Testes de Hipóteses (Estatísticas) sob a ótica bayesiana (Parte 1)

Prof. Caio Azevedo

Dados reais: estimação do número médio de acidentes

- Descrição: número de acidentes (com algum tipo de trauma para as pessoas envolvidas) em 92 dias durante o ano de 1961, medidos em algumas regiões da Suécia.
- Considerou-se apenas 43 dias, correspondendo àqueles em que não havia limite de velocidade.
- Vamos assumir que

  $X_i|\lambda \overset{i.i.d.}{\sim} Poisson(\lambda), i = 1, ..., 43.$

  que representa o número de acidentes observados no i-ésimo dia.
- Objetivo : testar hipóteses acerca de λ.

Dados reais: estimação do número médio de acidentes (cont.)

- Suponha que desejamos testar $H_0 : \lambda \leq \lambda_0$ vs $H_1 : \lambda > \lambda_0$, $\lambda \in \Theta = (0, \infty)$.
- Isto equivale à $\Theta = \Theta_0 \dot{\cup} \Theta_1$, em que $\Theta_0 = (0, \lambda_0]$ e $\Theta_1 = (\lambda_0, \infty)$.
- Sob o ponto de vista frequentista, podemos considerar a seguinte estatística do teste

$$Z = \frac{\overline{X} - \lambda_0}{\sqrt{\frac{\lambda_0}{n}}}$$

Dados reais: estimação do número médio de acidentes (cont.)

- Se $\lambda = \lambda_0$, $Z \approx N(0,1)$ para n suficientemente grande.
- Pode-se provar que, para que o teste tenha tamanho α, devemos rejeitar H_0, quando $z_c \geq z_{\text{crítico}}$, em que z_c é o valor calculado da estatística Z e

$$P(Z \geq z_{\text{crítico}} | \lambda = \lambda_0) = \alpha$$

considerando-se a distribuição N(0,1).

Testes de hipóteses sob a ótica bayesiana

- Como testar tais hipóteses (e outras) sob a ótica bayesiana?
- Essencialmente, a teoria de testes de hipóteses estatísticas, sob a ótica Bayesiana, baseia-se em calcular e comparar as probabilidades de ocorrência de cada hipótese à priori e à posteriori.
- Pode-se inserir na priori graus de credibilidade para cada hipótese, além de informações sobre o próprio parâmetro.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Beta

$$p_1(x|a,b) = \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}\mathbb{1}_{(0,1)}(x)$$

- Beta (incluindo os valores 0 e 1 no suporte)

$$p_2(x|a,b) = \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}\mathbb{1}_{[0,1]}(x)$$

- Mistura de uma beta com uma Bernoulli

$$\begin{aligned} p_3(x|a,b,\theta,\alpha) &= \alpha\frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}\mathbb{1}_{(0,1)}(x) \\ &+ (1-\alpha)\theta^{x}(1-\theta)^{1-x}\mathbb{1}_{\{0,1\}}(x) \end{aligned}$$

