)

Pode-se verificar que $\forall p \in (0,1)$,

$$f_1(p) - f_2(p) = f_4(p) - f_3(p)$$

Assim,

$E(U|P_1) - E(U|P_2) =$

$= \int_0^1 u(p) f_1(p)\,dp - \int_0^1 u(p) f_2(p)\,dp = \int_0^1 u(p)(f_1(p) - f_2(p))\,dp =$

$$= \int_0^1 u(p)\,(f_4(p) - f_3(p))\,dp = E(U \mid P_4) - E(U \mid P_3)$$

Assim,

$$E(U \mid P_1) - E(U \mid P_2) = E(U \mid P_4) - E(U \mid P_4)$$

Lista 2 no xerox!

Teoria da Decisão, Aula 11

21/09

$L(d,\theta) = -U(r(d,\theta))$

↳ perda (penalidade) na qual o indivíduo incorre ao escolher $d \in D$ quando o estado da natureza é $\theta \in \Theta$.

Para $d \in D$,

$$\rho(d, \mathbb{P}) = \int_{\Theta} L(d,\theta)\, d\mathbb{P}(\theta)$$ ⇒ risco da decisão d.

$$\rho^*(\mathbb{P}) = \inf\{\rho(d,\mathbb{P}) : d \in D\}$$ ⇒ risco de Bayes

$d^* \in D$ é uma decisão de Bayes se

$$\rho^*(d^*, \mathbb{P}) = \rho^*(\mathbb{P}).$$

Exemplo: $\Theta = \{0,1\}$, $\mathbb{P}_1(\theta=1) = 1/4$

$D = [0,1]$

$L(d,\theta) = |d-\theta|$

$$\rho(d,\mathbb{P}_1) = \int_{\Theta} L(d,\theta)\, d\mathbb{P}_1(\theta) = L(d,0)\mathbb{P}_1(\theta=0) + L(d,1)\mathbb{P}_1(\theta=1)$$

$$= |d-0| \cdot \frac{3}{4} + |d-1| \frac{1}{4}$$

$$= \frac{3}{4}d + (1-d)\frac{1}{4} = \frac{1}{4} + \frac{d}{2}.$$

$$\rho(d,\mathbb{P}_1) = \frac{1}{4} + \frac{d}{2}.$$

$\therefore \rho^*(\mathbb{P}_1) = \frac{1}{4}$. Como $\rho(0,\mathbb{P}_1) = 1/4$, 0 é decisão de Bayes.

Suponhamos ainda nesse exemplo $\mathbb{P}_2$ tal que $\mathbb{P}_2(\theta=1) = 4/5$.

$$\rho(d, \mathbb{P}_2) = \frac{1}{5} d + \frac{4}{5}(1-d) = \frac{4}{5} - \frac{3}{5} d$$

$$\rho^*(\mathbb{P}_2) = 1/5.$$

Como $\rho(1, \mathbb{P}_2) = \rho^*(\mathbb{P}_2)$, 1 é decisão de Bayes contra $\mathbb{P}_2$.

$$\mathbb{P}(\Theta=1) - (\mathbb{P}(\Theta=0) - \mathbb{P}(\Theta=1))d = \rho(d, \mathbb{P}).$$

Se $D = (0,1)$, então

$\rho^*(\mathbb{P}_1) = 1/4$, no entanto não existe decisão de Bayes.

$\rho^*(\mathbb{P}_2) = 1/5$, " " " " " " "

Ainda no Exemplo 1, considere $L(d,\theta) = (d-\theta)^2$, $D = [0,1]$, $\mathbb{P}_1$. Para $d \in D$.

$$\rho(d, \mathbb{P}_1) = (d-0)^2\, \mathbb{P}_1(\Theta=0) + (d-1)^2\, \mathbb{P}_1(\Theta=1)$$
$$= d^2 \cdot \frac{3}{4} + (1-d)^2 \frac{1}{4} \Rightarrow$$

$\rho(d, \mathbb{P}_1) = d^2 - \frac{d}{2} + \frac{1}{4}$, que atinge valor mínimo em $d^* = 1/4$.

Logo, $d^* = 1/4$ é a decisão de Bayes.

$d \in D$ Cada $d \in D$, induz

d $1-d$ um med.

$3/4$ $1/4$

$(\Theta, \mathcal{S}, \mathbb{P}) \to (\Omega, \mathcal{A}, \mathbb{P}_d)$

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

Comentários

(1) $D = \{$ ENCAPAR, NÃO ENCAPAR $\}$

$\quad\quad\quad\quad\quad\downarrow d_1 \quad\quad\quad \downarrow d_2$

$U(d,\theta)$

	$\theta=1$ (vende)	$\theta=0$ (não vende)
d_1	$v-c$	$-c$
d_2	v	0

v: preço máximo de venda

c: preço do custo p/ encapar

$P_{d_1}(\theta=1) = 9/10$

$P_{d_2}(\theta=1) = 2/10$

$$E(U|P_1) = (v-c)\frac{9}{10} + (-c)\frac{1}{10} = \frac{9v-10c}{10}$$

$$E(U|P_2) = v \cdot \frac{2}{10} + 0 \cdot \frac{8}{10} = \frac{2v}{10}$$

Encapar é a melhor ~~decisão~~ que não encapar se, e só se,

$$\frac{9v-10c}{10} > \frac{2v}{10} \iff c < \frac{7}{10}v.$$

(2) $D, \mathbb{P}, L$.

Seja $\lambda : \Theta \to \mathbb{R}$ tal que $\int_\Theta \lambda(\theta)\, d\mathbb{P}(\theta) < \infty$.

$$L' : D \times \Theta \to \mathbb{R}$$
$$(d,\theta) \longmapsto L'(d,\theta) = L(d,\theta) + \lambda(\theta)$$

$$\int_\Theta L'(d,\theta)\, d\mathbb{P} = \int_\Theta [L(d,\theta) + \lambda(\theta)]\, d\mathbb{P}(\theta) = \int_\Theta L(d,\theta)\, d\mathbb{P}(\theta) + \int_\Theta \lambda(\theta)\, d\mathbb{P}(\theta)$$

Para todo $a > 0$ e $\lambda : \Theta \to \mathbb{R}$ tal que $\int_\Theta \lambda(\theta)\, d\mathbb{P}(\theta) < \infty$, temos

$\forall\, d_1, d_2 \in D$, que

$$\int_\Theta L(d_1,\theta)\,d\mathbb{P}(\theta) \leq \int_\Theta L(d_2,\theta)\,d\mathbb{P}(\theta) \Leftrightarrow$$

$$\int_\Theta \overbrace{(aL(d,\theta) + \lambda(\theta))}^{L'(d_1,\theta)}\,d\mathbb{P}(\theta) \leq \int_\Theta \overbrace{(aL(d_2,\theta) + \lambda(\theta))}^{L^2(d_2,\theta)}\,d\mathbb{P}(\theta)$$

Em particular, se para $\theta \in \Theta$, $L(\cdot,\theta)$ é limitada inferiormente, podemos considerar

$$\lambda_0(\theta) = \inf\{L(d,\theta) : d \in D\}$$

Assim, definindo

$$L_0(d,\theta) = L(d,\theta) - \lambda_0(\theta), \text{ temos } L_0(d,\theta) \geq 0,\ \forall d, \forall \theta.$$

Do comentário anterior,

$$L_0(d,\theta) \geq 0, \ \forall (d,\theta) \in D \times \Theta \text{ e}$$

$$\int_\Theta L(d_1,\theta)\,d\mathbb{P}(\theta) \leq \int_\Theta L(d_2,\theta)\,d\mathbb{P}(\theta) \Leftrightarrow$$

$$\int_\Theta L_0(d_1,\theta)\,d\mathbb{P}(\theta) \leq \int_\Theta L_0(d_2,\theta)\,d\mathbb{P}(\theta), \ \forall d_1, d_2 \in D.$$

Exemplo: $D = \{d_1, d_2, d_3, d_4\}$, $\Theta = \{\theta_1, \theta_2, \theta_3\}$

$L(d,\theta)$	θ_1	θ_2	θ_3
d_1	2	3	9
d_2	-2	7	-3
d_3	10	7	-4
d_4	-3	8	0

$\lambda_0(\theta_1) = -3$

$\lambda_0(\theta_2) = 3$

$\lambda_0(\theta_3) = -4$

$L_0(d,\theta)$	θ_1	θ_2	θ_3
d_1	5	0	13
d_2	1	4	1
d_3	13	4	0
d_4	0	5	4

(3) Vimos ~~que~~, no Exemplo 1, que $\rho^*(\mathbb{P}_1) = 1/4$ $(\mathbb{P}_1(\theta=1) = 1/4)$ e $\rho^*(\mathbb{P}_2) = 1/5$ $(\mathbb{P}_2(\theta=1) = 4/5)$.

Fixado Θ, $\mathcal{F}$ e L, consideremos $\mathcal{P}_L$ o conjunto de todas as medidas de probabilidade em $(\Theta, \mathcal{F})$.

$\mathcal{P}_L$ é um conjunto convexo, pois $\forall \mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}_L$ e $\alpha \in [0,1]$,

$$\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2 \in \mathcal{P}_L.$$

Vamos olhar a transformação

$$\rho^* : \mathcal{P}_L \longrightarrow \mathbb{R}$$
$$\mathbb{P} \in \mathcal{P}_L \longmapsto \rho^*(\mathbb{P}) = \inf\{\rho(d,\mathbb{P}) : d \in D\}$$

Resultado. ρ^* é concava.

Sejam $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}_L$ e $\alpha \in [0,1]$

Para $d \in D$,

$$\rho(d, \alpha\mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) = \int_\Theta L(d,\theta)\, d(\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2)$$
$$= \alpha \int_\Theta L(d,\theta)\, d\mathbb{P}_1 + (1-\alpha)\int_\Theta L(d,\theta)\, d\mathbb{P}_2$$
$$= \alpha \rho(d,\mathbb{P}_1) + (1-\alpha)\rho(d,\mathbb{P}_2)$$

$$\rho^*(\alpha\mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) = \inf\{\rho(d, \alpha\mathbb{P}_1 + (1-\alpha)\mathbb{P}_2 : d \in D\}$$
$$= \inf\{\alpha\rho(d,\mathbb{P}_1) + (1-\alpha)\rho(d,\mathbb{P}_2) : d \in D\}$$
$$\geq \inf_{d\in D}\{\alpha\rho(d,\mathbb{P}_1)\} + \inf\{(1-\alpha)\rho(d,\mathbb{P}_2) : d \in D\}$$

$$= \alpha \rho^*(\mathbb{P}_1) + (1-\alpha)\rho^*(\mathbb{P}_2)$$

Logo,

$$\rho^*(\alpha \mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) \geq \alpha \rho^*(\mathbb{P}_1) + (1-\alpha)\rho^*(\mathbb{P}_2).$$

Exemplo: $D = [0,1]$

$\Theta = \{0,1\}$, $\mathcal{F} = \mathcal{P}(\{0,1\})$

$\mathcal{P}_L = \{ Ber(p) ; p \in [0,1] \}$

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

Seja $d \in D$ e $p \in [0,1]$. $\mathbb{P}_p$

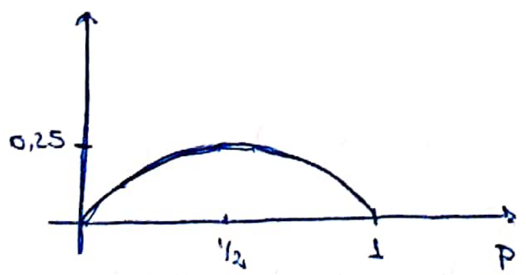
$$\rho(d, \mathbb{P}_p) = (d-0)^2 \, \mathbb{P}_p(\theta = 0) + (d-1)^2 \, \mathbb{P}_p(\theta = 1) =$$

$$= d^2(1-p) + (1-d)^2 p \Rightarrow$$

$$d^2 - pd^2 + p - 2pd + pd^2 = d^2 - 2pd + p$$

$\rho(d, \mathbb{P}_p) = d^2 - 2pd + p \Rightarrow d^* = p$ é decisão de Bayes contra $\mathbb{P}_p$.

$$\rho^*(\mathbb{P}_p) = p^2 - 2p^2 + p = p - p^2 = p(1-p).$$



ρ^* é estritamente côncava.

No exemplo anterior, considere agora $L(d,\theta) = |d-\theta|$.

Para $d \in D$,

$$\rho(d, \mathbb{P}_p) = |d-0| \, \mathbb{P}_p(\theta=0) + |d-1| \, \mathbb{P}_p(\theta = 1)$$

$$= d(1-p) + (1-d)p$$

$$= p + d(1-2p)$$

$p < 1/2 \xRightarrow{\text{reta crescente}} d^* = 0$ é decisão de Bayes contra $\mathbb{P}_p$

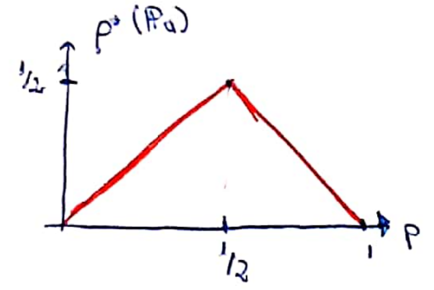
$p > 1/2 \xRightarrow{\text{decre}} d^* = 1$ " " " " " "

$p = 1/2 \xrightarrow{\text{const.}}$ toda $d \in D$ é decisão de Bayes

$p < 1/2$, $\rho^*(\mathbb{P}_d) = \rho(0, \mathbb{P}_d) = p$

$p > 1/2$, $\rho^*(\mathbb{P}_d) = \rho(1, \mathbb{P}_d) = p + 1 - 2p = 1 - p$

$p = 1/2$, $\rho^*(\mathbb{P}_d) = \rho\, 1/2$



Teoria da Decisão

Aula 12

23/09/15

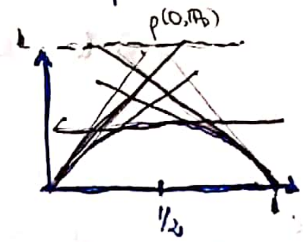
Ex. 1: $\Theta = \{0, 1\}$

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

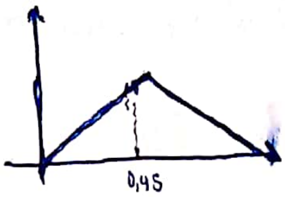
$D = [0, 1]$

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

$\rho^*(\mathbb{P}_p) = p(1-p)$



(B) $L(d, \theta) = |d - \theta|$



$$d^* = \begin{cases} 1, & p > 1/2 \\ 0, & p \leq 1/2 \end{cases}$$

Suponha que em (A), você acha que $\Theta = 1$ com probabilidade 0,3. Você declara $p' = 0,2$.

Risco "Verdadeiro": $\rho^*(0,3) = 0,3 \times 0,7 = 0,21$

Risco Incorrido: $\rho(0,2, 0,3) = \underset{d}{(0,2)^2} + (0,3)[1 - 2(0,2)] = 0,22$

Em (B), $\rho(d, \mathbb{P}_p) = d(1-p) + (1-d)p = d + p(1-2d)$ e

$\rho^*(0,3) = 0,3$

$\rho(0, 0,3) = 0,3$

Suponha agora que você acha que $\Theta = 1$ com probabilidade 0,45. Você declara $p' = 0,55$. Nesse caso

$\rho^*(0,45) = 0,45$

$\rho(1, 0,45) = 0,55$

Exemplo 2. $\Theta = \{0, 1, 2\}$

$\mathbb{P}_{(p_1, p_2)}(\Theta = 1) = p_1$; $\mathbb{P}_{(p_1, p_2)}(\Theta = 2) = p_2$ $\quad D = \mathbb{R}_+$

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

$$\rho(d, \mathbb{P}_{(p_1,p_2)}) = \mathbb{E}((d-\theta)^2 \mid \mathbb{P}_{(p_1,p_2)}) =$$

$$= \mathbb{E}\left((d - E(\theta \mid \mathbb{P}_{(p_1,p_2)}) + E(\theta \mid \mathbb{P}_{(p_1,p_2)}) - \theta)^2 \mid \mathbb{P}_{(p_1,p_2)}\right)$$

$$= (d - E(\theta \mid \mathbb{P}_{(p_1,p_2)}))^2 + \mathbb{E}\left((\theta - E(\theta \mid \mathbb{P}_{(p_1,p_2)}))^2 \mid \mathbb{P}_{(p_1,p_2)}\right)$$

$$+ 2\,(d - E(\theta \mid \mathbb{P}_{(p_1,p_2)}))\, \underbrace{\mathbb{E}\left((\theta - E(\theta \mid \mathbb{P}_{(p_1,p_2)}))^2 \mid \mathbb{P}_{(p_1,p_2)}\right)}_{0}$$

$\therefore$

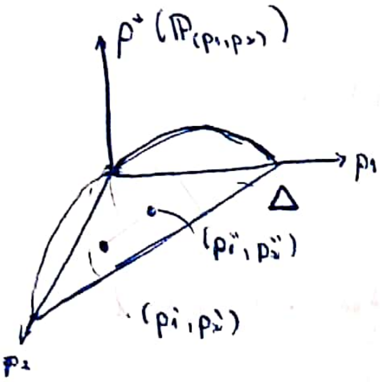
$$\Rightarrow \rho(d, \mathbb{P}_{(p_1,p_2)}) = \mathrm{Var}(\theta \mid \mathbb{P}_{(p_1,p_2)}) + (d - E(\theta \mid \mathbb{P}_{(p_1,p_2)}))^2$$

$\rho(d, \mathbb{P}_{(p_1,p_2)})$ atinge valor mínimo em $d^* = E(\theta \mid \mathbb{P}_{(p_1,p_2)})$.

Assim,

$$\rho^*(\mathbb{P}_{(p_1,p_2)}) = \rho(E(\theta \mid \mathbb{P}_{(p_1,p_2)}), \mathbb{P}_{(p_1,p_2)}) = \mathrm{Var}(\theta \mid \mathbb{P}_{(p_1,p_2)})$$

$$= 1p_1 + 4p_2 - (1p_1 + 2p_2)^2$$



Ou Hessiano, ou

$$(p_1', p_2'), (p_1'', p_2'') \in \Delta \{(q_1, q_2) \in \mathbb{R}_+^2 : q_1 + q_2 \leq 1\}$$

$$\rho^*(\mathbb{P}_{\alpha(p_1',p_2') + (1-\alpha)(p_1'',p_2'')}) \geq \alpha\, \rho^*(\mathbb{P}_{(p_1',p_2')}) + (1-\alpha)\, \rho^*(\mathbb{P}_{(p_1'',p_2'')})$$

$$= \alpha(1-\alpha)\left[(p_1' + 2p_2') - (p_1'' + 2p_2'')\right]^2 \ldots$$

Livros de Análise Convexa: Rockfellar, Análise Convexa

Schervish, Ferguson (Teoria da Decisão)
(Teoria da Est.)

$\mathcal{D} = \{d_1, d_2, \dots\}$

Definição: Ao procedimento que consiste em tomar a decisão $d_i \in \mathcal{D}$ com probabilidade α_i, $i \geq 1$, damos o nome de decisão ALEATORIZADA ou MISTA associada à medida $(\alpha_n)_{n \geq 1}$.

Notação:

$d_i \in \mathcal{D}$ é "aleatorizada" segundo uma medida degenerada em $\{i\}$. → Decisão Pura ou Não Aleatorizada

"$d_{(\alpha_n)_{n \geq 1}} = \sum_{i=1}^{\infty} \alpha_i d_i$"

$\mathcal{M}_0 = \{ d_{(\alpha)} : \alpha$ é medida de prob. em $(\mathcal{D}, \mathcal{P}(\mathcal{D}))\}$

Claro que $\mathcal{D} \subseteq \mathcal{M}_0$.

Fixada $\mathbb{P}$ sobre $(\Theta, \mathcal{F})$, $\inf\{\rho(d, \mathbb{P}) : d \in \mathcal{M}_0\} \leq \inf\{\rho(d, \mathbb{P}) : d \in \mathcal{D}\}$, onde para $d \in \mathcal{M}_0$, definimos

$$L(d_{(\alpha)}, \theta) = \sum_{n=1}^{\infty} \alpha_n L(d_n, \theta) \qquad (L \text{ estendida a } \mathcal{M}_0 \times \Theta)$$

e

$$\rho(d_{(\alpha)}, \mathbb{P}) = \int_{\Theta} L(d_{(\alpha)}, \theta)\, d\mathbb{P}(\theta)$$

Por outro lado, para $\underbrace{d}_{d_{(\alpha)}} \in \mathcal{M}_0$,

$$\rho(d_{(\alpha)}, \mathbb{P}) = \int_{\Theta} L(d_{(\alpha)}, \theta)\, d\mathbb{P}(\theta)$$

$$= \int_{\Theta} \left(\sum_{n=1}^{\infty} \alpha_n L(d_n, \theta) \right) d\mathbb{P}(\theta)$$

$$= \sum_{n=1}^{\infty} \alpha_n \int_{\Theta} L(d_n, \theta)\, dP(\theta)$$

$$= \sum_{n=1}^{\infty} \alpha_n \rho(d_n, P) \geq \sum_{n=1}^{\infty} \alpha_n \inf\{\rho(d, P) : d \in D\} =$$

$$= \inf\{\rho(d, P) : d \in D\}$$

Como para toda $d \in \mathcal{M}_0$,

$$\rho(d, P) \geq \inf\{\rho(d, P) : d \in D\}$$

Segue que

$$\inf\{\rho(d, P) : d \in \mathcal{M}_0\} \geq \inf\{\rho(d, P) : d \in D\}$$

Vamos considerar D e Θ finitos.

$D = \{d_1, \ldots, d_m\}$, $\Theta = \{\theta_1, \ldots, \theta_k\}$

Para $d \in D$, vamos considerar

$$\ell(d) = (L(d, \theta_1), L(d, \theta_2), \ldots, L(d, \theta_k))$$

Do mesmo modo, para $d_{(\alpha)} \in \mathcal{M}_0$,

$$\ell(d_{(\alpha)}) = \left(\sum_{j=1}^{m} \alpha_j L(d_j, \theta_1), \sum_{j=1}^{m} \alpha_j L(d_j, \theta_2), \ldots, \sum_{j=1}^{m} \alpha_j L(d_j, \theta_k)\right)$$

$$\ell(\mathcal{M}_0) = \{\ell(d) : d \in \mathcal{M}_0\}$$

Exemplo 1:

$k=2$ $\Theta = \{\theta_1, \theta_2\}$

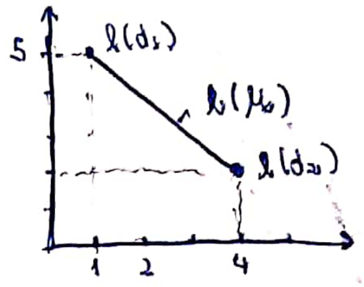
$m=2$ $\mathcal{D} = \{d_1, d_2\}$

$L(d,\theta)$

	d_1	d_2
θ_1	1	4
θ_2	5	2

$\ell(d_1) = (1,5)$

$\ell(d_2) = (4,2)$



$L(d_{(\alpha)}, \theta_1) = \alpha L(d_1, \theta_1) + (1-\alpha) L(d_2, \theta_2)$

Exemplo 2:

$k=2$, $m=4$, $\Theta = \{\theta_1, \theta_2\}$, $\mathcal{D} = \{d_1, d_2, d_3, d_4\}$

$L(d,\theta)$

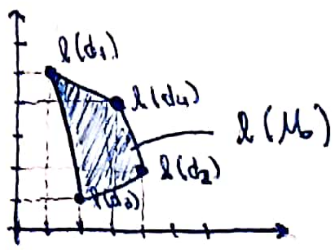
	d_1	d_2	d_3	d_4
θ_1	1	4	2	3
θ_2	5	2	1	4

$\ell(d_1) = (1,5)$

$\ell(d_2) = (4,2)$

$\ell(d_3) = (2,1)$

$\ell(d_4) = (3,4)$



vértices do polígono são chamado de pontos extremais envoltório convexo

Teoria da Decisão, Aula 14

Alguns outros resultados.

Resultado: d é a única decisão de Bayes contra $\mathbb{P}$ $\Rightarrow$ d é admissível

Suponha d decisão de Bayes contra $\mathbb{P}$ e δ inadmissível.
Então, $\exists\, d' \in \mathcal{D}$ tal que

$$L(d',\theta) \leq L(d,\theta),\ \forall \theta \in \Theta \text{ e } \exists\, \theta_0 \in \Theta \text{ tal que}$$

$$L(d',\theta_0) < L(d,\theta_1)$$

$$\rho(d',\mathbb{P}) = \int_\Theta L(d',\theta)\, d\mathbb{P} \leq \underbrace{\int_\Theta L(d,\theta)\, d\mathbb{P}}_{\rho^*(\mathbb{P})}$$

Logo, d' é decisão de Bayes e, portanto,

$$\{d,d'\} \subseteq \{d^* \in \mathcal{D} : \rho(d^*,\mathbb{P}) = \rho^*(\mathbb{P})\}$$

Resultado: Seja $\mathcal{D}_{\mathbb{P}} = \{d \in \mathcal{D} : \rho(d,\mathbb{P}) = \rho^*(\mathbb{P})\} \neq \emptyset$.
$\forall d_1, d_2 \in \mathcal{D}_{\mathbb{P}}$,

$$L(d_1,\cdot) = L(d_2,\cdot) \Rightarrow \forall d \in \mathcal{D}_{\mathbb{P}},\ d \text{ é admissível}$$

$\exists\, d' \in \mathcal{D}$ tal que

$$L(d',\theta) \leq L(d^*,\theta),\ \forall \theta \in \Theta \text{ e}$$
$$L(d',\theta_0) < L(d^*,\theta_0), \text{ para } \theta_0 \in \Theta$$

Do argumento exposto no resultado anterior, segue que $d' \in D_{\mathbb{P}}$. Logo, d' e $d^* \in D_{\mathbb{P}}$ são tais que

$$L(d', \Theta_0) \neq L(d^*, \Theta_0)!$$

Resultado: Suponha que $L(d, \cdot)$ é contínua para todo $d \in D$. Se $d^* \in D$ é de bayes contra algum $\mathbb{P}$ tal que $\mathbb{P}(A) > 0$, $\forall A \in \mathcal{F}$ dado, então d^* é admissível.

Suponhamos $d^* \in D$ inadmissível. Então $\exists d' \in D$ tal que

$$L(d', \Theta) \leq L(d^*, \Theta) \quad \forall \Theta \in \Theta \quad \text{e}$$

$$L(d', \Theta_0) < L(d^*, \Theta_0), \text{ p/ algum } \Theta_0 \in \Theta.$$

Seja $\varepsilon = L(d^*, \Theta_0) - L(d', \Theta_0) > 0$. Como $L(d', \cdot)$ é contínua, $\exists \delta_1 > 0$ tal que

$$|\Theta - \Theta_0| < \delta_1 \Rightarrow |L(d', \Theta) - L(d', \Theta_0)| < \varepsilon/4 \Rightarrow$$

$$L(d', \Theta) < L(d', \Theta_0) + \varepsilon/4. \qquad \text{(I)}$$

Do mesmo modo, $L(d^*, \cdot)$ é contínua e $\exists \delta_2 > 0$ tal que
$|\Theta - \Theta_0| < \delta_2 \Rightarrow |L(d^*, \Theta) - L(d^*, \Theta_0)| < \varepsilon/4 \Rightarrow$

$$L(d^*, \Theta) > L(d^*, \Theta_0) - \varepsilon/4 \qquad \text{(II)}$$

De (I) e (II)

$$|\Theta - \Theta_0| < \min\{\delta_1, \delta_2\} \Rightarrow \overset{\overset{L(d^*,\Theta_0)+\varepsilon/4}{\uparrow}}{L(d^*, \Theta)} - \overset{\overset{L(\delta^*,\Theta_0)+\varepsilon/4}{\uparrow}}{L(d', \Theta)} > \frac{\varepsilon}{2}$$

Seja $\mathbb{P}$ uma probabilidade tal como no enunciado. Assim,

$$\rho(d^*,\mathbb{P}) - \rho(d',\mathbb{P}) = \int_{\Theta} L(d^*,\theta)\,d\mathbb{P} - \int_{\Theta} L(d',\theta)\,d\mathbb{P}$$

$$= \int_{\Theta} [L(d^*,\theta) - L(d',\theta)]\,d\mathbb{P} = \int_{|\theta-\theta_0| \geq \underbrace{\min\{\delta_1,\delta_2\}}_{\geq 0}} [L(d^*,\theta)-L(d',\theta)]\,d\mathbb{P}$$

$$+ \int_{\underbrace{|\theta-\theta_0| < \min\{\delta_1,\delta_2\}}} \overbrace{[L(d^*,\theta)-L(d',\theta)]}^{> \varepsilon/2}\,d\mathbb{P} \geq \varepsilon/2 .$$

$$\geq \frac{\varepsilon}{2}\, \mathbb{P}(|\theta-\theta_0| < \min\{\delta_1,\delta_2\})$$

$$\geq \frac{\varepsilon}{2}\, \mathbb{P}(|\theta-\theta_0| < \min\{\delta_1,\delta_2\}) > 0$$

Logo, d^* não é decisão de Bayes contra $\mathbb{P}$.

$(\mathcal{D}, \Theta, L, \mathbb{P})$ → prob. de decisão

$\mathbb{P}$: incerteza; L: perda; Θ: parâmet.; $\mathcal{D}$: decisões

$\forall\, d \in \mathcal{D}$, calculo

$$\rho(d,\mathbb{P}) = \int_{\Theta} L(d,\theta)\,d\mathbb{P}$$

e tomo d^* tal que $\rho(d^*,\mathbb{P}) = \inf_{d\in\mathcal{D}} \rho(d,\mathbb{P})$

Exemplo 3.

	d_1	d_2	d_3	d_4	d_5	d_6	d_7
θ_1	0	0	2	3	4	10	6
θ_2	10	7	4	5	2	2	6

$\mathbb{P}(\theta=\theta_1) = p$, $\mathbb{P}(\theta=\theta_2) = 1-p$.

Decisão de Bayes

$p \leq 1/2 \rightarrow d_5$

$\frac{1}{2} \leq p \leq \frac{3}{5} \rightarrow d_3$

$p \geq 3/5 \rightarrow d_2$

Problemas de Decisão Com Dados, (Observações)

No exemplo 3, considere o experimento que consiste em observar X.

Suponhamos que $X / \Theta = \Theta_1 \sim \text{Bernoulli}(4/5)$
$X / \Theta = \Theta_2 \sim \text{Bernoulli}(1/3)$

Vamos considerar dois cenários.

- Pré-Experimentação
- Pós-Experimentação

CASO PÓS-EXPERIMENTAÇÃO:

$X = 0$. Temos então o problema de decisão $(D, \Theta, L, P_{\{X=0\}})$, onde $P_{\{X=0\}}$ é a medida de probabilidade sobre Θ dado $X=0$.

Pelo Teorema de Bayes,

$$P(\Theta = \Theta_1 / X = 0) = \frac{P(X=0 / \Theta = \Theta_1)\, P(\Theta = \Theta_1)}{\sum_{i=1}^{2} P(X=0 / \Theta = \Theta_i)\, P(\Theta_i)}$$

$$= \frac{1/5 \cdot p}{1/5 \cdot p + 2/3 (1-p)} = \frac{3p}{3p + 10(1-p)}$$

Para $d_2 \in D$, temos

$$\rho(d_2, \mathbb{P}_{\{X=0\}}) = \int_{\Theta} L(d_2, \theta)\, d\mathbb{P}_{\{X=0\}}(\theta) =$$

$$= L(d_2, \theta_1)\,\mathbb{P}(\theta = \theta_1 / X=0) + L(d_2, \theta_2)\,\mathbb{P}(\theta = \theta_2 / X=0)$$

$$= 0 \cdot \frac{3p}{3p + 10(1-p)} + 7 \cdot \frac{10(1-p)}{3p + 10(1-p)}$$

$\mathbb{P}_{\{X=0\}}(\theta)$

Como $\mathbb{P}(\theta = \theta_1 / X=0) = \dfrac{3p}{10-7p}$, temos que d_5 será decisão de Bayes contra $\mathbb{P}_{\{X=0\}}$ se

$$\mathbb{P}(\theta = \theta_1 / X = 0) = \frac{3p}{10-7p} \leq 1/2$$

$$\Leftrightarrow 6p \leq 10 - 7p \Leftrightarrow p \leq 10/13$$

Do msm modo, d_2 será decisão de Bayes contra $\mathbb{P}_{\{X=0\}}$ se

$$\frac{3p}{10-7p} \geq \frac{3}{5} \Leftrightarrow 15p \geq 30 - 21p \Leftrightarrow p \geq \frac{5}{6}.$$

Assim, temos

$p \leq \dfrac{10}{13}$, a decisão de Bayes contra $\mathbb{P}_{\{X=0\}}$ será d_5;

$\dfrac{10}{13} \leq p \leq \dfrac{5}{6}$, d_3 " " " "

$p \geq \dfrac{5}{6}$, d_2 será " " " " " ".

