



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

Versão temporária das aulas 1 e 2

<https://www.dropbox.com/s/50rthhr7e18bsge/EcolNum0T1T2.pptx?dl=0>

BAD FENIX

Docente > Gestão

Gestão de Páginas

- Ecologia Numérica
 - Ecologia Numérica(Tecnolo
 - Teóricas
 - Aula 1
 - Aula 2
 - Aula 3
 - Teórico-Práticas
 - Aula 1
 - Aula 2
 - Informação Geral
 - Informação Geral
 - PDFs
 - Outros recursos



Impor

9 ou Alunos de Ecologia Numérica (ENum-2) 1

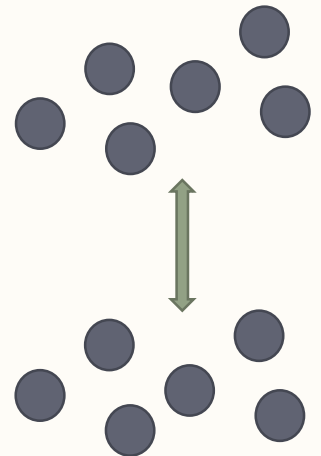
9 ou Alunos de Ecologia Numérica (ENum-2) 1

+ Criar

EN T1.pdf

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

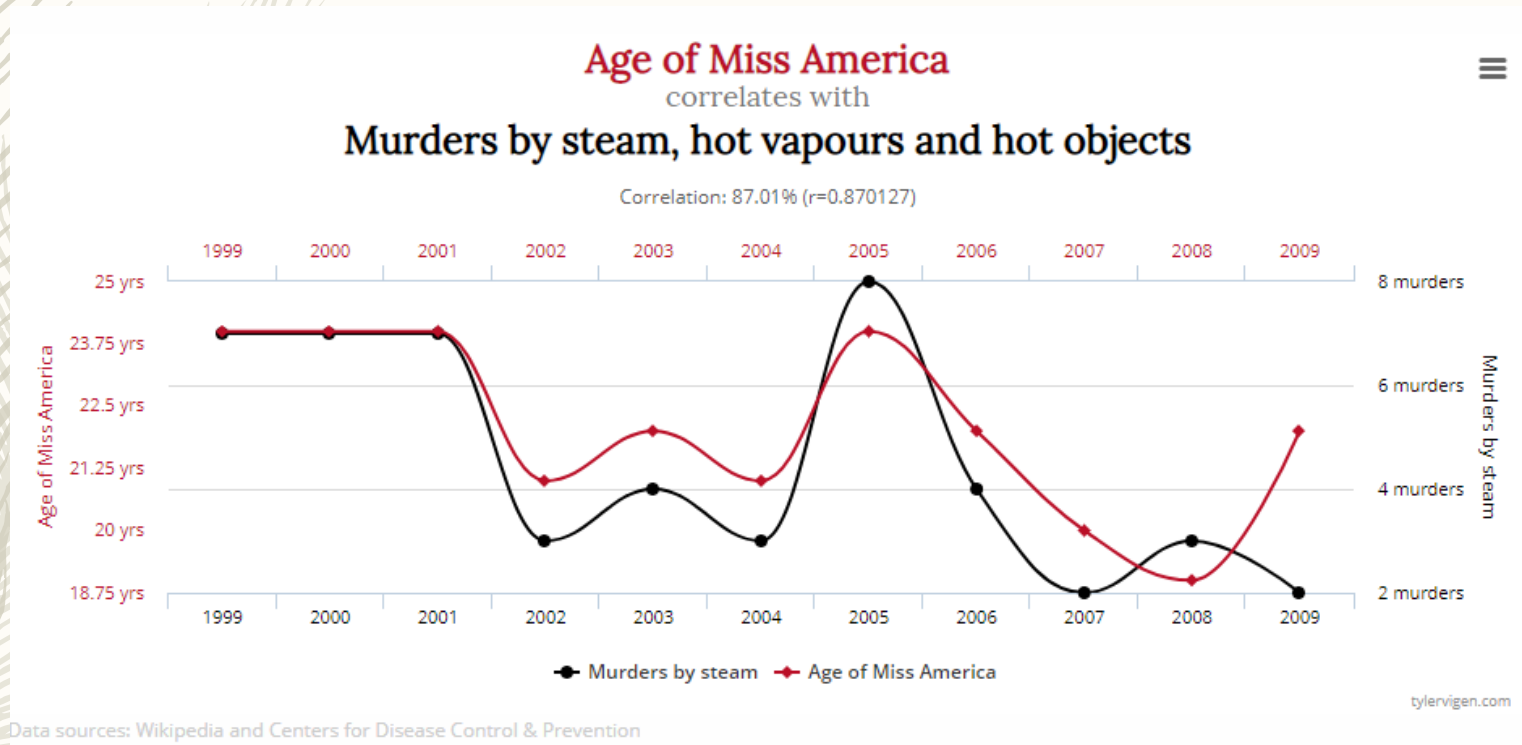
STUART H. HURLBERT
*Department of Biology, San Diego State University,
San Diego, California 92182 USA*

