

Statistiques et échantillonnage – Exercices

Traitement de données

- 1** Le tableau suivant provient du recueil de données effectué pendant trois ans par sept hôpitaux. Il s'agit du nombre d'admissions dues à des chutes de roller.

Sexe \ Âge	Hommes	Femmes	Total
9 ans et moins	160	183	
10 à 14 ans	694	312	
15 à 19 ans	229	47	
20 à 34 ans	174	127	
35 ans et plus	73	76	
Total			

1. Compléter le tableau.
2. Parmi les personnes hospitalisées suite à un accident de roller, déterminer la fréquence d'hommes.
3. Parmi l'ensemble des personnes hospitalisées, quelle est la fréquence d'hommes âgés de moins de 20 ans ?
4. Parmi les hommes hospitalisés, déterminer le pourcentage de personnes âgées de moins de 20 ans.
En déduire, parmi les hommes hospitalisés, la fréquence de personnes âgées d'au moins 20 ans.

- 2** Calculer la moyenne d'un test noté sur 5.

Note	1	2	3	4	5
Effectif	12	27	33	18	10

- 3** Dans une entreprise on a la répartition suivante des salaires.

Salaire S (en €)	Effectif
$1000 \leq S < 1200$	3
$1200 \leq S < 1400$	7
$1400 \leq S < 1700$	2

Déterminer une estimation du salaire moyen.

- 4** Maud veut calculer rapidement sa moyenne. Ses notes sont : 10 – 12 – 14 – 7,5 – 13 – 9,5 – 11 – 15.

Maud enlève 10 à chaque note ; elle calcule la moyenne \bar{y} des notes obtenues puis elle ajoute 10 pour obtenir sa moyenne.

1. Effectuer le calcul.
2. Calculer de la même façon la moyenne de Valentin : 9 – 8,5 – 16 – 7,5 – 12 – 13.

- 5** Dans une entreprise, le salaire moyen des 12 cadres est 2500 € et le salaire moyen des 28 ouvriers est de 1500 €. Calculer le salaire moyen d'un employé de l'entreprise.

- 6** Sur un test d'endurance effectué par 30 élèves d'une classe, la distance moyenne parcourue par les 18 garçons de la classe est de 1650 m, et la distance moyenne parcourue par les filles est 1400 m. La distance médiane parcourue par les filles est 1500 m et 1600 m pour les garçons.

1. Quelle est la moyenne pour la classe ?
2. Peut-on connaître la médiane de la classe ?

- 7** Calculer la moyenne de la série suivante.

120 000 000 000 015	120 000 000 000 008
119 999 999 999 990	119 999 999 999 986
120 000 000 000 007	120 000 000 000 000

- 8** Un professeur a corrigé 32 copies. La moyenne est 9,4 ; la médiane est 10,5 et les notes extrêmes sont 4 et 18.

1. Il envisage de remonter toutes les notes d'un point. Que deviendrait la moyenne, la médiane et l'étendue ?
2. Il envisage de remonter toutes les notes de 10 %. Mêmes questions.
3. Il corrige la copie d'un élève retardataire à laquelle il attribue 13. Calculer la nouvelle moyenne.

- 9** Vrai ou faux ? Justifier.

Dans une classe de 36 élèves, on étudie les notes du dernier contrôle d'anglais.

1. La note moyenne indique que la moitié de la classe a eu plus de cette note.
2. La note moyenne est la demi-somme des notes extrêmes.
3. Si on veut connaître la somme des notes, on multiplie la moyenne par 36.
4. Si on augmente toutes les notes de 2 points, la moyenne est augmentée de 2 points.
5. Si la plus haute note passe de 14 à 18,5, la moyenne est augmentée de 0,125 points.

- 10** Pour chacune des quatre séries déterminer la médiane et les quartiles.

Série 1 :

4,4	4,8	4,8	5,2	5,8	5,8	5,8	6,1	6,4
-----	-----	-----	-----	-----	-----	-----	-----	-----

Série 2 :

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

Série 3 :

x_i	0	1	2	3	4	5
n_i	18	21	22	9	5	2

Série 4 :

x_i	0	1	2	3	4	5
n_i	17	21	22	9	5	2

- 11** Le tableau suivant concerne la distance moyenne (en km) parcourue par les actifs ayant un emploi hors de leur commune de résidence pour les 96 départements de France métropolitaine.

Min	Q_1	Me	Q_3	Max
8,5	13,7	14,9	16,1	23,3

1. Dans combien de départements au moins la distance est-elle inférieure ou égale à 13,7 km ?
2. Interpréter la médiane.
3. La distance moyenne parcourue dans les Bouches-du-Rhône est 18,1 km. Comment peut-on situer ce département par rapport aux autres ?

- 12** On connaît la distribution des fréquences pour 57 mesures de longueur, en mètres, réalisées au cours d'une épreuve sportive.

Longueur	[0; 2[[2; 4[[4; 6[[6; 8[[8; 10]
Fréquence	0,14	0,26	0,32	0,23	0,05

Établir la répartition en effectif arrondi à 1, puis construire l'histogramme correspondant en prenant comme échelle 1 cm pour 1 mètre en abscisse et 1 cm pour un saut en ordonnée.

Notion d'échantillon

13 On lance une pièce et on observe la sortie de « pile » que l'on note P ou de « face » que l'on note F .

1. Donner un exemple d'échantillon de taille 10.
2. Quelle est la fréquence de « pile » dans cet échantillon ?

14 On tire des boules, avec remise, dans une urne contenant 2 boules rouges, 10 boules vertes et 11 boules noires.

1. Quelle est la probabilité de tirer une boule verte dans l'urne ?
2. Donner un exemple d'échantillon de taille 12.
3. Quelle est la fréquence du caractère « vert » dans cet échantillon ?

