

# SME0809 - Inferência Bayesiana - Distribuição Normal

Grupo 13 - Francisco Miranda - 4402962 - Heitor Carvalho - 11833351

Outubro 2021

## Caso 1: $\mu$ desconhecido e $\sigma$ conhecido

### Verossimilhança da distribuição

$$\mathcal{L}(y|\theta) = \prod_{i=1}^n p(y_i|\theta) = \prod_{i=1}^n e^{-1/2\sigma^2(y_i-\theta)^2}$$

### A priori da distribuição

$$p(\theta) \propto e^{(-1/2\tau_0^2)(\theta-\mu_0)^2}$$

### A posteriori da distribuição

A posteriori é computada assumindo-se que: 1. Cada observação é independentemente distribuída 2. Cada observação tem a mesma variância

$$p(\theta|y) = p(\theta) p(y|\theta) = p(\theta) \prod_{i=1}^n p(y_i|\theta) = e^{(-1/2\tau_0^2)(\theta-\mu_0)^2} \prod_{i=1}^n e^{(-1/2\sigma^2)(y_i-\theta)^2} = e^{(-\frac{1}{2})(1/\tau_0^2)(\theta-\mu_0)^2 + \frac{1}{\sigma^2} \sum_{i=1}^n (y_i-\theta)^2)}$$

Sendo,

$$\mu_n = \frac{\tau_0^{-2}\mu_0 + n\sigma^{-2}\bar{y}}{\tau_0^{-2} + \sigma^{-2}}, \quad \tau_n^{-2} = \tau_0^{-2} + n\sigma^{-2}$$

Podemos reescrever  $p(\theta|y)$  como:

$$p(\theta|y_n) \propto e^{(-1/2\tau_n^2)(\theta-\mu_n)^2}$$

Logo, para uma distribuição Normal com variância conhecida, a média a posteriori  $\mu_1$  pode ser interpretada como a média ponderada da média a priori e o valor observado  $y$ , sendo os pesos proporcionais às precisões de cada um.

Desse modo, a distribuição a posteriori da média  $\theta$  depende apenas da média amostral  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ , sendo assim,  $\bar{y}$  é uma estatística suficiente.

Portanto, para  $n$  observações, a posteriori apresenta a seguinte distribuição:

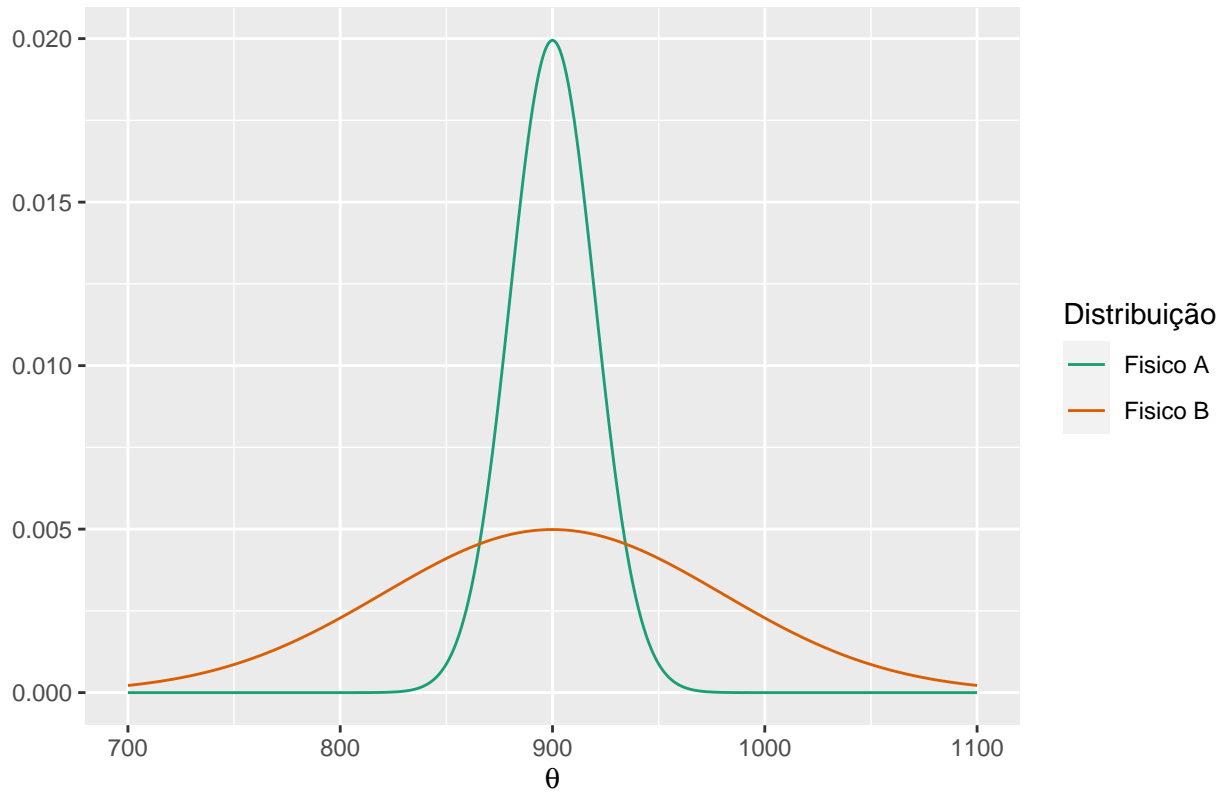
$$p(\theta|y_1, \dots, y_n) = p(\theta | \bar{y}) = \mathcal{N}(\theta|y_n, \tau_n^2)$$

$$f(x|\theta) = e^{-x^2}$$

- a) Faça um esboço do gráfico das distribuições prioris dos dois físicos em um mesmo sistema cartesiano.

Temos  $\theta_A \sim N(900, 20^2)$  e  $\theta_B \sim N(900, 80^2)$ . Assim:

**Distribuição a priori da grandeza estimada pelos físicos**



- b) Encontre a distribuição a posteriori para o físico A e para o físico B.

Como  $X|\theta \sim N(\theta, \sigma^2)$  com  $\sigma^2$  conhecido e  $\theta \sim N(\mu_0, \tau_0^2)$  então  $\theta|x \sim N(\mu_1, \tau_1)$ , sendo

$$\mu_1 = \frac{\tau_0^2 \mu_0 + \sigma^{-2} x}{\tau_0^{-2} + \sigma^{-2}}, \text{ e } \tau_1^{-2} = \tau_0^{-2} + \sigma^{-2}$$

