

Metodologia Ecológica

Verossimilhança Máxima

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

$$P(X) = \frac{n!}{X!(n-X)!} p^X (1-p)^{n-X}$$

Estimativa por verossimilhança

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

$$P(56) = \frac{100!}{56!(100-56)!} p^{56} (1-p)^{100-56}$$

$$P(56) = \frac{100!}{56!(44)!} p^{56} (1-p)^{44}$$

$$P(56) = (4,94 \times 10^{28}) p^{56} (1-p)^{44}$$

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”?

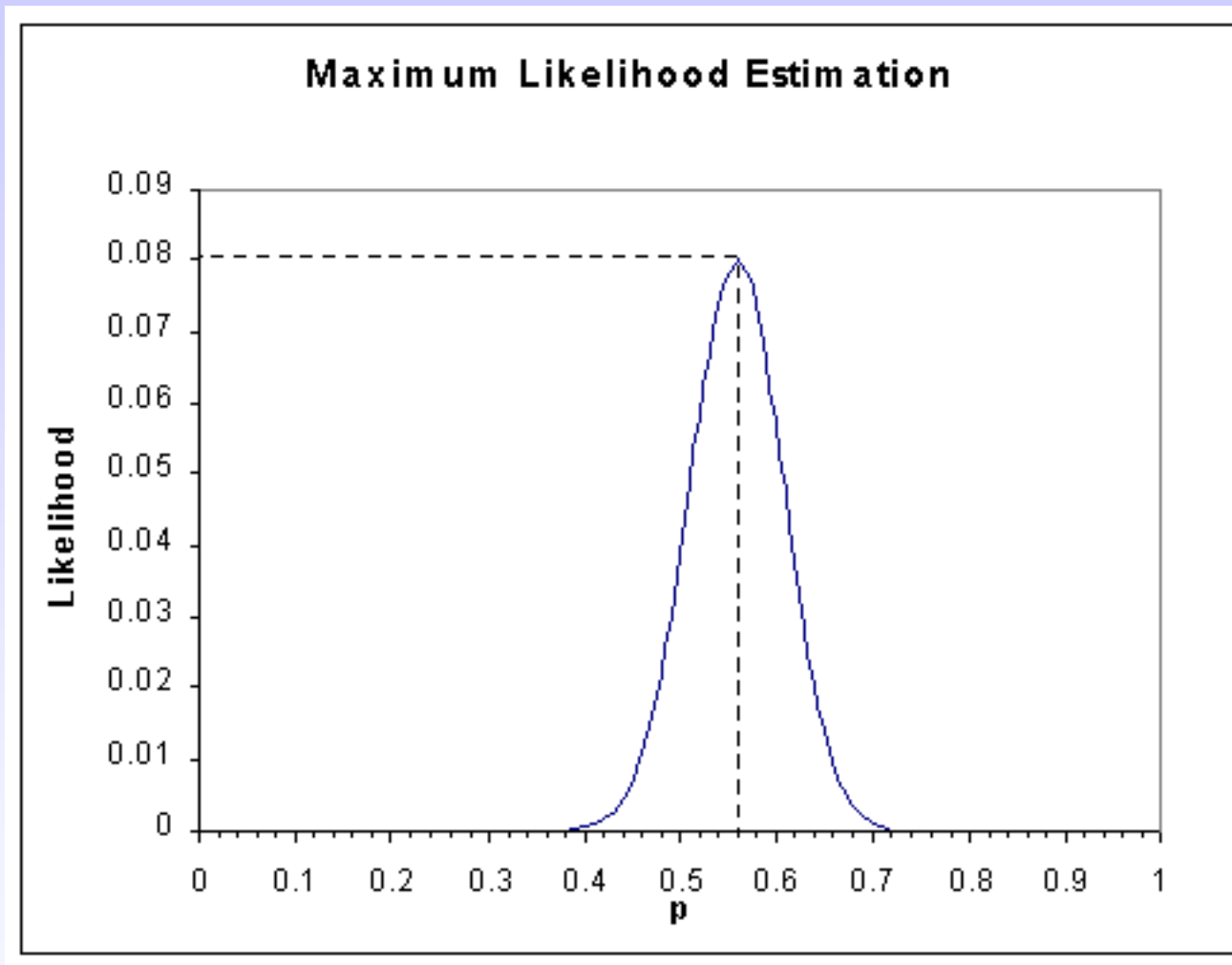
Se $p = 0,5$

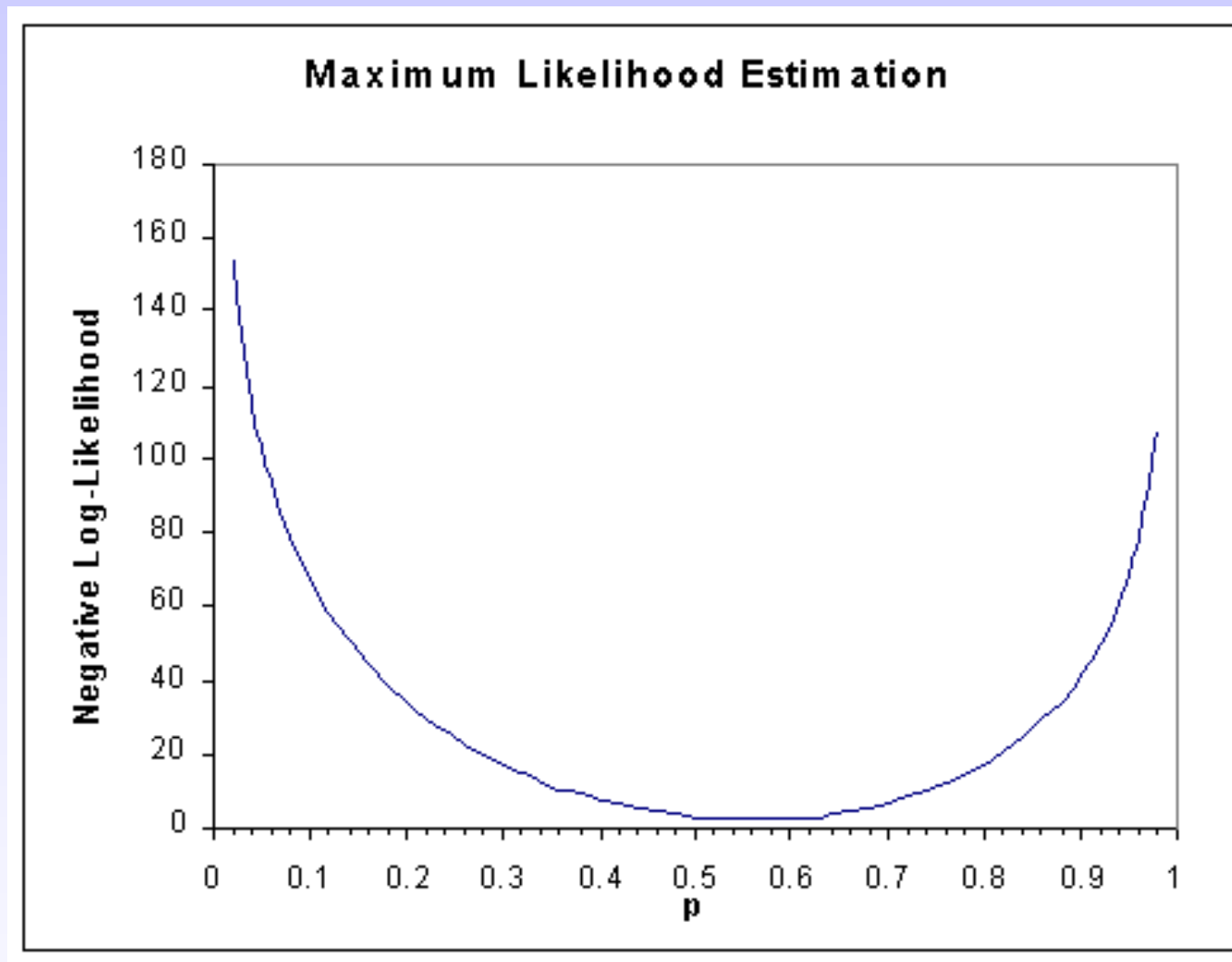
$$L(p = 0.5 | data) = \frac{100!}{56!44!} 0.5^{56} 0.5^{44} = 0.0389$$

