0 Tratamento Estatístico de dados experimentais

Objetivos

- Recordar conceitos associados ao tratamento de dados experimentais;
- Estabelecer fórmulas mais comuns;
- Procedimento gráfico: Ajuste de curvas a dados experimentais;

Esta apresentação destina-se *exclusivamente* a servir como recordação de conhecimentos já aprendidos. Para uma exposição extensiva do assunto, são indicados, entre outros:

- FUNDAMENTOS DA TEORIA DE ERROS, J. H. Vuolo, Ed Edgar Blücher Ltda, 1992.
- TRATAMENTO ESTATÍSTICO DE DADOS EM FÍSICA EXPERIMENTAL, O. A. M. Helene & Vito R.Vanin, Ed Edgar Blücher Ltda, 1981.
- TÓPICOS AVANÇADOS EM TRATAMENTO ESTATÍSTICO DE DADOS EM FÍSICA EXPERIMENTAL Vito R.Vanin & P.Gouffon, LAL-IFUSP, 1996.
- EXPERIMENTAL MEASUREMENTS: PRECISION, ERROR AND TRUTH, N.C.Barford, Addison-Wesley Publishing Co., Inc 1967.
- DATA REDUCTION AND ERROR ANALYSIS FOR THE PHYSICAL SCIENCES, P.R.Bevington, McGraw-Hill Book Co. 1969

Introdução Toda medida experimental resulta afetada de erro. Duas medidas de uma mesma grandeza, feitas nas mesmas condições, não apresentam os mesmos resultados. Quando isto acontece, invariavelmente está subestimada a acurácia do método e/ou do instrumento utilizado. Dado um conjunto de medidas de diâmetro de um prego, praticadas com um paquímetro,

$$0,95\,mm \quad 0,94\,mm \quad 0,96\,mm \quad 0,98\,mm \quad 0,99\,mm \quad 1,03\,mm$$

Qual destas medidas *melhor* representa o diâmetro do prego ?
A resposta: *nenhuma !*

Se acreditamos que os erros que ocorrem na medida são estatísticos, a melhor escolha para o diâmetro procurado será o *valor médio*:

$$\bar{x} = \frac{1}{n}\sum_{1}^{n} x_i \tag{1}$$

onde $\bar{x}$ representa o valor médio, n é o número de medidas e x_i é a i-ésima medida. A qualidade da medida pode ser aferida através do *desvio padrão* ou *desvio quadrático médio* σ:

$$\sigma = \sqrt{\sum_{1}^{n} \frac{(x_i - \bar{x})^2}{n-1}} \tag{2}$$

E o resultado deve ser indicado $(\bar{x} \pm \sigma)$

Para o exemplo apontado, deveríamos indicar para o diâmetro do prego

$$\bar{d} = 0,97500000.. \pm 0,0327108..$$

que o bom senso sugere

$$\bar{d} = 0,98 \pm 0,03$$

GUARDE APENAS O PRIMEIRO ALGARISMO SIGNIFICATIVO DE σ

Este é um procedimento tedioso, mas *necessário*. A maioria das máquinas de calcular incluem programas de cálculo de médias e desvio-padrão.

Aprenda a usar sua máquina !

$\bar{x} = \frac{1}{n}\sum_1^n x_i$ **é** a melhor escolha para representar o conjunto de pontos x_i:
Construamos uma função $D^2 = \sum_1^n (x_i - \bar{x})^2$. O *melhor* valor de $\bar{x}$ é aquele que torna mínima a soma das diferenças. Logo,

$$\frac{dD^2}{d\bar{x}} = -2\sum_1^n (x_i - \bar{x}) \qquad = 0$$

$$\sum_1^n x_i - \bar{x}\sum_1^n 1 = 0$$

$$\sum_1^n x_i - n\bar{x} = 0$$

$$\bar{x} = \frac{1}{n}\sum_1^n x_i$$

Propagação de Erros Raras vezes a medida que fazemos constitui o resultado final do processo de investigação que fazemos. O número obtido deve ser transformado matematicamente — somado, subtraído, dividido, multiplicado... — para dar a conhecer a resposta. Outros números intervenientes irão, com boa probabilidade, também apresentar uma incerteza, o que nos leva à questão: como determinar o número $\bar{z}$ que resulta da combinação de $\bar{x}$ e $\bar{y}$, i.e, $\bar{z} = f(\bar{x}, \bar{y})$?
Devemos distinguir entre *soma e subtração* e *multiplicação e divisão*. Na *adição e subtração*, se um dos termos é uma constante, desprovida de erro (como, por exemplo, π), o erro do termo remanescente determina o erro final. Já se ambos os termos apresentam erro, então pensamos nos erros como *vetores* e os somamos em quadratura. Dados $\bar{x} \pm \sigma_x$ + $\bar{y} \pm \sigma_y$, o melhor procedimento nos dá

