LDDR-Niveau I: Statistique

Exercice 8.1

Dans les situations suivantes, d'écrire la population/l'échantillon étudiés, et nommer et donner le type de la variable étudiée.

a) Afin de déterminer le profil socio-économique des ménages du canton de Neuchâtel, on note le nombre d'enfants par ménage pour un échantillon de 1000 ménages.

b) Selon l'Office fédéral de la statistique pour l'année 2013, en Suisse, 63,5% de la population parle le plus souvent l'allemand, 22,5% parle le français, 8,2% parle l'italien, 0,5% parle le romanche et 5,3% parle une autre langue.

c) On demande à chaque élève de 1MG son taux d'échec au semestre (le taux d'échec étant le nombre de notes insuffisantes divisé par le nombre total de notes) :
1. 0% 2. De 1% à 15,9% 3. De 16% à 49,9% 4. 50% et plus.

Exercice 8.2

Lors d'un sondage, on a demandé à 820 citoyens suisses leur opinion sur les accords bilatéraux Suisse-UE. Les réponses se répartissent comme suit.

Utilité	Effectif	Fréquence
Très utiles	95	
Utiles	342	
Nuisibles	210	
Très nuisibles	46	
Sans opinion	127	
Total		

a) Compléter le tableau distribution ci-dessus et représenter graphiquement la distribution par un diagramme approprié au type de variable.

b) Calculer le taux de confiance de ces accords, soit le pourcentage de personnes qui estiment les accords bilatéraux utiles ou très utiles.

Exercice 8.3

Sur une route où la vitesse est limitée à 80 km/h, on a mesuré la vitesse de 50 véhicules.

84	81	76	71	80	81	83	84	80	83
74	75	92	76	80	82	94	73	83	83
75	81	79	97	78	82	76	78	82	82
78	81	91	68	82	73	82	79	75	77
83	80	77	81	69	78	81	83	87	87

a) Grouper les données en classes et dresser un tableau de distribution.

b) Construire l'histogramme et le polygone des fréquences.

c) Compléter l'analyse suivante : « Une .... des véhicules respectent la limitation de vitesse de 80 km/h et ..... % roulent entre 80 et 85 km/h. En tenant compte d'une marge de tolérance de 5 km/h, ... % des véhicules sont amendables. »

Exercice 8.4

Les séries $S_1 = \{3,4,4,4,4,4,5\}$, $S_2 = \{1,3,4,4,5,5,6\}$ et $S_3 = \{1,1,4,4,6,6,6\}$ représentent les notes obtenues lors de trois tests dans une classe à l'effectif utopique de 7 élèves. Calculer l'écart-type de chaque série et interpréter.

Exercice 8.5

Dans une usine, lors d'un contrôle qualité, le diamètre, en mm, de 50 boulons tirés au hasard dans la production a été mesuré. Les résultats suivants ont été obtenus.

Répartition de 50 boulons selon leur diamètre.

Diamètre [mm]	**Effectifs**
[21.5; 21.8[	4
[21.8; 21.9[	6
[21.9; 22.0[	6
[22.0; 22.1[	13
[22.1; 22.2[	8
[22.2; 22.3[	7
[22.3; 22.5[	6
Total	50

a) Représenter l'histogramme de ces données ainsi que le polygone des fréquences cumulées.

b) Si la valeur nominale du diamètre des boulons est de 22 mm, calculer le pourcentage de boulons qui s'en écartent de plus de 0.3 mm. Vérifier la cohérence du résultat sur le polygone des fréquences cumulées.

Exercice 8.6

a. Développer puis regrouper l'expression $\sum_{i=1}^{3}(x_i - a)^2$

b. Écrire l'expression suivante à l'aide du symbole de sommation :

$$3ax_1y_1z_1 + 3ax_2y_2z_2 + \cdots + 3ax_ky_kz_k$$

c. Sachant que $\sum_{i=1}^{6} x_i = -4$ et que $\sum_{i=1}^{6} x_i^2 = 100$,calculer :

$\sum_{i=1}^{6}(2x_i + 3)$ $\qquad$ $\sum_{i=1}^{6} x_i(x_i - 1)$ $\qquad$ $\sum_{i=1}^{6}(x_i - 5)^2$

d. Les deux variables x et y prennent les valeurs x_i et y_i ci-dessous. Calculer :

x	2	−5	4	−8
y	−3	−8	10	6

1. $\sum_{i=1}^{4} x_i y_i$
2. $\sum_{i=1}^{4} x_i \cdot \sum_{i=1}^{4} x_i$
3. $\sum_{i=1}^{4} x_i^2 + y_i^2$
4. $\sum_{i=1}^{4} x_i^2 + \sum_{i=1}^{4} y_i^2$

Exercice 8.7

Reprendre la donnée de l'exercice 8.5.

a) Calculer le diamètre moyen des boulons et représenter la moyenne par un triangle sous l'axe horizontal de l'histogramme.

b) Calculer de diamètre médian. Marquer cette valeur par une barre verticale sur l'histogramme.

