

Teste de hipótese

por Reamostragem

Murillo F. Rodrigues

Mestrando em Genética e Biologia Evolutiva (IB/USP)

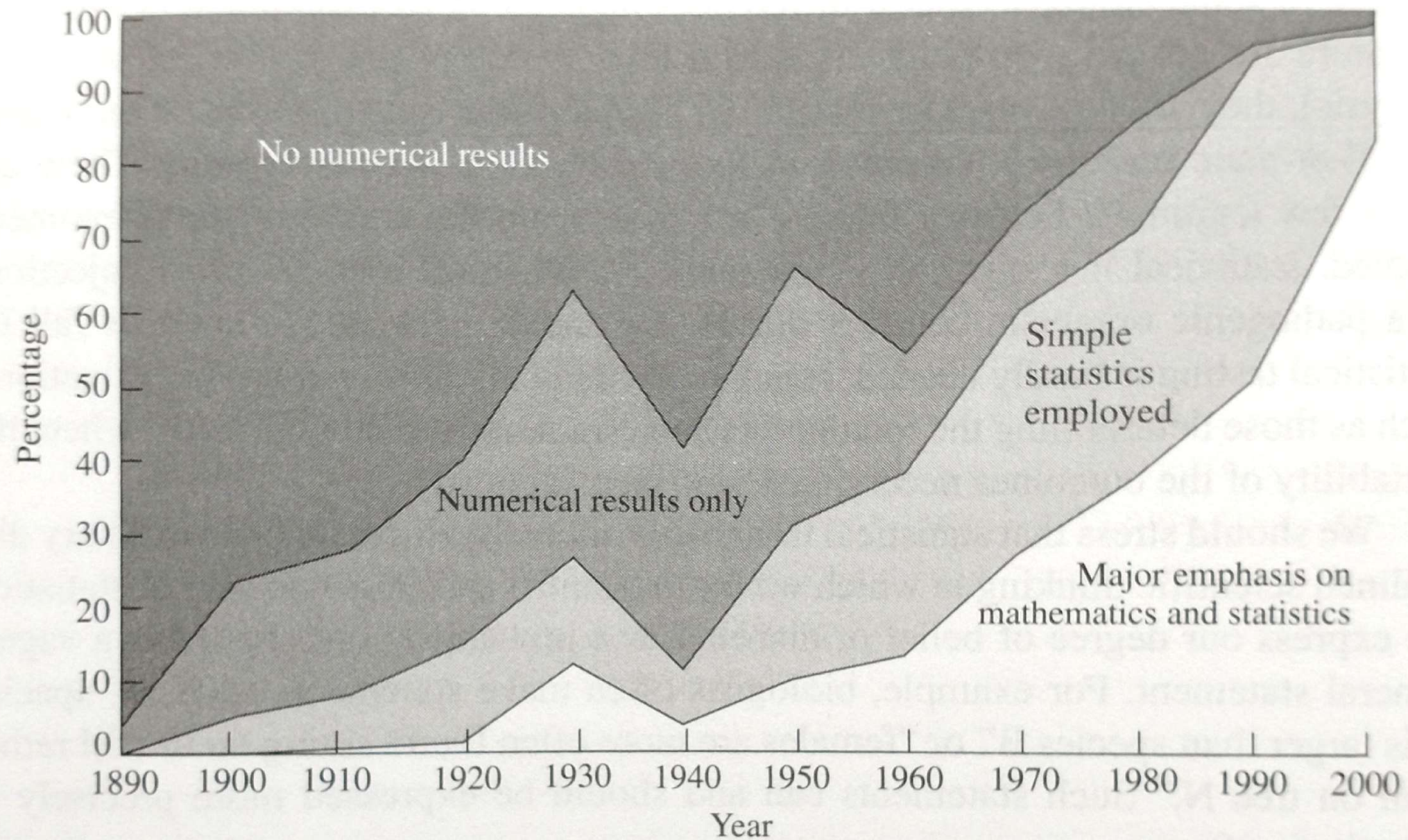


FIGURE 1.1 Proportions of articles involving numerical and statistical work in decennial issues of *The American Naturalist*.

Responda!

- Qual a aceleração com a qual um objeto cai na terra?
- Qual a massa molar do carbonato de cálcio?
- Quanto pesa uma andorinha?

Responda!

- Com qual aceleração um objeto cai na terra?
- Qual é a massa molar do carbonato de cálcio?
- Qual é a massa de uma andorinha?

À VARIAÇÃO É INERENTE À BIOLOGIA

Por que precisamos da Estatística?

Inferência

Origem: Wikipédia, a enciclopédia livre.



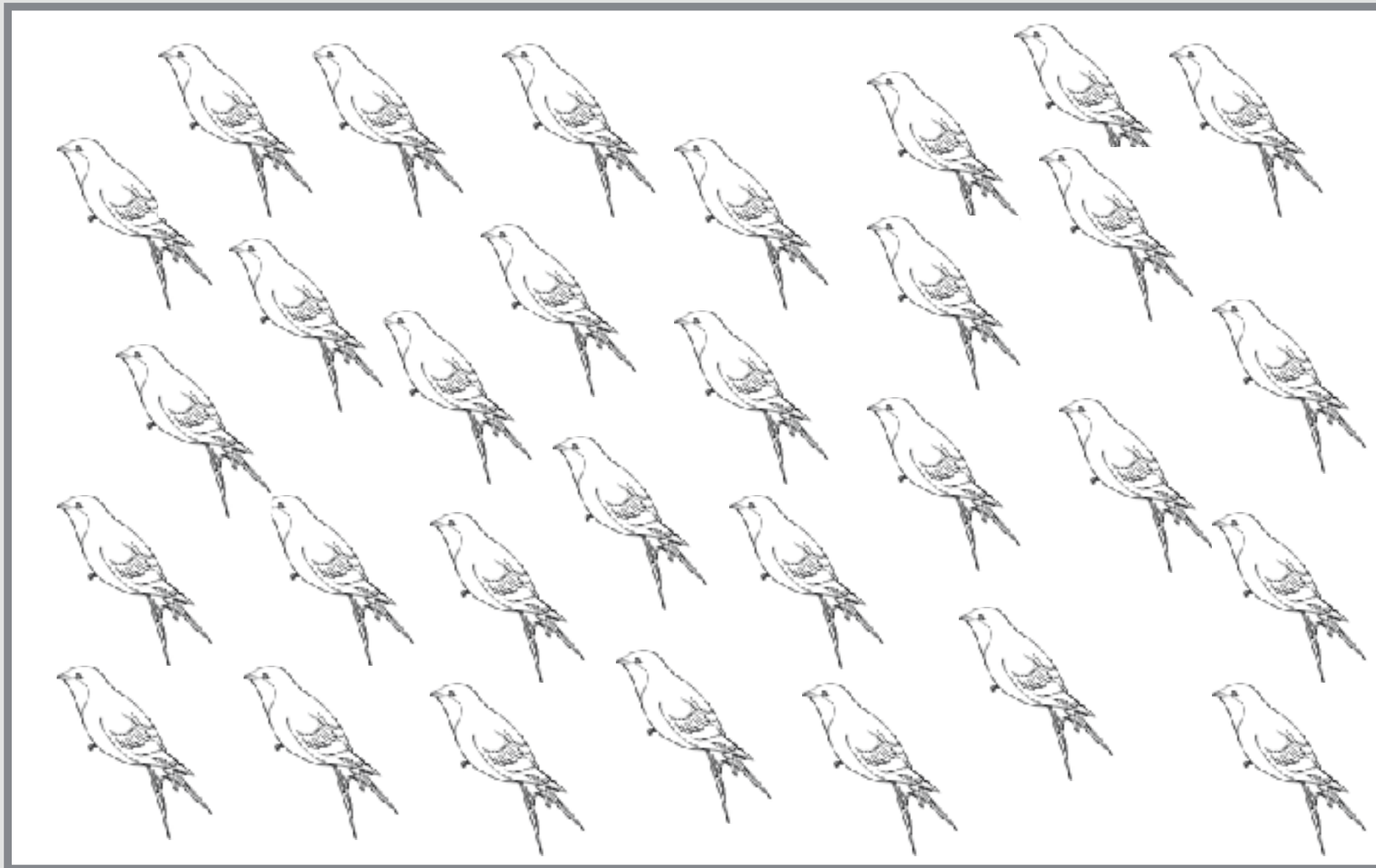
Esta página ou secção **não cita fontes confiáveis e independentes**, o que compromete sua credibilidade (desde julho de 2015). Por favor, **adicione referências** e **insira-as corretamente** no texto ou no **rodapé**. Conteúdo sem fontes poderá ser removido.

