Inferência Estatística com Abordagem Bayesiana

Rosangela Helena Loschi [1]

[1]Departamento de Estatística
Universidade Federal de Minas Gerais

11 de janeiro de 2022

TESTES de HIPÓTESES

- Teste de Jeffreys
- Teste de Significância Bayesiano Completo (FBST)

Teste de Hipóteses

Dentro do enfoque Bayesiano, o problema de decidir sobre qual hipótese $H_0, H_1, \ldots, H_k$ aceitar é conceitualmente mais simples.
Nos dois procedimentos de teste que veremos

- Escolhe-se a hipótese que minimiza uma dada função de perda.

O primeiro procedimento que veremos está baseado no seguinte princípio para tomada de decisão

- Calculamos as probabilidades *a posteriori* $P(H_i \mid \boldsymbol{x})$ de cada uma das hipóteses sob teste.
- Escolhe-se a hipótese mais provável *a posteriori*.
- Há um ponto de corte natural em nosso critério de decisão que é $1/2$.
- Isto não ocorre nos procedimentos clássicos onde devemos especificar o nível de significância para construírmos o critério de decisão.
- Se Θ_k não é unitário, então $P(H_k \mid \mathrm{x}) = \int_{\Theta_k} \pi(\theta \mid \mathrm{x}) d\theta$.

Teste de Hipóteses

Exemplo: Suponha que $X|\mu \overset{iid}{\sim} Normal(\mu, \sigma^2)$, σ^2 conhecido, e que a incerteza *a priori* sobre μ é bem descrita pela distribuição de Jeffreys. Teste as hipóteses $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$, μ_0 conhecido.

Solução: Neste caso, como a distribuição *a priori* é $\pi(\mu) \propto 1$, já sabemos que a distribuição *a posteriori* para μ é

$$\mu \mid \boldsymbol{x} \sim N(\bar{x}, \sigma^2/n).$$

Sendo assim, as probabilidades *a posteriori* de H_0 e H_1 são, respectivamente,

$$P(H_0 \mid \boldsymbol{x}) = P(\mu \leq \mu_0 \mid \boldsymbol{x}) = P\left(Z \leq \frac{n^{1/2}(\mu_0 - \bar{x})}{\sigma}\right) = \Phi\left(\frac{n^{1/2}(\mu_0 - \bar{x})}{\sigma}\right)$$

$$P(H_1 \mid \boldsymbol{x}) = P(\mu > \mu_0 \mid \boldsymbol{x}) = 1 - P\left(Z \leq \frac{n^{1/2}(\mu_0 - \bar{x})}{\sigma}\right) = 1 - \Phi\left(\frac{n^{1/2}(\mu_0 - \bar{x})}{\sigma}\right)$$

Se $P(H_1 \mid \boldsymbol{x}) > P(H_0 \mid \boldsymbol{x})$ decidimos por aceitar H_1.

Que decisao tomar se $\sigma^2 = 400$ e se observamos uma amostra com tamanho $n = 4$ em que $\bar{x} = 106$. Admita que *a priori* $\mu_0 = 100$.

Neste caso, temos que

$$P(H_0 \mid \boldsymbol{x}) = \Phi\left(\frac{100 - 106}{20/2}\right) = \Phi(-0,6) = 0,274 < 0,726 = P(H_1 \mid \boldsymbol{x})$$

Consequentemente, temos evidência a favor de H_1 e devemos aceitá-la.

OBS: 1)Note que a única amostra relevante para a nossa tomada de decisão foi a amostra observada. Amostras que poderiam ter sido observadas mas não o foram não são relevantes na decisão.

Teste de Hipóteses

2)Neste caso, estamos usando uma distribuição *a priori* não-informativa. Será que a decisão concide com a decisão tomada usando métodos clássicos?

- A tendência é pensar que, neste caso, as inferências Clássica e Bayesiana deveriam ser coincidentes pois a distribuição *a priori* é vaga.
- Cálculo do valor-P: Neste caso, a estatística de teste é $\bar{X}$

$$\begin{aligned} \bar{X} &\sim Normal(\mu_0, \sigma^2/n) \\ P - value &= P(\bar{X} > 106 \mid \mu = 100) = 1 - \Phi(0.6) = 0.2776 \end{aligned}$$

- $P - value > \alpha \Rightarrow$ Não se rejeita H_0. Isto ocorre para os níveis de significância convencionais ($\alpha = 0,01, 0,05, 0,10$).
- Temos uma decisão diferente da que tomamos utilizando os métodos Bayesianos.

Teste de Hipóteses

Formalmente, considere a situação habitual em teste de hipóteses em temos duas hipóteses sob teste:

$$H_0 : \theta \in \Theta_0 \;\times\; H_1 : \theta \in \Theta_1$$

onde $\{\Theta_0, \Theta_1\}$ é uma partição do espaço paramétrico Θ.

Neste caso, se $P(H_k)$ representa a probabilidade *a priori* para a hipótese H_k, $k = 0, 1$, *a posteriori* temos:

$$P(H_k \mid \mathrm{x}) = P(\mathrm{x} \mid H_k)P(H_k)/P(\mathrm{x}). \tag{1}$$

- Se $P(H_0 \mid \mathrm{x}) > P(H_1 \mid \mathrm{x})$ então aceitamos H_0.
- Outro critério: Decidimos a favor de H_0 se a *odds a posteriori* dada por
$$O(H_0, H_1 \mid \mathrm{x}) = \frac{P(H_0 \mid \mathrm{x})}{P(H_1 \mid \mathrm{x})}$$
for maior que 1.
- Se a *odds a posteriori* for 1, a hipótese H_0 é tão provável quanto H_1.

Teste de Hipóteses

Considerando as probabilidades dadas em (1) podemos escrever a *odds a posteriori* como

$$
\begin{aligned}
\frac{P(H_0 \mid x)}{P(H_1 \mid x)} &= \frac{P(x \mid H_0)}{P(x \mid H_1)} \times \frac{P(H_0)}{P(H_1)} \qquad (2) \\
O(H_0, H_1 \mid x) &= BF(H_0, H_1) \times O(H_0, H_1).
\end{aligned}
$$

onde

- $O(H_0, H_1) = \frac{P(H_0)}{P(H_1)}$ define a *odds a priori*.
- $BF(H_0, H_1) = \frac{P(x|H_0)}{P(x|H_1)}$ é o fator de Bayes que é visto como a *odds* trazida pelos dados a favor de H_0 contra H_1.

Teste de Hipóteses

- Note que o fator de Bayes é a razão entre as *odds a priori e a posteriori*

$$BF(H_0, H_1) = \frac{O(H_0, H_1 \mid \mathrm{x})}{O(H_0, H_1)}$$

- Termo criado por Jeffreys (1961).

