



# Ecologia Numérica

2018/2019

PRELIMINARES

---

Tiago A. Marques

[tamarques@ciencias.ulisboa.pt](mailto:tamarques@ciencias.ulisboa.pt)

# Ecologia Numérica: Horário

1.5 + 1.5 horas teóricas e 2 horas teórico-práticas por semana

	SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
	16	17	18	19	20	21	22
7 <sup>AM</sup>							
8				EN TP 2.2.19			
9							
10			EN T 2.2.21			EN TP 2.2.19	
11				Atendimento			
12 <sup>PM</sup>							
1		EN T 6.1.36					
2							
3		EN TP 2.2.19					
4							
5							
6							
7							
8							
~							

# Ecologia Numérica

- Horário de atendimento: quartas, das 10:30 às 12:30, sala 2.3.11A (enviar mail para [tamarques@ciencias.ulisboa.pt](mailto:tamarques@ciencias.ulisboa.pt) para agendar)



**OUR TEACHER  
IS MISSING...**

Writing Prompt: You arrive at school today and your teacher is not in the classroom. No one knows where she is or when she is coming back! Write a story telling what has happened to your teacher.



# Ecologia Numérica

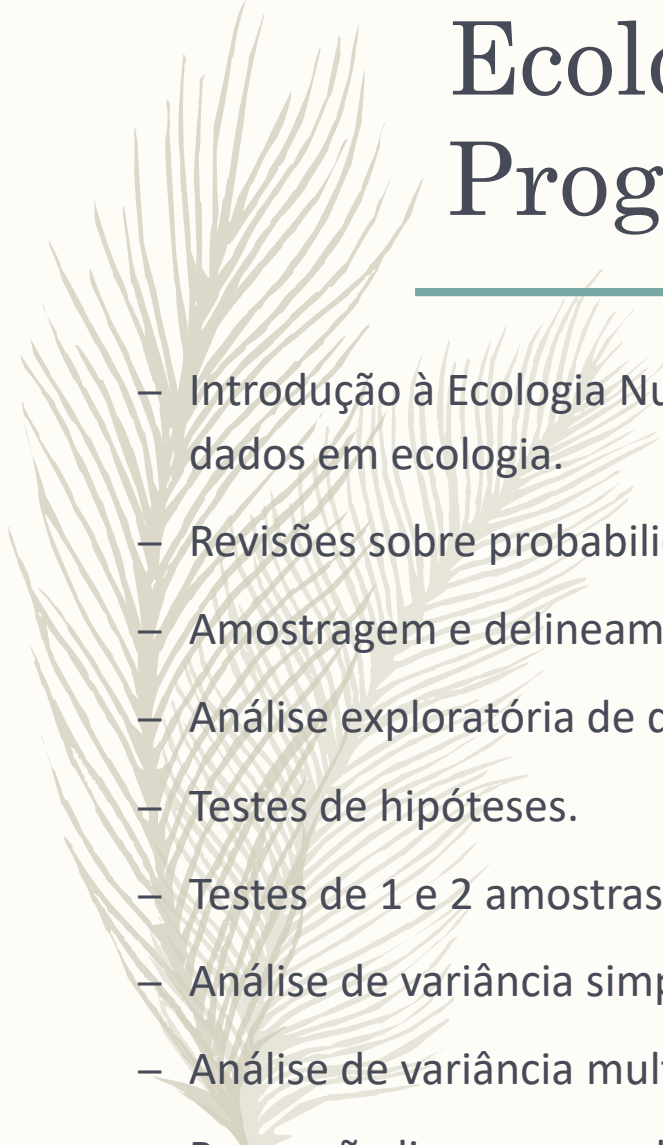
---

- Avaliação: 50% exame teórico + 50% exame teórico-prático com recurso a computador e software

Datas:

	1a Época	2a Época	Época Especial
Teórico	22 Jan 2019 – 16:30-19:30	08 Fev 2019 – 16:30-19:30	22 Jul 2019 – 14:30-16:30
Teórico-Prático	22 Jan 2019 – 8:00-16:00	08 Fev 2019 – 9:00-15:00	22 Jul 2019 – 11:30-13:30


Para ser possível a aprovação da disciplina, os alunos terão que garantir presença em pelo menos 10 das aulas teórico-práticas, não sendo exigida presença obrigatória nas aulas teóricas.



# Ecologia Numérica: Programa

---

- Introdução à Ecologia Numérica. O método científico. Tipos de estudos e de dados em ecologia.
- Revisões sobre probabilidades.
- Amostragem e delineamento experimental.
- Análise exploratória de dados.
- Testes de hipóteses.
- Testes de 1 e 2 amostras.
- Análise de variância simples e testes equivalentes não-paramétricos.
- Análise de variância multifactorial e com designs complexos.
- Regressão linear e correlação.
- Regressão não linear e modelos lineares generalizados.



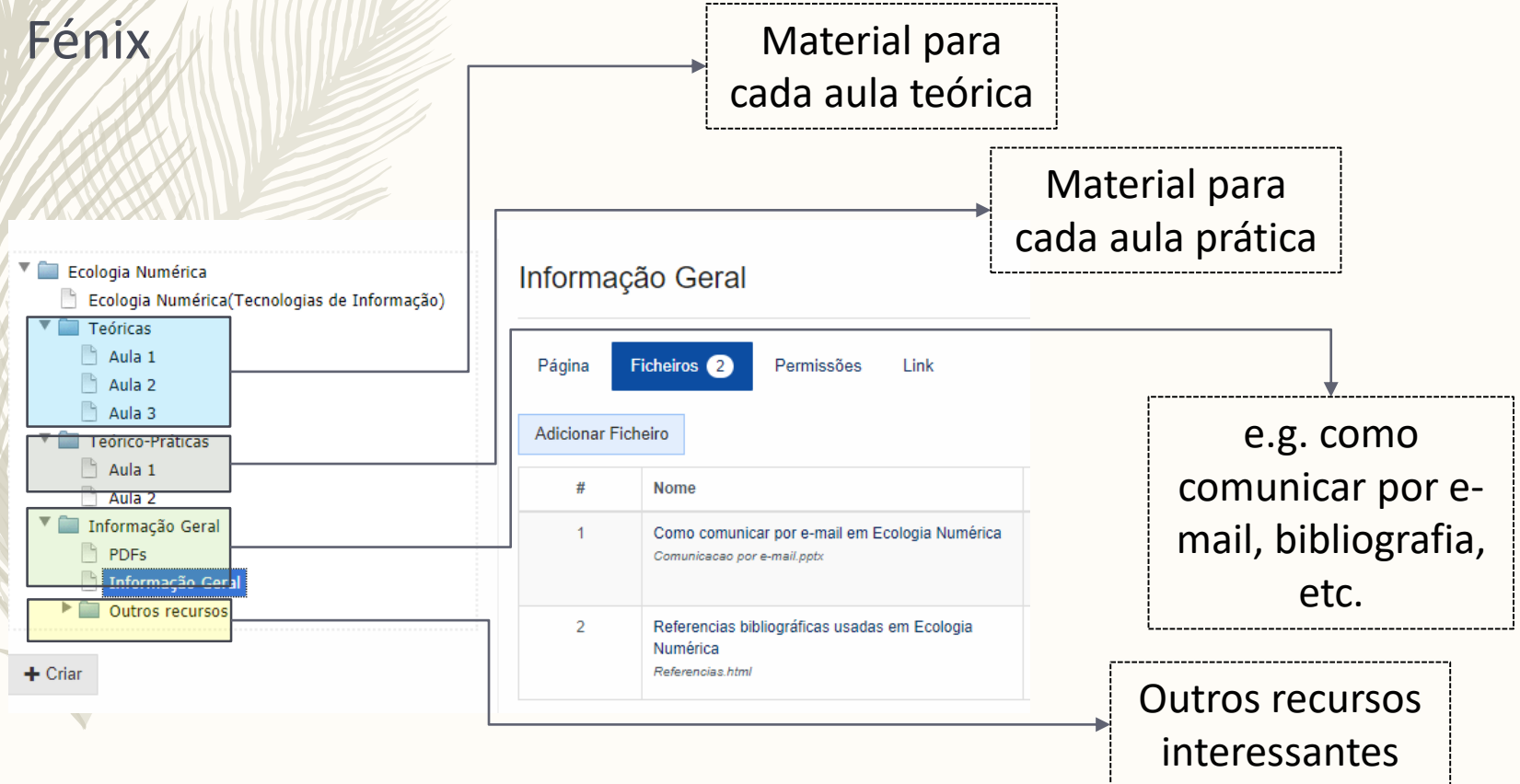
# Ecologia Numérica: Programa

---

- Análise de dados de frequências. Tabelas de contingência.
- Introdução à análise multivariada.
- Classificação hierárquica e não hierárquica.
- Análise discriminante.
- Análise de componentes principais.
- Análise de correspondências e análise canónica de correspondências.
- Escalamiento multidimensional.
- Critérios para a selecção de técnicas analíticas. Apresentação de resultados de análises estatísticas.

