Inferência Estatística com Abordagem Bayesiana

Rosangela Helena Loschi [1]

[1] Departamento de Estatística
Universidade Federal de Minas Gerais

19 de janeiro de 2021

MÉTODOS DE CONSTRUÇÃO DA DISTRIBUIÇÃO *A PRIORI*

- Distribuições conjugadas naturais
- **Distribuições de referências: Métodos Bayes-laplace e Jeffreys**

Distribuições referência ou não-informativas

Muitas vezes as distribuições conjugadas

- podem não representar bem o conhecimento inicial do indivíduo sobre o parâmetro ou
- podem não existir.

Outra forma muito comumente utilizada para fazer uma análise Bayesiana é o uso de **distribuições** *a priori* **não-informativas ou de referência**. Algumas justificativas para usá-las

(1) desconhecimento sobre o problema

(2) ter informação mas não querer utilizá-la

(3) querer **ser objetivo**.

Existem muitas classes de distribuições *a priori* não-informativas pois não há um consenso sobre o que é não-informação.

Dist. não-informativas: Sentido Bayes-Laplace

- Estão baseadas no **Princípio da Razão Insuficiente:** "se não há informação para diferenciar entre os valores de $\theta \in \Theta$ devemos atribuir a todos estes valores a mesma probabilidade´´.
- Sob este princípio: **a distribuição *a priori* de qualquer parâmetro é uma distribuição Uniforme no espaço paramétrico Θ.**

Exemplo:

- Se $\Theta = \{\theta_1, \ldots, \theta_N\} \Rightarrow \pi(\theta_i) = P(\theta = \theta_i) = 1/N$
- Se $\Theta = [a, b] \subset \mathbb{R} \Rightarrow \pi(\theta) = (b - a)^{-1}$
- Se $\Theta = \{\theta_1, \theta_2, \ldots\} \Rightarrow \pi(\theta_i) = P(\theta = \theta_i) \propto K$, constante. (Note que $\sum_{i=1}^{\infty} \pi(\theta) = \sum_{i=1}^{\infty} K = \infty$(imprópria)
- Se $\Theta = \mathbb{R} \Rightarrow \pi(\theta) \propto K$, constante.

Dist. não-informativas: Sentido Bayes-Laplace

Criticas a este procedimento:

(1)sob este princípio não há invariância

- Se o espaço paramétrico é $\Theta = (0,1)$, usando este princípio, *a priori*, $\theta \sim \mathrm{Uniforme}(0,1)$.
- Se a meta é espicificar a distribuição *a priori* para a reparametrização $\phi = -log(\theta)$ do modelo
    - Se usássemos o Princípio da Razão Insuficiente, a distribuição *a priori* para ϕ que é não-informativa no sentido de Bayes-Laplace deveria ser uma uniforme no intervalo $(0, \infty)$.
    - No entanto, como $\theta \sim Uniforme(0,1)$, usando técnicas de prob., provamos que $\phi = g(\theta) = \log(\theta) \sim Exponencial(1) \leftarrow$ nao seguimos o Princípio da Razão Insuficiente na especificação da distribuição de ϕ.

$$\begin{aligned} g^{-1}(\phi) &= e^{-\phi} \\ \pi(\phi) &= \pi(g^{-1}(\phi))|\frac{dg^{-1}(\phi)}{d\theta}| = 1 * |-e^{-\phi}|, \end{aligned}$$

se $g^{-1}(\phi) \in (0,1) \Rightarrow e^{-\phi} \in (0,1) \Rightarrow \phi \in (0, \infty)$.

Dist. não-informativas: Sentido Bayes-Laplace

(2) podemos construir distribuições *a priori* impróprias $\leftarrow$ não são medidas de probabilidade.

