

# Como estimar a incerteza no declive a partir do coeficiente de Pearson?

Manuela Ramos Silva<sup>1</sup>, Pablo Martín-Ramos<sup>2</sup>, Pedro Sidónio Pereira da Silva<sup>1</sup>

<sup>1</sup> CFisUC, Departamento de Física, Universidade de Coimbra, 3004-516 Coimbra, Portugal

<sup>2</sup> ESPH, Universidad de Zaragoza, Carretera de Cuarte s/n, 22071 Huesca, Espanha

psidonio@pollux.fis.uc.pt

## Introdução

A Física é uma ciência experimental: todos os conceitos físicos têm de ser verificados experimentalmente antes de os aceitarmos como leis da natureza. No estudo dessas leis, das mais simples às mais complexas, queremos muitas vezes determinar como uma quantidade que estamos a medir depende de outra. Muitas vezes essa dependência é linear ou pode ser linearizada, como por exemplo, a variação do período de um pêndulo com o comprimento de um fio, da diferença de potencial aos terminais de uma resistência com a corrente elétrica que a percorre, da temperatura crítica de um supercondutor com o campo magnético aplicado, do número de partículas alfa que volta para trás ao atingir uma folha de ouro em função da espessura dessa folha, da energia cinética máxima de fotoelétrons em função da frequência de luz incidente em superfícies metálicas, etc. O ajuste de uma reta aos pontos experimentais permite-nos extrair o declive da reta e a partir dele determinar uma grandeza física que se mantenha constante na experiência. Nos exemplos anteriores: a aceleração da gravidade, a resistência elétrica, o intervalo de energias proibidas entre estado supercondutor e estado normal, a carga do eletrão, a constante de Planck.

Embora nos laboratórios científicos exista muitas vezes *software* especificamente criado para permitir a análise dos resultados de cada experiência, nos laboratórios didáticos recorre-se muitas vezes à calculadora científica ou a folhas de cálculo simplificadas para fazer o ajuste de uma reta aos dados experimentais. Estas ferramentas permitem-nos de forma fácil obter os parâmetros da melhor reta, nomeadamente o declive, e também o coeficiente de Pearson ( $R$ ), cujo quadrado nos permite ajustar do ajustamento do modelo linear. Mais difícil é obter a incerteza associada ao valor do declive,

para, a partir desta, obter a incerteza na grandeza física que pretendemos determinar. É no entanto possível determinar a incerteza no declive a partir dos valores do próprio declive e do coeficiente de Pearson, como se demonstrará em seguida. Essa expressão é aplicada aos dados obtidos numa montagem simples para ilustração da lei de Ohm.

## Definições e cálculos

Consideremos o nosso conjunto de dados experimentais:  $N$  pares de valores  $(x, y)$  em que  $\bar{y}$  é a média do conjunto dos valores  $y_i$ ,  $\bar{y} = (\sum_{i=1}^N y_i)/N$  e  $\bar{x}$  a média para o conjunto de valores  $x_i$ ,  $\bar{x} = (\sum_{i=1}^N x_i)/N$ .

Chamaremos  $\hat{y}_i$  ao valor de  $y_i$  obtido através da reta que faz o melhor ajuste, ou seja  $\hat{y}_i(x_i) = m \cdot x_i + b$ . Vamos também definir as seguintes somas de quadrados:

$$\begin{aligned} SQ_T &= SQ_{yy} = \sum_{i=1}^N (y_i - \bar{y})^2 \\ SQ_E &= \sum_{i=1}^N (y_i - \hat{y}_i)^2 \\ SQ_{xx} &= \sum_{i=1}^N (x_i - \bar{x})^2 \\ SQ_{xy} &= \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \end{aligned} \quad (1)$$

O declive da reta que faz o melhor ajuste é

$$m = \frac{SQ_{xy}}{SQ_{xx}} \quad (2)$$

e a ordenada na origem  $b = \bar{y} - m\bar{x}$ .

