VEROSSIMILHANÇA MÁXIMA E SELEÇÃO DE MODELOS

Hipótese (θ_1)

Dado honesto

Qual a probabilidade de arremessarmos um dado **5 vezes** e obtermos **6 em todos os arremessos (x)**?

Hipótese (θ_1)

Dado honesto

$$\boldsymbol{P(x|\theta_1)} = P(P_1 * P_2 * \dots P_5)$$

$$(1/6)^5 = 1/7776$$

Hipótese (θ_1)

Dado honesto

$P(x|\theta_1) = 1/7776 = 0.00012$

Hipótese (θ_1)

Dado honesto

$P(x|\theta_1) = 1/7776 = 0.00012$

Hipótese (θ_2)

Dado viciado!!

+ VEROSSÍMIL

$P(x|\theta_2) = (0.8)^5 = 0.33$

$$\boldsymbol{L(\theta|x) = P(x|\theta)}$$

A **verossimilhança (*L*)** de um conjunto de **parâmetros (θ)**, dada uma **observação (x)** é igual a probabilidade desta mesma observação ter ocorrido dado os valores daqueles parâmetros.

$$L(\theta|x) = P(x|\theta)$$

$$\neq P(\theta|x)!!$$

A **verossimilhança (*L*)** de um conjunto de **parâmetros (θ)**, dada uma **observação (x)** é igual a probabilidade desta mesma observação ter ocorrido dado os valores daqueles parâmetros.

Observação (x): Barulhos vindos do forro da sua casa.

Hipótese (θ_1): *Gremlins* estão dando uma festa!

Observação (x): Barulhos vindos do forro da sua casa.

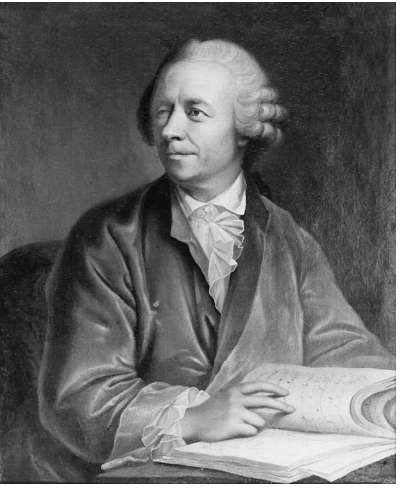
Hipótese (θ_1): *Gremlins* estão dando uma festa!

$P(x|\theta_1)$ é altamente **provável e verossímil** $L(\theta_1|x)$ ...

... mas $P(\theta_1|x)$ é altamente **improvável**

Daniel Bernoulli (1700-1782)

