# Estatísticas & Aquicultura.

Com: Carlos Antônio Zarzar

carlos.zarzar@ufopa.edu.br

# Estatísticas & Aquicultura.

Com: Carlos Antônio Zarzar

carlos.zarzar@ufopa.edu.br

Apresenta: Métodos computacionais



- Método de Monte Carlo;
- Método de Monte Carlo via Cadeia de Markov;
- Algoritmo de amostragem;
  - Gibbs;
  - Metropolis-Hasting;

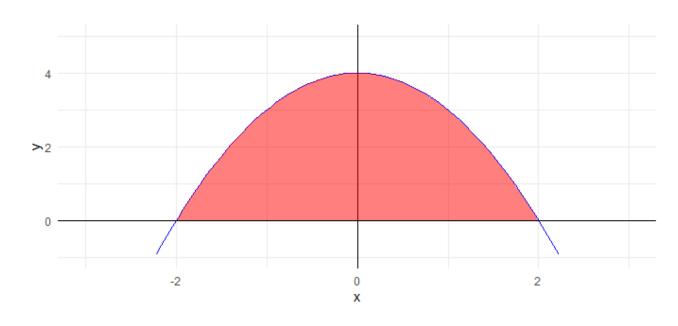
• Integração de Monte Carlo;

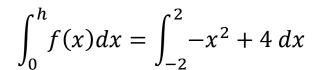
ou

• Método de Monte Carlo.

ЭU

• Método de Monte Carlo.





Integracao de Monte Carlo

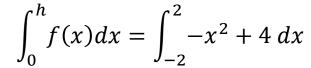
Monte Carlo é um método de integração numérico conhecido:

• Integração de Monte Carlo;

ou

Método de Monte Carlo.

Geração de variáveis aleatórias independentes com mesma distribuição



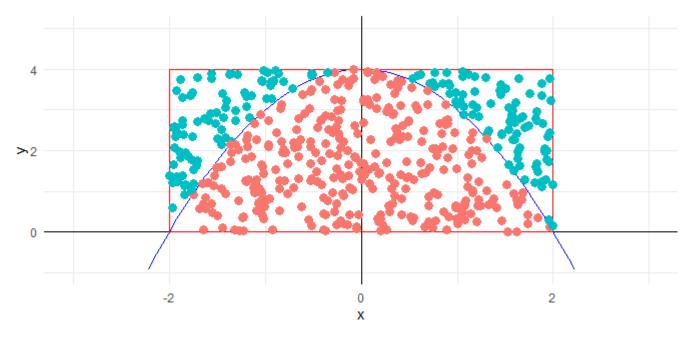
- Número de pontos <u>fora</u> da curva = 172;
- Número de pontos <u>abaixo</u> da curva = 328;
- Número de pontos <u>Totais</u> = 500;

$$\int_0^h f(x)dx = \int_{-2}^2 -x^2 + 4 \, dx$$

• Integração de Monte Carlo;

ou

Método de Monte Carlo.



Geração de variáveis aleatórias independentes com mesma distribuição

- Número de pontos <u>fora</u> da curva = 172;
- Número de pontos <u>abaixo</u> da curva = 328;
- Número de pontos **Totais** = 500;

$$Area f(x) = \frac{N umero}{N umero} 4$$

➤ Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$

Monte Carlo é um método de integração numérico conhecido:

• Integração de Monte Carlo;

OU

• Método de Monte Carlo.



- Número de pontos <u>fora</u> da curva = 172;
- Número de pontos <u>abaixo</u> da curva = 328;
- Número de pontos **Totais** = 500;

➤ Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

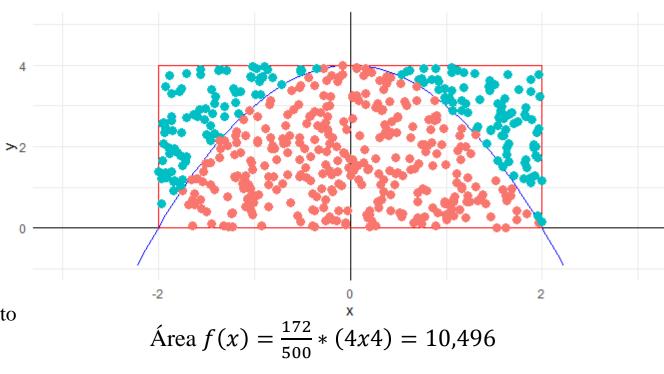
$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$

Monte Carlo é um método de integração numérico conhecido:

• Integração de Monte Carlo;

ou

• Método de Monte Carlo.



$$\int_0^h f(x)dx = \int_{-2}^2 -x^2 + 4 \, dx = 10,667$$

• Integração de Monte Carlo;

ou

Método de Monte Carlo.

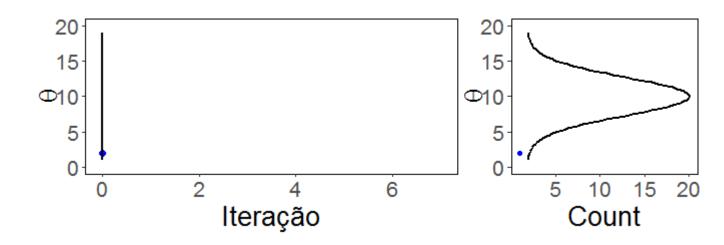
➤ Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$



• Integração de Monte Carlo;

ou

Método de Monte Carlo.

