

Corrigé Type : Examen d'Analyse des Données et Chimiométrie

Question 1 : Questions de cours (9,5 pts)

1. Différence entre l'analyse des données et la chimiométrie (2,5pts)

L'analyse des données est un ensemble général de méthodes statistiques pour traiter, visualiser et interpréter des données (moyenne, écart-type, régression, tests, etc.).

La chimiométrie est l'application spécifique de ces méthodes (et de méthodes avancées : ACP, PLS, chimiométrie multivariée, etc.) aux données chimiques pour résoudre des problèmes d'analyse, de calibration, de modélisation et d'optimisation en chimie analytique.

2. Paramètres qui représentent la précision (1,5 pts)

Les principaux paramètres sont :

- Écart-type (s) ou écart-type expérimental
- Variance (s^2)
- Coefficient de variation ($CV = s / \text{moyenne} \times 100 \%$)

Plus ces valeurs sont faibles → meilleure précision (répétabilité).

3. Différence entre la précision et l'exactitude (2pts)

- Précision : mesure de la répétabilité / dispersion des résultats autour de la moyenne (faible dispersion = haute précision).

- Exactitude : mesure de la proximité entre la valeur moyenne obtenue et la valeur vraie (ou valeur de référence) → liée au biais (faible biais = haute exactitude).

On peut avoir une bonne précision mais mauvaise exactitude (erresystématique), et vice versa.

4. Différents types d'erreurs dans les données expérimentales (1,5 pts)

- Erreur aléatoire(ou erreur accidentelle) : due au hasard, non prévisible, suit une distribution normale.

- Erreur systématique (ou erreur déterministe) : toujours dans le même sens, liée à un défaut de l'appareil, de la méthode, de l'étalonnage, etc.

- Erreur brute (ou grossière) : erreur humaine flagrante (mauvaise lecture, dilution erronée, etc.).

5. Caractéristiques permettant de distinguer chaque type d'erreur (2pts)

| Type d'erreur | Signe | Répétabilité des mesures | Effet sur la moyenne | Comment la détecter / corriger ? |

| Aléatoire | + et – (aléatoire) | Faible dispersion | N'annule pas l'erreur (moyenne $\approx$ valeur vraie) | Répéter les mesures, calculer s, CV |

| Systématique | Toujours + ou – | Bonne précision | Décalage systématique (biais) | Comparer à une valeur de référence, blanc, étalons |

| Brute | Très grande déviation | Une mesure aberrante | Fortement biaisée si non éliminée | Test de Grubbs/Dixon, rejet statistique des valeurs aberrantes |

Question 2 : Analyse de la concentration en fer dans un échantillon de sol (5,5 pts)

Un échantillon de sol est analysé pour déterminer la concentration en fer (en mg/kg). Les résultats des mesures sont les suivants : 120, 122, 121, 123, 124, 122, 121, 123, 122, 124.

1. Calculer la moyenne et l'écart type de ces mesures.
2. Déterminer la variance et le coefficient de variation.
3. Calculer les limites de confiance à 95% pour la concentration en glucose.
4. Interpréter les résultats en termes de précision et de fiabilité des mesures.
5. La valeur attendue est de 123.34 mg/kg.

- Y a-t-il une erreur systématique ?

La solution :

Données : 120, 122, 121, 123, 124, 122, 121, 123, 122, 124 mg/kg
(n = 10)

1. Moyenne et écart-type

Moyenne ($\bar{x}$) = 122.2 mg/kg (1pt)

Écart-type (s) = 1.32 mg/kg (1pt)

2. Variance et coefficient de variation

Variance (s^2) = 1.73 (mg/kg)² (0,75 pt)

Coefficient de variation CV = $(1.32 / 122.2) \times 100 \approx 1.08 \%$ (0,75 pt)

3. Limites de confiance à 95 % pour la concentration en fer (et non glucose – probable faute de frappe) (1pt)

t (95 %, ddl = 9) ≈ 2.262 (valeur standard, tableau donne ~ 2.26)

Intervalle de confiance : $\bar{x} \pm t \times (s / \sqrt{n})$

$\rightarrow 122.2 \pm 2.262 \times (1.32 / \sqrt{10}) \approx 122.2 \pm 0.94$

IC 95 % : 121.26 – 123.14 mg/kg

4. Interprétation en termes de précision et fiabilité (0,5 pt)

- Précision : très bonne (CV $\approx 1.08 \%$ est excellent pour une analyse de sol). Les mesures sont très répétables.

