

Aula de Exercícios 3: Mais sobre inferência bayesiana, análise de dados e métodos numéricos

Prof. Caio Azevedo

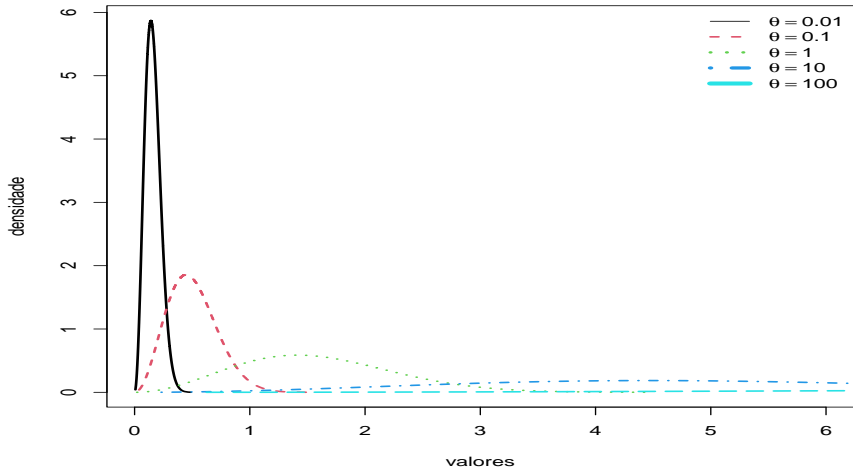
Distribuição de Maxwell

- Voltando aos dados sobre turbinas ([aqui](#)), consideraremos uma outra distribuição para dados positivo.
- Se $X|\theta \sim \text{Maxwell}(\theta)$, então

$$\sqrt{\frac{2}{\pi}} \frac{x^2}{\theta^3} \exp\left(\frac{-x^2}{2\theta^2}\right) \mathbb{1}_{\mathcal{R}^+}(x).$$

- Temos que: $\mathcal{E}(X|\theta) = 2\theta\sqrt{\frac{2}{\pi}}$ e $\mathcal{V}(X|\theta) = \frac{\theta^2(3\pi - 8)}{\pi}$.
- Vamos discutir sobre como realizar IB para θ .

Densidades da distribuição de Maxwell



Distribuição de Maxwell

- Apesar de ser natural fazer inferência diretamente para o parâmetro de interesse, as vezes pode ser mais útil (fácil, menos custoso etc...) considerar uma reparametrização, obter os resultados para esta e, depois, retornar para o parâmetro de interesse.
- Com efeito, nesse caso, não parece fácil obter uma família de prioris conjugadas para θ , com efeito, note que

$$L(\theta) \propto \theta^{-3n} \exp \left\{ -\frac{1}{\theta^2} \frac{\sum_{i=1}^n x_i^2}{2} \right\}.$$

Distribuição de Maxwell

- No entanto, definindo $\lambda = \theta^2$, temos que:

$$L(\lambda) \propto \lambda^{-3n/2} \exp \left\{ -\frac{1}{\lambda} \frac{\sum_{i=1}^n x_i^2}{2} \right\},$$

que se assemelha ao núcleo de uma distribuição

$IG(3n/2 - 1, \sum_{i=1}^n x_i^2/2)$. Assim, a família conjugada de prioris para o modelo $Maxwell(\lambda)$, é a distribuição $IG(a, b)$.

- Posteriori:

$$\begin{aligned} p(\lambda|\mathbf{x}) &\propto \lambda^{-3n/2} \exp \left\{ -\frac{1}{\lambda} \frac{\sum_{i=1}^n x_i^2}{2} \right\} \lambda^{a-1} \exp \left\{ -\frac{b}{\lambda} \right\} \mathbb{1}_{\mathcal{R}^+}(\lambda) \\ &= \lambda^{-3n/2-a-1} \exp \left\{ -\frac{1}{\lambda} \left(\frac{\sum_{i=1}^n x_i^2}{2} + b \right) \right\} \mathbb{1}_{\mathcal{R}^+}(\lambda) \end{aligned}$$