# Ecologia Numérica: Material de apoio às aulas

Fénix



# Bibliografía

- Borcard et al. 2011. Numerical Ecology with R. Springer
- Gotelli, N. & A. M. Ellison. 2004. A primer of ecological statistics. Sinauer Associates.
- Greenacre, M. & R. Primicerio. 2013. Multivariate analysis of ecological data. Fundación BBVA.
- Krebs, C. J. 1989. Ecological methodology. Harper-Collins Publishers.
- Bolker, B. 2007. Ecological Models and Data in R, rinceton University Press



# Bibliografia

- Legendre, L. & P. Legendre. 1998. Numerical ecology. Elsevier.
- Leps, J. & P. Smilauer. 2003. Multivariate Analysis of Ecological Data Using Canoco. Cambridge University Press.
- Manly, B. 1986. Multivariate statistical methods: a primer. Chapman & Hall.
- Murteira, B. 1993. Análise exploratória de dados. Estatística descritiva. McGraw-Hill.
- Pielou, E. C. 1984. Interpretation of ecological data: a primer on classification and ordination. John Wiley & Sons.
- Quinn, G. P. & M. J. Keough. 2002. Experimental design and data analysis for biologists. Cambridge University Press.
- Conover, W. J. 1980. Practical non-parametric statistics. 2nd edition. John Wiley & Sons
- Everitt, B. S. & G. Dunn. 2001. Applied Multivariate Data Analysis. Arnold Publication.
- Gauch, H. G. 1982. Multivariate analysis in community ecology. Cambridge University Press.
- Hair, J.; R. Anderson, R. Tatham & W. Black. 1995. Multivariate data analysis with readings. Prentice-Hall.

# Bibliografia

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

+ Criar

## Informação Geral

Página **Ficheiros 2** Permissões Link

Adicionar Ficheiro

#	Nome	Permissões
1	Como comunicar por e-mail em Ecologia Numérica <i>Comunicacao por e-mail.pptx</i>	Professores de Ecologia Numérica (ENum-2) 1 Semestre - 2018/2019 ou Alunos de Ecologia Numérica (ENum-2) 1 Semestre - 2018/2019
2	Referencias bibliográficas e páginas web <i>Referencias.html</i>	Público

# Bibliografia

Type to search...  Matching entries: 24/24

include abstract  use RegExp  ignore accents


Author	Title	Year	Journal/Proceedings	Reftype	DOI/URL
Legendre, P.	A brief history of the development of Numerical Ecology [Abstract] [BibTeX]	2018		www	URL

**Abstract:** A brief history of the development of Numerical Ecology. Uma apresentação recente por um dos fundadores da disciplina.

**BibTeX:**

```
@www{Legendre2018,
  author = {Pierre Legendre},
  title = {A brief history of the development of Numerical Ecology},
  year = {2018},
  url = {http://biol09.biol.umontreal.ca/PLcourses/Brief_history_of_NE.pdf}
}
```

Cooper, N. and Hsing, P.-Y.	A Guide to Reproducible Code in Ecology and Evolution [BibTeX]			book	URL
Steel, E.A., Kennedy, M.C., Cunningham, P.G. and Stanovick, J.S.	Applied statistics in ecology: common pitfalls and simple solutions [BibTeX]	2013	Ecosphere Vol. 4(9), pp. art115	article	DOI
BES	BES Guides to Better Science [BibTeX]	2018		www	URL
Anderson, D.R., Burnham, K.P., Gould, W.R. and Cherry, S.	Concerns about Finding Effects That Are Actually Spurious [BibTeX]	2001	Wildlife Society Bulletin Vol. 29, pp. 311-316	article	
Ullrich, M.D.	Environmental Statistics: A Handbook Using R	2015	Dis-Quintess	article	DOI



# Ecologia Numérica: Comunicação por e-mail

[tamarques@ciencias.ulisboa.pt](mailto:tamarques@ciencias.ulisboa.pt)

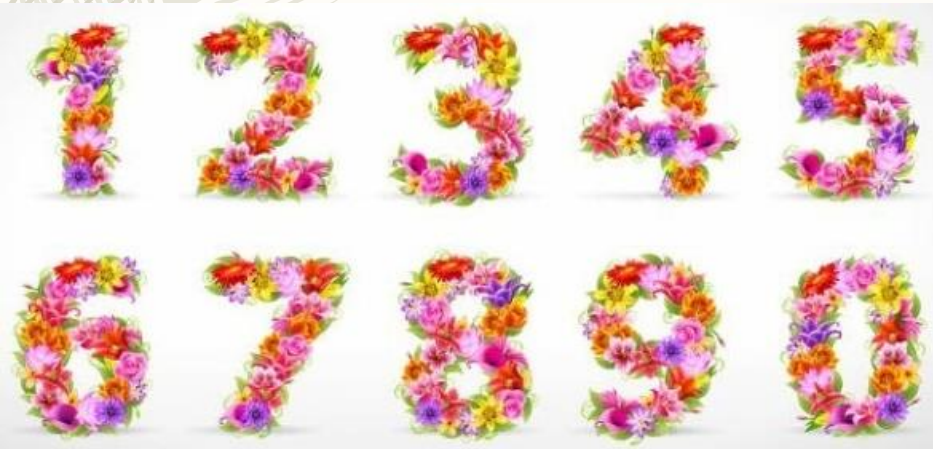
---

- Formatação do assunto do e-mail (sempre): EN-número de aluno-tópico-detahes
- Existirão tópicos específicos (e depois claro tópicos livres!):
  - Marcar um atendimento: EN-19549-atendimento-25 09 2018
  - Esclarecer uma duvida: EN-19549-duvida-teste de t
  - Enviar ficheiro de dados: EN-19549-dados-TiaMar19549.xlsx
  - Corrigir gralha nos slides: EN-19549-gralha-aula de \*\* \*\* \*\*\*\*\* Slide \*\*
  - Qualquer outro tópico a definir “EN-19549-sejaoqueforqueeudefina-blablaba”
  - Ou outro qualquer “EN-19549-outro-comentário sobre aula de ontem”
- Teste: cada um dos presentes vai-me mandar um mail teste
  - Assunto “EN-?????-teste1-whatever”.
  - No mail teste gostaria que me dessem qualquer informação que achem que me possa ser útil para a cadeira de EN, e.g.
    - *dados ou problemas que gostariam de ver usados como exemplos,*
    - *dificuldades que possam ter com a matemática/estatística,*
    - *etc.*

# Ecología Numérica

Objetivo principal:

pensar sobre números num contexto ecológico



# Quem sou eu?

---



- Biologia FCUL 1998 (fiz Ecologia Numérica em 1995/1996!)
- Mestrado, Probabilidade e Estatística, FCUL, 2001
- Doutoramento, Estatística, 2007, University of St Andrews (UStA)
- Senior Research Fellow como Ecological Statistician, UStA
- Investigador Integrado:
  - Centro de Estatística e Aplicações (CEAUL), FCUL
  - Center for Research into Ecological and Environmental Modelling (CREEM), UStA
- (até 03/09/2018) Professor auxiliar convidado no DEIO
- (desde 03/09/2018) Professor auxiliar convidado no DBA\*

\* a substituir o Prof. Henrique Cabral.

