

Tópicos adicionais

Estimação por máxima verossimilhança e
métodos bayesianos

Aishameriane Schmidt

Seminário para a disciplina de Econometria III

ESAG/UDESC

Professor Fernando Pozzobon

Florianópolis, novembro de 2018.

Função de Verossimilhança

0.2

0.4

0.6

u

λ

0.1

0.2

0.3

0.4

λ

O material está disponível em:

<https://tinyurl.com/econometria-udesc>

Para falar com a Aisha, utilize o aishamail:

aishameriane@gmail.com

Um tour do que vamos ver hoje

- O que é fazer inferência estatística?



Um tour do que vamos ver hoje

- O que é fazer inferência estatística?
- Relembrando o EMV e suas propriedades



Um tour do que vamos ver hoje

- O que é fazer inferência estatística?
- Relembrando o EMV e suas propriedades
- Quando o EMV dá ruim?



Um tour do que vamos ver hoje

- O que é fazer inferência estatística?
- Relembrando o EMV e suas propriedades
- Quando o EMV dá ruim?
- O que é inferência Bayesiana?
 - Definições básicas: verossimilhança, priori e posteriori.



Um tour do que vamos ver hoje

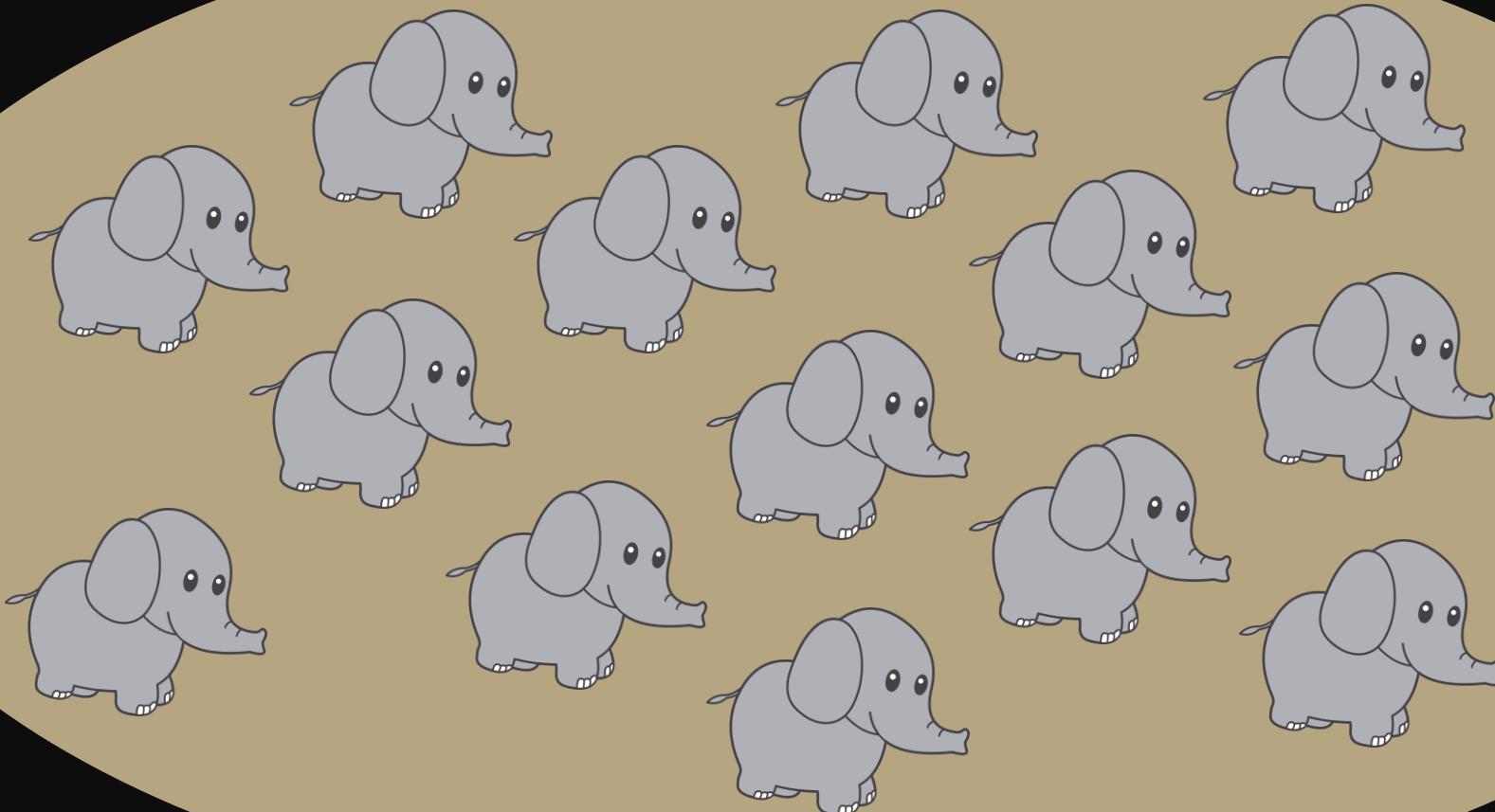
- O que é fazer inferência estatística?
- Relembrando o EMV e suas propriedades
- Quando o MLE dá ruim?
- O que é inferência Bayesiana?
 - Definições básicas: verossimilhança, priori e posteriori.
- Estimando a probabilidade de cara em uma moeda



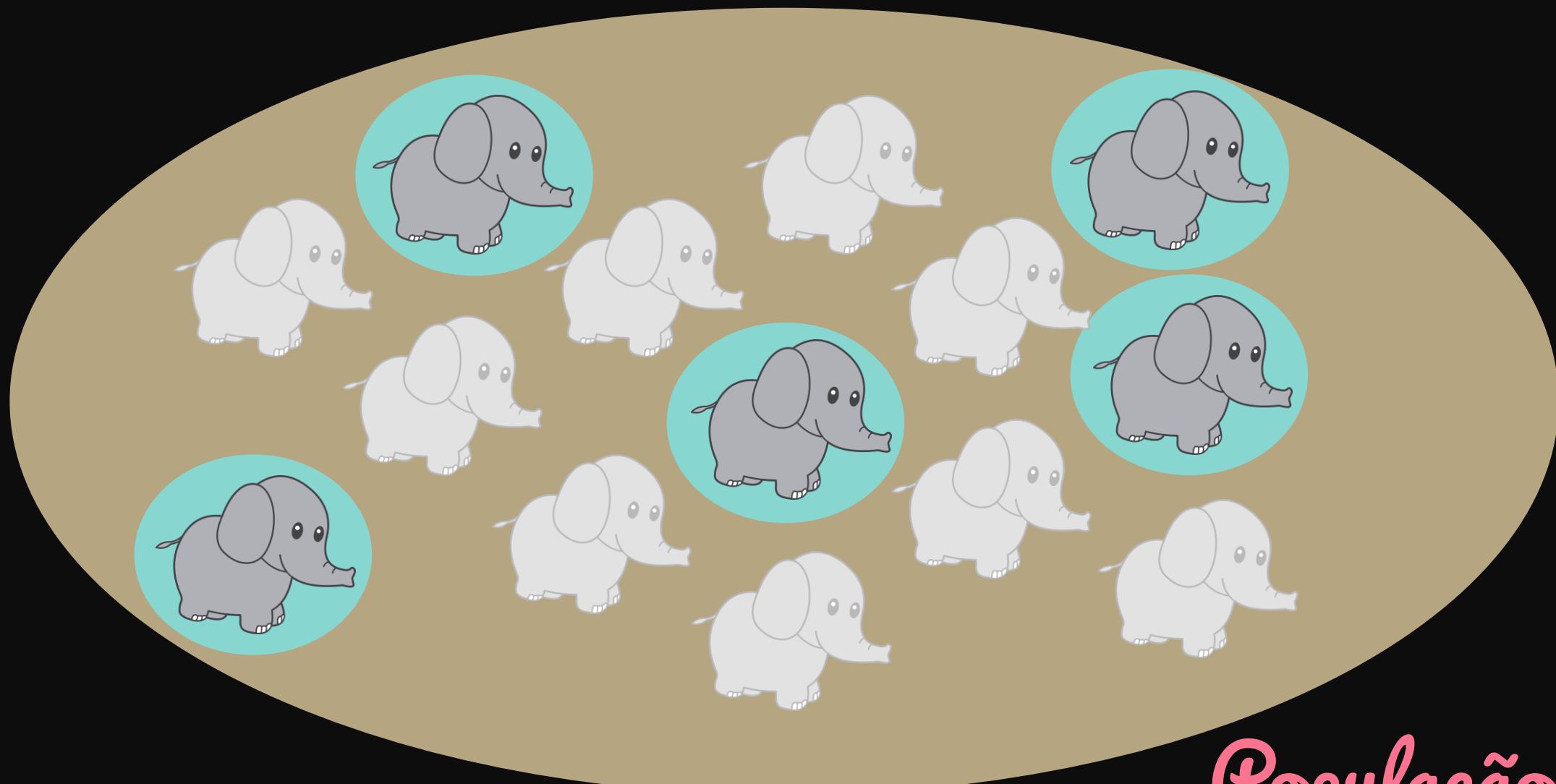
Problema fundamental da estatística



População



Amostra!