OBS: Se $\Theta_0 = \{\theta_0\}$ e $\Theta_1 = \{\theta_1\}$ são conjuntos unitário, isto é, estamos testando uma hipótese simples contra uma hipótese simples, o Fator de Bayes torna-se

$$BF(H_0, H_1) = \frac{P(\mathrm{x} \mid \theta_0)}{P(\mathrm{x} \mid \theta_1)}$$

que é a razão de verossimilhança usada nos testes clássicos.

Teste de Hipóteses

- Assim, o Fator de Bayes(que é um tipo de razão de verossimilhança) é, mais uma vez, responsável por introduzir a informação dos dados num contexto de teste de hipóteses.

Note da equação (3) que

$$\begin{aligned}
\frac{P(H_0 \mid \mathrm{x})}{1 - P(H_0 \mid \mathrm{x})} &= BF(H_0, H_1)\frac{P(H_0)}{P(H_1)} \\
P(H_0 \mid \mathrm{x}) &= BF(H_0, H_1)P(H_0)/P(H_1) - P(H_0 \mid \mathrm{x})BF(H_0, H_1)P(H_0)/P(H_1) \\
P(H_0 \mid \mathrm{x}) &= \frac{BF(H_0, H_1)P(H_0)/P(H_1)}{1 + BF(H_0, H_1)P(H_0)/P(H_1)}
\end{aligned}$$

$$P(H_0 \mid \mathrm{x}) = \left[1 + \left[BF(H_0, H_1)\frac{P(H_0)}{P(H_1)}\right]^{-1}\right]^{-1} \tag{3}$$

Teste de Hipóteses: Decisões com base no Fator de Bayes

- Em geral, se $BF(H_0, H_1) > 1$ então aceitamos H_0.
  - Se $P(H_0) = P(H_1) \Rightarrow$ aceitar H_0 se $P(H_0 \mid \boldsymbol{x}) > P(H_1 \mid \boldsymbol{x})$.
  - Se $P(H_0) \neq P(H_1) \Rightarrow$ aceitar H_0 se $O(H_0, H1 \mid \boldsymbol{x}) > O(H_0, H_1)$.
- A escala de Jeffreys fornece a força da evidência a favor de H_0 dada pelo fator de Bayes

$BF(H_0, H_1)$	força evidência	$BF(H_0, H_1)$	força evidência
< 1	contra	10 a 30	forte
1 a 3	fraca	30 a 100	muito forte
3 a 10	substancial	> 100	decisiva

Que riscos corremos de tomarmos decisões apenas com base no Fator de Bayes?

Efeito do Fator de Bayes nas decisões

Suponha que para um dado teste de hipóteses tenhamos $BF(H_0, H_1) = 1,875$. Isto significa que a evidência amostral favorece H_0.

- A *odds a posteriori* é $P(H_0|\mathrm{x})/P(H_1|\mathrm{x}) = 1,875$ revelando que *a posteriori* a evidência é a favor de H_0 e concluimos que a chande de H_0 é $1,875$ vezes a chance de H_1 (maior do que era *a priori*).
- Se $P(H_0) = 1/2$, *a priori*, não há evidência a favor de qualquer hipótese. A posteriori, teremos

$$P(H_0 \mid \mathrm{x}) = [1 + 1/(1,875)]^{-1} = 0,65217 > 0,5$$

o que é evidência a favor de H_0.

Efeito do Fator de Bayes nas decisões

- Se $P(H_0) = 1/3$, a *odds a priori* é $P(H_0)/P(H_1) = 1/2$.
  - A evidência *a priori* é contra H_0 e revela que a chance *a priori* de H_1 é 2 vezes a chance *a priori* de H_0.
  - A posteriori, teremos $P(H_0 \mid \mathrm{x}) = [1 + 1/(0,5 * 1,875)]^{-1} = 0,4839 < 0,5$ o que é evidência a favor de H_1.
  - Comparado com a informação *a priori*, há um ganho razoavelmente grande de informação a favor de H_0.
  - A *odds a posteriori* é $P(H_0|\mathrm{x})/P(H_1|\mathrm{x}) = 0,9375$ revelando que *a posteriori* a evidência continual a favor de H_1 mas ao misturamos as informações *a priori* e amostral concluimos que a chande de H_1 é apenas $1,0667$ vezes a chance de H_0 (menor do que era *a priori*).

Efeito do Fator de Bayes nas decisões

- Se $P(H_0) = 1/9 = 0,11$,
  - a *odds a priori* é $P(H_0)/P(H_1) = 1/8$. A evidência *a priori* é fortemente contra H_0 e revela que a chance *a priori* de H_1 é 8 vezes a chance *a priori* de H_0.
  - A posteriori, teremos $P(H_0 \mid \mathrm{x}) = 0,18987$. Ou seja, a evidência *a posteriori* sobre H_0 aumenta se comparado com a evidência *a priori* mas ainda temos que a evidência é a favor de H_1.
  - note que, neste caso, considerando apenas a evidência amostral teríamos decidido a favor de H_0.
- Se $P(H_0) = 8/9 = 0,8889$,
  - a *odds a priori* é $P(H_0)/P(H_1) = 8$. A evidência *a priori* é fortemente a favor H_0 e revela que a chance *a priori* de H_0 é 8 vezes a chance *a priori* de H_1.
  - A posteriori, teremos $P(H_0 \mid \mathrm{x}) = 0,9375$. Ou seja, a *odds a posteriori* é igual a 15 revelando que a evidência *a posteriori* a favor de H_0 é ainda mais forte do que se tinha *a priori*
  - note que, neste caso, as evidências amostral e *a priori* ambas são a favor de H_0.

Cálculo do Fator de Bayes para hipóteses compostas

- O Fator de Bayes não é uma razão de verossimilhança propriamente dita.
- No caso em que Θ_i não é um conjunto unitário, ela é um tipo de verossimilhança marginal.
- Suponha que Θ_i não seja um conjunto unitário. Neste caso, o cálculo do Fator de Bayes é feito da seguinte forma

$$P(\mathrm{x} \mid H_i) = \int_{\Theta_i} f(\boldsymbol{x} \mid \theta)\pi(\theta)d\theta$$

Exemplo: Suponha que $\mathrm{X}|\mu \overset{iid}{\sim} Normal(\mu, \sigma^2)$, σ^2 conhecido, e que a incerteza *a priori* sobre μ é bem descrita pela distribuição de Jeffreys. Determine o fator de Bayes se o interesse é testar as hipóteses $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$, μ_0 conhecido.