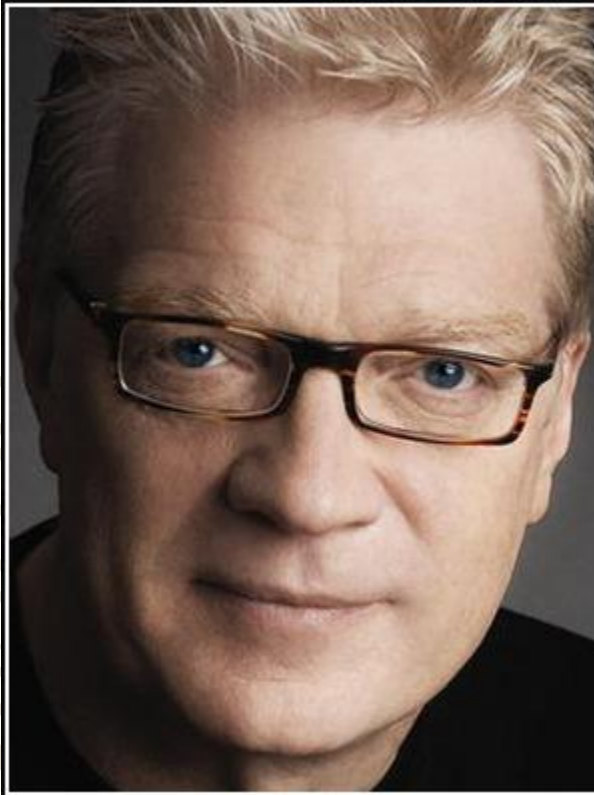
Nota importante: O crédito do material de EN é quase todo do Henrique. A responsabilidade dos erros é naturalmente minha.

# Quem sou eu?

- Interesses científicos:

- Amostragem de populações naturais.
- Estimação da abundância de populações naturais, em particular com base em dados de acústica passiva.
- Efeito de erros na medição das distâncias e distribuições não uniformes da densidade em amostragem por distâncias.
- Estatística ecológica
- Divulgação do uso da estatística aplicada à biologia





Teaching is a creative profession,  
not a delivery system. Great  
teachers do [pass on information],  
but what great teachers also do is  
mentor, stimulate, provoke, engage.

— *Ken Robinson* —

AZ QUOTES

<https://quotesurf.com/great-teachers-quotes>

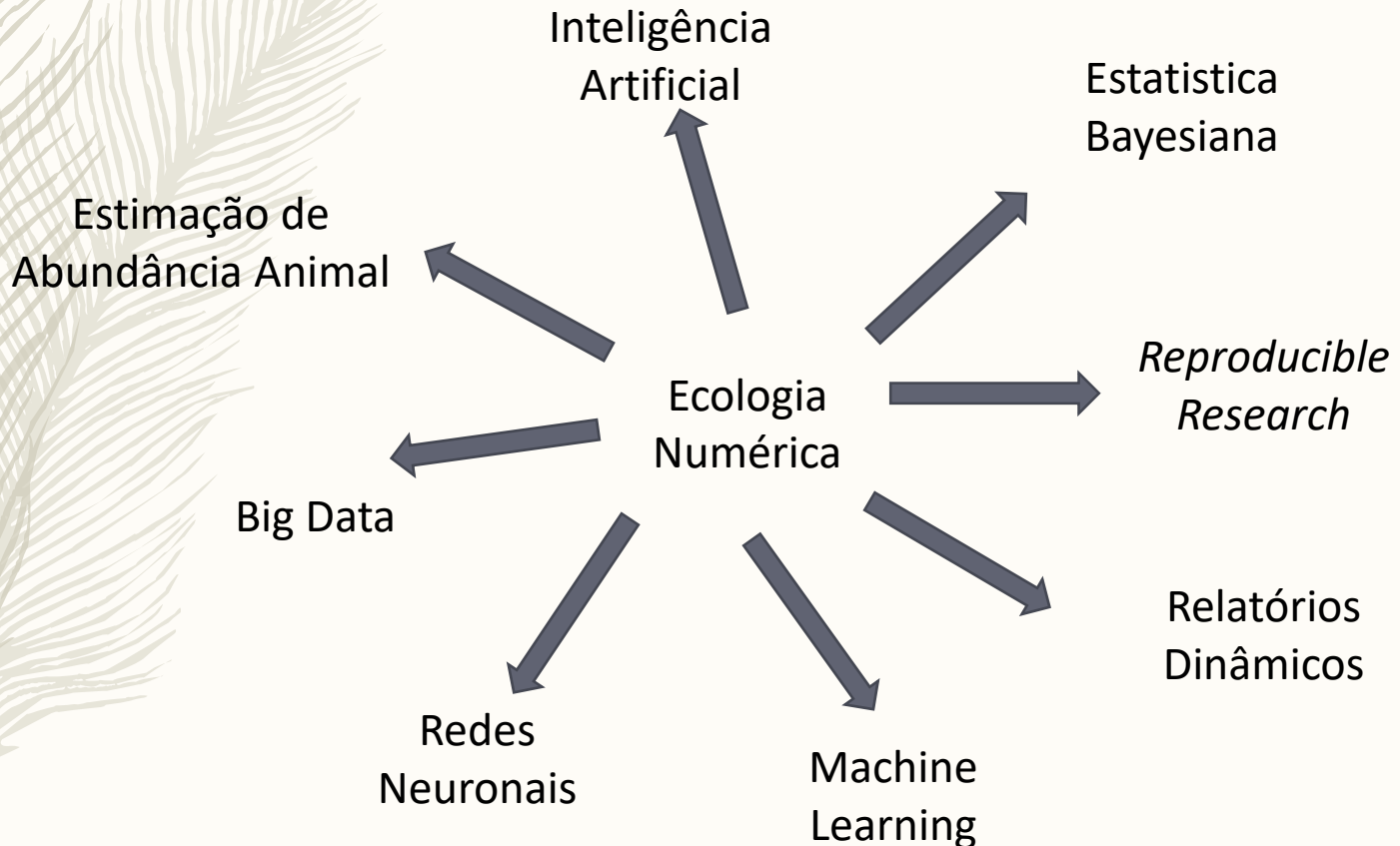
Todas as dúvidas são válidas, não há perguntas estúpidas. No que eu vos puder ajudar, sobre a Ecologia Numérica, a utilização da estatística em biologia ou o futuro, estou à disposição. Espero que aprendam comigo, e conto aprender convosco.



