EXTRAIT DE LA SECONDE EDITION DU LIVRE DE STATISTIQUE DESCRIPTIVE
COURS ET EXERCICES CORRIGES

EXERCICE 2

On a relevé lôâge des personnes présentes à une réunion:

25	29	41	40	26	25
33	28	35	40	27	28
35	29	33	26	35	32
40	21	33	33	25	28
28	35	32	33	32	25

Faire lôétude statistique c'est-à-dire
a) Expliquer quel est le caractère étudié. Est-il qualitatif ou quantitatif?
Est ï il discret ou continu?
b) transformer les données brutes en un tableau donnant les effectifs et les fréquences correspondant à chaque valeur prise par le caractère.

EXERCICE 2

a) Le caractère étudié est lôâge des personnes présentes à une réunion. Il est quantitatif continu.
b) Si on nôexige pas de conditions à suivre pour les classes dôamplitude, le choix reste subjectif.

lôâge des personnes	moins de 25	de 25 à moins de 30	de 30 à moins de 35	de 35 à moins de 40	de 40 au plus	Total
effectif $□_□$	1	13	8	4	4	30
fréquence $□_□$	$\frac{1}{30}$	$\frac{13}{30}$	$\frac{8}{30}$	$\frac{4}{30}$	$\frac{4}{30}$	□□ □ □ $□_□$ □ 1 □□ □

EXERCICE5

Le tableau suivant fournit les prix ($□_□$) et les quantités ($□_□$) de trois biens $□_□$ (avec i = 1, 2,3) en 2000, les indices élémentaires des prix ($□_□$) et des quantités ($□_□$) de ces trois biens en 2007 (base 100 en 2007).

Bien	$□_□$ 2000	$□_□$ 2000	$□_□$ 2007/2000	$□_□$ 2007/2000
$□_□$	4	5	200	25
$□_□$	10	4	120	125
$□_□$	8	5	75	100

1) Calculer les coefficients budgétaires $a_□^□$ (relatifs à la période de base).

…er la valeur de l'indice de Laspeyres des prix de 2007 base

…relatifs à la période courante.

EXERCICE 5

1. Les coefficients budgétaires sont les rapports des dépenses relatives à un bien ($P_i \times Q_i$) sur la période considérée, à la dépense totale effectuée au cours de la même période.
D'après le cours on a défini les coefficients budgétaires par :

$$a_i^t = \frac{P_{it} \times Q_{it}}{\sum_i P_{it} \times Q_{it}}$$

Soit

$$a_1^t = \frac{4 \times 5}{4 \times 5 + 10 \times 4 + 8 \times 5} = 0{,}2$$

$$a_2^t = \frac{10 \times 4}{4 \times 5 + 10 \times 4 + 8 \times 5} = 0{,}4$$

$$a_3^t = \frac{8 \times 5}{4 \times 5 + 10 \times 4 + 8 \times 5} = 0{,}4$$

EXERCICE 3

Dans un atelier, 25 ouvriers fabriquent des pièces d'un même type. Certaines de ces pièces sont mises au rebut pour défectuosité. La série suivante donne, pour chaque ouvrier, le nombre de pièces mises au rebut pour 200 pièces retenues:
3,11,8,7,4,6,8,4,5,17,15,5,12,7,10,9,3,7,13,16,11,9,8,7,9.

1. Quel est le caractère étudié? Est-il qualitatif ou quantitatif? continu ou discontinu?
2. Ranger ces observations par valeur non décroissante. Quelle est l'amplitude de cette série?
3. Classer les valeurs ci-dessus de 5 en 5 à partir de 0 (moins de 5, de 5 à 10 exclu,…).

Quel est l'effectif de chaque classe? Quelle est la fréquence de chaque classe?

EXERCICE 4

Une enquête en vue de la réduction du montant des allocations familiales, a été réalisée auprès d'une population de femmes de 40 ans. Cette enquête a donné les résultats suivants:

nombre d'enfants (x_i)	nombre de femmes (n_i)
0	10
1	20
2	20
3	30
4	20

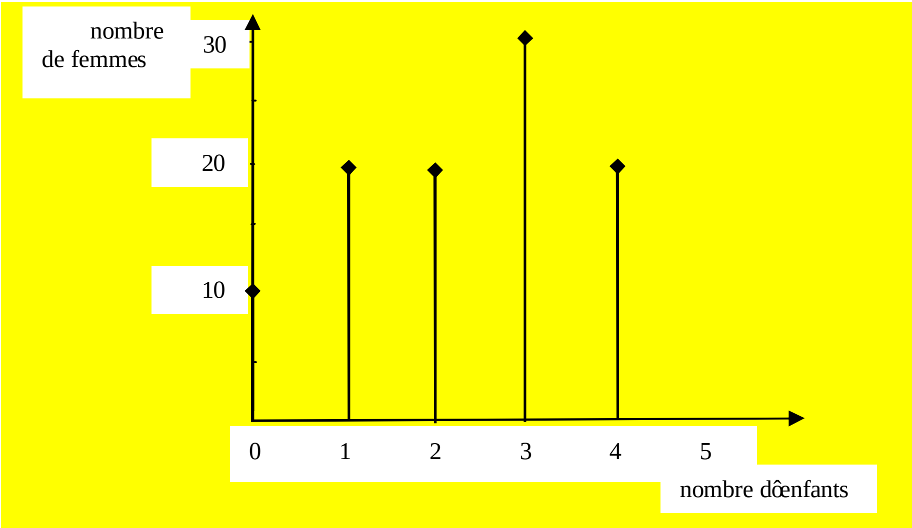
1) Caractériser la distribution.
2) Tracer le diagramme différentiel.
3) Définir et représenter la fonction de répartition (diagramme cumulé).
4) Donner la proportion des femmes ayant moins de 4 enfants.

EXERCICE 4

1.Population: 100 femmes de 40 ans; unité statistique: une femme; caractère: le nombre d'enfants (caractère quantitatif discret); modalités: au nombre de 5 (0, 1, 2, 3, 4).

2. La distribution statistique étant discrète, le diagramme différentiel est un diagramme en bâtons. Dans ce diagramme, on porte en abscisse les différentes modalités du caractère c'est à dire les différentes valeurs prises par la variable (0, 1, 2, 3, 4); en ordonnée seront indiqués soit les effectifs soit les fréquences relatives afférentes à chaque modalité.

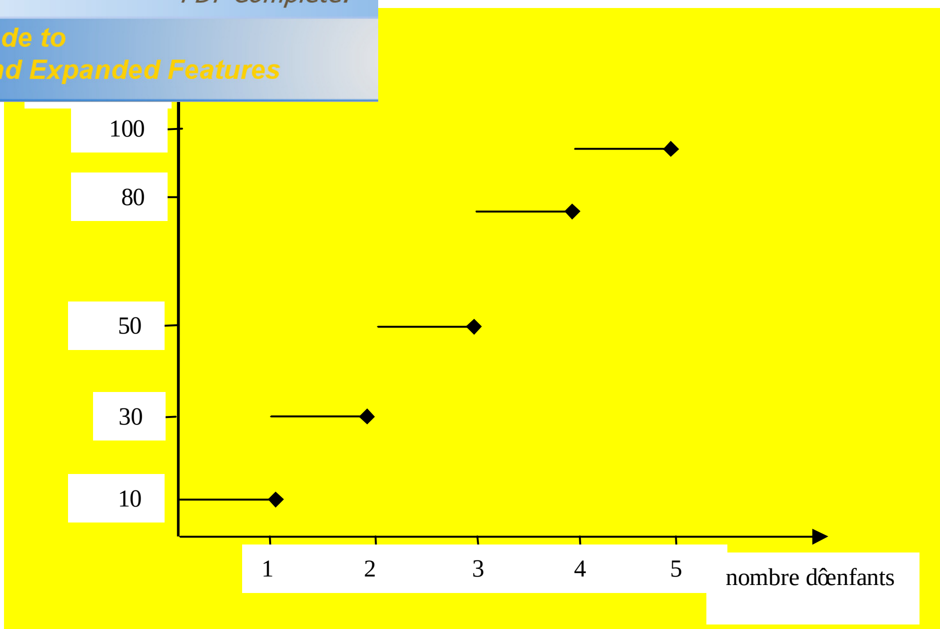
[illegible]	effectif ⍰	fréquence relative ⍰
[illegible]	10	0,1
1	20	0,2
2	20	0,2
3	30	0,3
4	20	0,2
Total	100	1



3. La fonction de répartition dõune variable X notée F est une application de lõensemble ⍰ des nombres réels dans lõensemble ⍰, qui à toute valeur donnée *x* de ⍰, associe le nombre dõindividus appartenant à la population pour lesquels la valeur de la variable est strictement inférieure à ⍰.

En termes de proportion, la fonction de répartition est une application de lõensemble ⍰ dans lõintervalle [0,1], qui à toute valeur donnée ⍰ de ⍰, associe la proportion des individus appartenant à lõensemble statistique pour lesquels la valeur de la variable est strictement inférieure à ⍰.

nombre dõenfants ⍰	effectif cumulé ⍰	fréquence cumulée ⍰
0	10	0,1
1	30	0,3
2	50	0,5
3	80	0,8
4	100	1



4. La proportion des femmes ayant moins de 4 enfants se lit directement dans le tableau: 0,8 ou 80%.

EXERCICE 5

Une compagnie de location de voiture s'intéresse au kilométrage effectué par ses véhicules au cours d'une journée. Elle a relevé le kilométrage de 55 de ses véhicules, regroupés dans le tableau suivant:

Trajet en kilométrage	[0,10[	[10,20[	[20,30[	[30,40[	[40,60[	[60,90[	[90,130[	[130,170[
Nombre de véhicules	3	8	10	20	6	3	3	2

a) Dresser le tableau statistique où figurera les classes; les centres des classes et les fréquences.
b) Tracer l'histogramme et le diagramme cumulé.

