Testes de Hipóteses (Estatísticas) sob a ótica bayesiana

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 (já analisados aqui).
- Considerou-se apenas 43 dias, correspondendo àqueles em que não havia limite de velocidade.
- Vamos assumir que $X_i|\lambda \stackrel{i.i.d.}{\sim} Poisson(\lambda), i = 1, ..., 43$.

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

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 $\Theta = \Theta_0 \dot{\bigcup} \Theta_1$, 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

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

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

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

em que $Z \approx N(0,1)$.

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 $\Theta = \Theta_0 \dot{\bigcup} \Theta_1$ 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): $P(\lambda \in \Theta_0|\boldsymbol{x}) = P(H_0|\boldsymbol{x})$ e $P(\lambda \in \Theta_1|\boldsymbol{x}) = P(H_1|\boldsymbol{x})$.
  - Chance à priori (entre as hipóteses): $O(H_1, H_0) = \dfrac{P(H_1)}{P(H_0)}$.
  - Chance à posteriori (entre as hipóteses): $O(H_1, H_0|\boldsymbol{x}) = \dfrac{P(H_1|\boldsymbol{x})}{P(H_0|\boldsymbol{x})}$.
  - Fator de Bayes: $B(\boldsymbol{x}) = \dfrac{O(H_1, H_0|\boldsymbol{x})}{O(H_1, H_0)} = \dfrac{P(H_1|\boldsymbol{x})P(H_0)}{P(H_0|\boldsymbol{x})P(H_1)}$.
- Em geral, se $O(H_1, H_0|\boldsymbol{x}) > c_1$ e/ou $B(\boldsymbol{x}) > c_2$, $c_i > 0, i = 1, 2$, 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 $B(\boldsymbol{x})$. Pode-se usar $O(H_1, H_0|\boldsymbol{x})$, 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 $O(H_1, H_0|\boldsymbol{x})$, principalmente quando se está trabalhando com prioris impróprias.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Essencialmente, em relação à questão 3), vamos considerar

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

- Contudo, as abordagens que serão discutidas poderão ser estendidas para outros casos.
- Note que, para as hipóteses acima, se λ for uma vac, então

$$P(H_0) = P(H_0|\boldsymbol{x}) = 0 \text{ e } P(H_1) = P(H_1|\boldsymbol{x}) = 1.$$

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: $F(\lambda_0|\boldsymbol{x}) = P(\lambda \leq \lambda_0|\boldsymbol{x})$ e $S(\lambda_0|\boldsymbol{x}) = P(\lambda > \lambda_0|\boldsymbol{x})$, respectivamente, a fda à posteriori e a fds à posteriori, no ponto λ_0.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Assim,

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

  e

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

- 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 $\lambda|\boldsymbol{x} \sim gama(a^*, (b^*)^{-1})$, em que $a^* \approx 1120,6784$ e $b^* \approx 43,02605$.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Podemos utilizar a função *pgamma* do R para calcular $O(H_1, H_0)$, $O(H_1, H_0|\boldsymbol{x})$ e $B(\boldsymbol{x})$.
- Neste caso, temos que: $O(H_1, H_0) = 0,4283$; $O(H_1, H_0|\boldsymbol{x}) < 0,0001$ e $B(\boldsymbol{x}) < 0,0001$ (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

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

- Agora, note que

$$
p(\lambda_0|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\lambda_0)p(\lambda_0)}{p(\boldsymbol{x})} = \frac{p(\boldsymbol{x}|\lambda_0)\alpha}{p(\boldsymbol{x})}.
$$

- Analogamente para λ_1.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Consequentemente,

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

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

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:

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

Testes de hipóteses sob a ótica bayesiana (cont.)

- Finalmente, temos que:

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

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

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,

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

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

Testes de hipóteses sob a ótica bayesiana (cont.)

- Logo:

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

- Portanto, analogamente às probabilidades a priori, teremos

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

e

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

Testes de hipóteses sob a ótica bayesiana (cont.)

- Portanto, temos que

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

- Finalmente,

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

- 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, $O(H_1, H_0) = \dfrac{1-\alpha}{\alpha}$.

Testes de hipóteses sob a ótica bayesiana (cont.)

- Além disso:

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

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

Testes de hipóteses sob a ótica bayesiana (cont.)