Cálculo do Fator de Bayes para hipóteses compostas

Solução:

$$
\begin{aligned}
P(\boldsymbol{x} \mid H_0) &= \int_{-\infty}^{\mu_0} (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}\right\} d\mu \\
&= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{\sum_{i=1}^n (x_i-\bar{x})^2}{2\sigma^2}\right\} \left(-\frac{2\pi\sigma^2}{n}\right)^{1/2} \\
&\quad \int_{-\infty}^{\mu_0} \left(\frac{n}{2\pi\sigma^2}\right)^{1/2} \exp\left\{-\frac{\sum_{i=1}^n (\mu-\bar{x})^2}{2\sigma^2}\right\} d\mu \\
&= (2\pi\sigma^2)^{-(n-1)/2} \sqrt{n} \exp\left\{-\frac{\sum_{i=1}^n (x_i-\bar{x})^2}{2\sigma^2}\right\} \Phi\left(\frac{(\mu_0-\bar{x})\sqrt{n}}{\sigma}\right)
\end{aligned}
$$

Analogamente, temos que

$$
\begin{aligned}
P(\boldsymbol{x} \mid H_1) &= \int_{\mu_0}^{\infty} (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}\right\} d\mu \\
&= (2\pi\sigma^2)^{-(n-1)/2} \sqrt{n} \exp\left\{-\frac{\sum_{i=1}^n (x_i-\bar{x})^2}{2\sigma^2}\right\} \left[1-\Phi\left(\frac{(\mu_0-\bar{x})\sqrt{n}}{\sigma}\right)\right]
\end{aligned}
$$

Cálculo do Fator de Bayes para hipóteses compostas

- Daí, o fator de Bayes é

$$BF(H_0, H_1) = \frac{\Phi\left(\frac{(\mu_0 - \bar{x})\sqrt{n}}{\sigma}\right)}{1 - \Phi\left(\frac{(\mu_0 - \bar{x})\sqrt{n}}{\sigma}\right)}$$

- Note que, neste caso, o fator de Bayes é igual a *odds a posteriori* pois vimos que

$$P(H_0 \mid \boldsymbol{x}) = P(\mu \leq \mu_0 \mid \boldsymbol{x}) = \Phi\left(\frac{n^{1/2}(\mu_0 - \bar{x})}{\sigma}\right)$$

$$P(H_1 \mid \boldsymbol{x}) = P(\mu > \mu_0 \mid \boldsymbol{x}) = 1 - \Phi\left(\frac{n^{1/2}(\mu_0 - \bar{x})}{\sigma}\right)$$

- Ou seja, é como se tivessemos assumindo que *a priori* $P(H_0) = P(H_1)$. Mas, *a priori* usamos uma distribuição imprópria. Então não tínhamos como calcular a *odds a priori*.

Cálculo do Fator de Bayes para hipóteses compostas

Exemplo (cont): No exemplo anterior, suponha que $n = 4$, $\bar{x} = 106$, $\sigma^2 = 400$ e $\mu_0 = 100$.

- O Fator de Bayes é
$$BF(H_0, H_1) = \frac{\Phi\left(\frac{(n)^{1/2}(\mu_0 - \bar{x})}{\sigma}\right)}{1 - \Phi\left(\frac{(n)^{1/2}(\mu_0 - \bar{x})}{\sigma}\right)} = \frac{\Phi(-0.6)}{1 - \Phi(-0.6)} = 0,3774$$
- A hipótese nula é $(0,3774)^{-1} = 2.65$ vezes menos provável que a hipótese alternativa e devemos rejeitar H_0.
- Como $BF(H_1, H_0) = [BF(H_0, H_1)]^{-1} = 2.65 \in (1, 3)$, temos Evidência Fraca a favor de H_1.

Teste de Hipóteses: A função de perda

Proposição: O teste cuja regra de decisão é "aceita-se H_0 se $P(H_0 \mid \mathrm{x}) > P(H_1 \mid \mathrm{x})$´´, é uma regra de Bayes quando a seguinte função perda é considerada

$$\begin{cases} L(\text{Aceitar } H_0, \theta) = \omega_1 1\{\theta \in \Theta_1\}, \\ L(\text{Rejeitar } H_0, \theta) = \omega_0 1\{\theta \in \Theta_0\}, \end{cases}$$

Prova: As perdas esperadas são

$$\begin{aligned} E(L(\text{Aceitar } H_0, \theta) \mid \mathrm{x}) &= \int \omega_1 1\{\theta \in \Theta_1\} \pi(\theta \mid \boldsymbol{x}) d\theta \\ &= \omega_1 \int_{\Theta_1} \pi(\theta \mid \boldsymbol{x}) d\theta = \omega_1 P(\theta \in \Theta_1 \mid \boldsymbol{x}) \\ E(L(\text{Rejeitar } H_0, \theta) \mid \mathrm{x}) &= \int \omega_0 1\{\theta \in \Theta_0\} \pi(\theta \mid \boldsymbol{x}) d\theta \\ &= \omega_0 \int_{\Theta_0} \pi(\theta \mid \boldsymbol{x}) d\theta = \omega_0 P(\theta \in \Theta_0 \mid \boldsymbol{x}) \end{aligned}$$

Teste de Hipóteses: A função de perda

Aceita-se H_0, se

$$
\begin{aligned}
E(L(\text{Aceitar } H_0, \theta) \mid \mathrm{x}) &< E(L(\text{rejeitar } H_0, \theta) \mid \mathrm{x}) \\
\Rightarrow \omega_1 P(\theta \in \Theta_1 \mid \boldsymbol{x}) &< \omega_0 P(\theta \in \Theta_0 \mid \boldsymbol{x}) \\
\Rightarrow \omega_1 [1 - P(\theta \in \Theta_0 \mid \boldsymbol{x})] &< \omega_0 P(\theta \in \Theta_0 \mid \boldsymbol{x}) \\
\Rightarrow \omega_1 &< [\omega_0 + \omega_1] P(\theta \in \Theta_0 \mid \boldsymbol{x}) \\
\Rightarrow P(H_0 \mid \mathrm{x}) &> \frac{\omega_1}{\omega_1 + \omega_0} \qquad (4)
\end{aligned}
$$

Efeito da escolha da Função de perda: Se $\omega_1 = k\omega_0$, k constante, aceitamos H_0 se

$$P(H_0 \mid \mathrm{x}) > k(k+1)^{-1}.$$

- Se $k = 1$ somos penalizados da mesma forma por aceitar H_0 quando H_0 é falsa e por rejeitar H_0 quando H_0 é verdadeira.
  - aceita-se H_0 se $P(H_0 \mid \mathrm{x}) > 0,500$.

Teste de Hipóteses: A função de perda

- Se $k = 2$ pagamos um preço duas vezes maior por aceitar H_0 quando H_0 é falsa do que por rejeitar H_0 quando H_0 é verdadeira.
  - aceita-se H_0 se $P(H_0 \mid \mathrm{x}) > 0,667$.
- Se $k = 100$ pagamos um preço cem vezes maior por aceitar H_0 quando H_0 é falsa do que por rejeitar H_0 quando H_0 é verdadeira.
  - aceita-se H_0 se $P(H_0 \mid \mathrm{x}) > 0,990$.
- Se $k = 1000$, aceita-se H_0 se $P(H_0 \mid \mathrm{x}) > 0,999$.