$$L(\theta|x) = P(x|\theta)$$

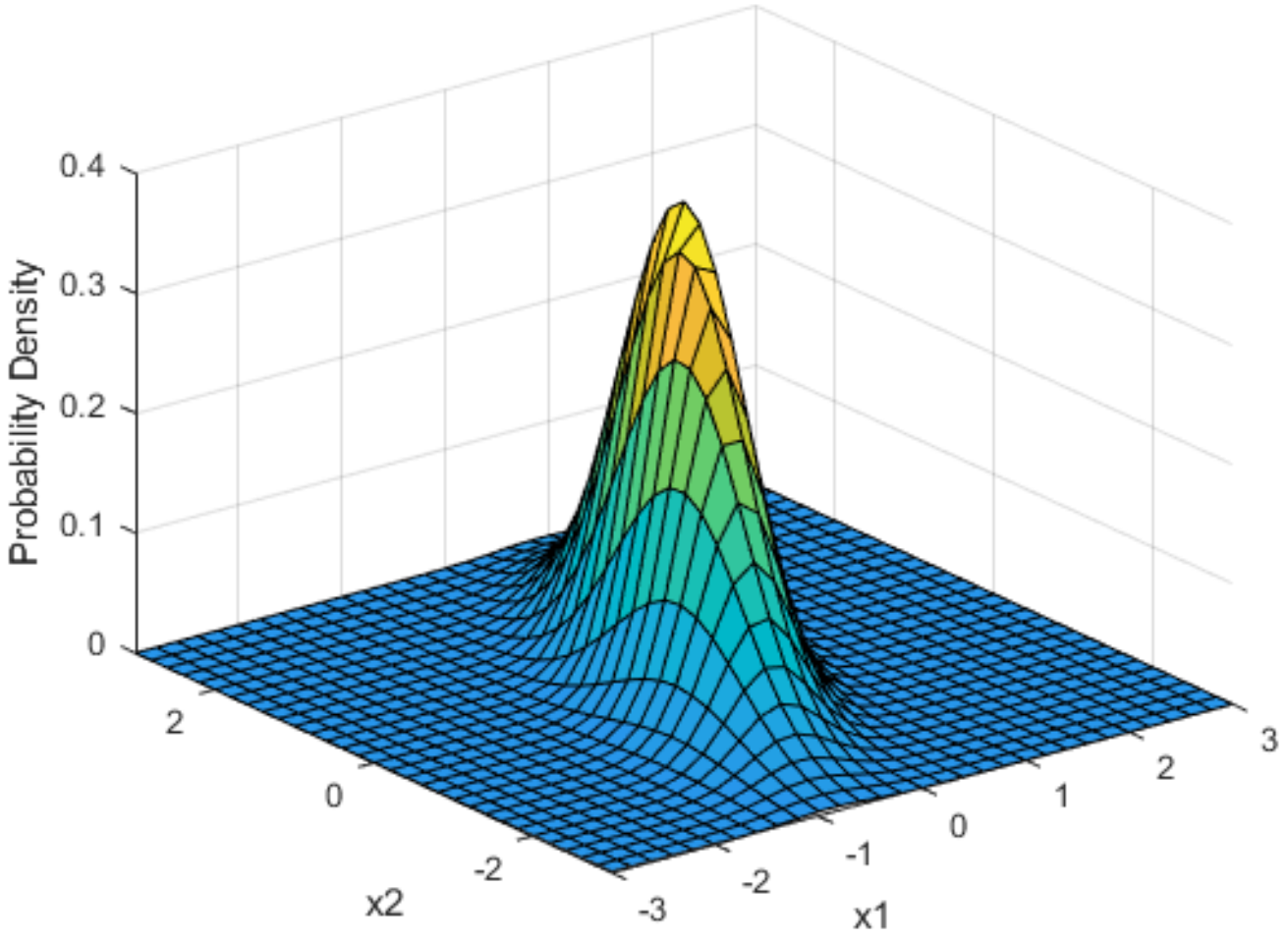


Leonhard Euler (1707-1783)

A verossimilhança pode ser usada para **estimar** um ou mais parâmetros (θ)

Ronald Fisher (1890-1962)

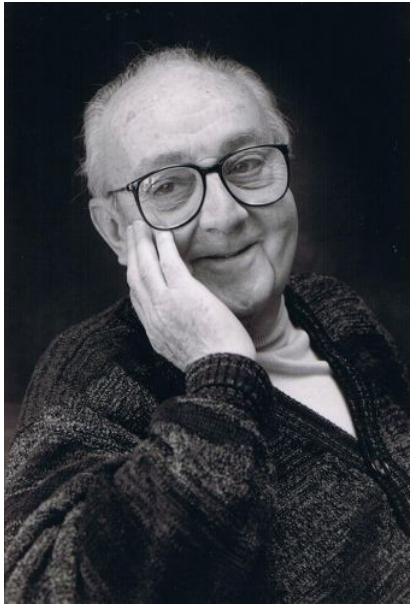
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$


Probability Density
0.4
0.3
0.2
0.1
0
2
0
-2
x2
-3
-2
-1
0
1
2
3
x1

Então... o modelo com **mais parâmetros** (+ ajustados)
é sempre o **mais verossímil**?

Mas afinal... o que são **modelos**?

“Essentially,
all models are wrong,
but some are useful.”



