**Lezione propedeutica al Laboratorio di Chimica Fisica I.**

**Teoria degli errori e uso di Excel**

**INTRODUZIONE:** una quantità fisica è sempre espressa da un **numero** con il corrispondente **errore** e da un’**unità di misura**. L’errore risulta importante perché dà un’indicazione del numero di **cifre significative** che bisogna tenere. Per esempio:

$$C=\frac{n}{V}=\frac{2.32∙10^{-2} mol}{0.51 L}=4.54901908∙10^{-2} M=(4.55\pm 0.01)∙10^{-2} M$$

Il numero 12000 ha solo due cifre significative. Invece, se scriviamo 12000.00 allora le cifre significative saranno tutte quelle scritte prima e dopo la virgola (7 nel caso specifico).

Nel caso, poi, degli zeri essi sono cifre significative solo se non sono all’inizio del numero (es.: 0.0059 ha 2 cifre significative; 5.9000 ne ha 5).

**Es.:** prendendo in considerazione l’**esperienza del VBC**, si otterranno dei valori di coefficiente di estinzione molare dell’ordine di (12580 ± 3740) dm3 mol-1 cm-1 che dovranno essere riportati come (13000 ± 4000) dm3 mol-1 cm-1.

Oppure, nel caso si debba riportare un valore molto piccolo come (0.000162 ± 0.00005) M, potrà essere riscritto come (0.00016 ± 0.00005) M o (16 ± 5)⋅10-4 M.

Ultimo esempio, nella **CMC conduttimetrica**, si possono avere delle pendenze (date da Excel) di circa (65.77561233 ± 0.62649162) che dovranno essere riportate come (65.8 ± 0.6).

**N.B.:** per le approssimazioni da 0 a 4, si arrotonda la cifra per difetto, da 5 a 9 per eccesso. Si approssima sempre a partire dal numero di partenza, non dall’approssimazione precedente (es.: 11.3456).

→ **Esercizio 1a (esercizio su cifre significative)**

L’errore sul valore di una grandezza si può esprimere in diversi modi: noi definiamo ora **l’errore relativo** percentuale e poi più avanti definiremo la deviazione standard (questi due errori valutano cose diverse)

$$E\_{r}= \frac{x\_{i}-x\_{esatto}}{x\_{esatto}}×100$$

→ **Esercizio 1b (esercizio su errore relativo)**

**CONCETTO DI ERRORE:** in ambito scientifico, se si ha un **errore sistematico** molto spesso l’esperimento stesso è pregiudicato; se si ha un **errore statistico**, invece, si ha l’impossibilità di ottenere lo stesso identico valore di una grandezza per misurazioni ripetute. Mediante una trattazione di tipo statistico è possibile conoscere il range di validità dei risultati affetti da questo secondo tipo di errore. Quindi, ipotizzando di avere molte misure (ripetizioni) di una data grandezza, sarà *y* il **valore più rappresentativo** (più ricorrente) cioè quello che minimizza la somma degli scarti (scarto = quadrato delle differenze tra il valore *y* e le mie N misure *xi*). Ricordando l’equazione che definisce lo scarto:

$$scarti=\sum\_{i=1}^{N}(x\_{i}-y)^{2}$$

Per trovare il minimo della funzione scarti faccio la derivata rispetto a *y* e la eguaglio a zero così trovo *y*:

$$\frac{d(scarti)}{dy}=-2\sum\_{i=1}^{N}(x\_{i}-y)=-2\sum\_{i=1}^{N}x\_{i}+N2y=0$$

$$y=\frac{\sum\_{i=1}^{N}x\_{i}}{N}=\overbar{x}$$

Una volta che ho trovato il valore più rappresentativo (che ho dimostrato essere il **valore medio**) devo stimare un **intervallo di confidenza**. Per fare questo è utile definire i concetti di **precisione** e **accuratezza**.



Preciso ma non accurato

Preciso e accurato

Né preciso né accurato

Accurato ma non preciso

Quindi, l’**accuratezza** dà un’indicazione di quanto il valore medio sia vicino al valore vero, la si può misurare mediante l’**errore relativo** percentuale. La **precisione**, invece, dice quanto i valori misurati siano vicini tra di loro (misurata mediante la **deviazione standard**, σ)

Quale errore dovrà essere fornito per il laboratorio? Erel% (accuratezza) o sigma (precisione)?