Quanto maior o valor de k, maior é o preço ω_1 que aceitamos pagar por **Aceitarmos** H_0 quando H_0 é falsa. Neste caso, exige-se uma maior é a evidência a favor de H_0 para podermos aceitá-la.

Teste de Hipóteses: Teste de Jeffreys para hipóteses precisas

- O teste que vimos é bastante simples e intuitivo se Θ_i, $i = 0, 1$, não é um conjunto unitário.
- No caso de testes para hipótese composta as decisões são tomadas considerando-se a evidência *a posteriori* sobre H_0 dada por $P(H_0 \mid \boldsymbol{x})$ a qual é calculada diretamente sem necessitarmos estabelecer a probabilidade de H_0 *a priori*.
- Se $\Theta_0 = \{\theta_0\}$ é unitário e θ é contínuo, então teremos que

$$P(H_0 \mid \mathrm{x}) = P(\theta = \theta_0 \mid \mathrm{x}) = 0.$$

- Isto gera um problema pois qualquer que seja a informação disponível, sempre rejeitaremos H_0.
- Uma solução apresentada para este problema foi dada por Jeffreys (1961).

Teste de Hipóteses: Teste de Jeffreys para hipóteses precisas

Se desejamos testar as hipóteses

$$H_0 : \theta = \theta_0 \quad \times \quad H_1 : \theta \neq \theta_0$$

onde θ_0 é um valor conhecido

- A estratégia proposta por Jeffreys consiste em
  - $*$ atribuir uma probabilidad *a priori* $P(H_0) = p \in (0, 1)$ para que H_0 seja verdadeira.
  - $*$ A distribuição *a priori* para θ torna-se então uma mistura discreta de medidas de probabilidade da seguinte forma

$$\pi^*(\theta) = p1\{\theta = \theta_0\} + (1 - p)\pi(\theta)1\{\theta \neq \theta_0\} \qquad (5)$$

  - $*$ $\pi(\theta)$ é a distribuição *a priori* para θ restrita a Θ_1

Teste de Hipóteses: Teste de Jeffreys para hipóteses precisas

Como vimos na expressão (3), a probabilidade *a posteriori* de H_0 é

$$P(H_0 \mid \mathrm{x}) = \left[1 + \left[BF(H_0, H_1)\frac{p}{1-p}\right]^{-1}\right]^{-1} \tag{6}$$

Neste contexto, o fator de Bayes é dado por

$$\begin{aligned} P(\mathrm{x} \mid H_0) &= f(\mathrm{x} \mid \theta_0) \\ P(\mathrm{x} \mid H_1) &= \int_{\Theta - \theta_0} f(\mathrm{x} \mid \theta, H_1)\pi(\theta \mid H_1)d\theta \\ &= \int_{\Theta} f(\mathrm{x} \mid \theta)\pi(\theta)d\theta = f_{H_1}(\mathrm{x}) \text{ (caso continuo)} \end{aligned} \tag{7}$$

onde $f_{H_1}(\mathrm{x})$ denota a distribuição preditiva *a priori* restrita a H_1.

Teste de Hipóteses: Teste de Jeffreys para hipóteses precisas

Assim, temos que

- o Fator de Bayes é

$$BF(H_0, H_1) = \frac{f(\mathrm{x} \mid \theta_0)}{f_{H_1}(\mathrm{x})}$$

- A probabilidade *a posteriori* para H_0 é

$$P(H_0 \mid \mathrm{x}) = \left[1 + \frac{1-p}{p} BF(H_0, H_1)^{-1}\right]^{-1}$$

- A distribuição preditiva *a priori* relativa a distribuição *a priori* para θ dada em (5) é

$$\begin{aligned} f(\boldsymbol{x}) &= \int_{\Theta} f(\boldsymbol{x} \mid \theta)\{p1\{\theta = \theta_0\} + (1-p)\pi(\theta)1\{\theta \neq \theta_0\}\}d\theta \\ &= pf(\boldsymbol{x} \mid \theta_0) + (1-p)f_{H_1}(\boldsymbol{x}) \end{aligned}$$

Exemplo: Considere o teste $H_0 : \theta = \theta_0$ *versus* $H_1 : \theta \neq \theta_0$.
Suponha que $X|\theta \overset{iid}{\sim} Normal(\theta, \sigma^2)$, σ^2 conhecido e que $\theta \mid Normal(m, b^2)$ seja a distribuição *a priori* para μ restrita a Θ_1.
Teste as hipóteses.

Solução: Comecemos calculando o fator de Bayes.

- Sob H_0 temos

$$f(\boldsymbol{x} \mid H_0) = f(\boldsymbol{x} \mid \theta_0) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{\frac{-\sum_{i=1}^{n}(x_i - \theta_0)^2}{2\sigma^2}\right\}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{\frac{-\sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \theta_0)^2}{2\sigma^2}\right\}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{\frac{-\sum_{i=1}^{n}(x_i - \bar{x})^2}{2\sigma^2}\right\} \exp\left\{\frac{-n(\bar{x} - \theta_0)^2}{2\sigma^2}\right\}$$

- Sob H_1 temos

$$
\begin{aligned}
f(\boldsymbol{x} \mid H_1) &= \int_{\Theta-\theta_0} \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{\frac{-\sum_{i=1}^n (x_i-\bar{x})^2}{2\sigma^2}\right\} \\
&\quad \times \exp\left\{\frac{-n(\bar{x}-\theta)^2}{2\sigma^2}\right\} \left(\frac{1}{2\pi b^2}\right)^{1/2} \exp\left\{\frac{-(\theta-m)^2}{2b^2}\right\} d\theta \\
&= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \left(\frac{1}{2\pi b^2}\right)^{1/2} \exp\left\{\frac{-\sum_{i=1}^n (x_i-\bar{x})^2}{2\sigma^2}\right\} \\
&\quad \times \int_{\Theta} \exp\left\{\frac{-n(\bar{x}-\theta_0)^2}{2\sigma^2}\right\} \exp\left\{\frac{-(\theta-m)^2}{2b^2}\right\} d\theta \\
&= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{\frac{-\sum_{i=1}^n (x_i-\bar{x})^2}{2\sigma^2}\right\} \left(\frac{\sigma^2/n}{b^2+\sigma^2/n}\right)^{1/2} \\
&\quad \times \exp\left\{\frac{-n(\bar{x}-m)^2}{2(\sigma^2+nb^2)}\right\}
\end{aligned}
$$

Teste de Hipóteses: Teste de Jeffreys para hipóteses precisas

- O fator de Bayes assume a seguinte forma:

$$BF(H_0, H_1) = \left[\frac{nb^2 + \sigma^2}{\sigma^2}\right]^{1/2} \exp\left\{\frac{n}{2}\left[\frac{(\bar{x} - m)^2}{nb^2 + \sigma^2} - \frac{(\bar{x} - \theta_0)^2}{\sigma^2}\right]\right\}$$

- Para o caso particular em que:
  + $n = 4$, $\bar{x} = 106$, $\sigma^2 = b^2 = 400$ y $\theta_0 = m = 100$
  + *A priori*, as duas hipóteses são equiprováveis, isto é, $p = 0.5$ (não-informativa para as hipóteses mas põe peso forte próximo de $\theta = 100$)

Teste de Hipóteses: Teste de Jeffreys para hipóteses precisas

O fator de Bayes torna-se

$$BF(H_0, H_1) = (n+1)^{1/2} \exp\left\{ -\frac{n^2(\bar{x}-\theta_0)^2}{2(n+1)\sigma^2} \right\} = 2.5912$$

- Escala de Jeffreys: evidência fraca a favor de H_0.

