

STT-1000 – Probabilités et statistique

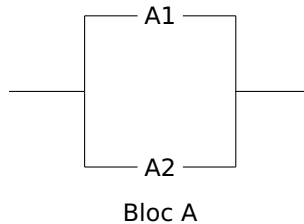
Recueil d'exercices

Le niveau de difficulté de chaque exercice est indiqué par des étoiles, de ★ (le plus facile) à ★★★★★ (le plus difficile). Cette échelle est estimée à partir des résultats obtenus par les étudiants des sessions précédentes; elle donne une indication utile, mais n'est pas une mesure exacte.

Module 1 - Introduction aux probabilités

Exercice

★ Considérons le montage électronique représenté ci-dessous.



Le bloc A contient deux éléments. Nous avons que

- L'élément A1 fonctionne avec une probabilité de 0,70.
- L'élément A2 fonctionne avec une probabilité de 0,70.

Pour que le montage fonctionne, au moins un élément du bloc A doit fonctionner. Les éléments d'un bloc fonctionnent de façon indépendante.

Quelle est la probabilité que le montage ne fonctionne pas ?

☐ 0,49 ☐ 0,70 ☐ 0,03 ☐ 0,91 ☐ 0,09

Exercice

★ Une population est formée à 40% d'enfants et à 60% d'adultes. On sait que parmi les enfants 95% aiment le chocolat, et que parmi les adultes seulement 70% aiment le chocolat. On choisit par hasard une personne parmi celles qui aiment le chocolat. Quelle est la probabilité que celle-ci soit un enfant ? Justifiez vos calculs.

Exercice

★ Soit deux événements A et B formant une partition de l'espace échantillonnal, avec $P(A) = 0,7$.

On a aussi que $P(D | A) = 0,02$, $P(D | B) = 0,05$.

Quelle expression permet de calculer $P(B | D)$?

☐ $\frac{0,3 \times 0,05}{0,02 + 0,05}$

☐ $\frac{0,05}{0,7 \times 0,02 + 0,3 \times 0,05}$

☐ $\frac{0,7 \times 0,02}{0,7 \times 0,02 + 0,3 \times 0,05}$

☐ $\frac{0,3 \times 0,05}{0,7 \times 0,02 + 0,3 \times 0,05}$

☐ $\frac{0,3 \times 0,05}{0,7 \times 0,02 + 0,7 \times 0,05}$

Exercice

★★ Soient A, B et C trois événements avec $P(A) = 0,5$, $P(B) = 0,5$, $P(C) = 0,2$, $P(A \cap B) = 0,25$. Nous avons aussi que C est indépendant de l'événement A et de l'événement B. Dites si chacun des énoncés ci-dessous est vrai ou faux. Votre réponse doit être justifiée en mots ou par des calculs. Un vrai ou faux sans justification n'aura aucun point.

(a) (3 points) Les événements A et B sont indépendants.

☐ VRAI ☐ FAUX

Justification :

(b) (3 points) $P(A \cup C) = 0,6$.

☐ VRAI ☐ FAUX

Justification :

(c) (3 points) Les événements A et B forment une partition de l'espace échantillonnal.

☐ VRAI ☐ FAUX

Justification :

Exercice

★★ Les questions (a) et (b) sont indépendantes.

(a) On nous donne les probabilités suivantes :

$$P(A) = 0,45$$

$$P(B) = 0,55$$

$$P(C) = 0,1$$

$$P(A \cap C) = 0$$

$$P(B \cap C) = 0$$

$$P(A|B) = 0,20$$

i. (3 points) Lequel des énoncés suivants est vrai ?

☐ $P(A \cup B) = 1$

☐ A , B et C sont mutuellement exclusifs.

☐ A et C sont indépendants.

☐ Comme $P(A) + P(B) = 1$, l'intersection de A et B doit être vide, c'est-à-dire $A \cap B = \emptyset$.

☐ A et C sont mutuellement exclusifs.

☐ Aucun des énoncés précédents n'est vrai.

ii. (4 points) A et B sont indépendants.

☐ VRAI ☐ FAUX

Justification :

(b) (3 points) Un certain item manufacturé est visuellement vérifié par deux inspecteurs à la suite. Le premier inspecteur détecte 90% des items défectueux. Parmi les item défectueux qui passent le premier inspecteur sans être détectés, 60% ne seront pas détectés par le deuxième inspecteur. Quelle est la probabilité qu'un item défectueux ne soit repéré par aucun inspecteur ?

☐ 90%

☐ 60%

☐ 10%

☐ 40%

☐ 54%

☐ 6%

☐ 9%

☐ 4%

☐ 36%

☐ 24%

Exercice

★★ Une compagnie d'assurances subdivise ses clients en assurance automobile en trois catégories : les jeunes (16 à 29 ans), les personnes d'âges moyen (30 à 59 ans) et les personnes âgées (60 ans et plus). Selon l'expérience de la compagnie, la probabilité qu'un jeune soit impliqué dans un accident pendant l'année est de 0.12, celle d'une personne d'âge moyen, de 0.06 et celle d'une personne âgée, de 0.10. La clientèle de la compagnie est formée de 30% de jeunes, de 45% de personnes d'âge moyen et de 25% de personnes âgées.

(a) (5 points) Si on choisit un client de cette compagnie au hasard, quelle est la probabilité que ce client ait un accident pendant l'année ?

(b) (5 points) Si on constate que le client choisi au hasard a eu un accident cette année, quelle est la probabilité que ce soit un jeune ?

Exercice

★★ Soit A , B et C trois événements avec $P(A) = 0,5$, $P(B) = 0,5$, $P(C) = 0,2$, $P(A \cap B) = 0,25$. Nous avons aussi que C est indépendant de l'événement A et de l'événement B . Dites si chacun des énoncés ci-dessous est vrai ou faux. Votre réponse doit être justifiée en mots ou par des calculs. Un vrai ou faux sans justification n'aura aucun point.

(a) (3 points) Les événements A et B sont indépendants.

☐ VRAI ☐ FAUX

Justification :

(b) (3 points) $P(A \cup C) = 0,6$.

☐ VRAI ☐ FAUX

Justification :

(c) (3 points) Les événements A et B forment une partition de l'espace échantillonnal.

☐ VRAI ☐ FAUX

Justification :

Exercice

★★

On considère l'espace échantillonnal

$$U = \{a, b, c, d, e, f\}$$

et les ensembles

$$A = \{a, b, c\}, \quad B = \{b, d, e\}, \quad C = \{c, e, f\}.$$

Calculez les ensembles suivants :

(a) (3 points) $A \cup B^c$

(b) (3 points) $(A \cup B) \cap C$

Exercice

★★ Soit deux événements A et B formant une partition de l'espace échantillonnal, avec $P(A) = 0,6$. On sait aussi que $P(D | A) = 0,05$ et $P(D | B) = 0,10$.

Quelle expression permet de calculer $P(A | D)$?

☐ $\frac{0,4 \times 0,10}{0,05 + 0,10}$

☐ $\frac{0,05}{0,6 \times 0,05 + 0,4 \times 0,10}$

☐ $\frac{0,6 \times 0,05}{0,6 \times 0,05 + 0,4 \times 0,10}$

☐ $\frac{0,4 \times 0,10}{0,6 \times 0,05 + 0,4 \times 0,10}$

☐ $\frac{0,6 \times 0,05}{0,6 \times 0,05 + 0,6 \times 0,10}$

Exercice

★★★ Soit A et B , deux événements de S . Nous savons que

$$0 < P(A) + P(B) < 1.$$

Laquelle des affirmations suivantes est tout le temps vraie ?

☐ A et B sont indépendants

☐ A et B forment une partition de l'espace échantillonnal

☐ $P(A \cup B) > 0$

☐ $P(A \cap B) > 0$

☐ A et B sont mutuellement exclusifs

Exercice

★★★ Soit $A \subset B \subset S$ où S est l'ensemble universel.

Laquelle des expressions suivantes est toujours vraie si $P(A) > 0$ et $P(B) > 0$?

☐ $P(A) > P(B)$ ☐ $P(A) + P(B) < 1$ ☐ $P(A|B) \geq P(A)$ ☐ $P(B|A) = 0$

Exercice

★★★ Après un certain nombre de tours pendant une partie de *Scrabble*, les 15 lettres suivantes n'ont toujours pas été choisies et sont donc encore dans le sac :

$\{a, e, i, g, h, l, n, o, p, t, t, v, v, w, x\}$.

- (a) (4 points) Si Alex doit piger 4 lettres au hasard dans le sac, quelle est la probabilité que ces lettres soient les lettres de son prénom, dans le bon ordre ? Laissez des traces de votre démarche.
- (b) (6 points) Supposons qu'Alex a pigé $\{i, g, t, w\}$. Il reste donc $\{a, e, h, l, n, o, p, t, v, v, x\}$ dans le sac. C'est maintenant au tour d'Eva de piger 4 lettres. Quelle est la probabilité que parmi ses 4 lettres, elle pige les lettres qui forment son prénom (peu importe l'ordre) ? Laissez des traces de votre démarche.

Exercice

★★★ Soient A , B et C trois événements de $\mathcal{U}$ tels que

- $\mathbb{P}(A) = 0.4$
- $\mathbb{P}(B) = 0.5$
- $\mathbb{P}(C) = 0.2$
- les événements A et B sont indépendants
- les événements B et C sont indépendants.
- les événements A et C sont mutuellement exclusifs.

Calculer $\mathbb{P}(A \cup B \cup C)$.

Exercice

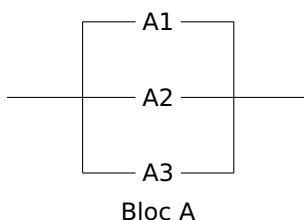
★★★ Une association est composée de 30 personnes. De combien de façons est-ce qu'on peut former un bureau exécutif composé d'un président, un vice-président, un trésorier et 4 membres ?

Exercice

★★★ Une usine fabrique 3 types de télévisions qu'on nomme A , B et C . On nous dit que 50% des télévisions fabriquées sont de type A et 20% de type B . On nous dit aussi que 5% des télévisions de type A , 10% des télévisions de type B et 15% des télévisions de type C sont défectueuses. Quelle est la probabilité qu'une télévision défectueuse soit de type B ?

Exercice

★★★ Considérons le montage électronique représenté ci-dessous.



Le bloc A contient trois éléments. Nous avons que

- L'élément $A1$ fonctionne avec une probabilité de 0,60.
- L'élément $A2$ fonctionne avec une probabilité de 0,50.
- L'élément $A3$ fonctionne avec une probabilité de 0,60.

Pour que le montage fonctionne, au moins un élément du bloc A doit fonctionner. Les éléments d'un bloc fonctionnent de façon indépendante.

Quelle est la probabilité que le montage ne fonctionne pas ?

☐ 0,08 ☐ 0,92 ☐ 0,18 ☐ 0,82 ☐ 0,63

Exercice

★★★ La tuberculose bovine est une maladie très contagieuse. Cette maladie est apparue dans le cheptel bovin d'un pays et touche 10% de ce cheptel. Des études statistiques ont montré que la probabilité pour un animal d'avoir un test positif à cette maladie sachant qu'il est malade est de 95% et que celle d'avoir un test négatif sachant qu'il n'est pas atteint par la maladie est de 90% . Définissons les événements suivants :

T : avoir un test positif à cette maladie pour un bovin choisi au hasard parmi le cheptel.

M : être malade pour un bovin choisi au hasard parmi le cheptel.

- (a) (4 points) Déterminez la probabilité d'avoir un test positif à cette maladie pour un bovin choisi au hasard parmi le cheptel sachant que celui-ci n'est pas malade.

- (b) (4 points) Calculez la probabilité pour un bovin choisi au hasard parmi le cheptel d'avoir un test positif à cette maladie.
- (c) (4 points) Calculez la probabilité qu'un bovin choisi au hasard parmi le cheptel soit malade sachant que le test est positif.

Exercice

★★★ On nous donne les probabilités suivantes :

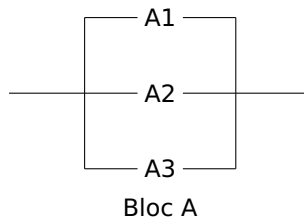
$$\begin{aligned}P(A) &= 0,45 \\P(B) &= 0,25 \\P(C) &= 0,30 \\P(C|A) &= 0,30 \\P(C|B) &= 0,30 \\P(A \cap B) &= 0,05\end{aligned}$$

Cochez tous les énoncés qui sont **vrais**, sachant que deux énoncés sont vrais. *Chaque énoncé qui est coché alors qu'il ne devrait pas l'être annule les points d'un élément coché qui devrait l'être.*

- ☐ $P(A \cup B) = 0,70$
- ☐ $P(A \cap \bar{B}) = 0,40$
- ☐ $P(A \cap C) = 0,1575$
- ☐ $P(A \cup C) = 0,75$
- ☐ A et B sont indépendants
- ☐ A et C sont indépendants
- ☐ Comme B et C sont indépendants, ils sont mutuellement exclusifs

Exercice

★★★ Considérons le montage électronique représenté ci-dessous.



Le bloc A contient trois éléments. Nous avons que

- L'élément A1 fonctionne avec une probabilité de 0,60.
- L'élément A2 fonctionne avec une probabilité de 0,50.
- L'élément A3 fonctionne avec une probabilité de 0,60.

Pour que le montage fonctionne, au moins un élément du bloc A doit fonctionner. Les éléments d'un bloc fonctionnent de façon indépendante.

Quelle est la probabilité que le montage ne fonctionne pas ?

- ☐ 0,08 ☐ 0,92 ☐ 0,18 ☐ 0,82 ☐ 0,63

Exercice

★★★

Un sac contient 40 billes rouges, 40 billes vertes et 20 billes bleues. Sur certaines de ces billes, une étoile a été dessinée. Nous savons que la probabilité de choisir une bille sans étoile parmi les billes rouges est de 0,875, la probabilité de choisir une bille qui est verte et qui a une étoile est de 0,1 et la probabilité de choisir une bille qui a une étoile parmi les billes bleues est de 0,5.

En utilisant les événements suivants :

- R : la bille est rouge
- V : la bille est verte
- B : la bille est bleue

— E : une étoile est dessinée sur la bille

- (a) (3 points) Tracez le diagramme de Venn qui représente cette situation. *Vous n'avez qu'à identifier les événements et l'espace échantillonnal. Il n'est pas nécessaire d'inscrire les probabilités.*
- (b) (8 points) Si vous avez pigé une bille avec une étoile, quelle est la probabilité que celle-ci soit bleue ?
- (c) i. (1 point) Les événements R et V sont indépendants.
☐ Vrai ☐ Faux
- ii. (1 point) Les événements R et E sont indépendants.
☐ Vrai ☐ Faux
- iii. (1 point) Les événements R, V et B forment une partition de l'espace échantillonnal.
☐ Vrai ☐ Faux

Exercice

★★★

On sait que $P(A) = 0,3$, $P(B) = 0,2$ et $P(A^c \cap B) = 0,1$. Calculez $P(A \cap B)$.

Exercice

★★★★

Douze personnes participent à une soirée micro-ouvert, dont Julie. Un programme de la soirée consiste à choisir 5 personnes parmi les 12 et à fixer un ordre de passage pour ces 5 personnes. Une contrainte est imposée : *la première personne à passer doit être l'une des trois personnes désignées comme animateurs*. Julie n'est pas animatrice.

Si on choisit un programme au hasard parmi tous les programmes respectant cette contrainte, quelle est la probabilité que Julie soit choisie et qu'elle passe en troisième position ?

Exercice

★★★★

Un magasin de jouets vend 3 types de figurines : des dragons, des licornes et des phénix. Parmi les figurines vendues, 50% sont des dragons, 30% sont des licornes et 20% sont des phénix. Certaines figurines sont défectueuses. On sait que 10% des dragons sont défectueux, la probabilité de choisir une figurine qui est une licorne et qui est défectueuse est de 0,06, et 25% des phénix sont défectueux.

En utilisant les événements suivants :

- D : la figurine est un dragon
- L : la figurine est une licorne
- P : la figurine est un phénix
- F : la figurine est défectueuse

- (a) (2 points) Tracez le diagramme de Venn qui représente cette situation. *Vous n'avez qu'à identifier les événements et l'espace échantillonnal. Il n'est pas nécessaire d'inscrire les probabilités.*
- (b) (5 points) Si vous avez acheté une figurine défectueuse, quelle est la probabilité que ce soit un phénix ?
- (c) i. (1 point) Les événements D et L sont indépendants.
☐ Vrai ☐ Faux
- ii. (1 point) Les événements D et F sont indépendants.
☐ Vrai ☐ Faux
- iii. (1 point) Les événements D, L et P forment une partition de l'espace échantillonnal.
☐ Vrai ☐ Faux

Exercice

★★★★ Soient A et B deux événements de $\mathcal{U}$ tels que

- $\mathbb{P}(\overline{A \cup B}) = 0.4$
- $\mathbb{P}(A \cap \overline{B}) = 0.2$
- $\mathbb{P}(\overline{A} \cap B) = 0.1$

Calculer $\mathbb{P}(A \cap B)$.

Exercice

★★★★ Une urne contient 25 balles. Parmi ces balles, 5 sont rouges, 5 sont vertes, 5 sont bleues, 5 sont jaunes et 5 sont noires. Pour chacune des couleurs, les balles sont numérotées de 1 à 5.

- (a) (2 points) Combien existe-t-il de sous-ensembles contenant 5 balles ?
- (b) (2 points) Combien existe-t-il de sous-ensembles contenant 5 balles, dont au moins une noire ?
- (c) (3 points) Combien existe-t-il de sous-ensembles contenant 5 balles et exactement 2 balles jaunes ?