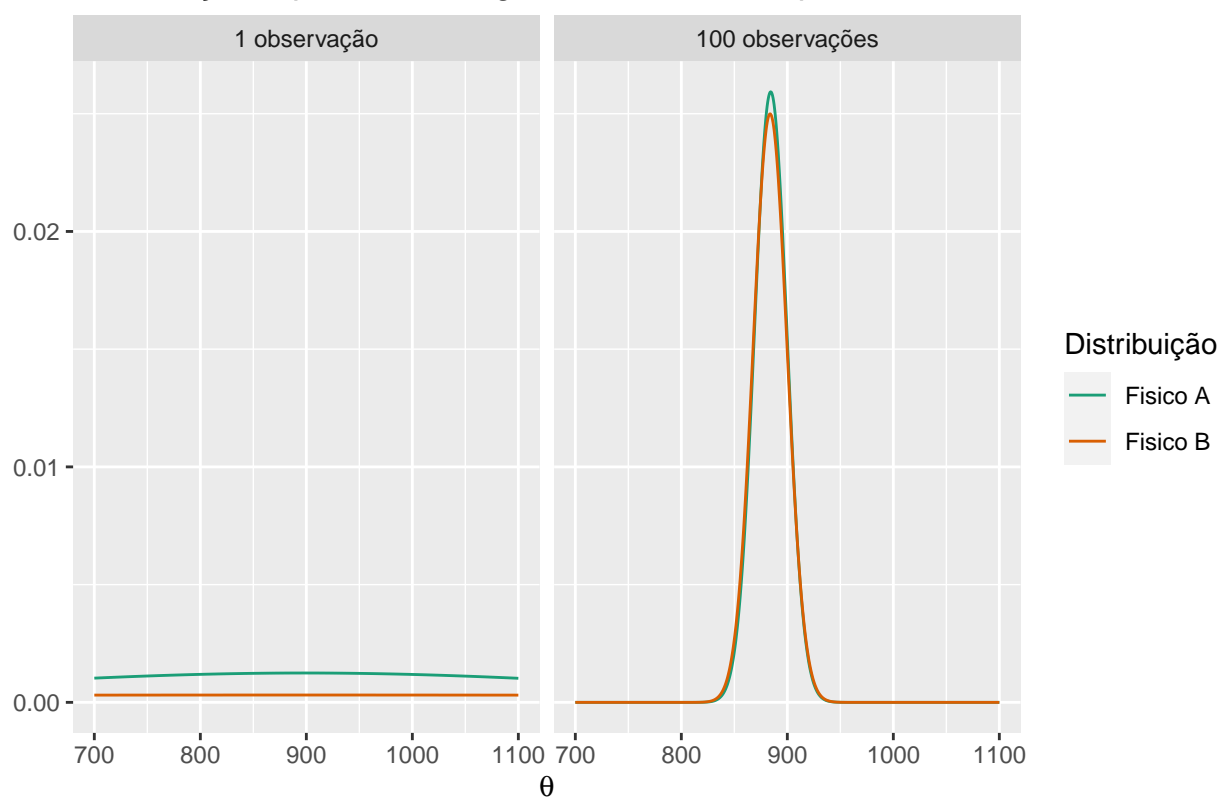
Assim, para 100 observações temos:

$$\theta_A \sim N(884.314, 15.3846), \quad \theta_B \sim N(883.7272, 15.9601)$$

Enquanto que para uma única observação, a posteriori é:

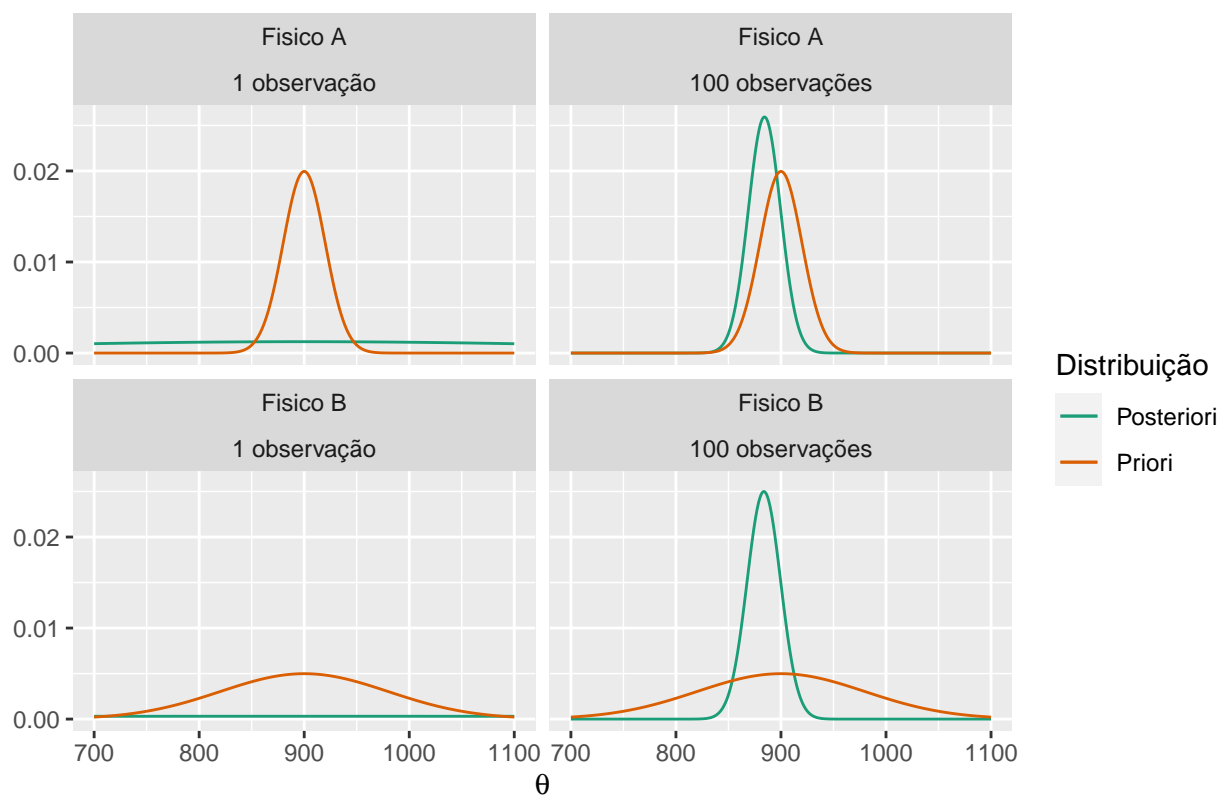
$$\theta_A \sim N(898.4, 320), \quad \theta_B \sim N(893.6, 1280)$$