— Encontre fontes: [Google](#) (notícias, livros e acadêmico)

🔗 **Nota:** Para outros significados, veja [Inferência \(desambiguação\)](#).

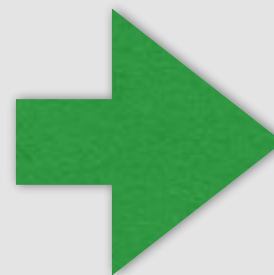
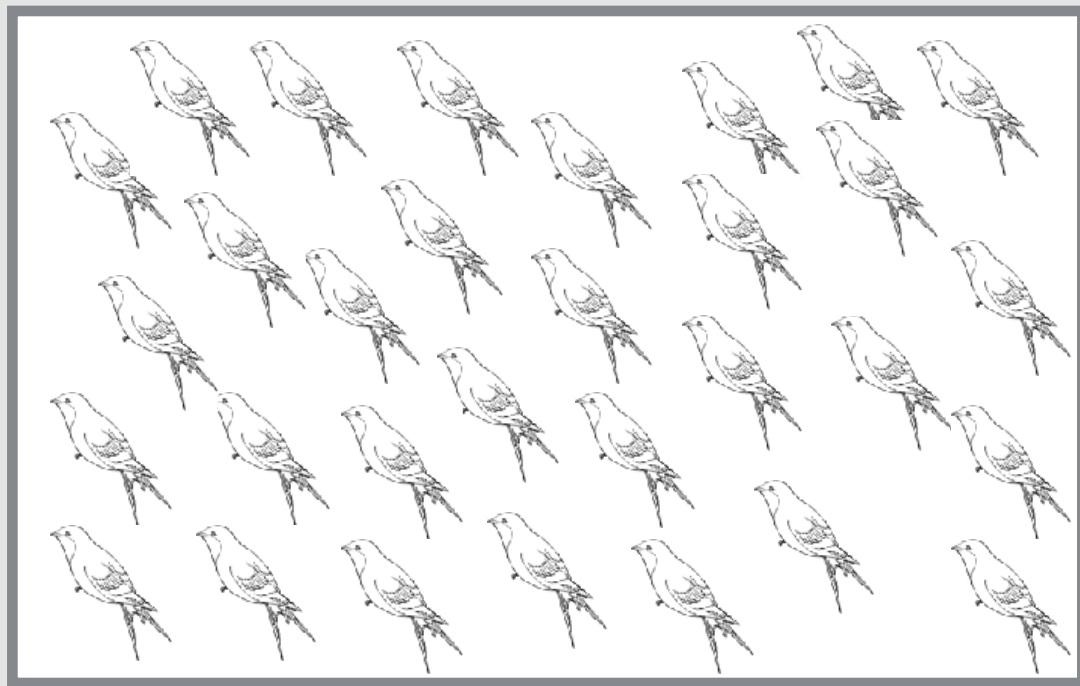
Em **Lógica**, **inferência** ou **ilação** é operação intelectual mediante a qual se afirma a verdade de uma **proposição** em decorrência de sua ligação com outras proposições já reconhecidas como verdadeiras. Consiste, portanto, em derivar conclusões a partir de **premissas** conhecidas ou decididamente verdadeiras. A conclusão também é chamada de **idiomática**.