Exercice

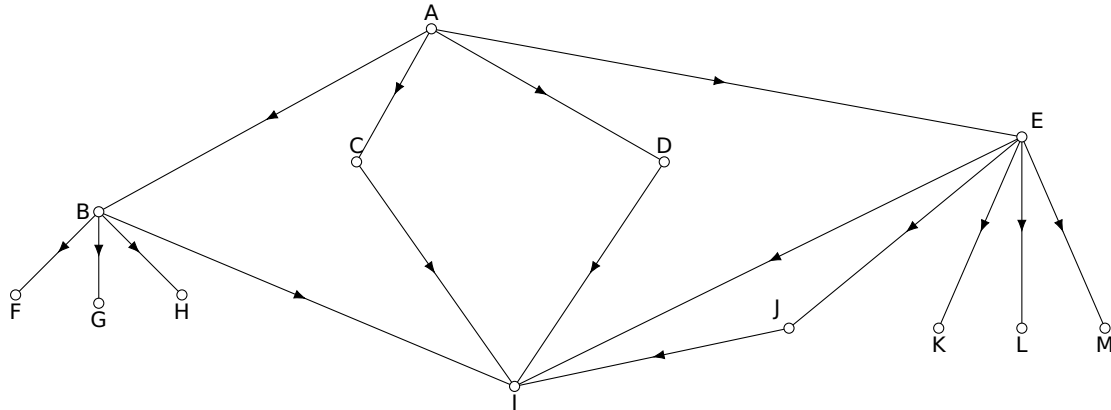
★★★★ Après un certain nombre de tours pendant une partie de *Scrabble*, les 15 lettres suivantes n'ont toujours pas été choisies et sont donc encore dans le sac :

$\{a, e, i, g, h, l, n, o, p, t, t, v, v, w, x\}$.

- (a) (4 points) Si Alex doit piger 4 lettres au hasard dans le sac, quelle est la probabilité que ces lettres soient les lettres de son prénom, dans le bon ordre ? Laissez des traces de votre démarche.
- (b) (6 points) Supposons qu'Alex a pigé $\{i, g, t, w\}$. Il reste donc $\{a, e, h, l, n, o, p, t, v, v, x\}$ dans le sac. C'est maintenant au tour d'Eva de piger 4 lettres. Quelle est la probabilité que parmi les 4 lettres qu'elle doit piger, elle pige les lettres qui forment son prénom (peu importe l'ordre) ? Laissez des traces de votre démarche.

Exercice

★★★★



Une randonneuse part du point A en choisissant au hasard parmi les sentiers AB, AC, AD et AE. Elle s'arrête seulement lorsqu'elle est arrivée à F, G, H, I, K, L ou M. Quelle est la probabilité qu'elle ne s'arrête pas au point M ?

- ☐ 1/7
- ☐ 1/8
- ☐ 1/5
- ☐ 1/20
- ☐ 1/24
- ☐ 1/16
- ☐ 6/7
- ☐ 7/8
- ☐ 4/5
- ☐ 19/20
- ☐ 23/24
- ☐ 15/16
- ☐ 10/11

Exercice

★★★★ Supposons que dans la population, il y a 51% d'hommes (H) et 49% de femmes (F), et que le daltonisme (D) est réparti de cette façon :

	Hommes (H)	Femmes (F)
Daltoniens (D)	0,04	XXX
Non daltoniens ($\bar{D}$)	0,47	YYY

Nous savons que la probabilité qu'une femme choisie au hasard soit daltonienne est de $\frac{1}{245}$.

- (a) (6 points) Quelle est la probabilité qu'une personne choisie au hasard soit une femme et soit non daltonienne ?

(b) (3 points) H et F sont indépendants.

☐ VRAI ☐ FAUX

Justification :

(c) (3 points) H et D sont indépendants.

☐ VRAI ☐ FAUX

Justification :

Exercice

★★★★

Une station de radio possède une banque de 14 chansons.

- 4 chansons sont du même artiste (artiste A) ;
- les 10 autres chansons sont toutes d'artistes différents, et différents de A.

On choisit au hasard 5 chansons distinctes pour une liste de lecture. Quelle est la probabilité que la liste contienne *exactement deux* chansons de l'artiste A ?

Exercice

★★★★★ Une personne a n clés sur son porte-clés. Parmi ces n clés, k clés déverrouillent une même serrure. La personne tente de déverrouiller cette serrure en utilisant une succession de clés, avec remise, c'est-à-dire qu'après avoir tenté avec une clé qui ne fonctionne pas, la clé est remise dans le lot. On note X la variable aléatoire du nombre de clés testées avant que la serrure soit déverrouillée. Que vaut $P(X = k)$?

☐ $\left(\frac{n-k}{n}\right)^{k-1} \frac{k}{n}$ ☐ $\binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k}$ ☐ $\binom{k-1}{n-1} \left(\frac{k}{n}\right)^n \left(1 - \frac{k}{n}\right)^{n-k}$ ☐ $\frac{k}{n}$

Exercice

★★★★★ Une entreprise de tourisme aérien propose des tours en hélicoptère. Elle dispose de 4 hélicoptères, de 4 pilotes et de 8 hôtesses de l'air. Chaque hélicoptère doit être opéré par 1 pilote et accompagné de 2 hôtesses de l'air pour assurer le confort et la sécurité des passagers.

De combien de façons différentes l'entreprise peut-elle attribuer les pilotes et les hôtesses aux hélicoptères ?

Exercice

★★★★★ Soit A et B , deux événements de $\mathcal{U}$, avec $P(A) > 0$ et $P(B) > 0$. Quel énoncé est **toujours** vrai ?

- ☐ Si $P(A) + P(B) = 1$, A et B forment une partition de $\mathcal{U}$.
- ☐ Si $P(A \cap B) = 0$, A et B sont indépendants.
- ☐ Si A et B sont indépendants, ils ne sont pas mutuellement exclusifs.
- ☐ Si $A \subset B$, le fait que B se réalise implique que A se réalise.

Module 2 - Variables aléatoires

Exercice

★★ On lance un dé *non équilibré* dont le résultat est la variable aléatoire X et a la fonction de masse suivante :

$$P_X(x) = \begin{cases} \frac{x}{21} & \text{si } x \in 1, 2, 3, 4, 5, 6, \\ 0 & \text{sinon.} \end{cases}$$

- (a) (2 points) Quelle est la probabilité que le résultat d'un lancer soit un nombre pair ?
- (b) (5 points) Quelle est l'espérance de la variable aléatoire X ? Laissez des traces de votre démarche.
- (c) (3 points) On lance le dé 10 fois. Soit Y la variable aléatoire qui compte le nombre total de «6» obtenus. Quelle est la distribution de Y ? Il faut le nom et tous les paramètres.

Exercice

★★ Soit la variable aléatoire X et sa fonction de répartition :

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ \sqrt{x} & \text{si } 0 \leq x < 1 \\ 1 & \text{si } 1 \leq x. \end{cases}$$

- (a) (5 points) En vous servant de la fonction de répartition, calculez $P(0,25 \leq X \leq 0,81 | X > 0,16)$. Laissez des traces de votre démarche.
- (b) (5 points) Quelle est la fonction de densité $f_X(x)$ de la variable aléatoire X ? Laissez des traces de votre démarche.

Exercice

★★ Laquelle des fonctions suivantes est une fonction de densité ?

$$\square f(a) = a, \text{ pour } 0 < a$$

$$\square f(b) = \begin{cases} 0 & \text{si } b < 0, \\ b & \text{si } 0 \leq b < 2, \\ 0 & \text{si } 2 \leq b \end{cases}$$

$$\square f(c) = \begin{cases} \sqrt{c} & \text{si } 0 \leq c \leq 1, \\ 0 & \text{sinon} \end{cases}$$

$$\square f(d) = \begin{cases} 0 & \text{si } d < 0, \\ d & \text{si } 0 \leq d \leq 1, \\ 1 & \text{si } 1 < d \end{cases}$$

Exercice

★★ Soit X une variable aléatoire ayant une espérance mathématique **nulle** et dont la fonction de masse est donnée par

x	-1	0	1
$p(x)$	a	b	$2b$

(a) (4 points) Montrez que $a = \frac{2}{5}$ et $b = \frac{1}{5}$.

(b) (4 points) Calculez la variance de X .

(c) (4 points) Calculez $E(-3X + 1)$ et $V(-3X + 1)$.

Exercice

★★ Laquelle des fonctions suivantes est une fonction de répartition ?

$$\square F(a) = a, \text{ pour } 0 < a$$

$$\square F(b) = \begin{cases} 0 & \text{si } b < 0, \\ b & \text{si } 0 \leq b < 1, \\ 1 & \text{si } 1 \leq b \end{cases}$$

$$\square F(c) = \begin{cases} \sqrt{c} & \text{si } 0 \leq c \leq 1, \\ 0 & \text{sinon} \end{cases}$$

$$\square F(d) = \begin{cases} 0 & \text{si } d < -1, \\ |d| & \text{si } -1 \leq d \leq 1, \\ 1 & \text{si } 1 < d \end{cases}$$

Exercice

★★★ Soit X une variable aléatoire discrète telle que $\mathbb{R}_X = \{0, 1, 2, 3\}$. Sa fonction de distribution de masse est donnée par :

x	0	1	2	3
$\mathbb{P}(X = x)$	a	$1/6$	b	$1/3$

où a et b sont des inconnus. On nous dit que $\mathbb{E}(X) = 2$. Calculer a et b .

Exercice

★★★ Soit X une variable aléatoire continue dont la fonction de répartition est donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ x^2 & \text{si } 0 \leq x < 1/2 \\ \frac{3x-1}{2} & \text{si } 1/2 \leq x < 1 \\ 1 & \text{si } 1 \leq x \end{cases}$$

(a) (4 points) Calculer $\mathbb{P}(X > 1/2 \mid 1/4 < X < 3/4)$.

(b) (4 points) Donner la fonction de densité f_X de X .

(c) (4 points) En déduire $\mathbb{E}(X)$.

Exercice

★★★ Une composante électrique est reliée à un circuit principal à l'aide de deux branchements. Les observations passées ont permis de dire que le nombre, noté X , de branchements défectueux sur une composante de ce type choisie aléatoirement, possède la distribution suivante :

x_i	0	1	2
$p(x_i)$	0.80	0.19	0.01

(a) (4 points) Quel est le nombre moyen de branchements défectueux par composante ?

- (b) (4 points) Si vous choisissez aléatoirement quatre de ces composantes, quelle est la probabilité que deux d'entre-elles aient exactement un branchement défectueux ?

Exercice

★★★★ On souhaite modéliser la durée de vie d'une composante informatique par une variable aléatoire continue. On observe que plus la composante est vieille plus le risque de bris à court terme est élevé. Quelle loi est la plus appropriée pour modéliser la durée de vie de cette composante ?

- ☐ Une loi exponentielle, car elle modélise bien les pannes aléatoires.
- ☐ Une loi uniforme, car la durée de vie de la composante est bornée.
- ☐ Une loi de Weibull, car son taux de panne peut dépendre de l'âge.
- ☐ Les lois exponentielle et Weibull sont équivalentes dans ce contexte.

Exercice

★★★★ Laquelle des fonctions suivantes est une fonction de masse de probabilité ?

- ☐ $P_X(a) = \begin{cases} a & \text{si } a > 0, \\ 0 & \text{sinon} \end{cases}$
- ☐ $P_X(b) = \begin{cases} 0,5 & \text{si } b \in \{0,1,2\}, \\ 0 & \text{sinon} \end{cases}$
- ☐ $P_X(c) = \begin{cases} 0,5 & \text{si } c = -2, \\ 0,25 & \text{si } c \in \{1,2\} \\ 0 & \text{sinon} \end{cases}$
- ☐ $P_X(d) = \begin{cases} d & \text{si } d \in \{1,2,3,4,5,6\}, \\ 0 & \text{sinon} \end{cases}$

Exercice

★★★★ Soit la variable aléatoire discrète X dont la fonction de probabilité est donnée dans le tableau suivant :

x	1	2	3	5	8	13	21
$p(x)$	0,05	0,1	0,01	0,3	0,5	0,01	0,03

Nous avons $p(x) = 0$ pour tous les x qui ne sont pas donnés dans le tableau.

- (a) (4 points) Remplissez le tableau suivant.

x	0	1	2	3	5	8	13	21	30
$F(x)$									

- (b) (4 points) Calculez l'espérance de X .
- (c) (4 points) Soit W une variable aléatoire qui vaut 2 si le résultat de X est pair et qui vaut 1 si celui-ci est impair. Donnez la fonction de probabilité de W .

Exercice

★★★★ Soit la variable aléatoire X dont la fonction de répartition est donnée par :

$$F(x) = \begin{cases} 0, & x < 0, \\ x^2, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

- (a) (3 points) Quelle est la valeur de m , la médiane de X ? C'est-à-dire, trouvez m , tel que $P(X < m) = 0,5$.
- (b) (4 points) Calculez la probabilité que X soit compris entre 0,2 et 0,3, sachant que X est supérieur à 0,1.

Suite de l'exercice 44 sur la feuille suivante

Suite de l'exercice 44

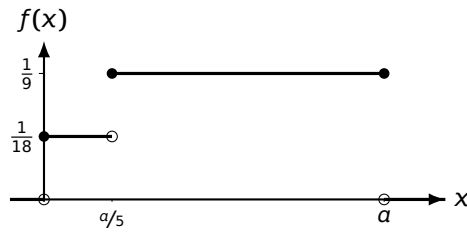
Suite de l'exercice 44

(c) (3 points) Donnez la fonction de densité de X .

Exercice

★★★★ Soit X une variable aléatoire continue de densité f_X définie ci-dessous.

$$f_X(x) = \begin{cases} \frac{1}{18} & \text{si } 0 \leq x < \frac{a}{5}, \\ \frac{1}{9} & \text{si } \frac{a}{5} \leq x \leq a, \\ 0 & \text{sinon.} \end{cases}$$



Utilisez cette densité pour répondre aux parties (a) et (b).

(a) (4 points) Que vaut a ?

(b) (5 points) Quelle est la probabilité que X prenne une valeur entre $\frac{a}{10}$ et $\frac{a}{2}$? Si vous n'avez pas trouvé la valeur de a , vous pouvez donner votre réponse en fonction de a .

Suite de l'exercice 45 sur la feuille suivante

Suite de l'exercice 45

Voici une autre densité. Utilisez-la pour répondre à la partie (c).

$$f_Y(y) = \begin{cases} \frac{1}{180} & \text{si } 0 \leq y < 20, \\ \frac{1}{90} & \text{si } 20 \leq y \leq 100, \\ 0 & \text{sinon.} \end{cases}$$



(c) (5 points) Quelle est l'espérance de Y ?

Exercice

★★★★ Soit X une variable aléatoire continue dont la fonction de répartition est donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ \frac{x^3}{8} & \text{si } 0 \leq x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

- (a) (4 points) Calculez $\mathbb{P}\left(X > 1 \mid X < \frac{3}{2}\right)$.
- (b) (4 points) Trouvez la fonction de densité $f_X(x)$.
- (c) (4 points) Calculez $\mathbb{E}(X)$.

Exercice

★★★★ Soit la fonction f définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} ce^x, & \text{si } x < 0, \\ 0, & \text{sinon,} \end{cases}$$

où $c \in \mathbb{R}$ est une constante donnée.

- (a) (3 points) Quelle doit être la valeur de la constante c pour que f soit une densité de probabilité d'une certaine variable aléatoire X ?
- (b) (3 points) Déterminer la fonction de répartition de X . *Si vous n'avez pas trouvé «c» à la question (a), vous pouvez exprimer la répartition avec «c».*
- (c) (5 points) Calculer l'espérance de X . *Si vous n'avez pas trouvé «c» à la question (a), vous pouvez exprimer l'espérance avec «c».*

Exercice

★★★★ En analyse de survie, nous définissons la fonction de survie $S(x)$ comme étant la probabilité de survivre plus de x unités de temps, c'est-à-dire $S(x) = P(X > x)$, où X est la variable aléatoire qui mesure la durée de vie. Soit la variable aléatoire X , dont la fonction de survie est donnée par

$$S(x) = \begin{cases} e^{-x^2} & \text{si } x \geq 0, \\ 1 & \text{sinon.} \end{cases}$$

- (a) (4 points) Donnez la fonction de répartition de X .
- (b) (2 points) Calculez $P(X \leq \sqrt{2} \mid X \leq \sqrt{3})$.

Module 3 - Lois discrètes

Exercice

★★ Un joueur de fléchette atteint la cible centrale avec une probabilité de 0.05. On suppose que les lancers sont indépendants.

- (a) (4 points) Calculez la probabilité qu'il atteigne la cible pour la première fois au septième lancer.
- (b) (4 points) Calculez la probabilité qu'il atteigne la cible pour la deuxième fois au dixième lancer.

Exercice

★★ Un vendeur d'ordinateurs a constaté que 30% des clients qui entrent dans son magasin, faisaient l'achat d'un ordinateur et que les clients achètent les ordinateurs indépendamment les uns des autres. Définissez clairement les variables aléatoires qui interviennent dans vos calculs pour répondre aux questions suivantes.

- (a) (4 points) Quatre clients sont dans le magasin. Quelle est la probabilité qu'au moins un d'entre eux achète un ordinateur ?
- (b) (4 points) Calculez la probabilité que 4 clients doivent entrer dans le magasin pour que le magasin vende son premier ordinateur ?
- (c) (4 points) Combien de clients, en moyenne, doivent entrer dans le magasin pour que le magasin vende son troisième ordinateur ?

Exercice

★★

Dans un projet de cartographie, un drone envoie automatiquement des relevés de position vers un serveur. Les relevés sont reçus indépendamment les uns des autres, à un taux constant au cours du temps. On suppose que le nombre de relevés reçus suit un processus de Poisson de taux $\lambda = 1,5$ relevé par heure. Charles surveille la réception des relevés.

