Curso de estatística básica

Aula 09

Modelos mistos e seleção de modelos

Pavel Dodonov
pdodonov@gmail.com
anotherecoblog.wordpress.com

Premissas de modelos lineares

Normalidade

Homogeneidade de variâncias

Variável explanatória fixa (sem erro)

Relações lineares

Independência

Premissas de modelos lineares

Normalidade
GLM (distribuições não-normais)
Transformar os dados

Homogeneidade de variâncias

Variável explanatória fixa (sem erro)

Relações lineares

Independência

Premissas de modelos lineares

Normalidade
GLM (distribuições não-normais)
Transformar os dados

Variável explanatória fixa (sem erro)

Independência

Homogeneidade de variâncias
Modelar a variância (GLS)
Transformar os dados

Relações lineares

Premissas de modelos lineares

Normalidade
GLM (distribuições não-normais)
Transformar os dados

Variável explanatória fixa (sem erro)
Bootstrap

Independência

Homogeneidade de variâncias
Modelar a variância (GLS)
Transformar os dados

Relações lineares

Premissas de modelos lineares

Normalidade
GLM (distribuições não-normais)
Transformar os dados

Homogeneidade de variâncias
Modelar a variância (GLS)
Transformar os dados

Variável explanatória fixa (sem erro)
Bootstrap

Relações lineares
Modelos aditivos, modelos não-lineares

Independência

Premissas de modelos lineares

Normalidade
GLM (distribuições não-normais)
Transformar os dados

Variável explanatória fixa (sem erro)
Bootstrap

Independência
Modelos mistos

Homogeneidade de variâncias
Modelar a variância (GLS)
Transformar os dados

Relações lineares
Modelos aditivos, modelos não-lineares

Premissas de modelos lineares

Normalidade
GLM (distribuições não-normais)
Transformar os dados

Variável explanatória fixa (sem erro)
Bootstrap

Independência
Modelos mistos

Homogeneidade de variâncias
Modelar a variância (GLS)
Transformar os dados

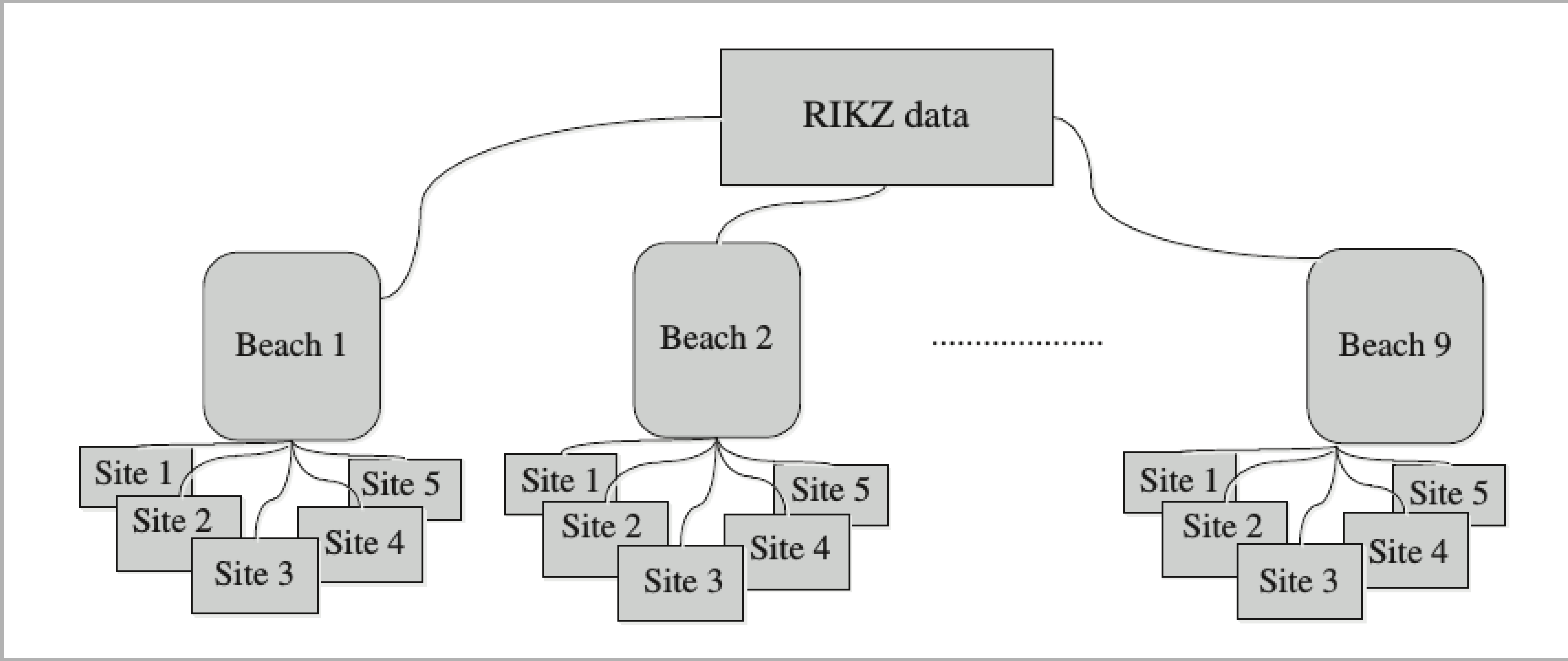
Relações lineares
Modelos aditivos, modelos não-lineares

Modelos mistos

Dados hierárquicos ou em múltiplos níveis –
existe algum tipo de pseudoreplicação