- Estamos violando o **Príncipio da Coerência**
- Não é visto como um problema sério se a distribuição *a posteriori* é própria.
- Neste caso,
  - teríamos inferência *a posteriori* mas não teríamos inferência *a priori*.
  - usaremos uma regra geral de condicionamento para construírmos a distribuição *a posteriori*

Exemplo: SeJa $X_1, \ldots, X_n$, dado μ, são variáveis aleatórias i.i.d com distribuição Normal(μ, σ^2), em que a variância σ^2 é conhecida. Para o parâmetro μ, temos que

$\ast$ o espaço paramétrico é $\mathbb{R}$.

$\ast$ a distribuição *a priori* não-informativa no sentido de Bayes-Laplace é $\pi(\mu) \propto 1$, $\forall \mu \in \mathbb{R} \leftarrow$ é imprópria.

Dist. não-informativas: Sentido Bayes-Laplace

- a função de verossimilhança é

$$
\begin{aligned}
f(\boldsymbol{x} \mid \mu) &= \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} \exp\left\{ \frac{-1}{2\sigma^2} (x_i - \mu)^2 \right\} \\
&= (2\pi\sigma^2)^{-n/2} \exp\left\{ \frac{-1}{2\sigma^2} \left(\sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2 \right) \right\} \\
&= (2\pi\sigma^2)^{-n/2} \exp\left\{ \frac{-1}{2\sigma^2} \sum_{i=1}^{n} x_i^2 \right\} \\
&\times \exp\left\{ \frac{-1}{2\sigma^2} \left(-2\mu \sum_{i=1}^{n} x_i + n\mu^2 \right) \right\} \\
&\propto \exp\left\{ -\frac{n}{2\sigma^2} \left(\mu^2 - 2\mu\bar{x} \right) \right\}
\end{aligned}
$$

- Segue que o núcleo da distribuição *a posteriori* para μ é

$$
\pi(\mu \mid \mathrm{x}) \propto \exp\left\{ -\frac{n}{2\sigma^2} (\mu^2 - 2\mu\bar{x}) \right\}.
$$

Dist. não-informativas: Sentido Bayes-Laplace

- Assim, a distribuição *a posteriori* para μ é

$$\mu \mid \mathrm{x} \sim \mathrm{Normal}(\bar{x}, \sigma^2/n)$$

  a qual é uma distribuição própria.
- Neste caso, não há inferência *a priori* mas sempre há inferência *a posteriori*.

Atenção:

- Quando utilizamos uma distribuição *a priori* **própria**, a distribuição *a posteriori* **é sempre própria.**
- Se consideramos uma distribuição *a priori* imprópria, não temos garantia de que a distribuição *a posteriori* será própria. Neste caso,
  - **devemos provar que a distribuição** *a posteriori* **é própria antes iniciarnos nossa análise.**

Dist. não-informativas: problema com dist. impróprias

Exemplo: Suponha que $X_i \mid \theta \overset{iid}{\sim} Binomial(1, \theta)$ e que a distribuição *a priori* para θ seja

$$\pi(\theta) \propto \theta^{-1}(1-\theta)^{-1} \leftarrow Beta(0,0)(impropria)$$

Se observamos $\sum_{i=1}^{n} x_i = n$ temos que a distribuição *a posteriori* para θ é

$$\pi(\theta \mid \boldsymbol{x}) \propto \theta^{-1}(1-\theta)^{-1}\theta^{n}(1-\theta)^{0} \propto \theta^{n-1}(1-\theta)^{-1}$$

Isto implica que a distribuição *a posteriori* para θ é uma distribução $Beta(n, 0)$ que também é imprópria. Não há inferência *a posteriori*.

A condição para termos uma distribuição *a posteriori* própria é termos pelo menos um zero na amostra.

