

Física Geral - Laboratório (2013/2)

Aula 2: Organização e descrição de dados e
parâmetros de dispersão e correlação



Lembrando: Bibliografia



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

Resumo: conjuntos de dados

Idades dos estudantes:

{18; 19; 18} (anos)

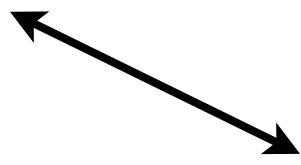
Medidas do comprimento de uma mesa:

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

Tipo sanguíneo dos estudantes de FG:

{'O-'; 'A-'; 'O+'}

...



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

Resumo: organizando conjuntos de dados em Histogramas

Para um conjunto de dados (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)

Resumo: organizando conjuntos de dados em Histogramas

Para um conjunto de dados (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)

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

Resumo: organizando conjuntos de dados em Histogramas

Para um conjunto de dados (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)

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

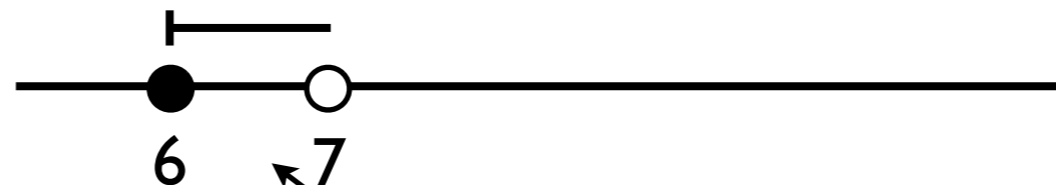


Classe de idades (anos)	Frequências
[6 , 7)	1
[7 , 8)	3
[8 , 9)	3
[9 , 10)	3
[10 , 11)	6
[11 , 12)	1
[12 , 13)	3
[13 , 14)	1
[14 , 15)	2
[15 , 16)	1

Resumo: organizando conjuntos de dados em Histogramas

Para um conjunto de dados (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)

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

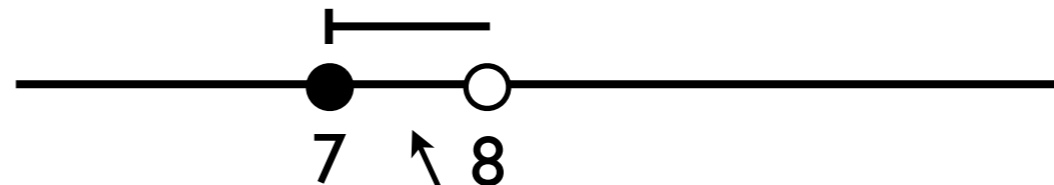


Classe de idades (anos)	Frequências
[6 , 7)	1
[7 , 8)	3
[8 , 9)	3
[9 , 10)	3
[10 , 11)	6
[11 , 12)	1
[12 , 13)	3
[13 , 14)	1
[14 , 15)	2
[15 , 16)	1

Resumo: organizando conjuntos de dados em Histogramas

Para um conjunto de dados (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)

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

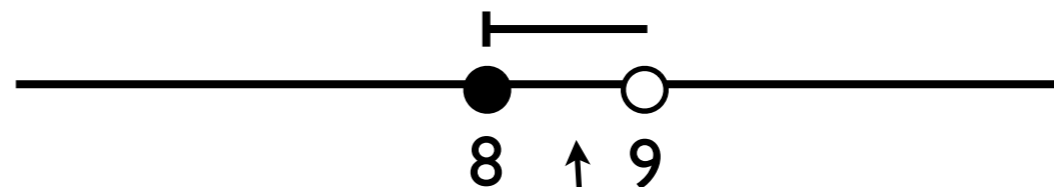


Classe de idades (anos)	Frequências
[6 , 7)	1
[7 , 8)	3
[8 , 9)	3
[9 , 10)	3
[10 , 11)	6
[11 , 12)	1
[12 , 13)	3
[13 , 14)	1
[14 , 15)	2
[15 , 16)	1

Resumo: organizando conjuntos de dados em Histogramas

Para um conjunto de dados (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)

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



Classe de idades (anos)	Frequências
$[6, 7)$	1
$[7, 8)$	3
$[8, 9)$	3
$[9, 10)$	3
$[10, 11)$	6
$[11, 12)$	1
$[12, 13)$	3
$[13, 14)$	1
$[14, 15)$	2
$[15, 16)$	1

Resumo: organizando conjuntos de dados em Histogramas

Para um conjunto de dados (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)

Escolha 1:

Classe de idades (anos)	Frequências
[6 , 7)	1
[7 , 8)	3
[8 , 9)	3
[9 , 10)	3
[10 , 11)	6
[11 , 12)	1
[12 , 13)	3
[13 , 14)	1
[14 , 15)	2
[15 , 16)	1

Resumo: organizando conjuntos de dados em Histogramas

Para um conjunto de dados (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)

Escolha 1:

Classe de idades (anos)	Frequências
[6 , 7)	1
[7 , 8)	3
[8 , 9)	3
[9 , 10)	3
[10 , 11)	6
[11 , 12)	1
[12 , 13)	3
[13 , 14)	1
[14 , 15)	2
[15 , 16)	1

Escolha 2:

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

Resumo: organizando conjuntos de dados em 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)

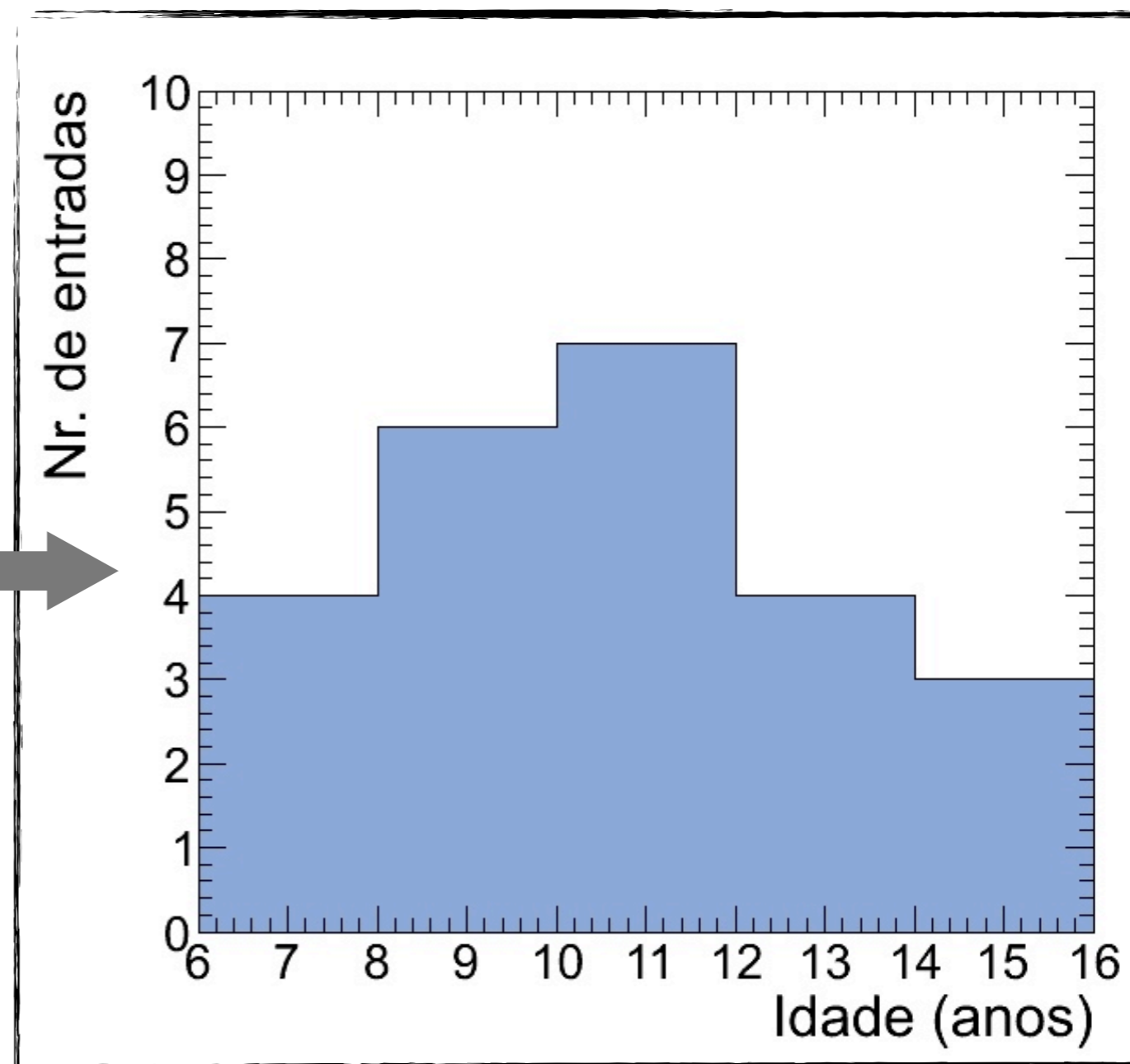
Classe de idades (anos)	Frequência
[6 , 8)	4
[8 , 10)	6
[10 , 12)	7
[12 , 14)	4
[14 , 16)	3

Resumo: organizando conjuntos de dados em 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



Resumo: parâmetros de posição

i) Média:

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

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

Dados em M classes (intervalos) com ponto médio $\{x_1, x_2, \dots, x_M\}$ e frequência $\{n_1, n_2, \dots, n_M\}$:

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

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

iii) *Média quadrática*:

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

iv) *Mediana* (Mesma quantidade de dados abaixo e acima da mediana):

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

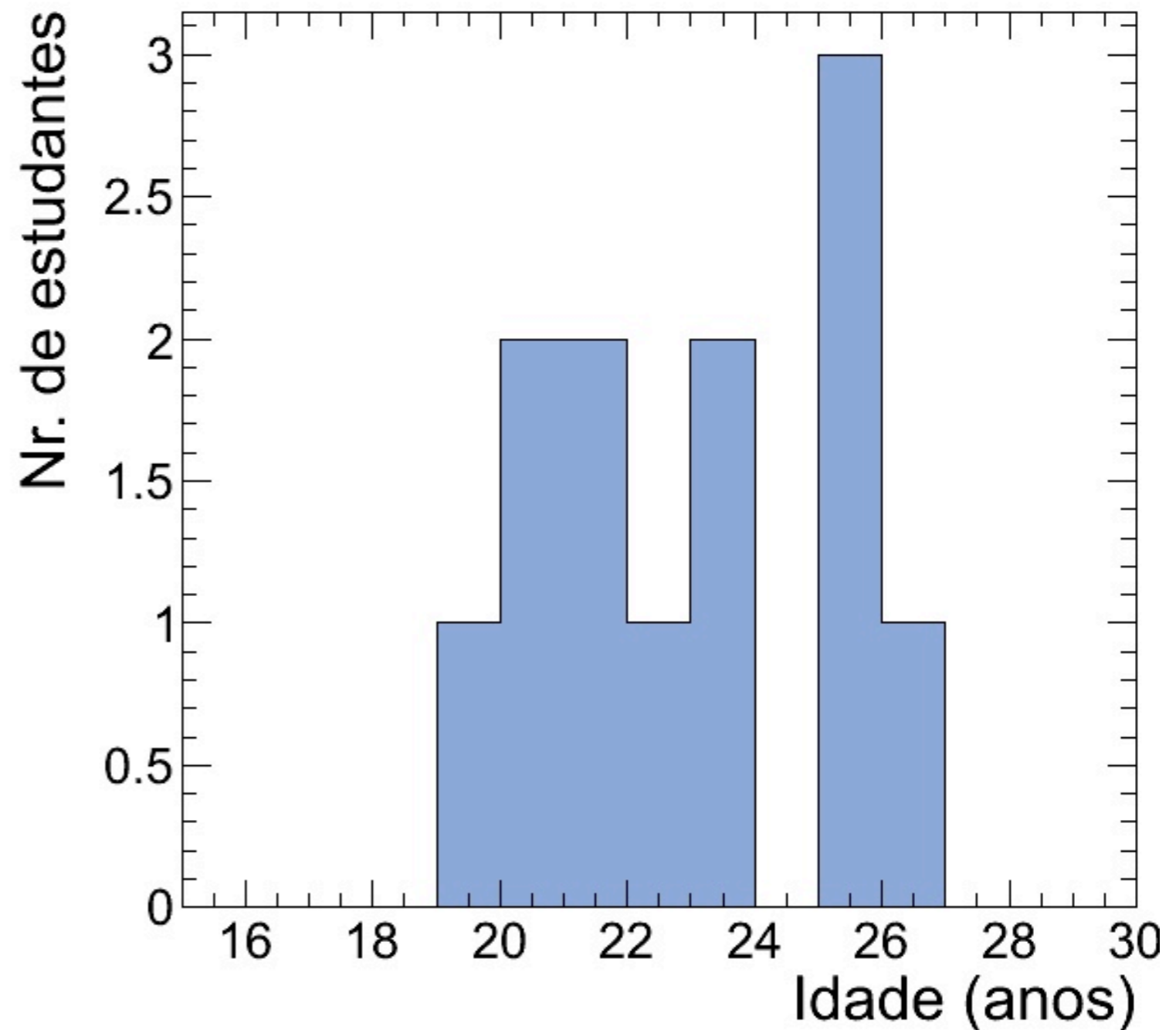
$$N(\text{par}) \rightarrow x_{\text{med}} = \frac{x_{N/2} + x_{(N/2+1)}}{2}$$

