Corrélation et causalité

Exercice 1 (Nuages de points et point moyen)

On donne ci-après quatre séries statistiques doubles et quatre nuages de points.

Série A

x_i	1	2	3	4	5	6
y_i	2	1	-1	0,5	2	1

Série C

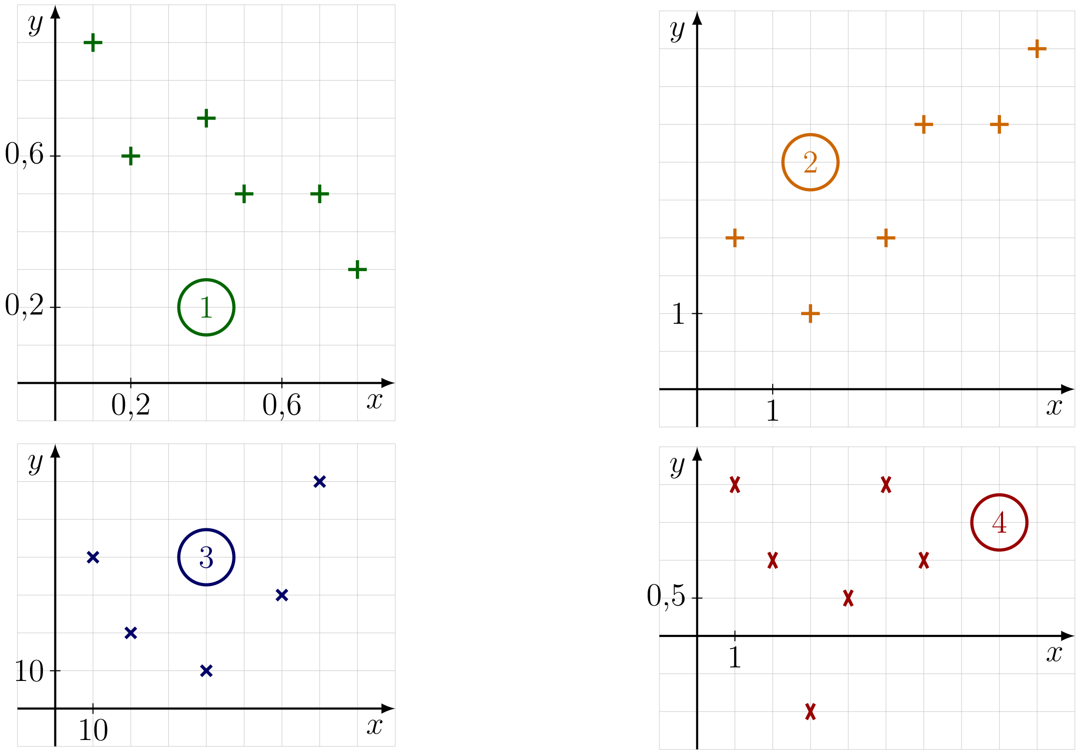
x_i	0,1	0,2	0,4	0,5	0,7	0,8
y_i	0,9	0,6	0,7	0,5	0,5	0,3

Série B

x_i	10	20	40	60	70
y_i	40	20	10	30	60

Série D

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



1. Associer chacune de ces séries et le nuage de points qui la représente.
2. Peut-on envisager un ajustement pour chacun de ces nuages de points ? Si oui, préciser de quel type.
3. Déterminer dans chacune des séries les coordonnées du point moyen et le placer sur le nuage correspondant.

Exercice 2 (Nuages de points et point moyen)

On considère les deux séries statistiques suivantes :

Série 1

x_i	0	1	2	3	4	5
y_i	52	96	145	213	345	567

Série 2

x_i	1970	1980	1990	2000	2010
y_i	0,1	1,2	2,3	4,5	6

1. En choisissant des unités adaptées dans chacun des cas, représenter le nuage de points dans un repère.
2. Dans chacun des cas, quel type d'ajustement paraît adapté pour ajuster le nuage ?
3. Déterminer dans chacune des séries les coordonnées du point moyen.

Exercice 3 (Nuages de points et point moyen)
On a noté, pour sic valeurs différentes du prix au kilogramme, le nombre de kilogrammes de cèpes vendus par un marchand. On obtient la série statistique suivante :

Prix x_i (en €/kg)	21,5	22	23,5	24	26	27,5
Quantité vendue y_i (en kg)	10	8,5	8	7	6,5	5

1. Représenter dans un repère le nuage de points.
2. Calculer les coordonnées du point moyen et le placer dans le nuage.
3. Interpréter les valeurs des coordonnées du point moyen dans le contexte de l'exercice.

Exercice 4 (Droite des moindres carrés)
Dans chaque cas, déterminer l'équation de la droite des moindres carrés de la série et le coefficient de corrélation, puis observer le nuage de points à l'aide de la calculatrice.

