Tratamento Estatístico de Dados em Física

Vito R. Vanin – Instituto de Física - USP

Outubro de 2024

Sumário

I. Introdução

Em um experimento, os resultados são obtidos dos valores medidos por meio de cálculos realizados no quadro teórico da Estatística, disciplina que também é usada na comparação desses resultados com os valores deduzidos da aplicação dos modelos físicos. Este texto apresenta os métodos e procedimentos adotados na análise dos dados e interpretação dos resultados obtidos dos experimentos com imagens do MEXI (Mecânica EXperimental com Imagens), que serão usados como exemplos sem uma introdução detalhada. Caso esteja lendo este texto sem haver realizado ao menos um desses experimentos, convém visitar a página http://www.fep.if.usp.br/~fisfoto/ a fim de familiarizar-se com os exemplos.

Não se pretende aqui cobrir todo o conteúdo da Estatística nem substituir um bom livro didático sobre o assunto, mas facilitar a elaboração dos modelos de tratamento de dados na análise dos experimentos; a bibliografia relaciona algumas obras que podem ser estudadas.

I.1 Estrutura do texto

O capítulo I traz uma introdução geral e o II resume os conceitos básicos na perspectiva de construção de modelos explicativos e preditivos das observações. Em seguida, no capítulo III, apresenta os aspectos gerais do cálculo do desvio-padrão de uma grandeza que é deduzida de outras grandezas medidas em que a flutuação estatística se manifesta, tarefa frequentemente chamada de "Propagação de Incertezas" (Bureau Internacional de Pesos e Medidas - BIPM, 2008) (Helene & Vanin, 1991). No capítulo IV, discutem-se os algoritmos práticos para efetuar esses cálculos, que são especializados para vários dos experimentos do MEXI no capítulo V, que será acrescido de novas seções à medida que se adicione material novo a essa página. A parte mais avançada do texto está no capítulo VI, que traz os métodos de determinação das linhas de tendência de uma grandeza física medida em função daquela que é controlada por quem conduz o experimento. O índice remissivo ao fim do texto pode facilitar encontrar as fórmulas desejadas na análise que estiver efetuando.

O texto explica a motivação e a racionalidade dos conceitos e definições, bem como apresenta deduções de vários resultados.

> Cada definição, conceito ou resultado que tem importância prática está escrito dentro de uma caixa como esta a fim de dar-lhe destaque, com a intenção de poupar quem lê de rever todo o texto para encontrar uma fórmula ou conceito.

I.2 Problemas que a estatística resolve

A Estatística é usada em muitas aplicações e campos de pesquisa. Seu quadro teórico evoluiu (e continua evoluindo) ao longo do tempo e, hoje, é adotado em todos os ramos da ciência. Há uma grande variedade de métodos, que refletem as particularidades das diferentes áreas de aplicação e pesquisa, mas o corpo central é comum a todas elas. Neste documento, porém, lidaremos com a estatística diretamente ligada com os procedimentos de medida dos valores de grandezas: a *inferência estatística*, que busca métodos gerais de extração das melhores estimativas dos valores das grandezas a partir dos valores obtidos nas medidas experimentais. A fim de mostrar a amplitude dessa especialidade, considere as situações relatadas a seguir.

i. O teor de glúten na farinha de trigo determina a elasticidade da massa do pão. Assim, um pão feito de farinha que contenha glúten em uma proporção menor que uma certa *fração mínima*[*] fica denso, sem uma textura convidativa. Por isso, o padeiro caprichoso necessita conhecer o teor de glúten na farinha com que faz seus pães.

ii. Um certo volume de gás pode ser forçado a passar por uma serpentina, de modo que absorva calor de um compartimento *frio*, que depois possa ser dissipado no ambiente externo, mais quente. Caso disponha de um modelo para esse gás, capaz de relacionar a energia necessária para forçá-lo a realizar um ciclo completo com o calor absorvido do compartimento frio sendo transportado ao ambiente, o engenheiro poderá dimensionar o maquinário da geladeira para que mantenha a temperatura em torno de 4º C no clima tropical.

iii. O consumo de energia elétrica cresce no país. Se dispuser de um modelo para esse consumo, o responsável pelo planejamento do setor elétrico poderá definir sua expansão adequadamente, sem que falte energia, o que travaria o crescimento, nem haja disponibilidade sem demanda, o que seria um desperdício dos recursos investidos.

Esses problemas, porém, não têm solução exata – é impossível conhecer o verdadeiro teor de glúten da farinha, não se consegue um modelo exato para um gás, muito menos para o consumo de energia elétrica do país. No entanto, o verdadeiro problema do padeiro é fazer um pão gostoso e o do engenheiro é projetar uma geladeira que funcione sem desperdiçar energia. O economista precisa planejar a construção de usinas hidrelétricas, eólicas, solares, térmicas, etc. para os próximos anos de modo que possa acelerar o processo quando houver indícios de falta (vai ser possível porque a economia cresceu mais que o previsto no modelo econométrico e haverá mais recursos) ou desacelerar, havendo indícios de sobra (vai precisar porque a economia cresceu menos que o previsto e recursos não devem ser usados em algo desnecessário).

O problema real tem solução por meio de medidas analisadas dentro dos quadros teóricos da química, física e economia, obtendo-se resultados numéricos por meio de cálculos que adotam o quadro teórico da inferência estatística. Da medida realizada, ao invés do valor verdadeiro da grandeza, obtêm-se uma estimativa da grandeza e da precisão com que foi medida. Embora não seja possível dispor do valor verdadeiro da grandeza, consegue-se determinar um *intervalo* com certa probabilidade de conter o valor verdadeiro; uma vez que reduzir a largura desse intervalo exige mais esforço e tempo na medida, deve-se estar atento ao propósito da tarefa para definir a precisão necessária. Nos exemplos do teor de glúten e do modelo do gás, intervalos com largura igual a alguns por cento do resultado final devem bastar; já para o consumo anual de energia elétrica do país, provavelmente se necessite um valor mais preciso que 1%, além de conhecer bem essa precisão.

Este texto aborda conceitos elementares da estatística a partir de exemplos de física, sem perda da generalidade – a economia, a química, e cada uma das ciências desenvolvem métodos algo diferentes, mas com base nestes mesmos conceitos.

---

[*] A determinação dessa *fração mínima* é um outro problema de inferência estatística...

II Conceitos elementares

Não é possível encontrar o valor exato de uma grandeza em um experimento de física, uma vez que ele só pode ser medido com um instrumento, que sempre tem limitações. No entanto, concebemos o mundo físico de maneira tal que muitas grandezas, tais como a carga do elétron e a constante universal dos gases, têm valores intrínsecos às suas naturezas, que chamaremos aqui de valores verdadeiros. Assim, em uma medida de uma dessas grandezas, conduzida adequadamente, obtém-se um valor *próximo* ao seu *valor verdadeiro*. Embora seja também impossível determinar a diferença entre o valor medido e o verdadeiro, uma vez que este último é desconhecido, podem-se definir grandezas que reflitam esta diferença de alguma maneira aproximada. Infelizmente, não se pode quantificar a incerteza simplesmente pela média da diferença entre o valor medido e o verdadeiro, porque a diferença, que pode ser tanto positiva quanto negativa, terá valor médio provavelmente nulo, o que obriga a elaborar mais esta questão.

A inferência estatística não trata de qualquer tipo de erro, apenas dos *erros aleatórios*, que são aqueles cujos valores não se pode antecipar. Um procedimento errado, como medir a distância entre dois pontos com uma trena mal esticada, leva a um valor certamente maior que o verdadeiro, portanto não é aleatório. Os termos *incerteza* e *precisão* têm os mesmos significados da linguagem cotidiana, mas aqui ligam-se à noção de erro aleatório, inerente ao processo de medida: quanto maior a incerteza de medida, menor a precisão, e vice-versa. É impossível evitar a incerteza em uma medida (ou ser totalmente preciso), de modo que encontrar o valor verdadeiro é impossível, também. Por isso, na medida de uma grandeza física, determinam-se *dois* números, um para estimar seu valor e outro para avaliar sua precisão.

O fato do resultado da medida de uma grandeza física x diferir do valor verdadeiro precisará ser levado em conta explicitamente nas fórmulas usadas para estimar a grandeza, o que pode gerar confusão quando se começa a estudar este assunto. Antecipando o que está nas próximas seções, em que se monta o quadro conceitual da inferência estatística, aparecerão nomes diferentes para valores que parecem expressar a mesma coisa – distinguiremos o valor verdadeiro da grandeza x_0 do resultado de uma medição em particular, x_i, da estimativa final, $\hat{x}$, e do valor esperado pelo modelo estatístico usado, $\langle x \rangle$. Todos esses nomes estão ligados a definições precisas, que serão fornecidas adiante.

II.1 A medida da grandeza e sua representação por histograma

Considere a determinação da aceleração local da gravidade de um objeto no experimento Queda Livre. Do conjunto de imagens A1, obtém-se o valor g_1 para essa grandeza, que deve ser próximo do valor verdadeiro g_0, que é uma constante física. A melhor maneira de estimar o quanto g_1 é diferente de g_0 é repetir a medição com *outros* conjuntos de imagens, ou seja, obter um conjunto de valores

$$\text{medida} = \{g_1, g_2, \dots, g_i, \dots, g_N\} = \{g_i, i = 1..N\} \qquad (2.1)$$

em que cada um dos valores g_i foi obtido de um conjunto de imagens diferente: g_2 foi obtido de A2, g_3 de A3 e assim por diante. Como proceder quando se dispõe de um único conjunto de imagens, que é o problema mais comum, é o assunto que ocupará todas as outras partes deste texto, mas antes é preciso estabelecer certos conceitos e definir alguns termos.

Esse conjunto de valores, proveniente de N repetições do procedimento de estimativa da aceleração da gravidade a partir de conjuntos de imagens diferentes, é uma *medida* da aceleração da gravidade. O nome *medida* também designa a estimativa do valor da grandeza que é deduzido desse conjunto de medições por meio dos procedimentos descritos a seguir.

Assim, *medida* designa tanto o resultado obtido para a grandeza de interesse quanto todo o conjunto de valores obtidos no experimento destinado a estimá-la. Um valor particular do conjunto de dados obtidos na medida, por exemplo g_2, pode ser nomeado uma *medição*.

Um valor da grandeza obtido no procedimento experimental (um dado) é uma medição. A palavra medida nomeia tanto a estimativa do valor da grandeza quanto o conjunto de dados (conjunto dos valores obtidos na repetição das medições do experimento).

Em um histograma, os valores obtidos são agrupados por faixas de valores. Para descrever a construção de um deles, considere a medida cujos dados estão na Tabela 1. Os valores extremos são 9,759 e 9,845 m/s^2, a partir do qual se define o tamanho dos intervalos em que a abscissa será dividida, ou seja, de cada *canal* do histograma[†], em 0,01 m/s^2.

O conjunto de valores de uma medida direta da grandeza de interesse pode ser representado graficamente por um *histograma*.

Tabela 1. Medições da aceleração da gravidade em São Paulo, em m/s^2, obtidas dos conjuntos de imagens A1 a A18 do experimento "Queda Livre" do MEXI.

9,818	9,772	9,819	9,794	9,790	9,800
9,759	9,806	9,845	9,803	9,835	9,799
9,817	9,776	9,796	9,780	9,844	9,762

A Tabela 2 relaciona as faixas e a frequência com que se obteve uma medição no canal correspondente. O histograma está na Figura 1, e consiste em representar as faixas na abscissa e a frequência na ordenada. Assim como em qualquer gráfico, abre-se mão dos detalhes que constam na Tabela 1 para formar ideia sobre a maneira com que os valores observados se distribuem entre os possíveis resultados experimentais.

Vê-se que os valores se distribuem em torno de 9,80 m/s^2 (estão *localizados* em 9,80 m/s^2), e a maioria se espalha por um intervalo de largura de alguns centésimos de m/s^2 em torno desse valor (a *dispersão* é de alguns centésimos de m/s^2).

[†] Em um histograma, a variável contínua representada na abscissa é discretizada pela divisão da abscissa em segmentos contíguos, que são os *canais*.

Tabela 2. Frequência com que os dados da tabela 1 caem nas faixas definidas na coluna da esquerda. Note que cada canal inclui o extremo inferior do intervalo e exclui o superior, que cai no canal seguinte.

Canal, em m/s^2	frequência
[9,75; 9,76)	1
[9,76; 9,77)	1
[9,77; 9,78)	2
[9,78; 9,79)	1
[9,79; 9,80)	4
[9,80; 9,81)	3
[9,81; 9,82)	3
[9,82; 9,83)	0
[9,83; 9,84)	1
[9,84; 9,85)	2

Ao fato de medições repetidas da mesma grandeza resultarem em valores diferentes por razões incontroláveis se dá o nome de *flutuação estatística*. Há diversos modelos para essa flutuação estatística, mas na grande maioria dos casos do MEXI e em muitos dos experimentos em física, é apropriado adotar o ***modelo normal***, em que a forma do histograma tende à de uma gaussiana quando o número de dados tende a infinito.

As duas próximas seções mostram as escolhas da estatística para esses parâmetros de localização e dispersão. A última seção deste capítulo retoma este exemplo, com a aplicação de todo o conteúdo que será desenvolvido nas demais seções.

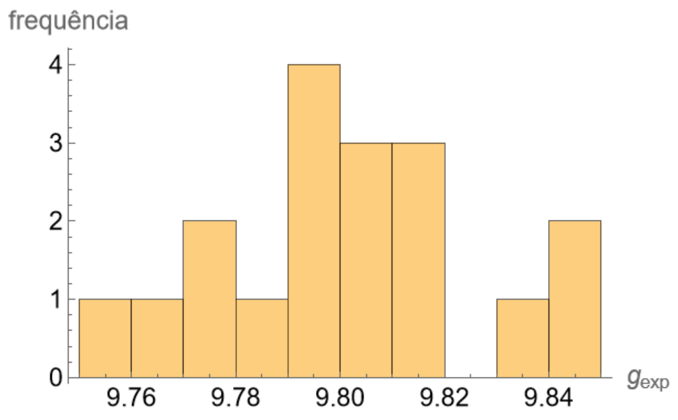

Figura 1. Histograma dos dados da tabela 1. As etiquetas numéricas da abscissa correspondem às fronteiras entre canais, que correspondem a intervalos fechados somente à esquerda, e os pequenos riscos verticais marcam os centros dos canais. Na ordenada, o número de medições que caíram no canal.

II.2 A estimativa da grandeza

As fórmulas desenvolvidas aqui são gerais, assim usaremos x para representar a grandeza, de modo que a medida é o conjunto dos dados $\{x_i, i = 1..N\}$. Estima-se o valor da grandeza por um parâmetro de localização.

Na maioria das medidas diretas da grandeza x, usa-se a média aritmética **(simples)**

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_i + \cdots + x_N}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (2.2)$$

como estimativa do parâmetro de localização.

A média é representada pelo símbolo $\bar{x}$, a fim de diferenciá-la do valor verdadeiro, x_0, e a soma é escrita como um somatório, frequentemente sem definir os valores extremos do iterador i , $\frac{1}{N}\sum_i x_i$, e às vezes até sem identificar o iterador, $\frac{1}{N}\sum x_i$.

Há condições sobre a maneira com que os dados se distribuem aleatoriamente em relação ao valor da grandeza para que a equação (2.2) forneça estimativas adequadas. Assim como nos fenômenos físicos, existem vários modelos para essa flutuação estatística. No *modelo normal*, mostra-se que a média é a melhor estimativa do valor da grandeza e que eliminar os valores extremos dos dados tende a *afastar* $\bar{x}$ de x_0. Assim, não se selecionam dados para estimar a grandeza. A única razão para descartar um dado é a ocorrência de erro na medição, quando o valor obtido deve ser removido do conjunto tanto se estiver longe quanto próximo de $\bar{x}$.