Atividade - Aula 1

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

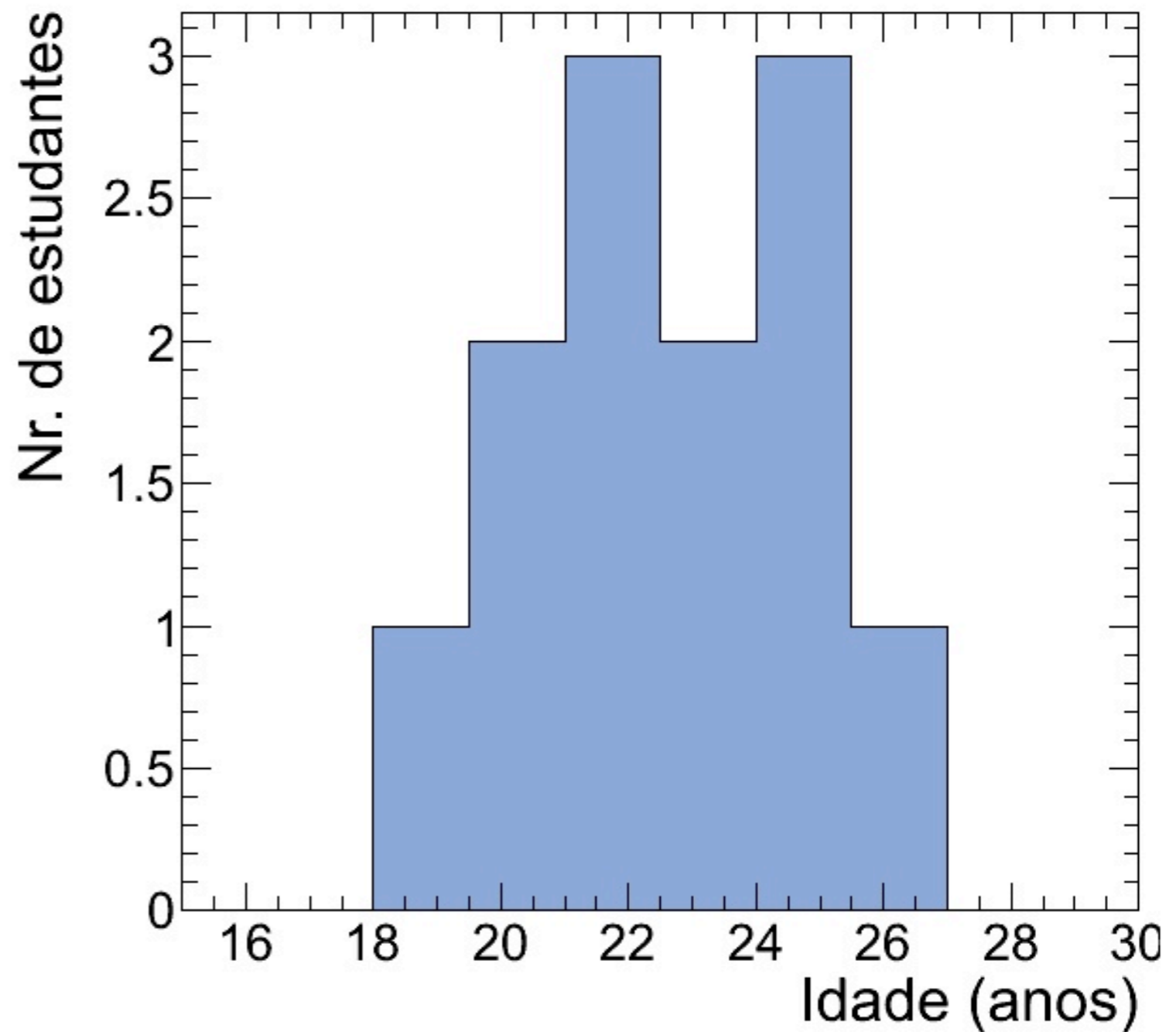
Atividade - Aula I

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



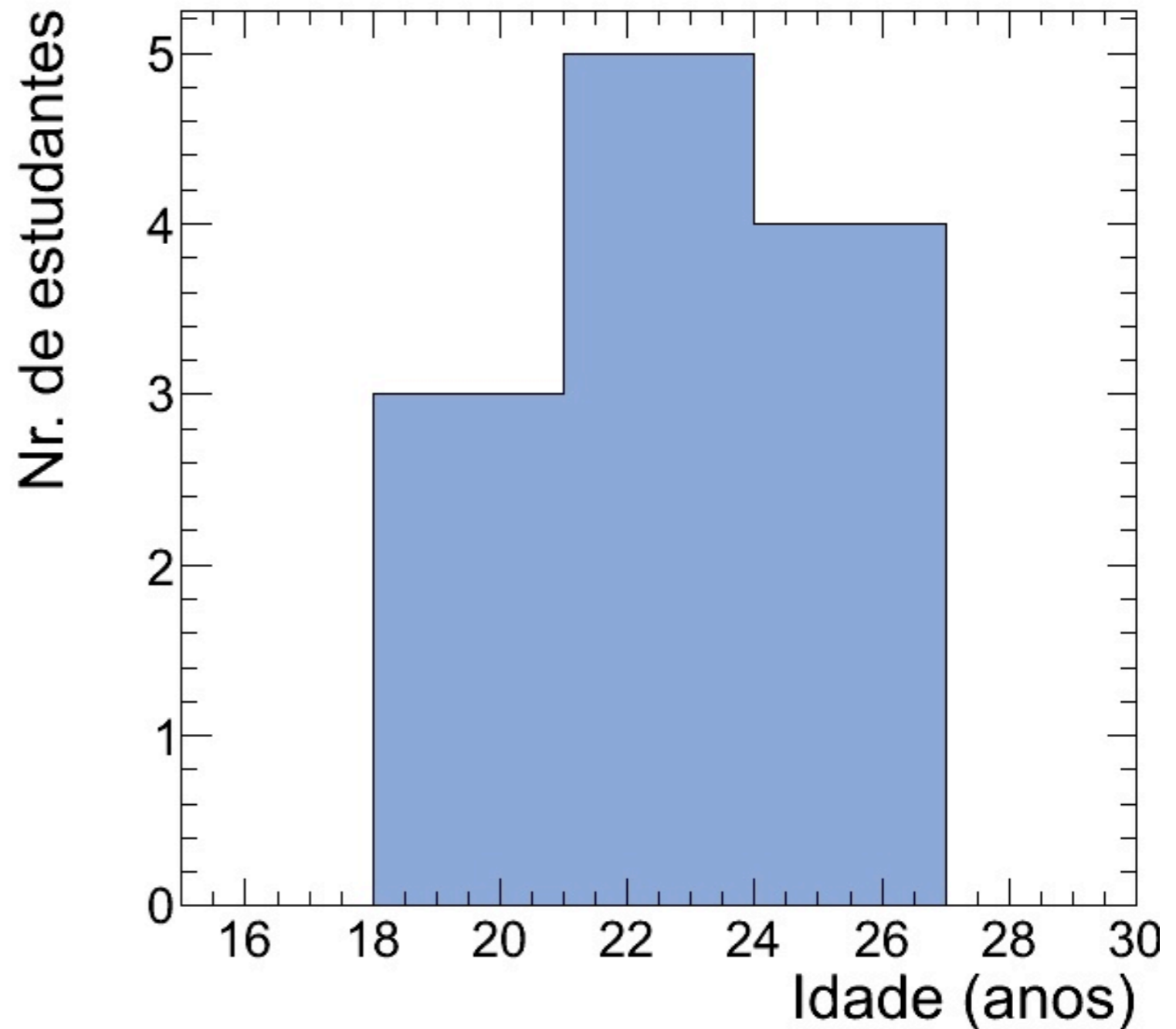
Atividade - Aula 1

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



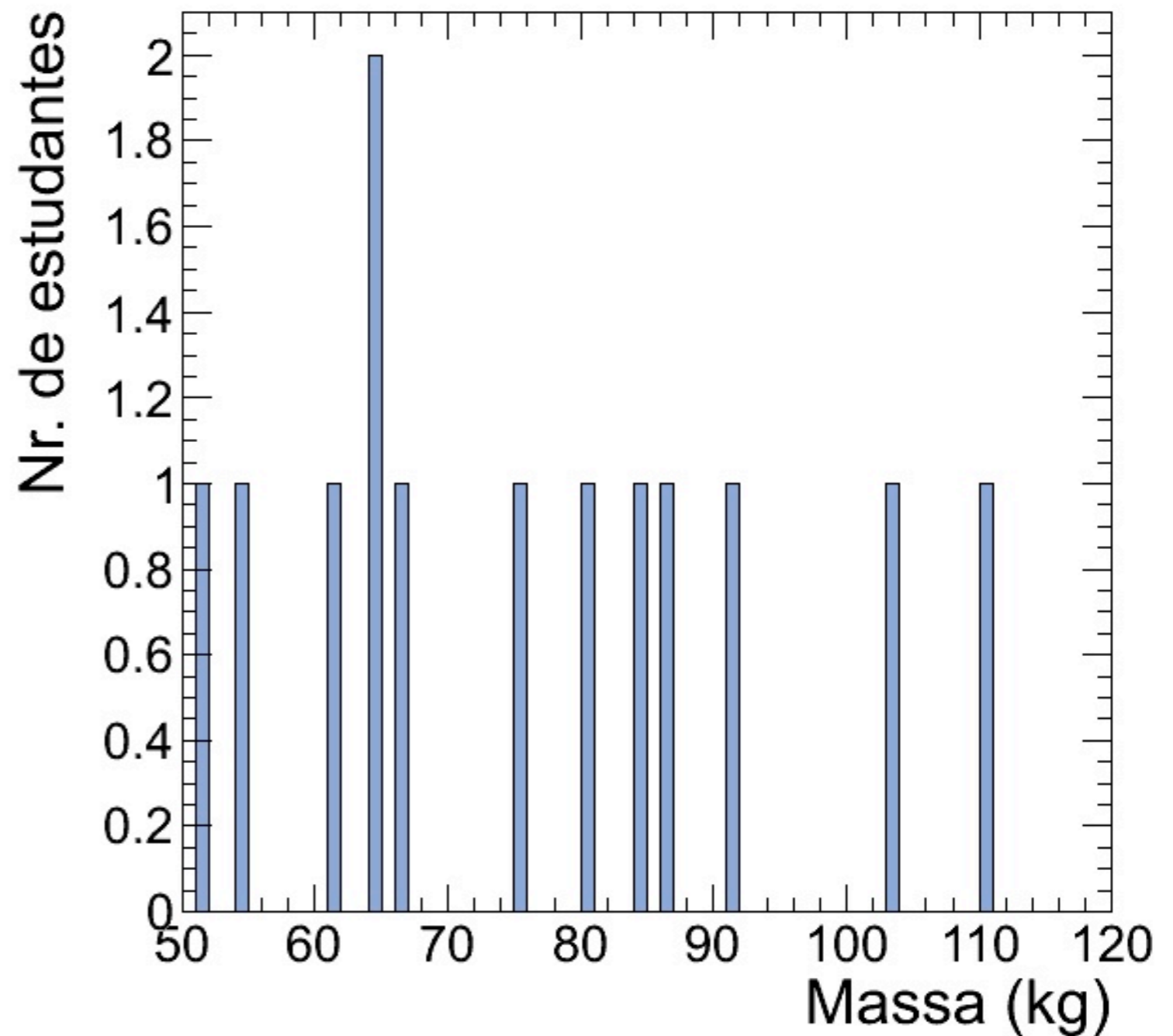
Atividade - Aula I

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



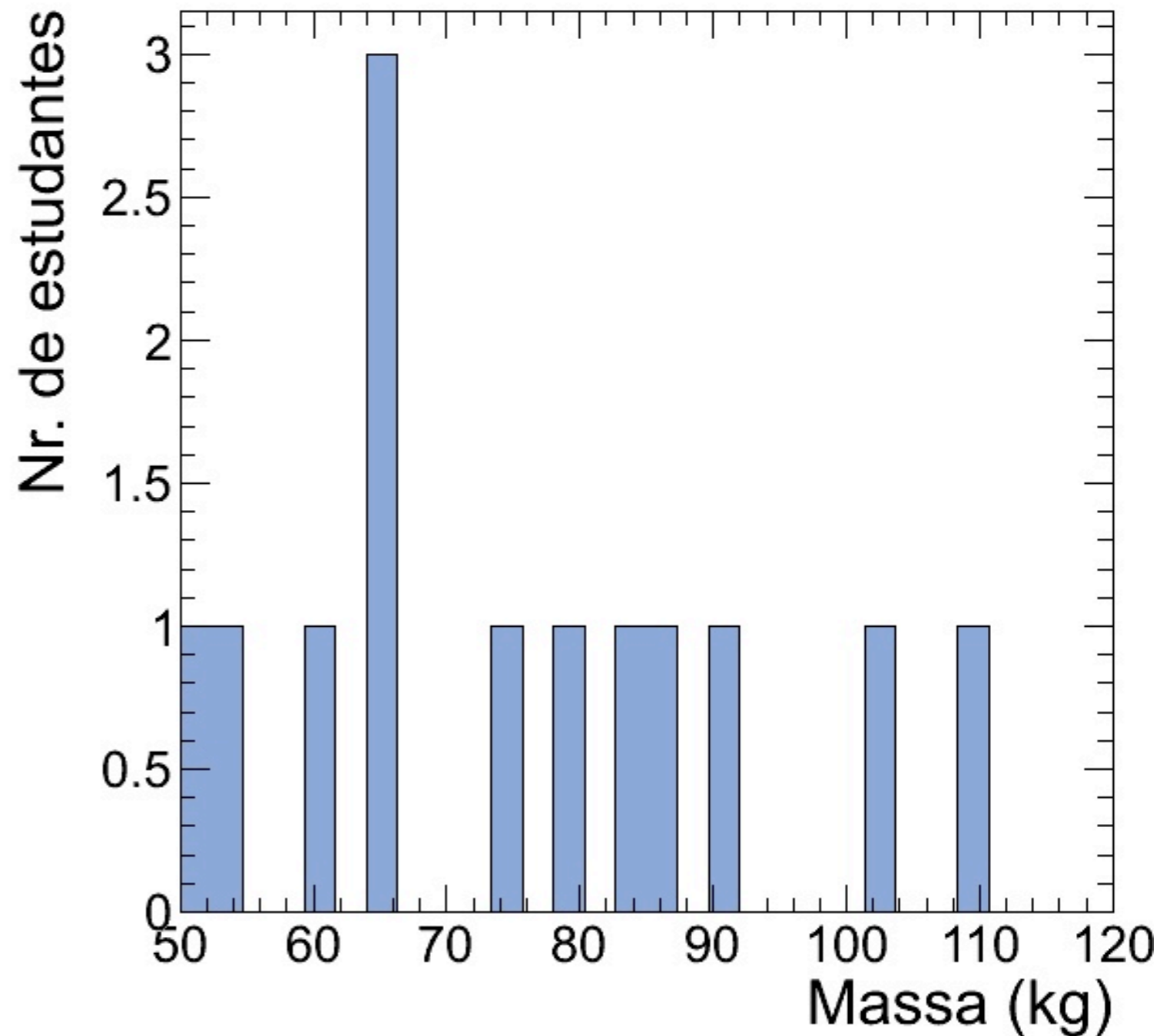
Atividade - Aula 1

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



