

Física Geral - Laboratório (2014/1)

Organização e descrição de dados



Física Geral - 2014/I

Bibliografia:



“Estimativas e Erros em Experimentos de Física”
(EdUERJ)

Dados e medidas

Dados: Valores ou qualificações de atributos dos elementos de um conjunto

Medidas: Dados numéricos associados a grandezas que descrevem um fenômeno ou sistema físico

Dados brutos

Exemplo de conjunto de dados:

1) Valores das idades de um grupo de estudantes de Física Geral

Estudante 1: 18 anos

Estudante 2: 19 anos

Estudante 3: 18 anos

Unidade: Anos

Dados brutos

Exemplo de conjunto de dados:

2) Valores das massas de um grupo de estudantes de Física Geral

Estudante 1: 60,2 Kg

Estudante 2: 72,4 Kg

Estudante 3: 65,6 Kg

Unidade: Quilograma (Kg)

Dados brutos

Exemplo de conjunto de dados:

3) Valores das alturas de um grupo de estudantes de Física Geral

Estudante 1: 172 cm

Estudante 2: 168 cm

Estudante 3: 180 cm

Unidade: Centímetro (cm)

Dados e medidas

Representação do conjunto de dados:

Idades dos estudantes = {18; 19; 18} (anos)

Massas dos estudantes = {60,2; 72,4; 65,6} (Kg)

Alturas dos estudantes = {172; 168; 180} (cm)

Em geral:

$\{x_1, x_2, x_3, \dots, x_N\} = \{\text{valor n}^\circ 1, \text{valor n}^\circ 2, \text{valor n}^\circ 3, \dots, \text{valor n}^\circ \text{'N'}\}$

Dados e medidas

Outros exemplos:

Medidas do comprimento de uma mesa:

$\{150,3; 152,0; 150,4; 151,8\}$ (cm)

Medidas de temperatura de uma sala:

$\{29,3; 28,6; 30,4\}$ ($^{\circ}\text{C}$)

Medidas da tensão da rede elétrica:

$\{115,2; 124,5; 128,3; 121,1\}$ (V)

Tipo sanguíneo dos estudantes de FG:

$\{\text{'O-'; 'A-'; 'O+'}\}$

Organizando um conjunto de dados: Tabelas

Tabelas: arranjos, ordenados ou não, de dados

Estudante de FG	Idade (anos)	Massa (Kg)	Altura (cm)
1	18	60,2	172
2	19	72,4	168
3	18	65,6	180

Mesa	Comprimento (cm)
1	150,3
2	152,0
3	150,4
4	151,8

Organizando um conjunto de dados: Classes e Histogramas

Classes: Intervalos em que um conjunto de dados é agrupado

Histogramas: Número de ocorrências ou frequência das classes de agrupamento de um conjunto de dados

Organizando um conjunto de dados: Classes e Histogramas

Classes: Intervalos em que um conjunto de dados é agrupado

Histogramas: Número de ocorrências ou frequência das classes de agrupamento de um conjunto de dados

- Passo nº 1: Definir classes de agrupamento de dados
- Passo nº 2: Computar frequências para cada classe de dados
- Passo nº 3: Representar graficamente frequências em forma de histogramas

Organizando um conjunto de dados: Classes e Histogramas

Classes: Intervalos em que um conjunto de dados é agrupado

Histogramas: Número de ocorrências ou frequência das classes de agrupamento de um conjunto de dados

- Passo n° 1: Definir classes de agrupamento de dados
- Passo n° 2: Computar frequências para cada classe de dados
- Passo n° 3: Representar graficamente frequências em forma de histogramas

Que tamanho de intervalo devemos usar para cada classe de frequência?

Organizando um conjunto de dados: Classes e Histogramas

Classes: Intervalos em que um conjunto de dados é agrupado

Histogramas: Número de ocorrências ou frequência das classes de agrupamento de um conjunto de dados

Exemplo:

Um conjunto maior de dados (idades): {10, 7, 10, 11, 10, 15, 8, 12, 14, 9, 6, 8, 7, 14, 10, 10, 7, 12, 12, 9, 13, 10, 9, 8} (anos)

24 elementos



Organizando um conjunto de dados: Classes e Histogramas

Um conjunto maior de dados (idades): {10, 7, 10, 11, 10, 15, 8, 12, 14, 9, 6, 8, 7, 14, 10, 10, 7, 12, 12, 9, 13, 10, 9, 8} (anos)

Organizando um conjunto de dados: Classes e Histogramas

Um conjunto maior de dados (idades): {10, 7, 10, 11, 10, 15, 8, 12, 14, 9, 6, 8, 7, 14, 10, 10, 7, 12, 12, 9, 13, 10, 9, 8} (anos)

Escolha 1:

Classe de idades (anos)	Frequências
6	1
7	3
8	3
9	3
10	6
11	1
12	3
13	1
14	2
15	1

Organizando um conjunto de dados: Classes e Histogramas

Um conjunto maior de dados (idades): {10, 7, 10, 11, 10, 15, 8, 12, 14, 9, 6, 8, 7, 14, 10, 10, 7, 12, 12, 9, 13, 10, 9, 8} (anos)

Escolha 1:

Classe de idades (anos)	Frequências
6	1
7	3
8	3
9	3
10	6
11	1
12	3
13	1
14	2
15	1

Escolha 2:

Classe de idades (anos)	Frequência
[6 - 8)	4
[8 - 10)	6
[10 - 12)	7
[12 - 14)	4
[14 - 16)	3

Organizando um conjunto de dados: Classes e Histogramas

Conjunto de idades:

{10, 7, 10, 11, 10, 15, 8, 12, 14,
9, 6, 8, 7, 14, 10, 10, 7, 12, 12,
9, 13, 10, 9, 8} (anos)

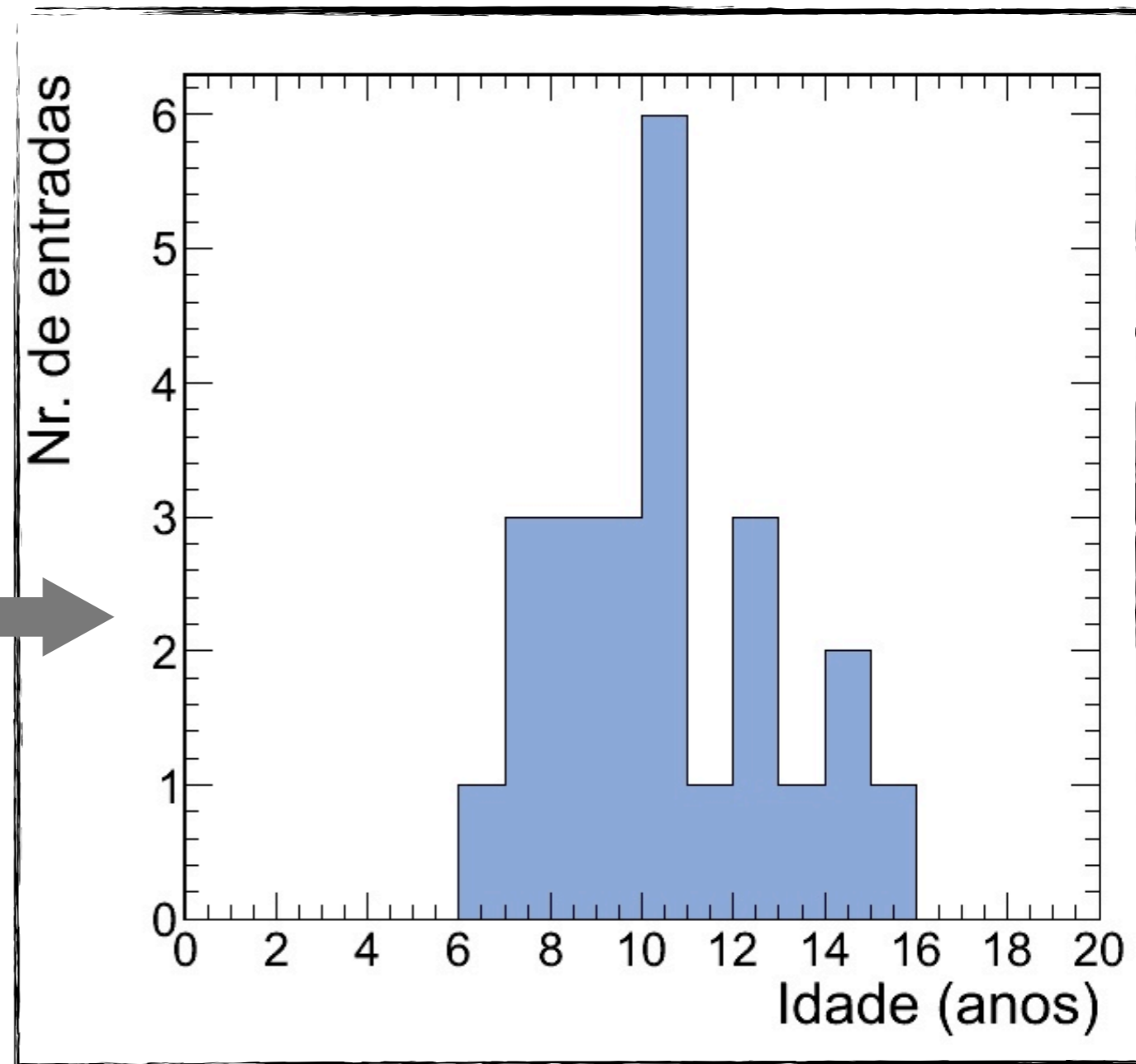
Classe de idades (anos)	Frequências
6	1
7	3
8	3
9	3
10	6
11	1
12	3
13	1
14	2
15	1

Organizando um conjunto de dados: Classes e Histogramas

Conjunto de idades:

{10, 7, 10, 11, 10, 15, 8, 12, 14, 9, 6, 8, 7, 14, 10, 10, 7, 12, 12, 9, 13, 10, 9, 8} (anos)

Classe de idades (anos)	Frequências
6	1
7	3
8	3
9	3
10	6
11	1
12	3
13	1
14	2
15	1



Organizando um conjunto de dados: Classes e Histogramas

Conjunto de idades:

{10, 7, 10, 11, 10, 15, 8, 12, 14, 9, 6, 8, 7, 14, 10, 10, 7, 12, 12, 9, 13, 10, 9, 8} (anos)

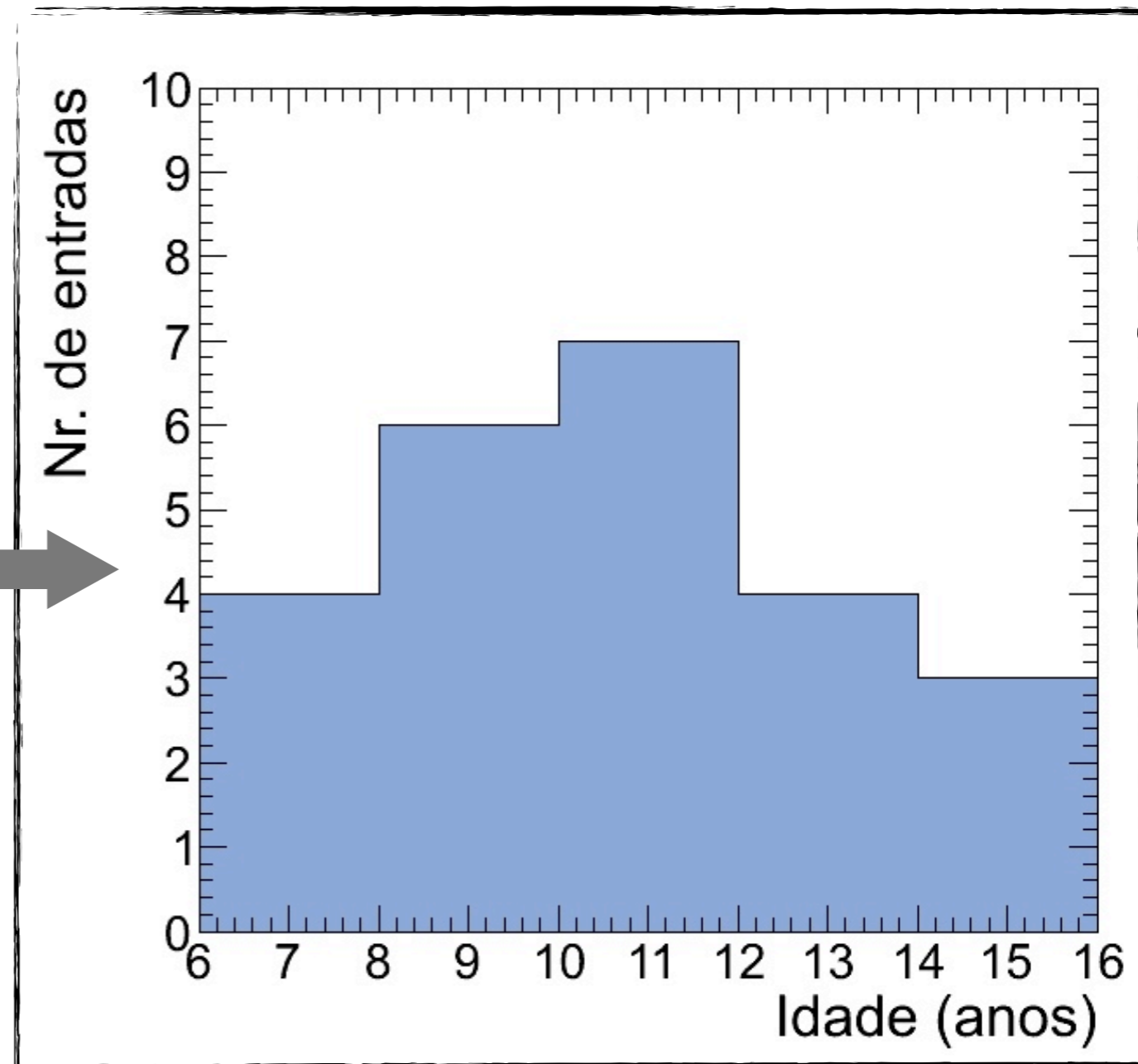
Classe de idades (anos)	Frequência
6 - 8	4
8 - 10	6
10 - 12	7
12 - 14	4
14 - 16	3

Organizando um conjunto de dados: Classes e Histogramas

Conjunto de idades:

{10, 7, 10, 11, 10, 15, 8, 12, 14, 9, 6, 8, 7, 14, 10, 10, 7, 12, 12, 9, 13, 10, 9, 8} (anos)

Classe de idades (anos)	Frequência
6 - 8	4
8 - 10	6
10 - 12	7
12 - 14	4
14 - 16	3

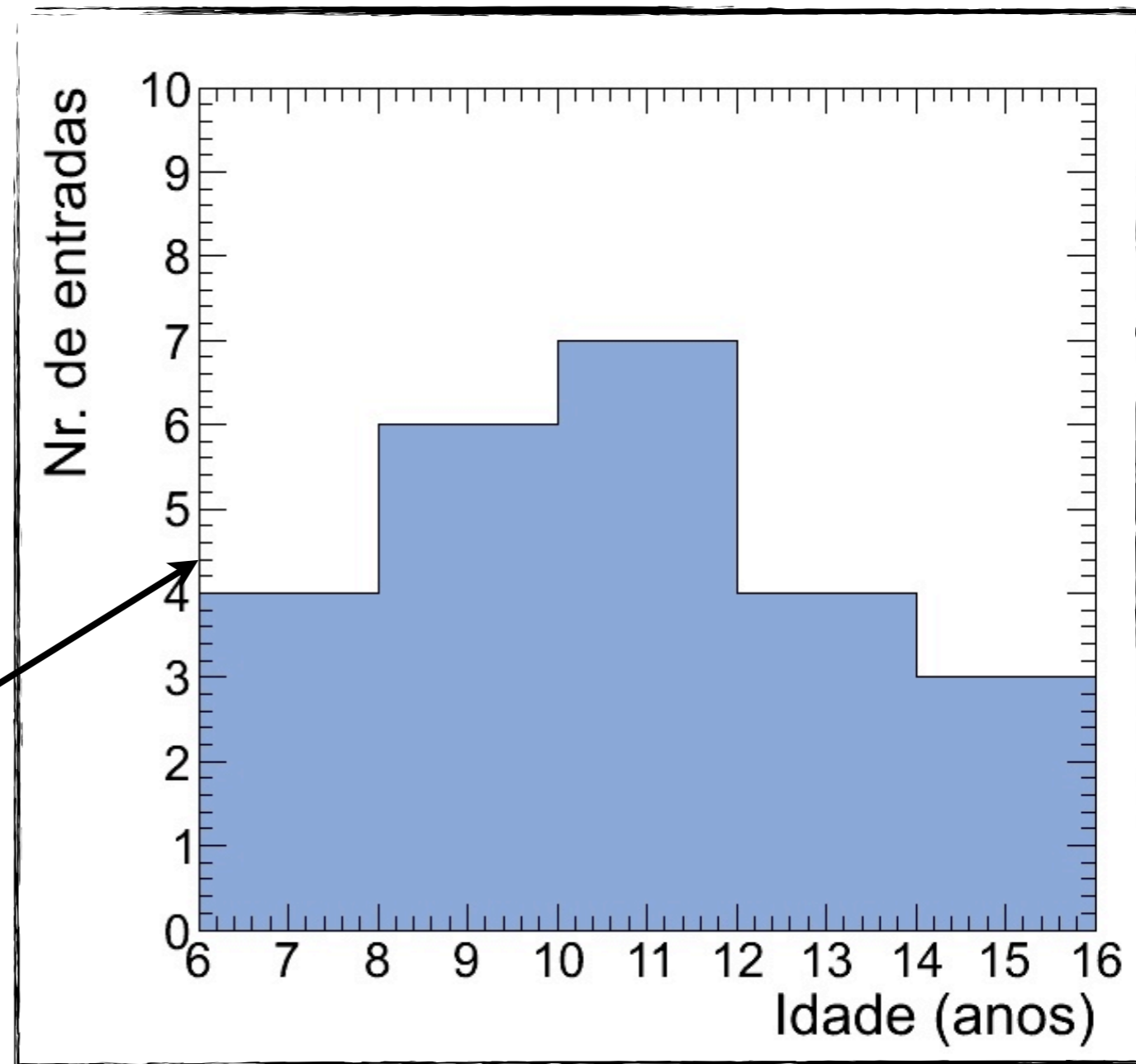
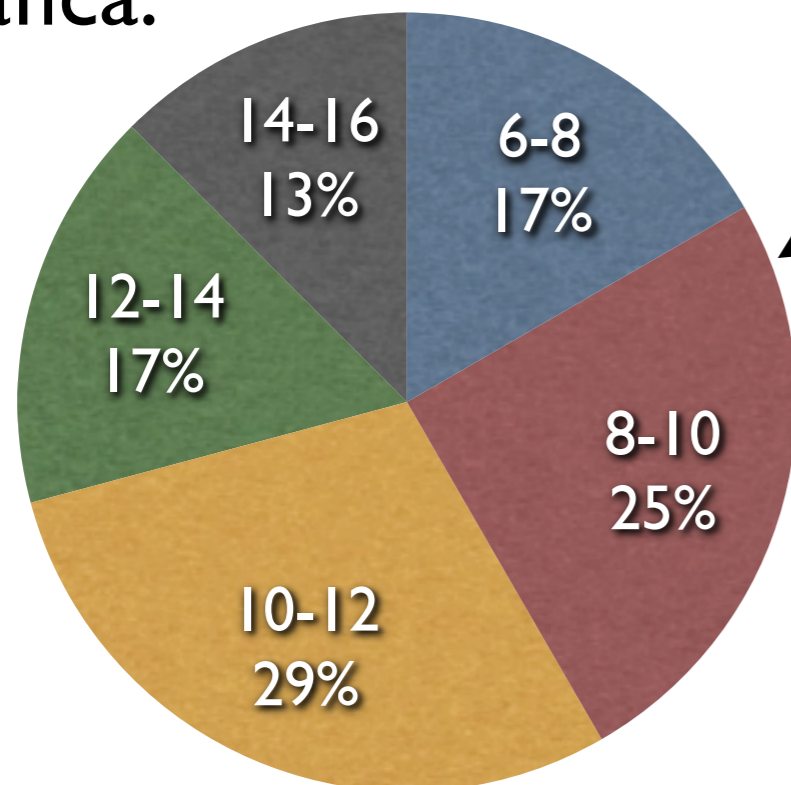


