Exercice n° 1 : STATISTIQUES SEANCE 1

Par définition,

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

En développant le terme mis au carré, c'est-à-dire

$$(x_i - \bar{x})^2 = x_i^2 - 2 x_i \bar{x} + \bar{x}^2,$$

nous obtenons

$$s^2 = \frac{1}{n} \left[\sum_{i=1}^{n} (x_i^2 - 2 x_i \bar{x} + \bar{x}^2) \right]$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - 2 \frac{\bar{x}}{n} \sum_{i=1}^{n} x_i + \frac{\bar{x}^2}{n} \sum_{i=1}^{n} 1$$

Etant donné que $\sum_{i=1}^{n} 1 = \underbrace{1 + 1 + \ldots + 1}_{n \times} = n$ et que par définition

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i ,$$

nous avons finalement,

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - 2 \bar{x} . \bar{x} + \frac{\bar{x}^2}{n} . n$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - 2 \bar{x}^2 + \bar{x}^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 .$$

Exercice n°2

Nous calculons successivement la nouvelle moyenne et la nouvelle variance.

1. Calcul de la nouvelle moyenne :

Avant modification, nous avions

$$\bar{x} = 5,9 = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}\sum_{i=1}^{10} x_i .$$

Soit $x_j = 8,5$, la valeur erronée, pour un certain $j \in \{1, \dots, n\}$.

Par conséquent,

$$5,9 = \frac{1}{10}\left(\sum_{i\neq j} x_i + x_j\right) = \frac{1}{10}\left(\sum_{i\neq j} x_i + 8,5\right) \quad (*)$$

Après modifications, nous avons,

$$\bar{x}' = \frac{1}{10}\sum_{i=1}^{10} x_i' = \frac{1}{10}\left(\sum_{i\neq j} x_i + x_j'\right) = \frac{1}{10}\left(\sum_{i\neq j} x_i + 6,5\right) \quad (**)$$

où $\bar{x}'$ est la nouvelle moyenne et x'_i les nouvelles données. Comme les x_i' sont égaux aux anciennes données, pour $i \neq j$, alors $x_i' = x_i$. Pour déterminer $\bar{x}'$, il nous faut déterminer $\left(\frac{1}{10}\sum_{i\neq j} x_i\right)$, ce qui est fourni par $(*)$. En effet, en isolant ce terme dans $(*)$, nous déduisons que

$$\frac{1}{10}\sum_{i\neq j} x_i = 5,9 - \frac{8,5}{10} .$$

En insérant cette expression dans $(**)$, nous obtenons finalement

$$\bar{x}' = 5,9 - \frac{8,5}{10} + \frac{6,5}{10}$$

$$= 5,9 - \frac{2}{10} = 5,9 - 0,2 = 5,7.$$

2 - Calcul de la nouvelle variance :

En considérant l'expression de la variance obtenue à l'exercice n° 1, nous savons qu'avant modifications

$$(*) \qquad 4,83 = s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i)^2 - \bar{x}^2 = \frac{1}{10}\left(\sum_{i\neq j} x_i^2 + (8,5)^2\right) - (5,9)^2$$

et qu'après modifications,

$$(**) \qquad s^{2\prime} = \frac{1}{10}\left[\sum_{i\neq j}(x_i)^2 + (6,5)^2\right] - (5,7)^2$$

Pour déterminer $s^{2\prime}$, il nous faut déterminer $\frac{1}{10}\sum_{i\neq j} x_i^2$, ce qui est fourni par $(*)$. En isolant ce terme dans $(*)$, nous obtenons

$$\frac{1}{10}\sum_{i\neq j}(x_i)^2 = 4,83 + (5,9)^2 - \frac{(8,5)^2}{10},$$

et en insérant cette expression dans $(**)$,

$$s^{2\prime} = 4,83 + (5,9)^2 - \frac{(8,5)^2}{10} + \frac{(6,5)^2}{10} - (5,7)^2$$
$$= 4,15.$$

Exercice n°3

Soient
$$x_i' = a x_i - b \quad , \quad i=1, \dots, n. \qquad (*)$$

1. Calcul de $\bar{x}$

Par définition,
$$\bar{x}' = \frac{1}{n} \sum_{i=1}^{n} x_i' \overset{(*)}{=} \frac{1}{n} \sum_{i=1}^{n} (a x_i - b)$$
$$= \frac{1}{n} \sum_{i=1}^{n} a x_i + \frac{1}{n} \sum_{i=1}^{n} (-b).$$

Etant donné que a et $(-b)$ sont des constantes, nous pouvons les sortir du signe somme, et donc
$$\bar{x}' = \frac{a}{n} \sum_{i=1}^{n} x_i + \frac{(-b)}{n} \sum_{i=1}^{n} 1$$
$$= a \bar{x} - b, \qquad (**)$$
c'est-à-dire, en d'autres termes,
$$\bar{x} = \frac{1}{a} (\bar{x}' + b)$$