Testes de hipóteses sob a ótica bayesiana

- Considere que $\Theta = \Theta_0 \dot{\cup} \Theta_1$ e que queremos testar

$$H_0 : \lambda \in \Theta_0 \text{ vs } H_1 : \lambda \in \Theta_1$$

- Especificamente no exemplo (lembre-se de que, neste caso, λ é uma vac), suponha que Θ_0 e Θ_1 são conjuntos infinitos não enumeráveis, por exemplo
  $\Theta_0 = (0, \lambda_0]$ e $\Theta_1 = (\lambda_0, \infty)$.
- Podemos testar tais hipóteses comparando-se as probabilidades à priori e a posteriori de cada uma delas serem verdadeiras.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Defina:
  - Probabilidades à priori (para cada hipótese): $P(\lambda \in \Theta_0) = P(H_0)$ e $P(\lambda \in \Theta_1) = P(H_1)$
  - Probabilidades à posteriori (para cada hipótese): $P(\lambda \in \Theta_0|\mathbf{x}) = P(H_0|\mathbf{x})$ e $P(\lambda \in \Theta_1|\mathbf{x}) = P(H_1|\mathbf{x})$.
  - Chance à priori (entre as hipóteses): $O(H_1, H_0) = \dfrac{P(H_1)}{P(H_0)}$.
  - Chance à posteriori (entre as hipóteses): $O(H_1, H_0|\mathbf{x}) = \dfrac{P(H_1|\mathbf{x})}{P(H_0|\mathbf{x})}$.
  - Fator de Bayes: $B(\mathbf{x}) = \dfrac{O(H_1, H_0|\mathbf{x})}{O(H_1, H_0)} = \dfrac{P(H_1|\mathbf{x})P(H_0)}{P(H_0|\mathbf{x})P(H_1)}$.
- Em geral, se $O(H_1, H_0|\mathbf{x}) > c_1$ e/ou $B(\mathbf{x}) > c_2$, $c_i > 0, i = 1, 2$, rejeita-se H_0.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Questões:

1. Como definir c_1, c_2 para que tenhamos uma probabilidade razoável de tomar a decisão correta?
2. Quando se utilizam prioris impróprias, à rigor, não se pode utilizar $O(H_1, H_0)$ e $B(\mathbf{x})$. Pode-se usar $O(H_1, H_0|\mathbf{x})$, desde que a posteriori seja própria.
3. Se λ corresponde à uma variável aleatória contínua (ou mista) e Θ_0 e/ou Θ_1 forem conjuntos finitos ou infinitos enumeráveis, os procedimentos acima podem se tornar inviáveis.
4. Se λ for discreto, em geral, podemos usar o procedimento acima sem problemas.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Em relação à questão 1), uma sugestão para o fator de Bayes é:

Valor	Evidência a favor de H_1
< 1	Contra
$[1; 3)$	Leve
$[3; 10)$	Moderada
$[10; 30)$	Forte
$[30; 100)$	Muito forte
≥ 100	Decisiva

a qual também pode ser usada para $O(H_1, H_0|\mathbf{x})$, principalmente quando se está trabalhando com prioris impróprias.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Essencialmente, em relação à questão 3), vamos considerar $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$.
- Contudo, as abordagens que serão discutidas poderão ser estendidas para outros casos.
- Note que, para as hipóteses acima, se λ for uma vac, então $P(H_0) = P(H_0|\mathbf{x}) = 0$ e $P(H_1) = P(H_1|\mathbf{x}) = 1$.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Basicamente, em relação à questão 3), existem quatro possibilidades para resolvê-la:

  1. Considerar que tais hipóteses não têm sentido prático e substituí-las por

  $$H_0 : \lambda \in (\lambda_0 - \Delta, \lambda_0 + \Delta) \text{ vs } H_1 \notin (\lambda_0 - \Delta, \lambda_0 + \Delta),$$

     para algum $\Delta > 0$.

Testes de hipóteses sob a ótica bayesiana (cont.)

2 Considerar que tais hipóteses têm sentido prático, ou seja o ponto λ_0 (ou o conjunto Θ_0) é relevante e, a priori, atribuir massa positiva para ele, por exemplo considerando:

$$p(\lambda) = \alpha h_0(\lambda)\mathbb{1}_{\Theta_0}(\lambda) + (1-\alpha)h_1(\lambda)\mathbb{1}_{\Theta_1}(\lambda)$$

em que $h_i(\lambda)$ é uma priori (não necessariamente uma distribuição de probabilidade) em $\Theta_i, i = 1, 2$. Se $\Theta_0 = \lambda_0$, então

$$p(\lambda) = \alpha \mathbb{1}_{\lambda_0}(\lambda) + (1-\alpha)h_1(\lambda)\mathbb{1}_{\Theta-\lambda_0}(\lambda)$$

em que $\alpha \in (0, 1)$ é conhecido.