População

Usamos os dados da *amostra*



Em conjunto com *estimadores*



Para obter *estimativas*



Dos *parâmetros populacionais*

Usamos os dados da *amostra*



Em conjunto com *estimadores*



Para obter *estimativas*



Dos *parâmetros populacionais*

Precisamos de métodos de
encontrar estimadores como
GMM, MLE, MQO, etc

Usamos os dados da *amostra*



Em conjunto com *estimadores*



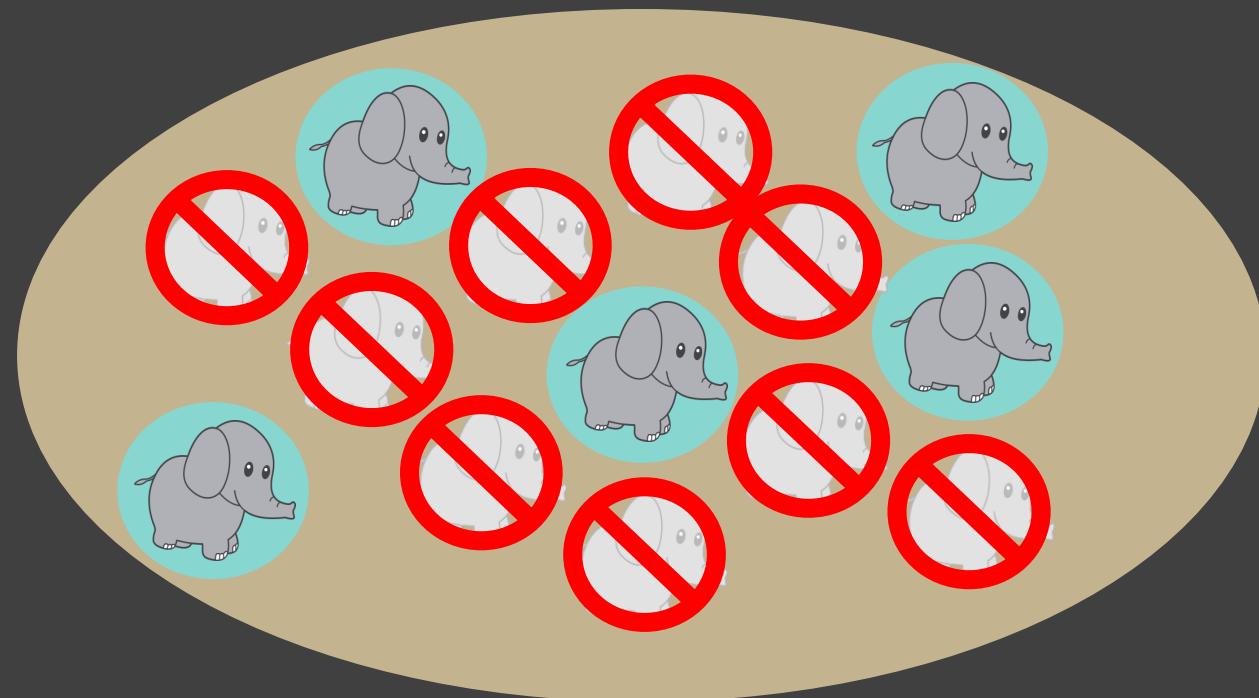
Para obter *(estimativas)*



Dos *parâmetros populacionais*

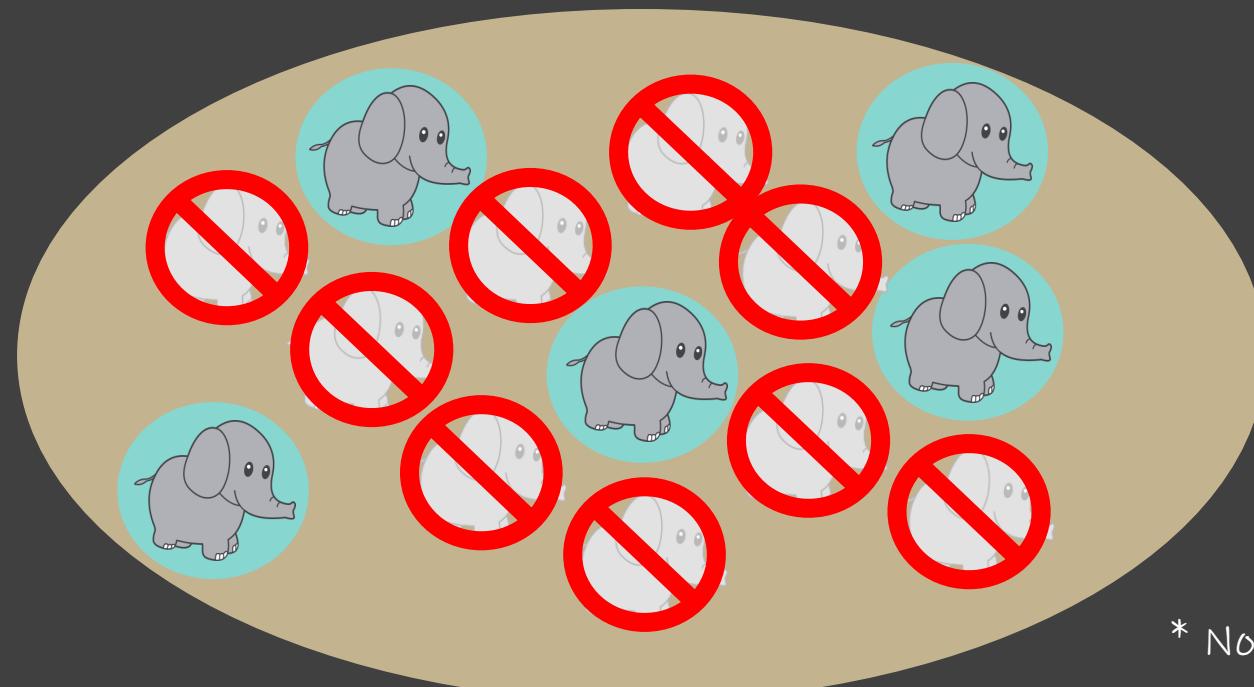
Nossos estimadores
devem produzir
estimativas com
propriedades boas

No "mundo real", nós geralmente **não temos como coletar mais dados** para comparar estimativas a partir de várias amostras.



No "mundo real", nós geralmente **não temos como coletar mais dados** para comparar estimativas a partir de várias amostras.

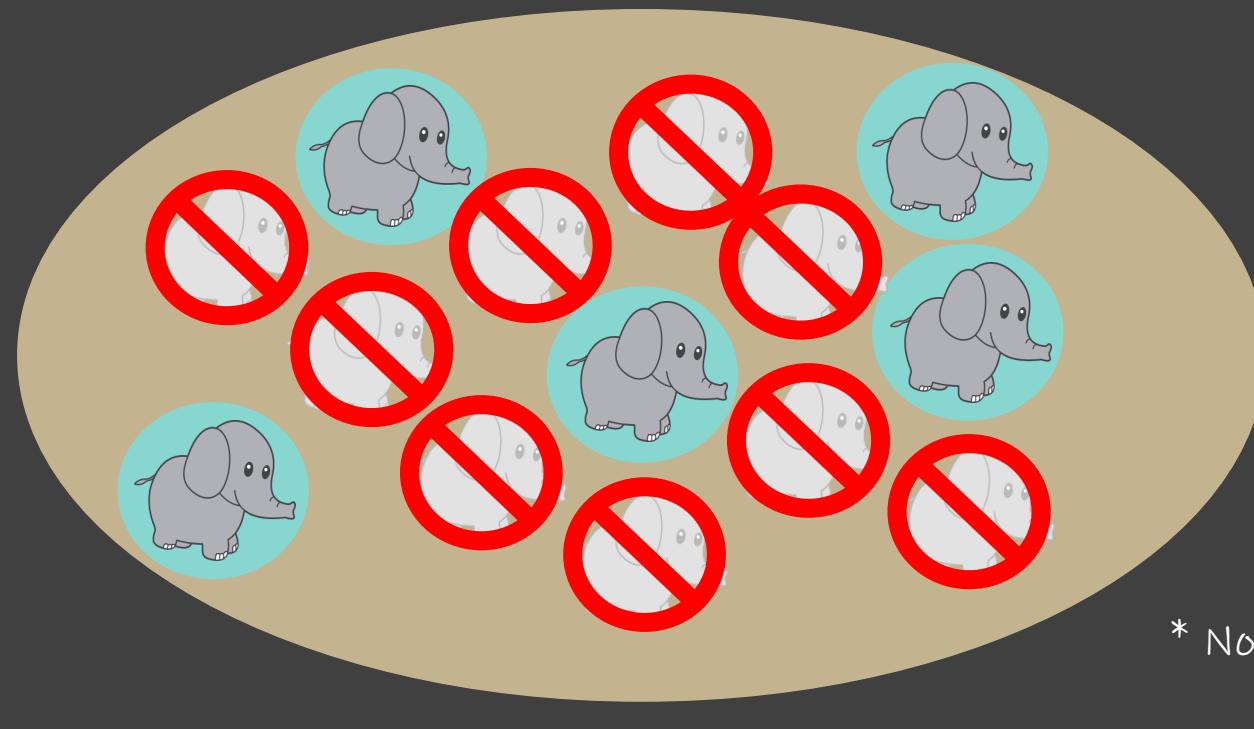
Nosso estimador irá produzir um número* (ou intervalo) e teremos que confiar que ele está sendo informativo a respeito da população de interesse.



