

2019

LABORATÓRIO DE FÍSICA I



Departamento de Físico-Química

Área de Física

Instituto de Química de Araraquara – UNESP

Engenharia de Bio-Processos e Biotecnologia

01/08/2019

“O que sabemos é uma gota e o que ignoramos é um oceano”

Isaac Newton

Agradecimentos

A elaboração desta apostila teve a colaboração dos professores Prof. Dr. Carlos de Oliveira Paiva Santos, Prof. Dr. Paulo Roberto Bueno, Prof. Dr. Marcelo Ornaghi Orlandi, Prof. Dr. Gustavo Troiano Feliciano, Prof. Dr. Anderson André Felix, do bolsista Prof. Adriano dos Santos e do técnico do laboratório de física Vinicius Orsi Valente.

SUMÁRIO

1. GRANDEZAS FÍSICAS E O SISTEMA INTERNACIONAL DE UNIDADES (SI).....	4
2. OS NÚMEROS.....	6
2.1 Os conceitos de exatidão e precisão	7
2.2 Algarismos significativos.....	8
2.3 Operações com algarismos significativos	9
3. ERROS E INCERTEZAS	10
3.1 Dispersão de conjunto de dados experimentais	12
3.2 Propagação de incertezas	14
3.3 Erro Percentual Relativo (E%)	16
4. REPRESENTAÇÃO GRÁFICA E REGRESSÃO LINEAR.....	16
4.1 Método Gráfico: Utilizando papel milimetrado.	18
4.2 Método Gráfico: Utilizando papel em escala logarítmica.....	20
4.3 Regressão linear pelo Método dos Mínimos Quadrados (MMQ).....	23
4.1.1 MMQ com incerteza em y.....	26
5. ROTEIROS PARA AS AULAS EXPERIMENTAIS.....	29
5.1 Elaboração dos relatórios	29

1. GRANDEZAS FÍSICAS E O SISTEMA INTERNACIONAL DE UNIDADES (SI)

Segundo o Vocabulário Internacional de Termos Fundamentais e Gerais da Metrologia (VIM), uma grandeza pode ser definida como sendo *a propriedade de um fenômeno, de um corpo ou de uma substância, que pode ser expressa quantitativamente sob a forma de um número e de uma referência*, que nada mais é do que uma unidade de medida. Quando dizemos que um prédio tem 15 metros de altura, a grandeza utilizada é o *comprimento*, sendo o “metro” a unidade de medida utilizada como referência. Outro exemplo seria a grandeza *tempo*, da qual utilizamos as unidades hora, minutos ou segundos para expressar a sua quantidade. Numa linguagem informal, podemos definir como grandeza *tudo aquilo que pode ser medido*, como a propriedade de um corpo ou a característica de um fenômeno (massa e velocidade, por exemplo). *Grandeza física* é um conceito específico de grandezas, que as limita no campo em que a Física atua.

No meio acadêmico, é muito comum a prática de comparações entre resultados e valores entre experimentos conduzidos em diferentes localidades. Seria inoportuno, então, adotar unidades de medidas próprias, pois ninguém conseguiria comparar os dados experimentais obtidos com outros resultados. Igualmente inconveniente seria quando alguém viajasse a outro país e encontrasse determinadas grandezas expressas em unidades diferentes e sem relação com aquelas que estavam habituados. Grandes problemas também ocorreriam na comparação de exames clínicos de dois laboratórios, os quais utilizassem unidades arbitrárias e sem relação. Para minimizar essa discrepância, que embora ainda exista, convencionou-se adotar unidades padrões, as quais todos podem utilizar para comparar resultados possibilitando, assim, comparar objetos ou fenômenos diretamente. Com o objetivo de resolver esta questão, nasceu o BIPM (Bureau Internacional de Pesos e Medidas) em 1875, outras organizações como a CGPM (Conferência Geral de Pesos e Medidas) e o CIPM (Comitê Internacional de Pesos e Medidas). Essas organizações prezavam desde o início pela uniformidade mundial das medidas e conformidade com o Sistema Internacional (SI), que é basicamente um conjunto de nomes e símbolos de grandezas em conformidade com a CGPM e é amplamente utilizado até hoje.

O SI é baseado nas sete grandezas de base que se encontram resumido na Tab. 1.1.

Tab. 1.1 Grandezas de base do SI

Grandeza	Símbolo da grandeza	Símbolo
Comprimento	$l, x, r, \text{ etc.}$	L
Massa	m	M
Tempo	t	T
Corrente elétrica	I, i	I
Temperatura termodinâmica	T	Θ
Quantidade de substância	n	N
Intensidade luminosa	I_v	J

Todas as outras grandezas são derivadas, o que significa que surgem após uma operação entre as grandezas de base. Por exemplo, a unidade para a grandeza derivada “velocidade” surge da divisão entre duas grandezas: o comprimento (distância) pelo o tempo.

Conforme a definição de grandeza, ela deve ter uma referência que geralmente é a unidade de medida. Desta forma, o SI possui, para cada grandeza de base, a unidade correspondente. Essa unidade, por sua vez, possui definições que não são fixas, pois são

constantemente revisadas devido ao avanço da ciência no campo das medições. De fato, as definições oficiais foram aprovadas pelo CGPM em 1889 e a mais recente, em 1983. Mais informações sobre estas convenções e definições visite o site do INMETRO (<http://www.inmetro.gov.br/noticias/conteudo/sistema-internacional-unidades.pdf>). A Tab. 1.2 resume as unidades do SI e as suas definições.

Tab. 1.2 Grandezas, unidades SI e suas definições.

Grandeza	Unidade SI (símbolo)	Definição
Massa	quilograma (kg)	Massa igual ao protótipo internacional do quilograma (protótipo constituído de platina-lítio).
Comprimento	metro (m)	O metro é o comprimento do trajeto percorrido pela luz no vácuo durante um intervalo de tempo de $1/299.792.458$ segundo.
Tempo	segundo (s)	O segundo é a duração de $9.192.631.770$ períodos da radiação correspondente à transição entre os dois níveis hiperfinos do estado fundamental do átomo de cézio 133.
Corrente elétrica	ampere (A)	O ampere é a intensidade de uma corrente elétrica constante que, se mantida em dois condutores paralelos, retilíneos, de comprimento infinito, de seção circular desprezível, e situados à distância de 1 metro entre si, no vácuo, produz entre estes condutores uma força igual a 2×10^{-7} newtons por metro de comprimento.
Temperatura	kelvin (K)	O kelvin, unidade de temperatura termodinâmica, é a fração $1/273,16$ da temperatura termodinâmica do ponto triplo da água.
Quantidade de substância	mol (mol)	O mol é a quantidade de substância de um sistema que contém tantas entidades elementares quantos átomos existem em $0,012$ quilograma de carbono 12.
Intensidade luminosa	candela (cd)	A candela é a intensidade luminosa, numa dada direção, de uma fonte que emite uma radiação monocromática de frequência 540×10^{12} hertz e que tem uma intensidade radiante nessa direção de $1/683$ watt por esferorradiano.

Assim como existem as grandezas derivadas, existem as unidades derivadas do SI. Por exemplo, no caso da velocidade, que seria o quociente das grandezas comprimento (distância) e tempo, a unidade SI seria metro por segundo, simbolicamente representado por m/s. Algumas unidades derivadas se encontram resumidas na Tab. 1.3.

Tab. 1.3 Algumas grandezas e unidades derivadas SI

Grandeza derivada		Unidade derivada SI	
Nome	Símbolo	Nome	Símbolo
Área	A	metro quadrado	m^2
Volume	V	metro cúbico	m^3
Velocidade	v	metro por segundo	m/s
Aceleração	a	Metro por segundo ao quadrado	m/s^2
Densidade, massa específica	ρ	quilograma por metro cúbico	kg/m^3
Densidade superficial	ρ_A	quilograma por metro quadrado	kg/m^2
Densidade de corrente	j	ampere por metro quadrado	A/m^2
Força	N	newton	$m.kg s^{-2}$
Pressão	Pa	pascal	N/m^2
Frequência	Hz	hertz	s^{-1}
Viscosidade dinâmica	$Pa.s$	pascal segundo	$m^{-1}.Kg.s^{-1}$

Em muitos casos, é conveniente representar as unidades utilizando múltiplos e submúltiplos, como mostrado na Tab. 1.4.

Tab. 1.4 Prefixos do SI

Fator	Nome do Prefixo	Símbolo	Fator	Nome do Prefixo	Símbolo
10^1	deca	da	10^{-1}	deci	d
10^2	hecto	h	10^{-2}	centi	c
10^3	kilo	k	10^{-3}	mili	m
10^6	mega	M	10^{-6}	micro	μ
10^9	giga	G	10^{-9}	nano	n
10^{12}	tera	T	10^{-12}	pico	p
10^{15}	peta	P	10^{-15}	femto	f
10^{18}	exa	E	10^{-18}	atto	a
10^{21}	zetta	Z	10^{-21}	zepto	z
10^{24}	yotta	Y	10^{-24}	yocto	y

Por exemplo, se a massa de um objeto é 0,001 g, pode-se representar esse valor utilizando o prefixo mili. Desta forma, $0,001g = 1mg$ (miligrama). Da mesma forma, 1.000 metros = 1km (quilômetro), $3 \times 10^{-6} = 3 \mu g$ (microgramas) e $1 TB = 10^{12}$ bytes.

