Chapitre 13 Statistiques

Faire savoir

A. L'essentiel du chapitre

I. Série statistique présentant un regroupement en classes.

Dans ce chapitre on étudiera des séries statistiques à modalités regroupées en classes.
Une telle série, comportant p classes de bornes : a_0 ; a_1 ; a_2 ; ; …ap ($a_0 < a_1 < \ldots < a_p$) est notée :
($[a_{i-1}$; ai[; ni) $_{1 \leq i \leq p}$, ou ni est l'effectif de la classe $[a_{i-1}$; ai[

I. 1 Effectifs cumulés, fréquences cumulées

Dans une série statistique ($[a_{i-1}$; ai[; ni) $_{1 \leq i \leq p}$:
L'effectif cumulé croissant de la classe $[a_{k-1}$; $a_k[$ est :

$$\sum_{i=1}^{i=k} n_i = n_1 + n_2 + \ldots + n_k$$

L'effectif cumulé décroissant de la classe $[a_k$; $a_p[$ est :

$$\sum_{i=k}^{i=p} n_i = n_k + n_{k+1} + \ldots + n_p .$$

La fréquence cumulée croissante (respectivement décroissante) s'obtient en divisant l'effectif cumulée croissant (respectivement décroissant) par l'effectif total..

Exemple

On a relevé la distance parcourue par chacun des 150 taxis d'une compagnie entre leur mise en circulation et leur première panne. Les résultats de cette enquête sont résumés dans le tableau ci- dessous.
Les distances étant exprimées en milliers de kilomètre.

Distance	[0 ; 5[	[5 ; 7[	[7 ; 9[	[9 ; 15[
Effectif (n_i)	15	78	36	21
Fréquence (f_i)	10%	52%	24%	14%

On dresse les tableaux des effectifs cumulés croissants et décroissants.

Distance parcourue	[0 ; 5[	[5 ; 7[	[7 ; 9[	[9 ; 15[
Effectif cumulé croissant	15	93	129	150
Effectif cumulé décroissant	150	135	57	21

Remarque

On obtient facilement, les résultats, des fréquences cumulées, en divisant les effectifs cumulés par l'effectif total.

Histogramme, polygone

Reprenons l'exemple précédent : la série de cet exemple est représentée graphiquement par l'histogramme ci-dessous :

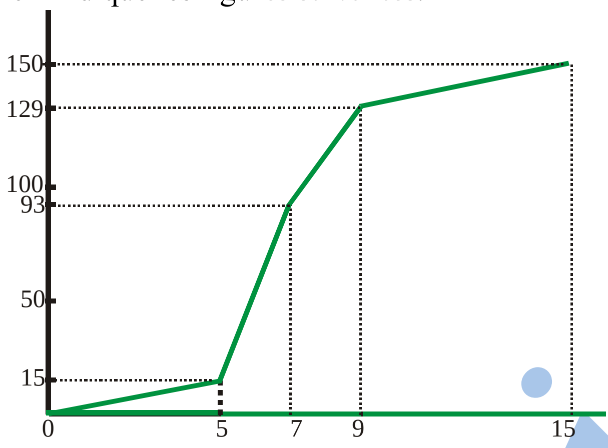
Dans le tableau précédent, la troisième colonne, par exemple, s'interprète ainsi : 129 taxis ont parcouru au plus 9 000 km et 57 taxis ont parcouru au moins 7000 km avant d'avoir leur première panne.

On en déduit le tableau suivant :

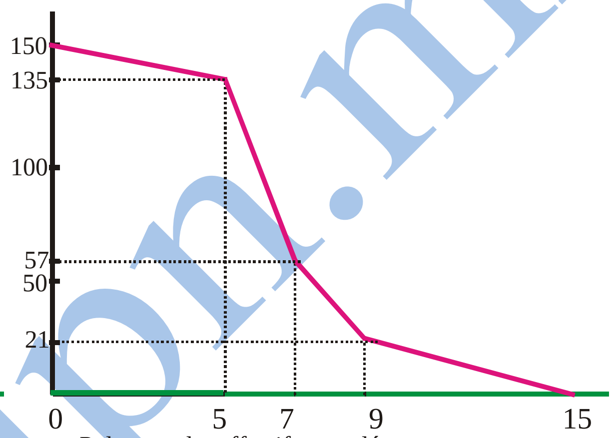
ai	0	5	7	9	15
Nombre de taxis ayant parcouru au plus ai($\times$ 1000 km)	0	15	93	129	150
Nombre de taxis ayant parcouru au moins ai($\times$ 1000 km)	150	135	57	21	0

Dans ce tableau, la somme des effectifs de chaque colonne est égale à l'effectif total.

Ces résultats peuvent être représentés par des polygones des effectifs cumulés croissants ou décroissants comme l'indique les figures suivantes.



Polygone des effectifs cumulés croissants (Fig.1)



Polygone des effectifs cumulés décroissants (Fig.2)

On construit de manière analogue, les polygones des fréquences cumulées croissantes et décroissantes.

I. 2) Caractéristiques de position

Classe modale

Définition

Soit une série statistique présentant un regroupement en classes. On appelle classe modale de cette série toute classe d'effectif maximal.

