

# Ecologia de Populações

## Verossimilhança Máxima

# O que é um modelo?

“Todos os modelos estão errados, mas alguns são úteis” (Box 1976)

Modelos matemáticos:  $y = b \cdot x$

Exemplo: Faturamento em loja de sorvete

Cada sorvete = 2 reais

Se vendeu 3 sorvetes, faturamento = 6 reais

Se vendeu 7 sorvetes, faturamento = 14 reais

Se vendeu 13 sorvetes, faturamento = 26 reais

**Generalizando:  $Y = 2 \cdot X$**

# O que é um modelo estatístico?

Modelos estatísticos:  $y = b \cdot x + \varepsilon$

Cada sorvete = *em média* 2 reais (depende do freguês!)

Se vendeu 3 sorvetes, faturamento = *em média* 6 reais

Se vendeu 7 sorvetes, faturamento = *em média* 14 reais

Se vendeu 13 sorvetes, faturamento = *em média* 26 reais

**Generalizando:  $Y = 2 \cdot X + \varepsilon$**

Relação  
espécies-  
área  
(Preston  
1962)

Iha	Area	Nespecies	LogArea	LogEspecies
Albemarle	5824.9	325	3.765	2.512
Charles	165.8	319	2.219	2.504
Chatham	505.1	306	2.703	2.486
James	525.8	224	2.721	2.350
Indefatigable	1007.5	193	3.003	2.286
Abingdon	51.8	119	1.714	2.076
Duncan	18.4	103	1.265	2.013
Narborough	634.6	80	2.802	1.903
Hood	46.6	79	1.669	1.898
Seymour	2.6	52	0.413	1.716
Barrington	19.4	48	1.288	1.681
Gardner	0.5	48	-0.286	1.681
Bindloe	116.6	47	2.067	1.672

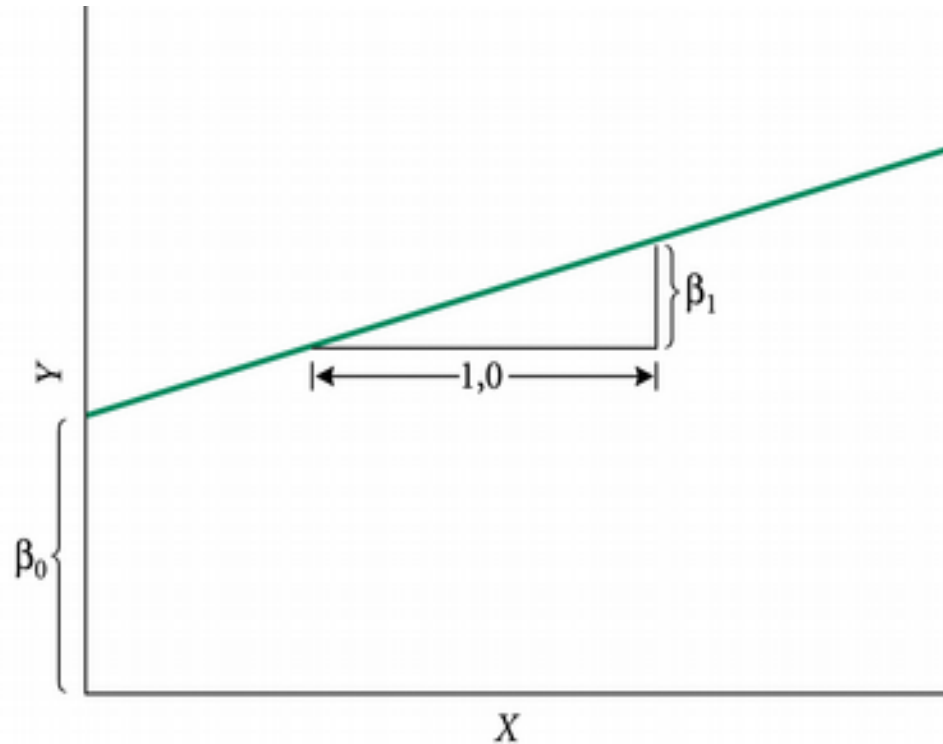
Definindo uma linha reta e seus dois parâmetros:  
inclinação e intercepto (Fig. 9.1)

Ajustando dados a um modelo linear:

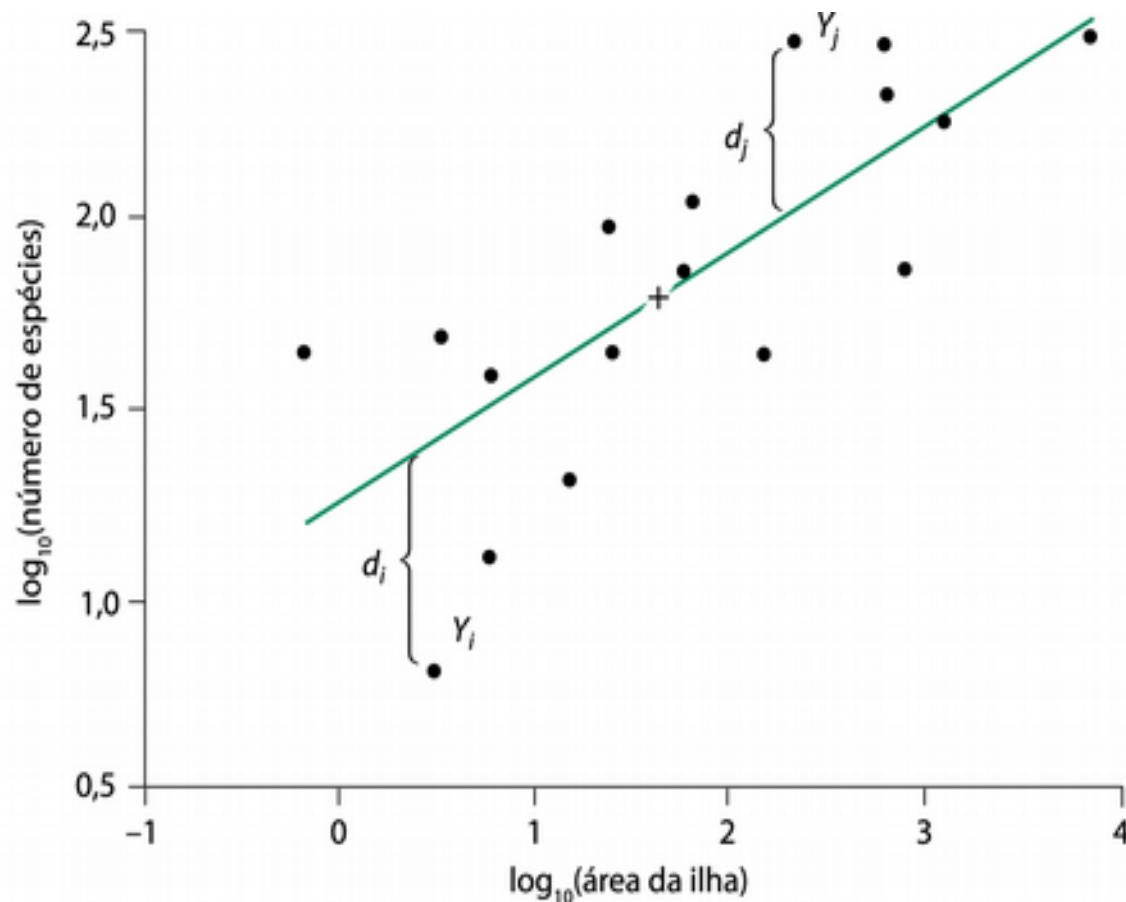
$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$\hat{Y}_i = \theta_0 + \theta_1 X_i$$

$$\varepsilon_i = (Y_i - \hat{Y}_i)^2$$



**Figura 9.1** Relação linear entre as variáveis  $X$  e  $Y$ . A linha é descrita pela equação  $Y = \beta_0 + \beta_1 X$ , onde  $\beta_0$  é o intercepto e  $\beta_1$  é a inclinação da linha. O intercepto  $\beta_0$  é o valor predito da equação quando  $X = 0$ . A inclinação da linha  $\beta_1$  é o aumento na variável  $Y$  associado com o de uma unidade da variável  $X$  ( $\Delta Y / \Delta X$ ). Se o valor de  $X$  é conhecido, o valor predito de  $Y$  pode ser calculado multiplicando  $X$  pela inclinação e somando o intercepto ( $\beta_0$ ).



**Figura 9.3** A soma dos quadrados dos resíduos é obtida somando os desvios quadrados ( $d_i$ ) de cada observação da linha de regressão ajustada. A estimativa do parâmetro dos mínimos quadrados garante que essa linha de regressão minimize a soma dos quadrados dos resíduos. O + marca o ponto central dos dados ( $\bar{X}$ ,  $\bar{Y}$ ). Essa linha de regressão descreve, ainda, a relação entre o logaritmo da área das ilhas e o do número de espécies de plantas das Ilhas de Galápagos, dados da Tabela 8.2. A equação da regressão é  $\log_{10}(\text{espécies}) = 1,320 + \log_{10}(\text{área}) \times 0,331$ ;  $r^2 = 0,584$ .