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

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 PLÂNTULAS	VEROSSIMILHANÇA	
(i)	(X=xi)	(μ=16)	(μ=35)
1	24	0.0144	0.0116
2	27	0.0034	0.0283
3	23	0.0216	0.0080
4	28	0.0019	0.0354
5	26	0.0057	0.0219
6	24	0.0144	0.0116
7	17	0.0934	0.0003
8	23	0.0216	0.0080
9	24	0.0144	0.0116
10	27	0.0034	0.0283
		1.6290E-20	1.8608E-20

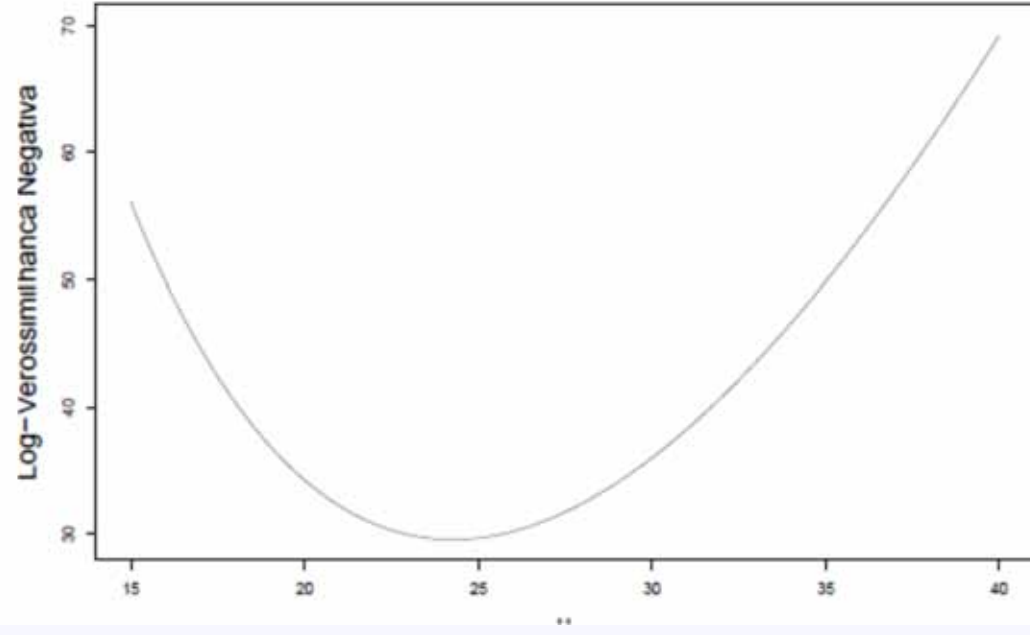
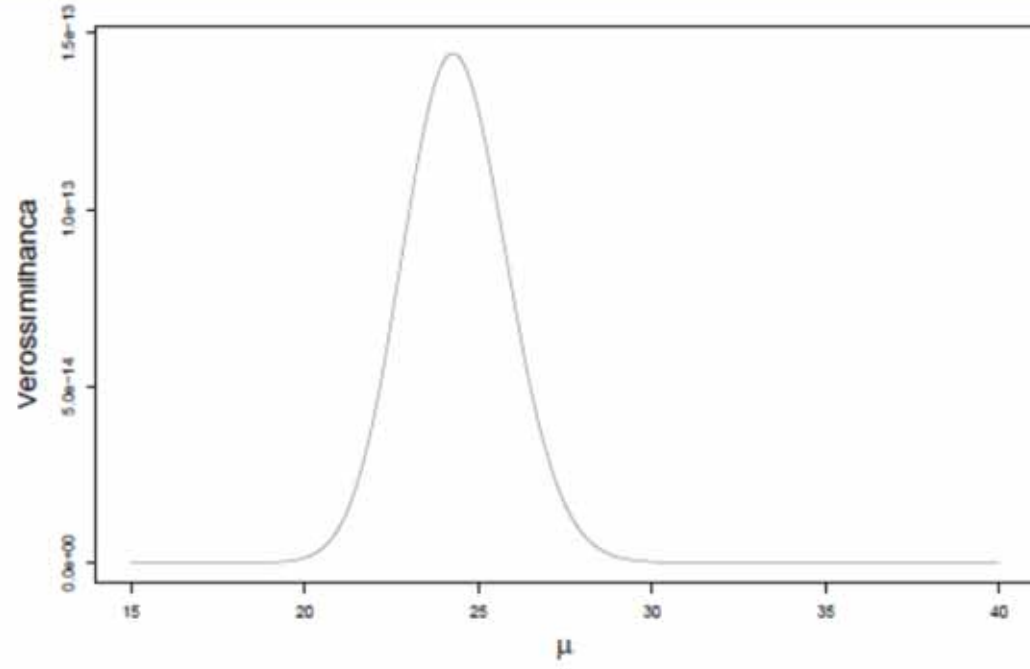
João Luís F. Batista