O desvio padrão do declive, que se calcula através da fórmula habitual de propagação de erros,

$$\sigma^2 = \left(\frac{\partial f}{\partial w}\right)^2 \Delta w^2 + \left(\frac{\partial f}{\partial v}\right)^2 \Delta v^2 + \left(\frac{\partial f}{\partial u}\right)^2 \Delta u^2 + \dots \quad (3)$$

é dado por

$$\sigma^2 = \frac{1}{N-2} \cdot \frac{SQ_E}{SQ_{xx}} \quad (4)$$

e  $R^2$ , que nos permite avaliar da qualidade do ajuste, é igual a

$$\frac{SQ_T - SQ_E}{SQ_T} \quad (5)$$

As somas de quadrados acima definidos podem escrever-se em função umas das outras, por exemplo,

$$\begin{aligned} SQ_T - SQ_E &= \sum_{i=1}^N (y_i - \bar{y})^2 - \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \sum_{i=1}^N (y_i - \bar{y})^2 - \sum_{i=1}^N (y_i - mx_i - \bar{y} + m\bar{x})^2 \\ &= \sum_{i=1}^N (y_i - \bar{y})^2 - \sum_{i=1}^N ((y_i - \bar{y}) - m(x_i - \bar{x}))^2 \\ &= \sum_{i=1}^N (y_i - \bar{y})^2 - \sum_{i=1}^N ((y_i - \bar{y})^2 - 2m(x_i - \bar{x})(y_i - \bar{y}) + m^2(x_i - \bar{x})^2) \\ &= +2m \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) - m^2 \sum_{i=1}^N (x_i - \bar{x})^2 \\ &= 2 \cdot \frac{SQ_{xy}}{SQ_{xx}} \cdot SQ_{xy} - \frac{SQ_{xy}^2}{SQ_{xx}^2} SQ_{xx} = \frac{SQ_{xy}^2}{SQ_{xx}} \end{aligned}$$

Usando esta igualdade é possível calcular  $\sigma$  usando  $m$  e  $R^2$ ,

$$\begin{aligned} \sigma^2 &= \frac{1}{N-2} \cdot \frac{SQ_E}{SQ_{xx}} = \frac{1}{N-2} \cdot \frac{1}{SQ_{xx}} \cdot SQ_E \cdot \frac{SQ_{xy}^2}{SQ_{xy}^2} \cdot \frac{SQ_{xx}}{SQ_{xx}} = \frac{1}{N-2} \cdot \frac{SQ_{xy}^2}{SQ_{xx}^2} \cdot SQ_E \cdot \frac{SQ_{xx}}{SQ_{xy}^2} \\ &= \frac{1}{N-2} \cdot m^2 \cdot \frac{SQ_E}{SQ_T - SQ_E} = \frac{1}{N-2} \cdot m^2 \cdot \frac{SQ_T - SQ_T + SQ_E}{SQ_T} \\ &= \frac{1}{N-2} \cdot m^2 \cdot \frac{1 - R^2}{R^2} \\ \Rightarrow \sigma &= m \sqrt{\frac{1}{N-2} \cdot \frac{1 - R^2}{R^2}} \quad (6) \end{aligned}$$

$(N-2)$  é o número de graus de liberdade, ou seja, o número de pontos experimentais independentes menos o número de parâmetros a estimar no ajuste.

### Aplicação

Como aplicação para a fórmula acima deduzida, vamos utilizar dois conjuntos de dados experimentais onde se verifica uma relação linear entre a diferença de potencial aplicada aos terminais de uma resistência e a intensidade da corrente eléctrica que a percorre. No primeiro ensaio usou-se uma fonte de tensão TTi (*Thurlby Thandar Instruments*)<sup>®</sup> PL330DP, um multímetro digital *digimess*<sup>®</sup> DM100 e uma resistência de 1 kΩ. No segundo ensaio, manteve-se o multímetro e usou-se a mesma resistência, mudando a fonte de tensão para o modelo *digimess*<sup>®</sup> HY3003-3 DC (Figura 1).

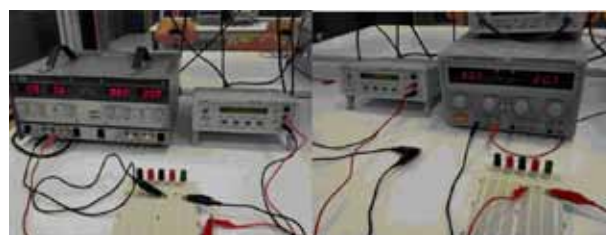


Fig. 1 - (esq.) Montagem experimental usando uma fonte de tensão TTi PL330DP e um multímetro DM100; (dir.) montagem experimental usando uma fonte de tensão HY3003-3 DC e um multímetro DM100.