### Distribuição a posteriori da grandeza estimada pelos físicos

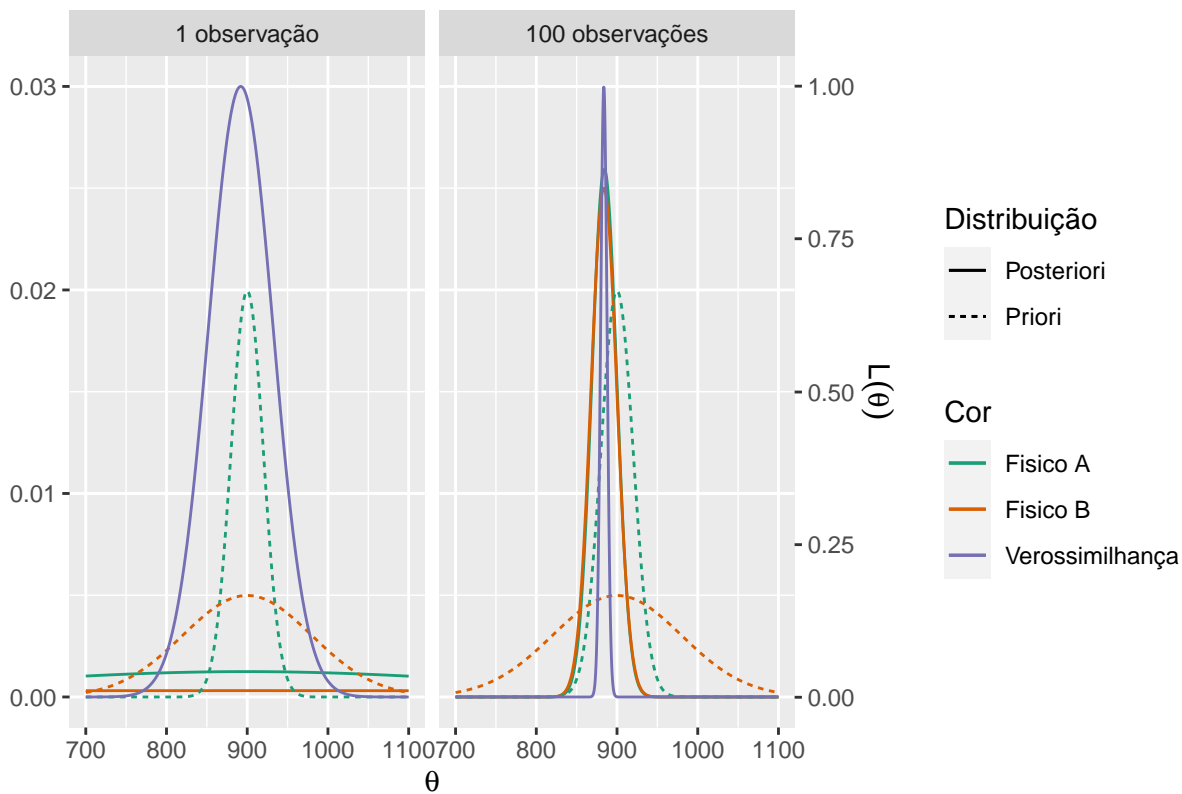


- c) Faça um esboço do gráfico das distribuições: a priori e a posteriori de cada um dos dois físicos em um mesmo sistema cartesiano.

## Distribuição da grandeza estimada pelos físicos



## Distribuição da grandeza estimada pelos físicos



- d) Observando o gráfico, qual físico aprendeu mais com o experimento? Justifique.

Aumentos nas precisões a posteriori em relação às precisões a priori com 100 observações:

- para o físico A: precisão( $\theta$ ) passou de  $\tau_0^{-2} = 0.0025$  a  $\tau_1^{-2} = 0.004225008$  (aumento de 70%).
- para o físico B: precisão( $\theta$ ) passou de  $\tau_0^{-2} = 0.00015625$  a  $\tau_1^{-2} = 0.003926$  (aumento de 2500%)

Com 1 observação:

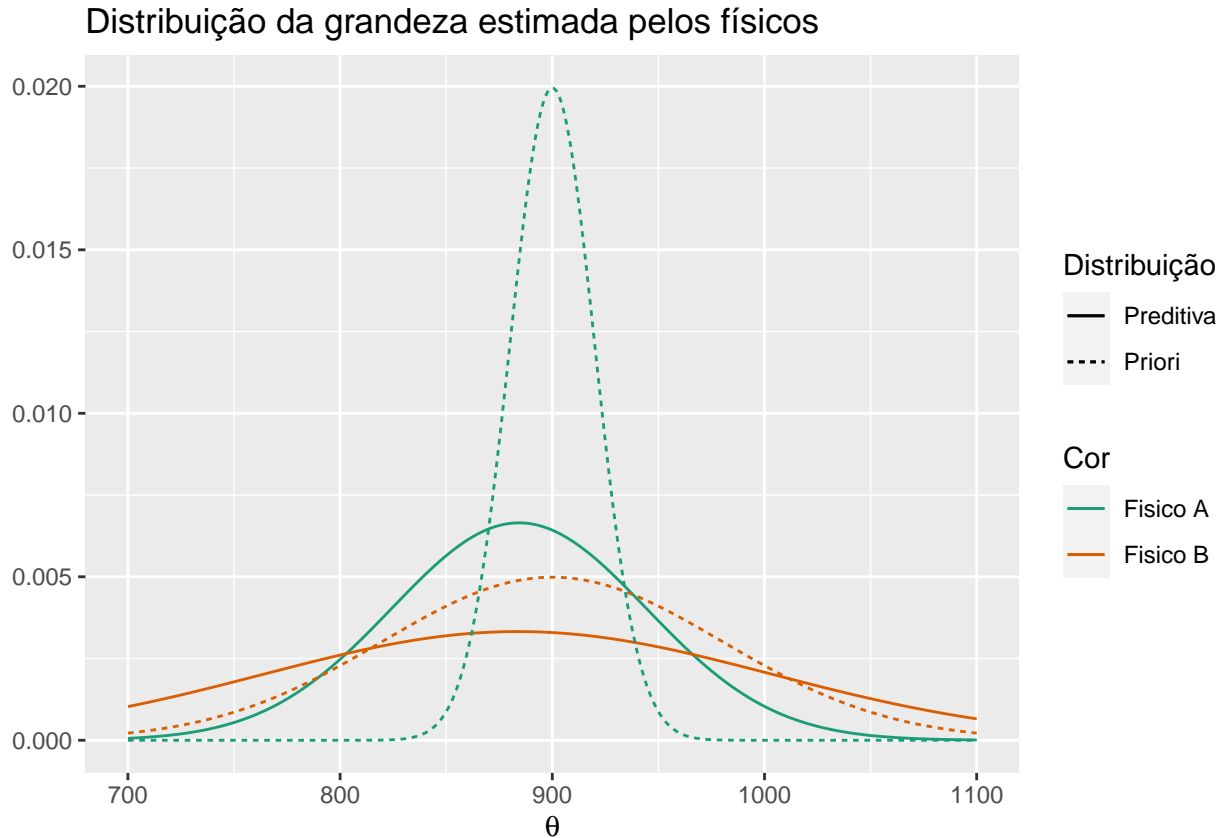
- para o físico A: precisão( $\theta$ ) passou de  $\tau_0^{-2} = 0.0025$  a  $\tau_1^{-2} = 6.1035 * 10^{-7}$ .
- para o físico B: precisão( $\theta$ ) passou de  $\tau_0^{-2} = 0.00015625$  a  $\tau_1^{-2} = 9.7656 * 10^{-6}$ .
- e) Construa uma tabela que contenha o resumo a priori e o resumo a posteriori.

