Erros Experimentais

– uma abordagem pedagógica[1] – Parte II

ISABEL M. A. FONSECA*

Resumo

A análise e tratamento dos erros sistemáticos e aleatórios é realizada por métodos distintos. O método relativo aos erros sistemáticos é o mais elaborado pois é necessário conhecer a sua origem para os poder corrigir. Já os erros aleatórios como apresentam origem subjectiva só podem ser detectados pela repetição das observações, podendo ser minimizados através de métodos que possibilitam a sua análise estatística.

Analisaremos também, com algum detalhe como é que as incertezas associadas aos valores medidos se propagam através do processo de medição para os resultados finais, fazendo uso das leis de propagação dos erros e das variâncias e do método directo.

1. Análise dos erros sistemáticos

Dum modo geral, podemos afirmar que a análise e tratamento dos erros de carácter sistemático que afectam a exactidão dos resultados, é um processo mais difícil e elaborado do que a análise e tratamento dos erros aleatórios que condicionam a precisão dos dados. Isto deve-se ao facto de a condição necessária para corrigir os erros sistemáticos, total ou parcialmente, é o conhecimento da sua origem. Já no caso dos erros aleatórios que apresentam origem subjectiva, esta questão não se levanta.

Como se referiu na Parte I deste artigo, o conceito de exactidão está associado à ideia de proximidade do valor correcto e, por conseguinte, a estimativa deste parâmetro exije que se conheça o valor correcto da grandeza que está a ser medida. Por exemplo, são considerados correctos os valores obtidos por instrumentos padrão fornecidos pelo NPL (National Physical Laboratory, USA). É o caso dos termómetros de mercúrio fabricados e certificados por este organismo e largamente utilizados como padrões primários para calibrar outros instrumentos de medição de temperatura.

O procedimento usado para estimar os erros sistemáticos depende muito da natureza da experiência em questão, no entanto, podemos referir duas situações mais comuns em que este procedimento é relativamente simples:

(a) Medição de uma grandeza conhecida

A determinação de uma grandeza para a qual existem na literatura um ou mais valores aceites como correctos, permite calcular a exactidão do valor medido por aplicação da expressão,

$$\Delta x = x_{obs} - x_c \quad (1)$$

Este procedimento é efectuado na maioria dos casos com o objectivo de determinar a exactidão inerente a um determinado método experimental, aparelho, etc, ou seja, quando se efectua uma calibração.

(b) Medição de uma grandeza desconhecida.

Esta situação é aquela com que o investigador se depara com muita frequência. Para inferir da correcção do resultado realiza a calibração do aparelho e do método, recorrendo a substâncias ou aparelhos de referência; a exactidão obtida com este procedimento é depois considerada como sendo a exactidão da grandeza sob medição.

Vamos agora imaginar uma experiência que se enquadre nesta situação. Considere-se a medição da temperatura de um banho termostatizado com um termómetro de resistência de platina.

A exactidão do resultado obtido pode ser determinada por comparação com a temperatura lida num padrão primário, por exemplo, um termómetro de mercúrio do NPL, cujo valor se considera ser correcto. Outro exemplo, pode ser a medição da viscosidade de um álcool, o pentanol, à temperatura T. Procede-se em primeiro lugar à calibração do viscosímetro com etanol, álcool para o qual existe muita informação na literatura, efectuando a medição à temperatura T. Por comparação do valor médio dos resultados obtidos com o valor correcto (tabelado na literatura), é possível inferir da exactidão do resultado usando a eq.(1).

Posteriormente, efectua-se a determinação da viscosidade do pentanol e atribui-se-lhe a exactidão obtida na calibração anterior.

A identificação e tratamento dos erros sistemáticos exige que o experimentalista aprenda a antecipar as possíveis fon-

* Departamento de Engenharia Química da Universidade de Coimbra, Pinhal de Marrocos, Pólo II, 3030-290, Coimbra, Portugal (fonseca@eq.uc.pt)

1 Segunda parte do artigo publicado no Boletim da Sociedade Portuguesa de Química, nº 95, 2004

tes desta categoria de erros e a garantir que são inferiores à precisão requerida (ver Exemplo 4 da Parte I deste artigo). Este procedimento implica um profundo envolvimento do investigador no processo experimental (método experimental, aparelho, etc) e vai-se aperfeiçoando com a prática.

