

[illegible]

# Inferência Bayesiana

Conceitos importantes para caminhar pelo **MCMC**:

- Qual a função de verossimilhança?
- Quais são os parâmetros do modelo?
- Conceito de superfície de verossimilhança.
- Conceito de busca heurística.

# Inferência Bayesiana

Conceitos importantes para caminhar pelo **MCMC**:

- Qual a função de verossimilhança?
- Quais são os parâmetros do modelo?
- Conceito de superfície de verossimilhança.
- Conceito de busca heurística.

Conceitos importantes da estatística **Bayesiana**:

- Teorema de Bayes.
- Distribuição a priori. (Prior)
- Distribuição a posteriori. (Posterior)

# Inferência Bayesiana

Teorema de Bayes

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

# Inferência Bayesiana

Teorema de Bayes

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(\theta|y)$  = distribuição posterior.

$P(y|\theta)$  = verossimilhança.

# Inferência Bayesiana

Teorema de Bayes

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(\theta|y)$  = probabilidade do modelo condicionada nas observações.

$P(y|\theta)$  = probabilidade das observações condicionada no modelo.

# Inferência Bayesiana

Teorema de Bayes

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

O objetivo da inferência usando a estatística Bayesiana é estimar a *distribuição posterior* dos parâmetros condicionada nas observações.

Diferente da máxima verossimilhança, estamos calculando a probabilidade dos parâmetros e não das observações.

# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$



# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(\theta|y)$  = probabilidade do modelo condicionada nas observações.

$P(\theta)$  = probabilidade marginal do modelo (prior).

$P(y|\theta)$  = probabilidade das observações condicionada no modelo.

$P(y)$  = probabilidade marginal das observações.

# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(\theta|y)$  = distribuição posterior (posterior).

$P(\theta)$  = distribuição a priori (prior).

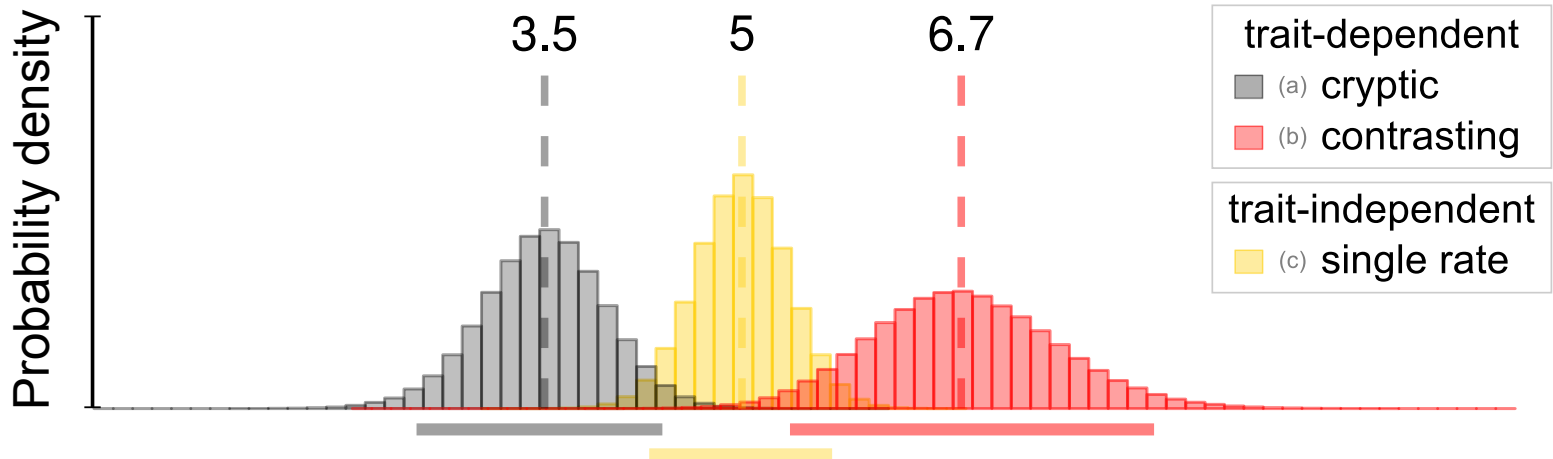
$P(y|\theta)$  = verossimilhança (likelihood).

$$P(y) = \sum_{\theta} P(\theta)P(y|\theta)$$

# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(\theta|y)$  = o principal objetivo de uma análise usando **MCMC**.

