Apostila de Tratamento de Dados

prof. R.Linares

17 de janeiro de 2013

Instituto de Física - Universidade Federal Fluminense

Resumo

A presente apostila visa servir como material de referência rápida nas análises estatísticas de dados experimentais dos cursos de Física Experimental 1-4. Muitas das informações contidas foram sintetizadas das referências indicadas ao final. É fortemente recomendado que os alunos busquem aprofundar os conceitos nestas referências. Correções e sugestões são bem-vinda para o email: **rlinares@if.uff.br.**

Sumário

1 A medida experimental

1.1 Inevitabilidade dos erros

Ao realizar uma medida experimental o resultado, expresso na forma do valor de uma determinada grandeza, é utilizado como critério de validação ou de contradição de um modelo teórico. No entanto, a medida está **sempre sujeita a erros** do próprio processo de medida. Não é necessário irmos muito longe para compreender a inevitabilidade de erros em um processo de medida. Vejamos, por exemplo, a determinação da altura de uma pessoa. Como você determinaria a altura de alguém? Uma idéia seria a de utilizar uma trena e medir diretamente a altura dessa pessoa em pé. Note que alguns problemas irão surgir como, por exemplo, em garantir que essa pessoa esteja perfeitamente ereta, afinal a altura medida irá depender se esta pessoa estiver ou não um pouco curvada. Para contornar isso, algum aluno esperto poderia propor que, ao invés de medir diretamente a altura, primeiro coloquemos essa pessoa encostada em uma parede plana e perpendicular ao piso e marquemos a posição do topo da cabeça com um lápis. Em seguida, mede-se a altura como a distância entre o piso e a marca de lápis na parede. Mas veja que podemos questionar o conceito da altura de uma pessoa: consideramos ou não os cabelos? Além disso, será que o peso da própria pessoa não iria comprimir ligeiramente a pele da sola dos pés de tal modo que a altura medida dessa forma seja um pouco menor da altura real? Este exemplo serve para ilustrar como uma medida simples é rica em detalhes que não são importantes em nosso cotidiano mas que em experimentos científicos são fundamentais. E por que não nos preocupamos com essas considerações em nosso dia-a-dia? Justamente por que a contribuição de cada uma dessas incertezas (se a pessoa esta reta, se consideramos os cabelos e a compressão da sola do pé) são pequenas e geralmente estamos satisfeitos em saber a altura com precisão de centímetros. Se você tem 1,80 m. ou

1,81 m. imagino que isso não represente um impacto em sua vida!

O exemplo acima é apenas um dos muitos exemplos que poderiam ser apresentados para convencer sobre a presença de erros. Analisemos, com olhar científico, como poderíamos determinar o período de 1 oscilação de um pêndulo simples. Um procedimento tradicional consiste em utilizar um cronômetro com o qual damos o *start* quando o pêndulo estiver em sua altura máxima e o *stop* quando ela voltar a essa mesma posição. Mas pensando por um instante, podemos garantir que o *stop* foi dado exatamente na mesma posição? E será que não existe o tempo de reação do observador entre observar o pêndulo e pressionar o *start* (e *stop*)? A vida experimental parece árdua já que temos sempre que conviver com a **inevitabilidade de erros durante o processo de medida.**

1.2 Valor verdadeiro e sua estimativa

Um aspecto importante que está subentendido é a existência de um valor verdadeiro x_o para a grandeza mensurada. A altura de uma pessoa e o período de um pêndulo simples são exemplos. Uma bolinha de alumínio, o mais esférico possível, também deve ter um único valor para a grandeza diâmetro. Enquanto isso, uma bolinha de papel amassado deve apresentar diâmetros diferentes dependendo da orientação do objeto. Por valor verdadeiro não deve ser compreendido como o "valor correto" da grandeza, a qual todos devem se esforçar para obté-lo. Lembre que é inevitável a presença de erros em uma medida. Por exemplo, ao medir a largura de uma folha de papel A4 com uma régua escolar, o valor obtido não é exatamente x_o mas sim x_1, que em si contém as distintas contribuições ao erro da medida. Uma 2ª medida do mesmo papel pode fornecer um outro valor x_2 diferente de x_1. Podemos escrever que

$$x_o = x_i + \epsilon_i$$

onde ϵ_i é o erro associado a i-ésima medida.

Se conhecessemos exatamente o erro cometido em cada medida poderíamos determinar o valor verdadeiro x_o da grandeza. Porém o melhor que podemos fazer é estimar o melhor valor para x_o. **O objetivo em uma medida experimetal é correlacionar o valor experimental com o valor verdadeiro da grandeza.** Essa correlação é feita com a determinação de uma região dentro da qual acredita-se que o valor verdadeiro esteja contido com um certo grau de confiança. É essa região que será definida como **intervalo de confiança.**

Para normalizar a linguagem, vamos diferenciar dado, medida e resultado de uma medida:

- **dado:** é o valor de uma observação
- **medida:** é o conjunto de dados
- **resultado:** valor que melhor representa um valor verdadeiro

Após uma medida devemos ser capazes de obter uma estimativa para o valor verdadeiro da grandeza. Mas apenas a estimativa não é suficiente. Uma questão a ser respondida é "*A estimativa do valor verdadeiro é boa?*". Não se pode afirmar que a estimativa obtida é exatamente o valor verdadeiro da grandeza e, portanto, precisamos estimar, agora, o quão longe o nosso resultado pode estar do valor verdadeiro. No final, devemos ser capazes de expressar o resultado como:

$$\text{Resultado} = x \pm \sigma$$

onde x é a melhor estimativa para o valor verdadeiro e σ a incerteza.

