TD n° 1
STATISTIQUE DESCRIPTIVE

A - ÉTUDE SUR LA CONSOMMATION JOURNALIÈRE D'UN ARTICLE

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

date (mars)	2	3	4	5	6	7	9	10	11	12	13	14
nombre d'articles vendus	**7**	**13**	**8**	**10**	**9**	**12**	**10**	**8**	**9**	**10**	**6**	**14**
date (mars)	16	17	18	19	20	21	23	24	25	26	27	28
nombre d'articles vendus	**7**	**15**	**9**	**11**	**12**	**11**	**12**	**5**	**14**	**11**	**8**	**10**
date (mars)	30	31										
nombre d'articles vendus	**14**	**12**										
date (avril)	1	2	3	4	6	7	8	9	10	11	13	14
nombre d'articles vendus	**8**	**5**	**7**	**13**	**12**	**16**	**11**	**9**	**11**	**11**	**12**	**12**
date (avril)	15	16	17	18	20	21	22	23	24	25	27	28
nombre d'articles vendus	**15**	**14**	**5**	**14**	**9**	**9**	**14**	**13**	**11**	**10**	**11**	**12**
date (avril)	29	30										
nombre d'articles vendus	**9**	**15**										

A1. Quelle est la variable statistique ? De quel type est-elle ? Comment peut-on organiser les données ?

A2. Regroupér les données en 6 classes d'amplitude 2. Indiquer pour chaque classe :
- Son effectif
- Sa fréquence exprimée en pourcentage.
- Ses fréquences cumulées croissantes et décroissantes, exprimées en pourcentage.

A3. Tracer sur un même graphique les courbes cumulatives croissantes et décroissantes des fréquences. L'abscisse du point d'intersection de ces deux courbes a-t-il une signification particulière ?

A4. a) En utilisant les touches statistiques de votre calculatrice, déterminer à partir de la <u>série classée</u> :
- La valeur moyenne de la série : $\bar{x}$.
- L'écart quadratique moyen de la série : s.

b) Déterminer à présent la valeur moyenne de la série à partir de la <u>série non classée</u>. Que constate-t-on ? Expliquer pourquoi.

A5. Déterminer le pourcentage approximatif de cas où le nombre d'articles vendus se situe dans l'intervalle $\left[\bar{x} - s, \bar{x} + s\right]$.

B - ÉTUDE D'UNE SÉRIE QUANTITATIVE

B1. On considère la série quantitative suivante :

7709	7710	7732	7746	7749	7750	7757	7762	7765	7767	7769	7771	7772
7772	7777	7780	7781	7783	7788	7790	7791	7792	7795	7796	7800	7801
7802	7804	7804	7805	7811	7812	7812	7817	7821	7823	7825	7826	7829
7832	7834	7839	7839	7841	7845	7850	7855	7860	7873	7889		

Quelle est l'étendue de la série? Regrouper les données en dix classes simples à manipuler.

B2. Tracer l'histogramme de la série. En déduire le mode. Représenter sur le même graphique le polygone des effectifs.

B3. Tracer la courbe cumulative des effectifs. En déduire graphiquement la valeur de la médiane. Retrouver cette valeur par le calcul.

B4. Calculer la moyenne et l'écart quadratique moyen :
- En utilisant votre calculatrice.
- Graphiquement en utilisant la méthode de la droite de Henry.

C - POUR VOUS TESTER

C1. Pour chacune des variables suivantes préciser si elle est :
- Qualitative.
- Quantitative discrète.
- Quantitative continue.

1. Revenu annuel. 2. Lieu de résidence. 3. Citoyenneté. 4. Âge.

5. Sexe. 6. Pointure en chaussures. 7. Couleur des yeux.

8. État matrimonial 9. Tour de taille 10. Nombre de langues parlées.

C2. Pour les sujets d'étude qui suivent, spécifier l'unité statistique, identifier la variable statistique sur laquelle porte l'étude ainsi que le type de variable. Préciser dans le cas où la variable est quantitative si elle est continue ou discrète.

Sujet de l'étude	Unité statistique	Variable statistique	Type de variable	Continue ou discrète
Temps d'exécution (en sec) d'un				

Sujet de l'étude	Unité statistique	Variable statistique	Type de variable	Continue ou discrète
programme en basic				
Absentéisme des ouvriers (en jours)				
Classification de la tâche d'un employé.				

C3. Les résultats qu'on obtient avec les courbes cumulatives comportent une erreur d'approximation. Quelle en est la cause?

C4. Quel est le concept probabiliste équivalent à la notion de courbe cumulative croissante?

C5. Quelle est l'hypothèse nécessaire au calcul de la valeur moyenne et de la valeur médiane dans le cas de séries classées ? En déduire si l'affirmation suivante est vraie ou fausse :

> « Si les données brutes ont tendance à se regrouper près de la limite inférieure de plusieurs classes, la moyenne calculée sera nettement supérieure à la véritable moyenne ».

C6. Quel est le principal défaut de la variance, en tant que caractéristique de dispersion?

C7. Calculer : $S = \sum_{i=1}^{n} (ax_i - a\bar{x} + a)$, sachant que $na = 2$.

C8. Préciser si chacune des affirmations suivante est vraie ou fausse :

1. $\sum_{i=1}^{n} (y_i - \bar{y}) = 0$.
2. La quantité : $\sum_{i=1}^{n} (y_i - a)^2$ est minimum lorsque $a = \bar{y}$.
3. $\sum_{i=1}^{n} (y_i + k) = (\sum_{i=1}^{n} y_i) + k$.
4. La médiane est influencée par les valeurs extrêmes d'une série.
5. La moyenne est influencée par les valeurs extrêmes d'une série.
6. Dans une distribution symétrique, la moyenne, la médiane et le mode sont confondus.

D - ÉLÉMENTS DE RÉPONSES AUX QUESTIONS DU C

Lorsque les réponses ne sont pas indiquées, vous pouvez vous référer aux pages du polycopié du chapitre 1 indiquées ci-dessous :

C1 - page 2	**C2** - page 2	**C3** - page 2	**C4** - page 4
C5 - page 2	**C6** - page 9	**C7** - S = 2	**C8 - 1**. page 6
C8 - 2. page 7	**C8 - 3**. Faux	**C8 - 4**. page 6	**C8 - 5**. page 6
C8 - 6. Vrai			

TD N°2
DÉNOMBREMENT - PROBABILITÉS
DISTRIBUTIONS DE PROBABILITÉS

A - AUTO - ÉVALUATION A PROPOS DU COURS

A1. Vrai ou faux ?

- Si la réalisation d'un événement A n'est pas influencée par celle d'un événement B, et inversement, A et B sont des événements incompatibles.
- Si l'événement A est inclus dans l'événement B, la probabilité de A est supérieure à celle de B.
- Un arrangement où l'ordre de présentation des éléments n'est pas pris en considération s'appelle permutation.
- Dans le cas d'une variable aléatoire continue : $p(a \leq x < b) = p(a \leq x \leq b)$.
- L'espérance mathématique d'une variable aléatoire indique la valeur moyenne de cette variable.
- Si on ajoute C à chaque valeur d'une variable aléatoire X :
  - → son espérance mathématique devient : $E(X) + C$.
  - → sa variance devient : $V(X) + C$.
- L'espérance mathématique d'une variable aléatoire centrée réduite est toujours égale à 1.
- Si deux variables aléatoires sont indépendantes, la covariance entre ces variables est nécessairement nulle.

A2. Que vaut $p(A \cap B)$, lorsque A et B sont :
- incompatibles ?
- indépendants ?

A3. Supposons que les probabilités de divers éléments se présentent selon le tableau ci-dessous, les événements A, B, C d'une part et D et E d'autre part étant incompatibles.

$\cap$	A	B	C	TOTAL
D			0.16	0.60
E	0.32		0.04	
TOTAL	0.40			

a) Indiquer sur le tableau les probabilités manquantes.

b) Déterminer $p(C)$; $p(\overline{C})$; $p(\overline{C} \cap A)$. Évaluer $p(D)$; $p(C \mid D)$; $p(C \cup D)$.

A4. Exercices d'application directe du cours

a) En Ile de France, chaque véhicule automobile a un numéro minéralogique comportant quatre chiffres (au plus) et trois lettres. Combien de véhicules peut-on ainsi immatriculer en Ile de France?

b) Combien d'anagrammes peut-on former avec les lettres du mot OIGNON?
Et avec le mot OGNON?(vous pouvez voir là l'amorce d'une réforme de l'orthographe!)

c) Quelles fonctions parmi les suivantes sont acceptables comme fonction de densité d'une variable discrète dont les valeurs possibles sont 0, 1, 2, 3 ?

- $p(0) = 1/4$ $p(1) = 3/8$ $p(2) = 1/16$ $p(3) = 3/16$
- $p(0) = 0$ $p(1) = 1/3$ $p(2) = 1/6$ $p(3) = 1/2$
- $p(0) = 1/5$ $p(1) = 1/4$ $p(2) = 1/3$ $p(3) = 1/2$
- $p(0) =1/4$ $p(1) = 1/2$ $p(2) = -1/4$ $p(3) = 1/2$

d) Quelles fonctions parmi les suivantes sont acceptables comme fonction de densité d'une variable continue?

- $\begin{cases} f(x) = 2/x^3 & \text{pour } 1 \le x \le 2 \\ f(x) = 0 & \text{ailleurs} \end{cases}$.
- $\begin{cases} f(x) = 2x^3 & \text{pour } -1 \le x \le 1 \\ f(x) = 0 & \text{ailleurs} \end{cases}$.
- $\begin{cases} f(x) = \dfrac{1}{2} \cos x & \text{pour } -\dfrac{\pi}{2} \le x \le \dfrac{\pi}{2} \\ f(x) = 0 & \text{ailleurs} \end{cases}$.
- $\begin{cases} f(x) = 2 \sin x & \text{pour } -\dfrac{\pi}{2} \le x \le \dfrac{\pi}{2} \\ f(x) = 0 & \text{ailleurs} \end{cases}$.

e) Une variable X a une espérance égale à 10 et un écart-type égal à 3. Déterminer $E(X^2)$.

B - DÉNOMBREMENT

B1. Contrôle de qualité

a) Dans un lot de vingt pièces fabriquées, on en prélève simultanément quatre. Combien de prélèvements différents peut-on ainsi obtenir ?

b) On suppose alors que sur les vingt pièces, quatre sont mauvaises. Dans combien de prélèvements :

1. les quatre pièces sont bonnes?
2. au moins une pièce est mauvaise?

3. une et une seule est mauvaise?
4. deux au moins sont mauvaises?

B2. De combien de manières peut-on ranger quatre paires de chaussettes dans trois tiroirs?

B3. Une multinationale décide de lancer un dentifrice pour chien. Le nom de ce nouveau produit indispensable doit comporter trois lettres.
- **a)** Combien de noms peut-on former avec toutes les lettres de l'alphabet?
- **b)** Combien de noms peut-on former comportant une consonne et deux voyelles?
- **c)** Combien de noms peut-on former comportant une consonne et deux voyelles différentes?

