

Université Med Khider Biskra
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie
Département des sciences Agronomiques
Module : Biométrie Dirigé par M^{me} MEBREK
TD N°4 (corrélation et la droite de régression) Master 1 (Hydropédologie)

Exercice 1:

Le calcul des caractéristiques d'un nombre de variables statistiques a donné les résultats suivants :

- $\text{Var}(X) = 23,2$; $\text{Var}(Y) = -11,95$; $\text{Var}(Z) = 2,23$; $\text{Var}(T) = 0$
- $\text{Cov}(X,Y) = -210$; $\sigma_x = 11,3$; $\sigma_y = 10,2$
- $E = 12$; $X = 18$; $M_e = 5$ (E est l'étendue et M_e est la médiane)

1. Indiquer dans chaque cas où est l'erreur et dites pourquoi ?
2. Si le coefficient de corrélation entre les deux variables X et Y est égal à zéro qu'est-ce que vous pouvez déduire sur la relation entre X et Y.
3. Lors du calcul de la moyenne arithmétique on remplace les classes par leur centre pourquoi ?

Exercice 2:

On a extrait au hasard un nombre d'épis de maïs et on a mesuré le poids et la longueur couverte de grains de chacun. Les longueurs ont été classées de centimètre en centimètre, pour certains épis on dispose de plusieurs poids pour une même longueur.

Les résultats sont :

Epis(cm)	7	8	9	10	11	12	13	14	15	16
Poids(g)	82	74/80/99	113/117	134	115/125/128/149	132/157	152/166	179	204	216

1. Calculer le coefficient de corrélation entre les deux paramètres. Expliquer ?
2. Etablir une régression linéaire entre le poids et la longueur.

Exercice 3 :

On veut déterminer s'il existe une relation linéaire entre les salaires Y et les années d'études X à travers les données suivantes :

X_i (années)	6	12	10	8	9
Y_i (en millier de DA)	10	20	17	12	11

1. Construire le nuage de points correspondant aux données.
2. Calculer le coefficient de corrélation des variables X et Y.
3. Déterminer la droite de régression de Y en X.

Exercice 4 :

Dans une ville de 100.000 habitants, une épidémie s'est déclenchée. On enregistre chaque jour le nombre de cas qui se sont déclarés. Au septième jour, le tableau suivant a été obtenu :

X_i	1	2	3	4	5	6	7
Y_i	4	13	38	106	330	965	2920

X_i représente le numéro du jour ; Y_i représente le nombre de cas enregistrés.

1. Ajuster la variable Y en fonction de la variable X à l'aide d'une fonction exponentielle de la forme $Y = B.A^X$
2. Combien de jours faut-il pour que tous les habitants de la ville soient atteints, si aucune mesure n'est prise pour stopper cette épidémie ?

Exercice 5 :

L'étude de deux variables X (le poids de 1000 grains en gramme) et Y (le rendement en qx/ha) pour une moyenne des plusieurs échantillons chez le blé dur a donné les résultats suivants :

X (g)	43	46	48	50	52	55	56	58	60	62
Y (qx/ha)	30	33	35	36	37	39	39	42	43	45

1. Calculer les variances de X et Y. **(2,5pts)**
2. Calculer la covariance. **(1,5pts)**
3. Calculer le coefficient de corrélation et interpréter (déduire le coefficient de détermination). **(2pts)**
Déterminer l'équation de la droite de régression ($Y = f(X)$). **(1pts)**

Exercice 6 :

On a pris deux variables X et Y, on a calculé pour un échantillon les quantités suivantes :

$$\bar{X} = 20 ; \bar{Y} = 12 ; \frac{1}{N} \sum x_i^2 = 436 ; \frac{1}{N} \sum y_i^2 = 153 \text{ et } \frac{1}{N} \sum x_i \cdot y_i = 2544$$

Calculer le coefficient de corrélation et interpréter.