Inferência Estatística com Abordagem Bayesiana

Rosangela Helena Loschi [1]

[1]Departamento de Estatística
Universidade Federal de Minas Gerais

18 de novembro de 2021

INFERÊNCIA NA FAMÍLIA NORMAL COM MÉDIA E VARIANCIA DESCONHECIDAS

Inferência na família $N(\mu, \sigma^2)$

Consideremos o caso em que $X_1, \ldots, X_n$, dado μ e σ^2, são variáveis aleatórias i.i.d com distribuição Normal(μ, σ^2), em que μ e σ^2 são desconhecidas.

Nossa meta é determinar

- a família conjugada natural
- a distribuição não informativa no Sentido Bayes-Laplace
- a distribuição não informativa no Sentido de Jeffreys.

Sob as suposições do problema temos que o núcleo da função de verossimilhança é

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

Inferência na família $N(\mu, \sigma^2)$: Conjugação

- Note que $g(\sigma^2)$ é o núcleo de uma distribuição Gama-Invertida e $h(\mu \mid \sigma^2)$ é o núcleo de uma distribuição normal.
- A fdp de $\sigma^2 \sim Gama - Inversa(a, d), a > 0, d > 0$ é

$$\pi(\sigma^2) = \frac{(a/2)^{d/2}}{\Gamma(d/2)}(\sigma^2)^{-\frac{d+2}{2}} \exp\left\{-\frac{a}{2\sigma^2}\right\}, \ \sigma^2 > 0.$$

- A esperança e variância de $\sigma^2 \sim Gama - Inversa(a, d)$ são, respectivamente,

$$\begin{aligned} E(\sigma^2) &= \frac{a}{d-2}, \ \text{se} \ d > 2 \\ V(\sigma^2) &= \frac{2a^2}{(d-2)^2(d-4)}, \ \text{se} \ d > 4. \end{aligned}$$

- Por construção, a família de distribuições para (μ, σ^2) que é candidata a ser conjugada da família $N(\mu, \sigma^2)$, (μ, σ^2) desconhecidos é

$$\begin{aligned} \mu \mid \sigma^2 &\sim Normal(m, V\sigma^2), \ m \in \mathbb{R}, V > 0 \qquad (1) \\ \sigma^2 &\sim Gama - Inversa(a, d), a > 0, d > 0 \qquad (2) \end{aligned}$$

Inferência na família $N(\mu, \sigma^2)$: Conjugação

- Esta distribuição *a priori* impõe uma relação de dependência entre μ e σ^2
- Temos que verificar se é uma família fechada por produto.

Cálculo da distribuição *a posteriori* conjunta de (μ, σ^2):

$$
\begin{aligned}
\pi(\mu, \sigma^2 \mid \boldsymbol{x}) \quad & \propto \quad \left(\frac{1}{\sigma^2}\right)^{n/2} \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right\} \exp\left\{-\frac{n}{2\sigma^2}\left(\mu^2 - 2\mu\bar{x}\right)\right\} \\
& \times \quad \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2V\sigma^2}(\mu^2 - 2\mu m + m^2)\right\} \\
& \times \quad \left(\frac{1}{\sigma^2}\right)^{\frac{d+2}{2}} \exp\left\{-\frac{a}{2\sigma^2}\right\} \\
& \propto \quad \left(\frac{1}{\sigma^2}\right)^{\frac{n+d+2}{2}} \exp\left\{-\frac{1}{2\sigma^2}(a + \sum_{i=1}^{n} x_i^2 + m^2/V)\right\} \\
& \times \quad \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2V\sigma^2}(\mu^2 - 2\mu m) - \frac{n}{2\sigma^2}\left(\mu^2 - 2\mu\bar{x}\right)\right\}
\end{aligned}
\tag{3}
$$

