Aplicações: Parte 1

Prof. Caio Azevedo

Variância conhecida

- Seja $X_1|\boldsymbol{\theta}, ..., X_n|\boldsymbol{\theta}, \boldsymbol{\theta} = (\mu, \sigma^2)$ uma amostra aleatória de $X|\boldsymbol{\theta} \sim N(\mu, \sigma^2)$.
- Se σ^2 conhecido, e $\mu \sim N(\alpha, \psi)$, (família conjugada) então $\mu|\mathbf{x} \sim N(\psi^* \alpha^*, \psi^*)$, em que

$$\psi^* = \left(\frac{n}{\sigma^2} + \frac{1}{\psi}\right)^{-1}; \alpha^* = \left(\frac{\alpha}{\psi} + \frac{n\overline{x}}{\sigma^2}\right)$$

- Distribuição preditiva a priori de uma única observação $X'|\mathbf{x} \sim N(\alpha, \sigma^2 + \psi)$
- Preditiva à posteriori à posteriori de uma única observação $X'_*|\mathbf{x} \sim N(\widehat{\psi}\widehat{\alpha}, \widehat{\psi} + \sigma^2)$.

Média conhecida

- Seja $X_1|\boldsymbol{\theta}, ..., X_n|\boldsymbol{\theta}, \boldsymbol{\theta} = (\mu, \sigma^2)$ uma amostra aleatória de $X|\boldsymbol{\theta} \sim N(\mu, \sigma^2)$.
- Se μ conhecido, e $\sigma^2 \sim IG(r, \gamma)$, (família conjugada) então $\sigma^2|\mathbf{x} \sim IG(r^*, \gamma^*)$, em que

$$r^* = \frac{n}{2} + r; \gamma^* = \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2 + \gamma$$

- Distribuição preditiva a priori para uma única observação $p(x'|xb) = \frac{\gamma^r}{\sqrt{2\pi}\Gamma(r)} \frac{\Gamma\left(\frac{1}{2}+r\right)}{\left[\frac{1}{2}(x'-\mu)^2+\gamma\right]^{r+1/2}}$.
- Distribuição preditiva à posteriori para uma única observação $p(x'|xb) = \frac{(\gamma^*)^{r^*}}{\sqrt{2\pi}\Gamma(r^*)} \frac{\Gamma\left(\frac{1}{2}+r^*\right)}{\left[\frac{1}{2}(x'-\mu)^2+\gamma^*\right]^{r^*+1/2}}$.

Ambos os parâmetros desconhecidos

- Família conjugada (normal inversa gama)

$$\begin{aligned} \mu|\sigma^2 &\sim N(\alpha, \sigma^2/\nu) \\ \sigma^2 &\sim \mathrm{IG}(r, \gamma) \end{aligned}$$

- Posteriori conjunta

$$\begin{aligned} \mu|\mathbf{x}, \sigma^2 &\sim N(\psi^* \alpha^*, \sigma^2/(\nu^*)) \\ \sigma^2|\mathbf{x} &\sim \mathrm{IG}(r^*, \gamma^*) \end{aligned}$$

em que $\nu^* = \nu + n$, $\gamma^* = \frac{1}{2}\left[\frac{n\nu}{n+\nu}(\overline{x} - \alpha)^2 + (n-1)s^2\right] + \gamma$, $r^* = \frac{n}{2} + r$.

- Além disso, $\mu|\mathbf{x} \sim t_{(2r^*)}\left(\psi^*\alpha^*, \sqrt{\frac{\gamma^*}{r^*\nu^*}}\right)$.

Cont.

- Ou seja,

$$p(\mu|\mathbf{x}) = \frac{\Gamma\left(\frac{2r^*+1}{2}\right)}{\Gamma\left(\frac{2r^*}{2}\right)\Gamma\left(\frac{1}{2}\right)}\left(\sqrt{2r^*}\delta^*\right)^{-1}\left[1 + \frac{(x-\psi^*\alpha^*)^2}{2r^*(\delta^*)^2}\right]^{-\frac{2r^*+1}{2}}$$

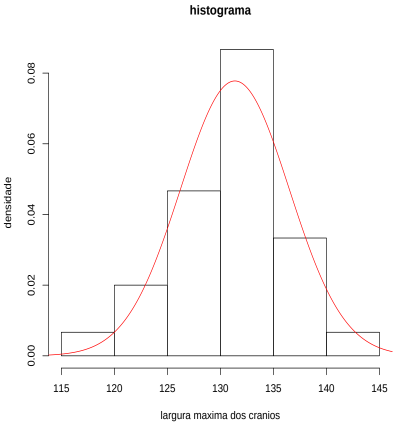
$$\delta^* = \sqrt{\frac{\gamma^*}{r^*\nu^*}}$$

- Distribuição preditiva à posteriori para uma única observação $X'|\mathbf{x} \sim t_{(2r^*)}\left(\alpha^*, \sqrt{\frac{\gamma^*}{r^*\nu^{**}}}\right)$, em que $\nu^{**} = \frac{\nu^*}{1+\nu^*}$.