II.3 A estimativa da precisão da medida

A precisão de uma medida se reflete na dispersão dos dados no histograma, assim ela é quantificada por um parâmetro de largura. Primeiro estima-se o valor do quadrado desse parâmetro, que tem o nome de *variância.*

A variância de uma medida é calculada pela fórmula

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \qquad (2.3)$$

A variância não tem a mesma dimensão física e muito menos a mesma unidade que a grandeza (usando $[grandeza]$ para representar o operador de extração da dimensão física da $grandeza$, $[\sigma_x^2] = [x^2] = [x]^2 \neq [x]$, a menos que x seja adimensional), mas sua raiz quadrada, sim. Desse modo, a grandeza usada para quantificar a largura do histograma (e a precisão da medida) é o desvio-padrão, que é a raiz quadrada **positiva** da *variância*.

A precisão de uma medida é quantificada pelo desvio-padrão

$$\sigma_x = \sqrt{\sigma_x^2} \qquad (2.4)$$

O desvio padrão σ_x é, portanto, definido positivo e quantifica a incerteza relacionada a *cada um* dos dados, ou seja, dá ideia da distância entre cada uma das medições do conjunto e o valor da grandeza.

Note que a média inclui todos os N dados, e, provavelmente, cerca de metade deles estará acima do valor verdadeiro, e a outra metade, abaixo. Por isso, as diferenças entre cada dado e o valor verdadeiro tendem a se cancelar no cálculo da média, que é uma soma, de modo que ela, provavelmente, cairá mais próxima do valor verdadeiro que a maioria dos dados. Então, o desvio-padrão da média é menor que o desvio-padrão da medida.

A precisão da média dos dados de uma medida é quantificada pelo desvio-padrão da média

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} \qquad (2.5)$$

A melhora da precisão da média de N dados em relação à precisão de cada dado por esse fator $\sqrt{N}$ da fórmula (2.5) é comum a muitos modelos da flutuação estatística (não todos) e pode ser mostrado de muitas maneiras.

II.4 Interpretação do desvio-padrão e do desvio-padrão da média

O desvio-padrão mede a dispersão dos dados. Os valores da média e do desvio-padrão dos dados da Tabela 1 (que estão no histograma da Figura 1) são 9,801 e 0,026 m/s^2. A tabela 3 mostra quantos dados estão nas faixas em torno da média com meia largura de um, dois e três desvios-padrões.

Tabela 3. Números de medições de g que estão nos intervalos em torno da média com meia largura igual a um, dois e três desvios padrões; calculado usando os dados da Tabela 1.

Metade da largura do intervalo	Intervalo (m/s^2)	frequência
Um desvio-padrão	[9,775 ; 9,827]	12
Dois desvios-padrão	[9,749 ; 9,853]	18
Três desvios-padrão	[9,723 ; 9,879]	18

No modelo normal, em uma medida com N dados, cada dado individual x_i se distribui em torno da média $\bar{x}$ de acordo com as seguintes probabilidades:

$$P(\bar{x} - \sigma_x < x_i < \bar{x} + \sigma_x) \sim 68\% \quad (N > 5)$$

$$P(\bar{x} - 2\sigma_x < x_i < \bar{x} + 2\sigma_x) \sim 95\% \quad (N > 10)$$

$$P(\bar{x} - 3\sigma_x < x_i < \bar{x} + 3\sigma_x) \gtrsim 99\% \quad (N > 20)$$

Fixando um valor de N, pode-se calcular essas probabilidades exatamente.

Já o desvio-padrão da média avalia a dispersão da média em relação ao valor verdadeiro. Há um teorema, conhecido como Teorema Central do Limite, que prova que a soma de um número grande de variáveis aleatórias com desvio-padrão finito[‡] se comporta como descrito no modelo normal. Assim, nas medidas com muitos dados, as probabilidades associadas aos intervalos em torno da média com meia-largura de um, dois, três desvios-padrões da média são as mesmas.

> No modelo normal em uma medida com N dados, a média $\bar{x}$ se distribui em torno do valor verdadeiro x_0 de acordo com as seguintes probabilidades
>
> $$P(x_0 - \sigma_{\bar{x}} < \bar{x} < x_0 + \sigma_{\bar{x}}) \sim 68\% \quad (N > 5)$$
> $$P(x_0 - 2\sigma_{\bar{x}} < \bar{x} < x_0 + 2\sigma_{\bar{x}}) \sim 95\% \quad (N > 10)$$
> $$P(x_0 - 3\sigma_{\bar{x}} < \bar{x} < x_0 + 3\sigma_{\bar{x}}) \gtrsim 99\% \quad (N > 20)$$

Estas relações quantificam as expectativas da diferença entre a estimativa $\bar{x}$ e o valor x_0. Ao isolar x_0 nas desigualdades acima, por exemplo a primeira delas:

$$\left.\begin{aligned} x_0 - \sigma_{\bar{x}} < \bar{x} \quad &\rightarrow\ x_0 < \bar{x} + \sigma_{\bar{x}} \\ \bar{x} < x_0 + \sigma_{\bar{x}} \quad &\rightarrow\ x_0 > \bar{x} - \sigma_{\bar{x}} \end{aligned}\right\} \ \bar{x} - \sigma_{\bar{x}} < x_0 < \bar{x} + \sigma_{\bar{x}}$$

essas mesmas relações podem ser reescritas de outra maneira.

> No modelo normal em uma medida com N dados, o *grau de confiança* com que um intervalo em torno da média $\bar{x}$ contém o valor verdadeiro x_0 é, aproximadamente
>
> $$P(\bar{x} - \sigma_{\bar{x}} < x_0 < \bar{x} + \sigma_{\bar{x}}) \sim 68\% \quad (N > 5)$$
> $$P(\bar{x} - 2\sigma_{\bar{x}} < x_0 < \bar{x} + 2\sigma_{\bar{x}}) \sim 95\% \quad (N > 10)$$
> $$P(\bar{x} - 3\sigma_{\bar{x}} < x_0 < \bar{x} + 3\sigma_{\bar{x}}) \gtrsim 99\% \quad (N > 20)$$

Estas probabilidades têm uma interpretação diferente daquela adotada nos quadros anteriores, uma vez que o valor verdadeiro é desconhecido, mas não flutua estatisticamente. Em relação aos intervalos deste último quadro, a frequência relativa com que a desigualdade é satisfeita no conjunto de medidas com o modelo normal é chamada de nível (ou grau) de confiança. É inerente ao quadro teórico da estatística que não se possa deduzir nada mais concreto que isso em relação a uma medida específica em análise, e temos que nos contentar que os resultados que obtivermos ao longo da vida vão se distribuir de acordo com esses graus de confiança, sem sabermos quais intervalos contém o valor verdadeiro e quais, não.

Assim, essas relações implicam que é provável que a média calculada não difira do valor verdadeiro por mais que um desvio-padrão, é bem provável que $\bar{x}$ não difira de x_0 mais que dois desvios-padrões, é muito provável que $\bar{x}$ não difira de x_0 mais que três desvios-padrões, e é quase certeza que $\bar{x}$ não difere de x_0 mais que cinco desvios-padrões.

[‡] A maioria das variáveis aleatórias observadas nas medidas físicas tem desvio-padrão finito, mas não todas. Por exemplo, a energia de um fóton emitido em uma transição entre níveis de energia discretos de um átomo (um fóton de luz ou um raio-x característico) tem variância infinita.

Essas desigualdades mostram que a ordem de grandeza do erro (erro é a diferença entre os valores medido e o verdadeiro) é próxima do desvio-padrão.

Em uma medida, o erro cometido:

- na medição de um dado, tem a ordem de grandeza do desvio-padrão do conjunto dos dados.
- na média de um conjunto de dados, tem a ordem de grandeza do desvio-padrão da média.

Em ambos os casos, o erro cometido pode ser maior ou menor que o respectivo desvio-padrão, mas poucas vezes será muito menor e raramente muito maior, embora nunca se possa saber seu valor para uma medição ou uma média em particular.

É preciso olhar os números dos quadros acima com cautela. O desvio-padrão é calculado a partir dos valores das medições, assim ele também tem uma flutuação estatística, que complica o cálculo dessas probabilidades, mas que podem ser feitos exatamente no modelo normal. Os números de dados indicados ($N > 5, 10, 20$) são os mínimos para que esses níveis de confiança sejam adotados. Em termos gerais, em um laboratório didático, meia dúzia de repetições da medição são suficientes para um resultado adequado, mas na decisão de vacinar contra a dengue, foram acompanhadas dezenas de milhares de pessoas para avaliar as probabilidades associadas à sua segurança e eficácia. Neste caso, porém, é completamente justificado despender muitos recursos, uma vez que as vidas de milhões de pessoas estão em jogo.

II.5 Estimativa do desvio-padrão quando não se pode repetir a medição

A melhor forma de quantificar a incerteza de uma medição, é calcular o desvio-padrão de uma medida com muitos dados obtidos nas mesmas condições e com instrumentos diferentes, embora produzidos da mesma maneira, e usar as fórmulas (1.3), (1.4) e (1.5). Muitas vezes, isso é impossível ou não é necessário, dado o propósito da medida. Adotaremos a seguinte regra (Vuolo, 1996)[§]:

Na medição com um instrumento graduado, estima-se o desvio-padrão em um valor entre a metade e um terço da menor divisão *que se consegue ler na escala*.

Essa regra envolve, então, a ideia de **erro máximo** na leitura, que é diferente do desvio-padrão, o qual define um intervalo em torno do valor da grandeza que contém cerca de 68% das medições (conforme seção anterior), longe portanto da **certeza**, associada a probabilidades próximas de 100 %. O exemplo abaixo ilustra o uso dessa regra.

Quando se usa uma régua graduada em **milímetros** para medir um objeto regular e em boa posição, é possível ter certeza que o valor está mais próximo de uma marca, da seguinte, ou do ponto central das duas marcas – pode-se ter certeza que o **erro máximo** da leitura é 0,5 mm. A regra do parágrafo precedente indica que o desvio-padrão está na faixa

$$\frac{0{,}5}{3} \leq \sigma_x \leq \frac{0{,}5}{2}$$

Esta regra simples não permite distinguir os diferentes valores dentro do intervalo [0,17; 0,25] representado acima, então escolhemos (arbitrariamente) $\sigma_x = 0{,}2$, um número redondo que

[§] Esta regra pode falhar, dependendo do procedimento de fabricação do instrumento; costuma funcionar para equipamentos de boa qualidade, mas não é possível ter certeza!

facilitará a discussão. O procedimento prático de leitura consiste em idealizar divisões de 0,2 mm da régua e avaliar qual delas está mais próxima da extremidade do objeto. Se sua leitura da régua for 7,82 cm, o resultado deverá ser representado como 7,82 ± 0,02 cm, significando que o comprimento do objeto provavelmente se encontra entre 7,80 cm e 7,84 cm (68% de confiança), entre 7,78 cm e 7,86 cm (95% de confiança) e com certeza está entre 7,76 cm e 7,88 cm. Se forem realizadas dez leituras diferentes, provavelmente três estarão mais longe do que um desvio-padrão, *não importa quão cuidadosa tenha sido sua avaliação*. Por isso, não se gasta muito tempo nessa avaliação. Neste exemplo, as leituras 7,80; 7,82 e 7,84 cm são equivalentes.

Este procedimento se baseia na ideia que **certeza**, na prática, significa cerca de 99% de confiança. Como a meia largura do intervalo de confiança associado a 99% de confiança é aproximadamente 3 desvios-padrões, divide-se o erro máximo por 3. Se, em uma outra medida, você entender que **certeza** significa cerca de 95% de confiança, divida o erro máximo por 2, uma vez que a meia largura do intervalo de confiança associado a 95% de confiança é 2 desvios-padrão. Como o nível de confiança associado à certeza é subjetivo, na prática adota-se um divisor entre 2 e 3. Por sua subjetividade, não é possível detalhar mais este critério, assim este procedimento deve ser reservado para os casos em que o resultado obtido não tenha impacto importante, ou não exista outra maneira de avaliar o desvio-padrão. De todo modo, usar esta aproximação é muito melhor do que ignorar a incerteza da medida.

Finalmente, como é possível estimar a grandeza a partir de um único dado (a medida é um conjunto com apenas uma medição, $N = 1$), usaremos uma notação particular para a estimativa da grandeza. Boa parte dos livros de estatística a identifica por um acento circunflexo sobre o símbolo da grandeza, assim:

$\hat{x}$ é a estimativa de x

que se lê "xis chapéu é a estimativa de x". Na medida com uma única medição $\{x_1\}$, adota-se $\hat{x} = x_1$ e, quando se tomam muitos dados, $\hat{x} = \bar{x}$.

II.6 Erro e tendenciosidade; a lei dos grandes números.

Considere uma grandeza x cuja medida está sujeita a flutuação estatística. Seu valor verdadeiro é x_0 e sua medida resulta no conjunto de dados

$$\{x_i, i = 1..N\}$$

O erro em cada medição é diferente, assim será identificado por ϵ_i, ou seja, cada dado se relaciona com o valor verdadeiro como

$$x_i = x_0 + \epsilon_i \qquad (2.6)$$

A média dos dados é

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = x_0 + \frac{1}{N}\sum_{i=1}^{N} \epsilon_i$$

Como os erros são aleatórios, tanto positivos quanto negativos, o segundo termo da expressão acima, que é a somatória dos N erros, não aumenta tão rápido quanto o número de dados, de modo que

$$\lim_{N\to\infty} \bar{x} = x_0 \qquad (2.7)$$

ou seja, a diferença entre a média e o valor verdadeiro tende a 0 quando N tende a infinito – esse é um teorema, conhecido como Lei dos Grandes Números, cujas condições de validade são frequentemente preenchidas pelos dados experimentais. Medidas que preencham essas condições de validade, mas que a diferença entre a média e o valor verdadeiro não tenda a 0 quando o número de dados cresce, são chamadas tendenciosas. Muitas vezes, determinar se uma medida é tendenciosa é um tema avançado do tratamento estatístico de dados, mas considere o exemplo de um cronômetro digital, cujo dígito menos significativo indica o número de segundos decorridos desde sua partida. Como ele só muda o dígito quando passa um segundo inteiro, a leitura sempre ignorará os décimos de segundo – a medição de um intervalo de tempo subestimará sua duração por um tempo entre 0 e 0,99999... s.

A equação (2.6) e o limite (2.7) expressam ideias centrais do modelo de dados usado na análise dos experimentos do MEXI. Como não existe maneira de provar que a grandeza buscada tenha um valor bem definido, adota-se como fato que ela tem um valor verdadeiro, x_0, que será impossível descobrir, uma vez que são necessárias infinitas medições (sem qualquer erro sistemático!) para isso. Os erros na observação do valor da grandeza são todos devidos ao instrumento, incluindo a pessoa que opera esse instrumento, no caso de escalas analógicas.

Vale a pena notar que há outros modelos. Na econometria, os dados são conhecidos exatamente, e os erros são devidos ao modelo adotado. Por exemplo, a energia elétrica produzida no país é exatamente conhecida, uma vez que as usinas medem esse valor com precisão muito melhor do que pode ser previsto, e qualquer modelo de consumo envolve tantas variáveis de natureza econômica, ambiental e relações sociais praticamente impossíveis de antecipar que não há como prever exatamente quanta energia terá que ser produzida para satisfazer as necessidades. Como a relação (2.6) continua valendo, os métodos estatísticos serão os mesmos, embora com interpretação diferente.

II.7 Precisão vs. acurácia

Aos tamanhos dos erros aleatórios e sistemáticos associam-se os conceitos de precisão e acurácia: quanto menor a variância, maior a precisão, e quanto menor o erro sistemático, maior a acurácia. Note-se que esses conceitos são independentes – a precisão, por ser ligada ao erro aleatório, está ligada a flutuações incontroláveis do processo de medida, enquanto a acurácia relaciona-se a desvios sistemáticos causados por características específicas desse processo. Por isso, é possível ter resultados precisos e inacurados, bem como resultados acurados de baixa precisão.

Quando o processo de medida é preciso, muitas vezes é possível determinar uma regra para o erro sistemático, e calibrar o aparelho, transformando sua leitura de modo a corresponder ao valor conhecido de padrões. A questão da calibração é um tema do tratamento de dados que está fora do escopo deste texto.