Exemple

Dans l'exemple précédent, la classe modale est [5 ; 7[.

Remarques

- une série statistique peut avoir plusieurs classes modales.
- le centre d'une classe modale est appelé mode de la série statistique

Médiane

Définition

Soit une série statistique présentant un regroupement en classes d'effectif total N.

On appelle médiane de cette série, le nombre réel M tel que : le nombre d'individus de modalité supérieure à M et le nombre d'individus de modalité inférieure à M soient tous égaux à $\frac{N}{2}$.

Exemple

Reprenons l'exemple précédent (distance parcourue par les taxis).

L'effectif total de la série est 150.

On représente sur le même graphique les polygones des effectifs cumulés croissants et décroissants.

Leur points d'intersection a pour ordonnée 75.

Soit x_0 son abscisse, on a, $x_0 \in [5 ; 7[$. Donc : $\frac{75-15}{x_0-5} = \frac{93-15}{7-5}$.

On en déduit : $x_0 = 5 + \frac{60}{39}$, d'où x = 6,54

On estime que la moitié des taxis ont parcouru au plus 6540 km avant la première panne.

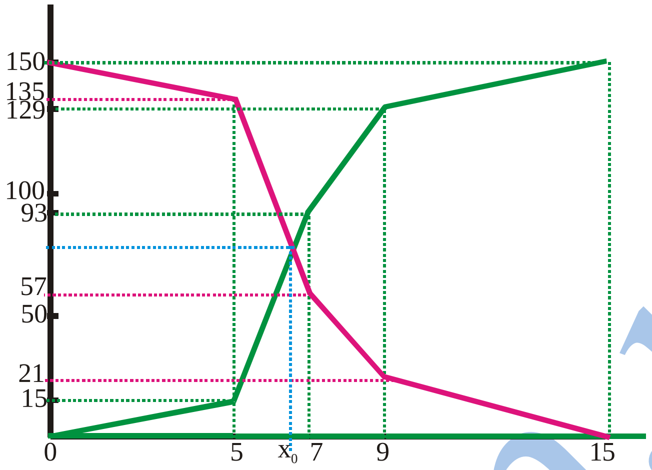


Fig.3

Moyenne

Définition

Soit $([a_{i-1} ; ai[; ni)_{1\leq i \leq p}$ une série statistique .

On appelle moyenne de cette série la moyenne $\overline{X}$ de la série statistique $(x_i ; n_i)_{1\leq i \leq p}$, ou x_i ; est le centre de la classe $[a_{i-1} ; ai[$.

Classe	Centre de la classe x_i	Effectif n_i	$n_i\ x_i$
[0 ; 5[	2,5	15	37,5
[5 ; 7[	6	78	468
[7 9 [	8	36	288
[9 ; 15[	12	21	252
Total		150	1045,5

Si N est l'effectif total de la série, on a : $\overline{X} = \frac{1}{N}\sum_{i=1}^{i=p} n_i x_i$.

Exemple

Reprenons toujours l'exemple précédent, on en déduit du tableau ci-contre que :

$\overline{X} = \frac{1045}{150}$ =6,97.

Les taxis parcourent donc en moyenne 6970 km avant leur premier panne.

Classe	Centre de la classe x_i	Effectif n_i	$n_i\ x_i$
[0 ; 5[	2,5	15	37,5
[5 ; 7[	6	78	468
[7 9 [	8	36	288
[9 ; 15[	12	21	252
Total		150	1045,5

Quartiles

Définitions

On appelle :

- Premier quartile d'une série statistique, et on désigne par Q1, la valeur telle que 25% des valeurs prises par la variable (donc 25% de l'effectif total N) lui soient inférieures, et 75% des valeurs lui soient supérieures.
- Troisième quartile d'une série statistique, et on désigne par Q3, la valeur telle que 75% des valeurs prises par la variable lui soient inférieures et 25% lui soient supérieures.

Le second quartile Q2 est, d'après la définition, la médiane M.

I. 3) Caractéristiques de dispersion

Définitions

Soit $([a_{i-1} ; a_i[, ni)_{1\leq i \leq p}$ une série statistique d'effectif N et de moyenne $\overline{X}$.

Pour tout nombre entier i tel que $1\leq i \leq p$, on désigne par xi le centre de la classe $[a_{i-1} ; a_i[$,

L'écart moyen est le nombre réel e_m tel que : $e_m = \frac{1}{N}\sum_{i=1}^{i=p} ni(xi-\overline{x})$.

Donc, $e_m = \dfrac{n_1|x_1 - \overline{x}| + n_2|x_2 - \overline{x}| + \dots n_p|xp - \overline{x}|}{N}$

La variance est le nombre réel V tel que : $V = \frac{1}{N}\sum_{i=1}^{i=p} ni(xi-\overline{x})^2$

Donc, $V = \dfrac{n_1(x_1 - \overline{x})^2 + n2(x_2 - \overline{x})^2 + \dots n_p(xp - \overline{x})^2}{N}$.

L'écart type est le nombre réel tel que : $\sigma = \sqrt{V}$.

Remarque

Dans la pratique, le calcul de la variance se fait à l'aide de la formule de Koenig : $V = \left(\frac{1}{N}\sum_{i=1}^{i=p} n_i x_i^2\right) - \overline{x}^2$.