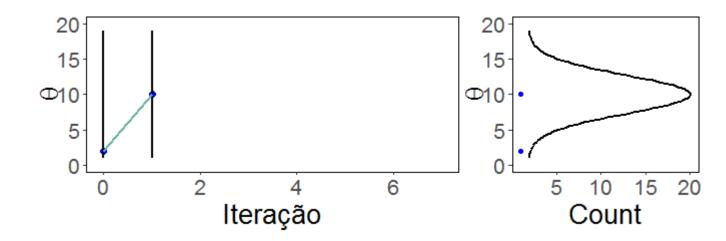
➤ Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$



• Integração de Monte Carlo;

ou

Método de Monte Carlo.

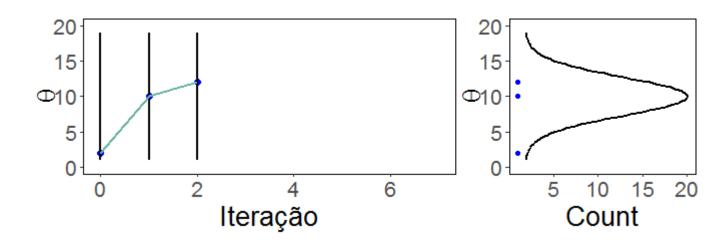
➤ Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$



• Integração de Monte Carlo;

ou

• Método de Monte Carlo.

➤ Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$



• Integração de Monte Carlo;

ou

Método de Monte Carlo.

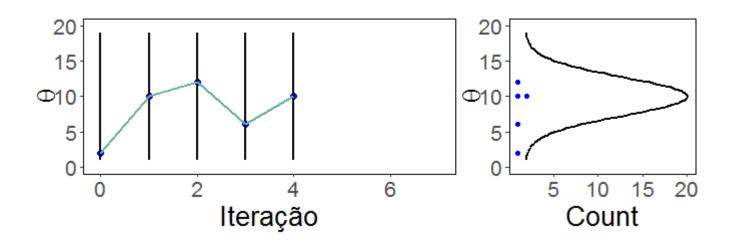
➤ Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n Y_i = \mu\right) = 1$$



• Integração de Monte Carlo;

ou

Método de Monte Carlo.

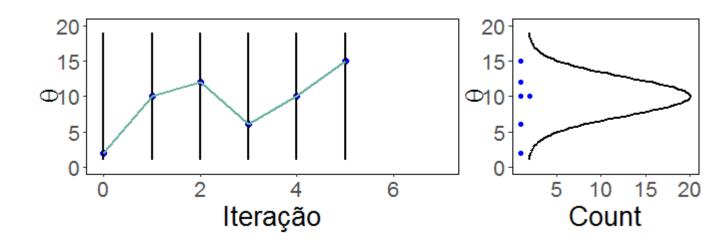
➤ Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$



• Integração de Monte Carlo;

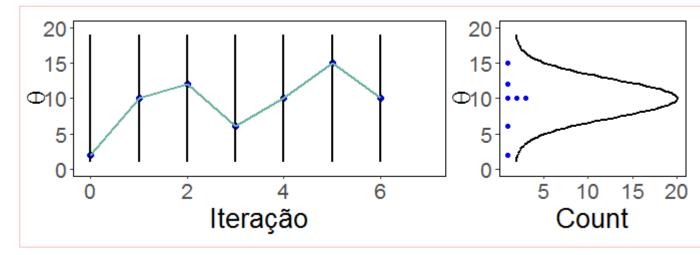
ou

Método de Monte Carlo.

Elei Forte dos Grandes Números Seja  $Y_1, Y_2, \dots, Y_n$  variáveis i.i.d. $\mu = E[Y] < \infty$ 

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$



• Integração de Monte Carlo;

ou

Método de Monte Carlo.

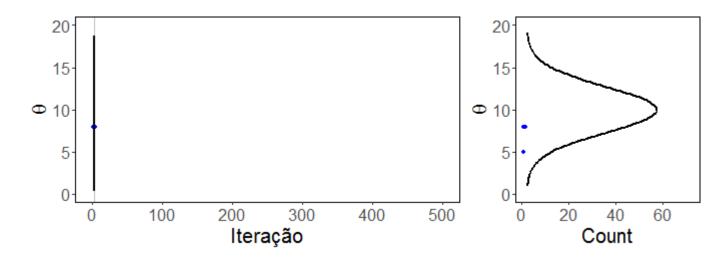
> Lei Forte dos Grandes Números

Seja  $Y_1, Y_2, \dots, Y_n$  variáveis *i.i.d.* 

$$\mu = E[Y] < \infty$$

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to \mu \quad \text{Convergindo quase certo} \\ n \to \infty$$