# Verossimilhança (*Likelihood*)

- Após um modelo e seus parâmetros serem definidos, e dados coletados, em geral o passo seguinte é estimar seu ajuste (*goodness of fit*)
- Isto é, os valores dos parâmetros do modelo que melhor se ajustam aos dados  
= Estimativa de parâmetros
- Método geral para estimativa de parâmetros: verossimilhança máxima (*maximum likelihood*)
- Quadrados mínimos (*least squares*): caso particular quando os resíduos seguem uma distribuição normal



# Verossimilhança Máxima

- Se a probabilidade de um evento  $x$  depende de parâmetros  $p$  de um modelo, escrevemos

$$P(x | p)$$

- mas quando falamos de verossimilhança

$$L(p | x)$$

- quer dizer, a verossimilhança dos parâmetros considerando o evento  $x$
- Quando estimamos um parâmetro por verossimilhança máxima queremos dizer que procuramos o valor do parâmetro que tenha a maior chance de produzir o evento.

	Viés da moeda favorecendo Cara				
Cara	0,1	0,3	0,5	0,7	0,9
0	0,59	0,17	0,03	0,00	0,00
1	0,33	0,36	0,16	0,03	0,00
2	0,07	0,31	0,31	0,13	0,01
3	0,01	0,13	0,31	0,31	0,07
4	0,00	0,03	0,16	0,36	0,33
5	0,00	0,00	0,03	0,17	0,59



<http://www.rasch.org/rmt/rmt1237.htm>

Suponha que em 5 jogadas da moeda foram obtidas 5 caras. Qual o deve ser o viés da moeda?

- Para estimarmos o viés da moeda, olhamos a probabilidade de obter 5 caras em todas as hipóteses. A hipótese onde esta frequência é a mais provável fornece uma estimativa do viés.
- No caso, 0,9 é a estimativa mais próxima do viés.
- O somatório dos valores de uma coluna é sempre 1,0. São as probabilidades dos resultados em cada hipótese, cada coluna uma hipótese.
- Como o somatório dos valores de uma linha é sempre diferente de 1, foi necessário diferenciar estes valores de probabilidades, adotando-se o termo verossimilhança (*likelihood*).

# Estimativa por verossimilhança

$n = 100$  (total de lançamentos da moeda)

$h = 56$  (total de caras)

OBTIVEMOS 56 CARAS E 44 CORÔAS. É UMA MOEDA “JUSTA”?

Precisamos de um modelo estatístico para descrever o fenômeno!

Variável aleatória binomial

$$P(X) = \frac{n!}{X!(n-X)!}$$

# Estimativa por verossimilhança

Mas no caso temos  $X = 56$  e  $n = 100$ , falta o valor de  $p$ , que estamos estimando.

$$P(56) = \frac{100!}{56!(100-56)!}$$

$$P(56) = \frac{100!}{56!(44)!}$$

$$P(56) = (4,94 \times 10^{-16})$$

# Estimativa por verossimilhança

Quando temos uma função do parâmetro  
(desconhecido),

mas as observações são fixas,

temos uma função de verossimilhança,

(não mais de densidade de probabilidade)

# Estimativa por verossimilhança

$n = 100$  (total de lançamentos da moeda)

$h = 56$  (total de caras)

OBTIVEMOS 56 CARAS E 44 CORÔAS. É UMA MOEDA “JUSTA”?

