

Fiche TD N° 01- Rappel sur Statistique Descriptive-

Exercice 01: On considère la série statistique suivante :

Valeur de X	$5+10^{1000}$	$9+10^{1000}$	$11+10^{1000}$	$17+10^{1000}$
Effectif	3	16	16	5

Calculer la moyenne arithmétique de cette série statistique.

Exercice 02:

Voici un tableau d'effectifs donnant la production mondiale d'or en 1937.

Continents	Europe	Asie	Afrique	Amérique	Océanie
Production en tonnes	176	87	431	350	56

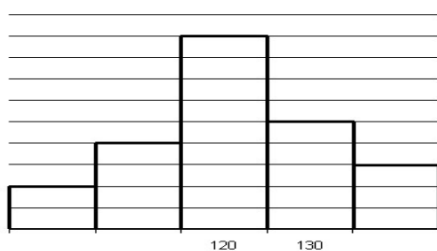
Préciser la population, sa taille, un individu de la population, la variable et son type.

Exercice 03: Des élèves ont participé à une dictée. Voici le nombre de fautes relevées pour chacun :

8	3	5	1	4	2	8	7	9	4
3	1	0	2	8	2	1	8	4	0
2	8	5	1	0	5	7	4	3	9
6	8	4	9	2	8	8	7	0	3
6	8	4	9	2	8	7	7	0	3

1. Expliquer la méthode de dépouillement de ces valeurs.
2. Déterminer la variable et son type. Quelle est la taille de l'échantillon ?
3. Déterminer le mode, Déterminer la médiane et les quartiles de cette série statistique ainsi que l'intervalle interquartile.
4. Déterminer la moyenne et l'écart-type.

Exercice 04: Le graphique ci-dessous représente l'histogramme des effectifs de la variable Q.I. (notée X) sur un échantillon de personnes. Les modalités de X sont regroupées en classes de même amplitude 10. La classe de centre 120 a pour effectif 63. Déterminer la fréquence de chaque classe. et Calculer la moyenne et l'écart-type



de X sur cet échantillon.

Exercice 05:

Une entreprise est constituée de deux usines, appelées A et B. Le tableau suivant récapitule les salaires en euros par catégorie de personnel et par usine :

Usine A	Salaires	Effectifs	Usine B	Salaires	Effectifs
Ouvriers	700	200	Ouvriers	900	60
Employés	1400	20	Employés	1600	40
Cadres	5300	10	Cadres	7300	20

Calculer la moyenne des salaires dans chacune des usines, dans l'entreprise. Vérifier que la moyenne des salaires dans l'entreprise est la moyenne des salaires moyens de chaque usine. ensuite Calculer la variance des salaires dans chacune des usines et dans l'entreprise. . Vérifier que la variance des salaires dans l'entreprise

est égale à la moyenne des variances des usines augmentée de la variance des moyennes calculées dans chaque usine.

Exercice 05:

Voici un tableau de données ponctuelles représentant une série statistique à deux variables (X, Y) . Donner le tableau à double entrée correspondant, ainsi que les distributions marginales de X et de Y . et Calculer la moyenne et l'écart-type de X et Y ?

Observation $n =$	1	2	3	4	5	6	7	8	9	10	11	12
X	1,5	0	0	2,5	2,5	2,5	0	1,5	1,5	2,5	2,5	1,5
Y	-1	1	-1	1	-1	0	1	-1	1	0	-1	1

Exercice 06: Soit X une v.a. admettant un moment d'ordre 2.

1. On pose $Y = 3X + 2$. Calculer $\text{Cov}(X, Y)$ et $\text{Var}(Y)$ en fonction de la variance de X .
2. On pose $Y = -3X + 2$. Calculer $\text{Cov}(X, Y)$ et $\text{Var}(Y)$ en fonction de $\text{Var}(X)$

Exercice 07:

Soit X une variable aléatoire dont la fonction de densité est donnée par :

$$f(x) = \begin{cases} x & \text{si } x \in [0; 1] \\ 2 - x & \text{si } x \in [1; 2] \\ 0 & \text{ailleurs} \end{cases}$$

1. Déterminer la fonction de répartition F de X . Tracer son graphe ainsi que celui de f .
2. Déterminer la médiane de X .
3. Calculer son espérance et sa variance.
4. Calculer $P(|X-1| < 1/2)$.

Exercice 08:

on prend cinq patients au hasard à qui il applique un traitement, et les cinq autres étant des témoins non traités à qui on administre un placebo. Les résultats concernant la durée de survie (exprimée en années) des malades étudiés sont présentés ci-après :

Malades traités (X)	6,5	4,2	17,8	7,9	13,2
Malades non traités (F)	6,7	0,4	2,9	1,2	5,6

- 1- Calculer la médiane (M) des données.
- 2- Compléter le tableau des effectifs ci-dessous :

	Durée de vie < M	Durée de vie > M	Σ
Malades traités			
Malades non traités			
Σ			

Exercice de TP:

Afin de déterminer s'il existe une relation entre l'âge des premiers mots d'un enfant et ses capacités intellectuelles, on a mesuré cet âge (en mois) et l'indice de Gesell (mesuré à l'âge scolaire) pour 21 enfants. Les données recueillies sont les suivantes pour les 21 enfants

: 15, 26, 10, 9, 15, 20, 18, 11, 8, 20, 7, 9, 10, 11; 11, 10, 12, 42, 17, 11, 10

pour l'âge, les valeurs correspondantes pour l'indice étant :

95, 71, 83, 91, 102, 87, 93, 100, 104, 94, 113, 96, 83, 84, 102, 100, 105, 57, 121, 86, 100.

1. Représenter le nuage de point et sa droite des moindres carrés. Qu'en pensez-vous ?
2. Il y a deux données éloignées (outliers) que vous indiquerez. Refaite le calcul de la droite des moindres carrés en éliminant l'une, puis l'autre puis les deux données éloignées. Qu'observez-vous ?
3. Analyser l'influence de ces données éloignées sur le choix de la droite. Sans la représentation graphique, comment pourriez-vous repérer une donnée éloignée ?