Bayesiana

* Lista 2

1)

i) Seja: H: A moeda é honesta; $\mathbb{P}(H) = \frac{1}{2} = 0,5$;

H^c: A moeda é falsa; $\mathbb{P}(H^c) = 1 - \frac{1}{2} = 0,5$;

C: O resultado do lançamento é cara;

C^c: " " " " " coroa.

Então; $\mathbb{P}(C|H) = 0,5$; $\mathbb{P}(C|H^c) = 1$ {Pois H^c tem 2 caras}.

Então; $\mathbb{P}(C) = \mathbb{P}(C|H)\mathbb{P}(H) + \mathbb{P}(C|H^c)\mathbb{P}(H^c)$

(Lei da probabilidade total)

$$= (0,5)(0,5) + 1 \cdot 0,5$$
$$= 0,75$$

ii) $$\mathbb{P}(H^c|C) = \underbrace{\frac{\mathbb{P}(C|H^c)\mathbb{P}(H^c)}{\mathbb{P}(C)}}_{\text{Teorema de Bayes}} = \frac{1 \cdot 0,5}{0,75} = \frac{2}{3}$$

iii) Se C = O resultado é cara; $C = C^1 = C$. Se repetirmos o experimento; $C \cdot C \cdot C \cdot \ldots \cdot C = C^n$ = O resultado é cara em n lançamentos.

Então; $\mathbb{P}(C^n) = \mathbb{P}(C^n|H)\mathbb{P}(H) + \mathbb{P}(C^n|H^c)\mathbb{P}(H^c)$

(Lei da prob. total)

$$= (0,5^n)(0,5) + \overset{1}{1^n} \cdot 0,5$$
$$= 0,5^{n+1} + 0,5.$$

(Continuação 1) iii))

Logo, temos que $P(H^c | C^n) = \dfrac{\overbrace{P(C^n | H^c)}^{(*)} P(H^c)}{P(C^n)}$ Teorema de Bayes

$(*) : P(C^n | H^c) = P(C | H^c) = 1$

Se só há caras em H^c; $(*) = 1 \ \forall n > 0$.

$= \dfrac{1 \cdot 0,5}{0,5^{n+1} + 0,5}$;

$\lim_{n \to \infty} \dfrac{0,5}{\underbrace{0,5^{n+1}}_{\to 0} + 0,5} = 1$. Isto é, quanto mais lançarmos a moeda, e mais caras encontrarmos obtermos, mais crível é a hipótese da moeda de fato ser falsa.

iv) Seja C^1 = O 1º lançamento resultou cara;
C^2 = O 2º lançamento é cara.

Temos: Em 1) ii); calculamos $P(H^c | C) = \frac{2}{3} = P(H^c | C^1)$.
Então $P(H | C) = 1 - 2/3 = 1/3 = P(H | C^1)$

Pelo teorema de Bayes, $P(C^2 | C^1) = P(C^2 | H^c) P(H^c | C^1) + P(C^2 | H) P(H | C^1)$;

Onde $P(C^2 | H^c) = 1$; e
$P(C^2 | H) = 1/2$, obviamente.

Logo, $P(C^2 | C^1) = 1 \cdot 2/3 + 0,5 \cdot 1/3$
$= 5/6$.

v) Os lançamentos serão independentes $\iff \mathbb{P}(C^j \cap C^k) = \mathbb{P}(C^j) \cdot \mathbb{P}(C^k)$

ou, $\mathbb{P}(C^j \mid C^k) = \mathbb{P}(C^j)$; o que não foi observado neste caso, visto que mostramos que $\mathbb{P}(C^2 \mid C^1) \neq \mathbb{P}(C^2)$

Para obter a independência condicional, precisaríamos saber se a moeda é honesta ou não.

2) $y_1, y_2, \ldots, y_n \sim \text{Bernoulli}(\theta)$,

$\theta \sim \text{Uniforme}(0, 1)$

i) Posteriori: $\mathbb{P}(\theta \mid y_1, y_2, \ldots, y_n) \propto \mathbb{P}(y_1, \ldots, y_n \mid \theta)\,\mathbb{P}(\theta)$.

Seja $\sum_{i=1}^{n} y_i = k$ o número de caras observadas na amostra; Então;

$\mathbb{P}(y_1, \ldots, y_n \mid \theta) = \theta^k (1-\theta)^{n-k}$ é a verossimilhança de uma Bernoulli.

E $\mathbb{P}(\theta) = 1$. Então;

$\mathbb{P}(\theta \mid y_1, \ldots, y_n) \propto \theta^k (1-\theta)^{n-k} \cdot 1$

Que tem a forma de uma Beta($\alpha = k+1$, $\beta = n-k+1$)

Se a posteriori de θ é uma Beta com estes parâmetros, então:

$$\mathbb{E}(\theta) \underset{(\text{Def.})}{=} \frac{\alpha}{\alpha+\beta} = \frac{k+1}{(k+1) + (n-k+1)} = \frac{k+1}{n+2};$$

E variância => $Var(\theta) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ (Def.)

$$= \frac{(K+1)(n-K+1)}{[(K+1)+(n-K+1)]^2[(K+1)+(n-K+1)+1]}$$

$$= \frac{(K+1)(n-K+1)}{(n+2)^2(n+3)}$$

ii) Expressar a esp. da forma $(1-\omega)E(\theta)+\omega\hat{\theta}$.

A priori; $\theta \sim Unif$; logo $E(\theta) = \dfrac{a+b}{2} = \dfrac{0+1}{2} = 0{,}5$.

E, o estimador $\hat{\theta}$ para a bernoulli é dado por $\dfrac{K}{n}$, i.e, a soma de sucessos, dividido pelo número de tentativas.

E, a esperança da posterior calculada em i) é $\dfrac{\alpha}{\alpha+\beta} = \dfrac{K+1}{n+2}$.