Modelos mistos

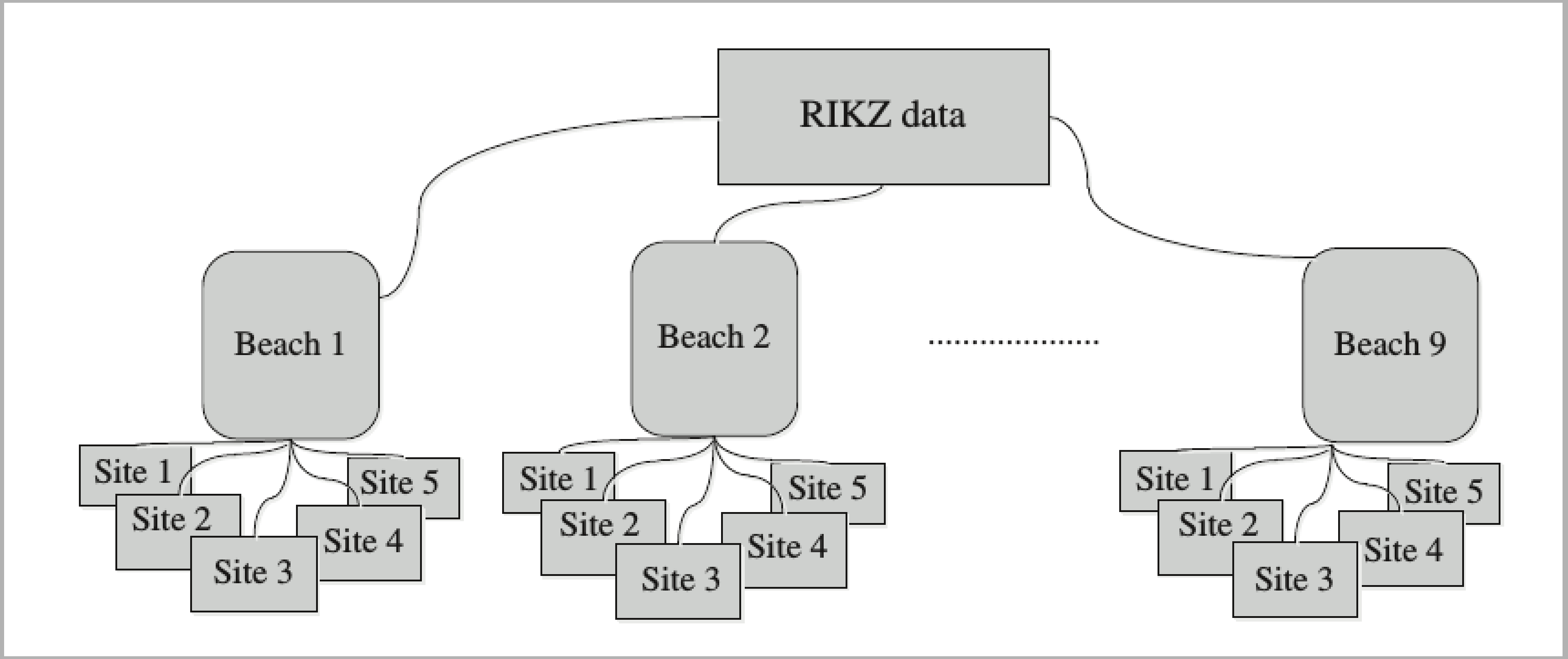
Dados hierárquicos ou em múltiplos níveis –
existe algum tipo de pseudoreplicação



Zuur et al. 2009

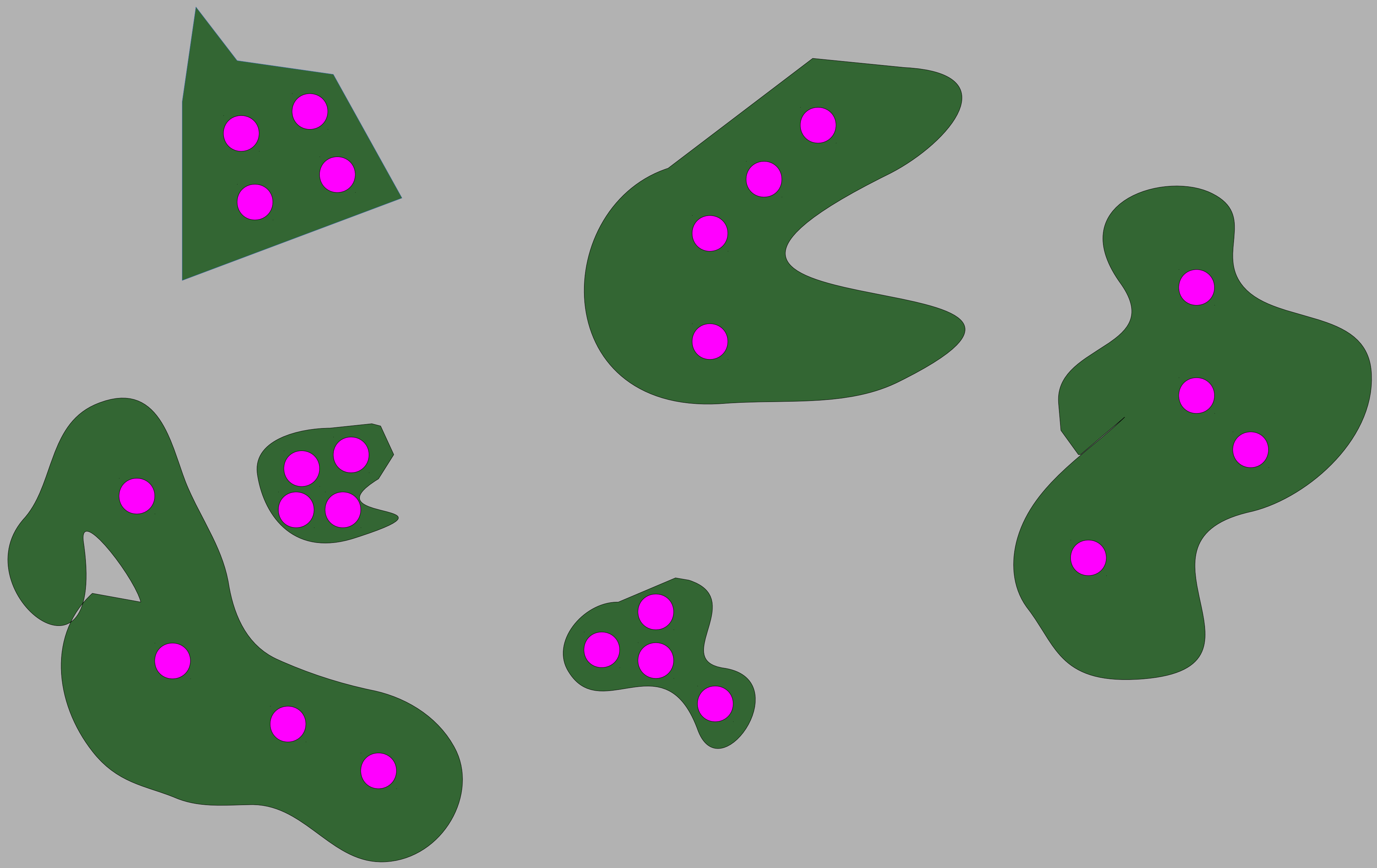
Modelos mistos

Dados hierárquicos ou em múltiplos níveis – existe algum tipo de pseudoreplicação