Inferência na família $N(\mu, \sigma^2)$: Conjugação

$$
\begin{aligned}
\pi(\mu, \sigma^2 \mid \boldsymbol{x}) &\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n+d+2}{2}} \exp\left\{-\frac{1}{2\sigma^2}(a + \sum_{i=1}^{n} x_i^2 + m^2/V)\right\} \\
&\times \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[\mu^2(n + 1/V) - 2\mu(n\bar{x} + m/V)\right]\right\} \\
&\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n+d+2}{2}} \exp\left\{-\frac{1}{2\sigma^2}(a + \sum_{i=1}^{n} x_i^2 + m^2/V)\right\} \\
&\times \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[\mu^2\left(\frac{nV+1}{V}\right) - 2\mu\left(\frac{nV\bar{x}+m}{V}\right)\right]\right\} \\
&\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n+d+2}{2}} \exp\left\{-\frac{1}{2\sigma^2}(a + \sum_{i=1}^{n} x_i^2 + m^2/V)\right\} \\
&\times \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{nV+1}{2V\sigma^2}\left[\mu^2 - 2\mu\frac{(nV\bar{x}+m)/V}{(nV+1)/V}\right]\right\}
\end{aligned}
\tag{4}
$$

Inferência na família $N(\mu, \sigma^2)$: Conjugação

$$
\begin{aligned}
\pi(\mu, \sigma^2 \mid \boldsymbol{x}) & \\
\propto & \left(\frac{1}{\sigma^2}\right)^{\frac{n+d+2}{2}} \exp\left\{-\frac{1}{2\sigma^2}(a + \sum_{i=1}^{n} x_i^2 + m^2/V) + \frac{nV+1}{2V\sigma^2}\left(\frac{nV\bar{x}+m}{nV+1}\right)^2\right\} \\
\times & \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{nV+1}{2V\sigma^2}\left[\mu^2 - 2\mu\frac{nV\bar{x}+m}{nV+1} + \left(\frac{nV\bar{x}+m}{nV+1}\right)^2\right]\right\} \\
\propto & \color{blue}{\left(\frac{1}{\sigma^2}\right)^{\frac{n+d+2}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(a + \sum_{i=1}^{n} x_i^2 + \frac{m^2}{V} - \frac{(nV\bar{x}+m)^2}{v(nV+1)}\right)\right\}} \\
\times & \color{red}{\left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{nV+1}{2V\sigma^2}\left[\mu - \left(\frac{nV\bar{x}+m}{nV+1}\right)\right]^2\right\}}
\end{aligned}
$$

- Na expressao acima a parte azul é o núclo de uma Gama-Inversa e a parte em vermelho é o núcleo de uma normal.
- Notamos o fechamento por produto da família de distribuições para (μ, σ^2) que constuímos.

Inferência na família $N(\mu, \sigma^2)$: Conjugação

- Assim, a família de distribuições para (μ, σ^2) que é conjugada natural da família de distribuições amostrais $N(\mu, \sigma^2)$, com ambos os parâmetros desconhecidos é:

$$\mu \mid \sigma^2 \sim N(m, V\sigma^2) \tag{5}$$

$$\sigma^2 \sim Gama - Inversa(a, d) \tag{6}$$

- A distribuição *a posteriori* é

$$\mu \mid \sigma^2, \boldsymbol{x} \sim N\left(\frac{nV\bar{x} + m}{nV + 1}, \frac{V}{nV + 1}\sigma^2\right) = N(m^*, V^*\sigma^2)$$

$$\sigma^2 \mid \boldsymbol{x} \sim Gama - Inversa\left(a + \sum_{i=1}^{n} x_i^2 + \frac{m^2}{V} - \frac{(nV\bar{x} + m)^2}{v(nV + 1)}, d + n\right)$$

$$= Gama - Inversa(a^*, d^*)$$

Inferência na família $N(\mu, \sigma^2)$: Conjugação

Comentários relevantes:

Usando propriedades de Esperança e esperança condicional obtemos:

- A média e a variância *a posteriori* de μ

$$
\begin{aligned}
E(\mu \mid \boldsymbol{x}) &= E_{\sigma^2|\boldsymbol{x}}(E(\mu \mid \sigma^2, \boldsymbol{x})) = E_{\sigma^2|\boldsymbol{x}}\left(\frac{nV\bar{x}+m}{nV+1}\right) = \frac{nV\bar{x}+m}{nV+1} \\
V(\mu \mid \boldsymbol{x}) &= E_{\sigma^2|\boldsymbol{x}}(V(\mu \mid \sigma^2, \boldsymbol{x})) + V_{\sigma^2|\boldsymbol{x}}(E(\mu \mid \sigma^2, \boldsymbol{x})) \\
&= \frac{V}{nV+1}E(\sigma^2 \mid \boldsymbol{x}) + V(\frac{nV\bar{x}+m}{nV+1}) = \frac{V}{nV+1}E(\sigma^2 \mid \boldsymbol{x})
\end{aligned}
$$