$X=1 \quad (P_{\{X=1\}})$

$$P(\theta=\theta_1 / X=1) = \frac{P(X=1/\theta=\theta_1)\,P(\theta=\theta_1)}{\sum_{i=1} P(X=1/\theta=\theta_i)\,P(\theta=\theta_i)}$$

$$= \frac{4/5\,p}{4/5\,p + \frac{1}{8}(1-p)} = \frac{12p}{12p+5(1-p)}$$

$$= \frac{12p}{5+7p}$$

Como $\overset{P_{\{X=1\}}}{P(\theta=\theta_1 / X=1)} = \frac{12p}{5+7p}$, temos que

d_5 será a decisão de Bayes contra $P_{\{X=1\}}$ se

$$P(\theta=\theta_1/X=1) = \frac{12p}{5+7p} \leq \frac{1}{2} \Leftrightarrow 24p \leq 5+7p \Leftrightarrow p \leq \frac{5}{17}$$

Do mesmo modo, d_2 será a decisão de Bayes contra $P_{\{X=1\}}$ se

$$\frac{12p}{5+7p} > \frac{3}{5} \Leftrightarrow 60p > 15+21p \Leftrightarrow p > \frac{5}{13}$$

Assim, temos

$p \leq \frac{5}{17}$, d_5 será a decisão de Bayes contra $P_{\{X=1\}}$

$\frac{5}{17} < p \leq \frac{5}{13}$, d_3 " " " " " " $P_{\{X=1\}}$

$p > \frac{5}{13}$, d_2 " " " " " $P_{\{X=1\}}$

Caso Pré-Experimentação.

Suponha que o exp. não foi conduzido e que X é também incerto (desconhecido).

Nesse caso, há incerteza sobre (θ, x), $\theta \in \Theta$, x possível observação do experimento. Além disso, devemos especificar agora uma regra de decisão (função) que associa a cada x uma decisão $d \in D$.

Seja $\mathcal{X}$ o cjto dos possíveis resultados do experimento.
(No exemplo 3, $\mathcal{X} = \{0,1\}$).

⋮

Temos assim um novo problema de decisão. Seja

$$\Delta = \{\delta : \mathcal{X} \to D\}$$

o conjunto de regra de decisões. Seja

$$\Omega = \Theta \times \mathcal{X}, \text{ munido de } \sigma(\Theta, \mathcal{X})$$

Seja $\mathbb{P}' : \sigma(\Theta \times \mathcal{X}) \to [0,1]$ uma probabilidade.
Como antes, seja $L : D \times \Theta \to \mathbb{R}$ e definamos

$$L' : \Delta \times \Omega \to \mathbb{R} \quad \text{por.}$$

Para $\delta \in \Delta$ e $(\theta, x) \in \Omega$, $L'(\delta, (\theta, x)) = L(\delta(x), \theta)$.

$$(\Delta, \Omega, L', \mathbb{P}')$$

Para cada $\delta \in \Delta$, seja

$$\bar{\rho}(\delta, \mathbb{P}') = \int_{\Omega} L'(\delta, (\theta, x))\, d\mathbb{P}'(\theta, x)$$ o Risco da Decisão

Função de Decisão δ contra $\mathbb{P}'$.

Do mesmo modo,

$$\bar{\rho}^*(\mathbb{P}') = \inf\{\bar{\rho}(\delta, \mathbb{P}') : \delta \in \Delta\}$$ é chamado Risco de Bayes

Teoria da Decisão

Aula 15 07/10/2015

- Problema Pós-Experimentação

Após observar $X=x$, você resolve o problema $(\mathcal{D}, \Theta, L, \mathbb{P}_{X=x})$

- Problema Pré-Experimentação

$$\Delta = \{\delta : \mathfrak{X} \to \mathcal{D}\}$$
↳ funções regra de decisão

$\Omega = \Theta \times \mathfrak{X}$ $\quad\quad$ $\mathbb{P}'$ sobre $\sigma(\Theta \times \mathfrak{X})$

$L' : \Delta \times \Omega \to \mathbb{R}$
$(\delta, (\theta, x)) \mapsto L'((\delta, (\theta, x))) = L(\delta(x), \theta)$

Para cada $\delta \in \Delta$,

$$\rho'(\delta, \mathbb{P}') = \int_{\Theta \times \mathfrak{X}} L'((\delta, (\theta, x)))\, d\mathbb{P}'(\theta, x) \to \text{risco da função de decisão } \delta \text{ contra } \mathbb{P}'$$

$$\rho'^{*}(\mathbb{P}') = \inf\{\rho'(\delta, \mathbb{P}') : \delta \in \Delta\} \to \text{Risco de Bayes}$$

Uma regra $\delta^{*} \in \Delta$ tal que $\rho'(\delta^{*}, \mathbb{P}') = \rho'^{*}(\mathbb{P}')$ é chamada Regra (ou função) de decisão de Bayes.

O procedimento de avaliar $\rho'(\delta, \mathbb{P}')$ para cada $\delta \in \Delta$ e então tomar $\delta^* \in \Delta$ e então tomar $\delta^* \in \Delta$ que atinge o Risco de Bayes.

$\rho^*(\mathbb{P}')$ é a chamada FORMA NORMAL do problema de decisão

$(\Delta, \Omega, L', \mathbb{P}')$.

Para $\delta \in \Delta$,

supor abs. cont. domina. pela λ.

$$\rho'(\delta, \mathbb{P}') = \int_{\Theta \times \mathfrak{X}} L'(\delta, (\theta, x))\, d\mathbb{P}' = \int_{\Theta \times \mathfrak{X}} L(\delta(x), \theta)\, d\mathbb{P}'$$

$$= \int_{\Theta \times \mathfrak{X}} L(\delta(x), \theta)\, f(\theta, x)\, d\theta\, dx$$

limitada de modo a poder trocar as int.

$$= \int_{\Theta} \left[\int_{\mathfrak{X}} L(\delta(x), \theta)\, f(x/\theta) f(\theta)\, dx \right] d\theta$$

$$= \int_{\Theta} \underbrace{\left[\int_{\mathfrak{X}} L(\delta(x), \theta)\, f(x/\theta)\, dx \right]}_{\substack{\text{Risco da função de decisão } \delta \in \Delta \\ \text{quando o estado da natureza é } \theta,\ R(\delta, \theta)}} f(\theta)\, d\theta$$

$$= \int_{\mathfrak{X}} \overset{* f(x)}{} \left[\int_{\Theta} L(\delta(x), \theta)\, \underline{\frac{f(x/\theta)\, f(\theta)}{f(x)}}\, d\theta \right] dx \Rightarrow$$

$$\rho'(\delta, \mathbb{P}') = \int_{\mathfrak{X}} f(x) \underbrace{\left[\int_{\Theta} L(\delta(x), \theta)\, f(\theta/x)\, d\theta \right]}_{\rho(\delta(x), \mathbb{P}_{X=x})} dx$$

$$\geq \int_{\mathcal{X}} f(x) \left[\int_{\Theta} L(d_x^*, \theta) f(\theta/x)\, d\theta \right] dx, \text{ onde}$$

d_x^* é decisão de Bayes do problema $(\mathcal{D}, \Theta, L, \mathbb{P}_{X=x})$.

Definindo $\delta^*: \mathcal{X} \to \mathcal{D}$ por $\delta^*(x) = d_x^*$, resulta que $\forall \delta \in \Delta$,

$$\rho'(\delta, \mathbb{P}') \geq \rho'(\delta^*, \mathbb{P}').$$

Logo, δ^* definido acima é regra de decisão de Bayes contra $\mathbb{P}'$.

Ao procedimento que consiste em resolver o problema $(\mathcal{D}, \Theta, L, \mathbb{P}_{X=x})$ para cada $x \in \mathcal{X}$ e então construir a função de decisão do problema $(\Delta, \Omega, L', \mathbb{P}')$ como descrito ao lado, dá-se o nome de Forma Extensiva do problema $(\Delta, \Omega, L', \mathbb{P}')$.

Retomando o exemplo da aula passada.

$\Theta = \{\theta_1, \theta_2\}$

$\mathbb{P}(\theta = \theta_1) = p$

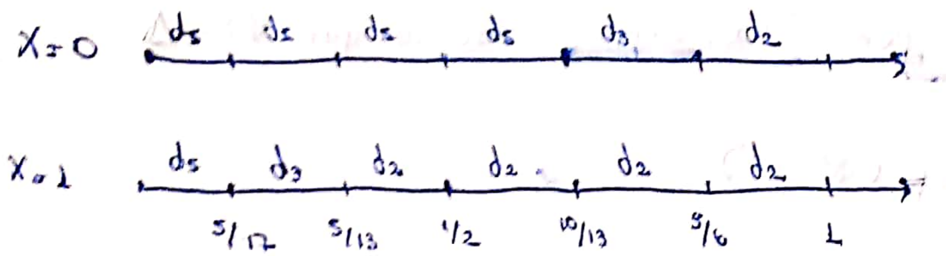
	d_1	d_2	d_3	d_4	d_5	d_6	d_7
θ_1	0	0	2	3	4	10	6
θ_2	10	7	4	5	2	2	6

$\mathcal{X} = \{0, 1\}$ $\quad X/\theta_1 \sim Ber(4/5)$

$\quad\quad\quad\quad\quad X/\theta_2 \sim Ber(1/3)$

$(\mathcal{D}, \Theta, L, P_{X=0})$ $\qquad$ $(\mathcal{D}, \Theta, L, P_{X=1})$

$$d_0^* = \begin{cases} d_2, & \text{se } p > 5/6 \\ d_3, & \\ d_5, & \text{se } p < 10/13 \end{cases} \qquad d_1^* = \begin{cases} d_2, & p > 5/13 \\ d_3, & \\ d_5, & p < 5/17. \end{cases}$$



Suponhamos $P(\Theta = \theta_2) = p = 4/10$. A sua regra de decisão de Bayes é

$$\delta^*(x) = \begin{cases} d_5, & x = 0 \\ d_2, & x = 1. \end{cases}$$

Se $P(\Theta = \theta_2) = 8/10$, então

$$\delta^*(x) = \begin{cases} d_3, & x = 0 \\ d_2, & x = 1 \end{cases}$$

Se $P(\Theta = \theta_2) = 9/10$, então

$$\delta^*(x) = d_2, \ \forall x \in \{0,1\}.$$

Analogamente, se $p = 2/10$,

$$\delta^*(x) = d_5, \ \forall x \in \{0,1\}$$

$$\rho'^*(\mathbb{P}_p) = \rho'(\delta^*, \mathbb{P}_p)$$

$$\sum_{x=0}^{1}\sum_{i=1}^{2} L(\delta^*(x),\theta_i)\, \mathbb{P}(\theta=\theta_i, X=x)$$

$$\sum_{x=0}^{1}\sum_{i=1}^{2} L(\delta^*(x;\theta_i))\, \mathbb{P}(\theta=\theta_i)\, \mathbb{P}(X=x/\theta=\theta_i)$$

$$= L(\delta^*(0),\theta_1)\, p\, \frac{1}{5} + L(\delta^*(0),\theta_2)(1-p)\frac{2}{3} + L(\delta^*(1),\theta_1)\, p\, \frac{4}{5}$$

$$+ L(\delta^*(1),\theta_2)(1-p)\frac{1}{3}$$

$$p < \frac{5}{17} \Rightarrow \frac{4p}{5} + 2(1-p)\frac{2}{3} + 4\,\frac{4p}{5} + 2(1-p)\frac{1}{3} = 4p + 2(1-p) = 2+2p$$

$$\frac{5}{17} < p < \frac{5}{13} \Rightarrow \frac{4p}{5} + 2(1-p)\frac{2}{3} + 2\,\frac{4p}{5} + 4\,\frac{(1-p)}{3} = \frac{40-4p}{15}$$

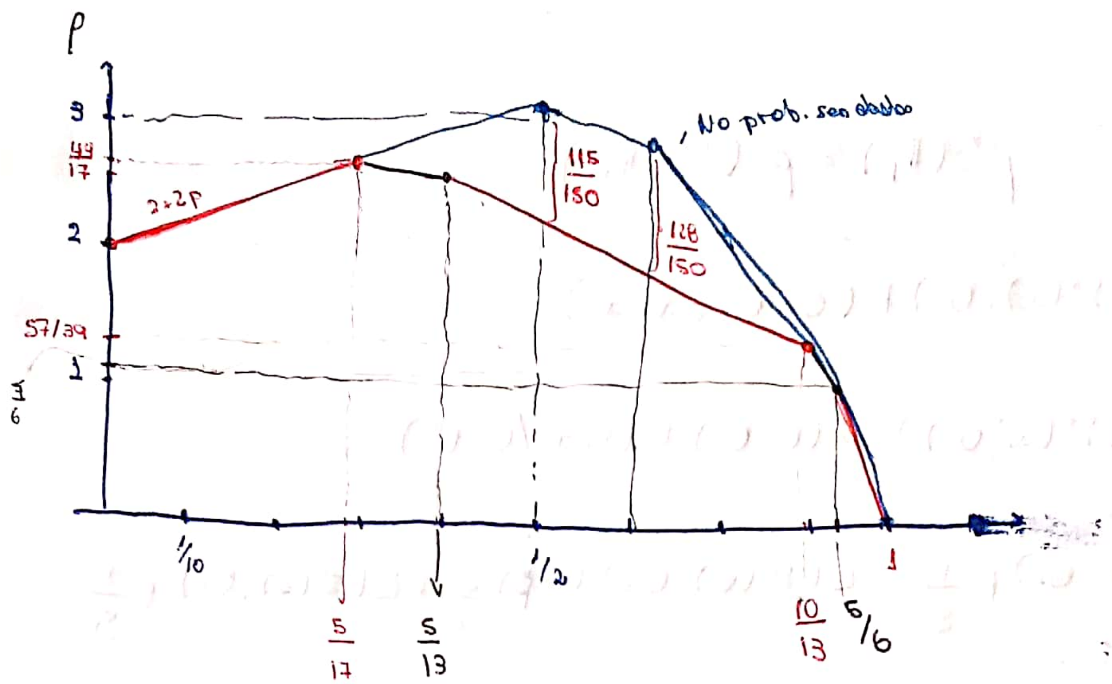
$$\frac{5}{13} < p < \frac{10}{13} \Rightarrow \frac{4p}{5} + 2\,\frac{2(1-p)}{3} + 0\,\frac{4p}{5} + 7\,\frac{(1-p)}{3} = \frac{4p}{5} + \frac{11(1-p)}{3}$$

$$= \frac{55-43p}{15}$$

$$\frac{10}{13} < p < \frac{5}{6} \rightarrow \frac{2p}{5} + 4\cdot\frac{2(1-p)}{3} + 0\cdot\frac{4p}{5} + 7\cdot\frac{(1-p)}{3} = \frac{2p}{5} + \frac{15(1-p)}{3}$$

$$= \frac{25-23p}{5}$$

$$p > \frac{5}{6} \rightarrow \underset{\theta_1}{0}\,\frac{p}{5} + \underset{\theta_2}{7}\,\frac{2}{3}(1-p) + \underset{\theta_1}{0}\,\frac{4}{5}p + \underset{\theta_2}{7}\,\frac{(1-p)}{3} = 7(1-p)$$



No problema sem dados,

$$\rho^*(\mathbb{P}_p) = \begin{cases} 2+2p & , p \leq 1/2 \\ 4-2p & , 1/2 \leq p \leq 3/5 \\ 7(1-p) & , p > 3/5 \end{cases}$$

Problema de Determinação do Tamanho Amostral e Da Regra da Decisão

$\overline{\Omega} = \Theta \times \mathfrak{X}$, $\mathfrak{X} = \prod_{i=1}^{\infty} \mathfrak{X}_i$, tudo que desconheço: sobre o qual há incerteza.

(posso vir observar)

$\Delta_n = \{\delta : \prod_{i=1}^{n} \mathfrak{X}_i \to D\}$

$\Delta_0 = D$

$\overline{\Delta} = \bigcup_{n=0}^{\infty} \{n\} \times \Delta_n$ → espaço de ações.

$\overline{\mathbb{P}}$ sobre $\sigma(\Theta \times \mathfrak{X})$, $\overline{L} : \overline{\Delta} \times \overline{\Omega} \to \mathbb{R}$

$((n, \delta^{(n)}), (\theta, x)) \to \overline{L}((n, \delta^{(n)}), (\theta, x)) = L(\delta^{(n)}(x), \theta) + C(n, \delta, \theta, x)$

Muito simplificadamente, vamos considerar

$$c(n, \delta, \theta, x) = c_0 \cdot n, \quad c_0 > 0 \text{ fixado!}$$

Para $(n, \delta) \in \bar{\Delta}$, temos

$$\bar{\rho}((n,\delta), \bar{P}) = \int_{\Theta \times \mathfrak{X}} \bar{L}((n,\delta),(\theta,x))\, d\bar{P}$$

$$= \int_{\Theta \times \mathfrak{X}} [L(\delta(x), \theta) + n c_0]\, d\bar{P} = n c_0 + \int_{\Theta \times \mathfrak{X}} L(\delta^{(n)}(x), \theta)\, d\bar{P}$$

Teoria da Decisão

14/10/15

Aula 16

Problema com custos:

$\mathcal{X} = \prod_{i=1}^{\infty} \mathcal{X}_i$

$\bar{\Omega} = \Theta \times \mathcal{X}$

$\Delta_n = \{\delta : \mathcal{X}_1 \times \dots \times \mathcal{X}_n \to D\}$, $\Delta_0 = D$

$\bar{\Delta} = \bigcup_{n=0}^{\infty} \{n\} \times \Delta_n$

$$\bar{L}((n,\delta),(\theta,x)) = L(\delta(x),\theta) + \underbrace{C(n,\delta,\theta,x)}_{nC_0 \, , \, C_0 > 0}$$

Avaliando $\bar{\rho}((n,\delta),\bar{\mathbb{P}})$.

$$\bar{\rho}((n,\delta),\bar{\mathbb{P}}) = \int_{\Theta \times \mathcal{X}} L(\delta(x),\theta)\, d\bar{\mathbb{P}} + n C_0$$

Exemplo:

$D = \{1,2,3\} = \Theta$

L

$d \backslash \theta$	1	2	3
1	0	1	2
2	1	0	1
3	2	1	0

$L(d,\theta) = |d-\theta|$

$\mathbb{P}(\theta = i) = p_i$, $i = 1,2,3$

Considerando o prob. sem dados $(\mathcal{D}, \Theta, L, P)$.

Decisão ótima d_1:

$$E[L(d_1,\theta)] \leq E[L(d_2,\theta)] \Leftrightarrow 1p_2 + 2(1-p_1-p_2) \leq p_1 + 1 - p_1 - p_2 ,$$

$$E[L(d_1,\theta)] \leq E[L(d_3,\theta)] \Leftrightarrow 1p_2 + 2(1-p_1-p_2) \leq 2p_1 + p_2$$

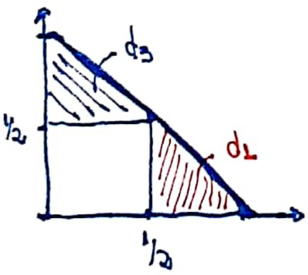
$$\left.\begin{aligned} &\Leftrightarrow 2p_1 \geq 1 \Rightarrow p_1 \geq 1/2. \\ &\Leftrightarrow 2 - 2p_2 \leq 4p_1 \Rightarrow p_1 \geq \frac{1-p_2}{2} \end{aligned}\right\} p_1 \geq \frac{1}{2}$$

Do mesmo modo, d_3 é decisão ótima se

$$E[L(d_3,\theta)] \leq E[L(d_1,\theta)] \Leftrightarrow p_3 \geq p_1$$

$$E[L(d_3,\theta)] \leq E[L(d_2,\theta)] \Leftrightarrow 2p_1 + p_2 \leq p_1 + p_3$$

$$\Leftrightarrow p_1 + p_2 \leq p_3 \Leftrightarrow p_3 \geq 1/2.$$



Suponhamos que estão disponíveis $X_1, X_2, X_3, \ldots$, que dado Θ, são c.i.i.d. $Exp(\Theta)$

Para $n \in \mathbb{N}$ e $(x_1, \ldots, x_n) \in \mathbb{R}^n_+$

$$P(\Theta = 1 \mid X_1 = x_1, \ldots, X_n = x_n) = \frac{f(x_1, \ldots, x_n / \Theta = 1)\, P(\Theta = 1)}{\sum_{j=1}^{3} f(x_1, \ldots, x_n / \Theta = j)\, P(\Theta = j)}$$

$$= \frac{\prod_{i=1}^{n} f(x_i / \Theta = 1)\, p_1}{\sum_{j=1}^{3} \prod_{i=1}^{n} \underbrace{f(x_i / \Theta = j)}_{j e^{-j x_i}}\, p_j}$$

$$\Rightarrow P(\theta = 1 \mid X_1 = x_1, \dots, X_n = x_n) = \frac{1^n e^{-\sum_{i=1}^{n} x_i} p_1}{\sum_{j=1}^{3} j^n e^{-j\sum_{i=1}^{n} x_i} p_j}$$

Suponhamos $p_1 = p_2 = p_3 = 1/3$:

- Após observar $X_1 = x_1, \dots, X_n = x_n$, a decisão será por d_1 se

$$P(\theta = 1 \mid X_1 = x_1, \dots, X_n = x_n) > 1/2 \Leftrightarrow e^{-\sum_{i=1}^{n} x_i} > 2^n e^{-2\sum x_i} + 3^n e^{-3\sum x_i}$$

$$\overset{(\times e^{3\sum x_i})}{\Leftrightarrow} e^{2\sum_{i=1}^{n} x_i} - 2^n e^{\sum x_i} - 3^n > 0$$

(*) $t = e^{\sum x_i}$

$t^2 - 2^n t - 3^n > 0 \Leftrightarrow$

$t < x_1$ ou $t > x_2$, com

$x_i = \frac{2^n \pm \sqrt{4^n + 4.3^n}}{2}$

x_1 é neg. Sobra $t > x_2$.

$$\overset{(*)}{\Leftrightarrow} e^{\sum x_i} > \frac{2^n + \sqrt{4^n + 4.3^n}}{2}$$

$$\Leftrightarrow \sum_{i=1}^{n} x_i > \underbrace{\log \frac{2^n + \sqrt{4^n + 4.3^n}}{2}}_{b}$$

- Decisão por d_3 se

$$P(\theta = 3 \mid X_1 = x_1, \dots, X_n = x_n) > 1/2 \Leftrightarrow 3^n e^{-3\sum x_i} > e^{-\sum x_i} + 2^n e^{-2\sum x_i}$$

$$\Leftrightarrow e^{2\sum x_i} + 2^n e^{\sum x_i} - 3^n < 0$$

$$\Leftrightarrow 0 < e^{\sum x_i} < \frac{-2^n + \sqrt{4^n + 4.3^n}}{2}$$

$$\Rightarrow \sum_{i=1}^{n} x_i < \underbrace{\log \frac{-2^n + \sqrt{4^n + 4.3^n}}{2}}_{a}$$

Assim,

$$\delta^{*(n)}(x_1, \dots, x_n) = \begin{cases} 1, & \sum_{i=1}^{n} x_i > b \\ 2, & a < \sum_{i=1}^{n} x_i < b \\ 3, & \sum_{i=1}^{n} x_i \leq a. \end{cases}$$

Por fim, devemos avaliar para cada $(n, \delta^{*(n)})$,

$$\bar{\rho}((n, \delta^{*(n)}), \bar{\mathbb{P}}) = \int_{\Theta \times \mathcal{X}} L(\delta^{*(n)}(x), \theta) \, d\bar{\mathbb{P}} + n C_0$$

$$\int_{\Theta \times \mathcal{X}} L(\delta^{*(n)}(x), \theta) \, d\bar{\mathbb{P}} = \sum_{i=1}^{3} \int_{\mathbb{R}_+^n} L(\delta^{*(n)}(x_1, \dots, x_n), \theta) f(x_1, \dots, x_n / \theta = i) \underbrace{\mathbb{P}(\theta = i)}_{1/3}$$

$$= \frac{1}{3} \left\{ \int_{\{\sum x_i < a\}} L(3,1) f(x / \theta = 1) dx + \int_{\{a < \sum x_i < b\}} \overbrace{L(2,1)}^{1} f(x / \theta = 1) dx + \int_{\{\sum x_i > b\}} \overbrace{L(1,1)}^{0} f(x / \theta = 1) dx \right.$$

$$+ \int_{\{\sum x_i < a\}} \overbrace{L(3,2)}^{1} f(x / \theta = 2) dx + \int_{\{a < \sum x_i < b\}} \overbrace{L(2,2)}^{0} f(x / \theta = 2) dx + \int_{\{\sum x_i > b\}} \overbrace{L(1,2)}^{1} f(x / \theta = 2) dx$$

$$\left. + \int_{\{\sum x_i < a\}} \overbrace{L(3,3)}^{0} f(x / \theta = 3) dx + \int_{\{a < \sum x_i < b\}} \overbrace{L(2,3)}^{1} f(x / \theta = 3) dx + \int_{\{\sum x_i > b\}} L(1,3) f(x / \theta = 3) dx \right\}$$

$\Longrightarrow$

$$\int_{\Theta \times \mathfrak{X}} L(\delta^{*(n)}(x), \theta)\, d\bar{P} =$$

$$\frac{1}{3}\left\{ P\left(a < \sum_{i=1}^{n} X_i < b / \theta = 1\right) + P\left(\sum_{i=1}^{n} X_i < a / \theta = 2\right) + P\left(\sum_{i=1}^{n} X_i > b / \theta = 2\right) \right.$$
$$\left. + P\left(a < \sum_{i=1}^{n} X_i < b / \theta = 3\right) + 2 \cdot P\left(\sum_{i=1}^{n} X_i < a / \theta = 1\right) + 2P\left(\sum_{i=1}^{n} X_i > b / \theta = 3\right) \right\}$$

Assim,

$$\bar{\rho}((n, \delta^{*(n)}), \bar{P}) = \int_{\Theta \times \mathfrak{X}} L(\delta^{*(n)}, \theta)\, d\bar{P} + \underbrace{0{,}01}_{C_0 = 0{,}01} n$$

Por fim, avaliando para cada $(n, \delta^{*(n)})$, $\bar{\rho}((n, \delta^{*(n)}), \bar{P})$, obtemos, como decisão ótima em $\bar{\Delta}$,

$$n = 15 \quad e \quad \delta^{(15)}(x_1, \dots, x_{15}) = \begin{cases} 1, & \sum_{i=1}^{15} x_i > 10{,}4103 \\ 3, & \sum_{i=1}^{15} x_i < 6{,}0689 \\ 2, & 6{,}0689 < \sum_{i=1}^{15} x_i < 10{,}4103 \end{cases}$$

No caso do problema sem dados e como $p_1 = p_2 = p_3 = 1/3$, a dec. ótima é d_2.

$$\rho(d_2, P) = \overset{1}{L(d_2, 1)} P(\theta = 1) + \overset{0}{L(d_2, 2)} P(\theta = 2) + \overset{1}{L(d_2, 3)} P(\theta = 3)$$
$$= \frac{2}{3}$$

Escolha de experimentos

$\xi \rightarrow$ conjunto de experimentos

Para cada $E \in \xi$, $\mathfrak{X}_E$ é o espaço amostral do experimento E.

E_0: não conduzir amostragem alguma

$$\Delta_E = \{\delta : \mathcal{X}_E \to D\}$$

$$\Delta_{E_0} = D$$

$$\tilde{\Delta} = \bigcup_{E \in \mathcal{E}} \{E\} \times \Delta_E$$

Custo : $C_E = C_E(-,-,-,-)$

Exemplo: (Exercício)

$\Theta = \{\theta_1, \theta_2\}$

$D = \{d_1, d_2\}$

θ \ d	d_1	d_2
θ_1	0	a_1
θ_2	a_2	0

Experimento 1:

Obs. X t.q. $X/\theta_1 \sim Ber(2/3)$

$X/\theta_2 \sim Ber(1/3)$

Custo: C_X

Experimento 2:

Obs. Y t.q. $Y/\theta_1 \sim Ber(3/4)$

$Y/\theta_2 \sim Ber(1/2)$

Custo: C_Y

-.-.-

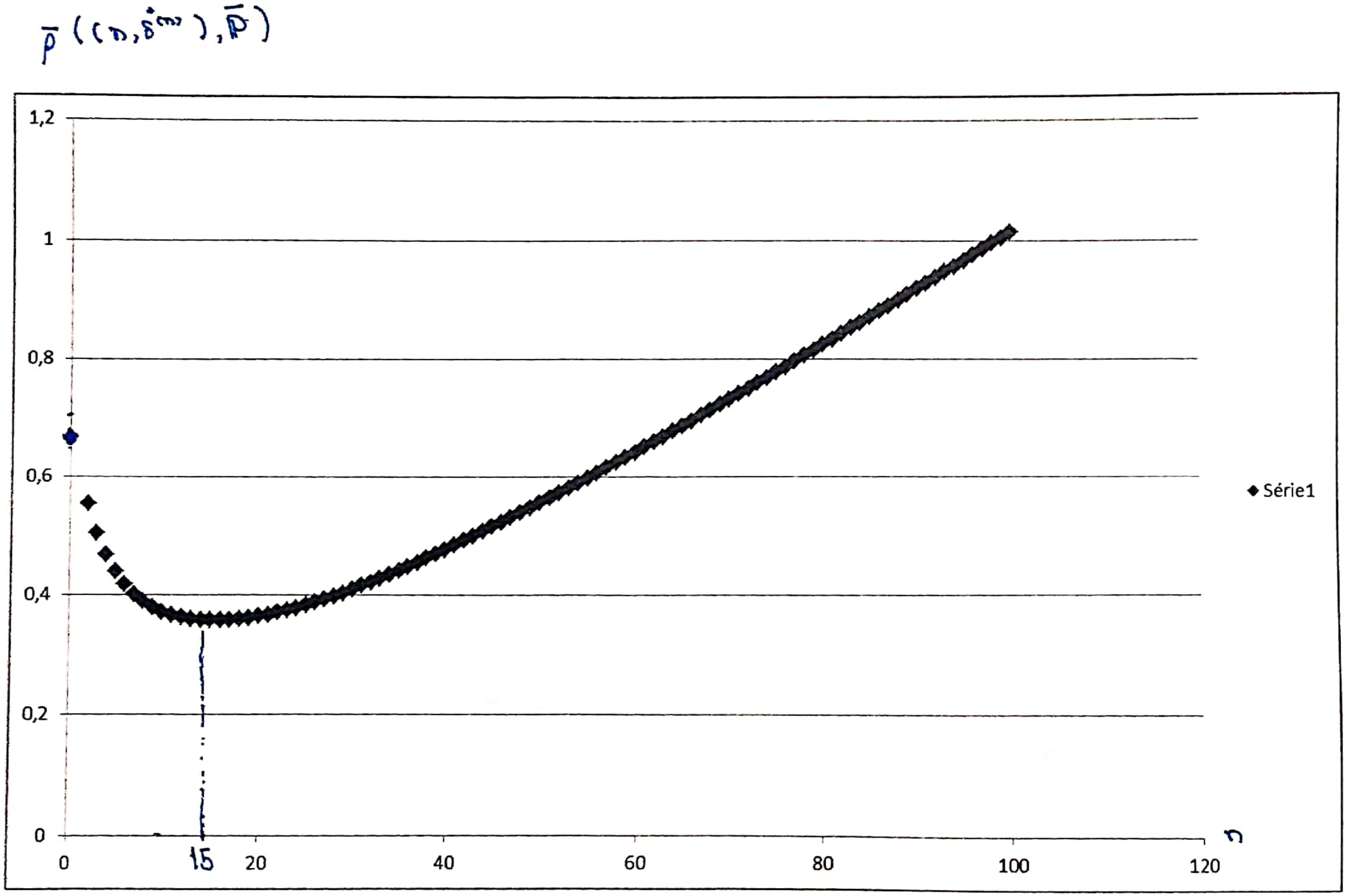
X, Y

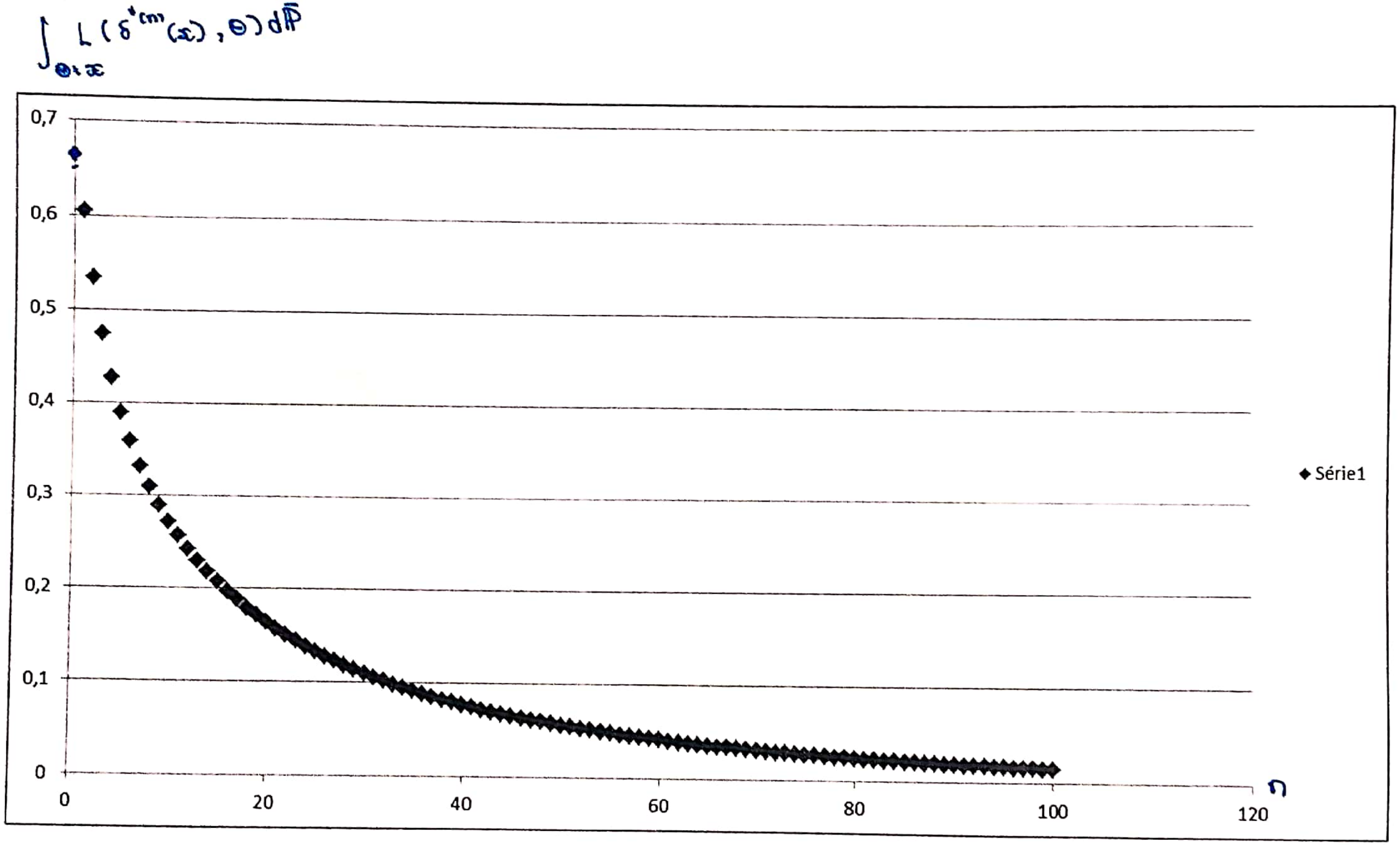
(1) Você é informado simultaneamente de X e Y.

$$f^{(1)}(\theta | x,y) \propto f(x,y|\theta) f(\theta)$$

$\theta = ?$

$f(\cdot) \xrightarrow[X \text{ e } Y]{} f^{(1)}(\cdot|x,y)$


$\bar{p}((n, \dot{s}^{m}), \bar{P})$
1,2
1
0,8
0,6
0,4
0,2
0
0
15
20
40
60
80
100
120
n
◆ Série1


$\int_{\Theta \times \mathcal{X}} L(\delta^{*(n)}(x), \theta) \, d\bar{P}$
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
20
40
60
80
100
120
n
Série1

(2) Você é informado primeiro de X e depois de Y.