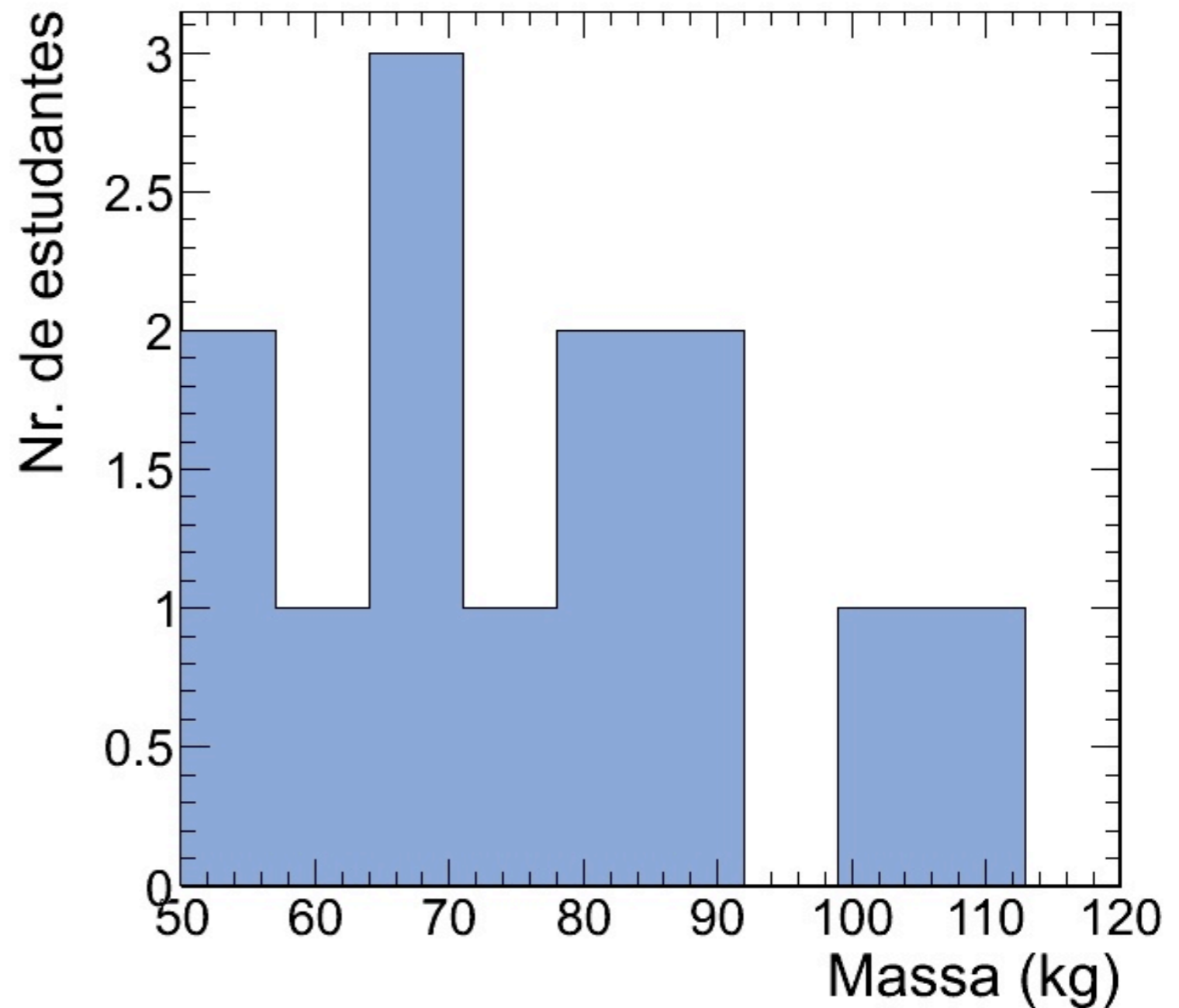
Atividade - Aula 1

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



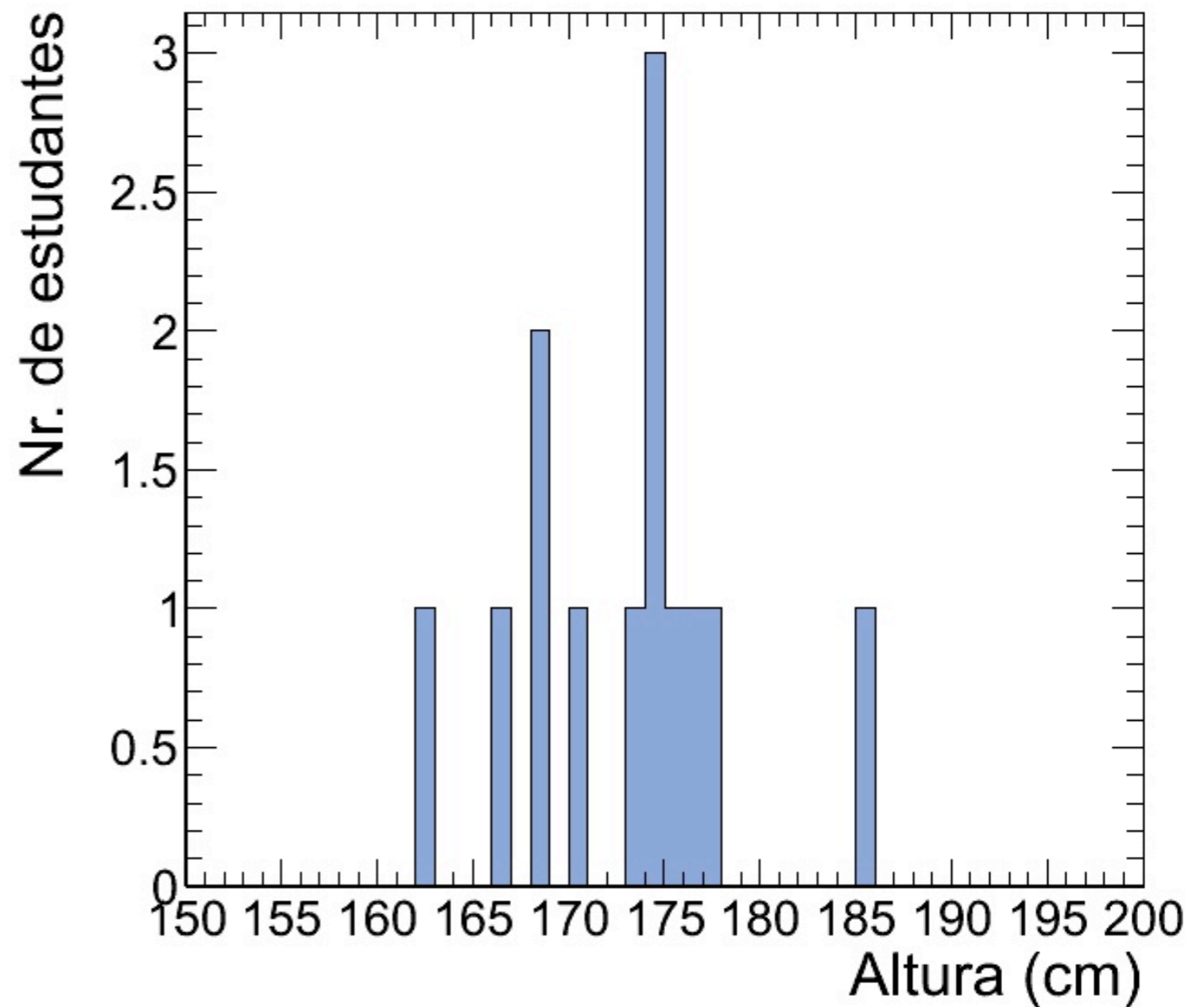
Atividade - Aula 1

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



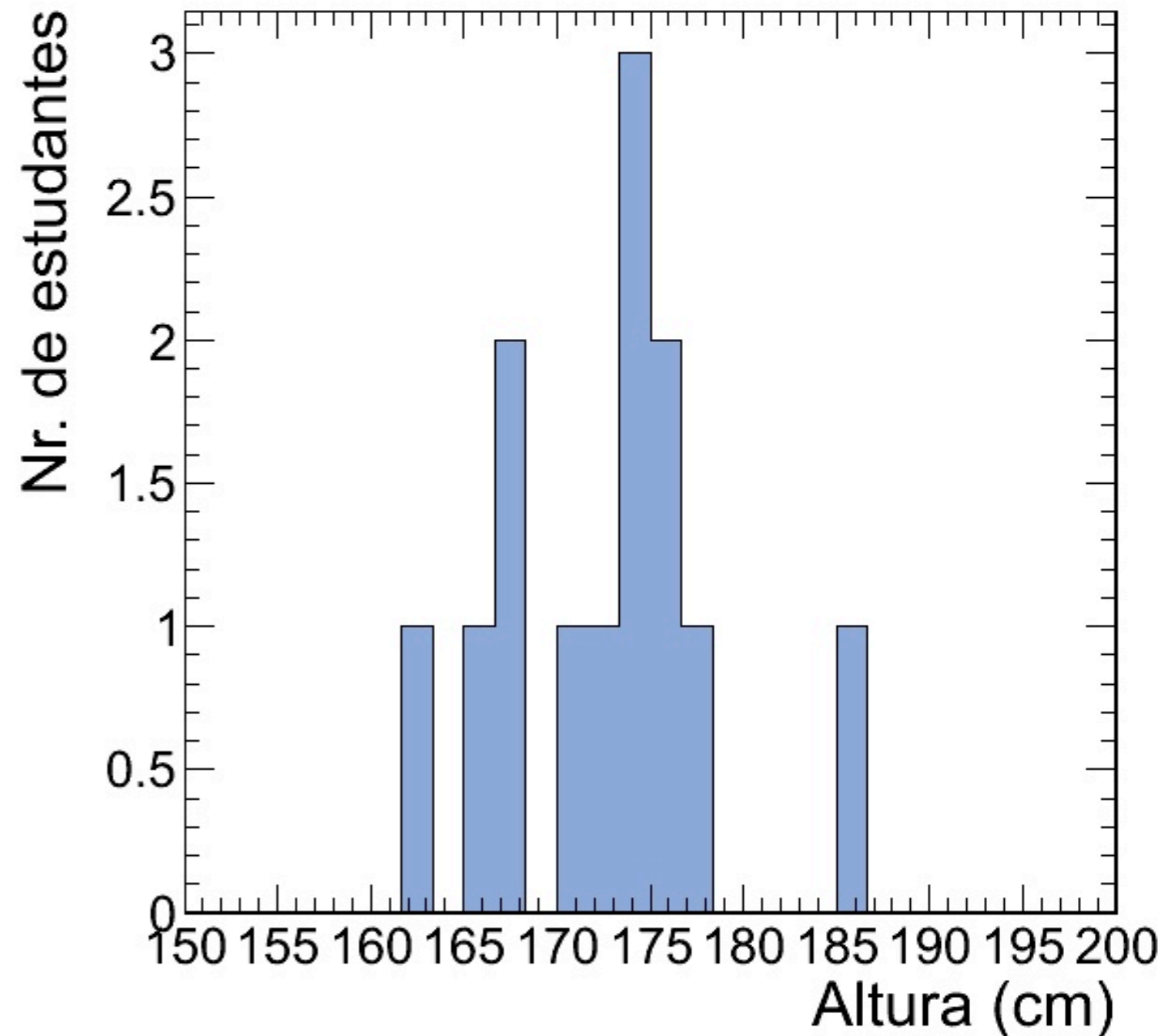
Atividade - Aula 1

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 - Aula 1

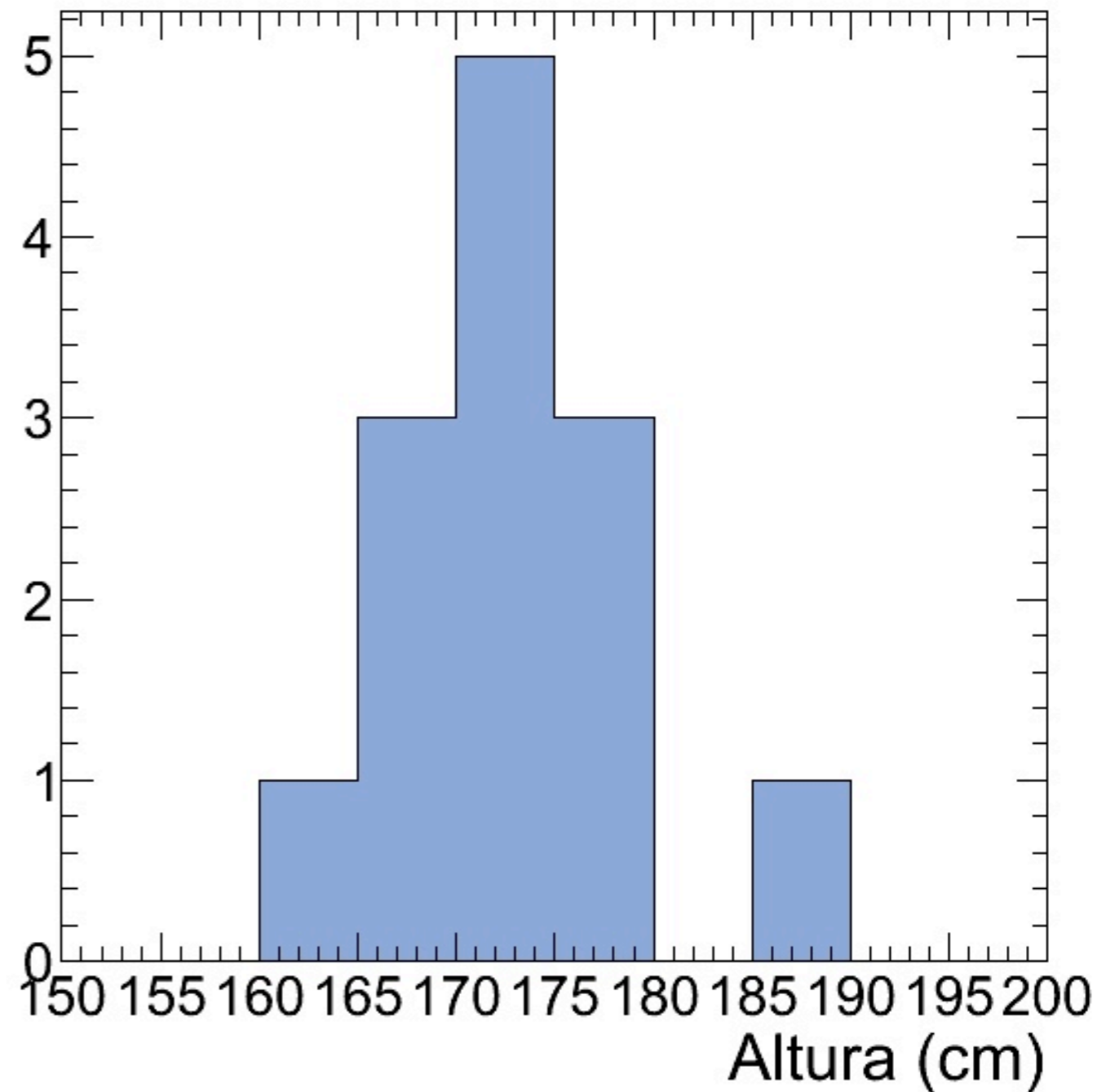
Estudante	Altura (cm)
1	174
2	185
3	174
4	170
5	168
6	173
7	162
8	176
9	166
10	168
11	177
12	174
13	175