2. OS NÚMEROS

A ciência exata é numérica. Ela tem a necessidade intrínseca de representar valores por meio de números e algarismos. Desta forma, resultados de experimentos, ensaios de laboratório de análise clínica, pesquisas de campo, rendimento de processos industriais e até testes de DNA são representados por meio da linguagem matemática e estatística. Para uma interpretação sem

equivocos, é necessário um prévio conhecimento de como abordar um conjunto de dados e “tratá-lo” adequadamente em planilhas e gráficos, bem como descrevê-lo de forma adequada e elegante para que possibilite uma visualização rápida e eficiente dos resultados.

Neste capítulo serão tratados os princípios básicos de como se abordar um conjunto de dados e como descrevê-lo de forma conveniente. Serão apresentados os conceitos de precisão e exatidão de números que, juntamente com as definições de dispersão de medidas da estatística descritiva, serão fundamentais para a interpretação da qualidade de um conjunto de dados.

2.1 Os conceitos de exatidão e precisão

Os números podem ser *exatos* ou *aproximados*. A diferença entre estas definições depende do conjunto que representa esses números. Por exemplo, num jogo de futebol há *exatamente* 11 jogadores de cada time e num jogo de xadrez, 16 peças brancas e 16 peças negras. Ou seja, *números exatos são aqueles que não apresentam incerteza*. Por outro lado, existem situações em que não é possível conhecer exatamente o valor de uma grandeza por restrições técnicas e operacionais, como em resultados de teste de paternidade, por exemplo. Neste caso, mesmo a margem de certeza sendo grande (99,9999%), o resultado tem chance de ser falso positivo em 0,0001%. Isso significa que, embora o número possa ser pequeno, de 1.000.000 de testes, um será falso positivo, ou seja, o exame afirmaria a paternidade equivocadamente. Além da margem de erro ocasionada pelo falso positivo, a probabilidade de *falso negativo* é ainda maior, na ordem de 1%. Isto significa que, em cada 100 casos em que o exame é negativo, um poderá estar errado, ou seja, a pessoa é realmente o pai. Isto significa que o exame de DNA não é tão preciso como comumente se divulga na mídia, uma vez que os valores do teste possuem um grau de incerteza inerente à própria técnica e metodologia do exame. Desta forma, pode-se definir como sendo *números aproximados aqueles em que não é possível conhecer com exatidão o real valor da grandeza, seja por limitações técnicas como operacionais*.

Com relação aos números aproximados, dois termos podem ser utilizados para descrever a confiança numérica: a exatidão e a precisão. A exatidão *está relacionada ao verdadeiro valor da grandeza* sendo estudada. A precisão é o quanto os valores obtidos em uma medição são reprodutíveis. Para ilustrar essa questão, imaginemos o seguinte problema:

Uma forma ilustrativa de definir precisão e exatidão pode ser conseguida ao desafiar seu amigo para uma partida de sinuca ou a um jogo de dardos (Fig. 2.1). Se o seu amigo possuir *alta precisão* no arremesso, mas “pouca pontaria” (baixa exatidão), ele conseguirá atingir uma mesma área do alvo do jogo de dardos em várias tentativas, mas longe do centro. Por outro lado, se ele possuir “boa pontaria”, ele estaria acertando o alvo central em vários arremessos, e sua exatidão seria alta.

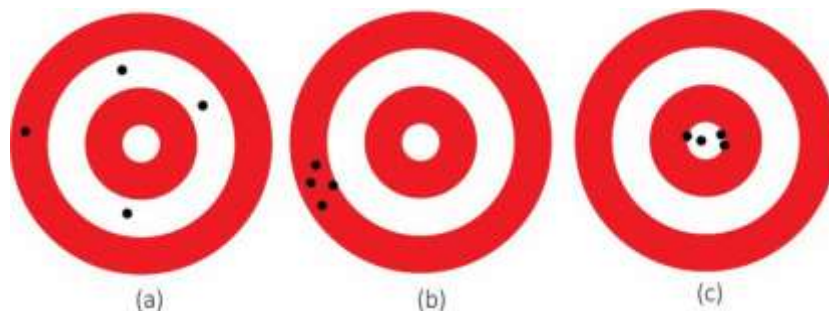


Figura 2.1. Ilustração da definição de precisão e exatidão em um jogo de dardos. (a) baixa precisão e baixa exatidão; (b) alta precisão e baixa exatidão; (c) alta precisão e alta exatidão.

2.2 Algarismos significativos

O problema da precisão requer que tenhamos uma forma conveniente de escrever um número. Expressar uma grandeza de forma correta é primordial para a interpretação inequívoca do conjunto de dados. Uma forma conveniente de expressar a precisão de um número é por meio dos *algarismos significativos*.

Antes de definir o que é um algarismo significativo voltemos ao problema da balança. A balança registra, em cada pesagem, quatro algarismos, como por exemplo, 48,24 kg. Esse número em questão possui *quatro algarismos significativos*. Portanto, pode-se definir *algarismos significativos como sendo a quantidade de números que representa um valor de uma medida ou grandeza, exceto os zeros à esquerda*. Quando maior for a quantidade de algarismos de um número, maior será a sua *precisão*. Entretanto, a quantidade de algarismos está inerentemente ligada à capacidade do dispositivo em medir a grandeza, bem como a metodologia empregada. Geralmente, a última casa representa o valor que é *incerto* ou aproximado. Isto porque, em qualquer medição, é importante registrar o maior número de algarismos que o dispositivo e o método de medida permitem. No caso da balança em questão, ela pode registrar os três primeiros algarismos, sendo o último, aproximado e incerto.

Para determinar o número de algarismos significativos de medições, seus dígitos devem ser contados, inicialmente pelo primeiro dígito (diferente de zero) à esquerda. *Os zeros à esquerda e aqueles que são colocados para posicionar a vírgula não devem ser considerados*.

Quando um número é escrito em notação exponencial, o número de algarismos significativos é determinado pelo coeficiente. Desta forma, 9×10^3 possui apenas um algarismo significativo, enquanto $9,0 \times 10^3$ e $9,02 \times 10^3$ possuem dois e três, respectivamente.

Uma questão importante e que deve ser resolvida com atenção são os números que terminam em zero. Por exemplo, imagine que as medições de duas determinadas grandezas registraram os seguintes valores: 30.000 e 45.000. Se o instrumento utilizado e a metodologia empregada possibilitam a obtenção dos valores da medição com aproximação de *mil*, os números em questão possuem apenas dois algarismos significativos, uma vez que os zeros existem apenas para posicionar o ponto e, portanto, não são significativos. Neste caso, os algarismos significativos seriam os algarismos *três* e *zero* para o primeiro número e *quatro* e *cinco* para o segundo. Entretanto, se estes números são expressos com aproximação até *cem*, eles têm três algarismos significativos, os dois já mencionados mais o primeiro zero, contando da esquerda para a direita. Se a aproximação fosse até *a dezena*, os números teriam quatro algarismos significativos, e assim sucessivamente. Sendo assim, uma forma conveniente de contornar esse problema é representá-

lo pela notação exponencial, utilizando o número de algarismos no coeficiente de acordo com a precisão da medida realizada.

Veja os exemplos na Tabela 2.1.

Tabela 2.1 Números e quantidade de algarismos significativos.

Número	Quantidade de algarismos significativos	Observação
8	1	-
8,7	2	-
8,78	3	-
8,781	4	-
2,01	3	O zero no meio do número deve ser considerado
70	1	O zero no final do número deve ser considerado se conhecer a precisão do número
0,001	1	Os zeros posicionados à esquerda e para posicionar a vírgula não devem ser considerados
80,01	4	O zero no meio do número deve ser considerado
0,89000	5	Dependerá da precisão do número.
1,8910	4 ou 5	Dependerá da precisão do número.
$2,1 \times 10^6$	2	O número de algarismos significativos é o número de algarismos do coeficiente
4×10^9	1	-
$2,99792458 \times 10^8$	9	Velocidade da luz no vácuo em $\text{m}\cdot\text{s}^{-1}$
$1,672623 \times 10^{-24}$	7	Massa do próton, em g
8.000	1 à 4	Não se sabe o número de algarismos significativos, pois não é conhecida a precisão do número.
$8,00 \times 10^3$	1 à 3	Dependerá da precisão do número.
$6,022137 \times 10^{23}$	7	Número de Avogrado, em mol^{-1}

2.3 Operações com algarismos significativos

É muito importante que o resultado de uma operação aritmética seja expresso com o número de algarismos significativos corretos para que ele não tenha maior e nem menor precisão do que a especificada pelas medições que originaram o cálculo. Para que esse inconveniente não ocorra, existem algumas regras que devem ser obedecidas de acordo com a operação em questão. São elas:

Regra para a adição e subtração: o número de dígitos à direita da vírgula no resultado calculado deve ser o mesmo do número com menos dígitos dos números somados ou subtraídos.

