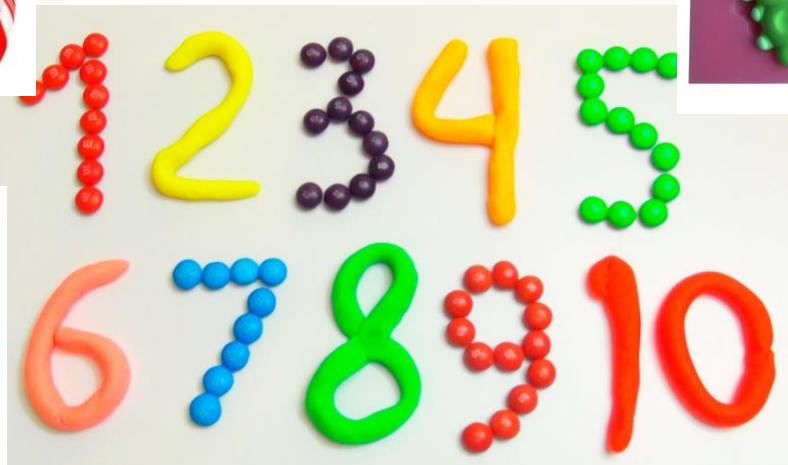


# Goodies\*



\* Goodies related to animals, plants and numbers... and toast...!

Se uma torrada com manteiga ou doce de um lado cair ao chão, qual é a probabilidade de ela cair com a face barrada virada para baixo?

A lógica diz... 0.5... mas será mesmo?



I never had a slice of bread,  
Particularly large and wide,  
That did not fall upon the floor,  
And always on the buttered side!

James Payn, 1884

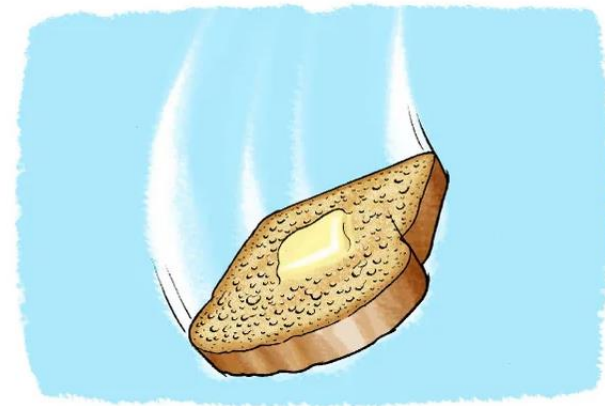
[https://en.wikipedia.org/wiki/Buttered\\_toast\\_phenomenon](https://en.wikipedia.org/wiki/Buttered_toast_phenomenon)

- Randomness
- Physics
- Perception (filters...)

# Falling Buttered Toast

An unfortunate Science Activity from Science Buddies

By Science Buddies on February 16, 2017



READ TH

POLICY & ET  
The Silenced:  
Muzzled by th  
21 minutes ago — C

CLIMATE  
A New Kind c  
September 17, 2015

BIOLOGY  
A Newly Iden  
Vanquishing  
1 hour ago — Simo

<https://www.scientificamerican.com/article/falling-buttered-toast/>

**lifestyle** food > eat

## Why your toast falls butter side down

IT'S a sticky issue that has plagued scientists - and breakfast tables - since 1835. But now researchers have found the answer as to why a piece of toast always seems to fall buttered side down.

NICK MCDERMOTT

Daily Mail • SEPTEMBER 5, 2013 12:59PM



advertisement

<https://www.news.com.au/lifestyle/food/why-your-toast-falls-butter-side-down/news-story/15852fc2a1b6bcfe3b748375a9cc1edf>

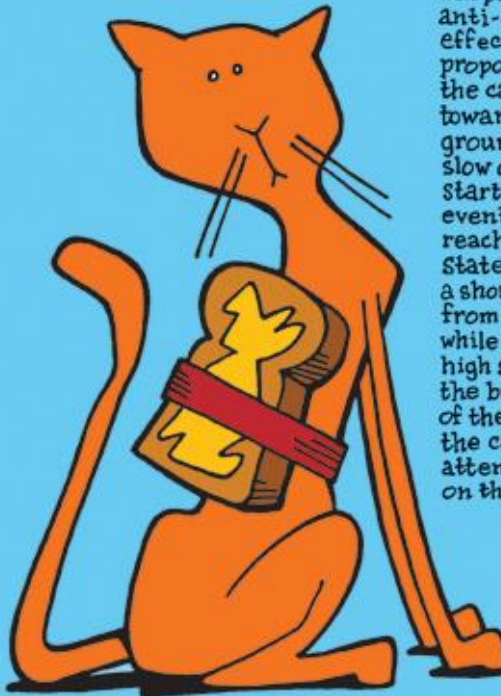
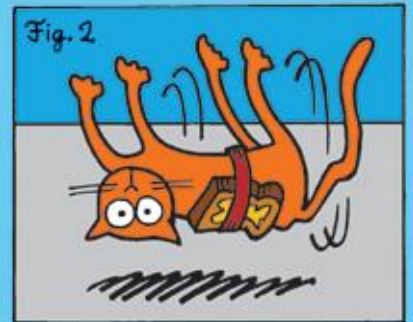
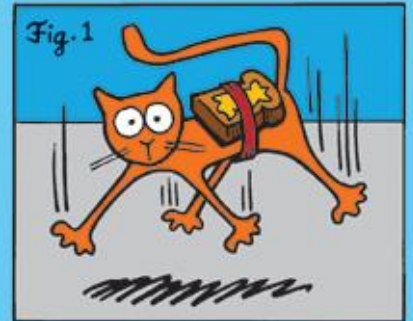
# BUTTERED CAT PARADOX

The buttered cat paradox is a paradox based on the tongue-in-cheek combination of two bits of folk wisdom:

- Cats always land on their feet.
- Buttered toast always lands buttered side down.

The paradox arises when one considers what would happen if one attached a piece of buttered toast (butter side up) to the back of a cat, then dropped the cat from a height.

Some people jokingly maintain that the experiment will produce an anti-gravity effect. They propose that as the cat falls towards the ground, it will slow down and start to rotate, eventually reaching a steady state of hovering a short distance from the ground while rotating at high speed as both the buttered side of the toast and the cat's feet attempt to land on the ground.





An approximate answer to the right  
problem is worth a good deal more  
than an exact answer to an  
approximate problem.

— *John Tukey* —

AZ QUOTES

[https://www.azquotes.com/author/14847-John\\_Tukey](https://www.azquotes.com/author/14847-John_Tukey)



Ciências  
ULisboa

**SPRINGER NATURE**

## Acesso aos eBooks da Springer Nature

**Até 30 de junho de 2020**, a comunidade de Ciências ULisboa tem acesso a duas coleções de eBooks da Springer Nature:

- Mathematics & Statistics;
- Earth & Environmental Sciences.

Para aceder, basta consultar a página da Biblioteca em <https://ciencias.ulisboa.pt/pt/ebooks>.

