

Física 12.º ano

00 Laboratório de pesagem de sementes

Versão: 1,0

José Pedro do Amaral

15 de setembro de 2025

1. Introdução

Até um relógio parado dá a hora certa duas vezes ao dia.

O objectivo principal destes laboratórios é o de tornar o método científico num ponto central de ensino em aulas de ciência. Porque o método científico é um (ou até o) elemento demarcante da Ciência, tentarei sempre de uma forma naturalmente limitada pela estrutura experimental, conduzir uma discussão que comece com a noção do conhecimento pré-existente e que eventualmente acabe na ciência e no seu método aplicados à experiência em curso. O objectivo recorrente para os alunos será a formulação de uma **hipótese** para cada experiência. Estas hipóteses deverão ser testadas no laboratório com os resultados experimentais e contra as teorias e leis aplicáveis. De forma relacionada e para uma exactidão de linguagem, também deverá ser notado que quando se usa as palavras hipótese, teoria e lei, estamos a falar de uma hierarquia de segmentos de conhecimento. Uma hierarquia em que a hipótese é conhecimento em teste, a teoria é conhecimento bem testado e aceite, e lei é conhecimento normativo de grande validade geral.

2. Preâmbulo — aproximação à estatística

Laboratório de Estatística com Sementes de Alfarroba

2.1 História do Quilate e da Semente de Alfarroba

A semente de alfarroba (*Ceratonia siliqua*) foi usada na Antiguidade como padrão de massa para a pesagem de pedras preciosas, devido a uma uniformidade putativa¹. Essa uniformidade tem sido debatida e contestada (Turnbull et al., 2006; Janzen, 1979). É possível que esta prática tenha dado origem à unidade *quilate* (do grego *kerátion*, via árabe *qīrāt*), definido actualmente como exactamente 200 mg. Este laboratório procura testar essa reputação de uniformidade à luz da estatística moderna, com base em medições reais

¹As sementes de alfarroba foram possivelmente a referência empírica original do quilate, usadas de facto em muitos mercados e bazares antes de haver pesos metálicos padronizados. O seu uso poderia não ter sido formalmente regulado, mas foi funcional, preciso em média e suficientemente fiável para servir como base prática da unidade, o que explica a sua persistência histórica e etimológica.

da massa de sementes de alfarroba. As sementes de alfarroba têm algumas características interessantes e mudam pouco de massa ao longo do tempo (Fig. 1).

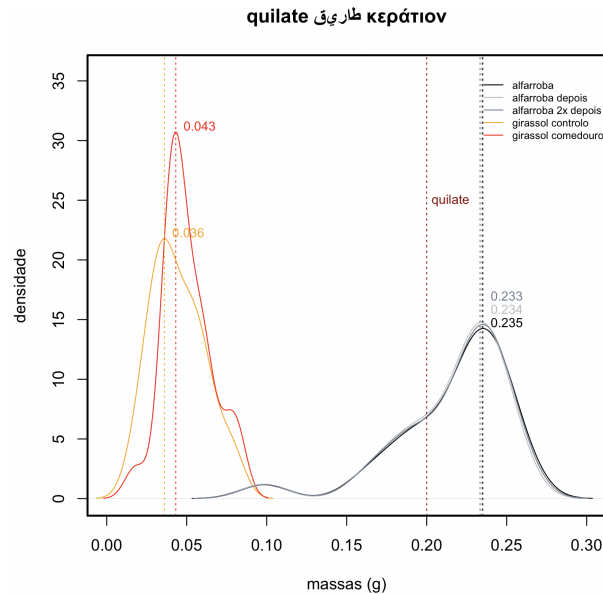


Figura 1: Comparação de massas de sementes de girassol e de alfarroba ($n = 23$ para cada grupo de amostras). Os valores de densidade são semelhantes aos de frequência mas calculados para que a área debaixo de cada curva seja 100%. As sementes de alfarroba pesadas para as três curvas são as mesmas mas pesadas em anos diferentes. A linha vertical de quilate corresponde ao valor SI de quilate que é 200 mg.

2.2 Tipos de Dados

Os dados que recolhemos em ciência podem ser de tipos diferentes, de acordo com a natureza daquilo que medem e com as operações que fazem sentido aplicar-lhes. Os **dados nominais** são aqueles que apenas servem para distinguir categorias, sem qualquer ordem implícita entre elas. São dados *qualitativos*: o nome da espécie de uma planta, a cor de uma flor, a marca de um produto. Nestes casos não faz sentido ordenar as categorias nem calcular médias; apenas podemos contar frequências ou proporções.

