L1 H. Perrot-J. Gamby

TD1 de Statistiques

Exercice 1 : chiffres d'affaires de supermarchés ("statnet")

On dispose pour tous les supermarchés d'une chaîne française de leurs chiffres d'affaires en milliers d'euros aux mois de décembre 2001 et décembre 2002 :

	Lyon	Paris	Marseille	Nice	Lille	Strasbourg	…
Déc. 2001	320 000	410 000	250 000	150 000	210 000	180 000	
Déc. 2002	380 000	380 000	280 000	200 000	210 000	170 000	

1) Quelle est la population étudiée : la clientèle, les dates, les supermarchés, les chiffres d'affaires ?
2) De quelle(s) variable(s) dispose-t-on : le chiffre d'affaires en décembre 2001, le chiffre d'affaires en décembre 2002, la localisation, la taille du supermarché ?
3) La variable "chiffre d'affaires en décembre 2001" est-elle quantitative ? discrète ? continue ? nominale ? ordinale ?
4) On compare les chiffres d'affaires aux deux dates disponibles, et on définit la variable "évolution" qui vaut "SUP", "STABLE", "INF" selon que le chiffre a augmenté, est resté identique ou a diminué entre les deux dates. Cette variable est-elle quantitative ? qualitative ? nominale ou ordinale ? En est-il de même si on note -1, 0, 1 selon l'évolution du chiffre ?

Exercice 2 : organisation du travail (B. Grais, "Statistique Descriptive", Dunod)

En vue d'établir rationnellement le nombre de postes de travail nécessaires pour assurer à sa clientèle un service satisfaisant, une agence de voyage a fait relever, minute par minute, le nombre d'appels téléphoniques reçus au cours d'une période de 30 jours. Cette opération a fourni, pour la tranche horaire de pointe qui se situe entre onze heures et midi, les résultats suivants :

nombre d'appels par minute	0	1	2	3	4	5	6	7	8 et +	total
nombre de minutes	93	261	416	393	308	174	93	42	20	1800

1) Déterminer la population et le caractère étudiés. Ce caractère est-il qualitatif ou quantitatif ? Existe-t-il une variable statistique ? Si oui, est-elle discrète ou continue ?
2) Représenter graphiquement la distribution des effectifs, le polygone des effectifs et la distribution des effectifs cumulés.
3) Représenter graphiquement la distribution des fréquences relatives et celle des fréquences relatives cumulées.

Exercice 3 : une affaire de tailles (M.R. Spiegel, "Théorie et Applic. de la Statistique", série Schaum)

Le tableau suivant donne la taille de 40 élèves d'un lycée à un centimètre près.

138	164	150	132	144	125	149	157
146	158	140	147	136	148	152	144
168	126	138	176	163	119	154	165
146	173	142	147	135	153	140	135
161	145	135	142	150	156	145	128

1) Le caractère étudié est-il représenté par une variable statistique discrète ou continue ?
2) Etablir la distribution des effectifs. Pour cela, après avoir défini l'étendue de la série statistique, on présentera les données sous forme d'un tableau faisant apparaître des classes d'étendue (amplitude) 9 cm, la borne inférieure de la première classe étant 118 cm.
3) Construire l'histogramme, le polygone des effectifs (fréquences) et le polygone des effectifs cumulés.
4) Comment en déduire l'histogramme des fréquences relatives et le polygone des fréquences relatives cumulées ?

Exercice 4 : une histoire de lampes (C. Labrousse, "Statistique", Dunod)

Le tableau suivant donne la distribution des fréquences de la durée de vie de 500 lampes.

durée de vie en heures	300 à 499	500 à 699	700 à 1099	1100 et plus
nombre de lampes	50	150	274	26

1) Tracer l'histogramme des fréquences relatives et le polygone des fréquences relatives.
2) Tracer le polygone des fréquences relatives cumulées.
3) Déterminer le nombre de lampes ayant une durée de vie inférieure à 900 heures.

TD2 de Statistiques

Exercice 1 : une série de notes

On considère la série statistique formée des notes suivantes : 12, 11, 5, 8, 15, 7, 10, 18, 9, 10, 12, 6, 17, 13, 15 et 8. Calculer la moyenne arithmétique, la médiane, les quartiles et le mode de cette série.

