Probabilidade e Inferência
Estatística II

Aula 01 - 15/08

Sergio Wechsler

email: swatgm@gmail.com

J. KADENE - Principle of Uncertainty

$? = \Theta$ $\qquad \Omega$. espaço paramétrico

X

$X_\theta \sim F_\theta \in \mathcal{F} = \{F_\theta(\cdot) : \theta \in \Omega\}$

$\uparrow$
θ um indice.

$X_1, X_2, \dots$ $\downarrow$ θ	$? = X_{n+1}$

$X_1, \dots, X_n \ \{0,1\}$

Teorema da Representação De Finetti: $P\{X_1 = x_1, \dots, X_n = x_n\} = \int_0^1 \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i}\, dF(\theta)$

Lema de Neymann - Pearson

$\Omega = \{\theta_0, \theta_1\}$

$f = \{f_0, f_1\}$ (densidades)

$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \end{cases}$

Lembrete: Teste $\varphi : \mathfrak{X} \to \{0,1\}$ $\quad$ 1: "rejeição"
$\qquad X \to \varphi(X)$ $\quad$ 0: "aceitação"

"melhor" teste
$\downarrow$
não existe

- tamanho: $\sup_{\theta \in \Omega_0} P\{\varphi(X) = 1 | \theta\} \leq \alpha$

Lema de Neyman - Pearson

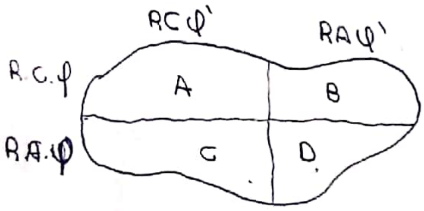
Considere o teste

$$\varphi(X) = \begin{cases} 1, & \text{quando } \overbrace{\dfrac{f_1(x)}{f_0(x)}}^{R.V} \geq k \\ 0, & \text{quando } \dfrac{f_1(x)}{f_0(x)} < k. \end{cases}$$

Então, $\forall \alpha \in (0,1)$

se $P\{\varphi(X) = 1 / \theta = \theta_0\} = \alpha$, então $P\{\varphi(X) = 0 / \theta = \theta_1\} \leq \text{Prob}\{\varphi'(X) = 0 / \theta = \theta_1\}$, para qualquer φ' com tamanho $\leq \alpha$.

"Prova" (qse sem palavras).



Queremos provar:

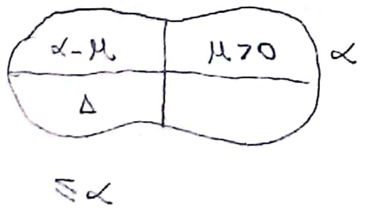
$P\{X \in C \cup D / \theta_1\} \leq \text{Prob}\{X \in B \cup D / \theta_1\} \Leftrightarrow$

$\text{Prob}\{X \in C / \theta_1\} \leq \text{Prob}\{X \in B / \theta_1\}$

$\Downarrow$

$$\int_C f_1(x)\,dx \leq k \int_C f_0(x)\,dx = k\Delta \leq k\mu = k \int_B f_0(x)\,dx \leq k \int_B \frac{1}{k} f_1(x)\,dx \quad \square$$

Se $\theta = \theta_0$



. se $\mu = 0 \Rightarrow \varphi' = \varphi$

$\alpha - \mu + \Delta \leq \alpha \Rightarrow \Delta \leq \mu$

→ Clássica

$\dfrac{\sup_{\theta_1} f_\theta(x)}{\sup_{\theta_0} f_\theta(x)}$

Bayes

$\dfrac{\int_{\theta_1} f_\theta(x) p_1(\theta)\,d\theta}{\int_{\theta_0} f_\theta(x) p_0(\theta)\,d\theta}$ →

FBST

chateados

. Changing

DOI. 10.100X/Sta.0000

Pericchi, Pereira

Prob. e Inf. Est. II

Aula 03 - 22/08

Teorema da Representação (var 0-1)

Parte I:

Sejam $X_1, X_2, \dots$ v.a.s com dist. P. Se P é permutável $\Leftrightarrow$

$\exists F$, função de distribuição em $[0,1]$ tal que $\forall n$, $\forall (x_1, \dots, x_n) \in \{0,1\}^n$,

$$P(X_1 = x_1, \dots, X_n = x_n) = \int_0^1 \boxed{\theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n - \sum_{i=1}^{n} x_i}} \, dF(\theta) =$$

$$= \begin{cases} \sum_i \boxed{\theta_i^{\sum x_i} (1-\theta_i)^{n-\sum x_i}} \, p(\theta_i) \\ \int_0^1 \boxed{\theta^{\sum x_i} (1-\theta)^{n-\sum x_i}} \, f(\theta) \, d\theta \end{cases}$$

Probabilidade e Inferência II

12/09 - Aula

$\mathcal{D} = \{d_0, d_1\}$

	θ_0	θ_1	$\max L(d,\theta)$
d_0	3	15	15
d_1	10	2	10

$$d^*_{minimax} = \arg\min_{d} \max L(d,\theta) \Rightarrow d^*_{minimax} = d_1$$

Na aula passada, fizeram pelas colunas. Erro. X

Supondo que θ_0 e θ_1 não são "estados da natureza",

$\downarrow \quad \downarrow$

decisões do jogador oposto

		θ^* θ_0	θ_1	
	d_0	3	15	15
d^*	d_1	10	2	10
		3	2	

Com a decisão minimax

$[3, 10]$

3: jog oposto ganha pelo menos 3

10: eu não perco mais que 10.

Mas posso ficar tentado a mudar minha jogada, visando perder só 3. O risco é se ele querer ganhar 15. Aí fudeu!