Dist. não-informativas *versus* distribuições conjugadas

No caso em que $X_1, \ldots, X_n$, dado μ, são variáveis aleatórias i.i.d com distribuição Normal(μ, σ^2), em que a variância σ^2 é conhecida, obtemos que

- A distribuição *a priori* conjugada da família normal com variância conhecida é $\mu \sim Normal(M, V)$ cujo núcleo é

$$\pi_C(\mu) \propto \exp\left\{-\frac{1}{2V}(\mu - M)^2\right\}$$

- Se assumirmos que $V \to \infty$ tinhamos uma distribuição vaga, que era dominada pelos dados quaquer que fosse n.
- Note que

$$\lim_{V \to \infty} \exp\left\{-\frac{1}{2V}(\mu - M)^2\right\} = 1$$

- Isto implica que se $V \to \infty$ o grau de informação trazido pela distribuição *a priori* conjugada $\pi_C(\mu)$ se aproxima daquele fornecido pela distribuições de Jeffreys (ver a seguir) e Bayes-Laplace.

Dist. não-informativas: Sentido Jeffreys

A distribuição *a priori* de Jeffreys é definida considerando-se a Informação de Fisher sobre θ.

A Informação de Fisher sobre θ denotada por $I(\theta)$

(1) mede a curvatura média da função de verossimilhança

(2) maior a curvatura da verossimilhança

- $\Rightarrow$ maior informação sumarizada pela verossimilhança (ou por $\boldsymbol{X}$) sobre θ
- $\Rightarrow$ maior é o valor da Informação de Fisher $I(\theta)$ no ponto θ.

Dist. não-informativas: Sentido Jeffreys

Definição(Informação de Fisher): Se $\boldsymbol{X}$ é um vetor/variável aleatória com função densidade (ou de probabilidade) $f(\boldsymbol{X} \mid \theta)$,

- se $\theta \in \Theta \subset \mathbb{R}$, a Informação de Fisher esperada é um número $I(\theta)$ dado por

$$I(\theta) = E_{\boldsymbol{X}|\theta}\left[-\frac{d^2}{d\theta^2}\ln f(\boldsymbol{X} \mid \theta)\right]$$

- se $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$, a Informação de Fisher esperada é uma matriz $I(\boldsymbol{\theta})$ dado por

$$I(\boldsymbol{\theta}) = E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[-\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\ln f(\boldsymbol{X} \mid \boldsymbol{\theta})\right]$$

  onde cada entrada ij da matriz $I(\boldsymbol{\theta})$ é dada por:

$$I_{ij}(\boldsymbol{\theta}) = E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln f(\boldsymbol{X} \mid \boldsymbol{\theta})\right]$$

Dist. não-informativas: Sentido Jeffreys

Definição: A distribuição *a priori* de Jeffreys é assim definida

- se $\theta \in \Theta \subset \mathbb{R}$,

$$\pi(\theta) \propto [I(\theta)]^{1/2}, \forall \theta \in \Theta.$$

- se $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$,

$$\pi(\boldsymbol{\theta}) \propto [\det I(\boldsymbol{\theta})]^{1/2}, \forall \boldsymbol{\theta} \in \Theta.$$

Dist. de Jeffreys: Modelo Bernoulli

Seja $X_1, \ldots, X_n$, dado θ, são i.i.d com distribuição Bernoulli(θ).

* Como $\Theta = (0, 1)$ a distribuição de Bayes-Laplace é a distribuição Uniforme$(0, 1)$ que pertence a família Beta.
* Cálculo da informação de Fisher para o modelo Bernoulli

$$\ln f(\boldsymbol{x} \mid \theta) = \sum_{i=1}^{n} x_i \ln \theta + (n - \sum_{i=1}^{n} x_i) \ln(1 - \theta)$$

$$\frac{d}{d\theta} \ln f(\boldsymbol{x} \mid \theta) = \frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{(n - \sum_{i=1}^{n} x_i)}{(1 - \theta)}$$

