

Estatística e Probabilidade

Encontro presencial

Etapa 5 - Unidades 8 e 9

Detaques sobre:

Teste de Hipóteses e Regressão Linear

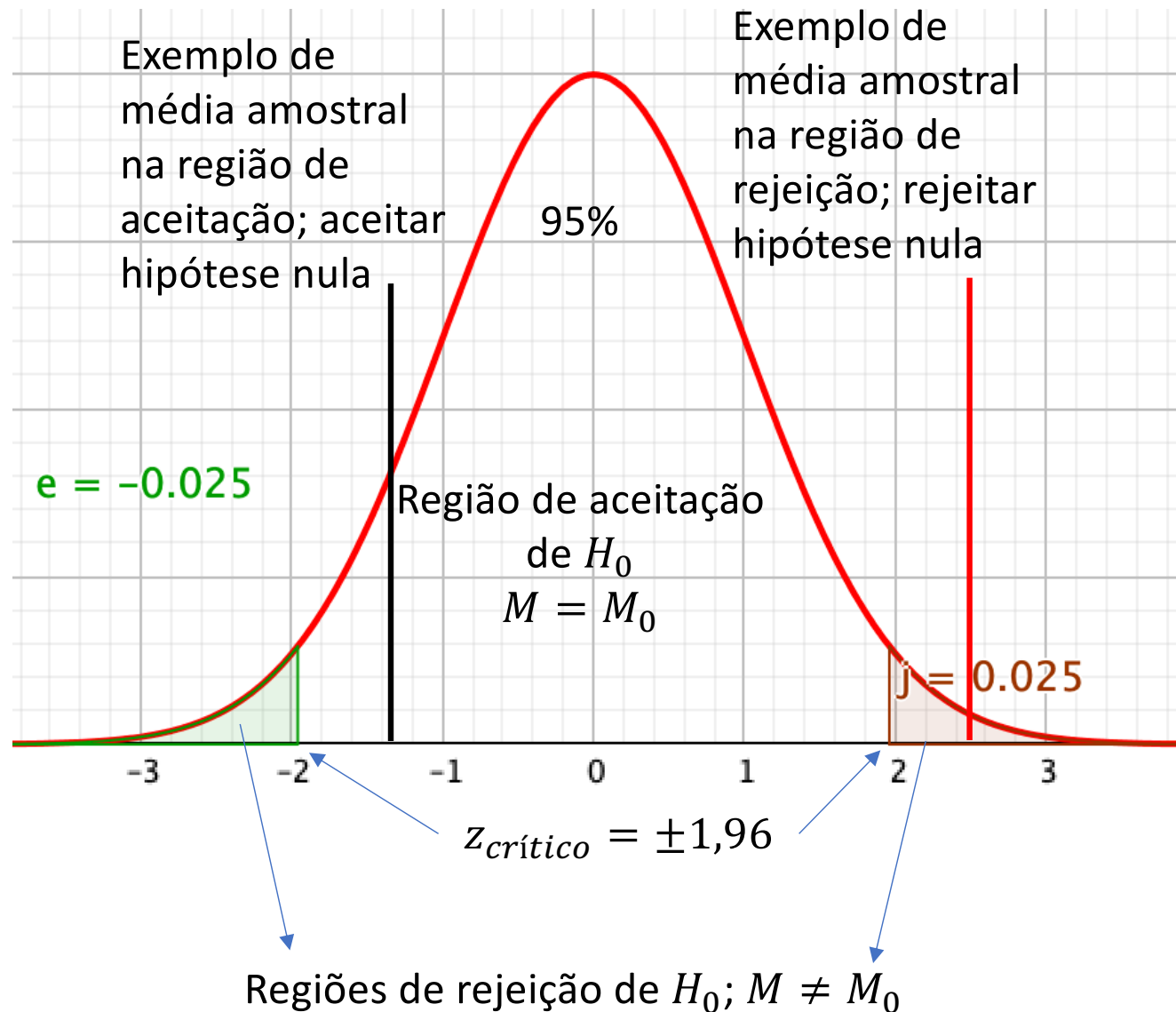
Prof. Simões

Teste de hipótese

- Consiste em testar uma hipótese sobre uma população, a partir dos dados de uma amostra
- Exemplo: um fabricante de máquinas de refrigerante automáticas afirma que uma máquina está calibrada para fornecer 340 ml de refrigerante por copo, com desvio padrão de 20 ml. O gerente de uma loja faz um teste com 35 copos, e obtém uma média de 350 ml. Testar a afirmação do fabricante, ao nível de significância de 5%, contra a afirmação de que o valor é diferente de 340 ml.