* No caso bayesiano obteremos uma densidade de probabilidade e não apenas um número. Já iremos discutir sobre isso.

No "mundo real", nós geralmente **não temos como coletar mais dados** para comparar estimativas a partir de várias amostras.

Nosso estimador irá produzir um número* (ou intervalo) e teremos que confiar que ele está sendo informativo a respeito da população de interesse.



Uma forma de "dormirmos tranquilos" é utilizar **estimadores com boas propriedades teóricas**

(mesmo que isso seja válido apenas sob certas condições de regularidade).

* No caso bayesiano obteremos uma densidade de probabilidade e não apenas um número. Já iremos discutir sobre isso.

O estimador de máxima verossimilhança

- ✓ Nem sempre é possível garantir que o estimador de MQO é Blue;



O estimador de máxima verossimilhança

- ✓ Nem sempre é possível garantir que o estimador de MQO é Blue;
- ✓ Nem sempre DÁ para calcular o estimador de MQO (nem tudo na vida é regressão linear): $y_i^\delta = \beta_0 + \beta_1 x_i + \varepsilon_i, \varepsilon_i \sim IID(0, \sigma^2)$



O estimador de máxima verossimilhança

- ✓ Nem sempre é possível garantir que o estimador de MQO é Blue;
- ✓ Nem sempre DÁ para usar o estimador de MQO (nem tudo na vida é regressão linear)
- ✓ Sob algumas condições, o estimador de MQO e o de MV serão iguais;



O estimador de máxima verossimilhança

- ✓ Nem sempre é possível garantir que o estimador de MQO é Blue;
- ✓ Nem sempre DÁ para usar o estimador de MQO (nem tudo na vida é regressão linear)
- ✓ Sob algumas condições, o estimador de MQO e o de MV serão iguais;
- ✓ O EMV é mais geral que o MQO podendo ser calculado para diferentes "modelos"



O estimador de máxima verossimilhança

- ✓ Nem sempre é possível garantir que o estimador de MQO é Blue;
- ✓ Nem sempre DÁ para usar o estimador de MQO (nem tudo na vida é regressão linear)
- ✓ Sob algumas condições, o estimador de MQO e o de MV serão iguais;
- ✓ O EMV é mais geral que o MQO podendo ser calculado para diferentes "modelos"



O estimador de MV, quando atendidas as hipóteses de regularidade, tem as seguintes propriedades:

- Consistência* (embora seja comum que seja viesado)
- Normalidade assintótica
- Eficiência assintótica
- Invariância

Nem tudo são flores na vida do EMV

- ✓ Precisamos ter especificado **corretamente** o modelo;



Nem tudo são flores na vida do EMV

- ✓ Precisamos ter especificado corretamente o modelo;
- ✓ As propriedades em **amostras finitas** podem ser muito diferentes das propriedades **assintóticas** que tanto gostamos;



Nem tudo são flores na vida do EMV

- ✓ Precisamos ter especificado corretamente o modelo;
- ✓ As propriedades em amostras finitas podem ser muito diferentes das propriedades assintóticas que tanto gostamos;
- ✓ Nem sempre dá para **escrever** facilmente a **função de verossimilhança**



Nem tudo são flores na vida do EMV

- ✓ Precisamos ter especificado corretamente o modelo;
- ✓ As propriedades em amostras finitas podem ser muito diferentes das propriedades assintóticas que tanto gostamos;
- ✓ Nem sempre dá para escrever facilmente a função de verossimilhança
- ✓ Algumas vezes você até consegue, mas **maximizar é um problema...**





Exemplo



Queremos saber se a moeda é honesta, ou seja,

$$\mathbb{P}(Panda) = \mathbb{P}(Não\ panda) = \frac{1}{2}$$



Vamos definir
 $\mathbb{P}(Panda) \stackrel{\text{def}}{=} \theta$

**Suponha que a moeda foi lançada 3 vezes e foram observados 3 pandas.
Qual seria sua estimativa para θ ?**



Precisamos da f.m.p. de y , o resultado do lançamento da moeda.

Se

$$y \sim Bernoulli(\theta),$$

então

$$p(y|\theta) = \theta^y \cdot (1 - \theta)^{1-y}$$



Precisamos da f.m.p. de y , o resultado do lançamento da moeda.

Se

$$y \sim Bernoulli(\theta),$$

então

$$p(y|\theta) = \theta^y \cdot (1 - \theta)^{1-y}$$

Para obter a verossimilhança, escrevemos a densidade conjunta:

$$\ell(\theta|y) = \theta^{\sum_{i=1}^n y} \cdot (1 - \theta)^{n - \sum_{i=1}^n y}$$



Precisamos da f.m.p. de y , o resultado do lançamento da moeda.

Se

$$y \sim Bernoulli(\theta),$$

então

$$p(y|\theta) = \theta^y \cdot (1 - \theta)^{1-y}$$

Para obter a verossimilhança, escrevemos a densidade conjunta:

$$\ell(\theta|y) = \theta^{\sum_{i=1}^n y_i} \cdot (1 - \theta)^{n - \sum_{i=1}^n y_i}$$

Tomamos o log e maximizamos para obter

$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^n y_i}{n}$$



Agora, podemos aplicar os dados da nossa amostra,

$$n = 3 \text{ e } \sum y = 3.$$



Agora, podemos aplicar os dados da nossa amostra,

$$n = 3 \text{ e } \sum y = 3.$$

$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^3 y_i}{3} = \frac{3}{3} = 1$$



Agora, podemos aplicar os dados da nossa amostra,

$$n = 3 \text{ e } \sum y = 3.$$

$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^3 y_i}{3} = \frac{3}{3} = 1$$

Nosso melhor palpite
sobre a probabilidade
de sair panda nessa
moeda é igual a 1 !?!



Created by Anniken & Andreas
from Noun Project



Agora, podemos aplicar os dados da nossa amostra,

$$n = 3 \text{ e } \sum y = 3.$$

$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^3 y_i}{3} = \frac{3}{3} = 1$$

O EMV sofre de excesso de ajuste, i.e., ele não atribui probabilidade aquilo que não aconteceu!