B4. **a)** Combien de mots de 7 lettres peut-on former avec les lettres A, B, C, D ,E, F, G, en les utilisant toutes ?
b) Combien y a-t-il de ces mots où les lettres CDE sont toujours ensemble
- Dans cet ordre?
- Dans un ordre quelconque?

B5. <u>Modèles en mécanique statistique</u>

En mécanique statistique, on étudie la distribution de particules dans l'espace.
Il est pratique de considérer que l'espace est subdivisé en petites cellules.
Trois modèles uniformes différents ont été proposés.
- Le modèle de Maxwell-Boltzmann (M-B), supposant que l'on peut distinguer les particules et ne limitant pas le nombre de particules par cellule.
- Le modèle de Bose-Einstein (B-E), supposant que l'on ne peut distinguer les particules et ne limitant pas le nombre de particules par cellule.
- Le modèle de Fermi-Dirac (F-D), supposant que l'on ne peut distinguer les particules, avec au plus une particule par cellule.

Donner, en fonction des différents modèles, le nombre de façons de répartir 3 particules dans trois cellules.

C - PROBABILITÉS

C1. Grand-père a trois bérets et une casquette. Quand il descend acheter sa baguette, il saisit un couvre-chef au hasard. Sachant qu'il prend une fois sur trois une baguette moulée et que deux fois sur cinq il oublie de chausser ses souliers, calculer la probabilité de le voir remonter en charentaises, un béret sur la tête et une baguette non moulée sous le bras.

C2. On lance trois dés non pipés. On note le nombre de points (1,2,3,4,5 ou 6) qui apparaît sur la face supérieure de chaque dé. Calculer la probabilité d'avoir :
a) trois 3. **b)** deux 2 et un 1. **c)** un 1, un 3, un 5.
d) la somme des points égale à 9. **e)** la somme des points égale à 10.

<u>Remarque</u> : Ces calculs ont été effectués à l'origine par Galilée pour expliquer la différence entre **d**) et **e**).

C3. Une urne contient une boule rouge, trois boules vertes et seize boules blanches. La boule rouge permet de gagner 10 euros, chaque boule verte permet de gagner 5 euros et les boules blanches ne rapportent rien. Un joueur tire simultanément cinq boules. Quelle est la probabilité pour que ce joueur gagne exactement 10 euros?

C4. Deux chasseurs aperçoivent simultanément un lapin et tirent en même temps. La probabilité que le premier tue le lapin est $\frac{4}{5}$, celle du second est $\frac{3}{4}$.

Quelle est la probabilité que le lapin soit tué?

C5. <u>Réfléchissez bien !</u>

1. On extrait treize cartes d'un jeu de 32 cartes : neuf noires et quatre rouges. On les met face contre table après les avoir mélangées. Avez-vous plus d'une chance sur deux de tirer deux noires ?
2. Pensez-vous que dans votre TD il y ait plus de 50% de chances que deux d'entre vous puissent fêter leur anniversaire le même jour ?

<u>D- DISTRIBUTIONS DE PROBABILITÉS</u>

D1. La distribution uniforme sur un intervalle $[a,b]$ décrit le phénomène qui consiste à prendre une valeur au hasard dans $[a,b]$. Elle a donc une fonction de densité constante sur l'intervalle $[a,b]$.

a) Déterminer la valeur de cette constante.

b) Quelle est la fonction de répartition de cette variable?

c) Quelle est la probabilité de tirer au hasard la valeur $\frac{a+b}{2}$ dans l'intervalle $[a,b]$?

d) Quelle est la probabilité de tirer au hasard une valeur de l'intervalle $[\frac{2a+b}{3}, \frac{a+2b}{3}]$?

e) <u>Application</u> : En arrondissant un nombre réel au nombre entier le plus proche, on introduit une erreur comprise entre -1/2 et +1/2. En supposant que cette erreur suit une loi uniforme continue, calculer :

- $p(\text{erreur} = 0)$;
- $p(\text{erreur} < 0)$;
- $p(|erreur| < 1/4)$.

D2. Pierre vous propose le jeu suivant : il vous paiera une somme en dollars égale au résultat que vous obtiendrez en lançant un dé ordinaire à 6 faces. Pour chaque coup joué, Pierre exige un montant de 3$.

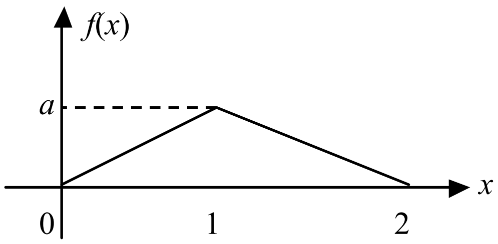
a) Montrez que si vous jouez un suffisamment grand nombre de fois, vous mettrez Pierre en faillite.

b) Pierre décide de changer les règles de son jeu; Il vous donnera 3$ si vous obtenez un cinq ou un six en lançant une fois le dé; Par contre, vous devrez lui verser une somme de 2.75$ si vous obtenez un 1, un 2, un 3 ou un 4. Acceptez-vous de jouer avec Pierre ?

D3. En supposant que le rayon R d'un cercle soit une variable de loi uniforme sur l'intervalle [0,1], déterminer :

a) L'espérance de la circonférence du cercle.
b) La variance de la circonférence du cercle.

D4. On considère la fonction f représentée ci-dessous.



a) Déterminer a pour que f soit la fonction de densité d'une variable continue X.
b) Quelles sont les valeurs de :

$p(X<1)$ $p(X>3/2)$ $p(1/2 \leq X \leq 3/2)$?

D5. On suppose que la distance entre deux stations-service consécutives le long d'une autoroute est de 50 km. On appelle X la distance parcourue par une automobile qui tombe en panne entre deux stations-service depuis le passage devant la première. X est donc comprise entre 0 et 50 km. On fait l'hypothèse qu'elle suit une loi uniforme sur [0,50].

a) La distance entre le lieu de la panne et la station -service la plus proche est une fonction de X, notée $g(X)$. Donner l'expression de $g(x)$ en fonction de la valeur x prise par la variable X, en prenant soin de distinguer deux cas.
b) Quelles sont l'espérance et la variance de $g(X)$.
c) Si une station-service demande un prix fixe de 40 euros plus 10 euros par km à parcourir pour effectuer un dépannage, quel est le coût moyen d'un dépannage? Quel est l'écart-type de ce coût?

D6. Charles ne supporte pas les chats et Sophie déteste les chiens. Charles n'élève pas plus d'un chien et Sophie pas plus d'un chat. La probabilité pour que Charles ait un chien est de 0.2. Si Charles n'a pas de chien, la probabilité pour que Sophie ait un chat de 0.1.
On note X le nombre de chiens de Charles et Y le nombre de chats de Sophie.

a) Calculer la probabilité pour qu'ils n'aient pas d'animaux.
Soit Z le nombre d'animaux du couple. La probabilité que Z soit égal à 1 est de 0.1.
b) Calculer la probabilité pour que Z soit égal à 2.
c) Déterminer $E(Z)$ et $\sigma(Z)$.
d) Établir la loi de probabilité du couple (X,Y).
Quelle est la loi de probabilité de Y?

e) Les variables X et Y sont-elles indépendantes?

E - EXERCICES SUPPLÉMENTAIRES

DÉNOMBREMENT

E1. Au loto, combien a-t-on de possibilités, en choisissant sept numéros (de 1 à 49), d'avoir :

a) Les six bons numéros plus le complémentaire?
b) Six numéros?
c) Cinq numéros plus le complémentaire?
d) Cinq numéros?
e) Quatre numéros?
f) Trois numéros?
Remarque : En dessous de cinq numéros le complémentaire ne joue aucun rôle.

E2. Combien existe-t-il de mains différentes au poker (donnes de 5 cartes parmi 32) comportant :

a) Un brelan d'as?
b) Un carré d'as?
c) Une paire d'as et une paire de rois?
d) Une quinte majeure (10,V,D,R,A)?
e) Un carré?
f) Au moins deux as?

E3. Julie a le choix entre quatre confitures différentes pour étaler sur une tartine, une biscotte et un toast. Combien a-t-elle de possibilités, sachant qu'elle peut éventuellement, en plus de la confiture, les beurrer?

E4. Un club de football est composé de 20 joueurs dont 3 gardiens de but. Combien d'équipes différentes de 11 joueurs peut-on former? (On ne tient pas compte de la place des joueurs sauf pour les gardiens qui ne peuvent jouer que dans les buts).

E5. Quatre bridgeurs sont installés en cours de partie, la répartition des cartes est la suivante :

♣A,V
♦2
♥D,10
♠A,R,D

♣2,3	♣R,D
♦V	♦A
♥9,8	♥A,R
♠V,8,7	♠6,4,2

♣10,9
♦R
♥V,7
♠10,5,3

Combien y a-t-il de manières différentes de terminer la partie, sachant que les plis doivent être de la même couleur?

Indication: On peut commencer par choisir un joueur, SUD par exemple, et compter de combien de façons il peut jouer ses cartes restantes. Puis, pour chaque carte jouée par SUD, compter combien de possibilités les autres joueurs ont de fournir à la couleur demandée.

PROBABILITÉS

E6. Les trois mousquetaires (donc quatre personnes) ont mélangé leurs bottes dans le couloir de l'auberge. D'Artagnan se lève le premier et prend deux bottes au hasard. Calculer la probabilité pour que :

a) Les deux bottes soient les siennes.
b) Les deux bottes forment une paire (une paire est la réunion d'un pied droit et d'un pied gauche).
c) Les deux bottes soient deux pieds droits.
d) Les deux bottes appartiennent à deux personnes différentes.

E7. D'un jeu de 32 cartes, on extrait 6 cartes. Quelle est la probabilité d'avoir exactement 2 dames et 3 trèfles?

E8. Dans une transat à la voile, cinquante concurrents sont engagés. Il y a 30 multicoques dont 20% menés par des femmes et 20 monocoques dont un seul des skippers est une femme. La probabilité de gagner pour un multicoque est supérieure de 10% à celle des monocoques. Les chances des bateaux de même structure sont identiques. Il y a exactement un vainqueur.

a) Calculer la probabilité pour que le vainqueur soit sur un monocoque.
b) Calculer la probabilité pour que le bateau qui franchit en tête la ligne d'arrivée soit un multicoque dirigé par une femme.
c) Calculer la probabilité pour que le vainqueur soit une femme.
d) Sachant que le vainqueur est une femme, calculer la probabilité pour que son bateau soit un monocoque.
e) Sachant que le premier bateau est un monocoque, calculer la probabilité pour que son skipper soit une femme.

E9. Un jeu de cubes se compose de treize cubes noirs et de deux cubes rouges. Combien faut-il en prendre simultanément pour que la probabilité d'avoir au moins un cube rouge dépasse 6/7.?

