UNIVERSITE DE BATNA_2

FACULTE DE MEDECINE

SERIE DE TD N° 2

DE BIOSTATISTIQUE

**1ère ANNEE, 22/23**

# STATISTIQUE DOUBLE

## Exercice 01 :

Pour étudier les mécanismes hormonaux de la puberté on a mesuré les concentrations de deux hormones : l'œstradiol et l'œstrone pour un groupe de 8 adolescentes. Les résultats sont :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **$x_i$ = concentration œstradiol pg/ml** | 7.5 | 16.5 | 22 | 30 | 39 | 54 | 69 | 77 |
| **$y_i$ = concentration œstrone pg/ml** | 9 | 18.5 | 21.5 | 27 | 32.5 | 48.5 | 57 | 58 |

On note par H le point moyen des quatre premiers points du nuage et par K le point moyen des quatre autres points.

1) Calculer les coordonnées des points H et K et déterminer la droite d'ajustement Y.
2) Utiliser la droite des moindres carrés ordinaires pour déterminer Y(X).
3) Calculer la covariance et le coefficient de corrélation linéaire.

## Solution de l'exo n° 1 :

1) Coordonnées du point H ($\bar{x}_H = 19$ ; $\bar{y}_H = 19$) et du point K ($\bar{x}_K = 59.75$ ; $\bar{y}_K = 49$)

Le point moyen des deux points H et K est ($\bar{x} = 39.375$ ; $\bar{y} = 34$).

La droite d'ajustement demandée (appelée aussi droite de Mayer) est celle qui passe par les deux points H et K et est de la forme y = A + B ∗ x ; le système suivant va être satisfait :

$$\begin{cases} A + B * x_h = y_h \\ A + B * x_k = y_k \end{cases} \Rightarrow \begin{cases} A + B * 19 = 19 \\ A + B * 59.75 = 49 \end{cases} \Rightarrow \{A = \frac{8.17}{1.63} = 5.012269939 \text{ ; } B = \frac{6}{8.15} = 0.736196319\}$$

La droite d'ajustement de Mayer aura donc la forme : $y = 5.0123 + 0.7362 * x$

2) Détermination de la droite des Moindres Carrés Ordinaires (MCO).

La calculatrice nous donne les paramètres suivants : $\bar{x} = 39.375$ ; $\bar{y} = 34$ ; $\sigma_x^2 = 554.546875$

$\sigma_y^2 = 298.5$ ; A = 5.321074357 ; B = 0.728353667 ⇒ $y = 5.3211 + 0.7284 * x$

3) Calcul de la covariance et du coefficient de corrélation.

Cov (x ; y) = $\sigma_{xy} = \frac{1}{N} \sum xy - \bar{x}\,\bar{y} = \frac{13941.25}{8} - 39.375 * 34\ \bar{y} = 403.90625$ ; r = 0.99274858

## Exercice 02 :

Dans le but de doser le cuivre dans une spécialité pharmaceutique, on évalue les critères de qualité d'une méthode d'analyse du cuivre par spectrophotométrie d'absorption atomique.

***QUESTION N°1 :*** Lors de l'étude de répétabilité de la méthode, on mesure 12 fois l'absorbance d'une même solution :

**0,524 0,520 0,516 0,532 0,533 0,528 0,514 0,527 0,536 0,512 0,517 0,535**

Calculer la moyenne, l'écart-type et le coefficient de variation de l'absorbance.

***QUESTION N°2 :*** Pour vérifier la linéarité de la méthode, on prépare 6 solutions étalons dont les concentrations sont régulièrement espacées entre 0 et 1 mg/mL :

| **Concentration** | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 1 |
|---|---|---|---|---|---|---|
| **Absorbance** | 0.036 | 0.254 | 0.422 | 0.627 | 0.785 | 0.980 |

Déterminer l'équation de la droite de régression qui décrit la courbe d'étalonnage.