Agora, podemos aplicar os dados da nossa amostra,

$$n = 3 \text{ e } \sum y = 3.$$

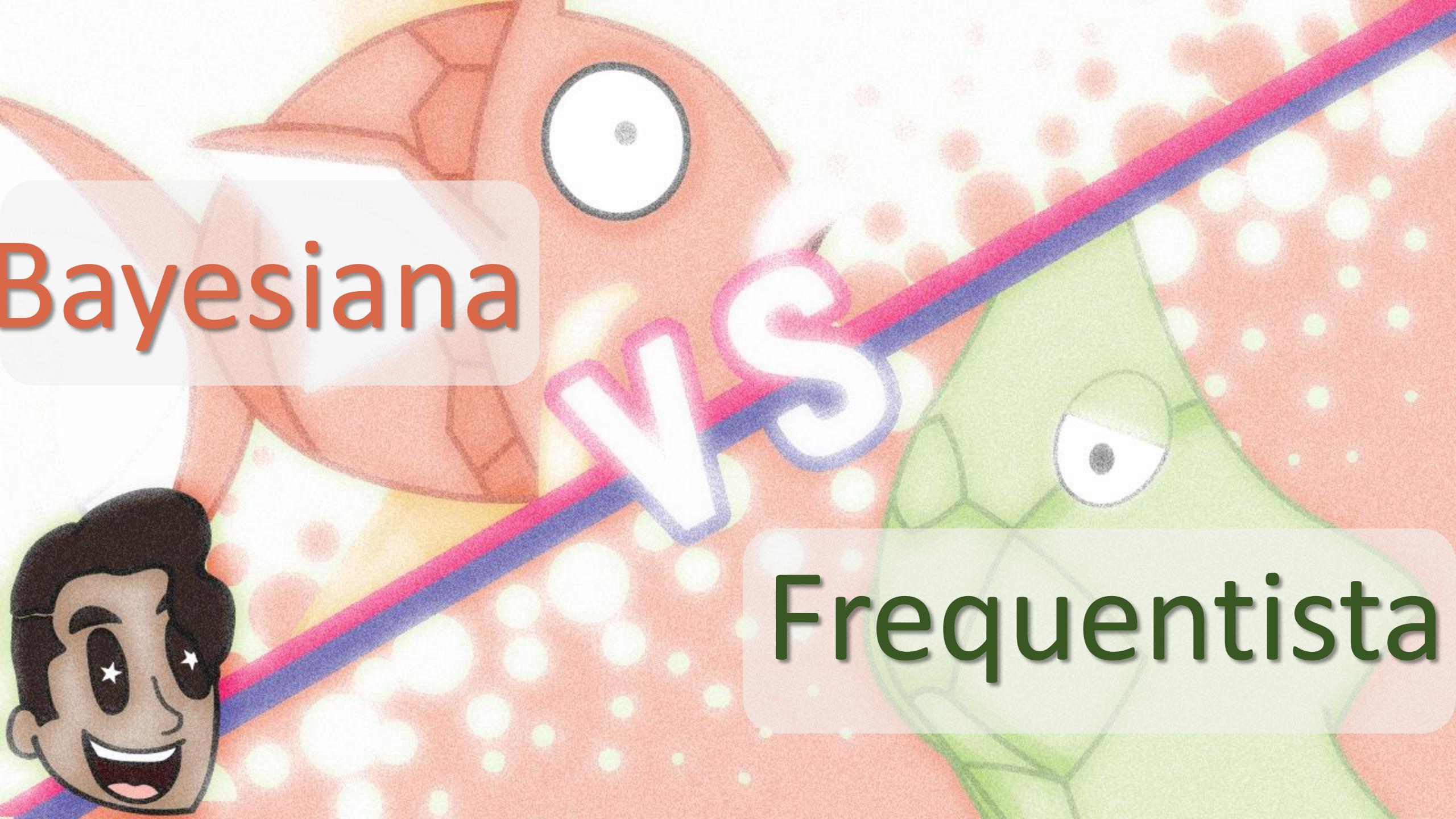
$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^3 y_i}{3} = \frac{3}{3} = 1$$



Então agora o que podemos fazer?

Podemos utilizar inferência bayesiana!





Bayesiana

VS

Frequentista

Inferência Clássica (ou frequentista)

Os parâmetros são quantidades fixas (e desconhecidas)

Parâmetros estáticos,
como o Metapod



Inferência Clássica (ou frequentista)

Os parâmetros são quantidades fixas (e desconhecidas)

Usamos amostras em combinação com estimadores para obter estimativas para estas quantidades

Parâmetros estáticos,
como o Metapod



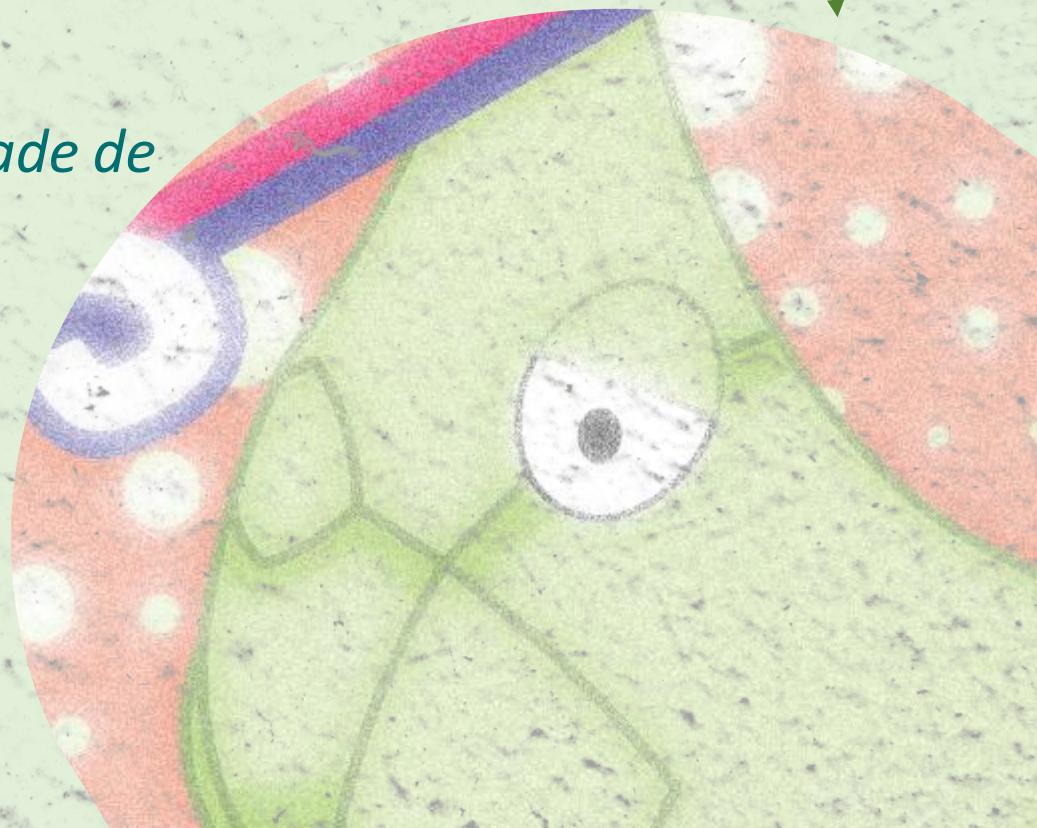
Inferência Clássica (ou frequentista)

Os parâmetros são quantidades fixas (e desconhecidas)

Usamos amostras em combinação com estimadores para obter estimativas para estas quantidades

A inferência se baseia em p-valores: *a probabilidade de encontrar um valor igual ou mais extremo assumindo que a hipótese nula é verdadeira*

Parâmetros estáticos,
como o Metapod



Inferência Clássica (ou frequentista)

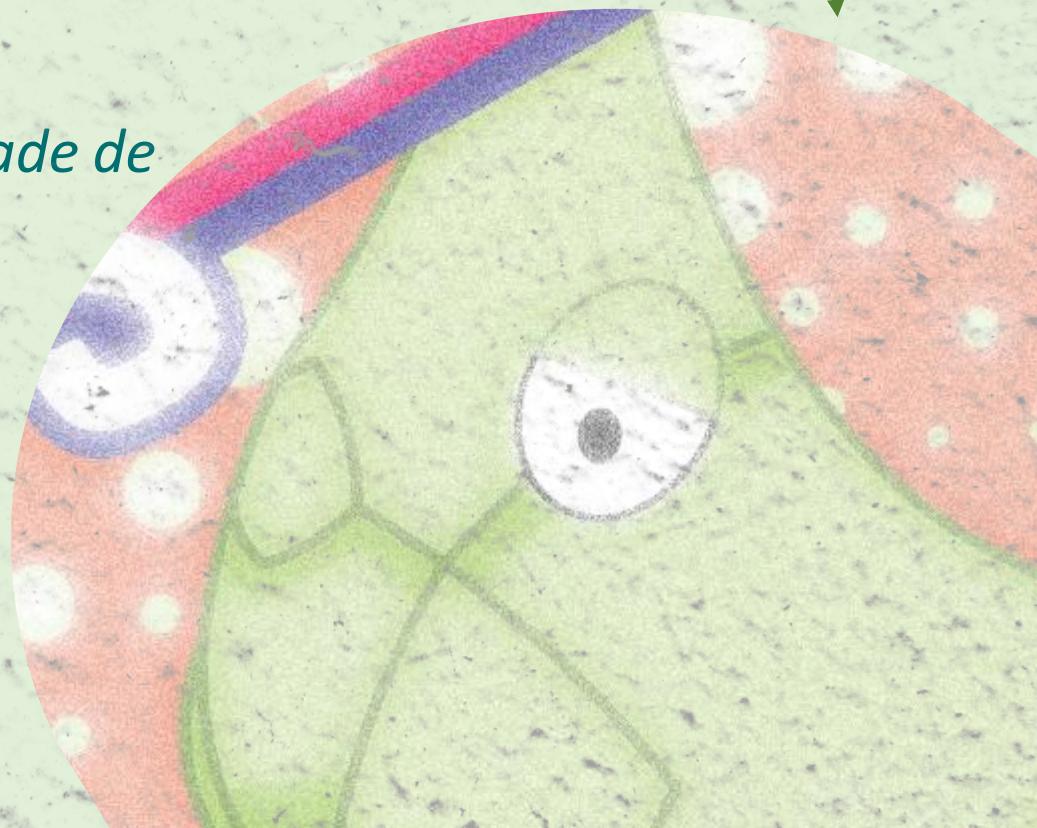
Os parâmetros são quantidades fixas (e desconhecidas)

Usamos amostras em combinação com estimadores para obter estimativas para estas quantidades

