



LAC-FMB

Liga
Acadêmica de
Ciências

A sessão começará às 18h30

Este documento foi gerado a partir de uma sessão da Liga Acadêmica de Ciências da Faculdade de Medicina da Bahia.

Todos os direitos estão reservados aos seus autores e à liga. Não edite o documento, não tire print. A utilização da sua imagem está proibida.

LAC-FMB

Regressão Linear Simples

Módulo de Bioestatística

Evelyn Sued

Sávio Gomes

Ligante Suporte: Thaís Barbosa



Sumário



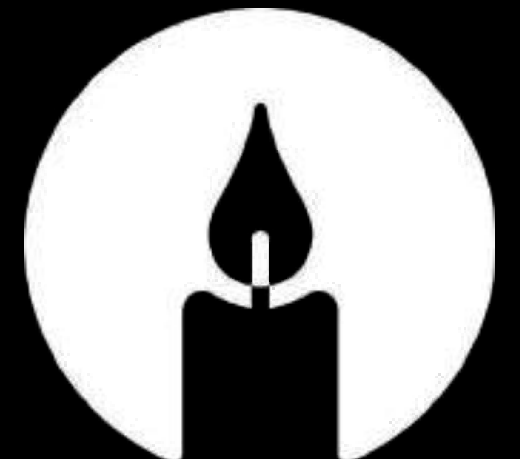
LAC-FMB
Liga Acadêmica de Ciências

Histórico da regressão

#01

Introdução

Módulo de Bioestatística



Histórico



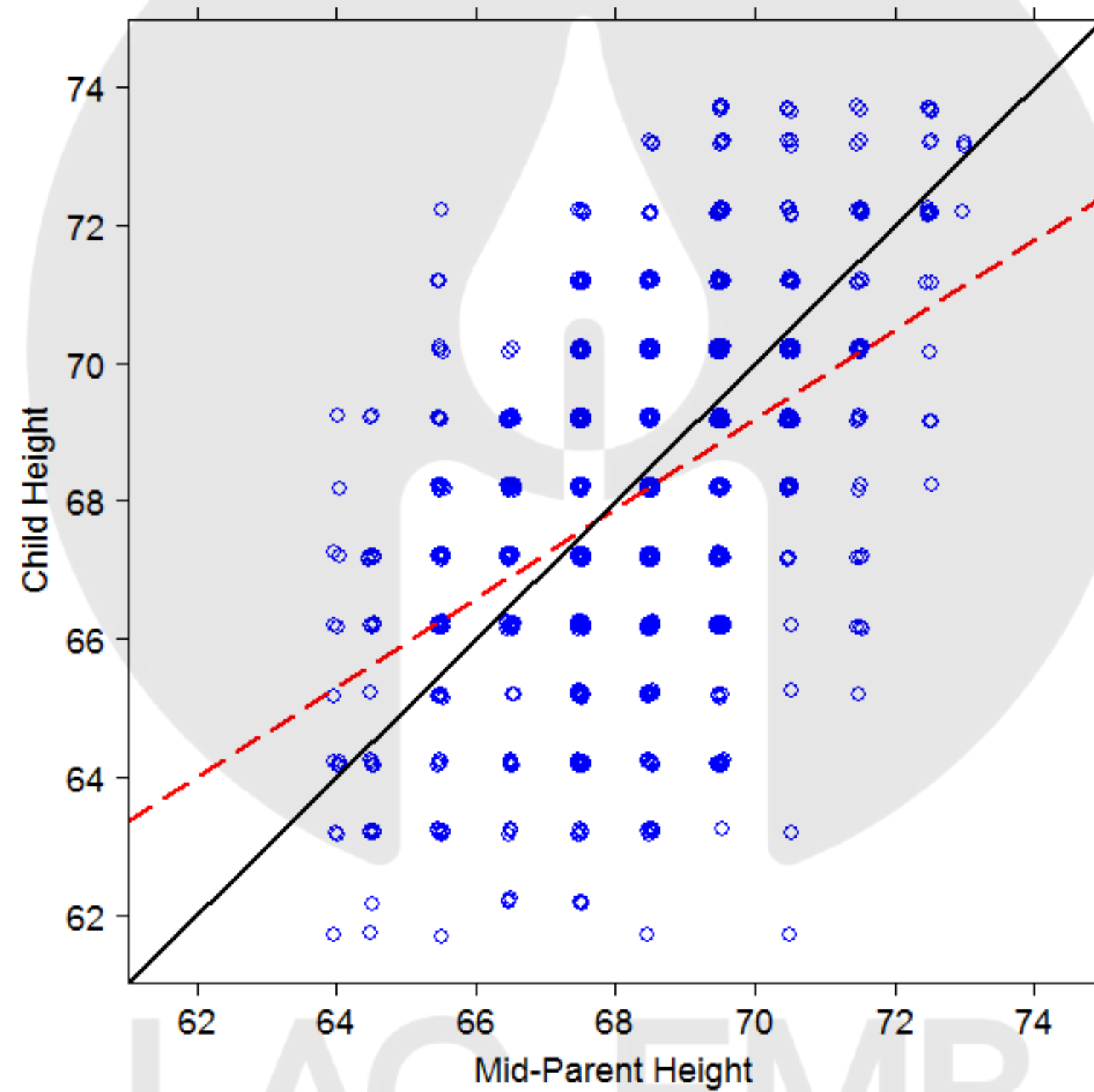
Francis Galton



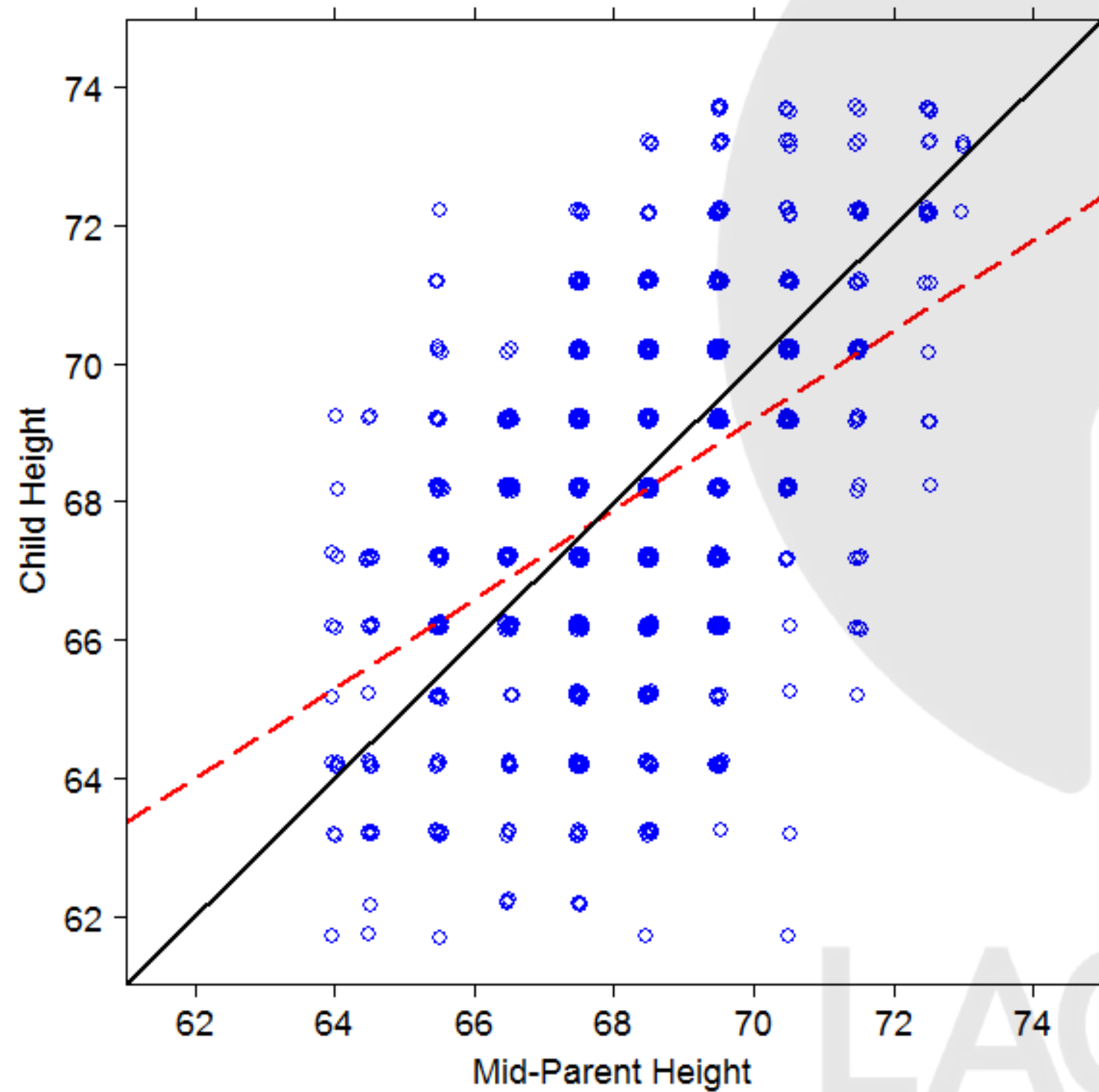
Histórico



Francis Galton



Histórico



“Filhos de pais altos tendem a ser mais baixos e filhos de pais baixos tendem a ser mais altos”



Francis Galton

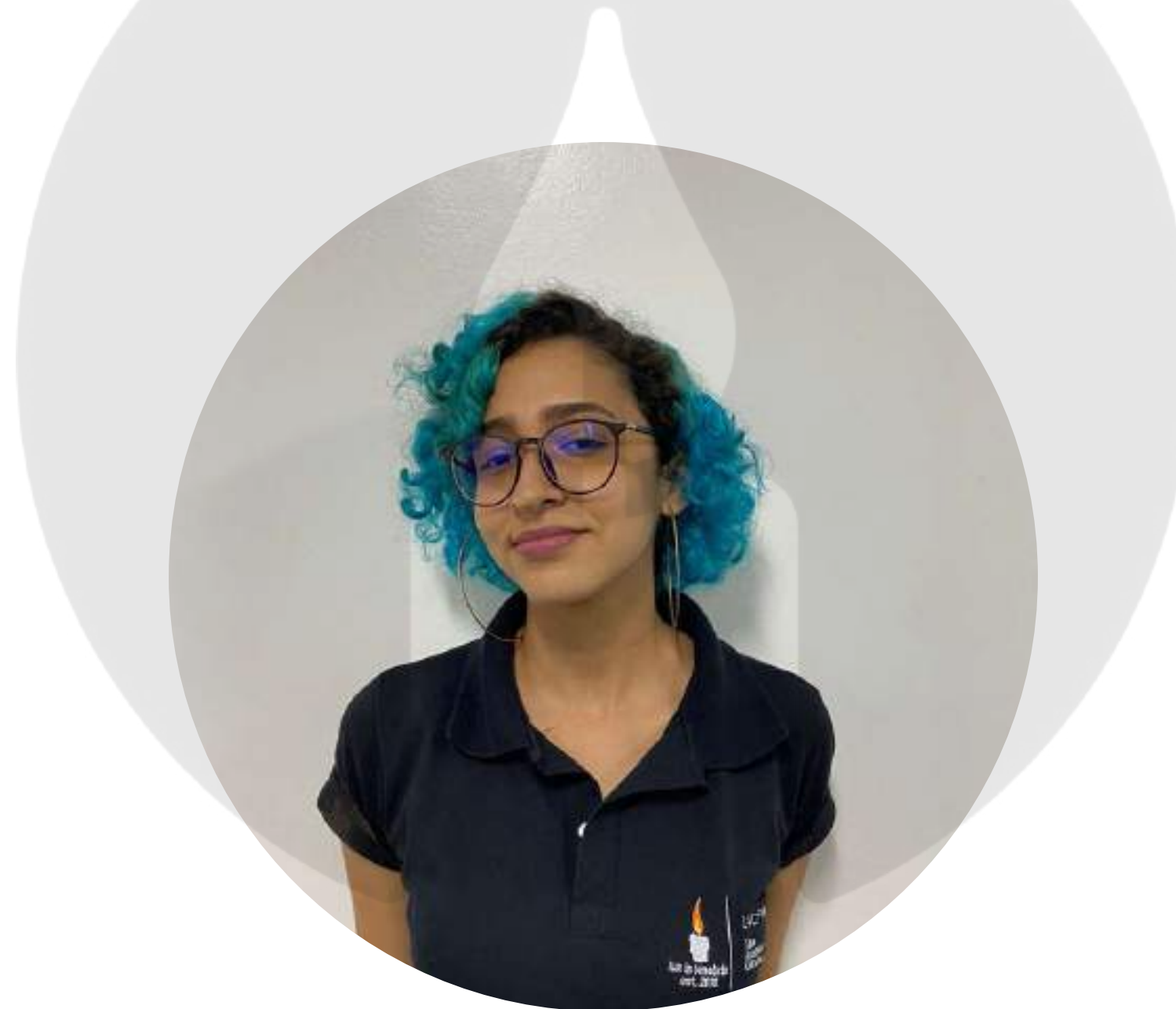




Beleza, Sávio. Mas o que é a
regressão linear simples?

LAC FMD

Contexto



Rebeca Borges



Contexto



Rebeca Borges

“A partir da quantidade de refeições diárias consigo prever a glicemia do meu paciente?”

LAC-FMB



Contexto



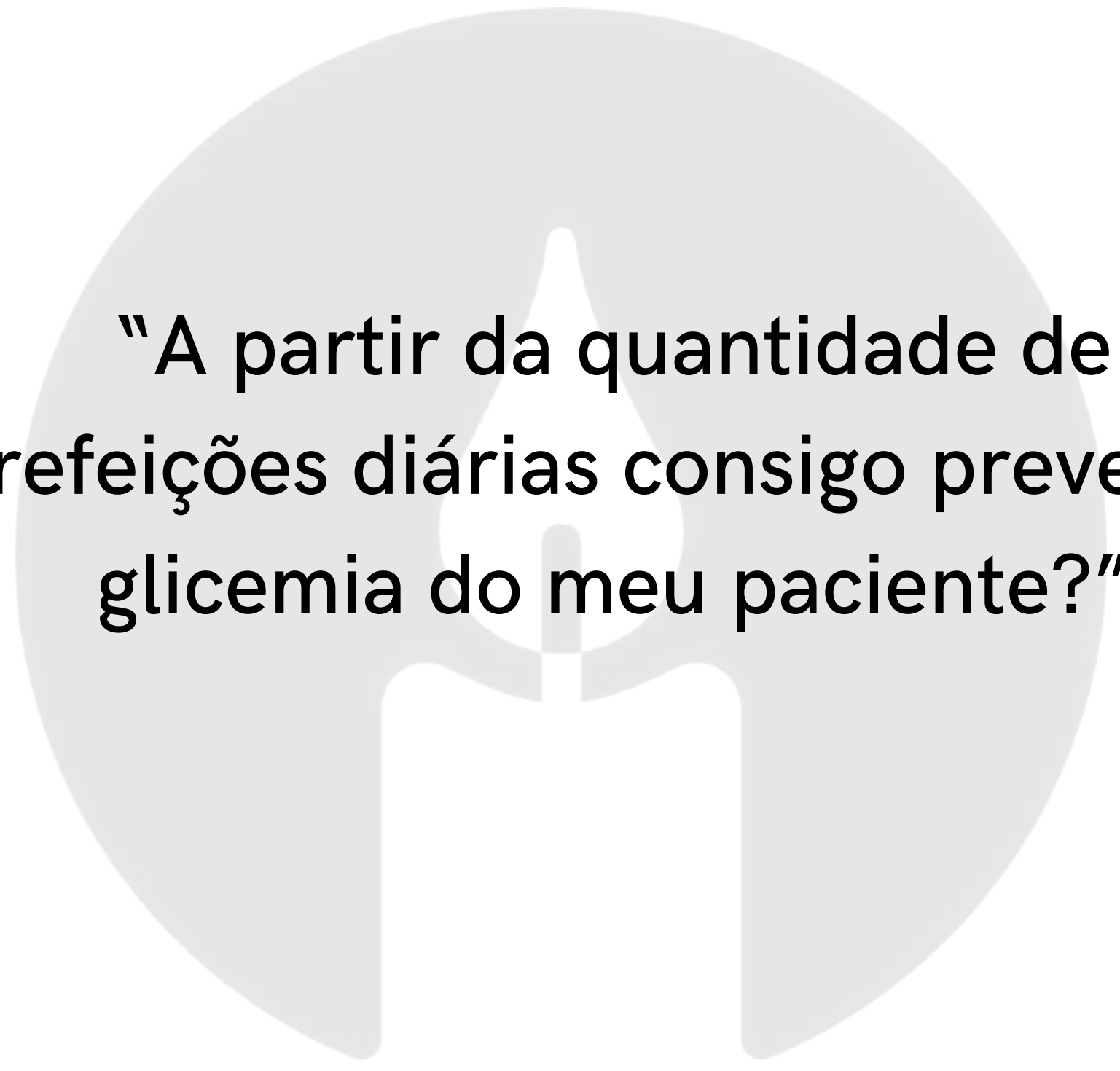
Guache

O que são variáveis independentes e dependentes?

“A partir da quantidade de refeições diárias consigo prever a glicemia do meu paciente?”

Rebeca Borges





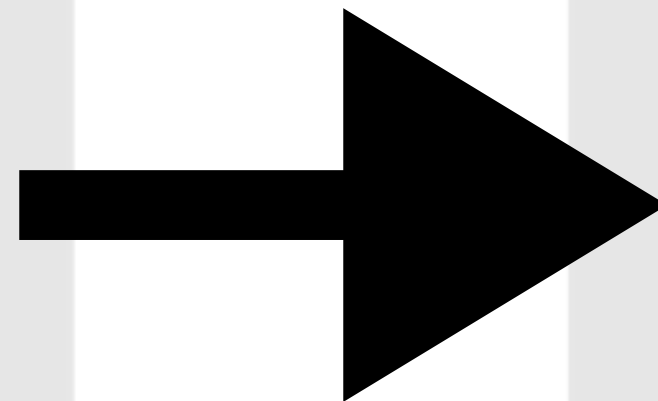
“A partir da quantidade de refeições diárias consigo prever a glicemia do meu paciente?”

LAC-FMB



“A partir da **quantidade de refeições diárias** consigo prever a glicemia do meu paciente?”

Q. de refeições
diárias



↓
Variável independente
ou preditora
(Previsora)



“A partir da **quantidade de refeições diárias** consigo prever a **glicemia** do meu paciente?”



REGRESSÃO LINEAR SIMPLES

“A partir da **quantidade de refeições diárias** consigo prever a **glicemia** do meu paciente?”



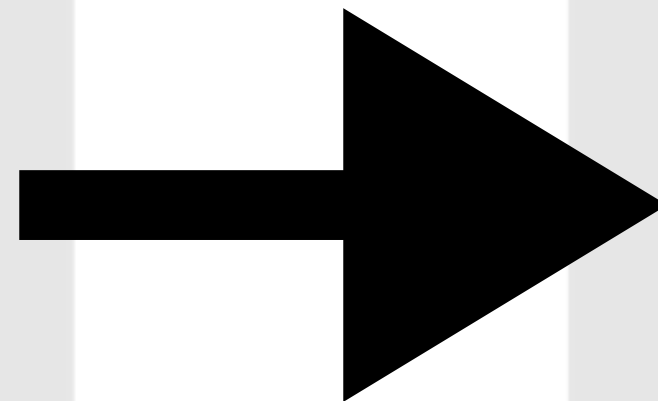
“A partir da quantidade de refeições diárias, de exercícios por semana e os níveis de cortisol consigo prever a glicemia do meu paciente?”

LAC-FMB



"A partir da **quantidade de refeições diárias, de exercícios por semana e os níveis de cortisol** consigo prever a glicemia do meu paciente?"

Q. de refeições diárias
Exercícios/semana
Níveis de cortisol



↓
Variável independente
ou preditora
(Previsora)



"A partir da **quantidade de refeições diárias, de exercícios por semana e os níveis de cortisol** consigo prever a **glicemia** do meu paciente?"

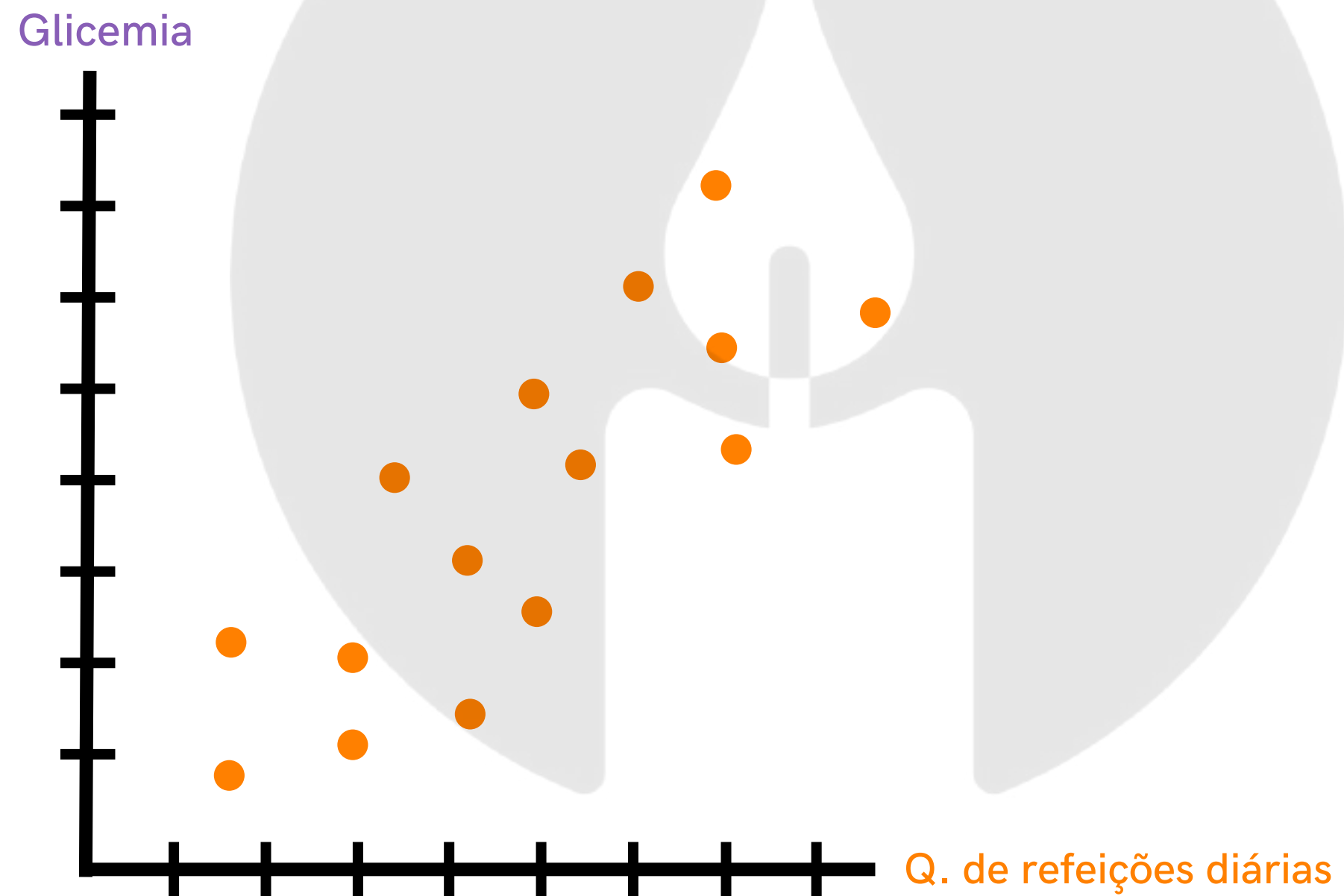


REGRESSÃO LINEAR MÚLTIPLA

"A partir da quantidade de refeições diárias, de exercícios por semana e os níveis de cortisol consigo prever a glicemia do meu paciente?"



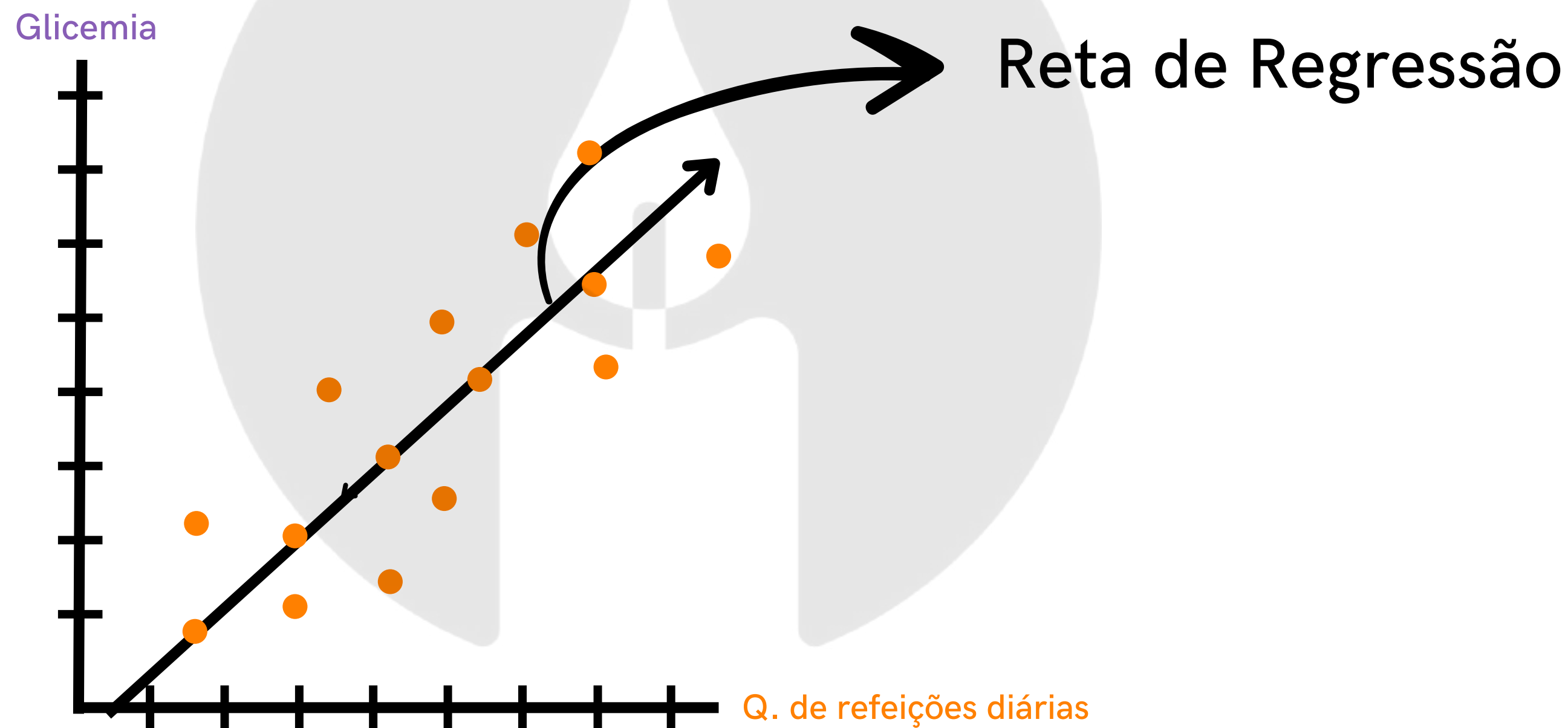
Reta de Regressão



LAC-FMB



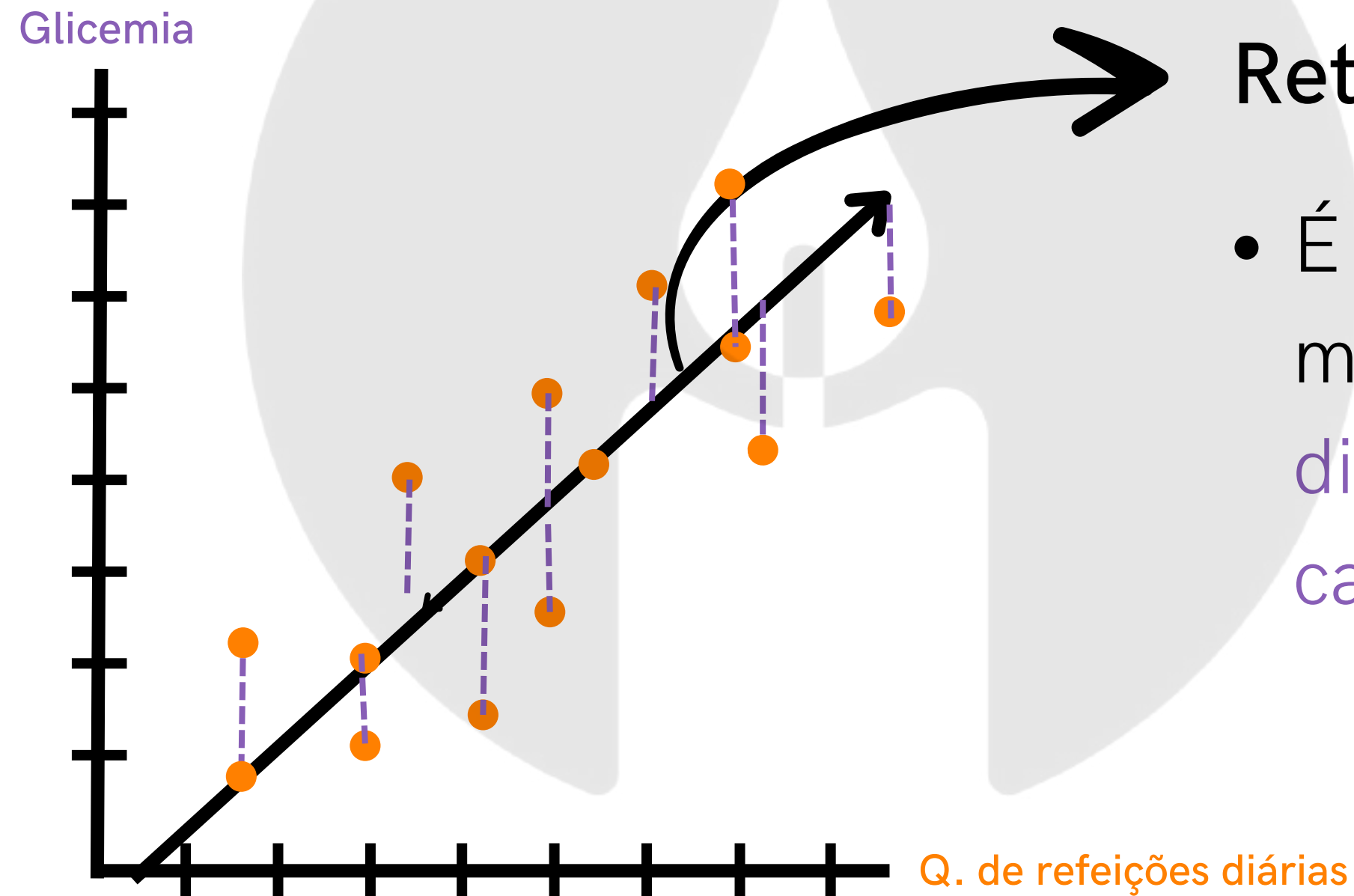
Reta de Regressão



LAC-FMB



Reta de Regressão



Reta de Regressão

- É a reta que tem a menor somatória da distância vertical de cada ponto (RESÍDUO)

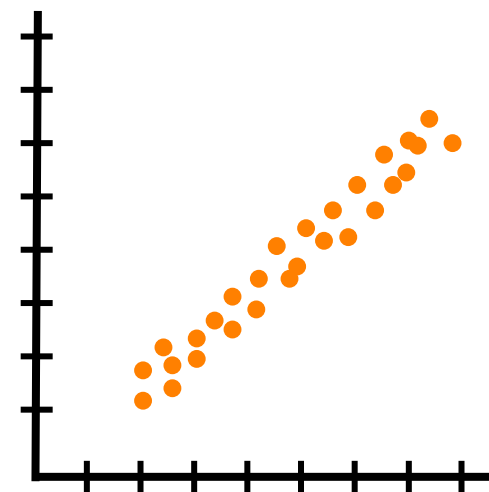
LAC-FMB



Correlação

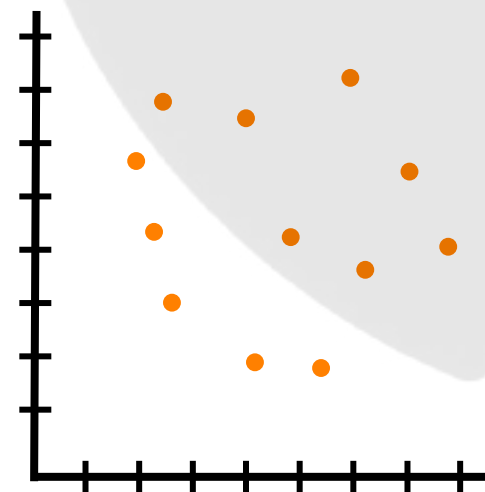
→ Quero saber se duas variáveis são relacionadas de maneira significativa.

Glicemia



Q. de refeições
diárias

Glicemia



Q. de refeições
diárias



Correlação

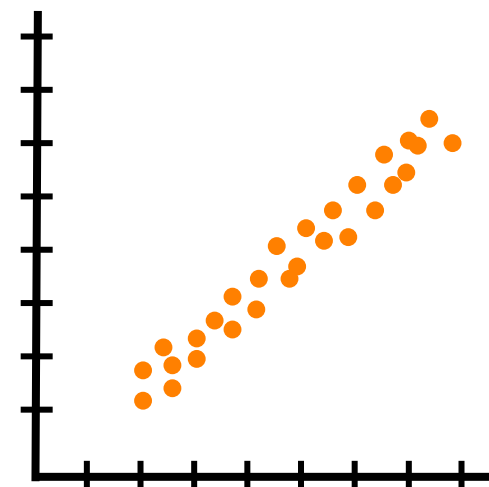
X

Regressão

→ Quero saber se duas variáveis são relacionadas de maneira significativa.

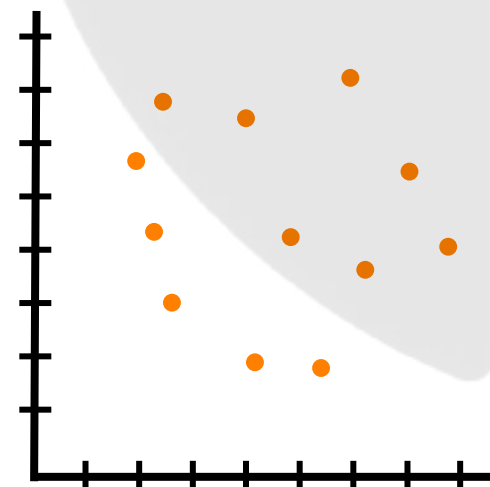
→ Quero saber se uma variável (independente) é uma boa preditora para prever o valor de outra (dependente)

Glicemia



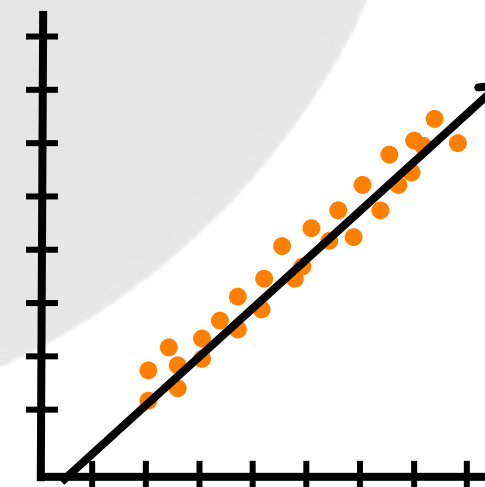
Q. de refeições
diárias

Glicemia



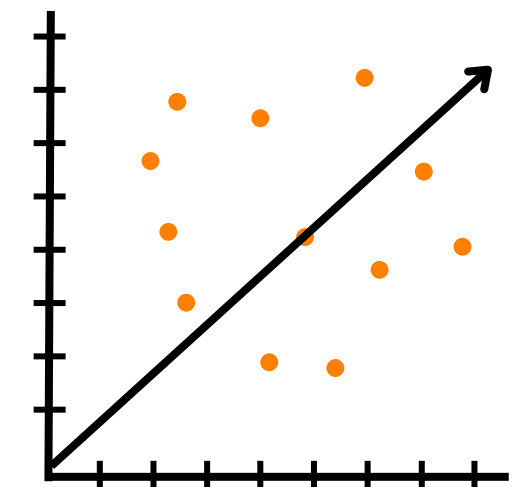
Q. de refeições
diárias

Glicemia



Q. de refeições
diárias

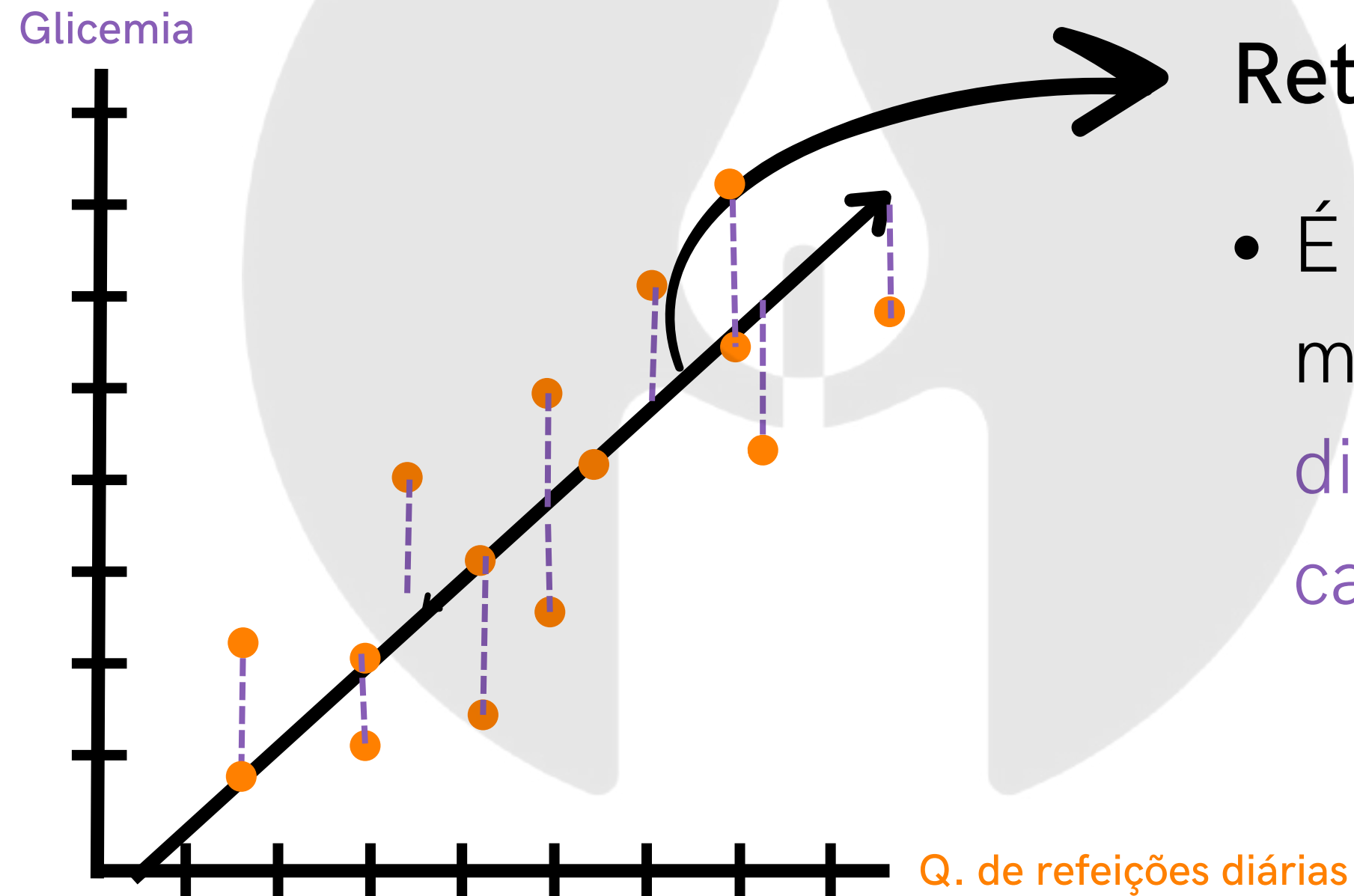
Glicemia



Q. de refeições
diárias



Reta de Regressão



Reta de Regressão

- É a reta que tem a menor somatória da distância vertical de cada ponto (RESÍDUO)