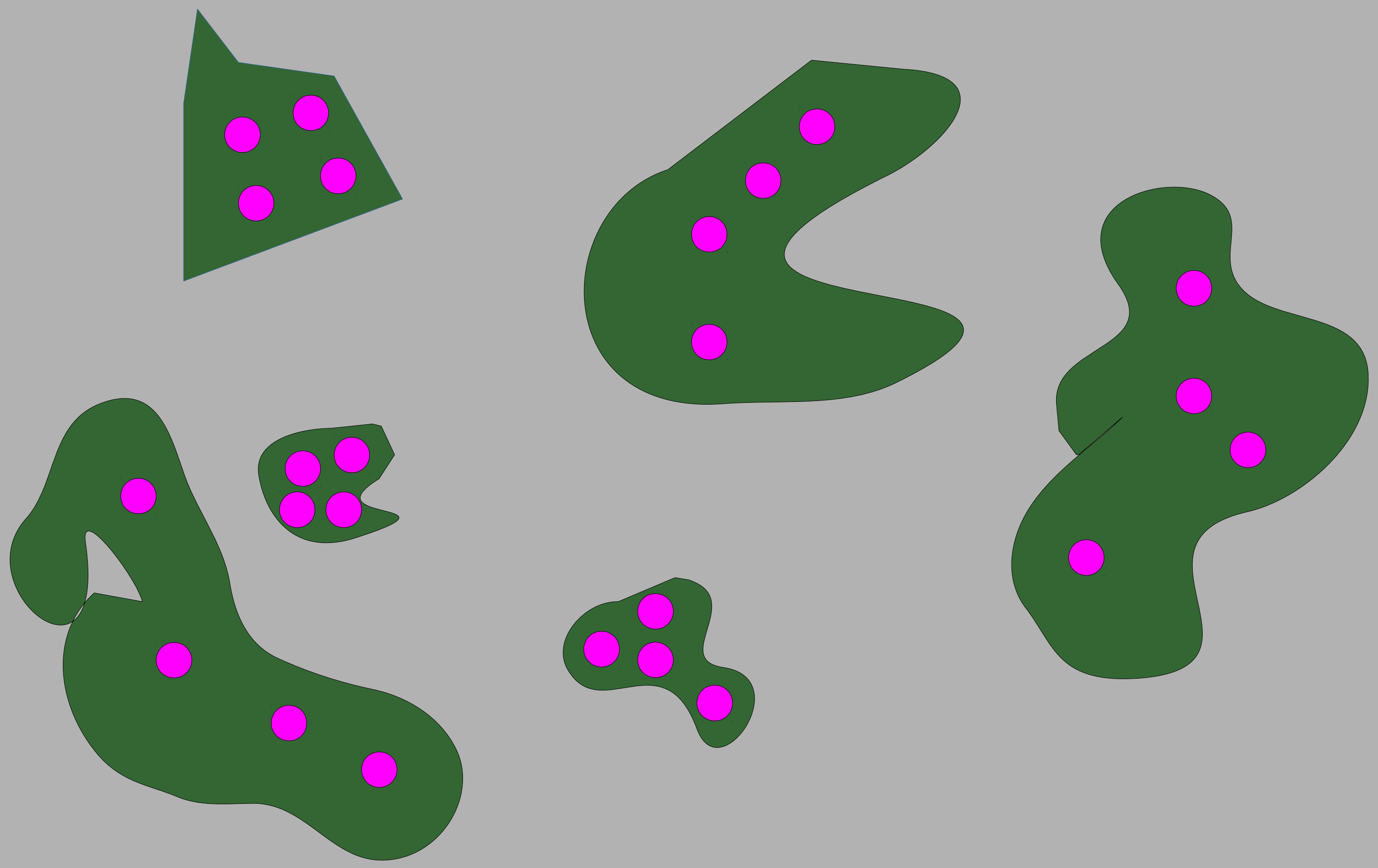


Zuur et al. 2009

Modelos mistos

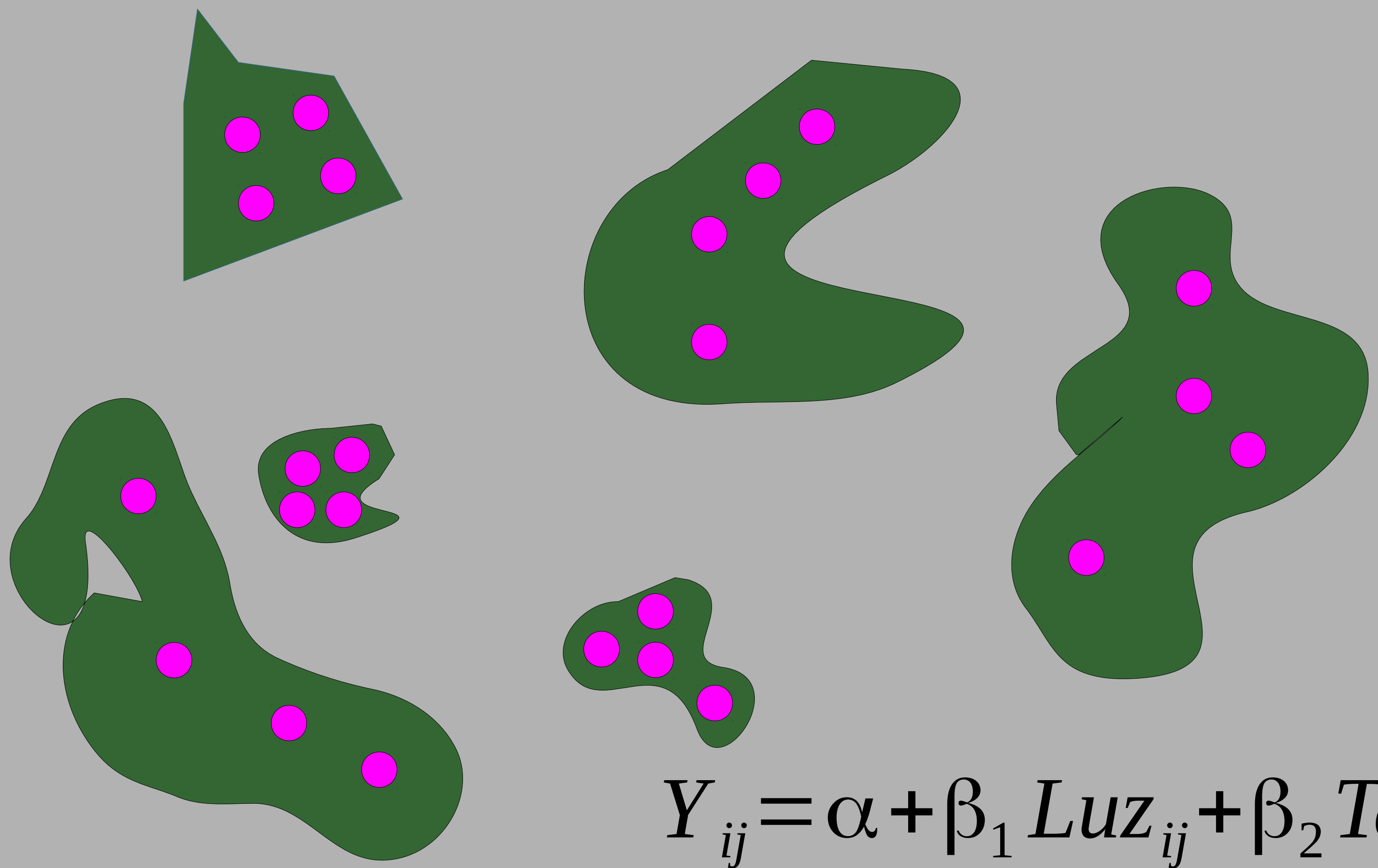


Modelos mistos



Biodiversidade ~ Tamanho do fragmento + Luminosidade na parcela

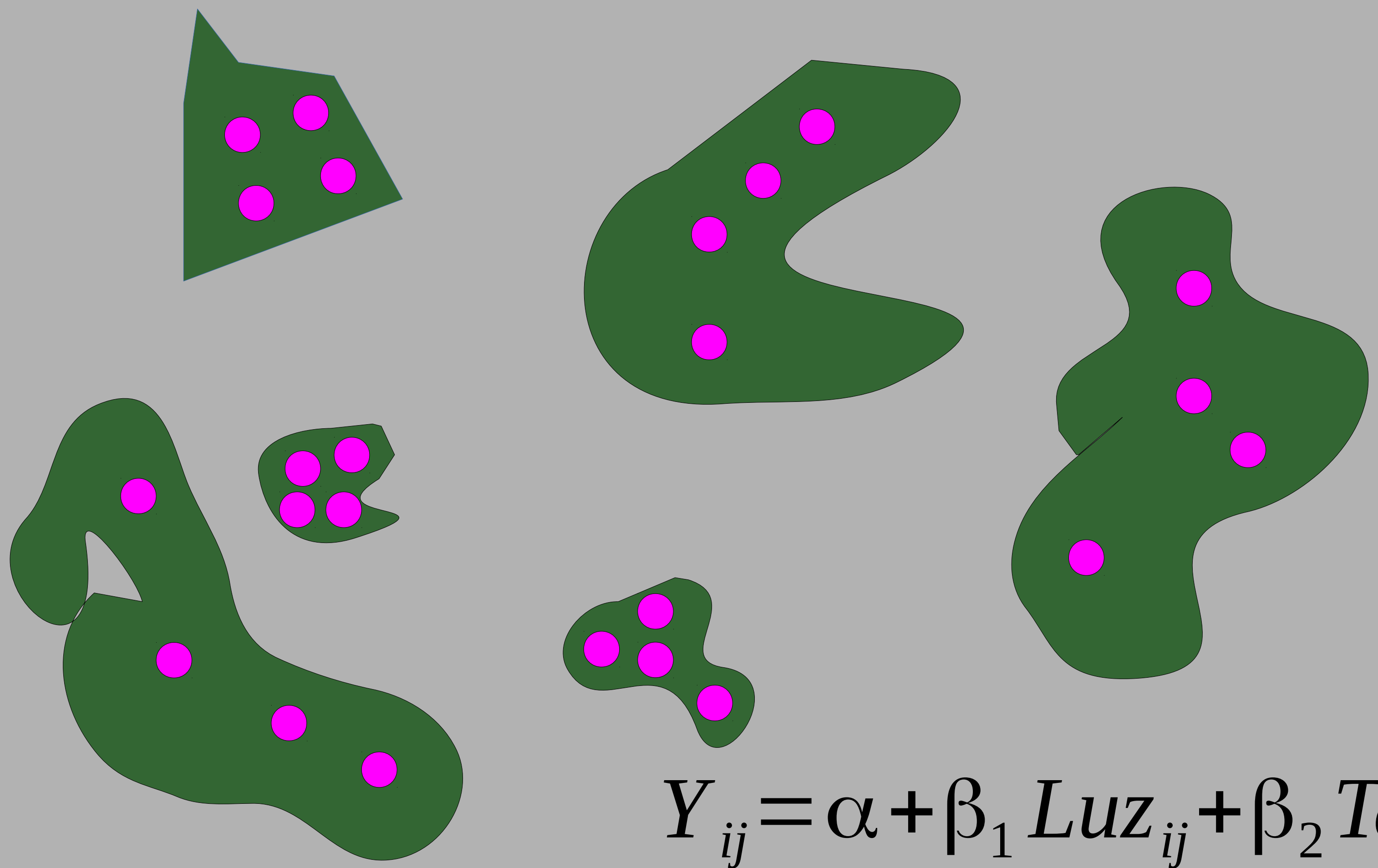
Modelos mistos



$$Y_{ij} = \alpha + \beta_1 Luz_{ij} + \beta_2 Tamanho_j + \epsilon_{ij}$$

Biodiversidade ~ Tamanho do fragmento + Luminosidade na parcela

Modelos mistos



$$Y_{ij} = \alpha + \beta_1 Luz_{ij} + \beta_2 Tamanho_i + \epsilon_{ij}$$

ij: j-ésima parcela do i-ésimo fragmento
i: i-ésimo fragmento
Biodiversidade ~ Tamanho do fragmento + Luminosidade na parcela

Análise em duas etapas

- 1) Aplicar uma regressão linear dentro de cada fragmento
  - Vai resultar em um coeficiente β_1 para cada fragmento

Y_{ij}

Análise em duas etapas

- 1) Aplicar uma regressão linear entre Y e luz dentro de cada fragmento
  - Vai resultar em um coeficiente β_1 para cada fragmento

Y_{ij}

- 2) Aplicar uma regressão linear entre estes coeficientes e o tamanho do fragmento

Análise em duas etapas

- Desvantagens
  - Resume os pontos do mesmo fragmento com um único valor

$$Y_{ij}$$

Análise em duas etapas

- Desvantagens
  - Resume os pontos do mesmo fragmento com um único valor
  - No segundo passo, trabalha com parâmetros de regressão, não com os dados observados Y_{ij}

Análise em duas etapas

- Desvantagens
  - Resume os pontos do mesmo fragmento com um único valor
  - No segundo passo, trabalha com parâmetros de Y_{ij} regressão, não com os dados observados
  - O tamanho amostral dentro de cada fragmento não é usado na segunda etapa

Análise em duas etapas

- Desvantagens
  - Resume os pontos do mesmo fragmento com um único valor
  - No segundo passo, trabalha com parâmetros de Y_{ij} regressão, não com os dados observados
  - O tamanho amostral dentro de cada fragmento não é usado na segunda etapa
    - (Mas N maior ainda fornece melhores estimativas do parâmetro --> menor erro amostrar --> é legal.)