Se  $p = 0,5$

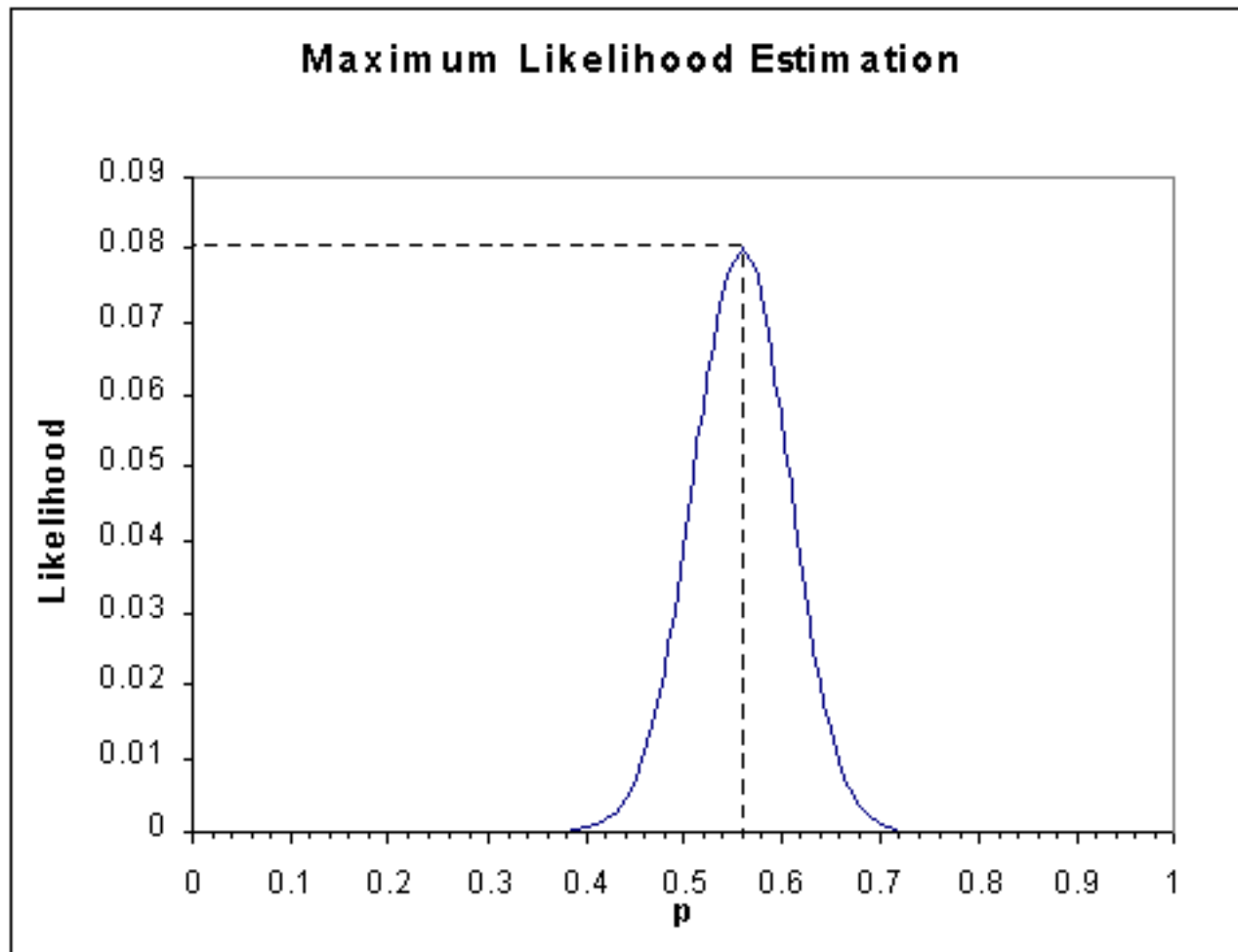
$$L(p = 0.5 | data) = \frac{100!}{56!44!} 0.5^{56} 0.5^{44} = 0.0389$$