Distribuição de Maxwell

- Posteriori: $\lambda|\mathbf{x} \sim IG(a^*, b^*)$, em que $a^* = \frac{3n}{2} + a$ e

$$b^* = \frac{\sum_{i=1}^n x_i^2}{2} + b.$$

- Portanto:

$$\mathcal{E}(\lambda|\mathbf{x}) = \frac{b^*}{a^*} = \frac{\sum_{i=1}^n x_i^2/2 + b}{3n/2 + a - 1}.$$

$$DP(\lambda|\mathbf{x}) = \frac{b^*}{(a^* - 1)(\sqrt{a^*} - 2)} = \frac{\sum_{i=1}^n x_i^2/2 + b}{(3n/2 + a - 1)\sqrt{3n/2 + a - 2}}.$$

Distribuição de Maxwell

- $IC(\lambda, \gamma)$. Temos que $2b^* \lambda^{-1} | \mathbf{x} \sim \chi^2_{(2a^*)}$. Assim

$$P(q_1 \leq 2b^* \lambda^{-1} \leq q_2 | \mathbf{x}) = \gamma \leftrightarrow P\left(\frac{2b^*}{q_2} \leq \lambda \leq \frac{2b^*}{q_1} | \mathbf{x}\right) = \gamma, (1)$$

em que $P(X \leq q_1) = \frac{1-\gamma}{2}$ e $P(X \geq q_2) = \frac{1-\gamma}{2}$, $X \sim \chi^2_{(2a^*)}$.

Logo

$$IC(\lambda, \gamma) = \left[\frac{2b^*}{q_2}; \frac{2b^*}{q_1} \right].$$

Distribuição de Maxwell

- Voltando ao parâmetro original (θ), queremos calcular: $\mathcal{E}(\theta|\mathbf{x})$, $DP(\theta|\mathbf{x})$ e $IC(\theta, \gamma)$.
- Notemos que ($\forall k > 0$):

$$\mathcal{E}(\theta^k|\mathbf{x}) = \mathcal{E}(\lambda^{k/2}|\mathbf{x}) = (b^*)^{k/2} \frac{\Gamma(a^* - k/2)}{\Gamma(a^*)}$$

$$\mathcal{E}(\theta|\mathbf{x}) = \mathcal{E}(\lambda^{1/2}|\mathbf{x}) = (b^*)^{1/2} \frac{\Gamma(a^* - 1/2)}{\Gamma(a^*)}$$

$$\mathcal{E}(\theta^2|\mathbf{x}) = \mathcal{E}(\lambda|\mathbf{x}) = \frac{b^*}{a^* - 1}$$

$$\begin{aligned} \mathcal{V}(\theta|\mathbf{x}) &= \frac{b^*}{a^* - 1} - b^* \left[\frac{\Gamma(a^* - 1/2)}{\Gamma(a^*)} \right]^2 \\ &= b^* \left[\frac{1}{a^* - 1} - \left(\frac{\Gamma(a^* - 1/2)}{\Gamma(a^*)} \right)^2 \right] \end{aligned}$$

Distribuição de Maxwell

- No caso do intervalo de credibilidade, temos, de (1), que:

$$\begin{aligned}P(q_1 \leq 2b^* \lambda^{-1} \leq q_2 | \mathbf{x}) &= \gamma \Leftrightarrow P\left(\frac{2b^*}{q_2} \leq \lambda \leq \frac{2b^*}{q_1} | \mathbf{x}\right) = \gamma, \\ &\Leftrightarrow P\left(\sqrt{\frac{2b^*}{q_2}} \leq \lambda^{1/2} \leq \sqrt{\frac{2b^*}{q_1}} | \mathbf{x}\right) = \gamma.\end{aligned}$$

- Logo: $IC(\lambda, \gamma) = \left[\sqrt{\frac{2b^*}{q_2}}; \sqrt{\frac{2b^*}{q_1}} \right]$.