- Assim,

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

- Logo,

$$P(H_0|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\lambda_0)\alpha}{p(\boldsymbol{x})} \text{ e } P(H_1|\boldsymbol{x}) = \frac{p_1(\boldsymbol{x})(1-\alpha)}{p(\boldsymbol{x})}$$

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

Testes de hipóteses sob a ótica bayesiana (cont.)

- Voltando ao exemplo, temos que:

$$p(\boldsymbol{x}|\lambda_0) = \frac{e^{-n\lambda_0}\lambda_0^{n\overline{x}}}{\prod_{i=1}^{n} x_i!} \text{ e } p_1(\boldsymbol{x}) = \int_0^{\infty} \frac{e^{-n\lambda}\lambda^{n\overline{x}}}{\prod_{i=1}^{n} x_i!} \frac{b^a}{\Gamma(a)} e^{-\lambda b}\lambda^{a-1} d\lambda.$$

- Logo,

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

- Neste caso é mais conveniente calcular $\ln B(\boldsymbol{x})$. Assim, temos que, se $\lambda_0 = 30$, $\ln B(\boldsymbol{x}) = 7,91 > log(100) = 4,61$. Portanto, seguindo o critério apresentado na tabela anterior, tem-se uma evidência "definitiva" em favor de H_1.

Exemplo Bernoulli

- $X_i|\theta \overset{iid}{\sim}$ Bernoulli(θ).
- Família conjugada: $\theta \sim \text{beta}(a,b)$.
- Posteriori: $\theta \sim \text{beta}(a^*,b^*)$, em que $a^* = n\overline{x} + a$ e $b^* = n(1-\overline{x}) + b$.
- Verossimilhança:

$$p(\boldsymbol{x}|\theta) = \theta^{n\overline{x}}(1-\theta)^{n(1-\overline{x})}. \tag{1}$$

- Distribuição preditiva, à priori, da amostra:

$$p(\boldsymbol{x}) = \int_0^1 p(\boldsymbol{x}|\theta)p(\theta)d\theta = \frac{\beta(a^*,b^*)}{\beta(a,b)}.$$

Exemplo Bernoulli

- Distribuição preditiva à posteriori, de uma nova observação (X_{n+1}) :

$$
\begin{aligned}
p(x_{n+1}|\theta) &\equiv p(x|\theta) = \int_0^1 p(x|\theta)p(\theta|\boldsymbol{x})d\theta \\
&= \int_0^1 \theta^x(1-\theta)^{1-x}\frac{\theta^{a^*-1}(1-\theta)^{b^*-1}}{\beta(a^*,b^*)}\mathbb{1}_{\{0,1\}}(x)d\theta \\
&= \frac{1}{\beta(a^*,b^*)}\int_0^1 \theta^{x+a^*-1}(1-\theta)^{1-x+b^*-1}d\theta \\
&= \frac{\beta(x+a^*,1-x+b^*)}{\beta(a^*,b^*)}\mathbb{1}_{\{0,1\}}(x)
\end{aligned}
$$

- Assim, $X_{n+1}|\theta \sim BB(1,a^*,b^*)$, em que $BB(m,\alpha,\beta)$ representa uma distribuição beta binomial de parâmetros (m,α,β).

Exemplo Bernoulli

- Vamos calcular as principais quantidades relativas à testes de hipóteses bayesianos.
- Hipóteses de interesse:
  1. $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$.
  2. $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$.
  3. $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$.
- As prioris não serão, necessariamente, as mesmas.

Exemplo Bernoulli (sob a família conjugada)

- Quantidades a Priori :

$$
\begin{aligned}
F(\theta) &= \int_0^{\theta} \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1} dz, \\
S(\theta) &= \int_{\theta}^{1} \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1} dz.
\end{aligned}
$$

- Quantidades a Posteriori:

$$
\begin{aligned}
F(\theta|\boldsymbol{x}) &= \int_0^{\theta} \frac{1}{\beta(a^*,b^*)} z^{a^*-1}(1-z)^{b^*-1} dz, \\
S(\theta|\boldsymbol{x}) &= \int_{\theta}^{1} \frac{1}{\beta(a^*,b^*)} z^{a^*-1}(1-z)^{b^*-1} dz.
\end{aligned}
$$

Exemplo Bernoulli (sob a família conjugada)

- As quantidades do slide anterior servem para testar o conjunto de hipóteses 2). Com efeito:

$$
\begin{aligned}
O(H_1, H_0) &= \frac{S(\theta_0)}{F(\theta_0)} = \frac{\displaystyle\int_{\theta_0}^{1} \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}dz}{\displaystyle\int_{0}^{\theta_0} \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}dz} \\
&= \frac{\displaystyle\int_{\theta_0}^{1} z^{a-1}(1-z)^{b-1}dz}{\displaystyle\int_{0}^{\theta_0} z^{a-1}(1-z)^{b-1}dz}
\end{aligned}
$$