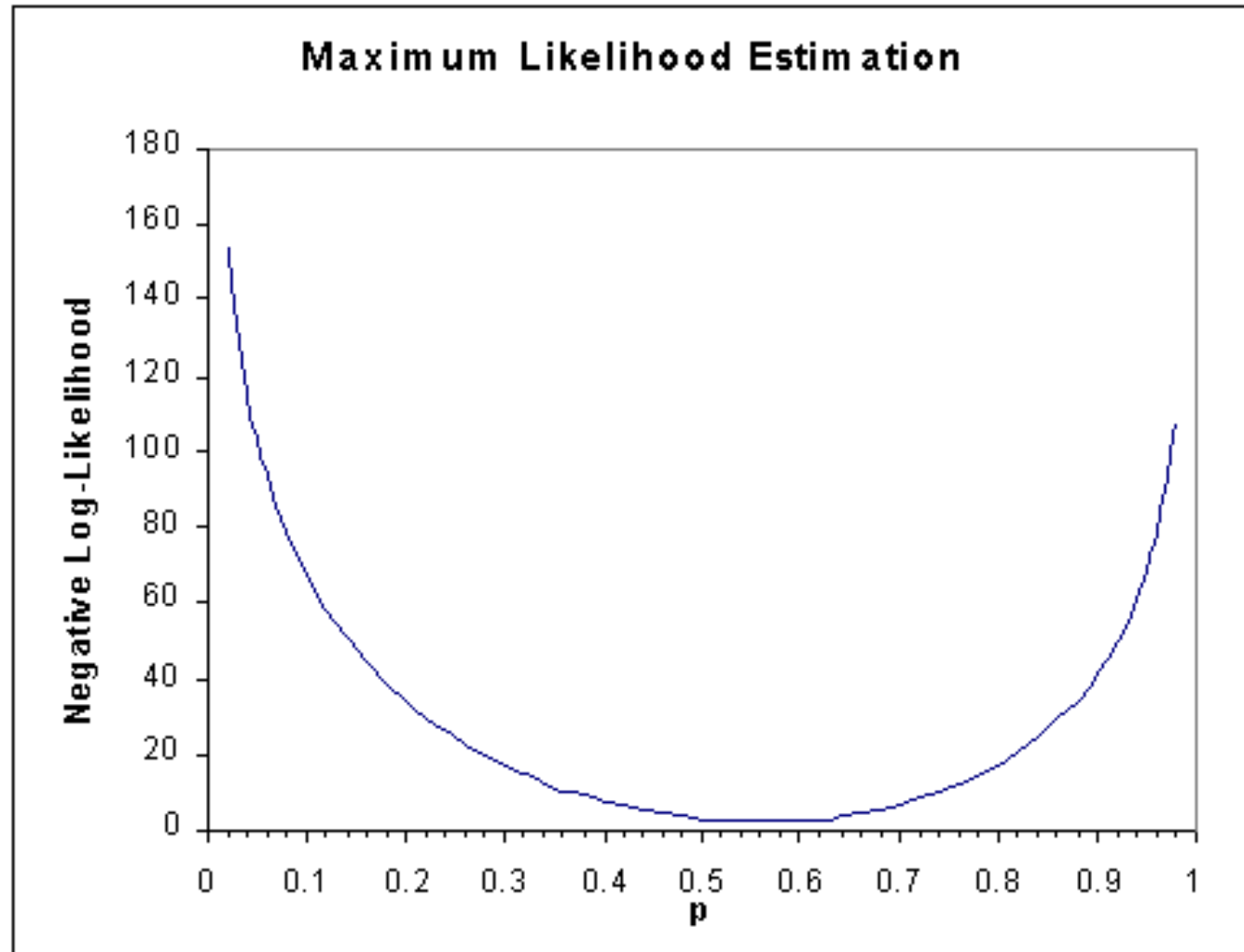
$$L(p = 0.52 | data) = \frac{100!}{56!44!} 0.52^{56} 0.48^{44} = 0.0581$$