Exemple

Il s'agit toujours de reprendre l'exemple du I.1.

Classe	Centre de la classe (x_i)	Effectif (n_i)	$\lvert x_i - \overline{x}\rvert$	$n_i\lvert x_i - \overline{x}\rvert$	x^2_i	$n_i\, x^2_i$
[0 ; 5[	2,5	15	4,47	67,05	6,25	93,75
[5 ;7[	6	78	0,97	75,66	36	2808
[7 ;9[	8	36	1,03	37,08	64	2304
[9 ;15[	12	21	5,03	105,63	144	3024
Total		150		285,42		8229,75

On en déduit du tableau que :

$e_m = \frac{285,42}{150} \approx 1,9028$; $V = \frac{8229,75}{150} - (6,97)^2 = 6,2841$; $\sigma = 2,5068$

II. Séries statistique à deux caractères

On peut, sur une population, étudier deux caractères quantitatifs.

La modalité associée à chaque individu est alors un couple de nombres réels. On construit ainsi une série statistique à deux caractères, ou série double.

II.1) Organisation des données

Exemple 1

On a relevé le poids (en kg) et la taille (en cm) de 30 personnes ; on a obtenu les résultats suivants :

Poids(x)	65	68	62	62	68	68	59	71	74	68	68	74	71	65	65
Taille (y)	165	177	174	168	165	171	165	177	174	171	165	174	174	174	171
Poids (x)	62	65	68	71	65	74	74	71	65	77	74	62	77	68	71
Taille (y)	174	174	171	171	174	168	177	174	165	180	177	168	180	171	174

Ces données permettent de définir deux séries statistiques à un caractère : $(x_i ; n_i)$ et $(y_j ; n_j)$ représentées par les deux tableaux suivants :

xi	59	62	65	68	71	74	77
ni	1	4	6	7	5	5	2

yj	165	168	171	174	177	180	…
nj	5	3	6	10	4	2	…

Soit X = $\{59 \; ; \; 62 \; ; \; 65 \; ; \; 68 \; ; \; 71 \; ; \; 74 \; ; \; 77\}$ et $y = \{165 \; ; \; 168 \; ; \; 171 \; ; \; 174 \; ; \; 177 \; ; \; 180\}$.

A chaque couple $(x_i ; y_j)$ de l'ensemble $X \times Y$, on associe le nombre d'élèves ayant le poids xi et la taille yj. Ce nombre est noté n_{ij}.

On définit ainsi une série statistique a deux caractères ; n_{ij} est appelé effectif de la modalité (xi ; yj).

Cette série, notée $(x_i ; y_j ; n_{ij})$ est représentée par le tableau à double entrée ci-contre :

Les totaux obtenus dans la dernière ligne du tableau sont les effectifs de la série (xi ; ni) ; ceux de la dernière colonne sont les effectifs de la série (yj ; nj).
Ces effectifs apparaissent ''en marge'' du tableau à double entrée.
Les séries statistiques $(x_i ; n_i)$ et $(y_j ; n_j)$ sont appelées séries marginales de la série double $(x_i ; y_j ; n_{ij})$.

y_j \ x_i	59	62	65	68	71	74	77	Total
165	1	0	2	2	0	0	0	5
168	0	2	0	0	0	1	0	3
171	0	0	1	4	1	0	0	6
174	0	2	3	0	3	2	0	10
177	0	0	0	1	1	2	0	4
180	0	0	0	0	0	0	2	3
Total	1	4	6	7	5	5	2	30

II.2) Nuage de points associés à une série double

Soit $(x_i ; y_j ; n_{ij})$ une série double.

Le plan étant muni d'un repère orthogonal.

L'ensemble de points Mij de coordonnées (xi ; yj) est appelé nuage de points associés à la série.

Dans le cas où les effectifs des modalités (xi ; yj) ne sont pas tous égaux, on représente ce nuage de points de deux façons :

- Représentation par points pondérés : on indique à côté de chaque point M_{ij}, l'effectif nij (Fig.1).
- Représentation par tâches : chaque point M_{ii} est remplacé par un disque dont l'aire est proportionnelle à n_{ij} .

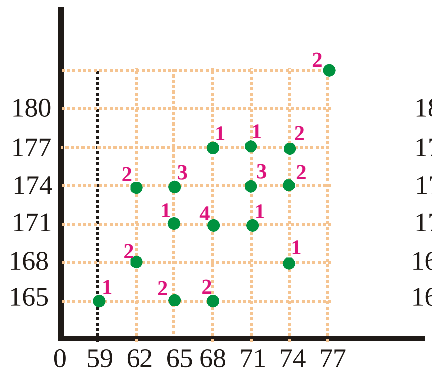


Fig. 1

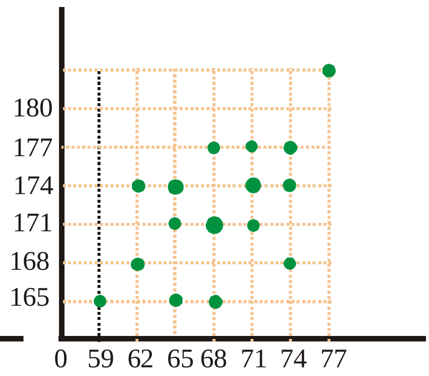


fig.2

Point moyen d'un nuage

Définition

Soit (xi ; yj ; nij) une série statistique à deux caractèrs quantitatifs.
On appelle point moyen du nuage de points représentant une série, le point de cordonnées $(\bar{x}\ ;\ \bar{y})$ ou $\bar{x}$ et $\bar{y}$ désignent les moyennes respectives des séries marginales $(x_i\ ;\ n_i)$ et $(y_j\ ;\ n_j)$.