Agora surge outra questão: *"É absolutamente seguro que o valor verdadeiro esteja contido nesse intervalo?"*. De modo geral não podemos garantir, com 100% de confiança, que o valor verdadeiro esteja contido no intervalo $[x-\sigma, x+\sigma]$. Para sermos completos na nossa resposta precisamos indicar a probabilidade do valor verdadeiro estar dentro do intervalo indicado. Essa probabilidade é indicada pelo **nível de confiança.**

Ao realizar um experimento espera-se também obter algum tipo de conclusão. Declarações de um número como resultado de um experimento puro e simplesmente são desinteressantes. Por exemplo:

A densidade do metal foi medida como $9,3 \pm 0,2 g/cm^3$.

A informação acima por si não fornece conteúdo adicional ao leitor. Uma conclusão deve informar ao leitor uma ou mais consequências daquela medida experimental ou estabelecer comparações com outros dados experimentais ou teorias. É nessa comparação de números que a análise de erros é importante.

Nas próximas seções serão abordados como fazer a melhor estimativa e que condições experimentais devem ser observadas.

Resumo da seção

1. Não é possível determinar o valor verdadeiro de uma grandeza.

2. É, então, necessário obter a melhor estimativa para o valor

verdadeiro.

3. em seguida, estimar um intervalo dentro do qual acreditamos que o valor verdadeiro esteja contido com um certo nível de confiança.

2 Tipos de erros

É comum classificar os erros em 2 categorias: erros aleatórios (ou estatístico) e erros sistemáticos. A Fig. ilustra os conceitos de erro aleatório e erro sistemático.

2.1 Erros aleatórios

Erros aleatórios estão associados à ideia de precisão e são causados por variações inerentes ao observador como, variações do instrumento de medida devido à flutuação da rede elétrica, umidade do ar, temperatura, etc. Se o resultado apresentado por um instrumento de medida depende da tensão da rede elétrica à qual ele está ligado e esta oscila em torno de um valor médio em intervalos de tempo muito maiores que o tempo de duração do experimento, de tal forma que tenhamos zerado o equipamento em uma condição e o estejamos usando em outra, então essa oscilação poderá ser causa de um erros sistemático. No entanto se

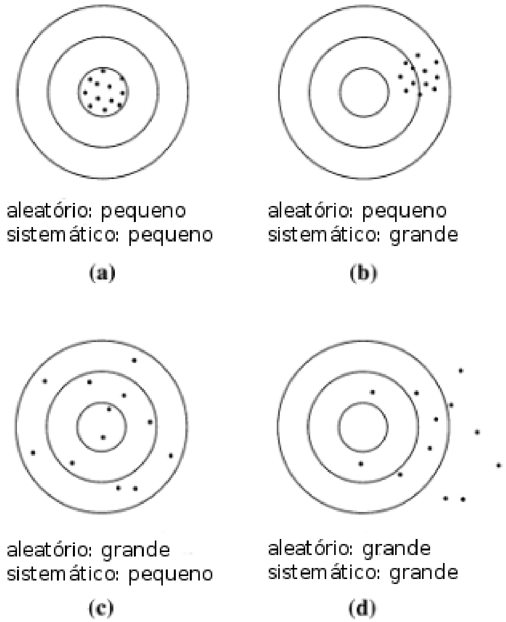

Figura 1: *Erros sistemáticos e aleatórios empregando alvos. (a) como todos os pontos atingiram pontos próximos, podemos dizer que os erros aleatórios são pequenos. Como a distribuição de pontos está concentrada no centro do alvo, os erros sistemáticos também são pequenos. (b) os erros aleatórios ainda são pequenos, mas os erros sistemáticos são bem maiores - os pontos estão sistematicamente fora do centro, a direita. (c) os erros aleatórios são grandes, mas não estão sistematicamente fora do centro. (d) ambos os erros aleatórios e sistemáticos são grandes. Extraído da ref. [1].*

a oscilação ocorre em intervalos de tempo muito menores que a duração da medida, então poderá ser causa de um erro aleatório.

2.2 Erros sistemáticos

Erros sistemáticos estão associados à idéia de acurácia e são causados por equipamentos incorretamente ajustados e/ou calibrados, uso de um procedimento incorreto pelo experimentador ou uma falha conceitual.

Os erros sistemáticos deve ser eliminados (ou reduzidos ao mínimo) pelo experimentador. Devemos assim observar se o instrumentos estão corretamente ajustados e calibrados e se os estamos usando da forma correta. Por exemplo, devemos "zerar" os aparelhos, ligá-los na fonte de tensão correta, usar as escalas de leitura convenientes, verificar se não apresentam nenhum defeito, etc.

Devemos também tomar cuidados para não cometer falhas conceituais, desprezando coisas não despreziveis. Por exemplo, se voce deseja medir o diâmetro de uma barra de Cadmio usada em um reator, perfeitamente cilindrica, e fizer o experimento a temperatura ambiente, o resultado obtido não sera o diâmetro da barra nas condiçoes de operação do reator. Para tornar o exemplo mais claro, suponha que voce obteve $D = 53,3$ mm a temperatura ambiente e o reator opera a 500°C. Como o coeficiente de dilatação linear do Cd é $3 \times 10^{-5}/°\text{C}$, o diâmetro da barra em operaçao sera $D_{op} \approx 53,3\left(1 + 3 \times 10^{-5} \times 475\right)\text{mm} = 54,1\text{mm}$, portanto diferente do diâmetro a temperatura ambiente, o que eventualmente enguiçara o reator.