p	L
0.48	0.0222
0.52	0.0581
0.56	0.0801
0.60	0.0576

p	L
0.50	0.0389
0.54	0.0739
0.58	0.0738
0.62	0.0378







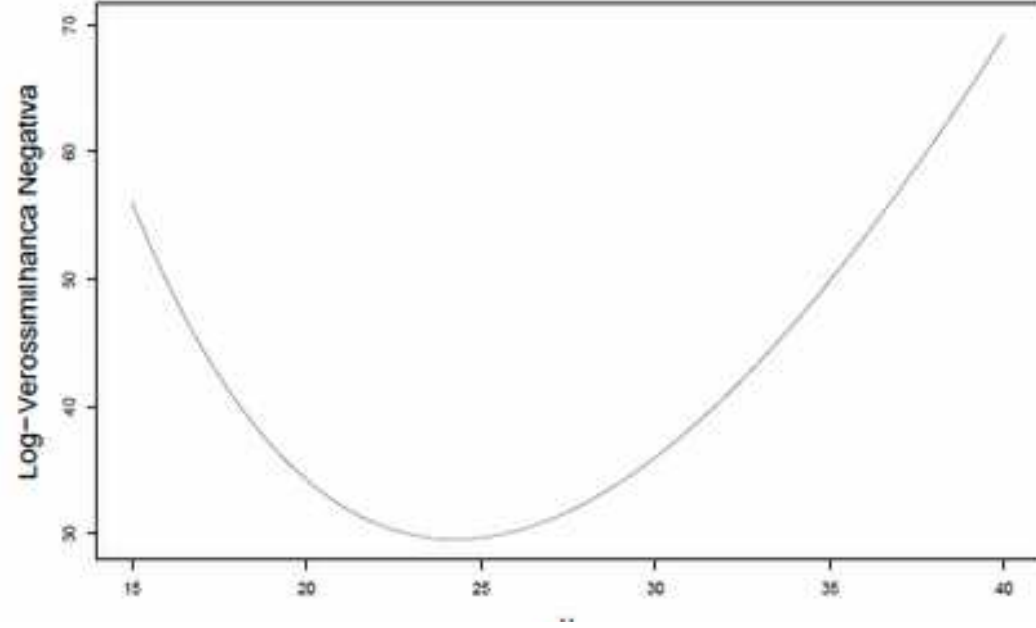
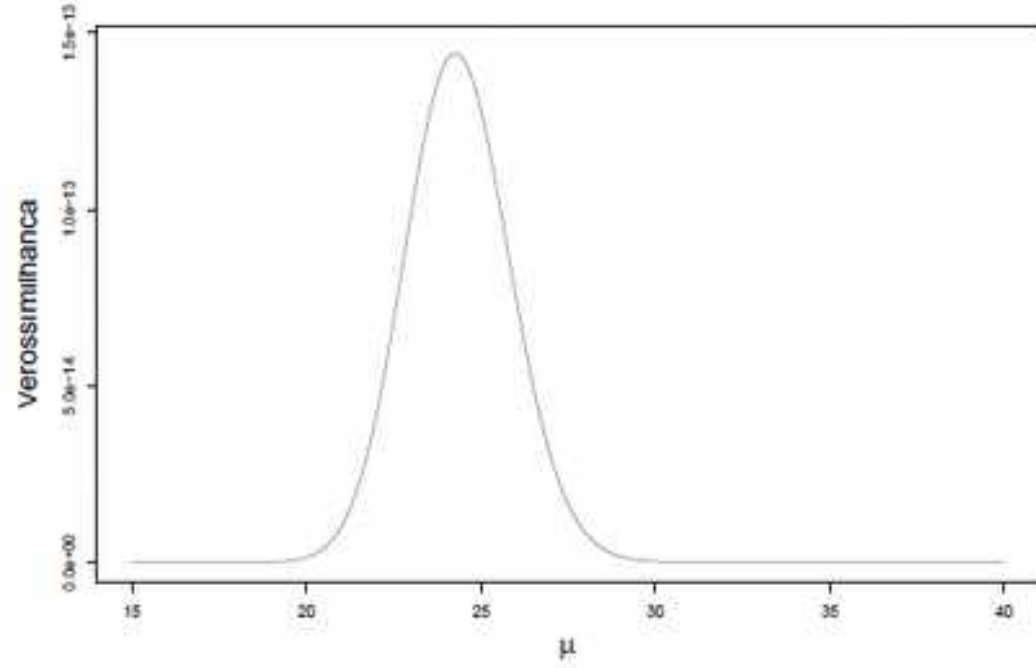
$$\mathcal{L}\{\mu|X_n\} = \prod_{i=1}^n \frac{e^{-\mu} \mu^{x_i}}{x_i!} = e^{-\mu} \prod_{i=1}^n \frac{\mu^{x_i}}{x_i!}$$