2. Calcul de s^2

Par définition,
$$s'^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i' - \bar{x}')^2 \overset{(*)(**)}{=} \frac{1}{n} \sum_{i=1}^{n} \left[(a x_i - b) - (a\bar{x} - b)\right]^2$$
$$= \frac{1}{n} \sum_{i=1}^{n} \left[a x_i - b - a\bar{x} + b\right]^2 = \frac{1}{n} \sum_{i=1}^{n} \left[a(x_i - \bar{x})\right]^2$$
$$= \frac{1}{n} \sum_{i=1}^{n} a^2 (x_i - \bar{x})^2 = a^2 \cdot \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$
$$= a^2 s^2,$$

c'est-à-dire, en d'autres termes,
$$s^2 = \frac{1}{a^2} s'^2$$

3. Calcul de m_3 :

Par définition,

$$m_3' = \frac{1}{n}\sum_{i=1}^{n}(x_i' - \bar{x}')^3 \overset{(**)}{\underset{(*)}{=}} \frac{1}{n}\sum_{i=1}^{n}\left[(ax_i - b) - (a\bar{x} - b)\right]^3$$

$$= \ldots = a^3 \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3$$

comme pour le calcul de s'^2

$$= a^3\, m_3,$$

c'est-à-dire, en d'autres termes,

$$m_3 = \frac{1}{a^3}\, m_3'.$$

4. Calcul de m_4 :

Par définition,

$$m_4' = \frac{1}{n}\sum_{i=1}^{n}(x_i' - \bar{x}')^4 \overset{(**)}{\underset{(*)}{=}} \frac{1}{n}\sum_{i=1}^{n}\left[(ax_i - b) - (a\bar{x} - b)\right]^4$$

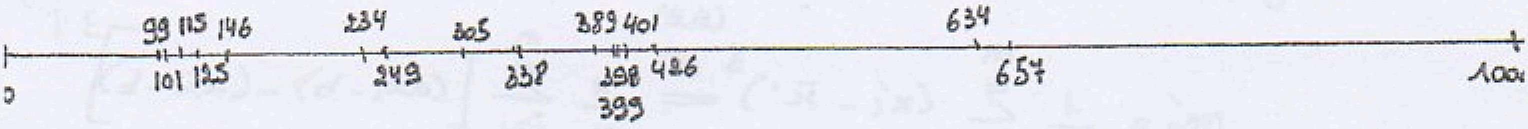
$$= \ldots = a^4 \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4$$

$$= a^4\, m_4,$$

c'est-à-dire,

$$m_4 = \frac{1}{a^4}\, m_4'.$$

Exercice n°4 :



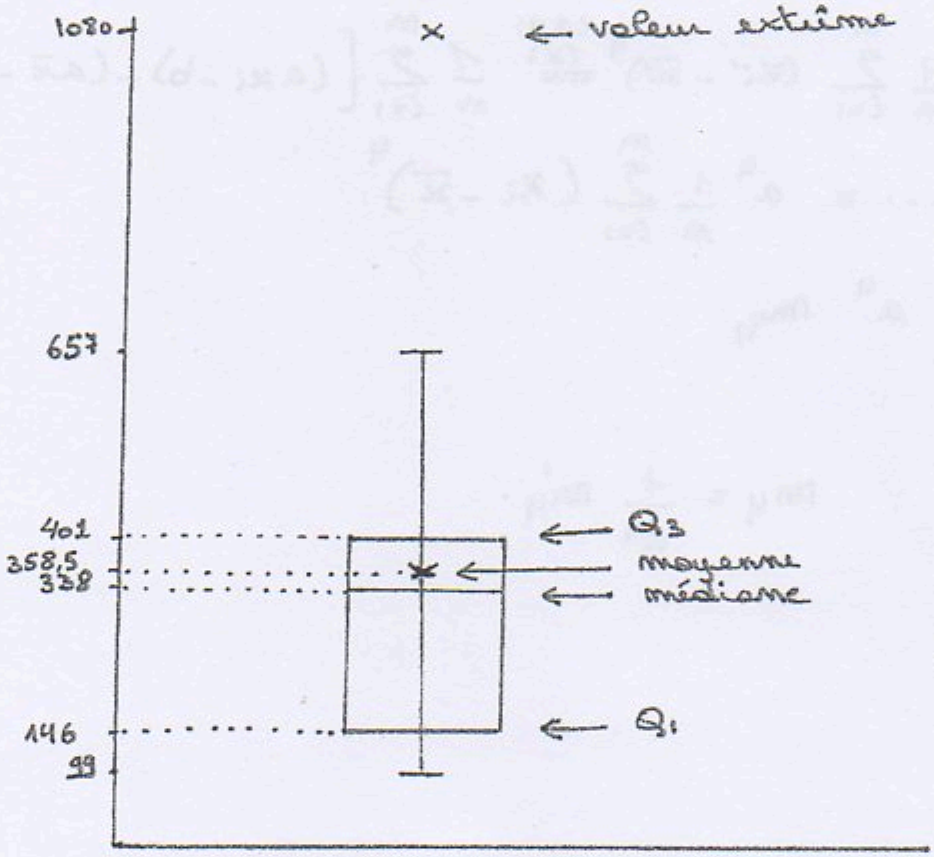
- médiane : $\tilde{x} = 338$
- 1er quartile : $Q_1 = 146$
- 3ème quartile : $Q_3 = 401$