- Fiabilité : l'intervalle de confiance est étroit ($\sim \pm 0.94$ mg/kg), donc la moyenne est fiable. On peut affirmer avec 95 % de confiance que la vraie concentration est entre 121.3 et 123.1 mg/kg.

5. Valeur attendue = 123.34 mg/kg (0,5 pt)

- Biais = $122.2 - 123.34 = -1.14$ mg/kg

$\rightarrow$ Oui, il y a une erreur systématique (négative, sous-estimation systématique d'environ 1.14 mg/kg).

La différence est statistiquement significative car l'IC 95 % (121.3–123.1) n'inclut pas 123.34.

Question 3 : Calcul de la droite de régression et de la concentration d'un analyte inconnu (5 pts)

La droite d'étalonnage obtenue est la suivante :

Concentration (mg/l)	Absorbance
2.0	0.120
4.0	0.240
6.0	0.360
8.0	0.480

1. Établir la droite de régression et calculer les erreurs associées à la pente et à l'ordonnée à l'origine.
2. Calculer la concentration d'un échantillon inconnu dont l'absorbance mesurée est de 0.300
3. t pour $n - 2 = 2$ degrés de liberté, calculer les limites de confiances de la pente et l'ordonnée à l'origine à 95%.
4. Calculer la limite de détection (LD) et la limite de quantification (LQ) ; sachant que : signal du blanc (y_B) : 0.010 et écart type du blanc (S_B) : 0.002

Données pour les questions 2 et 3 :

Tableau 1 Tableau qui représente les valeurs de z en fonction de degré de confiance

Degré de confiance %	Z
90	1,64
95	1,96
96	2
99	2,58

Tableau 2 Tableau qui représente les valeurs de t pour divers degrés de probabilité :

Degré de liberté	Intervalles de confiance	
	90 %	95 %
1	6,31	12,7
2	2,92	4,30
3	2,35	3.18
4	2,13	2,78
5	2,02	2.57
6	1,94	2,45
7	1,90	2,36
8	1,86	2,31
9	1,83	2,26

10	1,81	2,23
----	------	------

La solution :

Données d'étalonnage :

Concentration (mg/L)	Absorbance
2.0	0.120
4.0	0.240
6.0	0.360
8.0	0.480

1. Droite de régression et erreurs associées (1pt)

Équation : $A = 0.060 \times C - 0.000$ (ordonnée pratiquement nulle)

Pente (b) = $0.060 \text{ (mg/L)}^{-1}$

Ordonnée à l'origine (a) ≈ 0.000

Erreurs standard (approximation avec n=4) (1pt)

- Erreur sur la pente ≈ 0.000 (car points parfaitement alignés)

- Erreur sur l'ordonnée ≈ 0.000

2. Concentration pour absorbance = 0.300 (1pt)

$$C = (0.300 - 0) / 0.060 = 5.00 \text{ mg/L}$$

3. Limites de confiance à 95 % pour la pente et l'ordonnée (ddl = n-2 = 2) (1pt)

$$t(95\%, \text{ddl}=2) = 4.303 \text{ (tableau } \approx 4.30)$$

Comme les erreurs standard sont quasi nulles (ajustement parfait), les intervalles de confiance sont extrêmement étroits (presque le point estimé).

Pente : $0.060 \pm (0)$

Ordonnée : $0.000 \pm (0)$

4. Limite de détection (LD) et limite de quantification (LQ) (1pt)

$$y_B \text{ (signal du blanc)} = 0.010$$

$$S_B \text{ (écart-type du blanc)} = 0.002$$

Formules courantes (IUPAC / normes) :

$$- LD = y_B + 3 \times S_B = 0.010 + 3 \times 0.002 = \mathbf{0.016} \text{ (absorbance)}$$

$$\rightarrow \text{Concentration LD} = 0.016 / 0.060 \approx 0.27 \text{ mg/L}$$

$$- LQ = y_B + 10 \times S_B = 0.010 + 10 \times 0.002 = \mathbf{0.030} \text{ (absorbance)}$$

$$\rightarrow \text{Concentration LQ} = 0.030 / 0.060 = 0.50 \text{ mg/L}$$