Atividade - Aula 1

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

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

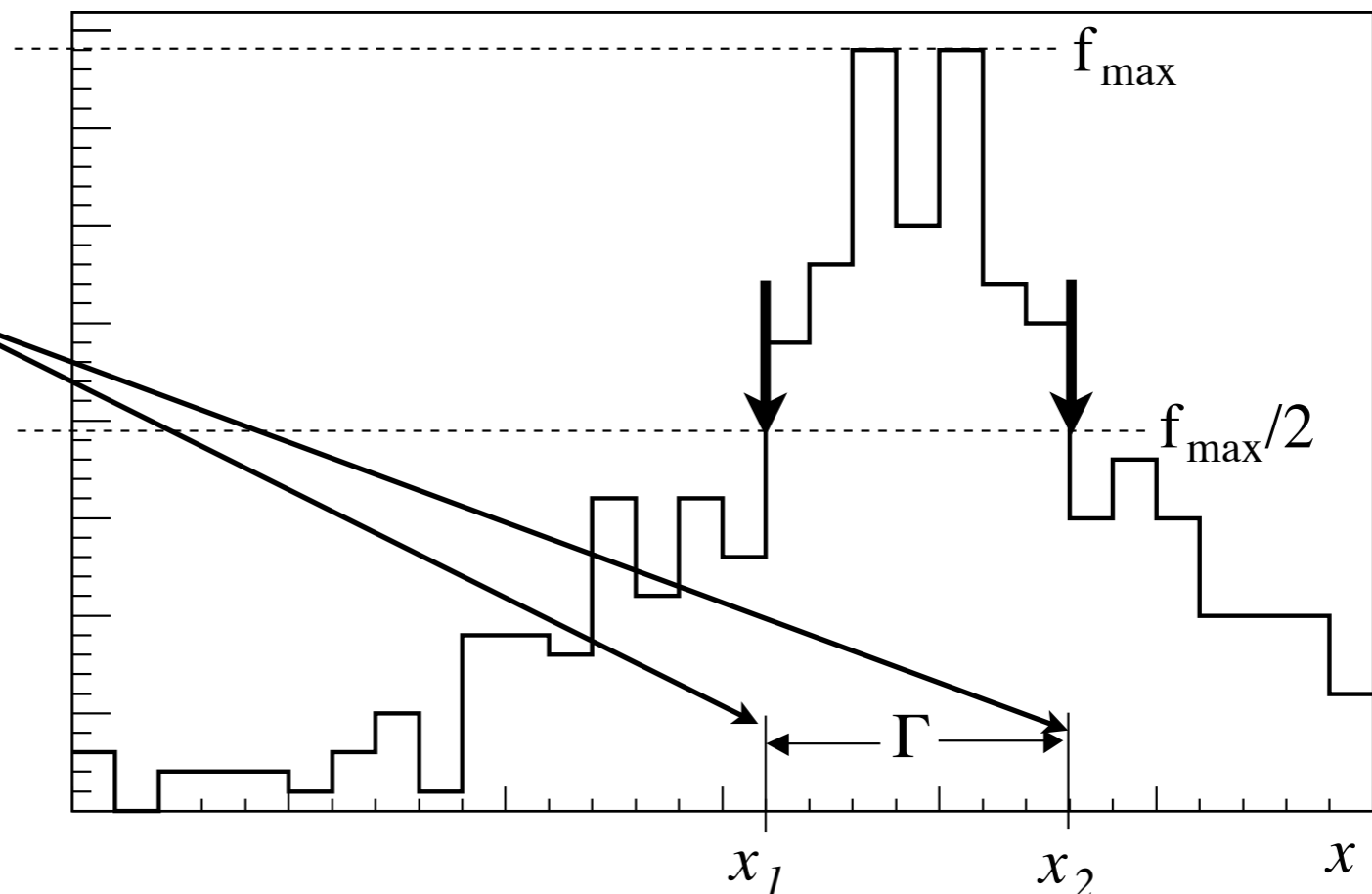

$$\sigma_x = \sqrt{\overline{x^2} - \bar{x}^2}$$

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

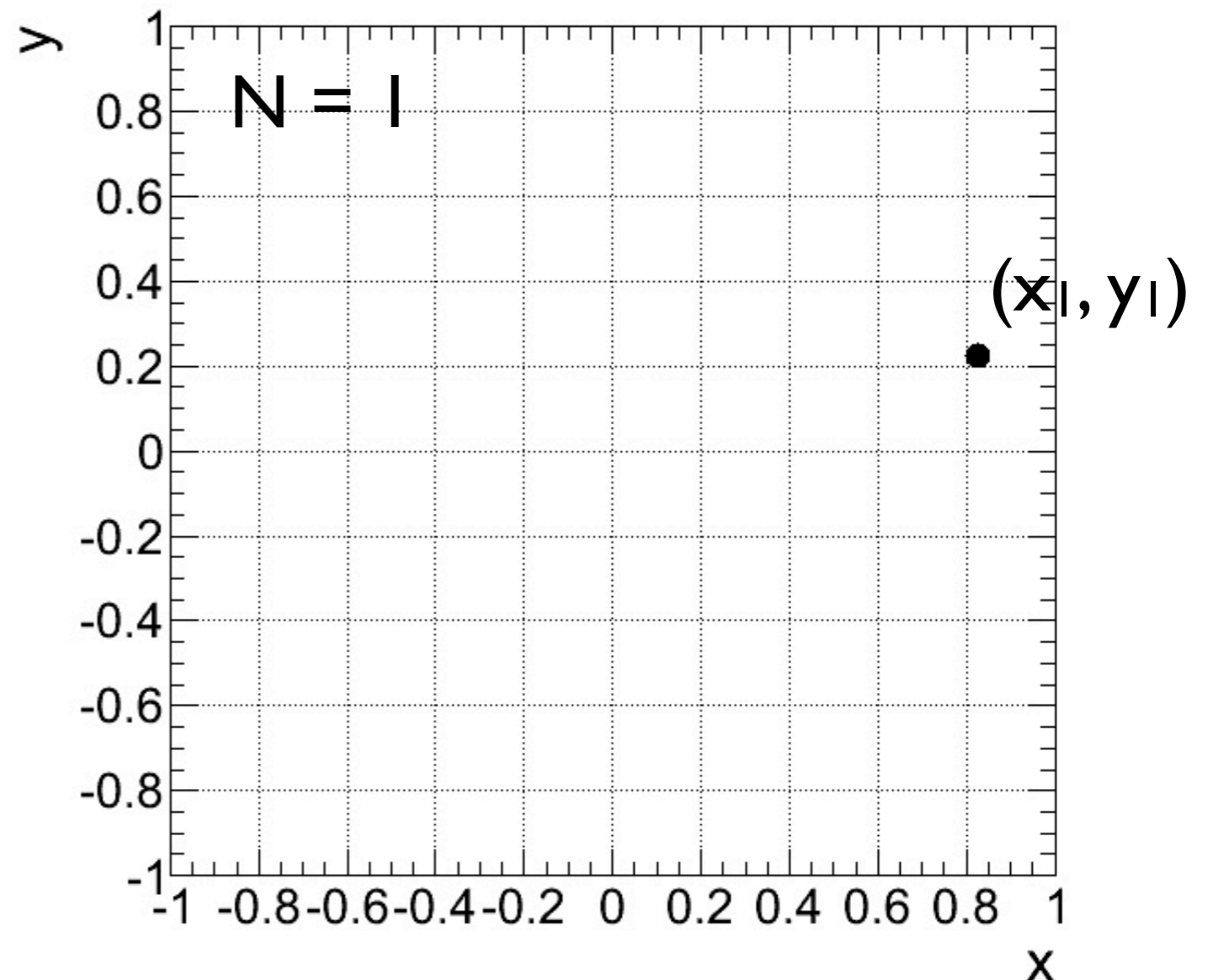
$$\Gamma = |x_2 - x_1|$$



Representando duas variáveis

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

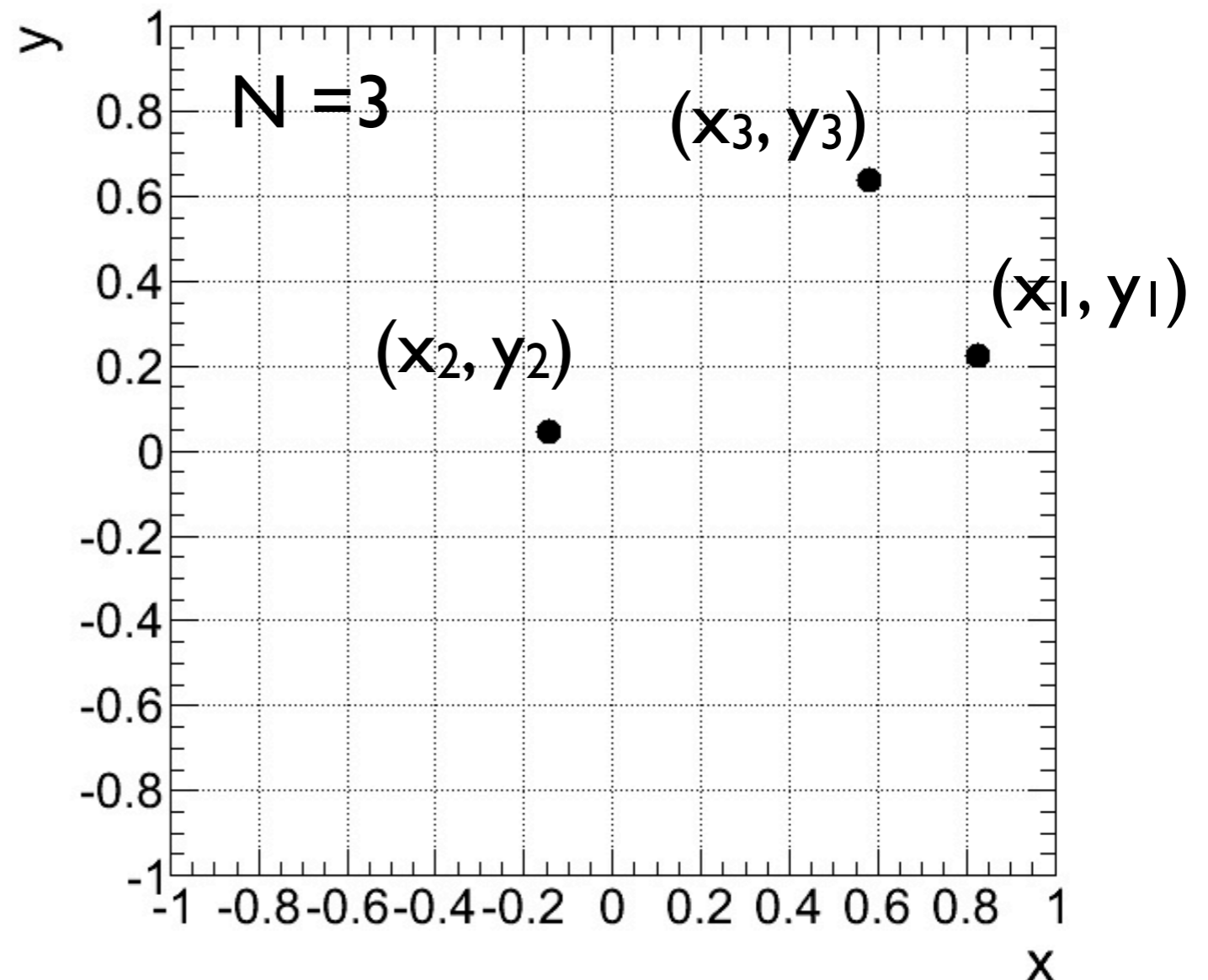
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

