Table des matières

1 Statistique descriptive

1.1 Exercice (Terminologie)

La variable statistique "couleur de maisons d'un quartier" est-elle : qualitative, quantitative, discrète, continue ?
La variable statistique "revenu brut" est-elle : qualitative, quantitative, discrète, continue ?
La variable statistique "nombre de maisons vendues par ville" est-elle : qualitative, quantitative, discrète, continue ?
Parmi ces assertions, préciser celles qui sont vraies, celles qui sont fausses :

1. On appelle variable, une caractéristique que l'on étudie.
2. La tâche de la statistique descriptive est de recueillir des données.
3. La tâche de la statistique descriptive est de présenter les données sous forme de tableaux, de graphiques et d'indicateurs statistiques.
4. En Statistique, on classe les variables selon différents types.
5. Les valeurs des variables sont aussi appelées modalités.
6. Pour une variable qualitative, chaque individu statistique ne peut avoir qu'une seule modalité.
7. Pour faire des traitements statistiques, il arrive qu'on transforme une variable quantitative en variable qualitative.
8. La variable quantitative poids d'automobile peut être reclassée en compacte, intermédiaire et grosse.
9. En pratique, lorsqu'une variable quantitative discrète prend un grand nombre de valeurs distinctes, on la traite comme continue.

Proposer des exemples de variable quantitative transformée en variable qualitative. Préciser les modalités de cette dernière.
Pour chacune des variables suivantes, préciser si elle est qualitative, quantitative discrète ou quantitative continue : (a) Revenu annuel. (b) Citoyenneté. (c) Distance. (d) Taille. (e) Lieu de résidence. (f) Age. (g) Couleur des yeux. (h) Nombre de langues parlées.
Pour les sujets d'étude qui suivent, spécifier : l'unité statistique, la variable statistique et son type :

1. Étude du temps de validité des lampes électriques.
2. Étude de l'absentéisme des ouvriers, en jours, dans une usine.
3. Répartition des étudiants d'une promotion selon la mention obtenue sue le diplôme du Bac.
4. On cherche à modéliser 1 le nombre de collisions impliquant deux voitures sur un ensemble de 100 intersections routières choisies au hasard dans une ville. Les données sont collectées sur une période d'un an et le nombre d'accidents pour chaque intersection est ainsi mesuré.

1.2 Exercice (indicateurs)

On dispose des résultats d'une enquête concernant l'âge et les loisirs d'une population de 20 personnes (S : sport, C : cuisine, V : voyager, L : lecture) :

Age	12	14	40	35	26	30	30	50	75	50	30	45	25	55	28	25	50	40	25	35
Loisir	S	S	C	C	S	V	V	L	L	L	V	C	C	C	S	L	L	C	V	V

1. Faire l'étude du caractère « âge » : dresser le tableau statistique (effectifs[1], effectifs cumulés), calculer les valeurs de tendance centrale et ceux de la dispersion et tracez le diagramme en bâtons et la boite à moustaches de cette distribution.
2. Faire l'étude du caractère « Loisir » dresser le tableau statistique, déterminer le mode et tracez le diagramme en bâtons et le diagramme à secteurs.

[1] n_i : le nombre d'individus qui ont le même x_i, ça s'appelle effectif partiel de x_i. L'effectif cumulé N_i d'une valeur est la somme de l'effectif de cette valeur et de tous les effectifs des valeurs qui précèdent.

1.3 Exercice (indicateurs, visualisation)

Le gérant d'un magasin vendant des articles de consommation courante a relevé pour un article particulier qui semble connaître une très forte popularité, le nombre d'articles vendus par jour. Son relevé a porté sur les ventes des mois de Mars et Avril, ce qui correspond à 52 jours de vente. Le relevé des observations se présente comme suit :

7 13 8 10 9 12 10 8 9 10 6 14 7 15 9 11 12 11 12 5 14 11 8 10 14 12 8
5 7 13 12 16 11 9 11 11 12 12 15 14 5 14 9 9 14 13 11 10 11 12 9 15.

1. Quel type est la variable statistique étudiée.
2. Déterminer le tableau statistique en fonction des effectifs, des fréquences, des effectifs cumulés et des fréquences cumulés.
3. Tracer le diagramme des bâtonnés associé à la variable X.
4. Soit Fx la fonction de répartition. Déterminer Fx .
5. Calculer le mode Mo et la moyenne arithmétique x.
6. Déterminer à partir du tableau puis à partir du graphe, la valeur de la médiane Me.
7. Calculer la variance et l'écart-type.