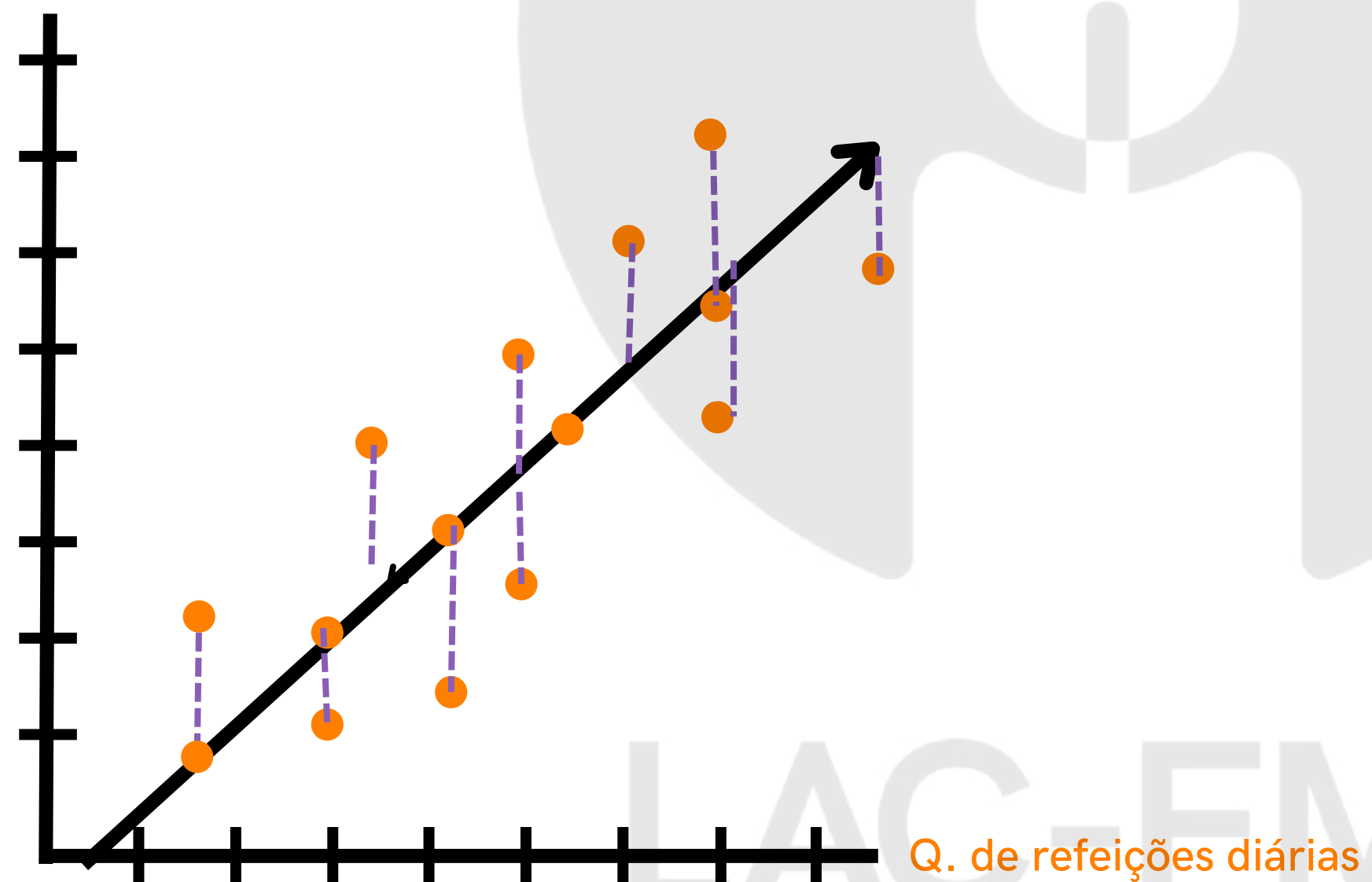
LAC-FMB



Reta de Regressão

$$Y = aX + b$$

Glicemia



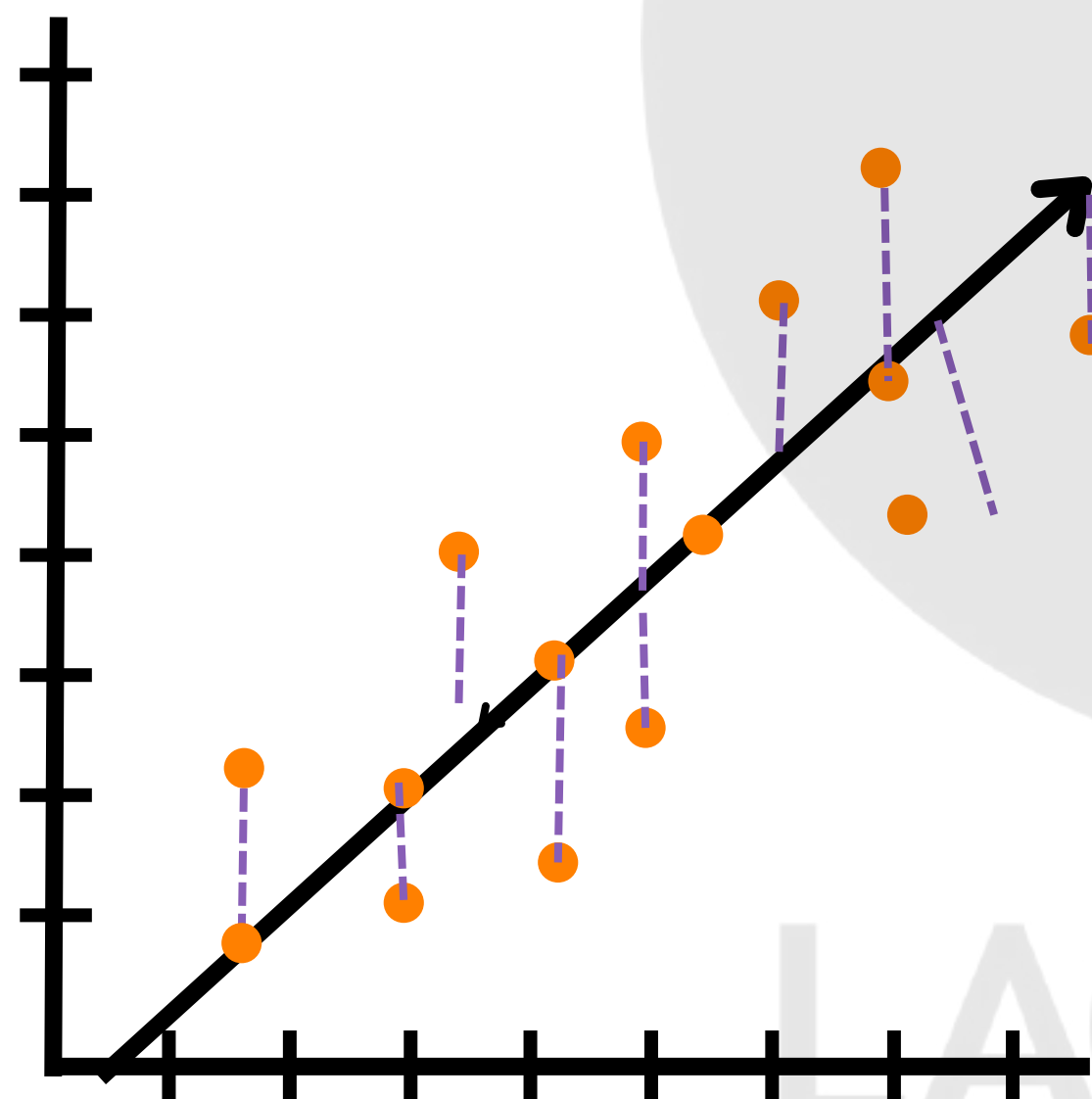
Q. de refeições diárias



Reta de Regressão

$$Y = aX + b$$

Glicemia



Q. de refeições diárias

Coeficientes da reta

a -coeficiente angular

b -coeficiente linear

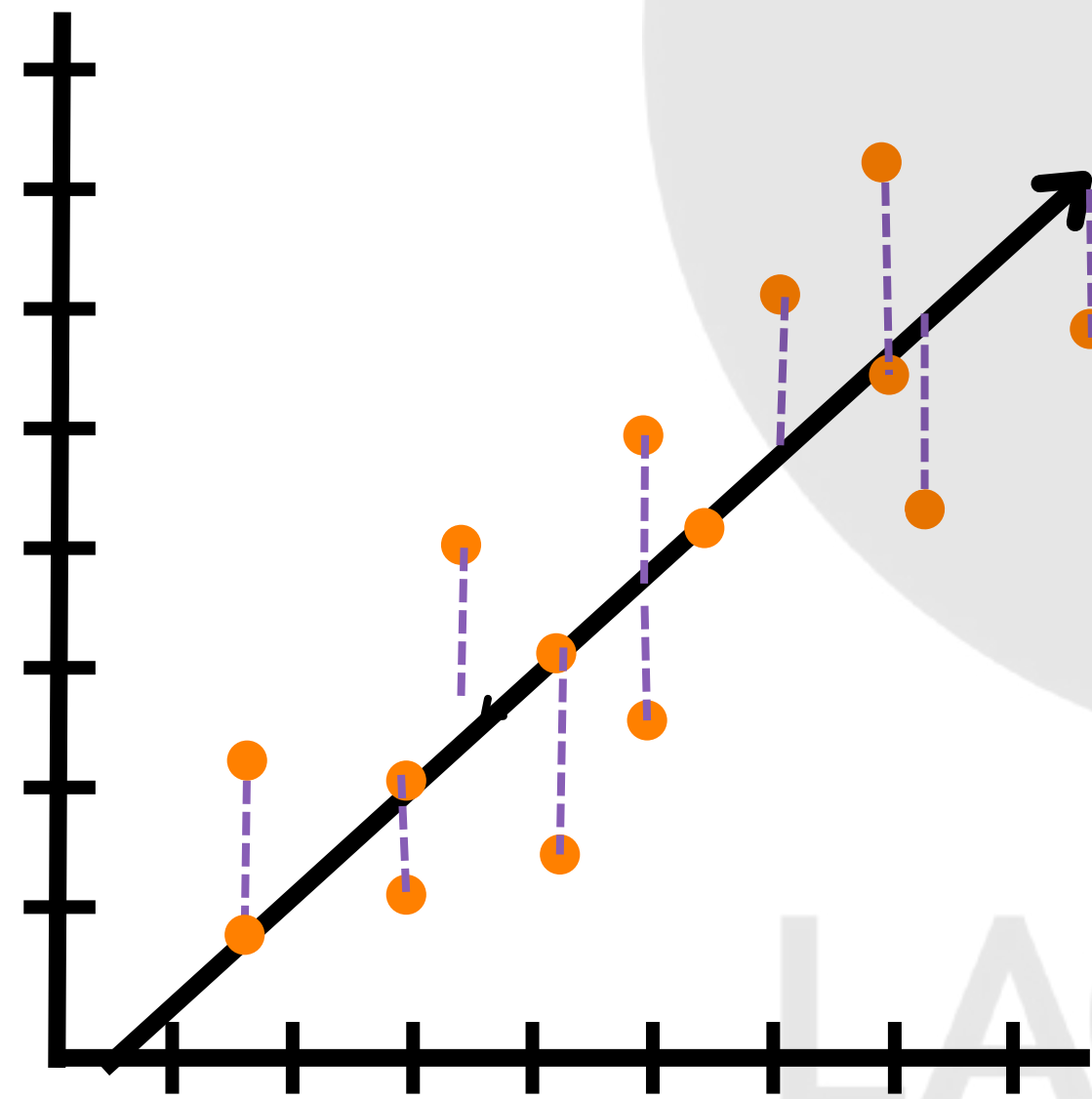


Reta de Regressão

$$y = \beta_1 x + \beta_0$$

$$\begin{array}{c} y = ax + b \\ \downarrow \quad \downarrow \\ y = \beta_1 x + \beta_0 \end{array}$$

Glicemia



Q. de refeições diárias

Coeficientes da reta

β_1 - coeficiente angular

β_0 - coeficiente linear

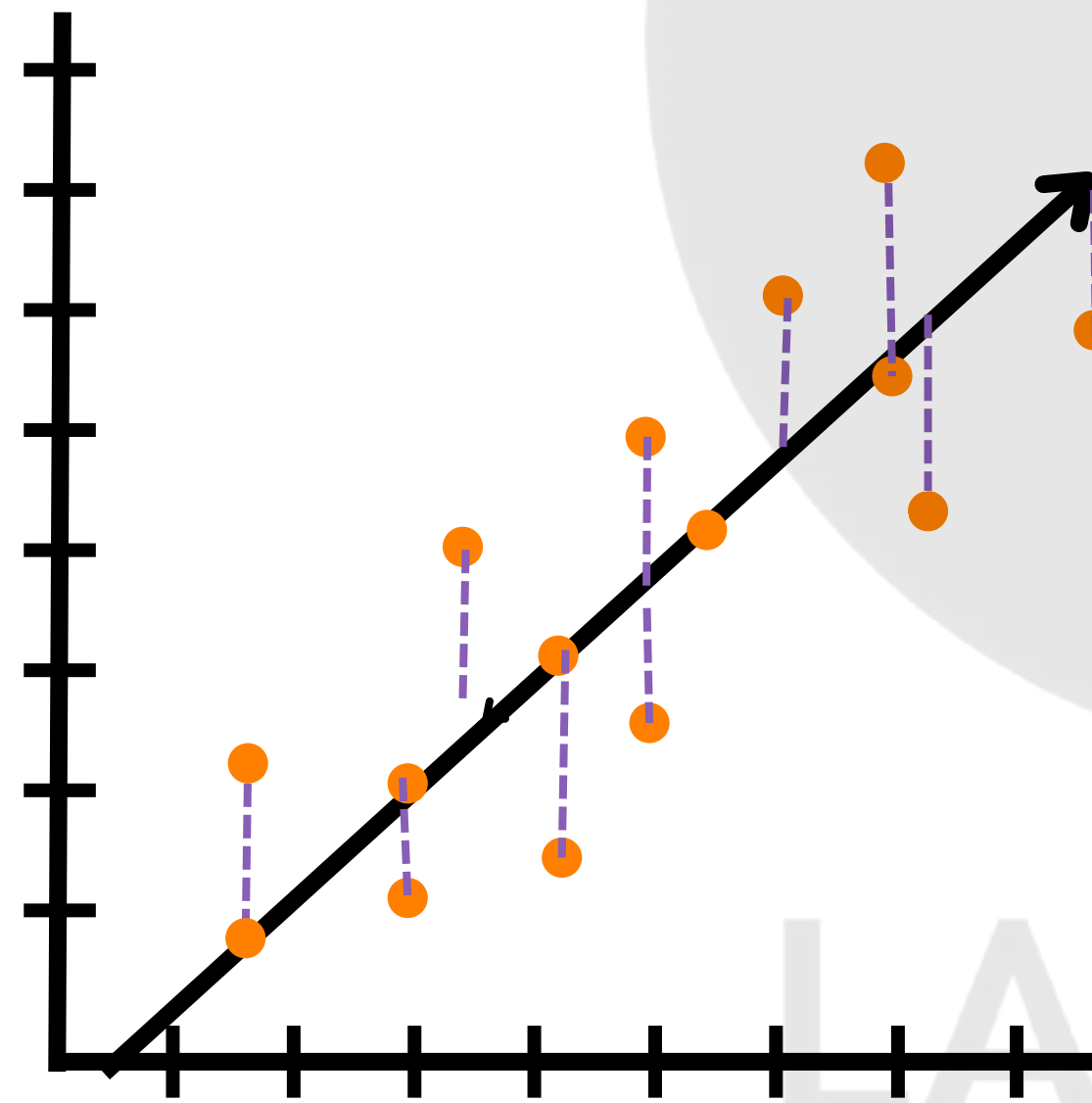


Reta de Regressão

$$y = \beta_1 x + \beta_0 + \epsilon_i$$

$$\begin{array}{l} y = ax + b \\ \downarrow \quad \downarrow \\ y = \beta_1 x + \beta_0 \end{array}$$

Glicemia



Q. de refeições diárias

Coeficientes da reta

β_1 -coeficiente angular

β_0 -coeficiente linear





Tá, mas por que eu quero
saber a equação da reta?

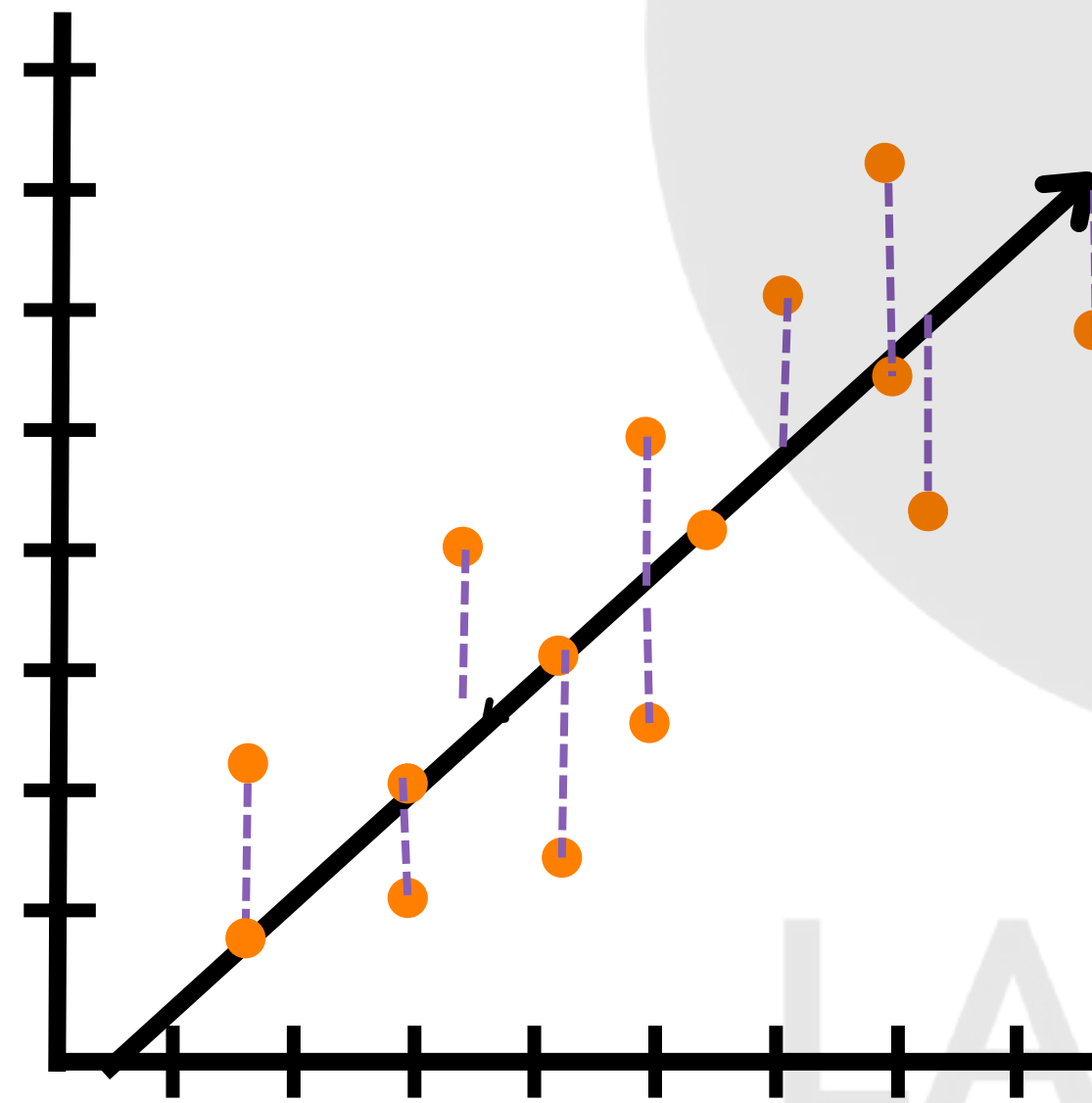
LAC FMD

Reta de Regressão

$$y = \beta_1 x + \beta_0$$

$$\begin{array}{c} y = ax + b \\ \downarrow \quad \downarrow \\ y = \beta_1 x + \beta_0 \end{array}$$

Glicemia



Coeficientes da reta

$$\beta_1 -50$$

$$\beta_0 -5$$

Q. de refeições diárias

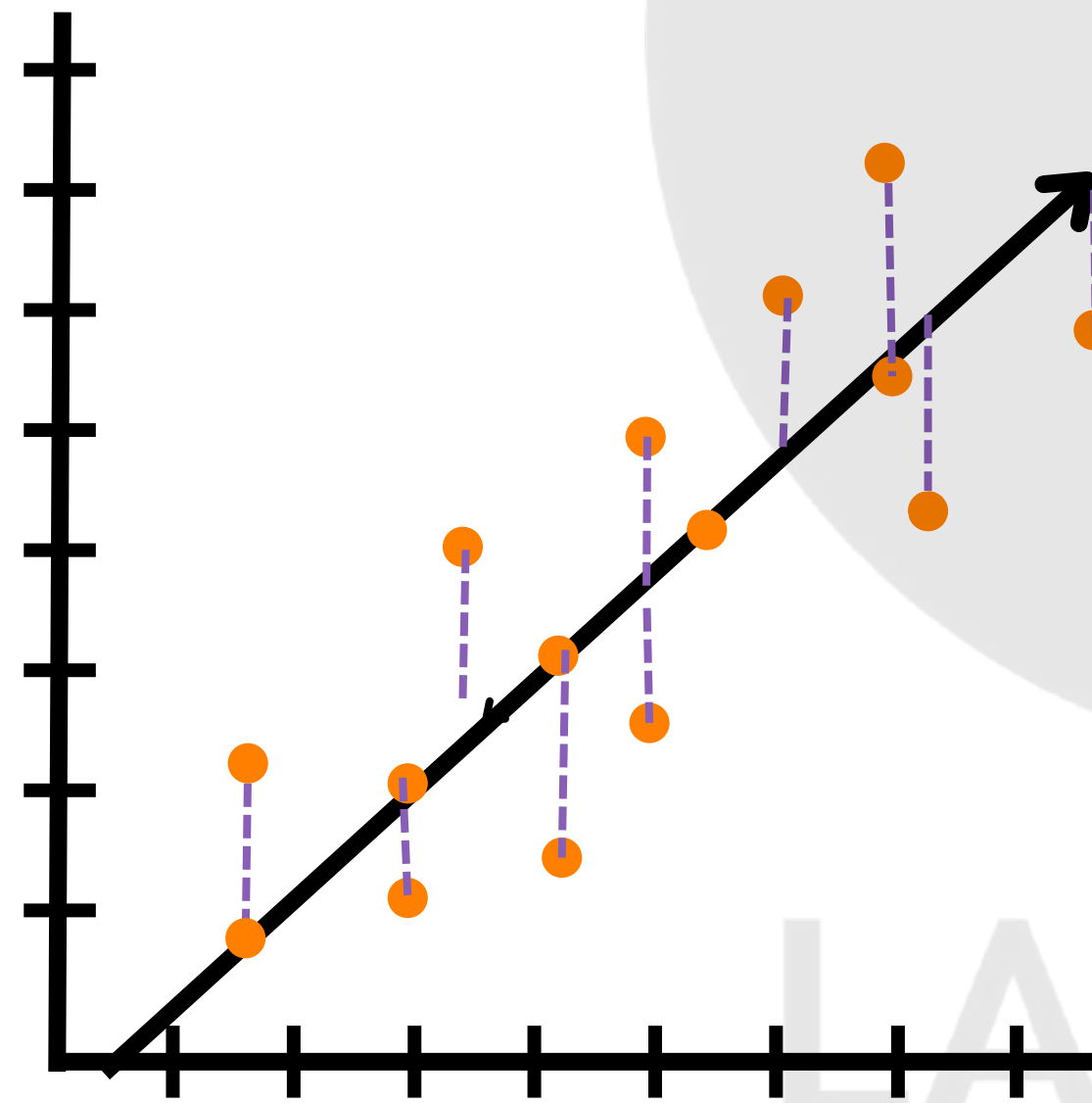


Reta de Regressão

$$y = \beta_1 x + \beta_0$$

$$\begin{array}{l} y = ax + b \\ \downarrow \quad \downarrow \\ y = \beta_1 x + \beta_0 \end{array}$$

Glicemia



Q. de refeições diárias

Coeficientes da reta

$$\beta_1 = 5$$

$$\beta_0 = 50$$

$$y = 5x + 50$$

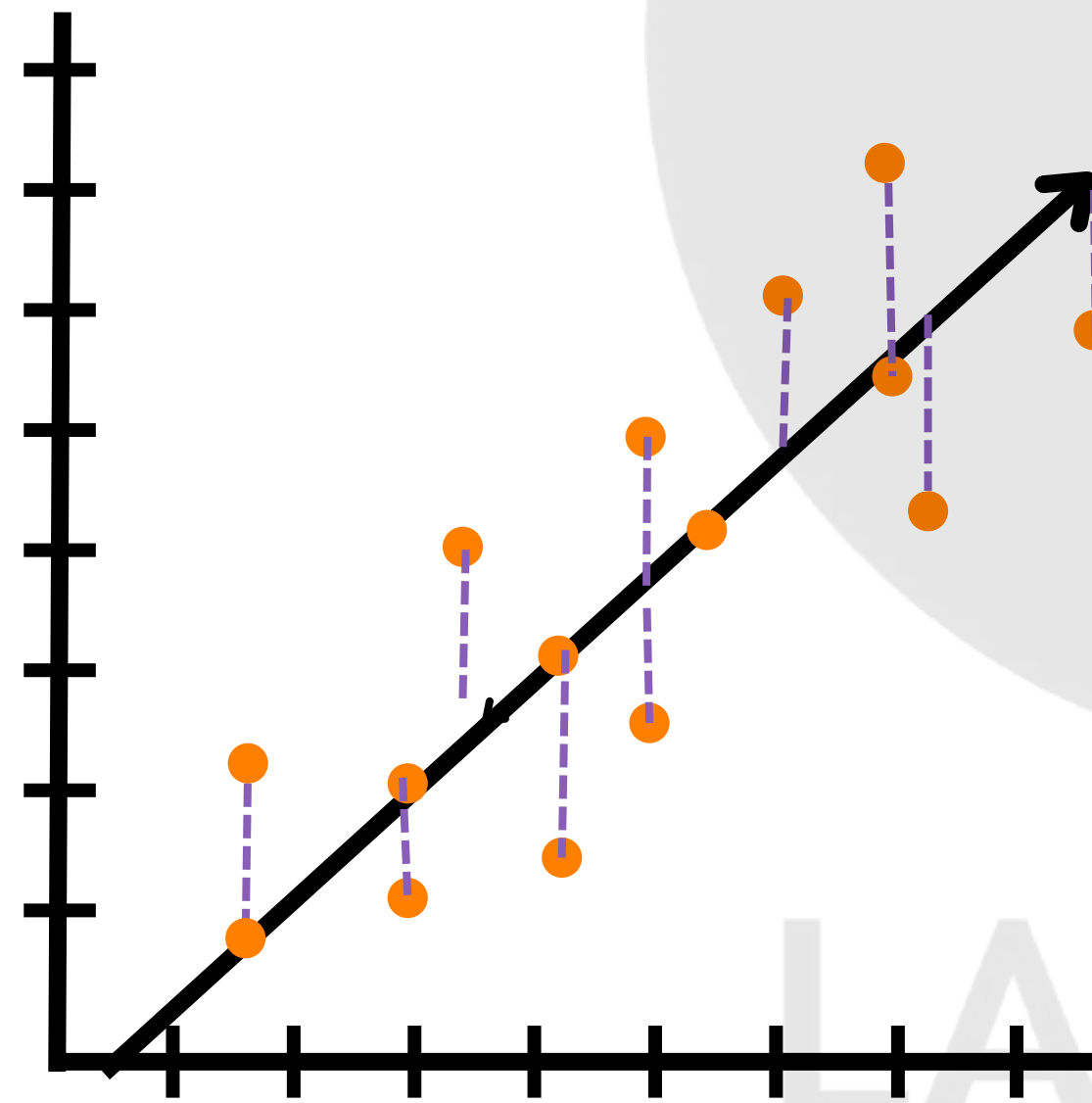


Reta de Regressão

$$y = \beta_1 x + \beta_0$$

$$\begin{array}{c} y = ax + b \\ \downarrow \quad \downarrow \\ y = \beta_1 x + \beta_0 \end{array}$$

Glicemia



Q. de refeições diárias

Coeficientes da reta

$$\beta_1 = 5$$

$$\beta_0 = 50$$

$$y = 5x + 50$$

$$y = 5 \cdot 3 + 50$$

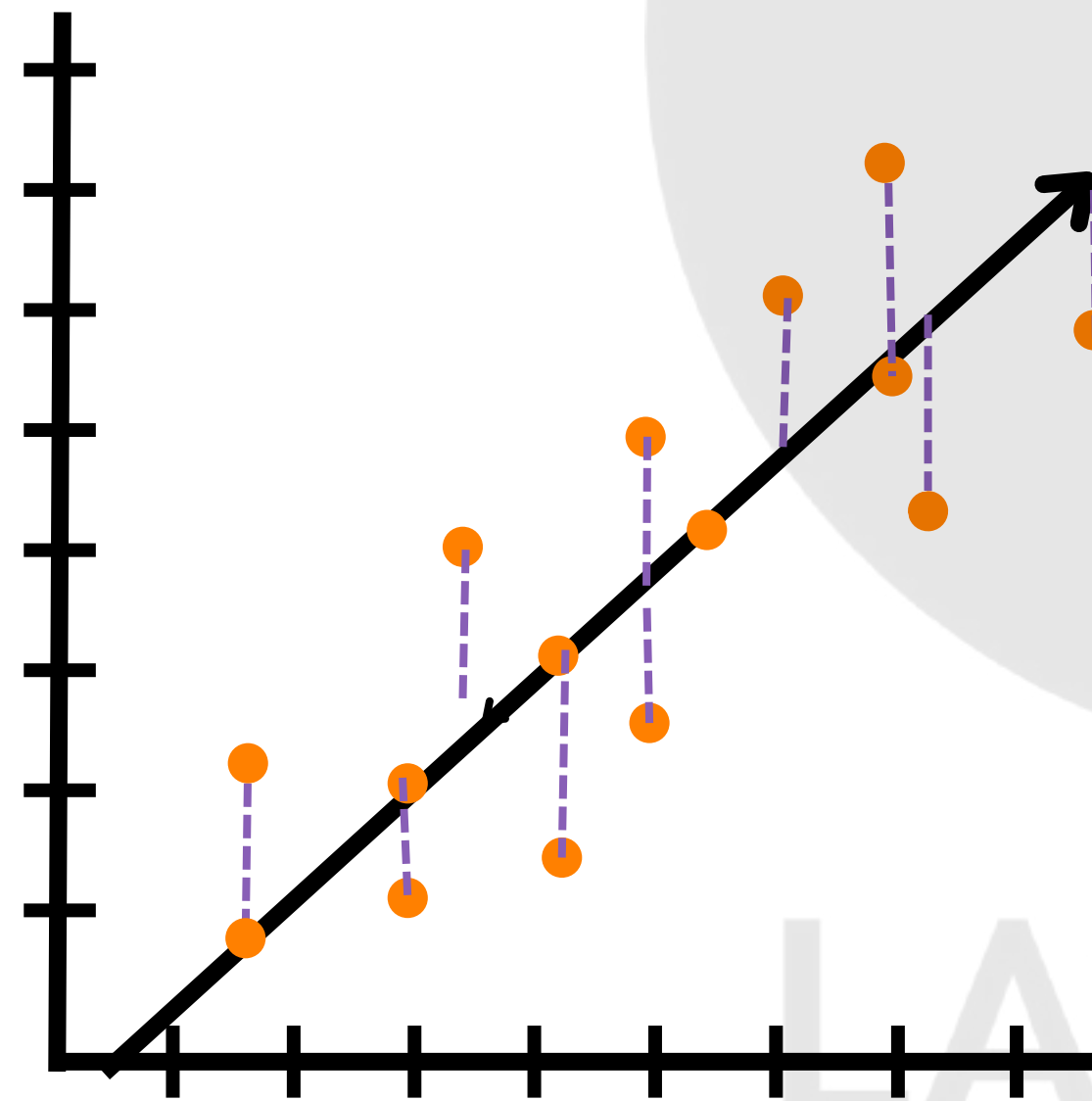


Reta de Regressão

$$y = \beta_1 x + \beta_0$$

$$\begin{array}{l} y = ax + b \\ \downarrow \quad \downarrow \\ y = \beta_1 x + \beta_0 \end{array}$$

Glicemia



Coeficientes da reta

$$\beta_1 -50$$

$$\beta_0 -5$$

$$y = 5x + 50$$

$$y = 5.3 + 50$$

$$65 = 5.3 + 50$$

Q. de refeições diárias

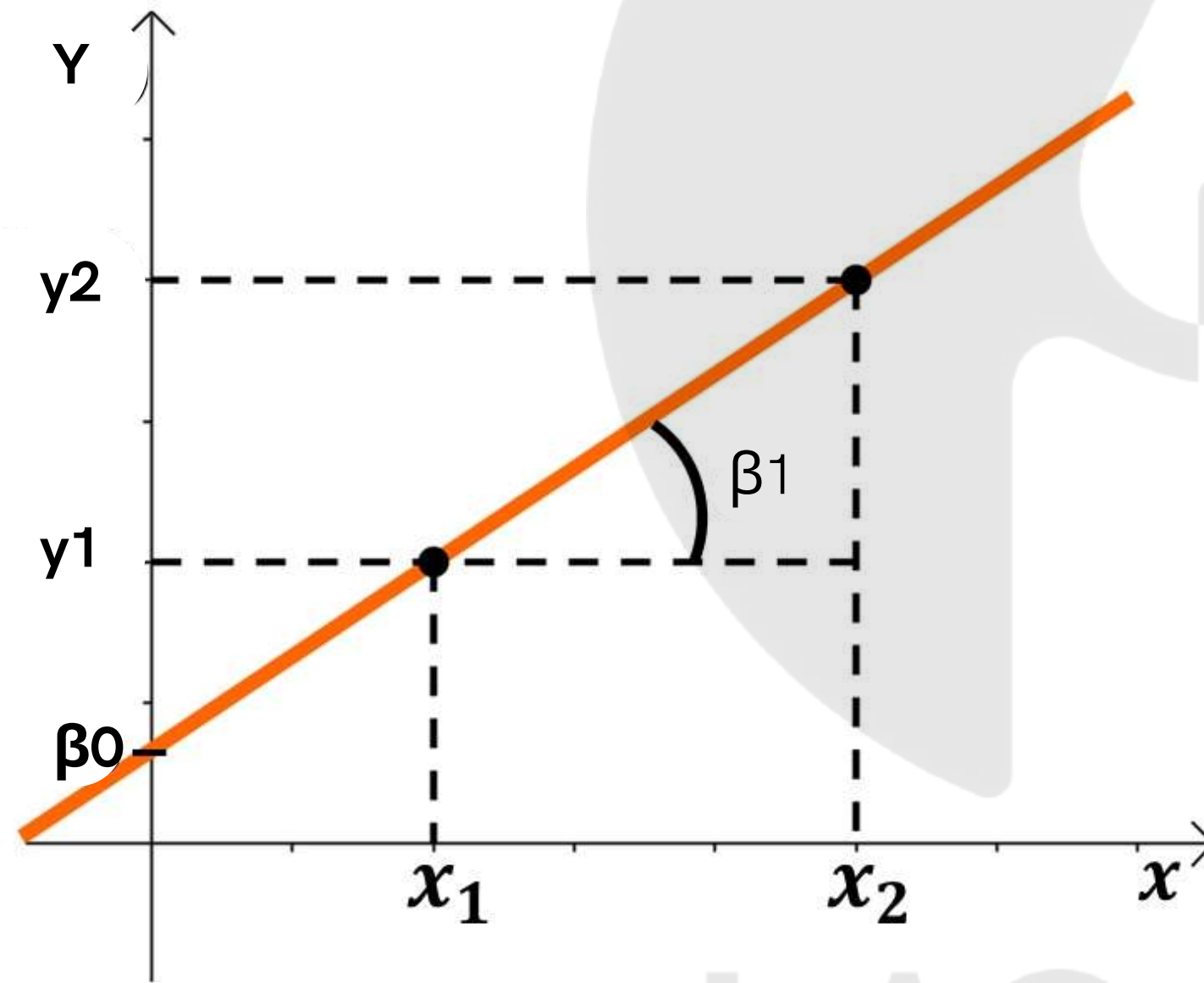




Mas então, como eu vou
achar os coeficientes da reta
de regressão?

LAC FMD

Clássico



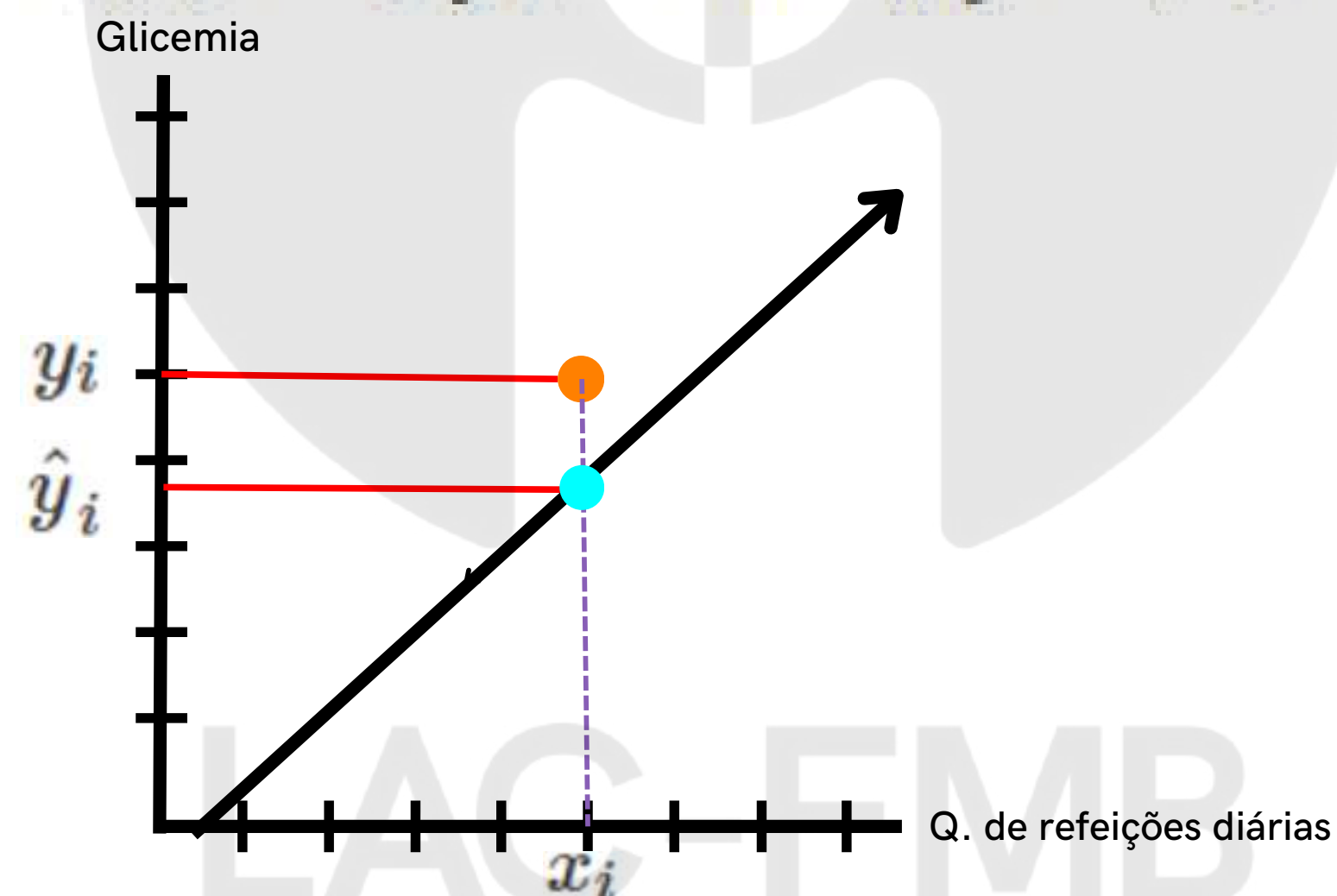
$$\begin{array}{l} y = ax + b \\ \quad \downarrow \quad \downarrow \\ y = \beta_1 x + \beta_0 \end{array}$$

$$\beta_1 = \frac{\Delta Y}{\Delta X} = \frac{y_2 - y_1}{x_2 - x_1}$$



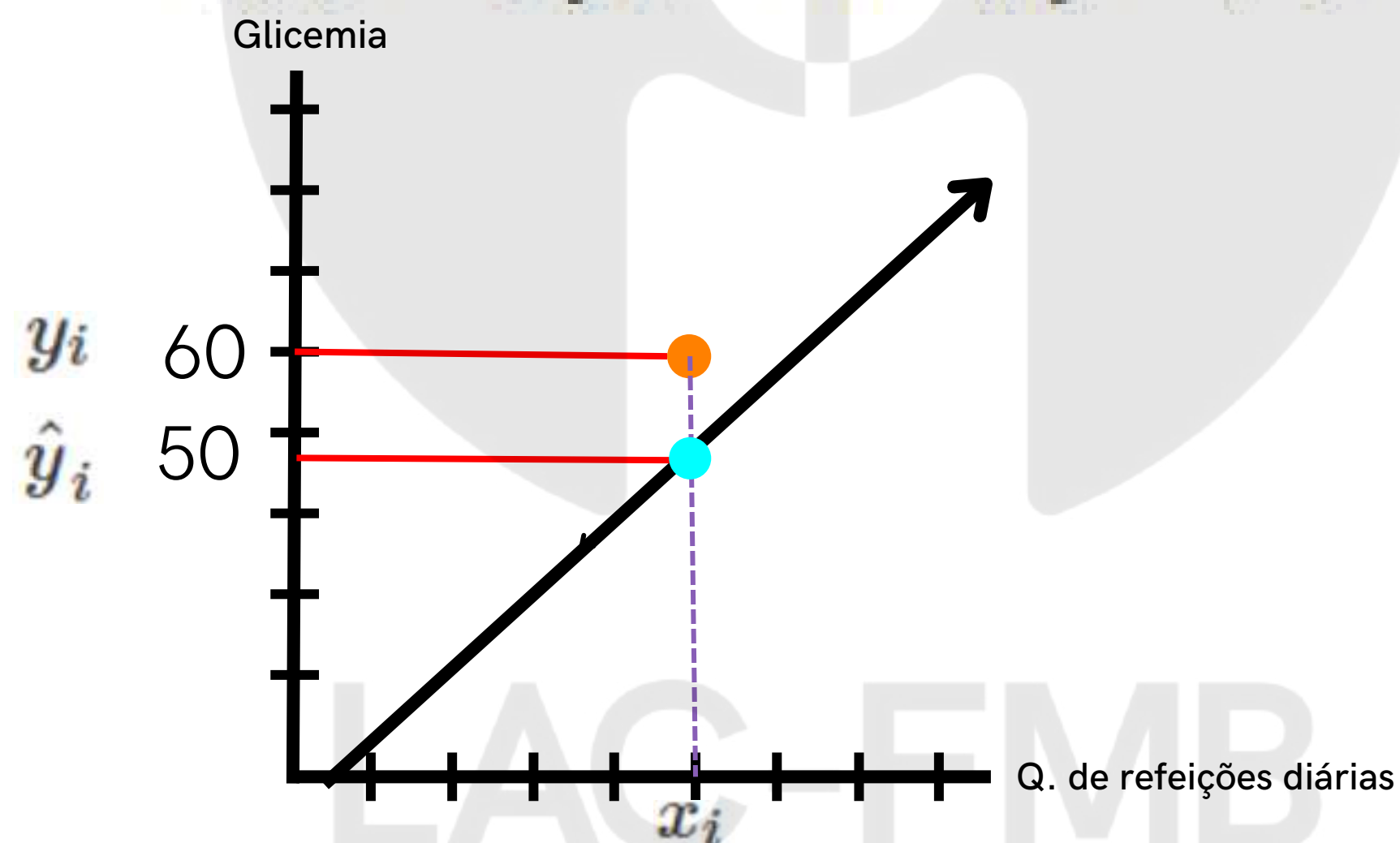
Método dos Mínimos Quadrados

2. Sendo $\hat{y}_i$ o ponto na reta de regressão mais próximo à y_i , um determinado ponto no gráfico, temos que a distância vertical do ponto à reta (RESÍDUO) é: $|y_i - \hat{y}_i|$, onde $\hat{y}_i = \beta_0 + \beta_1 x_i$



