Statistiques descriptives - Exercices

Exercice 1 (statistiques à une variable) :

On considère un lot de barres de métal à la sortie d'une machine . On effectue une série de 30 mesures et on obtient les résultats suivants, en centimètres :

92,3	92,6	92,3	92,6	92,4	92,5
92,5	92,7	92,5	92,2	92,8	92,1
92,4	92,3	92,6	92,5	92,1	92,4
92,8	92,1	92,3	92,4	92,2	92,3
92,2	92,4	92,7	92,3	92,5	92,2

1. Calculer les effectifs et remplir la deuxième ligne du tableau :

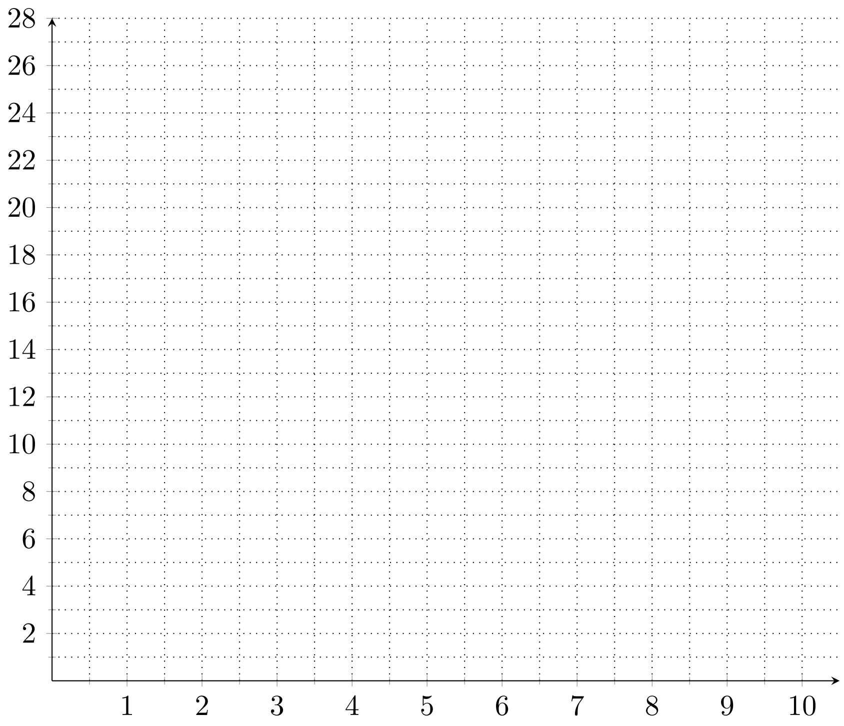
Longueur x_i	92,1	...	...	...
Effectif n_i	3	...	...	...
Fréquence f_i	0,1	...	...	...
Fréquence cumulée croissante fc_i	0,1	...	...	...

2. Construire le diagramme en bâtons correspondant (les x_i en abscisse et les n_i en ordonnée.)
3. Calculer les fréquences et les fréquences cumulées croissantes. Remplir les deux dernières lignes du tableau.
4. Calculer la moyenne et l'écart-type de cette série.
5. Calculer la médiane et les premier et troisième quartiles.
6. Représenter la série par un diagramme en boîte.

Exercice 2 (nuage, point moyen) :

Une entreprise fabrique des lots de circuits électroniques. Le tableau suivant indique le pourcentage y de circuits d'un lot qui ont une panne au cours de x semestres d'utilisation.

Nombre x_i de semestres	1	2	3	4	5	6	7	8	9	10
Pourcentage y_i	0	2	3	8	11	15	17	21	23	27