$\theta, ?$

$$f(\cdot) \xrightarrow{X} f^{(2)}(\theta / x) \xrightarrow{Y} f^{(2)}(\theta / x, y)$$

Obs: ñ haver regra de coerência no tempo, supercondicionamento ou regra de Jeffreys.

$$f^{(2)}(\theta | x, y) \propto f(y / \theta, x) f(\theta / x)$$

$$\propto \frac{f(x, y / \theta)}{f(x / \theta)} f(x / \theta) f(\theta)$$

$$\propto f(x, y / \theta) f(\theta) \propto f^{(1)}(\theta | x, y)$$

⋮

$$(\Delta, \Theta \times \mathcal{X}, L', P')$$

estimadores do parâmetro

$$L'(\delta, (\theta, x)) = L(\delta(x), \theta) = |\delta(x) - \theta|$$

$$= (\delta(x) - \theta)^2$$

Aula ~~18~~ 17

19/10

19/10/2015

ESTIMAÇÃO

(D, Θ, L, P)

Após observar $x \in \mathcal{X}$,

$(D, \Theta, L, P_{X=x})$ com $D = \Theta$.

Problemas pré-experimentação

$\Delta = \{\delta : \mathcal{X} \to D\}$

No ~~pb~~ problema de estimação $D = \Theta$, de modo que

$\Delta = \{\delta : \mathcal{X} \to \Theta\}$ conjunto de estimadores

$* \; (\Delta, \Theta_x \mathcal{X}, L', P')$

$L' : \Delta \times (\Theta_x \mathcal{X}) \to R_+$

$(\delta, (\theta, x)) \mapsto L'(\delta, (\theta, x)) = L(\delta(x), \theta)$

usualmente, L expressa a discrepância (distância) entre a estimativa $\delta(x)$ e o parâmetro (desconhecido) θ.

$L(d, \theta) = h(\theta)\, g(dist(d, \theta))$, g não-decrescente

Alguns exemplos: $(\Theta \subseteq R)$

(1) $L(d, \theta) = |d - \theta|$, perda absoluta.

(2) $L(d, \theta) = (d - \theta)^2$, ~~Peet~~ Perda Quadrática

(3) $L(d,\theta) = \dfrac{(d-\theta)^2}{|\theta|}$ $(\mathbb{P}(\theta=0)=0)$

Dissertação de Mestrado
Rosângela Helena Loschi
92 a 95 ...
Combining Probability Distribution
Genest e Zidek (1986)
Statistical Science

* Perda Quadrática : $\Theta \subseteq \mathbb{R}$

$L(d,\theta) = (d-\theta)^2$, $E(\theta^2) < \infty$

$\mathbb{E}[(d-\theta)^2] = \mathbb{E}[L(d,\theta)] \Rightarrow$

$\mathbb{E}[(d-\mathbb{E}(\theta)+\mathbb{E}(\theta)-\theta)^2] = \mathbb{E}[(d-\mathbb{E}(\theta))^2] + \cancel{2\mathbb{E}[(d-\mathbb{E}(\theta))(\mathbb{E}(\theta)-\theta)]^2}$

$= 2\mathbb{E}\left[(d-\mathbb{E}(\theta))(\mathbb{E}(\theta)-\theta)\right] + \mathbb{E}\left[(E(\theta)-\theta)^2\right] \Rightarrow$

$\mathbb{E}[L(d,\theta)] = (d-\mathbb{E}(\theta))^2 + \mathrm{Var}(\theta)$

Assim, $E(\theta) \in \Theta \Rightarrow d^* = \mathbb{E}(\theta)$

$\mathbb{E}(\theta) \notin \Theta \Rightarrow d^* \in D$ é tal que $|d^* - \mathbb{E}(\theta)| =$

$= \min\{|d-\mathbb{E}(\theta)| : d \in D\}$

$\mathbb{E}(\theta) \in \Theta \Rightarrow$

$\rho^*(\mathbb{P}) = \rho(\mathbb{E}(\theta),\mathbb{P}) = \mathrm{Var}(\theta)$

$\mathbb{E}(\theta) \notin \Theta \Rightarrow \rho^*(\mathbb{P}) = \rho(d^*,\mathbb{P}) = (d^*-\mathbb{E}(\theta))^2 + \mathrm{Var}(\theta)$

Após a observação de $X=x$.

Após a observação de $X=x$.

$(\Theta, \Theta, L, \mathbb{P}_{X=x})$

$d^*_x = \mathbb{E}(\theta/X=x)$

$\rho^*(\mathbb{P}_{X=x}) = \cancel{\rho(\mathbb{E}(\theta/X=x))} = \rho(\mathbb{E}(\theta/X=x),\mathbb{P}_{X=x}) = \mathrm{Var}(\theta/X$

* no problema pré-experimentação

$\delta^* : \mathcal{X} \to \Theta$

$x \in \mathcal{X} \mapsto \delta^*(x) = \mathbb{E}(\theta/X=x)$, isto é,

$\delta^*(x) = \mathbb{E}(\theta/X) \Rightarrow$ estimador de Bayes com relação à perda quadrática.

Em geral, para $\delta \in \Delta$,

$\rho(\delta,\mathbb{P}') = \displaystyle\int_{\Theta\times\mathcal{X}} L(\delta(x),\theta)\, d\mathbb{P}' = \int_{\mathcal{X}} f(x)\left[\int_{\Theta} L(\delta(x),\theta) f(\theta)\, d\theta\right] d$

Assim, ~~continuando~~ considerando a perda quadrática,

$\rho'(\delta^*) = \displaystyle\int_{\mathcal{X}} f(x)\left[\int_{\Theta} L(E(\theta/x),\theta)\, f(\theta/x)\, d\theta\right] dx$

$= \displaystyle\int_{\mathcal{X}} f(x)\Bigg[\underbrace{\int_{\Theta} (\theta - E(\theta/X=x))^2 f(\theta/x)\, d\theta}_{\to \mathrm{Var}(\theta/X=x)}\Bigg] dx \Rightarrow$

$$\rho^{*'}(\delta^*) = \mathbb{E}\left(Var(\theta/X)\right)$$

Exemplo: $X_1, \dots, X_n$ dado θ, c.i.i.d $N(\theta, \upsilon_0^2)$

Suponhamos, a priori $\theta \sim N(\theta_0, \sigma_0^2)$

Vemos que

$$\theta / X_1 = x_1, \dots, X_n = x_n \sim N\left(\frac{\upsilon_0^2 \theta_0 + n\sigma_0^2 \cdot \bar{x}}{\upsilon_0^2 + n\sigma_0^2}, \frac{\sigma_0^2 \cdot \upsilon_0^2}{\upsilon_0^2 + n\sigma_0^2}\right)$$

Após observar $x = (x_1, \dots, x_n) \in \mathbb{R}^n$,

$$\mathbb{E}(\theta / X = x) = \frac{\upsilon_0^2 \cdot \theta_0 + \sigma_0^2 \cdot n\bar{x}}{\upsilon_0^2 + n\sigma_0^2}$$

$$\rho^*\left(\mathbb{E}(\theta/x), \mathbb{P}_{X=x}\right) = Var(\theta / X = x) = \frac{\sigma_0^2 \cdot \upsilon_0^2}{\upsilon_0^2 + n\sigma_0^2}, \ \forall x \in \mathcal{X}$$

No problema pré-experimentação

$$\delta^*(x) = \mathbb{E}(\theta/X) = \frac{\upsilon_0^2 \theta_0 + n\sigma_0^2 \bar{X}}{\upsilon_0^2 + n\sigma_0^2}$$

$$\rho^{*'}(\mathbb{P}') = \rho'(\delta^*(x), \mathbb{P}') = \mathbb{E}\left(Var\left(\theta/X\right)\right) = \frac{\sigma_0^2 \upsilon_0^2}{\upsilon_0^2 + n\sigma_0^2}$$

Ex2. $X_1, \dots, X_n$ dado θ, c.i.i.d $Ber(\theta)$

$\theta \sim Beta(a, b)$, $a, b > 0$

Vemos que

$$\theta / X_1 = x_1, \dots, X_n = x_n \sim Beta\left(a + \sum_{i=1}^n x_i, b + n - \sum_{i=1}^n x_i\right)$$

Para $x = (x_1, \dots, x_n) \in \{0, 1\}^n$,

$$\mathbb{E}(\theta / X) = \frac{a + \bar{x}n}{a + b + n}$$

$$\rho^*(\mathbb{P}_{X=x}) = \rho\left(\mathbb{E}(\theta / X = x), \mathbb{P}_{X=x}\right) = Var(\theta/x) =$$

$$\frac{(a + n\bar{x})(b + n - n\bar{x})}{(a+b+n)^2 (a+b+n+1)}$$

* No problema pré-experimentação, o estimador de Bayes com relação à perda quadrática é

$$\mathbb{E}(\theta/X) = \frac{a + n\bar{X}}{a+b+n} = \frac{a+b}{a+b+n} \cdot \underbrace{\frac{a}{a+b}}_{\mathbb{E}(\theta)} + \frac{n}{a+b+n} \cdot \bar{X}$$

$$\rho^{*'}(\mathbb{P}') = \rho'\left(\mathbb{E}(\theta/X), \mathbb{P}'\right) = \mathbb{E}\left(Var(\theta/X)\right) =$$

$$Var(\theta) - Var\left(E(\theta/X)\right)$$

$$= \frac{ab}{(a+b)^2(a+b+n)} - Var\left(\frac{a+n\bar{X}}{a+b+n}\right) =$$

$$= Var(\theta) - \left(\frac{n}{a+b+n}\right)^2 . Var(\bar{X}) = \cancel{Var(\theta)} \left(\frac{n}{a+b+n}\right)^2 \cancel{Var(\bar{X})}$$

$$= Var(\theta) - \left(\frac{n}{a+b+n}\right)^2 \left[\mathbb{E}(Var(\bar{X}/\theta)) + Var(\underbrace{\mathbb{E}(\bar{X}/\theta)}_{Bernoulli(\theta)}) \right] =$$

$$= Var(\theta) - \left(\frac{n}{a+b+n}\right)^2 \left[\mathbb{E}\left(\frac{\theta(1-\theta)}{n}\right) + Var(\theta) \right]$$

$$\Rightarrow \underbrace{\rho^{*'}(\mathbb{P}')}_{\to 0 \text{ if } n\to\infty} = \frac{a.b}{(a+b)^2(a+b+1)} - \left(\frac{n}{a+b+n}\right)^2 \left[\frac{ab}{n(a+b)(a+b+1)} + \frac{ab}{(a+b)^2(a+b+1)} \right]$$

$$Var(\theta) = \underbrace{\mathbb{E}(Var(\theta/X))}_{\rho^{*}(\mathbb{P}')} + \underbrace{Var(\mathbb{E}(\theta/X))}_{Var(\delta^*)}$$

EX3:

θ brancas
$10-\theta$ verdes

$\theta = ?$

$\Theta = \{0, 1, \ldots, 10\}$

$\theta \sim Bin(10, 1/3)$

$\mathbb{E}(\theta) = 10 \cdot \frac{1}{3} \approx 3,33$

$\delta^* = 3$ distintos θ, sendo θ

$$\rho^*(\mathbb{P}) = \rho(3,\mathbb{P}) = (3-E(\theta))^2 + Var(\theta) = \left(3-\frac{10}{3}\right)^2 + \overset{\theta(1-\theta)}{10 \cdot \frac{1}{3} \cdot \frac{2}{3}} \Rightarrow \rho^*(\mathbb{P}) = \frac{21}{9}$$

$\theta \sim Beta(a,b)$, $u, v \in \mathbb{N}$

$$\mathbb{E}\left(\theta^u (1-\theta)^v\right) = \int_0^1 \theta^u (1-\theta)^v \frac{\Gamma(a,b)}{\Gamma(a)\Gamma(b)} \theta^{a-1}(1-\theta)^{b-1} d\theta =$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \frac{\Gamma(a+u+b+v)}{\Gamma(a+u)\Gamma(b+v)} \theta^{a+u-1} (1-\theta)^{b+v-1} d\theta \cdot \frac{\Gamma(a+u)\Gamma(b+v)}{\Gamma(a+u+b+v)} =$$

$$\mathbb{E}(\theta^u(1-\theta)^v) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+u)\Gamma(b+v)}{\Gamma(a+u+b+v)}$$

Assim,

$$\mathbb{E}(\theta(1-\theta)) = \frac{ab}{(a+b)(a+b+1)}$$, Resultado usado em $\rho^{*'}(\mathbb{P}')$

---

* Exemplo 3

* Suponhamos que $X|\theta \sim$ Hipergeométrica $(\overset{\text{brancas}}{\theta}, \overset{\text{verdes}}{10-\theta}, 2)$

Para $X = 0,1,2$, $x \in \{0,1,2\}$ $\qquad 10-\theta \geq 2-x$

$$\mathbb{P}(\theta=\theta/X=x) \propto \mathbb{P}(X=x/\theta=\theta)\mathbb{P}(\theta) \propto$$

$$\propto \frac{\binom{\theta}{x}\binom{10-\theta}{2-x}}{\binom{10}{2}} \cdot \binom{10}{\theta}\left(\frac{1}{3}\right)^\theta \left(\frac{2}{3}\right)^{10-\theta} \cdot \mathbb{I}(\theta)_{(0,1,\ldots,10)}$$

$$\propto \frac{\theta!}{(\theta-x)!}\frac{(10-\theta)!}{[8-(\theta-x)]!}\frac{(1/3)^{\theta}(2/3)^{10-\theta}}{\theta!(10-\theta)!}\ \mathbb{1}_{\{x,x+1,\ldots\}}(\theta)\ \mathbb{1}_{\{0,1,\ldots,8+x\}}(\theta)\ \mathbb{1}_{\{0,1,\ldots,10\}}(\theta)$$

$$= \binom{8}{\theta-x}\left(\frac{1}{3}\right)^{\theta-x}\left(2/3\right)^{8-(\theta-x)}\ \mathbb{1}_{\{x,\ldots,8+x\}}(\theta)$$

$= \theta-x\,/X=x \sim Bin(8, 1/3)$, logo

$$\mathbb{E}((\theta-x)/x) = 8\cdot\frac{1}{3} \Rightarrow \mathbb{E}(\theta/X=x) = x+\frac{8}{3} \quad {}^{2{,}667\,\approx\,3}$$

Assim, tomamos como estimativa de Bayes

$\boxed{x+3}$

$$\rho(x+3, \mathbb{P}_{X=x}) = (x+3-\mathbb{E}(\theta/x))^2 + Var(\theta/x) =$$

$$=(x+3-(x+\tfrac{8}{3})^2)+8\cdot\tfrac{1}{3}\cdot\tfrac{2}{3} \Rightarrow$$

$\rho(x+3, \mathbb{P}_{X=x}) = \frac{1}{9}+\frac{16}{9} = \frac{17}{9}$, $\forall x \in \mathcal{X}$, no

problema pré-experimentação,

$\delta^*(x) = X+3$ e

$$\rho^{*'}(\mathbb{P}') = \sum_{x=0}^{2}\mathbb{P}(X=x).\underbrace{\rho(x+3, \mathbb{P}_{X=x})}_{\sum_{\theta}L(x+3,\theta)\mathbb{P}(\theta/x)} =$$

$$= \sum_{x=0}^{2}\mathbb{P}(X=x).\frac{17}{9} \Rightarrow \rho^{*'}(\mathbb{P}') = \frac{19}{9}.$$

$\rho'(\delta, \mathbb{P}') = \int_{\Theta\times\mathcal{X}} L(\delta(x),\theta)\,d\mathbb{P}'$: risco do estimador δ

$\rho^*(\mathbb{P}') : \inf\{\rho'(\delta,\mathbb{P}) : \delta\in\Delta\}$: risco de Bayes

$L(d,\theta) = (d-\theta)^2\ (\Theta \subseteq \mathbb{R})$

$\delta^*(x) = \mathbb{E}(\theta/x)$

(2) PERDA ~~AB~~ ABSOLUTA $(\Theta \subset \mathbb{R})$

$$L(d,\theta) = |d-\theta| = \begin{cases}(d-\theta), & \theta< d\\ (\theta-d), & \theta\geq d\end{cases}$$

Dizemos que $m\in\mathbb{R}$ é uma mediana de $\mathbb{P}$ se

$\mathbb{P}((-\infty, m]) = 1/2$ e $\mathbb{P}(([m,\infty))) \geq 1/2$

(~~…~~ m é mediana de v.a X se $\mathbb{P}(X\leq m)\geq 1/2$ e $\mathbb{P}(X\geq m)\geq 1/2$.

Teoria da Decisão

Aula 18 21/10/15

$$\rho'(\delta, P') = \int_{\Theta \times \mathfrak{X}} L(\delta(x), \theta)\, dP' \quad \text{risco do estimador } \delta$$

$$\rho'^*(\delta, P') = \inf\{\rho'(\delta, P') : \delta \in \Delta\} : \text{risco de Bayes}$$

$L(d,\theta) = (d-\theta)^2$, $(\Theta = \mathbb{R})$

$\delta^*(x) = E(\theta / x)$

2. Perda Absoluta $(\Theta \subseteq \mathbb{R})$

$$L(d,\theta) = |d-\theta| = \begin{cases} (d-\theta), & \theta < d \\ (\theta-d), & \theta \geq d \end{cases}$$

Dizemos que $m \in \mathbb{R}$ é uma mediana de P se

$$P((-\infty, m]) \geq 1/2 \text{ e } P([m, \infty)) \geq 1/2$$

(m é mediana da v.a. X se $P(X \leq m) \geq 1/2$ e $P(X \geq m) \geq 1/2$)

- A priori. $(E(|\theta|) < \infty)$

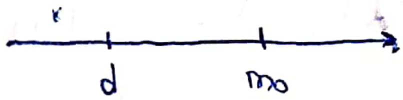
  Seja m_0 uma mediana da distribuição a priori de θ. Vamos avaliar

$$E[L(d,\theta)] - E[L(m_0, \theta)], \ \forall d \in \mathbb{R}$$

Primeiro, olhando para $d < m_0$,

$$E[L(d,\theta)] - E[L(m_0,\theta)] = E[|d-\theta|] - E[|m_0-\theta|]$$

$$= E[|d-\theta| - |m_0-\theta|]$$



$$= \int_{\{\theta<d\}} (d-m_0)\, dP + \int_{\{d\le\theta\le m_0\}} \overbrace{[(\theta-d)-(m_0-\theta)]}^{\geq -(m_0-d)}\, dP$$

$$+ \int_{\{\theta\geq m_0\}} (m_0-d)\, dP$$

$$\geq -(m_0-d)\,P(\theta<d) - (m_0-d)\,P(d\le\theta< m_0) + (m_0-d)\,P(\theta\geq m_0)$$

$$= (m_0-d)[\,P(\theta\geq m_0) - P(\theta\leq m_0)\,]$$

$$= \underbrace{(m_0-d)}_{m_0>d}\,[\,\underbrace{2P(\theta\geq m_0)}_{\geq 1/2} - 1\,] \geq 0 \quad \text{e, portanto,}$$

$$E[L(m_0,\theta)] \leq E[L(d,\theta)], \ \forall\, d<m_0$$

Do mesmo modo, verificamos (Exercício) que

$$E[L(m_0,\theta)] \leq E[L(d,\theta)], \ \forall\, d>m_0.$$

Logo, m_0 é uma decisão de Bayes no problema a priori.

Após a observação $X=x$, $Md(\theta|x)$, a mediana da dist. a posteriori é decisão de Bayes do problema $(\Theta, \Theta, L, \mathbb{P}_{X=x})$.

No problema pré-experimentação,

$\delta^*(X) = Md(\Theta / X)$ é estimador de Bayes com relação à perda absoluta.

Risco de Bayes

$$\rho^{1*}(\mathbb{P}') = \int_x f(x) \left[\int_\Theta |\Theta - Md(\Theta/x)| \, f(\Theta/x) \, d\Theta \right] dx$$

Suponha $L(d,\Theta) = \begin{cases} p_1 (d-\Theta), & \Theta < d \\ p_2 (\Theta - d), & \Theta \geq d \end{cases}$

$\delta^*(X) = $ perc. de ordem $\frac{p_2}{p_1+p_2}$ no caso subestima

$\delta^{*'}(X) = $ perc da ord $\frac{p_2}{p_1+p_2}$ no caso superest.

Exemplo 1: $X_1, \dots, X_n$, dado Θ, c.i.i.d. $N(\Theta, \sigma_0^2)$ e $\Theta \sim N(\Theta_0, \sigma_0^2)$

$$\Theta / X_1 = x_1, \dots, X_n = x_n \sim N\left(\frac{\sigma_0^2 \Theta_0 + \sigma_0^2 n\bar{x}}{\sigma_0^2 + n\sigma_0^2}, \underbrace{\frac{\sigma_0^2 \sigma_0^2}{\sigma_0^2 + n\sigma_0^2}}_{\gamma_x^2} \right)$$

$$Md(\Theta / x) = \frac{\sigma_0^2 \Theta_0 + \sigma_0^2 n\bar{x}}{\sigma_0^2 + n\sigma_0^2}$$

$$\rho(Md(\Theta/x), \mathbb{P}_{X=x}) = \int_{\mathbb{R}} |\Theta - Md(\Theta/x)| \, d\mathbb{P}_{X=x}$$

$$= \gamma_x \int_{\mathbb{R}} \left| \frac{\Theta - Md(\Theta/x)}{\gamma_x} \right| d\mathbb{P}_{X=x} = \gamma_x E(|Z|), \quad Z \sim N(0,1)$$

Como $E(|Z|) = \sqrt{\frac{2}{\pi}}$, então

$$\rho(Md(\theta|x), \mathbb{P}_{X=x}) = \sqrt{\frac{2}{\pi}} \sqrt{\frac{\sigma_0^2 \cdot \upsilon_0^2}{\upsilon_0^2 + n\sigma_0^2}} \quad , \forall x \in \mathbb{R}^n$$

No problema pré-experimentação,

$$\delta^*(X) = Md(\theta/X) = \frac{\upsilon_0^2 \theta_0 + \sigma_0^2 n \bar{X}}{\upsilon_0^2 + n\sigma_0^2}$$

$$\rho^*(\mathbb{P}) = \int_x f(x) \sqrt{\frac{2\sigma_0^2 \upsilon_0^2}{\pi(\upsilon_0^2 + n\sigma_0^2)}} dx = \sqrt{\frac{2\sigma_0^2 \upsilon_0^2}{\pi(\sigma_0^2 + n\sigma_0^2)}}$$

Exemplo 2: $X_1, \ldots, X_n$, dado θ, c.i.i.d. $\mathcal{U}(0,\theta)$, $\theta > 0$.
Suponha $\theta \sim$ Pareto(a,b), $a, b > 0$, $a > 1$.

Vimos que $\theta | X_1 = x_1, \ldots, X_n = x_n \sim$ Pareto$(a+n, \max\{b, x_1, \ldots, x_n\})$

A priori, queremos encontrar m_0 tal que $\mathbb{P}(\theta \geq m_0) = 1/2$.

$$\mathbb{P}(\theta \geq m_0) = \int_{m_0}^{\infty} \frac{ab^a}{\theta^{a+1}} d\theta = ab^a \left. \frac{\theta^{-a}}{-a} \right|_{m_0}^{\infty} = \cancel{a} b^a \frac{1}{\cancel{a}} \frac{1}{m_0^a} = \frac{1}{2} \Rightarrow$$

$$m_0^a = 2b^a \Rightarrow m_0 = \sqrt[a]{2}\, b \quad \text{, i.e. , } m_0 = 2^{1/a} b.$$

Após a observação $X = x$, a estimativa de Bayes é

$$Md(\theta/x) = 2^{\frac{1}{a+n}} \times \max\{b, x_1, \ldots, x_n\}$$

No problema pré-experimentação,

$$\delta^*(X) = Md(\Theta/X) = 2^{\frac{1}{a+n}} \max\{b, X_1, \dots, X_n\}$$

é o estimador de Bayes em relação à perda absoluta.

No problema pré-experimentação,

$$\rho^*(P') = \int_x f(x) \left[\int_\Theta \left| \Theta - 2^{1/a+n} \max\{b, x_1, \dots, x_n\} \right| f(\Theta / x)\, d\Theta \right] dx$$

$$= \int \dots \int_{\Theta > 2^{1/a+n} \max} \dots d\Theta\, dx + \int\int_{\Theta < 2^{1/a+n} \max} \dots d\Theta\, dx$$

$$\max\{b, x_1, \dots, x_n\} \frac{a+n}{a+n-1} \left(2^{\frac{1}{a+n}} - 1 \right)$$

$$= \frac{a+n}{a+n-1} \left(2^{1/a+n} - 1 \right) E\left[\max\{b, X_1, \dots, X_n\} \right]$$

$$= \frac{a+n}{a+n-1} \left(2^{1/a+n} - 1 \right) E\left[E\left[\max\{b, X_1, \dots, X_n\} / \Theta \right] \right]$$

Depois de longa jornada, temos

$$\rho^*(P') = \frac{ab\,[n(a+n) + (a-1)]}{(n+1)(a+n-1)(a-1)} \left(2^{1/a+n} - 1 \right).$$

Ex. 3:

Θ brancas
$10 - \Theta$ verdes

$\Theta \sim Bin(10, 1/3)$ $\quad X/\Theta \sim HG(\Theta, 10 - \Theta, 2)$

$\Theta - x / X = x \sim Bin(8, 1/3)$

Para o modelo Binomial $(8, 1/3)$,

Se $Y \sim Bin(8, 1/3)$, temos

$$P(Y \leq 2) = \frac{1024}{2187}$$

$$P(Y \leq 3) = \frac{4864}{6561} \geq 1/2$$

. 3 é mediana

$$P(Y \geq 3) = 1 - P(Y \leq 2) = 1 - \frac{1024}{2187} \geq \frac{1}{2} \quad e,$$

consequentemente 3 é mediana de Y.

Assim,

$$Md(\theta / x) = x + 3.$$

Antes do experimento,

$Md(\theta / X) = X + 3$ é o estimador de Bayes com relação à perda absoluta.

Nesse caso, o Risco de Bayes é dado por

$$\rho^{*}(\delta^{*}) = \sum_{x=0}^{2} \sum_{j=0}^{10} |(x+3) - j| \, P(\theta = j, X = x) =$$

$$= \sum_{x=0}^{2} \sum_{j=0}^{10} |(x+3) - j| \frac{\binom{j}{x}\binom{10-j}{2-x}}{\binom{10}{2}} \binom{10}{j} \left(\frac{1}{3}\right)^{j} \left(\frac{2}{3}\right)^{10-j}$$

(3) Perda 0-1

Θ é discreto: $\Theta = \{\theta_1, \theta_2, \dots\}$

$$L(d,\theta) = \begin{cases} 0, & d = \theta \\ 1, & d \neq \theta. \end{cases}$$

Para $d \in \mathcal{D} = \Theta$,

$$\mathbb{E}[L(d,\theta)] = \sum_{\theta \in \Theta} L(d,\theta)\, \mathbb{P}(\theta)$$

$$= \underbrace{L(d,d)}_{0}\, \mathbb{P}(d) + \sum_{\theta \neq d} \underbrace{L(d,\theta)}_{1}\, \mathbb{P}(\theta) \Rightarrow$$

$$\Rightarrow \mathbb{E}[L(d,\theta)] = 1 - \mathbb{P}(d)$$

Assim, a decisão de Bayes é $d^* \in \mathcal{D} = \Theta$ t.q.

$$\mathbb{P}(d^*) \geq \mathbb{P}(d), \ \forall d \in \mathcal{D} = \Theta.$$

Isto é, uma moda da distribuição de Θ é decisão de Bayes.

Após observação $X = x$, a estimativa de Bayes com relação à perda 0-1 é $Mo(\Theta / x)$.

Pré-experimentação: $\delta^*(X) = Mo(\Theta / X)$ é estimador de Bayes.

No ex. 3,

$\Theta - x \mid X = x \sim Bin(8, 1/3)$. moda prox. $np = 8/3 = 2,6\ldots$

$\mathbb{P}\binom{8}{2}(1/3)^2(2/3)^6 = 28 \cdot 2^6/3^8$

$\mathbb{P}\binom{8}{3}(1/3)^3(2/3)^5 = 56 \cdot 2^5/3^8$

Assim $x+2$ e $x+3$ são modas da distribuição de Θ/x

$\delta_1^*(X) = X+2$ e $\delta_2^*(X) = X+3$ são est. de Bayes com relação à perda 0-1.

Em geral, podemos considerar a seguinte perda.

$$L(d,\theta) = \begin{cases} 1, & |d-\theta| \geq c \\ 0, & |d-\theta| < c, \quad \underline{c>0} \end{cases}$$

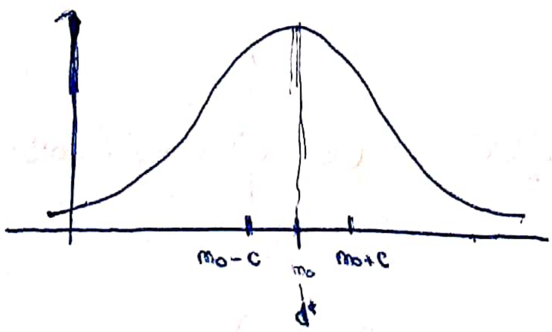
$$E[L(d,\theta)] = \int_{\Theta} L(d,\theta)\, d\mathbb{P} = \int_{\{|d-\theta|<c\}} 0 .\, d\mathbb{P} + \int_{\{|d-\theta|\geq c\}} 1\, d\mathbb{P}$$
$$= \mathbb{P}(|d-\theta| \geq c)$$
$$= 1 - \mathbb{P}(|d-\theta| < c) \Rightarrow$$

$$E[L(d,\theta)] = 1 - \mathbb{P}(d-c < \theta < d+c)$$

Assim, a decisão de Bayes será d^* tal que

$$\mathbb{P}(d^*-c < \theta < d^*+c) \geq \mathbb{P}(d-c < \theta < d+c), \ \forall\, d \in \mathcal{D}, \text{ i.e.}$$

d^* é o ponto central do intervalo modal de comprimento $2c$.



c_n sequ. decrescente t.q. $c_n \downarrow 0$ então (Em geral)

$$d^*(c_n) \to mo$$

No exemplo no caso multivariado

$\Theta = (\theta_1, \theta_2, \ldots, \theta_k) \quad (\Theta \subseteq \mathbb{R}^k)$

$d = (d_1, d_2, \ldots, d_k)$

$L(d,\theta) = (d-\theta)A(d-\theta)'$, $A_{k\times k}$ simétrica e tal que $xA'x \geq 0$, $\forall x \in \mathbb{R}^k$

(A é semidefinida positiva).

Suponha ainda que

$E[\theta_i^2] < \infty, \; \forall i = 1, \ldots, k$

Seja $\mu = (E(\theta_1), E(\theta_2), \ldots, E(\theta_k))$ vetor de média de Θ e

$\Sigma_{k\times k}$ a matriz de covariância de Θ $\quad (\Sigma_{ij} = Cov(\theta_i, \theta_j))$

$$E[L(d,\theta)] = E[(d-\theta)A(d-\theta)'] = E[(d-\mu+\mu-\theta)A(d-\mu+\mu-\theta)']$$

$$= E[(d-\mu)A(d-\mu)' + (d-\mu)A(\mu-\theta)' + (\mu-\theta)A(d-\mu) + (\mu-\theta)A(\mu-\theta)']$$

$$= \underbrace{(d-\mu)A(d-\mu)'}_{\geq 0 \text{ pois } A \text{ é sem def.} \;\Rightarrow\; d^* = \mu} + \underbrace{E[(\mu-\theta)A(\mu-\theta)']}_{\text{não dep. de } d}$$

Logo, como A é semi-definida positiva, μ é a decisão de Bayes contra $\mathbb{P}$.