Instrumentos digitais exige consideração especial. Em geral fabricantes de equipamentos de medida especificam uma tolerância; por exemplo, a tolerância de um multímetro digital é $\pm 1\%$. A precisão não pode ser melhor que a metade do último dígito no visor. A tolerância indicada pelo fabricante deve ser interpretada com cuidado.

3 Média e Desvio Padrão

A seção anterior procura levantar a questão fundamental em teoria de erros: estimar o melhor valor para x_o a partir de dados experimentais. Ao longo desta seção considere que x_i, $i = 1, 2, 3, .., N$, representa um conjunto de N dados relativos à mesma grandeza. Intuitivamente alunos da física experimental tendem utilizar a média dos dados como o valor mais realista para a grandeza mensurada.

De fato, a média é uma boa estimativa do valor verdadeiro da grandeza mensurada. Para isso, porém, dois aspectos devem

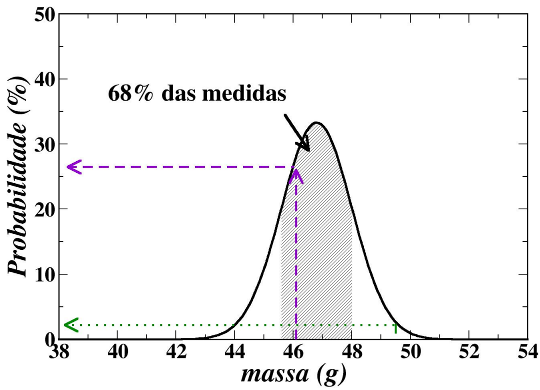

Figura 2: *Função de probabilidade gaussiana (normal) para a massa de um objeto. Numa distribuição normal, 68% dos valores estão contidos na área indicada.*

ser atentamente observados:

1. **os dados devem obedecer uma função de probabilidade gaussiana (também dita normal)**

2. **e devem ser estatisticamente independentes**

Uma função de probabilidade normal é ilustrada na Fig. 2. A primeira preposição requer que a probabilidade de realizar uma medida e obter o valor x_i seja dada por uma função normal centrada em x_o e com desvio padrão σ_o. Para compreender melhor isso, voltemos a Fig. 2. Na figura estão indicados os dados x_1 e x_2, que representam 2 dados obtidos em uma das experiências. O fato de que os dados seguem a função normal da figura implica que a probabilidade, em uma única observação, de obter o valor $x_1 = 46,1$g é de 26,5%. Consequentemente, a probabilidade de obter $x_2 = 49,5$g, em uma única observaçao, é de 2,2%. Se a forma da distribuição de probabilidades for diferente da normal isso deve ser levado em conta na análise do resultado.

Essa imposição sobre a função de probabilidade normal pode parecer bastante restritiva. No entanto é exatamente o contrário. Um teorema (teorema do valor central) prevê que a distribuição gaussiana é a distribuição obtida por um conjunto $N \to \infty$ de dados que são estatisticamente independentes.

A segunda preposição requer que os dados sejam independentes um dos outros. Em termos práticos, isso quer dizer que o valor obtido numa observação é completamente independente do valor observado anteriormente. Considere a medida do diâmetro de uma bolinha de metal utilizando um paquímetro. Nesse exemplo, para garantir que 2 observações sejam completamente independentes entre si devemos impor: i) as leituras sejam feitas por 2 observadores diferentes, para evitar fatores de erro humano como paralaxe, subjetividade no arredondamento, etc.; ii) utilizar paquímetros distintos, de fabricantes distintos. No laboratório didático é possivel utilizar diferentes paquímetros mas em laboratórios de pesquisa praticamente é impossivel cumprir completamente essa independência. A essência, no entanto, é que a contribuição dos distintos fatores deva ser pequena o suficiente para que os dados não contenham nenhum "vício" e possam ser considerados independentes.

Se os dados seguem uma distribuição normal e são estatisticamente independentes, podemos estimar o valor verdadeiro x_o de uma grandeza através da média dos dados

$$\overline{x} = \frac{1}{N}\sum_{i}^{N} x_i$$

Quando maior a quantidade de dados N mais rápido a média $\overline{x}$ tende ao valor verdadeiro.

O desvio padrão σ , obtidos a partir dos dados, é uma estimativa da *média da incerteza* e é calculado por

$$\sigma^2 = \frac{1}{N-1}\sum_{i}^{N} (x_i - \overline{x})^2$$

O desvio padrão é uma característica do arranjo experimental usado e pode depender dos instrumentos escolhidos, de variações da temperatura, do procedimento experimental, flutuações da rede elétrica, entre muitos outros fatores que fazem com que um dado obtido possa ser diferente de outro. Suponha que σ seja conhecido (através de uma medida anterior, por exemplo). Se você realizar uma única observação sob as mesmas condições, a incerteza associada a esse dado pode ser assumida como o próprio desvio padrão e o resultado seria escrito como

$$x_i \pm \sigma$$

A seguinte questão pode surgir nesse momento: *"Por que deveria medir mais vezes se com apenas uma medida já obtenho um dado x_i com sua respectiva incerteza σ na qual o valor verdadeiro x_o deve estar contido?"* Deixaremos essa questão para o aluno responder ao final desta seção.

