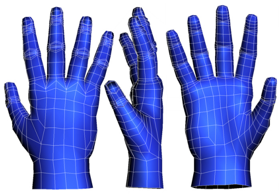
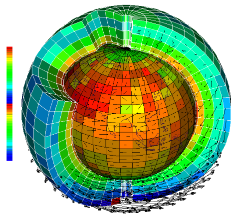
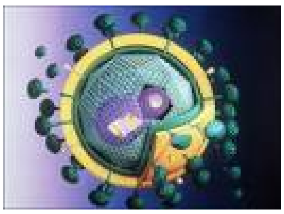
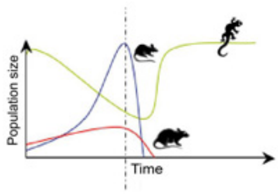
Mathématiques pour les Sciences de la Vie
Parcours Progressif

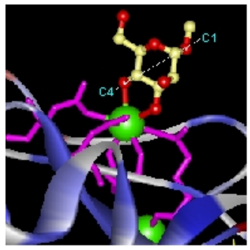
http://mathsv.univ-lyon1.fr





Probabilités et Statistique
Printemps 2021



Enseignants

Isabelle AMAT
Sandrine CHARLES
Anne-Béatrice DUFOUR
Christelle LOPES
Sylvain MOUSSET
Laurence MOUTON

LBBE

Biométrie et Biologie Évolutive
UMR CNRS 5558
http://lbbe.univ-lyon1.fr

Université Claude Bernard - Lyon 1

http://www.univ-lyon1.fr

Pourquoi étudier les mathématiques dès le début de la formation d'un biologiste ?

L'université, et particulièrement Lyon 1, a le devoir, et le souci, de former ses étudiants pour que leur insertion professionnelle dans le monde du travail leur permette de s'adapter aux nombreux changements qui ne manqueront pas de se produire dans les décennies à venir. Les futurs biologistes que vous serez devront être compétents dans plusieurs disciplines et pas seulement en biologie. Que ce soit la génétique, la génomique, l'écologie ou d'autres branches de la biologie, toutes nécessitent une approche quantitative, s'appuyant sur des données.

L'Université Claude Bernard Lyon 1 est réputée pour donner à ses étudiants biologistes une solide formation en mathématiques appliquées à la biologie. Il ne s'agit pas d'enseigner les mathématiques pour elles-mêmes, mais pour vous aider à résoudre des questions de nature biologique ; vos enseignants dans ce domaine possèdent une double compétence : en biologie et en mathématiques appliquées (modélisation et statistiques essentiellement). Nous savons tous que l'étude des mathématiques appliquées à des questions biologiques nécessite des efforts ; elle requiert également d'être appréhendée chaque année du cursus universitaire. C'est pourquoi nous commençons ce type d'enseignement très tôt, dès la première année, en le proposant ensuite chaque année. Une maturation progressive s'opère alors chez les étudiants qui découvrent, souvent en fin de cursus, la pertinence de cette approche, et en sont totalement convaincus lorsqu'ils sont confrontés aux demandes de leur employeur.

Domitien Debouzie, ancien Président de l'Université Lyon 1 (2002-2007).

Ont contribué à la réalisation de ce polycopié

Isabelle Amat		Alexis Avril
Marc Bailly-Bechet		Sandrine Charles
Marie Fablet		Vincent Foray
David Fouchet		Lucie Froissart
Aurore Gallot	Emmanuelle Gilot-Fromont	Laurent Guéguen
	Jean Lobry	
Janice Kielbassa	Christelle Lopes	Jodie Martin
Julien Martinez	Arnaud Mary	Sylvain Mousset
Laurence Mouton		Anne Nguyen
Pierre-François Pélisson	Olivier Raymond	Marie-Claude Venner
Samuel Venner	Raquel Tavares	Aurélien Vigneron

Probabilités simples. Probabilités conditionnelles. Théorème de Bayes

1 Séance 1

1.1 Course de chevaux (1-1*)

Dans une course de 20 chevaux quelle est la probabilité, en jouant 3 chevaux, de gagner le tiercé dans l'ordre ? Dans le désordre ?

1.2 Transistors (2-1*)

Un appareil contient six transistors dont deux exactement sont défectueux. On les identifie en testant les transistors l'un après l'autre. Le test s'arrête quand les deux transistors défectueux sont trouvés. Calculez la probabilité pour que le test :

1. soit terminé au bout de deux opérations ;
2. nécessite plus de trois opérations.

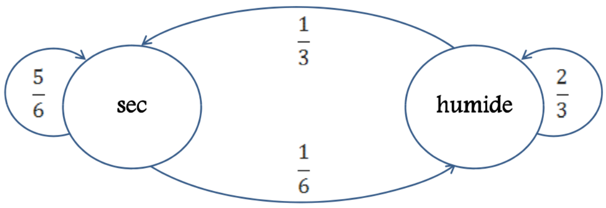
1.3 Groupes sanguins A, B, AB, O (2-1**)

Les groupes sanguins AB, A, B et O. A et B sont dominants par rapport à O.

1. Déterminer, en fonction des fréquences géniques p, q et r de A, B et O dans la génération F_0, les fréquences des quatre groupes sanguins dans la génération F_1.
2. Déterminer les fréquences géniques en F_1.

1.4 Météo en Markovie (2-3**)

La figure ci-dessous nous explique le comportement du temps en Markovie. Par exemple, s'il fait sec aujourd'hui, alors il fera sec demain avec une probabilité de $\frac{5}{6}$ ou il pleuvra avec une probabilité de $\frac{1}{6}$.



Aujourd'hui mercredi, le temps est sec. Quelle est la probabilité d'avoir samedi un temps humide ?

1.5 Dessert au resto U (2-4**)

On suppose que dans un restaurant universitaire on propose deux desserts à chaque repas. La probabilité que l'un des deux soit un yaourt est 0,4, une orange 0,8. La probabilité que les deux desserts soient un yaourt et une orange est 0,3.
Calculez la probabilité que l'on propose :

1. un yaourt et pas d'orange ?
2. une orange et pas de yaourt ?
3. ni yaourt ni orange ?

1.6 Pièce jetée dans un carré (2-3***)

Un jeu consiste à lancer une pièce (diamètre 3 cm) sur une table recouverte de carrés de 4 cm de côté. On gagne si la pièce tombe entièrement à l'intérieur d'un carré.

1. Quelle est la probabilité de gagner à ce jeu ?
2. En cas de gain le joueur reçoit 15 fois sa mise. Ce jeu est-il honnête ?

2 Séance 2

2.1 Irradiation de vers à soie (2-2*)

L'irradiation (rayon X) d'embryons de Vers à soie induit certaines anomalies. La probabilité d'une certaine anomalie est $p = \frac{1}{10}$.

a. Quelle est la probabilité de trouver au moins un embryon présentant cette anomalie sur dix disséqués ?

b. Combien faut-il en disséquer pour trouver au moins une anomalie avec une probabilité supérieure à 0,5 ?

2.2 Urnes et boules (2-5*)

Une urne contient 10 boules, 7 rouges et 3 vertes.
On tire 2 boules avec remise ; quelle est la probabilité de tirer 2 boules rouges ?
On tire 2 boules sans remise ; quelle est la probabilité de tirer 2 boules rouges ?

2.3 Pièces défectueuses (3-1**)

Trois machines A, B et C fournissent respectivement 50%, 30% et 20% de la production d'une usine. Les pourcentages de pièces défectueuses sont respectivement 3%, 4% et 5%. Quelle est la probabilité qu'une pièce prise au hasard dans la production, soit défectueuse ?

2.4 Petites ailes, yeux blancs (3-2**)

Dans une certaine population de mouches, 40% ont des petites ailes, 25% ont les yeux blancs, et 15% ont à la fois des petites ailes et les yeux blancs.
On extrait au hasard un individu de cette population.

1. S'il a des petites ailes, quelle est la probabilité qu'il ait aussi les yeux blancs ?
2. S'il a les yeux blancs, quelle est la probabilité qu'il n'ait pas de petites ailes ?
3. Quelle est la probabilité qu'il n'ait ni de petites ailes, ni les yeux blancs ?

2.5 Dépistage du cancer (3-4**)

Un test de dépistage du cancer est :

- Positif avec 96% de certitude
- Négatif avec 94% de certitude.

Le résultat de mon test est positif. Sachant qu'à mon âge 1 personne sur 145 a le cancer, quelle est la probabilité que j'aie effectivement le cancer ? Discutez qualitativement de ce résultat.

2.6 Jeu télévisé (3-6**)

Dans un jeu télévisé, un candidat doit choisir une question de repêchage en tirant un papier au hasard parmi trois papiers. Il y a

— Une question facile (3 chances sur 4 de donner la réponse exacte).

— Une question moyenne (2 chances sur 5 de donner la réponse exacte).

— Une question difficile (1 chance sur 5 de donner la réponse exacte).

Sachant que le candidat a donné la réponse exacte à la question qu'il a tirée, quelle est la probabilité conditionnelle que la question tirée ait été facile ?

2.7 Tests en série (3-2***)

Devant un certain "tableau clinique", on estime qu'une personne a 6 chances sur 10 d'être atteinte d'une maladie. On effectue alors deux tests biologiques.
Le premier est positif à 70% sur les malades et à 20% sur les non malades. Le second est positif à 90% sur les malades et à 30% sur les non malades.
On suppose que les deux tests sont indépendants. Quelle est la probabilité conditionnelle que le deuxième test soit positif si le premier l'a été ?

3 Séance 3

Question 1 Si deux événements A et B de l'univers des possibles Ω sont indépendants, alors:

- ☐ $P(A \cap B) = 1 - P(A \cup B)$
- ☐ $P(A \cap B) = P(A) \times P(B)$
- ☐ $P(A \cap B) = 0$
- ☐ $P(A \cap B) = P(A) + P(B)$

Question 2 Soit A un événement de l'univers des possibles Ω, alors

- ☐ $P(\bar{A}) = P(A) - 1$
- ☐ $P(\bar{A}) = P(A)^{-1}$
- ☐ $P(\bar{A}) = \dfrac{P(A)}{P(\Omega)}$
- ☐ $P(\bar{A}) = 1 - P(A)$

Question 3 Si A et B sont deux évènements tels que $A \cap B = \varnothing$, que vaut $p(A \cup B)$?

☐ $p(B)p(B|A)$ ☐ $p(A)p(B)$ ☐ $p(A)p(B|A)$ ☐ $p(A) + p(B)$

Question 4 Une roulette de casino possédant 37 cases numérotées de 0 à 36 est truquée de telle manière que la probabilité de tomber sur le chiffre 0 est 10 fois celle de tomber sur un autre numéro donné quelconque entre 1 et 36.
Quelle est la probabilité que le chiffre 25 sorte avec cette roulette?

☐ $\dfrac{1}{10}$ ☐ $\dfrac{1}{46}$ ☐ $\dfrac{1}{36}$ ☐ $\dfrac{1}{45}$

Question 5 On dispose de trois pièces de monnaie truquées. La probabilité de tomber sur pile est de:

- 0.8 pour la pièce 1,
- 0.3 pour la pièce 2,
- 0.6 pour la pièce 3.

On choisit une pièce au hasard (de façon équiprobable) et on la lance.
La pièce est tombée sur pile. Quelle est la probabilité d'avoir choisi la pièce 3?

☐ 0.6 ☐ 0.6471 ☐ 0.3333 ☐ 0.3529

Question 6 Soient A et B deux événements quelconques.
Laquelle des propositions suivantes est-elle toujours vraie?

- ☐ $P(\bar{A} \cap \bar{B}) = 1 - P(A \cap B)$
- ☐ $P(\bar{A} \cap \bar{B}) = P(A \cup B)$
- ☐ $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B})$
- ☐ $P(\bar{A} \cap \bar{B}) = P(\bar{A}) \times P(\bar{B})$

Question 7 $\{A, B, C, D\}$ forment un système complet d'événements de l'univers des possibles. E est un autre événement de l'univers des possibles.
Laquelle de ces propositions est fausse?