Organizando um conjunto de dados: Classes e Histogramas

Conjunto de idades:

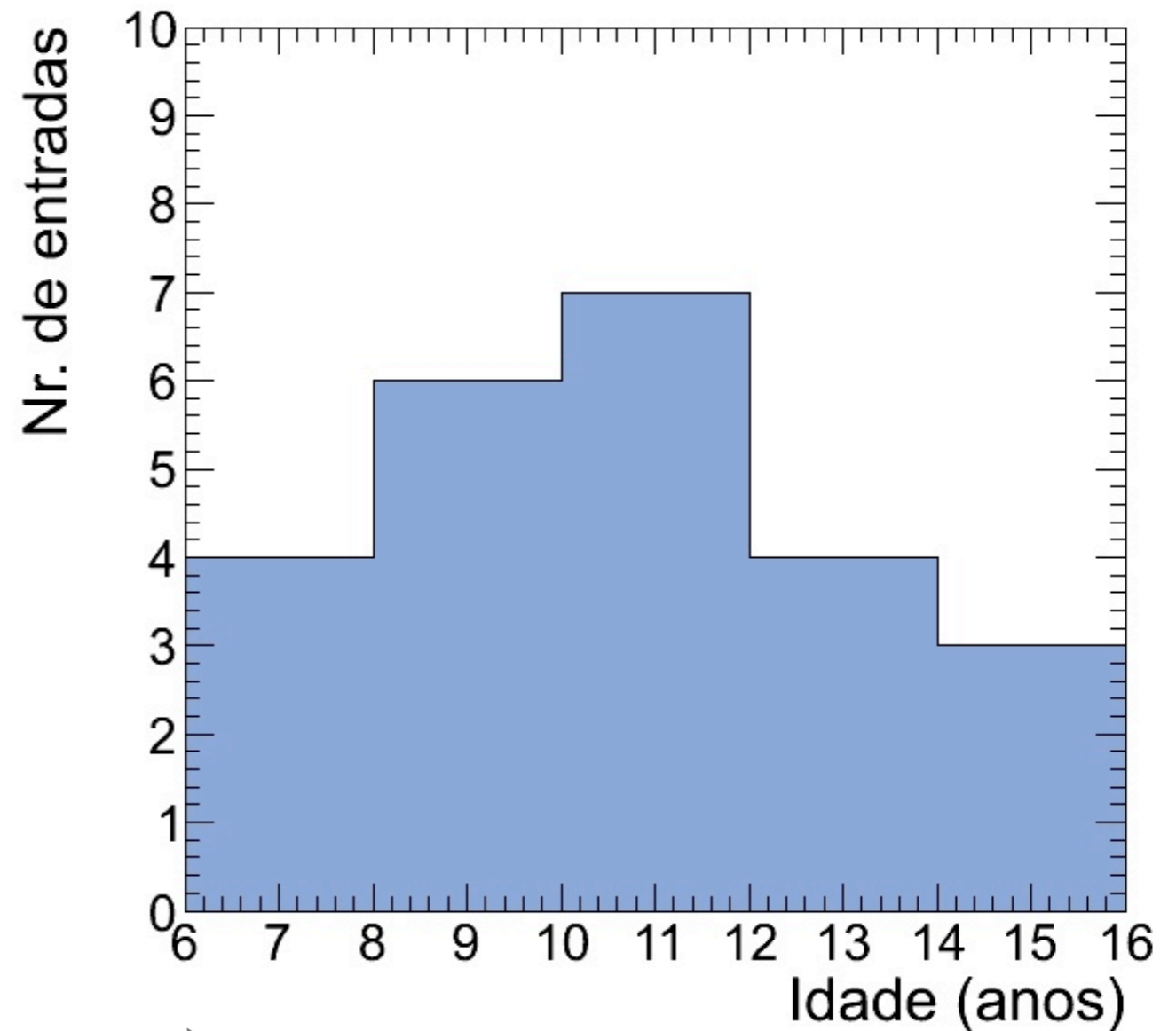
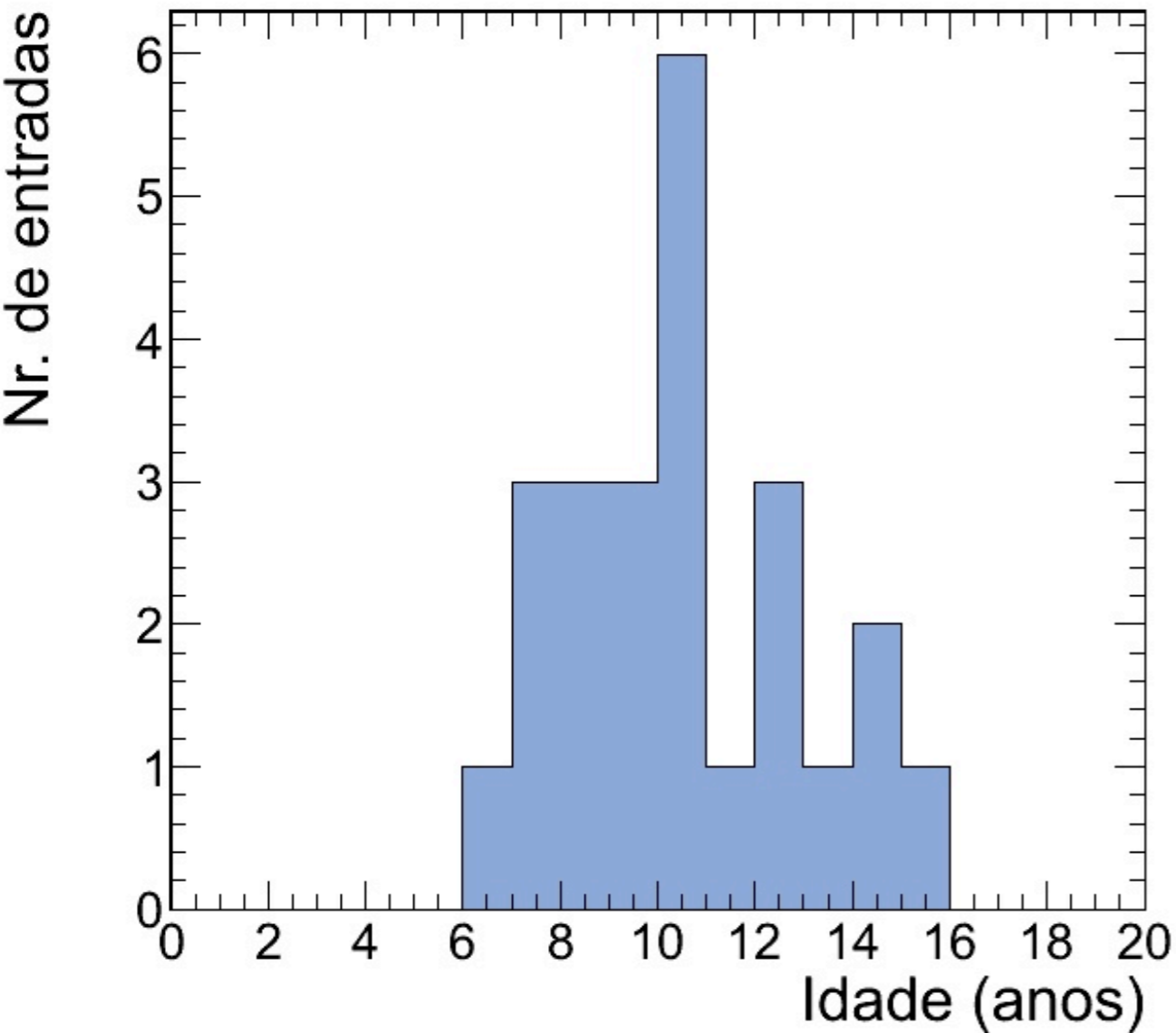
{10, 7, 10, 11, 10, 15, 8, 12, 14, 9, 6, 8, 7, 14, 10, 10, 7, 12, 12, 9, 13, 10, 9, 8} (anos)

Outra representação gráfica:



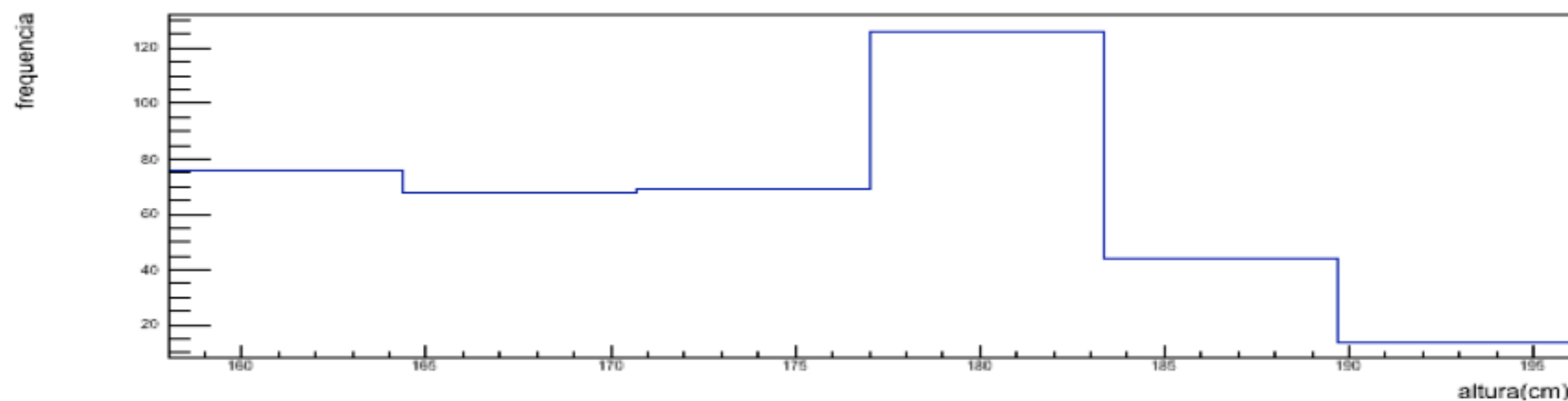
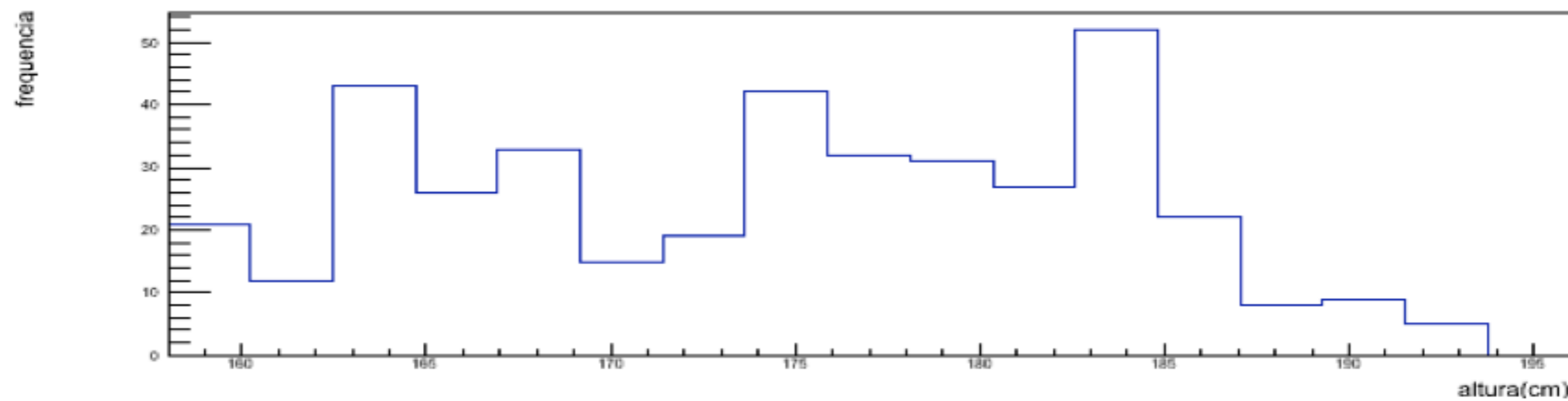
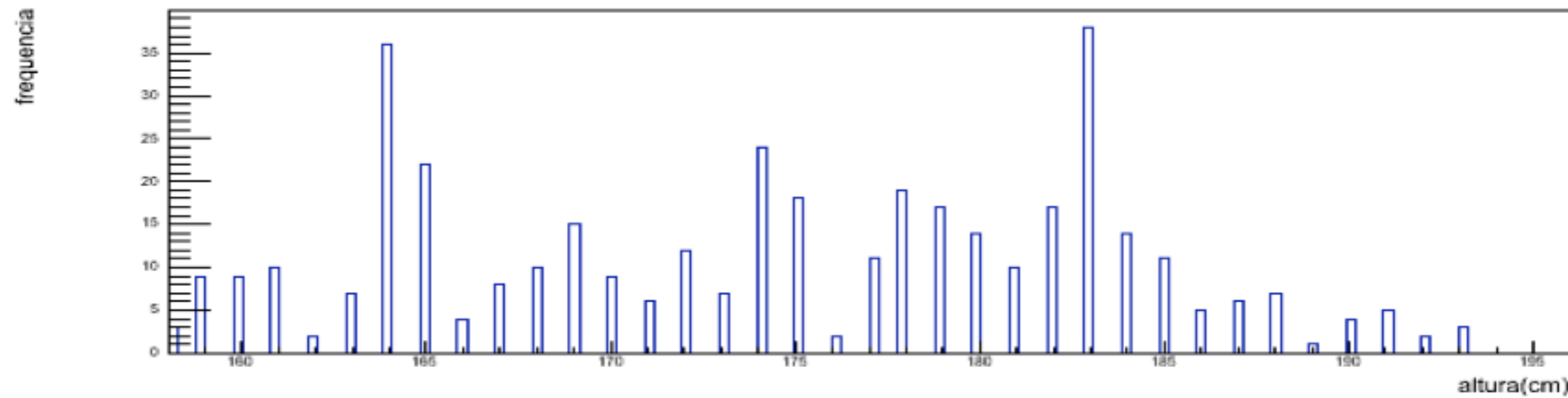
Organizando um conjunto de dados: Histogramas

Que tamanho de intervalo devemos usar para cada classe de frequência?



➔
Maior valor de intervalo

Um conjunto ainda maior de dados (valores de alturas de estudantes):



Maior valor de intervalo



Parâmetros de posição

Média: Valor médio de um conjunto de dados $\{x_1, x_2, x_3, \dots, x_N\}$

$$\bar{x} \equiv \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i$$

Parâmetros de posição

Média: Valor médio de um conjunto de dados $\{x_1, x_2, x_3, \dots, x_N\}$

Símbolo

$$\boxed{\bar{x}} \equiv \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i$$

Parâmetros de posição

Média: Valor médio de um conjunto de dados $\{x_1, x_2, x_3, \dots, x_N\}$

Símbolo

$$\boxed{\bar{x}} \equiv \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i$$

Cada elemento do conjunto de dados

Parâmetros de posição

Média: Valor médio de um conjunto de dados agrupados em M classes de frequência

Cada classe possui ponto médio $\{x_1, x_2, \dots, x_M\}$ e frequência $\{n_1, n_2, \dots, n_M\}$:

Parâmetros de posição

Média: Valor médio de um conjunto de dados agrupados em M classes de frequência

Cada classe possui ponto médio $\{x_1, x_2, \dots, x_M\}$ e frequência $\{n_1, n_2, \dots, n_M\}$:

$$\bar{x} \approx \frac{n_1 x_1 + n_2 x_2 + \dots + n_M x_M}{N} = \frac{1}{N} \sum_{j=1}^M n_j x_j$$

Parâmetros de posição

Média: Valor médio de um conjunto de dados agrupados em M classes de frequência

Cada classe possui ponto médio $\{x_1, x_2, \dots, x_M\}$ e frequência $\{n_1, n_2, \dots, n_M\}$:

$$\bar{x} \approx \frac{n_1 x_1 + n_2 x_2 + \dots + n_M x_M}{N} = \frac{1}{N} \sum_{j=1}^M n_j x_j$$

M : número de classes de frequência

N : número total de elementos

Parâmetros de posição

Média: Valor médio de um conjunto de dados agrupados em M classes de frequência

Cada classe possui ponto médio $\{x_1, x_2, \dots, x_M\}$ e frequência $\{n_1, n_2, \dots, n_M\}$:

$$\bar{x} \approx \frac{n_1 x_1 + n_2 x_2 + \dots + n_M x_M}{N} = \frac{1}{N} \sum_{j=1}^M n_j x_j$$