Exercice 2 : moyenne stable

On désire calculer la moyenne arithmétique d'une série statistique dont les valeurs x_i apparaissent progressivement au cours du temps. Soit m_n la moyenne obtenue à partir des n premières valeurs de x_i. Calculer la valeur de m_{n+1} à partir de la valeur de m_n et la $(n+1)^{ème}$ valeur x_{n+1} de la série.

Exercice 3 : les stars d'Hollywood (J.-P. Lecoutre, "Stat. Descriptive, exercices corrigés", Masson)

Une étude portant sur 60 stars d'Hollywood a permis de construire le tableau donnant la répartition en fonction de leur nombre de divorces :

nombre de divorces x_i	0	1	2	3	4	5	6	7 et +	total
effectif n_i	1	3	7	15	19	10	4	1	60

1) Représenter graphiquement cette distribution et calculer les effectifs cumulés.

2) Déterminer le mode, la médiane et les quartiles Q_1 et Q_3.

3) Calculer la moyenne arithmétique $\bar{x}$.

Exercice 4 : organisation du travail (B. Grais, "Statistique Descriptive", Dunod)

Calculer la moyenne arithmétique, le mode, la médiane et les quartiles de la série à caractère discret étudiée dans l'exercice 2 du TD1.

Exercice 5 : une histoire de lampes (C. Labrousse, "Statistique", Dunod)

Calculer la moyenne arithmétique, la médiane, les quartiles et le mode de la série à caractère continu et classes inégales étudiée dans l'exercice 4 du TD1.

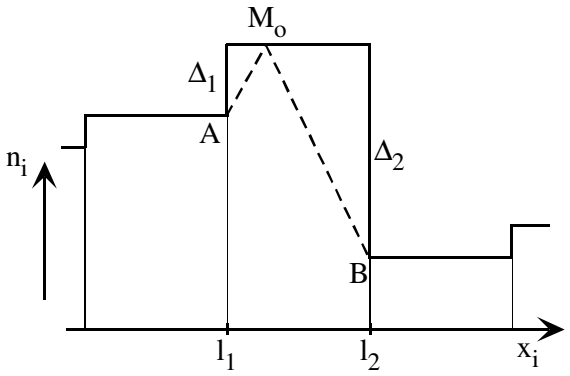
TSVP

Exercice 6 : propriétés du mode (série à caractère continu)

1) Démontrer que le mode d'une série classée, défini avec les notations du cours par la relation

$$M_o = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2}(l_2 - l_1)$$

correspond au point pour lequel les pentes des droites AM_o et M_oB sur l'histogramme sont égales en valeur absolue.



2) Montrer que si l'on construit une parabole passant par les milieux des sommets des trois rectangles de la figure, l'abscisse du point maximal de cette parabole est le mode donné par la relation ci-dessus.

Exercice 7 : une affaire de tailles (M.R. Spiegel, "Théorie et Applic. de la Statistique", série Schaum)

Calculer la moyenne arithmétique, le mode, la médiane et les quartiles de la série à caractère continu et classes égales étudiée dans l'exercice 3 du TD1.

TD3 de Statistiques

Exercice 1 : une série de notes

On considère la série statistique formée des notes suivantes : 12, 11, 5, 8, 15, 7, 10, 18, 9, 10, 12, 6, 17, 13, 15 et 8 dont on a calculé la moyenne arithmétique et les quartiles dans l'exercice 1 du TD2. Calculer l'écart moyen, la variance, l'écart-type, l'écart interquartile et le coefficient de variation de cette série.

Exercice 2 : les stars d'Hollywood (J.-P. Lecoutre, "Stat. Descriptive, exercices corrigés", Masson)

Calculer la variance, l'écart-type et le coefficient de variation de la série à caractère discret dont on a calculé la moyenne arithmétique dans l'exercice 3 du TD2.

Exercice 3 : un grand classique

Démontrer que la moyenne $f(a)$ des carrés des écarts par rapport à une valeur a quelconque : $f(a)=\frac{1}{n}\sum_{i=1}^{p} n_i(x_i-a)^2$ est minimale pour $a = \bar{x}$. Que vaut alors la valeur minimale de $f(a)$?

