

## Robustez Bayesiana

1

Problema : A inferência estatística sob o paradigma Bayesiano pode mostrar certa sensibilidade que a priori é subjetiva; isto pode tb. ocorrer com a verossimilhança...

Robustez Supondo que a distribuição à priori é  $f_0(\theta)$  e que a verossimilhança é  $f_0(x|\theta)$ . Na ausência de falta de especificação de  $f_0(\theta)$  é desejável que

- Seja  $\Gamma$  o conjunto de distribuições onde encontra-se a sua priori (correta) i.e.  $f_0(\theta) \in \Gamma$
- Seja  $\Lambda$  o conjunto de distribuições onde encontra-se a verossimilhança i.e.  $f_0(x|\theta) \in \Lambda$
- $\text{if } f(\theta) \in \Gamma \text{ e } f(x|\theta) \in \Lambda, \text{ complete } f(\theta|x) \in \Gamma^*$

Por construir  $f_0(\theta|x) \in \Gamma^*$  ( $f_0(\theta|x) \propto f_0(\theta) \cdot f_0(x|\theta)$ )

Def : se todos os inferências derivadas do cjto de posteriores  $\Gamma^*$

são "suficientemente" próximas às derivadas a partir de  $f_0(\theta|x)$

Dizemos que a inferência é robusta

(2)

Entorno de distribuições

Intérno de proximidade entre distribuições

Dado  $f_0(\theta)$  específico em  $P$ , entendemos  $P$  como um entorno de  $f_0(\theta)$

Exemplos (a) (classe de Variação Total)  $f(\theta) \in P$  se a probabilidade de que o(s) atribui a um determinado cte A "distá" daquele atribuído por  $f_0(\theta)$  em  $\epsilon$ :

$$|P(\theta \in A) - P_0(\theta \in A)| \leq \epsilon$$

usando  $P$                                     usando  $f_0$

Nota:  $P$  é naturalmente delimitado por  $\epsilon$ .

(b) (classe do quociente de densidades)  $f(\theta) \in P$  se  $\exists K$  cte :

$$f_0(\theta) \leq c f(\theta) \leq K \cdot f_0(\theta)$$

ou

$$1 \leq c \frac{f(\theta)}{f_0(\theta)} \leq K ; \text{ onde } c = \text{função de } K$$



Exemplo  $f_0(\theta) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{\theta^2}{2} \right\}; \theta \in (-\infty, +\infty)$

$$f(\theta) = \frac{1}{\pi(1+\theta^2)} \quad \text{cauchy}$$

Vejamos (b) e determinemos  $c \cdot (eK)$

$$\frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} \leq c \cdot \frac{1}{\pi(1+\theta^2)} \leq K \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} \quad \text{se} \quad \frac{1}{c} \leq \sqrt{\frac{2}{\pi}} \cdot \frac{e^{\theta^2/2}}{(1+\theta^2)} \leq \frac{K}{c}$$

Defina  $W(\theta) = \frac{e^{\theta^2/2}}{(1+\theta^2)}$

• limitada inferiormente em 0 por  $W(0)$



• mas não superiormente

[Exercício]

(3)

um terceiro caso  $\rightarrow$  família  $\epsilon$ -contaminante

Dado  $\epsilon > 0$

$$(c) \quad f(\theta) \in \Gamma \text{ se } f(\theta) = (1-\epsilon) f_0(\theta) + \epsilon g(\theta)$$

onde  $g(\theta) \in Q$  (família contaminante)

$$\text{ou } \mathcal{P} = \left\{ (1-\epsilon) f_0(\theta) + \epsilon g(\theta); g(\theta) \in Q \right\}$$

Nota 1 Se  $f(\theta) \in \mathcal{P} \Rightarrow f(\theta|x) = (1-\epsilon^*) f_0(\theta|x) + \epsilon^* g(\theta|x)$

onde  $f_0(\theta|x) \propto f_0(\theta) \cdot f(x|\theta)$ ;  $g(\theta|x) \propto g(\theta) \cdot f(x|\theta)$

$$\epsilon^* = \frac{\epsilon \cdot m_g(x)}{(1-\epsilon) m_0(x) + \epsilon m_g(x)}$$

$$(1-\epsilon) m_0(x) + \epsilon m_g(x)$$

$$\text{onde } m_g(x) = \int_{\#} g(\theta) f(x|\theta) d\theta \quad \& \quad m_0(x) = \int_{\#} f_0(\theta) f(x|\theta) d\theta$$

(H)

Nota 2 Embora a posteriori  $f(\theta|x)$  tenha uma forma compatível aos membros de uma família  $\epsilon^*$ -contaminante,  $\epsilon^*$  depende da lei contaminante  $g(\theta)$ . Não sendo assim uma genuína representante de (c).