É possivel demonstrar que a média de N dados também se distribuem em torno de x_o mas com o desvio padrão σ_m (desvio padrão da média) dado por

$$\sigma_m = \frac{\sigma}{\sqrt{N}} \tag{3.1}$$

Imediatamente vê-se que quanto maior a quantidade de dados N menor o valor de σ_m. Podemos apresentar o resultado de um experimento como

$$\overline{x} \pm \sigma_m$$

Para expressar completamente o resultado de uma medida são necessários 3 elementos (veja seção 1). Até o momento já destacamos como determinar o melhor valor para a grandeza x_o e estimamos um intervalo, $[\overline{x} - \sigma_m \,;\, \overline{x} + \sigma_m]$ no qual x_o deve estar contido. Agora nos falta dizer com que nível de confiança x_o está contido dentro deste intervalo. Antes de definir o nível de confiança observe que σ_m depende de N (eq. 3.1). Assim, o nível de confiança do intervalo $[\overline{x} - \sigma_m, \overline{x} + \sigma_m]$ irá depender da quantidade de dados observados (Tabela 1).

Vamos entender agora as informações contidas na Tabela1. Por exemplo, um aluno fez 3 observações do período de um pêndulo simples. O nível de confiança de que o valor verdadeiro do período de oscilação do pêndulo esteja contido dentro do intervalo $[\overline{x} - \sigma_m \,;\, \overline{x} + \sigma_m]$ é de 58% (veja na tabela). Se o intervalo for multiplicado por 1.32, $[\overline{x} - 1,32\sigma_m \,;\, \overline{x} + 1,32\sigma_m]$, o nivel de confiança aumenta para os 68%.

IMPORTANTE! Como norma, ao expressar os resultados sempre é definido o intervalo cujo nível de confiança é de 68%.

N	Nivel de Confiança	Fator (para intervalo conter 68%)
2	50%	1.84
3	58%	1.32
4	61%	1.20
5	63%	1.14
10	65%	1.06
20	67%	1.03
∞	68.3%	1.00

Tabela 1: *Nivel de confiaça do intervalo* $[\overline{x} - \sigma_m, \overline{x} + \sigma_m]$ *em função da quantidade de dados experimentais observados* (N).

Exemplo prático

Um grupo de estudantes mediu a aceleração da gravidade local e obteve os seguintes valores (em m/s^2): 9,75; 9,47; 10,22 e 10,05.

Assumindo que os dados seguem uma distribuição normal e são estatisticamente independentes, a média é uma boa estimativa para o valor verdadeiro da gravidade.

$$\overline{x} = \frac{1}{4}\,(9,75 + 9,47 + 10,22 + 10,05) = 9,87$$

O desvio padrão (σ), que determina como é a largura da distribuição normal a qual os dados seguem é

$$\sigma^2 = \frac{1}{(4-1)}[(9,75 - 9,87) + (9,47 - 9,87)^2 + (10,22 - 9,87)^2 + (10,05 - 9,87)^2] = 0,11$$

de tal modo que

$$\sigma = 0,33 \tag{3.2}$$

O desvio padrão da média (σ_m) será

$$\sigma_m = \frac{0,33}{\sqrt{4}} = 0,17$$

e o resultado ficaria g$= 9,87 \pm 0,17$ m/s^2.

Não sabemos o valor verdadeiro mas, estatisticamente, podemos afirmar que este encontra-se dentro do intervalo $[9,70\,;\,10,04]$ com 61% de confiança (veja Tabela 1). Por convenção define-se sempre o intervalo com nível de confiança de 68%. Para isso, temos que multiplicar σ_m por 1,20. Assim vamos escrever $\sigma'_m = 1,20 \times \sigma_m = 0,20$. Dessa forma o resultado final fica

$$g = 9,87 \pm 0,20 \mathrm{m/s^2} \quad \text{(grupo 1)}$$

Com o mesmo arranjo experimental outro grupo de alunos realizou apenas uma medida da gravidade e obteve 10,08 $\mathrm{m/s^2}$. A incerteza da medida pode ser estimada do próprio arranjo experimental mas suponhamos que esse grupo conheça o desvio padrão obtido pelo grupo anterior (eq. 3.2). O resultado dessa única medida ficaria

$$g = 10,08 \pm 0,33 \mathrm{m/s^2} \quad \text{(grupo 2)}$$

Os resultados são próximos entre si mas a precisão do grupo 1 é maior que a do grupo 2. Então, vale a pena medir mais vezes?

Resumo da seção

1. A média dos dados é uma boa estimativa para o valor verdadeiro de uma grandeza quando os dados seguem uma distribuição de probabilidade normal e são estatisticamente independente entre si.

2. O desvio padrão é uma característica do arranjo experimental como um todo, incluindo o instrumento de medida e o observador.

4 Números Significativos e Arredondamentos

Quando medimos a massa de um bloco e dizemos que é 0.5kg não estamos afirmando que a massa é exatamente 0.5000000kg. Por convenção estamos informando que a massa do bloco é conhecida até a primeira casa decimal e, portanto, seu valor verdadeiro esta entre 0.45kg e 0.55kg.

4.1 Determinando o n° de algarismos significativos

O número de algarismos significativos em um resultado é definido como:

1. o n° não-zero mais à esquerda é o *dígito mais significante.*

2. se **não há** ponto decimal, o dígito mais a direita e nao-zero é o *dígito menos significante.*

3. se **há** um ponto decimal, o dígito mais a direita, mesmo que seja um zero é o *dígito menos significativo.*

4. todos os dígitos entre o mais e o menos significativos são contados.