Testes para médias, bicaudal

- Exemplo de um teste de hipótese com 5% de significância



Passos:

- 1) Estabelecer as **hipóteses**
- 2) Com base na significância, calcular os **valores críticos** (máximo e mínimo aceitáveis)
- 3) Calcular o z da amostra, **valor de teste**
- 4) Comparar e **concluir**

Resolução

- Passo 1: estabelecer as hipóteses

$H_0 \Rightarrow M = 840 \text{ ml.}$ (hipótese nula; afirmação do fabricante está correta)

$H_1 \Rightarrow M \neq 840 \text{ ml.}$ (hipótese alternativa; afirmação do fabricante está incorreta)

- Passo 2: calcular os valores críticos $z_{crítico}$ (limites máx. e mín.) para significância de 5%

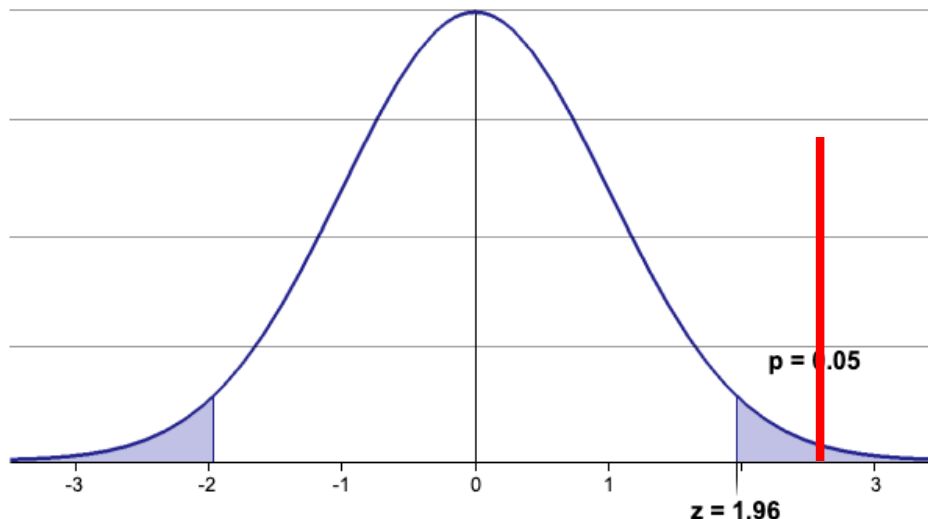
$$\frac{0,05}{2} = 0,025 \Rightarrow \text{corresponde a } 0,475 \text{ na tabela do } z$$

$$z_{crítico} = \pm 1,96$$

- Passo 3: calcular o valor de teste z_{teste}

$$z_{teste} = \frac{m - M_0}{\frac{\sigma}{\sqrt{n}}} = \frac{350 - 340}{\frac{20}{\sqrt{35}}} \Rightarrow z_{teste} = 2,50$$

- Passo 4: comparar e concluir



Conclusão: o valor de teste está na região de rejeição. Assim deve-se rejeitar a hipótese nula e aceitar a hipótese alternativa: a máquina está desregulada.