M : número de classes de frequência

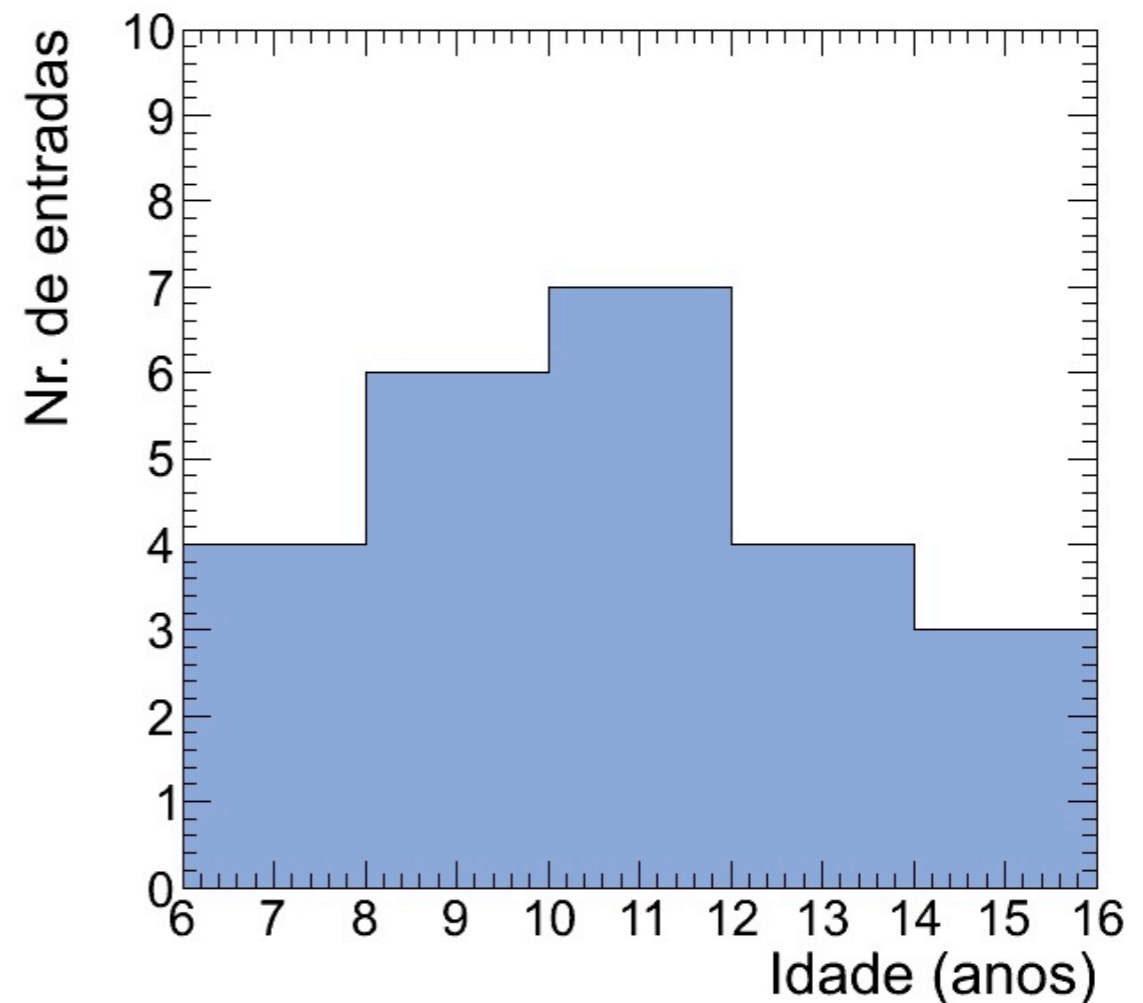
N : número total de elementos $\sum_{j=1}^M n_j = n_1 + n_2 + \dots + n_M = N$

Parâmetros de posição

Moda: Valor mais frequente de um conjunto de dados $\{x_1, x_2, x_3, \dots, x_N\}$

Símbolo: x_{mod}

Para dados agrupados em classes de frequências a moda é o ponto médio da classe de maior frequência

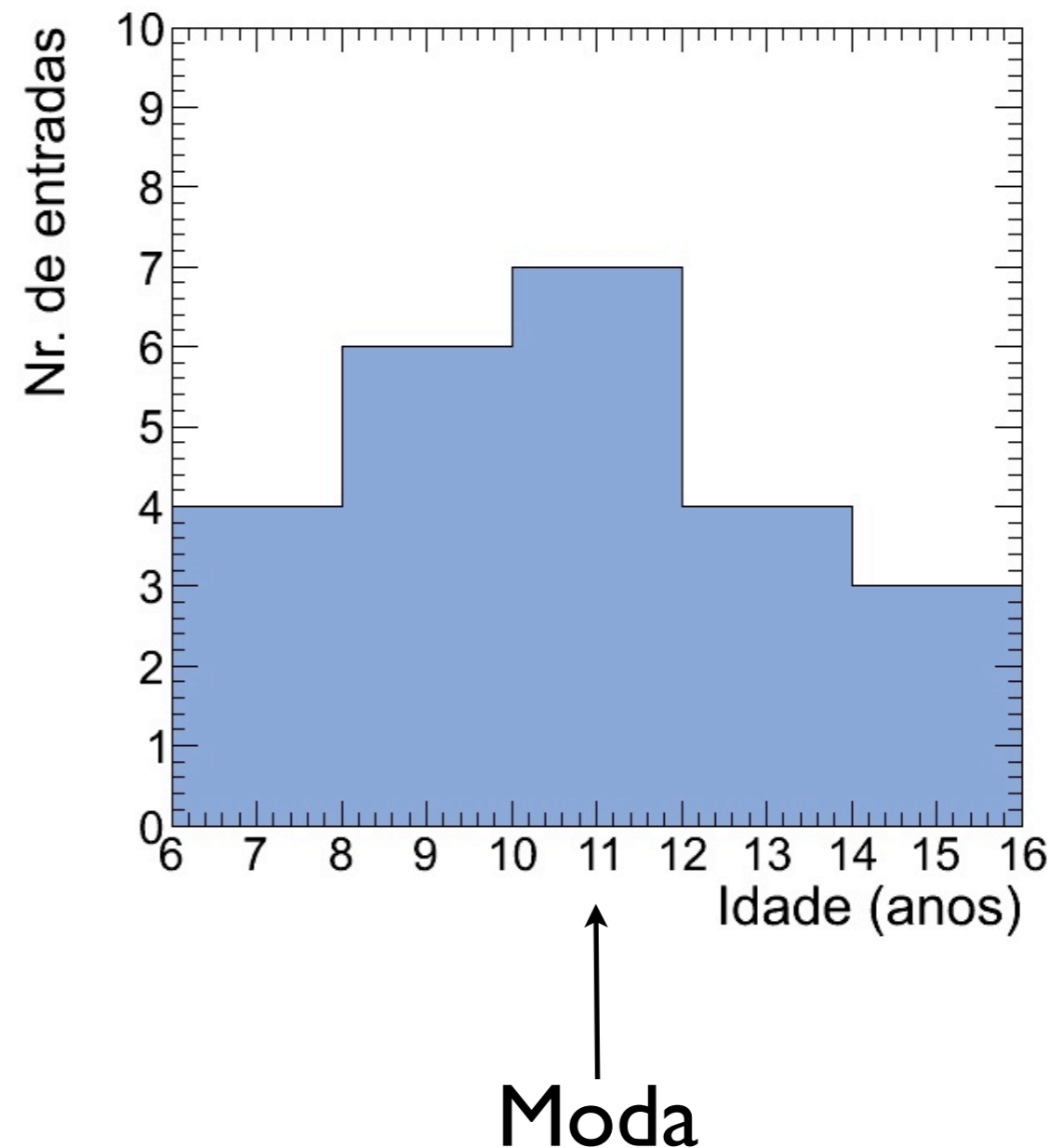


Parâmetros de posição

Moda: Valor mais frequente de um conjunto de dados $\{x_1, x_2, x_3, \dots, x_N\}$

Símbolo: x_{mod}

Para dados agrupados em classes de frequências a moda é o ponto médio da classe de maior frequência



Parâmetros de posição

Média quadrática: raiz quadrada da média dos quadrados dos dados:

$$x_{\text{rms}} \equiv \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_N^2}{N}} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$$

Parâmetros de posição

Média quadrática: raiz quadrada da média dos quadrados dos dados:

Símbolo

$$\boxed{x_{\text{rms}}} \equiv \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_N^2}{N}} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$$

Parâmetros de posição

Mediana: valor que divide uma distribuição ordenada de dados de forma que metade dos dados está acima, e metade abaixo deste valor

$$N(\text{ímpar}) \rightarrow x_{\text{med}} = x_{(N+1)/2}$$

$$N(\text{par}) \rightarrow x_{\text{med}} = \frac{x_{N/2} + x_{(N/2+1)}}{2}$$

Atividade de aula - Idade, massa e altura

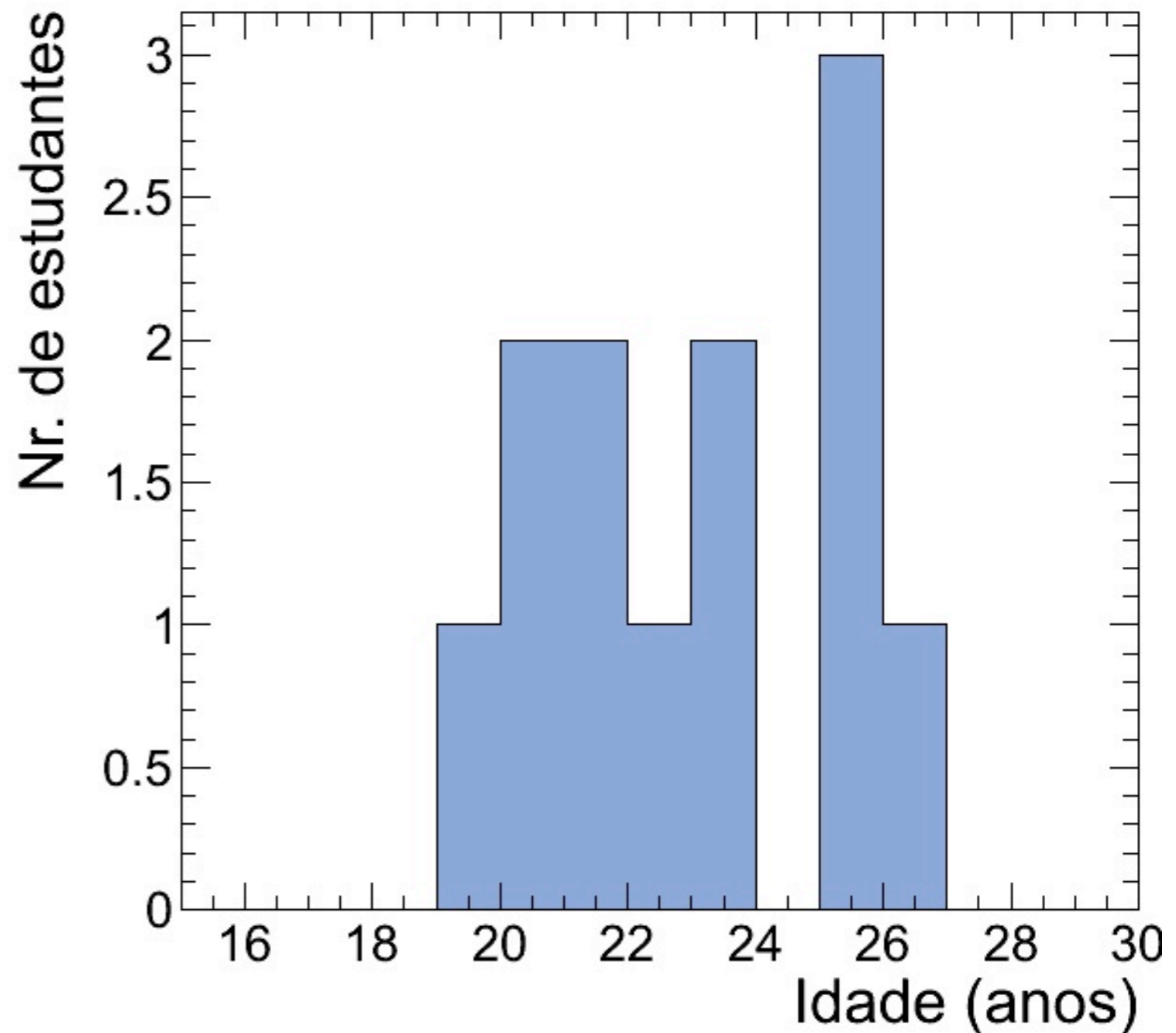
Estudante	Idade (anos)	Massa (kg)	Altura (cm)
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Atividade de aula - Idade, massa e altura

Estudante	Idade (anos)	Massa (kg)	Altura (cm)
1	22	64	174
2	21	110	185
3	20	75	174
4	23	80	170
5	25	61	168
6	30	86	173
7	20	54	162
8	25	84	176
9	19	51	166
10	26	64	168
11	21	66	177
12	23	103	174
13	25	91	175

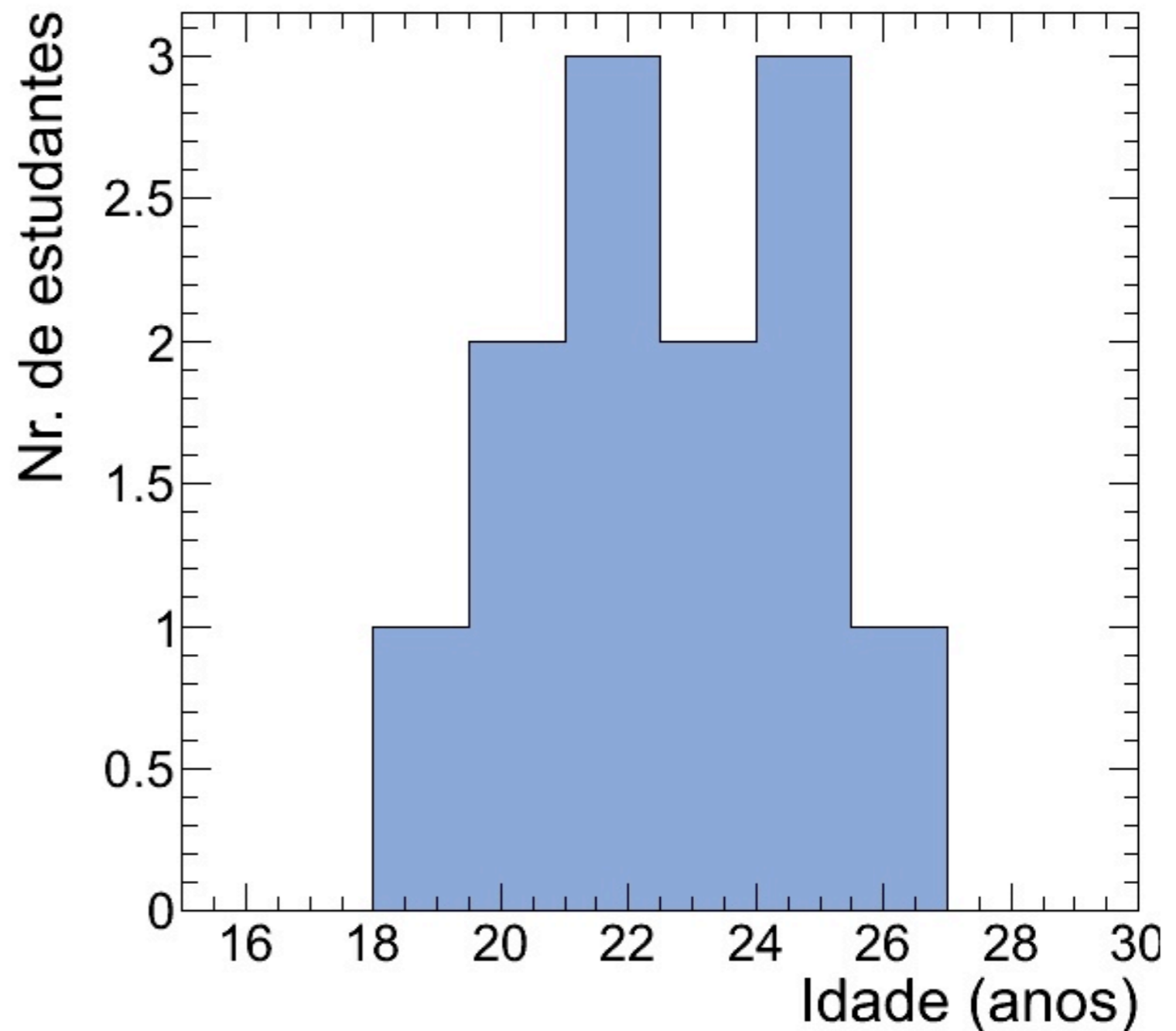
Atividade de aula - Idade, massa e altura

Estudante	Idade (anos)
1	22
2	21
3	20
4	23
5	25
6	30
7	20
8	25
9	19
10	26
11	21
12	23
13	25



Atividade de aula - Idade, massa e altura

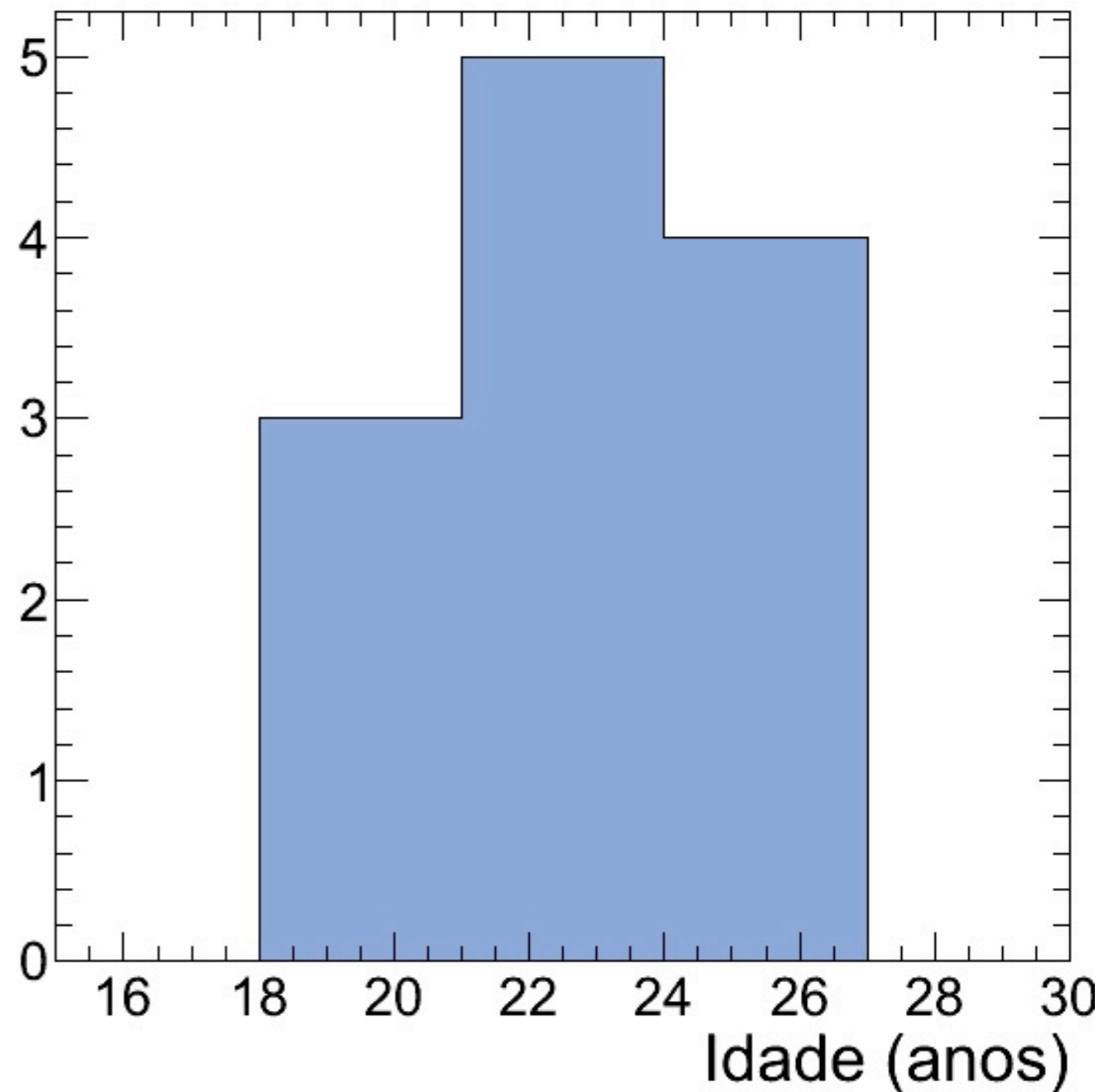
Estudante	Idade (anos)
1	22
2	21
3	20
4	23
5	25
6	30
7	20
8	25
9	19
10	26
11	21
12	23
13	25