- ☐ $P(E) = P(A \cap E) + P(B \cap E) + P(E|C)P(C) + P(E|D)P(D)$
- ☐ $P(E) = P(B)P(E|B) + P(\bar{B})P(E|\bar{B})$
- ☐ $P(E) = P(E \cap A) + P(E \cap B) + P(E \cap C) + P(E \cap D)$
- ☐ $P(E) = P(A|E)P(A) + P(B|E)P(B) + P(C|E)P(C) + P(D|E)P(D)$

Question 8 Dans un système complet d'événements, les événements sont toujours deux à deux:

☐ indépendants ☐ liés ☐ incompatibles ☐ complémentaires

Question 9 Soit A et B deux événements de l'univers des possibles tels que $P(B) \neq 0$ et $P(A) \neq 0$. Parmi les égalités suivantes, laquelle est toujours vérifiée?

- ☐ $P(A|B) + P(\bar{A}|B) = 1$
- ☐ $P(A \cap B) + P(\bar{A} \cap B) = P(A)$
- ☐ $P(A \cap B) + P(\bar{A} \cap B) = 1$
- ☐ $P(A|B) + P(\bar{A}|B) = P(B)$

Question 10 On lance simultanément 3 dés à 6 faces numérotées de 1 à 6, et on les somme. Quelle est la probabilité que la somme soit strictement inférieure à 5 ou strictement supérieure à 16?

☐ $\frac{8}{6^3}$ ☐ $\frac{20}{6^3}$ ☐ $\frac{20}{6^3}$ ☐ $\frac{1}{6^3}$

Question 11 On donne les probabilités suivantes d'être atteint de deux maladies données A et B:

— 17 personnes sur 100 sont atteintes de la maladie A.

— 12 personnes sur 100 sont atteintes de la maladie B.

— 78 personnes sur 100 ne sont atteintes d'aucune des deux maladies.

Quelle est la probabilité d'être atteinte à la fois des maladies A et B?

- ☐ 0.07
- ☐ 0.05
- ☐ Les données sont insuffisantes ou incohérentes
- ☐ 0.0204
- ☐ 0.29

Question 12 Soient A et B deux événements de l'univers des possibles Ω, alors:

- ☐ $P(A|B) = P(A \cup B) \times P(B)$
- ☐ $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$
- ☐ $P(A|B) = P(A \cap B) \times P(B)$
- ☐ $P(A|B) = \dfrac{P(A \cup B)}{P(B)}$

Question 13 Quatre espèces d'insectes morphologiquement non identifiables au stade larvaire V, G, E et P pondent dans les glands des mêmes chênes. Elles peuvent héberger une même bactérie endosymbiotique B.

— Les abondances relatives des espèces d'insectes sont 30% (espèce V), 38% (espèce G), 30% (espèce E) et 2% (espèce P).

— La prévalence (probabilité pour un insecte d'héberger la bactérie) de la bactérie B a été déterminée selon les espèces à 1% (espèce V), 5% (espèce G), 25% (espèce E), et 80% (espèce P).

Quelle est la probabilité qu'une larve ne soit pas infestée par B?

☐ 0.887 ☐ 0.113 ☐ 0.868 ☐ 0.722

Question 14 À bord du Titanic, la proportion de membres d'équipage était de 40%. Après le naufrage, la probabilité de survie des membres de l'équipage était de 0.24; celle des autres passagers était de 0.38.
Le Carpathia, premier navire sur les lieux de la catastrophe, a embarqué tous les survivants.
Quelle est la probabilité qu'un individu secouru par le Carpathia soit un membre d'équipage?.

☐ 0.296 ☐ 0.704 ☐ 0.144 ☐ 0.24

Question 15
Un fabricant d'alcootests propose à la gendarmerie nationale l'alcootest de caractéristiques techniques suivantes:

— Sensibilité: 90% des individus ivres sont déclarés positifs par ce test.

— Spécificité: 91% des individus sobres sont déclarés négatifs par ce test.

La population des automobilistes à tester comporte 2% de personnes ivres.
Quelle est la probabilité qu'une personne contrôlée positive par ce test soit réellement ivre?

☐ 0.8305 ☐ 0.9091 ☐ 0.9 ☐ 0.1695 ☐ 0.9802

Question 16 Des scientifiques ont étudié la présence de bactéries endosymbiotiques des genres *Buchnera* et *Wolbachia* chez un puceron du genre *Cinara*.
Dans la population étudiée, la probabilité pour un individu d'héberger des bactéries du genre *Wolbachia* était de 40%; la probabilité d'héberger des bactéries du genre *Buchnera* était de 5%; la probabilité d'héberger l'un ou l'autre de ces deux endosymbiotes était de 44%.
Quelle est la probabilité pour un individu d'héberger à la fois les deux genres de bactéries endosymbiotiques?

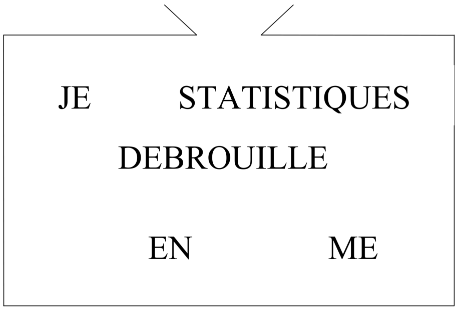
☐ 0.44 ☐ 0.02 ☐ 0.01 ☐ 0.0227

Variables aléatoires discrètes et loi Normale

4 Séance 4

4.1 Loi quelconque (1-1*)

Une épreuve consiste à tirer au hasard un mot de l'urne ci-dessous.



1. Ecrire l'ensemble Ω des résultats.
2. On définit les variables aléatoires suivantes :
   X : longueur du mot tiré
   Y : nombre de voyelles dans le mot tiré
   Ecrire la loi de probabilité de X et de Y.
   Calculer la moyenne et la variance de X et de Y.

4.2 3 lancers d'un dé (1-2**)

Une épreuve consiste à lancer un dé à 6 faces non pipé. Sur Ω on définit les deux événements suivants :

$$A = \{1, 2\}\,;\; B = \{3, 4, 5, 6\}.$$

Au cours de 80 séries de trois épreuves, on obtient les résultats suivants :

$(3A)$	$(2A, 1B)$	$(1A, 2B)$	$(3B)$
2	19	37	22

Comparer graphiquement les résultats observés aux résultats attendus.

4.3 Incendies dans une ville (1-7**)

Dans un village, il y a en moyenne deux incendies par an. Qelle est la probabilité pour qu'il y ait plus de quatre incendies dans l'année à venir ?

4.4 Groupe sanguin AB (1-10**)

Le groupe sanguin AB^- est présent chez 0.6% des individus. Lors d'une collecte de sang, on a recueilli 200 flacons. Quelle est la probabilité de trouver parmi ces 200 prélèvements :

1. Aucun flacon AB^-.
2. Au moins un flacon AB^-.

4.5 Groupe sanguin AB (1-11**)

Le groupe sanguin AB^- est présent chez 0.6% des individus. Combien faudrait-il faire de prélèvements pour que la probabilité de trouver au moins un flacon AB^- soit :

1. supérieure à 0.95.
2. supérieure à 0.99.

4.6 Test sur la prédisposition (1-7***)

Un test biologique T servant à détecter la prédisposition à une maladie M possède 4 résultats possibles, symbolisés par 0, +, ++ et +++. Les probabilités conditionnelles d'atteinte par la maladie M dans l'année qui suit le test sont les suivantes :

T	0	+	++	+++
P(M/T)	0,1	0,2	0,4	0,8

Les probabilités des résultats du test pour un individu pris au hasard sont les suivantes :

T	0	+	++	+++
P(T)	0,5	0,35	0,1	0,05

1. Calculer la probabilité P(M) pour qu'un individu, pris au hasard, soit atteint dans l'année par la maladie M.
2. Trois stratégies sont possibles (tous les coûts donnés sont des coûts annuels par individu) :
   a) On n'effectue ni test ni surveillance des individus ; dans ces conditions, le coût pour la collectivité de l'atteinte par M dans l'année est égal à 1000 U.
   b) On ne fait pas de test mais on effectue une surveillance médicale systématique dont le coût est 120 U ; dans ces conditions, la maladie M, diagnostiquée plus tôt, coûte seulement 450 U.
   c) On effectue un test systématique, dont le coût est de 20 U, et on limite la surveillance médicale (dont le coût est toujours 120 U) aux individus pour lesquels le résultat du test est ++ ou +++ (et la maladie M coûte alors 450 U ou 1000 U selon qu'elle survient chez un individu surveillé ou non).

On désigne par Xa, Xb et Xc les trois variables aléatoires égales au coût annuel par individu respectivement dans les trois stratégies a), b) et c). Calculer les trois espérances mathématiques $E(Xa)$, $E(Xb)$ et $E(Xc)$. Indiquer la meilleure stratégie.

5 Séance 5

5.1 n grand, p petit

Dans une population très nombreuse, des études régulières ont montré qu'il y avait 2% d'individus de type A. Calculez la probabilité dans un échantillon de 100 individus tirés au hasard d'obtenir :

1. Aucun individu de type A.
2. Au moins deux individus de type A.

5.2 Suspension bactérienne

Un liquide contient 10^5 bactéries par litre, réparties au hasard. On en prélève 1mm^3.

1. Quelle est la probabilité que ce prélèvement ne contienne aucune bactérie ?
2. Quelle est la probabilité qu'il contienne au moins trois bactéries ?

5.3 Récolte de pollen

Une abeille récolte du pollen en volant de fleur en fleur. Lorqu'elle rencontre une fleur qui n'a pas été visitée, elle la récolte. En moyenne, 98 % des fleurs ont déjà été visitées. L'abeille a visité 52 fleurs avant son retour à la ruche.

1. Quelle est la probabilité qu'elle ait récolté 2 fleurs ?
2. Quelle est la probabilité qu'elle ait récolté au moins une fleur ?
3. Combien aurait-elle dû visiter de fleurs pour être sûre à 99 % d'en récolter au moins une ?

5.4 convergenceloi-2

3000 canards en vol se répartissent aléatoirement sur les 60 étangs d'une zone marécageuse.

1. Quelle est la probabilité que plus de 50 canards se trouvent sur un étang donné ?
2. Quelle est la probabilité qu'il y en ait moins de 30 ?

5.5 Pièce de monnaie lancée 500 fois (3-1*)

Une pièce de monnaie est lancée 500 fois. Trouver la probabilité que le nombre de faces obtenues ne diffère pas de 250 de plus de 30.

5.6 Probabilité de germination

Dans un certain type de graine, la probabilité de germination est $p = 0.8$. Une personne sème 400 de ces graines. Calculez la probabilité pour qu'au moins 300 graines germent.

6 Séance 6

6.1 Loi normale centrée réduite (2-4**)

T suit une loi normale $\mathcal{N}(0, 1)$.

1. Calculer les probabilités des évènements $T < 2$, $-1 < T < 0.5$ et $4T > -3$.
2. Calculer les réels a et b tels que l'on ait $P(|T| < a) = 0.82$ ou $P(T < -b) = 0.61$.

6.2 Loi normale (2-5**)

X suit une loi normale $\mathcal{N}(\mu = -1, \sigma = 2)$.

1. Calculer les probabilités des évènements :
   — $X < -1$
   — $X > 1$
   — $-3 < X < 1$
2. Déterminer le réel $a > 0$ tel que $P(X > -a) = 0.4$.

6.3 Paramètres d'une loi normale (2-6**)

X suit une loi normale telle que $P(X < 3) = P(X > -1) = 0.8413$. Déterminer ses paramètres.