- (a) (4 points) Charles souhaite aller dîner pendant 40 minutes. Quelle est la probabilité qu'il ne manque *aucun* relevé pendant son absence ?
- (b) (4 points) S'il vient de recevoir un relevé, en moyenne combien de temps a-t-il avant de recevoir le *prochain* relevé ?
- (c) (4 points) Pour rédiger son article, Charles a besoin de 100 relevés. En moyenne, combien de temps faudra-t-il pour obtenir ces 100 relevés ?

Exercice

★★★ Jean est un vendeur dans une boutique et à chaque fois qu'un client se présente, Jean a une chance sur 5 d'effectuer une vente.

- (a) (4 points) Quelle est la probabilité de conclure 2 ventes avec les 5 prochains clients ?
- (b) (4 points) Quelle est la probabilité que le premier client avec qui Jean conclut une vente soit le deuxième ou le troisième client qui s'est présenté à la boutique ?
- (c) (4 points) Quelle est la probabilité de devoir servir 8 clients pour effectuer deux ventes ?

Exercice

★★★ Un enfant accepte de jouer au aux dominos avec son père jusqu'à ce qu'il perde une partie. L'enfant perd 9 parties sur 10, en moyenne. Soit W , le nombre de parties jouées, quelle est la loi de W ? Ici, la loi de Pascal fait référence à la loi binomiale négative.

- ☐ $W \sim \text{Bernoulli}(n = 10, p = 9/10)$
- ☐ $W \sim \text{Bernoulli}(p = 9/10)$
- ☐ $W \sim \text{Binomiale}(n = 10, p = 9/10)$
- ☐ $W \sim \text{Binomiale}(p = 9/10)$
- ☐ $W \sim \text{Géométrique}(n = 10, p = 9/10)$
- ☐ $W \sim \text{Géométrique}(p = 9/10)$
- ☐ $W \sim \text{Poisson}(\lambda = 9/10)$
- ☐ $W \sim \text{Pascal}(r = 10, p = 9/10)$
- ☐ $W \sim \text{Pascal}(p = 9/10)$

Exercice

★★★ On effectue un contrôle de fabrication sur des pièces dont une proportion $p = 0.02$ est défectueuse.

- (a) (4 points) On contrôle un lot de 1000 pièces : Soit X la variable aléatoire : « nombre de pièces défectueuses parmi 1000 ». Quelle est la loi de X ?
- (b) (4 points) En approximant cette loi par celle d'une loi normale, calculez la probabilité que X soit compris entre 18 et 22.

Exercice

★★★

Une voiture est stationnée sans avoir payé. Chaque jour, indépendamment des autres jours, il y a une probabilité $p = 0,15$ qu'elle reçoive une contravention. On suppose qu'au plus une contravention peut être donnée par jour. Le coût d'une contravention est de 65\$. Le coût d'un permis de stationnement mensuel est de 180\$.

- (a) (4 points) Quelle est la probabilité que la voiture ne reçoive *aucune* contravention au cours des 7 premiers jours ?
- (b) (4 points) Après combien de jours s'attend-on, en moyenne, à recevoir la première contravention ?
- (c) (4 points) En moyenne, après combien de jours s'attend-on à ce que le total des contraventions atteigne ou dépasse 180\$?

Exercice

★★★

On peut modéliser le nombre de jours avant la panne d'un serveur informatique par une loi géométrique de paramètre $p = 0,4$. Après 10 jours sans panne, que peut-on dire ?

- ☐ La panne est maintenant plus probable qu'au début.
- ☐ La panne est moins probable qu'au début.
- ☐ La probabilité de panne demain est la même qu'au premier jour.
- ☐ Il est certain qu'une panne surviendra bientôt.

Exercice

★★★ Dites si les énoncés suivants sont vrais ou faux. Aucune justification n'est requise.

(a) (1 point) Si $X \sim \text{Bin}(n, p)$, alors $E(X) = np$ et $\text{Var}(X) = np(1 - p)$.

☐ Vrai ☐ Faux

(b) (1 point) Si $X \sim \text{Exp}(\lambda)$, alors $\text{Var}(X) = \lambda^2$.

☐ Vrai ☐ Faux

Exercice

★★★ Les avions d'une grande compagnie aérienne atterrissent, dans un aéroport très achalandé, selon une loi de Poisson avec une moyenne de six avions à l'heure.

Définir clairement les variables aléatoires qui interviennent dans vos calculs pour répondre aux questions suivantes.

- (a) (4 points) Calculez la probabilité de voir atterrir 2 avions ou plus en 15 minutes.
- (b) (4 points) Calculez la probabilité de devoir attendre plus de 20 minutes pour voir atterrir un avion.
- (c) (4 points) Combien de temps, en moyenne, devriez-vous attendre pour voir atterrir, dans cet aéroport, le cinquième avion de cette compagnie.

Exercice

★★★★

Dans un projet de cartographie, un drone tente d'envoyer des relevés de position vers un serveur. À chaque tentative, le relevé est soit reçu avec succès, soit perdu. On suppose que les tentatives sont indépendantes et que la probabilité qu'un relevé soit reçu avec succès est constante et égale à $p = 0,6$.

Charles surveille la réception des relevés.

- (a) (4 points) Au cours des 10 prochaines tentatives, quelle est la probabilité que *exactement* 7 relevés soient reçus avec succès ?
- (b) (4 points) Quelle est la probabilité que le *premier* relevé reçu avec succès survienne à la 5^e tentative ?
- (c) (4 points) Chaque tentative d'envoi coûte 0,40\$, qu'elle soit réussie ou non. De plus, le lancement du drone entraîne un coût fixe de 10\$. On souhaite obtenir 50 relevés reçus avec succès. On note Z le coût total (en dollars) nécessaire pour atteindre cet objectif. Calculez $\mathbb{E}[Z]$.

Exercice

★★★★ La lune suscite un grand intérêt pour les pêcheurs car la qualité de la pêche dépend directement des différentes phases lunaires. Un pêcheur a remarqué les faits suivants pour chaque période d'une heure de pêche :

- Lorsque la lune est croissante, la probabilité de pêcher au moins un poisson est de cinq sur sept ;
- Lorsque la lune est décroissante, la probabilité de ne pêcher aucun poisson est de deux sur cinq.

La lune est considérée la moitié du temps comme croissante, et l'autre moitié comme décroissante.

- (a) (3 points) Quelle est la probabilité qu'en pêchant une heure, le pêcheur pêche au moins un poisson ?
- (b) (3 points) Si le pêcheur ne pêche aucun poisson après une heure de pêche, quelle est la probabilité que la lune soit croissante ?
- (c) (3 points) Si, en période de lune croissante, il pêche trois heures de suite, quelle est la probabilité qu'il ne pêche aucun poisson ?
- (d) (3 points) Si, en période de lune décroissante, il pêche aussi trois heures de suite, quelle est la probabilité qu'il pêche au moins un poisson ?

Exercice

★★★★ L'examen du code de la route comporte 40 questions. Chaque question propose 4 réponses possibles, dont une seule est correcte. Pour chaque question, le candidat connaît la bonne réponse avec probabilité p . S'il connaît la réponse, il répond correctement avec certitude. S'il ne la connaît pas, il choisit alors une réponse au hasard parmi les 4 possibilités. On suppose que les questions sont indépendantes.

Pour $1 \leq i \leq 40$, on définit l'événement

$$A_i = \{\text{le candidat répond correctement à la } i\text{-ème question}\}.$$

On note S le nombre total de bonnes réponses.

(a) (6 points) Calculez $P(A_i)$ en fonction de p pour $1 \leq i \leq 40$.

(b) (6 points) Déterminez la loi de S (en justifiant votre réponse).

Exercice

★★★★ Une personne a n clés sur son porte-clés. Parmi ces n clés, k clés déverrouillent une même serrure. La personne tente de déverrouiller cette serrure en utilisant une succession de clés, avec remise, c'est-à-dire qu'après avoir tenté avec une clé qui ne fonctionne pas, la clé est remise dans le lot. On note X la variable aléatoire du nombre de clés testées avant que la serrure soit déverrouillée. Que vaut $P(X = k)$?

$$\square \left(\frac{n-k}{n} \right)^{k-1} \frac{k}{n} \quad \square \binom{n}{k} \left(\frac{k}{n} \right)^k \left(1 - \frac{k}{n} \right)^{n-k} \quad \square \binom{k-1}{n-1} \left(\frac{k}{n} \right)^n \left(1 - \frac{k}{n} \right)^{n-k} \quad \square \frac{k}{n}$$

Module 4 - Lois continues, processus de Poisson et loi normale

Exercice

★★★ Un producteur de Vacherin Fribourgeois sait que le poids d'une de ses meules de fromage, X , est une variable aléatoire de loi normale $\mu = 15$ kg et $\sigma^2 = 4$ kg².

(a) (6 points) Quelle est le poids p , tel que $P(X > p) = 0,15866$? Laissez des traces de votre démarche.

(b) (8 points) Le producteur doit envoyer 100 meules à un supermarché. Une compagnie de transport lui charge 1500\$ si le poids des 100 meules est inférieur à 1800 kg ou 2000\$ si le poids est supérieur ou égal à 1800 kg. Quelle est l'espérance du prix que le producteur devra payer pour le transport avec cette compagnie ? Laissez des traces de votre démarche.

Exercice

★★★ Le nombre de tentatives d'infiltration dans un serveur sécurisé se fait selon un processus de Poisson. Le temps moyen entre deux tentatives est 12 minutes.

(a) (5 points) Quelle est la probabilité que le temps entre deux tentatives soit supérieur à 20 minutes. Laissez des traces de votre démarche.

(b) (5 points) Une équipe d'experts tente de répliquer les tentatives d'infiltration dans le but d'améliorer les défenses du serveur. Ils lancent une séquence de tentatives d'infiltration selon le même processus de Poisson. Quelle est l'espérance du temps total pour lancer les 8 tentatives d'infiltration ? Justifiez votre réponse.

Exercice

★★★ Yasmine s'intéresse à une variable aléatoire X suivant une loi Binomiale ($n = 70, p = 0,3$). Elle souhaite calculer $P(18 < X < 35)$, mais n'a pas accès à un ordinateur et n'a pas envie de faire tous les calculs à la main. En utilisant la méthode présentée en classe, calculez approximativement cette probabilité.

Exercice

★★★ Un employé d'un centre d'appel reçoit en moyenne 8 appels par heure. Pour chacune des sous-questions de cet exercice, vous aurez 1 point si vous identifiez la bonne loi et son(s) paramètre(s).

(a) (4 points) Quelle est la probabilité que l'employé reçoive 2 appels dans les 30 prochaines minutes ?

(b) (4 points) L'employé a reçu 3 appels dans les 7 dernières minutes. Quelle est la probabilité que le prochain appel survienne dans 10 minutes ?

(c) (4 points) L'employé a maintenant reçu 18 appels dans les 2 dernières heures. Quelle est l'espérance du temps avant que les 5 prochains appels surviennent ?

Exercice 67 BONUS

Soit la variable aléatoire X dont la fonction de répartition est donnée par

$$F_X(x) = \begin{cases} 1 - e^{-x}, & \text{si } 0 \leq x, \\ 0, & \text{sinon,} \end{cases}$$

et la variable aléatoire $Y = aX + b$, où $a, b \in \mathbb{R}$. Quelle est la fonction de répartition de Y ?

Exercice

★★★ Soit la variable aléatoire X qui suit une loi uniforme sur l'intervalle $[a, b]$. Nous savons que $E(X) = 4$ et $V(X) = 3$.

(a) (5 points) Donnez la fonction de densité de X .

(b) (4 points) Sachant que la fonction de répartition de X est donnée par $F(x) = \frac{x-1}{6}$ pour $x \in [a, b]$ et $F(x) = 0$ sinon, calculez $P(X > 3 | X > 2)$.

Exercice

★★★

On s'intéresse au temps avant d'obtenir une contravention. On suppose que le temps T_1 (en jours) avant la première contravention suit une loi exponentielle dont l'espérance est de 5 jours. Les temps entre les contraventions sont indépendants et de même loi.

Chaque contravention coûte 65\$ et un permis mensuel coûte 180\$.

- (a) (8 points) Considérons un mois de 30 jours. Quelle est la probabilité que le coût total des contraventions dépasse le coût du permis mensuel ?
- (b) (4 points) Soit T_3 le temps (en jours) avant la troisième contravention. Calculez $\mathbb{E}[T_3]$.

Exercice

★★★ Soit $X \sim \mathcal{N}(4, 9)$ et $Y \sim \mathcal{N}(2, 1)$, deux variables aléatoires indépendantes. Quelle est la distribution de $W = 3X - 2Y$?

☐ $\mathcal{N}(8, 10)$

☐ $\mathcal{N}(8, 29)$

☐ $\mathcal{N}(8, 85)$

☐ $\mathcal{N}(16, 85)$

☐ $\mathcal{N}(8, 130)$

Exercice

★★★★ Si X est une variable aléatoire normale de moyenne μ et de variance σ^2 , nous pouvons écrire $X \sim \mathcal{N}(\mu, \sigma^2)$. Laquelle des affirmations suivantes est vraie ?

☐ $Z = \frac{X - \mu}{\sigma^2} \sim \mathcal{N}(0, 1)$

☐ $Y = 3X + 2 \sim \mathcal{N}(3\mu + 2, 9\sigma^2 + 4)$

☐ La fonction de densité de X est symétrique

☐ La fonction de répartition de X est symétrique

Exercice

★★★★

Le poids des tomates d'un agriculteur suit une loi normale de moyenne 300 grammes et de variance 100 grammes². On prend des mesures sur 25 tomates choisies aléatoirement et on obtient une moyenne échantillonnale de 309 grammes et une variance échantillonnale de 118 grammes². Est-ce que la valeur de 118 grammes² est un paramètre ou une statistique ?

☐ Un paramètre ☐ Une statistique

Exercice

★★★★ Un médecin reçoit en moyenne 6 patients par heure.

- (a) (4 points) Quelle est la probabilité que 2 patients rentrent voir le médecin durant les 30 prochaines minutes ?
- (b) (4 points) Vous êtes le premier sur la liste des patients, quelle est la probabilité que vous attendiez plus que 12 minutes avant de voir le médecin ?
- (c) (4 points) Dans combien de temps la cinquième personne sur la liste d'attente peut-elle espérer voir le médecin ?

Exercice

★★★★ Soient X_1 , X_2 et X_3 trois variables aléatoires indépendantes telles que

— $X_1 \sim \mathcal{N}(1, 2)$

— $X_2 \sim \mathcal{N}(0, 3)$

— $X_3 \sim \mathcal{N}(1, 1)$

On pose $Y = X_1 + X_2 + 2X_3$.

- (a) (3 points) Donner la loi de Y .
- (b) (4 points) Calculer $\mathbb{P}(0 \leq Y \leq 6)$.
- (c) (4 points) Trouver m tel que $\mathbb{P}(Y > m) = 0.4$.

Exercice

★★★★ La variable aléatoire X suit une loi normale d'espérance μ et de variance μ^4 . Sachant que $P(X \leq 0) = 0,84134$ et que $P(X > 1) = 0,002275$, que vaut μ_X ?

Exercice

★★★★ Une certaine tablette de chocolat doit peser 200 grammes, avec une tolérance de ± 2 grammes. Elle est donc mise sur le marché si sa masse est comprise entre 198 g et 202 g.

On modélise la masse X (en grammes) d'une tablette par $X \sim \mathcal{N}(\mu, \sigma^2)$, et on suppose que la machine est bien centrée, i.e. $\mu = 200$. Le réglage des machines permet uniquement de modifier la valeur de σ .

(a) (4 points) Calculez la probabilité que la tablette soit mise sur le marché, i.e., $\mathbb{P}(198 \leq X \leq 202)$, si $\sigma^2 = 4$.

(b) (6 points) On souhaite régler la machine de manière à obtenir qu'une tablette choisie au hasard soit mise sur le marché avec une probabilité de $\mathbb{P}(198 \leq X \leq 202) = 0,95$. Déterminez la valeur de l'écart-type σ qui permet d'atteindre cet objectif.

Exercice

★★★★ Lequel des énoncés suivants est faux ?

- ☐ Si $X \sim \exp(\lambda)$, $P(X = 10 | X > 5) = P(X = 10)$.
- ☐ Une succession de variable aléatoires exponentielles est modélisée par la loi Gamma.
- ☐ On doit avoir des variables aléatoires normales pour pouvoir utiliser le théorème central limite.
- ☐ La fonction de densité de X qui suit une loi uniforme est constante sur $\mathbb{R}_X$.

Exercice

★★★★ Une file de clients arrive selon un processus de Poisson de taux 2 par heure. On s'intéresse au temps d'attente jusqu'au prochain client. Quelle est l'expression correcte pour calculer la probabilité que ce temps soit inférieur à 1 heure ?

- ☐ $P(T \leq 1)$, où $T \sim \text{Exp}(2)$
- ☐ $P(T = 1)$, où $T \sim \text{Exp}(2)$
- ☐ $P(X = 1)$, où $X \sim \text{Poisson}(2)$
- ☐ $P(X \leq 1)$, où $X \sim \text{Poisson}(2)$
- ☐ $P(T \leq 1)$, où $T \sim \text{Gamma}(2, 2)$

Exercice

★★★★ Soit $X \sim \mathcal{N}(10, 4)$ et $Y \sim \mathcal{N}(5, 1)$, deux variables aléatoires indépendantes.

On définit $Z = 2X - Y$. Quelle est la distribution de Z ? Soyez précis et laissez des traces de votre démarche.

Exercice

★★★★ Un fabricant de céréales sait que la masse X (en grammes) d'une boîte de céréales suit une loi normale de moyenne $\mu = 500$ g et de variance $\sigma^2 = 225$ g².

(a) (4 points) Quelle est la probabilité qu'une boîte choisie au hasard ait une masse entre 480 g et 520 g ?