Já quando a precisão da medida é tal que o desvio-padrão é muito maior que o erro sistemático, não há necessidade de efetuar nenhuma correção nas leituras, e o erro sistemático pode ser ignorado.

II.8 Valores esperados da grandeza e do desvio-padrão

A seção anterior trouxe uma operação formal com os dados, ao considerar uma medida com infinitos dados, impossível de obter, mas foi isso que permitiu relacionar a média com o valor verdadeiro, importante porque a média está ao alcance da pessoa que analisa o experimento, enquanto o valor verdadeiro é inacessível. Entender o significado dessa operação matemática (determinar uma grandeza em uma medida com infinitos dados) permite interpretar as fórmulas de cálculo do desvio-padrão das grandezas que não são medidas diretamente, mas funções de outras grandezas com desvios padrão conhecidos.

Define-se valor esperado como a média de uma quantidade infinita de dados. Essa operação na variável x é simbolizada $\langle x\rangle$, que se lê "valor esperado de xis". Assim

$$\langle x\rangle = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} x_i = x_0 \qquad (2.8)$$

Ao passar ao limite $N\to\infty$ a fórmula (2.3), encontra-se uma definição formal para a variância:

$$\lim_{N\to\infty}\sigma_x^2 = \lim_{N\to\infty}\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2 = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}(x_i-x_0)^2 = \langle (x-x_0)^2\rangle = \sigma_{x,0}^2 \qquad (2.9)$$

uma vez que, quando $N\to\infty$, $N-1\to N$ e $\bar{x}\to x_0$. Por corresponder ao valor obtido com infinitos dados, esse valor esperado é exato, dependendo apenas da exatidão do modelo de flutuação estatística adotado. Define-se o valor esperado da variância como

$$\sigma_{x,0}^2 = \langle (x-x_0)^2\rangle \qquad (2.9)$$

o que facilitará a construção das fórmulas que aproximam o desvio-padrão das grandezas que são calculadas como funções de várias grandezas medidas com incerteza.

II.9 Algarismos significativos

Como toda medida tem uma precisão finita, não tem sentido representar o valor de uma grandeza física com muitas casas decimais, porque aquelas mais à direita vão ser diferentes a cada repetição da medida. Essas casas a mais passam uma mensagem de precisão que não corresponde aos fatos, distraem a atenção de quem lê o número e são inúteis. Chamamos de *algarismos significativos* os dígitos que estão nas casas que não ficam mudando de maneira totalmente aleatória com a flutuação estatística[**]. A regra geral é:

O número de dígitos do valor registrado deve sugerir sua precisão. Caso o valor seja calculado a partir de outros, deve manter sua precisão e não sugerir uma precisão maior que a das parcelas e fatores.

[**] Veja Vuolo (Vuolo, 1996) para uma discussão sobre a flutuação estatística dos dígitos nas operações aritméticas.

Essa definição um pouco vaga já indica que existem vários conjuntos de regras convencionais para o uso dos algarismos significativos, sem que nenhum deles seja totalmente satisfatório, porque *um único valor* não consegue carregar toda a informação obtida pela medida. Uma representação adequada exige *dois valores*, a estimativa da grandeza e seu desvio-padrão, além do modelo de flutuação estatística adotado na interpretação dos dados, que, se não foi discriminado, indica que se deve adotar o *modelo normal* para interpretar o desvio-padrão.

A questão dos algarismos significativos, porém, não desaparece quando se informam os valores da grandeza e de seu desvio-padrão, uma vez que é necessário escolher quantos dígitos serão registrados. Assim, esta seção se divide em quatro partes, com as regras para:

i) interpretação dos algarismos usados na representação de uma grandeza.
ii) representação dos resultados de operações com algarismos significativos.
iii) arredondamento de números na base decimal.
iv) representação da grandeza por meio da sua estimativa e desvio-padrão.

Essas regras são adotadas por muita gente, mas não são universais. Assim, é importante entender sua motivação, para que você possa escolher o procedimento que adotará na divulgação dos seus resultados de medida, bem como entender outros procedimentos que foram e são usados na literatura. Note que usar procedimentos amplamente adotados ajudará as/os colegas a lerem e entenderem seus escritos de natureza científica.

II.9.i Interpretação dos algarismos usados na representação de uma grandeza

Quando uma grandeza física é representada por apenas um número e sua unidade, supõe-se que a precisão do método usado para determiná-la seja da ordem de grandeza do algarismo mais à direita do número exibido. A Tabela 4 traz, nas cinco primeiras linhas, exemplos em que esse enunciado não é ambíguo, e nas demais, em que a interpretação é dada por um conjunto de regras convencionais (Bevington, Robinson, & Keith, 1992):

1. O *dígito não nulo* mais à esquerda é o dígito **mais** significativo.
2. O dígito **menos** significativo depende de o registro conter ou não o separador decimal:
   - com vírgula → é o *dígito mais à direita*, mesmo que seja zero.
   - sem vírgula → é o *dígito não nulo mais à direita*.
3. Todos os dígitos entre o menos e o mais significativo contam como significativos, independentemente da posição do separador decimal.

Note em primeiro lugar que o número de algarismos significativos não depende do número de casas decimais, uma propriedade muito importante e necessária, uma vez que o significado do valor da grandeza não pode depender da unidade usada, e a posição do separador decimal, depende: 123 cm é o mesmo tamanho que 1,23 m, de modo que é necessário atribuir o mesmo significado às duas representações.

Essa independência de interpretação com a posição do separador decimal traz consequências engraçadas, como a diferença entre 580 e 580, . Por conta do significado da vírgula em uma sentença, não é hábito representar esse valor com três significativos dessa maneira em português, mas sim adotar a notação científica: $5{,}80 \times 10^2$, em que o zero é significativo – caso contrário, não haveria razão para escrevê-lo, enquanto o zero em 580 e 580, são ambos necessários para definir a ordem de grandeza do resultado, 10^3.

Tabela 4. Exemplos de contagem do número de algarismos significativos.

Valor decimal	Dígito menos significativo	Número de significativos
345	5	3
1200,0	0	5 ou 4 ½
0,0004	4	1
43	3	2
22,45	5	4 ou 3 ½
580	8	2
580,	0	3
0,00580	0	3
0,0058	8	2

Finalmente, quando o algarismo mais significativo for 1 ou 2, ele pode ser contado como ½ significativo. Isso porque uma variação de uma unidade (1→2 ou 2→3) modifica completamente o valor, o que não acontece tanto com os algarismos maiores que 3. Além disso, note que 0,099 tem dois significativos; atribuir três significativos a 0,101 é uma mudança muito grande, assim identificar o algarismo 1 na primeira posição como ½ significativo suaviza um pouco essa transição que, porém, é inevitável, e é uma das razões pelas quais as regras de uso de significativos não são totalmente consistentes. Essa questão também tem consequências na maneira de representar os resultados de operações matemáticas e será explicado na próxima subseção. Definir algarismo ½ significativo, bem como decidir se três é ½ ou inteiramente significativo, varia muito na literatura.

II.9.ii Significativos nos resultados de operações.

Adotam-se as seguintes regras para representar o resultado de cálculos aritméticos:

> Nas operações de soma e subtração, o resultado final deve ser escrito até a posição correspondente à posição do algarismo **menos** significativo de **maior valor absoluto.**

Por exemplo

$$25 + 1234 = 1259 \quad \text{e} \quad 0{,}00121 + 0{,}022 = 0{,}023$$

mas

$$2{,}5\times 10^2 + 1234 = 1{,}48\times 10^3$$

Nas operações de multiplicação e divisão, o resultado final deve ser escrito com o mesmo número de significativos do fator com menor número de significativos, por exemplo

$$(2{,}5 \times 10^4)\cdot 1234 = 30850000 = 3{,}1\times 10^7 \quad \text{e} \quad 0{,}00025 \times 1234 = 0{,}31$$

Estes exemplos foram escolhidos para não dar dúvidas, mas é fácil encontrar exemplos em que esse procedimento não funciona. Considere um corpo em queda livre que bate no chão no instante $t = 1{,}45$ s depois de abandonado sem apoio. Sua velocidade no instante de colisão é

$$v = 9{,}8 \cdot 1{,}45 = 14{,}21 \text{ m/s}$$

Assim escrito, o resultado sugere uma precisão relativa de 0,01/14,21 $\sim 0{,}1\% = 1‰$[††], enquanto a precisão relativa da aceleração da gravidade, g, é 0,1/9,8$\sim 1\%$, e a do tempo, t, é 0,01/1,45$\sim 1\%$. Assim, escrever a velocidade como

$$v = 14{,}2 \text{ m/s}$$

sugere uma precisão relativa de 0,1/14,2$\sim 1\%$, compatível com as precisões dos dois fatores usados.

Note, porém, que g tem dois significativos, mas t tem 2 ½ ou 3 significativos, conforme a regra adotada. Assim, muita gente não conta como significativo o primeiro algarismo quando ele é 1 ou 2. Independente da escolha para contagem dos significativos, há uma regra adicional para a multiplicação e subtração, que fica:

> Nas operações de multiplicação e divisão, o resultado final deve ser escrito com o mesmo número de significativos do fator com menor número de significativos desde que o dígito mais significativo do resultado não seja 1 ou 2, quando se retem um significativo a mais desde que não fique com mais significativos que qualquer parcela começando com 1 ou 2.

Enfim, essas regras são insatisfatórias com frequência. As dúvidas se reduzem quando as grandezas são representadas pelos pares de valor estimado e desvio-padrão, e se pode propagar esse desvio-padrão nos cálculos aritméticos – esse é o tema da seção seguinte. De todo modo, será necessário acudir a estas regras para representar o desvio-padrão da grandeza calculada a partir das grandezas medidas.

II.9.iii Arredondamento

A Tabela 5 mostra diversos exemplos de arredondamento, que embutem as seguintes regras:

> O algarismo menos significativo de um número é arredondado conforme o resto que está sendo abandonado, em unidades do valor decimal associado à posição desse algarismo. Se nessa unidade, esse resto, comparado com 0,5, for:
>
> - Menor que 0,5 → arredonda para baixo (abandona simplesmente)
> - Maior que 0,5 → arredonda para cima (acresce um no menos significativo)
> - Igual a 0,5 → arredonda de modo que o menos significativo seja um número par.

A razão para a última parte dessa regra, que leva a arredondar números terminados em 5 para cima metade das vezes e para baixo, a outra metade, é elaborada. Ao somar muitas parcelas arredondadas, esse procedimento evita a tendenciosidade para superestimar ou subestimar a soma, que existiria se 5 sempre arredondasse para cima ou para baixo.

[††] O símbolo ‰ = por mil é menos comum que o % = porcento, mas útil em certos casos.

Tabela 5. Exemplos de arredondamento de números de acordo com as regras enunciadas no texto.

valor	Meta de arredondamento	Valor arredondado
2,34999	0,001	2,350
2,34999	0,01	2,35
2,34999	0,1	2,3
2,34999 × 10^{-5}	10^{-8}	2,350 × 10^{-5}
7,65	0,1	7,6
7,55	0,1	7,6
7,549	0,1	7,5

II.9.iv Representação da grandeza pela sua estimativa e desvio-padrão

Nos casos em que a grandeza de interesse é obtida de uma medida com N dados, tal como na seção II.1:

$$\{g_1, g_2, \dots, g_i, \dots, g_N\} = \{g_i, i = 1..N\} \qquad (2.1)$$

representa-se a grandeza estimada pela média

$$\hat{g} = \bar{g}$$

e fornece-se o desvio-padrão da média, escrevendo

$$\bar{g} \pm \sigma_m \text{ ou } \bar{g}(\sigma_m)$$

de modo que apenas os algarismos significativos dos valores numéricos sejam exibidos.

A fim de seguir um procedimento uniforme em todos os experimentos, estabeleceremos uma regra, que é baseada na ordem de grandeza do algarismo menos significativo do desvio-padrão. A Tabela 6 mostra o desvio-padrão relativo do desvio-padrão em função do número de dados da medida, quando se adota o modelo normal.

Tabela 6. A coluna da direita apresenta a precisão relativa do desvio padrão da média, em função do número de dados na medida, que está na coluna da esquerda. A mesma tabela vale para a relação σ_σ/σ. Resultados válidos para o modelo normal de flutuação estatística dos dados da medida; consulte a bibliografia para sua dedução.

N	$\sigma_{\sigma_m}/\sigma_m$
5	0,35
10	0,24
100	0,07
1000	0,02

Em física, poucas medidas são repetidas milhares de vezes, o mais comum sendo uma dezena de repetições, de modo que adotaremos a seguinte regra prática, bastante comum:

> O resultado de uma medida é representado pelo valor estimado e seu desvio-padrão, com algarismos significativos de acordo com a seguinte regra:
>
> 1) registra-se o *desvio-padrão* com dois algarismos se o mais significativo for 1 ou 2, caso contrário registra-se apenas um algarismo significativo.
>
> 2) registra-se o *valor estimado* da grandeza até a mesma casa decimal que corresponde ao algarismo menos significativo do seu desvio-padrão.

Note que a Tabela 6 é mais abrangente que a regra. Caso verifique que o desvio-padrão foi deduzido de um número grande de dados, ignore a regra e registre tantos significativos quantos fizerem sentido.

II.10 Exemplo: uma medida da aceleração local da gravidade

A Tabela 1 da seção II.1 lista os dados de uma medida da aceleração local da gravidade, que está também representada pelo histograma da Figura 1. Esta seção traz todos os resultados dos diversos cálculos e procedimentos apresentados neste capítulo.

A estimativa da aceleração local da gravidade, calculada pela fórmula (2.2), é 9,80083 m/s^2. Os valores do desvio-padrão e do desvio-padrão da média, calculados pelas fórmulas (2.3) a (2.5), são 0,02570 e 0,00606, em m/s^2. Assim,

$$\bar{g} = 9{,}80083 \quad \text{e} \quad \sigma_m = 0{,}00606 \quad \text{em m/s}^2$$

Uma mudança de apenas 0,001 em qualquer dado da Tabela modifica os dois dígitos menos significativos do valor de $\bar{g}$ exibido acima, portanto não se deve relatar tantos dígitos assim. Neste caso, a medida de g teve 18 dados, portanto $\sigma_{\sigma_m}/\sigma_m \sim 0{,}2$ conforme a Tabela 6. Se usássemos 0,0061, sugeriríamos que a precisão relativa do desvio padrão é 0,0001/0,0061~2%, muito mais preciso que obtido. Assim, registra-se

$$\sigma_m = 0{,}006$$

com precisão relativa da ordem de 0,001/0,006=0,18=18%.

Uma vez escolhido que a representação numérica do desvio-padrão requer 3 dígitos depois da vírgula, a média deve ser registrada até essa mesma casa. Assim, adotando o procedimento do item II.9.iv,

$$g = 9{,}801(6) \;\; \text{ou} \;\; g = 9{,}801 \pm 0{,}006 \text{ m/s}^2$$

Na notação com parênteses, o desvio-padrão é o valor entre parênteses em unidades do dígito menos significativo da estimativa.

Esse resultado pode ser confrontado com o valor medido, por um gravímetro de precisão, da aceleração da gravidade no local em que o experimento foi efetuado (Instituto de Física da Universidade de São Paulo, campus Butantan, São Paulo, SP, Brasil): $g = 9{,}786$ m/s^2.

Sempre é interessante verificar se os dados se distribuem de acordo com o modelo normal. Para isso, construímos a Tabela 7, com base na Tabela 3, transformando frequência em frequência relativa e acrescentando uma coluna com os valores esperados da probabilidade.

Tabela 7. Porcentagens dos números de medições nos intervalos simétricos em torno da média com meias-larguras de um, dois e três desvios-padrões, comparados às probabilidades do modelo normal. Dados da Tabela 1.

Metade da largura do intervalo em torno da média	Intervalo (m/s^2)	Frequência relativa	Probabilidade no modelo normal
Um desvio-padrão	[9,775 ; 9,827]	67 %	$\sim$ 68%
Dois desvios-padrão	[9,749 ; 9,853]	100 %	$\sim$ 95%
Três desvios-padrão	[9,723 ; 9,879]	100%	$\gtrsim$ 99%

A figura abaixo acrescenta os valores das grandezas estatísticas ao histograma da Figura 1.