Modelos mistos

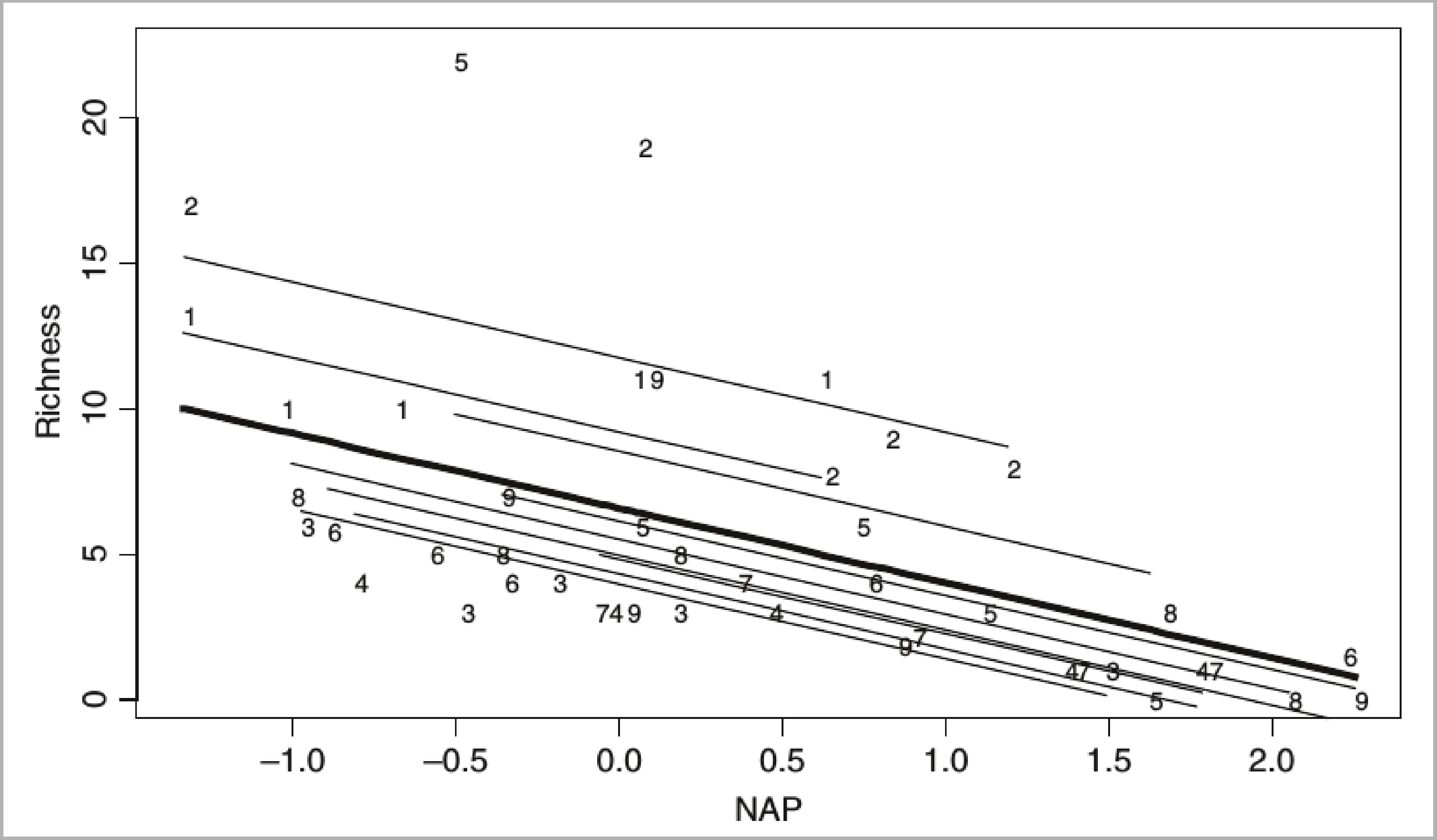
Junta os dois passos em um só

Misto por ter um componente fixo e um aleatório (que varia entre os fragmentos, etc)

Modelo do intercepto aleatório

- O intercepto varia aleatoriamente entre os fragmentos (segue uma distribuição normal)

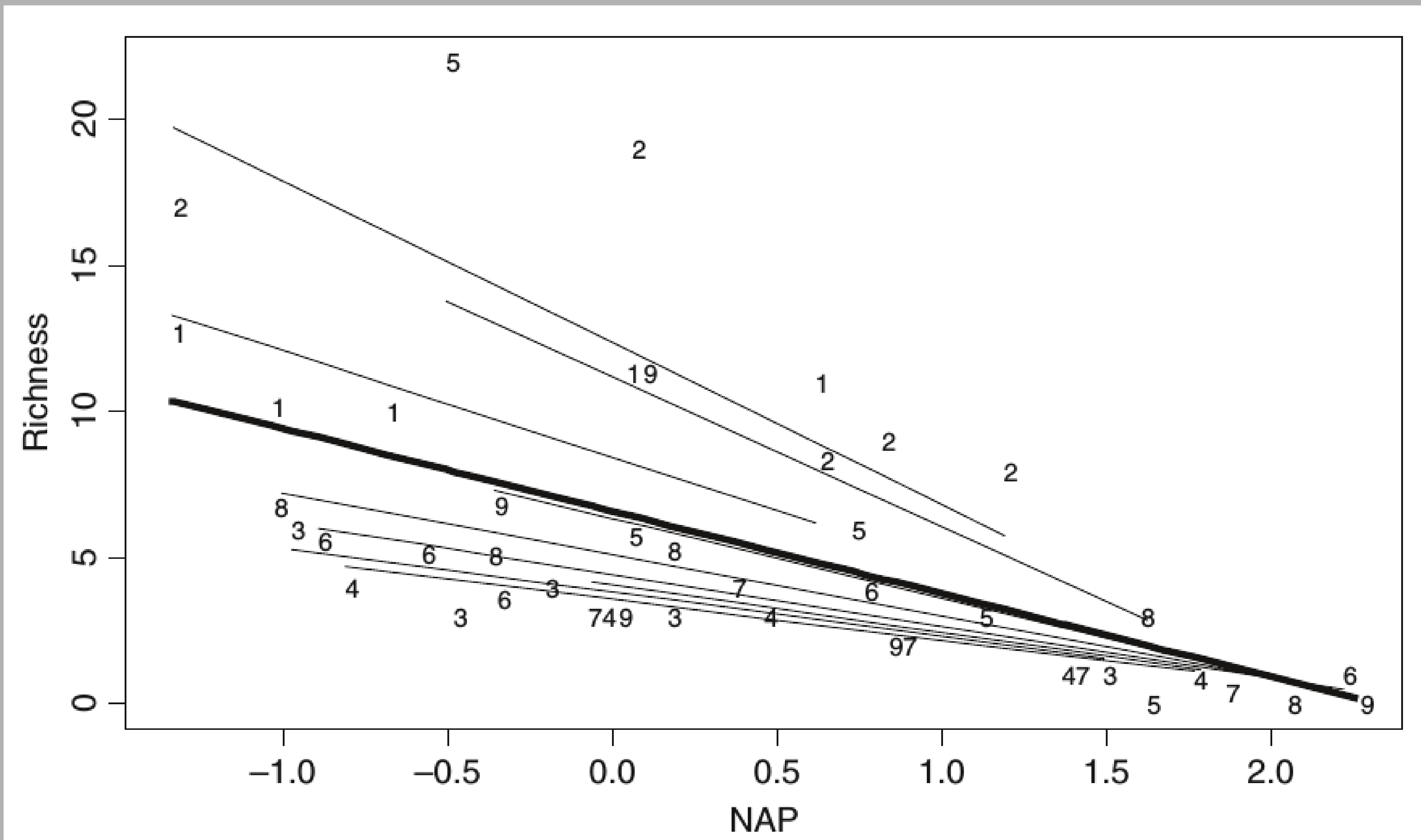
Modelo do intercepto aleatório



Modelo do intercepto e inclinação aleatórias

- O intercepto e a inclinação variam aleatoriamente entre os fragmentos (seguem distribuições normais)