6.4 Résistances (2-8**)

Une machine fabrique des résistances dont la valeur en ohms suit une loi normale $N(100, 3)$. Une seconde machine fabrique d'autres résistances dont la valeur en ohms suit une loi normale $N(200, 4)$.

a Quelle est la loi suivie en montant deux de ces résistances en série ?

b Entre quelles valeurs va se situer cette nouvelle résistance, avec une probabilité de 0.95 ?

6.5 Dé lancé 120 fois (3-2*)

Un dé non pipé est lancé 120 fois. Trouver la probabilité que la face "quatre" apparaisse au plus, quatorze fois.

7 Séance 7

Question 17 Soit une variable aléatoire X. Son espérance vaut -3.9 et sa variance vaut 6. On définit une nouvelle variable aléatoire Y telle que $Y = \dfrac{X+9}{3}$.
Quelle est l'espérance de Y?

☐ $\frac{5.1}{3}$ ☐ $\frac{5.1}{9}$ ☐ $\frac{15}{3}$ ☐ $\frac{-3.9}{3}$

Question 18 Soit une variable aléatoire X. Son espérance vaut 7.6 et sa variance vaut 7. On définit une nouvelle variable aléatoire Y telle que $Y = \dfrac{X+2.4}{7}$.
Quelle est la variance de Y?

☐ $\frac{9.4}{49}$ ☐ $\frac{7}{49}$ ☐ $\frac{7}{7}$ ☐ $\frac{9.4}{7}$

Question 19 Cette question possède une unique bonne réponse qui rapporte 1 point si elle est choisie. L'espérance de gain d'un étudiant qui choisirait de façon équiprobable une réponse au hasard parmi les 7 proposées est de 0 points. Quelle est la pénalité (en points) associée à une mauvaise réponse?

☐ $-\frac{1}{5}$ ☐ $-\frac{1}{7}$ ☐ $-\frac{1}{6}$ ☐ $-\frac{1}{9}$ ☐ $-\frac{1}{4}$ ☐ $-\frac{1}{8}$ ☐ $-\frac{1}{3}$

Question 20 Une compagnie d'assurances assure 470 navires pour une somme de 8 millions d'euros chacun. Chaque navire a, par an, une probabilité de couler égale à 3×10^{-3}. S'il ne coule pas, un navire est considéré en état de marche. On note X la variable aléatoire associée au nombre de navires assurés par cette compagnie coulés pendant un an.
Quelle loi la variable aléatoire X suit-elle?

☐ la loi de binomiale $\mathcal{B}(n = 24000, p = 3 \times 10^{-3})$
☐ la loi binomiale $\mathcal{B}(n = 470, p = 1.41)$.
☐ la loi de Poisson $\mathcal{P}(\lambda = 3 \times 10^{-3})$
☐ la loi de Poisson $\mathcal{P}(\lambda = 1.41)$.

Question 21 Une maladie congénitale rare appelée DICV (Déficit Immun Commun Variable) qui se traduit par un déficit en immunoglobulines affecte en France un nourrisson sur 100 000 naissances.
Dans un hôpital où naissent 180 bébés par an, quelle loi de distribution pourriez-vous proposer pour le nombre de nourrissons naissant atteints de DICV par an?

☐ La loi de Poisson $\mathcal{P}(\lambda = 1.8 \times 10^{-3})$
☐ La loi Normale $\mathcal{N}(\mu = 1.8 \times 10^{-3}, \sigma^2 = 1.8 \times 10^{-3})$
☐ La loi de Poisson $\mathcal{P}(\lambda = 10^{-5})$

Question 22 *Pinctada margaritifera* est une huître perlière utilisée pour la culture des perles noires de Tahiti. Pour produire une perle, l'huître est greffée avec un morceau du manteau d'une autre huître et un nucleus (petit morceau de coquillage qui sert de noyau aux sécrétions de nacre du greffon).
Dans une ferme d'huîtres perlières, le taux de succès de la greffe est de 33%. Un greffeur professionnel traite les huîtres par lots de 9.
Quelle est la probabilité pour qu'exactement 6 greffes aient échoué sur un lot de 9 perles?

☐ 0.6146 ☐ 0.0326 ☐ 0.0905 ☐ 0.2731

Question 23 Dans le cadre des travaux pratiques de génétique, des étudiants ont entrepris de réaliser les expériences originales de Gregor Mendel sur la coloration de l'albumen du pois. À la génération F2, ils attendent $\frac{3}{4}$ de pois jaunes et $\frac{1}{4}$ de pois verts.
36 étudiants ont compté le nombre de pois verts dans 36 échantillons indépendants de 160 pois de la génération F2.
Soit X la variable aléatoire "nombre de pois verts dans un échantillon de 160 pois de la génération F2".
Quelle est la variance de X?

☐ 9 ☐ 40 ☐ 6.75 ☐ 30

Question 24 La variable aléatoire X peut prendre quatre valeurs (2, 5, 6, 10) avec les probabilités suivantes:

Valeur k	2	5	6	10
Probabilité $P(X = k)$	0.4	0.2	0.3	0.1

Quelle-est l'espérance de X?

☐ 6.24 ☐ 4.6 ☐ 5.75 ☐ 1.15

Question 25 La variable aléatoire X peut prendre quatre valeurs (1, 3, 5, 6) avec les probabilités suivantes:

Valeur k	1	3	5	6
Probabilité $P(X = k)$	0.2	0.1	0.4	0.3

Quelle-est la variance de X?

☐ 3.41 ☐ 11.12 ☐ 3.6875 ☐ 21.9

Question 26 La variable aléatoire X suit la loi Normale $\mathcal{N}(\mu = 1.56, \sigma = 0.8)$.
Que vaut $P(1.44 \leqslant X \leqslant 2.32)$?

☐ 0.829 ☐ 0.389 ☐ 1.269 ☐ 0.56

Statistique descriptive et intervalle de confiance

8 Séance 8

8.1 Trajectométrie (1-1*)

On a mesuré la distance parcourue par un insecte en 30 secondes. Pour un lot de 10 insectes les résultats ont été les suivants (en mm) :

78-170-173-190-90-174-166-293-149-117

Calculez la moyenne et la variance des distances parcourues.

8.2 Galles du hêtre (1-3*)

La cécidomyie du hêtre provoque sur les feuilles de cet arbre des galles dont la distribution à été observée :

Nombre x de galles par feuille	0	1	2	3	4	5	6	7	8	9	10
Nombre n de feuilles portant x galles	482	133	46	24	6	5	2	1	0	1	0

a Représenter graphiquement ces données.

b Calculer la moyenne et la variance et l'écart type du nombre de galles par feuille

8.3 Distances entre domiciles (1-4*)

Une enquête concernant les distances entre domiciles des époux, au moment de leur mariage, a donné, dans le Finistère, les résultats suivants :

Distance (km)	Nbre Couples
0-2	138
2-4	384
4-6	210
6-8	103
8-10	63
10-12	28
12-14	20
14-16	19
16-18	12
18-20	9

1. Représenter graphiquement les données.

2. Calculer la moyenne et la variance des distances.

8.4 Hématimètre (1-3**)

Dans les 400 carrés d'un "hématimètre", on a compté 1872 levures, avec la répartition suivante :

Nombre de levures par carré	0 ou 1	2	3	4	5	6	7	8	9	10 ou plus
Nombre de carrés	20	43	53	86	70	54	37	18	10	9

1. Calculer le nombre moyen de levures par carré.
2. Quelle loi pourrait suivre le nombre de levures par carré ?

9 Séance 9

9.1 Taille des étudiants (2-1*)

Supposons que les tailles de 100 étudiants masculins de l'Université de BROLNEGE représentent un échantillon aléatoire des tailles des 4521 étudiants de cette université. Sachant que dans cet échantillon la taille moyenne vaut 1.73 m et l'écart-type 10 cm, déterminer un intervalle de confiance à 0.95 de la taille moyenne des étudiants de l'université.

9.2 Coquillages (2-8**)

Sur un échantillon de 83 coquillages on a relevé une cote déterminée. On a trouvé une moyenne de 52,17mm et un écart-type de 1,8mm.

1. Estimer la moyenne par un intervalle de confiance à 95%, à 98% et à 99%.

2. Que devrait être l'effectif de l'échantillon pour situer la moyenne dans un intervalle de 0,4mm avec une sécurité de 95% ?

9.3 Poids de cocons (2-2**)

Dans un lot de 500 cocons environ d'une race bivoltine de Bombyx mori, on a prélevé 20 cocons au hasard et on les a pesés. Les poids (en g) ainsi mesurés sont les suivants :

0,64	0,65	0,73	0,60	0,65	0,77	0,82	0,64	0,66	0,72
0,87	0,84	0,66	0,76	0,63	0,52	0,66	0,45	0,74	0,79

1. Calculer le poids moyen et la variance de cet échantillon.
2. Donner l'intervalle de confiance au seuil 5% du poids moyen d'un cocon de la population considérée

9.4 Variabilité d'une pesée (2-3**)

On admet qu'avec une balance utilisée dans un laboratoire, le résultat de la pesée d'un corps de masse μ comprise entre 50 g et 100 g est une variable aléatoire X de distribution normale $\mathcal{N}(\mu, \sigma = 0.08\,\text{g})$.

Dans le cas où $\mu = 72.37\,\text{g}$:

1. Calculez la probabilité de l'événement $72.30 < X < 72,50$.
2. Déterminez un intervalle $[\mu - h, \mu + h]$ tel que la probabilité que X prenne une valeur dans cet intervalle soit 0.99.

9.5 Elements figurés du sang (2-9**)

La mesure en cc du volume des éléments figurés (EF) dans 100cc de sang est en moyenne de 41.75 sur toute une population. On a effectué des mesures sur 100 sujets, résumées dans le tableau ci-dessous :

Volume des EF	[38.5 ;39.5]	[39.5 ;40.5]	[40.5 ;41.5]	[41.5 ;42.5]	[42.5 ;43.5]	[43.5 ;44.5]	[44.5 ;45.5]
Nombre d'individus	3	10	23	28	20	11	5

Ces 100 sujets sont-ils représentatifs de la population ?

9.6 Glucose (2-1***)

Des études statistiques montrent que le taux de glucose dans le sang est une variable Gaussienne X de moyenne 1 g/l et d'écart type 0.1 g/l.

On prend un échantillon de 9 individus. Quelles sont la moyenne et l'écart-type théorique ($\mu_{\bar{X}}$ et $\sigma_{\bar{X}}$) de la variable aléatoire $\bar{X}$? Quelle est la loi de $\bar{X}$?

On désire savoir si un traitement déplace le taux moyen de glucose. La moyenne observée sur un échantillon de 9 individus traités est $\bar{x} = 1.06$ g/l. Quelle hypothèse faites-vous ? Faites le test correspondant.

10 Séance 10

Question 27 Dans le cadre d'une étude écologique, les masses de 12 calamars géants issus d'une population du Pacifique ont été mesurées. Les observations sont les suivantes (masses données en kg).

192 229 171 99 134 136 160 73 125 171 170 118

Quelles sont la moyenne et la variance observées de la masse des calamars de cet échantillon ?

- ☐ $\bar{x} = 148.2$, $s^2 = 1664.8$
- ☐ $\bar{x} = 148.2$, $s^2 = 1816.2$
- ☐ $\bar{x} = 161.6$, $s^2 = 1816.2$
- ☐ $\bar{x} = 161.6$, $s^2 = 1664.8$

Question 28 Un biologiste a pesé $n = 80$ œufs de mésange bleue et a noté leurs masses x_i. On donne:

$$\sum_{i=1}^{80} x_i = 646.4 \qquad \sum_{i=1}^{80} x_i^2 = 5242.22$$

À partir de ces données, proposez une estimation de la moyenne et de la variance de la masse des œufs de mésange bleue dans cette population.