1.4 Exercice (indicateurs, visualisation)

On observe 100 fois le nombre d'arrivées (variable X) de clients à un bureau de poste pendant un intervalle de temps (10 minutes) et on obtient les valeurs suivantes :

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	5	5	5	5	5	5	5	6	6	6	6	6	6	6

1. Dresser le tableau statistique de la distribution de la variable X (effectifs cumulés, …).
2. Calculer les valeurs de tendance centrale de la distribution : la moyenne, le mode et les trois quartiles Q1, Q2 et Q3.
3. Calculer les valeurs de la dispersion de la distribution : variance, l'écart type et l'intervalle interquartile.
4. Tracer le diagramme en bâtons et la boite à moustaches de cette distribution.

1.5 Exercice (intervalles, indicateurs, visualisation)

Soient des résultats d'une enquête sur les loyers annuels des appartements dans un quartier.

Montant du loyer (x 1000)	Effectifs
[4, 6[	20
[6, 8[	40
[8, 10[	80
[10, 15[	30
[15, 20[	20
[20, 30[	10

1) Compléter le tableau statistique (valeurs centrales, effectifs cumulés, fréquence, fréquences cumulés)

1.6 Exercice (mise en œuvre avec R)

Mettre en œuvre l'ensemble de toutes les solutions des exercices de la fiche TD avec R.

2 Théorie des probabilités

2.1 Exercice : Notion de probabilité

1. Nous lançons une pièce de monnaie.
a) Décrivez les résultats possibles.
b) Décrivez l'univers Ω associé au lancer d'une pièce de monnaie.
c) Donnez deux événements associés à cette expérience aléatoire.

2. Nous lançons un dé cubique dont les faces sont numérotées de 1 à 6.
a) Décrivez les résultats possibles.
b) Décrivez l'univers Ω associé au lancer d'un dé cubique.
c) Donnez trois événements associés à cette expérience aléatoire.

3. Nous jetons trois fois un dé cubique dont les faces sont numérotées de 1 à 6. Décrivez l'univers Ω associé à cette expérience aléatoire.

2.2 Exercice : Notion de probabilité

Nous jetons simultanément deux dés cubiques de couleur différente dont les faces sont numérotées de 1 à 6. Nous notons par E l'événement « la somme des deux dés cubiques est un nombre impair », par F l'événement « au moins l'un des deux dés cubiques a une face égale à 1 », et par G « la somme des deux dés cubiques est égale à 5 ». Décrivez $E \cap F$, $E \cup F$, $F \cap G$, $E \cap F^c$ et $E \cap F \cap G$.

2.3 Exercice : Notion de probabilité

Trois joueurs α, β et γ jettent une pièce à tour de rôle. Le premier qui obtient pile a gagné. Nous admettrons que α joue d'abord, puis β et enfin γ. L'univers Ω associé à cette expérience aléatoire peut être décrit comme suit :

$$\Omega = \{1, 01, 001, 0001, \ldots\} \cup \{0000\ldots\}.$$

1. Donnez une interprétation des points de Ω.

2. Décrivez les événements suivants en terme de ces points :
a) premier événement : A = « α gagne » ;
b) deuxième événement : B = « β gagne » ;
c) troisième événement : $(A \cup B)^c$.

2.4 Exercice : Notion de probabilité

	Vrai	Faux
1. On peut associer deux ensembles fondamentaux différents à une même expérience.	☐	☐
2. L'égalité $P(A \cup B) = P(A) + P(B)$ implique que les événements A et B sont incompatibles.	☐	☐
3. L'information apportée par la connaissance de la réalisation d'un événement B augmente la probabilité d'un autre événement A, *i.e.* $P(A\|B) \geq P(A)$.	☐	☐
4. Si deux événements sont incompatibles, alors ils sont indépendants.	☐	☐
5. Si deux événements sont indépendants, alors leurs complémentaires le sont aussi.	☐	☐
6. Soit $\Omega = \{a, b, c, d\}$ avec équiprobabilité des événements élémentaires sur $\mathcal{P}(\Omega)$. Les événements $A = \{a, b\}$ et $B = \{a, c\}$ sont dépendants car ils sont réalisés simultanément quand l'événement élémentaire a est réalisé.	☐	☐