E10. Une urne contient trois boules blanches et quatre boules noires. On tire des boules une à une sans remettre les boules extraites dans l'urne.

a) Calculer la probabilité de n'avoir que des boules noires dans l'urne en 3,4,5,6 tirages.

b) Que vaut la somme de ces probabilités? Pourquoi?

E11. On jette ensemble quatre dés.

a) Quelle est la probabilité d'obtenir un carré(a,a,a,a), un brelan (a,a,a,b), deux paires (a,a,b,b), une paire (a,a,b,c), une disposition quelconque (a,b,c,d)?

b) Que vaut la somme de ces probabilités?

DISTRIBUTIONS DE PROBABILITÉS

E12. Durant la dernière année, un pâtissier a vendu jusqu'à 9 gâteaux d'anniversaire dans une même journée, mais il n'en a jamais vendu moins que 4. En analysant ses relevés de vente quotidiens, il constate qu'il a vendu 4 gâteaux dans 5% des jours; Il décide donc d'utiliser cette donnée comme probabilité de vendre 4 gâteaux dans une même journée (voici un bel exemple de probabilité empirique). Par la même méthode des fréquences relatives, il a obtenu les autres probabilités suivantes :

p(5 gâteaux) = 0.15 p(6 gâteaux) = 0.35 p(7 gâteaux) = 0.25

p(8 gâteaux) = 0.12 p(9 gâteaux) = 0.08

Combien de gâteaux d'anniversaire peut-il espérer vendre en une journée cette année, si les habitudes de ses clients ne changent pas?

E13. On lance un dé non pipé. Soit X le nombre de points de la face supérieure.

a) Quelle est la valeur moyenne de X ? Sa variance ?

b) On suppose qu'on reçoit 10 euros si on obtient 1, rien si on obtient 2 ou 3, et 4 euros si on obtient 5 ou 6. Quel est le prix honnête à payer pour lancer le dé une fois?

c) On suppose maintenant qu'on reçoive 18euros pour un 1 et rien sinon. Préférez-vous jouer au jeu décrit dans le **b)** où à celui-ci? Pourquoi?

d) On demande maintenant de miser 2 euros pour jouer au jeu du **b)** dans lequel les gains ont été divisés par deux. Quelle est l'espérance de votre gain net? Et sa variance?

E14. Déterminer l'espérance et la variance du volume d'un cube dont le côté suit une loi uniforme sur [0,1].

E15. Au restaurant, un individu a l'habitude de payer la note par le multiple de 1 euro immédiatement supérieur au montant de la facture et de laisser la différence en pourboire. Le pourboire en euros suit une loi uniforme continue de 0 à 1.

Calculer :

$p(0.3 \leq pourboire \leq 0.7)$ $p(pourboire < 0.2)$ $p(pourboire > 0.75)$.

E16. Anatole doit expédier deux objets dans un pays où un paquet sur dix n'arrive jamais à destination. Les objets valent respectivement 100 euros et 150 euros. Il hésite. Faut-il envoyer un seul paquet ou deux paquets séparés?

Soit X et Y les valeurs des envois arrivés à bon port par chacune des méthodes.

Déterminer les espérances mathématiques de ces deux variables aléatoires. Que peut-on en conclure?

E17. Un phénomène est décrit par une variable aléatoire continue dont la densité de probabilité est :

$$\begin{cases} f(x) = \dfrac{x}{k} & \text{si} \quad 0 < x \leq 10 \\ f(x) = \dfrac{20 - x}{k} & \text{si} \quad 10 < x \leq 20 \end{cases}$$

a) Déterminer la valeur de la constante k pour que f soit une densité de probabilité.
b) Tracer cette densité de probabilité.
c) Donner l'expression de sa fonction de répartition F.
d) Déterminer la probabilité que X prenne une valeur dans l'intervalle [5,15].
e) Calculer $F(10)$. Que représente cette probabilité?
f) A quoi correspond la valeur $x = 10$ comme caractéristique de position?

E18. Deux variables X et Y, dont les distributions sont données ci-dessous sont indépendantes.

x	0	1	2
$p(X=x)$	1/4	1/2	1/4

y	0	1	2	3
$p(Y=y)$	1/8	2/8	3/8	2/8

Déterminer :
a) La distribution conjointe de X et Y.
b) La distribution conditionnelle de Y suivant les valeurs de X.
c) $E(XY)$, $E(X+Y)$ et $V(X+Y)$.

F - RÉPONSES AUX EXERCICES DU E

DÉNOMBREMENT

E1. **a)** 1 **b)** 42 **c)** 252 **d)** 5166 **e)** 185115
f) 2 468 200

E2. **a)** 1512 **b)** 28 **c)** 864 **d)** 1024 **e)** 224 **f)** 21196.

E3. 512.

E4. 58344.

E5. 557 383 680.

PROBABILITÉS

E6. **a)** 1/28 **b)** 4/7 **c)** 3/14 **d)** 6/7.

E7. 15435.

E8. **a)** 0.019 **b)** 0.125 **c)** 0.143 **d)** 0.132 **e)** 0.05.

E9. $n<9$.

E10. **a)** en trois tirages : 0.029 ; en 4 tirages : 0.114 ; en 5 tirages : 0.286 ; en 6 tirages : 0.571
b) 1.

E11. **a)** carré : $\frac{1}{216}$; brelan : $\frac{20}{216}$; 2 paires : $\frac{15}{216}$; 1 paire : $\frac{120}{216}$
disposition quelconque : $\frac{60}{216}$ **b)** 1.

DISTRIBUTIONS DE PROBABILITÉS

E12. Entre 6 et 7 (6.48)

E13. **a)** 21/6 et 35/12; **b)** 3euros; **c)** celui du (b); **d)** -5 et 325.

E14. $E(\text{volume}) = \frac{1}{4}$ $V(\text{volume}) = \frac{9}{112}$.

E15. 0.4; 0.2; 0.25.

E16. $E(X) = E(Y) = 225$ euros : cela revient au même.

E17. **a)** 100; **c)** $\begin{cases} F(x) = \dfrac{x^2}{200} & \text{pour } 0 < x < 10 \\ F(x) = \dfrac{-x^2+40x-200}{200} & \text{pour } 10 \le x < 20\,; \\ F(x) = 1 & \text{pour } x \ge 20 \end{cases}$

d) 0.75; **e)** 0.50; **f)** mode, médiane, moyenne.

TD n° 3
PRINCIPALES DISTRIBUTIONS DE PROBABILITÉS

A - AUTO - ÉVALUATION A PROPOS DU COURS

- **Vrai ou faux ?**

1. Une variable qui suit une loi de Poisson peut prendre aussi bien de grandes valeurs que de petites valeurs (au fait, qu'appelle-t-on petites dans ce cas?).
2. A mesure que λ augmente, la distribution $p(\lambda)$ tend à devenir symétrique.
3. L'aire sous la courbe normale $N(m,\sigma)$ à l'extérieur de l'intervalle $[m-3\sigma, m+3\sigma]$ est négligeable.
4. La table de la loi normale centrée réduite permet d'obtenir directement la probabilité $p(-t<T<0)$ si t est une valeur prise par la variable aléatoire T distribuée suivant la loi $N(0,1)$.
5. Si la variable aléatoire X, nombre de succès au cours de n épreuves indépendantes, suit une loi binomiale $B(n,p)$ alors la variable aléatoire $\frac{X}{n}$, proportion de succès au cours des n épreuves, suit aussi une loi binomiale.
6. Si la variable aléatoire X suit une loi binomiale $B(n,p)$, alors la variable aléatoire $\frac{X}{n}$ a pour espérance p et pour variance $\frac{p(1-p)}{n}$.
7. Dans un lot de 2000 transistors, il y a 3% de transistors défectueux. Lorsque l'on prélève un échantillon de 100 transistors, la proportion de transistors défectueux est encore 3%.

- **Savez-vous répondre ?**

1. Quelle est la propriété remarquable de la loi de Poisson concernant ses paramètres descriptifs?
2. Si α est l'intensité d'un processus de Poisson observable pendant un intervalle de temps Δt, que représente $\alpha\Delta t$?
3. Si X suit une loi $N(m,\sigma)$, quelle est la probabilité que X soit supérieur à m?
4. Qu'appelle-t-on variable normale centrée réduite?
5. L'aire sous la courbe normale centrée réduite et au-dessus de l'intervalle $[-1.96, 1.96]$ est : **a)** 0.6826 **b)** 0.9544 **c)** 0.9500
6. Quelle est la valeur de t vérifiant $p(T<t)=0.8413$ lorsque T suit la loi $N(0,1)$?
7. Soient X_1, X_2, X_3, trois variables aléatoires normales indépendantes telles que : $E(X_1) = 100$, $Var(X_1) = 100$, $E(X_2) = 20$, $Var(X_2) = 4$, $E(X_3) = 50$, $Var(X_3) = 25$. On forme la combinaison linéaire : $Y = X_1 + 2X_2 - X_3$.
   a) Quelle est la distribution de Y?
   b) Déterminer $E(Y)$ et $Var(Y)$.
8. Quelle est l'espérance d'une variable de loi géométrique dont la variance est égale à 12?

9. Si une variable de loi exponentielle a une probabilité 1/2 d'excéder la valeur 1, quelle est son espérance?

B - EXERCICES A TRAITER EN TRAVAUX DIRIGES

B1. L'entreprise Luminex fabrique une lampe miniature qui est utilisée comme lampe-témoin sur des panneaux de contrôle électronique. Le responsable du contrôle de la fabrication a établi que 80% des lampes durent plus de 3000 heures. Des tests sont effectués sur des échantillons de taille $n = 15$. On s'intéresse au nombre de lampes dont la durée de vie est inférieure à 3000 heures dans un échantillon de taille 15.

a) Quelle est le nombre moyen de lampes qui ont une durée de vie inférieure à 3000 heures dans un échantillon de taille 15?

b) Quelle est la probabilité que toutes les lampes de l'échantillon durent plus de 3000 heures?

c) Quelle est la probabilité que 13 lampes et plus, dans un échantillon de taille 15, durent plus de 3000 heures?

B2. L'opératrice d'un standard téléphonique reçoit en moyenne 2 appels par minute. Les appels sont répartis au hasard dans le temps.

a) Quelle est la loi de probabilité régissant le nombre d'appels reçus en 5 minutes?

b) Quelle est la probabilité que ce nombre d'appels égale ou dépasse 15?

c) Quelle est la probabilité qu'il y ait entre 5 et 10 appels (bornes comprises) en 3 minutes?

B3. Une machine est conçue pour confectionner des paquets d'un poids de 500g. Elle a confectionné 1000 paquets mais ils n'ont pas exactement tous le même poids. On a constaté que la distribution des poids autour de la valeur moyenne de 500g avait un écart-type de 25g.

a) Combien de paquets pèsent entre 480g et 520g?

b) Combien de paquets pèsent entre 480g et 490g?

c) Combien pèsent plus de 450g?

d) Entre quelles limites symétriques par rapport à la moyenne sont comprises les 9/10 de cette production?