Regra para a multiplicação e divisão: o número de algarismos significativos no resultado calculado deve ser o mesmo que o menor número de algarismos significativos dos números envolvidos na operação.

Regra para operações com logaritmo: o número de dígitos após a vírgula decimal no logaritmo de um número é igual ao número de algarismos significativos do próprio número (número em que a operação logarítmica está sendo utilizado)

Após a realização de uma operação matemática é necessário se realizar o *arredondamento* do valor calculado. De forma geral, a redução do número de dígitos obedece as seguintes regras:

- a. Se o dígito a ser eliminado é maior que 5, o dígito precedente é aumentado de uma unidade (Ex. 5,56 é arredondado para 5,6).
- b. Se o dígito a ser eliminado é menor que 5, o dígito precedente é mantido (Ex. 3,34 é arredondado para 3,3).
- c. Se o número final for 5, normalmente se deixa par o último dígito do algarismo arredondado. Neste caso, 5,65 é arredondado para 5,6 e 5,75 é arredondado para 5,8. Entretanto, essa abordagem é arbitrária.

3. ERROS E INCERTEZAS

Medições de grandezas físicas não são valores absolutos ou verdadeiros uma vez que possuem erros¹ associados à metodologia empregada, ou seja, ao método² e ao procedimento de medição³.

Erros nas medições ocorrem por vários motivos. Um exemplo prático seria a realização, por 40 pessoas, de medições de massa de um objeto utilizando uma balança analítica. Provavelmente, dentre as pesagens, algumas não seriam idênticas às outras e não seria possível afirmar que uma medição esteja incorreta somente baseando-se no fato de que uma medida não é idêntica à outra. Na verdade, se for considerado que a metodologia empregada seja certa e que a balança esteja corretamente calibrada, todos os resultados estão corretos, apenas variam em torno do valor real (ou o mais próximo possível) da massa do objeto. Desta forma, o resultado é apenas uma estimativa e para que esteja completo precisa da incerteza associada a ele.

A definição de erro, entretanto, tem pouco significado prático. Sabendo-se que o *erro* é a subtração do valor medido do valor verdadeiro e é praticamente impossível calculá-lo, já que temos que conhecer previamente o valor verdadeiro do objeto de estudo. Entretanto, o conceito de erro é fundamental para compreender os motivos que os valores experimentais desviam dos valores verdadeiros.

¹ Importante notar que erro e incerteza não são sinônimos. Erro seria a diferença entre o valor medido (a) e o valor verdadeiro b (valor real da grandeza medida), $\text{Erro} = a - b$. Por outro lado, incerteza é a *melhor estimativa* do erro. Na prática, trabalha-se com incertezas.

² Segundo o VIM, método é a *descrição genérica duma organização lógica de operações utilizadas na realização duma medição*. Os métodos podem ser qualificados de vários modos como, por exemplo, método de definição “de zero”.

³ De acordo com o VIM, é a *descrição detalhada duma medição de acordo com um ou mais princípios de medição e com um dado método de medição, baseada num modelo de medição e incluindo todo cálculo destinado à obtenção dum resultado de medição*.

Tradicionalmente, o erro é constituído de três componentes: erros aleatórios, erros sistemáticos e erros grosseiros. *Erros aleatórios* são ocasionados por efeitos que se originam em variações temporais ou espaciais, imprevisíveis. Erro aleatório ocorre quando várias repetições são realizadas obtendo resultados que variam de forma imprevisível. *Erros sistemáticos*, por outro lado, são erros que variam de forma previsível ao longo de várias repetições como, por exemplo, oriundas da calibração incorreta do instrumento de medida. *Erros grosseiros* são erros normalmente associados à má execução do experimento e/ou medida ou até mesmo ao tratamento de dados errôneo.

Já a Incerteza é definida como sendo a dispersão dos valores referentes à medida de uma grandeza, pois ela reflete a falta de conhecimento exato do valor sendo medido. Conhecer as incertezas nas medições é muito importante uma vez que é uma indicação qualitativa dos resultados. Somente fornecendo as medidas juntamente com as incertezas, é possível comparar os resultados. Isto é, é por meio da incerteza que é possível quantificar a confiabilidade dos resultados: quanto maior a incerteza, menor a confiabilidade. Porém incerteza não deve ser confundida com erro. O cálculo do erro só é possível se for conhecido o valor verdadeiro do objeto em estudo. Por outro lado, o cálculo da incerteza não possui essa restrição e, por isso, tem maior significado prático e maior aplicabilidade que o erro.

As incertezas, por serem calculadas por duas formas distintas e são classificadas em tipos: incertezas do Tipo A e do Tipo B.

Incerteza do *Tipo A* são obtidos por métodos estatísticos. Para isso, as medidas devem ser realizadas repetidamente ou em uma série de observações da mesma grandeza física. Desta forma, quando calculamos a média e o desvio padrão de um conjunto de dados experimentais, estamos avaliando as incertezas do tipo A. A abordagem estatística básica da incerteza do tipo A será discutida na próxima seção.

As incertezas do *Tipo B* são avaliadas por meios que não sejam os adotados no Tipo A, ou seja, que não são realizadas por meio estatístico como, por exemplo, a incerteza de uma medida obtida por um certificado de calibração do instrumento utilizado. Elas são imprescindíveis quando é muito difícil ou desnecessário realizar uma série de observações. Um exemplo de incertezas do *Tipo B* que utilizaremos em nossos experimentos é a incerteza instrumental geralmente obtida pela metade da menor divisão da escala do instrumento de medida. Aprenderemos mais sobre este tipo de incerteza e como identificá-lo quando aprendermos a utilizar diferentes instrumentos de medidas, tais como a régua, o paquímetro e o micrômetro.

Um problema importante que normalmente encontramos é como propagar corretamente a incerteza das medidas. Por exemplo, queremos determinar a área de um terreno retangular e para isso dispomos de um instrumento de medição que possui incerteza associada ao resultado. Como a área de um terreno retangular é obtida pela multiplicação de seus lados, como poderemos estimar corretamente o erro associado à área, já que cada lado tem o seu erro associado? Será que deveríamos multiplicar os erros nesse caso?

Neste capítulo discutiremos como as incertezas do *Tipo A* são calculadas e a forma conveniente de propagá-las nos cálculos.

3.1 Dispersão de conjunto de dados experimentais

Vimos na seção 2 que registrar um valor com precisão e exatidão são fundamentais para análise de resultados. No caso de empresas e indústrias, checar seus instrumentos e realizar calibrações periódicas pode colaborar para diminuição de prejuízos e gastos.

Neste curso, as incertezas do Tipo A serão as utilizadas. Como foi visto, são as incertezas que estão relacionadas com parâmetros estatísticos de um conjunto de dados. Mais especificamente, com a *média*, *variância*, *desvio médio* e *desvio padrão*. Esses parâmetros são conhecidos como medidas de posição (a média) e dispersão (variância e desvio padrão).

Na maioria dos casos, a melhor estimativa para o valor esperado de uma grandeza (ou esperança, como também é conhecida) que varia aleatoriamente⁴ dentre várias observações independentes e realizadas nas mesmas condições experimentais é a *média aritmética*.

Matematicamente, a partir de um conjunto de dados (x_1, x_2, \dots, x_n) sendo x_n a n ésima observação da grandeza x , a média é definida como (Eq.3.1):

<i>Definição de média</i>	$\bar{x} = \frac{1}{n} \sum x_i \quad (3.1)$
---------------------------	----------------------------------------------

Porém, resumir um conjunto de dados por meio de apenas um valor esconde toda a informação sobre a variabilidade dos valores, não permitindo conhecer a qualidade dos resultados experimentais. São necessárias mais informações para melhores resultados.

Suponha, por exemplo, que um pesquisador, interessado em conhecer a tensão de ruptura de um material desenvolvido em seu centro de pesquisa, resolveu enviar cinco amostras do mesmo material para três laboratórios distintos (A, B e C) para a realização das medidas utilizando o *mesmo procedimento*. Os resultados são mostrados abaixo (Tab. 3.1):

Tab. 3.1: Ensaio de tensão de ruptura (σ) de cinco amostras do mesmo material realizadas em três laboratórios. As unidades estão em MPa.

Laboratório	1	2	3	4	5	Média
A	400	200	100	300	500	300
B	350	600	300	250	100	300
C	305	300	295	300	300	300

Pode-se observar que os resultados possuem a mesma média, mas nada informam sobre a variabilidade dos dados. Por exemplo, enquanto o laboratório C apresenta os valores com a menor dispersão (não variaram mais do que cinco unidades em relação à média), os laboratórios A e B variaram acima de 150 unidades. Desta forma, é necessário utilizar uma medida que sumarie a variabilidade de uma série de valores e que nos permita analisar e comparar resultados segundo algum critério.