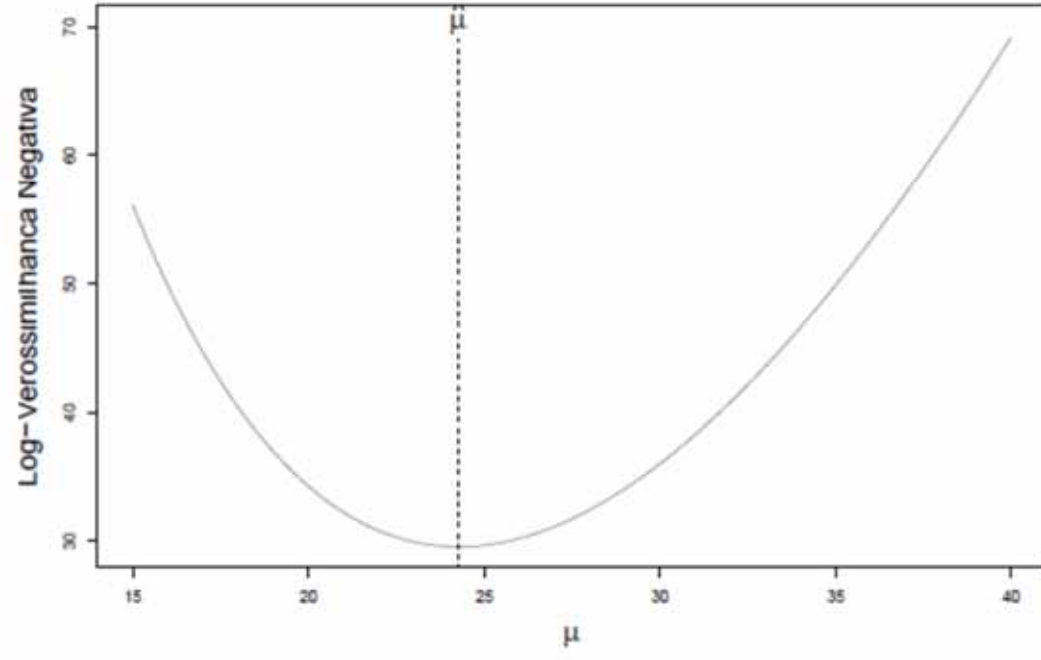
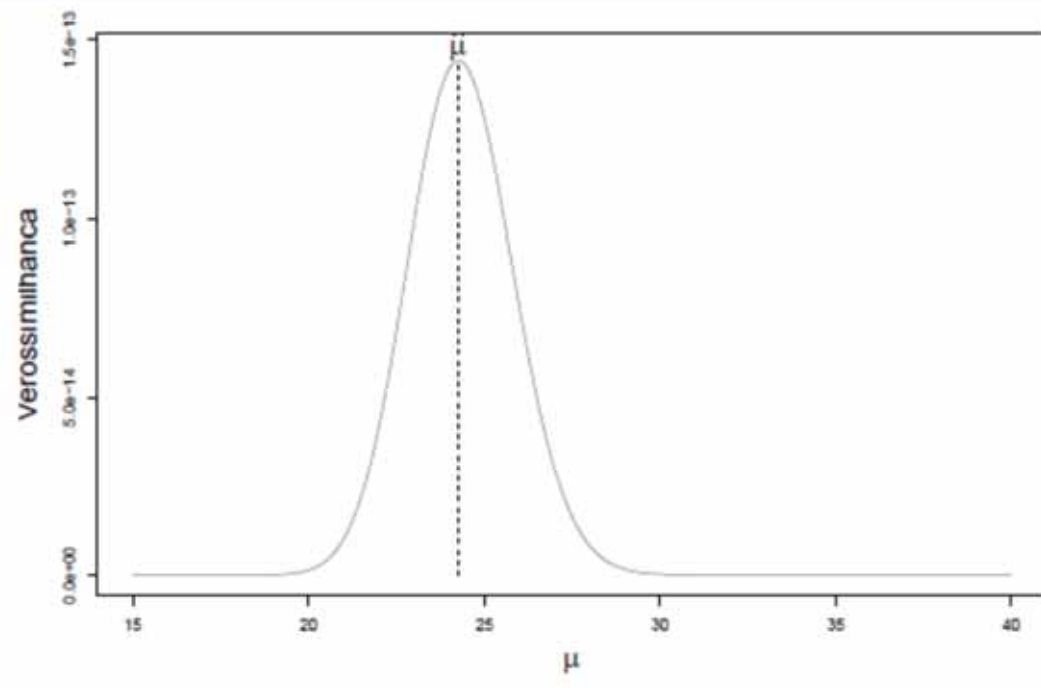
Prof. Marcus Vinícius Vieira – Instituto de Biologia UFRJ