2.5 Exercice : Variables aléatoires

	Vrai	Faux
1. Une variable aléatoire est une valeur numérique.	☐	☐
2. Toute application de $\mathcal{P}(\Omega)$ dans $\mathbb{N}$ est une variable aléatoire discrète.	☐	☐
3. L'espérance mathématique d'une v.a. discrète est la valeur qu'elle prend le plus fréquemment.	☐	☐
4. L'espérance mathématique d'une somme de deux v.a. discrètes est toujours la somme de leurs espérances.	☐	☐
5. Si deux v.a. ont la même espérance, alors elles ont la même variance.	☐	☐
6. La variance d'une somme de deux v.a. discrètes est toujours la somme de leurs variances.	☐	☐

2.6 Exercice : Variables aléatoires

Un joueur dispose d'un dé cubique truqué de telle sorte que la probabilité d'apparition d'un numéro est proportionnelle à ce dernier. Nous supposons que les faces sont numérotées de 1 à 6. Soit X la variable aléatoire réelle associée au lancer de ce dé.

1. Déterminez la loi de X.

2. Nous posons $Y = \frac{1}{X}\cdot$ Déterminez la loi de Y.

2.7 Exercice : Variables aléatoires

Soit X une v.a.d. telle que $\mathbb{P}([X=-1]) = \mathbb{P}([X=0]) = \mathbb{P}([X=1]) = 1/3$ et $\mathbb{P}([X=x]) = 0$ pour tout $x \in \mathbb{R} \setminus \{-1,0,1\}$.
Soit $Y = X^2$.
Montrez que les variables aléatoires X et Y ne sont pas corrélées. Sont-elles indépendantes ?

2.8 Exercice : Variables aléatoires continues

Soit f la fonction définie par

$$f(t) = \begin{cases} 0 & \text{si } t \notin [0,1] \\ 2t & \text{si } t \in [0,1] \end{cases}$$

1. a) Démontrez que la fonction f est une densité de probabilité.
b) Démontrez que la loi de probabilité définie par f admet une espérance mathématique et une variance que vous préciserez.
c) Déterminez la fonction de répartition associée à f.

2. Soit X une variable aléatoire de densité de probabilité f définie ci-dessus. Nous définissons $Y = 1 + X^2$. Déterminez la fonction de répartition de Y.

2.9 Exercice : Loi normale

Une variable aléatoire X suit la loi normale $\mathcal{N}(5,2;\, 0,8)$.

1. Calculez $\mathbb{P}(X \geqslant 4 | X \leqslant 5,2)$.

2. Une variable aléatoire Y indépendante de la variable X suit la loi normale $\mathcal{N}(\mu;\, 0,8)$.

- **a.** Quelle est la loi de la variable aléatoire $Y - 2X$?
- **b.** Déterminez μ sachant que $\mathbb{P}(Y \geqslant 2X) = 0,516$.

2.10 Exercice : Loi définie

Loi définie par sa f.r.

14. Soit X une v.a. de f.r. F définie par :

$$F(x) = \begin{cases} 0 & \text{si } x \leq 1 \\ \ln x & \text{si } 1 < x \leq \mathrm{e} \\ 1 & \text{si } \mathrm{e} < x \end{cases}$$

Calculer $E(X)$ et $V(X)$.

Analyse de l'énoncé et conseils. La loi est ici définie par sa f.r. dont on peut remarquer qu'elle est dérivable partout. Il s'agit donc d'une loi absolument continue dont on déterminera la densité en dérivant F et le calcul des moments s'effectuera alors comme dans les exercices précédents.

2.11 Exercice : Loi normale

Lecture de tables : loi normale

15. **a)** Si X suit une loi $N(35,5)$, calculer $P(X < 25)$, $P(37{,}5 < X < 40)$ et $P(32{,}5 < X < 37{,}5)$.

b) Calculer l'espérance et la variance d'une v.a. Y de loi normale, telle que $P(Y > -3) = 0{,}691\,5$ et $P(Y < 2) = 0{,}977\,2$.