Uma forma conveniente de realizar essa análise é verificar a dispersão dos valores em torno da média. Neste caso, pode-se abordar o problema de duas formas: (a) considerar o total

⁴ Grandeza que varia aleatoriamente dentre de um conjunto de dados e que obedece uma distribuição de probabilidades

dos desvios em valor absoluto ou (b) considerar o total dos quadrados dos desvios. Desta forma, teríamos para o laboratório C (Tab.3.2):

Tab.3.2: Abordagens para o cálculo de dispersão dos valores de tensão encontrados para o laboratório C

Abordagem	Fórmula	Resultado
(a)	$\sum_{i=1}^5 x_i - \bar{x} $	5+0+5+0+0 = 10
(b)	$\sum_{i=1}^5 (x_i - \bar{x})^2$	25+0+25+0+0 = 50

Entretanto, como nem sempre dois conjuntos de dados são representados pelo mesmo número de observações, a comparação entre os dois conjuntos pode ser prejudicada. Uma forma de contornar esse problema é exprimir os resultados obtidos para as abordagens (a) e (b) como médias. Desta forma, definem-se *desvio médio* (Eq. 3.2), e *variância* (Eq. 3.3) da seguinte forma:

Definição de desvio médio	$DM = \frac{1}{n} \sum_{i=1}^n x_i - \bar{x} $	(3.2)
Definição de variância	$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$	(3.3)

Sendo a variância uma medida que expressa um desvio quadrático médio, costuma-se utilizar o *desvio padrão* (σ), que é a raiz quadrada da variância (Eq. 3.4):

Definição de desvio padrão	$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$	(3.4)
-----------------------------------	--------------------------------------------------------------	-------

Quando se trata de amostras extraídas de uma pequena população de um conjunto de dados ou quando o conjunto de dados possui pequena quantidade de observações, os cálculos da variância e do desvio padrão sofrem uma pequena alteração, chamada de correção amostral. Neste caso, passa-se a designar variância amostral (s^2) e desvio padrão amostral (s) (Eqs. 3.5 e 3.6):

Definição de variância amostral	$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$	(3.5)
Definição de desvio padrão amostral	$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$	(3.6)

Na prática, a correção amostral tem efeito para conjuntos de até aproximadamente 30 observações. Acima desse valor, a correção não modificará o resultado final. Durante o curso trabalharemos com as definições amostrais.

Voltando ao exemplo da Tab. 3.1, os valores do desvio padrão amostral para cada conjunto de dados realizados em cada laboratório, utilizando a Eq. 3.6, seriam 158, 183 e 4, respectivamente. De fato, o laboratório C apresentou os resultados com menor dispersão.

Sabemos agora determinar a partir de n observações o desvio padrão de uma medida, isto é, sabemos estimar a partir da análise de n observações o erro que teríamos, com uma dada probabilidade, caso houvésemos realizado uma única determinação.

Entretanto, se realizarmos n determinações, um valor mais preciso da média e, conseqüentemente, uma dispersão mais precisa do conjunto de dados seriam alcançados. Com esse propósito, pode-se realizar vários conjuntos de n determinações, e então calcular os valores das respectivas médias e em seguida a média das médias e o desvio padrão da média das médias seria mais preciso. Em outras palavras, quanto maior o número de observações n , menor será o desvio padrão da média, ou seja, maior a precisão do resultado. Esta forma de cálculo do desvio padrão da média pode ser visto com uma forma de se ter um valor normalizado do desvio padrão amostral pelo número de determinações.

Definição de desvio padrão da média

$$d_m = \frac{s}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.7)$$

Outro ponto importante do ponto de vista prático é reconhecer que o equipamento de medição também possui erros inerentes à sua limitação na capacidade de medir. Se o desvio calculado para um conjunto de dados for *menor* que a incerteza do equipamento (geralmente obtida pela metade da menor divisão da escala) a incerteza do equipamento será a principal fonte de incerteza nas medidas. Entretanto, se o desvio calculado for *maior* que a incerteza do equipamento, a dispersão do conjunto de dados é que será a principal fonte de incerteza. Assim uma forma conveniente de trabalhar com ambas as incertezas é utilizando o chamado “desvio total” (Δx_{total}):

Definição total

$$\Delta x_{total} = \sqrt{d_m^2 + d_{inst}^2} \quad (3.8)$$

sendo d_m é o desvio padrão da média ($d = s/\sqrt{n}$) e d_{inst} o desvio do instrumento (que pode ser obtido no manual do instrumento. Entretanto, quando a informação não estiver disponível, pode-se adotar como sendo metade da menor divisão do instrumento. Neste caso, o desvio total é um valor que contém a contribuição dos desvios do instrumento e da dispersão do conjunto de dados.

3.2 Propagação de incertezas

Como visto na seção 1, existem as grandezas de base e as grandezas derivadas. Discutiuse que os resultados experimentais são valores aproximados do valor real e que as medidas de posição e dispersão, para os casos das incertezas do Tipo A são fundamentais para representar o conjunto de dados. A grande questão é saber como representar corretamente um valor de uma grandeza derivada quando ela é constituída de operações matemáticas que envolvem valores incertos. Um exemplo que ilustra o problema é calcular o deslocamento de um carro numa estrada que viaja a certa velocidade constante (e incerta) a partir de um tempo que se conhece com imprecisão (Fig. 3.1).

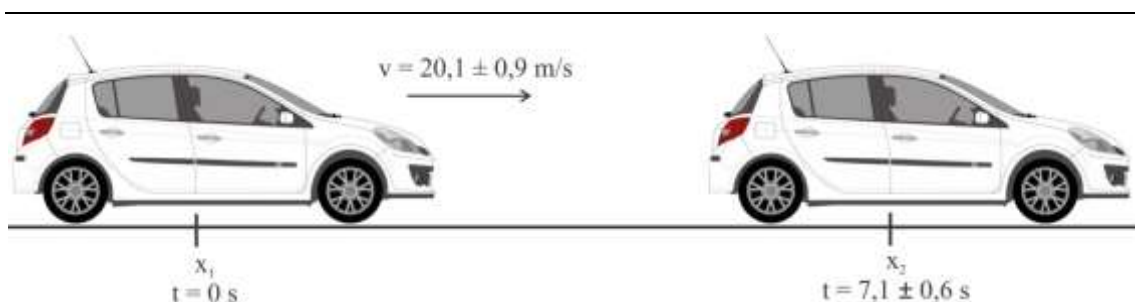


Figura 3.1. Problema de se conhecer o deslocamento do carro (Δx). O cálculo do deslocamento é dado pela relação $\Delta x = vt = (20,1 \pm 0,9)x(7,1 \pm 0,6)m$. Observe que a velocidade e o tempo possuem incertezas. Como calcular Δx neste caso?

O mesmo problema se estenderia para outras grandezas, como aceleração, força, densidade, área, volume, etc., que são originadas por operações matemáticas entre valores e que, geralmente, possuem incertezas. Uma forma de solucionar esse problema é *propagar a incerteza* até o valor final. Isto significa que cada valor da incerteza das grandezas envolvidas na operação matemática é levado em consideração, e que a incerteza do número final possui as componentes relativas de cada incerteza dos números calculadas. No caso da Fig. 3.1, a incerteza do deslocamento terá componentes das incertezas da velocidade e do tempo.

A abordagem matemática utilizada na propagação de incertezas foge do propósito da apostila e não deduziremos as equações envolvidas. Entretanto, as deduções envolvem o desenvolvimento de derivadas parciais. Para isso, se considera uma grandeza Y como função de outras grandezas a, b, c, \dots, k que possuem desvios padrões $\Delta a, \Delta b, \Delta c, \dots, \Delta k$ (Eq.3.9):

$$Y = f(a, b, c, \dots, k) \quad \text{Eq. 3.9}$$

O desvio padrão da grandeza Y (s_Y) é dada pela relação (Eq.3.10):

$$\Delta Y = \pm \left[\left(\frac{\partial Y}{\partial a} \right)^2 \Delta a^2 + \left(\frac{\partial Y}{\partial b} \right)^2 \Delta b^2 + \left(\frac{\partial Y}{\partial c} \right)^2 \Delta c^2 + \dots + \left(\frac{\partial Y}{\partial k} \right)^2 \Delta k^2 \right]^{1/2} \quad \text{Eq. 3.10}$$

Sendo $\left(\frac{\partial Y}{\partial a} \right)$ a derivada parcial da grandeza Y em função da grandeza a .

A Tab. 3.3 resume as fórmulas para o cálculo das propagações das incertezas para as quatro operações fundamentais. Considere a função $Y(a, b)$, sendo a e b grandezas com as médias \bar{a} e \bar{b} e com os respectivos desvios Δa e Δb ($a = \bar{a} \pm \Delta a$; $b = \bar{b} \pm \Delta b$).

Tab.3.3 Fórmulas para o cálculo das incertezas para as quatro operações fundamentais.