2. Análise estatística dos erros aleatórios

Os erros aleatórios devido às suas características só se revelam com a repetição das observações, podendo ser minimizados através da sua análise estatística. O procedimento adequado para analisar este tipo de erros é, por conseguinte, efectuar um conjunto significativo de medições da mesma grandeza, nas mesmas condições. O conjunto das n observações seguirá, em princípio, uma distribuição normal (para $n>30$). A melhor estimativa do valor da grandeza é a sua média, $\bar{x}$, na qual depositamos mais confiança do que em cada medição individual. Já referimos também que a incerteza média das medições individuais (precisão) é caracterizada pelo desvio padrão da amostra, σ, i.e., cada observação individual, x_i, apresenta a mesma incerteza

$$x_i \pm \sigma \tag{2}$$

Como é óbvio levanta-se agora a questão de saber qual é a incerteza associada ao valor médio, $\bar{x}$, que será naturalmente menor do que a incerteza das medições individuais. A expressão que define a incerteza no valor médio é,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \tag{3}$$

e designa-se por desvio padrão da média (ou desvio médio). Deste modo o valor da média deve ser afectado da respectiva incerteza (ou precisão):

$$\bar{x} \pm \sigma_{\bar{x}} \tag{4}$$

O facto de o denominador da eq.(4) ser $\sqrt{n}$ implica que aumentando o número de observações, o desvio padrão da média diminua, e, portanto, somos levados a concluir que se fizermos mais medições obteremos um valor de maior confiança, o que é verdade. No entanto, temos que ter em atenção que o factor $\sqrt{n}$ cresce muito lentamente com o aumento de n, o que determina que o aumento do número de medições não seja muitas vezes a estratégia mais acertada para melhorar a qualidade dos resultados. Na prática, o bom experimentalista, concluirá que para melhorar significativamente a precisão dos resultados é preferível dirigir os seus esforços no sentido de aperfeiçoar o processo experimental em vez de aumentar o número de medições. O exemplo 1 permite comparar a ordem de grandeza do desvio padrão de cada observação individual com o desvio padrão da média das observações.

Exemplo 1. Considere que a média duma amostra de 100 obervações duma determinada grandeza é $\bar{x} = 2.341$ (unidades) e que o desvio padrão é 0.180 (unidades). Então o desvio médio terá o valor,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.180}{\sqrt{100}}$$

$= 0.018$ (unidades)

Por conseguinte o resultado e a respectiva incerteza devem ser apresentados da seguinte forma:

$\bar{x} = 2.341 \pm 0.018$ (unidades)

Podemos verificar que a incerteza na média (desvio médio) é 10 vezes menor do que a incerteza de cada medição individual.

3. Propagação dos erros

A maioria das grandezas físicas não pode ser medida a partir de uma única observação directa. O procedimento envolve geralmente dois passos distintos. Um primeiro passo que consiste na determinação experimental de uma ou mais grandezas que podem ser medidas directamente. Num segundo passo, efectua-se então o cálculo da grandeza pretendida recorrendo aos valores medidos. Por exemplo, para determinarmos a área de um rectângulo medimos o seu comprimento, c, e a sua largura, l, e só depois calculamos a sua área, A, através da expressão, A = c × l.

Muitos outros exemplos podiam ser referidos em que a medição pressupõe estes dois passos distintos: a medição directa seguida dum processo de cálculo.

Sempre que uma experiência envolve estes dois passos a estimativa dos erros será também um procedimento com dois passos. Em primeiro lugar, devemos estimar os erros nas grandezas medidas directamente e, posteriormente, determinar como é que esses erros se propagam através dos cálculos para originar o erro no resultado final. Para determinar como é que os erros se propagam para o valor final utiliza-se uma ferramenta designada por lei de propagação dos erros. Uma outra forma desta lei designa-se por lei de propagação das variâncias (ou dos desvios padrões) e permite determinar como é que as incertezas (traduzidas pelos desvios padrões) se propagam dos dados para o resultado final. Esta lei aplica-se nos casos em que os dados resultam de medições repetidas, sendo possível determinar as precisões que lhe estão inerentes.