George Box (1919-2013)

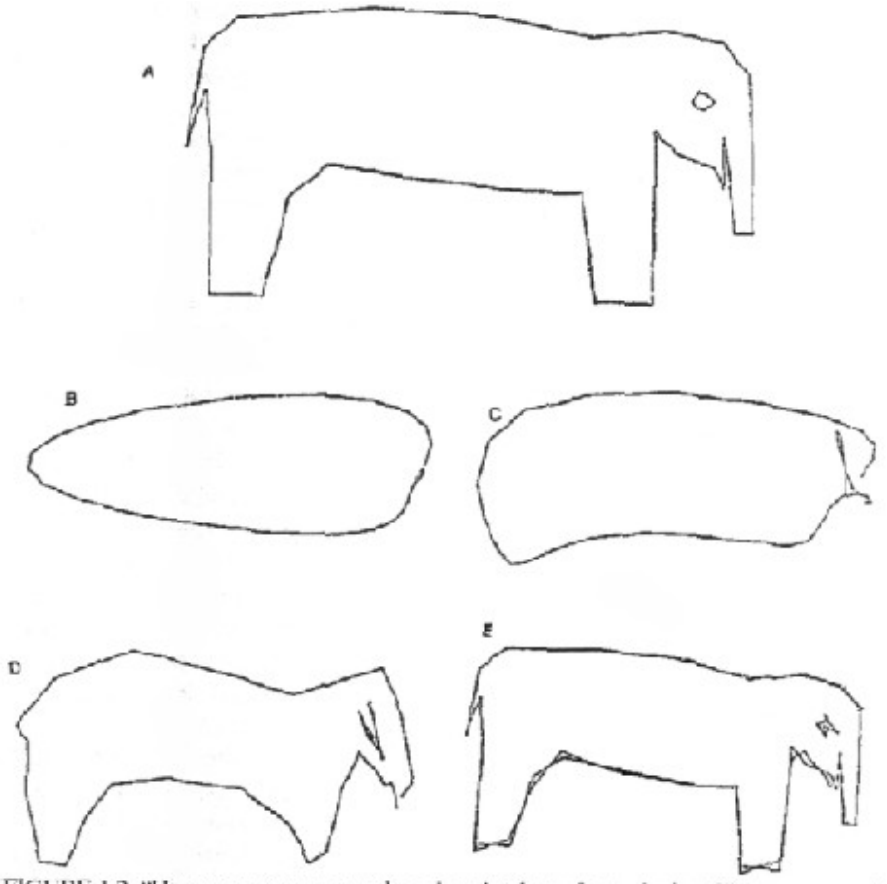


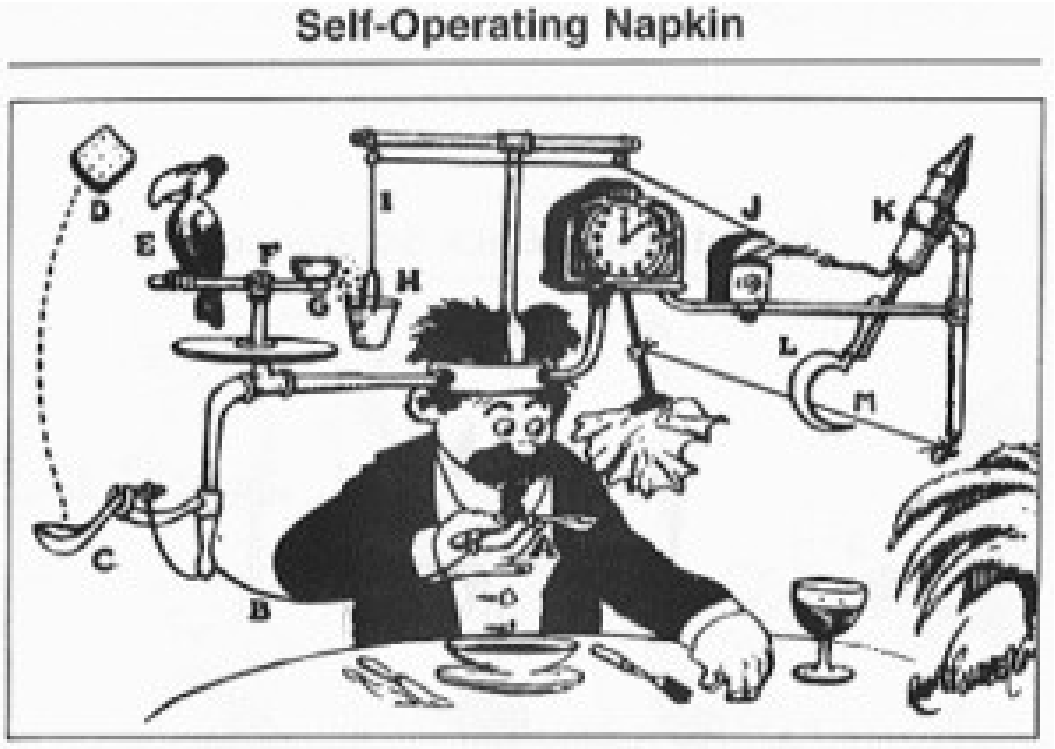
FIGURE 1.2. "How many parameters does does it take to fit an elephant?" was answered by Wei (1975). He started with an idealized drawing (A) defined by 36 points and used least squares Fourier sine series fits of the form $x(t) = \alpha_0 + \sum \alpha_i \sin(it\pi/36)$ and $y(t) = \beta_0 + \sum \beta_i \sin(it\pi/36)$ for $i = 1, \ldots, N$. He examined fits for $K = 5, 10, 20$, and 30 (shown in B–E) and stopped with the fit of a 30 term model. He concluded that the 30-term model "may not satisfy the third-grade art teacher, but would carry most chemical engineers into preliminary design."

A navalha de Occam

e o Princípio da parsimônia

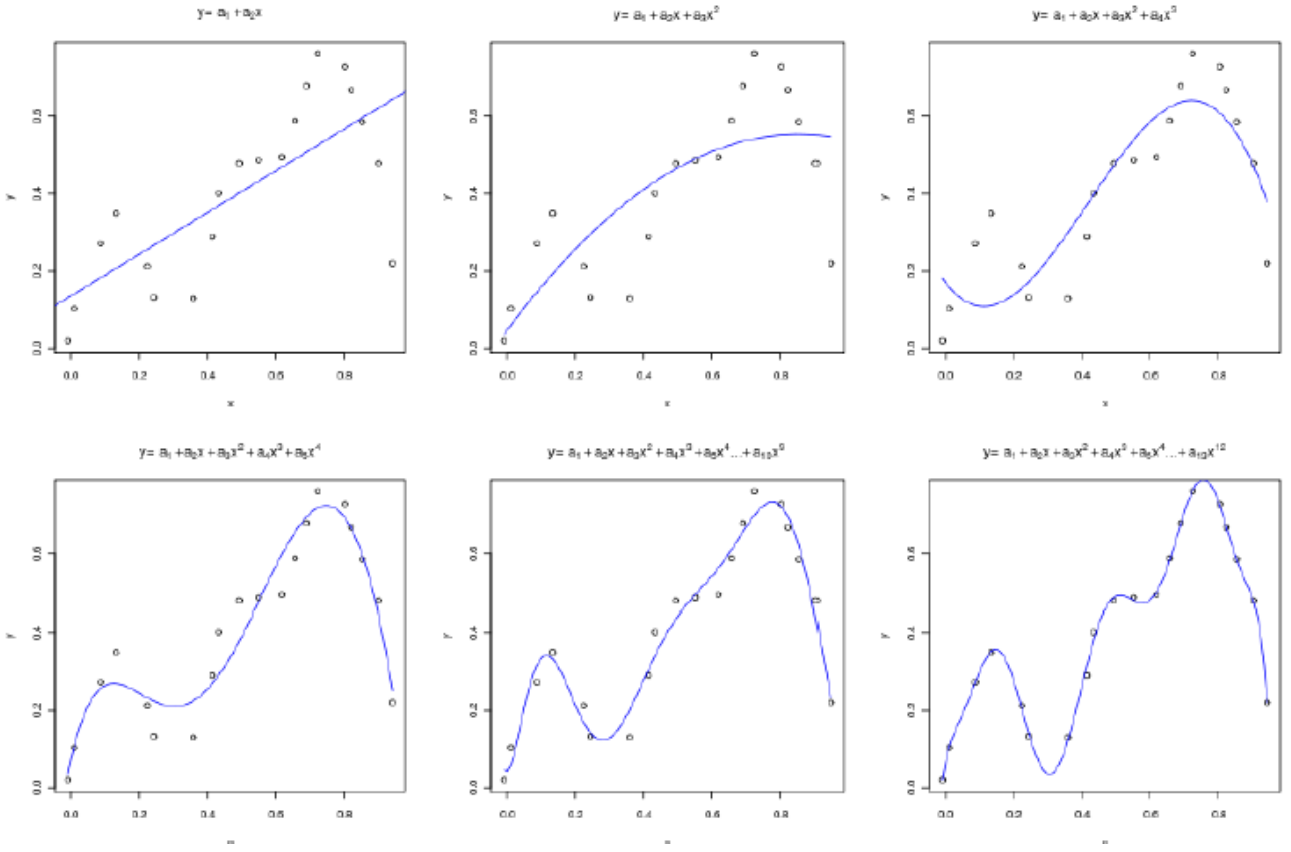
William de Occam (1285-1347)

