Inferência subjetiva

Motivação nas Exatas: (Teorema de Bayes) : Sejam A e B dois eventos desde a identidade $P(A)P(B/A) = P(B)P(A/B)$ resulta a forma

$$P(B/A) = \frac{P(A/B)}{P(A)} \cdot P(B) \qquad \text{se } P(A) > 0 .$$

Interpretação : Se B é o evento de interesse ; B tem uma probabilidade de acontecer : $P(B)$. Numa segunda etapa : A acontece com probabilidade $P(A)$ então a probabilidade de acontecer B (evento de interesse) muda para $P(B/A)$

$$P(B) \xrightarrow[\substack{\text{mudança de} \\ \text{espaço de} \\ \text{probabilidades}}]{} P(B/A)$$

O teorema de Bayes pode ser interpretado como um mecanismo de atualização de probabilidades, conforme a mudança de informação.

Observação : a probabilidade atual de acontecer de B : $P(B/A)$ cresce em relação à probabilidade antiga de acontecer de B : $P(B)$

Sempre que $P(A/B) > P(A)$.

Generalizando o teorema de Bayes:

Sejam $B_1, B_2, \ldots\ldots$ eventos mutuamente exclusivos e exaustivos

Então $P(B_r/A) = \frac{P(A/B_r)}{P(A)} \cdot P(B_r)$

$$P(B_r/A) = \frac{P(A/B_r)}{\sum_{s=1}^{\infty} P(A/B_s) \cdot P(B_s)} \cdot P(B_r)$$

Já que $P(A) = \sum_{s=1}^{\infty} P(A/B_s) P(B_s)$ (Teorema da probabilidade total).

Observação: a probabilidade $P(B_r/A)$ (logo de A ter acontecido com probabilidade $P(A)$) cresce em relação a $P(B_r)$ sempre q'

$$P(A/B_r) > \sum_{s=1}^{\infty} P(A/B_s) P(B_s)$$

— o —

Verossimilhança: lembrando o teorema de Bayes: $P(B/A) = \frac{P(A/B)}{P(A)} P(B)$

Sabendo q' o evento de interesse é B, a probabilidade $P(A/B)$ é conhecida como Verossimilhança de B dado A

Priori: corresponde ao valor $P(B)$ ou asignação de probabilidades feita sem vínculo ao evento A.

Posteriori: representada pelo valor $P(B/A)$. Atualização de probabilidades logo de ter sido incorporado o evento A.

desde o teorema de Bayes: se desejamos fazer inferência de θ, baseados na informação do dado x

$$dF(x)\, dF(\theta|x) = d\, F(x;\theta) = dF(\theta)\, dF(x|\theta)$$

$$dF(\theta|x) = \frac{dF(x|\theta)}{dF(x)} \cdot dF(\theta)$$

logo tendo as distribuições (contínuas ou discretas) correspondentes, resulta

$$f(\theta|x) = \frac{f(x|\theta)}{f(x)} \cdot f(\theta)$$

$$f(\theta|x) = \frac{f(x|\theta)}{\int f(\theta) f(x|\theta)\, d\theta}\, f(\theta)$$

Em geral como o dado x é conhecido então $f(x)$ é cte logo utilizamos a notação

$$f(\theta|x) \propto f(x|\theta)\, f(\theta)$$

onde f pode denotar uma densidade ou uma probabilidade pontual.

Exemplo: Seja x o número de sucessos numa série de n realizações com probabilidade de sucesso θ em cada uma delas. Isto é, x possui densidade binomial dado θ.

$$f(x|\theta) = \binom{n}{x} \theta^{x} (1-\theta)^{n-x} \qquad x = 0, 1, \ldots, n$$

lembre que o dado x é fixo, ou seja x é igual a algum dos valores $0, 1, \ldots, n$ (Na realização do experimento)

$f(x|\theta)$ é a verossimilhança de θ dado o x.