População



População e amostra

População

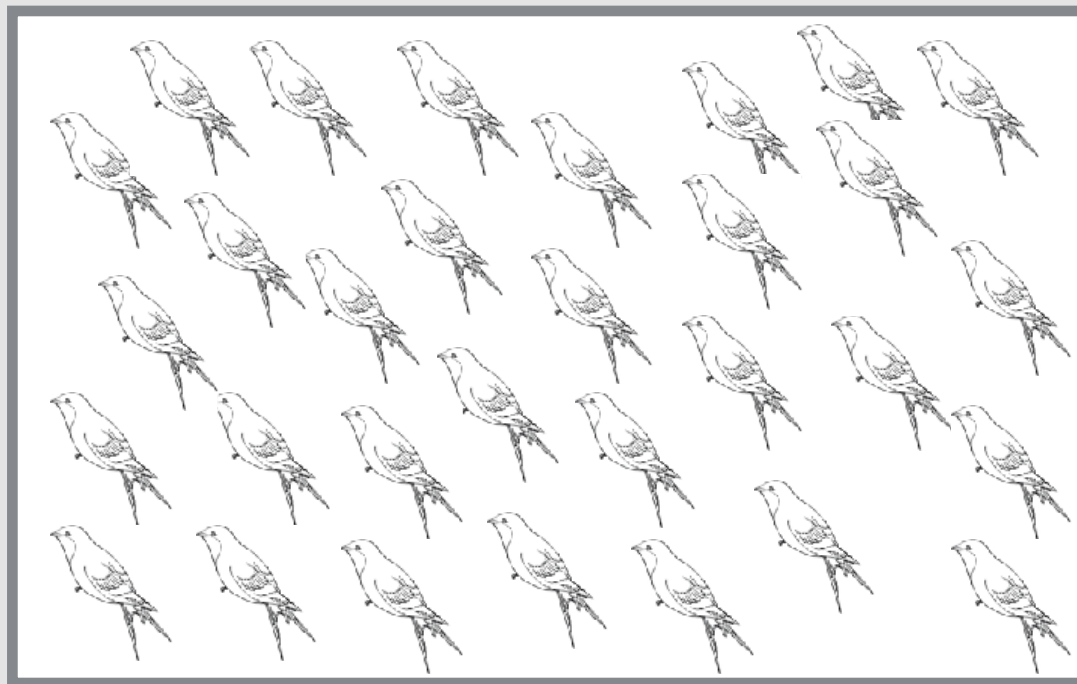


Amostra

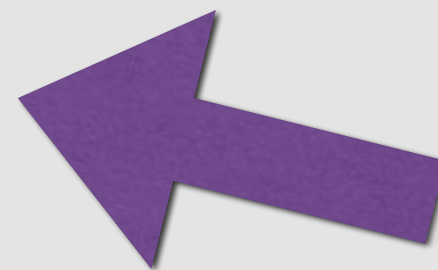
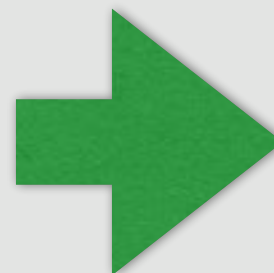
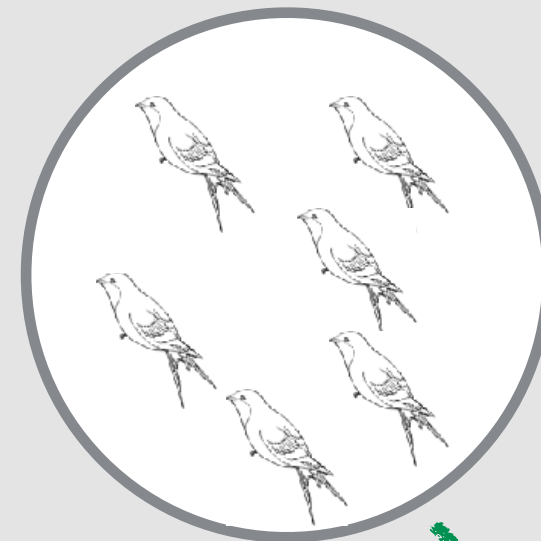


Podemos confiar no que calculamos?

População



Amostra



**Estatística de interesse
(e.g., média)**

**Inferência sobre
a população**

Previsão

Há relação entre manobrabilidade de andorinhas e o tipo de vegetação do ambiente em que vivem?

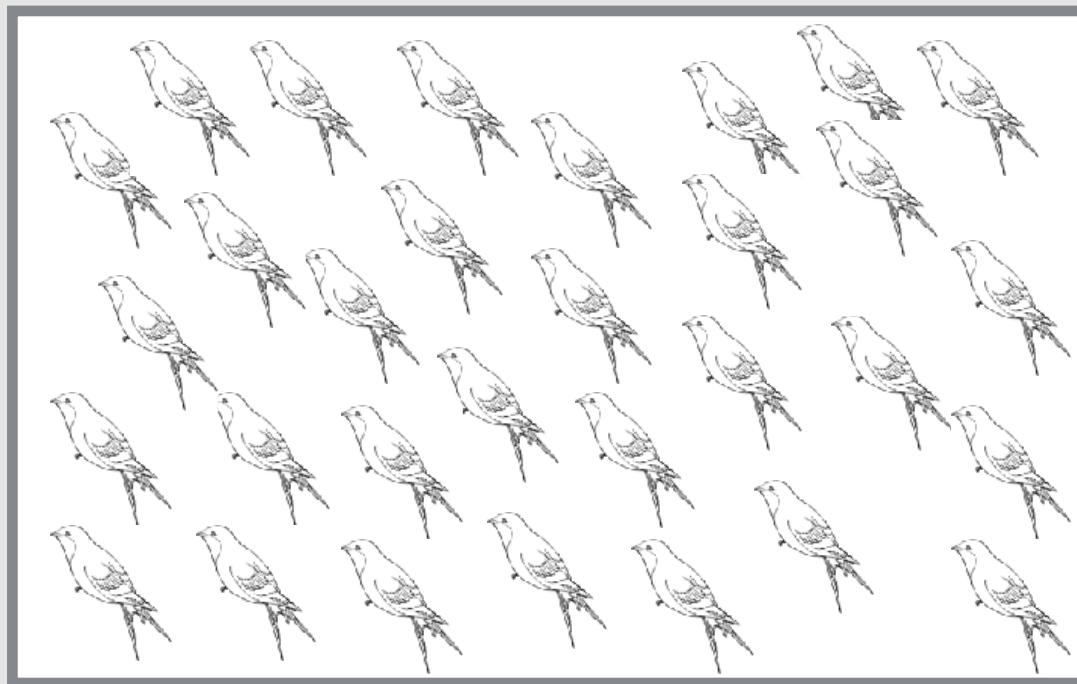
Andorinhas que vivem em ambientes com vegetação mais fechada têm maior manobrabilidade.

Andorinhas da mata atlântica têm, **em média**, menor comprimento de envergadura do que andorinhas do cerrado.

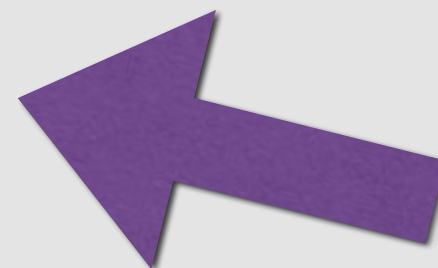
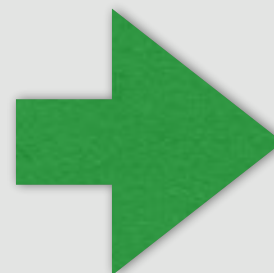
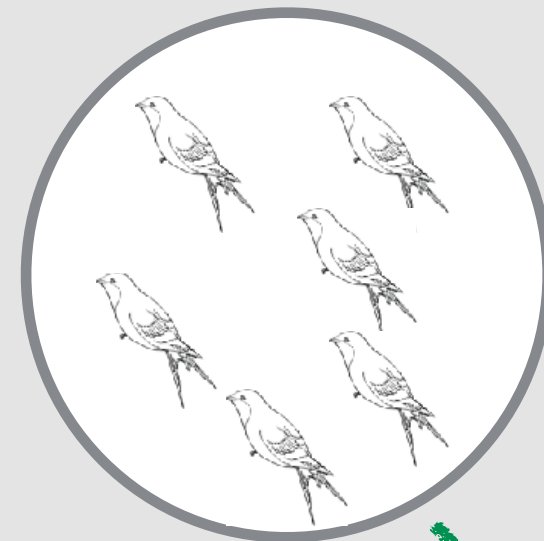
Estatística de interesse

Podemos confiar no que calculamos?