Possíveis variações na fórmula do z_{teste}

- Com desvio padrão **populacional**

$$z_{teste} = \frac{m - M_0}{\frac{\sigma}{\sqrt{n}}}$$

- Com desvio padrão **amostral**

$$n > 30$$



$$z_{teste} = \frac{m - M_0}{\frac{s}{\sqrt{n}}}$$

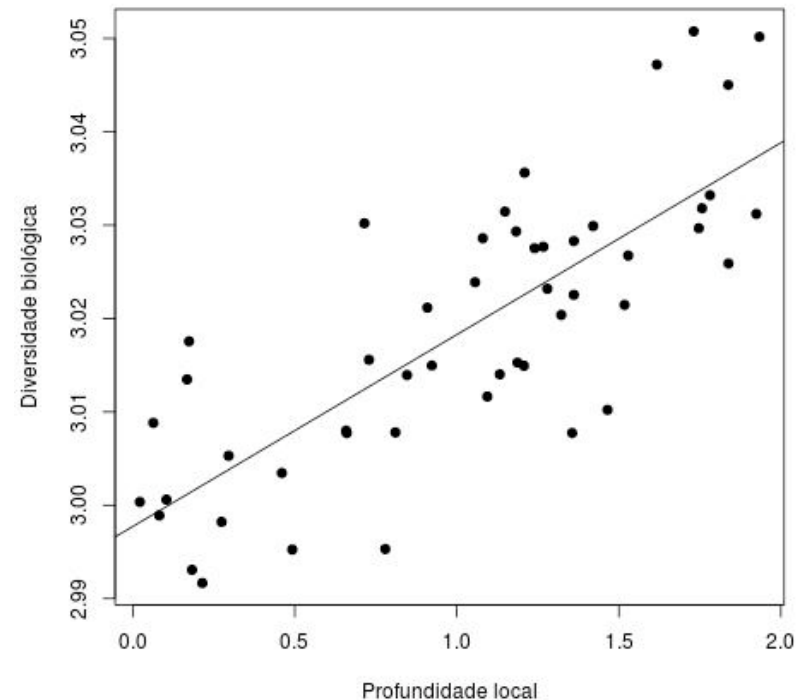
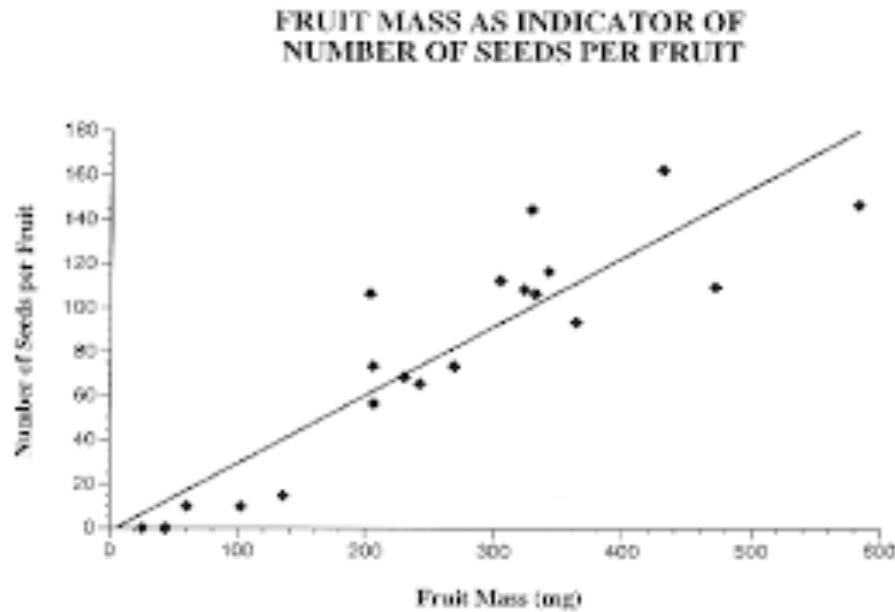
$$n \leq 30$$