A priori,

$$\rho^{*}(\mathbb{P}) = E[L(\mu,\theta)] = E[(\theta-\mu)A(\theta-\mu)']$$

$$= E\left[\sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}(\theta_i-\mu_i)(\theta_j-\mu_j)\right]$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}\, \underbrace{E[(\theta_i-E(\theta_i))(\theta_j-E(\theta_j))]}_{\Sigma_{ij} = Cov(\theta_i,\theta_j)} \Rightarrow$$

$$\rho^{*}(\mathbb{P}) = \text{traço}(A\Sigma)$$

Após observação $X=x$, o est. de Bayes é

$$\mu_x = (E(\theta_1/x), E(\theta_2/x), \dots, E(\theta_k/x))$$

e

$$\rho^{*}(\mu_x, \mathbb{P}_{X=x}) = tr(A\Sigma_x), \text{ onde } \Sigma_{x_{ij}} = Cov(\theta_i,\theta_j \mid x) = E(\theta_i\theta_j/x) - E(\theta_i/x)E(\theta_j/x).$$

Finalmente, pré-exp.,

$\delta^{*}(X) = \mu_X = (E(\theta_1/X), \dots, E(\theta_k/X))$ é estim. de Bayes contra L.

$\Sigma_{X_{ij}} = Cov(\theta_i,\theta_j/X)$

Risco de Bayes:

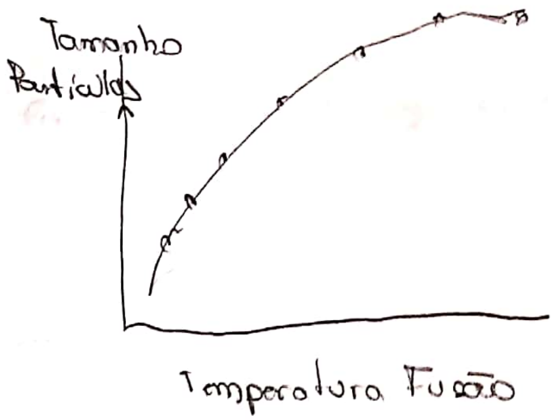
$$\rho^{**}(\mathbb{P}') = \int_{\mathcal{X}} f(x)\left[\int_{\Theta} L(\delta(x),\theta) f(\theta/x)\, d\theta\right] dx = E[tr(A\Sigma_X)]$$

Aula 19, Apr.

. Teoria da Decisão, Palestra Professor da Pol.

Risco em nanotecnologia!

4 d.C., vidros com nanopart. de ouro cobre
Faraday - tamanho da part. altera caract.
Feynman - Partículas pequenas devem ser ct."
Einsten - Cálc. tamanho partículas nanossacarídeos
1980, IBM desenv. tecnol p/ lidar com átomos
Ågimo ñ ganhou pelos nano cubos
Pol. da terra do gerom nanomateriais
- Alterar estrutura das partícula,
Termo de Energia Livre passa a ter impacto maior



Italiano
risco - riscare - navegar sobre rochas perigosas

Risco { Severidade / Probabilidade

www.hsenano.com → relatório de agências de

+100 000 000 ag. quími registrados

Aula 19, Apresentação

Controle: Por bandas

- Categoria 1 : Muito Bom
- " 2 : Bom
- " 3 : Ruim
- " 4 : Mto ruim

- Pequeno
- Médio
- Longo

- Seja $\Theta \in \Theta = $ [illegible], onde Θ é ~~[illegible]~~ número de nanop. em ~~[illegible]~~ a cada mm^3.

1 iguais
3 dif peq
5 dif grande
7 dif muit. grande
9 absurda

$$\begin{array}{c|cccc} & A & B & C & D \\ \hline A & 1 & a & b & c \\ B & 1/a & 1 & d & e \\ C & 1/b & 1/d & 1 & f \\ D & 1/c & 1/e & 1/f & 1 \end{array}$$

Bibl.: pessoas certas p/ dar as notas corretas

Indice de Cons.

$$\frac{\lambda_{max} - n}{n-1}$$

$$CR = \frac{CI}{RI(n)}$$

limite fagocitose: 15 nanomicros.
└ dissolve o material
ou não e fica encapsulado

Teoria da Decisão. Lista 4 na pasta

• Aula 20 04/11/15

$\Theta = (\Theta_1, \dots, \Theta_w)$
$d = (d_1, \dots, d_w)$

$$L(d,\Theta) = (d-\Theta) A (d-\Theta)'$$

$$\delta^*(X) = (E(\Theta_1/X), E(\Theta_2/X), \dots, E(\Theta_w/X))$$

Risco de Bayes $= E(tr\, A \Sigma_x)$

Ex, $X_1, \dots, X_n$, dado $\Theta = (\Theta_1, \Theta_2)$, c.i.i.d. $N(\Theta_1, v_0/\Theta_2)$, $v_0 > 0$.

$\Theta_2 \sim Gama(a,b)$, $a,b>0$
$\Theta_1/\Theta_2 \sim N\left(\mu, \frac{\sigma_0}{\Theta_2}\right)$, $\mu_0 \in R$, $\sigma_0 > 0$

Vimos que

$$\Theta_2 / X = x \sim Gama\left(a + n/2, b + n/2 \frac{(\bar{x}-\mu_0)}{v_0 + n\sigma_0} + \frac{1}{2v_0}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)$$

$$\Theta_1/\Theta_2, X = x \sim N\left(\frac{v_0\mu_0 + n\sigma_0\bar{x}}{v_0 + n\sigma_0}, \frac{\sigma_0 v_0}{(v_0 + n\sigma_0)\Theta_2}\right)$$

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Pelo resultado,

$$\delta^*(X) = \left(E(\Theta_1/X), E(\Theta_2/X) \right)$$

$$= \left(\frac{v_0 \mu_0 + n\sigma_0 \bar{X}}{v_0 + n\sigma_0}, \frac{a + n/2}{b + \frac{n}{2}\frac{(\bar{X}-\mu_0)^2}{v_0+n\sigma_0} + \frac{1}{2v_0}\sum_{i=1}^{n}(x_i-\bar{x})^2} \right)$$

$$E(\Theta_1/X=x) = E\left(E(\Theta_1/\Theta_2, X=x)\right) = \frac{v_0\mu_0 + n\sigma_0\bar{x}}{v_0 + n\sigma_0}$$

Risco de Bayes

$$= E(\mathrm{tr}\,\Sigma_x) = E\left[Var(\Theta_1/X) + Var(\Theta_2/X)\right]$$
$$= Var\,\Theta_1 - Var(E(\Theta_1/X)) + Var\,\Theta_2 - Var(E(\Theta_2/X))$$

Estimação Por Intervalo

- Schervish, Robert, Inoue, French-Rios

$\Theta \subseteq \mathbb{R}$,

$D = \{[a,b], a, b \in \mathbb{R}, a < b\}$

$$L([a,b],\theta) = h(b-a) + k\,\mathbb{I}_{[a,b]^c}(\theta)$$

h não-decrescente

Dentre as penalidades dessa "natureza", vamos considerar a seguinte:

$$L([a,b],\theta) = \begin{cases} b-a, & \theta \in [a,b] \\ b-a + c_1(a-\theta), & \theta < a \\ b-a + c_2(\theta-b), & \theta > b \end{cases}$$

Reescrevemos L da seguinte maneira.

$$L([a,b],\theta) = -a + c_1(a-\theta)\mathbb{I}_{(-\infty,a)}(\theta) + b + c_2(\theta-b)\mathbb{I}_{(b,\infty)}(\theta)$$

$$= (\theta-a) + c_1(a-\theta)\mathbb{I}_{(-\infty,a)}(\theta) + (b-\theta) + c_2(\theta-b)\mathbb{I}_{(b,\infty)}(\theta) \Rightarrow$$

$$L([a,b],\theta) = \overbrace{(\theta-a)\mathbb{I}_{[a,\infty)}(\theta) + (c_1-1)(a-\theta)\mathbb{I}_{(-\infty,a)}(\theta)}^{L_1(a,\theta)} + \underbrace{(b-\theta)\mathbb{I}_{(-\infty,b]}(\theta) + (c_2-1)(\theta-b)\mathbb{I}_{(b,\infty)}(\theta)}_{L_2(b,\theta)}$$

$$L([a,b],\theta) = L_1(a,\theta) + L_2(b,\theta)$$

Supondo $E(\theta) < \infty$, temos

$$E[L([a,b],\theta)] = E[L_1(a,\theta)] + E[L_2(b,\theta)]$$

$$L_1(a,\theta) = \begin{cases} 1\,(\theta-a), & a \leq \theta \\ (c_1-1)(a-\theta), & a > \theta \end{cases}$$

$a^* \rightarrow$ percentil de ordem $\dfrac{1}{1+(c_1-1)} = \dfrac{1}{c_1}$

Pode-se provar (Ex. L. 4) que $E[L_1(a,\theta)]$ atinge valor mínimo em

a^* = percentil de ordem $1/_{1+(c_1-1)} = 1/c_1$ de $\mathbb{P}$.

$c_1 > 1$ p/ estar bem def.

Do mesmo modo, escrevemos

$$L_2(b,\theta) = \begin{cases} (c_2-1)(\theta-b), & \theta > b \\ b-\theta & , \theta \leq b \end{cases}$$

Como no caso anterior, $E[L_2(b,\theta)]$ é mínimo em

$$b^* = \text{percentil de ordem } \frac{c_2-1}{c_2-1+1} = \frac{c_2-1}{c_2} \text{ de } \mathbb{P}$$

Assim, se $\frac{1}{c_1} \leq \frac{c_2-1}{c_2}$, $[a^*, b^*]$ é decisão de Bayes contra $\mathbb{P}$.

Em particular, se $c_1 = 2/\alpha = c_2$, $0 < \alpha < 1$, temos

$$\frac{1}{c_1} = \frac{\alpha}{2} \quad \text{e} \quad \frac{c_2-1}{c_2} = \frac{2/\alpha - 1}{2/\alpha} = 1 - \frac{\alpha}{2}.$$

Exemplo J: $X_1, \dots, X_n$, dado θ, c.i.i.d. $N(\theta, v_0^2)$

$\theta \sim N(\theta_0, \sigma_0^2)$

Vimos que

$$\theta \mid X_1 = x_1, \dots, X_n = x_n \sim N\left(\underbrace{\frac{v_0^2\theta_0 + n\sigma_0^2\bar{x}}{v_0^2 + n\sigma_0^2}}_{\mu_x}, \frac{\sigma_0^2 v_0^2}{v_0^2 + n\sigma_0^2} \right)$$

$$\left. \frac{\theta - \mu_x}{\sigma_x} \right| X = x \sim N(0,1)$$

$\Theta \subseteq \mathbb{R}$

Em muitos casos, optaremos por um conjunto (ou região) para Θ.
Suponha que a priori, a distribuição de θ, $\mathbb{P}: \mathcal{F} \to [0,1]$, absolutamente contínua (Lebesgue) ($\mathbb{P} \ll \lambda_1$).

Seja f a densidade de θ.

Definição: Dizemos que $A \in \mathcal{F}$ é uma região (conjunto) HIGHEST DENSITY de probabilidade γ, $0 < \gamma < 1$, se

1. $\mathbb{P}(\theta \in A) \geq \gamma$
2. $f(\theta') \geq f(\theta'')$, $\forall \theta' \in A$, $\forall \theta'' \in A^c$

Se para todo $x \in \mathcal{X}$, $\mathbb{P}_x \ll \lambda_1$, com densidade $f(\cdot|x)$, dizemos que $A \in \mathcal{F}$ atendendo (1) e (2) é uma região "HIGHEST POSTERIOR DENSITY" de probabilidade γ, ou simplesmente, região HPD.

Vamos considerar então um novo prob. de decisão em que

$D = \mathcal{F}$

$\Theta \subseteq \mathbb{R}^k$, $D = \mathcal{F}$

Para $A \in D$ e $\theta \in \Theta$

$L(A,\theta) = \lambda_k(A) + c\,\mathbb{I}_{A^c}(\theta)$

$$L(A,\theta) = \lambda_1(A) + c\,\mathbb{I}_{A^c}(\theta)$$

Para $A \in D$,

$$\begin{aligned} E[L(A,\theta)] &= E[\lambda_1(A) + c\,\mathbb{I}_{A^c}(\theta)] \\ &= \lambda_1(A) + c\,E[\mathbb{I}_{A^c}(\theta)] \\ &= \lambda_1(A) + c\,\mathbb{P}(\theta \in A^c) \\ &= \lambda_1(A) + c - c\,\mathbb{P}(\theta \in A) \end{aligned}$$

$\lambda(A) = \int_A d\lambda$

$$= c + \int_A 1 \, d\lambda_1 - c \int_A f \, d\lambda_1$$

$$= c + \int_A 1 - cf \, d\lambda_1$$

$$\mathbb{E}[L(A,\Theta)] = c + \int_A [1 - cf] \, d\lambda_1$$

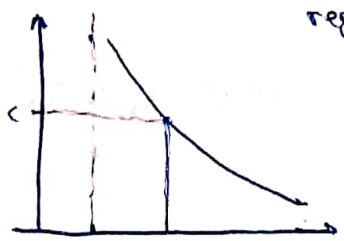
$\mathbb{E}[L(A,\Theta)]$ atinge valor mínimo em

$$A^* = \{\theta \in \Theta : 1 - cf(\theta) \leq 0\}, \text{ isto é,}$$

$$A^* = \{\theta \in \Theta : f(\theta) \geq 1/c\}$$

Voltando ao Ex 2.

Fixado $\gamma \in (0,1)$, $A^* = [\max\{b, x_{(n)}\}, \max\{b, x_{(n)}\} / ((1-\gamma)^{1/(a+n)})]$ é região HPD de prob. γ.



Exemplo 2: $X_1, \dots, X_n$, dado Θ, c.i.i.d. Ber$(\max\{\theta, 1-\theta\})$ (Não há identificabilidade! Prob: Pode não exist. $\tilde{\delta}_{MV}$, conv., entre outros)

$\Theta \sim$ Uniforme$(0,1)$. Para $x_1, \dots, x_n \in \{0,1\}$, temos

$$f(\theta \mid x_1, \dots, x_n) \propto P(X_1 = x_1, \dots, X_n = x_n \mid \theta) f(\theta) = \prod_{i=1}^{n} (\max\{\theta, 1-\theta\})^{x_i} [1 - \max\{\theta, 1-\theta\}]^{1-x_i} \mathbb{1}$$

$$= (\max\{\theta, 1-\theta\})^{\sum_{i=1}^n x_i} (1 - \max\{\theta, 1-\theta\})^{n - \sum_{i=1}^n x_i} \Rightarrow$$

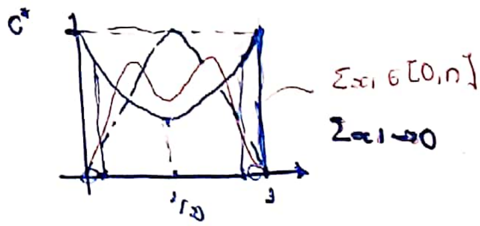
$$f(\theta|x_1,\dots,x_n) \propto \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}\,\mathbb{I}_{[1/2,1]} + (1-\theta)^{\sum x_i}\,\theta^{n-\sum x_i}\,\mathbb{I}_{[0,1/2)}(\theta) \quad \boxed{*}$$

Assim,

$$f(\theta|x_1,\dots,x_n) = \begin{cases} C^*\,\theta^{\sum x_i}(1-\theta)^{n-\sum x_i}, & \theta \in [1/2,1] \\ C^*\,\theta^{n-\sum x_i}(1-\theta)^{\sum x_i}, & \theta \in [0,1/2) \end{cases}$$

$$C^* = \left(\int_0^1 \boxed{*}\, d\theta\right)^{-1}$$

Para $x_i = 1$, $i = 1,\dots,n$, temos



$$f(\theta|x_1,\dots,x_n) = \begin{cases} C^*\,\theta^n, & \theta \in [1/2,1] \\ C\,(1-\theta)^n, & \theta \in [0,1/2) \end{cases}$$

Testes de Hipóteses — Aula ~~20~~ 21, 09/11

Θ, $\mathcal{F}$ → σ-álgebra de Θ

Vamos pegar $\Theta_0, \Theta_1 \in \mathcal{F}$ e $\Theta = \Theta_0 \cup \Theta_1$ e $\Theta_0 \cap \Theta_1 = \emptyset$

$H_0: \theta \in \Theta_0$,

$H_1: \theta \in \Theta_1$.

Neste caso: $D = \{\overset{0}{d_0}, \overset{1}{d_1}\}$

d_0: Não rejeitar H_0; d_1: rejeitar.

Consideremos as seguintes penalidades

	$\theta \in \Theta_0$	$\theta \in \Theta_1$
d_0	0	a_2
d_1	a_1	0

("mais simples função de perda")

Você observa $X = x$ $\left(\{d_0, d_1\}, \Theta, L, \mathbb{P}_{x=x}\right)$

$$\rho(d_0, \mathbb{P}_{x=x}) = \int_{\Theta} L(d_0,\theta)\, d\mathbb{P}_x = \int_{\Theta_0} 0\, d\mathbb{P}_x + \int_{\Theta_1} a_2\, d\mathbb{P}_x = a_2\, \mathbb{P}(\theta \in \Theta_1 | x)$$

Também

$$\rho(d_1, \mathbb{P}_x) = \int_{\Theta} L(d_1,\theta)\, d\mathbb{P}_x = \int_{\Theta_0} a_1\, d\mathbb{P}_x = a_1\, \mathbb{P}(\theta \in \Theta_0 | x).$$

→ Depende de ~~...~~ comparar os riscos

Ao observar $X=x$ decidimos por H_1 se

$$\boxed{\rho(d_1, \mathbb{P}_x) < \rho(d_0, \mathbb{P}_x)} \iff a_1 \mathbb{P}(\theta \in \Theta_0 | x) < a_2 \mathbb{P}(\theta \in \Theta_1 | x) \iff$$

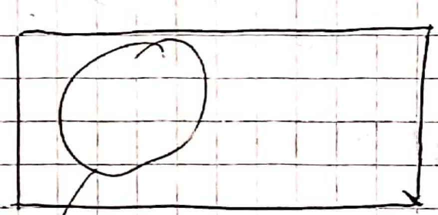
$$a_1 \mathbb{P}(\theta \in \Theta_0 | x) < a_2 - a_2 \mathbb{P}(\theta \in \Theta_0 | x) \iff$$

$$\iff \mathbb{P}(\theta \in \Theta_0 | x) < \frac{a_2}{a_1+a_2} \quad \left(\text{ou alt. } \mathbb{P}(\theta \in \Theta_1 | x) > \frac{a_1}{a_1+a_2}\right)$$

No problema pré-experimentação, a regra de decisão de Bayes (ou teste de Bayes) é dada por.

$$\varphi^*(x) = \begin{cases} 1, & \mathbb{P}(\theta \in \Theta_1 | x) > \frac{a_1}{a_1+a_2} \\ 0, & \mathbb{P}(\theta \in \Theta_1 | x) < \frac{a_1}{a_1+a_2} \end{cases}$$

$H_0: \theta \in \Theta_0$ e $H_1: \theta \in \Theta_1$



$\Psi = \{\varphi : \mathcal{X} \to \{0,1\}\}$

$\Psi' = \{\varphi \in \Psi : \text{tamanho } \varphi \leq \alpha\}$

Ex. 1: $\Theta = \{1/4, 1/2\}$

$X_1, \dots, X_n$ dado θ, c.i.i. Ber(θ)

Priori:

$\mathbb{P}(\theta = \frac{1}{4}) = \gamma = 1 - \mathbb{P}(\theta = \frac{1}{2})$, $\gamma \in \ell$

Duas hipóteses $H_0: \theta = \frac{1}{4}$ $H_1: \theta = \frac{1}{2}$

"Basta calcular um caso"

$$\mathbb{P}(\theta = \tfrac{1}{2} | X_1 = x_1, \dots) = \frac{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{\sum x_i}\left(\frac{1}{2}\right)^{n-\sum x_i}(1-\gamma)}{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{n-\sum x_i}(1-\gamma) + \left(\frac{1}{4}\right)^{\sum x_i}\left(\frac{3}{4}\right)^{n-\sum x_i}\gamma}$$

$$= \frac{2^n(1-\gamma)}{2^n(1-\gamma) + 3^{n-\sum x_i}\gamma} > \frac{a_1}{a_1+a_2} \iff 2^n(1-\gamma)a_2 > 3^{n-\sum x_i}\gamma a_1$$

$\vdots$

$\iff$

$$(n-\sum x_i)\log 3 < \log \frac{2^n(1-p)a_2}{p\,a_1} \iff$$

$$\vdots$$

$$\sum_{i=1}^{n} x_i > \frac{n\log(3) - \log \frac{2^n(1-p)a_2}{p\,a_1}}{\log 3}$$

$$= n - \frac{\log\left(\frac{a_2}{a_1}\frac{(1-p)}{p}\right)\cdot 2^n}{\log(3)} = K$$

Re-organizado para a interpretação

Logo o teste é dado por:

$$\varphi^*(x_1,\dots,x_n) = \begin{cases} 1, & \sum x_i > K \\ 0, & \text{c.c} \end{cases}$$

Ex. 2: $\Theta = [0,1]$, $\Theta_0 = [0,\frac{1}{2}]$, $\Theta_1 = (\frac{1}{2}, 1]$

$X_1, \dots, X_n$ dado θ, c.i.i.d $Ber(\theta)$

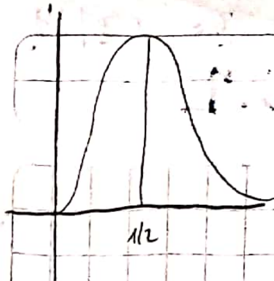
$\theta \sim Beta(a,b)$ com $a, b > 0$

Hipóteses: $H_0: \theta \leq \frac{1}{2}$, $H_1: \theta > \frac{1}{2}$

Resultado: $\theta | X = x \sim Beta(a + \sum x_i, b + n - \sum x_i)$

Rejeitamos H_0, se

$$\mathbb{P}\left(\theta > \frac{1}{2} \middle| x\right) > \frac{a_1}{a_2+a_1} \iff \int_{1/2}^{1} \frac{\Gamma(a+b+n)}{\Gamma(a+\sum x_i)\Gamma(b+n-\sum x_i)} \theta^{a+\sum x_i - 1} (1-\theta)^{b+n-\sum x_i} d\theta > \frac{a_1}{a_1+a_2}$$



Se $a_1 = a_2 = 1$, decidimos por H_1 se e só se

$$P\left(\theta > \frac{1}{2} \middle| x\right) \geq \frac{1}{2} \iff Md(\theta|x) > \frac{1}{2}.$$

Obs: Para $H_0: \theta \leq \theta_0$, $H_1: \theta > \theta_0$.

Logo $P(\theta > \theta_0 | x) > \frac{1}{2} \iff Med(\theta|x) > \theta_0$

Tipo de função de perda: $0-1-c$. Para as últimas hipóteses consideradas o teste utiliza a mediana a posteriori

Ex 3: $\Theta = \mathbb{R}_+^2$, $\theta = (\theta_1, \theta_2)$ $\quad H_0: \theta_1 > \theta_2$, $H_1: \theta_1 < \theta_2$

$X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$, dado θ, são cond ind tais que

$X_i | \theta \sim Poisson(\theta_1)$, $i = 1, \dots, n$ $\quad Y_j | \theta \sim Poisson(\theta_2)$, $j = 1, \dots, m$

Seja $\theta_i \sim Gamma(\mu_i, \nu_i)$, $i = 1, 2$ com $\theta_1 \perp\!\!\!\perp \theta_2$

$$f(\theta | x, y) \propto f(x, y | \theta) f(\theta) \propto \prod_{i=1}^{n_1} f(x_i|\theta) \prod_{j=1}^{n_2} f(y_j|\theta) f(\theta_1) f(\theta_2) \propto$$

$$\propto \theta_1^{\mu_1 + \sum x_i - 1} e^{-(\nu_1 + n_1)\theta_1} \quad * \text{ (Idem para } \theta_2)$$

Isto é: $\theta_1 \perp\!\!\!\perp \theta_2 | x, y$ com $\theta_1 | x, y \sim Gama(\mu_1 + \sum x_i, \nu_1 + n_1)$

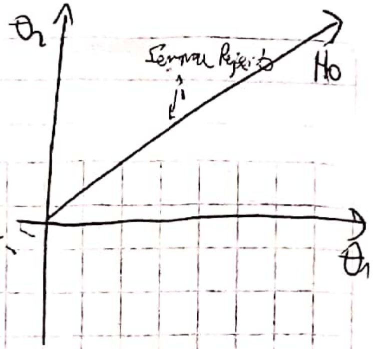
$\theta_2 | x, y \sim Gama(\mu_2 + \sum y_j, \nu_2 + n_2)$

Rejeitamos:

$$P(\theta_1 \geq \theta_2 | x, y) > \frac{a_1}{a_1+a_2}.$$

Ainda mais exemplo

$H_0: \theta_1 = \theta_2$, $H_1: \theta_1 \neq \theta_2$



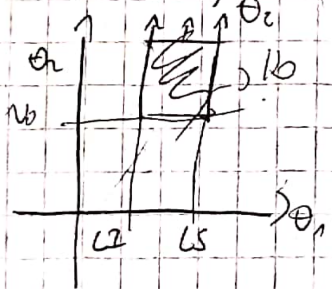
Existem outras alternativas:

$$\mathbb{P}(\theta_1 = \theta_2 \mid x, y) = 0, \quad \forall (x, y) \; !!!$$

Ex 4: $X_1, \dots, X_n$ dado $\theta = (\theta_1, \theta_2)$ c.i.i. $N(\theta, \frac{1}{\theta_2})$

$H_0: [LI, LS] \times [v_0, \infty]$

$H_1: (\quad)^c$



$\theta_2 \sim \Gamma(a', b')$

$\theta_1 \mid \theta_2 \sim N\left(\theta_0, \frac{\sigma_0}{\theta_2}\right)$

Vimos que: $\theta_2 \mid x_1 \dots x_n \sim \Gamma\left(a + \frac{n}{2}, b + \frac{n}{2} \frac{(\bar{x} - \theta_0)}{n\sigma_0 + v_0} + \frac{1}{2v_0} \sum (x_i - \bar{x})^2\right)$

$$\theta_1 \mid \theta_2, x_1, \dots, x_n \sim N\left(\frac{v_0\theta_0 + n\sigma_0\bar{x}}{v_0 + n\sigma_0}, \frac{\sigma_0 v_0}{(v_0 + n\sigma_0)\theta_2}\right)$$

Você decide por H_1 se e só se $\mathbb{P}(\theta \in \Theta_0 \mid x) < \frac{a_2}{a_1 + a_2}$.

11/11/2015

Outras possíveis Perdas

$\Theta = \mathbb{R}$

$\theta_0 \in \mathbb{R}$

$H_0: \theta \leq \theta_0$

$H_1: \theta > \theta_0$

$$\boxed{\begin{aligned} L(d_0, \theta) &= (\theta - \theta_0) I_{(\theta_0, \infty)}(\theta) \\ L(d_1, \theta) &= (\theta_0 - \theta) I_{(-\infty, \theta_0]}(\theta) \end{aligned}}$$

$$\rho(d_0, \mathbb{P}) = \int_{\Theta} L(d_0, \theta)\, d\mathbb{P} = \int_{\Theta_0} 0\, d\mathbb{P} + \int_{\Theta_1} (\theta - \theta_0)\, d\mathbb{P}$$

$$= \int_{(\theta_0, \infty)} \theta\, d\mathbb{P} - \theta_0 \mathbb{P}(\theta \in \Theta_1)$$

Analogamente

$$\rho(d_1, \mathbb{P}) = \int_{\Theta} L(d_1, \theta)\, d\mathbb{P} = \int_{\Theta_0} (\theta_0 - \theta)\, d\mathbb{P} + \int_{\Theta_1} 0\, d\mathbb{P} = \theta_0 \mathbb{P}(\theta \in \Theta_0) - \int_{(-\infty, \theta_0]} \theta\, d\mathbb{P}$$

Decisão por d_1 se:

$$\rho(d_1, \mathbb{P}) < \rho(d_0, \mathbb{P}) \iff \theta_0 \mathbb{P}(\theta \in \Theta_0) - \int_{(-\infty, \theta_0]} \theta\, d\mathbb{P} < \int_{(\theta_0, \infty)} \theta\, d\mathbb{P} - \theta_0 \mathbb{P}(\theta \in \Theta_1)$$

$$\iff \boxed{\theta_0 < E(\theta)}$$

Assim no problema com dados

$$\varphi^*(x) = \begin{cases} 1, & E(\theta \mid x) > \theta_0 \\ 0, & E(\theta \mid x) < \theta_0 \end{cases}$$

Teoria da Decisão

11/11/15 Aula 22

Observações Lista 4.

(1) $L(A,\theta) = |A| + k\,\mathbb{I}_{A^c}(\theta)$

(2) Estudar HPD em função de γ.

Outras penalidades possíveis:

$\Theta = \mathbb{R}$

$\theta_0 \in \mathbb{R}$ $\qquad L(d_0,\theta) = (\theta-\theta_0)\,\mathbb{I}_{(\theta_0,\infty)}(\theta)$

$H_0: \theta \leq \theta_0 \quad (\theta \in \Theta_0)$

$H_1: \theta > \theta_0 \quad (\theta \in \Theta_1) \qquad L(d_1,\theta) = (\theta_0-\theta)\,\mathbb{I}_{(-\infty,\theta_0]}(\theta)$

$$\rho(d_0,P) = \int_{\Theta} L(d_0,\theta)\,dP = \int_{\Theta_0} 0\,dP + \int_{\Theta_1} (\theta-\theta_0)\,dP$$

$$= \int_{(\theta_0,\infty)} \theta\,dP - \theta_0 P(\theta \in \Theta_1)$$

$$\rho(d_1,P) = \int_{\Theta} L(d_1,\theta)\,dP = \int_{\Theta_0} (\theta_0-\theta)\,dP + \int_{\Theta_1} 0\,dP = \theta_0 P(\theta \in \Theta_0) - \int_{(-\infty,\theta_0]} \theta\,dP$$

Decisão por d_1 se

$$\rho(d_1,P) < \rho(d_0,P) \Leftrightarrow \theta_0 P(\theta\in\Theta_0) - \int_{(-\infty,\theta_0]} \theta\,dP < \int_{(\theta_0,\infty)} \theta\,dP - \theta_0 P(\theta\in\Theta_1) \Leftrightarrow$$

$$\Leftrightarrow \theta_0 < E(\theta).$$

Assim, no problema com dados,

$$\varphi^*(x) = \begin{cases} 1, & E(\Theta/x) > \Theta_0 \\ 0, & E(\Theta/x) < \Theta_0 . \end{cases}$$

Ex 1. $X_1, \ldots, X_n$, dado Θ, c.i.i.d. $N(\Theta, v_0^2)$

$\Theta \sim N(\Theta_0, \sigma_0^2)$ $\quad H_0: \Theta \leq \Theta^* \times H_1: \Theta > \Theta^*$

$$\Theta / x_1, \ldots, x_n \sim N\left(\frac{v_0^2 \Theta_0 + n\sigma_0^2 \bar{x}}{v_0^2 + n\sigma_0^2}, \frac{v_0^2 \cdot \sigma_0^2}{v_0^2 + n\sigma_0^2} \right)$$

$\varphi^*(x) = 1$ se $E(\Theta/x) > \Theta^* \Leftrightarrow$

$$\frac{v_0^2 \Theta_0 + n\sigma_0^2 \bar{x}}{v_0^2 + n\sigma_0^2} > \Theta^* \Leftrightarrow \bar{x} > \frac{\Theta^* (v_0^2 + n\sigma_0^2) - v_0^2 \Theta_0}{n\sigma_0^2}$$

Ex. 2. $X_1, \ldots, X_n$, dado Θ, c.i.i.d. $Exp(\Theta)$

$\Theta \sim Gama(a, b)$, $a, b > 0$ $\qquad H_0: \Theta \leq \Theta_0$, $\Theta_0 > 0$

$\Theta / x_1, \ldots, x_n \sim Gama(a+n, b + \sum_{i=1}^n x_i)$ $\qquad H_1: \Theta > \Theta_0$

$$\varphi^*(x) = 1 \text{ se } E(\Theta/x) > \Theta_0 \Leftrightarrow \frac{a+n}{b + \sum x_i} > \Theta_0 \Leftrightarrow \sum_{i=1}^n x_i < (a+n)\Theta_0 - b.$$

Vimos também que para a perda 0-1-1,

$$\varphi'(x) = 1 \text{ se } Md(\Theta/x) > \Theta_0 \text{ é teste de Bayes}$$

para $H_0: \Theta \leq \Theta_0$ versus $H_1: \Theta > \Theta_0$.

Hipóteses Precisas

Suponhamos que $\Theta \subseteq \mathbb{R}^k$ e seja λ_k a medida de Lebesgue em $\mathcal{B}(\mathbb{R}^k)$.
Dizemos que a hipótese nula

$$H_0: \theta \in A \quad , \quad A \in \Theta \cap \mathcal{B}(\mathbb{R}^k)$$

é precisa se $\lambda_k(A) = 0$. e $\lambda_k(\Theta) > 0$.

Exemplo.

(1) $\Theta = \mathbb{R}$

$\Theta_0 = \{3\}$

$H_0: \theta \in \Theta_0 \ (\theta = 3)$

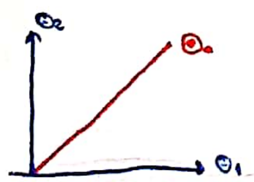
$H_1: \theta \notin \Theta_0$

Claro que $\lambda_1(\{3\}) = 0$. $\{3\}$ é hipótese precisa.