Operação	Cálculo de s_Y
Adição ($Y = a + b \rightarrow Y = \bar{a} + \bar{b} \pm \Delta Y$)	$\Delta Y = \pm \sqrt{\Delta a^2 + \Delta b^2}$
Subtração ($Y = a - b \rightarrow Y = \bar{a} - \bar{b} \pm \Delta Y$)	$\Delta Y = \pm \sqrt{\Delta a^2 + \Delta b^2}$
Produto ($Y = a \cdot b \rightarrow Y = \bar{a} \cdot \bar{b} \pm \Delta Y$)	$\Delta Y = \pm \sqrt{\bar{b}^2 \cdot \Delta a^2 + \bar{a}^2 \cdot \Delta b^2}$

Quociente ($Y = \frac{a}{b} \rightarrow Y = \frac{\bar{a}}{\bar{b}} \pm \Delta Y$)

$$\Delta Y = \pm \sqrt{\frac{\Delta a^2}{\bar{b}^2} + \frac{a^2}{\bar{b}^4} \Delta b^2}$$

No exemplo dado pela Fig. 3.1, tem-se que $\Delta x = vt = (20,1 \pm 0,9)x(7,1 \pm 0,6)$. A resposta do deslocamento seria $\Delta x = vt \pm \Delta Y_{\Delta x}$. Seguindo com o cálculo de $v \cdot t = 20,1x7,1 = 142,7$. E para a propagação da incerteza para o produto:

$$\Delta Y_{\Delta x} = \pm \sqrt{\bar{t}^2 s_v^2 + \bar{v}^2 s_t^2} = \sqrt{7,1^2 0,9^2 + 20,1^2 0,6^2} = 13,64$$

Portanto, $\Delta x = (1,4 \pm 0,1)x10^2 m$

Convém se atentar no uso de algarismos significativos. De forma geral, a incerteza é representada com apenas um algarismo significativo e o valor medido (ou calculado) representado de acordo com o número de casas decimais da incerteza obtida.

3.3 Erro Percentual Relativo (E%)

O erro percentual relativo (E%) é aquele em que se compara relativamente um valor medido ou obtido experimentalmente (Y_{exp}) em relação a um valor de referência (Y_{ref}). Este valor de referência é normalmente relacionado a valores teóricos e/ou previamente conhecidos. O erro percentual relativo indica o quão diferente é o valor experimental do valor de referência e tem por objetivo, por exemplo, avaliar a eficiência de determinado processo ou experimento e é dado por:

$$E\% = \left(\frac{|Y_{exp} - Y_{ref}|}{Y_{ref}} \right) \cdot 100\%$$

4. REPRESENTAÇÃO GRÁFICA E REGRESSÃO LINEAR

Um conjunto de dados pode ser representado em tabelas e gráficos. Entretanto, em determinadas ocasiões, a abordagem gráfica, além de fornecer rapidamente o comportamento do sistema de forma visual, permite obter maiores informações do objeto em estudo. Veja, por exemplo, os dados abaixo:

Tabela 4.1: Despesas de uma empresa.

Despesa	Valor em R\$
Comissões dos vendedores	40.000
Compra de matéria-prima	300.000
Energia/Água e telefone	10.000
Gastos com vendedores (combustível/hospedagem)	65.000

Impostos	60.000
Manutenção de computadores	3.000
Salários	50.000

As informações representadas desta forma apenas descrevem os valores de cada despesa. Entretanto, dificultam a análise administrativa uma vez que será necessário realizar cálculos posteriores com o objetivo de determinar, por exemplo, o porcentual que representa cada despesa. Uma forma simples de representar esses dados seria por meio de gráfico (Fig. 4.1).

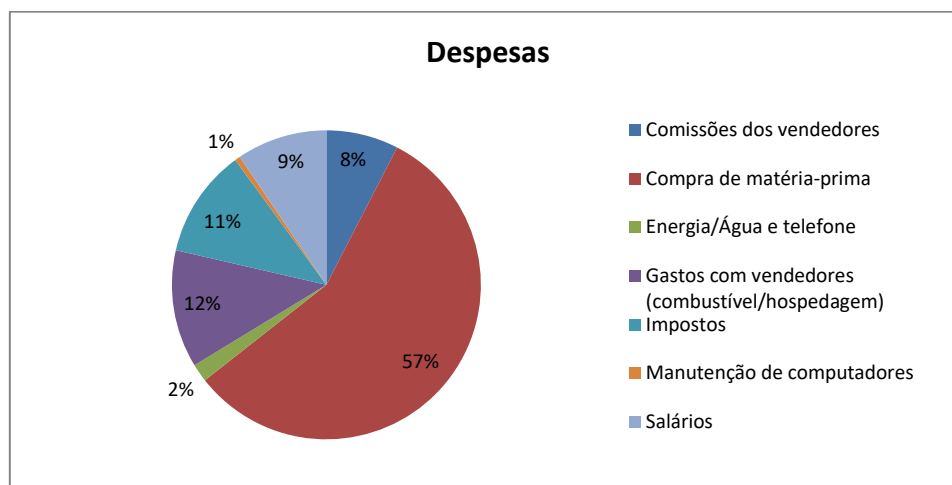


Fig. 4.1: Representação gráfica dos dados da Tab.4.1. Neste caso, os dados estão representados em um gráfico de “pizza”.

Como pode ser observado pelo gráfico da Fig. 4.1, a análise gráfica, neste caso, permite obter rapidamente melhores informações a respeito do sistema em estudo (despesas de uma empresa) quando comparado à análise da tabela.

A análise gráfica também permite obter informações que estão implícitas no conjunto de dados. Como exemplo ilustrativo, imagine uma empresa alimentícia faz o tratamento dos efluentes gerados de seu processo. A cada três horas, coletam-se os dados de volume de efluente que está sendo gerado em seu processo. Os dados estão representados na Tab.4.2.

Tab. 4.2: Valores de volume de efluente (m^3) em função da hora. A hora em “0” corresponde ao início das medições.

Hora	0	3	6	9	12	15	18	21
Volume de efluente (m^3)	50	100	175	225	300	350	400	465

Uma forma conveniente de representar esses dados é por meio do “gráfico de dispersão”, que representa como os dados estão dispersos no plano cartesiano. Entretanto, temos que levar em consideração alguns pontos importantes para antes de construir o gráfico.

Primeiramente, temos que conhecer qual é a variável dependente e a variável independente. Variável independente, como o próprio nome diz, *não depende* de nenhuma variável. Neste caso, o tempo em que foram coletados os valores de volume de efluente não depende de nenhuma outra variável e, portanto, é independente. Por outro lado, o volume de efluente foi coletado *em função* do tempo e, desta forma, ela *depende de outra grandeza*, que seria o tempo. Então, levando em consideração que a variável independente é representada no

eixo das abscissas (eixo x) e a variável dependente no eixo das ordenadas (eixo y), é possível construir o gráfico mostrado na Fig.4.2a.

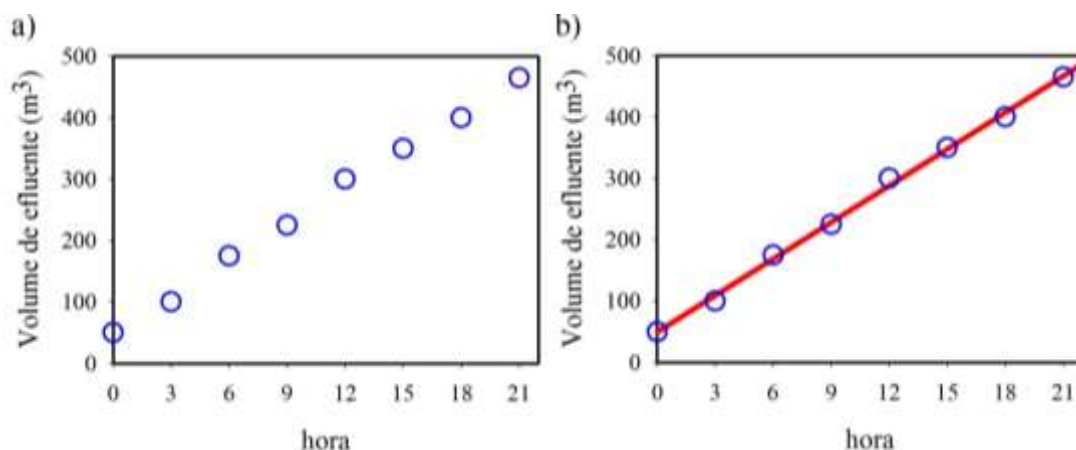


Fig.4.2: (a) À esquerda: Gráfico de dispersão dos valores de volume de efluente coletados por hora. Os círculos em azul representam os dados coletados de volume em função da hora. (b) À direita: Gráfico de dispersão contendo uma possível reta de regressão.

Por meio da análise do gráfico da Fig.4.2a, é possível observar que o volume de efluente cresce linearmente em função do tempo. De fato, pode-se traçar uma reta (dentre várias outras) que seria uma função do tipo $f(x) = ax + b$; ($a =$ coeficiente angular; $b =$ coeficiente linear), ligando todos os pontos do gráfico (Fig.4b).