15 Une urne contient 60 jetons blancs et 20 jetons rouges.

1. On tire successivement et sans remise 6 jetons de cette urne et l'on s'intéresse au nombre de jetons blancs obtenus. A-t-on constitué un échantillon ?
2. Même question si l'on tire successivement avec remise.

16 On a effectué plusieurs lancers d'un dé équilibré à six faces et on a obtenu les résultats suivants.

Nombre de lancers	1000	2000	5000	10000
Nombre de 1	194	371	839	1663
Fréquence de 1				

1. Compléter la ligne des fréquences.
2. Quelle est la fréquence théorique du résultat ?
3. Calculer l'écart de la fréquence obtenue avec la fréquence théorique dans chacun des cas.
4. Quel phénomène a-t-on mis en évidence ?

17 La formule =ALEA() permet de générer sur un tableur un réel de l'intervalle $[0;1[$ et la formule =ENT() permet de retirer les décimales d'un réel.

Proposer une formule permettant de générer au hasard :

1. un réel de l'intervalle $[0;5[$;
2. un réel de l'intervalle $[8;13[$;
3. un réel de l'intervalle $[10;20[$;
4. un entier compris entre 0 et 5 ;
5. un entier compris entre 8 et 12 ;
6. un entier compris entre 5 et 20.

Intervalle de fluctuation et prise de décision

18 Une urne contient 12 boules, dont 3 sont rouges. On tire successivement et avec remise 100 boules de l'urne.

1. Calculer l'intervalle de fluctuation de la fréquence de boules rouges observées sur un échantillon de taille 100.
2. Que peut-on dire du nombre de boules rouges que l'on aura observé ?

19 On lance 200 fois une pièce de monnaie équilibrée.

1. Calculer l'intervalle de fluctuation de la fréquence de « pile » observés sur un échantillon de taille 200.
2. Que peut-on dire du nombre de « pile » que l'on aura observé ?

20 26 % des français se déclarent allergiques aux pollens. On étudie la fréquence f des personnes allergiques dans un échantillon de 400 personnes ; on trouve 120 personnes allergiques.

1. Calculer l'intervalle de fluctuation de la fréquence du nombre de personne allergique dans les échantillons de taille 400.
2. Le nombre d'allergiques dans cet échantillon est-il anormal ?

21 Une station de ski familiale n'attire que 25 % de skieurs habitant hors du département. Souhaitant élargir sa clientèle, la station fait réaliser des travaux au cours de l'été suivant : nouveau télésiège et canon à neige.

L'hiver suivant, voulant connaître l'impact de ses investissements, 500 skieurs sont interrogés : 150 d'entre eux habitent hors du département.

Dites si les travaux de l'été ont eu un impact sur la fréquentation des skieurs habitant hors du département.

22 D'après l'Insee, la proportion de 0-17 ans en France était de 21,95 % en 2010.

La ville de Nice compte 64 242 jeunes âgés de 17 ans ou moins parmi 343 304 habitants.

1. Calculer la fréquence f des 0-17 ans à Nice.
2. Calculer l'intervalle de fluctuation.
3. Que peut-on en conclure ?

23 Dans un lycée où 65 % des élèves étudient l'espagnol, on interroge 125 élèves.

L'intervalle de fluctuation à 95 % de la fréquence d'élèves étudiant l'espagnol dans cet échantillon est (environ)

- a. $[64,91; 65,09]$
- b. $[0,56; 0,74]$
- c. $[0,64; 0,66]$
- d. $[0,54; 0,86]$

24 Une variété de souris présente des cancers spontanés avec un taux constant de 20 %. Trois chercheurs qui expérimentent un traitement ont obtenu les résultats suivants :

- le premier chercheur a décelé 14 cancers sur 100 souris ;
- le deuxième a décelé 28 cancers sur 200 souris ;
- le troisième a décelé 700 cancers sur 5000 souris.

Quels sont les traitements que l'on peut estimer positifs ?

25 On a lancé 200 fois un dé cubique à six faces et on a obtenu les fréquences suivantes.

Face	1	2	3	4	5	6
Fréquence	0,195	0,1	0,19	0,205	0,16	0,15

1. Si l'on suppose le dé équilibré, quelle probabilité p doit-on associer à chaque face ?
2. Peut-on estimer que le dé est-il équilibré ?
3. Reprendre ces questions en supposons que ces fréquences soient issues d'un échantillon de taille 400.

26 En 1976, dans un comté du Texas, Rodrigo Partida était condamné à huit ans de prison. Il attaqua ce jugement au motif que la désignation des jurés de ce comté était discriminante à l'égard des américains d'origine mexicaine. En effet, 79 % de la population de ce comté est d'origine mexicaine, et sur les 870 personnes convoquées pour être jurés lors d'une certaine période de référence, il n'y eut que 339 personnes d'origine mexicaine. Peut-on dire que la constitution des jurys est faite de façon aléatoire ?

27 Dans un secteur d'activité où il y a autant de femmes que d'hommes, une entreprise A de 92 personnes emploie 40 % de femmes, alors qu'une autre société B compte 46 % de femmes parmi ses 3500 salariés.

À votre avis, laquelle de ces deux entreprises respecte le moins bien la parité ?