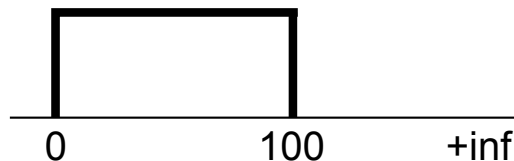
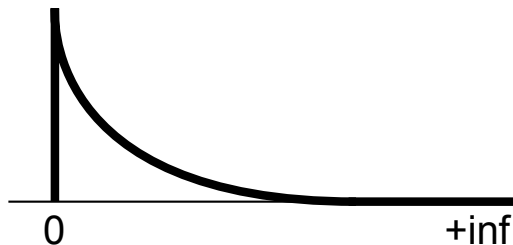


# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(\theta)$  = nosso conhecimento a priori dos valores possíveis dos parâmetros.

O prior pode ser qualquer distribuição que respeite as restrições dos parâmetros. Geralmente são separados em priors informativos ou 'não informativos'.



# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(\theta)$  = nosso conhecimento a priori dos valores possíveis dos parâmetros.

---

O prior é o elemento que mais gera discussão em análises Bayesianas.

Vamos explorar em mais detalhe a influência do prior nos nossos resultados.

# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(y|\theta)$  = função de verossimilhança.

---

Como vimos anteriormente ...

A função de verossimilhança calcula a verossimilhança de nossas observações condicionado aos valores dos parâmetros dos modelos.

# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(y)$  = probabilidade dos dados condicionada a TODOS os valores possíveis dos parâmetros.

---

$$P(y) = \sum_{\theta} P(\theta)P(y|\theta) \quad \text{para distribuições discretas}$$

ou

$$P(y) = \int P(\theta)P(y|\theta)d\theta \quad \text{para distribuições contínuas}$$

# Inferência Bayesiana

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$P(y)$  = probabilidade dos dados condicionada a TODOS os valores possíveis dos parâmetros.

$$P(y) = \sum_{\theta} P(\theta)P(y|\theta)$$

ou

$$P(y) = \int P(\theta)P(y|\theta)d\theta$$





# Inferência Bayesiana

$P(y)$  é somente possível de ser derivada para os modelos mais simples.

Essa quantidade é o principal motivo da inferência Bayesiana ter sido aplicada para uma ampla variedade de problemas somente após o desenvolvimento da computação.

Não conseguimos, na prática, calcular  $P(y)$ .

Usamos '*odds ratios*' ou 'razão de probabilidade'.

# Inferência Bayesiana

Vamos voltar a fórmula de Bayes ...

$$P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

$$P(y) = \int P(\theta)P(y|\theta)d\theta$$

Veja que  $P(y)$  não depende de um valor específico para os parâmetros do modelo, pois é uma *distribuição marginal de probabilidade* das observações.

Ou seja,  $P(y)$  não está condicionada a algum valor de parâmetro.

# Inferência Bayesiana

Comparando o valor de  $P(\theta|y)$  para  $\theta_1, \theta_2$ , temos:

$$P(\theta_1|y) = \frac{P(\theta_1)P(y|\theta_1)}{P(y)}$$

$$P(\theta_2|y) = \frac{P(\theta_2)P(y|\theta_2)}{P(y)}$$

Então a '*odds ratio*'  $\frac{P(\theta_1|y)}{P(\theta_2|y)}$  ...

# Inferência Bayesiana

A *odds ratio* para  $\theta_1, \theta_2$  :

$$\begin{aligned}\frac{P(\theta_1|y)}{P(\theta_2|y)} &= \frac{P(\theta_1)P(y|\theta_1)}{P(\theta_2)P(y|\theta_2)} \frac{P(y)}{P(y)} \\ &= \frac{P(\theta_1)P(y|\theta_1)}{P(\theta_2)P(y|\theta_2)}\end{aligned}$$

Podemos usar essa razão para comparar o fit de  $\theta_1$  em relação a  $\theta_2$  .

Mas somente para diferentes valores de parâmetros do MESMO MODELO.

# Introdução ao MCMC

A *odds ratio* para  $\theta_1, \theta_2$ :

$$\frac{P(\theta_1|y)}{P(\theta_2|y)} = \frac{P(\theta_1)P(y|\theta_1)}{P(\theta_2)P(y|\theta_2)}$$

Essa razão é uma peça fundamental do **MCMC**, pois permite a comparação relativa do fit entre dois pontos na superfície de verossimilhança.

O primeiro termo se chama prior ratio.

O segundo termo se chama likelihood ratio.