1.

x_i	1	5	8	10
y_i	3	6	10	15

2.

x_i	0	7	12	15	26
y_i	10	150	200	250	300

3.

x_i	50	120	150	210	250	300
y_i	57	40	37	30	25	21

Exercice 5 (Droite des moindres carrés)
On donne l'évolution d'une production y en fonction du nombre x d'années écoulées depuis 2015 :

x_i	0	1	2	3	4
y_i (en tonne)	2	2,3	2,9	3,5	3,8

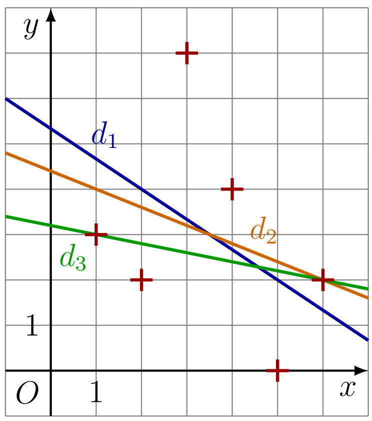
1. Déterminer l'équation de la droite de régression des moindres carrés.
2. L'ajustement précédent est-il adapté ?
3. En utilisant cet ajustement comme modèle :
   (a) Estimer la production en 2022 ;
   (b) Déterminer à partir de quelle année la production dépassera 7 tonnes.

Exercice 6 (Droite des moindres carrés)
Reprendre l'exercice 5 avec le tableau suivant :

x_i	0	1	2	3	4
y_i (en tonne)	1,5	2,1	2,9	3,4	3,9

Exercice 7 (Divers ajustements affines)
On considère la série statistique ci-dessous :

x_i	1	2	3	4	5	6
y_i	3	2	7	4	0	2



On a tracé ci-contre le nuage de points associé ainsi que les trois droites d'ajustement (voir le cours pour les détails) :
- droite des points extrêmes ;
- droite de Mayer ;
- droite des moindres carrés.

1. Identifier chacune des droites d_1, d_2 et d_3.
2. Démontrer que :
   (a) la droite des points extrêmes a pour équation $y = -0{,}2x + 3{,}2$.
   (b) la droite de Mayer admet pour équation $y = -\dfrac{2}{3}x + \dfrac{16}{3}$.
   (c) la droite des moindres carrés a pour équation $y = -0{,}4x + 4{,}4$.
3. Quelles observations peut-on faire ?

Exercice 8 (Autres ajustements)
Un restaurateur effectue une étude de marché dans l'intention d'installer un établissement proposant un « grand buffet à volonté » avec des tarifs allant de 15 à 35 euros.
À l'issue de cette étude, il dispose du tableau ci-dessous donnant le nombre de couverts potentiels y_i en fonction du tarif x_i (en euro).

x_i	15	17	20	22,5	25	30	35
y_i	500	350	270	190	140	75	50

1. Représenter sur la calculatrice le nuage de points associé à la série statistique $(x_i; y_i)$.
   Un ajustement affine semble-il adapté ?
2. (a) Compléter la tableau suivant, en arrondissant à 10^{-2} près.

x_i	15	17	20	22,5	25	30	35
$z_i = \ln(y_i)$							

   (b) À l'aide de la calculatrice, déterminer l'équation de la droite de régression de z en x par les moindres carrés, en arrondissant les coefficients à 10^{-3}, ainsi que le coefficient de corrélation.
   Cet ajustement affine est-il adapté ?
   (c) En déduire que la fonction f définie sur $[15; 35]$ par $f(x) = 2639\,\mathrm{e}^{-0,116x}$ permet de modéliser le nombre de couverts en fonction du tarif x (en euro).
   (d) Le restaurateur estime que si moins de 100 couverts sont servis, son projet ne sera pas viable. Déterminer le prix maximum qu'il doit fixer pour assurer la rentabilité de son restaurant.

Exercice 9 (Autres ajustements)
On donne l'évolution de la cote Argus, en euro, d'un monospace tous les 1er janvier à partir de 2010 :

Année	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Rang x_i	0	1	2	3	4	5	6	7	8	9
Cote y_i	24 163	22 180	17 284	15 681	13 840	12 215	10 148	9 221	7 393	6 050

1. (a) Déterminer l'équation réduite de la droite de régression de y en x par les moindres carrés, en arrondissant les coefficients à 10^{-3}, ainsi que le coefficient de corrélation.
   (b) En utilisant cet ajustement, estimer la cote Argus du monospace entre 2020 et 2025.
2. (a) Recopier et compléter le tableau suivant, en arrondissant à 0,001 près :

Rang x_i	0	1	...	9
$z_i = \ln(y_i)$	10,093		...	8,708

   (b) Déterminer l'équation de la droite de régression de z en x par les moindres carrés, en arrondissant les coefficients à 10^{-3}, ainsi que le coefficient de corrélation.
   (c) En déduire un ajustement exponentiel de la cote y en fonction du rang x de l'année.
   (d) En utilisant cet ajustement, estimer la cote Argus du monospace entre 2020 et 2025.
3. Lequel des ajustements précédents (affine et exponentiel) paraît le plus adapté ?
   Donner deux arguments.