Analyse de l'énoncé et conseils. La table 1, page 194, fournit les probabilités associées à la loi normale centrée et réduite. Il faut donc dans le cas général centrer la variable normale, c'est-à-dire retrancher son espérance, et la réduire, c'est-à-dire la diviser ensuite par son écart type. Les valeurs de $F(u)$ ne sont fournies que pour $u \geq 0$; la valeur de $F(-u)$ est obtenue en prenant le complément à 1 de celle de $F(u)$. Dans la seconde question, c'est à partir de la table 2, page 195, des fractiles que l'on devrait pouvoir déterminer les valeurs de la variable centrée-réduite associées aux probabilités données. Cependant ici ces probabilités figurent dans la table 1, ce qui permet de déterminer sans interpolation les valeurs exactes de la variable.

2.12 Exercice : Loi exponentielle

Loi exponentielle tronquée

18. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \mathrm{e}^{-(x-\theta)} & \text{si } x > \theta \\ 0 & \text{sinon} \end{cases}$$

où θ est un nombre réel donné.

a) Déterminer la f.r. F et la médiane de cette loi.

b) Soit $X_1, \ldots, X_n$ des v.a. indépendantes et de même loi que X et posons $m_n = \min\{X_1, \ldots, X_n\}$. Déterminer la f.r., puis la densité, de la v.a. m_n.

Analyse de l'énoncé et conseils. La médiane de la loi est le fractile d'ordre $1/2$, c'est-à-dire la valeur Md telle que $F(Md) = 1/2$. Pour déterminer la loi de la plus petite valeur d'un échantillon, c'est-à-dire de v.a. indépendantes et de même loi, il faut exprimer l'événement $(m_n > x)$ à l'aide d'événements associés aux variables X_i, $1 \leq i \leq n$. Ces événements seront indépendants, du fait de l'indépendance de ces v.a.

2.13 Exercice : Loi maximum

Loi du maximum

19. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \frac{1}{36}\left(9 - x^2\right) & \text{si} \quad -3 < x < 3 \\ 0 & \text{sinon} \end{cases}$$

Déterminer la f.r., puis la densité, de la v.a. $M_n = \max\{X_1, \ldots, X_n\}$, où $X_1, \ldots, X_n$ sont des v.a. indépendantes et de même loi que X.

Analyse de l'énoncé et conseils. La f.r. de M_n se détermine après avoir obtenu celle de X et en exprimant l'événement $(M_n < x)$ à l'aide d'événements indépendants associés aux variables indépendantes $X_1, \ldots, X_n$.

3 Statistiques inférentielles

3.1 Exercice : Recettes

Une usine produit deux objets A et B. La quantité d'objets A produite en une journée suit une loi normale de moyenne 2 000 et d'écart-type 500. La quantité d'objet B produite en une journée suit une loi normale de moyenne 1 000 et d'écart-type 300. Un objet A est vendu 30 euros et un objet B est vendu 50 euros. On suppose que tous les objets produits sont vendus.

1) Donner la loi de R représentant la recette réalisée en une journée. Donner sa moyenne et son écart-type.
2) Les dépenses liées à la production engendrent un coût fixe de 80 000 euros quelle que soit la production. Déterminer la probabilité que l'usine perde de l'argent en une journée.
3) On observe les recettes de cette usine sur un mois de 30 jours. Donner un intervalle de confiance à 5 % de la recette moyenne sur ces 30 jours.
4) Expliquer comment évoluerait cet intervalle de confiance si le nombre de jours d'observation augmentait.

3.2 Exercice : Sondage

Lors des dernières élections, le parti A a réalisé 52 % des suffrages. On réalise un sondage sur 1 000 personnes et 49 % de ces personnes affirment vouloir voter pour ce parti. On s'intéresse à cette baisse de pourcentage :

1) Poser les hypothèses d'un test unilatéral.
2) Déterminer la probabilité critique liée à l'observation.
3) Déterminer la valeur critique de proportion au risque 5%.
4) Conclure.

3.3 Exercice : Chiffre d'affaires

Le chiffre d'affaires journalier d'un magasin suit une loi normale de moyenne 1 000 euros et d'écart-type 500 euros. Le magasin change la présentation de sa vitrine et s'interroge sur l'effet de ce changement sur le chiffre d'affaire des 20 prochains jours.

1) Formulez un test d'hypothèse bilatéral sur la moyenne du chiffre d'affaires journalier pour déterminer un intervalle critique en dehors duquel il est possible d'affirmer au risque 1% que le changement de présentation a induit une modification du chiffre d'affaires.
2) On constate un chiffre d'affaires moyen de 1 200 euros sur ces 20 jours. Que conclure ?