Uma análise mais criteriosa e profunda do conjunto de dados possibilita extrair uma informação importante sobre uma característica do sistema: a *vazão de efluente*. A vazão (V) é definida como $V(m^3/t) = \text{volume } (m^3)/\text{tempo } (h)$. Pelo método gráfico, ela pode ser obtida facilmente por meio do coeficiente angular da reta. Ou seja, o conjunto de dados pode ser descrito por uma função linear, sendo possível agora estimar a quantidade de efluente coletado para qualquer tempo considerando um valor de *vazão de efluente* constante.

Neste ponto, chega-se a um importante problema: qual seria a melhor reta que descreve o comportamento linear de um conjunto de dados? Em outras palavras, quais seriam os melhores valores dos coeficientes angular (a) e linear (b) que melhor representariam o conjunto de dados da forma $f(x) = ax + b$, que fornecessem dados seguros a respeito daquilo que estamos deduzindo? Uma abordagem gráfica e matemática para resolver este problema serão apresentadas nas próximas seções.

4.1 Método Gráfico: Utilizando papel milimetrado.

Um dos papéis mais utilizados para a construção de gráficos é o de escala milimetrada. Ele possui as escalas principais a cada 1 mm e escalas secundárias a cada 10 mm. As escalas principais são encontradas de acordo com o conjunto dos dados e depende das dimensões do papel, que fornecerá em que orientação se deve plotar o gráfico no papel, ou seja, na forma retrato ou paisagem. Convém lembrar que em um gráfico no plano cartesiano, a variável independente (eixo das abscissas) está sempre orientada na horizontal e a variável dependente (eixo das ordenadas) no eixo vertical. Por exemplo, considere o conjunto de dados abaixo (Tab. 4.3), no qual a grandeza y foi obtida em função de variável independente x :

Tab. 4.3: Dados de y em função de x .

y	50	100	150	225	300
x	3	6	9	12	15

Imagine que a divisão das escalas no papel compreende a dimensão de 180 mm x 280 mm.

Primeiramente, devemos verificar qual a amplitude de valores das escalas. Como é comum construir os gráficos a partir da origem (0,0)⁵, realizamos os cálculos para y e para x :⁶

Amplitude no eixo $y = 300 - 0 = 300$.

Amplitude no eixo $x = 15 - 0 = 15$.

O processo de escolha da orientação do papel é bem intuitivo: Se a amplitude na escala x é maior, devemos trabalhar com o papel na orientação paisagem; se a amplitude na escala y é maior, devemos trabalhar com o papel na orientação retrato. No exemplo acima, teríamos que utilizar o papel na orientação retrato.

Após encontrarmos a orientação, devemos calcular os valores para as escalas principais utilizando uma “regra de três” entre as dimensões e as escalas. Utilizaremos aqui um exemplo com um papel de 280 mm x 180 mm para facilitar o entendimento, no entanto, este raciocínio é válido para papéis milimetrados de quaisquer dimensões:

280 mm equivalem a 300 (unidades de grandeza)

10 mm equivalem a **10,71** (unidades de grandeza)

Arredondando esse valor, encontra-se que a escala principal será de 11 (unidades de grandeza) para o eixo x . É importante notar que o arredondamento deve ser sempre para maior valor, fato que serve apenas para que todos os dados experimentais caibam no eixo.

Para encontrar o intervalo na escala secundária, realiza-se procedimento similar:

10 mm equivalem a 11 (unidades de grandeza)

1 mm equivale a **1,1** (unidade de grandeza)

Não é necessário que marque no eixo do gráfico os intervalos na escala secundária, apenas os da escala principal.

Para encontrar a variação no eixo y , utiliza-se o mesmo procedimento descrito anteriormente.

180 mm equivalem a 15 (unidades de grandeza)

⁵ Geralmente constrói-se o gráfico a partir da origem. Entretanto, convém ressaltar que, ao realizar este procedimento, não significa que exista a coordenada (0,0) no conjunto de dados experimentais e não se deve colocar um ponto na coordenada a menos que ela tenha sido resultado de dados experimentais.

⁶ Não é obrigatório que o gráfico seja construído a partir da origem.

10 mm equivalem a **0,83** (unidades de grandeza) = **1** (arredondando).

Para a escala secundária:

10 mm equivalem a 1 (unidades de grandeza)

1 mm equivale a **0,1** (unidade de grandeza)

Uma vez obtidas as escalas, constrói-se o gráfico marcando cada coordenada no gráfico de acordo com os dados fornecidos.

Nem sempre o conjunto de dados obtidos apresenta números inteiros. Quando isso não acontece, para facilitar os cálculos, pode multiplicar o eixo por 10, 100 ou mil (até mais se necessário). Neste caso, deve ser mencionada no eixo em questão a adequação adotada (se multiplicou por 10, colocar no eixo “x 10”, etc.).

Quando a **função é linear ou linearizada, escrita da forma $y = ax + b$** , é possível obter o coeficiente angular (a) e linear (b) de uma reta que se ajuste no conjunto de dados. Para isso, escolha duas coordenadas da **reta** como, por exemplo, a primeira coordenada entre o primeiro e o segundo dado experimental (x_1, y_1) e a segunda entre os dois últimos (x_2, y_2).

$$a = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \quad \text{Eq.4.1}$$

O coeficiente linear (b) é obtido por meio da interseção da curva com o eixo y , ou seja, na coordenada $(0; b)$.

Embora essa abordagem seja válida para calcular a e b , ela pode conter erros devido à inexatidão da reta traçada no gráfico, pois a reta traçada pode não ser a que melhor descreve os dados. O método que possibilita determinar os coeficientes da melhor reta que se ajuste ao conjunto de dados experimentais é o Método dos Mínimos Quadrados (MMQ), que será discutido adiante.

Para o uso do papel milimetrado pode ocorrer casos nos quais o seu conjunto de dados pode ser representado por funções do tipo:

$$f(x) = AX^2$$

sendo A constante. Em casos assim, para linearizar essa função para a forma linear **$f(x) = ax + b$** , chama-se o **X^2** de **x** . Logo, os valores do eixo das abscissas serão **$x = X^2$** , e o gráfico poderá ser plotado em um papel milimetrado de forma que se obtenha um gráfico que pode ser descrito por uma função linear do tipo **$f(x) = ax + b$** .

4.2 Método Gráfico: Utilizando papel em escala logarítmica.

Existem dois tipos de papel em escala logarítmica. O mono-log, em que apenas um dos eixos está em escala logarítmica e o di-log, em que ambas as escalas são logarítmicas.

Os papéis mono-log e di-log são ideais para construir gráficos de funções logarítmicas ou linearizadas de forma a se tornarem logarítmicas. Por exemplo, se a função $f(x) = AE^x$ (sendo A e E duas constantes) for plotada em papel milimetrado sem sofrer alguma transformação ela não resultará numa reta. Entretanto, utilizando o logarítmico em ambos os lados da função teríamos:

$$\log(f) = \log A + x \log(E) \quad \text{Eq.4.2}$$

A função passa a ser linear plotando o gráfico na forma de $\log(f)$ versus x , com coeficiente linear igual a $\log A$ e angular igual a $\log(E)$. Este é o chamado processo de linearização de uma função, utilizando papel milimetrado para plotar os dados e extrair os valores das constantes A e E .

Por outro lado, como esta função tem uma variável que é logarítmica (variável dependente) e outra em escala linear (variável independente), poderia se plotar estes dados diretamente em um papel mono-log. Plotando os dados experimentais no gráfico mono-log (observe que a escala é logarítmica e, portanto, não é necessário aplicar logaritmo nos valores de $f(x)$) os dados resultariam em um comportamento linear.

Considere outro exemplo. A função do tipo $f(x) = Ce^{dx}$ (sendo C e d constantes) pode ser linearizada aplicando logaritmo em ambos os lados da função. Como a função possui um exponencial em e , é conveniente escrever, nestes casos, o logaritmo natural ao invés do logaritmo em base 10. Neste caso, a linearização ficaria:

$$\ln f(x) = \ln C + d \ln(e)x = \ln C + dx \quad \text{Eq.4.3}$$

sendo $\ln C$ o coeficiente linear e d o coeficiente angular. Novamente, o papel mono-log seria o mais adequado.

Para calcular o coeficiente angular, basta utilizar a Eq. 4.18 se a escala utilizada for logaritmo em base 10 e Eq. 4.19 se for neperiano.

$$a = \frac{\Delta y}{\Delta x} = \frac{\log y_2 - \log y_1}{x_2 - x_1} \quad \text{Eq.4.4}$$

$$a = \frac{\Delta y}{\Delta x} = \frac{\ln y_2 - \ln y_1}{x_2 - x_1} \quad \text{Eq.4.5}$$

O coeficiente linear pode ser encontrado quando $x = 0$, ou seja, no ponto onde a reta intercepta o eixo y .

O uso do papel di-log se faz necessário em funções do tipo $f(x) = Ax^n$ (A e n são constantes), pois neste caso esta função tem as duas variáveis logarítmicas. Linearizando a função aplicando log em ambos os lados da função:

$$\log f(x) = \log A + n \log(x) \quad \text{Eq.4.6}$$

em que $\log A$ é o coeficiente linear e n o coeficiente angular. A função mostrada na Eq. 4.20 possui comportamento linear se os dados forem plotados na forma $\log f(x)$ versus $\log(x)$.

