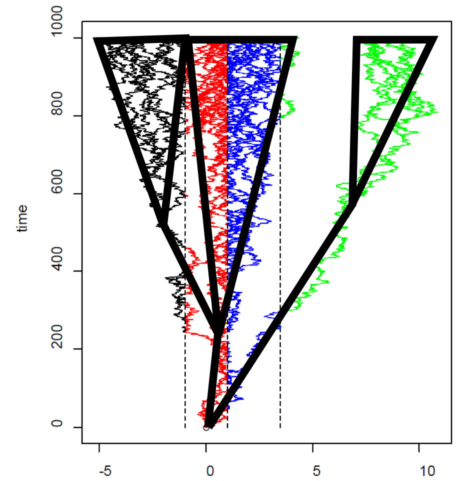
Função de verossimilhança

Número de parâmetros

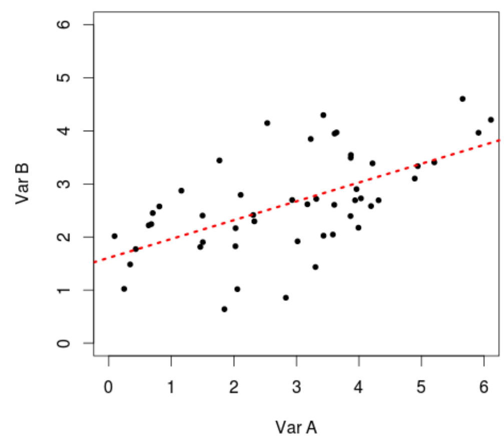
1

$$dX(t) = \sigma dB(t)$$



2

$$ax + b = y$$



Função de verossimilhança

Muda o valor do parâmetro ...

... verifica o ajuste nos dados.

Função de verossimilhança

Função de verossimilhança

Valores dos parâmetros

Verossimilhança do modelo

"Qual a chance de *observar estes dados* assumindo este modelo e seus valores de parâmetros?"

Função de verossimilhança

Valores dos parâmetros

$$P(y|\theta)$$

Verossimilhança do modelo

"Qual a chance de *observar estes dados* assumindo este modelo e seus valores de parâmetros?"

---

y = dados observados

θ = vetor de parâmetros

$$\theta = \{\theta_1, \theta_2, \theta_3, \ldots, \theta_i\}$$

Função de verossimilhança

Likelihood of a model = probabilidade dos dados dado o modelo.

$$P(y|\theta) \longrightarrow$$

"Qual a chance de *observar estes dados* assumindo este modelo e seus valores de parâmetros?"

---

y = dados observados

θ = vetor de parâmetros

$$\theta = \{\theta_1, \theta_2, \theta_3, \ldots, \theta_i\}$$

Função de verossimilhança

Likelihood of a model = probabilidade dos dados dado o modelo.