Portanto:

$$\frac{K+1}{n+2} = (1-\omega)E(\theta)+\omega\hat{\theta}$$

$$= (1-\omega)\frac{1}{2}+\frac{\omega K}{n} = \underbrace{\frac{(1-\omega)}{2}}_{\text{Priori}} + \underbrace{\frac{\omega K}{n}}_{\text{Verossimilhança}}$$

Vem da:

Desta forma, vemos que a esperança da posterior pode ser escrita como uma soma ponderada do conhecimento a priori e a verossimilhança dos dados, e ω deve ser tal que o aumento da amostra da verossimilhança torne o peso da priori quase insignificativo conforme n aumenta, i.e; $\frac{\text{constante}}{f(n)} = \omega$.

iii)

Preditiva: $\mathbb{E}_{\theta \sim Beta(K+1,\, n-K+1)}[\text{Amostral}(y|\theta)]$

Então,

$$P(y_{n+1} | y_1, \dots, y_n) = \int_0^1 \underbrace{P(y_{n+1}|\theta)}_{\substack{\text{Amostral} \\ \text{(Bernoulli)}}} \underbrace{P(\theta | y_1, \dots, y_n)}_{\substack{\text{Posteriori} \\ \text{(Beta)}}} d\theta$$

logo...

$$P(y_{n+1} | y_1, \dots, y_n) \overset{(*)}{=} \int_0^1 \theta^{y_{n+1}} (1-\theta)^{1-y_{n+1}} \frac{\theta^K (1-\theta)^{n-K}}{B(K+1, n-K+1)} d\theta$$

$$^{*}\text{Se } y_{n+1} = 1, (*) = \int_0^1 \frac{\theta^{K+1}(1-\theta)^{n-K}}{B(K+1, n-K+1)} d\theta =$$

$$= \frac{1}{B(K+1, n-K+1)} \int_0^1 \theta^{K+1}(1-\theta)^{n-K} d\theta = \frac{B(K+2, n-K+1)}{B(K+1, n-K+1)}$$

$$= \frac{\Gamma(\overset{K+1}{\cancel{K+2}})\,\Gamma(\cancel{n-K+1})}{\Gamma(K+2+n-K+1)} \Bigg/ \frac{\Gamma(\cancel{K+1})\,\Gamma(\cancel{n-K+1})}{\Gamma(n+2)} = \boxed{\frac{K+1}{n+2}}$$

6

Se $y_{n+1}=0$;

$$\mathbb{P}(y_{n+1}=0 \mid y_1,\dots,y_n) = \int_0^1 \theta^0 (1-\theta)^{1-0} \frac{\theta^k (1-\theta)^{n-k}}{B(k+1,n-k+1)} d\theta$$

$$= \frac{1}{B(k+1,n-k+1)} \int_0^1 \theta^k (1-\theta)^{n-k+1} d\theta$$

$$= \frac{B(k+1,n-k+2)}{B(k+1,n-k+1)} = \frac{\dfrac{\Gamma(k+1)\Gamma(n-k+2)}{\Gamma(n+3)}}{\dfrac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)}}$$

$$= \frac{\cancel{\Gamma(n-k+2)}^{\,n-k+1}\,\cancel{\Gamma(n+2)}}{\underbrace{\cancel{\Gamma(n+3)}}_{n+2}\,\cancel{\Gamma(n-k+1)}} = \boxed{\frac{n-k+1}{n+2}}$$

4) i) $PQ = \arg\min \mathbb{E}\left[(\theta - PQ)^2 \mid y_1,\dots,y_n\right]$

$=$ Média posterior $= \dfrac{k+1}{n+2}$

ii) Perda 0-1 = Moda, P/ $n=12$, $S=\sum_{i=1}^{12} y_i = 9$

Def Moda Beta $\dfrac{\alpha-1}{\alpha+\beta-2}$, $\alpha,\beta>1$

Então $\dfrac{9+1-1}{9+1+12-9+1-2} = \dfrac{9}{12} = 0{,}75$

iii) Perda absoluta = Mediana posteriori ($n=12$; $s=9$)

Def Mediana Beta $\approx \dfrac{\alpha - 1/3}{\alpha+\beta-2/3}$

$$\approx \frac{\overset{10}{9+1} - 1/3}{\underbrace{9+1+12-9+1}_{14}-2/3} = \frac{9{,}\overline{66}}{13.\overline{33}}$$

$\approx 0{,}725$

iv) HPD 99%

Será o intervalo quantílico da posteriori, i.é

no R: qbeta(.005, 10, 4); qbeta(.995, 10, 4)

$= \,]0{,}3793\ldots ; 0{,}9429\ldots[$

8) i)

Priori: $(p_1, p_2, p_3) \sim Dir(\alpha_1, \alpha_2, \alpha_3)$

Verossimilhança: $\mathbb{P}(x_1, x_2, x_3 \mid p_1, p_2, p_3) = \text{Multinom}\begin{pmatrix} p_1, p_2, p_3; \\ x_1, x_2, x_3, n \end{pmatrix}$

$$= \frac{n!}{x_1!\,x_2!\,x_3!}\, p_1^{x_1}\, p_2^{x_2}\, p_3^{x_3}$$

Posteriori:

∘ Tentei derivar esta posteriori, mas mostrou-se excessivamente complicado. Então vamos pela definição que a Dirichlet é a priori conjugada da multinomial.

Portanto a posterior será da forma:

$(p_1, \ldots, p_K) \mid (x_1, \ldots, x_K) \sim \text{Dirichlet}(\alpha_1 + x_1, \ldots, \alpha_K + x_K)$.

Neste caso.