Tabela 1 - Resultados Experimentais.

Ensaio 1		Ensaio 2	
Diferença de potencial (V)	Intensidade de corrente (mA)	Diferença de potencial (V)	Intensidade de corrente (mA)
0	-32,9x10 <sup>-3</sup>	0	5,92x10 <sup>-3</sup>
0,5	0,457	1,0	0,91
1,0	0,913	2,0	1,85
1,5	1,373	3,0	2,97
2,0	1,826	4,0	3,96
2,5	2,496	5,0	4,95
3,0	2,991	6,0	5,99
3,5	3,289	7,0	6,93
4,0	3,996	8,0	7,99
4,5	4,493	9,0	8,91
5,0	4,997	10,0	9,94

Os dados obtidos (Tabela 1) foram depois representados em dois gráficos distintos, quer usando a folha de cálculo Excel (Fig. 2), quer usando a calculadora científica TI-84 Plus (Fig. 3).

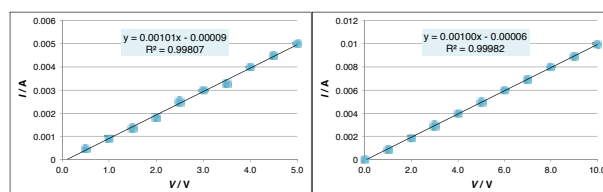


Fig. 2 - Gráficos obtidos usando a folha de cálculo Excel.



Fig. 3 - Gráfico e respectivo ajuste obtidos usando a calculadora gráfica.

Usando as opções destes dois programas, traçou-se a reta que corresponde ao melhor ajuste e obteve-se a equação da reta e o coeficiente de Pearson.

A partir dos valores do declive e do coeficiente de Pearson é possível, então, calcular a incerteza asso-

ciada ao declive para cada ensaio:

$$\text{Ensaio 1: } \sigma = 0,00101 \sqrt{\frac{1}{11-2} \times \frac{1-0,99802}{0,99802}} = 1,5 \times 10^{-5}$$

$$\text{Ensaio 2: } \sigma = 0,001 \sqrt{\frac{1}{11-2} \times \frac{1-0,99982}{0,99982}} = 5 \times 10^{-6}$$

Os resultados obtidos nos dois ensaios para o valor da resistência, são, respectivamente,  $R_{e1}=990 \Omega$  e  $R_{e2}=1000 \Omega$  e as incertezas associadas obtém-se a partir do desvio padrão de  $m$ , usando a fórmula de propagação de erros (3),

$$(\Delta R_e)^2 = \left( \frac{\partial}{\partial m} \left( \frac{1}{m} \right) \right)^2 (\Delta m)^2 = \frac{1}{m^4} (\Delta m)^2 \Rightarrow \Delta R_e = \frac{\Delta m}{m^2}$$

ou seja,  $\Delta R_{e1}=15 \Omega$  e  $\Delta R_{e2}=5 \Omega$ . Assim os resultados dos dois ensaios da medida da resistência eléctrica por utilização da lei de Ohm são  $R_{e1}=990 \pm 15 \Omega$  e  $R_{e2}=1000 \pm 5 \Omega$ , respectivamente.

Como nos dois ensaios foi usada exactamente a mesma resistência, é conveniente apresentar um único valor para  $R_e$ . Se se puder considerar que as incertezas associadas aos vários pares  $(V, I)$ , são iguais, independentemente do ensaio, então podem-se reunir todos os valores num único gráfico e ajustar-lhes uma reta, como mostra a Figura 4.