Qualquer uma das formas possibilita determinar qual é a variável responsável pela maior contribuição para a incerteza do resultado final.

3.1 Lei da propagação de erros (l.p.e.)

Para percebermos a forma analítica da lei da propagação dos erros vamos considerar o caso mais simples de uma função de uma só variável, $f(x)$. Vamos admitir que a grandeza x está afectada de um erro, Δx. O correspondente erro em f, Δf, é dado pela expressão,

$$\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \frac{df}{dx} \tag{5}$$

ou seja,

$$\Delta f \approx \frac{df}{dx}\Delta x \tag{6}$$

expressão que dá uma boa aproximação do erro em f, desde que o erro em x seja suficientemente pequeno.

A eq. (6) é a **lei da propagação dos erros** (l.p.e.) relativa a uma função de **uma só variável**, e evidencia que o erro em f é proporcional ao erro em x, sendo a constante de proporcionalidade a 1.ª derivada de f. A eq. (6) só é exacta se $f(x)$ for uma função linear em x. Para funções não lineares a equação só dará uma boa aproximação do erro em f se Δx for muito pequeno, requisito que é frequentemente satisfeito.

Para o caso de uma função linear, como por exemplo,

$$f(x) = a + b\,x \qquad (7)$$

a eq.(6) reduz-se a,

$$\Delta f = b\Delta x \qquad (8)$$

que dá um valor exacto do erro Δf, qualquer que seja o erro em x, mesmo grande; nesta situação, o erro em $f(x)$ é exactamente proporcional ao erro em x.

A figura 1, mostra a relação entre o erro em $f(x)$ e a aproximação obtida pelo l.p.e., no caso de a função não ser linear. A aproximação introduzida pela eq. (6) é obtida usando a tangente à curva em vez da própria curva, e é tanto mais realista quanto menor for o valor de Δx.

O exemplo 2 ilustra a aplicação de l.p.e. a uma função de uma só variável.

Exemplo 2. Determine o erro na área de um círculo de raio, r = 2.54 m, sabendo que o erro que lhe está implícito é Δr = 0.03 m.O valor da área do círculo é,

$$A = \pi r^2 = \pi x 2.54^2$$
$$= 20.268 m^2$$

Aplicando a lei de propagação de erro e tendo em atenção que a área não é uma função linear de r, obtém-se:

$$\Delta A \approx \left(\frac{\partial A}{\partial r}\right)\Delta r \approx 2\pi r \Delta r$$

Substituindo valores nesta última expressão,

$$\Delta A \approx 2\pi x 2.54 x 0.03$$
$$\approx 0.5 m^2$$

obtemos uma estimativa do erro que afecta a área. Podemos agora escrever o número de algarismos significativos correctos no valor desta grandeza, assim como do erro de que está afectada:

$$A = 20.3 \pm 0.5\ m^2$$

Vamos agora considerar o caso de a função f ser uma função de várias variáveis

$$f = f(x_1, x_2, \ldots x_k), \qquad (9)$$

Que apresentam erros Δx_i. O erro em f devido aos erros Δx_i pode ser calculado pela expressão:

$$\Delta f \approx \Delta x_1\left(\frac{\partial f}{\partial x_1}\right) + \Delta x_2\left(\frac{\partial f}{\partial x_2}\right) + \ldots + \Delta x_k\left(\frac{\partial f}{\partial x_K}\right) \qquad (10)$$

Esta última equação é denominada por lei de progação dos erros para uma função de multivariáveis. O valor de Δf será tanto mais correcto quanto menores forem os erros Δx_i. No caso de f ser linear em $(x_1, x_2, \ldots x_k)$, a expressão (10) é exacta, como já se referiu. A eq. (10) aparece por vezes com esta forma,

$$\Delta f \approx \Delta x_1\left|\frac{\partial f}{\partial x_1}\right| + \Delta x_2\left|\frac{\partial f}{\partial x_2}\right| + \ldots + \Delta x_k\left|\frac{\partial f}{\partial x_k}\right| \qquad (11)$$