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 plateau¹. Others contend that there is no evidence that athletes, male or female, are reaching the limits of their potential^{1,2}.

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-

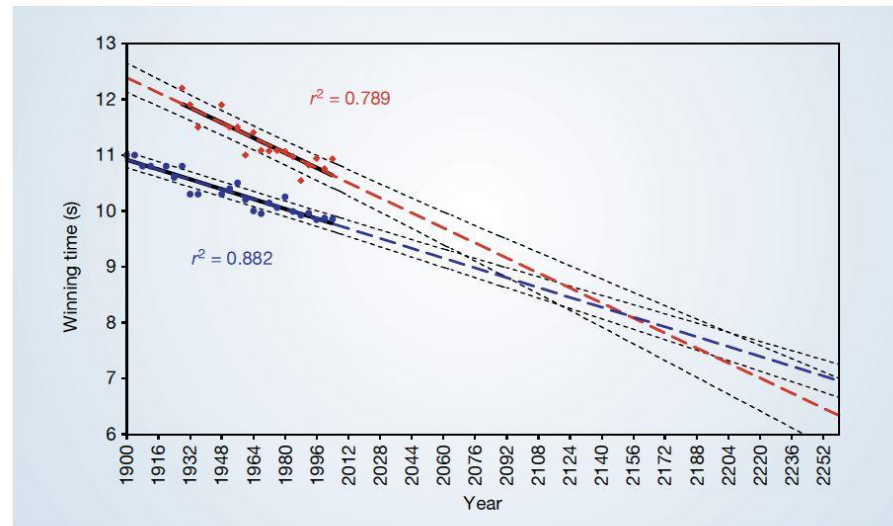


Figure 1 The winning Olympic 100-metre sprint times for men (blue points) and women (red points), with superimposed best-fit linear regression lines (solid black lines) and coefficients of determination. The regression lines are extrapolated (broken blue and red lines for men and women, respectively) 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



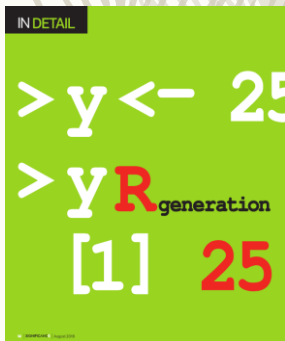
Biology students find holes in gap study