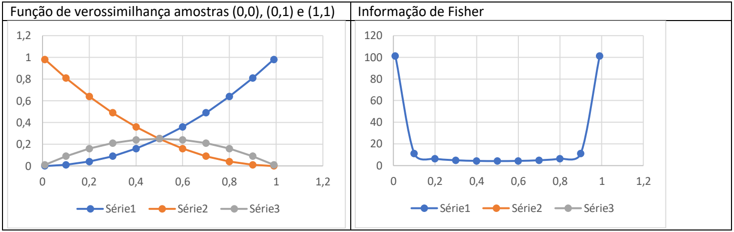
$$\frac{d^2}{d\theta^2} \ln f(\boldsymbol{x} \mid \theta) = -\frac{\sum_{i=1}^{n} x_i}{\theta^2} - \frac{(n - \sum_{i=1}^{n} x_i)}{(1 - \theta)^2}$$

Dist. de Jeffreys: Modelo Bernoulli

A informação de Fisher para o modelo Bernoulli é

$$I(\theta) \propto \theta^{-1}(1-\theta)^{-1}.$$

Informação de Fisher: caso bernoulli, n=2



Dist. de Jeffreys: Modelo Bernoulli

* Consequentemente, a distribuição de Jeffreys é

$$\pi(\theta) \propto \theta^{-1/2}(1-\theta)^{-1/2} \Rightarrow \theta \sim \mathrm{Beta}(1/2, 1/2).$$

* As distribuições de Jeffreys e Bayes-Laplace são próprias e ambas pertencem a família Beta.

Dist. não-informativas: Sentido Jeffreys

Observação Importante sobre a Informação de Fisher:

- Se $\boldsymbol{X} = (X_1, \ldots, X_n)$ são variáveis aleatórias independentes com $X_i \mid \theta \sim f_i(x_i \mid \theta)$, se $I_i(\theta)$ denota a Informação de Fisher calculada com base apenas na distribuição de X_i e $I(\theta)$ é a informação de Fisher calculada com base na amostra $\boldsymbol{X}$ então

$$I(\theta) = \sum_{i=1}^{n} I_i(\theta);$$

- Se $\boldsymbol{X} = (X_1, \ldots, X_n)$ são variáveis aleatórias independentes e identicamente distribuidas se $I(\theta)$ denota a Informação de Fisher calculada com base apenas na distribuição de cada X_i e $I_n(\theta)$ é a informação de Fisher calculada com base na amostra $\boldsymbol{X}$ então

$$I_n(\theta) = nI(\theta)$$

Dist. não-informativas: Sentido Jeffreys

Exemplo: Se $X_1, \ldots, X_n$, dado λ, são variáveis aleatórias i.i.d com distribuição exponencial com parametro λ cuja f.d.p. é

$$f(x_i \mid \lambda) = \lambda \exp\{-\lambda x_i\}, \quad x_i > 0. \tag{1}$$

(a)Encontre a distribuição *a priori* de Jeffreys para λ e (b) determine se a distribuição *a posteriori* é própria.

Solução: (a) Como a amostra é i.i.d. basta encontrarmos a informação de Fisher para X_i. Tomando o log da função de verossimilhança para X_i dada em (7) temos

$$lnf(x_i \mid \lambda) = ln\lambda - \lambda x_i$$

Derivando com respeito a λ temos

$$\frac{d}{d\lambda} lnf(x_i \mid \lambda) = 1/\lambda - x_i$$

$$\frac{d^2}{d^2\lambda} lnf(x_i \mid \lambda) = -1/\lambda^2.$$

Dist. não-informativas: Sentido Jeffreys

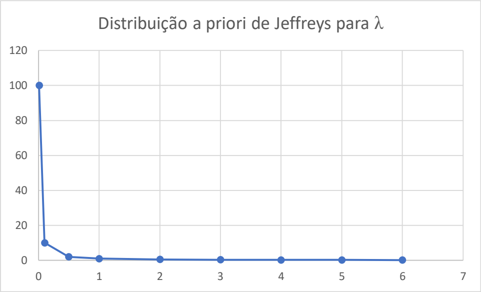
Dai, a Informação de Fisher sobre λ é

$$I(\lambda) = E_{X_i|\lambda}\left(-\frac{d^2}{d^2\lambda} lnf(x_i \mid \lambda)\right) = E_{X_i|\lambda}(\lambda^{-2}) = \lambda^{-2}$$

Por definição, a distribuição *a priori* de Jeffreys para λ é

$$\pi(\lambda) \propto I(\lambda)^{1/2} = (\lambda^{-2})^{1/2} = 1/\lambda, \quad \lambda > 0. \tag{2}$$

$\Rightarrow \lambda \sim Gama(0,0) \leftarrow$ É imprópria!