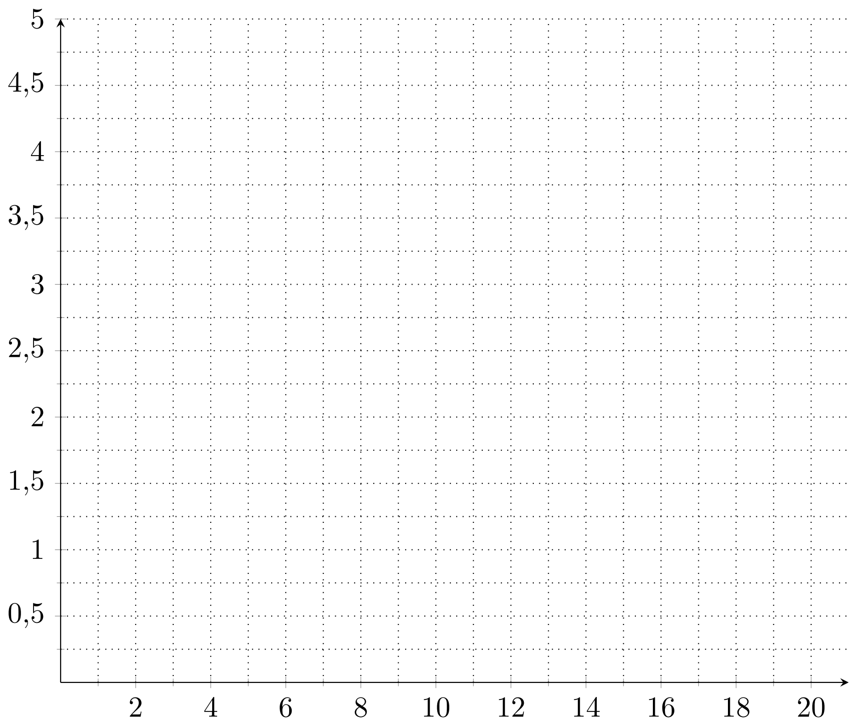
1. Représenter le nuage de points associé à cette série statistique.
2. Déterminer les coordonnées du point moyen G du nuage et le placer sur le graphique.
3. Au vu de la forme du nuage, un ajustement affine est-il pertinent ?
4. Si oui, tracer au jugé une droite D qui passe au plus près des points du nuage.
5. Déterminer l'équation de cette droite.
6. En utilisant cette droite, à combien peut-on estimer le pourcentage de circuits d'un lot qui ont une panne au cours de 16 semestres d'utilisation ?

Exercice 3 (droite de Mayer) :

Le mur d'une habitation est constitué par une paroi en béton et une couche de polystyrène d'épaisseur variable x ((en cm).

On a mesuré, pour une même épaisseur de béton, la résistance thermique y de ce mur en $^\circ$C.m^2/W (mètre carré fois degré C par Watt) pour différentes valeurs de x. On a obtenu les résultats suivants :

Épaisseur x_i	2	4	6	8	10	12	15	20
Résistance y_i	0,83	1,34	1,63	2,29	2,44	2,93	4,06	4,48



1. Représenter le nuage de points associé à cette série statistique.
2. Déterminer les coordonnées du point moyen G du nuage et le placer sur le graphique.
3. Déterminer les coordonnées de G_1, le point moyen des 4 premiers points du nuage (plus petites valeurs de x). Placer G_1 sur le graphique.
4. Déterminer les coordonnées de G_2, le point moyen des 4 derniers points du nuage (plus grandes valeurs de x). Placer G_2 sur le graphique.
5. Tracer la droite G_1G_2 et vérifier qu'elle passe par G.
6. Déterminer l'équation de cette droite.
7. En utilisant cette droite, à combien peut-on estimer la résistance thermique d'un mur comportant une épaisseur de 18 cm de polystyrène ?

Exercice 4 (droite de régression) – un classique :

Une machine fabrique en grande série des billes d'acier.

La moyenne des diamètres des billes produites en une semaine varie au cours du temps. La fabrication est jugée valable tant que cette moyenne reste dans l'intervalle [3,25 ; 3,32].

La semaine numérotée 0 est celle du réglage initial ; des contrôles hebdomadaires effectués lors des quatre premières semaines de fonctionnement ont donné les résultats suivants :

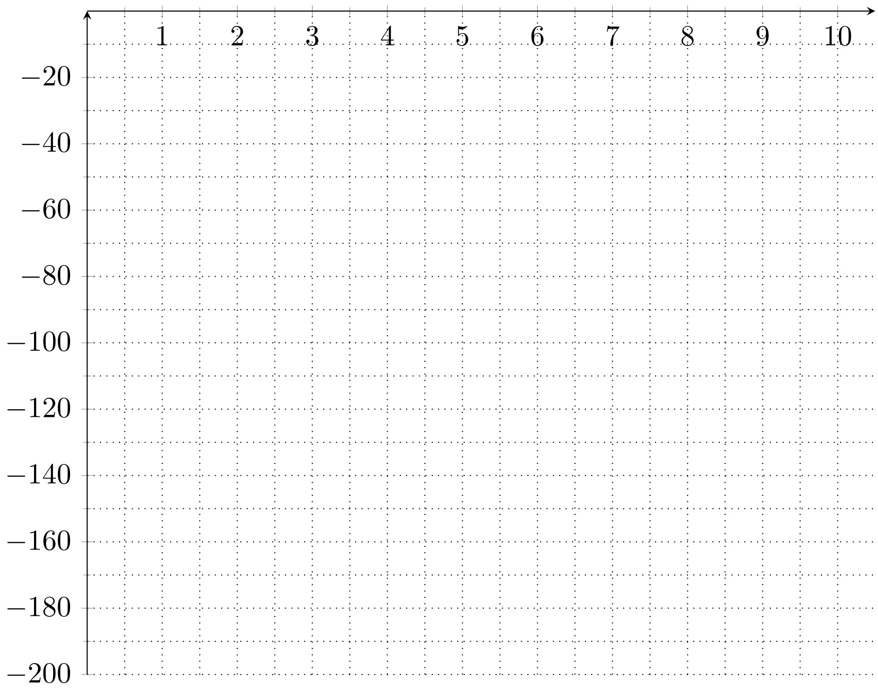
Semaine s	0	1	2	3	4
Moyenne m	3,32	3,32	3,31	3,29	3,27

