

# SME0809 - Inferência Bayesiana - Distribuição não informativa

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

Setembro 2021

```
# pacotes do R utilizados
library(tidyverse)
library(ggpubr)
library(effectsize) #padronizacao
set.seed(42)
```

Seja  $Y_1, \dots, Y_n$  uma a.a de  $Y \sim \text{Pois}(\theta)$ . Pede-se:

- a) encontre a distribuição a *priori* não informativa de Jeffreys

Temos que

$$p(y|\theta) = \frac{e^{-\theta} \theta^y}{y!}, \quad \theta > 0, \quad y = 0, 1, 2, \dots$$

Primeiramente, vamos obter a log-verossimilhança

$$\begin{aligned} \log(L(\theta)) &= \log \left( \prod_{i=1}^n p(y_i|\theta) \right) = \log \left( \prod_{i=1}^n \frac{e^{-\theta} \theta^{y_i}}{y_i!} \right) = \log \left( \frac{e^{-n\theta} \theta^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!} \right) = \\ &= -n\theta + \sum_{i=1}^n y_i \log(\theta) - \log \left( \prod_{i=1}^n y_i! \right) \end{aligned}$$

Tomando a segunda derivada, temos que:

$$\frac{\partial^2}{\partial \theta^2} \log(L(\theta)) = \frac{\partial^2}{\partial \theta^2} \left[ -n\theta + \sum_{i=1}^n y_i \log(\theta) - \log \left( \prod_{i=1}^n y_i! \right) \right] = \frac{\partial}{\partial \theta} \left[ -n + \frac{1}{\theta} \sum_{i=1}^n y_i \right] = -\frac{1}{\theta^2} \sum_{i=1}^n y_i$$

Assim, como  $J(\theta) = E \left( -\frac{\partial^2}{\partial \theta^2} \log(L(\theta)) \right)$  então

$$J(\theta) = E \left( \frac{1}{\theta^2} \sum_{i=1}^n y_i \right) = \frac{1}{\theta^2} E \left( \sum_{i=1}^n y_i \right) = \frac{n\theta}{\theta^2} = \frac{n}{\theta} \propto \frac{1}{\theta}$$

A distribuição a *priori* de Jeffreys é dada por  $\pi(\theta) \propto \sqrt{J(\theta)}$ . Logo,  $\pi(\theta) \propto \theta^{-1/2}$ .