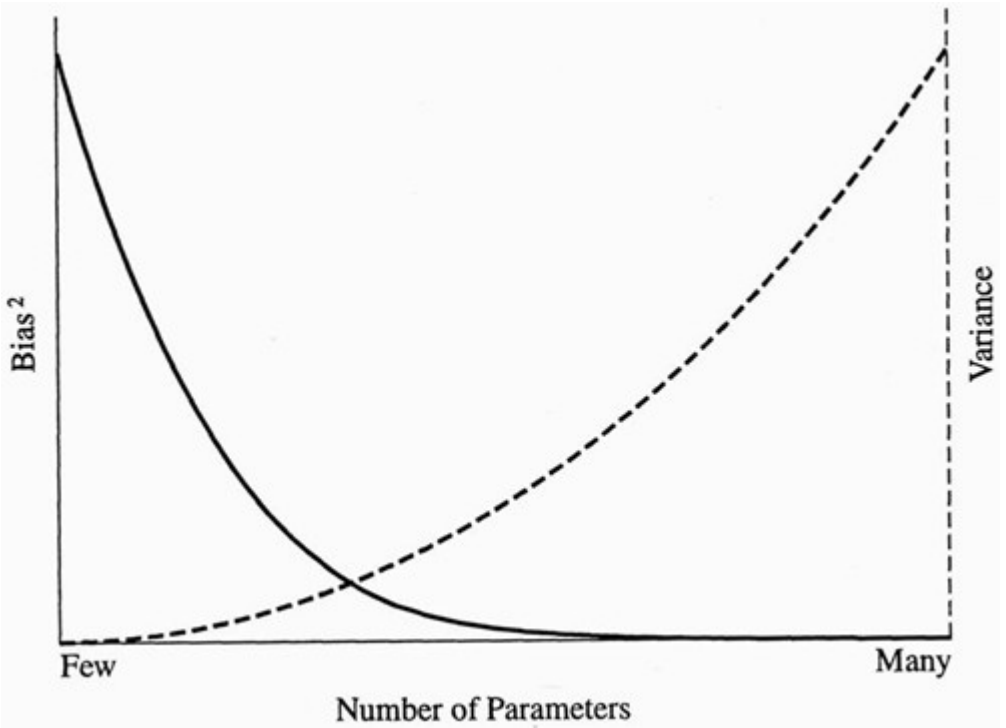
When you have two competing theories that make exactly the same predictions, the simpler one is the better.