# Ecologia Numérica

(e outras coisas que são importantes  
como desmistificar chavões)

---



# Quem são vocês?



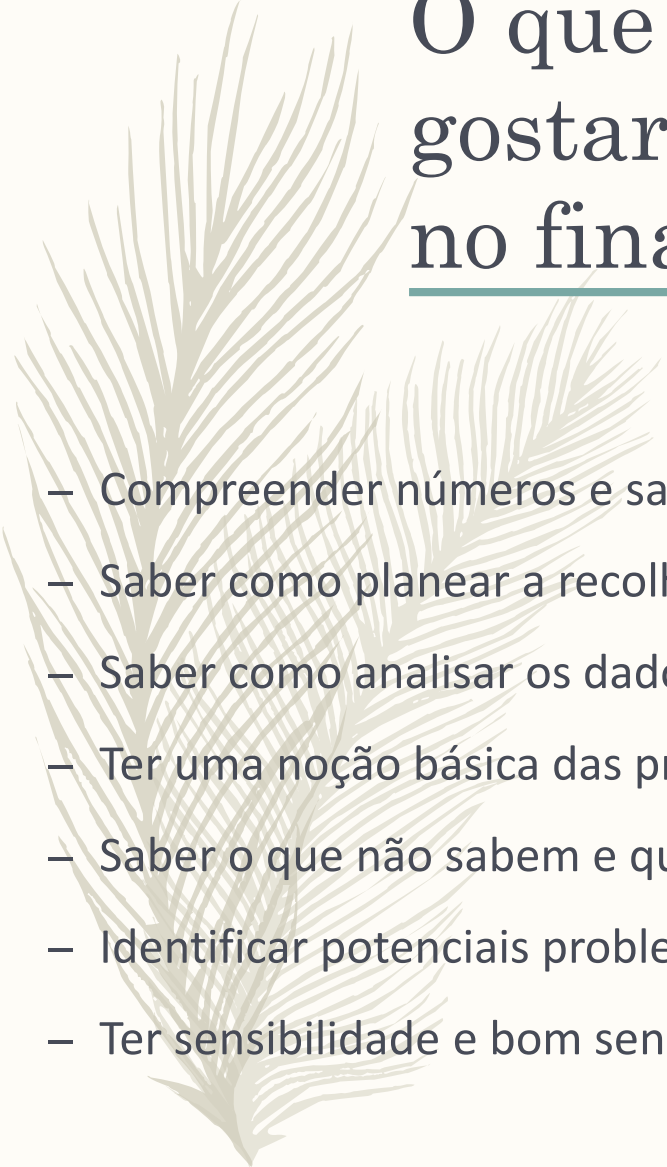
\*como fazer esta tag cloud? Por exemplo através de um relatório dinâmico...

# Contagem de mãos

---

1. Quem não sabe o que é uma variável aleatória?
2. Quem não sabe o que é um replicado?
3. Quem não sabe o que é um teste estatístico?
4. Quem não sabe o que é um valor P (P-value)?
5. Quem nunca trabalhou no R?
6. Quem não sabe o que é um relatório dinâmico ou *reproducible research*?
7. Quem não gostou de Bioestatística? (E porquê?)
8. Quem não tem medo de matemática?

\* Quem responder afirmativamente arrisca-se a que eu lhe pergunte para explicar à turma!



# O que quero que saibam (ou gostaria que soubessem!) fazer no final deste curso

- Compreender números e saber interpretá-los num contexto ecológico
- Saber como planear a recolha de dados para responder a uma questão ecológica
- Saber como analisar os dados resultantes
- Ter uma noção básica das principais técnicas de análise de dados ecológicos
- Saber o que não sabem e quando pedir ajuda para analisar um conjunto de dados
- Identificar potenciais problemas em ecologia relacionados com a análise de dados
- Ter sensibilidade e bom senso!!!



The best thing about being a  
statistician is that you get to play in  
everyone's backyard.

— *John Tukey* —

AZ QUOTES

<https://www.azquotes.com/quote/603405>

# Ecologia Numérica



---

Origem do termo: Ecologie Numérique (Legendre & Legendre, 1979)

Ecologia – é o ramo da biologia que lida com as relações entre os organismos e com as relações destes com o meio que os rodeia.

É sequer possível pensar em ecologia sem pensar em números? Sem pensar na necessidade de medir ou quantificar?

## A estatística nas revistas biológicas

Hoje em dia, ser biólogo implica algum tipo de publicação em revistas científicas da especialidade. Quantos papers de aplicações, no último número de cada revista conceituada na área da biologia/ecologia, usaram (pelo menos) uma técnica estatística como base para responder à(s) pergunta(s) colocada(s):

- ▶ Ecology - volume 92, número 9 - 13 em 13
- ▶ Biological Conservation - volume 144, número 11 - 19 em 19
- ▶ Conservation Biology - volume 25, número 5 - 14 em 14
- ▶ Journal of Animal Ecology - volume 80, número 1 - 25 em 25
- ▶ Journal of Applied Ecology, Nature, Science, PLOS biology -

TODOS

# Um inquérito (com quase tudo o que não se deve fazer...)

Enviado via e-mail, a todos os meus contactos biólogos, segunda-feira dia 17 Outubro às 11:14

"Ola a todos,

(desculpem o SPAM, espero que este mail vos encontrem bem)

(Se estão a receber isto e não são biólogos, desculpem, enganei-me, ignorem o resto)

Queria recolher uma amostra para uma aula que vou dar. Por isso preciso da vossa ajuda, para responder a 3 questões:

0. Em que ano completaste o curso?
1. Continuas a trabalhar em biologia?
2. Das cadeiras do tronco comum, qual aquela em que, se pudesses voltar atrás, gostarias de ter estado com mais atenção?
3. Das físicas, químicas, matemáticas e estatísticas, quais aquelas cujos conhecimentos necessitas/aplicas com mais frequência no teu trabalho?



# Um inquérito (com quase tudo o que não se deve fazer...)

Continuação do texto

" MUITO Obrigado

Bjs/Abr

T.

PS - aos que não forem do meu ano, ou universidade, se puderem passar isto a colegas do vosso ano, e reencaminharem para mim (ate terça à noite), agradecia. Se **alguem** estiver ao vosso lado, for **biologo**, e não receber isto, **pessam-lhe** para responder **tambem!**

PS2 - não quero bocas foleiras como " **alguem** como tu devia saber que não é assim que se faz um inquérito" ... isto é só para ilustrar uma ideia."

## Respostas ao inquérito

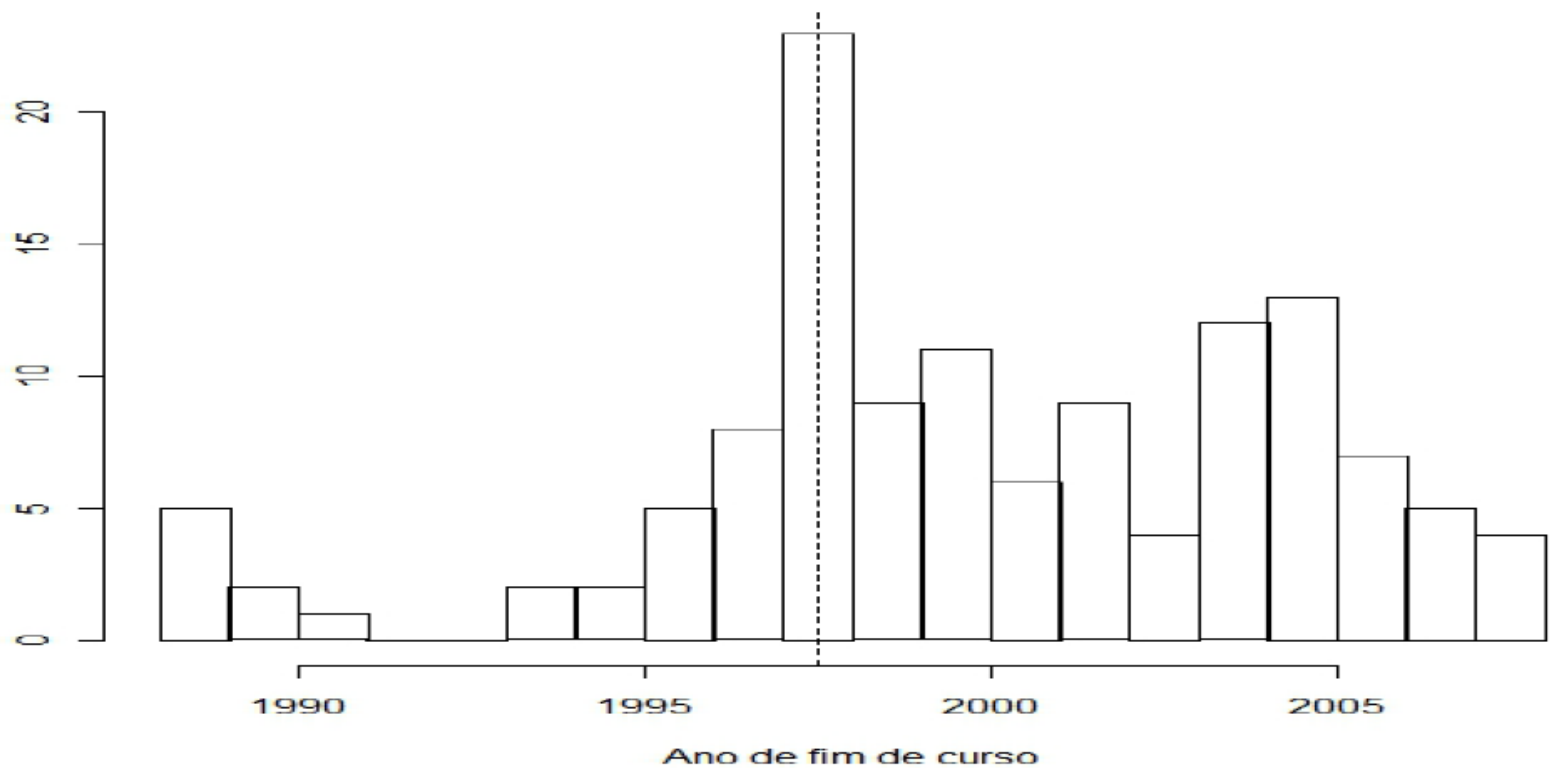
- 3 “mas ficaria mais à vontade se a pergunta fosse: Das cadeiras do tronco comum, qual aquela em que, se pudesses voltar atrás, gostarias de ter estado com mais atenção ou que tivesse sido ministrada de forma a ser mais útil?”
- 3 “Na realidade, todas (melhorei muito o meu snooker nessa altura); mas não sei se me teria sido muito útil, porque a maioria dessas disciplinas estão pouco adaptadas às nossas necessidades.”
- 3 “Acho que nenhuma em particular. O meu problema foi ter esquecido muita coisa desde então”
- 3 “Diria Bioestatística... mas só se tivesse sido reformulada para ser dada de outra forma.”
- 3 “Eu estive com muita atenção a todas as cadeiras!!! O que acontece é que na altura não soube tirar partido de algumas