Neste caso

$$\begin{aligned} P(\theta = 100 \mid \mathrm{x}) &= [1 + 1/2.5912]^{-1} \\ &= 0.719 > 0.5 = P(\theta = 100) \end{aligned}$$

Teste de Hipóteses: Debilidades do Teste de Jeffreys

Este Teste para Hipóteses precisas pode levar ao paradoxo de Jeffreys-Lindley.

- Se $b^2 \to \infty$, significa que escolhemos uma distribuição *a priori* pouco informativa para θ.
- Neste caso, $\lim_{b^2 \to \infty} BF(H_0, H_1)$ é

$$\lim_{b^2 \to \infty} \left[\frac{nb^2 + \sigma^2}{\sigma^2} \right]^{1/2} \exp\left\{ \frac{n}{2} \left[(\bar{x} - m)^2 nb^2 + \sigma^2 - \frac{(\bar{x} - \theta_0)^2}{\sigma^2} \right] \right\} = \infty.$$

- Consequentemente, $P(H_0 \mid \mathrm{x}) \to 1$ independentemente da informação da amostra.
- Ou seja, aceitamos H_0 sempre, independentemente, da evidência amostral.
- É um risco utilizar distribuição não-informativa para teste de hipóteses precisas com este procedimento.

Teste de Hipóteses: Teste de Jeffreys na presença de parâmetros de perturbação

Considere o teste $H_0 : \theta = \theta_0$ *versus* $H_1 : \theta \neq \theta_0$ e assuma que a distribuição de X depende de θ e γ, ambos desconhecidos.

- Neste caso, γ é um parâmetro de perturbação.
- Para testarmos hipóteses sobre θ devemos "eliminar"γ do cálculo da probabilidade *a posteriori* de H_0.
- Para isto, o Fator de Bayes deve ser calculado como segue

$$BF(H_0, H_1) = \frac{\int_\Gamma f(\mathrm{x} \mid \theta_0, \gamma)\pi(\gamma)d\gamma}{\int_\Theta \int_\Gamma f(\mathrm{x} \mid \theta, \gamma)\pi(\theta, \gamma)d\gamma d\theta}$$

- O cálculo da probabilidade *a posteriori* de H_0 é feito como mostrado anteriormente na expressão (3).

Teste de Hipóteses: Teste de Jeffreys na presença de parâmetros de perturbação

Exemplo: Considere o teste $H_0 : \theta = \theta_0$ *versus* $H_1 : \theta \neq \theta_0$. Se $X|\theta, \sigma^2 \overset{iid}{\sim} Normal(\theta, \sigma^2)$, θ, σ^2 desconhecidos e $\pi(\theta, \sigma^2) \propto 1/\sigma^2$, determine o fator de Bayes.

Solução: Para calcularmos o fator de BAyes necessitamos:

$$
\begin{aligned}
P(\boldsymbol{x} \mid H_0) &= \int_0^{\infty} f(\boldsymbol{x} \mid \theta_0, \sigma) 1/\sigma^2 d\sigma^2 \\
&= \int_0^{\infty} \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta_0)^2\right\} 1/\sigma^2 d\sigma^2 \\
&= \left(\frac{1}{2\pi}\right)^{n/2} \int_0^{\infty} \left(\frac{1}{\sigma^2}\right)^{(n+2)/2} \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta_0)^2\right\} d\sigma^2 \\
&= \left(\frac{1}{2\pi}\right)^{n/2} \Gamma(n/2) \left[\sum_{i=1}^{n}(x_i - \theta_0)^2/2\right]^{-(n)/2} \qquad (8)
\end{aligned}
$$

Teste de Hipóteses: Teste de Jeffreys na presença de parâmetros de perturbação

$$
\begin{aligned}
P(\boldsymbol{x} \mid H_1) &= \left(\frac{1}{2\pi}\right)^{n/2} \int_0^{\infty} \int_{-\infty}^{\infty} \left(\frac{1}{\sigma^2}\right)^{(n+2)/2} \exp\left\{\frac{-1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \theta)^2\right\} d\theta d\sigma^2 \\
&= \left(\frac{1}{2\pi}\right)^{n/2} \int_0^{\infty} \left(\frac{1}{\sigma^2}\right)^{(n+2)/2} \exp\left\{\frac{-1}{2\sigma^2} \sum_{i=1}^{n} (x_i)^2\right\} \\
&\qquad \int_{-\infty}^{\infty} \exp\left\{\frac{-n}{2\sigma^2}(\theta^2 - 2\theta\bar{x})\right\} d\theta d\sigma^2 \\
&= \left(\frac{1}{2\pi}\right)^{(n-1)/2} (1/n)^{1/2} \int_0^{\infty} \left(\frac{1}{\sigma^2}\right)^{(n+1)/2} \exp\left\{\frac{-1}{2\sigma^2}\left[\sum_{i=1}^{n} (x_i)^2 - n\bar{x}^2\right]\right\} \\
&\qquad \int_{-\infty}^{\infty} \left(\frac{1}{2\pi\sigma^2/n}\right)^{1/2} \exp\left\{\frac{-n}{2\sigma^2}(\theta - \bar{x})^2\right\} d\theta d\sigma^2 \\
&= \left(\frac{1}{2\pi}\right)^{(n-1)/2} (1/n)^{1/2} \Gamma\left((n-1)/2\right) \left[\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)/2\right]^{-(n-1)/2}
\end{aligned}
$$

Teste de Hipóteses: Teste de Jeffreys na presença de parâmetros de perturbação

Daí, o fator de Bayes é

$$
\begin{aligned}
BF(H_0, H_1) &= \frac{P(\boldsymbol{x} \mid H_0)}{P(\boldsymbol{x} \mid H_1)} \\
&= \frac{(n/2\pi)^{0.5}\Gamma(n/2)[\sum x_i^2 - n\bar{x}^2]^{(n-1)/2}}{\Gamma((n-1)/2)[\sum(x_i - \theta_0)]^{n/2}}
\end{aligned}
$$

Aproximações Computacionais para o fator de Bayes

Para determinar o fator de Bayes necessitamos calcular a distribuição preditiva *a priori* que é

$$
\begin{aligned}
f(x) &= \int_\Theta f(x \mid \theta)\pi(\theta)d\theta \\
&= E_\theta[f(\boldsymbol{x} \mid \theta)]
\end{aligned}
$$

Método 1 (Monte Carlo):

$P1$: Obtenha uma amostra $\theta_1, \ldots, \theta_T$ da distribuição *a priori* $\pi(\theta)$.