A inferência se baseia em p -valores: *a probabilidade de encontrar um valor igual ou mais extremo assumindo que a hipótese nula é verdadeira*

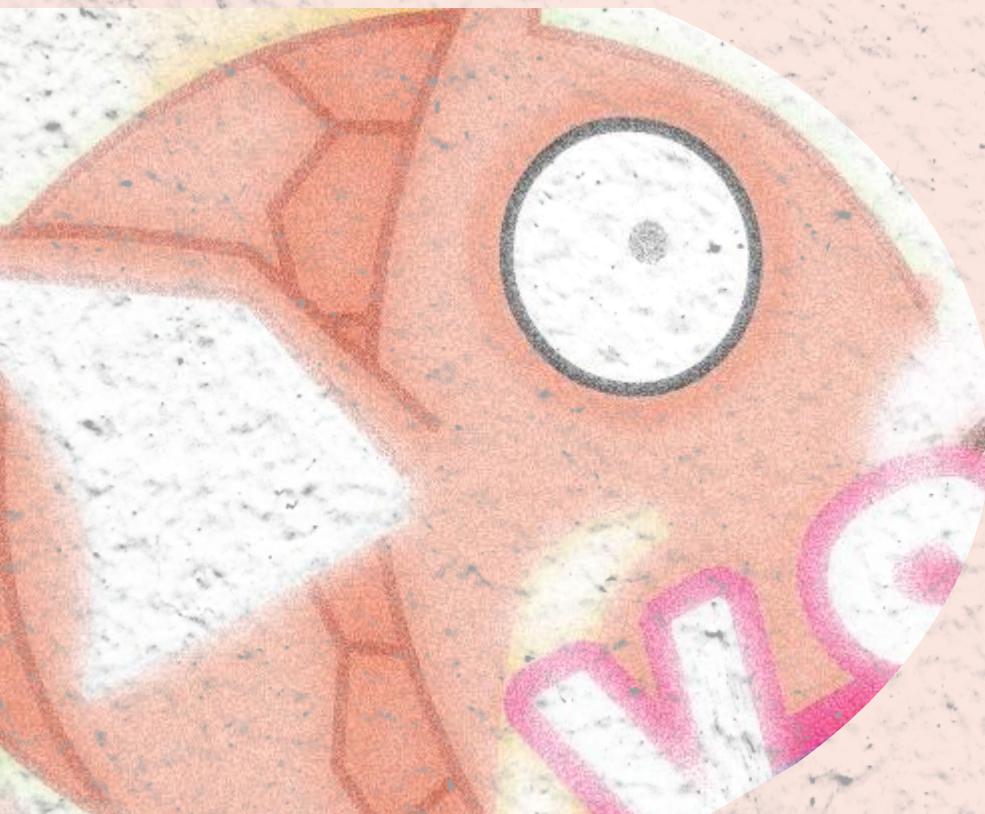
Usamos intervalos com $1 - \alpha\%$ de confiança:
Se fossem coletadas 100 amostras, esperaríamos que $1 - \alpha\%$ dos intervalos construídos contivessem o verdadeiro valor do parâmetro.

Parâmetros estáticos,
como o Metapod



Inferência Bayesiana

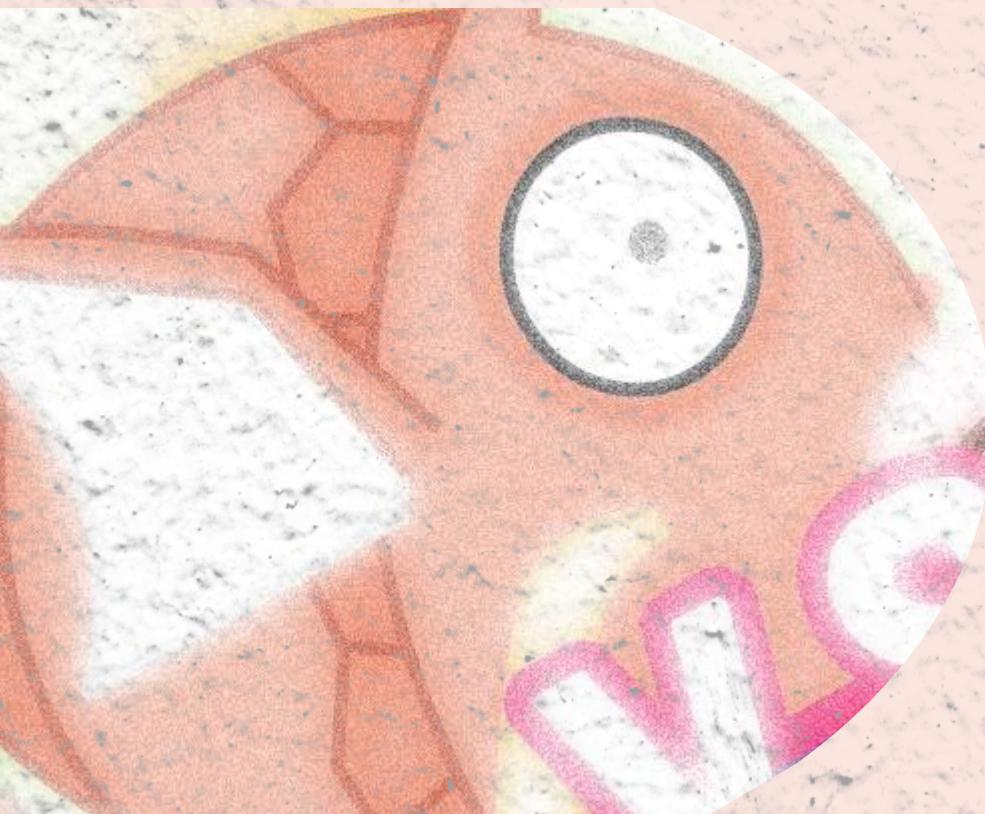
Os parâmetros seguem densidade de probabilidade e seus valores podem mudar de posição (mas não muito), como o Magikarp.



Os parâmetros são variáveis aleatórias

Inferência Bayesiana

Os parâmetros seguem densidade de probabilidade e seus valores podem mudar de posição (mas não muito), como o Magikarp.

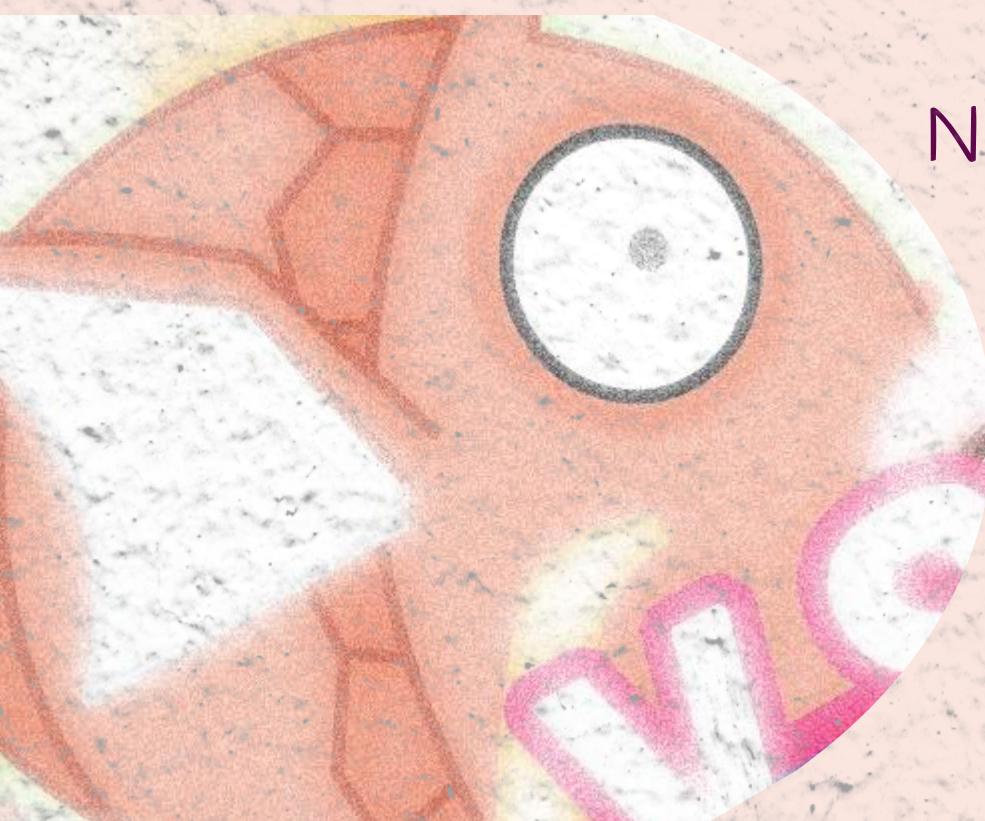


Os parâmetros são variáveis aleatórias

Usamos informações dos dados (verossimilhança) com informações a priori para obter a densidade a posteriori

Inferência Bayesiana

Os parâmetros seguem densidade de probabilidade e seus valores podem mudar de posição (mas não muito), como o Magikarp.