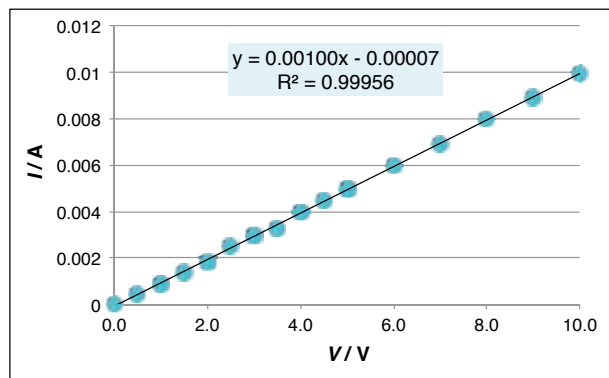


Fig. 4 - Gráfico obtido usando a folha de cálculo Excel.

Assim, obtemos o valor  $R_e=1000 \pm 7 \Omega$ .

Se nos dois ensaios fossem usados equipamentos muito diferentes, de tal forma que a incerteza experimental associada a cada medida fosse muito diferente, perderíamos as condições que nos permitiram deduzir a fórmula (1) mas podemos sempre obter o valor para  $R$  calculando a média ponderada dos valores  $R_{e1}$  e  $R_{e2}$ :

$R_e = 999 \Omega$ , com maior desvio para a média de  $9 \Omega$ , logo  $R_e = 999 \pm 9 \Omega$  (ver Tabela 2).

Tabela 2 - Cálculo da Resistência Elétrica.

	Ensaio 1	Ensaio 2
Valor do declive	0,00101	0,00100
Incerteza do declive	$1,5 \times 10^{-5}$	$5 \times 10^{-5}$
Resistência ( $\Omega$ )	990	1000
Incerteza associada	15	5
Média ponderada	$999 \pm 9$	
Resistência a partir do declive (todos os dados)	$1000 \pm 7$	

Um último comentário para a importância do traçado de um gráfico retratando as medidas experimentais: qualquer desvio da dependência linear é imediatamente aparente (esse desvio podia acontecer se danificássemos a resistência fazendo-lhe passar correntes muito elevadas). Além disso, o declive da melhor reta, ou seja o valor da resistência, é imune a erros sistemáticos, quer na fonte de tensão quer no amperímetro. Esses erros são aparentes na ordenada na origem, que surge diferente de zero (sendo zero o valor teoricamente previsto pela lei de Ohm).

#### Bibliografia

1. N. C. Barford, *Experimental Measurements: Precision, Error and Truth*, John Wiley, New York (1985)
2. I. F. Hughes, T. P. A. Hase, *Measurements and their Uncertainties, A Practical Guide to Modern Error Analysis*, Oxford University Press, Oxford (2010)
3. R. H. McCuen, *Statistical Methods for Engineers*, Prentice Hall, Englewood Cliffs, NJ (1985)
4. D. C. Montgomery, G. C. Runger, *Applied Statistics and Probability for Engineers*, Wiley, New York (2011)
5. J. R. Taylor, *An Introduction to Error Analysis*, University Science Books, California (1997)



**Manuela Ramos Silva** licenciou-se (1993) em Física e doutorou-se (2002) em Física Experimental na Universidade de Coimbra, onde é professora auxiliar com agregação. Trabalha na área da Física dos Materiais, com particular interesse em magnetismo molecular, tendo mais de 250 artigos publicados em revistas internacionais e mais de 3000 horas de aulas lecionadas.



**Pablo Martín-Ramos** é professor na Universidade de Saragoça. Concluiu o seu Mestrado em Engenharia Eletrotécnica em 2011 na Universidade de Columbia, Nova Iorque e o seu doutoramento em Engenharia de Telecomunicações em 2013 na Universidade de Valladolid. Atualmente, o seu trabalho é centrado em

materiais multifuncionais contendo lantanídeos e catalisadores heterógeneos para fotossíntese artificial. Tem mais de 60 trabalhos publicados em revistas internacionais.



**Pedro Sidónio Pereira da Silva** licenciou-se (1995) em Engenharia Física e finalizou o seu mestrado em Física Tecnológica em 2000 na Universidade de Coimbra. Concluiu o doutoramento (2012) em Física Experimental na mesma Universidade, onde trabalha atualmente como técnico superior (laboratório TAIL-

-UC). Entre 2000 e 2008 foi assistente no Instituto Politécnico de Castelo Branco. Trabalha na área da Física dos Materiais, com particular interesse em óptica não-linear, tendo mais de 70 artigos publicados em revistas internacionais.