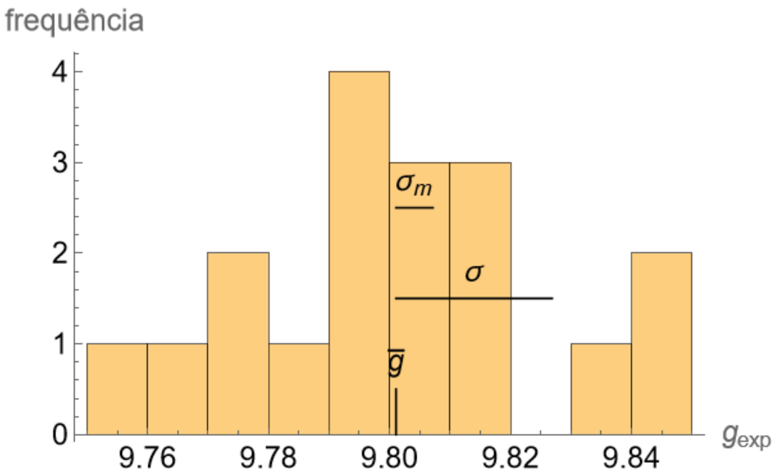

Figura 2. Histograma dos dados da medida de g (Tabela 1), representando os valores das grandezas estatísticas calculadas.

III. O desvio-padrão de grandezas deduzidas de outras

Certas grandezas são medidas diretamente; é o caso tanto da *posição* do corpo (medida na trena ou no quadriculado nas experiências do MEXI) quanto do *instante* em que o quadro do filme foi registrado (definido pelo cronômetro da filmadora e estampado no *time code*). Há outras grandezas, entretanto, cuja medição não é direta, como a *velocidade* de um carrinho ao longo de uma trajetória, uma vez que não existe uma “régua de velocidades”. Nos experimentos com imagens, porém, podemos calculá-las a partir das grandezas que a definem (posição e tempo). Embora existam fórmulas gerais para calcular o desvio-padrão de uma função de várias grandezas a partir dos seus desvios-padrão, elas são frequentemente trabalhosas de aplicar, além de haver muitas exceções. Por isso, o resto desta seção está destinado a orientá-lo e ajudá-lo a compreender o modo pelo qual os desvios-padrões são calculados em cada experimento. Esse assunto é muitas vezes chamado de *propagação de incertezas*.

Este capítulo desenvolve um procedimento de cálculo do desvio-padrão de grandezas deduzidas de outras a partir de alguns conceitos e casos particulares, incluindo uma fórmula de uso bastante amplo. O capítulo IV apresenta expressões simplificadas para cálculos recorrentes, enquanto o capítulo V traz os procedimentos necessários na análise de alguns experimentos em que as fórmulas do capítulo III não são válidas.

III.1 Quando uma das grandezas é constante

Em um experimento, é comum deduzir-se grandezas físicas por meio de fórmulas com muitos parâmetros, alguns deles constantes fundamentais (carga do elétron, constante de gravitação, constante dos gases) ou que são conhecidas com muita precisão naquele experimento, como o tempo nos quadros do MEXI, obtidos de um relógio de quartzo da filmadora, com precisão de microssegundos. Assim, certos valores que entram nas fórmulas podem ser considerados exatos, o que facilita não só realizar os cálculos, mas também interpretá-los. Quão pequeno precisa ser o desvio-padrão de uma grandeza para ser ignorado depende, muitas vezes, de considerar os resultados desta seção e da próxima, mas isso não invalida as propriedades básicas apresentadas a seguir.

Usamos exemplos do MEXI para tornar a discussão mais concreta, mas não é necessário conhecer a física subjacente nem as fórmulas do modelo físico do exemplo para entender as propriedades estatísticas apresentadas.

II.1.a Desvio-padrão de um valor multiplicado por uma constante

Na medida da quantidade de movimento p de um corpo de massa m com velocidade v,

$$p = m\,v$$

a massa costuma ser medida por uma balança, que facilmente tem precisão melhor que uma parte em mil, enquanto a velocidade depende da diferença de posições, que tem um erro considerável. Assim, pode-se ignorar a incerteza na massa, de modo que a incerteza em p provém exclusivamente da medida da velocidade, cujo desvio padrão será designado por σ_v.

Esse raciocínio permite entender que a grandeza m atua como um fator de escala – cada erro na medida de v produz um erro em p correspondente ao seu valor multiplicado por m, de modo que o desvio-padrão da quantidade de movimento fica

$$\sigma_p = m\,\sigma_v$$

Essa é uma regra geral: constantes multiplicativas fatoram no cálculo do desvio-padrão da grandeza deduzida de outra, sujeita a flutuação estatística. Escrevemos a regra em destaque usando valores genéricos para as grandezas:

Se $y = c\,x$, com c constante e sem erro, então $\hat{y} = c\,\hat{x}$ e $\sigma_y = |c|\sigma_x$ (3.1)

Como essa expressão é geral, a constante c pode ser negativa, mas o desvio-padrão é uma grandeza definida positiva, de modo que na expressão (3.1) usa-se o módulo de c.

III.1.b Desvio-padrão de uma grandeza somada com uma constante

No experimento de atrito variável, a componente da força de atrito ao longo do plano na direção do eixo de maior inclinação, f_y , é calculada a partir da componente da força resultante na direção y, F_y, pela fórmula

$$f_y = F_y + m\,g\,\mathrm{sen}\,\theta$$

a fim de considerar a contribuição do peso na força resultante. As medidas da massa m, da aceleração da gravidade g e do ângulo θ são muito mais precisas que as da força, F_y, de maneira que o desvio-padrão do termo $m\,g\,\mathrm{sen}\,\theta$ pode ser ignorado e o desvio-padrão da componente da força de atrito é igual ao da força resultante medida,

$$\sigma_{fy} = \sigma_{Fy}$$

Essa é outra regra geral: constantes aditivas podem ser ignoradas no cálculo do desvio-padrão da grandeza deduzida de outra, sujeita a flutuação estatística.

Se $y = x + c$, com c uma constante conhecida sem erro, então

$\hat{y} = \hat{x} + c$ e $\sigma_y = \sigma_x$ (3.2)

III.2 O desvio-padrão de uma função da grandeza medida com flutuação estatística

No experimento sobre conservação da energia, mede-se a velocidade v e deduz-se a energia cinética

$$E_c = \frac{m}{2}v^2$$

A massa é conhecida com muita precisão, de modo que o fator $m/2$ pode ser considerado constante, e calcula-se

$$\sigma_{Ec} = \frac{m}{2}\sigma_{v^2}$$

Neste caso, portanto, é necessário determinar o desvio-padrão do quadrado de uma grandeza sujeita a flutuação estatística.

Este problema é um caso particular de um problema mais geral, que é o de determinar o desvio-padrão de uma função $f(x)$, em que a grandeza x é medida com flutuação estatística. O caso $f(x) = x^2$ tem uma solução particular que ajuda a entender o problema geral, de modo que começamos por ele.

III.2.a Cálculo do desvio-padrão do quadrado da grandeza

Uma experimentadora quer saber o desvio-padrão da área de quadrados possivelmente idênticos e mede o tamanho de um lado de cada um deles; x representa o tamanho do lado, e sua medida está sujeita a flutuação estatística. Ela então obtém o conjunto de dados $\mathbb{X}$

$$\mathbb{X} = \{x_i, i = 1..N\}$$

Com esses dados, calcula a média $\bar{x}$, o desvio-padrão σ_x e o desvio-padrão da média σ_m, respectivamente com as fórmulas (2.2), (2.3) e (2.5).

A área de um quadrado[‡‡] se relaciona com o tamanho do lado pela expressão

$$S = x^2$$

O problema a resolver é determinar a estimativa da área, $\hat{S}$, e seu desvio-padrão, $\sigma_{\hat{S}}$.

Um procedimento possível (e correto) é construir o conjunto $\mathbb{S}$ das áreas, cujos elementos são $S_i = {x_i}^2$:

$$\mathbb{S} = \{{x_i}^2, i = 1..N\}$$

e, dos valores desse conjunto, calcular a média e o desvio-padrão σ_S, depois usar como estimativas da área e do seu desvio-padrão $\bar{S}$ e $\sigma_S/\sqrt{N}$. Ao invés de dar um exemplo com números e fazer a conta com uma calculadora, realizamos algebricamente este procedimento, que está na raiz de um método bastante comum de cálculo da variância de funções.

Quando x_0 é o valor verdadeiro do tamanho do lado e seleciona-se um dos dados de $\mathbb{X}$, x_k, pode-se usar a relação (1.6)

$$x_k = x_0 + \epsilon_k$$

em que ϵ_k é o erro da medição desse dado, suposto pequeno, ou seja, $|\epsilon_k| \ll x_0$. A fim de fazer uma figura para a representação geométrica deste problema e facilitar a discussão, supõe-se

$$\epsilon_k > 0$$

A ilustra a relação entre o quadrado de lado x_k (medido) e o quadrado de lado x_0 (verdadeiro, mas ao qual não temos acesso). Da figura, vê-se que

$$S_k \cong {x_0}^2 + 2x_0\epsilon_k$$

em que se ignorou o pequeno quadrado listrado de área ${\epsilon_k}^2$, uma vez que a área de cada um dos pequenos retângulos de erro é $x_0\epsilon_k \gg {\epsilon_k}^2$ porque $x_0 \gg \epsilon_k$ com $\epsilon_k > 0$.

[‡‡] Esta seção destina-se a explicar o procedimento de cálculo do desvio padrão da energia cinética, proporcional ao quadrado da velocidade. Nesta analogia com um quadrado no espaço, então, deseja-se que os lados sejam idênticos, o que não acontece, por exemplo, quando se corta um quadrado de papel com uma tesoura. Não é realista esperar que quadrados de papel tenham lados idênticos, mas para os propósitos desta subseção é essencial considerar quadrados ideais, com os lados idênticos e perpendiculares.

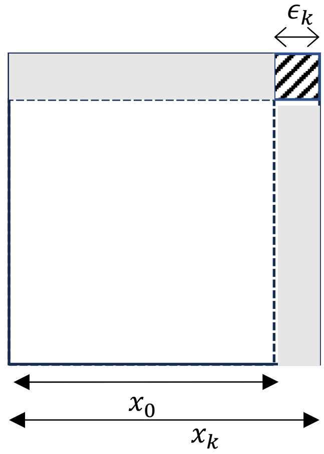

Figura 3. As linhas tracejadas delimitam o quadrado verdadeiro e as linhas cheias, o quadrado medido. A diferença entre eles é o erro na área do quadrado, formado pelos dois retângulos cinza e o pequeno quadrado hachurado.

Calculando o valor esperado, obtém-se

$$\langle S \rangle \cong {x_0}^2 = S_0$$

Para calcular a variância, usa-se a fórmula (2.9) e retira-se a restrição $\epsilon_k > 0$, que só foi necessária para fazer a figura; esse erro genérico vai ser simbolizado por ϵ, apenas. Assim

$${\sigma_S}^2 = \langle (S - S_0)^2 \rangle = \langle ({x_0}^2 + 2x_0\epsilon - S_0)^2 \rangle = \langle (2x_0\epsilon)^2 \rangle = 4\,{x_0}^2 \langle (\epsilon)^2 \rangle = 4{x_0}^2{\sigma_x}^2$$

Extraindo a raiz quadrada positiva para calcular o desvio-padrão de $S = x^2$, obtém-se

$$\sigma_S = 2x_0\sigma_x$$

Como não é possível conhecer o valor verdadeiro x_0, usa-se $\bar{x}$ no seu lugar, de modo que se adota

$$\sigma_S \cong 2|\bar{x}|\sigma_x \qquad (3.3)$$

Note que nessa fórmula aparece $|\bar{x}|$, a fim de garantir que o desvio-padrão seja positivo. Neste exemplo da medida da área de um quadrado a partir do seu lado, tanto x_0 quanto $\bar{x}$ são definidos positivos, mas a fórmula (3.3), escrita com o módulo do valor médio de x, vale para o quadrado de qualquer grandeza, mesmo negativa, o que amplia sua aplicabilidade. Neste caso particular do quadrado de uma grandeza, a fórmula (3.3) vale desde que a área do pequeno quadrado de lado ϵ possa ser ignorado, ou seja

$$x_0\epsilon_k \gg {\epsilon_k}^2 \Longrightarrow x_0 \gg |\epsilon_k|$$

Como $|\epsilon_k| \sim \sigma$, conforme a discussão da seção II.4, a condição de validade é

$$\sigma_x \ll x_0$$

uma vez que o erro ϵ é muito provavelmente da ordem de grandeza do desvio-padrão. Note que essa condição garante que $\bar{x} \cong x_0$, justificando também essa aproximação usada para obter a expressão (3.3). O quanto o desvio-padrão tem que ser menor que o valor da grandeza depende do problema, mas um fator 10 é suficiente na grande maioria dos problemas reais.

A próxima subseção usa a mesma linha de raciocínio para generalizar o procedimento para qualquer função de uma única grandeza sujeita à flutuação estatística.

II.2.b Cálculo do desvio-padrão de uma função de uma única variável

Na relação entre os desvios-padrão de x^2 e x da seção anterior, fórmula (3.3), o fator $2x_0$ vem da dependência da área do quadrado com o tamanho do lado: quando o lado de um quadrado aumenta de um certo valor ϵ, a área aumenta ao longo de dois lados, portanto em uma taxa que é duas vezes o comprimento do lado. Então, o desvio-padrão de uma grandeza y tal que

$$y = y(x)$$

dependerá de quanto y varia com o erro em x, ϵ, ou seja, depende da relação

$$\Delta y = \text{função}(\epsilon)$$

A função $y(x)$ pode ser expandida em série de Taylor em torno de x_0:

$$y = y(x_0) + \left.\frac{dy}{dx}\right|_{x=x0} (x - x_0) + \frac{1}{2!}\left.\frac{d^2y}{dx^2}\right|_{x=x0} (x - x_0)^2 + \cdots =$$

$$y_0 + \left.\frac{dy}{dx}\right|_{x=x0} \epsilon + \frac{1}{2!}\left.\frac{d^2y}{dx^2}\right|_{x=x0} \epsilon^2 + \cdots =$$

Quando o erro ϵ for pequeno e a função $y(x)$ bem comportada, os termos da ordem de ϵ^2 e superiores podem ser ignorados, portanto

$$y \cong y_0 + \left.\frac{dy}{dx}\right|_{x=x0} \epsilon \Longrightarrow y - y_0 \cong \left.\frac{dy}{dx}\right|_{x=x0} \epsilon$$

Efetuando o cálculo da variância de modo análogo ao do item anterior,

$$\sigma^2{}_y = \langle (y - y_0)^2 \rangle \cong \langle \left(\left.\frac{dy}{dx}\right|_{x=x0} \epsilon \right)^2 \rangle = \left(\left.\frac{dy}{dx}\right|_{x=x0} \right)^2 \langle (\epsilon)^2 \rangle = \left(\left.\frac{dy}{dx}\right|_{x=x0} \right)^2 \sigma_x{}^2$$

Por desconhecer x_0, é necessário aproximar

$$\left.\frac{dy}{dx}\right|_{x=x0} \cong \left.\frac{dy}{dx}\right|_{x=\bar{x}}$$

que, substituído na expressão anterior, dá a regra usual de cálculo do desvio-padrão:

Se $y = y(x)$ então $\hat{y} \cong y(\hat{x})$, $\sigma^2{}_y = \left(\left.\frac{dy}{dx}\right|_{x=\bar{x}} \right)^2 \sigma^2{}_x$ e $\sigma_y \cong \left| \left.\frac{dy}{dx}\right|_{x=\bar{x}} \right| \sigma_x$ (3.4)

desde que $\sigma_x \ll |x_0|$ e $y(x)$ seja *bem comportada* (exponencial e tangente nem sempre são bem comportadas para fins desta fórmula, o que é analisado na sub-seção III.3.c).

Nessa fórmula (3.4), o módulo do valor da derivada é necessário, uma vez que o desvio-padrão é a raiz positiva da variância.