- A esperança *a posteriori* de σ^2

$$
E(\sigma^2 \mid \boldsymbol{x}) = \frac{a(d-2)}{(d-2)(d+n-2)} + \frac{nS^2}{d+n-2} + \frac{n(m+\bar{x})^2}{(d+n-2)(nV+1)}
$$

onde $S^2 = \sum_{i=1}^{n}(xi-\bar{x})^2/n$ é o EMV para σ^2.

Inferência na família $N(\mu, \sigma^2)$: Conjugação

Comentários relevantes:

- Se $n \to \infty$ (forte informação dos dados) segue que

$$\begin{aligned} E(\mu \mid \boldsymbol{x}) &\to \bar{x} \\ V(\mu \mid \boldsymbol{x}) &\to 0 \\ E(\sigma^2 \mid \boldsymbol{x}) &\to S^2 \end{aligned}$$

- Se $V \to \infty$ (informação vaga sobre μ) segue que

$$\begin{aligned} E(\mu \mid \boldsymbol{x}) &\to \bar{x} \\ V(\mu \mid \boldsymbol{x}) &= \frac{1}{n} \lim_{V \to \infty} E(\sigma^2 \mid \boldsymbol{x}) \\ E(\sigma^2 \mid \boldsymbol{x}) &\to \frac{d-2}{d+n-2} \frac{a}{d-2} + \frac{n}{d+n-2} S^2 \\ &= \frac{d-2}{d+n-2} E(\sigma^2) + \frac{n}{d+n-2} S^2 \end{aligned}$$

Comentários relevantes:

- Se $V \to \infty$, $a \to 0$ e $d \to 0$ (informação vaga sobre ambos μ e σ^2) segue que

$$\begin{aligned} E(\mu \mid \boldsymbol{x}) &\to \bar{x} \\ V(\mu \mid \boldsymbol{x}) &\to \frac{1}{n-2} S^2 \\ E(\sigma^2 \mid \boldsymbol{x}) &\to \frac{n}{n-2} S^2 \end{aligned}$$

- Se $V = 0$ (informação contundente sobre μ) segue que

$$\begin{aligned} E(\mu \mid \boldsymbol{x}) &= m \\ V(\mu \mid \boldsymbol{x}) &= 0 \\ E(\sigma^2 \mid \boldsymbol{x}) &= \frac{(d-2)}{(d+n-2)} E(\sigma^2) + \frac{nS^2}{d+n-2} + \frac{n(m+\bar{x})^2}{(d+n-2)} \end{aligned}$$

Inferência na família $N(\mu, \sigma^2)$:conjugação

Se, *a priori*, temos que

$$\mu \mid \sigma^2 \sim Normal(m, V\sigma^2), \; m \in \mathbb{R}, V > 0 \qquad (7)$$

$$\sigma^2 \sim Gama-Inversa(a, d), a > 0, d > 0, \qquad (8)$$

- a distribuição *a priori* para μ é uma distribuição t-Student ($\mu \sim t(m, V, a, d)$) com fdp dada por:

$$\pi(\mu) = \left(\frac{1}{\pi}\right)^{1/2} \frac{\Gamma\left(\frac{d+1}{2}\right) a^{d/2}}{\Gamma\left(\frac{d}{2}\right) V^{1/2}} \left[a + \frac{(\mu - m)^2}{V}\right]^{-\frac{d+1}{2}}$$

- a média e variância *a priori* para μ são

$$E(\mu) = m, \qquad V(\mu) = \frac{aV}{d-2}$$

- O número de graus de liberdade é d.
- Como consequencia da conjugação, *a posteriori*, μ tem distribuição t-Student $\mu \mid \boldsymbol{x} \sim t(m^*, V^*, a^*, d^*)$ (ver slide 8 para os parâmetros).