$$P(y|\theta) \longrightarrow$$

Likelihood of a model
ou
Verossimilhança de um modelo

---

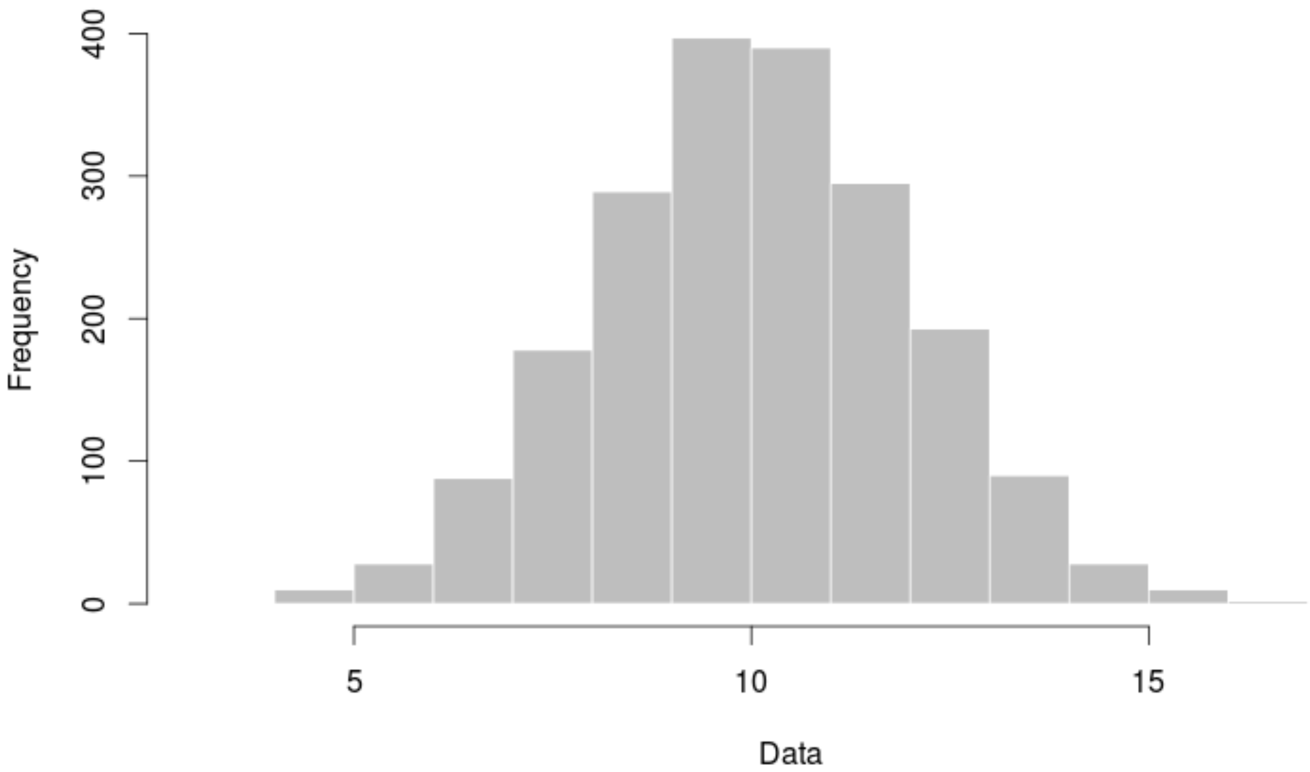
y = dados observados

θ = vetor de parâmetros

$$\theta = \{\theta_1, \theta_2, \theta_3, \ldots, \theta_i\}$$

Função de verossimilhança

Exemplo com uma distribuição normal



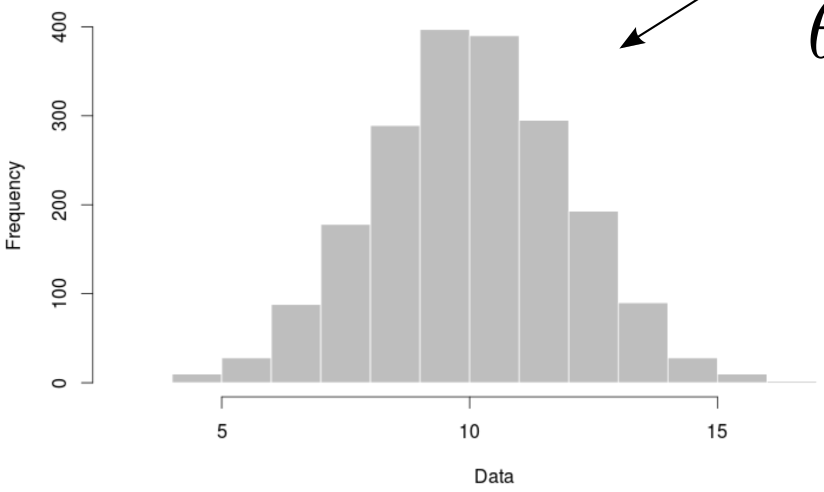
Função de verossimilhança

Ajustar um modelo de distribuição normal para os nossos dados.

Calcular $P(y|\theta)$, sendo que:

$$y = dados$$

$$\theta = \{\mu, \sigma^2\}$$



Função de verossimilhança

Como é a função de verossimilhança da distribuição normal?

$$P(y_1, y_2, \ldots, y_n | \mu, \sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} e^{\frac{-1}{2\sigma^2} \sum_{j=1}^{n} (y_j - \mu)^2}$$

onde $y = \{y_1, y_2, \ldots, y_n\}$

Usamos esta função para calcular a verossimilhança do modelo de distribuição normal com diferentes valores para os parâmetros: média e desvio padrão.

Função de verossimilhança

$$P(y_1, y_2, \ldots, y_n|\mu, \sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} e^{\frac{-1}{2\sigma^2}\sum_{j=1}^{n}(y_j - \mu)^2}$$

$$(2\pi\sigma^2)^{\frac{-n}{2}}$$

Essa quantidade fica MUITO pequena quando o valor de *n* é grande!!