B4. On veut transmettre un message électronique composé des digits 0 et 1. Les conditions imparfaites de transmission font en sorte qu'il y a une probabilité égale à 0.1 qu'un 0 soit changé en un 1 et un 1 en un 0 lors de la réception, et ce de façon indépendante pour chaque digit. Pour améliorer la qualité de la transmission, on propose d'émettre le bloc 00000 au lieu de 0 et le bloc 11111 au lieu de 1 et de traduire une majorité de 0 dans un bloc lors de la réception par 0 et une majorité de 1 par 1.

a) Quelle est la probabilité de recevoir une majorité de 1 si 00000 est émis?

b) Quelle est la probabilité de recevoir une majorité de 1 si 11111 est émis?

B5. Chaque année entre le 10 et le 14 août, la Terre traverse l'orbite de la comète Swift-Tuttle III, qui est jonchée de débris de toutes sortes (poussières microscopiques et cailloux de toutes tailles) qui pénètrent dans l'atmosphère terrestre à grande vitesse, occasionnant une pluie de météores ou d'« étoiles filantes ». On les appelle les

Perseïdes parce qu'elles semblent émaner de la constellation de Persée. Dans cette période, on peut apercevoir jusqu'à 40 étoiles filantes par heure.

a) Montrer que le passage d'étoiles filantes est régi par un processus de Poisson. On suppose alors que l'intensité de ce processus est 40 étoiles par heure.
b) Combien d'étoiles filantes peut-on voir en moyenne en un quart d'heure?
c) Quelle est la probabilité de voir passer au moins 10 étoiles en un quart d'heure?
d) Quel est le nombre maximum d'étoiles que l'on peut voir passer en un quart d'heure dans 95 % des cas?
e) Quel est le temps d'attente moyen entre deux passages successifs d'étoiles filantes?
f) Quelle est la probabilité d'attendre au moins 3 minutes entre deux passages successifs?

B6. Guy est distrait : quand il s'arrête pour prendre de l'essence, il y a une chance sur cinq pour qu'il reparte sans sa passagère, descendue pour visiter les lieux. Soit X la variable aléatoire égale au nombre d'étapes que Guy parcourt en compagnie de sa passagère.
a) Combien d'étapes Guy parcourt-il, en moyenne, en compagnie de sa passagère?
b) Quelle est la probabilité qu'il parcoure au moins trois étapes avec elle?
c) Quel est le nombre maximum d'étapes que peut comporter le voyage pour que la passagère arrive à destination dans la voiture de Guy avec une probabilité supérieure à 0.6?

B7. La distribution statistique suivante donne la répartition du nombre de jours sans accident, avec un accident,..., avec quatre accidents pour une période de 50 jours dans une ville.

Nombre d'accidents x_i	Nombre de jours n_i
0	21
1	18
2	7
3	3
4	1
Total	50

a) Calculer le nombre moyen d'accidents par jours pendant la période de 50 jours.
b) On veut ajuster la distribution précédente à l'aide d'une loi théorique. On choisit la loi de Poisson. Justifier ce choix. Calculer la probabilité pour qu'il arrive 0 accident, 1, 2, 3, 4 accidents. En déduire le nombre théorique Nth_i jours où auront lieu 0, 1, 2, 3, 4 accidents.
c) Comparer Nth_i et n_i ainsi que la variance de la loi théorique et la variance de la loi empirique. Conclure.
d) Combien y aura-t-il de jours sans accident pendant une période de 365 jours?

B8. COMPARAISON DU TEMPS DE FONCTIONNEMENT DE DEUX SYSTÈMES

- Deux éléments d'un système sont montés en parallèle. Le système fonctionne tant qu'au moins l'un des deux éléments fonctionne. On suppose que les temps de fonctionnement des éléments, X_1 et X_2, suivent une loi exponentielle et sont indépendants. X_1 et X_2 ont respectivement pour espérance 1000h et 1500h. Calculer et comparer :
  a) La probabilité que le premier élément fonctionne un temps inférieur à 900h.
  b) La probabilité que le deuxième élément fonctionne un temps inférieur à 900h.
  c) La probabilité que le système fonctionne un temps inférieur à 900h.
- Les deux éléments sont maintenant montés en série. Le système fonctionne seulement lorsque les deux éléments fonctionnent. Calculer la probabilité que le système fonctionne un temps inférieur à 900h. Comparer avec les probabilités calculées au problème précédent.

B9. Le diamètre d'un palier supportant un arbre de transmission doit excéder de 5 à 20 mm le diamètre de l'arbre pour que l'assemblage soit fonctionnel. En supposant que le diamètre de l'arbre suit une loi normale de moyenne 10 cm et d'écart-type 1 mm et que le diamètre du palier suit une loi normale de moyenne 11 cm et d'écart-type 2 mm, calculer la probabilité que l'assemblage d'un palier et d'un arbre pris au hasard de façon indépendante soit fonctionnel.

B10. Un transporteur aérien a observé que 10 % en moyenne des personnes ayant réservé un siège pour un vol ne se présentent pas au départ. S'il accepte jusqu'à 240 réservations alors qu'il ne dispose que de 230 sièges pour ce vol, quelle est la probabilité que toutes les personnes qui se présentent au départ aient un siège?

B11. Un compteur, placé au voisinage d'une source radioactive, enregistre en moyenne 5 impulsions par minute.

a) On effectue un comptage pendant deux minutes. Quelle est la probabilité d'observer un nombre r d'impulsions dans les trois cas suivants :
$r = 10$ $\quad 9 \le r \le 11$ $\quad 7 \le r \le 13$?

b) On effectue un comptage pendant six minutes. Quelle est la probabilité d'observer un nombre r d'impulsions lorsque :
$r = 30$? $\quad 28 \le r \le 32$?

B12. L'entreprise Simtech fabrique des tubes de verre pour l'entreprise Gescom. Gescom utilise ces tubes dans la fabrication d'un de ses produits. Gescom exige que les lots livrés par Simtech contiennent au plus 1% de défectueux. Elle considère également qu'un lot présentant 6% (ou plus) de tubes de verre défectueux est inacceptable et devrait avoir très peu de chances d'être accepté. Les lots sont habituellement constitués de 5000 tubes de verre. Le qualiticien de l'entreprise Gescom a mis au point le plan de contrôle suivant qui est utilisé pour réceptionner chaque livraison de Simtech.
A chaque livraison, prélever au hasard 200 tubes de verre. Si dans cet échantillon, on trouve 4 tubes de verre (ou moins) défectueux, le lot est considéré comme bon. Si l'on

trouve 5 tubes de verre (ou plus) défectueux, le lot est refusé et sera retourné à l'entreprise Simtech sans plus d'inspection

a) On suppose que la proportion de tubes défectueux dans le lot est de 1%. Identifier la loi de la variable aléatoire qui intervient dans ce plan de contrôle. Par quelle loi peut-on l'approcher ? Pourquoi ?

b) Avec ce plan de contrôle, quelle est la probabilité pour Simtech de se voir refuser un lot contenant 1% de défectueux au contrôle de réception de Gescom ?

c) Quelles sont les chances sur 1000 que Gescom accepte un lot comportant 6% de défectueux avec son plan de contrôle ?

B13. Une personne sur 100 est daltonienne. Au conseil de révision pour le service militaire, la visite médicale permet de recenser les daltoniens. On note F_n le pourcentage de daltoniens sur n conscrits. En utilisant une approximation de loi, déterminer une valeur de n à partir de laquelle ce pourcentage se trouve dans l'intervalle [0.009,0.011] avec une probabilité supérieure à 0.9.

C- EXERCICES SUPPLÉMENTAIRES

C1. A un jeu avec un adversaire, on a une probabilité p de gagner, q de perdre, et r de faire une partie nulle. Dans ce dernier cas, on détermine qui sera gagnant par tirage au sort en lançant une pièce de monnaie non pipée. On joue une suite de parties indépendantes dans les même conditions.

a) Quelle est la relation entre p, q, r ?

b) Déterminer la loi exacte des variables aléatoires suivantes :

- Le nombre de parties jusqu'à la première partie nulle.
- Le nombre de parties jusqu'à la première partie gagnée par tirage au sort.
- Le nombre de parties jusqu'à la première partie gagnée.

C2. Des billes produites par un manufacturier ont un diamètre (en cm) qui suit une loi de moyenne 1 et d'écart-type 0.02. Les billes qui ont un diamètre qui n'est pas compris entre 0.96 et 1.05 sont rejetées. Quelle est la proportion des billes qui sont rejetées ?

C3. Soit X une variable de loi normale $N(20,16)$. Trouver 2 nombres réels a et b tels que $p(a \leq X \leq b) = 0.50$.

Remarque : Il y a une infinité de solutions.

C4. Selon le Chevalier de Méré (1607-1685), la probabilité d'obtenir au moins 1 fois une face avec 1 point en lançant de façon indépendante 4 fois un dé est la même que la probabilité d'obtenir au moins une fois 2 faces avec un point chacune en lançant de façon indépendante 24 fois 2 dés. En supposant des dés standard à 6 faces non pipés, le raisonnement est le suivant : en lançant un dé, on a une chance sur 6 d'obtenir 1 face avec 1 point, donc en le lançant 4 fois on aura 4 chances sur 6 d'obtenir au moins 1 face avec 1 point; de façon analogue, en lançant 2 dés, on a une chance sur 36 d'obtenir 2 faces avec 1 point chacune, donc en les lançant 24 fois on aura 24 chances sur 36 d'obtenir au moins une fois 2 faces avec 1 point chacune. Or; on a 24/36=4/6=2/3. Mais en fait, le raisonnement est faux. Trouver l'erreur dans le raisonnement et déterminer les probabilités correctes des deux événements.

C5. L'entreprise Tubex fabrique des tubes de verre qui sont utilisés par une autre usine, filiale de Tubex. D'après le département d'assurance Qualité, la proportion de tubes de verre défectueux se situe à 5 %.

a) Quelle est la probabilité que le $4^{\text{ème}}$ tube de verre contrôlé s'avère le premier défectueux?

b) Combien de tubes doit-on contrôler, en moyenne, pour observer le premier tube défectueux?

C6. DATATION

Les principales méthodes de datation s'appuient sur l'hypothèse que le temps de vie d'un isotope radioactif, appelé isotope mère, suit une loi exponentielle avant de se désintégrer par l'éjection d'une ou plusieurs particules de son noyau en un isotope stable, appelé isotope fille. Le paramètre α de la loi exponentielle, appelé taux de désintégration, dépend du type d'isotope radioactif.

a) Montrer que la probabilité que l'isotope radioactif ne soit pas encore désintégré après un temps t est $e^{-\alpha t}$.

b) Si l'on suppose à l'origine un grand nombre N_0 d'isotopes radioactifs du même type, de taux de désintégration α, avec des temps de vie indépendants et si l'on suppose que le nombre d'isotopes radioactifs qui restent après un temps t est N, déterminer la relation approchée reliant N, N_0, α et t.

c) La demi-vie représente le temps nécessaire pour diminuer de moitié le nombre d'isotopes radioactifs dans un échantillon (roche ou parcelle d'origine organique). Cette demi-vie est notée $t_{1/2}$. α étant toujours le taux de désintégration des isotopes, trouver la relation entre α et $t_{1/2}$.