A distribuição marginal de μ

FATO: Suponha que

$$\begin{aligned} \mu \mid \sigma^2 &\sim Normal(m, V\sigma^2),\ \ m \in \mathbb{R}, V > 0 & (9) \\ \sigma^2 &\sim Gama - Inversa(a, d), a > 0, d > 0. & (10) \end{aligned}$$

Prove que a distribuição marginal para μ é uma distribuição t-Student ($\mu \sim t(m, V, a, d)$) com fdp dada por:

$$\pi(\mu) = \left(\frac{1}{\pi}\right)^{1/2} \frac{\Gamma\left(\frac{d+1}{2}\right) a^{d/2}}{\Gamma\left(\frac{d}{2}\right) V^{1/2}} \left[a + \frac{(\mu - m)^2}{V}\right]^{-\frac{d+1}{2}}$$

- este é um resultado de Cálculo de probabilidade.
- Para prová-lo devemos lembrar que toda distribuição pode ser obtida integrando-se uma distribuição conjunta.

A distribuição marginal de μ

$$
\begin{aligned}
\pi(\mu) &= \int_0^\infty \pi(\mu \mid \sigma^2)\pi(\sigma^2)d\sigma^2 \\
&= \int_0^\infty \left(\frac{1}{2\pi V\sigma^2}\right)^{1/2} \exp\left\{-\frac{(\mu-m)^2}{2V\sigma^2}\right\} \\
&\times \frac{(a/2)^{d/2}}{\Gamma(d/2)}\left(\frac{1}{\sigma^2}\right)^{(d+2)/2} \exp\left\{-\frac{a}{2\sigma^2}\right\} d\sigma^2 \\
&= \frac{(a/2)^{d/2}}{\Gamma(d/2)}\left(\frac{1}{2\pi V}\right)^{1/2} \int_0^\infty \left(\frac{1}{\sigma^2}\right)^{\frac{d+3}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(a+\frac{(\mu-m)^2}{V}\right)\right\} d\sigma^2 \\
&= \frac{(a/2)^{d/2}}{\Gamma(d/2)}\left(\frac{1}{2\pi V}\right)^{1/2} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\left[a+\frac{(\mu-m)^2}{V}\right]^{(d+1)/2}(1/2)^{(d+1)/2}} \\
&= \left(\frac{1}{\pi}\right)^{1/2} \frac{\Gamma((d+1)/2)}{\Gamma(d/2)} \frac{(a/2)^{d/2}}{(2V)^{1/2}} 2^{(d+1)/2} \left[a+\frac{(\mu-m)^2}{V}\right]^{-(d+1)/2} \\
&= \; = \left(\frac{1}{\pi}\right)^{1/2} \frac{\Gamma\left(\frac{d+1}{2}\right) a^{d/2}}{\Gamma\left(\frac{d}{2}\right) V^{1/2}} \left[a+\frac{(\mu-m)^2}{V}\right]^{-\frac{d+1}{2}}
\end{aligned}
$$

Inferência na família $N(\mu, \sigma^2)$:conjugação

Neste caso, a distribuição preditiva *a priori* para $\boldsymbol{X} = (X_1 \ldots, X_n)$ é obtida fazendo-se

$$f(\boldsymbol{X}) = \int_0^{\infty} \int_{-\infty}^{\infty} f(\boldsymbol{X} \mid \mu, \sigma^2) \pi(\mu \mid \sigma^2) \pi(\sigma\ 2) d\mu d\sigma^2.$$

Com isto, distribuição preditiva *a priori* para $\boldsymbol{X} = (X_1 \ldots, X_n)$ é é uma distribuição t-student n variada denotada por:

$$\boldsymbol{X} \sim t_n(m1_{n \times 1}, I_n + V1_{n \times n}, d, a)$$

com grau de liberdade d e média e variância dada por:

$$E(\boldsymbol{X}) = m1_{n \times 1} \qquad V(\boldsymbol{X}) = \frac{a}{d-2}(I_n + V1_{n \times n}),$$

onde I_n é a matriz identidade de ordem n e $1_{a \times b}$ denota uma matriz de uns de ordem $a \times b$.

Inferência na família $N(\mu, \sigma^2)$:Dist. de Bayes-Laplace

Neste caso,

* o espaço paramétrico é $\mathbb{R} \times \mathbb{R}_+$.
* a distribuição *a priori* não-informativa no sentido de Bayes-Laplace é $\pi(\mu, \sigma^2) \propto 1$, $\forall (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+ \leftarrow$ é imprópria.