# Introdução ao MCMC

A *odds ratio* para  $\theta_1, \theta_2$ :

$$\frac{P(\theta_1|y)}{P(\theta_2|y)} = \frac{P(\theta_1)P(y|\theta_1)}{P(\theta_2)P(y|\theta_2)}$$

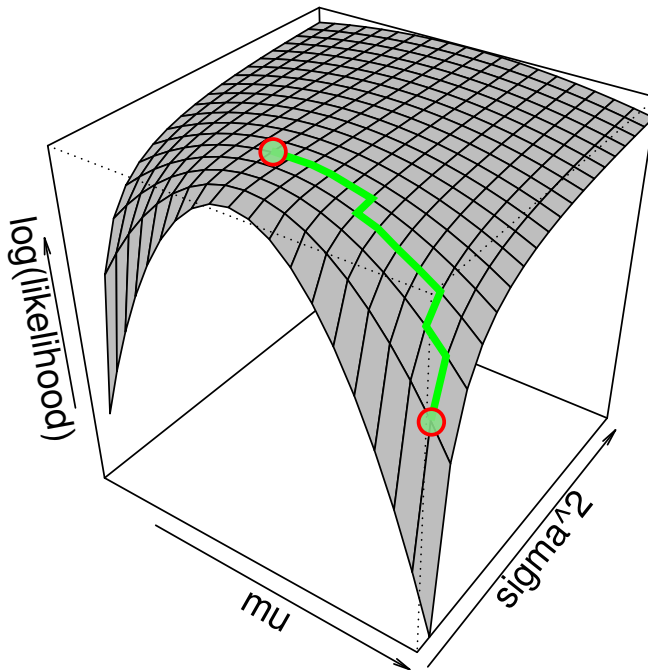
Essa razão é uma peça fundamental do **MCMC**, pois permite a comparação relativa do fit entre dois pontos na superfície de verossimilhança.

O primeiro termo se chama **prior ratio**.

O segundo termo se chama **likelihood ratio**.

Lembra-se da busca de ML?

# Maximizar uma função



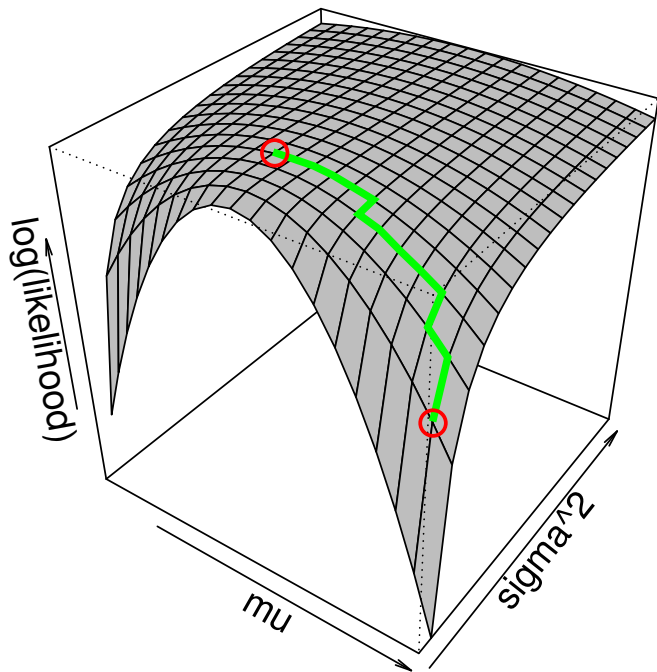
O método de busca mais simples tenta maximizar (ou minimizar) o valor de uma função.

Temos um valor inicial.

Caminhamos pela superfície de verossimilhança até encontrarmos uma combinação de parâmetros que maximiza a verossimilhança.

# Introdução ao MCMC

A '*odds ratio*' é a peça fundamental da máquina que caminha pela superfície de verossimilhança durante o **MCMC**.



$$\frac{P(\theta_1 | y)}{P(\theta_2 | y)} = \frac{P(\theta_1) P(y | \theta_1)}{P(\theta_2) P(y | \theta_2)}$$

Vamos usar o resultado desta razão para 'decidir' se estamos subindo na verossimilhança do modelo ou não.



# Introdução ao MCMC

*Odds ratio* para comparar o fit de  $\theta_1$  em relação a  $\theta_2$  .

$$\frac{P(\theta_1|y)}{P(\theta_2|y)} = 1; \theta_1 = \theta_2 \text{ ou equivalente}$$

$$\frac{P(\theta_1|y)}{P(\theta_2|y)} > 1; \theta_1 \text{ melhor fit que } \theta_2$$

$$\frac{P(\theta_1|y)}{P(\theta_2|y)} < 1; \theta_1 \text{ pior fit que } \theta_2$$

# Introdução ao MCMC

No próximo passo vamos entender como a fórmula de Bayes e a *odds ratio* é utilizada para amostrar a distribuição posterior dos parâmetros de um modelo.

Vou apresentar algoritmos de Markov chain Monte Carlo e vamos navegar pelos seus componentes.