(2)

$\Theta = \mathbb{R}_+^2$

$\Theta_0 = \{(\theta_1, \theta_2) \in \mathbb{R}_+^2 : \theta_1 = \theta_2\}$



$\lambda_2(\Theta_0) = 0$.

Duas abordagens ao problema de testar hipóteses nulas precisas.

(1) Teste de Jeffreys

(2) FBST (Full Bayesian Significance Test)

Voltando ao Ex 1.

$X_1, \dots, X_n$, dado Θ, c.i.i.d. $N(\Theta, v_0^2)$

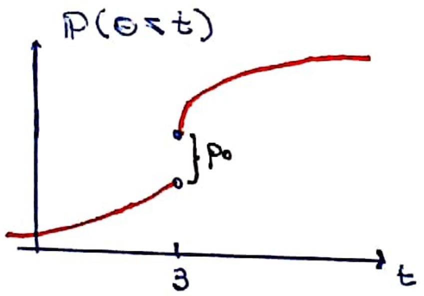
$H_0: \Theta = 3$

$H_1: \Theta \neq 3$

Vamos supor, a priori, uma medida de probabilidade mista $\mathbb{P}^*: \mathcal{B}(\mathbb{R}) \to [0,1]$

$$\mathbb{P}^*(A) = p_0 \, \mathbb{I}_A(3) + (1-p_0) \int_A \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}} \, d\theta$$

Alternativamente,



$\mathbb{P}^*(\Theta=3) = p_0$

$$\mathbb{P}^*(\Theta=3 / x_1, \dots, x_n) = \frac{f(x_1, \dots, x_n / \Theta=3)\, \mathbb{P}^*(\Theta=3)}{f(x_1, \dots, x_n / \Theta=3)\, \mathbb{P}^*(\Theta=3) + \int_{\mathbb{R}\setminus\{3\}} f(x_1, \dots, x_n \mid \Theta) \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}} d\theta \, (1-p_0)}$$

$$= \frac{e^{-\sum_{i=1}^n (x_i-3)^2/2v_0^2} \, p_0}{e^{-\sum_{i=1}^n (x_i-3)^2/2v_0^2} \, p_0 + (1-p_0) \int_{\mathbb{R}} e^{-\sum_{i=1}^n (x_i-\theta)^2/2v_0^2} \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-(\theta-\theta_0)^2/2\sigma_0^2} \, d\theta}$$

$$= \frac{p_0}{p_0 + \frac{(1-p_0)}{\sqrt{2\pi\sigma_0^2}} \int_{\mathbb{R}} e^{\frac{\sum(x_i-3)^2}{2v_0^2} - \frac{\sum(x_i-\theta)^2}{2v_0^2} - \frac{(\theta-\theta_0)^2}{2\sigma_0^2}} \, d\theta}$$

$$= \frac{p_0}{p_0 + \frac{(1-p_0)}{\sqrt{2\pi\sigma_0^2}} e^{\frac{9n - 6n\bar{x}}{2v_0^2} - \theta_0^2/2\sigma_0^2} \int_{\mathbb{R}} e^{-\frac{1}{2}\left\{\theta^2\left[n/v_0^2 + \frac{1}{\sigma_0^2}\right] - 2\theta\left[n\bar{x}/v_0^2 + \theta_0/\sigma_0^2\right]\right\}}}$$

Você decide por H_0 se

$$\mathbb{P}^*(\Theta = 3 \mid x) > \frac{a_2}{a_1 + a_2}$$

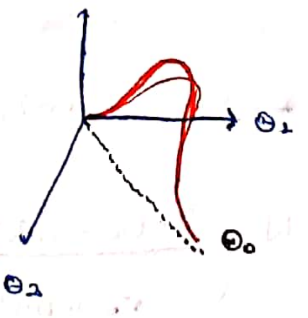
Exemplo 2.

$X_1, X_2, \ldots, X_{n_1}$, $Y_1, \ldots, Y_{n_2}$, dado $\Theta = (\Theta_1, \Theta_2)$, c. independentes tais que

$$X_i \mid \Theta \sim \text{Poisson}(\Theta_1),\ i = 1, \ldots, n_1$$
$$Y_j \mid \Theta \sim \text{Poisson}(\Theta_2),\ j = 1, \ldots, n_2$$

$H_0: \Theta_1 = \Theta_2$

$H_1: \Theta_1 \neq \Theta_2$



Vamos adotar a priori $\mathbb{P}^*$ dada por

$$\mathbb{P}^*(A) = p_0 \int_{A \cap \Theta_0} \underbrace{\frac{b_0^{a_0}}{\Gamma(a_0)} \theta^{a_0 - 1} e^{-b_0\theta}}_{f_1(\theta')} d\theta + (1-p_0) \int_A \underbrace{\frac{v_1^{\mu_1}}{\Gamma(\mu_1)} \theta_1^{\mu_1 - 1} e^{-v_1\theta_1} \cdot \frac{v_2^{\mu_2}}{\Gamma(\mu_2)} \theta_2^{\mu_2 - 1} e^{-v_2\theta_2}}_{f_2(\theta_1, \theta_2)} d\theta_1 d\theta_2.$$

$$\mathbb{P}^*(\Theta_0) = p_0.$$

$$\mathbb{P}^*(\Theta_1 = \Theta_2 \mid (x, y)) = \frac{p_0 \int_{\mathbb{R}_+} f(x, y \mid \theta') f_1(\theta') d\theta'}{p_0 \int_{\mathbb{R}_+} f(x, y \mid \theta') f_1(\theta') d\theta' + (1-p_0) \int_{\mathbb{R}^2_+} f(x, y \mid \theta_1, \theta_2) f_2(\theta_1, \theta_2) d\theta_1 d\theta_2}$$

$$= \frac{p_0 \int_{\mathbb{R}_+} e^{-(n_1+n_2)\theta'} \, \theta'^{\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j} \cdot \frac{b_0^{a_0}}{\Gamma(a_0)} \theta_1^{a_0-1} e^{-b_0\theta'} \, d\theta'}{+ (1-p) \int_{\mathbb{R}_+^2} e^{-n_1\theta_1} \theta_1^{\sum x_i} e^{-n_2\theta_2} \theta_2^{\sum^{n_2} y_j} \frac{v_1^{\mu_1}}{\Gamma(\mu_1)} \theta_1^{\mu_1-1} e^{-v_1\theta_1} \frac{v_2^{\mu_2}}{\Gamma(\mu_2)} \theta_2^{\mu_2-1} e^{-v_2\theta_2} \, d\theta_1 d\theta_2}$$

$$P(\theta_1=\theta_2 / (x,y)) = \frac{\frac{p_0 \, b_0^{a_0}}{\Gamma(a_0)} \frac{\Gamma(a_0+\sum^{n_1} x_i + \sum^{n_2} y_j)}{(b_0+n_1+n_2)^{a_0+\sum x_i+\sum y_j}}}{\frac{p_0 \, b_0^{a_0}}{\Gamma(a_0)} \frac{\Gamma(a_0+\sum x_i+\sum y_j)}{(b_0+n_1+n_2)^{a_0+\sum x_i+\sum y_j}} + \frac{(1-p)}{\Gamma(\mu_1)\Gamma(\mu_2)} \frac{\Gamma(\mu_1+\sum^{n_1} x_i)}{(v_1+n_1)^{\mu_1+\sum^{n_1} x_i}} \frac{\Gamma(\mu_2+\sum^{n_2} y_j)}{(v_2+n_2)^{\mu_2+\sum^{n_2} y_j}}}$$

$$\overset{?}{\gtrless} \frac{a_2}{a_1+a_2}$$

2. FBST

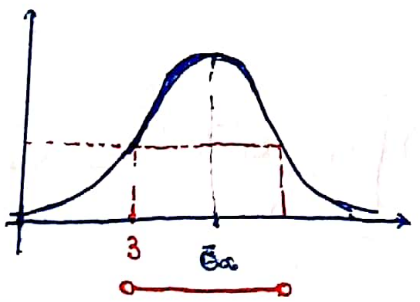
Voltando ao exemplo 1.

$H_0: \theta = 3$

$H_1: \theta \neq 3$

Após observar $x = (x_1, \dots, x_n)$,

$$\theta / X = x \sim N\left(\underbrace{\frac{v_0^2 \theta_0 + n\sigma_0^2 \bar{x}}{v_0^2 + n\sigma_0^2}}_{\bar{\theta}_x}, \frac{\sigma_0^2 v_0^2}{v_0^2 + n\sigma_0^2} \right)$$



$T_x = \{\theta \in \Theta : f(\theta|x) > f(3|x)\}$

conjunto tangente. Em geral, um HPD.

Evidência de Pereira-Stern. $Ev(\{3\}, x) = 1 - P(\theta \in T_x | x)$

Ex. 2.

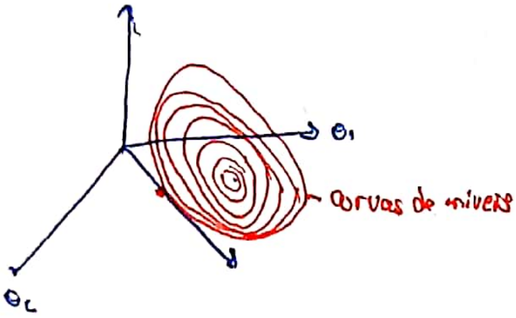
$\Theta_1 \perp\!\!\!\perp \Theta_2$ e $\Theta_i \sim \text{Gama}(\mu_i, v_i)$, $i=1,2$.

Então

$$\Theta_1 \perp\!\!\!\perp \Theta_2 \mid (x,y)$$

$\Theta_1 \mid (x,y) \sim \text{Gama}(\mu_1 + \sum^{n_1} x_i, v_1+n_1)$

$\Theta_2 \mid (x,y) \sim \text{Gama}(\mu_2 + \sum^{n_2} y_j, v_2+n_2)$



$$T_{(x,y)} = \left\{ \theta \in \Theta : f(\theta \mid (x,y)) > \sup_{\theta \in \Theta_0} f(\theta' \mid (x,y)) \right\}$$

Nesse caso,

$$Ev(\{\theta_1 = \theta_2\}, (x,y)) = 1 - P(\theta \in T_{(x,y)} \mid (x,y))$$

Entropy (1999). "Evidence and credibility: a full bayesian sig. teste ... "

Teoria da Decisão

Aula 25 23/11/15

Agrupamentos de Probabilidades

$(\Theta, \mathcal{F})$

Ind. 1 $\rightarrow P_1 : \mathcal{F} \rightarrow [0,1]$

Ind. 2 $\rightarrow P_2 : \mathcal{F} \rightarrow [0,1]$

$\vdots$

Ind. n $\rightarrow P_n : \mathcal{F} \rightarrow [0,1]$

$\mathcal{M}_p(\Theta, \mathcal{F})$: Conjunto de medidas de probabilidade sobre $(\Theta, \mathcal{F})$.

Definição: Um agrupamento de probabilidades ou opiniões ("opinion poo") é qualquer transformação de $(\mathcal{M}_p(\Theta, \mathcal{F}))^n$ em $\mathcal{M}_p(\Theta, \mathcal{F})$.

$$\mathbb{P}^{(G)} : (\mathcal{M}_p(\Theta, \mathcal{F})^n) \rightarrow \mathcal{M}_p(\Theta, \mathcal{F})$$
$$(\mathbb{P}_1, \dots, \mathbb{P}_n) \mapsto \mathbb{P}^{(G)}(\mathbb{P}_1, \dots, \mathbb{P}_n)$$

Exemplos:

1. $\mathbb{P}^{(G)}(\mathbb{P}_1, \dots, \mathbb{P}_n) = \mathbb{P}_1, \ \forall (\mathbb{P}_1, \dots, \mathbb{P}_n) \in (\mathcal{M}_p(\Theta, \mathcal{F}))^n$.

2. Agrupamento linear

$$\mathbb{P}^{(G)}(\mathbb{P}_1, \dots, \mathbb{P}_n) = \sum_{i=1}^{n} \alpha_i \mathbb{P}_i \quad , \quad \sum_{i=1}^{n} \alpha_i = 1.$$

3. Agrupamento Logarítmico

$\mathbb{P}_i \ll \mu$, $\mu : \mathcal{F} \to \mathbb{R}_+$, $i = 1, \dots, n$, e tais que

$$f_i = \frac{d\mathbb{P}_i}{d\mu}, \quad i = 1, \dots, n$$

Nessas condições,

$$f^{(G)}(\theta) = \frac{\prod_{i=1}^{n} (f_i(\theta))^{\alpha_i}}{\int_{\Theta} \prod_{i=1}^{n} (f_i(\theta))^{\alpha_i} \, d\mu(\theta) \; 0 < \infty} \quad \text{e,}$$

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

$$\mathbb{P}^{(G)}(A) = \int_A f^{(G)}(\theta) \, d\mu(\theta).$$

Suponha que, para cada $A \in \mathcal{F}$, $i \in \{1, 2, \dots, n\}$,

$$\mathbb{P}_i(A) = \int_A \lambda_i e^{-\lambda_i \theta} \, d\theta, \quad \lambda_i > 0.$$

No agrupamento linear, $\mathbb{P}^{(a)}$ possui densidade de probab. dada por

$$f^{(a)}(\theta) = \sum_{i=1}^{n} \alpha_i \lambda_i e^{-\lambda_i \theta} \, \mathbb{I}_{\mathbb{R}_+}(\theta)$$

No agrupamento logarítimo

$$f^{(G)}(\theta) \propto \prod_{i=1}^{n} (f_i(\theta))^{\alpha_i} = \prod_{i=1}^{n} (\lambda_i e^{-\lambda_i \theta} \, \mathbb{I}_{\mathbb{R}_+}(\theta))^{\alpha_i}$$

$$\propto e^{-(\sum_{i=1}^{n} \alpha_i \lambda_i)\theta} \, \mathbb{I}_{\mathbb{R}_+}(\theta)$$

$$\Rightarrow f^{(G)}(\theta) = \left(\sum_{i=1}^{n} \alpha_i \lambda_i\right) e^{-(\sum_{i=1}^{n} \alpha_i \lambda_i)\theta} \, \mathbb{I}_{\mathbb{R}_+}(\theta)$$

Propriedades Desejáveis Para Agrupamentos

Unanimidade

1. $P_1(\cdot) = P_2(\cdot) = \dots = P_n(\cdot) \Rightarrow P^{(G)}(\cdot) = P(\cdot)$
$= P(\cdot)$

$$P^{(G)}(A) = \sum_{i=1}^{n} \alpha_i P_i(A) = \sum_{i=1}^{n} \alpha_i P(A) = P(A) \rightarrow \text{Agrupamento linear satisf. unanimidade}$$

No agrupamento logarítmico, há unanimidade se $\sum_{i=1}^{n} \alpha_i = 1$.

2. Bayesianidade Externa

"Agrupar e, então, revisar incerteza é equivalente a revisar incerteza e, então, agrupar opniões"

Ex: $n=2$, $\Theta = [0,1]$.

$P_1 \sim Beta(2,1)$ $\quad \alpha_1 = \alpha_2 = 1/2$ (Agr. linear)
$P_2 \sim Beta(1,2)$

Suponhamos que os indivíduos concordam que X, dado Θ, é cond. $Bern(\Theta)$.

AGRUPANDO PRIMEIRO.

$$f^{(G)}(\theta) = \frac{1}{2} 2\theta \, \mathbb{I}_{[0,1]}(\theta) + \frac{1}{2} 2(1-\theta) \, \mathbb{I}_{[0,1]}(\theta) = \mathbb{I}_{[0,1]}(\theta)$$

Atualizando $f^{(G)}$ após observar $X=1$, temos

$$f^{(G)}(\theta / X=1) = 2\theta \, \mathbb{I}_{[0,1]}(\theta)$$

Por outro lado:

ATUALIZANDO AS INCERTEZAS DOS 2 INDIVÍDUOS:

$$f_1(\theta / X=1) = 3\theta^2 \, \mathbb{I}_{[0,1]}(\theta)$$
$$f_2(\theta / X=1) = 6\theta(1-\theta) \, \mathbb{I}_{[0,1]}(\theta)$$

Agrupando as opniões revisadas:

$$f^{(G)}(\theta / X=1) = \frac{1}{2} 3\theta^2 + \frac{1}{2} 6\theta(1-\theta) \, \mathbb{I}_{[0,1]}(\theta) = \left(3\theta - \frac{3\theta^2}{2}\right) \mathbb{I}_{[0,1]}(\theta)$$

Assim, o agrupamento linear não atende, em geral, Bayesionidade Externa. Tal propriedade é satisfeita pelo agrupamento logarítmico $f_1, f_2, \ldots, f_n$.

AGRUPANDO PRIMEIRO.

$$f^{(G)}(\theta) \propto \prod_{i=1}^{n} (f_i(\theta))^{\alpha_i}$$

Para cada $x \in \mathcal{X}$ e $\theta \in \Theta$, $f(x/\theta)$ é comum aos indivíduos e também é "verossimilhança" para o grupo. Atualizando $f^{(G)}(\cdot)$ ao observar $X=x$, temos

$$f^{(G)}(\theta / x) \propto f(x/\theta) f^{(G)}(\theta) = f(x/\theta) \prod_{i=1}^{n} (f_i(\theta))^{\alpha_i}$$

Por outro lado, cada indivíduo atualiza sua incerteza sobre θ primeiro.

$$f_i(\theta / x) \propto f(x/\theta) f_i(\theta)$$

$$\Rightarrow f_i(\theta / x) = k_i \, f(x/\theta) f_i(\theta), \quad k_i > 0, \quad i=1,\ldots,n$$

Em seguida, agrupando as opniões atualizadas, obtemos

$$f^{(G)}(\theta / x) \propto \prod_{i=1}^{n} (f_i(\theta/x))^{\alpha_i} = \prod_{i=1}^{n} (k_i f_i(x/\theta) f_i(\theta))^{\alpha_i}$$
$$\propto (f(x/\theta))^{\sum_{i=1}^{n} \alpha_i} \cdot \prod_{i=1}^{n} (f_i(\theta))^{\alpha_i}$$

Se $\sum_{i=1}^{n} \alpha_i = 1$, o agrupamento logarítmico atende Bayesianidade Externa.

3. PRESERVAÇÃO DE INDEPENDÊNCIA

Sejam $A, B \in \mathcal{F}$,

$$\mathbb{P}_i(A \cap B) = \mathbb{P}_i(A)\mathbb{P}_i(B),\ \forall i = 1, \ldots, n \Rightarrow \mathbb{P}^{(G)}(A \cap B) = \mathbb{P}^{(G)}(A)\,\mathbb{P}^{(G)}(B).$$

4. UNANIMIDADE FRACA

Se $A \in \mathcal{F}$ é tal que

$$\mathbb{P}_i(A) = \gamma,\ \forall i = 1, \ldots, n \Rightarrow \mathbb{P}^{(G)}(A) = \gamma.$$

CAMPOS DE COINCIDÊNCIA DE OPINIÕES DE DE FINETTI

$\Theta \subseteq \mathbb{R}$

Ind 1 $\rightarrow F_1$

Ind 2 $\rightarrow F_2$

} Por ora, foco em 2 indivíduos. F_i são contínuas!, $i = 1, 2, \ldots$

Ind. n $\rightarrow F_n$

$\exists\, (a, b) \subseteq \mathbb{R}$ t.q.

$$\mathbb{P}_1((a,b]) = F_1(b) - F_1(a) = F_2(b) - F_2(a) = \mathbb{P}_2((a,b])$$

Ideia da prova:

$m_1 = m_2$ ✓ (m_1 = mediana sob. P_1)

$m_1 \neq m_2$ $\mathcal{C} = \{(c,d) \subseteq \mathbb{R} : \mathbb{P}_1((c,d]) = 1/2\}$

$(-\infty, m_1)$, (1º quartil $\mathbb{P}_1$, 3º quartil $\mathbb{P}_1$), $(m_1, \infty) \in \mathcal{C}$

$\mathbb{P}_2((-\infty, m_1]) < 1/2$ ... $\mathbb{P}_2((m_1, \infty)) > 1/2$

cond. da transf. $\Rightarrow \exists\, (a,b)$ t.q. $\mathbb{P}_2((a,b]) = 1/2$

$\exists\ I_1 = (a,b]$ t.q. $P_1((a,b]) = P_2((a,b]) = 1/2$

$$\bar{F}_1(t) = \begin{cases} 2F_1(t), & t < a \\ 2\left[F_1(t+b-a) - 1/2\right], & t > a \end{cases}$$

$\exists\ (a',b') \subseteq \mathbb{R}$, $\bar{F}_1(b') - \bar{F}_1(a') = \bar{F}_2(b') - \bar{F}_2(a') = 1/2$

$b' < a$

$\bar{F}_1(b') - \bar{F}_1(a') = 2F_1(b') - 2F_1(a') = 1/2 \Rightarrow F_1(b') - F_1(a') = 1/4$
$\bar{F}_2(b') - \bar{F}_2(a') = 2F_2(b') - 2F_2(a') = 1/2 \Rightarrow F_2(b') - F_2(a') = 1/4.$

Distribuição . Consenso de M. H. DeGroot.

Ind. 1 $\rightarrow P_1^{(0)}$
Ind. 2 $\rightarrow P_2^{(0)}$
$\vdots$
Ind. n $\rightarrow P_n^{(0)}$

w_{ij}: o peso que o ind. i dá a opinião do ind. j sobre Θ.
Para todo i,j, $w_{ij} > 0$ e $\sum_{j=1}^{n} w_{ij} = 1$.

$n = 1$.

$$P_1^{(1)} = \sum_{j=1}^{n} w_{1j} P_j^{(0)}$$
$$P_2^{(1)} = \sum_{j=1}^{n} w_{2j} P_j^{(0)}$$
$$\vdots$$
$$P_n^{(1)} = \sum_{j=1}^{n} w_{nj} P_j^{(0)}$$

$n = 2$

$$P_1^{(2)} = \sum_{j=1}^{n} w_{1j} P_j^{(1)}, \ \ldots, \ P_n^{(2)} = \sum_{j=1}^{n} w_{nj} P_j^{(1)}$$

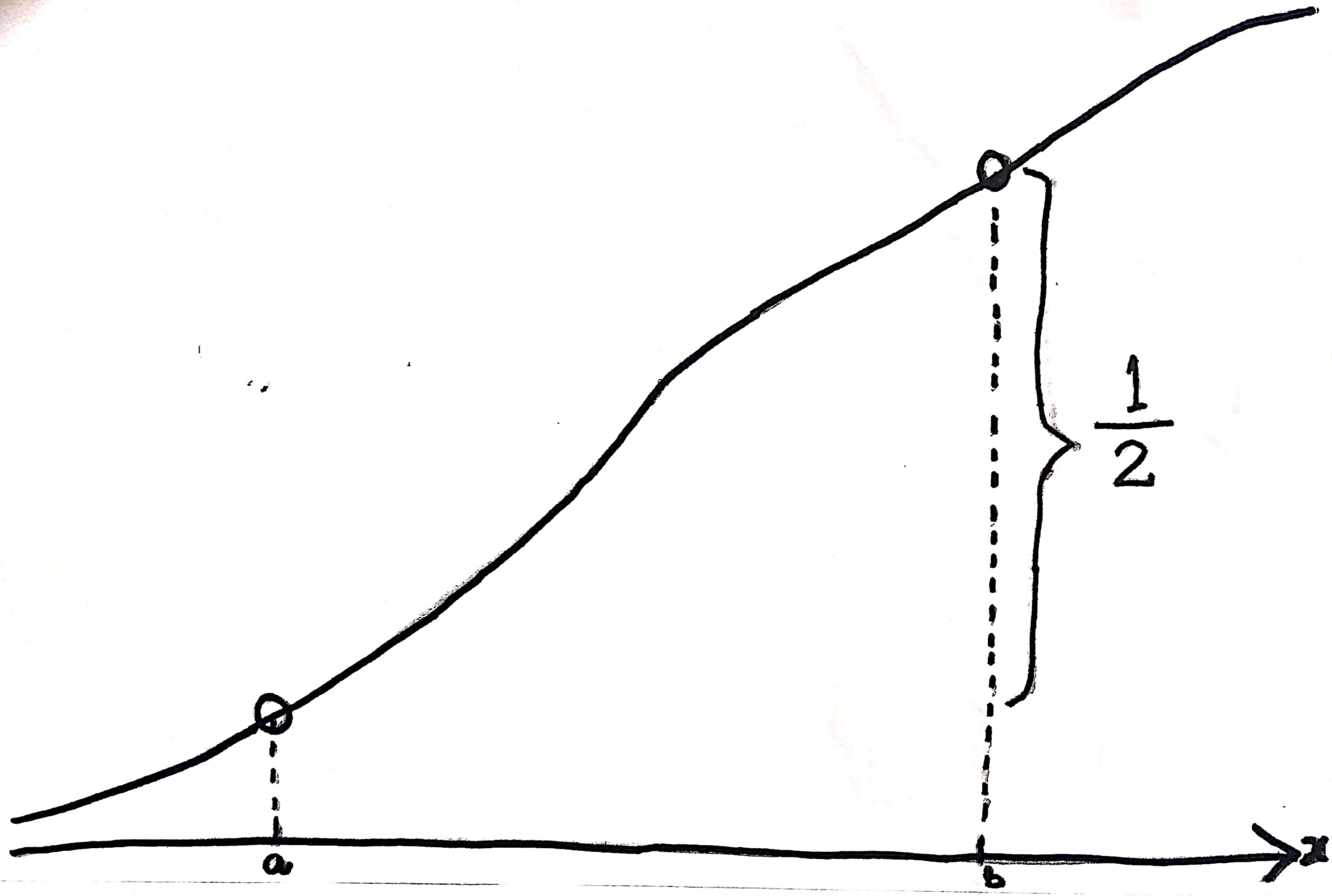
$\sum_{j=0}^{n} w_j^i P_j$

$\rightarrow \sum_{j=0}^{n} w_j^i P_j$ W^i recorrente e aperiodica.

$\vdots$

$\sum_{j=1}^{n} w_j^i P_j$

23/11/15



Teoria da Decisão

Aula . .   25/11/2015

Algoritmos ABC em Environmental Stress Screening

Algoritmo ABC: gerar pontos de uma dist. a posteriori.

Approximate Bayesian Computation

. ABC

. Alg. da Rejeição 1

Conheço a priori

Conheço a dist. dos dados

Prob. na posteriori   938,91

Gero

ESS : peças boas ou ruins

Lote N peças

Prop. p de peças de qualidade ruim

Taxa de falha das peças de qualidade boa : λ_b

" " " " " " " ruim : λ_r

Proc. de Estresse

1. Am. dos N peças
2. Submissão ao estresse
3. Verif. das que sobrev.
4. Veri do qual. das que falharam

Função cus

C_1: custo de estressar ruim e sobrev

cu " " " " e falhar

C_2: " " " boa e falhar

C_3: Custo operar de estressar qq e sobrev.

1^a Avaliação 05/10/2015

MAE 5753 - 2º semestre de 2015

1) Prove o seguinte resultado admitindo as suposições SP1 a SP4: se $(A_n)_{n\geq 1}$ é uma sequência decrescente de eventos e B é um evento tal que $\cap A_n \prec B$, então $B \precsim A_i$ apenas para um número finito de valores de i. Enuncie e prove o resultado análogo para uma sequência crescente de eventos $(A_n)_{n\geq 1}$.

2) Sejam p e m números fixados tais que $0 < p < 1$ e $m > 0$, e considere a seguinte situação: uma pessoa receberá uma quantia de m dólares que deverá alocar entre a ocorrência de um evento A, de probabilidade p, e seu complementar A^C. Sua recompensa será a quantia que foi alocada em A ou A^C, dependendo de qual evento efetivamente ocorrer. Em outras palavras, para todos os valores de x no intervalo [0,m], a pessoa poderá escolher uma entre todas as loterias do tipo $< x, p \,;\, m - x, 1 - p >$, isto é, uma loteria que premia com a recompensa x com probabilidade p e com a recompensa $m - x$ com probabilidade 1- p. Considerando todos os possíveis de valores de p e m, encontre a melhor alocação dos m dólares considerando que a função de utilidade U da pessoa para o ganho monetário é definida no intervalo [0,m] por

$$U(r) = 2(r/m_0) - (r/m_0)^2 \text{, onde } m_0 \geq m.$$

3) Prove o seguinte resultado a partir das suposições SP1 a SP4 e da suposição CP:

seja D um evento qualquer. Então, $D \sim \emptyset$ se, e somente se, $(A \mid D) \sim (B \mid D)$ para todos os eventos A e B.

4) Suponha que um indivíduo irá vender sorvete num jogo de futebol e deve decidir de antemão quanto deve encomendar dessa produto. Suponha que seu ganho é de m reais por cada litro de sorvete vendido, enquanto que sua perda é de c reais por cada litro encomendado mas não vendido. Admitindo que a demanda por esse tipo de sorvete (em litros) é uma variável aleatória absolutamente contínua Θ com f.d.p f e f.d. F, mostre que o lucro esperado desse indivíduo será maximizado encomendando uma quantidade α tal que $F(\alpha) = m/(m + c)$.

5) Suponha que o intervalo de tempo que um homem deve esperar a cada manhã para apanhar o ônibus que o leva para o trabalho é distribuído uniformemente no intervalo $(0,\Theta)$, onde o valor de Θ é desconhecido. Suponha, a priori, que Θ seja distribuído segundo o modelo Pareto de parâmetros $a = 1$ e $b > 0$. Por quantas manhãs esse homem deve registar o tempo de espera pelo ônibus que o leva ao trabalho para que ele seja capaz de especificar um intervalo de comprimento 0.01 unidades tal que a probabilidade de que $\log(\Theta)$ esteja nesse intervalo seja pelo menos 0,95 ? 299

2ª Avaliação 30/11/2015

MAE 5753 - 2º semestre de 2015

1) Suponha que $X_1,\dots,X_n$, dado $\boldsymbol{\theta}$, sejam condicionalmente independentes e identicamente distribuídas segundo o modelo Poisson de média $\boldsymbol{\theta}$. Suponha, a priori, que $\boldsymbol{\theta}$ seja distribuído segundo o modelo Gama de parâmetros **a** e **b**, **a**,**b**>0. Para estimar $\boldsymbol{\theta}$, considere a seguinte função de perda:

$$L(d,\theta) = (d-\theta)^2 / \theta = (d-\theta)^2\theta^{-1} .$$

Obtenha o estimador de Bayes para $\boldsymbol{\theta}$ nesse caso e mostre que o risco de Bayes é dado por 1/(b + n).

2) Suponha que $X_1,\dots,X_n$, dado $\boldsymbol{\theta}$, sejam condicionalmente independentes e identicamente distribuídas segundo o modelo Normal de média $\boldsymbol{\theta}$ e variância 9. Considere, a priori, que $P(\boldsymbol{\theta} = -1) = P(\boldsymbol{\theta} = 1) = 1/2$. Para testar as hipóteses $\mathbf{H_0:\theta = -1}$ versus $\mathbf{H_1:\theta = 1}$, considere a função de perda 0 – 1 – 1. Determine o teste de Bayes nesse caso.

3) Suponha que um parâmetro de interesse, $\boldsymbol{\theta}$, é distribuído, a priori, segundo $P(\boldsymbol{\theta} = \boldsymbol{\theta}_1) = \xi$ e $P(\boldsymbol{\theta} = \boldsymbol{\theta}_2) = 1 - \xi$, $0 < \xi < 1$. Para cada ponto amostral $\mathbf{x} \in \mathcal{X}$, seja $\boldsymbol{\xi(x)}$ a probabilidade a posteriori de $\{\boldsymbol{\theta} = \boldsymbol{\theta}_1\}$ dado $\mathbf{x}$. Mostre que $\mathbf{E[\xi(X)] = \xi}$.

$(\mathfrak{X}, \mathcal{B}(\mathfrak{X}), \mathbb{P})$

$x \in \mathfrak{X}$

$(\mathbb{P}(x$

$\theta: D \quad \mathbb{P}(\theta \dots$

4) Suponha que o parâmetro real θ deve ser estimado segundo a seguinte função de perda: $L(d,\theta) = k_1(\theta - d)$, se $d < \theta$, e $L(d,\theta) = k_2(d - \theta)$, se $d \geq \theta$, $k_1,k_2>0$. Suponha que $E[\ |\theta|\ ] <\infty$. Mostre que a decisão de Bayes, d^*, deve satisfazer

$$P(\theta \leq d^*) \geq k_1/(k_1 + k_2)^{-1} \quad \text{e} \quad P(\theta \geq d^*) \geq k_2/(k_1 + k_2)^{-1}.$$

5) Suponha que em um problema de decisão arbitrário, a decisão d^* é uma decisão de Bayes contra duas distribuições distintas P_1 e P_2 para o parâmetro θ. Mostre que para qualquer número α □ $(0,1)$, d^* é também decisão de Bayes contra a distribuição $\alpha P_1 + (1-\alpha) P_2$ para θ.

MAE 5753 - 2º semestre /2015

25/08

Pasta 4

Nº cópias 4

LISTA DE EXERCÍCIOS 1

Exercícios extraídos do livro "Optimal Statistical Decisions" de M. H DeGroot.

Exercícios do Capítulo 6 (página 83)

2, 3a, 3b, 3d, 4, 5, 6a, 6c

Exercícios do Capítulo 9 (páginas 186, 187 e 188)

21, 23, 30, 32, 34

this inequality is equivalent to the relation $AD \precsim BD$. These remarks naturally lead to the next assumption.