Os **dados ordinais** também representam categorias, mas nestes existe uma ordem natural entre eles. Por exemplo, as notas de satisfação numa escala de «mau, razoável, bom ou excelente» ou as classificações escolares do mesmo tipo. As categorias têm uma hierarquia, mas as distâncias entre elas não são necessariamente iguais; passar de «razoável» para «bom» não é a mesma coisa que passar de «bom» para «excelente». Podemos ordenar e comparar posições relativas, mas não calcular diretamente diferenças ou médias.

Por fim, os **dados contínuos**, resultam de medições em escalas numéricas que admitem subdivisões arbitrárias. Exemplos são a massa em gramas, a altura em centímetros, a temperatura em graus Celsius ou o tempo em segundos. Aqui faz sentido calcular médias e aplicar operações matemáticas porque as diferenças e proporções têm significado real. A massa de 240 mg é de facto o dobro da massa de 120 mg, e a média de várias cronometragens de um processo dá-nos uma boa estimativa da duração típica desse processo. Estes dados são dados *quantitativos*.

Em resumo, os dados nominais servem apenas para identificar categorias, os ordinais para ordenar categorias com significado hierárquico mas sem distâncias definidas, e os contínuos para quantificar grandezas de forma numérica, permitindo operações matemáticas.

2.3 Exactidão, Precisão e Medidas de Tendência Central

Distingue-se **exactidão** (proximidade ao valor verdadeiro) de **precisão** (reprodutibilidade das medições). Uma balança pode ser precisa mas não exacta, mas se for exacta será sempre também precisa² (Fig. 2).

2.3.1 Os alunos calcularão estas aproximações à tendência central da amostra:

Moda ($\hat{x}$) — valor mais frequente; sobretudo para $x \in \mathbb{Z}$. Com valores não-inteiros ($x \in \mathbb{R}$), o cálculo deverá ser feito a partir dum histograma de contentores (eq. 1; Fig. 4);

Mediana ($\tilde{x}$) — valor central da amostra ordenada. Se n for ímpar, a mediana será o valor central da amostra ordenada, mas se n for par, então a mediana será a média dos dois valores centrais (eq. 2);

Média ($\bar{x}$) — soma dos valores dividida pelo número de observações (eq. 3).

2.3.2 Calcularão também estas medidas de dispersão como aproximações à variabilidade da amostra:

Alcance — diferença entre o valor máximo e o mínimo;

Desvio médio absoluto (eq. 4) — média do somatório da diferença absoluta entre cada valor e a média;

Quartis — marcadores de 25, 50 e 75% dos valores ordenados da amostra. É de salientar que o quartil 50% é a mediana (Fig. 3).

²Competirá sempre ao cientista avaliar a utilidade dessa exactidão, um pouco como a do famoso relógio parado da epígrafe.

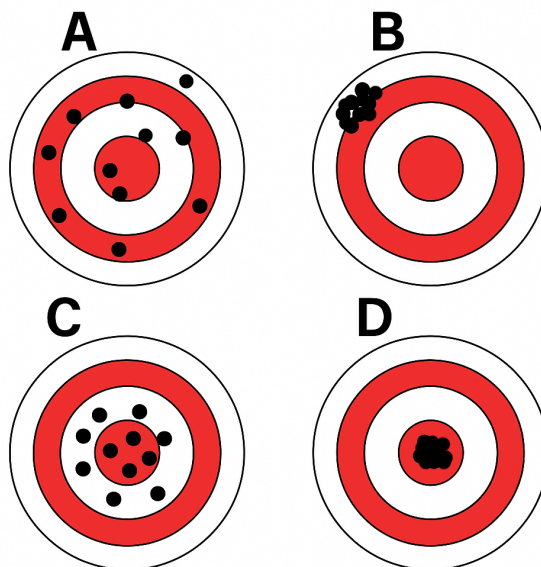


Figura 2: Exemplo de combinações de graus de precisão e de exactidão usando alvos e pontos de impacto ($n = 11$ por alvo). A. baixa exactidão e baixa precisão; B. baixa exactidão e alta precisão; C. alguma exactidão e alguma precisão; D. alta exactidão e alta precisão.