4.2 Critérios para Arredondamentos

Numa medida a casa do dígito menos significante é determinado pela incerteza. Por exemplo, $85,6342 \pm 0,04$ o dígito menos significante está na segunda casa decimal (3). Ao apresentar os resultados de uma medida ou na anotação dos valores de uma medida os arredondamentos devem ser realizados de acordo com a convenção definida aqui:

1. se o número seguinte for menor que 5, mantenha o dígito. Exemplo: $85,6342 \pm 0,04 \Rightarrow 85,63 \pm 0,04$

2. se o número seguinte for maior que 5, incremente o dígito. Exemplo: $152,63 \pm 1 \Rightarrow 153$

3. se o número seguinte for exatamente 5, incremente o dígito se este for ímpar. Exemplos: $8,3501 \pm 0,1 \Rightarrow 8,4 \pm 0,1$; $9,25 \pm 0,2 \Rightarrow 9,2 \pm 0,1$.

Essas convenções visam minimizar erros sistemáticos que podem surgir devido a arredondamentos. Considere os números 1,249 e 1,251. Ao arredondarmos para a segunda casa decimal obtemos

$$1,249 \Rightarrow 1,25$$
$$1,251 \Rightarrow 1,25$$

Como os dois números são próximos $(1,251 - 1,249 = 0,002)$, espera-se que o arredondamento deles leve aos mesmo valores. Ao arredondar para a primeira casa decimal é comum encontrar o seguinte:

$$\begin{aligned} 1,249 &\Rightarrow 1,2 \\ 1,251 &\Rightarrow 1,3 \end{aligned}$$

A diferença entre os dois números, que era de 0,002, passou a ser de 0,1! Se utilizarmos os critérios dispostos aqui conseguimos contornar parcialmente essa discrepância. No primeiro caso,

$$1,\overset{\downarrow}{2}49 \;\Rightarrow\; 1,2$$

que é o primeiro item da nossa convenção.

No segundo caso,

$$1,\overset{\downarrow}{2}51 \;\Rightarrow\; 1,2$$

que é o terceiro item da nossa convenção, já que o número seguinte é exatamente 5. Como o dígito da casa a ser arredondada (indicada com a seta) é par, não incrementamos o valor, permanecendo em 1,2. No final ambos os números continuam consistentes após o arredondamento.

Nossa convenção não é a prova de erros sistemáticos. Verifique os critérios de arredondamentos para os números 1,349 e 1,351, por exemplo. Você obterá que os valores arredondados serão 1,3 e 1,4, respectivamente.

Recomenda-se que os arredondamentos sejam feitos sempre ao final dos cálculos.

4.3 Manipulação com n° significativos

Soma e Subtração

Vamos somar os seguintes números: 312,17, 0,000628, 3,03, 2440,1 e 1535. O último dígito de cada um desses números é, supostamente, o duvidoso. Se representarmos o primeiro dígito desconhecido como d, a soma se torna:

$$
\begin{array}{r}
312{,}17d \\
0{,}000628d \\
3{,}03d \\
2440{,}1d \\
+1532{,}d \\
4290{,}dddddd
\end{array}
$$

A soma (subtração) de qualquer número com um número desconhecido d é também um número desconhecido. No caso acima temos que cada dígito à direita da vírgula decimal é também desconhecido. O resultado da soma é 4290, i.e., com 4 algarismos significativos. Nos poderiamos, desde o começo, ter excluido todos os dígitos à direita da virgula decimal (após arredondamento) e estabelecer a seguinte regra

Regra: selecione o termo cujo primeiro dígito incerto esteja o mais à esquerda possível (no exemplo acima, 1535). Despreze (após arredondamento) todos os dígitos nos outros termos que ocorram à direita daquele dígito.

Multiplicação e divisão

Suponha que as dimensões medidas de um retângulo sejam 14,53 cm e 35,61cm, e desejemos determinar a área do retângulo. Representando novamente os dígitos desconhecidos por x nos obtemos

$$\begin{array}{r} 14{,}53d \\ \times 35{,}61d \\ ddddd \\ 1453d \\ 8718d \\ 7265d \\ 4359d \\ 517,3ddddd \end{array}$$

A regra para adiçao foi utilizada para concluir que qualquer coluna com um d é desconhecida e portanto sua soma será representada por um d. Assim, vemos que só aparecem quatro dígitos significativos no resultado. Na multiplicaçao acima, cada um dos dois fatores tinha quatro algarismos significativos. Se repetirmos essa análise com outros fatores e com diferentes números de algarismos significativos chegaremos à conclusão de que o produto sempre tem o mesmo número de algarismo significativo que o fator com o menor número de algarismos significativos.

> **Regra:** identifique o fator com o menor número de algarismos significativos e, entao, trunque todos os outros fatores de tal modo que eles fiquem com o mesmo número de algarismos significativos que aquele fator.

5 Propagação de Erros

Frequentemente é necessário determinar o valor de uma grandeza w que é função de 1 ou mais dados experimentais. Ao longo dessa seção x, y e z representam variáveis experimentais cada um com suas respectivas incertezas

$$x \pm \sigma_x$$
$$y \pm \sigma_y$$
$$z \pm \sigma_z$$

O objetivo dessa seçao é descrever como se *propagam os erros* de cada variável para determinar o erro em w.

Estudantes que ainda não cursaram Cálculo II-B sigam diretamente para a seção 5.8.

Os detalhes que levam a *equação de propagação de erros* não serão apresentados aqui. Apesar disso, para sua validade, é importante observar durante as medidas que as variaveis (x,y e z) **não sejam correlacionadas.** Isso significa que a medida de uma dessas grandezas não dever trazer informação *a priori* de quanto deve ser o valor das outras 2 grandezas. Nessa condição, a equação de propagação de erros é

$$\sigma_w^2 \simeq \sigma_x^2 \left(\frac{\partial w}{\partial x}\right)^2 + \sigma_y^2 \left(\frac{\partial w}{\partial y}\right)^2 + \sigma_z^2 \left(\frac{\partial w}{\partial z}\right)^2 + ... \tag{5.1}$$

Vejamos como a equação acima é aplicada em diversos casos. Nos casos a seguir, a e b sao constantes.

