

TD-1 : Série univariée

Exercice 1. Réserves de pétroles

Voici un tableau donnant la liste des 20 pays ayant les plus grandes réserves de pétrole (en milliards de barils, pour le Venezuela).

Rang	Pays	Réserve
1	Venezuela	297.70
2	Saudi Arabia	268.35
3	Canada	174.4
4	Iran	157.30
5	Iraq	140.3
6	Kuwait	104.00
7	United Arab Emirates	97.80
8	Russia	80.00
9	Libya	48.00
10	Nigeria	37.20
11	United States	30.50
12	Kazakhstan	30.00
13	China	25.50
14	Qatar	25.30
15	Brazil	13.90
16	Bolivia	13.20
17	Angola	10.40
18	Mexico	10.40
19	Algeria	9.90
20	Azerbaijan	7.00

1. Quelle est la population étudiée ?
2. Préciser le type de la variable statistique ?
3. Calculez la moyenne \bar{x} , la variance V et l'écart-type s ?
4. Commenter la dispersion des données.
5. Calculer les trois quartiles. Comment se comparent la moyenne et la médiane ?
6. Quel coefficient permet de caractériser la forme des données ? Calculer ce coefficient et interpréter le résultat.
7. Faire les calculs nécessaires puis tracer la boîte à moustaches.
8. Bien réfléchir aux intervalles puis tracer l'histogramme. On indiquera sur l'axe de droite l'échelle correspondant à l'histogramme des densités de fréquences.
9. On considère maintenant aussi les 80 pays qui ont des réserves plus petites que celles de l'Azerbaïdjan.

- (a) Quelle est la population étudiée à présent ?
- (b) Donner un histogramme pour l'ensemble des 100 pays.
- (c) Estimer grossièrement la réserve moyenne par pays pour l'ensemble des 100 pays.

Exercice 2. On a noté les prix du pétrole brut du 20 nov au 8 déc 2014 (en \$) :

$$d = (75.6, 76.5, 75.7, 74.0, 73.7, 70.2, 65.9, 69.0, 67.0, 66.7, 65.9, 63.1, 62.9).$$

1. Quelle est la population étudiée ?
2. Quel est le type de la variable étudiée ?
3. Calculer la moyenne et l'écart-type.
4. Commenter la dispersion des données.
5. Commenter la forme des données.
6. Proposer deux méthodes pour la représentation graphique de ces données : une qui donne une information dans le temps - une autre qui donne une information sur la distribution des prix durant cette période.
7. Est-il possible de projeter ce prix sur le court terme ? sur le long terme ? Faites-le dans le cas d'une réponse positive. Quelle hypothèse fait-on ?

Exercice 3. On a relevé pour vingt individus, certaines caractéristiques biologiques et sociologiques. Le tableau ci-dessous résume les résultats obtenus :

Individu	Groupe sanguin	Âge	Nombre d'enfants
1	B	38	3
2	AB	42	4
3	A	16	0
4	A	51	2
5	O	28	1
6	A	34	1
7	O	21	1
8	A	54	4
9	AB	35	0
10	A	22	3
11	A	12	0
12	B	40	2
13	O	41	1
14	O	31	2
15	A	25	1
16	AB	64	6
17	A	60	1
18	O	55	0
19	B	49	2
20	A	42	2

1. Quel est l'effectif total ?
2. Pour chacune des variables statistiques, indiquer sa nature.
3. Déterminer les effectifs des variables « groupe sanguin » et « nombre d'enfants ».
4. On décide de regrouper les données individuelles de la variable « âge » selon le découpage suivant : $[0 ; 18[$, $[18 ; 25[$, $[25 ; 35[$, $[35 ; 49[$, $[49 ; 55[$ et $[55 ; 70[$. Déterminer la distribution des effectifs de la variable « âge » associée à ce découpage en classes.
5. Déterminer les fréquences de la variable « groupe sanguin ». Représenter graphiquement cette variable. Quel est le mode de cette variable ?

