Université de Batna 2
Faculté de Médecine
1ère année

TD 2 de Biostatistique
Statistique double

Exercice n° 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 :

Ajustement de Mayer :

La droite de Mayer passe par les deux points H (19 ; 19) et K (59.75 ; 49)

$$\begin{cases} 19 = 19a + b \\ 49 = 59.75a + b \end{cases}$$

La résolution de système nous donne les valeurs de couples (a ; b).

Donc l'équation de la droite de l'ajustement Mayer $y = 0.736x + 5.012$

2) Ajustement de la méthode des moindres carrés.

$$D_y(x): y = 0.728\,x + 5.3211$$

$$r = Cor(X,Y) = 0.992$$

$$\sigma_{xy} = Cov(X,Y) = \frac{\sum x_i y_i}{n} - \bar{X}\bar{Y} = \frac{13941.25}{8} - 39.375 \times 34 = 403.9$$

Exercice n° 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 ?

Question 1 : $\overline{Y} = 0.52 \quad \sigma_y = 8.19$

Question 2 :

Ajustement de la méthode des moindres carrés.

$$D_y(x): y = 0.93114\, x + 0.05176$$

$$r = Cor(X, Y) = 0.999236$$
$$r > 0.998$$

Oui linéaire

<u>Exercice n° 03 :</u>

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
AGE(Y)	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.

2) La 1ère droite de régression

Ajustement de la méthode des moindres carrés.

$$D_y(x): \ y = -35.78\, x + 167.38$$
$$r = Cor(X, Y) = -0.89$$

Pour prévoir le VEMS pour l'âge de 18 ans, on calcule tout d'abord la 2ème droite de régression :

$$D_x(y):\ x = -0.02253\, y + 4.42$$
$$r = Cor(X,Y) = Cor(Y,X) - 0.89$$
$$\hat{x} = -0.02253(18) + 4.42 = 4.015$$

(LA 2ième droite est plus logique car le VEMS se varie en fonction de l'âge)

Exercice n° 4

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					
[20 ; 25[		14	9	1	0	
[25 ; 30[		18	7	2	1	
[30 ; 35[		4	13	3	1	
[35 ; 40[		1	9	10	2	
[40 ; 45[		0	1	2	2	
Σ						

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

Solution :

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	14	9	1	0	24
[25 ; 30[	27.5	18	7	2	1	28
[30 ; 35[	32.5	4	13	3	1	21
[35 ; 40[	37.5	1	9	10	2	22
[40 ; 45[	42.5	0	1	2	2	5
Σ		37	39	18	6	100

Les fréquences

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	0.24
[25 ; 30[	27.5	0.18	0.07	0.02	0.01	0.28
[30 ; 35[	32.5	0.04	0.13	0.03	0.01	0.21
[35 ; 40[	37.5	0.01	0.09	0.10	0.02	0.22
[40 ; 45[	42.5	0	0.01	0.02	0.02	0.05
Σ		0.37	0.39	0.18	0.06	1

1)

La distribution marginale de x :

x_i	22.5	27.5	32.5	37.5	42.5
n_i	24	28	21	22	5

La distribution marginale de y :

y_j	19.5	24.5	29.5	34.5
n_j	37	39	18	6

Les moyennes et variances marginales sont :

$\bar{x} = 30.3 \;\; \bar{y} = 24.15 \;\;\; \sigma_x^2 = 36.66 \;\;\; \sigma_Y^2 = 19.6275$

$$r = Cor(X,Y) = 0.577$$

$$\sigma_{xy} = Cov(X,Y) = \frac{\sum_{i=1}^{5}\sum_{j=1}^{4} n_{ij}\, x_i y_j}{n} - \bar{X}\bar{Y} = 15.48$$

$$D_y(x)\colon y = 0.42x + 11.35$$

$$D_x(y)\colon x = 0.788y + 11.25$$

Exercice n° 5

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 = log dose	0	1	2	3
y	0.29	0.52	0.61	0.79

(Dose : unité arbitraire, y : fraction d'un effet maximum)

a) Déterminer les paramètres p et y0 de la relation effet-dose y= px + y0
b) Calculer le coefficient de corrélation linéaire.

Solution :

$$D_y(x)\colon \;\; y = 0.16x + 0.314$$

Relation linéaire entre (log dose)et (l'effet)

Exercice n° 6 :

Cinétique du premier ordre Un corps chimique se décompose selon une cinétique du premier ordre caractérisée par l'équation :

$$Q = Q_0 e^{-kt}$$

Où : Q désigne la quantité de corps restant à l'instant t ; Q0 la quantité initiale ; k la constante de vitesse de la décomposition.

On dispose des données expérimentales suivantes :

t (min)	1	2	3	4	5	6	7	8	9	10
Q (nanomoles)	416	319	244	188	144	113	85	66	50	41

Evaluer la constante de vitesse k.

Solution :

1ère méthode :

Vous choisissez la régression exponentielle :

$$\boldsymbol{y = ae^{bx}}$$

$$\mathrm{a} = 534.52$$

$$\mathrm{b} = -0.2605$$

$$\mathrm{r} = -0.99997$$

$$\mathrm{Q} = \mathrm{Q_0 e^{-kt}}$$

$\mathrm{Q} = 534.52\mathrm{e}^{0.261\mathrm{t}}$ k=0.261

2ème méthode :

Vous choisissez la régression linéaire

$$\mathrm{Q} = \mathrm{Q_0 e^{-kt}}$$

$$\mathbf{lin}\mathrm{Q} = \mathbf{lin}\mathrm{Q_0 e^{-kt}} = \boldsymbol{lin}\ \mathrm{Q_0} - k\boldsymbol{t}$$

On pose :

$$\begin{cases} y = \boldsymbol{lin}\ Q \\ x = t \\ a = -k \\ b = lin\ Q_0 \end{cases}$$

Vous cherchez la droite de régression de y en fonction de x :

$$y = ax + b = -0.2605x + 6.281377$$

$$a = -k = 0.261$$