Testes de hipóteses sob a ótica bayesiana (cont.)

3. Utilizar (como visto anteriormente), intervalos de credibilidade para testar as hipóteses.
4. Utilizar os testes de significância bayesianos plenos, FBST (*Full Bayesian Significance Tests*), Pereira and Stern (1999).

Observação: O procedimento 2), também pode ser adotado quando os conjuntos (Θ_0, Θ_1) apresentam probabilidade positiva.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Nos concentraremos basicamente nos três seguintes conjuntos de hipóteses:
  - $H_0 : \lambda \leq \lambda_0$ vs $H_1 : \lambda > \lambda_0$.
  - $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda = \lambda_1$.
  - $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$.
- Além disso, à menos que se mencione o contrário, consideraremos prioris próprias (no respectivo espaço paramétrico).

Testes de hipóteses sob a ótica bayesiana (cont.)

- Primeiramente, vamos considerar $H_0 : \lambda \leq \lambda_0$ vs $H_1 : \lambda > \lambda_0$.
- Admita uma priori própria com estrutura usual (sem se considerar uma mistura).
- Defina: $F(\lambda_0) = P(\lambda \leq \lambda_0)$ e $S(\lambda_0) = P(\lambda > \lambda_0)$, respectivamente, a fda à priori e a fds à priori, no ponto λ_0.
- Defina também: $F(\lambda_0|\mathbf{x}) = P(\lambda \leq \lambda_0|\mathbf{x})$ e $S(\lambda_0|\mathbf{x}) = P(\lambda > \lambda_0|\mathbf{x})$, respectivamente, a fda à posteriori e a fds à posteriori, no ponto λ_0.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Assim,

$$O(H_1, H_0) = \frac{S(\lambda_0)}{F(\lambda_0)}, O(H_1, H_0|\mathbf{x}) = \frac{S(\lambda_0|\mathbf{x})}{F(\lambda_0|\mathbf{x})}$$

  e

$$B(\mathbf{x}) = \frac{O(H_1, H_0|\mathbf{x})}{O(H_1, H_0)} = \frac{S(\lambda_0|\mathbf{x})F(\lambda_0)}{F(\lambda_0|\mathbf{x})S(\lambda_0)}.$$

- Tais resultados são aplicáveis se λ for uma va contínua, discreta ou mista.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Voltando ao exemplo do número de acidentes, vamos considerar que $\lambda \sim gama(a, b^{-1})$, em que $a \approx 0,6784$ e $b \approx 0,02605$ (pela análise feita anteriormente e para garantir que a priori seja própria).
- Suponha que queremos testar $H_0 : \lambda \leq 30$ e $H_1 : \lambda > 30$.
- Assim, como n $= 43$, temos que $\lambda|\mathbf{x} \sim gama(a^*, (b^*)^{-1})$, em que $a^* \approx 1120,6784$ e $b^* \approx 43,02605$.
- Podemos utilizar a função *qgamma* do R para calcular $O(H_1, H_0)$, $O(H_1, H_0|\mathbf{x})$ e $B(\mathbf{x})$.
- Neste caso, temos que: $O(H_1, H_0) = 0,4283$; $O(H_1, H_0|\mathbf{x}) < 0,0001$ e $B(\mathbf{x}) < 0,0001$ (não se rejeita H_0).$IC_B(\lambda, 0,95) = [24,54; 27,59]$

Testes de hipóteses sob a ótica bayesiana (cont.)