Físico	Media.pri	Media.pos.100	Media.pos.1	SD.pri	SD.pos.100	SD.pos.1
A	900	884.3140	898.4	20	15.3846	320
B	900	883.7272	893.6	80	15.9601	1280

- f) Encontre a distribuição preditiva e faça um esboço de seu gráfico.

A distribuição preditiva é dada por:

$$X \sim N(\mu_0, \tau_0^2 + \sigma^2)$$



## Caso 2: $\mu$ conhecido e $\sigma$ desconhecido

### Distribuições a priori

Seja  $Y_i$  uma amostra aleatória simples de uma distribuição  $Y \sim N(\theta, \sigma^2)$ , com  $\theta$  conhecido.

Primeiramente, vamos encontrar a função de verossimilhança de  $\sigma^2$ .

$$\mathcal{L}(y|\sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} e^{-(y_i - \theta)^2 / 2\sigma^2} \propto (\sigma^2)^{-\frac{n}{2}} e^{-\left(\frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \theta)^2\right)}$$

### Priori não informativa

$$\log(\mathcal{L}(y|\sigma^2)) \propto -\frac{n}{2} \log(\sigma^{-2}) - \sigma^{-2} \sum_{i=1}^n (y_i - \theta)^2$$

A distribuição a *priori* de Jeffreys é dada por  $\pi(\sigma^2) \propto \sqrt{J(\sigma^2)}$ .

$$\begin{aligned} J(\sigma^2) &\propto E\left(-\frac{\partial^2}{\partial \theta^2} \log(L(\theta))\right) = E\left(-\frac{\partial^2}{\partial \theta^2} \left(-\frac{n}{2} \log(\sigma^{-2}) - (\sigma^2)^{-1} \sum_{i=1}^n (y_i - \theta)^2\right)\right) \\ &= E\left(-\frac{\partial}{\partial \theta} \left(-\frac{n}{2} (\sigma^2)^{-1} + (\sigma^2)^{-2} \sum_{i=1}^n (y_i - \theta)^2\right)\right) = E\left(-\frac{n}{2\sigma^2} + 2(\sigma^2)^{-3} \sum_{i=1}^n (y_i - \theta)^2\right) \\ &= -\frac{n}{2\sigma^2} + 2\sigma^{-4} \sum_{i=1}^n (E(y_i) - \theta)^2 = -\frac{n}{\sigma^2} + 2\sigma^{-4} \sum_{i=1}^n (\theta - \theta)^2 = -\frac{n}{\sigma^2} \propto \sigma^{-2} \end{aligned}$$