Exemple 2

Les notes obtenues par 10 élèves aux épreuves de biologie et de physique d'un examen sont indiquées dans le tableau suivant :

Biologie (x)	9	12	5	6	9	14	3	6	12	14	
Physique (y)	10	13	8	10	13	17	5	8	16	16	

On remarque que dans cette série, les effectifs des modalités sont tous égaux à 1.

$\bar{x} = \dfrac{3+5+2\times6+2\times9+2\times12+2\times14}{10} = 9.$; $\bar{y} = \dfrac{5+2\times8+2\times10+2\times13+2\times16+17}{10} = 11,6.$

Le point moyen du nuage associé à la série est G(9 ; 11 , 6).

II.2) Ajustement linéaire

On considère dans ce paragraphe une série statistique à deux caractères x et y tel que l'effectif de chacune des modalités est égale à 1.
Une telle série sera notée $(x_i\ ;\ y_i)_{1\leq i\leq N}$, ou N est l'effectif total de la série et (xi ; yi) est la modalité du $i^{ème}$ individu.

Notion d'ajustement linéaire

Ajuster un nuage de points consiste à déterminer une courbe simple passant '' le plus près possible'' des points du nuage.
Si la courbe recherchée est une droite, l'ajustement est dit linéaire, il s'agit, en pratique, de déterminer deux droites appelées droites de régression.

Droites de régression

Définitions

Soit $(x_i\ ;\ y_i)_{1\leq i\leq N}$ une série statistique à deux caractères x et y , d'effectifs total N.
La covariance de cette série est le nombre réel, noté cov(x ;y), tel que :

- $\text{cov}(x\ ;y)=\left(\dfrac{1}{N}\sum_{i=1}^{N}x_iy_i\right)-\bar{x}.\bar{y}$.

Soit $(x_i\ ;\ y_i)_{1\leq i\leq N}$ série statistique à deux caractères x et y telle que : $V(x)\neq 0$.
La droite de régression de y en x passe par le point moyen du nuage associé à cette série et à pour coefficient directeur

- $\dfrac{\text{Cov}(x;y)}{V(x)}$.

Une équation de cette droite est : $y - \bar{y} = \dfrac{\mathrm{Cov}(x;y)}{V(x)}(x - \bar{x})$.

Soit $(x_i\,;\,y_i)_{1 \le i \le N}$ une série statistique à deux caractères x et y , telle que : $V(y) \neq 0$.
La droite de régression de x en y passe par le point moyen du nuage associé à cette série et à pour Equation : $x - \bar{x} = \dfrac{\mathrm{Cov}(x;y)}{V(y)}(y - \bar{y})$.

❖ **Coefficient de corrélation linéaire**

Définition

Soit $(x_i\,;\,y_i)_{1 \le i \le N}$ une série statistique à deux caractères x et y , telle que : $V(x) \neq 0$ et $V(y) \neq 0$.

Le coefficient de corrélation linéaire de cette série est le nombre réel r tel que : $r = \dfrac{\mathrm{Cov}(x;y)}{\sqrt{V(x)}.\sqrt{V(y)}}$.

Exemple

Reprenons l'exemple 2.
les calculs sont généralement disposés dans un tableau de la façon suivante :

x_i	y_i	$x_i.\,y_i$	x^2_i	y^2_i
9	10	90	81	100
12	13	156	144	169
5	08	40	25	64
6	10	60	36	100
9	13	117	81	169
14	17	238	196	289
3	5	15	9	25
6	8	48	36	64
12	16	192	144	256
14	16	224	196	256
90	116	1180	948	1492

On en déduit que : $\mathrm{Cov}(x;y) = \dfrac{1180}{10} - 9 \times 11{,}6$;

$\mathrm{Cov}(x;y) = 13{,}6$.

$V(x) = \dfrac{948}{10} - 9^2 = 13{,}8$.

La droite de régression de y en x est donc la droite (d) d'équation :

$y - 11{,}6 = \dfrac{13{,}6}{13{,}8}(x - 9)$, c'est-à-dire : $y = 0{,}986x + 2{,}73$.

On en déduit aussi que :

$V(y) = \dfrac{1492}{10} - 11{,}6^2 = 14{,}64$.

La droite de régression de x en y est donc la droite (d') d'équation :

$x - 9 = \dfrac{13{,}6}{14{,}64}(y - 11{,}6)$, c'est-à-dire : $y = 1{,}076x + 1{,}92$.

On en déduit aussi que le coefficient de corrélation linéaire est : $r = \dfrac{13{,}6}{\sqrt{13{,}8}.\sqrt{14{,}64}}$ ⊔ 0,957.