$$\text{Vejamos Nota 1: } f_0(\theta|x) = \frac{f_0(\theta) f(x|\theta)}{\int f_0(\theta) f(x|\theta) d\theta} = \frac{f_0(\theta) f(x|\theta)}{m_0(x)}$$

$$\text{analogamente } g(\theta|x) = \frac{g(\theta) f(x|\theta)}{m_g(x)}.$$

$$\begin{aligned} f(\theta|x) &= \frac{f(\theta) f(x|\theta)}{\int f(\theta) f(x|\theta) d\theta} = \frac{[(1-\epsilon) f_0(\theta) + \epsilon g(\theta)] f(x|\theta)}{(1-\epsilon) \int f_0(\theta) f(x|\theta) d\theta + \epsilon \int g(\theta) f(x|\theta) d\theta} = \\ &= \frac{(1-\epsilon) f_0(\theta) f(x|\theta) + \epsilon g(\theta) f(x|\theta)}{(1-\epsilon) m_0(x) + \epsilon m_g(x)} = \end{aligned}$$

(4)

$$f(\theta|x) = \frac{(1-\varepsilon) m_0(x)}{(1-\varepsilon)m_0(x) + \varepsilon m_g(x)} \cdot f_0(\theta|x) + \frac{\varepsilon m_g(x)}{(1-\varepsilon)m_0(x) + \varepsilon m_g(x)} \cdot g(\theta|x)$$

$\rightarrow \varepsilon^* = \frac{\varepsilon m_g(x)}{(1-\varepsilon)m_0(x) + \varepsilon m_g(x)}$