Assim,  $\pi(\sigma) \propto \sqrt{\sigma^{-2}} = \sigma^{-1}$ . Seu parâmetro  $\Phi$  de escala que faz com que  $\theta$  mude somente em locação pode ser obtido através do cálculo de

$$\phi \propto \int \pi(\sigma^2) d\sigma^2 = \int \frac{1}{\sigma^2} d\sigma^2 = \log|\sigma^2| + k \propto \log \sigma^2$$

$\phi$  é uma distribuição imprópria, pois  $\int_0^{+\infty} \log(\sigma^2) d\sigma^2$  é divergente. Assim, a *priori*

### Conjugadas Naturais

O suporte de nosso parâmetro de interesse  $\sigma > 0$  permite-nos adotar três distribuições de probabilidade estudadas durante o curso:

#### 1. Gama:

Se  $X \sim \text{Gama}(\alpha, \beta)$  então

$$f_X(x|\alpha, \beta) = \frac{\beta^\alpha x^{(\alpha-1)} e^{-\beta x}}{\Gamma(\alpha)} \propto x^{(\alpha-1)} e^{-\beta x}, \quad \alpha > 0, \beta > 0, x > 0$$

#### 2. Gama-Inversa:

Se  $X \sim \text{Gama-Inv}(\alpha, \beta)$  então

$$f_X(x|\alpha, \beta) = \frac{\beta^\alpha x^{-(\alpha+1)} e^{-\beta/x}}{\Gamma(\alpha)} \propto x^{-(\alpha+1)} e^{-\beta/x}, \quad \alpha > 0, \beta > 0, x > 0$$

### 3. Qui-Quadrado:

Se  $X \sim \chi^2(\nu)$  então

$$f_X(x|\nu) = \frac{x^{(\nu/2)-1} e^{-x/2}}{2^{\nu/2} \Gamma(\nu/2)} \propto x^{(\nu/2)-1} e^{-x/2}, \quad \nu > 0, x > 0$$

Note que as duas primeiras estão relacionadas via uma transformação simples e a última é um caso particular delas. Desprezadas as constantes não informativas, as três distribuições são da forma  $x$  elevado a uma potência vezes a exponencial de  $x$ . Dessa forma, as três distribuições servem como conjugada natural da Normal, em nosso caso. Neste trabalho, optou-se por utilizar a distribuição Gama Inversa.

Fazendo  $x = \sigma$ , temos uma *priori* da forma:

$$\sigma^{-(\alpha+1)} e^{-\beta/\sigma^2} \Rightarrow \pi(\sigma) \sim \text{Gama-Inv}(\alpha, \beta)$$

Se quisermos torná-la não informativa, basta utilizarmos  $\alpha \rightarrow 0, \beta \rightarrow 0$ .

### Distribuição a *posteriori*

$$\begin{aligned} \pi(\sigma|y) &\propto \mathcal{L}(y|\sigma^2) \pi(\sigma) = (\sigma^2)^{-\frac{n}{2}} e^{-\left(\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \theta)^2\right)} \sigma^{-(\alpha-1)} e^{-\beta/\sigma^2} \\ &= (\sigma^2)^{-(\alpha + \frac{n}{2} + 1)} e^{-\frac{1}{\sigma^2} \left(\beta + \frac{1}{2} \sum_{i=1}^n (y_i - \theta)^2\right)} \end{aligned}$$

Dessa forma,

$$\pi(\sigma^2|y) \sim \text{Gama-Inv}\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2} \sum_{i=1}^n (y_i - \theta)^2\right)$$