População



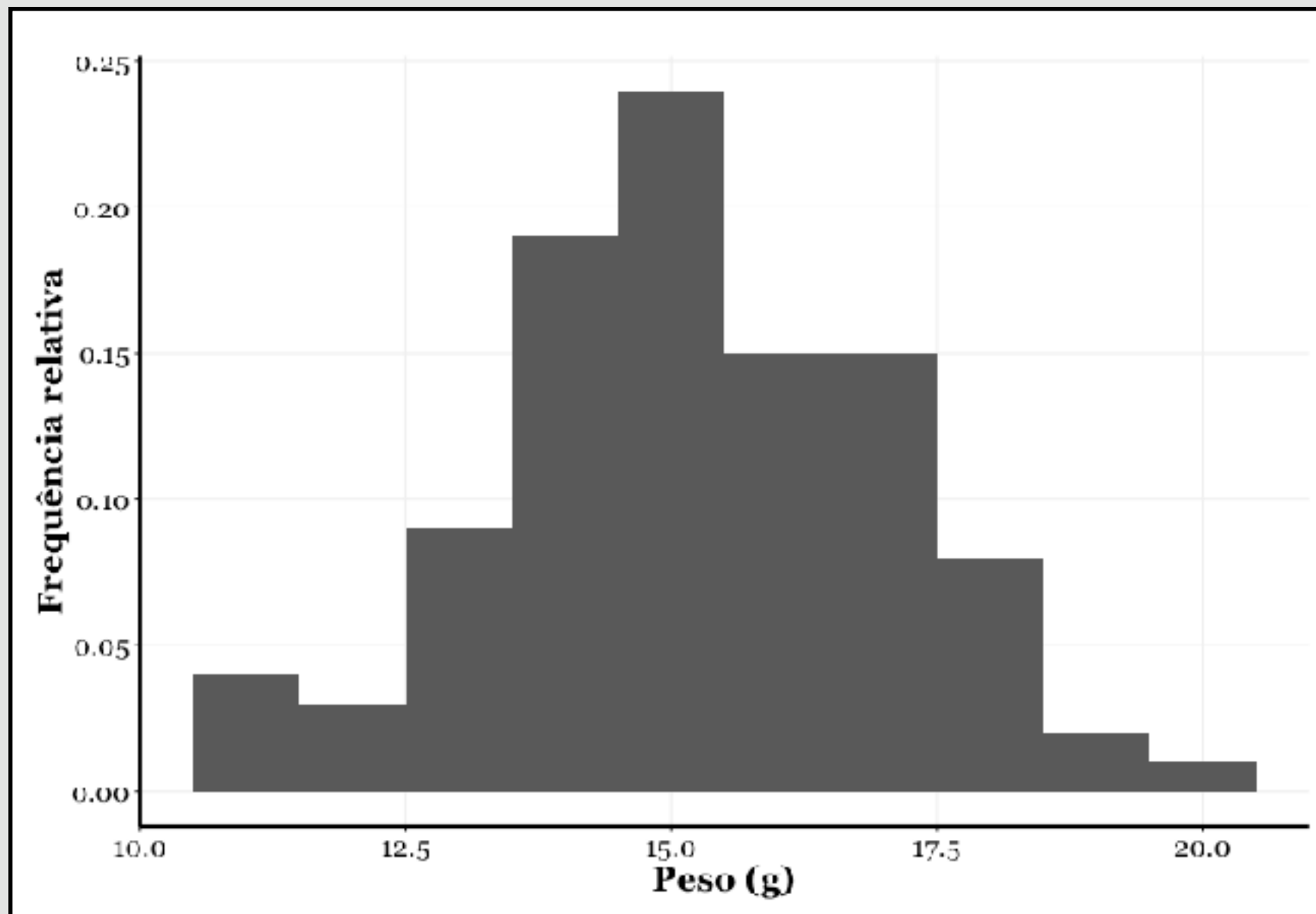
Amostra



**Inferência sobre
a população**

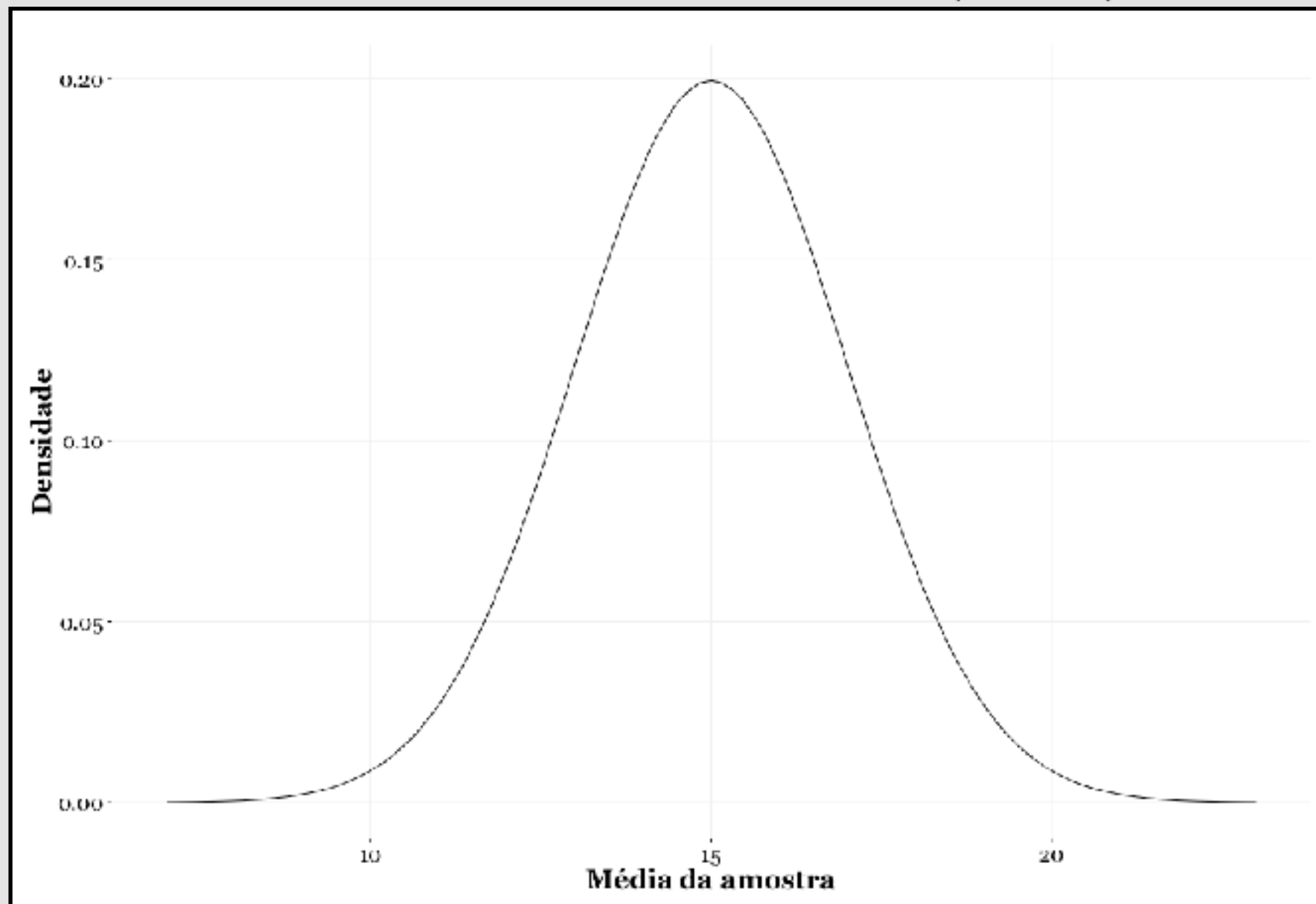
**Estatística de interesse
(e.g., média)**