3 $\quad v \quad$ 10

v = valor do jogo obtido caso as decisões sejam aleatorizadas.

Teorema do Minimax:

$$d^*_{B,p} = \arg\min_{d} \int_{\Theta} L(d,\theta)\, f(\theta)\, d\theta$$

decisão de Bayes contra a priori p

Wald

<u>Teorema</u> (Aposta com incerteza)

preferência

Se	$\precsim \subset \mathcal{D}\times\mathcal{D}$	$\star \subset \Theta\times\Theta$	$\circledast$

$L(d,\theta)$ $\quad$ Loschi,s (2008)

, $\exists\, L : \mathcal{D}\times\Theta \to \mathbb{R}$, e $\exists$ uma m. prob. p, tais que $\forall i, \forall j$.

$$d_i \precsim d_j \Leftrightarrow \int_{\Theta} L(d_i,\theta)\,p(\theta)\,d\theta < \int_{\Theta} L(d_j,\theta)\,p(\theta)\,d\theta$$

preferível

Kadane, pg. 268-270
Escolher o melhor bilhete

Corolário.
d_i que tem a menor perda esperada.
risco

$$d^*_{B,p} = \arg\min_{d} \int_{\Theta} L(d,\theta)\,p(\theta)\,d\theta$$

Probabilidade e Inferência II

Rubin, H ... A weak system of axioms. 17/09

Para que

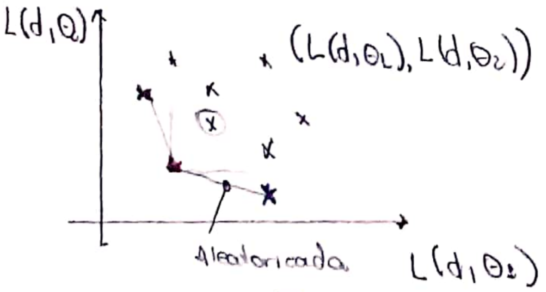
$$p(d) = \int_{\Theta} h(d,\theta)\,d\theta = \underbrace{\int_{\Theta} L(d,\theta)\,p(\theta)\,d\theta}_{\text{separatibilidade}}$$

temos a exigência de 3 axiomas (Ver Artigo ou Kadane).

•—•—∘

Consideremos agora,

$\Theta = \{\theta_1, \theta_2\}$



• Domina $\otimes$

$\Downarrow$

$\otimes$ é inadmissível

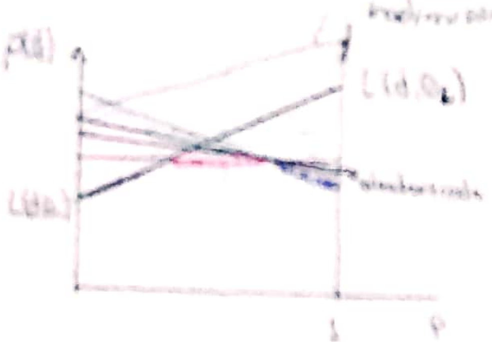
$\tilde{d}$ uma decisão aleatorizada

$$\tilde{d} = \begin{cases} d_1, & \alpha_1 \\ d_2, & \alpha_2 \\ \vdots & \\ d_k, & \alpha_k \end{cases}$$

$L(\tilde{d},\theta) = \sum_j \alpha_j L(d_j,\theta)$ (Paradoxo de Sant Petesburgo, Ver!)

Tudo isso no capítulo 8 do De Groot.

$\rho(d) = pL(d,\theta_1) + (1-p)L(d,\theta_2)$



Dependendo da priori, as 'retas' se alteram.

No entanto, sempre haverá

(Sob condições de regularidade)

Teorema: Uma decisão é admissível

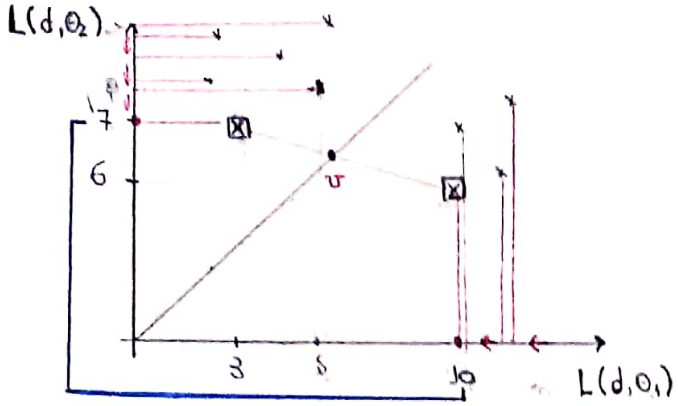
$\Leftrightarrow$

é decisão de Bayes contra alguma priori p.

$$\left(D, L, \Theta, \overset{\mathcal{X}}{\mathcal{X}}, \delta, \{F_\theta(\cdot/\theta), \theta \in \Theta\}, p\right)$$

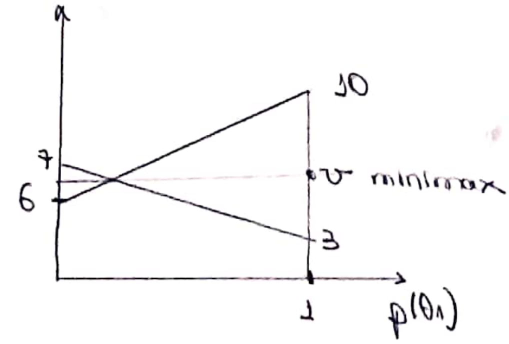
regra de decisão

19/09
H. Rubin

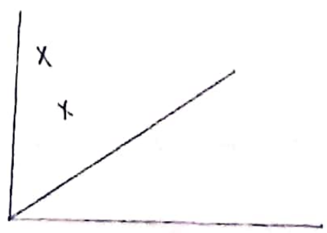


	θ_1	θ_2	
→ d_1	3	7	7
d_2	10	6	10
	3	6	

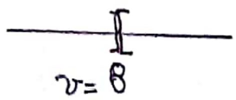
Quando dois seres estão envolvidos, as decisões podem ser alteradas ao longo das rodadas, o que dificulta a tomada de decisão.