(b) (6 points) On prélève un échantillon de 36 boîtes. Quelle est la probabilité que la masse moyenne de ces 36 boîtes dépasse 504 g ?

(c) (8 points) Le coût de production d'une boîte (en dollars) est une variable aléatoire dont on ne connaît pas la loi, mais dont l'espérance est 2\$ et la variance est 0,64\$². Approximez la probabilité que le coût total de production de 100 boîtes dépasse 212\$.

Exercice

★★★★ Un centre d'appels reçoit des appels selon un processus de Poisson de taux 4 par heure. On s'intéresse au temps d'attente jusqu'au prochain appel. Quelle est l'expression correcte pour calculer la probabilité que ce temps soit supérieur à 30 minutes ?

- ☐ $P(T > 0,5)$, où $T \sim \text{Exp}(8)$
- ☐ $P(T > 0,5)$, où $T \sim \text{Exp}(4)$
- ☐ $P(X > 2)$, où $X \sim \text{Poisson}(4)$
- ☐ $P(X > 0,5)$, où $X \sim \text{Poisson}(4)$
- ☐ $P(T > 0,5)$, où $T \sim \text{Gamma}(4, 4)$

Exercice

- ★★★★★ Le passage des voitures sur un certain viaduc est régi par un processus de Poisson d'intensité 60 par heure.
- (a) (7 points) Quelle est la probabilité que la prochaine voiture passe dans plus de 12 minutes, sachant qu'aucune voiture ne passera dans les 6 prochaines minutes ?
- (b) (3 points) Quelle est l'espérance du temps avant que 6 voitures passent sur le viaduc ?

Exercice

★★★★★ Soit trois éléments indépendants assemblés bout à bout. La longueur de chacun présente une distribution normale avec une moyenne de 2 pouces et un écart-type de 0,2 pouce. Or, les assemblages produits doivent avoir une longueur totale de 5,7 à 6,3 pouces. Quelle proportion d'entre eux satisfait à cette exigence ?

Exercice

★★★★★ Le diamètre X (exprimé en cm) des tomates livrées à une usine d'emballage américaine suit une loi normale $\mathcal{N}(7, \sigma^2)$ où l'écart-type σ de X est inconnu. Un tri automatique rejette toutes les tomates dont le diamètre n'est pas compris entre 6 cm et 8 cm.

- (a) (6 points) On constate que 10% des tomates livrées sont rejetées par ce procédé de tri. Calculez l'écart type σ .
- (b) (6 points) Le directeur veut réduire à 5% le pourcentage de tomates rejetées lors du tri. Ne pouvant agir sur les livraisons, il installe un système de tri qui rejette les tomates de diamètre inférieur à $7 - s$ ou supérieur à $7 + s$. Calculez s .

Exercice

★★★★★ On suppose que les requêtes arrivant sur un serveur informatique suivent un processus de Poisson homogène de taux $\lambda = 60$ requêtes par heure.

- (a) (5 points) On note T_1 le temps d'arrivée de la première requête (en minutes). Calculez $\mathbb{P}(1 \leq T_1 \leq 2 | T \geq 1)$.
- (b) (5 points) Quelle est la probabilité qu'aucune requête n'arrive entre 14h00 et 14h05 ?

Exercice

★★★★★ Un producteur de Gruyère sait que le poids d'une de ses meules de fromage, X , est une variable aléatoire qui suit une loi normale d'espérance $\mu = 30$ kg et de variance $\sigma^2 = 9$ kg². Il doit envoyer 144 meules à un festival. Le transporteur lui chargera 1500\$, à condition que le poids moyen des 144 meules soit inférieur ou égal à 30,295 kg. Dans le cas contraire, il lui chargera 2000\$.

- (a) (6 points) Quelle est la probabilité que le transporteur lui charge 1500\$?
- (b) (3 points) Quelle est l'espérance du prix qu'il en coûtera au producteur pour envoyer ses meules ?

Suite de l'exercice 86 sur la feuille suivante

Suite de l'exercice 86

- (c) (6 points) Le producteur sait que la quantité de lait nécessaire pour produire 0,5 kilogramme de Gruyère est d'espérance 6 litres et de variance 4 litres². Quelle est la probabilité qu'il lui faille moins de 415,36 litres pour produire une meule de 32 kilogrammes ?

Module 5 - Loi conjointe

Exercice

★★ Soit les variables aléatoires X et Y . Voici le tableau qui représente la fonction de masse conjointe de (X, Y) où a et b sont des probabilités.

$p(x, y)$		y			
		0	1	2	3
x	-1	a	$10/110$	$14/110$	b
	2	$23/110$	$9/110$	$1/110$	$30/110$
	4	$3/110$	0	$3/110$	$9/110$

On sait que $E[Y|X = -1] = \frac{31}{55}$.

- (a) (3 points) Que valent a et b ? Donnez vos réponses en fractions réduites.

$a =$ $b =$

- (b) (3 points) Est-ce que X et Y sont indépendantes ? Justifier votre réponse.

☐ Oui ☐ Non

- (c) (4 points) Calculez $P(X = -1|Y = 1)$.

Exercice

★★★ On considère deux variables aléatoires discrètes X_1 et X_2 pour lesquelles la fonction de probabilité de masse conjointe est donnée par

$p(x_1, x_2)$		x_2		
		0	2	4
x_1	0	0	$3/15$	$3/15$
	1	$2/15$	$6/15$	0
	2	0	$1/15$	0

- (a) (4 points) Donnez la distribution conditionnelle de X_2 sachant que $X_1 = 1$. Justifiez vos calculs.
- (b) (4 points) Sachant que $E(X_1) = 2/3$ et que $E(X_2) = 2$, calculez la covariance entre X_1 et X_2 . Laissez des traces de votre démarche.
- (c) (4 points) Montrez, sans utiliser le résultat obtenu en (b), que les variables aléatoires X_1 et X_2 ne sont pas indépendantes ?

Exercice

★★★ On lance deux dés standard à 6 faces et on définit les variables aléatoires X et Y comme suit :

- X vaut 1 si les deux dés affichent le même chiffre et 0 sinon.
- Y est le nombre de fois où le chiffre 6 apparaît.

- (a) (4 points) Dresser la fonction de masse conjointe de (X, Y) .
- (b) (4 points) Est-ce que X et Y sont indépendantes ? Justifier votre réponse.
- (c) (4 points) Calculer $E[Y|X = 1]$.

Exercice

★★★ L'examen A et l'examen B servent pour la classification à un même emploi. Nous savons que les notes de ces deux examens suivent des lois normales d'espérance $\mu_A = 1100$ et $\mu_B = 21$, respectivement, de variance $\sigma_A^2 = 40\,000$ et $\sigma_B^2 = 36$, respectivement et de fonctions de répartition F_A et F_B , respectivement. Nous établissons le pointage d'un candidat par la probabilité d'obtenir une note inférieure ou égale à la note du candidat en question (donc en évaluant $F_A(a)$ ou $F_B(b)$, pour un candidat ayant obtenu a à l'examen A ou b à l'examen B).

- (a) (4 points) Sophie a eu une note de 1300 à l'examen A alors que Maxime a eu une note de 24 à l'examen B. Lequel des deux a un meilleur pointage ?
- (b) (4 points) Quelle note aurait dû avoir Sophie afin d'obtenir le même pointage que Maxime ?

Exercice

★★★ Soit X et Y , deux variables aléatoires. Voici le tableau qui représente la fonction de masse conjointe de (X, Y) où les lettres $a, \dots, i$ sont des probabilités.

$p(x, y)$		x			$p_Y(y)$
		0	1	2	
y	-1	$1/16$	a	$1/16$	$4/16$
	0	b	$3/16$	c	d
	1	e	0	f	g
$p_X(x)$		$4/16$	h	i	

Nous avons aussi les deux informations suivantes :

— $P(X + Y = 1) = 6/16$

— $E(Y) = 0$

(a) (9 points) Que valent les probabilités suivantes ? Vous devez justifier vos résultats.

a =

d =

e =

(b) (4 points) Est-ce que X et Y sont indépendantes ? Justifier votre réponse.

(c) (4 points) Calculer $\mathbb{E}[Y|X = 1]$. Si vous n'avez pas trouvé **a**, **d** et/ou **e**, vous pouvez exprimer votre réponse en fonction d'une ou plusieurs de ces inconnues.

Exercice

★★★ Deux informaticiens testent un nouveau service de traitement de requêtes. On définit les variables aléatoires :

X : nombre de requêtes bien traitées dans un échantillon de 3

Y : nombre de requêtes mal traitées dans ce même échantillon.

On suppose que la probabilité qu'une requête soit bien traitée est $p = 0,8$ et que les requêtes sont indépendantes. Comme l'échantillon comporte exactement 3 requêtes, on a $X + Y = 3$.

La loi conjointe de (X, Y) est donnée par

$$p(x, y) = \begin{cases} \binom{3}{x} p^x (1-p)^y & \text{si } x + y = 3, \\ 0 & \text{sinon.} \end{cases}$$

(a) (4 points) Vérifiez que cette loi conjointe est bien une loi de probabilité.

(b) (3 points) Les variables X et Y sont-elles indépendantes ?

Suite de l'exercice 92 sur la feuille suivante

Suite de l'exercice 92

Suite de l'exercice 92

(c) (3 points) Calculez $P(X = 3 | Y = 0)$.

Exercice

★★★ Soit (X, Y) un couple de variables aléatoires discrètes dont la fonction de masse conjointe est donnée par le tableau suivant, où a et b sont des inconnues.

$p(x, y)$		y	
		0	1
x	0	0	a
	1	$1/6$	b
	2	$1/4$	$1/12$

On sait que $\mathbb{E}(X) = 1$.

(a) (6 points) Que valent a et b ?

(b) (4 points) Les variables X et Y sont-elles indépendantes ? Justifiez votre réponse.

(c) (6 points) Calculez $\mathbb{E}[Y | X = 2]$.

Exercice

★★★★ On considère deux variables aléatoires discrètes X_1 et X_2 pour lesquelles la fonction de probabilité de masse conjointe est donnée par

$p(x_1, x_2)$		x_2		
		0	2	4
x_1	0	0	$3/15$	$3/15$
	1	$2/15$	$6/15$	0
	2	0	$1/15$	0

(a) (4 points) Donnez la loi conditionnelle de X_2 sachant que $X_1 = 1$. Justifiez vos calculs.

(b) (4 points) Sachant que $E(X_1) = 2/3$ et que $E(X_2) = 2$, calculez la covariance entre X_1 et X_2 . Laissez des traces de votre démarche.

(c) (4 points) Montrez, sans utiliser le résultat obtenu en (b), que les variables aléatoires X_1 et X_2 ne sont pas indépendantes.

Exercice

★★★★ Soit l'expérience aléatoire qui consiste à lancer deux dés réguliers à 6 faces. On définit les variables aléatoires X et Y qui sont respectivement la somme et la différence absolue des deux dés.

Voici les probabilités conjointes de X et Y .

$p(x,y)$		x											
		2	3	4	5	6	7	8	9	10	11	12	
y	0	$1/36$	0	$1/36$	0	$1/36$	0	$1/36$	0	$1/36$	0	$1/36$	
	1	0	$1/18$	0	$1/18$	0	$1/18$	0	$1/18$	0	$1/18$	0	
	2	0	0	$1/18$	0	$1/18$	0	$1/18$	0	$1/18$	0	0	
	3	0	0	0	$1/18$	0	$1/18$	0	$1/18$	0	0	0	
	4	0	0	0	0	$1/18$	0	$1/18$	0	0	0	0	
	5	0	0	0	0	0	$1/18$	0	0	0	0	0	

(a) (4 points) Quelle est la probabilité que la somme des dés soit inférieure à 5 et que leur différence soit inférieure à 3 ?

(b) (6 points) Quelle est l'espérance de Y , sachant que $X = 6$?

(c) (4 points) Les variables aléatoire X et Y sont indépendantes. Une réponse sans bonne justification n'aura aucun point, même si elle est bonne.

☐ Vrai ☐ Faux

Justification :

Exercice

★★★★ Soit la loi conjointe de X et Y :

	$Y = 0$	$Y = 1$
$X = 0$	0,2	0,3
$X = 1$	0,1	0,4

Quelle est la valeur de $P(Y = 1 | X = 1)$? Laissez des traces de votre démarche.

Module 6 - Statistique descriptive

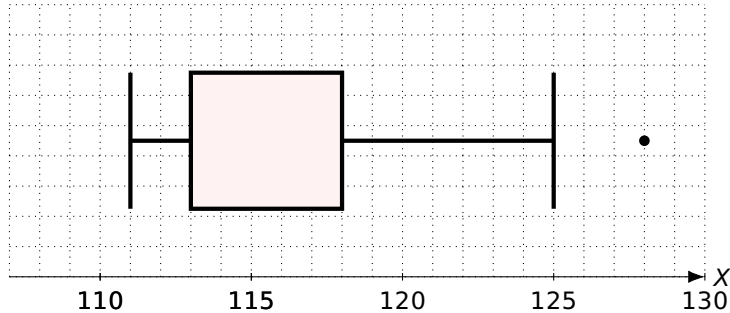
Exercice

★ Dans un laboratoire de microtechnique, Benjamin et Guillaume ont étudié le temps de réaction (en millisecondes) d'un certain type de détecteur de mouvements. Les 21 valeurs suivantes sont des temps de réaction collectés au cours d'une journée :

111	111	111	112	113	113	113
113	115	115	115	115	115	116
116	118	118	119	125	125	128

Le diagramme en boîte ci-dessous a été tracé à partir de ces résultats. Il manque toutefois quelque chose à ce diagramme (il ne s'agit pas d'un titre ou de noms d'axes).

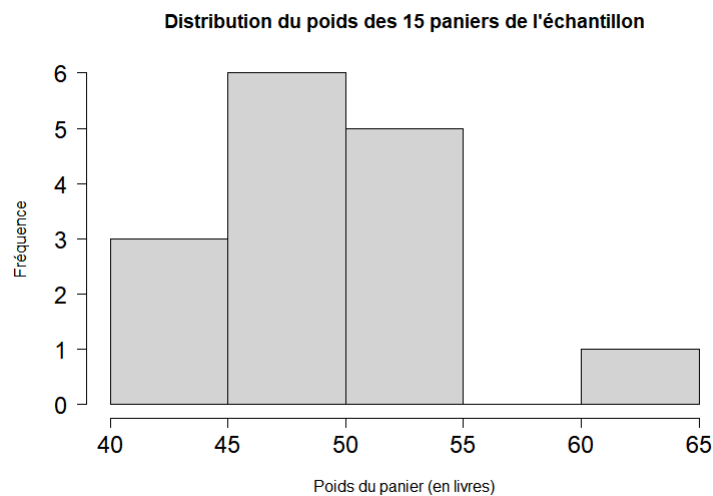
Ajoutez **sur le diagramme** l'élément manquant et dites ce que vous avez fait sous le diagramme.



Exercice

★★ Clara a un verger dans la municipalité de Saint-Coeur-de-Pomme et offre l'autocueillette de pommes. Elle veut savoir si les habitants de la municipalité de Saint-Coeur-de-Pomme aiment la sorte de pomme Vista Bella, qu'elle souhaite introduire dans son verger. Elle décide donc de faire un sondage. Elle choisit par échantillonnage aléatoire simple 100 personnes parmi les clients de son verger durant la fin de semaine du 8 au 10 octobre. Elle leur demande à tous s'ils aiment le goût de la pomme Vista Bella et note leurs réponses.

- (3 points) Est-ce qu'il y a une erreur de couverture ici? Justifiez votre réponse.
- (3 points) Expliquez ce que serait une erreur de mesure dans cet exemple.
- (4 points) Clara veut aussi savoir si ses clients mettent en moyenne 50 livres de pommes dans ses paniers de 50 livres. Pour cela, elle pèse un échantillon de 15 paniers de 50 livres de pommes remplis par ses clients. Les données obtenues sont présentées dans le graphique ci-dessous.



Utilisez le graphique ci-dessus pour estimer le poids moyen des 15 paniers de son échantillon. Laissez des traces de votre démarche.

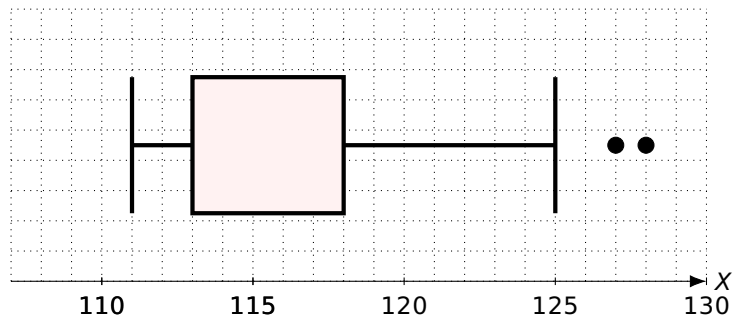
Exercice

★★ Dans un laboratoire de microtechnique, un technicien a étudié le temps de réaction (en millisecondes) d'un certain type de détecteur de mouvements. Les 21 valeurs suivantes sont des temps de réaction collectés au cours d'une journée par ordre croissant :

111	111	111	112	113	113	113
113	114	114	114	114	115	116
118	118	118	119	125	127	128

Le diagramme en boîte ci-dessous a été tracé à partir de ces résultats. Il manque toutefois quelque chose à ce diagramme (il ne s'agit pas d'un titre ou de noms d'axes).

- (a) (6 points) Ajoutez **sur le diagramme** l'élément manquant et dites ce que vous avez fait (ceci *doit* inclure des calculs) sous le diagramme.