Atividade de aula - Idade, massa e altura

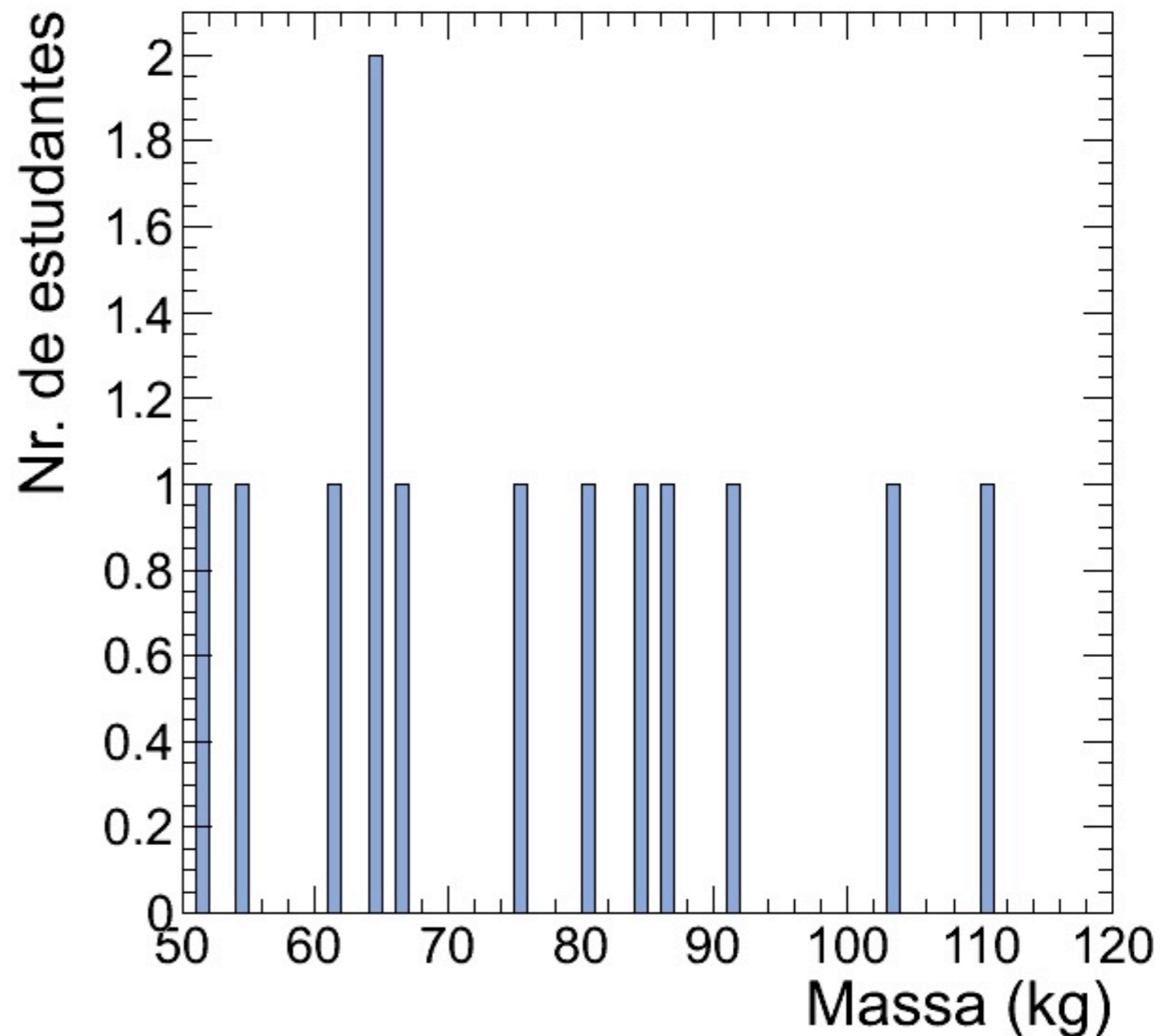
Estudante	Idade (anos)
1	22
2	21
3	20
4	23
5	25
6	30
7	20
8	25
9	19
10	26
11	21
12	23
13	25

Nr. de estudantes



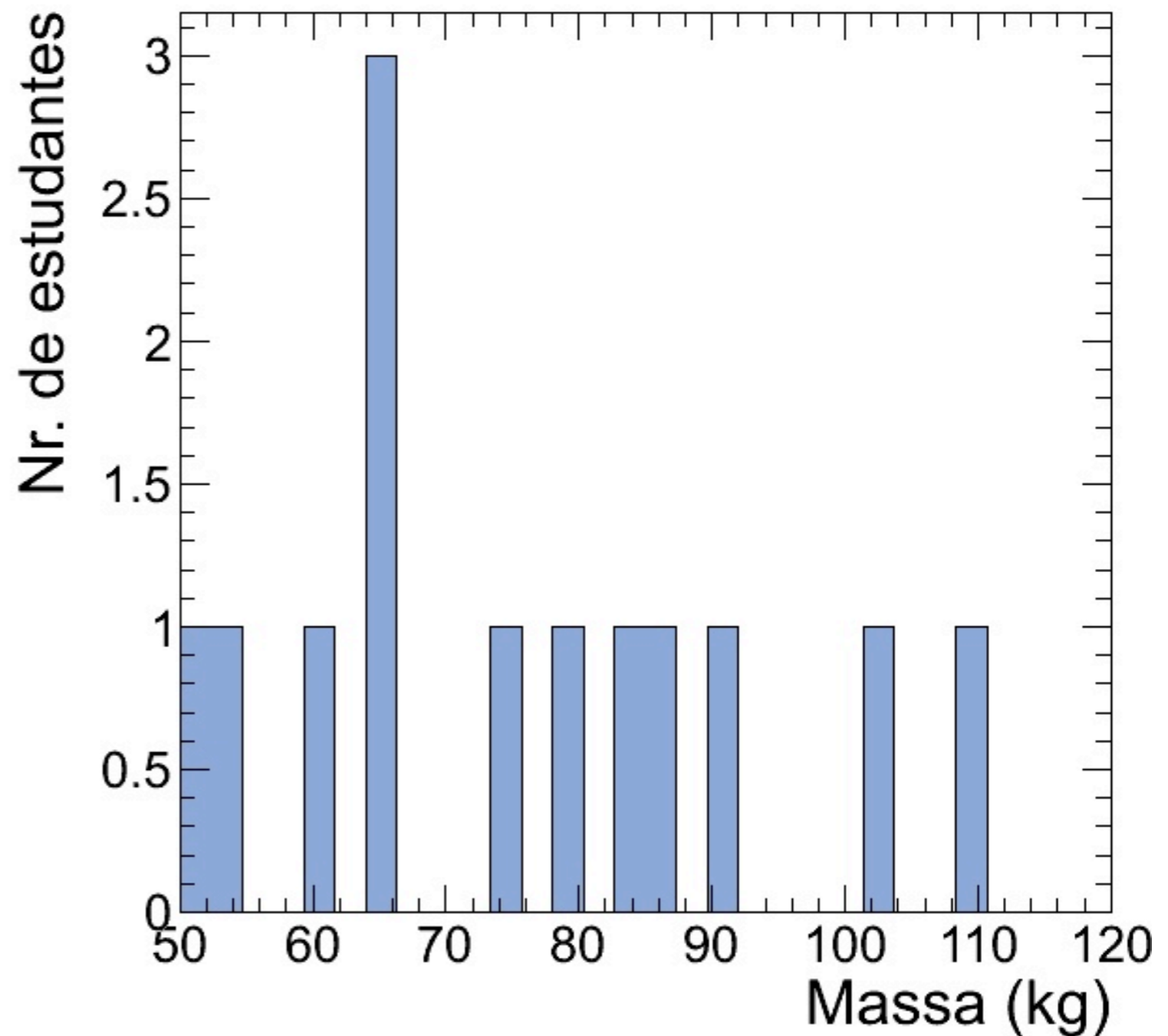
Atividade de aula - Idade, massa e altura

Estudante	Massa (kg)
1	64
2	110
3	75
4	80
5	61
6	86
7	54
8	84
9	51
10	64
11	66
12	103
13	91



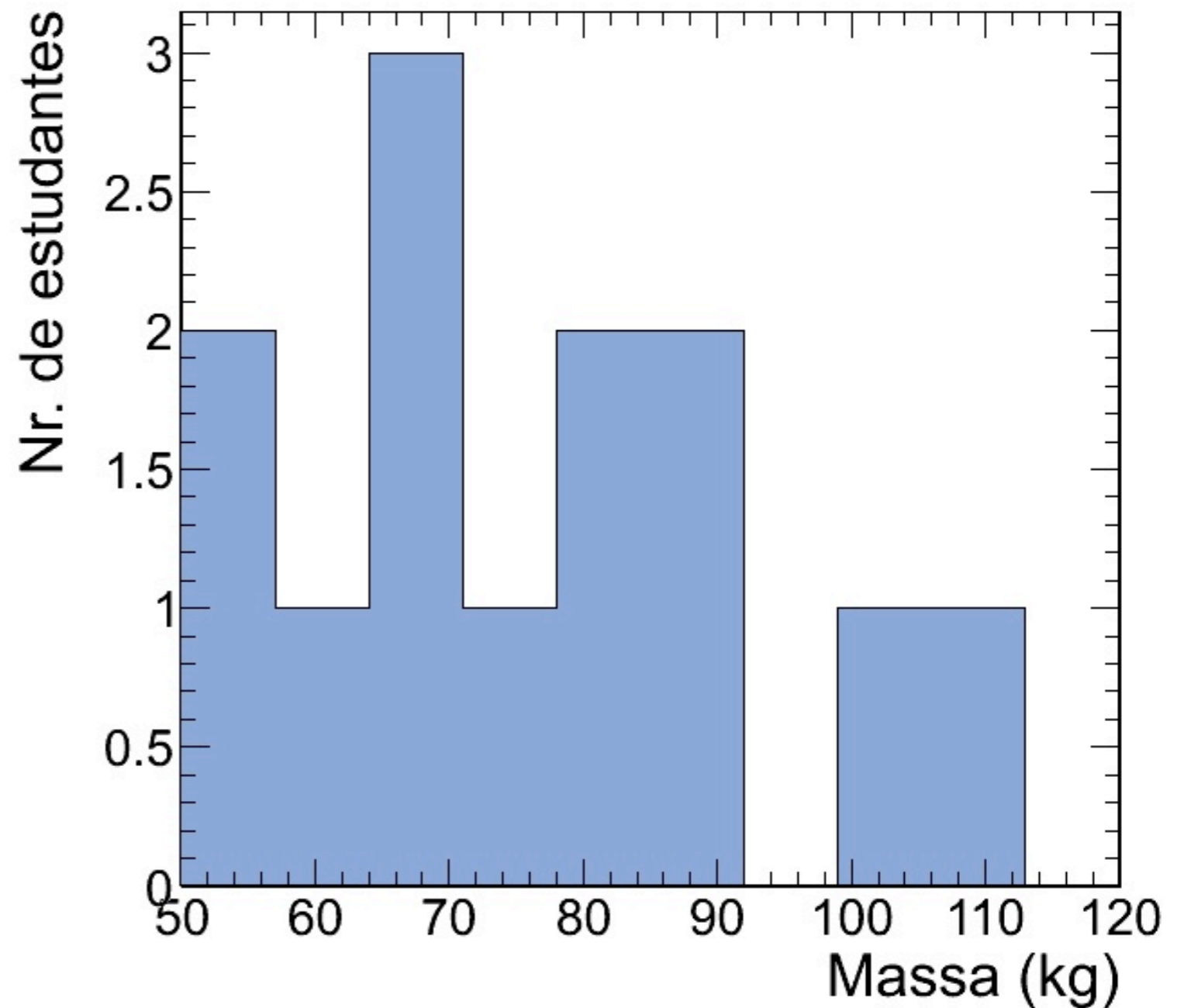
Atividade de aula - Idade, massa e altura

Estudante	Massa (kg)
1	64
2	110
3	75
4	80
5	61
6	86
7	54
8	84
9	51
10	64
11	66
12	103
13	91



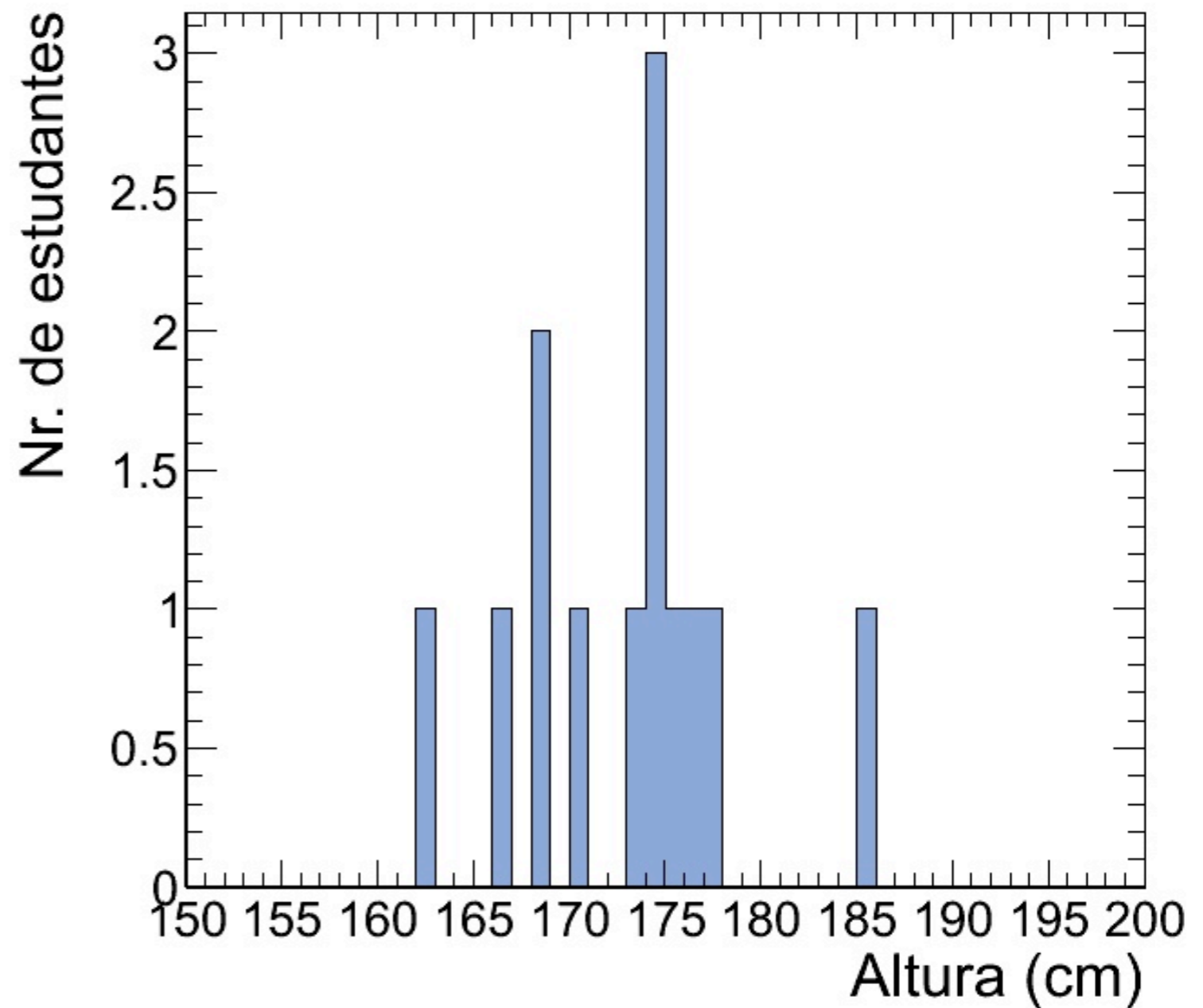
Atividade de aula - Idade, massa e altura

Estudante	Massa (kg)
1	64
2	110
3	75
4	80
5	61
6	86
7	54
8	84
9	51
10	64
11	66
12	103
13	91



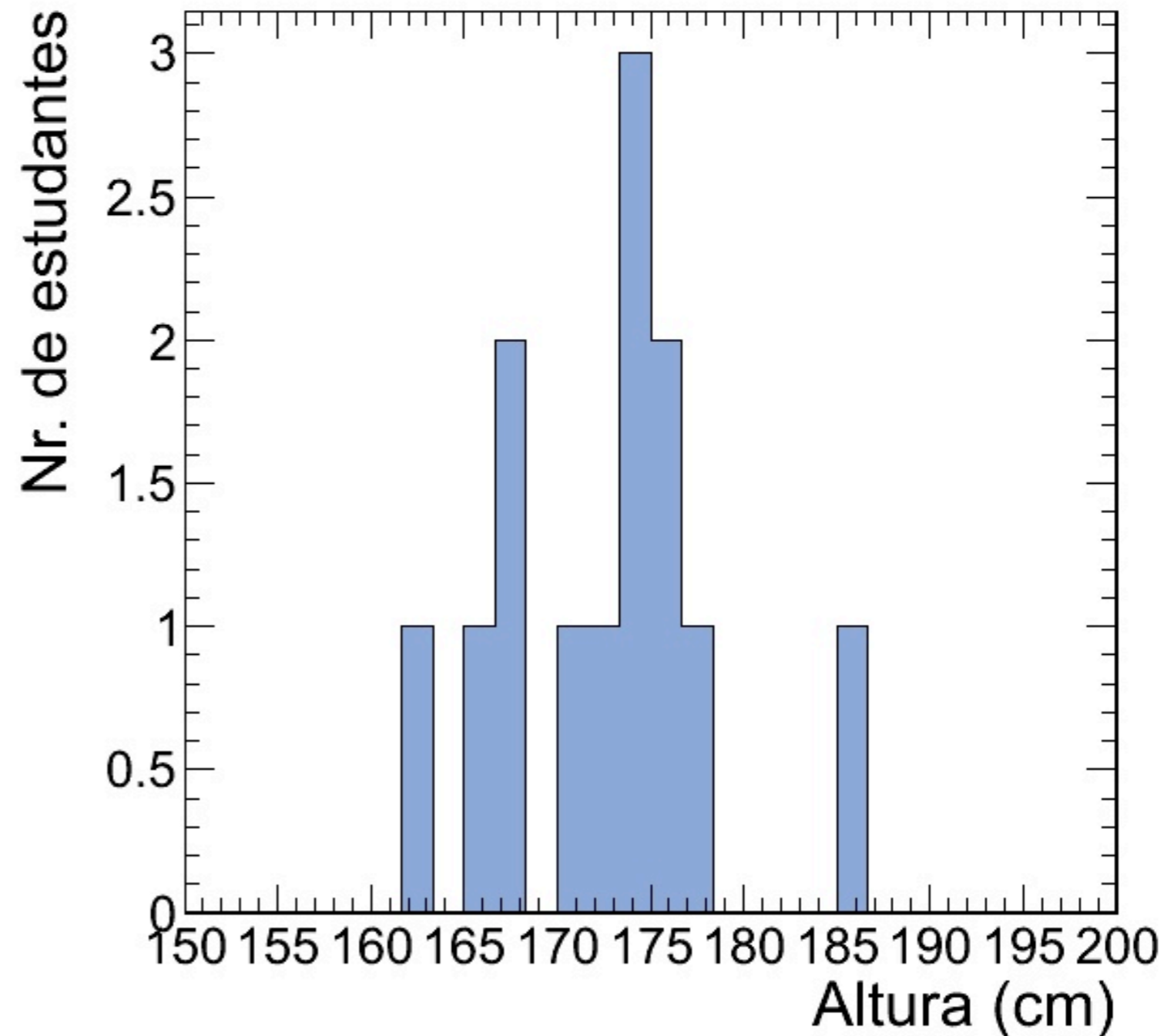
Atividade de aula - Idade, massa e altura