Exercice 4 : organisation du travail (B. Grais, "Statistique Descriptive", Dunod)

Calculer la variance, l'écart-type et le coefficient de variation de la série à caractère discret dont on a calculé la moyenne arithmétique dans l'exercice 4 du TD2.

Exercice 5 : une histoire de lampes (C. Labrousse, "Statistique", Dunod)

Calculer la variance, l'écart-type et le coefficient de variation de la série à caractère continu et classes inégales dont on a calculé la moyenne arithmétique dans l'exercice 5 du TD2.

Exercice 6 : une affaire de tailles (M.R. Spiegel, "Théorie et Applic. de la Statistique", série Schaum)

Calculer l'écart-type et le coefficient de variation de la série à caractère continu et classes égales dont on a calculé la moyenne arithmétique dans l'exercice 7 du TD2.

Exercice 7 : calcul stable de l'écart-type

On désire calculer l'écart-type d'une série statistique dont les valeurs x_i apparaissent progressivement au cours du temps. Soit σ_n l'écart-type obtenu à partir des n premières valeurs de x_i. Calculer la valeur de σ_{n+1} à partir de la valeur de σ_n, de la moyenne m_n des n premières valeurs de x_i, et de la $(n+1)^{\text{ème}}$ valeur x_{n+1} de la série.

Exercice 1 : ajustement linéaire

Les ventes annuelles d'un magasin exprimées en milliers d'euros sur les 6 dernières années sont données dans le tableau suivant (l'année 6 correspond à l'année 2004) :

année x_i	1	2	3	4	5	6
ventes y_i	22	33	39	47	60	79

1) Représenter graphiquement (échelles linéaires) l'évolution des ventes.

2) Construire manuellement une droite qui ajuste les données du tableau. Déterminer l'équation de cette droite.

3) Redémontrer les 2 équations qui permettent de déterminer les paramètres a et b de la droite des moindres carrés $y = a\,x + b$. Déterminer a et b pour ces données.

4) Calculer le coefficient de corrélation R.

5) En admettant que l'évolution constatée se poursuit de façon identique, estimer le montant des ventes au cours de l'année 2005.

Exercice 2 : ajustement par une fonction exponentielle $y = b\,a^x$

Les résultats d'une observation statistique entre 2 quantités x et y ont donné les valeurs suivantes :

x_i	2	5	7	11	15
y_i	20	300	4000	10^5	10^6

1) Représenter graphiquement (échelles semi-log) l'évolution des ventes.

2) Construire manuellement une droite qui ajuste les données du tableau avec ces échelles. Déterminer l'équation de cette droite et en déduire l'expression de y en fonction de x.

3) Déterminer l'équation de la droite des moindres carrés pour ces données. Calculer le coefficient de corrélation R.

Exercice 3 : ajustement par une fonction puissance $P = b\,V^a$ (M.R. Spiegel, "Théorie et Applications de la Statistique", série Schaum)

Le tableau suivant donne les valeurs expérimentales de la pression P d'une masse donnée de gaz pour différentes valeurs du volume V. D'après les principes de la thermodynamique, on a la relation $PV^\gamma = C$ où γ et C sont des constantes dépendant des conditions d'expérience.

Volume V_i en cm^3	54,3	61,8	72,4	88,7	118,6	194,0
Pression P_i en kg/cm^2	61,2	49,5	37,6	28,4	19,2	10,1

1) Représenter graphiquement (échelles log-log) les données du tableau et déterminer l'équation d'une droite ajustant ces données.