Distribuição de Maxwell

- Pelo **método do Jacobiano**, temos que:

$$\begin{aligned} p_{\theta|\mathbf{x}}(\theta|\mathbf{x}) &= p_{\lambda|\mathbf{x}}(\theta^2|\mathbf{x})2\theta \\ &= \frac{2 \left(\frac{1}{2} \sum_{i=1}^n x_i^2 + b\right)^{\frac{3n}{2}+a}}{\Gamma\left(\frac{3n}{2} + a\right)} \theta^{-3n-2a-1} \\ &\times \exp\left\{-\frac{1}{\theta^2} \left(\frac{\sum_{i=1}^n x_i^2}{2} + b\right)\right\} \mathbb{1}_{\mathcal{R}^+}(\theta) \end{aligned} \quad (2)$$

- Logo, $\theta|\mathbf{x} \sim ???$ (não corresponde à nenhuma distribuição “catalogada”).
- Vamos comparar a posteriori $(\theta|\mathbf{x})$, aproximada numericamente, com sua forma analítica exata (Equação (2)).

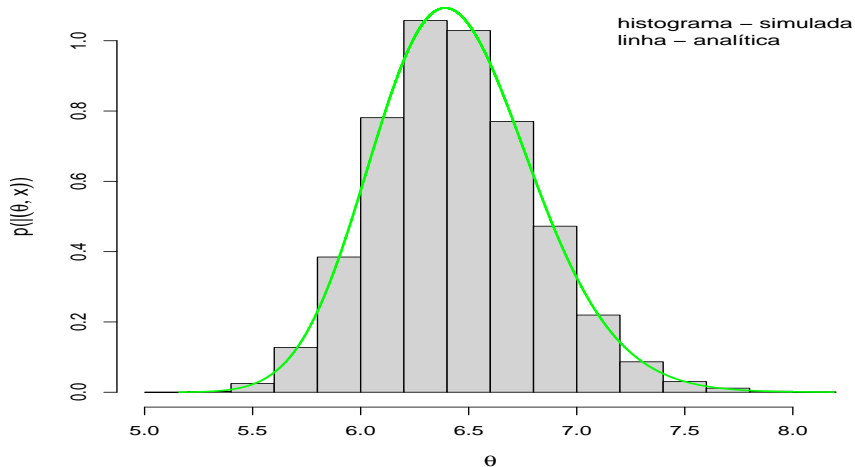
Distribuição de Maxwell

- Primeiramente, vamos determinar a e b , para $\lambda \sim IG(a, b)$.
- Vamos considerar uma priori vaga, com média próxima de alguma estimativa apropriada de λ . Lembre-se de que $\mathcal{E}(\lambda) = \frac{b}{a-1}$ e

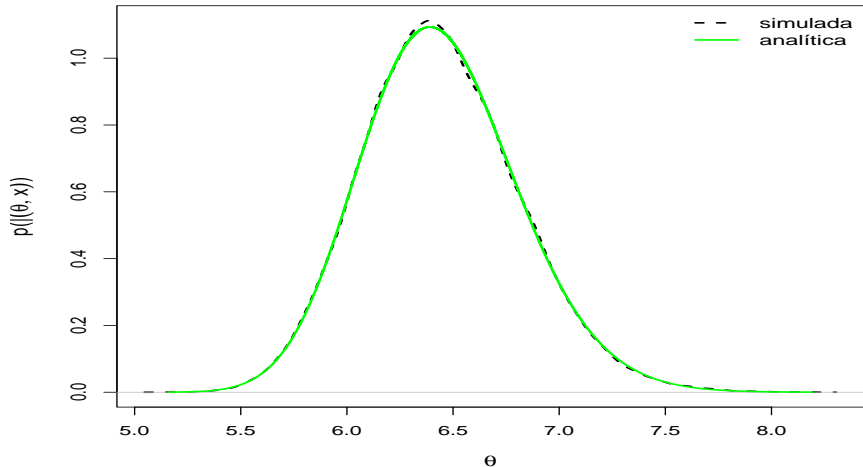
$$\mathcal{V}(\lambda) = \frac{b^2}{(a-1)^2(a-2)}.$$

- Por outro lado, a emm de θ é dada por $\tilde{\theta}_{MM} = \sqrt{\frac{\pi \bar{x}}{2 \bar{2}}}$. Em nosso caso: $\tilde{\theta}_{MM} \approx 7,84 \rightarrow \tilde{\lambda} \approx 61,40$, seria uma estimativa para λ .
- Aproximação numérica: histograma de $\theta^{(r)} = \sqrt{\lambda^{(r)}}$, $r = 1, 2, \dots, R$, simulados de $\lambda | \mathbf{x} \sim IG(a^*, b^*)$.
- Forma analítica: Equação (2).