- ☐ $\mu = 8.08$ et $\sigma^2 = 0.244$
- ☐ $\bar{x} = 8.08$ et $s^2 = 0.241$
- ☐ $\hat{\mu} = 8.08$ et $\hat{\sigma}^2 = 0.244$
- ☐ $\hat{x} = 8.08$ et $\hat{s}^2 = 0.241$

Question 29 Dans le numéro de septembre 2008 du journal *Genetics*, on a mesuré la taille de 33 articles en comptant le nombre de pages p imprimées par article. On obtient les données suivantes:

nombre de pages p	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
nombre d'articles n_p	1	1	2	2	4	2	3	2	6	5	0	2	2	0	1

On donne

$$\sum_{p=4}^{18} p n_p = 354 \qquad \sum_{p=4}^{18} p n_p^2 = 1258$$

$$\sum_{p=4}^{18} p^2 n_p = 4166 \qquad \sum_{p=4}^{18} n_p = 33$$

À partir de cet échantillon, estimez l'écart-type (exprimé en nombre de pages) de la taille d'un article dans le journal *Genetics*.

☐ $\hat{\sigma} = 3.39$ ☐ $s^2 = 11.17$ ☐ $s = 3.34$ ☐ $\hat{\sigma}^2 = 11.52$

Question 30 On a réalisé n mesures d'une variable aléatoire X dans un échantillon représentatif d'une population ($n \in \mathbb{N}$, $5 \leqslant n \leqslant 30$). On note $\bar{x}$ la moyenne et s^2 la variance de ces n mesures dans cet échantillon, $\hat{\sigma}$ est l'estimation de l'écart-type de X dans la population. Quelle est la formule donnant l'intervalle de confiance (centré) au risque α de l'espérance de X dans la population dont est tiré l'échantillon?

- ☐ $\bar{x} - t_{\alpha, n-1\,\text{ddl}} \dfrac{s}{\sqrt{n-1}} \leqslant \mu \leqslant \bar{x} + t_{\alpha, n-1\,\text{ddl}} \dfrac{s}{\sqrt{n-1}}$
- ☐ $\mu - t_{\alpha, n-1\,\text{ddl}} \dfrac{s}{\sqrt{n-1}} \leqslant \bar{x} \leqslant \mu + t_{\alpha, n-1\,\text{ddl}} \dfrac{s}{\sqrt{n-1}}$
- ☐ $\bar{x} - \varepsilon_\alpha \dfrac{s}{\sqrt{n-1}} \leqslant \mu \leqslant \bar{x} + \varepsilon_\alpha \dfrac{s}{\sqrt{n-1}}$
- ☐ $\bar{x} - t_{\alpha, n+1\,\text{ddl}} \dfrac{s}{\sqrt{n-1}} \leqslant \mu \leqslant \bar{x} + t_{\alpha, n+1\,\text{ddl}} \dfrac{s}{\sqrt{n-1}}$
- ☐ $\mu - \varepsilon_\alpha \dfrac{s}{\sqrt{n-1}} \leqslant \bar{x} \leqslant \mu + \varepsilon_\alpha \dfrac{s}{\sqrt{n-1}}$
- ☐ $\mu - t_{\alpha, n\,\text{ddl}} \dfrac{s}{\sqrt{n-1}} \leqslant \bar{x} \leqslant \mu + t_{\alpha, n\,\text{ddl}} \dfrac{s}{\sqrt{n-1}}$

Question 31
Pour un site web de partage de photographies, on veut estimer le nombre de photos téléchargées chaque minute par les internautes. Pour $n = 10$ intervalles d'une minute pris au hasard dans la journée, on a compté le nombre de photos téléchargées par les internautes.
On note x_i le nombre de photographies téléchargées durant le i^{eme} intervalle d'une minute. On donne les valeurs suivantes :

$$\sum_{i=1}^{10} x_i = 55522 \qquad \sum_{i=1}^{10} x_i^2 = 309209112 \qquad n = 10$$

En faisant l'approximation que le nombre de photos téléchargées par minute est normalement distribué, donnez un intervalle de confiance au risque de 5% de la moyenne du nombre de photos téléchargées par minute.

- ☐ [5333;5771]
- ☐ [5370;5734]
- ☐ [5321;5783]
- ☐ [5362;5742]
- ☐ [5352;5752]

Question 32 Lorsque l'on pèse un objet de masse m sur une balance, la valeur indiquée par la balance est $m + \varepsilon$ où ε est une variable aléatoire correspondant à l'erreur de mesure.
On fait l'hypothèse que sur une (mauvaise) balance de cuisine, cette erreur suit une loi normale d'espérance 0 g. Avec cette balance, on a mesuré 6 fois la masse d'un même objet. On a obtenu les mesures m_i suivantes (exprimées en kg).

1.524	1.498	1.568	1.528	1.454	1.569

On donne:

$$\sum_{i=1}^{6} m_i = 9.141 \qquad \sum_{i=1}^{6} m_i^2 = 13.935865$$

Donnez un intervalle de confiance à 90% de la masse m de l'objet pesé, exprimée en kg.

- ☐ $1.4875 \leqslant m \leqslant 1.5595$
- ☐ $1.4941 \leqslant m \leqslant 1.5529$
- ☐ $1.4967 \leqslant m \leqslant 1.5503$
- ☐ $1.4907 \leqslant m \leqslant 1.5563$

Tests de conformité à une moyenne théorique

11 Séance 11

11.1 Spécifications d'un médicament

Les spécifications d'un certain médicament indiquent que chaque comprimé doit contenir 2.5 g de substance active.
100 comprimés sont choisis au hasard dans la production puis analysés. Ils contiennent en moyenne $\bar{x} = 2.6$ g de substance active, avec un écart-type observé $s = 0.4$ g.
Peut-on dire que le médicament respecte les spécifications ($\alpha = 0.05$) ?

11.2 Circonférence crânienne

Dans un niveau sédimentaire, un anthropologue a découvert 10 crânes humains de mâles adultes et en a mesuré la circonférence. Il obtient les mesures suivantes :
56.6 54.9 57.7 56.8 57.7 55.6 57.7 56.1 54.8 55.9
Chez *Homo sapiens*, la circonférence du crâne est distribuée normalement autour d'une moyenne de 58 cm avec un écart-type de 1.5 cm. Les crânes trouvés peuvent-ils être des crânes d'*Homo sapiens* ?

12 Séance 12

12.1 Effet d'un traitement sur la masse des rongeurs

À la suite d'un traitement sur une variété de rongeurs, on prélève un échantillon de 5 animaux et on les pèse. On obtient les poids en g.

83 81 84 80 85

À la même époque, un grand nombre de mesures a permis d'établir que les rongeurs non traités avaient un poids moyen de 87.6 g.
Le poids moyen des rongeurs traités diffère-t-il significativement de cette norme au seuil 5% ? On suppose que le poids des rongeurs suit une loi normale.

12.2 Effet d'un traitement sur la pression artérielle

Chez un groupe de 10 malades, on expérimente les effets d'un traitement destiné à diminuer la pression artérielle systolique (en cm Hg).

Sujet n°	1	2	3	4	5	6	7	8	9	10
Avant traitement	15	18	17	20	21	18	17	15	19	16
Après traitement	12	16	17	18	17	15	18	14	16	18

Le traitement a-t-il une action significative au risque $\alpha = 5\%$? On suppose que la variable aléatoire égale à la différence des tensions artérielles suit une loi normale.

12.3 Masses des cerveaux

Les données suivantes ont été obtenues sur des échantillons d'individus d'une région d'Europe. Le caractère étudié est la masse (en g) du cerveau pour des sujets de 20 à 49 ans.

Masses des cerveaux des hommes								
Centre des classes	1170	1220	1270	1320	1370	1420	1470	Total
Effectifs	5	36	45	50	61	49	19	265
Masses des cerveaux des femmes								
Centre des classes	1070	1120	1170	1220	1270	1320	1370	Total
Effectifs	12	22	45	54	52	20	10	215

Calculez

1. Les moyennes observées $\bar{m}_h$ et $\bar{m}_f$ de la masse des cerveaux des hommes et des femmes dans cet échantillon.
2. Les écarts-types observés s_h et s_f de la masse des cerveaux des hommes et des femmes dans cet échantillon.
3. Les moyennes estimées $\hat{\mu}_h$ et $\hat{\mu}_f$ de la masse des cerveaux des hommes et des femmes dans cette population.
4. Les écarts-types estimés $\hat{\sigma}_h$ et $\hat{\sigma}_f$ de la masse des cerveaux des hommes et des femmes dans cette population.
5. Déterminez un intervalle de confiance au risque de 1% :
   pour la moyenne de la population des hommes.
   pour la moyenne de la population des femmes.
6. La masse moyenne du cerveau diffère-t-elle selon le sexe dans cette population au risque de 5% ?.

12.4 Poids moyens des épis de blé

Pour déterminer le poids moyen d'épis de blé appartenant à deux variétés, on a procédé à dix pesées pour chaque variété. Les moyennes obtenues ont été : $\bar{x}_1 = 170.7$cg et $\bar{x}_2 = 168.5$cg.
On admet que le poids de ces graines est distribué dans chaque variété suivant une loi de Gauss et que les variances de ces deux distributions peuvent-être considérées comme égales. Les estimations obtenues sur chaque échantillon sont $s_1^2 = 432.9$ et $s_2^2 = 182.7$.
La différence des moyennes est-elle significative au risque $\alpha = 0.05$?

Tests d'homogénéité de la moyenne

13 Séance 13

13.1 Taille des truites

On a mesuré la longueur de deux échantillons de truites. Le premier échantillon est composé de 50 truites d'élevage et donne $\bar{x}_a = 158.86\,\text{mm}$ et $s_a^2 = 37.18\,\text{mm}^2$. Le second échantillon est composé de 67 truites de rivière et donne $\bar{x}_b = 134.16\,\text{mm}$ et $s_b^2 = 36.92\,\text{mm}^2$.

1. Avec un risque de première espèce $\alpha = 5\%$, la longueur moyenne des truites d'élevage est-elle significativement différente de 157 mm ?
2. Même question avec un risque de première espèce $\alpha = 1\%$.
3. Les moyennes des deux populations diffèrent-elles de manière significative ?

13.2 Succès thérapeutique et fièvre (3**)

On s'intéresse aux résultats d'une thérapeutique chez des malades présentant ou non de la température. Le pourcentage de succès parmi 200 malades fiévreux a été de 72% alors qu'il a été de 88% parmi 100 malades non fiévreux.
Peut-on conclure, au risque 0,02, que les proportions de succès de la thérapeutique sont différentes selon la présence ou non de température chez les malades ?

13.3 Mesures de pucerons (2**)

Pour un groupe de 30 pucerons prélevés dans le Midi de la France, on a relevé les informations suivantes : la longueur x du cornicule gauche (organe excréteur) et la longueur y du troisième segment de l'antenne gauche.

x	400	420	440	460	480	500	520	540
y	320	356	413	418	485	464	457	517
	403	358	394	450	405	460	480	424
	327			379	414	446	511	476
	345			405	387	443		
				433	455	461		
						492		

1. Faire une représentation graphique des données.
2. Calculer, pour chaque variable, la moyenne, la variance et l'écart-type de l'échantillon.
3. Une étude similaire menée en Corse sur 32 pucerons a donné les résultats suivants :

	Moyenne	Variance
Longueur x du cornicule	334,18	1103,24
Longueur y du 3e segment	359,50	683,32

Les moyennes de chaque variable sont-elles significativement différentes au risque 5% entre les deux populations (Midi et Corse) ?

14 Séance 14

14.1 Ségrégation Mendelienne des allèles

On a effectué le coisement de balsamines blanches avec des balsamines pourpres. En première génération les fleurs sont toutes pourpres. On obtient en deuxième génération quatre catégories avec les effectifs suivants :

Couleurs	pourpre	rose	blanc lavande	blanc
Effectifs	1790	547	548	213

Peut-on accepter l'hypothèse de répartition mendélienne $\left(\frac{9}{16}; \frac{3}{16}; \frac{3}{16}; \frac{1}{16}\right)$.

