Série de TD N°02 – Statistique Descriptive

Exercice 1 : Le tableau suivant donne le nombre d'étudiants inscrits à un cours de mathématiques selon le type de baccalauréat :

Type de baccalauréat	Mathématiques	Sciences	Techniques
Nombre inscrits	120	260	20

1. Déterminer : la population, sa taille, les modalités, le caractère étudié et sa nature.
2. Représenter graphiquement cette série statistique en fonction des fréquences f_i.

Exercice 2 : On jette cent fois de suite cinq pièces de monnaie. A chaque jet, on a compté le nombre de "Pile" obtenu et relevé les résultats suivants :

1 3 5 4 1 2 2 4 1 3 3 3 4 2 2 3 1 4 3 2 0 2 4 1 3 5 0 3 2 2 2 3 4 4 1 1 3 3 4 2 2 2 3 3 4 1 1
1 2 2 3 4 4 2 3 3 3 2 2 2 1 1 4 3 2 2 2 3 2 3 3 4 1 3 2 3 2 3 1 5 0 2 2 3 2 2 3 2 3 2 0 3 2 2
1 2 4 3 4 2

1. **a)** Quelle est la population étudiée (en précisant sa taille et l'unité statistique) ?
   c) Quel est le caractère étudié (en précisant sa nature et ses modalités) ?
2. Dresser le tableau statistique de la distribution et la représenter graphiquement.
3. Calculer les effectifs cumulés et les représenter graphiquement.
4. Déterminer le fonction de répartition de cette distribution et tracer son graphe.
5. Calculer la médiane (M_e), le mode (M_0) , la moyenne ($\overline{X}$) et l'écart type (σ_X).
6. Calculer l'intervalle interquartile $Q_3 - Q_1$.
7. Quel est le nombre de jets où on a obtenu trois "Pile" ou plus ?

Exercice 3 : On donne la distribution de N logements d'un immeuble collectif des enseignants selon leurs superficies X (en m^2) :

Superficie X	[0, 40[	[40, 80[	[80, 120[
Effectif	4	n_2	6

1. Compléter le tableau sachant que la superficie moyenne est de 64 m^2.
2. Représenter graphiquement cette distribution ainsi que le mode.
3. Calculer le mode et l'écart type (σ_X) de cette distribution.
4. Comment s'appelle la valeur α de X telle que $F(\alpha) = 1 - F(\alpha)$, F étant la fonction de répartition de X . Calculer α (interpréter) ?.
5. Tracer la courbe des fréquences cumulées croissante et décroissante.
6. Déterminer la proportion de logements dont la superficie est inférieure à $\overline{X} + \sigma_X$.

Exercice 4 : Soit la distribution des 200 employés d'une entreprise selon leur salaire annuel exprimé en DA :

Salaire	[50,60[	[60,70[	[70,90[	[90,100[	[100,130[
Effectif	20	60	50	40	30

1. Tracer le polygône des fréquences et calculer le mode.
2. Calculer l'intervalle interquartile $Q_3 - Q_1$.

Corrigé Série de TD
Proba - Stat.

Exo 1 :

Type de Baccalauréat	Maths	Sciences	Techniques
Nombre d'inscrits	120	260	20
f_i	0,3	0,65	0,05
θ_i	108°	234°	18°

1) **Déterminer :**

Population : Les étudiants

Taille : 400

Modalités : Maths, sciences, Techniques.

Caractère étudié : Type de Baccalauréat

Nature : qualitative nominale.