Neste caso, o coeficiente angular (para escala log em base 10 e em base e) é obtido utilizando a relação Eq. 4.21 e Eq. 4.22, respectivamente.

$$a = \frac{\Delta y}{\Delta x} = \frac{\log y_2 - \log y_1}{\log x_2 - \log x_1} \quad \text{Eq.4.7}$$

$$a = \frac{\Delta y}{\Delta x} = \frac{\ln y_2 - \ln y_1}{\ln x_2 - \ln x_1} \quad \text{Eq.4.8}$$

O coeficiente linear pode ser obtido escolhendo uma coordenada da **reta** no plano do gráfico e substituindo os valores de $f(x)$ e n na equação da reta linearizada, determinando assim o valor do coeficiente linear pela resolução da equação da reta. Isto se deve ao fato de que em papel di-log não se tem o valor zero em ambos os eixos.

No papel dilog (assim como no papel monolog), as escalas principais estão separadas em décadas (Fig. 4.3). Isto significa que a escala inicia em 10^N (sendo N um número inteiro positivo, negativo ou nulo) e termina em 10^{N+1} . Por exemplo, se iniciar a escala com $N=0$ (que representa o número 1, uma vez que $10^0 = 1$), a próxima será 10 ($10^{0+1} = 10^1 = 10$). Sendo assim, as escalas dentro das décadas devem ser preenchidas com números inteiros de 2 a 9, que são os múltiplos de 10^N .

Para encontrar as escalas principais no papel, você perceberá que os traços começam a ficar cada vez mais próximos, até que há um salto. O primeiro será o valor 10^N , e o último traço antes do salto será 10^{N+1} (Fig. 4.3).

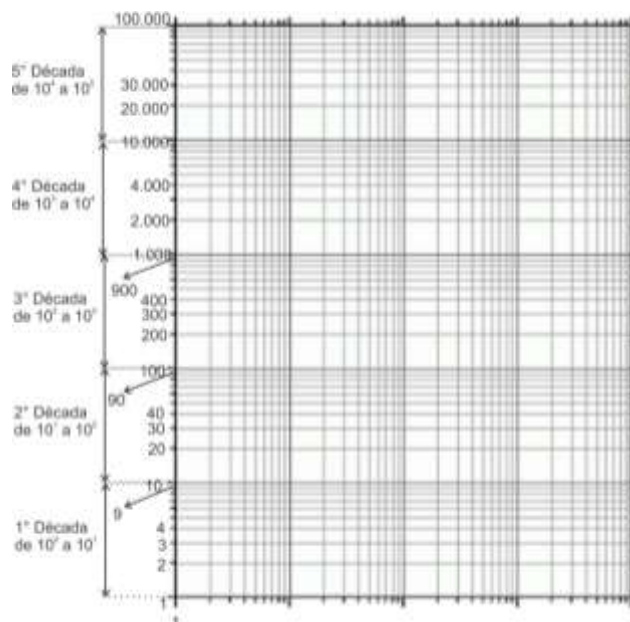


Fig. 4.3: Exemplo de papel em escala logarítmica (di-log) e como as décadas podem ser representadas no papel.

4.3 Regressão linear pelo Método dos Mínimos Quadrados (MMQ).

Uma forma simples de obter os coeficientes angular e linear da *melhor reta de regressão* que descreve o comportamento do sistema é pelo Método dos Mínimos Quadrados (MMQ). Este método baseia-se no fato de que a soma dos desvios ao quadrado, obtidos pela diferença entre o valor experimental e do modelo, se aproxima de zero.

Para melhor entender como o método se desenvolveu, considere um conjunto de dados experimentais representados na forma $T = \{(x_1, y_1); (x_2, y_2); \dots; (x_n, y_n)\}$. É necessário escrever a função Y na forma linear $f(x) = ax + b$. Assim, a soma dos resíduos ao quadrado (SRQ) seria (Eq.4.1):

$$SRQ = \sum_{i=1}^n (y_i - f(x_i))^2 \quad \text{Eq.4.9}$$

Desenvolvendo a relação, encontra-se:

$$SRQ = \sum_{i=1}^n y_i^2 - 2y_i f(x_i) + f(x_i)^2 \quad \text{Eq.4.10}$$

Como $f(x_i) = ax_i + b$, obtém-se:

$$SRQ = \sum_{i=1}^n y_i^2 - 2y_i(ax_i + b) + (ax_i + b)^2 \quad \text{Eq.4.11}$$

Desenvolvendo a relação matemática:

$$SRQ = \sum_{i=1}^n (y_i^2 - 2y_i ax_i - 2y_i b + a^2 x_i^2 + 2ax_i b + b^2) \quad \text{Eq.4.12}$$

Como desejamos encontrar os valores de a e b que minimizam a relação dada pela Eq.4.4., podemos encontrar a derivada parcial de SRQ em relação a cada coeficiente e igualarmos a zero. Desta forma, tem-se que:

$$\frac{\partial SRQ}{\partial b} = \frac{\partial SRQ}{\partial a} = 0 \quad \text{Eq.4.13}$$

Desenvolvendo para a relação $\frac{\partial SRQ}{\partial b}$:

$$\begin{aligned} \frac{\partial SRQ}{\partial b} &= \sum_{i=1}^n (-2y_i + 2ax_i + 2b) \\ \frac{\partial SRQ}{\partial b} &= 2 \left(\sum_{i=1}^n -y_i + a \sum_{i=1}^n x_i + b \sum_{i=1}^n 1 \right) \end{aligned} \quad \text{Eq.4.14}$$

Para simplificar a Eq.4.6, podemos utilizar as relações:

$$X = \sum_{i=1}^n x_i \quad \text{Eq.4.15}$$

$$Y = \sum_{i=1}^n y_i \quad \text{Eq.4.16}$$

Sendo que X e Y representam a soma dos valores experimentais em x e y , respectivamente.

Realizando a substituição das Eq.4.7 e Eq.4.8 na Eq.4.6, obtém-se:

$$\frac{\partial SRQ}{\partial b} = 2(aX - Y + bn) \quad \text{Eq.4.17}$$

Utilizando a Eq.4.5, a relação que calcula o coeficiente b é:

$$\frac{\partial SRQ}{\partial b} = 2(aX - Y + bn) = 0$$

$$b = \frac{Y - aX}{n} \quad \text{Eq.4.18}$$

Entretanto, para determinar b é necessário encontrar o valor de a . O coeficiente angular pode ser encontrado pela outra derivada parcial da Eq. 4.5:

$$\begin{aligned} \frac{\partial SRQ}{\partial a} &= \sum_{i=1}^n (-2y_i x_i + 2ax_i^2 + 2x_i b) \\ \frac{\partial SRQ}{\partial a} &= 2 \left(-\sum_{i=1}^n y_i x_i + a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i \right) \\ \frac{\partial SRQ}{\partial a} &= 2 \left(-\sum_{i=1}^n y_i x_i + a \sum_{i=1}^n x_i^2 + bX \right) \quad \text{Eq.4.19} \end{aligned}$$

Igualando a relação a zero e substituindo Eq. 4.10 na Eq.4.11:

$$\begin{aligned} 2 \left(-\sum_{i=1}^n y_i x_i + a \sum_{i=1}^n x_i^2 + bX \right) &= 0 \\ \left(-\sum_{i=1}^n y_i x_i + a \sum_{i=1}^n x_i^2 + bX \right) &= 0 \end{aligned}$$

$$\left(-\sum_{i=1}^n y_i x_i + a \sum_{i=1}^n x_i^2 + \frac{Y - aX}{n} X\right) = 0$$

$$\left(-n \sum_{i=1}^n y_i x_i + an \sum_{i=1}^n x_i^2 + YX - aX^2\right) = 0$$

$$a \left(n \sum_{i=1}^n x_i^2 - X^2\right) - n \sum_{i=1}^n y_i x_i + YX = 0$$

$$a = \frac{n \sum_{i=1}^n y_i x_i - YX}{n \sum_{i=1}^n x_i^2 - X^2} \quad \text{Eq.4.20}$$

Utilizando a mesma abordagem matemática, é possível calcular as constantes de outra forma. Ao invés de realizar a substituição da Eq.4.10 na Eq.4.11, pode-se obter um sistema de duas equações e duas variáveis que pode ser facilmente resolvido.

$$\begin{cases} a \sum_{i=1}^n x_i - \sum_{i=1}^n y_i + b \sum_{i=1}^n 1 = 0 \\ a \sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i x_i + b \sum_{i=1}^n x_i = 0 \end{cases} \quad \text{Sist.4.21}$$

Agora, serão calculados os melhores coeficientes da reta de regressão para o problema da vazão de efluentes apresentado acima. Primeiramente, são calculados os seguintes somatórios, de acordo com o sistema Sist.4.1:

$$\begin{aligned} \sum_{i=1}^n x_i &= 84 & \sum_{i=1}^n y_i x_i &= 29190 \\ \sum_{i=1}^n y_i &= 2065 & \sum_{i=1}^n x_i^2 &= 1260 \end{aligned}$$

Substituindo os valores no sistema Sist.4.1 e o resolvendo, são encontrados os coeficientes $a = 19,86$ e $b = 49,58$, sendo o coeficiente a a vazão do efluente. Desta forma, a função linear que melhor descreve o sistema é dada pela relação $f(x) = 19,86x + 49,58$.