$P2$: $\hat{f}(x) = \frac{1}{T}\sum_{i=1}^{T} f(x \mid \theta_i)$.

- Este método é simples mas não fornece uma boa aproximação.
- Sua aplicação não é possivel se utilizarmos distribuições *a priori* impróprias.

A distribuição preditiva *a priori* pode ser escrita como

$$\begin{aligned}
[f(\mathrm{x})]^{-1} &= \int_{\Theta} [f(\mathrm{x})]^{-1} \pi(\theta) d\theta \\
&= \int_{\Theta} [f(\mathrm{x} \mid \theta)]^{-1} f(\mathrm{x} \mid \theta) \pi(\theta) [f(\mathrm{x})]^{-1} d\theta \\
&= \int_{\Theta} [f(\mathrm{x} \mid \theta)]^{-1} \pi(\theta \mid \mathrm{x}) d\theta = E_{\theta \mid \boldsymbol{x}} [f(\mathrm{x} \mid \theta)]^{-1}]
\end{aligned}$$

Método 2 (Newton-Raftery(1994):

$P1$: Obtenha una amostra $\theta_1, \ldots, \theta_T$ de $\pi(\theta \mid \mathrm{x})$.

$P2$: $\hat{f}(\mathrm{x}) = \left[\frac{1}{T} \sum_{i=1}^{T} \frac{1}{f(\mathrm{x} \mid \theta_i)} \right]^{-1}$.

Este método pode ser instável por causa dos valores pequenos de $f(\mathrm{x} \mid \theta_i)$.

Aproximações Computacionais para o fator de Bayes

Seja $g(\theta)$ uma distribuição própria que seja uma boa aproximação de $\pi(\theta \mid \mathrm{x})$.

$$
\begin{aligned}
[f(\mathrm{x})]^{-1} &= \int_{\Theta} [f(\mathrm{x})]^{-1} g(\theta) d\theta \\
&= \int_{\Theta} \frac{g(\theta)}{f(\mathrm{x} \mid \theta)\pi(\theta)} \pi(\theta \mid \mathrm{x}) d\theta \\
&= E_{\theta|\boldsymbol{x}} \left[\frac{g(\theta)}{f(\mathrm{x} \mid \theta)\pi(\theta)} \right]
\end{aligned}
$$

Método 3 (Gelfand - Dey(1994)):

$P1$: Obtenha uma amostra $\theta_1, \ldots, \theta_T$ de $\pi(\theta \mid \mathrm{x})$.

$P2$: $\hat{f}(\mathrm{x}) = \left[\frac{1}{T} \sum_{i=1}^{T} \frac{g(\theta_i)}{f(\mathrm{x}|\theta_i)\pi(\theta_i)} \right]^{-1}$.

Fornece aproximações mais estáveis.

Teste de Significância Bayesiano Completo (FBST)

- Foi proposto formalmente por Pereira e Stern (1999, Entropy).
- Apresenta uma maneira mais interessante para testarmos hipóteses precisas, sem artifícios.
- A evidência sobre H_0 é obtida diretamente da distribuição *a posteriori* do parâmetro de interesse.
- Tem uma conexão com os intervalos HPD.

Teste de Significância Bayesiano Completo (FBST)

Considere as seguintes hipóteses

$$H_0 : \theta \in \Theta_0 \;\times\; H_1 : \theta \in \Theta_1$$

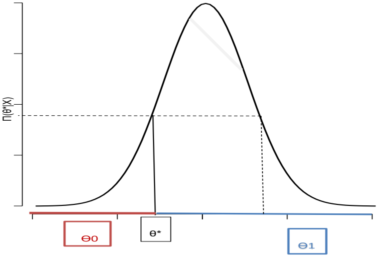
onde $\{\Theta_0, \Theta_1\}$ é uma partição do espaço paramétrico Θ.

- O FBST é um critério para decidir sobre H_0 considerando somente $\pi(\theta \mid \mathrm{x})$.
- Seja θ^* o valor em Θ_0 tal que

$$\pi(\theta^* \mid \mathrm{x}) = \max\{\pi(\theta \mid \mathrm{x}) : \theta \in \Theta_0\}$$

- Seja $T(\mathrm{x}) = \{\theta \in \Theta : \pi(\theta \mid \mathrm{x}) > \pi(\theta^* \mid \mathrm{x})\}$.

Teste de Significância Bayesiano Completo (FBST)



- A credibilidade de $T(\mathrm{x})$ é definida como a probabilidade *a posteriori* de $T(\mathrm{x})$, ou seja,

$$\kappa^* = \int_{T(\mathrm{x})} \pi(\theta \mid \mathrm{x}) d\theta = P(\theta \in T(\mathrm{x}) \mid \mathrm{x})$$

- A evidência *a posteriori* a favor de H_0 é

$$Ev(H_0, \mathrm{x}) = 1 - \kappa^*.$$

Teste de Significância Bayesiano Completo (FBST)

Se κ^* é grande e, consequentemente, $Ev(H_0, \text{x})$ é pequeno, temos que

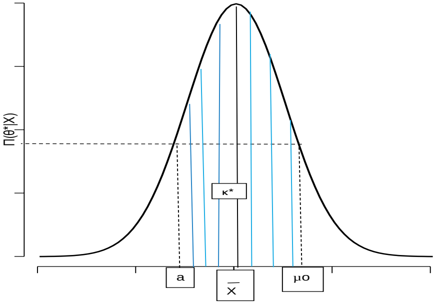
- os valores θ que pertencente à Θ_0 estão em uma região de baixa densidade *a posteriori*
- Evidência contra H_0.

A evidência *a posteriori* a valor de H_0, $Ev(H_0, \text{x})$, é um tipo de valor-P bayesiano.

Teste de Significância Bayesiano Completo (FBST)

Exemplo: Suponha que $X|\mu \stackrel{iid}{\sim} Normal(\mu, \sigma^2)$, σ^2 conhecido, e que, *a priori* $\pi(\mu) \propto 1$. Teste as hipóteses $H_0 : \mu = \mu_0$ *versus* $H_1 : \mu \neq \mu_0$, μ_0 conhecida.

- A distribuição *a posteriori* es $\mu \mid x \sim Normal(\bar{x}, \sigma^2/n)$.
- Se $\mu_0 > \bar{x}$ então $\theta^* = \mu_0$.