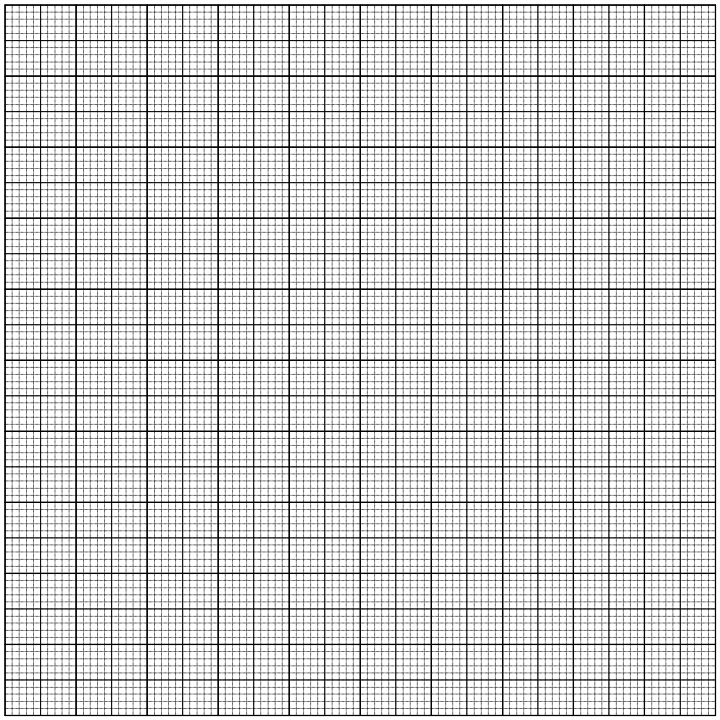
2) Déterminer l'équation de la droite des moindres carrés pour ces données. En déduire γ et C.

3) Quelle est la valeur du coefficient de corrélation R ?

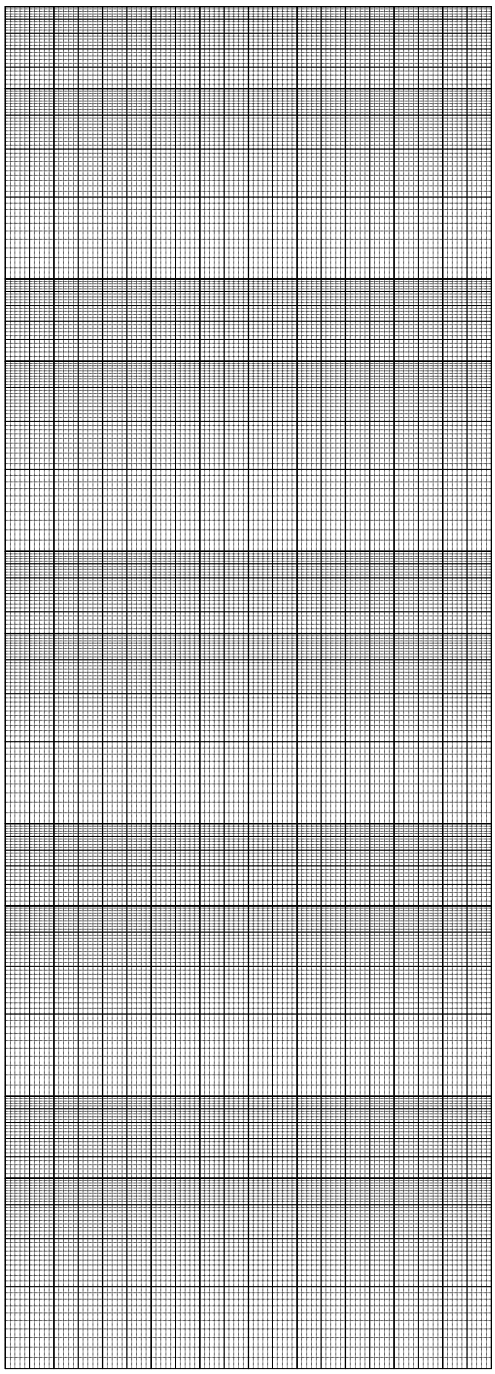
L1-L'2

TD4 de Statistiques (suite)

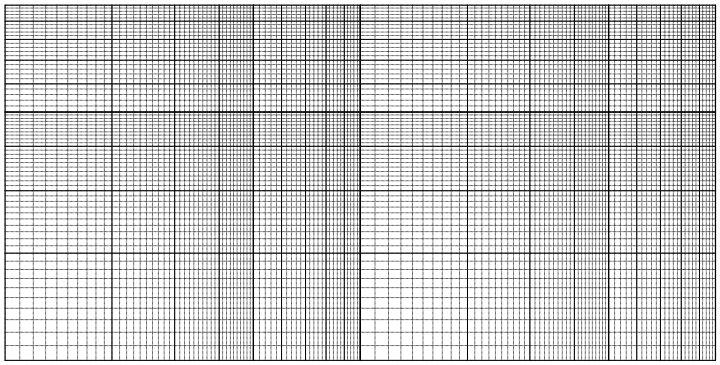
Exercice 1 :