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

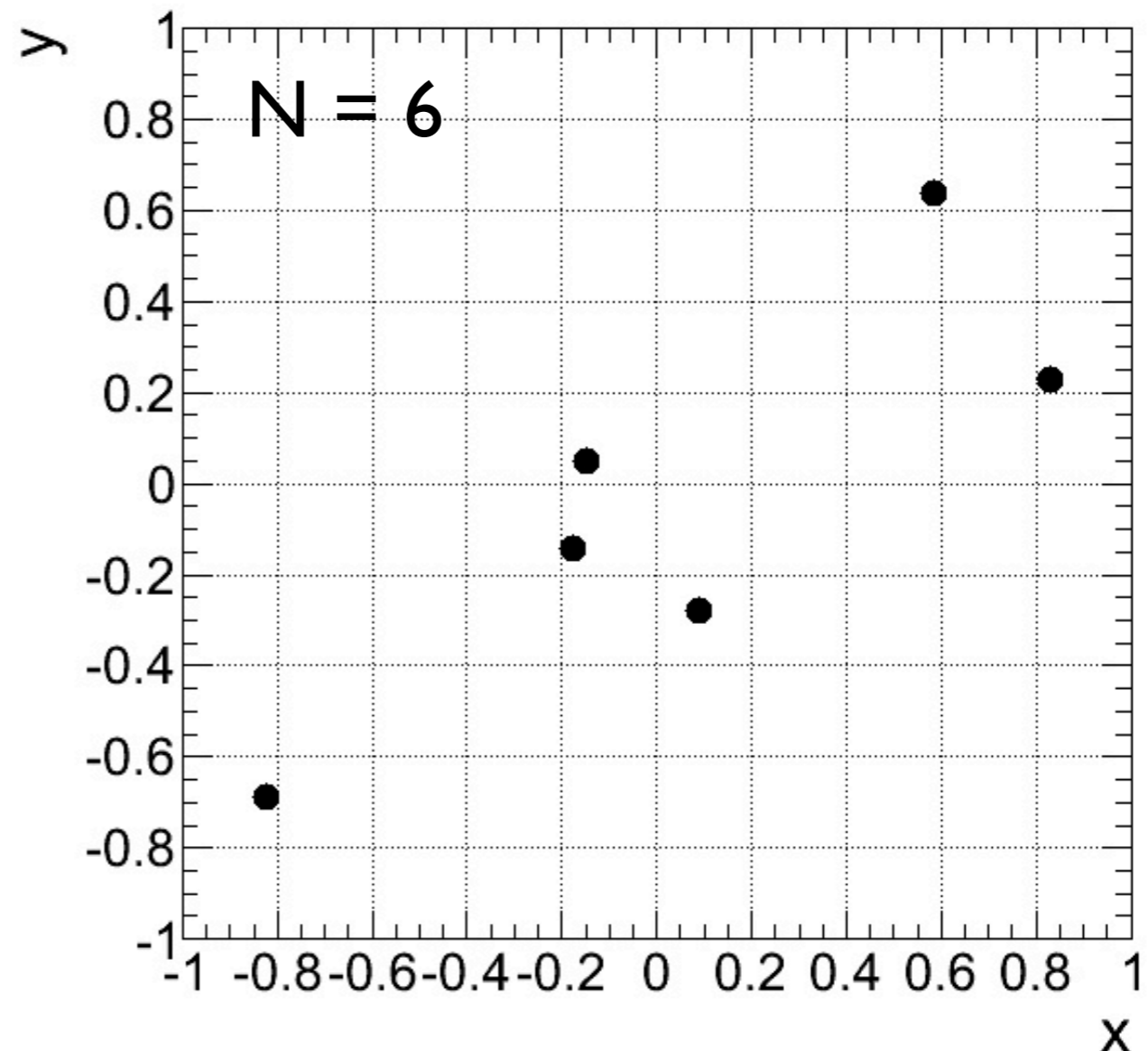
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

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

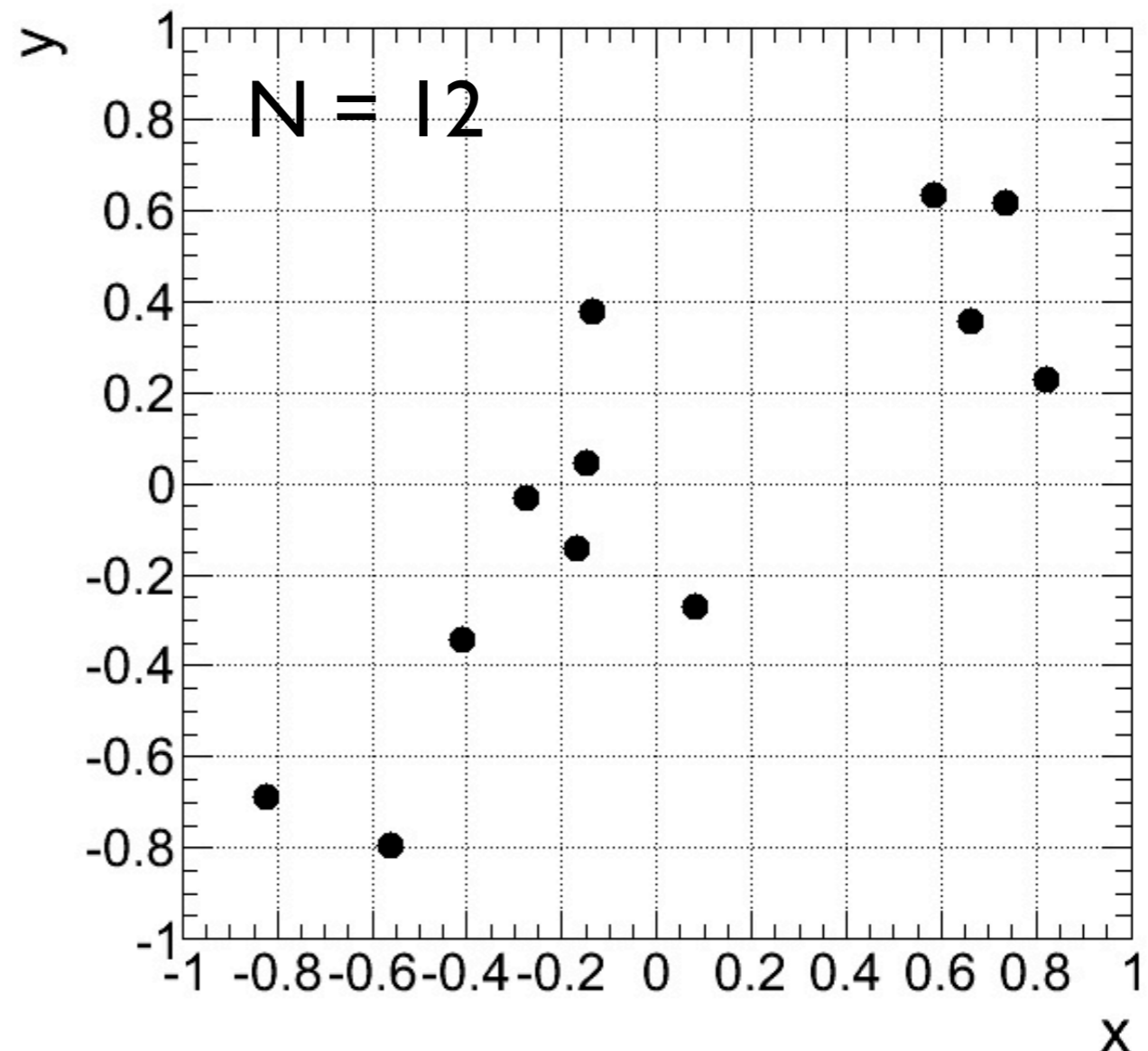
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

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

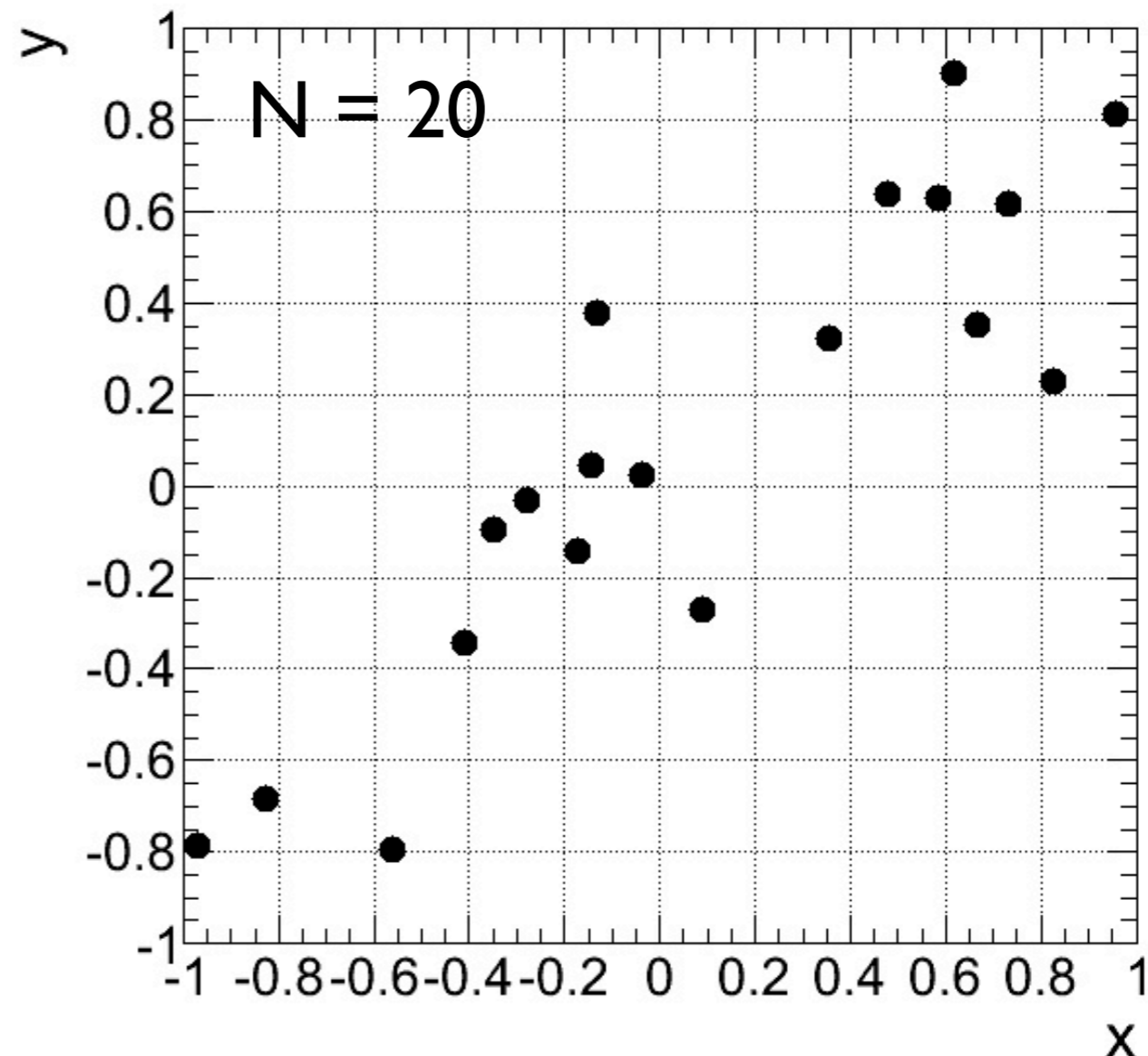
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

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

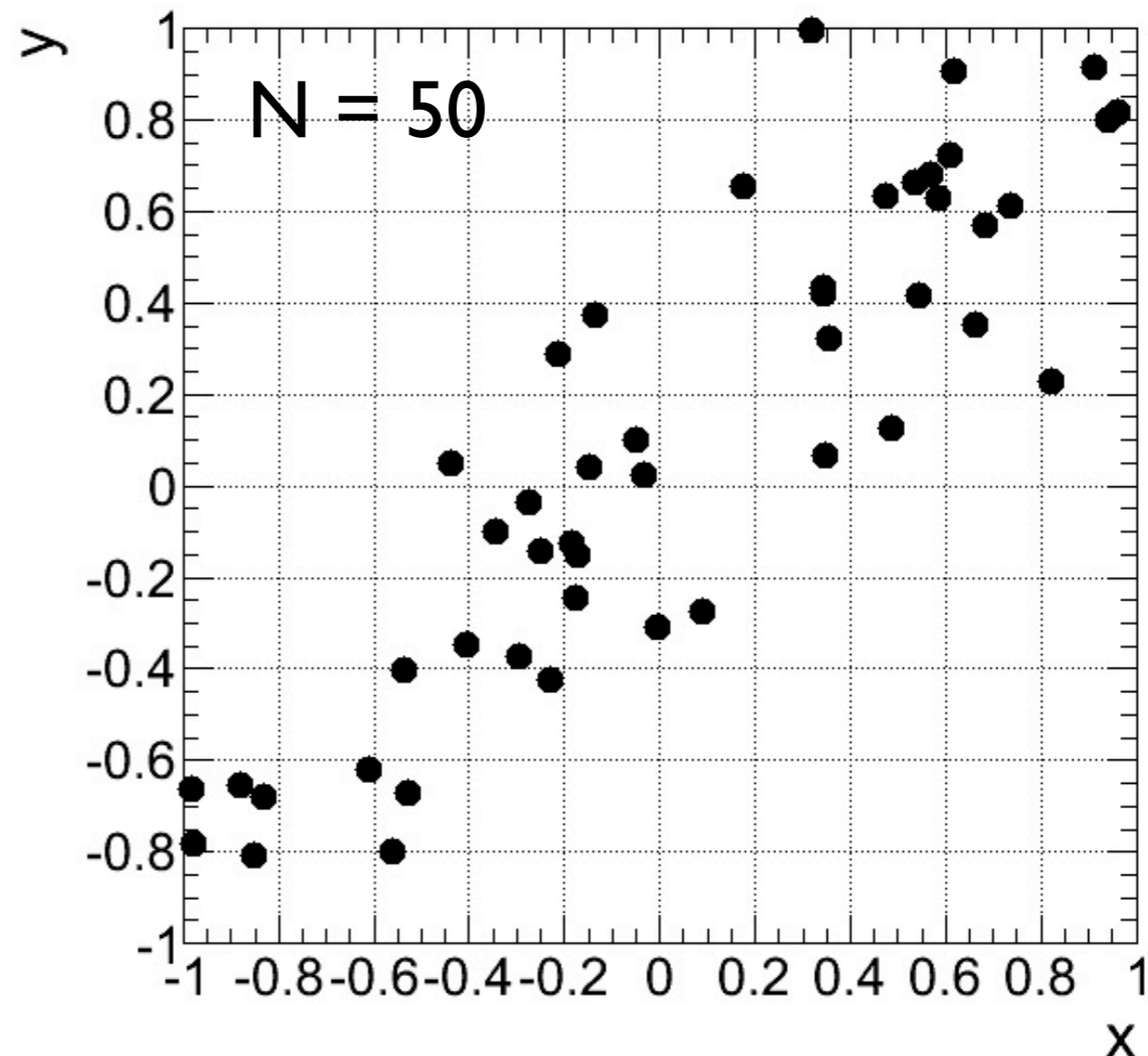
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