É importante notar que a abordagem do MMQ só se aplica adequadamente para sistemas que se *comportam linearmente*. Para uma função que não seja linear, o MMQ não se aplica. Entretanto, é possível *linearizar* uma função, a fim de torná-la apta para utilizar a regressão linear e obter os parâmetros desejados.

Um exemplo de linearização de função é o caso da queda livre de um corpo. A função teórica que descreve a posição de um corpo em queda livre, considerando que o corpo partiu do repouso, é:

$$s = s_0 - \frac{a}{2}t^2 \quad \text{Eq.4.22}$$

pode-se notar que ela é quadrática. Entretanto, podemos reescrever a Eq.4.13 de forma a torná-la linear, com a forma de:

$$y = y_0 + cx \quad \text{Eq.4.23}$$

Sendo $s = y$, $s_0 = y_0$, $-a/2 = c$ e $t^2 = x$. A função passa a ser linear na forma $y(x)$. Convertendo os dados experimentais de s para y e t^2 para x , pode-se utilizar a abordagem do MMQ para calcular os coeficientes y_0 e c e, por consequência, s_0 e a .

Coefficiente de determinação (r^2)

O coeficiente de determinação, também chamado de r^2 , é uma medida do ajustamento de um modelo estatístico linear, chamada Regressão linear, em relação aos valores observados para um determinado conjunto de dados. O valor de r^2 pode variar entre 0 e 1 e indica o quanto o modelo consegue se ajustar aos valores observados. Isto é, quanto maior o valor de r^2 , mais explicativo é o modelo e melhor ele se ajusta aos dados observados, em outras palavras, quanto mais próximo de 1 mais os dados podem ser descrito por um modelo estatístico linear. A equação que descreve este valor é dada por:

$$r^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

4.1.1 MMQ com incerteza em y .

Quando dados experimentais são coletados em replicata (mais de um experimento repetido nas mesmas condições), sempre existirá uma incerteza⁷ na variável dependente, ou seja, em y . Neste caso, existem duas possibilidades: a incerteza em y é igual para todo o conjunto de dados, ou seja, para cada valor de x existirá sempre a mesma incerteza (como no caso de se medir o espaço em função do tempo apenas uma vez, em que a incerteza no espaço, para cada valor de tempo, será a incerteza do instrumento); ou ela é diferente para cada valor de x . Para ambos os casos, as incertezas em y se propagam para os coeficientes angular e linear.

Quando o erro em y é constante, ou seja:

$$\begin{array}{c} \hline x_1, y_1 \pm \Delta y \\ x_2, y_2 \pm \Delta y \\ x_3, y_3 \pm \Delta y \\ \dots, \dots \\ x_n, y_n \pm \Delta y \\ \hline \end{array}$$

Os coeficientes da relação linear da equação $y = ax + b$ devem ser obtidos utilizando as relações:

$$a = \frac{1}{\beta} \left(N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i \right) \quad \text{Sist.4.2}$$

⁷ A incerteza, neste caso, é o desvio total, que é a combinação do desvio estatístico e instrumental.

$$b = \frac{1}{\beta} \left(\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i \sum_{i=1}^N x_i y_i \right)$$

$$\beta = N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2$$

Desta forma, os desvios nos coeficientes angular (Δa) e linear (Δb) são obtidos utilizando as seguintes equações:

$$(\Delta a)^2 = \frac{N}{\beta} (\Delta y)^2 \quad \text{Eq.4.15}$$

$$(\Delta b)^2 = \frac{(\Delta y)^2}{\beta} \sum_{i=1}^N x_i^2 \quad \text{Eq.4.16}$$

Por outro lado, quando os desvios em y são diferentes para cada medida:

$$\begin{array}{c} \overline{x_1, y_1 \pm \Delta Y_1} \\ x_2, y_2 \pm \Delta Y_2 \\ x_3, y_3 \pm \Delta Y_3 \\ \dots, \dots \\ \overline{x_n, y_n \pm \Delta Y_n} \end{array}$$

Os coeficientes da regressão linear de $y = ax + b$ devem ser obtidos utilizando as relações:

$$a = \frac{1}{\beta} \left(\sum_{i=1}^N \frac{1}{\Delta y_i^2} \sum_{i=1}^N \frac{x_i y_i}{\Delta y_i^2} - \sum_{i=1}^N \frac{x_i}{\Delta y_i^2} \sum_{i=1}^N \frac{y_i}{\Delta y_i^2} \right)$$

$$b = \frac{1}{\beta} \left(\sum_{i=1}^N \frac{x_i^2}{\Delta y_i^2} \sum_{i=1}^N \frac{y_i}{\Delta y_i^2} - \sum_{i=1}^N \frac{x_i}{\Delta y_i^2} \sum_{i=1}^N \frac{x_i y_i}{\Delta y_i^2} \right) \quad \text{Sist.4.3}$$

$$\beta = \sum_{i=1}^N \frac{1}{\Delta y_i^2} \sum_{i=1}^N \frac{x_i^2}{\Delta y_i^2} - \left(\sum_{i=1}^N \frac{x_i}{\Delta y_i^2} \right)^2$$

E os desvios nos coeficientes angular (Δa) e linear (Δb) são obtidos utilizando as seguintes equações:

$$(\Delta a)^2 = \frac{1}{\beta} \sum_{i=1}^N \frac{1}{\Delta y_i^2} \quad \text{Eq.4.17}$$

$$(\Delta b)^2 = \frac{1}{\beta} \sum_{i=1}^N \frac{x_i^2}{\Delta y_i^2}$$

Eq.4.18

5. ROTEIROS PARA AS AULAS EXPERIMENTAIS

5.1 Elaboração dos relatórios

Durante o curso de Laboratório Física I, no final de cada atividade prática, é solicitado a cada grupo elaborar um relatório das atividades desenvolvidas, que deverá ser entregue dentro do prazo estipulado pelo professor.

O relatório consiste em sintetizar e agrupar os dados obtidos, contextualizando e discutindo os resultados dentro dos objetivos da prática. Cada seção do relatório é importante, pois traz em si elementos que se ligam, trazendo sentido e fluidez no texto completo, além de oferecer uma apresentação atraente do conjunto dos dados e do trabalho desenvolvido. Desta forma, observar o objetivo de cada parte do relatório é importante para elaborá-lo com qualidade. A seguir, é apresentado um resumo dos componentes que serão exigidos no relatório, bem como a finalidade de cada item e a sua pontuação. Antes de continuar, entretanto, convém ressaltar que, em cada disciplina de natureza experimental da grade do curso de graduação, cada docente poderá exigir um formato de relatório diferenciado. Este apresentado será o formato solicitado na disciplina de Laboratório de Física I.

Tab. 5.1: Elementos exigidos no relatório

Parte externa		
	<i>Capa</i>	Deve conter o nome dos integrantes do grupo que participaram da prática, o curso e o nome do experimento.
Parte Interna		
Pré-texto	<i>Resumo em português</i>	Descrição dos objetivos do trabalho, os resultados e a conclusão. Deve ser apresentado como um parágrafo único.
	<i>Sumário</i>	Lista contendo cada parte do relatório com a referida página inicial que se inicia a seção.
Texto	<i>Introdução</i>	Conter o referencial teórico da prática desenvolvida. Deve possuir corpo de referências que corrobore as informações fornecidas.
	<i>Objetivos</i>	Descrever os objetivos da prática.
	<i>Metodologia</i>	Descrição dos materiais, métodos e equipamentos utilizados, que possibilite a compreensão e interpretação dos resultados, a reprodução do estudo e utilização do método por outros pesquisadores. Não é cópia do roteiro da prática.
	<i>Resultados e discussão</i>	Apresentação pormenorizada dos resultados obtidos, descrevendo-se comparações, comprovações e aplicações teóricas ou práticas.
	<i>Conclusão</i>	Parte final do texto, que deve conter a análise final do trabalho baseada nos resultados apresentados e alinhada com o objetivo do trabalho.
Pós-texto	<i>Referências</i>	Seguir norma ABNT – NBR 6023 – Informação e Documentação – Referências – Elaboração
<p>Capa, formatação e apresentação do texto: 2 pts. Formatação: forma de apresentação das tabelas, gráficos, legendas, figuras, espaçamento, margem, posição do texto e fonte (seguir as normas que estão no site http://www.iq.unesp.br/#!/biblioteca/normalizacao/monografias-e-relatorios/). Apresentação: impresso e grampeado (grampear a lateral esquerda com três grampos, de forma que permita o relatório ser aberto horizontalmente).</p> <p>Para maiores informações com relação ao formato das referências, acesse: http://www.biblioteca.iq.unesp.br/biblio/, em “serviços online” – “normalização”.</p>		