Durante o período acima mencionado, toda a comunidade de Ciências ULisboa pode consultar e fazer *download* dos livros que quiser, livremente.

A equipa da Biblioteca está ao dispor para qualquer esclarecimento adicional.

**Divulgado por:**

Faculdade de Ciências da Universidade de Lisboa

Faculty of Sciences of the University of Lisbon

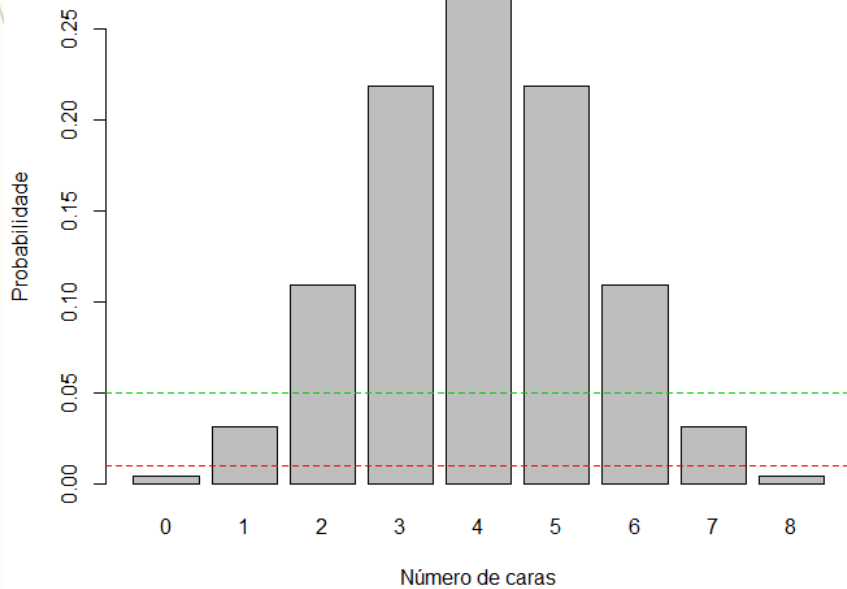
Campo Grande, 1749-016 Lisboa – Portugal

Gabinete de Organização e Gestão de Informação

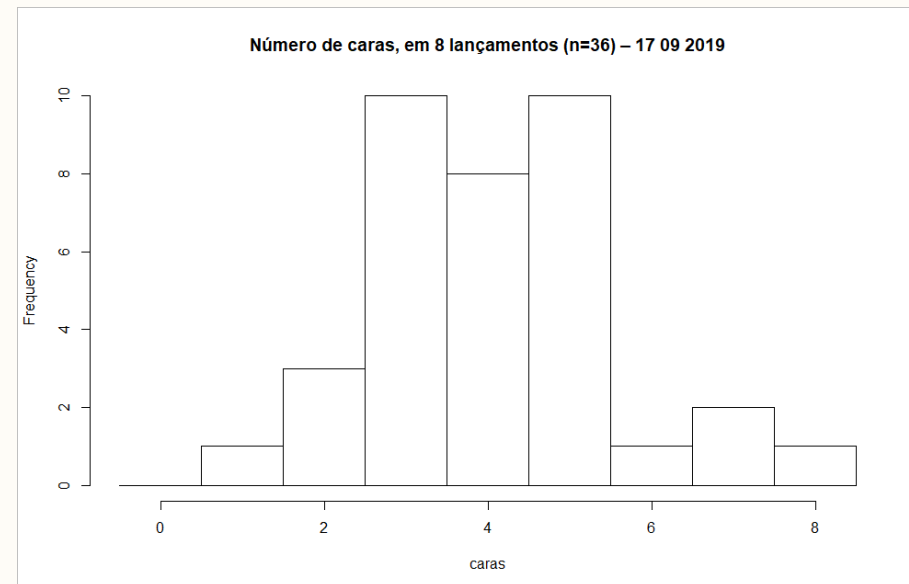
Tel. +351 217 500 000 - Ext: 25357 | 25302

[www.ciencias.ulisboa.pt](http://www.ciencias.ulisboa.pt)

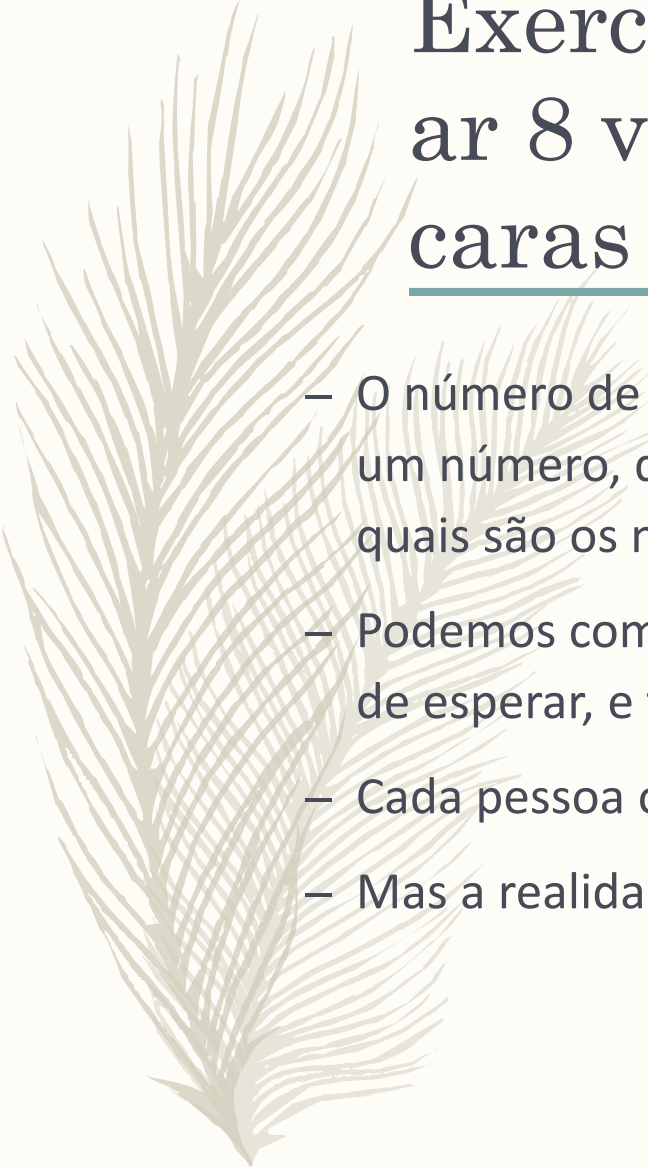
# Se eu lançar uma moeda e tirar 8 coroas seguidas, a minha moeda é viciada?



Distribuição exata



Amostra em sala de aula



# Exercício: lançar uma moeda ao ar 8 vezes, e contar o número de caras

---

- O número de caras é uma “variável aleatória” (cada pessoa obtém um número, desconhecido à partida, mas sabemos sem duvida quais são os números possíveis de observar)
- Podemos comparar com o número de caras obtido com o que seria de esperar, e tirar conclusões
- Cada pessoa obtém um resultado
- Mas a realidade é só uma: as moedas não são viciadas (espero eu!)