- Segue que o núcleo da distribuição *a posteriori* para (μ, σ^2) é

$$\begin{aligned} \pi(\mu, \sigma^2 \mid \mathrm{x}) &\propto \left(\frac{1}{\sigma^2}\right)^{n/2} \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right\} \exp\left\{-\frac{n}{2\sigma^2}\left(\mu^2 - 2\mu\bar{x}\right)\right\} \\ &\propto \left(\frac{1}{\sigma^2}\right)^{(n-1)/2} \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right\} \\ &\times \left(\frac{1}{\sigma^2}\right)^{1/2} \exp\left\{-\frac{n}{2\sigma^2}\left(\mu^2 - 2\mu\bar{x}\right)\right\} \end{aligned} \tag{11}$$

Inferência na família $N(\mu, \sigma^2)$:Dist. de Bayes-Laplace

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

Inferência na família $N(\mu, \sigma^2)$:Dist. de Bayes-Laplace

- *A posteriori*, temos que

$$\begin{aligned} \mu \mid \sigma^2, \boldsymbol{x} &\sim Normal(\bar{x}, \sigma^2/n) \\ \sigma^2 \mid \boldsymbol{x} &\sim Gamma - Inversa\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2, n-3\right) \end{aligned}$$

- Note que a distribuição *a posteriori* de σ^2 é própria somente se

$$\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 > 0 \quad \text{e} \quad n > 3$$

.

- Isto impõe que devemos ter mais que 3 elementos na amostra.
- Além disto, se $x_i = 0$ para todo i, teremos uma distribuição imprópria e não haverá inferência *a posteriori*.

Dist. de Jeffreys: Método de Jeffreys

Se $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$ existe dois métodos para encontrar a distribuição conjunta *a priori* de Jeffreys para $\boldsymbol{\theta}$

Método de Jeffreys:

- 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:

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

- a distribuição conjunta *a priori* de Jeffreys para $\boldsymbol{\theta}$ é

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

Inferência na família $N(\mu, \sigma^2)$:Método de Jeffreys

Como estamos considerando uma amostra iid da na família $N(\mu, \sigma^2)$, vamos calcular a informação de Fisher considerando apenas uma observação

- a verossimilhança

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

- o logaritmo da verossimilhança

$$\ln f = \ln f(x \mid \mu, \sigma^2) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2$$

- as derivadas parciais

$$\frac{\partial \ln f}{\partial \mu} = \frac{x-\mu}{\sigma^2} \qquad \frac{\partial^2 \ln f}{\partial \mu^2} = -\frac{1}{\sigma^2}$$

$$\frac{\partial \ln f}{\partial \sigma^2} = -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2(\sigma^2)^2} \qquad \frac{\partial^2 \ln f}{\partial (\sigma^2)^2} = \frac{1}{2(\sigma^2)^2} - \frac{(x-\mu)^2}{(\sigma^2)^3}$$

$$\frac{\partial \ln f}{\partial \mu \partial \sigma^2} = -\frac{x-\mu}{(\sigma^2)^2}$$

Inferência na família $N(\mu, \sigma^2)$:Método de Jeffreys

- A matriz de informação de Fisher para $\theta = (\mu, \sigma^2)$

$$
\begin{aligned}
I(\theta) &= det E_{x|\mu,\sigma} \begin{bmatrix} \frac{1}{\sigma^2} & \frac{x-\mu}{(\sigma^2)^2} \\ \frac{x-\mu}{(\sigma^2)^2} & -\frac{1}{2(\sigma^2)^2} + \frac{(x-\mu)^2}{(\sigma^2)^3} \end{bmatrix} \\
&= det \begin{bmatrix} E_{x|\mu,\sigma}\left(\frac{1}{\sigma^2}\right) & E_{x|\mu,\sigma}\left(\frac{x-\mu}{(\sigma^2)^2}\right) \\ E_{x|\mu,\sigma}\left(\frac{x-\mu}{(\sigma^2)^2}\right) & E_{x|\mu,\sigma}\left(-\frac{1}{2(\sigma^2)^2} + \frac{(x-\mu)^2}{(\sigma^2)^3}\right) \end{bmatrix} \\
&= det \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & E_{x|\mu,\sigma}\left(-\frac{1}{2(\sigma^2)^2} + \frac{(x-\mu)^2}{(\sigma^2)^3}\right) \end{bmatrix} \\
&= det \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \left(-\frac{1}{2(\sigma^2)^2} + \frac{\sigma^2}{(\sigma^2)^3}\right) \end{bmatrix} = det \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \left(\frac{1}{2(\sigma^2)^2}\right) \end{bmatrix} \\
&= \frac{1}{2(\sigma^2)^3} \qquad (12)
\end{aligned}
$$