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

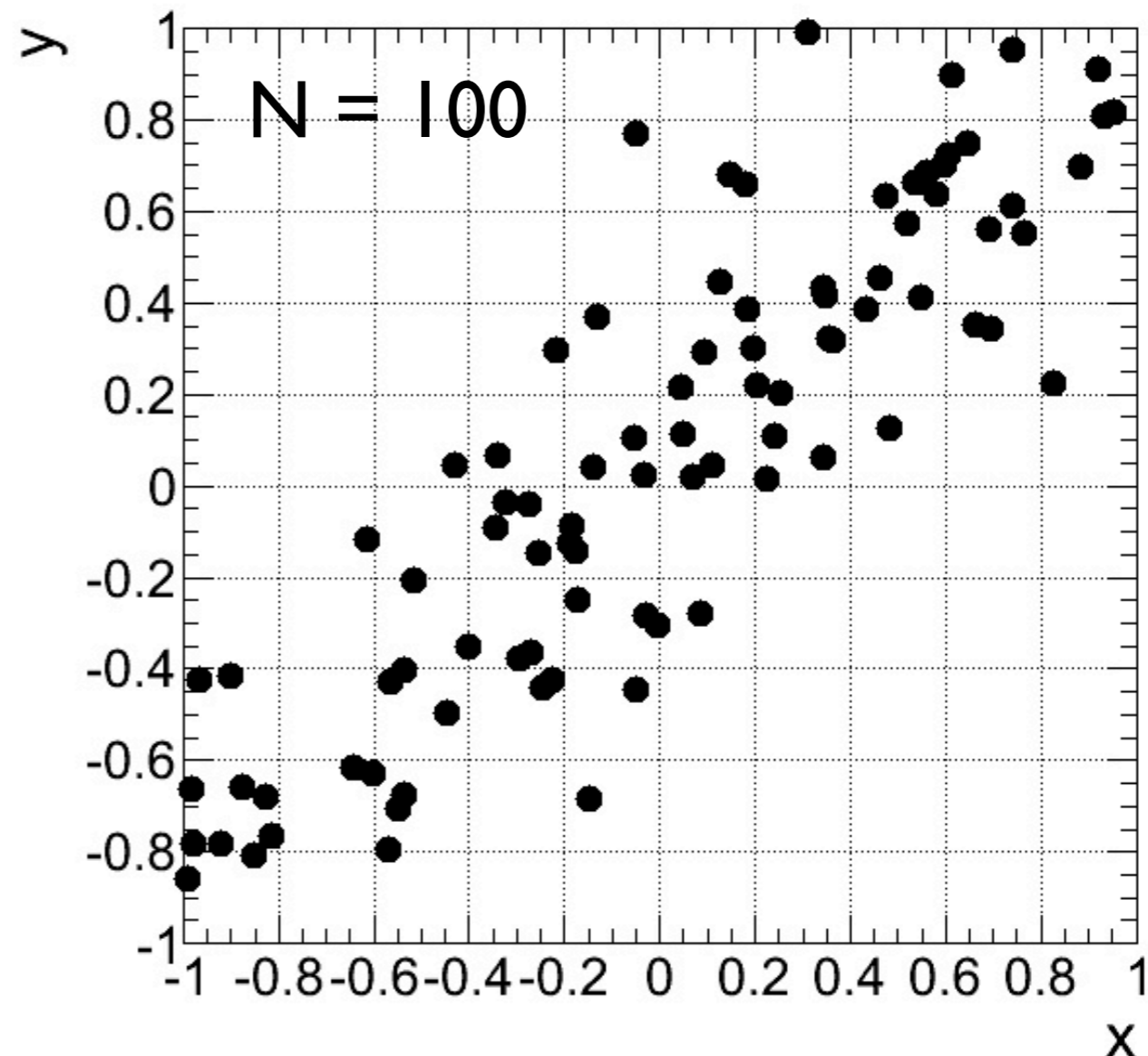
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



Representando duas variáveis

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

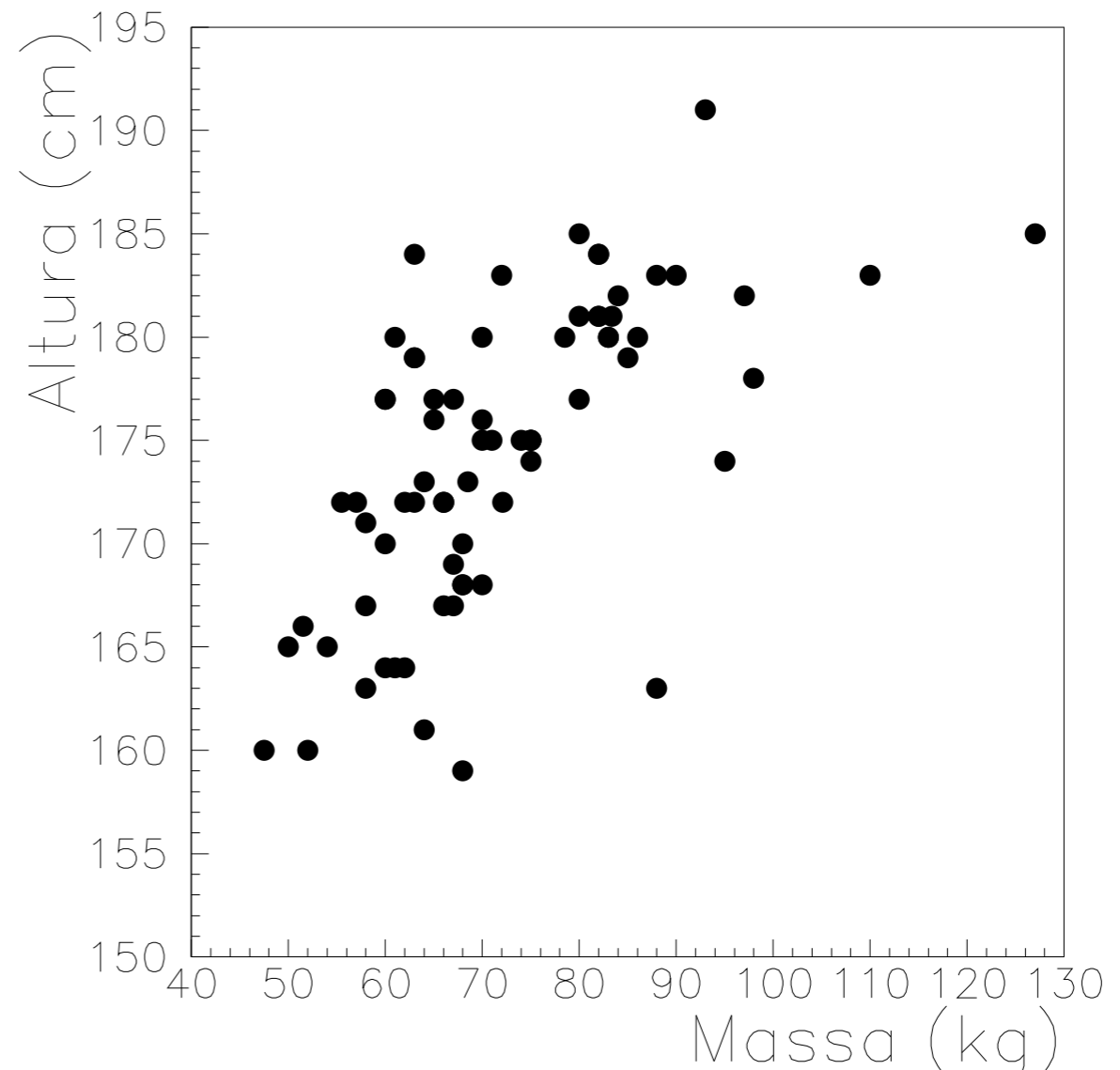
Exemplo: Considere um conjunto de dados de duas variáveis (x, y)



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:

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

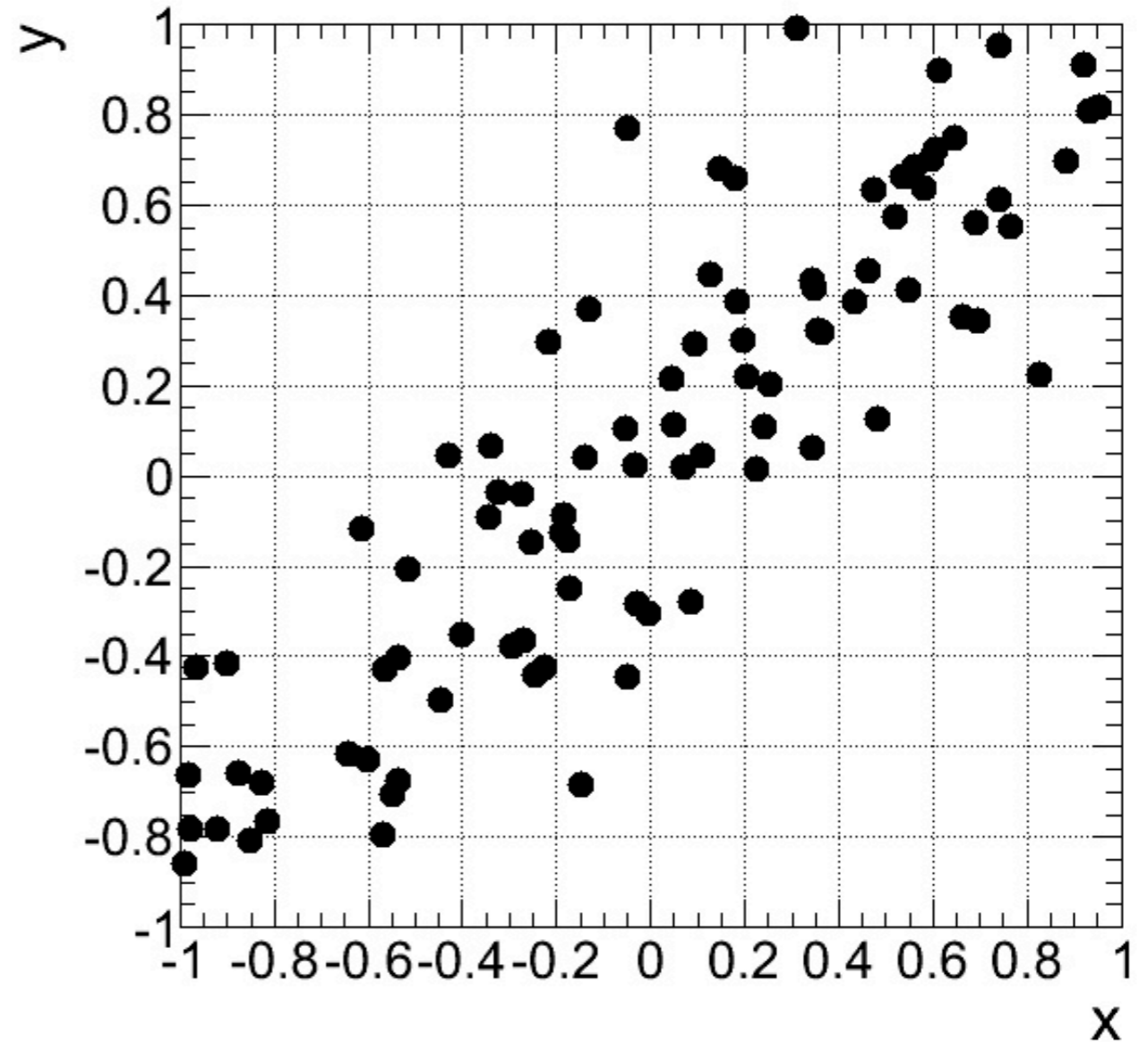
e que não importa a ordem das variáveis:

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

Parâmetros de correlação: covariância

Covariância:

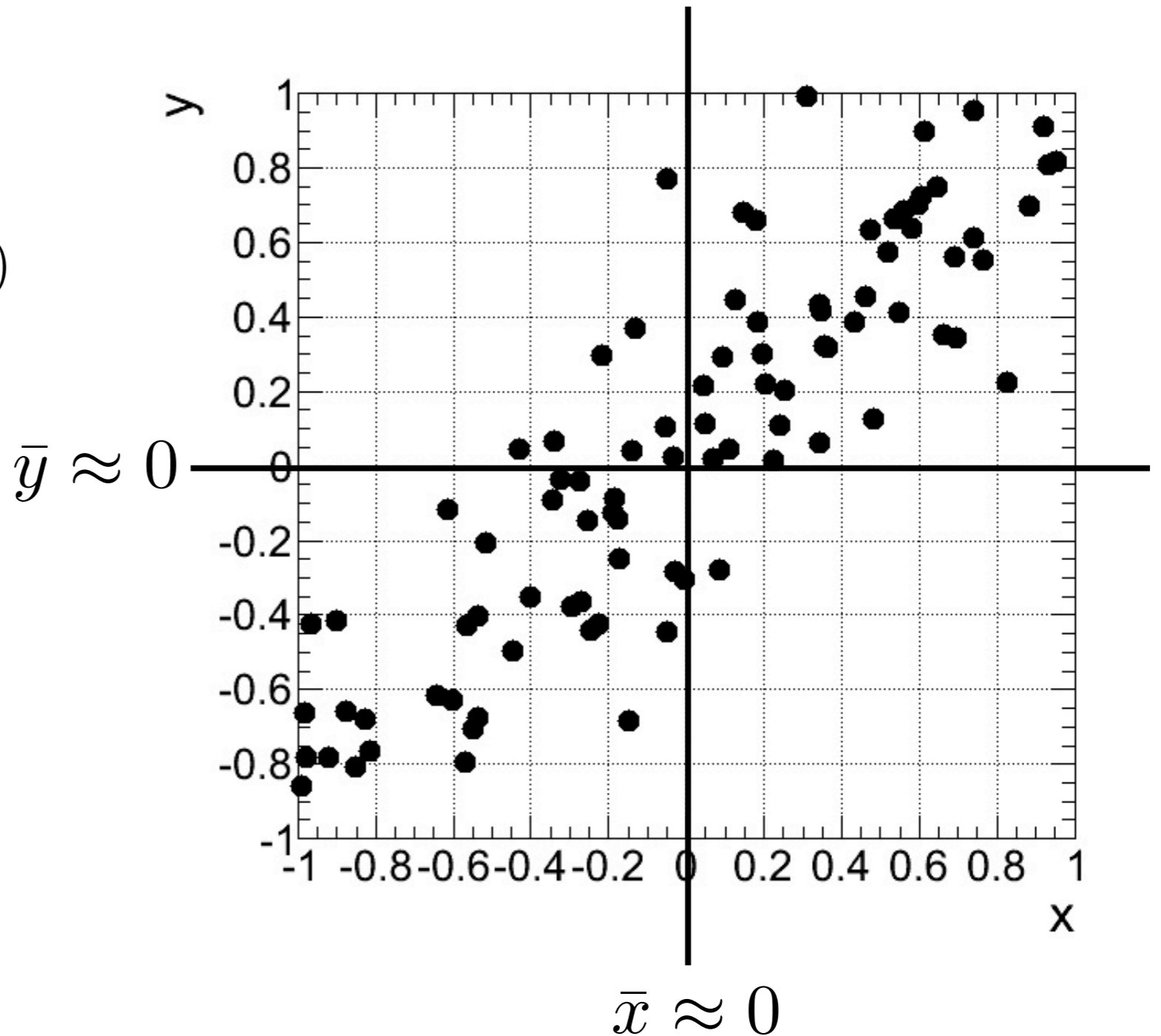
$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$