Confira que a fórmula (3.3) é um caso particular dessa regra geral (3.4).

III.3 Desvio-padrão de uma função de várias grandezas sujeitas a flutuação estatística

Nos experimentos com imagens, a velocidade v de um móvel é determinada indiretamente por meio de medições, em dois quadros, das posições x_2 e x_1, e dos respectivos tempos t_2 e t_1:

$$v = \frac{x_2 - x_1}{t_2 - t_1}$$

O desvio-padrão da velocidade dependerá dos desvios padrão das grandezas medidas diretamente, ou seja, das posições e dos respectivos instantes. Situações como essa são comuns, de modo que vamos desenvolver uma fórmula genérica para funções de várias variáveis com desvios-padrões conhecidos, mas antes mostraremos alguns casos particulares para ajudar na compreensão dessa fórmula genérica.

III.3.a A soma de variáveis sujeitas a flutuação estatística

Considere duas variáveis a e b, ambas sujeitas a flutuação estatística, e a grandeza u, que é a soma das duas:

$$u = a + b$$

A fim de fazer uso das definições da seção anterior, imagine infinitas repetições das medições. Para cada valor medido de a e b,

$$a = a_0 + \epsilon$$

$$b = b_0 + \delta$$

Substituindo essas expressões na definição de u

$$u = a + b = a_0 + \epsilon + b_0 + \delta \rightarrow$$

$$\langle u \rangle = a_0 + b_0 + \langle \epsilon \rangle + \langle \delta \rangle = a_0 + b_0 = u_0$$

desde que as medidas de a e b não sejam tendenciosas. A variância de u fica

$$\begin{aligned} \sigma^2{}_u = \langle (u - u_0)^2 \rangle = \langle (\epsilon + \delta)^2 \rangle = \langle \epsilon^2 + 2\epsilon\delta + \delta^2 \rangle = \langle \epsilon^2 \rangle + \langle \delta^2 \rangle + \langle 2\epsilon\delta \rangle \\ = \sigma_a^2 + \sigma_b^2 + \langle 2\epsilon\delta \rangle \end{aligned}$$

Aqui, há duas possibilidades:

i. o erro de medição da grandeza a não afeta a medição da grandeza b.

ii. o erro de medição de uma grandeza afeta a outra.

No segundo caso, o valor médio desse erro combinado precisa ser estimado. Essa grandeza tem o nome de covariância e é comum nas medidas em física e no tratamento de dados experimentais, mas lidar com ela está fora do escopo deste texto inicial. Já no primeiro caso, o último termo é nulo, $\langle 2\epsilon\delta \rangle = 0$, porque será a média de infinitos valores com tantas contribuições positivas quanto negativas. Nesse caso, diz-se que a e b são *estatisticamente independentes*, significando que o erro na observação de uma grandeza não afeta a outra.

Esse resultado pode ser generalizado para um número qualquer de grandezas *a, b, c,...*, dando origem à regra:

> Se $u = a + b + c + \cdots$, estima-se a grandeza por $\hat{u} = \hat{a} + \hat{b} + \hat{c} + \cdots$
>
> e seu desvio-padrão por $\sigma_u = \sqrt{\sigma_a^2 + \sigma_b^2 + \sigma_c^2 + \cdots}$ (3.5)
>
> desde que as grandezas $a, b, c, \ldots$ sejam estatisticamente independentes (a medida de uma delas não afeta a medida da outra).

III.3.b Desvio-padrão de uma função de várias variáveis sujeitas a flutuação estatística

Seja y uma grandeza que é função das grandezas $a, b, \ldots, z$, $y = f(a, b, \ldots, z)$. No exemplo da velocidade, y é a velocidade, enquanto que a, b, c e d são x_1, x_2, t_1 e t_2.

Cada uma das variáveis é medida com um erro, de modo que, para cada dado de uma delas, vale a expressão

$$a = a_0 + \epsilon_a$$

$$b = b_0 + \epsilon_b$$

e assim por diante para $c, d, \ldots, z$.

O cálculo a seguir combina as estratégias das subseções III.2.b e III.3.a. Começa-se com a expansão de f em torno dos valores verdadeiros das grandezas $a, b, \ldots, z$ pela expressão

$$f(a, b, \ldots, z) \cong f(a_0, b_0, \ldots, z_0) +$$
$$\left.\frac{\partial f}{\partial a}\right|_{\substack{a=a0\\ b=b0\\ \cdots\\ z=z0}} (a - a_0) + \left.\frac{\partial f}{\partial b}\right|_{\substack{a=a0\\ b=b0\\ \cdots\\ z=z0}} (b - b_0) + \cdots + \left.\frac{\partial f}{\partial z}\right|_{\substack{a=a0\\ b=b0\\ \cdots\\ z=z0}} (z - z_0)$$

que corresponde à generalização da série de Taylor para várias variáveis,[§§] e já está truncada nos termos de primeira ordem, de modo que esta será uma aproximação válida quando os desvios-padrões relativos forem pequenos. A fim de compactar a notação, define-se

$$\frac{\partial f}{\partial a_0} = \left.\frac{\partial f}{\partial a}\right|_{\substack{a=a0\\ b=b0\\ \cdots\\ z=z0}}, \frac{\partial f}{\partial b_0} = \left.\frac{\partial f}{\partial b}\right|_{\substack{a=a0\\ b=b0\\ \cdots\\ z=z0}}, \ldots, \frac{\partial f}{\partial z_0} = \left.\frac{\partial f}{\partial z}\right|_{\substack{a=a0\\ b=b0\\ \cdots\\ z=z0}}$$

Como as diferenças das variáveis em relação aos valores verdadeiros são os erros $\epsilon_a, \epsilon_b, \ldots$, a expressão acima fica

$$f(a, b, \ldots, z) \cong f(a_0, b_0, \ldots, z_0) + \frac{\partial f}{\partial a_0}\epsilon_a + \frac{\partial f}{\partial b_0}\epsilon_b + \cdots + \frac{\partial f}{\partial z_0}\epsilon_z$$

Isolando nessa expressão o erro em f

$$\epsilon_f = f(a, b, \ldots, z) - f(a_0, b_0, \ldots, z_0)$$

[§§] O símbolo $\frac{\partial f}{\partial a}$ representa a derivada parcial de *f* em relação a *a*, ou seja, a derivada da função *f* quando apenas *a* é tomada como variável, enquanto *b*, *c*, ..., *z* são consideradas constantes. Por exemplo, considere a função $f = f(a, b)$ definida por:

$$f(a, b) = a^2b^3 + 5a + 7b$$

As derivadas parciais da função f em relação a a, b são: $\begin{cases} \frac{\partial f}{\partial a} = 2ab^3 + 5 \\ \frac{\partial f}{\partial b} = 3a^2b^2 + 7 \end{cases}$

portanto as derivadas parciais calculadas no ponto $(\hat{a}, \hat{b})$ são: $\begin{cases} \left.\frac{\partial f}{\partial a}\right|_{\hat{a},\hat{b}} = 2\hat{a}\hat{b}^3 + 5 \\ \left.\frac{\partial f}{\partial b}\right|_{\hat{a},\hat{b}} = 3\hat{a}^2\hat{b}^2 + 7 \end{cases}$

obtém-se

$$\epsilon_f \cong \frac{\partial f}{\partial a_0}\epsilon_a + \frac{\partial f}{\partial b_0}\epsilon_b + \cdots + \frac{\partial f}{\partial z_0}\epsilon_z$$

A variância de y será, então

$$\sigma_y^2 = \langle \left(\frac{\partial f}{\partial a_0}\epsilon_a + \frac{\partial f}{\partial b_0}\epsilon_b + \cdots + \frac{\partial f}{\partial z_0}\epsilon_z\right)^2 \rangle$$

que, com a condição de que as grandezas $a, b, \ldots, z$ sejam estatisticamente independentes, fica

$$\sigma_y^2 \cong \left(\frac{\partial f}{\partial a_0}\right)^2 \sigma_a^2 + \left(\frac{\partial f}{\partial b_0}\right)^2 \sigma_b^2 + \cdots + \left(\frac{\partial f}{\partial z_0}\right)^2 \sigma_z^2$$

Esse resultado é uma combinação dos resultados das subseções III.2.d e III.3.a. Na prática, não se conhecem os valores verdadeiros, assim é preciso aproximar ainda mais o cálculo e substituir os valores verdadeiros pelas estimativas das grandezas: $\hat{a}, \hat{b}, \ldots, \hat{z}$.

Se $y = f(a, b, \ldots, z)$, então $\hat{y} = f\left(\hat{a}, \hat{b}, \ldots, \hat{z}\right)$ e sua variância é

$$\sigma_y^2 \cong \left(\frac{\partial f}{\partial \hat{a}}\right)^2 \sigma_a^2 + \left(\frac{\partial f}{\partial \hat{b}}\right)^2 \sigma_b^2 + \cdots + \left(\frac{\partial f}{\partial \hat{z}}\right)^2 \sigma_z^2 \qquad (3.6)$$

em que $\frac{\partial f}{\partial \hat{a}} = \left.\frac{\partial f}{\partial a}\right|_{\substack{a=\hat{a}\\ b=\hat{b}\\ \cdots\\ z=\hat{z}}}$ é a derivada parcial de f com respeito a a calculada para os valores estimados das grandezas, $\left(\hat{a}, \hat{b}, \ldots, \hat{z}\right)$, e analogamente para as demais, e desde que as grandezas $a, b, \ldots, z$ sejam estatisticamente independentes, $\sigma_a \ll |a_0|, \sigma_b \ll |b_0|, \ldots, \sigma_z \ll |z_0|$, e $f(a, b, \ldots, z)$ seja bem comportada (para exponencial e tangente, por exemplo, a fórmula pode não ser válida).

Vale a pena lembrar que os desvios-padrão σ_f, σ_a, σ_b, ..., σ_z são definidos positivos, obtidos como a raiz quadrada positiva da variância da variável correspondente. Não se pode tratar os desvios padrão como as médias (a média da soma é a soma das médias); nunca se somam ou subtraem desvios padrões, apesar de as variâncias serem aditivas.

III.3.c. Condições de validade da fórmula do desvio-padrão de uma função de várias variáveis

A fórmula (3.6) é uma boa aproximação desde que as seguintes condições sejam obedecidas:

i) As variáveis do conjunto $\{a, b, \ldots, z\}$ são estatisticamente independentes. Isso significa não apenas que a medida de qualquer uma delas não interfere na medida de todas as outras, como também que nenhuma das variáveis é simplesmente calculada a partir de uma ou algumas das variáveis desse conjunto. Se isso acontecer, é necessário expandir suas fórmulas na expressão de f **antes** de aplicar a fórmula (3.6); veja um exemplo na subseção IV.1.b. O que ocorre é que cada termo dessa equação precisa representar toda a dependência da função com relação à variável sobre a qual se calculou a derivada parcial; não pode haver dependências escondidas nas outras variáveis. Por outro lado, se alguma das grandezas $a, b, \ldots, z$ for medida diretamente a partir de outras grandezas que não estejam nesse conjunto, pode-se aplicar a fórmula (3.6) isoladamente a ela e usar o valor obtido para sua variância na fórmula (3.6) para calcular σ_y^2.
ii) Para cada grandeza $x = a, b, \ldots$ ou z, é preciso que valha a desigualdade $\sigma_x \ll |\hat{x}|$. Em muitas das fórmulas detalhadas adiante, essas grandezas aparecem no denominador, o que causará dificuldades matemáticas se ela puder ser nula, uma vez que seu inverso não estará definido.

Mesmo que $\hat{x}$ não seja nulo, quando $\sigma_x \approx |\hat{x}|$, há alguma probabilidade de uma dada observação de x dar resultado nulo. A variância de uma grandeza é definida pela relação (2.9), que considera infinitas observações: é o valor esperado da média de infinitas observações, assim nunca poderá conter um valor indefinido, mesmo que tenha baixa probabilidade de acontecer.

iii) A condição de suavidade da função $f(a, b, \ldots, z)$ vem da necessidade de truncar a expansão em série de Taylor. Por exemplo, se $f(a) = \tan a$, com $a_0 \approx \frac{\pi}{4}$ e $\sigma_a \sim \frac{\pi}{16}$, os valores de a acima de a_0 serão todos maiores que 1 e alguns muito grandes; já os valores abaixo de a_0 serão todos menores que 1 e no mínimo próximos a 0. Nessas condições, o valor esperado de $f(a)$, $\langle f(a) \rangle$,será muito maior que $f(a_0) = 1$, e a dedução da equação (3.6) usou $\langle f(a) \rangle = f(a_0)$, invalidando a aproximação. Porém, se $a_0 \approx \frac{\pi}{4}$ e $\sigma_a \lesssim \frac{\pi}{100}$, a fórmula (3.6) é uma aproximação aceitável para a variância de $f(a) = \tan a$.

IV Estimativa do desvio-padrão de funções de grandezas sujeitas a flutuação estatística – fórmulas recorrentes

Nas seções a seguir, estão os resultados da aplicação da fórmula (3.6) a algumas composições de variáveis mais comuns no cálculo de grandezas experimentais.

IV.1 Subtração de grandezas

A subseção III.3.a apresenta o cálculo do desvio-padrão da soma de duas grandezas. O resultado obtido lá vale também para a subtração, o que pode ser entendido de muitas formas. Aqui, isso está mostrado a partir da fórmula (3.6).

Considere a grandeza

$$u = a - b$$

em que a e b são grandezas sujeitas a flutuação estatística.

Aplicando a fórmula (3.6)

$$\sigma_u^2 = \left(\frac{\partial u}{\partial \hat{a}}\right)^2 \sigma_a^2 + \left(\frac{\partial u}{\partial \hat{b}}\right)^2 \sigma_b^2$$

em que

$$\frac{\partial u}{\partial \hat{a}} = \left.\frac{\partial u}{\partial a}\right|_{\substack{a=\hat{a}\\ b=\hat{b}}} = 1 \quad \text{e} \quad \frac{\partial u}{\partial \hat{b}} = \left.\frac{\partial u}{\partial b}\right|_{\substack{a=\hat{a}\\ b=\hat{b}}} = -1$$

obtém-se:

$$\sigma_u^2 = \sigma_a^2 + \sigma_b^2$$

ou seja, a fórmula da variância da soma é a mesma da variância da diferença – tanto faz somar ou subtrair, as incertezas de medição se acumulam, mas na soma, o resultado aumenta, na subtração, o resultado diminui. Assim, quando as duas grandezas têm mesmo sinal, o desvio-padrão **relativo** da sua diferença é muito maior que da soma.

Não destacamos este resultado, que se aplica a um número qualquer de variáveis somadas ou subtraídas, uma vez que corresponde à fórmula (3.5), em que não há qualquer restrição a grandezas negativas, e uma subtração pode ser entendida como a soma com o oposto da grandeza.