- A distribuição *a priori* de Jeffreys para (μ, σ^2) é

$$\pi(\mu, \sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{3/2} \leftarrow \text{impropria}$$

Inferência na família $N(\mu, \sigma^2)$:Método de Jeffreys

- A ditribuição *a posteriori* assumindo a distribuição *a priori* de Jeffreys é

$$
\begin{aligned}
\mu \mid \sigma^2, \boldsymbol{x} &\sim N\left(\bar{x}, \frac{\sigma^2}{n}\right) \\
\sigma^2 \mid \boldsymbol{x} &\sim Gama - Inversa\left(\sum_{i=1}^{n}(x_i - \bar{x})^2, n\right)
\end{aligned}
\tag{13}
$$

- É uma distribuição própria sempre que $n > 0$ e $\sum_{i=1}^{n}(x_i - \bar{x})^2 > 0$.
- Como no caso anterior, se $x_i = 0$ para todo i, não teremos inferência *a posteriori* pois tal distribuição será imprópria.

Dist. de Jeffreys: Método Jeffreys-Independente

Método Jeffreys-Independente: seja $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p) \in \Theta \subset \mathbb{R}^p$

- Assume-se que $\theta_1, \ldots, \theta_p$ são independentes
- Encontre a distribuição *a priori* $\pi(\theta_i)$ de cada θ_i considerando os demais θs conhecidos.
- A distribuição *a priori* de Jeffreys para $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ é

$$\pi(\boldsymbol{\theta}) \propto \pi(\theta_1) \ldots . \pi(\theta_p)$$

Inferência na família $N(\mu, \sigma^2)$:Dist. de Jeffreys-independente

Estamos considerando uma amostra iid da família $N(\mu, \sigma^2)$.

- assuma μ independente de σ^2.
- a verossimilhança

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

- o logaritmo da verossimilhança

$$\ln f = \ln f(x \mid \mu, \sigma^2) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2$$

- Informação de Fisher para μ

$$\begin{aligned} \frac{d\ln f}{d\mu} &= \frac{x-\mu}{\sigma^2} \qquad \frac{d^2\ln f}{d\mu^2} = -\frac{1}{\sigma^2} \\ I(\mu) &= \frac{1}{\sigma^2} \leftarrow \text{constante} \end{aligned}$$

- A distribuição de Jeffeys para μ

$$\pi(\mu) \propto 1$$

Inferência na família $N(\mu, \sigma^2)$:Dist. de Jeffreys-independente

- A informação de Fisher para σ^2

$$
\begin{aligned}
\frac{d \ln f}{\partial \sigma^2} &= -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2(\sigma^2)^2} \quad \frac{d^2 \ln f}{d(\sigma^2)^2} = \frac{1}{2(\sigma^2)^2} - \frac{(x-\mu)^2}{(\sigma^2)^3} \\
I(\sigma^2) &= E_{x|\mu,\sigma}\left(-\frac{1}{2(\sigma^2)^2} + \frac{(x-\mu)^2}{(\sigma^2)^3}\right) = \frac{1}{2(\sigma^2)^2} \qquad (14)
\end{aligned}
$$

- A distribução de Jeffreys para σ^2 é

$$\pi(\sigma^2) \propto \left(\frac{1}{2(\sigma^2)^2}\right)^{1/2} \propto \frac{1}{\sigma^2} \leftarrow \text{impropria}$$

- A distribução de Jeffeys conjunta para (μ, σ^2) usando o método Jeffreys-independente é

$$\pi(\mu, \sigma^2) = \pi(\mu)\pi(\sigma^2) \propto \frac{1}{\sigma^2} \leftarrow \text{impropria}$$

- é a distribuição não-informativa mais utilizada no caso da família Normal com média e variâncias desconhecidas.