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

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

PARCELA	NO. DE PLÂNTULAS	VEROSSIMILHANÇA	
(i)	(X=xi)	(μ=16)	(μ=35)
1	24	4.243	4.456
2	27	5.698	3.563
3	23	3.837	4.834
4	28	6.257	3.340
5	26	5.174	3.823
6	24	4.243	4.456
7	17	2.371	8.064
8	23	3.837	4.834
9	24	4.243	4.456
10	27	5.698	3.563
		45.600	45.390





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

$$u \sim N(0, \sigma^2) \quad P(u_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{u - \bar{u}}{\sigma}\right)^2}$$

$$L(\text{amostra}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(u_1/\sigma)^2} \times \dots \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(u_n/\sigma)^2}$$

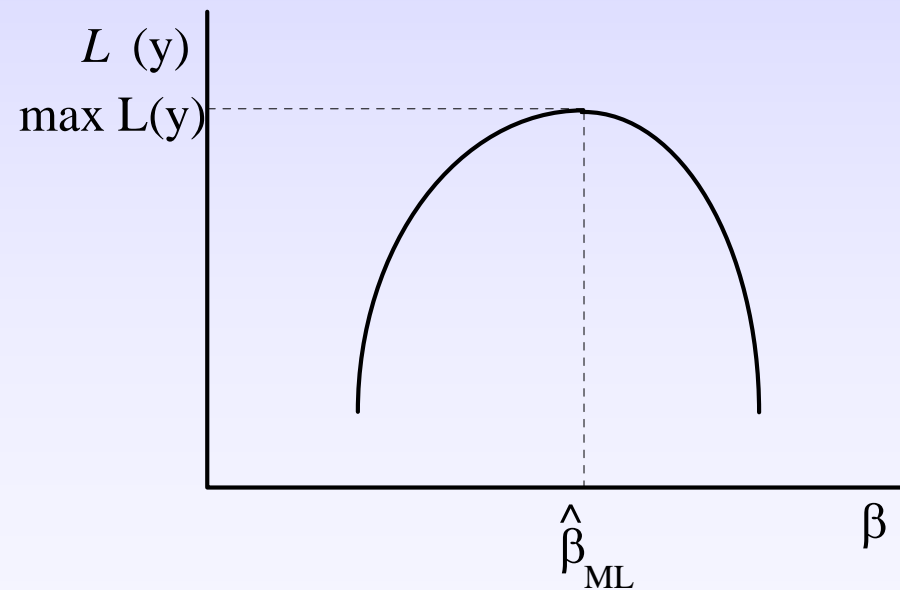
$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\sum_{i=1}^n u_i/\sigma\right)^2} \quad \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\sum (u_i - \bar{u})/\sigma\right)^2}$$