On admet que la fonction d'étalonnage peut être considérée comme linéaire si le coefficient de corrélation est supérieur à 0,998. La méthode est–elle linéaire ?

## Solution de l'exo n° 2 :

1) Calcul des paramètres : la moyenne ; l'écart-type et le coefficient de variation.

$\overline{x} = \frac{\sum x_i}{N} = \frac{6.294}{12} = 0.5245$ ; $\sigma = 0.008190441584$ ; $CV = \frac{\sigma}{\overline{x}} = 0.015615713$

2) Détermination de la droite de régression linéaire et le coefficient de corrélation.

La calculatrice donne A = 0.051761904 ; B = 0.931142857 et r = 0.999236013

La droite Y(X) étant la suivante $y = 0.0518 + 0.9311 * x$ et elle est linéaire car r > 0.998.

## ***Exercice 03 :***

On s'intéresse à la variation du VEMS en fonction de l'âge. Pour cela on mesure le VEMS de 10 sujets adultes (VEMS en litres ; le volume expiratoire maximum par seconde).

Les résultats sont indiqués dans le tableau suivant :

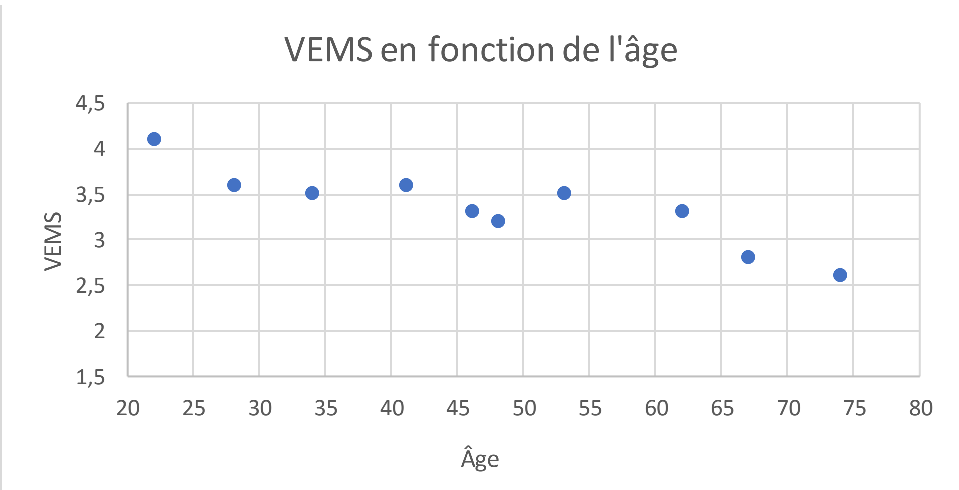
| **Sujet** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **VEMS** | 3.3 | 3.3 | 2.6 | 3.5 | 3.6 | 2.8 | 3.5 | 3.6 | 3.2 | 4.1 |
| **Âge** | 62 | 46 | 74 | 53 | 28 | 67 | 34 | 41 | 48 | 22 |

On envisage une relation linéaire entre le VEMS et l'âge.

a) Tracer le nuage de points
b) Trouver l'équation de la 1ère droite de régression.
c) Calculer le coefficient de corrélation entre les 2 variables.
d) Estimer le VEMS pour l'âge de 18 ans.

## Solution de l'exo n° 3 :

a) Le nuage de points est le suivant




b) L'équation de la droite de régression Y(X) où Y = VEMS et X = Âge est :