Assumption *CP* *For any three events A, B, and D, $(A|D) \precsim (B|D)$ if, and only if, $AD \precsim BD$.*

As we remarked above, when Assumptions SP_1 to SP_5 are made and there exists a probability distribution P which can be uniquely specified, it is sufficient for most purposes to apply Assumption CP only to events D such that $P(D) > 0$. However, in order to keep the assumptions independent of each other, we have imposed the slightly stronger requirement that Assumption CP applies to all events D. Furthermore, it now follows from Assumptions SP_1 and CP that for any three events A, B, and D, either $(A|D) \precsim (B|D)$ or $(B|D) \precsim (A|D)$. It is possible, of course, that both relations will be correct. In fact, both relations will always be correct when D is an event such that $P(D) = 0$ (see Exercise 6*a*).

By combining Assumption CP with Theorem 4 of Sec. 6.5, we obtain the following general theorem.

Theorem 1 *If the relation $\precsim$ satisfies Assumptions SP_1 to SP_5 and Assumption CP, then the function P defined by Eq. (3) of Sec. 6.4 is the unique probability distribution which has the following property: For any three events A, B, and D such that $P(D) > 0$, $(A|D) \precsim (B|D)$ if, and only if, $P(A|D) \precsim P(B|D)$.*

Further Remarks and References

Luce and Suppes (1965) give a good review of subjective probability, including the relatively small amount of work that has been done on the experimental measurement of subjective probabilities. An interesting and controversial group of papers dealing with the existence and measurement of subjective probabilities are those by Ellsberg (1961), Fellner (1961, 1963), Raiffa (1961), Brewer (1963), and Roberts (1963). The book by Fellner (1965) is also of interest here.

Exercises 7 to 9 at the end of this chapter require the actual evaluation of the reader's subjective probabilities of some specific events.

EXERCISES

1. Prove that the following results follow from Assumptions SP_1 to SP_3:

(*a*) If A, B, and D are any three events such that $A \precsim B$ and $B \prec D$, then $A \prec D$.

(*b*) If A and B are any two events, then $A \precsim B$ if, and only if, $A^c \succsim B^c$.

(*c*) If A and B are events such that $A \subset B$, then $A \precsim B$.

(*d*) If $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$ are events such that $B_iB_j = \emptyset$ for $i \neq j$ and $A_i \precsim B_i$ for $i = 1, \ldots, n$, then $\bigcup_{i=1}^n A_i \precsim \bigcup_{i=1}^n B_i$. If, in addition, $A_i \prec B_i$ for at least one value of i $(i = 1, \ldots, n)$, then $\bigcup_{i=1}^n A_i \prec \bigcup_{i=1}^n B_i$. (This is an extension of Theorem 2 of Sec. 6.2 since it is not assumed here that $A_1, \ldots, A_n$ are disjoint.) Furthermore, the statements remain true if the condition that $B_iB_j = \emptyset$ for $i \neq j$ is replaced by the weaker condition that $B_iB_j \sim \emptyset$ for $i \neq j$.

2. Suppose that Assumptions SP_1 to SP_3 are made. Let $A_1, \ldots, A_m$ be disjoint events, and let $B_1, \ldots, B_n$ also be disjoint events such that $\bigcup_{i=1}^m A_i = \bigcup_{i=1}^n B_i = S$, $A_1 \precsim \cdots \precsim A_m$ and $B_1 \precsim \cdots \precsim B_n$, and $m \leq n$. Prove that $B_1 \precsim A_m$. Prove also that if $m < n$, then $B_1 \prec A_m$.

3. Prove that the following results follow from Assumptions SP_1 to SP_4:

(*a*) If $A_1 \supset A_2 \supset \cdots$ is a decreasing sequence of events and if B is an event such that $\bigcap_{i=1}^\infty A_i \prec B$, then $A_i \succsim B$ for only a finite number of values of i $(i = 1, 2, \ldots)$. State and prove the analogous result for an increasing sequence of events $A_1 \subset A_2 \subset \cdots$.

(*b*) If $A_1 \subset A_2 \subset \cdots$ is an increasing sequence of events and $B_1 \supset B_2 \supset \cdots$ is a decreasing sequence of events such that $A_i \precsim B_i$ for $i = 1, 2, \ldots$, then $\bigcup_{i=1}^\infty A_i \precsim \bigcap_{i=1}^\infty B_i$.

(*c*) If $A_1 \supset A_2 \supset \cdots$ is a decreasing sequence of events and $B_1 \supset B_2 \supset \cdots$ is also a decreasing sequence of events such that $A_i \precsim B_i$ for $i = 1, 2, \ldots$, then $\bigcap_{i=1}^\infty A_i \precsim \bigcap_{i=1}^\infty B_i$. Furthermore, if $A_i \sim B_i$ for $i = 1, 2, \ldots$, then $\bigcap_{i=1}^\infty A_i \sim \bigcap_{i=1}^\infty B_i$. State and prove the analogous results for increasing sequences of events $A_1 \subset A_2 \subset \cdots$ and $B_1 \subset B_2 \subset \cdots$.

(*d*) If $A_1, A_2, \ldots$ and $B_1, B_2, \ldots$ are sequences of events such that $B_iB_j \sim \emptyset$ for $i \neq j$ and $A_i \precsim B_i$ for $i = 1, 2, \ldots$, then $\bigcup_{i=1}^\infty A_i \precsim \bigcup_{i=1}^\infty B_i$. If, in addition, $A_i \prec B_i$ for at least one value of i $(i = 1, 2, \ldots)$, then $\bigcup_{i=1}^\infty A_i \prec \bigcup_{i=1}^\infty B_i$. (Note that this is an extension of Theorem 6 of Sec. 6.2, since it is not assumed here that the events $A_1, A_2, \ldots$ are disjoint and the condition that $B_1, B_2, \ldots$ are disjoint is replaced by the weaker condition that $B_iB_j \sim \emptyset$ for $i \neq j$.)

(*e*) If $A_1, A_2, \ldots$ is a sequence of disjoint events and B is an event such that $B \succ \emptyset$, then $A_i \succsim B$ for only a finite number of values of i $(i = 1, 2, \ldots)$.

4. Let P^* be a given probability distribution on the set S of positive integers $\{1, 2, \ldots\}$ such that each integer in S is assigned a positive probability. Let $S_1 = \{1, 3, 5, \ldots\}$, and let $S_2 = \{2, 4, 6, \ldots\}$. Hence, every subset A of S can be expressed in the form $A = (AS_1) \cup (AS_2)$. Suppose that a relation $\precsim$ is defined between subsets of S as follows: If A and B are any two subsets of S, then $A \precsim B$ if either $P^*(AS_1) < P^*(BS_1)$ or $P^*(AS_1) = P^*(BS_1)$ and $P^*(AS_2) \leq P^*(BS_2)$. Show that the relation $\precsim$ satisfies Assumptions SP_1 to SP_3 but not Assumption SP_4.

5. Consider the problem described in Exercise 4, but suppose now that the relation $\precsim$ between subsets of S is defined as follows: If A and B are any two subsets of S, then $A \precsim B$ if either $P^*(A) < P^*(B)$ or $P^*(A) = P^*(B)$ and $P^*(AS_1) \leq P^*(BS_1)$. Show that the relation $\precsim$ satisfies Assumptions SP_1 to SP_3 but not Assumption SP_4.

6. Prove that the following results can be derived from Assumptions SP_1 to SP_4 and Assumption CP:

(*a*) If D is any event, then $D \sim \emptyset$ if, and only if, $(A|D) \sim (B|D)$ for all events A and B.

(*b*) If A and D are any events, then $(\emptyset|D) \precsim (A|D) \precsim (D|D)$.

(*c*) Let A, B, D, and E be events such that $ADE \sim BDE \sim \emptyset$. Then $(A|E) \precsim (B|E)$ if, and only if, $[(A \cup D)|E] \precsim [(B \cup D)|E]$.

15. Show that for any given positive numbers μ and σ^2, there is a unique gamma distribution which has mean μ and variance σ^2.

16. Suppose that when magnetic recording tape is manufactured by a certain process, the mean number W of defects on a 1,200-ft roll of tape is unknown, and suppose that the prior distribution of W is a gamma distribution whose mean is 2 and whose variance is 1. Suppose also that the number of defects on any roll of tape when $W = w$ has a Poisson distribution with mean w. Suppose further that after a random sample of rolls of tape has been taken and the number of defects on each roll has been counted, the mean of the posterior distribution of W is 1.6 and the variance is 0.16. Show that eight rolls of tape were included in the random sample and that the average number of defects per roll in the sample was 1.5.

17. Prove Theorem 2 of Sec. 9.4.

18. An unknown proportion W of the items produced by a certain machine are defective. Suppose that the prior distribution of W is a beta distribution with parameters $\alpha = 1$ and $\beta = 99$. Suppose also that items produced by the machine are selected at random and observed one at a time until exactly five defective items have been found. If, when sampling terminates, the mean of the posterior distribution of W is 0.02, show that 195 nondefective items were observed during the sampling process.

19. Suppose that in a large population of voters, the proportion W who belong to the Liberal Party is unknown, and suppose that the prior distribution of W is a beta distribution with parameters $\alpha = 1$ and $\beta = 10$.

(a) If, in a random sample of 1,000 voters, it is found that 123 belong to the Liberal Party, what is the posterior distribution of W?

(b) Suppose that instead of taking a random sample as in part a, voters are selected one at a time until exactly 123 have been found who belong to the Liberal Party. Suppose that a total of 1,000 voters had to be selected in order to accomplish this. What is the posterior distribution of W?

20. Prove Theorem 3 of Sec. 9.4.

21. The length of life of a lamp manufactured by a certain process has an exponential distribution with an unknown value of the parameter W. Suppose that the prior distribution of W is a gamma distribution for which the coefficient of variation is 0.5. A random sample of lamps is to be tested, and the length of life of each of the lamps is to be noted. If the coefficient of variation of the posterior distribution of W must be reduced to the value 0.1, show that 96 lamps should be tested.

22. Extend Theorem 3 of Sec. 9.4 so that it covers the case of a random sample from a gamma distribution with parameters α and W, where the value of α is specified ($\alpha > 0$) and the value of W is unknown.

23. Consider a normal distribution for which the value of the mean W is unknown and the variance is 4, and suppose that the prior distribution of W is a normal distribution whose variance is 9. How large a random sample must be taken from the given normal distribution in order to be able to specify an interval having a length of 1 unit such that the probability that W lies in this interval is at least 0.95? (*Answer:* $n = 62$.)

24. Prove Theorem 2 of Sec. 9.5.

25. Suppose that the value of the precision W of a normal distribution is unknown, and suppose that the distribution of W is a gamma distribution with parameters α and β. Let V denote the variance of the given normal distribution.

(a) Find the p.d.f. of V.

(b) Show that if $\alpha > 1$, $E(V) = \beta/(\alpha - 1)$.

(c) Show that if $\alpha > 2$, $\mathrm{Var}(V) = \beta^2/[(\alpha - 1)^2(\alpha - 2)]$.

26. Suppose that a random sample is to be taken from a normal distribution with a specified value of the mean and an unknown value of the precision W. Suppose also that the prior distribution of W is a gamma distribution and that the coefficient of variation of the posterior distribution of W must be reduced to the value 0.1. Show that this requirement will be satisfied, regardless of the value of the coefficient of variation of the prior distribution, if a sample of size $n = 200$ is taken.

27. Consider a normal distribution with an unknown value of the mean M and an unknown value of the precision R, and suppose that the prior joint distribution of M and R is as specified in Theorem 1 of Sec. 9.6. Find the conditional distribution of R when $M = m$.

28. Consider the conditions specified in Exercise 27. Suppose that the coefficient of variation of the prior distribution of R has the value 0.5. How large a random sample must be taken from the given normal distribution in order that the coefficient of variation of the posterior distribution of R will be reduced to the value 0.1? (*Answer:* $n = 192$.)

29. Consider the conditions specified in Exercise 27. Suppose that under the conditional posterior distribution of M when $R = 3$, the variance of M must be reduced to the value 0.01. Show that this requirement will be satisfied, regardless of the values of the parameters of the prior distribution, if a random sample of size $n = 34$ is taken from the given normal distribution.

30. The length of time for which a certain man must wait each morning for a bus taking him to work is uniformly distributed on the interval $(0, W)$, where the value of W is unknown and the prior distribution of W is a Pareto distribution with parameters $w_0 > 0$ and $\alpha = 1$. On how many mornings must the man observe his waiting time before he will be able to specify an interval having a length of 0.01 unit such that the probability that the unknown value of $\log W$ lies in this interval is at least 0.95? (*Answer:* $n = 299$.)

31. Consider the prior distribution of W specified in Exercise 30. On how many mornings must the man observe his waiting time in order that the coefficient of variation of the posterior distribution of W will be reduced to the value 0.01? (*Answer:* $n = 101$.)

32. Prove Theorem 2 of Sec. 9.7.

33. Consider a uniform distribution on the interval (W_1, W_2), where the values of W_1 and W_2 are unknown, and suppose that the prior joint distribution of W_1 and W_2 is a bilateral bivariate Pareto distribution with parameters $r_1 > 0$, $r_2 > 0$, and $\alpha = 2$. How large a random sample must be taken from the uniform distribution in order that the coefficient of variation of the posterior distribution of the random variable $W_2 - W_1$ will be reduced to the value 0.01? (*Answer:* $n = 140$.)

34. Suppose that a box contains N balls, of which an unknown number W are red and the rest are blue. Suppose also that the prior distribution of W is a hypergeometric distribution with parameters A, B, and N, where A and B are positive integers such that $A + B \geq N$.

(a) Now suppose that although the exact value of W is unknown, the statistician knows that $r \leq W \leq s$, where r and s are integers such that $0 < r \leq s < N$. Show that there are unique values of A and B such that, under the prior hypergeometric distribution,

$$\Pr\{r \leq W \leq s\} = 1, \qquad \Pr\{W = r\} > 0, \qquad \text{and} \qquad \Pr\{W = s\} > 0.$$

(b) Next, suppose that the statistician knows only that $0 \leq W \leq N$. Show that any prior hypergeometric distribution such that $A \geq N$ and $B \geq N$ will assign positive probability to each integer $0, 1, 2, \ldots, N$.

(c) Suppose that one ball is selected from the box at random. What is the probability, under the prior distribution, that it will be red?

(d) Suppose that n balls ($1 \leq n < N$) are selected at random from the box without replacement and that x of these balls are red. Show that the posterior distribution of the number of red balls among the $N - n$ balls remaining in the box is a hypergeometric distribution with parameters $A - x$, $B - (n - x)$, and $N - n$.

35. Suppose that there are k different types of items in a very large population and let W_i be the unknown proportion of the population that includes items of type i ($i = 1, 2, \ldots, k$). Suppose also that the prior distribution of $\mathbf{W} = (W_1, \ldots, W_k)'$ is a Dirichlet distribution with parametric vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)'$ such that $\alpha_1 + \cdots + \alpha_k = 6$. How large a random sample of items must be taken in order to be sure that no matter what the values of the individual components of the vector $\boldsymbol{\alpha}$ are and no matter what the observed outcomes are, the posterior variance of each proportion W_i ($i = 1, \ldots, k$) will be at most 0.005? (*Answer:* $n = 43$.)

36. Consider a bivariate normal distribution with an unknown mean vector $\mathbf{M} = (M_1, M_2)'$ and a precision matrix $\mathbf{r}$ which is known to be

$$\mathbf{r} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{4}{3} \end{bmatrix}.$$

Suppose that the prior distribution of $\mathbf{M}$ is a bivariate normal distribution for which the precision matrix is

$$\boldsymbol{\tau} = \begin{bmatrix} 1 & -1 \\ -1 & 6 \end{bmatrix}.$$

How large a random sample must be taken in order that the variance of the posterior distribution of the random variable $M_1 - M_2$ will be reduced to the value 0.01? (*Answer:* $n = 297$.)

37. Suppose that a random $k \times k$ symmetric, positive definite matrix $\mathbf{R}$ has a Wishart distribution with α degrees of freedom ($\alpha > k - 1$) and precision matrix $\boldsymbol{\tau}$. Show that the coefficient of variation of the determinant $|\mathbf{R}|$ is

$$\left[\frac{k(2\alpha - k + 3)}{(\alpha - k + 2)(\alpha - k + 1)}\right]^{\frac{1}{2}}.$$

Hint: See Exercise 14 of Chap. 5.

38. Consider a bivariate normal distribution with a specified mean vector and an unknown precision matrix $\mathbf{R}$. Suppose that the prior distribution of $\mathbf{R}$ is a Wishart distribution with 3 degrees of freedom and precision matrix $\boldsymbol{\tau}$. How large a random sample must be taken in order that the coefficient of variation of the posterior distribution of the determinant $|\mathbf{R}|$ will be reduced to the value 0.1? (*Answer:* $n = 399$.)

39. Consider again the conditions described in Exercise 38, and suppose that the elements of the 2×2 matrix $\mathbf{R}$ are

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}.$$

How large a random sample must be taken in order that the coefficient of variation of the posterior distribution of the random variable R_{11} will be reduced to the value 0.1? (*Answer:* $n = 197$.)

40. Prove that Eq. (3) of Sec. 9.11 is correct.

41. Consider a multivariate normal distribution with an unknown value of the mean vector $\mathbf{M}$ and an unknown value of the precision matrix $\mathbf{R}$, and suppose that the prior joint distribution of $\mathbf{M}$ and $\mathbf{R}$ is a multivariate normal-Wishart distribution as specified in Theorem 1 of Sec. 9.10. What is the conditional distribution of $\mathbf{R}$ when $\mathbf{M} = \mathbf{m}$?

42. Consider a bivariate normal distribution with an unknown value of the mean vector $\mathbf{M}$ and an unknown value of the precision matrix $\mathbf{R}$. Suppose that the prior joint distribution of $\mathbf{M}$ and $\mathbf{R}$ is a bivariate normal-Wishart distribution as specified in Theorem 1 of Sec. 9.10, and suppose that $\alpha = 3$ in this prior distribution. How large a random sample must be taken in order that the coefficient of variation of the posterior distribution of the determinant $|\mathbf{R}|$ will be reduced to the value 0.1? *Hint:* Note that the answer in this exercise must be the same as the answer in Exercise 38.

43. Consider the conditions specified in Exercise 42, and suppose again that $\alpha = 3$ in the prior joint distribution of $\mathbf{M}$ and $\mathbf{R}$. Suppose that a random sample is taken, and let $\boldsymbol{\mu}^*$ be the location vector and $\mathbf{T}^*$ the precision matrix of the posterior bivariate t distribution of the two-dimensional vector $\mathbf{M}$. Determine the size of a random sample that must be taken in order for the posterior distribution of $\mathbf{M}$ to satisfy the following equation:

$$\Pr[(\mathbf{M} - \boldsymbol{\mu}^*)'\mathbf{T}^*(\mathbf{M} - \boldsymbol{\mu}^*) \leq 10] \geq 0.95.$$

Hint: See expression (16) of Sec. 5.6. (*Answer:* $n = 5$.)

44. Prove Theorem 1 of Sec. 9.13.

45. Prove that if the joint distribution of the random vector $\mathbf{M}$ and the random variable W is a multivariate normal-gamma distribution as specified in Theorem 1 of Sec. 9.13, then the marginal distribution of $\mathbf{M}$ is a multivariate t distribution with 2α degrees of freedom, location vector $\boldsymbol{\mu}$, and precision matrix $(\alpha/\beta)\boldsymbol{\tau}$.

46. Consider five different normal distributions with unknown means $M_1, \ldots, M_5$ and with a common unknown precision W. Suppose that the prior joint distribution of $M_1, \ldots, M_5$ and W is as follows: For any given value $W = w$, the random variables $M_1, \ldots, M_5$ are independent and have the same normal distribution with mean 3 and precision $2w$. Furthermore, the marginal distribution of W is a gamma distribution with parameters $\alpha = 10$ and $\beta = 5$. Suppose also that a random sample of eight observations is taken from each of the five normal distributions. Let x_{ij} denote the value of the jth observation from the ith distribution, and for $i = 1, \ldots, 5$, let $\bar{x}_i$ denote the average of the eight observations from the ith distribution. Finally, let the values $\mu_1, \ldots, \mu_5$ and c be defined as follows:

$$\mu_i = \frac{3 + 4\bar{x}_i}{5} \qquad i = 1, \ldots, 5,$$

and

$$c = \frac{2.37}{600}\left[50 + 5\sum_{i=1}^{5}\sum_{j=1}^{8}(x_{ij} - \bar{x}_i)^2 + 8\sum_{i=1}^{5}(\bar{x}_i - 3)^2\right].$$

Show that under the posterior joint distribution of $M_1, \ldots, M_5$,

$$\Pr\left[\sum_{i=1}^{5}(M_i - \mu_i)^2 \leq c\right] = 0.95.$$

Hint: See expression (16) of Sec. 5.6.

47. Suppose that $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is a random sample of symmetric $k \times k$ matrices from a Wishart distribution with m degrees of freedom ($m > k - 1$) and an unknown value of the precision matrix $\mathbf{R}$. Suppose also that the prior distribution of $\mathbf{R}$ is a Wishart distribution with α degrees of freedom and precision matrix $\boldsymbol{\tau}$ such that $\alpha > k - 1$ and $\boldsymbol{\tau}$ is a $k \times k$ positive definite matrix. Show that the posterior distribution of $\mathbf{R}$ when $\mathbf{X}_i = \mathbf{x}_i$ ($i = 1, \ldots, n$) is a Wishart distribution with $\alpha + mn$ degrees of freedom and precision matrix $\boldsymbol{\tau} + \sum_{i=1}^{n}\mathbf{x}_i$.

2ª Lista de Exercícios

14/09

MAE 5753 - 2º Semestre/2015

Pasta 4

Nº cópias 4

Exercícios extraídos do livro "Optimal Statistical Decisions", de M.H. DeGroot

Exercícios 2, 3 e 4, página 115.

Exercício 7, página 116.

Exercícios 8(e) e 9, página 117.

Exercício 13, página 118.

tributions P_n and Q_n are bounded from below, they are subject to the conclusions of Lemma 3.

Now suppose that $P <^* Q$. By Assumption U_6, there must exist a positive integer n_0 such that $P_{n_0} <^* Q$. Also, by Lemma 2, $Q \lesssim^* Q_{n_0}$. Therefore, by Assumption U_2 and Lemma 2, there exists a number α $(0 < \alpha < 1)$ such that for $n = 1, 2, \ldots,$

$$P_{n_0} <^* \alpha Q_{n_0} + (1 - \alpha)P_{n_0} \lesssim^* Q \lesssim^* Q_n. \tag{15}$$

Hence, by Lemma 3, for $n = 1, 2, \ldots,$

$$c_{n_0} < \alpha d_{n_0} + (1 - \alpha)c_{n_0} \le d_n. \tag{16}$$

It now follows from Lemmas 1 and 2 that

$$E(U|P) \le c_{n_0} < \alpha d_{n_0} + (1 - \alpha)c_{n_0} \le E(U|Q). \tag{17}$$

The relation (17) establishes that $E(U|P) < E(U|Q)$.

Conversely, suppose that $E(U|P) < E(U|Q)$. Then there is a positive integer n_0 such that $c_{n_0} < E(U|Q) \le d_{n_0}$. Hence, there must exist a number α $(0 < \alpha < 1)$ such that for $n = 1, 2, \ldots,$

$$c_{n_0} < \alpha d_{n_0} + (1 - \alpha)c_{n_0} < E(U|Q) \le d_n. \tag{18}$$

It follows from Lemma 3 that for $n = 1, 2, \ldots,$

$$P_{n_0} <^* \alpha Q_{n_0} + (1 - \alpha)P_{n_0} \lesssim^* Q_n. \tag{19}$$

By Assumption U_6 and Lemma 2, we can now obtain the following relation:

$$P \lesssim^* P_{n_0} <^* \alpha Q_{n_0} + (1 - \alpha)P_{n_0} \lesssim^* Q. \tag{20}$$

Hence, $P <^* Q$.■

It was established in Theorem 2 of Sec 7.9 that the function U which has been constructed in this chapter and increasing linear transformations of U are the only functions which have the property described in Theorem 1.

Further Remarks and References

It should be remarked that no conditions have been imposed on the distributions in the class $\mathcal{P}$ which do not belong to $\mathcal{P}_E$. For some distributions $P \in \mathcal{P}$, the value of $E(U|P)$ can be regarded as $+\infty$. For such a distribution P, it is natural to assume that $P \gtrsim^* Q$ for any distribution $Q \in \mathcal{P}_E$, but it is not obvious how two such distributions should be compared with each other. Similar remarks apply to those distributions $P \in \mathcal{P}$ such that the value of $E(U|P)$ can be regarded as $-\infty$. However, there may still exist distributions $P \in \mathcal{P}$ such that $E(U|P)$ does not exist, either as a finite or as an infinite number. Although it could be assumed that the relation $\lesssim^*$ specifies a complete ordering of all distributions in $\mathcal{P}$, and further assumptions could be introduced which would assign a position in this ordering to each distribution outside of $\mathcal{P}_E$, it does not appear worthwhile to do so, and this subject will not be explored further. It should be emphasized that when the function U is bounded, then the classes $\mathcal{P}$, $\mathcal{P}_E$, and $\mathcal{P}_B$ are identical and the above questions do not arise.

The first axiomatic development of utility was given by Von Neumann and Morgenstern (1947). Other developments are due to Marschak (1950), Herstein and Milnor (1953), and Debreu (1960). Derivations are also given in the texts of Blackwell and Girshick (1954) and Chernoff and Moses (1959). Other interesting discussions and developments are given by Fishburn (1964, 1967a), Luce and Raiffa (1957), Luce and Suppes (1965), Pratt, Raiffa, and Schlaifer (1964), and Savage (1954). An indication of research in the theory of utility which emphasizes various aspects that have not been considered here is given in several of the papers in the collection edited by Thrall, Coombs, and Davis (1954) and in the papers of Aumann (1964), Koopmans, Diamond, and Williamson (1964), and Radner (1964).

Fishburn (1968) has published a bibliography on utility theory containing 315 items.

EXERCISES

1. Consider the possibility of your winning or losing any amount of money up to \$1,000. By using a technique of the type described in Sec. 7.5, determine several different points on your utility function and sketch that function as a curve over the interval from $-1{,}000$ to $+1{,}000$.

2. Consider an experiment in which a person first specifies a number of dollars x and then observes the value of a random variable Y. Suppose that if $Y \ge x$, he receives Y dollars as his reward. If $Y < x$, he receives a random reward X having some given probability distribution. It is assumed that the random variables X and Y are independent. Show that in order to maximize the expected utility of his reward, the person should specify a number x such that $U(x) = E[U(X)]$.

3. Suppose that two people, J and K, desire to make a bet. Person J will pay \$1 to person K if a specific event A occurs, and person K will pay x dollars to person J if the event A does not occur. Suppose that both J and K agree that $P(A) = p (0 < p < 1)$ and that the utility functions U_J and U_K of both persons are strictly increasing functions of monetary gain. Show that there is an amount x which is mutually agreeable to both J and K if, and only if,

$$U_J^{-1}\left[\frac{U_J(0) - pU_J(-1)}{1 - p}\right] < -U_K^{-1}\left[\frac{U_K(0) - pU_K(1)}{1 - p}\right].$$

Show also that when this relation is satisfied, any value of x which is between these two numbers will result in a mutually satisfactory bet.

4. Let $R = \{r_1, r_2, \ldots\}$ be a countable set of rewards, and let U be a utility function on R. Let $P_1, P_2, \ldots$ be a sequence of probability distributions on

R. For each distribution P_i in this sequence and each reward $r_j \in R$, let $p_{ij} = P_i\{\text{receiving } r_j\}$. Also, let $\alpha_1, \alpha_2, \ldots$ be a sequence of numbers such that $\alpha_i \geq 0$ for $i = 1, 2, \ldots$ and $\sum_{i=1}^{\infty} \alpha_i = 1$, and let

$$\bar{p}_j = \sum_{i=1}^{\infty} \alpha_i p_{ij} \qquad j = 1, 2, \ldots .$$

Finally, let $\bar{P}$ denote the probability distribution on R such that $\bar{p}_j = \bar{P}\{\text{receiving } r_j\}$ for $j = 1, 2, \ldots .$

Verify that $\bar{P}$ is a probability distribution. Also, show that if U is a bounded function, then the distribution $\bar{P}$ cannot be preferred to each of the distributions P_i $(i = 1, 2, \ldots)$. Give a simple example to show that if U is not bounded, then it might be true that $E(U|P_i)$ is finite for $i = 1, 2, \ldots$ but $E(U|\bar{P})$ is infinite.

5. For any real numbers r_1 and r_2 and for any probability p $(0 \leq p \leq 1)$, let $\langle r_1, p; r_2, 1-p\rangle$ denote a lottery which yields a reward of r_1 dollars with probability p and a reward of r_2 dollars with probability $1 - p$. Without referring to your utility function as sketched in Exercise 1, decide which lottery you prefer in each of the following pairs:

(a) $\langle 250, \frac{1}{4}; 0, \frac{3}{4}\rangle$ or $\langle 40, \frac{1}{2}; 70, \frac{1}{2}\rangle$;
(b) $\langle 400, \frac{1}{2}; -100, \frac{1}{2}\rangle$ or $\langle 150, \frac{2}{3}; 0, \frac{1}{3}\rangle$;
(c) $\langle 1{,}000, \frac{1}{2}; -1{,}000, \frac{1}{2}\rangle$ or $\langle 50, \frac{1}{2}; -50, \frac{1}{2}\rangle$.

Now find which lottery would be preferred in parts a to c, as determined by the sketch of your utility function made in Exercise 1. If these preferences do not agree with the decisions you have just made, revise your sketch.

6. Suppose that g is a real-valued function over some interval of the real line and that g can be differentiated twice throughout the interval. Prove that g is concave if, and only if, $g''(x) \leq 0$ at every point x in the interval.

7. Consider two boxes each of which contains both red balls and green balls. It is known that one-half the balls in box 1 are red and the other half are green. In box 2, the proportion X of red balls is not known with certainty, but this proportion has a probability distribution over the interval $0 \leq X \leq 1$.

(a) Suppose that a person is to select one ball at random from either box 1 or box 2. If that ball is red, he wins \$1; if it is green, he wins nothing. Show that under any utility function which is an increasing function of monetary gain, the person should prefer selecting the ball from box 1 if, and only if, $E(X) < \frac{1}{2}$.

(b) Suppose that a person can select n balls $(n \geq 2)$ at random from either of the boxes but that all n balls must be selected from the same box; suppose that each selected ball will be put back in the box before the next ball is selected; and suppose that he will receive \$1 for each red ball selected and nothing for each green ball. Also, suppose that his utility function U of monetary gain is strictly concave over the interval $[0, n]$, and suppose that $E(X) = \frac{1}{2}$. Show that the person should prefer to select the balls from box 1.

Hint: Show that if the balls are selected from box 2, then for any given value $X = x$, $E(U|x)$ is a concave function of x on the interval $0 \leq x \leq 1$. This can be done by showing that

$$\frac{d^2}{dx^2} E(U|x)$$

$$= n(n-1) \sum_{i=0}^{n-2} [U(i) - 2U(i+1) + U(i+2)] \binom{n-2}{i} x^i (1-x)^{n-2-i} < 0.$$

Then apply Jensen's inequality to $E(U|X)$.

(c) Consider again the problem in part (b), with the same utility function and the same distribution of X. Suppose now, however, that instead of receiving \$1 for each red ball and nothing for each green ball, the person will receive \$1 for each green ball and nothing for each red ball. Show that he should still prefer to select the balls from box 1.

8. Let p and m be fixed numbers such that $0 < p < 1$ and $m > 0$, and consider the following situation: A person is given a stake of m dollars which he can allocate between an event A of probability p and its complement A^c. He keeps as his reward the amount which he has allocated to either A or A^c, whichever actually occurs. In other words, for all values of x in the interval $0 \leq x \leq m$, he can choose among all lotteries of the form $\langle x, p; m - x, 1 - p\rangle$, as defined in Exercise 5. Being careful in each case to consider every possible pair of values of p and m, find the preferred allocation of the m dollars when the person's utility function U is defined on the interval $[0, m]$ of monetary gains as follows:

(a) $U(r) = r^\alpha$, where $\alpha > 1$.
(b) $U(r) = r$.
(c) $U(r) = r^\alpha$, where $0 < \alpha < 1$.
(d) $U(r) = \log r$.
(e) $U(r) = 2(r/m_0) - (r/m_0)^2$, where $m_0 \geq m$.

9. Consider a roulette wheel which is partitioned into k disjoint events $A_1, \ldots, A_k$ such that $P(A_i) = p_i$, for $i = 1, \ldots, k$, and $\sum_{i=1}^{k} p_i = 1$. Suppose that a person is given a stake of m dollars which he can allocate among the k events $A_1, \ldots, A_k$ and that he receives as his reward the amount x_i which he has allocated to the event A_i that actually occurs. In other words, he can choose among all lottery tickets which yield a reward x_i with probability p_i $(i = 1, \ldots, k)$, subject to the restriction that $x_i \geq 0$ for $i = 1, \ldots, k$ and $\sum_{i=1}^{k} x_i = m$. Suppose also that the person's utility function U, as a function of positive monetary rewards r, is defined as $U(r) = \log r$. Show that his preferred allocation is given by $x_i = mp_i$ for $i = 1, \ldots, k$. (Note that this exercise is a generalization of Exercise 8d.)