14.2 Yeux des drosophiles (1**)

Sur des drosophiles, on considère le caractère "couleur des yeux". Par croisement d'homozygotes "yeux rouges" AA et d'homozygotes "yeux bruns" aa, on obtient des hétérozygotes Aa. Si on croise ces hétérozygotes entre eux, on doit obtenir (loi de Mendel) $\frac{3}{4}$ d'yeux rouges (AA et Aa) et $\frac{1}{4}$ d'yeux bruns (aa). On a observé chez 309 drosophiles de cette seconde génération

— 241 "yeux rouges".

— 68 "yeux bruns".

1. Cette répartition est-elle conforme à la loi de Mendel ?
2. Combien aurait-il fallu observer de drosophiles aux yeux rouges pour décider de rejeter la conformité à la loi de Mendel avec un risque de 5% ?

14.3 Yesterday

La distribution des lettres entre voyelles et consonnes dans les textes anglais, estimée à partir de la liste des mots anglais contenus dans un dictionnaire informatisé est de 40.4% de voyelles et 59.6% de consonnes.
On a compté les lettres dans deux chansons des Beatles *Yesterday* et *Yellow submarine*. On a observé les résultats suivants :

	Yesterday	*Yellow submarine*
Nombre de voyelles	223	438
Nombre de consonnes	260	522

1. La distribution des lettres entre voyelles et consonnes dans l'échantillon constitué par la chanson *Yesterday* est-elle conforme à la distribution théorique estimée à partir de la concaténation de tous les mots du dictionnaire informatisé ?
2. La distribution des lettres entre voyelles et consonnes dans l'échantillon constitué par la chanson *Yellow submarine* est-elle conforme à la distribution théorique estimée à partir de la concaténation de tous les mots du dictionnaire informatisé ?
3. Quelles hypothèses pouvez-vous faire pour expliquer vos résultats ?
4. Dans le livre *Alice's Adventures in Wonderland* de Lewis Carroll, on a compté 49871 voyelles et 73119 consonnes. À la lumière de ce nouveau jeux de données, quelle hypothèse parmi les précédentes privilégieriez-vous ?

Test d'ajustement à une distribution théorique

15 Séance 15

15.1 Morts par ruade (4-4**)

Dans 10 unités de l'armée prussienne, pendant une période de 20 ans de 1775 à 1794, le nombre de morts par unité d'armée et par année résultant d'une ruade de cheval est donné dans le tableau suivant :

Nombre de morts par an	0	1	2	3	4
Fréquence	109	65	22	3	1

Peut-on considérer que cette mortalité accidentelle suit un processus de Poisson (c'est-à-dire ces données s'ajustent-elles à une distribution de Poisson) ?

15.2 Loi de Poisson dans les coopératives agricoles

Une enquète menée auprès du comptoir d'une coopérative agricole a permis d'étudier l'arrivée dans le temps des usagers de cette coopérative. Cette étude a été menée durant 150 heures.
Pendant l'unité de temps, soit une heure, on a noté :

Nombre d'usagers arrivés	0	1	2	3	4	5	6
Nombre d'occurrence de l'observation	37	46	39	19	5	3	1

Peut-on admettre, au risque de 5%, que le nombre d'usagers arrivant par heure dans cette coopérative agricole suit une loi de Poisson ?

15.3 Indice céphalique (4-1**)

Dans un groupe ethnique, on a observé la répartition de 500 individus, relativement à leur indice céphalique i (largeur du crâne exprimée en pourcentage de la longueur).

On a trouvé :	50 % de dolichocéphales	($i \leqslant 75$)
	40 % de mésocéphales	($75 < i \leqslant 80$)
	10 % de brachycéphales	($i > 80$)

Peut-on considérer que ce groupe appartient à une population où l'indice céphalique est distribué suivant une loi de Gauss (c'est à dire la loi normale) de moyenne $\mu = 74.34$ et d'écart-type $\sigma = 3.22$?

Sujets d'annales

16 Séance 16

16.1 *Marmota marmota*

La dynamique dans une population de marmottes alpines (*Marmota marmota*) est notamment la résultante des processus de mortalité et de naissance. La mortalité hivernale est très dépendante des réserves de graisse accumulées pendant l'été et donc de la masse corporelle en début d'automne. Des données de masse corporelle (en kg) juste avant l'hibernation ont été récoltées et réparties en 5 classes de masse comme indiqué dans le tableau suivant :

Classes de masse corporelle	Effectifs observés	Effectifs théoriques
]3,800 ; 4,200]	74	77,11
]4,200 ; 4,600]	106	101,33
]4,600 ; 5,000]	152	147,59
]5,000 ; 5,400]	94	99,68
]5,400 ; 5,800]	81	81,39

On désire savoir si la variable étudiée suit une loi normale dont les paramètres ont été estimés à partir de la masse moyenne et la variance de cet échantillon.

Question 33 Quel type de test est utilisé ?

- ☐ Test d'indépendance
- ☐ Test d'égalité
- ☐ Test de conformité
- ☐ Test d'ajustement

Question 34 Quelle loi suit la statistique du test sous l'hypothèse nulle ?

- ☐ Loi binomiale
- ☐ Loi normale
- ☐ Loi du χ^2
- ☐ Loi de Poisson

Question 35 Quel est le nombre de degrés de liberté de cette statistique ?

- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 2

Question 36 Quelle est la valeur de la statistique observée ?

- ☐ 0,798
- ☐ 9,49
- ☐ -0,03
- ☐ 5,99
- ☐ 1,96

Question 37 Quelle conclusion tirez-vous du test réalisé ($\alpha = 0{,}05$) ?

- ☐ On ne peut pas rejeter l'hypothèse selon laquelle la masse corporelle des marmottes suit une loi normale
- ☐ On ne peut pas rejeter l'hypothèse selon laquelle la masse corporelle des marmottes suit une loi du χ^2
- ☐ On ne peut pas rejeter l'hypothèse selon laquelle la masse corporelle des marmottes alpines suit une loi de Poisson
- ☐ On ne peut pas considérer que la masse corporelle des marmottes alpines suit une loi normale
- ☐ On ne peut pas considérer que la masse corporelle des marmottes alpines suit une loi du χ^2

Question 38 Quel est le risque associé à cette conclusion ?

- [ ] $1 - \beta$
- [ ] β inconnu
- [ ] 0,05
- [ ] 0,95

16.2 Population de la Vanoise

La conclusion tirée de la partie précédente sera adoptée pour la suite. Un échantillon aléatoire simple de 26 adultes a été prélevé dans la population de la Vanoise juste avant l'entrée en hibernation et ces adultes ont été pesés. Dans cet échantillon, on calcule $\bar{x} = 4,825$ kg et $s^2 = 1,724$ kg^2, respectivement la moyenne et la variance de la masse corporelle.

Question 39 Quelle est la fonction densité de probabilité d'une loi normale quelconque ?

- [ ] $\pi(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du$
- [ ] $\int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$
- [ ] $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$
- [ ] $\frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du$

Question 40 Quelle est l'estimation ponctuelle, $\hat{\mu}$, de μ pour cette population ?

- [ ] $\frac{n}{n-1}\bar{x}$
- [ ] $\bar{x}$
- [ ] $\frac{1}{n-1}\sum_{i=1}^{n} x_i$
- [ ] $\frac{n-1}{n}\bar{x}$

Question 41 Quelle est l'estimation ponctuelle, $\hat{\sigma}$, de l'écart-type de la population ?

- [ ] $1,724$
- [ ] $\sqrt{1,793}$
- [ ] $\sqrt{1,724}$
- [ ] $1,793$

Question 42 Quel est l'intervalle de confiance à 0,95 (c'est-à-dire avec $\alpha = 0,05$) de μ ?

- [ ] [4,11 ; 4,99]
- [ ] [4,69 ; 4,42]
- [ ] [4,28 ; 5,37]
- [ ] [4,31 ; 4,80]

Question 43 Quelle est l'interprétation biologique de cet intervalle de confiance ?

- [ ] On accepte l'hypothèse que $\mu = 4{,}825$ kg avec un risque β inconnu
- [ ] Il y a 95 % de chance pour que la moyenne μ soit égale à 4,825
- [ ] On accepte l'hypothèse que l'intervalle de confiance contienne μ
- [ ] L'intervalle de confiance a 95 % de chance de contenir la valeur moyenne μ de la population
- [ ] L'intervalle de confiance a 5 % de chance de contenir la valeur moyenne μ pour la population

Des pesées ont été effectuées sur un échantillon aléatoire simple de 19 marmottes issues de la population du Mercantour. On donne les éléments suivants : $\sum_{i=1}^{n} x_i = 95$ et $\sum_{i=1}^{n} x_i^2 = 509,01$. On désire savoir si on peut considérer que les masses moyennes dans les populations de Vanoise (μ_1) et du Mercantour (μ_2) sont identiques.

Question 44 Quelle proposition est nécessaire à l'application du test approprié ?

- [ ] Les effectifs doivent être grands (>30)
- [ ] Les variables suivent une loi normale et les effectifs sont petits
- [ ] Les variables suivent une loi normale et les variances σ_1^2 et σ_2^2 sont égales
- [ ] Les variances σ_1^2 et σ_2^2 sont égales
- [ ] Les variances dans les deux populations (respectivement σ_1^2 et σ_2^2) sont connues

Question 45 Quelle est la valeur absolue de la statistique observée ?

- [ ] -0,44
- [ ] 24,04
- [ ] 1,05
- [ ] 0,43

Question 46 Quelle conclusion ($\alpha = 0,05$) tirez-vous du test réalisé à la question précédente ?

- ☐ Les moyennes des deux échantillons sont différentes
- ☐ On ne peut pas conclure avec un risque β inconnu
- ☐ On ne peut pas conclure que les moyennes des deux échantillons sont différentes
- ☐ On ne peut pas conclure que les moyennes des deux populations sont différentes

16.3 Femelles gestantes et dominantes

Dans la population de Vanoise, la probabilité qu'une femelle adulte soit gestante est égale à 0,20. La probabilité qu'une femelle adulte soit dominante est égale à 0,28. La probabilité qu'une femelle dominante soit gestante est égale à 0,66.

Question 47 Que dire des événements ■ être une femelle gestante ■ et ■ être une femelle dominante ■ d'un point de vue probabiliste ?

- ☐ On ne peut pas savoir s'ils sont indépendants ou non
- ☐ Ils sont indépendants
- ☐ Ils sont incompatibles
- ☐ Ils ne sont pas indépendants

Question 48 Quelle est la probabilité qu'une femelle adulte soit gestante et dominante ?

☐ 0,1848 ☐ 0,4242 ☐ 0,1320 ☐ 0,3030

Question 49 Quelle est la probabilité qu'une femelle soit dominante si elle est gestante ?

☐ 0,093 ☐ 0,519 ☐ 0,040 ☐ 0,924

On a pesé 63 femelles gestantes, prélevées de façon aléatoire et simple dans la population de la Vanoise. Dans cet échantillon, la masse moyenne des femelles gestantes était de 4,628 kg et l'écart-type de 1,349 kg. On sait, par ailleurs, que la masse des femelles gestantes de *Marmota marmota* est une variable normale dont l'écart-type est de 1,55 kg. On voudrait savoir si on peut considérer que la masse moyenne des femelles gestantes dans la population de la Vanoise est égale à 4,755 kg (moyenne théorique de référence pour les femelles gestantes de *Marmota marmota*). On ne peut pas ici rejeter l'hypothèse d'égalité des deux écart-types.

Question 50 Quelle est l'hypothèse nulle que vous allez formuler ?