# Testes, decisões e erros

		Decisão	
		Moeda normal	Moeda viciada
Realidade	Moeda normal	<b>Decisão certa</b>	<b>Erro I</b>
	Moeda Viciada	<b>Erro II</b>	<b>Decisão certa</b>

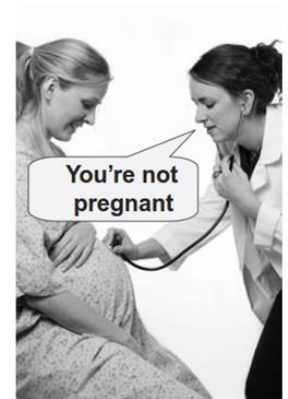
Na realidade a questão sobre qual é pior, um erro de tipo I ou II, é puramente filosófica, as opiniões podem divergir, e só pode ser avaliada caso a caso.

Em estatística, considera-se o erro de tipo I pior que o erro de tipo II.

**Type I error**  
(false positive)



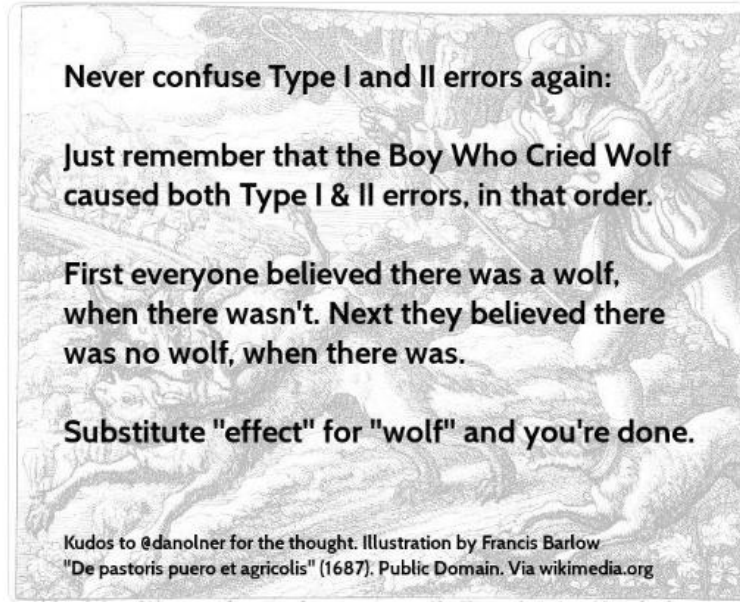
**Type II error**  
(false negative)



**Hugh Kearns**  
@ithinkwellHugh

Follow

Do you get confused by Type I and Type II errors.  
Remember the boy who cried wolf. Type I and Type II errors in order.  
[#PhDchat](#) [#ECRchat](#) [#PhDforum](#) [#MRES](#)



12:30 PM - 4 Jul 2019

160 Retweets 464 Likes

6 160 464

Tweet your reply

**Tiago André Marques** @TiagoALOMarques · now  
Replying to @ithinkwellHugh  
Really useful analogy and mnemonic. Will use it in class!



**Hugh Kearns**  
@ithinkwellHugh  
Procrastination, Perfectionism, Overcommitment and the Imposter Syndrome. Flinders University, Adelaide, iThinkWell. Lecturer, researcher and author.  
[Adelaide, South Australia](#)  
[facebook.com/ithinkwellhugh/](https://facebook.com/ithinkwellhugh/)  
Joined August 2015

Uma analogia legal...

É pior  
colocar um inocente na cadeia  
do que  
deixar um criminoso em liberdade?

Uma pessoa é considerada inocente até prova em contrário

*Beyond reasonable doubt...*

Na estatística as decisões fortes são aquelas que permitem rejeitar uma hipótese nula, em todos os outros casos podemos apenas não ter conseguido reunir provas suficientes para o fazer.