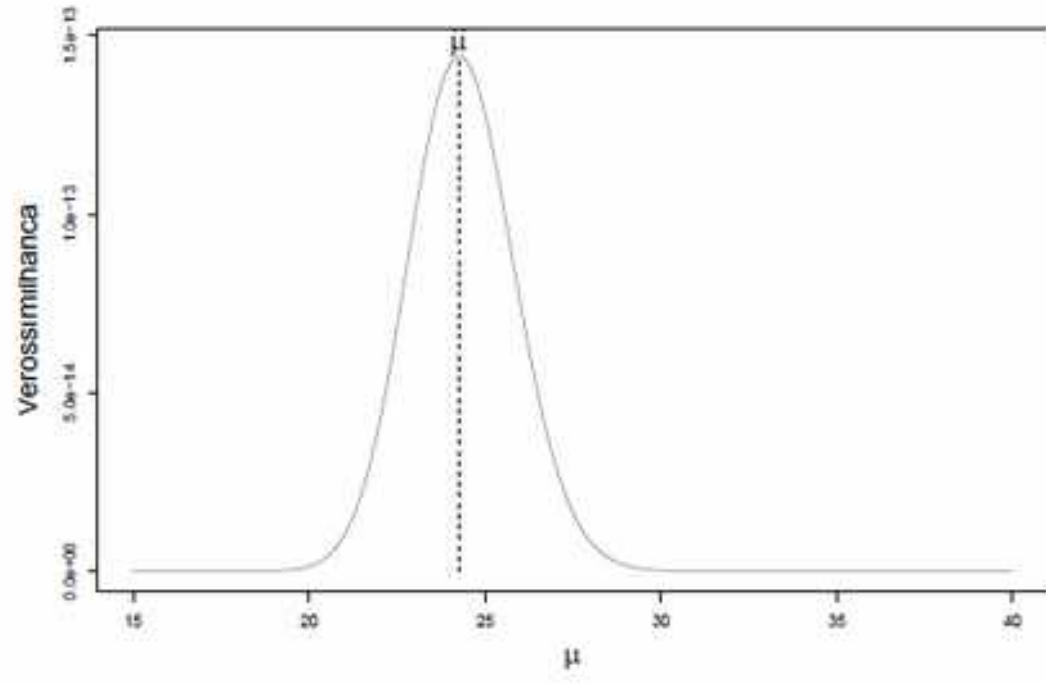
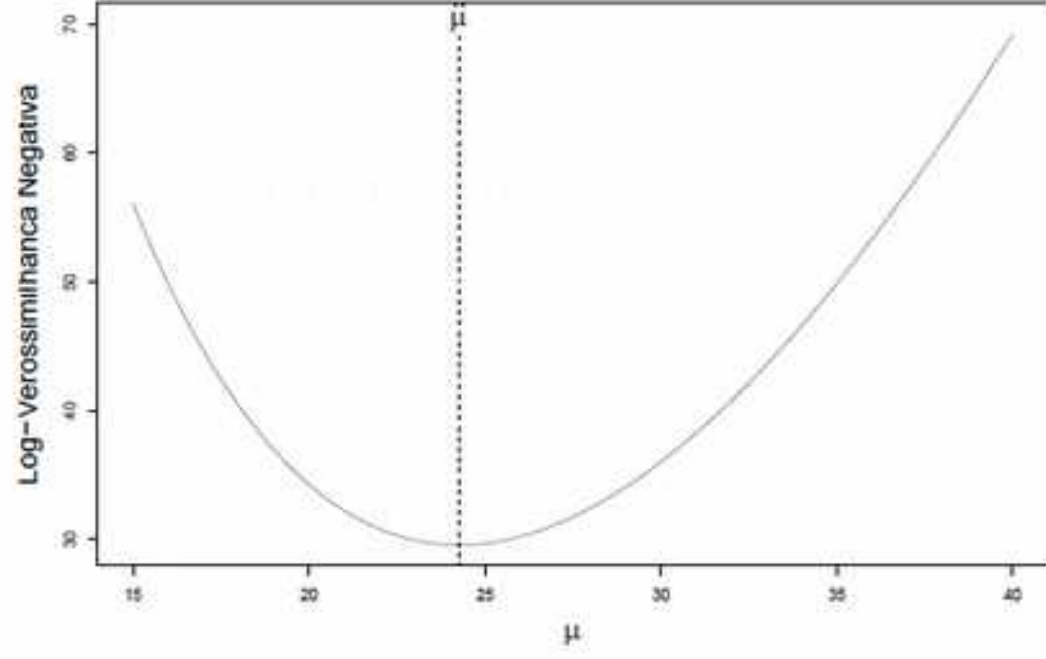
PARCELA	NO. DE PLÂNTAS
(i)	(X <sub>i</sub> )
1	24
2	27