## Respostas ao inquérito

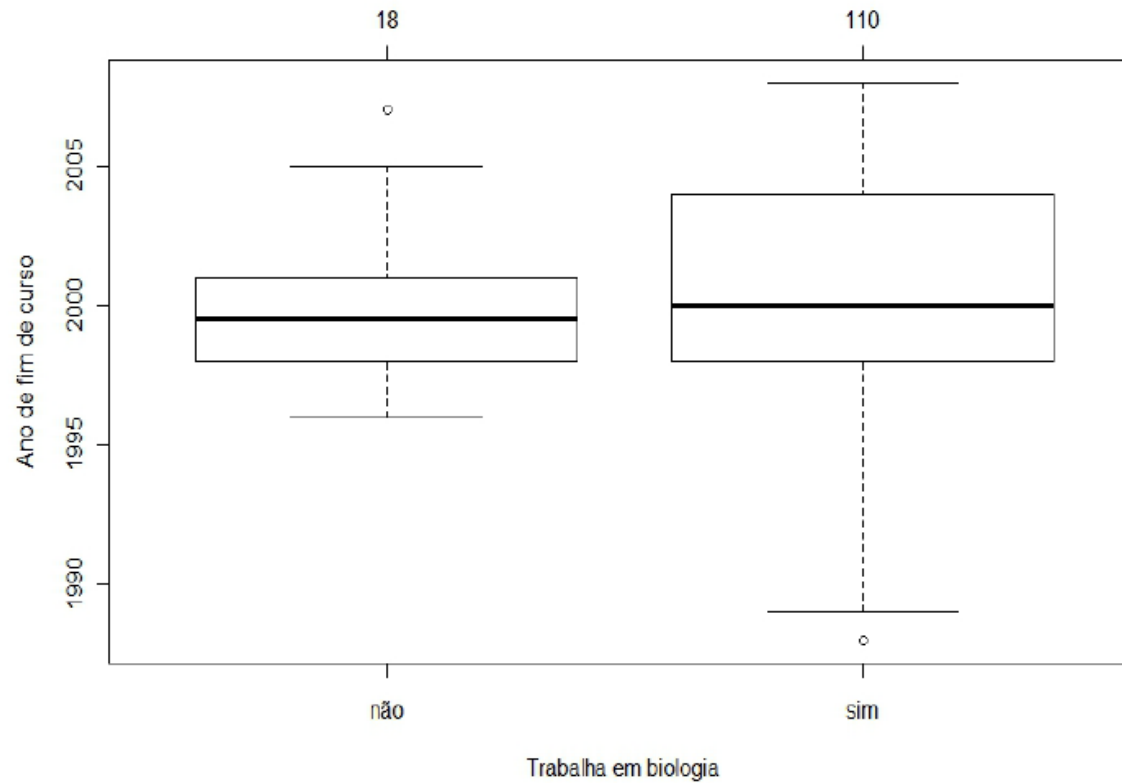
- 4 “por fazer actualmente divulgação de ciência alguma física, alguma química, mas as aprendizagens das cadeiras de matemática pouco ou nada me serviram.”
- 4 “Das do curso? Várias. Das mencionadas na pergunta, Estatística.”
- 4 “Das quatro opções que dás, são as estatísticas.”



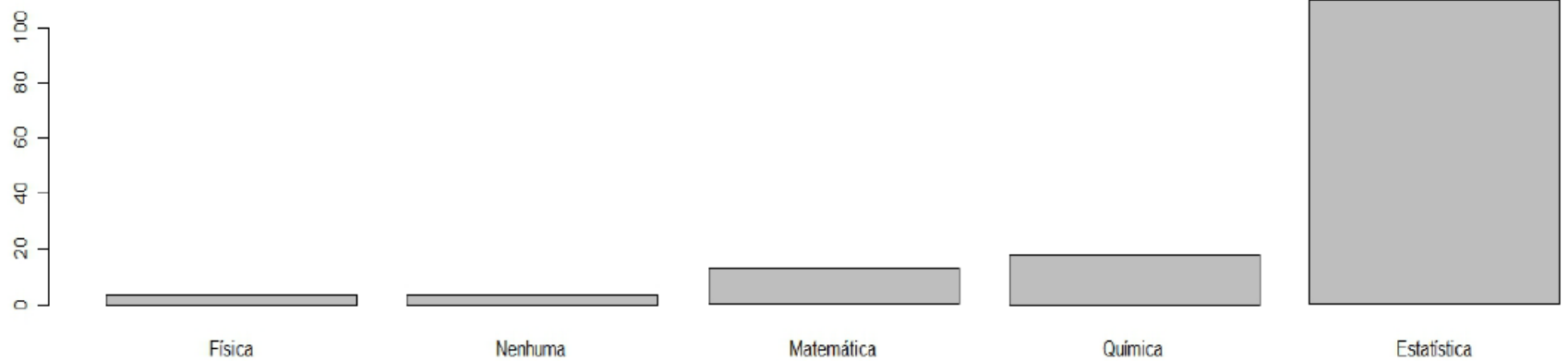
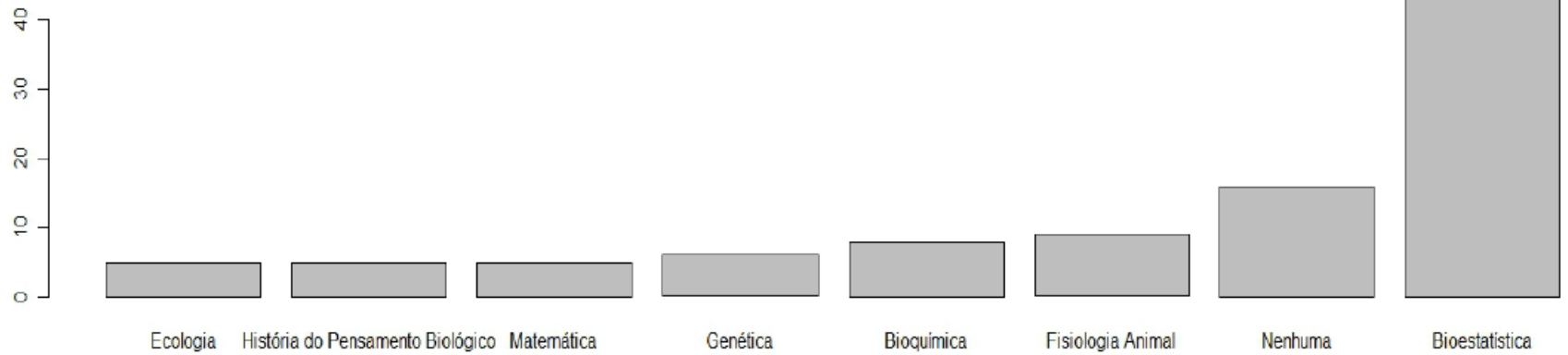
# Respostas ao inquérito - pergunta 1