Modelo do intercepto e inclinação aleatórias

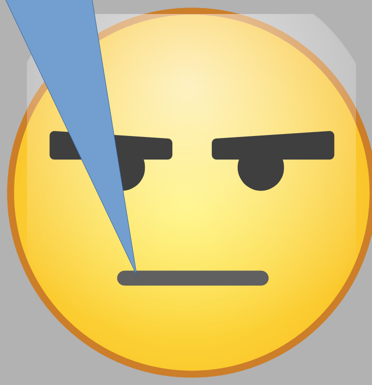


Comparando modelos

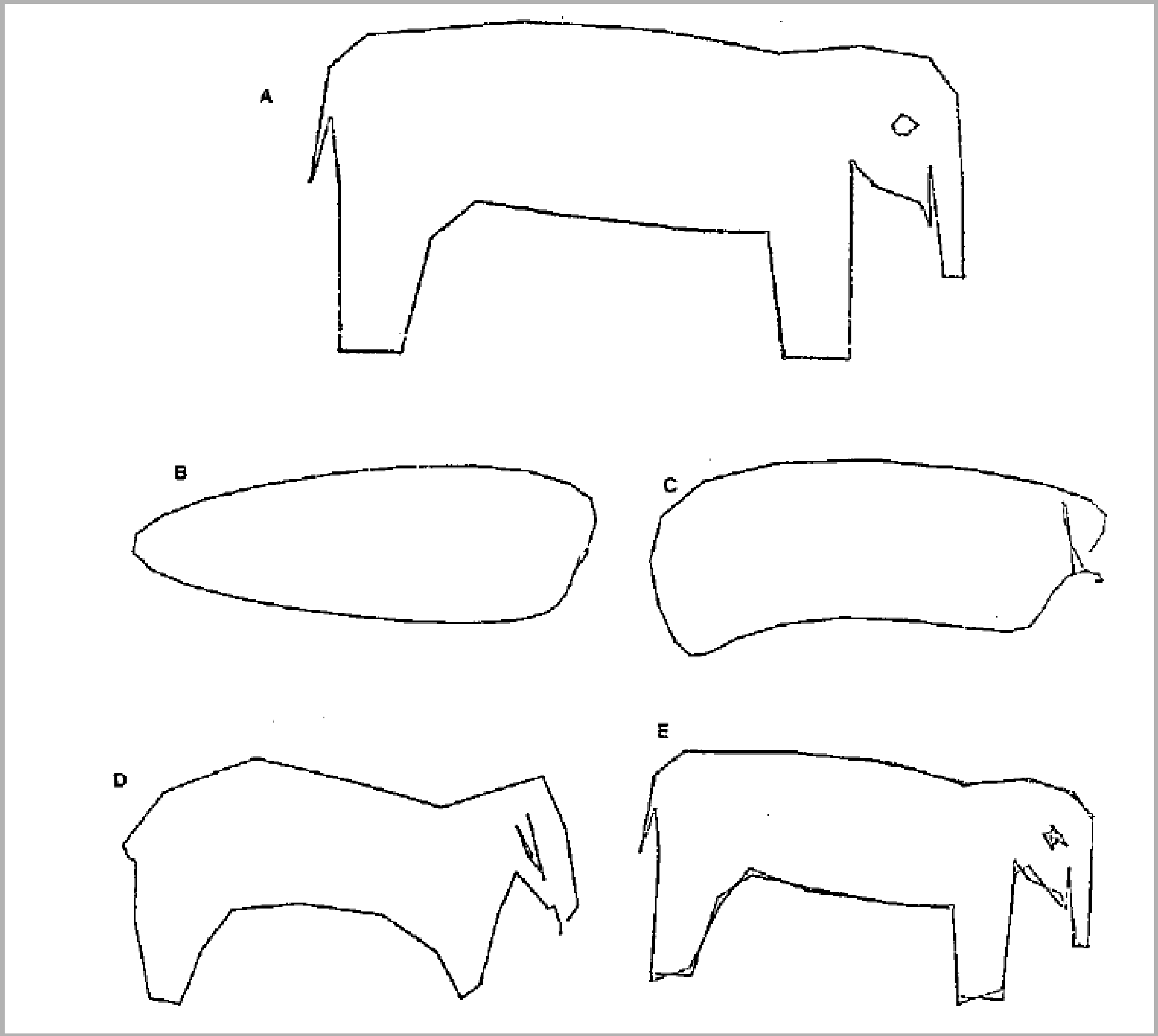
Parcimônia...

Rapaz, tive um ajuste perfeito!

Oxe, você usou parâmetros suficientes pra ajustar um elefante!



Ajustando um elefante.