· Posteriori:

$$\mathbb{P}(p_1, p_2, p_3 \mid x_1, x_2, x_3) \propto p_1^{\alpha_1-1} p_2^{\alpha_2-1} p_3^{\alpha_3-1} \cdot p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

$$= p_1^{\alpha_1-1+x_1} p_2^{\alpha_2-1+x_2} p_3^{\alpha_3-1+x_3}$$

Que é uma Dirichlet com parâmetros $[\alpha_1+x_1, \alpha_2+x_2, \alpha_3+x_3]$

· Posteriori marginais:

[Zamani, M. 2008]

As marginais serão $p_i \sim Beta(\alpha_i, \underset{\sim}{\alpha} - \alpha_i)$ Logo:

* $p_1 \mid (x_1, x_2, x_3) \sim Beta(\alpha_1+x_1, \alpha_2+x_2, \alpha_3+x_3)$.
* $p_2 \mid (x_1, x_2, x_3) \sim Beta(\alpha_2+x_2, \alpha_1+x_1, \alpha_3+x_3)$;
* $p_3 \mid (x_1, x_2, x_3) \sim Beta(\alpha_3+x_3, \alpha_1+x_1, \alpha_2+x_2)$.

ii) PQ de $p_j - p_i$ $(i, j = 1,2,3, \ i \neq j)$

Def. Esperança Dirichlet: $\mathbb{E}(X_i) = \dfrac{\alpha_i}{\sum_{i=1}^{k} \alpha_i}$

Então: * $\hat{p}_1 = \dfrac{\alpha_1+x_1}{\alpha_1+x_1+\alpha_2+x_2+\alpha_3+x_3}$

* $\hat{p}_2 = \dfrac{\alpha_2+x_2}{\alpha_1+x_1+\alpha_2+x_2+\alpha_3+x_3}$

* $\hat{p}_3 = \dfrac{\alpha_3+x_3}{\alpha_1+x_1+\alpha_2+x_2+\alpha_3+x_3}$

Como a esperança é linear, basta termos
$\mathbb{E}[p_j - p_i] = \mathbb{E}[p_j] - \mathbb{E}[p_i]$, Em que $\mathbb{E}[p_j]$ terá a mesma forma de $\mathbb{E}[p_i]$. $\forall i \neq j \in \{1, 2, 3\}$

11)
i) H_0) A droga não é melhor que o placebo : $\theta \leq 0{,}25$
H_1) A droga é superior ao placebo : $\theta > 0{,}25$
ii)
Priori: $U(0,1)$: $\mathbb{P}(\theta \leq 0{,}25) = 0{,}25$,
$\mathbb{P}(\theta > 0{,}25) = 0{,}75$
Chance relativa $= \dfrac{1 - 0{,}25}{0{,}25} = 3$

Verossimilhança:
A uniforme é a priori conjugada da Beta. Portanto, este será o formato da verossimilhança. Para $n = 20$ e $s = 8$, sob H_0), temos $V \sim \text{Beta}(8, 12)$

Posteriori:
Portanto, a posterior será justamente $P_0 \sim \text{Beta}(9, 13)$.
Logo, testando as hipóteses:

* $\mathbb{P}(\theta \leq 0{,}25 | 8) = \mathbb{P}[\text{Beta}(0{,}25 | 9, 13)$ $\{H_0\}$
No R: pbeta(0,25, 9, 13) $= 0{,}0561\ldots$
* $\{H_1\}$ $\mathbb{P}(\theta > 0{,}25 | 8) = 1 - \mathbb{P}[\text{Beta}(0{,}25 | 9, 13)$
No R: 1 - pbeta(0,25, 9, 13) $= 0{,}9438\ldots$

Logo, a chance relativa posteriori será:

No R: $(1 - pbeta(0.25, 9, 13))/(pbeta(.25, 9, 13))$

$= 16,81044.$

Portanto, o Fator de Bayes será:

$$\frac{16,81044}{3} = 5,60348$$

iii) $d = 1 \therefore$ Rejeitar; $d = 0 \therefore$ não rejeitar

F.P.V: 5x + custoso rejeitar H_0; H_0 verdadeira que não rejeitar H_0; H_0 é falsa

Função perda: $L(d, \theta)$; Risco = R.

Lança a droga / A droga é eficaz	Sim	Não
Sim	R_x	$R_y \rightarrow$ Ñ Rej H_0)
Não	$5R_y \rightarrow$ Rej. H_0)	R_x

Logo, a decisão "Ótima" será rejeitar H_0 caso:

$$\mathbb{P}(\theta > 0,25|S) > \frac{5R_y}{5R_y + R_y} = \frac{5}{6}$$

E, como $\mathbb{P}(\theta > 0,25|8) = 0,9438 > 5/6$,

A decisão é pela rejeição de H_0).

iv) A decisão de rejeitar H_0 está sendo fortalecida pela priori. Se tomarmos uma priori imprópria Beta(0,0); ela apenas multiplica a Verossimilhança por 1; (Não está no paradigma Bayesiano; mas na prática "dá no mesmo"); teríamos uma posteriori Beta(8,12) (= Verossimilhança).

No R: 1 - pbeta(.25, 8, 12) = 0,9225... > 5/6

A decisão seria a mesma neste caso (Rejeita H_0); Mas poderia mudar a decisão em outro exemplo...

Lista 3.1

1) Vero: $x_1, \dots, x_n \sim \text{Poisson}(\theta)$

$\theta \sim \text{Gama}(\alpha, \beta)$

i) F.D.P: $\dfrac{\theta^{k} e^{-\theta}}{k!}$; Análogo: $\dfrac{\theta^{x_i} e^{-\theta}}{x_i!}$

Então, a Verossimilhança será

$$L(x_i|\theta) = \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} = \frac{\theta^{\sum_i x_i} e^{-n\theta}}{\prod_{i=1}^{n} x_i!}$$

, e $\theta \equiv P(\theta) = \dfrac{\theta^{\alpha-1} e^{-\beta\theta}}{\Gamma(\alpha)} \propto \theta^{\alpha-1} e^{-\beta\theta}$

Então, a posteriori irá herdar as características da priori, com a convolução com a verossimilhança:

Posteriori:

$$\mathbb{P}(\theta \mid x_1, \dots, x_n) = \frac{\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta}}{\prod_{i=1}^{n} x_i!} \cdot \frac{\theta^{\alpha-1} e^{-\beta\theta}}{\Gamma(\alpha)}$$

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

$$\boxed{\propto \theta^{(\alpha+\sum_i x_i)-1} e^{-(\beta-n)\theta}} \quad (*)$$

Organizada da forma $(*)$, é fácil perceber que esta posteriori é uma Gama$\left(\alpha + \sum_{i=1}^{n} x_i ; \beta + n\right)$

$= \text{Gama}(\alpha^*, \beta^*)$.