Estudante	Altura (cm)
1	174
2	185
3	174
4	170
5	168
6	173
7	162
8	176
9	166
10	168
11	177
12	174
13	175



Atividade de aula - Idade, massa e altura

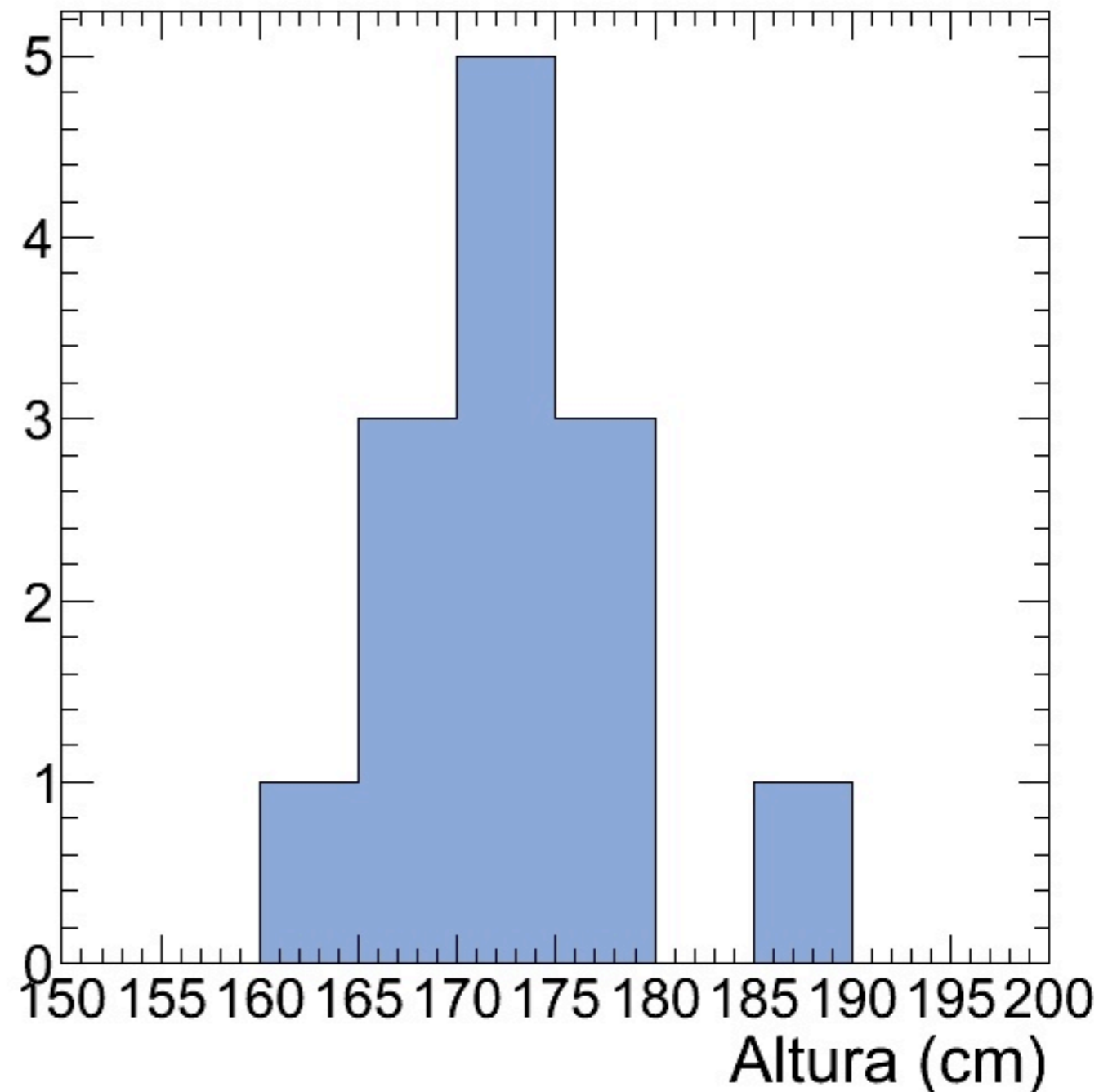
Estudante	Altura (cm)
1	174
2	185
3	174
4	170
5	168
6	173
7	162
8	176
9	166
10	168
11	177
12	174
13	175



Atividade de aula - Idade, massa e altura

Estudante	Altura (cm)
1	174
2	185
3	174
4	170
5	168
6	173
7	162
8	176
9	166
10	168
11	177
12	174
13	175

Nr. de estudantes



Parâmetros de dispersão

i) *Amplitude*: Diferença entre os valores máximo e mínimo de uma coleção de dados $\{x_1, x_2, \dots, x_N\}$

$$A = x_{\max} - x_{\min}$$

Parâmetros de dispersão

ii) *Desvio médio*: Média dos módulos dos desvios, em relação à média

$$\overline{|\delta x|} = \frac{1}{N} \sum_{i=1}^N |\delta x_i| = \frac{1}{N} \sum_{i=1}^N |x_i - \bar{x}| = \frac{|x_1 - \bar{x}| + \dots + |x_N - \bar{x}|}{N}$$

Parâmetros de dispersão

iii) *Variância*: Média dos quadrados dos desvios (δx_i)

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^N (\delta x_i)^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 = \frac{(x_1 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}$$


Note que a expressão para a variância pode ser simplificada por:

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 - \left(\frac{1}{N} \sum_{i=1}^N x_i \right)^2 = \overline{x^2} - \bar{x}^2$$

Parâmetros de dispersão

iv) *Desvio padrão*: Raiz quadrada da variância, ou média quadrática dos desvios

$$\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (\delta x_i)^2} = \sqrt{\frac{(x_1 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}}$$

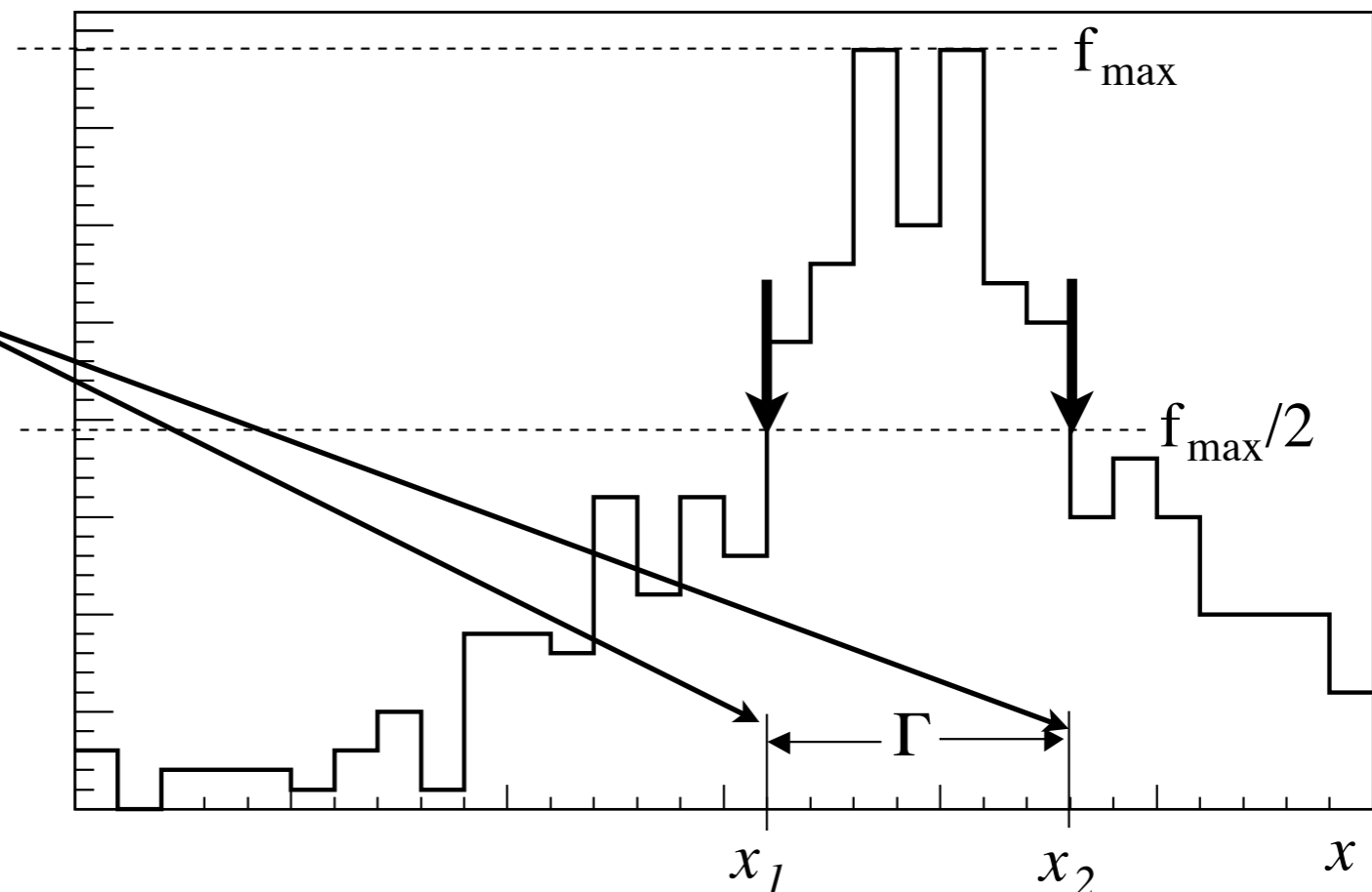

$$\sigma_x = \sqrt{\overline{x^2} - \bar{x}^2}$$

Parâmetros de dispersão

v) *Largura a meia altura*: Comprimento do intervalo limitado pelos valores (x_1, x_2) correspondentes à metade da frequência máxima

Símbolo: Γ

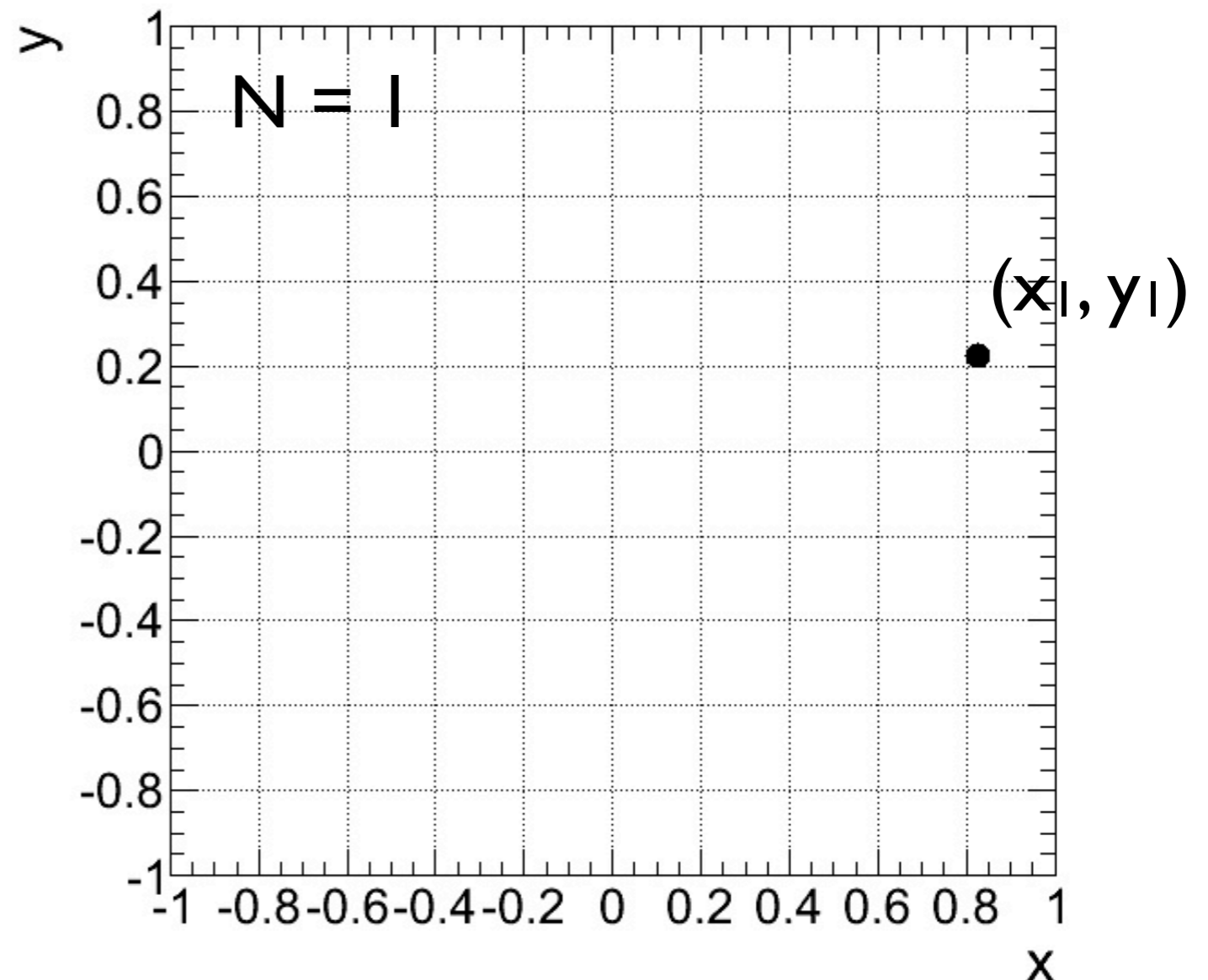
$$\Gamma = |x_2 - x_1|$$



Representando duas variáveis

Diagrama de dispersão: Gráfico representando medidas em duas variáveis $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

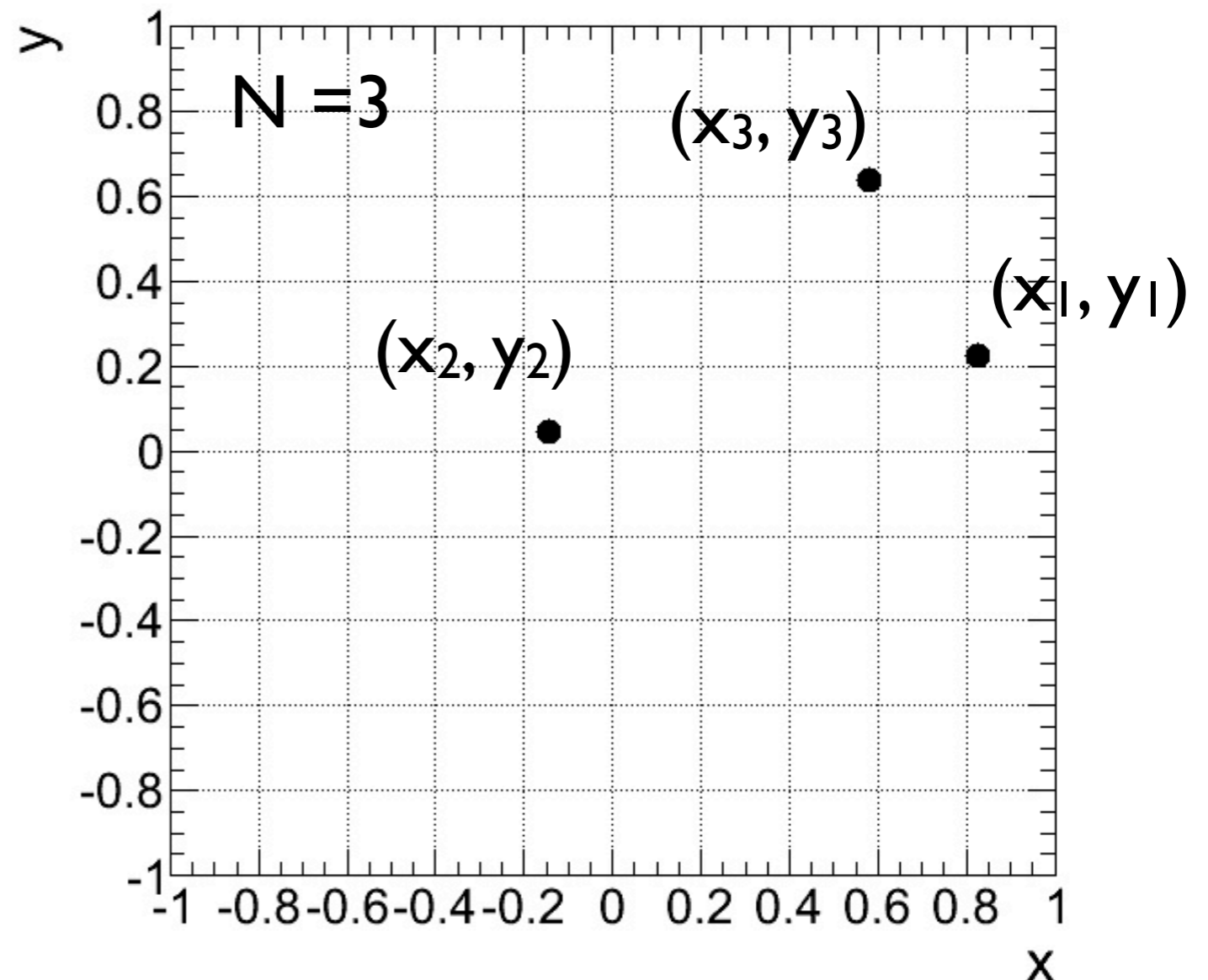
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