2) **Représentation graphique :**

$$\theta_i = \frac{n_i \times 360°}{n} = f_i \times 360°$$

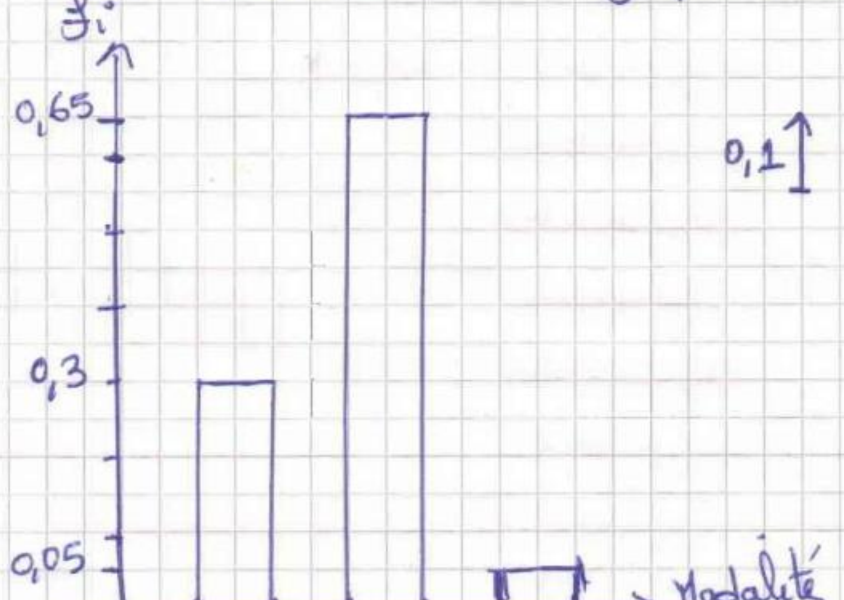


Diagramme en tuyaux d'orgues.

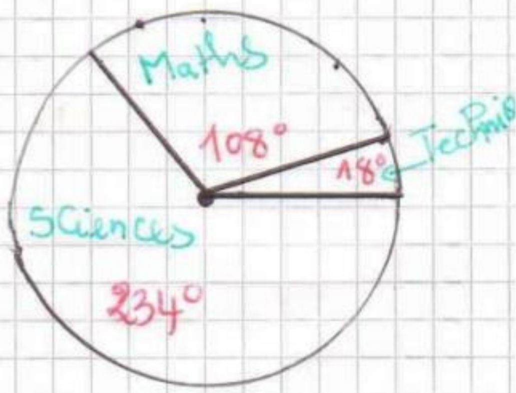


Diagramme circulaire

Ex 02 :

1) a) Population : 100 jetés

Taille : 100

Individu : un jet

b) Caractère : Le nombre de "Pile" obtenu.

Modalités : 0, 1, 2, 3, 4, 5

Nature : Quantitative Discrètes.

2) Tableau Statistique :

x_i	n_i	f_i	$N_i\nearrow$	$N_i\swarrow$	$F_i\nearrow$	$n_i x_i$	$n_i x_i^2$
0	4	0,04	4	100	0,04	0	0
1	15	0,15	19	96	0,19	15	15
2	34	0,34	53	81	0,53	68	136
3	29	0,29	82	47	0,82	87	261
4	15	0,15	97	18	0,97	60	240
5	3	0,05	100	3	1	15	75
Total	100	1	/	/	/	245	727

Représentation graphique :