Dist. de Jeffreys: Método Jeffreys-Independente

- A ditribuição *a posteriori* conjunta assumindo a distribuição *a priori* Jeffreys-independente é

$$\begin{aligned} \mu \mid \sigma^2, \boldsymbol{x} &\sim N\left(\bar{x}, \sigma^2/n\right) \\ \sigma^2 \mid \boldsymbol{x} &\sim Gama - Inversa\left(\sum_{i=1}^{n}(x_i - \bar{x})^2, n-1\right) \end{aligned} \tag{15}$$

- É própria desde que $n > 1$ e $\sum_{i=1}^{n}(x_i - \bar{x})^2 > 0$.

O Que leva o método Jeffreys-independente à popularidade?

Consideremos o caso em que $X_1, \ldots, X_n$, dado μ e σ^2, são variáveis aleatórias i.i.d com distribuição Normal(μ, σ^2), em que **μ e σ^2 são desconhecidas.**

Método de Jeffreys:

- A distribuição *a priori* de Jeffreys para (μ, σ^2) é

$$\pi(\mu, \sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{3/2}$$

- A distribuição *a posteriori* assumindo a distribuição *a priori* de Jeffreys é

$$\begin{aligned} \mu \mid \sigma^2, \boldsymbol{x} &\sim N\left(\bar{x}, \frac{\sigma^2}{n}\right) \\ \sigma^2 \mid \boldsymbol{x} &\sim Gama - Inversa\left(\sum_{i=1}^{n}(x_i - \bar{x})^2, n\right) \end{aligned} \qquad (16)$$

- Desta distribuição *a posteriori* decorre que

$$\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sigma^2} \mid \boldsymbol{x} \sim \chi^2_{(n)}$$

O Que leva o método Jeffreys-independente à popularidade?

Consideremos agora o caso em que $X_1, \ldots, X_n$, dado σ^2, são variáveis aleatórias i.i.d com distribuição Normal(μ, σ^2), em que **σ^2 é desconhecido e μ é conhecido.**

- A dist. *a priori* de Jeffreys para σ^2 é

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

- A dist. *a posteriori* assumindo a dist. *a priori* de Jeffreys é

$$\sigma^2 \mid \boldsymbol{x} \sim Gama - Inversa\left(\sum_{i=1}^{n}(x_i - \mu)^2, n\right)$$

- Desta distribuição *a posteriori* decorre que

$$\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{\sigma^2} \mid \boldsymbol{x} \sim \chi^2_{(n)}$$

- O método de Jeffreys fornece uma dist. *a priori* conjunta para (μ, σ^2) que induz uma dist. *a posteriori* para σ^2 que é insensível ao fato da média populacional μ ser ou não conhecida.

O Que leva o método Jeffreys-independente à popularidade?

Suponha que $X_1, \ldots, X_n$, dado μ e σ^2, são variáveis aleatórias i.i.d com distribuição Normal(μ, σ^2), com **μ e σ^2 são desconhecidas.**

Método Jeffreys-independente:

- A ditribuição *a priori* de Jeffreys para (μ, σ^2) é $\pi(\mu, \sigma^2) \propto \frac{1}{\sigma^2}$
- A ditribuição *a posteriori* conjunta assumindo a distribuição *a priori* Jeffreys-independente é

$$
\begin{aligned}
\mu \mid \sigma^2, \boldsymbol{x} &\sim N\left(\bar{x}, \sigma^2/n\right) \\
\sigma^2 \mid \boldsymbol{x} &\sim Gama - Inversa\left(\sum_{i=1}^{n}(x_i - \bar{x})^2, n-1\right)
\end{aligned}
\tag{17}
$$

- Desta distribuição *a posteriori* decorre que

$$
\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sigma^2} \mid \boldsymbol{x} \sim \chi^2_{(n-1)}
$$

- O Método Jeffreys-independente fornece uma distribuição *a priori* conjunta para (μ, σ^2) que induz uma distribuição *a posteriori* para σ^2 que é sensível ao fato da média populacional μ ser desconhecida.