Savoir-faire

A. Applications

Séries statistiques présentant un regroupement en classes

Exercice 1.

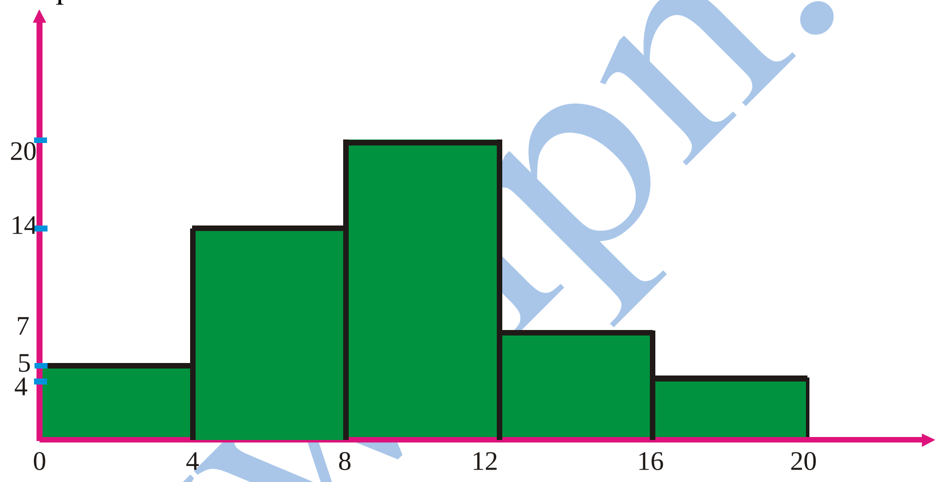
Les moyennes des notes obtenues par les 50 candidats à un concours se répartissent comme suit :

Classe	[0 ; 4[	[4 ; 8[	[8 ; 12[	[12 ; 16[	[16 ; 20[
Effectif	5	14	20	7	4

1) Construire un histogramme représentant cette série.
2) Dresser le tableau des effectifs cumulés croissants et décroissant, puis construire les polygones des effectifs cumulés croissants et décroissants.
3) Dresser les tableaux des fréquences cumulées croissantes et décroissantes.
4) Déterminer la classe modale, la médiane et le premier quartile de cette série statistique.
5) Calculer la moyenne $\overline{x}$ de cette série statistique.

Solution

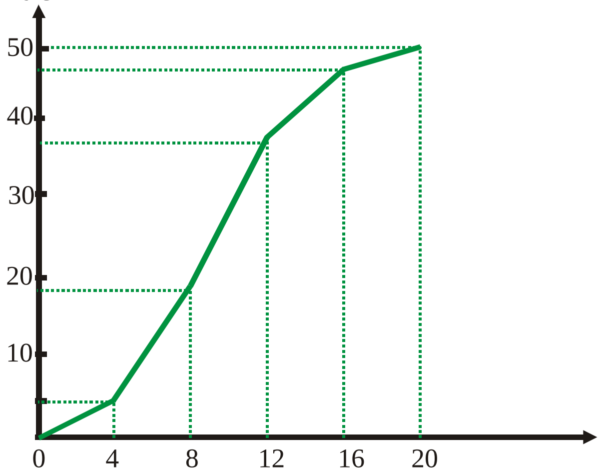
1) L'histogramme représentant cette série est le suivant :



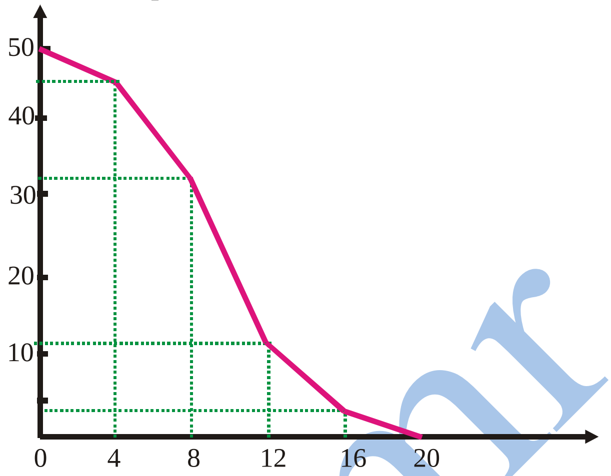
2) et 3, le tableau suivant donne la réponse aux questions posées

Classe	[0 ; 4[	[4 ; 8[	[8 ; 12[	[12 ; 16[	[16 ; 20[
Effectif	5	14	20	7	4
Effectif cumulé croissant	5	19	39	46	50
Effectif cumulé décroissant	50	45	31	11	4
Fréquence	10%	28%	40%	14%	8%
Fréquence cumulée croissante	10	38%	78%	92%	100%
Fréquence cumulée décroissante	100%	90%	62%	22%	8%

Les polygones des effectifs cumulés croissants et décroissants se présentent comme suit :