# Respostas ao inquérito - Em "relação à relação" entre a pergunta 1 e 2



# Respostas ao inquérito - pergunta 3 e 4



# Modelos, modelos...

---

Para descrever relações entre organismos, ou entre estes e o seu meio ambiente vamos precisar de modelos.

*All models are wrong, but some models are useful* – George P. Box

Desde o início temos um problema de dimensionalidade: os padrões que observarmos são devidos à influência de outras espécies, de factores ambientais (ou de uma combinação complexa de ambos).

A realidade não é univariada, não é multivariada, é na realidade *hipervariada*.

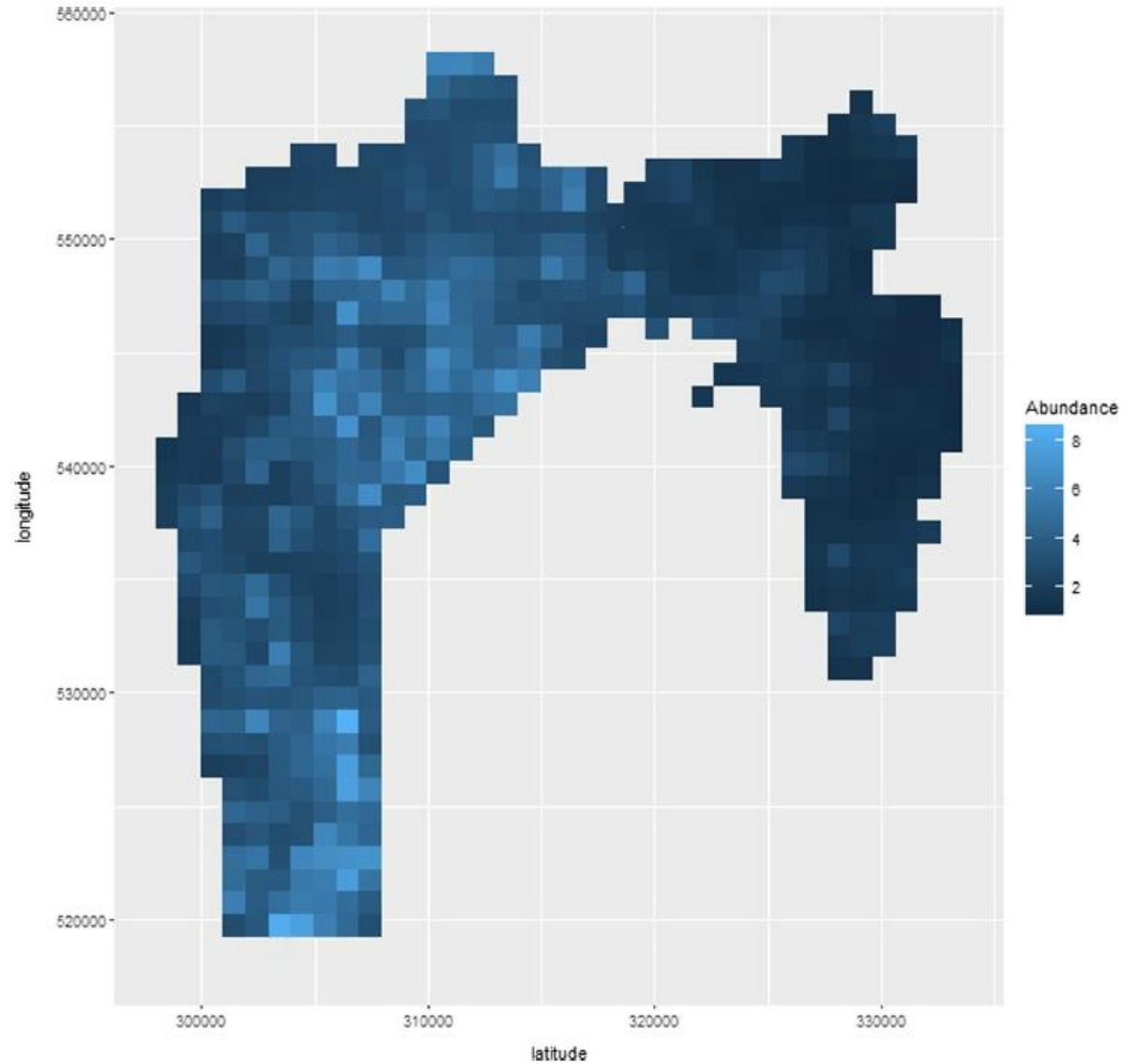


Fig. 4 Abundance distribution map of roe deer throughout our study area based on the DSM with geographical variables chosen for inference

Valente, A. M.; Marques, T. A.; Fonseca, C. & Torres, R. T. 2016 A new insight for monitoring ungulates: density surface modelling of roe deer in a Mediterranean habitat *European Journal of Wildlife Management* 62: 577-587



# Diferenças estatísticas e biológicas

---

Aumentando o tamanho da amostra até ao infinito, qualquer diferença se torna estatisticamente significativa...

Mas e então, o que quer isso dizer?

Há perguntas que não fazem grande sentido!

- As árvores da Floresta A são diferentes das da Floresta B? SÃO!
- Os peixes do rio C são diferentes dos do rio D? SÃO!
- A terra é redonda? NÃO!

É **apenas** o tamanho do efeito encontrado, o tamanho da diferença observada, que deve (e que pode!!!) ser interpretado!



## *Invited Paper:*

# THE ROLE OF HYPOTHESIS TESTING IN WILDLIFE SCIENCE

DOUGLAS H. JOHNSON,<sup>1</sup> U.S. Geological Survey, Northern Prairie Wildlife Research Center, Jamestown, ND 58401, USA

**Abstract:** Statistical testing of null hypotheses recently has come under fire in wildlife sciences (Cherry 1998; Johnson 1999; Anderson et al. 2000, 2001). In response to this criticism, Robinson and Wainer (2002) provide some further background information on significance testing; they argue that significance testing in fact is useful in certain situations. I counter by suggesting that such situations rarely arise in our field. I agree with Robinson and Wainer that replication is the key to scientific advancement. I believe, however, that significance testing and resulting *P*-values frequently are confused with issues of replication. Any single study can yield a *P*-value, but only consistent results from truly replicated studies will advance our understanding of the natural world.