Consideremos θ com uma distribuição Beta (p,q)

$$f(\theta) = \frac{1}{B(p,q)}\theta^{p-1}(1-\theta)^{q-1} \quad ; \quad 0 \leq \theta \leq 1$$

p e q são conhecidos (>0)

e $B(p,q) = \int_0^1 \theta^{p-1}(1-\theta)^{q-1}\,d\theta$.

$f(\theta)$ é a priori para θ.

Pelo teorema de Bayes $f(\theta/x) = \frac{f(x/\theta)}{f(x)} \cdot f(\theta)$ onde

$$f(x) = \int f(x/\theta) f(\theta)\,d\theta$$

Calculando $f(x) = \int f(x/\theta) f(\theta)\,d\theta = \int_0^1 \binom{n}{x}\theta^x(1-\theta)^{n-x} \cdot \frac{1}{B(p,q)}\theta^{p-1} \cdot (1-\theta)^{q-1}\,d\theta$

$$= \binom{n}{x}\frac{1}{B(p,q)}\int_0^1 \theta^{p+x-1}(1-\theta)^{n+q-x-1} \cdot d\theta$$

$$f(x) = \binom{n}{x}\frac{1}{B(p,q)} \cdot B(p+x\,;\,n+q-x)$$

finalmente $f(\theta/x) = \frac{f(x/\theta)}{f(x)} f(\theta) = \frac{\binom{n}{x}\theta^x(1-\theta)^{n-x}}{\binom{n}{x}\frac{B(p+x\,;\,n+q-x)}{B(p,q)}} \cdot \frac{1}{B(p,q)}\theta^{p-1} \cdot (1-\theta)^{q-1}$

$$= \frac{\theta^x(1-\theta)^{n-x} \cdot \theta^{p-1} \cdot (1-\theta)^{q-1}}{B(p+x\,;\,n+q-x)}$$

$$f(\theta/x) = \frac{\theta^{p+x-1} \cdot (1-\theta)^{n+q-x-1}}{B(p+x\,;\,n+q-x)}$$

logo $f(\theta/x)$ é uma Beta $(p+x\,;\,n+q-x)$

Inferência : θ tinha uma densidade Beta (p,q) p e q conhecidos >0. uma experiencia foi desenvolvida. e θ a posteriori (da experiencia) tem uma densidade Beta $(p+x\,;\,n+q-x)$

* Verossimilhança: $\frac{\partial}{\partial\theta}\ln f(x|\theta) \propto \frac{\partial}{\partial\theta}\left\{\ln(\theta^x)+\ln\left((1-\theta)^{n-x}\right)\right\}$

$$= \frac{x}{\theta} - \frac{(n-x)}{1-\theta} = \frac{x-\cancel{x\theta}-n\theta+\cancel{x\theta}}{\theta(1-\theta)}$$

$$= \frac{x-n\theta}{\theta(1-\theta)} = 0 \quad \text{sse} \quad x = n\theta$$

$$\text{sse} \quad \hat{\theta}_{MV} = \frac{x}{n}$$

a Verossimilhança favorece ao valor de θ : $\hat{\theta}_{MV} = \frac{x}{n}$

* Priori : $f(\theta) \propto \theta^{p-1}(1-\theta)^{q-1}$: calculo do valor de maior densidade

$$\frac{\partial f(\theta)}{\partial\theta} \propto (p-1)\theta^{p-2}.(1-\theta)^{q-1} + \theta^{p-1}.(q-1)(1-\theta)^{q-2}(-1)$$

$$= \theta^{p-2}.(1-\theta)^{q-2}.\left\{(p-1)(1-\theta)-(q-1)\theta\right\} = 0 \quad \text{não trivial } \left(\theta \neq {0 \atop 1}\right)$$