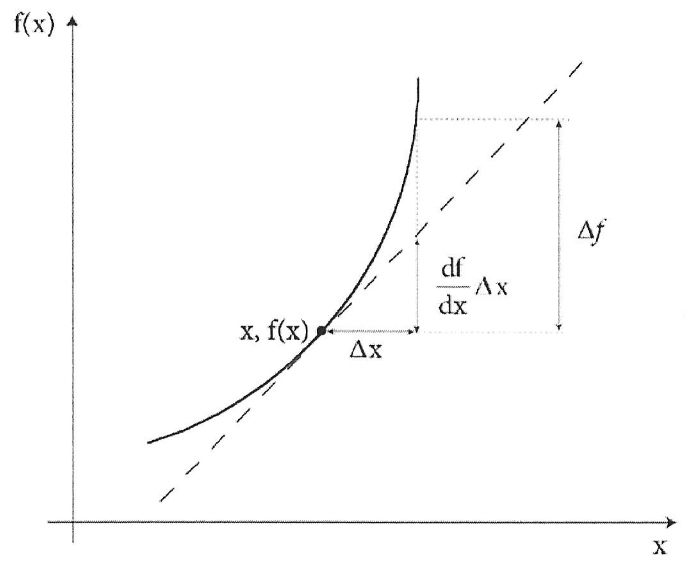
que constitui um limite superior do erro em f.

Vamos agora analisar algumas situações particulares em que não é possível na prática aplicar a eq. (10):

(a) Caso em que não é possível determinar as derivadas parciais presentes na eq.(10), uma vez que a variável f não é explicitável da expressão que a relaciona com x_i.

(b) A dependência de f das variáveis x_i não é traduzida apenas por uma equação, mas sim por um conjunto de equações dependentes, o que torna di-

Figura 1 *Relação entre os erros em x e em f(x) e a aproximação dada por* $\Delta f \approx \frac{df}{dx}\Delta x$, *eq.(6). Δx representa o erro em x e Δf o erro em f(x).*



fácil e muitas vezes impossível aplicar a eq. (10).

(c) A equação que relaciona f com as variáveis x_i é desconhecida

O problema poderá ser facilmente resolvido nos dois primeiros casos aplicando o método directo de cálculo dos erros que é válido qualquer que seja a dimensão dos erro. O procedimento deve ser o que se indica a seguir:

(i) calcula-se o valor f_0 de f para os valores de x_i, recolhidos experimentalmente.

(ii) calculam-se k valores de f definidos por:

$$f_i = f(x_1, x_2, \ldots\ldots x_i + \Delta x_i, \ldots. x_k), \qquad (12)$$

Onde a variável de ordem i é acrescentada do seu erro (que pode ser um valor positivo ou negativo).

(iii) determina-se o erro em f pela expressão,

$$\Delta f = \sum_i^k (fi - fo) \qquad (13)$$

Se os erros forem pequenos, este método conduz ao mesmo valor que a aplicação da lei de propagação dos erros, eq.(10). No caso de os erros já serem significativos o método directo permite obter estimativas do erro mais realistas.

A situação (c) é a mais complexa, podendo surgir numa experiência em que se determina f e x_i, que se sabem serem variáveis dependentes, embora se desconheça a forma funcional dessa dependência. Neste caso, podemos obter o erro Δf devido aos erros Δx_i afectando experimentalmente (se isso for possível) o valor de cada variável, x_i, individualmente, do respectivo erro e anotar o efeito produzido no valor médio da variável f. O erro total, Δf, poderá então ser calculado pela eq. (13). De qualquer modo, convém referir que cada situação deste tipo apresentará características particulares cabendo ao experimentalista imaginar e delinear o melhor procedimento que lhe permita fazer uma boa estimativa do erro na variável dependente.