Suite de l'exercice 99 sur la feuille suivante

Suite de l'exercice 99

Suite de l'exercice 99

- (b) Vous apprenez que la plus grande observation (128) est erronée. Vous ne connaissez pas sa valeur, mais vous savez que celle-ci est plus grande que 150. Nous notons cette valeur XXX et nous savons que $XXX > 150$. Le nouveau jeu de données est donc :

111	111	111	112	113	113	113
113	114	114	114	114	115	116
118	118	118	119	125	127	XXX

À partir de ce nouveau jeu de données, répondez aux questions suivantes. Nous nommons «jeu de données original» celui qui contient l'observation 128 et «nouveau jeu de données» celui qui contient l'observation XXX.

Inscrivez «Vrai» ou «Faux» dans les cases appropriées.

- i. (3 points) La médiane échantillonnale du nouveau jeu de données est supérieure à celle du jeu de données original.
☐ Vrai ☐ Faux
- ii. (3 points) La moyenne échantillonnale du nouveau jeu de données est supérieure à celle du jeu de données original.
☐ Vrai ☐ Faux
- iii. (3 points) La variance échantillonnale du nouveau jeu de données est supérieure à celle du jeu de données original.
☐ Vrai ☐ Faux

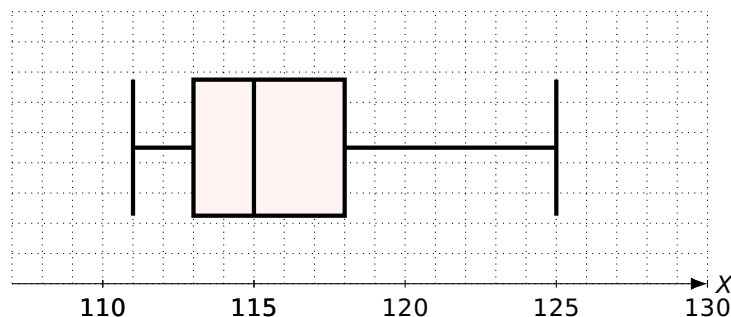
Exercice

★★ Dans un laboratoire de microtechnique, Audrey et Vincent ont étudié le temps de réaction (en millisecondes) d'un certain type de détecteur de mouvements. Les 21 valeurs suivantes sont des temps de réaction collectés au cours d'une journée :

111	111	111	112	113	113	113
113	115	115	115	115	115	116
116	118	118	119	125	125	128

Le diagramme en boîte ci-dessous a été tracé à partir de ces résultats. Il manque toutefois quelque chose à ce diagramme (il ne s'agit pas d'un titre ou de noms d'axes).

- (a) (3 points) Ajoutez **sur le diagramme** l'élément manquant et dites ce que vous avez fait (ceci *doit* inclure des calculs) sous le diagramme. Un ajout sur le diagramme sans justification n'aura aucun point.



Suite de l'exercice 100 sur la feuille suivante

Suite de l'exercice 100 . Suite de l'exercice 100

- (b) (6 points) Nous apprenons que la plus grande observation (128) est erronée. Nous ne connaissons pas sa valeur, mais nous savons que celle-ci est plus grande que 150. Nous notons cette valeur XXX et nous savons que $XXX > 150$. Le nouveau jeu de données est donc :

111	111	111	112	113	113	113
113	115	115	115	115	115	116
116	118	118	119	125	127	XXX

À partir de ce nouveau jeu de données, répondez aux questions suivantes. Nous nommons «jeu de données original» celui qui contient l'observation 128 et «nouveau jeu de données» celui qui contient l'observation XXX. Déterminez si les énoncés suivants sont vrais ou faux.

Une bonne réponse vaut 2 points, une mauvaise réponse vaut -1 point et une réponse vide vaut 0 point, pour un total minimum de 0 point.

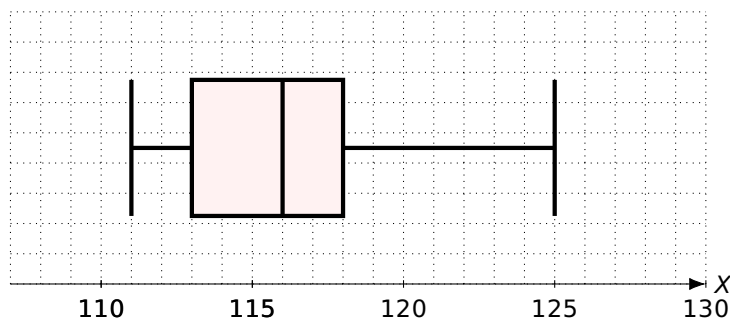
- La médiane échantillonnale du nouveau jeu de données est supérieure à celle du jeu de données original.
☐ Vrai ☐ Faux
- La moyenne échantillonnale du nouveau jeu de données est supérieure à celle du jeu de données original.
☐ Vrai ☐ Faux
- La variance échantillonnale du nouveau jeu de données est supérieure à celle du jeu de données original.
☐ Vrai ☐ Faux

Exercice

★★ Soit les 21 valeurs suivantes :

111	112	112	112	113	113	113
113	115	115	115	115	115	116
118	118	118	119	124	124	125

Le diagramme en boîte ci-dessous a été tracé à partir de ces résultats. Il y a toutefois au moins une erreur (il ne s'agit pas d'un titre ou de noms d'axes).



Quelle est cette erreur ?

- ☐ La médiane n'est pas au bon endroit.
- ☐ La moyenne échantillonnale n'est pas au bon endroit.
- ☐ Il devrait y avoir deux points : un à 105,5 et un à 125,5.
- ☐ La moustache de gauche devrait être à 105,5.
- ☐ La moustache de droite devrait être à 125,5.
- ☐ Il n'y a aucune erreur.

Exercice

★★★ Tracez un diagramme en boîte possible pour une distribution asymétrique à gauche, avec une médiane de 100, un écart interquartile de 20 et une seule donnée extrême. Le graphique n'a pas à être exactement à l'échelle, mais prenez soin d'inclure un axe gradué.

Plusieurs solutions possibles. S'assurer que la donnée extrême est à plus de 30 de la fin de la boîte.

Exercice

★★★ Parmi les statistiques descriptives suivantes, lesquelles sont des mesures de dispersion ? **Une réponse cochée alors qu'elle ne devrait pas l'être entraîne une pénalité de 1 points. Le minimum possible pour cette question est 0 point.**

- ☐ La moyenne échantillonnale
- ☐ La médiane échantillonnale
- ☐ Le mode échantillonnal
- ☐ La variance échantillonnale
- ☐ L'écart interquartile
- ☐ L'étendue

Exercice

★★★ Soit le jeu de données $x_1, \dots, x_n$, où n est un nombre impair. Nous supposons que toutes les valeurs sont différentes. La moyenne échantillonnale de ce jeu de données est $\bar{x}$ et sa médiane est $\tilde{x}$.

Dans chacune des situations suivantes, cochez toutes les réponses vraies.

(a) (2 points) Que pouvons-nous faire pour augmenter $\bar{x}$ sans augmenter $\tilde{x}$?

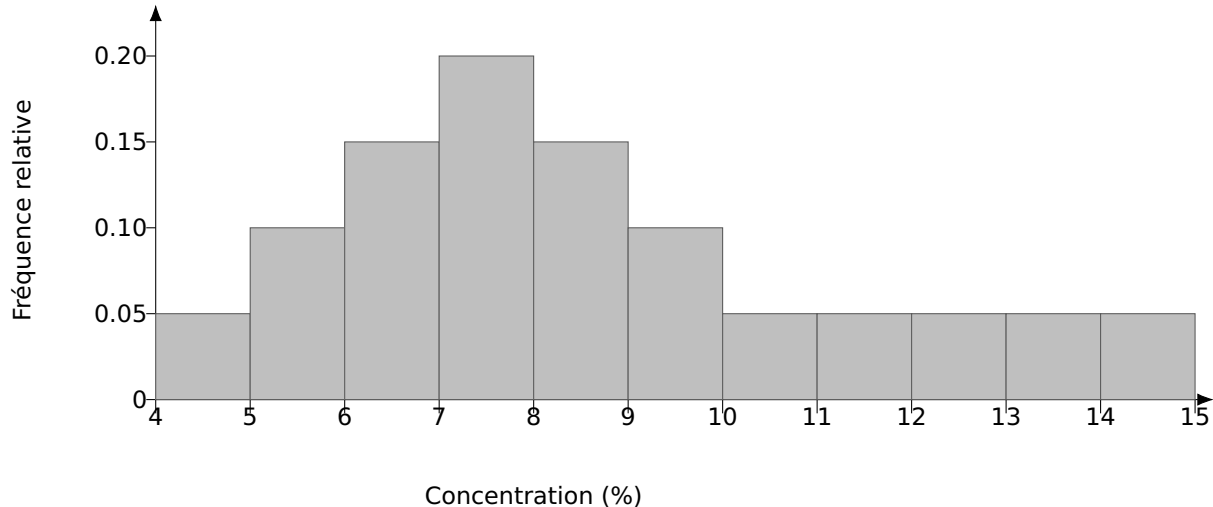
- ☐ Changer la valeur de $x_{(1)}$ pour une valeur supérieure à $x_{(n)}$
- ☐ Changer la valeur de $x_{(n)}$ pour une valeur inférieure à $x_{(1)}$
- ☐ Changer la valeur de $x_{(n)}$ pour une valeur supérieure à $x_{(n)}$
- ☐ Changer la valeur de $x_{(1)}$ pour une valeur supérieure à $x_{(1)}$, mais inférieure à $x_{(0,5n+0,5)}$

(b) (2 points) Que pouvons-nous faire pour augmenter $\bar{x}$ et $\tilde{x}$?

- ☐ Changer la valeur de $x_{(1)}$ pour une valeur supérieure à $x_{(n)}$
- ☐ Changer la valeur de $x_{(n)}$ pour une valeur inférieure à $x_{(1)}$
- ☐ Changer la valeur de $x_{(n)}$ pour une valeur supérieure à $x_{(n)}$
- ☐ Changer la valeur de $x_{(1)}$ pour une valeur supérieure à $x_{(1)}$, mais inférieure à $x_{(0,5n+0,5)}$

Exercice

★★★ Soit l'histogramme suivant :



Déterminez si les énoncés suivants sont vrais ou faux.

Une bonne réponse vaut 1 point, une mauvaise réponse vaut -1 point et une réponse vide vaut 0 point, pour un total minimum de 0 point.

(a) (2 points) Il est impossible de donner la taille de l'échantillon qui a servi à tracer cet histogramme avec l'information donnée.

- ☐ Vrai ☐ Faux

(b) (2 points) Cet histogramme est typique de données provenant d'une distribution où la moyenne est supérieure à la médiane.

- ☐ Vrai ☐ Faux

(c) (2 points) L'écart interquartile vaut au moins 5,5.

- ☐ Vrai ☐ Faux

(d) (2 points) La médiane échantillonnale est entre 7% et 9%, inclusivement.

☐ Vrai ☐ Faux

(e) (2 points) L'écart-type échantillonnal est de 11%.

☐ Vrai ☐ Faux

Exercice

★★★ Dites si les énoncés suivants sont vrais ou faux.

(a) (1 point) Soit $x_1, \dots, x_n$. Si la médiane échantillonnale est 25, nous sommes certains que la valeur de l'observation du milieu est 25.

☐ Vrai ☐ Faux

(b) (1 point) Soit $x_1, \dots, x_n$. Si la moyenne échantillonnale est 25, environ 50% des observations sont plus grandes que 25 et environ 50% sont plus petites que 25.

☐ Vrai ☐ Faux

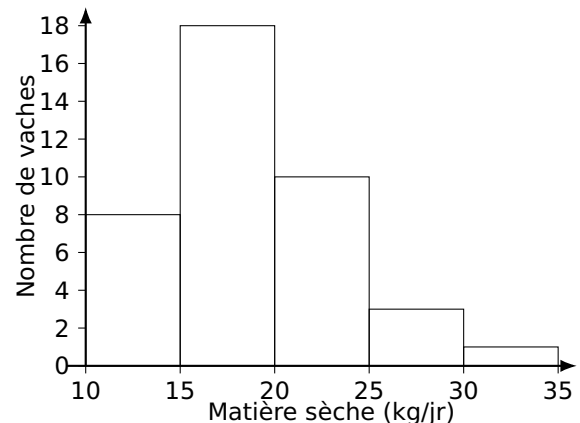
(c) (1 point) Soit $x_1, \dots, x_n$. Si $Q_1 = 12$, nous savons que le minimum de l'échantillon est 12.

☐ Vrai ☐ Faux

(d) (1 point) Soit $x_1, \dots, x_n$. Nous savons que $Q_3 > 0$.

☐ Vrai ☐ Faux

Exercice



★★★ Soit l'histogramme ci-contre.

Dans quelle classe de l'histogramme se situe le quantile d'ordre 0,30 ?

☐ Dans la classe [10,15[

☐ Dans la classe [15,20[

☐ Dans la classe [20,25[

☐ Dans la classe [25,30[

☐ Dans la classe [30,35[

Exercice

★★★ Dites si les énoncés suivants sont vrais ou faux. Aucune justification n'est requise.

(a) (1 point) Soit $x_1, \dots, x_n$. Si la médiane vaut 25, environ 50% des observations sont supérieures à 25 et environ 50% sont inférieures à 25.

☐ Vrai ☐ Faux

(b) (1 point) Soit $x_1, \dots, x_n$. Si on ajoute une constante $c > 0$ à chaque observation, l'écart-type ne change pas.

☐ Vrai ☐ Faux

(c) (1 point) Soit $x_1, \dots, x_n$. L'étendue est une mesure de dispersion résistante à la présence de valeurs aberrantes.

☐ Vrai ☐ Faux

(d) (1 point) Soit $x_1, \dots, x_n$. Si la variance échantillonnale S^2 vaut 0, alors toutes les observations sont identiques.

☐ Vrai ☐ Faux

Exercice

★★★

Supposons le petit échantillon suivant :

2, 4, 4, 6, 9.

Ajoutez une observation à ce jeu de données de façon à diminuer la moyenne sans modifier la médiane. Justifiez brièvement votre réponse.

Exercice

★★★

Vous voulez connaître l'âge des gens présents à une fête de quartier. Un ami questionne tous les individus présents et vous remet les résultats sous la forme suivante :

Classe d'âge	Fréquence
[0,5]	8
[5,10]	15
[10,25]	6
[25,40]	?
[40,100]	10

Le nombre d'individus âgés entre 25 et 40 ans est illisible. Votre ami avait toutefois pris le temps de faire un histogramme des données et vous informe que la hauteur du rectangle pour le groupe des individus âgés de 5 à 10 ans est de 0,06. Déduisez le nombre total de personnes présentes à la fête de quartier.

Module 7 - Estimation et intervalles de confiance

Exercice

★

(a) (6 points) On compare le temps moyen (en millisecondes) nécessaire pour exécuter une même tâche avec deux algorithmes.

- Algorithme A : $n_1 = 62$, $\bar{x}_1 = 72$, $s_1 = 10$
- Algorithme B : $n_2 = 60$, $\bar{x}_2 = 65$, $s_2 = 12$

On suppose que les temps d'exécution suivent des lois normales avec variances inconnues mais supposées égales.

Construisez un intervalle de confiance à 95 % pour la différence des moyennes $\mu_A - \mu_B$.

(b) i. (1 point) L'intervalle à 90% aurait été plus court.

☐ Vrai ☐ Faux

ii. (1 point) Puisque les tailles d'échantillons sont grandes, on aurait aussi pu utiliser l'intervalle qui considère les variances connues.

☐ Vrai ☐ Faux

Exercice

★★

On construit un intervalle de confiance à 95 % pour la variance d'une loi normale et on obtient

$$[4 ; 16].$$

Parmi les affirmations suivantes, cochez toutes celles qui sont correctes.

- ☐ La vraie variance a une probabilité de 0,95 d'être entre 4 et 16.
- ☐ Si on répétait l'échantillonnage, environ 95 % des intervalles construits contiendraient la vraie variance.
- ☐ Une valeur de 12 est plausible pour la variance de la population.
- ☐ La variance observée dans l'échantillon est égale à 10.
- ☐ 95 % des observations sont comprises entre 4 et 16.
- ☐ Un test d'hypothèse bilatéral avec $\alpha = 0.05$ rejetterait l'hypothèse $H_0 : \sigma^2 = 20$.

Exercice

★★★

Jean-Claude suppose que le temps d'attente à l'hôpital suit une exponentielle de moyenne 240 minutes. Il veut construire un intervalle de confiance bilatéral à 95% pour le vérifier. Il procède donc comme suit. Il veut prendre le temps d'attente de 16 personnes choisies au hasard et connaître $\bar{X}$ et S^2 . Il obtient $\bar{x} = 250$ et $s^2 = 59'536$. Comme il construit un intervalle de confiance bilatéral, il choisit $z_{\alpha/2} = 1,96$, à partir de sa table de la loi normale. Son intervalle de confiance à 95% est donc $[130,44; 369,56]$. Comme celui-ci contient 240, il se dit que cette valeur est plausible pour μ . Après avoir vérifié les résultats de Jean-Claude, vous constatez qu'il y a une erreur dans ce qu'il a fait. Quelle est cette erreur ?

Exercice

★★★ Si $X \sim \chi^2_{24}$, alors $P(X < 0) = P(X > 0) = 0.5$.

☐ Vrai ☐ Faux

Exercice

★★★

On a construit un intervalle de confiance pour la proportion de personnes qui préfèrent le chocolat noir au chocolat au lait. On a utilisé un échantillon de 2400 personnes, parmi lesquelles 960 ont répondu qu'elles préféraient le chocolat noir. On a obtenu l'intervalle $[37, 83\%; 42, 17\%]$. Quel est le niveau de confiance de cet intervalle ?

