

Statistique descriptive : Répétition n°1

1. On a répertorié le nombre (N) quotidien de patients admis avec un infarctus du myocarde (IM) dans deux hôpitaux A et B. Dans l'hôpital A, on a fait des relevés pendant 80 jours et dans l'hôpital B pendant 90 jours. Les données sont résumées dans le tableau ci-dessous.

N	0	1	2	3	4	5	6	7	8	9	10	11
Fréquence (hosp.A)	6	12	19	20	13	7	2	1	0	0	0	0
Fréquence (hosp.B)	1	1	7	20	15	18	18	3	4	1	0	2

- a) Quelle est la variable étudiée ? Quel est son type?
 - b) Séparément pour les deux hôpitaux, déterminer la distribution empirique des fréquences relatives et des fréquences relatives cumulées.
 - c) Comparer les deux séries à l'aide des diagrammes en barres.
 - d) Déterminer, séparément pour les deux hôpitaux, la moyenne, les trois quartiles ainsi que l'écart-type et l'écart interquartile.
 - e) Comparer les deux séries à l'aide des boîtes à moustaches.
2. Lors d'une campagne de prévention sur l'ostéoporose, trois variables ont été observées chez 20 femmes: l'âge (années), le poids (kg), le score Sahara (sans unité) et le facteur de risque "a déjà eu des fractures sans choc violent" (1=oui, 0=non). Le score Sahara détermine si la densité osseuse est normale (> -1) ou si elle est pathologique (≤ -1).

Age	Poids	Sahara	Fracture
59	92	0,1	0
74	71	2	1
65	72	1,7	0
71	57	0,8	0
65	60	-1,2	1
51	63	-1	1
54	102	0,2	0
71	59	1	0
63	104	0,1	1
70	55	-0,6	0
72	80	-1,6	1
62	92	-0,7	0
65	68	-1,8	0
69	100	-1,5	0

52	62	0,1	0
66	68	-0,4	0
74	54	-1,2	1
73	66	-1,6	1
68	71	-0,9	1
66	69	-1,1	0

- a) Quelle est le type de chacune de ces variables ?
 - b) Pour les variables « Age » et « Poids », déterminer le mode, la médiane, la moyenne, la variance et l'écart-type.
 - c) Mesurer le lien entre ces deux variables à l'aide du coefficient de Pearson.
 - d) Construire un histogramme de la variable « Sahara » (répartir préalablement les données en 8 classes d'amplitude constante). Quelle est la classe modale ?
 - e) Calculer, séparément pour les femmes à risque et les femmes sans risque (de fracture), la moyenne du score Sahara. Que peut-on conclure ?
3. Lors d'une enquête sur l'adhésion à l'Union Européenne en Suède, un échantillon de 195 personnes donne les résultats suivants : 130 votent " oui " ($X=1$) et les autres votent " non " ($X=0$) .
- a) Quelle est la variable étudiée ? Quel est son type ?
 - b) Calculer la moyenne, les trois quartiles et la variance de cette variable X.
4. On a mesuré le poids de raisins (Kg) récoltés en 2008 sur 20 pieds de vigne pris au hasard dans le Vignoble de Branceilles (Corrèze). Voici les observations:

2.2 2.3 2.4 3.2 3.1 2.0 4.1 5.2 2.5 3.5
6.5 5.4 2.3 3.4 4.6 4.0 3.3 2.5 3.5 4.2

- a) Quelle est la variable étudiée ? Quel est son type ?
- b) Déterminer le mode, la médiane, la moyenne, les trois quartiles ainsi que la variance, l'écart-type et l'écart interquartile de la série.
- c) Dessiner la boîte à moustaches de cette série de données.
- d) Répartir les 20 observations en les 5 classes $[2,3[$, $[3,4[$, $[4,5[$, $[5,6[$, $[6,7[$ dans un tableau qui reprendra les fréquences, fréquences cumulées, fréquences relatives et fréquences relatives cumulées.
- e) Représenter l'histogramme correspondant. Quelle est la classe modale ?