$\Rightarrow$ intervalle interquartile : $[146, 401]$.

$\Rightarrow |IQ| = 401 - 146 = 255$

- moyenne : $\bar{x} = \frac{1}{17} \sum_{i=1}^{17} x_i \simeq 358,5$.

Dessinons le box plot.



calculons
- $Q_3 + 1,5|IQ| = 401 + 1,5 \cdot 255 = 783,5$
- $Q_1 - 1,5|IQ| = 146 - 1,5 \cdot 255 = -236,5$.

La plus grande valeur inférieure à 783,5 est <u>657</u>.

La plus petite valeur supérieure à -236,5 est <u>99</u>.

Exercice n°5 :

	n_i	N_i	f_i	F_i
$C_1 = [20800, 22199]$	18	18	0,18	0,18
$C_2 = [22200, 22999]$	27	45	0,27	0,45
$C_3 = [23000, 23799]$	43	88	0,43	0,88
$C_4 = [23800, 24599]$	9	97	0,09	0,97
$C_5 = [24600, 25999]$	3	100	0,03	1

1. Dessiner l'histogramme et la fonction de distribution correspondants :

1.1. L'histogramme :

⚠ les longueurs des intervalles sont différentes. Pour l'histogramme nous en tenons compte en divisant les valeurs des n_i par la longueur des intervalles correspondants $|C_i|$.

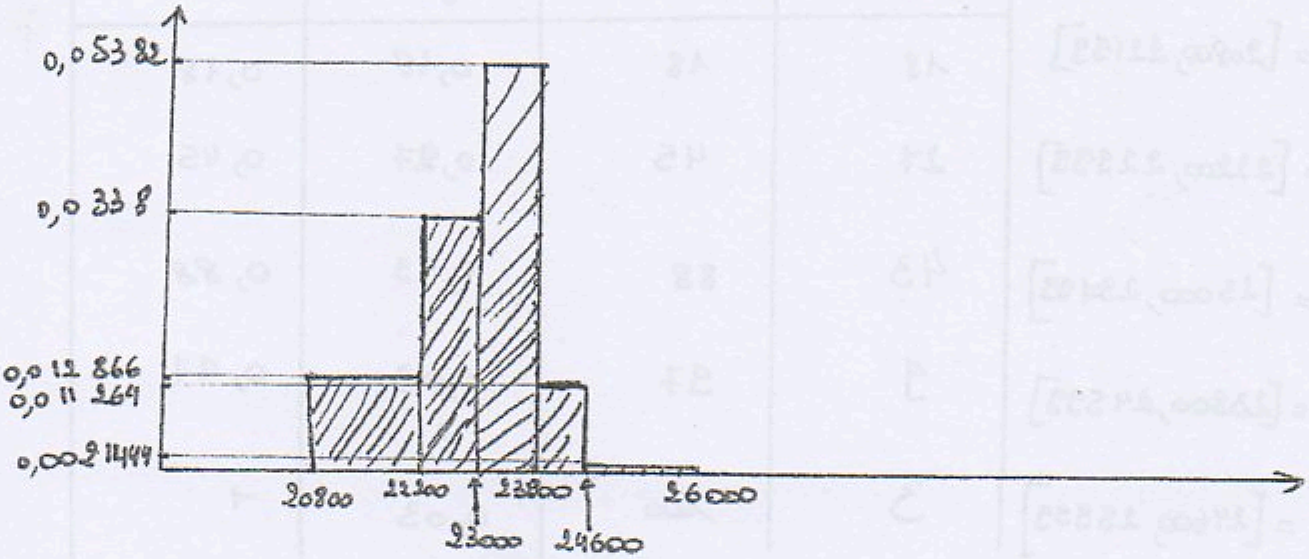
Tailles des classes : $|C_1| = 1400$ $|C_2| = |C_3| = |C_4| = 700$ $|C_5| = 1400$