* Média: Média Gama $= \frac{\alpha}{\beta}$ (Def.). Então:

$$\text{Média posteriori} = \frac{\alpha + \sum_i x_i}{\beta + n}$$

* Variância: Variância Gama $= \frac{\alpha}{(\beta)^2}$ (Def.). Então:

$$\text{Var. Posteriori} = \frac{\alpha + \sum_i x_i}{(\beta + n)^2}.$$

ii) $w\bar{x} + (1-w)\,\alpha^{*}\beta^{*-1}$ (**)

$\alpha^{*} = \alpha + \sum_i x_i$; $\beta^{*} = \beta + n$. Se $\bar{x} = \frac{1}{n}\sum_i x_i$,

Então $\sum_i x_i = n\bar{x}$.

$(**) = w\bar{x} + (1-w)\underbrace{(\alpha + n\bar{x})(\beta+n)^{-1}}_{\propto \frac{\alpha^{*}}{\beta^{*}}} \propto \theta^{\alpha+\sum_i x_i - 1} e^{-(\beta+n)\theta}$

; e absorvendo termos;

$\rightarrow \underbrace{w\bar{x}}_{\text{Dados}} + \underbrace{(1-w)\,\alpha\beta^{-1}}_{\text{Priori}}$. O que mostra justamente a posteriori como ponderações da verossimilhança com a priori ponderando para.

iii) Pelo fato de ser Gama, média $= \frac{\alpha}{\beta}$ e Var $\frac{\alpha}{\beta^2}$;

Teremos um domínio da priori, com a var. priori $\rightarrow 0$.

Daí, a posteriori $\approx$ priori!

iv) c tq Var. Posteriori $= \frac{\alpha + \sum_i x_i}{(\beta+n)^2} > \frac{\alpha}{\beta^2}$ (Var. Priori).

Sempre que $\bar{x} > c$.

Sol. $\frac{\alpha + n\bar{x}}{(\beta+n)^2} > \frac{\alpha}{\beta^2}$, $\frac{\alpha + n\bar{x}}{\beta^2 + n^2} > \frac{\alpha}{\beta^2}$, $\frac{\alpha+\bar{x}}{n\beta^2} > \frac{\alpha}{\beta^2}$.

$\rightarrow \frac{\alpha+\bar{x}}{n} > \alpha$; $\alpha + \bar{x} > \alpha n$; $\bar{x} > \alpha n - \alpha$

$\bar{x} > \alpha(n-1)$

Logo, o c crítico será $\underline{c = \alpha(n-1)}$.

Ou seja, $\forall\, n \geq 1$, se a média for maior que α (forma da priori); Então a variância posteriori será maior que a variância priori, levando ao resultado descrevendo em iii).

3) $X_1, \dots, X_n \sim \text{Poisson}(\theta)$;

$\theta \sim \text{Gama}(1, 1)$.

(a) PQ θ:

Posteriori $= \text{Gama}(\alpha+1, \beta+1)$.

Então; $\hat{\theta} = \dfrac{\alpha+1}{\beta+1}$.

(b) P. 0-1 θ:

Moda posteriori $= \dfrac{\alpha+1-1}{\beta+1} = \dfrac{\alpha}{\beta+1}$.

(c) $\bar{X} = 1{,}55$ e $n = 10$, temos.

$\frac{1}{10}\sum_{i=1}^{n} x_i = 1{,}55$; $\sum_i x_i = 15{,}5 = \alpha$; $\alpha^* = \alpha + 1$: $\alpha = 16{,}5$.

e, $\beta^* = \beta + 1 = n + 1 = 10 + 1 = 11$

. Mediana Gama (16.5, 11) Não tem forma fechada, No R : qgamma(0.5, 16.5, 11) = 1,4698...

(4) $n=10$; $\bar{x}=1{,}55$; HPD 95%

No R. qgamma(.025, 16.5, 11); qgamma(.975, 16,5, 11)

$=]\,0{,}8657\ldots, 2{,}3056\ldots[$ será o intervalo HPD 95%.

5) $x_1, \ldots, x_n \sim N(\mu, \phi^{-1})$;

$\mu \sim N(\mu_0, \gamma^{-1})$. ϕ^{-1} conhecida. $\sigma_\mu^2 = \frac{1}{\gamma}$

i) Posteriori de μ.

Como $\mu \sim N$ é a priori conjugada p/ $N(\mu, \phi^{-1})$ ϕ^{-1} conhecido, a Posteriori também será Normal, com média:

$$\frac{\bar{x} \cdot \frac{n}{\sigma^2} + \mu_0 \frac{1}{\sigma_\mu^2}}{\frac{n}{\sigma^2} + \frac{1}{\sigma_\mu^2}}$$

E variância: $\frac{1}{\frac{n}{\sigma^2} + \frac{1}{\sigma_\mu^2}}$; $\sigma_\mu^2 = \frac{1}{\gamma}$, $\sigma^2 = \frac{1}{\phi}$

Logo: $\mu \mid x_1, \ldots, x_n \sim N\left[\underbrace{\frac{\phi \bar{x} n + \mu_0 \gamma}{n\phi + \gamma}}_{=\mu^*}; \frac{1}{n\phi + \gamma}\right]$

ii): $w\bar{x} + (1+w)\mu_0$ (*): Como $E(\mu \mid x_1, \ldots, x_n) = \mu^*$

$$= \frac{\phi \bar{x} n + \gamma \mu_0}{n\phi + \gamma}, \text{ Se } w = \frac{\phi n}{n\phi + \gamma},$$

Temos que a forma (*) será uma ponderação da priori c/ a verossimilhança. (Média amostral + Méd. priori)

iii) S e $\bar{x}_m$ é condicionalmente indep.;

Então a preditiva terá a média da posteriori; e Variância da posteriori mais a variância de

$x_{n+1}, \ldots, x_{n+m} \Rightarrow \dfrac{\sigma^2}{m}$

Logo, a preditiva de $\bar{x}_m$ será:

$$\frac{1}{m}\sum_{i=1}^{m} x_i \,\Big|\, x_1, \ldots, x_n \sim N\left(\frac{\phi \bar{x} n + \mu_0 \gamma}{n\phi + \gamma} ; \frac{1}{n\phi + \gamma} + \frac{\sigma^2}{m}\right)$$

iv) Priori $p(\mu) \propto 1$ $\{$i.é.: $\gamma \to 0 \ (\sigma_\mu^2 \to \infty)\}$

Neste caso, a variância da priori "Explodiu", logo, a posteriori = Verossimilhança.