$$\text{sse} \quad (p-1)(1-\theta) = (q-1)\theta$$

$$(p-1)-\theta(p-1) = (q-1)\theta$$

$$p-1 = (q-1+p-1)\theta$$

$$p-1 = (q+p-2)\theta$$

$$\theta_{pri} = \frac{p-1}{q+p-2} \quad \text{Moda da priori}$$

a priori favorece ao valor $\theta = \frac{p-1}{q+p-2}$

* posteriori : $f(\theta/x)$ é Beta $(p+x \,;\, n+q-x)$

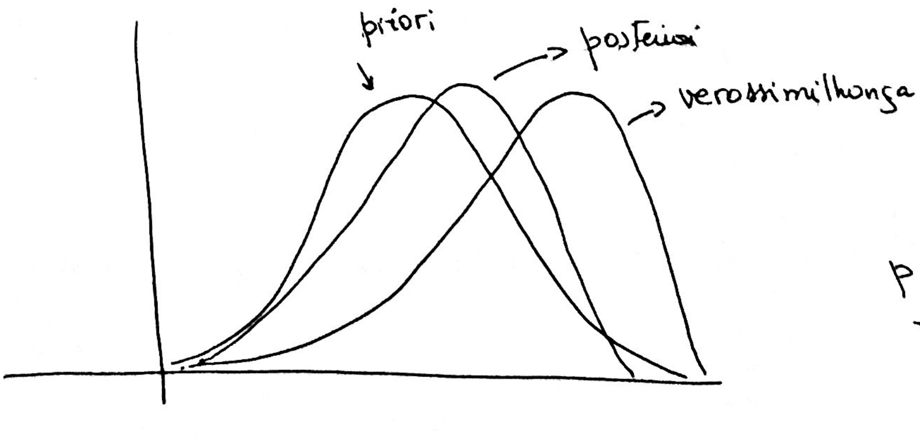
por analogia com a conta anterior ; a moda neste caso é

$$\theta = \frac{p+x-1}{p+x+n+q-x-2} = \frac{p+x-1}{n+p+q-2} = \theta_{pos}$$

Moda da posteriori.

$$\theta_{pos} = a \cdot \hat{\theta}_{MV} + (1-a)\theta_{pri} \quad \text{onde } a = \frac{n}{n+p+q-2}$$

Em relação ao θ a distribuição à posteriori tem toda a informação necessária para fazer algum tipo de inferencia de θ : resumenes como a media, moda (θ_{pos}), variância ... servem para recuperar a informação contida "codificada" pela posteriori $f(\theta/x)$.



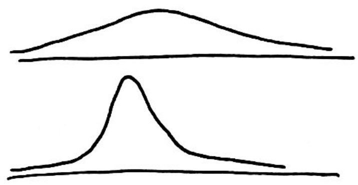
$p = q = 4$
$n = 6$
$x = 5$

A distribuição à posteriori contem toda a informação relativa ao θ. Resumenes simples media, mediana, moda, ... poderão responder (pontualmente) perguntas pontuais.

informação Vs dispersão

dispersão grande → informação fraca

" pequena → " forte



Procuramos q' os dados permitam fabricar uma posteriori de dispersão pequena (redução da incerteza)

Força da informação $= \frac{1}{\text{Var}} =$ precisão (se a precisão é grande a Var e pequena logo a informação forte)

ADunicamp

d s t q q s s

Quando a Inf Bayesiana é simples e qdo não:

a) fácil: suponho $x_1 \ldots x_n \sim N(\mu, \sigma^2)$

μ desconhecido, σ^2 conhecido.