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

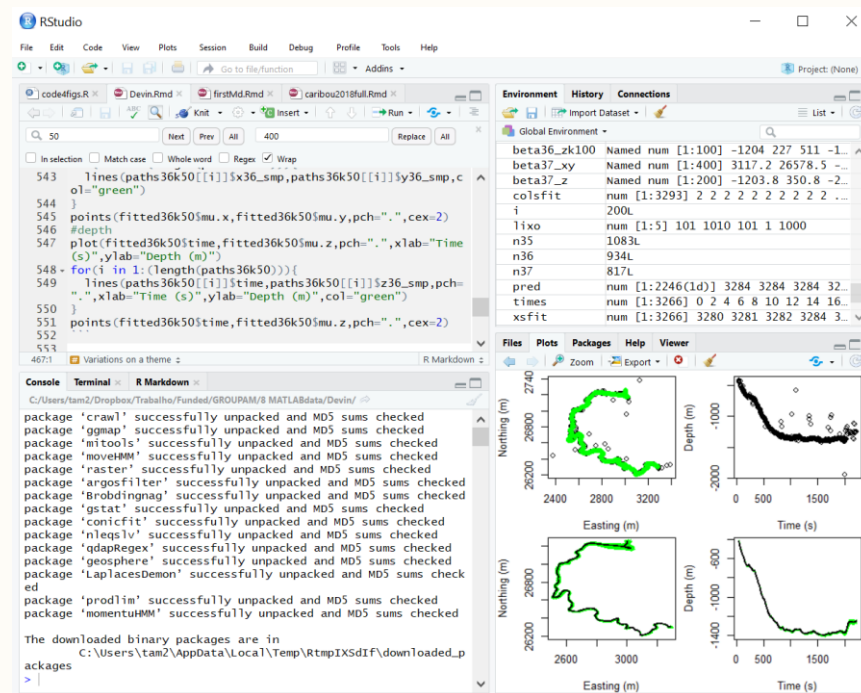
I've got the Power !

- 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,^{1,3} K. ULLAS KARANTH,² ARJUN M. GOPALASWAMY,² AND N. SAMBA KUMAR²