$$\bar{z} = (\bar{x} + \bar{y}) \pm \sqrt{\sigma_x^2 + \sigma_y^2}$$

Na *multiplicação e divisão*, temos $\bar{z} = \bar{x} \cdot \bar{y} \pm \sigma_z$ onde

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

É mais prático, no entanto, considerar a *relação diferencial*:

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

n	I	$V \pm \sigma_V$
	(A)	$(V \pm 0.5V)$
1	$0,2$	$1,0$
2	$0,4$	$2,2$
3	$0,6$	$3,0$
4	$0,7$	$3,5$
5	$0,8$	$4,0$
6	$1,2$	$6,0$
7	$1,4$	$7,0$
8	$1,6$	$7,8$
9	$1,8$	$9,2$
10	$2,0$	$11,0$

n	x_i	y_i	x_i^2	$x_i y_i$
1	$0,2$	$1,0$	$0,04$	$0,20$
2	$0,4$	$2,2$	$0,16$	$0,88$
3	$0,6$	$3,0$	$0,36$	$1,80$
4	$0,7$	$3,5$	$0,49$	$2,45$
5	$0,8$	$4,0$	$0,64$	$3,20$
6	$1,2$	$6,0$	$1,44$	$7,20$
7	$1,4$	$7,0$	$1,96$	$9,80$
8	$1,6$	$7,8$	$2,56$	$12,48$
9	$1,8$	$9,2$	$3,24$	$16,56$
10	$2,0$	$11,$	$4,00$	$22,00$
$\sum$	$10,7$	$54,7$	$14,89$	$76,57$

Tabela 1: **Tensões medidas sobre um resistor. Admitimos que a *corrente*, variável independente, não apresenta erros. Apenas os valores de tensão apresenta erros, que supomos iguais na região medida. Na tabela à direita estão os dados necessários para ajuste de uma reta pelo método dos mínimos quadrados.**

A generalização é imediata, quando temos mais de dois números envolvidos. Para $\bar{z} = \bar{u} \cdot \bar{v} \cdot \bar{x} \cdot \ldots$,

$$\left(\frac{\sigma_z}{z}\right)^2 = \left(\frac{\sigma_u}{u}\right)^2 + \left(\frac{\sigma_v}{v}\right)^2 + \left(\frac{\sigma_x}{x}\right)^2 + \ldots$$

Resumindo:

> Quando somando ou subtraindo grandezas afetadas de erro, combine as incertezas **absolutas** em quadratura. Quando multiplicando ou dividindo, combine as incertezas **fracionais**, ou relativas, em quadratura.

Ignore incertezas menores que $1/3$ das outras.

Ajuste de curvas Frequentemente praticamos uma série de medidas as quais sabemos pertencer a uma relação matemática. O caso mais simples é o de uma *relação linear*, que se traduz numa *reta* quando os resultados são apresentados de forma gráfica. Seja, por exemplo, determinar o melhor valor da resistência elétrica de um resistor sobre o qual, dada a corrente, medimos a tensão. Tomamos o erro cometido como constante — como é a situação prática mais encontrada. Sabemos que o resistor obedece à lei de Ohm, e portanto $V(I) = RI$. Em nossas medidas, admitimos que que a *variável independente* I não contem erros e as incertezas estão sobre a *variável dependente* V. Nosso problema é determinar *o melhor* valor para a resistência elétrica R. Neste exemplo, *todos* os pontos são utilizados para