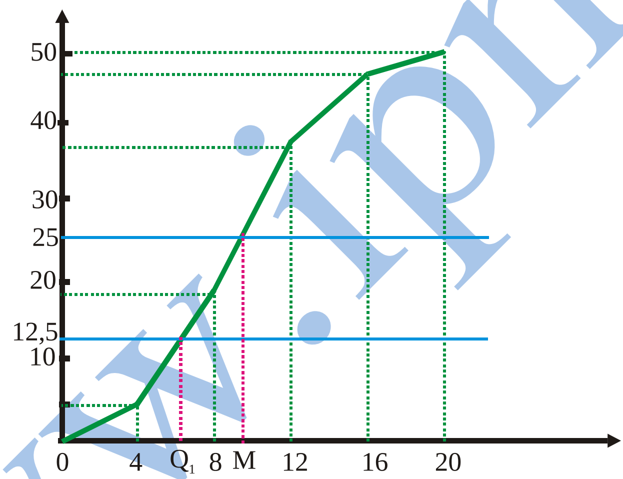


Polygone des effectifs cumulés croissants



Polygone des effectifs cumulés décroissants

4) la classe modale de cette série est [8 ; 12[.
Traçons les droites y = 25 et y = 12,5 sur le même graphique avec le polygone des effectifs cumulés croissants.



Nous constatons que $M \in [8\ ;\ 12[$ et que $Q_1 \in [4\ ;\ 8[$.

Donc : $\dfrac{25-19}{M-8} = \dfrac{39-19}{12-8} \Rightarrow \dfrac{6}{M-8} = 5 \Rightarrow M = \boxed{9,2}$

La médiane de cette série est M = $\boxed{9,2}$.

Nous avons aussi, $\dfrac{12,5-5}{Q_1-4} = \dfrac{19-5}{8-4} \Rightarrow \dfrac{7,5}{Q_1-4} = \dfrac{7}{2} \Rightarrow\ Q_1 = 4 + \dfrac{15}{7} \Rightarrow Q_1 \approx 6,14$,

Le premier quartile de cette série est $\boxed{Q_1 \approx 6,14}$

5)

Classe	Centre de la classe (x_i)	Effectif (n_i)	$n_i x_i$
[0 ; 4[	2	5	10
[4 ; 8[	6	14	84
[8 ; 12[	10	20	200
[12 ; 16[	14	7	98
[16 ; 20[	18	4	72
Total	50		464

La moyenne de cette série est $\overline{x} = \dfrac{464}{50} = 9,28$.

Exercice 2.

le tableau suivant donne la répartition en classes d'amplitudes 5 des tailles en cm de 40 individus.

Classe	Effectif (ni)	Centre de la classe (xi)	$\lvert x_i - \overline{x}\rvert$	$n_i \lvert x_i - \overline{x}\rvert$	$n_i\, x^2_i$
[155 ;160[	3				
[160 ;165[	8				
[165 ;170[	10				
[170 ;175[	10				
[175 ;180[	7				
[180 ;185[	2				
Total	40				

1) Calculer la moyenne de cette série et compléter le tableau.
2) Déterminer l'écart moyen et l'écart type de cette série.

Solution

La moyenne de la série est : $\overline{x} = \dfrac{157,5\times 3+162,5\times 8+167,5\times 10+177,5\times 7+182,5\times 2}{40}$.

$\overline{x} = \dfrac{462,5+1300+1675+1725+1242,5+365}{40} = \dfrac{6780}{40}$, donc, $\boxed{\overline{x} = 169,5}$.

Classe	Effectif (ni)	Centre de la classe (xi)	$\lvert x_i - \overline{x}\rvert$	$n_i \lvert x_i - \overline{x}\rvert$	$n_i\, x^2_i$
[155 ;160[	3	157,5	12	36	74918,75
[160 ;165[	8	162,5	7	56	211250
[165 ;170[	10	167,5	2	20	280562,5
[170 ;175[	10	172,5	3	30	297562,5
[175 ;180[	7	177,5	8	56	220543,75
[180 ;185[	2	182,5	13	26	66612,50
Total	40			224	1150950

2) $e_m = \dfrac{224}{40} = 5,6$; $V = \dfrac{1150950}{4040} - (169,5)^2 = 43,5$; $\sigma = \sqrt{V} = \sqrt{43,5} \approx 6,59$.

Exercice 3

Dans une maternité, on a relevé, pour chacune des vingt naissances, d'une journée, l'âge x de la mère et le poids y du nouveau-nés les résultats sont regroupés dans le tableau ci-dessous.

x	22	18	20	20	16	22	26	22	18	22	18	26	20	26	18	18	22	26	22	20
y	3,2	2,8	3,2	3,6	2,8	2,8	3	2,8	3	3	3,4	3,6	3,2	3	3	3,4	2,8	2,6	3	3,2

1) Présenter ces données dans un tableau à double entrée.
2) Déterminer les séries marginales associées aux caractères x et y.
3) Représenter par des tâches le nuage de points associés à cette série double.
4) Placer son point moyen G.

Solution

1) présentation des données dans un tableau à double entrée.