Exercice 2 :



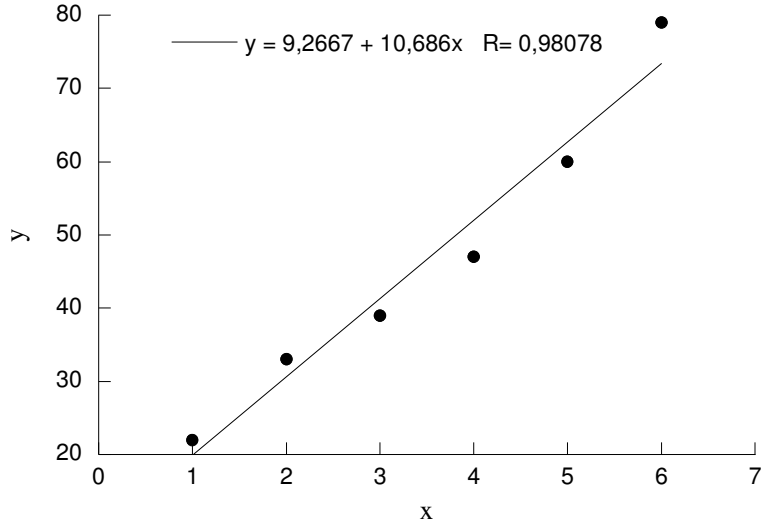
Exercice 3 :



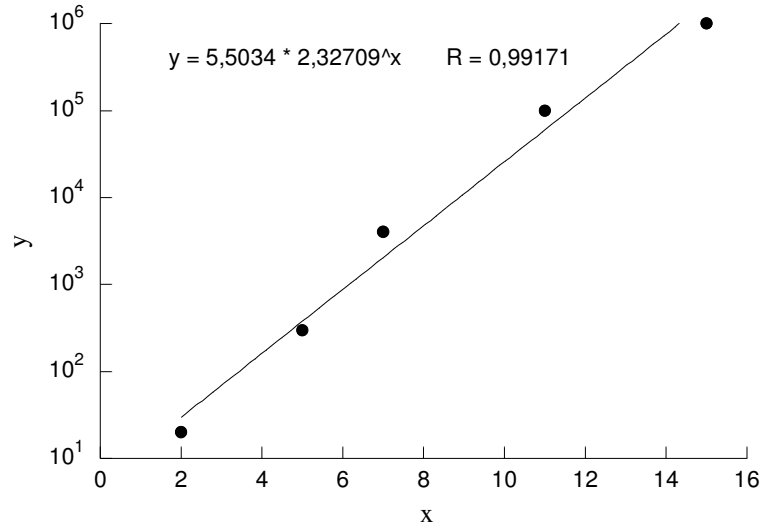
L1-L'2

Corrigé du TD4 de Statistiques

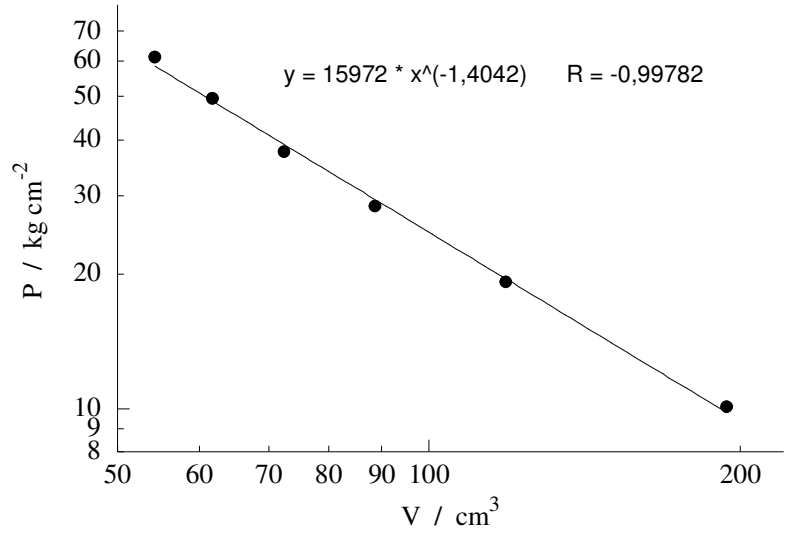
Exercice 1 :



Exercice 2 :



Exercice 3 :



Exercice 1 : changement de la période de référence

L'indice des prix à la consommation (IPC) pour l'ensemble des ménages (métropole + DOM), base 100 en janvier 1998, avait la valeur 84,9 en janvier 1990 et 110,7 en janvier 2005. Quelle est la valeur de l'indice en janvier 2005, base 100 en 1990 ?