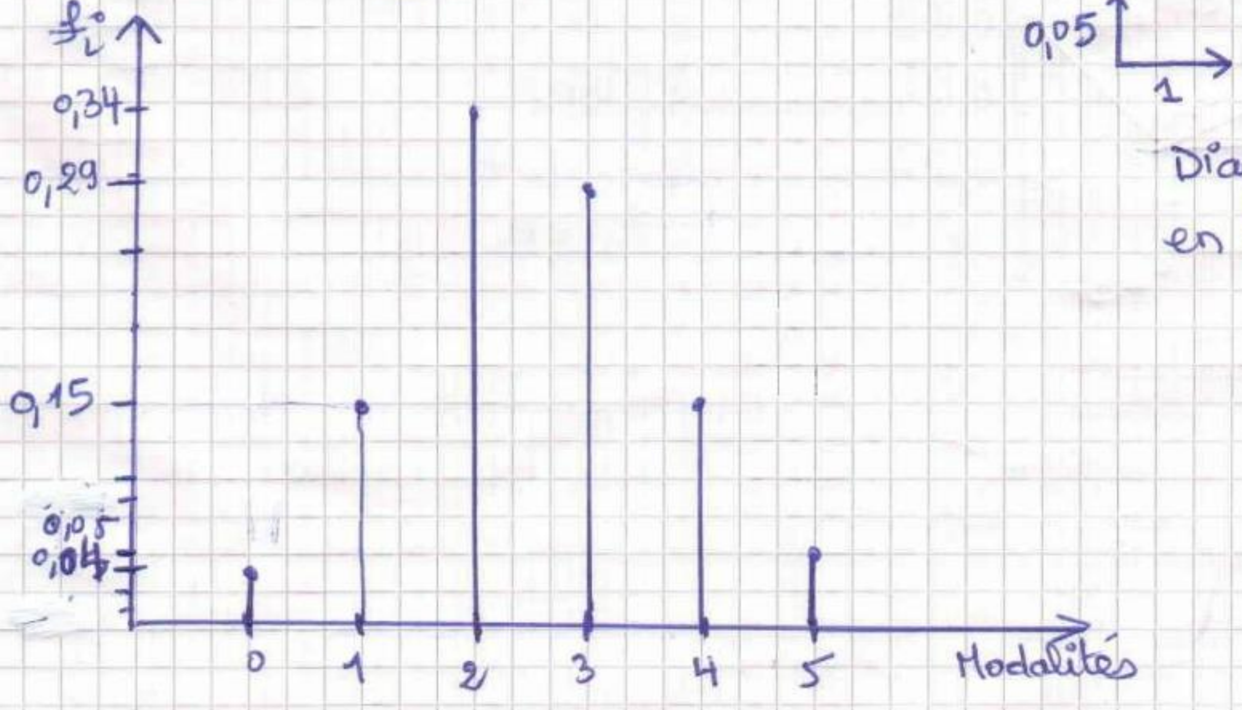
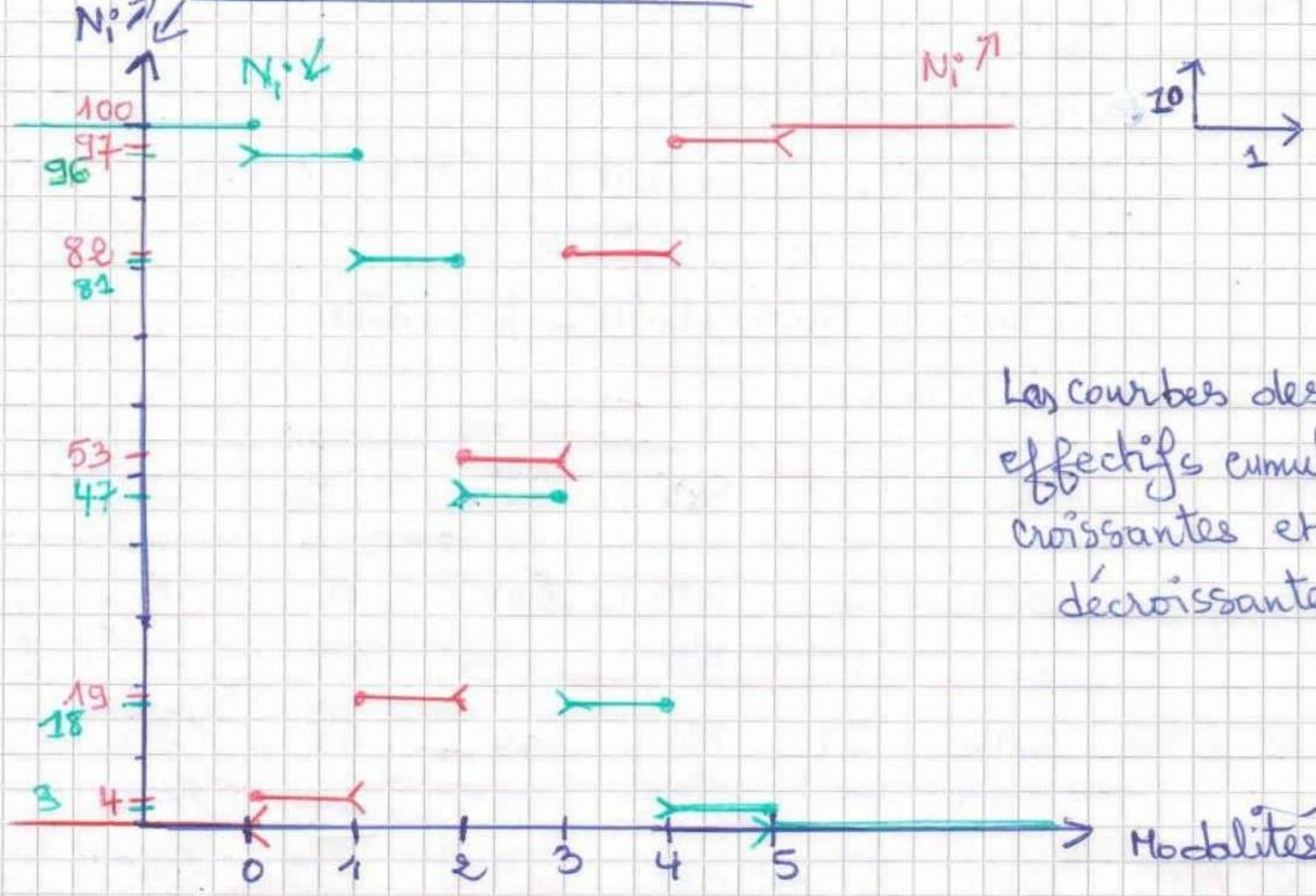


Diagramme en bâtons.