$i / \lvert C_i \rvert$
012866
0338
05382
011264
0021444

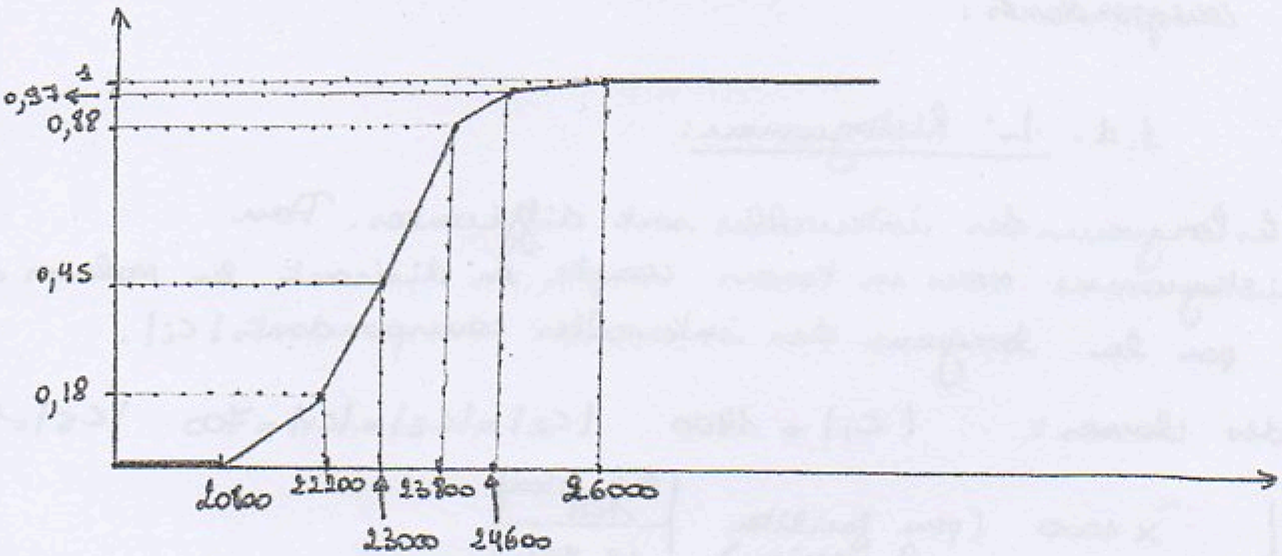
$\times 1000$ (pour faciliter le dessin)

$\frac{n_i}{\lvert C_i \rvert} \times 1000$
12,866
33,8
53,82
11,264
2,1444

stat TP1 p4



1.2. Fonction de distribution



2. Calcul de la moyenne et de l'écart-type

2.1. La moyenne : Elle est calculée en utilisant pour "x_i" la moitié de nos intervalles c_i !!!

Utilisons l'exercice 1.3. et effectuons une transformation linéaire de nos données :

$$x_i' = a x_i - b$$

$$= \frac{1}{1000} x_i - 21,4995$$

c_i	x_i	x_i'
$c_1 = [20800, 22199]$	21499,5	0
$c_2 = [22200, 22999]$	22599,5	1,1
$c_3 = [23000, 23799]$	23399,5	1,9
$c_4 = [23800, 24599]$	24199,5	2,7
$c_5 = [24600, 25999]$	25299,5	3,8

$$\bar{x}' = \frac{1}{100} \sum_{i=1}^{5} x_i' \, n_i \simeq 1,479.$$

Grâce à l'exercice 1.3, nous savons que $\bar{x} = \frac{1}{a} (\bar{x}' + b)$, ce qui donne :

$$\bar{x} = \frac{1}{\frac{1}{1000}} (1,479 + 21,4995) = 22\,970,5.$$

2.2. L'écart-type :

Calculons la variance pour les données x_i' :

$$s'^2 = \frac{1}{100} \sum_{i=1}^{5} n_i \, (x_i' - \bar{x}')^2$$

$$= 0,804459.$$

Grâce à l'exercice 1.3, nous savons que $s^2 = \frac{1}{a^2} s'^2$, ce qui donne pour l'écart-type

$$s = \sqrt{s^2} = \frac{1}{a} s' = \frac{1}{a} \sqrt{s'^2}$$

c'est-à-dire

$$\sigma = \frac{1}{\frac{1}{1000}} \cdot \sqrt{0{,}804459} \simeq 897.$$

Conclusion : L'ancienne moyenne était de 25 123 secondes et l'ancien écart-type de 130 secondes. La moyenne a donc baissé (résultat attendu) et l'écart-type a augmenté ce qui signifie qu'il y a une plus grande dispersion des temps d'autonomie autour de la moyenne. Il y a donc de très bonnes autonomies et de très mauvaises autonomies (par rapport à la moyenne) après 50 recharges et, globalement en moyenne, le temps d'autonomie a baissé.