Exercice 2 : introduction d'un nouveau bien dans un indice (J.-P. Lecoutre, "Statistique Descriptive, exercices corrigés", Masson)

Un indice des prix est calculé à partir de la consommation de trois biens, indiquée dans le tableau ci-après pour l'année de base 2003 (P = prix, Q = quantité) :

Bien	P	Q
B_1	3	60
B_2	1	8
B_3	4	20

L'indice des moyennes de Laspeyres des prix a eu pour valeur 104 en 2004. On décide alors d'introduire un quatrième bien (B_4), dont la consommation est 30 en 2003 et 26 en 2004, et dont le prix est 2 en 2003 et 2,5 en 2004, pour calculer un nouvel indice. Calculer la valeur de ce nouvel indice de Laspeyres en 2004.

Exercice 3 : réversibilité et transférabilité des indices (J.-P. Lecoutre, "Statistique Descriptive, exercices corrigés", Masson)

Pour comparer le coût de la vie dans certaines villes européennes, on se fonde sur le budget type d'un individu. Les prix (P) et quantités (Q) consommées de certains biens dans ces villes sont indiqués dans le tableau ci-après :

	Paris		Oslo		Rome	
	Q	P	Q	P	Q	P
logement	2	4	3	2	3	3
nourriture	4	2	5	2	3	4
chauffage	2	1	9	1	1	2

1) Calculer les indices des moyennes de Laspeyres et de Paasche à Paris en prenant Oslo pour base, puis à Oslo en prenant Paris pour base. Montrer que ces indices ne sont pas réversibles.

2) Calculer l'indice des moyennes de Laspeyres à Oslo en prenant Rome pour base, puis à Rome en prenant Paris pour base. Montrer que l'indice de Laspeyres n'est pas transférable.

Exercice 4 : indice de Fischer

1) Montrer que l'indice des moyennes de Fischer $F = \sqrt{LP}$ est réversible

2) Montrer que l'indice (base 1) des moyennes de la valeur $V = P\,Q$ est égal au produit des indices (base 1) de Fischer du prix et de la quantité.

Exercice 1 : distributions marginales et conditionnelles et leurs caractéristiques (CNAM, Paris, juin 1980, adaptation F. Huet)

Dans le but de voir s'il y avait une relation entre les résultats obtenus en 1ère année et en 2ème année de DEUG, le responsable d'une UFR de droit a tiré au hasard 250 dossiers d'étudiants admis à l'issue de la 1ère année et a observé leurs résultats en fin de 2ème année. Il a réalisé le tableau de contingence suivant :

résultats 2ème année (y) / résultats 1ère année (x)	1 ajourné	2 passable	3 assez bien	4 bien	5 très bien
1. passable	50	75	40	10	0
2. assez bien	10	20	20	0	0
3. bien	2	6	7	5	2
4. très bien	0	0	0	3	0

1) Comparer, à l'aide d'un même diagramme, la répartition des étudiants admis en fin de 1ère année et celle des étudiants en fin de 2ème année. Conclusion.

2) On désigne par x les notes obtenues par les étudiants de 1ère année et par y celles obtenues par les étudiants de 2ème année. Sachant qu'aucune note n'a été inférieure à 8 ou supérieure à 18 et que les mentions répondent à l'échelle des notes suivantes :

- ajourné : moins de 10
- passable : de 10 à moins de 12
- assez bien : de 12 à moins de 14
- bien : de 14 à moins de 16
- très bien : 16 et plus

calculer les distributions marginales des fréquences relatives $f_{i\bullet}$ et $f_{\bullet j}$ de x et y. Construire l'histogramme de la distribution marginale de y.

3) Calculer la distribution conditionnelle $f_{j|2}$ de y pour les étudiants ayant eu la mention assez bien en 1ère année. Montrer que : $f_{2j} = f_{j|2} \times f_{2\bullet}$.