1. Déterminer une valeur approchée à 10^{-3} près du coefficient de corrélation de la série statistique $s\,;m)$. Un ajustement affine est-il pertinent ?
2. Déterminer l'équation de la droite de régression de m en s (arrondir a et b à 10^{-2}).
3. Combien vaut alors, environ, le temps séparant deux réglages successifs de la machine ?

Exercice 5 (droite de régression) :

La température de fusion t (exprimée en degrés Celsius) de certains hydrocarbures varie avec le nombre x des atomes de carbone de leur molécule. Une série de 10 mesures donne les résultats suivants :

Nombre x	1	2	3	4	5	6	7	8	9	10
Température t	-190	-170	-180	-150	-130	-100	-90	-60	-50	-30

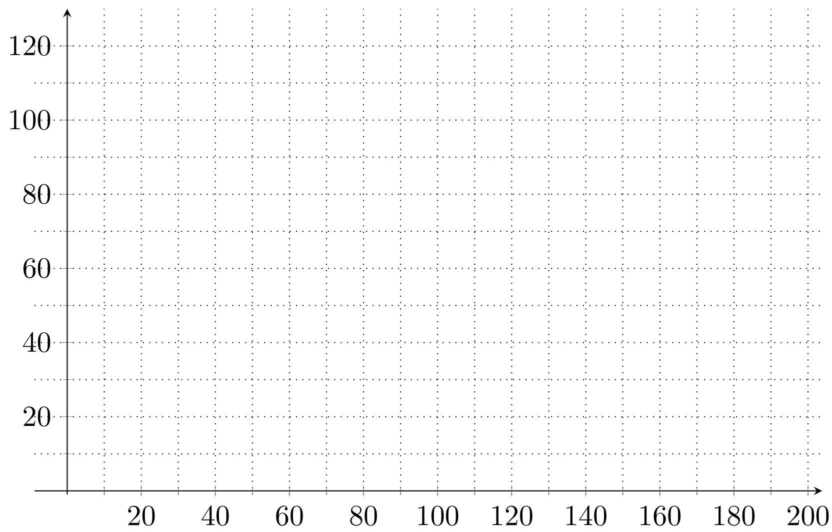


1. Représenter le nuage de points $M_i(x_i\,;t_i)$.
2. Déterminer les coordonnées du point moyen G du nuage et le placer sur le graphique.
3. Déterminer une valeur approchée à 10^{-3} près du coefficient de corrélation de la série statistique $x\,;t)$. Que peut-on en conclure ?
4. Déterminer l'équation de la droite de régression D de t en x (arrondir a à 10^{-2} et b à l'unité).
5. Tracer D sur le graphique.
6. Estimer à l'aide de la droite D la température de fusion de l'hydrocarbure $C_{12}H_{26}$.

Exercice 6 (nuage) :

Le tableau ci-dessous représente les précipitation moyenne en mm, pour chaque mois de l'année, pour les villes de Dakar et Nice. (Source : climate-data.org).

Mois	J	F	M	A	M	J	J	A	S	O	N	D
Précipitations à Dakar d (mm)	1	1	0	0	1	10	77	182	150	42	2	3
Précipitations à Nice n (mm)	74	72	70	67	52	41	18	34	67	113	118	85