- Suponha que desejamos testar agora $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda = \lambda_1$. Ou seja, $\Theta_0 = \lambda_0$ e $\Theta_1 = \lambda_1$.
- Neste caso, vamos assumir que

$$p(\lambda) = \alpha \mathbb{1}_{\lambda_0}(\lambda) + (1 - \alpha)\mathbb{1}_{\lambda_1}(\lambda)$$

- Neste caso, precisamos obter uma "nova" posteriori, pois, uma vez que a priori corresponde à uma mistura de distribuições, também o será a posteriori.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Assim, tem-se

$$\begin{aligned} p(\mathbf{x}) &= \sum_{\lambda} p(\mathbf{x}|\lambda)p(\lambda) = p(\mathbf{x}|\lambda_0)p(\lambda_0) + p(\mathbf{x}|\lambda_1)p(\lambda_1) \\ &= p(\mathbf{x}|\lambda_0)\alpha + p(\mathbf{x}|\lambda_1)(1-\alpha) \end{aligned}$$

- Agora, note que $p(\lambda_0|\mathbf{x}) = \frac{p(\mathbf{x}|\lambda_0)p(\lambda_0)}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\lambda_0)\alpha}{p(\mathbf{x})}$. Analogamente para λ_1

Testes de hipóteses sob a ótica bayesiana (cont.)

- Consequentemente,

$$
\begin{aligned}
p(\lambda|\mathbf{x}) &= \frac{1}{p(\mathbf{x})}\left[p(\mathbf{x}|\lambda_0)\alpha\mathbb{1}_{\lambda_0}(\lambda) + p(\mathbf{x}|\lambda_1)(1-\alpha)\mathbb{1}_{\lambda_1}(\lambda)\right] \\
&= p_0(\lambda|\mathbf{x})\mathbb{1}_{\lambda_0}(\lambda) + p_1(\lambda_1|\mathbf{x})\mathbb{1}_{\lambda_1}(\lambda)
\end{aligned}
$$

em que $p_i(\lambda|\mathbf{x}) = \dfrac{p(\mathbf{x}|\lambda_i)p(\lambda_i)}{p(\mathbf{x})}$.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Portanto, temos que:

$$O(H_1, H_0) = \frac{P(H_1)}{P(H_0)} = \frac{P(\lambda = \lambda_1)}{P(\lambda = \lambda_0)} = \frac{1-\alpha}{\alpha}$$

se $\alpha = 1/2$, então $O(H_1, H_0) = 1$.

- Além disso:

$$O(H_1, H_0|\mathbf{x}) = \frac{P(H_1|\mathbf{x})}{P(H_0|\mathbf{x})} = \frac{p(\mathbf{x}|\lambda_1)(1-\alpha)}{p(\mathbf{x})} \frac{p(\mathbf{x})}{p(\mathbf{x}|\lambda_0)\alpha} = \frac{p(\mathbf{x}|\lambda_1)(1-\alpha)}{p(\mathbf{x}|\lambda_0)\alpha}$$

Testes de hipóteses sob a ótica bayesiana (cont.)

- Finalmente, temos que:

$$B(\mathbf{x}) = \frac{O(H_1, H_0|\mathbf{x})}{O(H_1, H_0)} = \frac{p(\mathbf{x}|\lambda_1)}{p(\mathbf{x}|\lambda_0)}$$

neste caso, $B(.)$ coincide tanto como a estatística do teste da razão de verossimilhanças quanto com aquela relacionado ao teste MP.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Voltemos agora às seguintes hipóteses

$$H_0 : \lambda \leq \lambda_0 \text{ vs } H_1 : \lambda > \lambda_0.$$

Ou seja, $\Theta_0 = \{\lambda : \lambda \leq \lambda_0\}$ e $\Theta_1 = \{\lambda : \lambda > \lambda_0\}$.

- Agora, vamos assumir que

$$p(\lambda) = \alpha p_0(\lambda)\mathbb{1}_{\Theta_0}(\lambda) + (1-\alpha)p_1(\lambda)\mathbb{1}_{\Theta_1}(\lambda)$$

e, por simplicidade, considere que $p_i(.)$ é uma fdp em Θ_i, i=1,2.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Portanto, concluímos que

$$p(H_0) = P(\lambda \leq \lambda_0) = \alpha \underbrace{\int_{\Theta_0} p_0(\lambda) d\lambda}_{1} = \alpha$$

$$p(H_1) = P(\lambda > \lambda_0) = (1-\alpha) \underbrace{\int_{\Theta_1} p_1(\lambda) d\lambda}_{1} = 1 - \alpha$$

- Assim, vem que:

$$O(H_1, H_0) = \frac{P(H_1)}{P(H_0)} = \frac{1-\alpha}{\alpha}$$

Testes de hipóteses sob a ótica bayesiana (cont.)