5.1 Soma (Subtração) por uma constante

$$w = x + a$$

Nesse caso temos apenas uma variável e a eq.(5.1) se reduz a

$$\sigma_w^2 = \sigma_x^2$$

já que $\frac{\partial w}{\partial x} = 1$.

5.2 Produto por uma constante

$$w = ax$$

Novamente é uma função de 1 variável e a eq. (5.1) fica

$$\sigma_w^2 = \sigma_x^2 a^2 \ \Rightarrow \sigma_w = |a|\sigma_x$$

onde $\frac{\partial w}{\partial x} = a$.

5.3 Produto de 2 variáveis

$$w = axy \tag{5.2}$$

Esse é o primeiro exemplo que envolve 2 variáveis. A eq. (5.1) fica

$$\sigma_w^2 = \sigma_x^2 (ay)^2 + \sigma_y^2 (ax)^2$$

onde $\frac{\partial w}{\partial x} = ay$ e $\frac{\partial w}{\partial y} = ax$. A equação acima ainda pode ser escrita de uma forma mais simétrica se dividirmos ambos os lados por w^2

$$\frac{\sigma_w^2}{w^2} = \frac{\sigma_x^2(ay)^2 + \sigma_y^2(ax)^2}{(axy)^2} = \frac{\sigma_x^2}{x^2} + \frac{\sigma_y^2}{y^2}$$

em que no termo à direita foi usado a eq. (5.2). No final temos

$$\left(\frac{\sigma_w}{w}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2$$

Como seria a incerteza se w fosse o produto de 3 variáveis (i.e., $w = xyz$)? Fica

$$\left(\frac{\sigma_w}{w}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2 + \left(\frac{\sigma_z}{z}\right)^2$$

5.4 Potência

$$w = ax^b$$

Nesse caso a derivada parcial em x é

$$\frac{\partial w}{\partial x} = abx^{b-1} = \underbrace{ax^b}_{w} bx^{-1} = \frac{bw}{x}$$

e a eq. (5.1) fica

$$\sigma_w^2 = \sigma_x^2 \left(\frac{bw}{x}\right)^2$$

que pode ser reescrito como

$$\left(\frac{\sigma_w}{w}\right) = |b| \left(\frac{\sigma_x}{x}\right)$$

5.5 Exponencial

$$w = ae^{bx}$$

A derivada parcial em x fica

$$\frac{\partial w}{\partial x} = bae^{bx} = bw$$

e σ_w é dado por

$$\sigma_w^2 = \sigma_x^2 \left(bw\right)^2 \ \Rightarrow \ \frac{\sigma_w}{w} = \frac{\sigma_x}{|b|}$$

5.6 Logaritmo

$$w = a\ln(bx)$$

A derivada parcial em x fica

$$\frac{\partial w}{\partial x} = \frac{ab}{x}$$

e a eq.(5.1) fica

$$\sigma_w = \sigma_x \left(\frac{|ab|}{x} \right)$$

5.7 Cosseno

$$w = a\cos(bx)$$

A derivada com respeito a x é

$$\frac{\partial w}{\partial x} = -ab\mathrm{sen}(bx)$$

e

$$\sigma_w = \sigma_x ab\mathrm{sen}(bx)$$

Importante lembrar que nesse caso, como também para a função seno, a **incerteza deve ser calculada em radianos!**

5.8 Tabela de propagação de erros

A tabela a seguir sumariza as principais equações para a propagação de erros de acordo com a função.

função w	**Erro em** w (σ_w)
$w = x + a$	$\sigma_w = \sigma_x$
$w = ax$	$\sigma_w = a\sigma_x$
$w = axy$	$\left(\frac{\sigma_w}{w}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2$
$w = xyz$	$\left(\frac{\sigma_w}{w}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2 + \left(\frac{\sigma_z}{z}\right)^2$
$w = ae^{bx}$	$\frac{\sigma_w}{w} = \frac{\sigma_x}{b}$
$w = a\ln(bx)$	$\sigma_w = \sigma_x \left(\frac{ab}{x}\right)$
$w = a\cos(bx)$	$\sigma_w = \sigma_x ab\,\mathrm{sen}(bx)$
$w = a\,\mathrm{sen}(bx)$	$\sigma_w = \sigma_x ab\cos(bx)$

Tabela 2: *Tabela de propagaçao de erros para as funçoes mais comuns. Nas funçoes, a e b sao constantes.*

6 Comparação entre valores

Frequentemente desejamos comparar valores experimentais com teorias ou com outros valores medidos independentemente. No experimento para determinar a velocidade do som no ar (Experimento 7 do Laboratório Experimental III/XIX), um grupo de estudantes obteve o seguinte valor

$$345 \pm 5\text{m/s} \ \ \text{(grupo 1)}$$

O valor aceito para a velocidade do som, nas condições de temperatura e pressão normais, é

$$343\text{m/s} \ \ \text{(valor aceito)}$$

Ao apresentar o relatório o grupo deve indicar os valores experimental e aceito próximas uma das outras para permitir que o leitor rapidamente aprecie os resultados. Também deve incluir uma declaração explicita de que, tendo em vista que o valor aceito encontra-se dentro de sua margem de erro, sua medida é **compatível**.