Distribuição



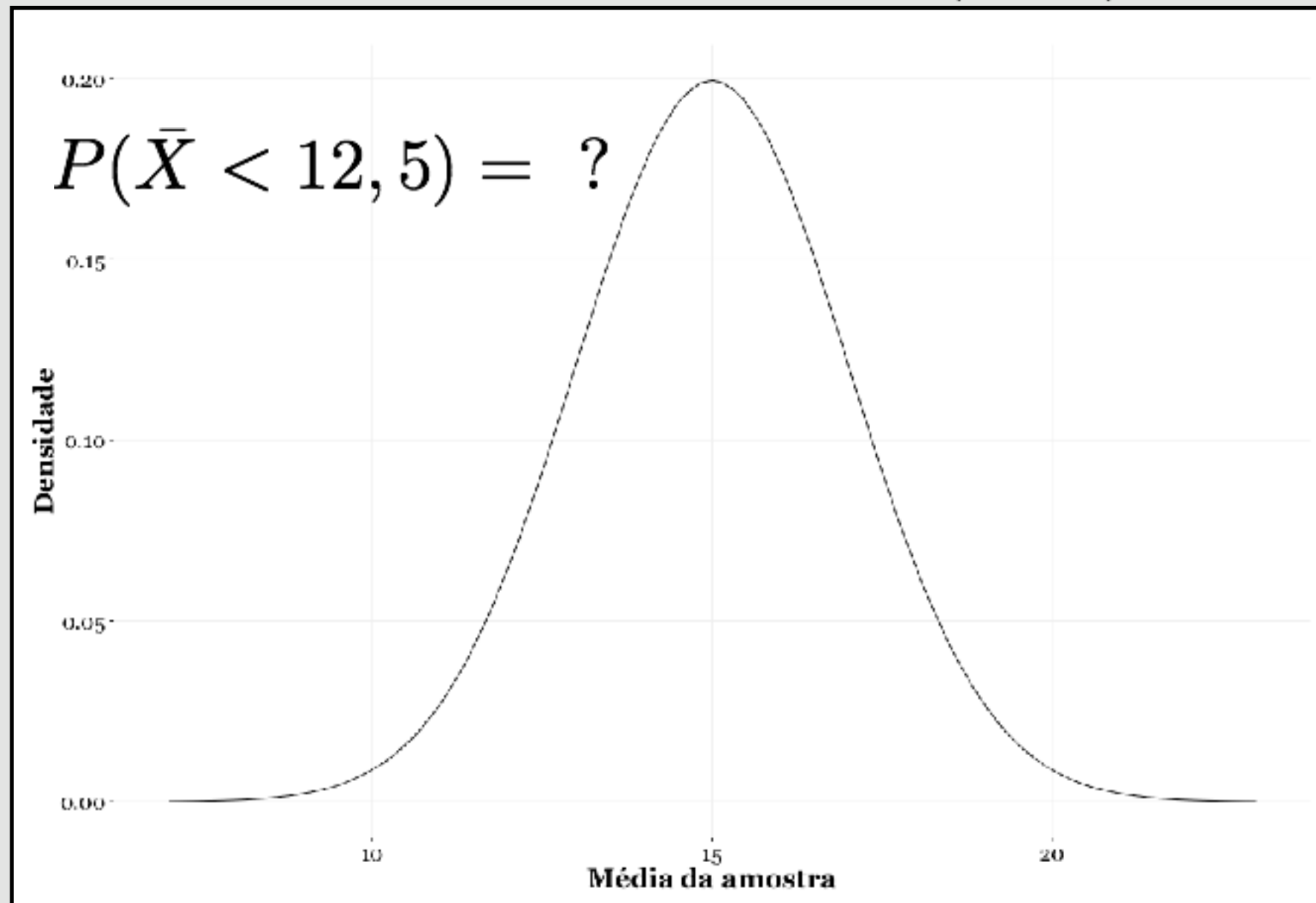
Normal

Média da amostra $\sim N(15, 2)$



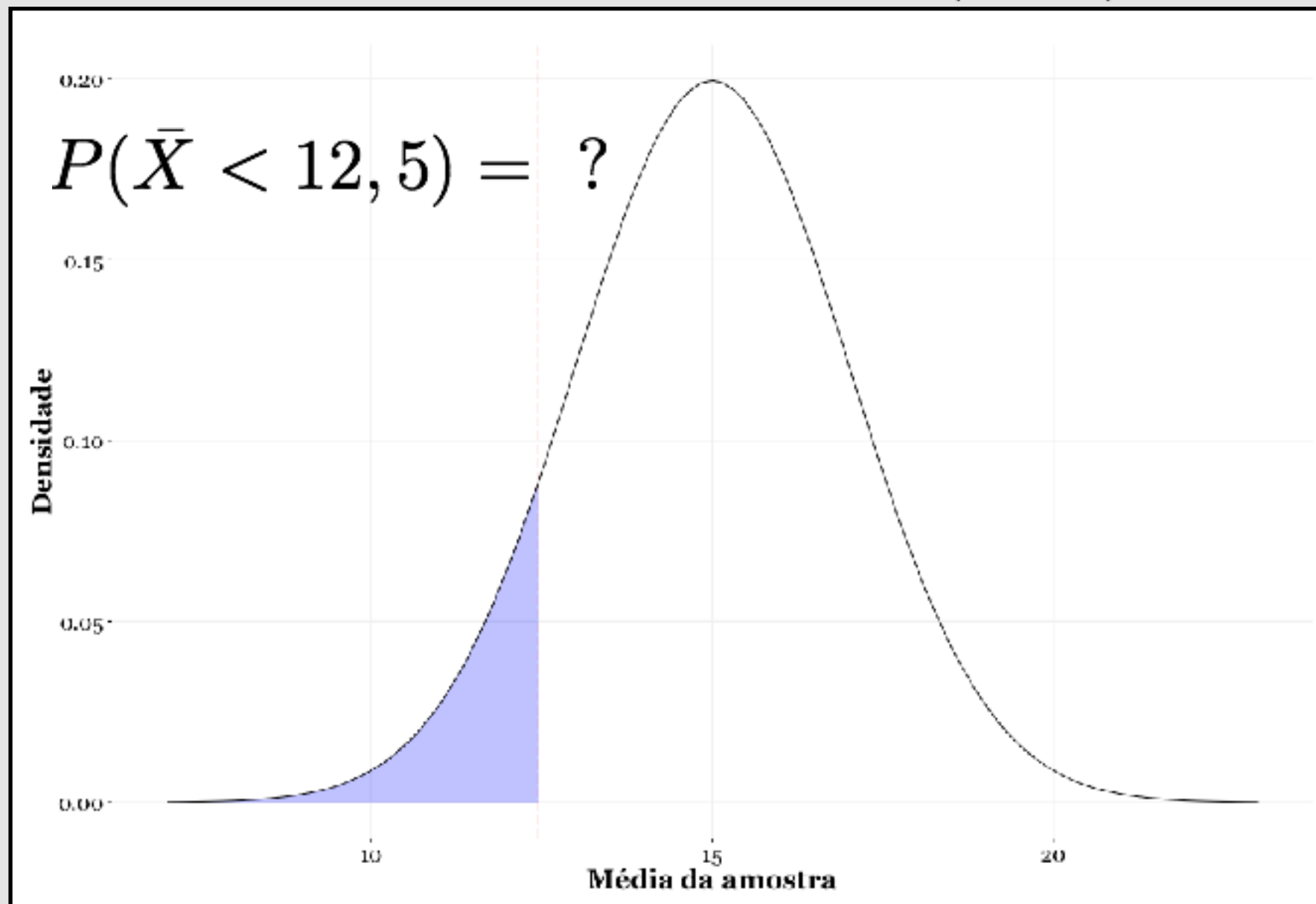
Normal

Média da amostra $\sim N(15, 2)$



Normal

Média da amostra $\sim N(15, 2)$




Teste de hipótese

1. Deduzo uma hipótese nula estatística
2. Calculo minha estatística de interesse para amostra
3. Calculo p-valor (com base numa distribuição conhecida *a priori*)
4. Decido com base em um nível de significância

E se ... ?

- Não existe ou não conheço uma distribuição conhecida para a minha estatística de interesse
- Meu N amostral é pequeno
- Alguma outra premissa dos testes paramétricos é quebrada

Resampling

 **Resampling Stats**

ABOUTCONTACT

ORDER

[Order online](#)

SOFTWARE

[Excel add-in version 4.0](#)

[Box Sampler](#)

BOOKS, ETC.

[Intro text online](#)

[Articles](#)

COURSES

[Online course](#)

TEACHING

[Teaching with RS](#)

Search

Intro Text Online

"Resampling: The New Statistics"

by Julian L. Simon

Second Edition published October 1997

This text grew out of chapters in the 1969 edition of *Basic Research Methods in Social Science* by the same author, and contains the first published example of what was later called the bootstrap. Simon is best known for his research in demography, population and the economics of natural resources, and gained fame when the noted biologist Paul Ehrlich selected five commodities and bet Simon that scarcity would drive their prices up over the period of the bet (in fact, their prices all dropped). *Resampling: The New Statistics* contains a number of examples in Resampling Stats, a computer program originated by Simon, but can be read on its own without the program.

Teste de hipótese

por Reamostragem

1. Calculo minha estatística de interesse para amostra
2. Reamostro com reposição (reordeno) os dados para construir um cenário nulo
3. Crio uma distribuição dos valores nulos
4. Calculo p-valor (com base no cenário nulo construído)
5. Decido com base em um nível de significância

Problemas

da Reamostragem

1. Dificuldade computacional
2. Variação dos resultados (inerente da reamostragem)
3. Domínio de inferência restrito
4. Demanda um cenário nulo adequado

Pergunta

Há relação entre manobrabilidade de andorinhas e o tipo de vegetação do ambiente em que vivem?

Pergunta (e hipótese)

Há relação entre manobrabilidade de andorinhas e o tipo de vegetação do ambiente em que vivem?

Andorinhas que vivem em ambientes com vegetação mais fechada têm maior manobrabilidade.

Previsão

Há relação entre manobrabilidade de andorinhas e o tipo de vegetação do ambiente em que vivem?