$y = 4.420383647 - 0.022534392 * x \Rightarrow y = 4.4204 - 0.0225 * x$

c) Le coefficient de corrélation r = – 0.897989181

d) L'estimation du VEMS pour x = 18 ans est : $\hat{y}(18) = 4.014764581$

## ***Exercice 04 :***

Le tableau suivant concerne les âges auxquels 100 couples se sont mariés :

| Classes | **Femmes Y** | **[17 ; 22[** | **[22 ; 27[** | **[27 ; 32[** | **[32 ; 37]** | **Σ** |
|---|---|---|---|---|---|---|
| **Maris X** | Centres | 19.5 | 24.5 | 29.5 | 34.5 | |
| **[20 ; 25[** | 22.5 | 0.14 | 0.09 | 0.01 | 0.00 | 0.24 24 |
| **[25 ; 30[** | 27.5 | 0.18 | 0.07 | 0.02 | 0.01 | 0.28 28 |
| **[30 ; 35[** | 32.5 | 0.04 | 0.13 | 0.03 | 0.01 | 0.21 21 |
| **[35 ; 40[** | 37.5 | 0.01 | 0.09 | 0.10 | 0.02 | 0.22 22 |
| **[40 ; 45]** | 42.5 | 0.00 | 0.01 | 0.02 | 0.02 | 0.05 5 |
| **Σ** | | 0.37 37 | 0.39 39 | 0.18 18 | 0.06 6 | 1 100 |

1) Compléter le tableau.
2) Calculer le tableau de contingence des fréquences.
3) Trouver les distributions marginales de X et de Y. Puis calculer les moyennes et les variances marginales.
4) Calculer la covariance entre X et Y ainsi que le coefficient de corrélation linéaire.

## Solution de l'exo n° 4 :

1) Voir tableau des données
2) Voir tableau des fréquences ci-dessus.
3) La distribution marginale, la moyenne et la variance de x (âge des maris) est :

| $c_i$ | 22.5 | 27.5 | 32.5 | 37.5 | 42.5 | **Σ** |
|---|---|---|---|---|---|---|
| $n_i$ | 24 | 28 | 21 | 22 | 5 | 100 |

$$\bar{x} = \frac{\sum x_i}{N} = \frac{3030}{100} = 30.3 \qquad \sigma_x^2 = \frac{\sum n_i x_i^2}{N} - \bar{x}^2 = \frac{95475}{100} - 918.09 = 36.66$$

La distribution marginale, la moyenne et la variance de y (âge des femmes) est :

| $c_i$ | 19.5 | 24.5 | 29.5 | 34.5 | **Σ** |
|---|---|---|---|---|---|
| $n_i$ | 37 | 39 | 18 | 6 | 100 |

$$\bar{y} = \frac{\sum y_i}{N} = \frac{2415}{100} = 24.15 \qquad \sigma_y^2 = \frac{\sum n_i y_i^2}{N} - \bar{y}^2 = \frac{60285}{100} - 583.2225 = 19.6275$$

4) Calcul de la covariance et du coefficient de corrélation.

Cov (x ; y) $= \frac{\sum\sum n_{ij} x_i y_i}{N} - \bar{x}\,\bar{y} = \frac{74722.5}{100} - 30.3 * 24.15 = 15.48$

Le coefficient de corrélation r = 0.577088251

La droite de régression Y(X) est la suivante y = A + B ∗ x = 11.35556465 + 0.422258592∗ x

## ***Exercice 05 :***

Pour juger de l'efficacité d'une drogue D dans la prévention d'une maladie. Au cours de l'étude de l'activité de la drogue D, on obtient les résultats suivants :

| **x** | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| **y** | 0.29 | 0.52 | 0.61 | 0.79 |

(x c'est le logarithme de la dose avec unité arbitraire, y : fraction d'un effet maximum)

Déterminer les paramètres p et $y_0$ de la relation effet–dose $\mathbf{y = px + y_0}$

Calculer le coefficient de corrélation linéaire.