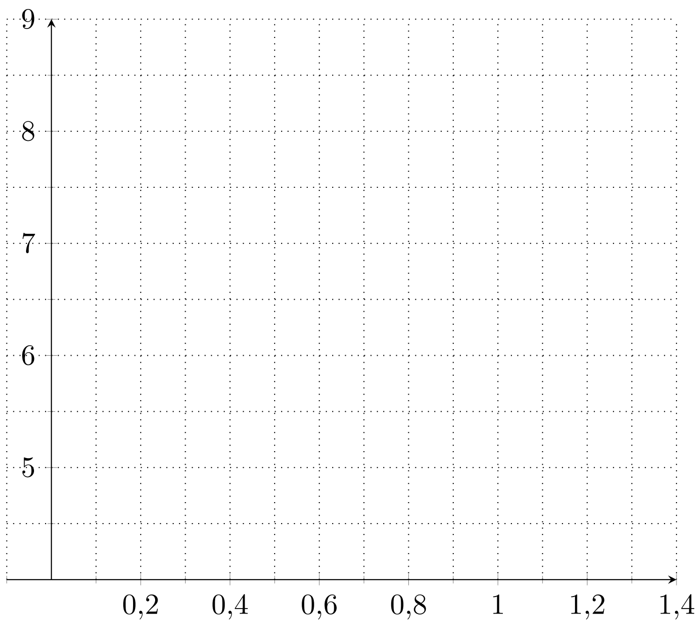
1. Représenter le nuage de point associé à la série $(d\,;n)$.
   Un ajustement affine est-il pertinent ?
2. Calculer le coefficient de corrélation. Que peut-on conclure ?

Exercice 7 (droite de régression - changement de variable) :

L'exercice concerne l'étude cinétique d'une réaction catalysée par la β-fructosidase.

La vitesse initiale v de la réaction a été mesurée en présence de différentes concentrations, notées s, de substrats dans des conditions identiques de pH et de température. Les résultats sont les suivants :

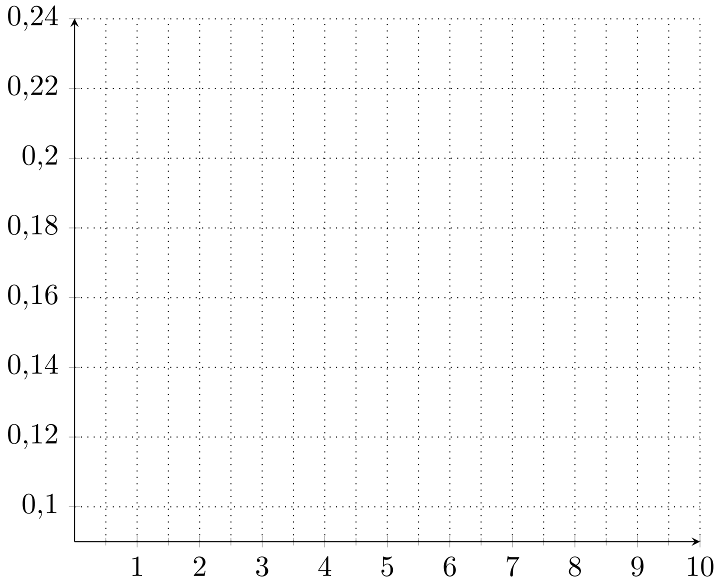
Concentration de substrat s_i en mmol.L^{-1}	0,1	0,2	0,3	0,5	1	1,3
Vitesse initiale v_i en μmol par minute	4,5	5,6	6,3	6,9	7,7	8,1



1. Représenter le nuage de point associé à la série statistique $(s\,;v)$. Que peut-on remarquer ?
2. On effectue les changements de variable : $x_i = \dfrac{1}{s_i}$ et $y_i = \dfrac{1}{v_i}$.

x_i	10					
y_i	0,22					

   a) Recopier et compléter le tableau ci-dessus (valeurs arrondies à 10^{-2} près).
   b) Représenter le nuage de points $M_i\,(x_i,\, y_i)$.



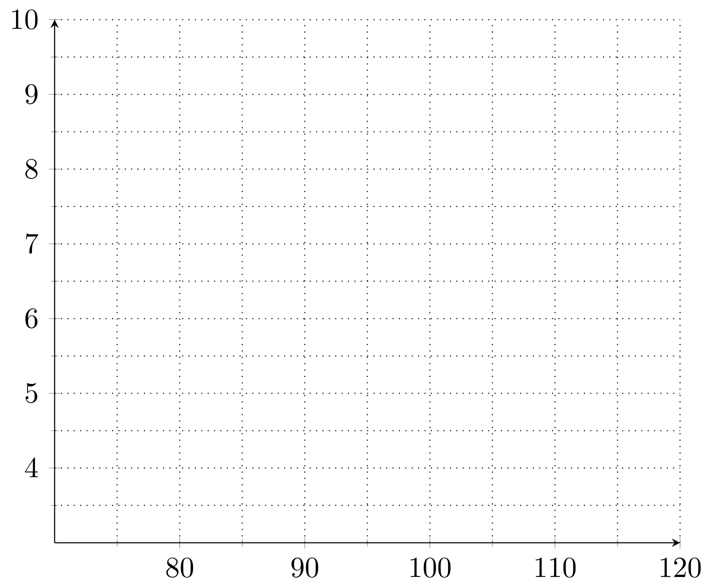
3. a) Par la méthode des moindres carrés, déterminer une équation de la droite de régression D de y en x. sous la forme $y = ax + b$ (a est arrondi à 10^{-4} près et b est arrondi à 10^{-2} près).
   b) Tracer la droite D sur le graphique précédent.
   c) En déduire une formule approchée donnant v en fonction de s.