P2

3) Représentation graphique des effectifs cumulées croissants et décroissants :

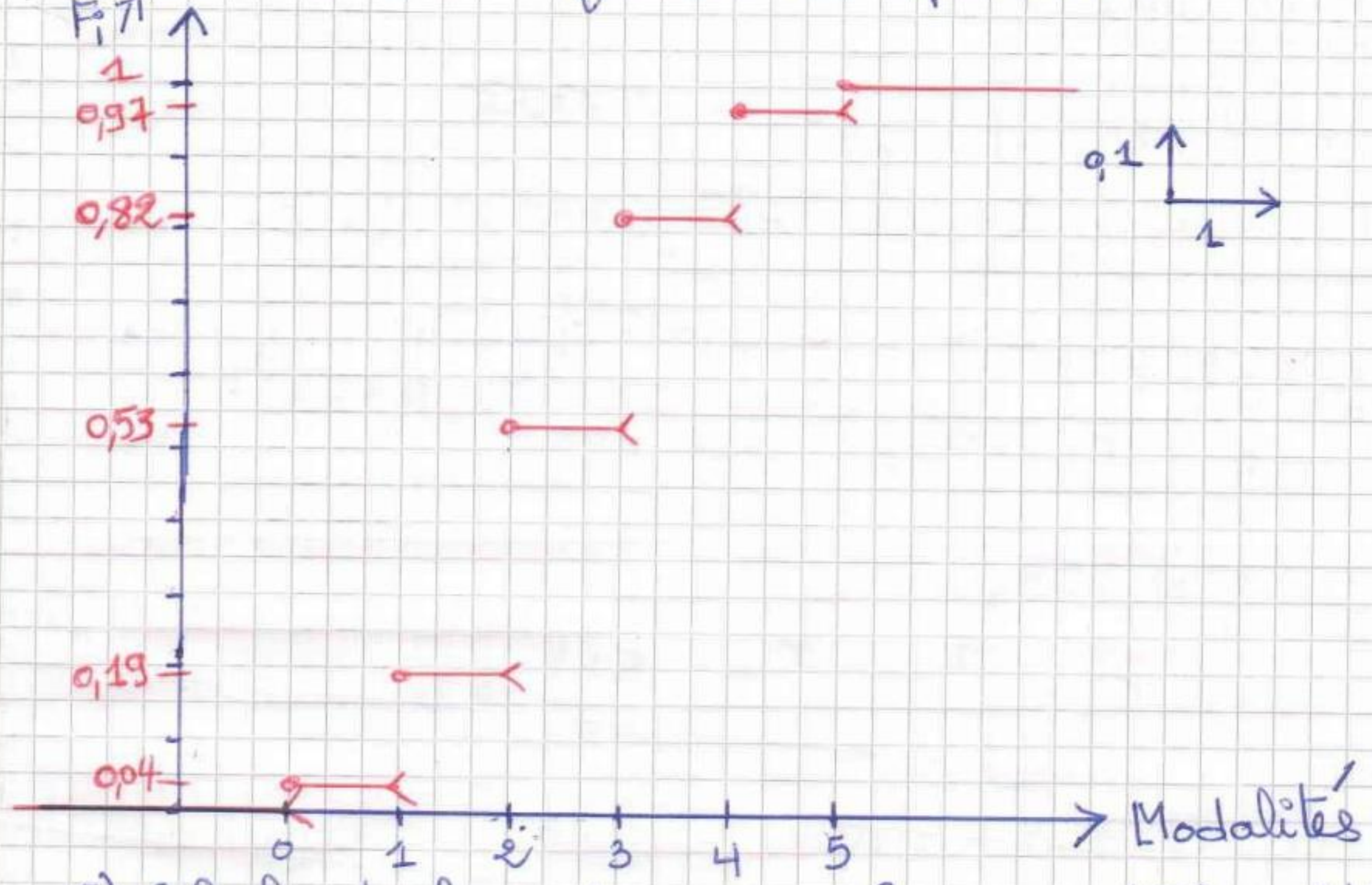


Les courbes des effectifs cumulés croissantes et décroissantes

4) La fonction de répartition :

$$F(x) = \begin{cases} 0 & x < 0 \\ 0,04 & 0 \leq x < 1 \\ 0,19 & 1 \leq x < 2 \\ 0,53 & 2 \leq x < 3 \\ 0,82 & 3 \leq x < 4 \\ 0,97 & 4 \leq x < 5 \\ 1 & x \geq 5. \end{cases}$$