Le nombre N d'isotopes radioactifs d'un type donné présent aujourd'hui dans un échantillon est obtenu directement par l'observation de ce nombre dans l'échantillon. Le nombre N_0 d'isotopes radioactifs de ce type présents originellement dans l'échantillon est obtenu principalement de deux façons :

- On compte le nombre $\widetilde{N}$ d'isotopes filles dans l'échantillon (lorsqu'on a de bonnes raisons de croire que ceux-ci n'ont pas été dispersés à l'extérieur et qu'il proviennent tous de la désintégration d'isotopes mères dans l'échantillon) et on l'ajoute au nombre d'isotopes toujours radioactifs dans l'échantillon pour obtenir : $N = N_0 + \widetilde{N}$
- On compte le nombre d'isotopes radioactifs du même type dans un échantillon de même nature mais d'âge 0 et on admet que ce nombre est celui à l'origine dans l'échantillon d'âge inconnu, donc égal à N_0 (cette méthode est employée par exemple pour le carbone-14 dont l'isotope fille, l'azote-14, se mélange instantanément à l'azote-14 environnant).

d) L'isotope fille de l'uranium-238 est le plomb-206. La demi-vie de l'uranium-238 est 4.51×10^9 années. Dans un échantillon on a dénombré 4×10^{12} isotopes d'uranium-238 et 10^{12} isotopes de plomb-206. Déterminer l'âge de l'échantillon.

C7. Un automobiliste doit dévisser dans le brouillard les boulons d'une roue de sa voiture. Il utilise une croix dont les quatre extrémités sont des clés de taille différentes indiscernables au toucher.

a) Il procède au hasard, sans méthode. Calculer la probabilité de faire trois essais pour trouver la bonne clé. Généraliser à n essais. On appelle X la variable aléatoire égale au nombre d'essais. Quelle sont son espérance et sa variance?

b) Il procède au hasard, en éliminant les extrémités déjà testées. Soit Y la variable aléatoire égale au nombre d'essais. Quelle est sa loi de probabilité? Calculer son espérance.

C8. Une firme construit un appareil électronique en grande série. Cet appareil est formé de divers éléments que l'on supposera infiniment résistants, sauf deux d'entre eux, que l'on baptisera E_1 et E_2, dont la durée de vie est courte.
Les durée de vie de E_1 et E_2 sont connues par leurs distributions de probabilité qui sont approximativement normales et dont les caractéristiques sont :

- Pour E_1 : durée moyenne de vie 1500h, écart-type 150h.
- Pour E_2 : durée moyenne de vie 1800h, écart-type 200h.

a) On prend un appareil au hasard et on le fait fonctionner. Quelle est la probabilité qu'il soit encore en marche sans avoir eu de panne au bout de 1400h? de 1600h?

b) Quelle est la probabilité qu'au bout de 1600h on ait dû changer un élément et un seul?

c) Quelle est la probabilité qu'au bout de 1600h on ait dû changer au moins un élément?

C9. Une voie d'accès est reliée à une autoroute en un point A. Les véhicules arrivent par la voie d'accès, au point A, au taux de 3 véhicules par minute et de l'autoroute, au taux de 7 véhicules par minute.

a) Quelle est la loi de probabilité régissant le nombre total de passages au point A en 1 minute ?

b) Quel est le nombre moyen de passages au point A en 5 minutes ?

c) Quel est le nombre total de passages au point A le plus probable au cours d'une période de 1 minute ? Est-il unique ?

d) Quelle est la probabilité pour qu'il se présente entre 5 et 10 véhicules (bornes incluses) au point A au cours d'une période de 1 minute ?

e) Quel est le temps moyen d'attente entre le passage de deux véhicules au point A ?

C10. Une confiture est qualifiée de « pur sucre » si elle contient entre 420 et 520g de sucre par kg. Un fabriquant vérifie 200 pots de 1 kg. Il trouve que le poids moyen de sucre est 465g avec un écart-type de 30g.

a) En considérant l'échantillon comme représentatif, calculer le pourcentage de la production du fabricant qui ne doit pas porter la mention « pur sucre » en considérant que le poids du sucre suit une loi normale.

b) Afin d'améliorer la qualité « pur sucre » le fabriquant souhaite éliminer 15% de sa production. Déterminer un intervalle [a,b] , centré sur la moyenne, tel que $p(a \leq X \leq b) = 0.85$.

c) Un magasin diététique lui propose d'acheter les pots à condition qu'ils aient moins de 495g de sucre, mais au minimum x_0 g. Déterminer x_0 sachant que le fabriquant refusera la vente au dessous de 20% de chute.

C11. Une société minière exploite deux gisements. Le premier a une production journalière moyenne de 2000 tonnes avec un écart-type de 150 tonnes. Le deuxième a une production journalière moyenne de 3000 tonnes avec un écart-type de 200 tonnes. Quelle est la probabilité que la production journalière moyenne de la société excède 5200 tonnes? (On admettra que les production journalière des deux gisements sont normales et indépendantes).

C12. En condition normale de fonctionnement, une machine produit des pièces défectueuses dans une proportion constante égale à 1/10000. Un client reçoit un lot tiré au hasard de 30000 pièces usinées par cette machine.

a) Quelle est la probabilité qu'il trouve strictement moins de 3 pièces défectueuses dans le lot?

b) Quelle est la probabilité qu'il y trouve au moins 7 pièces défectueuses?

C13. Dans la région de Trois- Rivières, il semble que 40% de la population soit en faveur de la mise en application d'un « ticket modérateur » pour les services de santé. Quelle est la probabilité que, dans un échantillon de 200 individus de cette région, une majorité soit en faveur de ce ticket modérateur?

C14. Supposons que l'erreur sur une mesure d'observation suive une loi $N(0,4)$.
Déterminer :

a) La probabilité que l'erreur sur une mesure d'observation soit positive.

b) L'espérance et la variance du nombre d'erreurs positives sur 100 mesures d'observation indépendantes.

c) La probabilité que 60 erreurs positives ou plus soient faites sur 100 mesures d'observation indépendantes.

C15. En Tapségonie, le ministre de l'Éducation décide de donner le bac à 80% des enfants dès la naissance. Pour chaque nouveau-né, on tire un numéro au hasard entre 1 et 10.

- Si c'est un multiple de 5, le bébé est recalé.
- Si c'est un sept, il obtient la mention « très bien ».
- Si c'est un multiple de 4, il obtient la mention « bien ».
- Si c'est un multiple de 3, il obtient la mention « assez bien ».
- Sinon, il obtient la mention « passable ».

a) Calculer la probabilité pour un nouveau-né d'obtenir le bac avec la mention « passable ».

Le village natal du ministre attend sept naissances pour l'année qui suit. On sait qu'il y aura trois filles et quatre garçons. Soit X la variable aléatoire égale au nombre de garçons bacheliers et Y la variable aléatoire égale au nombre de filles bachelières parmi ces bébés.

b) Déterminer les lois de probabilité de X et Y.

c) Calculer la probabilité d'avoir strictement plus de bachelières que de bacheliers.

d) Calculer la probabilité pour que ce village dépasse l'objectif du ministre.

e) La ville dont le ministre est maire attend six cents naissances. Calculer la probabilité pour qu'il y ait moins de 500 bacheliers parmi ces bébés.

D - RÉPONSES AUX EXERCICES DU C

C1 **a)** $p + q + r = 1$. **b)** $X_1 \rightsquigarrow G(1-p-q)$.

$X_2 \rightsquigarrow G(\frac{1}{2}(1-p-q))$.

$X_3 \rightsquigarrow G(\frac{1}{2}(1+p-q))$.

C2 0.971.

C3 En choisissant a et b symétriques par rapport à m : $a = 9.2$; $b = 30.8$.

C4 → Obtenir au moins une face avec un point en 4 lancers : $p_1 = 0.517$.
→ Obtenir au moins deux faces avec un point en 24 lancers : $p_2 = 0.491$.

C5 **a)** 0.042. **b)** 20.

C6 **b)** $N \cong N_0 e^{-\alpha t}$. **c)** $t_{1/2} = \dfrac{ln 2}{\alpha}$. **d)** 1.45×10^9 années.

C7 **a)** 0.14 ; $\left(\dfrac{3}{4}\right)^{n-1} \times \dfrac{1}{4}$; $E(X) = 4$; $Var(X) = 12$.
b) Y suit une loi uniforme de densité 0.25. $E(Y) = 0.25$.

C8 **a)** 0.7315; 0.2142. **b)** 0.6675. **c)** 0.7858.

C9 **a)** $P(10)$ **b)** 50 ; **c)** 9 ou 10; **d)** 0.5538; **e)** 6 mn.

C10 **a)** 0.89. **b)** [422,508]. **c)** 413g.

C11 0.2119.

C12 **a)** 0.4232; **b)** 0.0336.

C13 0.0015.

C14 **a)** 0.5. **b)** $E(Y) = 50$; $V(Y) = 25$. **c)** 0.179.

C15 **a)** 0.2. **b)** $X \rightsquigarrow B(4, 0.8)$; $Y \rightsquigarrow B(3, 0.8)$. **c)** 0.02.
d) 0.58. **e)** 0.977.

TD n° 4
INTERVALLES DE CONFIANCE

A - AUTO - ÉVALUATION A PROPOS DU COURS

- **Vrai ou faux ?**

1. Les paramètres d'une population sont des quantités aléatoires.
2. Lorsqu'un paramètre d'une population est estimé par un seul nombre, déduit des résultats de l'échantillon, ce nombre est appelé une estimation ponctuelle du paramètre.
3. « L'estimation » est la valeur numérique que prend l'estimateur selon les observations de l'échantillon.
4. L'estimation ponctuelle ne fournit aucune information concernant la précision de l'estimation effectuée.
5. Dans le processus d'estimation par intervalle de confiance, on peut dire avant la réalisation de l'expérience, que les limites de l'intervalle de confiance sont des valeurs fixes.
6. Un intervalle de confiance est toujours centré sur la valeur de l'estimateur du paramètre.
7. Plus le niveau de confiance associé à l'intervalle est élevé, plus l'amplitude de l'intervalle est grande.
8. Pour le même niveau de confiance et le même écart-type, plus la marge d'erreur est requise est faible, plus la taille de l'échantillon sera élevée.
9. A mesure que le nombre de degrés de liberté augmente, la distribution de Student tend à s'approcher de la loi normale centrée réduite.
10. Le nombre de degrés de liberté est une quantité qui est associée à une somme de carrés.