O exemplo 3 reporta-se à aplicação do método do cálculo directo dos erros a uma função de duas variáveis contendo incertezas conhecidas, sendo o resultado comparado com o obtido pela lei de propagação dos erros.

Exemplo 3. Determine uma estimativa do erro que afecta o valor da massa específica de um determinado composto sabendo que 10.5276 g do composto ocupa o volume de 5.394 cm³, a uma dada temperatura e pressão. Os erros que afectam estas grandezas são, respectivamente; Δm= 0.0004 g e ΔV= - 0.003 cm³.

Vamos resolver o problema por aplicação do método directo de cálculo dos erros e comparar o resultado com o obtido pela l.p.e.

Por aplicação directa da definição de massa específica obtém-se,

$$\rho_o = \frac{m}{V} = \frac{10.5276}{5.394} = 1.952 \text{ g m}^{-3}$$

Vamos então afectar m e V dos respectivos erros e calcular o valor da massa específica afectada do erro.

$$\rho_1 = \frac{10.5276 + 0.0004}{5.394 - 0.003} = 1.953 \text{ g cm}^{-3}$$

Assim o erro $\Delta\rho$ é,

$$\Delta\rho = \rho_1 - \rho_0 = 1.953 - 1.952 = 0.001\ g\ cm^{-3}$$

Vamos agora proceder à determinação de $\Delta\rho$ pela l.p.e., eq. (10):

$$\Delta\rho = \left(\frac{\partial\rho}{\partial m}\right)\Delta m + \left(\frac{\partial\rho}{\partial V}\right)\Delta V = \left(\frac{1}{V}\right)\Delta m + \left(-\frac{m}{V^2}\right)\Delta V$$

$$= \frac{1}{5.394}(0.0004) - \frac{10.5276}{5.394^2}(-0.003) = 0.001\ g\ cm^{-3}$$

Neste exemplo a eq.(10) dá uma estimativa correcta de $\Delta\rho$ devido ao facto de os erros Δm e ΔV serem pequenos.

3.2. Lei de propagação das variâncias (l.p.v.)

Vamos agora considerar uma variável G que é função das variáveis x e y:

$$G = g\,(x,y) \qquad (14)$$

Efectuaram-se medições das grandezas x e y repetidas vezes, obtendo-se um conjunto de n pares de valores, $(x_1,y_1),\ldots,(x_n, y_n)$; a partir dos valores de x_i determinaram-se a média, $\bar{x}$, e o desvio padrão, σ_x, da forma habitual e a partir dos valores de y_i calcularam-se também $\bar{y}$ e σ_y. Pretende-se determinar a incerteza no valor de G devido aos erros aleatórios, ou seja, o desvio padrão, σ_G. Vamos admitir que o valor mais provável de G é,

$$\bar{G} = g\left(\bar{x}, \bar{y}\right) \qquad (15)$$

e que a incerteza em G é portanto traduzida por,

$$\sigma_G^{\ 2} = \frac{1}{n}\sum\left(G_i - \bar{G}\right)^2 \qquad (16)$$

em que $G_i = g\,(x_i, y_i)$ é o valor de G no ponto de coordenadas (x_i, y_i). σ_{G^2} designa-se por variância de G, como é sabido. A aplicação da lei de propagação dos erros eq. (10), permite-nos obter o erro $(G_i - \bar{G})$ em função dos erros $(x_i - \bar{x})$ e $(y_i - \bar{y})$.

Por uma questão de simplificação de escrita considerámos no denominador de eq. (16) n em vez de $(n-1)$.

$$\left(G_i - \bar{G}\right) \approx \left(x_i - \bar{x}\right)\left(\frac{\partial G}{\partial x}\right) + \left(y_i - \bar{y}\right)\left(\frac{\partial G}{\partial y}\right) \tag{17}$$