Parâmetros de correlação: covariância

Covariância:

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

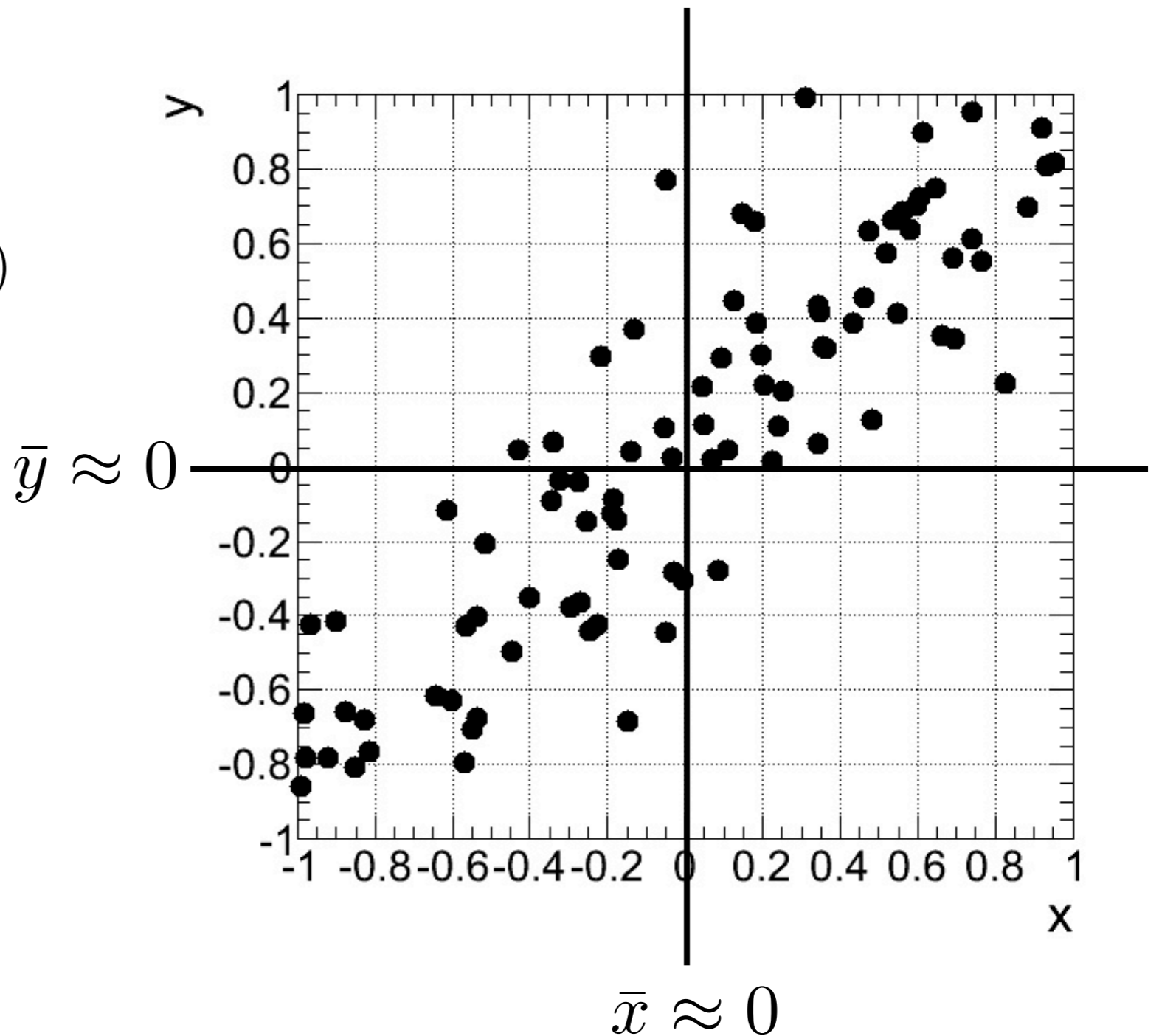


Parâmetros de correlação: covariância

Covariância:

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

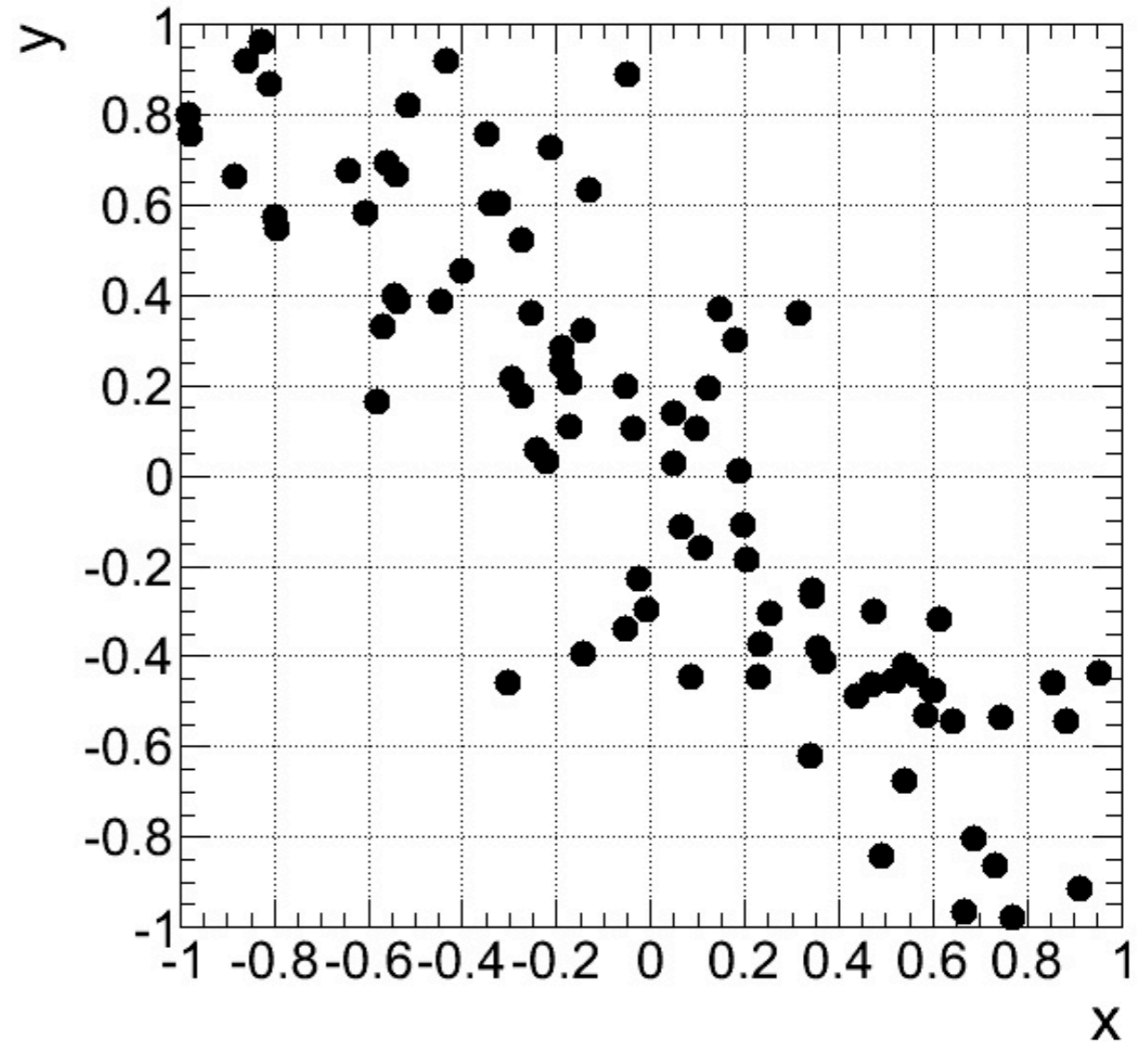
➔ $\sigma_{xy} > 0$



Parâmetros de correlação: covariância

Covariância:

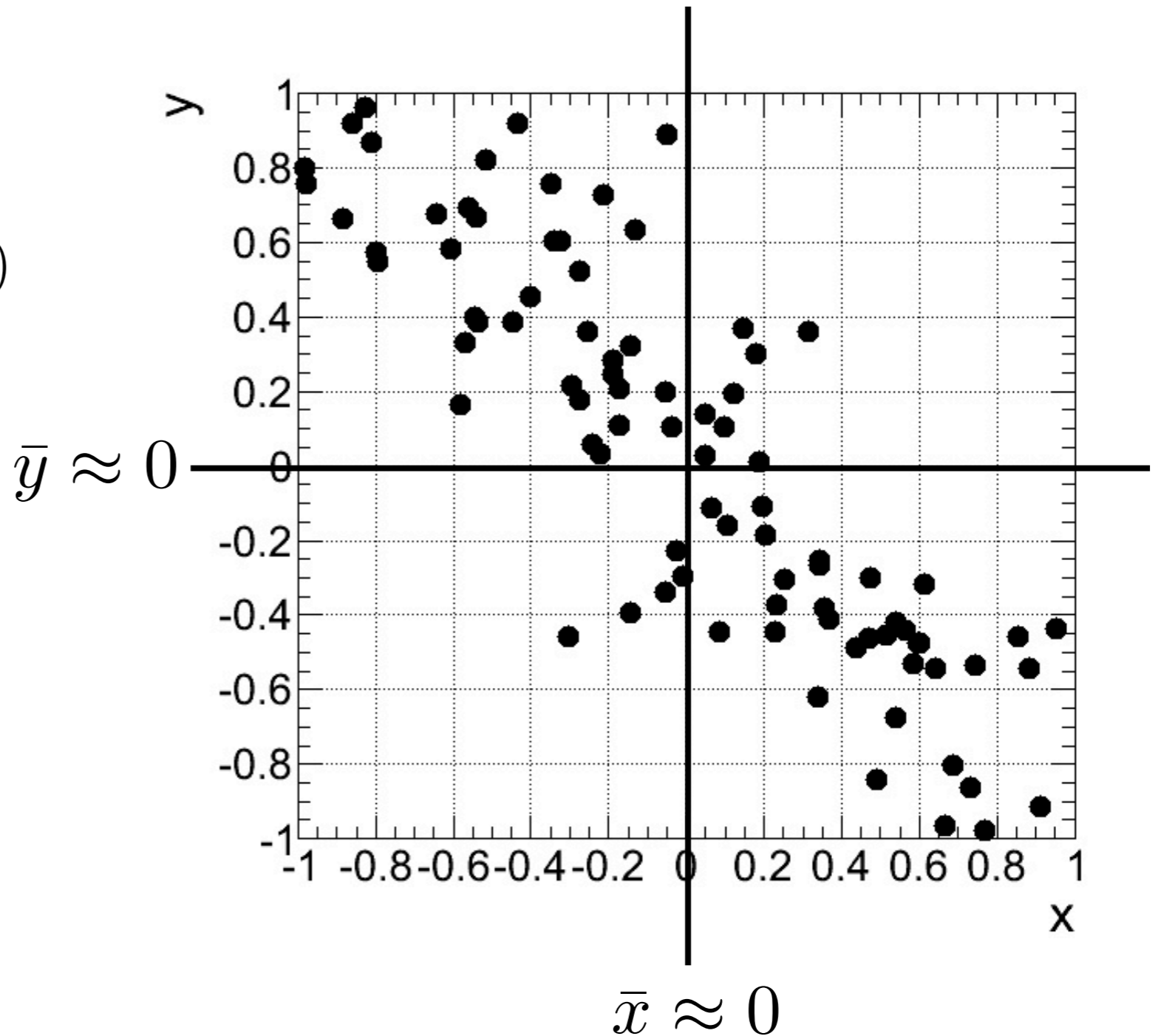
$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$



Parâmetros de correlação: covariância

Covariância:

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

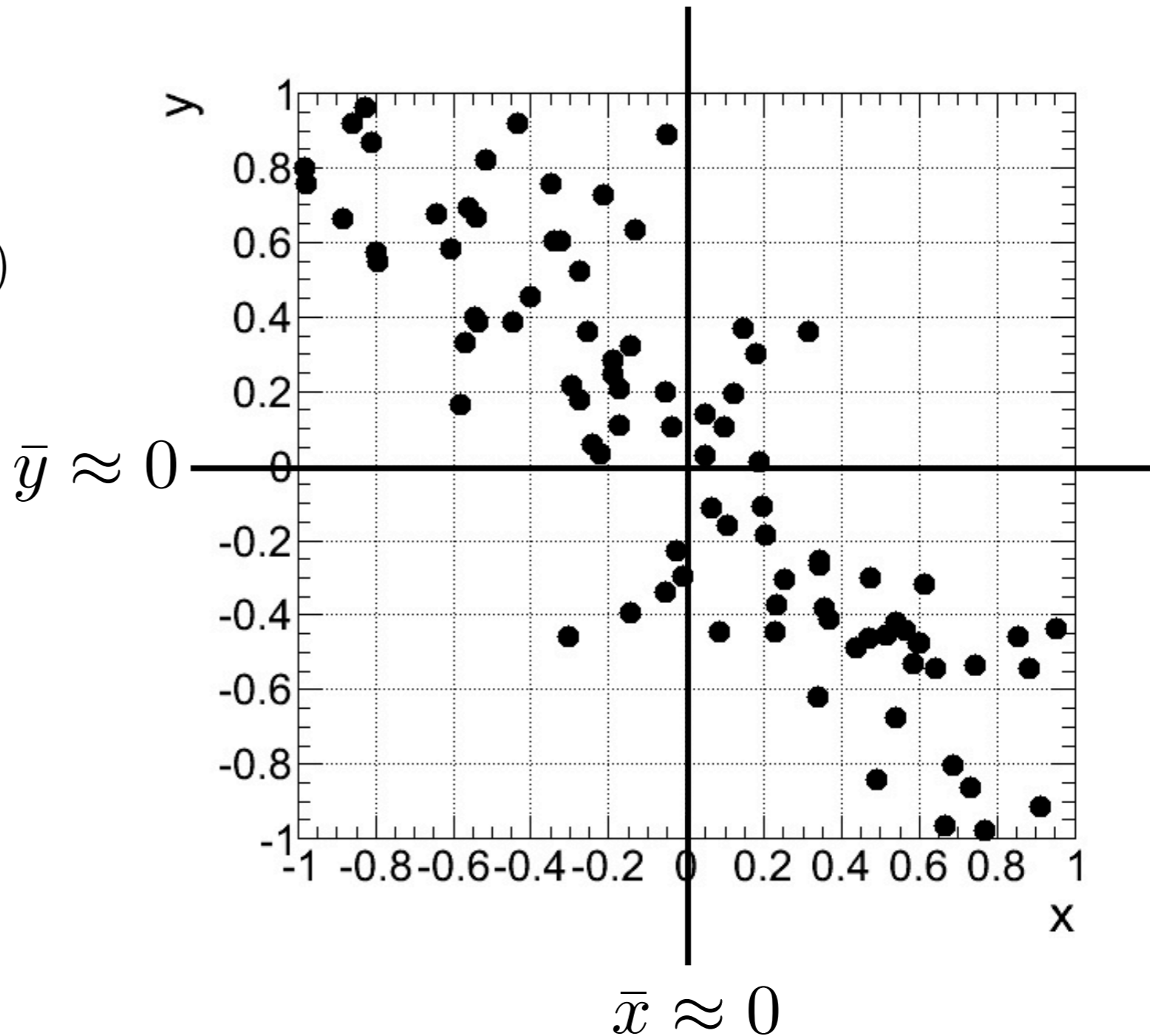


Parâmetros de correlação: covariância

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

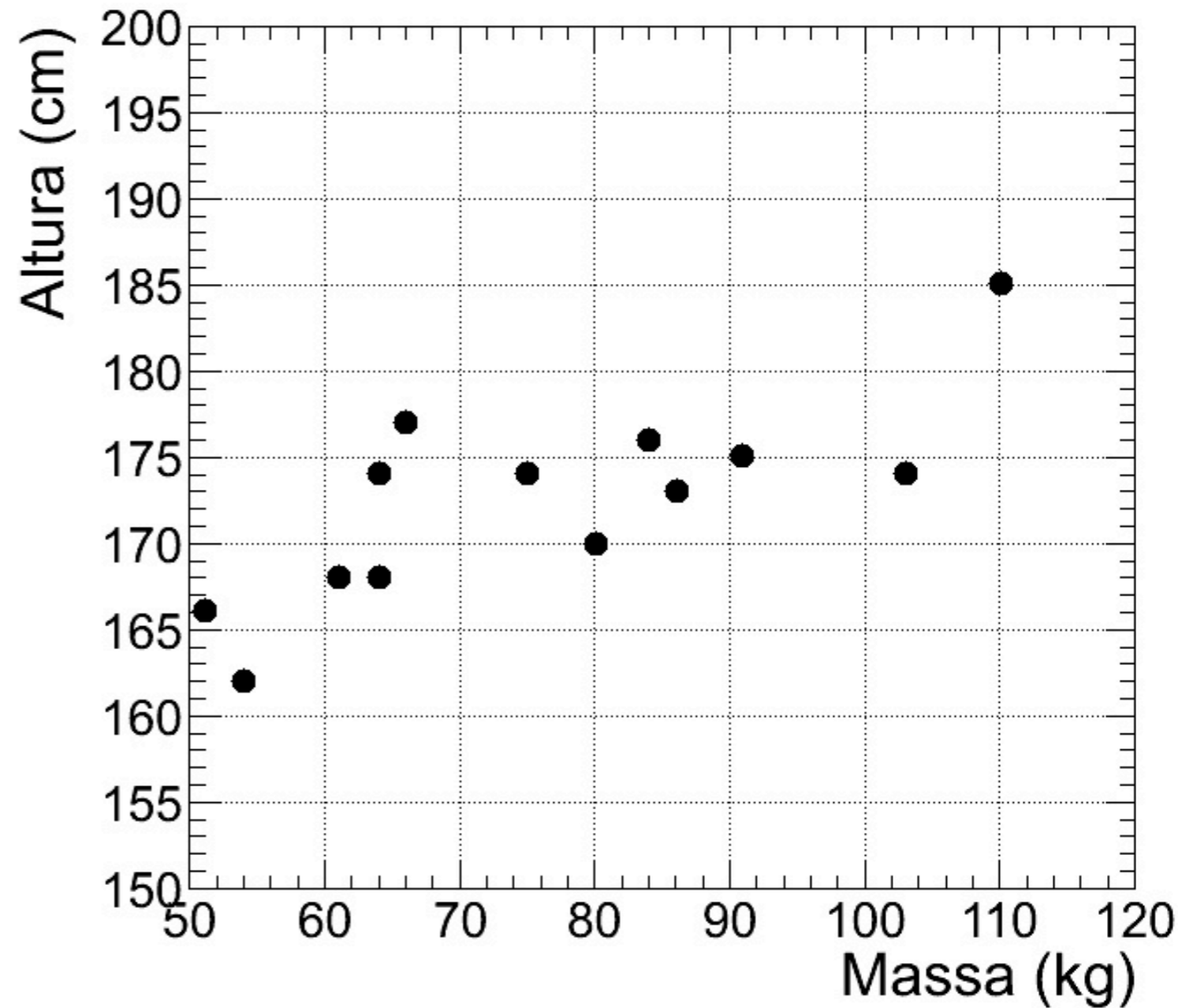
$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad -1 \leq r \leq 1$$

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

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

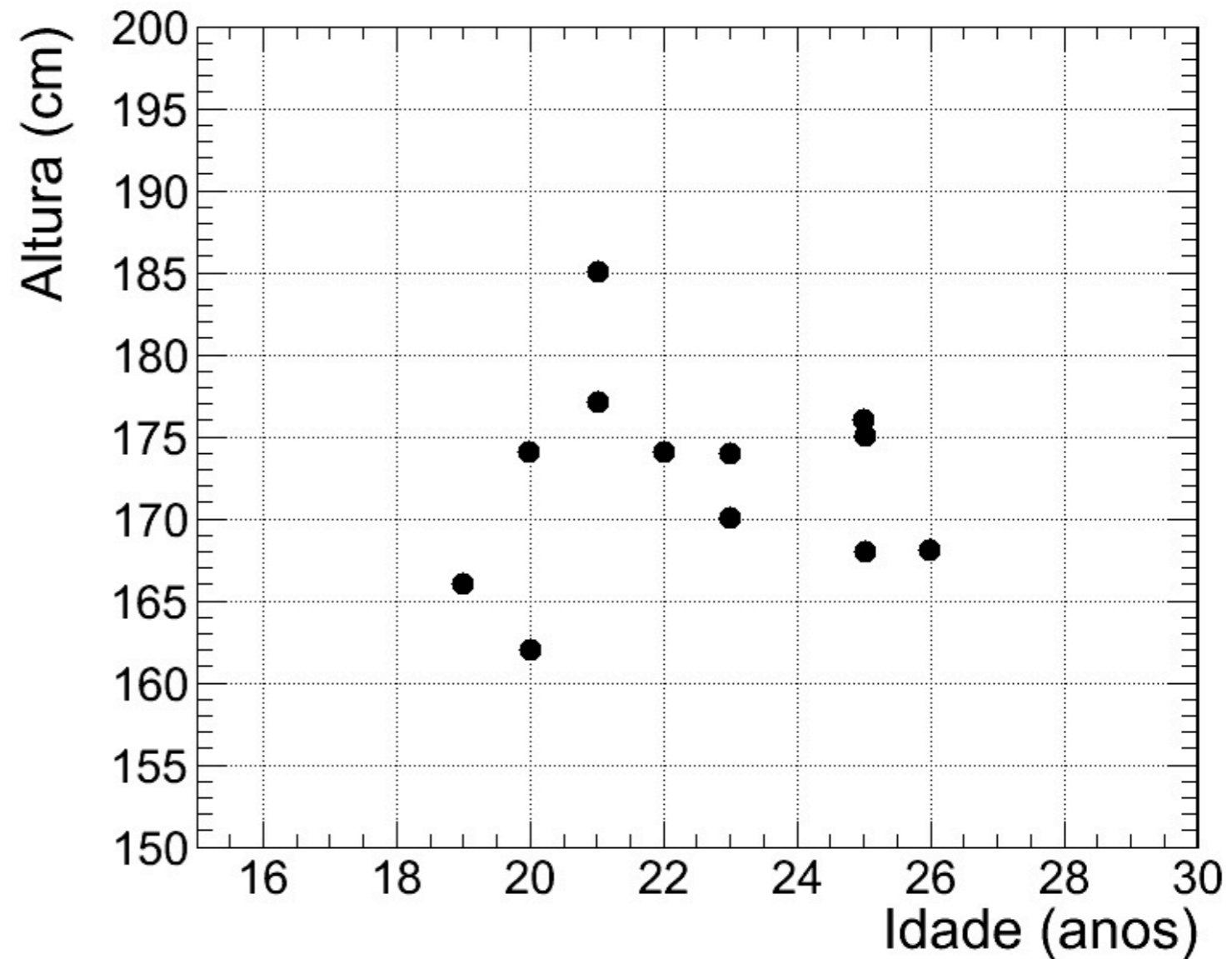
Atividade - Aula 1

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



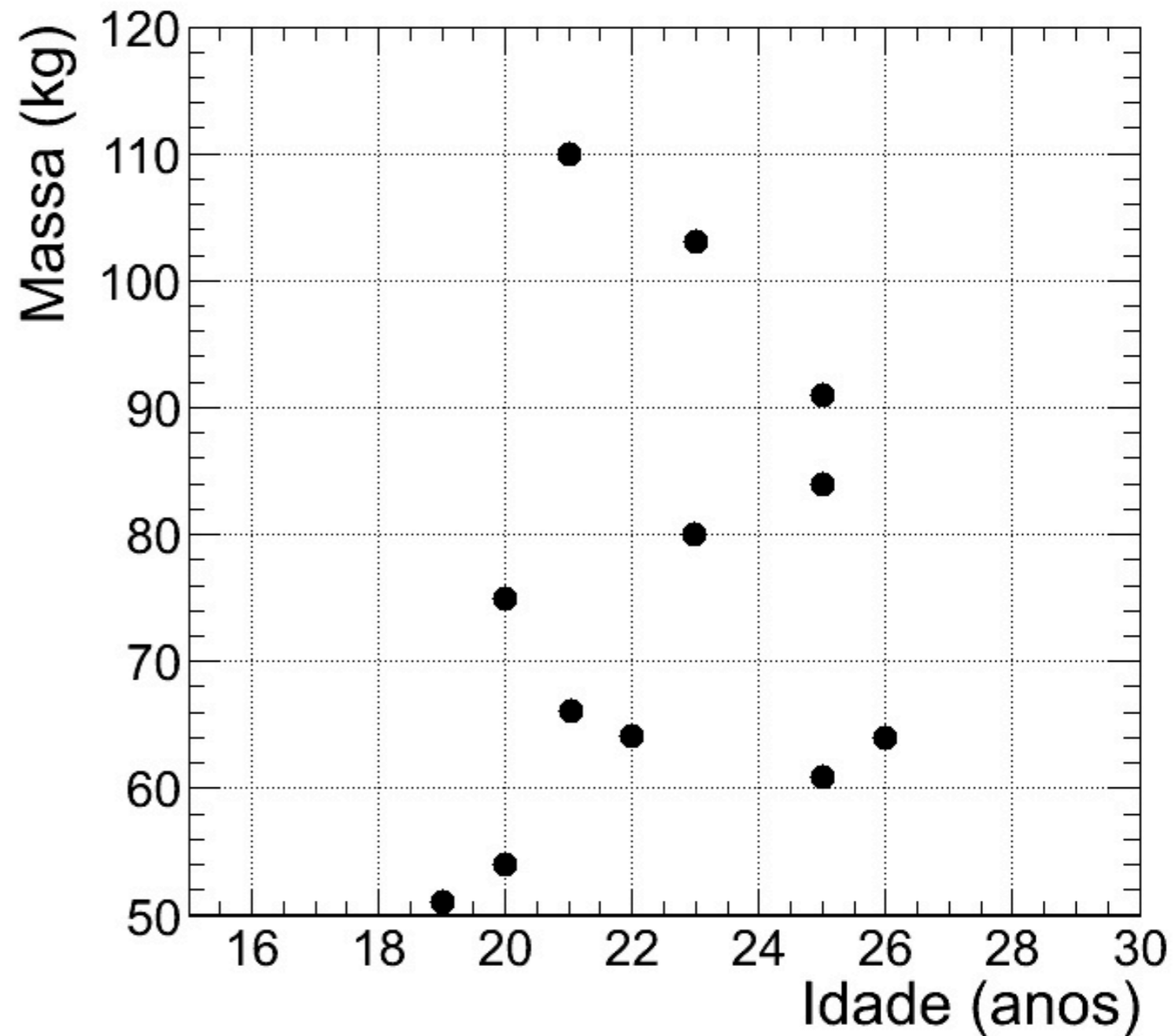
Atividade - Aula 1

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 - Aula 1

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



Atividade de aula

1- Com as coleções de dados das idades, massas e alturas dos estudantes da turma de Física Geral, determine:

- i) Amplitude, desvio médio, variância, desvio-padrão para cada variável;
- ii) Covariância de todos os pares de variáveis (massa x idade, altura x idade, altura x massa) e respectivo coeficiente de correlação

2- Exercícios 2.5.1 - 2.5.5 do livro “Estimativas e erros em Experimentos de Física”