Andorinhas que vivem em ambientes com vegetação mais fechada têm maior manobrabilidade.

Andorinhas da mata atlântica têm menor comprimento de envergadura do que andorinhas do cerrado.

Previsão

Há relação entre manobrabilidade de andorinhas e o tipo de vegetação do ambiente em que vivem?

Andorinhas que vivem em ambientes com vegetação mais fechada têm maior manobrabilidade.

Andorinhas da mata atlântica têm menor comprimento de envergadura do que andorinhas do cerrado.

Operacionalização das variáveis

Previsão

Há relação entre manobrabilidade de andorinhas e o tipo de vegetação do ambiente em que vivem?

Andorinhas que vivem em ambientes com vegetação mais fechada têm maior manobrabilidade.

Andorinhas **da mata atlântica** têm menor **comprimento de envergadura** do que andorinhas **do cerrado**.

Variável preditora vs. Variável resposta

Exemplo 1 - Manguezal e estabilidade

- Pergunta

O investimento em raízes de suporte por árvores de manguezal depende da estabilidade do solo em que se encontram?

- Hipótese

Árvores de manguezal em solos mais instáveis investem mais em raízes de suporte



Exemplo 1 - Manguezal e estabilidade

- Pergunta

O investimento em raízes de suporte por árvores de manguezal depende da estabilidade do solo em que se encontram?

- Hipótese

Árvores de manguezal em solos mais instáveis investem mais em raízes de suporte

- Previsão

A área coberta por raízes será maior, em média, em uma área que fica mais tempo inundada do que em uma área que fica menos tempo inundada.

Quais são as variáveis teóricas e quais são as operacionais? Qual é a variável preditora e qual é resposta?

Exemplo 1 - Manguezal e estabilidade

| Área | Tipo de Solo |
|------|--------------|
| 100 | Drenado |
| 150 | Drenado |
| 200 | Inundado |
| 250 | Inundado |

Média (Drenado) = 125
Média (Inundado) = 225

Estatística de interesse?

Diferença entre médias!
 $= 125 - 225 = -100$

Exemplo 1 - Manguezal e estabilidade

| Área | Tipo de Solo |
|------|--------------|
| 100 | Drenado |
| 150 | Inundado |
| 200 | Drenado |
| 250 | Inundado |

Embaralhar!

Média (Drenado) = 150
Média (Inundado) = 200

Estatística de interesse?

Diferença entre médias!
 $= 150 - 200 = -50$

Exemplo 1 - Manguezal e estabilidade

| Área | Tipo de Solo |
|------|--------------|
| 100 | Inundado |
| 150 | Inundado |
| 200 | Drenado |
| 250 | Drenado |

Embaralhar de novo!

Média (Drenado) = 225
Média (Inundado) = 125

Estatística de interesse?

Diferença entre médias!
 $= 225 - 125 = 100$

Exemplo 1 - Manguezal e estabilidade

| Área | Tipo de Solo |
|------|--------------|
| 100 | Inundado |
| 150 | Inundado |
| 200 | Drenado |
| 250 | Drenado |

Embaralhar de novo!