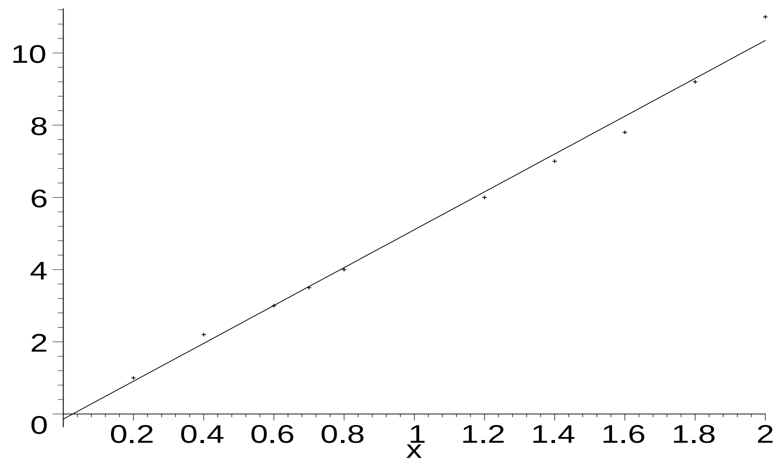


Figura 1: **Dados pontos experimentais que sabemos obedecer a uma relacão matemática conhecida, como determinar o melhor conjunto de parâmetros ?**

a determinação do *melhor* valor para a resistência elétrica do resistor. Que métodos podemos utilizar para isto ?

O método mais simples é o olho. Por mais surpreendente que possa parecer, um ajuste visual dos dados normalmente fornece bons resultados – e deve ser a primeira estimativa do processo. Um outro método consiste em traçar *retas de inclinações extremas* sobre os pontos e praticar algum tipo de média sobre os resultados. Entretanto, se queremos um método que forneça os mesmos resultados, ainda que aplicado por pessoas diferentes, temos que recorrer a processos matemáticos mais elaborados.

O mais famoso deles é denominado "Método dos mínimos quadrados" e consiste, em essência, em minimizar a diferença entre a curva a ser traçada e os pontos experimentais. Começamos construindo uma função χ^2 (leia-se "chi") matemática da *diferença*, ao quadrado, entre os pontos experimentais y_i e os pontos teóricos correspondentes y_{ti}

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - y_{ti})^2}{\sigma_i^2} = \sum_{i=1}^{n} \frac{(y_i - (ax_i + b))^2}{\sigma_i^2}$$

onde σ_i é a incerteza associada ao ponto medido y_i, n é o número de pontos experimentais, a e b são os coeficientes da reta que melhor ajusta os pontos — as nossas incógnitas.

Minimizamos χ^2 para determinar os melhores valores de a e b. Posto de outra forma, os *melhores* valores de a e b são exatamente aqueles que *tornam mínima a soma das diferenças entre os pontos experimentais e a reta procurada*.

Então,

$$\begin{aligned}
\frac{\partial \chi^2}{\partial a} &= -2\sum_{i=1}^{n} \frac{(y_i - ax_i - b)\, x_i}{\sigma_i^2} = 0 \\
\frac{\partial \chi^2}{\partial b} &= -2\sum_{i=1}^{n} \frac{(y_i - ax_i - b)}{\sigma_i^2} = 0
\end{aligned}$$

com solução

$$a = \frac{1}{\Delta}\left(\mathcal{S}_\sigma \mathcal{S}_{xy} - \mathcal{S}_x \mathcal{S}_y\right) \quad \text{e} \quad b = \frac{1}{\Delta}\left(\mathcal{S}_{x^2} \mathcal{S}_y - \mathcal{S}_x \mathcal{S}_{xy}\right)$$

onde

$$\Delta = \mathcal{S}_\sigma \mathcal{S}_{x^2} - \mathcal{S}_x^2$$

e

$$\mathcal{S}_\sigma = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \quad \mathcal{S}_x = \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \quad \mathcal{S}_y = \sum_{i=1}^{n} \frac{y_i}{\sigma_i^2} \quad \mathcal{S}_{xy} = \sum_{i=1}^{n} \frac{x_i y_i}{\sigma_i^2} \quad \mathcal{S}_{x^2} = \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2}$$

Exercício. Faça a dedução detalhada das fórmulas.

