

(A) Tratamento de dados de estatística de contagens

1. Uma distribuição $P(x)$ pode ser caracterizada pelo seu valor médio $\bar{x}$ e a sua dispersão σ (desvio padrão).

Só um número infinito de amostras ($N \rightarrow \infty$) de $P(x)$ nos permitiria obter $\bar{x}$ sem erro.

Se obtivermos uma só amostra, a cada valor x_i (canal i num multicanal) corresponde um número de contagens n_i . Mas, se se obtivessem N amostras da distribuição $P(x_i)$, as contagens n_i do canal i $n_i, n'_i, n''_i, \dots$ flutuariam em torno de um valor médio $\bar{n}_i$ com uma dispersão gaussiana: $\sigma = \sqrt{\bar{n}_i}$ (independentemente de $P(x_i)$ ser ou não gaussiana).

Portanto, usando uma só amostra, as contagens n_i do canal i são o estimador do valor médio desse canal $\bar{n}_i$ com um erro $\sqrt{n_i}$:

$$\text{canal } i \rightarrow n_i \pm \sqrt{n_i}$$

2. Suponhamos agora que a distribuição $P(x)$ é um pico isolado e nítido dum espectro de energia num multicanal. Para se obter o seu valor médio (ou centróide, ou baricentro) deve calcular-se:

$$\begin{cases} A = \sum_i n_i & (A \equiv \text{área do pico}) \\ \bar{c}_i = \frac{\sum_i c_i n_i}{\sum_i n_i} & (c_i \equiv n_i \text{ do canal } i) \end{cases}$$

A largura do pico, $\sigma_{\text{distr.}}$, permanece constante por mais contagens que se façam. Ela depende principalmente da dispersão que os detectores introduzem. O que se reduz ao fazermos $N \rightarrow \infty$ é o erro associado ao valor médio $\bar{c}_i$:

$$\sigma_{\bar{c}_i} = \frac{\sigma_{\text{distr.}}}{\sqrt{A}}$$

(ver dedução nos acetatos da aula teórica)

3. • Para obtermos o desvio padrão da distribuição, $\sigma_{\text{distr.}}$, parte-se do canal médio $\bar{c}_i$ e procuram-se os canais (e pico) c_1 e c_2 com contagens $s_1 = s_2 = \frac{1}{2}$ (em que s são as contagens do canal médio).

Então, a largura a meia altura do pico é:

$$\text{FWHM} = c_2 - c_1$$

e o desvio padrão da distribuição virá:

$$\sigma_{\text{distr.}} = \text{FWHM} / 2.355$$

• Este método prático é pouco preciso quando o pico cresce em poucos canais (Ex.: picos do "pulser"). Nesse caso, pode usar-se a definição de desvio-padrão:

$$\sigma_{\text{distr.}} = \sqrt{\frac{\sum_i (c_i - \bar{c})^2 n_i}{\sum_i n_i}}$$

③ Tratamento de dados experimentais de medições

No tratamento de medições baseadas em escalas graduadas (Ex: comprimentos, tempos, temperaturas, correntes, tensões, ...) o objectivo é obter um valor médio e o seu erro associado:

$$\Rightarrow \bar{x} \pm E_{\bar{x}} \quad \text{com } \bar{x} = \sum_{i=1}^n \frac{x_i}{n} \quad (\text{média dos } n \text{ ensaios})$$
$$e \quad E_{\bar{x}} = \begin{cases} = \sigma = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n}} & \begin{array}{l} \text{se } n \geq 10 \text{ ensaios} \\ \text{(desvio quadrático médio)} \\ \text{ou} \\ \text{(desvio padrão)} \end{array} \\ = \sum_{i=1}^n \frac{|x_i - \bar{x}|}{n} & \text{se } n < 10 \text{ ensaios} \\ & \text{(desvio médio)} \end{cases}$$

Quer dizer: - se o nº de ensaios n para a medição de uma grandeza já for um conjunto estatístico ($n \geq 10$) usa-se como estimador do erro a dispersão σ .