Função de verossimilhança

Na prática nós sempre calculamos o **logarítmo da verossimilhança**, pois as operações de exponenciação são transformadas em multiplicações.

$$\log(a^b) = b\log(a)$$

$$\log(ab) = \log(a) + \log(b)$$

Logarítimo da função de verossimilhança

$$\log P(y_1, y_2, \ldots, y_n|\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j - \mu)^2$$

Superfície de verossimilhança

Usamos o logarítimo da função de verossimilhança para calcular a verossimilhança dos valores dos parâmetros do modelo.

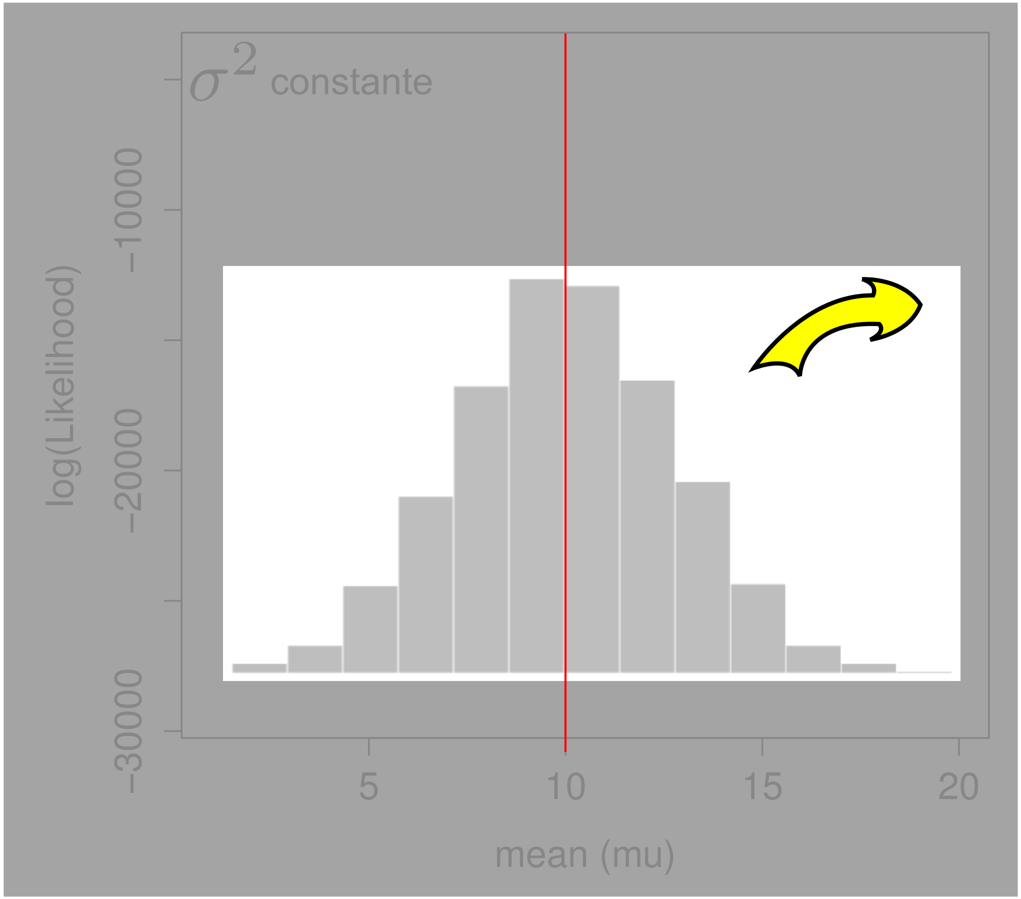
Se uma série de valores de parâmetros forem avaliadas podemos traçar um gráfico da verossimilhança em função dos valores de parâmetros.

Logarítimo da função de verossimilhança

$$\log P(y_1, y_2, \ldots, y_n | \mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j - \mu)^2$$

Superfície de verossimilhança

Gráfico da verossimilhança em função dos valores da média.