Extrait du cours

3. Moyenne harmonique

Si un train fait un trajet aller-retour entre deux villes à la vitesse moyenne v_1 pour l'aller et à la vitesse moyenne v_2 au retour, la vitesse moyenne du trajet complet n'est pas la moyenne arithmétique des deux vitesses, mais bien sûr une autre moyenne appelée moyenne harmonique.

La moyenne harmonique H d'une série statistique est égale à l'inverse de la moyenne arithmétique des inverses des valeurs observées.

Si n est le nombre d'observations $x_1, x_2, \ldots x_p$, la moyenne harmonique est:

$$H = \frac{1}{\frac{1}{n}\sum_{i=1}^{i=p}\frac{1}{x_i}} \quad \text{ou} \quad H = \frac{n}{\sum_{i=1}^{i=p}\frac{1}{x_i}}.$$

avec $n = \sum_{i=1}^{i=p} n_i = n_1 + n_2 + \ldots n_p$

Si la série est pondérée, on écrit:

$$H = \frac{\sum_{i=1}^{i=p} n_i}{\sum_{i=1}^{i=p}\frac{n_i}{x_i}} = \frac{1}{n\sum_{i=1}^{i=p}\frac{n_i}{x_i}} = \frac{1}{\sum_{i=1}^{i=n}\frac{n_i}{n}\cdot\frac{1}{x_i}} = \frac{1}{\sum_{i=1}^{i=p} f_i\cdot\frac{1}{x_i}}.$$

Soit $H = \dfrac{1}{\sum_{i=1}^{i=p} \dfrac{f_i}{x_i}}$.

<u>Exemple</u>

Un escargot monte un puits de longueur trente mètres en une heure, et en descendant sa vitesse double. Quelle est la vitesse moyenne de cet escargot?

En effet une approche intuitive conduit à dire que sa vitesse moyenne est $v_{moy} = \dfrac{30+60}{2} = 45$ mètres à lđheure .Ce qui est manifestement faux. En effet une vitesse est une grandeur qui dépend de deux paramètres, la distance et le temps et dans notre exemple, la distance du puits étant constante, et ce sont les temps qui varient. Soit d cette distance, v_1 et v_2 les deux vitesses, t_1 et t_2 les temps de parcours respectifs à la montée et la décente.

En montant $v_1 = \dfrac{d}{t_1}$ et en descendant $v_2 = \dfrac{d}{t_2}$, donc,

$$v_{moy} = \frac{d+d}{t_1+t_2} = \frac{2d}{\frac{d}{v_1}+\frac{d}{v_2}} = \frac{2}{\frac{1}{v_1}+\frac{1}{v_2}}.$$

Nous remarquons donc que la vitesse moyenne est ici la moyenne harmonique des vitesses.

Soit $v_{moy} = \dfrac{2}{\frac{1}{30}+\frac{1}{60}} = 40$ mètres/heure.

Lûusage de la moyenne harmonique est peu fréquent, cependant elle est utilisée dans les calculs des indices.

stique est la racine carrée de la moyenne arithmétique des carrées des valeurs observées $x_1, x_2, ... x_p$.

On la note:

$$Q = \sqrt{\frac{1}{n}\sum_{i=1}^{i=p} x_i^2}\ .$$

Si la série est pondérée, côest-à-dire affectée par des cî fficients de pondération, on écrit alors:

$$Q = \sqrt{\frac{\sum_{i=1}^{i=p} n_i x_i^2}{\sum_{i=1}^{p} n_i}} = \sqrt{\sum_{i=1}^{p} \frac{n_i}{n} \cdot x_i^2} = \sqrt{\sum_{i=1}^{p} f_i x_i^2}\ ,$$

On note aussi

$$Q = \left[\frac{\sum_{i=1}^{p} n_i x_i^2}{\sum_{i=1}^{p} n_i}\right]^{\frac{1}{2}} = \left[\sum_{i=1}^{p} f_i x_i^2\right]^{\frac{1}{2}}.$$

<u>Exemple 1</u>

Si un rectangle a pour côtés 3 et 7, le carré qui a même diagonale que le rectangle a pour côté la moyenne quadratique de 3 et 7, c'est-à-dire 5.38.