- se o nº de ensaios n é pequeno demais ($n < 10$) usa-se como estimador o desvio médio.

NOTA: Deve sempre estimar-se para além das divisões mais precisas do instrumento de medida que se está a usar. É a dispersão dos ensaios de medida ($i=1, \dots, n$) que surge a média e uma estimativa do erro (dada por um dos estimadores indicados) que já inclui a precisão do instrumento. Usar $\frac{1}{2}$ menor divisão da escala só se justificaria se só se fizesse um ensaio! E, neste caso, estar-se-ia a substituir o erro estatístico (que não avaliamos) por um erro sistemático. Ora estes erros devem ser sempre avaliados e reportados separadamente. Ou seja: $\bar{x} \pm E_{\text{estat}} \pm E_{\text{sistem.}}$

© Propagação de erros

1.

- Se todas as variáveis medidas representarem conjuntos estatísticos deve usar-se a expressão de propagação quadrática. Se, por exemplo, $z = f(x, y)$ vem:

$$\varepsilon_z^2 = \left(\frac{\partial f}{\partial x}\right)^2 \varepsilon_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \varepsilon_y^2 \quad \left(\begin{array}{l} \text{se } x \text{ e } y \text{ são} \\ \text{independentes} \end{array}\right)$$

$$\Rightarrow \varepsilon_z = \sqrt{\varepsilon_z^2}$$

- Se pelo menos uma variável não representar conjunto estatístico (isto é n.º ensaios $n < 10$) tem de usar-se a expressão de propagação linear:

$$\varepsilon_z = \frac{\partial f}{\partial x} \varepsilon_x + \frac{\partial f}{\partial y} \varepsilon_y \quad (x \text{ e } y \text{ independentes})$$

2. ► As variáveis x e y já são resultado de uma média de n ensaios e já têm pois o seu erro (estatístico):

$$\Rightarrow \bar{x} \pm \varepsilon_x \quad \text{e} \quad \bar{y} \pm \varepsilon_y$$

- As funções $\left(\frac{\partial f}{\partial x}\right)_x$ e $\left(\frac{\partial f}{\partial y}\right)_y$ devem ser calculadas nos pontos médios (só nesses pontos a expressão tem sentido físico). Usar outros pontos no intuito de majorar a propagação é querer confundir erro estatístico c/ erro sistemático. Ora este deve ser estimado e referido à parte:
- $$\Rightarrow \bar{z} \pm \varepsilon_{\text{estat}} \pm \varepsilon_{\text{sist.}} \quad \text{e/} \quad \bar{z} = f(\bar{x}, \bar{y})$$

• Aplicação ao ajuste dos mínimos quadráticos linear

Se a expressão teórica se puder exprimir na forma $y_{\text{teor}} = ax + b$, o estimador χ^2 escreve-se:

$$\chi^2 = \sum_i^n \frac{(y_i - ax_i - b)^2}{\sigma_i^2}$$

Preende minimizar-se a função χ^2 em ordem aos coeficientes a, e b, simultaneamente:

$$\begin{cases} \frac{\partial \chi^2}{\partial a} = \dots = 0 \\ \frac{\partial \chi^2}{\partial b} = \dots = 0 \end{cases}$$

$\Rightarrow$ de modo a extrair-se os valores de a e b e estimar os erros respectivos quadraticamente

Calculando, obtém-se:

$$\begin{cases} a = \frac{c_1 c_5 - c_3 c_4}{\Delta} \\ b = \frac{c_2 c_4 - c_1 c_3}{\Delta} \end{cases}, \quad \begin{cases} \sigma_a^2 = \frac{c_5}{\Delta} \\ \sigma_b^2 = \frac{c_2}{\Delta} \end{cases}$$

com: $\Delta = c_2 c_5 - c_3^2$ e:

$$\begin{cases} c_1 = \sum_i^n \frac{x_i y_i}{\sigma_i^2} \\ c_2 = \sum_i^n \frac{x_i^2}{\sigma_i^2} \\ c_3 = \sum_i^n \frac{x_i}{\sigma_i^2} \\ c_4 = \sum_i^n \frac{y_i}{\sigma_i^2} \\ c_5 = \sum_i^n \frac{1}{\sigma_i^2} \end{cases}$$