Diagrama de dispersão: Gráfico representando medidas em duas variáveis $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

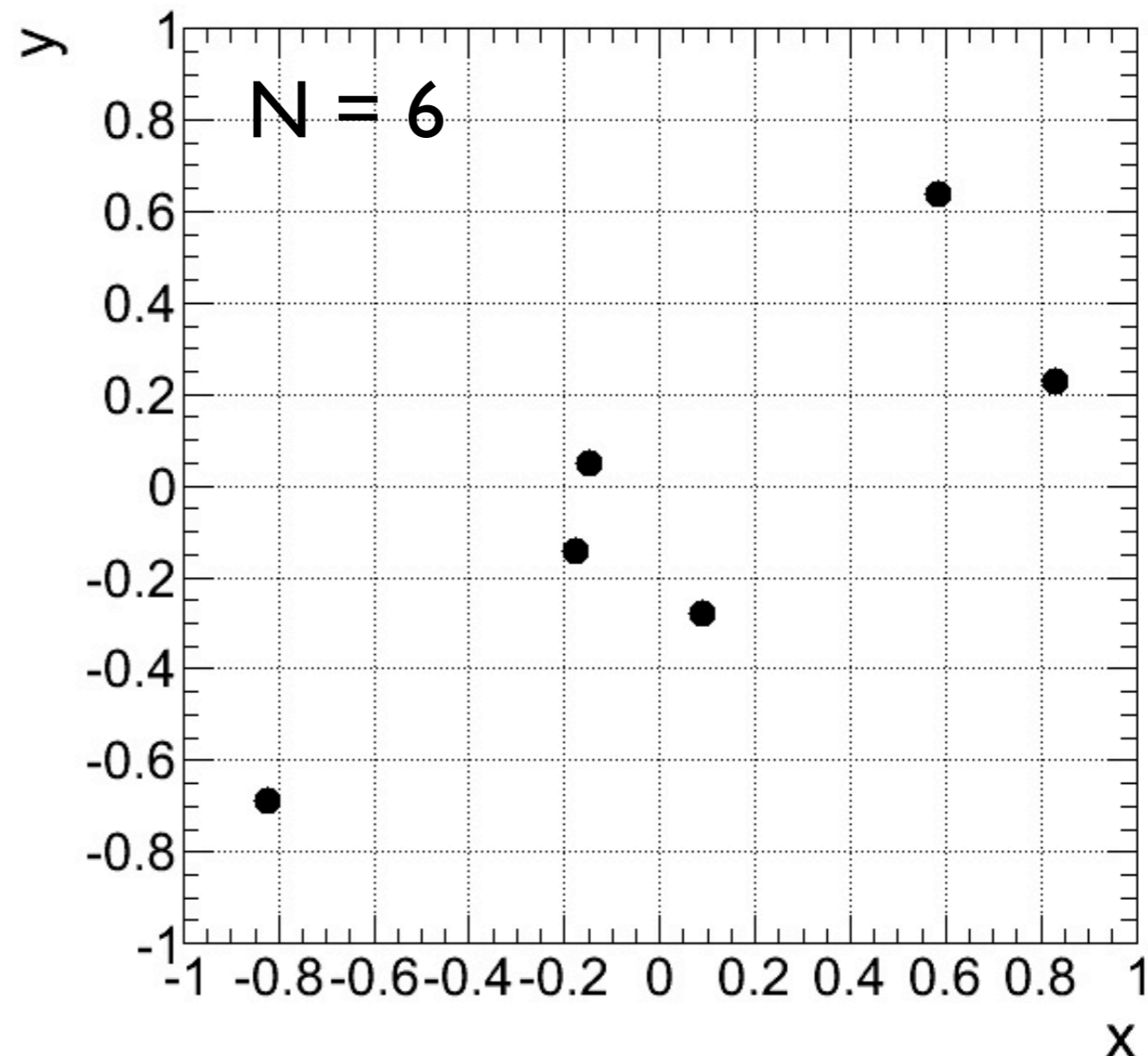
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

Diagrama de dispersão: Gráfico representando medidas em duas variáveis $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

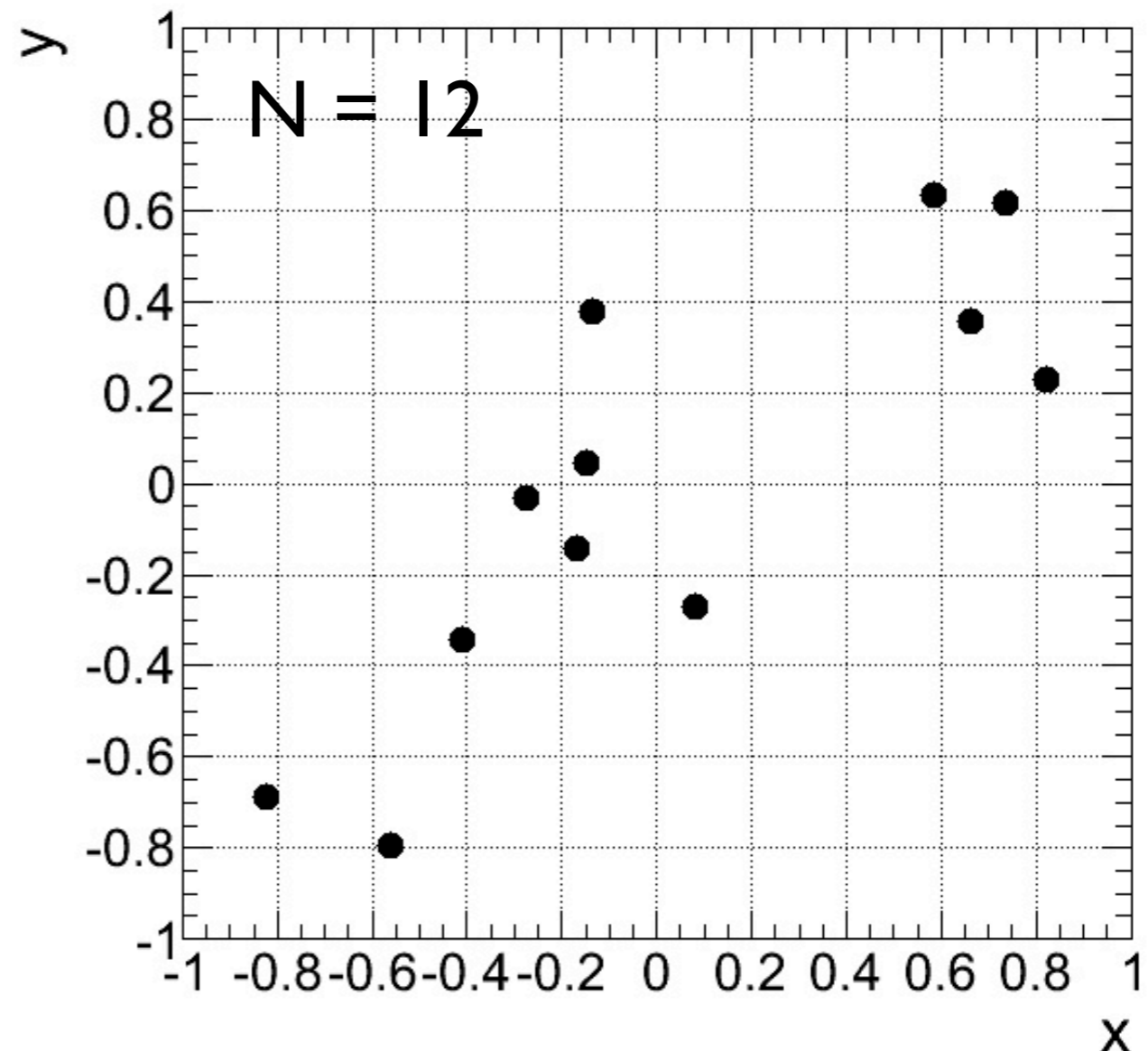
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

Diagrama de dispersão: Gráfico representando medidas em duas variáveis $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

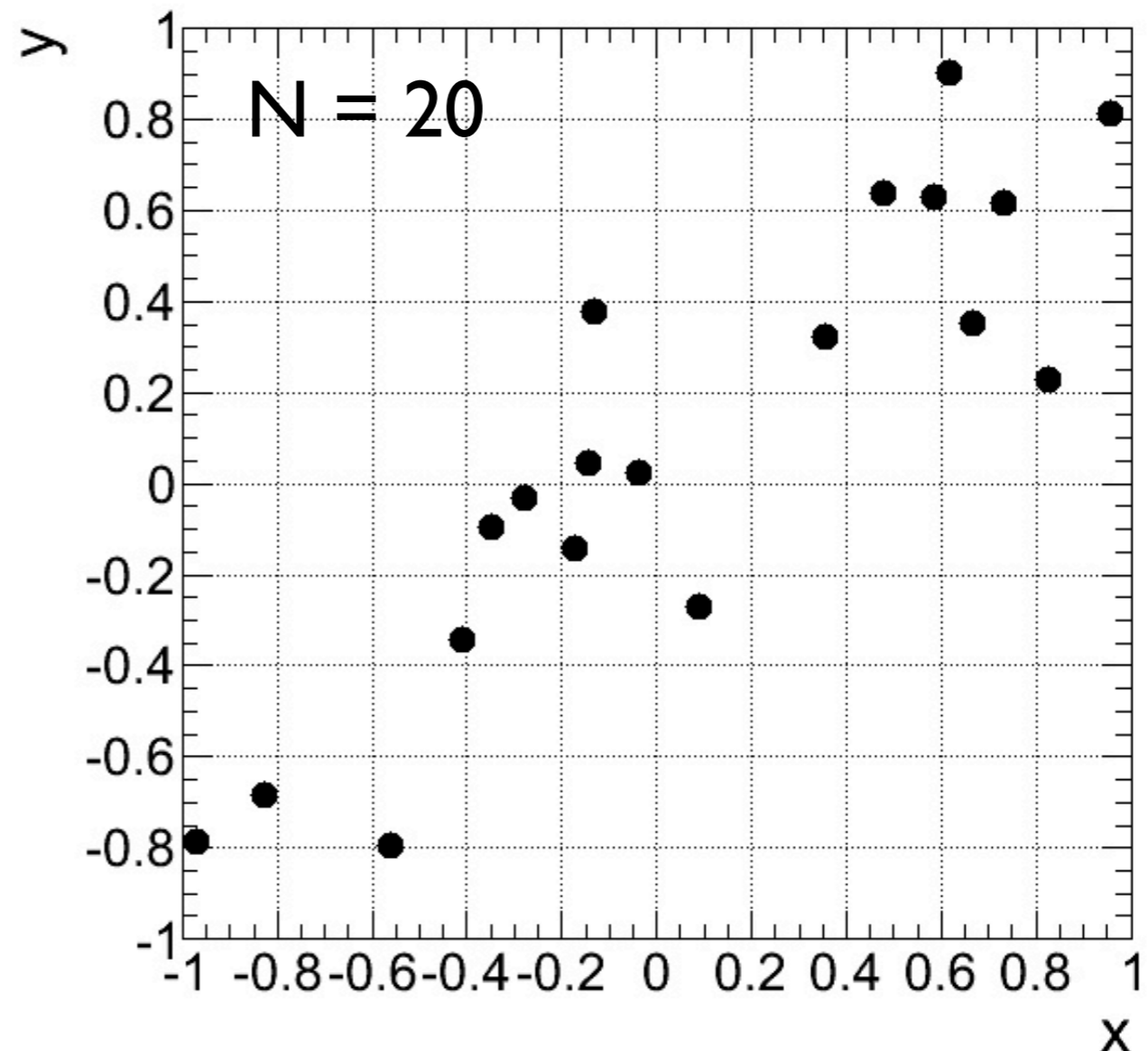
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

Diagrama de dispersão: Gráfico representando medidas em duas variáveis $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

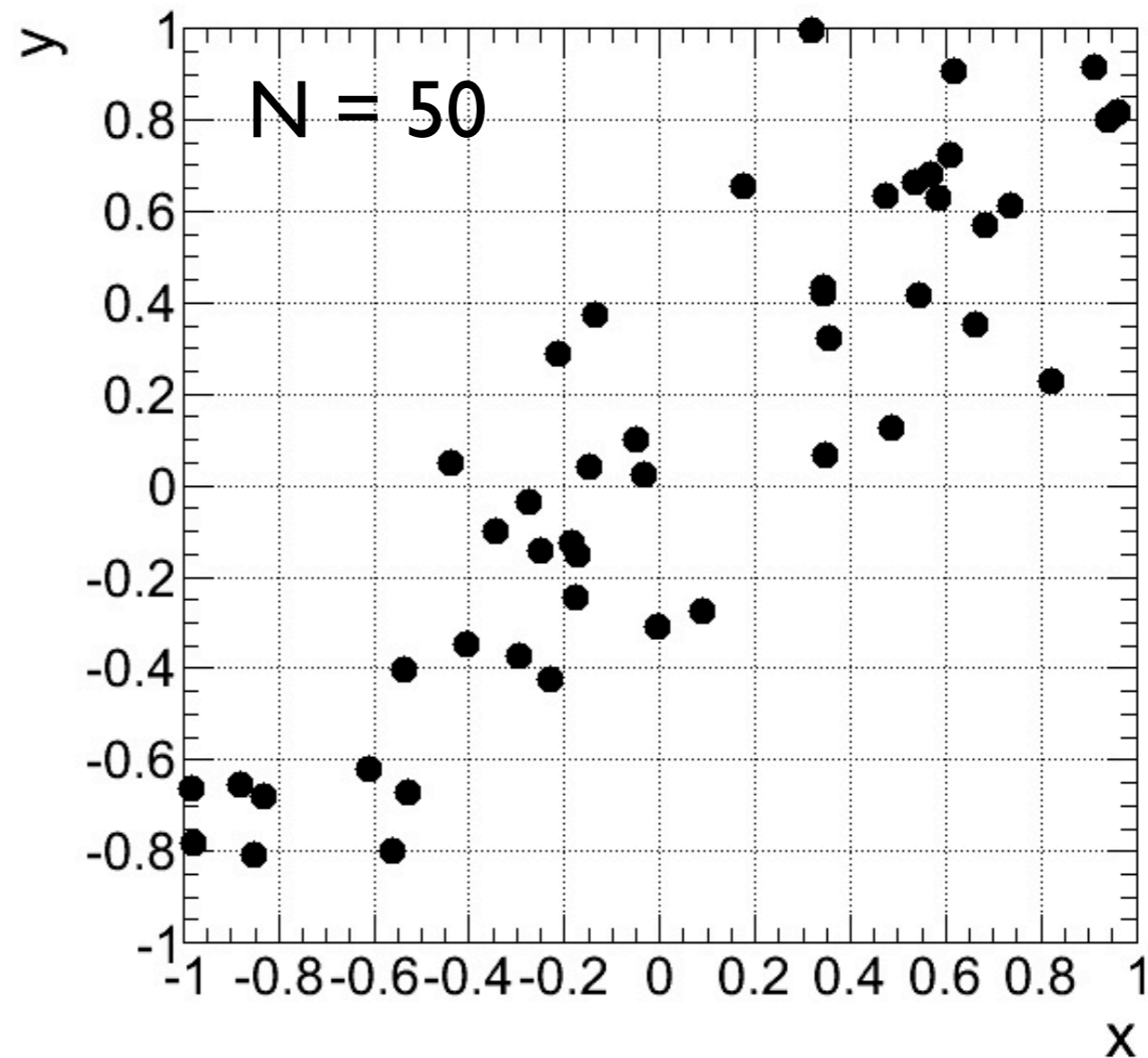
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

Diagrama de dispersão: Gráfico representando medidas em duas variáveis $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

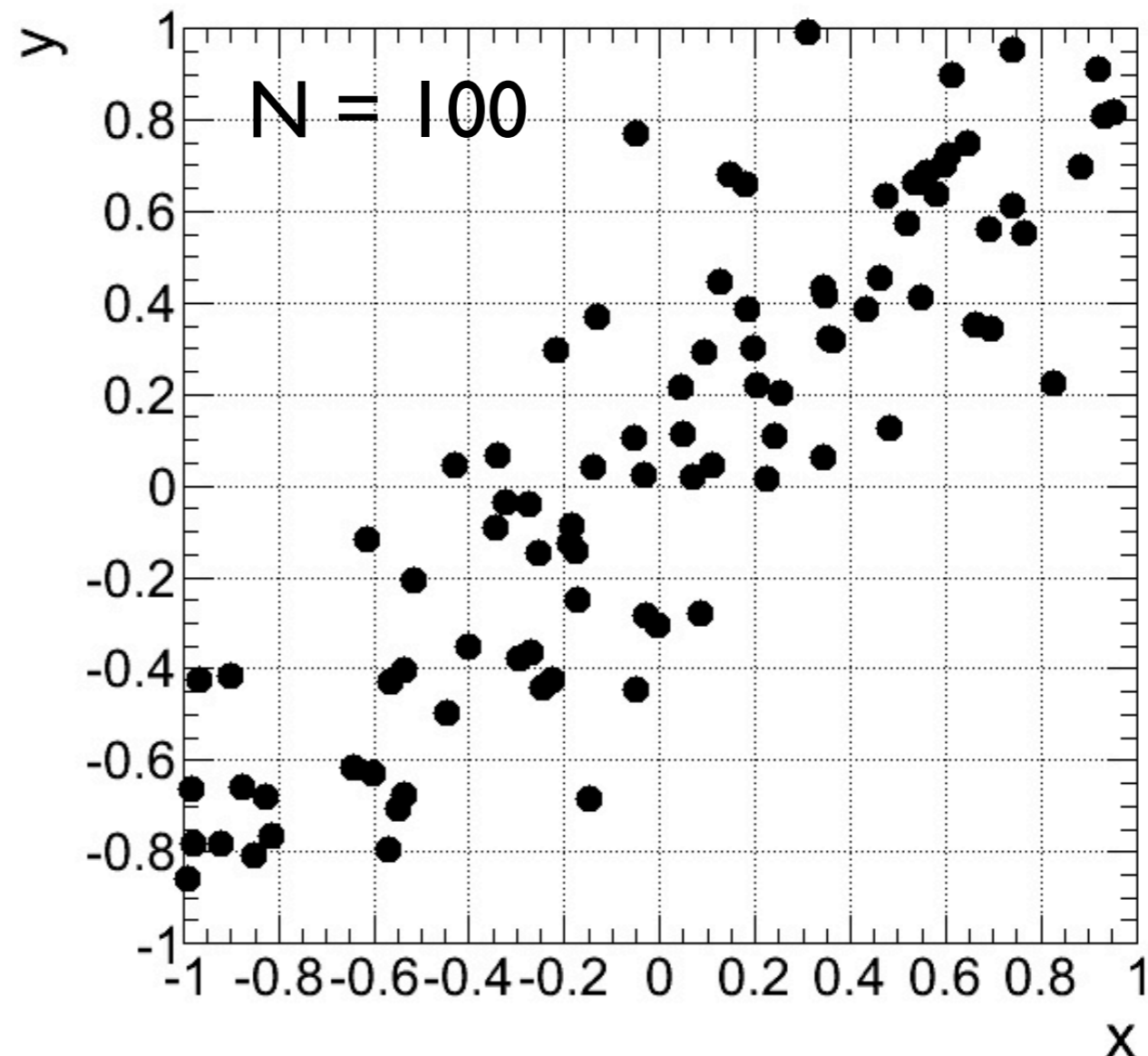
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