Exemplo Bernoulli (sob a família conjugada)

- Cont.:

$$
\begin{aligned}
O(H_1, H_0|\boldsymbol{x}) &= \frac{S(\theta_0|\boldsymbol{x})}{F(\theta_0|\boldsymbol{x})} = \frac{\displaystyle\int_{\theta_0}^{1} \frac{1}{\beta(a^*, b^*)} z^{a^*-1}(1-z)^{b^*-1} dz}{\displaystyle\int_{0}^{\theta_0} \frac{1}{\beta(a^*, b^*)} z^{a^*-1}(1-z)^{b^*-1} dz} \\
&= \frac{\displaystyle\int_{\theta_0}^{1} z^{a^*-1}(1-z)^{b^*-1} dz}{\displaystyle\int_{0}^{\theta_0} z^{a^*-1}(1-z)^{b^*-1} dz}
\end{aligned}
$$

Exemplo Bernoulli (sob a família conjugada)

- Com relação ao conjunto de hipóteses 1), vamos assumir que:

$$p(\theta) = \alpha \mathbb{1}_{\{\theta_0\}}(\theta) + (1-\alpha)\mathbb{1}_{\{\theta_1\}}(\theta).$$

- Logo:

$$p(\theta|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\theta)}{p(\boldsymbol{x})}\left[\alpha \mathbb{1}_{\{\theta_0\}}(\theta) + (1-\alpha)\mathbb{1}_{\{\theta_1\}}(\theta)\right]$$

em que $p(\boldsymbol{x}|\theta)$ é dado em (1) e

$$p(\boldsymbol{x}) = p(\boldsymbol{x}|\theta_0)\alpha + p(\boldsymbol{x}|\theta_1)(1-\alpha)$$

Exemplo Bernoulli (sob a família conjugada)

- Cont.: Temos que:

$$p(\theta|\boldsymbol{x}) = \frac{\theta^{n\overline{x}}(1-\theta)^{n(1-\overline{x})}}{p(\boldsymbol{x}|\theta_0)\alpha + p(\boldsymbol{x}|\theta_1)(1-\alpha)} \left[\alpha \mathbb{1}_{\{\theta_0\}}(\theta) + (1-\alpha)\mathbb{1}_{\{\theta_1\}}(\theta)\right].$$

- Portanto, vem que:

$$\begin{aligned} O(H_1, H_0) &= \frac{1-\alpha}{\alpha}, \\ O(H_1, H_0|\boldsymbol{x}) &= \frac{\theta_1^{n\overline{x}}(1-\theta_1)^{n(1-\overline{x})}}{\theta_0^{n\overline{x}}(1-\theta_0)^{n(1-\overline{x})}} \frac{\alpha}{1-\alpha}, \\ B(\boldsymbol{x}) &= \frac{\theta_1^{n\overline{x}}(1-\theta_1)^{n(1-\overline{x})}}{\theta_0^{n\overline{x}}(1-\theta_0)^{n(1-\overline{x})}}. \end{aligned}$$

Exemplo Bernoulli (sob a família conjugada)

- Cont.: Temos que:

$$p(\theta|\boldsymbol{x}) = \frac{\theta^{n\overline{x}}(1-\theta)^{n(1-\overline{x})}}{p(\boldsymbol{x}|\theta_0)\alpha + p(\boldsymbol{x}|\theta_1)(1-\alpha)}\left[\alpha \mathbb{1}_{\{\theta_0\}}(\theta) + (1-\alpha)\mathbb{1}_{\{\theta_1\}}(\theta)\right]$$

- Portanto, vem que:

$$\begin{aligned}
O(H_1, H_0) &= \frac{1-\alpha}{\alpha} \\
O(H_1, H_0|\boldsymbol{x}) &= \frac{\theta_1^{n\overline{x}}(1-\theta_1)^{n(1-\overline{x})}}{\theta_0^{n\overline{x}}(1-\theta_0)^{n(1-\overline{x})}}\frac{\alpha}{1-\alpha} \\
B(\boldsymbol{x}) &= \frac{\theta_1^{n\overline{x}}(1-\theta_1)^{n(1-\overline{x})}}{\theta_0^{n\overline{x}}(1-\theta_0)^{n(1-\overline{x})}}
\end{aligned}$$