Método dos Mínimos Quadrados

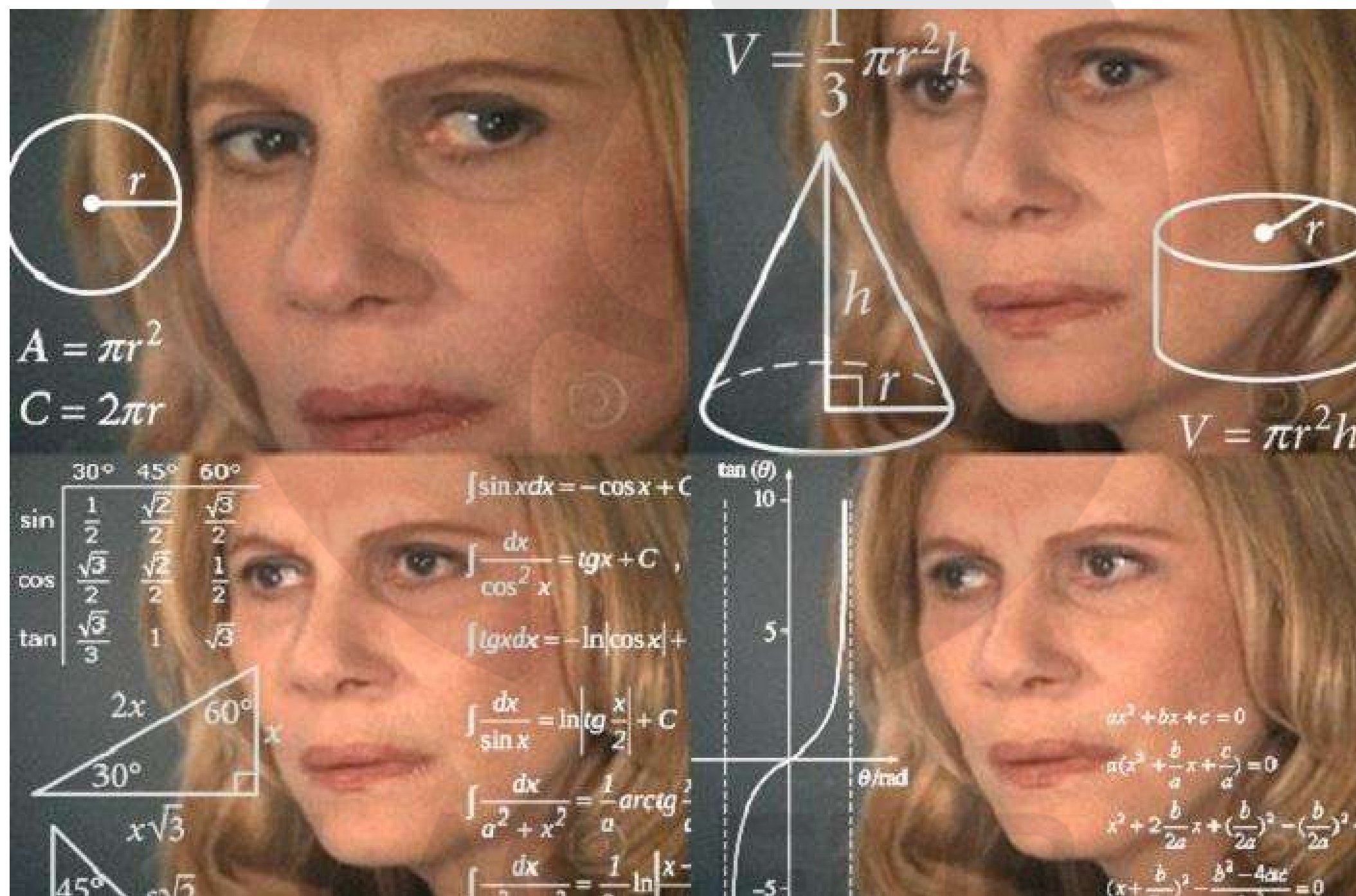
2. Sendo $\hat{y}_i$ o ponto na reta de regressão mais próximo à y_i , um determinado ponto no gráfico, temos que a distância vertical do ponto à reta (RESÍDUO) é: $|y_i - \hat{y}_i|$, onde $\hat{y}_i = \beta_0 + \beta_1 x_i$



$$\begin{array}{c} |y_i - \hat{y}_i| \\ \downarrow \quad \downarrow \\ 60 - 50 \\ 10 \\ \text{(RESÍDUO)} \end{array}$$



Método dos Mínimos Quadrados



Método dos Mínimos Quadrados

$$|y_i - \hat{y}_i|, \text{ onde } \hat{y}_i = \beta_0 + \beta_1 x_i$$

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2$$

LAC-FMB



Método dos Mínimos Quadrados

$$|y_i - \hat{y}_i|, \text{ onde } \hat{y}_i = \beta_0 + \beta_1 x_i$$

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2$$

LAC-FMB



Método dos Mínimos Quadrados

$$|y_i - \hat{y}_i|, \text{ onde } \hat{y}_i = \beta_0 + \beta_1 x_i$$

(RESÍDUO)

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2$$

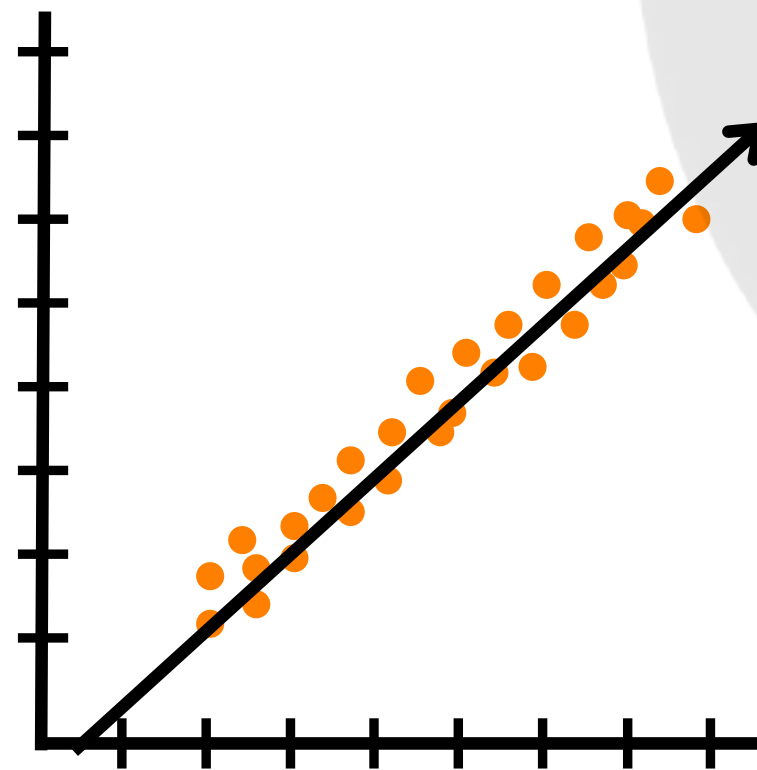
LAC-FMB



Método dos Mínimos Quadrados

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2$$

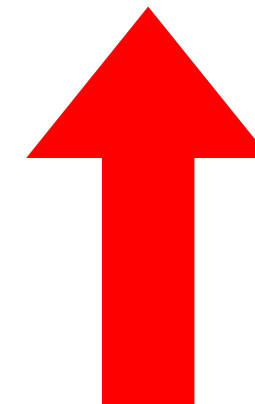
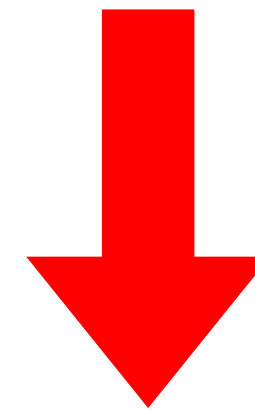
Glicemia



Q. de refeições
diárias

Soma do quadrado
dos resíduos

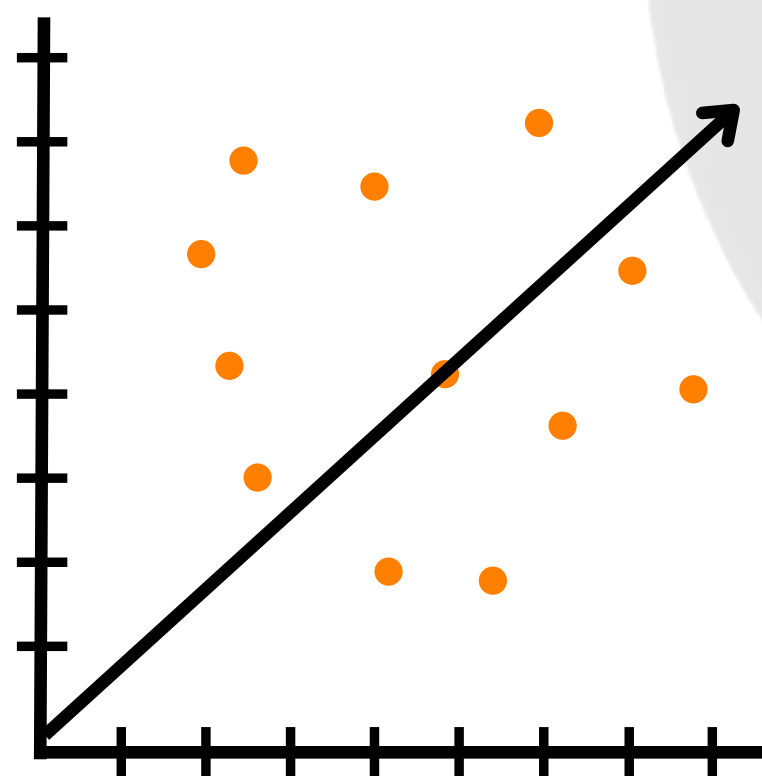
Precisão da
regressão



Método dos Mínimos Quadrados

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2$$

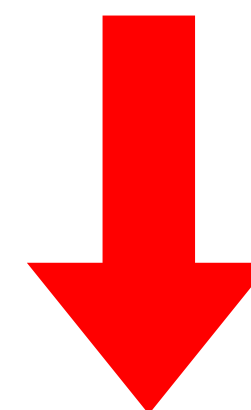
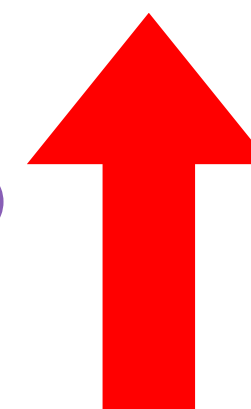
Glicemia



Q. de refeições
diárias

Soma do quadrado
dos resíduos

Precisão da
regressão



Certo, Sávio. Encontramos a
reta de melhor aderência, mas o
que fazer agora?





Certo, Sávio. Encontramos a
reta de melhor aderência, mas o
que fazer agora?

Avaliar quão bem ela adere aos dados!





**Analisar a qualidade de ajuste do
modelo!**



Como avaliar o melhor modelo?



LAC FMD



Como avaliar o melhor modelo?



Comparo com o modelo mais
básico: **a média.**



FIGHT

VS

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



FIGHT

Modelo
básico

VS

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



FIGHT

Modelo
básico

Média

VS

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



FIGHT

Modelo
básico

Média

VS

Melhor
modelo

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



FIGHT

Modelo
básico

Média

VS

Melhor
modelo

Linha de regressão

LAC-FMB

Liga Acadêmica de Ciências

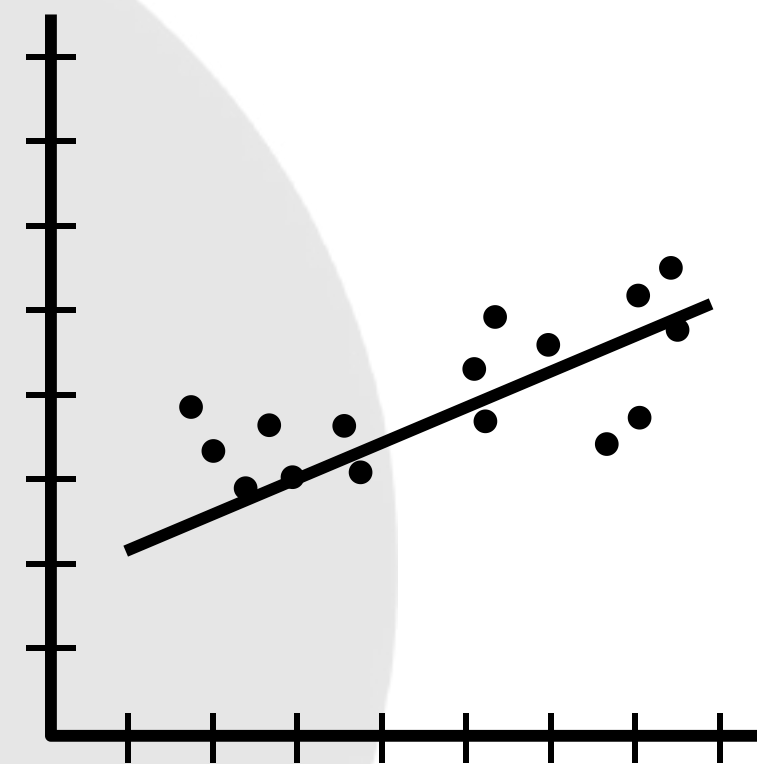
FIELD, A., 2009



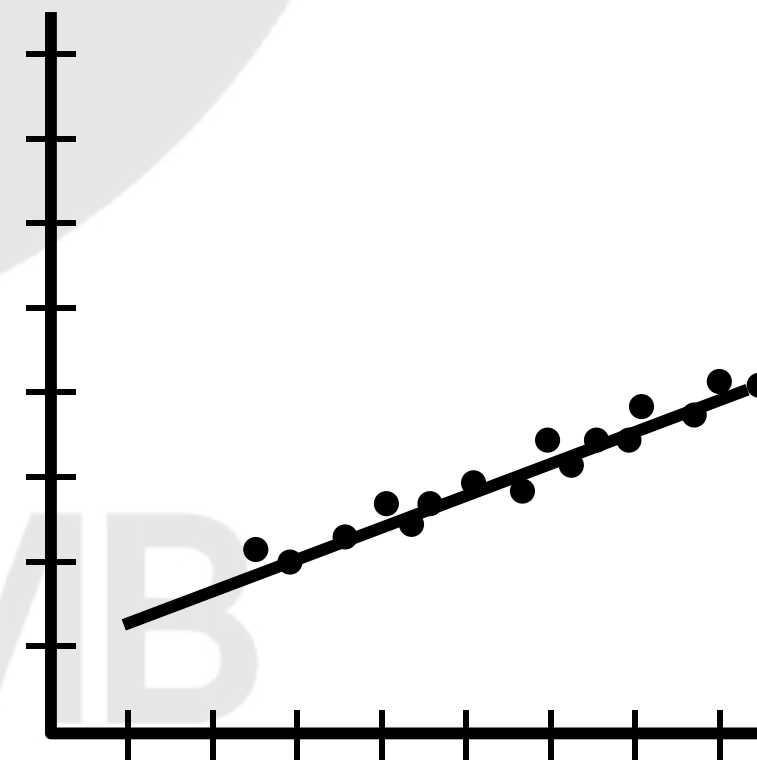
QUAL É O MELHOR MODELO?



1



2



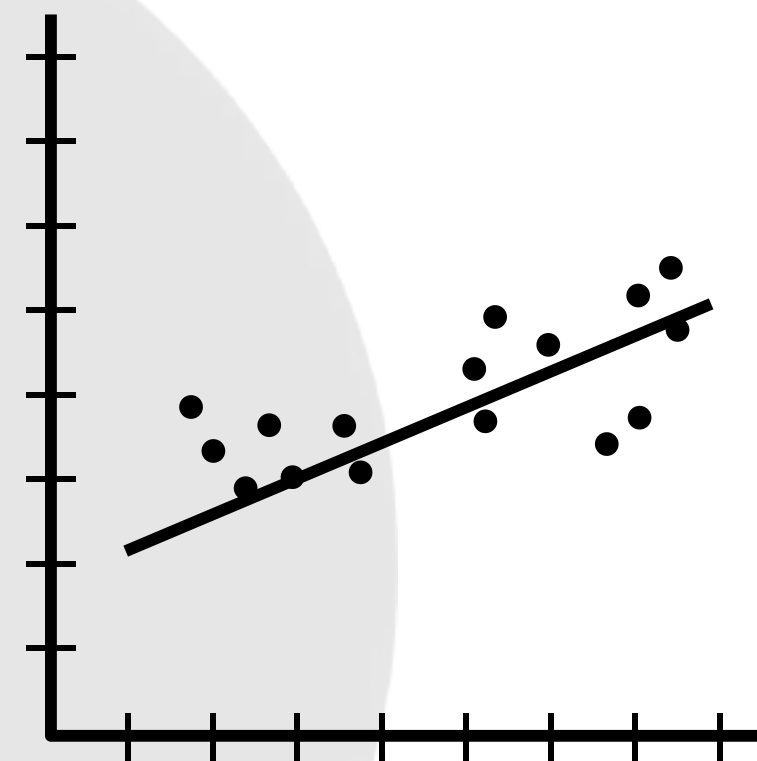
AC-FMB

Acadêmica de Ciências

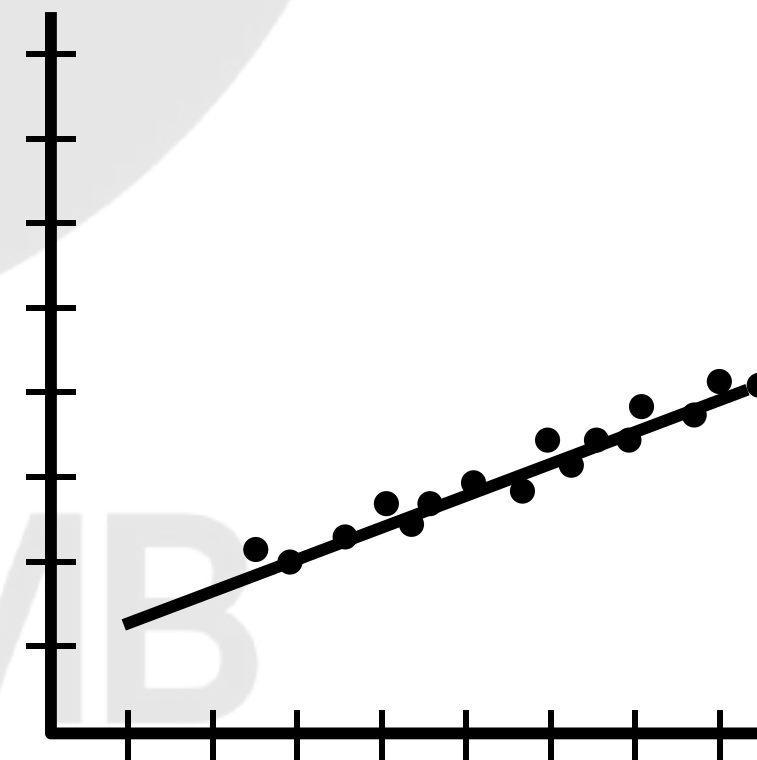
QUAL É O MELHOR MODELO?



1



2

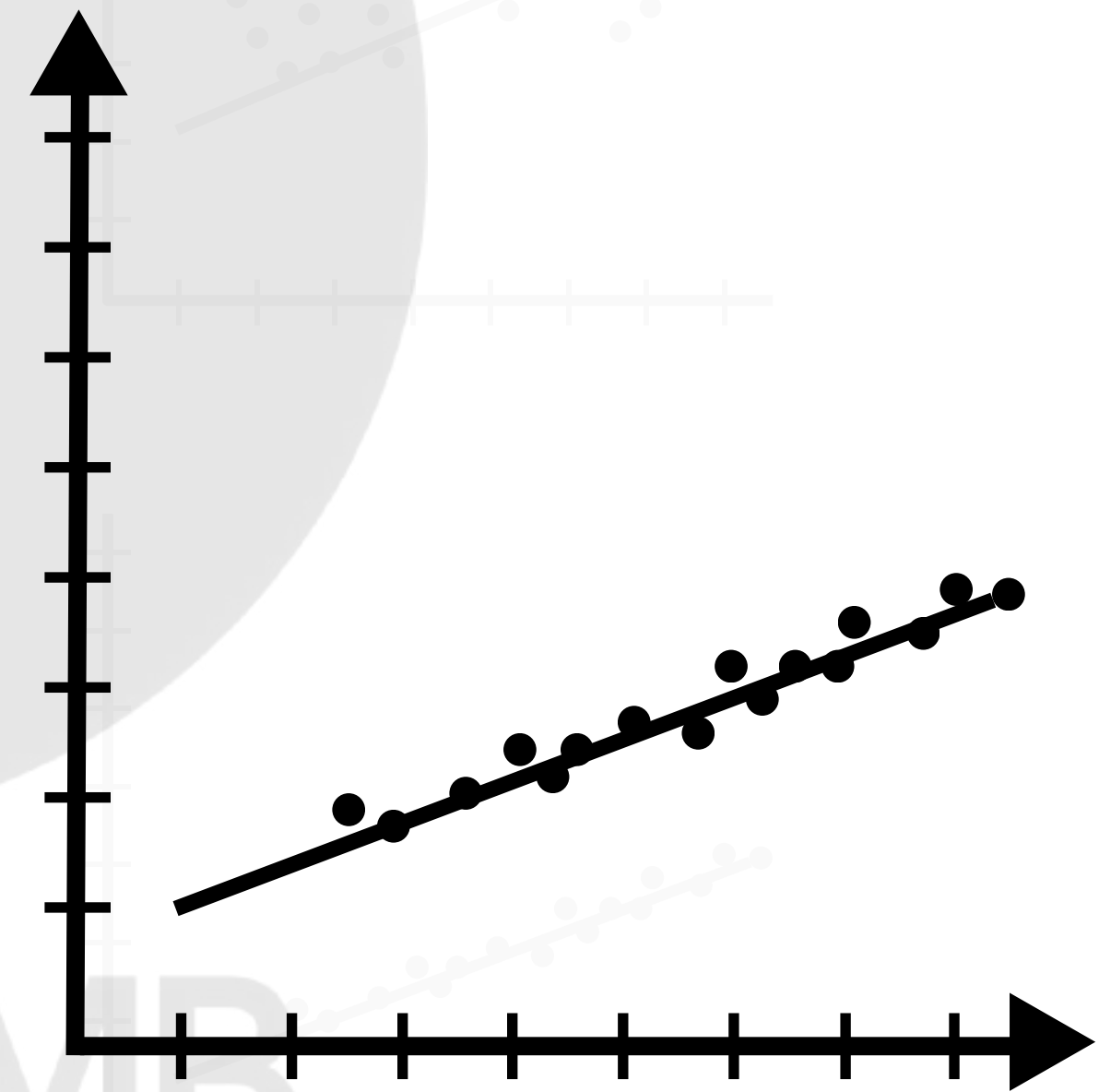


AC-FMB

Acadêmica de Ciências

Melhor modelo

“Adere a maior
quantidade de
dados.”



FIELD, A., 2009

LAC-FMIB

Liga Acadêmica de Ciências

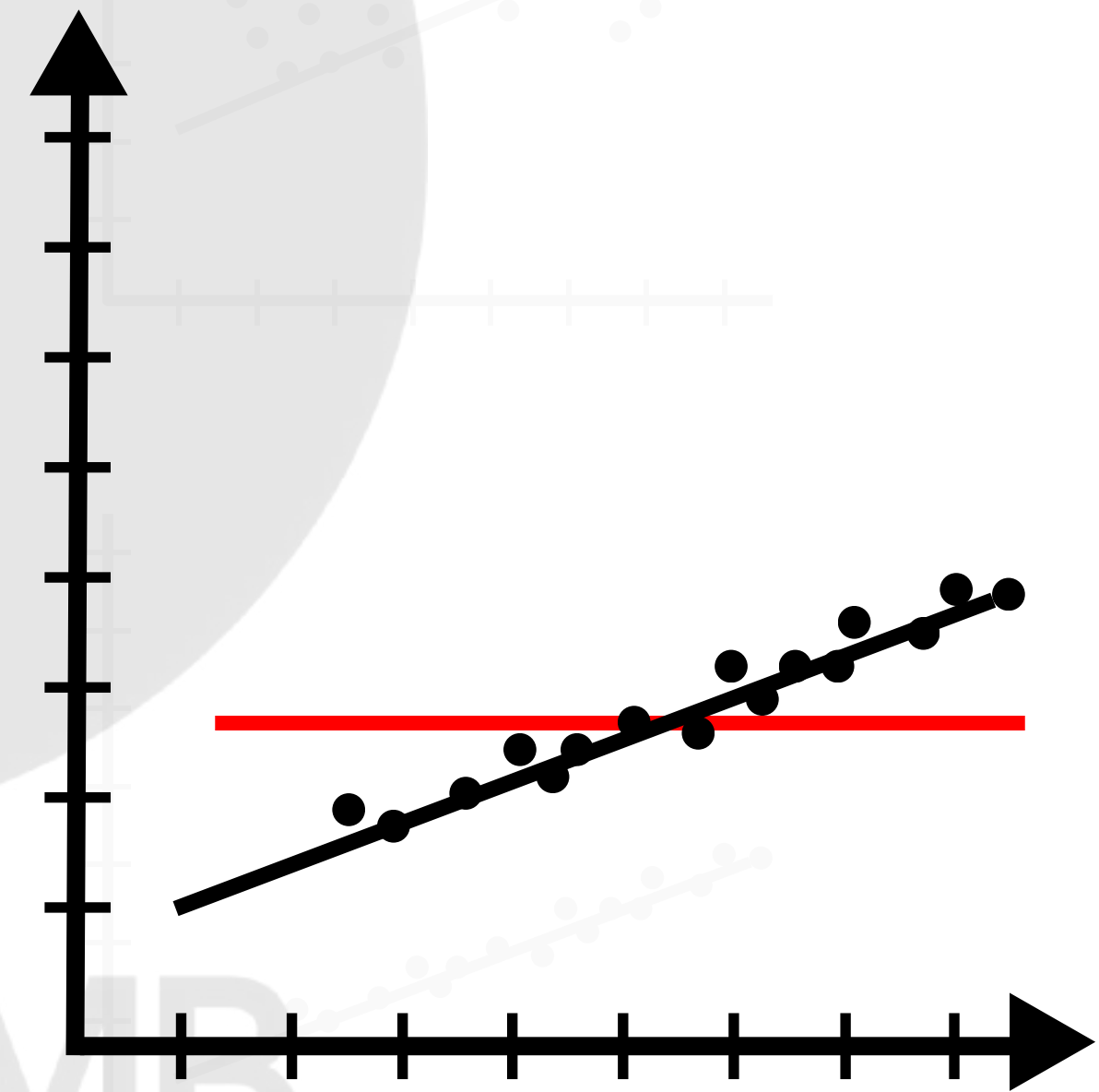
Estou confuso!!!

Mas por que a média **não** é
o melhor modelo?



Melhor modelo

“Adere a maior
quantidade de
dados.”



FIELD, A., 2009

LAC-FMIB

Liga Acadêmica de Ciências

Relembrando **MÉDIA** em **8** slides

DESVIOS

LAC-FMB

Liga Acadêmica de Ciências



Média

1

A média é um valor hipotético que pode ser calculado para qualquer conjunto de dados;

DESVIOS

LAC-FMB

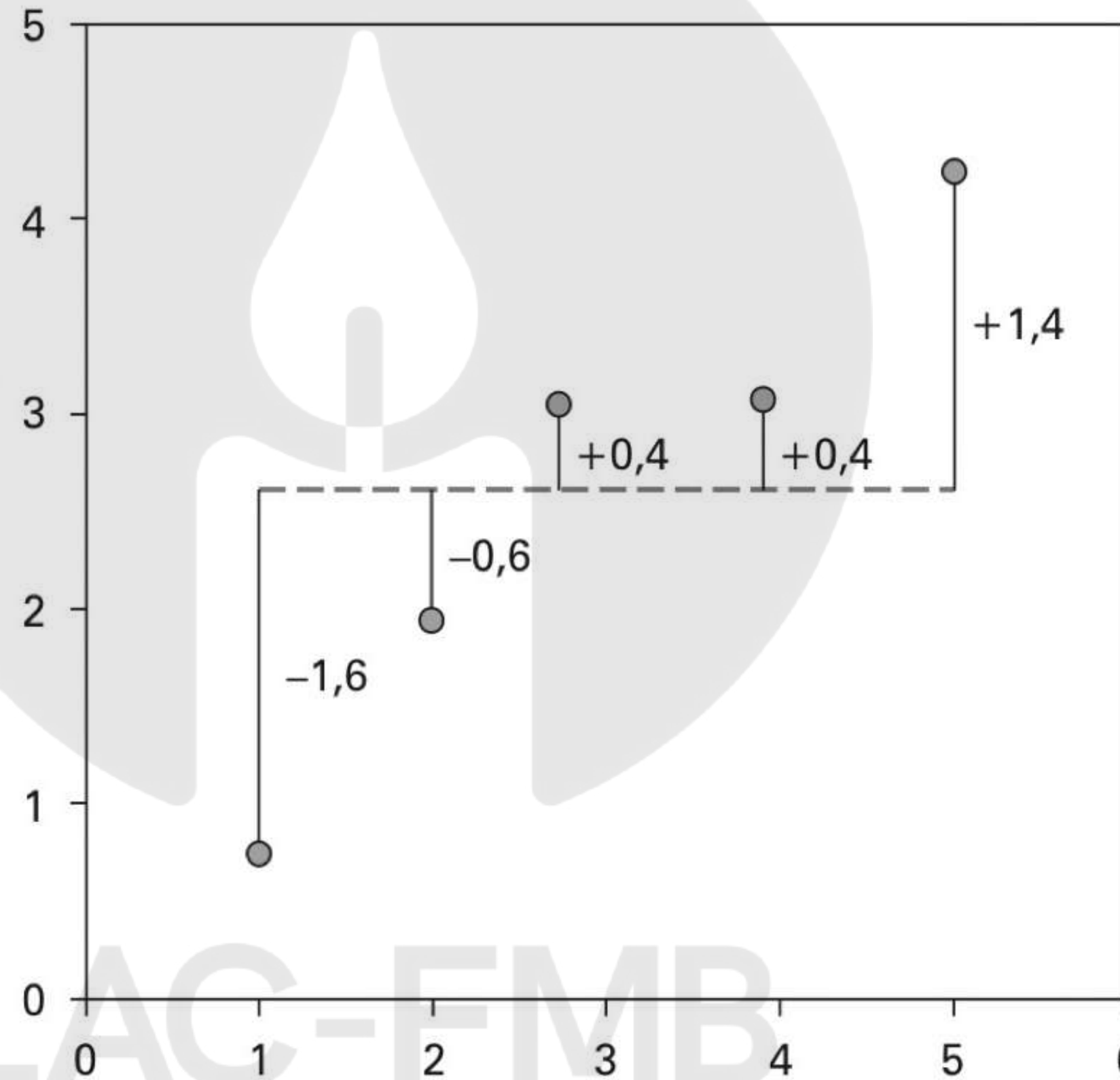
FIELD, A., 2009



Liga Acadêmica de Ciências

Desvios

2

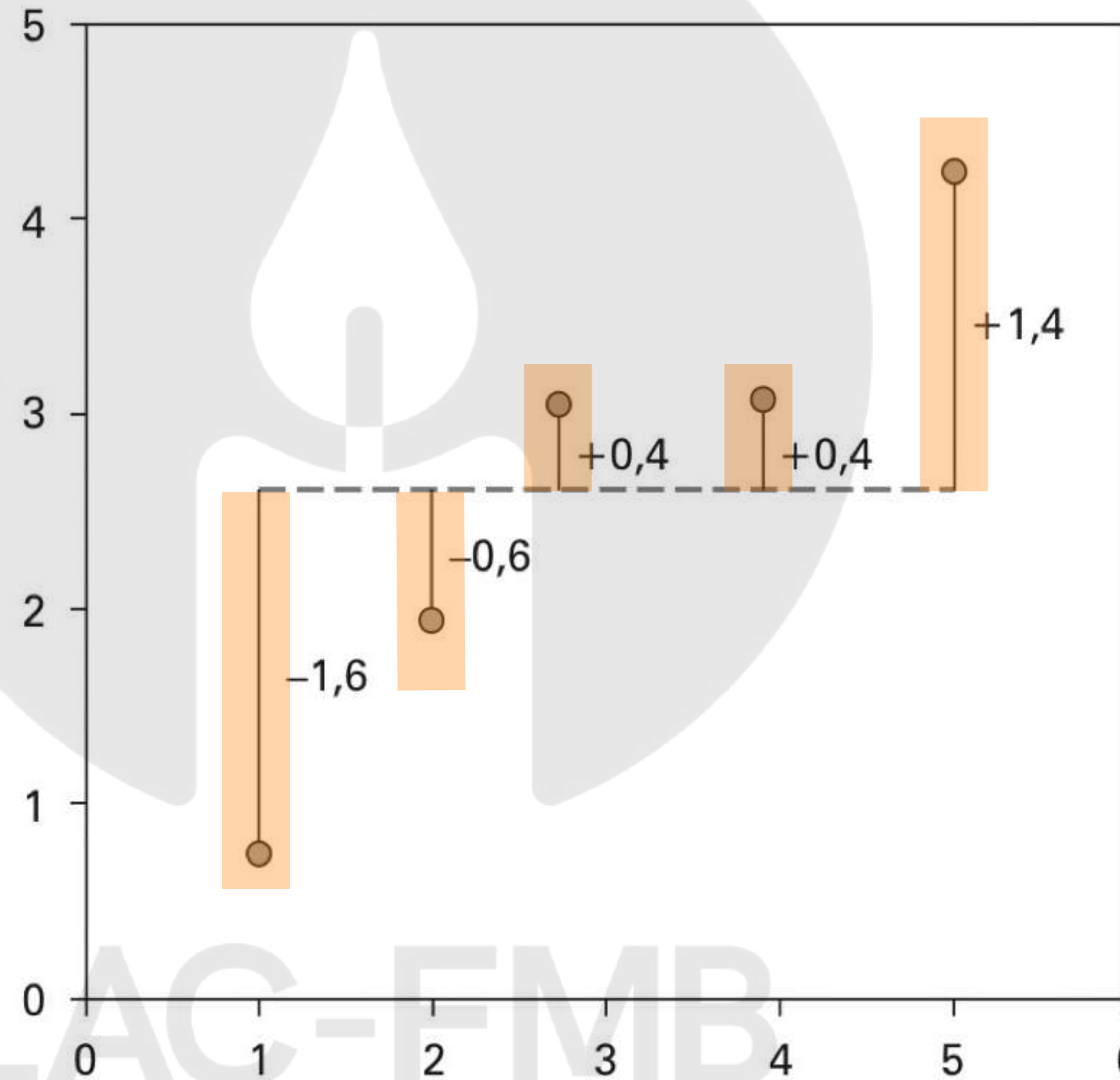


FIELD, A., 2009



Desvios

3



FIELD, A., 2009

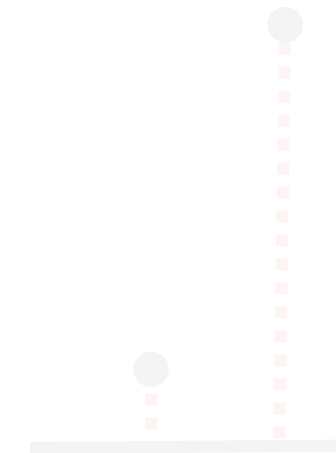
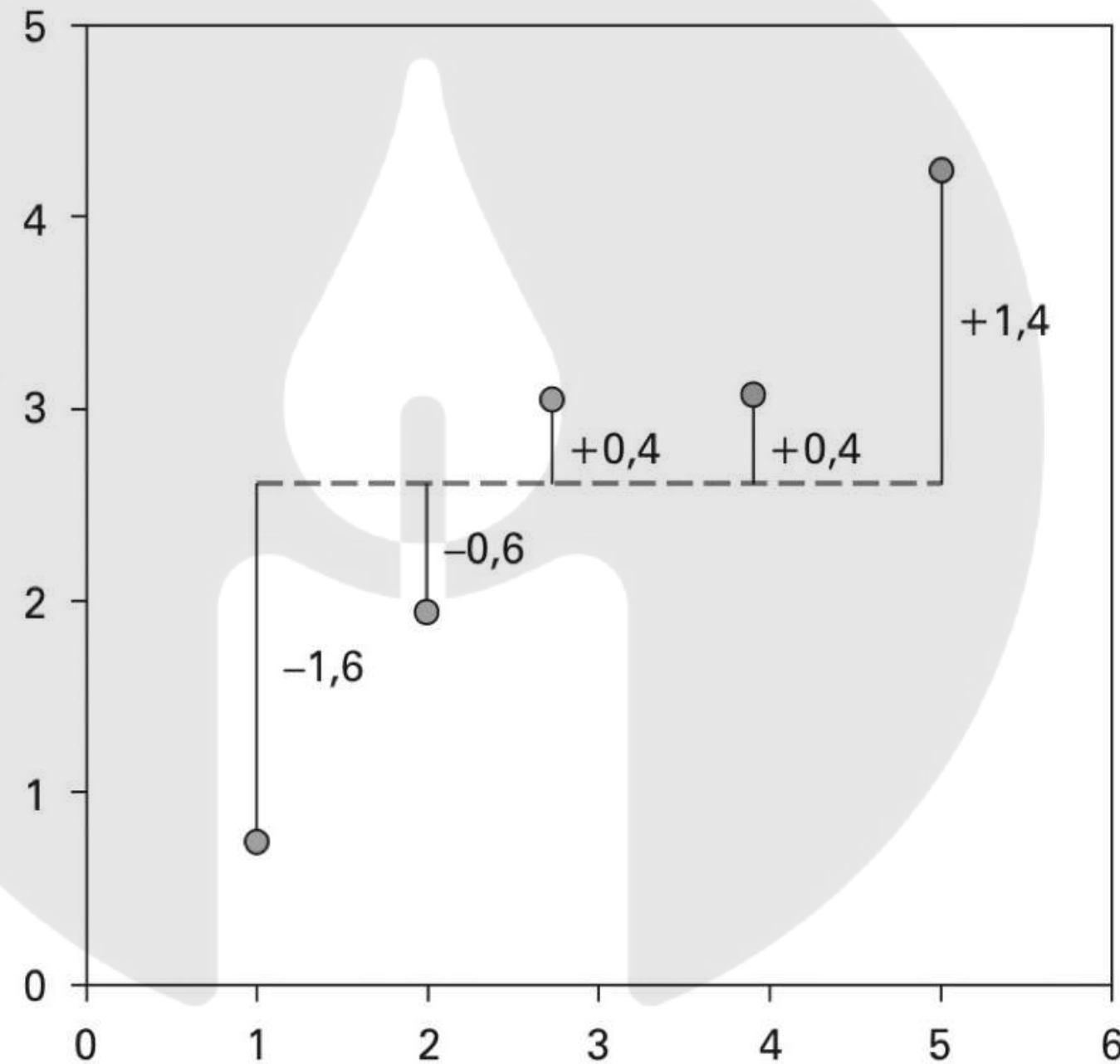


Como usar desvios para estimar
a **precisão do modelo?**

DESVIOS

Soma dos Desvios

5



DESVIOS

LAC-FMB

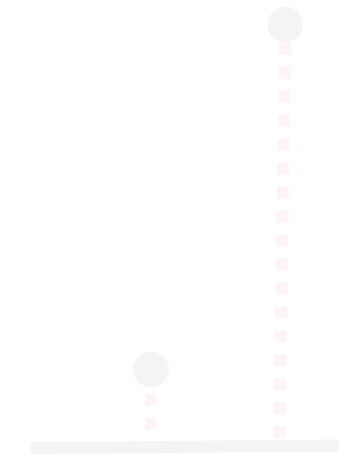
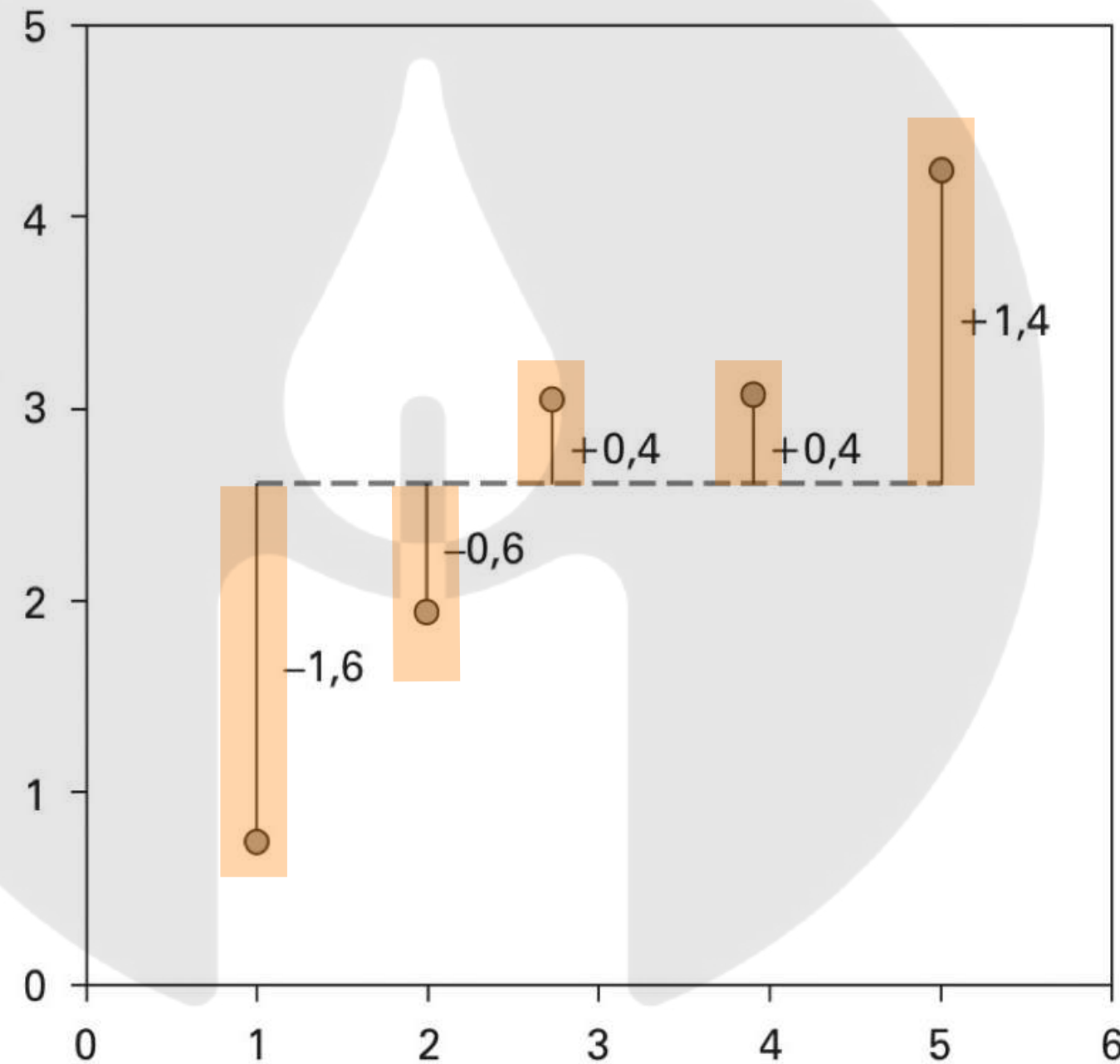
Liga Acadêmica de Ciências