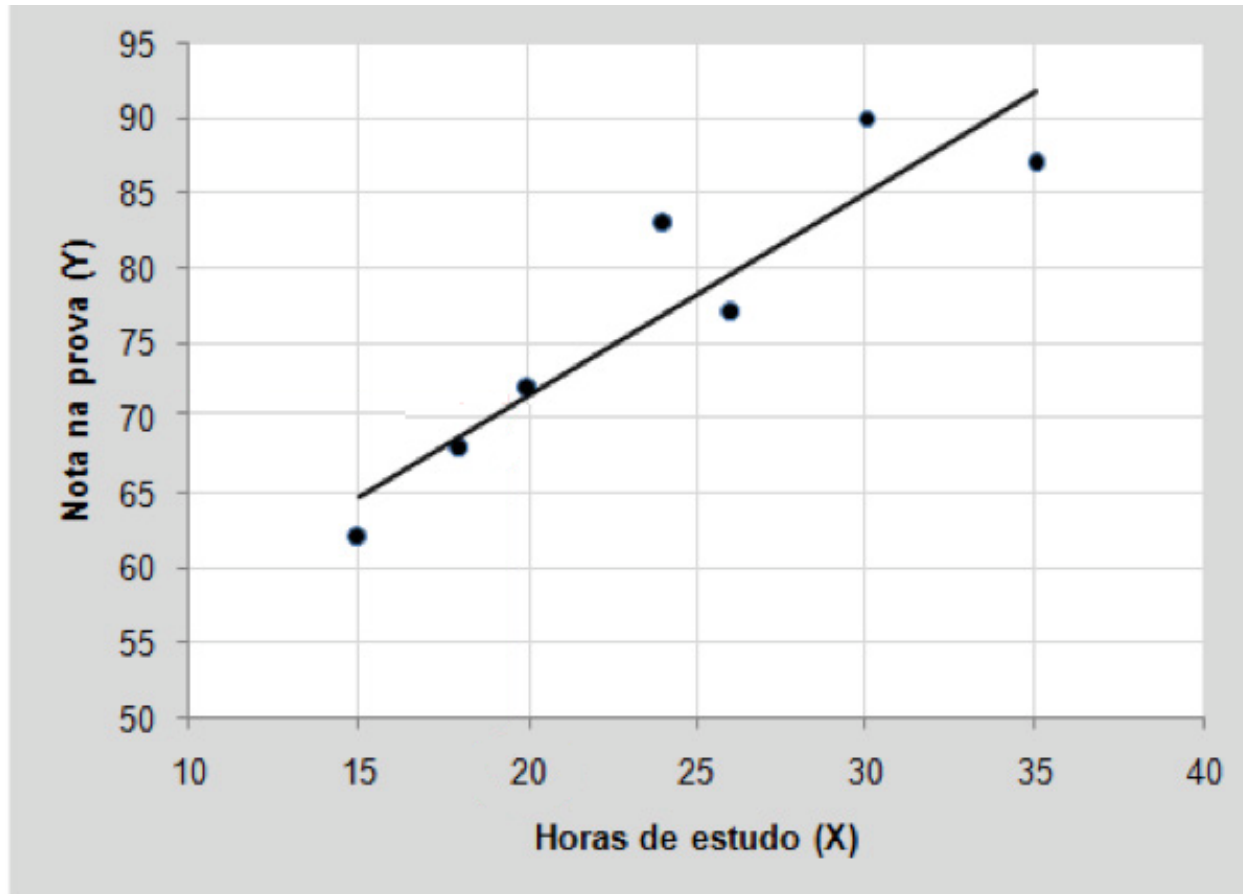
$$t_{teste} = \frac{m - M_0}{\frac{s}{\sqrt{n}}}$$

Regressão linear

- Consiste em um conjunto de cálculos que permitem determinar se duas variáveis estão relacionadas, e quão “forte” é essa relação



Regressão linear: coeficiente de correlação



$$s_x = \sqrt{\frac{\sum (x_1 - \bar{x})^2}{n - 1}}$$

$$s_y = \sqrt{\frac{\sum (y_1 - \bar{y})^2}{n - 1}}$$

Indica quão dependente é uma variável da outra.