Dados

- O conjunto de dados se refere à $n = 30$ observações nas quais as larguras máximas de crânios humanos, datadas do período pré-dinástico, foram medidas.
- Estimar a média e a variância.

Histograma e Box Plot

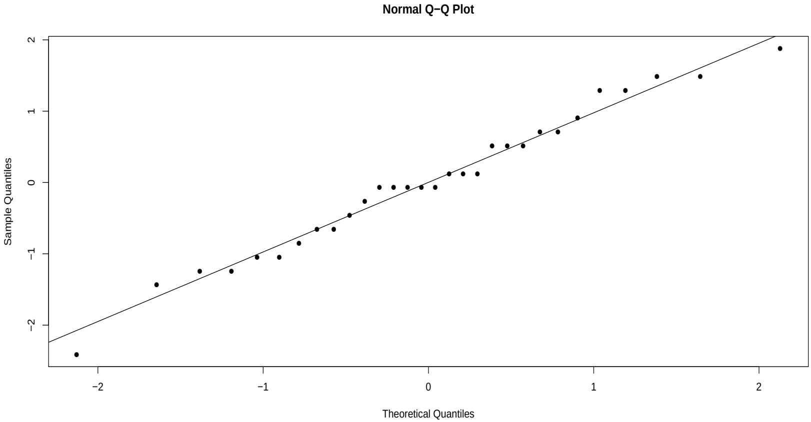


Análise Descritiva

- Medidas resumo.

Medida resumo	Estimativa
Média	131,37
Mediana	131,00
Desvio-padrão	5,12
CV(%)	3,90
Mínimo	119,00
Máximo	141,00

- p-valor para o teste de Kolmogorov - Smirnov (normalidade) : 0,8965


Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
−2
−1
0
1
2

Análise Inferencial: Suponda a variância conhecida

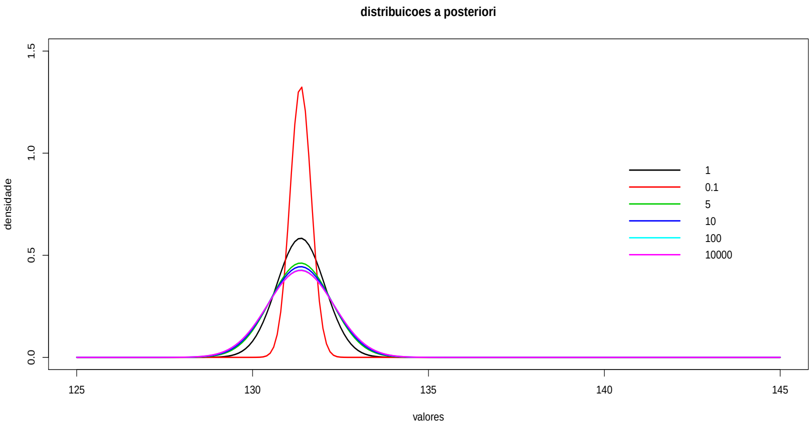
- Como escolher os valores dos hiperparâmetros (α, ψ)?
- Informações prévias.
- Estimativas de máxima verossimilhança, momentos etc.
- Inferência Bayesiana empírica:

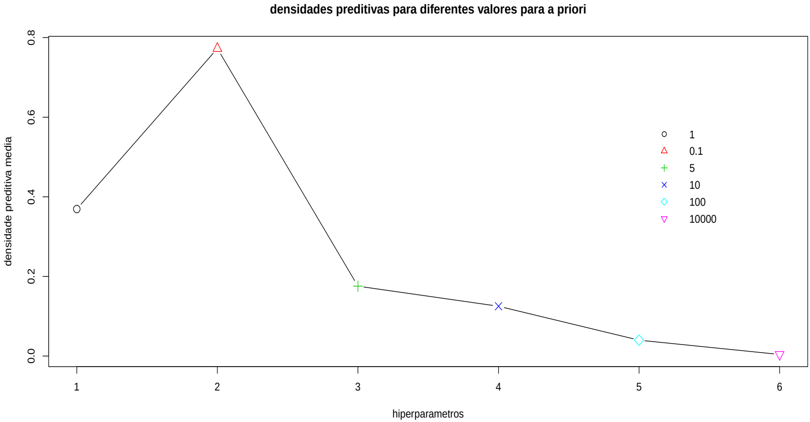
$$p(\mathbf{x}|\alpha, \psi) = \int_{\mathcal{R}} p(\mathbf{x}|\mu)p(\mu|\alpha, \psi)d(\mu)$$

- Definido um conjunto de valores, podemos escolher um par através de algum critérios de comparação de modelos (fator de Bayes, distribuições preditivas etc). Nesse caso, não podemos considerar prioris impróprias.