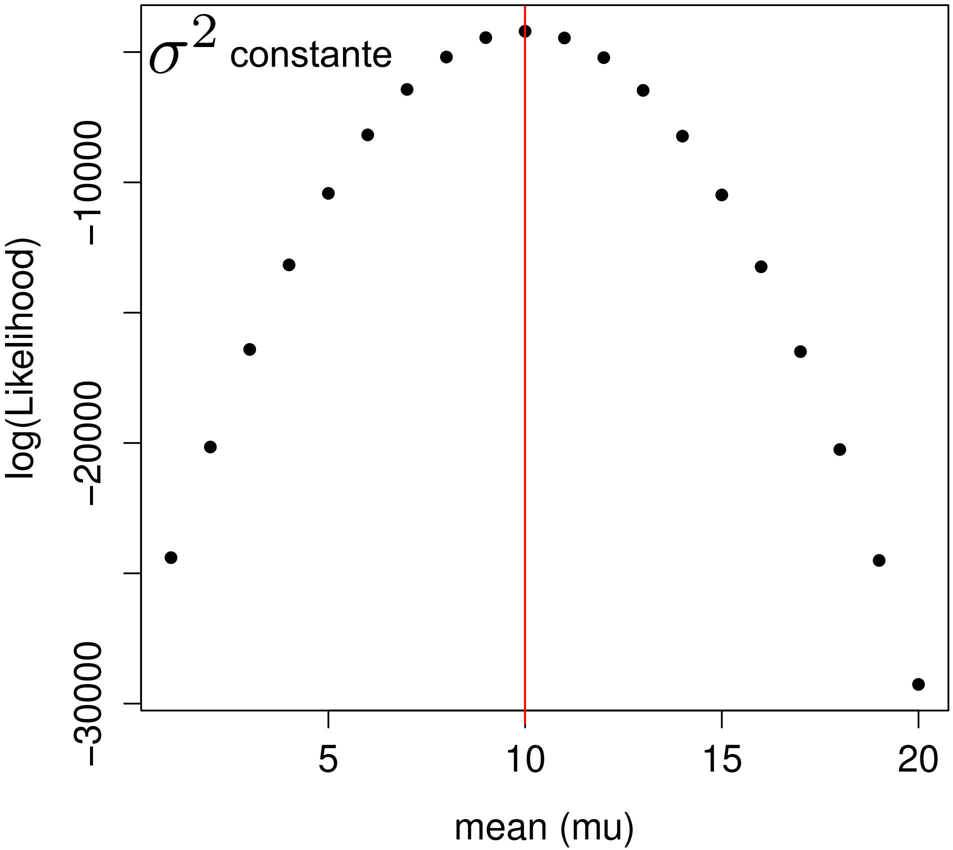


Distribuição dos dados...

Como é a curva do gráfico de verossimilhança?

Superfície de verossimilhança

Gráfico da verossimilhança em função dos valores da média.