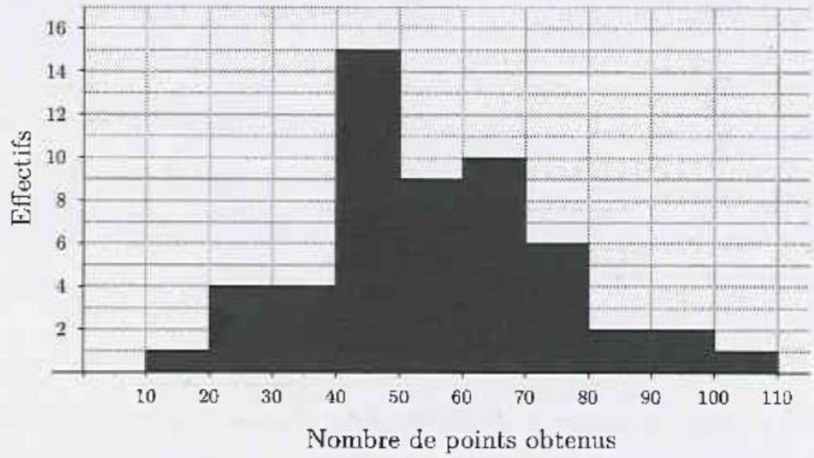
c) Déterminer la classe modale. Cette notion est-elle représentative ici ? Justifier la réponse.

d) Que peut-on conclure en comparant la moyenne, la médiane et la classe modale sur la forme de la distribution des diamètres des boulons ?

Exercice 8.8

Le nombre de points obtenus par les écoles de Suisse au concours de *Mathématiques sans Frontières* est représenté dans l'histogramme suivant :

Répartition de 54 écoles suisses selon le nombre de points à MSF



a) Calculer la moyenne $\bar{x}$ et l'écart-type σ de ces résultats et interpréter ces mesures. Marquer ces résultats sur le graphique de façon appropriée.

b) Quelle est la variable centrée réduite d'une école ayant obtenu 110 points ?
Quel est le nombre de points obtenus par une école qui présente une variable centrée réduite égale à -2 ?

Exercice 8.9

Le tableau ci-dessous décrit l'évolution du prix d'une baguette de pain dans une boulangerie.

Année	2010	2011	2012	2013	2014	2015
x	1	2	3	4	5	6
$p_{[CHF]}$	1.75	2	2.1	2.25	2.4	2.55

a. Représenter cette série statistique dans un repère orthonormée. Peut-on envisager un ajustement affine ?

b. On appelle G_1 le point moyen des 3 premiers relevés, G_2 le point moyen des 3 derniers relevés. Calculer G_1, G_2, puis donner l'équation de la droite d_1 d'ajustement de Mayer

c. Donner l'équation de la droite d_2 d'ajustement par la méthode des moindres carrés.

d. Comparer les deux méthodes et donner les prix prévisibles en 2018.

Exercice 8.10

On donne le tableau de valeurs ci-dessous.

x	1	2	3	4
y	1	3	6	7

a. En résolvant un système de 2 équations linéaires à deux inconnues, trouver la droite de régression $y = ax + b$ pour le nuage de points donné.

b. Vérifier le résultat précédent en utilisant les formules vues dans la théorie.

c. On peut démontrer qu'une droite de régression obtenue par la méthode des moindres carrés passe toujours par le point $P(\bar{x};\bar{y})$. Vérifier cette proposition avec les valeurs numériques ci-dessus.