## Solution de l'exo n° 5 :

En choisissant le mode « régression linéaire » la calculatrice donne la droite de régression Y(X) qui est la suivante : $y = y_0 + p * x = A + B * x = 0.314 + 0.159 * x$ avec $r = 0.987311051$

En choisissant le mode « régression logarithmique Néperien » avec les données suivantes :

| x | $e^0$ | $e^1$ | $e^2$ | $e^3$ |
|---|---|---|---|---|
| y | 0.29 | 0.52 | 0.61 | 0.79 |

La calculatrice donne la droite de régression Y(X) qui est : $y = y_0 + p * x = A + B * x$

Avec A = 0.314 et B = 0.159 et r = 0.987311051

**Avec les deux méthodes, on aboutit au même résultat.**

## ***Exercice 06 :*** *Cinétique du premier ordre.*

Un corps chimique se décompose selon une cinétique du premier ordre caractérisée par l'équation : $\mathbf{Q = Q_0\, e^{-kt}}$

Où : Q désigne la quantité de corps restant à l'instant t ; $Q_0$ la quantité initiale ; k la constante de vitesse de la décomposition. On dispose des données expérimentales suivantes :

| **t en minutes** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Q en nanomoles** | 416 | 319 | 244 | 188 | 144 | 113 | 85 | 66 | 50 | 41 |

Evaluer la constante de vitesse k.

## Solution de l'exo n° 6 :

*Une première méthode* : On choisit le mode « REGression EXPonentielle » avec la calculatrice KENKO ou bien le mode « **e^x** » avec la calculatrice CASIO ($5^{\text{ième}}$ choix) on obtient ce qui suit :

$A = 534.524494 = Q_0$ $B = -0.260530669 = -k$ $r = -0.999768735 \Rightarrow \mathbf{Q = Q_0\, e^{-k\,t}}$

$\Rightarrow \mathbf{Q = 534.52449\, e^{-0.26053\,t}} \Rightarrow \mathbf{k = 0.26053}$

*Une deuxième méthode* : On choisit le mode « **A B^x** » avec la calculatrice CASIO ($6^{\text{ième}}$ choix) on obtient ce qui suit :

$A = 534.524494 = Q_0$ $B = \mathbf{e^{-k}} = 0.7706425209 \Rightarrow \text{Ln}\, B = -k = -0.260530669$ et $k = 0.26053$

$r = -0.999768735 \Rightarrow \mathbf{Q = Q_0\, e^{-kt}} \Rightarrow \mathbf{Q = 534.52449\, e^{-0.26053\,t}} \Rightarrow \mathbf{k = 0.26053}$

### Une autre méthode avec la linéarisation avec le mode « REGression LINéaire » :

Les données avec le ***logarithme décimal*** sont les suivantes :

| **t en minutes** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Q en nanomoles** | log416 | log319 | log244 | log188 | log144 | log113 | log85 | log66 | log50 | log41 |

En utilisant le logarithme décimal : $\log Q = \log Q_0 + \log e^{-k\,t} = \log Q_0 + (-k \log e)\, t$

$\Rightarrow A = \log Q_0 \,;\, B = -k \log e \Rightarrow k = -\frac{B}{\log e}$

La calculatrice donne : $A = \log Q_0 = 2.727967611 \Rightarrow 10^A = 534.524494 \,;\, B = -0.113147032 \,;$

$r = -0.999768735 \Rightarrow k = -\frac{-0.113147032}{0.434294481} = 0.260530669 \Rightarrow \mathbf{k = 0.26053}$

Les données avec le ***logarithme Népérien*** sont les suivantes :

| **t en minutes** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Q en nanomoles** | Ln416 | Ln319 | Ln244 | Ln188 | Ln144 | Ln113 | Ln85 | Ln66 | Ln50 | Ln41 |

En utilisant le logarithme Népérien : $y = \text{Ln } Q = \text{Ln } Q_0 + \text{Ln } e^{-kt} = \text{Ln } Q_0 + (-k\,t) = A + B\,x$

$\Rightarrow A = \text{Ln } Q_0 = 6.281377555 \Rightarrow \mathbf{Q_0} = e^A =$ **534.524494** et $B = -k = -0.260530669 \Rightarrow$

$k = -B = 0.260530669 \Rightarrow$ **k = 0.26053**

***<u>Conclusion</u>*** **: La Constante de Vitesse k prend toujours la même valeur de 0.26053**