- **Savez vous répondre?**

1. Quel théorème important donne un résultat général concernant la distribution d'échantillonnage de la variable aléatoire $\overline{X}$? Quelle règle pratique est nécessaire pour appliquer ce théorème ?
2. Si la population que nous échantillonnons est distribuée selon une loi normale, que peut-on dire de la distribution de la moyenne $\overline{X}$? Existe-t-il une restriction sur la taille d'échantillon que l'on doit prélever de la population ?
3. Quel est l'axe de symétrie de la distribution de Student ? A celle de quelle distribution sa forme est-elle comparable ?
4. Quel est le paramètre de la loi de Student ?
5. A quoi a-t-on recours lorsqu'on doit estimer une valeur en indiquant la précision de l'estimation ?

B - EXERCICES A TRAITER EN TRAVAUX DIRIGES

B1. On mesure n fois une certaine grandeur physique G.
Les résultats des mesures sont corrigés des erreurs systématiques de façon à ne laisser subsister que les erreurs aléatoires.

Soient $G_1, G_2, G_3, \ldots\ldots, G_n$ ces résultats. La vraie valeur G_v de G est la moyenne des G_i dans une population infinie de mesures.

a) On suppose connue la précision de l'appareil de mesure. Cette précision est caractérisée par l'écart-type σ_{pop} des G_i dans une population infinie de mesures : $\sigma_{pop} = 4.3\,10^{-3}$ (unité arbitraire).

1) On effectue une seule détermination G_1 de G. On trouve $G_1 = 1.364$.
Donner la meilleure estimation possible de G_v.
Déterminer au seuil de 90% l'intervalle de confiance relatif à G_v.

2) On effectue 10 déterminations. On trouve :

1.365	1.371	1.368	1.359	1.362
1.366	1.365	1.367	1.363	1.364

Donner la meilleure estimation possible de G_v.
Déterminer au seuil de 90% l'intervalle de confiance relatif à G_v.

b) On ignore la précision de l'appareil de mesure. Déterminer au seuil de 90% l'intervalle de confiance de G_v à partir de la série de mesures de la question **a2**.

c) A partir de la série de mesures de la question **a2**, déterminer au seuil de 90% l'intervalle de confiance relatif à σ_{pop}.

B2. A la veille d'une consultation électorale concernant deux candidats, on a interrogé 100 électeurs constituant un échantillon représentatif. 58 d'entre eux ont déclaré avoir l'intention de voter pour le candidat DUPONT.

a) Indiquer avec une probabilité de 0.95 entre quelles limites se situe la proportion du corps électoral favorable à DUPONT au moment du sondage.
Peut-on en déduire, avec la même probabilité de 0.95 que DUPONT serait élu si les opinions ne se modifiaient pas ?

b) Avec une même fréquence observée d'électeurs favorables à DUPONT, quelle devrait être la taille minimum de l'échantillon pour pouvoir affirmer, avec un risque de 5% que DUPONT sera élu ?

B3. La firme Comtec vient de développer un nouveau dispositif électronique qui entre dans la fabrication d'appareils de traitement de texte. Avant de mettre en production ce nouveau dispositif, on veut effectuer des essais préliminaires pour être en mesure d'en estimer la fiabilité en termes de durée de vie. D'après le bureau de Recherche et Développement de l'entreprise, l'écart-type de la durée de vie de ce nouveau dispositif électronique serait de l'ordre de 100 heures.

a) Déterminer le nombre d'essais requis pour estimer, avec un niveau de confiance de 95%, la durée de vie moyenne d'une grande production de sorte que la marge d'erreur dans l'estimation n'excède pas 50 heures.

b) Pour le même niveau de confiance, quelle doit être le nombre d'essais de sorte que la marge d'erreur dans l'estimation de la durée de vie moyenne n'excède pas 20 heures ?

C EXERCICES SUPPLÉMENTAIRES

C1. Dans un atelier mécanique, on a vérifié le diamètre de tiges tournées autour d'un tour automatique. Le diamètre des tiges peut fluctuer selon le réglage du tour. Vingt tiges prélevées au hasard, ont été mesurées avec un micromètre de précision. Les résultats sont présentés dans le tableau ci-dessous.

39.5	39.9	39.7	40.0	40.65	39.1	39.4	38.4	40.0	41.5
42.6	41.1	37.8	38.5	40.0	41.3	39.4	41.2	38.4	40.8

a) En supposant que le diamètre des tiges est distribué selon une loi normale, estimer par intervalle de confiance le diamètre moyen des tiges de cette fabrication avec des niveaux de confiance de 90%, de 95% et de 99%.

b) Si le diamètre moyen des tiges doit être de 40 mm, doit-on envisager, selon les résultats de cet échantillon, de modifier le réglage du tour ?

C2. Dans une municipalité, on a effectué un sondage pour connaître l'opinion des contribuables sur un nouveau règlement d'emprunt. D'une liste informatisée de 6000 payeurs de taxes, on a prélevé, par tirage au sort, 150 noms. Sur ces 150, 45 étaient en faveur du nouveau règlement d'emprunt. Déterminer un intervalle de confiance pour p, la proportion vraie de contribuables de cette municipalité qui sont en faveur du nouveau règlement d'emprunt, avec les niveaux de confiance de 90%, de 95% et de 99%.

D RÉPONSES AUX EXERCICES DU C

C1 **a)** à 90% $39.48 \leq \bar{x} \leq 40.44$

à 95% $39.38 \leq \bar{x} \leq 40.54$

à 99% $39.17 \leq \bar{x} \leq 40.75$

b) Non

C2 à 90% $0.238 \leq p \leq 0.352$

à 95% $0.226 \leq p \leq 0.374$

à 99% $0.203 \leq p \leq 0.397$

TD n°5
LES TESTS D'HYPOTHÈSES

A - AUTO-EVALUATION A PROPOS DU COURS

- **Vrai ou faux?**

1. Lorsqu'on veut comparer deux variances, on a recours à la distribution de Fisher-Snédécor.
2. La quantité $F = \dfrac{\chi_1^{\ 2}/\nu_1}{\chi_2?/\nu_2}$ est distribuée selon la loi de Fisher avec ν_1 et ν_2 degrés de liberté.
3. La quantité F définie à la question précédente prend ses valeurs sur 3.
4. Lorsqu'on veut juger de la qualité d'ajustement d'une distribution théorique, on a recours à un test de Pearson ou test du khi-deux.
5. On dit que les écarts observés entre les fréquences théoriques et les fréquences observées sont significatifs lorsque la valeur calculée du χ^2 est plus faible que la valeur critique $\chi?_{\alpha,\nu}$.
6. Dans un tableau de contingence comportant r lignes et k colonnes, le nombre de degrés de liberté du khi-deux est $(r - 1)(k - 1)$.

- **Savez vous répondre ?**

1. Dans un test de comparaison de moyennes, quelle est la statistique qui convient pour le test ?
2. Dans un test d'ajustement, on perd toujours un degré de liberté dans le calcul du nombre de degrés de liberté du χ^2. Pourquoi ?
3. La distribution du khi-deux est-elle une distribution centrée ?
4. En supposant qu'il y ait concordance parfaite entre la distribution observée et la distribution théorique dans un tes du khi-deux, quelle valeur prendrait le χ^2 ?

B - EXERCICES A CHERCHER EN TRAVAUX DIRIGES

B1. On règle une machine pour fabriquer des pièces de diamètre 50 mm. On prélève sur la fabrication 10 pièces. On trouve dans cet échantillon un diamètre moyen de 50.2mm et un écart-type de 0.1mm. Peut-on en conclure au seuil de 5% que la machine est bien réglée ?

B2. On réceptionne deux lots d'acier et on veut déterminer s'ils sont de qualités semblables. Pour cela, on prélève dans chaque lot des éprouvettes que l'on soumet à des essais de traction. On trouve :

nombre d'éprouvettes	150	125

résistance moyenne à la rupture(kg/mm_)	105	101
écart-type sur la résistance à la rupture (kg/mm_)	12	14

Peut-on dire au seuil de 5% que ces deux lots sont semblables ?

B3. Sur un lot de 700 boulons soumis à des essais de rupture, 300 ont résisté. Sur un second lot de 225, 125 ont résisté. Peut-on admettre, au seuil de 5%, que ces deux lots appartiennent à la même population ?

B4. Une population de souris présente un taux de cancer spontané de 20%. On se demande si un traitement donné modifie ce taux. On applique ce traitement à 100 souris. On en trouve 11 cancéreuses.

a) Le traitement est-il efficace au seuil de 1% ? Au seuil de 5% ?

b) Sur combien de souris l'étude doit-elle porter pour qu'une différence de 2% entre la proportion de cancéreux dans l'échantillon et dans la population soit significative au seuil de 5% ?

B5. Deux méthodes de dosage de l'azote sont comparées du point de vue de leur précision. Une mesure répétée 25 fois en utilisant la méthode A donne :

$$\bar{x}_1 = 42.08 \quad \text{et} \quad s_1? = 5.16.$$

La même mesure répétée 30 fois en utilisant la méthode B donne :

$$\bar{x}_2 = 42.1 \quad \text{et} \quad s_2? = 1.96.$$

Peut-on dire au seuil de 5% que les deux méthodes ont la même précision ?

B6. La série statistique suivante obéit-elle à une loi normale ayant pour espérance mathématique et pour écart-type les valeurs de la moyenne et de l'écart quadratique moyen de la série (7800 et 39) ?

classes	[7700,7720[	[7720,7740[	[7740,7760[	[7760,7780[	[7780,7800[
effectifs	2	1	4	8	9
classes	[7800,7820[	[7820,7840[	[7840,7860[	[7860,7880[	[7880,7900[
effectifs	10	9	4	2	1

B7. On se propose de comparer les réactions produites par deux vaccins désignés par A et B. Pour cela, un groupe de 348 patients est divisé en deux sous-groupes. L'un est vacciné avec le vaccin A, l'autre avec le vaccin B. On trouve :

Vaccin	Réaction légère	Réaction moyenne	Ulcération	Abcès	Total
A	12	156	8	1	177
B	29	135	6	1	171

Les répartitions diffèrent-elles significativement au seuil de 5% ?

C - EXERCICE SUPPLÉMENTAIRE

Le tableau suivant donne le classement de 200 naissances en fonction de la parité (primipare = femme ayant un premier enfant ; multipare = femme ayant un deuxième, troisième, etc...enfant) et du poids du nouveau né :

	Poids < 3 kg	3 < Poids < 4 kg	Poids > 4 kg
Primipares	26	61	8
Multipares	20	63	22

1. Peut-on dire au seuil de 5% que la parité a une influence sur le poids du nouveau-né ?

2. Voulant préciser cette dépendance, on détermine la moyenne et l'écart-type du poids du nouveau né pour cet ensemble de naissances. On trouve :

Primipares	95 naissances	$P_{\text{moy}} = 3197$ g	$\sigma_p = 458$ g
Multipares	105 naissances	$P_{\text{moy}} = 3410$ g	$\sigma_p = 505$ g

En comparant moyennes et variances, toujours au seuil de 5%, aboutit-on à la même conclusion qu'au **1** ?