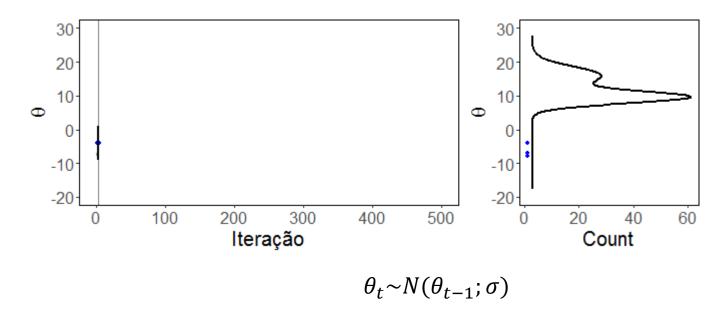
$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_{i}=\mu\right)=1$$



#### Monte Carlo Via Cadeia de Markov:

A posteriori: Bimodal Distribuição Normal mista

$$f(x) = 0.4 * N(17; 3) + 0.6 * N(3; 2)$$

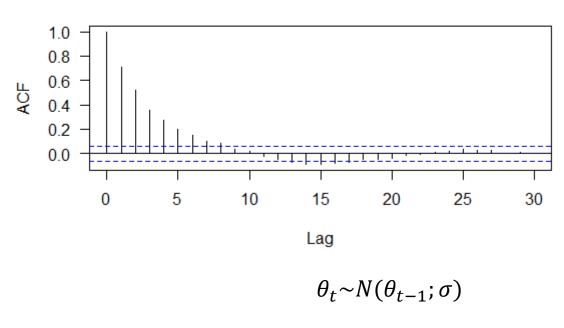


# Monte Carlo Via Cadeia de Markov:

# *A posteriori*: Bimodal Distribuição Normal mista

$$f(x) = 0.4 * N(17; 3) + 0.6 * N(3; 2)$$

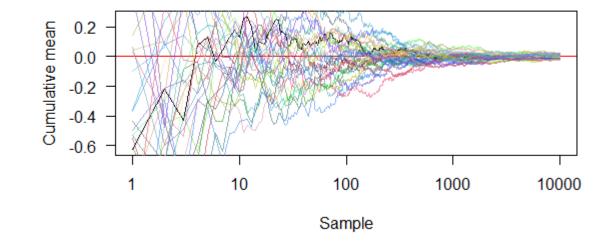
#### Intermediate



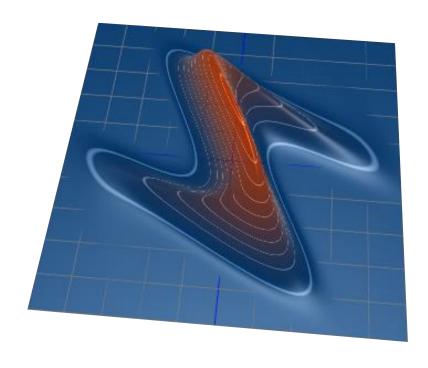
#### Monte Carlo Via Cadeia de Markov:

A posteriori: Bimodal Distribuição Normal mista

$$f(x) = 0.4 * N(17; 3) + 0.6 * N(3; 2)$$



Espaço paramétrico da distribuição *a posteriori* 

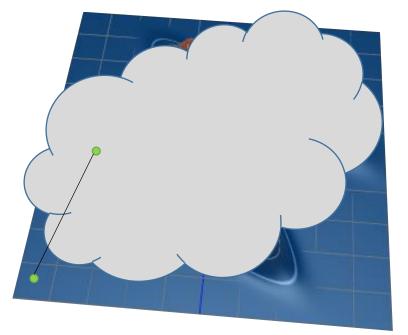


Desconhecido

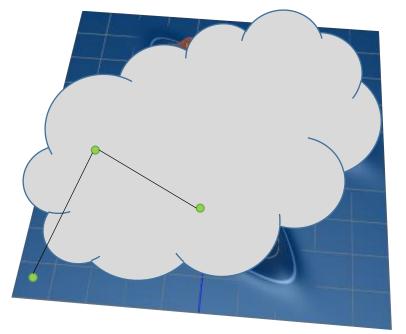




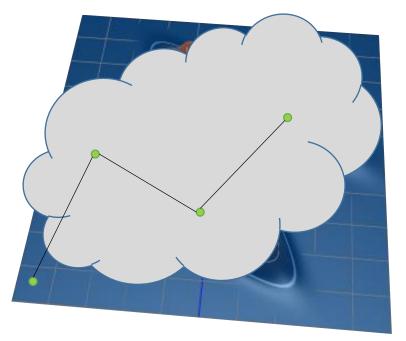
• Random walk MCMC



• Random walk MCMC



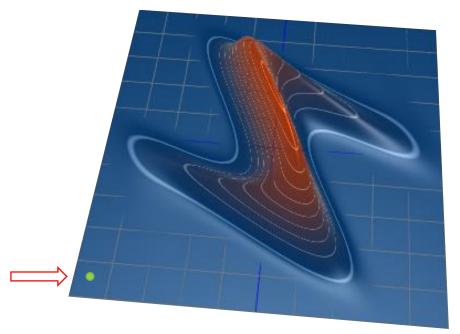
• Random walk MCMC



• Random walk MCMC

Custo computacional muito grande para amostrar todo o espaço paramétrico.