Exercice 8.11

On donne le tableau de valeurs ci-dessous.

x	1	3	4	6	8	9	11	14
y	1	2	4	4	5	7	8	9

a. Trouver la droite des moindres carrés de y en x ($y = ax + b$)

b. Trouver une nouvelle droite de régression obtenue en inversant les rôles de x et de y, c'est-à-dire la droite de régression de x en y. Cette droite sert à estimer x pour une valeur donnée de y.

c. Vérifier que les deux droites se coupent au point $P(\overline{x}; \overline{y})$

Exercice 8.12

Le nombre de bactéries y par unité de volume présent dans un bouillon de culture après x heures est donné par le tableau ci-dessous

x	0	1	2	3	4	5	6
y	32	47	65	92	132	190	275

a. Dessiner le nuage de points et constater qu'il n'y a pas de corrélation linéaire entre ces variables

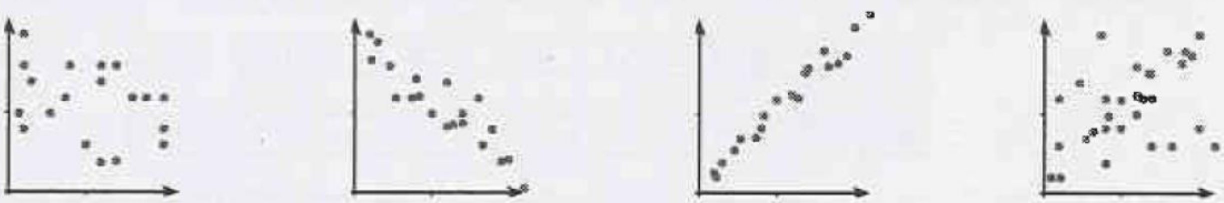
b. Effectuer le changement de variable $y' = \ln(y)$ puis dessiner le nouveau nuage de points de coordonnées x et y'. Constater que cette fois les points s'approchent d'une droite.

c. Calculer l'équation de la droite de régression de y' en x.

d. Trouver puis dessiner la courbe de régression de y en x sachant qu'elle est de la forme $y = a \cdot b^x$

Exercice 8.13

Associer les coefficients de corrélation suivants $r_1 = 0.44$, $r_2 = -0.91$, $r_3 = -0.36$ et $r_4 = 0.98$, au nuage de points dessiné ci-dessous



Exercice 8.14

Un test a permis de construire le tableau ci-dessous qui présente la distance de freinage d $_{[m]}$d'un véhicule sur route sèche en fonction de sa vitesse v $_{[km/h]}$.

v	40	50	60	70	80	90	100	110	120
d	20.29	28.42	35.57	45.75	58.94	70.12	95.15	98.17	113.19

a. Déterminer le coefficient de corrélation linaire entre v et d puis celui entre v et $z = \sqrt{d}$

b. Estimer la distance de freinage nécessaire à un véhicule circulant à 200 km/h

Exercice 8.15 *Idéalement avec un ordinateur...*

On veut étudier l'évolution des records de l'épreuve d'athlétisme du 100 mètres masculin. On donne dans le tableau suivant certains records, établis depuis 1900.

Année	1900	1912	1921	1930	1964	1983	1991	1999
x	0	12	21	30	64	83	91	99
t	10.80	10.60	10.40	10.30	10.06	9.93	9.86	9.79

Après étude, on effectue les changements de variables suivants : $a = e^{-0.00924x}$ et $b = \ln(t)$ pour modéliser notre problème. On obtient alors le tableau :

a	1.000	0.895	0.824	0.758	0.554	0.464	0.431	0.401
b	2.380	2.361	2.342	2.332	2.309	2.296	2.288	2.281

a. Calculer l'équation de la droite de régression de b en a obtenue par la méthode des moindres carrés.

b. En déduire que l'on peut modéliser une expression de t en fonction de x sous la forme suivante :

$$t = e^{a \cdot e^{-0.00924x}+b}$$

où a et b sont deux réels à déterminer.

c. A l'aide de cet ajustement, quel record du 100 mètres peut-on prévoir en 2010 ?

d. Calculer $\lim\limits_{x\to\infty} e^{0.154 \cdot e^{-0.00924x}+2.221}$

e. Que peut-on en conclure, en utilisant ce modèle, quant aux records du cent mètres masculin , à très long terme ?

Exercice 8.16 *

a. Selon le même modèle que dans la théorie, établir les équations permettant de trouver la parabole des moindres carrés $y = ax^2$ de régression de y en x.

b. Appliquer votre modèle et trouver la parabole de régression de y en x pour les données suivantes :

x	0.2	−0.7	0.5	0.6	−0.4
y	0.1	1	0.5	0.7	0.3