***JOURNAL OF WILDLIFE MANAGEMENT 66(2):272–276***

**Key words:** effect size, hypothesis test, null hypothesis, replication, significance test.



# Effect size, confidence interval and statistical significance: a practical guide for biologists

Shinichi Nakagawa<sup>1,\*</sup> and Innes C. Cuthill<sup>2</sup>

<sup>1</sup> *Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK* (E-mail: itchyshin@yahoo.co.nz)

<sup>2</sup> *School of Biological Sciences, University of Bristol, Bristol BS8 1UG, UK* (E-mail: i.cuthill@bristol.ac.uk)

*(Received 2 January 2007; revised 24 July 2007; accepted 27 July 2007)*

## ABSTRACT

Null hypothesis significance testing (NHST) is the dominant statistical approach in biology, although it has many, frequently unappreciated, problems. Most importantly, NHST does not provide us with two crucial pieces of information: (1) the magnitude of an effect of interest, and (2) the precision of the estimate of the magnitude of that effect. All biologists should be ultimately interested in biological importance, which may be assessed using the magnitude of an effect, but not its statistical significance. Therefore, we advocate presentation of measures of the magnitude of effects (i.e. effect size statistics) and their confidence intervals (CIs) in all biological journals.



# In My Opinion

---

## The need to get the basics right in wildlife field studies

*David R. Anderson*

tion" (Romesburg 1981, 1989, 1991, 1993). My objective is to focus attention on 2 major problems that seem fundamental to much of what we do in wildlife field studies: 1) the frequent use of convenience sampling and 2) the use of index values (usually raw counts) purporting to measure "relative abundance." These problems result in a lack of

*Invited Paper:*

## **SUGGESTIONS FOR PRESENTING THE RESULTS OF DATA ANALYSES**

DAVID R. ANDERSON,<sup>1,2</sup> Colorado Cooperative Fish and Wildlife Research Unit, Room 201 Wagar Building, Colorado State University, Fort Collins, CO 80523, USA

WILLIAM A. LINK, U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, MD 20708, USA

DOUGLAS H. JOHNSON, U.S. Geological Survey, Northern Prairie Wildlife Research Center, Jamestown, ND 58401, USA

KENNETH P. BURNHAM,<sup>1</sup> Colorado Cooperative Fish and Wildlife Research Unit, Room 201 Wagar Building, Colorado State University, Fort Collins, CO 80523, USA

**Abstract:** We give suggestions for the presentation of research results from frequentist, information-theoretic, and Bayesian analysis paradigms, followed by several general suggestions. The information-theoretic and Bayesian methods offer alternative approaches to data analysis and inference compared to traditionally used methods. Guidance is lacking on the presentation of results under these alternative procedures and on nontesting aspects of classical frequentist methods of statistical analysis. Null hypothesis testing has come under intense criticism. We recommend less reporting of the results of statistical tests of null hypotheses in cases where the null is surely false anyway, or where the null hypothesis is of little interest to science or management.

***JOURNAL OF WILDLIFE MANAGEMENT 65(3):373-378***

**Key words:** AIC, Bayesian statistics, frequentist methods, information-theoretic methods, likelihood, publication guidelines.





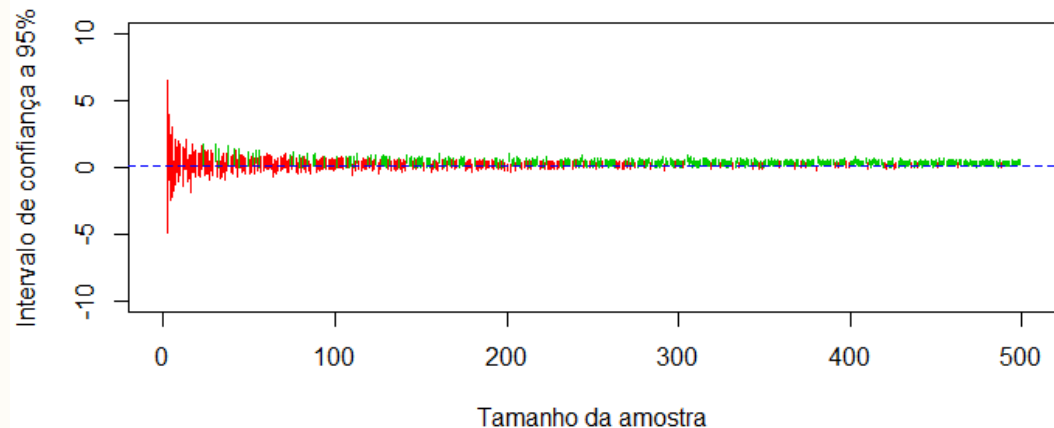
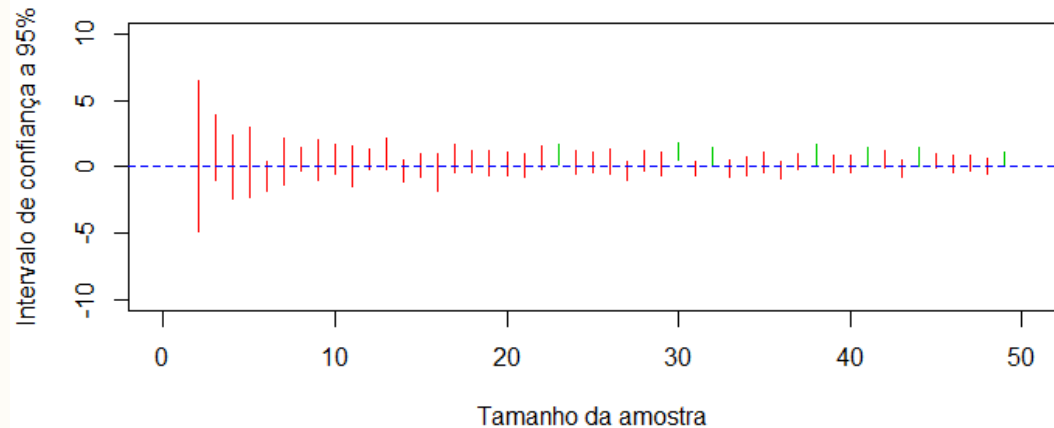
# Significância: estatística vs. biológica

---

- Toda e qualquer diferença se torna estatisticamente significativa desde que aumentemos suficientemente o tamanho da amostra
- Como tal, todas as diferenças serão no limite significativas
- Como podemos então interpretar uma diferença estatisticamente significativa?
- Qual a magnitude do efeito observado?
- Será biologicamente relevante que em media a espécie A produza 1.32 ovos e a espécie B 1.34 ovos, mesmo que isso seja estatisticamente significativo?

# Significância: estatística vs. biológica

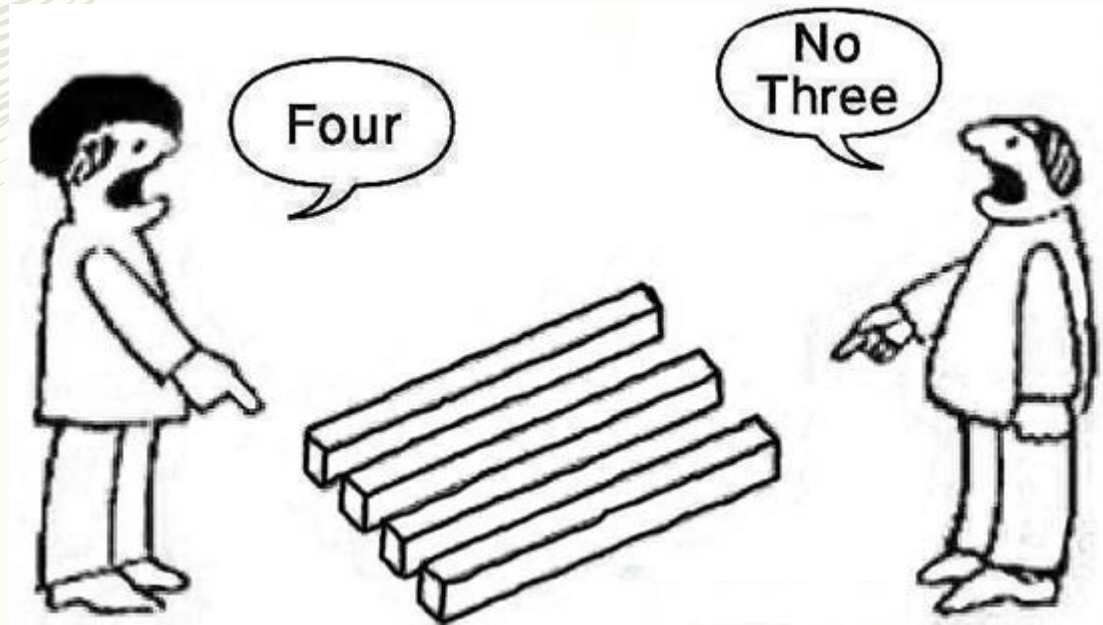
- Hipótese nula: média é 0
- Com tamanho baixo da amostra, não é fácil rejeitar
- Com amostra grande, é relativamente fácil
- (na realidade, a verdadeira média era 0.3!)
- À medida que o tamanho da amostra aumenta, conseguimos detectar a (verdadeira) diferença
- Mas será biologicamente relevante? É isso que faz com que seja sempre preciso um biólogo, o estatístico não sabe fazer isso!