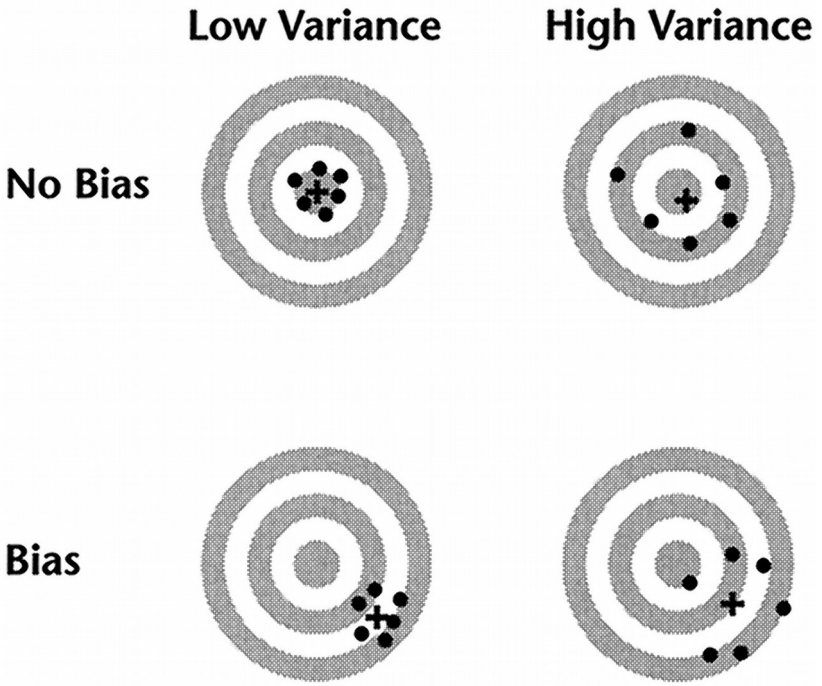


Rube Goldberg

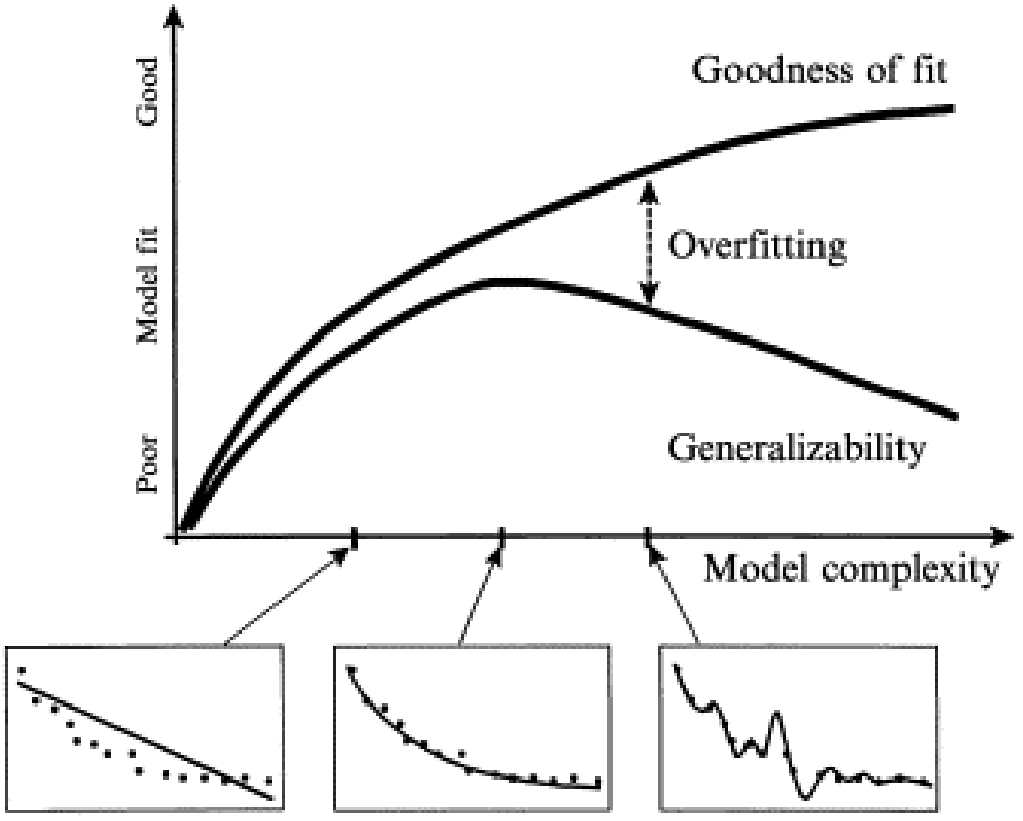
Sobreajuste (Overfitting)