Continuação

- Assumamos que $\sigma^2 = 26,31$ (como valor verdadeiro).
- Estimativa de máxima verossimilhança para $\widetilde{\alpha} = 131,37$ (média amostral) e diferentes valores para ψ.
- $\psi \in \{1, 0.1, 5, 10, 100, 1000\}$
- Distribuição preditiva para $\overline{X}$, $\overline{X} \sim N(\alpha, \psi + \sigma^2/n)$.
- Calcular os Fatores de Bayes $F_{ij}(\mathbf{x}) = \frac{p(\mathbf{x}|M_i)}{p(\mathbf{x}|M_j)}$, considerando a distribuição preditiva acima sob cada modelo.


distribuicoes a posteriori
densidade
0.0
0.5
1.0
1.5
125
130
135
140
145
valores
1
0.1
5
10
100
10000


densidades preditivas para diferentes valores para a priori
densidade preditiva media
0.0
0.2
0.4
0.6
0.8
1
2
3
4
5
6
hiperparametros
1
0.1
5
10
100
10000

Inferência Bayesiana

- Estimativas Bayesianas.

Hiperparâmetros	EAP	EPAP	$IC_B(95\%)$
$(131,37;1)$	131,37	0,68	[130,02;132,70]
$(131,37;0.01)$	131,37	0,30	[130,78;131,95]
$(131,37;5)$	131,37	0,86	[129,67;133,06]
$(131,37;10)$	131,37	0,90	[129,61;133,12]
$(131,37;100)$	131,37	0,93	[129,53;133,19]
$(131,37;1000)$	131,37	0,94	[129,53;133,20]

Análise Inferencial: Suponda a média conhecida

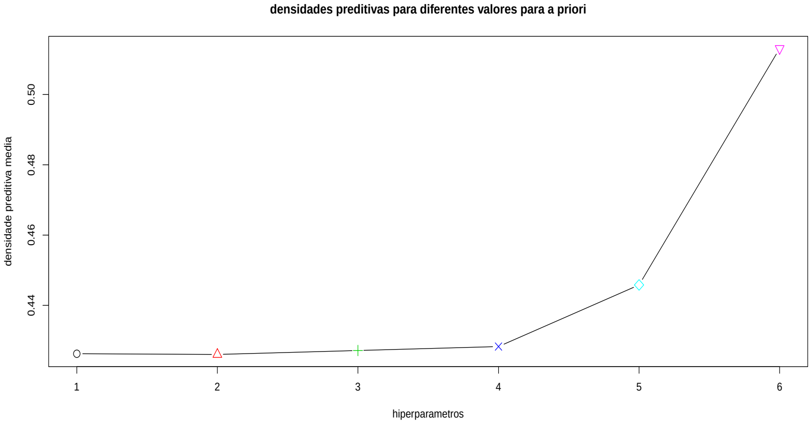
- Como escolher os valores dos hiperparâmetros (r, γ)?
- Informações prévias.
- Estimativas de máxima verossimilhança, momentos etc.
- Inferência Bayesiana empírica:

$$p(\mathbf{x}|r, \gamma) = \int_{\mathcal{R}^+} p(\mathbf{x}|\sigma^2) p(\sigma^2|r, \gamma) d(\sigma^2)$$

- Definido com conjunto de valores, escolher um par através de algum critérios de comparação de modelos (fator de Bayes, distribuições preditivas etc). Nesse caso, não podemos considerar prioris impróprias.