Exemplo Bernoulli (sob a família conjugada)

- Com relação ao conjunto de hipóteses 3), vamos assumir que:

$$p(\theta) = \alpha \mathbb{1}_{\{\theta_0\}}(\theta) + p(\theta)(1-\alpha)\mathbb{1}_{\Theta_1}(\theta),$$

em que $p(\theta) = \dfrac{1}{\beta(a,b)}\theta^{a-1}(1-\theta)^{b-1}\mathbb{1}_{\Theta_1}(\theta)$ e $\Theta_1 = (0,1) - \{\theta_0\}$.

- Logo, temos que:

$$p(\theta|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\theta)}{p(\boldsymbol{x})},$$

em que $p(\boldsymbol{x}) = \alpha p(\boldsymbol{x}|\theta_0) + (1-\alpha)\dfrac{\beta(a^*,b^*)}{\beta(a,b)}$.

Exemplo Bernoulli (sob a família conjugada)

- Assim, vem que:

$$
\begin{aligned}
O(H_1, H_0) &= \frac{(1-\alpha)}{\alpha}, \\
O(H_1, H_0|\boldsymbol{x}) &= \frac{p_1(\boldsymbol{x})(1-\alpha)}{p(\boldsymbol{x}|\theta_0)\alpha}, \\
B(\boldsymbol{x}) &= \frac{p_1(\boldsymbol{x})}{p(\boldsymbol{x}|\theta_0)},
\end{aligned}
$$

em que $p_1(\boldsymbol{x}) = \dfrac{\beta(a^*, b^*)}{\beta(a, b)}$ e $p(\boldsymbol{x}|\theta_0) = \theta_0^{n\overline{x}}(1-\theta_0)^{n(1-\overline{x})}$.

Exemplo: proporção de meninas que apresentaram menarca

- Milecer e Szczotka (1966) investigaram a idade do início da menstruação em 3918 garotas de Varsóvia.
- Para 25 médias de idade foram observadas a ocorrência ou não do início de períodos de menstruação em um dado número de adolescentes.
- Como exemplo, consideraremos somente as garotas com idade média de 13,08.
- Veja também aqui.

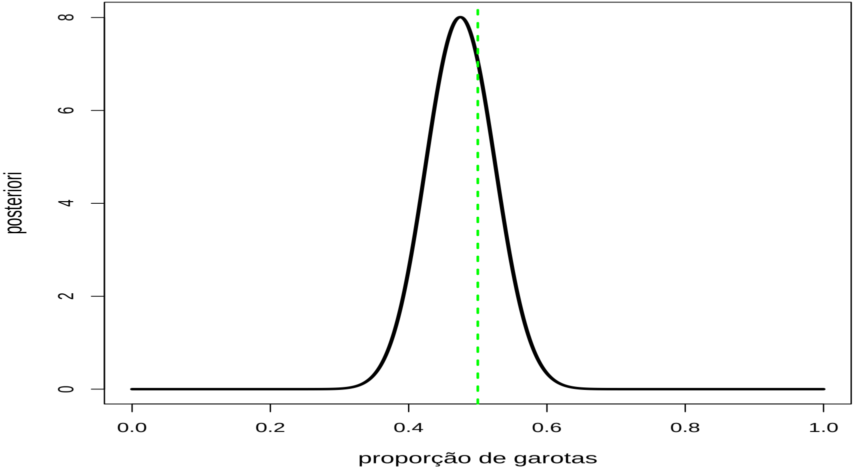
Exemplo: proporção de meninas que apresentaram menarca

- Nesse caso, temos que $X_i|\theta \overset{iid}{\sim} Bernoulli(\theta)$, sendo que $X_i, i = 1, 2, ..., 99$ é igual a 1 se a i-ésima garota apresentou menarca com idade em torno de 13 anos e 0, caso contrário.
- Temos que θ é a proporção de garotas que apresentaram menarca com idade em torno de 13 anos.
- Amostra: especificamente, das n = 99 entrevistadas, $\sum_{i=1}^{99} x_i = 47$ apresentaram menarca.

Cont.