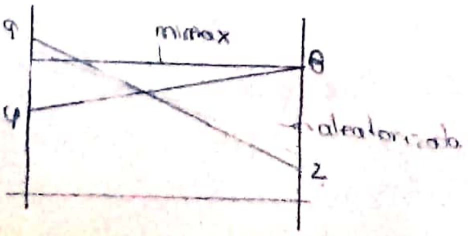


Outro caso:



2	9	9
4	8	8
2	8	





Chapter 8, M. DeGroot
Optimal Statistical Decision

$$L: \begin{array}{l} \mathcal{D} \times \Theta \to \mathbb{R} \\ (d,\theta) \mapsto L(d,\theta) \end{array}$$

$(\mathcal{D}, L, \Theta)_P$ $\mathcal{X}$ $X \sim F$ em $\mathcal{F} = \{F_X(\cdot|\theta) : \theta \in \Theta\}$ X e θ não são ind. $X \in \mathcal{X}$

DEF: Regra de Decisão $\underline{L} = L(\delta(X), \theta)$

$$\delta: \begin{array}{l} \mathcal{X} \longrightarrow \mathcal{D} \\ x \longmapsto \delta(x) \end{array}$$

$$\delta \in \Delta \rightsquigarrow (\Delta, \overset{L(\delta(x),\theta)}{\underline{L}(\delta,\theta)}, \mathcal{X} \times \Theta)$$

$$\delta^*_B = \arg\min_{\Delta} \mathbb{E}\{L(\delta(X),\theta)\} =$$

$$= \arg\min_{\Delta} \iint_{\Theta \times \mathcal{X}} L(\delta(x),\theta) p(x,\theta)\, dx\, d\theta$$

Teorema: δ^*_B é $\quad \delta^*: \begin{array}{l} \mathcal{X} \to \mathcal{D} \\ X \mapsto \delta(X) = d^*_{p(\theta|X)} \end{array}$

Probabilidade e Inferência Estatística II

24/09

Problema "original": (sem dados)

$(\mathcal{D}, L, \Theta, p)$

$L: \mathcal{D} \times \Theta \to \mathbb{R}$

Lema :

① Se $L'(d,\theta) = L(d,\theta) - \inf\limits_{d \in D} L(d,\theta)$, $\forall d \in \mathcal{D}$, então: para $\forall \theta \in \Theta$

(a) $L'(d,\theta) \geq 0$, $\forall d \in \mathcal{D}$, $\forall \theta \in \Theta$

(b) $\inf\limits_{d \in \mathcal{D}} L'(d,\theta) = 0$, $\forall \theta \in \Theta$

	θ_1	θ_2
d_1	3	7
d_2	5	-2

0	9
2	0

(c) $\forall$ priori em (partes de) Θ, $\forall d_1 \in \mathcal{D}$, $\forall d_2 \in \mathcal{D}$

$$\rho(d_1) < \rho(d_2) \Leftrightarrow \rho'(d_1) < \rho'(d_2)$$

Prova: $\Rightarrow$ $\rho(d_1) = \int_\Theta \left[L(d_1,\theta)\, p(\theta)\, d\theta < \int_\Theta \left[L(d_2,\theta)\, p(\theta)\, d\theta \right.\right.$
$\qquad -\inf L(d_1,\theta) \big] \qquad\qquad -\inf L(d_2,\theta) \big]$

Supondo que $\left| \int_\Theta \inf L(d,\theta)\, p(\theta)\, d\theta \right| < \infty$, $\left| \int_\Theta L(d,\theta)\, p(\theta)\, d\theta \right| < \infty$

temos: $\rho'(d_1) < \rho'(d_2)$. $\square$

Em geral, se $L' = a(\theta) + b\,L(d,\theta)$, com $b > 0$, $\left[\int_\Theta a(\theta)\, p(\theta)\, d\theta \right] < \infty$.

$\Rightarrow \rho(d) \sim \rho'(d)$

② Seja $d_\alpha = \alpha d_1 + (1-\alpha) d_2$, isto é,

$d_\alpha = d_1$, ... com prob α

$\qquad d_2$ ... com prob. $1-\alpha$.

Então, $p(d_\alpha) \geq \min(p(d_1), p(d_2))$
(qqer que seja p)

Prova: Para p fixa, e supondo que $p(d_1) \leq p(d_2)$.

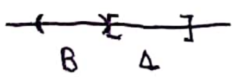
$$p(d_1) = p(\alpha d_1 + (1-\alpha) d_1) \leq \alpha p(d_1) + (1-\alpha) p(d_1) \leq$$

$$\leq \underbrace{p(\alpha d_1 + (1-\alpha) d_2)}_{p(d_\alpha)} \leq p(\alpha d_2 + (1-\alpha) d_2) = p(d_2)$$

E dessa forma, $p(d_1) \geq p(d_2)$.

El - Dash, N.

Felipe, 2. Desempate técnico.
Chance de gol.



. Nesse caso, $D($A ganhar $|$ Dados$) \to (1-\alpha)$

$X \in \mathfrak{X}$

$\mathcal{F} = \{ F_X(\cdot / \theta) : \theta \in \Theta \}$

Definição: Regra de decisão

$$\delta : \mathfrak{X} \to \mathcal{D}$$
$$x \mapsto \delta(x)$$

$\Delta = \{\delta, \delta$ é regra de decisão$\}$

$(\Delta, \underline{L}, (\Theta, \mathfrak{X}), p(\theta, x))$

[redacted] $\underline{L}(\delta, (\theta, x)) = L(\delta(x), \theta)$.