Conforme discutido no final da seção 3, a incerteza indica que o valor verdadeiro da grandeza encontra-se entre $x - \sigma$ e $x + \sigma$ com 68% de probabilidade. É bem possível que o valor verdadeiro esteja ligeiramente fora desse intervalo. Portanto uma medida pode ser considerada compatível mesmo se que o valor aceito esteja fora do intervalo estimado. Por exemplo, o grupo 2 obteve o seguinte valor:

$$336 \pm 5\text{m/s} \ \ \text{(grupo 2)}$$

Esse grupo certamente pode reivindicar que sua medida é consistente com o valor aceito.

Por outro lado, se o valor aceito estiver claramente fora das margens de erro há razão para pensar que algo deu errado. Por exemplo, o grupo 3 obteve o seguinte valor

$$364 \pm 4\text{m/s} \ \ \text{(grupo 3)}$$

A diferença entre o valor do grupo 3 e o valor aceito é de 21m/s; mais que 5 vezes a incerteza obtida. Com certeza o grupo deverá revisar suas medidas e cálculos. Excluindo-se erros de medidas e nos cálculos, outras possibilidades podem ser:

- incerteza subestimada
- erros de calibraçao nos instrumentos
- condiçées de temperatura e pressão distintas das normais

6.1 A média ponderada

Uma grandeza é medida várias vezes, por grupos de estudantes distintos, por turmas distantes e por períodos distintos. As vezes é interessante combinar essas medidas para produzir uma única melhor estimativa. Ao combinar distintas medidas queremos que os valores maior precisão tenham um peso maior que as medidas de menor precisão. Esse é o caso na qual utilizamos uma **média ponderada.**

Suponha que tenhamos um conjunto de N medidas

$$x_1 \pm \sigma_1;\ x_2 \pm \sigma_2;\ x_3 \pm \sigma_3;\ ...$$

A média ponderada é definida por

$$\text{média ponderada} = \frac{\sum \alpha_i x_i}{\sum \alpha_i} \qquad \text{onde } \alpha_i = \frac{1}{\sigma_i}$$

Antes porém deve-se tomar o cuidado para que todos os dados que compõem a média ponderada sejam compatíveis entre si.

7 Construção de gráficos

7.1 Considerações gerais

É sempre desejável visualizar a dependência entre duas grandezas como, por exemplo, a velocidade de um carrinho em função do tempo (Experimento 1, Física Experimental I). Ao construirmos o gráfico dos dados experimentais podemos observar claramente a dependência entre as variáveis e, além disso, extrair a aceleração média a partir da reta tangente aos dados.

Para a construção apropriada de um gráfico o princípio básico que deve guiar é: **apresentar as informações de forma clara ao leitor.**

Qualquer gráfico deve conter os seguintes elementos:

1. Título simples de modo a descrever suscintamente seu conteúdo.

2. Legenda dos eixos vertical (ordenadas) e horizontal (abscissas). Por descrição entende-se o nome ou símbolo da grandeza com sua respectiva unidade entre parênteses.

3. Uma escala em cada eixo. As divisões no papel milimetrado devem corresponder a números fáceis de serem tratados. Indicações numéricas dos eixos deve ser em diferenças de unidades de 1 em 1, 2 em 2, 5 em 5 ou um desses

números multiplicados por uma potência de 10. Evite números como 3, 7 ou 11. Uma boa escala permite a leitura imediata dos valores de um ponto contido no gráfico.

4. As incertezas são representadas como barras de erros em torno de cada ponto do gráfico.

5. Legenda dos dados. Quando mais de um conjunto de pontos é representado no gráfico é necessário diferenciar cada conjunto usando símbolos diferentes (como circulos, quadrados, triângulos, ...). A legenda é um quadro inserido no gráfico onde se coloca o símbolo ao lado de um texto curto que especifica qual conjunto de dados aquele símbolo representa.

As figuras (3) e (4) apresentam exemplos de gráficos com erros e outro correto. Qual deles facilita a leitura dos dados? Veja na Fig. (4) que é possivel identificar tendencias distintas para os comportamento dos dados do grupo 1 e do grupo 2.

7.2 Histogramas

(A ser implementado).

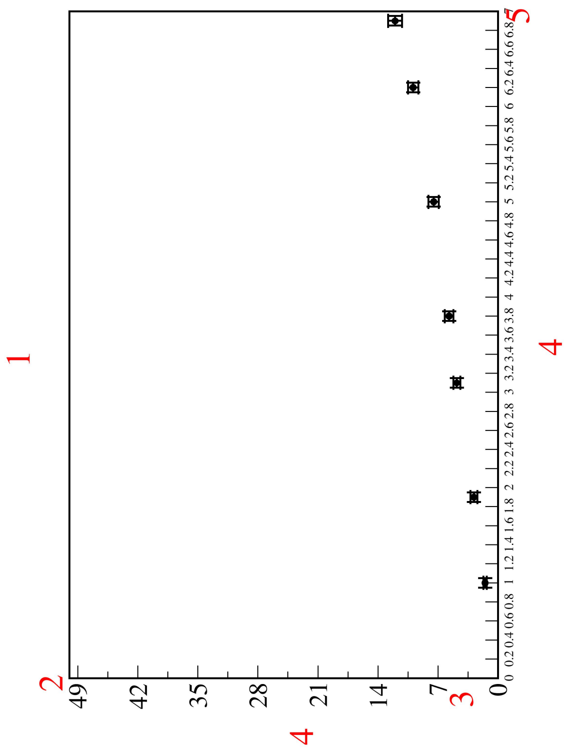

Figura 3: *Exemplo de um gráfico contendo erros de apresentação. Erro 1: ausência de um título. Erro 2: compromento do eixo inapropiada. Erro 3: escala do eixo inapropriada, dificultando a leitura dos valores. Erro 4: ausência das legendas nas escalas vertical e horizontal. Erro 5: escala do eixo inapropriada, apesar de permitir a leitura dos valores dos dados, polui a sua visualização.*