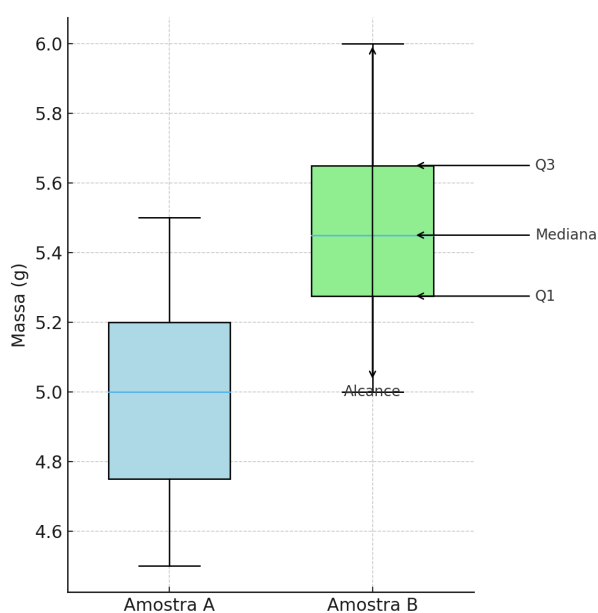


Figura 3: Comparação das massas de duas amostras, A e B, usando caixas de bigodes. Os bigodes indicam o alcance, as caixas indicam o intervalo entre o quartil 25% e o quartil 75%. O quartil 50% é a mediana.

2.3.3 Um pouco mais acerca do cálculo da moda

Quando os dados são contínuos, como massas em miligramas com casas decimais, a moda literal — isto é, o valor mais repetido — perde utilidade prática, porque cada medição tende a ser única ou quase. Nesses casos, não fará muito sentido trabalhar com uma

moda exacta, mas com uma moda aproximada, que será obtida a partir de classes ou de densidades.

Com esta abordagem das classes, divide-se primeiro a amostra em intervalos (os contentores representados no histograma por barras) e identifica-se o intervalo modal, que é aquele com maior frequência. A moda é então aproximada por qualquer ponto desse intervalo e, se necessário, poderá ainda ser refinada por interpolação dentro do próprio intervalo.

Pela abordagem das densidades, em vez de intervalos fixos recorre-se a uma estimativa de densidade suavizada, com métodos como o de *kernel density*, e a moda corresponderá ao ponto onde essa curva atinge o máximo local, o que é particularmente útil quando se tem poucos dados.

No caso concreto dos trinta valores da amostra que dou como exemplo, a moda discreta é de 242,2 mg, porque é o único valor que se repete (Tabela 1). Mas olhando para o histograma com seis contentores (Fig. 4), percebe-se que a classe entre cerca de 242 e 244 mg concentra mais valores, pelo que a moda por classes estaria nesse intervalo, em torno de 243 mg, o valor mais representado na distribuição. Conclui-se assim, para dados contínuos de laboratório, que é mais adequado definir a moda a partir do intervalo modal (ou da densidade estimada) do que pela simples repetição de valores.

Tabela 1: Valores de massas (g) usados para gerar o histograma da figura 4.

236,9	241,6	242,2	242,2	225,2	234,5
243,5	236,5	239,0	248,2	245,9	238,4
226,7	242,1	243,9	239,9	244,1	242,5
240,8	246,0	233,9	246,6	246,9	239,3
234,7	239,5	241,5	242,6	243,4	231,5

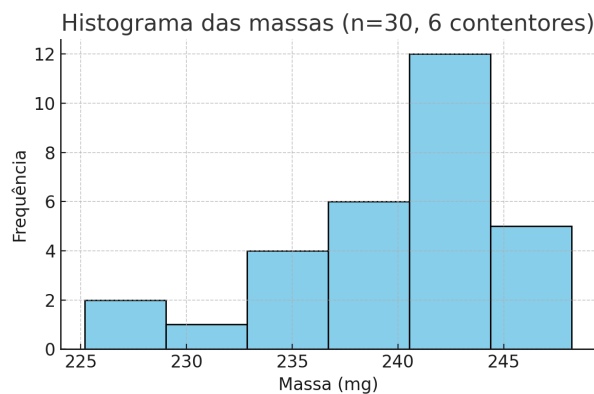


Figura 4: Histograma das frequências das massas de uma amostra. Com um $n = 30$, os valores estão divididos por seis contentores. A frequência indica o número de contagens no alcance do contentor.