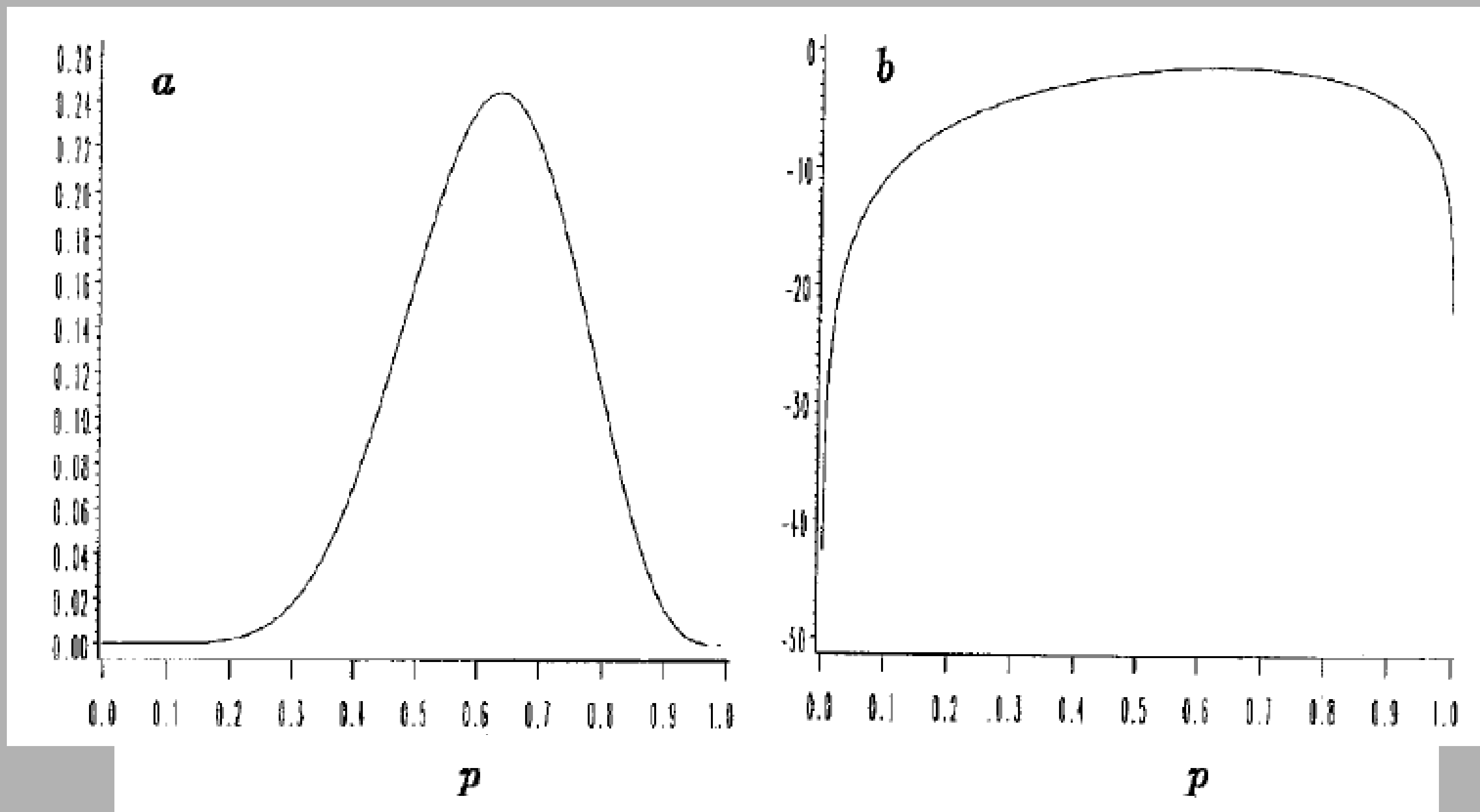
Burnham &
Anderson 2002

Complexidade do modelo

Model structure	Number of parameters $(K)^a$
$E(y) = ax^b$	3
$E(y) = a + b\log(x)$	3
$E(y) = a(x/(b+x))$	3
$E(y) = a(1 - e^{-bx})$	3
$E(y) = a - bc^x$	4
$E(y) = (a + bx)/(1 + cx)$	4
$E(y) = a(1 - e^{-bx})^c$	4
$E(y) = a\left(1 - [1 + (x/c)^d]^{-b}\right)$	5
$E(y) = a[1 - e^{-(b(x-c))^d}]$	5

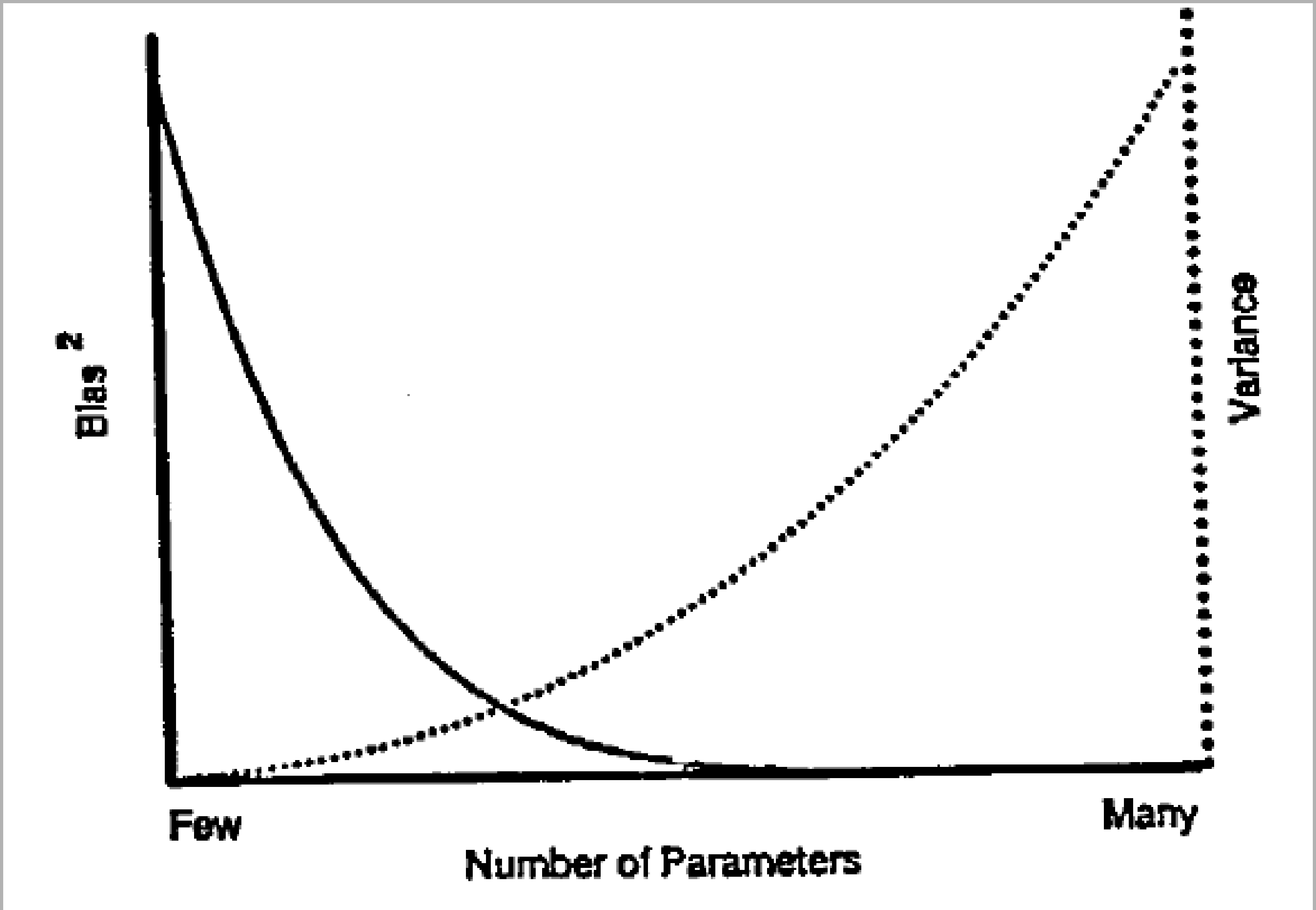
Burnham & Anderson 2002

Verossimilhança (*likelihood*)