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

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

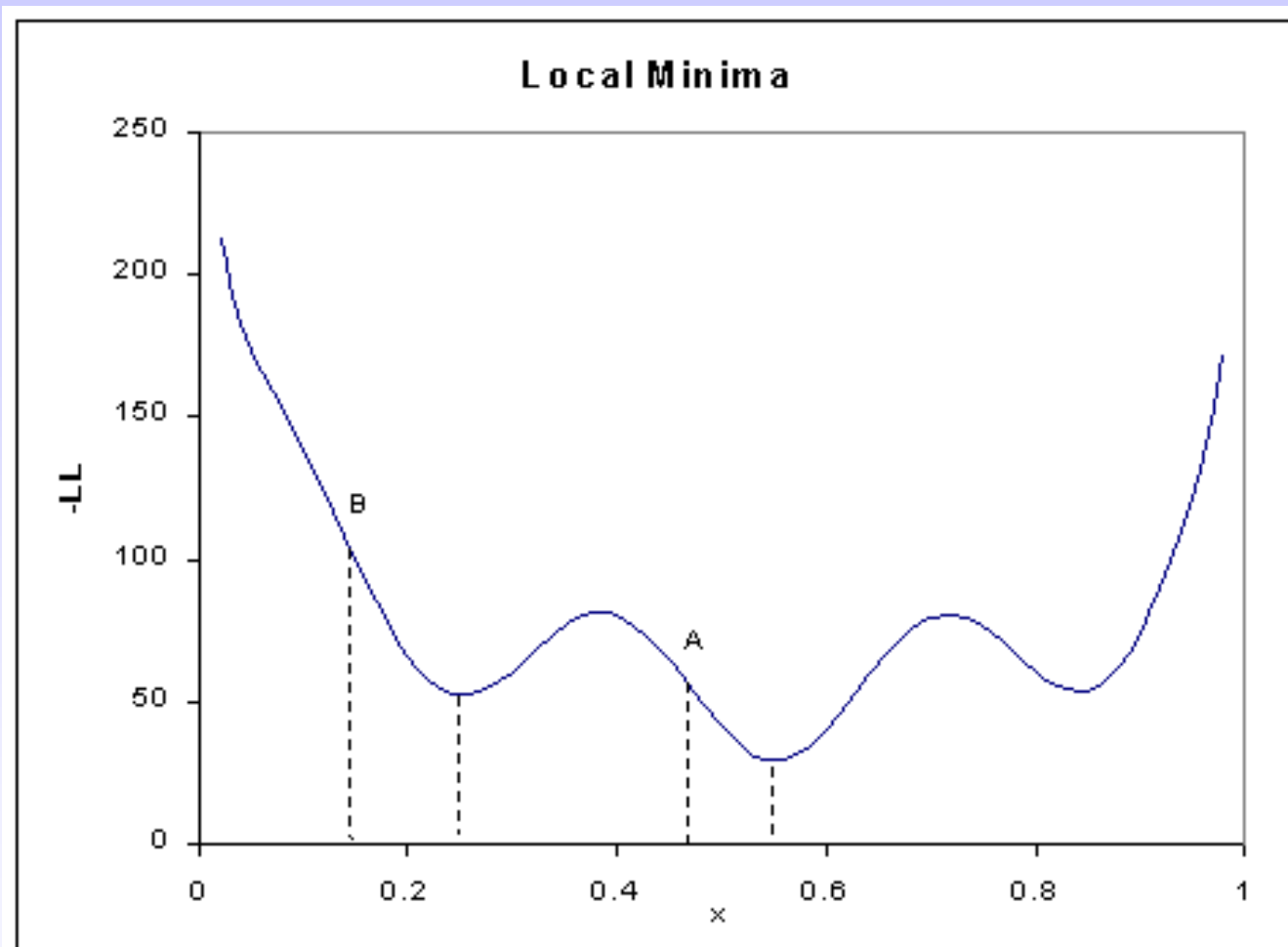
- $\text{Log } L = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \sum \left(\frac{u}{\sigma} \right)^2$

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



- A estimativa de σ^2 que maximiza sua verossimilhança é a Soma dos Quadrados dos Resíduos por $n = \text{RSS}/n$
- Que difere da estimativa de σ^2 pelos quadrados mínimos: $\text{RSS}/(n-k)$ (devido à correção do viés para amostras pequenas)
- Este viés não é importante para estimativas de VM, mas RSS é dividido por valores diferentes!

- $$\text{Log } L = -\frac{n}{2} \ln\left(\frac{2\pi \text{RSS}}{n}\right) - \frac{1}{2} \frac{\text{RSS}}{\text{RSS}/n} = -\frac{n}{2} \left(\ln\left(\frac{2\pi \text{RSS}}{n}\right) + 1 \right)$$



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