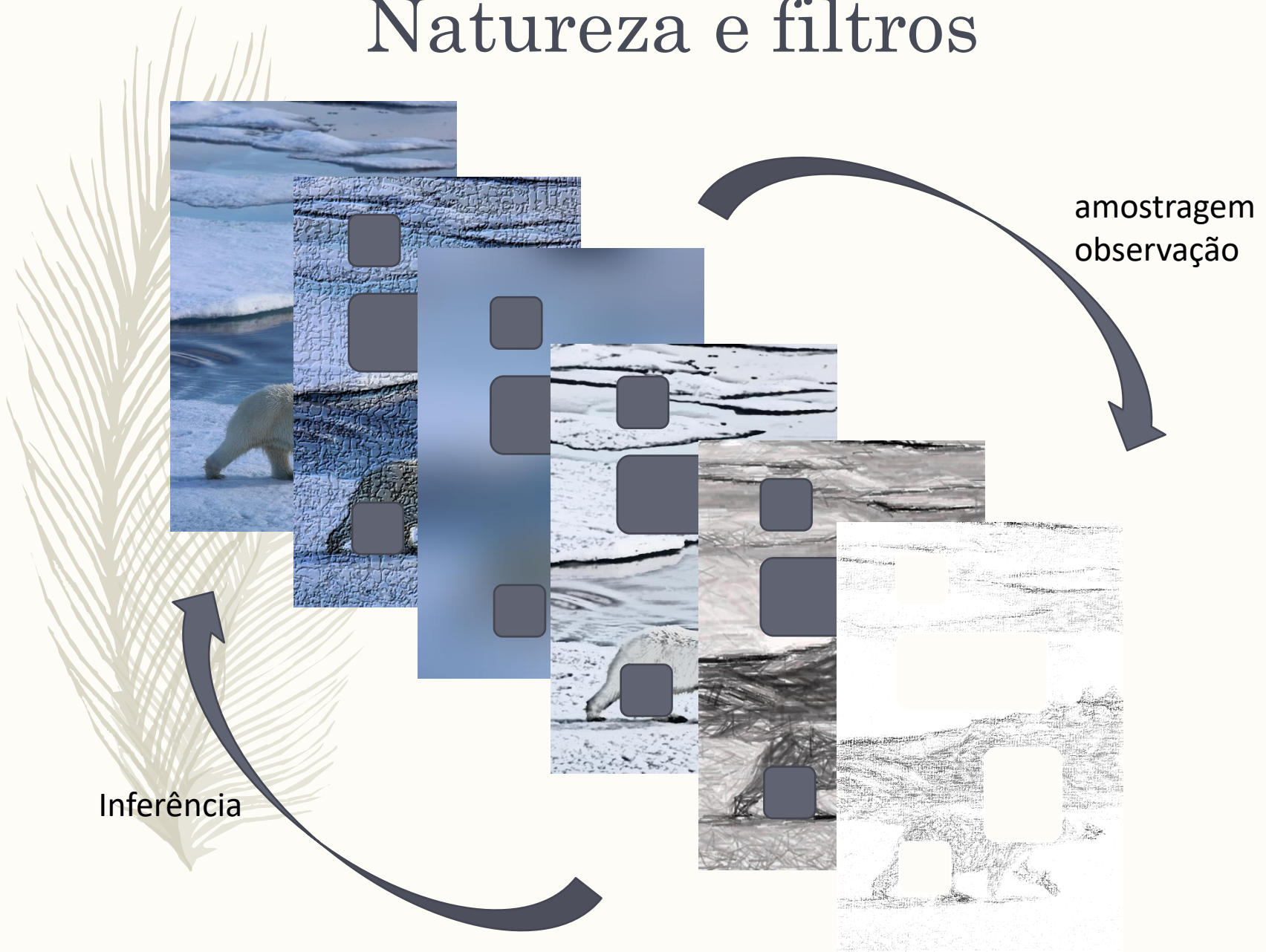
# Natureza e filtros



amostragem  
observação

Inferência

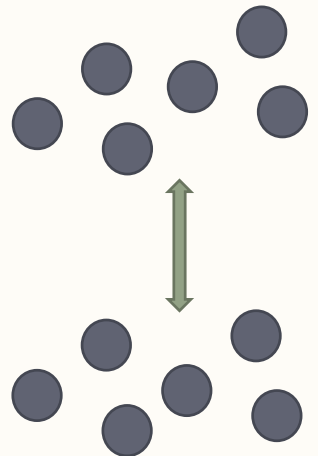
# Natureza e filtros



# A noção de variabilidade

---

- Perceber e compreender a variabilidade natural dos fenómenos biológicos é fundamental para poder dizer se uma observação é expectável (ou não!).
- Se eu quiser comparar um local no Alentejo com um local no Minho, em termos do número de espécies de plantas, qual é a variabilidade que está envolvida?
  - Temporal
  - Espacial
  - Replicação e pseudo-replicação



# Replicação e pseudoreplicação



- Se eu medir um par de folhas de uma mesma árvore, uma de exposição norte e outra de exposição sul, posso ver se norte e sul são diferentes?
- E se o fizer com 50 folhas diferentes?
- E se usar 50 folhas de 50 árvores diferentes?
- Outros exemplos possíveis?

*Ecological Monographs* 54(2), 1984, pp. 187-211  
© 1984 by the Ecological Society of America

## PSEUDOREPLICATION AND THE DESIGN OF ECOLOGICAL FIELD EXPERIMENTS

STUART H. HURLBERT  
*Department of Biology, San Diego State University,  
San Diego, California 92182 USA*

*Abstract.* Pseudoreplication is defined, as the use of inferential statistics to test for treatment effects with data from experiments where either treatments are not replicated (though samples may be) or replicates are not statistically independent. In ANOVA terminology, it is the testing for treatment

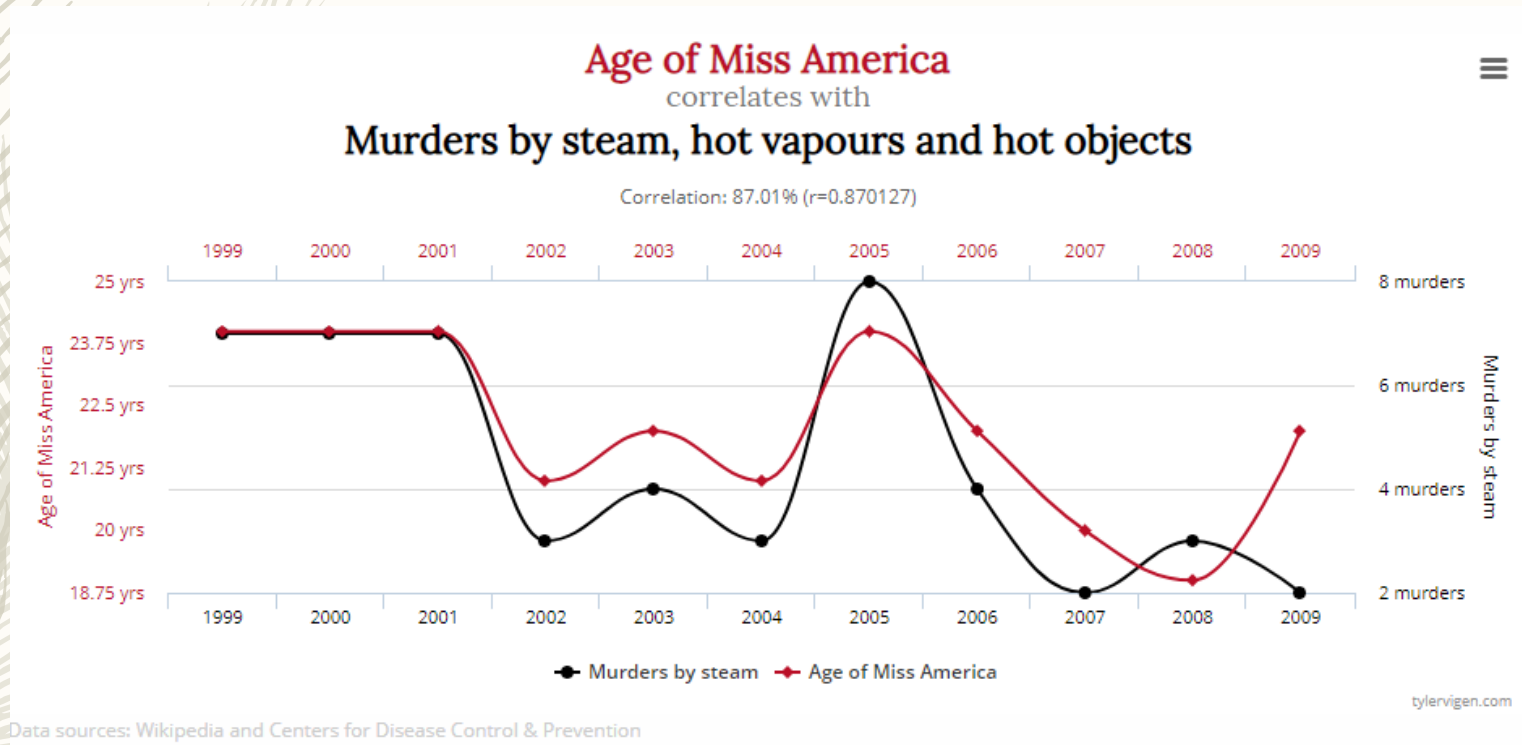
# Outros pecados capitais

---





# Correlação e causalidade



# Perigos da extrapolação

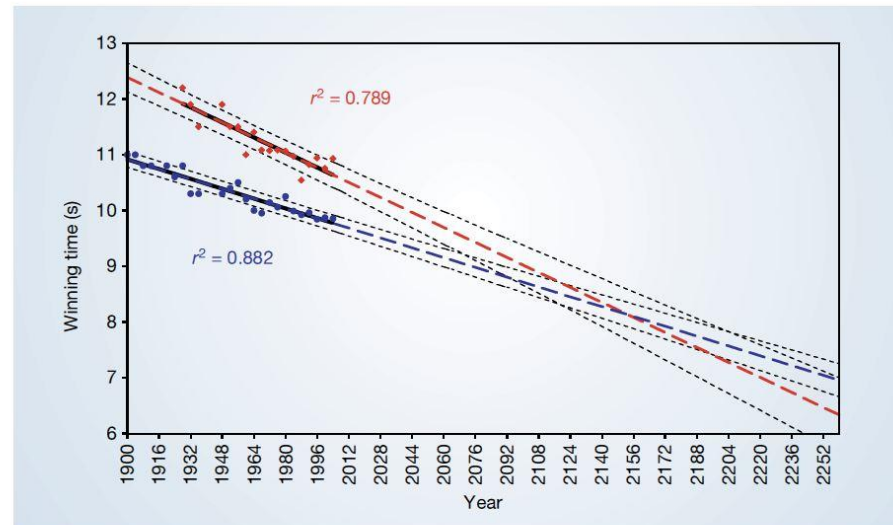
## Momentous sprint at the 2156 Olympics?