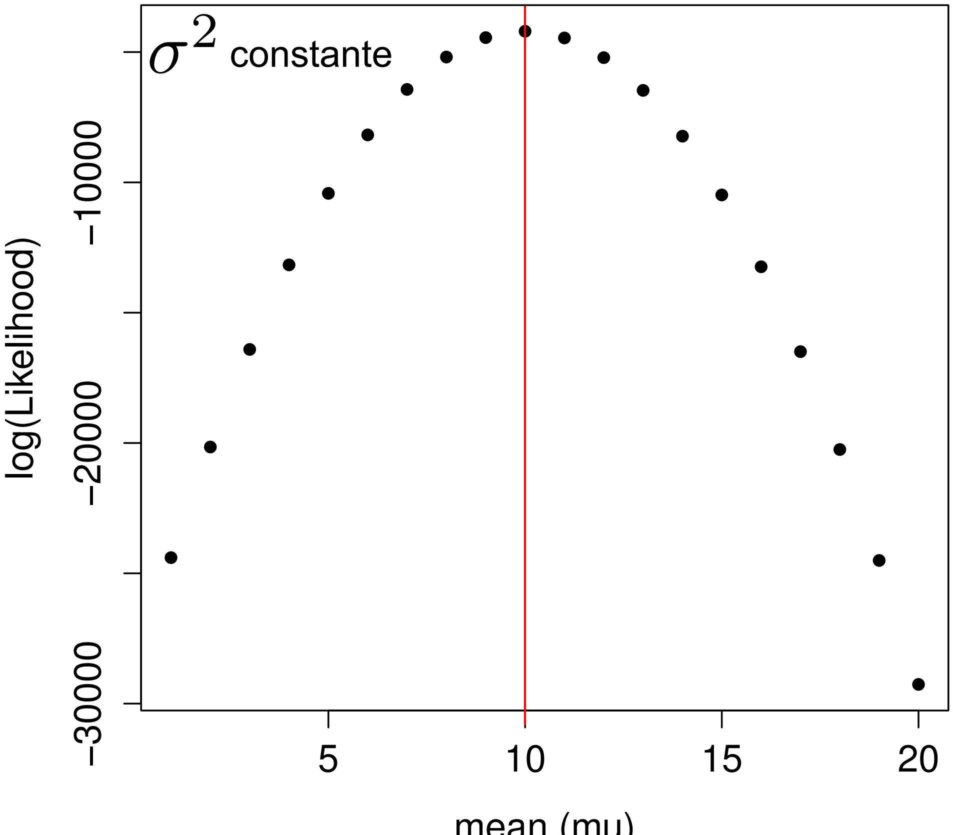


Distribuição dos dados...

Como é a curva do gráfico de verossimilhança?

Superfície de verossimilhança

Gráfico da verossimilhança em função dos valores da média.

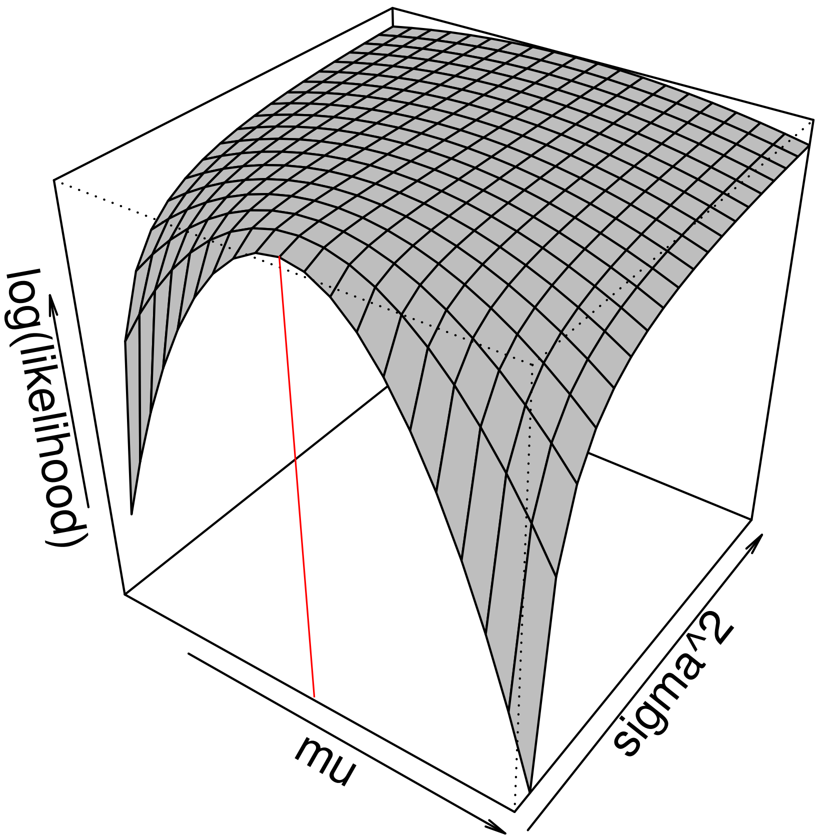


Esse gráfico mostra a verossimilhança do modelo com base em uma série de valores para a média.

A curva é o perfil de verossimilhança para μ condicionado em um valor de σ.

Superfície de verossimilhança

Superfície de verossimilhança para o modelo normal.

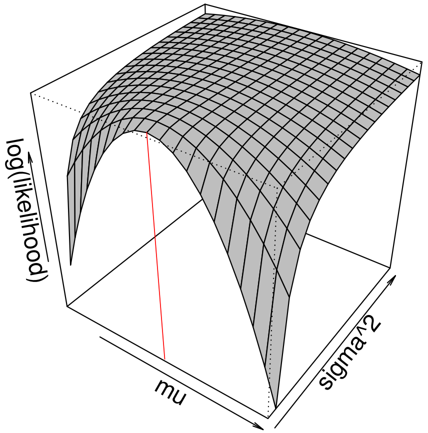


Valores de *log likelihood* para diversas combinações de μ e σ^2 produzem a superfície de likelihood.

Nosso objetivo principal é explorar métodos que *navegam* por essa superfície para achar as melhores combinações de valores de parâmetros.

MLE

Espere... Eu sei calcular a média e o
desvio padrão de uma distribuição!