La courbe de la fonction de répartition



5) Calcule de la Mediane (Me), la moyenne ($\bar{x}$) et l'écart type (σ_x) :

Mediane :

N = 100 , N est pair , $N = 2 \times p \Rightarrow p = 50$

$$Me = \frac{x_p + x_{p+1}}{2} = \frac{x_{50} + x_{51}}{2} = \frac{2+2}{2} = 2$$

$$\boxed{Me = 2}$$

Moyenne :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{6} n_i x_i = \frac{1}{100}(245) = 2.45$$

Variance :

$$V(x) = \frac{1}{n} \sum_{i=1}^{6} n_i x_i^2 - \bar{x}^2 = \frac{1}{100}(727) - (2.45)^2$$

$$= 7.27 - 6,0025 = 1,2675$$

L'écart type :

$$\sigma_x = \sqrt{V(x)} = \sqrt{1,2675} = 1,1258.$$

6) Calculer l'intervalle interquartile $Q_3 - Q_1$.

$Q_1 \Rightarrow \frac{1}{4} N = 25$

$$Q_1 = \frac{x_{25} + x_{26}}{2} = \frac{2+2}{2} = 2.$$

Q_3 : $\frac{3}{4} N = 75$

$$Q_3 = \frac{x_{75} + x_{76}}{2} = \frac{3+3}{2} = 3.$$

Donc $Q_3 - Q_1 = 3 - 2 = 1$.

7) Quel est le nombre de jets où on a obtenu 3 piles ou plus ?

$$Nb_{pile \geq 3} = n_4 + n_5 + n_6 = 29 + 15 + 3 = 47 \text{ jets}.$$

Exercice 3 :

1) Compléter le tableau sachant que la superficie moyenne est de $64\,m^2$:

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{3} n_i x_i \iff N\overline{X} = \sum_{i=1}^{3} n_i x_i$$

$$\iff (4+n_2+6)\times 64 = 4\times 20 + n_2\times 60 + 6\times 100$$

$$(10+n_2)\times 64 = 80 + n_2\times 60 + 600$$

$$640 + 64n_2 = 680 + 60n_2$$

$$64n_2 - 60n_2 = 680 - 640$$

$$4n_2 = 40$$

$$n_2 = \frac{40}{4} = 10$$

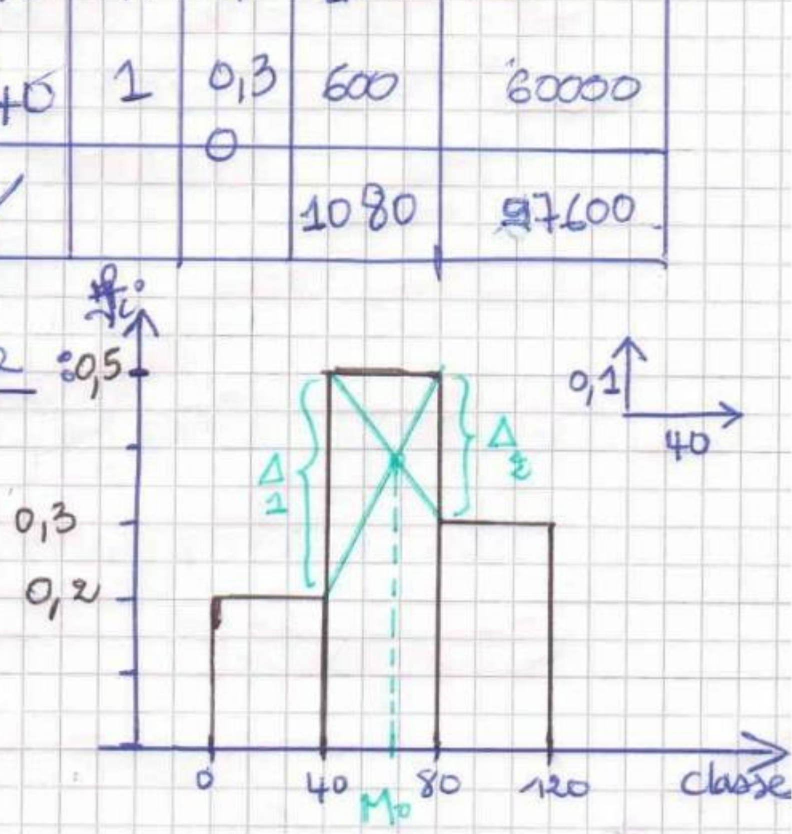
$\boxed{n_2 = 10}$

$$N = \sum_{i=1}^{3} n_i = 4 + 10 + 6 = 20 \qquad \boxed{N = 20}$$

Tableau Statistique :