Teste de Significância Bayesiano Completo (FBST)

- A densidade *a posteriori* avaliada em $\theta^* = \mu_0$ é

$$\pi(\theta^* \mid \mathrm{x}) = \pi(\mu_0 \mid \mathrm{x}) = \left(\frac{n}{2\pi\sigma^2}\right)^{1/2} \exp\left\{\frac{n}{2\sigma^2}(\mu_0 - \bar{x})^2\right\}.$$

- $T(\mathrm{x}) = \{\theta \in \Theta : a < \theta < \mu_0\}$, onde $|\bar{x} - a| = |\bar{x} - \mu_0|$.
- A credibilidade é

$$\begin{aligned} \kappa^* &= P(a < \theta < \mu_0 \mid \boldsymbol{x}) \\ &= \Phi\left(\frac{\sqrt{n}(\mu_0 - \bar{x})}{\sigma}\right) - \Phi\left(\frac{\sqrt{n}(a - \bar{x})}{\sigma}\right) \end{aligned}$$

- A evidência *a posterior*

$$Ev(H_0, \boldsymbol{x}) = 1 - \Phi\left(\frac{\sqrt{n}(\mu_0 - \bar{x})}{\sigma}\right) + \Phi\left(\frac{\sqrt{n}(a - \bar{x})}{\sigma}\right).$$

Teste de Significância Bayesiano Completo (FBST)

Assuma que $H_0 : \mu = 100$ *versus* $H_1 : \mu \neq 100$ e que para $n = 4$ observamos $\bar{x} = 106$. Admita que $\sigma^2 = 400$. Neste caso,

- A distribuição *a posteriori* é $\mu \mid \boldsymbol{x} \sim N(106, 400/4)$.
- a credibilidade é $\kappa^* = \Phi\left(\frac{112-106}{10}\right) - \Phi\left(\frac{100-106}{10}\right) = \Phi(0,6) - \Phi(-0,6) = 0,4514$
- Evidência sobre H_0 é $Ev(H_0, \mathrm{x}) = 0.5486$.
- Evidência a favor de H_0.

Exemplo: Seja X uma única observação da fdp

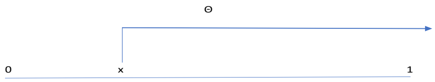
$$f(x \mid \theta) = \frac{2x}{\theta^2} 1\{0 < x < \theta\}, \theta > 0.$$

Assuma *a priori* que $\theta \sim U(0,1)$. (a) Teste se $\theta \leq \theta_0$, $\theta_0 \in (x, 1)$.

Solução: Precisamos calcular a distribuição *a posteriori* de θ.

$$\pi(\theta \mid \boldsymbol{x}) = \frac{\frac{2x}{\theta^2} 1\{0 < x < \theta\} 1\{0 < \theta < 1\}}{\int \frac{2x}{\theta^2} 1\{0 < x < \theta\} 1\{0 < \theta < 1\} d\theta} \tag{9}$$

O Domínio



Teste de Significância Bayesiano Completo (FBST)

A distribuição *a posteriori* de θ é

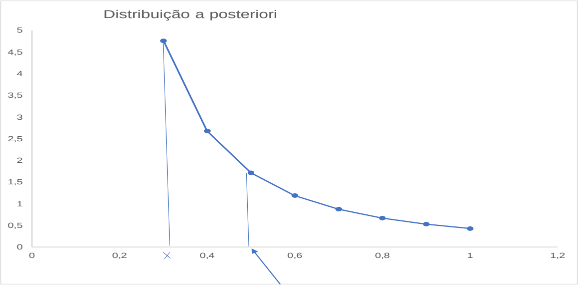
$$
\begin{aligned}
\pi(\theta \mid \boldsymbol{x}) &= \frac{\frac{2x}{\theta^2}1\{0 < x < \theta\}1\{0 < \theta < 1\}}{\int \frac{2x}{\theta^2}1\{0 < x < \theta\}1\{0 < \theta < 1\}d\theta} \\
&= \frac{\frac{1}{\theta^2}1\{x < \theta < 1\}}{\int \frac{1}{\theta^2}1\{x < \theta < 1\}d\theta} = \frac{\frac{1}{\theta^2}1\{x < \theta < 1\}}{\int_x^1 \frac{1}{\theta^2}d\theta} \\
&= \frac{x}{(1-x)\theta^2}1\{x < \theta < 1\} \qquad (10)
\end{aligned}
$$

Teste de Significância Bayesiano Completo (FBST)

Estudando o comportamento de $\pi(\theta \mid \boldsymbol{x})$.

$$\frac{d}{d\theta}\pi(\theta \mid \boldsymbol{x}) = -2\theta^{-3}\frac{x}{1-x} < 0, \forall\theta \in (x, 1)$$

- $\pi(\theta \mid \boldsymbol{x})$ é decrescente no intervalo $(x, 1)$.
- No teste a ser feito $\Theta_0 = (x, \theta_0)$



θ_o

Teste de Significância Bayesiano Completo (FBST)

Agora, no intervalo $\Theta_0 = (x, \theta_0)$ a distribuição *a posteriori* $\pi(\theta \mid \boldsymbol{x})$ é decrescente, o máximo de $\pi(\theta \mid \boldsymbol{x})$ será em $\theta = x$ e

$$\pi(\theta^* \mid x) = \pi(x \mid x) = [x(1-x)]^{-1}$$

Daí, temos que $T(x) = \phi \Rightarrow$ a credibilidade é $\kappa^* = 0$

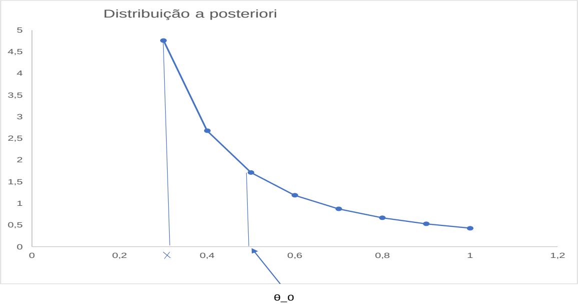
Consequentemente, A evidência de H_0 é

$$Ev(H_0, x) = 1 \quad \Rightarrow \quad \text{Aceitamos } H_0.$$

Teste de Significância Bayesiano Completo (FBST)

(b) Teste se $\theta \geq \theta_0$, $\theta_0 \in (x, 1)$.

Solução: Note que, neste caso, restrito a $\Theta_0 = (\theta_0, 1)$, a distribuição *a posteriori* é máximizada no ponto $\theta = \theta_0$.