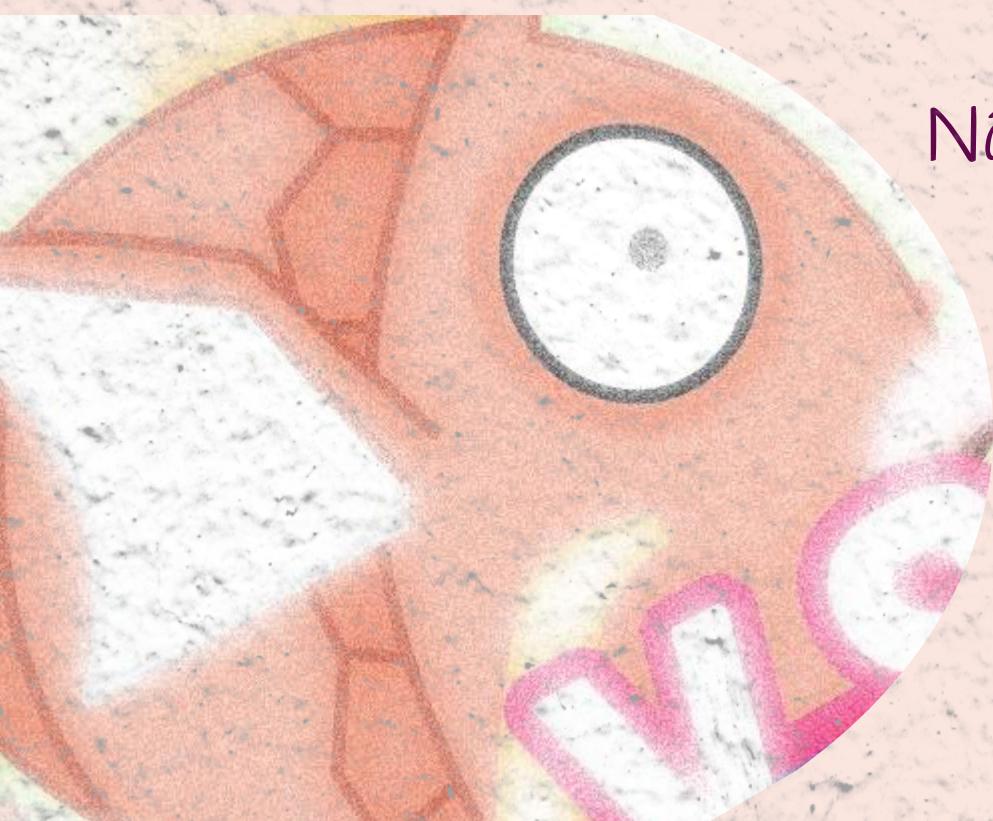


Os parâmetros são variáveis aleatórias
Usamos informações dos dados (verossimilhança) com informações a priori para obter a densidade a posteriori

Não existem p-valores aqui! As probabilidades para os parâmetros podem ser calculadas diretamente

Inferência Bayesiana

Os parâmetros seguem densidade de probabilidade e seus valores podem mudar de posição (mas não muito), como o Magikarp.



Os parâmetros são variáveis aleatórias
Usamos informações dos dados (verossimilhança) com informações a priori para obter a densidade a posteriori

Não existem p-valores aqui! As probabilidades para os parâmetros podem ser calculadas diretamente

Usamos intervalos de credibilidade:
A probabilidade do parâmetro estar no intervalo $[a;b]$ é igual a 0.95.



Definições Básicas

Teorema de Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$


 $P(A \cap B)$

Teorema de Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$



Informação de
A dado B

Teorema de Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Informação de
B dado A



Informação de
A dado B



Teorema de Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Informação de
B dado A

Informação de
A sem saber de B

Informação de
A dado B

Teorema de Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Informação dos dados
(verossimilhança)

Informação prévia sobre A
Priori para A

Posteriori para A



Exemplo (o retorno)



Queremos saber se a moeda é honesta, ou seja,

$$\mathbb{P}(\theta) = \mathbb{P}(1 - \theta) = \frac{1}{2}$$

**Para a estimação bayesiana,
precisaremos de uma priori**



Queremos saber se a moeda é honesta, ou seja,

$$\mathbb{P}(\theta) = \mathbb{P}(1 - \theta) = \frac{1}{2}$$

O que sabemos sobre moedas?



**Para a estimação bayesiana,
precisaremos de uma priori**



Podemos considerar que θ é uma quantia aleatória e escolher uma densidade a priori:

- ✓ Como θ é uma probabilidade, deve estar entre zero e um;



Created by Anniken & Andreas
from Noun Project



Podemos considerar que θ é uma quantia aleatória e escolher uma densidade a priori:

- ✓ Como θ é uma probabilidade, deve estar entre zero e um;
- ✓ Além disso, vamos supôr que θ assume qualquer valor neste intervalo;



Created by Anniken & Andreas
from Noun Project



Podemos considerar que θ é uma quantia aleatória e escolher uma densidade a priori:

- ✓ Como θ é uma probabilidade, deve estar entre zero e um;
- ✓ Além disso, vamos supôr que θ assume qualquer valor neste intervalo;
- ✓ Vamos assumir a moeda é (possivelmente) honesta, isto é, acreditamos que θ tem uma maior "chance" de ser 50%.



Created by Anniken & Andreas
from Noun Project



Podemos considerar que θ é uma quantia aleatória e escolher uma densidade a priori:



Created by Anniken & Andreas
from Noun Project

- ✓ Como θ é uma probabilidade, deve estar entre zero e um;
- ✓ Além disso, vamos supôr que θ assume qualquer valor neste intervalo;
- ✓ Vamos assumir a moeda é (possivelmente) honesta, isto é, acreditamos que θ tem uma maior "chance" de ser 50%.
- ✓ A f.d.p. de uma distribuição $Beta(\alpha, \beta)$ para θ é dada por

$$f_{\theta} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

Distribuição Beta(10,10)