Bias²
Variance
Few
Many
Number of Parameters


Low Variance
High Variance
No Bias
Bias

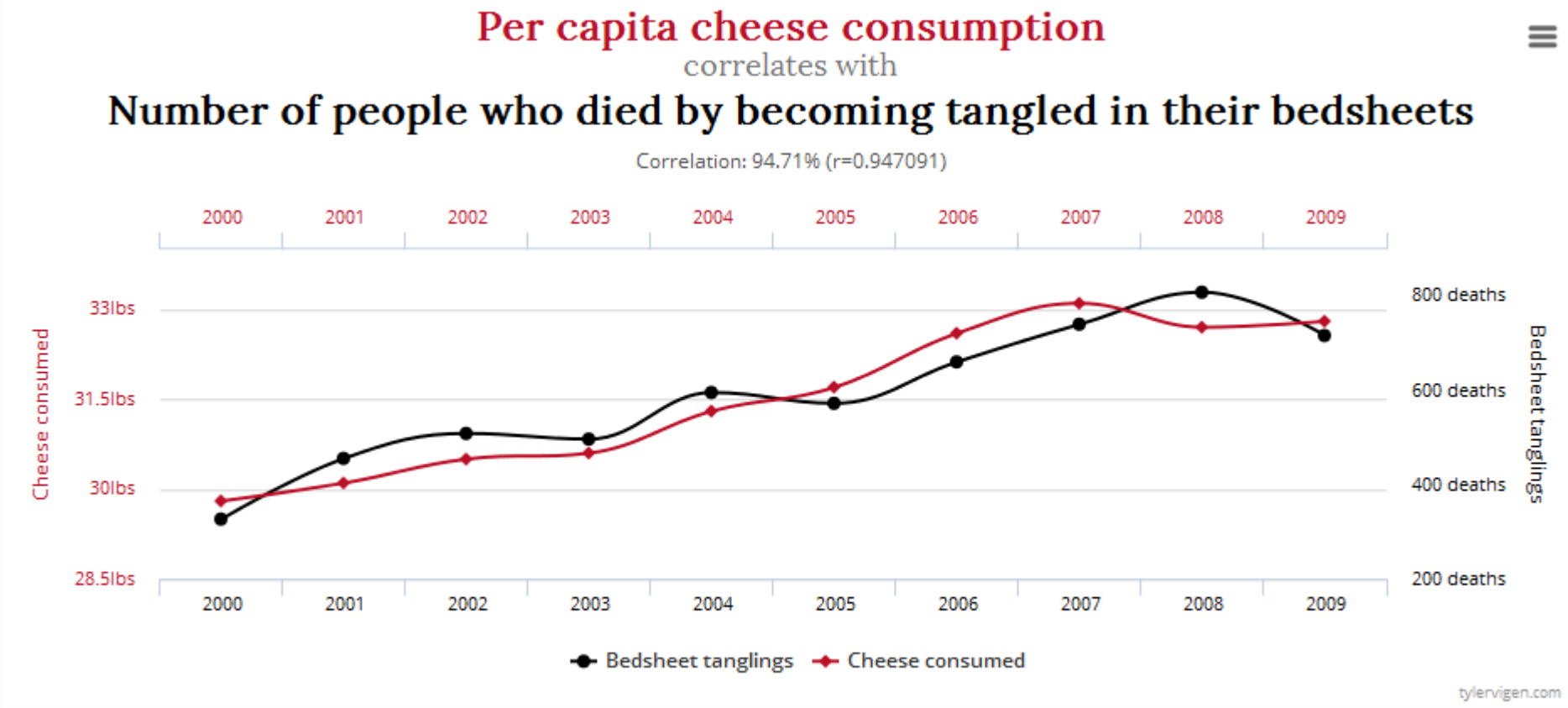
Na ecologia,
não queremos **apenas explicar** mas tambem **prever**.