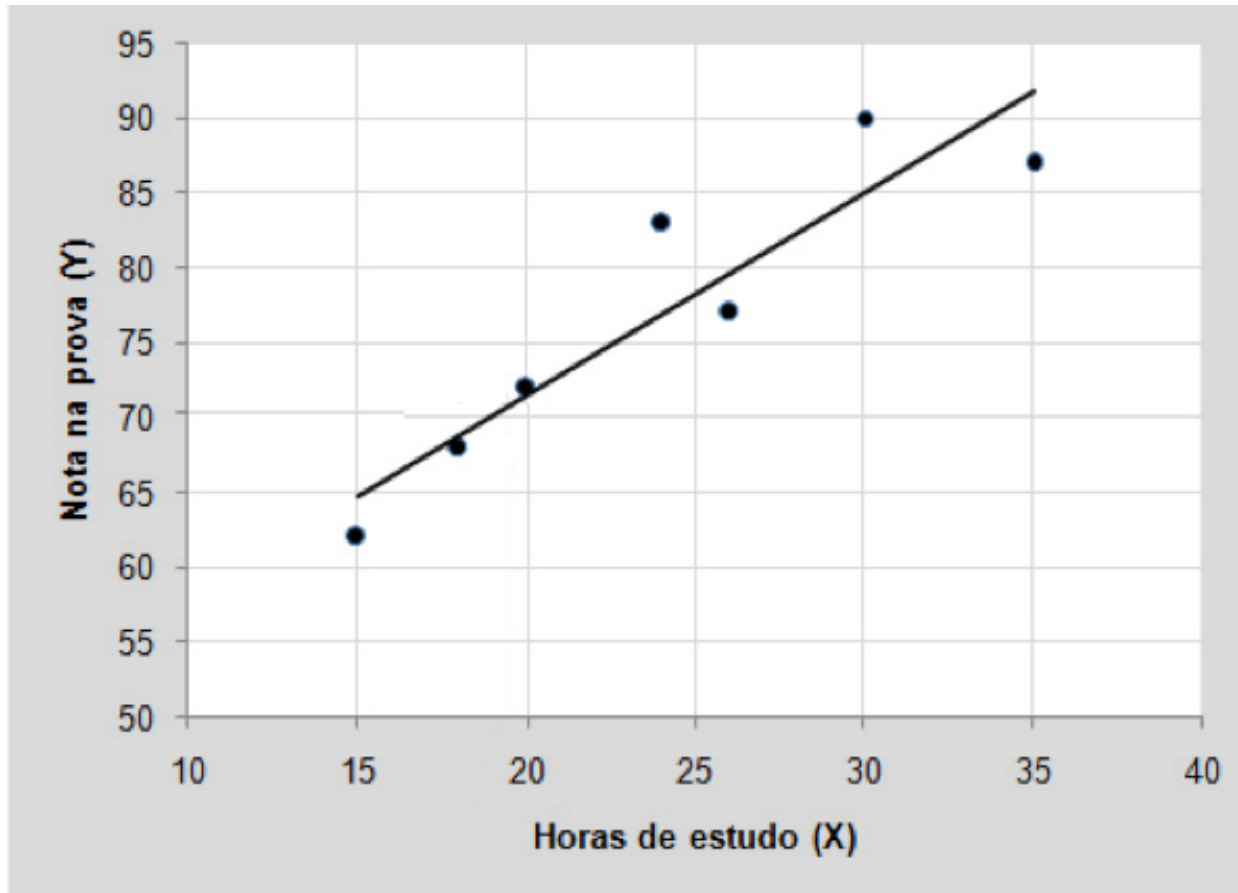
$$r = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{s_x \cdot s_y \cdot (n - 1)}$$

Regressão linear: coeficiente de correlação

Estudante	Horas de estudo (x_i)	Nota da prova (y_i)	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x}) \cdot (y_i - \bar{y})$
1	20	72	-4	-5	20
2	15	62	-9	-15	135
3	35	87	11	10	110
4	26	77	2	0	0
5	30	90	6	13	78
6	24	83	0	6	0
7	18	68	-6	-9	54
$\bar{x} = 24$		$\bar{y} = 77$			$\sum (x_i - \bar{x}) \cdot (y_i - \bar{y}) = 397$
$s_x = 7,0$		$s_y = 10,3$			

$$r = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{s_x \cdot s_y \cdot (n - 1)} = \frac{397}{7,0 \cdot 10,3 \cdot 6} = 0,92 \Rightarrow \text{correlação forte e positiva}$$

Regressão linear simples: equação da reta



$$y = A + Bx$$

$$B = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum x_i^2 - n \cdot \bar{x}^2}$$

$$A = \bar{y} - B \cdot \bar{x}$$

Regressão linear simples: equação da reta, exemplo

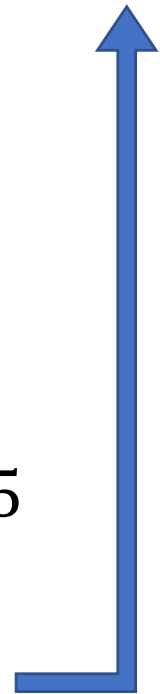
Estudante	Horas de estudo (x_i)	Nota da prova (y_i)	x_i^2	$x_i \cdot y_i$
1	20	72	400	1440
2	15	62	225	930
3	35	87	1225	3045
4	26	77	676	2002
5	30	90	900	2700
6	24	83	576	1992
7	18	68	324	1224
Média =>	24	77		
DP=>	7,0	10,3		
Σ =>			4326	13333