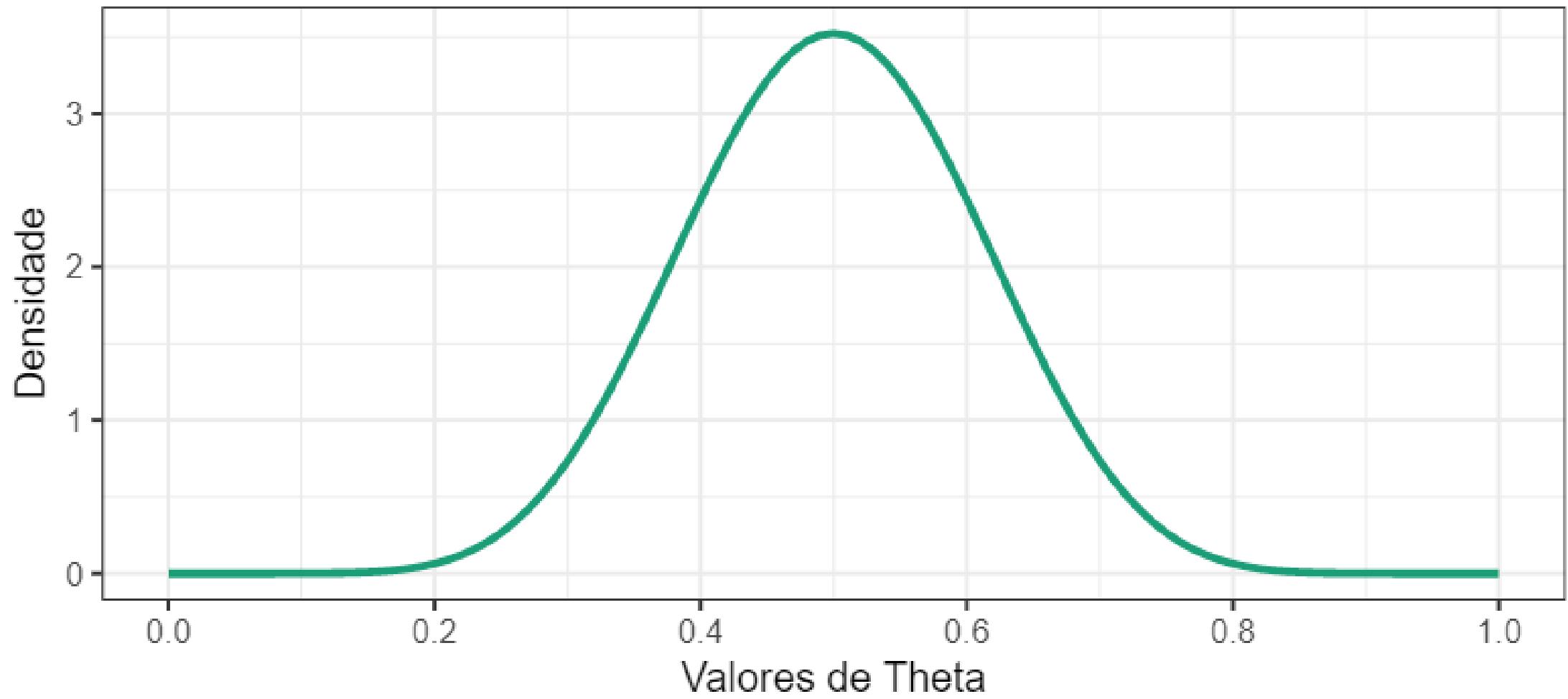
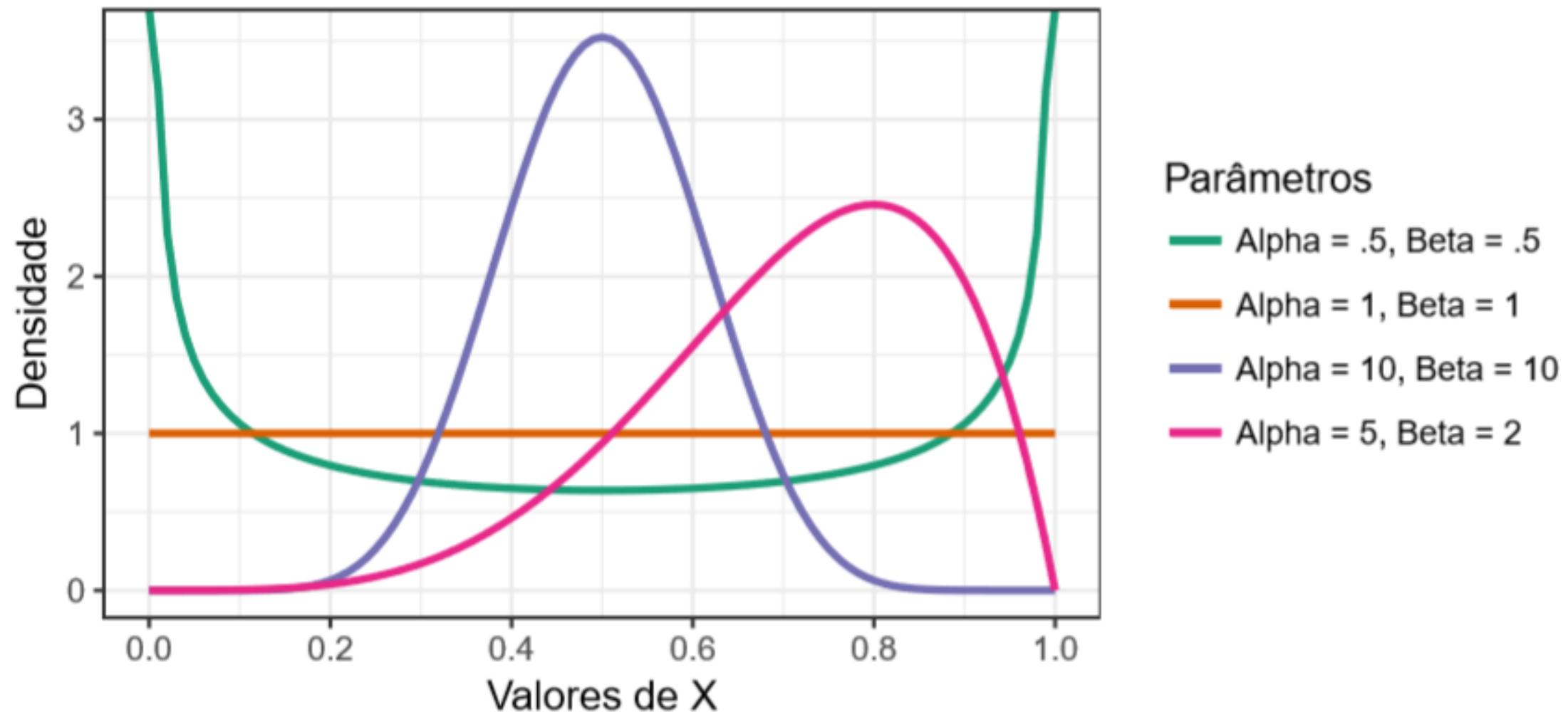


Figura 1 – Diferentes formas da distribuição $Beta(\alpha, \beta)$





Podemos considerar que θ é uma quantia aleatória e escolher uma densidade a priori:

- ✓ Obtemos a posteriori através do produto da verossimilhança com a priori



$$f(\theta|y) \propto \underbrace{\theta^{\alpha-1}(1-\theta)^{\beta-1}}_{\text{Priori}} \cdot \underbrace{\theta^{\sum_{i=1}^n y} \cdot (1-\theta)^{n-\sum_{i=1}^n y}}_{\text{Verossimilhança}}$$



Podemos considerar que θ é uma quantia aleatória e escolher uma densidade a priori:

- ✓ Obtemos a posteriori através do produto da verossimilhança com a priori



Created by Anniken & Andreas
from Noun Project

$$f(\theta|y) \propto \theta^{\alpha-1} (1-\theta)^{\beta-1} \cdot \theta^{\sum_{i=1}^n y_i} \cdot (1-\theta)^{n-\sum_{i=1}^n y_i}$$

$$\Rightarrow f(\theta|y) = \frac{\Gamma(\bar{\alpha} + \bar{\beta})}{\Gamma(\bar{\alpha})\Gamma(\bar{\beta})} \theta^{\bar{\alpha}-1}(1-\theta)^{\bar{\beta}-1}$$

* A conta inteira pode ser vista em <http://sistemabu.udesc.br/pergamenweb/vinculos/000042/000042c2.pdf>

$$\text{Com } \bar{\alpha} = \alpha + \sum y_i \quad e \quad \bar{\beta} = \beta + n - \sum y_i$$



Created by Anniken & Andreas
from Noun Project

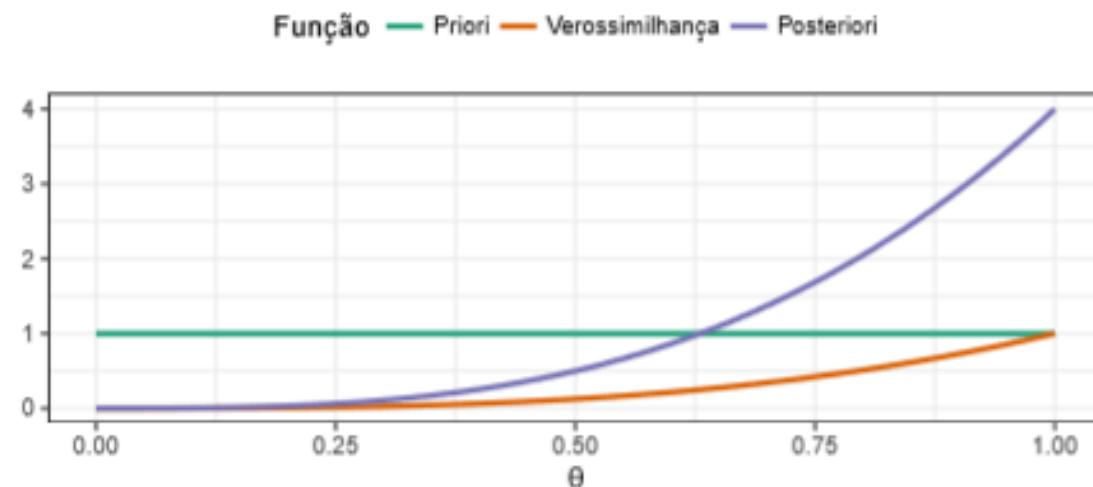
Se usarmos uma densidade a priori beta, a posteriori resultante neste caso também é beta

- ✓ Os parâmetros da posteriori são combinações dos hiperparâmetros da priori com os valores da verossimilhança.

$$\bar{\alpha} = \sum y + \underline{\alpha} \text{ e } \bar{\beta} = \underline{\beta} + n - \sum y$$

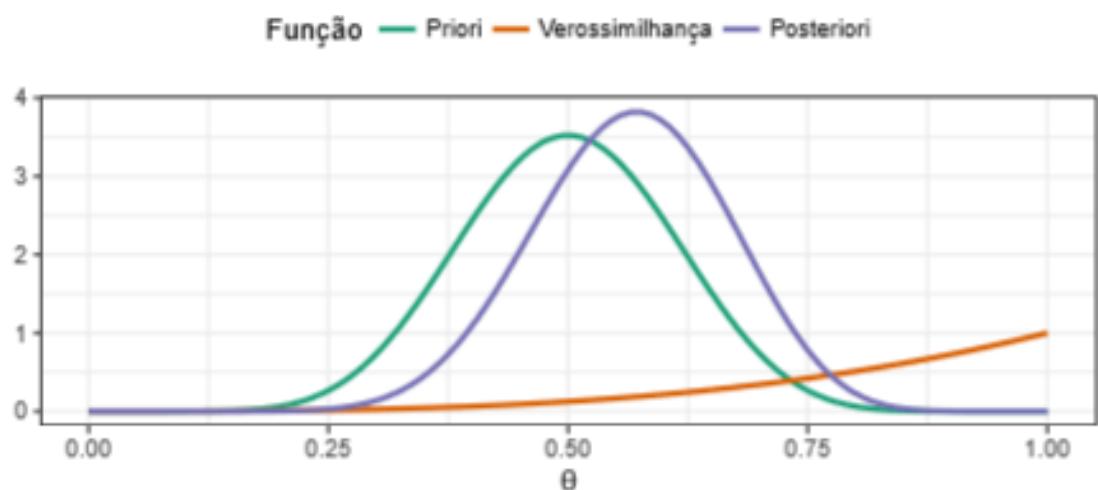
Para 3 pandas, temos que o valor esperado de θ a posteriori, se usarmos uma Beta(1,1) como priori, será 0.8.