Teorema: $\delta_p^*(X) = d^*_{p(\cdot/X)}$

Prova: (Supondo que as v.a.s de integ. possam ser permutáveis)

$$\int_{\mathfrak{X}}\int_{\Theta} L(d(x),\theta)\, p(x)\, \frac{p(\theta)\,p(x/\theta)}{p(x)}\, d\theta\, dx =$$

$$= \int_{\mathfrak{X}} p(x) \int_{\Theta} L(\delta(x),\theta)\, p(\theta/x)\, d\theta\, dx = E\{\rho_{p(\cdot/X)}(\delta(X))\} \geq E\{\rho(d^*(X))\}$$

Probabilidade e Inf. Estatística II

03/10

$\mathcal{D} = \{d_0, d_1\}$ $\quad \Theta = \{\theta_0, \theta_1\}$

$H_0: \theta = -1 \overset{N}{=} \theta_0$

$H_1: \theta = +1 \overset{N}{=} \theta_1$

$X = (X_1, \dots, X_n)$ v.a.s i.i.d. $N(\theta, 9)$

c.i.i.d, dado θ, $N(\theta, 9)$

	θ_0	θ_1
d_0	0	a
d_1	b	0

$P(H_0) = 1 - P(H_1)$

$P(\theta_0) = 1 - P(\theta_1)$

δ_B

$\rho(d_0) = a\, P(H_1)$

$\rho(d_1) = b\, P(H_0)$

$$d^* = \begin{cases} d_0 & \text{se } \dfrac{P(H_1/x)}{P(H_0/x)} \leq \dfrac{b}{a} \\[2ex] d_1, & \text{se } \dfrac{P(H_1/x)}{P(H_0/x)} \geq \dfrac{b}{a} \end{cases}$$

Mas

$$\frac{P(H_1/x)}{P(H_0/x)} = \frac{P(H_1)\, f(x/\theta_1)/f(x)}{P(H_0)\, f(x/\theta_0)/f(x)} = \frac{P(H_1)}{P(H_0)} \underbrace{R(x)}_{\text{R.V.}}$$

logo,

$$\delta(X) = 1 \quad \dots \quad \text{se } R(x) \geq \frac{b}{a} \frac{P(H_0)}{P(H_1)}$$

e

$$R(x) = \frac{\prod_{i=1}^{n} \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{(x_i-\theta_1)^2}{2\sigma_0^2}}}{\prod_{i=1}^{n} \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{(x_i-\theta_0)^2}{2\sigma_0^2}}} = \frac{e^{-\sum_{i=1}^{n}\frac{(x_i-\theta_1)^2}{2\sigma_0^2}}}{e^{-\sum_{i=1}^{n}\frac{(x_i-\theta_0)^2}{2\sigma_0^2}}} = \frac{e^{-\frac{\sum x_i^2}{2\sigma_0^2}}}{e^{-\frac{\sum x_i^2}{2\sigma_0^2}}} \frac{e^{\frac{2\theta_1}{\sigma_0^2}\sum_{i=1}^{n} x_i}}{e^{\frac{2\theta_0}{\sigma_0^2}\sum x_i}} \frac{e^{-\frac{n\theta_1^2}{2\sigma_0^2}}}{e^{-\frac{n\theta_0^2}{2\sigma_0^2}}} =$$

$$= e^{\frac{1}{\sigma_0^2}\left((\theta_1-\theta_0)\sum_{i=1}^{n} x_i - \frac{n}{2}(\theta_1^2-\theta_0^2)\right)}.$$

$$\Rightarrow R(x) \geq \frac{b}{a} \frac{P(H_0)}{P(H_1)} \Leftrightarrow$$

$$\frac{1}{\sigma_0^2}\left((\theta_1-\theta_0)\sum x_i - \frac{n}{2}(\theta_1^2-\theta_0)^2 \right) \geq \ln\left(\frac{b}{a}\frac{P(H_0)}{P(H_1)}\right) \Leftrightarrow$$

$$\sum X_i - \frac{n}{2}(\theta_0+\theta_1) \geq \frac{\sigma_0^2}{(\theta_1-\theta_0)} \ln\left(\frac{b}{a}\frac{P(H_0)}{P(H_1)}\right) \Leftrightarrow$$

$$\bar{X} \geq \frac{(\theta_0+\theta_1)}{2} + \frac{1}{n}\frac{\sigma_0^2}{(\theta_1-\theta_0)} \ln\left(\frac{b}{a}\frac{P(H_0)}{P(H_1)}\right)$$

No exemplo,

$\theta_0 = -1$, $\theta_1 = +1$, $\sigma_0^2 = 9$

Seja $P(H_0) = P(H_1) = 1/2$

Se $a = b$ e $P(H_0) = P(H_1)$, $\delta_B = 1$ se $\bar{X} \geq \left(\frac{\theta_0+\theta_1}{2}\right)$.