Dragagem de dados (*data dredging*)

"**Data dredging** is the use of **data mining** to uncover patterns in data that can be presented as statistically significant <u>without first devising a specific hypothesis</u> as to the underlying causality"

https://www.geckoboard.com/learn/data-literacy/statistical-fallacies/data-dredging/



http://tylervigen.com/spurious-correlations

Teoria da Informação + Verossimilhança

$$AIC = -2\ln(L) + 2k$$

Hirotugu Akaike (1927-2009)

Modelo	k	ln(*L*)	AIC	Δ_i	ω_i
x_1	β0, β1	-27	58	0	0.6224
$x_1 + x_2$	β0, β1, β3	-26,5	59	1	0.3775
x_2	β0, β1	-78	92	34	0.0001

Hirotugu Akaike (1927-2009)

$$\Delta_i = AIC_i - AIC_{min}$$

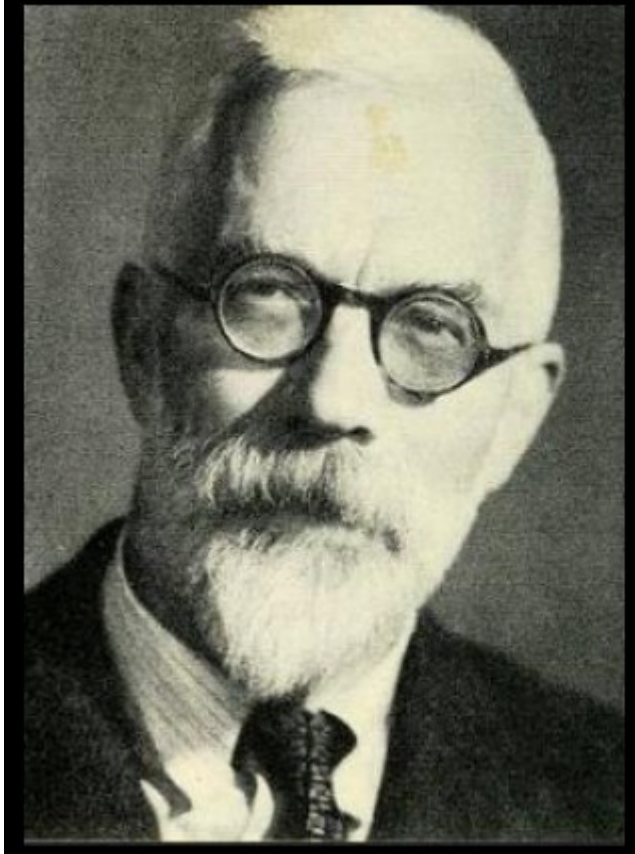
$$w_i = \frac{\exp\left(\frac{-1}{2}\Delta_i\right)}{\sum_{r=1}^{R}\exp\left(\frac{-1}{2}\Delta_r\right)}$$

Checklist

1. Defina claramente a sua **pergunta de estudo**;
2. Selecione seu conjunto de modelos/covariáveis) com base no seu **conhecimento biológico/ecológico**;
3. Seja **parcimonioso(a)**;

Checklist

1. NÃO compare amostras/variáveis dependentes diferentes;
2. NÃO compare variáveis dependentes em escalas diferentes!
3. Lembre-se de que não está testando significância ou rejeitando hipóteses.



To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of.

- Ronald Fisher -

quoteparrot.com