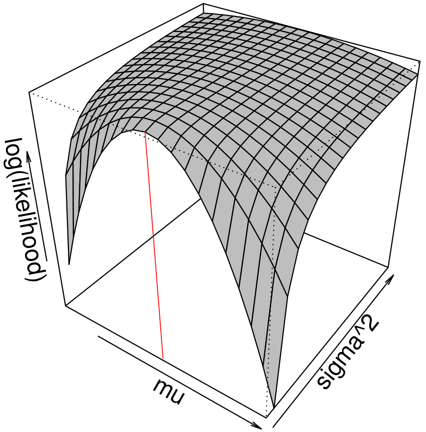


Sim, nós podemos calcular os parâmetros
que maximizam a função de verossimilhança.

Efetuamos uma *Maximum Likelihood Estimate*.

MLE

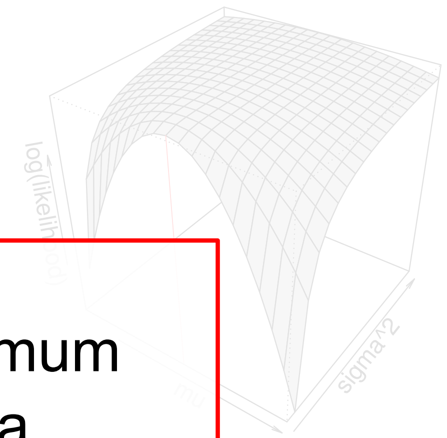
$$P(y|\mu,\sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} e^{\frac{-1}{2\sigma^2}\sum_{j=1}^{n}(y_j-\mu)^2}$$



Maximizando a verossimilhança usando algebra
nós derivamos a seguinte equação:

$$MLE = \bar{X} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

MLE



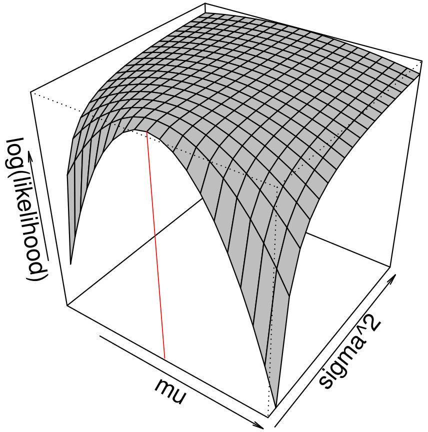
Mostrar a derivação da Maximum Likelihood Estimate para a distribuição normal.

$$MLE = \bar{X} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

MLE

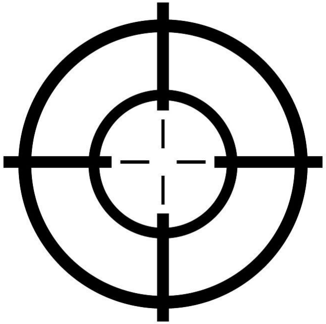
Podemos utilizar um **método de busca** pela MLE:

$$P(y|\mu,\sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} e^{\frac{-1}{2\sigma^2}\sum_{j=1}^{n}(y_j-\mu)^2}$$



Aplicamos uma **derivação direta** da MLE:

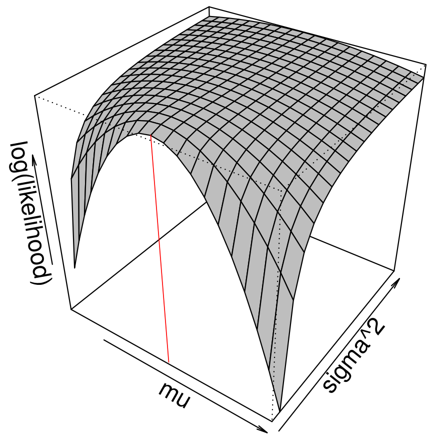
$$MLE = \bar{X} = \frac{1}{n}\sum_{j=1}^{n} x_j$$



MLE

Vamos utilizar métodos de busca

$$P(y|\mu,\sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} e^{\frac{-1}{2\sigma^2}\sum_{j=1}^{n}(y_j-\mu)^2}$$



A partir de agora vamos focar em métodos de busca.

Vamos explorar como funcionam buscas usando critérios de Máxima Verossimilhança e inferência Bayesiana.

Pontos importantes

Alguns pontos serão fundamentais:

- O ajuste de um modelo depende dos valores de parâmetros.
- Função de verossimilhança calcula o valor relativo de ajuste aos dados.
- Buscamos os valores de parâmetros que melhor se ajustam aos nossos dados.