Women sprinters are closing the gap on men and may one day overtake them.

The 2004 Olympic women's 100-metre sprint champion, Yuliya Nesterenko, is assured of fame and fortune. But we show here that — if current trends continue — it is the winner of the event in the 2156 Olympics whose name will be etched in sporting history forever, because this may be the first occasion on which the race is won in a faster time than the men's event.

The Athens Olympic Games could be viewed as another giant experiment in human athletic achievement. Are women narrowing the gap with men, or falling further behind? Some argue that the gains made by women in running events between the 1930s and the 1980s are decreasing as the women's achievements plateau<sup>1</sup>. Others contend that there is no evidence that athletes, male or female, are reaching the limits of their potential<sup>1,2</sup>.

In a limited test, we plot the winning times of the men's and women's Olympic finals over the past 100 years (ref. 3; for data set, see sup-



**Figure 1** The winning Olympic 100-metre sprint times for men (blue points) and women (red points), with superimposed best-fit linear regression lines (solid black lines) and coefficients of determination. The regression lines are extrapolated (broken blue and red lines for men and women, respectively) and 95% confidence intervals (dotted black lines) based on the available points are superimposed. The projections intersect just before the 2156 Olympics, when the winning women's 100-metre sprint time of 8.079 s will be faster than the men's at 8.098 s.

# Perigos da extrapolação



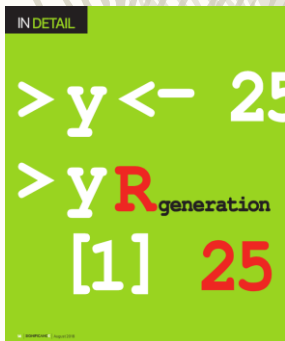
## Biology students find holes in gap study

*Sir* — We are students aged 16–18 in a Texas high school. Our biology teacher Vidya Rajan asked us to comment on the paper by A. J. Tatem and colleagues (*Nature* **431**, 525; 2004); we believe the projection on which it is based is riddled with flaws.

The idea of women running faster than men — although not novel (see B. J. Whipp and S. A. Ward *Nature* **355**, 25, 1992; and Correspondence *Nature* **356**, 21, 1992) — is interesting, but one cannot draw these conclusions based on generalization by extrapolation. Tatem *et al.* used a domain of 104 years to extrapolate to a domain of 252 years. It is not logical to say that the first 104 years will have data with exactly the same regression as the next 148 years. Using similar reasoning in 1992, Whipp and Ward suggested that women would run the marathon faster than

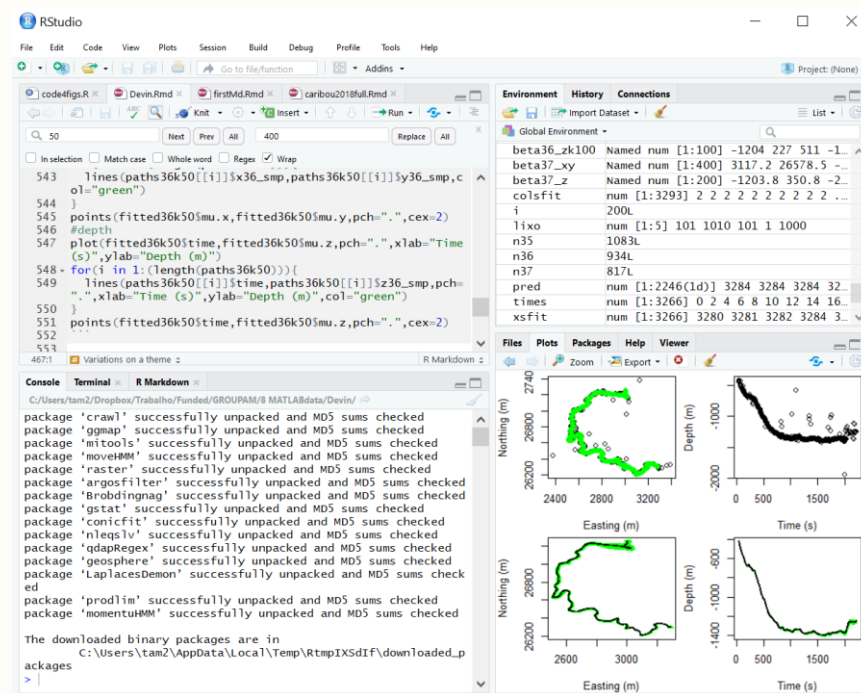
# I've got the Power !

- A base para as nossas análises de dados vai ser o ambiente R
- Ambiente integrado de análise de dados e linguagem de programação
- R Studio (uma interface para o R)
- Relatórios dinâmicos e R Markdown
- O conceito de *Reproducible Research*



The story of a statistical programming language that became a subcultural phenomenon. By **Nick Thieme**

Thieme, N. 2018 R generation. Significance 15:14-19



# A estatística como uma mais valia para um biólogo

---

- Se gostam da análise de dados então não hesitem e invistam em “**Ecologia Numérica**”
- Dominar esta área é meio caminho andado para garantir uma enorme variedade de opções em termos de empregabilidade futura
- Biólogos... há muitos!
- Estatísticos ecológicos são muito poucos, e por isso podem escolher o que fazem, e na realidade podem fazer quase tudo aquilo a que se propuserem.