$$= (1-\varepsilon^*) \cdot f_0(\theta|x) + \varepsilon^* \cdot g(\theta|x)$$

— o —

Sensibilidade da informação à posteriori  
 ~~~~~

Suponha que designemos  $E\{h(\theta)|x\}$  para uma posterior cuja proba  $f(\theta) \in \Gamma$ . O natural seria zêss o Suposto nessa onda (que  $f$  não é a representação fidedigna da incerteza em  $\theta$ ) as quantidades seguintes zêss de interesse

$$m = \sup_{(\star)} E\{h(\theta)|x\} \quad \text{e} \quad m = \inf_{(\star)} E\{h(\theta)|x\}$$

Note que  $E(h(\theta)|x) = \frac{\int_{\Theta} h(\theta) f(\theta) f(x|\theta) d\theta}{\int_{\Theta} f(\theta) f(x|\theta) d\theta}$

- Variando  $f$  em  $\Gamma \Rightarrow$  variar numerador e denominador

Pode proceder da seguinte forma: para obter  $m$  (\*) determinar o zero de um expressão. Vamos:

Visto que  $\int_{\Theta} h(\theta) f(\theta) f(x|\theta) d\theta \stackrel{(*)}{\leq} m \int_{\Theta} f(\theta) f(x|\theta) d\theta \quad \therefore$

$$\int_{\Theta} h^*(\theta, m) f(\theta) f(x|\theta) d\theta \leq 0 \quad \begin{array}{l} \text{onde } h^*(\theta|m) \\ = h(\theta) - m \end{array}$$

A88m Considera a expressão

$$t(m) = \sup_n \int h^*(\theta, m) f(\theta) f(x|\theta) d\theta$$

e  $m$  é tal que  $t(m) = 0$ .

— o —

No caso das famílias de contaminações  $\mathcal{E}$ , e de acordo com a Nota ① pag ⑤ temos

$$E\{R(\theta)/x\} = (1-\epsilon^*) E_0(h(\theta)/x) + \epsilon^* E_q(h(\theta)/x)$$

usando  $q(\theta|x)$   
 usando  $f_0(\theta|x)$       envolvendo  $q(\theta)$       ;)

No entanto usando as ideias da Nota 1 podemos ver que

$$E\{h(\theta)/x\} = \frac{a + \int h(\theta) q(\theta) f(x|\theta) d\theta}{b + \int q(\theta) f(x|\theta) d\theta}$$

sendo  $a = \frac{1-\epsilon}{\epsilon} \int h(\theta) f_0(\theta) f(x|\theta) d\theta$

$b = \frac{1-\epsilon}{\epsilon} \int f_0(\theta) f(x|\theta) d\theta$

cts  
não envolve  $q(\theta)$  ;)

veremos:  $E\{h(\theta)/x\} = \frac{(1-\epsilon) \int h(\theta) f_0(\theta) f(x|\theta) d\theta + \epsilon \int h(\theta) q(\theta) f(x|\theta) d\theta}{(1-\epsilon) \int f_0(\theta) f(x|\theta) d\theta + \epsilon \int q(\theta) f(x|\theta) d\theta}$

dividindo por  $\epsilon$

$$= \frac{\left( \frac{1-\epsilon}{\epsilon} \right) \int h(\theta) f_0(\theta) f(x|\theta) d\theta + \int h(\theta) q(\theta) f(x|\theta) d\theta}{\left( \frac{1-\epsilon}{\epsilon} \right) \int f_0(\theta) f(x|\theta) d\theta + \int q(\theta) f(x|\theta) d\theta}$$

$$= \frac{\left( \frac{1-\epsilon}{\epsilon} \right) \int f_0(\theta) f(x|\theta) d\theta + \int q(\theta) f(x|\theta) d\theta}{\left( \frac{1-\epsilon}{\epsilon} \right) \int f_0(\theta) f(x|\theta) d\theta + \int q(\theta) f(x|\theta) d\theta}$$

(6)

nossa problema de achar

$$m = \sup_{\Gamma} E\{h(\theta)/x\}$$

pode então ser novamente reformulado;

$$E\{h(\theta)/x\} = \frac{a + \int h(\theta)g(\theta)f(x|\theta)d\theta}{b + \int g(\theta)f(x|\theta)d\theta} \leq m$$

↓  
se, estamos fe

$$a + \int h(\theta)g(\theta)f(x|\theta)d\theta \leq mb + m \int g(\theta)f(x|\theta)d\theta$$

seu

$$\underbrace{\int [h(\theta) - m] g(\theta)f(x|\theta)d\theta}_{h^*(\theta, m)} \leq mb - a$$

Considerar

$$t_1(m) = \sup_Q \int h^*(\theta, m) g(\theta)f(x|\theta)d\theta$$

$$\text{e achar a solução } t_1(m) = mb - a$$

a qual equivale ao supremo de  $E\{h(\theta)/x\}$  em  $\Gamma$

de Ohmigas

→ Note que as operações são no domínio  $Q$

(7)

De acordo com a probabilidade Bayesiana o procedimento de escolha de um representante de  $\theta$  é feito mediante a incorporação de uma função de perda

Para introduzir as ideias suponha que deseja estimar  $\theta$  a partir da Bayesiana neste  $\theta$ .  $L(\theta, \delta) = (\theta - \delta)^2$

Sob pressupostos quodatores a regra de Bayes é tal que

$$E\{L(\theta, \delta^*)|x\} < E\{L(\theta, \delta)|x\}.$$

No caso contínuo:  $E(L(\theta, \delta)|x) = \int_{-\infty}^{\infty} (\theta - \delta)^2 \cdot \pi(\theta|x) d\theta$