La deviazione standard è la radice dello scarto medio (in cui al posto di y si mette x medio, a denominatore si mette N-1 quando ho tante misure, in genere più di 20; ad ogni modo, la funzione dev.st di Excel calcola sempre con N-1).

$$σ=\sqrt{\frac{\sum\_{i=1}^{N}(x\_{i}-\overbar{x})^{2}}{N-1}}$$

Dato che gli errori di tipo statistico sono casuali (seguono una distribuzione gaussiana), la probabilità che una nuova misura cada nell’intervallo tra il valor medio **–σ e il valor medio +σ è del 68.27%** (questo dovrà essere preso in considerazione nelle esperienze di laboratorio) e è prassi scrivere una grandezza sperimentale come **𝑥̅ ± 𝜎**. Nell’**intervallo tra il valor medio –3σ e il valor medio +3σ cade invece il 99.73%** dei valori. Se si hanno valori al di fuori di questo intervallo dovranno essere scartati in quanto statisticamente errati.

→ **Esercizio 2 (esercizio su valore medio e deviazione standard)**

**PROPAGAZIONE DELL’ERRORE:** quando si calcola il valore di una grandezza a partire da altre grandezze affette da errore, l’errore di questa sarà dipendente da tutti gli errori delle grandezze rientrate nel calcolo, in altre parole gli errori si propagano.

ESEMPIO: **C = n/V**. L’errore sulla pesata dà un errore su n mentre quello del matraccio su V. Quindi, si deve calcolare il ΔC dato da Δn e il ΔC dato da ΔV:

*i*) errore dato dall’errore sulla pesata (PM è esatto)

$$C\pm ΔC=\frac{n\pm Δn}{V}=\frac{n}{V}\pm \frac{Δn}{V} \rightarrow ΔC=\frac{Δn}{V}$$

*ii*) errore dato dall’errore sul volume

$$C\pm ΔC=\frac{n}{V\pm ΔV}$$

non posso dividere tra C e ΔC ma posso scrivere:

$$C\pm ΔC=\frac{n}{V\pm ΔV}=\frac{n}{V(1\pm \frac{ΔV}{V})}$$

e considerando che ΔV << V e che 1/(1+x) ≈ 1-x per x << 1, allora:

$$C\pm ΔC=\frac{n}{V}\left(1\mp \frac{ΔV}{V}\right)=\frac{n}{V}\mp \frac{nΔV}{V^{2}} \rightarrow ΔC=\frac{nΔV}{V^{2}}$$

*iii*) errore complessivo:

$$ΔC=\left|\frac{Δn}{V}\right|+\left|\frac{nΔV}{V^{2}}\right|$$

Si tiene in considerazione il modulo perché si vuole determinare l’errore massimo che si può commettere. L’errore totale viene quindi espresso come **somma dei valori assoluti degli errori dati dalle diverse grandezze** che rientrano nel calcolo. Se si fa la derivata della concentrazione rispetto alle moli e rispetto al volume, ci si rende conto che i due errori calcolati non sono altro che la derivata stessa moltiplicata per l’errore sulla grandezza:

$\frac{dC}{dn}=\frac{1}{V}$ errore dato dall’errore sulle moli: $ ΔC=\frac{Δn}{V}$

$\frac{dC}{dV}=-\frac{n}{V^{2}}$ errore dato dall’errore sul volume: $ ΔC=\frac{nΔV}{V^{2}}$

Quindi in generale si può dire che:

$$Δy=\sum\_{i}^{}\frac{∂y}{∂x\_{i}}Δx\_{i}$$

Questo modo di calcolare l’errore è diverso dalle formule che potreste avere già visto. Infatti, nel modo appena spiegato si determina un **errore assoluto**, più grande, “totale”, che comprende errori sia casuali che sistematici. L’alternativa è utilizzare delle formule che permettono di ricavare un errore che può essere definito come “più probabile”. Per rendere più chiara questa distinzione, si veda il seguente esempio:

$$y=\frac{a∙b}{c}$$

$$ Δy=\left|\frac{b}{c}∆a\right|+\left|\frac{a}{c}∆b\right|+\left|-\frac{a∙b}{c^{2}}∆c\right|$$



$∆y=\left|\frac{0.0050}{1.97}0.02\right|+\left|\frac{4.10}{1.97}0.0001\right|+\left|-\frac{4.10∙0.0050}{1.97^{2}}0.04\right|$→ **Δy = 0.0004**

Quindi, si conclude che:

$$\left|a\right|+\left|b\right|>\sqrt{a^{2}+b^{2}}$$

→ **Esercizio 3 (esercizio su propagazione errori)**

**Esempio bomba calorimetrica:**

$∆U = ∆U\_{Naftalene} + ∆U\_{filo}$ (ΔU filo trascurabile)

$∆U\_{Naftalene}= -m C\left(S\right) ∆T$ da cui si ricava C(S):

$$C\left(S\right) = -\frac{∆U\_{Naftalene}}{m ∆T}$$

$∆U\_{Naftalene}$ è considerato valore esatto, m e ΔT invece sono grandezze affette da errore. In particolare:

$∆∆T = \left(\frac{∂∆T}{∂T\_{i}}\right)\* ∆T\_{i} +\left(\frac{∂∆T}{∂T\_{f}}\right)\* ∆T\_{f} = \left|-1\*∆T\_{i}\right| + \left|∆T\_{f}\right| = 0.01 K + 0.01 K = 0.02 K$ (errore sul ΔT)

mentre l’errore sulla massa,Δm, è l’errore sulla bilancia. Quindi:

$$∆C\left(S\right) = \left(\frac{∂ C\left(S\right)}{∂ m}\right)\* ∆m + \left(\frac{∂ C\left(S\right)}{∂ T}\right)\*∆T =$$

$$= \left|\frac{-∆U\_{(teorico)}}{∆T} \* ∆m\right| + \left|\frac{-∆U\_{(teorico)}\* m}{∆T^{2}} \* ∆∆T\right|$$

**REGRESSIONE LINEARE:** in quasi tutte le esperienze del laboratorio, esistono delle **dipendenze lineari** da estrapolare. Esempio **CMC conduttimetrica**:



Ciò vuol dire definire la retta più rappresentativa per una serie di misure disposte in un grafico (cioè quella che minimizza la somma degli scarti di ciascun punto). Come si è fatto prima, anche in questo caso si definisce la funzione scarti come:

$$scarti=\sum\_{i=1}^{N}\left[y\_{i}-\left(mx\_{i}+q\right)\right]^{2}$$

che deve essere minimizzata rispetto a m e q. Per farlo devo imporre uguale a zero la derivata della funzione scarti rispetto a m e poi rispetto a q.



Dalle formule precedenti si ricavano, quindi, i valori di m e q. La **regressione lineare** permette di valutare l’errore complessivo per un set di punti, comprendendo tutti gli errori che concorrono (non si deve fare la propagazione dell’errore perché minimizza gli scarti quadratici medi). Durante le esperienze, mediante la funzione **REGR.LIN (di Excel, “scatola 3x2”)** si determineranno i valori di m e q, nonché il coefficiente di correlazione R (o di Pearson) fra le misure effettuate e quelle estrapolate tramite retta, cioè esprime quanto y dipenda linearmente da x. Solitamente si usa il **quadrato di R**, che risulta compreso fra 0 e 1 (se la retta è parallela all’asse x, R2 è <<1 perché y varia poco al variare di x, **esperienza tautomeria cheto-enolica**).

|  |  |
| --- | --- |
| 43500 | 0 |
| 100 | #N/D |
| 0.99998 | 0.006 |

 ↑

 err y

→ **Esercizio 4a (esercizio su rette di regressione, calcolo CMC)**

Per trovare l’ascissa dell’intersezione tra le due rette si eguagliano le equazioni delle due rette m1x+q1 = m2x+q2 e si trova x; mentre per calcolare Δx si applica la formula di propagazione degli errori in cui ci sono 4 variabili affette da errore: m1, m2, q1, q2.

$$x=\frac{q\_{2}-q\_{1}}{m\_{1}-m\_{2}}$$

$$∆x=\left|\frac{∆q\_{2}}{m\_{1}-m\_{2}}\right|+\left|-\frac{∆q\_{1}}{m\_{1}-m\_{2}}\right|+\left|\frac{(q\_{2}-q\_{1})∆m\_{1}}{(m\_{1}-m\_{2})^{2}}\right|+\left|-\frac{(q\_{2}-q\_{1})∆m\_{2}}{(m\_{1}-m\_{2})^{2}}\right|$$

→ **Esercizio 4b (esercizio su retta di Kraus a Bray)**

Trovare Ka dalla Kraus e Bray (Ka = q2/m) e ΔKa con la propagazione degli errori dove m e q sono le due variabili affette da errore:

$$ΔK\_{a}=\left|-\frac{q^{2}}{m^{2}}\right|∙Δm+\left|\frac{2q}{m}\right|∙Δq$$

**LEZIONE EXCEL:**

* come costruire un grafico (a dispersione, titolo grafico, titolo assi, aggiungere altre serie di punti)
* come aggiungere linea di tendenza (funzione regr.lin(VERO;VERO o FALSO;VERO) dipende se vincolata all’origine, impostare equazione su grafico e valore di R2, futura = retta che va oltre l’ultimo punto, verifica = retta che va prima del primo punto → CONTROL o COMMAND + SHIFT + INVIO)
* funzioni MEDIA(), DEV.ST()
* come mantenere costante una casella (F4, si deve avere il simbolo del dollaro es.: $A$1)
* come mettere le barre di errore sul grafico (solo lungo y, lungo x sono trascurabili)