Dist. não-informativas: Sentido Jeffreys

Exemplo (cont): (b) A função de verossimilhança é

$$f(\boldsymbol{x} \mid \lambda) = \prod_{i=1}^{n} \lambda \exp\{-\lambda x_i\}$$
$$= \lambda^n \exp\{-\lambda \sum_{i=1}^{n} x_i\} \quad (3)$$

Considerando a distribuição *a priori* encontrada no item (a), temos que a distribuição *a posteriori* para λ é:

$$\pi(\lambda \mid \boldsymbol{x}) \propto \lambda^n \exp\{-\lambda \sum_{i=1}^{n} x_i\} * 1/\lambda$$
$$\propto \lambda^{n-1} \exp\{-\lambda \sum_{i=1}^{n} x_i\} \quad (4)$$

Note de (4) que a distribuição *a posteriori* para λ é uma distribuição $Gama(n, \sum_{i=1}^{n} x_i)$. É própria se $n > 0$ e $\sum_{i=1}^{n} x_i > 0$.

Dist. não-informativas: Sentido Jeffreys

Observação: Considere o log da função de verossimilhança em (3)

$$lnf(\boldsymbol{x} \mid \lambda) = nln\lambda - \lambda \sum_{i=1}^{n} x_i \tag{5}$$

Derivando com respeito a λ temos

$$\frac{d}{d\lambda} lnf(x_i \mid \lambda) = n/\lambda - \sum_{i=1}^{n} x_i$$

$$\frac{d^2}{d^2\lambda} lnf(x_i \mid \lambda) = -n/\lambda^2.$$

Ou seja, como a amostra é iid, temos que

$$I_n(\lambda) = n(1/\lambda^2) = nI(\lambda),$$

como afirmado na teoria.

Dist. não-informativas: Sentido Jeffreys

A distribuição *a priori* de Jeffreys

* atribui mais peso aos valores de θ que são mais diferenciados pelo modelo estatístico.
* "Não é subjetiva´´ desde que não resume a informação inicial sobre θ.
* A única subjetividade envolvida em sua construção vem da subjetividade inherente à construção do modelo.
* Viola alguns principios Bayesianos (Princípio da Verosimilhança) pois considera o princípio da repetibilidade amostral.
* Também pode ser imprópria.
* Qual é o ganho deste método se comparado ao método de Bayes-Laplace?

Dist. não-informativas: Sentido Jeffreys

* É invariante sob transformações um-a-um.

Proposição: Se a distribuição *a priori* de θ é não informativa no sentido de Jeffreys e se $\phi = g(\theta)$, onde g é uma tranformação bijetiva de θ, então, *a priori*

$$\pi(\phi) \propto [I(\phi)]^{1/2}.$$

A prova segue utilizando-se o método do jacobiano e a seguinte propiedade da informação de Fisher

$$I(\phi) = I(\theta) \left(\frac{d\theta}{d\phi}\right)^2.$$

Dist. não-informativas: Sentido Jeffreys

Exemplo (cont.): Se $X_1, \ldots, X_n$, dado λ, são variáveis aleatórias i.i.d com distribuição exponencial com parametro λ cuja f.d.p. é

$$f(x_i \mid \lambda) = \lambda \exp\{-\lambda x_i\}, \quad x_i > 0. \tag{6}$$

Considere a seguinte reparametrização do modelo $\theta = 1/\lambda$ e encontre a distribuição *a priori* de Jeffreys para θ.