## Bayesian inference in camera trapping studies for a class of spatial capture–recapture models

J. ANDREW ROYLE,<sup>1,3</sup> K. ULLAS KARANTH,<sup>2</sup> ARJUN M. GOPALASWAMY,<sup>2</sup> AND N. SAMBA KUMAR<sup>2</sup>

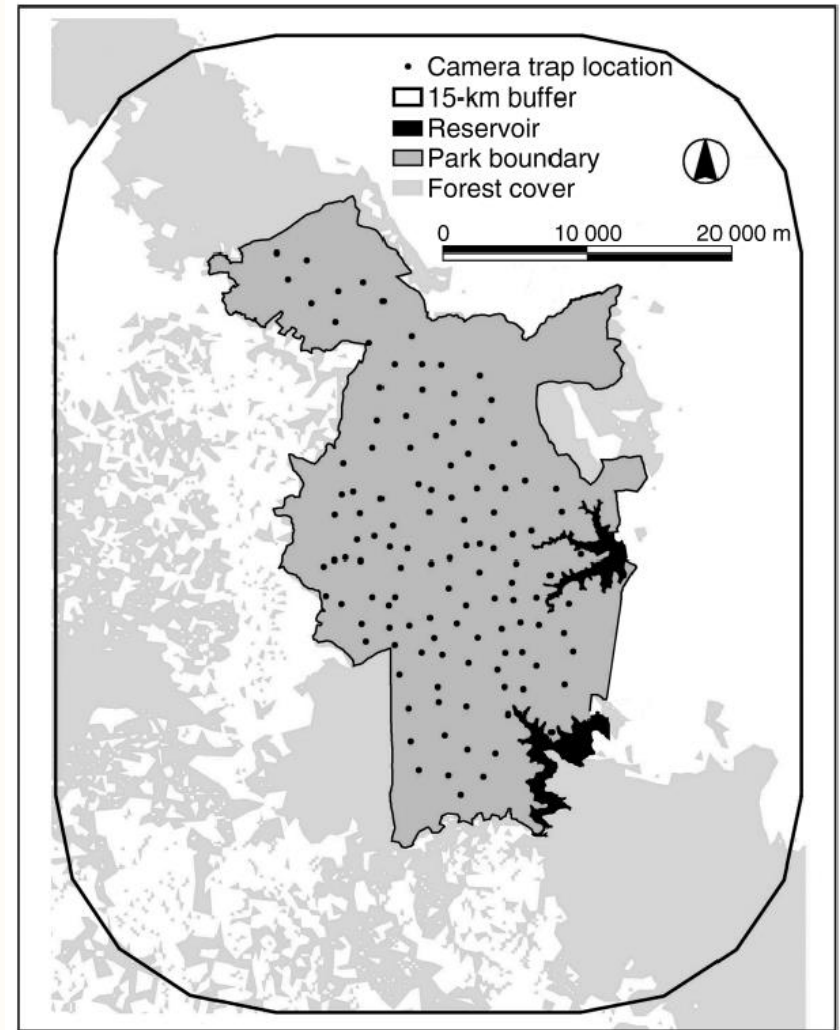


FIG. 1. Nagarahole Reserve tiger camera-trapping study area, Karnataka, southwestern India, with 15-km buffer outlined and non-habitat whited out.

APPLICATION

Program SPACECAP: software for estimating animal density using spatially explicit capture–recapture models

Arjun M. Gopalaswamy<sup>1,2,3\*</sup>, J. Andrew Royle<sup>4</sup>, James E. Hines<sup>4</sup>, Pallavi Singh<sup>1,2</sup>, Devcharan Jathanna<sup>1,2</sup>, N. Samba Kumar<sup>1,2</sup> and K. Ullas Karanth<sup>1,2,5</sup>

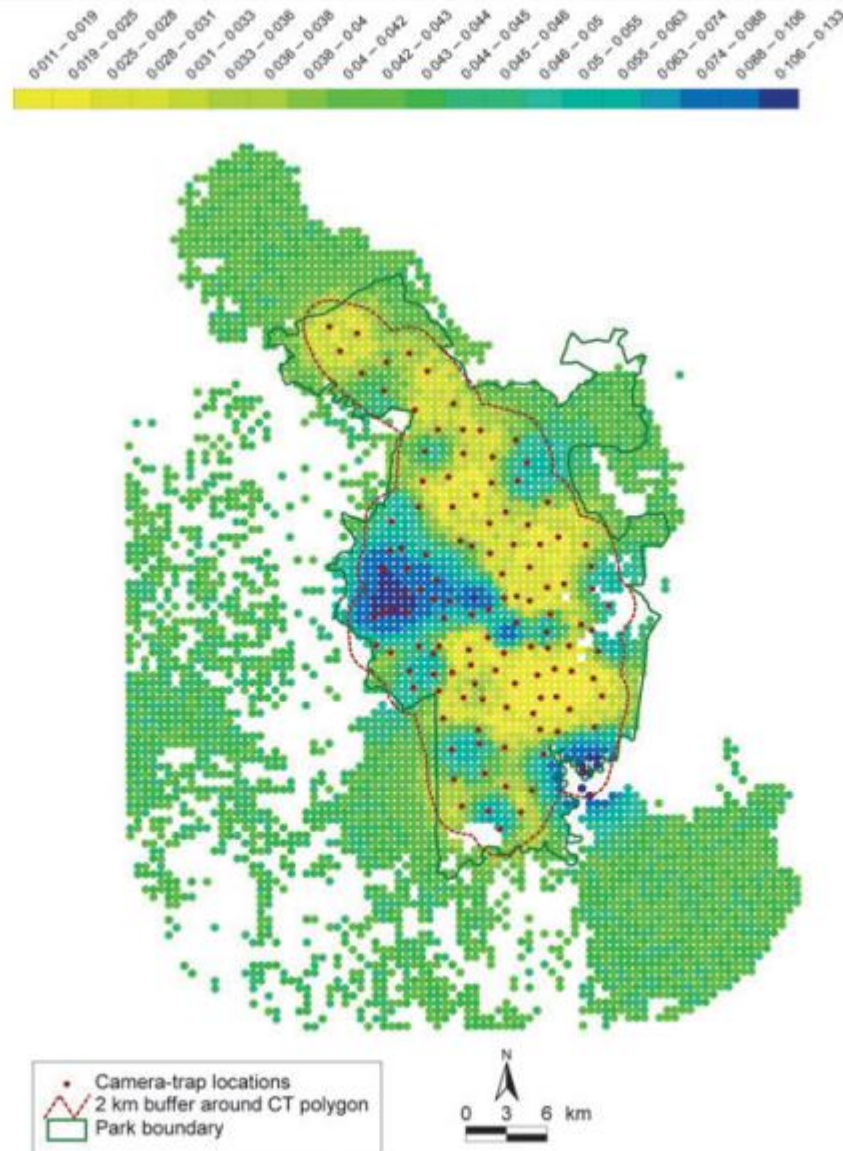


Fig. 2. A pixellated density map produced in SPACECAP showing relative animal densities. This map reports estimated tiger densities per pixel of size 0.336 km<sup>2</sup>. White regions represent unsuitable habitat.



# introdução à análise de dados o método científico

---

- Qual a necessidade dos estudos científicos efectuarem uma análise dos dados?
- Como foi a evolução da análise de dados nas últimas décadas?
- Como se deve desenvolver um estudo científico?





# introdução à análise de dados o método científico

---

- A análise dos dados é uma necessidade transversal em qualquer domínio científico
- A massificação dos processos analíticos de dados está intimamente relacionada com o advento da computação
- O refinamento e multiplicidade de técnicas também está associado às crescentes facilidades de cálculo
- Os programas de cálculo foram-se tornando progressivamente mais acessíveis e menos exigentes em termos de conhecimentos para a operação



# introdução à análise de dados o método científico

---

- Os dados resultam de observações e de experiências.
- Como são utilizados na investigação científica?
- Como se “faz” Ciência?