FIELD, A., 2009



Soma dos Desvios

6



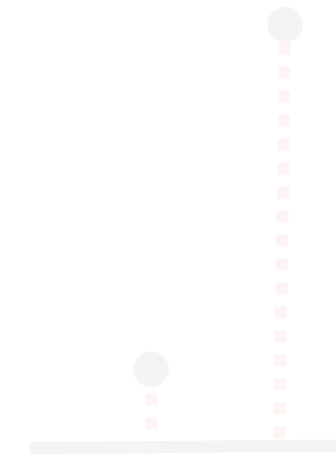
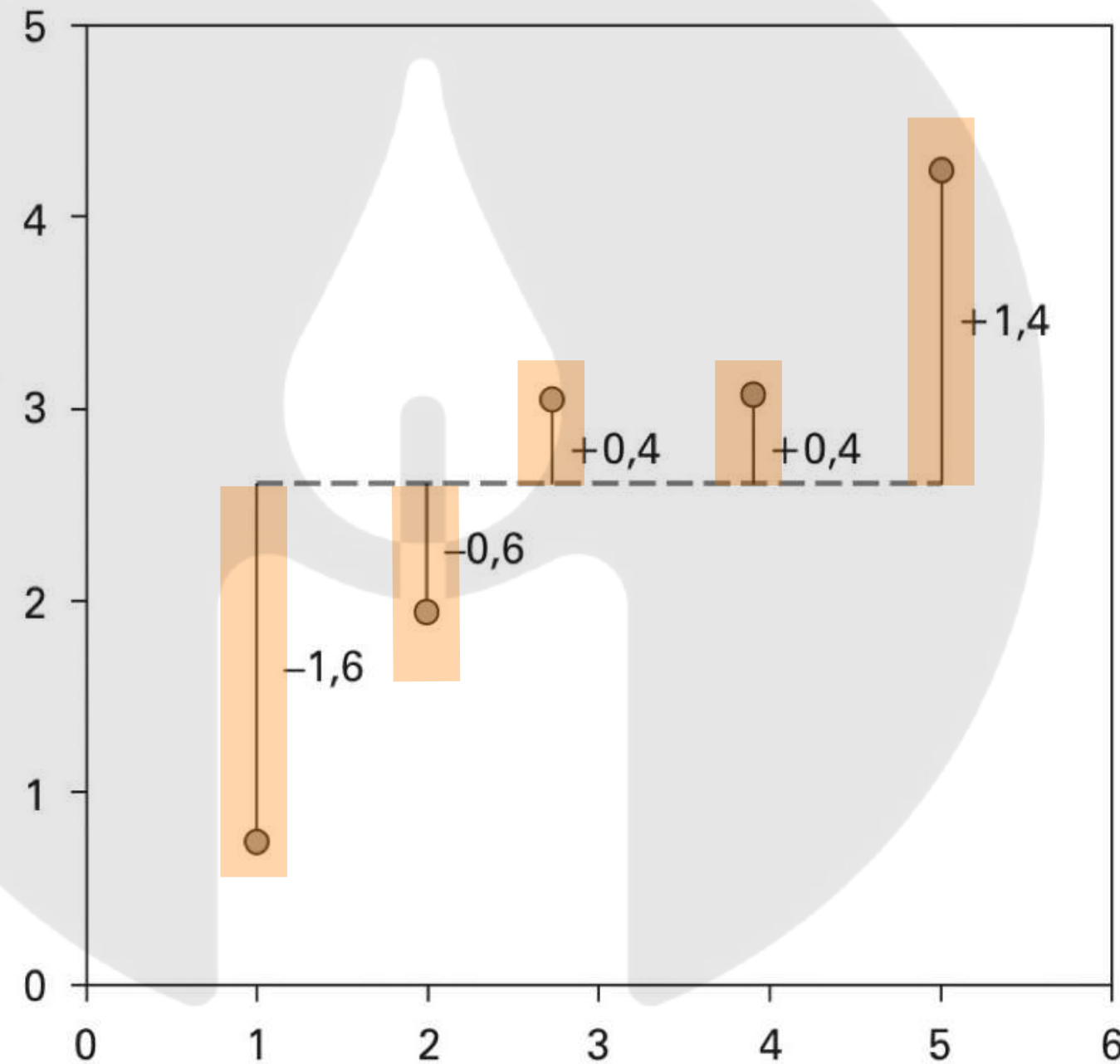
$$(-1,6) + (-0,6) + (0,4) + (0,4) + (1,4) =$$

FIELD, A., 2009



Soma dos Desvios

7



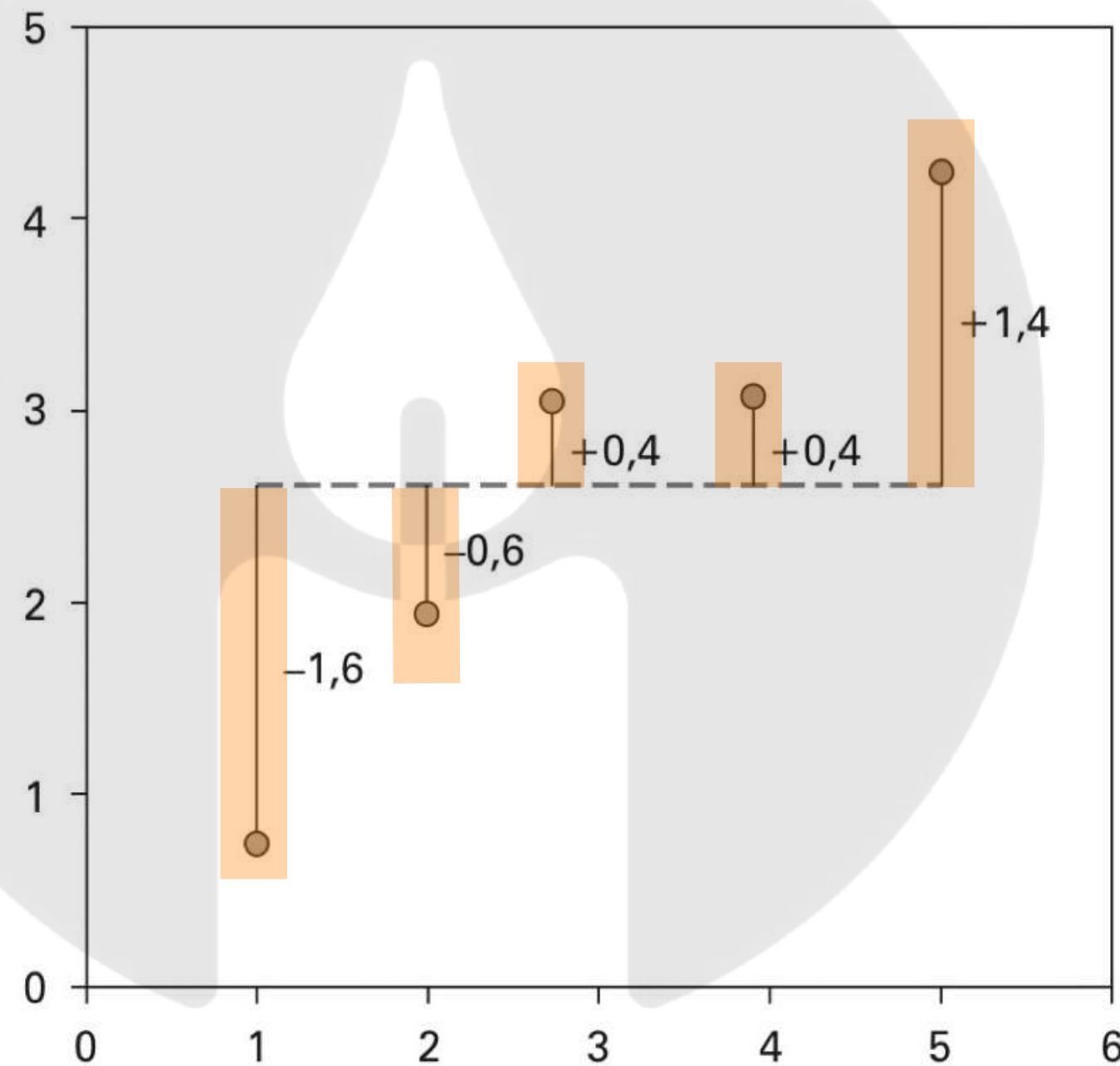
$$(-1,6) + (-0,6) + (0,4) + (0,4) + (1,4) = 0$$

FIELD, A., 2009



Soma dos Desvios ao Quadrado

8



$$(-1,6)^2 + (-0,6)^2 + (0,4)^2 + (0,4)^2 + (1,4)^2 = 5,20$$

FIELD, A., 2009





Voltando para **Regressão Linear Simples**

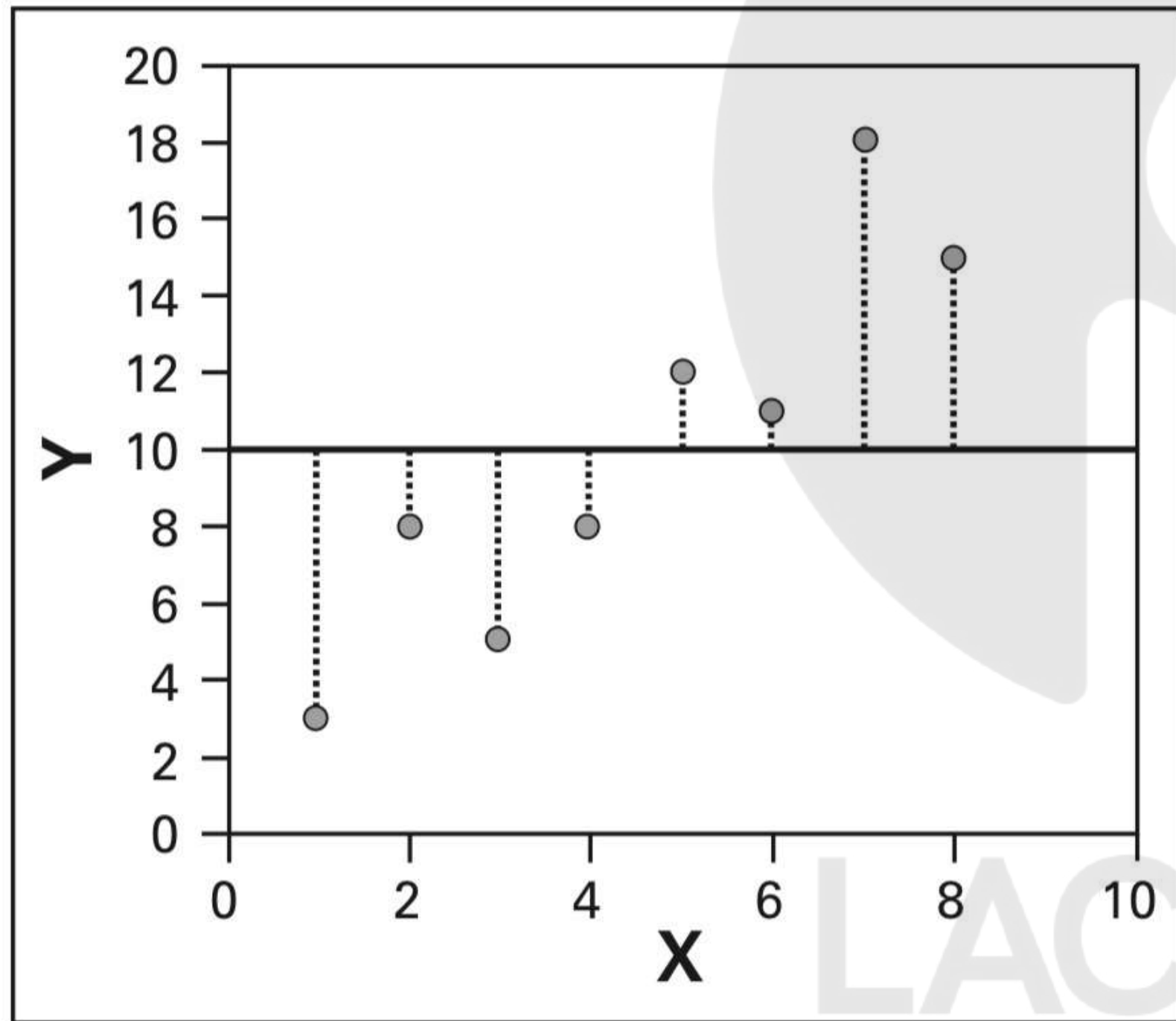
DESVIOS

LAC-FMB

Liga Acadêmica de Ciências



Soma Total dos Quadrados



SST

=

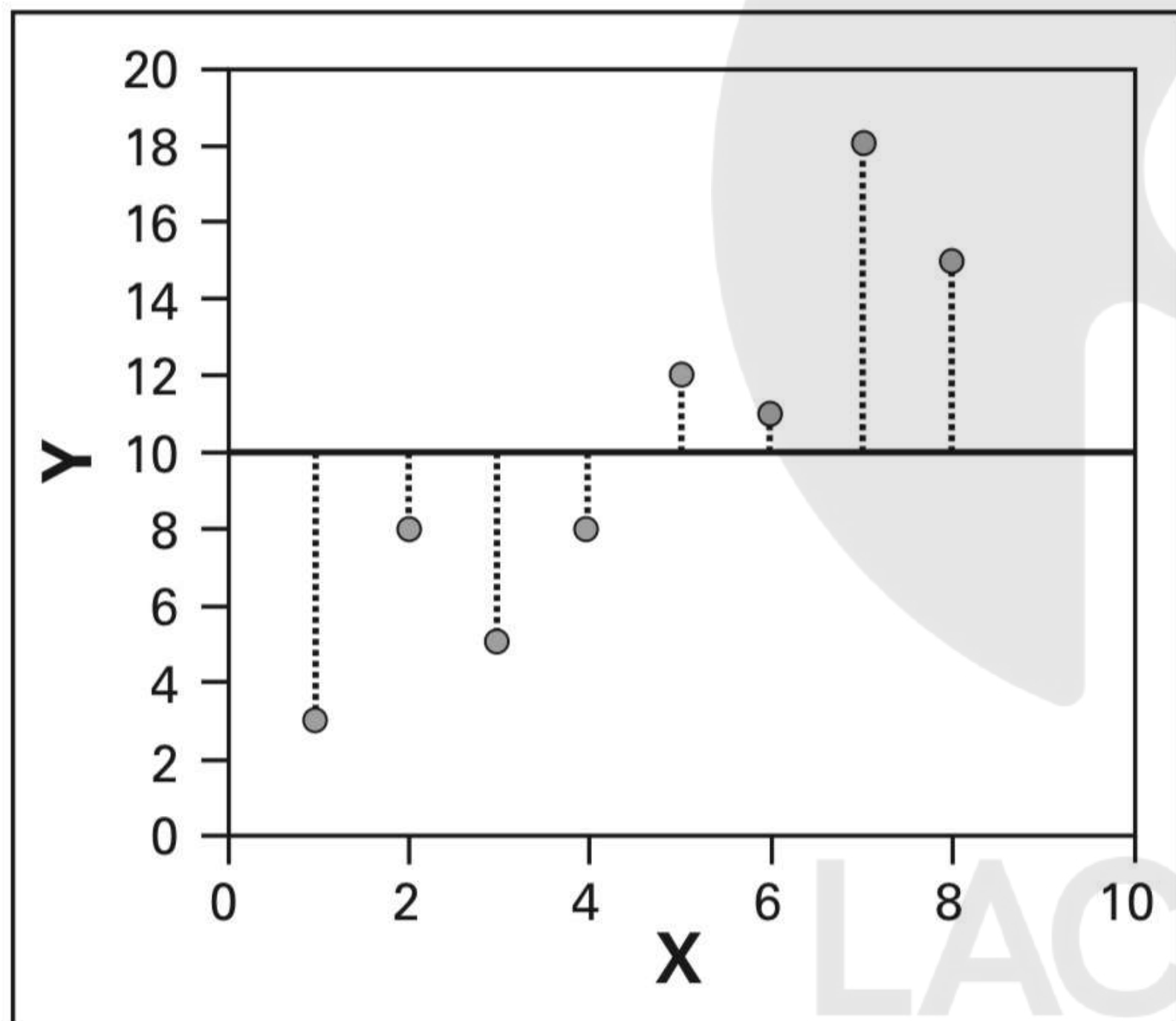
LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Soma Total dos Quadrados



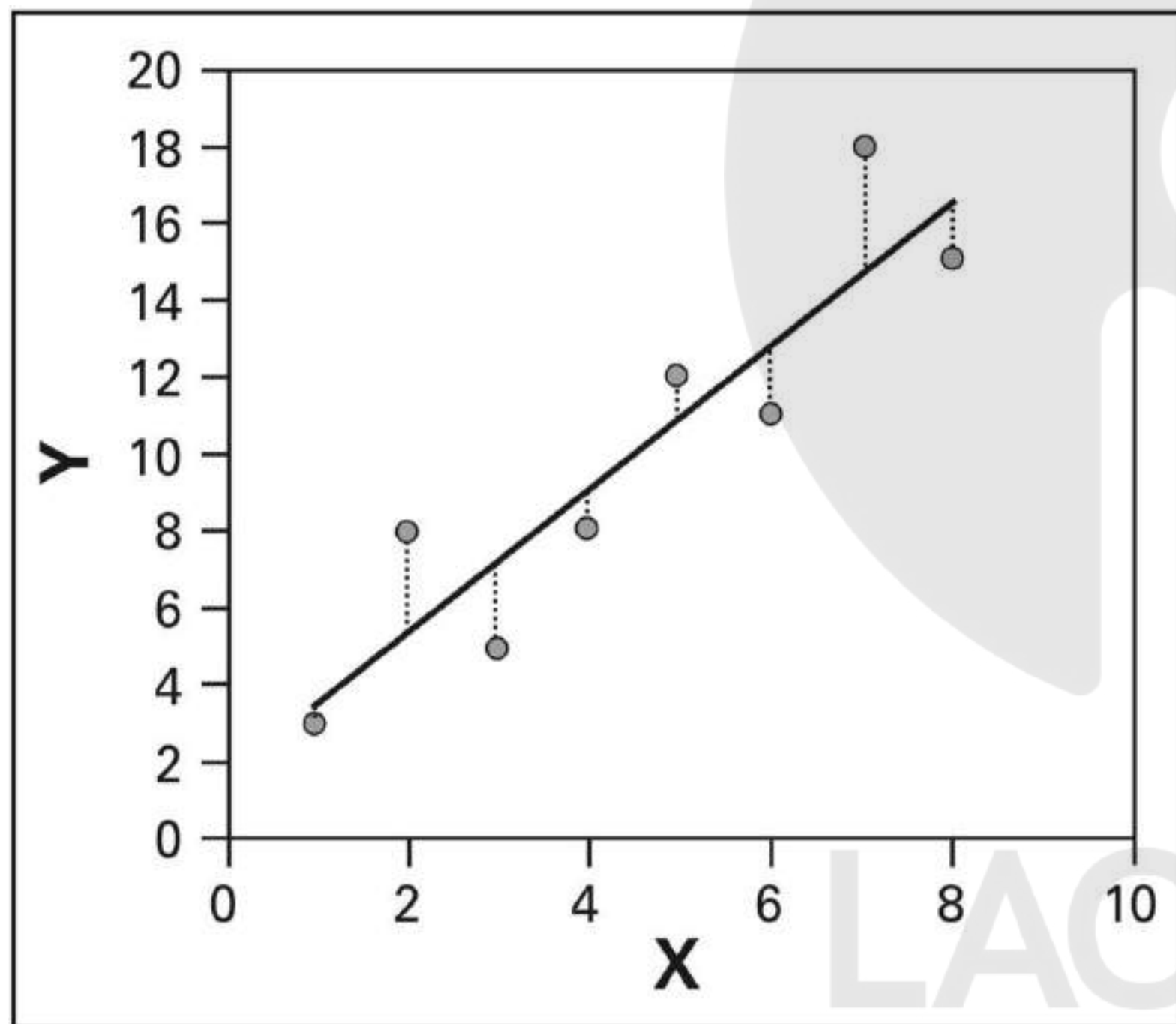
SST

=

Representa o grau de imprecisão quando o modelo mais básico é ajustado aos dados



Soma dos Quadrados Residual



SSR

=

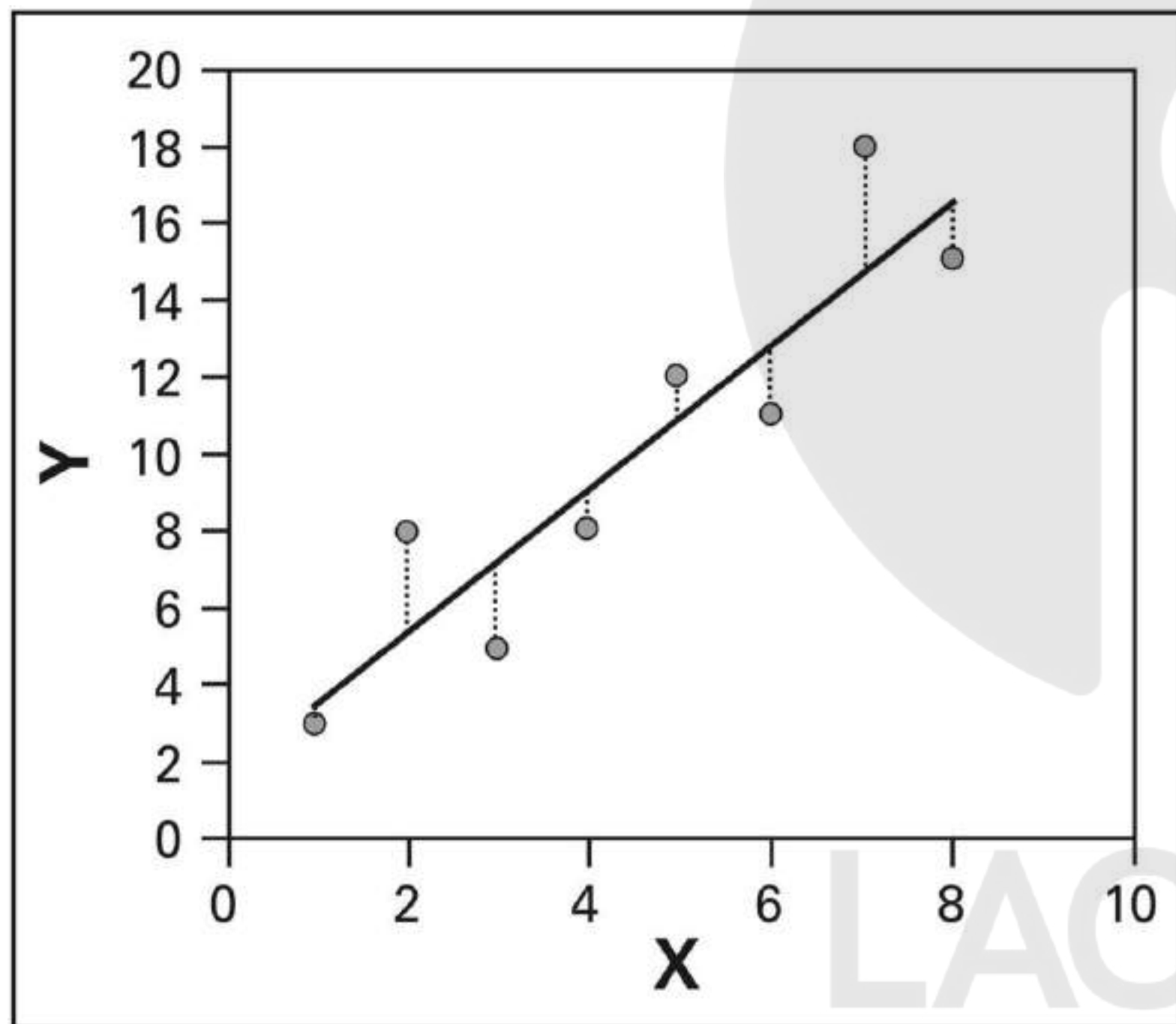
LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Soma dos Quadrados Residual



SSR

=

Representa o grau de imprecisão quando o melhor modelo é ajustado aos dados



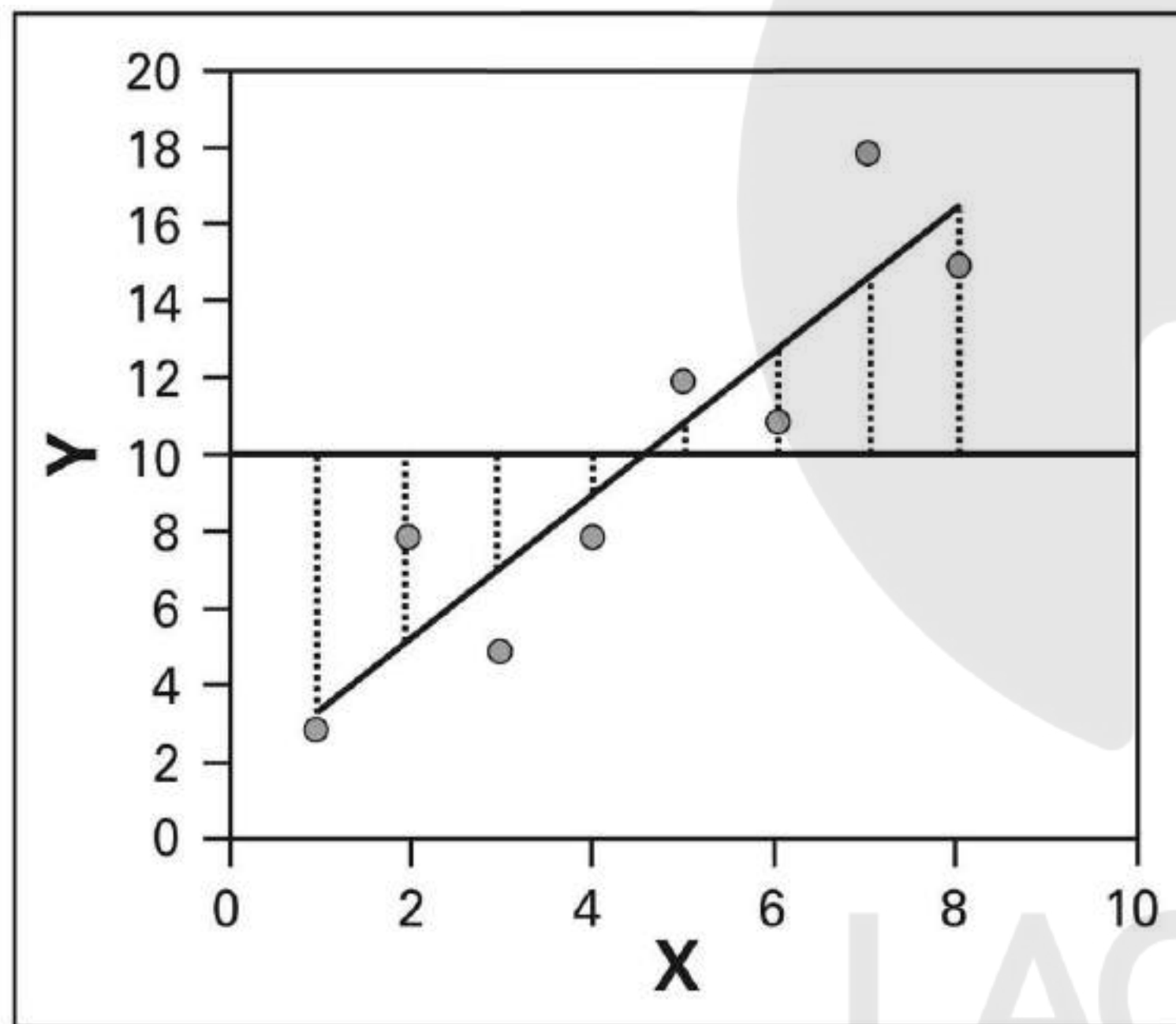
Como avaliar o melhor modelo?



Comparo com o modelo mais
básico: **a média.**

LAG FMD

Soma dos Quadrados do Modelo



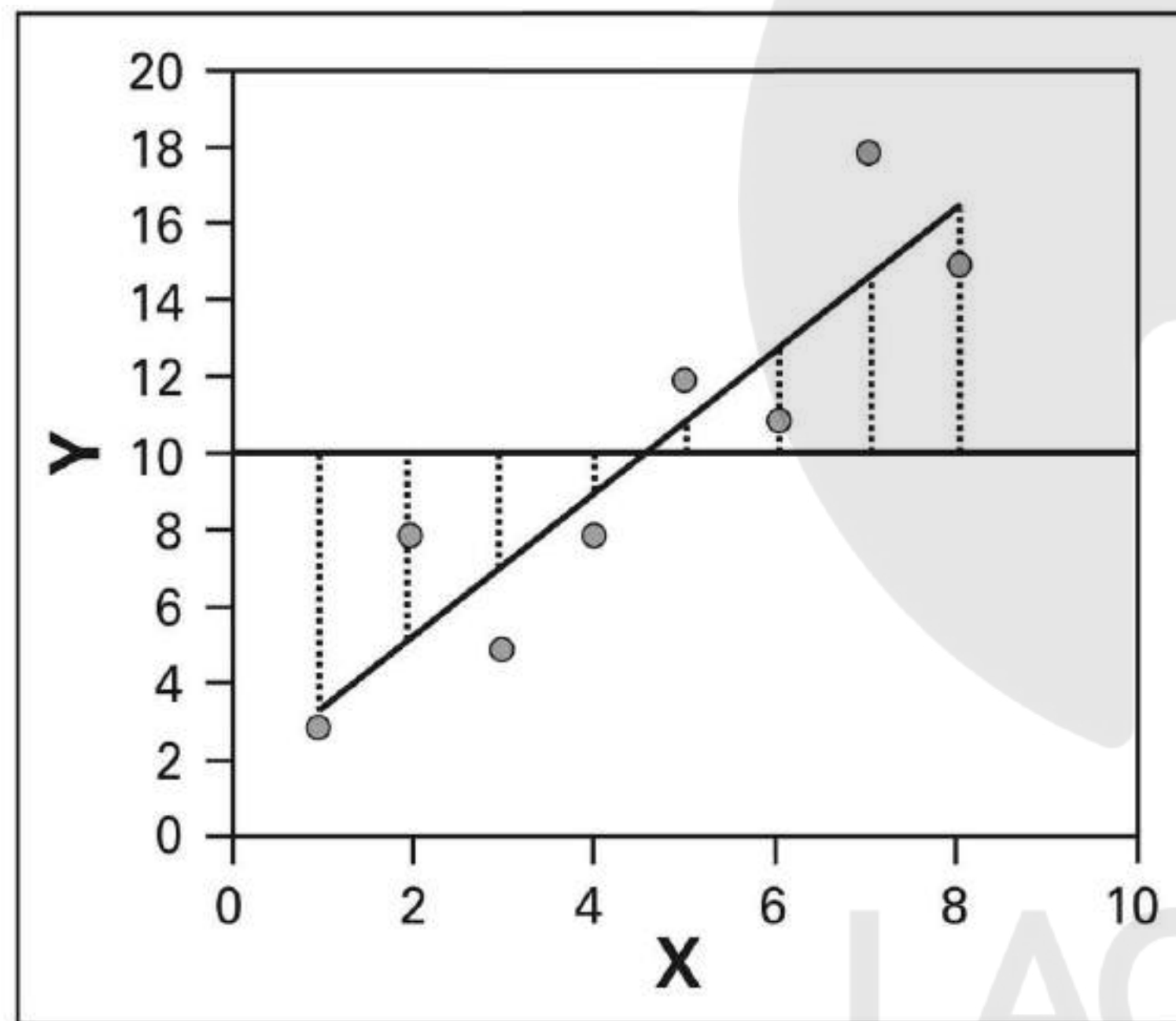
Soma Total dos
Quadrados

Soma dos Quadrados
Residual

$$SST - SSR = SSM$$
$$=$$



Soma dos Quadrados do Modelo



Soma Total dos
Quadrados

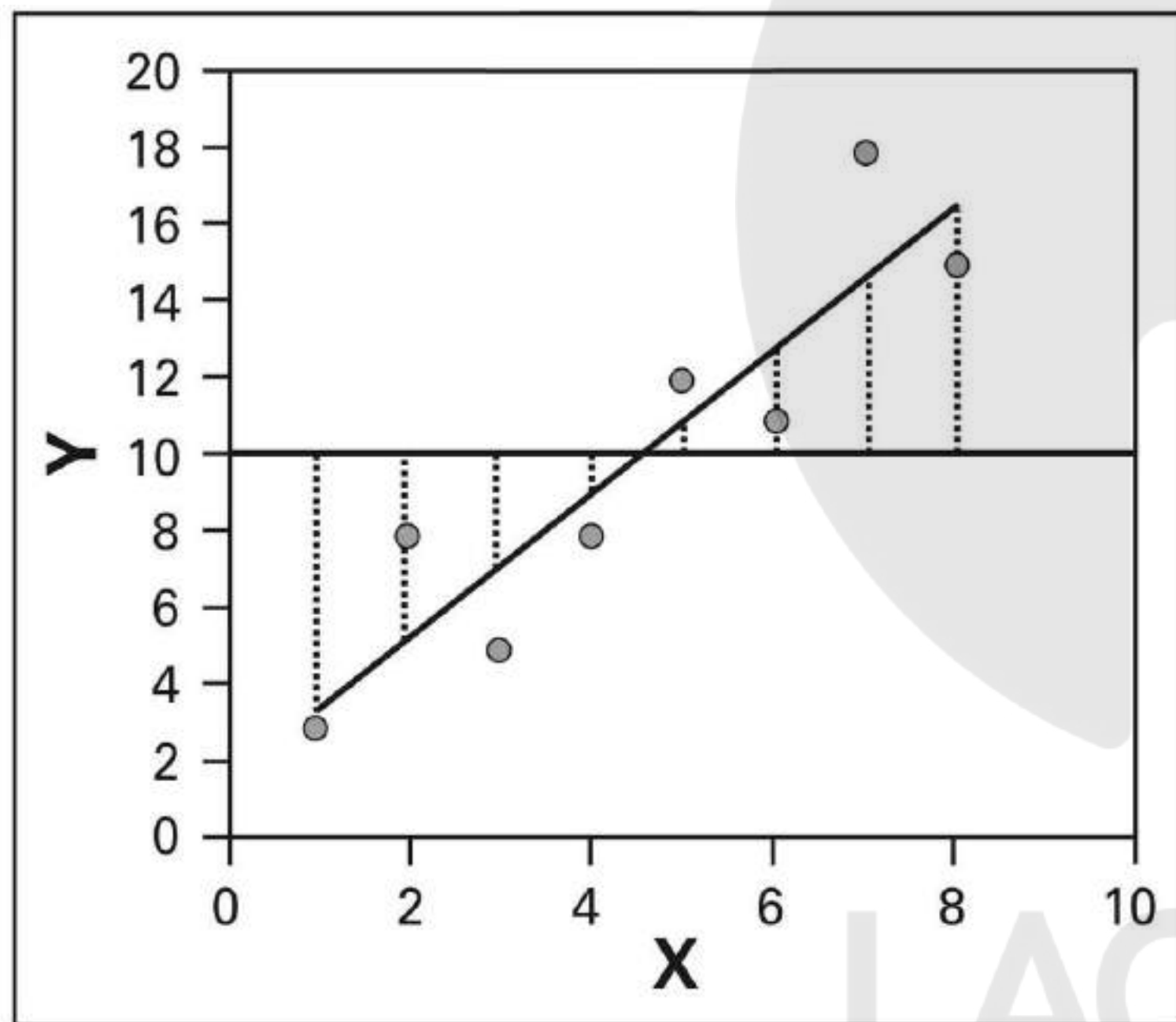
Soma dos Quadrados
Residual

$$SST - SSR = SSM$$
$$=$$

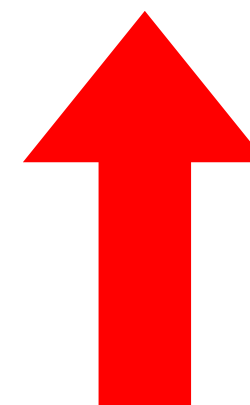
A **melhoria** na **previsão**
resultante da utilização da
linha de regressão em vez
da **linha média**



Soma dos Quadrados do Modelo



SSM

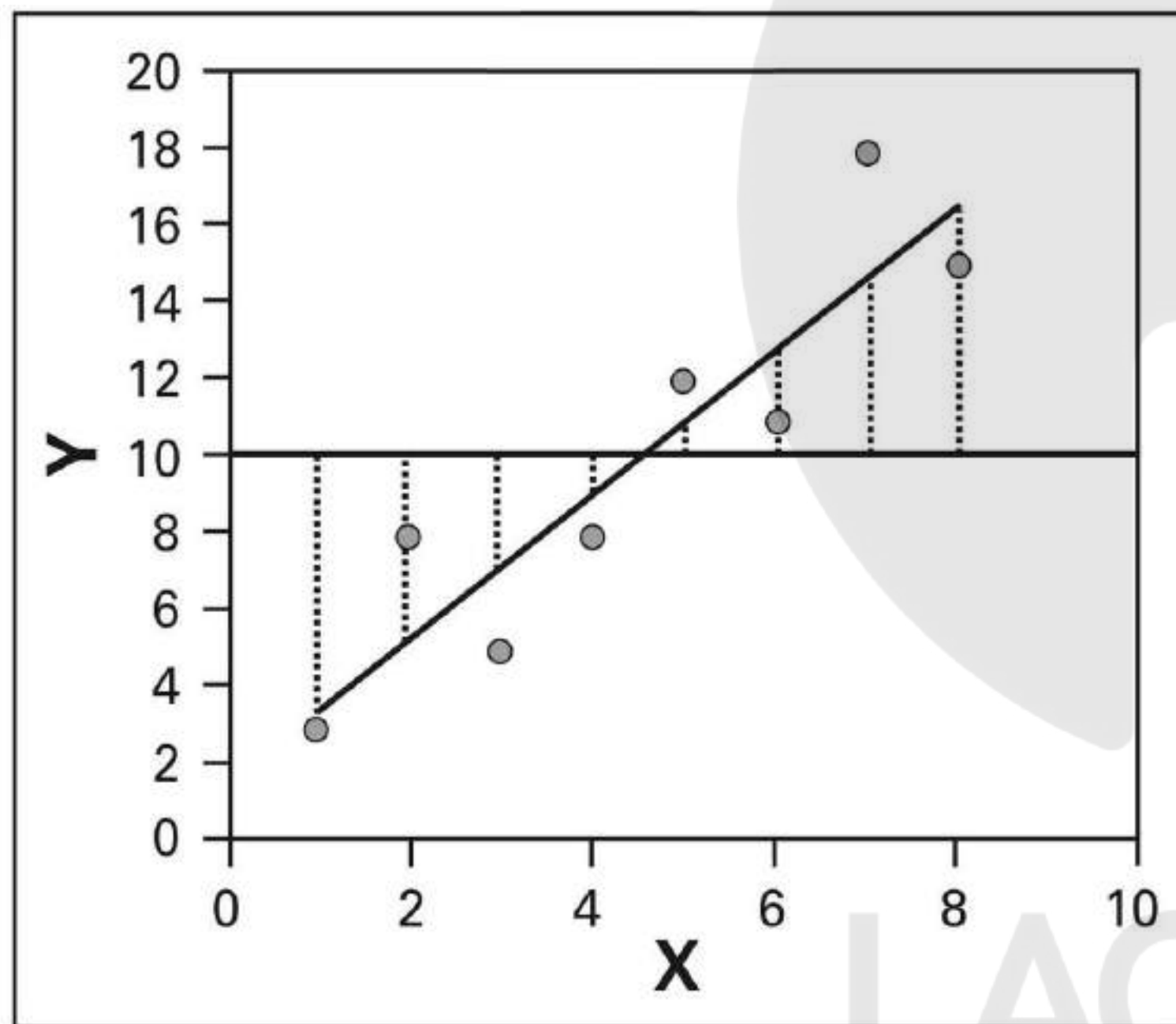


=

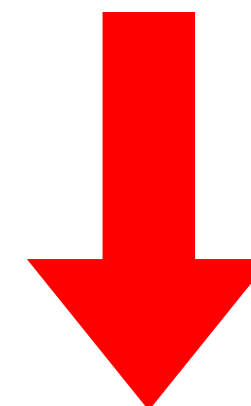
Modelo de regressão **BEM**
melhor do que a média



Soma dos Quadrados do Modelo



SSM



=

Modelo de regressão **UM**
POUCO melhor do que a
média



Soma dos Quadrados do Modelo



??????

