Université Norbert ZONGO Année Académique 2023-2024
Niveau : $MPCI\ L_1S_1$

TD de statistique descriptive

Exercice 1

1) a) Quelle différence y a t-il entre **la statistique** et **les statistiques** ?
   b) La statistique est-elle importante pour un pay ? Justifier votre reponse en prenant le cas du burkina Faso.
2) A travers un exemple, donner la différence entre une population mère et un échantillon en statistique.
3) a) Qu'appelle-t-on caractère dans une population ?
   b) Donner, avec précision, les deux grands types de caractères en statistique.
   c) Donner dans chacun des deux grands types de caractères 5 exemples si la population considérée est un groupe d'étudiant de l'université Norbert ZONGO.

Exercice 2

1) Dites si les caractères suivants sont quantitatives ou qualitatives
   a) La taille des agents d'une entreprise
   b) Le nombre d'actions vendues chaque jour à la bourse
   c) Les températures enregistrés chaque demi-heure dans un centre météorologique
   d) Le nombre de livre sur une étagère de bibliothèque
   e) Les provinces du Burkina Faso
   f) Le nombre des membres d'une famille
   g) Le nombre d'inscriptions à l'université Norbert ZONGO pendant un certain nombre d'années
   h) Le Statut civil d'un individu
   i) Le nombre de pétales d'une fleur
   j) La religion pratiquée
   k) Les types habitations
   l) L'échelle d'appréciation d'un produit
   m) Le revenue des employés
2) Pour chacun des caractères de la question précédente, identifier le type : nominal, ordinal, discret, continu.

Exercice 3

1) La formule de la variance d'une série statistique est donnée par

$$V(X) = \frac{1}{N}\sum_{i=1}^{n} n_i \left(x_i - \overline{x}\right)^2$$

avec $\overline{x}$ et N l'effectif total de la population.
Démontrer que :

$$V(X) = \frac{1}{N}\sum_{i=1}^{n} n_i x_i^2 - \overline{x}^2$$

2) Soit n un entier naturel et $(x_1, ..., x_n)$ un $n-$ uplet de réels. On souhaite trouver un réel x minimisant la des écarts au carré. On définit donc sur $\mathbb{R}$ la fonction G par :

$$G(x) = \sum_{i=1}^{n} (x - x_i)^2$$

   a) En écrivant $G(x)$ sous la forme d'un trinôme du second degré, demontrer que la fonction G admet un minimum sur $\mathbb{R}$ et indiquer en quelle valeur de x il est atteint
   b) Que représente d'un point de vue statistique la valeur de x trouver à la question précédente.

<u>Exercice 4</u>

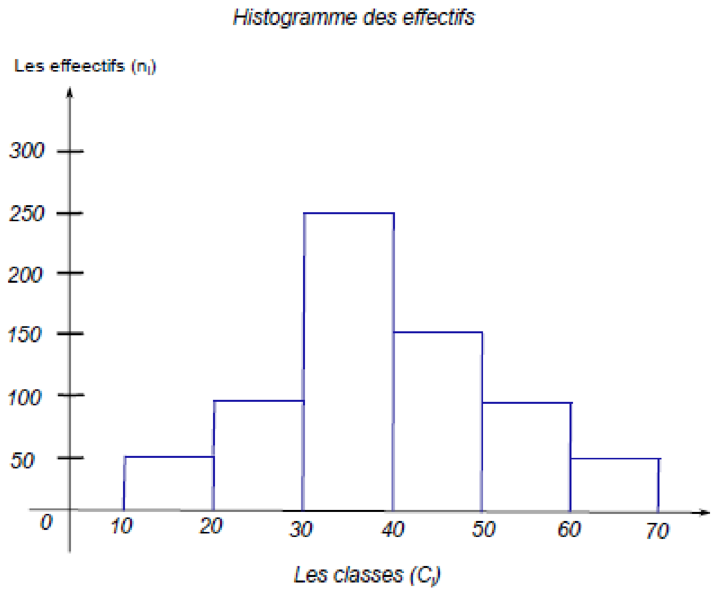
On observe le nombre d'arrivées (variable X) de clients à un bureau de poste pendant un intervalle de temps(10 mn) et on obtient les valeurs suivantes :