Se, $P(H_0) = P(H_1) = 1/2$

$$\delta_B = 1 \quad \text{se } \bar{X} \geq \frac{(\theta_0+\theta_1)}{2} + \frac{1}{n}\left\{ \frac{\sigma_0^2}{(\theta_1-\theta_0)} \ln\left(\frac{b}{a}\right)\right\}$$

Sendo

$n = 20$

$\sigma_0^2 = 9$

$$\bar{X} \geq \frac{1}{20}\left\{\frac{9}{2}\ln\left(\frac{b}{a}\right)\right\} = \frac{9}{40}\ln\left(\frac{b}{a}\right) \approx \frac{1}{4}\ln\left(\frac{b}{a}\right) = \frac{1}{4}\ln\left(\frac{1}{0{,}63}\right)$$

$\frac{b}{a} = (0{,}63)^{-1}$ $\Rightarrow \alpha = 0{,}05$ e $\frac{\alpha}{\beta} = 0{,}55$

$\downarrow$

$n = 100$ e $\alpha = 0{,}05$ $\Rightarrow \frac{\alpha}{\beta} = 195217$

n	$\theta_0 = 1$, $\theta_1 = -1$ $a = b = 1$, $P(H_0) = 1/2$ BAYES	$\alpha = 5\%$ N. Pearson	α/β	Bayes $a/b = 0,63$
20	0	0.10	0,55	0,24
⋮	⋮	⋮	⋮	⋮
100	0	-0.51	200 000	0,049
		↓	↓	↓
		-1	$+\infty$	Isso que o artigo irá mostrar!

15/10

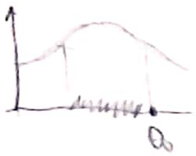
$H_0: \theta \in \Theta_0$ $\quad \dim \Theta_0 < \dim \Theta_1$ (hipótese precisa)

$H_1: \theta \in \Theta_1$

$f(\theta | x)$

$\sup\limits_{\theta \in \Theta_0} f(\theta | x)$

$T_x = \{ \theta \in \Theta : f(\theta|x) > \sup\limits_{\Theta_0} f(\theta | x) \}$



Começou o Café!

$EV_x(H_0) = 1 - P_r\{T_x | x\}$

$$\begin{cases} L(Ac\,H_0, \theta, x) = b + c\, I(\theta \in T_x) \\ L(Rej\,H_0, \theta, x) = a(1 - I(\theta \in T_x)) \end{cases}$$

$\Downarrow$

Rej $H_0 \Leftrightarrow EV < \frac{b+c}{a+c}$

Introdução às Probabilidade e Inferência II

17-10 - Aula 14

$$\_\ \_\ \_\ \_\ \_\ \_\ \_\ \_\ \_\ \_\ \_\ \frac{F}{S}$$

$\binom{11}{2}$ ou $\binom{11}{8}$

$\frac{11!}{2!\,9!}$ $\quad$ $\frac{11!}{3!\,8!}$

$N \sim$ Binomial Negativa $(k=9, p)$

$k=9,\ n=12$

$\underbrace{\underline{C}\ \ C\ \ \bar{C}\ \ \ldots}_{\substack{8C \\ 3\bar{C}}}\ \ \frac{C}{12}$

$P(N=12) \stackrel{\text{c}}{=} p\,P(R=8)$ $\qquad$ $R \sim$ Binomial $(8, p)$

$H_0\ \ p=p_0$

$H_1\ \ p>p_0$

$P(X \geq 9 / p_0)$ $\qquad$ $P(N \leq 12 / p_0)$

$$\sum_{x=9}^{12} P(X=x/p_0) \stackrel{\text{v}}{=} \sum_{n=9}^{12} P(N=n/p_0)$$

$X_1, \ldots, X_{100}$ i.i.d. $N(\theta, 1)$

$H_0: \theta = 0$ $\qquad$ $\alpha = 5\%$

$H_1: \theta \neq 0$

Sob H_0: $\sqrt{n}\,\bar{X} \sim N(0,1)$ $\quad$ Rej. $H_0 \leftrightarrow |\sqrt{n}\,\bar{X}| > 1{,}96$

$|\sqrt{100}\,\bar{x}| = 2{,}0 > 1{,}96 \ \ldots\ \circledast$

Abook

$|\sqrt{n}\,\bar{X}| > 2.10$

$X_1, \ldots, X_{100}$

$Y_1, \ldots, Y_{100}$

α/n

$\alpha \cdot 1/n \cdot 1$

P.S. $\Leftrightarrow$ PV   P = princípio   V: verossimilhança

P.C.   S: suficiência

Birbaum (1962)   C: condicionalidade

$E_1 = (x, \Theta, f_1) \quad x \quad L_x(\theta)$ se $L_x(\theta) \propto L_y(\theta) \Rightarrow Inf_x(\theta) = Inf_y(\theta)$ V. geral

$E_2 = (y, \Theta, f_2) \quad y \quad L_y(\theta)$

Paulo Marques ; Birbaun's Theorem Redux: "Afirma que resolveu o problema de Lekon".

Introdução À Probabilidade e Inferência Estatística II

22 de outubro, Aula 11

Permutabilidade

Def: Dizemos que X e Y são "permutáveis" quando $(X,Y) \sim (Y,X)$

Obs:

(1) A dist. de (X,Y) é que é "permutável"

(2)

(3) "Lembra" IID

I.I.D. $\Rightarrow$ permutáveis

$\not\Leftarrow$ "Exemplo das bolas brancas"

I.D. $\Leftarrow$ permutáveis

Permutabilidade

$$P(X=x) = \sum_y P(X=x, Y=y) = \sum_y P(X=y, Y=x) = P(Y=x)$$

Definição: Dizemos que $(X_1, \dots, X_n)$ são permutáveis quando $(X_1, \dots, X_n) \sim (X_{\tau_1}, \dots, X_{\tau_n})$, qualquer que seja a permutação $(\tau_1, \tau_2, \dots, \tau_n)$ de $(1, 2, \dots, n)$

Obs (1):

Definição: Dizemos que $(X_1, X_2, \dots)$ são permutáveis quando $(X_1, X_2, \dots, X_n)$ são permutáveis, qualquer que seja $n = 1, 2, \dots$

Fato: Se $X_1, X_2, \dots$ forem permutáveis, então, (por exemplo) $(X_1, X_2, X_3) \sim (X_{100}, X_{80}, X_{91})$

Idéia: Inferência Estocástica Palavra: "Faça!"