$$L\{\mu|X_n\} = n\mu - \log(\mu) \sum_{i=1}^n x_i + \sum_{i=1}^n \log(x_i!)$$

$$L\{\mu|X_n\} = n\mu - \log(\mu) k_1 + k_2$$

PARCELA	NO. DE PLÂN1
(i)	(X=xi)
1	24
2	27
...	...





$$Y = \beta_0 + \beta_1 x + u$$

$$u \sim N(0, \sigma^2)$$

$$P(x_i) =$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-1/2 \left[ (x_i - \mu) / \sigma \right]^2}$$

$$L(\text{amostra}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-1/2 \left[ (u_1 / \sigma) \right]^2}$$

$$= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-1/2 \left( \sum_{i=1}^n u \right)^2}$$

$$\log \left[ \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} u_i^2} \right]$$

$$\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n u_i^2} = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n \frac{u_i^2}{\sigma^2}}$$

$$\text{Log } L = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n u_i^2$$



$$\text{Log } L = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2} -$$

*EM ESTIMATIVAS DA MÉDIA:*

$$\varepsilon_i = Y_i - \text{média}(Y_i)$$

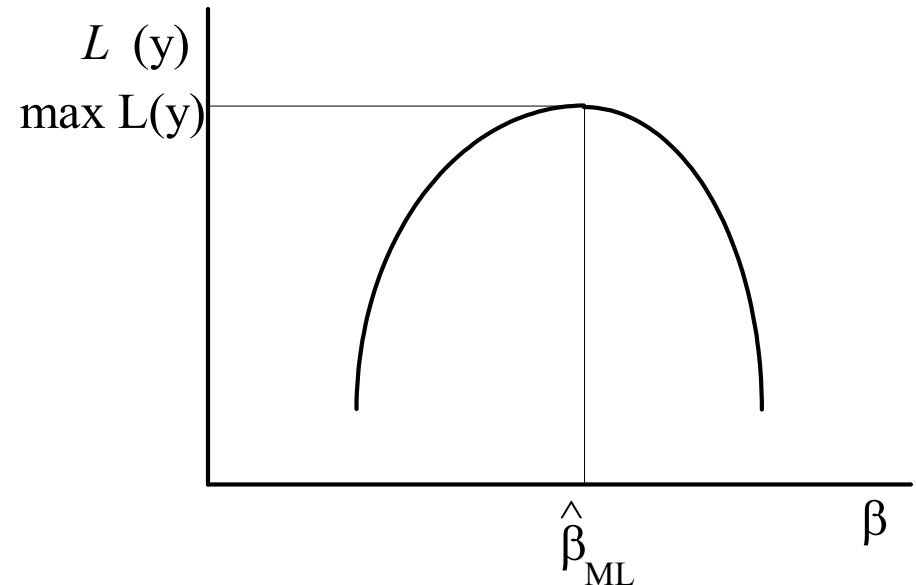
*EM MODELOS LINEARES:*

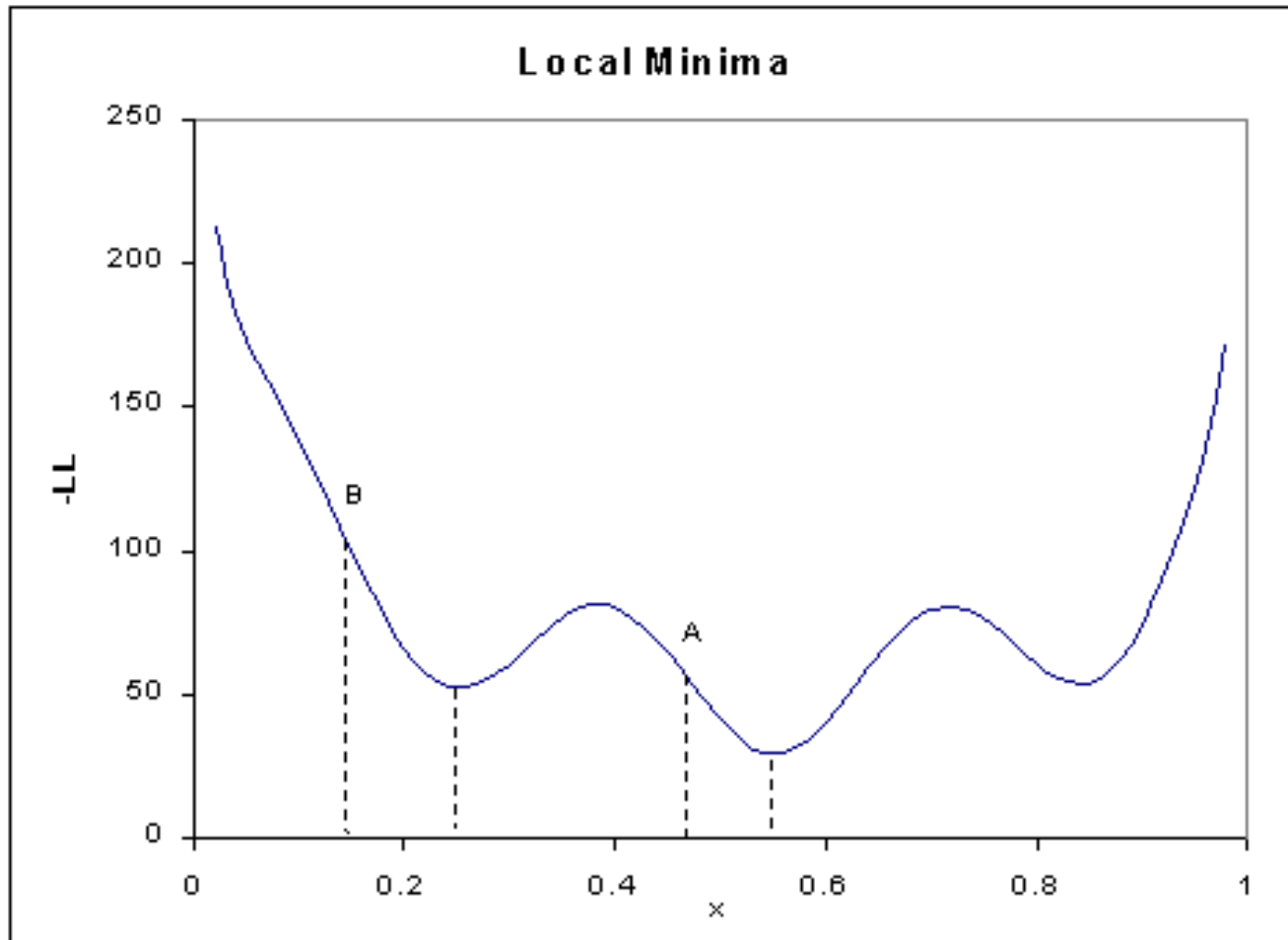
$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$\hat{Y}_i = \beta_0 + \beta_1 X_i$$

$$\varepsilon_i = Y_i - \hat{Y}_i$$

$$\varepsilon_i = Y_i - (\beta_0 + \beta_1 X_i)$$





# Lei ou princípio da verossimilhança

- Duas hipóteses, A e B
- Resultado  $x$  podem ocorrer segundo estas duas hipóteses e uma variável aleatória  $X$
- $P(A) = x > P(B) = x$
- Razão de verossimilhança:

mede a **força de evidência** a favor de uma hipótese

→ Mas não considera variância das estimativas dos modelos a novos conjuntos de dados!