Exercice 8 (droite de régression - changement de variable) :

La consommation y d'une voiture est donnée en fonction de sa vitesse moyenne x par le tableau suivant :

x en km/h	80	90	100	110	120
y en L/100 km	4	4,8	6,3	8	10

1. La consommation est-elle proportionnelle à la vitesse moyenne ?
2. a) Représenter le nuage de points correspondant à la série statistique $(x\,;y)$.



b) Calculer les coordonnées du point moyen G du nuage et le placer sur le graphique.

c) Déterminer (calculatrice) une équation de la droite de régression de y en x. Tracer cette droite sur le graphique.

d) En utilisant cet ajustement, estimer la consommation de la voiture pour une vitesse de 130 km/h (arrondir au dixième).

3. La forme du nuage permet d'envisager un ajustement exponentiel.
On pose $z = \ln(y)$ et on admet que la droite de régression obtenue pour les cinq points $(x\,;z)$ a pour équation : $z = 0{,}0234x - 0{,}508$.

a) À partir de cette équation, écrire y sous la forme $y = A\mathrm{e}^{Bx}$ (donner A et B arrondis à 10^{-4}).

b) Tracer sur le graphique la courbe d'équation $y = A\mathrm{e}^{Bx}$.

c) En utilisant cet ajustement, estimer la consommation de la voiture pour une vitesse de 130 km/h.

Exercice 9 (droite de régression - changement de variable) :

Un mobile est propulsé à très grande vitesse sur un axe, puis ralenti. On s'intéresse à la vitesse de ce mobile durant le freinage.

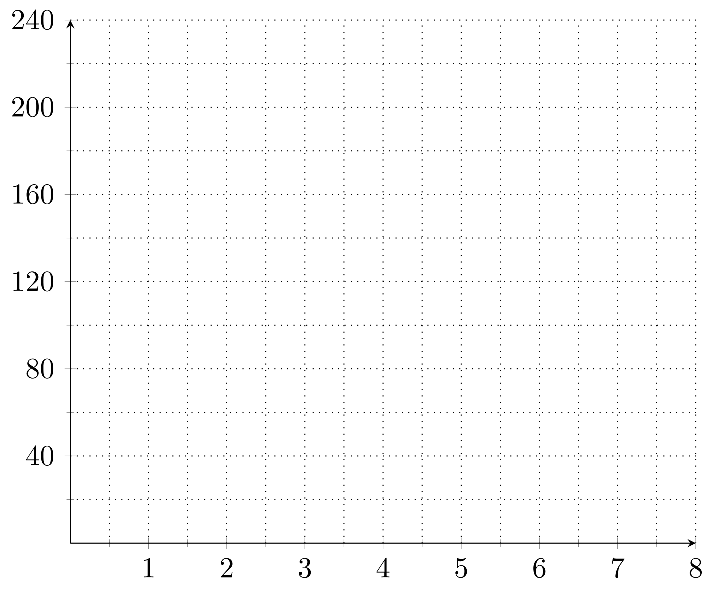
Les distances sont exprimées en mètres, les temps en secondes et les vitesses en mètres par seconde.

On arrondit les résultats au centième.

On a relevé les vitesses instantanées v_i de ce mobile aux instants t_i, l'indice i variant de 0 à 7.

t_i en s	0	1	2	3	4	5	6	7
v_i en m.s^{-1}	215	140	85	57	36	29	27	22

1. Représenter le nuage de points de cette série. Calculer le coefficient de corrélation.



Expliquer pourquoi on décide de ne pas réaliser un ajustement affine.

2. On pose $n_i = \ln(v_i - 15)$. Dresser le tableau de la série $(t\,;n)$.

3. Donner l'équation de la droite de régression de n et t. Calculer le coefficient de corrélation de la série $(t\,;n)$.

4. En déduire une expression de la vitesse v en fonction du temps sous la forme : $v = \alpha\mathrm{e}^{\beta t} + \gamma$, où α, β, γ sont des réels à déterminer.