Def. Dizemos que v.a.s permutáveis $(X_1, X_2, \dots, X_n)$ são "extendíveis" quando $\exists$ um processo permutável $(Y_1, Y_2, \dots)$ t.q.

$$(Y_1, \dots, Y_n) \sim (X_1, \dots, X_n)$$

Teorema da Representação

(I)

Seja $(X_1, X_2, \dots)$ em $\{0,1\}^\infty$ com medida P.

P é permutável $\Leftrightarrow \exists F$, função de dist. com $F(0-) = 0$, $F(1) = 1$, tal que $\forall n \geq 1$; $\forall x_i \in \{0,1\}$, temos

$\forall i = 1, 2, \dots, n$

$$P(X_1 = x_1, \dots, X_n = x_n) = \int_0^1 \theta^{\sum_1^n x_i} (1-\theta)^{n - \sum_1^n x_i} \, dF(\theta)$$

$$P(X_1 = x_1, \dots, X_n = x_n) = \theta^{\sum x_i} (1-\theta)^{n - \sum_1^n x_i}$$, (caso particular, cons. F degenerada).

(II) Se P é permutável, então

$$P\left\{ \frac{\sum^n X_i}{n} \rightarrow \right\} = 1$$

Johnson & Kotz

Breakthoughs in Statistics. (2 volumes)

(1987) Kyburg. Smokler. Foundation of Statistics

$$\lim_{n \to \infty} \frac{\sum^n X_i}{n} \stackrel{D}{=} \Theta$$

(III) $\Theta \sim F_P$

Stat

The ISI's Journal for the Rapid

Dissemination of Statistics Research

(wileyonlinelibrary.com) DOI: 10.100X/sta.0000

Changing the paradigm of fixed significance levels: Testing Hypothesis by Minimizing Sum of Errors Type I and Type II

Luis Pericchi[a*], Carlos Pereira[b]

Received 00 Month 2012; Accepted 00 Month 2012

apresentar objetivando

mais adequado

Our purpose, is to put forward a change in the paradigm of testing by generalizing a very natural idea exposed by Morris DeGroot (1975) aiming to an approach that is attractive to all schools of statistics, in a procedure better suited for the needs of science. DeGroot's seminal idea is to base testing statistical hypothesis on minimizing the weighted sum of type I plus type II error instead of of the prevailing paradigm which is fixing type I error and minimizing type II error. DeGroot's result is that in simple vs simple hypothesis the optimal criterion is to reject, according to the likelihood ratio as the evidence (ordering) statistics using a fixed threshold value, instead of a fixed tail probability. By defining expected type I and type II errors, we generalize DeGroot's approach and find that the optimal region is defined by the ratio of evidences, that is, averaged likelihoods (with respect to a prior measure) and a threshold fixed. This approach yields an optimal theory in complete generality, which the Classical Theory of Testing does not. This can be seen as a Bayes-Non-Bayes compromise: the criteria (weighted sum of type I and type II errors) is Frequentist, but the test criterion is the ratio of marginalized likelihood, which is Bayesian. We give arguments, to push the theory still further, so that the weighting measures (priors)of the likelihoods does not have to be proper and highly informative, but just predictively matched, that is that predictively matched priors, give rise to the same evidence (marginal likelihoods) using minimal (smallest) training samples.

The theory that emerges, similar to the theories based on Objective Bayes approaches, is a powerful response to criticisms of the prevailing approach of hypothesis testing, see for example Ioannidis (2005) and Siegfried (2010) among many others.

Keywords: Hypothesis Testing; Significance Levels; Errors Type I and Type II

[a]Department of Mathematics and Center for Biostatistics and Bioinformatics, University of Puerto Rico, Rio Piedras, San Juan, Puerto Rico.

[b]Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, SP, Brasil.

*Email: luis.pericchi@upr.edu

Seja $(X_1, X_2, \ldots)$ um processo estocástico em $\{0,1\}^\infty$ com distribuição $\mathbb{P}$

[illegible]

(I) $\mathbb{P}$ é permutável se, e somente se, $\exists\, F$ com $F(0)=0$, $F(1)=1$, tal que $\forall n$, $\forall (x_1, \ldots, x_n) \in \{0,1\}^n$,

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \underbrace{\int_0^1 \theta^{\sum x_i} (1-\theta)^{n - \sum x_i} \, dF(\theta)}_{\text{Representação paramétrica do P. estocástico}}$$

(II) Se $\mathbb{P}$ é permutável, temos

$$\mathbb{P}\left\{ \frac{\sum X_i}{n} \to \right\} = 1$$

(III) $\lim_{n \to \infty} \frac{\sum^n X_i}{n} \sim F$ (Exemplo: Se estamos na urna de Polya, temos uma beta. Se fosse i.i.d., teríamos a conv. para uma cte.)

Seja $(X_1, X_2, \ldots)$ em $\mathfrak{X}^\infty$ com dist. $\mathbb{P}$

(I) $\mathbb{P}$ é permutável se, e somente se,

$\forall n$, $\forall (x_1, \ldots, x_n)$,

$\exists\, \mu$, medida de prob. sobre $\mathfrak{F} = \{G \mid \text{dist. em } \mathfrak{X}\}$ tal que

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \int_{\mathfrak{F}} \prod_{i=1}^{n} G(x_i) \, d\mu(G)$$

Probabilidade e Inferência II

12 de Novembro de 2013

- Qualidade de Previsões Probabilísticas

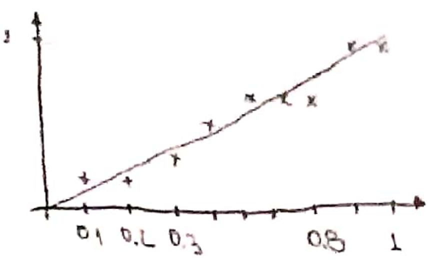
(1) Curvas de Calibração

P_1

P_2

P_N

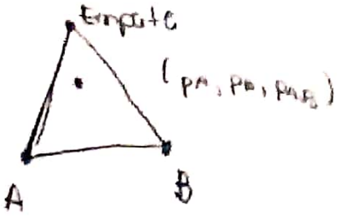
N grande



"Bem calibrado" quando a freq relativa de ocorrências é equivalente a proporção previamente prevista para aqueles dias

(2) Regra de BRIER

quanto menor a distância do ponto ocorrido



um pto no simplex quanto menor a dist. do ponto ao vértice ocorrido.