- ☐ H_0 : La masse moyenne des femelles gestantes dans la population est égale à la masse de 4,628 kg
- ☐ H_0 : La masse moyenne des femelles gestantes dans la population est égale à la masse théorique de 4,755 kg
- ☐ H_0 : La masse moyenne des femelles gestantes dans l'échantillon est égale à la masse théorique de 4,755 kg
- ☐ H_0 : La masse moyenne des femelles gestantes dans la population n'est pas égale à la masse théorique de 4,755 kg
- ☐ H_0 : La masse moyenne des femelles gestantes dans l'échantillon n'est pas égale à la masse de 4,628 kg

Question 51 Sous l'hypothèse nulle, quelle est la distribution de l'estimateur de la masse moyenne des femelles de deux ans ?

- ☐ Une loi normale de paramètres $(4,755,\ \sqrt{\frac{1{,}349^2}{62}})$
- ☐ Une loi normale de paramètres $(4,755,\ \sqrt{\frac{1{,}55}{63}})$
- ☐ Une loi normale de paramètres $(4,755,\ \sqrt{\frac{1{,}55^2}{63}})$
- ☐ Une loi normale de paramètres $(4,628,\ \sqrt{\frac{1{,}349}{62}})$
- ☐ Une loi normale de paramètres $(4,628,\ 1,349)$

Question 52 Quelle est la valeur absolue de la statistique observée ?

☐ 0,741 ☐ 0,809 ☐ 0,861 ☐ 0,650

Question 53 Quelle est la valeur seuil de rejet de la statistique ($\alpha = 0{,}02$) ?

☐ 2,326 ☐ 1,96 ☐ 0,75 ☐ 0,83

Question 54 Quelle est la probabilité de vous tromper dans votre conclusion ?

- ☐ La probabilité est égale à 0,02
- ☐ La probabilité est égale à 0,98
- ☐ La probabilité est inconnue et égale à $1 - \beta$, la puissance du test
- ☐ On ne connait pas la valeur de cette probabilité, il s'agit de β, risque de deuxième espèce
- ☐ La probabilité est égale à 0,05
- ☐ La probabilité est égale à 0,95

17 Séance 17

17.1 Première partie

Un entomologiste doit conduire une étude sur un insecte ravageur, la pyrale du maïs *Ostrinia nubilalis* (Lépidoptère). La première expérience a consisté à échantillonner 200 plants de maïs dans un champ cultivé de 500 000 plants afin d'estimer la distribution du nombre de larves (chenilles) de pyrale par plant. Le chercheur désirait obtenir un maximum de plants infestés par l'insecte et pour cela il a utilisé un indice morphologique (trou dans la tige du plant) sensé indiquer qu'au moins une larve de pyrale avait pénétré dans la tige du maïs (afin d'être sûr que la plupart des plants de son échantillon serait infesté par l'insecte).

Question 55 La stratégie d'échantillonnage utilisée par le scientifique vous parait :

- ☐ Bonne car chaque plant de maïs du champ avait la même probabilité d'être prélevé
- ☐ Bonne car la taille de l'échantillon est supérieure à 30
- ☐ Mauvaise car tous les plants du champ n'ont pas été analysés
- ☐ Mauvaise car chaque plant de maïs du champ n'avait pas la même probabilité d'être prélevé

Question 56 Quelle est la variable aléatoire à étudier ?

- ☐ Variable continue : nombre de larves par plant
- ☐ Variable discrète : nombre de plants sains
- ☐ Variable continue : nombre de plants sains
- ☐ Variable discrète : nombre de larves par plant

Les données obtenues sont résumées dans le tableau suivant :

X	0	1	2	3	4	5	6	7	8	≥ 9	Total
n	6	15	40	42	37	30	10	12	8	0	200
nx	0	15	80	126	148	150	60	84	64	0	727

Avec X : nombre de larves par plant et n = nombre de plants infestés par x larves. Pour optimiser la lutte contre ce ravageur, le chercheur voudrait savoir comment les larves sont réparties. Pour cela, il teste si la distribution de X suit une loi de probabilité caractérisée par un seul paramètre (moyenne=variance).

Question 57 Quelle est la bonne formulation de l'hypothèse nulle associée au test qu'il a utilisé ?

- ☐ La distribution de X dans le champ étudié suit une loi de Bernoulli
- ☐ La distribution de X dans l'échantillon suit une loi de Bernoulli
- ☐ La distribution de X dans l'échantillon suit une loi de Poisson
- ☐ La distribution de X dans le champ étudié suit une loi de Poisson

Question 58 Quelle est la moyenne de la distribution observée dans l'échantillon ?

- [ ] $3,63$
- [ ] $3,53$
- [ ] $3,43$
- [ ] $3,45$

Question 59 Sachant que pour la suite de nos calculs on a arrondi la moyenne de la distribution observée à la valeur $3,5$, quelle est la valeur de la probabilité attendue sous H_0 pour $X = 0$?

- [ ] $P(X = 0) = 0,0502$
- [ ] $P(X = 0) = 0,0302$
- [ ] $P(X = 0) = 0,1020$
- [ ] $P(X = 0) = 0,0042$

Question 60 Le nombre de plants non infestés ($X = 0$) attendu sous l'hypothèse nulle est obtenu en multipliant $P(X = 0)$ par :

- [ ] 200
- [ ] 727
- [ ] 927
- [ ] 6

L'ensemble des effectifs observés, n, et théoriques, n_t (= attendus sous H_0) est indiqué dans le tableau suivant.

x	0	1	2	3	4	5	6	7	8	≥ 9
n	6	15	40	42	37	30	10	12	8	0
n_t	6,04	21,14	37	43,16	37,76	26,44	15,42	7,7	3,38	1,96

Question 61 Quelle est la valeur observée de la statistique, χ^2_{obs}, du test ?

- [ ] $6,18$
- [ ] $15,18$
- [ ] $20,18$
- [ ] $8,18$

Question 62 Combien de paramètres avons-nous dû estimer pour calculer les probabilités attendues sous H_0 ayant permis de calculer les effectifs théoriques ?

- [ ] 2
- [ ] 3
- [ ] 1
- [ ] 0

Question 63 Quelle est le nombre de degrés de liberté (ddl) associé à la valeur seuil de la statistique utilisée ?

- [ ] 6
- [ ] 8
- [ ] 7
- [ ] 9

Question 64 En choisissant un risque $\alpha = 0,01$ quelle est la valeur théorique de la statistique du test utilisé ?

- [ ] $16,81$
- [ ] $18,48$
- [ ] $20,09$
- [ ] $21,67$

Question 65 Conclusion statistique à l'issue de votre test ?

- [ ] Non rejet de H_0, avec un risque β inconnu
- [ ] Non rejet de H_0, avec un risque β égal à $0,01$
- [ ] Rejet de H_0 avec un risque β inconnu
- [ ] Rejet de H_0 avec un risque α égal à $0,01$

Question 66 Conclusion biologique à l'issue de votre test ?

- ☐ Les données observées sont compatibles avec une distribution des larves/plant dans le champ étudié qui suivrait une loi de POISSON
- ☐ Les données observées sont compatibles avec une distribution des larves/plant dans le champ étudié qui suivrait une loi de BERNOULLI
- ☐ Les données observées ne sont pas compatibles avec une distribution des larves/plant dans le champ étudié qui suivrait une loi de POISSON
- ☐ Les données observées ne sont pas compatibles avec une distribution des larves/plant dans le champ étudié qui suivrait une loi de BERNOULLI

Question 67 Les conclusions des 2 questions précédentes vous paraissent-elles :

- ☐ Critiquables car l'échantillonnage réalisé sous-estime le nombre de plants sains (non attaqué par l'insecte)
- ☐ Critiquables car l'échantillonnage réalisé surestime le nombre de plants sains (non attaqué par l'insecte)
- ☐ Non critiquables car l'échantillonnage est représentatif de la population étudiée
- ☐ Non critiquables car la taille de l'échantillon est supérieure à 30

17.2 Deuxième partie

AFIN d'optimiser le dosage de l'insecticide à utiliser pour lutter contre la pyrale du maïs, le chercheur a pour objectif d'estimer le poids moyen (en mg) des larves à partir d'un échantillon de taille $n = 10$. Il obtient :

$$\sum_{i=1}^{n} x_i = 200 \text{ et } \sum_{i=1}^{n} {x_i}^2 = 100000$$

Question 68 Quelles sont les estimations de la moyenne et de l'écart-type du poids des larves dans le champ ?

- ☐ $\hat{\mu} = 20,0$; $\hat{\sigma} = 10,54$
- ☐ $\mu = 20,0$; $\sigma = 10,54$
- ☐ $\bar{x} = 20,10$; $s = 10,54$
- ☐ $\mu = 20,10$; $s = 10,00$

Question 69 Imaginons que l'on prélève un nombre élevé d'échantillons aléatoires simples de taille égale à celle de l'échantillon de l'étude. Parmi les formulations suivantes, laquelle est exacte ?

- ☐ Toutes les estimations sont différentes entre elles et en moyenne très différentes de la moyenne de la population
- ☐ Toutes les estimations sont différentes entre elles mais en moyenne proches de la moyenne de la population
- ☐ Toutes les estimations sont égales entre elles mais différentes de la moyenne de la population
- ☐ Toutes les estimations sont égales entre elles et identiques à la moyenne de la population

POUR intégrer la fluctuation d'échantillonnage dans l'analyse, on calcule un intervalle de confiance (niveau de confiance = 0,99 soit $\alpha = 0,01$).

Question 70 Que vaut cet intervalle de confiance ?

- ☐ $[9,72\ ;\ 30,38]$
- ☐ $[13,47\ ;\ 26,53]$
- ☐ $[9,17\ ;\ 30,83]$
- ☐ $[12,46\ ;\ 27,54]$

Question 71 Comment interprétez-vous cet intervalle de confiance ?

- ☐ μ a 0,5 % de chances de se trouver dans cet intervalle
- ☐ μ a 1 % de chances de se trouver dans cet intervalle
- ☐ μ a 95 % de chances de se trouver dans cet intervalle
- ☐ μ a 99 % de chances de se trouver dans cet intervalle

Question 72 Conclusion

- ☐ L'intervalle de confiance a 99 % de chance de contenir la valeur du poids moyen des larves de la population attaquant le champ cultivé
- ☐ L'intervalle de confiance a 1 % de chance de contenir la valeur du poids moyen des larves de la population attaquant le champ cultivé
- ☐ L'intervalle de confiance a 95 % de chance de contenir la valeur du poids moyen des larves de la population attaquant le champ cultivé
- ☐ L'intervalle de confiance a 5 % de chance de contenir la valeur du poids moyen des larves de la population attaquant le champ cultivé

Question 73 En gardant le même niveau de confiance que ci-dessus, comment faire pour diminuer la marge d'erreur de l'estimateur de la moyenne μ (= moyenne de la population) ?

- ☐ Réduire n, la taille de l'échantillon
- ☐ Augmenter σ l'écart-type du poids des larves de la population
- ☐ Augmenter l'erreur standard de l'estimateur de la moyenne μ
- ☐ Augmenter n, la taille de l'échantillon

17.3 Troisième partie

Le dernier objectif de l'étude est de savoir (au risque $\alpha = 0,05$) si le poids moyen des larves d'une seconde population de pyrale du maïs infestant un autre champ est égale ou différente de celle de la population infestant le premier champ (indiqué dans les parties précédentes). On indique que la longueur moyenne calculée à partir de l'échantillon prélevé dans le second champ (de taille $n = 10$) est égale à 25 mm.

Question 74 Quel test doit-on réaliser ?