- Por outro lado,

$$
\begin{aligned}
p(\mathbf{x}) &= \int_{\Theta} p(\mathbf{x}|\lambda)p(\lambda)d\lambda = \alpha \int_{\Theta} p(\mathbf{x}|\lambda)p_0(\lambda)\mathbb{1}_{\Theta_0}(\lambda)d\lambda \\
&+ (1-\alpha)\int_{\Theta} p(\mathbf{x}|\lambda)p_1(\lambda)\mathbb{1}_{\Theta_1}(\lambda)d\lambda \\
&= \alpha p_0(\mathbf{x}) + (1-\alpha)p_1(\mathbf{x})
\end{aligned}
$$

em que $p_i(\mathbf{x}) = \int_{\Theta_i} p(\mathbf{x}|\lambda)p_i(\lambda)d\lambda$, é uma tipo de distribuição preditiva à priori em Θ_i.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Logo:

$$\begin{aligned} p(\lambda|\mathbf{x}) &= \frac{1}{p(\mathbf{x})}\left[p(\mathbf{x}|\lambda)p_0(\lambda)\mathbb{1}_{\Theta_0}(\lambda) + p(\mathbf{x}|\lambda)p_1(\lambda)\mathbb{1}_{\Theta_1}(\lambda)\right] \\ &= \frac{p(\mathbf{x}|\lambda)}{p(\mathbf{x})}\left[p_0(\lambda)\mathbb{1}_{\Theta_0}(\lambda) + p_1(\lambda)\mathbb{1}_{\Theta_1}(\lambda)\right] \end{aligned}$$

- Portanto, analogamente às probabilidades à priori, teremos

$$P(H_0|\mathbf{x}) = \frac{\alpha}{p(\mathbf{x})}\int_{\Theta_0} p(\mathbf{x}|\lambda)p_0(\lambda)d\lambda = \frac{\alpha p_0(\mathbf{x})}{p(\mathbf{x})}$$

e

$$P(H_1|\mathbf{x}) = \frac{1-\alpha}{p(\mathbf{x})}\int_{\Theta_1} p(\mathbf{x}|\lambda)p_1(\lambda)d\lambda = \frac{(1-\alpha)p_1(\mathbf{x})}{p(\mathbf{x})}$$

Testes de hipóteses sob a ótica bayesiana (cont.)

- Portanto, temos que

$$O(H_1, H_0|\mathbf{x}) = \frac{(1-\alpha)p_1(\mathbf{x})}{p(\mathbf{x})}\frac{p(\mathbf{x})}{\alpha p_0(\mathbf{x})} = \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})}\frac{1-\alpha}{\alpha}$$

- Finalmente,

$$B(\mathbf{x}) = \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})}$$

- Tais resultados são aplicáveis se λ for uma va contínua, discreta ou mista.
- Consideraremos suas aplicações ao exemplo em questão mais à frente.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Finalmente, vamos considerar as as hipóteses:

$$H_0 : \lambda = \lambda_0 \text{ vs } H_1 : \lambda \neq \lambda_0.$$

- Neste caso, vamos assumir

$$p(\lambda) = \alpha \mathbb{1}_{\{\lambda_0\}}(\lambda) + (1-\alpha) p_1(\lambda) \mathbb{1}_{\Theta_1}(\lambda)$$

em que $\Theta_1 = \Theta - \{\lambda_0\}$.

- Note que, se λ for uma vac e $p_1(.)$ uma densidade em Θ, então

$$p_1(\lambda)\mathbb{1}_{\{\Theta\}}(\lambda) \equiv p_1(\lambda)\mathbb{1}_{\{\Theta_1\}}(\lambda).$$

Testes de hipóteses sob a ótica bayesiana (cont.)