Solução 1-Reparametrizando a fdp: Sob a re-parametrização dada temos que a fdp torna-se

$$f(x_i \mid \theta) = 1/\theta \exp\{-x_i/\theta\}, \quad x_i > 0.$$

Daí, $\ln f(x_i \mid \theta) = -\ln\theta - x_i/\theta \Rightarrow \frac{d}{d\theta} \ln f(x_i \mid \theta) = -1/\theta + x_i/\theta^2$
$\Rightarrow \frac{d^2}{d\theta^2} \ln f(x_i \mid \theta) = 1/\theta^2 - 2x_i/\theta^3$.
Com isto, temos que a Informação de Fisher sobre θ é

$$I(\theta) = -1/\theta^2 + 2E(X_i)/\theta^3 = -1/\theta^2 + 2\theta/\theta^3 = 1/\theta^2$$

Portanto, distribuição *a priori* de Jeffreys para θ é

$$\pi(\theta) \propto \theta^{-1}, \quad \theta > 0.$$

Dist. não-informativas: Sentido Jeffreys

Solução 2- usando a propriedade da Informação de Fisher:

A Informação de Fisher para λ é

$$I(\lambda) = 1/\lambda^2, \quad \lambda > 0.$$

Usando a propriedade da Informação de Fisher temos que

$$I(\theta) = I(\lambda)\left(\frac{d\lambda}{d\theta}\right)^2.$$

Agora $\theta = 1/\lambda \Rightarrow \lambda = 1/\theta$.

Substituindo na expressao anterior temos

$$I(\theta) = \left(\frac{1}{1/\theta}\right)^2 \left(\frac{d}{d\theta}(1/\theta)\right)^2 = (\theta)^2(-1/\theta^2)^2 = 1/\theta^2$$

Chegamos ao mesmo resultado mostrado na Solução 1 para a Informação de Fisher.

Dist. não-informativas: Sentido Jeffreys

Solução 3-Partindo da distribuição de Jeffreys para λ: Da equação (2) obtemos que a distribuição *a priori* de Jeffreys para λ é

$$\pi_L(\lambda) \propto 1/\lambda, \quad \lambda > 0.$$

Assim, utilizando resultados de calculo de probabilidade para encontramos a distribuição de θ temos que $\forall\theta : \theta = g(\lambda)$ para algum $\lambda > 0$, então $\pi_T(\theta) \propto \pi_L(g^{-1}(\theta)|\frac{d}{d\theta}g^{-1}(\theta)|$. Daí, temos que

$$\theta = 1/\lambda = g(\lambda) \Rightarrow \lambda = g^{-1}(\theta) = 1/\theta \Rightarrow \frac{d}{d\theta}g^{-1}(\theta) = -1/\theta^2$$

$$\pi_T(\theta) \propto \pi_L(1/\theta)| - 1/\theta^2| = \theta/\theta^2$$

Portanto, distribuição *a priori* de Jeffreys para θ é

$$\pi(\theta) \propto \theta^{-1}, \quad \theta > 0.$$

$\Rightarrow$ o principio de "não" informação é preservado!

Dist. não-informativas: Sentido Jeffreys

Alguns casos particulares de distribuição *a priori* de Jeffreys.

* Família exponencial uni-paramétrica: Se $f(x \mid \theta) = h(x) \exp\{\theta T(x) - \psi(\theta)\}$ $\theta \in \mathbb{R}$, então a distribuição *a priori* de Jeffreys é

$$\pi(\theta) \propto \left[\frac{d^2\psi(\theta)}{d\theta}\right]^{1/2} = \left[\psi''(\theta)\right]^{1/2}$$

.