As incertezas sobre a e b são dadas por

$$\sigma_a = \sqrt{\frac{\mathcal{S}_\sigma}{\Delta}} \qquad \sigma_b = \sqrt{\frac{\mathcal{S}_{x^2}}{\Delta}}$$

Vamos aplicar esta técnica aos dados experimentais da tabela 1. A forma mais fácil de praticar ajustes de retas está em utilizar as máquinas de calcular. As contas que se seguem constituem apenas um exemplo de aplicação.

$$\mathcal{S}_\sigma = 40 \quad N\frac{\mathcal{S}_x}{\mathcal{S}_\sigma} = 10,7 \quad N\frac{\mathcal{S}_y}{\mathcal{S}_\sigma} = 54,7 \quad N\frac{\mathcal{S}_{x^2}}{\mathcal{S}_\sigma} = 14,89 \quad N\frac{\mathcal{S}_{xy}}{\mathcal{S}_\sigma} = 76,57 \quad \Delta = \frac{481,11}{N^2}$$

Daí obtemos

$$\begin{aligned}
a &= \frac{10 \cdot 76,57 - 10,7 \cdot 54,7}{34,41} &&= 5,24295\ldots \\
b &= \frac{14,89 \cdot 54,7 - 10,7 \cdot 76,57}{34,41} &&= -0,13995\ldots \\
\sigma_a &= \sqrt{\frac{40}{505,56}} &&= 0,26954\ldots \\
\sigma_b &= \sqrt{\frac{14,89}{505,56}} &&= 0,34323\ldots
\end{aligned}$$

Os numeros devem ser indicados

$$a = 5,2 \pm 0,3 \qquad b = -0,1 \pm 0,3$$

A resposta, pois, é que o resistor tem valor $a = 5,2 \pm 0,3\,\Omega$.

---

21 22 23 24 25 26 27 28 ___/ ___/ 2009

Licenciatura em Física

Nome:________________________________ **N^o USP:** ☐ ☐☐☐ ☐☐☐

Companheiros:

________________________________ **1**

____

________________________________ **R**

____

EXPERIÊNCIA 0 Tratamento Estatístico de Dados Experimentais

Um capacitor é formado por duas placas circulares paralelas de diâmetro D separadas de ℓ.

☐ Um aluno mediu os seguintes valores, em centímetros, para o diâmetro de uma das armaduras:

$$16.4 \quad 16.2, \quad 16.3 \quad 16.5 \quad 16.4 \quad 16.3 \quad 16.2 \quad 16.3 \quad 16.1 \quad 16.4$$

Este conjunto de medidas é melhor indicado por $\boxed{\bar{D} = ... \pm .... cm}$

☐ Calcule a área da placa, expressa por $A = \pi D^2/4$.

- Expresse π de forma coerente. A máquina de calcular fornece $\pi = 3.141592654$ (implicando erro de $1/10^{10}$ que, obviamente não faz sentido em nosso caso). A forma correta de expressar π *para este caso!* é $\boxed{\pi = ...}$

☐ O mesmo estudante mediu, em milímetros, para a separação entre as placas,

$$3,20 \quad 3,22 \quad 3,23 \quad 2,96 \quad 3,00 \quad 2,95 \quad 2,90 \quad 2,92 \quad 2,96$$

Este conjunto de medidas é melhor indicado por $\boxed{\bar{\ell} = ... \pm .... mm}$

☐ A capacitância de um capacitor de placas paralelas, desprezando-se os efeitos de borda, é dada por $C = \varepsilon_o A/\ell$ onde ε_o, denominado *Permissividade do Vácuo* tem valor tabelado $8.854\,187\,817 \times 10^{-12}\ F/m$. Calcule a capacitância do arranjo:

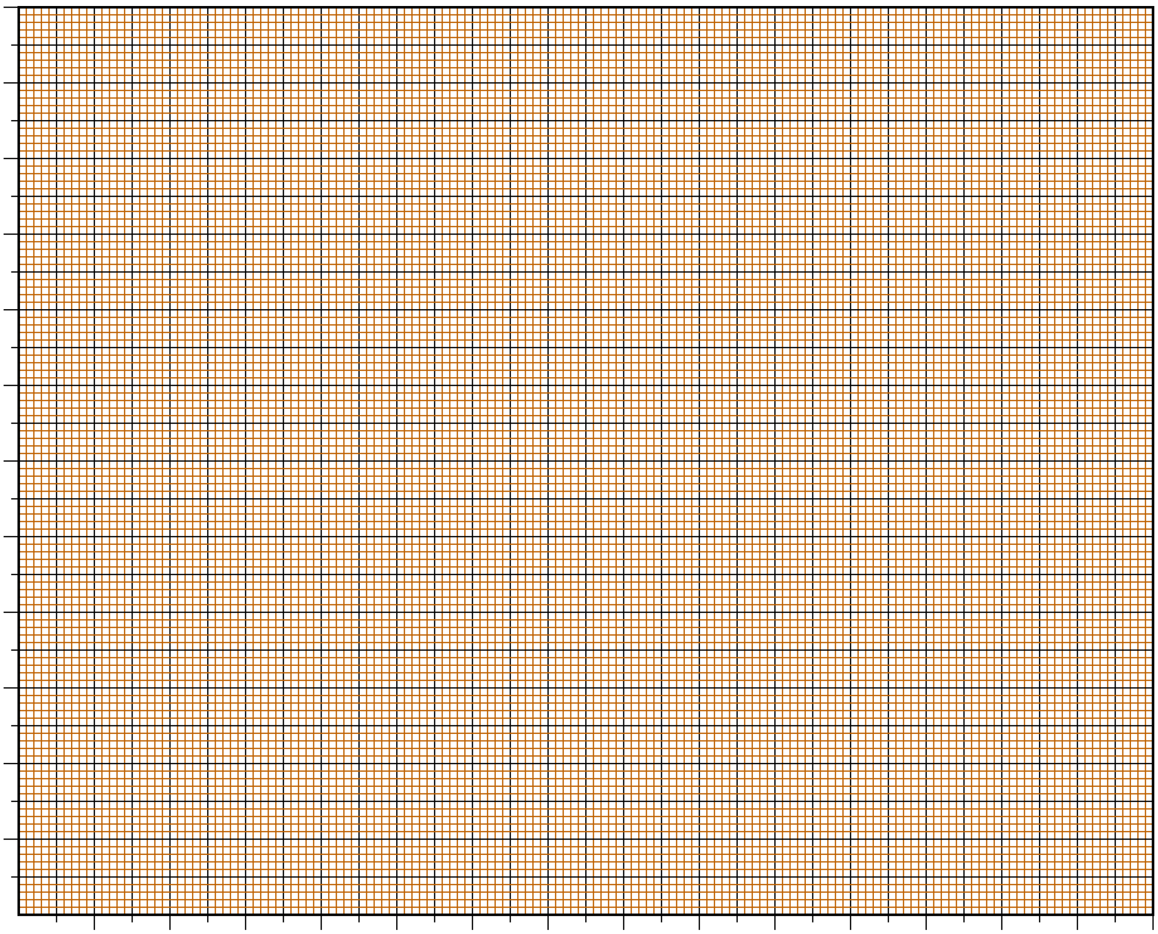
- Escreva ε_o com o número adequado de casas decimais $\varepsilon_o = ...\ F/m$
- Calcule a capacitância do capacitor: $C = (............ \pm ....) \times 10^{-......}\ F$

Mediu-se a *tensão* V entre e os terminais do capacitor e a *carga* Q armazenada neste capacitor, obtendo-se os valores da tabela. A relacão entre a carga armazenada num capacitor e a tensão entre seus terminais é dada por $Q = CV$. (Admita que as tensões são perfitamente conhecidas e o erro cometido na medida da carga é constante e igual $0,1$).

n		1	2	3	4	5	6	7	8	9
Tensão	(V)	500	600	700	800	900	1000	1100	1200	1300
Carga	(pC)	31.0	36.8	43.0	49.2	55.4	62.0	67.0	73.8	80.0

DETERMINAÇÃO DOS MELHORES VALORES DE CAPACITÂNCIA.

☐ **Pelo olho**. Construa um gráfico com V nas abscissas e Q nas ordenadas. Trace retas extremas aos dados, tal como melhor lhe parecer e estime o melhor valor para a capacitância, tomando a média entre os valores obtidos. Estime a incerteza.



Melhor valor para a capacitância: $C = (........ \pm .....) \times 10^{-\cdots}\,F$

☐ **Pelo método dos mínimos quadrados.**
Complete a tabele com os calculos necessários ao cálculo de coeficientes pelo método dos mínimos quadrados.

n	x_i	y_i	$1/\sigma_i^2$	x_i^2	$x_i y_i$
1					
2					
3					
4					
5					
6					
7					
8					
9					
$\sum$					

$S_\sigma =$ .... $S_x =$ .... $S_{x^2} =$ .... $S_y =$ .... $S_{xy} =$ ....

Melhor valor para a capacitância: $C = (\ldots\ldots\ldots \pm \ldots\ldots) \times 10^{-\cdots}\ F$

☐ ***OPCIONAL !*** **Pela Calculadora**. Caso voce tenha uma calculadora que lhe permite ajustar retas, Calcule o ajuste pela máquina.

☐ Discuta brevemente os resultados.