Exercice 1

Les résultats obtenus à partir du lancer d'un dé à 6 faces sur 100 expériences sont donnés dans le tableau suivant :

Résultats	1	2	3	4	5	6
Nombres d'observations	15	17	14	18	20	16

Représentation de la distribution des fréquences

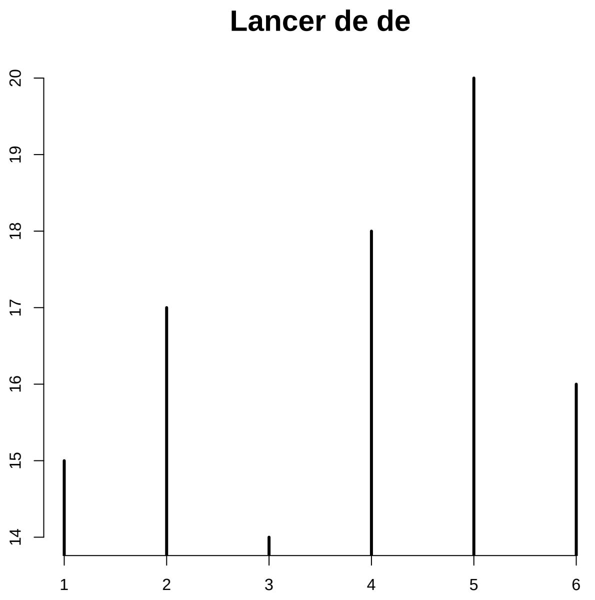


Figure 1: Diagramme en bâtons

1. Donner le mode (i.e. le résultat le plus fréquent).
   Rappel le mode ou valeur dominante désigne la valeur la plus représentée d'une variable quelconque dans une population d'objets, de personnes, de choses. Une répartition peut être unimodale ou plurimodale (bimodale, trimodale, ...), si deux ou plusieurs valeurs de la variable considérée émergent également.
   Ici le mode est 5.

2. Tracer la fonction cumulative croissante des fréquences relatives.

3. Donner la médiane et l'écart interquartile.
   $Q_1 = 2$.
   $Q_2 = 4$.
   $Q_3 = 5$.
   $EIQ = Q_3 - Q_1 = 3$ (écart interquartile).

4. Donner la moyenne des observations.

$$\text{Moyenne : } \overline{X} = \sum_{i=1}^{n} \frac{n_i}{n} x_i \Rightarrow \quad \overline{X} = 3.59$$

Moyenne et médiane sont des indicateurs-résumés de tendance centrale (ou "position").
La moyenne est sensible aux valeurs extrêmes, pas la médiane. Par exemple, une moyenne de salaires est "biaisée" s'il y a un très gros salaire.

5. Donner l'écart-type des observations.

$$\text{Ecart-type} : \sigma \simeq 1.69$$

Rappel En statistique, on appelle fréquence d'une valeur le quotient obtenu en divisant l'effectif de cette valeur par l'effectif total. Ce quotient est inférieur ou égal à 1 et est souvent exprimé en pourcentage.

Résultats	1	2	3	4	5	6
Nombres d'observations	15	17	14	18	20	16
Fréquences relatives	0.15	0.17	0.14	0.18	0.20	0.16
Fréquences relatives cumulées	0.15	0.32	0.46	0.64	0.84	1.00

Exercice 2

On donne un tableau statistique :

x_i	[20, 40[	[40, 60[	[60, 80[	[80, 100[	[100, 140[	[140, 200[
x_i^c	30	50	70	90	120	170
n_i	240	208	160	212	129	51
$\sum_i n_i \nearrow$	240	448	608	820	949	1000
$\sum_i n_i \searrow$	1000	760	552	392	180	51