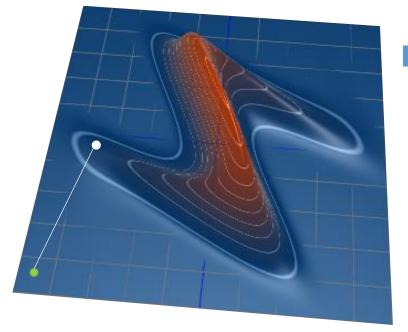




• Random walk MCMC

- <u>Método a amostragem:</u>
  - Gibbs
  - Metropolis-Hasting

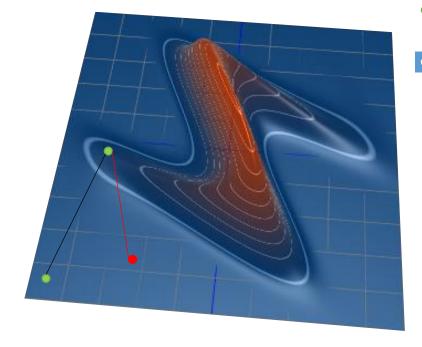
Etapa de aceitação ou rejeição de cada iteração do passeio aleatório.



- Random walk MCMC
- Amostrador em avaliação;

- Método a amostragem:
  - Gibbs
  - Metropolis-Hasting

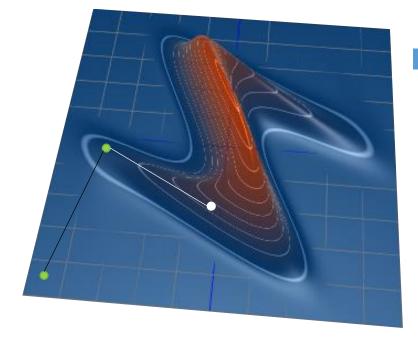
Etapa de aceitação ou rejeição de cada iteração do passeio aleatório.



- Random walk MCMC
- Amostrador em avaliação;
- Amostrador rejeitado;

- <u>Método a amostragem:</u>
  - Gibbs
  - Metropolis-Hasting

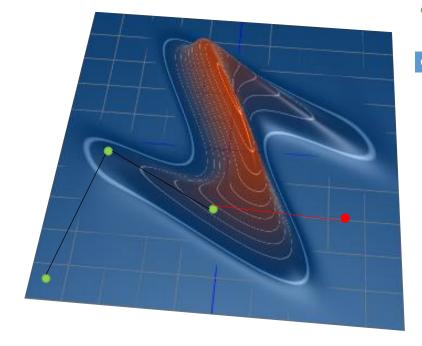
Etapa de aceitação ou rejeição de cada iteração do passeio aleatório.



- Random walk MCMC
- Amostrador em avaliação;
- Amostrador rejeitado;

- <u>Método a amostragem:</u>
  - Gibbs
  - Metropolis-Hasting

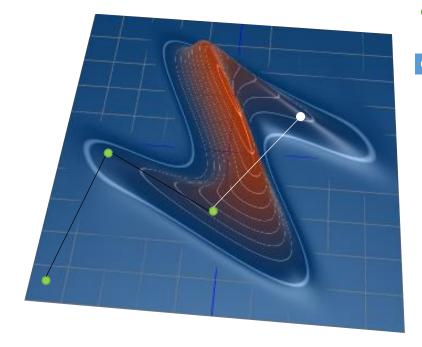
Etapa de aceitação ou rejeição de cada iteração do passeio aleatório.



- Random walk MCMC
- Amostrador em avaliação;
- Amostrador rejeitado;

- Método a amostragem:
  - Gibbs
  - Metropolis-Hasting

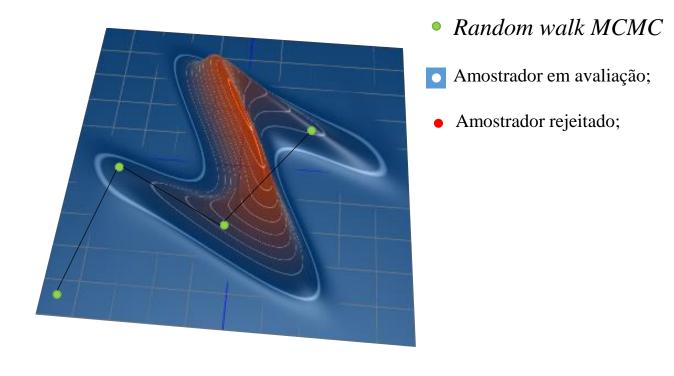
Etapa de aceitação ou rejeição de cada iteração do passeio aleatório.



- Random walk MCMC
- Amostrador em avaliação;
- Amostrador rejeitado;

- Método a amostragem:
  - Gibbs
  - Metropolis-Hasting

Etapa de aceitação ou rejeição de cada iteração do passeio aleatório.