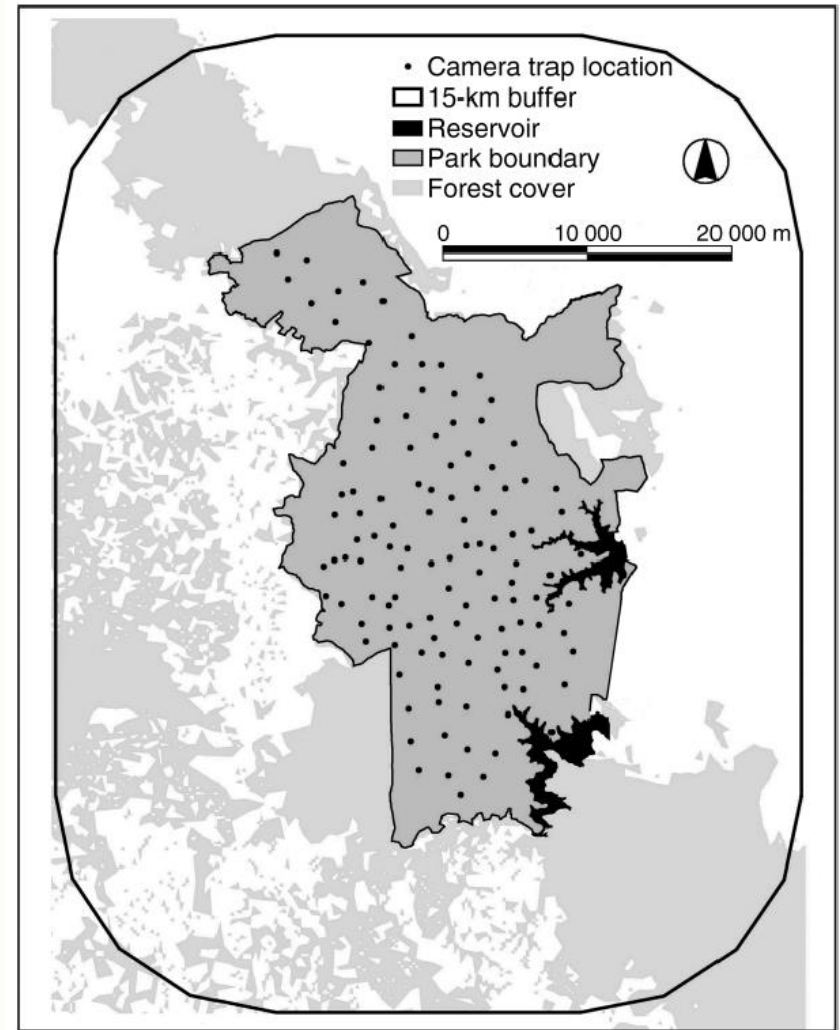


FIG. 1. Nagarahole Reserve tiger camera-trapping study area, Karnataka, southwestern India, with 15-km buffer outlined and non-habitat whited out.

APPLICATION

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

Arjun M. Gopalaswamy^{1,2,3*}, J. Andrew Royle⁴, James E. Hines⁴, Pallavi Singh^{1,2}, Devcharan Jathanna^{1,2}, N. Samba Kumar^{1,2} and K. Ullas Karanth^{1,2,5}

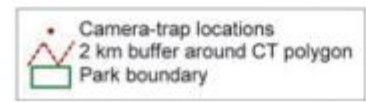
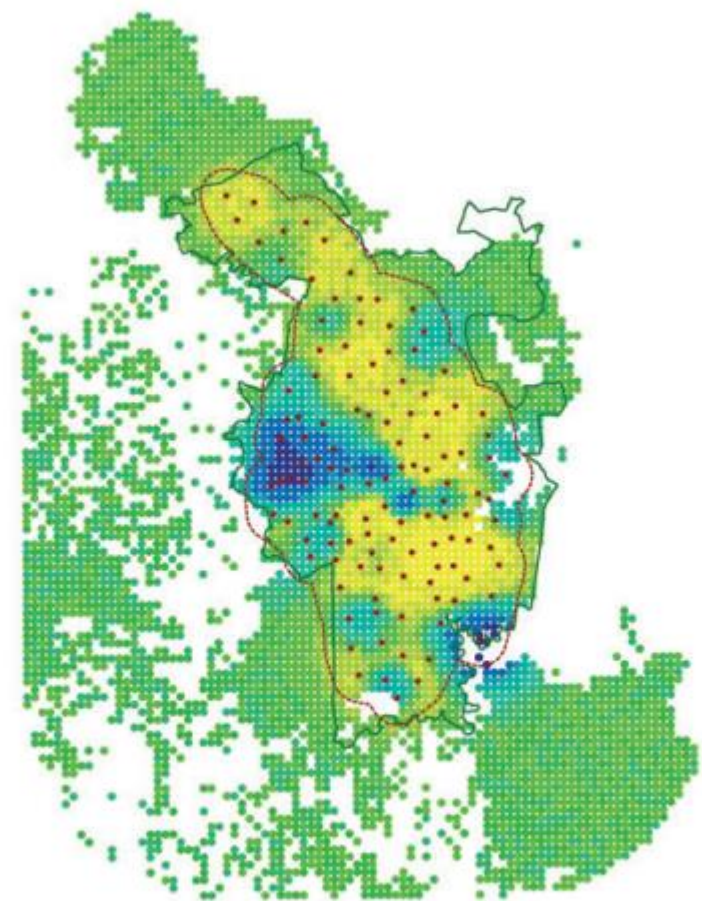


Fig. 2. A pixellated density map produced in SPACECAP showing relative animal densities. This map reports estimated tiger densities per pixel of size 0.336 km². White regions represent unsuitable habitat.