10. Suppose that a person's current fortune is x_0 dollars and that he has an opportunity to purchase a lottery ticket which yields either a reward of r dollars $(r > 0)$ or a reward of 0 dollars with equal probabilities. Suppose also that for any value of x $(x > 0)$, his utility $U(x)$ of having a fortune of x dollars is $U(x) = \log x$. Show that he should be willing to purchase the lottery ticket for any amount b such that

$$b < x_0 + \tfrac{1}{2}[r - (r^2 + 4x_0^2)^{\frac{1}{2}}].$$

11. Suppose that Mr. A and Mr. B are equally wealthy, each having a current fortune of x_0 dollars. Suppose also that both men have the same utility function U and that for any number x, their utility of having a fortune of x dollars is $U(x) = (x - x_0)^{\frac{1}{2}}$.

Suppose now that one of the two men receives, as a gift, a lottery ticket which yields either a reward of r dollars $(r > 0)$ or a reward of 0 dollars with equal probabilities. Show that there exists a number $b > 0$ having the following property: Regardless of which man receives the lottery ticket, he can sell it to the other man for b dollars and the sale will be advantageous to both men.

12. Let R be an abstract set of rewards on which a utility function U is defined which has the property stated in Theorem 1 of Sec. 7.10. For any reward $r \in R$, let the set $[r, \infty)$ be defined by Eq. (1) of Sec. 7.10. Suppose that P and Q

are distributions in the class $\mathcal{P}_E$ such that for every reward $r \in R$,

$$P\{[r, \infty)\} \leq Q\{[r, \infty)\}.$$

Show that $P \lesssim^* Q$.

This exercise can be rephrased as follows: Let X and Y be random variables with d.f.'s F and G, respectively. Suppose that both $E(X)$ and $E(Y)$ exist, and suppose that $F(x) \leq G(x)$ for every number x. Show that $E(X) \geq E(Y)$. [See Lehmann (1959), p. 73.]

13. Suppose that a person is going to sell Fizzy Cola at a football game and must decide in advance how much to order. Suppose that he makes a gain of m cents on each quart that he sells at the game but suffers a loss of c cents on each quart that he has ordered but does not sell. If it is assumed that the demand for Fizzy Cola at the game, as measured in quarts, is an absolutely continuous random variable X with p.d.f. f and d.f. F, show that his expected profit will be maximized if he orders an amount α such that $F(\alpha) = m/(m + c)$.

PART THREE

statistical decision problems

14/60

Pasta 4

Nº cópias 5

Lista de Exercícios 3

MAE 5753 - 2º semestre /2015

Exercícios extraídos do livro "Optimal Statistical Decisions", de M. H. De Groot.

Exercícios 3 e 4, página 149

Exercícios 5, 6 e 9, página 150

Exercícios 11 e 12, página 151

Exercícios 14, 15, 16, 17 e 19, página 152

Exercícios 21 e 22, página 153

Exercício 27, página 154

15 exercícios

the posterior g.p.d.f. $\xi(\cdot\,|x)$ of W, before Y is observed, as follows:

$$\xi(w|x) = \frac{g(x|w)\xi(w)}{\int_\Omega g(x|w')\xi(w')\,d\nu(w')}. \tag{2}$$

Furthermore, the conditional g.p.d.f. $h(\cdot\,|w, x)$ of Y when $W = w$ and $X = x$ is

$$h(y|w, x) = \frac{f(x, y|w)}{g(x|w)}. \tag{3}$$

Hence, for the second stage of the experiment, which involves the observation of Y, the g.p.d.f. given by Eq. (2) can be regarded as the prior g.p.d.f. of W, and the conditional g.p.d.f.'s given by Eq. (3) for each value $w \in \Omega$ form the appropriate family of distributions of Y. The posterior g.p.d.f. $\xi(\cdot\,|x, y)$ of W when $Y = y$ can now be computed as follows:

$$\xi(w|x, y) = \frac{h(y|w, x)\xi(w|x)}{\int_\Omega h(y|w', x)\xi(w'|x)\,d\nu(w')}. \tag{4}$$

If the g.p.d.f.'s given by Eqs. (2) and (3) are substituted in Eq. (4), the result is Eq. (1). This means that if the observations are made in more than one stage, then the posterior distribution can be computed in different stages by letting the posterior distribution after each stage serve as the prior distribution for the next stage. It also follows from this derivation that if the posterior distribution of W when $X = x$ and $Y = y$ is computed in two stages, the final result is the same regardless of whether X or Y is observed first.

The decision-making process can now be described in the following highly simplified, but helpful, way. At any given time, the statistician has a probability distribution for the parameter W. As time progresses, the statistician gains information about W from various sources and uses this information to revise his distribution for W. From time to time, when he must choose a decision whose consequences are related to W, he will select a decision that is optimal against his current distribution for W.

Some aspects of this description of decision making are realistic. As we live our lives, we revise our beliefs about various parameters as we learn about them, and when a decision must be chosen, we obviously make the choice on the basis of our current beliefs. However, in some situations, the choice of a decision at a certain time may affect the information that will become available and, hence, may affect the decisions that the statistician can choose at future times. Problems in which the statistician must consider the future and plan ahead appropriately are called *sequential decision problems*. Such problems will be treated in Chaps. 12 to 14.

EXERCISES

1. Consider a decision problem in which the parameter space $\Omega = \{w_1, w_2, w_3, w_4\}$, the decision space $D = \{d_1, d_2, d_3\}$, and the loss function L is specified by the accompanying table. Suppose that the p.f. ξ of the parameter W is such that $\xi(w_1) = \frac{1}{8}$, $\xi(w_2) = \frac{3}{8}$, $\xi(w_3) = \frac{1}{4}$, and $\xi(w_4) = \frac{1}{4}$. Show that decision d_3 is the Bayes decision against ξ.

Table for Exercise 1

	d_1	d_2	d_3
w_1	0	2	3
w_2	1	0	2
w_3	3	4	0
w_4	1	2	0

2. Consider a decision problem in which the parameter space $\Omega = \{w_1, w_2\}$, the decision space $D = \{d_1, d_2, d_3\}$, and the loss function L is specified by the accompanying table. Prove that decision d_3 is a Bayes decision against a given distribution of the parameter W if, and only if, $\frac{1}{3} \le \Pr(W = w_1) \le \frac{5}{9}$.

Table for Exercise 2

	d_1	d_2	d_3
w_1	0	10	4
w_2	8	0	3

✓ 3. Consider a decision problem in which the parameter space Ω contains all the numbers w in the interval $0 \le w \le 1$, the decision space D is the real line R^1, and for any numbers $w \in \Omega$ and $d \in D$, the loss $L(w, d)$ is defined by the equation

$$L(w, d) = 100(w - d)^2.$$

Suppose that the p.d.f. ξ of the parameter W is specified by the equation

$$\xi(w) = 2w \qquad 0 \le w \le 1.$$

Show that the value $d = \frac{2}{3}$ is the Bayes decision against ξ and that the Bayes risk is $\frac{50}{9}$.

✓ 4. Consider the decision problem described in Exercise 1, but suppose that the loss function L specified in that exercise is replaced by the new loss function L_0 specified by the table on page 150. Prove that against any distribution of the parameter W, the Bayes decisions when the loss function is L_0 are the same as the Bayes decisions when the loss function is L.

Table for Exercise 4

	d_1	d_2	d_3
w_1	4	6	7
w_2	0	−1	1
w_3	0	1	−3
w_4	1	2	0

5. Suppose that in an arbitrary decision problem, a certain decision d^* is a Bayes decision against two distinct distributions P_1 and P_2 of the parameter W. For any number α such that $0 < \alpha < 1$, show that d^* must also be a Bayes decision against the distribution $\alpha P_1 + (1 - \alpha)P_2$.

6. Consider the decision problem presented in Exercise 3. Suppose that statistician A believes that the p.d.f. of W is ξ, as specified in Exercise 3, but that statistician B believes that the p.d.f. of W is ξ_B, as specified by the equation

$$\xi_B(w) = 3w^2 \qquad 0 \le w \le 1.$$

How much additional risk does B think that A will incur because of A's incorrect belief about the distribution of W?

7. Consider a decision problem in which $\Omega = \{w_1, w_2\}$, $D = \{d_1, \ldots, d_5\}$, and the loss function L is given by the accompanying table. Against what distributions of the parameter W is the Bayes decision not unique?

Table for Exercise 7

	d_1	d_2	d_3	d_4	d_5
w_1	0	4	2	1	5
w_2	4	5	0	1	0

8. Consider a decision problem in which $\Omega = \{w_1, w_2\}$, $D = \{d_1, \ldots, d_7\}$, and the loss function L is given by the accompanying table. For each distribution ξ of the parameter W, find all Bayes decisions against ξ.

Table for Exercise 8

	d_1	d_2	d_3	d_4	d_5	d_6	d_7
w_1	1	6	0	2	7	3	4
w_2	10	1	13	8	0	5	4

9. Consider a decision problem in which both Ω and D have an infinite number of elements. Suppose that $\Omega = \{w_1, w_2, \ldots\}$, $D = \{d^*, d_1, d_2, \ldots\}$, and the loss function L is given by the accompanying table. Prove that d^* is the only admissible decision in the set D but that d^* is not a Bayes decision against any distribution of the parameter W.

Table for Exercise 9

	d^*	d_1	d_2	d_3	d_4	$\cdots$
w_1	$\frac{1}{2}$	0	0	0	0	$\cdots$
w_2	$\frac{1}{2}$	1	0	0	0	$\cdots$
w_3	$\frac{1}{2}$	1	1	0	0	$\cdots$
w_4	$\frac{1}{2}$	1	1	1	0	$\cdots$
w_5	$\frac{1}{2}$	1	1	1	1	$\cdots$
$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdots$

10. Suppose that there is probability $\frac{1}{10}$ that a signal is present in a certain system at any given time and probability $\frac{9}{10}$ that no signal is present in the system. Suppose that a measurement made on the system when a signal is present is normally distributed with mean 50 and precision 1 and a measurement made on the system when no signal is present is normally distributed with mean 52 and precision 1. Suppose that a measurement made on the system at a certain time has the value x. Show that the posterior probability that a signal is present is greater than the posterior probability that no signal is present if, and only if, $x < 51 - \frac{1}{2}\log 9$.

11. Consider a decision problem in which $\Omega = \{w_1, w_2\}$, $D = \{d_1, d_2\}$, and the loss function L is given by Table 8.2. Suppose that the statistician can observe a random variable Z whose conditional distribution when $W = w_1$ is normal with mean 0 and variance 1 and whose conditional distribution when $W = w_2$ is normal with mean 1 and variance 1. For any given prior probability $\xi = \Pr(W = w_1)$, let δ be the decision function such that $\delta(z) = d_1$ if

$$z \le \frac{1}{2} + \log \frac{\xi}{2(1 - \xi)}$$

and $\delta(z) = d_2$ otherwise. (*a*) Show that δ is a Bayes decision function against ξ. (*b*) Sketch the Bayes risk $\rho^*(\xi)$ as a function of ξ over the interval $0 \le \xi \le 1$.

12. Suppose that a statistician must decide whether a certain random variable has a uniform distribution on the interval (0, 1) or a uniform distribution on the interval $(0, \frac{1}{2})$. Each of these two distributions has probability $\frac{1}{2}$ of being the correct distribution. Suppose that the loss is 0 if a correct decision is made and is $a > 0$ if an incorrect decision is made. Suppose that the statistician can select the number of observations which will be made on the random variable but that the cost of each observation is $c > 0$. Show that the statistician should make n^* observations, where n^* is the nonnegative integer which minimizes the value $a2^{-(n+1)} + nc$.

13. Consider a decision problem in which $\Omega = \{w_1, w_2\}$, $D = \{d_1, d_2, d_3\}$, and the loss function L is given by the table on page 152. Suppose that an observation X is available with the following conditional distributions:

$$\Pr(X = 1|W = w_1) = \tfrac{3}{4}, \qquad \Pr(X = 0|W = w_1) = \tfrac{1}{4};$$
$$\Pr(X = 1|W = w_2) = \tfrac{1}{4}, \qquad \Pr(X = 0|W = w_2) = \tfrac{3}{4}.$$

Suppose that $\xi = \Pr(W = w_1)$. Find a Bayes decision function against each value of ξ $(0 \le \xi \le 1)$, and sketch the Bayes risk $\rho^*(\xi)$ as a function of ξ.

Table for Exercise 13

	d_1	d_2	d_3
w_1	0	10	3
w_2	10	0	3

14. Suppose that for the conditions in Exercise 13, before the statistician chooses a decision, he can observe the values of the random variables $X_1, \ldots, X_n$ which, for any given value $W = w_i$ $(i = 1, 2)$, are a random sample from the conditional distribution of X. For any given sample size n, find a Bayes decision function against each value of ξ.

15. Suppose that for the conditions in Exercise 14, $\xi = \frac{1}{2}$ and the cost c of each observation is $\frac{1}{10}$. Show that the optimal sample size n is 8 and that the minimum total risk which can be attained is 1.33.

16. Suppose that for the conditions in Exercises 13 and 14, the cost of each observation is not constant but that any observation whose value is 1 costs 0.15 and any observation whose value is 0 costs 0.05. Show that when $\xi = \frac{1}{2}$, the optimal sample size and minimum total risk will be precisely the same as those in Exercise 15.

17. Consider a decision problem in which $\Omega = \{w_1, w_2\}$, $D = \{d_1, d_2\}$, and the loss function L is given by the accompanying table. Suppose that $\Pr(W = w_1) = \Pr(W = w_2) = \frac{1}{2}$. Suppose also that the conditional distribution of an observation X when $W = w_1$ is normal with mean -1 and variance 9 and that the conditional distribution of X when $W = w_2$ is normal with mean 1 and variance 9. Suppose, further, that before the statistician chooses a decision, he can observe values of the random variables $X_1, \ldots, X_n$ which, for any given value of W, are a random sample from the conditional distribution of X just given. If each observation in the sample costs 1 unit, show that the optimal number of observations n is 42 and that the minimum total risk is 57.4.

Table for Exercise 17

	d_1	d_2
w_1	0	1,000
w_2	1,000	0

18. Consider k random variables $X_1, \ldots, X_k$. Suppose that the p.d.f. of exactly one of these random variables is g and the p.d.f. of each of the other $k - 1$ random variables is h but that it is not known which one of the random variables has p.d.f. g. For $i = 1, \ldots, k$, let ξ_i be the prior probability that X_i is the random variable whose p.d.f. is g. Here $\xi_i > 0$ for $i = 1, \ldots, k$ and $\Sigma_{i=1}^k \xi_i = 1$. (a) Suppose that the random variable X_1 is observed and found to have the value x. Find the posterior probability that the p.d.f. of X_1 is g. (b) Suppose that the random variable X_2 is observed and found to have the value x. Find the posterior probability that the p.d.f. of X_1 is g.

19. Consider two boxes A and B each of which contains both red balls and

green balls. It is known that in one of the boxes, $\frac{1}{2}$ of the balls are red and $\frac{1}{2}$ of the balls are green and that in the other box, $\frac{1}{4}$ of the balls are red and $\frac{3}{4}$ of the balls are green. Let the box in which $\frac{1}{2}$ of the balls are red be denoted as box W, and suppose that it is not known with certainty whether $W = A$ or $W = B$. Assume that $\Pr(W = A) = \xi$ and $\Pr(W = B) = 1 - \xi$, where ξ is a given number such that $0 < \xi < 1$.

Suppose that the statistician may select one ball at random from either box A or box B and that after observing its color, he must decide whether $W = A$ or $W = B$. Prove that if $\frac{1}{2} < \xi < \frac{2}{3}$, then in order to maximize the probability of making a correct decision, he should select the ball from box B. Prove also that if $\frac{2}{3} \le \xi \le 1$, then it does not matter from which box the ball is selected.

20. Consider a decision problem in which $\Omega = \{w_1, w_2\}$, $D = \{d_1, d_2\}$, and the loss function L is given by Table 8.4. Suppose that the statistician can observe either a random variable X or a random variable Y whose conditional distributions are as follows:

$$\Pr(X = 1|W = w_1) = \frac{2}{3}, \qquad \Pr(X = 0|W = w_1) = \frac{1}{3},$$
$$\Pr(X = 1|W = w_2) = \frac{1}{2}, \qquad \Pr(X = 0|W = w_2) = \frac{1}{2},$$

and

$$\Pr(Y = 1|W = w_1) = \frac{3}{4}, \qquad \Pr(Y = 0|W = w_1) = \frac{1}{4},$$
$$\Pr(Y = 1|W = w_2) = \frac{1}{2}, \qquad \Pr(Y = 0|W = w_2) = \frac{1}{2}.$$

Suppose also that the cost of observing X is the same as the cost of observing Y. Show that for any prior distribution of W and any values of the losses a_1 and a_2, the statistician should observe Y rather than X.

21. Let W be a parameter which takes the values w_1 and w_2 with prior probabilities specified by the equations $\Pr(W = w_1) = \xi$ and $\Pr(W = w_2) = 1 - \xi$. Suppose that an observation X is to be taken with conditional g.p.d.f.'s $f(\cdot\,|w_i)$ for $i = 1, 2$. Let $\xi(x)$ denote the posterior probability that $W = w_1$ when $X = x$. Prove that $E[\xi(X)] = \xi$, where the expectation is computed by assuming that W has the specified prior distribution.

22. For the conditions specified in Exercise 21, if it is assumed that $W = w_1$, prove that $E[\xi(X)] \ge \xi$. *Note:* This exercise can be interpreted as stating that, on the average, the posterior distribution will assign greater probability to the correct value of W than the prior distribution assigned to that value.

23. For the conditions specified in Exercise 21, suppose that the prior distribution of W is such that $\xi = 1 - \xi = \frac{1}{2}$. If it is assumed that $W = w_1$, prove that for any number ϵ $(0 < \epsilon < 1)$,

$$\Pr[\xi(X) \le \epsilon] \le \frac{\epsilon}{1 - \epsilon}.$$

Note: This exercise can be interpreted as stating that there is only a small probability that the posterior distribution will assign a small probability to the correct value of W.

24. Consider a set $\Omega = \{w_1, w_2, w_3\}$ containing three points, and let $\mathcal{P}$ be the set of all probability distributions (p_1, p_2, p_3) such that $p_i \ge 0$ $(i = 1, 2, 3)$ and $p_1 + p_2 + p_3 = 1$. Let $\mathcal{C}$ be the set of points either inside or on the boundary of an equilateral triangle which has a unit height. Also, let v_1, v_2, and v_3 denote the vertices of that triangle, and for $i = 1, 2, 3$, let S_i denote the side of the triangle that is opposite the vertex v_i. Show that the sum of the distances from any point in the triangle to the three sides of the triangle is 1. Then show that there is a one-to-one correspondence between the sets $\mathcal{P}$ and $\mathcal{C}$, in which any point $(p_1, p_2, p_3) \in \mathcal{P}$ corresponds to the point $\mathbf{x} \in \mathcal{C}$ whose distance from side S_i is p_i $(i = 1, 2, 3)$.

25. Suppose that in Exercise 24, corresponding to any fixed point $(p_1, p_2, p_3) \in \mathcal{P}$, a mass p_i is placed at the vertex v_i $(i = 1, 2, 3)$. Find the center of gravity $\mathbf{x}$ of this system of three mass points, and show that the position of $\mathbf{x}$ defines a one-to-one correspondence between $\mathcal{P}$ and $\mathcal{C}$. [For any point $\mathbf{x} \in \mathcal{C}$, the corresponding values (p_1, p_2, p_3) are called the *barycentric coordinates* of $\mathbf{x}$.]

26. Show that the correspondence between $\mathcal{P}$ and $\mathcal{C}$ defined in Exercise 25 is the same as the correspondence defined in Exercise 24.

27. Suppose that each of k statisticians has his own prior distribution for a certain parameter W, and let ξ_i be the g.p.d.f. which statistician i assigns to W $(i = 1, \ldots, k)$. Suppose also that an executive forms his opinion about W from the opinions of the k statisticians and that he assigns to W the g.p.d.f. ξ^* defined, at each point $w \in \Omega$, as follows:

$$\xi^*(w) = \alpha_1\xi_1(w) + \cdots + \alpha_k\xi_k(w).$$

Here $\alpha_1, \ldots, \alpha_k$ are weights such that $\alpha_i \geq 0$ $(i = 1, \ldots, k)$ and $\alpha_1 + \cdots + \alpha_k = 1$. The value of α_i reflects the relative weight that the executive gives to the opinion of statistician i. Suppose further that the k statisticians and the executive observe together the value of a random variable X whose conditional g.p.d.f. when $W = w$ is $f(\cdot \mid w)$. Show that the posterior g.p.d.f. of the executive will again be a linear combination of the posterior g.p.d.f.'s of the k statisticians, with new weights $\beta_1, \ldots, \beta_k$ which will depend on the observed value of X. Also, discuss the conditions under which the weight β_1 in the posterior g.p.d.f. will be greater than the weight α_1 in the prior g.p.d.f.

CHAPTER 9

conjugate prior distributions

9.1 SUFFICIENT STATISTICS

Consider a statistical problem in which a large amount of experimental data has been collected. The treatment of the data is often simplified if the statistician computes a few numerical values, or statistics, and considers these values as summaries of the relevant information in the data. In some problems, a statistical analysis that is based on these few summary values can be just as effective as any analysis that could be based on all the observed values. In this chapter we shall consider problems for which fully informative summaries of this type are available. Such summaries are known as *sufficient statistics.*

Suppose that W is a parameter which takes values in the space Ω. Also, suppose that X is a random variable, or random vector, which takes values in the sample space S. We shall let $f(\cdot \mid w)$ denote the conditional g.p.d.f. of X when $W = w$ $(w \in \Omega)$. It is assumed that the observed value of X will be available for making inferences and decisions relating to the parameter W. In this context, any function T of the observation X, whether or not T is a real-valued function, is called a *statistic.*

Loosely speaking, a statistic T is called a sufficient statistic if, for any prior distribution of W, its posterior distribution depends on the

4ª LISTA DE EXERCÍCIOS 4/11

Pasta 04

Nº cópias 05

MAE 5753 - 2º semestre / 2015

Exercícios do livro "Optimal Statistical Decisions", de M. H. De Groot.

Exercícios 1, 4, 6, 7, 8, 9, 10, 12, 13 e 19, páginas 260 a 263.

EXERCÍCIOS COMPLEMENTARES

Exercícios do livro "Theory of Statistics", de M. J. Schervish

Exercícios 4 e 5 (página 286), 27 e 31 (página 342) e 40 (página 343).

(1) $\Theta = \{\theta_1, \theta_2, \dots\} = \{\theta_n\}_{n \geq 1}$. $\mathcal{F} = \mathcal{P}(\Theta)$. $\mathbb{P} : \mathcal{F} \to [0,1]$: priori para θ.

$\mathcal{X}$: espaço amostral. $X|\theta$ com distribuição $\mathbb{P}(\cdot|\theta)$.

(a) Defina região HPD nesse caso.

(b) Considere o problema $(\Delta, \Omega, L', \mathbb{P}')$, onde $\Delta = \{\delta : \mathcal{X} \to \mathcal{P}(\Theta)\}$, $\Omega = \Theta \times \mathcal{X}$,

$\mathbb{P}'$ como acima e $L' : \Delta \times \Omega \to \mathbb{R}$ dado por

$$L'(\delta, (\theta, x)) = |\delta(x)| + k . \mathbb{I}_{\delta(x)^c}(\theta)$$, onde $|A|$ é o número de elementos de A, $A \subseteq \Theta$

Mostre que a decisão de Bayes associa a cada $x \in \mathcal{X}$ uma região HPD da posteriori de θ dado X.

(2) $\Theta = [0,1]$

$f(\theta) = 4|\theta - \frac{1}{2}| . \mathbb{I}_{[0,1]}(\theta)$. $X|\theta \sim Ber(\theta)$. Para $X = 1$, avaliar HPD em função de α.

(3) $\Theta = \mathbb{R}$

$$f(\theta) = \frac{1}{2}\frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} + \frac{1}{2}\frac{1}{\sqrt{2\pi}} e^{\frac{-(\theta-2)^2}{2}}$$. $X|\theta \sim N(\theta, 1/2)$. Obter posteriori de θ dado $X = x$.

Avaliar HPD em função de $x \in \mathbb{R}$ e $\alpha \in (0,1)$.

18 exercícios

EXERCISES

1. Suppose that $X_1, \ldots, X_n$ is a random sample from a Poisson distribution for which the value of the mean W is unknown. Suppose also that the prior distribution of W is a gamma distribution with parameters α and β ($\alpha > 1, \beta > 0$). Finally, suppose that it is desired to estimate the value of W and that the loss function L is specified by Eq. (13) of Sec. 11.2. Prove that the Bayes estimator of W is specified by Eq. (14) of Sec. 11.2, and prove that the Bayes risk is $1/(\beta + n)$.

2. Suppose that $X_1, \ldots, X_n$ is a random sample from a Bernoulli distribution for which the parameter W is unknown, and suppose that the prior distribution of W is a beta distribution with parameters α and β. If the value of W must be estimated when the loss L is specified by the equation $L(w, d) = (w - d)^2$ and if the sampling cost is c per observation, show that the optimal number of observations n is specified by the following equation:

$$n = \left[\frac{\alpha\beta}{c(\alpha + \beta)(\alpha + \beta + 1)}\right]^{\frac{1}{2}} - (\alpha + \beta).$$

3. Suppose that $X_1, \ldots, X_n$ is a random sample from a normal distribution with an unknown value of the mean W and a specified value of the precision r. Suppose also that the prior distribution of W is a normal distribution with mean μ and precision τ. If the value of W must be estimated when the loss L is specified by the equation $L(w, d) = (w - d)^2$ and if the sampling cost is c per observation, show that the optimal number of observations n is specified by the following equation:

$$n = \left(\frac{1}{cr}\right)^{\frac{1}{2}} - \frac{\tau}{r}.$$

4. Suppose that $X_1, \ldots, X_n$ is a random sample from a uniform distribution on the interval $(0, W)$, where the value of W is unknown and the prior distribution of W is a Pareto distribution with parameters w_0 and α ($\alpha > 2$). If the value of W must be estimated when the loss L is specified by the equation $L(w, d) = (w - d)^2$ and if the sampling cost is c per observation, show that the optimal number of observations n is specified by the following equation:

$$n = \left[\frac{2\alpha w_0^2}{c(\alpha - 2)}\right]^{\frac{1}{3}} - (\alpha - 1).$$

5. Let W be a random variable whose d.f. is G, and suppose that $E(|W|^\alpha) < \infty$, where $\alpha > 1$. Prove that the expectation $E(|W - d|^\alpha)$ is minimized when d is the unique number such that

$$\int_{w<d} (d - w)^{\alpha-1}\, dG(w) = \int_{w>d} (w - d)^{\alpha-1}\, dG(w).$$

6. Let W be a random variable with g.p.d.f. ξ, and suppose that ξ is symmetric with respect to the value v, so that $\xi(v + y) = \xi(v - y)$ for all values of y ($-\infty < y < \infty$). Suppose that Λ is a nonnegative convex function on the real line that is symmetric with respect to the value 0, and suppose also that for all values of d ($-\infty < d < \infty$),

$$R(d) = \int_{-\infty}^{\infty} \Lambda(w - d)\xi(w)\, d\nu(w) < \infty.$$

Prove that R is a convex function which is also symmetric with respect to the value v and that $R(d)$ is minimized at the value $d = v$.

7. Suppose that W is a parameter whose value must be estimated when the loss function L is specified by the following equation:

$$L(w, d) = \left(\frac{w - d}{d}\right)^2.$$

It is assumed that $E(W^2) < \infty$. If $E(W) \neq 0$, show that a Bayes estimate d is specified by the equation

$$d = \frac{E(W^2)}{E(W)}.$$

If $E(W) = 0$ and $E(W^2) > 0$, show that there is no Bayes estimate. Finally, show that in either case, the Bayes risk ρ^* is

$$\rho^* = \frac{\operatorname{Var}(W)}{E(W^2)}.$$

8. Suppose that W is a parameter whose value must be estimated when the loss function L is specified as follows:

$$L(w, d) = \begin{cases} k_1(w - d) & \text{for } d \leq w, \\ k_2(d - w) & \text{for } d \geq w. \end{cases}$$

Here k_1 and k_2 are positive constants and it is assumed that $E(|W|) < \infty$. Show that a number d will be a Bayes estimate if, and only if, the following relations are satisfied:

$$\Pr(W \leq d) \geq \frac{k_1}{k_1 + k_2} \quad \text{and} \quad \Pr(W \geq d) \geq \frac{k_2}{k_1 + k_2}.$$

9. Let $X_1, \ldots, X_n$ be a random sample from a Bernoulli distribution for which the parameter W is unknown. Suppose that the value of W must be estimated when the loss function L is

$$L(w, d) = \frac{(w - d)^2}{w(1 - w)}.$$

Suppose also that the prior distribution of W is the uniform distribution on the interval $(0, 1)$. Show that the Bayes estimator δ^* is specified by the equation $\delta^*(X_1, \ldots, X_n) = \bar{X}$ and show that the Bayes risk is $1/n$.

10. Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution for which the parameter W is unknown. Suppose that the value of $1/W$ must be estimated when the loss function L is

$$L(w, d) = \left(\frac{1}{w} - d\right)^2.$$

Suppose also that the prior distribution of W is a gamma distribution with parameters α and β such that $\alpha > 2$. (*a*) If the number of observations n is fixed, find a Bayes decision function and compute the Bayes risk. (*b*) If the cost per observation is c, show that the optimal number of observations n is specified by the following equation:

$$n = \frac{\beta}{[c(\alpha - 1)(\alpha - 2)]^{\frac{1}{2}}} - (\alpha - 1).$$

11. Suppose that $\mathbf{W} = (W_1, \ldots, W_k)'$ is a random vector for which the mean vector is $\boldsymbol{\mu}$ and the covariance matrix is $\boldsymbol{\Sigma}$. Show that for any fixed $k \times k$

matrix $\mathbf{A}$,

$$E[(\mathbf{W}-\boldsymbol{\mu})'\mathbf{A}(\mathbf{W}-\boldsymbol{\mu})] = \operatorname{tr}(\mathbf{A\Sigma}).$$

12. Suppose that the random vector $\mathbf{X} = (X_1, \ldots, X_k)'$ has a multinomial distribution for which the parameters are n and $\mathbf{W} = (W_1, \ldots, W_k)'$. Suppose that the value of $\mathbf{W}$ must be estimated when the loss function L is

$$L(\mathbf{w}, \mathbf{d}) = \sum_{i=1}^{k} (w_i - d_i)^2.$$

Finally, suppose that the prior distribution of $\mathbf{W}$ is a Dirichlet distribution for which the parametric vector is $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)'$. Find the Bayes estimator, and show that the value of the Bayes risk is

$$\frac{\alpha_0^2 - \Sigma_{i=1}^k \alpha_i^2}{\alpha_0(\alpha_0+1)(\alpha_0+n)},$$

where $\alpha_0 = \Sigma_{i=1}^k \alpha_i$.

13. Let $X_1, \ldots, X_n$ be a random sample from a normal distribution for which both the mean M and the precision R are unknown. Suppose that the value of M must be estimated and that when $M = m$, $R = r$, and the value of the estimate is d, the loss is

$$L(m, r, d) = (m-d)^2.$$

Suppose also that the prior joint distribution of M and R is a normal-gamma distribution as specified in Theorem 1 of Sec. 9.6, and assume that $\alpha > 1$ in this prior distribution. For any fixed number of observations n, find a Bayes decision function and show that the Bayes risk is $\beta/[(\alpha-1)(\tau+n)]$.

14. Suppose that $X_1, \ldots, X_n$ is a random sample from a Bernoulli distribution for which the parameter W is unknown, and suppose also that the prior distribution of W is a beta distribution with parameters α and β. Show that when

$$1 - \alpha < \sum_{i=1}^{n} X_i < n + \beta - 1,$$

a generalized maximum likelihood estimator $\hat{w}(X_1, \ldots, X_n)$ of W is specified by the equation

$$\hat{w}(X_1, \ldots, X_n) = \frac{\Sigma_{i=1}^n X_i + \alpha - 1}{n + \alpha + \beta - 2}.$$

15. Suppose that $X_1, \ldots, X_n$ is a random sample from a Poisson distribution for which the mean W is unknown, and suppose also that the prior distribution of W is a gamma distribution with parameters α and β. Show that when $\alpha + \Sigma_{i=1}^n X_i > 1$, a generalized maximum likelihood estimator $\hat{w}(X_1, \ldots, X_n)$ of W is specified by the equation

$$\hat{w}(X_1, \ldots, X_n) = \frac{\Sigma_{i=1}^n X_i + \alpha - 1}{n + \beta}.$$

16. Suppose that $X_1, \ldots, X_n$ is a random sample from a normal distribution for which both the mean M and the precision R are unknown. Suppose also that the prior joint distribution of M and R is a normal-gamma distribution as specified in

Theorem 1 of Sec. 9.6. If μ' and β' are as defined in that theorem, show that generalized maximum likelihood estimators $\hat{m}(X_1, \ldots, X_n)$ and $\hat{r}(X_1, \ldots, X_n)$ are specified by the equations $\hat{m}(X_1, \ldots, X_n) = \mu'$ and

$$\hat{r}(X_1, \ldots, X_n) = \frac{\alpha + (n-1)/2}{\beta'}.$$

17. Consider the problem presented in Sec. 11.6. Prove that the posterior probability that $W = w_1$ will be greater than the prior probability p if, and only if, the magnitude of the difference between the observed value of $\bar{x}$ and the value specified in Eq. (5) of Sec. 11.6 is less than the following number:

$$\frac{1}{n}\left\{(\tau+n)\left[\tau(w_1-\mu)^2 + \log\frac{\tau+n}{\tau}\right]\right\}^{\frac{1}{2}}.$$

18. Verify that Eq. (2) of Sec. 11.7 is correct.

19. Let $X_1, \ldots, X_n$ be a random sample from a Bernoulli distribution for which the value of the parameter W is unknown. Suppose that under the prior distribution of W, $\Pr(W = \frac{1}{2}) = p > 0$ and the remaining probability $1 - p$ is uniformly distributed over the interval $0 < W < 1$. Find the posterior probability that $W = \frac{1}{2}$ when $\Sigma_{i=1}^n X_i = y$, and show that this posterior probability is greater than the prior probability p if, and only if, the following inequality is satisfied:

$$\binom{n}{y}\left(\frac{1}{2}\right)^n > \frac{1}{n+1}.$$

20. Suppose that the random vector $\mathbf{X} = (X_1, \ldots, X_k)'$ has a multinomial distribution with parameters n and $\mathbf{W} = (W_1, \ldots, W_k)'$, where n is a specified positive integer and the values of the components of the vector $\mathbf{W}$ are unknown. Suppose also that the prior distribution of $\mathbf{W}$ satisfies the following three conditions: (a) $\Pr(W_1 = \frac{1}{2}) = p > 0$. (b) When $W_1 = \frac{1}{2}$, the joint distribution of $W_2, \ldots, W_k$ is a uniform distribution on the set of points $(w_2, \ldots, w_k)$ such that $w_i > 0$ $(i = 2, \ldots, k)$ and $\Sigma_{i=2}^k w_i = \frac{1}{2}$. (c) When $W_1 \neq \frac{1}{2}$, the joint distribution of $W_1, \ldots, W_k$ is a uniform distribution on the set of points $(w_1, \ldots, w_k)$ such that $w_i > 0$ $(i = 1, \ldots, k)$, $\Sigma_{i=1}^k w_i = 1$, and $w_1 \neq \frac{1}{2}$. Find the posterior distribution of the vector $\mathbf{W}$ when $X_i = x_i$ $(i = 1, \ldots, k)$, and show that

$$\frac{\Pr(W_1 = \frac{1}{2} \mid X_1 = x_1, \ldots, X_k = x_k)}{\Pr(W_1 \neq \frac{1}{2} \mid X_1 = x_1, \ldots, X_k = x_k)} = \frac{p}{1-p}\,\frac{n+k-1}{2^n(k-1)}\binom{n+k-2}{x_1}.$$

21. Let F be a distribution function on the real line for which the mean μ exists, and let the function T_F on the real line be defined by Eq. (9) of Sec. 11.8. (a) Prove that T_F is a nonnegative convex function which is strictly decreasing at any point s such that $T_F(s) > 0$. (b) Prove that T_F satisfies the relations (10) of Sec. 11.8.