Note que $\ln f(x \mid \theta) = \ln h(x) + \theta T(x) - \psi(\theta)$. Daí

$$\begin{aligned} \frac{d}{d\theta} \ln f(x \mid \theta) &= T(x) - \psi'(\theta) \\ \frac{d^2}{d\theta^2} \ln f(x \mid \theta) &= -\psi''(\theta) \end{aligned}$$

$$\Rightarrow I(\theta) = -E_{X|\theta}(-\psi''(\theta)) = -\psi''(\theta) \Rightarrow \pi(\theta) \propto [\psi''(\theta)]^{1/2}$$

Dist. não-informativas: Sentido Jeffreys

Exemplo: Se X, dado λ, tem distribuição exponencial com parâmetro λ então

$$f(x \mid \lambda) = \lambda \exp\{-\lambda x\} = \exp\{-\lambda x + ln\lambda\} \tag{7}$$

Neste caso, $\psi(\lambda) = -ln\lambda \Rightarrow \psi'(\lambda) = -1/\lambda \Rightarrow \psi''(\lambda) = 1/\lambda^2$. Usando o resultado anterior segue que a distribuição *a priori* de Jeffreys para λ é

$$\pi(\lambda) \propto \lambda^{-1}, \ \ \lambda > 0.$$

Dist. não-informativas: Sentido Jeffreys

- $*$ Família Localização: Se $f(x \mid \theta) = f(x - \theta)$ e, portanto, θ é um parâmetro de localização então $\pi(\theta) \propto 1$.

Exemplo: Suponha que $X \mid \mu \sim N(\mu, \sigma^2)$, com σ^2 conhecido. Então, a função de verossimilhança é

$$f(x|\mu) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \exp\left\{\frac{-1}{2\sigma^2}(X-\mu)^2\right\} \leftarrow f(x-\mu)$$

O logaritmo da função de verossimilhança é

$$\ln f(x|\mu) = \frac{1}{2}\ln\left(\frac{1}{2\pi\sigma^2}\right) - \frac{1}{2\sigma^2}(X-\mu)^2$$

Daí

$$\frac{d}{d\mu} f(x|\mu) = \frac{(X-\mu)}{\sigma^2} \Rightarrow \frac{d^2}{d\mu^2} f(x|\mu) = -\frac{1}{\sigma^2}$$

Dist. não-informativas: Sentido Jeffreys

A informação de Fisher com respeito a μ é

$$I(\mu) = -E_{X|\mu}\left(-\frac{1}{\sigma^2}\right) = \frac{1}{\sigma^2} \leftarrow \text{constante}$$

Por definição, a distribuição *a priori* de Jeffreys para μ é

$$\pi(\mu) \propto \left[\frac{1}{\sigma^2}\right]^{1/2} \propto 1$$

- É imprópria e coincide com a distribuição não-informativa no sentido de Bayes-Laplace.
- Como vimos, a distribuição *a posteriori* para μ é

$$\mu \mid \mathrm{x} \sim \mathrm{Normal}(\bar{x}, \sigma^2/n)$$

a qual é uma distribuição própria.

Dist. não-informativas: Sentido Jeffreys

- ∗ Família Escala :Se $f(x \mid \theta) = \frac{1}{\theta} f(\frac{x}{\theta})$ e, portanto, θ é um parâmetro de escala então $\pi(\theta) \propto 1/\theta$.

Exemplo: Suponha que $X \mid \sigma^2 \sim N(\mu, \sigma^2)$, com μ conhecido. Então, a função de verossimilhança é

$$f(x|\sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \exp\left\{\frac{-1}{2\sigma^2}(X-\mu)^2\right\} \leftarrow 1/\sigma^2 f(x/\sigma^2)$$

O logaritmo da função de verossimilhança é

$$\ln f(x|\sigma^2) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\sigma^2\right) - \frac{1}{2\sigma^2}(X-\mu)^2$$

Daí, a primeira e segunda derivadas com respeito a σ^2 são:

$$\frac{d}{d\sigma^2} f(x|\sigma^2) = \frac{-1}{2\sigma^2} + \frac{(X-\mu)^2}{2(\sigma^2)^2}$$

e

$$\frac{d^2}{d(\sigma^2)^2} f(x|\sigma^2) = \frac{1}{2(\sigma^2)^2} - \frac{(X-\mu)^2}{(\sigma^2)^3}$$