Média (Drenado) = 225
Média (Inundado) = 125

Estatística de interesse?

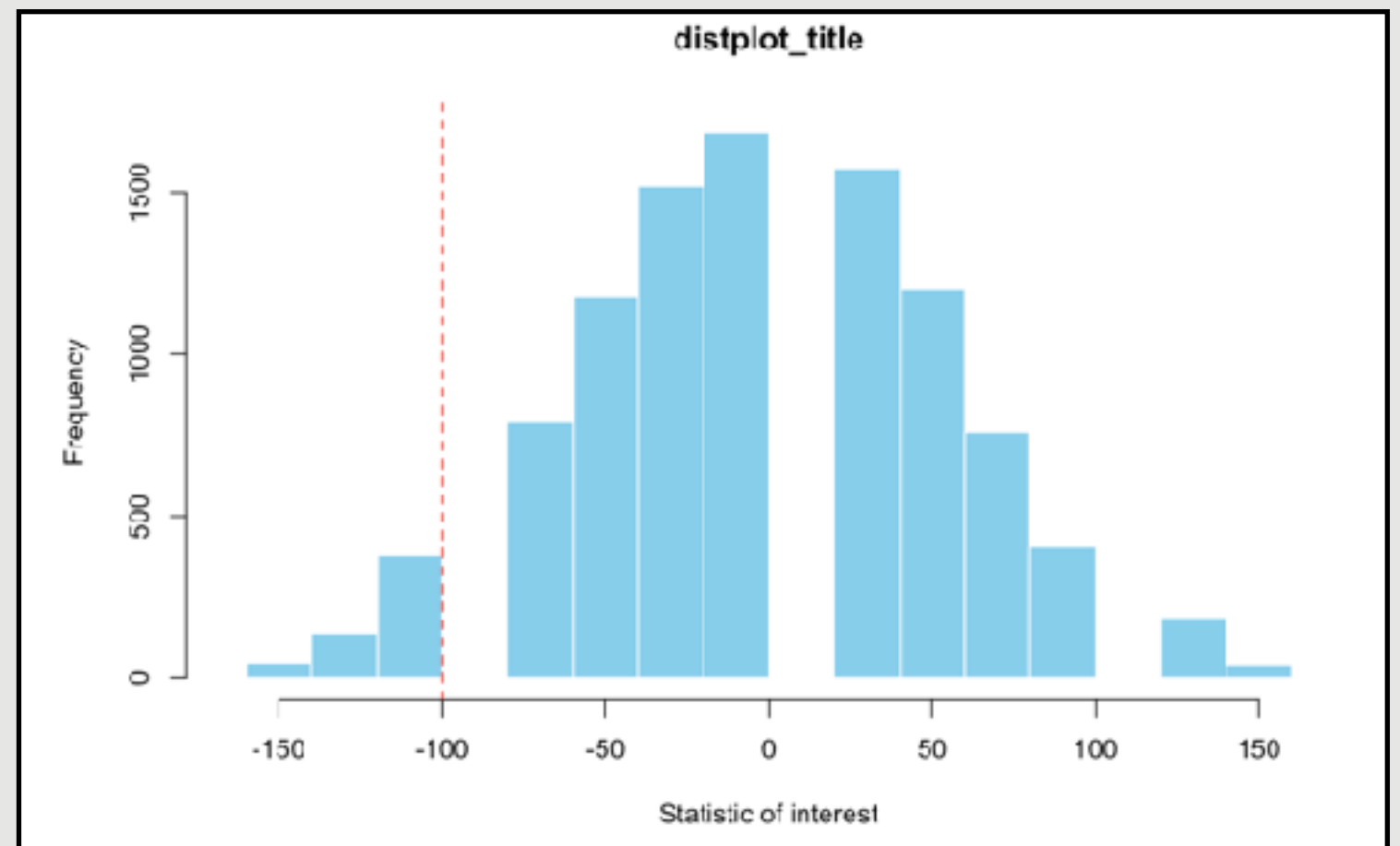
Diferença entre médias!
 $= 225 - 125 = 100$

Repetir muitas vezes!

Exemplo 1 - Manguezal e estabilidade

Diferença entre médias observada = -100

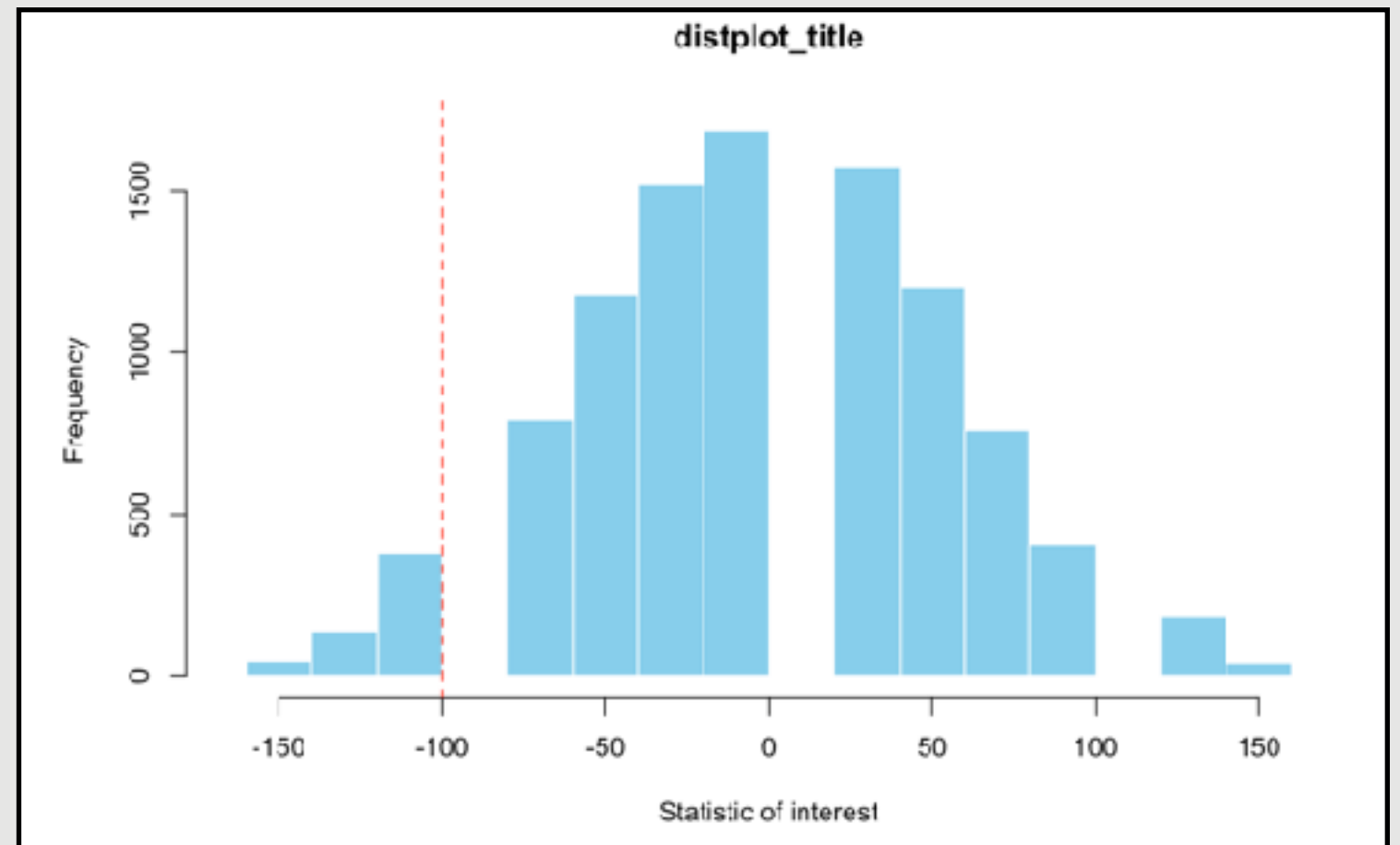
| Área | Tipo de Solo |
|------|--------------|
| 100 | Inundado |
| 150 | Inundado |
| 200 | Drenado |
| 250 | Drenado |