Exercice

★★★

On s'intéresse aux frais d'inscriptions des étudiants de 1er cycle à temps plein au Québec. Avec un échantillon aléatoire de $n = 121$ étudiants de la population, on obtient l'intervalle de confiance à 95% suivant : $[2\,550; 2\,630]$. Cochez tous les énoncés qui sont vrais parmi les suivants. (Aucune justification requise)

- ☐ Il y a 95% des étudiants de 1er cycle à temps plein au Québec qui paient des frais d'inscription entre 2 550 \$ et 2 630 \$
- ☐ Si on pigeait un échantillon de 1 000 étudiants pour construire un nouvel intervalle de confiance à 95% sur la moyenne, celui-ci aurait une plus grande probabilité de contenir la vraie valeur de μ que celui obtenu à partir de 121 étudiants.
- ☐ Si on pigeait un autre échantillon de 121 étudiants pour construire un intervalle de confiance à 95% sur la moyenne, l'intervalle obtenu aurait la même longueur que celui-ci.
- ☐ Si on pigeait plusieurs échantillons de 121 étudiants dans cette population, et qu'on construisait avec chacun d'eux l'intervalle de confiance à 95% pour μ , alors environ 95% de ces intervalles contiendraient la vraie valeur moyenne des frais d'inscription des étudiants de 1er cycle à temps plein au Québec.

Exercice

★★★ On suppose deux populations distribuées selon des lois normales de moyennes μ_1 et μ_2 et de variances σ_1^2 et σ_2^2 inconnues qui sont supposées égales. Vous avez les échantillons

$$X_{11}, \dots, X_{1n_1} \stackrel{i.i.d.}{\sim} N(\mu_1, \sigma_1^2),$$

$$X_{21}, \dots, X_{2n_2} \stackrel{i.i.d.}{\sim} N(\mu_2, \sigma_2^2),$$

où $n_1 = 8$ et $n_2 = 8$. Vous avez construit un intervalle de confiance à 95% pour $\mu_1 - \mu_2$ et obtenu $[0,1; 4,39]$. Lors de la présentation des résultats, vous réalisez que vous avez pris en note que $S_1^2 = 4,52$, mais vous ne trouvez plus la valeur de S_2^2 .

- (a) (8 points) Quelle est la valeur de S_2^2 ?
- (b) (2 points) Que pouvez-vous dire au sujet de $\mu_1 - \mu_2$?

Exercice

★★★ Clara se déplace en autobus pour aller travailler et elle note minutieusement le temps nécessaire pour ses déplacements. Voici ses cinq derniers temps de déplacements, en minutes :

22, 25, 24, 21, 33.

On peut supposer que ses temps de déplacements sont indépendants et suivent une loi normale de moyenne μ et de variance σ^2 .

- (a) (6 points) Donnez un intervalle de confiance à 90% pour la variance de ses temps de déplacement.
- (b) (2 points) Est-ce une bonne idée d'utiliser le point milieu de cet intervalle comme estimateur ponctuel pour σ^2 ? Justifiez votre réponse.
- (c) (2 points) Un collègue vous fait remarquer que la taille d'échantillon est très petite et que votre intervalle de confiance n'est probablement pas valide. Que lui répondez-vous ?

Exercice

★★★ Un intervalle de confiance de niveau 95% pour la moyenne théorique μ est $[5,11; 9,11]$. Pour chaque énoncé qui suit, indiquez si l'énoncé est vrai, faux ou si nous n'avons pas assez d'information pour le savoir.

- (a) (1 point) Au seuil 5%, on rejette $H_0 : \mu = 5$ pour $H_1 : \mu \neq 5$.

☐ Vrai ☐ Faux ☐ Nous n'avons pas assez d'information pour le savoir.

- (b) (1 point) Au seuil 1%, on rejette $H_0 : \mu = 5$ pour $H_1 : \mu \neq 5$.

☐ Vrai ☐ Faux ☐ Nous n'avons pas assez d'information pour le savoir.

(c) (1 point) Au seuil 10%, on rejette $H_0 : \mu = 5$ pour $H_1 : \mu \neq 5$.

☐ Vrai ☐ Faux ☐ Nous n'avons pas assez d'information pour le savoir.

(d) (1 point) Au seuil 5%, on rejette $H_0 : \mu = 8$ pour $H_1 : \mu \neq 8$.

☐ Vrai ☐ Faux ☐ Nous n'avons pas assez d'information pour le savoir.

Exercice

★★★ Aurélien suppose que le temps d'attente dans un bureau de la SAAQ suit une exponentielle de moyenne 240 minutes. Il veut construire un intervalle de confiance bilatéral à 95% pour le vérifier. Il note le temps d'attente de 16 personnes choisies au hasard et obtient $\bar{x} = 250$ et $s^2 = 59\,536$. Comme il construit un intervalle de confiance, il choisit $z_{\alpha/2} = 1,96$, à partir de sa table de la loi normale. Son intervalle de confiance de niveau approximatif 95% est donc $[130,44; 369,56]$. Étant donné que celui-ci contient la valeur 240, il se dit que cette valeur est plausible pour μ . Après avoir vérifié les résultats d'Aurélien, vous constatez qu'il y a une erreur dans ce qu'il a fait. Quelle est cette erreur?

Exercice

★★★ Un intervalle de confiance de niveau 95% pour l'espérance μ est $[5,11; 9,11]$. Pour chaque énoncé qui suit, indiquez si l'énoncé est vrai, faux ou si nous n'avons pas assez d'information pour le savoir.

(a) (1 point) Au seuil 5%, on ne rejette pas $H_0 : \mu = 5$ pour $H_1 : \mu \neq 5$.

☐ Vrai ☐ Faux ☐ Nous n'avons pas assez d'information pour le savoir.

(b) (1 point) Au seuil 1%, on rejette $H_0 : \mu = 4$ pour $H_1 : \mu \neq 4$.

☐ Vrai ☐ Faux ☐ Nous n'avons pas assez d'information pour le savoir.

(c) (1 point) Au seuil 10%, on rejette $H_0 : \mu = 4$ pour $H_1 : \mu \neq 4$.

☐ Vrai ☐ Faux ☐ Nous n'avons pas assez d'information pour le savoir.

(d) (1 point) Au seuil 5%, on rejette $H_0 : \sigma = 1$ pour $H_1 : \sigma \neq 1$.

☐ Vrai ☐ Faux ☐ Nous n'avons pas assez d'information pour le savoir.

Exercice

★★★ Annabelle suppose que le temps d'attente dans un bureau de la SAAQ suit une loi exponentielle de moyenne 240 minutes.

Elle veut construire un intervalle de confiance bilatéral à 95% pour le vérifier. Elle note alors le temps d'attente de 100 personnes choisies au hasard et obtient une moyenne $\bar{x} = 250$, et une variance $s^2 = 59\,536$. À partir de ces résultats, elle construit un intervalle de confiance en utilisant le quantile $z_{0,025}$, à partir de sa table de la loi normale, et obtient l'intervalle $[130,44; 369,56]$.

(a) (2 points) Écrivez les hypothèses du test qu'Annabelle devrait faire à partir de cet intervalle de confiance, en lien avec son interrogation.

(b) (2 points) Quelle serait la décision du test en (a)?

(c) (2 points) Le seuil du test sera-t-il exact? Pourquoi?

Exercice

★★★ Le directeur d'un grand magasin a noté le temps de passage de 18 clients à la caisse rapide lors d'une période d'affluence. Il obtient $\bar{x} = 46$ et $S^2 = 18$. Il suppose que les temps de passage suivent une loi normale d'espérance μ et de variance $\sigma^2 = 18$. Il construit un intervalle de confiance pour μ et obtient $[44,26; 47,74]$.

(a) (4 points) Quel est le niveau de l'intervalle de confiance?

(b) Pour chacun des énoncés suivants, dites s'il est vrai ou faux. Ici, on considère que le niveau trouvé en (a) est égal à $(1 - \alpha)$. Il n'est pas nécessaire de justifier vos réponses.

i. (1 point) La probabilité qu'un client choisi au hasard prenne entre 44,26 et 47,74 secondes est égale à $(1 - \alpha)$.

☐ Vrai ☐ Faux

ii. (1 point) Toutes choses étant égales par ailleurs, si on avait un échantillon de taille 81 au lieu de 18, l'intervalle de confiance serait de niveau $(1 - \alpha^*) > (1 - \alpha)$.

☐ Vrai ☐ Faux

iii. (1 point) Toutes choses étant égales par ailleurs, si $\bar{x}$ était égale à 146, l'intervalle de confiance calculé serait plus long.

☐ Vrai ☐ Faux

Exercice

★★★ Les sous-questions (a) et (b) sont indépendantes.

- (a) (5 points) Soit Z_1, Z_2, Z_3 et Z_4 , des variables aléatoires normales centrées-réduites indépendantes et la variable aléatoire

$$W = Z_1^2 + 2Z_2^2 + 3Z_3^2 - 2Z_4^2.$$

Que vaut $\text{Var}(W)$?

- (b) (5 points) Soit $X_1, \dots, X_9$ des réalisations d'une variable aléatoire normale d'espérance μ et de variance σ^2 .

Que vaut $P(\bar{X} > \mu + 0,625S)$, où S est l'écart-type échantillonnal ?

Exercice

★★★ Le directeur d'un grand magasin a noté le temps de passage de 18 clients à la caisse rapide lors d'une période d'affluence. Il obtient $\bar{x} = 46$ et $S^2 = 18$. Il suppose que les temps de passage suivent une loi normale d'espérance μ et de variance $\sigma^2 = 18$. Il construit un intervalle de confiance pour μ et obtient $[44,26; 47,74]$.

- (a) (4 points) Quel est le niveau de l'intervalle de confiance ?

(b) Pour chacun des énoncés suivants, dites s'il est vrai ou faux. Ici, on considère que le niveau trouvé en (a) est égal à $(1 - \alpha)$, il est donc possible de répondre sans avoir réussi la partie (a). Il n'est pas nécessaire de justifier vos réponses.

- i. (1 point) La probabilité qu'un client choisi au hasard prenne entre 44,26 et 47,74 secondes est égale à $(1 - \alpha)$.

☐ Vrai ☐ Faux

- ii. (1 point) Toutes choses étant égales par ailleurs, si on avait un échantillon de taille 81 au lieu de 18, l'intervalle de confiance serait de niveau $(1 - \alpha^*) > (1 - \alpha)$.

☐ Vrai ☐ Faux

- iii. (1 point) Toutes choses étant égales par ailleurs, si $\bar{x}$ était égale à 146, l'intervalle de confiance calculé serait plus long.

☐ Vrai ☐ Faux

Exercice

★★★

- (a) (7 points) On souhaite estimer la proportion d'élèves du secondaire qui utilisent l'intelligence artificielle pour faire leurs travaux scolaires. On veut construire un intervalle de confiance de niveau 90 % pour cette proportion, avec une marge d'erreur d'au plus 0,04. Quelle taille d'échantillon minimale faut-il utiliser ?

- (b) (2 points) Le directeur de l'école souhaite plutôt une marge d'erreur d'au plus 0,02. Est-ce que la taille d'échantillon proposée en (a) est suffisante pour ce faire ?

☐ Oui, car la taille d'échantillon dépend seulement du niveau de confiance.

☐ Oui, car la marge d'erreur diminue quand la proportion est plus petite.

☐ Non, car le niveau de confiance est trop élevé.

☐ Non, il faudrait augmenter la taille d'échantillon.

Exercice

★★★ Soit $X_1, \dots, X_{17}$ un échantillon aléatoire simple issu d'une distribution $\mathcal{N}(\mu = 5, \sigma^2 = 4)$, et soit S^2 la variance échantillonnale.

Trouvez l'espérance de la variable

$$Y = 4S^2.$$

Exercice

★★★ On a construit un intervalle de confiance pour la proportion de personnes qui préfèrent le chocolat noir au chocolat au lait. On a utilisé un échantillon de 2400 personnes, parmi lesquelles 960 ont répondu qu'elles préféraient le chocolat noir.

On a obtenu l'intervalle

$$[37,83\%; 42,17\%].$$

Quel est le niveau de confiance de cet intervalle ?

Exercice

★★★ Deux entreprises produisent des batteries rechargeables. On s'intéresse à la durée de vie moyenne, en heures, de leurs batteries.

Pour l'entreprise A, on prélève un échantillon de taille $n_A = 36$ et on obtient $\bar{x}_A = 82$. On sait que l'écart-type de la population vaut $\sigma_A = 12$.

Pour l'entreprise B, on prélève un échantillon de taille $n_B = 49$ et on obtient $\bar{x}_B = 76$. On sait que l'écart-type de la population vaut $\sigma_B = 14$.

Construisez un intervalle de confiance à 95% pour la différence des moyennes $\mu_A - \mu_B$.

Exercice

★★★★ Sachant que $\frac{(n-1)S^2}{\sigma^2}$ suit une loi du khi-carré à $n-1$ degrés de liberté, donnez $\text{Var}(S^2)$.

Exercice

★★★★ On souhaite effectuer une étude afin de savoir ce que les gens préfèrent entre la fondue et la raclette. Pour cette étude, on supposera que les réponses des gens $X_1, \dots, X_n$ sont indépendantes et identiquement distribuées selon une Bernoulli(p), où p représente la probabilité de choisir la fondue. Quelle doit être la taille d'échantillon minimale si on désire que notre intervalle de confiance à 90% pour p ait une longueur maximale de 0,04 ?

Exercice

★★★★ Nous nous intéressons aux prix des loyers des étudiants de l'Université Laval. Avec un échantillon aléatoire de $n = 241$ étudiants, nous obtenons l'intervalle de confiance à 95% pour la moyenne μ suivant : $[550; 1630]$. Cochez tous les énoncés qui sont vrais parmi les suivants. **Une réponse cochée alors qu'elle ne devrait pas l'être entraîne une pénalité de 1 point. Le minimum possible pour cette question est 0 point.**

- ☐ Il y a 95% des étudiants qui ont un loyer entre 550 \$ et 1630 \$
- ☐ Si on pigeait un échantillon de 100 étudiants pour construire un nouvel intervalle de confiance à 95% sur la moyenne, celui-ci aurait une plus faible probabilité de contenir la vraie valeur de μ que celui obtenu à partir de 241 étudiants.
- ☐ Si on pigeait un autre échantillon de 241 étudiants pour construire un intervalle de confiance à 95% sur la moyenne, l'intervalle obtenu aurait nécessairement la même longueur que celui-ci.
- ☐ Si on pigeait plusieurs échantillons de 121 étudiants dans cette population, et qu'on construisait avec chacun d'eux l'intervalle de confiance à 95% pour μ , alors environ 95% de ces intervalles contiendraient la vraie valeur moyenne des loyers des étudiants de l'Université Laval.

Exercice

★★★★ Sachant que $\frac{(n-1)S^2}{\sigma^2}$ suit une loi du khi-carré à $n-1$ degrés de liberté, donnez $\text{Var}(S^2)$. **Donnez les détails de votre démarche.**

Exercice

★★★★ Nous supposons deux populations distribuées selon des lois normales de moyennes μ_1 et μ_2 et de variances σ_1^2 et σ_2^2 inconnues qui sont supposées égales. Nous avons les échantillons

$$X_{11}, \dots, X_{1n_1} \stackrel{i.i.d.}{\sim} \mathcal{N}(\mu_1, \sigma_1^2), \quad X_{21}, \dots, X_{2n_2} \stackrel{i.i.d.}{\sim} \mathcal{N}(\mu_2, \sigma_2^2),$$

où $n_1 = 8$ et $n_2 = 8$. Nous avons construit un intervalle de confiance à 95% pour $\mu_1 - \mu_2$ et obtenu $[0,1; 4,39]$. Sachant que $S_1^2 = 4,52$, quelle est la valeur de S_2^2 ? **Donnez les détails de votre démarche.**

Exercice

★★★★ Soit $X_1, \dots, X_{100}$, des variables aléatoires indépendantes et identiquement distribuées selon une loi normale d'espérance μ et de variance $\sigma^2 = 400$. Nous désirons construire un intervalle de confiance de niveau 95% pour μ . Bien sûr, l'intervalle de confiance aura la forme d'un des intervalles vus au cours. Nous notons cet intervalle $[C_1; C_2]$.

- (a) (3 points) Que vaut $\mathbb{E}[C_1]$? **Donnez les détails de votre démarche.**
- (b) (3 points) Que vaut $\text{Var}[C_1]$? **Donnez les détails de votre démarche.**
- (c) (2 points) Quelle est la loi de probabilité de C_1 ?
- (d) (8 points) Quelle est la probabilité que C_1 soit inférieure ou égale à μ ? **Donnez les détails de votre démarche.**

Exercice

★★★★ Soit Z_1, Z_2 et Z_3 , des variables aléatoires normales centrées-réduites indépendantes et $W = Z_1^2 + 2Z_2^2 + 3Z_3^2$. Que vaut $\text{Var}(W)$? **Donnez les détails de votre démarche. Une réponse sans démarche ne donnera aucun point.**

$\text{Var}(W) =$

Exercice

★★★★ Soit $X_1, \dots, X_{100}$, des variables aléatoires indépendantes et identiquement distribuées selon une loi normale d'espérance μ et de variance $\sigma^2 = 400$. Nous désirons construire un intervalle de confiance de niveau 95% pour μ . Bien sûr, l'intervalle de confiance aura la forme d'un des intervalles vus au cours. Nous notons cet intervalle $[C_1; C_2]$.

- (a) (4 points) Que vaut $\mathbb{E}[C_2]$? **Donnez les détails de votre démarche.**
 (b) (4 points) Que vaut $\text{Var}[C_2]$? **Donnez les détails de votre démarche.**
 (c) (2 points) Quelle est la loi de probabilité de C_2 ?
 (d) (6 points) Quelle est la probabilité que C_2 soit inférieure ou égale à μ ? **Donnez les détails de votre démarche.**
Une réponse sans démarche ne donnera aucun point.

Réponse :

Exercice

★★★★ Les sous-questions (a) et (b) sont indépendantes.