"distância ao quadrado da previsão ao vértice que pode ter acontecido

(3) $(1/3, 1/3, 1/3)$

→ Meteorologia. Texto "Correg Melo"

Phelip Dawid. Prequencial Inference

(Previsões Sequenciais

Previsão P_1, P_2, P_3 . . . P_{n+1}

Observação x_1 x_2 x_3

Comparação!

(p_1, x_1)

(p_2, x_2)

Data Analysis.

"Não faz sentido prob. de empate ser maior que a prob do pior time ganhar"

A x B

Faz sentido

A x B

Não faz sentido

deveria ser

A x B

(Caso A melhor que B)

ou

A x B

(Caso B melhor que A.)

• Continuação demonstração Teorema Da Representação De Finetti (Caso Geral)

II) $\mathbb{P}\left\{ F_n \rightarrow \right\} = 1$, com $F_1, F_2, \ldots$ seq. de fções de distribuições empíricas

III) $\lim_{n \to \infty} F_n \sim \mu$, com

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ ) = \int_{\mathcal{F}} \prod_{i=1}^{n} F(x_i) \, d\mu(F)$$

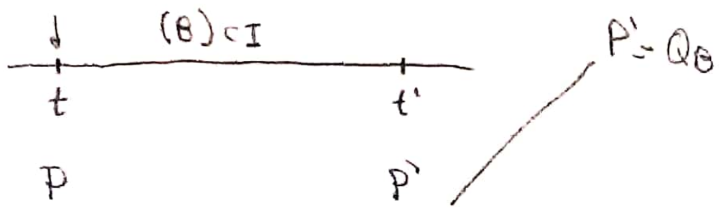
Em (II) se $\mathbb{P}\{F_n \rightarrow F\}$, temos a LGN (no caso ... teorema de Glivenko Cantelli.

$(\Omega, \mathcal{F}, P)$ $\quad B \in \mathcal{F},\ P(B) > 0$ $\quad$ 19 de Novembro

$$\underset{Q_B(A)}{\underline{P(A|B)}} = \frac{P(A \cap B)}{P(B)}, \quad \forall A \in \mathcal{F}$$

Teorema.

$\underset{Q_B(\cdot)}{\underline{P(\cdot|B)}}$ é uma medida de probabilidade em $\mathcal{F}$. $\quad Q(B) = 1$



Q_B

$\to Q_C$

$\to Q_D$

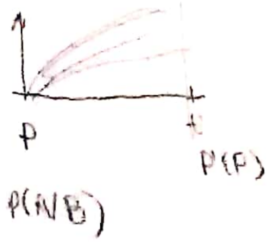
<u>Regra de Jeffreys</u>

$$P(A) \overset{?}{=} \sum_j P'(E_j)\, P(A|E_j)$$

Extremo
- $P'(E_j) = P(E_j)$ (caso de Bayes)
- $P'(A|E_j) = P(A|E_j)$

<u>Regra do Supercondicionamento</u> (*)

$P' = P'(P)$



Michael Goldstein

Gibbs for Kids

19 de novembro de 2013

$t \longrightarrow t^* \longrightarrow$ tempo

t: Coerência (bayesiana) $\Rightarrow$ Lab. $\Rightarrow P(\cdot \mid B)$

Receitas

1) Condicionamento bayesiano

$$P^*(A) = P(A/B), \forall A \in \mathcal{F}$$

2) Regra de Jeffrey

3) Super Condicionamento (Diaconis, Zabell)

4) Condicionamento Geral

(?) ...

Def: Coerência temporal (talvez impossível)

$\Updownarrow$

Coerência Coletiva

$\downarrow$

T. da Impossibilidade Arrow

- Problemas
- Receitas ⟵
- Soluções

Comparações

Exemplo 1. Duas moedas

$\Omega = \{CC, C\bar{C}, \bar{C}C, \bar{C}\bar{C}\}$

$\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4}$

$P^*(C_2) = 1 - P^*(C_2) = C.B. \quad P(C_2 / X = x)$

$$= \frac{P(C_2) P(X = x / C_2)}{P(X = x)}$$

x = Moeda caiu em pé

$P(X = x) = \ldots$

Exemplo 2. Roubaram a moeda

$\Omega = \{CC, \quad \bar{C}\bar{C}\}$

x = Roubaram a moeda

$$P^*(C_2) = \frac{P(C_2) P(X = x / C_2)}{P(X = x)}$$

C.B.

Não definido outra vez!

Exemplo 3. Pigmentos (Jeffrey)

A		B	
Green	0.3	Green	7/35
White	0.3	White	16/35
Blue	0.4	Blue	[illegible]/35

relâmpago

Não há nenhuma formula de Bayes que liga B à A. B não é obtida de A por cond. em algum evento

$\forall F \subset \Omega, \; P(\cdot | F) \neq P^*(\cdot)$.

A regra de Jeffrey resolve o problema do Exemplo 3.

19 de novembro

• Exemplo 4

Def Condicionamento bayesiano.