- Neste caso, temos que:

$$P(\lambda \in \Theta_0) = P(\lambda = \lambda_0) = \alpha$$

- Analogamente,

$$\begin{aligned} P(\lambda \in \Theta_1) &= \int_{\Theta} (1-\alpha) p_1(\lambda) \mathbb{1}_{\{\Theta_1\}}(\lambda) d\lambda = (1-\alpha) \underbrace{\int_{\Theta_1} p_1(\lambda) d\lambda}_{1} \\ &= 1-\alpha \end{aligned}$$

- Portanto, $O(H_1, H_0) = \frac{1-\alpha}{\alpha}$

Testes de hipóteses sob a ótica bayesiana (cont.)

- Além disso:

$$\begin{aligned} p(\mathbf{x}) &= \int_{\Theta} p(\mathbf{x}|\lambda)p(\lambda)d\lambda = p(\mathbf{x}|\lambda_0)p(\lambda_0) + (1-\alpha)\int_{\Theta_1} p(\mathbf{x}|\lambda)p(\lambda)d\lambda \\ &= \alpha p(\mathbf{x}|\lambda_0) + (1-\alpha)p_1(\mathbf{x}) \end{aligned}$$

em que $p_1(\mathbf{x}) = \int_{\Theta_1} p(\mathbf{x}|\lambda)p(\lambda)d\lambda$ é um tipo de distribuição preditiva à priori (sob H_1). À rigor, tal distribuição é igual à esta distribuição preditiva à priori obtida considerando-se todo o espaço paramétrico.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Assim,

$$p(\lambda|\mathbf{x}) = \frac{p(\mathbf{x}|\lambda)}{p(\mathbf{x})}\left[\alpha \mathbb{1}_{\{\lambda_0\}}(\lambda) + (1-\alpha)p_1(\lambda)\mathbb{1}_{\Theta_1}(\lambda)\right]$$

- Logo,

$$P(H_0|\mathbf{x}) = \frac{p(\mathbf{x}|\lambda_0)\alpha}{p(\mathbf{x})}$$

$$P(H_1|\mathbf{x}) = \frac{p_1(\mathbf{x})(1-\alpha)}{p(\mathbf{x})}$$

Testes de hipóteses sob a ótica bayesiana (cont.)

- Portanto, $O(H_1, H_0|\mathbf{x}) = \dfrac{p_1(\mathbf{x})}{p(\mathbf{x}|\lambda_0)} \dfrac{1-\alpha}{\alpha}$
- Logo, $B(\mathbf{x}) = \dfrac{p_1(\mathbf{x})}{p(\mathbf{x}|\lambda_0)}$

Testes de hipóteses sob a ótica bayesiana (cont.)

- Voltando ao exemplo, temos que:

$$p(\mathbf{x}|\lambda_0) = \frac{e^{-b\lambda_0}\lambda_0^{n\overline{x}}}{\prod_{i=1}^{n}!x_i}$$

e

$$p_1(\mathbf{x}) = \int_0^{\infty} \frac{e^{-b\lambda}\lambda^{n\overline{x}}}{\prod_{i=1}^{n}!x_i}\frac{b^a}{\Gamma(a)}e^{-\lambda b}\lambda^{a-1}d\lambda$$

- Logo,

$$B(\mathbf{x}) = \frac{b^a}{\Gamma(a)}\frac{\Gamma(n\overline{x}+a)}{(n+b)^{n\overline{x}+a}}\frac{e^{n\lambda_0}}{\lambda_0^{n\overline{x}}}$$

- Neste caso é mais conveniente calcular $\ln B(\mathbf{x})$, assim, temos que $\ln B(\mathbf{x}) = 7,91 > log(100) = 4,61$. Portanto, seguindo o critério apresentado na tabela anterior, tem-se uma evidência "definitiva" em favor de H_1.