SSM



=

Modelo de regressão **UM**

POUCO melhor do que a

média

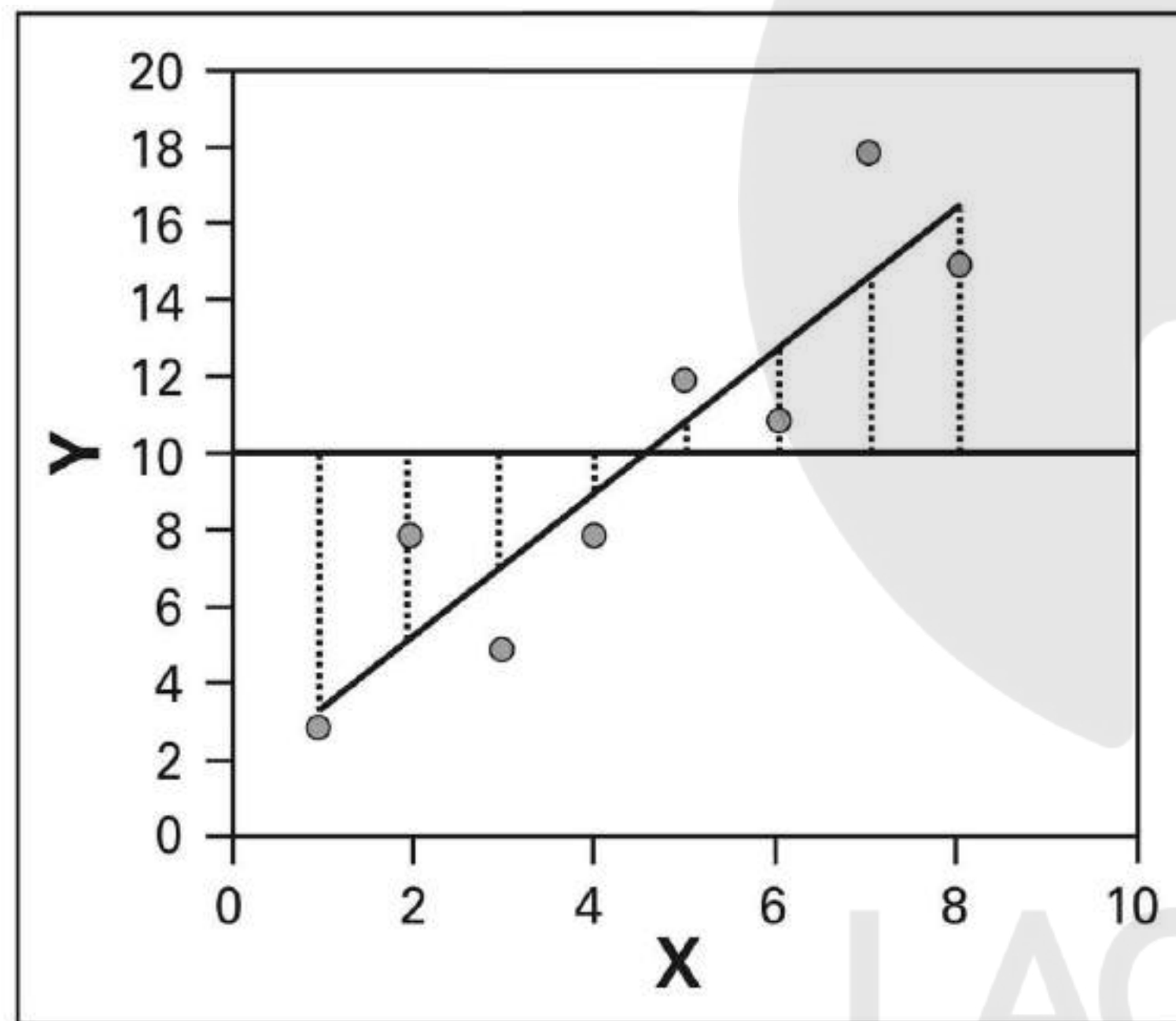
LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Soma dos Quadrados do Modelo

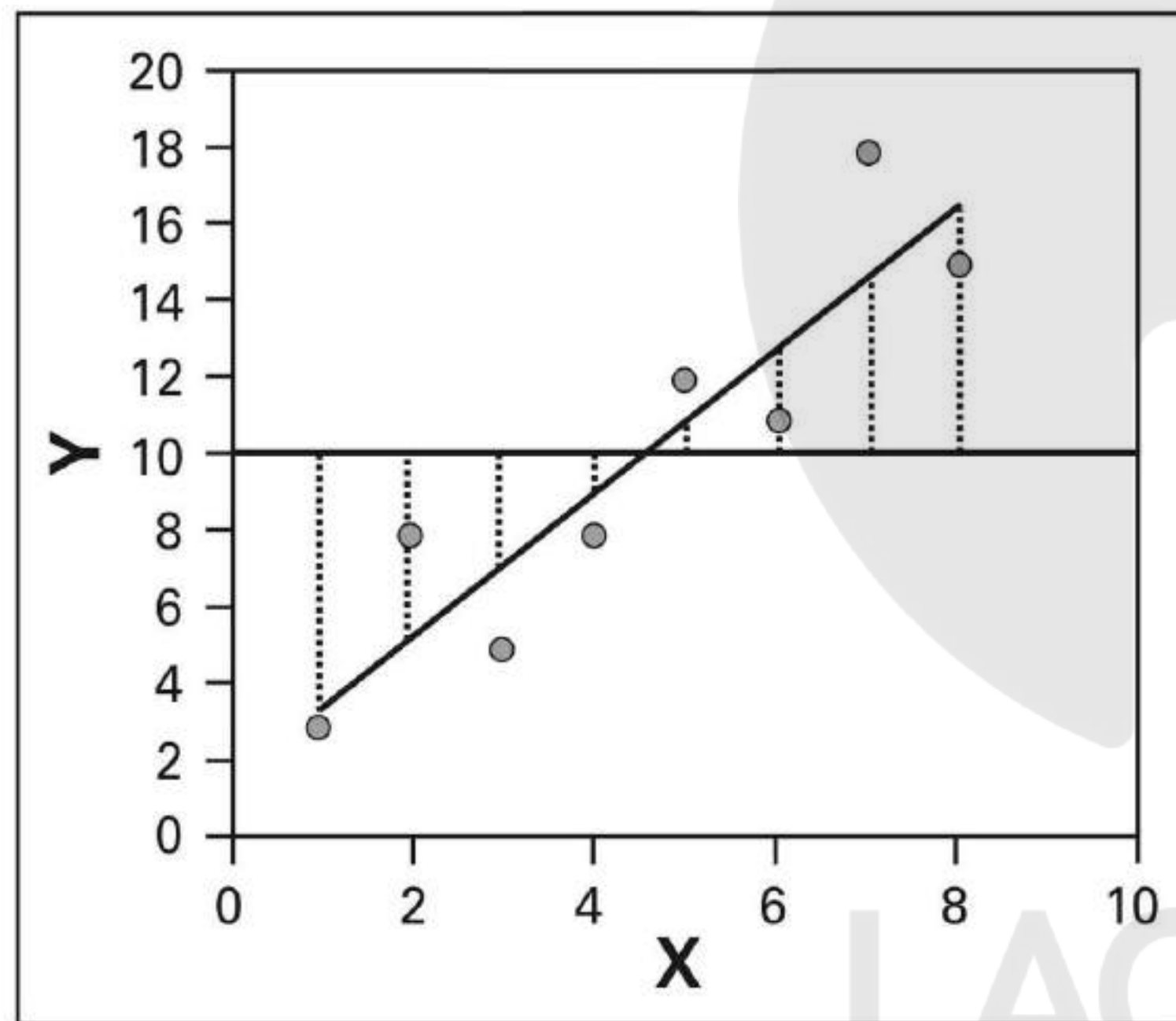


$$SST - SSR = SSM$$

=



Soma dos Quadrados do Modelo



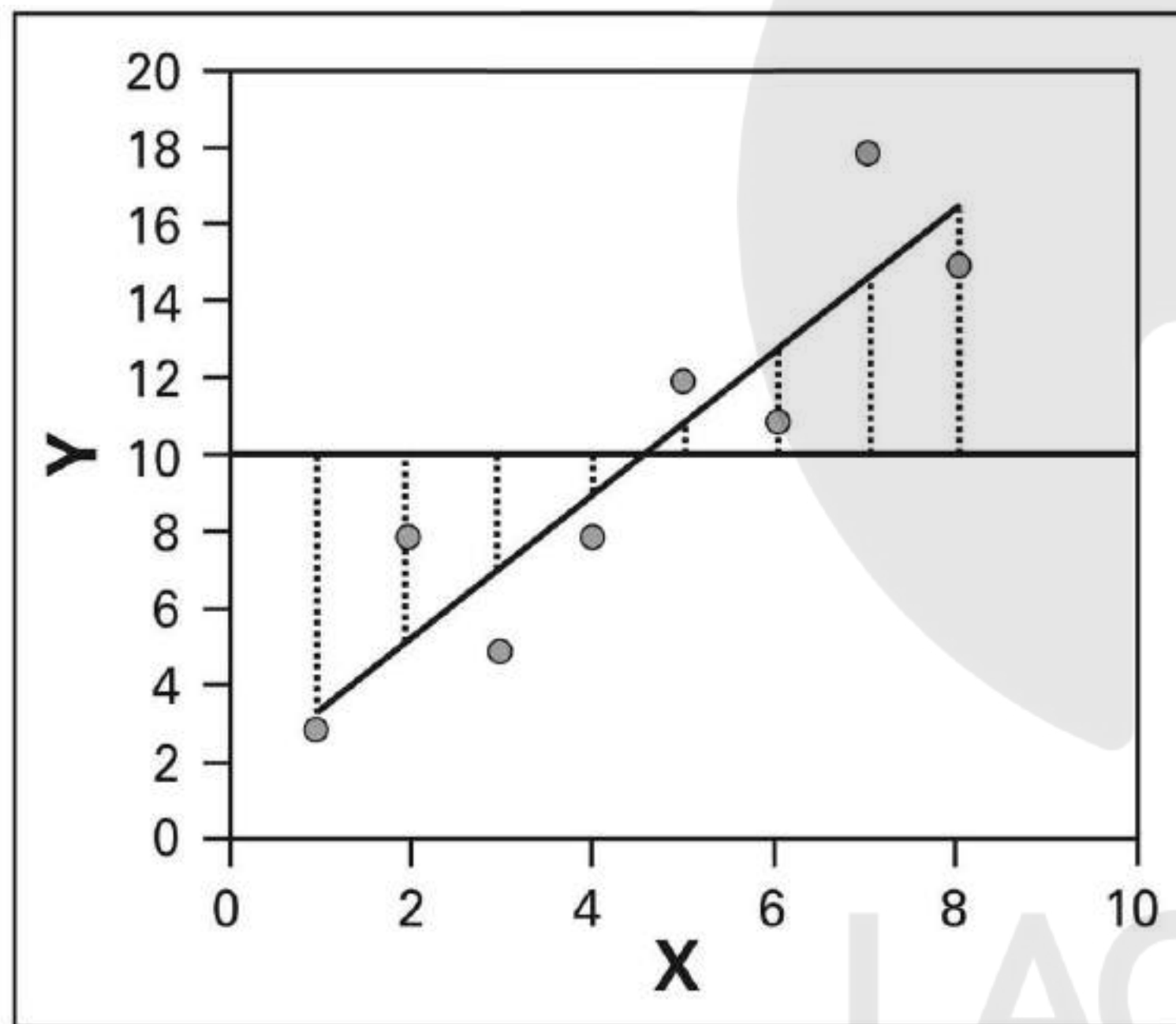
$$SST - SSR = SSM$$

$$100 - 95 = 5$$

=



Soma dos Quadrados do Modelo



$$SST - SSR = SSM$$

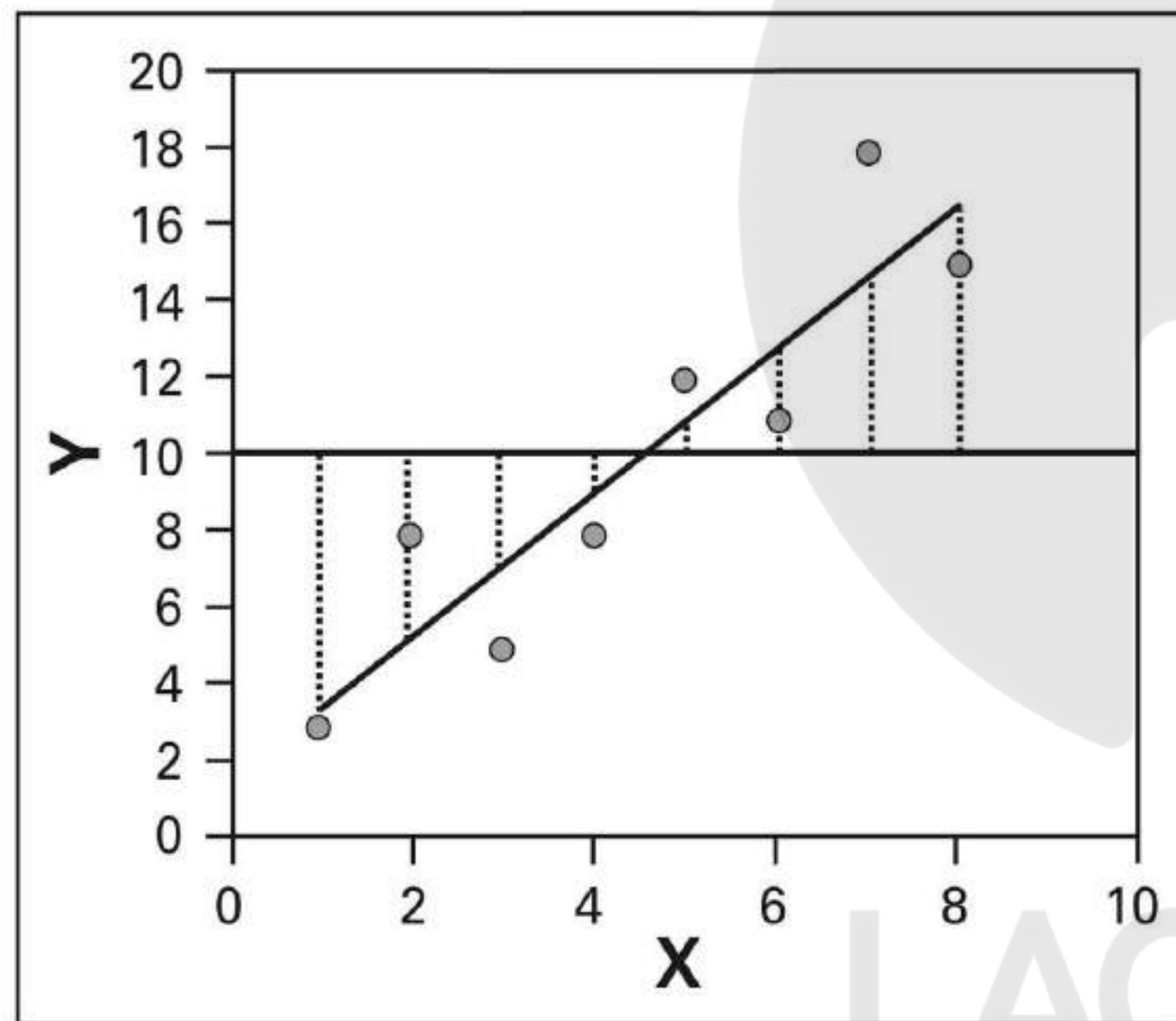
$$100 - 95 = 5$$

$$100 - 10 = 90$$

=



Soma dos Quadrados do Modelo



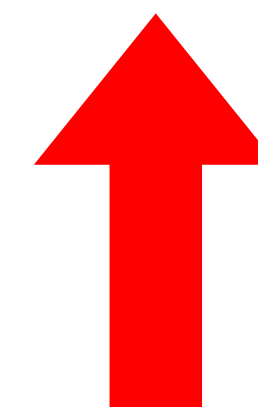
$$SST - SSR = SSM$$

$$100 - 95 = 5$$

$$100 - 10 = 90$$

=

SSM



AVALIANDO O MODELO

- Coeficiente de Ajuste do Modelo (R^2)
- Coeficiente de Correlação de Pearson (R)
- Teste F



Coeficiente de Ajuste do Modelo (R^2)

Representa o percentual de variação nos valores previstos que podem ser **explicados** pelo modelo.

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Coeficiente de Ajuste do Modelo (R^2)

Representa o percentual de variação nos valores previstos que podem ser **explicados** pelo modelo.

- Varia de 0 a 1
- Não infere **causalidade**!



Coeficiente de Ajuste do Modelo (R^2)

Representa o percentual de variação nos valores previstos que podem ser explicados pelo modelo.

Soma dos Quadrados
do Modelo

SSM

SST

Soma Total dos
Quadrados



Coeficiente de Ajuste do Modelo (R^2)

Representa o percentual de variação nos valores previstos que podem ser explicados pelo modelo.

Soma dos Quadrados
do Modelo

$$\frac{SSM}{SST} = \frac{SST - SSR}{SST}$$

Soma Total dos
Quadrados



Coeficiente de Ajuste do Modelo (R^2)

Representa o percentual de variação nos valores previstos que podem ser explicados pelo modelo.

Soma dos Quadrados
do Modelo

$$\frac{SSM}{SST} = \frac{SST - SSR}{SST} = R^2$$

Soma Total dos
Quadrados



Coeficiente de Ajuste do Modelo (R^2)

Representa o percentual de variação nos valores previstos que podem ser explicados pelo modelo.

$$\frac{SSM}{SST} = \frac{SST - SSR}{SST} = R^2$$

O ideal é que R^2 esteja cada mais próximo de 1!



Coeficiente de Ajuste do Modelo (R^2)

Representa o percentual de variação nos valores previstos que podem ser explicados pelo modelo.

$$\frac{SSM}{SST} = \frac{1 - 0}{1} = 1$$

Soma dos Quadrados Residual

O ideal é que R^2 esteja cada mais próximo de 1!



Coeficiente de Ajuste do Modelo (R^2)

Representa o percentual de variação nos valores previstos que podem ser explicados pelo modelo.

$$\frac{SSM}{SST} = \frac{1 - 0}{1} = 1$$

Modelo se ajustou MUITO bem!

O ideal é que R^2 esteja cada mais próximo de 1!



Coeficiente de Ajuste do Modelo (R^2)

Indica quanto da variável Y pode ser explicado pela variável X

Regressão linear

Peso (Y) — Altura (X)

$$R^2 = 0,500$$

LAC-FMB

Liga Acadêmica de Ciências



Coeficiente de Ajuste do Modelo (R^2)

Indica quanto da variável Y pode ser explicado pela variável X

Regressão linear

Peso (Y) — Altura (X)

$$R^2 = 0,500 = 50\%$$

LAC-FMB

Liga Acadêmica de Ciências



Coeficiente de Ajuste do Modelo (R^2)

Indica quanto da variável Y pode ser explicado pela variável X

Regressão linear

Peso (Y) — Altura (X)

$$R^2 = 0,500 = 50\%$$

Metade dos resultados da variável peso pode ser explicado pela variável altura.

LAC-FMB

Liga Acadêmica de Ciências



Coeficiente de Ajuste do Modelo (R^2)

Indica quanto da variável Y pode ser explicado pela variável X

Regressão linear

Peso (Y) — Altura (X)

$$R^2 = 50\%$$

● Dieta ● Exercício ● Genética



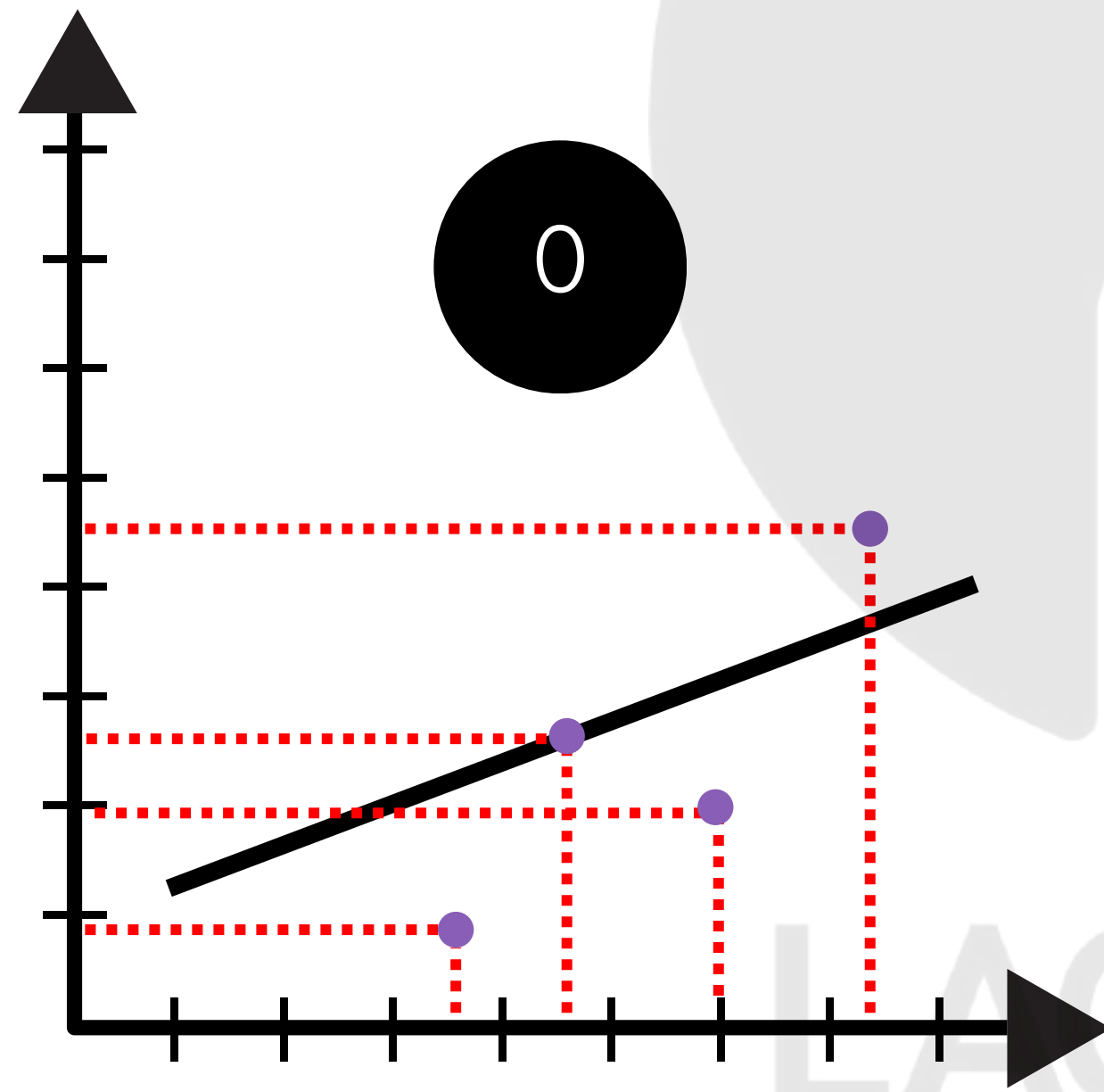
Coeficiente de Ajuste do Modelo (R^2)

LAC-FMB

Liga Acadêmica de Ciências



Coeficiente de Ajuste do Modelo (R^2)

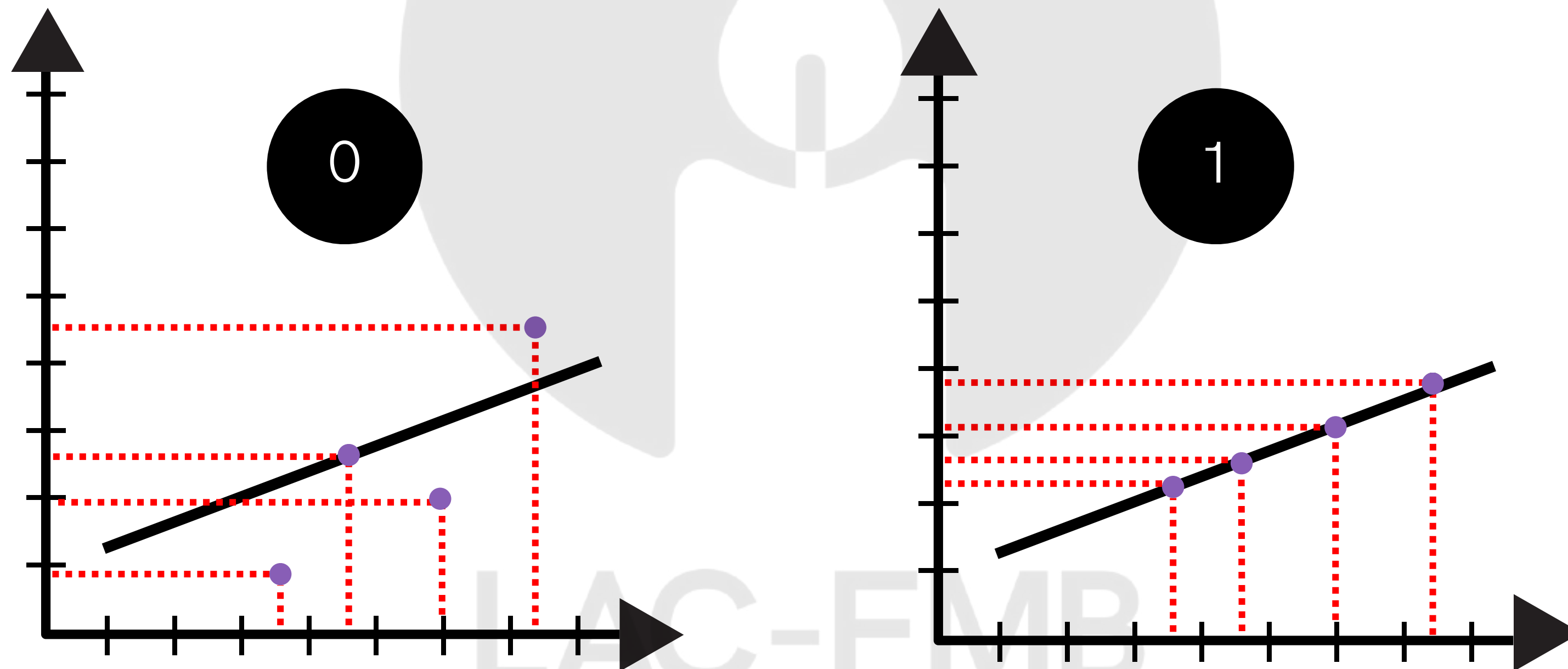


LAC-FMB

Liga Acadêmica de Ciências



Coeficiente de Ajuste do Modelo (R^2)



Coeficiente de Correlação de Pearson (R)

$$\sqrt{R^2} = R$$

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Coeficiente de Correlação de Pearson (R)

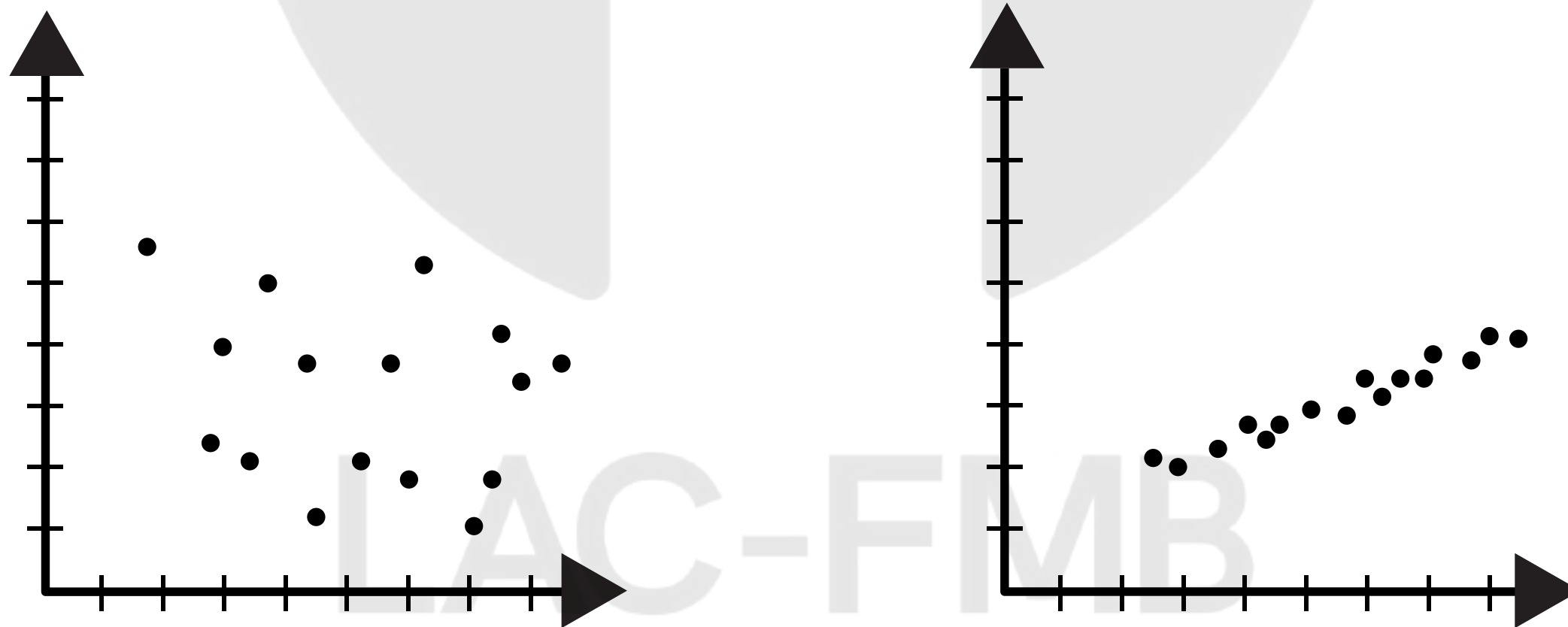
$$\sqrt{R^2} = R$$

- Avalia qualidade do modelo
- Dá uma boa medida do valor do relacionamento entre variáveis
- Fornece uma boa estimativa da aderência do modelo de regressão



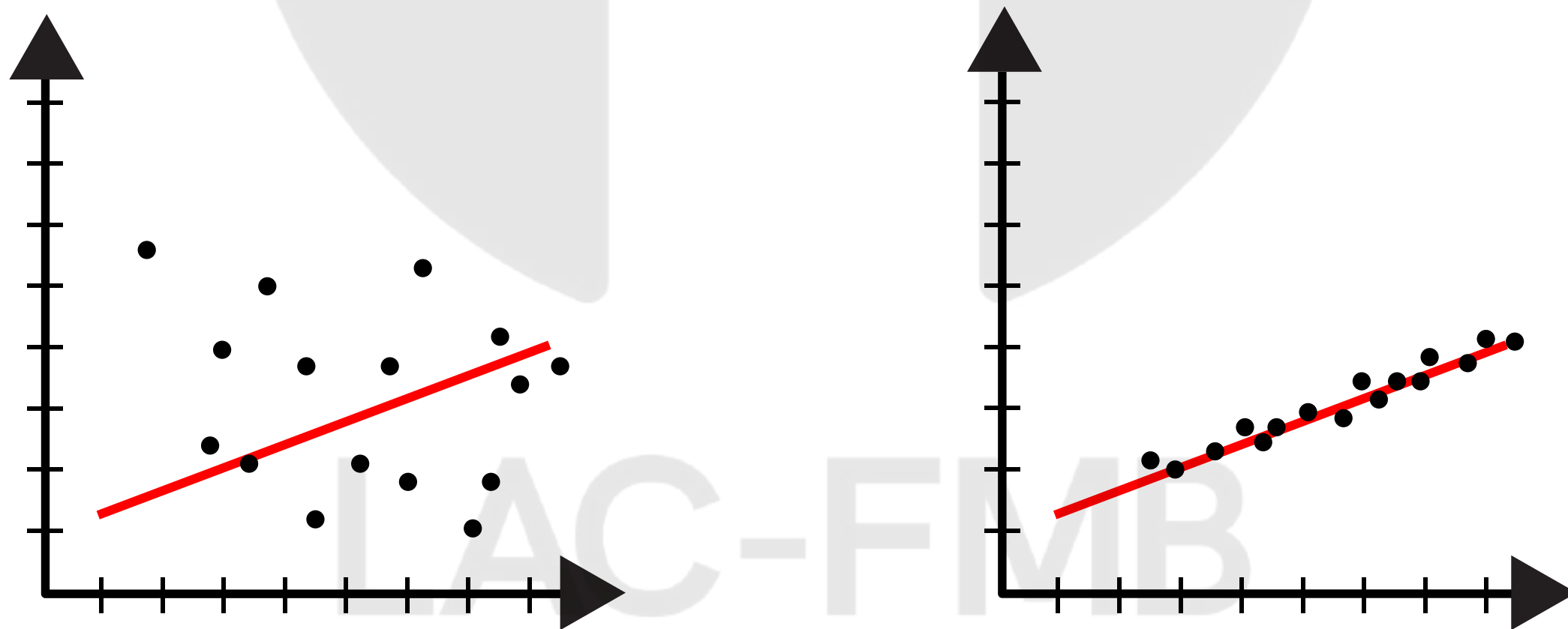
Coeficiente de Correlação de Pearson (**R**)

$$\sqrt{R^2} = R$$



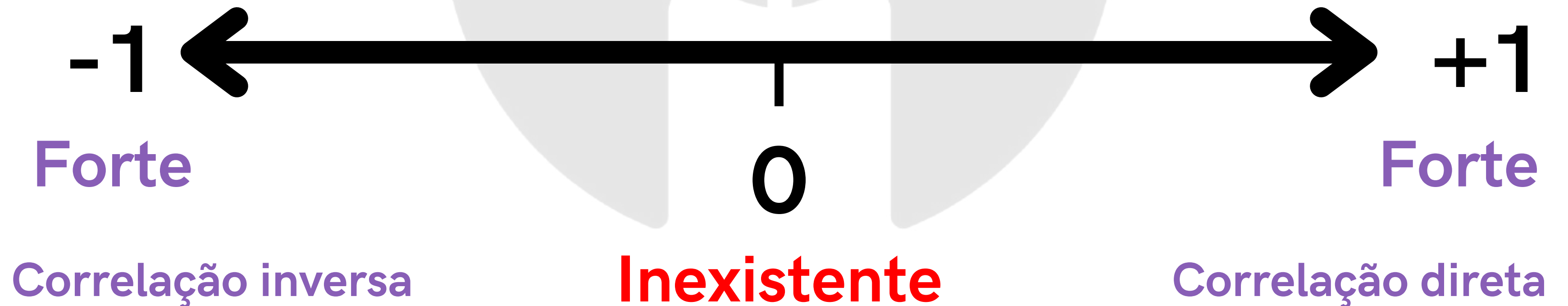
Coeficiente de Correlação de Pearson (**R**)

$$\sqrt{R^2} = R$$



Coeficiente de Correlação de Pearson (**R**)

Avalia se correlação entre as variáveis é estatisticamente significativa



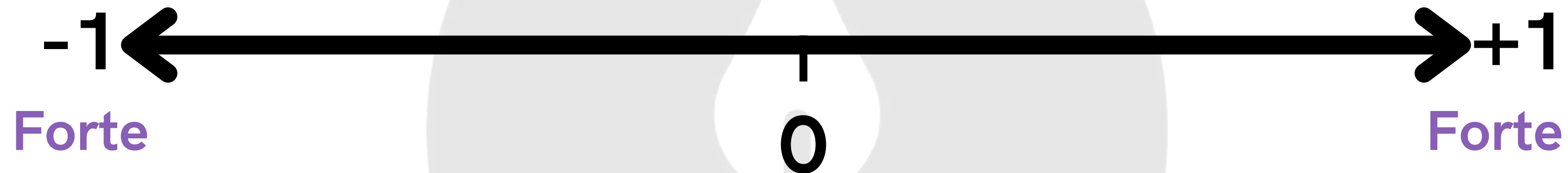
LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Coeficiente de Correlação de Pearson (**R**)



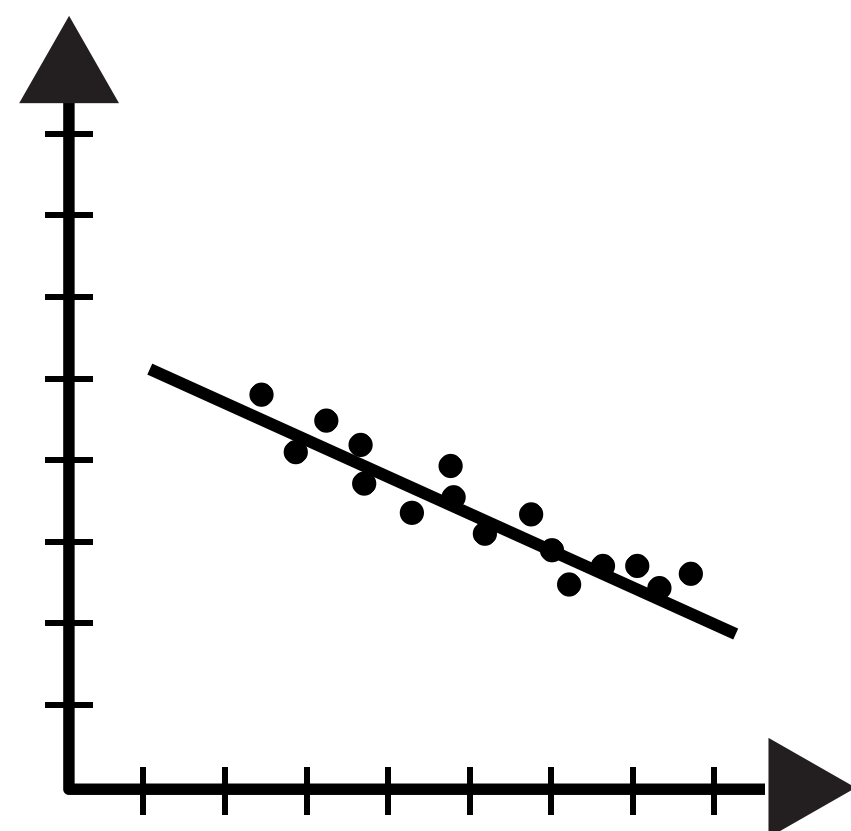
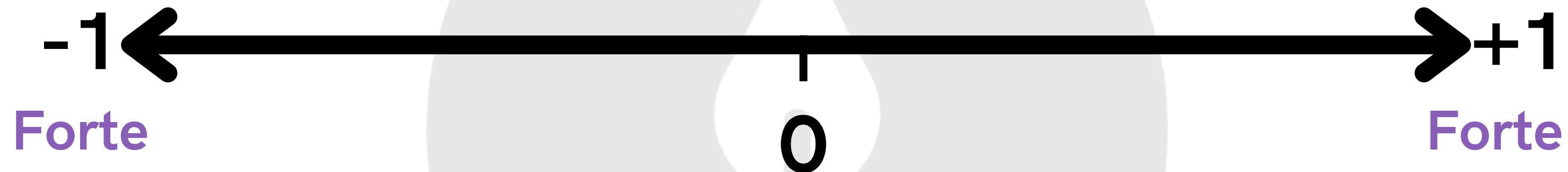
LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Coeficiente de Correlação de Pearson (R)



Correlação inversa

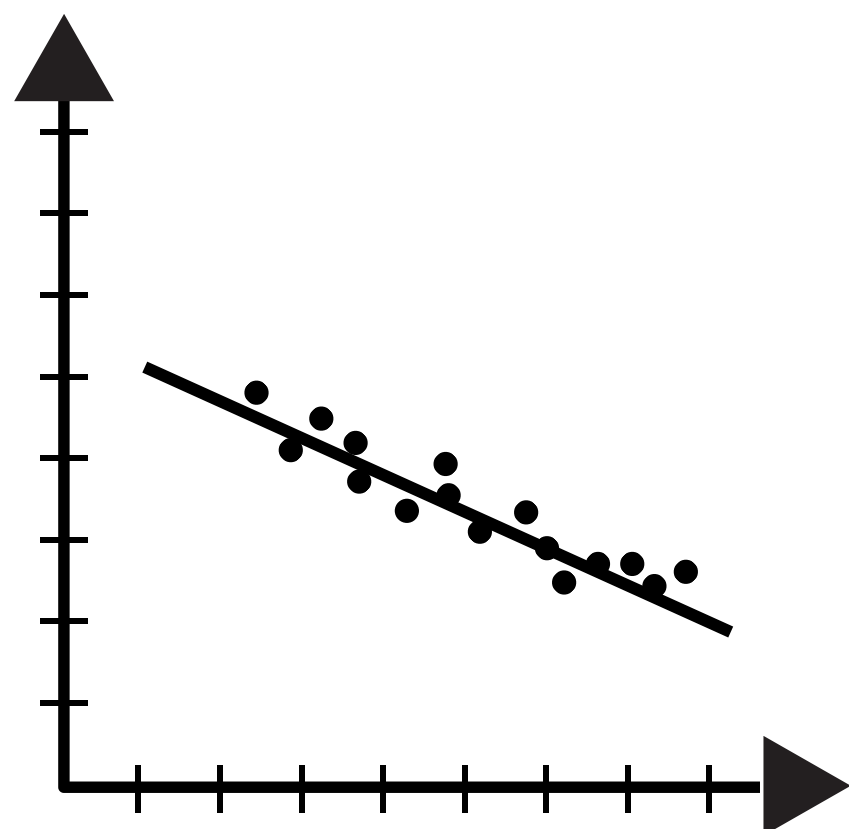
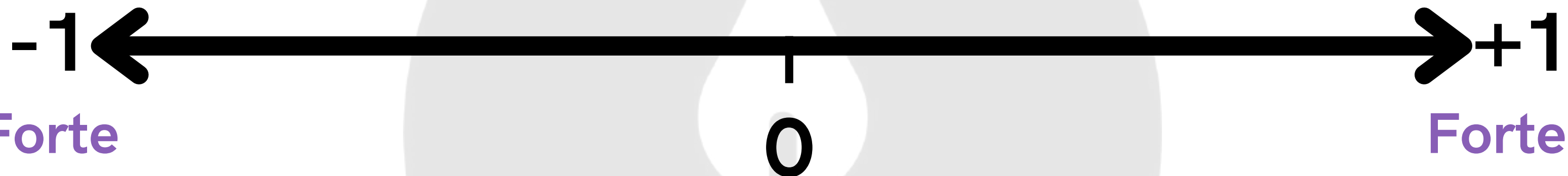
LAC-FMB

Liga Acadêmica de Ciências

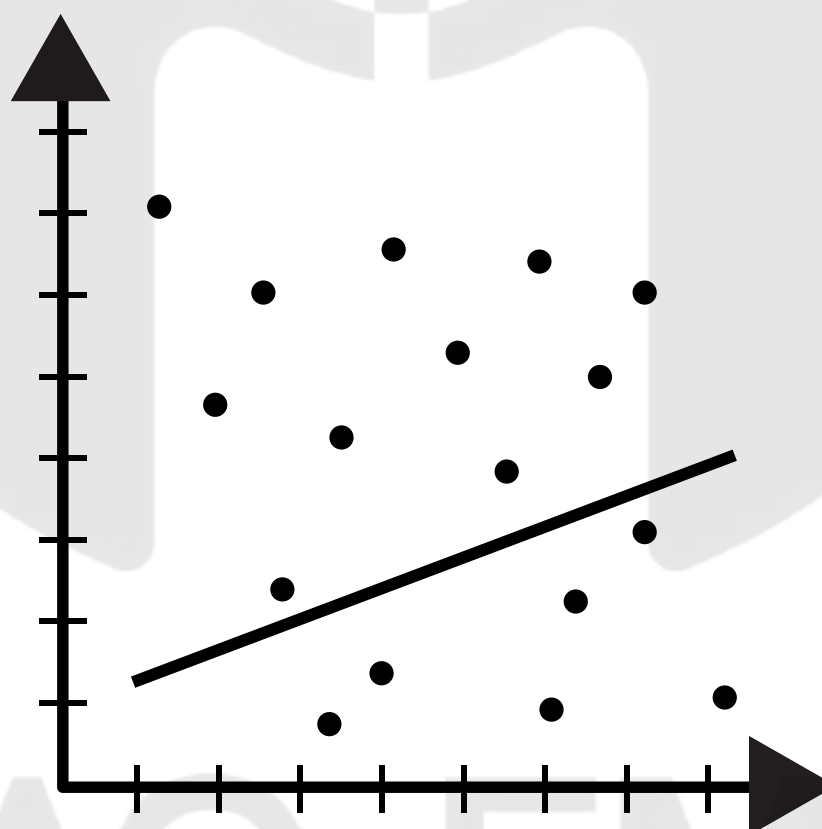
FIELD, A., 2009



Coeficiente de Correlação de Pearson (R)