Note que esta *priori* pode ser obtida a partir da conjugada natural Gama( $\alpha, \beta$ ), com  $\alpha = 1/2$  e  $\beta \rightarrow 0$ . Ilustramos o efeito de fixar  $\alpha$  e diminuir  $\beta$  abaixo:

```

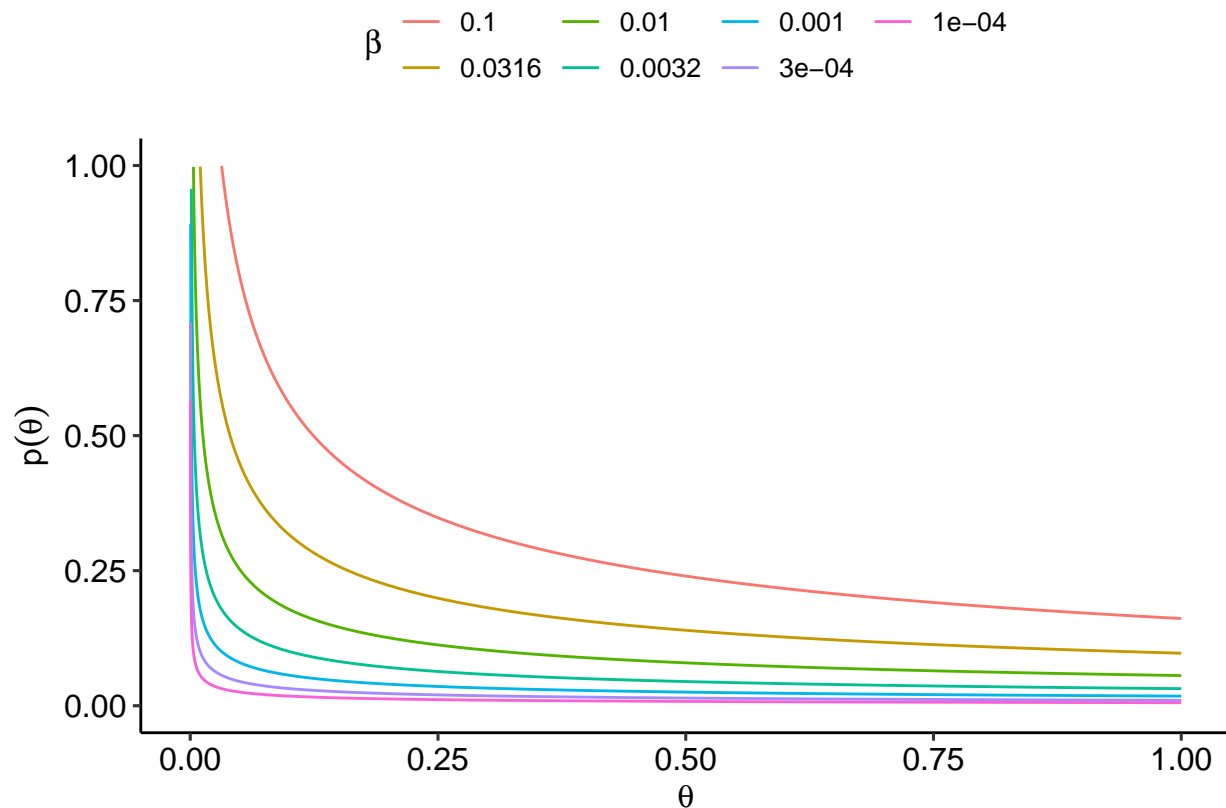
theta <- seq(0.0001, 1, 0.0001)

b <- seq(1, 4, 0.5)

b %>% map_dfr( ~tibble( x = theta,
                        y = dgamma(theta, 1/2, 10^(-.x)),
                        beta = as.factor( round( 10^(-.x), 4))) ) %>%

ggplot() + geom_line(aes(x = x, y = y, color = beta)) +
scale_y_continuous(limits = c(0,1)) + labs(color= expression(beta)) +
xlab(expression(theta)) + ylab(expression(p(theta))) +
theme_pubr()

```



Além disso,  $\pi(\theta)$  é uma distribuição imprópria pois  $\int_0^{+\infty} \theta^{-1/2} d\theta$  diverge.

- b) A função de verossimilhança na parametrização  $\theta$  muda em localização e escala? Justificar graficamente

Já obtivemos a função de verossimilhança em a):

$$\log(L(\theta)) = -n\theta + \sum_{i=1}^n y_i \log(\theta) - \log\left(\prod_{i=1}^n y_i!\right) \propto \log(\theta) \sum_{i=1}^n y_i - n\theta$$

Vamos agora realizar uma implementação e calcular a verossimilhança para diversas amostras aleatórias de tamanho 20 geradas de uma Poisson.

```

Ltheta <- function(theta, y) {
  exp( sum(y)*log(theta) - length(y)*theta)
}

theta <- seq(0,20, 0.01)

df <- tibble(
  L = theta %>% map_dbl(~Ltheta(.x, y = rpois(10,1))),
  theta = theta)

df <- NULL

for (lambda in c(1,4,9,16)) {
  df <- rbind(df, tibble(
    L = normalize(map_dbl(theta, Ltheta, y = rpois(20,lambda))),
    theta = theta,
    lambda = as_factor(lambda))
}

```

O gráfico com a verossimilhança normalizada fica:

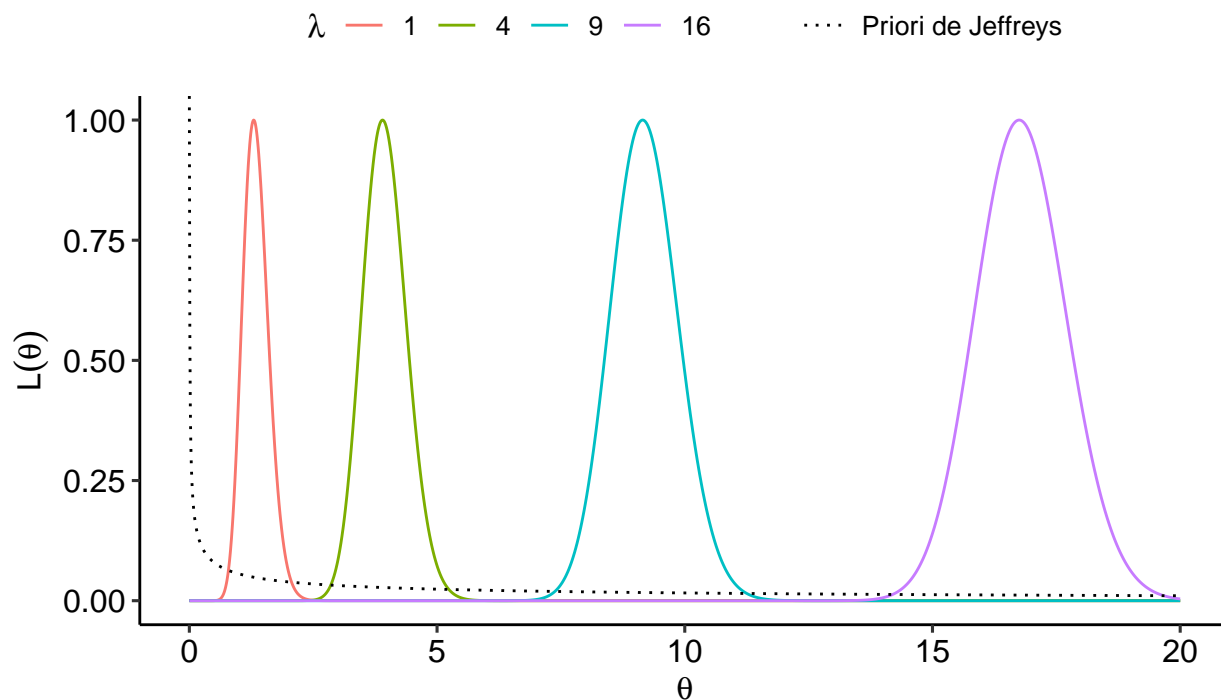
```

df %>% ggplot(aes(x = theta)) + geom_line(aes(y = L, color = lambda)) +
  geom_line(aes(y = dgamma(theta, 1/2, 0.01), linetype = "Priori de Jeffreys")) +
  labs(color= expression(lambda),
       title = expression("Verossimilhança normalizada de uma Poisson"~(lambda)),
       subtitle = "Amostra aléatória com n=20",
       x = expression(theta),
       y = expression(L(theta))) +
  scale_linetype_manual(name = " ", values = "dotted") +
  theme_pubr()

```

## Verossimilhança normalizada de uma Poisson ( $\lambda$ )

Amostra aléatória com  $n=20$



Dessa forma, vemos que a função de verossimilhança muda tanto em locação como escala, pois valores grandes de  $\lambda$  alocam a distribuição para direita, além de deixá-la mais achatada.

- c) caso a resposta ao item b) seja afirmativa, encontre a escala na qual a função de verossimilhança mude somente em locação. Mostre graficamente.

$$\phi \propto \int \pi(\theta) d\theta = \int \theta^{-1/2} d\theta = 2\sqrt{\theta} \Rightarrow \theta \propto \phi^2$$

```
df %>% ggplot(aes(x = sqrt(theta))) +
  geom_line(aes(y = sqrt(L), color = lambda)) +
  geom_line(aes(y = dgamma(theta, 1/2, 0.001), linetype = "Priori de Jeffreys")) +
  labs(color= expression(lambda),
       title = expression("Verossimilhança normalizada de uma Poisson"~(lambda)),
       subtitle = "Amostra aléatória com n=20",
       x = expression(Phi),
       y = expression(L(Phi))) +
  scale_linetype_manual(name = " ", values = "dotted") +
  theme_pubr()
```

## Verossimilhança normalizada de uma Poisson ( $\lambda$ )

Amostra aléatória com  $n=20$