IV.2 Produto e quociente

IV.2.a Produto de duas grandezas

Aplicando-se a fórmula (3.6) à quantidade de movimento linear

$$p = m\,v$$

em que m é a massa e v é a velocidade da partícula, obtém-se:

$$p(m, v) = m\,v \Rightarrow \sigma_p^2 = \left(\frac{\partial p}{\partial \hat{m}}\right)^2 \sigma_m^2 + \left(\frac{\partial p}{\partial \hat{v}}\right)^2 \sigma_v^2$$

Substituindo

$$\frac{\partial p}{\partial \widehat{m}} = \left.\frac{\partial p}{\partial m}\right|_{\substack{m=\widehat{m}\\ v=\hat{v}}} = \hat{v} \quad e \quad \frac{\partial p}{\partial \hat{v}} = \left.\frac{\partial p}{\partial v}\right|_{\substack{m=\widehat{m}\\ v=\hat{v}}} = \widehat{m}$$

obtém-se:

$$\sigma_p^2 = \bar{v}^2\sigma_m^2 + \bar{m}^2\sigma_v^2$$

Fatorando $\hat{p} = \widehat{m}\,\hat{v}$ no membro direito, obtém-se

$$\sigma_p^2 = \left(\frac{\widehat{m}\,\hat{v}}{\widehat{m}}\right)^2 \sigma_m^2 + \left(\frac{\widehat{m}\,\hat{v}}{\hat{v}}\right)^2 \sigma_v^2 = p^2\left[\left(\frac{\sigma_m}{\widehat{m}}\right)^2 + \left(\frac{\sigma_v}{\hat{v}}\right)^2\right]$$

que dá uma fórmula mais fácil de memorizar quando se isolam no membro esquerdo as grandezas que envolvem p:

$$\left(\frac{\sigma_p}{\hat{p}}\right)^2 = \left(\frac{\sigma_m}{\widehat{m}}\right)^2 + \left(\frac{\sigma_v}{\hat{v}}\right)^2 \qquad (4.1)$$

Em palavras, o quadrado do desvio-padrão relativo no produto é a soma dos quadrados dos desvios-padrão relativos dos fatores. Quando se usa a mesma unidade para σ_m e m, a razão $\frac{\sigma_m}{\bar{m}}$ não tem unidade, e o mesmo acontece com as outras razões da fórmula, de modo que não há necessidade de compatibilizar as unidades de p, m e v, o que facilita muito seu uso em comparação com a (3.6), em que cada fator de cada parcela terá, provavelmente, uma dimensão diferente, mas é preciso garantir que cada parcela termine com a mesma unidade física (atenção, não basta terem a mesma *dimensão* física, precisam ter a mesma *unidade*), para efetuar a soma.

III.2.b Quociente de duas grandezas

A fim de calcular o desvio-padrão associado à velocidade de um corpo,

$$v = \frac{x_2 - x_1}{t_2 - t_1}$$

primeiro agrupam-se as medidas de posição e tempo em duas grandezas separadas e independentes, $\Delta x = a$ e $\Delta t = b$, e então se aplica a fórmula (3.6), em que se substituem as variâncias de a e b. Assim, o valor estimado da velocidade, $\hat{v}$, é:

$$\hat{v}(\Delta x, \Delta t) = \frac{\Delta x}{\Delta t} = \frac{\hat{a}}{\hat{b}} \quad \xRightarrow{eq.\ (2.6)} \quad \sigma_v^2 = \left(\frac{\partial v}{\partial \hat{a}}\right)^2 \sigma_a^2 + \left(\frac{\partial v}{\partial \hat{b}}\right)^2 \sigma_b^2$$

Calculando as derivadas parciais

$$\frac{\partial v}{\partial \hat{a}} = \left.\frac{\partial v}{\partial a}\right|_{\substack{a=\hat{a}\\ b=\hat{b}}} = \frac{1}{\hat{b}} \quad \text{e} \quad \frac{\partial v}{\partial \hat{b}} = \left.\frac{\partial v}{\partial b}\right|_{\substack{a=\hat{a}\\ b=\hat{b}}} = -\frac{\hat{a}}{\hat{b}^2}$$

em que os valores de $\hat{a}$ e $\hat{b}$ são obtidos diretamente das leituras das posições, e suas variâncias são estimadas do instrumento de medida.

Substituindo essas derivadas na expressão anterior,

$$\sigma_v^2 = \left(\frac{1}{\hat{b}}\right)^2 \sigma_a^2 + \left(-\frac{\hat{a}}{\hat{b}^2}\right)^2 \sigma_b^2 = \frac{\hat{a}^2}{\hat{b}^2}\left(\frac{\sigma_a^2}{\hat{a}^2} + \frac{\sigma_b^2}{\hat{b}^2}\right) = \hat{v}^2\left(\frac{\sigma_a^2}{\hat{a}^2} + \frac{\sigma_b^2}{\hat{b}^2}\right)$$

que pode ser reescrita na mesma forma da equação (4.1), de modo que é possível enunciar uma mesma regra para os desvios padrão do produto e do quociente de duas variáveis: o quadrado do desvio-padrão relativo no produto ou no quociente é a soma dos quadrados dos desvios-padrão relativos das variáveis. A próxima subseção generaliza este resultado para um número qualquer de variáveis, que será o resultado destacado.

III.2.c Produto e quociente de muitas grandezas

Considere a expressão algébrica

$$u = \frac{a^{\alpha} b^{\beta} \dots}{g^{\gamma} h^{\eta} \dots} \qquad (4.2)$$

em que $a, b, \dots, g, h, \dots$ são sujeitas a flutuação estatística e $\alpha, \beta, \dots \gamma, \eta, \dots$ são constantes, conhecidas exatamente – que não precisam ser números inteiros, o resultado abaixo vale para qualquer tipo de constante.

Da aplicação da fórmula (3.6), fatorando u da expressão resultante como nos cálculos dos itens IV.2.a e IV.2.b, obtém-se a regra:

Se $u = \frac{a^{\alpha} b^{\beta} \dots}{g^{\gamma} h^{\eta} \dots}$ com $\alpha, \beta, \dots \gamma, \eta, \dots$ constantes → $\hat{u} = \frac{\hat{a}^{\alpha} \hat{b}^{\beta} \dots}{\hat{g}^{\gamma} \hat{h}^{\eta} \dots}$ e sua variância pode ser deduzida da

variância relativa $\left(\frac{\sigma_u}{\hat{u}}\right)^2 = \left(\alpha \frac{\sigma_a}{\hat{a}}\right)^2 + \left(\beta \frac{\sigma_b}{\hat{b}}\right)^2 + \cdots + \left(\gamma \frac{\sigma_g}{\hat{g}}\right)^2 + \left(\eta \frac{\sigma_h}{\hat{h}}\right)^2 + \cdots$ (4.3)

desde que $\sigma_a \ll |a_0|, \sigma_b \ll |b_0|, \dots, \sigma_g \ll |g_0|, \sigma_h \ll |h_0|, \dots$ e $a, b, \dots, g, h, \dots$ sejam estatisticamente independentes.

Note que a expressão (4.2) podia ter sido escrita como

$$u = a^{\alpha} b^{\beta} \dots g^{-\gamma} h^{-\eta} \dots \qquad (4.2')$$

o que não muda nada na fórmula da variância relativa, em que todos os termos estão ao quadrado, de modo que o sinal menos nos expoentes é irrelevante – quando as condições de validade da fórmula (3.6) são satisfeitas, a contribuição da flutuação estatística de uma das grandezas para o desvio-padrão relativo de u depende apenas do seu desvio-padrão relativo e do módulo do expoente, não importa se a proporcionalidade é direta ou inversa.

V. Aplicações aos Experimentos

Nesta seção estão os resultados da aplicação das fórmulas gerais à análise de alguns dos experimentos do MEXI , especialmente nos casos em que as fórmulas gerais do capítulo precedente não se aplicam e é preciso recorrer aos conceitos gerais para encontrar os procedimentos adequados. A fim de que este capítulo fique autocontido, para que possa ser consultado pontualmente, certos comentários já feitos no texto serão repetidos.

V.1 Desvio-padrão da velocidade e da aceleração

V.1.a Velocidade linear

A velocidade linear de um corpo nos experimentos com imagens é determinada a partir de medidas de posição, x_2 e x_1, e tempo, t_2 e t_1, em duas imagens distintas:

$$v = \frac{x_2 - x_1}{t_2 - t_1}$$

O desvio-padrão da velocidade dependerá, portanto, dos desvios-padrão das posições e dos respectivos instantes. A medida da posição é efetuada com a trena (ou quadriculado), enquanto os tempos são derivados da taxa de filmagem da câmera digital e da posição do quadro na sequência do vídeo. A base de tempo da filmadora é extremamente precisa, de modo que o desvio-padrão no intervalo de tempo pode ser ignorado, e o fator $\frac{1}{t_2 - t_1}$ é constante. Assim, pode-se usar a fórmula (3.1) e calcular

$$\sigma_v = \frac{1}{|t_2 - t_1|}\, \sigma_{x_2 - x_1}$$

Usando a fórmula (3.5) e a discussão da subseção IV.1.a para calcular a variância da diferença, obtém-se

$$\sigma_{x_2 - x_1} = \sqrt{\sigma_{x2}^2 + \sigma_{x1}^2}$$

e, como $\sigma_{x2}^2 = \sigma_{x1}^2 = \sigma_x^2$, obtém-se o resultado final

$$\sigma_v = \frac{\sqrt{2}}{|t_2 - t_1|}\, \sigma_x \qquad (5.1)$$

V.1.b Aceleração linear

Em um experimento em que imagens consecutivas do conjunto estão separadas pelo mesmo intervalo de tempo, o instante em que a imagem i é registrada pode ser descrita pela fórmula

$$t_i = t_0 + i\,\delta t, \quad i = 1..N$$

em que δt é o intervalo de tempo entre quadros consecutivos do conjunto, a aceleração linear no instante de tempo t_i é determinada a partir das medidas da velocidade v nos instantes t_{i+1} e t_{i-1}:

$$a(t_i) = \frac{v_{i+1} - v_{i-1}}{t_{i+1} - t_{i-1}} \qquad (5.2)$$

Aqui, o procedimento de usar a fórmula (3.6) não pode ser adotado diretamente, uma vez que as velocidades v_{i+1} e v_{i-1} tem um componente aleatório em comum:

$$v_{i+1} = \frac{x_{i+2} - x_i}{t_{i+2} - t_i} \quad e \quad v_{i-1} = \frac{x_i - x_{i-2}}{t_i - t_{i-2}}$$

Substituindo essas expressões em (5.2), notando que $t_{i+1} - t_{i-1} = t_{i+2} - t_i = t_i - t_{i-2} = 2\delta t$, obtém-se

$$a(t_i) = \frac{x_{i+2} - 2x_i + x_{i-2}}{(2\ \delta t)^2}$$

Ignorando a incerteza no intervalo de tempo como feito no item anterior, e calculando o desvio-padrão do numerador com a fórmula (3.5), obtém-se

$$\sigma_a = \frac{1}{(2\ \delta t)^2}\sqrt{\sigma_{x_{i+2}}^2 + 4\ \sigma_{x_i}^2 + \sigma_{x_{i-2}}^2}$$

Finalmente, como $\sigma_{x_{i+2}}^2 = \sigma_{x_i}^2 = \sigma_{x_{i-2}}^2 = \sigma_x^2$, obtém-se o resultado final

$$\sigma_a = \frac{\sqrt{6}}{(2\ \delta t)^2}\ \sigma_x \qquad (5.3)$$

em que $\delta t = t_{i+1} - t_i$ é o intervalo de tempo entre imagens **consecutivas** do conjunto analisado.

V.1.c Velocidade e aceleração angular a partir da coordenada angular

A velocidade angular ω é determinada experimentalmente como a razão da variação do ângulo φ entre dois quadros com o intervalo de tempo correspondente, $t_2 - t_1$. Toda a dedução realizada no item V.1.a pode ser repetida, trocando a coordenada linear x pela angular φ; neste caso, também pode-se ignorar a incerteza no tempo. Obtém-se

$$\sigma_\omega = \frac{\sqrt{2}}{|t_2 - t_1|}\sigma_\varphi \qquad (5.4)$$

Procedendo com a aceleração angular da mesma maneira que na subseção V.1.b, o resultado (5.3) vale para o desvio-padrão da aceleração angular, apenas trocando σ_x por σ_φ.

V.2 Energia cinética e energia potencial

No experimento da conservação da energia, a energia cinética do carrinho varia com o tempo. Em certos instantes, ela é bem maior que o desvio-padrão, noutros não, de modo que uma das condições de validade das fórmulas (3.6) e (4.3) nem sempre é satisfeita. Assim, o cálculo do desvio-padrão da energia cinética exige a consideração do seu valor em relação ao desvio-padrão.

V.2.a Quando a energia cinética é bem maior que o desvio-padrão

A energia cinética K de um corpo na mecânica clássica pode ser calculada a partir de sua massa m e de sua velocidade v:

$$K(m, v) = \frac{mv^2}{2} \qquad (5.5)$$

Se o desvio-padrão relativo da massa for comparável ao da velocidade, é preciso considerá-lo, de modo que

$$\sigma_K^2 = K^2\left[\left(\frac{\sigma_m}{m}\right)^2 + \left(2\frac{\sigma_v}{v}\right)^2\right] \Rightarrow \sigma_K = K\sqrt{\left(\frac{\sigma_m}{m}\right)^2 + \left(2\frac{\sigma_v}{v}\right)^2} \qquad (5.6)$$

Normalmente, o desvio-padrão da velocidade é bem maior que o da massa, o que simplifica o resultado para

$$\sigma_K = 2K\frac{\sigma_v}{|v|} \qquad (5.7)$$

Embora essa expressão possa ser simplificada algebricamente substituindo $K = \frac{mv^2}{2}$, não vale a pena, porque a fórmula resultante requer consistência no uso das unidades, enquanto esta última expressão permite o uso da velocidade tanto em cm/s quanto em m/s e da energia, joules ou ergs, sem qualquer preocupação com a unidade da massa.

V.2.b Quando a energia cinética é comparável ao desvio-padrão

A velocidade v provém de uma razão entre variações de posição, Δx, e tempo, $\Delta t.$ Quando $\Delta x{\sim}0$, então $v{\sim}0$ e a expressão (4.2) não pode ser aplicada, por conta do valor nulo no denominador. Examinando essa dificuldade com mais cuidado, verifica-se que a fórmula (3.6), da qual vem a (5.2) como caso particular, realmente não pode ser usada, porque, quando $v{\sim}0$, a condição da validade $\sigma_v \ll v_0$ é violada.

A solução deste problema depende do valor verdadeiro da velocidade, v_0, e está fora do escopo deste texto. A fim de obter uma estimativa, pode-se começar aplicando a fórmula (3.4), do desvio-padrão de uma grandeza que depende de uma única variável, já que neste caso pode-se ignorar a incerteza na massa. Substituindo K da expressão (5.5) na fórmula (3.4), obtém-se

$$\sigma_K = m|v|\,\sigma_v$$

Quando a velocidade é nula ou quase nula, sua medida dará um valor da ordem de grandeza do desvio-padrão. Substituindo-se $|v|$ na equação acima por σ_v, encontra-se a estimativa do quadro abaixo, que será adotada nos experimentos do MEXI.

Para $K = \frac{1}{2}mv^2$ com $\frac{\sigma_m}{m} \ll \frac{\sigma_v}{v}$ e $v \lesssim \sigma_v$ então $\qquad \sigma_K \approx m{\sigma_v}^2 \qquad (5.8)$

IV.2.c Energia potencial elástica

A energia potencial elástica *U* de um corpo na mecânica clássica pode ser calculada a partir da distância à posição de equilíbrio, x_e. Se x é a posição do corpo, $\delta = x - x_e$ e a energia potencial fica

$$U = \frac{k\delta^2}{2} \qquad (5.9)$$

em que k é a constante elástica. Nos experimentos do MEXI, x_0 é conhecido com grande precisão, de modo que o desvio-padrão de δ pode ser bem aproximado pelo de x, conforme seção III.1.b, ou seja

$$\sigma_\delta = \sigma_x$$

Além disso, o desvio-padrão relativo da constante elástica é menor que o da posição, de modo que o problema da estimativa do desvio-padrão de U é análogo ao da energia cinética, trocando m por k e x por v. Assim, as fórmulas serão similares e ficam, substituindo os símbolos correspondentes:

$$\text{Quando} \quad \delta \gg \sigma_x, \ \ \sigma_U = 2U\frac{\sigma_\delta}{\delta} = 2U\frac{\sigma_x}{\delta} \qquad (5.10)$$

$$\text{Quando} \quad \delta \sim \sigma_x, \ \ \sigma_U \approx k\sigma_x^2 \qquad (5.11)$$

V.3 Energia mecânica

A energia mecânica E de um corpo é a soma de suas energias cinética e potencial:

$$E(K,U) = K + U \qquad (5.12)$$

Usando a fórmula (3.5), da variância da soma, obtém-se

$$\sigma_E^2 = \sigma_K^2 + \sigma_U^2 \Rightarrow \sigma_E = \sqrt{\sigma_K^2 + \sigma_U^2} \qquad (5.13)$$

Cabe uma dúvida: a fórmula (3.5) requer que K e U sejam estatisticamente independentes, mas ambos dependem das posições. Porém, a velocidade em um instante é calculada da diferença das posições entre um instante antes e outro depois daquele em que é calculada a energia potencial, assim K e U não dependem dos mesmos dados – não há flutuação estatística comum aos dois termos, assim a independência estatística está garantida e a fórmula (5.13) está correta.