Diagrama de dispersão: Gráfico representando medidas em duas variáveis $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

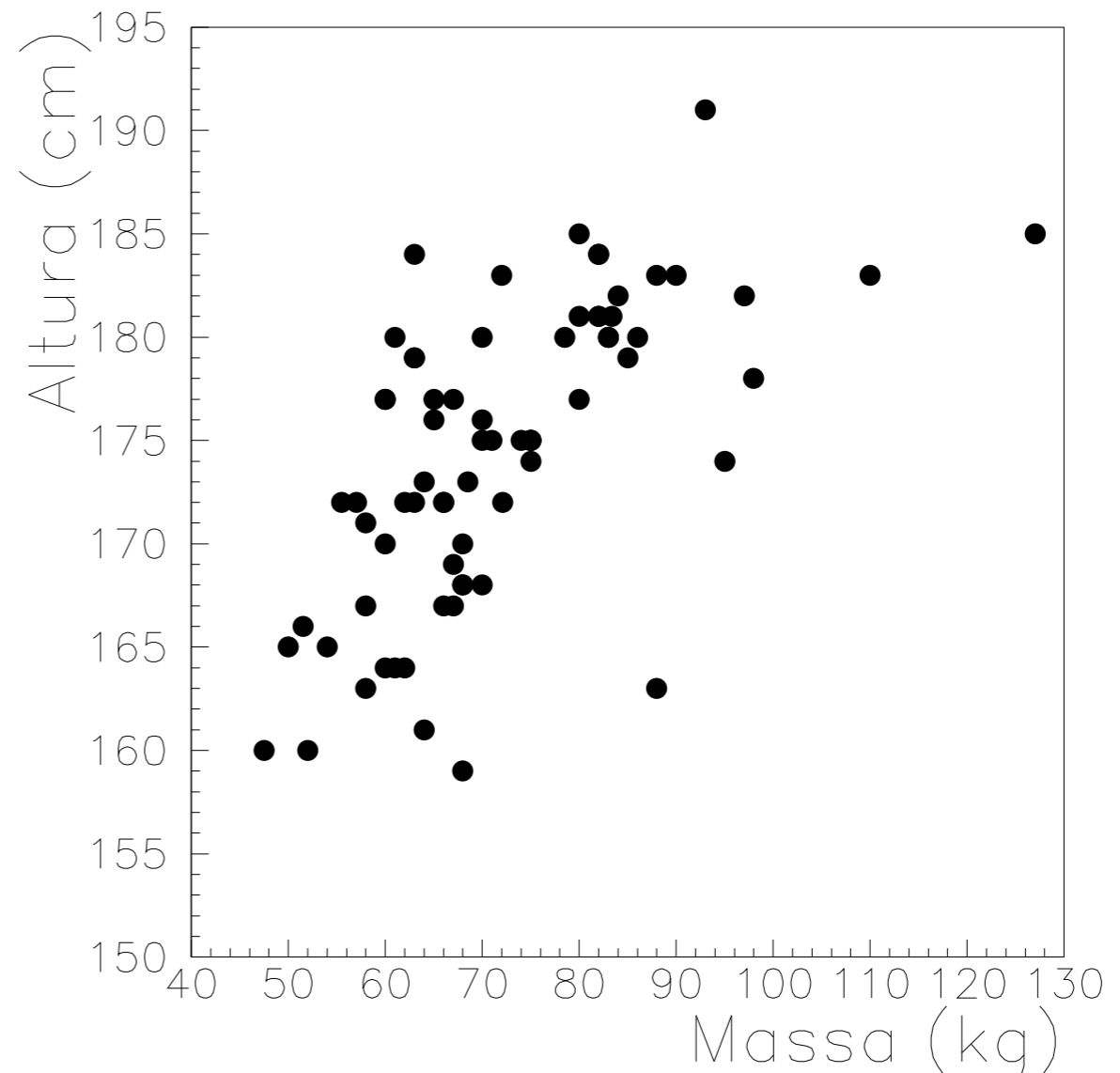
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

Diagrama de dispersão: Gráfico representando medidas em duas variáveis $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

Outro exemplo: dados de altura e massa de uma lista de estudantes:



Parâmetros de correlação

i) *Covariância*: média dos produtos dos desvios nas duas variáveis (δx_i e δy_i)

$$\begin{aligned}\sigma_{xy} &= \frac{1}{N} \sum_{i=1}^N \delta x_i \delta y_i = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) (y_i - \bar{y}) \\ &= \frac{(x_1 - \bar{x}) (y_1 - \bar{y}) + \dots + (x_N - \bar{x}) (y_N - \bar{y})}{N}\end{aligned}$$

Parâmetros de correlação

i) *Covariância*: média dos produtos dos desvios nas duas variáveis (δx_i e δy_i)

$$\begin{aligned}\sigma_{xy} &= \frac{1}{N} \sum_{i=1}^N \delta x_i \delta y_i = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \\ &= \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + \dots + (x_N - \bar{x})(y_N - \bar{y})}{N}\end{aligned}$$

Note que a expressão para a covariância pode ser simplificada por:

$$\sigma_{xy} = \overline{xy} - \bar{x}\bar{y}$$

Parâmetros de correlação

i) *Covariância*: média dos produtos dos desvios nas duas variáveis (δx_i e δy_i)

$$\begin{aligned}\sigma_{xy} &= \frac{1}{N} \sum_{i=1}^N \delta x_i \delta y_i = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \\ &= \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + \dots + (x_N - \bar{x})(y_N - \bar{y})}{N}\end{aligned}$$

Note que a expressão para a covariância pode ser simplificada por:

$$\sigma_{xy} = \overline{xy} - \bar{x}\bar{y}$$

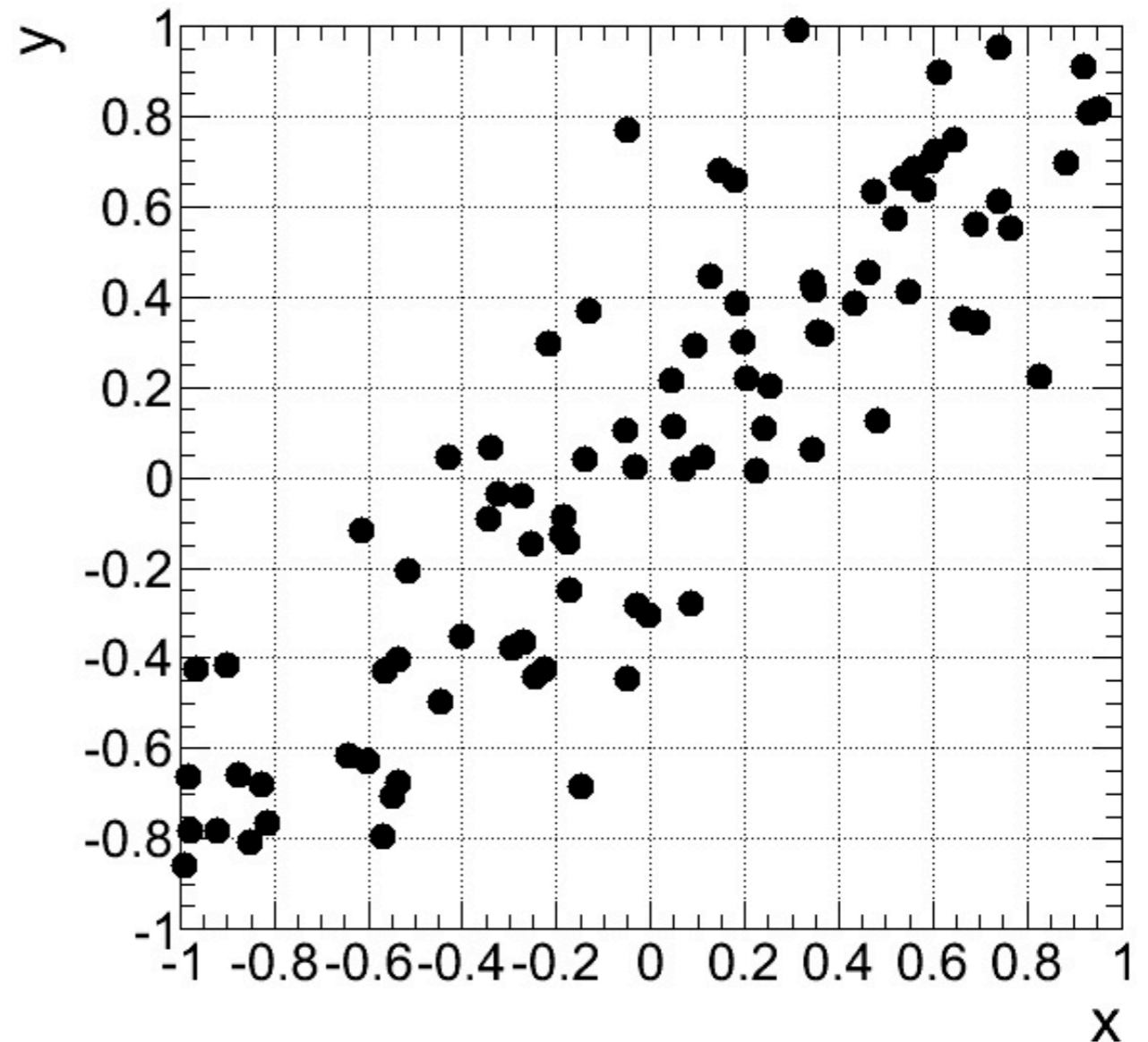
e que não importa a ordem das variáveis:

$$\sigma_{xy} = \sigma_{yx}$$

Parâmetros de correlação: covariância

Covariância:

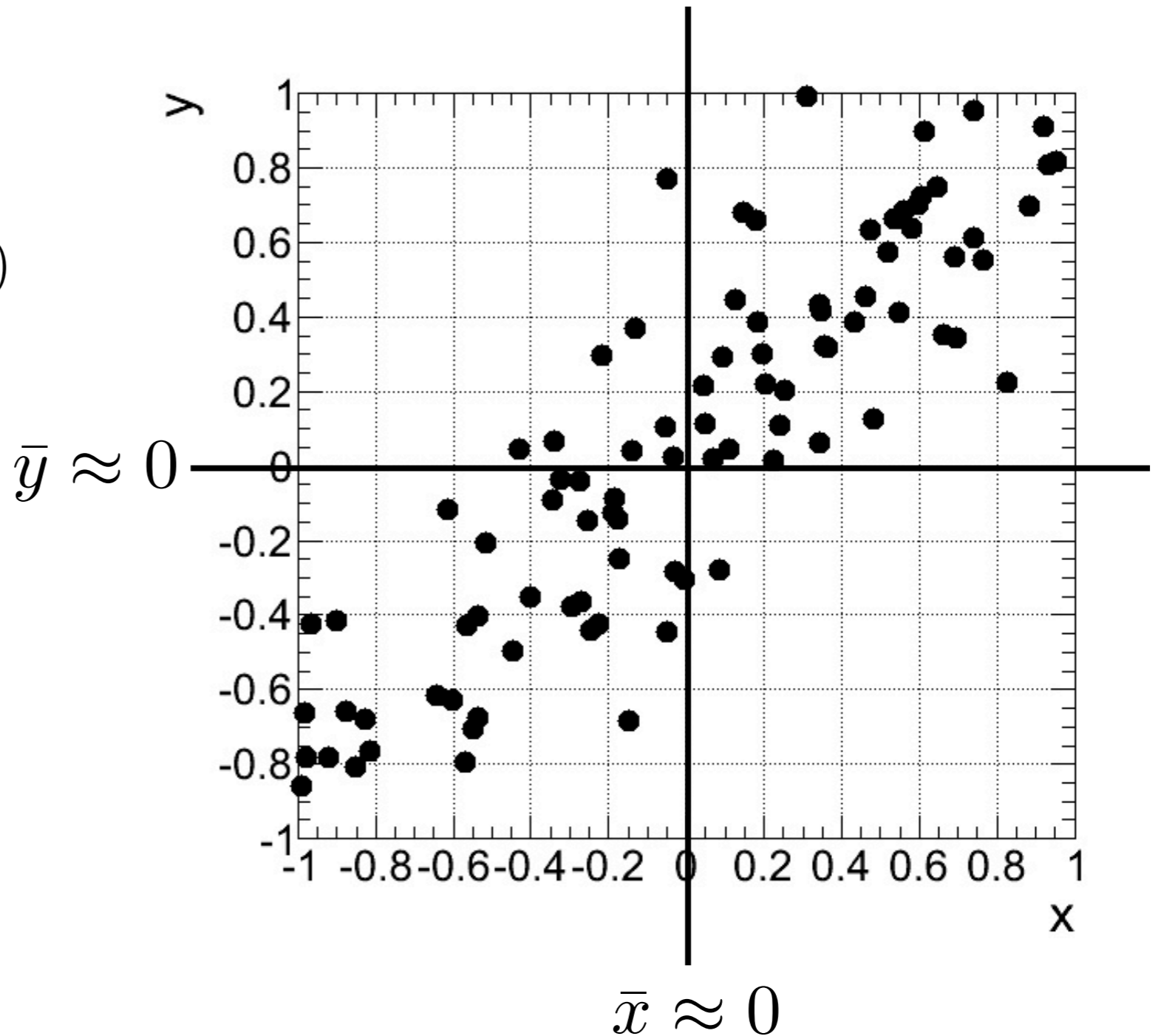
$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$



Parâmetros de correlação: covariância

Covariância:

$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

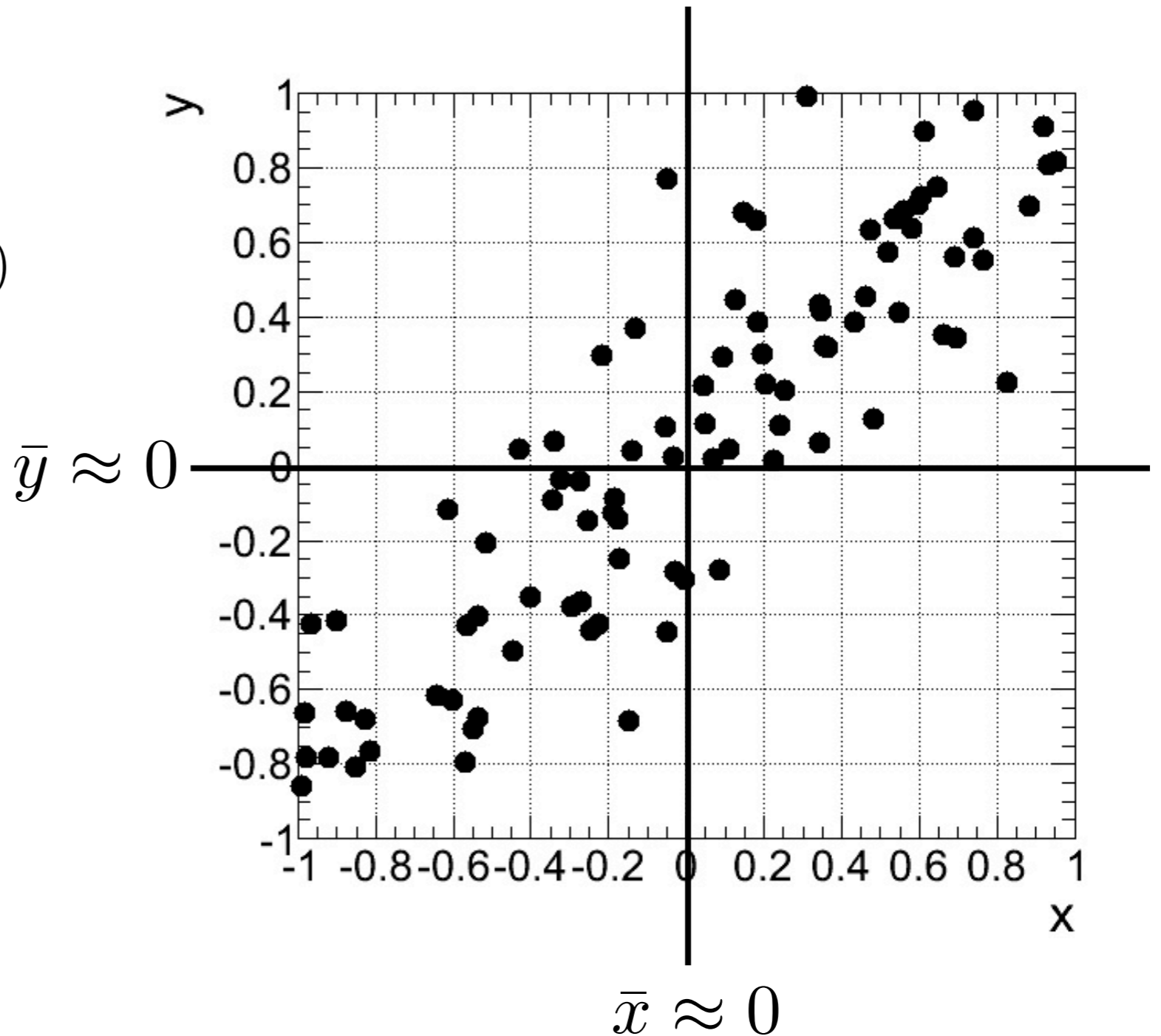


Parâmetros de correlação: covariância

Covariância:

$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

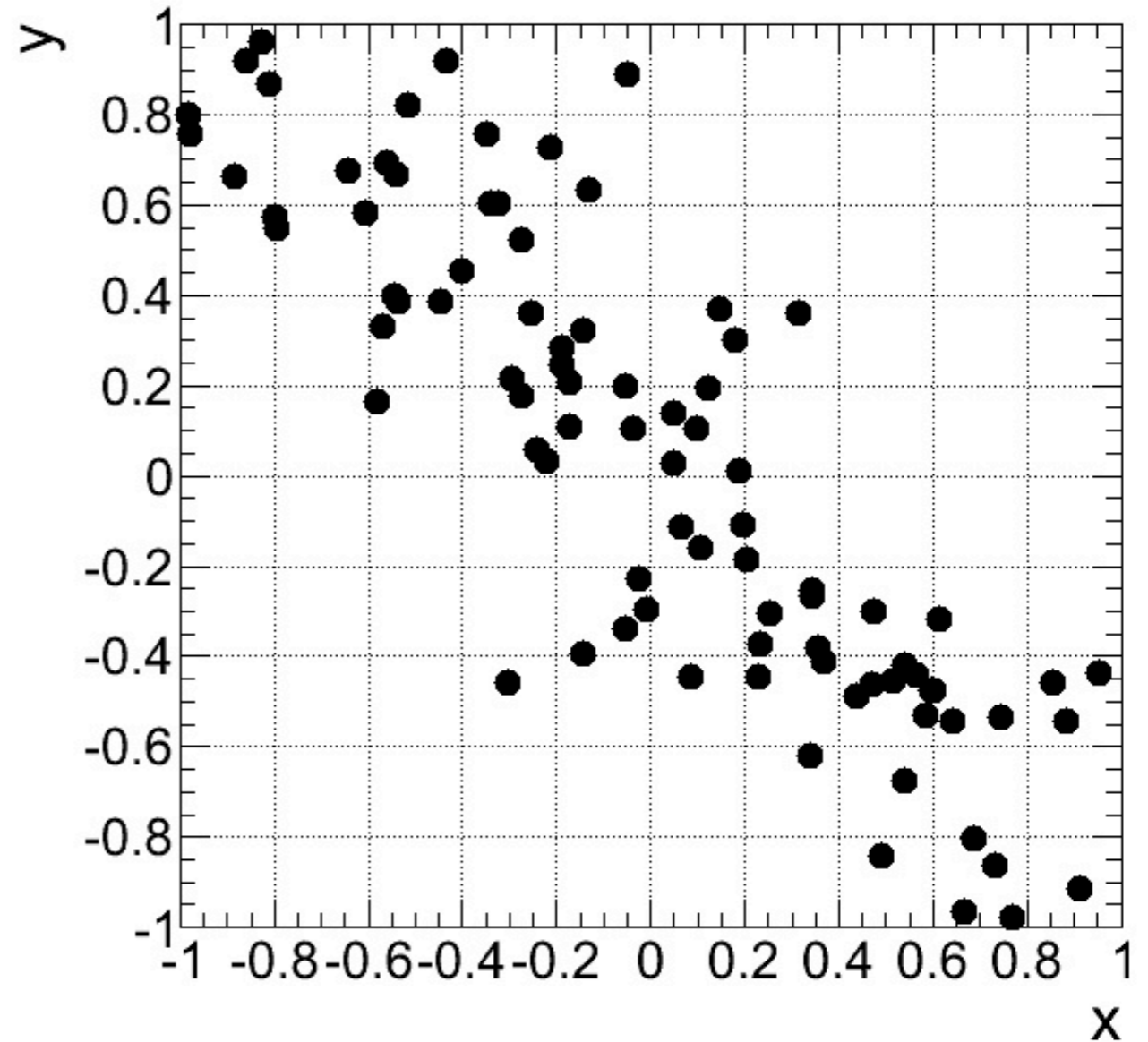
➔ $\sigma_{xy} > 0$



Parâmetros de correlação: covariância

Covariância:

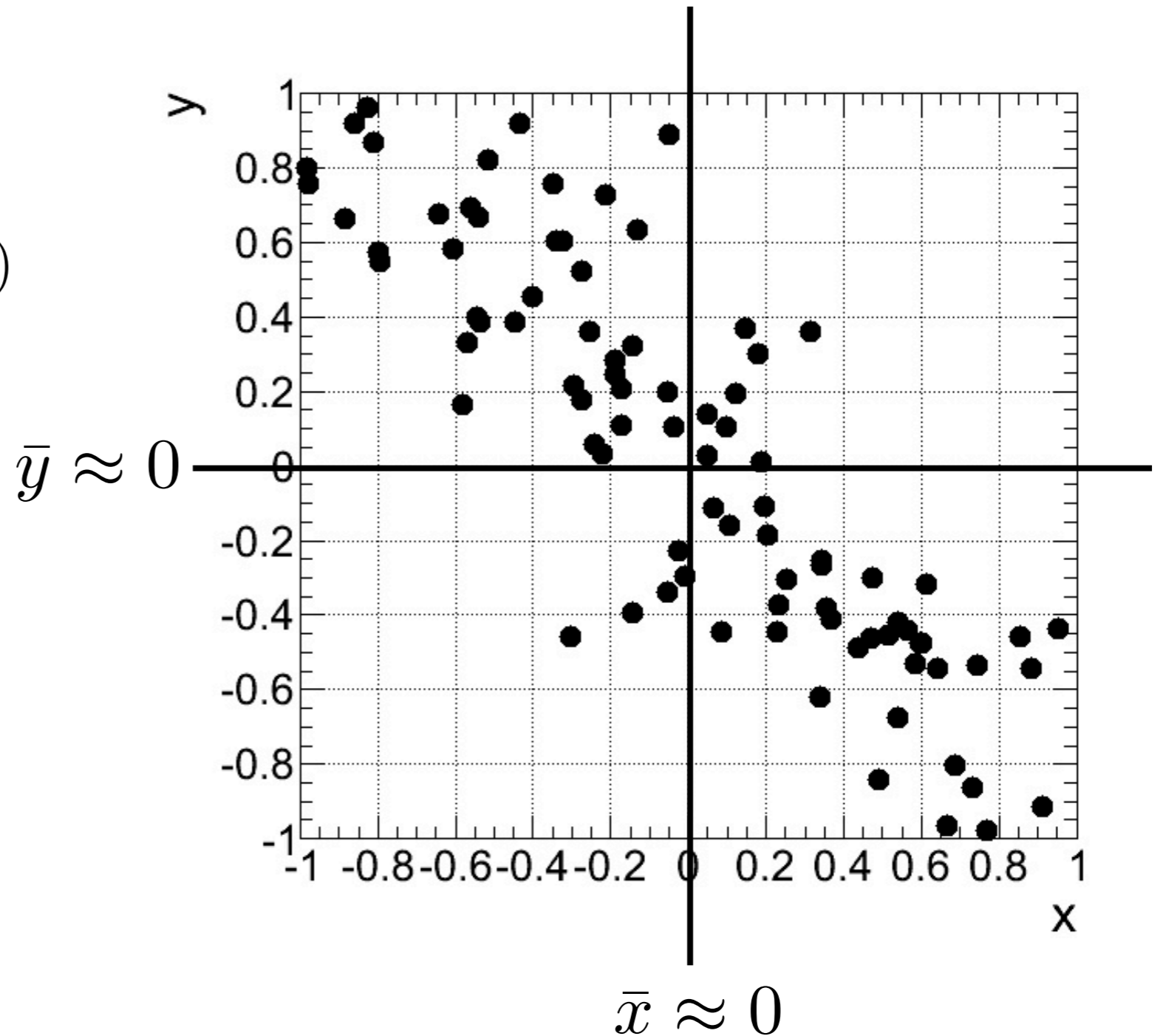
$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$



Parâmetros de correlação: covariância

Covariância:

$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

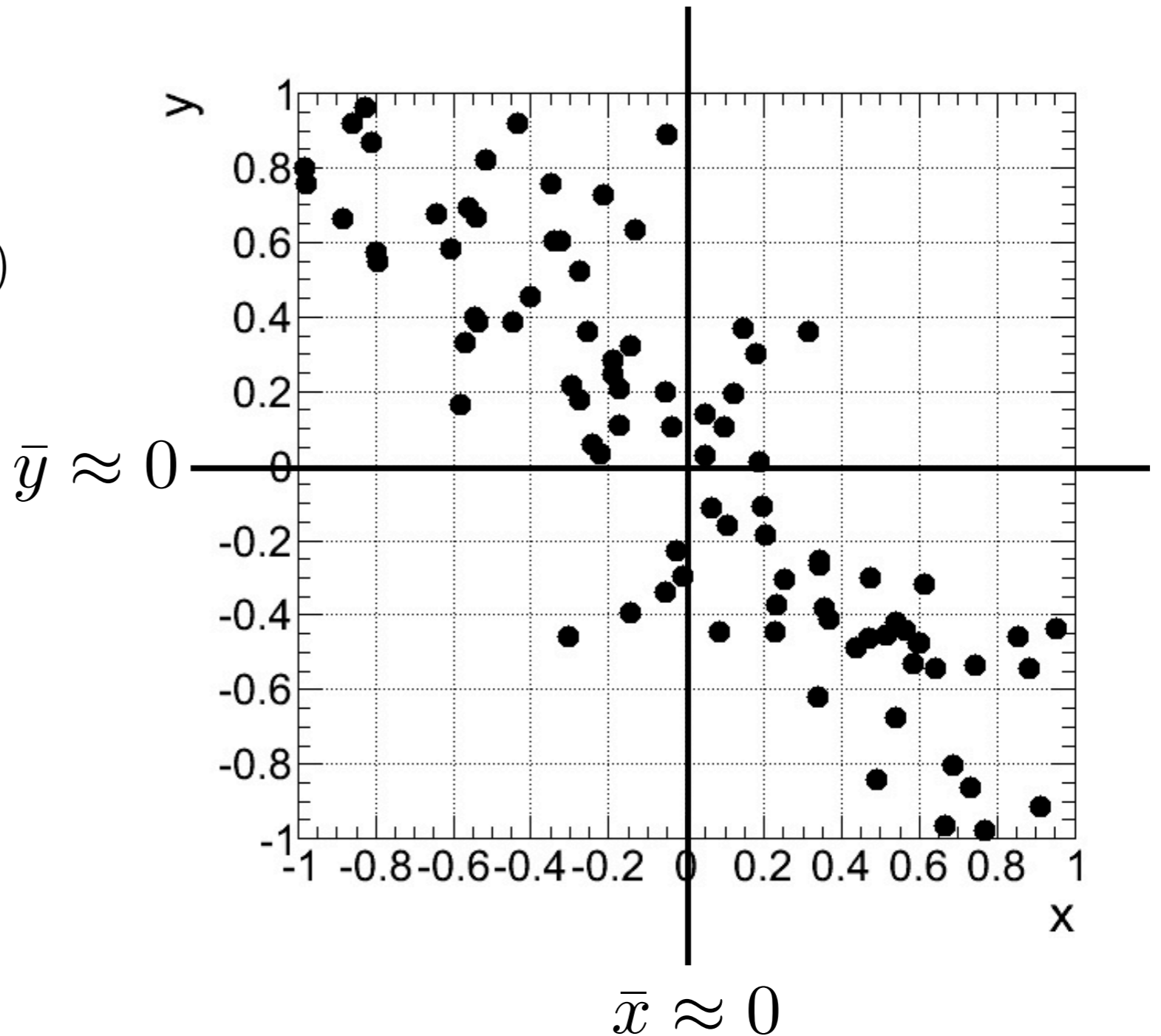


Parâmetros de correlação: covariância

Covariância:

$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

➔ $\sigma_{xy} < 0$



Parâmetros de correlação

ii) *Coeficiente de correlação linear de Pearson*: covariância entre duas variáveis, dividida por seus desvios padrão

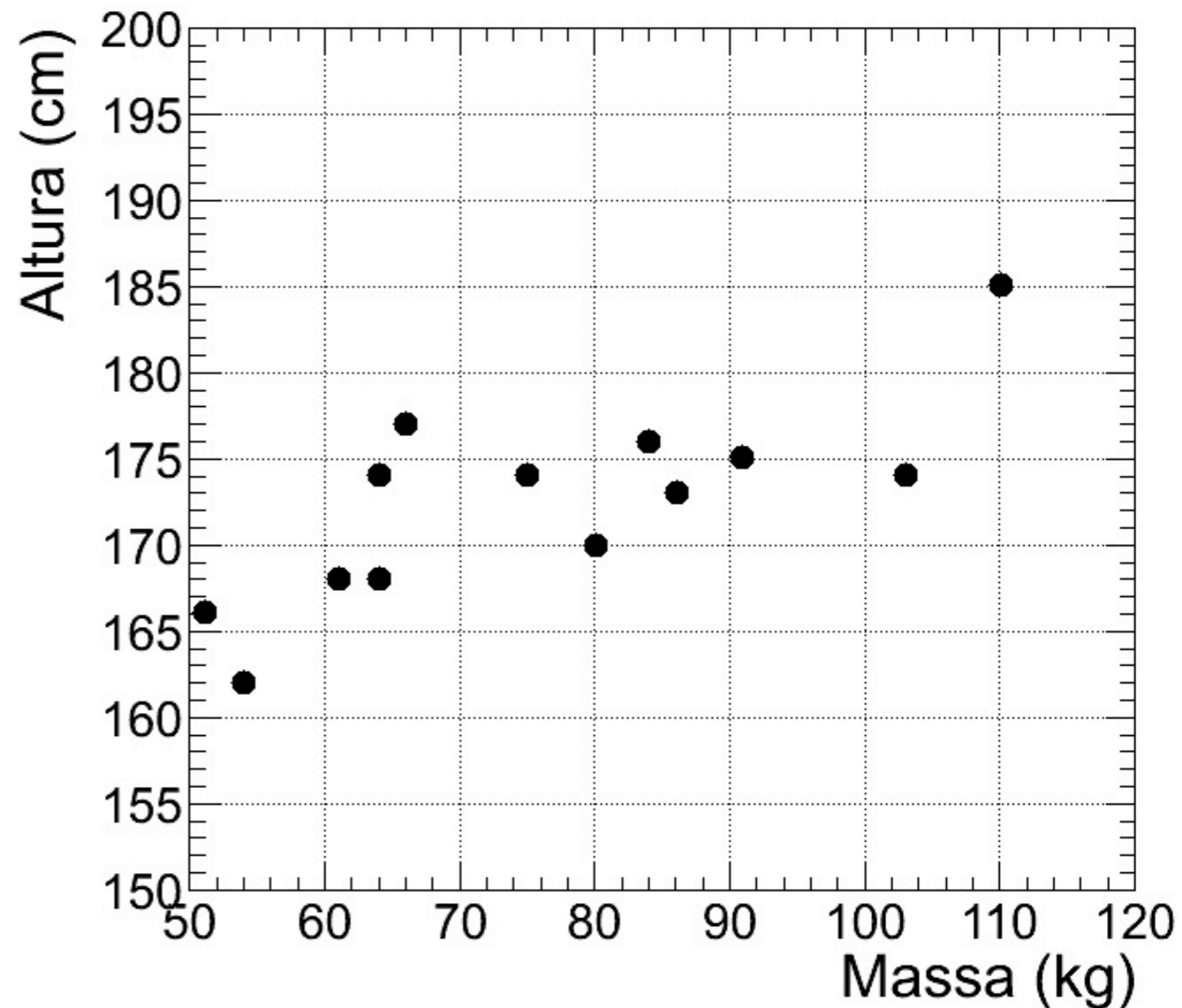
$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad -1 \leq r \leq 1$$

Correlação linear, perfeita e positiva: $r = 1$

Correlação linear, perfeita e negativa: $r = -1$

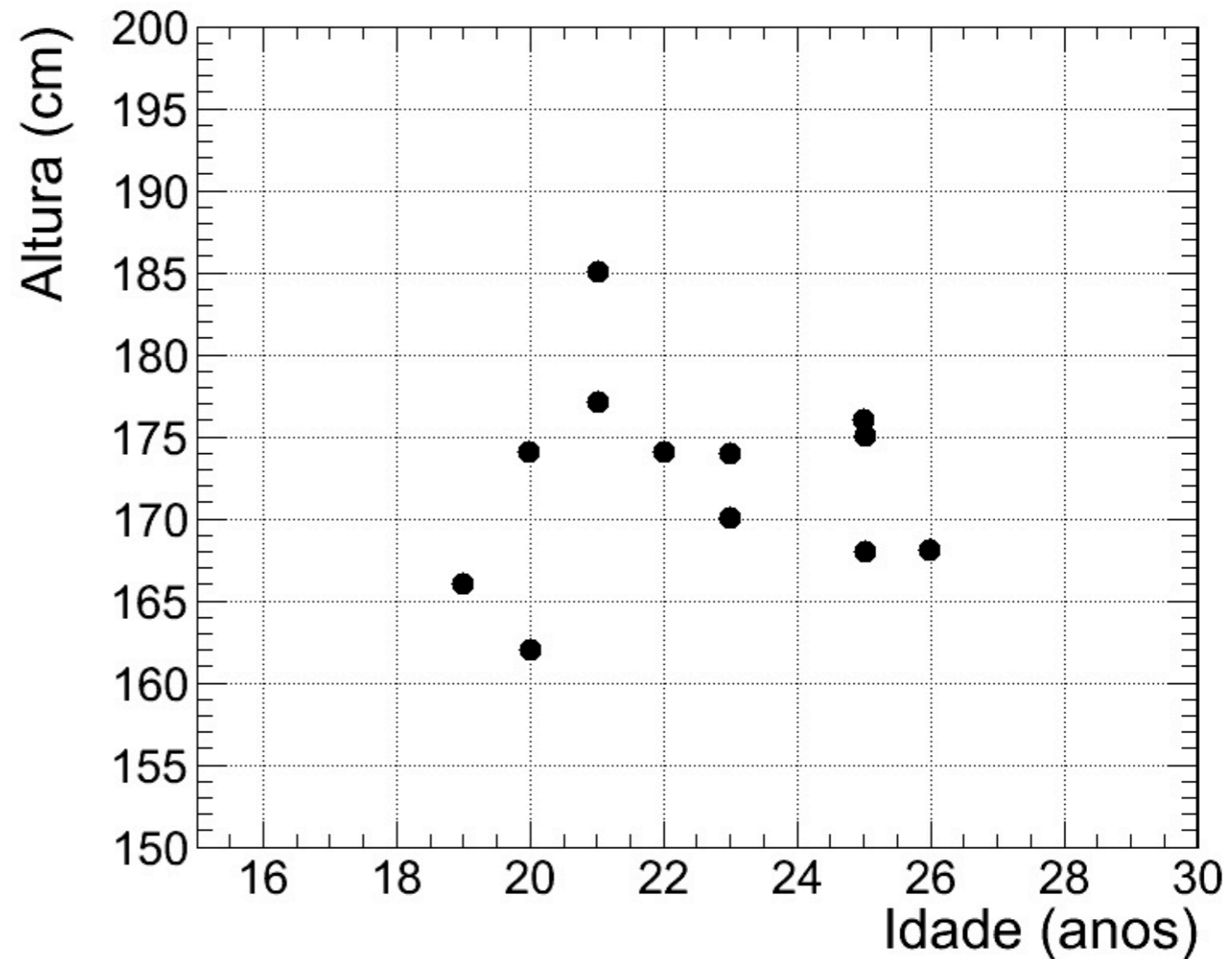
Atividade de aula - Idade, massa e altura

Estudante	Massa (kg)	Altura (cm)
1	64	174
2	110	185
3	75	174
4	80	170
5	61	168
6	86	173
7	54	162
8	84	176
9	51	166
10	64	168
11	66	177
12	103	174
13	91	175



Atividade de aula - Idade, massa e altura

Estudante	Idade (anos)	Altura (cm)
1	22	174
2	21	185
3	20	174
4	23	170
5	25	168
6	30	173
7	20	162
8	25	176
9	19	166
10	26	168
11	21	177
12	23	174
13	25	175



Atividade de aula - Idade, massa e altura

Estudante	Idade (anos)	Massa (kg)
1	22	64
2	21	110
3	20	75
4	23	80
5	25	61
6	30	86
7	20	54
8	25	84
9	19	51
10	26	64
11	21	66
12	23	103
13	25	91