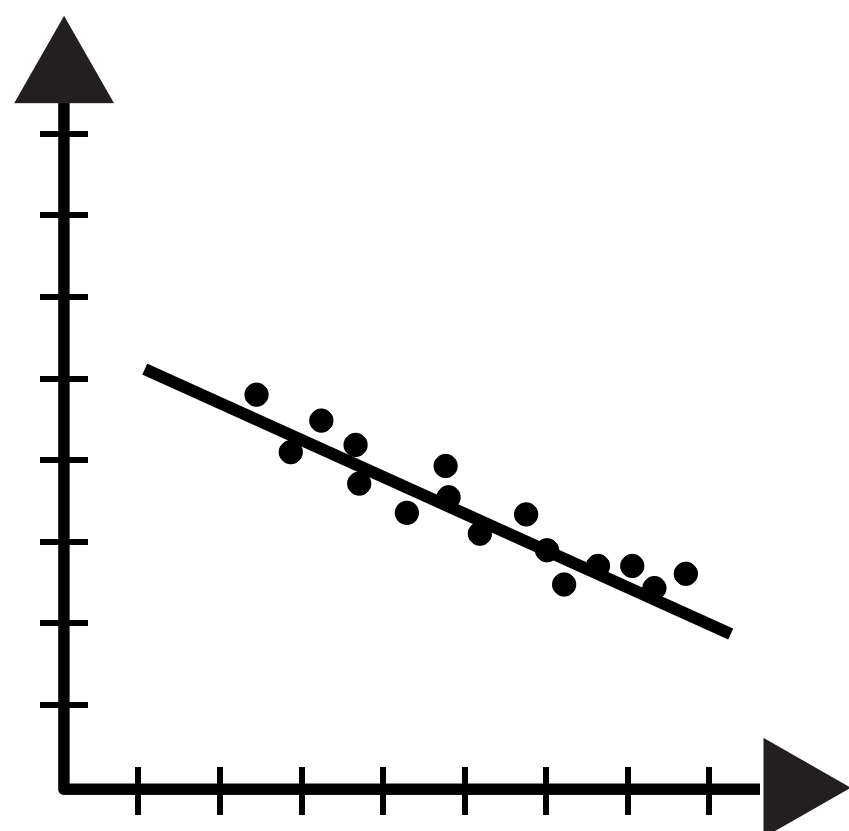
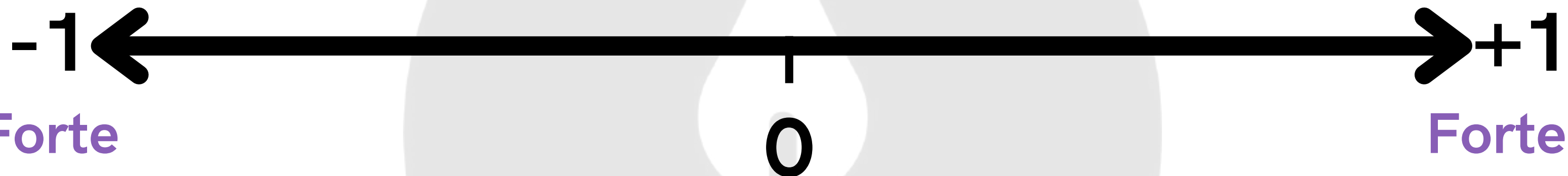
Correlação inversa



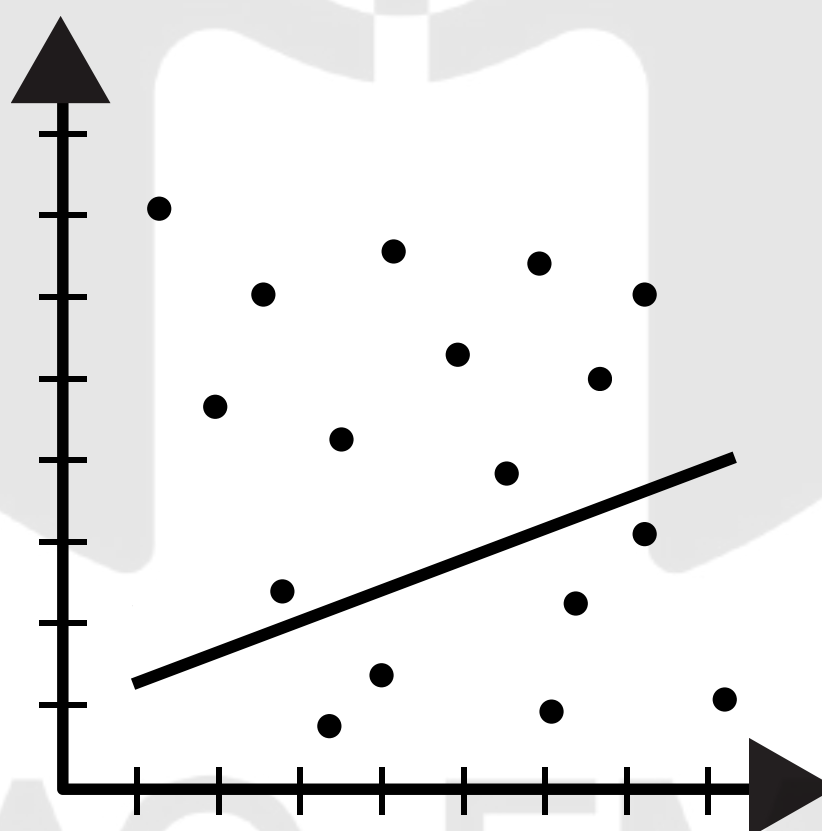
Inexistente



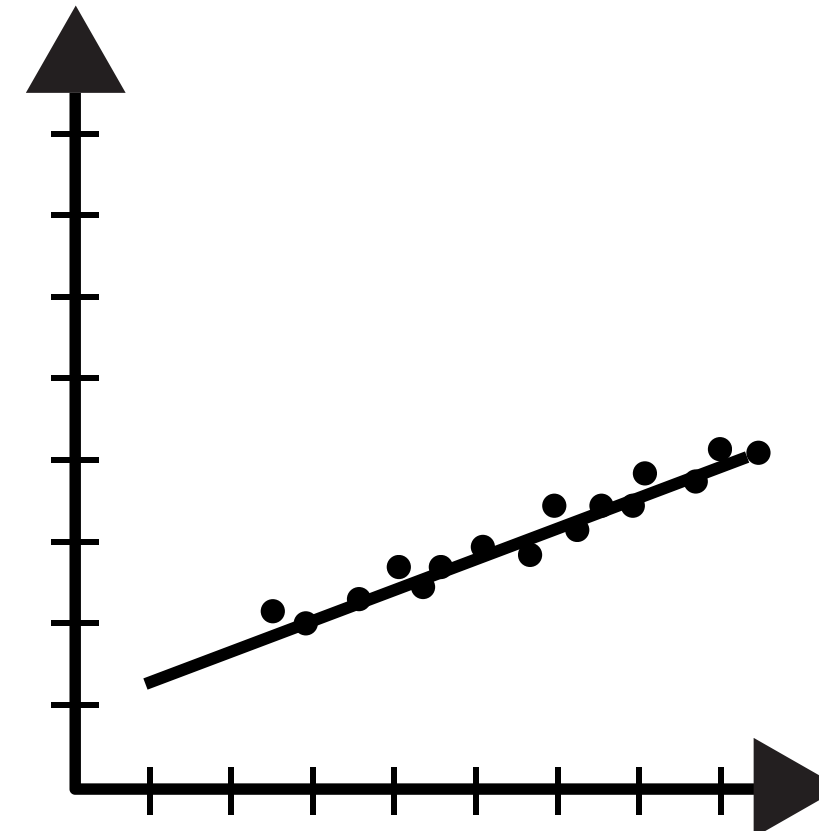
Coeficiente de Correlação de Pearson (**R**)



Correlação inversa



Inexistente



Correlação direta



Teste F

Avalia significância estatística do MODELO

LAC-FMB

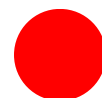
Liga Acadêmica de Ciências



Teste F

Avalia significância estatística do MODELO

Teste F



R^2

LAC-FMB

Liga Acadêmica de Ciências



Teste F

Avalia significância estatística do MODELO

Teste F

- "Modelo é estatisticamente significativo?"
- "Quão bem a variável independente do modelo consegue explicar a dependente?"

R^2

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Teste F

Avalia significância estatística do MODELO

Teste F

- "Modelo é estatisticamente significativo?"
- Valor de p
- "Quão bem a variável independente do modelo consegue explicar a dependente?"
- Porcentagem

R^2

LAC-FMB

Liga Acadêmica de Ciências



Teste F

Avalia significância estatística do MODELO

Teste F

R^2

- "Modelo é estatisticamente significativo?"
- Valor de p
- "Quão bem a variável independente do modelo consegue explicar a dependente?"
- 0 - 1
- Modelo pode ser estatisticamente significativo, mas variável independente não explicar muito bem a variável dependente!

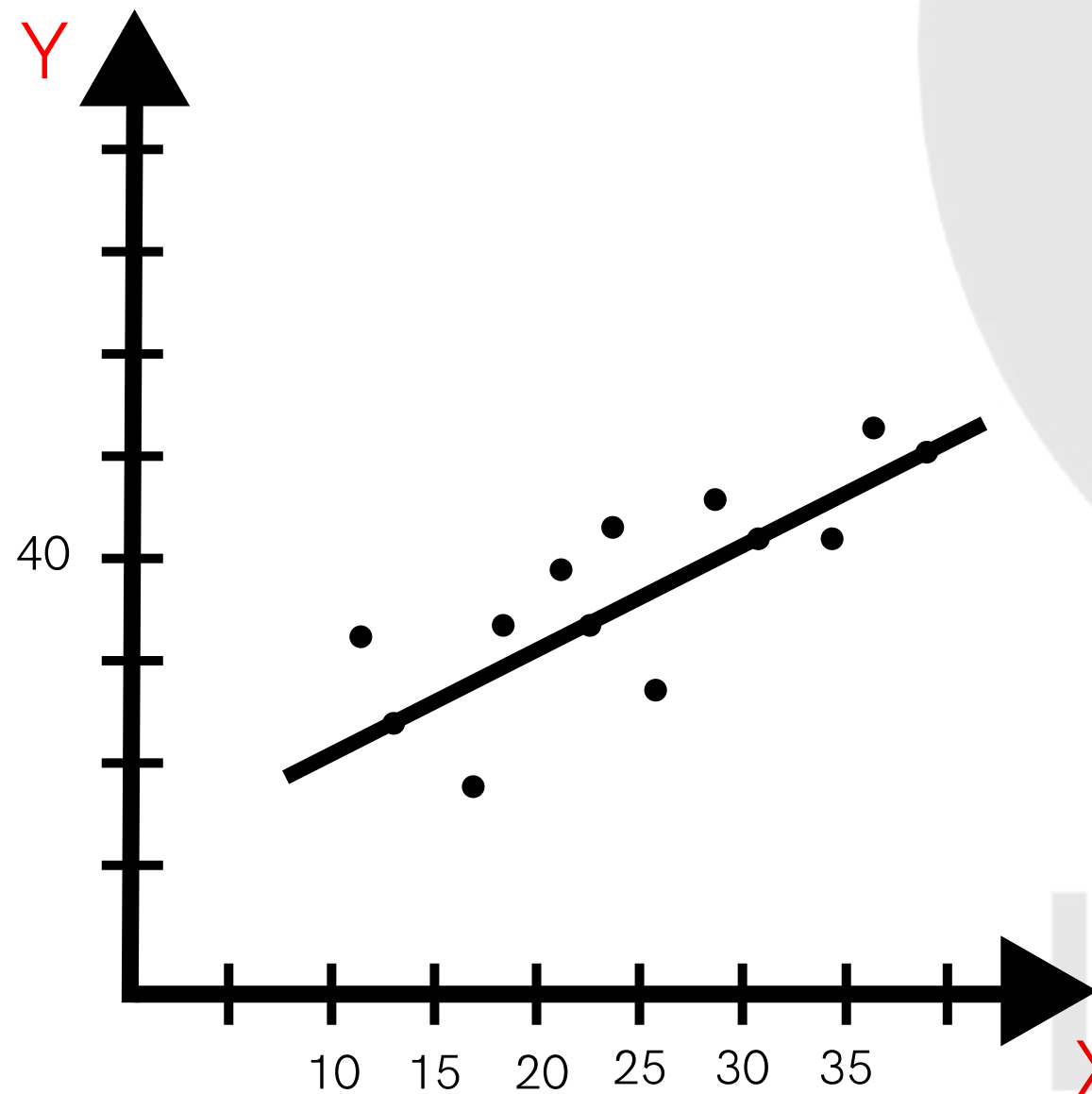


Pressupostos

- Normalidade dos resíduos
- Homocedasticidade
- Independência

Normalidade dos resíduos

Distribuição normal dos resíduos



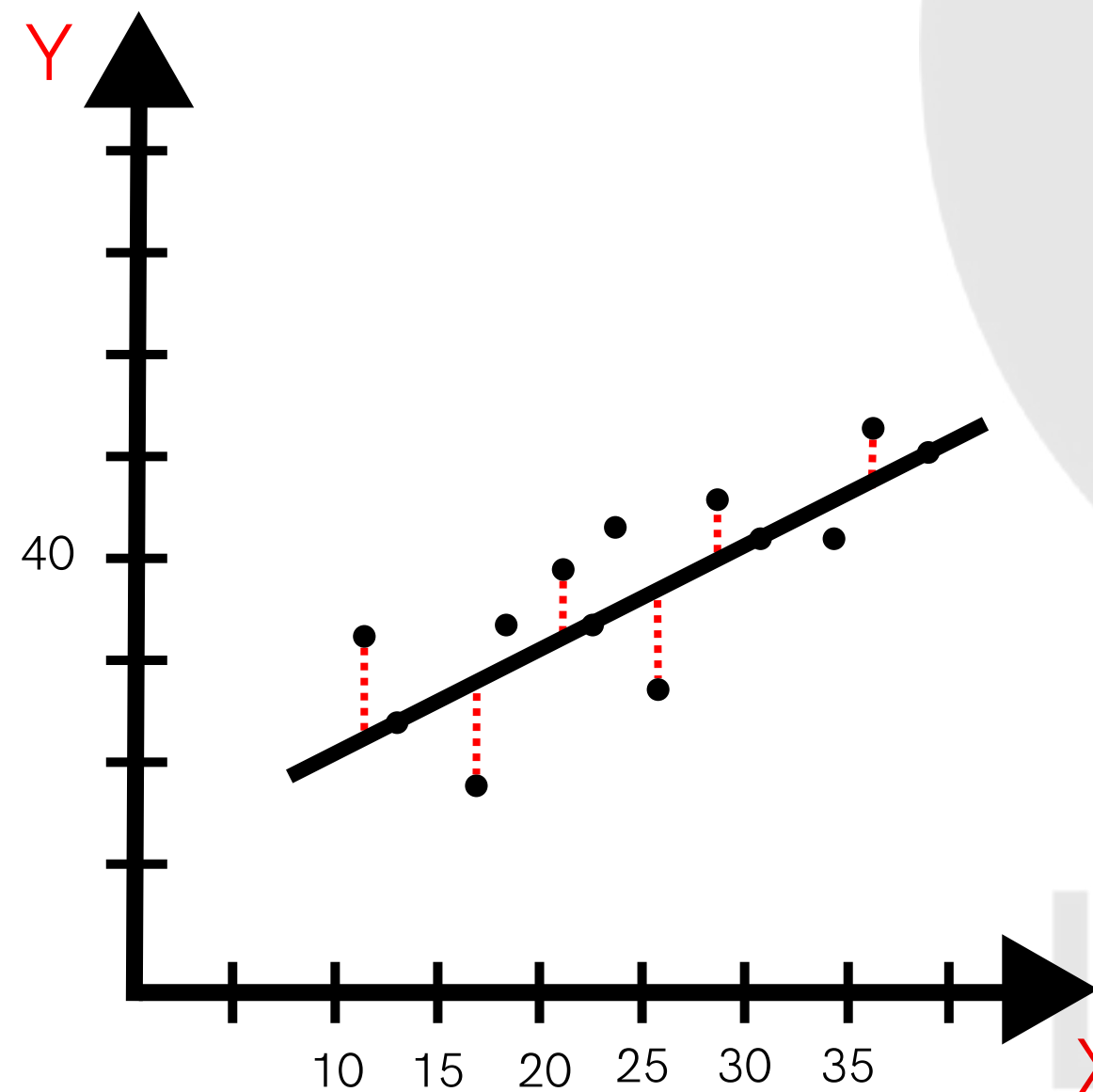
LAC-FMB

Liga Acadêmica de Ciências



Normalidade dos resíduos

Distribuição normal dos resíduos



Todos os pontos em cima da reta



Distribuição normal perfeita

LAC-FMB

Liga Acadêmica de Ciências

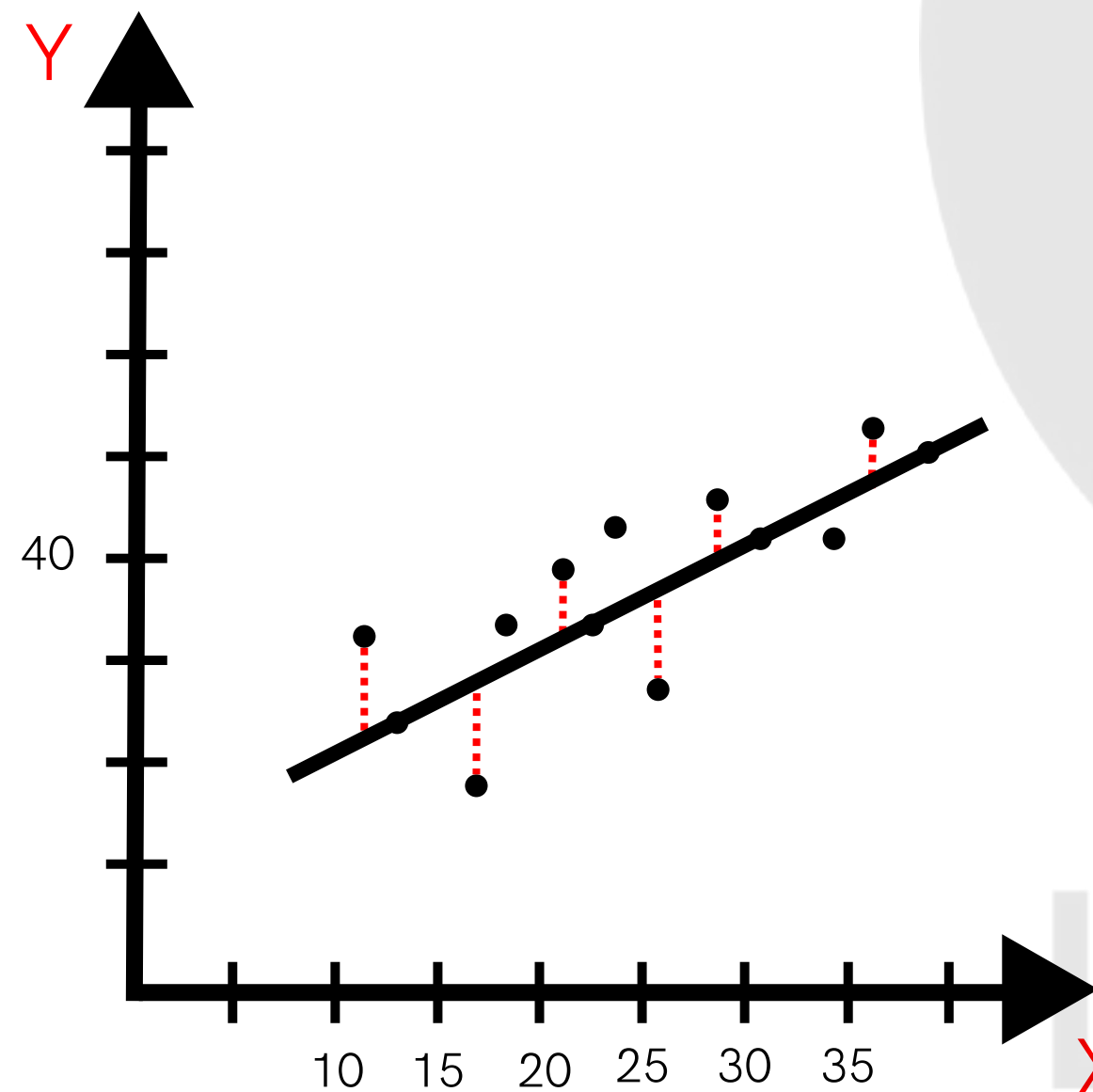
FIELD, A., 2009



Normalidade dos resíduos

Distribuição normal dos resíduos

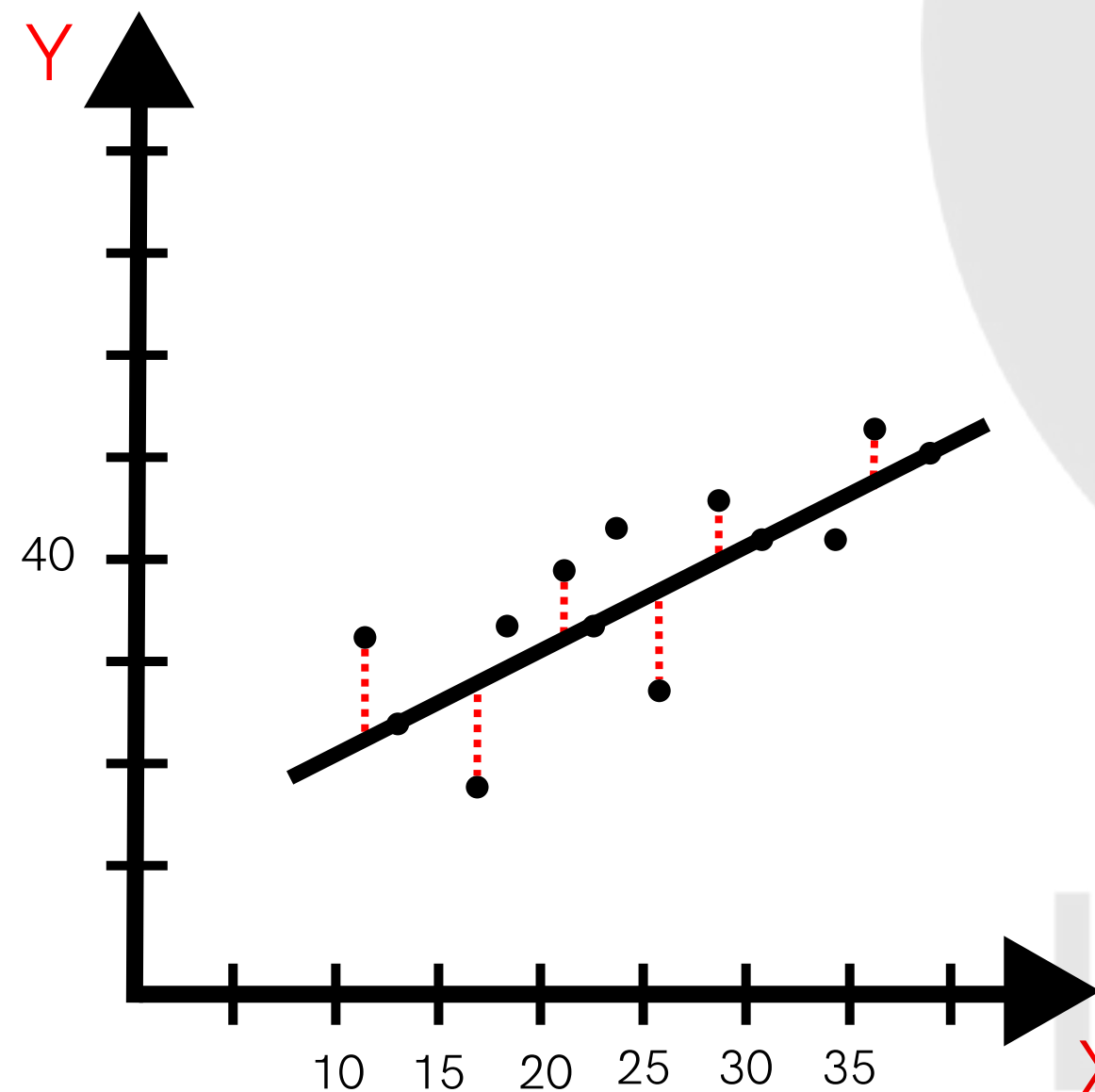
Teste Shapiro-Wilk e valor de p



Normalidade dos resíduos

Distribuição normal dos resíduos

Teste Shapiro-Wilk e valor de p



Normality Test (Shapiro-Wilk)

Statistic

p

0.990

0.740

LAC-FMB

Liga Acadêmica de Ciências

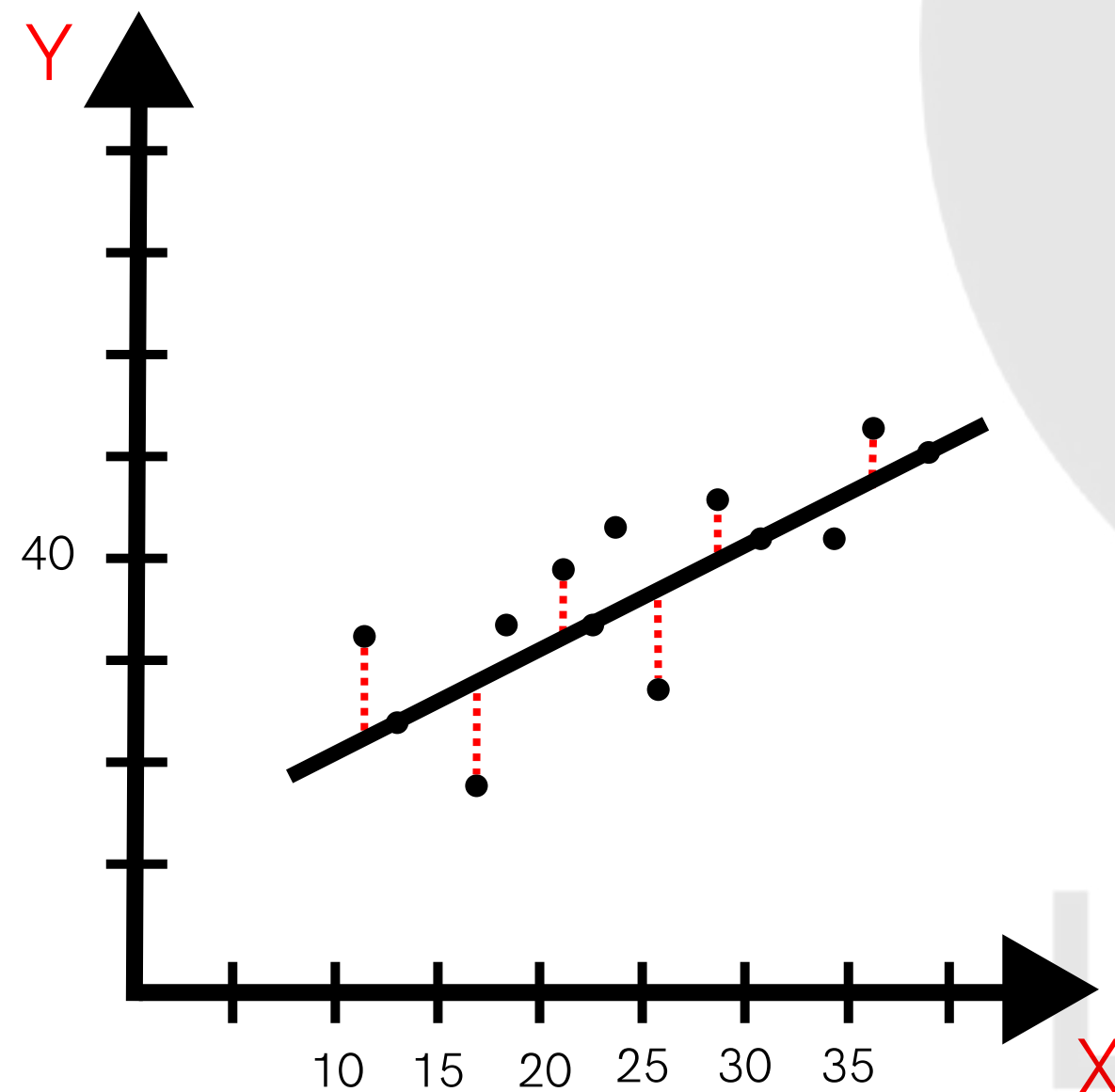
FIELD, A., 2009



Normalidade dos resíduos

Distribuição normal dos resíduos

Teste Shapiro-Wilk e valor de p



Normality Test (Shapiro-Wilk)

Statistic

p

0.990

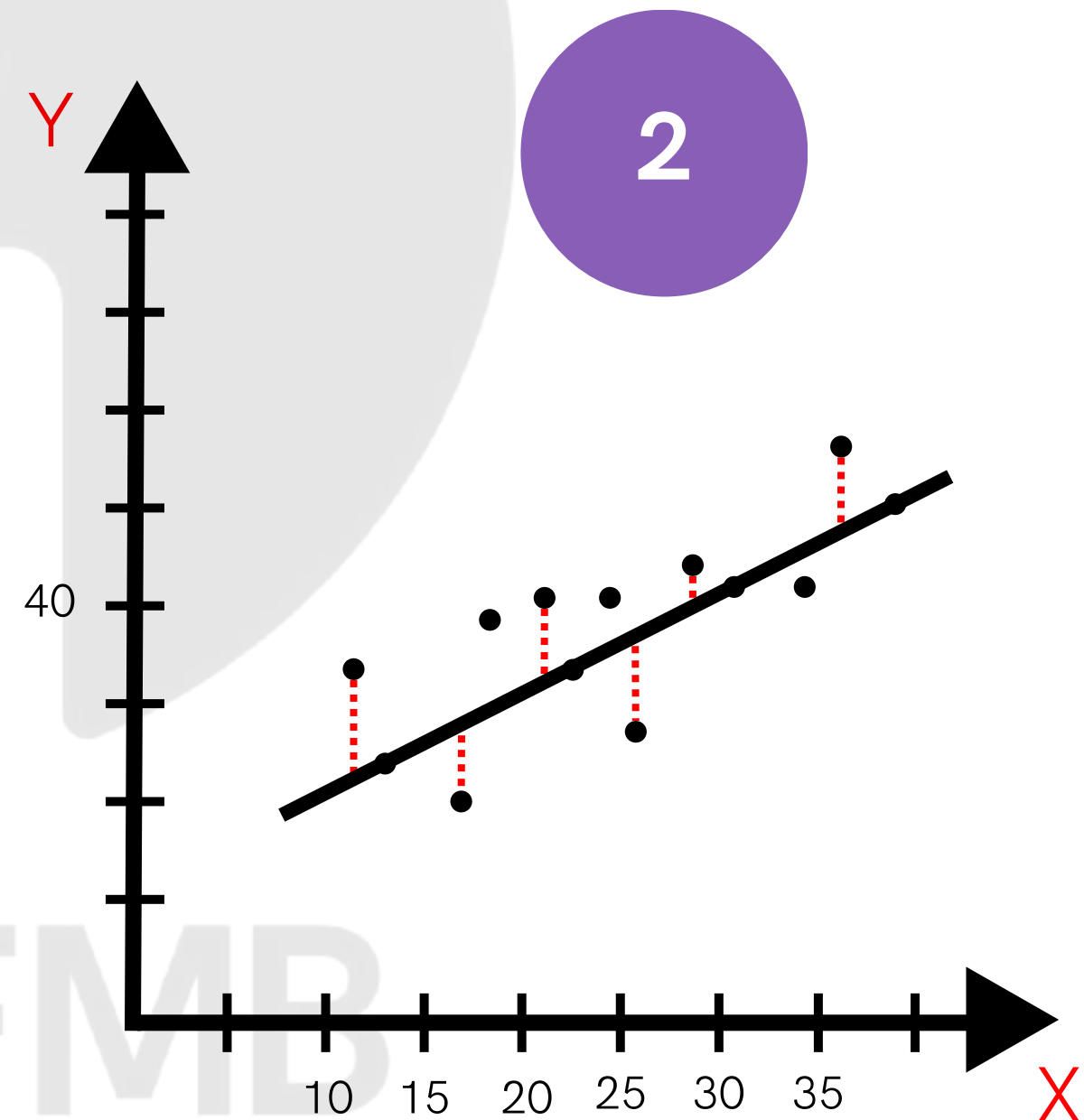
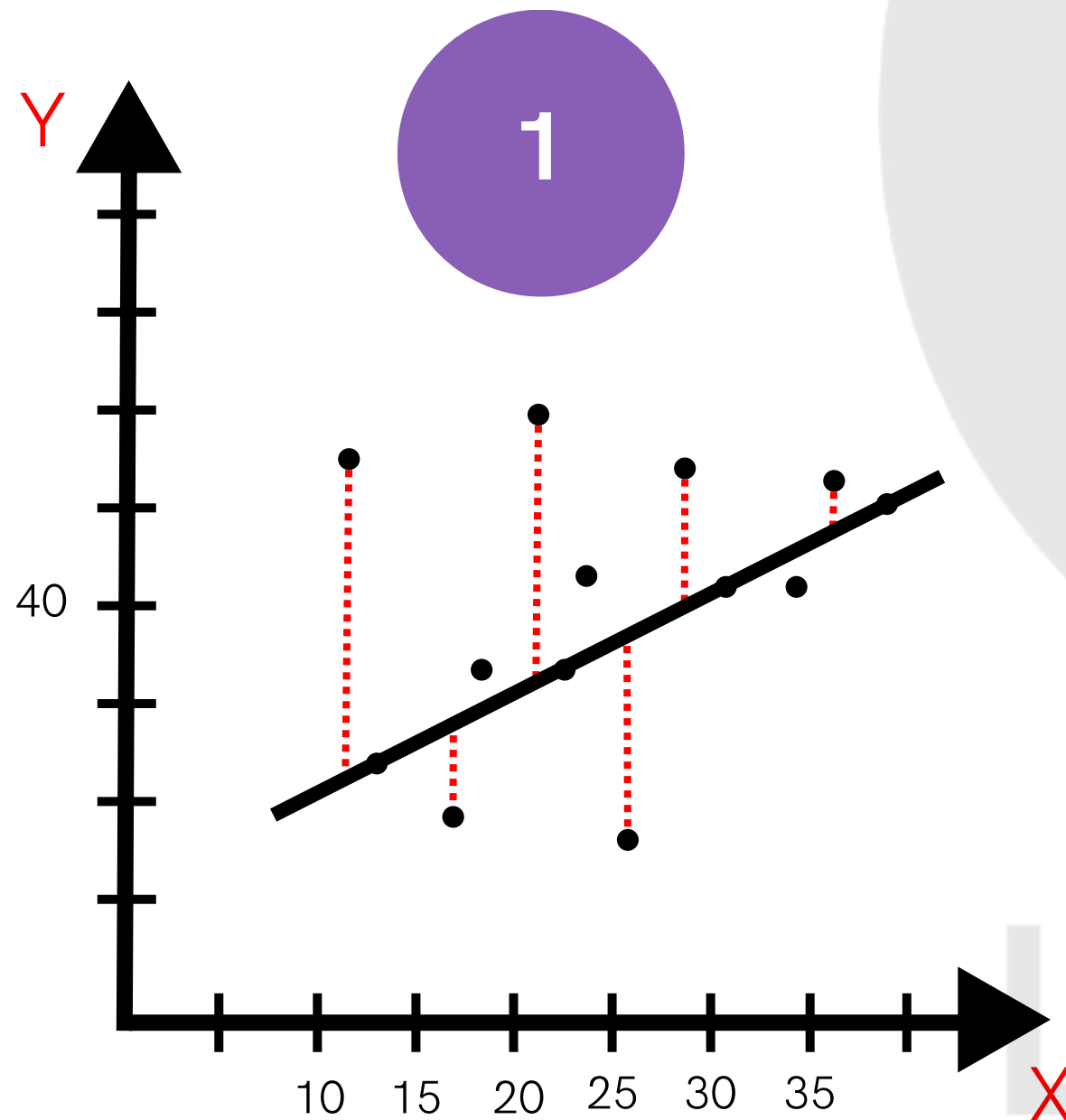
0.740

● Distribuição dos resíduos não se diferencia de uma maneira significativa de uma distribuição normal!



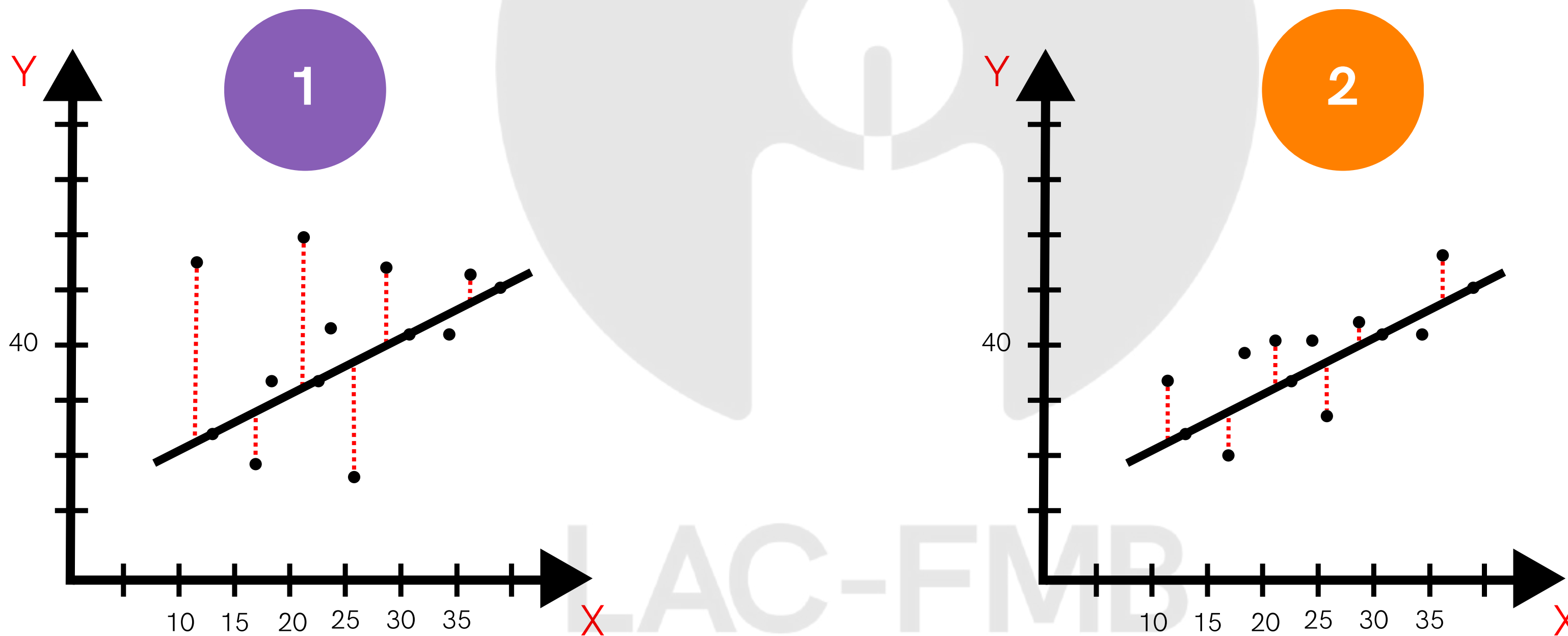
Normalidade dos resíduos

Qual a distribuição **MAIS NORMAL?**



Normalidade dos resíduos

Qual a distribuição **MAIS NORMAL?**



Homocedasticidade

Significa que os erros em um modelo de regressão têm a mesma dispersão, independentemente dos valores da variável independente.

LAC-FMB

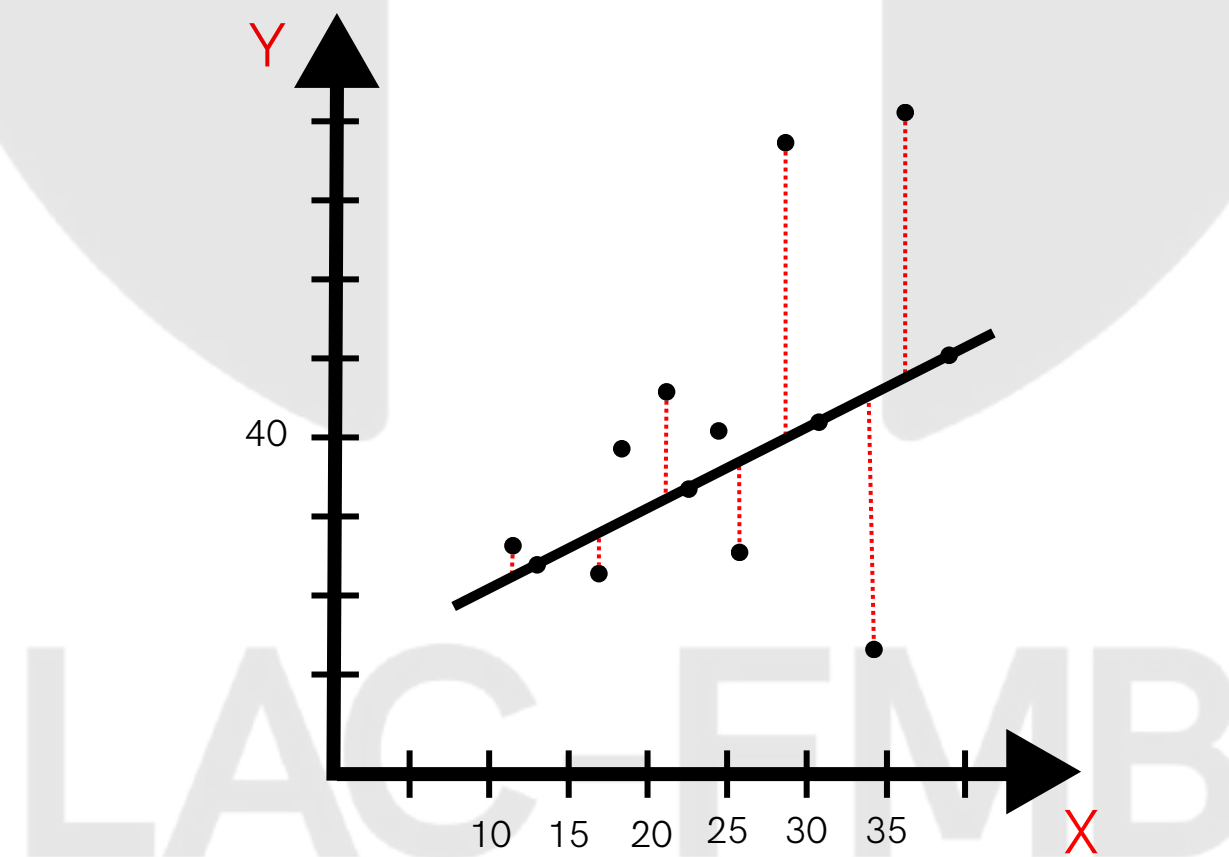
Liga Acadêmica de Ciências

FIELD, A., 2009



Homocedasticidade

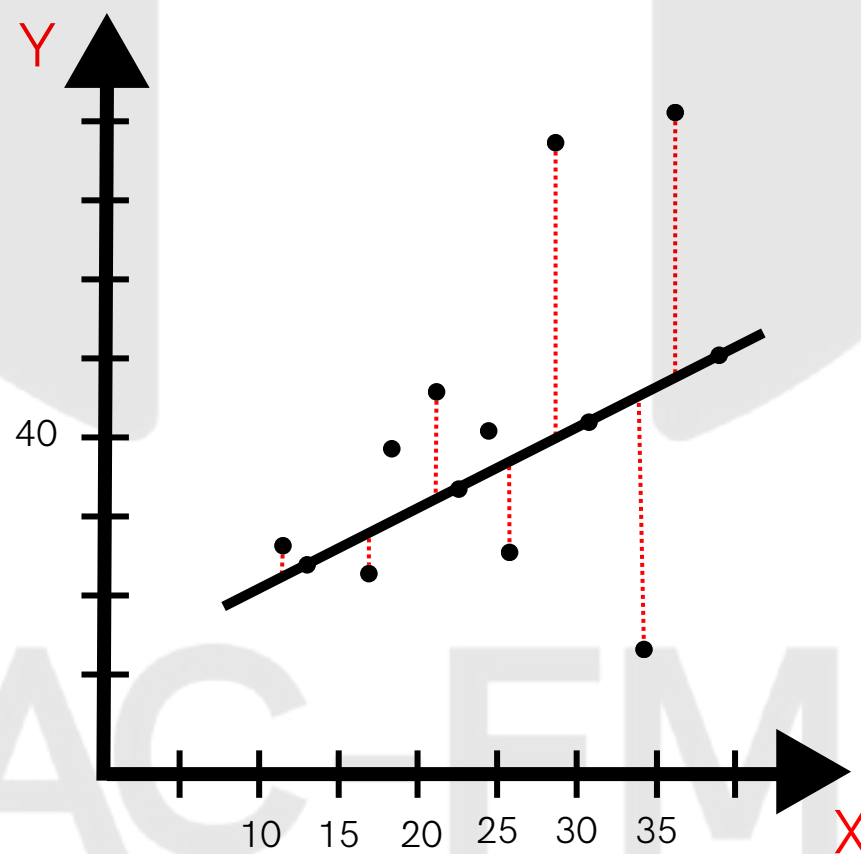
Significa que os erros em um modelo de regressão têm a mesma dispersão, independentemente dos valores da variável independente.