22. If F is the d.f. of a normal distribution for which the mean is μ and the precision is τ, prove that the function T_F satisfies Eq. (3) of Sec. 11.9.

23. Suppose that the joint distribution of two random variables X and Y is as follows: The conditional distribution of X when $Y = y$ $(-\infty < y < \infty)$ is a normal distribution with mean y and variance σ_1^2, and the marginal distribution of Y is a normal distribution with mean μ and variance σ_2^2. Show that the marginal distribution of X is a normal distribution with mean μ and variance $\sigma_1^2 + \sigma_2^2$.

24. The mileage Y that can be obtained from a certain gasoline is a random variable which depends on the amount x of a certain chemical in the gasoline. For any given value of x, the mileage Y has a normal distribution for which the mean is $m_1 + m_2x + m_3x^2$ and the precision is w. Here m_1, m_2, m_3, and w are the unknown

3. Let $X = (X_1, \dots, X_n)$ be such that the X_i are conditionally IID with $N(0, \sigma^2)$ distribution given $\Sigma = \sigma$ under the hypothesis H. Let $T(x) = \sqrt{n}\overline{x}/s$, where $\overline{x} = \sum_{i=1}^n x_i/n$ and $s^2 = \sum_{i=1}^n (x_i - \overline{x})^2/(n-1)$. Define $x \preceq y$ if $T(x) \le T(y)$. Find $p_H(x)$, showing that it is the same for all σ.

Section 4.2:

4. Suppose that $X \sim N(\theta, 1)$ given $\Theta = \theta$. Suppose that $L(\theta, 1) < L(\theta, 0)$ for all $\theta < \theta_0$ and that $L(\theta, 1) > L(\theta, 0)$ for all $\theta > \theta_0$. Prove that, for every prior there exists k such that the formal Bayes rule will be to choose action $a = 1$ if $X < k$.

5. Suppose that $X_1, \dots, X_n$ are conditionally IID with $N(\mu, \sigma^2)$ distribution given $\Theta = (\mu, \sigma)$. Use the improper prior having Radon–Nikodym derivative $1/\sigma$ with respect to Lebesgue measure on $(0, \infty) \times \mathbb{R}$. Let μ_0 and d be known values, and suppose that the loss function is

$$L(\theta, a) = \begin{cases} c & \text{if } a = 1 \text{ and } |\mu - \mu_0| \le d\sigma, \\ 1 & \text{if } a = 0 \text{ and } |\mu - \mu_0| > d\sigma, \\ 0 & \text{otherwise.} \end{cases}$$

Prove that the formal Bayes rule will be of the following form: Choose $a = 1$ if $|T| > k$ for some constant k, where $T = \sqrt{n}(\overline{X}_n - \mu_0)/S_n$ and

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, \quad S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

6. Suppose that $X \sim N(\theta, 1)$ given $\Theta = \theta$ and Θ has $\Pr(\Theta = \theta_0) = p_0$ and given $\Theta \ne \theta_0$, $\Theta \sim N(\theta_0, \tau^2)$. Prove that the posterior density of Θ with respect to the measure $\nu(A) = I_A(\theta_0) + \lambda(A)$, where λ is Lebesgue measure, is given by

$$f_{\Theta|X}(\theta|x) = \begin{cases} p_1 & \text{if } \theta = \theta_0, \\ \frac{(1-p_1)\tau}{\sqrt{2\pi(1+\tau^2)}} \exp\left[-\frac{\tau^2}{2(1+\tau^2)}(\theta - \theta_1)^2\right] & \text{if } \theta \ne \theta_0, \end{cases}$$

where

$$\theta_1 = \frac{x\tau^2 + \theta_0}{1 + \tau^2},$$

$$\frac{p_1}{1 - p_1} = \frac{p_0}{(1 - p_0)}\sqrt{1 + \tau^2}\exp\left\{-\frac{1}{2}\left[\frac{\tau^2}{1 + \tau^2}\right](x - \theta_0)^2\right\}.$$

7. Suppose that $X \sim N(\theta, 1)$ given $\Theta = \theta$. Let $H : \Theta = \theta_0$ and $A : \Theta \ne \theta_0$. Let the conditional prior given $\Theta \ne \theta_0$ be $N(\theta_0, \tau^2)$.

   (a) Prove that the Bayes factor is minimized if

$$\tau^2 = \begin{cases} (x - \theta_0)^2 - 1 & \text{if } |x - \theta_0| > 1, \\ 0 & \text{otherwise.} \end{cases}$$

   (b) Show that the minimum Bayes factor is $|x - \theta_0| \exp(\{-[x - \theta_0]^2 + 1\}/2)$ if $|x - \theta_0| > 1$, and is 1 if $|x - \theta_0| \le 1$.

8. In Example 4.21 on page 224, prove that

$$k = \lim_{\tau \to \infty} \frac{p_0}{\Phi\left(\frac{\sqrt{\pi/2}}{\tau}\right) - \Phi\left(\frac{-\sqrt{\pi/2}}{\tau}\right)}.$$

Section 4.3.1:

9. Let (α_0, α_1) be a point on the lower boundary of the risk set for a simple–simple hypothesis-testing problem. Prove that $\alpha_0 + \alpha_1 \le 1$.

10. In a simple–simple hypothesis-testing problem, prove that the minimax rule for a 0–1 loss function is any test that corresponds to the point where ∂_L intersects the line $y = x$.

11. Let $\Omega = \{0, 1\}$. Suppose that P_0 says that $X \sim U(-\sqrt{3}, \sqrt{3})$ and P_1 says that $X \sim N(0, 1)$. Let $\Omega_H = \{0\}$. Draw the risk set for a 0–1 loss function, and find the minimax rule.

12. In a simple–simple hypothesis-testing problem with 0–1 loss, show that a MP level α test has size α unless all tests with size α are inadmissible.

13. Return to the situation in Problem 28 on page 212. Consider the hypothesis $H : X \sim f_0$ versus $A : X \sim f_1$. Find all α such that the MP level α test is of the form "Reject H if $-d \le X \le d$," and write d as a function of α.

14.*Prove Proposition 4.46 on page 235.

15. In Example 4.49 on page 236, prove that the Bayes rule with size α has higher power than the unconditional size α test.

Section 4.3.2:

16. Suppose that the loss function is 0–1–c, that $\Omega_H = \{\theta_0\}$, and that $\Omega = \{\theta_0\} \cup \Omega_A$. Prove that a UMP level α test has no larger Bayes risk than any other size α test, no matter what prior we use.

Section 4.3.3:

17. Let $Y = |X|$ where $f_{X|\Theta}(x|\theta) = 1/(\pi\theta[1 + (x/\theta)^2])$. Suppose that $\Theta > 0$ for sure. Prove that the family of distributions for Y has MLR.

18. Let X have Cauchy distribution $Cau(\theta, 1)$ given $\Theta = \theta$.

   (a) Prove that the MP level α test of $H : \Theta = \theta_0$ versus $A : \Theta = \theta_1$ for $\theta_1 > \theta_0$ is essentially unique. That is, if ϕ and ψ are both MP level α tests, then $P_\theta(\phi(X) = \psi(X)) = 1$ for all θ.

   (b) Prove that there is no UMP level α test of $H : \Theta = \theta_0$ versus $A : \Theta > \theta_0$ for $0 < \alpha < 1$.

19. Let the parameter space be the open interval $\Omega = (0, 100)$. Let X_1 and X_2 be conditionally independent given $\Theta = \theta$ with $X_1 \sim Poi(\theta)$ and $X_2 \sim Poi(100 - \theta)$. We are interested in the hypothesis $H : \Theta \le c$ versus $A : \Theta > c$.

26. Consider the situation in Example 5.30 on page 308. Let squared error be the loss, that is, $L(\theta, a) = (\mu^2 - a)^2$. Show that the UMVUE dominates the MLE if $n \geq 2$. Find a formula for the difference in the risk functions. Also, find an estimator that dominates both the MLE and the UMVUE.

Section 5.1.4:

27. Consider the situation in Problem 6 on page 339. Find the likelihood function and show that a Bayesian would take the data at face value (that is, a Bayesian would calculate the same posterior as if the sample size had been fixed in advance at whatever value N turns out to be, no matter what the prior is).

28. Consider the situation in Problem 7 on page 339. If Λ and M are independent a priori with $\Lambda \sim \Gamma(a, b)$ and $\mathrm{M} \sim \Gamma(c, d)$, find the posterior mean of Λ given the data.

29. In Example 5.10 on page 299, find the posterior mean of Θ given $X = x$ for all x assuming that the prior for Θ is $Beta(\alpha_0, \beta_0)$.

30. Return to the situation of Problem 16 on page 140. If Θ has a prior density $f_\Theta(\theta) = ac^a I_{[c,\infty)}(\theta)/\theta^{a+1}$, find the posterior mean of Θ.

31.*Suppose that, conditional on N, $\{X_i\}_{i=1}^\infty$ are independent with the first N of them having $Ber(1/3)$ distribution and the rest having $Ber(2/3)$ distribution. The prior for N is $f_N(n) = 2^{-n}$ for $n = 1, 2, \ldots$.

(a) Find the posterior distribution of N given a finite sample $X_1, \ldots, X_n$, for known n.

(b) If $X_1 = 0, \ldots, X_n = 0$ is the observed finite sample, find the posterior mean of N.

Section 5.1.5:

32. Let $\mathcal{P}_0$ be the set of distributions on $(\mathbb{R}, \mathcal{B}^1)$ with finite variance. Let $T(P)$ be the standard deviation of the distribution P. Show that $IF(x; T, P) = (x-\mu)^2/[2\sigma] - \sigma/2$, where μ is the mean of P.

33. Let $\mathcal{P}_0$ be the class of distributions on $(\mathbb{R}, \mathcal{B}^1)$ with bounded support, and let $T(P)$ be the supremum of the support. Prove that the influence function for T is $IF(x; T, P) = 0$ if $x \leq T(P)$ and ∞ if $x > T(p)$.

34. Find the influence function for the $100\alpha\%$ trimmed mean at a continuous distribution P.

Section 5.2:

35. Prove Proposition 5.48 on page 316.

36. Let the parameter space be Ω with σ-field of subsets τ. Let X be a random quantity taking values in a set $\mathcal{X}$, and let X have conditional density $f_{X|\Theta}$ given Θ. Let $v : \mathcal{X} \to [0, \infty)$ be a measurable function such that the set

$$L(x) = \{\theta \in \Omega : f_{X|\Theta}(x|\theta) \geq v(x)\}$$

is in τ for all x. Let Θ have a prior distribution μ_Θ, and let μ_X denote the prior predictive distribution of X. Let $C : \mathcal{X} \to \tau$ be another set function such that $\mu_\Theta(C(x)) \leq \mu_\Theta(L(x))$, a.s. $[\mu_X]$. Prove that $\Pr(\Theta \in C(x)|X = x) \leq \Pr(\Theta \in L(x)|X = x)$, a.s. $[\mu_X]$.

37. Prove Proposition 5.56 on page 319.

38. Prove Proposition 5.61 on page 321.

39. Prove Proposition 5.79 on page 329.

40. Suppose that $\Omega = \mathbb{R}$ and that the posterior density of Θ given X is strongly unimodal. Let the action space be the set of all closed and bounded intervals $[a_1, a_2]$ in $\mathbb{R}$.

(a) Let the loss function be $L_l(\theta, [a_1, a_2]) = a_2 - a_1 + c\left(1 - I_{[a_1,a_2]}(\theta)\right)$. Prove that the formal Bayes rule is an HPD region.

(b) Let the loss function be $L_q(\theta, [a_1, a_2]) = (a_2 - a_1)^2 + c\left(1 - I_{[a_1,a_2]}(\theta)\right)$. Find the formal Bayes rule.

Section 5.3:

41. Suppose that one wished to construct a parametric bootstrap estimate in the second part of Example 5.80 on page 330.

(a) Explain how to construct the parametric bootstrap estimate using the $U(0, \theta)$ parametric family.

(b) Find the distribution of $R(X^*, \hat{F}_n)$ for the parametric bootstrap estimate.

(c) Will the parametric bootstrap estimate have the same problem that the nonparametric bootstrap estimate has?

42. How would one use the nonparametric bootstrap to find the bias and standard deviation for the sample correlation coefficient from a sample of n pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$?

43. Let $(x_1, Y_1), \ldots, (x_n, Y_n)$ be data pairs, and suppose that we entertain a regression model in which $\mathrm{E}(Y_i|\mathrm{B}_0 = \beta_0, \mathrm{B}_1 = \beta_1) = \beta_0 + \beta_1 x_i$. The y-intercept of the regression line is $x = -\beta_0/\beta_1$. Let $(\hat{\mathrm{B}}_0, \hat{\mathrm{B}}_1)$ be the usual least-squares regression estimator.

(a) How would one use the bootstrap to find the bias and standard deviation of the ratio $-\hat{\mathrm{B}}_0/\hat{\mathrm{B}}_1$?

(b) Suppose that one used the following formula for the approximate variance of the ratio of two random variables Z_0/Z_1:

$$\frac{\mathrm{Var}(Z_0)}{Z_1^2} + \frac{Z_0^2\mathrm{Var}(Z_1)}{Z_1^4} - 2\frac{Z_0\mathrm{Cov}(Z_0, Z_1)}{Z_1^3}.$$

Show how you would use this to find bootstrap confidence intervals for $-\mathrm{B}_0/\mathrm{B}_1$.

Aula 11

$$L(d,\theta) = -U(r(d,\theta))$$

↳ perda incorrida ao esc. $\theta \in D$ quando est. nat. é $\theta \in \Theta$.

Para $d \in D$,

$$\rho(d,\mathbb{P}) = \int_\Theta L(d,\theta)\,d\mathbb{P}(\theta) \Rightarrow \text{risco da decisão } d.$$

$$\rho^*(\mathbb{P}) = \inf\{\rho(d,\mathbb{P}) : d \in D\} \rightarrow \text{r. de Bayes.}$$

$d^* \in D$ é dec. de Bayes contra $\mathbb{P}$ se $\rho^*(d^*,\mathbb{P}) = \rho^*(\mathbb{P})$

Comentário

2. $D, \mathbb{P}, L$

Seja $\lambda : \Theta \to \mathbb{R}$ t.q. $\int_\Theta \lambda(\theta)\,d\mathbb{P}(\theta) \leq \infty$

$$L' : D \times \Theta \to \mathbb{R}$$
$$(d,\theta) \mapsto L'(d,\theta) = L(d,\theta) + \lambda(\theta)$$

$$\int_\Theta L'(d,\theta)\,d\mathbb{P} = \int_\Theta [L(d,\theta) + \lambda(\theta)]\,d\mathbb{P}(\theta)$$
$$= \int_\Theta L(d,\theta)\,d\mathbb{P}(\theta) + \int_\Theta \lambda(\theta)\,d\mathbb{P}(\theta$$

Para todo $a > 0$ e $\lambda : \Theta \to \mathbb{R}$ t.q. $\int_\Theta \lambda(\theta)\,d\mathbb{P}(\theta) < \infty$, temos $\forall d_1, d_2 \in D$, que

$$\int_\Theta L(d_1,\theta)\,d\mathbb{P}(\theta) \leq \int L(d_2,\theta)\,d\mathbb{P}(\theta) \Leftrightarrow$$

$$\int_\Theta \{\underbrace{a\,L(d_1,\theta) + \lambda(\theta)}_{L'(d_1,\theta)}\}\,d\mathbb{P}(\theta) \leq \int_\Theta \underbrace{a\,L(d_2,\theta) + \lambda(\theta)}_{L'(d_2,\theta)}\,d\mathbb{P}(\theta)$$

3. Fixado Θ, $\mathcal{F}$ e L (ñ negativa), seja $\mathcal{P}_L$ o conj. de todas as med. de prob. em $(\Theta, \mathcal{F})$.

$\mathcal{P}_L$ é convexo, pois $\forall\, \mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}_L$, $\alpha\mathbb{P}_1 + (1-\alpha)\mathbb{P}_2 \in \mathcal{P}_L$, $\alpha \in [0,1]$.

Vamos olhar a transf.

$$\rho^* : \mathcal{P}_L \to \mathbb{P}$$
$$\mathbb{P} \in \mathcal{P}_L \mapsto \rho^*(\mathbb{P}) = \inf\{\rho(d,\mathbb{P});\, d \in D\}$$

Resultado: ρ^* é convexa, i.e., $\rho^*(\alpha\mathbb{P}_1 + (1-\alpha)\mathbb{P}_2) \geq \alpha\rho^*(\mathbb{P}_1) + (1-\alpha)\rho^*(\mathbb{P}_2)$

Aula 12

Risco incorrido por A sob o ponto de vista de B:

$$\rho(d_A^*, \mathbb{P}_B) \quad , \text{ onde}$$

$d_A^* = \rho^*(\mathbb{P}_A)$. ~~Seg.~~ B, o risco adicional é

$$\rho(d_A^*, \mathbb{P}_B) - \rho^*(\mathbb{P}_B).$$

- Se $L(d, \Theta) = (d-\Theta)^2$, então

$$\rho(d, \mathbb{P}) = Var(\Theta / \mathbb{P}) + (d - \mathbb{E}(\Theta|\mathbb{P}))^2,$$

de modo que, $d^* = \mathbb{E}(\Theta/\mathbb{P})$ e,

$$\rho^*(\mathbb{P}) = Var(\Theta|\mathbb{P}).$$

$D = \{d_1, \dots\}$

Def. ~~Ao~~ proc. que cons. $d_i \in D$ com prob. α_i, $i \geq 1$, damos o nome de dec. aleat. ou mista assoc. à seq. $(\alpha_n)_{n \geq 1}$.

$$L(d_{(\alpha)}, \Theta) = \sum \alpha_n L(d_n, \Theta) \quad \text{"Lext. à } \mu_0 \times \Theta\text{"}$$

Aula 14.

d é admissível se

$$L(d, \Theta) \cancel{\geq} L(d', \Theta) \quad \forall \Theta \in \Theta \;, d' \in D \text{ e } d' \neq d \;\cancel{\ni}$$

$$\exists \Theta_0 \in D \text{ t.q. } L(d, \Theta_0) \leq L(d', \Theta_0) \quad \forall d' \neq d \in D.$$

- Se d é única dec. de Bayes, ela é admissível.

Problema de Decisão

$(D, \Theta, L, \mathbb{P})$

- D: decisões
- Θ: parâmetros
- L: perda
- $\mathbb{P}$: Prob.

$\forall d \in D$, calc

$$\rho(d, \mathbb{P}) = \int_\Theta L(d, \Theta)\, d\mathbb{P}$$

e tomo d^* t.q. $\rho(d^*, \mathbb{P}) = \inf_{d \in D} \rho(d, \mathbb{P})$.

Aula 14

Problema de Decisão com dados

- Pré - Experimentação

- Pós - Experimentação

- ~~Caso~~ Pós. Experimentação

$\underset{X=x}{\rightarrow}$ $(D, \Theta, L, \mathbb{P}_{\{X=x\}})$, onde $\mathbb{P}_{\{X=x\}}$ é a med. de prob. sob o todo $X=x$.

Para cada realização $X=x$, um problema de decisão distinto.
$X=x$ induz $\mathbb{P}_{\{X=x\}}$ distintas p/ x dist., o que leva à riscos dist.

- Caso Pré-Experimentação

Suponha que o exp. não foi conduzido e que X é também incerto (desconhecido).
Nesse casos há incerteza sobre (θ, x), $\theta \in \Theta$, x possível obs. do exp.
Além disso, devemos agora especificar uma regra (função) de decisão que associa a cada x uma decisão $d \in D$.
Se $\mathcal{X}$ o cjto dos pos. res. de exp. (No ex. 3 $\mathcal{X} = \{0,1\}$)

Temos um novo prob. de decisão.

$$\Delta : \{\delta : \mathcal{X} \to D\} \quad \text{cjto de regras de decisões}$$

$$\Omega : \Theta \times \mathcal{X}, \text{ munido de } \sigma(\Theta \times \mathcal{X})$$

$$\mathbb{P}' : \sigma(\Theta \times \mathcal{X}) \to [0,1] \text{ um prob.}$$

Como antes, $L : D \times \Theta \to \mathbb{R}$ e def.

$$L' : \Delta \times \Omega \to \mathbb{R} \text{ por}$$

P/ $\delta \in \Delta$ e $(\theta, x) \in \Omega$, $L'(\delta, (\theta, x)) = L(\delta(x), \theta)$.

$$(\Delta, \Omega, L', \mathbb{P}')$$

P/ cada $\delta \in \Delta$, $\bar{\rho}(\delta, \mathbb{P})$:

Pré. Exp.

$(\Delta, \Omega, L', \mathbb{P}')$

$\Delta = \{\delta : \mathfrak{X} \to D\}$ $\quad \mathbb{P}' : \sigma(\mathfrak{X} \times \Theta) \to [0,1]$

$\Omega = \mathfrak{X} \times \Theta$

$L'(\delta, (\theta, x)) = L(\delta(x), \theta)$

Para cada $\delta \in \Delta$,

$$\bar{\rho}(\delta, \mathbb{P}') = \int_{\underset{\Theta \times \mathfrak{X}}{\Omega}} L'(\delta, (\theta, x))\, d\mathbb{P}'(\theta, x)$$

risco da função decisão δ contra $\mathbb{P}'$

Do mesmo modo

$\bar{\rho}^*(\mathbb{P}') = \inf\{\bar{\rho}(\delta, \mathbb{P}') . \delta \in \Delta\}$ é chamado R. de Bayes

Aula 15

Pós:

Observar $X=x$, você resolve o prob. $(D, \Theta, L, \mathbb{P}_{X=x})$

Pré:

$\Delta = \{ \delta : \mathcal{X} \to D \}$ — funções regra de decisão

$\Omega = \Theta \times \mathcal{X}$

$\mathbb{P}'$ sobre $\sigma(\Theta \times \mathcal{X})$

$L' : \Delta \times \Omega \to \mathbb{R}$

$(\delta, (\theta, x)) \mapsto L'(\delta, (\theta, x)) = L(\delta(x), \theta)$

P/ cada $\delta \in \Delta$,

$\rho'(\delta, \mathbb{P}') = \int_{\Theta \times \mathcal{X}} L'(\delta, (\theta, x)) \, d\mathbb{P}'(\theta, x)$ → risco da f. decisão δ contra $\mathbb{P}'$

$\rho'^{*}(\mathbb{P}') = \inf \{ \rho'(\delta, \mathbb{P}'), \delta \in \Delta \}$ → Risco de Bayes

$\delta^{*} \in \Delta$ t.q. $\rho'(\delta^{*}, \mathbb{P}') = \rho'^{*}(\mathbb{P}')$ é regra de decisão de Bayes.

Prob. Det. Tamanho Am. e da Regra de Decisão

$\bar{\Omega} = \Theta \times \mathcal{X}$, $\mathcal{X} = \prod_{i=1}^{\infty} \mathcal{X}_i$: tudo que decorre sobre o qual há incerteza

$\Delta_n = \{ \delta : \prod_{i=1}^{n} \mathcal{X}_i \to D \}$

$\Delta_0 = D$

$\bar{\Delta} = \bigcup_{n=0}^{\infty} \{n\} \times \Delta_n$ → espaço de ações

$\bar{\mathbb{P}}$ sobre $\sigma(\Theta \times \mathcal{X})$, $\bar{L} : \bar{\Delta} \times \bar{\Omega} \to \mathbb{R}$

$((n, \delta^{(n)}), (\theta, x)) \to \bar{L}((n, \delta^{(n)}), (\theta, x))$

$= L(\delta^{(n)}(x), \theta) + \gamma(n, \delta, \theta, x)$

$L(\delta^{(n)}(x), \theta) + n c_0$, $c_0 > 0$

$$\bar{\rho}((n,\delta),\overline{\mathbb{P}}).$$

$$= \int_{\Theta\times\mathfrak{X}} \bar{L}((n,\delta),(\Theta,x))\, d\overline{\mathbb{P}}$$

$$= \int_{\Theta\times\mathfrak{X}} [L(\delta(x),\Theta) + n c_0]\, d\overline{\mathbb{P}}$$

$$= n c_0 + \int_{\Theta\mathfrak{X}} L(\delta^{(n)}(x),\Theta)\, d\mathbb{P}$$

$$\bar{\Omega} = \Theta \times \mathfrak{X}, \quad \mathfrak{X} = \prod_{i=1}^{n} \mathfrak{X}_i$$

$$\Delta_n = \{\delta : \prod_{i=1}^{n} \mathfrak{X}_i \to \mathcal{D}\}$$

$$\bar{\Delta} = \bigcup_{n=1}^{\infty} \{n\} \times \Delta_n$$

$$\bar{L} : \bar{\Delta} \times \Theta \to \mathbb{R}$$

$$((n,\delta),(\Theta,x)) \mapsto \bar{L}((n,\delta),(\Theta,x)) = L(\delta^{(n)}(x),\Theta) + C(n,\delta,\Theta,x)$$

$$n c_0.$$

$$\bar{\rho}((n,\delta),\overline{\mathbb{P}}) = \int_{\Theta\times\mathfrak{X}} \bar{L}((n,\delta),(\Theta,x))\, d\overline{\mathbb{P}}$$

$$= \int_{\Theta\times\mathfrak{X}} n c_0 + L(\delta^{(n)}(x),\Theta)\, d\overline{\mathbb{P}}$$

$$= n c_0 + \int_{\Theta\times\mathfrak{X}} L(\delta^{(n)}(x),\Theta)\, d\overline{\mathbb{P}}$$

Aula 16

com custos

$$\mathfrak{X} = \prod^{\infty} \mathfrak{X}_i$$

$$\bar{\Omega} = \Theta \times \mathfrak{X}$$

$$\Delta_n = \{ \delta : \mathfrak{X}_1 \times \cdots \times \mathfrak{X}_n \to D \} , \Delta_0 = D$$

$$\bar{\Delta} = \bigcup^{\infty} \{n\} \times \Delta_n$$

$$\overbrace{L((n, \delta), (\theta, x)) = L(\delta(x^{n}), \theta) + n c_0}^{\Gamma(n, \delta, \theta, x)} , c_0 > 0$$

Aula 17

Estimação

(D, Θ, L, P)

Após obs. $x \in \mathcal{X}$, $(D, \Theta, L, P_{X=x})$ com $D = \Theta$

Pro. Pré-Exp.

$$\Delta = \{\delta : \mathcal{X} \to D\}$$

No de est. $D = \Theta$, de modo qe

$$\Delta = \{\delta : \mathcal{X} \to \Theta\}$$ conj. dos estimadores

$(\Delta, \Theta \times \mathcal{X}, L', P')$

$$L' : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}_+$$
$$(\delta, (\theta, x)) \mapsto L'(\delta, (\theta, x)) = L(\delta(x), \theta)$$

Usualmente, L expressa a discrepância entre $\delta(x)$ e θ.

$$L(d, \theta) = h(\theta)\, g(dist(d, \theta))$$ g não decresc.

. Perda Quadrática $L(d, \theta) = (d - \theta)^2$

Se $E(\theta) \in \Theta \Rightarrow d^* = E(\theta)$
se $E(\theta) \notin \Theta \Rightarrow d^* \in D$ é t.q $|d^* - E(\theta)| = \min\{|d - E(\theta)| : d \in D\}$.

$\rho^*(P) = Var(\theta)$, se $E(\theta) \in \Theta$ e

$(d^* - E(\theta))^2 + Var(\theta)$ c.c.

Após a obs. de $X = x$,

$(\Theta, \Theta, L, P_{X=x})$

$d_x^* = E(\theta / X = x)$

$\rho^*(P_{X=x}) = Var(\theta / X)$

No prob. pré-exp.

$$\delta^* : \mathcal{X} \to \Theta$$
$$x \in \mathcal{X} \mapsto \delta^*(x) = \mathbb{E}(\Theta / X = x) \text{ i.e}$$

$\delta^*(x) = E(\Theta / X)$ - com rel. perda quadrática

Em geral, p/ $\delta \in \Delta$

$$\rho(\delta, P') = \int_{\Theta \times \mathcal{X}} L(\delta(x), \theta) \, dP'$$

$$= \int_{\mathcal{X}} f(x) \left[\int_{\Theta} L(\delta(x), \theta) \, f(\theta / x) \, d\theta \right] dx$$

Pai

$$\rho^*(\delta^*) = \int_{\mathcal{X}} f(x) \int_{\Theta} \underbrace{(\theta - E(\Theta / X = x))^2 f(\theta, x) \, d\theta}_{Var(\Theta / X = x)} \Big] \, dx$$

$$= \mathbb{E}\left(Var(\Theta / X)\right)$$

Testes de Hipóteses

Θ, $\mathcal{F}$ - σ-álgebra de Θ

Vamos tomar $\Theta_0, \Theta_1 \in \mathcal{F}$ e $\Theta = \Theta_0 \cup \Theta_1$, $\Theta_0 \cap \Theta_1 = \emptyset$

$$H_0 : \theta \in \Theta_0$$
$$H_1 : \theta \in \Theta_1$$

Neste caso

$$D = \{d_0, d_1\}$$

d_0: Ñ rej. H_0 ; d_1: rej.

	$\theta \in \Theta_0$	$\theta \in \Theta_1$
d_0	0	a_2
d_1	a_1	0

(mais simpl. f. de perda)

Obs. $X = x$ ($\{d_0, d_1\}$, Θ, L, $\mathbb{P}_{X=x}$)

$$\rho(d_0, \mathbb{P}_{X=x}) = \int_{\Theta} L(d_0, \theta) d\mathbb{P}_X = \int_{\Theta_1} a_2 \, d\mathbb{P}_X = a_2 \mathbb{P}_{X=x}(\theta \in \Theta_1) \quad \mathbb{P}(\theta \in \Theta_1 / x)$$

$$\rho(d_1, \mathbb{P}_{X=x}) = \int_{\Theta_0} a_1 \, d\mathbb{P}_x = a_1 \mathbb{P}(\theta \in \Theta_0 / x)$$

$$\rho(d_1, \mathbb{P}_x) < \rho(d_0, \mathbb{P}_x) \Leftrightarrow a_1 \mathbb{P}(\theta \in \Theta_0 / x) < a_2 \mathbb{P}(\theta \in \Theta_1 / x)$$

$$\Leftrightarrow a_1 \mathbb{P}(\theta \in \Theta_0 / x) < a_2 - a_2 \mathbb{P}(\theta \in \Theta_0 / x))$$

$$\Leftrightarrow \mathbb{P}(\theta \in \Theta_0 / x) < \frac{a_2}{a_1 + a_2}$$

$$\mathbb{P}(\theta \in \Theta_0 / x) < \frac{a_2}{a_1 + a_2}$$

No prob. Prev, a regra de dec. de Bayes (ou teste de Bayes) é dado

$$\varphi^*(x) = \begin{cases} 1, & P(\theta \in \Theta_1 | x) > a_1/a_1+a_2 \\ 0, & P(\theta \in \Theta_1 | x) < a_1/a_1+a_2 \end{cases}$$

Est. por intervalo

Θ

$D = \{[a,b], a, b \in \mathbb{R}, a < b\}$

$L([a,b], \theta) = b(b-a) + K \mathbb{I}_{[a,b]^c}(\theta)$

$D \quad \mathcal{J} = \sigma(\Theta)$ h. ñ fecn-

$L(A, \theta) = h_1(A) + c \mathbb{I}_{A^c}(\theta)$