4) Calculer les moyennes marginales $\bar{x}$ et $\bar{y}$ à partir des distributions marginales de x et y.

5) Calculer les moyennes conditionnelles $\bar{x}_j$ de x sachant que $y = y_j$ et retrouver la valeur de $\bar{x}$.

6) Calculer les moyennes conditionnelles $\bar{y}_i$ de y sachant que $x = x_i$ et retrouver la valeur de $\bar{y}$.

7) Tracer les courbes de régression $C_{y|x}$ et $C_{x|y}$ et montrer que $C_{x|y}$ a une pente plus forte que $C_{y|x}$.

8) Calculer l'écart-type de x, celui de y, puis la covariance et le coefficient de corrélation linéaire de x et y.

9) Montrer que la variance conditionnelle de x sachant y_1 s'écrit : $\sigma_1^2(x) = \sum_{i=1}^{4} (x_i - \bar{x}_1)^2 \, f_{i|1} = \overline{x_1^2} - \bar{x}_1^{\,2}$

et calculer la.

Série chronologique et moyenne mobile (CNAM, Paris, octobre 1977, adaptation F. Huet)

Soit la série chronologique y_t suivante donnant la consommation d'électricité en milliards de kWh, les chiffres étant arrondis au milliard de kWh le plus proche :

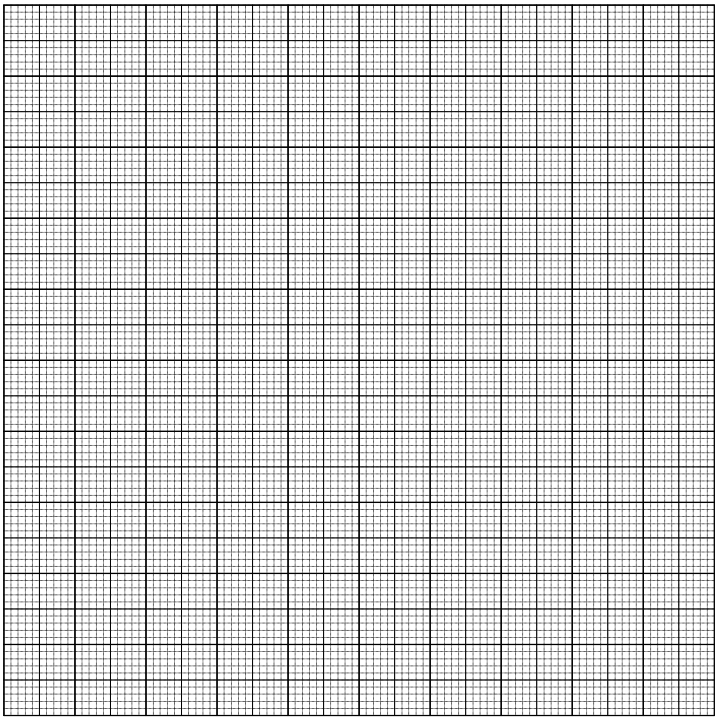
Année \ Trimestre	1	2	3	4
1974	-	14	13	16
1975	16	14	12	17
1976	18	15	14	18
1977	19	16	-	-

1) Représenter cette série sur papier millimétré. Quelles sont les composantes de la série ?

2) Vérifier que la moyenne mobile effectuée sur 4 points (la période des variations saisonnières est en effet le trimestre) donne les résultats présentés dans le tableau suivant.

Année \ Trimestre	1	2	3	4
1974	-	-	-	14,75
1975	14,63	14,63	15	15,38
1976	15,75	16,13	16,38	16,63
1977	-	-	-	-

3) Représenter la série des moyennes mobiles sur le graphique de la question 1.

4) Montrer que la droite d'ajustement de y_t par la méthode des moindres carrés a pour équation $y = 0{,}286\ t + 13{,}538$ et que le coefficient de corrélation linéaire a pour valeur approchée 0,53. Montrer que la droite d'ajustement des moyennes mobiles a une équation très proche ($y = 0{,}275\ t + 13{,}547$) et un coefficient de corrélation linéaire bien meilleur (environ 0,97).

5) Expliquer comment obtenir la série corrigée des variations saisonnières (CVS).