(Igualando assim paradigmas Bayesiano e Frequentista).

6) $x_1, \ldots, x_n \sim N(\mu, \phi^{-1})$, ϕ^{-1} Conhecida.

$\mu \sim N(\mu_0, \gamma^{-1})$

a) E.B. de μ de:

i) P.Q: Média posteriori $= \dfrac{\phi \bar{x} n + \mu_0 \gamma}{n\phi + \gamma}$ (**)

ii) Perda absoluta: Mediana;

P/ normal; Mediana = Média = (**).

iii) Perda 0-1: Moda:

P/ normal: Moda = Mediana = Média = (**).

(b) HPD $(1-\alpha)\cdot 100\%$ p/ μ

Será: $]\mu - z_{\frac{\alpha}{2}}, \mu - z_{\frac{\alpha}{2}}[$; Onde $z_{\frac{\alpha}{2}}$ é o quantil da normal padrão (0,1) p/ alfa desejado.

Lista 3.2

1) $X \sim Unif(\theta, \theta+1) \{-\infty < \theta < \infty\}$

(a) Def. Uniforme (a,b); $a \neq b$:

Se $b = a+1$; $Unif(a, a+1) = Unif(a,b)$

Se $a = \theta$; $Unif(\theta, \theta+1) = Unif(a,b)$

Logo, dado que θ seja constante, θ será apenas um parâmetro de locação de X; sem nenhuma perda de propriedade da distribuição $\square$

(b) Com priori $p(\theta) \propto 1$, a posteriori de θ será somente a verossimilhança. I.e., para uma amostra suficiente e ordenada de forma crescente; $x^{(1)}, \dots, x^{(n)}$;

$$\mathbb{P}(\theta \mid x_1, \dots, x_n) : \begin{cases} x^{(1)} \leq \theta \leq x^{(n)} - 1 = 1 \\ 0, \text{ c.c.} \end{cases}$$; Que é uma uniforme a posteriori.

(c) PQ = Média posteriori (Uniforme) $= \dfrac{(x^{(n)}-1) + x^{(1)}}{2}$

i)

* Perda absoluta: Mediana posterior
= Mediana uniforme
= Média uniforme

$$= \frac{\left(X^{(n)} - 1\right) + X^{(1)}}{2}$$

ii) Intervalo HPD de θ $100(1-\alpha)\%$:

$$= \left] X^{(1)} - \frac{\alpha}{2}\left(X^{(1)} - \left(X^{(n)} - 1\right)\right) ; X^{(n)} - 1 + \frac{\alpha}{2}\left(X^{(1)} - \left(X^{(n)} - 1\right)\right)\right[$$

Visto que a uniforme é uma distribuição simétrica.

iii) $\mathbb{P}(\theta > \theta_0 \mid X_1, \ldots, X_n) = \dfrac{X^{(1)} - \theta_0}{X^{(1)} - \left(X^{(n)} - 1\right)} = 1$

$\Longleftrightarrow X^{(1)} \leq \theta_0 \leq X^{(n)} - 1$; 0 c.c.

2) $X \sim Unif(0, \theta)$; $(0 < \theta < \infty)$

PDF Unif: $\begin{cases} \dfrac{1}{b-a} & ; x \in [a, b] \\ 0 & , c.c \end{cases}$, $\dfrac{1}{\theta - 0} = \dfrac{1}{\theta}$ é a PDF de X.

Então, $\sigma = \theta^{-1}$; $PDF(\sigma) = \theta \sim Unif\left(0, \frac{1}{\theta}\right)$.

Logo, p/ K constante. $Y \sim Unif\left(0, \frac{1}{K\theta}\right)$ terá outra escala. Logo, σ par. de escala.

(b) Priori: $P(\sigma) \propto \sigma^{-1}$.

Verossimilhança: P/uma amostra ordenada de X de tam. n; $x^{(1)}, \dots, x^{(n)}$,

$$L(\theta \mid X_1, \dots, X_n) = \prod_{i=1}^{n} \theta^{-1} I(0 \leq X^{(n)} \leq \theta)$$

$$= \theta^{-n} I(0 \leq X^{(n)} \leq \theta).$$

Com priori $\propto \theta^{-1}$; a Posteriori será:

$$P(\theta \mid X_1, \dots, X_n) \propto \int_{X^{(n)}}^{\infty} \theta^{-n-1} I(0 \leq X^{(n)} \leq \theta)\, d\theta$$

$$\propto \int_{X^{(n)}}^{\infty} \frac{1}{\theta^{n+1}} I[0 \leq X^{(n)} \leq \theta]\, d\theta$$

Sabendo que $\int \frac{1}{x^n}\, dx = \frac{x^{1-n}}{1-n} + C$ Nada Trivial!

$$P(\theta \mid X_1, \dots, X_n) = \frac{n \hat{\theta}^n}{\theta^{n+1}} I[0 \leq X^{(n)} \leq \theta] \left\{ \hat{\theta} = X^{(n)} \right.$$

Que tem a forma de uma Pareto ($X^{(n)}$ escala; n locação).