5.2. Mode des séries à caractère *continu*

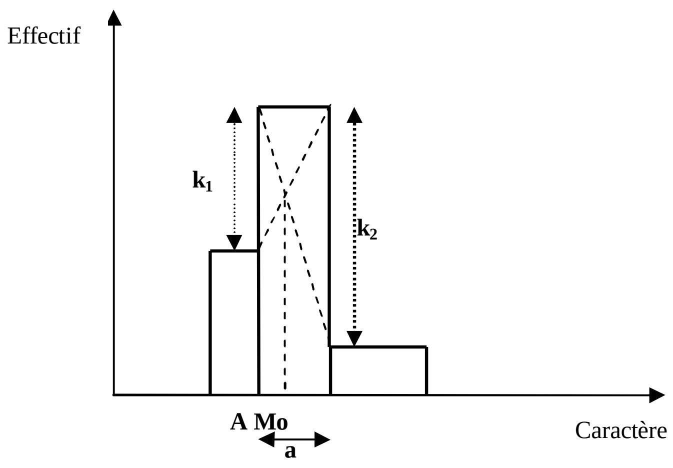
La détermination du mode est plus délicate pour les séries à caractère continu.

On perçoit immédiatement la classe correspondant à la fréquence la plus grande ou à lôeffectif le plus important. Cette classe porte le nom de classe modale; elle contient le mode.

Pour déterminer le mode, plusieurs procédés sont possibles.

-Si on ne sôintéresse pas à une valeur plus précise, il est possible de considérer que le mode de la série correspond au milieu de la classe modale.

-Si on désire plus de précision, le procédé le plus simple est un procédé graphique. Il consiste à estimer que le mode est déporté à lôintérieur de la classe modale en fonction des effectifs rectifiés des classes encadrant la classe modale. Le mode est alors obtenu de la façon indiquée sur la graphique ci-dessous.



Série continue
Détermination graphique du mode

…lculer le mode par la formule suivante, qui correspond à la …phique:

$$Mo = A + \frac{k_1}{k_1 + k_2} \times a.$$

A est la valeur du caractère correspondant à la valeur inférieure de classe modale; **a** est l'amplitude de la classe modale; k_1 et k_2 sont les différences entre les effectifs (ou les fréquences) rectifiés de la classe modale et des deux classes qui l'encadrent.

Exemple

Une étude portant sur la durée de vie d'une espèce biologique aquatique de la même race est consignée dans le tableau ci-dessous:

durée de vie en heures	nombre d'espèces en milliers
[0; 2 000[	8
[2 000; 4 000[	26
[4 000; 5 000[	20
[5 000; 6 000[	22
[6 000; 8 000[	18
[8 000; 10 000[	6

Déterminons la classe modale, et plus précisément le mode.
Les classes ne sont pas d'amplitudes égales; la classe modale n'est donc pas nécessairement celle où l'effectif est le plus élevé. En effet, la classe
[2 000; 4 000[dont l'effectif est 26 a une amplitude double de la classe
[5 000; 6 000[dont l'effectif est 22;
La classe dont la densité est maximum, est la classe modale. La densité (on dit aussi hauteur) est calculée à l'aide des effectifs et de l'amplitude. On a:

$d_i = \frac{n_i}{a_i}$, on obtient alors le tableau suivant:

durée de vie en heures	nombre d'espèces n_i	amplitude a_i	densité (ou hauteur) d_i
[0 ; 2 000[	8 000	2 000	4
[2 000 ; 4 000[	26 000	2 000	13
[4 000 ; 5 000[	20 000	1 000	20
[5 000 ; 6 000[	22 000	1 000	22
[6 000 ; 8 000[	18 000	2 000	9
[8 000 ; 10 000[	6 000	2 000	3

La classe modale est donc [5 000,6 000[. En appliquant la méthode proposée ci-dessus, on aura:

$$Mo = A + \frac{k_1}{k_1 + k_2} \times a,$$

$k_1 = 22 - 20 = 2$, et $k_2 = 22 - 9 = 13$.

2. Courbe de concentration- coefficient de concentration

2.1. Courbe de concentration

Nous illustrons notre propos par un exemple:
Soit la distribution suivante portant sur un ensemble de parcelles de terrains agricoles.

superficies des lots de terrains (en hectares)	nombre de lots
[0 ; 10[	8
[10 ; 20[	15
[20 ; 40[	9
[40 ; 70[	5
[70 ; 100[	3

Chapitre 7 Ajustements

1. Méthode dôajustement

Le nuage de point que nous avons sur les graphiques nous amène et dôune manière légitime à tracer à la main levée une courbe simple, régulière, telle que se compensent à peu près les écarts, positifs ou négatifs, pour une même abscisse, entre les ordonnées ajustées et les ordonnées vraies.

Ainsi par exemple dans le cas des deux graphiques ci-dessus un ajustement manuel peut être de la manière suivante.