

POLYTECH Lyon
Introduction aux méthodes probabilistes et statistiques

TD 1. Statistiques descriptives

Exercice 1 – D’après vous, les variables suivantes sont-elles des variables quantitatives discrètes, des variables quantitatives continues ou des variables qualitatives ? Quelles sont les représentations graphiques et calculs adaptés à chacune d’elles ?

- a) les notes des étudiants au prochain contrôle continu
- b) les notes des étudiants transcrites en notes anglosaxonnes (A, B, C,...)
- c) la taille (en cm) des étudiants de Lyon-1
- d) la mention au baccalauréat des étudiants de cette promotion
- e) le nombre d’années d’études supérieures avant l’admission à l’ISTIL des étudiants de cette promotion
- f) la catégorie socio-professionnelle des clients d’un opérateur téléphonique
- g) le poids des marchandises chargées dans les camions d’une plate-forme logistique d’une chaîne d’hypermarchés
- h) le nombre d’impacts de foudre par jour

Exercice 2 – Un atelier réalise le séchage de boues d’origine industrielle. Il obtient à la fin du processus des déchets (mesurés en kg). On a observé les poids suivants de déchets après le traitement de 100 kg :

4,7 4,3 4,5 4,9 4,2 4,7 4,0 4,2 5,0 3,9 4,6 4,6
4,8 4,4 4,2 4,6 4,3 4,9 4,0 4,5 4,1 4,4 4,3 4,3.

Notons x cette série numérique.

1. Opérer le dénombrement des différentes modalités du caractère et construire le tableau des effectifs et fréquences, puis tracer le diagramme cumulatif.
2. Rappeler la définition de la moyenne, la variance et l’écart-type. Les calculer pour notre exemple.
3. Définir la médiane et les quartiles, les déterminer. Tracer le boxplot.
4. On établit trois classes pour les modalités en prenant comme bornes 4,25 et 4,65. Écrire un tableau donnant les classes et leurs effectifs. En faire une représentation graphique (histogramme).
5. Supposons que la 9^e valeur soit 50 et non 5,0 (à cause d’une erreur d’unité dans la saisie des données). Que devient alors le boxplot ? Et le résumé numérique ?

Exercice 3 – On lance 100 fois un dé et on note les résultats :

face	1	2	3	4	5	6
nombre d’occurrences	13	16	18	16	13	24

1. Proposer deux manières de représenter graphiquement ces données.
2. Calculer le résumé numérique.
3. Tracer le boxplot.

Exercice 4 – Pour les données suivantes,

x	1	2	7	4	6
y	5	4	1	3	2

1. Tracer le nuage de points.
2. Deviner le signe et la valeur du coefficient de corrélation.

3. Calculer le coefficient de corrélation, la pente et l'ordonnée à l'origine de la droite de régression.

Exercice 5 – Un biologiste a récolté 10 lézards sauvages dans le sud-ouest des États-Unis. Après avoir mesuré leur longueur (en mm), on mesure leur vitesse de course (en m/s).

vitesse	1.28	1.36	1.24	2.47	1.94	2.52	2.67	1.29	1.56	2.66
longueur	179	157	169	146	143	131	159	142	141	130

1. Faire un diagramme et commenter les éventuelles observations inattendues.
2. Calculer le coefficient de corrélation ainsi que l'équation de la droite de régression linéaire.
3. Critiquer le modèle. Comment pourrait-on l'améliorer ?

Exercice 6 – Nous disposons d'un tableau de données qui comprend la quantité de goudrons et de monoxyde de carbone présents dans les cigarettes courantes.

1. Calculer un résumé numérique de la quantité de goudrons en prenant soin de définir chaque grandeur calculée. Puis tracez le boxplot. Observe-t-on des valeurs extrêmes ?
2. Tracer un histogramme, avec les classes $[1, 5[$, $[5, 10[$, $[10, 15[$, $[15, 20[$, $[20, 29.8]$?
3. On veut maintenant étudier la relation entre la quantité de goudrons et de monoxyde de carbone. Effectuer une régression linéaire des goudrons en fonction du monoxyde de carbone. Que pensez-vous de ce modèle ?

Marque	Goudrons (mg)	CO (mg)
Alpine	14,1	13,6
Benson & Hedges	16	16,6
Bull Durham	29,8	23,5
Camel Lights	8	10,2
Carlton	4,1	5,4
Chesterfield	15	15
Golden Lights	8,8	9
Kent	12,4	12,3
Kool	16,6	16,3
L & M	14,9	15,4
Lark Lights	13,7	13
Marlboro	15,1	14,4
Merit	7,8	10
Multi Filter	11,4	10,2
Newport Lights	9	9,5
Now	1	1,5
Old Gold	17	18,5
Pall Mall Light	12,8	12,6
Raleigh	15,8	17,5
Salem Ultra	4,5	4,9
Tareyton	14,5	15,9
VRAI	7,3	8,5
Viceroy Rich Light	8,6	10,6
Virginia Slims	15,2	13,9
Winston Lights	12	14,9

Exercice complémentaire– Voici un échantillon de salaires mensuels (net, en k€) d'ingénieurs de l'industrie :

1.2 1.5 1.6 1.7 1.7 1.8 1.8 1.9 2.0 2.1 2.2 2.2 2.3 2.3 2.4
 2.4 2.5 2.6 2.7 2.7 2.8 2.8 2.9 3.0 3.1 3.1 3.2 3.3 3.8 4.2

Réaliser une étude statistique de ces données.