1. Représenter cette série statistique.

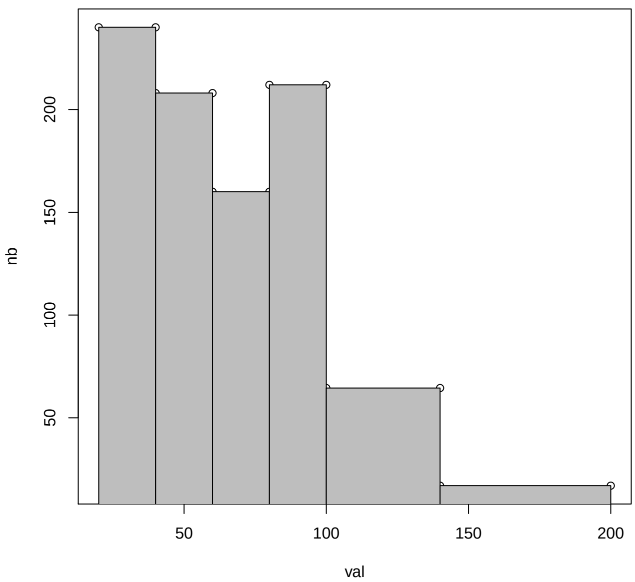
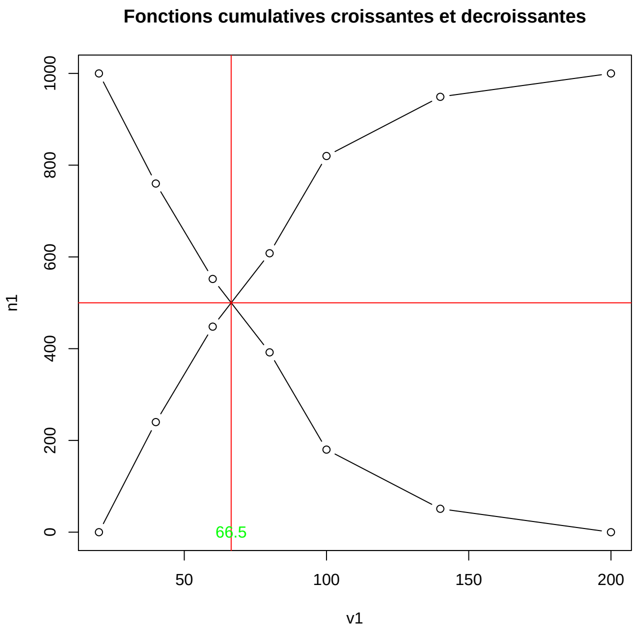


Figure 2: Histogramme - Intervalles de classe variables

2. Quelle est la classe modale.
Classe modale : [20, 40[

3. Représenter les effectifs cumulés croissants et décroissants. En déduire graphiquement la valeur de la médiane



4. Déterminer numériquement la médiane.
   Méthode détaillée :

   $N/2 = 500, \Rightarrow\ 500 \in [60,\ 80[$. On cherche l'équation de la droite $y = ax+b$ qui passe par $A(60, 448)$ et $B(80,\ 608)$.

448	=	60 a + b
608	=	80 a + b
160	=	20 a

   $a = 8 \ \Rightarrow\ b = -32$ donc $y = 8a - 32$. Pour $y = N/2$, on a $500 = 8\,x_{Q2} - 32 \Rightarrow \mathbf{x_{Q2} = 66.5}$

5. Déterminer à 0,001 près la moyenne, et l'écart-type de cette série :
   Xmean = 72.03., Var = 1380.379, std = 37.15345

6. Déterminer, de manière numérique, l'écart interquartile
   $N/4 = 250, \Rightarrow\ 250 \in [40,\ 60[$. On cherche l'équation de la droite $y = ax+b$ qui passe par $A(40, 240)$ et $B(60,\ 446)$.

240	=	40 a + b
448	=	60 a + b
208	=	20 a

   $a = 10.4 \ \Rightarrow\ b = -176$ donc $y = 10.4\,a - 176$.
   Pour $y = N/4$, on a $250 = 10.4\,x_{Q1} - 176 \Rightarrow \mathbf{x_{Q1} = 426/10.4 = 40.96}$

   $3N/4 = 750, \Rightarrow\ 250 \in [80,\ 100[$. On cherche l'équation de la droite $y = ax + b$ qui passe par $A(80, 608)$ et $B(100,\ 820)$.