classes	n_i	f_i	x_i	a_i	$F_i\nearrow$	$F_i\searrow$	$n_i x_i$	$n_i x_i^2$
[0, 40[	4	0,20	20	40	0,2	1	80	1600
[40, 80[	n_2 = 10	0,50	60	40	0,7	0,8	600	36000
[80, 120[	6	0,30	100	40	1	0,3	600	60000
Total	N=20	1	/	/			1080	97600

2) Représentation graphique



Histogramme

P6

Mode : La classe modale est $[40, 80[$

$$M_o = e_i + a_i \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

$$= 40 + 40 \frac{0,3}{0,3+0,2}$$

$$= 40 + (40 \times 0,6)$$

$$= 64$$

$\Delta_1 = 0,5 - 0,2 = 0,$

$\Delta_2 = 0,5 - 0,3 = 0,2$

$\Delta_1 + \Delta_2 = 0,3 + 0,2 = 0,5$

$$\boxed{M_o = 64}$$

3) calculer l'écart type :

$$V(x) = \frac{1}{N} \sum_{i=1}^{3} n_i x_i^2 - \bar{x}^2 = \frac{1}{20}(97600) - (64)^2$$

$$= 4880 - 4096 = 784 \qquad \boxed{V(x) = 784}$$

$$\sigma_x = \sqrt{V(x)} = \sqrt{784} = 28 \qquad \boxed{\sigma_x = 28}$$

4) Comment s'appelle la valeur de α :

On a F : étant la fonction de répartition

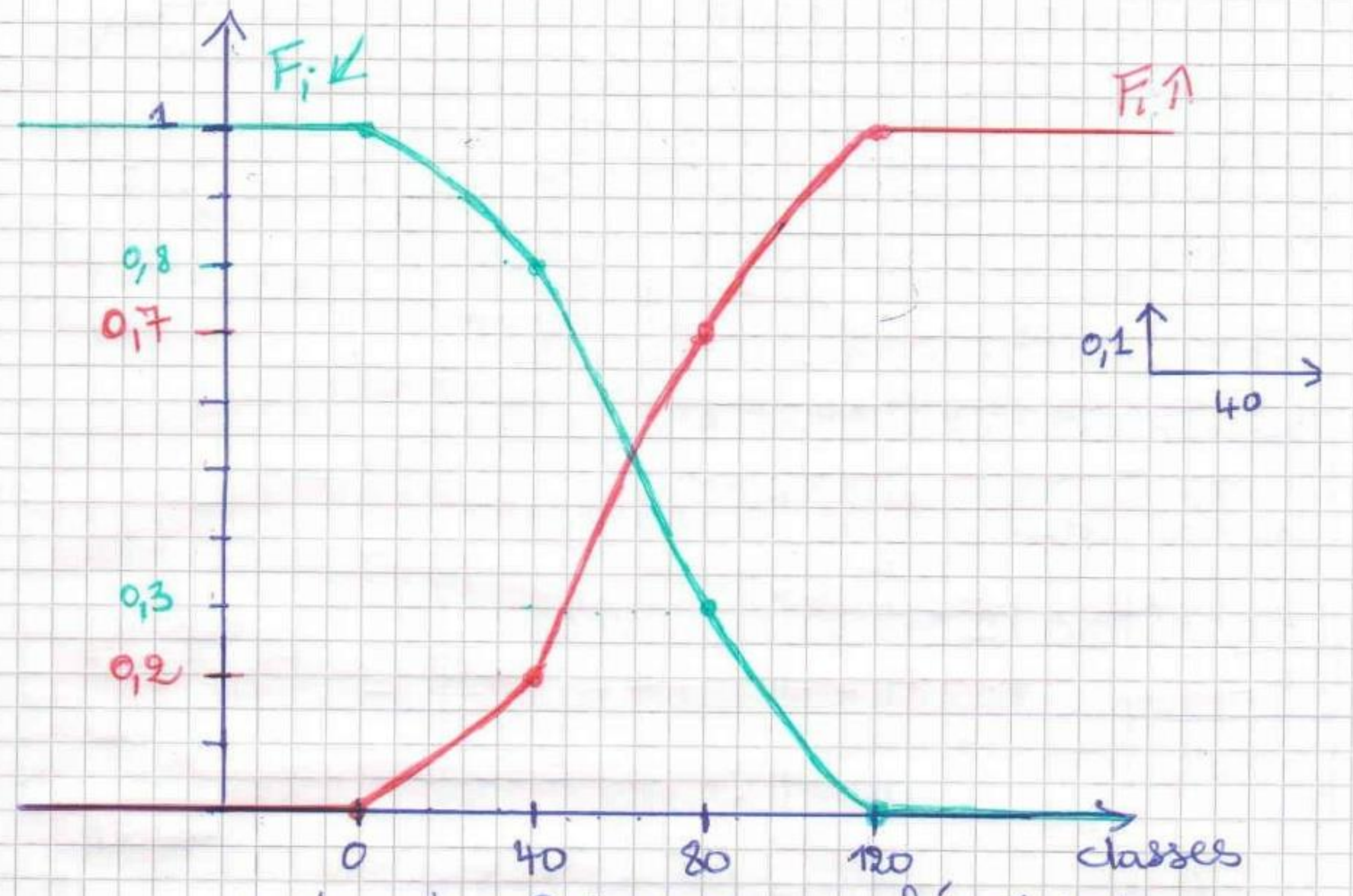
$$F(\alpha) = 1 - F(\alpha) \Leftrightarrow F(\alpha) + F(\alpha) = 1$$