# Fazer inferências sobre a realidade

---

- Mas o que é a realidade?







'Truth isn't truth,' says Rudy Giuliani. But nonsense is still nonsense

[fonte](#)



Mario Balotelli, playing for Manchester City, is shown a red card during a match against Arsenal.

- Twenty-nine teams involving 61 analysts used the same data set to address the same research question: whether soccer referees are more likely to give red cards to dark-skin-toned players than to light-skin-toned players.
- Analytic approaches varied widely across the teams
- Twenty teams (69%) found a statistically significant positive effect, and 9 teams (31%) did not observe a significant relationship.

# What is reality...?


---

- A single response (acorn count), three designed effects (species, site, and year) and 7 environmental variables
- “explain the variation in the response variable (acorn count) using the predictors available”
- responses from a skilled average self-reported statistical expertise of 6.7 on scale of 1 [low] to 10 [high]) diverse group of 24 ecologists
- no two final models included exactly the same set of predictors
- not a single predictor was included in every final model

# O acaso e a probabilidade



Se observar um pássaro ao acaso, qual é a probabilidade de não ver amarelo?



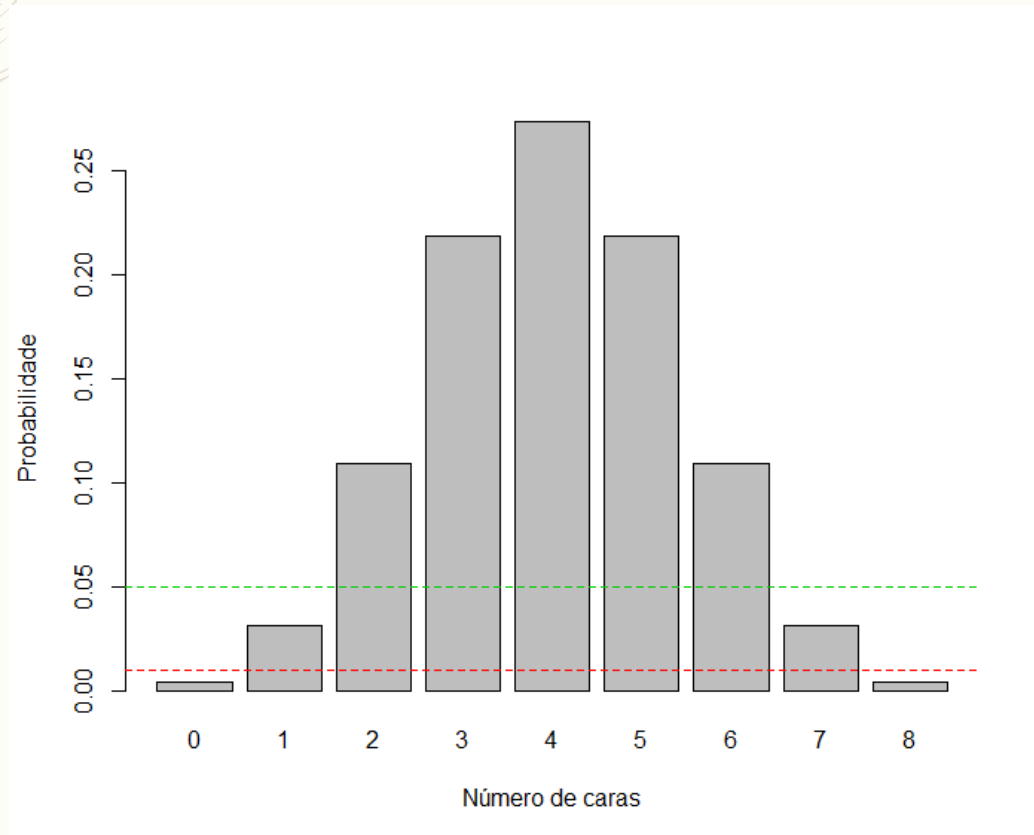
# Como avaliar se algo que observamos pode ser mais do que fruto do acaso?

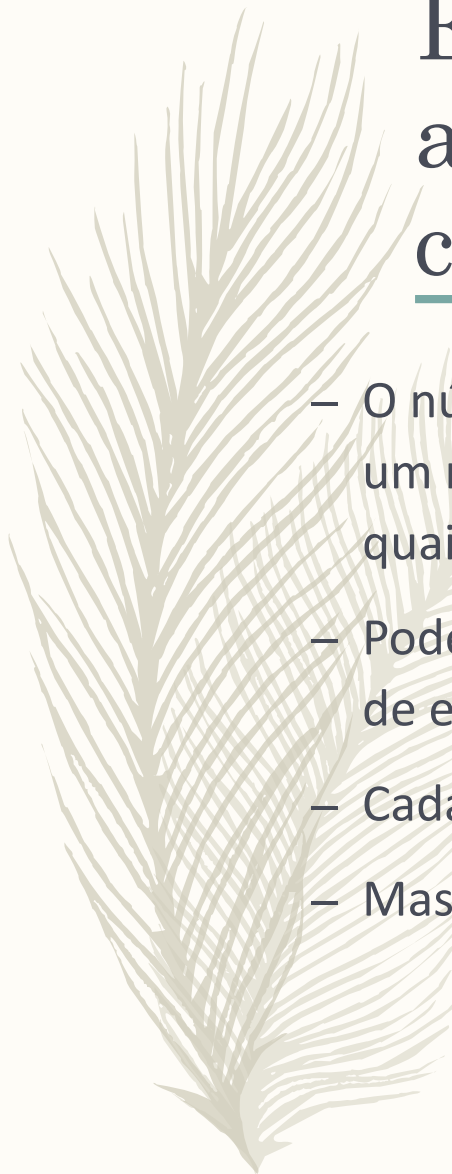
---

- Primeiro há que definir o que entendemos por acaso
- Qual a hipótese nula que assumimos?
  - As árvores do local A são iguais em tamanho às do local B
  - A temperatura não influencia o crescimento dos animais
  - A minha moeda não está viciada
- Depois recolhemos dados que contenham informação relevante sobre a hipótese formulada
- Por fim rejeitamos, ou não, a hipótese colocada

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

---





# 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>

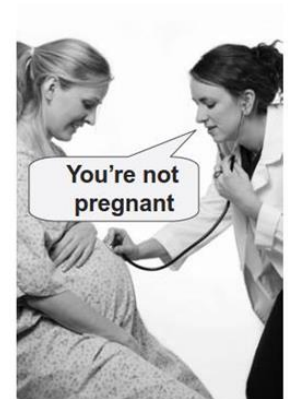
Na realidade a questão sobre qual é pior, um erro de tipo I ou II, é puramente filosófica, as opiniões podem divergir, e só pode ser avaliada caso a caso.

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

**Type I error**  
(false positive)



**Type II error**  
(false negative)



# Natureza e filtros



amostragem  
observação

Inferência



# Natureza e filtros