- (a) (5 points) Soit Z_1, Z_2, Z_3 et Z_4 , des variables aléatoires normales centrées-réduites indépendantes et la variable aléatoire

$$W = Z_1^2 + 2Z_2^2 + 3Z_3^2 - Z_4^2.$$

Que vaut $\text{Var}(W)$?

- (b) (5 points) Soit $X_1, \dots, X_{16}$ des réalisations d'une variable aléatoire normale d'espérance nulle et de variance égale à 49.

Que vaut $P(\bar{X} > 0,8215\sqrt{S^2})$?

Exercice

★★★★ Soit $X_1, \dots, X_{25}$ provenant d'une loi normale d'espérance $\mu = 0$ et de variance $\sigma^2 = 16$.

- (a) (5 points) Calculez $P(\bar{X} > 0,7495)$.
 (b) (5 points) Quelle est la valeur m , telle que la probabilité que la variance échantillonnale S^2 soit supérieure à m est de 0,05 ?

Module 8 - Tests d'hypothèses

Exercice

★★ Vous travaillez dans une usine qui produit des boulons, et le responsable de la qualité prétend que la variance de la longueur des boulons est de 4 mm². Vous avez des raisons de penser que la variance est en réalité plus élevée et décidez de mener un test d'hypothèse.

Vous prenez un échantillon aléatoire de 20 boulons et mesurez leurs longueurs. La variance des longueurs de l'échantillon est de 5,5 mm².

Effectuez un test d'hypothèse de seuil $\alpha = 1\%$ pour déterminer s'il existe suffisamment de preuves pour rejeter l'affirmation du responsable de la qualité. Vous pouvez supposer que les longueurs des boulons suivent une distribution normale. **Écrivez toutes les étapes du test et justifiez vos réponses !**

Exercice

★★ Un échantillon aléatoire de 800 personnes d'une région indique que 600 d'entre elles sont de la classe moyenne de revenus. Un échantillon de 1 000 personnes prélevé dans une autre région montre que 700 d'entre elles sont de la classes moyenne de revenus.

Tester, au seuil $\alpha = 5\%$, l'hypothèse selon laquelle les proportions de personnes de la classe moyenne de revenus dans les deux régions sont égales ou non.

Exercice

★★ Soit $X_1, \dots, X_{24}$ et $Y_1, \dots, Y_7$ deux échantillons aléatoires provenant respectivement de lois normales d'espérances μ_X et μ_Y et de variance σ_X^2 et σ_Y^2 , supposées inégales. On veut procéder à un test de seuil α avec les hypothèses

$$H_0 : \mu_X - \mu_Y = 0,$$

$$H_1 : \mu_X - \mu_Y > 0.$$

On obtient $\bar{x} = 18,18$, $\bar{y} = 12$, $s_X^2 = 120$ et $s_Y^2 = 28$. Quel est le seuil observé du test ?

Exercice

★★ Soit $X_1, \dots, X_n$ qui proviennent d'une loi normale d'espérance μ et de variance $\sigma^2 = 25$.

On fait le test d'hypothèses pour lequel $H_0 : \mu = 5$ et $H_1 : \mu > 5$ au seuil $\alpha = 2,5\%$.

Pour quelles valeurs de n rejette-t-on H_0 si $\bar{x} = 5,1$?

Exercice

★★ Nous faisons un test pour confronter $H_0 : \mu = 15$ à $H_1 : \mu \neq 15$. La règle de décision est de rejeter H_0 si $\bar{X} < 10$ ou si $\bar{X} > 20$. Dans chacun des cas suivants, déterminez si une erreur est commise et déterminez de quel type, le cas échéant.

- (a) (1 point) La vraie valeur de μ est 16 et nous obtenons $\bar{x} = 12$.
☐ Erreur de type I ☐ Erreur de type II ☐ Aucune erreur
- (b) (1 point) La vraie valeur de μ est 15 et nous obtenons $\bar{x} = 12$.
☐ Erreur de type I ☐ Erreur de type II ☐ Aucune erreur
- (c) (1 point) La vraie valeur de μ est 15 et nous obtenons $\bar{x} = 21$.
☐ Erreur de type I ☐ Erreur de type II ☐ Aucune erreur
- (d) (1 point) La vraie valeur de μ est 18 et nous obtenons $\bar{x} = 21$.
☐ Erreur de type I ☐ Erreur de type II ☐ Aucune erreur

Exercice

★★★ Une entreprise veut connaître la proportion des gens qui utilise leur produit phare afin de tester les hypothèses $H_0 : p = 0,40$ contre $H_1 : p < 0,4$. Elle sélectionne donc un échantillon de $n = 36$ personnes et obtient une estimation de $\hat{p} = 0,385$. En utilisant $\alpha = 0,05$, la compagnie rejettera H_0 si

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} < -Z_\alpha = -1,64.$$

- (a) (8 points) Quelle est la puissance du test si la vraie valeur est $p = 0,395$?
- (b) (2 points) La puissance du test serait-elle plus grande ou plus petite si on avait un échantillon de 25 observations au lieu de 10 observations ? (*Aucune justification requise*)
- ☐ La puissance serait plus grande avec $n = 25$.
☐ La puissance serait plus petite avec $n = 25$.
☐ La puissance serait la même avec $n = 25$ et $n = 10$.
☐ On n'a pas assez d'information pour répondre à la question.
- (c) (2 points) La puissance du test serait-elle plus grande ou plus petite s'il avait décidé de faire un test bilatéral ? (*Aucune justification requise*)
- ☐ La puissance serait plus grande avec un test bilatéral.
☐ La puissance serait plus petite avec un test bilatéral.
☐ La puissance serait la même avec un test bilatéral.
☐ On n'a pas assez d'information pour répondre à la question.

Exercice

★★★ Soit μ le temps moyen que prend une machine pour assembler des composants électroniques. Le responsable du contrôle de processus veut tester $H_0 : \mu = 20$ vs $H_1 : \mu \neq 20$ avec un échantillon aléatoire de 100 temps lors d'une journée donnée. La moyenne de l'échantillon est $\bar{x} = 21,7$ secondes et que la valeur-p du test est de 0,0294. Cochez tous les énoncés qui sont vrais parmi les suivants.

- ☐ Nous sommes certains de rejeter H_0 pour n'importe quel seuil $\alpha > 0$, car la valeur-p est inférieure à 0,05.
☐ Nous rejetons H_0 au seuil $\alpha = 0,01$
☐ Nous rejetons H_0 au seuil $\alpha = 0,1$
☐ Si à la place nous faisons un test unilatéral avec $H_1 > 20$, nous rejetterions H_0 au seuil $\alpha = 0,05$.
☐ Comme $\bar{x} \neq 20$, nous rejetons H_0 peu importe le seuil α .

Exercice

★★★ On s'intéresse au salaire des résidents d'une grande tour à condo. On sait que la distribution de ces salaires est bimodale. Notre hypothèse est que le salaire moyen est de 80000\$ et on veut tester si c'est bien le cas au seuil $\alpha = 0.05$. Un échantillon de 50 résidents nous donne un salaire moyen de 70000\$, avec un écart-type de 28000\$.

- (a) (4 points) Expliquez, en mots, ce que représente une erreur de type II dans le contexte ?
- (b) (8 points) Quel est le seuil observé (la valeur-p) pour ce test d'hypothèse ?

Exercice

★★★ Dans les situations suivantes, serait-il préférable d'utiliser un test de comparaison de moyennes pour des échantillons indépendants ou un test de comparaison de moyennes pour des échantillons appariés.

- (a) (1 point) On veut comparer le niveau en anglais des étudiants des cégeps de la région de Montréal à celui des étudiants des des cégeps de la région de Québec.

- (b) (1 point) On aimerait comparer la performance de deux de vélos de course en mesurant le temps qu'il faut pour monter le mont Bélair avec ces deux vélos.
- (c) (1 point) On veut comparer le temps moyen en secondes que prennent deux marques de four pour atteindre 200 celsius.
- (d) (1 point) On veut connaître la fréquence cardiaque moyenne d'adultes dans la quarantaine avant et après un effort soutenu de 10 minutes.

Exercice

★★★ Parmi les hypothèses suivantes, lesquelles sont valides dans la construction de tests comme on les présente dans le cours ?

- ☐ $H_0 : \mu = 12$ vs $H_1 : \mu > 12$
- ☐ $H_0 : \mu \leq 12$ vs $H_1 : \mu > 12$
- ☐ $H_0 : \bar{x} = 4$ vs $H_1 : \bar{x} \neq 4$
- ☐ $H_0 : \sigma^2 < 1$ vs $H_1 : \sigma^2 > 1$
- ☐ $H_0 : \sigma = 1$ vs $H_1 : \sigma \neq 1$

Exercice

★★★ Interpréter en contexte le résultat d'un test d'hypothèse.

Exercice

★★★ Une usine affirme qu'elle ne pollue pas le lac avoisinant et que la concentration moyenne d'un certain contaminant dans le sol au fond du lac ne dépasse pas 10 ppm. Un biologiste souhaite démontrer le contraire et utilise les hypothèses suivantes :

$$H_0 : \mu = 10$$

$$H_1 : \mu > 10.$$

Les études passées ont démontré que la variance théorique est égale à 16ppm². On peut de plus supposer que la concentration moyenne suit une loi normale.

Le biologiste utilisera 64 prélèvements et calculera $\bar{x}$, la moyenne des concentrations de ces 64 prélèvements.

Les sous-questions (a) et (b) sont indépendantes.

- (a) (6 points) Si le biologiste veut faire un test d'hypothèse au seuil 2,5%, à partir de quelle valeur de $\bar{x}$ rejettera-t-il H_0 ?
- (b) (6 points) Si la règle du biologiste de la compagnie était de rejeter H_0 lorsque $\bar{x} > 11$ et qu'on sait que $\mu = 12$, quelle est la puissance du test ?

Exercice

★★★ On sait par expérience qu'une certaine opération chirurgicale a 90% de chances de réussir. Cette opération est réalisée 400 fois chaque année dans une clinique. Le nombre N de réussites, dans une année, suit la loi binomiale de paramètres $n = 400$ et $p = 0.90$. Utilisez l'approximation normale pour N afin de répondre aux questions suivantes.

- (a) (4 points) Calculez la probabilité que la clinique réussisse au moins 345 opérations dans l'année.
- (b) (4 points) Calculez la probabilité que la clinique rate plus de 28 opérations dans l'année.

Suite de l'exercice 152 sur la feuille suivante

Suite de l'exercice 152

Suite de l'exercice 152

- (c) (4 points) L'assurance accepte de couvrir un certain nombre d'opérations ratées : ce nombre n'a que 1% de chances d'être dépassé. Quel est ce nombre ?

Exercice

★★★ Soit $X \sim \mathcal{N}(\mu; \sigma^2)$ et le test dont les hypothèses sont les suivantes :

$$H_0 : \mu = 10$$

$$H_1 : \mu < 10.$$

On sait que $\sigma^2 = 16$.

On utilise $X_1, \dots, X_{64}$ pour calculer $\bar{X}$.

Les sous-questions (a) et (b) sont indépendantes.

(a) (6 points) Si le seuil du test est de 2,5%, pour quelles valeurs de $\bar{x}$ rejette-t-on H_0 ?

(b) (6 points) Si la règle était de rejeter H_0 lorsque $\bar{x} < 8,5$ et qu'on sait que $\mu = 7$, quelle est la puissance du test ?

Exercice

★★★ Je veux tester les hypothèses suivantes en pigeant un échantillon de taille $n = 25$ dans une population normale de variance inconnue :

$$H_0 : \mu = 21 \quad \text{contre} \quad H_1 : \mu < 21.$$

J'obtiens un seuil observé de 0,8413.

Quel serait le seuil observé pour le test bilatéral avec la même hypothèse nulle ?

Exercice

★★★ Deux chaînes d'assemblage produisent des composantes électroniques. On veut vérifier si la chaîne A produit en moyenne des composantes plus lourdes que la chaîne B.

On prélève deux échantillons indépendants et on obtient les résultats suivants :

- Chaîne A : $n_A = 64$, $\bar{x}_A = 503$ mg, $\sigma_A = 8$ mg ;
- Chaîne B : $n_B = 100$, $\bar{x}_B = 500$ mg, $\sigma_B = 10$ mg.

On suppose que les deux populations sont normales et que les variances sont connues.

Au seuil $\alpha = 5\%$, testez

$$H_0 : \mu_A - \mu_B = 0 \quad \text{contre} \quad H_1 : \mu_A - \mu_B > 0.$$

Exercice

★★★

Caroline a trouvé une pièce de monnaie. Son intuition lui dit que la pièce tombe plus souvent sur le côté PILE que sur le côté FACE. Elle souhaite vérifier cette hypothèse et lance donc la pièce à 64 reprises dans les airs en notant le résultat obtenu.

- (a) (9 points) Au seuil de 5%, quelle est la puissance de son test statistique si, en vérité, la pièce de monnaie a une probabilité de 70% de tomber sur le côté PILE ?
- (b) (3 points) Caroline avait aussi pensé utiliser un seuil de 10%. Cochez la bonne réponse ci-dessous.

- ☐ La puissance aurait été plus petite avec un seuil de 10%.
- ☐ La puissance aurait été plus grande avec un seuil de 10%.
- ☐ La puissance aurait été identique avec un seuil de 10%.
- ☐ On ne peut pas savoir avec l'information donnée si la puissance aurait été plus grande, plus petite ou identique avec un seuil de 10%.

Exercice

★★★ On veut vérifier si un programme d'entraînement physique de 6 semaines réduit la pression artérielle systolique (en mmHg) d'adultes sédentaires. Six participants sont sélectionnés au hasard et leur pression artérielle est mesurée avant et après le programme.

Mesure	Participant						Moyenne
	1	2	3	4	5	6	
Avant	152	145	163	148	155	149	152
Après	146	140	155	143	151	147	147

On pose $d_i = \text{Avant}_i - \text{Après}_i$ et on vous informe que :

$$s_D^2 = 4 \quad \text{et} \quad \bar{d} = 5.$$

- (a) (5 points) Au seuil $\alpha = 5\%$, testez si le programme réduit la pression artérielle en moyenne.
(b) (2 points) Pourquoi est-il probablement préférable d'utiliser des échantillons appariés dans ce contexte? Une seule réponse est bonne.

- ☐ Car tous les participants ont naturellement la même pression artérielle de base.
☐ Car la mesure après le programme est plus précise que la mesure avant.
☐ Car nous voulons nous protéger contre la possibilité que les différences naturelles de pression artérielle entre les participants influencent la décision du test.
☐ Car nous rejetons H_0 au seuil $\alpha = 5\%$.
☐ Car il s'agit d'observations prises dans le temps.

Exercice

★★★★ Pourquoi est-il indiqué sur la feuille de formules que la loi de la statistique de test pour une proportion est approximative?

Exercice

★★★★ Parmi les hypothèses suivantes, **lesquelles** sont valides dans la construction de tests comme on les présente dans le cours? **Une réponse cochée alors qu'elle ne devrait pas l'être entraîne une pénalité de 1 points. Le minimum possible pour cette question est 0 point.**

- ☐ $H_0 : \mu = 12$ vs $H_1 : \mu > 12$
☐ $H_0 : \mu \leq 12$ vs $H_1 : \mu > 12$
☐ $H_0 : \bar{x} = 4$ vs $H_1 : \bar{x} \neq 4$
☐ $H_0 : \mu = 1$ vs $H_1 : \mu \neq 1$

Exercice

★★★★ Soit $X_1, \dots, X_{36}$, des variables aléatoires indépendantes et identiquement distribuées selon une loi normale d'espérance μ et de variance $\sigma^2 = 1296$. Nous désirons effectuer un **test unilatéral à droite** avec $H_0 : \mu = 10$ au seuil $\alpha = 0,025$. Nous avons les résultats suivants :

$$\bar{x} = 11,34 \quad s^2 = 754.$$

Si nous savons qu'en réalité $\mu = \mu_1$ et que la puissance du test est 0,13567, que vaut μ_1 ? **Donnez les détails de votre démarche.**

Exercice

★★★★ Des chercheurs veulent comparer les traitements A et B contre une maladie. Ils administrent le traitement A à un groupe de 45 personnes malades et le traitement B à un groupe de 36 autres personnes malades, puis ils comptent le nombre de guérisons dans chacun des groupes. Les chercheurs pensent que le traitement A est plus efficace que le traitement B et veulent vérifier cette affirmation en faisant un test avec les hypothèses suivantes :

$$H_0 : p_A = p_B, \\ H_1 : p_A > p_B,$$

où p_A et p_B représentent la probabilité de guérison dans les groupes A et B respectivement. S'il y a eu 18 guérisons dans le groupe B, quel doit être le nombre minimal de guérisons dans le groupe A pour que H_0 soit rejetée au seuil approximatif 15,866%? **Donnez les détails de votre démarche.**

$$\text{Coup de pouce : } \hat{p}(1 - \hat{p}) = \frac{-25\hat{p}_A^2 + 25\hat{p}_A + 14}{81}.$$

Exercice

★★★★ Supposons que les variables aléatoires $X_1, \dots, X_{20}$ suivent une loi exponentielle décalée de paramètre $\theta \geq 0$, dont la **fonction de densité** est donnée par

$$f_X(x) = \begin{cases} e^{-(x-\theta)}, & \text{si } \theta \leq x, \\ 0, & \text{sinon.} \end{cases}$$

Nous aimerions faire un test statistique pour savoir si $\theta = 0$. Nous avons donc les hypothèses
La statistique du test est

$$M = \min(X_1, \dots, X_{20}),$$

$$H_0 : \theta = 0 \quad H_1 : \theta > 0.$$

qui correspond à la valeur minimale de notre échantillon (nous pouvons aussi écrire $M = X_{(1)}$, dans la notation introduite au cours). Nous savons que sous H_0 , M suit une loi exponentielle de paramètre $n = 20$. Sa **fonction de répartition** est donnée par

$$F_M(m) = \begin{cases} 1 - e^{-20m}, & \text{si } 0 \leq m \\ 0, & \text{sinon.} \end{cases}$$