No entanto, quando se observa o gráfico da energia total com o tempo, percebe-se que há algo errado – as barras de incerteza são grandes, em comparação com a flutuação estatística. O que acontece é que os dados que entram no cálculo da energia em um instante entram também nas energias em instantes vizinhos, havendo, portanto, uma flutuação em comum. O cálculo dessa parte comum da variância – a covariância – está fora de escopo deste texto.

Este resultado é usado no experimento de conservação de energia.

VI. Curva de tendência

Em diversos experimentos, ajustam-se os parâmetros de uma reta ou parábola aos dados experimentais. Essa operação pode ser efetuada com uma planilha eletrônica, que dá o nome *curva de tendência* à função calculada com os parâmetros ajustados. O procedimento baseia-se no Método dos Mínimos Quadrados (MMQ), descrito em geral nas seções VI.2 e VI.3, e especializado para o ajuste dos parâmetros de uma reta nas seções seguintes.

VI.1 Definição do problema

Considere um experimento em que se observa uma grandeza física y que se relaciona com a variável x por uma função linear nos parâmetros a_0 e b_0:

$$y = a_0 f(x) + b_0 g(x) \qquad (6.1)$$

em que $f(x)$ e $g(x)$ são funções apenas de x, e a_0 e b_0 são constantes (isto é, independentes de x), cujos valores são relacionados ao sistema específico em estudo***. O propósito do experimento é encontrar estimativas $\hat{a}$ e $\hat{b}$ para esses parâmetros. O Método dos Mínimos Quadrados (MMQ) considera que a variável preditiva x é conhecida sem erro, e os desvios padrão da resposta y são todos iguais ou as razões entre eles são conhecidos. O texto abaixo considera que esses requisitos são satisfeitos pelos dados experimentais.

Quando se observam apenas dois dados, (x_1, y_1, σ_1) e (x_2, y_2, σ_2), em que σ_1 e σ_2 são os desvios padrão de y_1 e y_2, respectivamente, é possível resolver o sistema de duas equações a duas incógnitas que se obtém ao aplicar (6.1) a esse par de dados. A solução desse sistema fornece as estimativas $\hat{a}$ e $\hat{b}$, cujos desvios-padrão devem ser calculados a partir de σ_1 e σ_2 pelas fórmulas de propagação adequadas, conforme seção IV. Essa maneira, muitas vezes, não alcança a precisão necessária quando as incertezas de medida são significativas e não podem ser diminuídas. Além disso, esse procedimento não permite de maneira nenhuma testar se a função (6.1) é adequada para a descrição do fenômeno.

A prática em Física consiste em determinar N valores da grandeza y para diferentes valores de x, ou seja, determinar um conjunto de dados

$$\{(x_i, y_i, \sigma_i), i = 1, \ldots, N\}$$

em que o índice i identifica cada dado do conjunto de N pontos experimentais, e σ_i é o desvio padrão de y_i. O interesse em repetir a medição de y um número maior de vezes decorre da diminuição dos desvios-padrões de $\hat{a}$ e $\hat{b}$ por conta da redução da flutuação estatística quando essas estimativas são calculadas a partir de mais medições, da mesma forma que o desvio-padrão da média diminui com o número de dados, $\sigma_m = {}^{\sigma}/_{\sqrt{N}}$. A fim de testar se a relação (6.1) é válida, procura-se escolher valores de x distribuídos por toda a região de interesse dessa variável explicativa. Por exemplo, na queda livre de um corpo, procura-se observar sua posição desde o início até o final do movimento e não concentrar as medições nos primeiros e/ou nos últimos instantes do movimento.

*** É muito comum chamar x de variável explicativa ou preditiva e y de resposta. Usaremos esses nomes no texto, bem como os nomes variável independente para x e dependente para y.

Uma abordagem mais formal do MMQ, que inclui a demonstração das suas propriedades ótimas, pode ser encontrada em (Draper & Smith, 1998), e uma apresentação que inclui diversas aplicações específicas em (Helene, 2006).

VI.2. O Método dos Mínimos Quadrados

Para explicar o método, relaciona-se cada valor medido y_i com a função (6.1) e o correspondente erro experimental ϵ_i de maneira análoga à da eq. (2.6) na seção II.6:

$$y = a_0 f(x) + b_0 g(x) + \epsilon_i \qquad (6.2)$$

Como no capítulo II, esse erro, embora desconhecido, tem as seguintes propriedades:

i. $\langle \epsilon_i \rangle = 0$, que corresponde à hipótese de que as medições não são tendenciosas – os erros de leitura dos instrumentos para mais tendem a compensar os erros para menos.
ii. $\langle \epsilon_i^2 \rangle = \sigma_i^2$, que é a medida da dispersão dos dados experimentais e pode ser estimada com os métodos e fórmulas já descritos.

O procedimento consiste em minimizar a soma dos quadrados dos resíduos ponderados pelo inverso das variâncias dos dados, mais exatamente, a função

$$Q(a,b) = \sum_{i=1}^{N} \frac{\left(y_i - \left(a\, f(x_i) + b\, g(x_i)\right)\right)^2}{\sigma_i^2} \qquad (6.3)$$

Note que, nessa expressão, a e b são variáveis, o que é estranho, mas necessário, uma vez que a natureza não nos permite conhecer os seus valores verdadeiros, a_0 e b_0. Os valores $\hat{a}$ e $\hat{b}$, que minimizam $Q(a,b)$, são as estimativas dos parâmetros, e podem assumir quaisquer valores.

Como $Q(a,b)$ é um polinômio do 2º grau em a e b, pode-se desenvolver o polinômio que corresponde a essa parábola bidimensional e encontrar seu mínimo usando técnicas de álgebra linear (veja, por exemplo, o livro de Barone [1]). A maneira que vamos indicar aqui é calcular as derivadas parciais de $Q(a,b)$ em relação a a e b, que devem se anular para o par de valores $\hat{a}$ e $\hat{b}$ que minimizam a função e são as estimativas das grandezas físicas, ou seja,

$$\left.\frac{\partial Q}{\partial a}\right|_{\hat{a},\hat{b}} = 0$$

$$\left.\frac{\partial Q}{\partial b}\right|_{\hat{a},\hat{b}} = 0$$

As duas equações do sistema linear acima (a derivada de um polinômio do 2º grau é um polinômio do 1º grau) em função das duas incógnitas $\hat{a}$ e $\hat{b}$ são escritas de preferência em forma de matriz,

$$\begin{pmatrix} \sum \frac{y_i\, f(x_i)}{\sigma_i^2} \\ \sum \frac{y_i\, g(x_i)}{\sigma_i^2} \end{pmatrix} = \begin{pmatrix} \sum \frac{\left(f(x_i)\right)^2}{\sigma_i^2} & \sum \frac{f(x_i)\, g(x_i)}{\sigma_i^2} \\ \sum \frac{f(x_i)\, g(x_i)}{\sigma_i^2} & \sum \frac{\left(g(x_i)\right)^2}{\sigma_i^2} \end{pmatrix} \cdot \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \qquad (6.4)$$

em que todas as somatórias se estendem desde $i = 1$ até $i = N$. Essa fórmula matricial pode ser escrita em forma compacta como

$$\vec{D} = \mathbf{M} \cdot \vec{A} \qquad (6.5)$$

em que $\vec{D}$ é o vetor coluna do membro esquerdo de (6.4), cujos elementos são combinações lineares dos dados, $\mathbf{M}$ é a matriz quadrada, que depende somente dos valores da variável preditiva e dos desvios padrões, e $\vec{A}$ é o vetor das estimativas. A solução do sistema linear pode ser obtida por qualquer método. Por exemplo, definindo a matriz $\mathbf{V}$ como a inversa da matriz $\mathbf{M}$

$$\mathbf{V} = \mathbf{M}^{-\mathbf{1}} \qquad (6.6)$$

a solução de (6.4) é

$$\vec{A} = \mathbf{V} \cdot \vec{D} \qquad (6.7)$$

ou seja, os parâmetros ajustados são combinações lineares dos dados y_i. Os desvios-padrões das estimativas $\hat{a}$ e $\hat{b}$ podem ser calculados usando as fórmulas do capítulo III, verificando-se que suas variâncias são os elementos da diagonal da matriz $\mathbf{V}$ (Draper & Smith, 1998) (Helene & Vanin, 1991) (Helene, 2006), que, por isso, é chamada matriz das variâncias:

$$\sigma_{\hat{a}} = \sqrt{v_{11}} \quad \text{e} \quad \sigma_{\hat{b}} = \sqrt{v_{22}} \qquad (6.8)$$

É uma questão de álgebra estender as equações (6.2) a (6.8) para funções lineares com um número qualquer de parâmetros. Assim, o quadro abaixo resume as fórmulas para o caso geral, com uma notação que facilita seu uso.

Dada a função $y(x)$, linear em μ parâmetros α_ν,

$$y(x) = \sum_{\nu=1}^{\mu} \alpha_\nu h_\nu(x) = \alpha_1 h_1(x) + \alpha_2 h_2(x) + \cdots + \alpha_\mu h_\mu(x) \qquad (6.1')$$

as estimativas $\hat{\alpha}_\nu$ de mínimos quadrados são obtidas pela solução do sistema de equações

$$\vec{D} = \mathbf{M} \cdot \vec{\hat{\alpha}} \qquad (6.5')$$

(por exemplo, $\vec{\hat{\alpha}} = (\mathbf{M})^{-1}\, \vec{D}$) em que os elementos dos vetores e matrizes são

$$D_\nu = \sum \frac{y_i\, h_\nu(x_i)}{\sigma_i^2} \qquad (6.9)$$

e

$$M_{\nu\kappa} = \sum \frac{h_\nu(x_i)\, h_\kappa(x_i)}{\sigma_i^2} \qquad (6.10)$$

Os desvios-padrões são os elementos da diagonal da matriz $\mathbf{V} = (\mathbf{M})^{-\mathbf{1}}$.

VI.3 Origem e justificativa do Método dos Mínimos Quadrados

O método dos mínimos quadrados é devido a Gauss e Legendre (final do século XVIII e início do XIX), que o aplicaram na redução de dados de observações astronômicas; a prioridade de descoberta do método é uma questão interessante. Stigler (Stigler, 1981) sugere que Gauss descobriu o método antes de Legendre, que, porém, o apresentou em uma publicação que despertou o interesse dos astrônomos da época, o que Gauss teria tentado vários anos antes, mas não conseguido. Já Grosser (Grosser, 1962) relata que o método teria chamado a atenção quando da descoberta do asteroide Ceres, que ocorreu da maneira descrita no parágrafo seguinte, de acordo com esse livro e também com a referência (Forbes, 1971).

Após a descoberta de Urano, acreditava-se que deveria haver um planeta entre Marte e Júpiter. No início de 1801, um possível planeta nessa região foi avistado antes de esconder-se sob o Sol. Nenhuma das órbitas determinadas para esse planeta, com base nessas observações, permitiu reencontrá-lo no céu. Gauss interessou-se pelo problema e, em dois meses, desenvolveu o MMQ, com que previu a trajetória do objeto, que foi encontrado imediatamente, no final do ano. Esse objeto era Ceres, o maior dos planetas menores internos a Netuno, com uma órbita bastante excêntrica e que pertence ao cinturão de asteroides. Embora seja difícil ter certeza da razão pela qual os astrônomos erravam a previsão, é de se imaginar que escolhessem das observações aquelas que davam órbitas de acordo com suas expectativas. Esse episódio ilustra, então, a objetividade dos métodos estatísticos.

As propriedades que justificam o uso do MMQ decorrem da estrutura da fórmula (6.7) – cada parâmetro é estimado por uma combinação linear dos dados muito particular. Demonstra-se que os parâmetros estimados são não-tendenciosos e que têm as menores variâncias que podem ser obtidas por **quaisquer** combinações lineares dos dados, propriedades que *não* podem ser inferidas da forma da função $Q(a,b)$, eq. (6.3). No entanto, como $Q(a,b)$ é uma parábola, ela tem um único mínimo, o que indica a unicidade da estimativa. Essa também é uma propriedade importante, que contribui para a objetividade do método, uma vez que a estimativa não dependerá de quem realize a análise, mas apenas e tão somente dos dados obtidos e de suas propriedades estatísticas.

Finalmente, apenas para completar a discussão, quando os dados têm distribuição estatística de acordo com o modelo normal, pode-se mostrar que a estimativa da fórmula (6.7) é a de menor variância em termos absolutos – não há nenhuma combinação dos dados, mesmo não-linear, que permita a obtenção de estimativas de menor variância que as da fórmula (6.8).

VI.4 Interpretação do valor mínimo da soma dos quadrados dos resíduos

VI.4.i Propriedades estatísticas do valor mínimo da soma dos quadrados

O valor da função Q da equação (6.3) calculada com as estimativas $\hat{a}$ e $\hat{b}$ dos parâmetros, $Q(\hat{a},\hat{b})$, flutua estatisticamente, significando que, a cada repetição da medida, um valor diferente é obtido, mas tem um valor médio, que se relaciona com o número de dados.

Se a_0 e b_0 são os valores exatos dos parâmetros, então

$$Q(a_0,b_0)=\sum_{i=1}^{N}\frac{\left(y_i-\left(a_0f(x_i)+b_0g(x_i)\right)^2\right)^2}{\sigma_i^2}=\sum_{i=1}^{N}\frac{{\epsilon_i}^2}{\sigma_i^2}$$

em que se usou a relação (6.2) para identificar os erros ϵ_i. Ao calcular o valor médio de cada um dos membros da equação, obtém-se

$$\langle Q(a_0, b_0)\rangle = \sum_{i=1}^{N} \frac{\langle {\epsilon_i}^2\rangle}{\sigma_i^2} = \sum_{i=1}^{N} \frac{\sigma_i^2}{\sigma_i^2} = N \qquad (6.9)$$

Essa relação não é muito útil, uma vez que a_0 e b_0 são desconhecidos. No entanto, quando se substituem as estimativas conhecidas $\hat{a}$ e $\hat{b}$ no lugar de a_0 e b_0, obtém-se

$$Q(\hat{a}, \hat{b}) = \sum_{i=1}^{N} \frac{\left(y_i - \left(\hat{a}f(x_i) + \hat{b}g(x_i)\right)\right)^2}{\sigma_i^2} = \sum_{i=1}^{N} \frac{{\epsilon'_i}^2}{\sigma_i^2} \qquad (6.10)$$

em que ϵ'_i é o resíduo do ajuste no ponto x_i,

$$\epsilon'_i = y_i - \left(\hat{a}f(x_i) + \hat{b}g(x_i)\right) \qquad (6.11)$$

É possível calcular o valor médio da expressão (6.10) (Draper & Smith, 1998), obtendo-se uma expressão bastante parecida com a equação (6.9):

$$\langle Q(\hat{a}, \hat{b})\rangle = N - 2 \qquad (6.12)$$

Como Q é uma função definida positiva ou nula, o fato de seu valor médio ser nulo para $N = 2$ significa que Q vale zero sempre. Isso porque, nesse caso de $N = 2$, as estimativas $\hat{a}$ e $\hat{b}$ são a solução do sistema linear, de modo que os resíduos ϵ'_1 e ϵ'_2 para os dois dados são nulos e, portanto, sua soma quadrática é nula.

O resultado (6.12) pode ser generalizado para ajuste de μ parâmetros $\alpha_\nu, \nu = 1..\mu$

$$\langle Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu)\rangle = N - \mu \qquad (6.12')$$

Quando se faz *um* ajuste, porém, obtém-se *um* único valor $Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu)$, de modo que, se $N > \mu$, Q pode assumir qualquer valor no intervalo $[0, \infty)$. Quando os dados seguem o modelo normal, as probabilidades de ocorrência dos diferentes valores pode ser calculada e demonstra-se que a grandeza $Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu)$ flutua estatisticamente como a variável $\chi^2_{N-\mu}$ (diz-se qui-quadrado com $N - \mu$ graus de liberdade). A maneira exata dessa flutuação depende muito de $N - \mu$, mas seu desvio-padrão é $\sqrt{2(N-\mu)}$ e, para valores grandes de $N - \mu$, tem a forma da gaussiana, de modo que as probabilidades com que um valor individual Q se distribui em torno de $N - \mu$ são aproximadamente aquelas dos quadros da seção II.4. Esses resultados estão resumidos no quadro seguinte, e serão usados no resto desta seção.