Exercice 4. En 1990, lors du recensement effectué par l'INSEE, on a étudié la répartition des ménages français selon leur type. Les résultats furent les suivants (en milliers) :

Type de ménage	H. seul	F. seule	H. + enfant(s)	F. + enfant(s)	Couples	Autres
Effectif	2171	3673	225	1325	13702	443

1. Quelle est la population étudiée ?
2. Quel est le type de la variable « type du ménage » ?
3. Quel était le nombre de ménages en 1990 ?
4. Calculer la proportion d'adultes vivant seuls en 1990.
5. Quel est le mode de cette variable ?
6. Lors d'une enquête en 1999, on a de nouveau recensé la répartition des français selon leur type. Les résultats furent les suivants (en milliers) :

Type de ménage	H. seul	F. seule	H. + enfant(s)	F. + enfant(s)	Couples	Autres
Effectif	2942	4411	259	1444	14603	489

- (a) Quel était le nombre de ménages en 1999 ?
- (b) Calculer la proportion d'adultes vivant seul en 1999.
- (c) Représenter sur un même graphique les résultats observés en 1990 et 1999.

Exercice 5. On observe les données suivantes :

Nb. d'enfants	0	1	2	3	4	5	6	7	8
Nb. de ménages	45	71	49	24	20	18	9	3	2

1. Préciser la population étudiée : les enfants ou les ménages ? Préciser la variable étudiée. Quelle est la nature de la variable étudiée ?
2. Représenter la série par un graphique.
3. Calculer la moyenne et l'écart-type.

4. Commenter la dispersion.
5. Calculer les fréquences cumulées. Quelle est la proportion de ménages ayant moins de 6 enfants (au sens large) ?
6. Indiquer le mode.

Exercice 6. Le tableau suivant indique la durée de chômage (exprimée en mois) pour un échantillon de $n = 14845$ chômeurs

Classe	Amplitude	Effectif	Fréquence	Fréq. cumulée
[0; 1[2221		
[1; 2[2338		
[2; 3[1587		
[3; 5[1989		
[5; 7[1402		
[7; 9[950		
[9; 12[1035		
[12; 15[786		
[15; 18[803		
[18; 30[1734		
Total				

1. Préciser la population et la variable étudiée. Quelle est la nature de la variable étudiée ?
2. Compléter le tableau.
3. Quelle est le pourcentage de durées ne dépassant pas 18 mois ?
4. Représenter la fonction de répartition.
5. Calculer la médiane, puis la situer sur le graphique précédent.
6. Calculer la moyenne.
7. Préciser la classe modale.

Exercice 7. L'utilisation en vase clos d'un insecticide a donné les résultats suivants sur la durée de vie des insectes (exprimée en secondes) :

Durée de vie	[0; 4[[4; 8[[8; 12[[12; x_4 [[x_4 ; 22[[22; 30[[30; 42[
Fréquence	6%	p_2	p_3	17%	14%	11%	3%

Certaines données sont illisibles et ont été remplacées par x_4 , p_2 et p_3 .

- Déterminer p_2 et p_3 en sachant que le quatrième décile est égal à 9.5 secondes.
- Déterminer x_4 sachant que la durée de vie moyenne des insectes soumis à l'insecticide est égale à 13 secondes.
- Tracer l'histogramme de la distribution.
- Situer sur le graphique précédent la médiane et la calculer. Calculer les autres quartiles.
- Calculer l'écart-type.
- Représenter les différentes caractéristiques numériques à l'aide d'un box-plot.

Exercice 8. Pour une étude médicale, on mesure la températures maximale de patients infectés par la grippe

37.1	38.7	39.6	40.6	37.6	38.8	39.7	40.7	38	39.1
39.7	40.9	38.1	39.1	39.8	40.5	38.1	39.2	39.9	41.2
38.	39.3	39.9	41.4	38.3	39.3	40	41.5	38.4	39.4
40	41.7	38.5	39.5	40.1	41.8	38.6	39.6	40.4	41.9

- Quelle est la population étudiée ?
- Quel est le type de la variable statistique ?
- Remplir le tableau suivant des données regroupées

Températures	[37, 38[[38, 39[[39, 39.5[[39.5, 40[[40, 41[[41, 42[
Fréquences						
Fréquences cumulées						
Densités de fréquence						

- Calculer la moyenne à l'aide des données regroupées.
- Calculer l'écart type à l'aide des données regroupées.
- Tracer l'histogramme des densités de fréquences.