A derivada $\left(\frac{\partial G}{\partial x}\right)$ será determinada a partir dos valores médios das variáveis x e y, considerando y fixo e igual ao seu valor médio. Substituindo a eq. (17) na expressão (16) obtém-se a variância, σ_G^2,

$$\sigma_G^2 \approx \frac{1}{n}\sum_{i=1}^{n}\left[\left(x_i - \bar{x}\right)\left(\frac{\partial G}{\partial x}\right) + \left(y_i - \bar{y}\right)\left(\frac{\partial G}{\partial x}\right)\right]^2$$

$$\approx \frac{1}{n}\sum_{i=1}^{n}\left[\left(x_i - \bar{x}\right)^2\left(\frac{\partial G}{\partial x}\right)^2 + \left(y_i - \bar{y}\right)^2\left(\frac{\partial G}{\partial y}\right)^2 + 2\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)\left(\frac{\partial G}{\partial x}\right)\left(\frac{\partial G}{\partial y}\right)\right] \tag{18}$$

Os dois primeiros termos do segundo membro da eq. (18) podem ser expressos em termos das variâncias σ_x^2 e σ_y^2, que se definem pelas expressões seguintes:

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \tag{19}$$

e

$$\sigma_y^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 \tag{20}$$

O terceiro termo da eq. (18) representa a **covariância**, σ_{xy}, das variáveis x e y e tem uma definição semelhante à das eqs. (19) e (20),

$$\sigma_{xy} = \frac{1}{n}\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) \tag{21}$$

Tendo em conta as equações (19), (20) e (21), a variância em G, σ_G^2 pode ser expressa pela seguinte equação,

$$\sigma_G^2 \approx \sigma_x^2\left(\frac{\partial G}{\partial x}\right)^2 + \sigma_y^2\left(\frac{\partial G}{\partial y}\right)^2 + 2\sigma_{xy}\left(\frac{\partial G}{\partial x}\right)\left(\frac{\partial G}{\partial y}\right) \tag{22}$$

que se designa por lei de propagação das variâncias (ou lei da propagação dos desvios padrões). A contribuição mais significativa para σ_G^2 é devida aos dois primeiros termos da eq. (22). O terceiro termo é uma média dos termos cruzados que envolvem produtos dos desvios em x, $(x_i - \bar{x})$, e em y, $(y_i - \bar{y})$. Se as medições de x e de y forem independentes, podemos concluir que depois de efectuarmos muitas medições, a covariância σ_{xy} se aproxima do valor zero. Qualquer que seja o valor de y_i, a quantidade $(x_i - x)$, tem a mesma probabilidade de ser positiva ou negativa, o que determina que ao fim de um número infinito de medições os termos negativos e positivos se cancelem. Na prática o número de medições é finito e por isso σ_{xy} não será exactamente zero, mas sim um valor muito pequeno se os erros em x e y forem independentes e aleatórios. Nesta situação diz-se que as incertezas em x e y não estão correlacionadas, e a expressão (22) transforma-se em:

$$\sigma_G^2 \approx \sigma_x^2\left(\frac{\partial G}{\partial x}\right)^2 + \sigma_y^2\left(\frac{\partial G}{\partial y}\right)^2 \tag{23}$$

que é a lei da **propagação das variâncias** no caso das variáveis x e y serem independentes.

A seguir apresenta-se um exemplo de aplicação da lei de propagação das variâncias na forma (23).

Exemplo 4. Para determinar a área de um espelho rectangular mediram-se seis vezes o comprimento e seis vezes a largura tendo-se obtido os resultados seguintes:

***Tabela 1** Dimensão de um espelho rectangular em cm.*

c	125.5	125.4	125.3	125.4	125.6	125.4
l	60.8	60.7	60.5	60.8	60.4	60.6

Determine a incerteza da área do espelho.

Vamos determinar em primeiro lugar, as melhores estimativas das dimensões do espelho e as respectivas incertezas.