Continuação

- Assumamos que $\mu = 131,37$ (valor verdadeiro).
- Note que $\mathcal{E}(\sigma^2) = \frac{\gamma}{r-1}$ e $\mathcal{V}(\sigma^2) = \frac{\gamma^2}{(r-1)(r-2)}$. Podemos fixar os hiperparâmetros para valores de interesse de $\mathcal{E}(\sigma^2)$ e $\mathcal{V}(\sigma^2)$.
- De fato, neste caso, tem-se que $r = \frac{\mathcal{E}(\sigma^2)^2}{\mathcal{V}(\sigma^2)} + 2$ e $\gamma = \mathcal{E}(\sigma^2)\left[\frac{\mathcal{E}(\sigma^2)^2}{\mathcal{V}(\sigma^2)} + 1\right]$.
- Assim fixemos $E(\sigma^2) = 26,31$ (estimativa de máxima verossimilhança) e escolhamos alguma valor para $\mathcal{V}(\sigma^2) \in \{1, 0.1, 5, 10, 100, 1000\}$
- Distribibuição preditiva de $\overline{X}, \overline{X} \sim t_{2r}\left(\mu, \sqrt{\frac{\gamma}{rn}}\right)$.


densidades preditivas para diferentes valores para a priori
densidade preditiva media
0.44
0.46
0.48
0.50
1
2
3
4
5
6
hiperparametros

Inferência Bayesiana

- Estimativas Bayesianas.

Hiperparâmetros	EAP	EPAP	$IC_B(95\%)$
$(694, 17; 18236, 84)$	26,29	0,99	[24,42;28,30]
$(6923, 74; 182131, 66)$	26,31	0,32	[25,70;26,94]
$(140, 44; 3668, 42)$	26,22	2,12	[22,40;30,70]
$(71, 22; 1847, 37)$	26,15	2,85	[21,16;32,31]
$(8, 92; 208, 42)$	25,74	5,50	[17,14;38,51]
$(2, 70; 44, 52)$	25,52	6,44	[15,88;40,84]

Análise Inferencial: Ambos desconhecidos

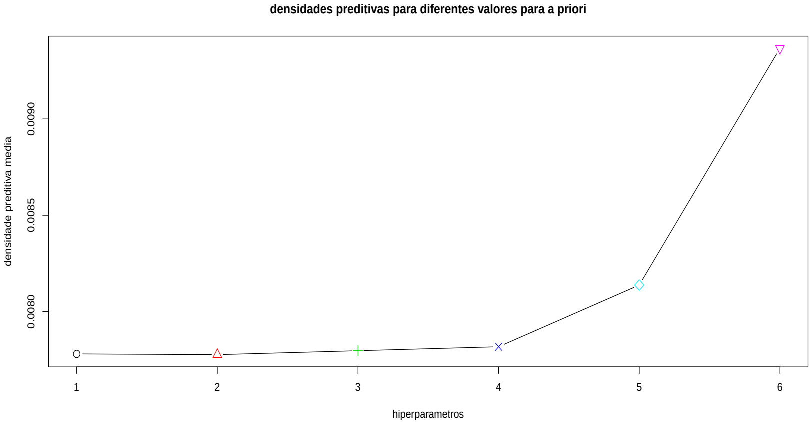
- Como escolher os valores dos hiperparâmetros (α, ν, r, γ)?
- Informações prévias.
- Estimativas de máxima verossimilhança, momentos etc.
- Inferência Bayesiana empírica:

$$p(\mathbf{x}|\alpha, \nu, r, \gamma) = \int_{\mathcal{R}^+} \int_{\mathcal{R}} p(\mathbf{x}|\mu, \sigma^2) p(\mu|\alpha, \nu, \sigma^2) p(\sigma^2|r, \gamma) d(\mu) d(\sigma^2)$$

- Definido com conjunto de valores, escolher um par através de algum critérios de comparação de modelos (fator de Bayes, distribuições preditivas etc). Nesse caso, não podemos considerar prioris impróprias.

Continuação

- Já vimos quais são os valores dos hiperparâmetros mais apropriados para cada uma das duas situações anteriores.
- $\alpha = 131,37$, $\nu = 100$ (equivalente à $\psi = 100$), $r = 6923,74$, $\gamma = 182131,66$. Usar os mesmos valores?
- Distribibuição preditiva de $\overline{X}, \overline{X} \sim t_{2r}\left(\alpha, \sqrt{\frac{\gamma}{r\left(\frac{1}{n}+\frac{1}{\nu}\right)^{-1}}}\right)$.
- Neste caso: $EAP(\mu) = 131,67$, $EPAP(\mu) = 0,67$, $IC_B[\mu; 0,95] = [130,05; 132,68]$.
- Neste caso: $EAP(\sigma^2) = 26,36$, $EPAP(\sigma^2) = 0,32$, $IC_B[\sigma^2; 0,95] = [25,74; 26,99]$.


densidades preditivas para diferentes valores para a priori
densidade preditiva media
0.0080
0.0085
0.0090
1
2
3
4
5
6
hiperparametros