$$\Leftrightarrow 2F(\alpha) = 1$$

$$\Leftrightarrow \boxed{F(\alpha) = \frac{1}{2}}$$

Donc α représente la Médiane de la distribution.

5) Tracer la courbe des fréquences cumulés croissoutes et décroissantes :



courbe des fréquences cumulées ↗ et ↙

6) Déterminer la proportion de logement dont la superficie est $< \bar{X} + \sigma_X$

$\bar{X} = 64$, $\sigma_x = 28$, $\bar{X} + \sigma_x = 64 + 28 = 92 \in [80, 120[$

$F(\bar{X} + \sigma_x) = F(92)$

$F(x) = F(e_{i-1}) + \frac{\frac{f_i}{}}{a_i}(x - e_{i-1}) \qquad x \in [e_{i-1}, e_i[$

$F(92) = F(80) + \frac{0,3}{40}(92 - 80)$

$= 0,7 + 0,09 = 0,79 \qquad \boxed{F(92) = 0,79}$

Donc, il y'a 79% des logements dont la superficie est inférieur à $(\bar{X} + \sigma_x = 92)$.

P8

Exercice 3 (suite) : question 4, Calcule de α.

On a : ***α = Me (la médiane)*** avec $F(Me) = 0.5$.

$Me \in [e_{i-1}, e_i[$ donc $Me \in [40, 80[$: classe modale.

$$Me = e_{i-1} + \frac{a_i}{f_i}\left(F(Me) - F(e_{i-1})\right)$$

$$= 40 + \frac{40}{0,5}(0.5 - F(40))$$

$$= 40 + 80\,(0.5 - 0.2)$$

$$= 40 + 80 \times 0.3$$

$$= 40 + 24 = 64$$

Donc

$$Me = 64.$$

Interprétation : 50% des logements (soit 10 logements) possèdent une superficie mois de 64m^2

et 50% d'autres ont une superficie plus de 64m^2 .