- <u>Método a amostragem:</u>
  - Gibbs

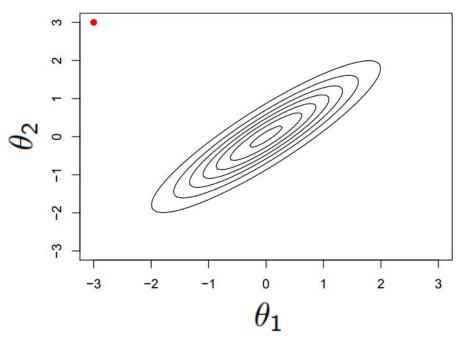
atenção

• Metropolis-Hasting

- Amostragem *a posteriori* é marginal por parâmetro;
- Lenta convergência;

(Principalmente espaços paramétricos de alta dimensão).

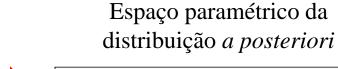
Espaço paramétrico da distribuição *a posteriori* 

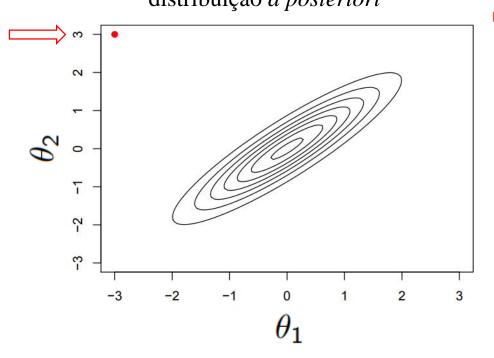


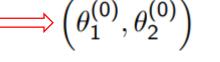
- Método a amostragem:
  - Gibbs atenção
  - Metropolis-Hasting

• Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} p(\theta_1|\theta_2, y)$$
Distribuição Alvo







$$heta_1^{(t)} \sim p\left(\theta_1 | \theta_2^{(t-1)}, y\right)$$
 $heta_2^{(t)} \sim p\left(\theta_2 | \theta_1^{(t)}, y\right)$ 

$$\theta_2^{(t)} \sim p\left(\theta_2|\theta_1^{(t)}, y\right)$$

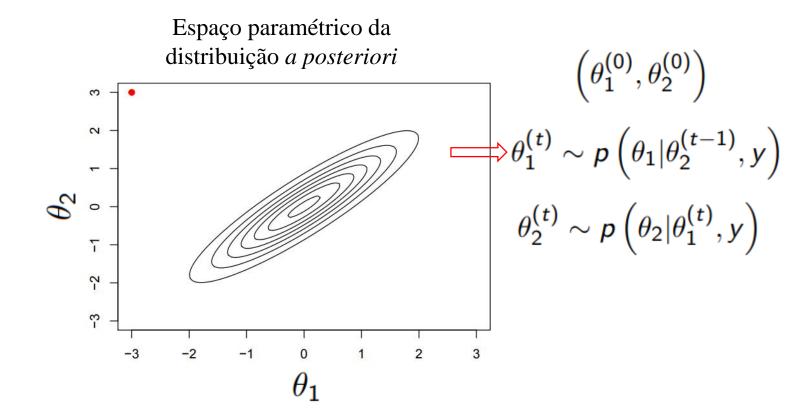
Método a amostragem:

Gibbs atenção

Metropolis-Hasting

Amostragem a posteriori é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{bmatrix} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{bmatrix}$$
Distribuição Alvo

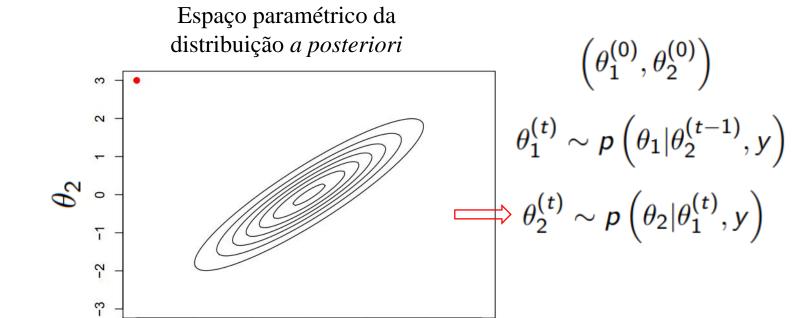


- <u>Método a amostragem:</u>
  - Gibbs atenção
  - Metropolis-Hasting

Amostragem *a posteriori* é marginal por parâmetro;

$$\begin{array}{ccc}
p(\theta|y) & \xrightarrow{\text{com}} & \theta = (\theta_1, \theta_2) & \xrightarrow{\text{amostrar}} & \begin{bmatrix}
p(\theta_1|\theta_2, y) \\
p(\theta_2|\theta_1, y)
\end{bmatrix}$$
Distribuição Alvo