A melhor estimativa do comprimento é dada pelo valor médio das várias observações:

$$\bar{c} = \frac{\sum c_i}{n}$$

$$= 125.43\text{cm}$$

A incerteza no valor médio, ou seja, o desvio médio, pode determinar-se pela equação:

$$\sigma_{\bar{c}} = \frac{\sigma_c}{\sqrt{n}}$$

$$= \frac{1}{\sqrt{n}}\left[\frac{\sum_{i=1}^{n}\left(c_i - \bar{c}\right)^2}{n}\right]^{1/2}$$

$$= \frac{1}{\sqrt{n}}\left[\frac{0.0604}{6}\right]^{1/2}$$

$$= 0.04\ cm$$

logo,

$\bar{c}$ = *125.43 ± 0.04cm*

Do mesmo modo se obtém:

$\bar{l}$ = *60.63 ± 0.07cm*

O valor da área pode então ser calculado:

$\bar{A} = \bar{c} \times \bar{l}$ = *125.43 x 60.63*

= *7605 cm²*

Aplicando a lei de propagação das variâncias, eq. (23).

$$\sigma_{\bar{A}} = \left[\sigma_{\bar{c}}^2\left(\frac{\partial \bar{A}}{\partial c}\right)^2 + \sigma_{\bar{l}}^2\left(\frac{\partial \bar{A}}{\partial l}\right)^2\right]^{1/2}$$

$$= \left(\sigma_{\bar{c}}^2\,\bar{l}^2 + \sigma_{\bar{l}}^2\,\bar{c}^2\right)^{1/2}$$

$$= \left(0.04^2 x\, 60.63^2 + 0.07^2 x\, 125.43^2\right)^{1/2}$$

$$= 9\ cm$$

Devemos portanto escrever os valores da área e respectiva incerteza da seguinte forma:

$$A = 7605 \pm 9\ cm$$

É evidente que o problema anterior poderia ter sido resolvido determinando o produto de cada valor do comprimento pela respectiva largura, obtendo-se assim um primeiro valor A .

Procedendo de igual modo para os restantes poderíamos no fim determinar $\bar{A}$ e depois a respectiva incerteza , $\sigma_{\bar{A}}$. O resultado obtido seria muito semelhante ao encontrado.

No caso de as medições de x e y não serem independentes, isto é, as incertezas em x e em y estiverem correlacionadas, a covariância pode não ser zero. É fácil de imaginar uma situação em que o valor de x medido e o correspondente valor de y estejam ambos subestimados, o que implica que $(x_i - \bar{x})$, e $(y_i - \bar{y})$ apresentam ambos o mesmo sinal e, por conseguinte, o seu produto seja sempre positivo. Podemos também considerar que as diferenças anteriores têm sinais opostos (uma das grandezas é sobrestimada e a outra é subestimada) o que determina que neste caso a covariância seja sempre negativa. Quando σ_{xy} não é zero, dizemos que os erros de x e y estão correlacionados e nesta situação temos que usar a expressão da lei de propagação das variâncias na forma (22) para obter um valor mais realista da incerteza σ_G. O exemplo 5 é demonstrativo duma situação em que a covariância de duas variáveis não deve ser desprezada, permitindo comparar, as incertezas obtidas pelas duas formas da lei da propagação das variâncias, equações (22) e (23).

Exemplo 5. As dimensões de uma mesa rectângular foram medidas por cinco experimentalistas tendo sido obtidos os valores apresentados da Tabela 2. Pretende-se determinar o perímetro da mesa e a respectiva incerteza.

Podemos calcular o perímetro da mesa aplicando a expressão que se indica a seguir. A incerteza em p pode ser obtida pela expressão (22). Assim,

$$p = 2x\left(\bar{c}+\bar{l}\right)$$
$$= 2\,x(123+52)$$
$$= 350\ \text{cm}^2$$

Utilizando os valores da tabela podemos determinar,

$$\sigma_c^2 = 2.0\ cm^2$$

e $\sigma_l^2 = 3.2\ cm^2$,

por aplicação da expressão (19).