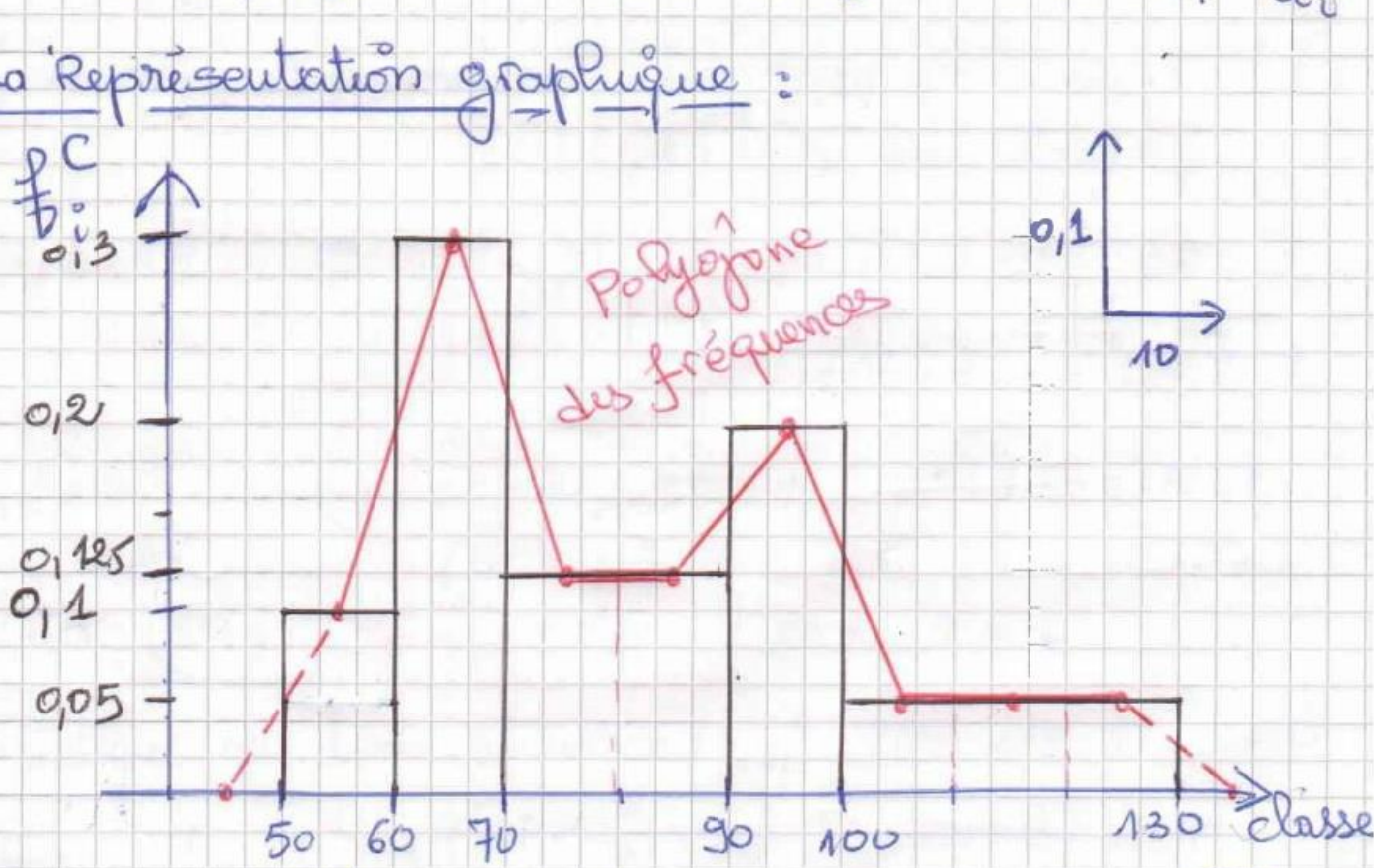
Exo4 :

1) **Le Tableau Statistique :**

Classes	n_i	f_i	a_i	$F_i \nearrow$	f_i^c
[50,60[	20	0,1	10	0,1	0,1
[60,70[	60	0,3	10	0,4	0,3
[70,90[	50	0,25	20	0,65	0,125
[90,100[	40	0,2	10	0,85	0,2
[100,130[	30	0,15	30	1	0,05
Total	200	1	/	/	/

$a = PGCD(a_i) = PGCD(10,20,30) = 10$, $f_i^c = \frac{f_i}{a_i} \times$

La Représentation graphique :



Histogramme

2) Calculer l'intervalle interquartile $Q_3 - Q_1$:

<u>$Q_1 = ?$:</u> $F(Q_1) = 0,25$ $\quad Q_1 \in [60, 70[$

$$Q_1 = e_{i-1} + \frac{a_i}{f_i}\left(F(Q_1) - F(e_{i-1})\right)$$

$$= 60 + \frac{10}{0,3}(0,25 - 0,1)$$

$$= 60 + 5 = 65 \qquad \boxed{Q_1 = 65}$$

<u>$Q_3 = ?$:</u> $F(Q_3) = 0,75$ $\quad Q_3 \in [90, 100[$

$$Q_3 = e_{i-1} + \frac{a_i}{f_i}\left(F(Q_3) - F(e_{i-1})\right)$$

$$= 90 + \frac{10}{0,2}(0,75 - 0,65)$$

$$= 90 + 5 = 95 \qquad \boxed{Q_3 = 95}$$

L'intervalle interquartile est

$$Q_3 - Q_1 = 95 - 65 = 30.$$