-3



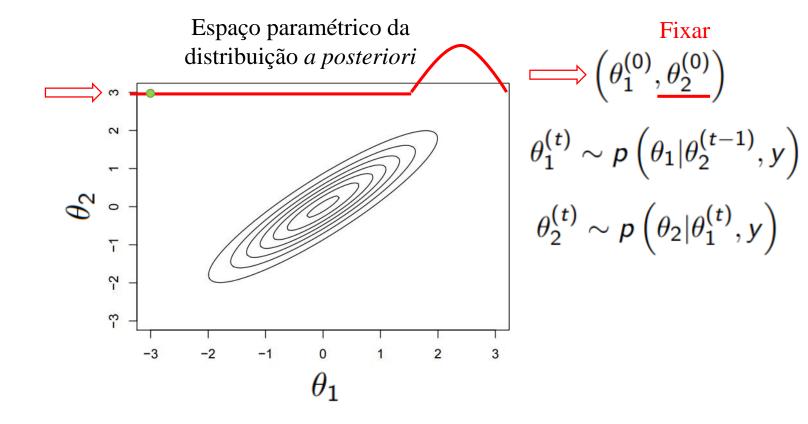
2

- <u>Método a amostragem:</u>
  - Gibbs atenção
  - Metropolis-Hasting

• Amostragem *a posteriori* é marginal por parâmetro;

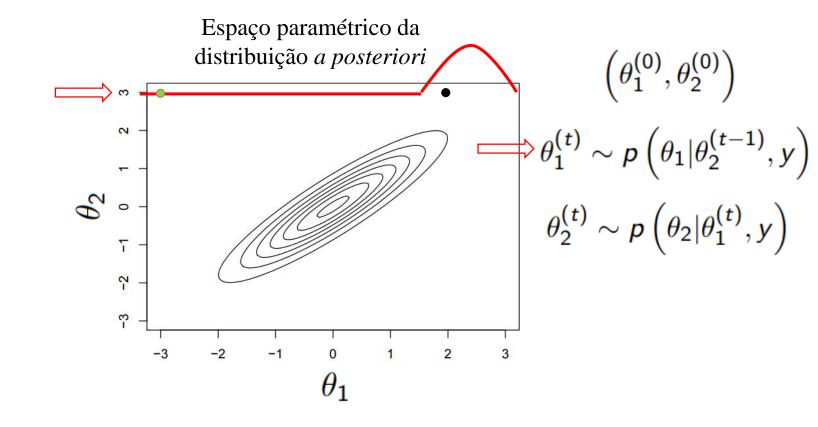
 $\theta_1$ 

$$\begin{array}{ccc}
p(\theta|y) & \xrightarrow{\text{com}} & \theta = (\theta_1, \theta_2) & \xrightarrow{\text{amostrar}} & \begin{bmatrix} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{bmatrix}
\end{array}$$
Distribuição Alvo



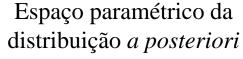
- <u>Método a amostragem:</u>
  - Gibbs atenção
  - Metropolis-Hasting

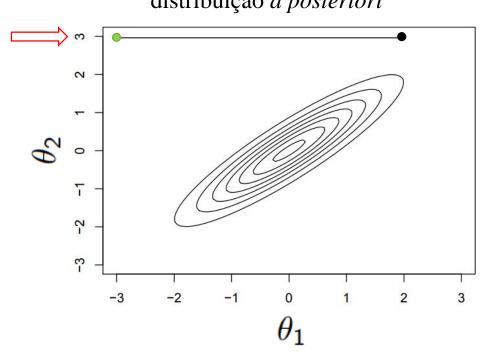
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} p(\theta_1)$$
Distribuição Alvo



- <u>Método a amostragem:</u>
  - Gibbs atenção
  - Metropolis-Hasting

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \int p(\theta_1|\theta_2|\theta_2)$$
Distribuição Alvo





$$\left(\theta_1^{(0)}, \theta_2^{(0)}\right)$$

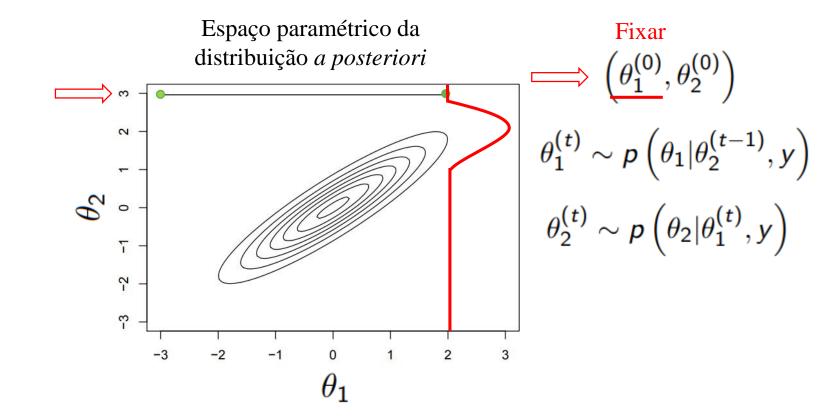
$$\theta_2^{(t)} \sim p\left(\theta_2|\theta_1^{(t)}, y\right)$$