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

- 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^o 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)



introdução à análise de dados o método científico

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



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)



introdução à análise de dados o método científico

Método indutivo

The Scientific Method - (Grossly Oversimplified)

Observation

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



introdução à análise de dados o método científico

Método hipotético-dedutivo

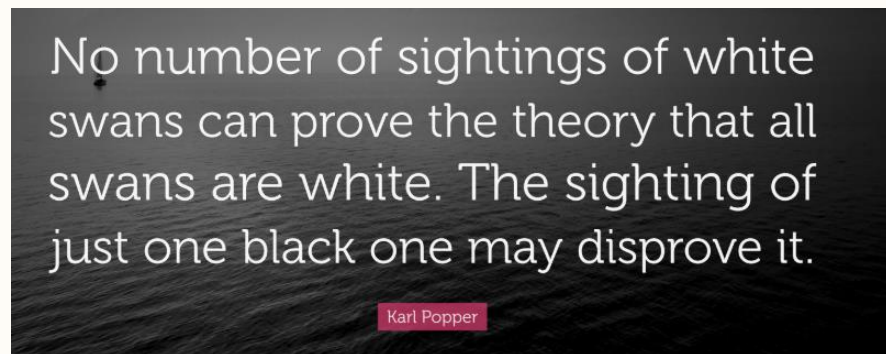
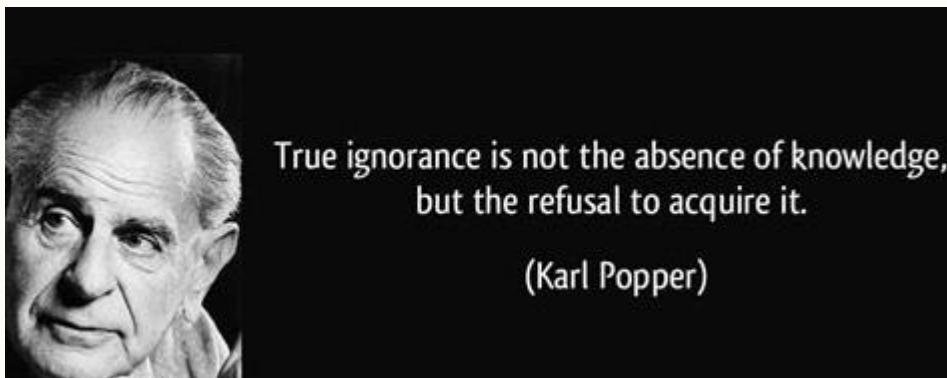
Baseia-se na:

Descrição de padrões

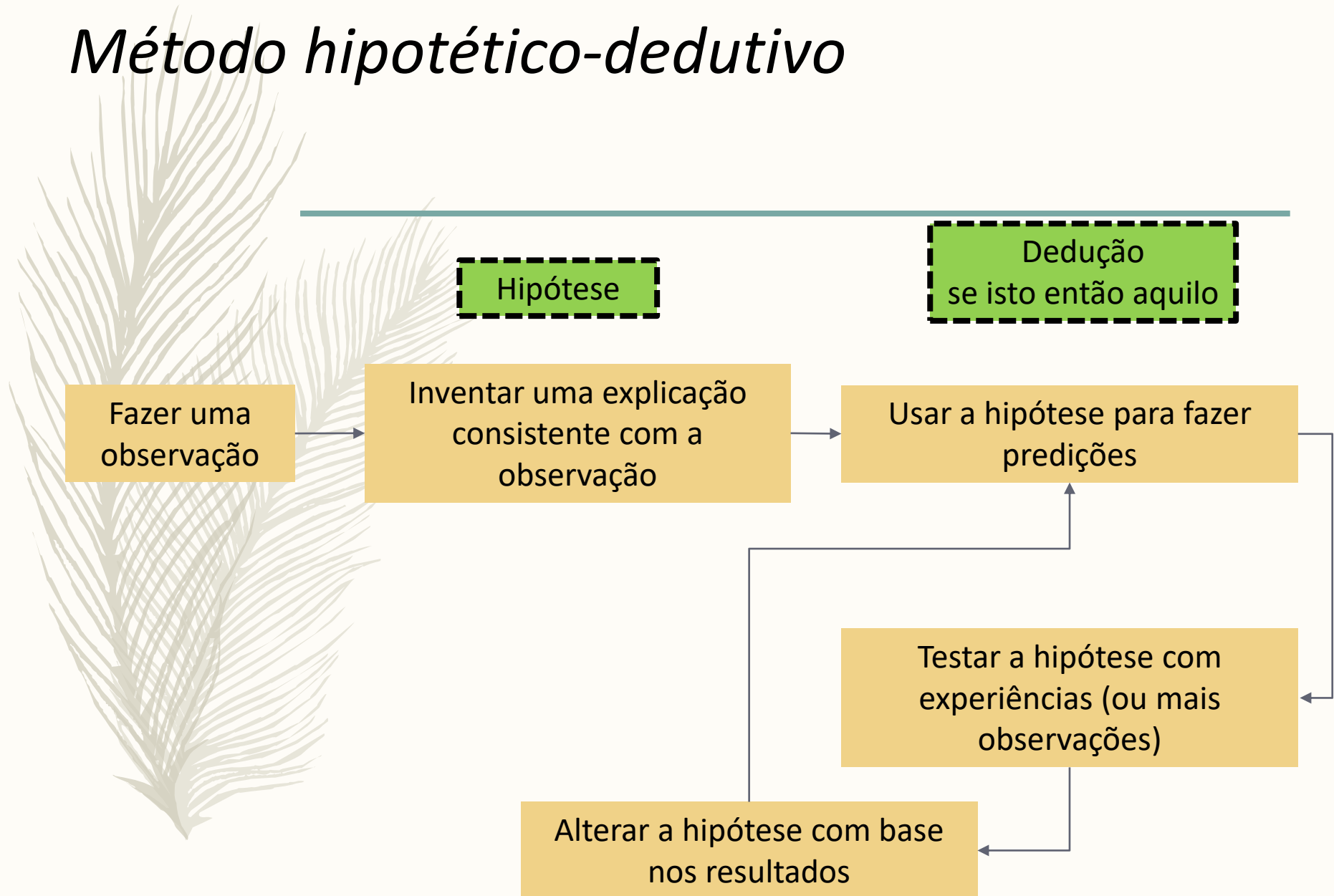
Dedução de um modelo

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 diferentes

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

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!

Exemplos:

Não há vida noutros planetas!

O medicamento ou o placebo são iguais!

Não há lince ibérico em Portugal!

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)



introdução à análise de dados o método científico

“Tipos” de estudos científicos

Unidisciplinares

Multidisciplinares

Interdisciplinares



“Tipos” de estudos científicos

- Descritivos vs. Experimentais
- Clássicos vs. Inovadores
- Fundamentais vs. Aplicados
- Importância regional vs. Importância global



introdução à análise de dados o método científico

Hipóteses sem dados não têm utilidade!

mas...

Dados sem hipóteses também não!



introdução à análise de dados o método científico

Que componentes deverá ter um
programa de investigação?

An approximate answer to the right question is worth a great deal more than...
a precise answer to the wrong question.

John Tukey

TPC: trabalho para casa

- Recolher um conjunto de dados “ecológicos”, com um tamanho de amostra pelo menos igual a 30, idealmente maior que 50
- Formular uma pergunta ecológica
- Registrar (pelo menos) duas variáveis que possam ser comparadas, e uma variável que possa ser relacionada com as anteriores
- Exemplos:
 - seleccionar 50 árvores. Recolher 2 folhas de cada árvore e o dap (diâmetro à altura do peito) de cada árvore
 - seleccionar 50 plantas com flores. Medir a altura ao solo da planta. Seleccionar a flor mais alta e a mais baixa. Contar quantos insectos há em cada uma das flores.
 - Seleccionar 50 pombos. Registrar se é macho ou fêmea. Registrar se está só ou acompanhado. Andar em direcção a cada pombo e registrar a distância a que ele “para e olha” para avaliar o perigo.