2.3.4 Erro da amostra

É fundamental para qualquer investigador, o facto de que quando se mede uma amostra de uma população, a média dessa amostra não é necessariamente igual à média verdadeira dessa população que queremos estimar (a média dita paramétrica). Chamamos a essa diferença entre a média da amostra e a média paramétrica, o erro da amostra, E (eq. 5).

O alcance, o desvio médio absoluto e os quartis são medidas de dispersão que descrevem como os dados individuais se distribuem à volta de um valor central. O alcance mostra apenas a diferença entre o valor máximo e o mínimo, sendo muito sensível a valores extremos, e por isso dá uma noção muito em bruto da variação possível mas não se relaciona de forma directa com o erro da amostra. O desvio médio absoluto, ao calcular a distância média dos dados em relação à média, fornece uma medida mais estável da variabilidade: quanto maior for o desvio médio absoluto, maior será também a incerteza na média obtida da amostra. Os quartis, em particular o intervalo interquartil, medem a dispersão da parte central dos dados e são mais robustos a valores extremos; também aqui, um intervalo interquartil maior implica maior variabilidade e, portanto, maior incerteza na média.

O erro da amostra, E , não descreve a dispersão dos dados em si, mas sim a variação da média da amostra em relação à média paramétrica. Apesar disso, está directamente relacionado com a dispersão: quanto maior for a variabilidade dos dados, maior será o erro esperado da amostra; quanto maior for o número de elementos da amostra, menor será esse erro. O alcance, o desvio médio absoluto e os quartis dizem-nos como variam os dados, enquanto que o erro da amostra diz-nos como varia a média quando tiramos amostras diferentes. Quando se recolhe uma amostra, calcula-se a média dessa amostra. Contudo, essa média depende das observações que por acaso foram feitas. Se se repetir a experiência e se medir outra amostra do mesmo tamanho na mesma população, então obter-se-á uma média nova da amostra com grande probabilidade de ser diferente da primeira.

2.4 Materiais, Fórmulas e Procedimento

2.4.1 Materiais

Sementes de alfarroba maduras (mínimo 20 por grupo);

Balanças de vários tipos;

Calculadora científica, folha de cálculo, outro programa estatístico (ex.: R, Numbers, LibreOffice Calc, ou GeoGebra) ou à mão.

2.4.2 Fórmulas

$$\text{Moda} = \hat{x} = \arg \max_{x_i} f(x_i) \quad (1)$$

$$\text{Mediana} = \tilde{x} = \begin{cases} x_{(\frac{n+1}{2})}, & \text{se } n \text{ for ímpar,} \\ \frac{1}{2} \left(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)} \right), & \text{se } n \text{ for par.} \end{cases} \quad (2)$$

$$\text{Média} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3)$$

$$\text{DMA} = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}| \quad (4)$$

$$\text{Erro da Amostra} = E = \bar{x} - \mu \quad (5)$$

2.4.3 Procedimento:

1. Pesar para determinar a massa de cada semente;
2. Registrar os dados e calcular as «estatísticas» descritivas;
3. Representar graficamente os dados (caixas de bigodes, Fig. 3; histograma, Fig. 4);
4. Aplicar os testes descritos;
5. Comparar a média obtida com o valor de 200 mg;
6. Discutir a precisão da semente de alfarroba como padrão histórico de massa.

2.5 Objetivos Educativos

Introduzir uma aproximação à estatística descritiva e inferencial;

Aplicar ferramentas matemáticas a dados reais contínuos e medidos pelos próprios alunos;

Desenvolver pensamento crítico e competências de representação gráfica;

Relacionar história da ciência com práticas laboratoriais actuais;

Preparar os alunos para a análise de dados dos outros laboratórios.

2.6 Orçamento Estimado

Sementes: 0€ por turma;

Balanças: reusáveis; já existentes nos laboratórios;

Métodos de cálculo: gratuitos.

2.7 Notas Finais

A uniformidade das sementes de alfarroba é um mito interessante — mas um mito que é cientificamente testável. Esta actividade combina história, medição, estatística e pensamento crítico, oferecendo uma excelente introdução à análise de dados em Física e Química.

Este projecto pode ser adaptado para outras sementes ou contextos agrícolas, ou estendido a outras variáveis (ex.: densidade, volume).