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

$$P^*(A) = P(A|B), \; \forall A \in \mathcal{F}$$

Def Equivalente

$$P^*(B) = 1 \iff P^*(\cdot/B) = P(\cdot/B)$$
$$\forall A \in \mathcal{F}$$

$(\Rightarrow)$

$$P^*(B) = P(B|B) = 1$$

$$P^*(A/B) = \frac{P^*(A)\,P^*(B/A)}{P^*(B)} = \frac{P(A/B)\,P^*(A\cap B)}{P^*(A)} = \frac{P(A/B)\,P(A\cap B/B)}{P(A/B)} =$$

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

$(\Leftarrow)$

$$P^*(A) = P^*(B)\,P^*(A/B) + P^*(\bar{B})$$
$$= P^*(A/B) = P(A/B).$$

Regra de Jeffrey

Partição de Ω (Discreto)

B B_1

B

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

$$P^*(A) = \sum_i P(A/B_i)\,P^*(B_i)$$

Se $P^*_j(B_j) = 1 \rightsquigarrow$ C.B.

$P^*(\ )$ é m. prob.

$$P^*_{J,B}(\Omega) = \sum P(\Omega / B_i) P(B_i) = 1$$

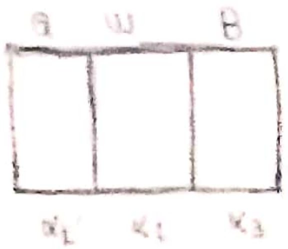
[illegible] (Σ [illegible])

Voltando ao exemplo do pijama

G $0.3 \rightarrow 7/85$

W $0.3 \rightarrow 16/85$

B $0.4 \rightarrow 12/85$



$\frac{7}{85} = \alpha_1 P(G/G)$

$\frac{16}{85} = \alpha_2 P(W/W)$

$\frac{12}{85} = \alpha_3 P(B/B)$

26 de Novembro de 2018

Atualização de Probabilidade $(\Omega, \mathcal{F}, P)$

(1) C. Bayesiano $(\Omega, \mathcal{F}, P^*)$, $P(E) > 0$, $E \in \mathcal{F}$

$$P^*(A) = P(A|E), \forall A \in \mathcal{F}$$

<u>Teorema</u>: $P^*(A) = P(A|B) \Leftrightarrow$

$$P^*(B) = 1 \wedge P^*(A|B) = P(A|B) \; \forall A \in \mathcal{F}$$ ✓

(2) <u>Regra de Jeffrey</u>

$(\Omega, \mathcal{G}, P^*)$ $\begin{cases} \mathcal{G} \in \mathcal{F} \\ (E_i), \text{ partição de } \Omega \\ P^*(E_i) \geq 0, \sum_i P^*(E_i) = 1 \end{cases}$

$$P^*(A) = \sum_i P^*(E_i) P(A|E_i), \forall A \in \mathcal{G}$$

1. $P(A|E_i) = 0$ qdo $P(E_i) = 0$, $\forall A \in \mathcal{G}, A \neq \Omega$

$P(\Omega|E_i) = 1$ qdo $P(E_i) = 0$

- $P^*(\Omega) = 1$

- $P^*(\Omega) = \sum_i P^*(E_i) \overset{1}{P(\Omega|E_i)} = \sum_i P^*(E_i) = 1$

- $P^*(A \cup B) = P^*(A) + P^*(B)$, se $A \cap B = \emptyset$
- $P^*(A) \geq 0$ ✓

<u>Teorema</u> $P^*(A) = \sum_i P^*(E_i) P(A|E_i), \forall A \in \mathcal{G}$

$$\Leftrightarrow$$

$$P^*(A|E_i) = P(A|E_i), \forall i, \forall A \in \mathcal{G}$$

Prova $(\Rightarrow)$

$$P^*(A \mid E_i) = \frac{P^*(A \cap E_i)}{P^*(E_i)} = \frac{\sum_j P^*(E_j)\, P(A \cap E_i \mid E_j)}{P^*(E_i)} =$$

$$= \frac{P(A \cap E_i \mid E_i)\, P^*(E_i)}{P^*(E_i)} = P(A \mid E_i)$$

$(\Leftarrow)$

$$P^*(A) = \sum_i P^*(E_i)\, P(A \mid E_i) = \dots$$

Caso Particular.

- Se $P^*(E_j) = 1$,

$$P^*(A) = P(A \mid E_j) \text{, i.e., C. Bayesiano}$$

- $P^*(E_i) = P(E_i), \forall i$

$$P^*(A) = P(A), \forall A \in \mathcal{F}$$

- Se não for satisf a cond J,

$$P^*(A) = \sum_i P^*(E_i)\, P^*(A \mid E_i) \rightarrow \text{carga todo conceito do zero}$$

Exemplos: Caso

$P(A \mid E_i) = 1$ fenómeno quando os E_i são os átomos de $(\Omega, \mathcal{F}, P)$

Exemplo Agora

G	3	7/35
B	4	16/35
W	3	12/35

E_1	E_2
G	$W \cup B$
0.2	0.8

$P^*(G) = 0.2 \cdot 1 + 0.8 \cdot 0 = 7/35$

$P^*(B) = 0.2 \cdot 0 + 0.8 \cdot \frac{16}{28} = \frac{16}{35}$

No condic. Bayesiano, um evento B aconteceu Dai, $P(B)=1$

- A regra de Jeffrey considera uma atribuição de probabilidades diferentes do zero e de 1 das partições $(E_i)_{i?,n}$

Se particionássemos

G	B	W
a	b	c

então

G	0.3	a
B	0.4	b
W	0.3	c

① Duas moedas Cair em pé.

$P(C_2 | X_1 = pé) = (1-\epsilon)\frac{1}{2} + \epsilon / 0$ pé

Cou T	pé
$1-\epsilon$	ϵ

② Roub. a moeda