

TD2 : Distribution, covariance et corrélation

Exercice 1 [Fréquences Cumulées Décroissantes]

Contrairement aux fréquences cumulées (croissantes), notées FCC, les fréquences cumulées décroissantes (FCD) commencent par la somme des poids $p_1 + \dots + p_n$ puis décroissent en $p_2 + \dots + p_n$, $p_3 + \dots + p_n$, et caetera ...

Une enquête sur les salaires mensuels des employés d'une entreprise a été effectuée et les résultats sont donnés par tranches salariales. Cependant, certaines données sont incomplètes :

Salaires mensuel	Fréquence	FCC	FCD
[1000; 1400[0,35		
[1400; 1800[0,60	
[1800; 2200[
[2200; 2600[0.10
	1.00		

Compléter le tableau en justifiant.

Exercice 2

Au sein d'une résidence, les logements se répartissent en fonction de leurs nombres de pièces de la façon suivante :

Pièce(s)	1	2	3	4	5	6
Logements	6	10	12	15	7	3

Nous considérerons ici que tous les logements sont pondérés de la même manière.

1. Calculer le nombre moyen de pièces des logements de cette résidence ainsi que l'écart-type de cette répartition.
2. Calculer les fréquences cumulées de cette répartition.
3. Calculer la médiane et les quartiles de cette répartition.
4. Tracer la fonction de répartition et retrouver graphiquement la médiane et les quartiles.
5. Tracer l'histogramme des effectifs et celui des fréquences cumulées. Faites apparaître sur ces deux diagrammes les valeurs de la moyenne et de la médiane.
6. Tracer le diagramme à moustaches relatif à ces observations.

Exercice 3

Voici les observations de deux variables couplées chez cinq individus pondérés par des poids p_k :

k	1	2	3	4	5
p_k	3/10	1/10	3/10	1/10	2/10
x_k	3	2	-4	-2	4
y_k	1	-2	3	2	-1

1. Réécrire ce tableau avec les données centrées $x_k - \bar{x}$ et $y_k - \bar{y}$.

2. Calculer $\text{Var}(x)$, $\text{Var}(y)$, $\text{Cov}(x, y)$ et $\rho(x, y)$.

Exercice 4

Soit un jeu de données associé à deux variables x et y observées sur n individus pondérés par des poids p_k normalisés.

1. Que vaut

$$\sum_{k=1}^n \sum_{l=1}^n p_k p_l (x_k - x_l)(y_k - y_l) ?$$

2. Montrer que

$$\text{Cov}(x, y) = \frac{\text{Var}(x + y) - \text{Var}(x - y)}{4} .$$

3. Pour quels couples de réels (a, b) a-t-on

$$\sum_{k=1}^n p_k (x_k - a)(y_k - b) = \text{Cov}(x, y) ?$$