- ☐ Test d'égalité de 2 moyennes avec $H_0\ :\ \bar{x}_1 = \bar{x}_2$
- ☐ Test d'égalité de 2 moyennes avec $H_0\ :\ \mu_1 = \mu_2$
- ☐ Test de conformité de moyenne avec $H_0\ :\ \bar{x} = \mu_0$
- ☐ Test de conformité de moyenne avec $H_0\ :\ \mu = \mu_0$

Question 75 En sachant que des tests ont permis de considérer que les hypothèses d'homoscédasticité et de normalité étaient acceptables, quelle statistique de test doit-on utiliser pour comparer les 2 longueurs moyennes ?

- ☐ Une statistique qui, sous H_0, suit une loi de STUDENT et intègre l'estimation d'une variance commune
- ☐ Une statistique qui, sous H_0, suit une approximativement une loi normale réduite et intègre les estimations des variances des 2 populations
- ☐ Une statistique qui, sous H_0, suit une loi de POISSON et intègre les variances des 2 populations
- ☐ Une statistique qui, sous H_0, suit une loi normale réduite et intègre les variances des 2 populations

Question 76 Sachant qu'on a obtenu une valeur absolue de la statistique calculée inférieure à celle de la valeur seuil (=théorique), quelle est la conclusion du test :

- ☐ Le poids moyen des larves varie entre les deux populations avec un risque de se tromper de 5 %
- ☐ Le poids moyen des larves varie entre les deux échantillons avec un risque de se tromper de 5 %
- ☐ Le poids moyen de larves est le même dans les deux populations avec un risque β de se tromper de 5 %
- ☐ On n'a pas réussi à mettre en évidence de différence de poids moyen de larves entre les deux populations : H_0 n'est pas rejetée, avec un risque β inconnu

Question 77 Avez-vous confiance dans votre conclusion à la question précédente ?

- ☐ Oui car la probabilité d'accepter H_0 alors qu'elle est fausse est probablement faible en raison d'une forte variance et d'une faible taille d'échantillon
- ☐ Non car la puissance du test réalisé est probablement trop faible en raison d'une grande variance et d'une faible taille d'échantillon
- ☐ Oui car la probabilité de rejeter H_0 alors qu'elle est fausse est probablement élevée en raison d'une forte variance et d'une faible taille d'échantillon
- ☐ Non car la puissance du test est probablement trop élevée en raison d'une grande variance et d'une faible taille d'échantillon

Autres exercices sur les probabilités pour vous entrainer

18 Probabilités simples. Probabilités conditionnelles. Théorème de Bayes.

18.1 Des urnes et des boules (1-1**)

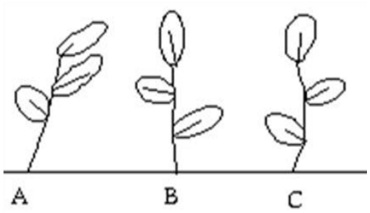
Un professeur décide de faire passer rapidement l'oral de "probabilités". L'étudiant est autorisé à répartir quatre boules, deux blanches et deux noires, entre deux urnes. Le professeur choisit alors au hasard une des urnes et en extrait une boule. Si la boule est noire, l'étudiant est reçu.
Comment répartiriez-vous les boules ?

18.2 Jet de dé (2-4*)

On jette un dé à 6 faces. Quel est le plus petit nombre de jets nécessaires pour obtenir au moins un as avec une probabilité supérieure à 0.5 ?

18.3 Contagion de plantes (2-1***)

On considère trois plantes A, B, C. La probabilité qu'une plante soit atteinte d'une virose (statisticose) est de 1/10. Chaque plante malade contamine chacune des plantes voisines avec la probabilité 1/5.



Quelle est la probabilité que la plante B soit malade ?

18.4 Match en 3 parties (3-1*)

Deux joueurs d'échecs A et B effectuent un match en trois parties. Pour chaque partie, et de façon indépendante des autres parties, on a :

— $P(\text{A gagne}) = 0.3$

— $P(\text{nulle})=0.2$

— $P(\text{B gagne})=0.5$

Sachant que le même joueur a gagné les 3 parties lors du match, quelle est la probabilité conditionnelle que ce soit le joueur A ?

18.5 Caractère albinos (3-3**)

Dans une population où le nombre de mâles est la moitié de celui des femelles, la fréquence du caractère albinos est de 6% chez les mâles et 0,36% chez les femelles.

1. Quelle est la probabilité qu'un individu pris au hasard soit albinos ?

2. Quelle est la probabilité qu'un albinos de la population soit un male ?

18.6 Daltonisme (3-5**)

On considère une population composée de 48% d'hommes et de 52% de femmes. La probabilité qu'un homme soit daltonien est 0.05, qu'une femme soit daltonienne est 0.0025 (on peut, peut-être, rappeler que le daltonisme est déterminé par un gène lié au chromosome X, et a donc une incidence différente suivant le sexe).
Quelle proportion de la population est daltonienne ?

18.7 2 maladies (3-3***)

On sait, par des enquêtes médicales, qu'un individu appartenant à une population donnée a la probabilité $\frac{1}{100}$ d'être atteint d'une affection A et la probabilité $\frac{1}{20}$ d'être atteint d'une autre affection B.
Si ces affections sont indépendantes,

1. Combien de sujets atteints au moins des deux affections doit-on s'attendre à trouver dans un échantillon de 10 000 sujets pris dans la population ?
2. Combien de sujets atteints de B doit-on trouver sur 500 sujets atteints de A ?

19 Variables aléatoires discrètes et loi normale

19.1 Vols aériens surréservés (1-2*)

Environ 5% des réservations pour les vols aériens sur une ligne donnée ne sont pas utilisées, c'est pourquoi une compagnie vend 100 billets pour 97 places.
Quelle est la probabilité que tous les passagers aient une place ?

19.2 2 dés pour 1 note (1-1**)

Pour fournir la note de fin d'année, le professeur X... lance deux dés et considère la plus petite des deux valeurs obtenues.
Il définit alors une variable aléatoire N, appelée "note" :

$$N : \text{ 3 fois la plus petite valeur obtenue.}$$

1. Écrire la loi de probabilité de N.
2. Calculer l'espérance $E(N)$ et la variance $V(N)$

19.3 Divisions d'un microorganisme (1-3**)

Un microorganisme se reproduit par divisions à intervalles réguliers. Suite à une première division, la suivante se produit 1 heure après (70% des cas) ou 2 heures après (30% des cas), indépendamment de l'histoire du microorganisme.
On isole un mircroorganisme après une division et on appelle X le nombre de microorganismes présents deux heures après.

1. Etablir la loi de probabilité de X.

2. Calculer E(X).

19.4 Impuretés dans des bouteilles (1-5**)

Avec 100 kg de verre liquide, on fabrique 100 bouteilles. Dans la masse en fusion sont réparties au hasard 100 impuretés.
Soit X la variable aléatoire "Nombre d'impuretés dans une bouteille".
Calculez :

a- $P(X = 0)$

b- $P(X \geqslant 2)$

c- $E(X)$

19.5 Suicides (1-8**)

Dans une grande ville, il y a en moyenne un suicide par jour.
Soit la variable aléatoire X : nombre de jours de l'année où au moins cinq personnes se suicident.
Calculez $E(X)$.

19.6 Monsieur Lune (1-1***)

Monsieur Lune affirme que, grâce à son nouvel ordinateur, il peut prédire le sexe des enfants à naître. Pour cette prédiction, il ne coûte que la somme de 20 euros, destinée à couvrir ses frais de gestion, et pour "prouver" sa bonne foi il rembourse en cas de prédiction erronée !

Soit X la variable aléatoire : gain de Monsieur Lune.

1. Ecrire la loi de probabilité de X.
2. Si Monsieur Lune trouve 1000 naïfs, combien peut-il espérer gagner ?

19.7 Jumeaux en Slovaquie (1-3***)

Les statistiques de naissance des jumeaux en Slovaquie en 1998 sont les suivantes :

{G,G}	874
{G,F,F,G}	954
{F,F}	811

On admet que la probabilité de naissance d'un garçon est p=$\frac{1}{2}$.

1. Montrer que la naissance de jumeaux n'est pas une expérience à deux épreuves indépendantes.
2. Soit f la proportion de vrais jumeaux et s la proportion de jumeaux de même sexe. Trouver la relation entre f et s. En déduire, pour la Tchécoslovaquie en 1958, la proportion de vrais jumeaux.

19.8 Loi normale (2-1*)

Une variable aléatoire suit une loi Normale de moyenne 4 et de variance 10.
Quelles sont les probabilités des événements :

1. $X < 4$
2. $-2,2 < X < 10,2$

19.9 Pupes de drosophiles (2-2**)

On a pesé les pupes d'un lot de 500 drosophiles dessechées, les poids étant répartis entre 425 μg et 565 μg. On a trouvé pour moyenne et variance de cet échantillon :

$$\bar{x} = 500\,\mu\text{g} \quad s^2 = 576\,\mu\text{g}^2 \tag{1}$$

En supposant que cet échantillon est extrait d'une population normale pour le caractère mesuré, calculer les effectifs théoriques de pupes de drosophiles dont le poids est :

— inférieur à 500 μg

— inférieur à 450 μg ou supérieur à 500 μg

— compris entre 400 et 500 μg

19.10 Pièce de monnaie lancée 500 fois (3-1*)

Une pièce de monnaie est lancée 500 fois. Trouver la probabilité que le nombre de faces obtenues ne diffère pas de 250 de plus de 30.

Autres exercices de statistique pour vous entrainer

20 Statistique descriptive et intervalle de confiance

20.1 Poule pondeuse (1-2*)

Le nombre d'œufs pondus en un an par une poule pondeuse a été relevé pendant 8 ans. Les résultats sont les suivants :

Année (t)	1	2	3	4	5	6	7	8
Nombre d'œufs pondus (n)	160	140	122	112	96	88	72	62

a) Représenter graphiquement ces données. Quelle relation entre n et t vous suggère cette représentation ?

b) Calculer la moyenne, la variance et l'écart type observés dans cet échantillon du nombre d'œufs pondus.

20.2 Taille des élèves (1-1**)

On a mesuré la taille (en cm). Les résultats sont les suivants :

138	164	150	132	144	125	149	157
146	158	140	147	136	148	152	144
168	126	138	176	163	119	154	165
146	173	142	147	135	153	140	135
161	145	135	142	150	156	145	128

a) Calculer la moyenne des tailles.

b) Regrouper les données en 10 classes puis en 5 classes seulement. Représenter graphiquement les résultats dans les 2 cas. Calculer la moyenne dans les 2 cas.

20.3 Accidents automobiles (1-2**)

On a enregistré, pendant 50 jours, le nombre d'accidents automobiles à PEARSON-GULCH. Les résultats sont les suivants :

Nombre d'accidents x_i	0	1	2	3	4
Nombre de jours (n_i) où s'est produit x_i accidents	21	18	7	3	1

a) Représenter graphiquement les données.

b) Calculer la moyenne et la variance du nombre d'accidents.

c) Quelle loi de probabilité pourrait suivre ce nombre d'accidents par jour ?

20.4 Soies sur les drosophiles (2-5**)

On a compté le nombre de soies sur le quatrième segment abdominal d'un lot de Drosophiles. Les résultats sont consignés dans le tableau suivant :

Nombre x de soies	16	17	18	19	20
Nombre de mouches ayant x soies	2	5	7	3	3

On demande :

a. De calculer la moyenne de cet échantillon,

b. De calculer l'intervalle de confiance de cette moyenne.

20.5 Urnes et boules (3-1*)

Une urne contient des boules rouges et des boules blanches. On tire successivement 128 boules en remettant chaque boule dans l'urne après tirage : 64 boules rouges ont été tirées.
Donner l'intervalle de confiance à 0.95 de la proportion des boules rouges dans l'urne.

20.6 Tabagisme (3-2*)

On a trouvé que le pourcentage de fumeurs d'une certaine population était de 60%. On prélève 100 personnes de cette population et on parie que le pourcentage de fumeurs sera compris entre 0.5 et 0.57. Quel est le risque de perdre le pari ?