(c) i) PQ = Média posteriori

$= \text{Média Pareto} = \frac{\alpha\, x_m}{\alpha - 1}; \ \alpha > 1$

$= \frac{n X^{(n)}}{n-1}; \ n > 1$

* P.A. = Mediana Posteriori

= Mediana Pareto = $x_m \sqrt[\alpha]{2}$

= $x^{(n)} \sqrt[n]{2}$

ii) HPD de θ p/ $100(1-\alpha)\%$ de credibilidade

Função quantil Pareto: $x_m (1-p)^{-\frac{1}{\alpha}}$

$= x^{(n)} (1-\alpha)^{-1/n}$ será o quantil superior p/ $100(1-\alpha)\%$; e $x^{(n)}$ será o quantil inferior $\forall \alpha$.

③ $X_1, \ldots, X_n | \theta \sim Ber(\theta)$

Def.: A priori de Jeffreys é proporcional a raiz quadrada do determinante da matriz de informação de Fisher.

$\therefore \pi_\theta(\underset{\sim}{\theta}) \propto \sqrt{\det I_F(\underset{\sim}{\theta})}$

E a matriz de informação de Fisher é a esperança do quadrado da derivada da log-verossimilhança em relação à θ, ou o negativo da esperança da segunda derivada da log verossimilhança com respeito a θ. I.e.:

$$I_F(\underset{\sim}{\theta}) = E\left[\left(\frac{\partial}{\partial \underset{\sim}{\theta}} \ln \mathcal{L}(\theta)\right. \bigg| \theta\right]$$; Em que $\mathcal{L}(\theta)$ é a log-verossimilhança de θ.

$$= -E\left[\frac{\partial^2}{\partial \underset{\sim}{\theta}} \ln \mathcal{L}(\theta) \bigg| \theta\right]$$

Para a Bernoulli(θ), a verossimilhança será:

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

$$= \theta^{\sum_i x_i}(1-\theta)^{n-\sum_i x_i}$$

Logo;

$$\mathcal{L}(\theta) = \sum_i x_i \ln(\theta) + \left(n - \sum_i x_i\right) \ln(1-\theta);$$

Com derivada

$$\frac{\partial}{\partial \theta}\mathcal{L}(\theta) = \frac{\sum_i x_i}{\theta} - \frac{n - \sum_i x_i}{1-\theta} \quad (*)$$

Derivando novamente;

$$\frac{\partial}{\partial \theta}(*) = -\frac{\sum_i x_i}{\theta^2} - \frac{n - \sum_i x_i}{(1-\theta)^2} \quad (**)$$

Sabendo que $\mathbb{E}(X) = \theta$, $X \sim$ Bernoulli, então

$$I_F(\theta) = -\mathbb{E}(**) =$$

$$= -\left[-\frac{n\theta}{\theta^2} - \frac{n - \theta n}{1-\theta^2}\right] = \frac{n}{\theta} + \frac{n(1-\theta)}{(1-\theta)^2} = n\left(\frac{1}{\theta} + \frac{1}{1-\theta}\right)$$

Logo, a priori de Jeffreys será a raiz quadrada de $I_F(\theta)$

$$\Rightarrow \pi_\theta(\theta) \propto \sqrt{n\left(\frac{1}{\theta}+\frac{1}{1-\theta}\right)} = \sqrt{\frac{n}{\theta(1-\theta)}} \overset{n \text{ fixo}}{\propto} \sqrt{\frac{1}{\theta(1-\theta)}} = \frac{1}{\sqrt{\theta(1-\theta)}}$$

$$\propto \frac{1}{\theta^{1/2}(1-\theta)^{1/2}} \sim Beta(1/2, 1/2),$$

Que está no suporte da Beta ($\alpha, \beta > 0$). Logo, esta priori é própria.

(b) Verossimilhança: $\theta^{\sum_i x_i}(1-\theta)^{n-\sum_i x_i}$

Posteriori:

$$\mathbb{P}(\theta | x_1, \ldots, x_n) \propto \frac{\theta^{\sum_i x_i}(1-\theta)^{n-\sum_i x_i}}{\theta^{1/2}(1-\theta)^{1/2}}$$

$$\propto \theta^{\sum_i x_i - 1/2}(1-\theta)^{n-\sum_i x_i - 1/2}$$

Que é uma Beta $\left(\sum_i x_i - 1/2 + 1, n - \sum_i x_i - 1/2 + 1\right)$

PQ: Média Posteriori

$= $ Média Beta $= \frac{\alpha}{\alpha+\beta}$

$$= \frac{\sum_i x_i + 1/2}{\sum_i x_i + 1/2 + n - \sum_i x_i + 1/2} = \frac{\sum_i x_i + 1/2}{n+1}$$

Para achar o HPD $100(1-\alpha)\%$ de credibilidade para esta distribuição, basta tomar os quantis que contenham $100(1-\alpha)$ da densidade da distribuição. Para isso, deve-se saber n e $\sum_1 x_i$.

Lista 4

1) $y_1, \dots, y_n \sim N(\mu, 1/\tau)$ * Amostra | $\lambda_0, \alpha_0, \beta_0 > 0$

$\mu | \tau \sim N(\mu_0, \lambda_0 \tau)$ * Priori μ

$\tau \sim Gama(\alpha_0, \beta_0)$ * Priori τ

(a) · Verossimilhança:

$$L(y|\mu,\tau) = \prod_{i=1}^{n} \sqrt{\frac{\tau}{2\pi}} \exp\left(-\frac{\tau}{2}(y_i-\mu)^2\right)$$

$$= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}(y_i-\mu)^2\right)$$

· Priori $\mu|\tau$

$$L(\mu|\tau) = \prod_{i=1}^{1} \sqrt{\frac{\lambda_0\tau}{2\pi}} \exp\left(-\frac{\lambda_0\tau}{2}(\mu-\mu_0)^2\right)$$

$$= \sqrt{\frac{\lambda_0\tau}{2\pi}} \exp\left(-\frac{\lambda_0\tau}{2}(\mu-\mu_0)^2\right)$$

· Priori τ

$$L(\tau) = \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \tau^{\alpha_0-1} \exp(-\beta_0\tau).$$