$$f(x_i \mid \mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

$$L = \prod_{i=1}^{n} f(x_i/\mu) = \frac{1}{(2\pi)^{n/2} \cdot \sigma^{n \cdot \frac{2}{2}}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2\right\}$$

$$= \frac{1}{(2\pi)^{n/2} \cdot \sigma^{\frac{2n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n} x_i^2 - 2\mu x_i + \mu^2\right]\right\}$$

$$= \frac{1}{(2\pi)^{n/2} \cdot \sigma^{\frac{2n}{2}}} \exp\left\{\underbrace{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2}_{\text{cte } (\sigma^2 \text{ conhecido})} + \frac{2\mu}{2\sigma^2}\sum_{i=1}^{n} x_i - \frac{\mu^2 n}{2\sigma^2}\right\}$$

$$\propto \exp\left\{\frac{\mu}{\sigma^2} n\bar{x} - \frac{\mu^2 n}{2\sigma^2}\right\} = \exp\left\{-\frac{n}{2\sigma^2}\left(\mu^2 - 2\mu\bar{x}\right)\right\}$$

vejamos $(\mu - \bar{x})^2 = \mu^2 - 2\mu\bar{x} + \bar{x}^2$

$$\propto \exp\left\{-\frac{n}{2\sigma^2}\left(\mu^2 - 2\mu\bar{x} + \bar{x}^2\right)\right\} = \exp\left\{-\frac{n}{2\sigma^2}(\mu - \bar{x})^2\right\}$$

ie $f(x_1 \; x_n/\mu) \underset{\text{em } \mu}{\propto} \exp\left(-\frac{n}{2\sigma^2}[\mu - \bar{x}]^2\right)$

Se consideramos a priori $f(\mu) \propto \exp\left\{-\frac{1}{2\gamma}(\mu-m)^2\right\}$

$$\Rightarrow \pi(\mu|x_1 \ldots x_n) \propto \exp\left\{-\frac{1}{2\gamma}(\mu-m)^2 - \frac{n}{2\sigma^2}(\mu-\bar{x})^2\right\}$$

temos

$$-\frac{1}{2\gamma}\left[\mu^2 - 2\mu m + m^2\right] - \frac{n}{2\sigma^2}\left[\mu^2 - 2\mu\bar{x} + \bar{x}^2\right] \propto$$

$$-\left(\frac{1}{2\gamma} + \frac{n}{2\sigma^2}\right)\mu^2 + 2.\mu\left[\frac{m}{2\gamma} + \frac{n\bar{x}}{2\sigma^2}\right]$$

$$= -\frac{1}{2}\left[\frac{1}{\gamma} + \frac{n}{\sigma^2}\right]\mu^2 + 2\frac{1}{2}\mu\left[\frac{m}{\gamma} + \frac{n\bar{x}}{\sigma^2}\right]$$

$$= -\frac{1}{2}\left[\frac{1}{\gamma} + \frac{n}{\sigma^2}\right]\left\{\mu^2 - 2\mu.\frac{\left(\frac{m}{\gamma} + \frac{n\bar{x}}{\sigma^2}\right)}{\left(\frac{1}{\gamma} + \frac{n}{\sigma^2}\right)}\right\}.$$

$$V^* = \left(\frac{1}{\gamma} + \frac{n}{\sigma^2}\right)^{-1}; \quad m^* = \frac{\left(\frac{m}{\gamma} + \frac{n\bar{x}}{\sigma^2}\right)}{\left(\frac{1}{\gamma} + \frac{n}{\sigma^2}\right)}$$

posteriori com media m^* e Var V^*.

Ver que ocorre q.do $n \to \infty$ com m^*

$$V^* = \frac{1}{\frac{1}{\gamma} + \frac{n}{\sigma^2}} \xrightarrow[n\to\infty]{} 0 \; ; \left(\text{l'Hopital em } m^*\right) \; m^* \xrightarrow[n\to\infty]{} \frac{\bar{x}/\sigma^2}{1/\sigma^2} = \bar{x}$$

b) . modelado

Considere a priori no problema anterior Logística

$$f(\mu) \propto \exp(p\cdot\mu)\,\{1+\exp(\mu)\}^{-(p+q)}$$

i.e. $\pi(\mu/x)$ não tem forma fechada.

obs → p/ se tratar de 1 parâm μ basta a integração numérica.