Análise dimensional

Ex.: $[x] = [y] = L = [\sigma_x] = [\sigma_y]$

$\therefore [b] = L ; [a] = 1$

Logo: $[\sigma_b^2] = \frac{[c_2]}{[\Delta]}$

$L^2 = \frac{1}{L^{-2}}$

Acumulação de dados obtidos nas mesmas condições

Tratamento de n amostras sequenciais

- Havendo n amostras, faz-se a soma das obtidas em sequência:

$$\begin{cases} N_1 \pm E_{N_1} & E_{N_1} = \sqrt{N_1} \\ \vdots \\ N_n \pm E_{N_n} & E_{N_n} = \sqrt{N_n} \end{cases}$$

($E_1, \dots, E_n$ são $\sigma_1, \dots, \sigma_n$ de distr. normais)

⇒ Soma das amostras com erro global:

$$N = \sum_{i=1}^n N_i \quad \text{e} \quad E_N = \sigma_N = \sqrt{N}$$

A soma das amostras e propagação do erro (quadrático) é:

$$\begin{aligned} \sigma_N^2 &= \sigma_1^2 + \dots + \sigma_n^2 = N_1 + \dots + N_n = N \\ \Rightarrow \sigma_N &= \sqrt{N} \end{aligned}$$

ou seja, dá exatamente a mesma que a avaliação global do erro: $N \pm \sqrt{N}$

Comparação entre o resultado experimental e resultados anteriores

- As seguintes definições não são necessárias:

Para uma grandeza x ($x_{\text{tabelado}} < x_{\text{medido}}$):

- desvio à exatidão:

$$\frac{\Delta x}{x} \quad (\Delta x = x_{\text{tab}} - x_{\text{medido}})$$

- desvio à precisão:

$$\frac{\sigma_x}{x} \quad (\equiv \eta_x \text{ erro relativo})$$

(Nota: os denominadores podem $\bar{n}$ ser ≈ 5)

➤ ~~Outra~~ **Do** nosso ponto de vista de físicos experimentais, basta usar uma grandeza $\bar{q}$ aglomera as duas:

$$\boxed{\frac{\Delta x}{\sigma_x} \equiv n^\circ \sigma} \left(\begin{array}{l} \text{afastamento} \\ \text{valor tab.} < \text{valor} \\ \text{medido} \end{array} \right) \text{ em } \underline{n^\circ \text{ de } \sigma}$$

Combinação de diferentes resultados experimentais

Se possuírmos N diferentes amostras $x_1, x_2, \dots, x_N$ da mesma distribuição, de desvios padrão $\sigma_1, \sigma_2, \dots, \sigma_N$ (os métodos experimentais podem ser diferentes), para obtermos a média das médias das amostras devemos usar a

média ponderada:
$$\bar{x} = \frac{\sum_i^N x_i / \sigma_i^2}{\sum_i^N 1 / \sigma_i^2}$$

e a sua

variância:
$$\sigma^2(\bar{x}) = \frac{1}{\sum_i^N 1 / \sigma_i^2}$$

$\Rightarrow$ Deve dar-se menos peso às amostras com maiores dispersões (medidas com instrumentos menos precisos).

A variância $\sigma^2(\bar{x})$ obtém-se directamente da média ponderada, aplicando a esta a expressão quadrática de propagação dos erros.

Se todas as amostras têm a mesma dispersão σ , vem

$$\bar{x} = \frac{\sum_i^N x_i}{N}$$

$$\sigma^2(\bar{x}) = \frac{\sigma^2}{N}$$

$$\text{ou} \quad \sigma(\bar{x}) = \frac{\sigma}{\sqrt{N}}$$