# introdução à análise de dados o método científico

---

*Aspectos filosóficos*

## Método dedutivo

1. Termos de referência (ex<sup>o</sup> ponto, recta)
2. Definições (triângulo, quadrado)

é uma afirmação que descreve um termo

3. Axiomas (dois pontos definem uma recta)

uma afirmação que é tida como verdadeira, que serve como base ou pressuposto para novos argumentos

4. Teorema

uma afirmação provada com base em outras afirmações mais simples que se assumem como verdadeiras (os axiomas)



## Método dedutivo

The Deductive Method - Mathematics

axiom

axiom

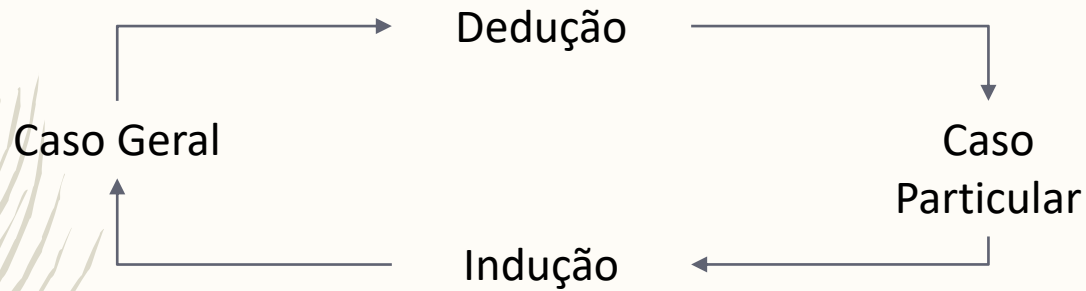
Start with simple, true statements called axioms (or postulates) along with definitions and undefined terms..



## Método dedutivo

Axiomas ou teoremas errados poderiam gerar uma infinidade de outros partindo de premissas incorrectas.

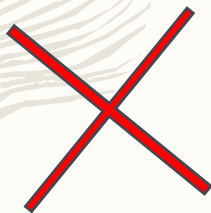
Mas se as premissas forem correctas a conclusão deverá ser correcta



No método dedutivo começamos por premissas que “assumimos” ser verdadeiras, e tiramos conclusões que serão verdadeiras (se os pressupostos forem verdadeiros)

Lógica 1

Todos os alunos detestam Matemática  
O Sebastião detesta matemática  
O Sebastião é um aluno



Lógica 2

Todos os alunos detestam matemática  
O Sebastião é um aluno  
O Sebastião detesta matemática





# introdução à análise de dados o método científico

---

## Método indutivo

1. Parte de observações da Natureza
2. Tenta a partir destas observações formular leis e teorias



Francis Bacon  
1561-1626



Galileo Galilei  
1564-1642

1. Esta espécie de baleia respira ar
2. Esta outra espécie de baleia respira ar
3. E esta terceira também
4. Indução: todas as baleias respiram ar (resultado final pode ser errado)



## Método indutivo

The Scientific Method - (Grossly Oversimplified)

Observation

The basis of the scientific method is making OBSERVATIONS -  
sometimes in the form of EXPERIMENTS.



## Dedução ou indução?

Factos:

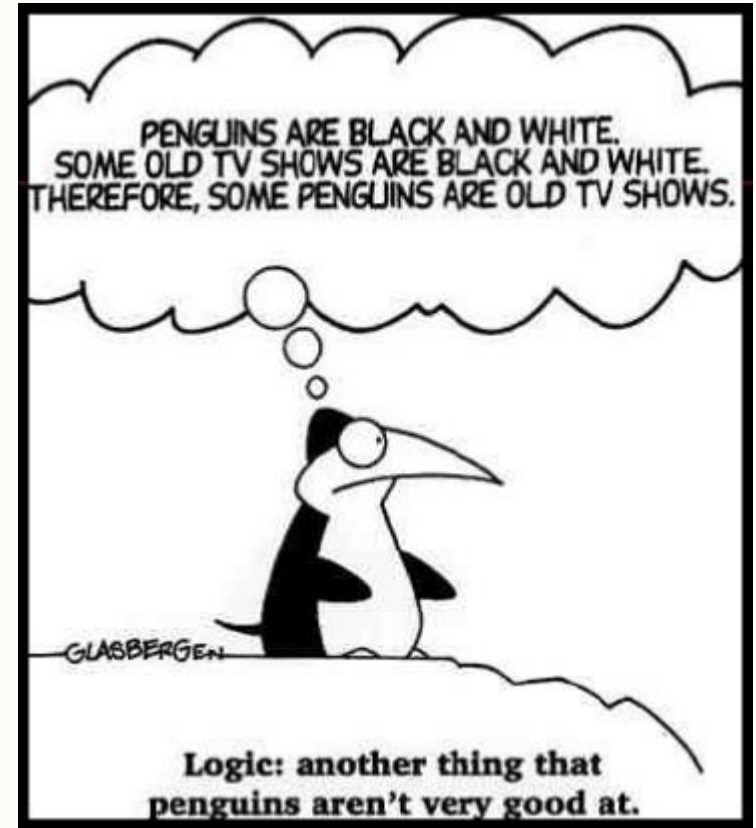
“A aula de apresentação foi uma seca!”

“A primeira TP foi uma seca”

“A aula de hoje está a ser uma seca”

Raciocínio:

A aula de amanhã vai ser uma seca...





# introdução à análise de dados o método científico

---

## *Método hipotético-dedutivo*

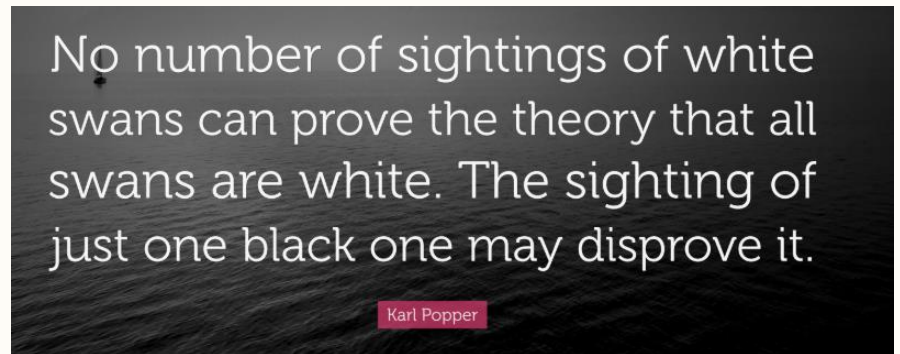
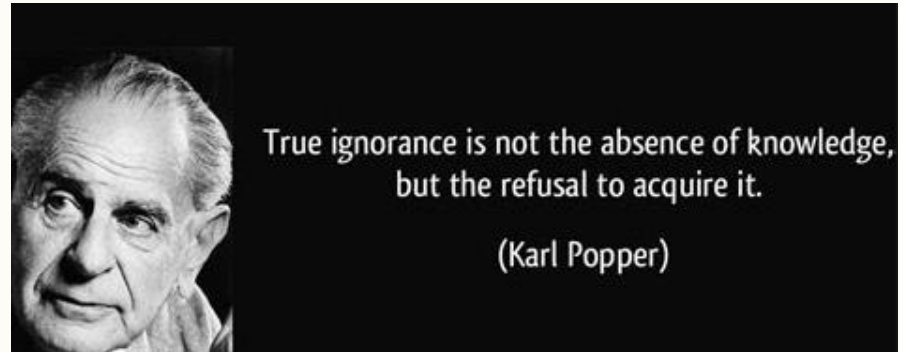
Baseia-se na:

Descrição de padrões

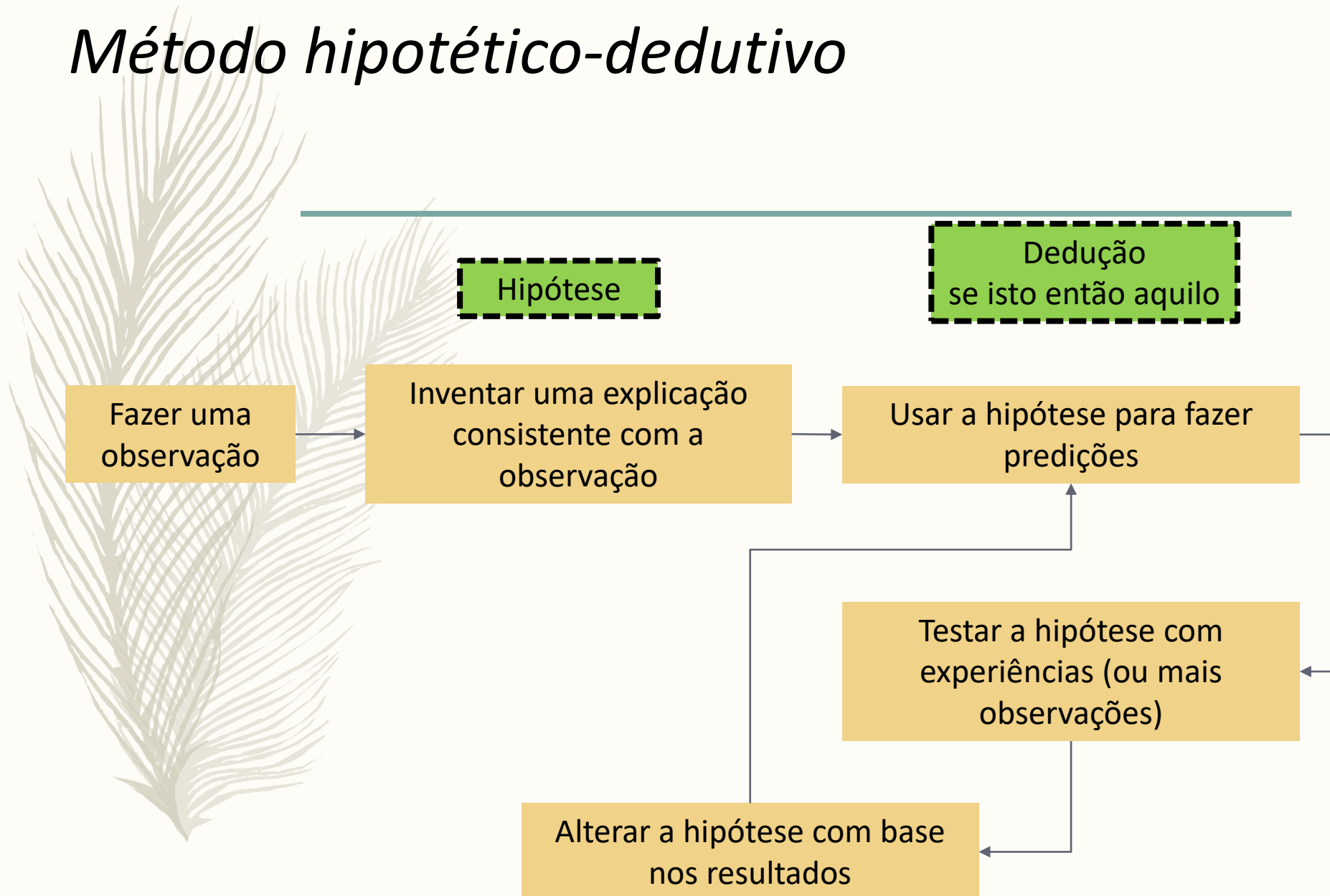
Dedução de um modelo

Formulação de hipóteses

Testes – procura da falsidade das hipóteses



# *Método hipotético-dedutivo*



# Questão: Os macroinvertebrados aquáticos de um rio respondem à velocidade da corrente?

---

Processo iterativo:

Passo 1

Fazer uma observação: levantar umas quantas pedras na zona de corrente, ir a uma zona de remanso e fazer o mesmo, observar os animais encontrados

Hipótese: As comunidades são iguais

Teste: recolher 50 amostras num rápido e 50 amostras num pego, e contar quantos exemplares de cada espécie se encontram

Resultado: De facto há 3 vezes mais animais na zona de corrente, as comunidades não devem ser iguais

# Questão: Os macroinvertebrados aquáticos de um rio respondem à velocidade da corrente?

---

Processo iterativo:  
Passo 2

Hipótese: há uma zona de clara fronteira entre pego e rápido

Teste: recolher 10 amostras em cada um de 10 lugares, entre um pego e um rápido

Resultado: Há uma tendencia gradual para aumentar a abundancia com a corrente, mas nenhuma clara fronteira

# Questão: Os macroinvertebrados aquáticos de um rio respondem à velocidade da corrente?

---

Processo iterativo:  
Passo 3

Hipótese: há uma mudança linear no número de animais com a velocidade da corrente

Teste: correlacionar abundância com velocidade da corrente

Resultado: a correlação não é linear

# Questão: Os macroinvertebrados aquáticos de um rio respondem à velocidade da corrente?

---

Processo iterativo:  
Passo 4

Hipótese: Há uma relação complexa entre abundância e corrente

Etc.



# introdução à análise de dados o método científico

---

## *Método hipotético-dedutivo*

Quais as hipóteses a formular?

Poderão ser quaisquer?

*Exemplos:*

*Há vida noutros planetas!*

*Há lince ibérico em Portugal!*





## Método hipotético-dedutivo

Este modelo assenta exclusivamente no princípio da refutação de hipóteses. A hipótese é construída por forma a ser refutada/rejeitada!

Exemplos:

Não há vida noutros planetas!

Não há lince ibérico em Portugal!

O medicamento ou o placebo são iguais!

O crescimento desta planta é independente da temperatura!

Há o mesmo número de zebras na savana e na floresta!



# introdução à análise de dados o método científico

---

## *Como surgem as questões a investigar*

- **Curiosidade** (exº como as aves migratórias encontram os seus territórios?)
- **Observações casuais** (exº um peixe salta da água. Porquê?)
- **Observações exploratórias** (exº quando colocamos vários indivíduos num aquário há comportamentos agressivos entre eles)
- **Investigação prévia** (exº estudos anteriores levantam novas questões)