Homocedasticidade

Significa que os erros em um modelo de regressão têm a mesma dispersão, independentemente dos valores da variável independente.

● Teste de Levene



Independência

Todos os valores da variável de saída são independentes entre si.

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Independência

Um resíduo não é influenciado por outro resíduo.

LAC-FMB

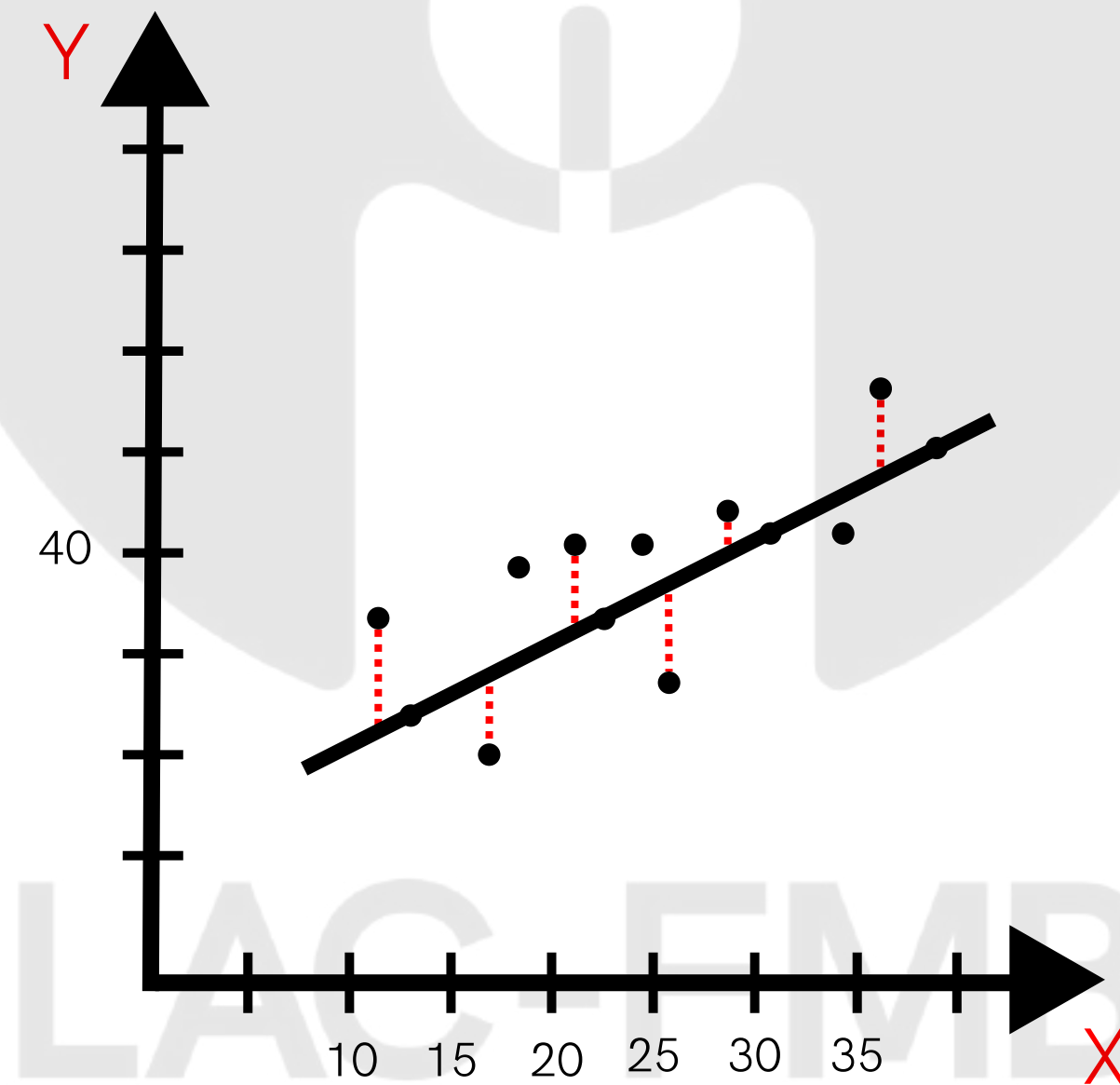
Liga Acadêmica de Ciências

FIELD, A., 2009



Independência

Um resíduo não é influenciado por outro resíduo.

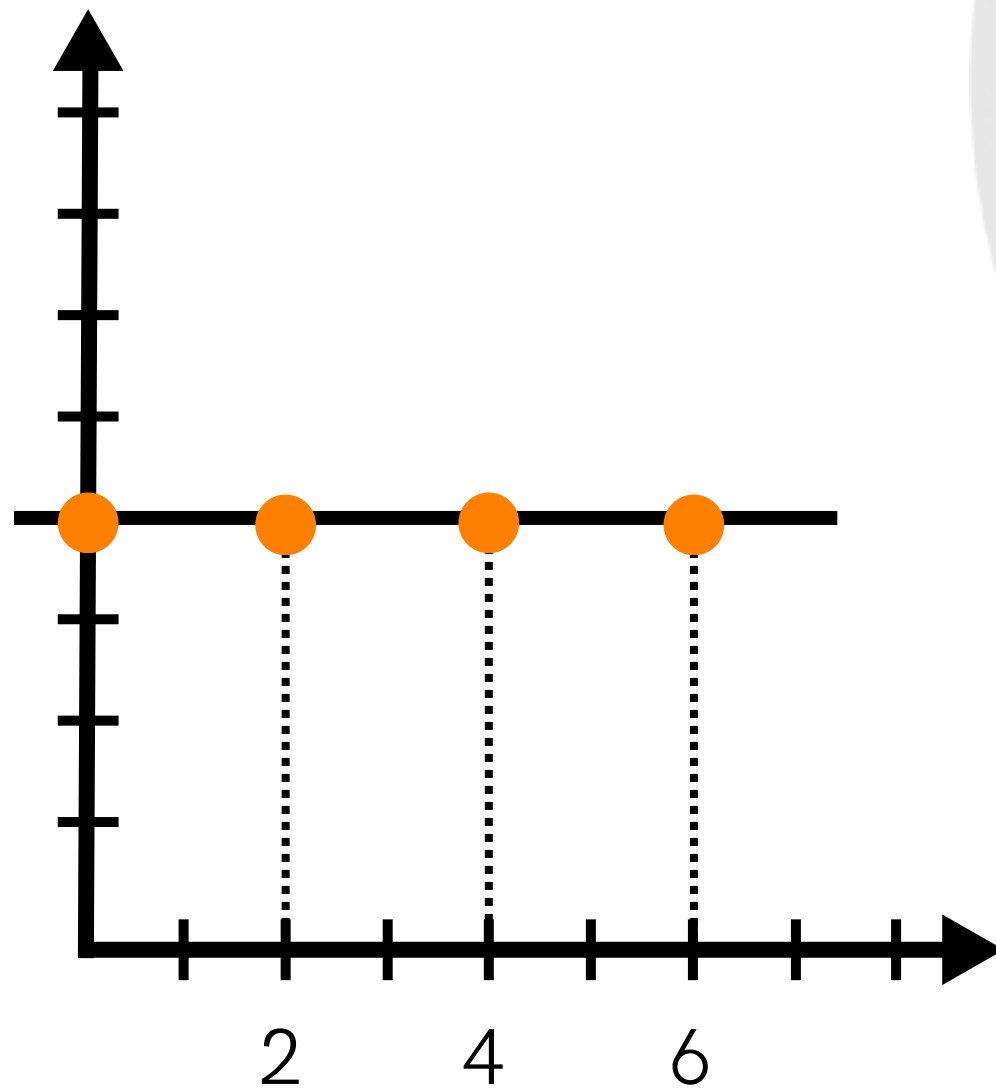


INTERPRETANDO PREVISÕES INDIVIDUAIS



Teste t

$$y = \beta_1 x + \beta_0$$



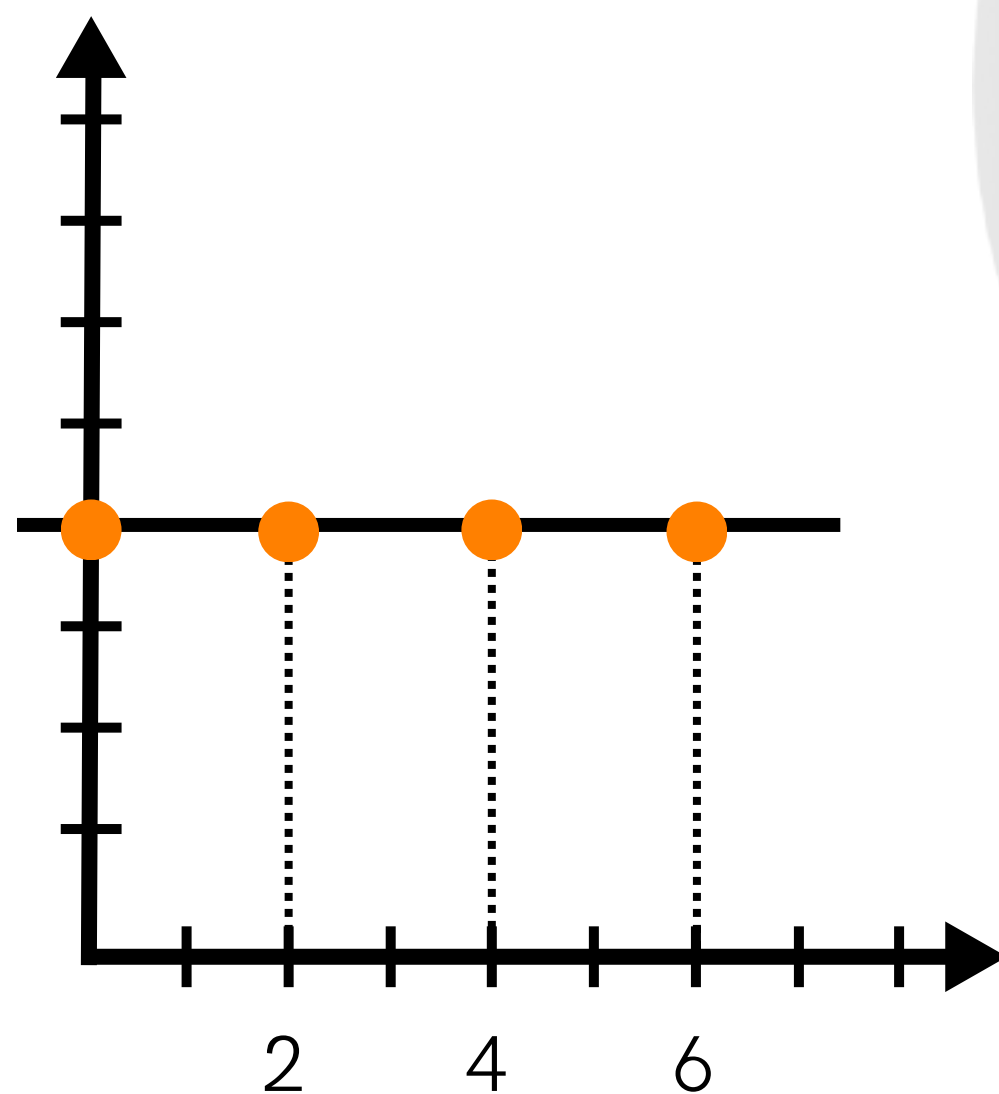
LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$y = \beta_1 x + \beta_0$$



- Representa a mudança na saída resultante de uma unidade no previsor.

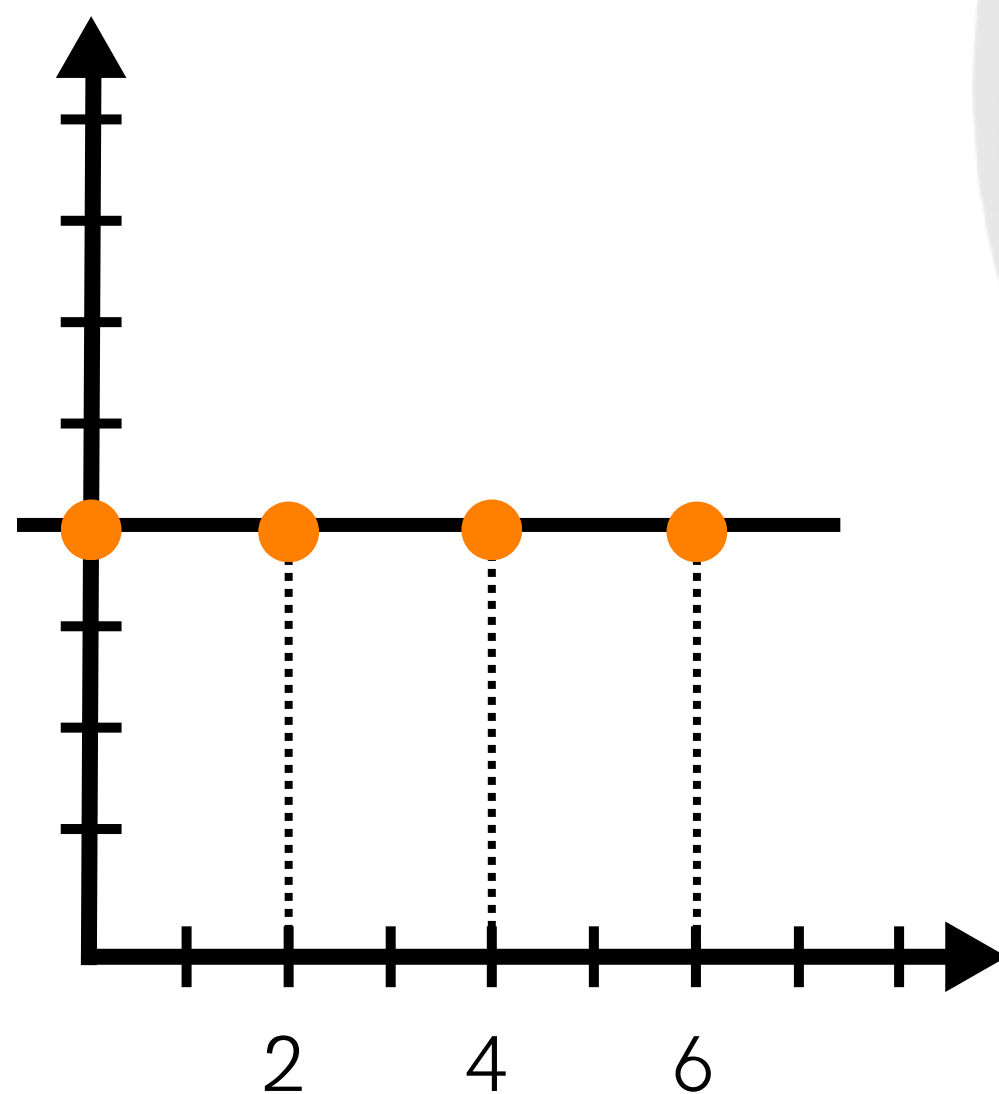
LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$y = \beta_1 x + \beta_0$$



● Representa a mudança na saída resultante de uma unidade no previsor.

● $\beta_1 = 0$

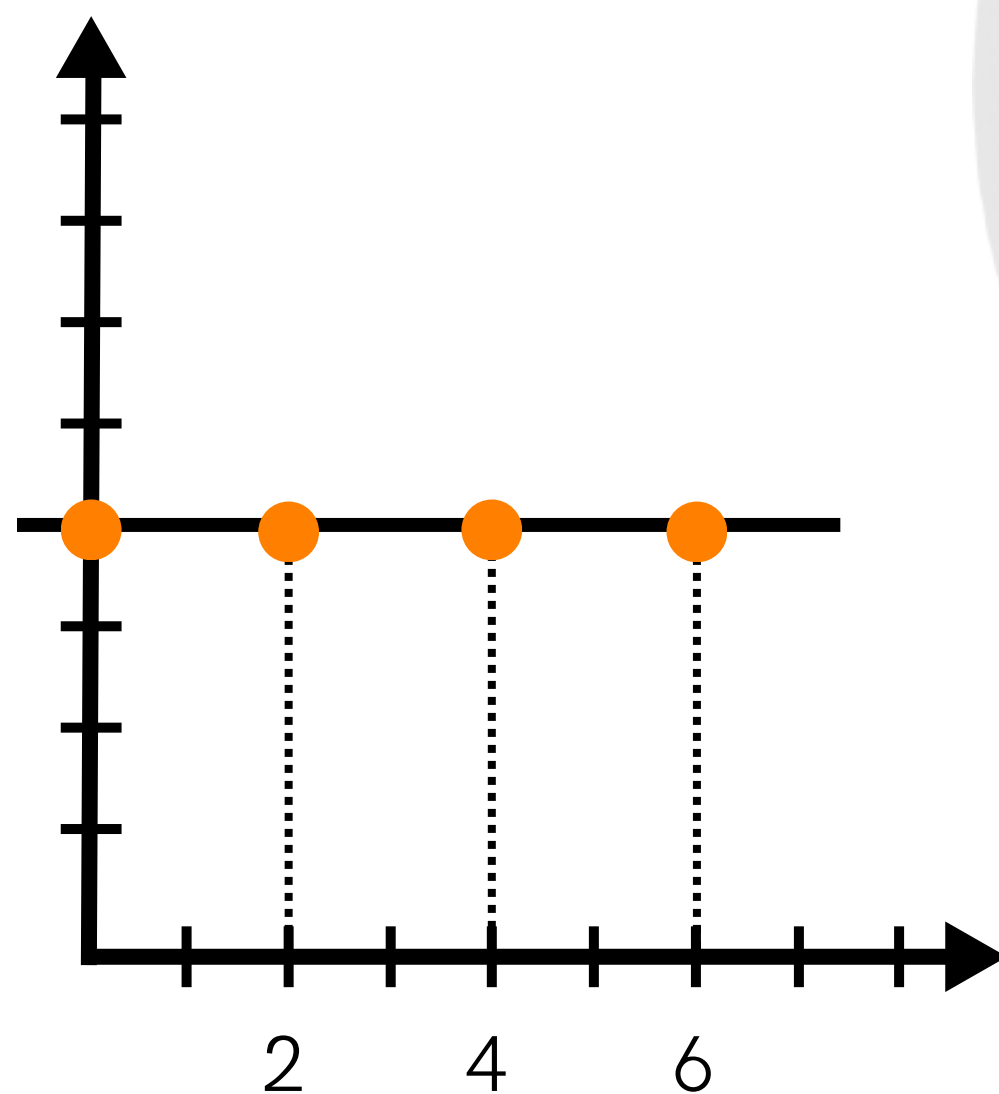
LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$y = \beta_1 x + \beta_0$$



- Representa a mudança na saída resultante de uma unidade no previsor.

- $\beta_1 = 0$

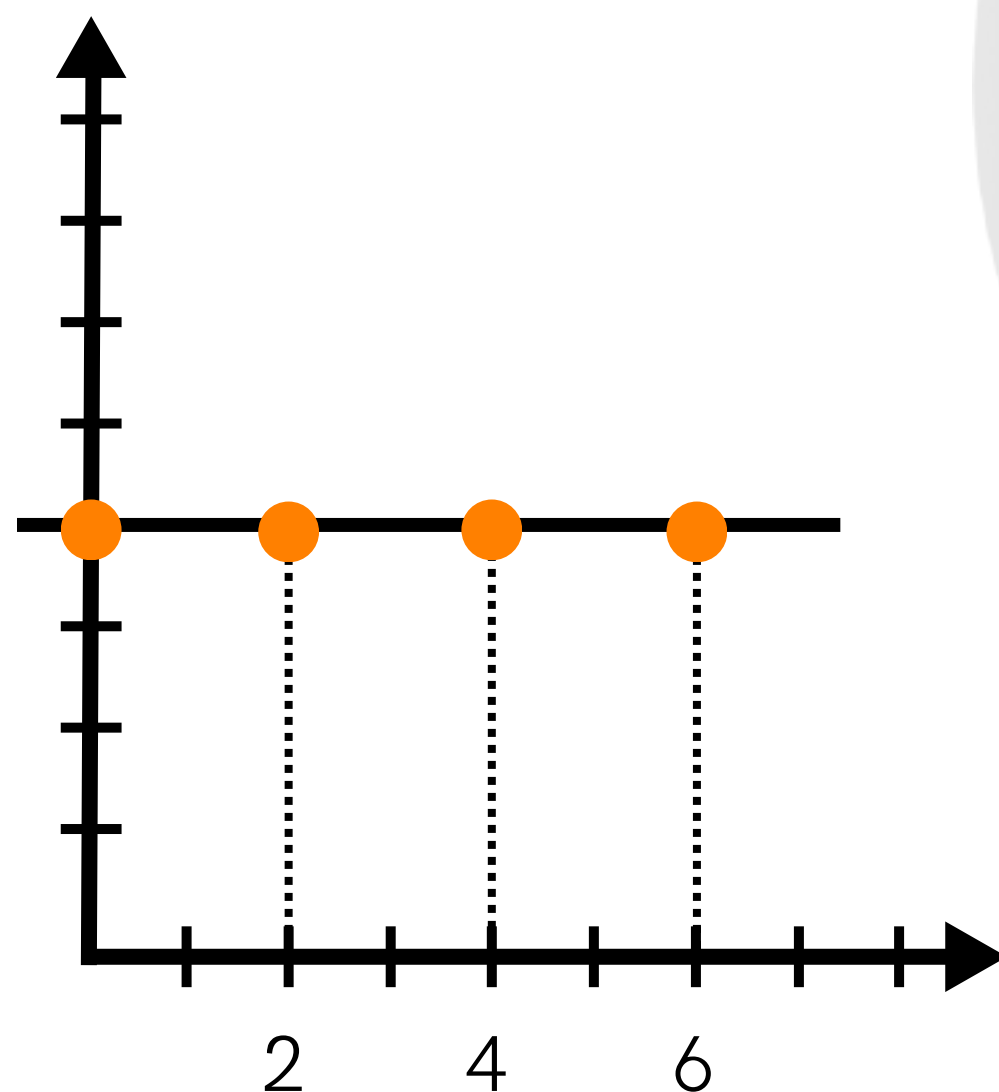
LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$y = \beta_1 x + \beta_0$$



● Representa a mudança na saída resultante de uma unidade no previsor.

● $\beta_1 = 0 \quad = \quad Y = \beta_0$

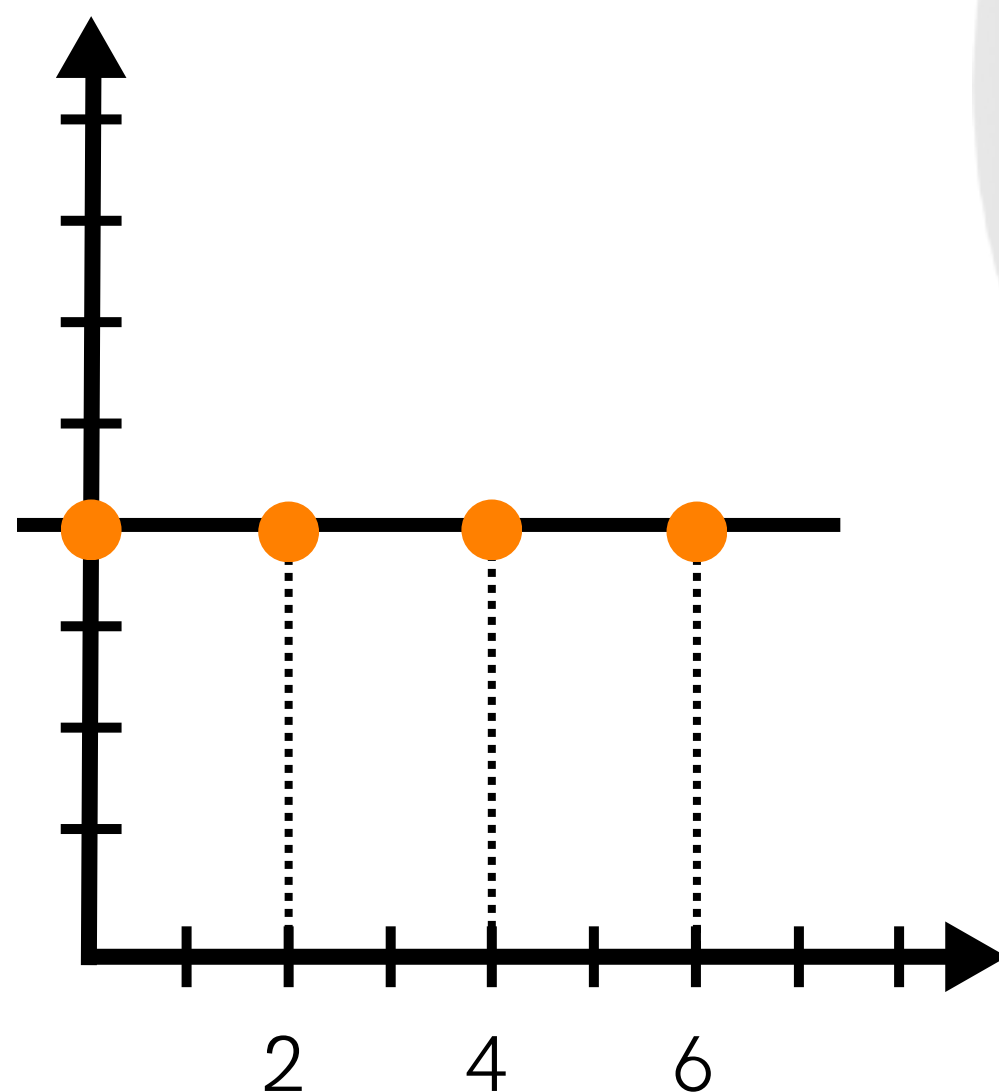
LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$y = \beta_1 x + \beta_0$$



- Representa a mudança na saída resultante de uma unidade no previsor.
- $\beta_1 = 0 \quad \Rightarrow \quad Y = \beta_0$
- Modelo ruim = média

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Testa a hipótese nula de que o valor de β_1 é zero: dessa forma, se for significativo aceitamos a hipótese que o valor de β_1 é significativamente diferente de zero e que a variável previsora contribui significativamente para a nossa habilidade de estimar o valor de saída.

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Teste t

Testa a hipótese nula de que o valor de β_1 é zero: dessa forma, se for significativo aceitamos a hipótese que o valor de β_1 é significativamente diferente de zero e que a variável previsora contribui significativamente para a nossa habilidade de estimar o valor de saíd.

LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$



**Qual a variável independente
e qual a variável dependente?**

LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

Variável independente (X)

- **Altura**

Variável dependente (Y)

- **Peso**

LAC-FMB

Liga Acadêmica de Ciências



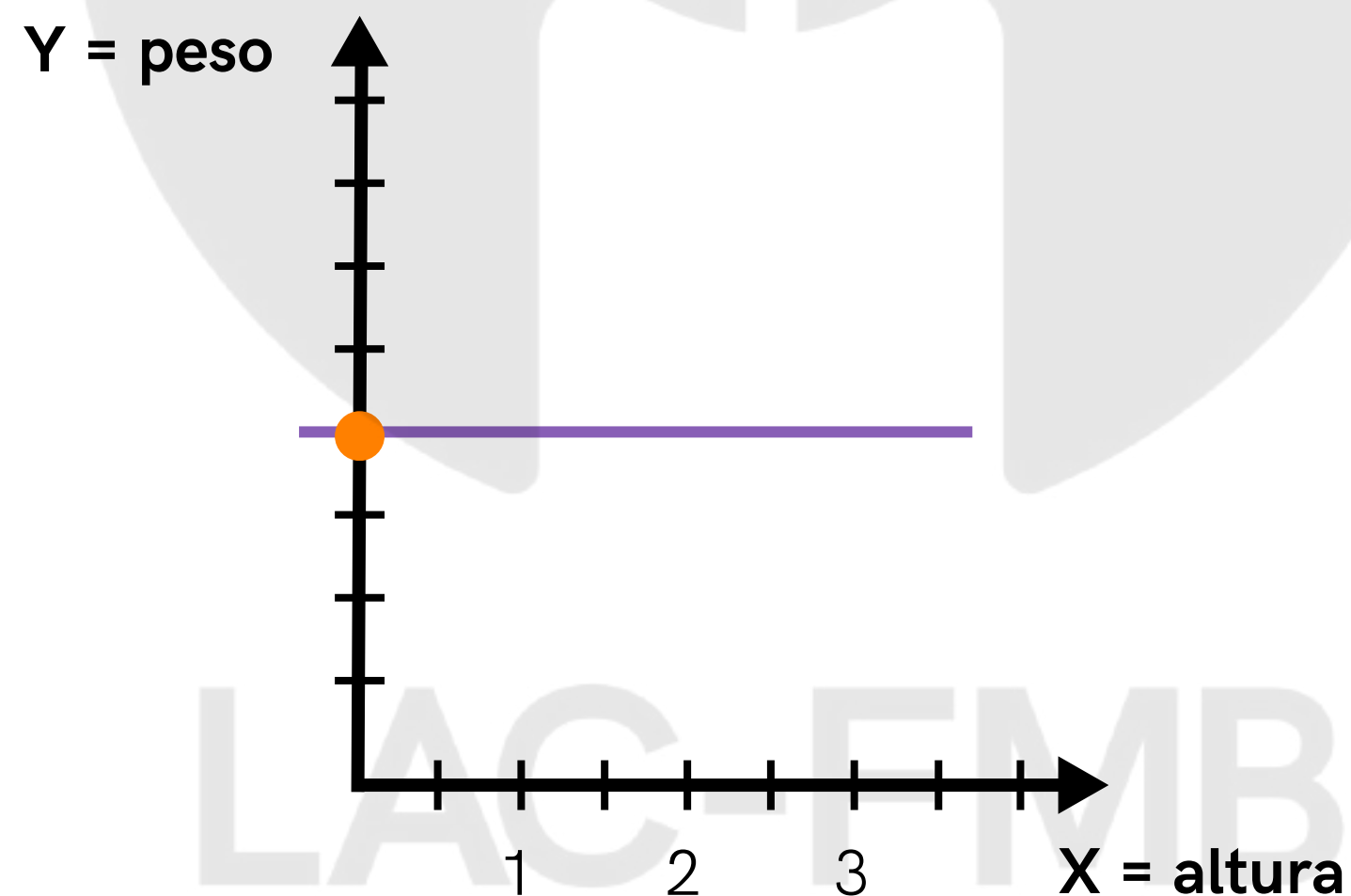
Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

Hipótese Nula

$$H_0: \beta_1 = 0$$

Altura **não** é uma variável relevante para o peso



LAC-FMB

Liga Acadêmica de Ciências



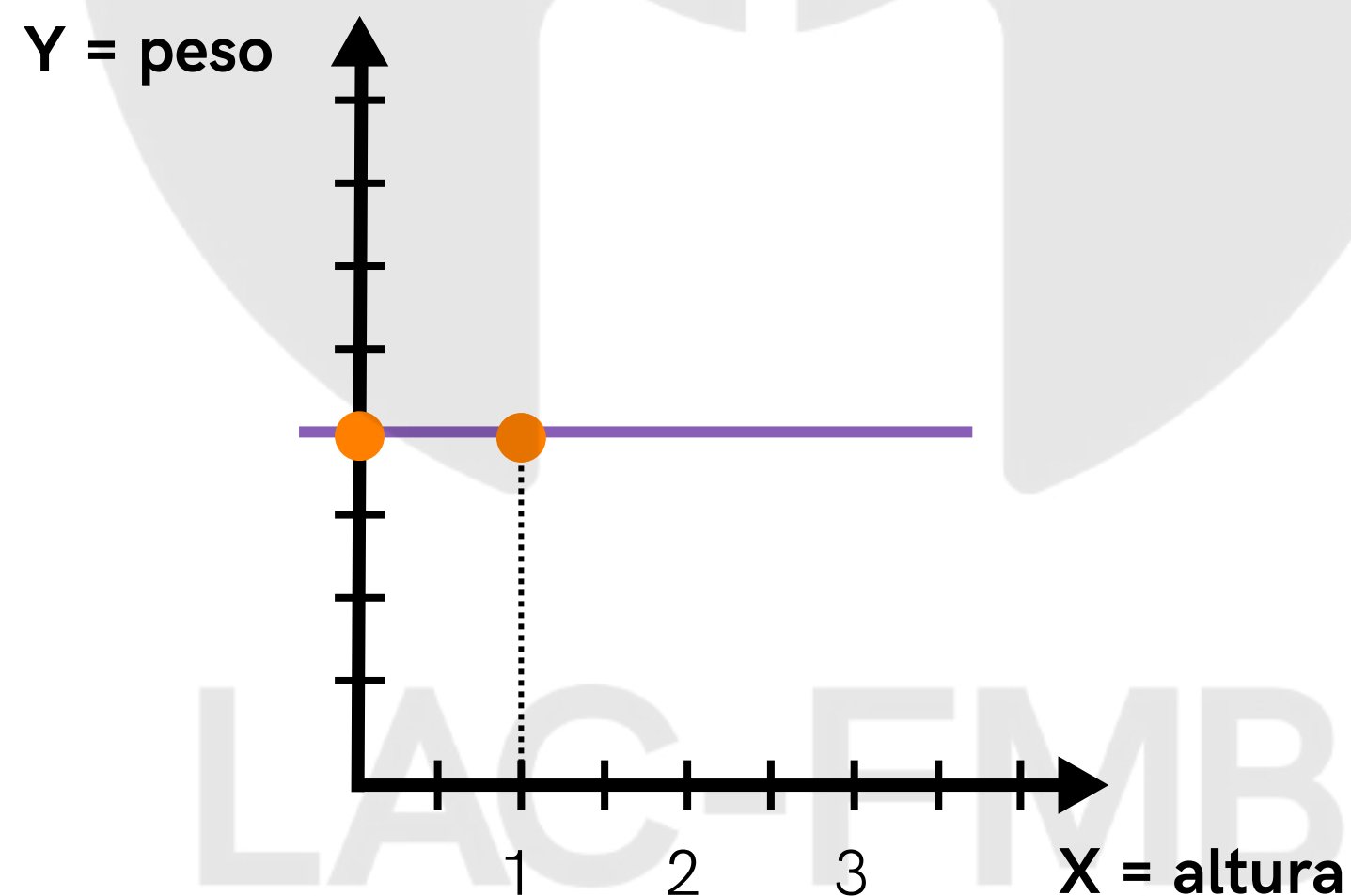
Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

Hipótese Nula

$$H_0: \beta_1 = 0$$

Altura **não** é uma variável relevante para o peso



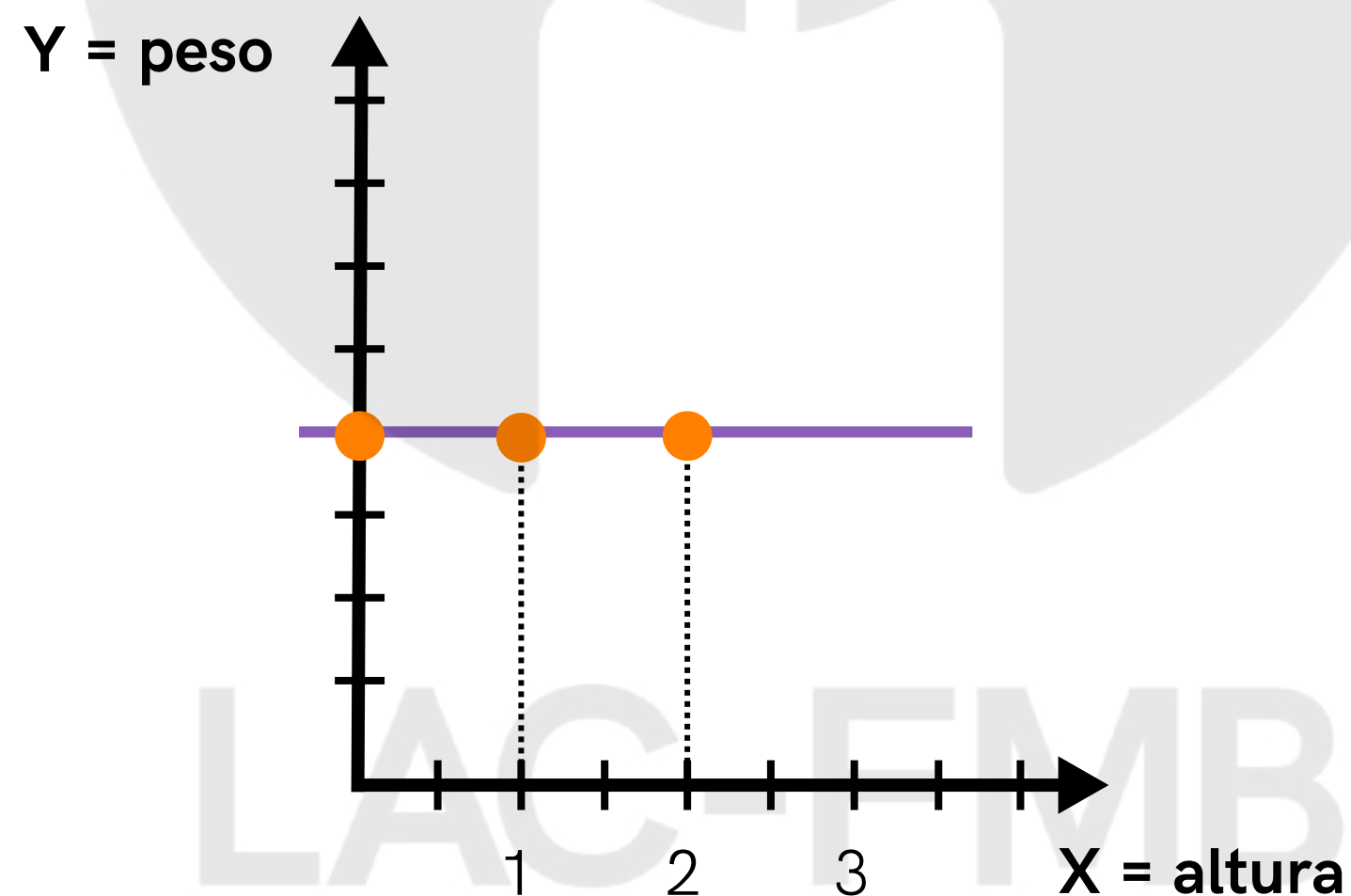
Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

Hipótese Nula

$$H_0: \beta_1 = 0$$

Altura **não** é uma variável relevante para o peso



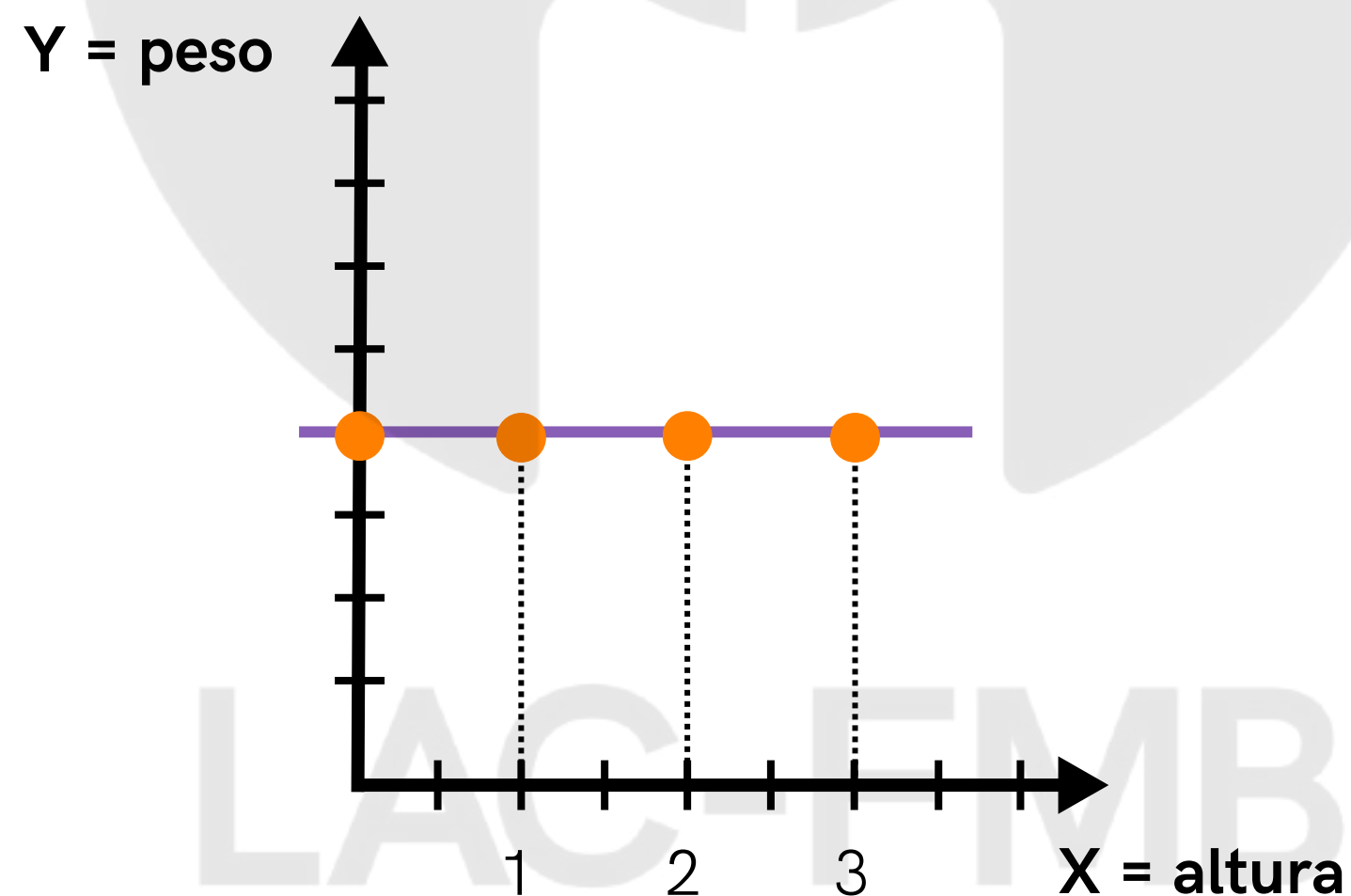
Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

Hipótese Nula

$$H_0: \beta_1 = 0$$

Altura **não** é uma variável relevante para o peso



Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

Hipótese Nula

$H_0: \beta_1 = 0$ Altura **não** é uma variável relevante para o peso

Hipótese Alternativa

$H_A: \beta_1 \neq 0$ Altura é uma variável relevante para o peso

LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

Rejeita a hipótese nula

Hipótese Alternativa

HA: $\beta_1 \neq 0$ Altura é uma variável relevante para o peso

LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$\text{Peso} = \beta_1 \text{ altura} + \beta_0$$

Rejeita a hipótese nula

Hipótese Alternativa

HA: $\beta_1 \neq 0$ Altura é uma variável relevante para o peso

Boa capacidade preditiva para esse modelo



Teste t

Quero saber se o índice de gordura corporal (IGC) de uma pessoa é afetado pelo número de vezes que ela consome fastfood na semana. Com 95% de intervalo de confiança e a partir dos dados de regressão, faça a análise:

$$\text{IGC} = \beta_1 \text{ fastfood} + \beta_0$$

LAC-FMB

Liga Acadêmica de Ciências



Teste t

$$\text{IGC} = \beta_1 \text{ fastfood} + \beta_0$$

$H_0: \beta_1 = 0$ ➡ O consumo de fastfood **NÃO É** relevante para explicar o IGC de um indivíduo

$H_A: \beta_1 \neq 0$ ➡ O consumo de fastfood **É** relevante para explicar o IGC de um indivíduo

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			

LAC-FMB

Liga Acadêmica de Ciências

FIELD, A., 2009



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} = \frac{\hat{\beta}_1 - \beta_{ho}}{\hat{\beta}_1}$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			

Estimativa para parâmetro

Valor da H0

$$t_{obs} = \frac{\hat{\beta}_1 - \beta_{ho}}{\hat{\beta}_1}$$

Erro Padrão associado a estimativa



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			

$$t_{obs} = \frac{\hat{\beta}_1 - \beta_{ho}}{\hat{\beta}_1} = \frac{0,0672 - 0}{0,0059} = 11,3705$$



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			

$$t_{obs} = \frac{\hat{\beta}_1 - \beta_{ho}}{\hat{\beta}_1} = \frac{0,0672 - 0}{0,0059} = 11,3705$$



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} = 11,3705$$

$$\text{Grau de Liberdade} = n - 2$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} = 11,3705$$

$$\text{Grau de Liberdade} = n - 2$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000

GRAU DE LIBERDADE?????