D - RÉPONSES A L'EXERCICE DU C

1. oui.

2. les moyennes sont différentes au seuil de 5%, et les écarts-types semblables.

TD n° 6
RÉGRESSION ET CORRÉLATION LINÉAIRES

A - AUTO - ÉVALUATION A PROPOS DU COURS

- **Vrai ou faux ?**

1. Dans un problème de régression, pour chaque valeur x_i, la composante $\alpha+\beta x_i$ a un comportement aléatoire.

2. Dans le modèle linéaire simple de régression, on emploie le qualificatif « simple » parce que le modèle ne comporte qu'une seule variable dépendante.

3. Pour un échantillon donné, il peut exister plusieurs droites d'ajustement qui permettent de minimiser $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ par la méthode des moindres carrés.

4. La droite de régression ne passe pas toujours par le point $(\bar{x}, \bar{y})$.

5. La variance $Var(B)$ de la variable B a tendance à augmenter lorsque la variance σ^2 des résidus diminue.

6. Si la dispersion des valeurs de la variable explicative est faible, la variance $Var(B)$ est élevée.

7. La valeur du coefficient de corrélation r varie entre 0 et 1.

8. Lorsqu'on veut tester l'hypothèse selon laquelle il y a absence de corrélation linéaire entre deux variables aléatoires, on pose comme hypothèse nulle $H_0 : \rho = 0$.

9. La distribution d'échantillonnage du coefficient de corrélation r est symétrique dans le cas où $\rho = 0.5$.

10. Si on effectue des observations de 2 variables indépendantes, on s'attend à ce que le coefficient de corrélation linéaire soit proche de 0.

11. Un coefficient de corrélation linéaire négatif signifie une faible corrélation.

12. Si deux variables sont corrélées linéairement, alors nécessairement la variation de l'une est la cause de la variation de l'autre.

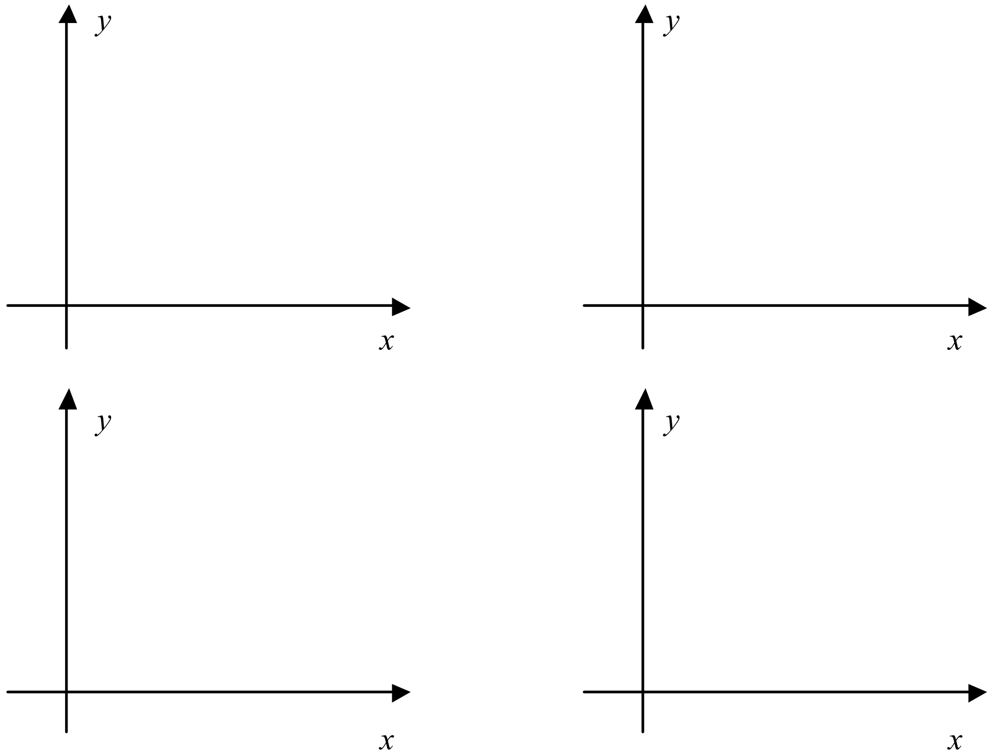
13. Si, à un ensemble de n points (x_i, y_i) pour $i = 1, \ldots, n$, on ajoute m points tous égaux à $(\bar{x}, \bar{y})$, alors le coefficient de corrélation linéaire reste le même.

- **Savez-vous répondre ?**

1. Dans une étude de régression, quel est le nom de la variable qui possède un caractère aléatoire?

2. Attribuer à chacun des graphiques suivants, les énoncés qui vous semblent plausibles :

a) Modèle linéaire d'ordre un avec pente positive.
b) Modèle linéaire d'ordre un avec pente négative
c) Aucune liaison.
d) Modèle linéaire dont l'ordre est supérieur à un.



3. Selon les hypothèses du modèle linéaire simple, préciser ce que valent les expressions suivantes :

$E(\varepsilon_i) =$ $V(\varepsilon_i) =$ $E(Y_i) =$

$V(\alpha + \beta x_i) =$ $V(Y_i) =$

4. Dans le modèle de régression linéaire simple, quelle hypothèse fondamentale précise-t-on concernant la variance des erreurs ε_i ?

5. Lorsqu'on connaît le coefficient b de la droite de régression, quelle expression nous permet de calculer rapidement a?

6. Sous quelles conditions la variable : $T = \dfrac{B - \beta}{V(B)}$ est-elle distribuée selon une loi de Student à $(n\text{-}2)$ degrés de liberté?

7. Si on veut tester l'hypothèse nulle $H_0 : \alpha = 0$ (la droite de régression passe par l'origine), quelle condition est requise quant aux observations pour que ce test ait une signification réelle?

8. Soit les données suivantes qui représentent les productions mensuelles (x_i en milliers d'unités) et les coûts de production correspondants (y_i en milliers de dollars) d'une entreprise au cours d'une année (i = 1, 2, ..., 12) :

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	5	4	7	8	6	7	9	8	6	7	5	3
y_i	70	60	90	80	70	80	90	75	80	75	75	60

a) L'équation de la droite de régression de y en x est $y = 47.42 + 4.48x$. Interpréter l'ordonnée à l'origine de cette droite.

b) Estimer le coût de production de 10000 unités à l'aide de la droite déterminée à la question précédente. Peut-on se fier à ce nombre?

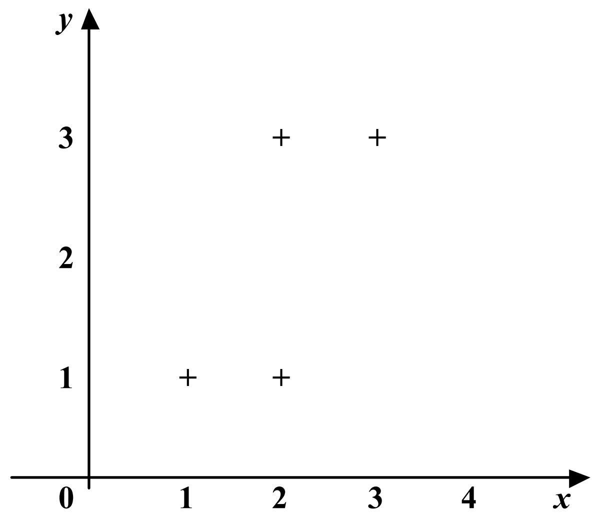
c) On détermine que :

- la variance des coûts mensuels de production expliquée par la droite de régression est $s_E^{\,2} = 57.33$
- la variance résiduelle de ces coûts est $S_R^{\,2} = 27.91$

Quelle conclusion tirer à la lumière de ces renseignements ?

B - <u>EXERCICES A TRAITER EN TRAVAUX DIRIGÉS</u>

B1. Voici la représentation graphique de 4 points dans le plan :



a) Déterminer la droite de régression de y en x. Représentez la sur la figure.

b) Calculer le coefficient de corrélation.

c) On ajoute les points de coordonnées (5,5), (6,5), (6,7), (7,7) à ceux déjà tracés. Sans faire de calculs, répondre aux questions suivantes :

- Est-ce que la droite de régression de y en x reste la même? Sinon, sa pente est-elle plus grande ou plus petite?
- Vous attendez-vous à un coefficient de corrélation inférieur, égal ou supérieur à celui calculé au **b)**?

B2. Sous polarisation directe le courant d'une diode Schottky est donné par la relation :

$$I = I_s \exp(qV / nkT). \qquad (1)$$

I_S courant de saturation; q charge élémentaire; T température thermodynamique;
k constante de Boltzmann, n paramètre de non-idéalité; V tension appliquée.
On mesure le courant d'une diode Schottky en fonction de la tension directe, et on trouve :

V(en mv)	100	150	200	250	300	350	400
I (en A)	$1.66 .10^{-4}$	$1.11 .10^{-3}$	$7.5 .10^{-3}$	$5.02\ 10^{-2}$	$3.37 .10^{-1}$	2.26	15.20

1. Transformer la relation (1) de manière à obtenir une relation linéaire entre une fonction de I et V.
2. En utilisant la méthode des moindres carrés déterminer le paramètre de non idéalité n et le courant de saturation I_S (on a $kT / q = 25mV$).

B3. On cherche à déterminer si deux variables aléatoires sont liées entre elles. Pour cela, on mesure ces deux variables pour 8 événements. On obtient :

X	52	47	42	35	34	36	37	46	51
Y	30	29	27	23	22	23	24	25	28
X	45	50	53	50	45	48	40	43	38
Y	26	27	28	29	28	27	24	25	23

a) Calculer le coefficient de corrélation.
b) Déterminer au seuil de 5% si les deux variables sont liées entre elles.
c) Donner les équations des deux droites de régression :

$$y = a + bx \qquad \text{et} \qquad x = a' + b'y.$$

B4. L'agence de publicité ABC constate qu'il existe une relation entre les budgets de publicité de ses clients (en milliers de francs) et le montant de leurs ventes (en centaines de milliers de francs).

Client	1	2	3	4	5	6	7	8
Montant des ventes (en centaines de milliers de francs)	6	7	9	9	7	8	6	12
budgets de publicité (en milliers de francs)	45	80	70	85	60	55	75	90

a) Déterminer l'équation de régression suivant le critère des moindres carrés.
b) En effectuant un test sur la pente peut-on conclure qu'il existe une relation entre les variables au seuil de 5%?
c) Calculer les coefficients de détermination et de corrélation. Comment les interpréter?

C - EXERCICES SUPPLÉMENTAIRES

C1.

a) Étant donné un ensemble de points $(x_1, y_1), \ldots\ldots, (x_n, y_n)$, déterminer par la méthode des moindres carrés la droite de la forme $y = bx$ qui approche le mieux ces points.

b) Effectuer les calculs pour les points suivants.