Teste de Significância Bayesiano Completo (FBST)

Como $\pi(\theta \mid \boldsymbol{x})$ é decrescente, para todos os pontos $\theta \in \Theta_1 = (x, \theta_0)$ teremos

$$\pi(\theta \mid \boldsymbol{x}) < \pi(\theta^* \mid \boldsymbol{x}) = \pi(\theta_0 \mid \boldsymbol{x}) = \frac{x}{(1-x)\theta_0^2}$$

$$T(X) = (x, \theta_0) = \Theta_1$$

Daí, a credibilidade é $\kappa^* = \int_x^{\theta_0} \frac{x}{(1-x)\theta^2} d\theta = \frac{\theta_0 - x}{(1-x)\theta_0}$

A evidência de H_0 é $Ev(H_0, x) = 1 - \frac{\theta_0 - x}{(1-x)\theta_0}$

- Se $x = 1/3$ e $\theta_0 = 2/3$ temos que $Ev(H_0, x) = 1/4 \Rightarrow$ Aceitamos H_0.
- Se $x = 0,5$ e $\theta_0 = 0,99$ temos que $Ev(H_0, x) \approx 0,01 \Rightarrow$ Rejeitamos H_0.

FBST: função de perda

Proposição: o FBST é uma regra de Bayes se a seguinte função perda é considerada

$$\begin{cases} L(\text{Aceitar } H_0, \theta) = b + c1\{\theta \in T(\mathrm{x})\}, \\ L(\text{rejeitar } H_0, \theta) = a[1 - 1\{\theta \in T(\mathrm{x})\}], \end{cases}$$

onde a, b e c são números positivos.
Prova: As perdas esperadas são

$$\begin{aligned} E(L(\text{Aceitar } H_0, \theta) \mid \mathrm{x}) &= \int_\Theta b + c1\{\theta \in T(\mathrm{x})\}\pi(\theta \mid \boldsymbol{x})d\theta \\ &= b\int_\Theta P(\theta \mid \boldsymbol{x})d\theta + c\int_{T(x)} P(\theta \mid \boldsymbol{x})d\theta \\ &= b + cP(\theta \in T(x) \mid \boldsymbol{x}) = b + c\kappa^* \\ E(L(\text{Rejeitar } H_0, \theta) \mid \mathrm{x}) &= \int_\Theta a[1 - 1\{\theta \in T(\mathrm{x})\}]\pi(\theta \mid \boldsymbol{x})d\theta \\ &= a\int_\Theta P(\theta \mid \boldsymbol{x})d\theta - a\int_{T(x)} P(\theta \mid \boldsymbol{x})d\theta \\ &= a - aP(\theta \in T(x) \mid \boldsymbol{x}) = a - a\kappa^* \end{aligned}$$

FBST: função de perda

Aceita-se a hipóteses nula se

$$
\begin{aligned}
E(L(\text{Aceitar } H_0, \theta) \mid \mathrm{x}) &< E(L(\text{rejeitar } H_0, \theta) \mid \mathrm{x}) \\
b + c\kappa^* &< a - a\kappa^* \\
\Rightarrow \kappa^* &< \frac{a-b}{c+a}
\end{aligned}
$$

Como $\kappa^* = 1 - Ev(H_0, \mathrm{x})$, aceita-se a hipóteses nula se

$$Ev(H_0, \mathrm{x}) > \frac{b+c}{b+a}.$$

FBST:Como escolher a, b e c?

Para entender os efeitos de a, b e c na tomada de decisão, consideremos alguns casos extemos.

- Se $T^*(\mathrm{x}) = \Theta$ temos forte evidência contra H_0 (H_0 é falsa) pois $Ev(H_0, \mathrm{x}) = 0$ e a perda é

$$\begin{cases} L(\text{Aceitar } H_0, \theta) = b + c \\ L(\text{rejeitar } H_0, \theta) = 0. \end{cases}$$

- Se $T^*(\mathrm{x}) = \phi(vazio)$ temos forte evidência a favor H_0 (H_0 é verdadeira) pois $Ev(H_0, \mathrm{x}) = 1$ e a perda é

$$\begin{cases} L(\text{Aceitar } H_0, \theta) = b \\ L(\text{rejeitar } H_0, \theta) = a. \end{cases}$$

Consequentemente, uma boa escolha é considerar a e c grandes e $b \to 0$.

FBST e sua conexão com as regiões HPD

- As Regiões de mais alta densidade *a posteriori* (regiões HPD) são usualmente consideradas para se decidir sobre H_0.
- Se $R(\mathrm{x}) = \{\theta \in \Theta : \pi(\theta|\mathrm{x}) \geq c_\gamma\}$ é uma região HPD para θ com probabilidade $(1-\gamma)$, onde c_γ é a maior constante tal que $P(\theta \in R(\mathrm{x})|\mathrm{x}) \geq 1-\gamma$.
  - aceitamos H_0 se θ_0 pertence a $R(\mathrm{x})$.
- Este critério de decisão não tem qualquer equivalência com o critério fornecido pelo teste de Jeffreys.
- No entanto, existe uma equivalência entre este critério de decisão e decisões tomadas utilizando o FBST.

FBST e sua conexão com as regiões HPD

A região HPD com probabilidade $1-\gamma$ e o FBST produzem a mesma decisão se

- $Ev(H_0, \text{x}) > \gamma \Rightarrow$ aceitamos H_0.
  - Se $H_0 : \theta = \theta_0$ e a região HPD é um interlavo (l, L), se $\theta_0 \in (l, L)$, aceitamos H_0. Neste caso, temos que $\kappa^* < 1-\gamma$ $\Rightarrow Ev(H_0, \text{x}) > \gamma$ que leva à aceitação de H_0 pelo FBST.

- $Ev(H_0, \text{x}) < \gamma \Rightarrow$ Rejeitamos H_0.
  - Se $H_0 : \theta = \theta_0$ e a região HPD é um interlavo (l, L), se $\theta_0 \notin (l, L)$, rejeitamos H_0. Neste caso, temos que $\kappa^* > 1-\gamma$ $\Rightarrow Ev(H_0, \text{x}) < \gamma$ que leva à rejeição de H_0 pelo FBST.

FBST: Aproximação computacional para a evidência

Suponha que o kernel da distribuição *a posteriori*

$$\pi(\theta \mid \mathrm{x}) \propto f(\mathrm{x} \mid \theta)\pi(\theta)$$

seja conhecido.

$P1$: Determine θ^* e $f(\mathrm{x} \mid \theta^*)\pi(\theta^*)$.

$P2$: Obtenha uma amostra $\theta_1, \ldots, \theta_T$ de $\pi(\theta \mid \mathrm{x})$.

$P3$: Calcule $f(\mathrm{x} \mid \theta_i)\pi(\theta_i)$, $i = 1, \ldots, n$.

$P3$: a evidência *a posteriori* é aproximada por:

$$\hat{E}(H_0, \mathrm{x}) = 1 - \frac{\sum_{i=1}^{T} 1\{f(\mathrm{x} \mid \theta_i)\pi(\theta_i) > f(\mathrm{x} \mid \theta^*)\pi(\theta^*)\}}{T}$$

O passo $P1$ é fácil se H_0 é uma hipótese precisa mais é necessário uma rotina para encontrar o máximo em Θ_0, caso contrário.