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1990 2013

Included observations: 24

Excluded observations: 0

Total observations: 24

Number of variables: 3

Number of lags: 0

Number of instruments: 0

R-squared: 0,62013

Std. Error

t-Statistic

Prob.

0,0367

6,103542

0,0000

0,0059

11,37056

0,0000

GRAU DE LIBERDADE EM 8 SLIDES

Grau de Liberdade = $n - 2$

LAC-FMB

Liga Acadêmica de Ciências



Grau de Liberdade

1

GL = N° amostra - Parâmetros estimados

Média

$$\frac{11,3705}{1} - 2 = 4$$

LAC-FMB

Liga Acadêmica de Ciências



Grau de Liberdade

2

GL = N° amostra - Parâmetros estimados

Média

2

•

•

•

•

=

4

LAC-FMB

Liga Acadêmica de Ciências



Grau de Liberdade

3

GL = N° amostra - Parâmetros estimados

Média

$$\frac{2}{11,3705} \cdot \frac{7}{11,3705} = 4$$

LAC-FMB

Liga Acadêmica de Ciências



Grau de Liberdade

4

GL = N° amostra - Parâmetros estimados

Média

$$\frac{2}{7} \cdot \frac{3}{1} = 4$$

LAC-FMB

Liga Acadêmica de Ciências



Grau de Liberdade

5

GL = N° amostra - Parâmetros estimados

Média

$$\frac{2}{7} \cdot \frac{3}{5} = 4$$

LAC-FMB

Liga Acadêmica de Ciências



Grau de Liberdade

6

GL = N° amostra - Parâmetros estimados

Média

$$\frac{2}{7} \cdot \frac{3}{5} \cdot X = 4$$

LAC-FMB

Liga Acadêmica de Ciências



Grau de Liberdade

7

GL = N° amostra - Parâmetros estimados

Média

$$(2 + 7 + 3 + 5 + X) = 4$$

5

LAC-FMB

Liga Acadêmica de Ciências



Grau de Liberdade

8

GL = N° amostra - Parâmetros estimados

Média

$$\frac{(2 + 7 + 3 + 5 + X)}{5} = 4 \quad | \quad X = 3$$

5

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} = 11,3705$$

Voltando para Regressão Linear Simples

Dependent Variable: IGC

Method: Least Squares

Sample:

Included in Model

Excluded from Model

Stepwise Selection Criteria

Model

Sum of Squares

Mean Square

F-Statistic

Prob. > F

Adjusted R-Squared

Change in Sum of Squares

Change in Mean Square

Change in F-Statistic

Change in Prob. > F

Change in Adjusted R-Squared

Change in Sum of Squares

Change in Mean Square

Change in F-Statistic

Change in Prob. > F

Change in Adjusted R-Squared

Std. Error

t-Statistic

Prob.

0,0367

6,103542

0,0000

0,0059

11,37056

0,0000

R-squared

2013

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} = 11,3705$$

$$\text{Grau de Liberdade} = n - 2$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} = 11,3705$$

$$\text{Grau de Liberdade} = n - 2 = 47 - 2 = 45$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			

LAC-FMB

Liga Acadêmica de Ciências



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} = 11,3705$$

$$\text{Grau de Liberdade} = n - 2 = 47 - 2 = 45$$

$$\text{Nível de significância} = 5\%$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			



Te

Mé

HC

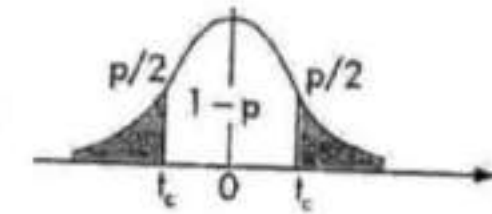
HA

tok

G

Distribuição t-Student

Distribuição t-Student
Corpo da tabela fornece os valores t_c tais que $P(-t_c < t < t_c) = 1-p$
Para $v > 120$, usar a aproximação normal



ob.

0000

0000

Graus de liberdade: v	p=90%	80%	70%	60%	50%	40%	30%	20%	10%	5%	4%	3%	2%	1%	0,5%	0,25%	0,20%	0,10%
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	15,895	21,205	31,821	63,657	127,321	254,647	318,309	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	4,849	5,643	6,965	9,925	14,089	19,962	22,327	31,599
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	3,482	3,896	4,541	5,841	7,453	9,465	10,215	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	2,999	3,298	3,747	4,604	5,598	6,758	7,173	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	2,757	3,003	3,365	4,032	4,773	5,604	5,893	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	2,612	2,829	3,143	3,707	4,317	4,981	5,208	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,517	2,715	2,998	3,499	4,029	4,595	4,785	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,449	2,634	2,896	3,355	3,833	4,334	4,501	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,398	2,574	2,821	3,250	3,690	4,146	4,297	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,359	2,527	2,764	3,169	3,581	4,005	4,144	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,328	2,491	2,718	3,106	3,497	3,895	4,025	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,303	2,461	2,681	3,055	3,428	3,807	3,930	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,282	2,436	2,650	3,012	3,372	3,735	3,852	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,264	2,415	2,624	2,977	3,326	3,675	3,787	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,249	2,397	2,602	2,947	3,286	3,624	3,733	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,235	2,382	2,583	2,921	3,252	3,581	3,686	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,224	2,368	2,567	2,898	3,222	3,543	3,646	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,214	2,356	2,552	2,878	3,197	3,510	3,610	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,205	2,346	2,539	2,861	3,174	3,481	3,579	3,883



Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} = 11,3705$$

$$\text{Grau de Liberdade} = n - 2 = 47 - 2 = 45$$

$$\text{Nível de significância} = 5\%$$

$$t_{crítico} = 2,014$$

Dependent Variable: IGC

Method: Least Squares

Sample: 1 47

Included Observations: 47

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0,2240	0,0367	6,103542	0,0000
FASTFOOD	0,0672	0,0059	11,37056	0,0000
R-squared	0,62013			



Teste t

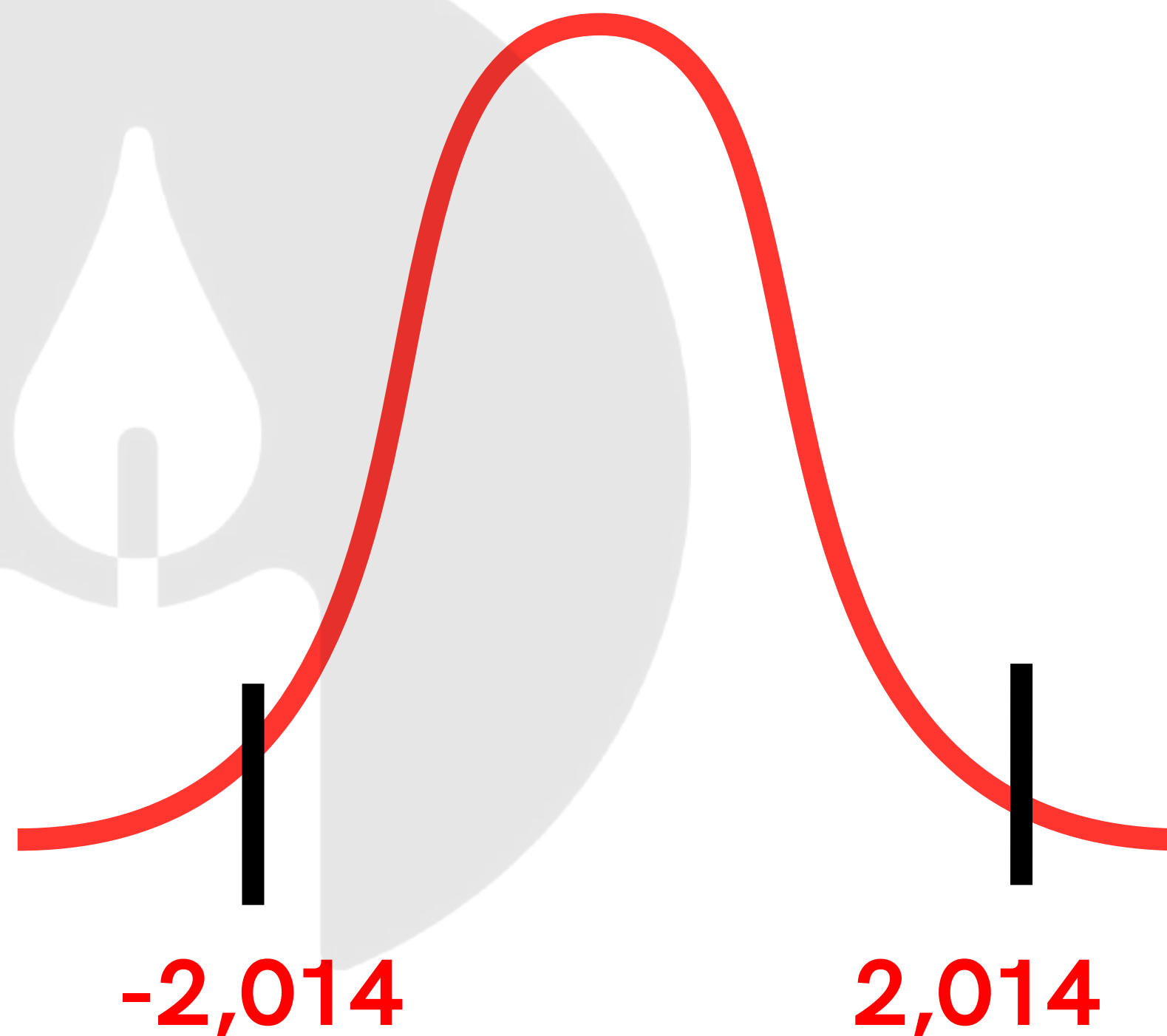
Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{\text{obs}} = 11,3705$$

$$t_{\text{crítico}} = 2,014$$



LAC-FMB

Liga Acadêmica de Ciências



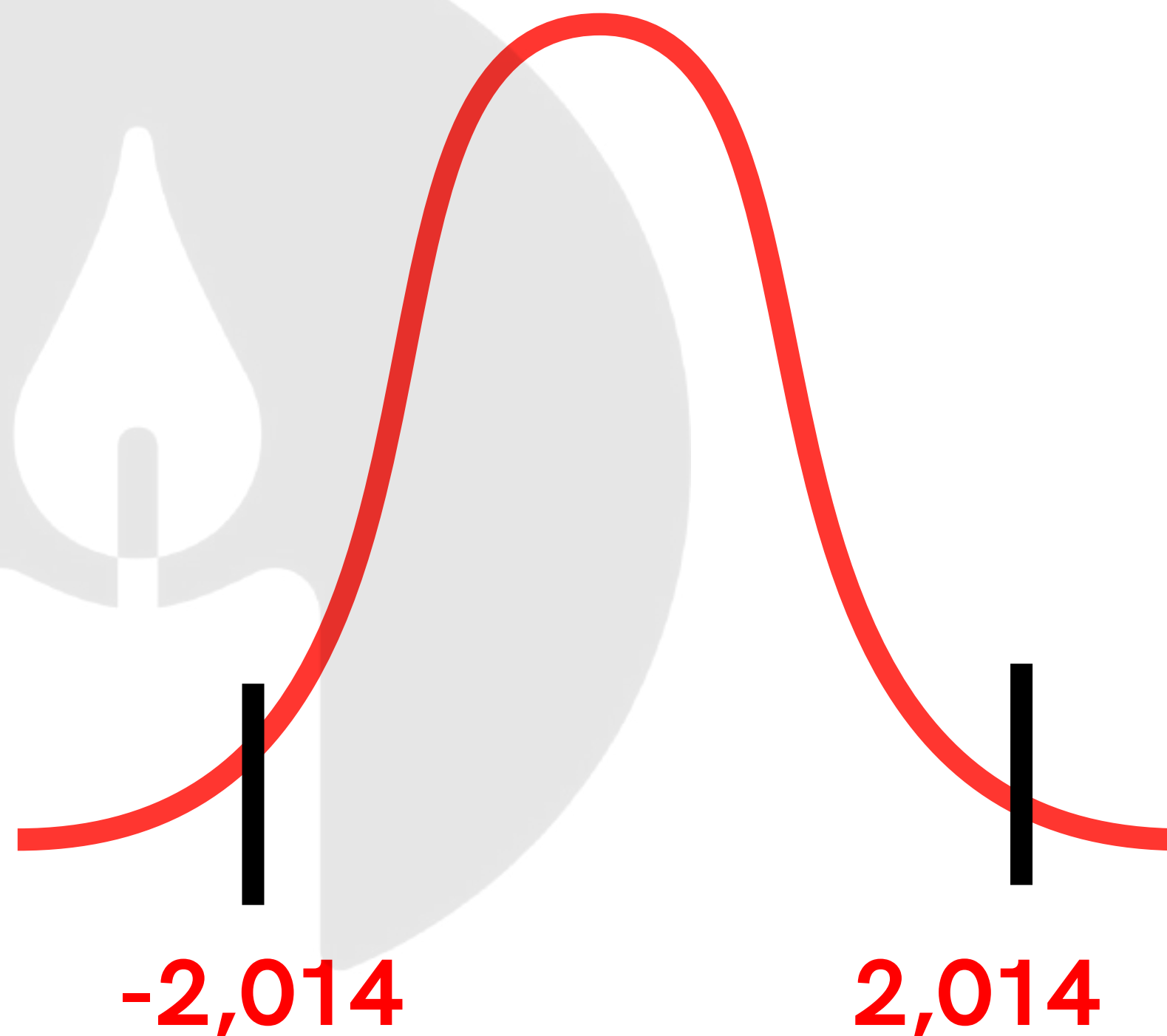
Teste t

Método t observado

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$t_{obs} > t_{crítico}$$



REJEITA HIPÓTESE NULA!



Teste t

Método t observado

Com 95% de confiança, há evidências de que o consumo de fastfood é relevante para explicar o IGC de um indivíduo!

REJEITA HIPÓTESE NULA!



PRÁTICA!

Software JAMOV



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Almeida	35.017	105.34
Pedro Conneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barreira	33.589	99.35
Thiago da dae	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Regressão Linear Simples

O consumo de bombom é um bom preditor para prever o peso das pessoas?

LAC-FMB

Liga Acadêmica de Ciências



NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Almeida	35.017	105.34
Pedro Conneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Matriz de correlações

Correlação

		Bombom (x/mês)	Peso
Bombom (x/mês)	R de Pearson	—	
	p-value	—	
Peso	R de Pearson	0.664	—
	p-value	< .001	—

LAC-FMB

Liga Acadêmica de Ciências



NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.980	106.00
Guustavo	27.483	86.18
Isaíel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Matriz de correlações



Qual o significado desse $R = 0.664$?

Correlação

		Bombom (x/mês)	Peso
Bombom (x/mês)	R de Pearson	—	
	p-value	—	
Peso	R de Pearson	0.664	—
	p-value	< .001	—



NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.980	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Almeida	35.017	105.34
Pedro Carneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Matriz de correlações

Que o consumo de bombom se correlaciona diretamente de maneira relativamente forte com a variável peso.

Correlação

		Bombom (x/mês)	Peso
Bombom (x/mês)	R de Pearson	—	
	p-value	—	
Peso	R de Pearson	0.664	—
	p-value	< .001	—

LAC-FMB

Liga Acadêmica de Ciências

JAMOV



Ajuste do Modelo

NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Carneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Medidas de Ajustamento do Modelo

Modelo	R	R ²	F	Teste ao Modelo Global		
				gl1	gl2	p
1	0.664	0.440	38.6	1	49	< .001

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001



NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.980	106.00
Guustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Carneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Ajuste do Modelo



Qual o significado desse $R^2 = 0.440$?

Medidas de Ajustamento do Modelo

Modelo	R	R ²	Teste ao Modelo Global			
			F	gl1	gl2	p
1	0.664	0.440	38.6	1	49	< .001

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001



NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Carneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Ajuste do Modelo

44% dos resultados da variável peso podem ser explicado pelo consumo de bombom.

Medidas de Ajustamento do Modelo

Modelo	R	R ²	Teste ao Modelo Global			
			F	gl1	gl2	p
1	0.664	0.440	38.6	1	49	< .001

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001



Ajuste do Modelo

NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Guustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conrado	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Medidas de Ajustamento do Modelo

Modelo	R	R ²	Teste ao Modelo Global			
			F	gl1	gl2	p
1	0.664	0.440	38.6	1	49	< .001

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001



NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.930	106.00
Guustavo	27.483	86.18
Isaíel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Carneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Ajuste do Modelo



Qual o significado do teste F com valor de $p < 0.001$?

Medidas de Ajustamento do Modelo

Modelo	R	R ²	Teste ao Modelo Global			
			F	gl1	gl2	p
1	0.664	0.440	38.6	1	49	< .001

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001



NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Guustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conrado	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Ajuste do Modelo

Que esse modelo é estatisticamente significativo.

Medidas de Ajustamento do Modelo

Modelo	R	R ²	Teste ao Modelo Global			
			F	gl1	gl2	p
1	0.664	0.440	38.6	1	49	< .001

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.980	106.00
Guustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Ajuste do Modelo



Qual o significado do teste de Shapiro-Wilk com valor de $p = 0.765$?

Verificação de Pressupostos

Teste à Normalidade (Shapiro-Wilk)

Estatística	p
0.985	0.765



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conrado	39.231	123.10
Pedro Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barlow	33.589	99.35
Thiago da Silva	39.187	108.00
Thiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Ajuste do Modelo

Valor de $p > 0.05$

A distribuição dos resíduos não se diferencia de uma maneira significativa de uma distribuição normal!

Verificação de Pressupostos

Teste à Normalidade (Shapiro-Wilk)

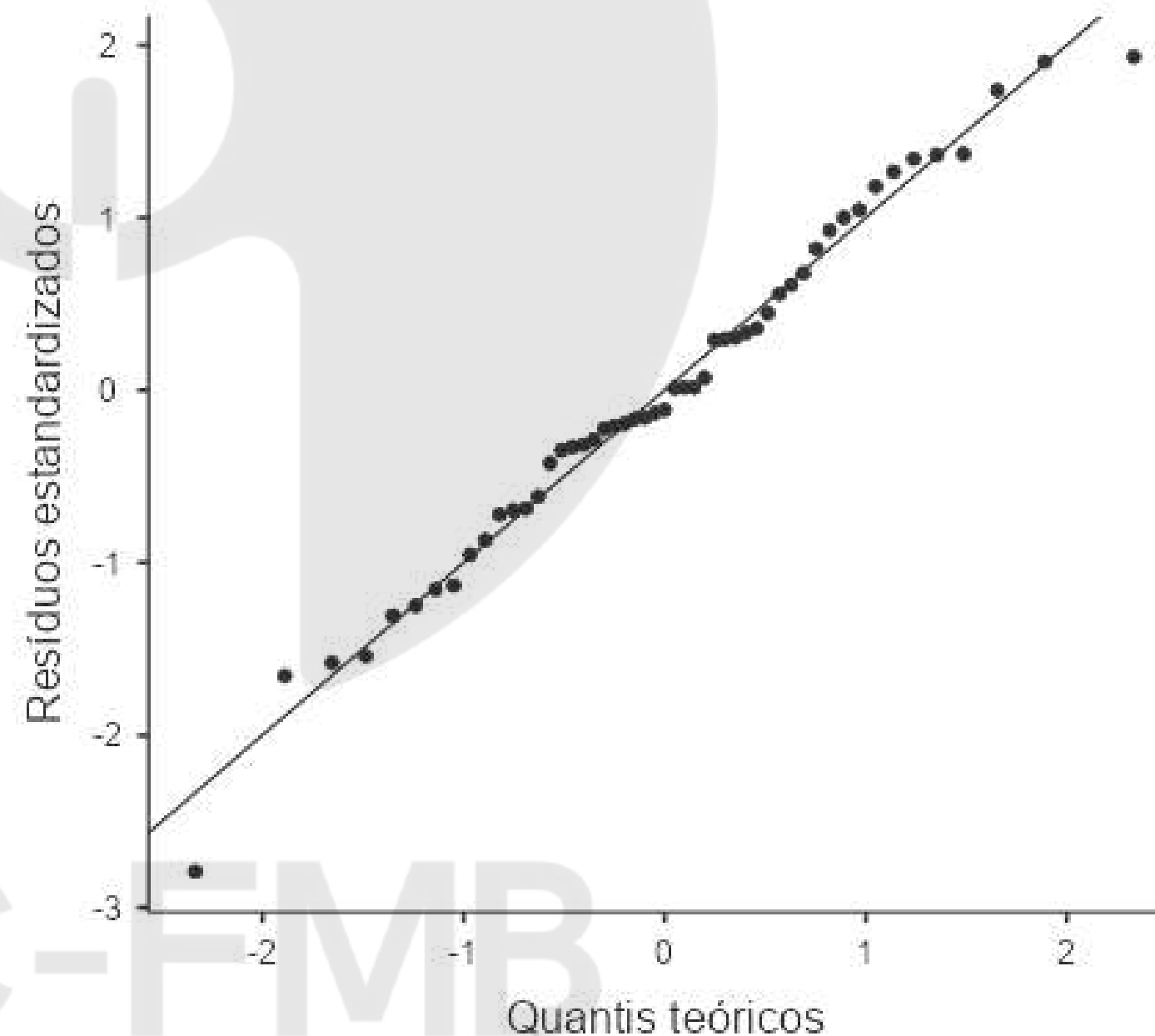
Estatística	p
0.985	0.765



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Israelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conrado	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barlow	33.589	99.35
Thiago da dae	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Ajuste do Modelo

Gráfico Q-Q



NOMES	BOMBOM(x/mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Almeida	35.017	105.34
Pedro Conneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Equação da regressão

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001

LAC-FMB

Liga Acadêmica de Ciências



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conrado	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Equação da regressão

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001

$$y = a + bx$$

$$y = 32.95 + 1.78x$$



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.960	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conrado	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Equação da regressão

Coeficientes do Modelo - Peso

Preditor	Estimativas	Erro-padrão	t	p
Intercepto	32.95	8.487	3.88	< .001
Bombom (x/mês)	1.78	0.286	6.21	< .001

$$y = a + bx$$

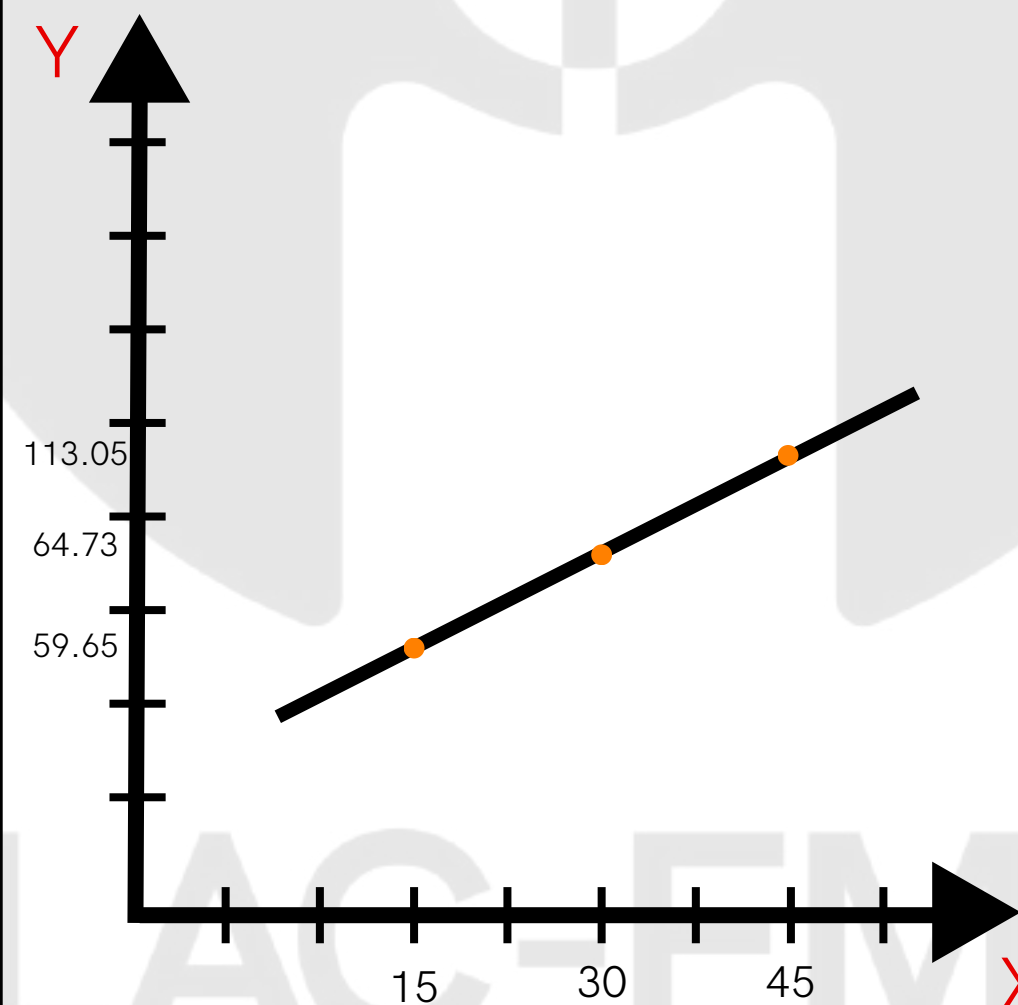
$$y = 32.95 + 1.78x$$



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.980	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conrado	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago the doc	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Equação da regressão

$$y = a + bx$$

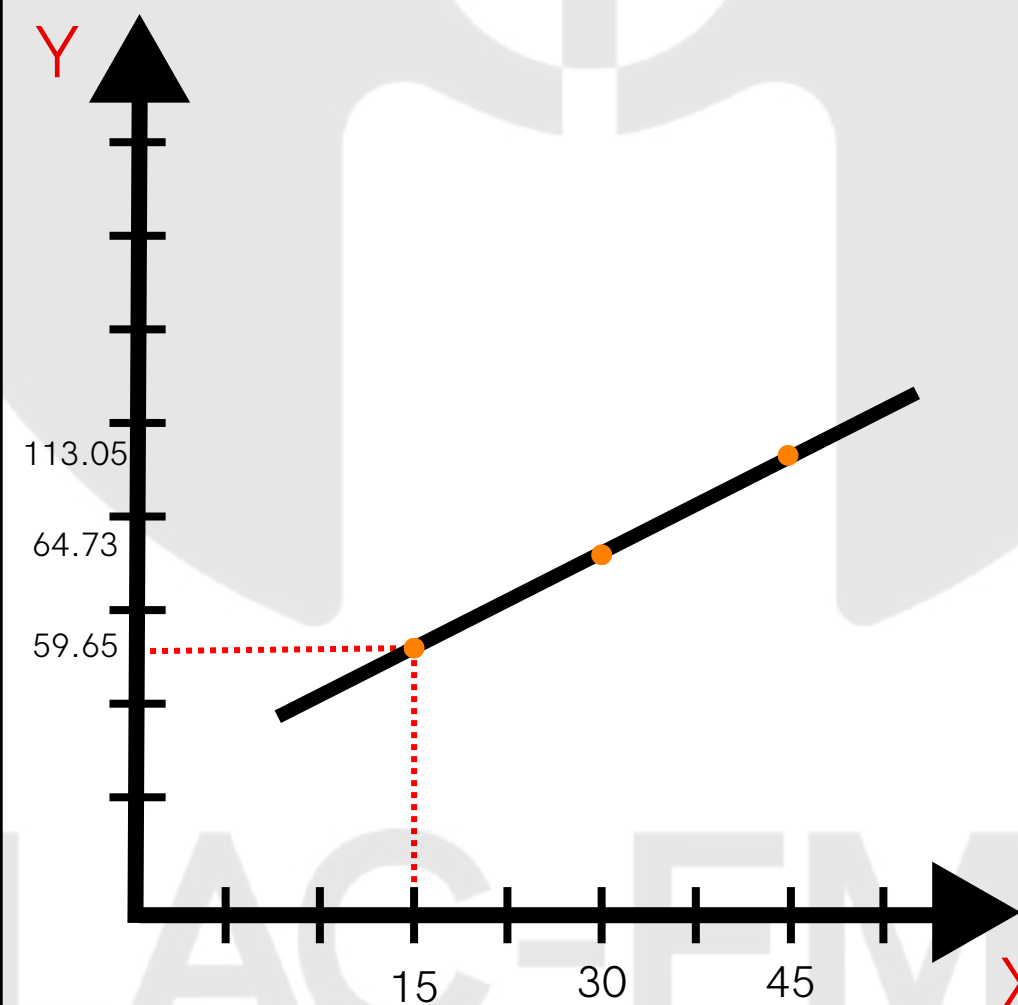


NOMES	BOMBOM(x em mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Dorcid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.980	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conrado	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago da dae	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Equação da regressão

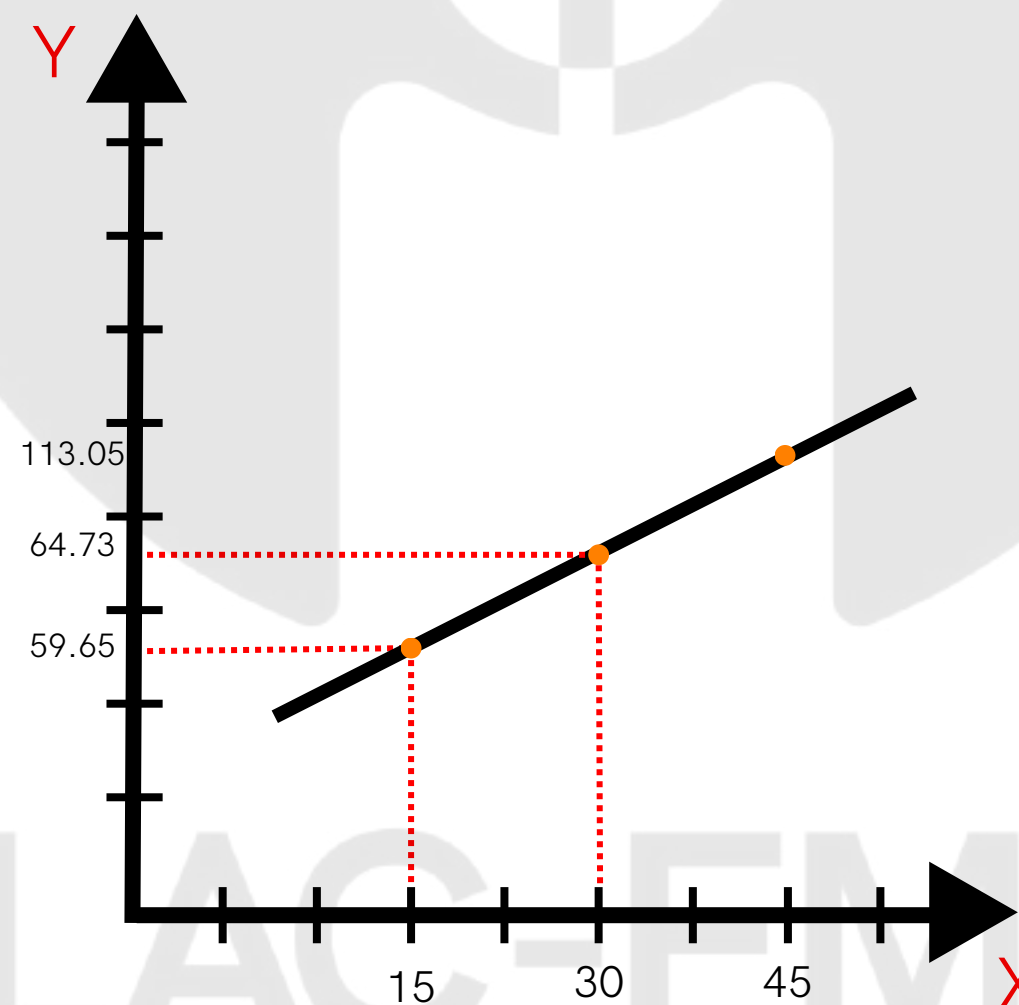
$$y = a + bx$$

$$59.65 = 32.95 + 1.78 \cdot 15$$



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Darvid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.980	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Conneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérvio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago da dae	39.187	108.00
Tiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Equação da regressão



$$y = a + bx$$

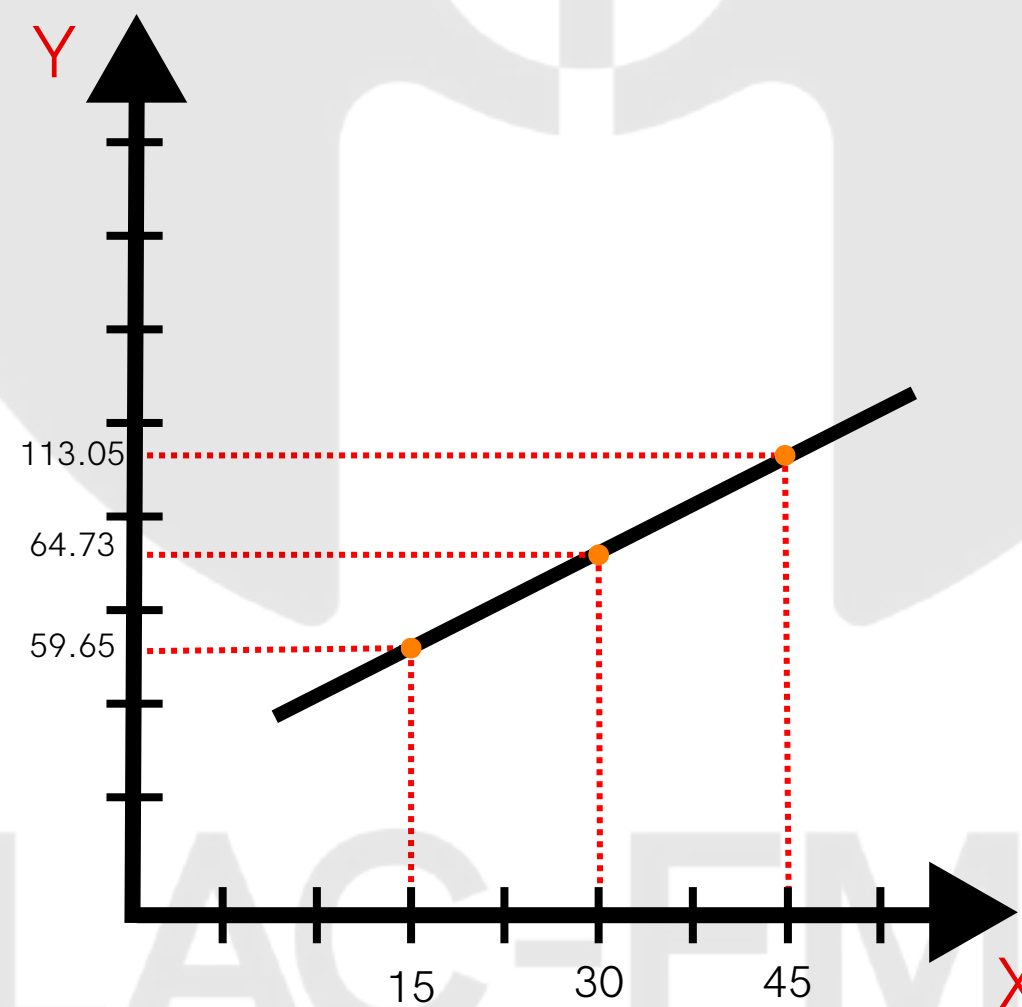
$$59.65 = 32.95 + 1.78 \cdot 15$$

$$64.73 = 32.95 + 1.78 \cdot 30$$



NOMES	BOMBOM(x1 mês)	PESO(kg)
Amanda	22.420	68.12
Ana Beatriz	22.377	73.00
Bea	37.618	113.64
Doniel	24.919	106.00
Darvid	31.894	88.00
Evelyn	18.924	68.00
Guilherme	40.980	106.00
Gustavo	27.483	86.18
Israel	40.018	95.00
Isabelly	24.173	73.99
Luiz F.	22.779	70.17
Luiz V.	24.660	89.00
Marina	19.340	53.19
Murilo	30.152	62.00
Pedro Alencar	35.017	105.34
Pedro Carneiro	39.231	123.10
Petelo Borges	22.203	67.26
Petela Lelis	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérgio	21.647	89.00
Thais Barilona	33.589	99.35
Thiago da Mac	39.187	108.00
Thiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R.	33.685	90.00
Carl Sagon	20.422	58.94

Equação da regressão



$$y = a + bx$$

$$59.65 = 32.95 + 1.78 \cdot 15$$

$$64.73 = 32.95 + 1.78 \cdot 30$$

$$113.05 = 32.95 + 1.78 \cdot 45$$



Regressão Linear Simples

Uma variável (X) é utilizada para
prever outra variável (Y)!

NOMES	BOMBA(X) (mls)	PESO(Y) (kg)
Amanda	22.420	68.12
Domini	24.919	101.00
David	31.894	88.00
Evelyn	16.924	68.00
Guilherme	40.920	101.00
Guatiero	27.483	86.18
Israel	40.018	95.00
Isabella	24.173	73.99
Marina	19.340	53.19
Muriel	30.152	62.00
Pedro Almeida	35.017	109.34
Pedro Gustavo	39.171	88.00
Patricia Dantas	22.173	68.00
Patricia Lela	23.307	55.00
Samira	25.158	86.00
Santiago	24.157	77.00
Sérgio	21.647	89.00
Thais Bastina	33.589	99.35
Thiago do sac	39.187	108.00
Thiago Silva	28.531	112.00
Vinicius Leal	15.323	50.00
Vinicius R	33.685	90.00
Carl Sagers	20.422	58.94



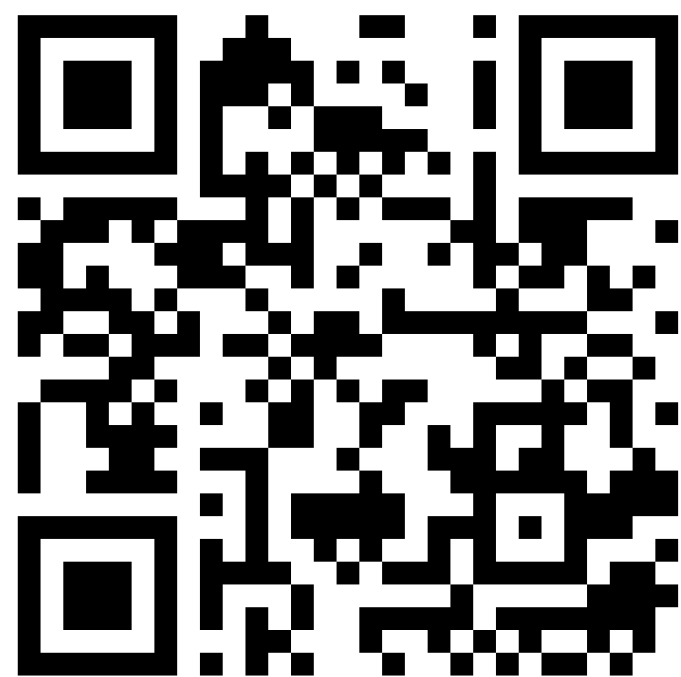
Referências



1. SENN, S. Francis Galton and regression to the mean. Significance, v. 8, n. 3, p. 124-126, 25 ago. 2011.
2. HULLEY, S. B. et al. Delineando a pesquisa clínica - 4.ed. [s.l.] Artmed Editora, 2015.
3. FIELD, A.; VIALILORÍ. Descobrimos a estatística usando o SPSS (2a. ed.). [s.l.] Grupo A - Bookman, 2000.
4. NETO, A. Bioestatística Sem Sgredos. 1a edição ed. [s.l: s.n.].
5. WHEELAN, C. Estatística. [s.l.] Editora Schwarcz - Companhia das Letras, 2016.
6. JAMOVÍ. jamovi - Stats. Open. Now. Disponível em: <<https://www.jamovi.org/>>.

Obrigado!

Avalie a sessão Encontre-nos aqui



@lacfmb



Academicast



academiafmb.com



linkme.bio/lacfmb



LAC-FMB

Liga Acadêmica de Ciências

LAC-F

Liga Acadêmica de