Portanto, a posterior será a convolução dessas três f. d. , $\mathcal{J}_q$:

$$\mathbb{P}(\mu, \tau \mid y) \propto \mathbb{P}(y \mid \mu, \tau)\,\mathbb{P}(\mu \mid \tau)\,\mathbb{P}(\tau)$$

$$\propto \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}(y_i-\mu)^2\right) \cdot \left(\frac{\lambda_0 \tau}{2\pi}\right)^{1/2} \exp\left(-\frac{\lambda_0\tau}{2}(\mu-\mu_0)^2\right) \cdot \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}\tau^{\alpha_0-1}\exp(-\beta_0\tau)$$

$$\propto \tau^{n/2}\,\tau^{1/2}\,\tau^{\alpha-1}\beta_0^{\alpha_0}\exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}(y_i-\mu)^2-\frac{\lambda_0\tau}{2}(\mu-\mu_0)^2-\beta_0\tau\right)$$

$$\propto \tau^{\left(\frac{n+1}{2}\right)+\alpha_0-1}\exp\left\{-\frac{\tau}{2}\left[\left(\sum_{i=1}^{n}(y_i-\mu)^2\right)-\lambda_0(\mu-\mu_0)^2\right]-\beta_0\tau\right\}$$

Sabendo que $\int_{-\infty}^{\infty} e^{-x^2}dx=\sqrt{\pi}$; e $\int_{-\infty}^{\infty} e^{-\frac{(x+b)^2}{c^2}}dx = c\sqrt{\pi}$;

definem a integral gaussiana, completaremos o quadrado em exp(...) para tentar integrar

Ignorando por enquanto $\tau^{\frac{(n+1)}{2}+\alpha_0-1}$, temos:

$$\exp\left(-\frac{\tau}{2}\left(\left(\sum_{i=1}^{n}(y_i-\mu)^2\right)-\lambda_0(\mu-\mu_0)^2\right)-\beta_0\tau\right)$$

$$=\exp\left(-\frac{\tau}{2}\left(\sum_{i=1}^{n}y_i^2-2\mu\sum_{i=1}^{n}y_i+n\mu^2+\lambda_0\mu_0^2-2\lambda_0\mu\mu_0+\lambda_0\mu^2\right)-\beta_0\tau\right)$$

$$= \exp\left(-\frac{\tau}{2}\left(\mu^2(n+\lambda_0) - 2\mu\left(\sum_{i=1}^{n} y_i - \lambda_0\mu_0\right) + \sum_{i=1}^{n} y_i^2 + \lambda_0\mu_0^2\right) - \beta_0\tau\right)$$

Completando quadrado

$$\left.\underbrace{\mu^2(n+\lambda_0)} - \underbrace{2\mu\left(\sum_{i=1}^{n} y_i - \lambda_0\mu_0\right)} + \underbrace{\sum_{i=1}^{n} y_i^2 + \lambda_0\mu_0^2}\right\} ax^2+bx+c$$

$x = \mu$; $x^2 = \mu^2$; $b = -2\left(\sum_{i=1}^{n} y_i - \lambda_0 y_0\right)$; c

$a = (n+\lambda_0)$

$x = \mu$

Então, queremos $a(x-h)^2+K$,

$\hookrightarrow \underbrace{(n+\lambda_0)}(\mu-h)^2+K$

$h = \dfrac{-b}{2a}$ e $K = c - \dfrac{b^2}{4a}$,

$$(n+\lambda_0)\left(\mu - \frac{2\left(\sum_{i=1}^{n} y_i - \lambda_0\mu_0\right)}{2(n+\lambda_0)}\right)^2 + \sum_{i=1}^{n} y_i^2 + \lambda_0\mu_0^2 - \frac{4\left(\sum_{i=1}^{n} y_i - \lambda_0\mu_0\right)^2}{4(n+\lambda_0)}$$

$$= (n+\lambda_0)\left(\mu - \frac{\left(\sum_{i=1}^{n} y_i - \lambda_0\mu_0\right)}{n+\lambda_0}\right)^2 + \sum_{i=1}^{n} y_i^2 + \lambda_0\mu_0^2 - \frac{\left(\sum_{i=1}^{n} y_i + \lambda_0\mu_0\right)^2}{n+\lambda_0}$$

* Voltando

$$\exp\left(-\frac{\tau}{2}\left(\mu^2(n-\lambda_0)-2\mu\left(\sum_{i=1}^{n}y_i-\lambda_0\mu_0\right)+\sum_{i=1}^{n}y_i^2+\lambda_0\mu_0^2\right)-\beta_0\tau\right)$$

$$=\exp\left\{-\frac{\tau}{2}(n+\lambda_0)\left(\mu-\frac{\sum_{i=1}^{n}y_i-\lambda_0\mu_0}{n+\lambda_0}\right)^2+\sum_{i=1}^{n}y_i^2+\lambda_0\mu_0^2-\frac{\left(\sum_{i=1}^{n}y_i+\lambda_0\mu_0\right)^2}{n-\lambda_0}-\beta_0\tau\right\}$$

Logo;

$$\int_{-\infty}^{\infty}\tau^{\left(\frac{n+1}{2}\right)+\alpha_0-1}\exp(\,''\,)\,d\mu$$

$$\propto\int_{-\infty}^{\infty}\exp\left(\frac{-\tau(n+\lambda_0)}{2}\left(\mu-\frac{\sum_{i=1}^{n}y_i-\lambda_0\mu_0}{n+\lambda_0}\right)\right)d\mu=(*)$$

Lembrando que p/ integral gaussiana; $\int_{-\infty}^{\infty}e^{-\frac{(x-b)^2}{c^2}}dx=c\sqrt{\pi}$; $b=\frac{\sum_{i=1}^{n}y_i-\lambda_0\mu_0}{n+\lambda_0}$;

$c=\left(\frac{\tau(n+\lambda_0)}{2}\right)^{-1}$;

Temos que

$$(*)=\sqrt{\frac{2\pi}{\tau(n+\lambda_0)}}$$

Esta é a marginal $\mathbb{P}(\mu|D)$.