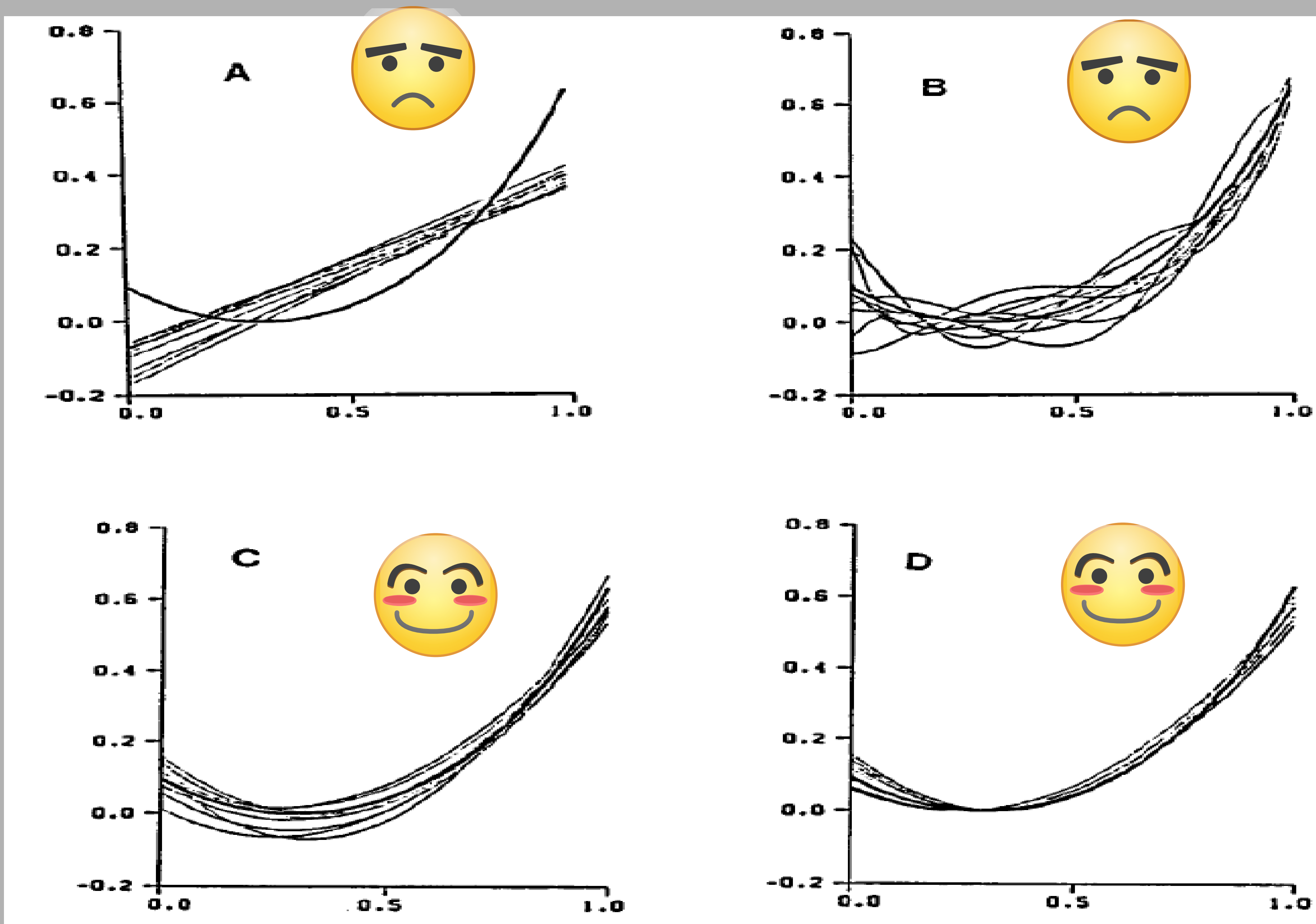


Plots of the binomial likelihood (a) and log-likelihood (b) function, given n 11 penny flips and the observation that y 7 of these were heads. - Burnham and Anderson 2002

Parcimônia



Parcimônia



Critério de Informação de Akaike

$$AIC = -2\log\left(L\left(\hat{\theta}|y\right)\right) + 2K$$

Verossimilhança
(ajuste)

Número de parâmetros
(complexidade)

Critério de Informação de Akaike

$$AIC = -2\log\left(L\left(\hat{\theta}|y\right)\right) + 2K$$

$$\Delta AIC = AIC_j - AIC_{min}$$

Critério de Informação de Akaike

$$AIC = -2\log\left(L\left(\hat{\theta}|y\right)\right) + 2K$$

$$\Delta AIC = AIC_j - AIC_{min}$$

AIC do modelo j

AIC do melhor modelo

Critério de Informação de Akaike

$$AIC = -2\log\left(L\left(\hat{\theta}|y\right)\right) + 2K$$

$$\Delta AIC = AIC_j - AIC_{min}$$

Δ_j	Evidence ratio
2	2.7
4	7.4
8	54.6
10	148.4
15	1,808.0
20	22,026.5

Quão mais provável é o melhor modelo em relação ao pior modelo

Burnham and Anderson 2002

Critério de Informação de Akaike

$$AIC = -2\log\left(L\left(\hat{\theta}|y\right)\right) + 2K$$

$$\Delta AIC = AIC_j - AIC_{min}$$

Δ_i	Level of Empirical Support of Model i
0-2	Substantial
4-7	Considerably less
> 10	Essentially none.

Burnham and
Anderson 2002

Critério de Informação de Akaike

$AIC_i - AIC_j$	Relative likelihood (j:i)	Interpretation
Reference 1		
$>1-2$	$>1.6-2.7$	significant difference between models i and j
Reference 2		
4.2	8	strong enough difference to be of general scientific interest
6.9	32	"quite strong" evidence in favor of model j
Reference 3		
0–4.6	1–10	limited support for model j
4.6–9.2	10–100	moderate support
9.2–13.8	100–1000	strong support
>13.8	>1000	very strong support
Reference 4		
0–2	1–2.7	substantial support of model i
4–7	7.4–33.1	considerably less support
>10	>148	essentially no support
Reference 5		
0 to 4–7	1 to 7.4–33.1	model i is plausible
7–14	33.1–1097	value judgments for hypotheses in this region are equivocal
>14	>1097	model i is implausible

Murtaugh 2014

Critério de Informação de Akaike

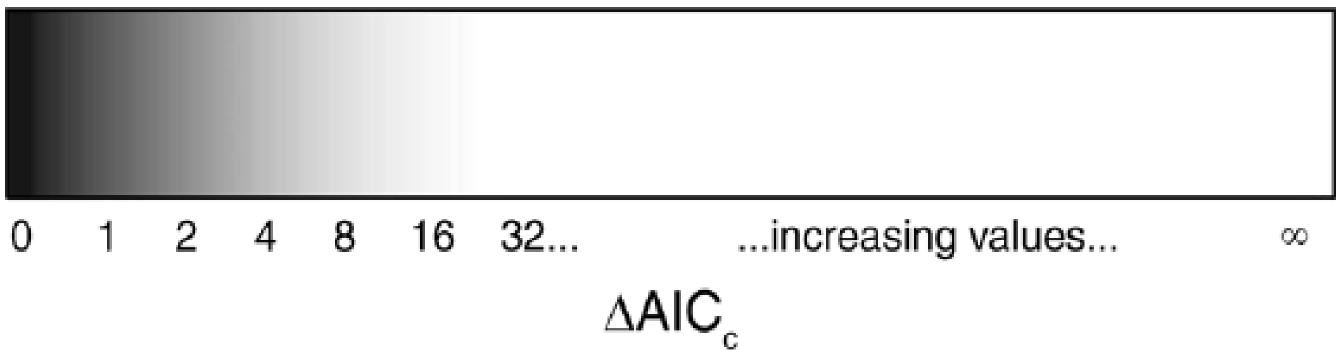


FIG. 3. Interpretation of ΔAIC, from Burnham et al. (2011). “Plausible hypotheses are identified by a narrow region in the continuum where $\Delta <$ perhaps four to seven (black and dark gray). The evidence in the light grey area is inconclusive and value judgments for hypotheses in this region are equivocal. Implausible models are shown in white, $\Delta >$ about 14.” (The authors define Δ, or ΔAIC_c, as the difference between the value of AIC_c for a focal model and the minimum value of AIC_c in a group of models, where AIC_c is a modification of AIC that includes a correction for small sample size.)

Murtaugh 2014

Modelos aninhados

- Um modelo é uma simplificação do outro

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i$$

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{1i} X_{2i} + \epsilon_i$$

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i$$

$$Y_i = \beta_0 e^{(\beta_1 X_i)} + \epsilon_i$$

Modelo nulo

- Ou modelo contendo apenas o intercepto (*intercept-only model*)
- Está aninhado nos outros modelos.

$$Y_i = \beta_0 + \epsilon_i$$