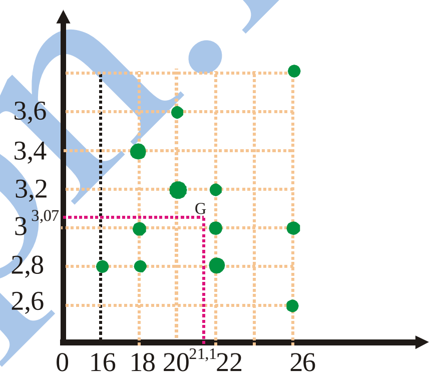
x / y	16	18	20	22	26	Total
2,6	0	0	0	0	1	1
2,8	1	1	0	3	0	5
3	0	2	0	2	2	6
3,2	0	0	3	1	0	4
3,4	0	2	0	0	0	2
3,6	0	0	1	0	1	2
Total	1	5	4	6	4	20

Les séries marginales :

x_i	16	18	20	22	26
n_i	1	5	4	6	4

y_j	2,5	2,8	3	3,2	3,4	3,6
n_j	1	5	6	4	2	2

3) Représentation du nuage de points associés à cette série.



$$\bar{x} = \frac{16\times1+18\times5+20\times4+22\times6+26\times4}{20} = 21,1,$$

$$\bar{y} = \frac{2,6\times1+2,8\times5+3\times6+3,2\times4+3,4\times2+3,6\times2}{20} = 3,07\ .$$

Le point moyen est G(21,1 ; 3,07).

Exercice 4

On considère la série (x_i ; y_i) déterminée par le tableau :

xi	-3	-1	0	0
y_i	1	-1	4	2

Déterminer une équation de la droite de régression de y en x et une équation de la droite de régression de x en y.

Calculer le coefficient de corrélation de cette série.

Solution

x_i	y_i	x_iy_i	x^2_i	y^2_i
-3	1	-3	9	1
-1	-1	1	1	1
0	4	0	0	16
5	2	10	25	4
1	6	8	35	22

$\bar{x} = \frac{1}{4}$; $\bar{y} = \frac{6}{4} = \frac{3}{2}$;

$\mathrm{Cov}(x;y) = \frac{8}{4} - \frac{1}{4} \times \frac{3}{2} = 2 - \frac{3}{8} = \frac{13}{8}$

$V(x) = \frac{35}{4} - \frac{1}{16} = \frac{139}{16}$; $\qquad V(y) = \frac{22}{4} - \frac{9}{4} = \frac{13}{4}$.

La droite de régression de y en x a pour équation : $y - \bar{y} = \frac{\mathrm{Cov}(x;y)}{V(x)}\left(x - \bar{x}\right)$.

Donc : $y - \frac{3}{2} = \frac{\frac{13}{8}}{\frac{139}{16}}\left(x - \frac{1}{4}\right) \Rightarrow y = \frac{26}{139}x - \frac{26}{556} + \frac{3}{2} \Rightarrow y = \boxed{y = \frac{26}{139}x + \frac{202}{139}}$

La droite de régression de x en y a pour équation :

$$x - \bar{x} = \frac{\mathrm{cov}(x;y)}{V(y)}\left(y - \bar{y}\right) \Rightarrow x - \frac{1}{4} = \frac{\frac{13}{8}}{\frac{13}{4}}\left(y - \frac{3}{2}\right) = \frac{1}{2}y - \frac{3}{4} \Rightarrow \frac{1}{2}y = x - \frac{1}{4} + \frac{3}{4} \Rightarrow \boxed{y = 2x + 1}.$$

Le coefficient de corrélation est $r = \frac{\mathrm{Cov}(x\ ;\ y)}{\sqrt{V(x)}.\sqrt{V(y)}} = \frac{\frac{13}{8}}{\sqrt{\frac{139}{16}}.\sqrt{\frac{13}{4}}} \simeq 0{,}306$.

B. Exercices

1. 1) On considère la série statistique présentée dans le tableau suivant :

Classe	[1 ; 2[	[2 ; 4[	[4 ; 5[	[5 ; 6[	[6 ; 9[
Effec-	25	15	16	12	32

a) Construire un histogramme représentant cette série.
b) Dresser le tableau des effectifs cumulés croissants et décroissants, puis construire les polygones des effectifs cumulés croissants et décroissants.
c) Dresser le tableau des fréquences cumulées croissantes et décroissantes.
d) Déterminer la classe modale de cette série.
c) Calculer la moyenne, la variance et l'écart type de cette série.

2. Le tableau suivant donne les notes sur 100 obtenues par 120 étudiants d'une université à leur examen final.

Classe	[30;40]	[40;50]	[50;60]	[60;70]	[70;80]	[80;90]	[90;100]
Effectif	1	3	11	21	43	32	9

a) Dresser le tableau des fréquences cumulées croissantes et décroissantes.
Construire les polygones des fréquences cumulées croissantes et décroissantes.
b) Calculer la médiane, le premier quartile et le troisième quartile de cette série.

3. On considère une série dont les effectifs cumulés décroissants sont donnés dans le tableau suivant :

Classe	[0 ;4 [	[4 ;6 [	[6 ;8 [	[8 ;10[	
Eff.cu. dec	60	45	24	6	

a) Construire le polygone des effectifs cumulés décroissants.
b) Déterminer graphiquement, la médiane de cette série, vérifier le résultat par le calcul.
c) Calculer la moyenne de cette série.