La table suivante donne quelques valeurs de cette fonction de répartition.

m	$F_M(m)$
0,0005	0,01
0,0025	0,05
0,0053	0,10
0,1151	0,90
0,1498	0,95
0,2303	0,99

Nous avons l'échantillon $x_1, \dots, x_{20}$ suivant :

0,1613	0,1933	0,2980	0,3331	0,3423	0,4392	0,4472	0,4580	0,5289	0,5530
0,5570	0,7737	0,8574	0,9376	1,2509	1,4475	1,9064	2,6019	2,8685	2,9829

- (a) (3 points) Au seuil $\alpha = 5\%$, quelle est la **décision** du test. **Une décision sans justification n'obtiendra aucun point, même si la décision est bonne.**
- (b) (3 points) **Calculez** la valeur-p du test en ne gardant que 4 chiffres après la virgule. **Donnez les détails de votre démarche.**

Exercice

★★★★ Soit les échantillons

$$Y_1, \dots, Y_n \stackrel{i.i.d.}{\sim} \mathcal{N}(\mu_Y, \sigma_Y^2),$$

$$X_1, \dots, X_m \stackrel{i.i.d.}{\sim} \mathcal{N}(\mu_X, \sigma_X^2),$$

où $n = 8$, $m = 8$ et $\sigma_Y^2 = \sigma_X^2$ sont inconnues. Pour le test dont les hypothèses sont

$$H_0 : \mu_Y - \mu_X = 0,$$

$$H_1 : \mu_Y - \mu_X > 0,$$

la valeur-p est de 0,01. Sachant que $\bar{y} = 6,824$, $\bar{x} = 4,2$ et $s_Y^2 = 4,5$, quelle est la valeur de s_X^2 ? **Donnez les détails de votre démarche. Une réponse sans démarche ne donnera aucun point.**

$$s_X^2 = \boxed{}$$

Exercice

★★★★ Une entreprise fabrique des barres de chocolat et s'intéresse à la variabilité de la quantité de sucre dans les barres. Elle choisit au hasard 11 barres. Dans l'échantillon, la quantité moyenne de sucre par barre est de 20,05 g et la variance échantillonnale est de 7,995 g². Sachant que X , la quantité de sucre dans une barre, est une variable aléatoire qui suit une loi normale d'espérance μ et de variance σ^2 , elle fait le test suivant, au seuil $\alpha = 2,5\%$,

$$H_0 : \sigma^2 = 5$$

$$H_1 : \sigma^2 > 5.$$

- (a) (3 points) À partir de quelle valeur de S^2 rejette-t-on H_0 ?
☐ 1,960 ☐ 2,131 ☐ 10,24 ☐ 15,99 ☐ 20,48
- (b) (2 points) Quelle est la valeur de la statistique du test sous H_0 , arrondie à 3 décimales ?
☐ 1,960 ☐ 2,131 ☐ 10,24 ☐ 15,99 ☐ 20,48
- (c) (2 points) Quel est le seuil observé du test ?
☐ $\alpha^* = 0,01$.

- ☐ $\alpha^* = 0,025.$
- ☐ $\alpha^* = 0,05 .$
- ☐ $\alpha^* = 0,1.$

(d) (3 points) Si en réalité $\sigma^2 = 6,4$ et que la règle de décision était plutôt de rejeter H_0 si $\frac{10S^2}{5} > 25,6$, quelle est l'expression de la puissance du test ?

- ☐ $P\left(\frac{10S^2}{6,4} > 25,6\right)$
- ☐ $P\left(\frac{10S^2}{5} > 25,6\right)$
- ☐ $P\left(\frac{10S^2}{6,4} > 20\right)$
- ☐ $P\left(\frac{10S^2}{5} > 20\right)$

Exercice

★★★★ Soit une loi normale d'espérance μ et de variance $\sigma^2 = 64$. Nous voulons faire un test d'hypothèses au seuil $\alpha = 0,05$ pour confronter

$$H_0 : \mu = 600,$$

$$H_1 : \mu \neq 600.$$

Les valeurs critiques du test sont donc $-z_{0,025} = -1,96$ et $z_{0,025} = 1,96$. Avec un échantillon aléatoire de $n = 241$ étudiants, nous obtenons un seuil observé de 0,004. Cochez tous les énoncés qui sont vrais parmi les suivants. **Une réponse cochée alors qu'elle ne devrait pas l'être entraîne une pénalité de 1 point. Le minimum possible pour cette question est 0 point.**

- ☐ Nous savons que $|z_0| > 1,96$.
- ☐ Nous savons que $|z_0| < 1,96$.
- ☐ Nous rejetons H_0 .
- ☐ Nous ne rejetons pas H_0 .
- ☐ Nous sommes certains que l'intervalle de confiance à 95% pour μ contient 600.
- ☐ Nous sommes certains que l'intervalle de confiance à 95% pour μ ne contient pas 600.
- ☐ Nous sommes certains que l'intervalle de confiance à 99% pour μ contient 600.
- ☐ Nous sommes certains que l'intervalle de confiance à 99% pour μ ne contient pas 600.
- ☐ Nous ne pouvons rien dire sur ce test car nous ne connaissons pas la statistique observée, z_0 .

Exercice

★★★★ Soit $X_1, \dots, X_{25}$, des variables aléatoires indépendantes et identiquement distribuées selon une loi normale d'espérance μ et de variance $\sigma^2 = 625$. Nous désirons effectuer un **test unilatéral à droite** avec $H_0 : \mu = 9$ au seuil $\alpha = 0,025$. Nous avons les résultats suivants :

$$\bar{x} = 11,34$$

$$s^2 = 754.$$

Si nous supposons qu'en réalité la vraie moyenne $\mu = \mu_1$ et que la puissance du test est 0,13567, que vaut μ_1 ? **Donnez les détails de votre démarche. Une réponse sans démarche ne donnera aucun point.**

$\mu_1 =$

Exercice

★★★★ On demande à 3600 Suisses s'ils préfèrent la fondue ou la raclette. On obtient que 1849 personnes préfèrent la fondue. Y a-t-il suffisamment de preuves pour dire que la majorité des Suisses préfèrent la fondue, au seuil approximatif $\alpha = 0,05$? **Écrivez toutes les étapes du test et justifiez vos réponses !**

Utilisez l'approximation suivante : $\sqrt{\frac{\frac{1849}{3600} \left(1 - \frac{1849}{3600}\right)}{3600}} \approx 0,008$

Module 9 - Régression linéaire simple

Exercice

★★★

Afin d'estimer le volume de bois disponible dans la forêt qui mène à La Barillette, nous aimerions modéliser le volume de bois d'un arbre à l'aide de sa circonférence. Nous disposons du volume (Y) et de la circonférence (X) de 31 arbres. Afin d'étudier le modèle $Y_i = \beta_0 + \beta_1 X_i + E_i$, en supposant que les postulats de régression linéaire sont respectés, nous avons résultats suivants :

$$\sum_{i=1}^{31} x_i = 930; \quad \sum_{i=1}^{31} x_i^2 = 28\,900; \quad \sum_{i=1}^{31} x_i y_i = 113\,880; \quad \sum_{i=1}^{31} y_i = 3\,596; \quad \sum_{i=1}^{31} y_i^2 = 446\,716.$$

- (a) (8 points) Donnez les estimations de β_0 et β_1 . **Donnez les détails de votre démarche. Des réponses sans démarche ne donneront aucun point.**

$$\hat{\beta}_0 = \boxed{} \quad \hat{\beta}_1 = \boxed{}$$

Suite de l'exercice 168 sur la feuille suivante

Suite de l'exercice 168

Suite de l'exercice 168

- (b) (2 points) Sophie aimerait estimer le volume de bois de l'arbre devant sa maison. Cet arbre a une circonférence de 28. La prévision ponctuelle du volume de bois de cet arbre est de 104. Choisissez l'intervalle de valeurs ayant 95% de chances de contenir le volume de bois de cet arbre de circonférence 28.

- ☐ $104 \pm t_{0,05;29} \sqrt{MS_E \left[\frac{1}{n} + \frac{(28 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $104 \pm t_{0,025;1} \sqrt{MS_E \left[\frac{1}{n} + \frac{(28 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $104 \pm t_{0,025;29} \sqrt{MS_E \left[\frac{1}{n} + \frac{(28 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $104 \pm t_{0,025;29} \sqrt{MS_E \left[1 + \frac{1}{n} + \frac{(28 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $104 \pm t_{0,025;1} \sqrt{MS_E \left[1 + \frac{1}{n} + \frac{(28 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $104 \pm t_{0,05;29} \sqrt{MS_E \left[1 + \frac{1}{n} + \frac{(28 - \bar{x})^2}{S_{XX}} \right]}$

- (c) (1 point) Si Sophie avait un arbre ayant une circonférence de 35, l'intervalle de valeurs calculé à la question précédente serait...

- ☐ plus court
- ☐ de même longueur
- ☐ plus long
- ☐ on ne peut pas savoir la différence de longueur

- (d) (1 point) La valeur-p du test de significativité du lien linéaire entre X et Y est égale à 0,00014. Cochez l'énoncé qui est vrai.

- ☐ Il n'y a pas de lien linéaire significatif entre X et Y.
- ☐ Il y a un lien linéaire significatif entre X et Y.
- ☐ Nous ne pouvons rien dire car nous ne connaissons pas le seuil du test.
- ☐ Nous ne pouvons rien dire car nous ne connaissons pas t_0 .

Exercice

★★★

On a ajusté un modèle de régression linéaire pour le prix de vente d'une maison, en milliers de dollars, en fonction de l'évaluation de cette maison, aussi en milliers de dollars, à l'aide d'un jeu de données de 92 observations. On a obtenu les valeurs suivantes :

$$\hat{\beta}_0 = 20,942, \quad \hat{\beta}_1 = 1,06873, \quad \hat{\sigma}^2 = 1075.$$

De plus, on vous indique que

$$\bar{x} = 201,7504 \quad \text{et} \quad S_{XX} = 5209571.$$

- (a) (3 points) Interprétez la pente estimée, en contexte.
- (b) (6 points) Un ami vient de mettre en vente sa maison, qui a été évaluée à 280 milliers de dollars. Construisez un intervalle de confiance à 95% pour le prix de vente de sa maison.
- (c) (4 points) Au seuil de 5%, peut-on conclure qu'il existe une relation linéaire significative entre l'évaluation d'une maison et son prix de vente ?
- (d) (points) On veut maintenant vérifier la validité des postulats du modèle de régression.
- i. (1 point) Parmi les graphiques ci-dessous lequel permet de vérifier le postulat de normalité des erreurs ?
- ☐ Graphique A
- ☐ Graphique B
- ☐ Graphique C
- ii. (1 point) Que concluez-vous quant au postulat de normalité des erreurs ?

- ☐ Le postulat de normalité est vérifié.
- ☐ Le postulat de normalité n'est pas vérifié.
- ☐ Le graphique ne permet pas de conclure ; il faudrait une analyse plus poussée.

Exercice

★★★★

Afin d'estimer le volume de bois disponible dans la forêt entre Burtigny et Saint-Oyens, nous aimerions modéliser le volume de bois d'un arbre à l'aide de sa circonférence. Nous disposons du volume (Y) et de la circonférence (X) de 31 arbres. Afin d'étudier le modèle $Y_i = \beta_0 + \beta_1 X_i + E_i$, en supposant que les postulats de régression linéaire sont respectés, nous avons résultats suivants :

$$\sum_{i=1}^{31} x_i = 410,70; \quad \sum_{i=1}^{31} x_i^2 = 5\,736,55; \quad \sum_{i=1}^{31} x_i y_i = 13\,887,86; \quad \sum_{i=1}^{31} y_i = 935,30; \quad \sum_{i=1}^{31} y_i^2 = 36\,324,99.$$

- (a) (7 points) Donnez l'équation de la droite de régression permettant de prédire le volume de bois d'un arbre de circonférence précise. **Donnez les détails de votre démarche.**

Suite de l'exercice 170 sur la feuille suivante

Suite de l'exercice 170

Suite de l'exercice 170

- (b) (3 points) Calculez la prévision ponctuelle du volume de bois d'un arbre ayant une circonférence de 11. **Donnez les détails de votre démarche.**
- (c) (2 points) Alex aimerait estimer le volume de bois de l'arbre devant sa maison. Cet arbre a une circonférence de 14. La prévision ponctuelle du volume de bois de cet arbre est environ 33,98. Choisissez l'intervalle de valeurs ayant 95% de chances de contenir le volume de bois de cet arbre de circonférence 14.

- ☐ $33,98 \pm t_{0,05;29} \sqrt{MS_E \left[\frac{1}{n} + \frac{(14 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $33,98 \pm t_{0,025;1} \sqrt{MS_E \left[\frac{1}{n} + \frac{(14 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $33,98 \pm t_{0,025;29} \sqrt{MS_E \left[\frac{1}{n} + \frac{(14 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $33,98 \pm t_{0,025;29} \sqrt{MS_E \left[1 + \frac{1}{n} + \frac{(14 - \bar{x})^2}{S_{XX}} \right]}$
- ☐ $33,98 \pm t_{0,025;1} \sqrt{MS_E \left[1 + \frac{1}{n} + \frac{(14 - \bar{x})^2}{S_{XX}} \right]}$

- (d) (1 point) Si Alex avait un arbre ayant une circonférence de 18, l'intervalle de valeurs calculé à la question précédente serait...

- ☐ plus court
- ☐ de même longueur
- ☐ plus long
- ☐ on ne peut pas savoir la différence de longueur

- (e) (2 points) La valeur-p du test de significativité du lien linéaire entre X et Y est égale à 0,00014. Cochez l'énoncé qui est vrai.

- ☐ Il n'y a pas de lien linéaire significatif entre X et Y .
- ☐ Il y a un lien linéaire significatif entre X et Y .
- ☐ Nous ne pouvons rien dire car nous ne connaissons pas le seuil du test.
- ☐ Nous ne pouvons rien dire car nous ne connaissons pas t_0 .

Exercice

★★★★ Un limnologue a étudié certaines propriétés physico-chimiques de l'eau du lac Mistassini en 1973. En particulier, il a étudié la quantité d'oxygène dissous Y (en ppm) à diverses profondeurs X (en pieds), à l'aide de la méthode de Winkler. Les résultats obtenus sont les suivants :

X	10	20	30	50	110
Y	10.4	10.4	10.5	10.9	11.4

On a calculé pour vous, à partir de ces données, les quantités suivantes :

$$\sum_{i=1}^5 x_i = 220, \quad \sum_{i=1}^5 y_i = 53.6, \quad \sum_{i=1}^5 x_i^2 = 16\,000, \quad \sum_{i=1}^5 y_i^2 = 575.34, \quad \sum_{i=1}^5 x_i y_i = 2426.$$

- (a) (4 points) Trouvez l'équation de la droite de régression estimée.
- (b) (4 points) Donnez une estimation pour la quantité moyenne d'oxygène dissous lorsque la profondeur du lac est de 10 pieds.

Suite de l'exercice 171 sur la feuille suivante

Suite de l'exercice 171

Suite de l'exercice 171

- (c) (6 points) Les données présentent-elles de l'évidence, au seuil de 5%, qu'il y a une relation linéaire significative entre la quantité d'oxygène dissous et la profondeur du lac ?
- (d) (2 points) Calculez le coefficient de détermination et dire ce que ce nombre signifie.

Exercice

★★★★ Un limnologue a étudié certaines propriétés physico-chimiques de l'eau du lac Mistassini en 1973. En particulier, il a étudié la quantité d'oxygène dissous Y (en ppm) à diverses profondeurs X (en pieds), à l'aide de la méthode de Winkler. Les résultats obtenus sont les suivants :

X	10	20	30	50	110
Y	10.4	10.4	10.5	10.9	11.4

On a calculé pour vous, à partir de ces données, les quantités suivantes :

$$\sum_{i=1}^5 x_i = 220, \quad \sum_{i=1}^5 y_i = 53.6, \quad \sum_{i=1}^5 x_i^2 = 16\,000, \quad \sum_{i=1}^5 y_i^2 = 575.34, \quad \sum_{i=1}^5 x_i y_i = 2426.$$

- (a) (4 points) Trouvez l'équation de la droite de régression estimée.
- (b) (4 points) Donnez une estimation pour la quantité moyenne d'oxygène dissous lorsque la profondeur du lac est de 10 pieds.

Suite de l'exercice 172 sur la feuille suivante

Suite de l'exercice 172

Suite de l'exercice 172

- (c) (6 points) Les données présentent-elles de l'évidence, au seuil de 5%, qu'il y a une relation linéaire significative entre la quantité d'oxygène dissous et la profondeur du lac ?

Exercice

★★★★

On étudie la relation entre le nombre d'heures passées à pratiquer un logiciel de programmation (x , en heures) et le score obtenu à un test (y , en points) chez $n = 12$ élèves.

Un modèle de régression linéaire simple a été ajusté. On obtient :

$$\hat{y} = 41,3 + 2,6x, \quad \hat{\sigma}^2 = 4, \quad S_{xx} = 25, \quad S_{yy} = 200.$$

- (a) (3 points) Pour une observation donnée, on a $x = 4$ et $y = 53$. Calculez le résidu associé à cette observation.
- (b) (6 points) Au seuil 5 %, peut-on affirmer qu'il y a un lien linéaire entre le temps d'étude et le score obtenu au test ?

Suite de l'exercice 173 sur la feuille suivante

Suite de l'exercice 173

(c) (3 points) Calculez le coefficient de détermination R^2 et interprétez le dans le contexte.