608	=	80 a + b
820	=	100 a + b
212	=	20 a

$a = 10.6 \Rightarrow b = -240$ donc $y = 10.6\,a - 240$.
Pour $y = 3N/4$, on a $750 = 10.6\,x_{Q3} - 240 \Rightarrow \mathbf{x_{Q3} = 990/10.6 = 93.40}$

$$IEQ = x_{Q3} - x_{Q1} = 93.40 - 40.96 = 52.40$$

7. Dans cette question on prendra pour valeur de la moyenne 72 et pour valeur de l'écart-type 37,15 Déterminer le pourcentage de l'effectif total dans l'intervalle $\pm 1\sigma$.

$$[Xmean - \sigma, Xmean + \sigma] = [72.03 - 37.15345, 72.03 + 37.15345] = [34.877, 109.183] \Rightarrow$$

$$Xmin = 34.87\,(\approx 35) \in [20, 40] \Rightarrow 40 - 35 = 5 \Rightarrow \approx 25\% \text{ de l' effectif de la classe } = 60$$

$$Xmax = 109\,(\approx 110) \in [100, 140] \Rightarrow 110-100. \approx 10 \Rightarrow 10 \approx 1/4 = 25\% \text{ de l'effectif de la classe } \approx 129*$$

$$\text{Total\_effectif} = 60 + 208 + 160 + 212 + 32.5 = 672.5 \quad \Rightarrow p \approx 67.25\%$$

Exercice 3

Le taux de triglycérides est observé chez 250 hommes de 20 à 30 ans. On relève les résultats suivants :

Triglycérides (g/l)	[0.0, 0.6]	[0.6, 0.8]	[0.8, 1.0]	[1.0, 1.2]	[1.2, 1.4]	[1.4, ∞]
Centres des classes	0.5	0.7	0.9	1.1	1.3	1.5
Nombres d'observations	5	32	86	89	32	6
Fréquences relatives	0.020	0.128	0.344	0.356	0.128	0.024
Fréquences relatives cumulées	0.020	0.148	0.492	0.848	0.976	1.000
Eff. cumul. Croiss.	5	37	123	212	244	250
Eff. cumul. Decroiss.	250	245	213	127	38	5

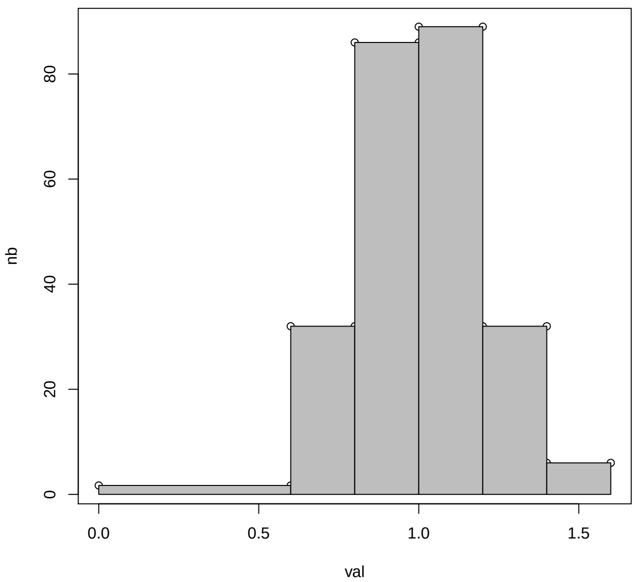
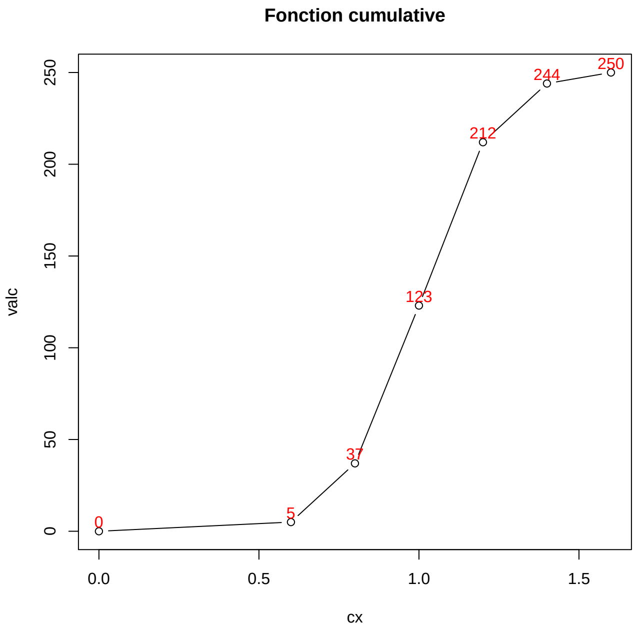