$$y = A + Bx$$

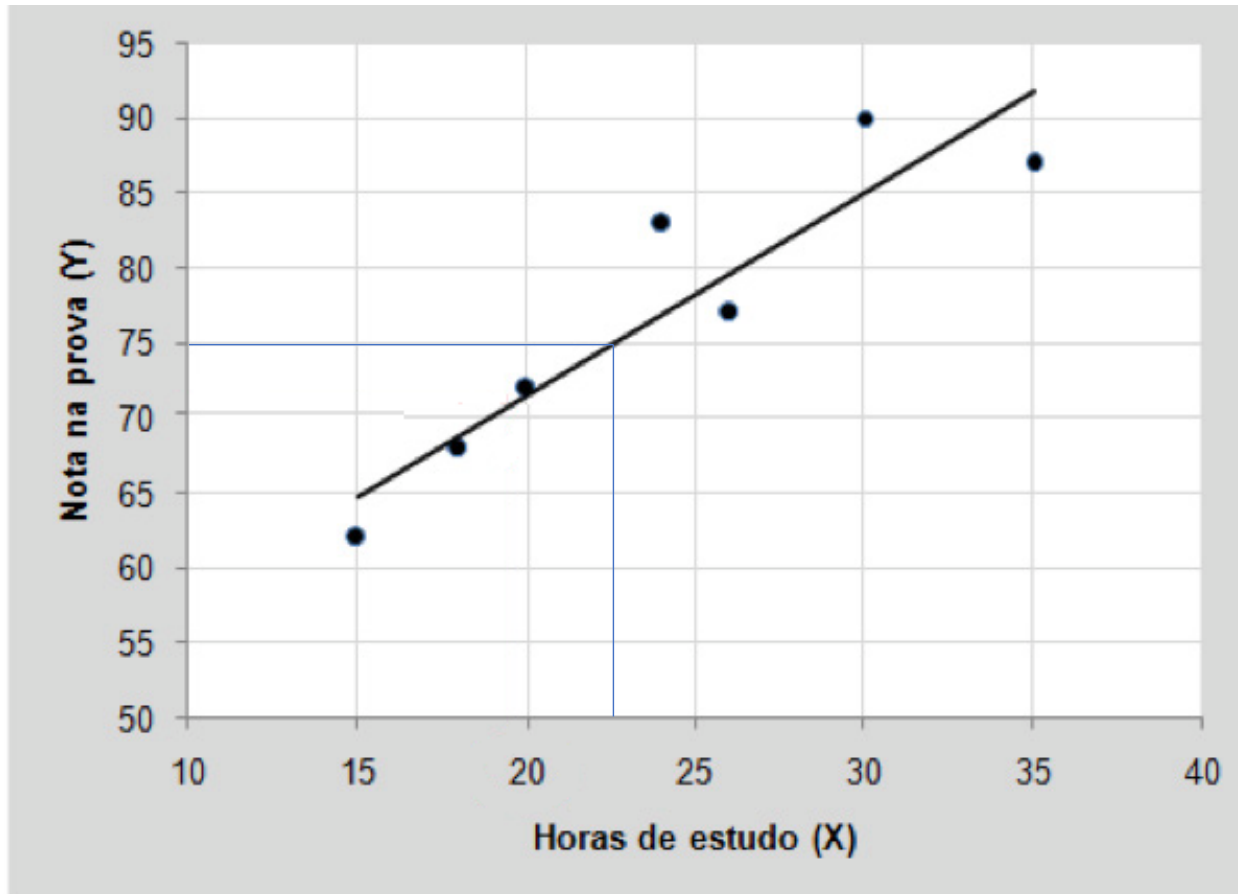
$$y = 44,6 + 1,35x$$

$$B = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum x_i^2 - n \cdot \bar{x}^2} = \frac{13333 - 7 \cdot 24 \cdot 77}{4326 - 7 \cdot 24^2} \Rightarrow B = 1,35$$

$$A = \bar{y} - B \cdot \bar{x} = 77 - 1,35 \cdot 24 \Rightarrow A = 44,6$$



Regressão linear simples: equação da reta



Com isso, podemos, por exemplo, interpolar valores. Por exemplo, que nota tiraria um aluno que estudou 22,5 horas?

$$y = 44,6 + 1,35x$$

$$y = 44,6 + 1,35 \cdot 22,5$$

$$y = 75$$

Regressão linear com a Cassio fx ou HP10s



- **MODE 3 1:** Vai para o modo de regressão linear
- **SHIFT CLR 1 = :** limpa os registradores estatísticos
- **20 , 72 M+ :** armazena o primeiro par de dados (atenção, não é o ponto decimal)
- 15 , 62 M+ : etc até n=7
- **SHIFT S-VAR ►► 1 = :** A
- **SHIFT S-VAR ►► 2 = :** B
- **SHIFT S-VAR ►► 3 = :** r

$$y = A + Bx$$