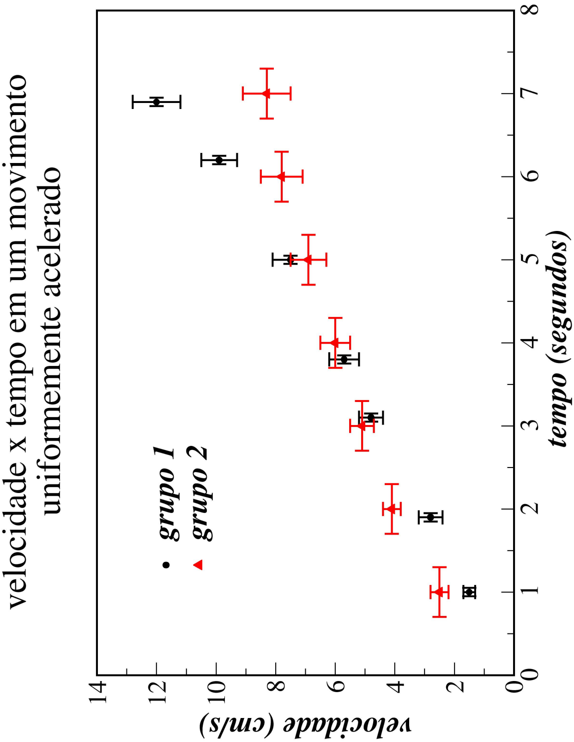

Figura 4: *Exemplo de um gráfico correto.*

8 Ajustes de retas

Os dados do grupo 1 da Fig.(4) indicam um comportamento proporcional entre velocidade e tempo. Para verificar essa proporcionalidade você poderia desenhar uma reta que passa pelos dados através da origem. Em outras situações os dados podem apresentar um comportamento linear, em que se comportam como uma reta que passa fora da origem. Em alguns experimentos os dados obtidos apresentam um comportamento quadrático ou exponencial. Em ambos os casos é possível linearizar as grandezas de tal forma a permitir uma análise linear. Nesse tipo de análise estamos procurando a melhor reta que descreva o comportamento dos dados. Essa reta é escrita como

$$y = A + Bx$$

e em geral os coeficientes linear A e angular B contém informações relevantes.

8.1 Método dos Mínimos Quadrados

O método dos Mínimos Quadrados (MMQ) resulta da minimização do quadrado da distância entre os valores experimentais de y_i e os valores calculados como $y_i^{'} = f(x_i)$. Não será deduzido aqui as equações que determinam os parâmetros A e B pelo método dos mínimos quadrados. No entanto é importante conhecer as consideração que levarão às equações abaixo:

1. As incertezas em x são muito pequenas, de tal modo a considerarmos $\sigma_x \approx 0$. Por mais que σ_x não seja nulo, a incerteza relativa em x é muito menor que em y (i.e., $\sigma_x/x \ll \sigma_y/y$).

2. As incertezas em y são da mesma ordem de grandeza, de tal modo que $\sigma_1 \approxeq \sigma_2 \approxeq \sigma_i \approxeq \sigma_y$.

Os coeficientes são dados por (onde por simplicidade os índices i foram omitidos):

$$A = \frac{\sum y \sum x^2 - \sum x \sum xy}{N \sum x^2 - \left(\sum x\right)^2}$$

e

$$B = \frac{N \sum xy - \sum x \sum y}{N \sum x^2 - \left(\sum x\right)^2}$$

As incertezas de A e B (σ_A e σ_B, respectivamente) são:

$$\sigma_A = \sigma_y \sqrt{\frac{\sum x^2}{N \sum x^2 - \left(\sum x\right)^2}}$$

e

$$\sigma_A = \sigma_y \sqrt{\frac{N}{N \sum x^2 - \left(\sum x\right)^2}}$$

onde $\sigma_y = \sqrt{\frac{1}{N-2} \sum \left(y_i - A - Bx_i\right)^2}$.

8.2 Tabela para o MMQ

As equações que determinam os coeficientes e suas respectivas incertezas pelo método dos mínimos quadrados são fáceis de serem calculadas se houver organização nos cálculos de cada uma das somatórias. A tabela abaixo serve de auxílio nessas contas. Note que basicamente precisamos de $\sum x$, $\sum y$, $\sum xy$ e $\sum x^2$.

Referências

[1] TAYLOR, J.R. (1982). Introdução à análise de erros: o estudo de incertezas em medições físicas. 2ª edição. Trad. Waldir Leite Roque. Porto Alegre, Bookman, 2012.

[2] BEVINGTON, P.R. (1969). Data reduction and error analysis for the physical sciences. 3rd edition. New York, The McGraw-Hill, 2003.

dado	x_i	y_i	$x_i y_i$	x_i^2	$(A - Bx_i) = \beta$	$(y_i - \beta)^2$
1						
2						
3						
4						
5						
6						
$\vdots$						
SOMA						

Tabela 3: *Tabela de auxilio para calculo de coeficientes pelo metodo dos minimos quadrados.*

[3] HELENE, O.A.M. e VANIN, V.R. (1981). Tratamento estatístico de dados em física experimental. 2ª edição. São Paulo, Edgard Blucher, 1991.