Figure 3: Histogramme - Intervalles de classe variables

1. Donner la classe modale (i.e. la classe la plus fréquente).
   Classe modale : [1.0, 1.2[.

2. Tracer la fonction cumulative croissante des fréquences relatives.



3. Donner la médiane et l'écart interquartile.
   Aide : pour avoir des classes de même longueur, on remplacera dans le calcul de la moyenne et de l'écart-type, la première par [0.4; 0.6[et la dernière par [1.4; 1.6[.

   • 2ème méthode de calcul des quartiles

   $\frac{Q_1-1}{0{,}25-0{,}492} = \frac{0{,}8-1}{0{,}148-0{,}492} \simeq 0,581 \Longrightarrow Q_1 \simeq 0,859.$

   1ere Méthode $N/4 = 62.5$. On cherche l'équation de la droite $y = ax + b$ qui passe par $A(0.8, 37)$ et $B(1,\ 123)$. $a = 430;\ \Rightarrow\ b = -317$ donc $y\ =\ 430\,x\ -\ 317$.

37	=	0.8 a + b
123	=	a + b
86	=	0.2 a

   Pour $y = N/4$, on a $62.5\ =\ 430\,x_{Q1}\ -\ 317\ \Rightarrow\ \mathbf{x_{Q1}\ =\ 0.88}$

   • 2ème méthode de calcul des quartiles
   $\frac{Q_2-1}{0{,}5-0{,}492} = \frac{1{,}2-1}{0{,}848-0{,}492} \simeq 0,561 \Longrightarrow Q_2 \simeq 1,004$ (médiane).

   1ere Méthode $N/2\ =\ 125,\ \Rightarrow\ \ 250$. On cherche l'équation de la droite $y\ =\ ax + b$ qui passe par $A(1, 123)$ et $B(1.2,\ 212)$. $a = 445;\ \Rightarrow\ b = -322$ donc $y\ =\ 445\,x\ -\ 322$.
   Pour $y = N/2$, on a $125\ =\ 445\,x_{Q1}\ -\ 322\ \Rightarrow\ \mathbf{x_{Q2}\ =\ 1.004}$

   • 2ème méthode de calcul des quartiles
   $\frac{Q_3-1}{0{,}75-0{,}492} = \frac{1{,}2-1}{0{,}848-0{,}492} \simeq 0,561 \Longrightarrow Q_3 \simeq 1,144.$

123	=	1 a + b
212	=	1.2 a + b
89	=	0.2 a

1ere Méthode $3N/4 = 187.5$. On cherche l'équation de la droite $y = ax + b$ qui passe par $A(1, 123)$ et $B(1.2,\ 212)$ (voir Q2).

- Ecart interquartile $Q_3 - Q_1 \simeq 0,285$

4. Donner la moyenne des observations avec 4 décimales.
   Moyenne : $Xmean = \frac{1}{N} \sum_{i=1}^{n} n_i\, x_i \ \simeq 1,0032$.

5. Donner l'écart-type des observations avec 4 décimales.
   Ecart-type : $\simeq 0,2025$.