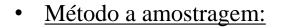
• <u>Método a amostragem:</u>

Gibbs atenção

• Metropolis-Hasting

$$\begin{array}{ccc}
p(\theta|y) & \xrightarrow{\text{com}} & \theta = (\theta_1, \theta_2) & \xrightarrow{\text{amostrar}} & \begin{bmatrix}
p(\theta_1|\theta_2, y) \\
p(\theta_2|\theta_1, y)
\end{bmatrix}$$
Distribuição Alvo

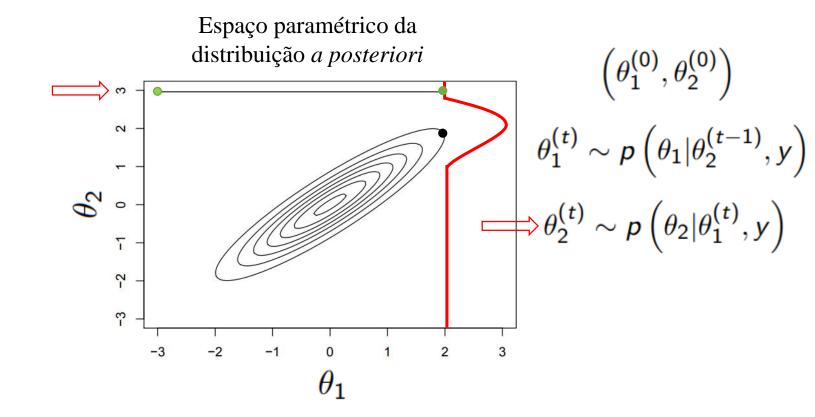




Gibbs atenção

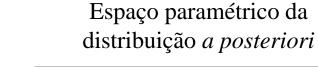
• Metropolis-Hasting

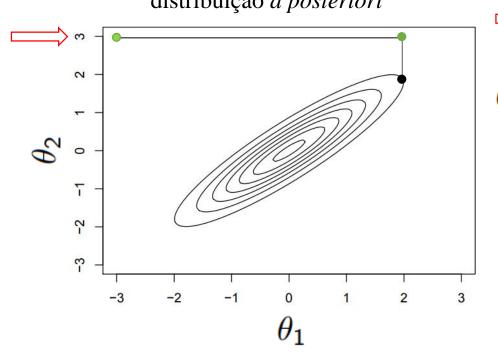
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, \theta_2) & \text{otherwise} \\ p(\theta_2|\theta_1, \theta_2) & \text{otherwise} \end{cases}$$
Distribuição Alvo

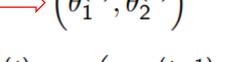


- <u>Método a amostragem:</u>
  - Gibbs atenção
  - Metropolis-Hasting

$$\begin{array}{ccc}
p(\theta|y) & \xrightarrow{\text{com}} & \theta = (\theta_1, \theta_2) & \xrightarrow{\text{amostrar}} & \begin{bmatrix}
p(\theta_1|\theta_2, y) \\
p(\theta_2|\theta_1, y)
\end{bmatrix}$$
Distribuição Alvo





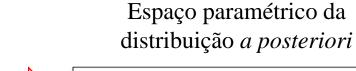


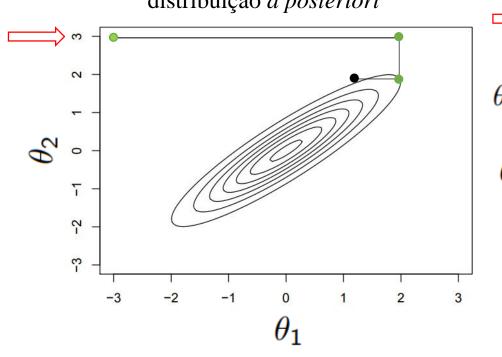
$$^{t)} \sim p\left(\theta_1|\theta_2^{(t-1)}, y\right)$$

$$\theta_2^{(t)} \sim p\left(\theta_2|\theta_1^{(t)}, y\right)$$

- <u>Método a amostragem:</u>
  - Gibbs atenção
  - Metropolis-Hasting

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \sqrt{\frac{p(\theta_1|\theta_2, y_1)}{p(\theta_2|\theta_1, y_2)}}$$
Distribuição Alvo





$$\Longrightarrow \left(\theta_1^{(0)}, \theta_2^{(0)}\right)$$

$$h_1^{(t)} \sim p\left(\theta_1 | \theta_2^{(t-1)}, y\right)$$

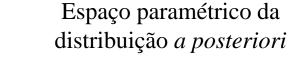
$$\theta_2^{(t)} \sim p\left(\theta_2|\theta_1^{(t)}, y\right)$$

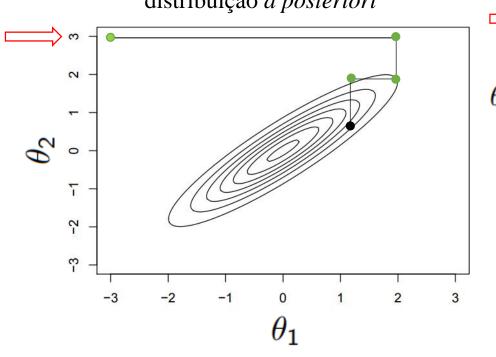
• <u>Método a amostragem:</u>

Gibbs atenção

• Metropolis-Hasting

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$
Distribuição Alvo





$$\Longrightarrow \left(\theta_1^{(0)}, \theta_2^{(0)}\right)$$

$$egin{aligned} egin{pmatrix} egin{pmatrix} (t) \ \lambda \end{bmatrix} \sim p\left( heta_1 | heta_2^{(t-1)}, y 
ight) \end{aligned}$$