Até aqui, os desvios padrão foram supostos conhecidos, mas os resultados apresentados não dependem desse fato, de modo que eles podem ser usados de duas maneiras muito distintas. A primeira é usar a relação (6.12') para determinar as variâncias dos dados quando elas são desconhecidas a menos de um fator, que está discutido na subseção VI.4.ii. Outra, que depende de conhecer as variâncias dos dados, é usar a informação sobre as propriedades de flutuação estatística de $Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu)$ para validar a função modelo usada, discutido na subseção VI.4.iii.

Considere a função $y(x)$, cujos μ parâmetros α_ν são ajustados pelo MMQ

$$y(x) = \sum_{\nu=1}^{\mu} \alpha_\nu h_\nu(x) = \alpha_1 h_1(x) + \alpha_2 h_2(x) + \cdots + \alpha_\mu h_\mu(x) \qquad (6.1')$$

A soma dos quadrados dos resíduos ponderados, usando os parâmetros ajustados $\hat{\alpha}_\nu$,

$$Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu) = \sum_{i=1}^{N} \frac{\left(y_i - \sum_{\nu=1}^{\mu} \hat{\alpha}_\nu h_\nu(x_i)\right)^2}{\sigma_i^2}$$

tem valor médio

$$\langle Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu)\rangle = N - \mu \qquad (6.12')$$

com $\mathrm{var}\left(Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu)\right) = 2(N-\mu)$, e $Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu)$ segue o modelo normal quando $N - \mu$ é grande.

VI.4.ii Estimativa do desvio-padrão dos dados

A relação (6.12') permite determinar um único valor desconhecido. Assim, esta aplicação somente pode ser efetuada em dois casos:

a. $\sigma_i = \sigma$, todo i
b. $\sigma_i = \gamma_i\, \sigma$ em que todos os $\gamma_i, i = 1..N$ são conhecidos

O caso a é muito comum, mas o b abrange o caso a fazendo $\gamma_i = 1$ para todo i, assim começamos por ele. O valor mínimo de Q, substituindo os desvios padrão do caso b, é

$$Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu) = \frac{1}{\sigma^2}\sum_{i=1}^{N} \frac{\left(y_i - \sum_{\nu=1}^{\mu} \hat{\alpha}_\nu h_\nu(x_i)\right)^2}{\gamma_i^2}$$

Substituindo $Q(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_\mu)$ pelo seu valor médio da relação (6.12') e isolando σ^2, obtém-se

$$\hat{\sigma}^2 = \frac{1}{N-\mu}\sum_{i=1}^{N} \frac{\left(y_i - \sum_{\nu=1}^{\mu} \hat{\alpha}_\nu h_\nu(x_i)\right)^2}{\gamma_i^2}$$

que permite estimar todas as variâncias, uma vez que os valores γ_i são conhecidos.

Trocando γ_i por 1, obtém-se a estimativa da variância dos dados quando eles são todos iguais.

Quando as variâncias dos dados são todas iguais a um valor desconhecido, esse valor pode ser estimado a partir do ajuste dos parâmetros da função modelo pelo MMQ como

$$\hat{\sigma}^2 = \frac{1}{N-\mu}\sum_{i=1}^{N}\left(y_i - \sum_{\nu=1}^{\mu} \hat{\alpha}_\nu h_\nu(x_i)\right)^2 \qquad (6.13)$$

com a hipótese que a função $y(x) = \sum_{\nu=1}^{\mu} \alpha_\nu h_\nu(x)$ representa corretamente a dependência da resposta y na variável explicativa do fenômeno em estudo, x.

VI.4.iii. Uso do valor mínimo da soma dos quadrados para testar o ajuste

Quando os desvios padrões dos dados são conhecidos, as relações desenvolvidas na seção VI.4.i podem ser usadas de forma a realizar um teste de hipótese rigoroso, dentro do quadro teórico da teoria estatística (Helene & Vanin, 1991). Aqui, vamos nos limitar aos aspectos qualitativos desse teste.

Valores de $Q(\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_\mu)$ muito menores que $N - \mu$ são devidos a desvios-padrões superestimados. Já valores muito maiores que $N - \mu$ sugerem que o modelo seja inadequado, quando se deve buscar outra função para relacionar y com x no lugar da (6.1'), ou, então, os desvios-padrões estão subestimados.

A dificuldade está em definir o que significa "muito menor" ou "muito maior". Como indicado no quadro da seção VI.4.i, $\langle Q(\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_\mu)\rangle = N - \mu$ e o desvio padrão de $Q(\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_\mu)$ é $\sqrt{2(N-\mu)}$, com uma flutuação estatística que se assemelha à do modelo normal quando há muitos dados, o que permite usar as probabilidades dos quadros da seção II.4:

> No ajuste de μ parâmetros a uma medida com N dados, o valor $\hat{Q} = Q(\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_\mu)$ flutua estatisticamente em torno do valor $\nu = N - \mu$, com desvio-padrão $\sqrt{2\nu}$. As probabilidades dos intervalos de um, dois e três desvios-padrões são, aproximadamente:
>
> $$P\left(\nu - \sqrt{2\nu} < \hat{Q} < \nu + \sqrt{2\nu}\right) \sim 68\% \quad (\nu > 5)$$
>
> $$P\left(\nu - 2\sqrt{2\nu} < \hat{Q} < \nu + 2\sqrt{2\nu}\right) \sim 95\% \quad (\nu > 10)$$
>
> $$P\left(\nu - 3\sqrt{2\nu} < \hat{Q} < \nu + 3\sqrt{2\nu}\right) \gtrsim 99\% \quad (\nu > 20)$$
>
> em que a faixa de valores em que vale a aproximação normal está indicada em cada caso.

Assim, diferenças menores que um ou dois desvios-padrão são muito prováveis e não sugerem enganos na função modelo nem nos desvios padrão; diferenças entre dois e três desvios padrão acontecem de vez em quando (em média, uma vez a cada 20 repetições do procedimento) e maior que três, raramente. Portanto, um valor de $Q(\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_\mu)$ diferente da média mais que três desvios-padrão é motivo para suspeitar de enganos na função modelo, ou nos desvios-padrão dos y_i, ou nesses dois elementos da análise dos dados.

Como resultado deste teste, pode acontecer que se conclua que os desvios-padrão dos dados estão, provavelmente, subestimados ou superestimados. Quando isso acontece os desvios padrões dos parâmetros ajustados também estarão subestimados ou superestimados, uma vez que as expressões algébricas da equação (6.8), com que se calculam as variâncias dos parâmetros, dependem diretamente das variâncias dos dados.

> No MMQ, as estimativas das variâncias da fórmula (6.8) somente são válidas quando $\left|Q(\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_\mu) - (N-\mu)\right| < 3\sqrt{2(N-\mu)}$.

VI.5 Estimativa dos parâmetros de uma reta

Considere a função afim que relaciona uma variável preditiva (ou independente) x à variável resposta (ou dependente) y

$$y = mx + b$$

em que m e b são constantes reais.

A medida da resposta para um conjunto de N valores x_i da variável preditiva,

$$\{(x_i, y_i, \sigma_i), i = 1, \ldots, N\}$$

em que o índice i identifica cada dado do conjunto e σ_i é o desvio padrão de y_i, permite determinar estimativas $\widehat{m}$ e $\widehat{b}$ pelo MMQ, usando a fórmula (6.4), em que se trocam

$$f(x_i) \rightarrow x_i, \quad g(x_i) \rightarrow 1.$$

No que segue nesta seção, também usamos que as variâncias de todos os dados são iguais, o que corresponde a diversos experimentos do MEXI:

$$\sigma_i = \sigma \quad , \text{todo } i \qquad (6.14)$$

Substituindo a fórmula (6.14) na (6.4), pode-se fatorar ${\sigma_i}^2 = \sigma^2$ nos denominadores das matrizes, que aparecem nos dois lados da equação e podem ser cancelados, obtendo-se

$$\begin{pmatrix} \sum y_i\, x_i \\ \sum y_i \end{pmatrix} = \begin{pmatrix} \sum {x_i}^2 & \sum x_i \\ \sum x_i & N \end{pmatrix} \cdot \begin{pmatrix} \widehat{m} \\ \widehat{b} \end{pmatrix} \qquad (6.15)$$

em que todos os somatórios vão de $i = 1$ até N. Então

$$\widehat{m} = \frac{N \sum y_i\, x_i - (\sum y_i)(\sum x_i)}{N \sum {x_i}^2 - (\sum x_i)^2} \qquad (6.16)$$

$$\widehat{b} = \frac{(\sum y_i)(\sum {x_i}^2) - (\sum y_i\, x_i)(\sum x_i)}{N \sum {x_i}^2 - (\sum x_i)^2} \qquad (6.17)$$

As variâncias podem ser calculadas com a fórmula (3.6), usando que $\sigma_i = \sigma$, todo i. Os resultados são:

$$\sigma_m^2 = \sigma^2 \, \frac{N}{N \sum {x_i}^2 - (\sum x_i)^2} \qquad (6.18)$$

$$\sigma_b^2 = \sigma^2 \, \frac{\sum {x_i}^2}{N \sum {x_i}^2 - (\sum x_i)^2} \qquad (6.19)$$

Os mesmos resultados são obtidos com a fórmula (6.8), em que é preciso lembrar de incluir no cálculo da matriz **M** as variâncias dos dados, que não aparecem na equação (6.15), onde cancelaram com o mesmo fator comum a todos elementos do vetor $\vec{D}$, que aparece do lado esquerdo dessa relação.

A seção seguinte especializa o resultado (6.18) para o caso em que se observa a variável independente a intervalos igualmente espaçados, comum nos experimentos do MEXI.

VI.6 O desvio padrão da inclinação de uma reta

Ao usar o MMQ para ajustar m e b da função $y = mx + b$ a um conjunto de dados $\{(x_i, y_i, \sigma_i), i = 1..N\}$ em que o desvio padrão da variável resposta y_i é σ_i, a variância de m pode ser calculada de acordo com a fórmula (6.8), obtendo-se

$$\sigma_m^2 = \frac{\sum \frac{1}{\sigma_\mathrm{i}^2}}{\sum \frac{1}{\sigma_\mathrm{i}^2} - \left(\sum \frac{x_i}{\sigma_\mathrm{i}^2}\right)^2}$$

Quando todos $\sigma_i = \sigma$, essa expressão se reduz à (6.18), que pode ser escrita de maneira mais simples quando se define uma coordenada central x_c como

$$x_c = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad (6.20)$$

e coordenadas relativas δ_i

$$\delta_i = x_i - x_c \qquad (6.21)$$

O denominador da eq. (6.18), depois de substituir ambas as definições, reduz-se a

$$N \sum_{i=1}^{N} (\delta_i + x_c)^2 - N^2 {x_c}^2 = N \sum_{i=1}^{N} {\delta_i}^2 + 2N x_c \sum_{i=1}^{N} \delta_i = N \sum_{i=1}^{N} {\delta_i}^2$$

A primeira identidade vem do cancelamento do último termo do membro esquerdo com o último termo na expansão da soma dos quadrados de $\delta_i + x_c$, e a segunda, do fato de que $\sum_{i=1}^{N} \delta_i = 0$. Usando esse resultado no denominador da eq. (6.18), após alguma álgebra se obtém

$$\sigma_m^2 = \frac{\sigma^2}{\sum_{i=1}^{N} {\delta_i}^2} \qquad (6.22)$$

Este resultado mostra que a incerteza em m depende apenas da *incerteza* dos dados y (não em seus valores) e da dispersão dos valores de x escolhido para realizar a medida.

Uma expressão ainda mais interessante pode ser obtida sempre que a variável x é amostrada uniformemente, usando um intervalo constante Δx entre medições adjacentes, condição comum a muitos dos experimentos do MEXI. Para simplicidade algébrica, escolhemos N ímpar, então podemos expressar esse conjunto de dados por

$$x_i = x_c + i\,\Delta x \qquad (6.23)$$

em que i é um número inteiro no intervalo $-\frac{N-1}{2} \leq i \leq \frac{N-1}{2}$. Para facilitar as manipulações algébricas, definimos um número inteiro $\nu = \frac{N-1}{2}$ que entra apenas nos cálculos intermediários. Usando a relação (6.23) para calcular o denominador da eq. (6.22), segue-se

$$\sum_{i=1}^{N} {\delta_i}^2 = \Delta x^2 \sum_{i=-\nu}^{\nu} i^2 = \Delta x^2 \frac{1}{3} \nu(\nu+1)(2\nu+1) = \Delta x^2 \frac{1}{12}(N^3 - N)$$

Para N suficientemente grande, o último N entre parênteses pode ser ignorado. Substituindo o denominador da eq. (6.22) pela expressão resultante, obtém-se

$$\sigma_m^2 \cong \frac{12\,\sigma^2}{N\,(N\,\Delta x)^2} \qquad (6.24)$$

Embora a expressão (6.24) resolva o problema, é interessante destacar o papel da escolha do intervalo de medição na incerteza em m. Se x_o é o menor valor observado de x, o maior valor é

$$x_f = (N-1)\Delta x + x_o \quad \Leftrightarrow \quad x_f - x_o = (N-1)\Delta x \approx N\,\Delta x$$

uma boa aproximação quando N é grande, o que frequentemente é o caso. Substituindo este resultado na fórmula (6.24), obtém-se

$$\sigma_m^2 \cong \frac{12\,\sigma^2}{N\left(x_f - x_o\right)^2}$$

O desvio padrão da inclinação de uma reta, obtida pelo MMQ por ajuste a N dados igualmente espaçados no intervalo $\left[x_0, x_f\right]$ da variável preditiva x, vale

$$\sigma_m \approx \frac{\sigma}{x_f - x_0}\sqrt{\frac{12}{N}}$$

em que σ é o desvio-padrão de cada valor da resposta y_i.

Esta expressão mostra claramente que a incerteza na inclinação m depende apenas do intervalo de valores em que se observa x, a incerteza nos dados e o número de pontos observados, portanto não depende da reta com os parâmetros ajustados ter diferenças pequenas ou grandes em cada ponto experimental, ou seja, de parecem bem ou mal ajustado. A avaliação da qualidade do ajuste é independente dos desvios-padrão e deve ser feita conforme a seção VI.4.iii.

VII Bibliografia

Barone, M. (2002). *Álgebra Linear* (3a ed.). São Paulo: IME-USP.

Bevington, P., Robinson, R., & Keith, D. (1992). *Data Reduction and Error Analysis for the Physical Sciences* (2nd ed.). New York: Mc. Graw Hill.

Bureau Internacional de Pesos e Medidas - BIPM. (2008). *Guide to the Expression of Uncertainty in Measurement (GUM).* Paris: BIPM. Retrieved 12 31, 2023, from https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf

Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis* (3rd ed.). Hoboken, New Jerswy: Wiley.

Forbes, E. G. (1971). Gauss and the Discovery of Ceres. *Journal for the History of Astronomy, 2*, pp. 195-199.

Grosser, M. (1962). *The Discovery of Neptune.* Cambridge, Massachusetts (EUA): Harvard University Press.

Helene, O. (2006). *Método dos Mínimos Quadrados com Formalismo Matricial.* São Paulo: Livratia da Física.

Helene, O. A., & Vanin, V. R. (1991). *Tratamento Estatístico de Dados em Física Experimental* (2a ed.). São Paulo: Edgard Blucher.

Stigler, S. M. (1981). Gauss and the Invention of Least Squares. *Annals of Statistics, 9*, pp. 465-474. doi:10.1214/aos/1176345451

Vuolo, J. H. (1996). *Fundamentos da Teoria de Erros* (2a ed.). São Paulo: Edgard Blücher.

Índice remissivo