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6

1) Dresser le tableau statistique de la distribution de la variable X (effectifs et effectifs cumulés, fréquence et fréquence cumulées).
2) Calculer les valeurs de tendance centrale de la distribution : la moyenne (arithmétique, géométrique, harmonique), le mode et les trois quartiles Q_1, Q_2, Q_3.
3) Calculer les valeurs de la dispersion de la distribution : variance, écart type, l'intervalle interquartile, l'écart moyen et le coefficient de variation.
4) Déterminer le coefficient de dissymétrie de Pearson. Interpréter.
4) Tracer le diagramme en bâtons et la boite à moustaches de cette distribution.

Exercice 5

Dans une gare routière, on évalue le temps d'attente des voyageurs en minutes. Voici l'histogramme des fréquences absolues de cette variable.



1) Déterminer la variable statistique X et son type et sa population.
2) Déterminer le nombre de voyageurs.
3) Depuis le graphe, déterminer le tableau statistique.
5) Calculer la médiane M_e, le 1^{er} Q_1 et le 3^e Q_3 quartile puis les interpréter.
6 Déterminer le mode M_0.
6) Calculer les moyennes arithmétique, géométrique et harmonique
7) Calculer la variance et l'écart type.
8) Calculer le coefficient de dissymétrie de Pearson. Interpréter

Exercice 6

On a relevé la population d'une grande ville sur 50 ans tous les 5 ans. Les résultats son présentés dans le tableau suivant :

Annés x_i	0	5	10	15	20	25	30	35	40	45	50
Population en milliers y_i	19,4	19,4	27,6	40,3	50	59	69	87	132	166	216

1) Calculer le coefficient de corrélation entre x et y puis déduire que x et y ne sont pas linéairement liées.
2) On effectue le changement de variable $z = ln(y)$.
   a) Réaliser un nouveau tableau pressentant les valeurs prises par les variables x et z.

b) Représenter un nuage de points à partir des donnés des variables x et z.

c) Un ajustement linéaire semble être possible? Confirmer ce résultat en calculant le coefficient de corrélation entre x et z.

d) Déterminer une équation de la droite d'ajustement de z en x par la méthode des moindres carrées.

Réaliser un nouveau tableau pressentant les valeurs prises par les variables x et z.

3) a) Enduire la relation qui lie y et x.

b) Suivant ce modèle, estimer le nombre d'années sachant que le nombre d'habitants est 70000.

c) En admettant que le modèle mathématique reste valable en dehors du domaine d'étude , extrapoler le nombre d'habitants 5 ans après l'étude.

Exercice 7

Une expérience a été réalisée sur 250 personnes pour étudier la relation qui existe entre l'âge X et le temps de sommeil Y, le tableau suivant a été obtenu :

X/Y	[5 :7[	[7 :9[	[9 :11[	[11 :15[
[1 :3[	0	0	3	36
[3 :11[	0	5	11	27
[11 :19[	3	9	36	17
[19 :31[	0	26	24	5
[31 :59[	24	17	7	0

1) Les variables X et Y sont t-elles indépendante?
2) Calculer f_{42} et f_{23}.
3) Déterminer le pourcentage des individus dont leur temps de sommeil ne vaut pas 11h.
4) Déterminer l'age moyen des individus dont le temps sommeil se trouve dans l'intervalle $[9, 11[$.
4) Déterminer la tension moyenne des individus dont l'âge se trouve dans l'intervalle $[11, 19[$.
5) Calculer les moyennes marginales et les écart types marginaux de X et Y.
6) Déterminer la covariance et le coefficient de corrélation linéaire.
7) Déterminer la droite de régression de Y en fonction de X.
8) Estimer le temps de sommeil d'une personne de 66 ans.
9) Estimer l'âge d'une personne dont le temps de sommeille est de 8,5h..