Para a marginal $\mathbb{P}(\tau|D)$, consideramos constante a marginal anterior, tal que

$$\mathbb{P}(\tau|D) \propto \tau^{\frac{(n+1)}{2}+\alpha_0-1} \exp\left(-\frac{\tau}{2}\left(\sum_{i=1}^{n} y_i^2+\lambda_0\mu_0^2-\frac{\left(\sum_{i=1}^{n} y_i+\lambda_0\mu_0\right)^2}{n+\lambda_0}\right)+\beta_0\right)$$

Gama: $\frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x)$, $x \sim \text{Gama}(\alpha, \beta)$

Se $\tau \sim$ Gama;

$$\beta = \beta_0 + \frac{\left(\sum_{i=1}^{n} y_i^2-\lambda_0\mu_0^2-\frac{\left(\sum_{i=1}^{n} y_i+\lambda_0\mu_0\right)^2}{n+\lambda_0}\right)}{2}$$

$$\alpha = \alpha_0 + \frac{n+1-1}{2} = \alpha_0 + \frac{n}{2}$$

Logo: $\tau|D \sim \text{Gama}\left(\alpha_0+\frac{n}{2}, \beta_0+\sum_{i=1}^{n} y_i^2-\lambda_0\mu_0^2-\frac{\left(\sum_{i=1}^{n} y_i+\lambda_0\mu_0\right)^2}{n+\lambda_0}\right)$

Por fim: Como este é um modelo Normal-Gama;

Então: $\mu \mid \tau, D \sim \text{Normal}\left(\frac{\bar{x}\cdot\frac{n}{\sigma^2}+\mu_0\frac{1}{\sigma_\mu^2}}{\frac{n}{\sigma^2}+\frac{1}{\sigma_\mu^2}}, \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\sigma_\mu^2}}\right)$

$$\Rightarrow \mu|\tau, D \sim N\left(\frac{\bar{y}n+\mu_0\lambda_0}{n+\lambda_0}; \frac{1}{(\lambda_0+n)\tau}\right) \quad \square$$

(b) A posteriori segue gráfica com a priori não informativa, onde que continua no suporte da normal com esta convolução

Neste caso, com a variância da prior escaladada, teremos a convergência do método Bayesiano com o método frequentista.

(2) (c) Por definição, a Dirichlet só converge $\iff \alpha_i > 0 \, \forall i$.

Prova: Com f. d. priori

$$\pi(\theta) \propto \prod_{i=1}^{K} \theta_i^{-1}$$

E verossimilhança Multinomial

$$P(n|\theta) \propto \prod_{i=1}^{K} \theta_i^{n_i}$$

A posteriori será:

$$P(\theta|n) \propto \prod_{i=1}^{K} \theta_i^{n_i-1}$$

Se, para algum $n_i = 0$; teremos algum termo $\frac{1}{n_i}$; que resultará na divergência da integral $\square$

(3) $A = 467$; $W = 315$; $O = 109$; $n = 891$

Então: $X \sim Multinom\left(\frac{467}{891}, \frac{315}{891}, \frac{109}{891}\right)$.

Tomando uma priori (conjugada) Dirichlet "não informativa" $\alpha = (1, 1, 1)$.

Assim, teremos uma posteriori Dirichlet $(467+1, 315+1, 109+1)$.

Logo, $P(\theta | x) \sim Dirichlet(468, 316, 110)$.

(a) $P(\theta_i) > 0,5 \Rightarrow$ Via simulação no R, podemos fazer:

```
> DirichletReg::rdirichlet(10000, c(468, 316, 110))
```

Que irá retornar uma matriz de proporções. Observando estas, vimos que a probabilidade simulada de alguma proporção ser $\geq 0,5$ é $\approx 0,9183$. Dadas, todas de θ_1, i.e., Agnelo (Seed utilizada: 150167636 [minha matrícula])

(b) $P(\theta_1) > 0,5 \approx 0,92$, conforme descrito acima.

(c) Tomando a média das simulações acima, vemos que

* Média Agnelo $\approx 0,5232744$
* Média Waliam $\approx 0,353678$

$\rightarrow$ Diferença Média $\approx 0,17$.

Lista 5

(2) $\eta \sim N(0,1)$; $\eta = \ln\left(\frac{p}{1-p}\right)$

logo, $p = \dfrac{e^{\eta}}{1+e^{\eta}}$

Como a verossimilhança da Bernoulli é dada por:

$$P(y|p) = \binom{n}{y} p^{y} (1-p)^{n-y}$$

Fixando $n = 12$; $y = 9$;

$$P(9|p) = \binom{12}{9} p^{9} (1-p)^{3}$$

$$= \frac{12!}{(12-9)!\,9!} p^{9} (1-p)^{3}$$

$$= \frac{12 \cdot 11 \cdot 10 \cdot \cancel{9!}}{3!\;\cancel{9!}} p^{9} (1-p)^{3}$$

$$= 220\, p^{9} (1-p)^{3}$$

Que neste caso, será

$$P(9) = 220 \cdot \frac{e^{\eta+9}}{1+e^{\eta}} \cdot \left(1 - \frac{e^{\eta}}{1+e^{\eta}}\right)^{3} ; \eta \sim N(0,1)$$

Implementar no R seria simples como:

P/ verificar $\sigma^2 = 4, \sigma^2 = 9$; basta setar > sd = 2 (ou) > sd = 3.

```
> sd = 1 ; eta = rnorm(1, sd = sd)
> 220*((exp(eta)/(1+exp(eta)))^9)*((1-(exp(eta)/ 1+exp(eta))))^3)
```

P/ verificar para $\sigma^2 = 4$ ou $\sigma^2 = 9$; basta mudar o valor de sd:

'> sd = 2' | ou '> sd = 3'

Notamos o que foi garantido ao longo do curso sobre posterioris; ou seja, quanto mais alta é a variância da priori, menos informativa ela é. Neste caso, os valores da posteriori são cada vez mais incertos.