20.7 Formule leucocytaire (3-3*)

Une formule leukocytaire, déterminée par l'examen de 500 éléments blancs, a donné comme pourcentage de lymphocytes 30 %. Peut-on considérer ce pourcentage comme significativement différent de 25% ?

20.8 Vote (3-1**)

Un échantillon de 100 votants choisis au hasard parmi tous les votants d'un bureau donné indique que 55% d'entre eux votent pour le candidat A.

— Donner l'intervalle de confiance à 95% de la proportion de tous les votants en faveur du candidat A.

— Quelle aurait dû être la taille de l'échantillon pour conclure au niveau 95% que le candidat A serait élu ?

20.9 Yeux des drosophiles (3-2**)

Sur des Drosophiles, on considère le caractère "couleur des yeux". Par croisement d'homozygotes "yeux rouges" AA et d'homozygotes "yeux bruns" aa, on obtient des hétérozygotes Aa. Si on croise ces hétérozygotes entre eux, on doit obtenir (loi de Mendel) 3/4 de "yeux rouges" (AA et Aa), et 1/4 de "yeux bruns" (aa). On a observé, chez 309 drosophiles de cette seconde génération, 241 "yeux rouges" et 68 " yeux bruns".

Cette répartition est-elle conforme à la loi de Mendel ?

Combien aurait-il fallu observer de drosophiles " yeux rouges", pour décider de rejeter la conformité avec la loi de Mendel, au risque 5% ?

20.10 Pouvoir germinatif (3-3**)

Un laboratoire d'agronomie a effectué une étude sur le maintien du pouvoir germinatif des graines de *Papivorus subaquaticus* après une conservation de 3 ans. Sur un lot de 80 graines, 47 ont germé.
Donner, au risque d'erreur 5%, un intervalle de confiance symétrique pour la probabilité de germination des graines de *Papivorus subaquaticus* après une conservation de 3 ans.

20.11 Sondage (3-1***)

Au second tour d'une élection présidentielle, deux candidats sont en présence.
Combien faut-il interroger d'électeurs pour obtenir une estimation à 1% près des intentions de vote ?

21 Tests d'hypothèses

21.1 Conformité de pièces (2-1**)

Dans une entreprise on fabrique des pièces métalliques. Le poids moyen de ces pièces est de 10 g. Les calculs faits sur des statistiques antérieures ont montré que l'écart-type du poids de ces pièces est de 0.3 g.
Au cours d'une certaine période on fabrique un très grand nombre de pièces. Comme il n'est pas possible de contrôler que toutes les pièces pèsent en moyenne 10 g, on en pèse 100 prises au hasard dans la production de la période considérée. Le poids moyen de ces 100 pièces est $\bar{x} = 10.05$ g.

1. Que penser de ce résultat ?
2. Est-il en contradiction avec la moyenne jusque là admise ?

21.2 Hauteurs d'arbres (1**)

Dans deux types de forêts distincts, on a mesuré les hauteurs respectivement de 13 et 14 peuplements choisis au hasard et indépendamment (*cf. Tableau*), dans le but de vérifier si les hauteurs moyennes des deux types sont ou ne sont pas égales.

Type 1	Type 2
23,4	22,5
24,4	22,9
24,6	23,7
24,9	24,0
25,0	24,4
26,2	24,5
26,3	25,3
26,8	26,0
26,8	26,2
26,9	26,4
27,0	26,7
27,6	26,9
27,7	27,4
	28,5

Quelle est votre conclusion ?

21.3 Familles de 5 enfants (2**)

On s'intéresse à la composition des familles de 5 enfants. L'observation de 320 familles de 5 enfants a permis de dresser le tableau :

Nombre de filles dans la famille	0	1	2	3	4	5
Nombre de familles	18	56	110	88	40	8

A partir de ces résultats, testez l'hypothèse H_0 : "dans chaque famille, les naissances successives sont indépendantes entre elles et la probabilité de naissance d'une fille est constamment égale à 0,5".

21.4 Succès thérapeutique et âge (4-3**)

On veut savoir si les résultats d'une thérapeutique sont les même suivant l'âge du malade. Les malades testés ont été répartis en quatre tranches d'âges. On a noté, pour chaque groupe, le nombre de succès de la thérapeutique. On a obtenu les résultats suivants :

âge x	Effectif du groupe	Nombre de Succès
$x < 30$	200	150
$30 \leqslant x < 40$	110	75
$40 \leqslant x < 50$	150	80
$x \geqslant 50$	80	45

Quelle conclusion tirez-vous ?

21.5 Maladies des bovins (4**)

On connaît les effectifs d'individus affectés pour les trois maladies principales rencontrées chez les bovins dans trois régions du Western Wollega en Ethiopie.

	Fièvre aphteuse	PPCB	Trypanomose
Mana Sibu	150	240	210
Nole Kaba	140	130	280
Anfilo	220	180	140

Peut-on conclure, au risque de 0,05, à une différence de répartition des maladies dans les trois régions du Western Wollega ?

Annexes :

A Tables de la loi normale

$$\text{Loi normale : table de la fonction } \Pi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du.$$

La table donne les valeurs de $\Pi(x)$ pour x positif. Lorsque x est négatif, il faut prendre le complément à l'unité de la valeur lue dans la table.

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767

x	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
$\Pi(x)$	0.9772	0.9821	0.9861	0.9893	0.9918	0.9938	0.9953	0.9965	0.9974	0.9981

Exemple : pour $x = +1.37$ $\quad \Pi(x) = 0.9147$
pour $x = -1.37$ $\quad \Pi(x) = 0.0853$

Loi normale : table de l'écart-réduit.

La table donne, pour une probabilité α, la valeur ε telle que la probabilité que l'écart-réduit égale ou dépasse en valeur absolue ε vaut α.

α	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	∞	2.576	2.326	2.170	2.054	1.960	1.881	1.812	1.751	1.695
0.10	1.645	1.598	1.555	1.514	1.476	1.440	1.405	1.372	1.341	1.311
0.20	1.282	1.254	1.227	1.200	1.175	1.150	1.126	1.103	1.080	1.058
0.30	1.036	1.015	0.994	0.974	0.954	0.935	0.915	0.896	0.878	0.860
0.40	0.842	0.824	0.806	0.789	0.772	0.755	0.739	0.722	0.706	0.690
0.50	0.674	0.659	0.643	0.628	0.613	0.598	0.583	0.568	0.553	0.539
0.60	0.524	0.510	0.496	0.482	0.468	0.454	0.440	0.426	0.412	0.399
0.70	0.385	0.372	0.358	0.345	0.332	0.319	0.305	0.292	0.279	0.266
0.80	0.253	0.240	0.228	0.215	0.202	0.189	0.176	0.164	0.151	0.138
0.90	0.126	0.113	0.100	0.088	0.075	0.063	0.050	0.038	0.025	0.013

La probabilité α s'obtient par addition des nombres écrits en marge.
Exemple : pour $\alpha = 0.00 + 0.05 = 0.05$, $\varepsilon = 1.960$.

Table pour les petites valeurs de probabilité :

α	0.001	0.000 1	0.000 01	0.000 001	0.000 000 1	0.000 000 01	0.000 000 001
ε	3.29053	3.89059	4.41717	4.89164	5.32672	5.73073	6.10941

B Table de la loi de Student (t)

Table de t.

La table donne, pour une probabilité α et un nombre de degrés de liberté (d.d.l.) donnés, la valeur x telle que la probabilité que t égale ou dépasse x en valeur absolue vaut α.

d.d.l. \ α	0.90	0.50	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	0.158	1.000	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	0.142	0.816	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	0.137	0.765	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.134	0.741	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.132	0.727	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.131	0.718	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.130	0.711	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.130	0.706	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.129	0.703	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.129	0.700	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.129	0.697	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.128	0.695	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.128	0.694	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.128	0.692	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.128	0.691	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.128	0.690	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.128	0.689	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.127	0.688	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.127	0.688	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.127	0.687	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.127	0.686	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.127	0.686	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.127	0.685	1.060	1.319	1.714	2.069	2.500	2.807	3.768
24	0.127	0.685	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.127	0.684	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.127	0.684	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.127	0.684	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.127	0.683	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.127	0.683	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.127	0.683	1.055	1.310	1.697	2.042	2.457	2.750	3.646
∞	0.126	0.674	1.036	1.282	1.645	1.960	2.326	2.576	3.291

Exemple : Une variable aléatoire suivant une loi de t avec 10 degrés de liberté a une probabilité $\alpha = 0.05$ d'être égale ou supérieure à 2.228 en valeur absolue.

C Table de la loi de Pearson (χ^2)

Table de χ^2.

La table donne, pour une probabilité α et un nombre de degrés de liberté (d.d.l.) donnés, la valeur x telle que la probabilité que χ^2 égale ou dépasse x vaut α.

d.d.l. \ α	0.90	0.50	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	0.016	0.455	1.074	1.642	2.706	3.841	5.412	6.635	10.828
2	0.211	1.386	2.408	3.219	4.605	5.991	7.824	9.210	13.816
3	0.584	2.366	3.665	4.642	6.251	7.815	9.837	11.345	16.266
4	1.064	3.357	4.878	5.989	7.779	9.488	11.668	13.277	18.467
5	1.610	4.351	6.064	7.289	9.236	11.070	13.388	15.086	20.515
6	2.204	5.348	7.231	8.558	10.645	12.592	15.033	16.812	22.458
7	2.833	6.346	8.383	9.803	12.017	14.067	16.622	18.475	24.322
8	3.490	7.344	9.524	11.030	13.362	15.507	18.168	20.090	26.124
9	4.168	8.343	10.656	12.242	14.684	16.919	19.679	21.666	27.877
10	4.865	9.342	11.781	13.442	15.987	18.307	21.161	23.209	29.588
11	5.578	10.341	12.899	14.631	17.275	19.675	22.618	24.725	31.264
12	6.304	11.340	14.011	15.812	18.549	21.026	24.054	26.217	32.909
13	7.042	12.340	15.119	16.985	19.812	22.362	25.472	27.688	34.528
14	7.790	13.339	16.222	18.151	21.064	23.685	26.873	29.141	36.123
15	8.547	14.339	17.322	19.311	22.307	24.996	28.259	30.578	37.697
16	9.312	15.338	18.418	20.465	23.542	26.296	29.633	32.000	39.252
17	10.085	16.338	19.511	21.615	24.769	27.587	30.995	33.409	40.790
18	10.865	17.338	20.601	22.760	25.989	28.869	32.346	34.805	42.312
19	11.651	18.338	21.689	23.900	27.204	30.144	33.687	36.191	43.820
20	12.443	19.337	22.775	25.038	28.412	31.410	35.020	37.566	45.315
21	13.240	20.337	23.858	26.171	29.615	32.671	36.343	38.932	46.797
22	14.041	21.337	24.939	27.301	30.813	33.924	37.659	40.289	48.268
23	14.848	22.337	26.018	28.429	32.007	35.172	38.968	41.638	49.728
24	15.659	23.337	27.096	29.553	33.196	36.415	40.270	42.980	51.179
25	16.473	24.337	28.172	30.675	34.382	37.652	41.566	44.314	52.620
26	17.292	25.336	29.246	31.795	35.563	38.885	42.856	45.642	54.052
27	18.114	26.336	30.319	32.912	36.741	40.113	44.140	46.963	55.476
28	18.939	27.336	31.391	34.027	37.916	41.337	45.419	48.278	56.892
29	19.768	28.336	32.461	35.139	39.087	42.557	46.693	49.588	58.301
30	20.599	29.336	33.530	36.250	40.256	43.773	47.962	50.892	59.703

Exemple : Une variable aléatoire suivant une loi de χ^2 avec 3 degrés de liberté a une probabilité $\alpha = 0.90$ d'être égale ou supérieure à 0.584.