2.8 Hipótese e Resultados

2.8.1 Hipótese:

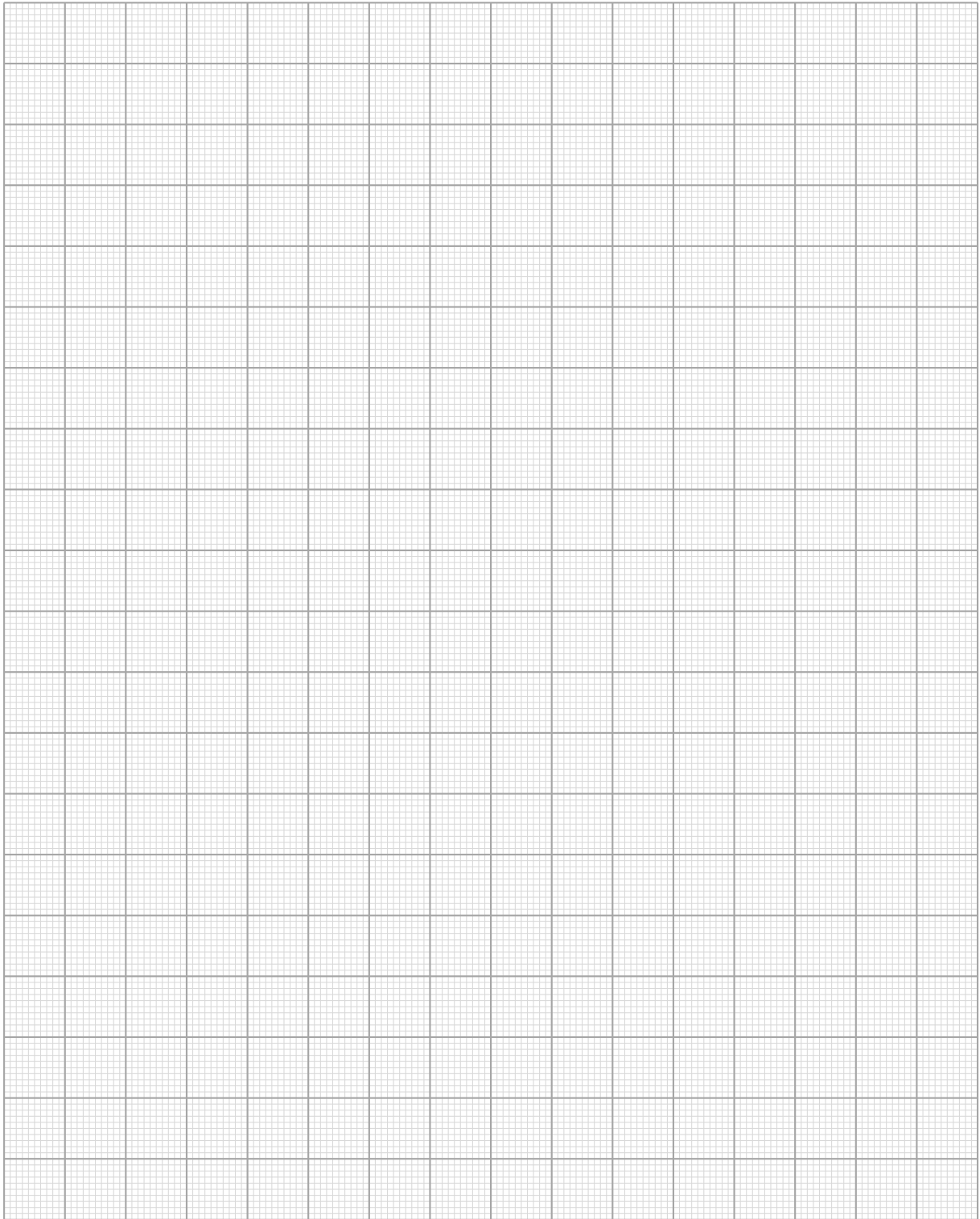
2.8.2 Resultados:

n.º	massa	cálculos
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		

2.8.3 Rejeição ou falha de rejeição de hipótese? Porquê?

2.8.4 A massa das sementes de alfarroba não varia? Como poderiam ter sido usadas como padrões de massa?

2.9 Gráficos:



3. Referências

- Janzen, D.H. 1979. How accurate was the carat? *Nature*, **280**: 10.
- Turnbull, L.A., Santamaria, L., Martorell, T., Rallo, J., and Hector, A. 2006. Seed size variability: from carob to carats. *Biology Letters*, **2**: 397–400.