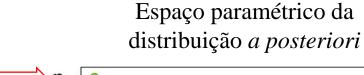
$$\theta_2^{(t)} \sim p\left(\theta_2|\theta_1^{(t)}, y\right)$$

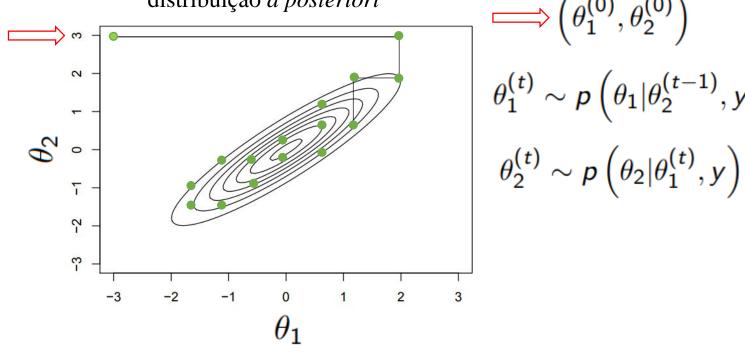
• <u>Método a amostragem:</u>

Gibbs atenção

• Metropolis-Hasting

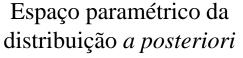
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{bmatrix} p(\theta_1|\theta_2) \\ p(\theta_2|\theta_1) \end{bmatrix}$$
Distribuição Alvo

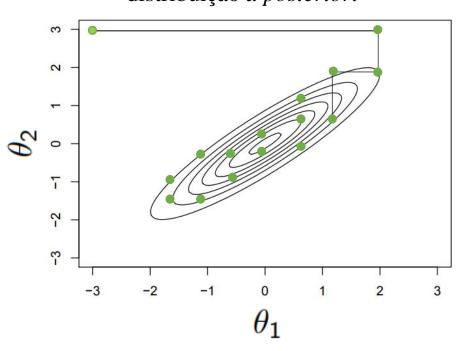


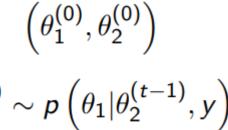


- <u>Método a amostragem:</u>
  - Gibbs atenção
  - Metropolis-Hasting

$$\begin{array}{ccc}
p(\theta|y) & \xrightarrow{\text{com}} & \theta = (\theta_1, \theta_2) & \xrightarrow{\text{amostrar}} & p(\theta_1, \theta_2) \\
\text{Distribuição Alvo} & & & & & & \\
\end{array}$$







$$egin{aligned} heta_1^{(t)} &\sim p\left( heta_1| heta_2^{(t-1)},y
ight) \ heta_2^{(t)} &\sim p\left( heta_2| heta_1^{(t)},y
ight) \end{aligned}$$

- Método a amostragem:
  - Gibbs

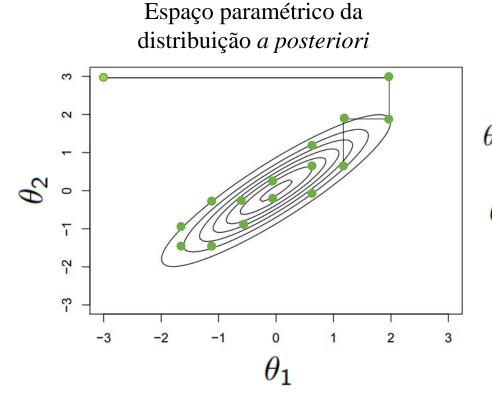
atenção

Metropolis-Hasting

- Amostragem *a posteriori* é marginal por parâmetro;
- Lenta convergência;

(Principalmente espaços paramétricos de alta dimensão).





$$egin{align} \left( heta_1^{(0)}, heta_2^{(0)}
ight) \ &\sim p\left( heta_1| heta_2^{(t-1)},y
ight) \ \end{pmatrix}$$

$$heta_1^{(t)} \sim p\left(\theta_1 | \theta_2^{(t-1)}, y\right)$$
 $heta_2^{(t)} \sim p\left(\theta_2 | \theta_1^{(t)}, y\right)$ 

- Método a amostragem:
  - Gibbs
  - Metropolis-Hasting

atenção

- Amostragem a posteriori conjunta;
- Lenta convergência;

Proporção de teste

(aceitação/rejeição) por

iteração ainda é muito alta.

## Estatísticas & Aquicultura.

Com: Carlos Antônio Zarzar

carloszarzar\_@hotmail.com

## Obrigado! Bons estudos!

"Hierarquia, as vezes é necessário, entretanto nem sempre é preciso."