4.

vie en heures, regroupées en huit classes, d'un lot de piles d'une même marque pour lampes à poche.

D.v	[50 ;60[	[60 ;70[	[70 ;80[	[80 ;90[	[90 ;100[	[100 ;110[	[110;120[	[120;130 [
Ef	10	15	35	40	30	20	15	5

a) Calculer la moyenne, la variance et l'écart type de cette série.
b) Calculer la médiane, le premier quartile et le troisième quartile de cette série.

6. La série suivante donne la répartition des tailles en cm de 36 élèves d'une classe de 6ème année.

Tai.	[145 ;155 [	[155 ;160 [	[160 ; 165[	[165 ;170 [	170 ;175 [	[175 ;18 [
Effe.	3	5	6	8	8	6

a) Représenter cette série par l(histogramme des effectifs.
b) Déterminer la (ou les) classe(s)modale (s) de cette série et calculer sa moyenne.
c) Construire le polygone des effectifs cumulés croissants. Déterminer la médiane de cette série.

7. Une étude statistique portant sur un caractère x a permis de construire l'histogramme ci-dessous.
a) Quelle est la (ou les) classe(s) modale(s) de la série statistique correspondante ?
b) Construire le polygone des fréquences cumulées croissantes et décroissantes, en déduire la médiane.
c) Calculer la moyenne et l'écart type de cette série.
c) Sachant, que l'effectif de la population est 250, dresse le tableau des effectifs.

8. On dispose de la statistique suivante relative aux exploitations agricoles d'une région, classées d'après leur superficie en hectares.

Sup. en H	[0 ; 1[	[1 ; 2[	[2 ; 3[	[3 ; 51[	[5 ; 10	[10 ;20[	[20 ; 40[
effectif	**26**	**35**	**29**	**53**	**91**	**71**	**25**

Représenter l'histogramme relatif à cette distribution.
Dans un repère orthogonal (O ; I ; J) , représenter les

On considère une série dont les fréquences sont données dans le tableau suivant :

Classe	[0 ; 3[	[3 ; 5 [	[5 ; 7 [	[7 ;10[	
Eff.cu.. dec	10%	25%	35%	30%	

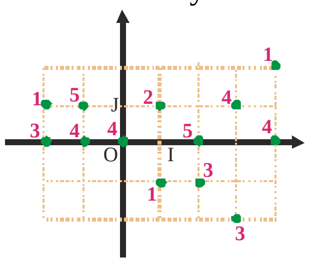
a) Calculer la moyenne, la variance et l'écart type de cette série.

5 Le tableau suivant donne les durées de

x_i \ y_i	0	1	2	4	5
0	0	3	0	0	0
1	0	4	2	2	0
2	3	5	4	4	2
3	1	4	0	1	0
5	0	2	2	0	1

a) Représenter par des points pondérés le nuage de points associé à cette série.
b) Déterminer et placer le point moyen du nuage.

10 Dans le plan muni d'un repère (O ; I ; J), on a représenté par des points pondérés le nuage associé à une série statistique à deux caractères x et y.



a) Dresser les tableaux des effectifs des deux séries marginales.
b) Déterminer et placer le point moyen du nuage.

11 Soit la série statistique double :

x	3	5	6	8	9	11
y	2	3	4	6	5	8

Représenter le nuage de points du tableau.
Calculer les coordonnées du point moyen G.
Déterminer une équation de la droite de régression de y en x et une équation de la droite de régression de x en y.

polygones cumulatifs croissants et décroissants.
Déterminer les coordonnées du point d'intersection de ces deux polygones.
Déterminer le pourcentage des exploitations agricoles de superficies : inférieur à 5 hectares ; supérieure à 10 hectares ; comprises entre 2 et 20 hectares.

9. On a relevé le nombre x de sœurs et le nombre y de frères de chacun des 40 élèves d'une classe.
On a obtenu la série double suivante :

12 Le tableau suivant donne le poids y en kg d'un nourrisson, x jours après sa naissance.

x_i	5	7	10	14	18	22	26
y_i	3,61	3,70	3,75	3,85	3,90	4,05	4,1

a) Représenter le nuage de points associés à cette à cette série.
b) Déterminer une équation de la droite de régression de y en x et représenter cette droite sur le graphique.
c) Donner une estimation du poids du nourrisson 30 jours après sa naissance.

13 Le tableau suivant donne la tension artérielle moyenne y en fonction de l'âge x d'une population

Age (x_i)	36	42	48	54	60	66
Tension (y_i)	11,8	14	12,6	15	15,5	15,1

a) Déterminer une équation pour chacune des deux droites de régression.
b) Calculer le coefficient de corrélation linéaire.

14 On a étudié trois caractères x ; y et z sur une population de cinq individus et on a obtenu les résultats suivants :

x_i	1	2	3	4	5
yi	40	45	35	35	30
zi	0	-3	1	4	-1

a) Représenter sur deux graphiques différents les nuages de points associés respectivement aux séries
(x_i ; y_i) et
(x_i ; z_i).
b) Déterminer une équation de la droite de régression de y en x et une équation de la droite de régression de z en x.
c) Calculer les coefficients de corrélation linéaire des séries (x_i ; y_i) et (x_i ; z_i).