Exemplo 1 - Manguezal e estabilidade

Diferença entre médias observada = -100

| Área | Tipo de Solo |
|------|--------------|
| 100 | Inundado |
| 150 | Inundado |
| 200 | Drenado |
| 250 | Drenado |

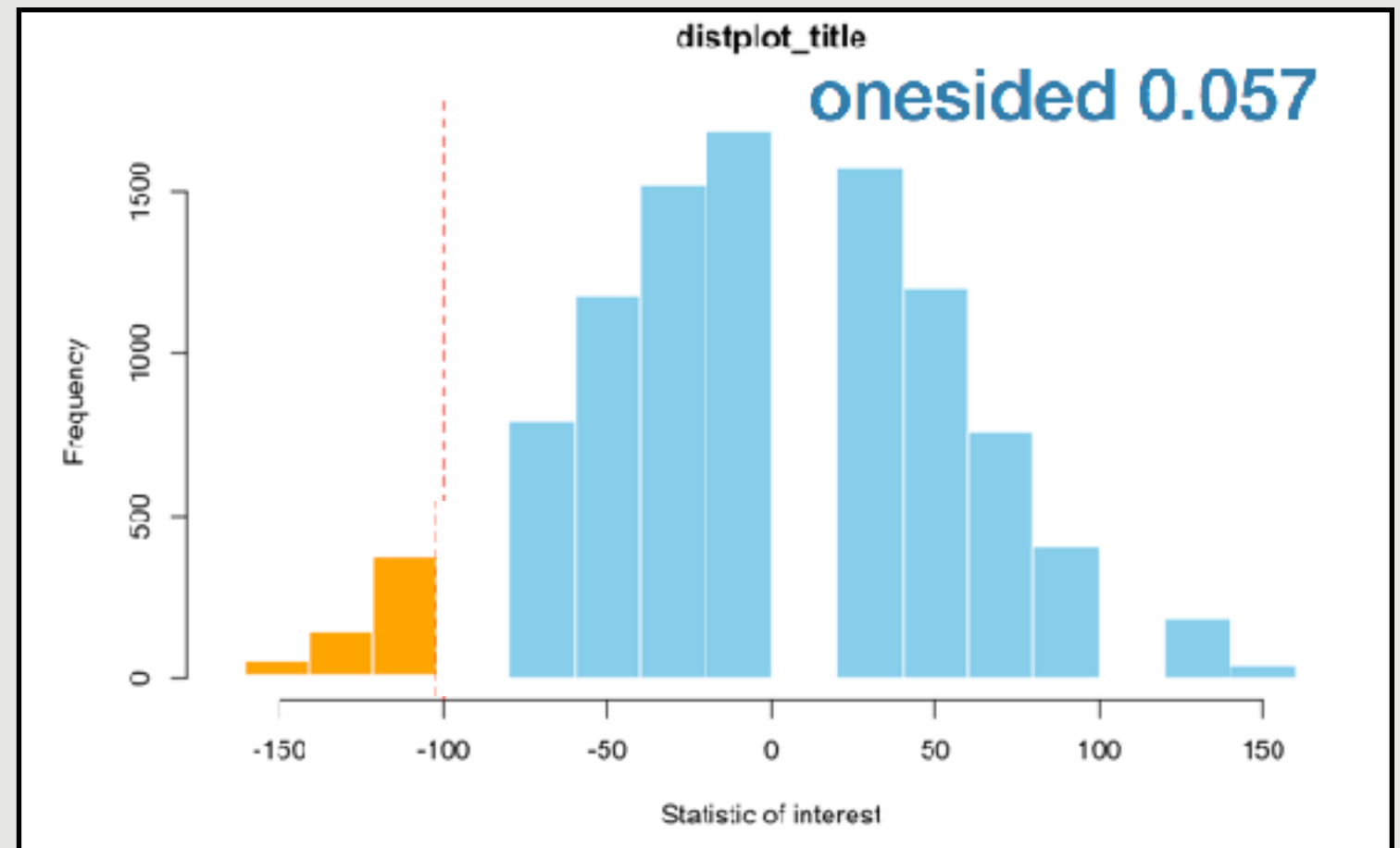


- **Previsão**
A **área coberta por raízes** será maior, em média, em uma área que fica **mais tempo inundada** do que em uma área que fica menos tempo inundada.

Exemplo 1 - Manguezal e estabilidade

Diferença entre médias observada = -100

| Área | Tipo de Solo |
|------|--------------|
| 100 | Inundado |
| 150 | Inundado |
| 200 | Drenado |
| 250 | Drenado |



- **P-valor**
A probabilidade de obter uma estatística de interesse igual ou mais extrema que a observado sob a hipótese nula

Exemplo 2 - Mangues equilibristas

- Pergunta

O torque causado pelo peso da copa resulta em árvores de mangue com mais raízes?

- Hipótese

Quanto maior o torque causado pela copa, mais raízes a árvore terá



Exemplo 3 - Formigas protetoras

- Pergunta

Formigas respondem mais intensamente a dano em folhas mais jovens?

- Hipótese

Dano nas folhas velhas leva a menos formigas recrutadas quando comparado a dano nas folhas jovens



Considerações finais



Teste de hipótese

por Reamostragem

1. Calculo minha estatística de interesse para amostra
2. Reamostro com reposição (reordeno) os dados para construir um cenário nulo
3. Crio uma distribuição dos valores nulos
4. Calculo p-valor (com base no cenário nulo construído)
5. Decido com base em um nível de significância

Agradecimentos e Créditos

- Ayana Martins
- Paulo Inácio Prado
- Alexandre Adalardo de Oliveira
- Comissão da EcoEscola