(0, 1)	(1, 2)	(1, 1)	(-2, 0)	(-2, -2)	(2, 1)	(0, -1)	(2, 0)

c) Représenter le nuage de points et la droite sur un graphique.

C2. On fait passer un examen écrit à 10 conducteurs pour vérifier leurs connaissances sur la conduite automobile (compréhension des signes routiers...). Le tableau ci-dessous indique le résultat obtenu à l'examen ainsi que le nombre d'années d'expérience de la conduite automobile.

Résultats (%)	60	65	70	67	75	75	80	78	85	82
nombre d'années	6	4	7	9	4	10	6	7	4	8

a) Calculer le coefficient de corrélation linéaire entre ces deux variables.
b) Est-ce que la valeur de r est suffisamment élevée, au seuil $\alpha = 0.05$, pour conclure que ces deux variables sont corrélées?

C3. Une étude sur le taux de mortalité due au cancer du poumon et sur la consommation de cigarettes dans certains états des États-Unis a permis d'obtenir le tableau ci-dessous.

État	consommation annuelle de cigarettes par habitant	Taux de mortalité due au cancer du poumon pour 100000 habitants
Delaware	3400	24
Indiana	2600	20
Iowa	2200	17
Montana	2400	19
New Jersey	2900	26
Washington	2100	20
Moyenne	2600	21

a) Tester au niveau de 5% l'hypothèse que la consommation de cigarettes ne modifie pas le taux de mortalité due au cancer du poumon.
b) La décision est-elle la même avec un niveau de signification de 1%?

C4. Si un échantillon de 2 variables couplées est de taille 30, quelle est la plus grande valeur absolue du coefficient de corrélation qui permet de conclure à l'indépendance possible des 2 variables en cause au niveau de 10%?

D - RÉPONSES AUX EXERCICES DU C :

C1 **a)** $b = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^{2}}$ **b)** $y = \dfrac{1}{2}x$ remarque : La droite de régression de y en x a pour équation : $y = 0.13 + 0.48x$ si on ne lui impose pas de passer par 0.

C2 **a)** $r = -\,0.0359$.
b) non.

C3. **a)** on rejette H_0.
b) on accepte H_0.

C4. 0.3060.

TD n° 7
PROCESSUS STOCHASTIQUES

A – CHAINES DE MARKOV

A1. Une unité de production dispose de deux machines automatiques fonctionnant indépendamment et ayant chacune une fiabilité p au cours d'une journée. Lorsqu'elle tombe en panne elle est réparée pendant la nuit et se retrouve donc en état de marche le lendemain. Mais une seule machine peut être réparée à la fois. Soit X_n le nombre de machines en panne au début de la n-ième journée.

a) Dessiner le graphe des transitions correspondant.
b) Etablir la matrice des transitions.
c) On suppose les deux machines sont au départ en état de marche. Calculer les probabilités de l'état 1, puis celles de l'état 2.

A2. Soit un dispositif technique comprenant deux éléments montés en parallèle et fonctionnant indépendamment. Chaque élément a une fiabilité égale à p au cours d'une journée, et ne peut pas être réparé. Soit X_n le nombre de machines en panne au début de la n-ième journée.

a) Dessiner le graphe des transitions correspondant et établir la matrice des transitions.
b) Déterminer la distribution stationnaire de cette chaîne de Markov. Est-elle une distribution limite?
c) Montrer que cette chaîne de Markov est absorbante.
d) On suppose maintenant que $p = 0{,}9$. Combien de jours faut-il en moyenne pour que le dispositif technique ne fonctionne plus, si les deux machines sont au départ en état de marche?

A3. Les données de l'exercice **A1** restent valables mais maintenant on ne répare plus une machine tombée en panne que le lendemain.

a) Dessiner le graphe des transitions correspondant.
b) Etablir la matrice des transitions.
c) Déterminer la distribution stationnaire de cette chaîne de Markov. Est-elle une distribution limite?

A4. Soit une chaîne de Markov à deux états dont la matrice de transition est :

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

où $0 \leq a \leq 1$, $0 \leq b \leq 1$ et $0 < a + b < 2$.

a) Calculer P^n.
b) En déduire la distribution notée $\pi(n)$ de X_n.
c) Montrer que la chaîne de Markov définie par P est convergente et déterminer la limite π.
d) Existe-t'il une distribution stationnaire? Est-elle unique?

e) Que se passe-t'il dans le cas où $a + b \in \{ 0, 1 \}$?

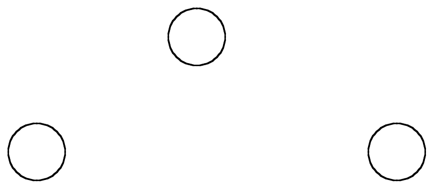
A5. Considérons une chaîne de Markov de matrice $P = \begin{pmatrix} 3/4 & 1/4 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}$

Déterminer le régime permanent du processus stochastique ainsi défini.

A6. Soit la chaîne de Markov définie par $P = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

a) Déterminer les classes transitoires et récurrentes de cette chaîne de Markov
b) Existe-il une distribution stationnaire? une distribution limite?

A7. Soit la chaîne de Markov définie par le graphe suivant :



a) Etablir la matrice des transitions.
b) Montrer que la chaîne de Markov ainsi définie est absorbante.
c) Déterminer la distribution du nombre de transitions jusqu'à l'absorption en partant de l'état 1. En déduire le temps moyen d'absorption en partant de l'état 1.
d) Retrouver directement ce dernier résultat.

A8. Dans une production en série, les articles passent par 3 étapes de fabrication; ils sont inspectés à la fin de chaque étape. Ils peuvent alors présenter 3 états possibles : totalement défectueux (probabilité p; l'article est jeté), légèrement défectueux (probabilité q; il passe une seconde fois par la même étape), en bon état (probabilité r avec $p + q + r = 1$).
Déterminer, en prenant $p = 0{,}1$ et $q = 0{,}3$, la durée moyenne jusqu'à ce qu'un article quitte la machine ainsi que la probabilité qu'en quittant la machine cet article soit en bon état.

A9. Un scarabée se déplace sur les arêtes d'un tétraèdre régulier dont les sommets sont numérotés de 1 à 4. Initialement il est en 1 et une arête est parcourue en une unité de temps. De chaque sommet le choix de parcourir une des trois arêtes voisines est équiprobable. Calculer le temps moyen nécessaire τ pour atteindre le sommet 4 pour la première fois, puis la probabilité p d'atteindre 4 si l'on rajoute de la colle au sommet 2.

A10. Dans une station service équipée d'une seule pompe, on a observé que le nombre de voitures arrivant au cours du n-ième intervalle de temps $[n-1, n]$ est une variable aléatoire Y_n de distribution $p (Y_n = 0) = p (Y_n = 1) = 0{,}4$ et $p (Y_n = 2) = 0{,}2$ pour tout $n \geq 1$. On suppose que chaque service dure exactement une unité de temps; les services commencent (si au moins une voiture est présente) aux instants 1, 2 ...La station dispose de deux places d'attente. On note X_n le nombre de voitures se trouvant dans la station immédiatement après l'instant n.

a) Vérifier que (X_n) forme bien une chaîne de Markov.
b) Dessiner le graphe des transitions correspondant.
c) Etablir la matrice des transitions.
d) Pendant quel pourcentage de temps, la station est-elle en moyenne inoccupée?
e) Si au départ la station est inoccupée, quelle est la durée moyenne jusqu'à ce que les deux places soient occupées pour la première fois?

B – PROCESSUS STOCHASTIQUES A TEMPS CONTINU

B1. On suppose que la durée de vie d'un type de dispositif technique est exponentielle de paramètre $\lambda = 1h^{-1}$.Dés qu'un dispositif tombe en panne, il est immédiatement remplacé par un dispositif identique.
Quelle est la probabilité d'avoir plus de 3 pannes en 2 heures?
Quelle est la distribution de l'instant d'occurrence de la première panne sachant que le deuxième dispositif fonctionne encore au bout de 3 heures?
Si l'on dispose de 9 dispositifs, quelle est la probabilité que le temps total de service soit au moins égal à 12 h?

B2. Dans un système d'attente du type M/M/1 un client arrive en moyenne toutes les 12 minutes et la durée moyenne de service est de 8 minutes.

a) Calculer la probabilité pour que 2 clients au moins attendent d'être servis.
b) Calculer le nombre moyen de clients dans le système?
c) Déterminer le temps de séjour moyen d'un client dans le système?
d) Quelle est la probabilité pour qu'un client attende moins de 30 minutes avant d'être servi? qu'il séjourne moins de 30 minutes dans le système?
e) Le taux d'arrivée des clients augmente de 20%. Calculer le pourcentage de l'augmentation du nombre et du temps de séjour moyen des clients dans le système.

B3. Dans un système d'attente du type M/M/1 le système accueille en moyenne 20 clients par heure.

a) Quelle doit être la durée moyenne de service pour que la probabilité qu'un client qui arrive et doit attendre ne dépasse pas 0,5 ?
b) Quelle est alors la probabilité qu'un client qui arrive et doit attendre ait devant lui une file d'attente formée de n clients?

B4. Dans un cabinet médical, l'arrivée des patients est décrite par un processus de Poisson de paramètre $\lambda = 4h^{-1}$. La durée de traitement suit une loi exponentielle de moyenne 12 min. La salle d'attente ne contenant qu'un client, calculer la probabilité qu'une personne qui arrive soit traitée :

B5. Un atelier comprend 3 machines automatiques qui sont entretenues par un opérateur qui ne peut réparer qu'une machine à la fois. La durée de bon fonctionnement de chaque machine suit une distribution exponentielle de paramètre λ, et le temps de réparation une distribution exponentielle de paramètre μ. Pour $\lambda = \frac{\mu}{3}$ calculer :

- la probabilité que l'opérateur soit libre
- le nombre moyen de machines qui fonctionnent.

B6. L'arrivée des clients dans une banque suit un processus de poisson dont le taux moyen est de 9 clients par heure. La durée de service par client suit une loi exponentielle de moyenne 10 min.

a) Calculer le nombre minimal a de guichets nécessaires pour assurer un régime stationnaire de ce processus, puis le temps d'attente moyen si $s = a$ ou si $s = a+1$.

b) Dans cette question $s = a + 1$.

1. Quelle est la probabilité qu'un client n'attende pas ?
2. Calculer la probabilité que le temps d'attente soit inférieur à 5 minutes.

B7. Un central téléphonique d'une petite entreprise comprend deux lignes d'entrée. Les appels suivent un processus de Poisson dont le taux moyen est λ, la durée des appels suit une loi exponentielle de paramètre μ.

a) Calculer la distribution stationnaire du système.

b) Si $\lambda = \mu$ calculer le pourcentage d'appels perdus.

c) On remplace les opérateurs par un répondeur. On a calculé que $\lambda = 10$ et on veut que le pourcentage d'appels perdus n'excède pas 10%. Déterminer la durée maximale du message à enregistrer.