Figura 2 – Triplots para duas prioris diferentes no exemplo (2.3)



(a) Priori Beta(1,1) e posteriori Beta(4,1)

Figura 2 – Triplots para duas prioris diferentes no exemplo (2.3)



(b) Priori Beta(10,10) e posteriori Beta(23,10)

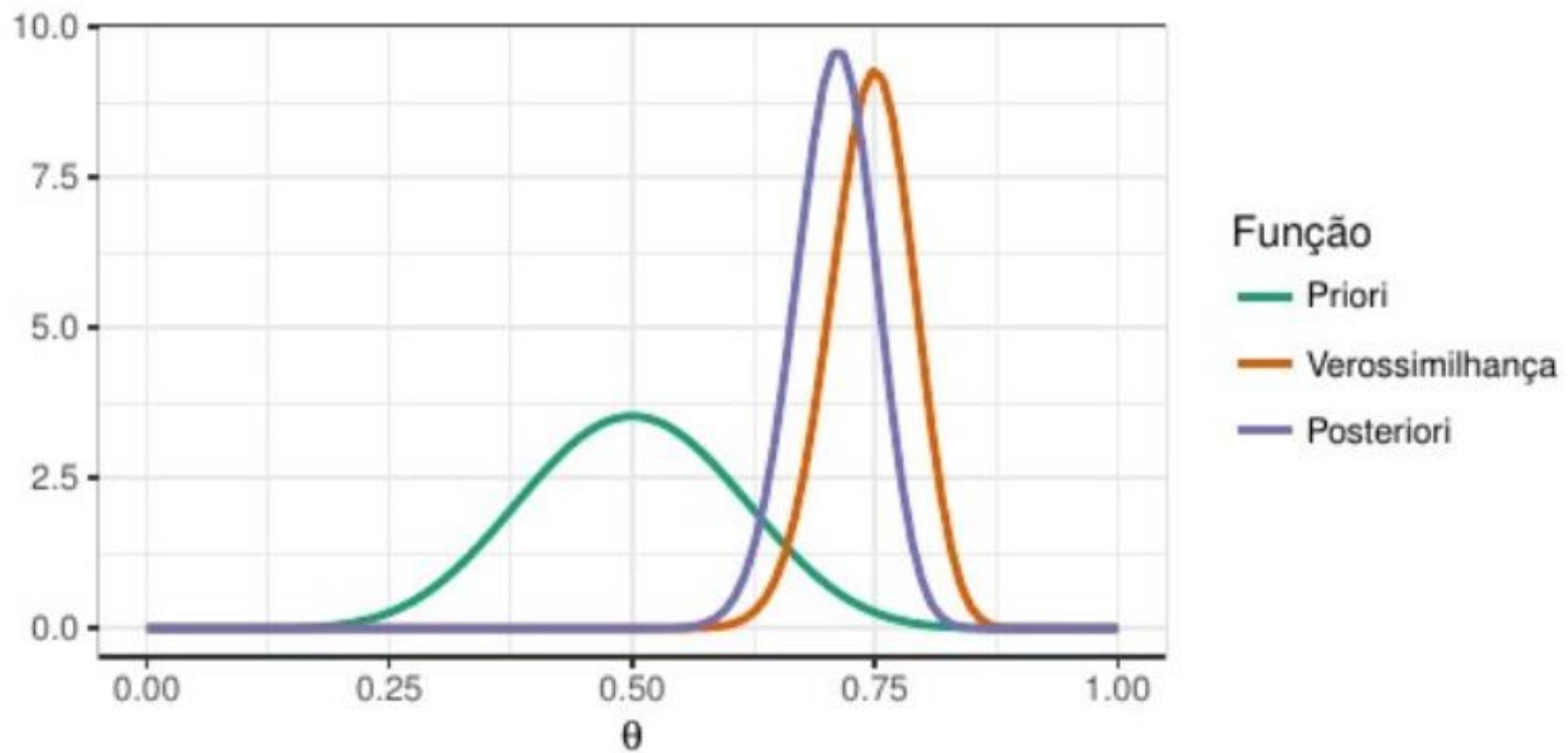


Figura 15: Triplot para priori Beta(10,10) e verossimilhança Binomial com $n = 100$ e $\sum y = 75$.



Aplicações contemporâneas

Bayesiana na vida dos economistas

- ✓ Métodos bayesianos permitem incorporar prioris de maneira mais intuitiva aos modelos

Bayesiana na vida dos economistas

- ✓ Métodos bayesianos permitem incorporar prioris de maneira mais intuitiva aos modelos
- ✓ Modelos atuais, especialmente modelos macroeconômicos, são complexos e nem sempre métodos clássicos são possíveis de serem utilizados

Bayesiana na vida dos economistas

- ✓ Métodos bayesianos permitem incorporar prioris de maneira mais intuitiva aos modelos
- ✓ Modelos atuais, especialmente modelos macroeconômicos, são complexos e nem sempre métodos clássicos são possíveis de serem utilizados
- ✓ Bancos Centrais utilizam metodologia bayesiana para estimação de modelos VAR, DSGE, etc.

Bayesian Methods in Applied Econometrics, or, Why Econometrics Should Always and Everywhere Be Bayesian

Christopher A. Sims
Princeton University
sims@princeton.edu

August 6, 2007

Professor Christopher Sims,
prêmio Nobel em Economia (2011):
<http://www.princeton.edu/~sims/>

UNDERSTANDING NON-BAYESIANS

ABSTRACT.

I. INTRODUCTION

Once one becomes used to thinking about inference from a Bayesian perspective, it becomes difficult to understand why many econometricians are uncomfortable with that way of thinking. But some very good econometricians are either firmly non-Bayesian or (more commonly these days) think of Bayesian approaches as a “tool” which might sometimes be appropriate, sometimes not. This paper tries to articulate the counterarguments to a Bayesian perspective. There are some counterarguments that are frequently expressed, but are not hard to dismiss. Others, though, correspond to types of application where convenient, seemingly sensible, frequentist tools exist, while Bayesian approaches are either not yet developed or seem quite incon-

PROBABILITY MODELS FOR MONETARY POLICY DECISIONS

CHRISTOPHER A. SIMS



FEDERAL RESERVE BANK of ATLANTA

RESEARCH & DATA ECONOMY MATTERS BANKING & PAYMENTS NEWS & EVENTS EDUCATION COMMUNITY DEV

CONFERENCES & EVENTS PRESS RELEASES SPEECHES

CONFERENCE & EVENTS

ESOBE 2018 New Orleans - October 11-12, 2018



European Seminar on Bayesian Econometric.

New Orleans Branch – Federal Reserve Bank of Atlanta
525 Saint Charles Avenue
New Orleans, LA 70130

Call for Papers

We invite papers applying Bayesian econometrics and statistical methods to financial time series analysis, risk management, economic growth analysis, measurement of policy effectiveness, individual decision making in marketing, labor market analysis, forecasting in monetary policy, just to name a few.

Authors should submit an extended abstract or completed paper by July 13, 2018. Acceptance notifications will be sent by August 17, 2018.

[Click here to submit papers](#)

All presenters must register by September 10, 2018, in order to guarantee a spot on the program. After this deadline date participants will be charged an additional \$100 fee.

Registration fees are:
Non-students: \$300
Students: \$150

Apresenta_o_SIN....pdf

Associação Brasileira de Estatística
XIV EBEB - Encontro Brasileiro de Estatística Bayesiana - Rio de Janeiro

Login

Taxes Fazer cadastro XIV EBEB Atividades / Resumos Comissões Submissões Mais ▾

XIV EBEB

De 5 a 8 de março de 2018, IMPA, Rio de Janeiro, Brasil



Dois exemplos de trabalhos de econometria bayesiana: [Pôster na ESOBE](#) e [Pôster na EBEB](#).

A cartoon illustration of a blue submarine with a smiling face, wearing purple goggles. It has two periscopes on top. The word "Obrigada!" is overlaid in white script text.

Obrigada!