$$= \int_{-\infty}^{\infty} (\theta^2 - 2\delta\theta + \delta^2) \pi(\theta|x) d\theta = \int_{-\infty}^{\infty} \theta^2 \pi(\theta|x) d\theta - 2\delta E(\theta|x) + \delta^2$$

$$\frac{d}{d\delta} E\{L(\theta, \delta)|x\} = -2E(\theta|x) + 2\delta = 0 \quad \text{see } \boxed{\delta^* = E(\theta|x)}$$

the mínimo p.  $L \geq 0$

$\rightarrow$  exemplo ~~pag 9~~  $\rightarrow$  novo  $\rightarrow$

Considera em complemento a perda multilinear

$$L_{k_1, k_2}(\theta, \delta) = \begin{cases} k_2(\theta - \delta) & \text{se } \theta > \delta \\ k_1(\delta - \theta) & \text{se } \theta \leq \delta \end{cases}$$

$$E(L_{k_1, k_2}(\theta, \delta)|x) = \int_{-\infty}^{\delta} k_1(\delta - \theta) \pi(\theta|x) d\theta + \int_{\delta}^{\infty} k_2(\theta - \delta) \pi(\theta|x) d\theta$$

$$= k_1 \delta \int_{-\infty}^{\delta} \pi(\theta|x) d\theta - k_1 \int_{-\infty}^{\delta} \theta \pi(\theta|x) d\theta + k_2 \int_{\delta}^{\infty} \theta \pi(\theta|x) d\theta - k_2 \delta \int_{\delta}^{\infty} \pi(\theta|x) d\theta$$

Derivando em  $\delta$

$$0 = k_1 \int_{-\infty}^{\delta} \pi(\theta|x) d\theta + k_1 \delta \pi(\delta|x) - k_1 \delta \pi(\delta|x) + k_2 (-1) \cdot \delta \pi(\delta|x) -$$

$$- k_2 \int_{\delta}^{\infty} \pi(\theta|x) d\theta - k_2 \delta (-1) \pi(\delta|x)$$

(8)

$$0 = k_1 P(\theta \leq \delta | x) - k_2 P(\theta > \delta | x)$$

$$= k_1 P(\theta \leq \delta | x) - k_2 \{1 - P(\theta \leq \delta | x)\}$$

$$= k_1 P(\theta \leq \delta | x) - k_2 + k_2 P(\theta \leq \delta | x)$$

$$\frac{k_2}{k_1 + k_2} = P(\theta \leq \delta | x) \rightarrow \text{solución} \rightarrow \text{de Bayes}.$$

por ejemplo si  $k_1 = k_2 = k \Rightarrow$  la regla de Bayes es la mediana de  $\theta$  en posterior

Ejemplo

$$x_1, x_n \sim \text{Ba}(\theta), \quad \pi(\theta) = \text{Beta}(\alpha, \beta) \\ = \theta^{\alpha-1} \cdot (1-\theta)^{\beta-1}$$

$$\prod_{i=1}^n f(x_i | \theta) \cdot \pi(\theta) \propto \theta^{x_1} \cdots \theta^{x_n} \cdot (1-\theta)^{1-x_1} \cdots (1-\theta)^{1-x_n} \propto \theta^{\alpha-1} \cdot (1-\theta)^{\beta-1}$$

$$= \theta^{\sum_{i=1}^n x_i + \alpha - 1} \cdot (1-\theta)^{n - \sum_{i=1}^n x_i + \beta - 1}$$

i.e.  $\pi(\theta | x_1, \dots, x_n)$  es

$$\text{Beta}(\sum x_i + \alpha, n - \sum x_i + \beta)$$

Regla de Bayes

$$\frac{\alpha}{\alpha + \beta} \text{ (prior)}$$

(a) Sobre punto quodrofico  
(Medie)

$$= \frac{\sum x_i + \alpha}{m + \alpha + \beta}$$

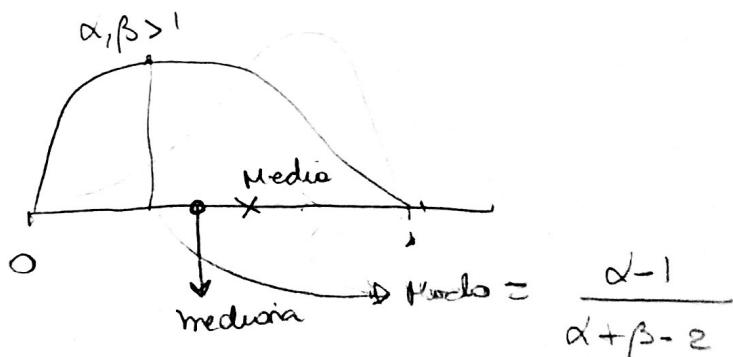
(b) Sobre punto multilinear ( $k_1 = k_2 = k$ )  $\approx \frac{\sum x_i + \alpha - \frac{1}{3}}{m + \alpha + \beta - \frac{2}{3}}$

$\alpha = 2, \beta = 3 \rightarrow$  (Mediana)

$$\text{Medio} = \frac{\alpha}{\alpha + \beta} = 0,4$$

$$\text{Mediana} = \frac{\alpha - \frac{1}{3}}{\alpha + \beta - \frac{2}{3}} = 0,33$$

$$\text{Modo} = \frac{\alpha - 1}{\alpha + \beta - 2} = 0,33$$



(9)

toma excepto 4/5 o Método Bay

$$\pi(\mu) \text{ é } N(m, \sigma^2) \quad \xrightarrow{\text{conocido}}$$

$$\pi(\mu | x_1, \dots, x_n) \text{ é } N(m^*, \sigma^2)$$

$$m^* = \frac{\frac{m}{\sigma^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\sigma^2} + \frac{n}{\sigma^2}}$$

regras de Bayes  
sob pressão  
que os dados

considere que a priori não fosse bem formulada em termos do  
medio  $m$  temos  $\pi(\mu) \text{ é } N(m + \Delta_0, \sigma^2)$

$$m^* = \frac{\frac{m + \Delta_0}{\sigma^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\sigma^2} + \frac{n}{\sigma^2}} = \frac{\frac{m}{\sigma^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\sigma^2} + \frac{n}{\sigma^2}} + \frac{\frac{\Delta_0}{\sigma^2}}{\frac{1}{\sigma^2} + \frac{n}{\sigma^2}}$$

$$= m^* + \frac{\Delta_0 (\sigma^2 \cdot n)}{\sigma^2 [n + 1]} = m^* + \sqrt{\frac{\Delta_0 \cdot \sigma^2}{(n + 1)}} \quad (\rightarrow \text{pequeno quando } \Delta_0 \gg \sigma^2)$$

ou seja se a priori é formulada com grande incerteza  
corroborase o problema

O que diz que a decisão de Bayes é robusta sob  
certas considerações