- Hipóteses de interesse
  1. $H_0 : \theta = 0,50$ vs $H_1 : \theta = 0,45$.
  2. $H_0 : \theta \leq 0,50$ vs $H_1 : \theta > 0,50$.
  3. $H_0 : \theta = 0,50$ vs $H_1 : \theta \neq 0,50$.
- Vamos considerar $\alpha = 1/2$ e, quando pertinente, $\theta \sim beta(1,1)$:
  1. Hipóteses 1: $O(H_1, H_0) = 1$, $O(H_1, H_0|\boldsymbol{x}) = B(\boldsymbol{x}) = 1,004$.
  2. Hipóteses 2: $O(H_1, H_0) = 1$, $O(H_1, H_0|\boldsymbol{x}) = 0,4464$, $B(\boldsymbol{x}) = 0,4464$.
  3. Hipóteses 3: $O(H_1, H_0) = 1$, $O(H_1, H_0|\boldsymbol{x}) = B(\boldsymbol{x}) = 0,1417$.
- Com relação às hipóteses 2 e 3, temos fortes evidência à favor de H_0, enquanto que para a hipótese 1, temos, essencialmente, uma indecisão. $HPD(\theta; 0,95) = [0,379; 0.572]$.

Posteriori



Paralelo com a abordagem frequentista

- Considere, novamente, as hipóteses $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1 = \Theta_0^c$. Então, do ponto de vista clássico (frequentista), a mensuração da qualidade é feita em termos dos erros do tipo I e II e a decisão é tomada de acordo com as priporiedades da região crítica (embora também se possa avaliar a qualidade dos testes bayesianos através de critérios frequentistas).
- Do ponto de vista Bayesiano, temos que decidir entre H_0 e H_1 considerando as informações a priori de $\theta(\pi(.))$ e a distribuição a posteriori $\theta(\pi(.|\boldsymbol{x}))$.

Paralelo com a abordagem frequentista

- Como vimos anteriormente, dada a distribuição a posteriori, podemos calcular, por exemplo:

$$P(H_0|\boldsymbol{x}) = P(\theta \in \Theta_0|\boldsymbol{x}) = P(H_0 \text{ é verdadeira}|\boldsymbol{x})$$

$$P(H_1|\boldsymbol{x}) = P(\theta \in \Theta_1|\boldsymbol{x}) = P(H_1 \text{ é verdadeira}|\boldsymbol{x})$$

- Em termos dos conceitos de testes de hipóteses frequentistas a RC (Região Crítica) é da forma

$$RC = \{\boldsymbol{x} \in \mathcal{X} : P(\theta \in \Theta_0^c|\boldsymbol{x}) \geq c\}.$$

- Consequentemente, a região de aceitação é dada por:

$$RC^c = \{\boldsymbol{x} \in \mathcal{X} : P(\theta \in \Theta_0^c|\boldsymbol{x}) < c\}.$$

Paralelo com a abordagem frequentista

- Exemplo: Seja $X_1, ..., X_n$ uma aa de $X \sim N(\theta, \sigma^2)$ e $\theta \sim N(\mu, \tau^2)$, com μ, τ^2, σ^2 conhecidos. Considere as hipóteses:

$$H_0 : \theta \leq \theta_0 \text{ vs } H_1 : \theta > \theta_0.$$

- Neste caso, temos que: $\theta | \boldsymbol{X} = \boldsymbol{x} \sim N((1-\eta)\overline{x} + \eta\mu, \eta\tau^2)$, em que $\eta = \dfrac{\sigma^2}{n\tau^2 + \tau^2}$.

Paralelo com a abordagem frequentista

- Por outro lado, não se rejeita H_0 se

$$
\begin{aligned}
P(\theta \in \Theta_0|\boldsymbol{x}) \geq c &= P(\theta \leq \theta_0|\boldsymbol{x}) \geq c = \Phi\left(\frac{\theta_0 - (1-\eta)\overline{x} - \eta\mu}{\sqrt{n}\tau}\right) \geq c \\
&= \frac{\theta_0 - (1-\eta)\overline{x} - \eta\mu}{\sqrt{n}\tau} \geq \Phi^{-1}(c) \\
&= -\overline{x} \geq \frac{-\theta_0 + \eta\mu}{1-\eta} + \Phi^{-1}(c) \\
&\rightarrow \overline{x} \leq \frac{\theta_0 - \eta\mu}{1-\eta} + \Phi^{-1}(c)
\end{aligned}
$$

- Assim, a RC é dado por:

$$
RC = \left\{\boldsymbol{x} \in \mathcal{R}^n : \overline{x} > \frac{\theta_0 - \eta\mu}{1-\eta} + \Phi^{-1}(c)\right\}
$$