Dist. não-informativas: Sentido Jeffreys

A informação de Fisher com respeito a σ^2 é

$$\begin{aligned} I(\sigma^2) &= -E_{X|\sigma^2}\left(\frac{1}{2(\sigma^2)^2} - \frac{(X-\mu)^2}{(\sigma^2)^3}\right) = \\ &= -\frac{1}{2(\sigma^2)^2} + \frac{E_{X|\sigma^2}(X-\mu)^2}{(\sigma^2)^3} = \frac{1}{2(\sigma^2)^2} \end{aligned} \tag{8}$$

Por definição, a distribuição *a priori* de Jeffreys para σ^2 é

$$\pi(\sigma^2) \propto \left[\frac{1}{2(\sigma^2)^2}\right]^{1/2} \propto \frac{1}{\sigma^2}$$

Análise de referência nos Datos Genéticos

* Dados: 34 individuos com Síndrome de Down (6, 22 e 6, respectivamente, com 1, 2 e 3 alelos distintos no loco de interesse).
* frequências alélicas: 0.12, 0.45, 0.09, 0.31, 0.01 e 0.02.

Para $\phi \in (0,1)$, a verossimilhança

$$P(\boldsymbol{Y} = \boldsymbol{y} \mid \phi) = \frac{n!}{y_1!y_2!y_3!}[\theta_1(\phi)]^{y_1}[\theta_2(\phi)]^{y_2}[\theta_3(\phi)]^{y_3}, \qquad (9)$$

* Como o espaço paramétrico para ϕ é o intervalo $(0,1)$, a distribuição *a priori* para ϕ, não-informativa no sentido de Bayes-Laplace é

$$\phi \sim Uniforme(0,1)$$

.

A informação de Fisher para ϕ associada ao modelo em (9) é

$$I(\phi) = n\left(\frac{a^2}{\theta_1(\phi)} + \frac{b^2}{\theta_2(\phi)} + \frac{c^2}{\theta_3(\phi)}\right),$$

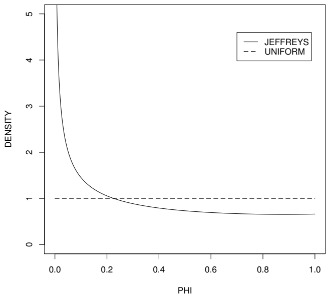
onde $a = \sum_{i=1}^m p_i^3 - \sum_{i=1}^m p_i^2$ e, $b = \sum_{i,j=1}^m p_i^2 p_j - \sum_{i,j=1}^m p_i p_j$ e $c = \sum_{i,j,k=1}^m p_i p_j p_k$, para $i \neq j \neq k$. Consequentemente, a distribuição de Jeffreys para ϕ é

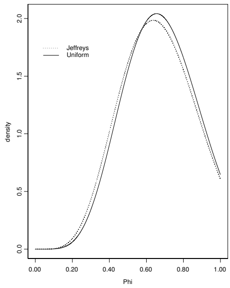
$$\pi(\phi) \propto \left(\frac{a^2}{\theta_1(\phi)} + \frac{b^2}{\theta_2(\phi)} + \frac{c}{\phi}\right)^{1/2}. \tag{10}$$

Esta distribuição é própria.

Análise de referência nos Datos Genéticos

Distribuições *a priori* de referência e *a posteriori* para ϕ.





mi Apesar das distribuições *a priori* serem diferentes, as informações que trazem sobre ϕ é fraca, gerando distribuições *a posteriori* muito similares.

Análise de referência nos Datos Genéticos

Table: Estimativas *a posteriori* para ϕ

Dist. *a prior*	Resumos *a posteriori*		
	Media	variância	Moda
Jeffreys	0.6417	0.0322	0.6374
Uniforme	0.6549	0.0305	0.6552

* Estimador de máxima verossimilhança e variância assintótica são 0.6552 e 0.0481, respectivamente.