Podemos também observar que os valores mais elevados de c estão correlacionados com os valores mais baixos de l e vice-versa, porque as diferenças $(c-\bar{c})$ e $(l-\bar{l})$ têm sempre sinais opostos. Esta correlação implica que os produtos $(c-\bar{c})(l-\bar{l})$ da coluna mais à direita da Tabela 2 são sempre negativos ou zero. Por conseguinte, a covariância definida por (21) é negativa:

$$\sigma_{cl} = \frac{1}{n}\sum_{i=1}^{n}\left(c-\bar{c}\right)\left(l-\bar{l}\right)$$
$$= \left(\frac{1}{5}\right)(-12)$$
$$= -2.4\ cm^2$$

Podemos agora determinar σ_p pela expressão (22) tendo em atenção que

$$\left(\frac{\partial p}{\partial c}\right) = \left(\frac{\partial p}{\partial l}\right) = 2.$$

Deste modo,

$$\sigma_p = \left[\sigma_c^2\left(\frac{\partial p}{\partial c}\right)^2 + \sigma_l^2\left(\frac{\partial p}{\partial l}\right)^2 + 2\sigma_{cl}\left(\frac{\partial p}{\partial c}\right)\left(\frac{\partial p}{\partial l}\right)\right]^{1/2}$$

$$= (2 \times 2^2 + 3.2 \times 2^2 - 2 \times 2.4 \times 2 \times 2)^{1/2}$$

$$= 1\ \text{cm}$$

Se considerarmos incorrectamente que as incertezas em c e l não estão correlacionadas, ou seja, se tratarmos as medições do comprimento e da largura da mesa como independentes, então para determinarmos σ_p devemos usar a expressão (23) :

$$\sigma_p = \left[\sigma_c^2\left(\frac{\partial p}{\partial c}\right)^2 + \sigma_l^2\left(\frac{\partial p}{\partial l}\right)^2\right]^{1/2}$$
$$= \left(2 \times 2^2 + 3.2 \times 2^2\right)^{1/2}$$
$$= 4\ cm$$

Com este exemplo podemos constatar que a correlação das incertezas nas variáveis medidas pode originar uma diferença significativa na incerteza que se propagou. As incertezas positivas que se observam em c são acompanhadas por incertezas negativas em l e vice-versa. Por esta razão quando se adicionam c e l as incertezas tendem a cancelar-se resultando uma incerteza no perímetro de apenas 1 cm. O recurso à l.p.v. na forma (23) sobrestima a incerteza no valor da área e portanto não deverá ser utilizada neste caso.

Tabela 2 *Resultados das medições das dimensões de uma mesa rectângular (em cm) obtidos por cinco observadores.*

	c	l	$(c-\bar{c})$	$(l-\bar{l})$	$(c-\bar{c})(l-\bar{l})$
A	125	50	2	-2	-4
B	121	55	-2	3	-6
C	123	51	0	-1	0
D	122	53	-1	1	-1
E	124	51	1	-1	-1

Conclusão

O tratamento dos erros de carácter sistemático é um processo complexo que exige que se identifique a priori a sua origem. Estes erros poderão ser eliminados (total ou parcialmente) pela introdução de factores correctivos ou ensaios em branco.

Os erros aleatórios porque possuem natureza indeterminada só podem ser detectados pela repetição das experiências e subsequente análise estatística. Frequentemente o resultado final de uma experiência é calculado a partir de grandezas obtidas experimentalmente. Por esta razão, a determinação do erro do resultado final é função dos erros das observações directas. As leis de propagação dos erros e das variâncias permitem determinar como as incertezas das grandezas medidas se propagam para o resultado final.

Em alguns casos, não é possível recorrer a estas leis havendo então necessidade de usar o método directo de determinação de incertezas.

Bibliografia

Aconselha-se aos alunos a leitura dos seguintes livros:

L. Lyons, *A Practical Guide to Data Analysis for the Physical Sciences*, Cambridge University Press, 1991.

P.R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, 1969.

G.F.P. Box, W.G. Hunter, J.S. Hunter, *Statistics for Experiments*, John Wiley, 1978

Para os professores sugere-se também a leitura dos seguintes artigos que abordam os assuntos com mais profundidade:

W.J. Youden, How to evaluate accuracy, Mat. Res. & Std. **1** (1961) 268.

C. Eisenhart, *Realistic evaluation of the precision and accuracy of instrument calibration systems*, J. Research, N.B.S., **67c** (1963) 161.