Posteriori de θ : analítica \times simulada



Posteriori de θ : analítica \times simulada



Comparação entre os modelos gama e Maxwell

- Vamos comparar “somente” as curvas das densidades (gama e Maxwell) uma vez que cálculo da distribuição preditiva para o modelo gama é um pouco complicado.
- Também utilizaremos alguns critérios de informação (estatísticas de comparação de modelos), como AIC, BIC etc, utilizando a log-posteriori no lugar da logverossimilhança ([aqui](#)).
- Para utilização dos critérios de informação, do ponto de vista bayesiano, veja [aqui](#).

Estatísticas de comparação de modelos

- Veremos as seguintes medidas:
 - AIC (“Akaike Information Criteria” - Critério de Informação de Akaike).
 - BIC (“Bayesian Information Criteria” - Critério de informação Bayesiano, também conhecido como SIC).
 - AIC_c (“Corrected Akaike Information Criteria” - critério de Informação de Akaike corrigido).
 - SABIC (“Sample Adjusted BIC” - BIC ajustado pelo tamanho da amostra).
 - HQNIC (“HannanQuinn information criterion” - Critério de informação de de Hanna-Quinn).

Cont.

- As estatísticas mencionadas anteriormente, para o i -ésimo modelo, são dadas, respectivamente, por:

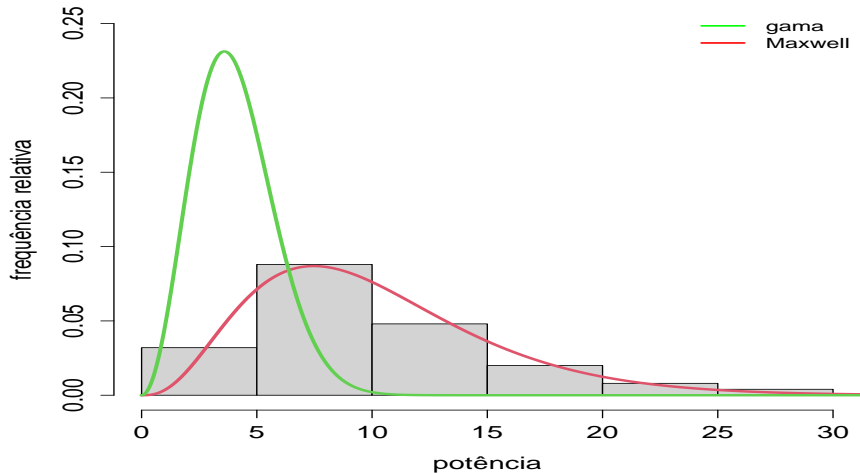
$$AIC_i = -2l_i(\tilde{\theta}_i) + 2k ; BIC_i = -2l_i(\tilde{\theta}_i) + k \ln(n)$$

$$AIC_{C_i} = AIC_i + \frac{2k(k+1)}{n-k-1} ; SABIC_i = -2l_i(\tilde{\theta}_i) + k \ln\left(\frac{n+2}{24}\right)$$

$$HQIC_i = -2l_i(\tilde{\theta}_i) + 2k \ln(\ln(n))$$

que $l_i(\tilde{\theta}_i) \equiv \ln p_i(\theta_i | \mathbf{x})$ denota a log-verossimilhança do i -ésimo modelo avaliada em alguma estimativa bayesiana (p.e. EAP), k é o número de parâmetros e n é o número de observações. Portanto, o modelo que apresentar **os menores valores**, será o modelo “melhor ajustado” aos dados.

Histograma dos dados e densidades estimadas



Critérios de Informação (CI)

CI	gama	Maxwell
AIC	17043,98	21777,77
BIC	17045,89	21779,68
AICc	17044,06	21777,85
SABIC	17042,75	21776,54
HQCIC	17044,71	21778,50
CAICe	17046,89	21780,68
-2log-post	17041,98	21775,77

Comentários

- O modelo gama foi substancialmente melhor do que o modelo Maxwell (densidades e critérios de informação).
- O modelo gama tem dois parâmetros (maior flexibilidade), enquanto que modelo Maxwell tem somente um.
- Exercício: comparar os modelos usando as respectivas distribuições preditivas.